Regularity of the Brakke Flow

Dissertation
zur Erlangung des Grades eines Doktors der Naturwissenschaften
am Fachbereich Mathematik und Informatik der Freien Universität Berlin

vorgelegt von
Ananda Lahiri

Berlin 2014

Erstgutachter: Prof. Dr. Klaus Ecker
Zweitgutachter: Dr. habil. Felix Schulze

Tag der Disputation: 11.07.2014

Contents

1 Introduction

Mean Curvature Flow This thesis deals with regularity of the mean curvature flow, which is the gradient flow of the area-functional. The solutions have many features in common with the heat flow and in particular there is a smoothing effect, which is the central idea behind the results presented here. Applications of the mean curvature flow are numerous and wide spread among different areas of research including physics, material science, image processing and mathematical finance. Specifically it is used to find minimal surfaces, remove noise inaccuracy from images or for the description of grain growth in metals.

The mean curvature flow was introduced in K.A. Brakke's work 'The motion of a surface by its mean curvature' in 1978 [B]. He described the flow in a very general setting using families of varifolds and prescribing their behaviour on test-functions. Besides proving existence of non-trivial solutions and perpendicularity of the generalised mean curvature vector, Brakke's book also contains a comprehensive regularity theory. His main regularity theorem says that a weak mean curvature flow is smooth in some neighbourhood of almost every point in space-time. Since then, there have been many new developments. Starting with Huisken's work in 1984 (see [H1]) many results were proven especially for the smooth flow. One important achievement is the monotonicity formula, first formulated by Huisken in [H2], later localized by Ecker (see [E1]) and generalised to varifolds by Tom Ilmanen in [I2]. Another such formula will be derived here in order to establish a distance estimate. A further interesting result is White's theorem from [W4], where he showed local regularity in case the Gaussian density ratios are close to one.

Later the weak mean curvature flow was reformulated as a level set problem by Chen, Giga and Goto [CGG] as well as by Evans and Spruck [ES]. Another weak characterization was given by Ilmanen [I] in form of weak set-theoretic sub-solutions. These new approaches led to further regularity results using stronger assumptions, in particular for convex flows.

Although there has been a lot of progress describing the mean curvature flow in special situations, Brakke's conclusions are still the best one can get in the general case. There are two other main sources dealing with these results. In 2004, Ecker published a book [E4] about Brakke's regularity theory focussing on the special case where smooth solutions of mean curvature flow develop singularities for the first time. In 2012, Kasai and Tonegawa proved Brakke's main regularity theorem for a more general flow [KT]. Although Ecker's book included a lot of Brakke's original ideas, especially in the appendix, and [KT] makes full use of the popping soap film lemma [B, 6.6], one of Brakke's central techniques, the main idea to obtain local regularity in the original work, is not presented in either at all. Brakke uses a cylindrical heat kernel to explicitly construct families of graphs and under certain flatness assumptions these move almost by mean curvature flow. Somehow it seems very natural to simulate mean curvature flow of almost flat surfaces by heat diffusion, but this approach does not appear to have been used for mean curvature flow anywhere else.

Summary The main purpose of this thesis is to give a simpler proof of Brakke's regularity theorems [B, 6.10, 6.12]. We use the original approach fixing the numerous, often non trivial, gaps in Brakke's arguments, and try to improve some of his calculations and estimates. Among other things we make use of techniques and results developed in more recent years specifically Huisken's monotonicity formula from [H2]. In doing this, we maintain Brakke's central idea of approximating mean curvature flow by linear heat diffusion. Moreover, we elaborate on this approach to make it more adaptive for potential further applications.

We consider n-surfaces in $\mathbb{R}^{n+k}$ for integers $n \geq 2, k \geq 1$. The mean curvature flow in the smooth case is given by

$$\frac{\partial F}{\partial t}(t,p) = \vec{H}\left(F(t,p)\right), \tag{1.1}$$

where $F_t = F(t,\cdot) : N \to \mathbb{R}^{n+k}$ is a family of immersions, N is an n-dimensional smooth manifold and $\vec{H}$ is the mean curvature vector. Using the evolution of the area element due to Huisken [H1], equality (1.1) yields

$$\frac{d}{dt}\int_{\mathbb{R}^{n+k}} \phi \, d\mu_t \leq -\int_{\mathbb{R}^{n+k}} |\vec{H}|^2 \phi \, d\mu_t + \int_{\mathbb{R}^{n+k}} \vec{H} \cdot D\phi \, d\mu_t \tag{1.2}$$

for all $\phi \in C_c^1\left(\mathbb{R}^{n+k}, \mathbb{R}^+\right)$, and where μ_t is the induced measure from $M_t = F_t(N)$, i.e. $\mu_t = \mathscr{H}^n \llcorner M_t$. Following Brakke [B], inequality (1.2) can be generalised to families of integral varifolds, which is the motivation for Definition 3.4. A solution of this more general flow will be called a Brakke flow. In particular solutions of smooth mean curvature flow always induce Brakke flows. However the converse is not true. One aim of this thesis is to give criteria for Brakke flows to actually be induced by smooth mean curvature flows. Note that in the smooth case, (1.2) actually holds with equality and for all $\phi \in C_c^1\left(\mathbb{R}^{n+k}, \mathbb{R}\right)$, but for the generalisation one only demands inequality to obtain compactness and existence results.

Below we give a list of our main results using intuitive geometric formulations. For the precise statements see the corresponding theorems.

1. A comparison theorem relating mean curvature flow and heat flow (see Theorem 4.15). This is a much stronger version of [B, 6.8]. In particular we are able to get rid of the slab condition and the mean curvature term that appear in Brakke's version. Moreover our result does not assume a sign on the test function.

2. If a Brakke flow in some region is contained in a narrow enough slab and also the area ratio in suitable balls is controlled by certain bounds, then in a smaller region it is actually smooth and graphical (see Theorem 8.4, originally [B, 6.10]). Note that in contrast to Brakke we don't assume unit density. Moreover Brakke's proof contains a major gap in the usage of the clearing out lemma, which he needs to obtain height bounds. We correct his argument and also give an alternative proof with a height estimate derived from Huisken's monotonicity formula.

3. The general regularity theorem 9.7 (originally [B, 6.12]), says that at a time, where no sudden loss of area occurs, the singular set of a Brakke flow has top-dimensional measure zero. We provide a new streamlined version of his proof incorporating ideas from [E2], which make it much shorter and more transparent.

For items 2 and 3 there are alternative versions due to Kasai and Tonegawa (see [KT] and [T]) which precede our work. They consider a more general flow, using techniques, which are different to both Brakke's and ours. In particular the explicit constructions via the heat kernel are replaced by indirect blow up arguments.

Besides proving these results we apply Theorem 8.4 to Brakke flows, which start from a "very plane-like "varifold. A varifold μ is considered very plane-like in a cylinder $C_R(y)$, if $\mathrm{spt}\mu \cap C_R(y)$ can be written as a graph over $\mathbb{R}^n$ except for a set which has small μ-measure compared to R^n. Moreover on the graphical part of $\mathrm{spt}\mu$ the graph function has to satisfy a certain height and gradient bound. (see Definition 11.1). Note that the measure and the height bound are seen in relation to R, so in a larger cylinder the varifold may be more plane-like. Within this framework of varifolds we can show the following new results:

1. Consider a Brakke flow starting from a varifold which is very plane-like in $C_R(y)$. Then there are two possibilities:
   (1) After some time there exists a period of time where the flow is smooth and graphical inside a smaller cylinder.
   (2) At some later time there exists a smaller cylinder with μ-measure zero.
   (See Theorem 11.7).

2. As a special case, we consider a Brakke flow in any slab, starting from a varifold which can be written as a graph over the whole $\mathbb{R}^n$, except for a compact set and which satisfies a certain gradient bound on the graphical part. Such a Brakke flow is very plane-like for large enough cylinders around every y, so by the previous result it will become smooth and graphical or there will be a cylinder which does not intersect the flow and will eventually grow infinitely large, see Theorem 11.17.

3. In the smooth case we prove a similar result where we assume the flow is very plane-like for a period of time, but in the plane-like condition we allow for arbitrary gradient on the graphical part. See Propositions 12.13 and 12.16.

4. Independently of the "plane-like "-setting we consider a smooth mean curvature flow, which is graphical in some cylinder $C_R(0)$ for a period of time $[-R^2, 0]$. Suppose that the gradient of the graphical representation is bounded by some L. Theorem 12.11 then says that we can extend the graphical representation on the smaller cylinder $C_{\delta R}(0)$ onto the short period of time $[0, \delta^2 R^2]$, where δ is bounded from below depending on L.

Methods Here we present our specific approach and explain how the individual results are related. Also the main calculations are shown in simplified form omitting most of the specific form of the error terms and writing everything in the co-dimension-1 case. Equations in which we leave out error terms feature the symbol $\approx$ instead of an equality sign. Usually these left out terms will be controlled by tilt-, height- and mean curvature-excess.

Sections 2 and 3 give an introduction to our setting. In particular we state an exact definition of a Brakke flow, see 3.1 and 3.4. Afterwards we derive the basic continuity properties and the behaviour of time dependent test functions, see Proposition 3.8.

In sections 4 - 8 we prove local regularity for Brakke flows based on the following calculation: Consider the cylindrical heat kernel

$$\Psi(t,x) := (2n)^{-\frac{n}{2}} e^{\frac{|\hat{x}|^2}{4t}}$$

(actually we use a truncated version to make things localized, see Definitions 4.1 and 4.2). The main trick is to carry out a convolution with this kernel, then use that a Brakke flow almost evolves by heat diffusion and approximate the result via Taylor expansion. Consider a Brakke flow (μ_t), such that $\mathrm{spt}\mu_t$ can be approximated by Lipschitz functions in some weak sense. Let

$f : [-T, T] \times \mathbb{R}^n \to \mathbb{R}$ be a family of such approximations, $r, T \in (0, \infty)$ and $p, q \in (0, T)$. By the properties of the heat kernel, we obtain

$$\int_{B_r^n(0)} f(0, \hat{y}) d\hat{y} \approx \int_{B_r^n(0)} \int x_{n+1} \Psi(p, \hat{x} - \hat{y}) d\mu_0(x) d\hat{y} \tag{1.3}$$

and a Taylor expansion yields

$$\begin{aligned} &\int_{B_r^n(0)} \int x_{n+1} \Psi(p + q, \hat{x}) d\mu_{-q}(x) d\hat{y} \\ &\approx \int_{B_r^n(0)} \hat{y} \cdot \int f(-q, \hat{x}) D\Psi(p + q, \hat{x}) d\hat{x} d\hat{y} \\ &\quad + \int_{B_r^n(0)} \int f(-q, \hat{x}) \Psi(p + q, \hat{x}) d\hat{x} d\hat{y}. \end{aligned} \tag{1.4}$$

We remark that (1.3) becomes more precise the smaller p is chosen, whereas (1.4) becomes more precise the larger $p + q$ is chosen. The main observation is that for a Brakke flow we have

$$\begin{aligned} &\int_{B_r^n(0)} \int x_{n+1} \Psi(p, \hat{x} - \hat{y}) d\mu_0(x) d\hat{y} \\ &\approx \int_{B_r^n(0)} \int x_{n+1} \Psi(p + q, \hat{x}) d\mu_{-q}(x) d\hat{y}, \end{aligned} \tag{1.5}$$

which becomes less precise for larger q. Then combining (1.3)-(1.4) yields

$$\begin{aligned} \int_{B_r^n(0)} f(0, \hat{y}) d\hat{y} \approx &\int_{B_r^n(0)} \hat{y} \cdot \int f(-q, \hat{x}) D\Psi(p + q, \hat{x}) d\hat{x} d\hat{y} \\ &+ \int_{B_r^n(0)} \int f(-q, \hat{x}) \Psi(p + q, \hat{x}) d\hat{x} d\hat{y}. \end{aligned} \tag{1.6}$$

This basically says that we can use $f(-q, \cdot)$ to define an affine function which is a good approximation to $f(0, 0)$ in some integral sense.

The basic estimates for (1.3)-(1.4) will be done in section 4. It turns out that (1.3)-(1.4) can be controlled by the maximal height, the tilt-excess and the mean curvature-excess of μ. In section 7 we show that a flow in a narrow slab which satisfies certain area ratio bounds has small tilt- and mean curvature-excess such that we can use the results from section 4. To use (1.6) effectively we need a mean curvature version of standard $L^\infty - L^2$-estimate to obtain a point-wise bound from the integral one. This is done in sections 5 and 6. Then, in section 8, we finally combine the results from sections 4 - 7 to derive (1.6) and use it to obtain $C^{1,\alpha}$-regularity.

In section 4 we start with transferring some basic properties of the usual heat kernel to the truncated heat kernel we use. Then we deal with relation (1.5), which is based on the main idea of estimating the difference between Brakke flow and heat flow. The major calculation is done in Proposition 4.11. This yields an estimate for the evolution equation of the heat kernel for a Brakke flow in terms of tilt- and mean curvature-excess (plus good terms). Furthermore, since $\int \Delta_M \Psi d\mu = 0$ we obtain

$$\int e_{n+1} \cdot \vec{H} \ \Psi d\mu = \int x_{n+1} \Delta_M \Psi d\mu.$$

Combining this with the Brakke flow equation 3.4 and the evolution equation 4.11, already yields the heat diffusion result 4.15. This can now be used to specify the relation in (1.5). A similar approach is used by Brakke in [B, 6.8, 6.9] but with different calculations, as he is not using the Δ_M at all and therefore relies on Lipschitz approximations to bring in $\Delta_{\mathbb{R}^n}$.

Next we want to obtain an $L^\infty - L^2$-estimate. In section 5 we follow Brakke's approach, using a clearing out lemma inside large balls. There we verify the argumentation that Brakke uses in [B, 6.9. pages 195-196], which is incomplete there, resulting in Lemma 5.14. To do this, we use Lemma 5.7, which is a new version of a clearing out lemma that considers the intersection of a ball and a cylinder.

A second method is based on the more up-to-date $L^\infty - L^2$-estimate 6.8, found in chapter 4 of [E4]. This chapter deals with some consequences of Huisken's monotonicity formula. The results in [E4] are proven for smooth mean curvature flow, but those we need carry over to Brakke flow without great effort, see section 6.

In the slab setting we have small height, if we choose the slab narrow enough. Besides that it is essential to have small tilt- and mean curvature-excess, for Theorem 4.15 to be useful, for Lipschitz approximations to be close and for many other details not mentioned here, to work out. Also, we require bounds on the area ratio of the solution to get any Lipschitz approximations at all. These will all be provided by Theorem 7.7, which is the final result of section 7. We observe that for a varifold lying in a narrow slab, the area in a cylinder $E = R^{-n}\mu_t(C_R(0))$ is decreasing as long as E is in $(h, \omega_n - h) \cup (\omega_n + h, 2\omega_n - h)$. Here $\omega_n = \mathscr{L}^n(B_1^n(0))$ and h is the height of the slab, which is assumed to be small. Moreover the rate of decrease is determined by a certain differential inequality,

$$\overline{D}E(t) \leq -Q^{-1}R^{-2} \min\left\{|E(t)|^{\frac{n-1}{n}}, h^{-\frac{2}{3}}|E(t)|^{\frac{4}{3}}, 1\right\}, \tag{1.7}$$

where Q is a large constant. Solving (1.7) yields the following: If the initial area is smaller then $2\omega_n - h$, and if much later there is area greater than h

left, then in the time in between the area was very close to ω_n. Moreover, if the area stays almost constant for some time there cannot be a lot of mean curvature. This establishes a bound on mean curvature-excess, and by the fact that we assume that the flow stays in a slab, we also get bounds on tilt-excess. This has been done by Brakke in [B, 6.6] and at the beginning of the proof of [B, 6.9]. It was rewritten in [KT] in a more detailed way.

With Theorem 7.7 and the $L^\infty - L^2$-estimate 6.8 established, we can attack local regularity in section 8. Consider a Brakke flow inside a slab with plane-like mass in $[-T, T] \times C_1(0)$. By the local flatness theorem 7.7, there exists a smaller period of time, where we have good Lipschitz approximations and small tilt-excess. Actually, we only have those in an integral sense in time. The slab condition yields $|x_{n+1}| \le h$. Suppose h and δ are very small. Theorem 4.15 can be used to obtain

$$\int_{-15\delta^2}^{\delta^2} \int_{B^n_{4\delta}} \left| \int x_{n+1} \Psi(p, \hat{x} - \hat{y}) d\mu_t - \int x_{n+1} \Psi(p - t + q, \hat{x} - \hat{y}) d\mu_{-q} \right| d\hat{y}\, dt$$
$$\le \delta^6 h,$$

for $p < \delta^2 < q \le \delta^\epsilon$. Furthermore, the error term for the Taylor expansion in (1.5) can be estimated by

$$\left| \sum_{u,v,w} \hat{y}_u \hat{y}_v \int_0^1 \int f(-q, \hat{x}) D^2_{uv} \Psi(p - t + q, \hat{x} - \theta \hat{y}) d\hat{x}\, d\theta \right| \le \delta^2 q^{-1} h,$$

for $y \in B^n_{4\delta}(0)$ and $t \in (-15\delta^2, \delta^2)$. Also (1.3) is stated more precisely as

$$\int_{B^n_{4\delta}} f(0, \hat{y}) d\hat{y} - \int_{B^n_{4\delta}} \int x_{n+1} \Psi(p, \hat{x} - \hat{y}) d\mu_0(x) d\hat{y} \le \delta^3 h$$

for p small enough depending on δ. For $q = \delta^\epsilon$, calculation (1.6) then yields an affine subspace $A_t = \{x_{n+1} = a_t + b_t \cdot \hat{x}\}$ such that

$$\int_{-15\delta^2}^{\delta^2} \int_{B^n_{4\delta}} |f(0, y) - a_t - b_t \cdot y| dy\, dt \le \delta^{2-\epsilon} h. \tag{1.8}$$

Here $a_t = \int f(-q, \hat{x}) \Psi(p - t + q, \hat{x}) d\hat{x}$ and $b_t = \int f(-q, \hat{x}) D\Psi(p - t + q, \hat{x}) d\hat{x}$. By the properties of the heat kernel, we can estimate $|a_t + y \cdot b_t - a_0 - y \cdot b_0| \le \delta^2 q^{-1} h$, so that (1.8) holds with fixed a_0, b_0. Actually, we even have

$$\int_{-15\delta^2}^{\delta^2} \int_{C_{4\delta}} |x_{n+1} - a_0 - b_0 \cdot \hat{x}| d\mu_t\, dt \le \delta^{2-\epsilon} h.$$

Note that $d(x, A_0) \leq |x_{n+1} - a_0 - b_0 \cdot \hat{x}|$, thus we can use Corollary 6.8 to obtain $d(x, A_0) \leq \delta^{2-\epsilon} h$ for all $x \in C_{2\delta}(0) \cap \operatorname{spt}\mu_t$, $t \in [-\delta^2, \delta^2]$. This leads to Lemma 8.1, which says that in the above situation, one can find new coordinates, such that in a δ-smaller region the height is bounded by $\delta^{2-\epsilon} h$. So for every space-time point (t_0, x_0), which is far enough inside the interior of $(-T, T) \times C_R(0)$, we obtain a sequence of contracting slabs containing shrinking neighbourhoods of (t_0, x_0), which finally converge to the tangent space $T_{x_0}\mu_{t_0}$. As a result, Brakke's local regularity theorem 8.4 (originally [B, 6.10]) states that every Brakke flow, which locally lies in a slab and with plane-like mass for a given time interval, is actually a smooth graphical mean curvature flow for some time in a specific subinterval. This is an elaboration of Brakke's [B, 6.9 and 6.10]. Note that in this thesis we only prove local $C^{1,\alpha}$-regularity, for the smoothness and the fact that solutions move by smooth mean curvature flow we refer to [B] and [T]

In section 9, we apply the local regularity result in the general case. The main tool is Lemma 9.5, which can be used to state the following: Provided that there is no jump decrease in area at time $t = 0$, for almost all of the density one points x where the tangent space $T_x\mu_0$ exists, we can find a small neighbourhood where we can apply Brakke's local regularity theorem 8.4. In addition, for almost all of the density zero points of μ_0, we can find a small neighbourhood where we can apply our clearing out lemma 5.7. So at time $t = 0$ almost all points are regular, which is the statement of Brakke's general regularity theorem 9.7 (originally [B, 6.12]). Though most of the calculations appear in [B, 6.12], this is almost a completely new proof incorporating the approach in [E2].

Another interesting observation is the expansion of holes under the Brakke flow which is observed in section 10. We show that for a Brakke flow satisfying an a-priori height-excess bound, the area increase inside a growing cylinder is restricted by that bound, which is a replication of Brakke's [B, 6.5]. In the slab setting this actually says that holes are growing arbitrarily fast if the slab is narrow enough, see Proposition 10.6. A similar result can be found in White's [W3], based on a different approach.

In section 11, we apply Brakke's local regularity theorem 8.4 to Brakke flows starting from a very plane-like varifold. A varifold μ is considered very plane-like in a cylinder $C_R(y)$, if $\mu \cap C_R(y)$ can be written as a graph over $\mathbb{R}^n$ except for a bad set S having small μ-measure. Moreover the graphical part has to satisfy certain height and gradient bounds. (see Definition 11.1). Use of the clearing out lemma, yields that after a short time the flow is contained in a small slab, which will be arbitrarily narrow if the height bound on the graphical part and the bound on the measure of S are small enough. By the flatness of the graphical part, we obtain area ratio bounds close to ω_n there, and if the measure of S is small enough the area ratio of μ_0 in $C_R(y)$ has to be close to ω_n. This also yields a bound on the overall measure of μ_t in smaller cylinders for later times. Then at a certain later time s_0, either there exists a cylinder $C_{2\delta R}(0)$ which contains no μ_{s_0}-measure, or using the clearing out lemma we have a lower area ratio bound for $\mu_{s_0-\epsilon}(C_{\delta R}(0))$. In the second case, by Theorem 8.4, the flow has to be graphical in a small neighbourhood for a small time period in between 0 and $s_0 - \epsilon$. This yields the statement of Theorem 11.7.

In the smooth case with co-dimension-1, these results can be improved, which is something that we do in section 12. It turns out that the gradient bound can be replaced by assuming that the flow is plane-like for a certain time interval. The key result here is Theorem 12.11, which says that a graphical representation can be extended to later times and additionally yields, that the Lipschitz constant of the extended graphical representation is small, if the flow lies in a narrow enough slab.

This thesis was supported by the SFB 647 Raum-Zeit-Materie. I would like to thank my supervisor Klaus Ecker for his advice and patience. Moreover I would like to thank Ahmad Afuni, Ann Björner, Theodora Bourni, Apostolos Damialis, Hanne Hardering, Felix Jachan, Mathew Langford, Tobias Marxen, Ulrich Menne, Kashif Rashul, Oliver Schnürer and Felix Schulze.

Notation We denote the canonical basis of $\mathbb{R}^{n+k}$ by $\mathbf{e}_1, \ldots, \mathbf{e}_{n+k}$. For an n-dimensional subspace T of $\mathbb{R}^{n+k}$ we denote by $\pi_T : \mathbb{R}^{n+k} \to T$ the projection onto T. Set $T^\perp = \{x \in \mathbb{R}^{n+k} : \ x \cdot v = 0 \ \forall v \in T\}$ and $\pi_T^\perp := \pi_{T^\perp}$.

Sometimes we make the following identifications: $\mathbb{R}^n = \mathbb{R}^n \times \{0\}^k$ and $\mathbb{R}^k = \{0\}^n \times \mathbb{R}^k$, also we may identify $\hat{y} \in \mathbb{R}^n$ with $(\hat{y}, 0) \in \mathbb{R}^n \times \{0\}^k$, this should be clear from the context. For $x \in \mathbb{R}^{n+k}$ set $\hat{x} = (\hat{x}, 0) = \pi_{\mathbb{R}^n}(x) = \pi_{\mathbb{R}^n \times \{0\}^k}(x)$.

For $R \in (0, \infty)$ and $x_0 \in \mathbb{R}^{n+k}$ set

$$\begin{aligned} C_R(x_0) &:= \left\{x \in \mathbb{R}^{n+k} : |\hat{x} - \hat{x}_0| < R\right\}, \\ B_R(x_0) &:= \left\{x \in \mathbb{R}^{n+k} : |x - x_0| < R\right\}. \end{aligned}$$

For $m \in \mathbb{N}$, $R \in (0, \infty)$ and $y_0 \in \mathbb{R}^m$ set

$$\begin{aligned} B_R^m(x_0) &:= \{y \in \mathbb{R}^m : |y - y_0| < R\}, \\ S_R^{m-1}(x_0) &:= \{y \in \mathbb{R}^m : |y - y_0| = R\}. \end{aligned}$$

We denote the n-dimensional volume of the unit ball by $\omega_n := \mathscr{L}^n(B_1^n(0))$.

For $m \in \mathbb{N}$ consider $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$ and $A = (a_{ij})_{1 \le i,j \le j} \in \mathbb{R}^{m \times m}$. The matrix A operates on $\mathbb{R}^m$ in the usual way $A(v) := Av = \sum_{ij=1}^m a_{ij} v_j \mathbf{e}_i$. The norm is defined by

$$|v|^2 := \sum_{i=1}^m v_i^2, \qquad |A|^2 := \sum_{ij=1}^m a_{ij}^2.$$

The operator norm is defined by

$$\|A\|_{op} := \sup_{v \in S_1^{m-1}(0)} |A(v)| = \sup_{v \in \mathbb{R}^m \setminus \{0\}} |A(v)||v|^{-1}.$$

For a time dependent function $f : I \times \mathbb{R}^m \to \mathbb{R}$ and a family $(\mu_t)_{t \in I}$ of Radon measures in $\mathbb{R}^m$ we often abbreviate

$$\int f d\mu_t := \int_{\mathbb{R}^m} f(t, x) d\mu_t(x).$$

For a relation $\sim \in \{=, \le, \ge, <, >\}$ we define the Kronecker δ by

$$\delta_{i \sim j} := \begin{cases} 1 & \text{if } i \sim j \\ 0 & \text{else} \end{cases}, \qquad \delta_{ij} := \delta_{i=j} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{else} \end{cases}$$

for $i, j \in \mathbb{N}$.

Quantities that only depend on n and/or k are considered constant. Such a constant may be denoted by C_n, in particular the value of C_n may change in each line. Note that C_n may depend on k. We will not always mention dependence on n and/or k.

2 Varifolds

In the first part of this section we recall the definition of an integral varifold and some of its basic properties. For an excellent introduction to varifolds we refer to [S] or [Sch]. Afterwards we present the cylindrical growth lemma from [B], see Lemma 2.8 . One important feature of varifolds, is that they can be approximated by Lipschitz functions, if they satisfy certain area and height-excess bounds, see Theorem 2.9. Appropriate theorems can be found in [KT], [B] or [Sch].

2.1 Definition. Let μ be a Radon measure on $\mathbb{R}^{n+k}$.

1. For $A \subset \mathbb{R}^{n+k}$ and $\phi \in C_c^0\left(\mathbb{R}^{n+k}, \mathbb{R}\right)$ define the measures $\mu \llcorner A$ and $\mu \llcorner \phi$ by
$$\mu \llcorner A(B) := \mu\left(A \cap B\right),$$
$$\mu \lfloor \phi(B) := \int_B \phi(x) d\mu(x)$$
for every $B \subset \mathbb{R}^{n+k}$. Then $\mu \llcorner A$ and $\mu \llcorner \phi$ are also Radon measures. Also set $\mu(\phi) := \mu\lfloor\phi(\mathbb{R}^{n+k})$.

2. Define the *support* of μ by
$$\text{spt}\mu := \left\{x \in \mathbb{R}^{n+k} : \forall r > 0 \;\; \mu\left(B_r(x)\right) > 0\right\}.$$

3. If $\mu\left(\mathbb{R}^{n+k} \setminus U\right) = 0$ for some $U \subset \mathbb{R}^{n+k}$, we say μ is a Radon measure in U. In particular if μ is a Radon measure in U we have $\text{spt}\mu \subset \overline{U}$.

4. Define the *n-dimensional density* of μ in $x \in \mathbb{R}^{n+k}$ by
$$\Theta^n\left(\mu, x\right) := \lim_{r \searrow 0} \omega_n^{-1} r^{-n} \mu\left(B_r(x)\right),$$
if this limit exists.

5. If for an $x \in \mathbb{R}^{n+k}$ there exists an n-dimensional subspace $T \subset \mathbb{R}^{n+k}$ and a $\theta_x \in (0, \infty)$ such that for all $\phi \in C_c^0\left(\mathbb{R}^{n+k}\right)$
$$\lim_{\lambda \searrow 0} \lambda^{-n} \int_{\mathbb{R}^{n+k}} \phi\left(\frac{z-x}{\lambda}\right) d\mu(z) = \theta_x \int_T \phi(y) d\mathscr{H}^n(y),$$
we call $T_x\mu := T$ the *approximate tangent space* of μ in x with *multiplicity* θ_x. Note that if $T_x\mu$ exists it has to be unique.

2.2 Definition. Let $U \subset \mathbb{R}^{n+k}$.

1. A Radon measure μ in U is called a *rectifiable n-varifold*, if for μ-almost every $x \in \mathbb{R}^{n+k}$ the approximate tangent space exists.

2. A rectifiable n-varifold μ in U is called an *integral n-varifold*, if for μ-almost every $x \in \mathbb{R}^{n+k}$ we have $\Theta^n(\mu, x) \in \mathbb{Z}^+ \cup \{0\}$.

3. An integral n-varifold μ in U is called a *unit density n-varifold*, if for μ-almost every $x \in \mathbb{R}^{n+k}$ we have $\Theta^n(\mu, x) \in \{0, 1\}$.

2.3 Remark. Let $U \subset \mathbb{R}^{n+k}$ and μ be a rectifiable n-varifold in U, i.e. for μ-almost every $x \in \mathbb{R}^{n+k}$ the approximate tangent space $T_x\mu$ exists with multiplicity $\theta_x \in (0, \infty)$.

1. Then $\theta_x = \Theta^n(\mu, x)$ for μ-almost every $x \in \mathbb{R}^{n+k}$, in particular the limit in Definition 2.14 exists. Moreover
$$\lim_{\lambda \searrow 0} \lambda^{-n} \int_{\mathbb{R}^{n+k}} \phi\left(\frac{z-x}{\lambda}\right) d\mu(z) = \Theta^n(\mu, x) \int_{T_x\mu} \phi(y) d\mathscr{H}^n(y), \quad (2.1)$$
for all $\phi \in C_c^0\left(\mathbb{R}^{n+k}\right)$for μ-almost every $x \in \mathbb{R}^{n+k}$.

2. If μ is a unit density n-varifold, then for $\mathscr{H}^n$-almost every $x \in \mathbb{R}^{n+k}$ we either have $\Theta^n(\mu, x) = 1$ and the approximate tangent space exists with multiplicity 1 or $\Theta^n(\mu, x) = 0$

2.4 Definition. For a subset $U \subset \mathbb{R}^{n+k}$ a rectifiable n-varifold μ in U and $x \in \mathrm{int}(U)$ such that the approximate tangent space $T_x\mu$ exists (which is the case for $\mathscr{H}^n$-almost every x) we define

- $\nabla^\mu \phi(x) := \pi_{T_x\mu}(D\phi(x))$ for every $\phi \in C^1(U, \mathbb{R})$.
- $\mathrm{div}_\mu X(x) := \sum_{i=1}^{n+k} \nabla^\mu (X_i(x)) \cdot \mathbf{e}_i$ for every $X \in C^1\left(U, \mathbb{R}^{n+k}\right)$.

This can now be used to define the mean curvature vector on μ

2.5 Definition. Consider an open subset $U \subset \mathbb{R}^{n+k}$, a rectifiable n-varifold μ in U and $M = \{\Theta^n(\mu, x) > 0\}$. Suppose there exists a locally μ-integrable function $\vec{H} : M \to \mathbb{R}^{n+k}$ such that for every $X \in C_c^1\left(U, \mathbb{R}^{n+k}\right)$
$$\int_U \mathrm{div}_\mu X d\mu = -\int_U \vec{H} \cdot X d\mu, \quad (2.2)$$
then $\vec{H}$ is called the *(generalized) mean curvature vector* of μ (in U). Suppose $\vec{H}$ exists, let $\phi \in C^2(U, \mathbb{R})$ and $x \in U$ such that the approximate tangent space $T_x\mu$ exists, then we define the Laplace Beltrami operator in x by
$$\Delta_\mu \phi(x) := \mathrm{div}_\mu(D\phi(x)) + \vec{H}(x) \cdot D\phi(x). \quad (2.3)$$

2.6 Remark. Let $U \subset \mathbb{R}^{n+k}$ be open and μ be a rectifiable n-varifold in U with mean curvature vector $\vec{H}$ and let $\phi \in C^2(U, \mathbb{R})$, $X \in C^1\left(U, \mathbb{R}^{n+k}\right)$

1. If $\mu = \mathscr{H}^n \llcorner N$ for an n-dimensional C^1-manifold N, then $\nabla^\mu \phi(x) = \nabla^N \phi(x)$ and $\mathrm{div}_\mu \phi(x) = \mathrm{div}_N \phi(x)$ for all $x \in N$, where ∇^N and div_N are defined in the usual way for manifolds. If N is a C^2-manifold with $\partial N \cap U = \emptyset$ and μ is a unit density n-varifold, then $\vec{H}$ equals the usual mean curvature vector defined on N.

2. For an open subset $V \subset U$ the measure $\mu \llcorner V$ is a rectifiable n-varifold in V with mean curvature vector $\vec{H} \llcorner V$

3. As in the smooth case for $\upsilon \in C^2(U, \mathbb{R})$ and $f \in C^2(\mathbb{R}, \mathbb{R})$ at points x where the approximate tangent space $T_x\mu$ exists we can calculate

$$\begin{aligned} \mathrm{div}_\mu(\phi X) &= \nabla^\mu \phi \cdot X + \phi \, \mathrm{div}_\mu X, \\ \Delta_\mu(\phi \upsilon) &= \phi \Delta_\mu \upsilon + \upsilon \Delta_\mu \phi + 2 \nabla^\mu \phi \cdot \nabla^\mu \upsilon, \\ \Delta_\mu(f(\phi)) &= f'(\phi) \Delta_\mu \phi + f''(\phi) |\nabla^\mu \phi|^2. \end{aligned}$$

4. For $x \in U$ such that the approximate tangent space $T_x\mu$ exists we can calculate

$$\mathrm{div}_\mu(X(x)) = \sum_{i=1}^{n} (DX(x)\tau_i) \cdot \tau_i,$$

where $(\tau_i)_{1 \le i \le n}$ is an orthonormal basis of $T_x\mu$. In particular for the identity map we have

$$\mathrm{div}_\mu(x) = n.$$

5. For $x \in U$ such that the approximate tangent space $T_x\mu$ exists we can calculate

$$\mathrm{div}_\mu(D\phi(x)) = \Delta_{\mathbb{R}^{n+k}} \phi(x) - \sum_{l=1}^{k} \sum_{i,j=1}^{n+k} \frac{\partial^2 \phi(x)}{\partial x_i \partial x_j} (\nu_l \cdot \mathbf{e}_i)(\nu_l \cdot \mathbf{e}_j),$$

where $(\nu_l)_{1 \le l \le k}$ is an orthonormal basis of $T_x\mu^\perp$.

6. For $X \in C^1\left(U, \mathbb{R}^{n+k}\right)$ with $\mathrm{spt}\mu \cap \mathrm{spt}X \subset\subset U$ equality (2.2) holds.

7. If $\mathrm{spt}\mu \cap \mathrm{spt}\phi \subset\subset U$ equalities (2.2) and (2.3) yield

$$\int_U \Delta_\mu \phi \, d\mu = 0.$$

In the smooth case the mean curvature vector is always normal to the surface. This fact carries over to integral n-varfolds as proven by Brakke in [B, 5.8]. We will frequently use this fact.

2.7 Theorem (Perpendicularity Of Mean Curvature, [B, 5.8])**.** *Let $U \subset \mathbb{R}^{n+k}$ open and μ an integral n-varifold in U with mean curvature vector $\vec{H}$, then*

$$\vec{H} \perp T_x\mu \tag{2.4}$$

for μ almost every $x \in U$.

Now we can estimate how changing the radius of a cylinder varies its measure. This lemma is taken from [B, 6.4]. We filled in some details for the proof.

2.8 Lemma (Cylindrical Growth Rates, [B, 6.4])**.** *Let $R_2, \alpha_0, \beta_0 \in (0,\infty)$, $R_1 \in (0, R_2)$, $U \subset \mathbb{R}^{n+k}$ open and μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$ and*

$$\mathrm{spt}\mu \cap C_{R_2}(0) \subset\subset U. \tag{2.5}$$

For $\phi \in C_c^3([-1,1], \mathbb{R}^+)$ and $\rho \in [r, R_2]$ set

$$\rho^{-n} \int_{C_\rho(0)} |\vec{H}(x)|^2 \Phi_\rho(x) d\mu(x) =: \tilde{\alpha}_\phi(\rho)^2, \tag{2.6}$$

$$\rho^{-n} \int_{C_\rho(0)} |\pi_{T_xM} - \pi_{\mathbb{R}^n}|^2 \, \Phi_\rho(x) \, d\mu(x) =: \beta_\phi(\rho)^2, \tag{2.7}$$

where $\Phi_\rho(x) := \phi(\rho^{-1}|\hat{x}|)$. Then the following holds:

1. *Suppose $\tilde{\alpha}_\phi(\rho) \leq \alpha_0$ and $\beta_\phi(\rho) \leq \beta_0$ for all $\rho \in [R_1, R_2]$, then*

$$\begin{aligned}&\left| R_2^{-n} \int_{C_{R_2}(0)} \Phi_{R_2} \, d\mu - R_1^{-n} \int_{C_{R_1}(0)} \Phi_{R_1} \, d\mu \right| \\ &\leq n\beta_0^2 \log\left(R_1^{-1} R_2\right) + \alpha_0\beta_0(R_2 - R_1) + 2\beta_0^2.\end{aligned}$$

2. *Assume ϕ is monotonically non-increasing on $[0,1]$. Suppose $\tilde{\alpha}_\phi(R_2) \leq R_2^{-1}\alpha_0$ and $\beta_\phi(R_2) \leq \beta_0$, then*

$$\begin{aligned}&\left| R_2^{-n} \int_{C_{R_2}} \Phi_{R_2} \, d\mu - R_1^{-n} \int_{C_{R_1}} \Phi_{R_1} \, d\mu \right| \\ &\leq R_1^{-n} R_2^n \left(n\beta_0^2 \log\left(R_1^{-1} R_2\right) + \alpha_0\beta_0 R_2^{-1}(R_2 - R_1) + 2\beta_0^2 \right).\end{aligned}$$

Proof. Set $U = C_{R_2}(0)$, at $x \in \mathrm{spt}\mu$ where $T_x\mu$ exists set

$$\pi_n := \pi_{\mathbb{R}^n}, \quad \pi_k := \pi_{\mathbb{R}^k}, \quad \pi_x := \pi_{T_x\mu}, \quad \pi_x^\perp := \pi_{T_x\mu^\perp}.$$

For $x \in \mathrm{spt}\mu$ for which the approximate tangent space $T_x\mu$ exists calculate

$$\begin{aligned} \mathrm{div}_\mu\left(\Phi_\rho(x)\hat{x}\right) &= \sum_{i=1}^{n+k} \mathbf{e}_i \cdot \left(\nabla^M\left(\Phi_\rho(x)\hat{x}_i\right)\right) \\ &= \sum_{i=1}^{n} \mathbf{e}_i \cdot \left(\hat{x}_i\ \nabla^M \Phi_\rho(x) + \Phi_\rho(x)\pi_x(\mathbf{e}_i)\right) \\ &= \hat{x} \cdot \nabla^M \Phi_\rho(x) + \Phi_\rho(x) \sum_{i=1}^{n} \mathbf{e}_i \cdot \pi_x(\mathbf{e}_i) \\ &= \pi_x(\hat{x}) \cdot D\Phi_\rho(x) + \Phi_\rho(x) \left(n - \sum_{j=n+1}^{n+k} |\pi_x(\mathbf{e}_j)|^2\right). \end{aligned} \tag{2.8}$$

In the last step we used $\sum_{i=1}^{n+k} |\pi_x(\mathbf{e}_i)|^2 = n$ as $T_x\mu$ is an n-dimensional subspace. With the tilt-excess characterization from Remark A.11 equality (2.8) becomes

$$\mathrm{div}_\mu\left(\Phi_\rho(x)\hat{x}\right) = \pi_x(\hat{x}) \cdot D\Phi_\rho(x) + \Phi_\rho(x)\left(n - \frac{1}{2}|\pi_n - \pi_x|^2\right) \tag{2.9}$$

for all $\rho \in (0, R_2]$ and all $x \in \mathrm{spt}\mu$ where $T_x\mu$ exists.

By definition of the mean curvature vector (2.2) we have

$$\int_U \mathrm{div}_\mu\left(\Phi_\rho(x)\hat{x}\right) d\mu = -\int_U \Phi_\rho(x)\hat{x} \cdot \vec{H}(x) d\mu(x) \tag{2.10}$$

for all $\rho \in (0, R_2]$. Here we used that $\mathrm{spt}\phi \subset\subset [-1, 1]$ and $\mathrm{spt}\mu \cap C_{R_2}(0) \subset\subset U$, so $\mathrm{spt}\mu \cap \mathrm{spt}\Phi_\rho \subset\subset U$. By Theorem 2.7 we can use Remark A.7.1 for the term on the right of (2.10) to estimate

$$\begin{aligned} \left|\int_U \mathrm{div}_\mu\left(\Phi_\rho(x)\hat{x}\right) d\mu\right| &= \left|\int_U \Phi_\rho(x)\left(\pi_n(x) - \pi_x(x)\right) \cdot \vec{H}(x) d\mu(x)\right| \\ &\le \int_U \Phi_\rho(x)|\hat{x}|\,|\pi_n - \pi_x|\,|\vec{H}(x)| d\mu(x) \\ &\le \rho \int_U \Phi_\rho(x)\,|\pi_n - \pi_x|\,|\vec{H}(x)| d\mu(x) \end{aligned}$$

for all $\rho \in (0, R_2]$. Now Combine this with (2.9) and use Hölders inequality to obtain

$$\begin{aligned}
&\rho^{-1}\int_U \left(\pi_x(\hat{x})\cdot D\Phi_\rho(x) + \Phi_\rho(x)\left(n - \frac{1}{2}\left|\pi_n(x)-\pi_x(x)\right|^2\right)\right) d\mu(x) \\
&\leq \int_U \Phi_\rho(x)\left|\pi_n - \pi_x\right| |\vec{H}_t(x)| d\mu(x) \\
&\leq \left(\int_U \Phi_\rho(x)\left|\pi_n-\pi_x\right|^2 d\mu(x) \int_U \Phi_\rho(x)|\vec{H}(x)|^2 d\mu(x)\right)^{\frac{1}{2}}
\end{aligned}$$

for all $\rho \in (0, R_2]$. Then by definitions (2.6) and (2.7) and the assumed bounds on $\tilde{\alpha}_\phi(\rho)$ and $\beta_\phi(\rho)$

$$\begin{aligned}
&\rho^{-1}\int_U \left(\pi_x(\hat{x})\cdot D\Phi_\rho(x) + \Phi_\rho(x)\left(n - \frac{1}{2}\left|\pi_n(x)-\pi_x(x)\right|^2\right)\right) d\mu(x) \\
&\leq \left(\tilde{\alpha}_\phi(\rho)^2\rho^n\beta_\phi(\rho)^2\rho^n\right)^{\frac{1}{2}} \leq \alpha_0\beta_0\rho^n
\end{aligned} \tag{2.11}$$

for all $\rho \in [R_1, R_2]$.

As $\Phi_\rho(x) = \phi\left(\rho^{-1}|\hat{x}|\right)$ we have

$$D\Phi_\rho(x) = \phi'\left(\rho^{-1}|\hat{x}|\right)\frac{\hat{x}}{\rho|\hat{x}|} = \phi'\left(\rho^{-1}|\hat{x}|\right)\frac{|\hat{x}|}{\rho^2}\frac{\rho\hat{x}}{|\hat{x}|^2} = -\frac{\partial}{\partial\rho}\left(\Phi_\rho(x)\right)\frac{\rho\hat{x}}{|\hat{x}|^2}$$

for all $\rho \in (0, R_2]$ and all $x \in \mathbb{R}^{n+k}$. Then we can calculate

$$\begin{aligned}
&\left|\frac{\partial}{\partial\rho}\left(\frac{1}{\rho^n}\int_U \Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right)\right| \\
&= \rho^{-n}\left|\int_U \frac{\partial}{\partial\rho}\left(\Phi_r(x)\right)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x) - \frac{n}{\rho}\int_U \Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right| \\
&= \rho^{-n-1}\int_U \left(\pi_x(\hat{x})\cdot D\Phi_\rho(x) + n\Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right) d\mu
\end{aligned}$$

and combining this with (2.11) establishes

$$\begin{aligned}
&\left|\frac{\partial}{\partial\rho}\left(\frac{1}{\rho^n}\int_U \Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right)\right| \\
&\leq \alpha_0\beta_0 + \rho^{-n-1}\int_U \Phi_\rho(x)\left|n - \frac{1}{2}\left|\pi_n-\pi_x\right|^2 - n\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right| d\mu
\end{aligned} \tag{2.12}$$

for all $\rho \in [R_1, R_2]$. By Remark A.14 we obtain

$$\left|\frac{1}{2}\left|\pi_n-\pi_x\right|^2 - n\left(1-\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right)\right| \leq n\left|\pi_n-\pi_x\right|^2 \tag{2.13}$$

for all $x \in \operatorname{spt}\mu$ where $T_x\mu$ exists. Here we used $|a-b| \leq \max\{a,b\}$ for $a, b > 0$. Then by (2.7) and $\beta_\phi(\rho) \leq \beta_0$

$$\begin{aligned}
&\rho^{-n-1}\int_U \Phi_\rho(x)\left|n - \frac{1}{2}\left|\pi_n(x)-\pi_x(x)\right|^2 - n\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right| d\mu \\
&\leq n\rho^{-n-1}\int_U \Phi_\rho(x)\left|p(x)-p_x(x)\right|^2 d\mu(x) = n\beta_\phi(\rho)^2\rho^{-1} \leq n\beta_0^2\rho^{-1}
\end{aligned}$$

for all $\rho \in [R_1, R_2]$. Inserting this into (2.12) produces

$$\left|\frac{\partial}{\partial r}\left(\frac{1}{\rho^n}\int_U \Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right)\right| \leq \alpha_0\beta_0 + \frac{n}{\rho}\beta_0^2$$

for all $\rho \in [R_1, R_2]$ and integrating with respect to ρ then yields

$$\left|\left[\frac{1}{\rho^n}\int_U \Phi_\rho(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right]_{R_1}^{R_2}\right| \leq \alpha_0\beta_0(R_2-R_1) + n\beta_0^2\log\left(\frac{R_2}{R_1}\right). \tag{2.14}$$

Using again Remark A.14 we can estimate $|1-|\hat{x}|^{-2}|\pi_x(\hat{x})|^2| \leq |\pi_n - \pi_x|^2$. Combined with (2.7) and $\beta_\phi(\rho) \leq \beta_0$ we obtain

$$\rho^{-n}\int_U \Phi_\rho(x)\left|1 - \frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right| d\mu(x) \leq \beta_0^2 \tag{2.15}$$

for all $\rho \in [R_1, R_2]$. Combining (2.14) and (2.15) finally establishes

$$\begin{aligned}
&\left|R_2^{-n}\int_U \Phi_{R_2}\, d\mu - R_1^{-n}\int_U \Phi_{R_1}\, d\mu\right| \\
&\leq \left|R_2^{-n}\int_U \Phi_{R_2}(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x) - R_1^{-n}\int_U \Phi_{R_1}(x)\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}d\mu(x)\right| \\
&\quad + 2\max_{\rho\in\{R_1,R_2\}}\rho^{-n}\int_U \Phi_\rho(x)\left|1-\frac{|\pi_x(\hat{x})|^2}{|\hat{x}|^2}\right| d\mu(x) \\
&\leq \alpha_0\beta_0(R_2-R_1) + n\beta_0^2\log\left(R_1^{-1}R_2\right) + 2\beta_0^2.
\end{aligned}$$

For statement 2. note that as ϕ is monotonically non-increasing on $[0,1]$ we have $\Phi_\rho(x) \leq \Phi_{R_2}(x)$ for all $x \in \mathbb{R}^{n+k}$ for all $\rho \in (0, R_2]$. This lets us estimate

$$\begin{aligned}
(\tilde{\alpha}_\phi(\rho))^2 &\leq \rho^{-n}R_2^n(\tilde{\alpha}_\phi(R_2))^2 \leq R_1^{-n}R_2^n(\tilde{\alpha}_\phi(R_2))^2 \leq R_1^{-n}R_2^n(R_2^{-1}\alpha_0)^2 \\
(\beta_\phi(\rho))^2 &\leq \rho^{-n}R_2^n(\beta_\phi(R_2))^2 \leq R_1^{-n}R_2^n(\beta_\phi(R_2))^2 \leq R_1^{-n}R_2^n\beta_0^2
\end{aligned}$$

for all $\rho \in [R_1, R_2]$. Thus we can use statement 1. with α_0 replaced by $R_1^{-\frac{n}{2}}R_2^{-\frac{n}{2}-1}\alpha_0$ and β_0 replaced by $R_1^{-\frac{n}{2}}R_2^{-\frac{n}{2}}\beta_0$ which yields statement 2. $\square$

For the local regularity iteration lemma 8.1 we need a Lipschitz approximation for integral varifolds like in [B, 5.4]. Note that we only need the "one sheet"-case.

2.9 Theorem (Lipschitz Approximation, [B, 5.4])**.** *For all $l, \lambda \in (0,1)$ there exist $C \in (1,\infty), \gamma_0 \in (0,1)$ such that for all $R, \alpha, \beta \in (0,\infty)$ and $\gamma \in [0,\gamma_0]$ the following holds: Let μ be an integral n-varifold in $B_{7R}(0)$ with L^2-integrable mean curvature vector $\vec{H}$ and suppose μ satisfies:*

$$(3R)^{-n}\mu\left(B_{3R}(0)\right) \le (2-\lambda)\omega_n \tag{2.16}$$

$$R^{-n}\mu\left(B_R(0)\right) \ge \lambda\omega_n \tag{2.17}$$

$$R^{-n+2}\int_{B_{7R}(0)} |\vec{H}(x)|^2 d\mu(x) \le \alpha^2 \tag{2.18}$$

$$R^{-n}\int_{B_{7R}(0)} \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 d\mu(x) \le \beta^2 \tag{2.19}$$

$$R^{-n-2}\int_{B_{7R}(0)} \left|\pi_{\mathbb{R}^k}(x)\right|^2 d\mu(x) \le \gamma^2. \tag{2.20}$$

Then there exists a Lipschitz map $f : B^n_R(0) \to \mathbb{R}^k$ with

$$\operatorname{lip}(f) \le l, \quad \sup|f(y)| \le C\gamma^{\frac{2}{n+2}}R \tag{2.21}$$

such that for

$$Y := \left\{\hat{y} \in B^n_R(0) : \ f(\hat{y}) \in B^k_R(0),\ \Theta^n\left(\mu, (\hat{y}, f(\hat{y}))\right) = 1\right\} \tag{2.22}$$

$$X := \left\{x \in B^n_R(0) \times B^k_R(0) : \exists \hat{y} \in Y\ x = (\hat{y}, f(\hat{y}))\right\} \tag{2.23}$$

we can estimate

$$\mu\left(B^n_R(0) \times B^k_R(0) \setminus X\right) + \mathscr{L}^n\left(B^n_R(0) \setminus Y\right) \le CR^nE, \tag{2.24}$$

where $E := \left(\alpha^{\frac{2n}{n-2}}\delta_{n\ge3} + \beta^2 + \gamma^2\right)$. Here $\delta_{n\ge3} := 1$, if $n \ge 3$ and 0 otherwise.

2.10 Remark. In the above statement we actually can choose

$$\tilde{Y} := \left\{\hat{y} \in Y : \ T_{F(\hat{y})}\mu \text{ and } T_{F(\hat{y})}\tilde{\mu} \text{ exist with } T_{F(\hat{y})}\mu = T_{F(\hat{y})}\tilde{\mu}\right\}, \tag{2.25}$$

where $F(\hat{x}) = (\hat{x}, f(\hat{x}))$ and $\tilde{\mu} = \mathscr{H}^n \llcorner \operatorname{graph}(f)$. Then $\mathscr{L}^n(Y \setminus \tilde{Y}) = 0$, such that (2.24) still holds for Y replaced by $\tilde{Y}$.

2.11 Remark. Theorem 2.9 can be used to estimate the integral over the varifold by the integral over the graph of f. Set $F(\hat{x}) = (\hat{x}, f(\hat{x}))$, and consider an L^1-integrable function

$$\phi : B^n_R(0) \times B^k_R(0) \cap [\mathrm{spt}\mu \cup \mathrm{graph}(f)] \to \mathbb{R}$$

then we have

$$\left| \int_{B^n_R(0)\times B^k_R(0)} \phi(x) d\mu(x) - \int_{B^n_R(0)} \phi(F(\hat{x})) JF(\hat{x}) d\mathscr{L}^n(\hat{x}) \right| \leq C_n R^n \sup |\phi| \, E.$$

If $l \leq 1$ we even have

$$\left| \int_{B^n_R(0)\times B^k_R(0)} \phi(x) d\mu(x) - \int_{B^n_R(0)} \phi(F(\hat{x})) d\mathscr{L}^n(\hat{x}) \right| \leq C_n R^n \sup |\phi| \, E.$$

Here in both estimates $\sup |\phi|$ is the essential supremum of $|\phi|$ over the set $B^n_R(0) \times B^k_R(0) \cap [\mathrm{spt}\mu \cup \mathrm{graph}(f)]$.

Proof. For a proof of Theorem 2.9 see [B, 5.4] or [Sch, 18.1]. Remark 2.10 is from [Sch, 18.2]. Remark 2.11 we will prove below. By (2.22) and (2.23) we have

$$\int_X \phi(x) d\mu(x) = \int_Y \phi(F(\hat{y})) JF(\hat{y}) d\mathscr{L}^n(\hat{y}).$$

So we can estimate

$$\begin{aligned}
&\left| \int_{B^n_R(0)\times B^k_R(0)} \phi(x) d\mu(x) - \int_{B^n_R(0)} \phi(F(\hat{x})) JF(\hat{x}) d\mathscr{L}^n(\hat{x}) \right| \\
&= \left| \int_{B^n_R(0)\times B^k_R(0)\setminus X} \phi(x) d\mu(x) - \int_{B^n_R(0)\setminus Y} \phi(F(\hat{x})) JF(\hat{x}) d\mathscr{L}^n(\hat{x}) \right| \\
&\leq \sup |\phi| \, \mu\left(B^n_R(0) \times B^k_R(0) \setminus X\right) + \sup |\phi| \, \mathscr{L}^n\left(B^n_R(0) \setminus Y\right).
\end{aligned}$$

Then the first inequality follows with (2.24).

For the second inequality use Remark A.12 to estimate

$$|1 - JF(\hat{x})| \leq C_n |\pi_{T_{F(\hat{x})}\tilde{\mu}} - \pi_{\mathbb{R}^n}|^2 \tag{2.26}$$

for almost every $\hat{x} \in B^n_R(0)$, where $\tilde{\mu} = \mathscr{H}^n \llcorner \mathrm{graph}(f)$. In particular this is bounded by a constant, so with (2.24) we have

$$\begin{aligned}
&\int_{B^n_R(0)} |\phi(F(\hat{x}))(1 - JF(\hat{x}))| \, d\mathscr{L}^n(\hat{x}) \\
&\leq \int_Y |\phi(F(\hat{x}))(1 - JF(\hat{x}))| \, d\mathscr{L}^n(\hat{x}) + C_n R^n \sup |\phi| \, E.
\end{aligned} \tag{2.27}$$

Using (2.26) we can also estimate

$$\int_Y |1 - JF(\hat{x})| \, d\mathscr{L}^n(\hat{x}) \leq C_n \int_Y |\pi_{T_{F(\hat{x})}\tilde{\mu}} - \pi_{\mathbb{R}^n}|^2 d\mathscr{L}^n(\hat{x}).$$

Then with Remark 2.10 and assumption (2.19) we obtain

$$\int_Y |1 - JF(\hat{x})| \, d\mathscr{L}^n(\hat{x}) \leq C_n \int_{B_{7R}(0)} |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, d\mu(x) \leq C_n \beta^2 R^n,$$

where we also used graph$f \subset B^n_R(0) \times B^k_R(0) \subset B_{7R}(0)$ and $JF \geq 1$. Then with (2.24) and $\beta^2 \leq E$

$$\int_Y |\phi(F(\hat{x}))(1 - JF(\hat{x}))| \, d\mathscr{L}^n(\hat{x}) \leq C_n R^n \sup |\phi| \, E.$$

Now combine this with the first statement of Remark 2.11 and (2.27) to establish the second statement. □

3 Brakke Flow

In this section we define the Brakke flow via the Brakke variation. Afterwards we derive the almost continuity property of this flow. Then we deal with the behaviour of time dependent test functions integrated over a Brakke flow. In the end we use barrier functions, to see how area and height bounds at a starting time transfer to later times. Most of this is from [B] chapter 2-3, though we had to change the order slightly.

A Brakke flow will be a family of Radon measures which satisfies inequality (1.2) in a generalized interpretation. This shall be made precise below. In the previous section we introduced the mean curvature vector on varifolds, this suggests the following definition for the right hand side of (1.2):

3.1 Definition. For an open subset $U \subset \mathbb{R}^{n+k}$, a Radon measure μ in U and $\phi \in C^{0,1}\left(U, \mathbb{R}^+\right)$, we define the Brakke varitation $\mathscr{B}(\mu, \phi)$ as follows.

- If μ is a rectifiable n-varifold with mean curvature vector $\vec{H}$ and $\vec{H}$ is L^2-integrabel on μ we set

$$\mathscr{B}\left(\mu, \phi\right) := -\int_U \phi|\vec{H}|^2 d\mu + \int_U \pi^{\perp}_{T_x\mu}\left(D\phi\right) \cdot \vec{H} d\mu,$$

  this is called the *non-singular case.*

- else we set $\mathscr{B}\left(\mu, \phi\right) := -\infty$, this is called the *singular case.*

3.2 Remark.

1. For an integral n-varifold μ we can write for the Brakke variation in the non-singular case

$$\mathscr{B}\left(\mu, \phi\right) = -\int_U \phi|\vec{H}|^2 d\mu + \int_U D\phi \cdot \vec{H} d\mu.$$

   Here we used $\vec{H}(x) \perp T_x\mu$ for μ-almost every x in U, by Theorem 2.7.

2. For $\phi \in C^2_c\left(U, \mathbb{R}^+\right)$ we can estimate $\phi^{-1}|D\phi|^2 \leq 2|D^2\phi|$ on $\{\phi > 0\}$ and setting $\phi^{-1}(x)|D\phi(x)|^2 = 0$ outside $\{\phi > 0\}$ yields a continuous function, see Proposition A.6. Then with Young's inequality we can estimate for the Brakke variation in the non-singular case

$$\begin{aligned}\mathscr{B}\left(\mu, \phi\right) &\leq -\int_U \phi|\vec{H}|^2 d\mu + \int_U \phi|\vec{H}|^2 + \frac{|D\phi|^2}{4\phi} d\mu \\ &\leq \int_U \frac{|D\phi|^2}{4\phi} d\mu \leq \int_U \frac{1}{2}|D^2\phi| d\mu.\end{aligned}$$

   Note that in the singular case we trivially have $-\infty \leq \int_U \frac{|D\phi|^2}{4\phi} d\mu$.

Next we consider the left hand side of (1.2). Usually the integral of a test function over varying varifolds will not be differentiable, but the upper derivative will exist, that is:

3.3 Definition. For a function $f : (a,b) \to \mathbb{R}$ and a point $t_0 \in (a,b)$, the upper derivative from the right is defined by

$$\overline{D}_t f(t_0) := \limsup_{h \to 0} \; h^{-1} \left(f(t_0 + h) - f(t_0) \right)$$

We allow this to be $\pm\infty$, so it always exists.

Now we can define the Brakke flow by:

3.4 Definition. For $t_1 \in \mathbb{R}, t_2 \in (t_1, \infty)$ and an open subset $U \subset \mathbb{R}^{n+k}$ let $(\mu_t)_{t \in [t_1,t_2]}$ be a family of Radon measures in U. We call $(\mu_t)_{t \in [t_1,t_2]}$ a *Brakke flow* in U, if the following holds

1. For every test function $\phi \in C_c^{0,1}(U, \mathbb{R}^+)$ we have for all $t \in [t_1, t_2]$
$$\overline{D}_t \mu_t(\phi) \leq \mathscr{B}(\mu_t, \phi), \tag{3.1}$$
where $\mathscr{B}$ is the Brakke variation from Definition 3.1.
2. For almost every $t \in (t_1, t_2)$ the Radon measure μ_t is an integral n-varifold.

3.5 Remark.

1. Note that by definition a Brakke flow $(\mu_t)_{t \in [t_1,t_2]}$ in U satisfies $\mu_t\left(\mathbb{R}^{n+k} \setminus U\right) = 0$ for all $t \in [t_1, t_2]$. For an open subset $V \subset U$ the restriction $(\mu_t \llcorner V)_{t \in [t_1,t_2]}$ is a Brakke flow in V, although (μ_t) itself may be not. Of course you could define a Brakke flow in U for measures with support outside U, but all expressions we use only consider the restriction to U, which we do not want to write all the time.
2. Let $\phi \in C^1\left(\mathbb{R}^{n+k}, \mathbb{R}^+\right)$ with $K := \bigcup_{t \in [t_1,t_2]} \mathrm{spt}\mu_t \cap \mathrm{spt}\phi \subset\subset U$, then inequality (3.1) holds for all $t \in [t_1, t_2]$. To see this multiply ϕ with a cut-off function $\zeta \in C_c^\infty(U, [0,1])$, which satisfies $K \subset \{\zeta = 1\}$. Then (3.1) holds for $\phi\zeta \in C_c^1(U, \mathbb{R}^+)$ and as $\zeta \equiv 1$ in the terms of (3.1), we can ignore the ζ.
3. A Brakke flow $(\mu_t)_{t \in [t_1,t_2]}$ can always be extended to the time interval $[t_1, t_2 + T]$ for arbitrary $T \in (0, \infty)$ by setting $\mu_t \equiv 0$ for all $t \in (t_2, t_2 + T]$. By [B, 4.29] there actually exist non-trivial Brakke flows for any initial integral n-varifold μ_{t_1}.

4. Usually condition 2 is not included in the definition, so we basically look at integral Brakke flows but omit to write integral. Some of the results still hold without assuming condition 2.

3.6 Lemma. *For all open subsets $U \subset \mathbb{R}^{n+k}$ and all $t_1 \in \mathbb{R}, t_2 \in (t_1, \infty)$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U and $V \subset\subset U$, then there exists an $M \in (1, \infty)$ with $\mu_t(V) \le M$ for all $t \in [t_1, t_2]$. Note that M depends on V and $(\mu_t)_{t\in[t_1,t_2]}$.*

Proof. Fix a $\delta \in (0, 1)$ and extend the flow to the time interval $[t_1, t_2 + \delta]$ by setting $\mu_t \equiv 0$ for all $t \in (t_2, t_2 + \delta]$. The extended flow than still is a Brakke flow.

As V is compactly contained in U there exists a $\phi \in C^2_c(U, [0, 1])$ with $\phi(x) = 1$ for all $x \in V$. Now let $t \in [t_1, t_2]$ be arbitrary. By the definition of upper derivative from the right we find $h_t \in (0, \delta)$ such that $h^{-1} (\mu_{t+h}(\phi) - \mu_t(\phi)) \le \overline{D}_t\mu_t(\phi) + 1$ for all $h \in (0, h_t]$. Using Definition 3.4 and Remark 3.2.2 we can estimate

$$\begin{aligned} \mu_{t+h}(\phi) &\le \mu_t(\phi) + h\left(\mathscr{B}(\mu_t, \phi) + 1\right) \\ &\le \mu_t(\phi) + 2h\left(\int_U |D^2\phi| d\mu_t + 1\right) \le \mu_t(\phi) + hM_t \end{aligned} \tag{3.2}$$

for all $h \in (0, h_t]$ and $M_t := 2\left(\int_U |D^2\phi| d\mu_t + 1\right) \in [0, \infty)$. This yields a covering $[t_1 + h_{t_1}, t_2] \subset \bigcup_{t\in[t_1,t_2]}(t, t + h_t)$, so by compactness there exist $s_2, \dots, s_N$ with $[t_1 + h_{t_1}, t_2] \subset \bigcup_{i=2}^N (s_i, s_i + h_i)$, for an $N \in \mathbb{N}$ and $h_i = h_{s_i}$. Set $s_1 := t_1$ to obtain $(t_1, t_2] \subset \bigcup_{i=1}^N (s_i, s_i + h_i)$.

W.l.o.g. assume $s_i < s_{i+1}$ for $i = 1, \dots, N - 1$. By iteration of (3.2) we can estimate for every $i = 1, \dots, N$

$$\mu_t(\phi) \le \mu_{t_1}(\phi) + \sum_{j=1}^{i} M_{s_j}$$

for all $t \in (s_i, s_i + h_i)$. Here we used $h_i \le 1$ for all $i = 1, \dots, N$. By the covering feature this yields

$$\mu_t(\phi) \le \mu_{t_1}(\phi) + M$$

for all $t \in (t_1, t_2]$, where $M := \mu_{t_1}(V) + \sum_{i=1}^N M_{s_i}$. Then the result follows by $V \subset \mathrm{spt}\phi$. □

We can use Lemma 3.6 to derive some of the continuity properties in [B, 3.10]. Note that we do not need barrier functions here.

3.7 Proposition (Continuity Properties, [B, 3.10]). *Consider an open subset $U \subset \mathbb{R}^{n+k}$ and $t_1 \in \mathbb{R}, t_2 \in (t_1, \infty)$. Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U and $\phi \in C_c^2(U, \mathbb{R}^+)$ then the following holds:*

1. $\overline{D}_t\mu_t(\phi) := \limsup_{h\to 0} h^{-1}(\mu_{t+h}(\phi) - \mu_t(\phi)) \le L < \infty$ *for all* $t \in [t_1, t_2]$, *where* $L \in \mathbb{R}$ *may depend on* ϕ *as well as the whole flow* (μ_t).
2. $\lim_{\delta\searrow 0} \mu_{t+\delta}(\phi) \le \mu_t(\phi) \le \lim_{\delta\searrow 0} \mu_{t-\delta}(\phi)$ *for all* $t \in [t_1, t_2]$
3. *There exists a countable set* $S \subset [t_1, t_2]$ *such that for all* $t \in [t_1, t_2] \setminus S$
$$\lim_{h\to 0} \mu_{t+h}(\vartheta) = \mu_t(\vartheta) \quad \forall \vartheta \in C_c^0(U).$$

Proof. Consider $\phi \in C_c^2(U, \mathbb{R}^+)$. By regularity of ϕ there exists $L_1 \in (1, \infty)$ such that $\sup |D^2\phi| \le L_1$. Set $V := \operatorname{spt}\phi$, then Lemma 3.6 yields an $L_2 \in (1, \infty)$ such that $\mu_t(V) \le L_2$ for all $t \in [t_1, t_2]$. Using Definition 3.4 and Remark 3.2 we can estimate

$$\overline{D}_t\mu_t(\phi) \le \mathscr{B}(\mu_t, \phi) \le \int_U |D^2\phi| d\mu_t \le L_1 L_2 =: L$$

for all $t \in [t_1, t_2]$. Then by Proposition A.19 we have $\mu_{s_2}(\phi) - \mu_{s_1}(\phi) \le (s_2 - s_1)L$ for all $t_1 \le s_1 < s_2 \le t_2$. This establishes statement 1.

Now consider $g : [t_1, t_2] \to \mathbb{R}$ given by $g(t) := \mu_t(\phi) - Lt$. Then g is monotonically non-increasing and $\lim_{\delta\searrow 0} g(t + \delta)$ as well as $\lim_{\delta\searrow 0} g(t - \delta)$ always exist, so $\lim_{\delta\searrow 0} g(t+\delta) \le g(t) \le \lim_{\delta\searrow 0} g(t-\delta)$, which gives statement 2.

For the continuity let $\phi \in C_c^2(U, \mathbb{R}^+)$ and consider again $g(t) := \mu_t(\phi) - Lt$. Denote by $S \subset [t_1, t_2]$ the set of times t where g is not continuous. As g is monotonically non-increasing it "jumps down" at all discontinuities in t, so we can decompose $S = \bigcup_{n=1}^{\infty} S_{\frac{1}{n}}$, where S_ϵ is the set of times where g abruptly decreases by at least ϵ

$$S_\epsilon := \left\{ t \in [t_1, t_2] : \lim_{\delta\searrow 0} \mu_{t+\delta} + \epsilon < \lim_{\delta\searrow 0} \mu_{t-\delta}(\phi) \right\}.$$

By the monotonicity of g each of these S_ϵ consists of discrete points, so S has to be countable. Thus we proved that there exists a countable set $S_\phi \subset [t_1, t_2]$ such that $t \to \mu_t(\phi)$ is a continuous function for all $t \in [t_1, t_2] \setminus S$

Now take a countable $A \subset C_c^2(U, \mathbb{R}^+)$ such that A is dense in $C_c^0(U, \mathbb{R}^+)$. Then for almost every $t \in [t_1, t_2]$ we know that $\lim_{h\to 0} \mu_{t+h}(\phi) = \mu_t(\phi) \ \ \forall \phi \in A$. For $\vartheta \in C_c^0(U, (-\infty, \infty))$ we can write $\vartheta = \vartheta^+ - \vartheta^-$ where $\vartheta^+(x) := \max\{0, \vartheta(x)\}$, $\vartheta^-(x) := \max\{0, -\vartheta(x)\}$. Then approximating ϑ^+, ϑ^- by functions from A yields the result. $\square$

Now we can consider the behaviour of time varying test functions. This is basically [B, 3.5].

3.8 Proposition (Time Varying Test Functions, [B, 3.5])**.** *For all open subsets $V, U \subset \mathbb{R}^{n+k}$, every $t_1 \in \mathbb{R}, t_2 \in (t_1, \infty)$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U and $\phi \in C^1([t_1,t_2] \times V, \mathbb{R}^+)$, $\phi_t := \phi(t,\cdot)$ with $\phi_t \in C^1_c(V, \mathbb{R}^+)$ for all $t \in [t_1,t_2]$ and*

$$\bigcup_{t\in[t_1,t_2]} \mathrm{spt}\mu_t \cap V \subset\subset U. \tag{3.3}$$

Then we have for almost every $s \in (t_1, t_2)$

$$\overline{D}_t \mu_t(\phi_t)\Big|_{t=s} \leq \mathscr{B}(\mu_s, \phi_s) + \mu_s\left(\frac{d}{dt}\Big|_{t=s} \phi_t\right). \tag{3.4}$$

In particular for every $t_1 \leq a < b \leq t_2$

$$\mu_b(\phi_b) - \mu_a(\phi_a) \leq \int_a^b \left(\mathscr{B}(\mu_t, \phi_t) + \int_U \frac{\partial \phi}{\partial t}(t,x)d\mu_t(x)\right)dt. \tag{3.5}$$

3.9 Remark.

1. By (3.5) the Brakke flow has to be non-singular for almost every time $t \in [t_1,t_2]$. In particular the mean curvature vector $\vec{H}$ is defined and L^2-integrable on U for almost every time $t \in [t_1, t_2]$.

2. The rather technical condition (3.3) is necessary because our test functions are not assumed to have compact support in U. In particular we later want to use $V = C_\rho(a)$, $a \in \mathbb{R}^{n+k}$, $\rho \in (0,\infty)$.

3. Inequality (3.5) can be used as a definition of Brakke flow which is done in [KT]. Their definition also includes an extra term.

4. Brakke's original proof contains a major gap. He tries to estimate $\mu_{s+h}\left(\frac{\partial}{\partial t}\phi(s,\cdot)\right)$ using (3.1), which is in general not possible as $\frac{\partial}{\partial t}\phi(s,\cdot)$ may have negative values.

Proof. Note that inequality (3.5) follows from inequality (3.4) by Proposition A.19. For $s \in (t_1,t_2)$ set $E_s := \overline{D}_t\mu_t(\phi(t,\cdot))\Big|_{t=s}$ and calculate

$$\begin{aligned} E_s &= \limsup_{h\to 0}\ h^{-1}\left[\mu_{s+h}(\phi_{s+h}) - \mu_s(\phi_s)\right] \\ &= \limsup_{h\to 0}\ h^{-1}\left[\mu_{s+h}(\phi_{s+h}) - \mu_{s+h}(\phi_s) + \mu_{s+h}(\phi_s) - \mu_s(\phi_s)\right] \\ &\qquad + \mu_{s+h}\left(\frac{\partial\phi}{\partial t}(s,\cdot)\right) - \mu_{s+h}\left(\frac{\partial\phi}{\partial t}(s,\cdot)\right), \end{aligned}$$

so we arrive at the inequality

$$
\begin{aligned}
E_s \leq & \limsup_{h\to 0}\ h^{-1}\left[\mu_{s+h}\left(\phi_s\right)-\mu_s\left(\phi_s\right)\right] \\
& +\limsup_{h\to 0}\ h^{-1}\int_U\left[\phi_{s+h}(x)-\phi_s(x)-h\frac{\partial\phi}{\partial t}(s,x)\right]d\mu_{s+h}(x) \\
& +\limsup_{h\to 0}\mu_{s+h}\left(\frac{\partial\phi}{\partial t}(s,\cdot)\right).
\end{aligned}
\tag{3.6}
$$

By Definition 3.4 we can estimate

$$
\limsup_{h\to 0}\ h^{-1}\left[\mu_{s+h}\left(\phi_s\right)-\mu_s\left(\phi_s\right)\right]=\overline{D}_t\mu_t(\phi_s)\Big|_{t=s}\leq\mathscr{B}\left(\mu_s,\phi_s\right). \tag{3.7}
$$

Here we needed to be allowed to use (3.3) to use (3.1), see Remark 3.5.2. Set

$$
K:=\bigcup_{t\in[t_1,t_2]}\operatorname{spt}\mu_t\cap V\subset\subset U.
$$

By Lemma 3.6 there exists an $M\in(0,\infty)$ such that $\sup_{t\in[t_1,t_2]}\mu_t(K)\leq M$. This lets us estimate

$$
\begin{aligned}
&\int_U\left[\phi_{s+h}(x)-\phi_s(x)-h\frac{\partial\phi}{\partial t}(s,x)\right]d\mu_{s+h}(x) \\
&=\int_U\int_0^h\left[\frac{\partial\phi}{\partial t}(s+r,x)-\frac{\partial\phi}{\partial t}(s,x)\right]dr\ d\mu_{s+h}(x) \\
&=Mh\sup_{r\in(-|h|,|h|)}\sup_{x\in K}\left|\frac{\partial\phi}{\partial t}(s+r,x)-\frac{\partial\phi}{\partial t}(s,x)\right|.
\end{aligned}
$$

Then by the continuity of the derivative

$$
\limsup_{h\to 0}\ h^{-1}\int_U\left[\phi_{s+h}(x)-\phi_s(x)-h\frac{\partial\phi}{\partial t}(s,x)\right]d\mu_{s+h}(x)=0. \tag{3.8}
$$

Inserting (3.7) and (3.8) into (3.6) we arrive at

$$
\overline{D}_t\mu_t\left(\phi(t,\cdot)\right)\Big|_{t=s}\leq\mathscr{B}\left(\mu_s,\phi(s,\cdot)\right)+\limsup_{h\to 0}\mu_{s+h}\left(\frac{\partial\phi}{\partial t}(s,\cdot)\right)
$$

for every $s\in[t_1,t_2)$. At times where μ_s is continuous the last term is $\mu_s\left(\frac{\partial\phi}{\partial t}(s,\cdot)\right)$ and according to Lemma 3.7 this is the case for almost every $s\in(t_1,t_2)$. $\square$

Next we derive a certain estimate of the Brakke variation for a cylindrical cut-off function in terms of mean curvature- and height-excess. It particularly shows that, if the mean curvature-excess is large compared to the height-excess one gets a negative bound from above on the Brakke variation and thus for a Brakke flow the measure of the cut-off function is decreasing. This is taken from [B, 6.5].

3.10 Lemma (Variation Bound, [B, 6.5])**.** *There exists a constant $C \in (1,\infty)$ such that for $R, \gamma \in (0,\infty)$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Suppose*

$$R^{-n-2} \int_{C_R(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu(x) \leq \gamma^2. \tag{3.9}$$

Then the estimate

$$R^{-n+2} \mathscr{B}\left(\mu, \phi_R^2\right) \leq -\frac{1}{2} R^{-n+2} \int_U |\vec{H}|^2 \phi_R^2 \, d\mu + CM^2\gamma^2 \tag{3.10}$$

holds for all $\phi \in C_c^2([-1,1],\mathbb{R}^+)$ with $\phi' \equiv 0$ on $[-\frac{1}{2},\frac{1}{2}]$, where $\phi_R(x) = \phi\left(R^{-1}|\hat{x}|\right)$ and $M := \max\{\sup|\phi|, \sup|\phi'|, \sup|\phi''|, 1\}$

Proof. Set $\Phi := \phi_R$, then $\max\{\sup|\Phi|, R\sup|D\Phi|, R^2\sup|D^2\Phi|\} \leq M$. Set

$$\alpha^2 := R^{-n+2} \int_U |\vec{H}|^2 \Phi^2 d\mu.$$

By definition of the Brakke variation and Remark A.7.1 we have

$$\begin{aligned} R^{-n+2} \mathscr{B}\left(\mu, \Phi^2\right) &= R^{-n+2} \int_U \vec{H} \cdot D\left(\Phi^2\right) d\mu - \alpha^2 \\ &\leq 2R^{-n+2} \int_U \Phi\, |D\Phi|\, |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|\, |\vec{H}| d\mu - \alpha^2, \end{aligned}$$

where we used $D\Phi \in \mathbb{R}^n \times \{0\}^k$ and $\vec{H} \perp T_x\mu$ by Theorem 2.7. With Young's inequality it then follows that

$$R^{-n+2} \mathscr{B}\left(\mu, \Phi^2\right) \leq 2R^{-n+2} \int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, |D\Phi|^2 \, d\mu - \frac{3}{4}\alpha^2, \tag{3.11}$$

By Lemma A.13 we can estimate

$$\begin{aligned} &\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, |D\Phi|^2 \, d\mu \\ &\leq C_n \Bigg(\int_U |\pi_{\mathbb{R}^k}(x)|^2 \, |D\, |D\Phi(x)||^2 \, d\mu(x) \\ &\qquad + \sqrt{\int_U |\vec{H}(x)|^2 \Phi^2(x) d\mu(x) \int_U |\pi_{\mathbb{R}^k}(x)|^2 \frac{|D\Phi(x)|^4}{\Phi^2(x)} d\mu(x)} \Bigg). \end{aligned}$$

Here we used Lemma A.13 with $f = \Phi$, $g = |D\Phi|$ and $h = \Phi^{-1}|D\phi|^2$. By Proposition A.6.1 h is actually well defined if we sustain with 0 outside $\{\Phi > 0\}$. Proposition A.6 also yields $|D\,|D\Phi||^2 \le |D^2\Phi|^2 \le R^{-4}M^2$ and $\phi^{-1}|D\Phi|^2 \le 2\sup|D^2\Phi| \le 2R^{-2}M$. Combined with Young's inequality this lets us estimate

$$\begin{aligned}
&\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2\, |D\Phi|^2\, d\mu \\
&\le C_n R^4 M^2 \int_{C_R(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu(x) + \frac{1}{8}\int_U |\vec{H}(x)|^2\Phi^2(x)d\mu(x) \\
&\le \left(C_n M^2\gamma^2 + \frac{\alpha^2}{8}\right) R^{n-2}.
\end{aligned}$$

Inserting this into (3.11) establishes the result. □

A very important tool are barrier functions, introduced in the next Lemma, which is from [B, 3.6]

3.11 Lemma (Barrier Function Lemma, Brakke [B, 3.6])**.** *Let μ be an integral n-varifold in $B_{2R}(x_0)$, $x_0 \in \mathbb{R}^{n+k}$, $R \in (0,\infty)$ and $f \in C^2(\mathbb{R}, \mathbb{R}^+)$ with $f'' \ge 0$ and $f(t) = 0$ for $t \ge R^2$. For $(t,x) \in [t_0,\infty) \times \mathbb{R}^{n+k}$, $t_0 \in \mathbb{R}$ set $r(t,x) := |x - x_0|^2 + 2n(t - t_0)$. Then*

$$\mathscr{B}(\mu, f(r(s,\cdot))) \le -\int_{B_{2R}(x_0)} \frac{d}{ds} f(r(s,x))d\mu(x)$$

for all $s \in (t_0 - \frac{3R^2}{2n}, \infty)$

Proof. For $s \in (t_0 - \frac{3R^2}{2n}, \infty)$ set $r_s(x) = |x - x_0|^2 + 2n(s - t_0)$. We may assume μ has L^2-integrable mean curvature vector $\vec{H}$, or else $\mathscr{B}(\mu, f(r(s,\cdot))) = -\infty$, which directly implies the result. By the definition of mean curvature and divergence (see Definitions 2.4 and 2.5) calculate

$$\begin{aligned}
&\int_{\mathbb{R}^{n+k}} D(f(r_s(x))) \cdot \vec{H} d\mu = -\int_{\mathbb{R}^{n+k}} \operatorname{div}_\mu(D(f(r_s(x))))d\mu \\
&= -\int_{\mathbb{R}^{n+k}} \operatorname{div}_\mu(2f'(r_s(x))(x - x_0))d\mu(x) \\
&= -2\int_{\mathbb{R}^{n+k}} \nabla^\mu f'(r_s(x)) \cdot (x - x_0)d\mu(x) - 2\int_{\mathbb{R}^{n+k}} f'(r_s(x))\operatorname{div}_\mu x\, d\mu(x) \\
&= -4\int_{\mathbb{R}^{n+k}} f''(r_s(x))|\pi_{T_x\mu}(x - x_0)|^2 d\mu(x) - 2n\int_{\mathbb{R}^{n+k}} f'(r_s(x))d\mu(x)
\end{aligned}$$

thus we can estimate

$$\begin{aligned}\int_{\mathbb{R}^{n+k}} D(f(r_s(x))) \cdot \vec{H} d\mu(x) &\leq -2n \int_{\mathbb{R}^{n+k}} f'(r_s(x)) d\mu(x) \\ &= -\int_{\mathbb{R}^{n+k}} \frac{d}{ds} f(r(s,x)) d\mu_t(x).\end{aligned}$$

Then the result follows by Definition 3.1, as the $|\vec{H}|^2$-integral is always negative. Note that for $s \in (t_0 - \frac{3R^2}{2n}, \infty)$, $r_s(x) < R^2$ implies $|x - x_0|^2 < 4R^2$, so $\mathrm{spt}(f \circ r_s) \subset B_{2R}(x_0)$ for all $s \in (t_0 - \frac{3R^2}{2n}, \infty)$. □

This can now be used to establish local area and height bounds

3.12 Lemma. *For an open subset $U \subset \mathbb{R}^{n+k}$, $t_1 \in \mathbb{R}$, $t_2 \in (t_1, \infty)$, a Brakke flow $(\mu_t)_{t \in [t_1, t_2]}$ in U, $x_0 \in \mathbb{R}^{n+k}$, $R \in (0, \infty)$ with $B_R(x_0) \subset\subset U$ and $r(t,x) := |x - x_0|^2 + 2n(t - t_1)$ the following holds:*

1. *Let $f \in C^2([0,\infty), \mathbb{R}^+)$ with $f'' \geq 0$ and $f(t) = 0$ for $t \geq R^2$ then*

   $$\int_U f(r(s_2, x)) d\mu_{s_2}(x) \leq \int_U f(r(s_1, x)) d\mu_{s_1}(x)$$

   holds for all $t_1 \leq s_1 \leq s_2 \leq t_2$.

2. *For $f_R(r) := (\{1 - R^{-2} r\}_+)^3$ and $\kappa \in (0,1)$ this implies*

   $$\mu_{t_1}(B_R(x_0)) \geq (\kappa - \kappa^2)^3 \mu_t\left(B_{(1-\kappa)R}(x_0)\right)$$

   for all $t \in [t_1, t_1 + (2n)^{-1} \kappa R^2] \cap [t_1, t_2]$

3. *For $\delta \in (0, \infty)$ and $K \subset U$ compact with $d(K, \partial U) > 2\delta$ there exists $N = N(K, \delta) \in \mathbb{N}$ such that*

   $$N \mu_{t_1}(K_\delta) \geq \mu_t(K)$$

   for all $t \in [t_1, t_1 + (4n)^{-1} \delta^2] \cap [t_1, t_2]$, where $K_\delta := \{x \in \mathbb{R}^{n+k} d(x, K) < \delta\}$.

4. *Consider $\delta \in (0, 6^{-1}]$ and $v \in \mathbb{R}^{n+k}$ with $|v| = 1$. If we have*

   $$\mathrm{spt}\mu_{t_1} \cap B_R(x_0) \subset \{x \in U : (x - x_0) \cdot v \leq 0\}, \tag{3.12}$$

   then

   $$\mathrm{spt}\mu_t \cap B_{\frac{R}{2}}(x_0) \subset \{x \in U : (x - x_0) \cdot v \leq \delta R\} \tag{3.13}$$

   for all $t \in [t_1, t_1 + (6n)^{-1} \delta R^2] \cap [t_1, t_2]$.

5. *Suppose* $U = \mathbb{R}^{n+k}$ *and* $v \in \mathbb{R}^{n+k}$ *with* $|v| = 1$. *If we have*
$$\mathrm{spt}\mu_{t_1} \subset \left\{x \in \mathbb{R}^{n+k} : (x - x_0) \cdot v \leq 0\right\}, \tag{3.14}$$
then
$$\mathrm{spt}\mu_t \subset \left\{x \in \mathbb{R}^{n+k} : (x - x_0) \cdot v \leq 0\right\} \tag{3.15}$$
for all $t \in [t_1, t_2]$.

Proof. 1. Combine Lemma 3.11 and inequality (3.5) from Proposition 3.8.

2. Note that for the derivatives of f we have $f_R' = -2R^{-2}\left(\{1 - R^{-2}r\}_+\right)^2$ and $f_R'' = -4R^{-4}\left(\{1 - R^{-2}r\}_+\right)$. By Result 1. with $s_1 = t_1$ and $s_2 = t$ this yields
$$\int_U f(r(t,x))d\mu_t(x) \leq \int_U f(r(t_1,x))d\mu_{t_1}(x) \leq \mu_{t_1}\left(B_R(x_0)\right),$$
where we used $\mathrm{spt}f(r(t_1,\cdot)) \subset B_R(x_0)$. For $t \in [t_1, t_1 + (2n)^{-1}\kappa R^2] \cap [t_1, t_2]$ and $x \in B_{(1-\kappa)R}(x_0)$ calculate
$$f(r(t,x)) \geq \left(\{1 - ((1-\kappa)^2 + \kappa)\}_+\right)^3 = \left(\kappa - \kappa^2\right)^3,$$
which verifies the result.

3. Consider the covering $K \subset \bigcup_{x \in K} B_{\frac{\delta}{2}}(x)$. By compactness of K there exists an $N_0 \in \mathbb{N}$ and $x_1, \ldots, x_N$ such that $K = \bigcup_{i=1}^{N_0} B_{\frac{\delta}{2}}(x)$. Then we can use result 2. with $x_0 = x_i$, $\kappa = \frac{1}{2}$ and $R = \delta$ to estimate
$$\mu_t(K) \leq \sum_{i=1}^{N_0} \mu_t\left(B_{\frac{\delta}{2}}(x_i)\right) \leq 8\sum_{i=1}^{N_0} \mu_{t_1}\left(B_\delta(x_i)\right) \leq 8N_0\mu_{t_1}(K_\delta)$$
for all $t \in [t_1, t_1 + (4n)^{-1}\delta^2] \cap [t_1, t_2]$. Here we used $8\left(\frac{1}{2} - \frac{1}{4}\right)^3 = \frac{1}{8}$. Then the result follows for $N = 8N_0$.

4. Set $T := \{x \in \mathbb{R}^{n+k} : x \cdot v = 0\}$. Let $s \in [t_1, t_1 + (6n)^{-1}\delta^2 R^2] \cap [t_1, t_2]$ and $y \in \mathrm{spt}\mu_s \cap B_{\frac{R}{2}}(x_0)$ be arbitrary. Set
$$\begin{aligned} a &:= y - ((y - x_0) \cdot v)v \\ a_0 &:= a + 3^{-1}Rv. \end{aligned}$$
Note that a is the projection of y onto $T + x_0$. The idea is to define a ball above a in v-direction, which does not intersect $\mathrm{spt}\mu_{t_1}$, then by

statement 2 a slightly smaller ball will not intersect $\mathrm{spt}\mu_t$ at later times, which then implies a height bound for y. We can calculate $(a-x_0)\cdot v = 0$ and $(a_0 - x_0)\cdot v = 3^{-1}R$ which yields

$$(x - x_0)\cdot v = 3^{-1}R + (x - a_0)\cdot v \tag{3.16}$$

for all $x \in \mathbb{R}^{n+k}$. Also we see

$$|a - x_0| = |y - x_0 - ((y - x_0)\cdot v)v| = |\pi_T(y - x_0)| \leq 2^{-1}R$$

and

$$\begin{aligned} |a_0 - x_0| &= \sqrt{|\pi_T(a_0 - x_0)|^2 + |(a_0 - x_0)\cdot v|^2} \\ &\leq \sqrt{|a - x_0|^2 + 3^{-2}R^2} \leq \frac{2}{3}R. \end{aligned}$$

Then $B_{3^{-1}R}(a_0) \subset B_R(x_0) \subset\subset U$. By (3.16) we see that

$$(x - x_0)\cdot v \geq 3^{-1}R - |x - a_0|$$

for all $x \in \mathbb{R}^{n+k}$. Combined with assumption (3.12) this yields $\mathrm{spt}\mu_{t_1} \cap B_{3^{-1}R}(a_0) = \emptyset$. Then use statement 2 with $x_0 = a_0$, $\kappa = 3\delta$ and R replaced by $3^{-1}R$ to obtain

$$\mathrm{spt}\mu_t \cap B_{(3^{-1}-\delta)R}(a_0) = \emptyset$$

for all $t \in [t_1, t_1 + (6n)^{-1}\delta R^2] \cap [t_1, t_2]$. In particular

$$|y - a_0| \geq (3^{-1} - \delta)R. \tag{3.17}$$

By definition of a_0 we see

$$y - a_0 - ((y - a_0)\cdot v)v = y - a_0 - ((y - a)\cdot v - 3^{-1}R)v = 0.$$

In particular this means

$$|y - a_0| = (a_0 - y)\cdot v, \text{ or } |y - a_0| = (y - a_0)\cdot v.$$

Combined with (3.17) and (3.16) with $x = y$ we can conclude

$$(y - x_0)\cdot v \leq \delta R, \text{ or } (y - x_0)\cdot v \geq (3^{-1} + 3^{-1} - \delta)R.$$

The second case contradicts $y \in B_{\frac{R}{2}}(x_0)$, as $|y - x_0| \geq (y - x_0)\cdot v$ and $\delta \leq 6^{-1}$. Thus we obtain the height bound and as s, y where arbitrary this establishes the result.

5. Let $s \in [t_1, t_2]$, $y \in \mathrm{spt}\mu_s$ and $\epsilon \in (0,1)$ be arbitrary. Set $R := \max\{\epsilon^{-2}, (6n(s-t_1))^2, 4|y-x_0|, 4\}$, then use statement 4. with $\delta = R^{-\frac{3}{2}} < \frac{1}{6}$ to obtain

$$\mathrm{spt}\mu_t \cap B_{\frac{R}{2}}(x_0) \subset \left\{x \in U : (x-x_0)\cdot v \leq R^{-\frac{1}{2}}\right\}$$

for all $t \in [t_1, t_1 + (6n)^{-1}\sqrt{R}] \cap [t_1, t_2]$. By the choice of R we have $s \in [t_1, t_1 + (6n)^{-1}\sqrt{R}]$, $y \in B_{\frac{R}{2}}(x_0)$ and $R^{-\frac{1}{2}} \leq \epsilon$. Thus $(y - x_0)\cdot v \leq \epsilon$ and as s, y, ϵ where arbitrary, this establishes the result.

□

In the smooth case a Brakke flow can be characterized by the mean curvature flow equation.

3.13 Definition. Let $(M_t)_{t\in[t_1,t_2]}$ be an immersed family of n-manifolds in $\mathbb{R}^{n+k}$, that is $M_t = F_t(N)$ for an n-dimensional manifold N and a smooth family of immersions $F_t = F(t,\cdot) : N \to \mathbb{R}^{n+k}$. $(M_t)_{t\in[t_1,t_2]}$ is called a *(smooth) mean curvature flow*, if for every $t \in (t_1, t_2)$ and $p \in N$

$$\left(\frac{\partial F}{\partial t}(t,p)\right)^{\perp} = \vec{H}\left(F(t,p)\right). \tag{3.18}$$

If F_t is actually an proper embedding for all $t \in [t_1, t_2]$, we call $(M_t)_{t\in[t_1,t_2]}$ an *embedded mean curvature flow*

3.14 Remark. For an open subset $U \subset \mathbb{R}^{n+k}$ every properly embedded manifold M defines a Radon measure μ_M in U via $\mu_M = \mathscr{H}^n \llcorner (M \cap U)$. This associated measure is an integral n-varifold and satisfies $\mu\left(\mathbb{R}^{n+k} \setminus U\right) = 0$.

For an embedded mean curvature flow $(M_t)_{t\in[t_1,t_2]}$ as above we can calculate

$$\frac{d}{dt}\int_{\mathbb{R}^{n+k}} \phi(t,\cdot)d\mu_t = \int_{\mathbb{R}^{n+k}} \left(-|\vec{H}|^2\phi(t,\cdot) + \vec{H}\cdot D\phi(t,\cdot)\right)d\mu_t \tag{3.19}$$

for every $\phi \in C^1_c(U, \mathbb{R})$. Here $\mu_t = \mu_{M_t}$ and we used the evolution of the area element under mean curvature flow, which was first calculated in [H1]. This formula is found in [E4].

If $U \cap \left(\bigcup_{[t_1,t_2]} \partial M_t\right) = \emptyset$, the smooth mean curvature vector restricted to U defines a generalised mean curvature vector for μ_t. Then (3.19) implies the Brakke flow equation (3.1), such that the associated measure forms a Brakke flow in U.

The reverse is in general false. However, a Brakke flow $(\mu_t)_{t\in[t_1,t_2]}$ in U, for which the μ_t are generated by a family of embedded $C^{1,\alpha}$-graphs with certainly small $C^{1,\alpha}$-norm actually creates a smooth mean curvature flow via $M_t = \mathrm{spt}\mu_t \cap U$. This has been done by Brakke as part of [B, 6.10], but the proof contains many gaps and small errors. There exists a new proof in [T], see in particular [T, 6.3].

4 Heat Diffusion

Consider a Brakke flow $(\mu_t)_{t\in[t_1,t_2]}$ in an open subset $U \subset \mathbb{R}^{n+k}$. The aim of this section is to show that evolution by Brakke flow is somehow close to linear heat diffusion. A convolution of some function $g \in C^\infty\left(\mathbb{R}^{n+k}\right)$ with the cylindrical heat kernel Ψ on μ_t would evolve for time τ by the Brakke flow by just changing the measure. However, we could also evolve by changing the parameter in the heat kernel, which is like moving by linear heat diffusion.

$$\begin{tikzcd} & \int g(x)\Psi(s,x)d\mu_{t+\tau}(x) \\ \int g(x)\Psi(s,x)d\mu_t(x) \arrow[ur, "\text{Brakke flow}"] \arrow[dr, "\text{heat diffusion}"'] & ||\ ? \\ & \int g(x)\Psi(s+\tau,x)d\mu_t(x) \end{tikzcd}$$

Figure 1: evolution for time τ

The main result of this section is Theorem 4.15, which estimates the difference of these two outcomes. Later we will use this for $g(x) = x_{n+j}$ and Brakke flows in a certainly narrow slab. In this case the two evolutions almost behave identically.

This section is based on [B, 6.8 and 6.9]. In particular Theorem 4.15 is like an integrated version of [B, 6.8], but it is much more general and the proof is very different. We will calculate the mean curvature evolution equation for our heat kernel, which yields Proposition 4.11. Combining this with the Brakke flow equation and the definition of the mean curvature vector, in the form of Lemma 4.13, already produces Theorem 4.15. Note that the results from this section are not covered in [KT]. In [KT] a completely different approach is taken as far as [B, 6.8 and 6.9] is concerned. Here we decided to adhere to Brakke's original method which is interesting in its own right.

We start with the definition and some basic properties of the heat kernel. As we are looking at the flow only locally we have to modify the heat kernel with a cut-off function ζ.

4.1 Definition. Fix $\zeta \in C^\infty\left([0,\infty),[0,1]\right)$ with

$$\zeta(r) = \begin{cases} 1 & \text{for } 0 \le r \le 1 - 2^{-n-9} \\ 0 & \text{for } 1 \le r \end{cases}$$

and such that $\max\{\sup|\zeta'|, \sup|\zeta''|\} \le \sigma_1$, for some constant $\sigma_1 \in (0,\infty)$.

4.2 Definition. For $t, \rho \in (0, \infty)$ and $\hat{x} \in \mathbb{R}^n$ we define the cylindrical heat kernel Ψ by

$$\begin{aligned}\Psi(t,\hat{x}) &:= \sigma_2\, t^{-\frac{n}{2}} e^{-\frac{|\hat{x}|^2}{4t}} \\ \Psi_\rho(t,\hat{x}) &:= \Psi(t,\hat{x})\zeta\left(\rho^{-1}|\hat{x}|\right)\end{aligned}$$

where $\sigma_2^{-1} := \int_{\mathbb{R}^n} e^{-\frac{|\hat{x}|^2}{4}} dx$, so $\int_{\mathbb{R}^n} \Psi_\rho(t,\hat{x}) d\mathscr{L}^n(\hat{x}) \le 1$ for every $t > 0$.

4.3 Remark. Note that σ_1 and σ_2 are considered absolute constants and dependence on σ_1, σ_2 will not be denoted explicitly in the sequel.

4.4 Remark. For the truncated heat kernel we calculate the following derivatives:

$$\begin{aligned}\frac{\partial \Psi_\rho}{\partial t}(t,\hat{x}) &= \left(\frac{|\hat{x}|^2}{4t^2} - \frac{n}{2t}\right)\Psi_\rho(t,\hat{x}) \\ \frac{\partial \Psi_\rho}{\partial x_i}(t,\hat{x}) &= -\frac{\hat{x}_i}{2t}\Psi_\rho(t,\hat{x}) + \frac{\hat{x}_i}{\rho|\hat{x}|}\zeta'(\rho^{-1}|\hat{x}|)\Psi(t,\hat{x}) \\ \frac{\partial^2 \Psi_\rho}{\partial x_i \partial x_j}(t,\hat{x}) &= \left(\frac{\hat{x}_i\hat{x}_j}{4t^2} - \frac{\delta_{ij}}{2t}\right)\Psi_\rho(t,\hat{x}) + \frac{\hat{x}_i\hat{x}_j}{\rho^2|\hat{x}|^2}\zeta''(\rho^{-1}|\hat{x}|)\Psi(t,\hat{x}) \\ &\quad + \left(\frac{\delta_{ij}}{\rho|\hat{x}|} - \frac{\hat{x}_i\hat{x}_j}{\rho|\hat{x}|^3} - \frac{\hat{x}_i\hat{x}_j}{t\rho|\hat{x}|}\right)\zeta'(\rho^{-1}|\hat{x}|)\Psi(t,\hat{x}) \\ \Delta_{\mathbb{R}^{n+k}}\Psi_\rho(t,\hat{x}) &= \left(\frac{|\hat{x}|^2}{4t^2} - \frac{n}{2t}\right)\Psi_\rho(t,\hat{x}) + \rho^{-2}\zeta''(\rho^{-1}|\hat{x}|)\Psi(t,\hat{x}) \\ &\quad + \left(\frac{n-1}{\rho|\hat{x}|} - \frac{|\hat{x}|}{t\rho}\right)\zeta'(\rho^{-1}|\hat{x}|)\Psi(t,\hat{x}).\end{aligned}$$

Here $1 \le i, j \le n$.

4.5 Remark. 1. $\Psi(t,\hat{x}) \le \sigma_2 t^{-\frac{n}{2}}$ for all $t \in (0,\infty)$ and all $\hat{x} \in \mathbb{R}^n$.

2. For every $P \in [0,\infty)$ exists $\kappa \in (0,1)$ such that the following holds: Let $t, r \in (0,\infty)$, $\hat{x}_0 \in \mathbb{R}^n$ and $\hat{x} \in \mathbb{R}^n \setminus B_r(x_0)$, then we can estimate

$$\Psi(t,\hat{x}) \le \sigma_2 t^{-\frac{n}{2}} e^{-\frac{r^2}{4t}}.$$

Suppose $r^{-2}t \le \kappa$ for κ small enough depending on P and n, then we can estimate further

$$\Psi(t,\hat{x}) \le \sigma_2 r^{-n}(t^{-1}r^2)^{\frac{n}{2}} e^{-\frac{r^2}{4t}} \le \sigma_2 r^{-n}(r^{-2}t)^P.$$

3. Statement 2 will be used to estimate all the extra-terms arising from the cut-off function in the derivatives above. Note that the derivatives of $\zeta(\rho^{-1}(x))$ are 0 for $x \in B_{\frac{\rho}{2}}(0)$, so statement 2 says that, if ρ is large compared to t, cutting-off only produces small extra-terms.

4.6 Lemma. *For every $P \in [0, \infty)$ there exists a $\Lambda \in (1, \infty)$ such that for all $t \in (0, \infty)$ and all $\hat{x}_0 \in \mathbb{R}^n$*

$$\int_{\mathbb{R}^n} |\hat{x} - \hat{x}_0|^P \Psi(t, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \leq \Lambda t^{\frac{P}{2}}.$$

Proof. By the transformation of variables $\hat{y} = t^{-\frac{1}{2}}(\hat{x} - \hat{x}_0)$ we obtain

$$\begin{aligned}\int_{\mathbb{R}^n} |\hat{x} - \hat{x}_0|^P t^{-\frac{n}{2}} e^{-\frac{|\hat{x}-\hat{x}_0|^2}{4t}} d\mathscr{L}^n(\hat{x}) &= \int_{\mathbb{R}^n} |\sqrt{t}\hat{y}|^P t^{-\frac{n}{2}} e^{-\frac{t|\hat{y}|^2}{4t}} t^{\frac{n}{2}} d\mathscr{L}^n(\hat{y}) \\ &= t^{\frac{P}{2}} \int_{\mathbb{R}^n} |\hat{y}|^P e^{-\frac{|\hat{y}|^2}{4}} d\mathscr{L}^n(\hat{y}).\end{aligned}$$

Then for $\Lambda := \sigma_2 \int_{\mathbb{R}^n} |\hat{y}|^P e^{-\frac{|\hat{y}|^2}{4}} d\mathscr{L}^n(\hat{y}) + 1$ the result follows. □

4.7 Lemma. *There exists a $C \in (1, \infty)$ such that for all $q_0 \in (0, \infty)$, $q \in (q_0, \infty)$ and all $\hat{x}_0 \in \mathbb{R}^n$*

$$\int_{\mathbb{R}^n} |\Psi(q, \hat{x} - \hat{x}_0) - \Psi(q_0, \hat{x} - \hat{x}_0)| \, d\mathscr{L}^n(\hat{x}) \leq C \log\left(1 + q_0^{-1}(q - q_0)\right).$$

Proof. Use the fundamental theorem of calculus and Fubini's theorem to obtain

$$\begin{aligned}&\int_{\mathbb{R}^n} |\Psi(q, \hat{x} - \hat{x}_0) - \Psi(q_0, \hat{x} - \hat{x}_0)| \, d\mathscr{L}^n(\hat{x}) \\ &= \int_{\mathbb{R}^n} \left| \int_0^{q-q_0} \frac{\partial}{\partial t} \Psi(q_0 + s, \hat{x} - \hat{x}_0) ds \right| d\mathscr{L}^n(\hat{x}) \\ &\leq \int_0^{q-q_0} \int_{\mathbb{R}^n} \left| \frac{\partial}{\partial t} \Psi(q_0 + s, \hat{x} - \hat{x}_0) \right| d\mathscr{L}^n(\hat{x}) ds.\end{aligned} \tag{4.1}$$

With Lemma 4.6 we can estimate

$$\begin{aligned}&\int_{\mathbb{R}^n} \left| \frac{\partial}{\partial t} \Psi(q_0 + s, \hat{x} - \hat{x}_0) \right| d\mathscr{L}^n(\hat{x}) \\ &\leq \int_{\mathbb{R}^n} \left(\frac{|\hat{x} - \hat{x}_0|^2}{4(q_0 + s)^2} + \frac{n}{2(q_0 + s)} \right) \Psi(q_0 + s, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \leq C_n (q_0 + s)^{-1}.\end{aligned}$$

Inserting this into (4.1) yields

$$\begin{aligned}\int_{\mathbb{R}^n} |\Psi(q,\hat{x}-\hat{x}_0)-\Psi(q_0,\hat{x}-\hat{x}_0)|\, d\mathscr{L}^n(\hat{x}) &\le C_n \int_0^{q-q_0} (q_0+s)^{-1} ds \\ &= C_n \log\left(q_0^{-1}q\right)\end{aligned}$$

This establishes the result. $\square$

4.8 Lemma. *For every $P_0 \in (0,\infty)$ there exists a $\kappa_0 \in (0,1)$ such that for all $r,t \in (0,\infty)$, $x_0 \in \mathbb{R}^{n+k}$ with $r^{-2}t \le \kappa_0$ the following holds:*

1. *We have*

$$\int_{\mathbb{R}^n \setminus B_r(\hat{x}_0)} \Psi(t,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \le (r^{-2}t)^{P_0}.$$

2. *For every $R \in (r,\infty)$ and every Radon measure μ with $\mu(C_R(x_0)) < \infty$ we have*

$$\int_{C_R(x_0)\setminus C_r(x_0)} \Psi(t,\hat{x}-\hat{x}_0) d\mu(x) \le (r^{-2}t)^{P_0} r^{-n} \mu(C_R(x_0)))$$

Proof. Writing the integral in spherical coordinates we obtain

$$\int_{\mathbb{R}^n \setminus B_r(\hat{x}_0)} \Psi(t,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) = \sigma_2\omega_n \int_r^\infty \rho^{n-1} t^{-\frac{n}{2}} e^{-\frac{\rho^2}{4t}} d\rho$$

then transforming $s = t^{-\frac{1}{2}}\rho$ yields

$$\int_{\mathbb{R}^n \setminus B_r(\hat{x}_0)} \Psi(t,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) = \sigma_2\omega_n \int_{t^{-\frac{1}{2}}r}^\infty s^{n-1} e^{-\frac{s^2}{4}} ds.$$

For the last integral we estimate

$$\int_{t^{-\frac{1}{2}}r}^\infty s^{n-1} e^{-\frac{s^2}{4}} ds \le \int_{t^{-\frac{1}{2}}r}^\infty \frac{s^{-2P_0-1}}{2P_0\sigma_2\omega_n} ds \le \frac{t^{P_0}}{r^{2P_0}\sigma_2\omega_n},$$

where we used that $s^{n-1}e^{-\frac{s^2}{4}} \le \frac{1}{2P_0\sigma_2\omega_n} s^{-2P_0-1}$ for $s \ge t^{-\frac{1}{2}}r$ large enough, which can be achieved for small enough κ_0 depending on P_0. This establishes result 1.

For statement 2. we can directly estimate

$$\begin{aligned}\int_{C_R(x_0)\setminus C_r(x_0)} \Psi(t,\hat{x}-\hat{x}_0) d\mu(x) \qquad &\le \sigma_2 \int_{C_R(x_0)\setminus C_r(x_0)} t^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} d\mu(x) \\ \le \sigma_2 t^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} \mu(C_R(x_0)).\end{aligned}$$

For $r^{-2}t \leq \kappa_0$ small enough depending on P_0 we have

$$\sigma_2 t^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} = \sigma_2 r^{-n} \left(r^{-2}t\right)^{-\frac{n}{2}} e^{-\frac{r^2}{4t}} \leq r^{-n} \left(r^{-2}t\right)^{P_0}$$

which establishes result 2. $\square$

4.9 Lemma. *For every $P_1 \in (0,\infty)$ there exists a $\kappa_1 \in (0,1)$ such that for all $\rho, t \in (0,\infty)$ with $\rho^{-2}t \leq \kappa_1$ and every $x_0 \in \mathbb{R}^{n+k}$ the following holds: Let μ be a Radon measure with $\mu(C_\rho(x_0)) < \infty$ then*

$$\begin{aligned}&\int_{C_\rho(x_0)} |D\Psi(t,\hat{x}-\hat{x}_0)\zeta\left(\rho^{-1}|\hat{x}-\hat{x}_0|\right) - D\Psi_\rho(t,\hat{x}-\hat{x}_0)|d\mu(x)\\ &\leq (\rho^{-2}t)^{P_1}\rho^{-n-1}\mu(C_\rho(x_0))\end{aligned} \tag{4.2}$$

and

$$\begin{aligned}&\int_{C_\rho(x_0)} |D^2\Psi(t,\hat{x}-\hat{x}_0)\zeta\left(\rho^{-1}|\hat{x}-\hat{x}_0|\right) - D^2\Psi_\rho(t,\hat{x}-\hat{x}_0)|d\mu(x)\\ &\leq (\rho^{-2}t)^{P_1}\rho^{-n-2}\mu(C_\rho(x_0))\end{aligned} \tag{4.3}$$

Proof. We may assume $x_0 = 0$. Use Remark 4.4 to estimate

$$\begin{aligned}D_1 &:= \int_{C_\rho(0)} |D\Psi(t,\hat{x})\zeta\left(\rho^{-1}|\hat{x}|\right) - D\Psi_\rho(t,\hat{x})|d\mu(x)\\ &\leq n\max_i \int_{C_\rho(0)} \frac{|\hat{x}_i|}{\rho|\hat{x}|}\zeta'\left(\rho^{-1}|\hat{x}|\right)\Psi(t,\hat{x})d\mu(x),\end{aligned}$$

as well as

$$\begin{aligned}D_2 &:= \int_{C_\rho(0)} |D^2\Psi(t,\hat{x})\zeta\left(\rho^{-1}|\hat{x}|\right) - D^2\Psi_\rho(t,\hat{x})|d\mu(x)\\ &\leq n^2\max_{i,j}\Bigg[\int_{C_\rho(0)} \frac{|\hat{x}_i\hat{x}_j|}{\rho^2|\hat{x}|^2}\zeta''\left(\rho^{-1}|\hat{x}|\right)\Psi(t,\hat{x})d\mu(x)\\ &\qquad + \int_{C_\rho(0)}\left(\frac{\delta_{ij}}{\rho|\hat{x}|} + \frac{|\hat{x}_i\hat{x}_j|}{\rho|\hat{x}|^3} + \frac{|\hat{x}_i\hat{x}_j|}{t\rho|\hat{x}|}\right)\zeta'\left(\rho^{-1}|\hat{x}|\right)\Psi(t,\hat{x})d\mu(x)\Bigg].\end{aligned}$$

By definition of ζ we have $\zeta'(\rho^{-1}|\hat{x}|) = \zeta''(\rho^{-1}|\hat{x}|) = 0$ for all $x \in C_{\frac{\rho}{2}}(0)$ and all $x \in \mathbb{R}^{n+k} \setminus C_\rho(0)$. Also we have $\max\{\sup|\zeta'|, \sup|\zeta''|\} \leq \sigma_1$ so we can estimate for D_1 and D_2

$$\rho^i D_i \leq C_n\left(1 + \frac{\rho^2}{t}\right)\int_{C_\rho(0)\setminus C_{\frac{\rho}{2}}(0)} \Psi(t,\hat{x})d\mu(x)$$

for $i \in \{1,2\}$. Note that we estimated $\sigma_1 \leq C_n$. Now we can use Lemma 4.8.2 with $P_0 = P_1 + 2$ and $R = 2r = \rho$ to find a $\kappa_0 \in (0,1)$ depending on P_1 such that

$$\rho^i D_i \leq C_n 2^{2P_1+4} \left(1 + \frac{\rho^2}{t}\right) (\rho^{-2}t)^{P_1+2} \rho^{-n} \mu(C_\rho(0)) \tag{4.4}$$

for $i \in \{1,2\}$. Here we chose $\kappa_1 \leq 4\kappa_0$, so $\rho^{-2}t \leq \kappa_1 \leq 4\kappa_0$. Also for κ_1 small enough depending on P_1 the inequality $\rho^{-2}t \leq \kappa_1$ leads to

$$C_n 2^{2P_1+4} \left(1 + \frac{\rho^2}{t}\right) (\rho^{-2}t)^2 \leq C_n 2^{2P_1+6} \kappa_1 \leq 1.$$

Thus (4.4) becomes

$$\rho^i D_i \leq (\rho^{-2}t)^{P_1} \rho^{-n} \mu(C_\rho(0))$$

for $i \in \{1,2\}$, which establish the result. $\square$

If the time parameter goes to zero, the truncated heat kernel converges to the Dirac delta distribution. This is shown in the following proposition.

4.10 Proposition. *For every $P \in (0,\infty)$ there exists a $\kappa \in (0,1)$ such that for all $\rho, s, M \in (0,\infty)$ and every $\hat{x}_0 \in \mathbb{R}^n$ the following holds:*

1. *For $f : B^n_\rho(\hat{x}_0) \to [-M, M]$ and $r \in \left(0, \frac{\rho}{2}\right]$ with $r^{-2}s \leq \kappa$ we have*

$$\begin{aligned}&\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x}) \Psi_\rho(s, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \\ &\leq \sup_{B^n_r(\hat{x}_0)} |f(\hat{x}_0) - f(\hat{x})| + M \left(r^{-2}s\right)^P.\end{aligned}$$

2. *For a continuous function $f \in C^0(B^n_\rho(\hat{x}_0))$ we obtain*

$$f(\hat{x}_0) = \lim_{t \searrow 0} \int_{\mathbb{R}^n} f(\hat{x}) \Psi_\rho(t, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}).$$

3. *If $\rho^{-2}s \leq \kappa$ we can estimate*

$$\int_{\mathbb{R}^n} \Psi(s, \hat{x} - \hat{x}_0) - \Psi_\rho(s, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \leq \left(\rho^{-2}s\right)^P.$$

Proof. Use $\int_{\mathbb{R}^n} \Psi(s,\hat{x}) d\mathscr{L}^n(\hat{x}) = 1$ and $\Psi_\rho(\hat{x}) = \Psi(\hat{x})\zeta\left(\rho^{-1}|\hat{x}|\right)$ to calculate

$$
\begin{aligned}
&\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x})\Psi_\rho(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(x) \right| \\
&= \left| \int_{\mathbb{R}^n} f(\hat{x}_0)\Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) - \int_{\mathbb{R}^n} f(\hat{x})\Psi_\rho(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \\
&= \left| \int_{\mathbb{R}^n} \big(f(\hat{x}_0) - f(\hat{x})\zeta(\rho^{-1}|\hat{x}-\hat{x}_0|) \big) \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \right|.
\end{aligned}
$$

Here we set $f \equiv 0$ outside $B^n_\rho(\hat{x}_0)$. Partitioning $\mathbb{R}^n$ into $B_r(x_0)$ and $\mathbb{R}^n \setminus B_r(x_0)$ we can estimate

$$
\begin{aligned}
&\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x})\Psi_\rho(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \\
&\le \int_{\mathbb{R}^n} \left| f(\hat{x}_0) - f(\hat{x})\zeta(\rho^{-1}|\hat{x}-\hat{x}_0|) \right| \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \\
&\le \sup_{\hat{x}\in B_r(\hat{x}_0)} \left| f(\hat{x}_0) - f(\hat{x})\zeta(\rho^{-1}|\hat{x}-\hat{x}_0|) \right| \int_{B_r(\hat{x}_0)} \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \\
&\quad + \sup_{\hat{x}\in \mathbb{R}^n} \left| f(\hat{x}_0) - f(\hat{x})\zeta(\rho^{-1}|\hat{x}-\hat{x}_0|) \right| \int_{\mathbb{R}^n\setminus B_r(\hat{x}_0)} \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}).
\end{aligned}
$$

As $\zeta \equiv 1$ on $B^n_{\frac{\rho}{2}}(x_0) \supset B^n_r(x_0)$ and $|f| \le M$ we obtain

$$
\begin{aligned}
&\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x})\Psi_\rho(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \\
&\le \sup_{\hat{x}\in B_r(\hat{x}_0)} |f(\hat{x}_0) - f(\hat{x})| + 2M\omega_n \int_{\mathbb{R}^n\setminus B_r(\hat{x}_0)} \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x})
\end{aligned}
\tag{4.5}
$$

By Lemma 4.8.1 with $P_0 = P+1$ we find a $\kappa_0 \in (0,1)$ depending on P such that

$$
\int_{\mathbb{R}^n\setminus B_r(\hat{x}_0)} \Psi(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \le \left(r^{-2}s\right)^{P+1}
$$

for $r^{-2}s \le \kappa_0$. Inserting this into (4.5) yields

$$
\begin{aligned}
&\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x})\Psi_\rho(s,\hat{x}-\hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \\
&\le \sup_{\hat{x}\in B_r(\hat{x}_0)} |f(\hat{x}_0) - f(\hat{x})| + 2M\left(r^{-2}s\right)^{P+1}
\end{aligned}
\tag{4.6}
$$

for $r^{-2}s \leq \kappa_0$. For κ small enough depending on κ_0 which in turn depends on P, this establishes statement 1.

To prove statement 2. let $\epsilon \in (0,1)$ be arbitrary. By continuity of f there exists $r \in \left(0, \frac{\rho}{2}\right]$ such that

$$\sup_{\hat{x} \in B_r(\hat{x}_0)} |f(\hat{x}_0) - f(\hat{x})| \leq \frac{\epsilon}{2}.$$

Using statement 1. with $P = 1$, $s = t$ and $M := \sup_{B^n_\rho(\hat{x}_0)} |f|$ yields a $\kappa \in (0,1)$ such that

$$\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x}) \Psi_\rho(t, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \leq \frac{\epsilon}{2} + M r^{-2} t$$

for all $t \in (0, \kappa r^2]$. Set $\delta = \min\{\kappa, M^{-1}\frac{\epsilon}{2}\} r^2$. Then for all $t \in (0, \delta]$

$$\left| f(\hat{x}_0) - \int_{\mathbb{R}^n} f(\hat{x}) \Psi_\rho(t, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \leq \epsilon$$

and as ϵ was arbitrary this establishes statement 2.

In Order to show statement 3. we apply (4.6) with $f \equiv 1$ and $r = \frac{\rho}{2}$ to obtain

$$\left| 1 - \int_{\mathbb{R}^n} \Psi_\rho(s, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) \right| \leq 2^{2P+3} \left(\rho^{-2}s\right)^{P+1}$$

for $\rho^{-2}s \leq 4\kappa_0$. For κ small enough $2^{2P+3}\left(\rho^{-2}s\right)^{P+1} \leq \left(\rho^{-2}s\right)^P$. Then $\int_{\mathbb{R}^n} \Psi(s, \hat{x} - \hat{x}_0) d\mathscr{L}^n = 1$ verifies the wanted estimate. Here the choice of κ depends on P and κ_0 which in turn also depends on P. $\square$

We know that Ψ satisfies the heat equation in $\mathbb{R}^{n+k}$. For a sufficiently flat varifold mean curvature flow almost coincides with heat diffusion. Consequently the evolution equation of the heat kernel is controlled by curvature and tilt-excess.

4.11 Proposition. *For every $P \in (0,\infty)$ there exists a $\kappa \in (0,1)$ such that for all $\rho, s, K \in (0,\infty)$ with $\rho^{-2}s \leq \kappa$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Suppose $\mu\left(C_\rho(x_0)\right) < \infty$. Then for every $g : U \to [-K, K]$*

$$\begin{aligned}
&\int_U \left| g(x) \left(\frac{\partial}{\partial t} - \Delta_\mu \right) \Psi_\rho(s, \hat{x}) \right| d\mu(x) \\
&\leq \frac{1}{2} \int_U |g(x)| |\vec{H}(x)|^2 \Psi_\rho(s, \hat{x}) d\mu(x) + (\rho^{-2}s)^P K \rho^{-n-2} \mu\left(C_\rho(0)\right) \\
&\quad + \int_U |g(x)| \frac{|\hat{x}|^2 + s}{s^2} \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 \Psi_\rho(s, \hat{x}) d\mu(x).
\end{aligned}$$

4.12 Remark. This formula is also valid for the non-truncated heat kernel. In this case we can leave out the $(\rho^{-2}s)^P \rho^{-n-2}\mu\left(C_\rho(0)\right)$-term, but have to replace U by $\mathbb{R}^{n+k}$ and we have to assume that the integrals on the right hand side exist.

Proof. For $x \in \operatorname{spt}\mu$ such that $T_x\mu$ exists let $\{\nu_j\}_{1\le j\le k}$ be an orthonormal basis of $T_x\mu^\perp$. By Remark 2.6.5 we have

$$\begin{aligned}
&\int_U \left(\frac{\partial}{\partial t} - \Delta_\mu\right)\Psi_\rho \, d\mu \\
&= \int_U \left(\left(\frac{\partial}{\partial t} - \Delta_{\mathrm{R}^{n+k}}\right)\Psi_\rho - D\Psi_\rho \cdot \vec{H} + \sum_{l=1}^{k} \nu_l \cdot D^2\Psi_\rho(\nu_l)\right) d\mu.
\end{aligned} \tag{4.7}$$

To calculate the first term on the right hand side of (4.7) use Remark 4.4. This yields

$$\begin{aligned}
&\left(\frac{\partial}{\partial t} - \Delta_{\mathrm{R}^{n+k}}\right)\Psi_\rho(s,\hat{x}) \\
&= -\rho^{-2}\zeta''(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) - \left(\frac{n-1}{\rho|\hat{x}|} - \frac{|\hat{x}|}{s\rho}\right)\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x})
\end{aligned} \tag{4.8}$$

for all $x \in \mathbb{R}^{n+k}$. Also by Remark 4.4 we can calculate

$$D\Psi_\rho(s,\hat{x}) \cdot \vec{H}(x) = -\frac{1}{2s}\hat{x} \cdot \vec{H}\Psi_\rho(s,\hat{x}) + \frac{1}{\rho|\hat{x}|}\hat{x} \cdot \vec{H}\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) \tag{4.9}$$

for all $x \in \operatorname{spt}\mu$. For $x \in \operatorname{spt}\mu$ such that $T_x\mu$ exists calculate

$$\sum_{l=1}^{k}\sum_{i,j=1}^{n} \hat{x}_i\hat{x}_j\left(\nu_l \cdot \mathbf{e}_i\right)\left(\nu_l \cdot \mathbf{e}_j\right) = \sum_{l=1}^{k}\left(\nu_l \cdot \hat{x}\right)^2 = \left|\pi_{T_xV}^{\perp}\left(\hat{x}\right)\right|^2. \tag{4.10}$$

Proposition A.11 implies

$$\sum_{l=1}^{k}\sum_{i,j=1}^{n} \delta_{ij}\left(\nu_l \cdot \mathbf{e}_i\right)\left(\nu_l \cdot \mathbf{e}_j\right) = \sum_{i=1}^{n}\sum_{l=1}^{k}\left(\nu_l \cdot \mathbf{e}_i\right)^2 = \frac{1}{2}\left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2. \tag{4.11}$$

Using Remark 4.4 in combination with (4.10) and (4.11) we can calculate

$$
\begin{aligned}
\sum_{l=1}^{k} \nu_l \cdot D^2\Psi_\rho(s,\hat{x})(\nu_l) &= \sum_{l=1}^{k}\sum_{i,j=1}^{n} \frac{\partial^2\Psi_\rho}{\partial x_i \partial x_j}(s,\hat{x})\,(\nu_l\cdot \mathbf{e}_i)\,(\nu_l\cdot\mathbf{e}_j) \\
&= \left(\frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{4s^2} - \frac{\left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2}{4s}\right)\Psi_\rho(s,\hat{x}) \\
&\quad + \frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{\rho^2|\hat{x}|^2}\zeta''(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) \\
&\quad + \left(\frac{\left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2}{2\rho|\hat{x}|} - \frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{\rho|\hat{x}|^3} - \frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{s\rho|\hat{x}|}\right)\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x})
\end{aligned}
\tag{4.12}
$$

for all $x \in \mathrm{spt}\mu$ such that $T_x\mu$ exists.

Inserting (4.8), (4.9) and (4.12) into (4.7) yields

$$
\begin{aligned}
&\int_U \left(\frac{\partial}{\partial t} - \Delta_\mu\right)\Psi_\rho(s,\hat{x})d\mu(x) \\
&= \int_U \Bigg[\left(\frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{4s^2} - \frac{\left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2}{4s} + \vec{H}(x)\cdot\frac{\hat{x}}{2s}\right)\Psi_\rho(s,x) \\
&\qquad + \frac{\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2 - |\hat{x}|^2}{\rho^2|\hat{x}|^2}\zeta''(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) \\
&\qquad - \vec{H}(x)\cdot\frac{\hat{x}}{\rho|\hat{x}|}\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) \\
&\qquad + \frac{\left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2 - 2(|\hat{x}|^{-2} + s^{-1})\left|\pi^{\perp}_{T_x\mu}(\hat{x})\right|^2}{2\rho|\hat{x}|}\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x}) \\
&\qquad + \left(\frac{|\hat{x}|}{s\rho} - \frac{n-1}{\rho|\hat{x}|}\right)\zeta'(\rho^{-1}|\hat{x}|)\Psi(s,\hat{x})\Bigg]d\mu(x).
\end{aligned}
$$

By definition of ζ we have $\zeta'(\rho^{-1}|\hat{x}|) = \zeta''(\rho^{-1}|\hat{x}|) = 0$ for all $x \in \bar{C}_{\frac{\rho}{2}}(0)$ and all $x \in U \setminus C_\rho(0)$. Also we have $\max\{\sup|\zeta'|, \sup|\zeta''|\} \le \sigma_1$, $\sup_U|g| \le K$

and $|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}|^2 \le 2n$ (by Proposition A.11), so we can estimate

$$\begin{aligned}
&\int_U \left| g(x) \left(\frac{\partial}{\partial t} - \Delta_\mu \right) \Psi_\rho(s,\hat{x}) \right| d\mu(x) \\
&\le \int_U |g(x)| \left| \left(\frac{\left|\pi^\perp_{T_x\mu}(\hat{x})\right|^2}{4s^2} - \frac{|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2}{4s} \right) \Psi_\rho(s,\hat{x}) \right. \\
&\qquad\qquad \left. + \vec{H}(x) \cdot \left(\frac{\hat{x}}{2s} \Psi_\rho(s,x) - \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \Psi(s,\hat{x}) \right) \right| d\mu(x) \\
&\quad + C_n \left(\rho^{-2} + s^{-1} \right) K \int_{C_\rho(0) \setminus C_{\frac{\rho}{2}}(0)} \Psi(s,\hat{x}) d\mu(x)
\end{aligned} \tag{4.13}$$

Note that σ_1 is controlled by C_n. By 4.8.2 with $P_0 = P + 2$, $\hat{x}_0 = 0$ and $R = 2r = \rho$ we find a $\kappa_0 \in (0,1)$ such that

$$\int_{C_\rho(0) \setminus C_{\frac{\rho}{2}}(0)} \Psi(s,\hat{x}) d\mu(x) \le C_n 2^{2P+4} (\rho^{-2}s)^{P+2} \rho^{-n} \mu\left(C_\rho(0) \right). \tag{4.14}$$

Here we chose $\kappa \le 4\kappa_0$, so $\rho^{-2}s \le \kappa \le 4\kappa_0$. Inserting (4.14) into (4.13) yields

$$\begin{aligned}
&\int_U \left| g(x) \left(\frac{\partial}{\partial t} - \Delta_\mu \right) \Psi_\rho(s,\hat{x}) \right| d\mu(x) \\
&\le \int_U |g(x)| \left(\frac{\left|\pi^\perp_{T_x\mu}(\hat{x})\right|^2}{4s^2} + \frac{|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2}{4s} \right) \Psi_\rho(s,\hat{x}) d\mu(x) \\
&\quad + \int_U |g(x)| \left| \vec{H}(x) \cdot \left(\frac{\hat{x}}{2s} \Psi_\rho(s,x) - \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \Psi(s,\hat{x}) \right) \right| d\mu(x) \\
&\quad + C_n 2^{2P+4} \left(\rho^{-2} + s^{-1} \right) (\rho^{-2}s)^{P+2} \rho^{-n} K \mu\left(C_\rho(0) \right).
\end{aligned} \tag{4.15}$$

As $\rho^{-2}s \le \kappa \le 1$ we have $\left(\rho^{-2} + s^{-1} \right) (\rho^{-2}s)^2 \le 2\kappa\rho^{-2}$. Also we can estimate $\left|\pi^\perp_{T_x\mu}(\hat{x})\right|^2 = |\hat{x} - \pi_{T_x\mu}(\hat{x})|^2 \le |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 |\hat{x}|^2$. Then (4.15) becomes

$$\begin{aligned}
&\int_U \left| g(x) \left(\frac{\partial}{\partial t} - \Delta_\mu \right) \Psi_\rho(s,\hat{x}) \right| d\mu(x) \\
&\le \int_U \frac{|\hat{x}|^2 + s}{4s^2} |g(x)| \left| \pi_{T_x\mu} - \pi_{\mathbb{R}^n} \right|^2 \Psi_\rho(s,\hat{x}) d\mu(x) \\
&\quad + \int_U |g(x)| \left| \vec{H}(x) \cdot \left(\frac{\hat{x}}{2s} \Psi_\rho(s,x) + \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \Psi(s,\hat{x}) \right) \right| d\mu(x) \\
&\quad + C_n 2^{2P+4} \kappa (\rho^{-2}s)^P \rho^{-n-2} K \mu\left(C_\rho(0) \right).
\end{aligned} \tag{4.16}$$

By Theorem 2.7 we can use Remark A.7 to calculate for almost every $x \in \mathrm{spt}\mu$ that $\vec{H} \cdot \hat{x} = \vec{H} \cdot (\pi_{\mathbb{R}^n}(\hat{x}) - \pi_{T_x\mu}(\hat{x}))$. Then with Young's inequality we can estimate

$$\left| \vec{H}(x) \cdot \frac{\hat{x}}{2s} \right| \leq \frac{1}{4} |\vec{H}(x)|^2 + |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \frac{|\hat{x}|^2}{2s^2}.$$

Hence

$$\begin{aligned} &\int_U |g(x)| \left| \vec{H}(x) \cdot \frac{\hat{x}}{2s} \Psi_\rho(s, \hat{x}) \right| d\mu(x) \\ &\leq \frac{1}{4} \int_U |g(x)| |\vec{H}(x)|^2 \Psi_\rho(s, \hat{x}) d\mu(x) \\ &\quad + \int_U |g(x)| \frac{|\hat{x}|^2}{2s^2} |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \Psi_\rho(s, \hat{x}) d\mu(x). \end{aligned} \tag{4.17}$$

Using again Young's inequality we also have

$$\left| \vec{H}(x) \cdot \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \right| \leq \frac{1}{4} |\vec{H}(x)|^2 \zeta(\rho^{-1}|\hat{x}|) + \frac{(\zeta'(\rho^{-1}|\hat{x}|))^2}{\rho^2 \zeta(\rho^{-1}|\hat{x}|)}.$$

Note that by Proposition A.6.1 the function $\zeta^{-1}(\zeta')^2$ is well defined and can be controlled by $\sup|\zeta''| \leq \sigma_1$. Also use again that $\mathrm{spt}(\zeta'(\rho^{-1}|\pi_{\mathbb{R}^n}(\cdot)|)) \subset C_\rho(0) \setminus C_{\frac{\rho}{2}}(0)$ and $\sup_U |g| \leq K$. Thus we obtain

$$\begin{aligned} &\int_U |g(x)| \left| \vec{H}(x) \cdot \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \right| \Psi(s, \hat{x}) d\mu(x) \\ &\leq \frac{1}{4} \int_U |g(x)| |\vec{H}(x)|^2 \Psi_\rho(s, \hat{x}) d\mu(x) + 2\sigma_1 \rho^{-2} K \int_{C_\rho(0) \setminus C_{\frac{\rho}{2}}(0)} \Psi(s, \hat{x}) d\mu(x). \end{aligned}$$

With (4.14) this yields

$$\begin{aligned} &\int_U |g(x)| \left| \vec{H}(x) \cdot \frac{\hat{x}}{\rho|\hat{x}|} \zeta'(\rho^{-1}|\hat{x}|) \right| \Psi(s, \hat{x}) d\mu(x) \\ &\leq \frac{1}{4} \int_U |g(x)| |\vec{H}(x)|^2 \Psi_\rho(s, \hat{x}) d\mu(x) \\ &\quad + C_n 2^{2P+4} \kappa (\rho^{-2}s)^P \rho^{-n-2} K \mu(C_\rho(0)), \end{aligned} \tag{4.18}$$

where we used again $\rho^{-2}s \leq \kappa \leq 1$. Then inserting (4.17) and (4.18) into (4.16) establishes the result for κ small enough depending on P. □

From the definition of the mean curvature vector on varifolds (see Definition 2.5) we obtain the following lemma:

4.13 Lemma. *For every $\rho, s \in (0,\infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with mean curvature vector $\vec{H}$ and*

$$\mathrm{spt}\mu \cap C_\rho(0) \subset\subset U. \tag{4.19}$$

Then for every $g \in C^2(U)$

$$\begin{aligned}&\int_U g(x)\Delta_\mu \Psi_\rho(s,\hat{x})d\mu(x)\\ &= \int_U \left(Dg(x)\cdot \vec{H}(x) + \mathrm{div}_\mu Dg(x)\right)\Psi_\rho(s,\hat{x})d\mu(x).\end{aligned} \tag{4.20}$$

Proof. First we can calculate using equality (2.3) and Remark 2.6.3

$$\begin{aligned}&\int_U g(x)\Delta_\mu \Psi_\rho(s,\hat{x})d\mu(x) - \int_U \Delta_\mu\left(g(x)\Psi_\rho(s,\hat{x})\right)d\mu(x)\\ &= -\int_U \left(\Psi_\rho(s,\hat{x})\Delta_\mu g(x) + 2\nabla^\mu g(x)\cdot\nabla^\mu\Psi_\rho(s,\hat{x})\right)d\mu(x)\\ &= -\int_U \left(\mathrm{div}_\mu Dg(x) + Dg(x)\cdot\vec{H}(x)\right)\Psi_\rho(s,\hat{x})d\mu(x)\\ &\quad - 2\int_U Dg(x)\cdot\nabla^\mu\Psi_\rho(s,\hat{x})d\mu(x).\end{aligned}$$

By assumption (4.19) we can use equality (2.2) and combine this with Remark 2.6.3 to obtain

$$\begin{aligned}&\int_U Dg(x)\cdot\nabla^\mu\Psi_\rho(s,\hat{x})d\mu(x)\\ &= \int_U \mathrm{div}_\mu\left(\Psi_\rho(s,\hat{x})Dg(x)\right) - \mathrm{div}_\mu Dg(x)\Psi_\rho(s,\hat{x})d\mu(x)\\ &= -\int_U \left(\mathrm{div}_\mu Dg(x) + Dg(x)\cdot\vec{H}(x)\right)\Psi_\rho(s,\hat{x})d\mu(x).\end{aligned}$$

Combining these two calculations already establishes

$$\begin{aligned}&\int_U g(x)\Delta_\mu \Psi_\rho(s,\hat{x})d\mu(x) - \int_U \Delta_\mu\left(g(x)\Psi_\rho(s,\hat{x})\right)d\mu(x)\\ &= \int_U \left(\mathrm{div}_\mu Dg(x) + Dg(x)\cdot\vec{H}(x)\right)\Psi_\rho(s,\hat{x})d\mu(x).\end{aligned}$$

By assumption (4.19) we can use Remark 2.6.7 to see that the Laplace term vanishes, which establishes the result. □

4.14 Lemma. *For every $P_1 \in (0,\infty)$ there exists a $\kappa_1 \in (0,1)$ such that, for every $s, \rho, K \in (0,\infty)$ with $\rho^{-2}s \leq \kappa_1$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Suppose $\mu\left(C_\rho(x_0)\right) < \infty$, then for every $g : U \to [-K, K]$*

$$\begin{aligned}
&\int_U g\vec{H}(x) \cdot D\Psi_\rho(s,\hat{x})d\mu(x) \\
&\leq \int_U \frac{|g|}{2}|\vec{H}|^2\Psi_\rho(s,\hat{x})d\mu(x) + K\left(\rho^{-2}s\right)^{P_1}\rho^{-n-2}\mu\left(C_\rho(0)\right) \\
&\quad + K\int_U \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 \frac{|\hat{x}|^2}{2s^2}\Psi_\rho(s,\hat{x})d\mu(x).
\end{aligned}$$

Proof. Set $\psi(s,x) := \Psi_\rho(s,\hat{x})$ for $x \in \mathbb{R}^{n+k}$. By Theorem 2.7 and Young's inequality we can estimate

$$\begin{aligned}
\vec{H} \cdot D\psi &= |\vec{H} \cdot \left(\pi_{T_x\mu}(D\psi) - \pi_{\mathbb{R}^n}(D\psi)\right)| \\
&\leq \frac{1}{2}|\vec{H}|^2\psi + \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 \frac{|D\psi|^2}{2\psi}.
\end{aligned}$$

Note that by Proposition A.6.1 the term $\psi^{-1}|D\psi|^2$ is well defined. Thus we have

$$\begin{aligned}
&\int_U g\vec{H} \cdot D\psi \, d\mu \\
&\leq \int_U \frac{|g|}{2}|\vec{H}|^2\psi \, d\mu + K\int_U \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 \psi^{-1}|D\psi|^2 d\mu,
\end{aligned} \tag{4.21}$$

where we estimated $\sup_U |g|$ by K. Using Remark 4.4 yields

$$\begin{aligned}
&\left(\Psi_\rho(s,\hat{x})\right)^{-1}|D\Psi_\rho(s,\hat{x})|^2 \\
&\leq \left(\Psi_\rho(s,\hat{x})\right)^{-1}\left(\frac{|\hat{x}|^2}{2s^2}\Psi_\rho^2(s,\hat{x}) + \frac{2}{\rho^2}\left(\zeta'(\rho^{-1}|\hat{x}|)\right)^2\Psi^2(s,\hat{x})\right) \\
&\leq \frac{|\hat{x}|^2}{2s^2}\Psi_\rho(s,\hat{x}) + \frac{2}{\rho^2}\left(\zeta(\rho^{-1}|\hat{x}|)\right)^{-1}\left(\zeta'(\rho^{-1}|\hat{x}|)\right)^2\Psi(s,\hat{x})
\end{aligned}$$

for all $\hat{x} \in \mathbb{R}^n$. By definition of ζ we have $\operatorname{spt}(\zeta'(\rho^{-1}|\pi_{\mathbb{R}^n}(\cdot)|)) \subset C_\rho(0)\backslash C_{\frac{\rho}{2}}(0)$. Combine this with Proposition A.6.1 and with the relation $|\zeta''(\rho^{-1}|\hat{x}|)| \leq \sigma_1$

to obtain

$$
\begin{aligned}
&\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, \psi^{-1} |D\psi|^2 d\mu \\
&\le \int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, \frac{|\hat{x}|^2}{2s^2} \psi d\mu(x) \\
&\quad + \frac{2\sigma_1}{\rho^2} \int_{C_\rho(0)\setminus C_{\frac{\rho}{2}}(0)} |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, \Psi(s,\hat{x}) d\mu(x).
\end{aligned}
\tag{4.22}
$$

Now by Lemma 4.8.2 with $P_0 = P_1 + 1$, $R = 2r = \rho$ and $t = s$ there exists a $\kappa_0 \in (0,1)$ depending on P_1 such that

$$
\int_{C_\rho(0)\setminus C_{\frac{\rho}{2}}(0)} \Psi(s,\hat{x}) d\mu(x) \le 2^{2P_1+2} (\rho^{-2}s)^{P_1+1} \rho^{-n} \mu\left(C_\rho(x_0)\right)
$$

if $\rho^{-2}s \le 4\kappa_0$, so choose $\kappa_1 \le 4\kappa_0$. As $\left|\pi_{T_x\mu_{t_0}} - \pi_{\mathbb{R}^n}\right|^2 \le 4n^2$ (by Proposition A.11) we can insert this into (4.22) to conclude

$$
\begin{aligned}
&\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, \psi^{-1} |D\psi|^2 d\mu \\
&\le \int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, \frac{|\hat{x}|^2}{2s^2} \psi d\mu(x) + C_n 2^{2P_1+2} (\rho^{-2}s)^{P_1+1} \rho^{-n-2} \mu\left(C_\rho(0)\right)
\end{aligned}
$$

Now use $\rho^{-2}s \le \kappa_1$, so for κ_1 small enough depending on P_1, we can estimate $C_n 2^{2P_1+2}(\rho^{-2}s) \le 1$. Then with (4.21) the result follows. $\square$

By Proposition 4.11, Lemma 4.13 and Lemma 4.14 we can control the right hand side of (3.5) for functions of the form $\phi = g\Psi$, for $g \in C^2(U,[0,K])$, $K \in (1,\infty)$. This lets us estimate the difference between evolution by Brakke flow and evolution by increasing the heat kernel parameter. The following result is an improvement of [B, 6.8].

4.15 Theorem (Heat Diffusion, [B, 6.8])**.** *For every $P_0 \in (0,\infty)$ there exists a $\kappa_0 \in (0,1)$ such that, for all $t_0 \in \mathbb{R}$, $\rho, p, q, \in (0,\infty)$ with $\rho^{-2}(p+q) \le \kappa_0$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[0,q]}$ be a Brakke flow in U with*

$$
\bigcup_{t\in[0,q]} \operatorname{spt}\mu_t \cap C_\rho(0) \subset\subset U. \tag{4.23}
$$

Then for every $h \in C^2(U, [-M, M])$

$$
\begin{aligned}
&\left|\int_U h(x)\Psi_\rho(p,\hat{x})d\mu_q(x) - \int_U h(x)\Psi_\rho(p+q,\hat{x})d\mu_0(x)\right| \\
&\leq 4M\Bigg[\int_0^q\int_U |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2 \frac{|\hat{x}|^2+p+q-t}{(p+q-t)^2}\Psi_\rho(p+q-t,\hat{x})d\mu_t(x)dt \\
&\qquad + \left|\int_U \Psi_\rho(p+q,\hat{x})d\mu_0(x) - \int_U \Psi_\rho(p,\hat{x})d\mu_q(x)\right| \\
&\qquad + (\rho^{-2}(p+q))^{P_0}\rho^{-n-2}\int_0^q \mu_t\left(C_\rho(0)\right)dt\Bigg] \\
&\quad + \int_0^q\int_U |\mathrm{div}_{\mu_t} Dh(x)|\,\Psi_\rho(p+q-t,\hat{x})d\mu_t(x)dt.
\end{aligned}
$$

4.16 Remark.

- If $h \geq 0$ we leave out the second term ($\int \Psi_R(s_2)d\mu_0 - \int \Psi_R(s_0)d\mu_\tau$), but in exchange only obtain the estimate without absolute value. This is more like the version Brakke has in [B, 6.8, 6.9].
- For result [B, 6.8] Brakke needs certain area ratio bounds as well as small tilt- and height-excess, as he uses Lipschitz approximations in the proof. Our theorem can be applied to any Brakke flow. However for the right hand side to be optimal, one needs small tilt-excess.
- For affine h the divergence term vanishes. Actually here the theorem will only be used with $h(x) := x_{n+j} + a$ for $j = 1, \ldots, k$, $a \in \mathbb{R}$

Proof. For $x \in \mathbb{R}^{n+k}$ and $t \in (-\infty, p+q)$ set

$$
\begin{aligned}
\psi(t,x) &:= \Psi_\rho(p+q-t,\hat{x}) \\
\overline{g}(x) &:= h(x) + M \\
\underline{g}(x) &:= -h(x) + M.
\end{aligned}
$$

Note that both $\sup_U |\overline{g}|$ and $\sup_U |\underline{g}|$ are bounded by $2M$. For $g = \overline{g}$ or $g = \underline{g}$, as g is positive and by (4.23) we can use Proposition 3.8 with $\phi = g\psi$ to estimate

$$
\begin{aligned}
\mathscr{D}(g) &:= \int_U g(x)\Psi_R(p,\hat{x})d\mu_q(x) - \int_U g(x)\Psi_R(p+q,\hat{x})d\mu_0(x) \\
&\leq \int_0^q\int_U \left(\vec{H}\cdot D(g\psi) - |\vec{H}|^2 g\psi + \frac{\partial}{\partial t}(g\psi)\right)d\mu_t\,dt.
\end{aligned}
\tag{4.24}
$$

Here we used that at almost every time the Brakke variation is not singular. Fix a time $t \in [0,q]$ where μ_t has L^2-integrable mean curvature $\vec{H}$ in U. In view of Lemma 4.13 with $s = p+q-t$ we can estimate

$$\begin{aligned}\int_U \left(\vec{H}\cdot D(g\psi) + g\frac{\partial}{\partial t}\psi\right) d\mu_t \leq & \int_U g\left(\Delta_\mu + \frac{\partial}{\partial t}\right)\psi \, d\mu_t \\ & + \int_U g\vec{H}\cdot D\psi d\mu_t \\ & + \int_U |\mathrm{div}_{\mu_t} Dg|\,\psi\, d\mu_t.\end{aligned} \tag{4.25}$$

Using the bound on the evolution equation of the heat kernel derived in Proposition 4.11 with $P = P_0$, $\mu = \mu_t$ and $s = p+q-t$ we can estimate

$$\begin{aligned}&\int_U g\left(\Delta_\mu + \frac{\partial}{\partial t}\right)\psi\, d\mu_t \\ &\leq \int_U \frac{g}{2}|\vec{H}|^2\psi\, d\mu_t + 2M\left(\rho^{-2}(p+q)\right)^{P_0}\rho^{-n-2}\mu_t\left(C_\rho(0)\right) \\ &\quad + 2M\int_U |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2\,\frac{|\hat{x}|^2 + p+q-t}{(p+q-t)^2}\psi\, d\mu_t(x).\end{aligned} \tag{4.26}$$

Here we used $p+q-t \leq p+q \leq \kappa_0 \leq \kappa$, where κ is from Proposition 4.11. Note that $\frac{\partial}{\partial t}\psi(t,x) = -\frac{\partial}{\partial t}\Psi_\rho(s,\hat{x})$. By Lemma 4.14 with $P_1 = P_0$, $\mu = \mu_{t_0+t}$ and $s = p+q-t$ we can estimate

$$\begin{aligned}&\int_U g\vec{H}\cdot D\psi d\mu_t \\ &\leq \int_U \frac{g}{2}|\vec{H}|^2\psi d\mu_t + 2M\left(\rho^{-2}(p+q)\right)^{P_0}\rho^{-n-2}\mu_t\left(C_\rho(0)\right) \\ &\quad + M\int_U |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2\,\frac{|\hat{x}|^2}{(p+q-t)^2}\psi(t,x) d\mu_t(x).\end{aligned} \tag{4.27}$$

Here we used $p+q-t \leq p+q \leq \kappa_0 \leq \kappa_1$, where κ_1 is from Lemma 4.14. Inserting (4.26) and (4.27) into (4.25) yields

$$\begin{aligned}&\int_U \left(\vec{H}\cdot D(g\psi) + g\frac{\partial}{\partial t}\psi\right) d\mu_t \\ &\leq \int_U g|\vec{H}|^2\psi\, d\mu_t + 4M\left(\rho^{-2}(p+q)\right)^{P_0}\rho^{-n-2}\mu_t\left(C_\rho(0)\right) \\ &\quad + 3M\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2\,\frac{|\hat{x}|^2 + p+q-t}{(p+q-t)^2}\psi\, d\mu_t(x) \\ &\quad + \int_U |\mathrm{div}_{\mu_t} Dh|\,\psi\, d\mu_t\end{aligned}$$

for all $t \in [0,q]$ where μ_t has L^2-integrable mean curvature $\vec{H}$ in U. Here we also used that $Dg = Dh$. We substitude this into (4.24) to conclude

$$\begin{aligned}\mathscr{D}(g) \le 4M\Bigg[&\int_0^q \int_U |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2 \frac{|\hat{x}|^2 + p + q - t}{(p+q-t)^2} \Psi_R(p+q-t, x)\, d\mu_t(x)dt \\ &+ \left(\rho^{-2}(p+q)\right)^{P_0} \rho^{-n-2} \int_0^q \mu_t\left(C_\rho(0)\right) dt \Bigg] \\ &+ \int_0^q \int_U |\mathrm{div}_{\mu_t} Dg(x)|\, \Psi_\rho(p+q-t, x)\, d\mu_t(x)dt\end{aligned}$$

for $g = \overline{g}$ or $\underline{g}$. To turn this into an estimate for h instead of g note that

$$\begin{aligned}&\int_U h(x)\Psi_\rho(p,\hat{x})d\mu_q(x) - \int_U h(x)\Psi_\rho(p+q,\hat{x})d\mu_0(x) \\ &\le \mathscr{D}(\overline{g}) + M\left|\int_U \Psi_\rho(p+q,\hat{x})d\mu_0(x) - \int_U \Psi_\rho(p,\hat{x})d\mu_q(x)\right|\end{aligned}$$

and

$$\begin{aligned}&\int_U h(x)\Psi_\rho(p+q,\hat{x})d\mu_0(x) - \int_U h(x)\Psi_\rho(p,\hat{x})d\mu_q(x) \\ &\le \mathscr{D}(\underline{g}) + M\left|\int_U \Psi_\rho(p+q,\hat{x})d\mu_0(x) - \int_U \Psi_\rho(p,\hat{x})d\mu_q(x)\right|,\end{aligned}$$

which establishes the result. □

We will use Theorem 4.15 only in integrated form. Note that in the form above we need point-wise small tilt to control the right hand side. Doing one more integration in space can be used to bring in the tilt-excess, as the following lemma shows (when considered with $\phi = |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2$).

4.17 Lemma. *Let $\rho, r, t \in (0,\infty)$, $y_0 \in \mathbb{R}^{n+k}$ and μ be a rectifiable n-varifold in $\mathbb{R}^{n+k}$. Suppose $\phi \in L^1_\mu(\mathrm{spt}\mu, \mathbb{R}^+)$ and $\vartheta \in C^0_c([-\rho,\rho],\mathbb{R}^+)$. Then we can estimate*

$$\begin{aligned}&\int_{B^n_r(\hat{y}_0)} \int_{\mathbb{R}^{n+k}} \phi(x)\vartheta(|\hat{x}-\hat{y}|)\Psi(t,\hat{x}-\hat{y})d\mu(x)d\mathscr{L}^n(\hat{y}) \\ &\le \int_{C_{r+\rho}(y_0)} \phi(x)d\mu(x) \int_{\mathbb{R}^n} \vartheta(|\hat{y}|)\Psi(t,\hat{y})d\mathscr{L}^n(\hat{y}).\end{aligned}$$

Proof. Calculate using Fubini and $\operatorname{spt}(\vartheta(|\cdot|)) \subset B^n_\rho(0)$

$$\begin{aligned}
&\int_{B^n_r(\hat{y}_0)} \int_{\mathbb{R}^{n+k}} \phi(x)\vartheta(|\hat{x}-\hat{y}|)\Psi(t,\hat{x}-\hat{y})d\mu(x)d\mathscr{L}^n(\hat{y}) \\
&= \int_{C_{r+\rho}(y_0)} \left(\phi(x) \int_{B^n_r(\hat{y}_0)} \vartheta(|\hat{x}-\hat{y}|)\Psi(t,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{y})\right) d\mu(x).
\end{aligned}$$

As $\phi, \vartheta, \Psi \geq 0$ we then can estimate

$$\begin{aligned}
&\int_{B^n_r(\hat{y}_0)} \int_{\mathbb{R}^{n+k}} \phi(x)\vartheta(|\hat{x}-\hat{y}|)\Psi(t,\hat{x}-\hat{y})d\mu(x)d\mathscr{L}^n(\hat{y}) \\
&\leq \int_{C_{r+\rho}(y_0)} \left(\phi(x) \int_{\mathbb{R}^n} \vartheta(|\hat{x}-\hat{y}|)\Psi(t,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{y})\right) d\mu(x).
\end{aligned}$$

Now the inner integral is actually independent of x which yields the result. $\square$

With Lemma 4.17 we obtain the following version of Theorem 4.15. This will be made use of in Lemma 8.1.

4.18 Lemma. *There exists a $C \in (1,\infty)$ such that for every $P_0 \in (0,\infty)$ there exists a $\kappa_0 \in (0,1)$ such that for all $t_0 \in \mathbb{R}$, $\rho, r, p, q, \gamma \in (0,\infty)$ with $\rho^{-2}(p+q) \leq \kappa_0$ and every open subset $V \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_0,t_0+q]}$ be a Brakke flow in V, and $j \in \{1,\dots,k\}$ with*

$$\bigcup_{t\in[t_0,t_0+q]} \operatorname{spt}\mu_t \cap C_{r+\rho}(0) \subset \{x \in C_{r+\rho}(0),\ |x_{n+j}| \leq \gamma\rho\} \subset\subset V. \tag{4.28}$$

Then

$$\begin{aligned}
&\int_{B^n_r(0)} \left| \int_V x_{n+j}\Psi_\rho(p,\hat{x}-\hat{y})d\mu_{t_0+q}(x) \right. \\
&\qquad \left. - \int_V x_{n+j}\Psi_\rho(p+q,\hat{x}-\hat{y})d\mu_{t_0}(x) \right| d\mathscr{L}^n(\hat{y}) \\
&\leq C\gamma\rho \Bigg[p^{-1} \int_{t_0}^{t_0+q} \int_{C_{r+\rho}(0)} |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2\, d\mu_t(x)dt \\
&\qquad + \left(\rho^{-2}(p+q)\right)^{P_0} r^n \rho^{-n-2} \int_{t_0}^{t_0+q} \mu_t\left(C_{\rho+r}(0)\right) dt \\
&\qquad + \max_{\tilde{q}\in\{0,q\}} \sup_{\hat{y}\in B^n_r(0)} r^n \left| \int_V \Psi_\rho(p+q-\tilde{q},\hat{x}-\hat{y})d\mu_{t_0+\tilde{q}}(x) - 1 \right| \Bigg].
\end{aligned}$$

Proof. First we consider the case $t_0 = 0$. For given $V \subset \mathbb{R}^{n+k}$ set

$$U := \{x \in V,\ |x_{n+j}| \leq 2\gamma\rho\}\,, \tag{4.29}$$

then by (4.28) we have $\mu_t = \mu_t \llcorner U$. In particular $(\mu_t)_{t\in[t_0,t_0+q]}$ is a Brakke flow in U as well. Note that (4.28) also holds with V replaced by U. The set U is introduced, because we want a small bound for $h(x) = x_{n+j}$, which we do not have in V.

For $t \in [0, q]$ define $\phi_t \in L^1_\mu(\mathrm{spt}\mu_t)$ by

$$\phi_t(x) := |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2$$

for $x \in \mathrm{spt}\mu_t$ such that $T_x\mu_t$ exists, which is the case for almost every $x \in \mathrm{spt}\mu_t$. For $t \in [0, q]$ and $a \in [0, \rho]$ define $\vartheta_t \in C^\infty_c([0, \rho], \mathbb{R}^+)$ by

$$\vartheta_t(a) := \frac{a^2 + p + q - t}{(p + q - t)^2}\zeta(\rho^{-1}a),$$

where ζ is the cut-off function from Definition 4.1.

For $\hat{y} \in B^n_r(0)$ set $y = (\hat{y}, 0)$. We can use Theorem 4.15 with $h(x) = x_{n+j}$ and translated in space by y to obtain

$$\begin{aligned}
&\left|\int_U x_{n+j}\Psi_\rho(p, \hat{x} - \hat{y})d\mu_q(x) - \int_U x_{n+j}\Psi_\rho(p + q, \hat{x} - \hat{y})d\mu_0(x)\right| \\
&\leq C_n\gamma\rho\Bigg[\int_0^q \int_U \phi_t(x)\vartheta_t(|\hat{x} - \hat{y}|)\Psi(p + q - t, \hat{x} - \hat{y})d\mu_t(x)ds \\
&\qquad + \left(\rho^{-2}(p + q)\right)^{d_0} \rho^{-n-2} \int_0^q \mu_{t_0+s}\left(C_\rho(y)\right)ds \\
&\qquad + \left|\int_U \Psi_\rho(p + q, \hat{x} - \hat{y})d\mu_0(x) - \int_U \Psi_\rho(p, \hat{x} - \hat{y})d\mu_q(x)\right|\Bigg]
\end{aligned} \tag{4.30}$$

Here we used that by (4.28) and (4.29) we have $\sup_{U\cap C_\rho(y)} |x_{n+j}| \leq 2\gamma\rho$. Also (4.28) holds with V replaced by U, which guarantees (4.23) with 0 replaced by y. Note that $D^2x_{n+j} = 0$ so the divergence-term vanishes. By Lemma 4.17 with $y_0 = 0$ we can estimate for $t \in [0, q]$

$$\begin{aligned}
&\int_{B^n_r(0)} \int_U \phi_t(x)\vartheta_t(|\hat{x} - \hat{y}|)\Psi(p + q - t, \hat{x} - \hat{y})d\mu_t(x)d\mathscr{L}^n(\hat{y}) \\
&\leq \int_{C_{r+\rho}(0)} \phi_t(x)d\mu_t(x) \int_{\mathbb{R}^n} \frac{|\hat{y}|^2 + p + q - t}{(p + q - t)^2}\Psi_\rho(p + q - t, \hat{y})d\mathscr{L}^n(\hat{y}),
\end{aligned}$$

where we inserted our definition of ϑ_t. As $\Psi_\rho \leq \Psi$ we can use Lemma 4.6 to estimate the second integral which yields

$$\begin{aligned}&\int_{B_r^n(0)} \int_U \phi_t(x)\vartheta_t(|\hat{x}-\hat{y}|)\Psi(p+q-t,\hat{x}-\hat{y})d\mu_t(x)d\mathscr{L}^n(\hat{y}) \\ &\leq C_n p^{-1} \int_{C_{r+\rho}(0)} |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2 \, d\mu_t(x)\end{aligned} \tag{4.31}$$

for all $t \in [0,q]$, where we inserted our definition of ϕ_t and we roughly estimated $(p+q-t)^{-1} \leq p^{-1}$. We can directly estimate for $t \in [0,q]$

$$\int_{B_r^n(0)} \mu_t\left(C_\rho(y)\right) d\mathscr{L}^n(\hat{y}) \leq \omega_n r^n \mu_t\left(C_{\rho+r}(0)\right) \tag{4.32}$$

and also by the triangle inequality

$$\begin{aligned}&\left|\int_U \Psi_\rho(p+q,\hat{x}-\hat{y})d\mu_0(x) - \int_U \Psi_\rho(p,\hat{x}-\hat{y})d\mu_q(x)\right| \\ &\leq 2 \max_{\tilde{q}\in\{0,q\}} \left| \int_U \Psi_\rho(p+q-\tilde{q},\hat{x}-\hat{y})d\mu_{\tilde{q}}(x) - 1\right|\end{aligned} \tag{4.33}$$

for all $y \in C_r^n(0)$.

Then in view of $\mu_t = \mu_t \llcorner V = \mu_t \llcorner U$ we can integrate (4.30) in $\hat{y}$ and then switch the order of integration to use estimates (4.31), (4.32) and (4.33), which establishes the result for $t_0 = 0$.

Now for arbitrary $t_0 \in \mathbb{R}$ consider $\tilde{\mu}_t := \mu_{t-t_0}$. As $(\mu_t)_{t\in[t_0,t_0+q]}$ is a Brakke flow in V , $(\tilde{\mu}_t)_{t\in[0,q]}$ is a Brakke flow in V as well. Applying the already established statement to $(\tilde{\mu}_t)$ implies the statement for (μ_t). □

We want to use that for small parameter the heat kernel converges to the Dirac delta function, as we showed in Proposition 4.10.2. The difference between $f(\hat{y})$ and $\int f(\hat{x})\Psi_\rho(p,\hat{x}-\hat{y})d\hat{x}$ can be bounded in terms of $|f|, |Df|$ and p, which we prove in the next lemma. This result is based on a calculation from [B, 6.9].

4.19 Lemma ([B, 6.9])**.** *There exists a $C \in (1,\infty)$ such that for every $P \in (0,\infty)$ there exists a $\kappa \in (0,1)$ such that for all $\rho, r, p \in (0,\infty)$ with $\rho^{-2}p \leq \kappa$ the following holds: Consider $g \in C^{0,1}\left(B_{r+\rho}^n(0)\right)$ then we can estimate*

$$\begin{aligned}&\int_{B_r^n(0)} \left|g(\hat{y}) - \int_{B_\rho^n(\hat{y})} g(\hat{x})\Psi_\rho(p,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x})\right| d\mathscr{L}^n(\hat{y}) \\ &\leq Cp^{\frac{1}{2}} \int_{B_{r+\rho}^n(0)} |Dg(\hat{x})|d\mathscr{L}^n(\hat{x}) + (\rho^{-2}p)^P \sup_{B_{r+\rho}^n(0)} |g| r^n.\end{aligned} \tag{4.34}$$

Proof. For $\hat{y} \in B_r(0)$ set

$$\phi(\hat{y}) := \left| g(\hat{y}) - \int_{B^n_\rho(\hat{y})} g(\hat{x})\Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right|. \tag{4.35}$$

By Proposition 4.10.2 and the fundamental theorem of calculus we can calculate

$$\begin{aligned}
&\int_{B^n_\rho(\hat{y})} g(\hat{x})\Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) - g(\hat{y}) \\
&= \int_{B^n_\rho(\hat{y})} g(\hat{x})\Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) - \lim_{q \searrow 0} \int_{B^n_\rho(\hat{y})} g(\hat{x})\Psi_\rho(q, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \\
&= \int_0^p \int_{B^n_\rho(\hat{y})} g(\hat{x}) \frac{d}{ds} \Psi_\rho(s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) ds.
\end{aligned}$$

Here we used that the last integral exists by Remark 4.4 and Lemma 4.6. Thus for ϕ from (4.35) we have

$$\begin{aligned}
\phi(\hat{y}) \leq &\int_0^p \left| \int_{B^n_\rho(\hat{y})} g(\hat{x}) \Delta_{\mathbb{R}^n} \Psi_\rho(s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| ds \\
&+ \int_0^p \left| \int_{B^n_\rho(\hat{y})} g(\hat{x}) \left(\frac{\partial}{\partial t} - \Delta_{\mathbb{R}^n} \right) \Psi_\rho(s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| ds
\end{aligned} \tag{4.36}$$

for all $\hat{y} \in B_r(0)$. Note that the second integral would vanish for the usual (non-truncated) heat kernel. But we already showed how to estimate the extra terms from the cut-off function. Using Lemma 4.9 with $\mu = \mathscr{L}^n$, $x_0 = (\hat{y}, 0)$ and $P_1 = P$ we find a κ_1 depending on P such that we can estimate

$$\begin{aligned}
&\left| \int_{B^n_\rho(\hat{y})} g(\hat{x}) \left(\Delta_{\mathbb{R}^n} \Psi_\rho(s, \hat{x} - \hat{y}) - \Delta_{\mathbb{R}^n} \Psi(s, \hat{x} - \hat{y}) \zeta\left(\rho^{-1} |\hat{x} - \hat{y}|\right) \right) d\mathscr{L}^n(\hat{x}) \right| \\
&\leq C_n (\rho^{-2} p)^P \rho^{-2} \sup_{B^n_{r+\rho}(0)} |g|
\end{aligned}$$

for all $\hat{y} \in B_r(0)$ and all $s \in (0, p]$, where we used $\rho^{-2} s \leq \rho^{-2} p \leq \kappa \leq \kappa_1$, for $\kappa \leq \kappa_1$. Thus we can calculate

$$\begin{aligned}
&\left| \int_{B^n_\rho(\hat{y})} g(\hat{x}) \left(\frac{\partial}{\partial t} - \Delta_{\mathbb{R}^n} \right) \Psi_\rho(s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| \\
&\leq C_n (\rho^{-2} p)^P \rho^{-2} \sup_{B^n_{r+\rho}(0)} |g|
\end{aligned} \tag{4.37}$$

for all $\hat{y} \in B_r(0)$ and all $s \in (0, p]$. Here we used that

$$\frac{\partial}{\partial t}\Psi_\rho(s,\hat{z}) = \frac{\partial}{\partial t}\Psi(s,\hat{z})\zeta\left(\rho^{-1}|\hat{z}|\right) = \Delta_{\mathbb{R}^n}\Psi(s,\hat{z})\zeta\left(\rho^{-1}|\hat{z}|\right)$$

for all $s \in (0,\infty)$ and all $\hat{z} \in \mathbb{R}^n$. Inserting (4.37) into (4.36) yields

$$\begin{aligned}\phi(\hat{y}) \leq &\int_0^p \left| \int_{B^n_\rho(\hat{y})} g(\hat{x})\Delta_{\mathbb{R}^n}\Psi_\rho(s,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x})\right| ds \\ &+ C_n(\rho^{-2}p)^{P+1}\sup_{B^n_{r+\rho}(0)}|g|\end{aligned} \tag{4.38}$$

for all $\hat{y} \in B_r(0)$ and all $s \in (0, p]$. Now integration by parts yields

$$\begin{aligned}&\int_{B^n_\rho(\hat{y})} g(\hat{x})\Delta_{\mathbb{R}^n}\Psi_\rho(s,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x}) \\ &= -\int_{B^n_\rho(\hat{y})} Dg(\hat{x})\cdot D\Psi_\rho(s,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x}).\end{aligned}$$

By Remark 4.4 we know $|D\Psi_\rho(s,\hat{z})| = \left(\frac{|\hat{z}|}{2t}\zeta(\rho^{-1}|\hat{z}|) + \rho^{-1}\zeta'(\rho^{-1}|\hat{z}|)\right)\Psi(s,\hat{z})$ for all $\hat{z} \in \mathbb{R}^n$. Then we can estimate using Lemma 4.17 with $\mu = \mathscr{L}^n$, $\phi = |Dg|$, $\vartheta(a) = \frac{a}{2t}\zeta(\rho^{-1}a) + \rho^{-1}\zeta'(\rho^{-1}a)$ and $\hat{y}_0 = \hat{y}$

$$\begin{aligned}&\int_{B^n_r(0)}\left|\int_{B^n_\rho(\hat{y})} g(\hat{x})\Delta_{\mathbb{R}^n}\Psi_\rho(s,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x})\right| d\mathscr{L}^n(\hat{y}) \\ &\leq \int_{B^n_r(0)}\int_{B^n_\rho(\hat{y})} |Dg(\hat{x})||D\Psi_\rho(s,\hat{x}-\hat{y})|d\mathscr{L}^n(\hat{x})d\mathscr{L}^n(\hat{y}) \\ &\leq \int_{B^n_{r+\rho}(0)}|Dg(\hat{x})|d\mathscr{L}^n(\hat{x})\int_{\mathbb{R}^n}|D\Psi_\rho(s,\hat{y})|d\mathscr{L}^n(\hat{y})\end{aligned} \tag{4.39}$$

for all $s \in (0,\infty)$. Using Lemma 4.9 with $\mu = \mathscr{L}^n$, $x_0 = 0$ and $P_1 = 1$ we find a κ_1 such that we can estimate

$$\int_{\mathbb{R}^n}|D\Psi_\rho(s,\hat{y})|d\mathscr{L}^n(\hat{y}) \leq \int_{\mathbb{R}^n}|D\Psi(s,\hat{y})|d\mathscr{L}^n(\hat{y}) + C_n\rho^{-3}p$$

for all $s \in (0,p]$, where we used $\rho^{-2}s \leq \rho^{-2}p \leq \kappa \leq \kappa_1$ for $\kappa \leq \kappa_1$. Here we also estimated $\zeta \leq 1$. As $D\Psi(s,\hat{y}) = (2s)^{-1}\hat{y}\Psi(s,\hat{y})$ we can then use Lemma 4.6 to estimate

$$\int_{\mathbb{R}^n}|D\Psi_\rho(s,\hat{y})|d\mathscr{L}^n(\hat{y}) \leq C_n s^{-\frac{1}{2}} + C_n\rho^{-3}p \leq C_n s^{-\frac{1}{2}}$$

for all $s \in (0, p]$, where we used $\rho^{-2} p \leq \kappa \leq 1$. Inserting this into (4.39) yields

$$\int_{B_r^n(0)} \left| \int_{B_\rho^n(\hat{y})} g(\hat{x}) \Delta_{\mathbb{R}^n} \Psi_\rho(s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| d\mathscr{L}^n(\hat{y})$$
$$\leq C_n s^{-\frac{1}{2}} \int_{B_{r+\rho}^n(0)} |Dg(\hat{x})| d\mathscr{L}^n(\hat{x})$$

for all $s \in (0, p]$. Then with (4.38) we conclude

$$\int_{B_r^n(0)} \phi(\hat{y}) d\mathscr{L}^n(\hat{y}) \leq C_n \int_0^p s^{-\frac{1}{2}} ds \int_{B_{r+\rho}^n(0)} |Dg(\hat{x})| d\mathscr{L}^n(\hat{x})$$
$$+ C_n(\rho^{-2} p)^{P+1} \sup_{B_{r+\rho}^n(0)} |g| r^n.$$

Now estimate $C_n(\rho^{-2}p) \leq C_n \kappa \leq 1$ for κ small depending on P. Also calculate $\int_0^p s^{-\frac{1}{2}} ds = 2p^{\frac{1}{2}}$. In view of (4.35) this establishes the result. $\square$

5 Clearing Out

A very important local result for mean curvature flow is the clearing out lemma, which essentially states that regions in which the flow has small mass will develop parts which do not contain the flowing surfaces at all. The clearing out lemma first appeared in Brakke's original work [B, 6.3]. Here we derive an improved version, see Lemma 5.7, based on Brakke's original calculations in [B, 6.3]. This will be used to verify an L^∞-estimate in form of Proposition 5.10, which can in turn be used in the proof of Theorem 8.4 as an alternative to Corollary 6.8. In the proof of [B, 6.9] Brakke uses a statement similar to Proposition 5.10, but his argumentation contains a major gap. In particular it is unclear why the conditions he states there allow the usage of his clearing out lemma [B, 6.3].

We will need the following two results which can be found in [All, 5.1.3] and [EG, 1.5.2] respectively.

5.1 Lemma ([All, 5.1.3])**.** *Let $R \in (0,\infty), \lambda \in [0,\infty)$ and let μ be an integral n-varifold in $\mathbb{R}^{n+k}$ with*

$$\|\delta\mu\| \left(\overline{B_r(0)}\right) \leq \lambda\mu \left(\overline{B_r(0)}\right) \tag{5.1}$$

for all $r \in (0,R)$. Then $\exp(\lambda r) r^{-n}\mu\left(\overline{B_r(0)}\right)$ is non-decreasing on $(0,R)$. Recall that the first variation of μ is defined by

$$\|\delta\mu\| (A) := \left\{ \int_A \operatorname{div}_\mu X d\mu, \;\; X \in C^1_c\left(A, \mathbb{R}^{n+k}\right), |X(x)| \leq 1 \;\; \forall x \in A \right\}$$

5.2 Theorem (Besicovitch Covering Theorem)**.** *For every $N \in \mathbb{N}$ there exists a constant $K \in \mathbb{N}$ such that the following holds: For $U \subset \mathbb{R}^N$ consider a family of balls $(B_u)_{u\in U}$, where $B_u := B_{r_u}(u)$ with $r_u \in (0,R)$, $R \in (0,\infty)$. Then there exist K subsets $V_l \subset U$ such that $U \subset \bigcup_{l=1}^K \bigcup_{u\in V_l} \overline{B_u}$ and the closures of balls in the same V_l are disjoint, that means $\overline{B_u} \cap \overline{B_v} = \emptyset$ for all $u, v \in V_l, u \neq v$ for all $l = 1,\ldots,K$.*

The following lemma is a generalization of a part of the proof of [B, 6.3]. In particular Brakke uses a fixed test-function ϕ and only considers $q = 1$. Note that the possibility to choose $q = \frac{1}{2}$ will be crucial in proving Lemma 5.7.

5.3 Lemma ([B, 6.3])**.** *For every $m \in (2,\infty)$ there exists a $\delta \in (0,1]$ such that, for all $\rho \in (0,\infty)$, $q \in (0,1]$ and every open subset $U \subset \mathbb{R}^{n+k}$ the* *following holds: Let $x_0 \in U$, μ be a rectifiable n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$ and $\Theta^n(\mu, x_0) \geq 1$. Then for every $\phi \in C_c^{0,1}(U, \mathbb{R}^+)$ there exists a radius $r \in (0, \infty)$ such that*

$$\begin{aligned} &\int_{\overline{B_r(x_0)}} |\vec{H}|^2 \phi^m d\mu + \int_{\overline{B_r(x_0)}} |\nabla^\mu \phi|^2 \phi^{m-2} d\mu + \rho^{-2} \int_{\overline{B_r(x_0)}} \phi^{m-q} d\mu \\ &\geq \delta \rho^{-2} \xi^{-\alpha} \int_{\overline{B_r(x_0)}} \phi^m d\mu \end{aligned} \tag{5.2}$$

for $\xi := \rho^{-n} \int_U \phi^m d\mu$, where $\alpha = \frac{2q}{nq+2m}$.

5.4 Remark. Note that although the proof is done by contradiction, we obtain an explicit lower bound for δ, see (5.5). Also note that r may depend on ρ. For Lemma 5.7 $B_\rho(x_0)$ will be the support of ϕ.

Proof. Set $B_r = B_r(x_0)$. We want to prove the lemma by contradiction. Assume the statement is false, then there exists an $m \in (2, \infty)$ such that, for every $\delta > 0$ there exist $\rho \in (0, \infty)$, $q \in (0, 1]$, an open subset $U \subset \mathbb{R}^{n+k}$, a rectifiable n-varifold μ in U with L^2-integrable mean curvature vector $\vec{H}$ and a point $x_0 \in U$ with $\Theta^n(\mu, x_0) \geq 1$. Also there exists a test function $\phi \in C_c^{0,1}(U, \mathbb{R}^+)$ such that for every $r \in (0, \infty)$

$$\begin{aligned} &\int_{\overline{B_r}} |\vec{H}|^2 \phi^m d\mu + \int_{\overline{B_r}} |\nabla^\mu \phi|^2 \phi^{m-2} d\mu + \rho^{-2} \int_{\overline{B_r}} \phi^{m-q} d\mu \\ &< \delta \rho^{-2} \xi^{-\alpha} \int_{\overline{B_r}} \phi^m d\mu. \end{aligned} \tag{5.3}$$

In particular as all terms are positive each term on the left hand side is smaller than the one on the right hand side. Multiply (5.3) with r^{-n}, then letting $r \searrow 0$ we obtain with the third term on the left hand side of (5.3) that

$$\rho^{-2} \phi(x_0)^{m-q} \Theta^n(\mu, x_0) \omega_n < \delta \rho^{-2} \xi^{-\alpha} \phi(x_0)^m \Theta^n(\mu, x_0) \omega_n,$$

where we used the definition of density and continuity of ϕ. As $\Theta^n(\mu, x_0) > 0$ we can conclude

$$\phi(x_0) > \delta^{-\frac{1}{q}} \xi^{\frac{\alpha}{q}}. \tag{5.4}$$

We want to use Allard's monotonicity lemma on the varifold $\mu \lfloor \phi^m$. For some $X \in C_c^1\left(\overline{B_r}, \mathbb{R}^{n+k}\right)$ with $|X(x)| \leq 1 \;\; \forall x \in \overline{B_r}$ Remark 2.63, property (2.2) and Hölder's inequality can be used to obtain

$$\begin{aligned}
&\int_{\overline{B_r}} (\operatorname{div}_\mu X)\, \phi^m d\mu \\
&= \int_{\overline{B_r}} \operatorname{div}_\mu (X\phi^m)\, d\mu - \int_{\overline{B_r}} X \cdot \nabla^\mu \phi^m d\mu \\
&\le \int_{\overline{B_r}} |\vec{H}| \phi^m d\mu + m \int_{\overline{B_r}} |\nabla^\mu \phi| \phi^{m-1} d\mu \\
&\le \left(\left(\int_{\overline{B_r}} |\vec{H}|^2 \phi^m d\mu \right)^{\frac{1}{2}} + \left(m \int_{\overline{B_r}} |\nabla^\mu \phi|^2 \phi^{m-2} d\mu \right)^{\frac{1}{2}} \right) \left(\int_{\overline{B_r}} \phi^m d\mu \right)^{\frac{1}{2}}.
\end{aligned}$$

Thus using (5.3) yields

$$\int_{\overline{B_r}} (\operatorname{div}_\mu X)\, \phi^m d\mu < (1+m)\sqrt{\delta \xi^{-\alpha}} \rho^{-1} \int_{\overline{B_r}} \phi^m d\mu = \lambda \mu \lfloor \phi^m (\overline{B_r}),$$

for $\lambda := (1+m)\sqrt{\delta\xi^{-\alpha}}\rho^{-1}$ and for every $r \in (0, \infty)$. Then Lemma 5.1 tells us that $\exp(\lambda r) r^{-n} \int_{\overline{B_r}} \phi^m d\mu$ is non-decreasing in r. Letting $r \searrow 0$ we obtain

$$\exp(\lambda r) r^{-n} \int_{\overline{B_r}} \phi^m d\mu \ge \omega_n \Theta^n(\mu, x_0) \phi(x_0)^m,$$

again by definition of density and continuity of ϕ. For $r = \lambda^{-1}$ we then obtain with estimate (5.4)

$$\exp(1)\lambda^n \int_U \phi^m d\mu > \omega_n \Theta^n(\mu, x_0) \delta^{-\frac{m}{q}} \xi^{\frac{\alpha m}{q}},$$

where we also used $\mu\left(\overline{B_r} \setminus U\right) = 0$. By definition of ξ, λ and as $\Theta^n(\mu, x_0) \ge 1$ this yields

$$\exp(1)(1+m)^n \delta^{\frac{n}{2}} \xi^{-\frac{\alpha n}{2}+1} > \omega_n \delta^{-\frac{m}{q}} \xi^{\frac{\alpha m}{q}},$$

which implies

$$\delta^{\frac{n}{2}+\frac{m}{q}} > \omega_n \exp(1)^{-1}(1+m)^{-n} \xi^{\frac{\alpha m}{q}+\frac{\alpha n}{2}-1} = \omega_n \exp(1)^{-1}(1+m)^{-n}, \tag{5.5}$$

where we used the definition of α for the last equality. In view of (5.5) we obtain a contradiction for δ small enough. $\square$

Combining this lemma with the Besicovitch covering theorem 5.2 yields the following Sobolev-type inequality. This is also a generalization of part of the proof of [B, 6.3].

5.5 Lemma ([B, 6.3]). *For every $m \in (2,\infty)$ there exists a $\Lambda_0 \in (1,\infty)$ such that for all $q \in (0,1]$, $\rho \in (0,\infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be a rectifiable n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$ and let $\phi \in C_c^{0,1}(U,\mathbb{R}^+)$ then we have*

$$\rho^{\alpha n}\left(\int_U \phi^m d\mu\right)^{1-\alpha} \leq \Lambda_0 \int_U \left(\rho^2\phi^m|\vec{H}|^2 + \rho^2\phi^{m-2}|\nabla^\mu\phi|^2 + \phi^{m-q}\right) d\mu,$$

where $\alpha := \frac{2q}{nq+2m} < 1$.

Proof. Set $\xi := \rho^{-n}\int_U \phi^m d\mu$. By Lemma 5.3 we obtain a δ depending only on m, n such that for every $x \in \mathbb{R}^{n+k}$ with $\Theta^n(\mu,x) \geq 1$ there exists a radius $r(x) \in (0,\infty)$ such that for $B_x = B_{r(x)}(x)$

$$\int_{\overline{B_x}} |\vec{H}|^2\phi^m + |\nabla^\mu\phi|^2\phi^{m-2} + \rho^{-2}\phi^{m-q} d\mu \geq \delta\rho^{-2}\xi^{-\alpha}\int_{\overline{B_x}} \phi^m d\mu. \tag{5.6}$$

As we assumed that ϕ has compact support there exists an $R_0 \in (0,\infty)$ such that $\mathrm{spt}\phi \subset B_{R_0}(0)$. For $x \in B_{2R_0}(0)$ with $\Theta^n(\mu,x) \geq 1$ set $\tilde{r}(x) := \min\{r(x), 3R_0\}$, then (5.6) also holds for $B_x = B_{\tilde{r}(x)}(x)$. For $x \notin B_{2R_0}(0)$ set $\tilde{r}(x) := R_0$, then (5.6) for $B_x = B_{\tilde{r}(x)}(x)$ trivially becomes $0 \geq 0$. So we may assume $r(x) \in (0, 3R_0)$ for all $x \in \mathbb{R}^{n+k}$ with $\Theta^n(\mu,x) \geq 1$.

Now set $A := \left\{x \in \mathbb{R}^{n+k},\ \Theta(\mu,x) \geq 1\right\}$ and consider the family $\left(\overline{B_x}\right)_{x\in A}$. By Theorem 5.2 there exists a constant $K \in \mathbb{N}$ depending only on $n+k$ and subsets $A_i \subset A$, $i = 1,\ldots,K$ such that $\overline{B_x} \cap \overline{B_y} = \emptyset$ for all $x \neq y,\ x,y \in A_i$ and $A \subset \left(\bigcup_{i=1}^K \bigcup_{x\in A_i} \overline{B_x}\right)$. Then we can estimate using (5.6)

$$\begin{aligned}
\int_U \phi^m d\mu &\leq \sum_{i=1}^K \sum_{x\in A_i} \int_{\overline{B_x}} \phi^m d\mu \\
&\leq \delta^{-1}\rho^2\xi^\alpha \sum_{i=1}^K \sum_{x\in A_i} \int_{\overline{B_x}} |\vec{H}|^2\phi^m + |\nabla^\mu\phi|^2\phi^{m-2} + \rho^{-2}\phi^{m-q} d\mu \\
&\leq \delta^{-1}\xi^\alpha K \int_U R^2|\vec{H}|^2\phi^m + \rho^2|\nabla^\mu\phi|^2\phi^{m-2} + \phi^{m-q} d\mu.
\end{aligned}$$

Now divide by ξ^α to verify the result for $\Lambda_0 = \delta^{-1}K$. □

For a Brakke flow and a test function ϕ satisfying inequality (5.8) stated below, Lemma 5.5 can be used to obtain a differential inequality (see (5.13)) for the integral of the function. Solving this inequality yields that $\mathrm{spt}\phi$ will become empty after some time. This is again a generalization of a part of Brakke's proof for [B, 6.3].

5.6 Proposition ([B, 6.3]). *For every $m \in (2, \infty)$ and $q \in (0, 1]$ there exists a $\Lambda \in (1, \infty)$ such that for $\alpha := \frac{2q}{nq+2m}$, for all $\rho \in (0, \infty)$, $\eta \in [0, \infty)$, ,$\lambda \in [0, 1]$ $t_0 \in \mathbb{R}$, $T \in (\Lambda\lambda^{-1}\eta^{\alpha}\rho^2, \infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_0,t_0+T]}$ be a Brakke flow in U and $\phi \in C^{0,1}([t_0, t_0 + T] \times U, \mathbb{R}^+)$. For $\phi_t = \phi(t, \cdot)$ suppose*

$$\bigcup_{t\in[t_0,t_0+T]} \mathrm{spt}\phi_t \subset\subset U. \tag{5.7}$$

Also assume

$$\begin{aligned} &\int_U \left(\frac{\partial}{\partial t}(\phi(t, x)^m) - \mathrm{div}_{\mu_t} D(\phi(t, x)^m)\right) d\mu_t(x) \\ &\leq -\lambda \int_U \left(|\nabla^{\mu_t}\phi(t, x)|^2\phi(t, x)^{m-2} + \rho^{-2}\phi(t, x)^{m-q}\right) d\mu_t(x) \end{aligned} \tag{5.8}$$

for almost every $t \in [t_0, t_0 + T]$ and

$$\rho^{-n} \int_U \phi(t_0, x)^m d\mu_{t_0}(x) \leq \eta. \tag{5.9}$$

Then

$$\mathrm{spt}\mu_t \cap \mathrm{spt}\phi_t = \emptyset. \tag{5.10}$$

for every $t \in [t_0 + \Lambda\lambda^{-1}\eta^{\alpha}\rho^2, t_0 + T]$.

Proof. Set $\xi(t) := \rho^{-n} \int_U \phi_t^m d\mu_t$. Let $t \in (t_0, t_0 + T)$ be a non-singular time, i.e. $\mathscr{B}(\mu_t, \phi_t^m) < \infty$. In particular the mean curvature vector $\vec{H}$ is defined and L^2-integrable on U. At such a time we can use Definition 3.1 and (2.2) to estimate

$$\mathscr{B}(U, \mu_t, \phi_t^m) + \int_U \frac{\partial}{\partial t}(\phi_t^m) d\mu_t \leq \int_U -|\vec{H}|^2\phi_t^m + \frac{\partial}{\partial t}(\phi_t^m) - \mathrm{div}_{\mu_t} D(\phi_t^m) d\mu_t.$$

With (5.8) and as $\lambda \leq 1$ we can then estimate further

$$\begin{aligned} &\mathscr{B}(U, \mu_t, \phi_t^m) + \int_U \frac{\partial}{\partial t}(\phi_t^m) d\mu_t \\ &\leq -\lambda \int_U \left(|\vec{H}|^2\phi_t^m + |\nabla^{\mu_t}\phi|^2\phi_t^{m-2} + \rho^{-2}\phi_t^{m-q}\right) d\mu_t. \end{aligned} \tag{5.11}$$

By Lemma 5.5 there exists a Λ_0 depending on m, q such that

$$\rho^{n-2} (\xi(t))^{1-\alpha} \leq \Lambda_0 \int_U |\vec{H}|^2\phi_t^m + |\nabla^{\mu}\phi_t|^2\phi_t^{m-2} + \rho^{-2}\phi_t^{m-q} d\mu_t \tag{5.12}$$

Combining (5.11) and (5.12) with Proposition 3.8 we conclude

$$\overline{D}\xi(t) \leq -\lambda\Lambda_0^{-1}\rho^{-2}\left(\xi(t)\right)^{1-\alpha} \leq 0 \tag{5.13}$$

for almost every $t \in [t_0, t_0 + T]$. Here we used that at almost every time $t \in (t_0, t_0 + T)$ the Brakke variation is non-singular. Also we needed (5.7) to apply Proposition 3.8. Note that $\xi(t) \geq 0$ for all $t \in [t_0, t_0 + T]$. Let $s \in [t_0, t_0 + T]$ be such that $\xi(t) > 0$ for all $t \in [t_0, s]$. Then with Proposition A.19 and Proposition A.20 we can estimate

$$\xi(s)^\alpha - \xi(t_0)^\alpha \leq \int_{t_0}^{s} \overline{D}_t\left(\xi(p)^\alpha\right) dp = \alpha \int_{t_0}^{s} \xi(p)^{\alpha-1}\overline{D}_t\xi(p)dp,$$

so by (5.13) we have

$$\xi(s)^\alpha - \xi(t_0)^\alpha \leq -\alpha\lambda\Lambda_0^{-1}\rho^{-2}(s - t_0).$$

Hence with (5.9) we obtain

$$s \leq t_0 + \alpha^{-1}\Lambda_0\lambda^{-1}(\xi(t_0)^\alpha - \xi(s)^\alpha)\rho^2 < t_0 + \Lambda\lambda^{-1}\eta^\alpha\rho^2,$$

where we set $\Lambda := \alpha^{-1}\Lambda_0$. So there has to exist an $s_0 \in [t_0, t_0 + \Lambda\lambda^{-1}\eta^\alpha\rho^2]$ such that $\xi(s_0) = 0$. Then Proposition A.19 and (5.13) imply $\xi(t) = 0$ for all $t \in [s_0, t_0 + T]$, which establishes the result. $\square$

Now we shall insert a specific test-function into Proposition 5.6 to obtain a clearing out result. Due to our generalizations we can chose a different Φ to the one Brakke chooses in [B, 6.3].

5.7 Lemma (Clearing Out Lemma, [B, 6.3])**.** *There exists a constant $C \in (1, \infty)$ such that for $\sigma = \frac{1}{n+12}$ for all $r, R \in (0, \infty), \eta \in [0, \infty)$, $x_0 \in \mathbb{R}^{n+k}$, $t_0 \in \mathbb{R}$, $T \in [C\eta^{2\sigma}Rr, \infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ with $C_r(x_0) \cap B_R(x_0) \subset\subset U$ the following holds: Let $(\mu_t)_{t\in[t_0,t_0+T]}$ be a Brakke flow in U and suppose*

$$(Rr)^{-\frac{n}{2}} \int_U \Phi^3 d\mu_{t_0} \leq \eta, \tag{5.14}$$

where

$$\Phi(x) := \left\{1 - R^{-2}|x - x_0|^2\right\}_+ \left\{1 - r^{-2}|\hat{x} - \hat{x}_0|^2\right\}_+. \tag{5.15}$$

Then for all $t \in [t_0 + C\eta^{2\sigma}Rr, t_0 + T]$

$$\mu_t(C_{r(t)}(x_0) \cap B_{R(t)}(x_0)) = 0, \tag{5.16}$$

where $r(t) := \sqrt{\{r^2 - 4n(t - t_0)\}_+}$ and $R(t) := \sqrt{\{R^2 - 4n(t - t_0)\}_+}$. In the special case where $R = r$ we obtain the same result with $\sigma = \frac{1}{n+6}$

5.8 Remark. Note that this is a new result. Only the special case where $R = r$ is from [B, 6.3]. In particular for the proof of Proposition 5.10 it is crucial to choose $R >> r$.

Proof. Consider the functions $f, g : \mathbb{R} \times \mathbb{R}^{n+k} \to \mathbb{R}^+$

$$\begin{aligned} f(t,x) &:= \frac{1}{R^2}\left\{R(t)^2 - |x - x_0|^2\right\}_+, \\ g(t,x) &:= \frac{1}{r^2}\left\{r(t)^2 - |\hat{x} - \hat{x}_0|^2\right\}_+. \end{aligned}$$

We want to use the previous lemma with $\phi(t,x) = f(t,x)g(t,x)$. Set again $\phi_t = \phi(t,\cdot), f_t = f(t,\cdot), g_t = g(t,\cdot)$. Let $m \geq 3$, such that ϕ^m is C^2 in both space and time. Calculate

$$\begin{aligned} &\frac{\partial}{\partial t}\left(\phi(t,x)^m\right) = m\phi_t^{m-1}\frac{\partial}{\partial t}(f(t,x)g(t,x)) \\ &= m\phi_t^{m-1}\left(R^{-2}\left(2R(t)\frac{-2n}{R(t)}\right)g(t,x) + r^{-2}\left(2r(t)\frac{-2n}{r(t)}\right)f(t,x)\right), \end{aligned}$$

so

$$\frac{\partial}{\partial t}\left(\phi(t,x)^m\right) = -4nm\left(R^{-2}g_t(x) + r^{-2}f_t(x)\right)\phi_t^{m-1}(x) \tag{5.17}$$

for every $(t,x) \in \mathbb{R} \times \mathbb{R}^{n+k}$. Consider $t \in (0,T)$ where the measure μ_t is a rectifiable n-varifold with L^2-integrable mean curvature vector $\vec{H}$, which is the case for almost every $t \in (0,T)$. For μ_t-almost every point in U the approximate tangent space exists. Then we can calculate at such a point

$$\mathrm{div}_{\mu_t} D(\phi_t^m) = m(m-1)|\nabla^{\mu_t}\phi_t|^2\phi_t^{m-2} + m\left(\mathrm{div}_{\mu_t} D\phi_t\right)\phi_t^{m-1}. \tag{5.18}$$

For $\mathrm{div}_{\mu_t} D\phi_t$ we can calculate using Remark 2.6.3

$$\begin{aligned} \mathrm{div}_{\mu_t} D\phi_t &= \mathrm{div}_{\mu_t}(f_t Dg_t + g_t Df_t) \\ &= f_t \mathrm{div}_{\mu_t} Dg_t + g_t \mathrm{div}_{\mu_t} Df_t + 2\nabla^{\mu_t} f_t \cdot \nabla^{\mu_t} g_t \\ &= -2r^{-2} f_t \mathrm{div}_{\mu_t}(\hat{x}) - 2R^{-2} g_t \mathrm{div}_{\mu_t}(x) + 2\nabla^{\mu_t} f_t \cdot \nabla^{\mu_t} g_t, \end{aligned}$$

which yields the estimate

$$\mathrm{div}_{\mu_t} D\phi_t \geq -2n\left(R^{-2}g_t + r^{-2}f_t\right) + 2\nabla^{\mu_t} f_t \cdot \nabla^{\mu_t} g_t. \tag{5.19}$$

Also estimate

$$2\left(\nabla^{\mu_t} f_t \cdot \nabla^{\mu_t} g_t\right)\phi = 2g_t\nabla^{\mu_t} f_t \cdot f_t\nabla^{\mu_t} g_t \leq \left|g_t\nabla^{\mu_t} f_t + f_t\nabla^{\mu_t} g_t\right|^2,$$

so

$$2\left(\nabla^{\mu_t} f_t \cdot \nabla^{\mu_t} g_t\right)\phi \leq |\nabla^{\mu_t}\phi_t|^2. \tag{5.20}$$

Inserting (5.19) and (5.20) into (5.18) yields

$$\begin{aligned}\operatorname{div}_{\mu_t} D(\phi_t^m) \geq{}& m(m-1)|\nabla^{\mu_t}\phi_t|^2\phi_t^{m-2} - m|\nabla^{\mu_t}\phi_t|^2\phi_t^{m-2}\\ &- 2nm\left(R^{-2}g_t + r^{-2}f_t\right)\phi_t^{m-1}\end{aligned} \tag{5.21}$$

for almost every $t \in [t_0, t_0+T]$, for μ_t-almost every $x \in U$ and for all $m \in [3,\infty)$. Now set $m = 3$. By combining (5.21) with (5.17) we obtain

$$\begin{aligned}\frac{\partial}{\partial t}\left(\phi(t,x)^3\right) - \operatorname{div}_{\mu_t}\left(\phi_t(x)^3\right) \leq{}& \Big(-12n\left(R^{-2}g_t + r^{-2}f_t\right)\phi_t^2 - 6|\nabla^{\mu_t}\phi_t|^2\phi_t\\ &+ 6n\left(R^{-2}g_t + r^{-2}f_t\right)\phi_t^2 + 3|\nabla^{\mu_t}\phi_t|^2\phi_t\Big)\Big|_x\\ ={}& \Big(-6n\left(R^{-2}g_t + r^{-2}f_t\right)\phi_t^2 - 3|\nabla^{\mu_t}\phi_t|^2\phi_t\Big)\Big|_x\end{aligned}$$

thus as ϕ is positive

$$\frac{\partial}{\partial t}\left(\phi(t,x)^3\right) - \operatorname{div}_{\mu_t}\left(\phi_t(x)^3\right) \leq -\Big((Rr)^{-1}\phi_t^{\frac{5}{2}} + |\nabla^{\mu_t}\phi_t|^2\phi_t\Big)\Big|_x \tag{5.22}$$

for almost every $t \in [t_0, t_0+T]$ and for μ_t-almost every $x \in U$. Here we used that

$$R^{-2}g_t + r^{-2}f_t \geq \sqrt{R^{-2}g_t r^{-2}f_t} = (Rr)^{-1}\phi_t^{\frac{1}{2}}.$$

Then the result follows from Proposition 5.6 with $m = 3, q = \frac{1}{2}$, $\lambda = 1$ and $\rho = \sqrt{Rr}$. Note that $\operatorname{spt}\phi_t = C_{r(t)}(x_0) \cap B_{R(t)}(x_0)$ and $\phi(t_0,\cdot) = \Phi$.

In the special case $R = r$ the same calculation with $\phi(t,x) = f(t,x)$ yields a slightly better estimate. In this case we can use Proposition 5.6 with $q = 1$ which yields the better σ. □

For a Brakke flow starting with an integral height bound the clearing out lemma yields a point-wise height bound.

5.9 Lemma. *There exists a $C \in (1,\infty)$ such that for $\sigma := \frac{1}{n+6}$ for all $R \in (0,\infty)$, $l, \Gamma, \eta \in [0,\infty)$ and $\delta \in (0,(4n)^{-1})$, $t_0 \in \mathbb{R}$ the following holds: Let $(\mu_t)_{t\in[t_0,t_0+\delta R^2]}$ be a Brakke flow in $C_{2R}(0)$ satisfying*

$$\mu_{t_0}\left(C_{2R}(0) \setminus (B^n_{2R}(0) \times B^k_\Gamma(0))\right) \leq \eta R^n. \tag{5.23}$$

Then

$$\operatorname{spt}\mu_t \cap C_R(0) \subset \{x \in C_R(0) : |\pi_{\mathbb{R}^k}(x)| \leq 4n\delta R + \Gamma\} \tag{5.24}$$

for all $t \in [t_0 + C\eta^{2\sigma}R^2, t_0 + \delta R^2]$. Note that this interval is empty unless η is sufficiently small.

Proof. Let $s \in [t_0 + C\eta^{2\sigma}R^2, t_0 + \delta R^2]$ and $y \in \mathrm{spt}\mu_s \cap C_R(0)$ be arbitrary. First suppose

$$|\pi_{\mathbb{R}^k}(y)| \geq R + \Gamma$$

Then (5.23) yields

$$\mu_{t_0}(B_R(y)) \leq \eta R^n.$$

Thus we can use Lemma 5.7 with $r = R$ and $x_0 = y$ to obtain

$$\mu_t(B_{R(t)}(y)) = 0$$

for all $t \in [t_0 + C_n\eta^{2\sigma}, t_0 + \delta R^2]$, where $R(t) := \sqrt{R^2 - 4n(t - t_0)}$. As $\delta < (4n)^{-1}$ and by assumption $s \in [t_0 + C_n\eta^{2\sigma}, t_0 + \delta R^2]$ we can estimate $R(s) \geq \sqrt{R^2 - 4n\delta R^2} > 0$. In particular $y \notin \mathrm{spt}\mu_s$, which is a contradiction.

Now suppose

$$0 < |\pi_{\mathbb{R}^k}(y)| < R + \Gamma, \tag{5.25}$$

then set

$$\begin{aligned} v &:= |y - \hat{y}|^{-1}(y - \hat{y}) \\ a_0 &:= \hat{y} + (\Gamma + R)v. \end{aligned}$$

We want to use the clearing out lemma around a_0 with $r = R$. To do so we have to show that $B_R(a_0)$ has small μ_{t_0}-measure. First note that as $v \in \{0\}^n \times \mathbb{R}^k$ and $\hat{y} \in B^n_R(0)$ we have

$$|\pi_{\mathbb{R}^k}(a_0)| = \Gamma + R \quad \text{and} \quad |\hat{a}_0| \leq R.$$

In particular we see

$$B_R(a_0) \subset C_{2R}(0) \setminus (B^n_{2R}(0) \times B^k_\Gamma(0)).$$

Then (5.23) yields

$$\mu_{t_0}(B_R(a_0)) \leq \eta R^n.$$

Thus we can use Lemma 5.7 with $r = R$, $x_0 = a_0$ to obtain

$$\mu_t(B_{R(t)}(a_0)) = 0$$

for all $t \in [t_0 + C_n\eta^{2\sigma}, t_0 + \delta R^2]$, where $R(t) := \sqrt{R^2 - 4n(t-t_0)}$. As $\delta \leq (4n)^{-1}$ we can estimate

$$R(t) \geq \sqrt{R^2 - 4n\delta R^2} = \sqrt{1-4n\delta}R \geq (1-4n\delta)R = R - 4n\delta R$$

for all $t \in [t_0 + C_n\eta^{2\sigma}, t_0 + \delta R^2]$. In particular

$$\mu_s(B_{R-4n\delta R}(a_0)) = 0.$$

Thus by choice of s and y we have

$$|y - a_0| \geq R - 4n\delta R. \tag{5.26}$$

By definition of a_0 and v we can calculate

$$|y - a_0| = |\pi_{\mathbb{R}^k}(y - a_0)| = |\pi_{\mathbb{R}^k}(y) - (\Gamma + R)v| = ||\pi_{\mathbb{R}^k}(y)| - (\Gamma + R)|.$$

Thus (5.26) implies

$$||\pi_{\mathbb{R}^k}(y)| - (\Gamma + R)| \geq R - 4n\delta R. \tag{5.27}$$

A case distinction in (5.27) then yields

$$|\pi_{\mathbb{R}^k}(y)| \leq 4n\delta R + \Gamma, \quad \text{or} \quad |\pi_{\mathbb{R}^k}(y)| \geq 2R + \Gamma - 4n\delta R > R + \Gamma,$$

where we used $\delta < (4n)^{-1}$. The second case contradicts our assumption (5.25). Thus we obtain the height bound and as s, y where arbitrary this establishes the result. $\square$

A similar approach is now used to prove Brakke's height estimate, which is a different version of Lemma 5.9. Here we already assume a hight bound which then will be improved. Note that Brakke does not state a result like Proposition 5.10, but the proof here mostly follows a calculation from [B, 6.9]. However, the argumentation in [B, 6.9] contains a major gap, as Brakke indirectly uses Lemma 5.7 in our generalized form, though he only proved the usual spherical clearing out lemma (Lemma 5.7 with R=r).

5.10 Proposition (Height Estimate, [B, 6.9])**.** *There exists a $c \in (0,1)$ such that for all $\lambda, \delta \in (0,1]$, $\rho \in (0,\infty)$, $h \in (0, \lambda^{-1}]$, $s_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+k}$, $\Lambda := \lambda^{-1}\rho$ and every open subset $U \subset \mathbb{R}^{n+k}$ with $C_{4\sqrt{n}\rho}(y_0) \cap B_{2\Lambda}(y_0) \subset\subset U$ the following holds: Let $(\mu_t)_{t\in[s_0-2\delta\rho^2, s_0+\rho^2]}$ be a Brakke flow in U and $j \in \{1, \ldots k\}$ and suppose there exists an $s \in [s_0 - 2\delta\rho^2, s_0 - \delta\rho^2]$ such that*

$$\bigcup_{t\in[s,s_0+\rho^2]} \mathrm{spt}\mu_t \cap C_{4\sqrt{n}\rho}(y_0) \subset \{x \in U, (x - y_0)\cdot \mathbf{e}_{n+j} \leq h\rho\}, \tag{5.28}$$

$$\int_{C_{4\sqrt{n}R}(y_0)} \{\Lambda - |x - y_0 - \Lambda \mathbf{e}_{n+j}|\}_+ \, d\mu_s(x) \leq c\lambda^3\delta^{\frac{n+12}{2}}h^{-2}\rho^{n+1}. \tag{5.29}$$

Then for all $t \in [s_0, s_0 + \rho^2]$

$$\mathrm{spt}\mu_t \cap B_\rho(y_0) \subset \{x \in U, (x - y_0) \cdot \mathbf{e}_{n+j} \leq (24n + 1)\lambda\rho\}\,. \tag{5.30}$$

5.11 Remark. Basically λ denotes the new height bound, whereas h is the height bound we already have. The time one has to wait until the new height bound comes into effect is related to δ. In particular for fixed δ and h the point-wise height estimate (5.30) becomes smaller the smaller the integral in (5.29) is. Note that (5.29) is similar to an integral height bound.

Proof. We may assume $(24n + 1)\lambda \leq h$ and as $h \leq \lambda^{-1}$ this yields $(24n + 1)\lambda^2 \leq 1$. First note that by (5.28) and $h\rho \leq \lambda^{-1}\rho = \Lambda$ we know that $(x - y_0 - \Lambda\mathbf{e}_{n+j}) \cdot \mathbf{e}_{n+j}$ is negative, so

$$\begin{aligned}(x - y_0) \cdot \mathbf{e}_{n+j} &= (x - y_0 - \Lambda\mathbf{e}_{n+j}) \cdot \mathbf{e}_{n+j} + \Lambda \\ &= \Lambda - |(x - y_0 - \Lambda\mathbf{e}_{n+j}) \cdot \mathbf{e}_{n+j}|\end{aligned} \tag{5.31}$$

for all $x \in \mathrm{spt}\mu_t \cap C_{4\sqrt{n}\rho}(y_0)$, $t \in [s, s_0 + \rho^2]$. Then using (5.31) we can calculate

$$\begin{aligned}&(x - y_0) \cdot \mathbf{e}_{n+j} \\ &= \Lambda - \left(|(x - y_0 - \Lambda\mathbf{e}_{n+j}|^2 - \sum_{1\leq i\leq n+k,\, i\neq n+j} |(x - y_0) \cdot \mathbf{e}_i|^2\right)^{\frac{1}{2}} \\ &\leq \Lambda - \sqrt{|(x - y_0 - \Lambda\mathbf{e}_{n+j}|^2 - |x - y_0|^2}\end{aligned} \tag{5.32}$$

for all $x \in \mathrm{spt}\mu_t \cap C_{4\sqrt{n}\rho}(y_0)$, $t \in [s, s_0 + \rho^2]$. We want to show that for all $t \in [s_0, s_0 + \rho^2]$

$$\mathrm{spt}\mu_t \cap C_\rho(y_0) \cap B_{\Lambda_0}(y_0 + \Lambda\mathbf{e}_{n+j}) = \emptyset \tag{5.33}$$

for $\Lambda_0 := \Lambda - 12n\lambda\rho > 0$. Suppose (5.33) would be true, then in view of (5.32) we can estimate for all $t \in [s_0, s_0 + \rho^2]$ and for all $x \in \mathrm{spt}\mu_t \cap B_\rho(y_0)$

$$(x - y_0) \cdot \mathbf{e}_{n+j} \leq \Lambda - \sqrt{|x - (y_0 + \Lambda\mathbf{e}_{n+j})|^2 - |x - y_0|^2} \leq \Lambda - \sqrt{\Lambda_0^2 - \rho^2}.$$

By definition of Λ_0 and by $\Lambda = \lambda^{-1}\rho$ we can estimate further

$$\begin{aligned}(x - y_0) \cdot \mathbf{e}_{n+j} &\leq \Lambda\left(1 - \sqrt{(1 - 12n\lambda^2)^2 - \lambda^2}\right) \leq \Lambda\left(1 - \sqrt{1 - (24n + 1)\lambda^2}\right) \\ &\leq (24n + 1)\Lambda\lambda^2 \leq (24n + 1)\lambda\rho.\end{aligned}$$

which establishes (5.30). Here we used $(24n+1)\lambda^2 \le 1$. So it remains to show (5.33).

To verify (5.33) we will use Lemma 5.7 with $R = \Lambda$, $r = 4\sqrt{n}\rho$, $t_0 = s$ and $x_0 = y_0 + \Lambda \mathbf{e}_{n+j}$, so we have to consider the test function

$$\Phi(x) := \left\{1 - \Lambda^{-2}|x - y_0 - \Lambda \mathbf{e}_{n+j}|^2\right\}_+ \left\{1 - (4\sqrt{n}\rho)^{-2}|\hat{x} - \hat{y}_0|^2\right\}_+ .$$

Set

$$\eta := (4\sqrt{n}\Lambda\rho)^{-\frac{n}{2}} \int_U \Phi^3 d\mu_s, \tag{5.34}$$

then Lemma 5.7 yields

$$\mathrm{spt}\mu_t \cap C_{r(t)}(y_0) \cap B_{R(t)}(y_0 + \Lambda \mathbf{e}_{n+j}) = \emptyset \tag{5.35}$$

for all $t \in [s + C_n \eta^{\frac{2}{n+12}} \Lambda\rho, s_0 + \rho^2]$, where $r(t) := \sqrt{16n\rho^2 - 4n(t-s)}$ and $R(t) := \sqrt{\Lambda^2 - 4n(t-s)}$. But for $t \in [s + C_n \eta^{\frac{2}{n+12}} \Lambda\rho, s_0 + \rho^2]$ we have $t - s \le s_0 + \rho^2 - (s_0 - 2\delta\rho^2) \le 3\rho^2$, so we can estimate

$$\begin{aligned} R(t) &\ge \sqrt{\Lambda^2 - 12n\rho^2} = \sqrt{1 - 12n\lambda^2}\Lambda \ge \left(1 - 12n\lambda^2\right)\Lambda = \Lambda_0 \\ r(t) &\ge \sqrt{16n\rho^2 - 12n\rho^2} \ge \rho \end{aligned}$$

for all $t \in [s + C_n \eta^{\frac{2}{n+12}} \Lambda\rho, s_0 + \rho^2]$. Thus if $s + C_n \eta^{\frac{2}{n+12}} \Lambda\rho \le s_0$ then (5.35) implies (5.33), which establishes the result. By assumption $s \le s_0 - \delta\rho^2$, so it suffices to show $C_n \eta^{\frac{2}{n+12}} \Lambda\rho \le \delta\rho^2$. In view of (5.34) it remains to prove

$$C_n \left((\Lambda\rho)^{-\frac{n}{2}} \int_U \Phi^3 d\mu_s \right)^{\frac{2}{n+12}} \le \Lambda^{-1}\delta\rho. \tag{5.36}$$

To establish inequality (5.36) we calculate for $x \in \mathrm{spt}\mu_s \cap C_{4\sqrt{n}\rho}(y_0) \cap B_\Lambda(y_0 + \Lambda \mathbf{e}_{n+j})$ using (5.31) and (5.28)

$$\begin{aligned} \Phi(x)^3 &\le \left\{1 - \Lambda^{-2}|x - y_0 - \Lambda \mathbf{e}_{n+j}|^2\right\}_+^3 \\ &= \left\{\Lambda^{-6}(\Lambda + |x - y_0 - \Lambda \mathbf{e}_{n+j}|)^3(\Lambda - |x - y_0 - \Lambda \mathbf{e}_{n+j}|)^3\right\}_+ \\ &\le \Lambda^{-6}(\Lambda + \Lambda)^3 |(x - y_0)\cdot \mathbf{e}_{n+j}|^2 \left\{\Lambda - |x - y_0 - \Lambda \mathbf{e}_{n+j}|\right\}_+ \\ &\le \Lambda^{-3} h^2\rho^2 \left\{\Lambda - |x - y_0 - \Lambda \mathbf{e}_{n+j}|\right\}_+ . \end{aligned}$$

Then with assumption (5.29) we obtain

$$\begin{aligned} C_n \int_U \Phi^3 d\mu_s &\le C_n \Lambda^{-3} h^2 \rho^2 \int_{C_{4\sqrt{n}R}(y_0)} \left\{\Lambda - |x - y_0 - \Lambda \mathbf{e}_{n+j}|\right\}_+ d\mu_s(x) \\ &\le C_n c \lambda^6 \delta^{\frac{n+12}{2}} \rho^n, \end{aligned}$$

where we used $\Lambda = \lambda^{-1}\rho$ and $\mathrm{spt}\Phi \subset C_{4\sqrt{n}R}(y_0) \cap B_\Lambda(y_0 + \Lambda \mathbf{e}_{n+j})$. This can now be used to estimate

$$C_n \left((\Lambda\rho)^{-\frac{n}{2}} \int_U \Phi^3 d\mu_s \right)^{\frac{2}{n+12}} \leq C_n \left(c\lambda^{6+\frac{n}{2}} \right)^{\frac{2}{n+12}} \delta = C_n c^{\frac{2}{n+12}} \delta\lambda.$$

Now as $\lambda = \Lambda^{-1}\rho$ this implies (5.36) for c small enough. As before (5.36) then establishes the result. □

5.12 Lemma. *For every $\Lambda \in (0,\infty)$ and every $j \in \{1, \dots k\}$ we have*

$$\{\Lambda - |x - \Lambda \mathbf{e}_{n+j}|\}_+ \leq \left\{ x \cdot \mathbf{e}_{n+j} - (2\Lambda)^{-1} |\hat{x}|^2 \right\}_+ \tag{5.37}$$

for all $x \in \mathbb{R}^{n+k}$.

For the proof we will use the following fact

5.13 Remark. For $a, b, c \in \mathbb{R}^+$ with $a^2 \geq b^2 + c^2$ we can estimate

$$a - \sqrt{b^2 + c^2} \leq \sqrt{a^2 - b^2} - c. \tag{5.38}$$

To verify this compute

$$a^2c^2 \leq b^2(a^2 - b^2 - c^2) + c^2a^2 = (b^2 + c^2)(a^2 - b^2),$$

which implies

$$\begin{aligned} (a+c)^2 = a^2 + 2ac + c^2 &\leq a^2 + 2\sqrt{b^2 + c^2}\sqrt{a^2 - b^2} + c^2 + b^2 - b^2 \\ &= \left(\sqrt{b^2 + c^2} + \sqrt{a^2 - b^2} \right)^2 \end{aligned}$$

and this verifies the result.

Proof. To illustrate this statement note that $\hat{x} \to (2\Lambda)^{-1} |\hat{x}|^2$ is a paraboloid and $B_\Lambda(\Lambda \mathbf{e}_{n+j})$ the best fitting ball through 0. Consider the functions $f, g \in C^\infty(B^n_\Lambda(0), \mathbb{R}^+)$

$$\begin{aligned} f(\hat{x}) &:= (2\Lambda)^{-1} |\hat{x}|^2 \\ g(\hat{x}) &:= \Lambda - \sqrt{\Lambda^2 - |\hat{x}|^2}. \end{aligned} \tag{5.39}$$

For the derivatives we can calculate

$$\begin{aligned} \frac{\partial}{\partial x_i} f(\hat{x}) &= \Lambda^{-1} \hat{x} \cdot \mathbf{e}_i \\ \frac{\partial}{\partial x_i} g(\hat{x}) &= \left(\Lambda^2 - |\hat{x}|^2 \right)^{-\frac{1}{2}} \hat{x} \cdot \mathbf{e}_i \end{aligned}$$

for $1 \le i \le n$ and $\hat{x} \in B^n_\Lambda(0)$. Note that $f(0) = g(0) = 0$. Also for $\hat{v} \in B^n_\Lambda(0)$ and $\theta \in [0,1]$ we can estimate

$$D(g-f)(\theta\hat{v}) \cdot \hat{v} = \left(\Lambda^2 - \theta^2|\hat{v}|^2\right)^{-\frac{1}{2}} \theta|\hat{v}|^2 - \Lambda^{-1}\theta|\hat{v}|^2 \ge 0.$$

Then by Taylor's formula

$$(g-f)(\hat{v}) = (g-f)(0) + \int_0^1 D(g-f)(\theta\hat{v}) \cdot \hat{v}\, d\theta \ge 0,$$

so we conclude

$$g(\hat{x}) \ge f(\hat{x}) \tag{5.40}$$

for all $\hat{x} \in B^n_\Lambda(0)$.

Next we want to show

$$\Lambda - |x - \Lambda\mathbf{e}_{n+j}| \le x \cdot \mathbf{e}_{n+j} - g(\hat{x}) \tag{5.41}$$

for all $x \in B_\Lambda(\Lambda\mathbf{e}_{n+j})$. Inequality (5.41) basically says that for a point inside a ball the shortest distance to the boundary is smaller than the distance to the lower boundary point on the same vertical line. For $x \in \mathbb{R}^{n+k}$ we can estimate

$$\begin{aligned}
\Lambda - |x - \Lambda\mathbf{e}_{n+j}| &= \Lambda - \left(\sum_{i=1}^n |x \cdot \mathbf{e}_i|^2 + \sum_{i=1}^k |(x - \Lambda\mathbf{e}_{n+j}) \cdot \mathbf{e}_{n+i}|^2\right)^{\frac{1}{2}} \\
&\le \Lambda - \sqrt{|\hat{x}|^2 + |(x - \Lambda\mathbf{e}_{n+j}) \cdot \mathbf{e}_{n+j}|^2} \\
&= \Lambda - \sqrt{|\hat{x}|^2 + (\Lambda - x \cdot \mathbf{e}_{n+j})^2}.
\end{aligned}$$

If $x \in B_\Lambda(\Lambda\mathbf{e}_{n+j})$ we have $(\Lambda - x \cdot \mathbf{e}_{n+j})^2 + |\hat{x}|^2 \le |x - \Lambda\mathbf{e}_{n+j}|^2 \le \Lambda^2$, so we can use (5.38) to obtain

$$\Lambda - |x - \Lambda\mathbf{e}_{n+j}| \le \sqrt{\Lambda^2 - |\hat{x}|^2} - (\Lambda - x \cdot \mathbf{e}_{n+j}) = x \cdot \mathbf{e}_{n+j} - g(\hat{x})$$

for all $x \in B_\Lambda(\Lambda\mathbf{e}_{n+j})$, where we used definition (5.39) in the last step. Thus we proved (5.41) which in view of (5.40) yields

$$\left\{\Lambda - |x - \Lambda\mathbf{e}_{n+j}|\right\}_+ \le x \cdot \mathbf{e}_{n+j} - f(\hat{x})$$

for all $x \in B_\Lambda(\Lambda\mathbf{e}_{n+j})$ and by definition (5.39) this verifies (5.37). Here we used that for points outside $B_\Lambda(\Lambda\mathbf{e}_{n+j})$ the left hand side of (5.37) is zero and the right hand side is positive. $\square$

The form in which Proposition 5.10 will be used is the following:

5.14 Lemma. *There exist $c \in (0,1)$ and $C \in (1,\infty)$ such that for all $\tau \in (0,1]$, $R \in (0,\infty)$, $h \in (0,1]$, $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+k}$ and every open subset $V \subset \mathbb{R}^{n+k}$ with $B_{2\tau^{-1}R}(y_0) \subset\subset V$ the following holds: Let $(\mu_t)_{t\in[t_0-2R^2,t_0+R^2]}$ be a Brakke flow in U and T be an n-dimensional subspace in $\mathbb{R}^{n+k}$ with*

$$\bigcup_{t\in[t_0-2R^2,t_0+R^2]} \mathrm{spt}\mu_t \cap C^T_{4\sqrt{n}R}(y_0) \subset \left\{x \in V, \left|\pi_T^\perp(x-y_0)\right| \leq hR\right\}. \tag{5.42}$$

Suppose there exists a $t_1 \in [t_0-2R^2, t_0-R^2]$ and an orthonormal basis $(\nu_j)_{1\leq j\leq k}$ of the co-space $T^\perp$ such that

$$R^{-n-1}\int_{C^T_{4\sqrt{n}R}(y_0)} \left\{|(x-y_0)\cdot\nu_j| - \frac{|\pi_T(x-y_0)|^2}{2\tau^{-1}R}\right\}_+ d\mu_{t_1}(x) \leq c\frac{\tau^3}{h^2} \tag{5.43}$$

for every $j \in \{1,\dots k\}$. Then for all $t \in [t_0, t_0+R^2]$

$$\mathrm{spt}\mu_t \cap B_R(y_0) \subset \left\{x \in V, \left|\pi_T^\perp(x-y_0)\right| \leq C\tau R\right\}. \tag{5.44}$$

Here $C^T_r(x_0) := \{x \in \mathbb{R}^{n+k} : |\pi_T(x-x_0)| \leq r\}$.

Proof. Fix an arbitrary $j \in \{1,\dots k\}$ and a sign $\star \in \{+,-\}$. Fix an associated rotation $S \in SO(n+k)$ with $S(\mathbb{R}^n) = T$, $S(\mathbb{R}^k) = T^\perp$ and in particular $S(\mathbf{e}_{n+j}) = \star\nu_j$. Consider the Brakke flow $(\tilde{\mu}_t)_{t\in[t_0-3R^2,t_0+R^2]}$ defined by

$$\tilde{\mu}_t(A) := \mu_t(S(A)),$$

for all $A \subset \mathbb{R}^{n+k}$, where $S(A) := \{S(a), a \in A\}$. Assumptions (5.42) and (5.43) imply for the rotated flow

$$\bigcup_{t\in[t_0-2R^2,t_0+R^2]} \mathrm{spt}\tilde{\mu}_t \cap C_{4\sqrt{n}R}(y_0) \subset \left\{x \in U, |\pi_{\mathbb{R}^k}(x-y_0)| \leq hR\right\}, \tag{5.45}$$

$$R^{-n-1}\int_{C_{4\sqrt{n}R}(y_0)} \left\{|(x-y_0)\cdot\mathbf{e}_{n+j}| - \frac{|\hat{x}-\hat{y}_0|^2}{2\tau^{-1}R}\right\}_+ d\tilde{\mu}_{t_1}(x) \leq c\frac{\tau^3}{h^2} \tag{5.46}$$

where $U := S^{-1}(V)$. By Lemma 5.12 with $\Lambda = \tau^{-1}R$ estimate (5.46) implies

$$R^{-n-1}\int_{C_{4\sqrt{n}R}(y_0)} \left\{\Lambda - |x-y_0-\Lambda\mathbf{e}_{n+j}|\right\}_+ d\tilde{\mu}_{t_1}(x) \leq c\tau^3h^{-2}. \tag{5.47}$$

Here we used $(x-y_0)\cdot\mathbf{e}_{n+j} \le |(x-y_0)\cdot\mathbf{e}_{n+j}|$. Now we can apply Proposition 5.10 with $\delta = 1$, $\lambda = \tau$, $\rho = R$, $s_0 = t_0$ and $s = t_1$. Note that (5.45) implies (5.28), as $(x-y_0)\cdot\mathbf{e}_{n+j} \le |\pi_{\mathbb{R}^k}(x-y_0)|$. Proposition 5.10 then yields

$$\mathrm{spt}\tilde{\mu}_t \cap B_R(y_0) \subset \{x \in U, (x-y)\cdot\mathbf{e}_{n+j} \le C\tau R\}\,.$$

for all $t \in [t_0, t_0+R^2]$ for some $C \in (1,\infty)$. Thus by definition of $(\tilde{\mu}_t)$ we obtain

$$\mathrm{spt}\mu_t \cap B_R(y_0) \subset \{x \in V, \star(x-y)\cdot\nu_j \le C\tau R\}\,.$$

As $j \in \{1,\dots k\}$ and $\star \in \{+,-\}$ were arbitrary this establishes (5.44) $\square$

6 Monotonicity Formula

We want to prove Huisken's monotonicity formula 6.2 taken from [H2] for the spherical heat kernel on Brakke flows, which has been established by Ilmanen in [I2]. This will lead to the $L^\infty - L^2$-estimate Corollary 6.8, an important tool for proving local regularity later. Here we follow [E4, chapter 4], where smooth flows are considered, but the results carry over to our case.

To prove the monotonicity formula we need the spherical heat kernel Φ. Later we will use the spherically shrinking cut-off function φ defined below, to obtain local estimates.

6.1 Definition. Let $x_0 \in \mathbb{R}^{n+k}$, $t_0 \in \mathbb{R}$, $\rho \in (0,\infty)$ be fixed. For $x \in \mathbb{R}^{n+k}$ and $t \in (-\infty, t_0)$ set

$$\Phi_{(t_0,x_0)}(t,x) := (4\pi(t_0-t))^{-\frac{n}{2}}\, exp\left(\frac{|x-x_0|^2}{4(t-t_0)}\right).$$

$$\varphi_{(t_0,x_0),\rho}(t,x) := \left\{1-\rho^{-2}\left(|x-x_0|^2+2n(t-t_0)\right)\right\}_+^3.$$

6.2 Theorem (Monotonicity Formula, [H2])**.** *Consider an open subset $U \subset \mathbb{R}^{n+k}$, $(t_0,x_0) \in \mathbb{R}\times U$ and $s_0 \in (-\infty,t_0)$. Let $(\mu_t)_{t\in[s_0,t_0]}$ be a Brakke flow in U and let $f \in C^2\left([s_0,t_0]\times\mathbb{R}^{n+k},\mathbb{R}^+\right)$ with*

$$\bigcup_{t\in[s_0,t_0]} \operatorname{spt} f(t,\cdot) \subset\subset U. \tag{6.1}$$

Then for $s_0 < s_1 < s_2 < t_0$

$$\begin{aligned}&\int_U f(s_2,x)\Phi_{(t_0,x_0)}(s_2,x)d\mu_{s_2}(x) - \int_U f(s_1,x)\Phi_{(t_0,x_0)}(s_1,x)d\mu_{s_1}(x)\\ &\le \int_{s_1}^{s_2}\int_U\left(\left(\frac{\partial}{\partial t}-\Delta_{\mu_t}+\vec{H}\cdot D\right)f-\left|\vec{H}-\vartheta_{(t_0,x_0)}\right|^2 f\right)\Phi_{(t_0,x_0)}d\mu_t\,dt,\end{aligned}$$

where $\vartheta_{(t_0,x_0)}(t,x) = \frac{\pi^\perp_{T_x\mu_t}(x-x_0)}{2(t-t_0)}$.

Applying this theorem with $f=\varphi$ yields a local monotonicity formula for the heat kernel. This was discovered by Ecker, see [E4].

6.3 Remark ([E4, 4.8])**.** Under the above assumptions for $\rho \in (0,\infty)$ with $B_\rho(x_0) \subset\subset U$ and $f = \varphi_{(t_0,x_0),\rho}$ we can use Definition 2.5 and Remark 2.6 to estimate

$$\left(\frac{\partial}{\partial t}-\Delta_{\mu_t}+\vec{H}\cdot D\right)\varphi_{(t_0,x_0),\rho}(t,x)\le 0$$

for almost every $t \in [s_0, t_0]$ and almost every $x \in \mathrm{spt}\mu_t$. Then by Theorem 6.2

$$\int_U \Phi_{(t_0,x_0)}\varphi_{(t_0,x_0),\rho}\, d\mu_{s_2} \leq \int_U \Phi_{(t_0,x_0)}\varphi_{(t_0,x_0),\rho}\, d\mu_{s_1} \tag{6.2}$$

for all $s_0 < s_1 < s_2 < t_0$.

Proof of Theorem 6.2. We fix (t_0, x_0) throughout the proof and just write Φ, omitting the index. Consider times $t \in (s_1, s_2)$ where μ_t is integral and has L^2-integrable mean curvature vector $\vec{H}$. Let ∇, div and Δ be with respect to μ_t. Then we can calculate for the heat kernel by Definition 2.5

$$\begin{aligned}\left(\vec{H}D + \Delta\right)\Phi &= \mathrm{div}D\Phi + 2\vec{H}\cdot D\Phi \\ &= \mathrm{div}D\Phi + \frac{\left|\nabla^\perp\Phi\right|^2}{\Phi} - \left|\vec{H} - \frac{\nabla^\perp\Phi}{\Phi}\right|^2 \Phi + |\vec{H}|^2\Phi\end{aligned} \tag{6.3}$$

at points $x \in \mathrm{spt}\mu_t$ where $\vec{H} \perp T_x\mu_t$, which are μ_t-almost all x due to Theorem 2.7. Using Definition 6.1 and definition of ϑ_{t_0,x_0} one can directly calculate

$$\begin{aligned}\frac{\partial\Phi}{\partial t} + \mathrm{div}D\Phi + \frac{\left|\nabla^\perp\Phi\right|^2}{\Phi} &= 0, \\ \nabla^\perp\Phi(t,x) = \frac{\pi^\perp_{T_x\mu_t}(x - x_0)}{2(t-t_0)}\Phi(t,x) &= \vartheta_{(t_0,x_0)}(t,x)\Phi(t,x).\end{aligned}$$

Combining this with (6.3) yields

$$\begin{aligned}&\left(\frac{\partial}{\partial t} + \vec{H}\cdot D + \Delta\right)\Phi - |\vec{H}|^2\Phi \\ &= -\left|\vec{H} - \frac{\nabla^\perp\Phi}{\Phi}\right|^2\Phi = -\left|\vec{H} - \vartheta_{(t_0,x_0)}\right|^2\Phi\end{aligned} \tag{6.4}$$

for almost every time $t \in (s_1, s_2)$ at μ_t-almost every point $x \in \mathrm{spt}\mu_t$.

Now we integrate in time over (s_1, s_2) and use inequality (3.5) to obtain

$$\begin{aligned}&\int_U f(s_2,x)\Phi(s_2,x)d\mu_{s_2}(x) - \int_U f(s_1,x)\Phi(s_1,x)d\mu s_1(x) \\ &\leq \int_{s_1}^{s_2}\left(\int_U \left(\frac{\partial}{\partial t} + \vec{H}\cdot D\right)(f\Phi) - |\vec{H}|^2 f\Phi d\mu_t\right)dt \\ &\leq \int_{s_1}^{s_2}\int_U \Bigg[\Phi\left(\frac{\partial}{\partial t} - \Delta + \vec{H}\cdot D\right)f \\ &\qquad + \left(\left(\frac{\partial}{\partial t} + \vec{H}\cdot D + \Delta\right)\Phi - |\vec{H}|^2\Phi\right)f\Bigg]d\mu_t\, dt\end{aligned}$$

and together with equality (6.4) this establishes the result. $\square$

6.4 Definition (Gaussian Density, [H2]). Consider $t_1, t_2 \in \mathbb{R}$ with $t_1 < t_2$ and an open subset $U \subset \mathbb{R}^{n+k}$. Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U. For all $t_0 \in (t_1, t_2]$, $\rho \in (0, \infty)$ and $x_0 \in \mathbb{R}^{n+k}$ with $B_\rho(x_0) \subset\subset U$ set

$$\Theta(\mu, t_0, x_0) := \lim_{t\nearrow t_0} \int_U \Phi_{(t_0,x_0)}(t, x)\ \varphi_{(t_0,x_0),\rho}(t, x)\ d\mu_t(x). \tag{6.5}$$

This is called the *Gaussian density* of (μ_t) at (t_0, x_0). Actually this limit always exists and is independent of ρ. Note that the Gaussian density is different from the density Θ^n in Definition 2.1.4.

Proof. By Remark 6.3 we have that $t \to \int_{\mathbb{R}^{n+k}} \Phi_{(t_0,x_0)}\varphi_{(t_0,x_0),\rho}\ d\mu_t$ is monotonically decreasing for $t < t_0$. Also it is bounded from below by 0, so the limit for $t \nearrow t_0$ has to exist.

Now let $0 < \rho_1 < \rho_2 < \infty$ with $B_{\rho_2}(x_0) \subset\subset U$. In particular there exists $R \in (\rho_2, \infty)$ such that $B_R(x_0) \subset\subset U$. For arbitrarily small $\epsilon \in (0, 1)$ consider $t \in (t_0 - \epsilon\rho_1^2, t_0)$. We may assume ϵ is small enough such that

$$\mathrm{spt}\varphi_{(t_0,x_0),\rho_i}(t, \cdot) \subset B_R(x_0) \subset\subset U \tag{6.6}$$

for $i \in \{1, 2\}$. For $x \in B_{\sqrt{\epsilon}\rho_1}(x_0)$ we can estimate

$$|\varphi_{(t_0,x_0),\rho_i}(t, x) - 1| = 1 - (1 - \rho_i^{-2}(|x - x_0|^2 + 2n(t - t_0)))^3 \le C_n\epsilon \tag{6.7}$$

for $i \in \{1, 2\}$, where we used that $\rho_i^{-2}\rho_1^2 \le 1$ and that ϵ small enough. The Gaussian density difference between $\rho = \rho_1$ and $\rho = \rho_2$ can be estimated by

$$\begin{aligned} D :=& \int_U \Phi_{(t_0,x_0)} \left|\varphi_{(t_0,x_0),\rho_2} - \varphi_{(t_0,x_0),\rho_1}\right| d\mu_t \\ \le& \int_U \Phi_{(t_0,x_0)} \left(\left|\varphi_{(t_0,x_0),\rho_2} - 1\right| + \left|\varphi_{(t_0,x_0),\rho_1} - 1\right|\right) d\mu_t. \end{aligned}$$

Thus with (6.7) and (6.6)

$$D \le C_n\epsilon \int_{B_R(x_0)} \Phi_{(t_0,x_0)} d\mu_t + 2 \int_{B_R(x_0)\setminus B_{\sqrt{\epsilon}\rho_1}(x_0)} \Phi_{(t_0,x_0)} d\mu_t. \tag{6.8}$$

Note that by Lemma 3.6 and (6.6) there exists an $M \in (1, \infty)$ such that $\mu_t(B_R(x_0)) \le M$ for all $t \in [t_1, t_2]$. Then as Φ is bounded, the first integral in (6.8) is bounded by a constant times ϵ. The second integral in (6.8) can be estimated by

$$\int_{B_R(x_0)\setminus B_{\sqrt{\epsilon}\rho_1}(x_0)} \Phi_{(t_0,x_0)} d\mu_t \le C_n(t_0 - t)^{-\frac{n}{2}} \exp\left(-\frac{\epsilon\rho_1^2}{4(t_0 - t)}\right) M$$

and this expression tends to 0 for $t \nearrow t_0$. Thus we can find a $\delta \in (0,1)$ such that for all $t \in (t_0 - \delta, t_0)$ the estimate $D \leq C_n \epsilon$ holds. As ϵ is arbitrarily small, this establishes the independence of ρ. $\square$

Next we prove a lower bound on the Gaussian density like [E4, 4.20]. This time it actually is more difficult than in the smooth case, see also [W2].

6.5 Proposition. *For every $T \in (0,\infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[-T,0]}$ be a Brakke flow in U. Suppose μ_{t_0} is an integral n-varifold for some $t_0 \in (-T,0]$, then for $\mathscr{H}^n$-almost every $x_0 \in$* $\mathrm{spt}\mu_{t_0}$

$$\Theta(\mu, t_0, x_0) \geq 1.$$

Proof. As μ_{t_0} is an integral n-varifold we have $\Theta^n(\mu_{t_0}, x) \geq 1$, $T_x\mu_{t_0}$ exists and (2.1) holds for $\mathscr{H}^n$-almost every $x \in \mathrm{spt}\mu_{t_0}$. Let $x_0 \in U$ be such a point and let $\epsilon \in (0,1)$ be given. By Remark 2.3 for every ball $B_r(x_0)$, $r \in (0,\infty)$ we can calculate

$$\begin{aligned}
&\lim_{\delta\searrow 0} \int_{B_r(x_0)} \Phi_{(t_0+\delta,x_0)}(t_0,x)\, d\mu_{t_0}(x) \\
&= (4\pi)^{-\frac{n}{2}} \lim_{\delta\searrow 0} \left(\sqrt{\delta}\right)^{-n} \int_{B_r(x_0)} e^{-\left|\frac{x_0-x}{16\sqrt{\delta}}\right|^2} d\mu_{t_0}(x) \\
&= (4\pi)^{-\frac{n}{2}} \Theta^n(\mu_{t_0}, x_0) \int_{T_{x_0}\mu_{t_0}} e^{-\frac{|y|^2}{4}} d\mathscr{L}^n(y) = \Theta^n(\mu_{t_0}, x_0) \geq 1,
\end{aligned}$$

where we identified $T_{x_0}\mu_{t_0}$ with $\mathbb{R}^n$. So there exists $\delta_1 \in (0, (2n)^{-1}\epsilon\rho^2)$ such that

$$\int_{B_{\sqrt{\epsilon}\rho}(x_0)} \Phi_{(t_0+\delta,x_0)}\, d\mu_{t_0} \geq 1 - \epsilon \tag{6.9}$$

for all $\delta \in (0,\delta_1)$. There exists $\rho \in (0, \sqrt{T + t_0})$ such that $B_\rho(x_0) \subset\subset U$. By Definition 6.4 we can find a $t \in (t_0 - \rho^2, t_0)$ such that

$$\Theta(\mu, t_0, x_0) \geq \int_{\mathbb{R}^{n+k}} \Phi_{(t_0,x_0)}(t,x)\varphi_{(t_0,x_0),\rho}(t,x)\, d\mu_t(x) - \epsilon. \tag{6.10}$$

By the continuity of our test functions we can choose $\delta \in (0,\delta_1)$ such that

$$\begin{aligned}
&\int_{\mathbb{R}^{n+k}} \Phi_{(t_0+\delta,x_0)}(t,x)\varphi_{(t_0+\delta,x_0),\rho}(t,x)\, d\mu_t(x) \\
&\leq \int_{\mathbb{R}^{n+k}} \Phi_{(t_0,x_0)}(t,x)\varphi_{(t_0,x_0),\rho}(t,x)\, d\mu_t(x) + \epsilon.
\end{aligned} \tag{6.11}$$

Now combine estimates (6.10) and (6.11) to obtain

$$\begin{aligned}\Theta(\mu, t_0, x_0) &\geq \int_{\mathbb{R}^{n+k}} \Phi_{(t_0,x_0)}(t,x)\varphi_{(t_0,x_0),\rho}(t,x)d\mu_t(x) - \epsilon \\ &\geq \int_{\mathbb{R}^{n+k}} \Phi_{(t_0+\delta,x_0)}(t,x)\varphi_{(t_0+\delta,x_0),\rho}(t,x)d\mu_t(x) - 2\epsilon.\end{aligned}$$

Then with the monotonicity (6.2) we can conclude

$$\Theta(\mu, t_0, x_0) \geq \int_{\mathbb{R}^{n+k}} \Phi_{(t_0+\delta,x_0)}(t_0,x)\varphi_{(t_0+\delta,x_0),\rho}(t_0,x)d\mu_{t_0}(x) - 2\epsilon. \tag{6.12}$$

Now use that $\delta \leq \delta_1 \leq (2n)^{-1}\epsilon\rho$, so for $x \in B_{\sqrt{\epsilon}\rho}(x_0)$ we can estimate

$$\varphi_{(t_0+\delta,x_0),\rho}(t_0,x) = (1-\rho^{-2}(|x-x_0|^2 + 2n\delta))^3 \geq (1-2\epsilon)^3.$$

Inserting this into (6.12) we can use (6.9) to finally estimate

$$\Theta(\mu, t_0, x_0) \geq (1-2\epsilon)^3 \int_{B_{\sqrt{\epsilon}\rho}(x_0)} \Phi_{(t_0+\delta,x_0)}(t_0,x)d\mu_{t_0}(x) - 2\epsilon \geq (1-2\epsilon)^4 - 2\epsilon,$$

and for $\epsilon \searrow 0$ this establishes the result. □

The monotonicity formula can now be used to prove a mean value inequality. The proof we will give follows [E4, 4.25] and [E5, 2.1]. In [KT, 6.5] a similar result can be found for a more general flow but with fixed function f.

6.6 Proposition (Mean Value Inequality, [E4, 4.25], [E5, 2.1], [KT, 6.5])**.** *There exists a constant $C \in (1,\infty)$ such that for all $T \in (0,\infty)$, $t_0 \in \mathbb{R}$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_0-T,t_0]}$ be a Brakke flow in U such that μ_{t_0} is an integral n-varifold. Let $f \in C^2([t_0-T,t_0]\times U,\mathbb{R})$ be such that for almost every $t_0 - T \leq t \leq t_0$ for μ_t-almost every $x \in U$*

$$\left(\frac{\partial}{\partial t} - \Delta_{\mu_t} + \vec{H}\cdot D\right) f(t,x) \leq 0. \tag{6.13}$$

Then for all $\rho \in (0, 2^{-1}\sqrt{T})$ and $a \in \mathbb{R}^{n+k}$ with $B_{2\rho}(a) \subset U$ the inequality

$$|f(s,y)| \leq C\rho^{-n-2}\int_{t_0-4\rho^2}^{t_0}\int_{B_{2\rho}(a)} |f(t,x)|d\mu_t(x)dt \tag{6.14}$$

holds, for all $s \in [t_0-\rho^2, t_0]$ and all $y \in \operatorname{spt}\mu_s \cap B_\rho(a)$.

6.7 Remark. Here only the L^2-version of (6.14) will be proven. Then we refer to [E5, 2.1], where an advanced calculus trick due to Schoen and Bartnik is used to derive the L^1-version from its corresponding L^2-estimate.

Proof. Fix $a \in \mathbb{R}^{n+k}$ and $\rho \in (0, 2^{-1}\sqrt{T})$ with $B_{2\rho}(a) \subset U$. First we will assume $f(t,x) \geq 0$ for all $(t,x) \in [t_0 - T, t_0] \times U$. We want to show

$$f(s,y)^2 \leq C\rho^{-n-2} \int_{t_0-4\rho^2}^{t_0} \int_{B_{2\rho}(a)} f(t,x)^2 d\mu_t(x)\, dt \tag{6.15}$$

for all $s \in [t_0 - \rho^2, t_0]$ and all $y \in \mathrm{spt}\mu_s \cap B_\rho(a)$. Fix $s_0 \in [t_0 - 2\rho^2, t_0]$ and $y_0 \in \mathrm{spt}\mu_{s_0} \cap B_\rho(a)$ with $\Theta(\mu, s_0, y_0) \geq 1$. We want to show (6.15) holds for $(s,y) = (s_0, y_0)$.

Consider a time t such that μ_t is integral and has L^2-integrable mean curvature vector $\vec{H}$. Let ∇ and Δ be with respect to μ_t, then by Remark 2.6.3 we can calculate at almost every point

$$\left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) f^2 = 2f \left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) f - 2\left|\nabla f\right|^2. \tag{6.16}$$

Furthermore for a $\phi \in C^2\left([-T,0] \times \mathbb{R}^{n+k}, \mathbb{R}\right)$ we can estimate using Young's inequality

$$\nabla f^2 \cdot \nabla \phi^2 = 4f\phi \nabla f \cdot \nabla \phi \leq \phi^2 \left|\nabla f\right|^2 + 4f^2 \left|\nabla \phi\right|^2. \tag{6.17}$$

Combining (6.16) and (6.17) we obtain

$$\begin{aligned}
&\left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) f^2\phi^2 \\
&\leq \phi^2 \left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) f^2 + f^2 \left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) \phi^2 - 2\nabla f^2 \cdot \nabla \phi^2 \\
&\leq 2f\phi^2 \left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) f + f^2 \left[\left(\frac{\partial}{\partial t} - \Delta + \vec{H} \cdot D\right) \phi^2 + 8\left|\nabla \phi\right|^2\right].
\end{aligned}$$

at almost every point in $\mathrm{spt}\mu_t$. Integrating in space and time we can drop the first term, as it is negative by assumption (6.13) and as we assumed $f \geq 0$, such that by Theorem (6.2) we have

$$\begin{aligned}
&\int_U f^2\phi^2\Phi_{(s_0,y_0)} d\mu_s - \int_U f^2\phi^2\Phi_{(s_0,y_0)} d\mu_{s_0-\rho^2} \\
&\leq \int_{s_0-\rho^2}^{s} \int_U f^2 \Phi_{(s_0,y_0)} C_\phi \, d\mu_t \, dt =: I
\end{aligned} \tag{6.18}$$

for all $s \in (s_0 - \rho^2, s_0)$, where $C_\phi := \left(\frac{\partial}{\partial t} - \Delta_{\mu_t} + \vec{H} \cdot D\right) \phi^2 + 8 \left|\nabla^{\mu_t} \phi\right|^2$. Now choose $\phi(x) = \phi_0(x - y_0)$ where

$$\phi_0(t, x) = \begin{cases} 1, & (t, x) \in (s_0 - \frac{\rho^2}{4}, 0) \times B^{n+k}_{\frac{\rho}{2}}(0) \\ 0, & t \leq s_0 - \frac{\rho^2}{2} \text{ or } |x| \geq \rho \end{cases}$$

and smooth in between such that $C_\phi \leq \rho^{-2} C_n$. Then the $\mu_{s_0 - \rho^2}$-integral in (6.18) vanishes, so (6.18) becomes $\int_U f^2 \phi^2 \Phi_{(s_0, y_0)} d\mu_s \leq I$. Furthermore $C_\phi = 0$ on $(s_0 - \frac{\rho^2}{4}, 0) \times B_{\frac{\rho}{2}}(x_0)$ and outside $B_\rho(x_0)$. Then we can estimate the integral I from (6.18) by

$$\begin{aligned} I &\leq \frac{C_n}{\rho^2} \left(\int_{s_0 - \rho^2}^{s_0 - \frac{\rho^2}{4}} \int_{B_{\frac{\rho}{2}}} f^2 \Phi_{(s_0, y_0)} \, d\mu_t \, dt + \int_{s_0 - \rho^2}^{s_0} \int_{B_\rho \setminus B_{\frac{\rho}{2}}} f^2 \Phi_{(s_0, y_0)} \, d\mu_t \, dt \right) \\ &\leq \frac{C_n}{\rho^2} \left(\int_{s_0 - \rho^2}^{s_0 - \frac{\rho^2}{8n}} \int_{B_\rho} f^2 \Phi_{(s_0, y_0)} \, d\mu_t \, dt + \int_{s_0 - \frac{\rho^2}{8n}}^{s_0} \int_{B_\rho \setminus B_{\frac{\rho}{2}}} f^2 \Phi_{(s_0, y_0)} \, d\mu_t \, dt \right) \end{aligned}$$

where all the balls are centred in y_0. Now with Definition 6.1 and as $r \to r^{-\alpha} e^{-\beta r^{-1}}$ is monotonously increasing on $(0, \beta \alpha^{-1}]$ we can estimate for $t \in \left[s_0 - \frac{\rho^2}{8n}, s_0\right)$ and $x \notin B_{\frac{\rho}{2}}(y_0)$

$$\begin{aligned} \Phi_{(s_0, y_0)}(t, x) &\leq C_n (s_0 - t)^{-\frac{n}{2}} \exp\left(- \frac{\rho^2}{16(s_0 - t)} \right) \\ &\leq C_n \left(\frac{\rho^2}{8n} \right)^{-\frac{n}{2}} \exp\left(- \frac{1}{2n} \right) \leq C_n \rho^n \end{aligned}$$

Also Definition 6.1 implies $\Phi_{(s_0, y_0)}(t, x) \leq C_n \rho^n$ for all $t \in \left[s_0 - \rho^2, s_0 - \frac{\rho^2}{8n}\right]$. Then we obtain for I

$$I \leq \frac{C_n}{\rho^{n+2}} \int_{s_0 - \rho^2}^{s_0} \int_{B_\rho} (f(t, x))^2 \, d\mu_t(x) dt.$$

In view of (6.18) and by definition of ϕ this lets us conclude

$$\begin{aligned} \int_{B_{\frac{\rho}{2}}(y_0)} f^2 \Phi_{(s_0, y_0)} d\mu_s &= \int_{B_{\frac{\rho}{2}}(y_0)} f^2 \phi^2 \Phi_{(s_0, y_0)} d\mu_s \\ &\leq \frac{C_n}{\rho^{n+2}} \int_{s_0 - \rho^2}^{s_0} \int_{B_\rho(y_0)} f^2 \, d\mu_t \, dt \end{aligned} \tag{6.19}$$

for all $s \in (s_0 - \frac{\rho^2}{4}, s_0)$. By continuity of f we find $r \in (0, 4^{-1}\rho)$ such that

$$\inf_{[s_0-r^2,s_0]\times B_{2r}(y_0)} f(t,x)^2 \geq \frac{1}{2} f(s_0,y_0)^2.$$

Furthermore for $s \in (s_0 - \frac{r^2}{4n}, s_0)$ by Definition 6.1 we have $\mathrm{spt}\varphi_{(s_0,y_0),r}(s,x) \subset B_{2r}(y_0) \subset B_{\frac{\rho}{2}}(y_0)$, as well as $\varphi_{(s_0,y_0),r}(s,x) \leq 2$ for all $x \in \mathbb{R}^{n+k}$. Thus with (6.19) we conclude

$$\begin{aligned} &f(s_0,y_0)^2 \int_{B_{2r}(y_0)} \Phi_{(s_0,y_0)} \varphi_{(s_0,y_0),r}(s,x) d\mu_s(x) \\ &\leq 4 \int_{B_{\frac{\rho}{2}}(y_0)} f^2 \Phi_{(s_0,y_0)} d\mu_s \leq \frac{C_n}{\rho^{n+2}} \int_{t_0-4\rho^2}^{t_0} \int_{B_{2\rho}(a)} f^2 \, d\mu_t \, dt \end{aligned}$$

for $s \in (s_0 - \frac{r^2}{4n}, s_0) \subset (s_0 - \frac{\rho^2}{4}, s_0)$, where we used $4r \leq \rho$. Also we used $B_\rho(y) \subset B_{2\rho}(a)$ and $[s_0 - \rho^2, s_0] \subset [t_0 - 4\rho^2, t_0]$. Then by Definition 6.4 for $s \nearrow s_0$, as we assumed $\Theta(\mu, s_0, y_0) \geq 1$ this establishes estimate (6.15) for $(s,y) = (s_0,y_0)$.

As (s_0,y_0) was arbitrary this shows (6.15) holds for all $s \in [t_0 - 2\rho^2, t_0]$ and all $y \in \mathrm{spt}\mu_s \cap B_\rho(a)$ with $\Theta(\mu,s,y) \geq 1$. Due to continuity of f and Proposition 6.5 we can extend this to all $y \in \mathrm{spt}\mu_s \cap B_\rho(a)$, $s \in [t_0 - \rho^2, t_0]$.

To see this let $s \in [t_0 - \rho^2, t_0]$ and let $y \in \mathrm{spt}\mu_s \cap B_\rho(a)$ with $\Theta(\mu,s,y)$ arbitrary. For $\epsilon \in (0,1)$ choose $r \in (0,\rho)$ such that

$$(f(s,y))^2 \leq \inf_{[s-r^2,s+r^2]\times B_r(y)} (f(s_0,y_0))^2 + \epsilon, \tag{6.20}$$

which is always possible by continuity of f. As $y \in \mathrm{spt}\mu_s \cap B_\rho(a)$ we have $\mu_s(B_\delta(y)) > 0$ for all $\delta \in (0,\infty)$, see Remark 2.1. Then use Lemma 3.12.2 with $\kappa = \frac{1}{2}$, $x_0 = y$ and $R = 2\delta$ to see

$$\mu_{s_0}(B_{2\delta}(y)) > 0$$

for all $s_0 \in (s - n^{-1}\delta, s]$ for all $\delta \in (0,\infty)$. Then choose δ small enough such that $2\delta \leq r < \rho$ and $B_{2\delta}(y) \subset B_\rho(a)$. There exists $s_0 \in (s - n^{-1}\delta, s]$ where μ_{s_0} is integral and using Proposition 6.5 we find $y_0 \in \mathrm{spt}\mu_{s_0} \cap B_{2\delta}(y)$ with $\Theta(\mu, s_0, y_0) \geq 1$, in particular (6.15) holds for $(s,y) = (s_0,y_0)$. Also, by choice of δ we have $(s_0,y_0) \in [s - r^2, s + r^2] \times B_r(y)$. Then we can use (6.20) and (6.15) to estimate

$$(f(s,y))^2 \leq (f(s_0,y_0))^2 + \epsilon \leq \frac{C_n}{\rho^{n+2}} \int_{t_0-4\rho^2}^{t_0} \int_{B_{2\rho}(a)} f^2 \, d\mu_t \, dt + \epsilon$$

and for $\epsilon \searrow 0$ we obtain (6.15) for all $s \in [t_0-\rho^2, t_0]$ and all $y \in \operatorname{spt}\mu_s \cap B_\rho(a)$.

So we showed the L^2-version of (6.14) for positive functions f. To obtain the L^1-estimate we refer to [E5, 2.1]. Note that the trick used there is just a calculus iteration argument. In particular smoothness of the surfaces or the flow equation are not needed. This yields

$$|f(s,y)| \leq C\rho^{-n-2} \int_{t_0-4\rho^2}^{t_0} \int_{B_{2\rho}(a)} |f(t,x)| d\mu_t(x) dt \tag{6.21}$$

holds, for all $s \in [t_0-\rho^2, t_0]$ and all $y \in \operatorname{spt}\mu_s \cap B_\rho(a)$ in the case $f \geq 0$.

Now consider $f \in C^2([t_0-T, t_0] \times U, \mathbb{R})$ without sign condition (but satisfying (6.13)). For $\epsilon \in (0,1)$ consider $\sqrt{f^2+\epsilon^2}$, then by Remark 2.6.3

$$\begin{aligned}
&\left(\frac{\partial}{\partial t} - \Delta_{\mu_t} + \vec{H} \cdot D\right) \sqrt{f^2+\epsilon^2} \\
&= \frac{f}{\sqrt{f^2+\epsilon^2}} \left(\frac{\partial}{\partial t} - \Delta_{\mu_t} + \vec{H} \cdot D\right) f - \frac{\epsilon^2}{(f^2+\epsilon^2)^{\frac{3}{2}}} |\nabla^\mu f|^2 \leq 0
\end{aligned}$$

for almost every time $t \in [t_0-T, t_0]$ and μ_t-almost every point in U, where we used that f satisfies (6.13). Thus $\sqrt{f^2+\epsilon^2}$ is positive and satisfies (6.13), so by (6.21) we obtain

$$\sqrt{f(s,y)^2+\epsilon^2} \leq C\rho^{-n-2} \int_{t_0-4\rho^2}^{t_0} \int_{B_{2\rho}(a)} \sqrt{f(t,x)^2+\epsilon^2} d\mu_t(x) dt$$

for all $s \in [t_0-\rho^2, t_0]$ and all $y \in \operatorname{spt}\mu_s \cap B_\rho(a)$. Then for $\epsilon \to 0$ follows the result, as $\sqrt{f^2+\epsilon^2} \to |f|$. □

As a corollary we obtain a distance estimate like in [E4, 4.26]. There already exists a similar result for weak mean curvature flow, see [KT, 6.5].

6.8 Corollary (Distance Estimate, [E4, 4.26], [KT, 6.5])**.** *There exists a constant $C \in (1,\infty)$ such that for all $r \in (0,\infty)$, $t_0 \in \mathbb{R}$, $x_0, y_0, v \in \mathbb{R}^{n+k}$ and every open subset $U \subset \mathbb{R}^{n+k}$ with $B_{2r}(x_0) \subset U$ the following holds: Let $(\mu_t)_{t\in[t_0-r^2,t_0]}$ be a Brakke flow in U then*

$$|(y-y_0) \cdot v| \leq Cr^{-n-2} \int_{t_0-4r^2}^{t_0} \int_{B_{2r}(x_0)} |(x-y_0) \cdot v| d\mu_t(x)\, dt \tag{6.22}$$

for all $s \in [t_0-r^2, t_0]$ and all $y \in \operatorname{spt}\mu_s \cap B_r(x_0)$.

Proof. Consider the function $f(x) = (x - y_0) \cdot v$ then f is C^∞ and satisfies

$$\left(\frac{\partial}{\partial t} - \Delta\mu_t + \vec{H} \cdot D\right)(x - x_0) \cdot v = -\mathrm{div}_{\mu_t}((x - y_0) \cdot v) = 0$$

where we used Definition 2.4 and Definition 2.5. Thus the result follows directly from Proposition 6.6 with $\rho = r$ and $a = x_0$ $\square$

7 Bounds On Area Ratio

In order to approximate Brakke flow by heat diffusion and for many other purposes as well, we prefer our surface to be flat, in a sense which will be made precise later. Here we consider Brakke flows which are contained in a slab, or with bounded height-excess. We show that in this setting mild area ratio bounds can imply almost flatness. We follow [B, 6.6 and 6.9], which is also covered in [KT, chapter 6]. First we inspect the evolution of the area ratio in a cylinder, in order to find that it is decreasing if it is not already close to the area ratio of a plane, see Proposition 7.5. Assuming an upper bound on the area ratio at the beginning and a lower bound later, we obtain for some time in between that the surface almost has area ratio like a plane, which directly implies bounds on mean curvature- and tilt-excess. This leads to Theorem 7.7. The only small difference to [KT] and [B] is the usage of variable test functions, which approximate the characteristic function of the cylinder arbitrarily well. Using these we can state Theorem 7.7 for cylinders directly, while the analogous statements in [KT] and [B] use fixed cylindrical cut-off functions instead.

7.1 Definition. Consider $\zeta \in C^\infty([0,\infty),[0,1])$ from Definition 4.1. For $R \in (0,\infty)$ and $p \in [1,\infty)$ we define

$$\zeta_{R,p}(x) := \zeta\left(\left(R^{-1}|\hat{x}|\right)^p\right)$$

for all $x = (\hat{x},\tilde{x}) \in \mathbb{R}^n \times \mathbb{R}^k$. Note that $\zeta_{R,p}$ is defined on $\mathbb{R}^{n+k}$ although it only depends on the $\mathbb{R}^n$-components. Moreover set

$$\varpi_p := \int_{B_1^n(0)\times\{0\}^k} \zeta_{1,p}^2 d\mathscr{H}^n.$$

For a rectifiable n-varifold μ in $U \subset \mathbb{R}^{n+k}$ we are interested in the difference

$$E = E(\mu,R,p) := R^{-n}\int_U \zeta_{R,p}\, d\mu_t - \varpi_p.$$

7.2 Lemma. *There exists a $C \in (1,\infty)$ such that for all $R \in (0,\infty)$ and $p \in [1,\infty)$ the following holds:*

1. *for all $x \in \mathbb{R}^{n+k}$*

$$\zeta_{R,p}(x) = \begin{cases} 1 & \text{for } 0 \le |\hat{x}| \le (1-p^{-1}2^{-n-8})R \\ 0 & \text{for } R \le |\hat{x}|\,. \end{cases}$$

2. $\max\left\{Rp^{-1}\sup|D\zeta_{R,p}|, R^2p^{-2}\sup|D^2\zeta_{R,p}|\right\} \leq C.$

3. $R^{-n}\int_{B^n_R(0)\times\{0\}^k}\zeta^2_{R,p}d\mathscr{H}^n = \varpi_p.$

4. $(1-p^{-1}2^{-8})\omega_n \leq \varpi_p \leq \omega_n.$

Proof. 1. Let $|\hat{x}| \leq (1-p^{-1}2^{-n-8})R$. By the binomial theorem we have

$$\begin{aligned}\left(R^{-1}|\hat{x}|\right)^p &\leq (1-p^{-1}2^{-n-8})^p\\ &= 1-2^{-n-8}+\binom{p}{2}(p^{-1}2^{-n-8})^2+\sum_{q=3}^{p}\binom{p}{q}(-p^{-1}2^{-n-8})^q.\end{aligned}$$

Now by $\binom{p}{q} \geq \binom{p}{q+1}p^{-1}$ we see that the last sum is negative which yields

$$\left(R^{-1}|\hat{x}|\right)^p \leq 1-2^{-n-8}+\binom{p}{2}p^{-2}(2^{-n-8})^2 \leq 1-2^{-n-9}.$$

By Definition 4.1 we have $\zeta(r)=1$ for $r\in[0,1-2^{-n-9}]$, so we proved statement 1.

2. Calculate for $i,j\in\{1,\dots,n\}$

$$\frac{\partial}{\partial x_i}\zeta_{R,p}(x) = \zeta'\left((R^{-1}|\hat{x}|)^p\right)pR^{-p}|\hat{x}|^{p-2}\hat{x}_i$$

and

$$\begin{aligned}\frac{\partial^2}{\partial x_i\partial x_j}\zeta_{R,p}(x) =&\zeta''\left(R^{-1}|\hat{x}|\right)p^2R^{-2p}|\hat{x}|^{2p-4}\hat{x}_i\hat{x}_j\\ &+\zeta'\left((R^{-1}|\hat{x}|)^p\right)pR^{-p}\left((p-2)|\hat{x}|^{p-4}\hat{x}_i\hat{x}_j+|\hat{x}|^{p-2}\delta_{ij}\right).\end{aligned}$$

Statement 2 then follows by the properties of ζ (see Definition 4.1). In particular we only have to consider the case $\frac{R}{2}\leq|\hat{x}|\leq R$.

3. Property 3 follows from the transformation of variables $\hat{y}=R^{-1}\hat{x}$ inside the integral.

4. To prove property 4 estimate for $r=1-p^{-1}2^{-n-8}$

$$\begin{aligned}r^n\omega_n &= \int_{B^n_r(0)}1d\mathscr{L}^n = \int_{B^n_r(0)}\zeta^2_{1,p}d\mathscr{L}^n \leq \int_{B^n_1(0)}\zeta^2_{1,p}d\mathscr{L}^n\\ &= \varpi_p \leq \int_{B^n_1(0)}1d\mathscr{L}^n = \omega_n,\end{aligned}$$

where we used $r\leq 1$ and Statement 1. Also note that

$$r^n \geq (1-p^{-1}2^{-n-8})^n \geq 1-2^np^{-1}2^{-n-8} = 1-p^{-1}2^{-8}.$$

□

The aim of this section is to find criteria that imply a bound on E. Though $E = 0$ does not mean that we have a plane, a small E indicates a flat shape and allows for Lipschitz approximations. Furthermore a Brakke flow with bounded E for some time, can only have limited curvature integral in that period, which leads to our excess bounds in this section's main Theorem 7.7.

Applying Lemma 3.10 with $\phi_R = \zeta_{R,p}$ yields the following statement:

7.3 Lemma. *There exists a constant $C \in (1, \infty)$ such that for all $R, \gamma \in (0, \infty)$, $p \in [1, \infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Suppose*

$$R^{-n-2} \int_{C_R(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu(x) \le \gamma^2. \tag{7.1}$$

Then

$$R^{-n+2} \mathscr{B}\left(U, \mu, \zeta_{R,p}^2\right) \le -\frac{1}{2} R^{-n+2} \int_U |\vec{H}|^2 \zeta_{R,p}^2 d\mu + C p^4 \gamma^2. \tag{7.2}$$

Note that $\frac{d}{dt} E(\mu_t, R, p) = \frac{d}{dt} R^{-n+2} \mathscr{B}\left(U, \mu, \zeta_{R,p}^2\right)$ so if we can bound the right hand side of (7.2) by E, this will lead to a differential equation for E. The biggest step in this direction is the next Lemma, which is based on the first half of Brakke's proof of the popping soap film lemma [B, 6.6].

7.4 Lemma (Area Ratio Derivative, [B, 6.6])**.** *For every $q \in [1, \infty)$ there exists a $Q \in (1, \infty)$ such that for every $R \in (0, \infty)$, every $\gamma \in (0, Q^{-1}]$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Suppose*

$$\operatorname{spt}\mu \cap C_R(0) \subset\subset U \tag{7.3}$$

$$R^{-n-2} \int_{C_R(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu \le \gamma^2 \tag{7.4}$$

$$\left| R^{-n} \int_U \zeta_{R,q}^2 d\mu - \varpi_q \right| =: |E| \in \left[Q\gamma^2, (1 - (2q)^{-1})\omega_n\right]. \tag{7.5}$$

Then

$$R^{-n+2} \mathscr{B}\left(U, \mu, \zeta_{R,q}^2\right) \le -Q^{-1} \begin{cases} \min\left\{\gamma^{-\frac{2}{3}} |E|^{\frac{4}{3}}, 1\right\} & \text{if } n \le 2, \\ \min\left\{|E|^{\frac{n-2}{n}}, \gamma^{-\frac{2}{3}} |E|^{\frac{4}{3}}, 1\right\} & \text{if } n > 2. \end{cases} \tag{7.6}$$

Proof. We consider the mean curvature-excess

$$\alpha^2 := R^{-n+2} \int_U |\vec{H}|^2 \zeta_{R,q}^2 d\mu. \tag{7.7}$$

We want to show that there exists a $\delta = \delta(q) \in (0,1)$ such that

$$\alpha^2 \geq \delta \begin{cases} \min\left\{\gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}}, 1\right\} & \text{if } n \leq 2, \\ \min\left\{|E|^{\frac{n-2}{n}}, \gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}}, 1\right\} & \text{if } n > 2. \end{cases} \tag{7.8}$$

Showing (7.8) is the main step of the proof. By definition of the right hand side of (7.8) it suffices to consider small α^2 and then show that $|E|$ is bounded from above in terms of α^2. In order to show this we first use Lemma 2.8 to transfer the area ratio bounds from (7.5) to smaller balls with radii $\frac{R}{3}, \frac{R}{9}$. This leads to a Lipschitz approximation in $B_{r_0}(0), r_0 = \frac{R}{9}$, which then lets us get better estimates on the area ratio, in this smaller ball. Using again Lemma 2.8 yields the upper bound on $|E|$ in terms of α^2. Once (7.8) is verified Lemma 7.3 can be used to establish the result.

To prove (7.8) we assume $\alpha^2 \leq \delta$ and lead this assumption to the conclusion

$$\alpha^2 \geq \delta \begin{cases} \gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}} & \text{if } n \leq 2, \\ \min\left\{|E|^{\frac{n-2}{n}}, \gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}}\right\} & \text{if } n > 2, \end{cases}$$

where we will choose δ small depending on q. Also we may assume $\gamma^2 \leq \delta$, which we can always achieve as $\gamma \leq Q^{-1}$ and we can choose $Q \geq \delta^{-\frac{1}{2}}$. Moreover the tilt-excess

$$\beta^2 := R^{-n} \int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \zeta_{R,p}^2 d\mu \tag{7.9}$$

can be estimated by the height- and curvature-excess due to Lemma A.13 with $f = g = h = \zeta_{R,q}$. This yields

$$\beta^2 = R^{-n}\beta_g^2 \leq C_n \left(\alpha\gamma + R^{-n} \int_U |\pi_{\mathbb{R}^k}(x)|^2 |\nabla^\mu \zeta_{R,q}| d\mu(x)\right),$$

where we used $\mathrm{spt}\zeta_{R,q} \subset C_R(0)$ to estimate $\alpha_f^2 \leq R^{n-2}\alpha^2$ and $\gamma_h^2 \leq R^{n+2}\gamma^2$. Using $|D\zeta_{R,q}| \leq q^2 R^{-2}\sigma_1$ and (7.4) we obtain

$$\beta^2 \leq C_n \left(\alpha\gamma + q^2\gamma^2\right) \leq C_n q^2 \delta \tag{7.10}$$

where we used our bounds on α and γ for the second estimate. Thus all the three excesses are bounded by $\alpha^2+\beta^2+\gamma^2$ for which

$$\alpha^2+\beta^2+\gamma^2 \leq C_n q^2 \delta \tag{7.11}$$

holds.

We want to approximate μ in $B_{r_0}(0)$ by a Lipschitz function, for some $r_0 \in (0,R)$. In order to use Theorem 2.9 we have to show

$$\mu\left(B_{r_0}(0)\right) \geq \lambda \omega_n r_0^n \tag{7.12}$$
$$\mu\left(B_{3r_0}(0)\right) \leq (2-\lambda)\omega_n (3r_0)^n \tag{7.13}$$

for some $\lambda \in (0,1)$.

Define

$$r_0 := \frac{R}{9}, \quad r_1 := (1-q^{-1}2^{-n-8})r_0, \quad r_2 := 3(1+q^{-1}2^{-n-4})r_0.$$

By Lemma 2.8.2 with $R_2 = R$, $R_1 = r_i$, $\Phi_R = \zeta_{R,q}$, $\alpha_0 = \alpha$, $\beta_0 = \beta$ we obtain

$$\begin{aligned}&\left| R^{-n}\int_U \zeta_{R,q}^2 d\mu - r_i^{-n}\int_U \zeta_{r_i,q}^2 d\mu \right| \\ &\leq r_i^{-n}R^n\left(\left(n\log\left(\frac{R}{r_i}\right)+2\right)\beta^2 + R^{-1}(R-r_i)\alpha\beta\right)\end{aligned}$$

for $i \in \{1,2\}$. Here we had to use that (7.3) equals (2.5). Then we can estimate

$$\left| R^{-n}\int_U \zeta_{R,q}^2 d\mu - r_i^{-n}\int_U \zeta_{r_i,q}^2 d\mu \right| \leq C_n \left(\frac{R}{r_i}\right)^n \left(\beta^2+\alpha\beta\right) \tag{7.14}$$

for $i \in \{1,2\}$, where we used $\log(Rr_i^{-1}) \leq \log(18) \leq 3$. As the height-excess is small due to (7.4), we can estimate the measure of a ball from below by the measure of a cylinder with slightly smaller radius. In particular for $B_{r_0}(0)$ and $C_{r_1}(0)$ we obtain for the set $A := \{x \in C_{r_1}(0) : |\pi_{\mathbb{R}^k}(x)|^2 > r_0^2 - r_1^2\}$ that

$$\mu\left(C_{r_1}(0) \setminus B_{r_0}(0)\right) \leq \mu\left(A\right). \tag{7.15}$$

Here we used that for $x \in C_{r_1}(0)$ with $|\pi_{\mathbb{R}^k}(x)|^2 \leq r_0^2 - r_1^2$ we can estimate $|x|^2 = |\pi_{\mathbb{R}^k}(x)|^2 + |\hat{x}| \leq r_0^2$ so $x \notin C_{r_1}(0) \setminus B_{r_0}(0)$. Using (7.4) the measure of A implies a lower bound on γ by

$$R^{n+2}\gamma^2 \geq (r_0^2 - r_1^2)\mu\left(A\right),$$

so we can estimate with (7.15)

$$\mu\left(C_{r_1}(0)\setminus B_{r_0}(0)\right) \leq (r_0^2-r_1^2)^{-1}\gamma^2 R^{n+2} \leq C_n q\gamma^2 R^n. \tag{7.16}$$

Here we estimated $r_0^2 - r_1^2 \geq (1-(1-q^{-1}2^{-n-8})^2)r_0^2 \geq (C_n q)^{-1} r_0^2$ and we also used $R^2 = 81 r_0^2$ by definition of r_0. Now we can use estimate (7.14) and (7.16) to bound $\mu\left(B_{r_0}(0)\right)$ from below

$$\begin{aligned}\mu\left(B_{r_0}(0)\right) &\geq \int_U \zeta_{r_1,q}^2 d\mu - \mu\left(C_{r_1}(0)\setminus B_{r_0}(0)\right)\\ &\geq r_1^n R^{-n}\int_U \zeta_{R,q}^2 d\mu - C_n R^n\left(\beta^2+\alpha\beta\right) - C_n q\gamma^2 R^n.\end{aligned}$$

Here the first step holds because $\mathrm{spt}\zeta_{r_1,q}\subset C_{r_1}(0)$. The bounds (7.5) and (7.11) yield $R^{-n}\int_U \zeta_{R,q}^2 d\mu \geq \varpi_q - (1-(2q)^{-1})\omega_n$ and $\beta^2+\alpha\beta+\gamma^2 \leq C_n q^2\delta$, so we can estimate

$$\mu\left(B_{r_0}(0)\right) \geq r_1^n\left(\varpi_q - (1-(2q)^{-1})\omega_n\right) - C_n R^n q^3\omega_n\delta.$$

Using $r_1 := (1-q^{-1}2^{-n-8})r_0$ and $R := 9r_0$ we obtain

$$\begin{aligned}\mu\left(B_{r_0}(0)\right) &\geq r_0^n\left[(1-q^{-1}2^{-n-8})^n\left(\varpi_q-\omega_n+(2q)^{-1}\omega_n\right) - C_n q^3\omega_n\delta\right]\\ &\geq \left[(1-q^{-1}2^{-8})(-q^{-1}2^{-8}+(2q)^{-1}) - C_n q^3\delta\right]\omega_n r_0^n\\ &\geq \left[2^{-3}q^{-1} - C_n q^3\delta\right]\omega_n r_0^n \geq 2^{-4}q^{-1}\omega_n r_0^n,\end{aligned}$$

where we had to choose δ small enough depending on q. Here we also used $\varpi_q \geq (1-q^{-1}2^{-8})\omega_n$. So we verified (7.12) for $\lambda \leq 2^{-4}q^{-1}$. For the upper bound we can analogously estimate with (7.14)

$$\mu\left(B_{3r_0}(0)\right) \leq \int_U \zeta_{r_2,q}^2 d\mu \leq r_2^n R^{-n}\int_U \zeta_{R,q}^2 d\mu + C_n R^n\left(\beta^2+\alpha\beta\right).$$

Here the first step holds because $(1-q^{-1}2^{-n-8})r_2 \geq 3r_0$, so $\mathrm{spt}\zeta_{r_2,q}^2 \supset C_{3r_0}$. Then the bounds (7.5) and (7.11) yield $R^{-n}\int_U \zeta_{R,q}^2 d\mu \leq \varpi_q + (1-(2q)^{-1})\omega_n$ and $\beta^2+\alpha\beta+\gamma^2 \leq C_n q^2\delta$, so we can estimate

$$\mu\left(B_{3r_0}(0)\right) \leq r_2^n\left(\varpi_q + (1-(2q)^{-1})\omega_n\right) + C_n R^n q^2\omega_n\delta.$$

Using $r_2 := 3(1+q^{-1}2^{-n-4})r_0$ and $R := 9r_0$ we obtain

$$\begin{aligned}\mu\left(B_{3r_0}(0)\right) &\leq (3r_0)^n\left[(1+q^{-1}2^{-n-4})^n(\varpi_q+\omega_n-(2q)^{-1}\omega_n) + C_n q^2\delta\omega_n\right]\\ &\leq (3r_0)^n\omega_n\left[(1+q^{-1}2^{-4})(2-(2q)^{-1}) + C_n q^2\delta\right]\\ &\leq (2-2^{-2}q^{-1}+C_n q^2\delta)\omega_n(3r_0)^n \leq (2-2^{-3}q^{-1})\omega_n(3r_0)^n,\end{aligned}$$

where we had to choose δ small enough depending on q. Here we also needed $\varpi_q \leq \omega_n$. So (7.12) and (7.13) hold for $\lambda = 2^{-4}p^{-1}$. To apply Theorem 2.9 with α, β and γ we see that in view of $\mathrm{spt}\mu \cap B_{7r_0}(0) \subset U$ equations (7.4), (7.7) (7.9) imply (2.18), (2.19) and (2.20). For Q large enough also $\gamma \leq \gamma_0$ is verified, such that we can use Theorem 2.9 in the ball $B_{r_0}(0)$. This yields the existence of a Lipschitz function $f : B^n_{r_0}(0) \to \mathbb{R}^k$ with $\mathrm{lip}(f) \leq 1$ and

$$\mu\left(B^n_{r_0}(0) \times B^k_{r_0}(0) \setminus X\right) + \mathscr{L}^n\left(B^n_{r_0}(0) \setminus Y\right) \leq C_n r_0^n \left(\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3} + \beta^2 + \gamma^2\right),$$

where $X = \left\{x \in C_{r_0}(0) : \exists \hat{y} \in B^n_{r_0}(0), (\hat{y}, f(\hat{y})) = x, \Theta\left(\mu, x\right) = 1\right\}$ and $Y = \pi_{\mathbb{R}^n}\left(X\right)$. Now we can approximate each integral on the varifold by the integral on the graph of f. In particular for $\zeta^2_{r_1,q}$ we can estimate

$$\begin{aligned}&\left|\int_{B_{r_0}(0)} \zeta^2_{r_1,q} d\mu - \int_{B^n_{r_0}(0)} \zeta^2_{r_1,q}(\hat{y}, 0) d\mathscr{L}^n(\hat{y})\right| \\ &\leq C_n \left(\frac{R}{9}\right)^n \left(\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3} + \beta^2 + \gamma^2\right),\end{aligned}$$

see Remark 2.11. By definition of ϖ_q we have $r^n \varpi_q = \int \zeta^2_{r,q} d\mathscr{L}^n$ for all $r \in (0, R)$, so

$$\left|r_1^{-n} \int_{B_{r_0}(0)} \zeta^2_{r_1,q} d\mu - \varpi_q\right| \leq C_n \left(\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3} + \beta^2 + \gamma^2\right), \tag{7.17}$$

where we used $r_1 \leq r_0$. We want such an estimate for U instead of $B_{r_0}(0)$. Use $B_{r_0}(0) \subset \{\zeta_{r_1,q} = 1\} \subset \mathrm{spt}\zeta_{r_1,q} \subset C_{r_1}(0)$ to see

$$\int_{B_{r_0}(0)} \zeta^2_{r_1,q} d\mu \leq \int_U \zeta^2_{r_1,q} d\mu \leq \int_{B_{r_0}(0)} \zeta^2_{r_1,q} d\mu + \mu\left(C_{r_1}(0) \setminus B_{r_0}(0)\right).$$

Then with (7.17), (7.16) and $r_1 := (1 - q^{-1}2^{-n-8})9^{-1}R$ we obtain

$$\left|r_1^{-n} \int_U \zeta^2_{r_1,q} d\mu - \varpi_q\right| \leq C_n q \left(\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3} + \beta^2 + \gamma^2\right). \tag{7.18}$$

We want (7.18) for R instead of r_1. By (7.14) with r_1 we can estimate

$$\left|R^{-n} \int_U \zeta^2_{R,q} d\mu - r_1^{-n} \int_U \zeta^2_{r_1,q} d\mu\right| \leq 4n r_1^{-n} R^n \left(\alpha\beta + \beta^2\right) \leq C_n \left(\alpha\beta + \beta^2\right),$$

where we used $r_1 := (1 - q^{-1}2^{-n-8})9^{-1}R$. Combining this with (7.18) yields

$$|E| = \left|R^{-n} \int_U \zeta^2_{R,q} d\mu - \varpi_q\right| \leq C_n q \left(\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3} + \alpha\beta + \beta^2 + \gamma^2\right).$$

By the first inequality of (7.10) we have $\beta \leq C_n q(\sqrt{\alpha\gamma} + \gamma)$, so we can estimate $\alpha\beta + \beta^2 \leq C_n q^2(\alpha\gamma + \gamma^2 + \alpha^{\frac{3}{2}}\gamma^{\frac{1}{2}})$ and arrive at

$$|E| \leq C_n q^3 \max\left\{\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3}, \alpha^{\frac{3}{2}}\gamma^{\frac{1}{2}}, \gamma^2\right\}.$$

Assumption (7.5) yields the lower bound $|E| \geq Q\gamma^2$. So for Q large enough depending on q we can rule out that the maximum is attained for γ^2. Thus we have

$$|E| \leq C_n q^3 \max\left\{\alpha^{\frac{2n}{n-2}}\delta_{n\geq 3}, \alpha^{\frac{3}{2}}\gamma^{\frac{1}{2}}\right\}$$

and this inequality establishes the desired lower curvature bound (7.8) for δ small depending on q, n, k.

Now we can use (7.8) to obtain a bound on the Brakke variation of $\zeta_{R,p}^2$. Use that by Lemma 7.3

$$R^{-n+2}\mathscr{B}\left(U, \mu, \zeta_{R,q}^2\right) \leq -\frac{\alpha^2}{2} + C_0 q^4\gamma^2 \tag{7.19}$$

for a constant $C_0 \in (1, \infty)$. We want to show that the last term in (7.19) is smaller than $\frac{\alpha^2}{4}$ to bound the Brakke variation by $-\frac{\alpha^2}{4}$ from above.

By (7.5) we can estimate $|E| \geq Q\gamma^2$, so (7.8) yields

$$\alpha^2 \geq \delta \begin{cases} \min\left\{Q\gamma^2, 1\right\} & \text{if } n \leq 2, \\ \min\left\{(Q\gamma^2)^{\frac{n-2}{n}}, Q\gamma^2, 1\right\} & \text{if } n > 2. \end{cases}$$

By assumption we know $Q\gamma^2 \leq Q^{-1} \leq 1$, so the 1 cannot be the minimum. Also for $n \geq 3$ we have $\frac{n-2}{n} \leq 1$, thus $(Q\gamma^2)^{\frac{n-2}{n}} \geq Q\gamma^2$, so for Q large enough

$$\alpha^2 \geq \delta Q\gamma^2 \geq 4C_0 q^4\gamma^2,$$

where C_0 is the constant from (7.19). Thus (7.19) combined with (7.8) implies

$$R^{-n+2}\mathscr{B}\left(U, \mu, \zeta_{R,q}^2\right) \leq -\frac{\alpha^2}{4} \leq -\frac{\delta}{4} \begin{cases} \min\left\{\gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}}, 1\right\} & \text{if } n \leq 2, \\ \min\left\{|E|^{\frac{n-2}{n}}, \gamma^{-\frac{2}{3}}|E|^{\frac{4}{3}}, 1\right\} & \text{if } n > 2, \end{cases}$$

which establishes the result, if $Q \geq 4\delta^{-1}$. Note that δ only depends on q, n, k. $\square$

Now we can use this to derive a differential inequality for E. Solving this inequality establishes bounds on $\mu_t(\zeta_{R,p}^2)$ in a certain time interval. This is done in the next proposition, which is a reformulation of Brakke's popping soap film lemma [B, 6.6].

7.5 Proposition (Decreasing Area Ratio, [B, 6.6]). *For every $p \in [1,\infty)$ there exist $P \in (1,\infty)$ and $\gamma_0 \in (0,1)$ such that for all $R \in (0,\infty)$, $\gamma \in (0,\gamma_0]$, all $t_1, t_2 \in \mathbb{R}$ with $t_2 - t_1 > PR^2$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U with*

$$\bigcup_{t\in[t_1,t_2]} \mathrm{spt}\mu_t \cap C_R(0) \subset\subset U \tag{7.20}$$

$$\sup_{t\in[t_1,t_2]} R^{-n-2} \int_{U\cap C_R(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \leq \gamma^2. \tag{7.21}$$

Then

$$R^{-n}\int_U \zeta_{R,p}^2 d\mu_{t_1} \leq (2-p^{-1})\omega_n \Longrightarrow R^{-n}\int_U \zeta_{R,p}^2 d\mu_{s_1} \leq \varpi_p + P\gamma^2 \tag{7.22}$$

$$R^{-n}\int_U \zeta_{R,p}^2 d\mu_{t_2} \geq p^{-1}\omega_n \Longrightarrow R^{-n}\int_U \zeta_{R,p}^2 d\mu_{s_2} \geq \varpi_p - P\gamma^2 \tag{7.23}$$

holds for all $s_1 \in [t_1 + PR^2, t_2]$ and all $s_2 \in [t_1, t_2 - PR^2]$.

Proof. For $t \in [t_1, t_2]$ set

$$E(t) := R^{-n}\int_U \zeta_{R,p}^2 d\mu_t - \varpi_p. \tag{7.24}$$

By Lemma 7.4 applied with $q = p$ there exists a Q depending on p such that

$$R^{-n+2}\mathscr{B}\left(U,\mu_t,\zeta_{R,p}^2\right) \leq -Q^{-1}\begin{cases}\min\left\{\gamma^{-\frac{2}{3}}|E(t)|^{\frac{4}{3}},1\right\} & \text{if } n\leq 2,\\ \min\left\{|E(t)|^{\frac{n-2}{n}},\gamma^{-\frac{2}{3}}|E(t)|^{\frac{4}{3}},1\right\} & \text{if } n>2,\end{cases}$$

for all $t \in [t_1,t_2]$ with $Q\gamma^2 \leq |E(t)| \leq (1-(2p)^{-1})\omega_n$. Note that we need to choose $\gamma_0 \leq Q^{-1}$ in order to apply Lemma 7.4. With the Brakke flow inequality (3.1) we then have for all $n \geq 1$

$$\begin{aligned}\overline{D}E(t) &= R^{-n}\overline{D}\left(\int_U \zeta_{R,p}^2 d\mu_t\right)\\ &\leq -Q^{-1}R^{-2}\min\left\{|E(t)|^{\frac{n-1}{n}},\gamma^{-\frac{2}{3}}|E(t)|^{\frac{4}{3}},1\right\}\end{aligned} \tag{7.25}$$

for all $t \in [t_1,t_2]$ with $Q\gamma^2 \leq |E(t)| \leq (1-(2p)^{-1})\omega_n$. Here we used that $|E(t)|^{\frac{n-1}{n}} \leq |E(t)|^{\frac{n-2}{n}}$ for $|E(t)| \leq 1$ and when $|E(t)| \geq 1$ the expression $|E(t)|$ cannot be the minimum. For the $n \leq 2$ case we used that taking a minimum over a larger set, only makes it smaller. Note that assumption

(7.20) is necessary in order to apply (3.1). Inequality (7.25) implies that $E(t)$ is monotonically decreasing for certain t. Consider the set

$$V := [-(1-(2p)^{-1})\omega_n, -Q\gamma^2] \cup [Q\gamma^2, (1-(2p)^{-1})\omega_n].$$

If $E(t)$ is once smaller than $v \in V$ it will stay below this value, i.e.

$$\exists d_1 \in [t_1, t_2] : \ E(d_1) \leq v \Longrightarrow E(t) \leq v \ \ \forall t \in [d_1, t_2], \tag{7.26}$$

$$\exists d_2 \in [t_1, t_2] : \ E(d_2) > v \Longrightarrow E(t) > v \ \ \forall t \in [t_1, d_2] \tag{7.27}$$

for all $v \in V$. To see (7.26) let $v \in V$, $d_1 \in [t_1, t_2]$ and consider the set $J := \{d_0 \in [d_1, t_2] : E(t) \leq v \ \forall t \in [d_1, d_0]\}$. By Proposition 3.7.2 $E(t)$ cannot "jump up" at any time t. Hence J is closed. Consider a $d_0 \in J$. If $E(d_0) < v$ we can use again Proposition 3.7.2 to find a $\delta \in (0,1)$ such that $[d_0, d_0 + \delta] \cap [d_1, t_2] \subset J$. If $E(d_0) = v$ use (7.25) to find such a δ. Thus J is closed and open inside $[d_1, t_2]$. As $[d_1, t_2]$ is connected this proves (7.26). Then (7.26) implies (7.27) via a contra position argument.

For an interval $I = [a, b] \subset [t_1, t_2]$ consider the properties:

$$Q\gamma^2 \leq E(t) \leq (1-(2p)^{-1})\omega_n \quad \forall t \in [a, b], \tag{7.28}$$

$$-Q\gamma^2 \geq E(t) \geq -(1-(2p)^{-1})\omega_n \quad \forall t \in [a, b]. \tag{7.29}$$

The proof is based on the following observation: There exists a $P \in (1, \infty)$ such that for every $I = [a, b] \subset [t_1, t_2]$ for which either (7.28) or (7.29) holds, we can estimate

$$b - a \leq PR^2. \tag{7.30}$$

To prove this let $I = [a, b] \subset [t_1, t_2]$ be such that either (7.28) or (7.29) holds. Then inequality (7.25) holds for all $t \in I$. In particular E is monotonically decreasing and does not change sign on I. Let $I_1, I_2, I_3 \subset I$ be the parts, where the min of (7.25) is 1, $|E(t)|^{\frac{n-1}{n}}$ or $\gamma^{-\frac{2}{3}}|E(t)|^{\frac{4}{3}}$ respectively.

$$\begin{aligned} I_1 &:= \left\{ s \in I \ : 1 \leq |E(s)| \leq (1-(2p)^{-1})\omega_n \right\} \\ I_2 &:= \left\{ s \in I \ : \gamma^{\frac{2n}{3n+1}} \leq |E(s)| \leq 1 \right\} \\ I_3 &:= \left\{ s \in I \ : Q\gamma^2 \leq |E(s)| \leq \gamma^{\frac{2n}{3n+1}} \right\}. \end{aligned}$$

By monotonicity of E the sets I_1, I_2, I_3 are each intervals themselves with $I_1 \cup I_2 \cup I_3 = I$. We can solve the ODE inequality on each of this intervals separately which will give an upper bound for $b - a$. For $I_1 = [a_1, b_1]$ estimate by (7.25) and with Proposition A.19

$$E(b_1) - E(a_1) \leq \int_{a_1}^{b_1} \overline{D}E(t)dt \leq -Q^{-1}(b_1 - a_1)R^{-2}.$$

Either (7.31) or (7.32) imply a bound of $|E(b_1) - E(a_1)|$ by $(1-(2p)^{-1})\omega_n$, so we obtain

$$b_1 - a_1 \leq \omega_n Q R^2. \tag{7.31}$$

For $I_2 = [a_2, b_2]$ we will distinguish the cases where I satisfies (7.28) hence $E > 0$ and where I satisfies (7.29) hence $E < 0$. By (7.25) with Propositions A.19 and A.20 estimate for $E > 0$

$$\begin{aligned}
&E(b_2)^{\frac{1}{n}} - E(a_2)^{\frac{1}{n}} \leq \int_{a_2}^{b_2} \overline{D}\left(E(t)^{\frac{1}{n}}\right) dt \\
&\leq n^{-1} \int_{a_2}^{b_2} E(t)^{\frac{-n+1}{n}} \overline{D} E(t) dt \leq -(nQ)^{-1}(b_2 - a_2) R^{-2}.
\end{aligned}$$

Here we used that for $f(r) = r^{\frac{1}{n}}$, $r > 0$ the derivative satisfies $Df(r) > 0$. Analogously we can estimate for $E < 0$

$$\begin{aligned}
&(-E(b_2))^{\frac{1}{n}} - (-E(a_2))^{\frac{1}{n}} \geq \int_{a_2}^{b_2} \underline{D}\left((-E(t))^{\frac{1}{n}}\right) dt \\
&\geq -n^{-1} \int_{a_2}^{b_2} (-E(t))^{\frac{-n+1}{n}} \overline{D} E(t) dt \geq (nQ)^{-1}(b_2 - a_2) R^{-2}.
\end{aligned}$$

Here we used that for $f(r) = (-r)^{\frac{1}{n}}$, $r < 0$ the derivative satisfies $Df(r) < 0$. In both cases we can estimate

$$b_2 - a_2 \leq nQR^2 \left| |E(b_2)|^{\frac{1}{n}} - |E(a_2)|^{\frac{1}{n}} \right| \leq \sqrt[n]{\omega_n} nQR^2, \tag{7.32}$$

where we used $|E(t)| \leq (1-(2p)^{-1})\omega_n$ for all $t \in I$.

For $I_3 = [a_3, b_3]$ we will again distinguish the cases where I satisfies (7.28) hence $E > 0$ and where I satisfies (7.29) hence $E < 0$. By (7.25) with Propositions A.19 and A.20 estimate for $E > 0$

$$\begin{aligned}
&E(b_3)^{-\frac{1}{3}} - E(a_3)^{-\frac{1}{3}} \geq \int_{a_3}^{b_3} \underline{D}\left(E(t)^{-\frac{1}{3}}\right) dt \\
&\geq -\frac{1}{3} \int_{a_3}^{b_3} E(t)^{-\frac{4}{3}} \overline{D} E(t) dt \geq (3Q)^{-1}(b_3 - a_3)\gamma^{-\frac{2}{3}} R^{-2}.
\end{aligned}$$

Here we used that for $f(r) = r^{-\frac{1}{3}}$, $r > 0$ the derivative satisfies $Df(r) < 0$. Analogously we can estimate for $E < 0$

$$\begin{aligned}
&(-E(b_3))^{-\frac{1}{3}} - (-E(a_3))^{-\frac{1}{3}} \leq \int_{a_3}^{b_3} \overline{D}\left(E(t)^{-\frac{1}{3}}\right) dt \\
&\leq \frac{1}{3} \int_{a_3}^{b_3} E(t)^{-\frac{4}{3}} \overline{D} E(t) dt \leq -(3Q)^{-1}(b_3 - a_3)\gamma^{-\frac{2}{3}} R^{-2}.
\end{aligned}$$

Here we used that for $f(r) = (-r)^{-\frac{1}{3}}$, $r < 0$ the derivative satisfies $Df(r) > 0$. In both cases we can estimate

$$b_3 - a_3 \leq 3QR^2\gamma^{\frac{2}{3}} \left| |E(b_3)|^{-\frac{1}{3}} - |E(a_3)|^{-\frac{1}{3}} \right| \leq 3nQ^{\frac{2}{3}}R^2, \tag{7.33}$$

where we used $|E(t)| \geq Q\gamma^2$ for all $t \in I$. Combining (7.31), (7.32) and (7.33) we see $b - a \leq C_n QR^2$, thus we established (7.30) for some P depending on Q which depended on p. Note that for larger P estimate (7.30) remains true.

Now we can use (7.30) to verify the statements (7.22) and (7.23). Suppose $R^{-n} \int_U \zeta_{R,p}^2 d\mu_{t_1} \leq (2-p^{-1})\omega_n$, then with $\varpi_p \geq (1-p^{-1}2^{-8})\omega_n$ we can estimate

$$E(t_1) = R^{-n} \int_U \zeta_{R,p}^2 d\mu_{t_1} - \varpi_p \leq (2 - p^{-1} - 1 + p^{-1}2^{-8})\omega_n < (1 - (2p)^{-1})\omega_n.$$

By (7.26) this yields $E(t) \leq (1 - (2p)^{-1})\omega_n$ for all $t \in [t_1, t_2]$. Thus by (7.30) in view of (7.28) there has to exist $d_1 \in [t_1, t_1 + PR^2]$ such that $E(d_1) < Q\gamma^2$, so by (7.26) we obtain $E(t) \leq Q\gamma^2$ for all $[d_1, t_2]$. As $d_1 \leq t_1 + PR^2$ and by definition of E this verifies (7.22) for $P \geq Q$. Suppose $R^{-n} \int_U \zeta_{R,p}^2 d\mu_{t_2} \geq p^{-1}\omega_n$, then with $\varpi_p \leq \omega_n$ we can estimate

$$E(t_2) = R^{-n} \int_U \zeta_{R,p}^2 d\mu_{t_2} - \varpi_p \geq p^{-1}\omega_n - \omega_n > -(1 - (2p)^{-1})\omega_n.$$

By (7.27) this yields $E(t) \geq -(1 - (2p)^{-1})\omega_n$ for all $t \in [t_1, t_2]$. Thus by (7.30) in view of (7.29) there has to exist $d_2 \in [t_2 - PR^2, t_2]$ such that $E(d_2) > -Q\gamma^2$, so by (7.27) we obtain $E(t) \geq -Q\gamma^2$ for all $[t_1, d_2]$. As $d_2 \geq t_2 - PR^2$ and by definition of E this verifies (7.23) for $P \geq Q$. □

In the slab setting this can be used to obtain a time interval, where the flow is almost flat. This is the way in which Brakke uses the popping soap film lemma in the proof of [B, 6.9].

7.6 Lemma. *For every $q \in [1, \infty)$ there exists a $Q \in (1, \infty)$ such that, for all $K \in (1, \infty)$ there exists a $\eta_0 \in (0, 1)$ such that for all $\rho \in (0, \infty)$, $\eta \in (0, \eta_0]$, $s_1, s_2 \in \mathbb{R}$ with $s_2 - s_1 > 2QR^2$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t \in [s_1, s_2]}$ be a Brakke flow in U with*

$$\operatorname{spt}\mu_t \cap C_\rho(0) \subset \{x \in C_\rho(0), |\pi_{\mathbb{R}^k}(x)| \leq \eta\rho\} \subset\subset U \tag{7.34}$$

$$\rho^{-n}\mu_t\,(C_\rho(0)) \leq K \tag{7.35}$$

for all $t \in [s_1, s_2]$. Suppose

$$\rho^{-n} \int_U \zeta_{\rho,q}^2 d\mu_{s_1} \leq (2 - q^{-1})\omega_n\ , \quad \rho^{-n} \int_U \zeta_{\rho,q}^2 d\mu_{s_2} \geq q^{-1}\omega_n \tag{7.36}$$

Then for all $s \in [s_1 + QR^2, s_2 - QR^2]$

$$\left|\rho^{-n}\int_U \zeta_{\rho,q}^2 d\mu_s - \varpi_p\right| \leq QK\eta^2, \tag{7.37}$$

$$\rho^{-n}\int_{s-\rho^2}^{s+\rho^2}\int_U |\vec{H}|^2\zeta_{\rho,q}^2 d\mu_t\, dt \leq QK\eta^2, \tag{7.38}$$

$$\rho^{-n-2}\int_{s-\rho^2}^{s+\rho^2}\int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2\, \zeta_{\rho,q}^2 d\mu_t\, dt \leq QK\eta^2. \tag{7.39}$$

Proof. For $t \in [s_1, s_2]$ set

$$\alpha_t^2 := \rho^{-n+2}\int_U |\vec{H}|^2\zeta_{\rho,q}^2 d\mu_t \tag{7.40}$$

$$\beta_t^2 := \rho^{-n}\int_U |\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}|^2\, \zeta_{\rho,q}^2 d\mu_t, \tag{7.41}$$

where α is only defined for almost every t. We want to use Proposition 7.5. The slab condition (7.34) and the cylindrical area ratio bound (7.35) directly yield a bound for the height-excess, namely

$$\rho^{-n-2}\int_{U\cap C_\rho(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t(x) \leq K\eta^2 \tag{7.42}$$

for all $t \in [s_1, s_2]$. Proposition 7.5 with $p = q$, $\gamma^2 = K\eta^2$, $t_1 = s_1$, $t_2 = s_2$ and $R = \rho$ yields P and γ_0 depending on q such that

$$\left|\rho^{-n}\int_U \zeta_{\rho,q}^2 d\mu_s - \varpi_p\right| \leq PK\eta^2 \tag{7.43}$$

for all $s \in [s_1 + P\rho^2, s_2 - P\rho^2]$. Here we used that (7.36) implies the validity of the assumptions in the statements (7.22) and (7.23). Hence the conclusions in (7.22) and (7.23) imply (7.43). Note that in order to use Proposition 7.5 we choose $Q \geq P = P(q)$ and $K\eta^2 \leq K\eta_0^2 \leq \gamma_0^2$ for η_0 small depending on K and $\gamma_0 = \gamma_0(q)$. Inequality (7.43) then directly implies (7.37).

To prove (7.38) we first note that by (7.43) we have

$$-2PK\eta^2 \leq \left(\rho^{-n}\int_U \zeta_{\rho,q}^2 d\mu_{s+\rho^2} - \rho^{-n}\int_U \zeta_{\rho,q}^2 d\mu_{s-\rho^2}\right).$$

Then by (3.1) and Proposition A.19 we obtain

$$-2PK\eta^2 \leq \rho^{-n}\int_{s-\rho^2}^{s+\rho^2} \mathscr{B}(U, \mu_t, \zeta_{\rho,q}^2)dt \tag{7.44}$$

for all $s \in [s_1 + P\rho^2 + \rho^2, s_2 - \rho^2 - P\rho^2]$. Now by Lemma 7.3 with $R = \rho$, $\gamma^2 = K\eta^2$ and $p = q$ combined with estimate (7.42) and in view of definition (7.40) we have

$$\rho^{-n+2}\mathscr{B}(U, \mu_t, \zeta_{\rho,q}^2) \leq -\frac{\alpha_t^2}{2} + C_n q^4 K\eta^2$$

for almost every $t \in [s_1, s_2]$. Combining this with (7.44) yields

$$-2PK\eta^2 \leq -\frac{1}{2}\rho^{-2}\int_{s-\rho^2}^{s+\rho^2} \alpha_t^2 dt + C_n q^4 M\eta^2$$

for all $s \in [s_1 + (P+1)\rho^2, s_2 - (P+1)\rho^2]$. Thus we obtain

$$\rho^{-2}\int_{s-\rho^2}^{s+\rho^2} \alpha_t^2 dt \leq C_n(P + q^4)K\eta^2 \tag{7.45}$$

for all $s \in [s_1 + (P+1)\rho^2, s_2 - (P+1)\rho^2]$. In view of (7.40) this implies (7.38) for Q large enough depending on q. Note that P only depends on q and constants.

Now as usual, bounds on mean curvature-excess and on height imply a bound on tilt-excess. Using Lemma A.13 with $f = g = h = \zeta_{\rho,q}$ yields

$$\beta_t^2 = \rho^{-n}\beta_g^2 \leq C_n\left(\alpha_t\sqrt{K}\eta + \rho^{-n}\int_U |\pi_{\mathbb{R}^k}(x)|^2|\nabla^{\mu_t}\zeta_{\rho,q}|^2 d\mu_t(x)\right)$$

for almost every $t \in [s_1, s_2]$, where we used (7.42) to estimate $\gamma_h \leq \rho^{n+2}K\eta^2$ and we used $\alpha_f^2 = \rho^{n-2}\alpha_t^2$. Then again by (7.42) and using $|D\zeta_{\rho,q}| \leq q\rho^{-1}\sigma_1$ as well as Young's inequality we obtain

$$\beta_t^2 \leq C_n\left(\alpha_t^2 + Kq^2\eta^2\right)$$

for almost every $t \in [s_1, s_2]$. Thus integrating over time we can use (7.45) to obtain

$$\rho^{-2}\int_{s-\rho^2}^{s+\rho^2} \beta_t^2 dt \leq C_n\left(P + q^4\right)K\eta^2$$

for all $s \in [s_1 + (P+1)\rho^2, s_2 - (P+1)\rho^2]$. In view of (7.41) this verifies (7.39) for Q large enough depending on q. Note again that P only depends on q and constants. □

Now combine this with the cylindrical growth lemma 2.8, to obtain nice density ratio estimates for smaller radii as well. This is the form in which Brakke's popping soap film lemma enters the calculations in [B, 6.9], although it is never formulated as an own statement.

7.7 Proposition (Local Flatness, [B, 6.9], [KT, 5.7])**.** *For every $\lambda \in (0,1]$ there exists a $\Lambda \in (1,\infty)$ such that, for every $\tau \in (0,1]$ and every $M \in [1,\infty)$ there exists a $h_0 \in (0,1)$ such that for all $h \in (0,h_0]$,$R \in (0,\infty)$, $y_0 \in \mathbb{R}^{n+k}$ $t_1, t_2 \in \mathbb{R}$ with $t_2 - t_1 > 2\Lambda R^2$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U with*

$$\mathrm{spt}\mu_t \cap C_{(1+\lambda)R}(y_0) \subset \left\{x \in C_{(1+\lambda)R}(y_0), |\pi_{\mathbb{R}^k}(x-y_0)| \leq hR\right\} \subset\subset U, \quad (7.46)$$

$$R^{-n}\mu_t\left(C_{(1+\lambda)R}(y_0)\right) \leq M. \quad (7.47)$$

for all $t \in [t_1,t_2]$. Suppose

$$R^{-n}\mu_{t_1}\left(C_{(1+\lambda)R}(y_0)\right) \leq (2-\lambda)\omega_n\ , \quad R^{-n}\mu_{t_2}\left(C_R(y_0)\right) \geq \lambda\omega_n. \quad (7.48)$$

Then for all $s \in [t_1+\Lambda R^2, t_2-\Lambda R^2]$ and every $r \in [\tau R, R]$

$$\left|r^{-n}\mu_t\left(C_r(y_0)\right) - \omega_n\right| \leq \lambda\omega_n, \quad (7.49)$$

$$R^{-n}\int_{s-R^2}^{s+R^2}\int_{C_R(y_0)} |\vec{H}|^2 d\mu_t\, dt \leq \Lambda M h^2, \quad (7.50)$$

$$R^{-n-2}\int_{s-R^2}^{s+R^2}\int_{C_R(y_0)} \left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 d\mu_t\, dt \leq \Lambda M h^2. \quad (7.51)$$

Proof. We may assume $y_0 = 0$. For given $\lambda \in (0,1]$ and $R \in (0,\infty)$ set $q := 2^n\lambda^{-1}$ and $R_0 := (1+\lambda)R$ and for $t \in [t_1,t_2]$ set

$$\alpha(t)^2 := R_0^{-n+2}\int_U |\vec{H}|^2\zeta_{R_0,q}^2 d\mu_t, \quad (7.52)$$

$$\beta(t)^2 := R_0^{-n}\int_U \left|\pi_{T_x\mu_t} - \pi_{\mathbb{R}^n}\right|^2\zeta_{R_0,q}^2 d\mu_t, \quad (7.53)$$

where α is only defined for almost every t. The idea is to use Lemma 7.6 with $\rho = R_0$ to obtain plane-like area ratio for the flow inside C_{R_0}. By Lemma 2.8 these area ratios can be transferred to the smaller cylinder C_r but they may become worse. Using Lemma 7.6 again but this time with $\rho = r$ we obtain plane-like area ratios for C_r. By choice of q and Lemma 7.2.1 we have

$$\left\{\zeta_{(1+2^{-n-1}\lambda)r_0,q} = 1\right\} \supset C_{(1-q^{-1}2^{-n-8})(1+2^{-n-1}\lambda)r_0}(0) \supset C_{r_0}(0) \quad (7.54)$$

for all $r_0 \in (0,\infty)$, where we calculated

$$(1-q^{-1}2^{-n-8})(1+2^{-n-1}\lambda) = (1-2^{-2n-8}\lambda)(1+2^{-n-1}\lambda) \geq 1,$$

as $q := 2^n\lambda_1^{-1}$. In particular for $r_0 = R$ we obtain

$$C_{R_0}(0) \supset \mathrm{spt}\zeta_{R_0,q} \supset \{\zeta_{R_0,q} = 1\} \supset C_R(0), \quad (7.55)$$

where we used $R_0 = (1+\lambda)R \geq (1+2^{-n-1}\lambda)R$

We want to apply Lemma 7.6 with $s_1 = t_1$, $s_2 = t_2$, $K = M$, $\rho = R_0$ and $\eta = h$. We immediately see that (7.46) implies (7.34) and (7.47) implies (7.35). By (7.48), $q := 2^n\lambda^{-1}$ and $\lambda \leq 1$ we can estimate

$$
\begin{aligned}
R_0^{-n}\int_U \zeta_{R_0,q}^2 d\mu_{t_1} &\leq R^{-n}\mu_{t_1}\left(C_{R_0}(0)\right) \leq (2-\lambda)\omega_n = (2-2^{-n}q^{-1})\omega_n,\\
R_0^{-n}\int_U \zeta_{R_0,q}^2 d\mu_{t_2} &\geq (1+\lambda)^{-n}R^{-n}\mu_{t_2}\left(C_R(0)\right) \geq 2^{-n}\lambda\omega_n = q^{-1}\omega_n
\end{aligned}
$$

which verifies (7.36). Here we had to use (7.55) for the first estimate in the second line. To apply Lemma 7.6 as above we also have to choose $h_0 \leq \eta_0$ and $\Lambda \geq (1+\lambda)^2 Q$, where $Q \in (1,\infty)$ depends on q and η_0 depends on M and q. Note that q is determined by λ. Lemma 7.6 then yields

$$
\left| R_0^{-n}\int_U \zeta_{R_0,q}^2 d\mu_s - \varpi_q \right| \leq \frac{\varpi_q}{2} \tag{7.56}
$$

and

$$
R_0^{-2}\int_{s-R_0^2}^{s+R_0^2} \alpha(t)^2 dt \leq QMh^2, \quad R_0^{-2}\int_{s-R_0^2}^{s+R_0^2} \beta(t)^2 dt \leq QMh^2 \tag{7.57}
$$

for all $s \in [t_1 + QR_0^2, t_2 - QR_0^2]$. Note that by $t_2 - t_1 \geq 2\Lambda \geq 2Q$ the time interval is non-empty. To obtain (7.56) we estimated $QMh^2 \leq \frac{\varpi_q}{2}$, as $h \leq h_0$ for h_0 small depending on λ and M. In view of definitions (7.52) and (7.53), as $R \leq R_0 \leq 2R$ and by (7.55) the inequalities in (7.57) already verify (7.50) and (7.51) for $\Lambda \geq 2^{n+2}Q$.

Next we want to use Lemma 2.8.2 at times where mean curvature- and tilt-excess are small. By (7.57) we have

$$
\int_{s-R_0^2}^{s+R_0^2} \left(\alpha(t)^2 + \beta(t)^2\right) dt \leq QMh^2 R_0^2
$$

for all $s \in [t_1 + QR_0^2, t_2 - QR_0^2]$. In particular for $\Lambda \geq (Q+1)(1+\lambda)^2$ this holds for $s = t_1 + (Q+1)R_0^2$ and $s = t_2 - (Q+1)R_0^2$, so we can find

$$
\begin{aligned}
a_1 &\in \left(t_1 + QR_0^2, t_1 + (Q+2)R_0^2\right),\\
a_2 &\in \left(t_2 - (Q+2)R_0^2, t_2 - QR_0^2\right)
\end{aligned} \tag{7.58}
$$

with

$$
\alpha(a_i)^2 + \beta(a_i)^2 \leq 4QMh^2, \quad i \in \{1,2\}. \tag{7.59}
$$

Now for the varifolds μ_{a_i}, $i \in \{1,2\}$ use Lemma 2.8.2 with $R_2 = R_0$, $R_1 = r_0$ and $\Phi_\rho = \zeta_{\rho,q}$ to obtain

$$\begin{aligned}&\left| R_0^{-n} \int_U \zeta_{R_0,p}^2 d\mu_{a_i} - r_0^{-n} \int_U \zeta_{r_0,p}^2 d\mu_{a_i} \right| \\ &\leq r_0^{-n} R_0^n \left(\left(n \log\left(\frac{R_0}{r_0}\right) + 2\right) \beta(a_i)^2 + \frac{R_0 - r_0}{R_0} \alpha(a_i)\beta(a_i)\right)\end{aligned}$$

for all $r_0 \in (0, R_0)$, $t \in [t_1, t_2]$, $i \in \{1,2\}$. Here we had to use that as $R_2 = R_0 = (1+\lambda)R$ we see that (7.46) implies (2.5). Thus we can estimate with (7.59)

$$\left| R_0^{-n} \int_U \zeta_{R_0,p}^2 d\mu_{a_i} - r_0^{-n} \int_U \zeta_{r_0,p}^2 d\mu_{a_i} \right| \leq C_n r_0^{-n-1} R^{n+1} Q M h^2 \tag{7.60}$$

for all $r_0 \in (0, R_0)$, $t \in [t_1, t_2]$ $i \in \{1,2\}$. Here we used $R_0 \leq 2R$, $\alpha(a_i)\beta(a_i) \leq \alpha(a_i)^2 + \beta(a_i)^2$, and we estimated $\log\left(\frac{R_0}{r_0}\right) \leq r_0^{-1} R_0$. Now use (7.60) with $r_0 = r \in [\tau R, R]$ and $r_0 = r_2 := (1 + 2^{-n-1}\lambda) r$ at times a_1, a_2 and combine this with (7.56) to obtain

$$r^{-n} \int_U \zeta_{r,q}^2 d\mu_{a_1} \leq \frac{3}{2}\varpi_q + C_n \tau^{-n-1} Q M h^2 \leq \frac{5}{3}\omega_n, \tag{7.61}$$

$$r^{-n} \int_U \zeta_{r,q}^2 d\mu_{a_2} \geq \frac{1}{2}\varpi_q - C_n \tau^{-n-1} Q M h^2 \geq \frac{1}{3}\omega_n, \tag{7.62}$$

$$r_2^{-n} \int_U \zeta_{r_2,q}^2 d\mu_{a_1} \leq \frac{3}{2}\varpi_q + C_n (1 + 2^{-n-1}\lambda)^{-n-1} \tau^{-n-1} Q M h^2 \leq \frac{5}{3}\omega_n, \tag{7.63}$$

$$r_2^{-n} \int_U \zeta_{r_2,q}^2 d\mu_{a_2} \geq \frac{1}{2}\varpi_q - C_n (1 + 2^{-n-1}\lambda)^{-n-1} \tau^{-n-1} Q M h^2 \geq \frac{1}{3}\omega_n, \tag{7.64}$$

where we used $r \geq \tau R$ and we had to choose h_0 small depending on Q, M, τ. Note that Q depends on q which is determined by λ. Here we also used that by (7.58) a_1 and a_2 are contained in $[t_1 + QR_0^2, t_2 - QR_0^2]$, which is necessary for (7.56). Moreover we used $\frac{7}{8}\omega_n \leq \omega_q \leq \omega_n$.

Next we want to apply Lemma 7.6 in the smaller scale $r_0 \in \{r, r_2\}$. Our assumptions (7.46) and (7.47) imply

$$\operatorname{spt}\mu_t \cap C_{r_0}(0) \subset \left\{x \in U, |\pi_{\mathbb{R}^k}(x)| \leq \tau^{-1} h r_0\right\} \subset\subset U, \tag{7.65}$$

$$r_0^{-n} \mu_t\left(C_{r_0}(0)\right) \leq \tau^{-n} M, \tag{7.66}$$

for all $t \in [a_1, a_2]$ for $r_0 \in \{r, r_2\}$, where we used $r \geq \tau R$ and $1 \leq 1 + 2^{-n-1}\lambda$. We want to apply Lemma 7.6 with $s_1 = a_1$, $s_2 = a_2$, $K = \tau^{-n} M$, $\eta = \tau^{-1} h$

one time we will choose $\rho = r$ and the second time $\rho = r_2$. Statements (7.65) and (7.66) imply (7.34) and (7.35) respectively. Estimates (7.61) and (7.62) yield (7.36) for $\rho = r$, as $q^{-1} \leq \frac{1}{3}$. Similarly estimates (7.63) and (7.64) yield (7.36) for $\rho = r_2$. We have to choose $h_0 \leq \tau\eta_0$ to use $\eta = \tau^{-1}h$ as height bound, where η_0 depends on τ, M and q. Note that η_0 is different from the one above because of the larger choice of K. By (7.58) we have $a_2 - a_1 \geq t_2 - t_1 - 2(Q+2)R_0^2 \geq (\Lambda - 2(Q+2))R_0^2$, so for $\Lambda \geq 2(Q+3)$ the time interval is large enough. Then by Lemma 7.6 we obtain

$$\left| r^{-n} \int_U \zeta_{r,q}^2 d\mu_s - \varpi_q \right| \leq \tau^{-n-2} Q M h^2 \tag{7.67}$$

$$\left| r_2^{-n} \int_U \zeta_{r_2,q}^2 d\mu_s - \varpi_q \right| \leq \tau^{-n-2} Q M h^2. \tag{7.68}$$

for all $s \in [a_1 + QR_0^2, a_2 - QR_0^2]$. Actually we would get different time intervals but we coarsely estimated r and r_2 by R_0. Note that we applied Lemma 7.6 with the same q as above, so Q is the same as before. By (7.67) we can now estimate for r

$$\begin{aligned} r^{-n}\mu_s\left(C_r(0)\right) &\geq r^{-n} \int_U \zeta_{r,q}^2 d\mu_s \geq \varpi_q - \tau^{-n-2} Q M h^2 \\ &\geq (1 - q^{-1}2^{-8})\omega_n - \tau^{-n-2} Q M h^2 \geq (1-\lambda)\omega_n \end{aligned} \tag{7.69}$$

for all $s \in [a_1 + QR_0^2, a_2 - QR_0^2]$, where we used $q = 2^n\lambda^{-1}$, Lemma 7.2.4 and $h \leq h_0$ for h_0 small depending on λ, τ, Q, M. Using (7.54) with $r_0 = r$ we have $\{\zeta_{r_2,p} = 1\} \supset C_r(0)$, as $r_2 = (1 + 2^{-n-1}\lambda)r$. Then by (7.68) we can estimate

$$\begin{aligned} r^{-n}\mu_s\left(C_r(0)\right) &\leq (1 + 2^{-n-1}\lambda)^n r_2^{-n} \int_U \zeta_{r_2,q}^2 d\mu_s \\ &\leq (1 + 2^{-1}\lambda)\left(\varpi_q + \tau^{-n-2} Q M h^2\right) \leq (1+\lambda)\omega_n \end{aligned} \tag{7.70}$$

for all $s \in [a_1 + QR_0^2, a_2 - QR_0^2]$, where we used $\varpi_q \leq \omega_n$ and $h \leq h_0$ for h_0 small depending on λ, τ, Q, M.

Inequalities (7.69) and (7.70) then imply (7.49) for all $s \in [a_1 + QR_0^2, a_2 - QR_0^2]$. By (7.58) we see that $[a_1 + QR_0^2, a_2 - QR_0^2] \supset [t_1 + 2(Q+1)R_0^2, t_2 - 2(Q+1)R_0^2]$, so as $R_0 = (1+\lambda)R$, we can choose $\Lambda \geq 2(Q+1)(1+\lambda)^2$ to establish the result. Note that Q depends on q which is determined by λ. $\square$

In the above setting the bounds in (7.48) remain valid for points $y \in B_{\frac{\lambda}{2}}(y_0)$ in slightly weaker form, which is shown in the next lemma.

7.8 Lemma. *For every $\lambda_1 \in (0,1]$ there exists a $\Lambda_1 \in (1,\infty)$ such that, for every $M \in [1,\infty)$ there exists a $\eta_1 \in (0,1)$ such that for all $R_0 \in (0,\infty)$, $\eta \in [0,\eta_1]$, $t_1, t_2 \in \mathbb{R}^{n+k}$ with $t_2 - t_1 > 2\Lambda_1 R_0^2$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in $B_{(1+2\lambda_1)R_0}(0)$ with*

$$\mathrm{spt}\mu_t \subset \left\{x \in \overline{B_{(1+2\lambda_1)R_0}(0)}, |\pi_{\mathbb{R}^k}(x)| \le \eta R_0\right\}, \tag{7.71}$$

$$R_0^{-n}\mu_t\left(B_{(1+2\lambda_1)R_0}(0)\right) \le M \tag{7.72}$$

for all $t \in [t_1,t_2]$ and

$$R_0^{-n}\mu_{t_1}\left(B_{(1+2\lambda_1)R_0}(0)\right) \le (2-\lambda_1)\omega_n\ , \quad R_0^{-n}\mu_{t_2}\left(B_{R_0}(0)\right) \ge \lambda_1\omega_n. \tag{7.73}$$

Set $\rho_0 := 2^{-2}R_0$, then for all $y \in B^n_{\lambda_1 R_0}(0) \times \{0\}^k$ the estimates

$$\rho_0^{-n}\mu_t\left(B_{(1+2^{-n-2}\lambda_1)\rho_0}(y)\right) \le \frac{3}{2}\omega_n\ , \quad \rho_0^{-n}\mu_t\left(B_{\rho_0}(y)\right) \ge \frac{1}{2}\omega_n \tag{7.74}$$

hold for all $t \in [t_1 + \Lambda_1 R_0^2, t_2 - \Lambda_1 R_0^2]$.

Proof. We want to apply Proposition 7.7 with $R = \left(1 + \frac{\lambda_1}{4}\right)R_0$. For given $\lambda_1 \in (0,1]$ fix an arbitrary $y \in B^n_{\lambda_1 R}(0) \times \{0\}^k$. As $\eta \le \eta_1$ for $\eta_1 < 2^{-1}\lambda_1$ we see by (7.71)

$$(\mathrm{spt}\mu_t \cap C_{(1+\frac{\lambda_1}{2})R_0}(y)) \subset (\mathrm{spt}\mu_t \cap C_{(1+\frac{3}{2}\lambda_1)R_0}(0)) \subset\subset B_{(1+2\lambda_1)R_0}(0)$$

for all $t \in [t_1,t_2]$. Now estimate

$$(1+2^{-n-4}\lambda_1)\left(1+\frac{\lambda_1}{4}\right) \le 1 + \frac{\lambda_1}{4} + 2^{-n-3}\lambda_1 \le 1 + \frac{\lambda_1}{2}.$$

Then for $R := \left(1+\frac{\lambda_1}{4}\right)R_0$ we have

$$\mathrm{spt}\mu_t \cap C_{(1+2^{-n-4}\lambda_1)R}(y) \subset\subset B_{(1+2\lambda_1)R_0}(0) \tag{7.75}$$

for all $t \in [t_1,t_2]$. In view of (7.75) we can use (7.73) to estimate

$$\begin{aligned}R^{-n}\mu_{t_1}\left(C_{(1+2^{-n-4}\lambda_1)R}(y)\right) &\le \left(1+\frac{\lambda_1}{4}\right)^{-n} R_0^{-n}\mu_{t_1}\left(B_{(1+2\lambda_1)R_0}(0)\right)\\ &\le (2-\lambda_1)\omega_n \le (2-2^{-n-4}\lambda_1)\omega_n.\end{aligned}$$

Also by (7.73) we have

$$R^{-n}\mu_{t_2}\left(C_R(y)\right) \ge \left(1+\frac{\lambda_1}{4}\right)^{-n} R_0^{-n}\mu_{t_1}\left(B_{R_0}(0)\right) \ge 2^{-n}\lambda_1\omega_n \ge 2^{-n-4}\lambda_1\omega_n$$

where we used $R \geq R_0$ and $\lambda_1 \leq 1$. This allows us to apply Proposition 7.7 with $U = B_{(1+2\lambda_1)R_0}(0)$, $\lambda := 2^{-n-4}\lambda_1$, $h = \eta$, $R := \left(1 + \frac{\lambda_1}{4}\right) R_0$, $\tau = \frac{1}{8}$ and $y_0 = y$. As $R \geq R_0$ we have $\eta R \geq \eta R_0$, so in view of (7.75) assumption (7.71) implies (7.46). Proposition 7.7 then yields a Λ depending on λ_1 and a h_0 depending on M and λ_1 such that, if $\eta_1 \leq h_0$ we have

$$|r^{-n}\mu_t\left(B_r(y)\right) - \omega_n| \leq 2^{-n-4}\lambda_1\omega_n \tag{7.76}$$

for all $r \in (2^{-3}R, R)$ and all $t \in [t_1 + \Lambda_1 R^2, t_2 - \Lambda_1 R^2]$. Here we chose $\Lambda_1 \geq 4\Lambda$. Note that then, as $2R_0 \geq R$

$$[t_1 + \Lambda R^2, t_2 - \Lambda R^2] \supset [t_1 + \Lambda_1 R_0^2, t_2 - \Lambda_1 R_0^2].$$

Now consider $\rho_0 = 2^{-2}R_0$. As $R = (1 + \frac{\lambda_1}{4})R_0$ and $\lambda_1 \in (0, 1]$ we see $\rho_0 \in (2^{-3}R, R)$. Using (7.76) with $r = \rho_0$ we obtain

$$\rho_0^{-n}\mu_t\left(B_{\rho_0}(y)\right) \geq (1 - 2^{-n-4}\lambda_1)\omega_n \geq \frac{1}{2}\omega_n$$

for all $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$, where we used $\lambda_1 \leq 1$. Also we have

$$\frac{R}{8} \leq \frac{R_0}{4} \leq (1 + 2^{-n-2}\lambda_1)\rho_0 \leq (1 + 2^{-2}\lambda_1)R_0 = R$$

as $\lambda_1 \leq 1$, $\rho_0 = 2^{-2}R_0$ and $R = (1 + \frac{\lambda_1}{4})R_0$. Thus we can use (7.76) with $r = (1 + 2^{-n-2}\lambda_1)\rho_0$ to obtain

$$\begin{aligned}
&\rho_0^{-n}\mu_t\left(B_{(1+2^{-n-2}\lambda_1)\rho_0}(y)\right) \\
&= (1 + 2^{-n-2}\lambda_1)^n \left((1 + 2^{-n-2}\lambda_1)\rho_0\right)^{-n} \mu_t\left(B_{(1+2^{-n-2}\lambda_1)\rho_0}(y)\right) \\
&\leq (1 + 2^{-2}\lambda_1)(1 + 2^{-n-4}\lambda_1)\omega_n \leq (1 + 2^{-2})(1 + 2^{-4})\omega_n \leq \frac{3}{2}\omega_n
\end{aligned}$$

for all $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$, where we used $\lambda_1 \leq 1$. This establishes the result. $\square$

8 Local Regularity

In this section we establish Theorem 8.4, which is our version of Brakke's local regularity theorem [B, 6.10]. This result states that, if a Brakke flow in a suitable region is contained in a narrow enough slab and its area ratios are controlled by certain bounds, then in a smaller region it is actually a smooth graphical mean curvature flow. All our later results like Theorem 9.7 or Theorem 11.7 will use this fact in some way. The proof of Theorem 8.4 is based on an iteration argument stated in Lemma 8.1, which is based on [B, 6.9]. Note that the original proof in [B, 6.9] contains a serious gap in the usage of the clearing out lemma in order to obtain a supremum bound on the height. This gap we have fixed in section 5. We also give an alternative proof replacing the height estimate from section 5 by Corollary 6.8 from section 6. Besides this the overall strategy is very similar as in [B]. There also exists a new proof of local regularity, using very different techniques, see [KT].

The key to proofing local regularity is the following iteration lemma. It states that under certain conditions, if a Brakke flow in some region for a large enough time is contained in a narrow slab with respect to $\mathbb{R}^n$ then we can find a subspace T such that in a smaller region for a smaller time interval the flow is in a more narrow slab with respect to T.

8.1 Lemma (Iteration Lemma, [B, 6.9])**.** *For every $\lambda_0 \in (0, 2^{-n-5}]$ there exists a $\Lambda_0 \in (1, \infty)$ such that, for every $\epsilon \in (0, 1]$ there exists a $\delta_0 \in (0, 1)$ such that, for every $\delta \in (0, \delta_0]$ there exists a $\eta_0 \in (0, 1)$ such that, for all $R_0 \in (0, \infty)$ and $\eta \in [0, \eta_0]$ the following holds: Let $(\mu_t)_{t\in[-\Lambda_0 R_0^2, \Lambda_0 R_0^2]}$ be a Brakke flow in $B_{3R_0}(0)$ with*

$$\mathrm{spt}\mu_t \subset \left\{ x \in \overline{B_{3R_0}(0)},\ |\pi_{\mathbb{R}^k}(x)| \leq \eta R_0 \right\}, \tag{8.1}$$

$$R_0^{-n}\mu_t\left(B_{(1+\lambda_0)R_0}(0)\right) \leq \frac{3}{2}\omega_n\ , \quad R_0^{-n}\mu_t\left(B_{R_0}(0)\right) \geq \frac{1}{2}\omega_n \tag{8.2}$$

for all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. Then there exist an n-dimensional subspace $T \subset \mathbb{R}^{n+k}$ with $|\pi_T - \pi_{\mathbb{R}^n}| \leq \delta^{-\epsilon}\eta$, and a point $z \in \{0\}^n \times \mathbb{R}^k$ with $|z| \leq \sqrt{k}\eta R_0$ such that

$$\mathrm{spt}\mu_t \cap B_{16\delta R_0}(z) \subset \left\{ x \in \mathbb{R}^{n+k}, |\pi_T^\perp(x-z)| \leq \delta^{2-\epsilon}\eta R_0 \right\} \tag{8.3}$$

$$(\delta R_0)^{-n}\mu_t\left(B_{(1+2\lambda_0)\delta R_0}(z)\right) \leq \frac{5}{4}\omega_n \tag{8.4}$$

$$(\delta R_0)^{-n}\mu_t\left(B_{(1-2\lambda_0)\delta R_0}(z)\right) \geq \frac{3}{4}\omega_n \tag{8.5}$$

for all $t \in [-\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2]$.

Proof. For given $R_0 \in (0,\infty)$ and Λ_0 set

$$\rho := 2^{-4}R_0,$$
$$r_0 := \sqrt{2\Lambda_0}\delta R_0.$$

For given $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$ set

$$\alpha_t^2 := R_0^{-n+2}\int_{C_{R_0}(0)} |\vec{H}(x)|^2 d\mu_t(x), \tag{8.6}$$

$$\beta_t^2 := R_0^{-n}\int_{C_{R_0}(0)} |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 \, d\mu_t(x). \tag{8.7}$$

Note that α_t is only defined for almost every $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$.

By (8.1) and $\eta \le \eta_0$ we can choose η_0 small enough depending on δ such that

$$\mathrm{spt}\mu_t \cap C_{(1+2^{-1}\lambda_0)r}(0) \subset \mathrm{spt}\mu_t \cap B_{(1+\lambda_0)r}(0) \tag{8.8}$$

for all $r \in [2^{-1}\delta R_0, R_0]$ and all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. Using (8.8) with $r = R_0$ combined with (8.1) and (8.2) implies

$$R_0^{-n-2}\int_{C_{R_0}(0)} |\pi_{\mathbb{R}^k}(x)|^2 \, d\mu_t(x) \le \frac{3}{2}\omega_n\eta^2 \tag{8.9}$$

for all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. First we use the Proposition 7.7 to obtain bounds on mean cruvature- and tilt-excess, as well as area ratios close to ω_n. In particular this yields good Lipschitz approximations.

For $r \in (0, R)$ we want to use Proposition 7.7 with $U = B_{3R_0}(0)$, $R = R_0$, $y_0 = 0$, $\lambda = 2^{-1}\lambda_0$, $h = \eta$, $t_1 = -\Lambda_0 R^2$, $t_2 = \Lambda_0 R^2$, $M = \frac{3}{2}\omega_n$ and $\tau = 2^{-1}\delta R_0$. In view of (8.8), we see that (8.1) and (8.2) verify (7.46), (7.47) and (7.48). Theorem 7.7 then yields a Λ depending on λ_0 and a h_0 depending on λ_0 and δ such that

$$\left| r^{-n}\mu_s\left(C_r(0)\right) - \omega_n \right| \le 2^{-1}\lambda_0\omega_n, \tag{8.10}$$

$$R_0^{-2}\int_{s-R_0^2}^{s+R_0^2} \left(\alpha_t^2 + \beta_t^2\right) dt \le \frac{3}{2}\omega_n\Lambda_0\eta^2 \tag{8.11}$$

for all $r \in (2^{-1}\delta R_0, R_0]$ and $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. Here we chose $\eta_0 \le h_0$ and $\Lambda_0 \ge 2\Lambda$. Note that then $h = \eta \le h_0$ and

$$[-\Lambda_0 R_0^2 + \Lambda R_0^2, \Lambda_0 R_0^2 - \Lambda R_0^2] \supset [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2].$$

Now for some $x \in \{0\}^n \times \mathbb{R}^k \cap B_{\sqrt{k}\eta R_0}(0)$ we can use (8.10) to obtain density ratios like in (8.4) and (8.5). As $\delta \leq \delta_0$ we have

$$[-\Lambda_0 \delta R_0^2, \Lambda_0 \delta R_0^2] \subset [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2] \tag{8.12}$$

for $\delta_0 \leq 2^{-2}$. Set $r_1 := (1+2\lambda_0)\delta R_0$ and $r_2 := (1-3\lambda)(1+\lambda_0)^{-1}\delta R_0$. Using (8.10) with $r = r_1$ we can estimate

$$\begin{aligned}(\delta R_0)^{-n}\mu_s\left(B_{(1+2\lambda_0)\delta R_0}(x)\right) &\leq (1+2\lambda_0)^n r_1^{-n}\mu_s\left(C_{r_1}(x)\right)\\ &\leq (1+2\lambda_0)^n(1+\lambda_0)\omega_n\end{aligned}$$

for all $x \in \{0\}^n \times \mathbb{R}^k$ and all $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. Now as $\lambda_0 \leq 2^{-n-5}$, we can calculate $(1+2\lambda_0)^n(1+\lambda_0) \leq 1 + 2^{n+3}\lambda_0 \leq 1 + 2^{-2}$, to see

$$(\delta R_0)^{-n}\mu_s\left(B_{(1+2\lambda_0)\delta R_0}(x)\right) \leq \frac{5}{4}\omega_n \tag{8.13}$$

for all $x \in \{0\}^n \times \mathbb{R}^k$ and all $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. Using (8.10) with $r = r_2$ and (8.8) with $r = r_2$ we can estimate

$$\begin{aligned}(\delta R_0)^{-n}\mu_s\left(B_{(1-3\lambda_0)\delta R_0}(0)\right) &= (1-3\lambda_0)^{-n}(1+\lambda_0)^n r_2^{-n}\mu_s\left(B_{(1+\lambda_0)r_2}(0)\right)\\ &\geq r_2^{-n}\mu_s\left(C_{r_2}(0)\right) \geq (1-\lambda_0)\omega_n \geq \frac{3}{4}\omega_n\end{aligned}$$

for all $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$, where we also used $\lambda_0 \leq 2^{-1}$. Now as $\eta \leq \eta_0$ we have $\sqrt{k}\eta R_0 \leq \lambda_0 \delta R_0$ for η_0 small depending on λ_0 and δ. Then

$$(\delta R_0)^{-n}\mu_s\left(B_{(1-2\lambda_0)\delta R_0}(x)\right) \geq (\delta R_0)^{-n}\mu_s\left(B_{(1-3\lambda_0)\delta R_0}(0)\right) \geq \frac{3}{4}\omega_n \tag{8.14}$$

for all $x \in B_{\sqrt{k}\eta R_0}(0)$ and all $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$.

Consider μ_s for $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. We want to use Theorem 2.9 with $R = 2\rho = \frac{R_0}{8}$ and $l = 1$. Inequality (8.10) implies (2.16) and (2.17) for $R = 2\rho$. In view of $B_{14\rho}(0) \subset C_{R_0}(0)$ and definitions (8.6), (8.7) and estimate (8.9), the estimates (2.18), (2.19) and (2.20) hold for $\alpha = C_n\alpha_s$, $\beta = C_n\beta_s$ and $\gamma = C_n\eta$. Using $\eta \leq \eta_0$ we can also achieve $C_n\eta \leq \gamma_0$ for η_0 small enough. Then for every $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$ Theorem 2.9 and Remark 2.10 yield the existence of a Lipschitz function $f_s : B^n_{2\rho}(0) \to \mathbb{R}^k$ and $F_s(\hat{y}) := (\hat{y}, f_t(\hat{y}))$ such that

$$\operatorname{lip}(f_s) \leq 1, \ \ \sup|f_s| \leq \eta R_0 \tag{8.15}$$

and such that we can estimate

$$\mu\left(C_{2\rho}(0) \setminus X_s\right) + \mathscr{L}^n\left(B^n_{2\rho}(0) \setminus Y_t\right) \leq C_n\rho^n E_s, \tag{8.16}$$

where for $M_s := \text{graph}(f_s)$

$$\begin{aligned} Y_s := \big\{\hat{y} \in B^n_{2\rho}(0) :\ & y := F_t(\hat{y}) \in C_{2\rho}(0) \cap U,\ \Theta^n(\mu_s, y) = 1, \\ & T_y M_s \text{ and } T_y\mu_s \text{ exist with } T_y M_s = T_y \mu_s\big\}, \\ X_s := \{x \in C_{2\rho}(0) : & \exists \hat{y} \in Y\ x = (\hat{y}, f(t,\hat{y}))\}. \end{aligned} \tag{8.17}$$

Here the error term E_s is given by

$$E_s = \left(\alpha_s^{\frac{2n}{n-2}} \delta_{n\geq 3} + \beta_s^2 + \eta^2\right). \tag{8.18}$$

Note that for $\alpha_s \leq 1$, we have $\alpha_s^{\frac{2n}{n-2}} \delta_{n\geq 3} \leq \alpha_s^2$. In order to obtain the height bound in (8.15), we had to cut-off the function f one would normally obtain from Theorem 2.9 by setting

$$f_s(\hat{y}) = \begin{cases} \eta R_0 & \text{if } f(\hat{y}) > \eta R_0 \\ f(\hat{y}) & \text{if } |f(\hat{y})| \leq \eta R_0 \\ -\eta R_0 & \text{if } f(\hat{y}) < \eta R_0 \end{cases}$$

for all $\hat{y} \in B^n_{2\rho}(0)$. This does not increase the Lipschitz constant and in view of (8.1) it does not change the sets X_s, Y_s, so (8.16) remains valid. To obtain (8.16), we also used (8.1), in order to see that $B^n_{2\rho}(0) \times B^k_{2\rho}(0) \cap \text{spt}\mu_s \subset C_{2\rho}(0)$.

The Lipschitz approximation will be used to write an integral over μ_s as an integral over $\mathbb{R}^n$. For $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$, $r \leq 2\rho$ and an L^1-integrable function

$$\phi : C_r(0) \cap [\text{spt}\mu_s \cup \text{graph}(f_s)] \to \mathbb{R}$$

we can use Remark 2.11 to estimate

$$\left| \int_{C_r(0)} \phi(x) d\mu_s(x) - \int_{B^n_r(0)} \phi(F_s(\hat{x})) d\mathscr{L}^n(\hat{x}) \right| \leq C_n R_0^n \sup|\phi|\, E_s, \tag{8.19}$$

where $\sup|\phi|$ is the essential supremum of $|\phi|$ over the set

$$C_r(0) \cap [\text{spt}\mu_s \cup \text{graph}(f_s)].$$

Here we used Remark 2.11 with $\phi\big|_{C_r(0)}$ and that by (8.1) we have $B^n_r(0) \times B^k_{2\rho}(0) \cap \text{spt}\mu_s = C_r(0) \cap \text{spt}\mu_s$. Also we used $16\rho = R_0$.

For given $\epsilon \in (0,1]$, $\delta \in (0, \delta_0]$ and $N \in (1,\infty)$ consider the parameters

$$\begin{aligned} p &:= \delta^N R_0^2, \\ q_0 &:= \delta^\epsilon R_0^2, \end{aligned}$$

where N will be chosen large depending on n, k. To define suitable T and z, we need to find a time s_0 near $-7\Lambda_0\delta^2 R_0^2 - q_0$, where the Lipschitz approximation is actually good. As $\delta \leq 1$ and for $\Lambda_0 \geq 8$ we have $q_0 = \delta^\epsilon R_0^2 \leq 2^{-3}\Lambda_0 R_0^2$. Let $\delta_0 \leq 2^{-3}$ then $7\Lambda_0\delta^2 R_0^2 \leq 2^{-3}\Lambda_0 R_0^2$, so

$$-7\Lambda_0\delta^2 R_0^2 - q_0 \in [-2^{-2}\Lambda_0 R_0^2, 2^{-2}\Lambda_0 R_0^2]. \tag{8.20}$$

Then by inequality (8.11) and $\delta \leq 1$ there exists

$$s_0 \in \big(-(7\Lambda_0 + 1)\delta^2 R_0^2 - q_0, -7\Lambda_0\delta^2 R_0^2 - q_0\big) \tag{8.21}$$

with

$$\alpha_{s_0}^2 + \beta_{s_0}^2 \leq C_n \Lambda_0 \delta^{-2}\eta^2. \tag{8.22}$$

Statements (8.20), (8.21) and $\delta \leq \delta_0 \leq 2^{-3}$, $\Lambda_0 \geq 1$ imply

$$[s_0, \Lambda_0\delta^2 R_0^2] \subset [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2], \tag{8.23}$$

in particular a Lipschitz approximation exists for all $t \in [s_0, \Lambda_0\delta^2 R_0^2]$. By (8.22) and $\eta \leq \eta_0$ we have $\alpha_{s_0} \leq 1$, for η_0 small enough depending on δ and Λ_0, then

$$E_{s_0} \leq \delta^{-3}\eta^2, \tag{8.24}$$

where we also used $\delta \leq \delta_0$ and chose δ_0 small depending on Λ_0. Here E_t is the error term defined in (8.18). This lets us define T and z as follows: For $i \in \{1, \ldots, n\}$, $j \in \{1, \ldots, k\}$ and $\hat{y} \in \mathbb{R}^n$ set

$$t_j(\hat{y}) := \sum_{i=1}^{n} \hat{y}_i t_{ij}, \quad t_{ij} := -\int_{B_\rho^n(0)} f_{s_0} \cdot \mathbf{e}_j \frac{\partial \Psi_\rho}{\partial x_i}(q_0, \hat{x}) d\mathscr{L}^n(\hat{x}), \tag{8.25}$$

$$z_{n+j} := \int_{B_\rho^n(0)} f_{s_0} \cdot \mathbf{e}_j \Psi_\rho(q_0, \hat{x}) d\mathscr{L}^n(\hat{x}). \tag{8.26}$$

Then set $T := \{\hat{x} + \sum_{j=1}^k t_j(\hat{x})\mathbf{e}_{n+j},\ \hat{x} \in \mathbb{R}^n\}$ and $z := (0, z_{n+1}, \ldots, z_{n+k})$. In view of (8.23) there exists a Lipschitz function $f_{s_0} : B_{2\rho}^n(0) \to \mathbb{R}^k$ satisfying (8.15) and (8.16). By choice of s_0 the error term E_{s_0} is small, which will let us estimate $|t_{ij}|$ and $|z_{n+j}|$.

To estimate the tilt of T let $i \in \{1, \ldots, n\}, j \in \{1, \ldots, k\}$. By Lemma 4.9 with measure $\mathscr{L}^n$ and $P_1 = 1$, we can estimate using also (8.15)

$$\begin{aligned} |t_{ij}| &\leq \eta R_0 \int_{B_\rho^n(0)} |D\Psi_\rho(q_0, \hat{y})| \mathscr{L}^n(\hat{y}) \\ &\leq C_n \eta \left(R_0 \int_{B_\rho^n(0)} |D\Psi(q_0, \hat{y})| \mathscr{L}^n(\hat{y}) + R_0^{-2} q_0 \right) \end{aligned}$$

where we used $16\rho = R_0$ and $\rho^{-2}q_0 = C_n\delta^\epsilon \le C_n\delta_0^\epsilon \le \kappa_1$ for δ_0 small enough depending on ϵ. Also we estimated $\zeta \le 1$. As $D\Psi(q_0,\hat{y}) = (2q_0)^{-1}\hat{y}\Psi(q_0,\hat{y})$, we can use Lemma 4.6 to estimate

$$|t_{ij}| \le C_n\eta\left(q_0^{-\frac{1}{2}}R_0 + R_0^{-2}q_0\right) \le C_n\delta^{-\frac{\epsilon}{2}}\eta \tag{8.27}$$

for all $i \in \{1,\dots,n\}, j \in \{1,\dots,k\}$, where we used $q_0 = \delta^\epsilon R_0^2$ and $\delta \le 1$. To estimate the excess of T, we define $a_i = \mathbf{e}_i + \sum_{j=1}^k t_{ij}\mathbf{e}_{n+j}$ for $1 \le i \le n$. Then we have $T = \operatorname{span}(a_i)_{1\le i\le n}$ and our bound for the $|t_{ij}|$ yields $|a_i - \mathbf{e}_i| \le C_n\delta^{-\frac{\epsilon}{2}}\eta$. Thus by Lemma A.10.2 with $B = \mathbb{R}^n$ we obtain

$$|\pi_T - \pi_{\mathbb{R}^n}| \le \delta^{-\epsilon}\eta, \tag{8.28}$$

where we used $\delta \le \delta_0$, so $C_n\delta^{-\frac{\epsilon}{2}} \le \delta^{-\epsilon}$ for δ_0 small enough depending on ϵ.

To estimate $|z|$ let $j \in \{1,\dots,k\}$. As $\zeta \le 1$ and $\int_{\mathbb{R}^n}\Psi(q_0,\hat{y})\mathscr{L}^n(\hat{y}) = 1$, we can estimate using (8.15)

$$|z_{n+j}| \le \eta R_0 \int_{B_\rho^n(0)} \Psi(q_0,\hat{y})\mathscr{L}^n(\hat{y}) \le \eta R_0$$

for all $j \in \{1,\dots,k\}$ and in view of $\hat{z} = 0$ this yields

$$|z| \le \sqrt{k}\eta R_0. \tag{8.29}$$

Thus (8.25) and (8.26) define T and z that are close to $\mathbb{R}^n$ and 0 as supposed. In particular we have $z \in \{0\}^n \times \mathbb{R}^k \cap B_{\sqrt{k}\eta R_0}(0)$, so in view of (8.12) the estimates (8.13) and (8.14) imply (8.4) and (8.5).

The main part is now to show that T and z actually provide a smaller slab containing the Brakke flow. We can either use the distance estimate from the monotonicity section Corollary 6.8 or the improving height estimate from the clearing out section Lemma 5.12, which will both require to estimate pretty similar integrals.

1. With Corollary 6.8:
Fix an arbitrary $j \in \{1,\dots,k\}$. We want to use Corollary 6.8 with $v := \mathbf{e}_{n+j} - \sum_{i=1}^n t_{ij}\mathbf{e}_i$ and $x_0 = y_0 = z$. Thus

$$|(x-z)\cdot v| = \Phi(x) := |x_{n+j} - t_j(\hat{x}) - z_{n+j}| \tag{8.30}$$

for all $x \in \mathbb{R}^{n+k}$, where we used $\hat{z} = 0$. Note that by (8.27), (8.29) and $\delta \le 1$ we obtain that Φ satisfies the height bound

$$|\Phi(x)| \le C_n\delta^{-\frac{\epsilon}{2}}\eta R_0$$

for all $x \in \mathbb{R}^n$ with $|x_{n+j}| \leq \delta^{-\frac{\epsilon}{2}}\eta R_0$, so by (8.1) and (8.15)

$$|\Phi(x)| \leq C_n \delta^{-\frac{\epsilon}{2}}\eta R_0 \tag{8.31}$$

for all $x \in \mathrm{spt}\mu_t \cup \mathrm{graph}(f_t)$ for all $t \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. In view of (8.30), Corollary 6.8 with $t_0 = \Lambda_0\delta^2 R_0^2$ and $r = r_0 = \sqrt{2\Lambda_0}\delta R_0$ yields

$$\begin{aligned} S &:= \sup_{[-\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2]} \sup_{spt\mu_t \cap B_{16\delta R_0}(z)} \Phi \\ &\leq C_n r_0^{-n-2} \int_{-7\Lambda_0\delta^2 R_0^2}^{\Lambda_0\delta^2 R_0^2} \int_{B_{4r_0}(z)} \Phi d\mu_t \, dt. \end{aligned} \tag{8.32}$$

where we used $\Lambda_0 \geq 2^7$, so $r_0 \geq 16\delta R_0$.

Our aim is to show that S is smaller then $\delta^{1-\epsilon}\eta r_0$. To estimate the right hand side of (8.32), we want to bring in our Lipschitz approximations. This is only useful for times s where the error term E_s can be estimated. We obtain an L^1-bound on E_s, but only integrating over times where $\alpha \leq 1$. In view of (8.20) we have

$$[-7\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2] \subset [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2], \tag{8.33}$$

as $\delta \leq 1$ and $q_0 \geq 0$. Then we can use (8.11) to obtain

$$\int_{-7\Lambda_0\delta^2 R_0^2}^{\Lambda_0\delta^2 R_0^2} \left(\alpha_t^2 + \beta_t^2\right) dt \leq C_n \Lambda_0 \eta^2 R_0^2. \tag{8.34}$$

Here we also used that $8\Lambda_0\delta R_0^2 \leq 2R_0^2$ as $\delta \leq \delta_0$, for δ_0 small depending on Λ_0. Set

$$I_\alpha := \left\{ t \in [-7\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2] : \alpha_t \leq 1 \right\}, \tag{8.35}$$

so for all $t \in I_\alpha$ we can estimate $\alpha_t^{\frac{2n}{n-2}} \leq \alpha_t^2$. This will allow us to bound E_t for $t \in I_\alpha$. The Lipschitz approximations above live in the cylinder $C_{2\rho}(0)$, so we want the balls $B_{4r_0}(z)$ to be contained in this cylinder. As $\rho = 2^{-4}R_0$, and $\delta \leq \delta_0$, for δ_0 small depending on Λ_0 we have $8\sqrt{n\Lambda_0}\delta R_0 \leq 2^{-1}\rho$. Also as $|z| \leq \sqrt{k}\eta R_0$ and $\eta \leq \eta_0$, for η_0 small depending on k, we have $|z| \leq 2^{-1}\rho$. Thus we conclude

$$B_{4r_0}(z) \subset B_{8\sqrt{n\Lambda_0}\delta R_0}(z) \subset B_\rho(0) \subset B_{R_0}(0), \tag{8.36}$$

where we used that $r_0 = \sqrt{2\Lambda_0}\delta R_0$. Here including the second ball looks a bit out of place, but will be of use later. In view of (8.33) for all $t \in I_\alpha$ a

Lipschitz approximation $f_s : B^n_{2\rho}(0) \to \mathbb{R}^k$ satisfying (8.15) and (8.16) exists. By (8.36) we can use these Lipschitz functions to describe the integral on the right hand side of (8.32). Also by (8.35) we have $\alpha_t^{\frac{2n}{n-2}} \leq \alpha_t^2$ for $t \in I_\alpha$, so the error term defined in (8.18), can be estimate with inequality (8.34), which leads to

$$\int_{I_\alpha} E_t \, dt \leq C_n \Lambda_0 \eta^2 R_0^2. \tag{8.37}$$

Note that by (8.35) and $r_0 = \sqrt{2\Lambda_0}\delta R_0$ we have

$$|I_\alpha| \leq 4r_0^2 = 8\Lambda\delta^2 R_0^2 \leq R_0^2, \tag{8.38}$$

as $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. By (8.34) and definition of I_α we can estimate

$$\left|[-7\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2] \setminus I_\alpha\right| \leq \int_{[-7\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2]\setminus I_\alpha} \alpha_t^2 dt \leq C_n \Lambda_0 \eta^2 R_0^2.$$

Combined with (8.31) this yields

$$\begin{aligned} &\int_{-7\Lambda_0\delta^2 R_0^2}^{\Lambda_0\delta^2 R_0^2} \int_{B_{4r_0}(z)} \Phi d\mu_t \, dt - \int_{I_\alpha} \int_{B_{4r_0}(z)} \Phi d\mu_t \, dt \\ &\leq C_n \Lambda_0 \delta^{-\frac{\epsilon}{2}} \eta^3 R_0^3 \sup_{t \in [-7\Lambda_0\delta^2 R_0^2, \Lambda_0\delta^2 R_0^2]} \mu_t(B_{4r_0}(z)) \leq C_n \Lambda_0 \delta^{-\frac{\epsilon}{2}} \eta^3 R_0^{n+3}, \end{aligned} \tag{8.39}$$

where we used (8.2) and (8.36), to estimate the measure of the ball. In view of (8.31), Φ is bounded on both $\mathrm{spt}\mu_t$ and $\mathrm{graph} f_t$. Then by (8.19) and $\hat{z} = 0$, we obtain

$$\begin{aligned} &\int_{I_\alpha} \int_{C_{4r_0}(z)} \Phi d\mu_t \, dt - \int_{I_\alpha} \int_{B^n_{4r_0}(0)} \Phi(F_t(\hat{y})) d\mathscr{L}^n(\hat{y}) \, dt \\ &\leq \int_{I_\alpha} C_n \delta^{-\frac{\epsilon}{2}} \eta R_0^{n+1} E_t \, dt \leq C_n \Lambda_0 \delta^{-\frac{\epsilon}{2}} \eta^3 R_0^{n+3} \end{aligned} \tag{8.40}$$

where we used (8.37) to estimate E_t. Here we also used $4r_0 = 4\sqrt{2\Lambda_0}\delta R_0 \leq 2^{-3} R_0 = 2\rho$, as $\delta \leq \delta_0$, for δ_0 small depending on Λ_0. Inserting (8.39) and (8.40) into (8.32), we conclude

$$S \leq C_n r_0^{-n-2} \int_{I_\alpha} \int_{B^n_{4r_0}(0)} \Phi(F_t(\hat{y})) d\mathscr{L}^n(\hat{y}) \, dt + \delta^{2-\epsilon} \eta R_0, \tag{8.41}$$

where we used $\Phi \geq 0$, $r_0 = \sqrt{2\Lambda_0}\delta R_0$, $\epsilon \geq 0$, $\Lambda_0 \geq 1$, $\eta \leq \eta_0$ and η_0 small depending on δ and Λ_0. For $\hat{y} \in B^n_\rho(0)$ set

$$f^j_t(\hat{y}) := f_t(\hat{y}) \cdot \mathbf{e}_j.$$

Then by definition (8.30) we have

$$\Phi(F_t(\hat{y})) = \left| f^j_t(\hat{y}) - t_j(\hat{y}) - z_{n+j} \right|$$

and by the triangle inequality we can estimate for $t \in I_\alpha$ and $\hat{y} \in B^n_\rho(0)$

$$\Phi(F_t(\hat{y})) \leq \Phi_1(t, \hat{y}) + \Phi_2(t, \hat{y}) + \Phi_3(t, \hat{y}) + \Phi_4(\hat{y}), \tag{8.42}$$

where the Φ_i are defined by

$$\Phi_1(t, \hat{y}) := \left| f^j_t(\hat{y}) - \int_{B_\rho(\hat{y})} f^j_t \Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| \tag{8.43}$$

$$\begin{aligned} \Phi_2(t, \hat{y}) := \Bigg| &\int_{B_\rho(\hat{y})} f^j_t(\hat{y}) \Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \\ &- \int_{B_\rho(\hat{y})} f^j_{s_0}(\hat{y}) \Psi_\rho(p + t - s_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x})(x) \Bigg| \end{aligned} \tag{8.44}$$

$$\begin{aligned} \Phi_3(t, \hat{y}) := \Bigg| &\int_{B_\rho(\hat{y})} f^j_{s_0}(\hat{y}) \Psi_\rho(p + t - s_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \\ &- \int_{B_\rho(\hat{y})} f^j_{s_0} \Psi_\rho(q_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \Bigg| \end{aligned} \tag{8.45}$$

$$\Phi_4(\hat{y}) := \left| \int_{B_\rho(\hat{y})} f^j_{s_0} \Psi_\rho(q_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) - t_j(\hat{y}) - z_{n+j} \right|. \tag{8.46}$$

In view of (8.41), in order to establish $S \leq \delta^{2-\epsilon}\eta R_0$, we can estimate the space-time integrals over these Φ_i.

In order to estimate Φ_1, we want to use that for small parameter the heat kernel converges to the Dirac delta function, as we showed in Proposition 4.10.2. In view of (8.43), we can use Lemma 4.19 with $g = f^j_t$, $r = \rho$ and $P = 3$ to obtain

$$\begin{aligned} &\int_{B^n_\rho(0)} \Phi_1(t, \hat{y}) d\mathscr{L}^n(\hat{y}) \\ &\leq C_n p^{\frac{1}{2}} \int_{B^n_{2\rho}(0)} |Df^j_t(\hat{x})| d\mathscr{L}^n(\hat{x}) + (\rho^{-2}p)^3 \eta R_0 \rho^n \end{aligned} \tag{8.47}$$

for all $t \in I_\alpha$, where we used that by $p = \delta^N R_0^2 = 2^8 \delta^N \rho^2$, we can estimate $\rho^{-2} p \leq 2^8 \delta_0^N$, so for δ_0 small we have $\rho^{-2} p \leq \kappa$. Also we used (8.15) to estimate $\sup |f_t^j| \leq \eta R_0$. By (8.15) we also have $\mathrm{lip} f_t \leq 1$, so by Proposition A.12.1 we can estimate

$$|Df_t^j(\hat{x})| \leq C_n |\pi_{T_{F_t(\hat{x})} M_t} - \pi_{\mathbb{R}^n}|$$

for all $\hat{x} \in B_{2\rho}^n(0)$, where $M_t = \mathrm{graph} f_t$. So with (8.16) we can estimate

$$\int_{B_{2\rho}^n(0)} |Df_t^j(\hat{x})| d\mathscr{L}^n(\hat{x}) \leq C_n \int_{B_{2\rho}^n(0) \cap Y_t} |\pi_{T_{F_t(\hat{x})} M_t} - \pi_{\mathbb{R}^n}| d\mathscr{L}^n(\hat{x}) + C_n E_t R_0^n.$$

Then by the definition of Y_t in (8.17) and as $JF \geq 1$, we have

$$\int_{B_{2\rho}^n(0)} |Df_t^j(\hat{x})| d\mathscr{L}^n(\hat{x}) \leq C_n \int_{C_{2\rho}^n(0) \cap X_t} |\pi_{T_x \mu_t} - \pi_{\mathbb{R}^n}| d\mu_t(x) + C_n E_t R_0^n.$$

Thus Hölder's inequality and the definition of β_t in (8.7) yield

$$\int_{B_{2\rho}^n(0)} |Df_t^j(\hat{x})| d\mathscr{L}^n(\hat{x}) \leq C_n \beta_t R_0^n + C_n E_t R_0^n,$$

where we used $16\rho = R_0$. Inserting this into (8.47) we obtain

$$\int_{B_\rho^n(0)} \Phi_1(t, \hat{y}) d\mathscr{L}^n(\hat{y}) \leq C_n p^{\frac{1}{2}} (\beta_t + E_t) R_0^n + C_n p^3 \eta R_0^{n-5} \tag{8.48}$$

for all $t \in I_\alpha$. Then integrating in time yields

$$\int_{I_\alpha} \int_{B_\rho^n(0)} \Phi_1(t, \hat{y}) d\mathscr{L}^n(\hat{y}) dt \leq C_n p^{\frac{1}{2}} \int_{I_\alpha} (\beta_t + E_t) dt R_0^n + C_n p^3 \eta |I_\alpha| R_0^{n-5},$$

where we used $16\rho = R_0$. Use again Hölder's inequality combined with (8.34), (8.37) and (8.38) to estimate

$$\begin{aligned}
&\int_{I_\alpha} \int_{B_\rho^n(0)} \Phi_1(t, \hat{y}) d\mathscr{L}^n(\hat{y}) dt \\
&\leq C_n p^{\frac{1}{2}} \left(\sqrt{|I_\alpha| \int_{I_\alpha} \beta_t^2 dt} + \int_{I_\alpha} E_t dt \right) R_0^n + C_n p^3 \eta |I_\alpha| R_0^{n-5} \\
&\leq C_n \Lambda_0 p^{\frac{1}{2}} \left(\eta + \eta^2 \right) R_0^{n+2} + C_n p^3 \eta R_0^{n-3} \leq C_n \Lambda_0 R_0^{n+2} p^{\frac{1}{2}} \eta
\end{aligned}$$

where we used $p = \delta^N R_0^2 \le R_0^2$ as $\delta \le 1$. Now use that $p^{\frac{1}{2}} = \delta^{\frac{N}{2}} R_0 \le \delta^{n+6} R_0$ for $N \ge 2n+12$. Then we conclude

$$(\delta R_0)^{-n-2} \int_{I_\alpha} \int_{B_\rho^n(0)} \Phi_1(t,\hat{y}) d\mathscr{L}^n(\hat{y}) dt \le C_n \delta^3 \eta R_0, \tag{8.49}$$

where we choose δ small depending on Λ_0.

For Φ_2 we can use the relation between Brakke flow and convolution of the height function of our family of varifolds with the heat kernel. In view of (8.21), $p = \delta^N R_0^2$, $q_0 = \delta^\epsilon R_0^2$, $\delta \le \delta_0$ and $16\rho = R_0$, we can estimate

$$\begin{aligned} p + t - s_0 &\le p + q_0 + (8\Lambda_0 + 1)\delta^2 R_0^2 \le C_n(\delta^N + \Lambda_0 \delta^2) R_0^2 + q_0 \\ &\le 2q_0 = 2\delta^\epsilon R_0^2 = 2^9 \delta_0^\epsilon \rho^2 \end{aligned} \tag{8.50}$$

for all $t \in (-\infty, \Lambda_0 \delta^2 R_0^2]$, for δ_0 small enough depending on ϵ and Λ_0. Note that by (8.21) and (8.35), estimate (8.50) holds for $t = s_0$ and all $t \in I_\alpha$. By definition of the heat kernel (see Definition 4.2) we can estimate

$$|\Psi_\rho(p + s - s_0, \hat{z})| \le C_n p^{-\frac{n}{2}} \tag{8.51}$$

for all $\hat{z} \in \mathbb{R}^n$ and all $s \in [s_0, \infty)$. Then by (8.19) we can estimate

$$\begin{aligned} &\Bigg| \int_{C_\rho(y)} x_{n+j} \Psi_\rho(p + s - s_0, \hat{x} - \hat{y}) d\mu_s(x) \\ &\quad - \int_{B_\rho^n(\hat{y})} f_s^j(\hat{x}) \Psi_\rho(p + s - s_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \Bigg| \\ &\le C_n \eta p^{-\frac{n}{2}} R_0^{n+1} E_s \le C_n \eta \delta^{-\frac{n}{2}N} R_0 E_s \le E_s R_0 \end{aligned} \tag{8.52}$$

for all $s \in [s_0, \infty)$ and all $y \in C_\rho^n(0)$, where we used (8.1), (8.15) and (8.51) to estimate the sup in (8.19). Also we used $p = \delta^N R_0^2$ and $\eta \le \eta_0$ for η_0 small depending on δ. By (8.21) and (8.35), we have $t \in [s_0, \infty)$ for all $t \in I_\alpha$. In view of (8.44) we then can estimate using (8.52) for $s = t$ and $s = s_0$

$$\begin{aligned} \Phi_2(t,\hat{y}) \le &\Bigg| \int_{B_{3R_0}(0)} x_{n+j}(\hat{x}) \Psi_\rho(p, \hat{x} - \hat{y}) d\mu_t(\hat{x}) \\ &\quad - \int_{B_{3R_0}(0)} x_{n+j}(\hat{x}) \Psi_\rho(p + t - s_0, \hat{x} - \hat{y}) d\mu_{s_0}(\hat{x}) \Bigg| \\ &+ (E_t + E_{s_0}) R_0 \end{aligned} \tag{8.53}$$

for all $t \in I_\alpha$ and $\hat{y} \in B_\rho^n(0)$. Here we used that by (8.1), $16\rho = R_0$ and $\eta \le 1$ we have

$$C_\rho(y) \cap \operatorname{spt}\Psi_\rho(p, \cdot - \hat{y}) \cap \operatorname{spt}\mu_t = B_{3R_0}(0) \cap \operatorname{spt}\Psi_\rho(p, \cdot - \hat{y}) \cap \operatorname{spt}\mu_t \tag{8.54}$$

for all $y \in \mathbb{R}^{n+k}$ and all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. Integrating (8.53) over $\hat{y}$, the difference on the right hand side can be estimated using Lemma 4.18 with $V := B_{3R_0}(0)$, $\gamma = 16\eta$ $P_0 = \epsilon^{-1}(n+5)$, $r = \rho$, $t_0 = s_0$ and $q = t - s_0$. Doing so we obtain

$$
\begin{aligned}
&\int_{B_\rho^n(0)} \Phi_2(t, \hat{y}) d\mathscr{L}^n(y) \\
&\leq C_n \eta \rho \Bigg[p^{-1} R_0^n \int_{s_0}^t \beta_s^2 ds + (\rho^{-2}(p + t - s_0))^{\frac{n+5}{\epsilon}} \rho^{-2} \int_{s_0}^t \mu_s\left(C_{2\rho}(0)\right) ds \\
&\qquad + \sup_{y \in B_\rho^n(0)} \rho^n \left| \int_{B_{3R_0}(0)} \Psi_\rho(p, \hat{x} - \hat{y}) d\mu_t(x) - 1 \right| \\
&\qquad + \sup_{y \in B_\rho^n(0)} \rho^n \left| \int_{B_{3R_0}(0)} \Psi_\rho(p + t - s_0, \hat{x} - \hat{y}) d\mu_{s_0}(x) - 1 \right| \Bigg] \\
&\quad + C_n (E_t + E_{s_0}) R_0^{n+1}
\end{aligned}
\tag{8.55}
$$

for all $t \in I_\alpha$, where we used the definition of β_s in (8.7) and $2\rho \leq R_0$ to estimate the tilt term. Also we used that by (8.50) we have $p + t - s_0 \leq C_n \delta^\epsilon \leq \kappa_0 \rho^2$ for δ_0 small depending on ϵ and κ_0. Note that by $16\rho = R_0$ and by choosing $\gamma = 16\eta$ assumption (8.1) implies (4.28). By (8.23) we have $s_0 \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$. For $t \in I_\alpha$ we see by (8.21), (8.35) and $q_0 = \delta^\epsilon R_0^2$ that $t \in [s_0, s_0 + R_0^2]$ for δ small depending on Λ_0. Thus we can use (8.11) to estimate

$$
p^{-1} R_0^n \int_{s_0}^t \beta_s^2 ds \leq C_n \Lambda_0 p^{-1} \eta^2 R_0^{n+2} \leq \delta^{n+5} R_0^n, \tag{8.56}
$$

for all $t \in I_\alpha$, where we used $p = \delta^N R_0^2$ and $\eta \leq \eta_0$ and we chose η_0 small depending on δ, Λ_0 and N. In view of (8.8), (8.2) and $16\rho = R_0$ we can estimate $\mu_s\left(C_{2\rho}(0)\right) \leq C_n R_0^n$. Combining this with (8.50) we obtain

$$
\begin{aligned}
&(\rho^{-2}(p + t - s_0))^{\frac{n+5}{\epsilon}} \rho^{-2} \int_{s_0}^t \mu_s\left(C_{2\rho}(0)\right) ds \\
&\leq C_n (\rho^{-2}(p + t - s_0))^{\frac{n+5}{\epsilon} + 1} R_0^n \leq C_n \delta^{n+5} R_0^n
\end{aligned}
\tag{8.57}
$$

for all $t \in I_\alpha$, where we used $16\rho = R_0$ and $\delta, \epsilon \leq 1$. By (8.19) we can estimate

$$
\begin{aligned}
&\left| \int_{C_\rho(y)} \Psi_\rho(p + t - s, \hat{x} - \hat{y}) d\mu_s(x) - 1 \right| \\
&\leq \left| \int_{B_\rho^n(\hat{y})} \Psi_\rho(p + t - s, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) - 1 \right| + C_n p^{-\frac{n}{2}} R_0^n E_s
\end{aligned}
$$

for all $s \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$, $t \in [s, \infty)$ and $y \in C_\rho^n(0)$, where we used (8.51) to estimate the sup in (8.19). By Proposition 4.10.3 with $P = \epsilon^{-1}(n+5)$ and $\hat{x}_0 = \hat{y}$ this can be estimated by

$$\begin{aligned}&\left| \int_{C_\rho(y)} \Psi_\rho(p+t-s, \hat{x}-\hat{y}) d\mu_s - 1 \right| \\ &\leq \left(\rho^{-2}(p+t-s)\right)^{\frac{n+5}{\epsilon}} + C_n p^{-\frac{n}{2}} R_0^n E_s \leq C_n \delta^{n+5} + \eta^{-1} E_s\end{aligned}$$

for all $t \in [s_0, \Lambda_0 \delta R_0^2]$, $s \in \{t, s_0\}$ and all $y \in C_\rho^n(0)$. Here we used that $p + t - s \leq C_n \delta^\epsilon \rho^2 \leq \kappa \rho^2$, for δ_0 small enough depending on ϵ and κ, which follows by (8.50) for $s = s_0$ and by $p = \delta^N R_0^2$ for $s = t$. Also we used $p = \delta^N R_0^2$ and $\eta \leq \eta_0$ for η_0 small depending on δ and N to estimate $C_n p^{-\frac{n}{2}} R_0^n \leq \eta^{-1}$. In particular for all $t \in I_\alpha$ we have

$$\left| \int_{B_{3R_0}(0)} \Psi_\rho(p+t-s_0, \hat{x}-\hat{y}) d\mu_{s_0} - 1 \right| \leq C_n \delta^{n+5} + \eta^{-1} E_{s_0} \tag{8.58}$$

$$\left| \int_{B_{3R_0}(0)} \Psi_\rho(p, \hat{x}-\hat{y}) d\mu_t - 1 \right| \leq C_n \delta^{n+5} + \eta^{-1} E_t \tag{8.59}$$

for all $\hat{y} \in B_\rho^n(0)$, where we used that by (8.21) and (8.35) we have $I_\alpha \subset [s_0, \Lambda_0 \delta^2 R_0^2]$ and that by (8.54) integrating this function over $B_{3R_0}(0)$ is the same as integrating over $C_\rho((\hat{y}, 0))$. Inserting (8.56), (8.57), (8.58) and (8.59) into (8.55) we obtain

$$\int_{B_\rho^n(0)} \Phi_2(t, y) d\mathscr{L}^n(y) \leq C_n \eta \delta^{n+5} R_0^{n+1} + C_n (E_t + E_{s_0}) R_0^{n+1} \tag{8.60}$$

for all $t \in I_\alpha$, where we used $16\rho = R_0$. The error term E_{s_0} is small due to (8.24) and integrating in time we can use (8.37) to estimate E_t, thus we have

$$\begin{aligned}\int_{I_\alpha} \int_{B_\rho^n(0)} \Phi_2(t, \hat{y}) d\mathscr{L}^n(y) dt &\leq C_n R_0^{n+1} \left(\eta \delta^{n+5} R_0^2 + \int_{I_\alpha} E_t dt + E_{s_0} R_0^2 \right) \\ &\leq C_n R_0^{n+3} \left(\eta \delta^{n+5} + \Lambda_0 \eta^2 + \delta^{-3} \eta^2 \right),\end{aligned}$$

where we used $|I_\alpha| \leq R_0^2$ by (8.38). Then as $\eta \leq \eta_0$ for η_0 small depending on δ and Λ_0, we can conclude

$$(\delta R_0)^{-n-2} \int_{I_\alpha} \int_{B_\rho^n(0)} \Phi_2(t, \hat{y}) d\mathscr{L}^n(y) dt \leq C_n \delta^3 \eta R_0. \tag{8.61}$$

For Φ_3 we want to use Lemma 4.7. To do so, we have to get rid of the cut-off part of our heat kernel. First, in view of (8.45), we can use (8.15) to estimate

$$\Phi_3(t,\hat{y}) \leq \eta R_0 \int_{B^n_\rho(0)} |\Psi_\rho(p+t-s_0,\hat{x}-\hat{y}) - \Psi_\rho(q_0,\hat{x}-\hat{y})|\, d\mathscr{L}^n(\hat{x}) \quad (8.62)$$

for all $t \in I_\alpha$ and all $\hat{y} \in B^n_\rho(0)$. Proposition 4.10.3 with $P = 3\epsilon^{-1}$ and $\hat{x}_0 = \hat{y}$ yields

$$\int_{B^n_\rho(0)} |\Psi(q,\hat{x}-\hat{y}) - \Psi_\rho(q,\hat{x}-\hat{y})|\, d\mathscr{L}^n(\hat{x}) \leq \left(\rho^{-2}q\right)^{\frac{3}{\epsilon}} \leq C_n\delta^3$$

for all $q \in (0, 2q_0]$ and all $\hat{y} \in B^n_\rho(0)$, where we used $q_0 = \delta^\epsilon R_0^2 \leq C_n\delta^\epsilon\rho^2$, so as $\delta \leq \delta_0$ we have $\rho^{-2}q \leq \kappa$ for δ_0 small enough depending on κ and ϵ. By (8.50) we have $p+t-s_0 \leq 2q_0$, thus (8.62) becomes

$$\begin{aligned}\Phi_3(t,\hat{y}) \leq \eta R_0 &\int_{B^n_\rho(0)} |\Psi(p+t-s_0,\hat{x}-\hat{y}) - \Psi(q_0,\hat{x}-\hat{y})|\, d\mathscr{L}^n(\hat{x}) \\ &+ C_n\delta^3\eta R_0\end{aligned}$$

for all $t \in I_\alpha$ and all $\hat{y} \in B^n_\rho(0)$. Now we can use Lemma 4.7 with $q = p+t-s_0$. Note that by (8.21) and (8.35), we can estimate $p+t-s_0 > q_0$ for every $t \in I_\alpha$. Then with Lemma 4.7 we obtain

$$\begin{aligned}\Phi_3(t,\hat{y}) &\leq \left(\log\left(1+q_0^{-1}(p+t-s_0-q_0)\right) + C_n\delta^3\right)\eta R_0 \\ &\leq \left(q_0^{-1}(p+t-s_0-q_0) + C_n\delta^3\right)\eta R_0\end{aligned}$$

for all $t \in I_\alpha$ and all $\hat{y} \in B^n_\rho(0)$. Now use (8.21) and (8.35), to estimate $p+t-s_0-q_0 \leq C_n\Lambda_0\delta^2R_0^2$ for all $t \in I_\alpha$. Then as $q_0 = \delta^\epsilon R_0^2$, we arrive at

$$\Phi_3(t,\hat{y}) \leq C_n\Lambda_0\delta^{2-\epsilon}\eta R_0$$

for all $t \in I_\alpha$ and all $\hat{y} \in B^n_\rho(0)$. By (8.38) we have $|I_\alpha| \leq 4r_0^2$, thus we conclude

$$r_0^{-n-2}\int_{I_\alpha}\int_{B^n_{4r_0}(0)} \Phi_3(t,\hat{y})d\mathscr{L}^n(\hat{y})\, dt \leq C_n\Lambda_0\delta^{2-\epsilon}\eta R_0. \quad (8.63)$$

Here we also used (8.36) to see that $B^n_{4r_0}(0) \subset B^n_\rho(0)$.

For Φ_4 we can use that μ_{s_0} becomes flat by convolution with the heat kernel with high parameter. By definition of t_j and z_{n+j} (see (8.25) and (8.26)) we have

$$\begin{aligned}&\int_{B^n_\rho(\hat{y})} f^j_{s_0}\Psi_\rho(q_0,\hat{x}-\hat{y})d\mathscr{L}^n(\hat{x}) - t_j(\hat{y}) - z_{n+j} \\ &= \int_{B^n_\rho(\hat{y})} f^j_{s_0}\left(\Psi_\rho(q_0,\hat{x}-\hat{y}) - \Psi_\rho(q_0,\hat{x}) + \hat{y}\cdot D\Psi_\rho(q_0,\hat{x})\right) d\mathscr{L}^n(\hat{x})\end{aligned} \quad (8.64)$$

for all $\hat{y} \in B^n_\rho(0)$. Due to Taylor's expansion theorem the last integrand can be estimated in terms of the second derivative

$$\begin{aligned}
&|\Psi_\rho(q_0, \hat{x} - \hat{y}) - \Psi_\rho(q_0, \hat{x}) + \hat{y} \cdot D\Psi_\rho(q_0, \hat{x})| \\
&\leq \left| \int_0^1 (1-\theta) \left(\sum_{i,j=1}^n \hat{y}_i \hat{y}_j \frac{\partial^2 \Psi_\rho}{\partial x_i \partial x_j}(q_0, \hat{x} - \theta \hat{y}) \right) d\theta \right| \\
&\leq |\hat{y}|^2 \int_0^1 \left| D^2 \Psi_\rho(q_0, \hat{x} - \theta \hat{y}) \right| d\theta.
\end{aligned}$$

Then in view of (8.46) we can use (8.64) and (8.15) to estimate

$$\Phi_4(q_0, \hat{y}) \leq \eta R_0 |\hat{y}|^2 \int_{B^n_\rho(\hat{y})} \int_0^1 \left| D^2 \Psi_\rho(q_0, \hat{x} - \theta \hat{y}) \right| d\theta d\mathscr{L}^n(\hat{x}) \tag{8.65}$$

for all $\hat{y} \in B^n_\rho(0)$. In order to estimate the second derivatives, we first get rid of the cut-off part of our heat kernel. Use Lemma 4.9 with $\mu = \mathscr{L}^n$ and $P_1 = 1$ to obtain

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left| D^2 \Psi_\rho(q_0, \hat{x} - \hat{x}_0) \right| d\mathscr{L}^n(\hat{x}) \\
&\leq \int_{B^n_\rho(\hat{x}_0)} \left| D^2 \Psi(q_0, \hat{x} - \hat{x}_0) \right| d\mathscr{L}^n(\hat{x}) + C_n q_0 \rho^{-4}
\end{aligned}$$

for all $\hat{x}_0 \in \mathbb{R}^n$. Here we used that by $q_0 = 2^8 \delta^\epsilon \rho^2$, and $\delta \leq \delta_0$ we can estimate $\rho^{-2} q_0 \leq \kappa_1$, for δ_0 small enough depending on ϵ and κ_1. Also we used $\mathrm{spt}\Psi_\rho(q_0, \cdot - \hat{x}_0) \subset B^n_\rho(\hat{x}_0)$. Then we can differentiate the heat kernel and use Lemma 4.6 to estimate

$$\begin{aligned}
&\int_{\mathbb{R}^n} \left| D^2 \Psi_\rho(q_0, \hat{x} - \hat{x}_0) \right| d\mathscr{L}^n(\hat{x}) \\
&\leq \int_{B^n_\rho(\hat{x}_0)} \left((4q_0)^{-2} |\hat{x} - \hat{x}_0|^2 + q_0^{-1} \right) \Psi(q_0, \hat{x} - \hat{x}_0) d\mathscr{L}^n(\hat{x}) + C_n q_0 \rho^{-4} \\
&\leq C_n q_0^{-1} + C_n q_0 \rho^{-4} \leq C_n \delta^{-\epsilon} R_0^{-2}
\end{aligned}$$

for all $\hat{x}_0 \in \mathbb{R}^n$, where we used $q_0 = \delta^\epsilon R_0^2$ and $\delta \leq 1$. Inserting this into (8.65) with $\hat{x}_0 = \theta \hat{y}$ we obtain

$$\Phi_4(q_0, \hat{y}) \leq C_n \eta R_0^{-1} |\hat{y}|^2 \delta^{-\epsilon} \tag{8.66}$$

for all $\hat{y} \in B^n_\rho(0)$. For $\hat{y} \in B^n_{4r_0}(0)$ we have $|\hat{y}|^2 \leq C_n \Lambda_0 \delta^2 R_0^2$, as $r_0 = \sqrt{2\Lambda_0} \delta R_0$. Then we can conclude

$$r_0^{-n-2} \int_{I_\alpha} \int_{B^n_{4r_0}(0)} \Phi_4(q_0, \hat{y}) d\mathscr{L}^n(\hat{y})\, dt \leq C_n \Lambda_0 \delta^{2-\epsilon} \eta R_0. \tag{8.67}$$

Here we used (8.36) to see that $B^n_{4r_0}(0) \subset B^n_\rho(0)$ and also we used that by (8.38) we have $|I_\alpha| \leq 4r_0^2$.

In view of (8.42) we can insert (8.49),(8.61), (8.63)and (8.67) into (8.41) to obtain

$$S \leq C_n \Lambda_0 \delta^{2-\epsilon} \eta R_0,$$

where we used $\Lambda_0 \geq 1$ and $B^n_{4r_0}(0) \subset B^n_\rho(0)$, by (8.36). By (8.30) and (8.32) this yields

$$\sup_{[-\Lambda_0\delta^2 R_0^2, \Lambda_0 \delta^2 R_0^2]} \sup_{x \in spt\mu_t \cap B_{16\delta R_0}(z)} |x_{n+j} - t_j(\hat{x}) - z_{n+j}| \leq C_n \Lambda_0 \delta^{2-\epsilon} \eta R_0$$

for every $1 \leq j \leq k$. Note that by definition of T (see (8.25)), we have

$$|\pi_T^\perp(x)| = \inf_{v \in T} |x - v| = \inf_{\hat{w} \in \mathbb{R}^n} \left| x - \left(\hat{w} + \sum_{j=1}^k t_j(\hat{w}) \mathbf{e}_{n+j} \right) \right|$$

for all $x \in \mathbb{R}^{n+k}$. Thus we can can estimate for all $x \in \mathrm{spt}\mu_t \cap B_{16\delta R_0}(z)$ and all $t \in [-\Lambda_0 \delta^2 R_0^2, \Lambda_0 \delta^2 R_0^2]$

$$\begin{aligned} |\pi_T^\perp(x - z)| &\leq \left| x - z - \left(\hat{x} + \sum_{j=1}^k t_j(\hat{x}) \mathbf{e}_{n+j} \right) \right| \\ &= \left| \sum_{j=1}^k \left(x_{n+j} - t_j(\hat{x}) - z_{n+j} \right) \mathbf{e}_{n+j} \right| \leq C_n \Lambda_0 \delta^{2-\epsilon} \eta R_0, \end{aligned}$$

where we used $\hat{z} = 0$. As $\delta \leq \delta_0$ and δ_0 small depending on ϵ and Λ_0 we conclude

$$|\pi_T^\perp(x - z)| \leq \delta^{2-2\epsilon} \eta R_0,$$

for all $x \in \mathrm{spt}\mu_t \cap B_{16\delta R_0}(z)$ and all $t \in [-\Lambda_0 \delta^2 R_0^2, \Lambda_0 \delta^2 R_0^2]$. This establishes (8.3) with ϵ replaced by 2ϵ, which completes the result. Next we give an alternative method how to establish (8.3), which is closer to Brakke's original work.

2. With Lemma 5.14:
In this case we need the extra assumption that μ_t is a Brakke flow in $B_{\delta^\epsilon \eta^{-1} R_0}(0)$. Also (8.1) has to be changed to

$$\mathrm{spt}\mu_t \cap C_{3R_0}(0) \subset \{x \in C_{3R_0}(0),\ |\pi_{\mathbb{R}^k}(x)| \leq \eta R_0\} \tag{8.68}$$

for all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. Note that these assumptions are only stronger and all the previous statements remain true. It turns out that we have to choose T and z a bit different, depending on a time t_1 defined below:

Let $\delta_0 \le 2^{-2}$, then $4\delta^2 R_0^2 \le 2^{-2} R_0$, so $-4\Lambda_0\delta^2 R_0^2 \in [-2^{-2}\Lambda_0 R_0^2, 2^{-2}\Lambda_0 R_0^2]$. Then by inequality (8.11) and $\delta \le 1$ there exists $t_1 \in [-5\Lambda_0\delta^2 R_0^2, -3\Lambda_0\delta^2 R_0^2]$ with

$$\alpha_{t_1}^2 + \beta_{t_1}^2 \le C_n \Lambda_0 \delta^{-2} \eta^2. \tag{8.69}$$

For $\delta \le \delta_0 \le 2^{-2}$ we have $t_1 \in [-2^{-1}\Lambda_0 R_0^2, 2^{-1}\Lambda_0 R_0^2]$, in particular a Lipschitz approximation exists at time t_1. By (8.69) we have $\alpha_{t_1} \le 1$, as $\eta \le \eta_0$ for η_0 small depending on δ and Λ_0. Thus

$$E_{t_1} \le \delta^{-3}\eta^2, \tag{8.70}$$

where we also used $\delta \le \delta_0$ and chose δ_0 small depending on Λ_0. Here E_t is the error term defined in (8.18).

This lets us define $\bar{T}$ and $\bar{z}$ as follows: Set

$$\bar{q} := p + t_1 - s_0.$$

For $j \in \{1, \dots, k\}$ and $\hat{y} \in \mathbb{R}^n$ set

$$\bar{t}_j(\hat{y}) := \sum_{i=1}^n \hat{y}_i \bar{t}_{ij}, \ \ \bar{t}_{ij} := -\int_{B_\rho^n(0)} f_{s_0} \cdot \mathbf{e}_j \frac{\partial \Psi_\rho}{\partial x_i}(\bar{q}, \hat{x}) d\mathscr{L}^n(\hat{x}), \tag{8.71}$$

$$\bar{z}_{n+j} := \int_{B_\rho^n(0)} f_{s_0} \cdot \mathbf{e}_j \Psi_\rho(\bar{q}, \hat{x}) d\mathscr{L}^n(\hat{x}). \tag{8.72}$$

Then set $\bar{T} := \{\hat{x} + \sum_{j=1}^k \bar{t}_j(\hat{x})\mathbf{e}_{n+j}, \ \hat{x} \in \mathbb{R}^n\}$ and $\bar{z} := (0, \bar{z}_{n+1}, \dots, \bar{z}_{n+k})$. Note that by (8.21), $p = \delta^N R_0^2$, $q_0 = \delta^\epsilon R_0^2$ and $\delta \le \delta_0$ for δ_0 small depending on Λ_0 we have

$$\bar{q} \le (\delta^N + (4\Lambda_0 + 1)\delta^2) R_0^2 + q_0 \le 2q_0$$
$$\bar{q} \ge (\delta^N + 2\Lambda_0) R_0^2 + q_0 \ge q_0,$$

so basically all estimates above hold with q_0 replaced by $\bar{q}$, except some used to estimate Φ_3. In particular (8.27), (8.28), (8.29) and (8.36) still hold for T and z replaced by $\bar{T}$ and $\bar{z}$.

To use Lemma 5.14 we need to verify a height bound in the tilted cylinder

$$C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z}) = \{x \in \mathbb{R}^{n+k} : |\pi_{\bar{T}}(x - \bar{z})| \le \sqrt{32n\Lambda_0}\delta R_0\}. \tag{8.73}$$

For $x \in \mathbb{R}^{n+k}$ we can estimate

$$\begin{aligned}|\hat{x}| &= |x - \pi_{\mathbb{R}^k}(x)| = |x - \bar{z} - \pi_{\mathbb{R}^k}(x - \bar{z})| \\ &= |\pi_{\bar{T}}(x - \bar{z}) + \pi_{\bar{T}}^{\perp}(x - \bar{z}) - \pi_{\mathbb{R}^k}(x - \bar{z})| \\ &\leq \left|\pi_{\bar{T}}^{\perp} - \pi_{\mathbb{R}^k}\right| |x - \bar{z}| + |\pi_{\bar{T}}(x - \bar{z})|\end{aligned}$$

thus in view of Remark A.7.2 we can use (8.28) and (8.29) to obtain

$$|\hat{x}| \leq \delta^{-\epsilon}\eta|x| + C_n\delta^{-\epsilon}\eta^2 R_0 + |\pi_{\bar{T}}(x - \bar{z})|.$$

In particular this yields

$$C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z}) \cap B_{\delta^{\epsilon}\eta^{-1}R_0}(0) \subset C_{2R_0}(\bar{z}),$$

where we estimated $1 + C_n\delta^{-\epsilon}\eta^2 + \sqrt{32n\Lambda_0}\delta \leq 2$, by $\delta \leq \delta_0$, $\eta \leq \eta_0$ and for δ_0 small depending on Λ_0, as well as η_0 small depending on δ and ϵ. Then (8.68) yields

$$\mathrm{spt}\mu_t \cap C^{\bar{T}}_{32\sqrt{n\Lambda_0}\delta R_0}(\bar{z}) \subset \{x \in C_{3R_0}(0),\ |\pi_{\mathbb{R}^k}(x - \bar{z})| \leq \eta R_0\} \tag{8.74}$$

for all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$. By Remark A.7.3 we have

$$|\pi_{\bar{T}}^{\perp}(y)| \leq 2|\pi_{\mathbb{R}^k}(\pi_{\bar{T}}^{\perp}(y))|$$

for all $y \in \mathbb{R}^{n+k}$, where we used $|\pi_{\bar{T}}^{\perp} - \pi_{\mathbb{R}^k}| = |\pi_{\bar{T}} - \pi_{\mathbb{R}^n}| \leq \delta^{-\epsilon}\eta \leq 2^{-1}$, due to (8.28) and as $\eta \leq \eta_0$ for η_0 small depending on δ and ϵ. Then estimate for $y \in \mathbb{R}^{n+k}$

$$\begin{aligned}|\pi_{\bar{T}}^{\perp}(y)| \leq 2|\pi_{\mathbb{R}^k}(\pi_{\bar{T}}^{\perp}(y))| &= 2|\pi_{\mathbb{R}^k}(y) - \pi_{\mathbb{R}^k}(\pi_{\bar{T}}(y))| \\ &= 2|\pi_{\mathbb{R}^k}(y) - \pi_{\bar{T}}(y) + \pi_{\mathbb{R}^n}(\pi_{\bar{T}}(y))| \\ &\leq 2\left(|\pi_{\mathbb{R}^k}(y)| + |\pi_{\bar{T}} - \pi_{\mathbb{R}^n}||\pi_{\bar{T}}(y)|\right).\end{aligned}$$

In particular for $x \in C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z})$ and in view of (8.28) this yields

$$|\pi_{\bar{T}}^{\perp}(x - \bar{z})| \leq 2|\pi_{\mathbb{R}^k}(x - \bar{z})| + 2\delta^{-\epsilon}\eta\sqrt{32n\Lambda_0}\delta R_0 \leq 2|\pi_{\mathbb{R}^k}(x - \bar{z})| + \eta R_0,$$

where we used $\delta \leq \delta_0$ and δ_0 small depending on ϵ and Λ_0. Thus (8.74) implies

$$\mathrm{spt}\mu_t \cap C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z}) \subset \left\{x \in C_{3R_0}(0),\ |\pi_{\bar{T}}^{\perp}(x - \bar{z})| \leq 3\eta R_0\right\} \tag{8.75}$$

for all $t \in [-\Lambda_0 R_0^2, \Lambda_0 R_0^2]$.

To use Lemma 5.14, we need a suitable orthonormal basis $(\nu_j)_{1\le j\le k}$ of $\bar{T}^\perp$. Set

$$N_j := \mathbf{e}_{n+j} - \sum_{i=1}^n \bar{t}_{ij}\mathbf{e}_i, \tag{8.76}$$

$j \in \{1,\dots,k\}$, these N_j form a basis of $\bar{T}^\perp$ and will be later used to calculate. By (8.27) we have $|N_j - \mathbf{e}_{n+j}| \le \delta^{-\epsilon}\eta$ for all $j \in \{1,\dots,k\}$, as $\delta \le \delta_0$ for δ_0 small depending on ϵ. Then by Lemma A.10.1 with $B = \mathbb{R}^n$ there exists an orthonormal basis $(\nu_j)_{1\le j\le k}$ of $\bar{T}^\perp$ such that

$$\max\{|N_j - \mathbf{e}_{n+j}|, |\nu_j - \mathbf{e}_{n+j}|, |\nu_j - N_j|\} \le C_n\delta^{-\epsilon}\eta \tag{8.77}$$

for all $j \in \{1,\dots,k\}$. We want to use Lemma 5.14 with $\tau = \delta^{1-2\epsilon}\eta$, $h = 3\eta$, $t_0 = -\Lambda_0\delta^2R_0^2$ and $R = r_0 = \sqrt{2\Lambda_0}\delta R_0$. Fix a $j \in \{1,\dots,k\}$. In order to verify (5.43), we have to show that

$$\bar{S} := (\delta R_0)^{-n}\int_{C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z})} \left\{|\nu_j\cdot(x-\bar{z})| - \frac{\eta\,|\pi_T(x-\bar{z})|^2}{2\delta^{2\epsilon}R_0}\right\}_+ d\mu_{t_1} \tag{8.78}$$

is smaller than $c_0\delta^{4-6\epsilon}\eta R_0$ for some $c_0 \in (0,1)$, where we used $\Lambda_0 \ge 1$.

In view of (8.75) and (8.36), we see that

$$\mathrm{spt}\mu_t \cap C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z}) \subset \mathrm{spt}\mu_t \cap B_{8\sqrt{n\Lambda_0}\delta R_0}(0) \subset C_\rho(0). \tag{8.79}$$

For the first inclusion of (8.79) we estimated for $x \in \mathrm{spt}\mu_t \cap C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z})$

$$\begin{aligned}|x| &\le |x-\bar{z}| + C_n\eta R_0 \le |\pi_{\bar{T}}(x-\bar{z})| + |\pi^\perp_{\bar{T}}(x-\bar{z})| + C_n\eta R_0\\ &\le \sqrt{32n\Lambda_0}\delta R_0 + C_n\eta R_0 \le 8\sqrt{n\Lambda_0}\delta R_0,\end{aligned}$$

where we used (8.75) and $|\bar{z}| \le C_n\eta R_0$ by (8.29). Also we used $\Lambda_0 \ge 1$ and $\eta \le \eta_0$ for η_0 small depending on δ. The second inclusion of (8.79) then follows from (8.36).

For $x \in \mathrm{spt}\mu_{t_1} \cap C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z})$ and $j \in \{1,\dots,k\}$ in view of the definition of N_j in (8.76), we can use (8.77) and (8.75), to estimate

$$\begin{aligned}&|\nu_j\cdot(x-\bar{z}) - (x_{n+j} - \bar{t}_j(\hat{x}) - \bar{z}_{n+j})| = |\nu_j\cdot(x-\bar{z}) - N_j\cdot(x-z)|\\ &\le |\nu_j - N_j|\left|\pi^\perp_{\bar{T}}(x-\bar{z})\right| \le C_n\delta^{-\epsilon}\eta^2R_0.\end{aligned}$$

Also by (8.28),(8.79) and $\pi_{\mathbb{R}^n}(\bar{z}) = 0$, we have

$$\begin{aligned}&\left||\pi_{\bar{T}}(x-\bar{z})|^2 - |\hat{x}|^2\right|\\ &= |(|\pi_{\bar{T}}(x-\bar{z})| + |\pi_{\mathbb{R}^n}(x-\bar{z})|)\,(|\pi_{\bar{T}}(x-\bar{z})| - |\pi_{\mathbb{R}^n}(x-\bar{z})|)|\\ &\le 2\,|\pi_{\bar{T}} - \pi_{\mathbb{R}^n}|\,|x-\bar{z}|^2 \le C_n\Lambda_0\delta^{2-\epsilon}\eta R_0^2,\end{aligned}$$

thus we have

$$\begin{aligned}
&\left\{|\nu_j \cdot (x-\bar{z})| - \delta^{-2\epsilon}\eta \left|\pi_{\bar{T}}(x-z)\right|^2 R_0^{-1}\right\}_+ \\
&\le \left\{|x_{n+j} - \bar{t}_j(\hat{x}) - \bar{z}_{n+j}| - \delta^{-2\epsilon}\eta|\hat{x}|^2 R_0^{-1}\right\}_+ + C_n\delta^{-\epsilon}\eta^2 R_0,
\end{aligned} \tag{8.80}$$

for all $x \in \mathrm{spt}\mu_{t_1} \cap C^{\bar{T}}_{\sqrt{32n\Lambda_0}\delta R_0}(\bar{z})$ and all $j \in \{1,\dots,k\}$, where we used $\delta \le \delta_0$, for δ_0 small depending on ϵ and Λ_0.

Similar to (8.30) set this time

$$\bar{\Phi}(x) := |x_{n+j} - \bar{t}_j(\hat{x}) - \bar{z}_{n+j}|\,.$$

As (8.27) and (8.29) hold for t_{ij} and z_{n+j} replaced by $\bar{t}_{ij}$ and $\bar{z}_{n+j}$, we have that (8.31) holds for Φ replaced by $\bar{\Phi}$. By (8.80) and (8.79), the $\tilde{S}$ from (8.78) can be estimated by

$$\begin{aligned}
\bar{S} \le (\delta R_0)^{-n} &\int_{B_{8\sqrt{n\Lambda_0}\delta R_0}(0)} \left\{\bar{\Phi}(x) - \frac{\eta|\hat{x}|^2}{2\delta^{2\epsilon}R_0}\right\}_+ d\mu_{t_1}(x) \\
&+ C_n\delta^{-\epsilon}\eta^2 R_0.
\end{aligned} \tag{8.81}$$

As we have a Lipschitz approximation at time t_1, we can use (8.19), to write this as an integral over $\mathbb{R}^n$. To do so, we have to bound the integrand. By (8.31) we obtain

$$\bar{\Phi}(x) - \frac{\eta|\hat{x}|^2}{2\delta^{2\epsilon}R_0} \le C_n\left(\delta^\epsilon\eta + \Lambda_0\delta^{2-2\epsilon}\eta\right)R_0 \le C_n\delta^\epsilon\eta R_0$$

for all $x \in \mathrm{spt}\mu_{t_1} \cap B_{8\sqrt{n\Lambda_0}\delta R_0}(0)$, where we used $\delta \le \delta_0$ for δ_0 small depending on Λ_0 and ϵ. Then combining (8.19) with (8.81) yields

$$\begin{aligned}
\bar{S} \le (\delta R_0)^{-n} &\int_{B^n_{8\sqrt{n\Lambda_0}\delta R_0}(0)} \left\{\bar{\Phi}(F_{t_1}(\hat{y})) - \frac{\eta|\hat{y}|^2}{2\delta^{2\epsilon}R_0}\right\}_+ \mathscr{L}^n(\hat{y}) \\
&+ C_n\left(\delta^\epsilon\eta(\delta R_0)^{-n}\rho^n E_{t_1} + \Lambda_0\delta^{-\epsilon}\eta^2\right)R_0,
\end{aligned}$$

where we estimated $8\sqrt{n\Lambda_0}\delta R_0 \le 2^{-4}R_0 = \rho$, as $\delta \le \delta_0$ for δ_0 small depending on Λ_0. Now by (8.70), (8.79), $\eta \le \eta_0$ and η_0 small depending on δ we obtain

$$\bar{S} \le (\delta R_0)^{-n} \int_{B^n_\rho(0)} \left\{\bar{\Phi}(F_{t_1}(\hat{y})) - \frac{\eta|\hat{y}|^2}{2\delta^{2\epsilon}R_0}\right\}_+ \mathscr{L}^n(\hat{y}) + C_n\Lambda_0\delta^{-\epsilon}\eta^2 R_0. \tag{8.82}$$

Like in (8.42) we can use the triangle inequality to estimate

$$\left\{\bar{\Phi}(F_{t_1}(\hat{y}) - \frac{\eta|\hat{y}|^2}{2\delta^{2\epsilon}R_0}\right\}_+ \le \Phi_1(t_1,\hat{y}) + \Phi_2(t_1,\hat{y}) + \left\{\bar{\Phi}_4(\hat{y}) - \frac{\eta|\hat{y}|^2}{2\delta^{2\epsilon}R_0}\right\}_+ \tag{8.83}$$

for all $\hat{y} \in B^n_\rho(0)$. Here $\Phi_1(t_1, \hat{y})$, $\Phi_2(t_1, \hat{y})$ are defined as in (8.43), (8.44), where now we only look at the fixed time t_1, i.e.

$$\Phi_1(t_1, \hat{y}) := \left| f^j_t(\hat{y}) - \int_{B_\rho(\hat{y})} f^j_{t_1} \Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \right| \tag{8.84}$$

$$\begin{aligned} \Phi_2(t_1, \hat{y}) := \Bigg| &\int_{B_\rho(\hat{y})} f^j_{t_1}(\hat{y}) \Psi_\rho(p, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) \\ &- \int_{B_\rho(\hat{y})} f^j_{s_0}(\hat{y}) \Psi_\rho(p + t_1 - s_0, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x})(x) \Bigg|. \end{aligned} \tag{8.85}$$

The quantity $\bar{\Phi}_4$ is defined by

$$\bar{\Phi}_4(\hat{y}) := \left| \int_{B_\rho(\hat{y})} f^j_{s_0} \Psi_\rho(\bar{q}, \hat{x} - \hat{y}) d\mathscr{L}^n(\hat{x}) - \bar{t}_j(\hat{y}) - \bar{z}_{n+j} \right|. \tag{8.86}$$

To see (8.83) we used $\bar{q} = p + t_1 - s_0$. Note that as $t_1 \in [-5\Lambda_0 \delta^2 R_0^2, -3\Lambda_0 \delta^2 R_0^2]$ and by (8.69) we have $t_1 \in I_\alpha$ defined in (8.35). Thus by (8.48)

$$\int_{B^n_\rho(0)} \Phi_1(t_1, \hat{y}) d\mathscr{L}^n(\hat{y}) dt \le C_n p^{\frac{1}{2}} (\beta_{t_1} + E_{t_1}) R_0^n + C_n p^3 \eta R_0^{n+5}.$$

Then we can use (8.69) and (8.70) to estimate

$$\begin{aligned} \int_{B^n_\rho(0)} \Phi_1(t_1, \hat{y}) d\mathscr{L}^n(\hat{y}) dt &\le C_n p^{\frac{1}{2}} (\Lambda_0 \delta^{-1} \eta + \delta^{-3} \eta^2) R_0^n + C_n p^3 \eta R_0^{n+5} \\ &\le C_n p^{\frac{1}{2}} \Lambda_0 \delta^{-3} \eta R_0^{n+1} \le C_n \Lambda_0 \delta^{\frac{N}{2} - 3} \eta R_0^{n+1}, \end{aligned}$$

where we used $p = \delta^N R_0^2$ and $\delta \le 1$. Then as $\delta \le \delta_0$ for δ_0 small depending on Λ_0 and for $N \ge 2n + 18$, we conclude

$$\int_{B^n_\rho(0)} \Phi_1(t_1, \hat{y}) d\mathscr{L}^n(\hat{y}) dt \le \delta^{n+5} \eta R_0^{n+1}. \tag{8.87}$$

Also we have by (8.60)

$$\int_{B^n_\rho(0)} \Phi_2(t_1, y) d\mathscr{L}^n(y) \le C_n \eta \delta^{n+5} R_0^{n+1} + C_n (E_{t_1} + E_{s_0}) R_0^{n+1}$$

so with (8.70) and (8.24) we obtain

$$\int_{B^n_\rho(0)} \Phi_2(t_1, y) d\mathscr{L}^n(y) \le C_n \left(\delta^{n+5} + \delta^{-3} \eta \right) \eta R_0^{n+1} \le C_n \delta^{n+5} \eta R_0^{n+1}, \tag{8.88}$$

where we used $\eta \leq \eta_0$ for η_0 small depending on δ.

By definitions (8.71) and (8.72), we have

$$\begin{aligned}\bar{\Phi}_4(\hat{y}) &= \int_{B^n_\rho(\hat{y})} f^j_{s_0}\Psi_\rho(\bar{q},\hat{x}-\hat{y}) - t_j(\hat{y}) - z_{n+j}\, d\mathscr{L}^n(\hat{x}) \\ &= \int_{B^n_\rho(\hat{y})} f^j_{s_0}\left(\Psi_\rho(\bar{q},\hat{x}-\hat{y}) - \Psi_\rho(\bar{q},\hat{x}) + \hat{y}\cdot D\Psi_\rho(\bar{q},\hat{x})\right) d\mathscr{L}^n(\hat{x}).\end{aligned}$$

Similar calculations as those which led to (8.66) will yield

$$\bar{\Phi}_4(\hat{y}) \leq C_n \eta R_0^{-1}|\hat{y}|^2\delta^{-\epsilon}$$

for all $\hat{y} \in B^n_\rho(0)$. Here we used $\bar{q} \in [q_0, 2q_0]$, so $\bar{q} = C_n\delta^\epsilon R_0^2$. Thus we see

$$\left\{\bar{\Phi}_4(\hat{y}) - \frac{\eta|\hat{y}|^2}{2\delta^{2\epsilon}R}\right\}_+ = 0 \tag{8.89}$$

for all $\hat{y} \in B^n_\rho(0)$, where we used $\delta \leq \delta_0$ for δ_0 small depending on ϵ. In view of (8.83), we can insert (8.87),(8.88) and (8.89) into (8.82), to obtain

$$\tilde{S} \leq C_n\delta^5\eta R_0. \tag{8.90}$$

Now we can use Lemma 5.14 with $\tau = \delta^{1-2\epsilon}\eta$, $h = 3\eta$, $t_0 = -\Lambda_0\delta^2R_0^2$, $y_0 = z$ and $R = \sqrt{2\Lambda_0}\delta R_0$. By (8.29) we have $|\bar{z}| \leq C_n\eta R_0$, so $B_{\delta^{2\epsilon}(3\eta)^{-1}R_0}(\bar{z}) \subset\subset B_{\delta^{2\epsilon}\eta^{-1}R_0}(0)$, as $\delta \leq 1$, $\eta \leq \eta_0$ and η_0 small. Inclusion (8.75) directly establishes (5.42). In view of (8.78), inequality (8.90) implies (5.43). Then Lemma 5.14 yields

$$\mathrm{spt}\mu_t \cap B_{\sqrt{2\Lambda_0}\delta R_0}(\bar{z}) \subset \{x \in \mathbb{R}^{n+k}, |\pi^\perp_{\bar{T}}(x-\bar{z})| \leq C_n\sqrt{\Lambda_0}\delta^{2-2\epsilon}\eta R_0\}.$$

For $\Lambda_0 \geq 2^7$, we see $B_{\sqrt{2\Lambda_0}\delta R_0}(\bar{z}) \supset B_{16\delta R_0}(\bar{z})$. Also as $\delta \leq \delta_0$ for δ_0 small depending on Λ_0, we can estimate $C_n\sqrt{\Lambda_0}\delta^{2-2\epsilon}\eta R_0 \leq \delta^{1-3\epsilon}\eta r_0$. So we established the result with ϵ replaced by 3ϵ.

Note that because of the different assumption (8.68) we actually have

$$\mathrm{spt}\mu_t \cap C^{\bar{T}}_{15\delta R_0}(\bar{z}) = \mathrm{spt}\mu_t \cap B_{16\delta R_0}(\bar{z}) \tag{8.91}$$

thus we can exchange the set $B_{16\delta R_0}(\bar{z})$ in (8.3) by $C^{\bar{T}}_{15\delta R_0}(\bar{z})$. This is very important, if one wants to iterate this version of the Lemma, because this now implies the different condition (8.68) for the next step. To see inclusion (8.91) consider $x \in \mathrm{spt}\mu_t \cap C^{\bar{T}}_{15\delta R_0}(\bar{z})$. As $\sqrt{n\Lambda_0} \geq 1$, we can use (8.75), to calculate

$$|x-\bar{z}| = |\pi_{\bar{T}}(x-\bar{z}) + \pi^\perp_{\bar{T}}(x-\bar{z})| \leq 15\delta R_0 + 3\eta R_0 \leq 16\delta R_0,$$

where we used $\eta \leq \eta_0$ for $\eta_0 \leq 3^{-1}\delta$. $\square$

8.2 Remark. Brakke claims the unit density hypothesis is needed to obtain the above result. He says it is crucial to estimate the difference between f_t and the convolution $f_t * \psi$. In our proof this is done in the calculation for $\Phi_1(t,y)$. However, we do not appear to need the unit density hypothesis for our proof.

To iterate Lemma 8.1 properly one wants the centre points to have a fixed projection onto $\mathbb{R}^n$. The form in which Lemma 8.1 is used is the following:

8.3 Lemma. *For every $\lambda_0 \in (0, 2^{-n-5}]$ there exists a $\Lambda_0 \in (1,\infty)$ such that, for every $\alpha \in (0,1)$ there exists a $\beta_0 \in (0,1)$ such that, for every $\beta \in (0,\beta_0]$ there exists a $\gamma_0 \in (0,1)$ such that, for all $\rho \in (0,\infty)$, $t_0 \in \mathbb{R}$, $\hat{y}_0 \in \mathbb{R}^n$, $\gamma \in [0,\gamma_0]$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_0-\Lambda_0\rho^2, t_0+\Lambda_0\rho^2]}$ be a Brakke flow in U. Let A be an n-dimensional subspace of $\mathbb{R}^{n+k}$ with $|\pi_A - \pi_{\mathbb{R}^n}| \le \frac{1}{4}$. Let $a \in \{\hat{y}_0\} \times \mathbb{R}^k$ with $B_{3\rho}(a) \subset U$. Suppose*

$$\mathrm{spt}\mu_t \cap C_{3\rho}(a) \subset \left\{x \in B_{5\rho}(a),\ |\pi_A^\perp(x-a)| \le \gamma\rho\right\}, \tag{8.92}$$

$$\rho^{-n}\mu_t\left(B_{(1+\lambda_0)\rho}(a)\right) \le \frac{3}{2}\omega_n\ , \quad \rho^{-n}\mu_t\left(B_\rho(a)\right) \ge \frac{1}{2}\omega_n \tag{8.93}$$

for all $t \in [t_0 - \Lambda_0\rho^2, t_0 + \Lambda_0\rho^2]$. Then there exist an n-dimensional subspace A^* of $\mathbb{R}^{n+k}$ with $|\pi_A - \pi_{A^*}| \le \delta^{\alpha-1}\gamma$ and a point $a^* \in \{\hat{y}_0\}\times\mathbb{R}^k$ with $|a^*-a| \le 3\sqrt{k}\gamma R_0$ such that

$$\mathrm{spt}\mu_t \cap C_{3\beta\rho}(a^*) \subset \left\{x \in B_{5\beta\rho}(a^*),\ |\pi_{A^*}^\perp(x-a^*)| \le \gamma\beta^{1+\alpha}\rho\right\}, \tag{8.94}$$

$$(\beta\rho)^{-n}\mu_t\left(B_{(1+\lambda_0)\beta\rho}(a^*)\right) \le \frac{3}{2}\omega_n\ , \quad (\beta\rho)^{-n}\mu_t\left(B_{\beta\rho}(a^*)\right) \ge \frac{1}{2}\omega_n \tag{8.95}$$

for all $t \in [t_0 - \Lambda_0\beta^2\rho^2, t_0 + \Lambda_0\beta^2\rho^2]$.

Note that the cylinders in (8.92) and (8.94)are ordinary cylinders with respect to $\mathbb{R}^n$, whereas the height bounds are with respect to the n-dimensional subspaces A and A^*.

Proof. Fix a rotation $S \in SO(n+k)$ with $S(\mathbb{R}^n) = A$, $S(\mathbb{R}^k) = A^\perp$. Consider the Brakke flow $(\tilde{\mu}_t)_{t\in[-\Lambda_0\rho^2,\Lambda_0\rho^2]}$ in $B_{3\rho}(0)$ defined by

$$\tilde{\mu}_t(B) := \mu_{t+t_0}(S(B)+a) \llcorner B_{3\rho}(0),$$

for all $B \subset \mathbb{R}^{n+k}$, where $S(B)+a = \{S(b)+a, b \in B\}$. Here we used $B_{3\rho}(a) \subset U$, to see that $S(U+a) \supset B_{3\rho}(0)$. This is necessary for $(\tilde{\mu}_t)$ to be a Brakke flow in $B_{3\rho}(0)$. Then $(\tilde{\mu}_t)$ satisfies

$$\mathrm{spt}\tilde{\mu}_t \cap B_{3\rho}(0) \subset \{x \in B_{5\rho}(0),\ |\pi_{\mathbb{R}^k}(x)| \le \gamma\rho\}, \tag{8.96}$$

$$\rho^{-n}\tilde{\mu}_t\left(B_{(1+\lambda_0)\rho}(0)\right) \le \frac{3}{2}\omega_n\ , \quad \rho^{-n}\tilde{\mu}_t\left(B_\rho(0)\right) \ge \frac{1}{2}\omega_n \tag{8.97}$$

for all $t \in [-\Lambda_0\rho^2, \Lambda_0\rho^2]$. Note that we used $B_{3\rho}(0) \subset C_{3\rho}(0)$. Thus we can use Lemma 8.1 with $\epsilon = 1-\alpha$, $\delta = \beta$, $\eta_0 = \gamma_0$, $\eta = \gamma$ and $R_0 = \rho$. Lemma 8.1 then yields an n-dimensional subspace T with $|\pi_{\mathbb{R}^n} - \pi_T| \leq \beta^{\alpha-1}\gamma$ and a point $z \in \{0\}^n$ with $|z| \leq \sqrt{k}\gamma\rho$ such that

$$\mathrm{spt}\tilde{\mu}_t \cap B_{16\beta\rho}(z) \subset \left\{x \in B_{16\beta\rho}(z),\ |\pi_T^\perp(x-z)| \leq \delta^{1+\alpha}\gamma\rho\right\}, \quad (8.98)$$

$$(\beta\rho)^{-n}\tilde{\mu}_t\left(B_{(1+2\lambda_0)\beta\rho}(z)\right) \leq \frac{5}{4}\omega_n\ , \quad (\beta\rho)^{-n}\tilde{\mu}_t\left(B_{(1-2\lambda_0)\beta\rho}(z)\right) \geq \frac{3}{4}\omega_n \quad (8.99)$$

for all $t \in [-\Lambda_0\beta^2\rho^2, \Lambda_0\beta^2\rho^2]$. Now set $A^* := S(T)$, then by Remark A.8.3

$$|\pi_A - \pi_{A^*}| = \left|\pi_{S^{-1}(A)} - \pi_{S^{-1}(A^*)}\right| = |\pi_{\mathbb{R}^n} - \pi_T| \leq \beta^{\alpha-1}\gamma.$$

Set $b = S(z) + a$, then $|b-a| = |S(z)| \leq \sqrt{k}\gamma\rho$. Statements (8.98) and (8.99) imply for (μ_t)

$$\mathrm{spt}\mu_t \cap B_{16\beta\rho}(b) \subset \left\{x \in B_{16\beta\rho}(b),\ |\pi_{A^*}^\perp(x-b)| \leq \delta^{1+\alpha}\gamma\rho\right\}, \quad (8.100)$$

$$(\beta\rho)^{-n}\mu_t\left(B_{(1+2\lambda_0)\beta\rho}(b)\right) \leq \frac{5}{4}\omega_n\ , \quad (\beta\rho)^{-n}\mu_t\left(B_{(1-2\lambda_0)\beta\rho}(b)\right) \geq \frac{3}{4}\omega_n \quad (8.101)$$

for all $t \in [t_0 - 2\Lambda_0\beta^2\rho^2, t_0 + 2\Lambda_0\beta^2\rho^2]$.

Note that b may not be in $\{\hat{y}\} \times \mathbb{R}^k$, but shifting b a bit yields a suitable a^* as we will see below. We can estimate

$$|\pi_{\mathbb{R}^n} - \pi_{A^*}| \leq |\pi_{\mathbb{R}^n} - \pi_A| + |\pi_A - \pi_{A^*}| \leq 2^{-2} + \beta^{\alpha-1}\gamma \leq 2^{-1} \quad (8.102)$$

as $\gamma \leq \gamma_0$ for γ_0 small depending on α and β. Thus using Proposition A.9.4, there exists a unique intersection point in $\{\hat{y}\} \times \mathbb{R}^k \cap A^* + b$, so we can define a^* by

$$\{a^*\} = \{\hat{y}\} \times \mathbb{R}^k \cap A^* + b.$$

Combining $|a-b| \leq \sqrt{k}\gamma\rho$ and $\pi_{\mathbb{R}^n}(a) = \pi_{\mathbb{R}^n}(a^*) = \hat{y}$ with $a^* \in A^* + b$ and (8.102), we can estimate

$$\begin{aligned}
|a^* - b| &= |\pi_{\mathbb{R}^k}(a^* - b) + \pi_{\mathbb{R}^n}(a^* - b)| \\
&\leq \left|\left(\pi_{\mathbb{R}^k} - \pi_{A^*}^\perp\right)(a^* - b)\right| + |\pi_{\mathbb{R}^n}(a - b)| \\
&\leq \frac{|a^* - b|}{2} + \sqrt{k}\gamma\rho.
\end{aligned}$$

So we have $|a^* - b| \leq 2\sqrt{k}\gamma\rho$, which yields $|a^* - a| \leq |b - a| + 2\sqrt{k}\gamma\rho \leq 3\sqrt{k}\gamma\rho$. Also we see $|a^* - b| \leq \lambda_0\beta\rho$, as $\gamma \leq \gamma_0$ for γ_0 small depending on λ_0 and β, thus (8.101) implies (8.95).

Now it remains to show that the height bound (8.100) actually holds for the cylinder. Let $x \in \mathrm{spt}\mu_t \cap C_{3\beta\rho}(a^*)$, by $|a^* - a| \le 3\sqrt{k}\gamma\rho$ we can estimate

$$|x - a^*| \le |x - a| + |a - a^*| \le |\pi_A(x - a)| + \left|\pi_A^\perp(x - a)\right| + 3\sqrt{k}\gamma\rho.$$

Use (8.92) and Remark A.7.3 to obtain

$$|x - a^*| \le \frac{4}{3}\,|\pi_{\mathbb{R}^n}(\pi_A(x - a))| + \gamma\rho + 3\sqrt{k}\gamma\rho.$$

Using $\hat{a} = \hat{y}$ and again (8.92) yields

$$\begin{aligned}|x - a^*| &\le \frac{4}{3}\,|\pi_{\mathbb{R}^n}(x - a)| + \frac{4}{3}\left|\pi_{\mathbb{R}^n}(\pi_A^\perp(x - a))\right| + (1 + 3\sqrt{k})\gamma\rho \\ &\le \frac{4}{3}\,|\hat{x} - \hat{y}| + 2\gamma\rho + (1 + 3\sqrt{k})\gamma\rho,\end{aligned}$$

so as $x \in C_{3\beta\rho}(a^*)$ we can conclude

$$|x - a^*| \le 4\beta\rho + (3 + 3\sqrt{k})\gamma\rho \le 5\beta\rho$$

where we used $\gamma \le \gamma_0$ for γ_0 small depending on β. Thus we showed

$$\mathrm{spt}\mu_t \cap C_{3\beta\rho}(a^*) \subset B_{5\beta\rho}(a^*) \subset B_{16\beta\rho}(a^*),$$

so (8.100) implies (8.94), which completes the statement. Actually we never make full use of the factor 16 in (8.100) and could change it to a factor 5 in Lemma 8.1. $\square$

For a point $\hat{y} \in \mathbb{R}^n$ Lemma 8.3 can now be used to obtain a sequence of points in $\{\hat{y}\} \times \mathbb{R}^k$ and subspaces of $\mathbb{R}^{n+k}$ which converge to a point in $\mathrm{spt}\mu_t$ and its tangent space. This will define a function f with $\mathrm{graph}(f) = \mathrm{spt}\mu_t$ and by the type of convergence we obtain $C^{1,\alpha}$-regularity for f.

8.4 Theorem (Local Regularity Theorem,[B, 6.10]). *For every $\lambda \in (0,1]$ and $\alpha \in (0,1)$ there exists a $\Lambda \in (1,\infty)$ such that for every $K \in [1,\infty)$ there exists a $h_0 \in (0,1)$ such that for all $R \in (0,\infty)$, $h \in [0,h_0]$, $t_1, t_2 \in \mathbb{R}^{n+k}$ with $t_2 - t_1 > 2\Lambda R^2$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in $B_{(1+2\lambda)R}(0)$ with*

$$\mathrm{spt}\mu_t \subset \left\{x \in \overline{B_{(1+2\lambda)R}(0)}, |\pi_{\mathbb{R}^k}(x)| \le hR\right\}, \tag{8.103}$$

$$R^{-n}\mu_t\left(B_{(1+2\lambda)R}(0)\right) \le K \tag{8.104}$$

for all $t \in [t_1, t_2]$ *and*

$$R^{-n}\mu_{t_1}\left(B_{(1+2\lambda)R}(0)\right) \le (2-\lambda)\omega_n\ , \quad R^{-n}\mu_{t_2}\left(B_R(0)\right) \ge \lambda\omega_n. \tag{8.105}$$

Then there exists a smooth function $f : [t_1 + \Lambda R^2, t_2 - \Lambda R^2] \times B^n_{\lambda R}(0) \to \mathbb{R}^k$ *such that for* $M_t = \mathrm{graph}(f(t,\cdot))$ *we have*

$$(\mathrm{spt}\mu_t \cap C_{\lambda R}(0)) = M_t. \tag{8.106}$$

Furthermore the estimates

$$|f(t,\hat{x}) - f(s,\hat{y})| \le \Lambda h\left(R^{-\alpha}|t-s|^{\frac{1+\alpha}{2}} + |\hat{x}-\hat{y}|\right), \tag{8.107}$$

$$\left|\pi_{T_xM_t} - \pi_{T_yM_s}\right| \le \Lambda h R^{-\alpha}\left(|t-s|^{\frac{\alpha}{2}} \ + |\hat{x}-\hat{y}|^{\alpha}\right) \tag{8.108}$$

hold for all $t, s \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ *and all* $\hat{x}, \hat{y} \in B^n_{\lambda R}(0)$*, where* $x = (\hat{x}, f(t,\hat{x}))$*,* $y = (\hat{y}, f(t,\hat{y}))$*.*

8.5 Remark. Here we will only prove the $C^{1,\alpha}$- regularity. For the higher regularity we refer to [T] and [LSU].

Proof. We want to apply Lemma 8.3. By assumption we have a height bound and by Lemma 7.8 we obtain area ratio bounds as well. For given $R \in (0,\infty)$ set $\rho_0 := 2^{-2}R$. Using Lemma 7.8 with $\eta = h$, $\lambda_1 = \lambda$, $M = K$ and $R_0 = R$ yields

$$\rho_0^{-n}\mu_t\left(B_{(1+2^{-n-2}\lambda)\rho_0}(y)\right) \le \frac{3}{2}\omega_n\ , \quad \rho_0^{-n}\mu_t\left(B_{\rho_0}(y)\right) \ge \frac{1}{2}\omega_n \tag{8.109}$$

for all $y \in B^n_{\lambda R_0}(0) \times \{0\}^k$ for all $t \in [t_1 + 2^{-1}\Lambda R^2, t_2 - 2^{-1}\Lambda R^2]$. Here we had to chose $h_0 \le \eta_1$ and $\Lambda \ge 2\Lambda_1$. Note that then

$$[t_1 + \Lambda_1 R^2, t_2 - \Lambda_1 R^2] \supset [t_1 + 2^{-1}\Lambda R^2, t_2 - 2^{-1}\Lambda R^2].$$

Now temporarily fix $y_0 \in B^n_{\lambda R}(0) \times \{0\}^k$ and $t_0 \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$. For given $\Lambda_0 \in (1,\infty)$ we can choose $\Lambda \ge 2\Lambda_0$ and obtain from (8.109)

$$\rho_0^{-n}\mu_t\left(B_{(1+2^{-n-5}\lambda)\rho_0}(y_0)\right) \le \frac{3}{2}\omega_n\ , \quad \rho_0^{-n}\mu_t\left(B_{\rho_0}(y_0)\right) \ge \frac{1}{2}\omega_n \tag{8.110}$$

for all $t \in [t_0 - \Lambda_0\rho_0^2, t_0 + \Lambda_0\rho_0^2]$. Here we calculated as $\rho_0 \le R$

$$\begin{aligned} t_0 + \Lambda_0\rho_0^2 &\le t_2 - \Lambda R^2 + \Lambda_0 R^2 \le t_2 - 2^{-1}\Lambda R^2 \\ t_0 - \Lambda_0\rho_0^2 &\ge t_1 + \Lambda R^2 - \Lambda_0 R^2 \ge t_1 + 2^{-1}\Lambda R^2. \end{aligned}$$

Consider $\alpha \in (0,1)$ and $\Lambda_0 \in (1,\infty)$, where Λ_0 may depend on λ and α. Let $\delta \in (0,1)$ be a variable we will choose later depending on α and Λ_0. Note that our choice of Λ will depend on δ. Let $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ then set

$$\begin{aligned} \rho_m := \delta^m \rho_0, \quad \eta_m := 4\delta^{\alpha m} h, \quad \tau_m := \Lambda_0 \delta^{2m} \rho_0^2, \\ z_0(t_0, \hat{y}_0) := y_0, \quad T_0(t_0, \hat{y}_0) := \mathbb{R}^n. \end{aligned} \tag{8.111}$$

Iterating Lemma 8.3 with $\lambda_0 = 2^{-n-5}\lambda$ and $\beta = \delta$ then yields $\Lambda_0 = \Lambda_0(\alpha, \lambda) \in (1,\infty)$, $\beta_0 = \beta_0(\alpha,\lambda) \in (0,1)$ and $\gamma_0 = \gamma_0(\delta,\alpha,\lambda) \in (0,1)$ such that for $\Lambda \geq 2\Lambda_0$, $\delta \leq \beta_0$ and $h_0 \leq \gamma_0$ the following holds: For every $m \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ there exist an n-dimensional subspace $T_m = T_m(t_0, \hat{y}_0)$ of $\mathbb{R}^{n+k}$ with

$$|\pi_{T_m} - \pi_{\mathbb{R}^n}| \leq 8\delta^{-1} h \leq 2^{-3} \tag{8.112}$$

and a point $z_m = z_m(t_0, \hat{y}_0)$ in $\{\hat{y}_0\} \times \mathbb{R}^k$ with

$$B_{3\rho_m}(z_m) \subset B_R(0) \tag{8.113}$$

such that

$$(\mathrm{spt}\mu_t \cap C_{3\rho_m}(z_m)) \subset \left\{x \in B_{5\rho_m}(z_m), \left|\pi_{T_m}^{\perp}(x - z_m)\right| \leq \eta_m \rho_m\right\}, \tag{8.114}$$

$$\rho_m^{-n} \mu_t \left(B_{(1+\lambda_0)\rho_m}(z_m)\right) \leq \frac{3}{2}\omega_n, \quad \rho_m^{-n} \mu_t \left(B_{\rho_m}(z_m)\right) \geq \frac{1}{2}\omega_n \tag{8.115}$$

for all $t \in [t_0 - \tau_m, t_0 + \tau_m]$, and such that the following recursions hold

$$\left|\pi_{T_{m+1}} - \pi_{T_m}\right| \leq \delta^{-\alpha}\eta_m, \quad |z_{m+1} - z_m| \leq 3\sqrt{k}\eta_m\rho_m. \tag{8.116}$$

To prove this use induction. For $m = 0$ by definitions in (8.111) statement (8.115) directly follows from (8.110). For (8.114) use (8.103) and

$$\mathrm{spt}\mu_t \cap C_{3\rho_0}(y_0) \subset B_{4\rho_0}(y_0),$$

as $4\rho_0 = R$ and $h \leq h_0$ for $h_0 \leq 2^{-2}$.

Now assume for all $l \in \{0, \ldots, m\}$ we can find T_l, z_l such that (8.112)-(8.116) hold. Then use Lemma 8.3 with $\lambda_0 = 2^{-n-5}\lambda$, $a = z_m$, $A = T_m$, $\beta = \delta$, $\gamma = \eta_m$ and $\rho = \rho_m$ to obtain $z_{m+1} := a^*$ and $T_{m+1} := A^*$ such that (8.114), (8.115) and (8.116) hold for $m+1$. By (8.116) we have

$$|\pi_{T_m+1} - \pi_{\mathbb{R}^n}| \leq \sum_{l=0}^{m+1} \left|\pi_{T_{l+1}} - \pi_{T_l}\right| \leq 4\delta^{-1} h \sum_{l=0}^{m+1} \delta^{\alpha l} \leq 8\delta^{-1} h \leq 2^{-3}, \tag{8.117}$$

where we used $\delta^\alpha \leq 2^{-1}$, for δ small depending on α and $h \leq h_0$ for h_0 small depending on δ. This implies (8.112) for $m+1$. By (8.116) we also have

$$|z_{m+1} - z_m| \leq 3\sqrt{k}\eta_m \rho_m = 12\sqrt{k}\delta^{\alpha m} h \rho_m \leq 2^{-1}\rho_m,$$

as $\delta \leq 1$ and $h \leq h_0$ for h_0 small. For $\delta \leq 2^{-1}$ we see $\rho_{m+1} \leq 2^{-1}\rho_m$ so $B_{3\rho_{m+1}}(z_{m+1}) \subset B_{3\rho_m}(z_m)$, which implies (8.113) Thus we established (8.112)- (8.116) for all $m \in \mathbb{N}_0$.

As $\hat{y}_0 \in B^n_{\lambda R}(0)$ and $t_0 \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ where arbitrary, we can now define $f : [t_1 + \Lambda R^2, t_2 - \Lambda R^2] \times B^n_{\lambda R}(0) \to \mathbb{R}^k$ by

$$\begin{aligned} f(t,\hat{y}) &:= \lim_{m\to\infty} \pi_{\mathbb{R}^k}(z_m(t,\hat{y})) \\ F(t,\hat{y}) &:= (\hat{y}, f(t,\hat{y})) \end{aligned} \tag{8.118}$$

for all $(t,\hat{y}) \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2] \times B^n_{\lambda R}(0)$ which is well defined, as by (8.116) the z_m form a Cauchy sequence, thus the limes exists. In particular we can estimate

$$|f(t,\hat{y}) - \pi_{\mathbb{R}^k}(z_m(t,\hat{y}))| \leq 3\sqrt{k}\sum_{l=m}^{\infty} \eta_l \rho_l \leq C_n \delta^{(1+\alpha)m} h \rho_0 \sum_{l=0}^{\infty} \delta^{\alpha l}$$

and as $\delta^\alpha \leq 2^{-1}$, for δ small depending on α, this yields

$$|f(t,\hat{y}) - \pi_{\mathbb{R}^k}(z_m(t,\hat{y}))| \leq \delta^{(1+\alpha)m} h \rho_0 \tag{8.119}$$

for all $\hat{y} \in B^n_{\lambda R}(0)$ and all $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$.

We want to show that for every $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ the graph of $f(t,\cdot)$ is indeed equal to the varifold μ_t inside $C_{\lambda R}(0)$. First suppose $y \in \mathrm{spt}\mu_t \cap C_{\lambda R}(0)$, then for every $m \in \mathbb{N}$ it is obviously true that $y \in C_{3\rho_m}(\hat{y})$. By (8.114) this implies $y \in B_{5\rho_m}(z_m(t,\hat{y}))$ for all $m \in \mathbb{N}$. In view of definition (8.118) letting $m \to \infty$ then yields $y = (\hat{y}, f(t,\hat{y}))$. Here we used $\lim_{m\to\infty} \rho_m = \delta^m \rho_0 = 0$.

Second suppose $\hat{y} \in B^n_{\lambda R}(0)$ and let $\epsilon > 0$ be arbitrary. By using again that $\lim_{m\to\infty} \rho_m = 0$ combined with definition (8.118) there exists $m \in \mathbb{N}$ such that $\rho_m \leq \frac{\epsilon}{2}$ and $|f(t,\hat{y}) - \pi_{\mathbb{R}^k}(z_m(t,\hat{y}))| \leq \frac{\epsilon}{2}$. Then we can use the density ratio bound (8.115) at $z_m(t,\hat{y})$, to estimate

$$\mu_t\left(B_\epsilon((\hat{y}, f(t,\hat{y}))\right) \geq \mu_t(B_{\rho_m}(z_m(t,\hat{y})) \geq \frac{\omega_n}{2}\rho_m^n > 0$$

and as ϵ was arbitrary, we conclude $(\hat{y}, f(t,\hat{y})) \in \mathrm{spt}\mu_t$. Thus we established (8.106).

Now that we have established that $\mathrm{spt}\mu_t$ is a graph, we can attack continuity. First we want to show that f is $C^{0,\frac{1+\alpha}{2}}$ in the t variable. It suffices to prove this locally. Fix $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$, $\hat{y} \in B^n_{\lambda R}(0)$ and consider the iteration $T_m = T_m(t, \hat{y}), z_m = z_m(t, \hat{y})$ for $m \in \mathbb{N}_0$. For any $s \in (t - \rho_0^2, t + \rho_0^2) \cap [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$, there exists an $m \in \mathbb{N}$ such that

$$\delta^{2(m+1)}\rho_0^2 \leq |t - s| \leq \delta^{2m}\rho_0^2 = \tau_m.$$

In view of (8.112) we can use Remark A.7.3 to estimate

$$\begin{aligned}|f(t,\hat{y}) - f(s,\hat{y})| &= |\pi_{\mathbb{R}^k}(F(t,\hat{y}) - F(s,\hat{y}))| \leq 2\left|\pi^{\perp}_{T_m}(F(t,\hat{y}) - F(s,\hat{y}))\right| \\ &\leq 2\left(\left|\pi^{\perp}_{T_m}(F(t,\hat{y}) - z_m)\right| + \left|\pi^{\perp}_{T_m}(F(s,\hat{y}) - z_m)\right|\right).\end{aligned}$$

By choice of m and (8.114), we see that both $F(t,\hat{y})$ and $F(s,\hat{y})$ are contained in $\left\{x \in \mathbb{R}^{n+k} : \left|\pi^{\perp}_{T_m}(x - z_m)\right| \leq \eta_m\rho_m\right\}$. Note that we can use (8.114) here, because (8.106) is already established. Thus we obtain

$$|f(t,\hat{y}) - f(s,\hat{y})| \leq 4\eta_m\rho_m \leq C_n\delta^{\alpha m + m}h\rho_0 \leq C_n\delta^{-\alpha-1}h\rho_0^{-\alpha}|t-s|^{\frac{1+\alpha}{2}}.$$

For $\Lambda \geq C_n\delta^{-\alpha-1}$ and as $4\rho_0 = R$ this establishes one part of (8.107).

Next we want to show that f is differentiable. Fix $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$, $\hat{y} \in B^n_{\lambda R}(0)$ and consider the iteration $T_m = T_m(t,\hat{y}), z_m = z_m(t,\hat{y})$ for $m \in \mathbb{N}_0$. For $i \in \{1, \dots, n\}$ let p_i^m be the unique intersection point in $\{\mathbf{e}_i\} \times \mathbb{R}^k \cap T_m$, such a unique point exists by Proposition A.9.4, as by (8.112) we have $\|\pi_{T_m} - \pi_{\mathbb{R}^n}\|_{op} < 1$. We claim

$$\frac{\partial}{\partial y_i} f(t,\hat{y}) = g_i(t,\hat{y}) := \lim_{m\to\infty} \pi_{\mathbb{R}^k}(p_i^m). \tag{8.120}$$

To show (8.120) we first have to verify that the limit on the right hand side exists. We know $p_i^m \in T_m$ and $\pi_{\mathbb{R}^n}(p_i^m) = \mathbf{e}_i$. In view of (8.112), we can use Remark A.7.3 to bound $|p_i^m|$ by

$$|p_i^m| = |\pi_{T_m}(p_i^m)| \leq 2|\pi_{\mathbb{R}^n}(\pi_{T_m}(p_i^m))| = 2|\pi_{\mathbb{R}^n}(p_i^m)| = 2. \tag{8.121}$$

Using $\pi_{\mathbb{R}^n}(p_i^m) = \mathbf{e}_i$ and again Remark A.7.3, we then obtain

$$\begin{aligned}|p_i^{m_1} - p_i^{m_2}| &= \left|\pi_{T_{m_1}}(p_i^{m_1}) - \pi_{T_{m_2}}(p_i^{m_2})\right| \\ &\leq \left|\pi_{T_{m_1}}(p_i^{m_1} - p_i^{m_2})\right| + \left|\pi_{T_{m_1}} - \pi_{T_{m_2}}\right||p_i^{m_2}| \\ &= \left|\left(\pi_{T_{m_1}} - \pi_{\mathbb{R}^n}\right)(p_i^{m_1} - p_i^{m_2})\right| + \sum_{l=m_1}^{m_2-1}\left|\pi_{T_{l+1}} - \pi_{T_l}\right||p_i^{m_2}| \\ &\leq \frac{1}{8}|p_i^{m_1} - p_i^{m_2}| + 2\sum_{l=m_1}^{m_2-1}\left|\pi_{T_{l+1}} - \pi_{T_l}\right|.\end{aligned}$$

Thus with (8.116) we conclude

$$|p_i^{m_1} - p_i^{m_2}| \le C_n \delta^{-\alpha} \sum_{l=m_1}^{m_2-1} \eta_l \le C_n \delta^{(m_1-1)\alpha} h \sum_{l=0}^{\infty} \delta^{\alpha l} \le C_n \delta^{(m_1-1)\alpha} h \quad (8.122)$$

for all $m_1, m_2 \in \mathbb{N}_0$ with $m_1 \le m_2$, where we used $\delta^\alpha \le 2^{-1}$, for δ small depending on α. Thus p_i^m is a Cauchy sequence, so g_i is well defined. Furthermore we can estimate the rate of convergence by

$$|g_i(t,y) - \pi_{\mathbb{R}^k}(p_i^m)| \le C_n \delta^{(m-1)\alpha} h. \quad (8.123)$$

To show that g_i actually is the derivative in $\mathbf{e}_i$-direction take some arbitrary small $\xi_0 \in (0, \rho_0 - |y|)$. For any $\xi \in (-\xi_0, +\xi_0) \setminus \{0\}$ there exists $m \in \mathbb{N}_0$ such that

$$\delta^{m+1}\rho_0 \le |\xi| \le \delta^m \rho_0 = \rho_m.$$

We want to show that the difference quotient of f converges to g. Note that

$$\pi_{\mathbb{R}^n}(z_m + \xi p_i^m) = \hat{y} + \xi \mathbf{e}_i = \pi_{\mathbb{R}^n}(F(t, \hat{y} + \xi \mathbf{e}_i)).$$

Using (8.123) and (8.119) we obtain

$$\begin{aligned}
&\left|\xi^{-1}\left(f(t,\hat{y}+\xi\mathbf{e}_i) - f(t,\hat{y})\right) - g_i(t,\hat{y})\right| \\
&\le |\xi|^{-1}\left|f(t,\hat{y}+\xi\mathbf{e}_i) - \pi_{\mathbb{R}^k}(z_m) - \xi\pi_{\mathbb{R}^k}(p_i^m)\right| + C_n\delta^{(m-1)\alpha}h \\
&= |\xi|^{-1}\left|\pi_{\mathbb{R}^k}\left(F(t,\hat{y}+\xi\mathbf{e}_i) - z_m - \xi p_i^m\right)\right| + C_n\delta^{(m-1)\alpha}h,
\end{aligned}$$

where we used $\delta^\alpha \le 1$ and $\delta^{m+1}\rho_0 \le |\xi|$. In view of (8.112) we can use Remark A.7.3 to estimate further

$$\begin{aligned}
&\left|\xi^{-1}\left(f(t,\hat{y}+\xi\mathbf{e}_i) - f(t,\hat{y})\right) - g_i(t,\hat{y})\right| \\
&\le 2|\xi|^{-1}\left|\pi_{T_m}^{\perp}\left(F(t,\hat{y}+\xi\mathbf{e}_i) - z_m - \xi p_i^m\right)\right| + C_n\delta^{(m-1)\alpha}h.
\end{aligned}$$

By choice of m we have $|\xi| \le \rho_m$, so in view of (8.114)

$$F(t,\hat{y}+\xi\mathbf{e}_i) \in \left\{x \in \mathbb{R}^{n+k} : \left|\pi_{T_m}^{\perp}(x - z_m)\right| \le \eta_m\rho_m\right\}.$$

Also we know $p_i^m \in T_m$. Thus we can conclude

$$\begin{aligned}
&\left|\xi^{-1}\left(f(t,\hat{y}+\xi\mathbf{e}_i) - f(t,\hat{y})\right) - g_i(t,\hat{y})\right| \\
&\le 2|\xi|^{-1}\eta_m\rho_m + C_n\delta^{(m-1)\alpha}h \le C_n\delta^{-2}h|\xi|^\alpha\rho_0^{-\alpha},
\end{aligned}$$

where we used $\delta, \alpha \leq 1$, $\delta^{m+1}\rho_0 \leq |\xi|$ and $\eta_m \rho_m = 4\delta^{m+\alpha m} h \rho_0$. For $\xi_0 \to 0$ the last expression becomes arbitrary small, which shows that $g_i(t, \hat{y})$ is indeed the derivative of $f(t, \hat{y})$ in i-direction, which establishes (8.120).

To get a Lipschitz bound for f note that, as $T_0 = \mathbb{R}^n$, we have $p_i^0 = \mathbf{e}_i$. Then by (8.122), we see

$$|p_i^m - \mathbf{e}_i| \leq C_n \delta^{-\alpha} h \tag{8.124}$$

for all $i \in \{1, \ldots, n\}$ and all $m \in \mathbb{N}_0$. Hence (8.120) implies

$$\frac{\partial}{\partial y_i} f(t, \hat{y}) \leq C_n \delta^{-\alpha} h$$

for all $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ and $\hat{y} \in B^n_{\lambda R}(0)$. Thus for $\Lambda \geq \delta^{-\alpha}$ this establishes the second part of (8.107).

Fix again arbitrary $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ and $\hat{y} \in B^n_{\lambda R}(0)$. We want to conclude that $T_m(t, \hat{y})$ converges to $T_{F(t,\hat{y})} M_t$. Note that by differentiability of f the tangent space $T_{F(t,\hat{y})} M_t$ exists. In view of equality (8.120) we have

$$\frac{\partial}{\partial y_i} F(t, \hat{y}) = \lim_{m \to \infty} p_i^m.$$

In particular by (8.122)

$$\left| \frac{\partial}{\partial y_i} F(t, \hat{y}) - p_i^m \right| \leq C_n \delta^{(m-1)\alpha} h \leq 1 \tag{8.125}$$

for all $m \in \mathbb{N}_0$ and all $i \in \{1, \ldots, n\}$, where we chose h small. By definition of the p_i^m, we have $p_i^m \in T_m(t, \hat{y})$ for all $i \in \{1, \ldots, n\}$ and all $m \in \mathbb{N}_0$. Then with inequality (8.124) we can estimate

$$|p_i^m \cdot p_j^m - \delta_{ij}| = |(p_i^m - \mathbf{e}_i) \cdot p_j^m + \mathbf{e}_i \cdot (p_j^m - \mathbf{e}_j)| \leq C_n \delta^{-\alpha} h \leq C_{A.10}^{-1}$$

for all $i, j \in \{1, \ldots, n\}$ and all $m \in \mathbb{N}_0$, where we used $|p_j^m| \leq 2$ and we chose h small depending on δ and $C_{A.10}$. Here $C_{A.10}$ denotes the constant from Lemma A.10. As the $\frac{\partial}{\partial y_i} F(t, \hat{y})_{1 \leq i \leq n}$ form a basis of $T_{F(t,\hat{y})} M_t$, we can now use Lemma A.10.1 and estimate (8.125) to obtain

$$\left| \pi_{T_{F(t,\hat{y})} M_t} - \pi_{T_m(t,\hat{y})} \right| \leq C_n \delta^{(m-1)\alpha} h \tag{8.126}$$

for all $m \in \mathbb{N}_0$, all $t \in [t_1 + \Lambda R^2, t_2 - \Lambda R^2]$ and all $\hat{y} \in B^n_{\lambda R}(0)$. In particular we can conclude that $T_{F(t,\hat{y})} M_t = \lim_{m \to \infty} T_m(t, \hat{y})$.

To obtain $C^{1,\alpha}$-regularity, fix arbitrary $t,s \in [t_1+\Lambda R^2, t_2-\Lambda R^2]$ and $\hat{x},\hat{y} \in B^n_{\lambda R}(0)$. We want to show

$$\left|\pi_{T_{F(t,\hat{x})}\mu_t} - \pi_{T_{F(s,\hat{y})}\mu_s}\right| \leq C_n \delta^{-2\alpha} h \rho_0^{-\alpha} \left(|t-s| + |\hat{x}-\hat{y}|^2\right)^{\frac{\alpha}{2}}. \tag{8.127}$$

If $|t-s|+|\hat{x}-\hat{y}|^2 > \rho_0^2$, inequality (8.127) directly follows from (8.126) with $m=0$ vie the triangle inequality. So we may assume there exists $m \in \mathbb{N}_0$ such that

$$\delta^{2(m+1)}\rho_0 \leq |t-s| + |\hat{x}-\hat{y}|^2 \leq \delta^{2m}\rho_0^2.$$

Consider the tangent spaces $A_1 = T_m(t,\hat{x})$ and $A_2 = T_m(s,\hat{y})$. Let $A_1 = \operatorname{span}(\tau_i)_{1\leq i\leq n}$ with $\tau_i \cdot \tau_j = \delta_{ij}$. Changing the τ_i a bit will give a basis for A_2 such that we can use Lemma A.10 to estimate $|\pi_{A_1} - \pi_{A_2}|$. For any $i \in \{1,\dots,n\}$ set $\tilde{\tau}_i := \pi_{A_2}(\tau_i)$ then calculate by the triangle inequality

$$\begin{aligned}
|\rho_m\tau_i - \rho_m\tilde{\tau}_i| = &\left|\pi_{A_2}^{\perp}(\rho_m\tau_i)\right| \\
&\leq |z_m(t,\hat{x}) + \rho_m\tau_i - F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i))| \\
&\quad + |F(t,\hat{x}) - z_m(t,\hat{x})| \\
&\quad + \left|\pi_{A_2}^{\perp}\left(F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(s,\hat{y})\right)\right| \\
&\quad + \left|\pi_{A_2}^{\perp}\left(F(t,\hat{x}) - z_m(s,\hat{y})\right)\right|.
\end{aligned} \tag{8.128}$$

By (8.119) we have

$$|F(t,\hat{x}) - z_m(t,\hat{x})| = |f(t,\hat{x}) - \pi_{\mathbb{R}^k}(z_m(t,\hat{x}))| \leq \delta^{\alpha m} h \rho_m, \tag{8.129}$$

where we used $\pi_{\mathbb{R}^n}(z_m(t,\hat{x})) = \hat{x}$ and $\rho_m = \delta^m\rho_0$. By choice of m we have $|\hat{x}-\hat{y}| \leq \rho_m$ and $|t-s| \leq \tau_m$ thus

$$\begin{aligned}
|\pi_{\mathbb{R}^n}(F(t,\hat{x}) - z_m(s,\hat{y}))| &= |\hat{y}_1 - \hat{y}_2| \leq \rho_m \\
|\pi_{\mathbb{R}^n}(F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(s,\hat{y}))| &= |\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i) - \hat{y}| \leq 2\rho_m \\
|\pi_{\mathbb{R}^n}(F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(t,\hat{x}))| &= |\rho_m\pi_{\mathbb{R}^n}(\tau_i)| \leq \rho_m.
\end{aligned}$$

So we can use (8.114) to see

$$\begin{aligned}
\left|\pi_{A_2}^{\perp}(F(t,\hat{x}) - z_m(s,\hat{y}))\right| &\leq \eta_m\rho_m \\
\left|\pi_{A_2}^{\perp}(F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(s,\hat{y}))\right| &\leq \eta_m\rho_m \\
\left|\pi_{A_1}^{\perp}(F(t,\hat{x}+\rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(t,\hat{x}))\right| &\leq \eta_m\rho_m.
\end{aligned} \tag{8.130}$$

Note that we can use (8.114) here, because (8.106) is already established. In view of (8.112), we can use Remark A.7.3 to estimate

$$\begin{aligned}
&|z_m(t,\hat{x}) + \rho_m\tau_i - F(t,\hat{x} + \rho_m\pi_{\mathbb{R}^n}(\tau_i))| \\
&= |\pi_{\mathbb{R}^k}(z_m(t,\hat{x}) + \rho_m\tau_i - F(t,\hat{x} + \rho_m\pi_{\mathbb{R}^n}(\tau_i)))| \\
&\leq 2\left|\pi_{A_1}^{\perp}(z_m(t,\hat{x}) + \rho_m\tau_i - F(t,\hat{x} + \rho_m\pi_{\mathbb{R}^n}(\tau_i)))\right| \\
&= 2\left|\pi_{A_1}^{\perp}(F(t,\hat{x} + \rho_m\pi_{\mathbb{R}^n}(\tau_i)) - z_m(t,\hat{x}))\right|,
\end{aligned}$$

where we used $\tau_i \in A_1$. Hence by (8.130)

$$|z_m(t,\hat{x}) + \rho_m\tau_i - F(t,\hat{x} + \rho_m\pi_{\mathbb{R}^n}(\tau_i))| \leq 2\eta_m\rho_m. \tag{8.131}$$

Inserting (8.129), (8.130) and (8.131) into (8.128) yields

$$|\rho_m\tau_i - \rho_m\tilde{\tau}_i| \leq C_n(\eta_m + \delta^{\alpha m}h)\rho_m = C_n\delta^{\alpha m}h\rho_m$$

for all $i \in \{1, \ldots, n\}$, thus we can use Proposition A.10.1 to conclude

$$|\pi_{A_1} - \pi_{A_2}| \leq C_n\delta^{\alpha m}h.$$

Recall that $A_1 = T_m(t,\hat{y}_1)$ and $A_2 = T_m(t,\hat{y}_2)$, so with (8.126) we obtain

$$\begin{aligned}
\left|\pi_{T_{F(t,\hat{x})}\mu_t} - \pi_{T_{F(s,\hat{y})}\mu_s}\right| &\leq \left|\pi_{T_{F(t,\hat{x})}\mu_t} - \pi_{A_1}\right| + |\pi_{A_1} - \pi_{A_2}| + \left|\pi_{A_2} - \pi_{T_{F(s,\hat{y})}\mu_s}\right| \\
&\leq C_n\delta^{\alpha(m-1)}h,
\end{aligned}$$

where we used $\delta^{-\alpha} \geq 1$. By choice of m, we have

$$\delta^{\alpha(m-1)} \leq \delta^{-2\alpha}\rho_0^{-\alpha}\left(|t-s| + |\hat{x} - \hat{y}|^2\right)^{\frac{\alpha}{2}},$$

so we verified (8.127). Note that $(a+b)^p \leq 2^p(a^p + b^p)$ for $a, b, p \in [0,\infty)$, thus for $\Lambda \geq C_n\delta^{-\alpha}$ (8.127) establishes (8.108). $\square$

9 General Regularity

Here we want to give a new, shorter proof of Brakke's general regularity theorem [B, 6.12]. The basic idea remains, that for almost every point in a unit density varifold we either have density one and the existence of a tangent space or density zero. Then looking close enough the conditions for either the local regularity theorem or the clearing out lemma should be satisfied, which establishes the required regularity.

We recall the definition of our usual cut-off function, which we shall use here to cut-off spherically rather than cylindrically.

9.1 Definition. Recall $\zeta \in C^\infty([0,\infty),[0,1])$ from Definition 4.1. We will use

$$\zeta(r) = \begin{cases} 1 & \text{for } 0 \le r \le 1 - 2^{-n-9} \\ 0 & \text{for } 1 \le r \end{cases}$$

and $\max\{\sup|\zeta'|, \sup|\zeta''|\} \le \sigma_1$.

First we need another version of Theorem 8.4, where the absolute height bound (8.103) is replaced by an integral one. This follows easily with Corollary 6.8.

9.2 Theorem (Local Regularity Theorem (2nd Version), [B, 6.11])**.** *For every $\lambda_0 \in (0, 2^{-2}]$ there exist a $\Lambda_0 \in (1,\infty)$ and a $\gamma_0 \in (0,1)$ such that for all $t_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+k}$ and every $R \in (0,\infty)$ the following holds: Let $(\mu_t)_{t\in[t_0-\Lambda_0R^2, t_0+\Lambda_0R^2]}$ be Brakke flow in $B_{(2+2\lambda_0)R}(0)$ and T an n-dimensional subspace of $\mathbb{R}^{n+k}$ with*

$$R^{-n-1}\int_{B_{(2+2\lambda_0)R}(y_0)} \left|\pi_T^\perp(x-y_0)\right| d\mu_t(x) \le \gamma_0 \tag{9.1}$$

$$R^{-n}\mu_t\left(B_{(1+2\lambda_0)R}(y_0)\right) \le (2-\lambda)\omega_n, \quad R^{-n}\mu_t\left(B_R(y_0)\right) \ge \lambda\omega_n \tag{9.2}$$

for all $t \in [t_0 - \Lambda_0R^2, t_0 + \Lambda_0R^2]$. Then there exists a smooth function $f : [t_0 - R^2, t_0 + R^2] \times B^n_{\lambda R}(\hat{x}_0) \to \mathbb{R}^k$ with

$$(\mathrm{spt}\mu_t \cap C_{\lambda R}(y_0)) = S(\mathrm{graph}(f(t,\cdot))), \tag{9.3}$$

for all $t \in [t_0 - R^2, t_0 + R^2]$ and some $S \in SO(n+k)$

Proof. We may assume $y_0 = 0$ and $T = \mathbb{R}^n$. Consider $t \in [t_0 - \Lambda_0R^2 + R^2, t_0 + \Lambda_0R^2]$ and $x_0 \in \mathrm{spt}\mu_t \cap B_{(1+2\lambda_0)R}(0)$, by Corollary 6.8 with $y_0 = 0$, $v = \mathbf{e}_{n+j}$ and $r = 2^{-1}R$ combined with (9.1) we obtain

$$|x_0 \cdot \mathbf{e}_{n+j}| \le C_nR^{-n-2}\int_{t_0-R^2}^{t_0}\int_{B_R(x_0)} |x \cdot \mathbf{e}_{n+j}| d\mu_t(x) \le C_n\gamma_0 R$$

for all $j \in \{1, \dots, k\}$, where we used $y = x_0$. Thus we obtain

$$\mathrm{spt}\mu_t \cap B_{(1+2\lambda_0)R}(0) \subset \{x \in \mathbb{R}^{n+k} : |\pi_{\mathbb{R}^k}(x)| \le C_n \gamma_0 R\}$$

for all $t \in [t_0 - \Lambda_0 R^2 + R^2, t_0 + \Lambda_0 R^2]$. Then we can apply Theorem 8.4 with $t_1 = t_0 - \Lambda_0 R^2 + R^2$, $t_2 = t_0 + \Lambda_0 R^2$, $\lambda = \lambda_0$, $\alpha = 2^{-1}$ and Brakke flow $(\mu_t \llcorner B_{1+2\lambda_0)R}(0))_{t\in[t_1,t_2]}$, which yields a $\Lambda = \Lambda(\lambda_0) \in (1, \infty)$ and an $h_0 = h_0(\Lambda_0, \lambda_0) \in (0, 1)$ such that, if $C_n\gamma_0 \le h_0$, we obtain (9.3) for the time interval $[t_0 + (\Lambda + 1 - \Lambda_0)R^2, t_0 - (\Lambda - \Lambda_0)R^2]$. So choosing $\gamma_0 \le C_n^{-1} h_0$ and $\Lambda_0 \ge \Lambda + 2$ establishes the result. $\square$

The Brakke flow allows the solution to "jump" in the sense of a sudden local loss of area. In such a case we cannot expect to obtain regularity, so in the following we will rule out these "jump-decreases" by an extra assumption.

9.3 Definition. For $t_1, t_2 \in \mathbb{R}$ with $t_1 < t_2$, an open subset $U \in \mathrm{R}^{n+k}$, and a time $t_0 \in (t_1, t_2)$, a Brakke flow $(\mu_t)_{t\in[t_1,t_2]}$ in U is called *continuous at time* t_0, if for every $\phi \in C_c^2(U, \mathbb{R}^+)$

$$\lim_{\delta\to 0} \mu_{t_0+\delta}(\phi) = \mu_{t_0}(\phi). \tag{9.4}$$

9.4 Lemma. *For all $R, \tau \in (0, \infty)$, $x_0 \in \mathbb{R}^{n+k}$ and every open subset $U \in \mathrm{R}^{n+k}$ with $B_R(x_0) \subset\subset U$ the following holds: Let $(\mu_t)_{t\in[-\tau,\tau]}$ be a Brakke flow in U, which is continuous at time* 0. *Then*

$$\lim_{\delta\searrow 0} \int_{-\delta}^{\delta} \int_{B_R(x_0)} |\vec{H}|^2 d\mu_t \, dt = 0. \tag{9.5}$$

Proof. First suppose $B_{3R}(x_0) \subset\subset U$. Consider a $\delta \in (0, \tau)$. Look at the test function $\varphi : (-\delta, +\delta) \times \mathbb{R}^{n+k} \to \mathbb{R}^+$

$$\varphi(t, x) = \left\{1 - \frac{|x - x_0|^2 + 2nt}{4R^2}\right\}_+.$$

Note that for $t \ge -2n^{-1}R^2$ the support of φ lies in $B_{3R}(x_0)$. For almost every $t \in (-\delta, +\delta)$ there exists an L^2-integrable mean curvature vector $\vec{H}$ on U. For these t we can calculate with Remark 2.6

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\varphi^3\right) - \mathrm{div}_{\mu_t} D\left(\varphi^3\right) &= -\frac{3n}{2R^2}\varphi^2 - 3\varphi\nabla^{\mu_t}(\varphi^2)\cdot D\varphi - 3\varphi^2 \mathrm{div}_{\mu_t}(D\varphi) \\
&= -\frac{3n}{2R^2}\varphi^2 - 6\varphi\left|\nabla^{\mu_t}\varphi\right|^2 - \frac{3}{2R^2}\varphi^2 \mathrm{div}_{\mu_t}(x) \\
&= -6\varphi\left|\nabla^{\mu_t}\varphi\right|^2 \le 0
\end{aligned}$$

at almost every $x \in B_{3R}(x_0)$, which yields

$$\int_{\mathbb{R}^{n+k}} D\left(\varphi^3\right) \cdot \vec{H} d\mu_t = -\int_{\mathbb{R}^{n+k}} \operatorname{div}_{\mu_t} D\left(\varphi^3\right) d\mu_t \leq -\int_{\mathbb{R}^{n+k}} \frac{\partial}{\partial t}\left(\varphi^3\right) d\mu_t.$$

Thus

$$\mathscr{B}\left(\mu_t, \varphi(t,\cdot)^3\right) + \int_U \frac{\partial}{\partial t}\left(\varphi(t,x)^3\right) d\mu_t(x) \leq -\int_U \varphi(t,x)^3 |\vec{H}(x)|^2 d\mu_t(x)$$

for almost every $t \in (-\delta, +\delta)$. Then with Proposition 3.8 we obtain

$$\mu_\delta\left(\varphi(\delta,\cdot)^3\right) - \mu_{-\delta}\left(\varphi(-\delta,\cdot)^3\right) \leq -\int_{-\delta}^{\delta}\int_U \varphi(t,x)^3 |\vec{H}(x)|^2 d\mu_t(x) dt. \tag{9.6}$$

For $\delta \leq (2n)^{-1}R^2$ and $t \in [-\delta, \delta]$ we can estimate

$$\begin{aligned} \left|\varphi(t,x)^3 - \varphi(0,x)^3\right| &\leq \left(\varphi(0,x) + (2R^2)^{-1}n\delta\right)^3 - \varphi(0,x)^3 \\ &\leq 2^3(2R^2)^{-1}n\delta \leq C_n R^{-2}\delta. \end{aligned}$$

Here we used $\varphi(0,x) \leq 1$. Also for $x \in B_R(x_0)$ and $\delta \leq (2n)^{-1}R^2$ we can estimate for $t \in [-\delta, \delta]$

$$\varphi(t,x)^3 \geq \left(\frac{3R^2 - 2n\delta}{4R^2}\right)^3 \geq 2^{-3}.$$

Thus (9.6) implies

$$\begin{aligned} 8\int_{-\delta}^{\delta}\int_{B_R(x_0)} |\vec{H}|^2 d\mu_t\, dt \leq &\left(\mu_{-\delta}\left(\varphi(0,\cdot)^3\right) - \mu_\delta\left(\varphi(0,\cdot)^3\right)\right) \\ &+ C_n R^{-2}\delta\left(\mu_{-\delta}\left(B_{3R}(x_0)\right) + \mu_\delta\left(B_{3R}(x_0)\right)\right), \end{aligned} \tag{9.7}$$

for all $0 < \delta \leq \min\{\tau, (2n)^{-1}R^2\}$. Here we used $\operatorname{spt}\varphi(t,\cdot) \subset B_{3R}(x_0)$ for all $t \geq -2n^{-1}R^2$. For $\delta \searrow 0$ the first difference on the right hand side of (9.7) goes to 0, due to the continuity at time 0. Furthermore by Lemma 3.6 there exists $M \in (0,\infty)$ such that,

$$\mu_t\left(B_{3R}(x_0)\right) \leq M$$

for all $t \in [-\tau, \tau]$, so for $\delta \searrow 0$ the second difference on the right hand side of (9.7) goes to 0 as well. So we obtain

$$\limsup_{\delta \searrow 0} \int_{-\delta}^{\delta}\int_{B_R(x_0)} |\vec{H}|^2 d\mu_t\, dt \leq 0.$$

But for almost every $t \in (-\delta, \delta)$ the integral $\int_{\mathbb{R}^{n+k}} |\vec{H}|^2 d\mu_t$ is well defined and positive, which implies that the limit exists and equals 0.

If now $B_R(x_0) \subset\subset U$ there exists $r < 3^{-1} d(\partial U, B_R(x_0))$. Then we can cover $B_R(x_0)$ by a finite collection of $B_r(x_i)$, $x_i \in B_R(x_i)$, $i = 1, \dots, N$. For each i we have $B_{3r}(x_i) \subset\subset U$, so we can use the previous conclusions inside these balls and estimate

$$\int_{-\delta}^{\delta} \int_{B_R(0)} |\vec{H}|^2 d\mu_t \, dt \le \sum_{i=1}^{N} \int_{-\delta}^{\delta} \int_{B_r(x_i)} |\vec{H}|^2 d\mu_t \, dt \to 0$$

for $\delta \to 0$. $\square$

The main ingredient for the general regularity theorem will be the next technical lemma. It basically says that if one considers the set of points x where μ_t "jumps" in a δ small parabolic ball around (t_0, x), then the measure of this set should vanish for $\delta \searrow 0$. This lemma does not appear in [B], but the main calculation, the one that leads to (9.10) below, is taken from [B, 6.12].

9.5 Lemma. *For all $R, L, \tau \in (0, \infty)$ the following holds: Let $(\mu_t)_{t \in [-R^2, R^2]}$ be a Brakke flow in $B_{2R}(0)$ which is continuous at time 0. Consider the set*

$$\begin{aligned} D(\tau, \delta) &:= \{x \in B_R(0),\ \mathscr{D}(x, \delta) \ge \tau\}, \ \textit{with} \\ \mathscr{D}(x, \delta) &:= \sup_{\phi \in C_c^{0,1}(B_\delta(x), [0,1]), \mathrm{lip}(\phi) \le \delta^{-1} L} \ \sup_{t \in (-\delta^2, \delta^2)} \left|\delta^{-n} \mu_t(\phi) - \delta^{-n} \mu_0(\phi)\right| \end{aligned}$$

for $\delta \in (0, R)$. Then $\mathscr{H}^n \left(\bigcap_{\delta \in (0,R)} D(\tau, \delta) \right) = 0$.

Now in addition assume there exists a subset $A \subset B_R(0)$ and a collection of functions $\{\vartheta_{\delta,x} \in C_c^{0,1}(B_\delta(x), [0,1]), x \in A, \delta \in (0, R)\}$ with

$$\mathrm{lip}(\vartheta_{\delta,x}) \le \delta^{-1} L \quad \textit{and} \quad \lim_{\delta \searrow 0} \delta^{-n} \mu_0(\vartheta_{\delta,x}) =: \varrho(x) \in \mathbb{R}$$

for all $x \in A$ and all $\delta \in (0, R)$. Consider the set

$$E(\tau, \delta) := \left\{ x \in A,\ \sup_{t \in (-\delta^2, \delta^2)} \left|\delta^{-n} \mu_t(\vartheta_{\delta,x}) - \varrho(x)\right| \ge \tau \right\}$$

for $\delta \in (0, R)$. Then $\mathscr{H}^n \left(\bigcap_{\delta \in (0,R)} E(\tau, \delta) \right) = 0$.

For the proof we need Vitali's covering theorem found in [EG, 1.5.1] which says:

9.6 Theorem (Vitali Covering Theorem)**.** *Let $R_0 \in (0, \infty)$ and $(B_{r_i}(x_i))_{i\in I}$ be a family of balls $r_i \in (0, R_0), x_i \in \mathbb{R}^n$ for an arbitrary set of indices I. Then there exists a countable subset $J \subset I$ such that $\overline{B_{r_i}(x_i)} \cap \overline{B_{r_j}(x_j)} = \emptyset$ for all $i, j \in J$ and*

$$\bigcup_{i\in I} \overline{B_{r_i}(x_i)} \subset \bigcup_{i\in J} \overline{B_{5r_i}(x_i)}.$$

Proof of 9.5. Let $R, L, \tau \in (0, \infty)$ be given. Consider $\delta \leq 2^{-1}R$, then

$$\bigcup_{x\in D(\tau,\delta)} B_\delta(x) \supset \overline{D(\tau, \delta)}. \tag{9.8}$$

Thus by Theorem 9.6 we obtain a finite covering

$$\bigcup_{i=1}^{N} B_{5\delta}(b_i) \supset \overline{D(\tau, \delta)}$$

such that, the $B_\delta(b_i)$ are disjoint and $b_i \in D(\tau, \delta)$ for all $i \in \{1, \ldots, N\}$, where $N \in \mathbb{N}$ depends on δ. Here Vitali's theorem first yields a countable covering, but as $\overline{D(\tau, \delta)}$ is compact, we can choose finitely many balls, which already cover $\overline{D(\tau, \delta)}$.

By definition of $D(\tau, \delta)$ for every $i \in \{1, \ldots, N\}$ there exist a $\phi_i \in C_c^{0,1}(B_\delta(b_i), [0, 1])$ with $\mathrm{lip}(\phi_i) \leq \delta^{-1}L$ and a $t_i \in (-\delta^2, \delta^2)$ such that $\frac{\tau}{2}\delta^n \leq |\mu_{t_i}(\phi_i) - \mu_0(\phi_i)|$. Then by (3.1) and Proposition A.19 we can estimate

$$\begin{aligned}
\frac{\tau}{2}\delta^n &\leq |\mu_{t_i}(\phi_i) - \mu_0(\phi_i)| \\
&\leq \mu_{-\delta^2}(\phi_i) - \mu_{\delta^2}(\phi_i) + 2 \sup_{-\delta^2 \leq s_1 < s_2 \leq \delta^2} |\mu_{s_1}(\phi_i) - \mu_{s_2}(\phi_i)| \\
&\leq \mu_{-\delta^2}(\phi_i) - \mu_{\delta^2}(\phi_i) + 2 \int_{-\delta^2}^{\delta^2} \int_{B_\delta(b_i)} |D\phi_i| \, |\vec{H}| d\mu_t \, dt
\end{aligned}$$

for every $i \in \{1, \ldots, N\}$, where we used $\mathrm{spt}\phi_i \subset B_\delta(b_i)$ and that $D\phi$ exists almost everywhere. Combined with (9.8) this lets us estimate the $\mathscr{H}^n_{10\delta}$-measure of $D(\tau, \delta)$ by

$$\begin{aligned}
&\mathscr{H}^n_{10\delta}(D(\tau, \delta)) \leq \mathscr{H}^n_{10\delta}\left(\bigcup_{i=1}^{N} B_{5\delta}(b_i)\right) \leq 5^n\omega_n \sum_{i=1}^{N} \delta^n \\
&\leq C_n\tau^{-1}\left(\sum_{i=1}^{N} (\mu_{-\delta^2}(\phi_i) - \mu_{\delta^2}(\phi_i)) + \int_{-\delta^2}^{\delta^2} \int_{B_\delta(b_i)} |D\phi_i| \, |\vec{H}| d\mu_t \, dt\right).
\end{aligned}$$

Then using $|D\phi_i(x)| \leq \delta^{-1}L$ for almost every $x \in B_\delta(b_i)$ and the disjointness of the $B_\delta(b_i)$ yields

$$
\begin{aligned}
&\mathscr{H}^n_{10\delta}\left(D(\tau,\delta)\right) \\
&\leq C_n\tau^{-1}\left(\sum_{i=1}^N \mu_{-\delta^2}\left(\phi_i\right) - \mu_{\delta^2}\left(\phi_i\right) + L\delta^{-1}\int_{-\delta^2}^{\delta^2}\int_{B_{2R}(0)} |\vec{H}| d\mu_t\, dt\right).
\end{aligned}
\tag{9.9}
$$

Now use the radial cut-off function $\zeta_0(x) = \zeta\left(\frac{|x|}{2R}\right)$ for $x \in \mathbb{R}^{n+k}$, where ζ is from Definition 9.1. By definition of ζ_0 and $\delta \leq 2^{-1}R$ we have

$$
B_{2R}(0) \supset\supset \mathrm{spt}\zeta_0 \supset \{\zeta_0 = 1\} \supset B_{\frac{3}{2}R}(0) \supset B_\delta(b_i) \supset \mathrm{spt}\phi_i
$$

for all $i \in \{1,\dots,N\}$. So the disjointness of the $B_\delta(b_i)$ yields

$$
\zeta_0(x) - \sum_{i=1}^N \phi_i(x) \geq 0
$$

for all $x \in B_{2R}$. In particular this can be used as the test function in (3.1). Then by (3.1) and Proposition A.19 we obtain

$$
\begin{aligned}
&\sum_{i=1}^N \left(\mu_{-\delta^2}\left(\phi_i\right) - \mu_{\delta^2}\left(\phi_i\right)\right) \\
&= \mu_{-\delta^2}\left(\zeta_0\right) - \mu_{\delta^2}\left(\zeta_0\right) + \mu_{\delta^2}\left(\zeta_0 - \sum_{i=1}^N \phi_i\right) - \mu_{-\delta^2}\left(\zeta_0 - \sum_{i=1}^N \phi_i\right) \\
&\leq \mu_{-\delta^2}\left(\zeta_0\right) - \mu_{\delta^2}\left(\zeta_0\right) + \int_{-\delta^2}^{\delta^2}\int_{B_{2R}(0)} \left|D\left(\zeta_0 - \sum_{i=1}^N \phi_i\right)\right| |\vec{H}| d\mu_t\, dt.
\end{aligned}
$$

As $|D\phi_i| \leq \delta^{-1}L$ and for $\delta \leq (\sigma_1 L)^{-1}R$ also $|D\zeta_0| \leq \sigma_1(2R)^{-1} \leq \delta^{-1}L$ we can estimate $\left|D\left(\zeta_0 - \sum_{i=1}^N \phi_i\right)\right| \leq 2L\delta^{-1}$. Thus (9.9) becomes

$$
\begin{aligned}
&\mathscr{H}^n_{10\delta}\left(D(\tau,\delta)\right) \\
&\leq C_n\tau^{-1}\left(\left(\mu_{-\delta^2}\left(\zeta_0\right) - \mu_{\delta^2}\left(\zeta_0\right)\right) + L\delta^{-1}\int_{-\delta^2}^{\delta^2}\int_{V} |\vec{H}| d\mu_t\, dt\right)
\end{aligned}
\tag{9.10}
$$

for all $\delta \in (0, (\sigma_1 L + 1)^{-1}R)$, where $V := \mathrm{spt}\zeta_0$. Now let ϵ be given, then there exists a $\delta_1 \in (0,1)$ depending on ϵ such that

$$
\mathscr{H}^n\left(\bigcap_{\delta\in(0,R)} D(\tau,\delta)\right) \leq \mathscr{H}^n_{10\delta_0}\left(\bigcap_{\delta\in(0,R)} D(\tau,\delta)\right) + \epsilon
\tag{9.11}
$$

for all $\delta_0 \in (0, \delta_1]$. Also as the flow is continuous at time 0 there exists $\delta_2 \in (0,1)$ depending on ϵ such that

$$\mu_{-\delta^2}(\zeta_0) - \mu_{\delta^2}(\zeta_0) \leq \epsilon \tag{9.12}$$

for all $\delta \in (0, \delta_2]$. As $V = \text{spt}\zeta_0 \subset\subset B_{2R}$, by Lemma 3.6 there exists an $M \in (1,\infty)$ such that $\mu_t(V) \leq M$ for all $t \in [-R^2, R^2]$. Then use Hölder's estimate and Lemma 9.4 to obtain a $\delta_3 \in (0,1)$ depending on ϵ such that

$$\delta^{-1} \int_{-\delta^2}^{\delta^2} \int_B |\vec{H}| d\mu_t \, dt \leq C_n \sqrt{M} \left(\int_{-\delta^2}^{\delta^2} \int_B |\vec{H}|^2 d\mu_t \, dt \right)^{\frac{1}{2}} \leq C_n \sqrt{M} \epsilon \tag{9.13}$$

for all $\delta \in (0, \delta_3]$. Then combining (9.10)-(9.13) we conclude for $\delta_0 = \min\{\delta_1, \delta_2, \delta_3, (\sigma_1 L + 1)^{-1} R\}$

$$\mathscr{H}^n \left(\bigcap_{\delta \in (0,R)} D(\tau, \delta) \right) \leq \mathscr{H}^n_{10\delta_0} \left(D(\tau, \delta_0) \right) + \epsilon \leq C_n (1 + \tau^{-1}) M \epsilon$$

for all $\epsilon \in (0,1)$ and letting ϵ go to 0 establishes the first result.

For the second part consider the set

$$A(r, \delta) := \left\{ x \in A, \ \left| \delta^{-n} \mu_0(\vartheta_{\delta,x}) - \varrho(x) \right| \geq r \right\}$$

for $r \in (0, \infty)$. Then for given $\tau \in (0,\infty)$ we have

$$E(\tau, \delta) \subset \left(D(2^{-2}\tau, \delta) \cup A(2^{-2}\tau, \delta) \right). \tag{9.14}$$

To see this consider $x \in E(\tau, \delta) \setminus A(2^{-2}\tau, \delta)$. As $x \in E(\tau, \delta)$ there has to exist a $t \in (-\delta^2, \delta^2)$ such that $|\delta^{-n} \mu_t(\vartheta_{\delta,x}) - \varrho(x)| \geq 2^{-1}\tau$. Then, as $x \notin A(2^{-2}\tau, \delta)$, we can estimate

$$\begin{aligned} \left| \delta^{-n} \mu_t(\vartheta_{\delta,x}) - \delta^{-n} \mu_0(\vartheta_{\delta,x}) \right| &\geq \left| \delta^{-n} \mu_t(\vartheta_{\delta,x}) - \varrho(x) \right| - \left| \varrho(x) - \delta^{-n} \mu_0(\vartheta_{\delta,x}) \right| \\ &\geq 2^{-1}\tau - 2^{-2}\tau = 2^{-2}\tau \end{aligned}$$

and as $\vartheta_{\delta,x} \in C^{0,1}_c(B_\delta(x), [0,1])$ with $\text{lip}\vartheta_{\delta,x} \leq \delta^{-1} L$ we see that $x \in D(2^{-2}\tau, \delta)$.

As we assumed $\lim_{\delta \searrow 0} \delta^{-n} \mu_0(\vartheta_{\delta,x}) = \varrho(x)$ we see that $\bigcap_{\delta \in (0,R)} A(2^{-2}\tau, \delta) = \emptyset$. Thus (9.14) yields

$$\bigcap_{\delta \in (0,R)} E(\tau, \delta) \subset \bigcap_{\delta \in (0,R)} D(2^{-2}\tau, \delta).$$

We have already seen that $\bigcap_{\delta \in (0,R)} D(2^{-2}\tau, \delta)$ has Hausdorff measure 0, which establishes the result. $\square$

Now we have all the ingredients to prove Brakke's general regularity theorem

9.7 Theorem (General Regularity Theorem, [B, 6.12])**.** *Let $\rho \in (0,\infty)$ and $(\mu_t)_{t\in[-\rho^2,\rho^2]}$ be a Brakke flow in $B_{2\rho}(0)$ which is continuous at time 0 and suppose μ_0 is a unit density n-varifold. Then there exists a set S with $\mathscr{H}^n(S)=0$ such that the following holds: For every $x \in B_\rho(0)\setminus S$ there exists an $r\in(0,\rho)$ such that either $\operatorname{spt}\mu_t \cap B_r(x) = \emptyset$ or $\operatorname{spt}\mu_t\cap B_r(x) = F_t(B^n_r(0))$ for a smooth family of embeddings $F_t : (-r^2,r^2)\times B^n_r(0)\to\mathbb{R}^{n+k}$.*

Proof. Let S be the set of all $x\in B_\rho(0)$ for which there exists no such r. We want to show $\mathscr{H}^n(S)=0$. Consider the sets

$$\begin{aligned} V &= \{x\in B_\rho(0),\ \Theta^n(\mu_0,x)=\theta_x = 1\ \wedge \exists\, T_x\mu_0\}\\ W &= \{x\in B_\rho(0),\ \Theta^n(\mu_0,x)=0\}. \end{aligned}$$

As μ_0 is a unit density varifold we have $\mathscr{H}^n\left(B_\rho(0)\setminus(V\cup W)\right)=0$ by Remark (2.3). So it suffices to consider $S\cap V$ and $S\cap W$ and prove they both have measure 0. We will do this by using the local regularity theorem and the clearing out lemma respectively. Use the radial cut-off function $\zeta_{r,x}(y)=\zeta\left(r^{-1}|y-x|\right)$ for $x,y\in\mathbb{R}^{n+k}$, where ζ is from Definition 9.1. Set $\varpi := \int_{\mathbb{R}^n\times\{0\}^k}\zeta^2_{1,0}(y)d\mathscr{H}^n(y)$, then $(1-2^{-8})\omega_n\le\varpi\le\omega_n$.

First we want to show $\mathscr{H}^n(S\cap V)=0$. For $\Lambda\in(1,\infty)$, $\gamma\in(0,1)$ and $r\in(0,\Lambda^{-1}\rho)$ consider

$$\begin{aligned} V_1(r) &= \left\{x\in V,\ \sup_{t\in[-\Lambda^2r^2,\Lambda^2r^2]} r^{-n-1}\int_{B_r(x)}\left|\pi^\perp_{T_x\mu_0}(y-x)\right|\zeta_{r,x}\,d\mu_t\ge\gamma\right\}\\ V_2(r) &= \left\{x\in V,\ \sup_{t\in[-\Lambda^2r^2,\Lambda^2r^2]}\left|r^{-n}\mu_t\left(\zeta_{r,x}\right)-\varpi\right|\ge\frac{\omega_n}{4}\right\}. \end{aligned}$$

By Theorem 9.2 we can fix $\Lambda\in(1,\infty)$ and $\gamma\in(0,1)$ such that

$$S\cap V\subset\bigcap_{r\in(0,(6\Lambda)^{-1}\rho^2)}\left(V_1(6r)\cup V_2((1+2^{-n-3})r)\cup V_2(r)\right). \tag{9.15}$$

To see this take an $x\in V$ which is not in the larger set on the right. So there exists $r\in(0,(6\Lambda)^{-1}R^2)$ such that x is not in $V_1(3r)\cup V_2((1+2^{-n-3})r)\cup V_2(r)$ which yields

$$\begin{aligned} \int_{B_{3r}(x)}\left|\pi^\perp_{T_x\mu_0}(y-x)\right|d\mu_t\le\int_{B_{6r}(x)}\left|\pi^\perp_{T_x\mu_0}(y-x)\right|\zeta_{6r,x}(y)d\mu_t\le\gamma(6r)^{n+1},\\ \mu_t\left(B_{(1+2\lambda_0)r}(x)\right)\le\mu_t\left(\zeta_{(1+2^{-n-3})r,x}\right)\le\frac{5\omega_n}{4}(1+2^{-n-3})^nr^n\le\frac{3\omega_n}{2}r^n,\\ \mu_t\left(B_r(x)\right)\ge\mu_t\left(\zeta_{r,x}\right)\ge\left(\varpi-\frac{\omega_n}{4}\right)r^n\ge\frac{\omega_n}{2}r^n \end{aligned}$$

for all $t \in [-\Lambda^2 r^2, \Lambda^2 r^2]$ and $2\lambda_0 := (1-2^{-n-8})(1+2^{-n-3}) - 1 \in (0, 2^{-n-3})$. Here we used

$$B_\delta(x) \supset \mathrm{spt}\zeta_{\delta,x} \supset \{\zeta_{\delta,x} = 1\} \supset B_{(1-2^{-n-8})\delta}(x) \supset B_{2^{-1}\delta}(x).$$

Then by Theorem 9.2 with $y_0 = x$, $T = T_x\mu_0$ and $R = r$ there exist Λ_0 and γ_0 such that, if $\Lambda^2 \geq \Lambda_0$ and $\gamma \leq 6^{-n-1}\gamma_0$ we obtain that $x \notin S$. This implies (9.15).

By Remark 2.3 and as we have density 1 in x we obtain

$$\begin{aligned}
&\lim_{r\searrow 0} r^{-n-1}\mu_0\left(\left|\pi^\perp_{T_x\mu_0}(\cdot - x)\right|\zeta_{r,x}\right)\\
&= \lim_{r\searrow 0} r^{-n}\int_{\mathbb{R}^{n+k}}\left|\pi^\perp_{T_x\mu_0}\left(\frac{y-x}{r}\right)\right|\zeta\left(\frac{|y-x|}{r}\right)d\mathscr{H}^n(y)\\
&= \int_{T_x\mu_0}\left|\pi^\perp_{T_x\mu_0}(y)\right|\zeta(|y|)\,d\mathscr{H}^n(y) = 0,
\end{aligned}$$

for all $x \in V$. Then using Lemma 9.5 with $R = \rho$, $\tau = \Lambda^{-n}\gamma r$, $L = \Lambda(1+\sigma_1)$, $\delta = \Lambda r$, $\varrho(x) = 0$ and $\vartheta_{\delta,x}(y) = \Lambda\delta^{-1}\left|\pi^\perp_{T_x\mu_0}(y-x)\right|\zeta_{\Lambda^{-1}\delta,x}(y)$ yields that $\bigcap_{r\in(0,\Lambda^{-1}\rho)} V_1(r)$ has Hausdorff measure 0. Similarly by Remark 2.3 and as we have density 1 in x we obtain

$$\begin{aligned}
\lim_{r\searrow 0} r^{-n}\mu_0(\zeta_{r,x}) &= \lim_{r\searrow 0} r^{-n}\int_{\mathbb{R}^{n+k}}\zeta\left(\frac{|y-x|}{r}\right)d\mu_0\\
&= \int_{T_x\mu_0}\zeta(|y|)\,d\mathscr{H}^n(y) = \int_{\mathbb{R}^n\times\{0\}^k}\zeta_{1,0}\,d\mathscr{H}^n = \varpi,
\end{aligned}$$

Then using Lemma 9.5 with $R = r$, $\tau = 2^{-2}\Lambda^{-n}\omega_n$, $L = \Lambda\sigma_1$, $\delta = \Lambda r$, $\varrho(x) = \varpi$ and $\vartheta_{\delta,x} = \zeta_{\Lambda^{-1}\delta,x}$ yields that $\bigcap_{r\in(0,\Lambda^{-1}\rho)} V_2(r)$ has Hausdorff measure 0. Thus in view of (9.15) we conclude $\mathscr{H}^n(S \cap V) = 0$.

In the same way we can show $\mathscr{H}^n(S \cap W) = 0$. For $\beta \in (0,1)$ and $r \in (0,\rho)$ consider

$$W(r) = \left\{x \in W, \sup_{t\in[-r^2,r^2]} \delta^{-n}\mu_t(\zeta_{r,x}) \geq \beta\right\}.$$

By Lemma 5.7 we can fix $\beta \in (0,1)$ such that

$$S \cap W \subset \bigcap_{r\in(0,2^{-1}\rho)} W(2r). \tag{9.16}$$

To see this take an $x \in W$ which is not in the set on the right. Then there exists $r \in (0,\infty)$ such that $x \notin W(2r)$ which yields

$$\mu_t(B_r(x)) \leq \mu_t(\zeta_{2r,x}) \leq \beta r^n$$

for all $t \in [-2r^2, 2r^2]$. Then by Lemma 5.7 with $R = r$, $\eta = \beta$, $x_0 = x$ and $t_0 = -(8n)^{-1}r^2$, we obtain that for all $t \in [-(8n)^{-1}r^2 + C_n\beta^{2\sigma}r^2, 2r^2]$ we have

$$\mu_t(B_{r(t)}(x)) = 0,$$

where $r(t) = \sqrt{r^2 - 4n(t - t_0)}$ and $\sigma = \frac{1}{n+6}$. In particular if we choose β small enough such that $C_n\beta^{2\sigma} \leq (8n)^{-1}$, we obtain that $\mu_0(B_{r(0)}(x)) = 0$ and as $r(0) = \sqrt{r^2 - 2^{-1}r^2} \geq 2^{-1}r$ we see $x \notin S$. This implies (9.16).

By definition of density (see Definition 2.1) and as we have density 0 in x we obtain

$$0 \leq \lim_{r \searrow 0} r^{-n}\mu_0\left(\zeta_{r,x}\right) \leq \lim_{r \searrow 0} r^{-n}\mu_0\left(B_r(x)\right) = 0,$$

for all $x \in W$. Then using Lemma 9.5 with $R = \rho$, $\tau = \gamma$, $L = \sigma_1$, $\delta = r$, , $\varrho(x) = 0$ and $\vartheta_{\delta,x} = \zeta_{\delta,x}$ yields that $\bigcap_{r\in(0,\rho)} W(r)$ has Hausdorff measure 0. Thus in view of (9.16) we conclude $\mathscr{H}^n(S \cap W) = 0$ which establishes the result. $\square$

10 Opening Holes

In this section we study how the area ratio of a Brakke flow inside a growing cylinder behaves. Brakke observed that the increase can be controlled by a bound on the height integral, see [B, 6.5]. We will reproduce his calculations in a slightly different form, where we specify the exact dependence on the growth rate of the cylinder. This leads to Proposition 10.4, which will be used to show that for Brakke flows in narrow slabs, holes can open arbitrarily fast, see Proposition 10.6.

10.1 Definition. Recall $\zeta \in C^\infty([0,\infty),[0,1])$ from Definition 4.1. For $R \in (0,\infty)$ we define

$$\zeta_R(x) := \zeta\left(\left(R^{-1}|\hat{x}|\right)\right)$$

for all $x \in \mathbb{R}^{n+k}$. Also set $\varpi := \int_{B_1^n(0)\times\{0\}^k} \zeta_1^2 d\mathscr{H}^n$.

This is as in Definition 7.1 with $p = 1$. Then as in Lemma 7.2 we have again

10.2 Lemma. *There exists a $C \in (1,\infty)$ such that for all $R \in (0,\infty)$ the following holds:*

1. *for all $x \in \mathbb{R}^{n+k}$*

$$\zeta_R(x) = \begin{cases} 1 & \text{for } 0 \le |\hat{x}| \le (1-2^{-n-8})R \\ 0 & \text{for } R \le |\hat{x}|\,. \end{cases}$$

2. $\max\left\{R\sup|D\zeta_R|, R^2\sup|D^2\zeta_R|\right\} \le C$.
3. $R^{-n}\int_{B_R^n(0)\times\{0\}^k} \zeta_R^2 d\mathscr{H}^n = \varpi$.
4. $(1-2^{-8})\omega_n \le \varpi \le \omega_n$.

First we derive a bound for the time derivative of a growing test function on a varifold. This is based on the first part of [B, 6.5].

10.3 Lemma. *There exists a constant $C \in (1,\infty)$ such that for all $R_1, \lambda, \gamma^2 \in (0,\infty)$, $t_1 \in \mathbb{R}$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Set $R(t) := \sqrt{R_1^2 + \lambda(t-t_1)}$ and let $t_0 \in (t_1 - \lambda^{-1}R_1^2, \infty)$ be such that*

$$\operatorname{spt}\mu \cap C_{R(t_0)}(0) \subset\subset U \tag{10.1}$$

$$\int_{C_{R(t_0)}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu \le \gamma^2 R(t_0)^{n+2}. \tag{10.2}$$

Then we can estimate

$$\begin{aligned}
&\int_U \frac{\partial}{\partial t}\left(\zeta^2_{R(t)}\right)\Big|_{t=t_0} d\mu \\
&\le \frac{1}{2}\int_U |\vec{H}|^2 \zeta^2_{R(t_0)} d\mu + \frac{\lambda n}{2R(t_0)^2}\int_U \zeta^2_{R(t_0)} d\mu + C\lambda(\lambda+1)^3\gamma^2 R(t_0)^{n-2}.
\end{aligned} \tag{10.3}$$

Proof. Fix a $t_0 \in (t_1 - \lambda^{-1}R_1^2, \infty)$ such that (10.1), (10.2) hold and set$R = R(t_0)$. By definition 10.1 and $R(t)$ we can calculate

$$\begin{aligned}
2\frac{\partial}{\partial t}\Big|_{t=t_0} \zeta^2_{R(t)}(x) &= 4\zeta_R(x)\zeta'\left(R^{-1}|\hat{x}|\right)|\hat{x}|\frac{\partial}{\partial t}\Big|_{t=t_0} R(t)^{-1} \\
&= -2\zeta_R(x)\zeta'\left(R^{-1}|\hat{x}|\right)|\hat{x}|R^{-3}\lambda = -R^{-2}\lambda D\zeta_R^2(x)\cdot\hat{x}
\end{aligned}$$

for all $x \in \mathbb{R}^{n+k}$. At a point x where the approximate tangent space exists we can use Remark 2.6.4 to obtain

$$\operatorname{div}_\mu(\zeta_R^2\hat{x}) = \zeta_R^2 \operatorname{div}_\mu(\hat{x}) + \nabla^\mu \zeta_R^2 \cdot \hat{x} \le n\zeta_R^2 + \nabla^\mu\zeta_R^2\cdot\hat{x}.$$

So we can estimate

$$\begin{aligned}
&\frac{\partial}{\partial t}\Big|_{t=t_0} \zeta^2_{R(t)}(x) = -\frac{\lambda}{2R^2}\left(\nabla^\mu\zeta_R^2\cdot\hat{x} + \left(D\zeta_R^2 - \pi_{T_x\mu}(D\zeta_R^2)\right)\cdot\hat{x}\right) \\
&\le \frac{\lambda}{2R^2}\left(n\zeta_R^2 - \operatorname{div}_\mu(\zeta_R^2\hat{x}) - \left(D\zeta_R^2 - \pi_{T_x\mu}(D\zeta_R^2)\right)\cdot(\hat{x} - \pi_{T_x\mu}(\hat{x}))\right)
\end{aligned}$$

for almost every $x \in U$, where we used $D\zeta_R^2 - \pi_{T_x\mu}(D\zeta_R^2) \in T_x^\perp\mu$. Then by Definition 2.5 and (10.1) we have

$$\begin{aligned}
\int_U \frac{\partial}{\partial t}\left(\zeta^2_{R(t)}\right)\Big|_{t=t_0} d\mu \le \frac{\lambda}{2R^2}\bigg(&\int_U \zeta_R^2(x)\vec{H}(x)\cdot\hat{x}\, d\mu(x) + n\int_U \zeta_R^2 d\mu \\
&+ \int_U |\pi_{\mathbb{R}^n} - \pi_{T_x\mu}|^2\, |D\zeta_R^2(x)||\hat{x}| d\mu(x)\bigg),
\end{aligned} \tag{10.4}$$

where we used $D\zeta_R^2 = \pi_{\mathbb{R}^n}(D\zeta_R^2)$. Note that by (10.1) we can treat ζ_R like a function in $C_c^\infty(U)$ here. By Theorem 2.7 we can use Remark A.7.1 and combine this with Young's inequality to obtain

$$\begin{aligned}
\int_U \zeta_R^2(x)\vec{H}(x)\cdot\hat{x}\, d\mu(x) &= \int_U \zeta_R^2(x)\vec{H}(x)\cdot(\pi_{\mathbb{R}^n} - \pi_{T_x\mu})\hat{x}\, d\mu(x) \\
&\le \frac{R^2}{2\lambda}\int_U \zeta_R^2|\vec{H}|^2 d\mu + \frac{\lambda}{2}\int_U |\pi_{\mathbb{R}^n} - \pi_{T_x\mu}|^2\, \zeta_R^2 d\mu,
\end{aligned}$$

where we estimated $|\hat{x}| \le R$ as $\text{spt}\zeta_R \subset C_R(0)$. Note that $\zeta_R^2 \le \zeta_R^{\frac{3}{2}}$ and by Proposition A.6.1 also $|D\zeta_R^2||\hat{x}| \le 2\zeta\sqrt{|D^2\zeta_R|\zeta}R \le 2\sqrt{\sigma_1}\zeta_R^{\frac{3}{2}}$ for all $x \in \text{spt}\zeta_R$. Thus (10.4) becomes

$$\begin{aligned}\int_U \frac{\partial}{\partial t}\left(\zeta_{R(t)}^2\right)\Big|_{t=0} d\mu \le \frac{1}{4}\int_U \zeta_R^2|\vec{H}|^2 d\mu + \frac{n\lambda}{2R^2}\int_U \zeta_R^2 d\mu \\ + \frac{\lambda(4\sqrt{\sigma_1}+\lambda)}{4R^2}\int_U \left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2 |\zeta_R^{\frac{3}{2}} d\mu.\end{aligned} \tag{10.5}$$

Using Lemma A.13 with $f = \zeta_R$, $g = \zeta_R^{\frac{3}{4}}$ and $h = \zeta_R^{\frac{3}{2}}$, we have

$$\begin{aligned}&\int_U \left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2 \zeta_R^{\frac{3}{2}} d\mu \\ &\le C_n\left(\int_U |\vec{H}|^2\zeta_R^2 d\mu \int_U |\pi_{\mathbb{R}^k}(x)|^2\zeta_R^3 d\mu\right)^{\frac{1}{2}} + C_n\int_U |\pi_{\mathbb{R}^k}(x)|^2|\nabla^\mu\zeta_R^{\frac{3}{4}}|^2 d\mu.\end{aligned}$$

By Proposition A.6.1 we can estimate $|\nabla^\mu\zeta_R^{\frac{3}{4}}| \le \zeta^{-\frac{1}{4}}\sqrt{|D^2\zeta_R|\zeta} \le \sqrt{\sigma_1}R^{-1}\zeta_R^{\frac{1}{4}}$. Then with $2\sqrt{ab} \le a+b$ we obtain

$$\begin{aligned}&\int_U \left|\pi_{\mathbb{R}^n} - \pi_{T_x\mu}\right|^2 \zeta_R^{\frac{3}{2}} d\mu \\ &\le \frac{R^2}{\lambda(4\sqrt{\sigma_1}+\lambda)}\int_U |\vec{H}|^2\zeta_R^2 d\mu + C_n R^{-2}(\lambda(1+\lambda)+1)\int_U |\pi_{\mathbb{R}^k}(x)|^2\zeta_R^{\frac{1}{2}} d\mu,\end{aligned}$$

where we used $\zeta_R^3 \le \zeta_R^{\frac{1}{2}}$ and σ_1 is constant. Inserting into (10.5) yields

$$\begin{aligned}\int_U \frac{\partial}{\partial t}\left(\zeta_{R(t)}^2\right)\Big|_{t=0} d\mu \le \frac{1}{2}\int_U \zeta_R^2|\vec{H}|^2 d\mu + \frac{n\lambda}{2R^2}\int_U \zeta_R^2 d\mu \\ + C_n R^{-4}\lambda(1+\lambda)(\lambda(1+\lambda)+1)\int_{C_{R(t_0)}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu,\end{aligned} \tag{10.6}$$

where we used $\text{spt}\zeta_R \subset C_R(0)$. Finally we can estimate

$$C_n R^{-4}\lambda(1+\lambda)(\lambda(1+\lambda)+1) \le C_n\lambda(1+\lambda)^3 R^{-4}.$$

Then with assumption (10.2) and as $R = R(t_0)$ inequality (10.6) establishes the result. $\square$

Now we can prove a bound for the measures inside expanding cylinders. This is from [B, 6.5]. Here we give some more details and explicitly state how the cylinder growth effects the measure bound.

10.4 Proposition ([B, 6.5]). *There exists a constant $C \in (1,\infty)$ such that for all $R_1, \lambda, \gamma^2 \in (0,\infty)$, $t_1 \in \mathbb{R}$, $t_2 \in (t_1,\infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[t_1,t_2]}$ be a Brakke flow in U with*

$$\bigcup_{t\in[t_1,t_2]} \operatorname{spt}\mu_t \cap C_{R(t_2)}(0) \subset\subset U, \tag{10.7}$$

$$\sup_{t\in[t_1,t_2]} R(t)^{-n-2} \int_{C_{R(t)}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \le \gamma^2, \tag{10.8}$$

where $R(t) := \sqrt{R_1^2 + \lambda(t-t_1)}$. Then

$$R(t)^{-n}\mu_t\left(\zeta^2_{R(t)}\right) \le R_1^{-n}\mu_0\left(\zeta^2_{R_1}\right) + C(\lambda^{-1}+\lambda^3)\gamma^2 \log\left(R_1^{-1}R(t)\right) \tag{10.9}$$

for all $t \in [t_1,t_2]$.

Proof. For $t \in [t_1,t_2]$ set

$$E(t) := R(t)^{-n}\int_U \zeta^2_{R(t)} d\mu_t.$$

We want to derive a differential inequality for $E(t)$. By Lemma 10.3 with $t_0 = s$, and Lemma 3.10 with $\phi = \zeta$ we can estimate

$$\mathscr{B}\left(\mu_s, \zeta^2_{R(s)}\right) \le -\int_U \frac{\partial}{\partial t}\zeta^2_{R(s)} d\mu_s + \frac{\lambda n}{2}R(s)^{-2}E(s) + C_n(1+\lambda)^4\gamma^2 R(s)^{n-2}$$

for almost every $s \in [t_1,t_2]$. Thus by Proposition 3.8 we obtain

$$\overline{D}\mu_s(\zeta^2_{R(s)}) \le \frac{\lambda n}{2}R(s)^{-2}E(s) + C_n(1+\lambda)^4\gamma^2 R(s)^{n-2}$$

for almost every $s \in [t_1,t_2]$. Then with $R'(s) = \frac{\lambda}{2}R(s)^{-1}$ we conclude

$$\begin{aligned}\overline{D}E(s) &= -nR(s)^{-n-1}R'(s)\mu_s(\zeta^2_{R(s)}) + R(s)^{-n}\overline{D}\mu_s(\zeta^2_{R(s)})\\ &\le C_n(1+\lambda)^4\gamma^2 R(s)^{-2} = C_n(\lambda^{-1}+\lambda^3)\gamma^2 R'(s)R(s)^{-1}\end{aligned}$$

for almost every $s \in [t_1,t_2]$. Now this differential inequality can be integrated using Proposition A.19 to see

$$\begin{aligned}E(t) - E(t_1) &\le \int_{t_1}^t \overline{D}E(s)ds\\ &\le C_n(\lambda^{-1}+\lambda^3)\gamma^2 \int_{t_1}^t R'(s)R(s)^{-1}ds\\ &= C_n(\lambda^{-1}+\lambda^3)\gamma^2 \int_{t_1}^t \frac{\partial}{\partial s}\log(R(s))ds\\ &= C_n(\lambda^{-1}+\lambda^3)\gamma^2 \log\left(R_1^{-1}R(t)\right)\end{aligned}$$

for all $t \in [t_1,t_2]$ and by definition of $E(t)$ this establishes the result. $\square$

Now this can be combined with the clearing out lemma 5.7 to show that small holes become larger.

10.5 Lemma. *There exist constants $C \in (1, \infty)$ and $\gamma_1 \in (0,1)$such that, for all $\delta \in (0,1]$, $\rho \in (0,\infty), \gamma_0 \in (0,\gamma_1]$, $\sigma = \frac{1}{n+12}$, for all $s_1, s_2 \in \mathbb{R}$ with $s_2 - s_1 \in (C\gamma^{2\sigma}\rho^2, \rho]$ and every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let $(\mu_t)_{t\in[s_1,s_2]}$ be a Brakke flow in U with*

$$\mathrm{spt}\mu_t \cap C_{3\rho}(0) \subset B_{4\rho}(0) \subset\subset U, \tag{10.10}$$

$$\int_{C_{3\rho}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \leq \gamma_0^2 \delta^{n+3}\rho^{n+2}, \tag{10.11}$$

$$\mu_{s_1}\left(C_{\delta\rho}(0)\right) = 0 \tag{10.12}$$

for all $t \in [s_1, s_2]$. Then

$$\mathrm{spt}\mu_t \cap B_{3\rho}(0) \cap C_\rho(0) = \emptyset \tag{10.13}$$

for all $t \in [s_1 + C\gamma^{2\sigma}\rho^2, s_2]$.

Proof. Consider an arbitrary $s_0 \in [s_1 + \gamma^{2\sigma}\rho^2, s_2]$. Set $\lambda := (9\rho^2 - \delta^2\rho^2)(s_0 - s_1)^{-1}$ and $\rho(t) := \sqrt{\delta^2\rho^2 + \lambda|t - s_1|}$. Then we see $\rho(s_0) = 3\rho$ and by (10.11)

$$\rho(t)^{-n-2}\int_{C_{\rho(t)}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \leq \delta^{-n-2}\rho^{-n-2}\int_{C_{3\rho}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \leq \delta\gamma_0^2$$

for all $t \in [s_1, s_0]$. Thus we can use Proposition 10.4 with $t_1 = s_1$, $t_2 = s_0$, $\gamma^2 = \delta\gamma_0^2$ and $R_1 = \delta\rho$ to obtain

$$(3\rho)^{-n}\mu_{s_0}(\zeta_{3\rho}^2) \leq (\delta\rho)^{-n}\mu_{s_1}(\zeta_{\delta\rho}^2) + C_n(\lambda^{-1} + \lambda^3)\delta\gamma_0^2\log(\delta^{-1}).$$

As $C_{2r} \subset \{\zeta_{3r} = 1\} \subset \mathrm{spt}\zeta_{3r} \subset C_{3r}$ and with (10.12) we obtain

$$\rho^{-n}\mu_{s_0}(C_{2\rho}(0)) \leq C_n(\lambda^{-1} + \lambda^3)\gamma_0^2$$

By choice of s_0 and λ we have $1 \leq \lambda \leq C_n\gamma_0^{-2\sigma} \leq C_n\gamma_0^{-\frac{1}{6}}$ so we can estimate

$$\rho^{-n}\mu_{s_0}(C_{2\rho}(0)) \leq C_n\gamma_0^{2-\frac{1}{2}} \leq C_n\gamma_0. \tag{10.14}$$

Now we can use Lemma 5.7 with $R = 4\rho$, $r = 2\rho$, $x_0 = 0$ and $\eta = C_n\gamma_0$. Note that $\mathrm{spt}\Phi \subset C_{2\rho}(0)$, so (10.14) implies (5.14). Then by Lemma 5.7 we obtain a constant C_0 such that

$$\mathrm{spt}\mu_t \cap B_{R(t)} \cap C_{r(t)} = \emptyset \tag{10.15}$$

for all $t \in [s_0 + C_0\gamma_0^{2\sigma}\rho^2, s_2]$, where $R(t) = \sqrt{16\rho^2 - 4n(t-s_0)}$ and $r(t) = \sqrt{4\rho^2 - 4n(t-s_0)}$. Note that the time interval may be empty. As $\gamma_0 \le \gamma_1$ we can estimate

$$2C_0\gamma_0^{2\sigma}\rho^2 \le (4n)^{-1}\rho^2$$

for γ_1 small depending on constants. Thus we obtain

$$\begin{aligned} R(s_0 + 2C_0\gamma_0^{2\sigma}\rho^2) &\ge 3\rho \\ r(s_0 + 2C_0\gamma_0^{2\sigma}\rho^2) &\ge \rho. \end{aligned}$$

So if $s_0 + 2C_0\gamma_0^{2\sigma}\rho^2 \le s_2$, equality (10.15) yields

$$\mathrm{spt}\mu_t \cap B_{3\rho} \cap C_\rho = \emptyset$$

for $t = s_0 + 2C_0\gamma_0^{2\sigma}\rho^2$. Then as s_0 was arbitrary in $[s_1 + \gamma_0^{2\sigma}\rho^2, s_2]$, we conclude that (10.13) holds for all $t \in [s_1 + (2C_0+1)\gamma_0^{2\sigma}\rho^2, s_2]$ and for $C \ge (2C_0+1)$ this establishes the result. □

To get rid of the bound (10.11) for all time, we can use Remark 3.12 which lets us replace this assumption by a mass bound and a slab condition at the starting time. In this form the statement is very similar to White's expending hole theorem, see [W3, 4.1].

10.6 Proposition (Opening Holes)**.** *There exists a constant $C \in (1,\infty)$ such that, for every $\beta \in (0,1]$ there exists a $h_0 \in (0,1)$ such that, for all $r \in (0,\infty), M \in [1,\infty), h \in (0,h_0]$, $s_0 \in \mathbb{R}$ and $\Lambda = Mh^{-2}$ the following holds: Let $(\mu_t)_{t\in[s_0,s_0+n^{-1}r^2]}$ be a Brakke flow in $B_{3\Lambda r}(0)$ and suppose*

$$\mathrm{spt}\mu_{s_0} \subset \left\{ x \in \overline{B_{3\Lambda r}(0)} : |\pi_{\mathbb{R}^k}(x)| \le M^{-\frac{1}{2}}hr \right\}, \tag{10.16}$$

$$\mu_{s_0}(B_{5r}(0)) \le Mr^n, \tag{10.17}$$

$$\mu_{s_0}(C_{\beta r}(0)) = 0. \tag{10.18}$$

Then

$$\mathrm{spt}\mu_t \cap C_r(0) \cap B_{\Lambda r}(0) = \emptyset \tag{10.19}$$

for all $t \in [s_0 + Ch^\sigma r^2, s_0 + n^{-1}r^2]$, $\sigma = \frac{1}{n+12}$.

Proof. First we want to establish a height bound for later times. Fix an arbitrary $v \in \{0\}^n \times \mathbb{R}^k$ with $|v| = 1$ and set $x_0 = M^{-\frac{1}{2}}hrv$. Then (10.16) yields

$$\mathrm{spt}\mu_{t_0} \subset \left\{ x \in \overline{B_{2\Lambda}(0)} : (x - x_0)\cdot v \le 0 \right\}. \tag{10.20}$$

By Lemma 3.12.4 with $U = B_{3\Lambda r}(0)$, $t_1 = s_0$, $t_2 = s_0 + r^2$, $R = 2\Lambda r$ and $\delta = M^{-\frac{3}{2}}h^3$ we obtain

$$\mathrm{spt}\mu_t \cap B_{\Lambda r}(0) \subset \left\{x \in B_{2\Lambda r}(0) : x \cdot v \leq M^{-\frac{1}{2}}hr + 2M^{-\frac{3}{2}}h^3\Lambda r\right\} \tag{10.21}$$

for all $t \in [s_0, s_0 + 4(6n)^{-1}M^{-\frac{3}{2}}h^3\Lambda^2 r^2] \cap [s_0, s_0 + n^{-1}r^2]$. Here we estimated $M^{-\frac{3}{2}}h^3 \leq 6^{-1}$, as $M \geq 1$ and $h \leq h_0$, for h_0 small enough. Now as $\Lambda = Mh^{-2}$ we have

$$2M^{-\frac{3}{2}}h^3\Lambda r = 2M^{-1}hr$$
$$4(6n)^{-1}M^{-\frac{3}{2}}h^3\Lambda^2 r^2 \geq (6n)^{-1}M^{\frac{1}{2}}h^{-1}r^2 \geq n^{-1}r^2$$

where we used $M \geq 1$ and $h \leq h_0$, for h_0 small enough. Then as v was arbitrary (10.21) yields

$$\mathrm{spt}\mu_t \cap B_{\Lambda r}(0) \subset \left\{x \in B_{\Lambda r}(0) : |\pi_{\mathbb{R}^k}(x)| \leq 3M^{-\frac{1}{2}}hr\right\} \tag{10.22}$$

for all $t \in [s_0, s_0 + n^{-1}r^2]$. As $3M^{-\frac{1}{2}}h \leq 1$ for $h_0 \leq 3$ this implies

$$\mathrm{spt}\mu_t \cap C_{3r}(0) \cap B_{\Lambda r}(0) \subset B_{4r}(0) \subset\subset B_{\Lambda r}(0) \tag{10.23}$$

for all $t \in [s_0, s_0 + n^{-1}r^2]$.

Now we want to establish a measure bound for later times. By Lemma 3.12.2 with $U = B_{3\Lambda r}(0)$, $t_1 = s_0$, $t_2 = s_0 + r^2$, $x_0 = 0$ $R = 5r$ and $\kappa = \frac{1}{5}$, we can estimate using (10.17)

$$\mu_t\left(B_{4r}(0)\right) \leq C_n \mu_0\left(B_{5r}(0)\right) \leq C_n M r^n \tag{10.24}$$

for all $t \in [s_0, s_0+n^{-1}r^2]$. Combined with (10.22) and (10.23) this establishes the integral height bound

$$\int_{C_{3r}(0)} |\pi_{\mathbb{R}^k}(x)|^2 d\mu_t \leq C_n h^2 r^{n+2} \leq \beta^{n+3} h r^{n+2},$$

for all $t \in [s_0, s_0 + n^{-1}r^2]$, where we used $h \leq h_0$ for h_0 small depending on β. Then we can apply Lemma 10.5 to the restricted flow $(\mu_t \llcorner B_{\Lambda r}(0))$ with $\rho = r$, $s_1 = s_0$, $s_2 = s_0 + n^{-1}r^2$, $\delta = \beta$ and $\gamma_0^2 = h$. Note that due to the restriction (10.23) verifies (10.10).

By Lemma 10.5 we obtain constants C_0 and γ_1 such that, for $\sqrt{h} \leq \sqrt{h_0} \leq \gamma_1$ and $C \geq C_0$ we have

$$\mathrm{spt}\mu_t \cap C_r(0) \cap B_{3r}(0) = \emptyset \tag{10.25}$$

for all $t \in [s_0 + Ch^\sigma r^2, s_0 + n^{-1}r^2]$. Then the result follows from (10.22) and $3M^{-\frac{1}{2}}h \leq 1$ as $h \leq h_0$ for h_0 small enough. $\square$

11 Plane-Like Varifolds

A further application of Brakke's local regularity theorem is to show that Brakke flows become graphical, provided the starting varifold is somehow "plane-like". First we introduce certain parameters that measure how "plane-like "a varifold is, see Definition 11.1. If the starting varifold is "plane-like enough ", that is if the parameters are chosen appropriately, this yields a height bound and the necessary upper area ratio bounds required for Brakke's local regularity theorem 8.4. Now their are two possibilities. Either the lower measure bound holds as well such that the theorem yields regularity, or there exists a cylinder inside of which the flow has very small measure. In this case this cylinder becomes empty after short time, that is void of the flow. This is the main result of this section stated in theorem 11.7.

11.1 Definition. Let $\rho \in (0,\infty), l, h, \xi \in [0,\infty)$, $y_0 \in \mathbb{R}^{n+k}$ and $S \subset B^n_\rho(\hat{y}_0)$. An integral n-varifold μ in $\mathbb{R}^{n+k}$ is called *locally $(\rho, S, l, \gamma, \xi)$-plane-like around y_0*, if there exists a function $f \in C^{0,1}\left(B^n_\rho(\hat{y}_0) \setminus S, \mathbb{R}^k\right)$ with

$$(\mathrm{spt}\mu \cap C_\rho(y_0)) \setminus \left(S \times \mathbb{R}^k\right) = \mathrm{graph}(f), \tag{11.1}$$

such that the following assumptions are satisfied:

$$\mathrm{lip}(f) \leq l, \quad \sup|f - \pi_{\mathbb{R}^k}(y_0)| \leq \gamma\rho, \tag{11.2}$$

$$\mu(S \times \mathbb{R}^k) \leq \xi\rho^n. \tag{11.3}$$

A manifold M is called *locally $(\rho, S, l, \gamma, \xi)$-plane-like around y_0*, if the associated measure $\mu = \mathscr{H}^n \llcorner M$ is locally $(\rho, S, l, \gamma, \xi)$-plane-like around y_0.

11.2 Remark.

1. The varifold is more plane-like the smaller l, γ and ξ are. A small ξ means the varifold is more graph-like, while small l and γ induce flatness.

2. Note that for R, y_0 and μ fixed one can choose different S to obtain plane-likeness with different l, γ, ξ. Choosing S larger might increase ξ but maybe allows smaller l and γ.

3. Let $r, M, \Gamma \in (0,\infty)$ and μ be an integer n-varifold in the slab $\mathbb{R}^n \times B^k_\Gamma(0)$ with $\mu\left(C_R(0)\right) \leq Mr^n$. Suppose there exist $R \in (r,\infty)$, $l \in (0,\infty)$ and a function $f \in C^{0,1}\left(B^n_R(0) \setminus B^n_r(0), B^k_\Gamma(0)\right)$ with $\mathrm{lip}(f) \leq l$ and

$$\mathrm{spt}\mu \cap C_R(0) \setminus C_r(0) = \mathrm{graph}(f).$$

Then μ is $(R, B^n_r(0), l, \Gamma R^{-1}, Mr^nR^{-n})$-plane-like around 0. In particular the last two parameters become arbitrary small for large R.

4. Suppose μ is $(\rho, S, l, \gamma, \xi)$-plane-like around y_0. Then μ is also $(\delta\rho, S \cap B^n_{\delta\rho}(\hat{y}_0), l, \delta^{-1}\gamma, \delta^{-n}\xi)$-plane-like around y_0 for every $\delta \in (0, 1]$.

In case μ has bounded mean curvature the measure $\mu(S \times \mathbb{R}^k)$ in (11.3) can be bounded by $\mathscr{L}^n(S)$ which is often nicer to estimate. This is a direct consequence of the monotonicity formula in [S, 4.3.2].

11.3 Proposition. *For every $p \in (n, \infty)$ there exists a $P \in (1, \infty)$ such that, for all $\rho, M \in (0, \infty)$, $l, \gamma, \alpha \in [0, \infty)$, $\delta \in [0, \frac{1}{2}]$ and every open subset $U \in \mathbb{R}^{n+k}$ the following holds: Let μ be an integral n-varifold in U with generalised mean curvature vector $\vec{H}$. Suppose $\mu(U) \leq M\rho^n$, $\int_U |\vec{H}|^p d\mu \leq \alpha\rho^{n-p}$ and*

$$(\mathrm{spt}\mu \cap C_\rho(0)) \subset \{x \in C_\rho(0),\ |\pi_{\mathbb{R}^k}(x)| \leq \gamma\rho\} \subset\subset U. \tag{11.4}$$

Also suppose there exists a function $f \in C^{0,1}\left(B^n_\rho(0) \setminus B^n_{\delta\rho}(0), B^k_{\gamma\rho}(0)\right)$ with $\mathrm{lip}(f) \leq l$ and

$$\mathrm{spt}\mu \cap C_\rho(0) \setminus C_{\delta\rho}(0) = \mathrm{graph}(f). \tag{11.5}$$

Then μ is $(\rho, B^n_{\delta\rho}(0), l, \gamma, \xi)$-plane-like around 0 for $\xi = P(M+\alpha)(\gamma+\delta)^k\delta^{n-k}$.

11.4 Remark. In case $n > k$ the measure bound ξ becomes arbitrary small for small δ and fixed M, α, γ.

Let us recall the monotonicity formula:

11.5 Theorem (Monotonicity Formula, [S, 4.3.2])**.** *For every $R \in (0, \infty)$, $\alpha \in [0, \infty)$, $p \in (n, \infty)$ and every open subset $U \subset \mathbb{R}^{n+k}$ with $B_R(0) \subset\subset U$ the following holds: Let μ be an integral n-varifold in U with mean curvature vector $\vec{H}$ and suppose $\int_{B_R(0)} |\vec{H}|^p\, d\mu \leq \Gamma^p$, then*

$$\left(r^{-n}\mu\left(B_r(0)\right)\right)^{\frac{1}{p}} - \left(R^{-n}\mu\left(B_R(0)\right)\right)^{\frac{1}{p}} \leq \frac{\Gamma}{p-n}\left(R^{\frac{p-n}{p}} - r^{\frac{p-n}{p}}\right)$$

for all $r \in (0, R)$.

Proof of Proposition 11.3. First we need to cover the set $B^n_{\delta\rho}(0) \times B^k_{\gamma\rho}(0)$ by balls. There exists an $N \in \mathbb{N}$ with $N - 1 \leq \gamma\delta^{-1} \leq N$. Define points $x_a := \sum_{j=1}^k a_j\delta\rho\mathbf{e}_{n+j}$ for $a = (a_1, \ldots, a_k) \in A := \{-N, \ldots, N\}^k$. Then for $r := \sqrt{1+k}\delta\rho$ we have

$$B^n_{\delta\rho}(0) \times B^k_{\gamma\rho}(0) \subset \bigcup_{a \in A} B_r(x_a). \tag{11.6}$$

To see this let $y \in B^n_{\delta\rho}(0) \times B^k_{\gamma\rho}(0)$. For $j \in \{1, \dots, k\}$ choose any $a_j \in [\delta^{-1}\rho^{-1}y_{n+j} - 1, \delta^{-1}\rho^{-1}y_{n+j} + 1] \cap \{-N, \dots, N\}$, then

$$\begin{aligned}|y - x_a| &= \sqrt{|\hat{y}|^2 + \sum_{j=1}^{k} |y_{n+j} - a_j\delta\rho|^2} \\ &\le \delta\rho\sqrt{1 + \sum_{j=1}^{k} |\delta^{-1}\rho^{-1}y_{n+j} - a_j|^2} \le \sqrt{1+k}\,\delta\rho.\end{aligned}$$

For each ball $B_r(x_a)$ we can use Theorem 11.5 with $R = \rho$ and $\Gamma^p = \alpha\rho^{n-p}$, to estimate

$$\left(r^{-n}\mu\left(B_r(x_a)\right)\right)^{\frac{1}{p}} \le M^{\frac{1}{p}} + \frac{\alpha^{\frac{1}{p}}\rho^{\frac{n-p}{p}}}{p-n}\rho^{\frac{p-n}{p}},$$

so

$$\begin{aligned}\mu\left(B_r(x_a)\right) &\le r^n\left(M^{\frac{1}{p}} + \frac{\alpha^{\frac{1}{p}}}{n-p}\right)^p \le 2^p r^n (M + (n-p)^{-p}\alpha) \\ &\le C_p(M+\alpha)\delta^n\rho^n\end{aligned}$$

for all $a \in A$ and some $C_p \in (1, \infty)$ depending on n and p, where we used $r := \sqrt{1+k}\,\delta\rho$. Then with (11.6) and (11.4) we obtain

$$\begin{aligned}\mu\left(C_{\delta\rho}(0)\right) &\le \sum_{a\in A}\mu\left(B_r(x_a)\right) \le C_p|A|(M+\alpha)\delta^n R^n \\ &\le C_p N^k(M+\alpha)\delta^n R^n \le C_p(M+\alpha)(\gamma\delta^{-1}+1)^k\delta^n R^n,\end{aligned}$$

where we used $N \le 1 + \gamma\delta^{-1}$. For P large depending on n, k and p, this establishes the result. $\square$

As we showed in section 5, the clearing out lemma can be used to obtain a height bound, if the starting varifold has small measure above a certain height. This can easily be applied to Brakke flows starting from a locally plane-like varifold with small ξ.

11.6 Lemma. *There exists a $C \in (1, \infty)$ such that for $\sigma := \frac{1}{n+6}$ for all $\rho \in (0, \infty)$, $l \in [0, \infty)$, $\gamma, \xi \in (0, 1)$ $\tau \in (0, (16n)^{-1})$, $s_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+k}$ and every $S \subset B^n_\rho(\hat{y}_0)$ the following holds: Let $(\mu_t)_{t\in[s_0, s_0+\tau\rho^2]}$ be a Brakke flow in $C_\rho(y_0)$. Suppose μ_{t_0} is locally $(\rho, S, l, \gamma, \xi)$-graph-like around y_0 then*

$$\left(\mathrm{spt}\mu_t \cap C_{\frac{\rho}{2}}(y_0)\right) \subset \left\{x \in B_{\frac{\rho}{2}}(y_0) : |\pi_{\mathbb{R}^k}(x - y_0)| \le (16n\tau + \gamma)\rho\right\} \quad (11.7)$$

for all $t \in [s_0 + C\xi^{2\sigma}\rho^2, \tau\rho^2]$. Note that this interval is empty unless ξ is sufficiently small.

Proof. We may assume $y_0 = 0$. As μ_{s_0} is locally $(\rho, S, l, \gamma, \xi)$-graph-like around 0 we can use $\sup |f| \le \gamma\rho$ and (11.1) to see

$$\mu_{s_0}\left(C_\rho(0) \setminus (B^n_\rho(0) \times B^k_{\gamma\rho}(0))\right) \le \mu_{s_0}\left(C_\rho(0) \cap (S \times \mathbb{R}^k)\right)$$

thus by (11.3) we have

$$\mu_{s_0}\left(C_\rho(0) \setminus (B^n_\rho(0) \times B^k_{\gamma\rho}(0))\right) \le \xi\rho^n$$

then using Lemma 5.9 with $\eta = \xi$, $\delta = 4\tau$, $R = \frac{\rho}{2}$, $t_0 = s_0$ and $\Gamma = \gamma\rho$ yields the result. □

11.7 Theorem (Local Graph Or Hole Alternative)**.** *There exist constants $C \in (1, \infty)$, $l, \delta_0 \in (0, 1)$ such that for all $\delta \in (0, \delta_0)$, $\rho \in (0, \infty)$, $\gamma \in (0, \delta^{\frac{3}{2}})$, $\xi \in (0, (\delta_0\delta)^{n+6})$, $s_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+k}$ and every $S \subset B^n_\rho(\hat{y}_0)$ the following holds:*

Let $(\mu_t)_{t\in[s_0,s_0+3\delta^2\rho^2]}$ be a Brakke flow in $C_\rho(y_0)$ and suppose μ_{s_0} is locally $(\rho, S, l, \gamma, \xi)$-plane-like around y_0. Then (at least) one of the following two statements holds:

1. *For $I := [s_0 + \delta^2\rho^2, s_0 + 2\delta^2\rho^2]$ and an $f \in C^\infty\left(I \times B^n_{\delta_0\delta\rho}(\hat{y}_0), \mathbb{R}^k\right)$, with $\mathrm{lip}(f) \le C\sqrt{\delta}$, $\sup |f| \le C\delta^{\frac{3}{2}}\rho$ and*

$$\mathrm{spt}\mu_t \cap C_{\delta_0\delta\rho}(y_0) = \mathrm{graph}(f(t, \cdot))$$

   for all $t \in I$.

2. *$\mu_{s_0+3\delta^2\rho^2}\left(C_{\delta_0\delta\rho}(y_0)\right) = 0$.*

11.8 Remark.

1. Let C be the constant from A.4, then the only conditions on l are $l \le 2$ and $\sqrt{1 + (kC)l^2} < \frac{8}{5}$. With slight modifications in the proof, already $\sqrt{1 + (kC)l^2} < 2$ would be enough. In particular in the case $k = 1$, we can choose l close to $\sqrt{3}$. Note that for higher l we may need smaller δ.

2. Even for Brakke flows starting from a smooth locally graphical manifold this is an interesting result, as it is not clear that such flows stay graphical at all.

3. Both alternatives can be true for the same flow, for example if the flow first becomes graphical and then vanishes abruptly. Actually this is the only example we can think of.

Proof. We may assume $y_0 = 0$. Set $2\sigma := \frac{2}{n+6} = \left(\frac{n}{2}+3\right)^{-1}$. Let $\delta \in (0, \delta_0]$, $\rho \in (0,\infty)$, $\gamma \in (0, \delta^{\frac{3}{2}})$, $\xi \in (0, (\delta_0\delta)^{n+6})$, $s_0 \in \mathbb{R}$ and $S \subset B^n_\rho(0)$ be given. First we want to establish a height bound. As μ_{s_0} is locally $(\rho, S, l, \gamma, \xi)$-plane-like we can use Lemma 11.6 with $y_0 = 0$ and $\tau = \delta^{\frac{3}{2}}$ to obtain

$$\left(\mathrm{spt}\mu_t \cap C_{\frac{\rho}{2}}(0)\right) \subset \left\{x \in C_{\frac{\rho}{2}}(0) : |\pi_{\mathbb{R}^k}(x)| \leq (16n\delta^{\frac{3}{2}} + \gamma)\rho\right\} \tag{11.8}$$

for all $t \in [s_0 + C_n\xi^{2\sigma}\rho^2, s_0 + 3\delta^2\rho^2]$, where we used $\delta^{\frac{3}{2}} \leq (16n)^{-1}$ and $\delta^{\frac{3}{2}} \geq 3\delta^2$, as $\delta \leq \delta_0$ for δ_0 small enough. Here we had to extend the Brakke flow by setting $\mu_t = \emptyset$ for all $t \in (t_0 + 3\delta^2\rho^2, \delta^{\frac{3}{2}}\rho^2]$, to use Lemma 11.6 with our choice of τ. Consider some constant $\Lambda_0 \in (1,\infty)$, which we will fix later depending only on n and k. Set

$$R_0 := (2\Lambda_0)^{-\frac{1}{2}}\delta\rho$$

and use $\xi^{2\rho} \leq \delta_0\delta^2$, $\gamma \leq \delta^{\frac{3}{2}}$ to obtain

$$\left(\mathrm{spt}\mu_t \cap C_{\frac{\rho}{2}}(0)\right) \subset \left\{x \in B_{\frac{\rho}{2}}(0) : |\pi_{\mathbb{R}^k}(x)| \leq C_2\sqrt{\Lambda_0\delta}R_0\right\} \tag{11.9}$$

for all $t \in [s_0 + C_1\delta_0\Lambda_0R_0^2, s_0 + 6\Lambda_0R_0^2]$ for constants $C_1, C_2 \in (1,\infty)$

Next we want to establish a measure bound. As μ_{s_0} is locally $(\rho, S, l, \gamma, \xi)$-plane-like around 0, there exists a function $g \in C^{0,1}\left(B^n_\rho(0) \setminus S, \mathbb{R}^k\right)$ with $\mathrm{lip}(g) \leq l$ and $\sup|g| \leq \gamma\rho$. Set $G(\hat{y}) := (\hat{y}, g(\hat{y}))$, by Proposition A.4 the Jacobian of G is bounded by

$$JG(\hat{y}) \leq \sqrt{1 + C_nl^2}$$

for all $\hat{y} \in B^n_\rho(0)$. Then with (11.1), (11.3) and by $\xi \leq \delta^{n+6}$ we obtain

$$\begin{aligned}\mu_{s_0}(B_r(0)) &\leq \mu_{s_0}\left(C_r(0) \setminus \left(S \times \mathbb{R}^k\right)\right) + \mu_{s_0}\left(S \times \mathbb{R}^k\right) \\ &\leq \omega_nr^n\sqrt{1 + C_nl^2} + \delta^{n+6}\rho^n\end{aligned} \tag{11.10}$$

for all $r \in (0,\rho]$. Note that $\sqrt{24n\Lambda_0}R_0 = (12n)^{-\frac{1}{2}}\delta\rho \leq \rho$, as $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. Then applying (11.10) with $r = \sqrt{24n\Lambda_0}R_0$ yields

$$\begin{aligned}\mu_{s_0}\left(B_{\sqrt{24n\Lambda_0}R_0}(0)\right) &\leq \omega_n(\sqrt{24n\Lambda_0})^nR_0^n\sqrt{1 + C_nl^2} + \delta^{n+6}\rho^n \\ &\leq C_n(1 + \delta^6)\Lambda_0^{\frac{n}{2}}R_0^n, \leq C_n\Lambda_0^{\frac{n}{2}}R_0^n\end{aligned}$$

where we used $R_0 = C_n\Lambda^{-\frac{1}{2}}\delta\rho$ and $l, \delta \leq 1$. Thus by Remark 3.12.2 with $R = \sqrt{24n\Lambda_0}R_0$ and $\kappa = \frac{1}{2}$ we can conclude

$$\begin{aligned}R_0^{-n}\mu_t(B_{2R_0}(0)) &\leq R_0^{-n}\mu_t\left(B_{\sqrt{6n\Lambda_0}R_0}(0)\right) \\ &\leq C_nR_0^{-n}\mu_{s_0}\left(B_{\sqrt{24n\Lambda_0}R_0}(0)\right) \leq C_3\Lambda_0^{\frac{n}{2}}\end{aligned} \tag{11.11}$$

for all $t \in [s_0, s_0 + 6\Lambda_0 R_0^2]$ and a constant $C_3 \in (1, \infty)$, where we used $\Lambda_0 \geq 1$.

Consider

$$t_1 := s_0 + C_1 \delta_0 \Lambda_0 R_0^2,$$
$$\lambda_0 := 2^{-n-4}.$$

We want to show $\mu_{t_1} B_{(1+2\lambda_0)R_0}(0) \leq \frac{3}{2}$. By choice of λ_0 we see $(1+4\lambda_0)R_0 \leq 2\Lambda_0^{-\frac{1}{2}}\delta\rho \leq \rho$, as $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. Then applying (11.10) with $r = (1+4\lambda_0)R_0$ yields

$$\begin{aligned}\mu_{s_0}\left(B_{(1+4\lambda_0)R_0}(0)\right) &\leq \omega_n(1+4\lambda_0)^n R_0^n \sqrt{1+C_n l^2} + \delta^{n+6}\rho^n \\ &\leq \omega_n(1+2^{-2})\sqrt{1+C_n l^2} R_0^n + \delta_0^6 \Lambda_0^{\frac{n}{2}} R_0^n,\end{aligned}$$

where we used $R_0 \geq \Lambda_0^{-\frac{1}{2}}\delta\rho$ and $\delta \leq \delta_0$. Thus for δ_0 small depending on Λ_0 and l small depending on n and k, we obtain

$$R_0^{-n}\mu_{s_0}\left(B_{(1+4\lambda_0)R_0}(0)\right) \leq \frac{3}{2}\omega_n. \tag{11.12}$$

There exists a cut-off function $\phi \in C_c^\infty\left(B_{(1+4\lambda_0)R_0}(0), [0,1]\right)$ with

$$B_{(1+4\lambda_0)R_0}(0) \supset \mathrm{spt}\phi \supset \{\phi = 1\} \supset B_{(1+2\lambda_0)R_0}(0) \tag{11.13}$$

and $\sup|D^2\phi| \leq C_n R_0^{-2}$. By (3.1) in view of Proposition A.19 and Remark 3.2.2 in view of $\sup|D^2\phi| \leq C_n R_0^{-2}$ we can estimate

$$\begin{aligned}\mu_{t_1}(B_{(1+2\lambda_0)R_0}(0)) &\leq \mu_{t_1}(\phi) \leq \mu_{s_0}(\phi) + \int_{s_0}^{t_1} \mathscr{B}(\mu_t, \phi)dt \\ &\leq \mu_{s_0}(\phi) + C_n\delta_0\Lambda_0 \sup_{t\in[s_0,t_1]} \mu_t(\mathrm{spt}\phi) \\ &\leq \mu_{s_0}\left(B_{(1+4\lambda_0)R_0}(0)\right) + C_n\delta_0\Lambda_0 \sup_{t\in[s_0,t_1]} \mu_t(B_{(1+4\lambda_0)R_0}(0)),\end{aligned}$$

where we used $t_1 - s_0 = C_1\delta_0\Lambda_0 R_0^2$ and (11.13). By choice of λ_0 we have $(1+4\lambda_0)R_0 \leq 2R_0$ and for δ_0 small enough we have $t_1 \leq 6\Lambda_0 R_0^2$. Then with (11.11) and (11.12) we obtain

$$R_0^{-n}\mu_{t_1}(B_{(1+2\lambda_0)R_0}(0)) \leq \frac{3}{2}\omega_n + C_n\delta_0\Lambda_0^{1+\frac{n}{2}} \leq \frac{7}{4}\omega_n \leq (2-\lambda_0)\omega_n, \tag{11.14}$$

where we used $\lambda_0 = 2^{-n-4}$ and we chose δ_0 small depending on Λ_0

If we would have a lower bound on $\mu_{s_0+5\Lambda_0 R_0^2}(B_{R_0}(0))$, then we could apply the local regularity theorem. Let $C_{5.7}$ be the constant from Lemma 5.7. We set

$$\lambda := \min\{\lambda_0, \omega_n^{-1}(8nC_{5.7})^{-\frac{1}{2\sigma}}\},$$
$$t_2 := s_0 + 6\Lambda_0 R_0^2 - (8n)^{-1}R_0^2.$$

Case 1: Suppose

$$R_0^{-n}\mu_{t_2}(B_{R_0}(0)) \geq \lambda\omega_n. \tag{11.15}$$

Then we can apply Theorem 8.4 for our flow restricted to B_{2R_0} with $R = R_0$, $\alpha = \frac{1}{2}$, $t_1 = s_0 + C_1\delta_0^2\Lambda_0 R_0^2$, $t_2 = s_0 + 6\Lambda_0 R_0^2 - (8n)^{-1}R_0^2$, $M = C_3\Lambda_0^{\frac{n}{2}}$ and $h = C_2\sqrt{\Lambda_0}\delta$. To see this let Λ depending on n and k and h_0 depending on Λ_0 be from Theorem 8.4 corresponding to our choice of λ and M. Note that as $\lambda \leq \lambda_0$, estimate (11.11) implies (8.104) by choice of M and (11.14) combined with (11.15) imply (8.105). Moreover our height estimate (11.9) implies (8.103), which also uses $\frac{\rho}{2} \geq (1+2\lambda)(2\Lambda_0)^{-\frac{1}{2}}\delta\rho = (1+2\lambda)R_0$, as $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. Furthermore if $\Lambda_0 \geq \Lambda$ we can estimate

$$t_1 + \Lambda R_0^2 = t_0 + (C_1\delta_0^2\Lambda_0 + \Lambda)R_0^2 \leq s_0 + 2\Lambda_0 R_0^2 = s_0 + \delta^2\rho$$
$$t_2 - \Lambda R_0^2 = s_0 + (6\Lambda_0 - \Lambda - (8n)^{-1})R_0^2 \geq s_0 + 4\Lambda_0 R_0^2 = s_0 + 2\delta^2\rho,$$

where we used $R_0 = (2\Lambda_0)^{-\frac{1}{2}}\delta\rho$, $\Lambda_0 \geq 1$ and δ_0 small depending on n and k. Thus we choose $\Lambda_0 \geq \Lambda$ and $\delta_0 \leq C_2^{-2}h_0^2\Lambda_0^{-1}$, then Theorem 8.4 yields a smooth function $f \in C^\infty\left(I \times C_{\lambda R_0}(0), \mathbb{R}^k\right)$, with $\operatorname{lip}(f) \leq C_n\sqrt{\Lambda_0}\delta$, $\sup|f| \leq C_n\sqrt{\Lambda_0}\delta R_0$ and

$$\operatorname{spt}\mu_t \cap C_{\lambda R_0}(0) \cap B_{2R_0}(0) = \operatorname{graph}(f(t,\cdot))$$

for all $t \in I$. Here the intersection with $B_{2R_0}(0)$ is because we had to restrict our flow. Actually this intersection is obsolete as in view of (11.9) we see that

$$\operatorname{spt}\mu_t \cap C_{\lambda R_0}(0) \subset B_{2R_0}(0),$$

where we used $R_0 = (2\Lambda)^{-\frac{1}{2}}\delta\rho \leq 2^{-2}\rho$, $\lambda \leq 1$ and $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. As $R_0 = (2\Lambda_0)^{-\frac{1}{2}}\delta\rho$ we can choose δ_0 small and C large enough depending on Λ_0 and λ to establish the result in this case.

Case 2: Suppose

$$R_0^{-n}\mu_{t_2}(B_{R_0}(0)) < \lambda\omega_n.$$

Then we can use Lemma 5.7 with $R = r = R_0$, $t_0 = t_2 = s_0 + 6\Lambda_0 R_0^2 - (8n)^{-1} R_0^2$, $x_0 = 0$ and $\eta = \lambda\omega_n$ to obtain

$$\mu_t \left(B_{R(t)}(0)\right) = 0 \tag{11.16}$$

for all $t \in [t_2 + C_{5.7}(\lambda\omega_n)^{2\sigma} R_0^2, s_0 + 6\Lambda_0 R_0^2]$, where $R(t) = \sqrt{R_0^2 - 4n(t - t_2)}$. By choice of λ and t_2, we have

$$t_2 + C_{5.7}(\lambda\omega_n)^{2\sigma} R_0^2 \leq t_2 - (8n)^{-1} = s_0 + 6\Lambda_0 R_0^2$$

Thus (11.16) holds for $t = s_0 + 6\Lambda_0 R_0^2$ and

$$R(s_0 + 6\Lambda_0 R_0^2) = \sqrt{R_0^2 - 2^{-1} R_0^2} \geq 2^{-1} R_0.$$

As $R_0 = (2\Lambda)^{-\frac{1}{2}} \delta\rho \geq 2\delta_0 \delta\rho$ for δ_0 small depending on Λ_0 this establishes

$$\mu_{s_0 + 3\delta^2\rho^2} \left(B_{2\delta_0\delta\rho}(0)\right) = 0. \tag{11.17}$$

In view of (11.9) we see that

$$\mathrm{spt}\mu_{s_0 + 3\delta^2\rho^2} \cap C_{\delta_0\delta\rho}(0) \subset B_{2\delta_0\delta\rho}(0),$$

where we used $R_0 = (2\Lambda)^{-\frac{1}{2}} \delta\rho$ and $\delta \leq \delta_0$ for δ_0 small depending on Λ_0. So (11.17) establishes the second alternative of our statement. □

The main drawback of Theorem 11.7 is that one only gets regularity, if one can exclude the appearance of empty cylinders, that is cylinders which do not contain the flow. A theorem in [W1] can help overcome this problem under certain conditions. There White describes the topological changes that may appear for a level set flow.

11.9 Definition. Consider $t_1 \in \mathbb{R}$, $t_2 \in (t_1, \infty)$. For a Brakke flow $B = (\mu_t)_{t \in [t_1, t_2]}$ in $\mathbb{R}^{n+1}$ set

$$\mathcal{M}(B) = \left\{(t, x) \in [t_1, t_2] \times \mathbb{R}^{n+1} : x \in \mathrm{spt}\mu_t\right\}.$$

For an integral n-varifold μ set

$$\mathbf{B}(\mu) := \left\{(\mu_t)_{t \in [t_1, t_2]},\ \text{Brakke flow in } \mathbb{R}^{n+1} \text{ with } \mu_{t_1} = \mu\right\}.$$

A closed subset $\mathcal{M} \subset [t_1, t_2] \times \mathbb{R}^{n+1}$ is called a *level set flow* in $\mathbb{R}^{n+1}$, if there exists an integral n-varifold μ such that

$$\mathcal{M} = \bigcup_{B \in \mathbf{B}(\mu)} \mathcal{M}(B).$$

A Brakke flow $B = (\mu_t)_{t \in [t_1, t_2]}$ in $\mathbb{R}^{n+1}$ *corresponds to a level set flow* if $\mathcal{M}(B)$ is a level set flow.

11.10 Theorem (Topological Change For Level Set Flow,[W1, 5.2])**.** *Consider $t_1 \in \mathbb{R}$ and $t_2 \in (t_1, \infty)$. Let $\mathcal{M}$ be a closed subset of space-time $[t_1, t_2] \times \mathbb{R}^{n+1}$ and let $\mathcal{W}$ be its complement. For $s, s_1, s_2 \in [t_1, t_2]$ set*

$$W[s] := \{(t, x) \in \mathcal{W} : t = s\},$$
$$W[s_1, s_2] := \{(t, x) \in \mathcal{W} : s_1 \leq t \leq s_2\}.$$

Suppose $\mathcal{M}$ is a level set flow, then the following holds: If X and Y are in different connected components of $W[t_1]$, then they are in different connected components of $W[t_1, t_2]$.

11.11 Remark.

1. Note that the level set flow is only defined in the case of co-dimension one.
2. For conditions implying that a Brakke flow corresponds to a level set flow, see [I, 11.4].

11.12 Definition. Consider a subset $U \subset \mathbb{R}^m$ for some $m \in \mathbb{N}$. We say two points $a, b \in U$ are *path connected* in U, if there exists a continuous function $\gamma : [0, 1] \to U$ with $\gamma(0) = a$ and $\gamma(1) = b$.

Consider an integral n-varifold μ in $\mathbb{R}^{n+1}$ and $a, b \in \mathbb{R}^{n+1} \setminus \mathrm{spt}\mu$. We say an integral n-varifold μ *separates* a and b, if a and b are not path connected in $\mathbb{R}^{n+1} \setminus \mathrm{spt}\mu$

11.13 Corollary. *There exist constants $C \in (1, \infty)$, $l, \delta_0 \in (0, 1)$ such that for all $\delta \in (0, \delta_0)$, $\rho, \Gamma \in (0, \infty)$, $\gamma \in (0, \delta^{\frac{3}{2}})$, $\xi \in (0, \delta_0 \delta^{n+6})$, $s_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^{n+1}$ and every $S \subset B^n_\rho(\hat{y}_0)$ the following holds:*

Let $(\mu_t)_{t \in [s_0, s_0 + 3\delta^2\rho^2]}$ be a Brakke flow in $\mathbb{R}^n \times [-\Gamma, \Gamma]$ that corresponds to a level set flow. Suppose μ_{s_0} is locally $(\rho, S, l, \gamma, \xi)$-plane-like around y_0 and separates $(\hat{y}_0, -2\Gamma)$ and $(\hat{y}_0, 2\Gamma)$.

Then for $I := [s_0 + \delta^2\rho^2, s_0 + 2\delta^2\rho^2]$ and $r := \delta_0\delta\rho$ there exists a smooth function $f \in C^\infty\left(I \times B^n_r(\hat{y}_0), \mathbb{R}^k\right)$, with $\mathrm{lip}(f) \leq C\sqrt{\delta}$, $\sup |f| \leq C\delta^{\frac{3}{2}}\rho$ and

$$\mathrm{spt}\mu_t \cap C_r(y_0) = \mathrm{graph}(f(t, \cdot))$$

for all $t \in I$.

Proof. By Theorem 11.7 applied to $\mu_t \llcorner C_R(y_0)$ we immediately obtain the result or

$$\mu_{s_0 + 3\delta^2\rho^2}\left(C_{\delta_0\delta\rho}(y_0)\right) = 0, \tag{11.18}$$

which we will lead to a contradiction.

Let $a := (\hat{y}_0, -2\Gamma)$ and $b := (\hat{y}_0, 2\Gamma)$, by (11.18) we see, that a and b are path connected in $\mathbb{R}^{n+1} \setminus \mathrm{spt}\mu_{s_0+3\delta^2\rho^2}$. In particular $(s_0 + 3\delta^2\rho^2, a)$ and $(s_0 + 3\delta^2\rho^2, b)$ belong to the same connected component of $W[s_0 + 3\delta^2\rho^2]$. As the whole flow is contained in $\mathbb{R}^n \times [-\Gamma, \Gamma]$ we have $(s_0 + 3\delta^2\rho^2, a)$ is connected to (s_0, a) and $(s_0 + 3\delta^2\rho^2, b)$ is connected to (s_0, b) in $W[s_0, s_0 + 3\delta^2\rho^2]$, thus (s_0, a) and (s_0, b) belong to the same connected component of $W[s_0, s_0 + 3\delta^2\rho^2]$. As the Brakke flow corresponds to a level set flow we can use Theorem 11.10 to obtain that a and b are path connected in $\mathbb{R}^{n+1} \setminus \mathrm{spt}\mu_{s_0}$ as well. This contradicts the fact that μ_{s_0} separates a and b. □

Now we want to transfer our local results to Brakke flows that are defined in all of $\mathbb{R}^{n+k}$.

11.14 Definition. Let $l, \Gamma, \Xi \in [0, \infty)$ and $S \subset \mathbb{R}^n$. An integral n-varifold μ in $\mathbb{R}^{n+k}$ is called *globally* (S, l, Γ, Ξ)*-plane-like*, if there exists a function $f \in C^{0,1}\left(\mathbb{R}^n \setminus S, \mathbb{R}^k\right)$ with

$$\mathrm{spt}\mu \setminus \left(S \times \mathbb{R}^k\right) = \mathrm{graph}(f), \tag{11.19}$$

such that the following assumptions are satisfied:

$$\mathrm{lip}(f) \le l, \quad \sup |f| \le \Gamma, \tag{11.20}$$

$$\sup_{\hat{y} \in \mathbb{R}^n} \mu\left((B_R^n(\hat{y}) \cap S) \times \mathbb{R}^k\right) \le \Xi R^{n-1} \quad \forall R \in [\Xi, \infty). \tag{11.21}$$

A manifold M_t is called *globally* (S, l, Γ, Ξ)*-plane-like*, if the associated measure $\mu_t = \mathscr{H}^n \llcorner M_t$ is globally (S, l, Γ, Ξ)-plane-like.

11.15 Remark.

1. Let $l, \Gamma, \Xi, \in [0, \infty)$ and $S \subset \mathbb{R}^n$. Suppose μ is globally (S, l, Γ, Ξ)-plane-like. For $\rho \in [\Xi, \infty)$ and $\hat{y} \in \mathbb{R}^n$ estimate

$$\mu\left((B_\rho^n(\hat{y}) \cap S) \times \mathbb{R}^k\right) \le \Xi\rho^{n-1} \le \rho^{-1}\Xi\rho^n,$$

where we used $\rho \ge \Xi$ and (11.21). This implies that μ is locally $(\rho, B_\rho^n(\hat{y}) \cap S, l, \rho^{-1}\Gamma, \rho^{-1}\Xi)$-plane-like around every $y \in \mathbb{R}^n \times \{0\}^k$.

2. For an integral n-varifold μ an $S \subset \mathbb{R}^n$ with $\mu(S \times \mathbb{R}^k) \in [0, \infty)$, we can set $M := \mu(S \times \mathbb{R}^k)$ and estimate

$$\sup_{y \in \mathbb{R}^n} \mu\left((B_R^n(\hat{y}) \cap S) \times \mathbb{R}^k\right) \le \mu(S \times \mathbb{R}^k) = M^{\frac{n}{n}} \le \sqrt[n]{M} R^{n-1}$$

for all $R \in [\sqrt[n]{M}, \infty)$ and all $\hat{y} \in \mathbb{R}^n$. So (11.21) is verified for $\Xi \ge \sqrt[n]{M}$.

3. Property (11.21) basically says that S is $n-1$-dimensional and μ does not concentrate in $S \times \mathbb{R}^k$. In particular (11.21) can hold although $\mu(S \times \mathbb{R}^k) = \infty$.

11.16 Lemma. *There exists a $C \in (1, \infty)$ such that for all $l, \Gamma, \Xi \in [0, \infty)$, $T \in [C\Xi^2, \infty)$ and $S \subset \mathbb{R}^n$ the following holds: Let $(\mu_t)_{t \in [0,T]}$ be a Brakke flow in $\mathbb{R}^{n+k}$. Suppose μ_0 is globally (S, l, Γ, Ξ)-plane-like, then*

$$\mathrm{spt}\mu_t \subset \left\{ x \in \mathbb{R}^{n+k} : |\pi_{\mathbb{R}^k}(x)| \leq C\Xi + \Gamma \right\} \tag{11.22}$$

for all $t \in [C\Xi^2, T]$

Proof. We want to apply Lemma 11.6. Set $\sigma = \frac{1}{n+6}$, let $\rho \in [\Xi, \infty)$ and $y \in \mathbb{R}^n \times \{0\}^k$. By Remark 11.15.1 we have μ_0 is $(\rho, B^n_\rho(\hat{y}) \cap S, l, \rho^{-1}\Gamma, \rho^{-1}\Xi)$-plane-like around y. Thus we can use Lemma 11.6 with $s_0 = 0$, $y_0 = y$, $\gamma = \rho^{-1}\Gamma$, $\xi = \rho^{-1}\Xi$ and $\tau = (32n)^{-1}$ to obtain a $C_1 \in (1, \infty)$ such that

$$\left(\mathrm{spt}\mu_t \cap C_{\frac{\rho}{2}}(y)\right) \subset \left\{ x \in C_{\frac{\rho}{2}}(y) : |\pi_{\mathbb{R}^k}(x)| \leq \left(\frac{1}{2} + \rho^{-1}\Gamma\right)\rho \right\}$$

for all $t \in [C_1 \Xi^{2\sigma} \rho^{-2\sigma+2}, (32n)^{-1}\rho^2]$, if $T \geq (32n)^{-1}\rho^2$. Now choose $\rho := (32nC_1)^{\frac{1}{2\sigma}}\Xi = (32nC_1)^{2n+12}\Xi \in [\Xi, \infty)$, this yields

$$\left(\mathrm{spt}\mu_{t_1} \cap C_{\frac{\rho}{2}}(y)\right) \subset \left\{ x \in C_{\frac{\rho}{2}}(y) : |\pi_{\mathbb{R}^k}(x)| \leq (32nC_1)^{2n+12}\Xi + \Gamma \right\}$$

for $t_1 := (32n)^{-1}\rho^2 = (32n)^{-1}(32nC_1)^{4n+24}\Xi^2$, if $T \geq (32n)^{-1}(32nC_1)^{4n+24}\Xi^2$. Thus for $C := (32nC_1)^{4n+24}$ and as y was arbitrary, we obtain a global slab bound at time t_1

$$\mathrm{spt}\mu_{t_1} \subset \left\{ x \in \mathbb{R}^{n+k} : |\pi_{\mathbb{R}^k}(x)| \leq C\Xi + \Gamma \right\},$$

where we used $\sqrt{C} \leq C$. Now we can use that a global slab bound is maintained forever. Namely use Lemma 3.12.5 with $t_2 = T$, $x_0 = \pm(C\Xi + \Gamma)\mathbf{e}_{n+j}$ and $v = \pm\mathbf{e}_{n+j}$ for all $j \in \{1, \ldots, k\}$. This establishes the result as $t_1 \leq C\Xi^2$. $\square$

11.17 Proposition. *For every $\epsilon \in (0,1)$ there exist $\Lambda \in (1, \infty)$, $l \in (0,1)$ such that, for all $R \in (0, \infty)$, $\Gamma, \Xi \in (0, R]$, $s_0 \in \mathbb{R}$ and $S \subset \mathbb{R}^n$ the following holds: Let $(\mu_t)_{t \in [s_0, s_0 + 3\Lambda R^2]}$ be a Brakke flow in in $\mathbb{R}^{n+k}$ with μ_{s_0} is globally (S, l, Γ, Ξ)-plane-like. Then (at least) one of the following two statements holds:*

1. *For* $I = [s_0 + \Lambda R^2, s_0 + 2\Lambda R^2]$ *there exists an* $f \in C^\infty\left(I \times \mathbb{R}^n, \mathbb{R}^k\right)$ *with* $\mathrm{lip} f \leq \epsilon$ *and*

$$\mathrm{spt}\mu_t = \mathrm{graph}(f(t,\cdot))$$

for all $t \in I$.

2. *There exists a* $y \in \mathbb{R}^{n+k}$ *with* $\mu_{s_0+3\Lambda R^2}\left(C_R(y)\right) = 0$

Proof. We may assume $s_0 = 0$. Take $\delta_0, l \in (0,1)$ from Theorem 11.7. For $\delta \in (0, \delta_0]$ set

$$\rho := \delta^{-2n-12} R$$

and let $y \in \mathbb{R}^n \times \{0\}^k$, then as $\Xi \leq R \leq \rho$ and $\delta \leq 1$ we can use Remark 11.15.1, to see that μ_0 is locally $(\rho, B^n_\rho(\hat{y}) \cap S, l, \delta^{2n+12}, \delta^{2n+12})$-plane-like around y. Thus for every $y \in \mathbb{R}^n \times \{0\}^k$ we can use Theorem 11.7, which leads to one of the following two alternatives

Alternative 1:

For $I_0 = [\delta^2\rho^2, 2\delta^2\rho^2]$, for every $y \in \mathbb{R}^n \times \{0\}^k$ there exists a smooth function $f_y \in C^\infty\left(I \times C_{\delta_0\delta\rho}(y), \mathbb{R}^k\right)$ with $\mathrm{lip}(f_y) \leq C_n\sqrt{\delta}$ and $\mu_t \cap C_{\delta^2\rho}(y)$ for all $t \in I$. Then all these functions have to be parts of just one function f defined on all $\mathbb{R}^n$ with $\mathrm{graph}(f(t,\cdot)) = \mathrm{spt}\mu_t$ for all $t \in I$.

Alternative 2:

There exists a $y \in \mathbb{R}^n \times \{0\}^k$ such that $\mu_{3\delta^2\rho^2}\left(C_{\delta_0\delta\rho}(y)\right) = 0$.

Now fix δ small enough depending on ϵ such that $\mathrm{lip}(f_y) \leq \epsilon$. Then we can choose $\Lambda = \delta^{-4n-22}$ to establish the result. $\delta_0\delta\rho \geq R$ as $\delta \leq \delta_0$ and by definition of ρ. □

12 Graphical Hypersurfaces

Here we want to apply the results from the previous section to smooth mean curvature flow of hypersurfaces in $\mathbb{R}^{n+1}$.

It turns out that for smooth mean curvature flow in a sufficiently narrow slab the gradient decreases. This will be used in the proof of Proposition 12.13, which is similar to Theorem 11.7, but without the small bound on the Lipschitz constant of the starting surface. To compensate for this we have to assume that the flow stays plane-like for a certain time. The key result here is Theorem 12.11, which says that a graphical representation can be extended to later times and additionally yields, that the Lipschitz constant of the extended graphical representation is small, if the flow lies in a narrow enough slab. This is a consequence of White's smooth regularity theorem 12.8, where the idea to use this was suggested to the author by Felix Schulze.

Recall the following definitions:

12.1 Definition. For an open subset $\Omega \subset \mathbb{R}^n$ consider an embedding $F \in C^2(\Omega, \mathbb{R}^{n+1})$ and $M = F(\Omega)$. Let $\nu \in C^2(M, \mathbb{R}^{n+1})$ be a normal field on M. For $p \in \Omega$ and $x = F(p)$ we define:

1. The first fundamental form $g(p) \in \mathbb{R}^{n\times n}$ by

$$\begin{aligned} g_{ij}(p) &:= \frac{\partial}{\partial x_i}F(p) \cdot \frac{\partial}{\partial x_i}F(p), \\ g^{ij}(p) &:= \left(g^{-1}(p)\right)_{ij} \end{aligned}$$

   for all $p \in \Omega$ and all $1 \le i, j \le n$.

2. The second fundamental form $A(x) : \mathbb{R}^{n\times n}$ by

$$\begin{aligned} A_{ij}(x) &:= \frac{\partial^2}{\partial x_i \partial x_j}F(p) \cdot \nu(x), \\ A^i_j(x) &:= \sum_{l=1}^n g^{il}(p) A_{lj}(x), \\ \|A(x)\|^2 &:= \sum_{i,j=1}^n A^i_j(x) A^j_i(x) \end{aligned}$$

   for all $x \in M$ and all $1 \le i, j \le n$.

3. At points $x \in M$ where $\nu(x) \cdot e_{n+1} \neq 0$ we set

$$v(x) := \left(\nu(x) \cdot e_{n+1}\right)^{-1}$$

12.2 Remark. In the above setting with $F(\hat{y}) = (\hat{y}, f(\hat{y}))$ for some $f \in C^2(\Omega, \mathbb{R})$, we have for all $w \in \mathbb{R}^n$, $\hat{y} \in \Omega$ and $y = F(\hat{y})$

$$\begin{aligned} v(y) &= \sqrt{1 + |Df(\hat{y})|^2} \\ |D^2 f(\hat{y})| &\leq \left(1 + |Df(\hat{y})|^2\right)^{\frac{3}{2}} \|A(y)\| \,. \end{aligned}$$

For a proof see for example [CM, 4.1]

The following elementary results show that small curvature of M yields bounds on the first two derivatives of parametrisations of M over its tangent space. The next statement can be found in [CM, 2.4], which we formulate slightly differently, as we do not use intrinsic balls.

12.3 Proposition ([CM, 2.4])**.** *There exist $C \in (1, \infty)$ and $\epsilon_0 \in (0, 1)$ such that, for every $\epsilon \in (0, \epsilon_0]$ and every $R \in (0, \infty)$ the following holds: Let M be an embedded C^2-hypersurface in $\mathbb{R}^{n+1}$ with $0 \in M$ and $T_0 M = \mathbb{R}^n$. Suppose M satisfies $\partial M \cap B_{2R}(0) = \emptyset$ and*

$$\sup_{M \cap B_{2R}(0)} \|A\| \leq \epsilon R^{-1} \tag{12.1}$$

Then there exists a $g \in C^2(B^n_R(0), \mathbb{R})$ with $\operatorname{graph}(g) \subset M$ *and*

$$g(0) = 0, \quad Dg(0) = 0, \tag{12.2}$$

$$\max\{R^{-1} \sup |g|, \sup |Dg|, R \sup |D^2 g|\} \leq C\epsilon. \tag{12.3}$$

Proof. Consider radii $r \in (0, R)$, such an r is called proper, if there exists a $g_r \in C^2(B^n_r(0), \mathbb{R})$ with $g_r(0) = 0$, $Dg_r(0) = 0$, $\operatorname{graph}(g_r) \subset M$ and $E(g_r) \leq 5\epsilon$, where

$$E(g) := \max\{R^{-1} \sup |g|, \sup |Dg|, R \sup |D^2 g|\}.$$

Consider the set

$$I := \{r \in (0, R] : r \text{ is proper}\}$$

By Proposition A.16 and Remark A.17 there exist r_0 and $g_{r_0} \in C^2(B^n_R(0), \mathbb{R})$ with $g_{r_0}(0) = 0$, $Dg_{r_0}(0) = 0$ and $\operatorname{graph}(g_{r_0}) \subset M$. Then by Remark 12.2 and by (12.1) we can estimate

$$|D^2 g_{r_0}(0)| \leq \epsilon R^{-1}$$

and by continuity we can choose r_0 a bit smaller, such that $E(g_{r_0}) \leq 2\epsilon$. Thus there exists an $r_0 \in I$.

Let $(r_m)_{m\in\mathbb{N}}$ be a sequence in I with $r_m \to r \in (0, R]$, we want to show $r \in I$. Let g_m be the corresponding functions. As $g_m(0) = 0$ and $\operatorname{graph}(g_m) \subset M$ for all $m \in \mathbb{N}$, we have

$$g_{m_1}(\hat{y}) = g_{m_2}(\hat{y})$$

for all $\hat{y} \in B^n_{\tilde{r}}(0)$, $\tilde{r} = \min\{r_{m_1}, r_{m_2}\}$ for all $m_1, m_2 \in \mathbb{N}$. Then for every $\hat{y} \in B^n_r(0)$ there exists $m \in \mathbb{N}$ with $\hat{y} \in B^n_{r_m}(0)$. Set $g_r := g_m(\hat{y})$, which is well defined by above considerations. As $E(g_m) \leq 5\epsilon$ for all m we have $E(g_r) \leq 5\epsilon$. By $g_m(0) = 0$ and $Dg_m(0) = 0$ for all m also $g_r(0) = 0$ and $Dg_r(0) = 0$. Furthermore as $\operatorname{graph}(g_m) \subset M$ for all m and as M is closed, we have $\operatorname{graph}(f_r) \subset M$. Thus r is proper.

Now we want to show I is also open (relative in $(0, R]$). As $(0, R]$ is connected, this would establish the result.

Fix $r \in I$. There exists a function $g_r \in C^2(B^n_r(0), \mathbb{R})$ with $g_r(0) = 0$, $Dg_r(0) = 0$, $\operatorname{graph}(g_r) \subset M$ and $E(g_r) \leq 5\epsilon$. Let $x \in \partial(\operatorname{graph}(g_r)) \cap M$. As $\epsilon \leq \epsilon_0$, we can estimate $|g_r(x)| \leq E(g_r)r \leq 5\epsilon r \leq r$, for $\epsilon_0 \leq 5^{-1}$. Then we have $x \in B_{2R}(0)$ and as M contains no boundary points in $B_{2R}(0)$, we know x is an inner point of M. Also as x is close to $\operatorname{graph}(g_r)$ the normal cannot be perpendicular, so

$$\|\pi_{\mathbb{R}^n} - \pi_{T_xM}\|_{op} < 1.$$

Thus we can use Proposition A.16 and Remark A.17.2, to obtain a small δ_x such that, $M \cap B_{\delta_x}(x)$ can be written as a graph over $\mathbb{R}^n$. Then as $x \in \partial(\operatorname{graph}(g_r)) \cap M$ was arbitrary and by compactness of $\partial(\operatorname{graph}(g_r))$ there exists a small $\delta \in (0, 1)$, such that g_r can be extended to some $g_{r+\delta} \in C^2\left(B^n_{r+\delta}(0), \mathbb{R}\right)$ with

$$\operatorname{graph}(g_{r+\delta}) \subset M.$$

It remains to show that $E(g_{r+\delta}) \leq 5\epsilon$. To do this we will show that $E(g_r)$ is actually bounded by $4\epsilon R^{-1}$ which then yields $E(g_{r+\delta}) \leq 5\epsilon$ for small δ.

By Remark 12.2 we have

$$|D^2 g_r(\hat{y}| \leq \left(1 + |Dg(\hat{y})|^2\right)^{\frac{3}{2}} \|A(y)\|$$

for all $\hat{y} \in B^n_r(0)$. Then $E(g_r) \leq 5\epsilon$ and (12.1) yield

$$|D^2 g_r(\hat{y}| \leq 4\epsilon R^{-1},$$

for all $\hat{y} \in B^n_r(0)$, where we used $|Dg(\hat{y})| \leq 5\epsilon \leq 1$, as $\epsilon \leq \epsilon_0$ for $\epsilon_0 \leq 5^{-1}$. Using $g_r(0) = 0$, $Dg_r(0) = 0$ and the mean value formula we obtain

$$|Dg_r(\hat{y}| \leq 4\epsilon, \quad |g_r(\hat{y}| \leq 4\epsilon R$$

for all $\hat{y} \in B^n_r(0)$. Thus $E(g_r) \leq 4\epsilon$, which yields $E(g_{r+\delta}) \leq 5\epsilon$, for δ small enough.

So I is open (relative in $(0, R]$) and as $(0, R]$ is connected this establishes the result. $\square$

In the case where $M = \text{grap}(f)$ Proposition 12.3 can be used to show that the tilt of the tangent space yields a lower bound on the radius of the cylinder in which one can parametrize over the tangent space.

12.4 Corollary. *There exist $C \in (1, \infty)$ and $\epsilon_0 \in (0, 1)$ such that, for all $\epsilon \in (0, \epsilon_0]$, $R, L \in (0, \infty)$ and every $x_0 \in \mathbb{R}^{n+1}$ the following holds: Let $f \in C^2\left(B^n_{2R}(x), \mathbb{R}\right)$ with $(\hat{x}_0, f(\hat{x}_0)) = x_0$ and $|Df(\hat{x}_0)| \leq L$. Suppose $M = \text{graph}(f)$ satisfies*

$$\sup_M \|A\| \leq \epsilon R^{-1}.$$

Let $(b_i)_{1 \leq i \leq n}$ be an orthonormal basis of $T_{x_0}M$. Then there exists a local parametrisation $g \in C^2\left(B^n_R(0), \mathbb{R}\right)$ with $g(0) = 0$ and

$$x = \pi_{T_{x_0}M}(x) + g\left(x_b\right)\nu(x_0) + x_0 \tag{12.4}$$

$$\sum_{i=1}^n \hat{y}_i b_i + g(\hat{y})\nu(x_0) + x_0 \in M \tag{12.5}$$

for all $x \in M \cap C_r(x_0)$ and all $\hat{y} \in B^n_R(0)$, where $x_b := \sum_{i=1}^n \left((x - x_0) \cdot b_i\right) \mathbf{e}_i$ and $r = (2(1 + L^2))^{-\frac{1}{2}} R$. Furthermore g satisfies

$$g(0) = 0, \quad Dg(0) = 0, \tag{12.6}$$

$$\max\left\{R^{-1} \sup |g| + \sup |Dg| + R \sup |D^2 g|\right\} \leq C\epsilon. \tag{12.7}$$

Proof. Define $S \in SO(n+1)$ by

$$S(x) = \sum_{i=1}^n x_i b_i + x_{n+1}\nu(x_0).$$

Then $S(\mathbb{R}^n) = T_0 M$ and $S(e_{n+1}) = \nu(x_0)$.

Let ϵ_0 be from Proposition 12.3. As $\epsilon \leq \epsilon_0$ we can apply Proposition 12.3 to the manifold $S^{-1}(M - x_0)$ to obtain a function $g \in C^2\left(B^n_R(0), \mathbb{R}\right)$, which satisfies (12.6), (12.7) and $\text{graph}(g) \subset S^{-1}(M - x_0)$. In particular

$$S(\hat{y}, g(\hat{y})) + x_0 = \sum_{i=1}^n \hat{y}_i b_i + g(\hat{y})\nu(x_0) + x_0 \in M \tag{12.8}$$

for all $\hat{y} \in B_R^n(0)$. This already verifies (12.5).

Set $N := S(\text{graph}(g)) + x_0$, we know $N \subset M$ and want to show $M \cap C_r(0) \subset N$ for $r^2 = (2(1+L^2))^{-1}R^2$. Consider the set

$$Y := \{\hat{y} \in B_r^n(\hat{x}_0) : (\hat{y}, f(\hat{y})) \in N\}.$$

As $g(0) = 0$ we see $\hat{x}_0 \in N$, so $\hat{x}_0 \in Y$. Consider $\hat{y} \in Y$, as $y = (\hat{y}, f(\hat{y}))$ we have

$$\left\| \pi_{\mathbb{R}^n} - \pi_{T_yM} \right\|_{op} < 1.$$

Thus we can use Proposition A.16 and Remark A.17.2, to obtain a δ such that $B_\delta^n(\hat{y}) \cap B_r^n(\hat{x}_0) \subset Y$. This shows that Y is relatively open in $B_r^n(0)$, and as $B_r^n(0)$ is connected it suffices to show Y is relatively closed to obtain $Y = B_r^n(0)$ and thus the result.

Consider a sequence $(\hat{y}_m)_{m\in\mathbb{N}}$ in Y with $\lim_{n\to\infty} \hat{y}_m = \hat{y}_0$, $\hat{y}_0 \in B_r^n(\hat{x}_0)$. For $m \in \mathbb{N}$ set $y_m := (\hat{y}_m, f(\hat{y}_m))$, we want to exclude the case that (y_m) converges to a point in ∂N. For $m \in \mathbb{N}$ we know $\hat{y}_m \in Y$, so there exists $\hat{w}_m \in B_R^n(0)$ such that

$$y_m = S(\hat{w}_n, g(\hat{w}_m)) + x_0.$$

For $\hat{w}_m$ we can estimate

$$\begin{aligned} |\hat{w}_m|^2 &\le |(\hat{w}_m, g(\hat{w}_m))|^2 = |S(\hat{w}_m, g(\hat{w}_m))|^2 = |y_m - x_0|^2 \\ &= |\hat{y}_m - \hat{x}_0|^2 + |f(\hat{y}_m) - x_0 \cdot \mathbf{e}_{n+1}|^2 \le r^2 + |f(\hat{y}_m) - f(\hat{x}_0)|^2 \end{aligned}$$

for all $m \in \mathbb{N}$, where we used $f(\hat{x}_0) = x_0 \cdot \mathbf{e}_{n+1}$ and $\hat{y}_m \in N \subset B_r^n(\hat{x}_0)$. Then we can use the mean value formula and $|Df(\hat{x}_0)| \le L$, to obtain

$$|\hat{w}_m|^2 \le r^2 + |Df(\hat{x}_0)|^2 |\hat{y}_m - \hat{x}_0|^2 \le (1+L^2)r^2 = \frac{R^2}{2}$$

for all $m \in \mathbb{N}$, where we used $r^2 = (2(1+L^2))^{-1}R^2$. Thus a subsequence of the $\hat{w}_m$ converges to some $\hat{w}_0 \in B_{\frac{R}{2}}^n(\hat{x}_0)$. In particular $g(\hat{w}_0)$ is defined. By continuity of f and g we can conclude

$$(\hat{y}_0, f(\hat{y}_0)) = \lim_{m\to\infty} y_m = \lim_{m\to\infty} S(\hat{w}_n, g(\hat{w}_m)) + x_0 = S(\hat{w}_0, g(\hat{w}_0)) + x_0,$$

so $(\hat{y}_0, f(\hat{y}_0)) \in N$, which shows that Y is closed in $B_r^n(\hat{x}_0)$. As we already showed Y is open, non-empty and $B_r^n(0)$ is connected this establishes the result. $\square$

Also for manifolds in a narrow slab Proposition 12.3 can be used to obtain a tilt bound from small curvature.

12.5 Corollary. *There exist $C \in (1,\infty)$ and $\epsilon_1 \in (0,1)$ such that, for all $\epsilon \in (0,\epsilon_1]$, $R \in (0,\infty)$ and every $x_0 \in \mathbb{R}^{n+1}$ the following holds: Let M be an embedded C^2-hypersurface in $\mathbb{R}^{n+1}$ with $x_0 \in M$ and*

$$M \subset \mathbb{R}^n \times [x_0 \cdot \mathbf{e}_{n+1} - \epsilon R, x_0 \cdot \mathbf{e}_{n+1} + \epsilon R]. \tag{12.9}$$

Suppose M also satisfies $\partial M \cap C_{2R}(x_0) = \emptyset$ and

$$\sup_{M \cap C_{2R}(x_0)} \|A(x)\| \leq \epsilon R^{-1}. \tag{12.10}$$

Then $\left|\pi_{T_{x_0}M} - \pi_{\mathbb{R}^n}\right| \leq C\epsilon$.

Proof. Let $(b_i)_{1\leq i\leq n}$ be an orthonormal basis of $T_{x_0}M$. Define $S \in SO(n+1)$ by

$$S(x) = \sum_{i=1}^n x_i b_i + x_{n+1}\nu(x_0).$$

Then $S(\mathbb{R}^n) = T_{x_0}M$ and $S(e_{n+1}) = \nu(x_0)$.

Let ϵ_0 be from Proposition 12.3. Let $\epsilon_1 \leq \epsilon_0$, then we can apply Proposition 12.3 to the manifold $S^{-1}(M - x_0)$, to obtain a function $g \in C^2\left(B^n_R(0), \mathbb{R}\right)$ with $g(0) = 0$, $\sup|g| \leq C_n\epsilon R$ and $\operatorname{graph}(g) \subset S^{-1}(M - x_0)$. In particular

$$S(\hat{y}, g(\hat{y})) + x_0 = \sum_{i=1}^n \hat{y}_i b_i + g(\hat{y})\nu(x_0) + x_0 \in M \tag{12.11}$$

for all $\hat{y} \in B^n_R(0)$.

For $i \in \{1,\dots,n\}$ and $r = \frac{R}{2}$ define $w_i \in \mathbb{R}^n \times \{0\}^k$ by

$$w_i := \pi_{\mathbb{R}^n}\left(S\left(r\mathbf{e}_i, g(r\mathbf{e}_i)\right)\right) = \pi_{\mathbb{R}^n}\left(rb_i + g(r\mathbf{e}_i)\nu(x_0)\right).$$

Then we can estimate

$$\begin{aligned} |rb_i - w_i| &\leq |g(r\mathbf{e}_i)\nu(x_0)| + |rb_i + g(r\mathbf{e}_i)\nu(x_0) - w_i| \\ &= |g(r\mathbf{e}_i)| + |S\left(r\mathbf{e}_i, g(r\mathbf{e}_i)\right) \cdot e_{n+1}|. \end{aligned}$$

By (12.11) we have $S\left(r\mathbf{e}_i, g(r\mathbf{e}_i)\right) + x_0 \in M$, so by the slap condition (12.9) and with $\sup|g| \leq C_n\epsilon R$, we obtain

$$|rb_i - w_i| \leq C_n\epsilon r$$

for all $i \in \{1,\dots,n\}$, where we used $r = \frac{R}{2}$. Now set $\tilde{w}_i := r^{-1}w_i$ for $i = 1,\dots,n$. Then $|b_i - \tilde{w}_i| \leq C_n\epsilon$. In particular as $\epsilon \leq \epsilon_1$ for ϵ_1 small enough the $\tilde{w}_i$ form a basis of $\mathbb{R}^n$. Then Lemma A.10.2 with $T = \mathbb{R}^n$ and $B = T_{x_0}M$ establishes the result. $\square$

We want to show: If (M_t) is a smooth mean curvature flow (see Definition 3.13) which is graphical inside a cylinder for a long period of time, then it stays graphical in a much smaller cylinder for a little longer. To do so we need local gradient and curvature estimates established in [EH2], see also [E4].

12.6 Proposition ([EH2, 2.1])**.** *Let $R, T \in (0,\infty)$, $t_0 \in \mathbb{R}$, $x_0 \in \mathbb{R}^{n+1}$ and $(M_t)_{t\in[t_0,t_0+T]}$ be a smooth mean curvature flow in $\mathbb{R}^{n+1}$. Suppose for $t \in [t_0, t_0+(2n)^{-1}R^2]$ we have that v is well-defined inside $B_{\sqrt{R^2-2n(s-t_0)}}(x_0)$ for all $s \in [t_0, t]$. Then the estimate*

$$v(x)\left(1-\frac{|x-x_0|^2+2n(t-t_0)}{R^2}\right) \le \sup_{M_{t_0}\cap B_R(x_0)} v$$

holds for all $x \in M_t \cap B_{\sqrt{R^2-2n(t-t_0)}}(x_0)$.

12.7 Proposition ([EH2, 3.1])**.** *There exists $C \in (1,\infty)$ such that for all $R \in (0,\infty)$, the following holds: Let $(M_t)_{t\in[-4R^2,0]}$ be a smooth mean curvature flow in $\mathbb{R}^{n+1}$ such that*

$$M_t \cap C_{2R}(0) \cap B_{3R}(0) = \operatorname{graph}(f_t)$$

for some $f_t : B^n_{2R}(0) \to \mathbb{R}$ for all $t \in [-4R^2, 0]$. Then the estimate

$$\|A_t(\hat{x}, f_t(\hat{x}))\|^2 \le C\left((t+4R^2)^{-1}+R^{-2}\right) \sup_{(s,\hat{x})\in[-4R^2,t]\times B^n_{2R}(\hat{x}_0)} v_s(\hat{x}, f_s(\hat{x}))^4$$

holds for all $\hat{x} \in B^n_R(0)$ and all $t \in (-4R^2, 0]$.

The main ingredient to sustain the graphical representability of (M_t) is the regularity result by White from [W4]. The version presented here is taken from [E4, 5.6]

12.8 Theorem ([W4])**.** *There exist constants $c \in (0,1)$ and $C \in (1,\infty)$ such that for all $R \in (0,\infty)$, $t_0 \in \mathbb{R}$ and $x_0 \in \mathbb{R}^{n+k}$ the following holds: Let $(M_t)_{t\in[t_0-8R^2,t_0]}$ be a smooth mean curvature flow in $\mathbb{R}^{n+1}$. Suppose $x_0 \in M_{t_0}$ and for some $\rho \in (2R,\infty)$*

$$\int_{M_t} \Phi_{(s,x)}\varphi_{(s,x),\rho} \le 1+c \tag{12.12}$$

for all $(s,x) \in [t_0-4R^2, t_0] \times B_{2R}(x_0)$ and $t \in [s-4R^2, s)$. Then

$$\|A_t(x)\| \le CR^{-1} \tag{12.13}$$

for all $t \in [t_0-R^2, t_0]$ and $x \in M_t \cap B_R(x_0)$. Here Φ and φ are from Definition 6.1.

Using these results we can prove, that graphical representation for a long time on a large region implies curvature bounds on a smaller region for a shorter time period. This new time period where the curvature is bounded exceeds the time where we assumed a graphical representation.

12.9 Lemma. *There exists a $C \in (1,\infty)$ such that, for all $L \in [1,\infty)$, $r \in (0,\infty)$, $t_1 \in \mathbb{R}$ the following holds: Let $(M_t)_{t\in[t_1-C^2L^6r^2,t_1+r^2]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$. Suppose there exists an $f \in C^\infty([t_1 - C^2L^6r^2, t_1] \times B^n_{CL^3r}(0), \mathbb{R})$ with $\sup|Df| \le L$ and*

$$\operatorname{graph}(f(t,\cdot)) = M_t \cap C_{CL^3r}(0) \quad \forall t \in \left[t_1 - C^2L^6r^2, t_1\right]. \tag{12.14}$$

Then

$$\sup_{x\in M_t\cap B_r(0)} \|A_t(x)\| \le r^{-1} \quad \forall t \in \left[t_1 - r^2, t_1 + r^2\right]. \tag{12.15}$$

Proof. We may assume $t_1 = 0$. As $|Du| \le L$ we have $v \le 2L$, thus Proposition 12.7 with $R = 2^{-1}CL^3r$ yields

$$\|A_t(\hat{x}, f_t(\hat{x}))\|^2 \le C_n \left((t + C^2L^6r^2)^{-1} + C^{-2}L^{-6}r^{-2}\right) L^4$$

for all $\hat{x} \in B^n_{2^{-1}CL^3r}(0)$ and all $t \in (-C^2L^6r^2, 0]$. In particular this implies

$$\|A_t(\hat{x}, f_t(\hat{x}))\| \le C_n C^{-1}L^{-1}r^{-1}$$

for all $\hat{x} \in B^n_{2^{-1}CL^3r}(0)$ and all $t \in (-2^{-1}C^2L^6r^2, 0]$. Consider a $\gamma \in (0,1)$ which we will choose small depending only on n. Set

$$R_0 = \gamma^{-1}r \quad \text{and} \quad \rho = \gamma^{-2}r.$$

We can choose C large enough depending on γ such that

$$\|A_t(\hat{x}, f_t(\hat{x}))\| \le \gamma(L\rho)^{-1} \tag{12.16}$$

for all $\hat{x} \in B^n_{16L\rho}(0)$ and all $t \in (-8R_0^2, 0]$, where we estimated $L^3 \ge L \ge 1$.

We want to use Theorem 12.8 with $t_0 = r^2$, $x_0 = 0$, $R = R_0 = \gamma^{-1}r$ and $\rho = \gamma^{-2}r$, so consider arbitrary $(s,x) \in (r^2 - 4R_0^2, r^2] \times B_{R_0}(0)$ and $t \in (s - 4R_0^2, s)$. We want to verify (12.12). Set $\tau = t - r^2$ then $\tau \in (-8R_0^2, 0]$ and by the monotonicity formula (6.2) we can estimate

$$\int_{M_t} \Phi_{(s,x)}\varphi_{(s,x),\rho} \le \int_{M_\tau} \Phi_{(s,x)}\varphi_{(s,x),\rho}. \tag{12.17}$$

By choice of s, τ, $r = \gamma R_0 = \gamma^2\rho$ and by definition of φ (see Definition 6.1) we can estimate

$$2n(s-\tau) \le 2n(4R_0^2 + r^2) \le 10n\gamma\rho \le \rho,$$

thus $\mathrm{spt}\varphi_{(s,x),\rho}(t,\cdot) \subset B_{2\rho}(x) \subset B_{4\rho}(0)$ and

$$\sup |\varphi_{(s,x),\rho}(\tau,\cdot)| \le (1+10n\gamma)^3 \le 1 + C_n\gamma,$$

where we chose γ small depending on n. Then (12.17) yields

$$\int_{M_t} \Phi_{(s,x)}\varphi_{(s,x),\rho} \le (1+C_n\gamma)\int_{B_{4\rho}(0)} \Phi_{(s,x)} d\mu_\tau. \tag{12.18}$$

As $\tau \in (-8R^2, 0]$ and as we have small curvature at this time, one would expect the heat kernel integral on the right hand side of (12.18) to be lower than $1+c$ as in (12.12), this will be shown next.

In view of (12.16) for γ small enough we can use Corollary 12.4 for the manifold $M \llcorner C_{16L\rho}(0)$ and with $R = 8L\rho$, $\epsilon = 8\gamma$, $x_0 = 0$ to obtain a parametrisation $g_\tau \in C^2\left(B^n_{8L\rho}(0), \mathbb{R}\right)$ with

$$g_\tau(0) = 0, \quad Dg_\tau(0) = 0, \quad \sup |D^2 g_\tau| \le C_n\gamma L^{-1}\rho^{-1} \tag{12.19}$$

and

$$x = \sum_{i=1}^n (x\cdot b_i)b_i + g_\tau(x_b)\nu_\tau(0) \tag{12.20}$$

for all $x \in M_\tau \cap C_{4\rho}(0)$, where $(b_i)_{i=1,\dots,n}$ is an orthonormal basis of T_0M_τ, $x_b := \sum_{i=1}^n (x\cdot b_i)\mathbf{e}_i$ and $\nu_\tau(0)$ is the normal pointing upwards. Here we estimated $r = (2(1+L^2)^{-\frac{1}{2}}R \ge (4L^2)^{-\frac{1}{2}}8L\rho = 4\rho$, where we used $L \ge 1$. In view of (12.19) the mean value formula yields

$$\sup_{B^n_{4\rho}(0)} \sqrt{1+|Dg_\tau|^2} \le \sqrt{1+C_nL^{-2}\gamma^2}. \tag{12.21}$$

Using (12.20), we can calculate

$$\int_{B_{4\rho}(0)} \Phi_{(s,x)} d\mu_\tau = \int_{B^n_{4\rho}(0)} \Phi_{(s,(x_a, g_\tau(\hat{x})))}\left(\tau, (\hat{y}, g_\tau(\hat{y}))\right) Jg_\tau(\hat{y}) d\mathscr{L}^n(\hat{y}).$$

In view of the definition of the Jacobian in co-dimension one we obtain with (12.21) and by definition of Φ (see Definition 6.1)

$$\begin{aligned}\int_{B_{4\rho}(0)} \Phi_{(s,x)} d\mu_\tau &\le \sup_{B^n_{4R_0}(0)} \sqrt{1+|Dg_\tau|^2}\int_{\mathbb{R}^n} \Phi_{(s,(x_a,0))}(\tau,(\hat{y},0)) d\mathscr{L}^n(\hat{y}) \\ &\le \sqrt{1+C_nL^{-2}\gamma^2} \le 1 + C_n\gamma^2,\end{aligned}$$

where we used $L \geq 1$ and $C_n\gamma^2 \leq 1$ for γ small enough. Then (12.18) yields

$$\int_{M_t} \Phi_{(s,x)}\varphi_{(s,x),\rho} \leq (1+C_n\gamma)(1+C_n\gamma^2)$$

and for γ small enough this implies (12.12). Then Theorem 12.8 yields

$$\|A_t(x)\| \leq C_n R_0^{-1}$$

for all $t \in [r^2 - R_0^2, r^2]$ and $x \in M_t \cap B_{R_0}(0)$. Now as $r = \gamma R_0$ we can choose γ small to establish the result. $\square$

We want to use this to show that we actually can expand the graphical representation to later times. We will need the following smoothness estimate from [E4, 3.22]:

12.10 Proposition. *Fore every $C_0 \in (1,\infty)$ there exists a $C_1 \in (1,\infty)$ such that for all $R \in (0,\infty)$, $t_0 \in \mathbb{R}$ the following holds: Let $(M_t)_{t\in[t_0-4R^2,t_0]}$ be a smooth mean curvature flow in $\mathbb{R}^{n+1}$. If*

$$\|A_t(x)\|^2 \leq C_0 R^{-2}$$

for all $x \in M_t \cap B_{2R}(0)$ and $t \in [t_0 - 4R^2, t_0]$, then

$$\left\|\nabla^{M_t} A(t,x)\right\|^2 \leq C_1 R^{-4}$$

holds for all $x \in M_t \cap B_R(0)$ and $t \in [t_0 - R^2, t_0]$.

12.11 Theorem (Staying Graphical)**.** *There exists a $C \in (1,\infty)$ such that for all $L \in [1,\infty)$, $\rho, \Gamma \in (0,\infty)$, $s_0 \in \mathbb{R}$ and $a = (\hat{a}, a_{n+1}) \in \mathbb{R}^{n+1}$ the following holds: Let $(M_t)_{t\in[s_0-C^2L^8\rho^2, s_0+\rho^2]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$. Suppose there exists an $f \in C^\infty\left([s_0 - C^2L^8\rho^2, s_0] \times B^n_{CL^4\rho}(\hat{a})\right)$ with $\sup|Df| \leq L$, $\sup|f - a_{n+1}| \leq \Gamma$ and*

$$\operatorname{graph}(f(t,\cdot)) = M_t \cap C_{CL^4\rho}(a) \quad \forall t \in \left[s_0 - C^2L^8\rho^2, s_0\right]. \tag{12.22}$$

Then there exists a $g \in C^\infty\left([s_0 - \rho^2, s_0 + \rho^2] \times B^n_\rho(\hat{a})\right)$ with $\sup|g - a_{n+1}| \leq \Gamma + \rho$ and

$$\operatorname{graph}(g(t,\cdot)) = M_t \cap C_\rho(a) \quad \forall t \in \left[s_0 - \rho^2, s_0 + \rho^2\right]. \tag{12.23}$$

Also g satisfies

$$\sqrt{1 + \sup|Dg|^2} \leq 2\sqrt{1 + \sup|Df|^2}. \tag{12.24}$$

If in addition $\Gamma \leq L^{-1}\rho$, then $\sup|Dg| \leq CL^{-1}$.

Proof. We may assume $a = 0$ and $s_0 = 0$. Let $M_t = \phi_t(N)$ for an n-dimensional manifold N and a smooth family of embeddings $\phi_t = \phi(t, \cdot) : N \to \mathbb{R}^{n+k}$. For given $\rho \in (0, \infty)$ set

$$R_0 = (2\sqrt{n} + 4)\rho.$$

Let $x \in C_{R_0}(0)$. For C large enough we see that

$$B_{12nL^4\rho}(x) \subset C_{CL^4\rho}(0).$$

As $M_t \cap C_{CL^4\rho}(0) \subset \mathbb{R}^n \times [-\Gamma, \Gamma]$ we can use Lemma 3.12.4 with $t_0 = -\rho^2$, $x_0 = (\hat{x}, 0) \pm \Gamma \mathbf{e}_{n+1}$, $v = \pm \mathbf{e}_{n+1}$, $R = 12nL^4\rho$ and $\delta = (12n)^{-1}L^{-8}$ to obtain

$$M_t \cap B_{6nL^4\rho}(x) \subset \mathbb{R}^n \times [-\Gamma - L^{-4}\rho, \Gamma + L^{-4}\rho]$$

for all $t \in [-\rho^2, \rho^2]$ and all $x \in C_{R_0}(0)$. Thus

$$M_t \cap C_{R_0}(0) \subset \mathbb{R}^n \times [-\Gamma - L^{-4}\rho, \Gamma + L^{-4}\rho] \tag{12.25}$$

for all $t \in [-\rho^2, \rho^2]$.

Let C_1 be from Proposition 12.10 for $C_0 = 2^{-1}$. By Lemma 12.9 with $r = 8C_1^2 LR_0$ and $t_1 = \rho^2 - r^2$ we obtain

$$\|A_t(x)\| \leq \left(8C_1^2 LR_0\right)^{-1}$$

for all $t \in [\rho^2 - 2(8C_1^2 LR_0)^2, \rho^2]$ and $x \in M_t \cap C_{8C_1^2 LR_0}(0)$. Here we had to choose C big depending on C_1 and the constant from Lemma 12.9. Then Proposition 12.10 with $R = 4C_1^2 LR_0$ and $t_0 = \rho^2$ yields

$$\left\|\nabla^{M_t} A_t(x)\right\| \leq (LR_0)^{-2}$$

for all $t \in [\rho^2 - 4C_1^4 LR_0^2, \rho^2]$ and $x \in M_t \cap C_{4C_1^2 LR_0}(0)$. This implies bounds on the mean curvature as well, so we have

$$\|A(t,x)\| \leq (LR_0)^{-1} \tag{12.26}$$

$$\left\|\nabla^{M_t} A(t,x)\right\| \leq (LR_0)^{-2} \tag{12.27}$$

$$|H(x)| \leq (LR_0)^{-1} \tag{12.28}$$

$$|\nabla^{M_t} H(x)| \leq (LR_0)^{-2} \tag{12.29}$$

for all $t \in [-\rho^2, \rho^2]$ and $x \in M_t \cap C_{4R_0}(0)$. Here we estimated $C_1 \geq 1$ and $R_0^2 \geq 2\rho^2$.

Temporarily fix $x \in M_s \cap C_{R_0}(0)$, $s \in [-\rho^2, \rho^2]$ then there exists $p \in N$ with $\phi(s, p) = x$. We want to show

$$\phi(t, p) \in C_{3R_0}(0) \tag{12.30}$$

for all $t \in [-\rho^2, s]$. Set

$$I := \left\{ \tau \in [-\rho^2, s] : \phi(t,p) \in \overline{C_{3R_0}(0)} \;\; \forall\, t \in [\tau, s] \right\}.$$

By definition of p and by continuity of ϕ, we have that I is non-empty and closed. Consider $\tau \in I$ then by integrating (3.18) and using (12.28), we can estimate

$$|\phi(\tau,p) - \phi(s,p)| \leq \int_\tau^s |H(\phi(t,p))| dt \leq (s-\tau)\,(LR_0)^{-1} \leq R_0,$$

where we used $s - \tau \leq 2\rho^2 \leq R_0^2$ and $L \geq 1$. As $\phi(s,p) \in C_{R_0}(0)$ we see $\phi(\tau,p) \in C_{2R_0}(0)$, so continuity of ϕ implies that I is relative open. Thus, as $[-\rho^2, s]$ is connected we have $I = [-\rho^2, s]$. This establishes (12.30). The bound on Df yields $v(\phi(-\rho^2,p)) \leq 2L$, so

$$\nu_{-\rho^2}(\phi(-\rho^2,p))) \cdot e_{n+1} = (v(\phi(-\rho^2,p)))^{-1} \geq (2L)^{-1}. \tag{12.31}$$

From Huisken [H1] we know $\frac{d}{dt}\nu(\phi(t,p)) = \nabla^{M_t} H(\phi(t,p))$, so in view of (12.30), we can use (12.29), to estimate

$$\begin{aligned}\left|\nu_s(x) - \nu_{-\rho^2}(\phi(-\rho^2,p))\right| &= \int_{-\rho^2}^s \left|\nabla^{M_t} H(\phi(t,p))\right| dt \\ &\leq (s+\rho^2)\,(LR_0)^{-2} \leq (4L)^{-1},\end{aligned}$$

where we used $s + \rho^2 \leq 2\rho^2 \leq 4^{-1}R_0^2$ and $L \geq 1$. So in view of (12.31) we showed $|\nu_s(x) \cdot e_{n+1}| \geq (4L^{-1})$ and by Proposition A.9 this yields

$$\|\pi_{T_x M_s} - \pi_{\mathbb{R}^n}\|_{op} < 1 \tag{12.32}$$

for all $x \in M_s \cap C_{R_0}(0)$, $s \in [-\rho^2, \rho^2]$.

Temporary fix $s \in [-\rho^2, \rho^2]$. As the normal always has a non-zero e_{n+1}-component everywhere in $C_{R_0}(0)$, we obtain that the number of sheets has to be constant there, so

$$N(\hat{y}) := \sharp\left\{x \in M_s \cap \{\hat{y}\} \times [-\Gamma-\rho, \Gamma+\rho]\right\} = m_0 \tag{12.33}$$

for every $\hat{y} \in B^n_{R_0}(0)$ for a fixed $m_0 \in \mathbb{N} \cup \{0\}$. This we want to prove now. First note that the height bound $\Gamma + R$ in (12.33) is no restriction, as by (12.25) and $L \geq 1$, we have $M_s \cap C_{R_0}(0) \subset B^n_{R_0}(0) \times [-\Gamma-\rho, \Gamma+\rho]$.

We observe that $N(\hat{y})$ has to be finite. To see this note that $\{\hat{y}\} \times [-\Gamma - \rho, \Gamma+\rho]$ is bounded. Also in view of (12.32), we can use Proposition A.16 to see that the x in $N(\hat{y})$ have to be discrete, so $N(\hat{y})$ has to be finite for every $\hat{y} \in B^n_{R_0}(0)$.

For $m \in \mathbb{N} \cup \{0\}$ consider the set

$$Y(m) := \left\{ \hat{y} \in B^n_{R_0}(0) : N(\hat{y}) = m \right\}.$$

We want to show that $Y(m)$ is open and closed for every m thus $N(\hat{y})$ would have to be constant.

To show that $Y(m)$ is open let $\hat{y} \in Y(m)$. Then there exist different $x_1, \ldots, x_m \in \{\hat{y}\} \times [-\Gamma - \rho, \Gamma + \rho]$ with $x_i \in M_s$. In view of (12.32), we can use Proposition A.16 to obtain an $r_i > 0$ and a function $g_i \in C^\infty\left(B^n_{2r_i}(\hat{x}_i)\right)$ with $g(\hat{x}_i) = x_i$, such that

$$x = (\hat{x}, g_i(\hat{x})) \tag{12.34}$$

$$(\hat{v}, g_i(\hat{v})) \in M \tag{12.35}$$

for all $x \in B_{2r_i}(x_i)$ and for all $\hat{v} \in B^n_{2r_i}(x_i)$. To see this use Proposition A.16 with $S = Id$, $x_0 = x_i$ and $g_i(\hat{v}) = g(\hat{v} - \hat{x}_i) + x_i$.

Now consider the set

$$W := \{\hat{y}\} \times \left([-\Gamma - \rho, \Gamma + \rho] \setminus \bigcup_{i=1}^{n} (x_i \cdot \mathbf{e}_{n+1} - 2r_i, x_i \cdot \mathbf{e}_{n+1} + 2r_i) \right).$$

This describes all the points on the $\{\hat{y}\}$-axis, which are certainly away from the x_i. By choice of the x_i we see that $W \cap M_s = \emptyset$, so as M_s is closed for every $x \in W$ there exists a radius $\tilde{r}_x$ such that $B_{\tilde{r}_x}(x) \cap M_s = \emptyset$. As W is compact we find an $\tilde{r} \in (0,1)$ such that $B_{\tilde{r}}(x) \cap M_s = \emptyset$ for all $x \in W$. Then by definition of W we have

$$M_s \cap C_{\tilde{r}}((\hat{y}, 0)) \setminus \bigcup_{i=1}^{n} B_{2r_i}(x_i) = \emptyset. \tag{12.36}$$

Set $r = \min\{\tilde{r}, r_1, \ldots, r_m\}$, by (12.34), (12.35) and (12.36) we see

$$N(\hat{x}) = m$$

for all $\hat{x} \in B^n_r(\hat{y})$, this implies $Y(m)$ is open for every $m \in \mathbb{N} \cup \{0\}$.

Now let $\hat{y} \notin Y(m)$, then $\hat{y} \in Y(\tilde{m})$ for some $\tilde{m} \neq m$. But as we just showed there exists an $r \in (0,1)$ such that $N(\hat{x}) = \tilde{m}$ for all $\hat{x} \in B^n_r(\hat{y})$ thus $Y(m)$ has to be closed and as $B^n_{R_0}(0)$ is connected this implies (12.33).

Equality (12.33) holds for every $s \in [-\rho^2, \rho^2]$. In view of assumption (12.22) we see $m_0 = 1$ for all times $s \in [-\rho^2, 0]$. However, the number of sheets cannot jump, so we have $m_0 = 1$ for all the times. Then $M_t \cap C_{R_0}(0)$ is a graph for all $t \in [-\rho^2, \rho^2]$, such that we can find g as stated. Note that g is actually defined on $B^n_{R_0}(0)$ and satisfies the height bound from (12.33).

To obtain the gradient bound (12.24) consider $x \in M_s \cap C_\rho(0)$, $s \in [0, \rho^2]$.sup By definition of R_0 we have $B_{\sqrt{4n}\rho}(x) \subset C_{R_0}(0)$, in particular v is well defined in $M_t \cap B_{\sqrt{4n}\rho}(x)$ for all $t \in [0, s]$. By Proposition 12.6 with $t_0 = 0$, $x = x_0$ and $R = \sqrt{4n}\rho$ we can estimate

$$\frac{1}{2}v(x) \leq v(x)\left(1 - \frac{2ns}{4n\rho^2}\right) \leq \sup_{M_{s_0} \cap C_{4n\rho^2}(0)} v \leq \sqrt{1 + \sup |Df|^2}.$$

Then $\sqrt{1 + \sup |Dg|^2} \leq 2\sqrt{1 + \sup |Df|^2}$, where the sup of g is over the set $B^n_\rho(0)$. This completes the first result.

For the second part we assume

$$\Gamma \leq L^{-1}\rho.$$

Note that

$$\sup |Dg| \leq 2\sqrt{1 + \sup |Df|^2} \leq 4L. \tag{12.37}$$

Let ϵ_1 be from Corollary 12.5. If $L \leq 4\epsilon_1^{-1}$, we can estimate $\sup |g| \leq 4L \leq 16\epsilon_1^{-2}L^{-1}$, thus the second result follows for $C \geq C_n\epsilon_1^{-2}$.

Now suppose $L \geq 4\epsilon_0^{-1}$. Consider $x \in M_t \cap C_{3\rho}(0)$, then (12.25) and (12.26) imply (12.9) and (12.10) with $\epsilon = L^{-1}$ and $R = \rho$. Here we use $R_0 \geq \rho$ and $L \geq 1$. Then Corollary 12.5 yields

$$|\pi_{T_xM_t} - \pi_{\mathbb{R}^n}| \leq C_nL^{-1} \tag{12.38}$$

for all $x \in M_t \cap C_{3\rho}(0)$ and all $t \in [-\rho^2, \rho^2]$. By Proposition A.12.1 we can estimate

$$\sup |Dg(t,x)| \leq C_n\, |\pi_{T_xM_t} - \pi_{\mathbb{R}^n}|\, (1 + \sup |Dg(t,x)|)^2.$$

Combining this with (12.38) we arrive at

$$\sup |Dg(t,x)| \leq C_nL^{-1}(1 + \sup |Dg(t,x)|). \tag{12.39}$$

Now by (12.37) and (12.39) we obtain $\sup |Dg(t,x)| \leq C_n$. Then using (12.39) again yields $\sup |Dg(t,x)| \leq C_nL^{-1}$. $\square$

Now recursively using Theorem 12.11 we see that the time interval for which we obtain graphical representation can be arbitrarily large, if we start with graphical representation in a large enough cylinder.

12.12 Proposition. *There exists a $C \in (1,\infty)$ such that for all $L_0 \in [1,\infty)$, $\rho_0, \Gamma_0 \in (0,\infty)$, $\tau \in (0,1]$, $t_0 \in \mathbb{R}$, $b = (\hat{b}, b_{n+1}) \in \mathbb{R}^{n+1}$ the following holds: Let $(M_t)_{t\in[t_0-\tau^2\rho_0^2, t_0+\rho_0^2]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$. Suppose there exists an $f \in C^\infty([t_0-\tau^2\rho_0^2, t_0] \times B^n_{CL_0^8\tau^{-1}\rho_0}(\hat{b})$ with $\sup|Df| \le L_0$, $\sup|f - b_{n+1}| \le \Gamma_0$ and*

$$\operatorname{graph}(f(t,\cdot)) = M_t \cap C_{CL_0^8\tau^{-1}\rho_0}(b) \quad \forall t \in \left[t_0 - \tau^2\rho_0^2, t_0\right]. \tag{12.40}$$

Then exists a $g \in C^\infty\left([t_0, t_0+\rho_0^2] \times B^n_{\rho_0}(\hat{b})\right)$ with $\sup|g - b_{n+1}| \le \Gamma_0 + \rho_0$ and

$$\operatorname{graph}(g(t,\cdot)) = M_t \cap C_{\rho_0}(b) \quad \forall t \in \left[t_0, t_0+\rho_0^2\right]. \tag{12.41}$$

Also g satisfies

$$\sqrt{1+\sup|Dg|^2} \le 2\sqrt{1+\sup|Df|^2}. \tag{12.42}$$

If in addition $\Gamma_0 \le C^{-1}\tau L_0^{-5}\rho_0$, then $\sup|Dg| \le L_0^{-1}$.

Proof. We may assume $t_0 = 0$ and $b = 0$. Let $C_{12.11}$ be the constant from Theorem 12.11. Set

$$L := 4C_{12.11}L_0.$$

Choose $N \in \mathbb{N}$ such that

$$C_{12.11}^2 L^8\tau^{-2} \le N < 2C_{12.11}^2 L^8 \tau^{-2}, \tag{12.43}$$

then set

$$\rho^2 := N^{-1}\rho_0^2.$$

The idea is to iterate the previous theorem with radius ρ. With each iteration step we can continue g by a time interval of length ρ^2, but inside a cylinder with radius decreased by $C_{12.11}L^4\rho$. For $i \in \{0,\dots,N\}$ set

$$r_i := 2^{-2}CL_0^8\tau^{-1}\rho_0 - i(C_{12.11}L^4\rho + \sqrt{4n}\rho_0). \tag{12.44}$$

This will be the radius of the cylinder in which we have graphical representation after the i-th step. By Definition of ρ and (12.43) we have $N\rho = \sqrt{N}\rho_0 \le 2C_{12.11}L^4\tau^{-1}\rho_0$. Then we can estimate

$$r_i \ge r_N \ge 2^{-2}CL_0^8\tau^{-1}\rho_0 - C_{12.11}^2 L^8\tau^{-1}\rho_0 - \sqrt{4n}\rho_0 \ge \rho_0 \tag{12.45}$$

for all $i \in \{0, \ldots, N\}$, where we used $L = C_n L_0$, $L_0 \geq 1$ and we chose C large enough.

We claim that for every $i \in \{0, \ldots, N\}$ the following holds: There exists a function $g_i \in C^\infty\left([(i-2)\rho^2, i\rho^2] \times B^n_{r_i}(0)\right)$ with $\sup |g_i| \leq \Gamma_0 + i\rho$,

$$\operatorname{graph}(g_i(t, \cdot)) = M_t \cap C_{r_i}(0) \quad \forall t \in \left[-\tau^2 \rho_0^2, i\rho^2\right] \tag{12.46}$$

and

$$\sqrt{1 + \sup |Dg_i|^2} \leq 2\sqrt{1 + \sup |Df|^2}. \tag{12.47}$$

We will prove this by induction. For $i = 0$ we can use $g_0 = f \llcorner C_{r_0}$, which satisfies (12.46) and (12.47).

Now suppose our claim holds for some $i \in \{0, \ldots, N-1\}$. In particular by (12.47) and $\sup |Df| \leq L_0$ we have $\sup |Dg_i| \leq 4L_0 \leq L$. Fix an arbitrary $y \in B^n_{r_{i+1} + \sqrt{4n}\rho_0}(0) \times \{0\}$. By (12.44) we have

$$C_{C_{12.11} L^3 \rho}(y) \subset C_{r_i}(0).$$

Also by $\rho^2 = N\rho_0^2$ and (12.43) we see that

$$i\rho^2 - C_{12.11}^2 L^6 \rho^2 \geq -\tau^2 \rho_0^2.$$

Theorem 12.11 with $f = g_i$, $\Gamma = \Gamma_0 + i\rho$, $s_0 = i\rho^2$ and $a = y$ then yields a function $g_y \in \left([(i-1)\rho^2, (i+1)\rho^2] \times B^n_\rho(\hat{y})\right)$ with

$$\operatorname{graph}(g_y(t, \cdot)) = M_t \cap C_\rho(y) \quad \forall t \in \left[(i-1)\rho^2, (i+1)\rho^2\right]$$

and as $y \in B^n_{r_{i+1} + \sqrt{4n}\rho_0}(0) \times \{0\}$ was arbitrary and the g_y overlap we can assemble them to one $\tilde{g} \in \left([(i-1)\rho^2, (i+1)\rho^2] \times C^n_{r_{i+1} + \sqrt{4n}\rho_0}(0)\right)$ with

$$\operatorname{graph}(\tilde{g}(t, \cdot)) = M_t \cap C_{r_{i+1} + \sqrt{4n}\rho_0}(0) \quad \forall t \in \left[(i-1)\rho^2, (i+1)\rho^2\right].$$

As by induction hypothesis g_i is defined for all times $[-\tau^2 r^2, (i+1)\rho^2]$ we obtain a $g_{i+1} \in \left([-\tau^2 R^2, (i+1)\rho^2] \times B^n_{R_{i+1} + \sqrt{4n}R}(0)\right)$ that satisfies (12.46) actually on the larger cylinder $C_{R_{i+1} + \sqrt{4n}\rho_0}(0)$.

To verify (12.47) consider $x \in M_s \cap C_{r_{i+1}}(0)$, $s \in [0, (i+1)\rho^2]$. By definition of r_{i+1} we have $B_{\sqrt{4n}\rho_0}(x) \subset C_{r_{i+1} + \sqrt{4n}\rho_0}(0)$, in particular v is well defined in $M_t \cap B_{\sqrt{4n}\rho}(x)$ for all $t \in [0, (i+1)\rho^2]$. By Proposition 12.6 with $x = x_0$, $t_0 = 0$ and $R = \sqrt{4n}\rho_0$ we can estimate

$$\frac{1}{2} v(x) \leq v(x) \left(1 - \frac{2ns}{4n\rho_0^2}\right) \leq \sup_{M_{s_0} \cap C_{4n\rho_0^2}(x)} v \leq \sqrt{1 + \sup |Df|^2}$$

where we used $s \leq N\rho^2 = \rho_0^2$. Then $\sqrt{1+\sup|Dg_{i+1}|^2} \leq 2\sqrt{1+\sup|Df|^2}$, where the sup of g_{i+1} is over the set $B^n_{R_{i+1}}(0)$. This completes the induction argument and thus yields the first result.

Now suppose $\Gamma_0 \leq C^{-1}\tau L_0^{-5}\rho_0$. By assumption (12.40) we have

$$M_{t_0-\tau^2\rho_0^2} \cap C_{CL_0^8\tau^{-1}\rho_0}(0) \subset \mathbb{R}^n \times [-\Gamma_0, \Gamma_0].$$

Then we can use Lemma 3.12.4 with $t_0 = -\tau^2\rho_0^2$, $x_0 = \pm\Gamma_0\mathbf{e}_{n+1}$, $v = \pm\mathbf{e}_{n+1}$, $R = CL_0^8\tau^{-1}\rho_0$ and $\delta = 12nC^{-2}L_0^{-16}\tau^2$ to obtain

$$\begin{aligned} &M_t \cap B_{2^{-1}CL_0^8\tau^{-1}\rho_0}(0) \\ &\subset \mathbb{R}^n \times [-\Gamma_0 - 12nC^{-1}L_0^{-8}\tau\rho_0, \Gamma_0 + 12nC^{-1}L_0^{-8}\tau\rho_0] \end{aligned} \tag{12.48}$$

for all $t \in [-\tau^2\rho_0^2, -\tau^2\rho_0^2 + 2\rho_0^2] \cap [-\tau^2\rho_0^2, \rho_0^2]$. Here we used that as $L_0 \geq 1$, $\tau \leq 1$ and that for C large enough $\delta = 12nC^{-2}L_0^{-16}\tau^2 \leq 6^{-1}$. Then with $\Gamma_0 \leq C^{-1}\tau L_0^{-5}\rho_0$ we can estimate

$$\Gamma_0 + 12nC^{-1}L_0^{-8}\tau\rho_0 \leq C_nC^{-1}L_0^{-5}\tau\rho_0 \leq C_nC^{-1}L^{-1}\rho \leq L^{-1}\rho,$$

where we used $N\rho^2 = \rho_0^2$, $L = 4L_0$, $C, L_0 \geq 1$, and estimate (12.43). Also we chose C large enough. Then (12.48) yields

$$M_t \cap C_{2^{-2}CL_0^8\tau^{-1}\rho_0}(x) \subset \mathbb{R}^n \times [-L^{-1}\rho, L^{-1}\rho]$$

for all $t \in [-\tau^2\rho_0^2, \rho_0^2]$, where we also used $\tau \leq 1$. Note that for r_i from (12.44), we have $r_i \leq 2^{-2}CL_0^8\tau^{-1}\rho_0$ for all $i \in \{0, \dots, N\}$. So all the g_i in the above induction argument satisfy the bound $\sup|g_i| \leq L^{-1}\rho$. Then in each induction step we can additionally use the second statement of Theorem 12.11 to obtain $\sup|Dg_i| \leq C_{12.11}L^{-1} \leq L_0^{-1}$ for all $i \in \{1, \dots, N\}$. This establish the result. $\square$

Now we can use these statements to prove new versions of Theorem 11.7 and Proposition 11.17. We start with the local result (recall Definition 11.1).

12.13 Proposition. *For every $\epsilon \in (0,1)$ there exist $\lambda, \kappa \in (0,1)$ such that for all $R \in (0,\infty)$, $L_0 \in [1,\infty)$, $\beta \in (0,1]$, $\gamma_0 \in (0, \beta\lambda L_0^{-6}]$, $\xi_0 \in (0,\lambda]$, $r \in (0, \lambda L_0^{-1}R]$ $t_0 \in \mathbb{R}$, $x_0, a_0 \in \mathbb{R}^{n+1}$ the following holds:*

Let $(M_t)_{t\in[t_0-\beta^2R^2, t_0+3\kappa R^2]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$ with $\bigcup_{t\in[t_0-\beta^2R^2,t_0+3\kappa R^2]} \partial M_t \cap C_{2R}(x_0) = \emptyset$. Suppose for all $t \in [t_0 - \beta^2R^2, t_0]$ we have: M_t is locally $(R, S_t, L_0, \gamma_0, \xi_0)$-plane-like around x_0 for some $S_t \subset B^n_R(0) \cap B^n_r(\hat{a}_0)$. Then one of the following two statements holds:

1. *For* $I := [t_0 + \kappa R^2, t_0 + 2\kappa R^2]$ *there exists an* $g \in C^\infty(I \times B^n_{\lambda R}(\hat{x}_0))$, *with* $\mathrm{lip}(g) \le \epsilon$, $\sup|g| \le \epsilon R$ *and*

$$\mathrm{spt}\mu_t \cap C_{\lambda R}(x_0) = \mathrm{graph}(g(t,\cdot))$$

for all $t \in I$.

2. $\mu_{t_0+3\kappa\rho^2}(C_{\lambda R}(x_0)) = 0$.

12.14 Remark. We believe that an argumentation like in Corollary 11.13 can be used to exclude the appearance of empty cylinders as in alternative (2). In particular a result like Theorem 11.10 should hold for smooth mean curvature flows as well.

Proof. We may assume $t_0 = 0$ and $x_0 = 0$.

Let l be the constant from Theorem 11.7 we want to show that M_0 actually is locally (R, S, l, γ_0, ξ)-plane-like around 0 for some small $S \subset B^n_R(0)$ and small ξ. So we have to show that the Lipschitz constant is actually smaller at least on the major part of $B^n_R(0)$. To see this we will use Theorem 12.11. There the statement for the smaller gradient is obtained by a small height bound, which we have on the graphical part.

As for all $t \in [-R^2, 0]$ we have M_t is locally $(R, S_t, L_0, \gamma_0, \xi_0)$-plane-like around 0 for some subset $S_t \subset B^n_R(0) \cap B^n_r(\hat{a}_0)$ there exists a function $f \in C^\infty([-R^2, 0] \times B^n_R(0) \setminus B^n_r(\hat{a}_0))$ with $\mathrm{lip}(f) \le L_0$, $\sup|f| \le \gamma_0 R$ and

$$M_t \cap C_R(0) \setminus C_r(a) = \mathrm{graph} f(t,\cdot) \cap C_R(0) \setminus C_r(a_0). \tag{12.49}$$

Let $C_{12.11}$ be the constant from Theorem 12.11. Consider the radius

$$r_1 := C_{12.11}^{-1} L_0^{-5} \sqrt{\lambda} \beta R$$

By choice of variables the sup bound of f implies

$$\sup|f| \le \gamma_0 R = \gamma_0 C_{12.11} L_0^5 \lambda^{-\frac{1}{2}} \beta^{-1} r_1 \le C_{12.11} \sqrt{\lambda} L_0^{-1} r_1 \le \sqrt[3]{\lambda} L_0^{-1} r_1, \tag{12.50}$$

where we used $\gamma_0 \le \beta \lambda L_0^{-6}$ and we chose λ small depending on $C_{12.11}$. Consider

$$y \in B^n_{\frac{R}{2}}(0) \setminus B^n_{r_1 + \sqrt{\lambda} L_0^{-1} R}(\hat{a}_0) \times \{0\}.$$

By choice of r_1 we have $C_{12.11} L_0^4 r_1 = \sqrt{\lambda} L_0^{-1} \beta R$. Then by $\beta \le 1$, $L_0 \ge 1$ and for $\sqrt{\lambda} \le 2^{-1}$ we see

$$C_{C_{12.11} L_0^4 r_1}(y) \subset C_R(0) \setminus C_r(a_0).$$

By choice of r_1 we also have $-C_{12.11}^2 L_0^8 r_1^2 = -\beta^2\lambda L_0^{-2} R^2 \geq -\beta^2 R^2$, as $\lambda, L_0^{-1} \leq 1$. In view of (12.49) and (12.50), we can use Theorem 12.11 with $a = y$, $s_0 = 0$, $L = \lambda^{-\frac{1}{3}} L_0$ and $\rho = r_1$ to obtain $|Df(0,\hat{y})| \leq C_n \sqrt[3]{\lambda} L_0^{-1}$. Then as y was arbitrary we have

$$|Df(0,\hat{y})| \leq l \tag{12.51}$$

for all $\hat{y} \in B^n_{\frac{R}{2}}(0) \setminus B^n_{r+\sqrt{\lambda}L_0^{-1}R}(\hat{a}_0)$, where we used $L_0 \geq 1$ and we chose λ small depending on C_n and l. This establishes the desired Lipschitz bound but only outside $B^n_{r+\sqrt{\lambda}L_0^{-1}R}(\hat{a}_0)$. Next we want to establish a bound for the μ_0 measure of this set.

As M_0 is locally $(R, S_0, L_0, \gamma, \xi_0)$-plane-like around 0 for some $S_0 \subset B^n_R(0) \cap B^n_r(\hat{a}_0)$ we obtain an $\tilde{f} \in C^\infty(B^n_R(0) \setminus S_0)$ with $\mathrm{lip}(f) \leq L_0$ and

$$M_0 \cap C_R(0) \setminus S_0 \times \mathbb{R} = \mathrm{graph}\tilde{f} \tag{12.52}$$

and also

$$\mu_0(S_0 \times \mathbb{R}) \leq \xi_0 R^n. \tag{12.53}$$

Consider the set

$$S = B^n_{\frac{R}{2}}(0) \cap B^n_{r+\sqrt{\lambda}L_0^{-1}R}(\hat{a}_0).$$

With (12.52) and (12.53) we can estimate

$$\begin{aligned}\mu_0(S \times \mathbb{R}) &\leq \mu_0(S_0 \times \mathbb{R}) + \int_{S\setminus S_0} \sqrt{1 + |Df(\hat{y})|^2} d\mathscr{L}^n(\hat{y}) \\ &\leq \xi_0 R^n + C_n L_0 \left(r + \sqrt{\lambda}L_0^{-1}R\right)^n.\end{aligned}$$

Thus by $\xi_0 \leq \lambda$, $r = \lambda L_0^{-1} R$ and $L_0 \geq 1$

$$\begin{aligned}\mu_0(S \times \mathbb{R}) &\leq \xi_0 R^n + C_n L_0 \left(r + \sqrt{\lambda}L_0^{-1}R\right)^n \\ &\leq C_n\sqrt{\lambda}R^n = C_n\sqrt{\lambda}2^{-n}R^n.\end{aligned} \tag{12.54}$$

In view of (12.49), (12.51) and (12.54), we see that for $S = B^n_{r+\sqrt{\lambda}L_0^{-1}R}(0)$ and $\xi = C_n\sqrt{\lambda}$, we have $M_{-\tau^2R^2}$ is locally $(2^{-1}R, S, l, 2\gamma_0, \xi)$-plane-like around 0. For λ small enough depending on δ we have

$$2\gamma_0 \leq 2\lambda L_0^{-5} \leq \delta^{\frac{3}{2}} \quad \text{and} \quad \xi = C_n\sqrt{\lambda} \leq \delta^{2n+12}.$$

As $\bigcup_{t\in[s_0-R^2,s_0+2\tau R^2]} \partial M_t \cap C_{2R}(0) = \emptyset$ and (M_t) moves by smooth mean curvature flow, we have that $\mu_t = \mathscr{H}^n \llcorner M_t \cap C_{2R}(0)$ is a Brakke flow. Then we can use Theorem 11.7 with $y_0 = 0$, $s_0 = 0$, $\gamma = 2\gamma_0$ and $\rho = 2^{-1}R$. Choose $\delta \leq \delta_0$ and small enough depending on ϵ, to obtain the desired gradient and height bound for g. Then set $\kappa := \delta_0\delta$ and choose $\lambda \leq \delta_0\sqrt{\kappa}2^{-1}$. This establishes the result. $\square$

Like for Brakke flows the local result yields a result for globally graph-like flows (recall Definition 11.14).

12.15 Lemma. *For every $\epsilon \in (0,1)$ there exists a $\Lambda_0 \in (1,0)$ such that for all $R_0 \in (0,\infty)$, $L_0 \in [1,\infty)$, $\beta_0 \in (0,1]$, $T_0 \in (R_0^2,\infty)$, $\Gamma_0 \in (0,\beta_0 L_0^{-6}R_0]$, $\Xi_0 \in (0,R_0]$, $t_0 \in \mathbb{R}$, $a_0 \in \mathbb{R}^{n+1}$ the following holds: Let $(M_t)_{t\in[t_0-\beta_0^2\Lambda_0^2R_0^2,t_0+T_0]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$ without boundary. Suppose for all $t \in [t_0-\beta_0^2\Lambda_0^2R_0^2,t_0]$ we have: M_t is globally (S_t,L_0,Γ_0,Ξ_0)-graph-like for some $S_t \subset B^n_{R_0}(\hat{a}_0)$. Then one of the following two statements holds:*

1. *For $I = [t_0+R_0^2, t_0+T_0]$ there exists an $f \in C^\infty\left(I\times\mathbb{R}^n,\mathbb{R}^k\right)$ with* $\mathrm{lip} f \leq \epsilon$ *and*
$$\mathrm{spt}\mu_t = \mathrm{graph}(f(t,\cdot))$$
for all $t \in I$.

2. *There exists a $(t,y) \in [t_0,t_0+R_0^2] \times \mathbb{R}^{n+1}$ with $\mu_t\left(C_{R_0}(y)\right) = 0$*

Proof. For given $\epsilon \in (0,1)$ let λ,κ be from Proposition 12.13. Set
$$R := \Lambda_0 R_0$$
and consider arbitrary $y \in \mathbb{R}^n \times \{0\}$, $t \in [t_0-\beta_0^2\Lambda_0^2R_0^2,t_0]$. As $\Xi_0 \leq R_0 \leq R$ we can use Remark 11.15.1, to see that μ_t is locally $(R, B^n_R(\hat{y}) \cap S_t, L_0, \beta_0 L_0^{-6}\Lambda_0^{-1}, \Lambda_0^{-1})$-plane-like around y. Here we used $\Gamma_0 \in (0,\beta_0 L_0^{-6}R_0]$ and $\Xi_0 \in (0,R_0]$. Note that
$$B^n_R(\hat{y}) \cap S_t \subset B^n_R(\hat{y}) \cap B^n_{\Lambda_0^{-1}R}(\hat{a}_0).$$
If we choose
$$\Lambda_0 \geq \lambda^{-1},$$
all the conditions for Proposition 12.13 used with $\gamma = \beta_0 L_0^{-6}\Lambda_0^{-1}$, $\xi = \Lambda_0^{-1}$) and $r = \Lambda_0^{-1}$ are satisfied. Then using Proposition 12.13 for every $y \in \mathbb{R}^n \times \{0\}^k$, leads to one of the following two cases

Alternative 1:
For $I_0 = [t_0+\kappa R^2, t_0+2\kappa R^2]$ and for every $y \in \mathbb{R}^n$ there exists a smooth

function $f_y \in C^\infty\left(I \times C_{\lambda R}(y), \mathbb{R}^k\right)$ with $\text{lip}(f_y) \leq \epsilon$ and $\text{graph}(f_y(t,\cdot)) = \text{spt}\mu_t$ for all $t \in I$. Then all these functions have to be restrictions of a singel function f defined on all $\mathbb{R}^n$ with $\text{graph}(f(t,\cdot)) = \text{spt}\mu_t$ for all $t \in I$.

Moreover we can extend f to later times such that $\text{graph}(f(t,\cdot))$ moves by smooth mean curvature flow for all $t \in [t_0 + \kappa R^2, t_0 + T]$, see [EH1, 4.4]. As also (M_t) is a smooth mean curvature flow and $M_{t_0+\kappa R^2}$ coincides with $\text{graph}(f(t_0 + \kappa R^2, \cdot))$ they have to be the same for all later times as well.

Alternative 2:
There exists a $y \in \mathbb{R}^n$ such that $\mu_{t_0+3\kappa R^2}\left(C_{\lambda R}(y)\right) = 0$.

As $\Lambda_0 \geq \lambda^{-1}$ we have $\lambda R \geq R_0$. Also choose $\Lambda_0 \geq \sqrt{3\kappa}$. Then we have $3\kappa R^2 \leq R_0^2$ which establishes the result. □

In Lemma 12.15 we assume that the flow is globally plane-like for a long time. Actually this condition can be weakened.

12.16 Proposition. *There exists a $P \in (1,\infty)$ such that for every $\epsilon \in (0,1)$ there exists a $\Lambda \in (1,0)$ such that for all $r \in (0,\infty)$, $L \in [1,\infty)$, $\beta \in (0,1]$, $T \in (\Lambda^2\beta^{-2}L^P r^2, \infty)$, $\Gamma \in (0,r]$, $\Xi \in (0,r]$, $t_0 \in \mathbb{R}$, $a \in \mathbb{R}^{n+1}$ the following holds: Let $(M_t)_{t\in[-\beta^2r^2,T]}$ be an embedded mean curvature flow in $\mathbb{R}^{n+1}$ without boundary. Suppose for all $t \in [-\beta^2 r^2, 0]$ we have: M_t is globally (S_t, L, Γ, Ξ)-plane-like for some $S_t \subset B_r^n(\hat{a})$. Then one of the following two statements holds:*

1. *For $I = [\Lambda^2 L^P \beta^{-2} r^2, T]$ there exists an $f \in C^\infty(I \times \mathbb{R}^n)$ with $\text{lip} f \leq \epsilon$ and*

$$\text{spt}\mu_t = \text{graph}(f(t,\cdot))$$

for all $t \in I$.

2. *There exists a $(t, y) \in [0, \Lambda^2 L^P \beta^{-2} r^2] \times \mathbb{R}^{n+1}$ with $\mu_t\left(C_{\Lambda r}(y)\right) = 0$*

Proof. Let $\epsilon \in (0,1)$ be given and let Λ_0 be the quantity from Lemma 12.15 corresponding to this ϵ. As M_0 is globally (S_t, L, Γ, Ξ)-plane-like we can use Lemma 11.16 to obtain

$$\text{spt}\mu_t \subset \left\{x \in \mathbb{R}^{n+1} : |x_{n+1}| \leq C_n \Xi + \Gamma\right\}$$

for all $t \in [C_n \Xi^2, T]$. In particular as $\Gamma \leq r$ and $\Xi \leq r$, there exists a constant $C_1 \in (1,\infty)$ such that

$$|x_{n+1}| \leq C_1 4^{-6} r \tag{12.55}$$

for all $x \in M_t$ for all $t \in [C_1 r^2, T]$. Here we chose $\Lambda^2 \geq C_1$, so this time interval is non-empty. We want to use Proposition 12.12 with

$$\rho_0 := 2C_1\Lambda_0 L^6 r$$
$$\tau := (2C_1\Lambda_0 L^6)^{-1}\beta.$$

First calculate

$$\tau^2\rho_0^2 = \beta^2 r^2. \tag{12.56}$$

As for all $t \in [-\beta^2 r^2, 0]$ we have M_t is globally (S_t, L, Γ, Ξ)-plane-like for some $S_t \subset B_r^n(\hat{a})$ there exists an $f \in C^\infty([-\beta^2 r^2, 0] \times \mathbb{R}^n \setminus B_r^n(\hat{a}))$ with $\mathrm{lip}(f) \leq L$, $\sup|f| \leq \Gamma$ and

$$M_t \cap \mathbb{R}^{n+1} \setminus C_r(a) = \mathrm{graph} f(t,\cdot) \cap \mathbb{R}^{n+1} \setminus C_r(a). \tag{12.57}$$

for all $t \in [-\beta^2 R^2, 0]$. Let $C_{12.12}$ be the constant from Proposition 12.12. Considering the radius

$$\rho_1 := C_{12.12} L^8 \tau^{-1}\rho_0 = 4C_1^2 C_{12.12}\Lambda_0 L^{20}\beta^{-1} r = C_n\Lambda_0 L^{20}\beta^{-1} r \tag{12.58}$$

and $y \in \mathbb{R}^{n+1} \setminus C_{r+\rho_1}(a)$, we have

$$C_{\rho_1}(y) \subset \mathbb{R}^{n+1} \setminus C_r(a). \tag{12.59}$$

Then by (12.56), (12.57), (12.58) and (12.59), we can use Proposition 12.12 to obtain a $g_y \in C^\infty\left([0,\rho_0^2] \times B_{\rho_0}^n(\hat{y})\right)$ with $\mathrm{lip}(g_y) \leq 4L$ and

$$M_t \cap C_{\rho_0}(y) = \mathrm{graph} g_y(t,\cdot) \cap C_{\rho_0}(y).$$

for all $t \in [0, \rho_0^2]$. As y was arbitrary in $\mathbb{R}^{n+1} \setminus C_{r+\rho_1}(a)$ and by (12.57) we obtain a function $g \in C^\infty\left([0,\rho_0^2] \times \mathbb{R}^n \setminus B_{r+\rho_1}^n(\hat{a})\right)$ with $\mathrm{lip}(g) \leq 4L$ and

$$M_t \cap \mathbb{R}^{n+1} \setminus C_{r+\rho_1}(a) = \mathrm{graph} f(t,\cdot) \cap \mathbb{R}^{n+1} \setminus C_{r+\rho_1}(a). \tag{12.60}$$

for all $t \in [-\beta^2 r^2, \rho_0^2]$.

Next we want to bound the measure inside the set $C_{r+\rho_1}(a)$ for all times $t \in [0, \rho_0^2]$. By Remark 3.12.2 with $R = 4\sqrt{n}\rho_1$ and $\kappa = \frac{1}{2}$ we obtain

$$\mu_t\left(B_{2\sqrt{n}\rho_1}(a)\right) \leq 8\mu_0\left(B_{4\sqrt{n}\rho_1}(a)\right)$$

for all $t \in [0, \rho_0^2]$, where we used $\rho_0 \leq \rho_1$. Then using (12.55) and $2C_1 r \leq \rho_0 \leq \rho_1$ this yields

$$\mu_t\left(C_{r+\rho_1}(a)\right) \leq \mu_t\left(B_{2\sqrt{n}\rho_1}(a)\right) \leq 8\mu_0\left(B_{4\sqrt{n}\rho_1}(a)\right) \tag{12.61}$$

for all $t \in [2^{-1}\rho_0^2, \rho_0^2]$.

As M_0 is globally (S_0, L, Γ, Ξ)-bounded for some $S_0 \subset B_r^n(\hat{a})$ there exists an $f_0 \in C^\infty(\mathbb{R}^n \setminus B_r^n(\hat{a}))$ with $\mathrm{lip}(f_0) \leq L$ and

$$M_0 \cap \mathbb{R}^{n+1} \setminus C_r(a) = \mathrm{graph} f_0 \cap \mathbb{R}^{n+1} \setminus C_r(a) \tag{12.62}$$

and as $\Xi \leq r$ we also have

$$\mu_0 (S_0 \times \mathbb{R}) = \mu_0 ((S_0 \cap B_r^n(\hat{a})) \times \mathbb{R}) \leq r^n. \tag{12.63}$$

With (12.62) and (12.63) we can estimate

$$\begin{aligned} \mu_0 \left(C_{4\sqrt{n}\rho_1}(a)\right) &\leq \mu_0 (S_0 \times \mathbb{R}) + \int_{B_{4\sqrt{n}\rho_1}(\hat{a})\setminus S_0} \sqrt{1 + |Df(\hat{y})|^2} d\mathscr{L}^n(\hat{y}) \\ &\leq r^n + C_n L \rho_1^n. \end{aligned}$$

By $\rho_1 = C_n \Lambda_0 L^{20} \beta^{-1} r$ we obtain

$$\mu_0 \left(C_{4\sqrt{n}\rho_1}(a)\right) \leq C_n \left(\Lambda_0 L^{21} \beta^{-1} r\right)^n,$$

where we used $L \geq 1$. Thus by (12.61) we have

$$\mu_t \left(C_{r+\rho_1}(a)\right) \leq \left(C_2 \Lambda_0 L^{21} \beta^{-1} r\right)^n$$

for all $t \in [2^{-1}\rho_0^2, \rho_0^2]$, for some constant $C_2 \in (1, \infty)$. In particular we see

$$\sup_{\hat{y} \in \mathbb{R}^n} \mu_t \left(\left(B_\rho^n(\hat{y}) \cap B_{r+\rho_1}^n(\hat{a})\right) \times \mathbb{R}^k\right) \leq C_2 \Lambda_0 L^{21} \beta^{-1} r \rho^{n-1} \tag{12.64}$$

for all $\rho \in [C_2 \Lambda_0 L^{21} \beta^{-1} r, \infty)$ and all $t \in [0, \rho_0^2]$.

Now we have all the ingredients to finally use Lemma 12.15. By (12.55), (12.60) and (12.64) we can conclude that for all $t \in [2^{-1}\rho_0^2, \rho_0^2]$ the manifold M_t is globally $(S, 4L, \Gamma_0, \Xi_0)$-bounded with

$$\begin{aligned} S &:= B_{r+\rho_1}^n(\hat{a}) \\ \Gamma_0 &:= C_1 4^{-6} r \\ \Xi_0 &:= C_2 \Lambda_0 L^{21} \beta^{-1} r. \end{aligned}$$

Now set

$$\begin{aligned} R_0 &:= C_2 \Lambda_0 L^{21} \beta^{-1} r \\ \beta_0 &:= \beta (C_2 \Lambda_0)^{-1} C_1 L^{-15}. \end{aligned}$$

Then we can estimate

$$\beta_0^2\Lambda_0^2R_0^2 = C_1^2\Lambda_0^2L^{12}r^2 = 2^{-2}(2C_1\Lambda_0L^6r)^2 \leq 2^{-1}\rho_0^2$$

so the time interval $[2^{-1}\rho_0^2, \rho_0^2]$ where the flow is globally plane-like is long enough. To obtain the height bound we calculate

$$\Gamma_0 = C_1 4^{-6}r = C_1 4^{-6}L^{-6}\Lambda_0 L^{21}\beta^{-1}\beta\Lambda_0^{-15}L^{-9}r = (4L)^{-6}\beta_0 R_0.$$

So we can use Lemma 12.15 with $t_0 = \rho_0^2$. This establishes the result for $\Lambda \geq C_2\Lambda_0$ and $P = 42$. $\square$

A Appendix

A.1 Lipschitz Functions

A.1 Definition. A function $f : \Omega \to \mathbb{R}^m$ for $\Omega \subset \mathbb{R}^n$ is called Lipschitz (continuous), if there exists an $L \in \mathbb{R}^+$ such that:

$$|f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in \Omega. \tag{A.1}$$

Then $\mathrm{lip}(f)$ denotes the smallest such L.

A.2 Remark. Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \to \mathbb{R}^k$ a Lipschitz function. Then there exists $F : \mathbb{R}^n \to \mathbb{R}^k$ with $F(x) = f(x)\ \forall x \in \Omega$ and $\mathrm{lip}(F) \leq \sqrt{k}\,\mathrm{lip}(f)$.

A.3 Theorem (Rademacher)**.** *Every Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable $\mathscr{L}^n$-almost everywhere, that means there exists a set $\Omega \subset \mathbb{R}^n$ with $\mathscr{L}^n(\mathbb{R}^n \setminus \Omega) = 0$ and for every $x \in \Omega$ there exists $Df_x : \mathbb{R}^n \to \mathbb{R}$ with*

$$\lim_{h \to 0} h^{-1}\left(f(hv + x) - f(x) - hDf_x \cdot v\right) = 0 \tag{A.2}$$

for every $v \in S^{n-1}$

A.4 Proposition. *There exists a constant $C \in (1, \infty)$ such that, for every $\Omega \subset \mathbb{R}^n$ the following holds: Let $f \in C^{0,1}\left(\Omega, \mathbb{R}^k\right)$ and define $F \in C^{0,1}\left(\Omega, \mathbb{R}^{n+k}\right)$ by $F(y) := (x, f(y))$. Then for almost every $x \in \Omega$ the Jacobian $JF = \sqrt{\det(DF^T DF)}$ is well defined and satisfies the inequality*

$$1 \leq JF(x)^2 \leq 1 + CL^{2n} \sum_{j=1}^{k} |Df_j(x)|^2, \tag{A.3}$$

where $L := \max\{\mathrm{lip}(f), 1\}$

Proof. By Theorem A.3 DF and hence JF are well defined almost everywhere. Thinking of DF^T as a matrix with columns a_i we obtain

$$DF^T = (a_1 \cdots a_n \cdots a_{n+k}) = \left(e_1 \cdots e_n\ Df_1^T \cdots Df_k^T\right),$$

so the columns of DF^T are the n basis vectors of $\mathbb{R}^n$ and the gradients of f_j. To calculate the Jacobian we use a formula for the product of matrices

$$\det(DF^T DF) = \sum_{1 \leq l_1 < \ldots < l_n \leq n+k} \left(\det\left(A^{l_1 \cdots l_n}\right)\right)^2,$$

where $A^{l_1\cdots l_n} = (a_{l_1} \cdots a_{l_n})$, which means $A^{l_1\cdots l_n}$ is the matrix containing the columns l_j of DF^T. The summand $l_j = j$ gives $A^{1\cdots n} = E_n$, which leads to

$$\det(DF^T DF) = 1 + \sum_{\substack{1\le l_1<\ldots<l_n\le n+k \\ n+1\le l_n}} \left(\det\left(A^{l_1\cdots l_n}\right)\right)^2 . \tag{A.4}$$

This verifies the lower bound. So it remains to consider matrices $A^{l_1\cdots l_n}$ with $n+1 \le l_n$, which are those containing at least one gradient vector as a column. Let $l_n = n+j$, such that $A^{l_1\cdots l_n} = \left(a_{l_1} \cdots a_{l_{n-1}} Df_j^T\right)$ and calculate the determinant through developing by the last column

$$\det A^{l_1\cdots l_n} = \sum_{1\le i\le n} (-1)^{i+j} \frac{df_j}{dx_i} \det\left(A^{l_1\cdots l_n}\right)'_{ij} ,$$

where $\left(A^{l_1\cdots l_n}\right)'_{ij}$ is constructed by erasing the jth column and the ith line from $A^{l_1\cdots l_n}$. The determinant of this matrix can now be estimated directly by using the Leibniz formula. As all entries of $\left(A^{l_1\cdots l_n}\right)'_{ij}$ are bounded by L we obtain

$$\det\left(A^{l_1\cdots l_n}\right)'_{ij} = \sum_{\sigma\in S_n} (-1)^{i+j} \mathrm{sign}(\sigma)\cdot a_{l_1\sigma(l_1)} \cdots a_{l_{n-1}\sigma(l_{n-1})} \le (n-1)! L^{n-1}.$$

This establishes an estimate for the determinant of $dF^* dF$. Inserting this into (A.4) yields

$$\begin{aligned}
\det(DF^T DF) &\le 1 + \sum_{j=1}^{k} \left(\sum_{\substack{1\le l_1<\ldots<l_n\le n+k \\ n+j=l_n}} \left(\sum_{1\le i\le n} (n-1)! L^{n-1} \frac{df_j}{dx_i} \right)^2 \right) \\
&\le 1 + (n+k)!(n!)^2 L^{2n} n^2 \sum_{j=1}^{k} |Df_j|^2,
\end{aligned}$$

which establishes the result. $\square$

A.5 Proposition. *There exists a constant $C \in (1,\infty)$ such that for every open subset $\Omega \subset \mathbb{R}^n$ the following holds: Let $f \in C^{0,1}\left(\Omega, \mathbb{R}^k\right)$ and the n-rectifiable set $M = \mathrm{graph}(f)$. Then*

$$|Df_m(\hat{x})|^2 \le 2nL^2 |\nabla^M x_{n+m}|^2 \tag{A.5}$$

for every $1 \le m \le k$, for almost all $x \in M$, where $L := \max\{\mathrm{lip}(f), 1\}$

Proof. Let $1 \leq m \leq k$ and $1 \leq i \leq n$ be fixed. Consider $F(x) = (x, f(x))$, then $\partial_i F := \frac{\partial}{\partial x_i} F$ is tangential. So for every $x \in M$ we have an orthonormal basis $\tau_1, \ldots, \tau_n$ of $T_x M$ with $\tau_1 = \frac{\partial_i F}{|\partial_i F|}$. This lets us calculate

$$|\nabla^M x_{n+m}|^2 = |\nabla^M f_m(\hat{x})|^2 \geq |\langle Df_m(\hat{x}), \tau_1 \rangle \tau_1|^2 = \frac{|\frac{\partial}{\partial x_i} f_m(\hat{x})|^2}{|\partial_i F|^2}.$$

Then by summing over i and by $|\partial_i F|^2 \leq 1 + L^2$ we obtain the result. □

A.6 Proposition (Zheng, [I, 6.6]). *There exists a constant $C \in (1, \infty)$ such that for every open subset $\Omega \subset \mathbb{R}^n$ the following holds: Let $f \in C_c^2(\Omega, \mathbb{R}^+)$ Then the following estimates hold*

1. *For $\hat{x} \in \{f > 0\}$ the estimate*

$$\frac{|Df(\hat{x})|^2}{f(\hat{x})} \leq 2 \sup |D^2 f|$$

holds. In particular g defined by $g(\hat{x}) = \frac{|Df(\hat{x})|^2}{f(\hat{x})}$ for $\hat{x} \in \{f > 0\}$ and $g(\hat{x}) = 0$ for $\hat{x} \in \Omega \setminus \{f > 0\}$ is in $C_c^2(\Omega, \mathbb{R}^+)$

2. *For every $\hat{x} \in \Omega$ we can estimate*

$$|D|Df(\hat{x})||^2 \leq |D^2 f(\hat{x})|^2$$

Proof. 1. See [I, 6.6]

2. Calculate

$$\begin{aligned} |D|Df(\hat{x})||^2 &= \sum_{i=1}^n \left(\frac{\partial}{\partial x_i} |D(f(\hat{x}))| \right)^2 = \sum_{i=1}^n \left(\frac{Df(\hat{x})}{|Df(\hat{x})|} \frac{\partial}{\partial x_i} Df(\hat{x}) \right)^2 \\ &\leq \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} Df(\hat{x}) \right|^2 = |D^2 f(\hat{x})|^2, \end{aligned}$$

which establishes the result.

□

A.2 Projections And Tilt

Let T, B be n-dimensional subspaces of $\mathbb{R}^{n+k}$ then $|\pi_T - \pi_B|$ is called the tilt between two subspaces T and B. Let μ be a rectifiable n-varifold in $\mathbb{R}^{n+k}$. For $U \subset \mathbb{R}^{n+k}$ the term

$$\int_U |\pi_{T_x \mu} - \pi_T|^2 \, d\mu(x)$$

is called the tilt-excess of μ in U.

A.7 Remark. Consider two n-dimensional subspaces T_1, T_2 of $\mathbb{R}^{n+k}$

1. If $x \in T_1$ and $y \in T_2^\perp$ we can estimate the scalar product in the following way

$$\begin{aligned} x \cdot y &= (x - \pi_{T_2}(x)) \cdot y + \pi_{T_2}(x) \cdot y = (\pi_{T_1}(x) - \pi_{T_2}(x)) \cdot y + 0 \\ &\leq |\pi_{T_1} - \pi_{T_2}| \, |x||y| \end{aligned}$$

2. For the tilt the following identity holds:

$$\begin{aligned} |\pi_{T_1} - \pi_{T_2}| &= \left| -\pi_{T_1}^\perp + \pi_{T_2}^\perp + \pi_{T_1}^\perp + \pi_{T_1} - \pi_{T_2}^\perp - \pi_{T_2} \right| \\ &= \left| -\pi_{T_1}^\perp + \pi_{T_2}^\perp + E_{n+k} - E_{n+k} \right| = \left| \pi_{T_1}^\perp - \pi_{T_2}^\perp \right| \end{aligned}$$

3. Suppose $|\pi_{T_1} - \pi_{T_2}| \leq \epsilon \in (0, 1)$. Then we can calculate

$$|\pi_{T_1}(x)| - |\pi_{T_2} \circ \pi_{T_1}(x)| \leq |\pi_{T_1}(x) - \pi_{T_2} \circ \pi_{T_1}(x)| \leq \epsilon |\pi_{T_1}(x)|,$$

so $|\pi_{T_1}(x)| \leq (1 - \epsilon)^{-1} |\pi_{T_2} \circ \pi_{T_1}(x)|$.

A.8 Remark. For an isometry $S \in SO(n + k)$ the following holds:

1. Let T be an n-dimensional subspace of $\mathbb{R}^{n+k}$. Let $(t_i)_{1 \leq i \leq n}$ be an orthonormal basis of T, then $(S(t_i))_{1 \leq i \leq n}$ is an orthonormal basis of $S(T)$, thus we can calculate

$$\pi_{S(T)}(x) = \sum_{i=1}^{n} (S(t_i) \cdot x) S(t_i) = S \left(\sum_{i=1}^{n} (t_i \cdot S^T(x)) t_i \right) = S \left(\pi_T(S^T x) \right).$$

2. Let $A, B \in \mathbb{R}^{n+k \times n+k}$. As $S(\partial B_1(0)) = \partial B_1(0)$ we can calculate

$$\|A - B\|_{op} = \|S \circ A - S \circ B\|_{op} = \|A \circ S - B \circ S\|_{op}.$$

In particular if $A, B \in SO(n + k)$, we have

$$\|A - B\|_{op} = \left\| B^T \circ A - E_{n+k} \right\|_{op} = \left\| B^T - A^T \right\|_{op}$$

3. Let T_1, T_2 be n-dimensional subspaces of $\mathbb{R}^{n+k}$. Combining statements 1 and 2 we can calculate

$$\begin{aligned} \|\pi_{T_1} - \pi_{T_2}\|_{op} &= \left\| S \circ (\pi_{T_1} - \pi_{T_2}) \circ S^T \right\|_{op} \\ &= \left\| S \circ \pi_{T_1} \circ S^T - S \circ \pi_{T_2} \circ S^T \right\|_{op} = \left\| \pi_{S(T_1)} - \pi_{S(T_2)} \right\|_{op} \end{aligned}$$

A.9 Proposition. *Let T_1, T_2 be m-dimensional subspaces of $\mathbb{R}^{n+k}$*

1. *If there exists a $v \in T_1 \cap T_2^{\perp}$ with $v \neq 0$, then there also exists a $w \in T_1^{\perp} \cap T_2$ with $w \neq 0$,*

2. *We always have $\|\pi_{T_1} - \pi_{T_2}\|_{op} \leq 1$ and $\|\pi_{T_1} - \pi_{T_2}\|_{op} = 1$, if and only if there exists a $v \in T_1^{\perp} \cap T_2$ with $v \neq 0$.*

3. *Let $(b_i)_{1 \leq i \leq m}$ be an orthonormal basis of T_1 and suppose $\|\pi_{T_1} - \pi_{T_2}\|_{op} < 1$, then the vectors $(\pi_{T_2}(b_i))_{1 \leq i \leq m}$ form a basis of T_2.*

4. *Consider $x \in T_2$ and suppose $\|\pi_{T_1} - \pi_{T_2}\|_{op} < 1$, then there exists exactly one point in $T_1 \cap \left\{x + v, v \in T_2^{\perp}\right\}$*

5. *In particular for $f \in C^{0,1}\left(\Omega, \mathbb{R}^k\right)$, $\Omega \subset \mathbb{R}^{n+k}$, $M := \operatorname{graph}(f)$ and $\hat{x} \in \Omega$, $x = (\hat{x}, f(\hat{x}))$ such that $Df(\hat{x})$ and T_xM exist, we have $\|\pi_{T_xM} - \pi_{\mathbb{R}^n}\|_{op} < 1$.*

Proof. 1. Suppose there exists a $v \in T_1 \cap T_2^{\perp}$, $v \neq 0$. As T_1 is m-dimensional and $T_2^{\perp}$ is $n+k-m$-dimensional and both subspaces share one direction we have that $T_1 \cup T_2^{\perp}$ is at most $n+k-1$-dimensional. Thus there exists a $w \in \left(T_1 \cup T_2^{\perp}\right)^{\perp}$ with $w \neq 0$. Then $w \in T_1^{\perp} \cap T_2$.

2. Let $v \in \mathbb{R}^{n+k}$ with $|v| = 1$, then we can find an orthonormal basis $(b_i)_{1 \leq i \leq n+k}$ of $\mathbb{R}^{n+k}$ with $b_1 = v$. Calculate

$$
\begin{aligned}
&\sum_{i=1}^{n+k} \left(\left(v - \pi_{T_1^{\perp}}(v) - \pi_{T_2}(v)\right) \cdot b_i\right)^2 \\
&= \sum_{i=1}^{n+k} \left(\delta_{1i} - \left(\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)\right) \cdot b_i\right)^2 \\
&= \sum_{i=1}^{n+k} \left(\delta_{1i} - 2\delta_{1i} \left(\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)\right) \cdot b_i + \left(\left(\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)\right) \cdot b_i\right)^2\right) \\
&= 1 - 2\left(\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)\right) \cdot v + \sum_{i=1}^{n+k} \left(\left(\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)\right) \cdot b_i\right)^2 \\
&= 1 - 2\left(|\pi_{T_1^{\perp}}(v)|^2 + |\pi_{T_2}(v)|^2\right) + |\pi_{T_1^{\perp}}(v) + \pi_{T_2}(v)|^2
\end{aligned}
$$

and combining this with the parallelogram law, we obtain

$$
\sum_{i=1}^{n+k} \left((\pi_{T_1}(v) - \pi_{T_2}(v)) \cdot b_i\right)^2 \leq 1 - |\pi_{T_1^{\perp}}(v) - \pi_{T_2}(v)|^2.
$$

As the $|\cdot|$-norm is independent of the choice of basis this yields

$$|\pi_{T_1}(v) - \pi_{T_2}(v)| = \sqrt{1 - |\pi_{T_1^\perp}(v) - \pi_{T_2}(v)|^2} \tag{A.6}$$

for all $v \in \mathbb{R}^{n+k}$ with $|v| = 1$. This directly implies $\|\pi_{T_1} - \pi_{T_2}\|_{op} \leq 1$.

Now let $v \in T_1^\perp \cap T_2$, $v \neq 0$, then

$$|\pi_{T_1}(v) - \pi_{T_2}(v)| = |v|,$$

so $\|\pi_{T_1} - \pi_{T_2}\| \geq 1$ and as we already know ≤ 1, we have equality.

If on the other side $\|\pi_{T_1} - \pi_{T_2}\| = 1$, then there exists a $v \in \mathbb{R}^{n+k}$ with $|v| = 1$ such that $|\pi_{T_1}(v) - \pi_{T_2}(v)| = 1$. This requires equality in (A.6), so we have

$$\pi_{T_1^\perp}(v) = \pi_{T_2}(v) \quad \text{and} \quad \pi_{T_2^\perp}(v) = \pi_{T_1}(v)$$

thus $\pi_{T_1^\perp}(v) \in T_1^\perp \cap T_2$ and $\pi_{T_1}(v) \in T_2^\perp \cap T_1$. Here we used that T_1 and T_2 are exchangeable in (A.6). As either $\pi_{T_1^\perp}(v) \neq 0$ or $\pi_{T_1}(v) \neq 0$ and by statement 1, this establishes statement 2.

3. Let $(b_i)_{1\leq i\leq m}$ be an orthonormal basis of T_1 and suppose $(\pi_{T_2}(b_i))_{1\leq i\leq m}$ do not form a basis of T_2. Then there exists a v in T_2 with $v \cdot \pi_{T_2}(b_i) = 0$ for all $i \in \{1, \dots, m\}$. As $v \in T_2$ this yields

   $$v \cdot b_i = \pi_{T_2}(v) \cdot b_i = v \cdot \pi_{T_2}(b_i) = 0$$

   for all $i \in \{1, \dots, m\}$, so $v \in T_1^\perp \cap T_2$, thus by statement 2, we have $\|\pi_{T_1} - \pi_{T_2}\| = 1$.

4. Let $x \in T_2$, $\|\pi_{T_1} - \pi_{T_2}\| < 1$ and let $(b_i)_{1\leq i\leq m}$ be an orthonormal basis of T_1. By statement 3 the vectors $(\pi_{T_2}(b_i))_{1\leq i\leq m}$ form a basis of T_2, so there exists an $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ such that $\sum_{i=1}^m \alpha_i \pi_{T_2}(b_i) = x$. Then the vector $w := \sum_{i=1}^m \alpha_i b_i$ is in T_1 and $\pi_{T_2}(w) = x$, thus $w \in T_1 \cap \{x + v, v \in T_2^\perp\}$, so this set is not empty.

   Now let $x \in T_2$ and suppose there exist $a, b \in T_1 \cap \{x + v, v \in T_2^\perp\}$ with $a \neq b$. Then $b - a \in T_1 \cap T_2^\perp$ and $b - a \neq 0$, thus by statement 2 $\|\pi_{T_1} - \pi_{T_2}\| = 1$.

5. Recall that $(e_i, \frac{\partial}{\partial x_i} f(\hat{x}))$ span T_xM. As each of this vectors has components in $\mathbb{R}^n$ we have $T_xM \cap \mathbb{R}^k = \{0\}$, so by statement 2 we can conclude $\|\pi_{T_xM} - \pi_{\mathbb{R}^n}\| < 1$.

$\square$

A.10 Lemma. *There exists $C \in (1, \infty)$ such that for every $\epsilon \in [0, 1]$ the following holds: Let $T = \operatorname{span}(t_i)_{1 \le i \le m}$ and $B = \operatorname{span}(b_i)_{1 \le i \le m}$ be m-dimensional subspaces of $\mathbb{R}^{n+k}$ with $|b_i - t_i| \le \epsilon$ for all $i \in \{1, \dots, m\}$, then the following holds:*

1. *If the $(b_i)_{1 \le i \le m}$ are orthonormal, then there exists an orthonormal basis $(\tau_i)_{1 \le i \le n}$ of T with $|b_i - \tau_i| \le C\epsilon$ for all $i \in \{1, \dots, m\}$.*
2. *If $|b_i \cdot b_j - \delta_{ij}| \le \frac{1}{C}$ for all $i, j \in \{1, \dots, m\}$, we have $|\pi_T - \pi_B| \le C\epsilon$.*

Proof. First we consider the case that the $(b_i)_{1 \le i \le m}$ are orthonormal. We may assume $b_i = \mathbf{e}_i$ for $i \in \{1, \dots, m\}$ and $B = \mathbb{R}^m$. We are going to prove the statement by induction. In case $m = 0$, we have $T = \mathbb{R}^m = \{0\}$ and everything trivially holds. Suppose both statements hold for $m - 1$ with constant $\tilde{C}$. Define $\tilde{T} = \operatorname{span}(t_i)_{1 \le i \le m-1}$ and

$$v := t_m - \pi_{\tilde{T}}(t_m).$$

By induction hypothesis there exists an orthonormal basis $(\tau_i)_{1 \le i \le m-1}$ of $\tilde{T}$ with $|\mathbf{e}_i - \tau_i| \le C\epsilon$ for all $i \in \{1, \dots, m-1\}$. By definition of v we directly see $v \cdot \tau_i = 0$ for all $i \in \{1, \dots, m-1\}$. Also by induction hypothesis we have $|\pi_{\tilde{T}} - \pi_{\mathbb{R}^{m-1}}| \le \tilde{C}\epsilon$. This lets us calculate

$$\begin{aligned} |v - \mathbf{e}_m| &\le |t_m - \mathbf{e}_m| + |\pi_{\tilde{T}}(t_m)| \le \epsilon + |\pi_{\tilde{T}} - \pi_{\mathbb{R}^{m-1}}|\,|t_m| + |\pi_{\mathbb{R}^{m-1}}(t_m)| \\ &\le \epsilon + \tilde{C}\epsilon(1 + \epsilon) + |\pi_{\mathbb{R}^{m-1}}(t_m - \mathbf{e}_m)| \le 2(1 + \tilde{C})\epsilon, \end{aligned}$$

where we used $\epsilon \le 1$ and $|\mathbf{e}_m - t_m| \le \epsilon$. In particular $|1 - |v|| \le 2(1 + \tilde{C})\epsilon$ so for $\tau_m := |v|^{-1}v$ we can estimate

$$|\mathbf{e}_m - \tau_m| = |\mathbf{e}_m - |v|^{-1}v| \le |\mathbf{e}_m - v| + |v - |v|^{-1}v| \le 4(1 + \tilde{C})\epsilon,$$

thus $(\tau_i)_{1 \le i \le m}$ provides the desired orthonormal basis. To estimate the tilt calculate for $x \in \mathbb{R}^{n+k}$ with $|x| = 1$

$$\begin{aligned} |\pi_T(x) - \pi_{\mathbb{R}^n}(x)| &\le |\pi_{\tilde{T}}(x) - \pi_{\mathbb{R}^{m-1}}(x)| + |(\tau_m \cdot x)\tau_m - (\mathbf{e}_m \cdot x)\mathbf{e}_m| \\ &\le \tilde{C}\epsilon + |(\tau_m \cdot x)(\tau_m - \mathbf{e}_m)| + |((\tau_m \cdot x) - (\mathbf{e}_m \cdot x))\mathbf{e}_m| \\ &\le \tilde{C}\epsilon + 8(1 + \tilde{C})\epsilon \end{aligned}$$

which establishes the estimate for m. Thus we showed result (1) and result (2) for some constant $C_1 \in (1, \infty)$ in the special case where the $(b_i)_{1 \le i \le m}$ are orthonormal.

Now suppose the $(b_i)_{1\leq i\leq m}$ satisfy $|b_i \cdot b_j - \delta_{ij}| \leq \frac{1}{C}$ for all $i \in \{1, \dots, m\}$. We can orthogonalize the (t_i) and (b_i) using the Gram-Schmidt process

$$\tau_1 := |t_1^m|^{-1} t_1, \quad \tilde{\tau}_i := t_i - \sum_{j=1}^{i-1} (t_i \cdot \tau_j)\tau_j, \ \tau_i = |\tilde{\tau}_i|^{-1}\tilde{\tau}_i$$
$$\nu_1 := |t_1^m|^{-1} b_1, \quad \tilde{\nu}_i := b_i - \sum_{j=1}^{i-1} (b_i \cdot \nu_j)\nu_j, \ \nu_i = |\tilde{\nu}_i|^{-1}\tilde{\nu}_i$$

for $i = 2, \dots, m$. Then $(\tau_i^m)_{1\leq i\leq n}$ forms an orthonormal basis of T and $(\nu_i^m)_{1\leq i\leq n}$ forms an orthonormal basis of B. We want to show by induction

$$|\tau_i - \nu_i| \leq (16C_1)^i \epsilon \tag{A.7}$$
$$|\nu_i - b_i| \leq (16(n+k))^i C^{-1} \tag{A.8}$$

for all $i \in \{1, \dots, m\}$. For $i = 1$ we can estimate

$$|\tau_1 - \nu_1| \leq \left| |b_1|^{-1} (t_1 - b_1) \right| + \left| \left(|t_1|^{-1} - |b_1|^{-1} \right) t_1 \right| \leq 4\epsilon,$$

where we used $|b_1| \geq \sqrt{1 - \frac{1}{C}} \geq \frac{1}{2}$ for $C \geq \frac{4}{3}$. Also we have

$$|b_1 - \nu_1| \leq ||b_1| - 1| = \frac{||b_1|^2 - 1|}{|b_1| + 1} \leq C^{-1}.$$

Now suppose (A.7) holds for all $j \in \{1, \dots, i-1\}$. Set $T_i := \mathrm{span}(\tau_j)_{1\leq j\leq i-1}$ and $B_i := \mathrm{span}(\nu_j)_{1\leq j\leq i-1}$. By induction hypothesis (A.7) and by result (2) for orthonormal vectors, we have

$$|\pi_{T_i} - \pi_{B_i}| \leq C_1 (16C_1)^{i-1} \epsilon.$$

Then we can estimate

$$\begin{aligned} |\tilde{\tau}_i - \tilde{\nu}_i| &= |t_i - \pi_{T_i}(t_i) - b_i + \pi_{B_i}(b_i)| \\ &\leq |t_i - b_i| + |\pi_{T_i}(t_i) - \pi_{T_i}(b_i)| + |\pi_{T_i}(b_i) - \pi_{B_i}(b_i)| \\ &\leq 2\epsilon + 2C_1(16C_1)^{i-1}\epsilon, \end{aligned}$$

where we used $|b_i| \leq \sqrt{1 + C^{-1}} \leq 2$. So we have

$$|\tilde{\tau}_i - \tilde{\nu}_i| \leq 4^{-1}(16C_1)^i \epsilon, \quad ||\tilde{\tau}_i| - |\tilde{\nu}_i|| \leq 4^{-1}(16C_1)^i \epsilon. \tag{A.9}$$

By induction hypothesis (A.7) we can calculate

$$\begin{aligned}|\tilde{\nu}_i - b_i| &= \left|\sum_{j=1}^{i-1}(b_i \cdot b_j + b_i \cdot (\nu_j - b_j))\nu_j\right| \\ &\le (i-1)\left(C^{-1} + 2(16(n+k))^{i-1}C^{-1}\right),\end{aligned}$$

where we used $|b_i| \le \sqrt{1+C^{-1}} \le 2$. As $i \le m \le n+k$ this yields

$$|\tilde{\nu}_i - b_i| \le 4^{-1}(16(n+k))^i C^{-1}, \quad ||\tilde{\nu}_i| - 1| \le \frac{1}{2}, \tag{A.10}$$

where we chose C large depending on n and k for the second estimate. Combining (A.9) and (A.10) we can estimate

$$|\tau_i - \nu_i| \le \left||\tilde{\nu}_i|^{-1}(\tilde{\tau}_i - \tilde{\nu}_i)\right| + \left|\left(|\tilde{\tau}_i|^{-1} - |\tilde{\nu}_i|^{-1}\right)\tilde{\tau}_i\right| \le (16C_1)^i\epsilon,$$

where we used (A.10) to estimate $|\tilde{\nu}_i| \ge \frac{1}{2}$. By (A.10) we can also estimate

$$|\nu_i - b_i| \le \left||\tilde{\nu}_i|^{-1}(b_i - \tilde{\nu}_i)\right| + \left|\left(1 - |\tilde{\nu}_i|^{-1}\right)b_i\right| \le (16(n+k))^i C^{-1}$$

where we used $|b_i| \le \sqrt{1+C^{-1}} \le 2$. This completes our induction argument, which establishes the result. □

A.11 Proposition. *In case $T = \mathbb{R}^n$ (which we identify with $\mathbb{R}^n \times \{0\}^k$) and if $x \in \mathrm{spt}\mu$ admits an approximate tangent space $T_x\mu$, the following identity holds*

$$|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 = 2\sum_{j=1}^{k}|\nabla^\mu x_{n+j}|^2 = 2\sum_{j=1}^{k}|\pi_{T_x\mu}(\mathbf{e}_{n+j})|^2 = 2\sum_{l=1}^{n}\sum_{j=1}^{k}|\tau_l \cdot \mathbf{e}_{n+j}|^2,$$

where $(\tau_i)_{1\le i\le n}$ is an orthonormal basis of $T_x\mu$. Another identity for $|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|$ is

$$|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 = 2\sum_{i=1}^{n}\left|\pi^\perp_{T_x\mu}(\mathbf{e}_i)\right|^2 = 2\sum_{l=1}^{k}\sum_{i=1}^{n}|\nu_l \cdot \mathbf{e}_i|^2,$$

where $(\nu_j)_{1\le j\le k}$ is an orthonormal basis of $T_x\mu^\perp$.

Proof. The first equality is obtained by the following calculation where $p =$

π_T and $p_x = \pi_{T_x\mu}$.

$$
\begin{aligned}
\left|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}\right|^2 &= \operatorname{tr}\left((p_x - p)^*(p_x - p)\right) \\
&= \sum_{i=1}^{n+k} \mathbf{e}_i \cdot \left((p_x - p)^*(p_x - p)\mathbf{e}_i\right) \\
&= \sum_{i=1}^{n+k} (p_x - p)\mathbf{e}_i \cdot (p_x - p)\mathbf{e}_i \\
&= \sum_{i=1}^{n+k} \left(|p_x(\mathbf{e}_i)|^2 - 2p_x(\mathbf{e}_i) \cdot p(\mathbf{e}_i) + |p(\mathbf{e}_i)|^2\right) \\
&= \sum_{j=1}^{k} |p_x(\mathbf{e}_{n+j})|^2 + \sum_{i=1}^{n} \left(|p_x(\mathbf{e}_i)|^2 - 2p_x(\mathbf{e}_i) \cdot \mathbf{e}_i\right) + \sum_{i=1}^{n+k} |p(\mathbf{e}_i)|^2 \\
&= \sum_{j=1}^{k} |p_x(\mathbf{e}_{n+j})|^2 - \sum_{i=1}^{n} |p_x(\mathbf{e}_i)|^2 + \sum_{i=1}^{n+k} |p_x(\mathbf{e}_i)|^2 \\
&= 2\sum_{j=1}^{k} |p_x(\mathbf{e}_{n+j})|^2 = 2\sum_{j=1}^{k} \left|p_x\left(D(x_{n+j})\right)\right|^2 = 2\sum_{j=1}^{k} \left|\nabla^\mu x_{n+j}\right|^2 .
\end{aligned}
$$

This establishes the first identity. For the second identity we calculate

$$
\begin{aligned}
&\sum_{i=1}^{n}\sum_{j=1}^{k} \left|\tau_i \cdot \mathbf{e}_{n+j}\right|^2 - \sum_{i=1}^{n}\sum_{j=1}^{k} \left|\nu_j \cdot \mathbf{e}_i\right|^2 \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n+k} \left|\tau_i \cdot \mathbf{e}_j\right|^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} \left|\tau_i \cdot \mathbf{e}_j\right|^2 - \sum_{i=1}^{n}\sum_{j=1}^{k} \left|\nu_j \cdot \mathbf{e}_i\right|^2 \\
&= n - \sum_{l=1}^{n} \left(\sum_{i=1}^{n} \left|\tau_i \cdot \mathbf{e}_l\right|^2 + \sum_{j=1}^{k} \left|\nu_j \cdot e_l\right|^2 \right).
\end{aligned}
$$

Now use that $(\tau_i)_{1\le i\le n}$ and $(\nu_j)_{1\le j\le k}$ together span all of $\mathbb{R}^{n+k}$, so $\sum_{i=1}^{n} \left|\tau_i \cdot \mathbf{e}_l\right|^2 + \sum_{j=1}^{k} \left|\nu_j \cdot \mathbf{e}_l\right|^2 = |\mathbf{e}_l|^2$. Then

$$
\sum_{i=1}^{n}\sum_{j=1}^{k} \left|\tau_i \cdot \mathbf{e}_{n+j}\right|^2 - \sum_{i=1}^{n}\sum_{j=1}^{k} \left|\nu_j \cdot \mathbf{e}_i\right|^2 = n - \sum_{l=1}^{n} |\mathbf{e}_l|^2 = n - n = 0.
$$

This establishes the second identity. $\square$

Now combine this and A.11 with A.4 and A.5 to obtain:

A.12 Proposition. *There exists a constant $C \in (1,\infty)$ such that, for every $\Omega \subset \mathbb{R}^n$ the following holds: For $f \in C^{0,1}\left(\Omega, \mathbb{R}^k\right)$ set $F(y) := (y, f(y))$, $\mu := \mathscr{H}^n \llcorner \operatorname{graph}(f)$ and $L := \max\{\operatorname{lip}(f), 1\}$. Then at points $\hat{y} \in \Omega$ where f is differentiable and hence the tangent space $T_{F(\hat{y})}\mu$ exists the following is true:*

1. *We can estimate the tilt by $|Df|$ and vice versa, i.e.*
$$L^{-1}C^{-1}|Df(\hat{y})| \le |\pi_{T_{F(\hat{y})}\mu} - \pi_{\mathbb{R}^n}| \le C|Df(\hat{y})|$$
2. *The Jacobian $JF = \sqrt{\det(DF^T DF)}$ satisfies the inequality*
$$1 \le JF(\hat{y})^2 \le 1 + CL^{2n+2}\left|\pi_{T_{F(\hat{y})}\mu} - \pi_{\mathbb{R}^n}\right|^2. \tag{A.11}$$

Proof. For the first inequality of statement 1 combine A.5 and A.11. To show the second inequality note that $T_{F(y)}\mu = \operatorname{span}(e_i, D_i f(y))_{1\le i\le n}$ and we can estimate $|(e_i, 0) - (e_i, D_i f(y))| \le |Df(y)|$, so Proposition A.10.2 yields the estimate.

Statement 2 follows from the first inequality of statement 1 combined with A.4. □

A.13 Lemma (Tilt Bound Lemma, [B, 5.5])**.** *There exists a constant $C \in (0,\infty)$ such that for every open subset $U \subset \mathbb{R}^{n+k}$ the following holds: Let μ be an integer n-varifold in U with L^2-integrable mean curvature vector $\vec{H}$. Consider $g \in C^1_c(U,\mathbb{R})$, $f, h \in C^0_c(U,\mathbb{R})$ with $g^2 \le fh$. Then the estimate*
$$\beta_g^2 \le C\left(\alpha_f\gamma_h + \xi_g^2\right)$$
holds, where
$$\begin{aligned}
\alpha_f^2 &:= \int_U |\vec{H}(x)|^2 f(x)^2 d\mu(x),\\
\beta_g^2 &:= \int_U |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 g(x)^2 d\mu(x),\\
\gamma_h^2 &:= \int_U |\pi_{\mathbb{R}^k}(x)|^2 h(x)^2 d\mu(x),\\
\xi_g^2 &:= \int_U |\pi_{\mathbb{R}^k}(x)|^2 |\nabla^\mu g(x)|^2 d\mu(x).
\end{aligned}$$

Proof. Consider $X(x) := g(x)^2\pi_{\mathbb{R}^k}(x)$, then use Remarks 2.6 and A.11 to estimate
$$\begin{aligned}
2\operatorname{div}_\mu X(x) &= 4g(x)\nabla^\mu g(x)\cdot\pi_{T_x\mu}\left(\pi_{\mathbb{R}^k}(x)\right) + 2g(x)^2\sum_{j=n+1}^{n+k}\mathbf{e}_j\left(\pi_{T_x\mu}(\mathbf{e}_j)\right)\\
&\ge -4g(x)|\nabla^\mu g(x)|\,|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|\,|\pi_{\mathbb{R}^k}(x)| + g(x)^2\,|\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2
\end{aligned}$$

for all $x \in \operatorname{spt}\mu \cap U$ where $T_x\mu$ exists. By integrating we obtain

$$\beta_g^2 \leq 2 \int_U \operatorname{div}_\mu X \, d\mu + 4\beta_g \xi_g.$$

If now $\beta_g^2 \leq 8\beta_g\xi_g$ then $\beta_g^2 \leq 64\xi_g^2$ which yields the result for $C \geq 64$. Else we have $\beta_g^2 \leq 2\int_U \operatorname{div}_\mu X d\mu + \frac{1}{2}\beta_g^2$. Then by (2.2) and Hölder's inequality

$$\beta_g^2 \leq 4 \int_U \operatorname{div}_\mu X \, d\mu \leq 4 \int_U g(x)^2 |\pi_{\mathbb{R}^k}(x)||\vec{H}(x)| d\mu(x) \leq 4\alpha_f \gamma_h,$$

which establishes the result for $C \geq 4$. $\square$

A.14 Remark. Let μ be a rectifiable n-varifold in $\mathbb{R}^{n+k}$ and $x \in \operatorname{spt}\mu$ such that $T_x\mu$ exists, then we can estimate

$$\begin{aligned} 1 - \frac{|\pi_{T_x\mu}(\hat{x})|^2}{|\hat{x}|^2} &= \left(|\hat{x}|^2 - |\pi_{T_x\mu}(\hat{x})|^2\right) |\hat{x}|^{-2} \\ &= \left(|\hat{x} - \pi_{T_x\mu}(\hat{x})|^2 + 2\hat{x} \cdot \pi_{T_x\mu}(\hat{x}) - 2|\pi_{T_x\mu}(\hat{x})|^2\right) |\hat{x}|^{-2} \\ &= \frac{|\hat{x} - \pi_{T_x\mu}(\hat{x})|^2}{|\hat{x}|^2} \leq |\pi_{T_x\mu} - \pi_{\mathbb{R}^n}|^2 . \end{aligned}$$

A.3 Local Parametrization

A.15 Definition. For an C^1-regular n-manifold M in $\mathbb{R}^{n+k}$ and $x \in M$ set

$$SO(M, x) := \{S \in SO(n+k) : \ S(\mathbb{R}^n) = T_x M \text{ and } S(\mathbb{R}^k) = T_x M^\perp\}$$

where we identified $\mathbb{R}^n$ with $\mathbb{R}^n \times \{0\}^k$ and $\mathbb{R}^k$ with $\{0\}^n \times \mathbb{R}^k$.

A.16 Proposition. *Let $m \in \mathbb{N} \cup \{0\}$, M an embedded C^{1+m}-regular n-manifold in $\mathbb{R}^{n+k}$, and $x_0 \in M$. Let $S \in SO(n+k)$ with*

$$\left\|\pi_{\mathbb{R}^n} - \pi_{S^{-1}(T_{x_0}M)}\right\|_{op} < 1. \tag{A.12}$$

Then there exists an $r \in (0, \infty)$ such that the following holds:

1. *For all $x_1, x_2 \in M \cap B_r(x_0)$ we have that $x_1 - x_2 \in S(\mathbb{R}^k)$ implies $x_1 = x_2$*

2. *There exists a $g \in C^{1+m}\left(B_r^n(0), \mathbb{R}^k\right)$ with $g(0) = 0$ and*

$$S\left(\hat{v}, g\left(\hat{v}\right)\right) + x_0 \in M \tag{A.13}$$

$$x = S\left(p(x), g\left(p(x)\right)\right) + x_0, \tag{A.14}$$

for all $\hat{v} \in B^n_r(0)$ and all $x \in M \cap B_r(x_0)$, where $p(x) := \pi_{\mathbb{R}^n} \circ S^{-1}(x - x_0)$. In particular $M_0 :=$ graph(g) is a C^{1+m}-regular n-manifold in $\mathbb{R}^{n+k}$ with

$$S\left(M_0 \cap B_r(0)\right) + x_0 = M \cap B_r(x_0) \tag{A.15}$$

$$S(T_x M_0) = T_{x_0+S(x)} M \tag{A.16}$$

for all $x \in M_0$.

3. *g is unique in the sense that for every $\tilde{g} \in C^{1+m}\left(B^n_r(0), \mathbb{R}^k\right)$ which satisfies* (A.13) *actually $\tilde{g} = g$*

A.17 Remark. In the above setting:

1. There always exists $S_{x_0} \in SO(n+k)$ such that $S_{x_0}(\mathbb{R}^n) = T_{x_0}M$ and $S_{x_0}(\mathbb{R}^k) = T_{x_0}M^\perp$. For $S = S_{x_0}$ assumption (A.12) is always satisfied and the resulting g can be seen as a parametrization over the tangent space. In that case (A.16) implies $Dg(0) = 0$

2. In case $\left\|\pi_{\mathbb{R}^n} - \pi_{T_{x_0}M}\right\|_{op} < 1$ we can use $S = Id$ and obtain a parametrisation over $B^n_r(\hat{x}_0)$.

Proof. First we need a local parametrization around x_0. As M is embedded there exists an open subset $U \subset \mathbb{R}^{n+k}$ with $x_0 \in U$ and a diffeomorphism $\Psi \in C^{1+m}\left(B_1(0), U\right)$ with $\Psi(0) = x_0$, $\Psi(B_1(0)) = U$ and

$$M \cap U = \Psi(B^n_1(0)).$$

As $D\Psi(0)$ has rank $= n + k$ we have

$$d_1 := \inf_{v \in \partial B_1(0)} D\Psi(0) \cdot v \in (0, \infty). \tag{A.17}$$

By assumption (A.12) there exists a $d_2 \in (0, 1)$ with

$$\left\|\pi_{\mathbb{R}^n} - \pi_{S^{-1}(T_x M)}\right\|_{op} < 1 - 2d_2. \tag{A.18}$$

As Ψ is continuous we can choose $\rho \in (0, 1)$ such that

$$|D\Psi(v) - D\Psi(0)| \le d_1 d_2 \tag{A.19}$$

for all $v \in B_\rho(0)$. Set $U_1 := \Psi(B_\rho(0))$ then U_1 is open and contains x_0, so there exists an $r_1 \in (0, \infty)$ such that $B_{r_1}(x_0) \subset U_1$.

Suppose there exist $x_1, x_2 \in M \cap B_{r_1}(x_0)$ with $x_1 \neq x_2$ and $x_2 - x_1 \in S(\mathbb{R}^k)$. This will lead to a contradiction. Let $\hat{v}_1, \hat{v}_2 \in B^n_\rho(0)$ with $\Psi(\hat{v}_1) = x_1$

and $\Psi(\hat{v}_2) = x_2$, hence $\hat{v}_1 \neq \hat{v}_2$ and $\Psi(\hat{v}_2) - \Psi(\hat{v}_1) \in S(\mathbb{R}^k)$. To get $\hat{v}_1, \hat{v}_2$ we actually need that M is embedded. By the mean value theorem there exists a $t \in (0, 1)$ such that

$$\Psi(\hat{v}_2) - \Psi(\hat{v}_1) = D\Psi(\hat{v}_1 + (\hat{v}_2 - \hat{v}_1)t) \cdot (\hat{v}_2 - \hat{v}_1).$$

Set $\hat{w} := \hat{v}_1 + (\hat{v}_2 - \hat{v}_1)t$ and $\hat{a} := |\hat{v}_2 - \hat{v}_1|^{-1}(\hat{v}_2 - \hat{v}_1)$. Then $D\Psi(\hat{w}) \cdot \hat{a} \in S(\mathbb{R}^k)$ and $D\Psi(0) \cdot \hat{a} \in T_xM$. Here we used the fact that the $\left(\frac{\partial}{\partial x_i}\Psi(0)\right)_{1 \le i \le n}$ span T_xM. As $|\hat{a}| = 1$ we obtain with (A.19)

$$d_1 d_2 \geq |D\Psi(\hat{w}) - D\Psi(0)||\hat{a}| \geq |D\Psi(\hat{w}) \cdot \hat{a} - D\Psi(0) \cdot \hat{a}|$$

Now set $\hat{a}_1 = S^{-1}(D\Psi(0) \cdot \hat{a})$ and $\tilde{a}_2 = S^{-1}(D\Psi(\hat{w}) \cdot \hat{a})$. Then $\hat{a}_1 \in S^{-1}(T_xM)$, $\tilde{a}_2 \in \mathbb{R}^k$ and $|\hat{a}_1 - \tilde{a}_2| \leq d_1 d_2$. Also by (A.17) we have $d_1 \leq |\hat{a}_1|$. With (A.18) we can then conclude

$$\begin{aligned}
|\hat{a}_1| &= |\pi_{S^{-1}(T_xM)}(\hat{a}_1)| = |\pi_{S^{-1}(T_xM)}(\hat{a}_1) - \pi_{\mathbb{R}^n}(\hat{a}_1) + \pi_{\mathbb{R}^n}(\hat{a}_1)| \\
&\leq |(\pi_{S^{-1}(T_xM)} - \pi_{\mathbb{R}^n})\hat{a}_1| + |\hat{a}_1 - \tilde{a}_2| \\
&\leq \left\|\pi_{S^{-1}(T_xM)} - \pi_{\mathbb{R}^n}\right\|_{op} |\hat{a}_1| + d_1 d_2 \\
&< (1 - 2d_2)|\hat{a}_1| + d_2|\hat{a}_1| = (1 - d_2)|\hat{a}_1|.
\end{aligned}$$

This yields a contradiction, as $d_2 > 0$ by assumption and $|\hat{a}_1| > 0$. Thus we established the first statement for $r \in (0, r_1)$

For the second part set $\psi = \Psi\big|_{B_1^n(0)}$ and consider the function $\Phi \in C^{1+m}\left(B_1^n(0) \times \mathbb{R}^k, \mathbb{R}^{n+k}\right)$ defined by

$$\Phi(\hat{y}, \tilde{y}) := S^{-1}(\psi(\hat{y}) - x_0) + \tilde{y}. \tag{A.20}$$

As $\psi(0) = x_0$, we see $\Phi(0) = 0$. We want to use the inverse function theorem, so we need to show $D\Psi(0)$ is invertible, hence calculate

$$\frac{\partial}{\partial x_i}\Phi(\hat{y}, \tilde{y}) = \begin{cases} S^{-1}\frac{\partial}{\partial x_i}\psi(\hat{y}) & \text{if } 1 \leq i \leq n, \\ e_i & \text{if } n + 1 \leq i \leq n + k \end{cases}$$

for all $\hat{y} \in B_1^n(0)$. Set $b_i := S^{-1}\left(\frac{\partial}{\partial x_i}\psi(0)\right)$ for $i = 1, \dots, n$. The matrix $D\Phi(0)$ is invertible if $b_1, \dots, b_n, e_{n+1}, \dots, e_{n+k}$ are linearly independent. Suppose this is not the case, then there exists an $\alpha \in \mathbb{R}^{n+k} \setminus \{0\}$ such that

$$\sum_{i=1}^{n} \alpha_i b_i = \sum_{j=1}^{k} \alpha_{n+j} e_{n+j} =: z.$$

For the vector z we know $z \in \mathbb{R}^k$ and as the e_{n+j} are linearly independent $z \neq 0$. As the $\left(\frac{\partial}{\partial x_i}\psi(0)\right)_{1\leq i\leq n}$ form a basis of T_xM, the vectors $(b_i)_{1\leq i\leq n}$ form a basis of $S^{-1}(T_xM)$, so $z \in S^{-1}(T_xM) \cap \mathbb{R}^k$. Then by Proposition A.9.2 we have $\left\|\pi_{\mathbb{R}^n} - \pi_{S^{-1}(T_xM)}\right\|_{op} = 1$ which is a contradiction, so such a z cannot exist. Hence $D\Phi(0)$ is invertible.

By the inverse function theorem there exist a $\delta \in (0,1)$ and a function $\Phi^{-1} \in C^{1+m}(U_2, B_\delta(0))$, where $U_2 := \Phi(B_\delta(0))$, such that $\Phi^{-1}(\Phi(y)) = y$ for all $y \in B_\delta(0)$ and $\Phi(\Phi^{-1}(v)) = v$ for all $v \in U_2$. Also $U_2 \subset \mathbb{R}^{n+k}$ is open and $0 \in U_2$, so there exists $r_2 \in (0,\infty)$ such that $B_{r_2}(0) \subset U_2$. Then define $G \in C^{1+m}\left(B^n_{r_2}(0), \mathbb{R}^{n+k}\right)$ and $g \in C^{1+m}\left(B^n_{r_2}(0), \mathbb{R}^k\right)$ by

$$G(\hat{y}) := S^{-1}\left(\psi\left(\pi_{\mathbb{R}^n}(\Phi^{-1}(\hat{y},0))\right) - x_0\right), \tag{A.21}$$

$$g(\hat{y}) := \pi_{\mathbb{R}^k}\left(G(\hat{y})\right). \tag{A.22}$$

As $\psi(0) = x_0$ and $\Phi(0) = 0$ we see $G(0) = 0$. We would like to have $(\hat{y}, g(\hat{y})) = G(\hat{y})$ for all $\hat{y} \in B^n_{r_2}(0)$. To get this, consider $\hat{y} \in B^n_{r_2}(0)$ and set $(\hat{v}, \tilde{v}) = \Phi^{-1}(\hat{y}, 0)$. Then calculate using the definition of Φ (A.20)

$$G(\hat{y}) = S^{-1}\left(\psi\left(\hat{v}\right) - x\right) = \Phi(\hat{v}, 0)$$

and

$$\pi_{\mathbb{R}^n}\left(G(\hat{y})\right) = \pi_{\mathbb{R}^n}\left(\Phi(\hat{v},0)\right) = \pi_{\mathbb{R}^n}\left(\Phi(\hat{v},\tilde{v})\right) = \pi_{\mathbb{R}^n}\left(\hat{y},0\right) = \hat{y}.$$

So indeed $(\hat{y}, g(\hat{y})) = G(\hat{y})$ for all $\hat{y} \in B^n_{r_2}(0)$. Definition (A.21) then yields

$$S(\hat{y}, g(\hat{y})) + x_0 = S(G(\hat{y}) + x_0 = \psi\left(\pi_{\mathbb{R}^n}(\Phi_1^{-1}(\hat{y},0))\right) \in M$$

for all $\hat{y} \in B^n_{r_2}(0)$, which verifies (A.13) for $r \in (0, r_2)$. To prove (A.14) recall $r_1 \in (0,\infty)$ from the first statement. By setting $r_3 = \min\{r_1, r_2\}$ we obtain $p(B_{r_3}(x_0)) = B^n_{r_3}(0) \subset B^n_{r_2}(0)$ so $g(p(x))$ is defined for $x \in B_{r_3}(x_0)$. For $x_1 \in M \cap B_{r_3}(x_0)$ set $x_2 := S(p(x_1), g(p(x_1))) + x_0$, we need to show that x_1 and x_2 actually are equal. By (A.13) we know $x_2 \in M$, then with definition of x_2 and p follows

$$\begin{aligned}
S^{-1}(x_1 - x_2) &= S^{-1}\left(x_1 - x_0 - S(p(x_1), g(p(x_1)))\right) \\
&= S^{-1}\left(x_1 - x_0\right) - (p(x_1), g(p(x_1))) \\
&= \left(\pi_{\mathbb{R}^n}\left(S^{-1}\left(x_1 - x_0\right)\right) - p(x_1), \pi_{\mathbb{R}^k}\left(S^{-1}\left(x_1 - x_0\right)\right) - g(p(x_1))\right) \\
&= \left(0, \pi_{\mathbb{R}^k}\left(S^{-1}\left(x_1 - x_0\right)\right) - g(p(x_1))\right) \in \{0\}^n \times \mathbb{R}^k.
\end{aligned}$$

Thus statement 1 yields $x_1 = x_2$, which verifies (A.14) for $r \in (0, \min\{r_1, r_2\}]$.

To see (A.15), note that $S(B_r(0)) + x_0) = B_r(x_0)$. Consider an $x \in S(M_0 \cap B_r(0)) + x_0$, then there exists $v \in B_r^n(0)$ with $x = S(v, g(v)) + x_0$, so by (A.13) $x \in M$. If $x \in M \cap B_r(x_0)$ then by (A.14) and $p(B_r(x_0)) = B_r^n(0)$ we have $x = S(p(x), g(p(x))) + x_0 \in S(M_0 \cap B_r(0)) + x_0$.

For (A.15) take $x \in M_0$ and consider a curve $c_1 : (-1, 1) \to M_0$ with $c_1(0) = x$. By (A.13) $c_2(t) := S(c_1(t)) + x_0$ defines a curve in M and $c_2(0) = S(x) + x_0$. We can calculate $c_2'(0) = S(c_1'(0))$, so $S(T_xM_0) \subset T_{x_0+S(x)}M$ and as both are n-dimensional subspaces we actually have equality.

For the uniqueness suppose there exists a $\tilde{g} \in C^{1+m}\left(B_r^n(0), \mathbb{R}^k\right)$ that satisfies (A.13). Consider $v \in B_r^n(0)$, then by (A.13) $x_1 = S(v, g(v)) + x_0$ and $x_2 = S(v, \tilde{g}(v)) + x_0$ are both in M. We can calculate

$$\pi_{\mathbb{R}^n}(S^{-1}(x_1 - x_2)) = v - v = 0$$

so by statement 1 we have $x_1 = x_2$, hence $S^{-1}(x_1 - x_0) = S^{-1}(x_2 - x_0)$ and so $\tilde{g}(v) = g(v)$. As this holds for all $v \in B_r^n(0)$ we actually have $\tilde{g} = g$. This also shows that g is independent of the choice of ψ above. □

A.4 Differential Inequalities

To rigorously deal with the differential equations in the Propositions 5.6, 7.5 and 10.4, we need the following propositions. First recall the upper and lower derivative, which are just bounds on the differential quotient.

A.18 Definition. For a function $f : (a, b) \to \mathbb{R}$ and a point $t_0 \in (a, b)$, the upper and lower derivative are given by

$$\overline{D}f(t_0) := \limsup_{h\to 0} \frac{f(t_0 + h) - f(t_0)}{h}$$
$$\underline{D}f(t_0) := \liminf_{h\to 0} \frac{f(t_0 + h) - f(t_0)}{h}.$$

We allow this to be $\pm\infty$, so it always exists.

For these derivatives the following version of the fundamental theorem of integration holds:

A.19 Proposition. *Let $f : (a, b) \to \mathbb{R}$ measurable*

1. *If $\overline{D}f(t) \in [-\infty, M]$ for some $M \in \mathbb{R}$ and all $t \in (a, b)$. Then for all $a < t_1 < t_2 < b$ we have*

$$f(t_2) - f(t_1) \le \int_{t_1}^{t_2} \overline{D}f(s)ds.$$

2. If $\underline{D}f(t) \in [M, +\infty]$ for some $M \in \mathbb{R}$ and all $t \in (a, b)$. Then for all $a < t_1 < t_2 < b$ we have

$$f(t_2) - f(t_1) \geq \int_{t_1}^{t_2} \underline{D}f(s)ds.$$

Proof. For the first statement suppose the assumption holds for some negative M. For $n \in \mathbb{N}$ define

$$g_n(t) := n\left(f(t + n^{-1}) - f(t)\right).$$

Then the g_n are negative measurable functions and by Fatou's Lemma we obtain

$$\limsup_{n\to\infty} \int_{t_1}^{t_2-\epsilon} g_n(s)ds \leq \int_{t_1}^{t_2-\epsilon} \left(\limsup_{n\to\infty} g_n(s)\right) ds$$

for every $\epsilon > 0$. By defintion of g_n we clearly have $\limsup_{n\to\infty} g_n(s) \leq \overline{D}f(s)$. Furthermore by negativity of the upper derivative, f is monotonically decreasing and we can estimate

$$\begin{aligned}\int_{t_1}^{t_2-\epsilon} g_n(s)ds &= n\int_{t_2-\epsilon}^{t_2-\epsilon+n^{-1}} f(s)ds - n\int_{t_1}^{t_1+n^{-1}} f(s)ds \\ &\geq f(t_2 - \epsilon + n^{-1}) - f(t_1) \geq f(t_2) - f(t_1)\end{aligned}$$

for $n > \epsilon^{-1}$. So for every $\epsilon > 0$ we have

$$f(t_2) - f(t_1) \leq \int_{t_1}^{t_2-\epsilon} \overline{D}f(s)ds.$$

Now we can use the monotone convergence theorem with $g_n := \overline{D}f\chi_{[t_1,t_2-n^{-1}]}$, where here χ is the cut-off function on the interval. The g_n converge pointwise to $\overline{D}f$ on (t_1, t_2), so the theorem then states convergence of the integral for $\epsilon \searrow 0$, which yields the result.

If M is positive we can just look at $h(t) := f(t) - Mt$. Then h has negative upper derivative and applying the proposition to h implies the result on f. For the second statement just look at $-f$ and use the first statement, then the result follows from $\overline{D}(-f) = -\underline{D}f$. □

Furthermore we will need the following chain rule:

A.20 Proposition. *For intervals $(a_1, b_1), (a_2, b_2) \subset \mathbb{R}$ and functions $g : (a_1, b_1) \to (a_2, b_2), f : (a_2, b_2) \to \mathbb{R}$ consider $h : (a_1, b_1) \to \mathbb{R}$ given by $h(t) := f(g(t))$. Let $t_0 \in (a_1, b_1)$ be such that g is continuous in t_0 and f is differentiable in $g(t_0)$*

1. *If* $Df(g(t_0)) > 0$*, we have* $\overline{D}h(t_0) = Df(g(t_0))\overline{D}g(t_0)$ *and* $\underline{D}h(t_0) = Df(g(t_0))\underline{D}g(t_0)$
2. *If* $Df(g(t_0)) < 0$*, we have* $\overline{D}h(t_0) = Df(g(t_0))\underline{D}g(t_0)$ *and* $\underline{D}h(t_0) = Df(g(t_0))\overline{D}g(t_0)$

(Note that the upper and lower derivatives are allowed to be $\pm\infty$*)*

Proof. Set $s := g(t_0)$. As f is differentiable in s we have for all $\delta \in (-\delta_0, \delta_0)$

$$f(s+\delta) = f(s) + \delta Df(s) + \delta r(\delta)$$

for some $r : (-\delta_0, \delta_0) \to \mathbb{R}$ with $\lim_{\delta\to 0} r(\delta) = 0$. Then we can calculate

$$\frac{h(t_0+\delta) - h(t_0)}{\delta} = \frac{g(t_0+\delta) - g(t_0)}{\delta} \left(Df(g(t_0)) + r(g(t_0+\delta) - g(t_0))\right).$$

If $Df(s) > 0$, we have that the term in brackets is positive for small δ. Taking the $\limsup_{\delta\to 0}$ yields

$$\overline{D}h(t_0) = \overline{D}g(t_0)\left(Df(g(t_0)) + \lim_{\delta\to 0} r(g(t_0+\delta) - g(t_0))\right)$$

and by the continuaty of g in t_0 the error term vanishes in the limit. The same calcultaion works with $\liminf_{\delta\to 0}$. To get the $Df < 0$ statemnts just use the previous on $-f$ and combine this with $\overline{D}(-h) = -\underline{D}h$. □

A.5 References

References

[All] W. K. Allard, On the first variation of a varifold, *Annals of Math.* **95**, 417-491 (1972)

[B] K. A. Brakke, *The Motion of a Surface by its Mean Curvature*, Math. Notes Princeton, NJ, Princeton University Press (1978)

[CGG] Y. G. Chen,Y. Giga and S. Goto, Uniqueness and existence of viscosity solutions of generalised mean curvature fLow *J. Differ. Geom.* **33**, 749-786 (1991)

[CM] T. H. Colding and W. P. Minicozzi: *A Course in Minimal Surfaces* American Mathematical Society 2011.

[E1] K. Ecker, Local techniques for mean curvature flow, *Proc. Centre for Mathematics and its Applications, Australian NationalUniversity*, Vol. 26, 107-119 (1991)

[E2] K. Ecker, On regularity for mean curvature flow, *Calc. Var.* **3**, 107-126 (1995)

[E3] K. Ecker, A local monotonicity formula for mean curvature flow, *Annals of Math.* **154**, 503-525 (2001)

[E4] K. Ecker, *Regularity Theory for Mean Curvature Flow*, Birkhäuser Verlag, Boston-Basel-Berlin (2004)

[E5] K. Ecker, Partial regularity at the first singular time for hypersurfaces evolving by mean curvature flow, *Math. Ann.* **356**, 217-240 (2012)

[EH1] K. Ecker and G. Huisken Mean curvature evolution by entire graphs, *Annals of Math.* **139**, 453-471 (1989)

[EH2] K. Ecker and G. Huisken, Interior estimates for hypersurfaces moving by mean curvature *Invent. Math.* **105**, 547-569 (1991)

[EG] L. Evans and R. Gariepy R, *Measure Theory and Fine Properties of Functions* CRC Press, (1992)

[ES] L. Evans and J. Spruck, Motion of level sets by mean curvature I *Journal of Differential Geometry* **33**, 635-681 (1991)

[H1] G. Huisken, Flow by mean curvature of convex surfaces into spheres', *J. Differ. Geom.* **20**, 237-266 (1984)

[H2] G. Huisken Asymptotic behaviour for singularities of the mean curvature flow, *J. Differ. Geom.* **31**, 285-299 (1990)

[I] T. Ilmanen, Elliptic regularization and partial regularity for motion by mean curvature, *Mem. Amer. Math. Soc.* **108** (1994)

[I2] Tom Ilmanen, Singularities of mean curvature flow of surfaces, preprint (1995)

[KT] K. Kasai and Y. Tonegawa, A genereal regularity theory for weak mean curvature flow, published online in *Calc. Var. Partial Differential Equations*

[LSU] O. A. Layzhenskaya, V. A. Solonnikov and N. N. Uraltseva, *Linear and Quasilinear Equations of Parabolic Type*, Transl., II.Ser., Am. Math. Soc. **23** (1968)

[S] L. M. Simon, *Lectures on Geometric Measure Theory*, Proc. of the CMA, Vol.3 (1983)

[Sch] R. Schätzle, Geometrische Maßtheorie Vorlesungsskript (2006/2007)

[T] Y. Tonegawa, A Second derivative Hölder estimate for weak mean curvature flow, *Advances in Calculus of Variations* **7** no.1, 91-138 (2014)

[W1] B. White, The topology of hypersurfaces moving by mean curvature flow, *Communications in analysis and geometry* **3**, 317-333 (1995)

[W2] B. White, Stratification of minimal surfaces, mean curvature flows and harmonic maps, *J. reine und angewandte Math.* **488** , 1-35 (1997)

[W3] B. White, The size of the singular set in mean curvature flow of mean-convex sets, *J. Amer. Math. Soc.* **13**, no.3, 665-695 (2000)

[W4] B. White, A Local Regularity Theorem for Mean Curvature Flow, *Ann. of Math.* **161**, 148-1519 (2005)

A.6 Abstract

English:
This work is about the regularity of the Brakke flow. A Brakke flow is a family of Radon measures $(\mu_t)_{t\in[t_1,t_2]}$ in $\mathbb{R}^{n+k}$, such that the inequality

$$\overline{D}_t\mu_t(\phi) \leq \mathscr{B}(\mu_t,\phi)$$

holds for all $\phi \in C_c^{0,1}(U,\mathbb{R}^+)$. Here $\mathscr{B}(\mu_t,\phi)$ denotes the Brakke variation defined by

$$\mathscr{B}(\mu_t,\phi) := -\int_U \phi|\vec{H}|^2 d\mu + \int_U D\phi\cdot\vec{H}d\mu,$$

if this expression exists and $\mathscr{B}(\mu_t,\phi) := -\infty$ else. Moreover we assume, that the μ_t are integral at almost every time.

The central Result is Brakke's local regularity theorem, which considers Brakke flows that lie in a slab $\text{spt}\mu_t \cap B_1(0) \subset \mathbb{R}^n \times B_h^k(0)$. If this slab is narrow enough, i.e. h small and the values $\mu_{t_1}(B_1(0))$, $\mu_{t_2}(B_1(0))$ are not too far from $\mathscr{L}^n(B_1(0))$, then there exists a small ball such that $\text{spt}\mu_t \cap B_\delta(0)$ is smooth and graphical for all times $t \in [t_1 + C(n,k), t_2 - C(n,k)]$.

Now consider an arbitrary Brakke flow in $U \subset \mathbb{R}^{n+k}$, then for almost every time t_0, we have that for all $\phi \in C_c^{0,1}(U,\mathbb{R}^+)$ the function $t \to \mu_t(\phi)$ is continuous in t_0 Let t_0 be such a time and suppose μ_{t_0} has density one almost everywhere. Brakke's general regularity theorem states, that at time t_0 the singular set has $\mathscr{H}^n$-measure zero, this means, for almost every point $x \in U$ there exists a space-time-neighbourhood where the Brakke flow is smooth. This result is primarily based on the fact, that for almost every point with a tangent space, we can find a small neighbourhood, where the local regularity theorem can be applied.

In the last part we consider Brakke flows, for which the starting varifold μ_{t_1} restricted to $C_1(0) \setminus \tilde{S}$ is graphical. If $\tilde{S}$ has small enough μ_{t_1}-measure and if the graphical part satisfies certain gradient- and height-bounds, then one can use the local regularity theorem to show, that $(\mu_s)_{s\in[t_1+C,t_1+2C]}$ is completely graphical, or there exists a cylinder $C_\delta(0)$ with μ_{s_0}-measure zero, where $s_0 = t_1 + 3C$.

Deutsch:
Diese Arbeit befasst sich mit der Regularität des Brakke Flusses. Bei einem Brakke Fluss handelt es sich um eine Familie von Radon maßen $(\mu_t)_{t\in[t_1,t_2]}$ im $\mathbb{R}^{n+k}$, die der Ungleichung

$$\overline{D}_t\mu_t\left(\phi\right) \leq \mathscr{B}\left(\mu_t,\phi\right)$$

für alle $\phi \in C_c^{0,1}\left(U,\mathbb{R}^+\right)$ genügen. Hierbei bezeichnet $\mathscr{B}\left(\mu_t,\phi\right)$ die Brakke Variation gegeben durch

$$\mathscr{B}\left(\mu_t,\phi\right) := -\int_U \phi|\vec{H}|^2 d\mu + \int_U D\phi\cdot\vec{H}d\mu,$$

falls dieser Ausdruck definiert ist und $\mathscr{B}\left(\mu_t,\phi\right) := -\infty$ sonst. Darüber hinaus nehmen wir an, dass μ_t für fast alle Zeiten eine integrale Varifaltigkeit ist.

Zentrales Ergebnis ist Brakkes lokales Regularitätstheorem, dabei werden Brakke Flüsse betrachtet die lokal in einer horizontalen Röhre liegen $\mathrm{spt}\mu_t \cap B_1(0) \subset \mathbb{R}^n \times B_h^k(0)$. Ist nun die Röhre schmal genug, also h klein, und sind die Werte $\mu_{t_1}(B_1(0))$, $\mu_{t_2}(B_1(0))$ nicht zu weit weg von $\mathscr{L}^n(B_1(0))$, so gibt es eine kleine Kugel in der $\mathrm{spt}\mu_t \cap B_\delta(0)$ graphisch ist für alle Zeiten $t \in [t_1 + C(n,k), t_2 - C(n,k)]$.

Hat man nun einen beliebigen Brakke Fluss in $U \subset \mathbb{R}^{n+k}$, so gilt für fast alle Zeiten t_0, dass für alle $\phi \in C_c^{0,1}\left(U,\mathbb{R}^+\right)$ die Abbildung $t \to \mu_t(\phi)$ stetig ist in t_0. Sei t_0 ein solcher Zeitpunkt und nehmen wir weiter an μ_{t_0} habe fast überall Dichte Eins. Brakkes allgemeines Regularitätstheorem besagt, dass zum Zeitpunkt t_0 die singuläre Menge $\mathscr{H}^n$-maß Null hat, das heisst, dass für fast alle Punkte $x \in U$ eine kleine Raum-Zeit-Umgebung existiert in der der Fluss glatt ist. Dieses Ergebnis beruht im wesentlichen darauf, dass sich für fast alle Punkte mit Tangentialraum eine kleine Umgebung finden lässt, in der das lokale Regulartätstheorem angewandt werden kann.

Im letzten Teil betrachten wir Brakke Flüsse, deren Anfangsvarifaltigkeit μ_{t_1} eingeschränkt auf $C_1(0) \setminus \tilde{S}$ graphisch ist. Ist dass μ_{t_1}-maß von $\tilde{S}$ klein genug und genügt der graphische Teil von μ_{t_1} bestimmten Gradienten- und Höhen-schranken, so lässt sich mit Hilfe des lokalen Regulartätstheorems zeigen, dass der Fluss $(\mu_s)_{s\in[t_1+C,t_1+2C]}$ komplett graphisch ist, oder ein Zylinder $C_\delta(0)$ existiert der μ_{s_0}-maß Null besitzt, wobei $s_0 = t_1 + 3C$.

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig und nur unter Benutzung der angegebenen Hilfsmittel angefertigt zu haben. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen und ist noch nicht veröffentlicht.

Berlin, 08. April 2014

____

Ananda Lahiri