Epsilon-regularity for the Brakke flow with boundary

Carlo Gasparetto*

Abstract

We prove that, if a Brakke flow with boundary is close enough to a stationary half-plane with density one, then it is $C^{1,\alpha}$. Our approach is based on viscosity techniques introduced by Savin in the context of elliptic equations. The same techniques can be used to give a proof of Brakke's (interior) regularity theorem which is alternative to the original one. We also prove a constancy theorem for the Brakke flow.

1 Introduction

In this paper, we state and prove a Brakke-type theorem for the mean curvature flow with boundary, that is a flow of m-dimensional surfaces in $\mathbb{R}^d$ so that at every point the normal component of the velocity is equal to the mean curvature and the boundary is fixed. A weak notion of such a flow has been recently introduced in [20] by using integral varifolds, as devised by Brakke [5]. The objects in question are called *integral Brakke flows with boundary.*

In short, given a $(m-1)$-dimensional submanifold Γ, an integral Brakke flow with boundary Γ is a collection $\{V_t\}_{t\in I}$ of m-dimensional integral varifolds with the constraint that the first variation of V_t is a measure whose singular part with respect to $||V_t||$ behaves like $\mathcal{H}^{m-1}\llcorner\Gamma$ and the varifolds satisfy an evolution equation that encodes the information on the velocity. A precise definition will be given in Section 2.

The main result of this paper is that, if a Brakke flow in a ball of radius 1 is close enough (in some appropriate topology) to a unit-density half plane (which is a stationary solution to the mean curvature flow with a prescribed straight boundary), then the Brakke flow becomes smooth up to the boundary in a smaller ball and after some fixed waiting time. Namely, we prove the following

Theorem 1.1 (ε-regularity)**.** *Let Γ be a $C^{1,\alpha}$, $(m-1)$-dimensional submanifold of B_1 and let $\{V_t\}_{t\in[-\Lambda,0]}$ be an integral Brakke flow with boundary Γ in $B_1\times[-\Lambda,0]$. Assume the following:*

1. $0\in\operatorname{supp}||V_0||$*;*
2. *the mass of V_t is close to the measure of a m-dimensional half disk, namely*
$$\|V_t\|\leq\frac{3}{4}\omega_m \text{ for every } t\in[-\Lambda,0],$$
where ω_m is the Lebesgue measure of a ball of radius 1 in $\mathbb{R}^m$.
3. *there exists a half-plane S^+ such that, for every $t\in[-\Lambda,0]$:*
$$\operatorname{supp}\|V_t\|\subset\{x\in\mathbb{R}^d:\ \operatorname{dist}(x,S^+)\leq\varepsilon\}.$$

If ε and Λ are small enough, then there exist small constants η,β and a family $\{N_t\}_{t\in(-\eta^2,0]}$ of $C^{1,\beta}$ surfaces with boundary Γ such that
$$\|V_t\|\llcorner B_\eta=\mathcal{H}^m\llcorner N_t$$
for every $t\in(-\eta^2,0]$.

*Mathematics Area, SISSA, Trieste (Italy). cgaspare@sissa.it

We briefly comment on the assumptions. The key assumptions are Item 2 and Item 3, which describe how the Brakke flow is close to being a half-plane (with a straight boundary). Item 1, on the other hand, prevents a "pathological" behavior of Brakke flows, which is the possibility of a sudden loss of mass (see, for example, [15, Section 2.3]). All the assumptions will be made more rigorous in Section 4.

As a corollary of Theorem 1.1 and of the maximum principle that we will prove in Subsection 3.2, $\{N_t\}$ is a viscosity solution to the mean curvature flow. Therefore more regularity may be proved, *a posteriori*.

A central point in our work is that, under appropriate assumptions, the support of an integral Brakke flow with boundary satisfies a maximum principle. In order to fix ideas, assume that the support of the flow is the graph of some function $u : \mathbb{R}^m \to \mathbb{R}^{d-m}$. Then it can be proved that $|u|$ is a viscosity subsolution (in a suitable sense which we will describe at a later stage) to

$$\partial_t \varphi - \mathcal{M}^+(D^2\varphi) \leq 0,$$

where $\mathcal{M}^+$ is a Pucci maximal operator. We may therefore exploit this property to adopt a technique developed by Savin in [12] in the framework of elliptic equations and later adapted by Wang in [17] to parabolic equations, which we now summarize in our case. The key step in proving Theorem 1.1 is proving the following *improvement of flatness*:

Proposition 1.2 (Improvement of flatness)**.** *Under the assumption of Theorem 1.1, there exist $\eta > 0$ and a half plane T^+ close to S^+ such that, for every $t \in (-\eta^2, 0]$,*

$$\operatorname{supp} \|V_t\| \cap B_\eta \subset \left\{ x \in \mathbb{R}^d : \operatorname{dist}(x, T^+) \leq \frac{\varepsilon}{2}\eta \right\}.$$

In summary, if the Brakke flow is "ε-flat" at scale 1, then it becomes "$\eta\varepsilon/2$-flat" at scale η, for some η small and universal; from this, proving $C^{1,\alpha}$-regularity is classical.

The proof of Proposition 1.2 is based on a contradiction and compactness argument. Assume one can find a sequence of flatter and flatter Brakke flows for which the conclusion of Proposition 1.2 does not hold. Then appropriate rescalings of the supports of such flows converge in a suitable sense to the graph of a solution to the heat equation. The desired improvement of flatness is a straightforward consequence of classical Schauder estimates. The above convergence is achieved via a Harnack-type inequality, in the spirit of [17], and a barrier argument that describes the behavior of the Brakke flow near the boundary.

Theorem 1.1 answers a question left open in [20, Remark 11.2], that is whether an integral Brakke flow with boundary that has a tangent flow which is a unit-density half-plane is smooth in a backward neighborhood. The reader should also compare our results with the regularity theorems proved in [20]. The latter are proved under the additional assumption that the flow is *standard*: namely the flow has to be smooth at every point where a tangent flow is a unit-density half-plane (see [20, Definition 11.1]). Since we only prove backward regularity, our result does not guarantee (as it should not be expected) that an integral Brakke flow with boundary satisfying the assumptions of Theorem 1.1 is actually standard.

The first ε-regularity theorem for the mean curvature flow (without boundary) was proved in [5] and then refined in [9], where the authors extended the result to mean curvature flow in general ambient manifolds. Both those proof are variational and rely on L^2 energy estimates, somehow in the spirit of [1]. We think that a variational proof of Theorem 1.1 may be performed, by adapting the arguments in [2] and in [4] to account for the presence of the boundary. As mentioned, our proof is based on an argument first developed in [12] for elliptic equations and then adapted to parabolic equations in [17]. This method was used in [11] to prove an Allard-type theorem for minimal surfaces. Although an adaptation of the same techniques to the mean curvature flow seems quite natural, to the best of the author's knowledge this paper is the first instance in which these techniques are used for the mean curvature flow.

The regularity of mean curvature flow with boundary has been briefly investigated also in [21, 18]. One should also see [7], where the author defines a Brakke flow with a free boundary condition. Another definition of Brakke flow with fixed boundary has been investigated in [14].

1.1 Structure of the paper

In Section 2, we collect some notations that will be used throughout the paper and some well known facts about rectifiable measures. We then recall the definition of Integral Brakke flow with boundary, as stated in [20].

Section 3 is dedicated to collecting some known results about Integral Brakke flows and to adapting them to the case of an integral Brakke flow with fixed boundary. In Subsection 3.3, we state and prove a constancy theorem for the Brakke flow both with and without boundary, which, to the best of the author's knowledge, has not appeared in the literature yet. We think this result is of independent interest.

The core of the paper is Section 4, where we state and prove the improvement of flatness described in Proposition 1.2, which will later yield the desired $C^{1,\beta}$ regularity. The proof of this result is described in Subsection 4.2. The aforementioned barrier argument and Harnack-type inequality, which are crucial for obtaining the desired compactness, are discussed in Section 5 and Section 6, respectively.

Finally, the proof of Theorem 1.1 is given in Section 7, where we iterate the improvement of flatness to obtain the desired regularity.

Acknowledgements

The author is extremely grateful to Guido De Philippis for his crucial help in writing this paper and to Luigi De Masi for countless discussions and comments. This material is based upon work partially supported by INDAM-GNAMPA.

2 Preliminaries, notation and definitions

Throughout the paper, we consider fixed two positive integers m and d such that $m \leq d$. All the constants taken in consideration in the present work depend, in general, on these two parameters, although we will mostly avoid stating such dependency.

For the present section, we introduce two generic positive integers $k \leq n$ to define some objects in full generality.

2.1 Space-time

By $\mathbb{R}^{n,1}$ we denote the space $\{(x,t)\colon x \in \mathbb{R}^n \text{ and } t \in \mathbb{R}\}$. We use upper-case letters to denote points in $\mathbb{R}^{n,1}$, for example $X = (x,t)$.

For any couple $X = (x,t)$ and $Y = (y,s)$ of points in $\mathbb{R}^{n,1}$, we let

$$\rho(X,Y) = |x-y| + |t-s|^{1/2};$$

ρ is a metric on $\mathbb{R}^{n,1}$ (see, for example [10, Exercise 8.5.1]) and the topology that ρ induces on $\mathbb{R}^{n,1}$ coincides with the euclidean topology of $\mathbb{R}^{n+1}$. In particular, if $d_H(E,F)$ is the Hausdorff distance between E and F with respect to ρ and K is a compact subset of $\mathbb{R}^{n,1}$, then the space of non-empty closed subsets of K is a compact metric space, when endowed with the metric d_H.

If $x \in \mathbb{R}^n$ and $r > 0$, we set $B^n_r(x) = \{y \in \mathbb{R}^n \colon |y - x| < r\}$. When the dimension of the space is clear, we omit its indication and simply write $B_r(x)$. We also omit the indication of the center of the ball, whenever it coincides with 0, so that $B_r = B_r(0)$. If $(x,t) \in \mathbb{R}^{n,1}$, we define the parabolic cylinder

$$Q^n_r(x,t) = B^n_r(x) \times (t - r^2, t],$$

where the apex n indicates the dimension of the space component; as above, its indication will be omitted when no confusion shall arise. Lastly, we let $Q_r = Q_r(0,0)$.

$\partial_p(U \times (a,b))$ denotes the *parabolic boundary* of the cylinder $U \times (a,b)$, where $U \subset \mathbb{R}^n$:

$$\partial_p(U \times (a,b)) := \big(\overline{U} \times \{a\}\big) \cup \big(\partial U \times (a,b)\big).$$

We define the measures $\mathcal{L}^{n,1}$ and $\mathcal{H}^{s,1}$ (for any $0 \le s \le n$) on $\mathbb{R}^{n,1}$ by

$$\mathcal{L}^{n,1}(E \times F) = \mathcal{L}^n(E) \times \mathcal{L}^1(F), \qquad \mathcal{H}^{s,1}(E \times F) = \mathcal{H}^s(E) \times \mathcal{L}^1(F)$$

for $E \subset \mathbb{R}^n$ and $F \subset \mathbb{R}$, where $\mathcal{L}^n$ is the Lebesgue measure in $\mathbb{R}^n$ and $\mathcal{H}^s$ is the s-dimensional Hausdorff measure in $\mathbb{R}^n$.

For any function $f : \mathbb{R}^{n,1} \to \mathbb{R}^k$, we denote by $\nabla f(x,t)$ the gradient of the function $f(\cdot, t)$ computed at x and by $\partial_t f(x,t)$ the derivative of $f(x, \cdot)$ computed at t, whenever they are defined.

Lastly, for a set $E \subset \mathbb{R}^n$ and $x \in \mathbb{R}^n$, we let $\chi_E(x) = 0$ if $x \notin E$ and $\chi_E(x) = 1$ if $x \in E$.

2.2 Linear functions and subspaces of the euclidean space

We let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be the canonical orthonormal basis of $\mathbb{R}^n$.

We define the Grassmannian $\mathrm{Gr}(k,n)$ as the space of (unoriented) k-dimensional linear subspaces of $\mathbb{R}^n$; we identify $S \in \mathrm{Gr}(k,n)$ with the endomorphism $S\colon \mathbb{R}^n \to \mathbb{R}^n$ representing the orthogonal projection onto S. When no confusion shall arise and an orthonormal basis $\{\zeta_1, \dots, \zeta_k\}$ of S is fixed, we identify S with $\mathbb{R}^k$ via the canonical bijection

$$\iota : S \ni x \mapsto (x \cdot \zeta_1, \dots, x \cdot \zeta_k) \in \mathbb{R}^k :$$

therefore by Sx we denote both the point $Sx \in S \subset \mathbb{R}^n$ and its image via ι. In particular, when $S = \mathrm{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $X = (x,t) \in \mathbb{R}^{n,1}$, we will often use the notation $x' = Sx = (x \cdot \mathbf{e}_1, \dots, x \cdot \mathbf{e}_k) \in \mathbb{R}^k$.

We also let $S : \mathbb{R}^{n,1} \to \mathbb{R}^{k,1}$ be the map $S(x,t) = (Sx, t)$ and, in the case $S = \mathrm{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$, for $X = (x,t) \in \mathbb{R}^{n,1}$ we let $X' = (x', t)$.

Lastly, if S and T are two endomorphisms of $\mathbb{R}^n$, we define the scalar product between S and T by

$$S\colon T = \sum_{i,j=1}^{n} S_{ij} T_{ij},$$

where (S_{ij}) is the representation of S as a $n \times n$ matrix such that

$$S_{ij} = (S\mathbf{e}_i) \cdot \mathbf{e}_j.$$

We also let $|S| = \sqrt{S\colon S}$.

2.3 Hölder regularity

We point out some facts and definitions on Hölder regularity for several objects. In what follows, $\kappa \in (0,1)$ is a fixed parameter.

- **Functions on $\mathbb{R}^n$.** A function $u : \mathbb{R}^n \supset U \to \mathbb{R}^k$ is said to be in $C^{1,\kappa}(U; \mathbb{R}^k)$ if $u \in C^1(U; \mathbb{R}^k)$ and

$$[\nabla u]_{C^{0,\kappa}(U)} = \sup_{\substack{x,y \in U \\ x \neq y}} \frac{|\nabla u(x) - \nabla u(y)|}{|x-y|^\kappa} < \infty.$$

  We also let

$$\|u\|_{C^{1,\kappa}(U)} = \|u\|_{L^\infty(U)} + [\nabla u]_{C^{0,\kappa}(U)}.$$

- **Functions on $\mathbb{R}^{n,1}$.** Let $\Omega \subset \mathbb{R}^{n,1}$. We say that $u : \Omega \to \mathbb{R}^k$ is in $C^{1,\kappa}(\Omega; \mathbb{R}^k)$ if it is everywhere differentiable with respect to the space variable and the quantity

$$\|u\|_{C^{1,\kappa}(\Omega)} := \sup_{\Omega} |u| + \sup_{\substack{X,Y \in \Omega \\ X \neq Y}} \frac{|\nabla u(X) - \nabla u(Y)|}{\rho(X,Y)^\kappa} + \sup_{\substack{(x,t),(x,s) \in \Omega \\ s \neq t}} \frac{|u(x,t) - u(x,s)|}{|t-s|^{(1+\kappa)/2}}$$

  is finite.

- **Submanifolds.** We say that a k-dimensional, properly embedded submanifold Γ of some open set $U \subset \mathbb{R}^n$ is $C^{1,\kappa}$ if there exists some $\kappa > 0$ such that, for every $x, y \in \Gamma$, it holds

$$[\Gamma]_{C^{1,\kappa}(U)} := \sup_{\substack{x,y\in\Gamma \\ x\neq y}} \frac{|T_x\Gamma - T_y\Gamma|}{|x-y|^\kappa} < \infty,$$

where $T.\Gamma \in \mathrm{Gr}(k,n)$ is the tangent space to Γ and $|T_x\Gamma - T_y\Gamma|$ should be intended as in Subsection 2.2.

Before stating the next definition, we introduce the following notation denoting a closed half-space in $\mathbb{R}^m$:

$$\mathbb{R}^m_+ = \{x' \in \mathbb{R}^m : x'_m \geq 0\}.$$

Definition 2.1 ($C^{1,\kappa}$-regular single sheet flow with boundary). *Let $U \subset \mathbb{R}^d$ be an open connected set, $I \subset \mathbb{R}$ be a non-empty interval and let Γ be a $(m-1)$-dimensional, properly embedded submanifold of U. We say that $N = \bigcup_{t\in I} N_t \times \{t\} \subset U \times I$ is a $C^{1,\kappa}$-regular single sheet flow with boundary Γ in $U \times I$ if there exist $x' \in \mathbb{R}^m_+$, $r > 0$ and a family of functions*

$$f_t : \overline{B_r(x')} \cap \mathbb{R}^m_+ \to \overline{U} \text{ for every } t \in I$$

such that:

- *for every $t \in I$, N_t is homeomorphic to $\overline{B_r(x')} \cap \mathbb{R}^m_+$ via f_t;*
- *for every $t \in I$, $f_t(\partial B_r(x') \cap \mathbb{R}^m_+) \subset \partial U$;*
- *$f_t(B_r(x') \cap \partial\mathbb{R}^m_+) = \Gamma$ for every $t \in I$;*
- *the map $(t,y) \mapsto f_t(y)$ is $C^{1,\kappa}$ in $(\overline{B_r(x')} \cap \mathbb{R}^m_+) \times I$.*

In particular, if N is a $C^{1,\kappa}$-regular single sheet flow with boundary, then N_t is a $C^{1,\kappa}$, m-dimensional submanifold of U with boundary Γ. Notice that it well may be that $B_r(x') \cap (\mathbb{R}^m_+)^c = \varnothing$. Clearly, in this case, $\Gamma = \varnothing$.

2.4 Integral Varifolds

We adopt most of the terminology from [20]. Let $U \subset \mathbb{R}^d$ be an open set and let $\mathcal{M}(U)$ be the set of non-negative Radon measures on U; if φ is continuous and compactly supported on U, we let $M(\varphi) = \int \varphi(x)\, dM(x)$. Let $\mathcal{M}_m(U)$ be the set of m-dimensional rectifiable non-negative Radon measures on U. Namely, $M \in \mathcal{M}_m(U)$ if and only if there exist a m-dimensional rectifiable set E and a non-negative function $\theta \in L^1_{loc}(\mathcal{H}^m \llcorner E)$ such that

$$M(\varphi) = \int_E \theta(x)\varphi(x)\, d\mathcal{H}^m(x) \qquad \text{for all } \varphi \in C_c(U).$$

We also let $\mathcal{IM}_m(U)$ be the set of those $M \in \mathcal{M}_m(U)$ such that their density $\theta(x)$ is a non-negative integer at M-a.e. x. If $M \in \mathcal{M}_m(U)$, then for M-a.e. x the approximate tangent space $T_xM \in \mathrm{Gr}(m,d)$ is well defined (see, for instance, [13, Chapter 3]). A m-dimensional varifold on U is a Radon measure on $U \times \mathrm{Gr}(m,d)$ (see [13, Chapter 8]). In particular, to each $M \in \mathcal{M}_m(U)$ we may associate a m-dimensional varifold $\mathrm{Var}(M)$ by

$$\mathrm{Var}(M)(\varphi) = \int \varphi(x, T_xM)\, dM(x) \qquad \text{for all } \varphi \in C_c(U \times \mathrm{Gr}(m,d)).$$

Such an object is called a rectifiable varifold (see [13, Chapter 4]); $\mathrm{Var}(M)$ is said to be integral if and only if $M \in \mathcal{IM}_m(U)$. If $M \in \mathcal{M}_m(U)$, we say that $\mathrm{Var}(M)$ has bounded first variation if there exists $C > 0$ such that, for every smooth vector field $F : U \to \mathbb{R}^d$ with compact support in U, it holds

$$\int T_xM : \nabla F(x)\, dM(x) \leq C \, \|\nabla F\|_\infty .$$

If $\mathrm{Var}(M)$ has bounded first variation, then there exists a M-locally integrable vector field H_M, a Radon measure β_M that is singular with respect to M and a β_M-locally integrable unit vector field ζ_M such that, for every $F \in C^1_c(U;\mathbb{R}^d)$, it holds

$$\int T_x M : \nabla F(x)\, dM(x) = -\int H_M \cdot F\, dM + \int F \cdot \zeta_M\, d\beta_M. \tag{2.1}$$

In the following, we will often denote

$$\mathrm{div}_S\, F(x) = S \colon \nabla F(x).$$

When $M \in \mathcal{M}_m(U)$, we also let

$$\mathrm{div}_M\, F(x) := \mathrm{div}_{T_x M}\, F(x) = T_x M \colon \nabla F(x),$$

whenever it is well defined.

Definition 2.2. *Let Γ be a properly embedded $(m-1)$-dimensional submanifold of $U \subset \mathbb{R}^d$. We let $\mathcal{V}_m(U,\Gamma)$ be the space of those $M \in \mathcal{IM}_m(U)$ such that $\mathrm{Var}(M)$ has bounded first variation and the following hold true:*

1. *$\beta_M(E) \leq \mathcal{H}^{m-1}(E \cap \Gamma)$ for every $E \subset U$;*
2. *$H_M(x)$ and $T_x M$ are perpendicular at M-a.e. x.*

As mentioned in the remark following [20, Definition 6], Item 2 is actually redundant, as it can be derived from [5, §5].

As in [20], for $M \in \mathcal{V}_m(U,\Gamma)$ we let

$$\nu_M(x) = \lim_{r \searrow 0} \frac{1}{\omega_{m-1} r^{m-1}} \int_{B_r(x)} \zeta_M\, d\beta_M \tag{2.2}$$

where the limit exists, and $\nu_M(x) = 0$ otherwise. Notice that the requirement $\beta_M \leq \mathcal{H}^{m-1} \llcorner \Gamma$ in Definition 2.2 yields $|\nu_M| \leq 1$ $\mathcal{H}^{m-1} \llcorner \Gamma$-a.e.. Moreover, by [2, §3.1], $\nu_M(y) \perp \Gamma$ for $\mathcal{H}^{m-1}$-a.e. $y \in \Gamma$.

In the following, whenever Γ is a $(m-1)$-dimensional submanifold of $\mathbb{R}^d$, by a small abuse of notation we denote by Γ the Hausdorff measure $\mathcal{H}^{m-1} \llcorner \Gamma$, if no confusion shall arise.

2.5 Integral Brakke flows with boundary

Let $U \subset \mathbb{R}^d$ be an open set, $I \subset \mathbb{R}$ be a non-empty interval and let Γ be a properly embedded $(m-1)$-dimensional submanifold of U.

Definition 2.3 (Integral Brakke flow)**.** *A m-dimensional integral Brakke flow with boundary Γ in $U \times I$ is a collection $\mathbf{M} = \{M_t \colon t \in I\} \subset \mathcal{M}(U)$ such that the following hold true:*

1. *for almost every t, $M_t \in \mathcal{V}_m(U,\Gamma)$;*
2. *if $I' \subset\subset I$ and $U' \subset\subset U$, then $\int_{I'} \int_{U'} (1 + |H_{M_t}|^2)\, dM_t\, dt < +\infty$,*
3. *if $[a,b] \subset I$ and u is a non-negative, compactly supported, C^1 function on $U \times I$, then*

$$\int u(\cdot,a)\, dM_a - \int u(\cdot,b)\, dM_b \geq \int_a^b \int \left(u|H_{M_t}|^2 - H_{M_t} \cdot \nabla u - \partial_t u\right) dM_t\, dt. \tag{2.3}$$

We denote by $\mathcal{BF}_m(U \times I, \Gamma)$ the set of all m-dimensional integral Brakke flows in $U \times I$ with boundary Γ.

When $\Gamma = \varnothing$, we drop its indication and simply write $\mathcal{BF}_m(U \times I)$; notice that in this case $\beta_{M_t} = 0$ for a.e. t, and the definition agrees with the one of integral Brakke flow (without boundary) given, for instance, in [15].

Given $\mathbf{M} \in \mathcal{BF}_m(U \times I, \Gamma)$, we define its *space-time mass measure* M by

$$\int \varphi(x,t)\, dM(x,t) = \int \int \varphi(x,t)\, dM_t\, dt$$

for every $\varphi \in C_c(U \times I)$. We define the space-time track of $\mathbf{M}$ to be the closed set

$$\Sigma_{\mathbf{M}} = \operatorname{Clos}\left(\bigcup_{t \in I} \operatorname{supp} M_t \times \{t\} \right)$$

and we let $\Sigma_{\mathbf{M}}(t)$ be the slice at time t of $\Sigma_{\mathbf{M}}$, namely $\Sigma_{\mathbf{M}}(t) = \{x \in \mathbb{R}^n \colon (x,t) \in \Sigma_{\mathbf{M}}\}$. It is straightforward to check that $\operatorname{supp} M \subset \Sigma_{\mathbf{M}}$. Under reasonable assumptions, the opposite inclusion holds true as well: we further discuss this point in Lemma 3.5. Whenever no confusion may arise, we write Σ and Σ_t in place of $\Sigma_{\mathbf{M}}$ and $\Sigma_{\mathbf{M}}(t)$, respectively.

Remark 2.4 (Scaling properties)**.** A Brakke flow $\mathbf{M} \in \mathcal{BF}_m(U \times I, \Gamma)$ may be translated and parabolically dilated while preserving the requirements in Definition 2.3. For $x_0 \in \mathbb{R}^n$ and $r > 0$, let $T_{x_0,r}(y) = (y - x_0)/r$. By $(T_{x_0,r})_\sharp \mu$ we denote the push-forward of $\mu \in \mathcal{M}(\mathbb{R}^n)$ through $T_{x_0,r}$. Then $\mathbf{M}' = \{M'_s\}$ given by

$$M'_s = r^{-m} (T_{x_0,r})_\sharp M_{t_0 + r^2 s}$$

is a Brakke flow in $\frac{U - x_0}{r} \times \frac{I - t_0}{r^2}$ with boundary $\frac{1}{r}(\Gamma - x_0)$. In this case, we will write

$$\mathbf{M}' = \mathcal{D}_r(\mathbf{M} - X_0)$$

where, as usual, $X_0 = (x_0, t_0)$.

3 Properties of Integral Brakke flows with boundary

We collect some known results about Integral Brakke flows, which we will use throughout the rest of the paper.

3.1 Monotonicity properties

We denote by $\Psi : \mathbb{R}^d \times (-\infty, 0) \to \mathbb{R}$ the m-dimensional backward heat kernel

$$\Psi(x,t) = \frac{1}{(4\pi(-t))^{m/2}} \exp\left(- \frac{|x|^2}{4(-t)} \right).$$

We also pick a smooth cut-off function $\phi \in C_c^\infty([0,2))$ such that $\phi \equiv 1$ in $[0,1]$, $|\phi'| \leq 2$ and $0 \leq \phi \leq 1$ everywhere, which from now on we consider fixed. ϕ being chosen, for $R > 0$ we set

$$\Psi_R(x,t) = \Psi(x,t) \phi\left(\frac{|x|}{R} \right).$$

Proposition 3.1 (Huisken monotonicity formula)**.** *There exists a universal constant $C > 0$ such that, if $\mathbf{M} \in \mathcal{BF}_m(U \times (-T, 0), \Gamma)$ and $B_{2R} \subset U$, then for every $-T < s \leq t < 0$ it holds*

$$\begin{aligned} \int \Psi_R(x,t)\, dM_t - \int \Psi_R(x,s)\, dM_s \\ \leq \int_s^t \int \nu_{M_\tau} \cdot \nabla \Psi_R(\cdot, \tau)\, d\Gamma\, d\tau \qquad (3.1) \\ + C \frac{t-s}{R^2} \sup_{\tau \in [s,t]} \frac{M_\tau(B_{2R})}{R^m}, \qquad (3.2) \end{aligned}$$

where ν_{M_τ} is defined in (2.2).

Proof. See [20, Theorem 6.1]. $\square$

In several points of the present work, we are going to need some precise bounds on (3.1) and (3.2). While in most cases we will assume a uniform bound of the form

$$\sup_t \sup_{B_r(x)} \frac{M_t(B_r(x))}{r^m} \le E_1 < \infty$$

which takes care of (3.2), estimating (3.1) requires some more attention. What we prove in the following lemma is that, at a small enough scale, (3.1) is close to $\frac{1}{2}$ if $0 \notin \Gamma$, otherwise it is very small.

Lemma 3.2. *For every $\delta > 0$, there exist small positive constants Λ and c with the following property. Let $U \subset \mathbb{R}^d$ be open and let Γ be a $C^{1,\alpha}$ submanifold of U. Then, for every $R \le c/[\Gamma]_{C^{1,\alpha}(U)}$ and for every $(x,t) \in U \times \mathbb{R}$ such that $B_{2R}(x) \subset U$, it holds*

$$\int_{t-\Lambda R^2}^{t} \int |T_y\Gamma^{\perp} \nabla \psi_R(y-x, s-t)| \, d\Gamma(y) \, ds \le \frac{1}{2}\chi_{\Gamma^c}(x) + \delta.$$

The proof of Lemma 3.2 is somehow cumbersome and is therefore postponed to Appendix A.

Exploiting the above result, we may prove a sort of *clearing-out lemma*, in the spirit of [15, Proposition 3.6]. Namely, we prove that, provided we have some control on (3.2) and (3.1), if a point (x,t) is in the space-time track of $\mathbf{M}$, then M_s cannot be too small in a backward neighborhood of (x,t).

Before proceeding with this result, we introduce the following terminology:

Definition 3.3 (Maximal density ratio)**.** *A Brakke flow $\mathbf{M}$ (possibly with boundary) in $U \times I$ is said to have* bounded maximal density ratio *in $U' \times I'$, where $U' \subset U$ and $I' \subset I$, if*

$$\mathrm{mdr}(\mathbf{M}, U' \times I') := \sup_{B_r(x) \subset U'} \sup_{t \in I'} \frac{M_t(B_r(x))}{r^m} < \infty.$$

Proposition 3.4 (Clearing-out lemma)**.** *For every $K < \infty$ there exist positive constants c_1, c_2 with the following property. Let Γ be a $C^{1,\alpha}$ submanifold of U and let $\mathbf{M} \in \mathcal{BF}_m(U \times (a,b), \Gamma)$ be such that*

$$\mathrm{mdr}(\mathbf{M}, U \times (a,b)) \le K.$$

If $(x,t) \in \Sigma_{\mathbf{M}}$, and R is small enough depending on Γ, then

$$M_{t-c_1R^2}(B_{4R}(x)) \ge c_2 R^m.$$

Proof. The proof of the case without a boundary can be found, for example, in [15, Proposition 3.6]. For the sake of completeness, we sketch the proof along the same line in the case of an Integral Brakke flow with boundary.

Corresponding to $\delta = 1/4$, choose Λ and c as in Lemma 3.2. Let $(x,t) \in \Sigma_t$ and let $R \le c/[\Gamma]_{C^{1,\alpha}(U)}$.

We first assume that $x \in \operatorname{supp} M_t$ and that $M_t \in \mathcal{V}_m(U, \Gamma)$, so that in particular $M_t = \theta(\cdot)\mathcal{H}^m \llcorner E$ for some m-rectifiable set E. Then there exists $y \in B_R(x)$ such that

$$1 \le \theta(y) = \lim_{\tau \nearrow 0} M_t(\Psi_R(\cdot - y, \tau)). \tag{3.3}$$

Therefore, by Proposition 3.1, for any $t_1 < t$, it holds

$$M_{t_1}(\Psi_R(\cdot - y, t_1 - t)) \ge \theta(y) - CK\frac{t-t_1}{R^2} - \int_{t_1}^{t} \int \nu_{M_s} \cdot \Psi_R(\cdot - y, s-t) \, d\Gamma \, ds.$$

We now choose c_1 so small that both $CKc_1 \le \frac{1}{8}$ and $c_1 \le \Lambda$ and we set $t_1 = t - c_1R^2$. Then, using Lemma 3.2, we obtain

$$M_{t_1}(\Psi_R(\cdot - y, -c_1R^2)) \ge \theta(y) - \frac{1}{8} - \left(\frac{1}{2} + \frac{1}{4}\right) \ge \frac{1}{8},$$

where the second inequality is given by (3.3). Notice that, for every $z \in \mathbb{R}^d$, simple computations yield

$$\Psi_R(z-y,-c_1R^2) \leq CR^{-m}\chi_{B_{2R(y)}}(z) \leq CR^{-m}\chi_{B_{3R(x)}}(z)$$

for some $C>0$ universal. Hence, by integrating the above inequality in $M_{t-c_1R^2}$, we obtain

$$M_{t-c_1R^2}(B_{3R}(x)) \geq \frac{R^m}{C} M_{t-c_1R^2}(\Psi_R(\cdot - y, -c_1R^2)) \geq \frac{R^m}{8C},$$

as desired.

If $x \notin \operatorname{supp} M_t$ or $M_t \notin \mathcal{V}_m(U,\Gamma)$, then one can find a sequence of points (x_i,t_i) such that $M_{t_i} \in \mathcal{V}_m(U,\Gamma)$, $x_i \in \operatorname{supp} M_{t_i}$ and such that $(x_i,t_i) \to (x,t)$. It is then sufficient to choose R_i so that $t_i - c_1R_i^2 = t - c_1R^2$ to obtain, for i large enough,

$$M_{t-c_1R^2}(B_{4R}(x)) \geq M_{t-c_1R^2}(B_{3R}(x_i)) \geq c_2R^m.$$

□

We now state two important consequences of Proposition 3.4.

Lemma 3.5. *Let $\mathbf{M} \in \mathcal{BF}_m(U\times I,\Gamma)$ have bounded maximal density ratio in $U \times I$ and let $\Gamma \in C^{1,\alpha}(U)$. Then*

$$\Sigma_{\mathbf{M}} = \operatorname{supp} M.$$

Proof. The inclusion $\operatorname{supp} M \subset \Sigma$ is trivial. For the opposite inclusion, notice that, for a point $(x,t) \in \Sigma$ and for every $r>0$ small enough, Proposition 3.4 gives

$$M_{t-cr^2}(B_r(x)) \geq cr^m.$$

It is now sufficient to integrate this inequality in r to obtain that for every $r>0$ small enough, there is a set of the form

$$A_r = \left\{(y,s)\colon |y-x| \leq \theta\sqrt{t-s} \leq r\right\}$$

for some positive θ, c depending only on $\operatorname{mdr}(\mathbf{M})$ such that $M(A_r)>0$, hence $(x,t) \in \operatorname{supp} M$, as claimed. □

Lemma 3.6. *Let $\mathbf{M} \in \mathcal{BF}_m(U\times I,\Gamma)$ have bounded maximal density ratio and let $\Gamma \in C^{1,\alpha}(U)$. Then*

$$M \geq \mathcal{H}^{m,1} \llcorner \Sigma_{\mathbf{M}}.$$

For the proof of the above lemma, we refer the reader to Appendix B.

3.2 Maximum principle

In the present subsection, we assume that $U \subset \mathbb{R}^d$ is open and $I \subset \mathbb{R}$ is an interval of the form $(a,b]$. We also let Γ be a $(m-1)$-dimensional, $C^{1,\alpha}$ submanifold of U.

The main result of the present section is the following maximum principle.

Proposition 3.7 (Maximum principle)**.** *Let $\mathbf{M} \in \mathcal{BF}_m(U\times I,\Gamma)$ have bounded maximal density ratio.*

If there exist $u \in C^2(U\times I)$ and a point $(x_0,t_0) \in \Sigma \setminus \partial_p(U\times I)$ with $x_0 \notin \Gamma$ such that $u|_{\Sigma\cap\{t\leq t_0\}}$ has a local maximum at (x_0,t_0) and $\nabla u(x_0,t_0) \neq 0$, then

$$\partial_t u(x_0,t_0) - \inf_{\substack{T\in \mathrm{Gr}(m,d)\\ T\perp \nabla u(x_0,t_0)}} T\colon D^2u(x_0,t_0) \geq 0.$$

Proof. This proposition is a corollary of the results in [8, Section 13], see also [3]; for the reader's convenience, we give a self-contained proof, in the spirit of, for example, [19].

We may assume, without loss of generality, that $(x_0,t_0)=(0,0)$ and that $u|_{\Sigma\cap\{t\leq 0\}}$ has a strict local maximum at $(0,0)$ (if not, replace u by $u-|x|^4-|t|^2$).

Step 1. We first prove that

$$\partial_t u(0,0) - \operatorname{trace}_m D^2 u(0,0) \geq 0,$$

where $\operatorname{trace}_m D^2 u$ is the sum of the m smallest eigenvalues of $D^2 u$. Assume the result does not hold. In particular, since $u \in C^2$, we can choose $\rho > 0$ and $\varepsilon > 0$ small so that:

(i) $\partial_t u - \operatorname{trace}_m D^2 u < -\varepsilon < 0$ in Q_ρ;

(ii) $B_\rho \cap \Gamma = \varnothing$;

(iii) $u > \varepsilon > 0$ in $\Sigma \cap Q_{\rho/2}$ and $u < 0$ in $\Sigma \cap \{t \leq 0\} \setminus Q_\rho$;

where the last point holds up to adding a constant to u. We now let $\varphi(x,t) = (u^+(x,t))^4$, where $u^+ = \max\{u, 0\}$, and we use φ as a test function for (2.3). Since $\varphi(\cdot, -\rho^2) = 0$ by assumption, we have

$$\begin{aligned}
0 &\leq \int \varphi(\cdot, 0)\, dM_0 \\
&= \int \varphi(\cdot, 0)\, dM_0 - \int \varphi(\cdot, -\rho^2)\, dM_{-\rho^2} \\
&\leq \int_{-\rho^2}^0 \int \Big(-|H|^2 \varphi + H \cdot \nabla \varphi + \partial_t \varphi \Big)\, dM_t\, dt
\end{aligned}$$

where the last inequality is given by (2.3) and we have set $H(\cdot, t) = H_{M_t}(\cdot)$ for a.e. t. We now use the fact that $\operatorname{supp} \varphi \subset \Gamma^c$, thus

$$\int H \cdot \nabla \varphi\, dM_t = -\int \operatorname{div}_{M_t} \nabla \varphi\, dM_t$$

for a.e. t. Since the term $|H|^2 \varphi$ is non-negative, we obtain from the above chain of inequalities:

$$0 \leq \int_{-\rho^2}^0 \int \Big(-\operatorname{div}_{M_t} \nabla \varphi + \partial_t \varphi \Big)\, dM_t\, dt.$$

Some straightforward computations show that

$$\operatorname{div}_{M_t} \nabla \varphi = 4(u^+)^3 \operatorname{div}_{M_t} \nabla u \geq 4(u^+)^3 \operatorname{trace}_m D^2 u$$

and $\partial_t \varphi = 4(u^+)^3 \partial_t u$. Therefore

$$0 \leq \int_{-\rho^2}^0 \int 4(u^+)^3 \big(\partial_t u - \operatorname{trace}_m D^2 u \big)\, dM_t\, dt \leq -4\varepsilon^4 M(Q_{\rho/2})$$

where the last inequality is given by Item i and Item iii above. In particular, it must be

$$M(Q_{\rho/2}) = 0;$$

however, by Lemma 3.5, $(0,0) \in \Sigma = \operatorname{supp} M$, thus we reach a contradiction.

Step 2. We now prove the general result. It is sufficient to show that one can find a m-dimensional subspace $\bar{T}$ of $\mathbb{R}^d$ such that

$$\partial_t u(0,0) - \bar{T} \colon D^2 u(0,0) \geq 0$$

and $\bar{T} \nabla u(0,0) = 0$. Without loss of generality, assume that $u(0,0) = 0$. Let $\psi_j(z) = z + \frac{j}{2} z^2$ and let

$$u_j(X) = \psi_j(u(X)).$$

Then, for every j, $u_j|_{\Sigma \cap \{t \leq 0\}}$ has a local maximum at $(0,0)$. Therefore, by Step 1, there is a m-dimensional subspace T_j of $\mathbb{R}^d$ such that, at $(0,0)$,

$$0 \leq \partial_t u_j - T_j \colon D^2 u_j = (\partial_t u - T_j \colon D^2 u) + j T_j \colon (\nabla u \otimes \nabla u).$$

Up to a subsequence, which we do not relabel, we have that $T_j \to \bar{T}$ for some m-dimensional subspace $\bar{T}$, and

$$\bar{T} \colon (\nabla u \otimes \nabla u) \leq \liminf_j \frac{1}{j}(\partial_t u - T_j \colon D^2 u) = 0,$$

thus $\bar{T} \perp \nabla u$. On the other hand, since $jT_j \colon (\nabla u \otimes \nabla u) \geq 0$, we have

$$\bar{T} \colon D^2 u \leq \liminf_j T_j \colon D^2 u \leq \liminf_j (\partial_t u - jT_j \colon (\nabla u \otimes \nabla u)) \leq \partial_t u,$$

as desired. □

Given a upper-semicontinuous function $u : \mathbb{R}^{m,1} \to [0,1] \cup \{-\infty\}$ and a smooth function $\varphi : \mathbb{R}^{m,1} \to \mathbb{R}$, we say that φ *touches u from above at* $(x_0', t_0) \in \mathbb{R}^{m,1}$ if there exists $r > 0$ such that

$$\begin{cases} \varphi(x', t) \geq u(x', t) & \text{for every } (x', t) \in Q_r^m(x_0', t_0), \\ \varphi(x_0', t_0) = u(x_0', t_0). \end{cases}$$

We recall the definition of Pucci's maximal operator (see, for instance, [6, Section 2.2]). For a symmetric matrix $N \in \mathbb{R}^{d \times d}$, we let

$$\mathcal{M}^+(N) := \mathcal{M}^+\Big(N, \frac{1}{2}, 2\Big) = \frac{1}{2} \sum_{\lambda_i < 0} \lambda_i + 2 \sum_{\lambda_i > 0} \lambda_i, \tag{3.4}$$

where $\lambda_i = \lambda_i(N)$ are the eigenvalues of N. The following result is a consequence of Proposition 3.7.

Corollary 3.8. *Let $\mathbf{M} \in \mathcal{BF}_m(\mathbb{R}^{d,1})$ have bounded maximal density ratio. For every $(x', t) \in \mathbb{R}^{m,1}$, let*

$$u(x', t) = \sup\{|z| \colon z \in \mathbb{R}^{d-m} \text{ and } (x', z) \in \Sigma_{\mathbf{M}}(t)\}$$

with the convention $\sup \varnothing = -\infty$ and assume that $u \leq 1$ everywhere. There is $\delta > 0$ universal such that, whenever a smooth function $\varphi : \mathbb{R}^{m,1} \to \mathbb{R}$ touches u from above at $X_0' = (x_0', t_0)$ and $\max\{|D^2\varphi(X_0')|, |\nabla\varphi(X_0')|\} \leq \delta$, it holds

$$\partial_t \varphi(X_0') - \mathcal{M}^+(D^2\varphi(X_0')) \leq 0.$$

Proof. We assume $x_0' = 0$ and $t_0 = 0$. Notice that, since Σ is closed and $u(0,0) = \varphi(0,0)$, the supremum in the definition of u is attained and, without loss of generality, we may assume that the contact point is $x_0 = \varphi(0,0)\mathbf{e}_d \in \Sigma_0$. We let $S = \operatorname{span}\{\mathbf{e}_1, \ldots, \mathbf{e}_m\}$ and $S' = \operatorname{span}\{\mathbf{e}_{m+1}, \ldots, \mathbf{e}_{d-1}\}$, so that $\mathbb{R}^d = S + S' + \operatorname{span}\{\mathbf{e}_d\}$. Consider the function

$$H(x, t) = \frac{1}{4}\left|S'x\right|^2 + x \cdot \mathbf{e}_d - \varphi(Sx, t).$$

By assumption, in a neighborhood of $(x_0, 0)$ it holds

$$|S^\perp x| \leq \varphi(Sx, t) \leq 2$$

for every $(x, t) \in \Sigma$; therefore it can be checked that $H|_{\Sigma \cap \{t \leq 0\}} \leq 0$ in the same neighborhood. Since $H(x_0, 0) = 0$, $H|_{\Sigma \cap \{t \leq 0\}}$ has a local maximum at $(x_0, 0)$. Hence, by Proposition 3.7, it holds

$$\partial_t H(x_0, 0) - \inf_{T \perp \nabla H(x_0, 0)} T \colon D^2 H(x_0, 0) \geq 0. \tag{3.5}$$

We now estimate the two summands in the above inequality. In order to do so, we first remark that

$$\nabla H(x_0, 0) = \begin{pmatrix} -\nabla\varphi(0,0) \\ 0 \\ 1 \end{pmatrix} \qquad D^2 H(x_0, 0) = \begin{pmatrix} -D^2\varphi(0,0) & 0 & 0 \\ 0 & \frac{1}{2} I_{S'} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Consider $T \in \mathrm{Gr}(m,d)$ and an orthonormal basis $\zeta_1, \dots, \zeta_m$ of T. Then

$$\begin{aligned} T \colon D^2 H &= \sum_{i=1}^m \left\langle D^2 H \zeta_i ; \zeta_i \right\rangle \\ &= \sum_{i=1}^m \left(- \left\langle D^2 \varphi(Sx,t) S\zeta_i ; S\zeta_i \right\rangle + \frac{1}{2} \left| S' \zeta_i \right|^2 \right) \\ &\geq - \sum_{i=1}^m \left\langle D^2 \varphi(Sx,t) S\zeta_i ; S\zeta_i \right\rangle . \end{aligned}$$

In particular, $S \colon D^2 H(x,t) = -\Delta\varphi(Sx,t)$ and

$$\begin{aligned} T \colon D^2 H &= S \colon D^2 H + (T - S) \colon D^2 H \\ &\geq -\Delta\varphi - |T - S||D^2\varphi|, \end{aligned}$$

Now, if $|T - S| \leq c_1$, then the above inequality yields that, for some small c_1 universal,

$$T \colon D^2 H(x_0, 0) \geq -\mathcal{M}^+(D^2\varphi(0,0)).$$

On the other hand, if $|T - S| \geq c_1$, then we may choose an orthonormal basis $\zeta_1, \dots, \zeta_m$ of T such that $|S^\perp \zeta_1| \geq c_2$ for some c_2 universal. Since we are also assuming $T \perp \nabla H(x_0, 0)$, we have

$$0 = \zeta_1 \cdot \nabla H(x_0, 0) = -S\zeta_1 \cdot \nabla\varphi(0,0) + \zeta_1 \cdot \mathbf{e}_d$$

thus, in particular, $|\zeta_1 \cdot \mathbf{e}_d| \leq |\nabla\varphi| \leq \delta$ and

$$|S'\zeta_1| \geq |S^\perp \zeta_1| - |\zeta_1 \cdot \mathbf{e}_d| \geq c_2 - \delta \geq \frac{c_2}{2},$$

provided $\delta \leq c_2/2$. Therefore

$$T \colon D^2 H(x_0, 0) \geq - \sum_{i=1}^m \left\langle D^2\varphi(0,0)\zeta_i ; \zeta_i \right\rangle + \frac{1}{2}|S'\zeta_1|^2 \geq -\Delta\varphi(0,0) + \frac{c_2^2}{8} \geq -C\delta + \frac{c_2^2}{8}$$

for some C universal, since $|D^2\varphi(0,0)| \leq \delta$ by assumption. We may choose δ smaller, if needed, so that

$$-\mathcal{M}^+(D^2\varphi(0,0)) \leq C\delta \leq -C\delta + \frac{c_2^2}{8}.$$

Therefore, whether $|T - S| \leq c_1$ or not, it holds

$$T \colon D^2 H(x_0, 0) \geq -\mathcal{M}^+(D^2\varphi(0,0)).$$

We conclude the proof by remarking that

$$\partial_t H(x_0, 0) = -\partial_t \varphi(0,0),$$

thus (3.5) gives the desired result.

$\square$

Remark 3.9. With some more accurate computations, one may show that, actually, at the contact point φ satisfies the following inequality:

$$\partial_t \varphi - \sqrt{1 + |\nabla\varphi|^2}\, \mathrm{div}\left(\frac{\nabla\varphi}{\sqrt{1 + |\nabla\varphi|^2}} \right) \leq 0.$$

However, the weaker result proved in Corollary 3.8 will be sufficient for the rest of the paper.

3.3 Constancy theorem

Theorem 3.10 (Constancy theorem)**.** *Let Γ be a (possibly empty) $(m-1)$-dimensional properly embedded submanifold of B_1 and let $\mathbf{M} \in \mathcal{BF}_m(Q_1, \Gamma)$ have bounded maximal density ratio in Q_1. Assume that there is a $C^{1,\beta}$-regular single sheet flow N with boundary Γ in Q_1 such that*

$$\Sigma_{\mathbf{M}} \subset N$$

and that $(0,0) \in \Sigma_{\mathbf{M}}$.

1. *If $\Gamma \neq \varnothing$, then*
$$M = \mathcal{H}^{m,1} \llcorner N;$$
2. *if $\Gamma = \varnothing$, then there is $\theta : (-1, 0] \to \mathbb{N} \setminus \{0\}$ such that*
$$M_t = \theta(t)\mathcal{H}^m \llcorner N_t$$
for almost every $t \in (-1, 0]$.

Proof. First of all, given N as in the statement, let us fix the ball $B_0 := B_r(x') \subset \mathbb{R}^m_+$ and the family of homeomorphisms $\{f_t\}_{t\in(1,0]}$ as in Definition 2.1.

Step 1. First, we prove that $\Sigma = N$. Assume, by contradiction, that there is $X_0 = (x_0, t_0) \in N \cap \Sigma^c$. By Proposition 3.4, there is also a point $X_1 = (x_1, t_1) \in \Sigma \setminus \Gamma$ so close to $(0, 0)$ that $t_0 < t_1 < 0$. Since $(y, t) \mapsto f_t(y)$ is $C^{1,\beta}$, there exist $r > 0$ and a continuous curve $[t_0, t_1] \ni t \mapsto p(t) \in \mathbb{R}^d$ such that

$$\begin{cases} p(t) \in N_t \text{ for every } t; \\ p(t_0) = x_0 \text{ and } p(t_1) = x_1; \\ B_r(p(t)) \subset B_1 \setminus \Gamma \text{ and } B_r(x_0) \subset \Sigma^c_{t_0}. \end{cases}$$

To prove that such a curve and radius exists, take an affine curve $q : [t_0, t_1] \to B_0$ such that $q(t_0) = f^{-1}_{t_0}(x_0)$ and $q(t_1) = f^{-1}_{t_1}(x_1)$. Then set $p(t) = f_t(q(t))$.

Since $X_1 \in \Sigma$, the continuity of $t \mapsto p(t)$ and Proposition 3.4 imply that there must be a time $\bar t$ such that $B_r(p(t)) \subset \Sigma^c_t$ for every $t \le \bar t$ and a point $\bar x \in \partial B_r(p(\bar t)) \cap \Sigma_{\bar t}$.

Let us now consider a sequence $r_j \searrow 0$ and define the dilations

$$\mathbf{M}^j = \mathcal{D}_{r_j}(\mathbf{M} - (\bar x, \bar t)).$$

Since $\mathbf{M}$ has bounded maximal density ratio, the compactness theorems in [20, Section 10] yield that, up to passing to a subsequence, $\mathbf{M}^j$ converges to a limit Brakke flow $\mathbf{M}^\infty$. Since $\Sigma_{\mathbf{M}} \subset N$, we also have $\Sigma_{\mathbf{M}^\infty} \subset T := T_{\bar x} N_{\bar t}$. We claim, that, actually, $\Sigma_{\mathbf{M}^\infty}$ is included in the half-plane $T^+ = T_{\bar x}(N_{\bar t} \setminus B_r(p(\bar t))) \subset T$. Indeed, if that were not the case, then there would be a point $(z, \tau) \in (T \setminus T^+) \times (-\infty, 0]$ and $\rho > 0$ such that $M^\infty_\tau(B_\rho(z)) > 0$. By weak convergence, this yields

$$\liminf_j \frac{M_{\bar t + r_j^2 \tau}(B_{\rho r_j}(\bar x + r_j z))}{r_j^m} > 0.$$

This is, however, impossible, since by continuity of $t \mapsto p(t)$, we have

$$B_{\rho r_j}(\bar x + r_j z) \subset B_r(p(\bar t)) \subset (\Sigma_{\bar t + r_j^2 \tau})^c$$

for j large enough. Therefore $\Sigma_{\mathbf{M}^\infty} \subset T^+$. We finally show that this violates the maximum principle. First of all, notice the fact that $\mathbf{M}$ has bounded maximal density ratio implies that $(0, 0) \in \Sigma_{\mathbf{M}^\infty}$. Without loss of generality, say $T^+ = \{x_{m+1} = \cdots = x_d = 0 \text{ and } x_m > 0\}$ and let

$$f(x, t) = A|T^\perp x|^2 - |T(x + (1 - t)\mathbf{e}_m)|$$

for some $A > 0$. Then $f|_{\Sigma_{\mathbf{M}^\infty}}$ has a local maximum at $(0, 0)$, but the conclusion of Proposition 3.7 is violated if A is chosen large enough, depending on m. This concludes Step 1.

Step 2. By Step 1, Proposition 3.7 and since N is $C^{1,\beta}$, $\{N_t\}_{t\in(-1,0]}$ is a viscosity solution to the mean curvature flow. In particular, classical Schauder estimates imply that, for every $t \in (-1, 0]$, $N_t\backslash\Gamma$ is smooth. Now, by Definition 2.3, for almost every t, there is an integer-valued function $\theta_t \in L^1_{loc}(N_t \setminus \Gamma)$ so that

$$M_t = \theta_t(\cdot)\mathcal{H}^m \llcorner N_t.$$

By testing (2.1) with vector fields $X \in C^1_c(B_1 \setminus \Gamma; \mathbb{R}^d)$ that are tangent to N_t, one deduces that for almost every t, $\theta_t(\cdot)$ is an integer-valued $W^{1,1}_{loc}$ function on $N_t \setminus \Gamma$. Since N_t is connected, $\theta_t(\cdot)$ must be constant for almost every t. Moreover, $\theta_t \neq 0$ for almost every t, otherwise Step 1 would fail.

This concludes the proof in the case $\Gamma = \varnothing$. Otherwise, if $\Gamma \neq \varnothing$, then it holds $\beta_{M_t} = \theta_t\mathcal{H}^{m-1}\llcorner\Gamma$, which yields $\theta_t = 1$ for almost every t. $\square$

4 Improvement of flatness

This section is the core of the present work. We prove that if a Brakke flow with boundary is sufficiently flat in Q_1, then its flatness can be improved at a smaller universal scale. This is going to allow us to prove the desired $C^{1,\beta}$ regularity: see Section 7.

We introduce the following notation. We fix a m-dimensional subspace of $\mathbb{R}^d$, which we denote by S, and a $(m-1)$-dimensional subspace of $\mathbb{R}^d$, which we denote by Γ_0, such that $\Gamma_0 \subset S$. Up to changing coordinates in $\mathbb{R}^d$, we shall assume for the rest of the present section that $S = \operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ and that $\Gamma_0 = \operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_{m-1}\}$. We also let $S^+ = S \cap \{x_m > 0\}$.

Given a $(m-1)$-dimensional submanifold Γ of B_1, we write $\Gamma \in \mathcal{F}_\alpha(\delta, B_R)$ if the following hold true:

- Γ is a $C^{1,\alpha}$ submanifold of B_R and $[\Gamma]_{C^{1,\alpha}(B_R)} \leq \delta R^{-\alpha}$.
- $0 \in \Gamma$ and $T_0\Gamma = \Gamma_0$.

In passing, we remark that if $\Gamma \in \mathcal{F}_\alpha(\delta, B_R)$ and $\theta > 0$, then $\theta\Gamma \in \mathcal{F}_\alpha(\delta, B_{\theta R})$.

Moreover, if $\Gamma \in \mathcal{F}_\alpha(\delta, B_R)$, and δ is smaller than some constant depending only on α, then there exists $\gamma : \Gamma_0 \cap B_R \to \Gamma_0^\perp$ such that $|\gamma(0)| = |\nabla\gamma(0)| = 0$, $||\gamma||_{C^{1,\alpha}(B_R)} \leq \delta R^{-\alpha}$ and

$$\Gamma = \{x + \gamma(x)\colon x \in \Gamma_0 \cap B_R\} \cap B_R;$$

given $\Gamma \in \mathcal{F}_\alpha(\delta, B_R)$, we will always implicitly define γ as above.

The following is the main result of the present section.

Theorem 4.1 (Improvement of flatness)**.** *For every E_0 and α, there exist constants $\Lambda, \varepsilon_0, \eta, \beta$ (small) and C (large) with the following property. Let $\varepsilon \leq \varepsilon_0$, $\Gamma \in \mathcal{F}_\alpha(\varepsilon, B_1)$ and $\mathbf{M} \in \mathcal{BF}_m(B_1 \times [-\Lambda, 0], \Gamma)$ be such that $(0, 0) \in \Sigma_{\mathbf{M}}$,*

$$\Sigma_{\mathbf{M}} \subset \{(z, \tau)\colon \operatorname{dist}(z, S^+) \leq \varepsilon\},$$

$$\sup_{t\in[-\Lambda,0]} M_t(B_1) \leq E_0$$

and

$$\int_{B_1} \Psi(\cdot, -\Lambda)\, dM_{-\Lambda} \leq \frac{3}{4}. \tag{4.1}$$

Then there exists a half plane T^+ of the form

$$T^+ = \{x + w\zeta : x \in \Gamma_0, w > 0\} \tag{4.2}$$

for some $\zeta \in \Gamma_0^\perp$ with $|\zeta - \mathbf{e}_m| \leq C\varepsilon$, such that

$$\Sigma_{\mathbf{M}} \cap Q_\eta \subset \Big\{(x, t)\colon \operatorname{dist}(x, T^+) \leq \eta^{1+\beta}\varepsilon\Big\}. \tag{4.3}$$

The proof of Theorem 4.1 is based on a contradiction and compactness argument. If one assumes the conclusion does not hold, then it is possible to find a sequence of Brakke flows which are flatter and flatter and satisfy the other assumptions of Theorem 4.1, for which, however, no half-plane of the form (4.2) can be found so that the flatness improves at any smaller scale. However, for such flows, one shows that, after an appropriate rescaling, the space-time tracks must converge in the Hausdorff distance to the graph of a solution to the heat equation. It is then sufficient to use Schauder estimates for the heat equation with Dirichlet boundary condition to recover the conclusion.

The central point of the proof is to obtain the desired compactness. This is achieved via the two following results. The first one provides a control over the oscillations near Γ of the space-time support of a Brakke flow satisfying the assumptions of Theorem 4.1.

Proposition 4.2 (Boundary behaviour)**.** *For every E_0 and α, there exist small constants c_1 and r_1 with the following property. Let $\mathbf{M}$ and Γ satisfy the assumptions of Theorem 4.1. Then*

$$\Sigma \cap Q_{r_1} \subset \left\{ (x - \gamma(x'')) \cdot \mathbf{e}_m \geq -\varepsilon^2 + c_1 \frac{|S^\perp(x - \gamma(x''))|^2}{2\varepsilon^2} \right\}.$$

Here, x'' denotes the point $(x_1, \dots, x_{m-1}, 0, \dots, 0) \in \Gamma_0$.

With the above results at hand, we may prove that, if the Brakke flow is flat enough, then assumption (4.1) gives a Holder-type modulus of continuity in parabolic cylinders whose radii are controlled from below by some power of the flatness ε.

Proposition 4.3 (Decay of oscillations)**.** *For every E_0 and α, there exist small constants ς, C_2 and r_2 with the following property. Let $\mathbf{M}$ and Γ satisfy the assumptions of Theorem 4.1 and let $(x,t), (y,s) \in \Sigma \cap Q_{r_2}$. If $\min\{x_m, y_m\} \geq 2\varepsilon$ and*

$$\rho := \rho((x', t), (y', s)) \geq C_2 \varepsilon^\varsigma,$$

then

$$|S^\perp(x - y)| \leq C_2 \varepsilon \rho^\varsigma.$$

The two above results are sufficient to prove, via an Arzelà-Ascoli-type argument, the convergence in the Hausdorff distance which we have described.

Before proceeding, it is worth spending a few words on how the constants in Theorem 4.1 will be chosen.

- We fix Λ once and for all in Proposition 4.4; it will be needed to prove that $\mathbf{M}$ has bounded maximal density ratio in a smaller parabolic cylinder, Q_{r_3}.
- Propositions 4.2, 4.3 and 4.4 hold true provided ε_0 is small enough (depending on E_0). We will therefore always assume that this is the case. The final value of ε_0 will not be determined explicitly, as Theorem 4.1 is proved by compactness.
- The constants r_1 and r_2 chosen in Propositions 4.2 and 4.3 are chosen smaller than r_3 (determined in Proposition 4.4) and they depend on E_0 and α. These two constants will give upper bounds for η. We will then give a further upper bound for η coming from the regularity properties of the heat equation.
- Lastly, the constants C and β depend only on α and on regularity properties for the heat equation.

We now briefly describe the rest of the present section. The proof of Theorem 4.1 is given in Subsection 4.2. In Subsection 4.1, we state and prove some lemmas which will be useful in the following. The proofs of Proposition 4.2 and Proposition 4.3 are postponed to Section 5 and Section 6, respectively.

4.1 Preliminaries to the proof of Theorem 4.1

Some remarks on the assumptions of Theorem 4.1 will be needed for the proofs of Proposition 4.2, Proposition 4.3 and, ultimately, of Theorem 4.1 itself. We begin by showing that (4.1) propagates in the interior of the domain.

Proposition 4.4 (Propagation of small density)**.** *For every E_0 and α, there is r_3 small with the following property. Let $\mathbf{M}$ and Γ satisfy the assumptions of Theorem 4.1. Then, for every $(x,t) \in Q_{r_3}$ and for every $\tau \in (-r_3^2, 0)$, it holds*

$$\int_{B_{r_3}(x)} \Psi(\cdot - x, \tau)\, dM_t \leq \frac{7}{8} + \frac{1}{2}\chi_{\Gamma^c}(x).$$

Proof. We fix positive constants $r_3 \leq \frac{1}{8}$, ε, Λ and δ, all of which we will determine later; we always assume that r_3 is much smaller than Λ. For simplicity of notation, in this proof we set $r = r_3$. For $(x,t) \in Q_r$ and $\tau \in (-r^2, 0)$, we let $t_0 = t - \tau$. Then, by Proposition 3.1, it holds

$$\begin{aligned}
\int_{B_r(x)} \Psi(\cdot - x, t - t_0)\, dM_t &\leq \int \Psi_{1/8}(\cdot - x, t - t_0)\, dM_t \\
&\leq \int \Psi_{1/8}(\cdot - x, -\Lambda - t_0)\, dM_{-\Lambda} \\
&\quad + \int_{-\Lambda}^{t} \int \nu_M \cdot \nabla \Psi_{1/8}(\cdot - x, s - t_0)\, d\Gamma\, d\tau + CE_0(t - \Lambda).
\end{aligned} \tag{4.4}$$

By Lemma 3.2, if ε and Λ are small enough and r is much smaller than Λ, then

$$\begin{aligned}
&\int_{-\Lambda}^{t} \int \nu_M \cdot \nabla \Psi_{1/8}(\cdot - x, s - t_0)\, d\Gamma\, d\tau \\
&\quad \leq \int_{t_0 - 2\Lambda}^{t_0} \int |T_y\Gamma^{\perp} \nabla \Psi_{1/8}(y - x, s - t_0)|\, d\Gamma(y)\, d\tau \\
&\quad \leq \frac{1}{2}\chi_{\Gamma^c}(x) + \delta.
\end{aligned}$$

We then take Λ even smaller so that $CE_0(t - \Lambda) \leq CE_0(\Lambda + r^2) \leq 2CE_0\Lambda \leq \delta$ (recall that r is much smaller than Λ).

So far, we have fixed ε and Λ depending only on E_0 and δ, and we have assumed that r is much smaller than Λ. The last step is to choose r even smaller in order to bound (4.4) from above. To this end, we let L be the Lipschitz constant of Ψ restricted to $\mathbb{R}^d \times (-\infty, -\Lambda/2]$. Since r is much smaller than Λ, then $-\Lambda - t_0 \leq -\Lambda/2$ and we can estimate, for every $y \in B_{1/4}(x)$,

$$\begin{aligned}
\Psi_{1/8}(y - x, -\Lambda - t_0) &\leq \Psi(y - x, -\Lambda - t_0) \\
&\leq \Psi(y, -\Lambda) + L(|x| + |t_0|) \\
&\leq \Psi(y, -\Lambda) + 2Lr.
\end{aligned}$$

Let now $b = b(\Lambda) > 0$ be so small that $\Psi(y, -\Lambda) \geq b$ if $|y| \leq 1/2$. In particular, assuming that $r \leq 1/4$, for every $y \in B_{1/4}(x)$, it holds

$$\Psi_{1/8}(y - x, -\Lambda - t_0) \leq \left(1 + \frac{2Lr}{b}\right)\Psi(y, -\Lambda)$$

The same bound holds, trivially, for any y such that $|y - x| \geq 1/4$. We now choose r even smaller, if needed, so that $\frac{2Lr}{b} \leq \delta$. Therefore we may bound

$$\int \Psi_{1/8}(\cdot - x, t_0 + \Lambda)\, dM_{-\Lambda} \leq (1 + \delta) \int_{B_1} \Psi(\cdot, -\Lambda)\, dM_{-\Lambda} \leq \frac{3}{4}(1 + \delta),$$

which yields the desired conclusion, up to choosing δ small universal. □

Corollary 4.5 (Bound on $\mathrm{mdr}(\mathbf{M})$)**.** *Under the assumptions of Proposition 4.4, there exist E_1 universal such that, for every $t \in [-r_3^2, 0]$ and every $B_r(x) \subset B_{r_3}$, it holds*

$$M_t(B_r(x)) \leq E_1 r^m. \tag{4.5}$$

In particular, for every $(x,t) \in \Sigma_{\mathbf{M}} \cap Q_{r_3}$ and for every $r > 0$ small enough, it holds

$$M_{t-c_1 r^2}(B_r(x)) \geq c_2 r^m \tag{4.6}$$

for some c_1, c_2 small universal.

Proof. Let x, t and r as in the statement. Then

$$M_t(B_r(x)) \leq C r^m \int_{B_r(x)} \Psi(\cdot - x, -r^2)\, dM_t \leq 2C r^m.$$

(4.6) follows from (4.5) and Proposition 3.4. □

4.2 Proof of Theorem 4.1

As stated earlier, we are going to argue by contradiction and compactness. Namely, we fix E_0 and α, we let Λ be as specified in Proposition 4.4 and we assume there exist $\varepsilon_j \searrow 0$ and two sequences $\{\Gamma^j\}$, $\{\mathbf{M}^j\}$ such that, for every j, $\mathbf{M}^j$ and Γ^j satisfy the assumptions of Theorem 4.1 with ε_0 replaced by ε_j.

In particular, we assume that $\Gamma^j \in \mathcal{F}_\alpha(\varepsilon_j, B_1)$ and

$$\Sigma_{\mathbf{M}^j} \subset \{(z,\tau) \colon \operatorname{dist}(z, S^+) \leq \varepsilon_j\}. \tag{4.7}$$

We also assume, for the sake of contradiction, that for no j (4.3) is satisfied for any choice of T^+, η and β.

In the following, we let $\gamma^j : \Gamma_0 \cap B_1 \to \Gamma_0^\perp$ be such that $\Gamma^j \cap B_1 \subset \operatorname{graph} \gamma^j$, as in the definition of $\mathcal{F}_\alpha(\varepsilon_j, B_1)$, and we let $\Sigma^j := \Sigma_{\mathbf{M}^j}$. We also fix $r_0 = \min\{r_1, r_2, r_3\}$, so that the conclusions of Propositions 4.2, 4.3, 4.4 and of Corollary 4.5 hold true in Q_{r_0}.

Lemma 4.6 (Compactness and convergence to hyperplane)**.** *There exists a subsequence (not relabeled) such that, for almost every $t \in (-r_0^2, 0]$,*

$$M_t^j \rightharpoonup \mathcal{H}^m \llcorner S^+$$

as Radon measures in B_{r_0}.

Proof. By the Arzelà-Ascoli theorem, $\gamma^j \to 0$ in C^1 up to subsequences. By Corollary 4.5, we may apply the compactness theorems proven in [20, Theorems 10.1 and 10.2] and find a further subsequence (not relabeled) and $\mathbf{M}^\infty \in \mathcal{BF}_m(Q_{r_0}, \Gamma_0)$ such that, for every $t \in (-r_0^2, 0]$,

$$M_t^j \rightharpoonup M_t^\infty.$$

In particular, the weak convergence stated above and (4.7) yield

$$M_t^\infty((S^+)^c) = 0$$

for every $t \in (-r_0^2, 0]$. Moreover, by (4.6), $(0,0) \in \Sigma_{\mathbf{M}^\infty}$. By Theorem 3.10, the proof is concluded. □

Before stating the next result, we define some objects that we will use in the rest of the subsection. First of all, let $F_\varepsilon : \mathbb{R}^d \to \mathbb{R}^d$ be the map

$$F_\varepsilon(x) = \left(Sx, \frac{1}{\varepsilon} S^\perp x \right);$$

with a small abuse of notation, we use the same notation for the map $F_\varepsilon\colon \mathbb{R}^{d,1}\to\mathbb{R}^{d,1}$ such that $F_\varepsilon(x,t)=(F_\varepsilon(x),t)$. We now define

$$\widetilde{\Sigma}^j = F_{\varepsilon_j}(\Sigma^j).$$

Notice that, by (4.7), $\widetilde{\Sigma}^j\subset\{(x,t)\colon |S^\perp x|\le 1\}$ for every j. For $j\in\mathbb{N}$ and $(x',t)\in Q_{r_0}$, we define

$$u^j(x',t)=\left\{z\in\overline{B_1^{d-m}}\colon ((x',z),t)\in\widetilde{\Sigma}^j\right\};$$

notice that such a set may well be empty or have more than one element. We also define $\widetilde{\gamma}^j=F_{\varepsilon_j}\circ\gamma^j$; it is clear that

$$\widetilde{\gamma}^j\cdot\mathbf{e}_m\to 0 \qquad \text{in } C^{1,\alpha}.$$

Furthermore, since $||\widetilde{\gamma}^j||_{C^{1,\alpha}(B_{r_0})}\le 1$, by the Arzelà-Ascoli theorem and up to passing to a subsequence (which we do not relabel) we may find $g\colon B^{m-1}_{r_0}\to\overline{B_1^{d-m}}$ such that, for every $0<\varsigma<\alpha$,

$$S^\perp\widetilde{\gamma}^j\to g \qquad \text{in } C^{1,\varsigma}$$

and $\|g\|_{C^{1,\varsigma}}\le 1$.

In order to keep the notation light, in the following we denote by $E=\overline{B^m_{r_0}}\times\overline{B_1^{d-m}}\times[-r_0^2,0]\subset\mathbb{R}^{d,1}$ and $E'=S(E)=\overline{Q^m_{r_0}}\subset\mathbb{R}^{m,1}$. We also let $E'_+=E'\cap\{x_m\ge 0\}$.

Lemma 4.7 (Uniform convergence)**.** *There exists a subsequence (not relabeled) and $u\colon E'_+\to\overline{B_1^{d-m}}$ with the following properties:*

(i) it holds

$$d_H(\widetilde{\Sigma}_j\cap E;\operatorname{graph}u)\to 0 \tag{4.8}$$

as $j\to\infty$.

(ii) For every $(x'',t)\in\overline{Q^{m-1}_{r_0}}$ it holds $u((x'',0),t)=g(x'')$.

(iii) For every $X',Y'\in E'_+$,

$$|u(X')-u(Y')|\le 2C_2\rho(X',Y')^\varsigma,$$

where C_2 and ς are as in Proposition 4.3.

In (4.8), by $\operatorname{graph}u$ we mean the set $\{(x',u(x',t),t)\colon (x',t)\in E'_+\}\subset E$.

Proof. Step 1: Hausdorff convergence. By Lemma 4.6, $\widetilde{\Sigma}^j\cap E\neq\varnothing$ eventually. Thus one may extract a subsequence (not relabeled) so that $\widetilde{\Sigma}^j\cap E$ converges in the Hausdorff distance to some closed set $\widetilde{\Sigma}\subset E$. Since, by assumption, $\widetilde{\Sigma}^j\subset\{x_m\ge-\varepsilon_j\}$, it must also be $\widetilde{\Sigma}\subset\{x_m\ge 0\}$. We define the set-valued function

$$u(x',t)=\left\{y\in\overline{B_1^{d-m}}\colon ((x',y),t)\in\widetilde{\Sigma}\right\} \tag{4.9}$$

for $(x',t)\in E'_+$.

Step 2: $u(x',t)\neq\varnothing$ for every $(x',t)\in E'_+$. Assume, by contradiction, that there exists $(x',t)\in E'_+\setminus S(\widetilde{\Sigma})$ (recall the notation $S(x,t)=(Sx,t)=(x',t)$). Then, since $S(\widetilde{\Sigma})$ is closed, there exists an open neighborhood U' of (x',t) such that $U'\subset(S(\widetilde{\Sigma}))^c$. If we let $U=S^{-1}(U')\subset\mathbb{R}^{d,1}$, then by Lemma 4.6 and Fatou's lemma

$$0<\mathcal{H}^{m,1}(U\cap(S^+\times\mathbb{R}))\le\liminf_j M^j(U),$$

thus $M^j(U)>0$ eventually. In particular, by taking smaller and smaller neighborhoods, one can pick a subsequence $j_\ell\to\infty$ and a sequence $X_\ell\in\Sigma^{j_\ell}$ so that $S(X_\ell)\to(x',t)$. By using the maps F_{ε_j} defined above, we rescale in the directions of $S^\perp$ and find that, up to subsequences, there exists $z\in\overline{B_1^{d-m}}$ such that

$$\widetilde{\Sigma}^{j_\ell}\ni F_{\varepsilon_{j_\ell}}(X_\ell)\to((x',z),t).$$

By Step 1, $((x', z), t) \in \widetilde{\Sigma}$, which contradicts the fact that $u(x', t) = \varnothing$.

Step 3: $u((x'', 0), t) = \{g(x'')\}$**.** Let $(x'', t) \in \overline{Q^{m-1}_{r_0}}$. If $y \in u((x'', 0), t)$, then by Step 1 there exists a sequence $(x_j, t_j) \in \widetilde{\Sigma}^j$ such that $x_j \to ((x'', 0), y)$ and $t_j \to t$. In particular, by Proposition 4.2, it holds

$$
\begin{aligned}
|S^\perp(x_j - \widetilde{\gamma}^j(x''_j))| &= \frac{1}{\varepsilon_j}\big|S^\perp\big(F^{-1}_{\varepsilon_j}(x_j) - \gamma^j(x''_j)\big)\big| \\
&\le C|x_j \cdot \mathbf{e}_m + \varepsilon_j + \varepsilon_j^2|^{1/2} \\
&\longrightarrow 0
\end{aligned}
$$

as $j \to \infty$. Since $S^\perp \widetilde{\gamma}^j$ converges uniformly to g and $S^\perp x_j \to y$, it must be

$$
u((x'', 0), t) = \{g(x'')\}.
$$

Step 4: $u(x', t)$ **is a singleton and Item iii holds true.** For $i = 1, 2$, let $X_i = (x_i, t_i) \in \widetilde{\Sigma}$. Let also $\rho := \rho(S(X_1), S(X_2))$ and, without loss of generality, assume $(x_2)_m \ge (x_1)_m$.

Case 1: $(x_1)_m = 0$. By Step 1 and Proposition 4.2, we have

$$
|S^\perp x_2 - g(x''_2)| \le C(x_2)_m^{1/2} \le C\rho^{1/2}.
$$

Moreover, $|S^\perp x_1 - g(x''_2)| = |g(x''_1) - g(x''_2)| \le C\rho$. Thus

$$
|S^\perp x_2 - S^\perp x_1| \le C\rho^{1/2} + C\rho \le C\rho^\varsigma.
$$

Case 2: $(x_1)_m > 0$ *and* $\rho = 0$. In this case, we prove that $S^\perp(x_1) = S^\perp(x_2)$. Fix ω much smaller than $(x_1)_m$. By Steps 1 and 2, we may pick j large enough and three points $Y_1, Y_2, W = (w, \tau) \in \widetilde{\Sigma}^j$ such that $\rho(X_i, Y_i) \le \omega$ and $\rho(S(W), S(X_i)) \ge 2\omega$. Up to choosing j larger, we may assume that $\omega \ge C\varepsilon_j^\varsigma$ and $(y_i)_m \ge (x_i)_m - \omega \ge 2\varepsilon_j$. Therefore, by Proposition 4.3, since $\rho(S(W), S(Y_i)) \ge \omega$, we estimate

$$
\begin{aligned}
&|S^\perp(x_1 - x_2)| \\
&\quad \le |S^\perp(x_1 - y_1)| + |S^\perp(x_2 - y_2)| + |S^\perp(y_1 - w)| + |S^\perp(y_2 - w)| \\
&\quad \le 2\omega + C\omega^\varsigma.
\end{aligned}
$$

Since $\omega > 0$ is arbitrary, it holds $S^\perp(x_1) = S^\perp(x_2)$. In particular, $u(x', t)$ is a singleton for every $(x', t) \in E'_+$. With a small abuse of notation, from here onwards, we will denote by $u(x', t) \in \mathbb{R}^{d-m}$ the only element of the set defined in (4.9).

Case 3: $(x_1)_m > 0$ *and* $\rho > 0$. By Steps 1 and 2, we may choose j large enough and two points Y_1, Y_2 such that the following hold true:

1. $Y_1, Y_2 \in \widetilde{\Sigma}^j$;
2. for $i = 1, 2$, $\rho(X_i, Y_i) < \rho/8$;
3. $C\varepsilon_j^\varsigma \le \rho/2$;
4. for $i = 1, 2$, $(y_i)_m \ge 2\varepsilon_j$.

Then, by Proposition 4.3, it holds

$$
\begin{aligned}
|u(SX_1) - u(SX_2)| &\le |u(SX_1) - S^\perp(y_1)| + |u(SX_2) - S^\perp(y_2)| + |S^\perp(y_1 - y_2)| \\
&\le 2\frac{\rho}{8} + C\rho^\varsigma \le 2C\rho^\varsigma,
\end{aligned}
$$

as desired.

□

The rest of the proof consists in proving that u defined in Lemma 4.7 solves the heat equation in the interior of E'_+. To this end, we recall some facts about the heat equation. First, recall that $E'_+ = \overline{Q^m_{r_0}} \cap \{x_m \geq 0\}$ and let us introduce the sets

$$\begin{gathered}\operatorname{Int}_p E'_+ = E'_+ \setminus \partial_p E'_+,\\ (E'_+)_r = \{x' \in \mathbb{R}^m \colon |x'| \leq r_0 - r \text{ and } x'_m \geq r\} \times [-r_0^2 + r^2, 0].\end{gathered}$$

Notice that $\operatorname{Int}_p E'_+ = \bigcup_{r>0} (E'_+)_r$.

Lemma 4.8 (Interior regularity for the heat equation)**.** *Let $g \in C(\partial_p E'_+)$. Then there exists $h \in C^\infty(\operatorname{Int}_p E'_+) \cap C(E'_+)$ such that*

$$\begin{cases} \partial_t h - \Delta h = 0 & \text{in } \operatorname{Int}_p E'_+ \\ h = g & \text{on } \partial_p E'_+. \end{cases}$$

Moreover, for every $r > 0$ there exists $C > 0$ such that, for every $(x', t) \in (E'_+)_r$, it holds

$$\max\{|h(x',t)|, |\nabla h(x',t)|, |D^2 h(x',t)|, |\partial_t h(x',t)|\} \leq C \, \|g\|_{L^\infty(\partial_p E'_+)} \, .$$

We now proceed with the proof of Theorem 4.1.

Lemma 4.9. *Let u be as in Lemma 4.7. Then $u \in C^\infty(\operatorname{Int}_p E'_+; \mathbb{R}^{d-m}) \cap C(E'_+; \mathbb{R}^{d-m})$ and*

$$u_t - \Delta u = 0$$

in Ω.

Proof. We take as a model the proof of [11, Lemma 2.4]. We show that u is equal to the solution $h : E'_+ \to \mathbb{R}^{d-m}$ to the boundary value problem

$$\begin{cases} \partial_t h - \Delta h = 0 & \text{in } \operatorname{Int}_p E'_+ \\ h = u & \text{on } \partial_p E'_+. \end{cases}$$

whose existence is guaranteed by Lemma 4.8. If not, there exist r, ω small and positive so that the function

$$E'_+ \ni (x', t) \mapsto |u(x', t) - h(x', t)| + \omega |x'|^2$$

achieves its maximum at $(x'_0, t_0) \in (E'_+)_{2r}$. Since $\widetilde{\Sigma}^j$ converges in the Hausdorff distance to u, for some large j we may find $X_1 = (x_1, t_1) \in \Sigma^j$ such that $(x'_1, t_1) \in (E'_+)_r$ and the restriction to Σ^j of

$$H(x,t) := \left| \frac{S^\perp x}{\varepsilon_j} - h(Sx, t) \right| + \omega |Sx|^2$$

achieves its maximum at X_1.

We claim that, if ε_j is small enough, depending on r and ω, then for every m-dimensional subspace T, it holds $T : D^2 H(X_1) > \partial_t H(X_1)$. This would contradict Proposition 3.7, thus concluding the proof. To prove the claim, we define $f(x,t) = \frac{1}{\varepsilon_j} S^\perp x - h(Sx, t)$ and, with some straightforward computations, we write

$$H(x,t) = G_1(x,t) + G_2(x,t),$$

where

$$\begin{gathered} G_1(x,t) = |f(X_1)|^2 + 2 f(X_1) \cdot (f(x,t) - f(X_1)) + \omega |Sx|^2, \\ G_2(x,t) = |f(x,t) - f(X_1)|^2. \end{gathered}$$

Notice that, by Lemma 4.8, there exists C depending on r such that

$$|D^2 G_1(X_1)| \leq C(\omega + |f(X_1)||D^2 h(SX_1)|) \leq C.$$

Then, just as in [11], it is easy to show that, if $|T - S| \le c\omega$, then

$$T : D^2G_1(X_1) > \partial_t H(X_1)$$

and $D^2G_2(X_1) \ge 0$, thus in this case $T : D^2H(X_1) > \partial_t H(X_1)$. On the other hand, if $|T - S| \ge c\omega$, then there exists a unit-vector $\nu \in T$ such that $S^\perp \nu \ge c\omega$. In particular, since $D^2G_2 = 2\nabla f(X_1)\nabla f(X_1)^T$, it holds

$$T : D^2G_2(X_1) \ge -|S\nu|^2|\nabla h(SX_1)|^2 + \frac{1}{\varepsilon^2}|S^\perp\nu|^2 \ge \frac{c\omega^2}{\varepsilon_j^2}.$$

We now conclude by remarking that $\partial_t H(X_1) = 2f(X_1)\cdot\partial_t h(SX_1) \le C$ and $T : D^2G_1(X_1) \ge -|D^2G_1(X_1)| \ge -C$, thus

$$T : D^2H(X_1) \ge -C + \frac{c\omega^2}{\varepsilon_j^2} > \partial_t H(X_1),$$

provided ε_j is chosen small enough depending on ω and C (which, in turns, is a large constant depending on r). □

Once proven that u is a solution to the heat equation, it is sufficient to apply the following classical estimate:

Lemma 4.10 (Boundary regularity for the heat equation)**.** *For every $\alpha \in (0,1)$, there exist positive constants C and β with the following property. Let $u \in C^2(\operatorname{Int}_p E'_+) \cap C(E'_+)$ be such that*

$$\partial_t u - \Delta u = 0 \qquad \textit{in } \operatorname{Int}_p E'_+.$$

Assume, moreover, that for all t, $u(\cdot,t)|_{\{x_m=0\}} = g \in C^{1,\alpha}(B_{r_0} \cap \{x_m = 0\})$, that $|g(0)| = |Dg(0)| = 0$ and that $|u| \le 1$ everywhere. Then there exists a linear operator $L : \mathbb{R}^m \to \mathbb{R}^{d-m}$ with $|L| \le C$ such that, for every $\eta \in (0, 1/4)$,

$$|u(x', t) - L(x')| \le C\eta^{1+\beta}$$

in $(B^m_\eta \cap \{x_m \ge 0\}) \times (-\eta^2, 0]$.

Proof. See [16, Theorem 2.1]. □

Remark 4.11. From the fact that $g \in C^{1,\alpha}$ and that $Dg(0) = 0$, it follows that $L(x') = 0$ if $x_m = 0$.

Conclusion of the proof of Theorem 4.1. By Lemma 4.9 and Lemma 4.10, there exists $L : \mathbb{R}^m \to \mathbb{R}^{d-m}$ linear such that $L(x') = 0$ if $x'_m = 0$, $|L| \le C$ and, for every η small, it holds

$$\widetilde{\Sigma} \cap (B^m_{2\eta} \times B^{d-m}_1 \times (-4\eta^2, 0]) \subset \{(x,t) : |S^\perp x - L(Sx)| \le C\eta^{1+\beta}\}.$$

We fix η small, to be specified later and we choose j sufficiently large so that the Hausdorff distance between $\widetilde{\Sigma}$ and $\widetilde{\Sigma}^j$ is smaller than $\eta^{1+\beta}$. We now let $T = \{x \in \mathbb{R}^d : S^\perp x = \varepsilon_j L(Sx)\}$. Then it holds

$$\Sigma^j \cap Q_\eta \subset \{|T^\perp x| \le C'\varepsilon_j \eta^{1+\beta}\}.$$

Moreover, by Proposition 4.2 and the fact that $|\gamma_m(x'')| \le \varepsilon |x''|^{1+\alpha}$, it holds

$$\Sigma^j \cap Q_\eta \subset \{x_m \ge -\varepsilon_j \eta^{1+\alpha} - \varepsilon_j^2\},$$

provided $\eta \le r_2$. We choose j large enough so that $\varepsilon_j^2 \le \eta^{1+\beta}$. Since β can be chosen smaller than α, we have

$$\Sigma^j \cap Q_\eta \subset \{x_m \ge -2\varepsilon_j \eta^{1+\beta}\} \cap \{|T^\perp x| \le C\varepsilon_j \eta^{1+\beta}\}.$$

Up to choosing j larger, the above inclusion yields

$$\Sigma^j \cap Q_\eta \subset \{\operatorname{dist}(\cdot, T^+) \le 2C\eta^{1+\beta}\}.$$

We conclude the proof by choosing $\beta' > \beta$ and η so small that $2C\eta^{1+\beta} \le \eta^{1+\beta'}$ and we recover (4.3) (with β' instead of β). This contradicts the assumption made at the beginning of the present subsection, thus concluding the proof. □

5 Boundary behavior

We now prove Proposition 4.2. The setting is the following. Let E_0 and α be given and let r_3 be the constant given in Proposition 4.4. Assume $\mathbf{M}$ and Γ satisfy the assumptions of Theorem 4.1. Then $\mathrm{mdr}(\mathbf{M}, Q_{r_3}) < \infty$, therefore Proposition 4.2 follows from the following, more general, statement:

Proposition 5.1 (Boundary behavior at scale R)**.** *There exist c and ε_1 depending only on α with the following property. Let $0 < \delta < \varepsilon \leq \varepsilon_1$, $\Gamma \in \mathcal{F}_\alpha(\delta, B_R)$ and $\mathbf{M} \in \mathcal{BF}_m(Q_R, \Gamma)$ be such that*

$$\Sigma \cap Q_R \subset \{(x,t) \colon \operatorname{dist}(x, S^+) \leq \varepsilon R\}$$

and

$$\mathrm{mdr}(\mathbf{M}, Q_R) < \infty. \tag{5.1}$$

Then

$$\Sigma \cap Q_{R/2} \subset \left\{ (x,t) \colon x_m \geq \gamma_m(x'') - R\delta^2 + cR\frac{|S^\perp(x - \gamma(x''))|^2}{2(\varepsilon R)^2} \right\}.$$

Remark 5.2. The role of (5.1) is to guarantee that the maximum principle (Proposition 3.7) holds true.

Proof. By a simple rescaling argument, it is sufficient to prove the result in the case $R = 1$. We fix c small and $\varepsilon_1 \leq c$, to be specified later. By contradiction, assume there exist $0 < \delta \leq \varepsilon \leq \varepsilon_1$, Γ and $\mathbf{M}$ as above, and a point $(\bar{x}, \bar{t}) \in \Sigma \cap Q_{1/2}$ such that

$$0 < \omega := \frac{c}{2\varepsilon^2} |S^\perp(\bar{x} - \gamma(\bar{x}''))|^2 - \delta^2 + \gamma_m(\bar{x}'') - \bar{x}_m.$$

We show that, if this is the case, then we may build a family of surfaces sliding in the direction of $\mathbf{e}_m$ that touch Σ at some point where the conclusion of Proposition 3.7 fails.

In order to do so, we first define the functions $g : \mathbb{R}^{d-m} \to \mathbb{R}$ and $h : \mathbb{R}^m \to \mathbb{R}$ as

$$g(z) = c\frac{|z - S^\perp \gamma(\bar{x}'')|^2}{2\varepsilon^2}$$
$$h(y) = P(y'') - |y'' - \bar{x}''|^2 - y_m,$$

where

$$P(y'') = \gamma_m(\bar{x}'') + \nabla\gamma_m(\bar{x}'') \cdot (x'' - \bar{x}'') - \delta^2 - C|y'' - \bar{x}''|^2,$$

and C is a constant depending only on α chosen so that

$$P(x'') \leq \gamma_m(x'') \tag{5.2}$$

(to show that such C depending only on α exists, use the fact that $\gamma \in C^{1,\alpha}(B_R)$ and Young's inequality). Then, choose a smooth function $f \colon \mathbb{R} \to \mathbb{R}$ such that $f(-1) = -4c$, $f|_{t \geq -1/4} \geq -\frac{\omega}{2}$, $f < 0$ everywhere and $f'(t) \leq 8c$ everywhere. We now set

$$H(x,t) = g(S^\perp x) + h(Sx) + f(t).$$

This way, the zero-level set of H is a surface sliding in the $\mathbf{e}_m$-direction. Notice that

$$H(\bar{x}, \bar{t}) = \omega + f(\bar{t}) > 0. \tag{5.3}$$

We now show that, if $(x,t) \in \Sigma \cap ((\Gamma \times \mathbb{R}) \cup \partial_p Q_1)$, then $H(x,t) \leq 0$.

1. If $x \in \Sigma_{-1}$, then
   - $g(S^\perp x) \leq c(\varepsilon + \delta)^2/(2\varepsilon^2) = 2c$, since $|S^\perp x| \leq \varepsilon$ and $|\gamma(x'')| \leq \delta \leq \varepsilon$;
   - by (5.2), $h(Sx) \leq \gamma_m(x'') - x_m \leq 2\varepsilon$.

The two above facts, along with the assumption $f(-1) = -4c$, yield

$$H(x, -1) \le 2c + 2\varepsilon - 4c \le 0$$

provided $\varepsilon \le c$.

2. If $x \in \partial B_1 \cap \Sigma_t$, then $|S^\perp x| \le \varepsilon$ and $x_m \ge -\varepsilon$, thus

$$|x''| \ge \sqrt{1 - \varepsilon^2 - x_m^2} \ge \frac{3}{4} - x_m,$$

provided ε is small enough. In particular, $|x'' - \bar{x}''| \ge \frac{1}{4} - x_m$. Hence:

   - since $\|\gamma\|_{C^{1,\alpha}(B_1)} \le \delta$, we have

   $$h(Sx) \le 2\delta - (C+1)|x'' - \bar{x}''|^2 - x_m \le 2\delta - (C+1)\left(\frac{1}{4} - x_m\right)^2 - x_m;$$

   - as in Item 1, $g(S^\perp x) \le 2c$;
   - $f(t) \le 0$;

Therefore

$$H(x, t) \le 2c + 2\delta - (C+1)\left(\frac{1}{4} - x_m\right)^2 - x_m \le 2c + 2\delta - \frac{C}{4(1+C)} \le 0$$

provided $C \ge 1$ and c, δ are small enough.

3. Lastly, for every $x \in \Gamma$ and $t \in (-1, 0)$, under the assumptions $\delta \le \varepsilon$ and $c \le 1$, it holds

$$\begin{aligned} g(S^\perp x) &= \frac{c}{2\varepsilon^2}|S^\perp(\gamma(x'') - \gamma(\bar{x}''))|^2 \\ &\le \frac{c}{2\varepsilon^2} \|\nabla\gamma\|_\infty^2 \, |x'' - \bar{x}''|^2 \\ &\le |x'' - \bar{x}''|^2. \end{aligned}$$

Since $f \le 0$ and $h(Sx) \le \gamma_m(x'') - |x'' - \bar{x}''|^2 - x_m$, we have

$$H(x, t) \le \gamma_m(x'') - x_m = 0.$$

Points 1-3 above and (5.3) show that there must exist $Y = (y, s) \in Q_1 \cap \Sigma$ with $y \notin \Gamma$ such that $H|_{\{t \le s\}}$ has a local maximum at (y, s).

We now show that one can choose c even smaller, if needed, so that the existence of such a point would contradict the maximum principle. Indeed, since $|S^\perp y| \le \varepsilon$, if c is small enough then $|\nabla h(Sy)|^2 / |\nabla g(S^\perp y)|^2 \ge \varepsilon$, thus

$$|S^\perp \nabla H(Y)| = |\nabla g(S^\perp y)| \le (1 - \varepsilon)|\nabla H(Y)|.$$

Therefore, if T is a m-dimensional subspace of $\mathbb{R}^d$ such that $T \perp \nabla H(Y)$, then

$$T\colon S^\perp \ge \varepsilon$$

and

$$T\colon D^2 H(Y) = T\colon \begin{pmatrix} D^2 h & 0 \\ 0 & D^2 g \end{pmatrix} \ge -|D^2 h(Sy)| + \varepsilon |D^2 g(S^\perp y)|.$$

Now, simple computations show that, up to multiplications by constants depending only on m and α, $|D^2 h(Sy)| \le 1$ and $|D^2 g(S^\perp y)| \ge \frac{c}{\varepsilon^2}$ where c is the constant we fixed at the beginning of the proof. Therefore, if ε is much smaller than c, then $T\colon D^2 H(y) \ge \frac{c}{2\varepsilon}$. However, by Proposition 3.7, it holds

$$\inf_{T \perp \nabla H(Y)} T\colon D^2 H(Y) \le \partial_t H(Y) = f'(s) \le 8c,$$

which is a contradiction. $\square$

6 Decay of oscillations: proof of Proposition 4.3

In the present section, we prove Proposition 4.3.

We begin by giving the following definition:

Definition 6.1. *Let $u : \mathbb{R}^{m,1} \to [-\infty, 1]$ be an upper-semicontinuous function. Assume that, whenever a smooth function $\varphi : \mathbb{R}^{m,1} \to \mathbb{R}$ touches u from above at some $(x'_0, t_0) \in U \times I$ (according to the terminology set in Subsection 3.2) and $|\nabla\varphi(x'_0, t_0)|, |D^2\varphi(x'_0, t_0)|$ are smaller than some fixed universal constant δ_0, then*

$$\partial_t \varphi - \mathcal{M}^+(D^2\varphi) \le 0 \tag{6.1}$$

at (x'_0, t_0) (see Subsection 3.2 for the definition of $\mathcal{M}^+$). Then u is said to be a viscosity subsolution to (6.1) *in $U \times I$.*

The reader should notice that the classical definition of viscosity solution is slightly different than ours, in that the test function φ usually has no restrictions on the magnitude of $|\nabla\varphi|$ and $|D^2\varphi|$ at the touching point.

The proof of Proposition 4.3 is achieved in three steps:

1. First of all, one sees that the support of a $\mathbf{M}$ behaves, in some sense, like the graph of a viscosity subsolution to (6.1), as in the definition above; this was proved in Corollary 3.8.
2. By exploiting the results in [17], one shows that, if a Σ has a point far enough from S, then the mass of $\mathbf{M}$ near that point cannot be too small.
3. If Σ does not have the decay of oscillations stated in Proposition 4.3, then by the previous step the mass of $\mathbf{M}$ in some parabolic cylinder must be large; this contradicts the small density assumption (4.1).

Before proceeding, we introduce some notations that we are going to use in the present subsection. Given $\theta \in (0, 1)$, we define the set

$$\mathcal{P}_1^\theta = \left\{ (x', t) \in \mathbb{R}^{m,1} \colon |x'|^2 < \frac{-t}{\theta^2} < 1 \right\}.$$

One should compare these sets with those which, in [17], are called "parabolic balls". Our definition slightly differs from theirs; notice that with our choice $\mathcal{P}_1^\theta \subset B_1^m \times (-\theta^2, 0)$.

Lemma 6.2 (Measure estimate, [17])**.** *For every $\theta > 0$ and $\mu \in (0, 1)$, there exist small constants η', r with the following property. Let $u \colon \mathbb{R}^{m,1} \to [-\infty, 1]$ be a viscosity solution to* (6.1) *in $B_1^m \times (-\theta^2, 0)$ and assume that*

$$u(Y_0) \ge 1 - \eta'$$

for some $Y_0 \in B_r^m \times (-\theta^2 r^2, 0)$. Then

$$\mathcal{L}^{m,1}(\{u \ge 1 - \mu\} \cap \mathcal{P}_1^\theta) \ge (1 - \mu)\mathcal{L}^{m,1}(\mathcal{P}_1^\theta). \tag{6.2}$$

Proof. This result corresponds, essentially, to [17, Lemma 4.3]. Apart from some trivial adjustment of constants, there are two caveats:

- The results in [17] are stated with the classical definition of viscosity solutions, where no bound on the test function at the touching point is required. However, it is easy to see that the results are valid for our definition of viscosity solution, as well.
- In our setting, we allow u to be merely upper-semicontinuous and, possibly, take infinite values, while in [17] u is required to be continuous. This minor point can be easily overcome by looking at the sup-convolution of u:

  $$u_\delta(x, t) = \sup \left\{ u(y, s) - \frac{1}{\delta}(|x - y|^2 + (t - s)^2) \right\},$$

  which conserves the property of being a viscosity subsolution to (6.1) and for which (6.2) holds true, by [17, Lemma 4.3]. Letting $\delta \searrow 0$ gives the desired conclusion.

$\square$

Before stating the next result, we fix some further notations. For any closed set $\Sigma \subset \mathbb{R}^{d,1}$ and any $\Omega \subset \mathbb{R}^{d,1}$, we let

$$\operatorname{osc}(\Sigma, \Omega) = \inf \Big\{ h > 0\colon \text{there is } y \in \mathbb{R}^d \text{ such that } \Sigma \cap \Omega \subset \{x\colon |S^\perp(x-y)| \le h\} \Big\}.$$

We also let

$$C_r = \big\{ x \in \mathbb{R}^d\colon |Sx| < r \big\}.$$

Lemma 6.3 (Harnack inequality)**.** *For every $\delta \in (0,1)$, there exist small constants $\varepsilon_2, \theta, r, \eta$ with the following property. Let $\varepsilon \le \varepsilon_2$ and $\mathbf{M} \in \mathcal{BF}_m(C_1 \times (-\theta^2, 0])$ be such that:*

$$\Sigma \subset \{|S^\perp x| \le \varepsilon\}, \tag{6.3}$$

$$\int_{C_1} \Psi(\cdot, t)\, dM_t \le 2 - \delta \qquad \textit{for all } t \in (-\theta^2, 0), \tag{6.4}$$

and

$$\operatorname{mdr}(\mathbf{M}, C_1 \times (-\theta^2, 0]) < \infty. \tag{6.5}$$

Then

$$\operatorname{osc}(\Sigma, C_r \times (-\theta^2 r^2, 0]) \le (1-\eta)\varepsilon. \tag{6.6}$$

The proof of the above result involves some technical estimates. It is therefore convenient to give an overview of the strategy. If (6.6) does not hold, then one finds two points Y_1 and Y_2 in Σ that are far enough in $S^\perp$. By applying Lemma 6.2 twice, we find that in $C_1 \times (-\theta^2, 0)$ the mass of $\mathbf{M}$ must be almost that of two m-dimensional disks. This contradicts (6.4), which encodes the fact that the mass of $\mathbf{M}$ must not exceed by too much that of a single disk.

Proof of Lemma 6.3. Let $\delta \in (0,1)$ be given. Fix θ and μ, which we will specify later, and let r and η' be chosen accordingly as in Lemma 6.2. Moreover, fix ε much smaller than μ and $\eta \le \eta'$, to be specified later. Assume, by contradiction, that there exist $\mathbf{M} \in \mathcal{BF}_m(C_1 \times (-\theta^2, 0])$ that satisfies the assumptions of the present result with the choices made above, and two points $Y_1 = (y_1, s_1), Y_2 = (y_2, s_2) \in \Sigma \cap (C_r \times (-\theta^2 r^2, 0])$ with $|S^\perp y_1 - S^\perp y_2| \ge 2(1-\eta)\varepsilon$. For every $(x', t) \in B_1^m \times (-\theta^2, 0]$ and for $i = 1, 2$, let

$$u_i(x', t) = \frac{1}{2\varepsilon} \sup \Big\{ |z - S^\perp y_i|\colon z \in S^\perp \text{ and } (x', z) \in \Sigma_t \Big\}.$$

Notice that u_1 and u_2 are upper-semicontinuous and, for every (x', t), either $u_1(x', t), u_2(x', t) \in [0,1]$ or $u_1 = u_2 = -\infty$. By Corollary 3.8 and (6.5), both u_1 and u_2 are viscosity subsolutions to (6.1). Moreover,

$$u_1(Sy_2, s_2) \ge \frac{1}{2\varepsilon} \big| S^\perp y_2 - S^\perp y_1 \big| \ge 1 - \eta \ge 1 - \eta',$$

hence, by Lemma 6.2,

$$\mathcal{L}^{m,1}(\{u_1 \ge 1 - \mu\} \cap \mathcal{P}_1^\theta) \ge (1-\mu)\mathcal{L}^{m,1}(\mathcal{P}_1^\theta).$$

With the same argument, one also obtains

$$\mathcal{L}^{m,1}(\{u_2 \ge 1 - \mu\} \cap \mathcal{P}_1^\theta) \ge (1-\mu)\mathcal{L}^{m,1}(\mathcal{P}_1^\theta). \tag{6.7}$$

We now want to estimate

$$\int_{C_1 \times (-\theta^2, 0)} \Psi\, dM.$$

We first define, for $i = 1, 2$, the sets

$$A_i = \{(x,t) \in \mathbb{R}^{d,1}\colon (Sx, t) \in \mathcal{P}_1^\theta, |S^\perp(x - y_i)| \le \varepsilon/2 \text{ and } t \le -2\varepsilon^2/\delta\}.$$

Notice that $A_1 \cap A_2 = \varnothing$ and, by (6.3), for M-a.e. $(x,t) \in A_i$, it holds

$$\begin{aligned}
\Psi(x,t) &= \exp\left(\frac{|S^\perp x|^2}{4t}\right)\Psi'(Sx,t)\\
&\geq \exp\left(-\frac{\varepsilon^2}{8\varepsilon^2/\delta}\right)\Psi'(Sx,t)\\
&\geq \left(1-\frac{\delta}{8}\right)\Psi'(Sx,t),
\end{aligned}$$

where $\Psi'(x',t) := \Psi((x',0),t)$.

Therefore we have

$$\begin{aligned}
&\int_{C_1\times(-\theta^2,0)} \Psi\, dM\\
&\quad\geq \int_{A_1} \Psi\, dM + \int_{A_2} \Psi\, dM\\
&\quad\geq \left(1-\frac{\delta}{8}\right)\left(\int_{A_1} \Psi'(Sx,t)\, dM(x,t) + \int_{A_2} \Psi'(Sx,t)\, dM(x,t)\right).
\end{aligned} \tag{6.8}$$

Moreover, by Lemma 3.6 and by the coarea formula,

$$\int_{A_i} \Psi'(Sx,t)\, dM(x,t) \geq \int_{S(A_i\cap\Sigma)} \Psi'(x',t)\, d\mathcal{L}^{m,1}(x',t). \tag{6.9}$$

We may assume that μ and η are smaller that some universal constant so that, if $z \in \mathbb{R}^{d-m}$ is such that $\frac{1}{2\varepsilon}|z - S^\perp y_2| \geq 1-\mu$, then

$$|z - S^\perp y_1| \leq \frac{\varepsilon}{2}.$$

In particular, we have

$$S(A_1\cap\Sigma) \supset \{u_2 \geq 1-\mu\} \cap \mathcal{P}_1^\theta \cap \{t \leq -2\varepsilon^2/\delta\}$$

which, together with (6.7), yields that $S(A_1\cap\Sigma)$ covers a large portion of $\mathcal{P}_1^\theta$: namely

$$\begin{aligned}
\mathcal{L}^{m,1}\Big(S(A_1\cap\Sigma)\Big) &\geq \mathcal{L}^{m,1}\Big(\mathcal{P}_1^\theta \cap \{u_2 \geq 1-\mu\} \cap \{t \leq -2\varepsilon^2/\delta\}\Big)\\
&\geq \mathcal{L}^{m,1}\Big(\mathcal{P}_1^\theta \cap \{u_2 \geq 1-\mu\}\Big) - \frac{2\varepsilon^2}{\delta}\\
&\geq (1-2\mu)\mathcal{L}^{m,1}(\mathcal{P}_1^\theta),
\end{aligned}$$

provided $\varepsilon^2 \leq c\delta\mu$ for some c small universal.

We are now ready to choose μ, depending on δ, so that the above inequality and the fact that $\Psi \in L^1(\mathcal{L}^{m,1}\llcorner\mathcal{P}_1^\theta)$ yield

$$\begin{aligned}
&\int_{S(A_1\cap\Sigma)} \Psi'\, d\mathcal{L}^{m,1}\\
&\quad\geq \int_{\mathcal{P}_1^\theta} \Psi'\, d\mathcal{L}^{m,1} - \frac{\delta\theta^2}{8}\\
&\quad= \theta^2 \int_{B_1^m} \Psi'(\cdot,-\theta^2)\, d\mathcal{L}^m - \frac{\delta\theta^2}{8}.
\end{aligned} \tag{6.10}$$

Finally, we also choose θ small such that

$$\int_{B_1^m} \Psi'(\cdot,-\theta^2)\, d\mathcal{L}^m \geq \int_{\mathbb{R}^m} \Psi'(\cdot,-\theta^2)\, d\mathcal{L}^m - \frac{\delta}{8} = 1-\frac{\delta}{8}. \tag{6.11}$$

By (6.10) and (6.11), it holds

$$\int_{S(A_1\cap\Sigma)} \Psi'\, d\mathcal{L}^{m,1} \geq \theta^2\left(1-\frac{\delta}{4}\right). \tag{6.12}$$

The same argument can be repeated for A_2, thus giving

$$\int_{S(A_2\cap\Sigma)} \Psi'\, d\mathcal{L}^{m,1} \geq \theta^2\left(1-\frac{\delta}{4}\right). \tag{6.13}$$

We conclude the proof by combining (6.8), (6.9), and (6.12), (6.13), obtaining

$$\begin{aligned}&\int_{C_1\times(-\theta^2,0)} \Psi\, dM \\ &\qquad\geq 2\theta^2\left(1-\frac{\delta}{8}\right)\left(1-\frac{\delta}{4}\right) \\ &\qquad\geq \theta^2\left(2-\frac{3}{4}\delta\right)\end{aligned}$$

which contradicts (6.4). This concludes the proof. □

A simple rescaling argument allows one to iterate Lemma 6.3 and obtain the following

Proposition 6.4. *For every $\delta\in(0,1)$ there exist C (large) and ς (small) with the following property. Let $\mathbf{M}\in\mathcal{BF}_m(C_R\times(-R^2,R^2))$ be such that*

$$\mathrm{mdr}(\mathbf{M}, C_R\times(-R^2,R^2)) < \infty \tag{6.14}$$

and assume that, for every $(x,t)\in C_{R/2}\times(-R^2/4,R^2/4))$ and every $s\in(t-R^2/4,t)$, it holds

$$\int_{C_{R/2}(x)} \Psi(\cdot - x, s-t)\, dM_s \leq 2-\delta. \tag{6.15}$$

If $\varepsilon=\mathrm{osc}(\Sigma, C_R\times(-R,R))$, then for any couple $(x,t),(y,s)\in C_{R/2}\times(-R^2/4,R^2/4)\cap\Sigma$ such that $\rho=\rho(X',Y')\geq CR^{1-\varsigma}\varepsilon^{\varsigma}$, it holds

$$|S^\perp(x-y)| \leq C\varepsilon\left(\frac{\rho}{R}\right)^{\varsigma}.$$

Proof. We prove the result for $R=1$, as the general case follows by replacing $\mathbf{M}$ with $\mathcal{D}_R\mathbf{M}$.

Let $\varepsilon_2,\theta,r,\eta$ be the constants given in Lemma 6.3 in correspondence to δ. Consider the rescaled flows $\mathbf{M}^k=\mathcal{D}_{r^k}(\mathbf{M}-X)$. By induction, the assumptions of Lemma 6.3 are in place for every integer k such that

$$\left(\frac{1-\eta}{r}\right)^k \varepsilon \leq \varepsilon_2. \tag{6.16}$$

Therefore, scaling back to the original flow, we see that for those k:

$$\mathrm{osc}\left(\Sigma_{\mathbf{M}}, C_{r^k}(x)\times(t-\theta^2r^{2k},t]\right) \leq (1-\eta)^k\varepsilon.$$

Let now $X=(x,t)$ and $Y=(y,s)$ be two points in $C_{1/2}\times(-1/4,1/4)\cap\Sigma$ and let $\rho=\rho((x',t),(y',s))$. Without loss of generality, we may assume that $t\geq s$. Furthermore, by taking $C\geq 2/\theta$, we may clearly reduce ourselves to the case $\rho\leq\theta/2$. By choosing ς small enough and C larger than the choice made above, if necessary, we infer from $\rho\geq C\varepsilon^{\varsigma}$ that there exists $k\in\mathbb{N}$ satisfying (6.16) such that $r^{k+1}\leq 2\rho/\theta\leq r^k$. Thus

$$Y\in C_{2\rho}(x)\times(t-4\rho^2,t]\subset C_{r^k}(x)\times(t-\theta^2r^{2k},t]$$

hence it must be $|S^\perp(x-y)|\leq 2(1-\eta)^k\varepsilon$. We conclude the proof by taking C larger and ς smaller, if needed, so that $2(1-\eta)^k\leq C\rho^{\varsigma}$. □

We finally prove Proposition 4.3.

Proof of Proposition 4.3. Let $r_2 = \frac{1}{2}\min\{r_1, r_3\}$, where r_1 and r_3 are given in Propositions 4.2 and 4.4, respectively. Let also $X = (x,t), Y = (y,s)$ be two points in $\Sigma \cap Q_{r_2}$. Without loss of generality, we assume that $R := x_m \geq y_m \geq 2\varepsilon$. Let $\rho = \rho((x',t),(y',s))$; finally, let ς be the constant determined in Proposition 6.4 corresponding to $\delta = \frac{1}{2}$. We shall distinguish two cases.

If $\rho \leq \frac{R}{8}$, then we may find $t' \in (-r_2^2, 0]$ such that $X, Y \in C_{R/4}(x) \times (t' - R^2/16, t' + R^2/16)$ and $C_{R/4}(x) \subset \Gamma^c$. Since $R \leq 1$, the assumption $\rho \geq C\varepsilon^\varsigma$ yields $\rho \geq CR^{1-\varsigma}\varepsilon^\varsigma$. By Proposition 4.4 and Corollary 4.5, (6.14) and (6.15) hold true, thus Proposition 6.4 applies and we obtain

$$|S^\perp(x-y)| \leq C\left(\frac{\rho}{R}\right)^\varsigma \operatorname{osc}(\Sigma, \mathcal{U}(X)),$$

where $\mathcal{U}(X)\colon = C_{R/4}(x) \times (t' - R^2/16, t' + R^2/16)$. By Proposition 4.2, we may estimate

$$\operatorname{osc}(\Sigma, \mathcal{U}(X)) \leq 2C\varepsilon\sqrt{2R + \varepsilon + \varepsilon^2} + CR\,\|\nabla\gamma\|_\infty \leq C\varepsilon R^{1/2},$$

since $\varepsilon \leq R/2$ and $\|\nabla\gamma\|_\infty \leq C\varepsilon$. Thus

$$|S^\perp(x-y)| \leq C\varepsilon R^{1/2-\varsigma}\rho^\varsigma \leq C\varepsilon\rho^\varsigma,$$

since ς can be chosen smaller than 1/2.

On the other hand, if $\rho \geq \frac{R}{8}$, then it is sufficient to use Proposition 4.2 twice and the fact that $\|\nabla\gamma\|_\infty \leq C\varepsilon$ to estimate:

$$\begin{aligned}
|S^\perp(x - \gamma(x''))| &\leq C\varepsilon(R + \varepsilon + \varepsilon^2)^{1/2} \leq C\varepsilon\rho^{1/2},\\
|S^\perp(y - \gamma(y''))| &\leq C\varepsilon(2R + \varepsilon + \varepsilon^2)^{1/2} \leq C\varepsilon\rho^{1/2},\\
|S^\perp(\gamma(y'') - \gamma(x''))| &\leq C\varepsilon\rho.
\end{aligned}$$

We therefore conclude

$$|S^\perp(x-y)| \leq C\varepsilon\rho^{1/2} \leq C\varepsilon\rho^\varsigma,$$

which is the desired result. □

7 $C^{1,\beta}$ regularity

In the present section, we prove the following ε-regularity theorem:

Theorem 7.1 ($C^{1,\beta}$ regularity). *For every E_0 and α, there are small constants $\varepsilon_0, \Lambda, \eta$ and β with the following property. Let $\Gamma \in \mathcal{F}_\alpha(\varepsilon_0, B_1)$, $\mathbf{M} \in \mathcal{BF}_m(B_1 \times [-\Lambda, 0], \Gamma)$ be such that $(0,0) \in \Sigma_{\mathbf{M}}$,*

$$\Sigma_{\mathbf{M}} \subset \{(z,\tau)\colon \operatorname{dist}(z, S^+) \leq \varepsilon_0\},$$
$$\sup_{t\in[-\Lambda,0]} M_t(B_1) \leq E_0$$

and

$$\int_{B_1} \Psi(\cdot, -\Lambda)\, dM_{-\Lambda} \leq \frac{3}{4}.$$

Then $\Sigma_{\mathbf{M}} \cap Q_\eta$ is a $C^{1,\beta}$-regular single sheet flow.

Before proving the above result, we record the following consequence of Theorem 4.1

Proposition 7.2 (Iteration of the improvement of flatness). *Under the assumptions of Theorem 4.1, for every $X = (x,t) \in \Sigma \cap Q_\eta$:*

- *If $x \in \Gamma$, then there exists a m-dimensional half plane T_X^+ such that $\partial T_X^+ = T_x\Gamma$ and*

  $$\Sigma \cap Q_{\eta^k}(X) \subset \{\operatorname{dist}(\cdot, x + T_X^+) \leq 2\varepsilon\eta^{k(1+\beta)}\}$$

  for every $k \in \mathbb{N}$;

- *If $x \notin \Gamma$, then there exists a m-dimensional plane T_X such that*

$$\Sigma \cap Q_{\eta^k}(X) \subset \{\mathrm{dist}(\cdot, x + T_X) \leq 2\varepsilon\eta^{k(1+\beta)}\}$$

  for every $k \in \mathbb{N}$.

Proof. This result is a straightforward consequence of an iteration of Theorem 4.1. Namely, given $X \in \Sigma \cap Q_\eta$ with $x \in \Gamma$, we may find a sequence of half-planes T_k^+ such that

$$\Sigma \cap Q_{\eta^k} \subset \{\mathrm{dist}(\cdot, T_k^+) \leq \eta^{k(1+\beta)}\varepsilon\}.$$

Moreover, $|T_k^+ - T_{k-1}^+| \leq C\varepsilon\frac{\eta^{k(1+\beta)}}{\eta^k}$ for some C depending only on E_0 and α. Therefore, $\{T_k^+\}$ converges to some half plane T_X^+ for which the conclusion of the proposition holds true.

For the case $x \notin \Gamma$, one may see [15] or replicate the techniques of the previous sections. □

Remark 7.3. Given $x \in \Gamma$ and $T_{(x,t)}^+$ as in Proposition 7.2, throughout the rest of the present section, we let $T_{(x,t)}$ be the m-dimensional plane obtained by reflecting $T_{(x,t)}^+$ across $T_x\Gamma$. We remark the following conclusion of Theorem 1.1: there is C depending only on E_0 and α such that, for every $X \in \Sigma \cap Q_\eta$, it holds

$$|T_X - S| \leq C\varepsilon.$$

We are now ready to prove Theorem 7.1.

Proof of Theorem 7.1. Up to a rotation, we may assume, without loss of generality, that $T_{(0,0)}$ defined in Proposition 7.2 coincides with the plane S that satisfies the assumptions of the present result.

Step 1: Hölder decay of the tangent planes. We prove that, for any two $X = (x,t), Y = (y,s) \in \Sigma \cap Q_{\eta/4}$, it holds

$$|T_X - T_Y| \leq C_0\varepsilon\rho(X,Y)^\beta$$

for some C_0 universal large enough.

Let $\rho = \rho(X,Y)$ and, without loss of generality, say $t \geq s$. By Proposition 7.2, since $Q_\rho(Y) \subset Q_{2\rho}(X)$,

$$\Sigma \cap Q_\rho(Y) \subset \Big\{(z,\tau) : \mathrm{dist}(z - x, T_X) \leq \varepsilon(2\rho)^{1+\beta}\Big\}.$$

We may apply Proposition 7.2 to the flow $\mathcal{D}_\rho(\mathbf{M} - Y)$. Then, by Remark 7.3, we have

$$|T_Y - T_X| \leq C\rho^\beta,$$

as desired.

Step 2: Σ is included in the graph of a $C^{1,\beta}$ function over S. Let $X \in \Sigma \cap Q_\eta^d$ and, for simplicity of notation, let $T = T_X$ as defined in Remark 7.3; recall that $|T_X - S| \leq C\varepsilon$. For any other point $Y \in \Sigma \cap Q_\eta^d$, we may write

$$\begin{aligned}
|S^\perp(x-y)| &\leq |T^\perp(x-y)| + |S^\perp - T^\perp||x-y| \\
&\leq C\varepsilon\rho(X,Y)^{1+\beta} + C\varepsilon\rho(X,Y) \\
&\leq 2C\varepsilon\rho(X,Y).
\end{aligned}$$

If ε is smaller than some universal constant, we conclude

$$|S^\perp(x-y)| \leq 3C\varepsilon\rho(SX, SY).$$

In passing, we record that the above inequality, together with Proposition 7.2, yields

$$|T_X^\perp(x-y)| \leq C\varepsilon\rho(SX,SY)^{1+\beta}. \tag{7.1}$$

Now, it may be checked by direct computations that

$$(I - S^\perp T)S^\perp(x-y) = S^\perp TS(x-y) + S^\perp T^\perp(x-y).$$

By Remark 7.3, $|S-T| \le C\varepsilon$, thus $(I - S^\perp T)$ is invertible and $|(I-S^\perp T)^{-1}| \le 2$ provided ε is small enough. In particular, by letting $L = (I - S^\perp T)^{-1}S^\perp$, we have $|L| \le 2$ and

$$\left|S^\perp x - S^\perp y - LT(Sx - Sy)\right| \le |LT^\perp(x-y)| \le 2|T^\perp(x-y)| \le C\varepsilon\rho(SX,SY)^{1+\beta}. \qquad (7.2)$$

We may repeat the discussion above for any other couple of points X, Y. In particular, by (7.1), $\Sigma \cap Q_\eta$ is included in the graph of some function $u : S(\Sigma \cap Q^d_\eta) \to \mathbb{R}^{d-m}$. Moreover, by (7.2), u is differentiable, and for every $X \in Q^d_\eta \cap \Sigma$, we have

$$\nabla u(Sx)v' = (I - S^\perp T_X)^{-1}S^\perp T_X S^{-1}(v')$$

for every $v' \in \mathbb{R}^m$. In particular, provided ε is small enough, it holds

$$|\nabla u(SX) - \nabla u(SY)| \le C|T_X - T_Y| \le C\varepsilon\rho(SX,SY)^\beta$$

for every $X, Y \in \Sigma \cap Q^d_\eta$. It is straightforward to see that u can be extended to Q^m_η and

$$\|u\|_{C^{1,\alpha}(Q_\eta)} \le C\varepsilon.$$

Step 3: conclusion. We let

$$E_+ := \{(x,t) \in Q^d_\eta \colon x_m > \gamma_m(x'')\}, \qquad E_- := \{(x,t) \in Q^d_\eta \colon x_m < \gamma_m(x'')\}.$$

By Step 2 and Theorem 3.10, since $\Gamma \cap E_+ = \varnothing$, either $(\Sigma \setminus \Gamma) \cap E_+ = \operatorname{graph} u \cap E_+$ or $(\Sigma \setminus \Gamma) \cap E_+ = \varnothing$. The same holds true for the component of $\Sigma \setminus \Gamma$ inside E_-; however, by Proposition 5.1 applied with $R = \eta$, it must be $(\Sigma \setminus \Gamma) \cap E_- = \varnothing$. On the other hand, since $(0,0) \in \Sigma$, by Proposition 3.4 it cannot be $(\Sigma \setminus \Gamma) \cap E_+ = \varnothing$ as well. Therefore

$$\Sigma \cap Q_\eta = \overline{E_+} \cap \operatorname{graph} u$$

and

$$M = \mathcal{H}^{m,1} \llcorner (\operatorname{graph} u \cap E_+).$$

We conclude the proof by showing that $(x', u(x',t)) \in \Gamma$ whenever $x' \in S(\Gamma)$. Indeed, if that was not the case for some $(x',t) \in S(\Gamma) \times (-\eta^2, 0]$, then, by continuity of u, $\operatorname{graph} u$ would detach from Γ in a set of positive $\mathcal{H}^{m-1,1}$ measure. This however would violate the fact that $M_t \in \mathcal{V}_m(B_\eta, \Gamma)$ for almost every t, thus reaching a contradiction. □

A Proof of Lemma 3.2

Up to rescaling and translating, it is sufficient to prove that there exist A and Λ small and positive, depending only on δ and α, such that, if Γ is a $(m-1)$-dimensional properly embedded submanifold of B_2 with $[\Gamma]_{C^{1,\alpha}(B_2)} \le A$, then

$$\int_{-\Lambda}^0 \int |T_y\Gamma^\perp \nabla\Psi_1(y,t)| \, d\Gamma(y)\, dt \le \frac{1}{2}\chi_{\Gamma^c}(0) + \delta. \qquad (A.1)$$

For brevity, we denote by Γ_y the space $T_y\Gamma$. Throughout the proof, C will denote constants (possibly changing from one expression to another) depending only on m, d, α.

Case 1: $0 \in \Gamma$. We start by remarking that

$$\begin{aligned}&\int_{-\Lambda}^0 \int |\Gamma_y^\perp \nabla\Psi_1(\cdot,t)| \, d\Gamma\, dt \\ &\qquad \le \int_{-\Lambda}^0 \int_{B_2} |\Gamma_y^\perp \nabla\Psi(\cdot,t)| \, d\Gamma(y)\, dt + C\Gamma(B_2)\int_{-\Lambda}^0 \frac{e^{1/(4t)}}{(-t)^{m/2}}\, dt.\end{aligned}$$

If A is smaller than some universal constant, then $\Gamma(B_2) \leq C$, thus we may take Λ small depending on δ so that the last term in the above inequality is smaller than $\delta/2$. Therefore we reduce ourselves to proving that, if A is small, then

$$I_1 := \int_{-\Lambda}^0 \int_{B_2} |\Gamma_y^\perp \nabla \Psi(\cdot, t)| \, d\Gamma \, dt \leq \frac{\delta}{2}.$$

Since $[\Gamma]_{C^{1,\alpha}(B_2)} \leq A$ is small, for every $(y,t) \in \Gamma \times (-\infty, 0)$:

$$|\Gamma_y^\perp \nabla \Psi(y,t)| \leq C \frac{e^{|y|^2/t}}{(-t)^{1+m/2}} |\Gamma_y^\perp y| \leq CA \frac{e^{|y|^2/t}}{(-t)^{1+m/2}} |y|^{1+\alpha}.$$

We then use the fact that, if A is smaller than some universal constant, then $\Gamma \cap B_2$ is the graph over Γ_0 of some function $\gamma : \mathbb{R}^{m-1} \to \mathbb{R}^{d-m+1}$ such that $\|\gamma\|_{C^{1,\alpha}} \leq CA$. In particular, by using the area formula and the fact that the $||\nabla\gamma||_{L^\infty(B_2)} \leq 1$ for A small enough, we obtain

$$\int_{B_2} |\Gamma_y^\perp \nabla \Psi(y,t)| \, d\Gamma(y) \leq CA \frac{1}{(-t)^{1+m/2}} \int_{\mathbb{R}^{m-1}} |y|^{1+\alpha} e^{|y|^2/t} \, d\mathcal{L}^{m-1}(y) = CA(-t)^{\frac{\alpha}{2}-1}$$

for some C depending only on m and α. Therefore, assuming $\Lambda \leq 1$,

$$I_1 \leq CA \int_{-1}^0 (-t)^{\frac{\alpha}{2}-1} \, dt \leq CA.$$

We conclude the proof in the case $0 \in \Gamma$ by choosing $A \leq \frac{\delta}{2C}$.

Case 2: $0 \notin \Gamma$. Let E_Γ be the m-dimensional Hausdorff measure restricted to the exterior cone

$$C_\gamma := \{\lambda y \colon \lambda \geq 1 \text{ and } y \in \Gamma\}$$

with multiplicity, as defined in [20, Section 7]. With similar computations to those in the proof of [20, Theorem 7.1], we may show that

$$\begin{aligned}
&\int_{-\Lambda}^0 \int |\Gamma_y^\perp \nabla \Psi_1(y,t)| \, d\Gamma(y) \, dt \\
&\qquad \leq -\lim_{\tau \nearrow 0} \int \Psi_1(\cdot, \tau) \, dE_\Gamma + \int \Psi_1(\cdot, -\Lambda) \, dE_\Gamma + C\Lambda E_\Gamma(B_2) \\
&\qquad = \int \Psi_1(\cdot, -\Lambda) \, dE_\Gamma + C\Lambda E_\Gamma(B_2),
\end{aligned} \tag{A.2}$$

Where the last equality comes from the fact that $0 \notin \Gamma$.

In order to prove (A.1), we argue by contradiction: assume there is a sequence $\{\Gamma^j\}$ with $0 \notin \Gamma^j$ such that $\left\|\Gamma^j\right\|_{C^{1,\alpha}(B_2)} \leq \frac{1}{j}$ for which the left-hand side of (A.2) is greater than $\frac{1}{2} + \delta$. One may show that, up to extracting a subsequence, E_{Γ^j} converges weakly to $\mathcal{H}^m \llcorner S^+$, where S^+ is some m-dimensional half plane such that $0 \notin \mathrm{Int}(S^+)$. Therefore

$$\limsup_{j \to \infty} \left\{ \int \Psi_1(\cdot, -\Lambda) E_{\Gamma^j} + C\Lambda E_{\Gamma^j}(B_2) \right\} \leq \int_{S^+} \Psi_1(\cdot, -\Lambda) \, d\mathcal{H}^m + C\Lambda \mathcal{H}^m(\overline{B_2} \cap S^+).$$

Since $0 \notin \mathrm{Int}(S^+)$, the integral in the right-hand side of the above inequality is smaller than $\frac{1}{2}$ for every choice of Λ. On the other hand, Λ may be chosen so small that

$$C\Lambda \mathcal{H}^m(\overline{B_2} \cap S^+) \leq \frac{\delta}{2},$$

which contradicts the assumption made above, thus concluding the proof.

B Proof of Lemma 3.6

We refer the reader to [9, Lemma 9.4] for a detailed proof of Lemma 3.6. Since some minor modifications are needed, in this section we sketch the outline of the proof.

Let $U \subset \mathbb{R}^d$ be open, $I \subset \mathbb{R}$ be a non-empty interval, Γ be a $(m-1)$-dimensional $C^{1,\alpha}$ submanifold of U and let $\mathbf{M} \in \mathcal{BF}_m(U \times I, \Gamma)$. We assume that $\mathbf{M}$ satisfies a bound of the form

$$\mathrm{mdr}(\mathbf{M}, U \times I) \leq E_1 < \infty \tag{B.1}$$

and we let $\Sigma = \Sigma_{\mathbf{M}}$ be its space-time support.

Before proceeding, by virtue of Proposition 3.4, we fix small constants c_1, c_2 and R_0, depending on E_1 and Γ, such that for every $(x,t) \in \Sigma$ and every $R \leq R_0$ such that $B_R(x) \times (t - c_1 R^2, t) \subset\subset U \times I$, it holds

$$M_{t-c_1R^2}(B_{R/2}(x)) \geq c_2 R^m.$$

By Definition 2.3, for almost every $t \in I$ there exist a m-dimensional rectifiable set $E \subset U$ and a positive, integer valued function $\theta_t : E_t \to \mathbb{N}$ such that $M_t = \theta_t(\cdot)\mathcal{H}^m \llcorner E_t$ We choose a time t as above, with the additional condition that $s \mapsto M_s(\varphi)$ is continuous at t for every $\varphi \in C_c(U)$. By [15, Proposition 3.3], almost every $t \in I$ satisfies the latter condition.

We claim that, for every such t and for every $B_{3r}(x_0) \subset\subset U$, it holds

$$\mathcal{H}^m((\Sigma_t \setminus E_t) \cap B_r(x_0)) = 0; \tag{B.2}$$

this clearly implies (3.4).

In order to prove (B.2), we argue by contradiction. Assume that there is $(x_0, t_0) \in U \times I$ and $r > 0$ such that $B_{3r} \subset\subset U$ and $\mathcal{H}^m(A \cap B_r(x_0)) > 0$. Without loss of generality, we may take $x_0 = 0$, $t = 0$ and set $A := \Sigma_0 \setminus E_0$.

Let

$$A_k := \big\{x \in A \cap B_r \colon M_0(B_R(x)) \leq c_2 R^m/2 \text{ for all } R \in (0, r/k)\big\}.$$

Since, for $\mathcal{H}^m$-a.e. $x \in E_0^c$, it holds

$$\lim_{R \searrow 0} \frac{M_0(B_R(x))}{R^m} = 0,$$

we have

$$0 < \mathcal{H}^m(A \cap B_r) = \mathcal{H}^m\Big(\bigcup_{k \in \mathbb{N}} A_k\Big).$$

Therefore we may find $k \in \mathbb{N}$ such that $b_0 := \mathcal{H}^m(A_k) > 0$.

By standard measure-theoretic arguments, it is not hard to show that there exists c small universal such that, for every R small enough, we may find $N \in \mathbb{N}$ and a finite collection of points $\{x_j\}_{j=1}^N \subset A_k$ such that $\{B_R(x_j)\}$ are mutually disjoint and

$$NR^m \geq cb_0. \tag{B.3}$$

By definition of A, since $x_j \in A_k$, we have

$$M_0\Big(\bigcup_{j=1}^N B_R(x_j)\Big) \leq Nc_2 \frac{R^m}{2}. \tag{B.4}$$

On the other hand, by Proposition 3.4 and the fact that $x_j \in \Sigma_0$, we have

$$M_{-c_1R^2}\Big(\bigcup_{j=1}^N B_{R/2}(x_j)\Big) \geq Nc_2 R^m. \tag{B.5}$$

We now fix a cut-off function $\varphi \in C_c^\infty(B_1)$ such that $\varphi \in [0,1]$ everywhere, $\varphi|_{B_{1/2}} \equiv 1$ and $|\nabla\varphi| \leq 4$. Then, given R small, we let $\varphi_0(x) = \varphi(x/(2r))$, $\varphi_j(x) = \varphi((x - x_j)/R)$ and

$$\widetilde{\varphi} = \varphi_0 - \sum_{j=1}^N \varphi_j.$$

Then clearly $\widetilde{\varphi} \in [0,1]$ everywhere and $|\nabla\widetilde{\varphi}| \le C/R$. Notice, moreover, that

$$\sum_{j=1}^{N} \chi_{B_{R/2}(x_j)} \le \sum_{j=1}^{N} \varphi_j \le \sum_{j=1}^{N} \chi_{B_R(x_j)}$$

For brevity, set $s = -c_1 R^2$. By (B.4) and (B.5), we have

$$\begin{aligned} M_0(\varphi_0) - M_s(\varphi_0) &= \big(M_0(\widetilde{\varphi}) - M_s(\widetilde{\varphi})\big) + \Big(M_0\big(\sum \varphi_j\big) - M_s\big(\sum \varphi_j\big)\Big) \\ &\le \big(M_0(\widetilde{\varphi}) - M_s(\widetilde{\varphi})\big) + (Nc_2R^m/2 - Nc_2R^m) \\ &\le \big(M_0(\widetilde{\varphi}) - M_s(\widetilde{\varphi})\big) - c_3 b_0 \end{aligned} \tag{B.6}$$

for some c_3 small, where (B.3) was used in the last inequality.

We now estimate, by using Definition 2.3,

$$\begin{aligned} M_0(\widetilde{\varphi}) - M_s(\widetilde{\varphi}) &\le \int_s^0 \int H \cdot \nabla\widetilde{\varphi}\, dM_t\, dt \\ &\le \Big(\int_s^0 \int_{B_{2r}} |H|^2\Big)^{1/2} \Big(\int_s^0 \int |\nabla\widetilde{\varphi}|^2\Big)^{1/2}. \end{aligned} \tag{B.7}$$

By (B.1) and the fact that $s = -c_1 R^m$, we have, for some C large,

$$\int_s^0 \int |\nabla\widetilde{\varphi}|^2\, dM_t\, dt \le (-s)||\nabla\widetilde{\varphi}||_\infty M_t(B_{2r}) \le CE_1 r^m,$$

therefore (B.6) and (B.7) yield

$$\int_{-c_2R^2}^0 \int_{B_{2r}} |H|^2\, dM_t\, dt \ge \frac{1}{CE_1r^m}\Big(M_0(\varphi_0) - M_{-c_1R^2}(\varphi_0) + c_3 b_0\Big).$$

By assumption, $t \mapsto M_t(\varphi_0)$ is continuous at 0. Thus we may choose R so small that the right-hand side of the above inequality is larger than $c\frac{b_0}{E_1r^m}$ for some c small enough.

Finally, we consider the function $\widehat{\varphi} = \varphi(x/(3r))$. By Definition 2.3 and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} M_0(\widehat{\varphi}) - M_{-c_1R^2}(\widehat{\varphi}) &\le \int_{-c_1R^2}^0 \int_{B_{3r}} \big(-\widehat{\varphi}|H|^2 + \nabla\widehat{\varphi}\cdot H\big)\, dM_t\, dt \\ &\le -\int_{-c_1R^2}^0 \int_{B_{3r}} \frac{1}{2}\widehat{\varphi}|H|^2\, dM_t\, dt + \int_{-c_1R^2}^0 \int_{B_{3r}} \frac{|\nabla\widehat{\varphi}|^2}{2\widehat{\varphi}}\, dM_t\, dt \\ &\le -\frac{1}{2}\int_{-c_1R^2}^0 \int_{B_{2r}} |H|^2\, dM_t\, dt + CE_1r^mR^2 \\ &\le -c\frac{b_0}{E_1r^m} \end{aligned}$$

provided R is chosen small enough. This contradicts the continuity of $t \mapsto M_t(\widehat{\varphi})$ at 0, thus concluding the proof.

References

[1] William K. Allard. "On the first variation of a varifold". In: *Ann. of Math. (2)* 95 (1972), pp. 417–491.

[2] William K. Allard. "On the first variation of a varifold: boundary behavior". In: *Ann. of Math. (2)* 101 (1975), pp. 418–446.

[3] Luigi Ambrosio and H. Mete Soner. “A measure-theoretic approach to higher codimension mean curvature flows”. In: *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* 25.1-2 (1997), pp. 27–49.

[4] Theodora Bourni. “Allard-type boundary regularity for $C^{1,\alpha}$ boundaries”. In: *Adv. Calc. Var.* 9.2 (2016), pp. 143–161.

[5] Kenneth A. Brakke. *The motion of a surface by its mean curvature.* Vol. 20. Mathematical Notes. Princeton University Press, Princeton, N.J., 1978.

[6] Luis A. Caffarelli and Xavier Cabré. *Fully nonlinear elliptic equations.* Vol. 43. American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, RI, 1995.

[7] Nick Edelen. “The free-boundary Brakke flow”. In: *J. Reine Angew. Math.* 758 (2020), pp. 95–137.

[8] Or Hershkovits and Brian White. *Avoidance for Set-Theoretic Solutions of Mean Curvature-Type Flows.* 2018. URL: `https://arxiv.org/abs/1809.03026`.

[9] Kota Kasai and Yoshihiro Tonegawa. “A general regularity theory for weak mean curvature flow”. In: *Calc. Var. Partial Differential Equations* 50.1-2 (2014), pp. 1–68.

[10] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces.* Vol. 12. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1996.

[11] O. Savin. “Viscosity solutions and the minimal surface system”. In: Harv. Univ. Cent. Math. Sci. Appl. Ser. Math. 2 (2018), pp. 135–145.

[12] Ovidiu Savin. “Small perturbation solutions for elliptic equations”. In: *Comm. Partial Differential Equations* 32.4-6 (2007), pp. 557–578.

[13] Leon Simon. *Lectures on geometric measure theory.* Vol. 3. Proceedings of the Centre for Mathematical Analysis, Australian National University. Australian National University, Centre for Mathematical Analysis, Canberra, 1983.

[14] Salvatore Stuvard and Yoshihiro Tonegawa. “An existence theorem for Brakke flow with fixed boundary conditions”. In: *Calc. Var. Partial Differential Equations* 60.1 (2021), Paper No. 43, 53.

[15] Yoshihiro Tonegawa. *Brakke’s mean curvature flow.* SpringerBriefs in Mathematics. An introduction. Springer, Singapore, 2019.

[16] Lihe Wang. “On the regularity theory of fully nonlinear parabolic equations. II”. In: *Comm. Pure Appl. Math.* 45.2 (1992), pp. 141–178.

[17] Yu Wang. “Small perturbation solutions for parabolic equations”. In: *Indiana Univ. Math. J.* 62.2 (2013), pp. 671–697.

[18] Brian White. “A local regularity theorem for mean curvature flow”. In: *Ann. of Math. (2)* 161.3 (2005), pp. 1487–1519.

[19] Brian White. “Controlling area blow-up in minimal or bounded mean curvature varieties”. In: *J. Differential Geom.* 102.3 (2016), pp. 501–535.

[20] Brian White. “Mean curvature flow with boundary”. In: *Ars Inven. Anal.* 2021 (2021). Id/No 4, p. 43.

[21] Brian White. “The topology of hypersurfaces moving by mean curvature”. In: *Comm. Anal. Geom.* 3.1-2 (1995), pp. 317–333.