

# A SHORT PROOF OF ALLARD'S AND BRAKKE'S REGULARITY THEOREMS

GUIDO DE PHILIPPIS, CARLO GASPARETTO, AND FELIX SCHULZE

ABSTRACT. We give new short proofs of Allard's regularity theorem for varifolds with bounded first variation and Brakke's regularity theorem for integral Brakke flows with bounded forcing. They are based on a decay of flatness, following from weighted versions of the respective monotonicity formulas, together with a characterization of non-homogeneous blow-ups using the viscosity approach introduced by Savin.

## 1. INTRODUCTION

Allard's and Brakke's  $\varepsilon$ -regularity theorems are key tools in the study of, respectively, minimal surfaces and mean curvature flows, and they can be roughly stated as follows:

*If a  $m$ -dimensional minimal surface (resp. the space-time track of a mean curvature flow) is sufficiently flat in the ball of radius 1 (resp. parabolic cylinder of radius 1) and its area is roughly the one of the unit  $m$ -dimensional disk (resp. the weighted Gaussian density of a unit disk) then in a smaller ball (resp. parabolic cylinder) it can be written as the graph of a smooth function which enjoys suitable a-priori estimates.*

See Theorems 2.1 and 3.8 below for the rigorous statement. Note these results are also relevant in the smooth category, since the scale and the regularity of the graphical parametrization of the surface only depend on the a priori assumption of (weak) closeness to the  $m$ -dimensional unit disk.

The original proofs by Allard and Brakke are modeled on the pioneering ideas introduced by De Giorgi in the regularity theory for co-dimension 1 area minimizing surfaces, [7] and on their implementation done by Almgren in [3]. In particular the proof is, roughly speaking, divided into the following steps:

- (a) Under the desired assumption it is possible to show that most of the surface can be covered by the graph of a Lipschitz function, whose  $W^{1,2}$  norm can be estimated by the difference in area between the surface and the plane.
- (b) Since the minimal surface equation (respectively the mean curvature flow) linearizes on the Laplacian equation (resp. the heat equation), this function is close to a harmonic (resp. caloric) function which enjoys strong a priori estimates.
- (c) These estimates can be pulled back to the minimal surface (resp. space time track of the mean curvature flow) to show that the initial assumptions are satisfied also in the ball of radius  $1/2$ . A suitable iteration provides then the conclusion.

In this approach the two main difficulties lie in the approximation procedure in step (a) and in proving that the closeness in step (b) is in a strong enough topology to be

---

*Key words and phrases.* Minimal surfaces, mean curvature flow, regularity theory.

**Acknowledgments.** G.D.P. has been partially supported by the NSF grant DMS 2055686 and by the Simons Foundation. C.G. is supported by the European Research Council (ERC), under the European Union's Horizon 2020 research and innovation program, through the project ERC VAREG - Variational approach to the regularity of the free boundaries (grant agreement No. 853404) and partially supported by INDAM-GNAMPA. Part of this work was carried out while C.G. was a PhD student at SISSA-Trieste.

able to pull-back the estimates from the linearized equation. Among various references, we refer the reader to [1] for a very clear account of Allard's theorem and to [11] for a simplified proof of Brakke theorem. Inspired by the work of Caffarelli and Cordoba [6], in [12] Savin provided a viscosity type approach to the above proofs, in which step (a) is completely avoided and step (b) is replaced by a partial Harnack inequality, obtained via Aleksandrov-Bakelmann-Pucci (ABP) type estimates, see [14] and [18] for the extension to the minimal surface system and to parabolic equations, respectively. This approach has been recently exploited by the second named author to prove a boundary version of Brakke's regularity theorem, [8]. We also mention that Savin's partial Harnack inequality have been a crucial ingredient in the proof of the De Giorgi conjecture on solutions of the Allen Cahn equation, [13].

In this note we show how combining both the "variational" and the "viscosity" approach, it is possible to obtain a very short and self-sufficient proof of both Allard's and Brakke's theorems. The key observation is that while viscosity techniques are very robust in allowing to pass to the limit in the equation under  $L^\infty$  convergence (a key step in Savin's approach) the ABP estimate can be replaced by a simple variational argument based on the fact that coordinates are harmonic (resp. caloric) when restricted to the minimal surface (resp. mean curvature flow) and that for harmonic functions the mean value inequality can be easily obtained by testing the weak formulation of the equation with a suitable truncation of the fundamental solution, see for instance [5].

This short note is organized as follows: in Section 2 we prove Allard's theorem and in Section 3 we prove Brakke's theorem. In order to make the note self-contained, we conclude with an appendix where we record the proof of the maximum principle for varifolds and Brakke flows. Although the proofs are done in the natural context of varifolds and Brakke flows, we invite the reader to take in mind the simple case of a smooth surface with zero mean curvature and a smooth mean curvature flow.

## 2. ALLARD'S REGULARITY THEOREM

In the following, we denote by  $\mathcal{M}_m(U)$  the space of  $m$ -dimensional rectifiable Radon measures on  $U$ , namely those Radon measures  $M$  on  $U$  for which there is a  $m$ -dimensional rectifiable set  $E \subset U$  with  $\mathcal{H}^m(E) < \infty$  and  $M \ll \mathcal{H}^m \llcorner E$ . If  $M \in \mathcal{M}_m(U)$ , then  $M = \Theta^m(M, \cdot) \mathcal{H}^m$ , where

$$\Theta^m(M, x) = \lim_{r \searrow 0} \frac{M(B_r(x))}{\omega_m r^m}$$

wherever the limit exists. For  $M$ -almost every  $x$ , the approximate tangent plane to  $M$  at  $x$  is well defined and we denote it by  $T_x M \in \text{Gr}(m, d)$ . For  $S \in \text{Gr}(m, d)$  and  $F \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ , we introduce the notation

$$\text{div}_S F(x) = \sum_{i=1}^m \nabla F(x) \eta_i \cdot \eta_i,$$

where  $\{\eta_i\}_{i=1}^m$  is any orthonormal basis of  $S$ . Next, for every  $p \in (1, +\infty]$ , we let  $\mathcal{M}_m^p(U)$  be the set of those measures  $M \in \mathcal{M}_m(U)$  such that, for every  $W \subset\subset U$  there is  $h_W \in \mathbb{R}$  such that

$$(2.1) \quad \int \text{div}_{T_x M} F(x) \, dM(x) \leq h_W \left( \int |F|^{p'} \, dM \right)^{1/p'}$$

for every  $F \in C_c^1(W; \mathbb{R}^d)$ , where  $p'$  is the conjugate exponent of  $p$ . To each  $M \in \mathcal{M}_m^p(U)$  we associate a vector field  $H_M \in L_{loc}^p(M; \mathbb{R}^d)$  such that

$$(2.2) \quad \int \text{div}_{T_x M} F \, dM = - \int H_M \cdot F \, dM$$

for every  $F \in C_c^1(U; \mathbb{R}^d)$ . We call  $H_M$  the *generalized mean curvature vector* of  $M$ . Whenever the indication of  $M$  is unnecessary, we write  $H$  in place of  $H_M$ .

The goal of the present section is proving the following version of Allard's theorem:

**Theorem 2.1** (Allard's regularity theorem). *For every  $\alpha \in (0, 1)$ , there is  $\delta_0 > 0$  with the following property. Let  $M \in \mathcal{M}_m^\infty(B_1)$  and assume that  $0 \in \text{supp } M$ ,  $\Theta^m(M, x) \geq 1$  for  $M$ -almost every  $x$ ,*

$$M(B_1) \leq (1 + \delta_0)\omega_m \quad \text{and} \quad \|H_M\|_{L^\infty(M)} \leq \delta_0.$$

*Then  $\text{supp } M \cap B_{1/2}$  is the graph of some function  $u \in C^{1,\alpha}(B_1^m; \mathbb{R}^{d-m})$ .*

If  $H$  is more regular, then Schauder estimates entail higher regularity for  $\text{supp } M$ , as well. We shall prove Theorem 2.1 in Subsection 2.2. The next subsection is dedicated to proving a decay property of the oscillations of  $\text{supp } M$ , which allows us to prove an improvement of flatness (see Theorem 2.8 below).

**2.1. Decay of oscillations.** In the present subsection, we assume the following:

**Assumption 2.2.**  *$M \in \mathcal{M}_m^\infty(B_R)$  is such that:*

- (1)  $\Theta^m(M, x) \geq 1$  for  $M$ -almost every  $x \in B_R$ ;
- (2) for every  $r \in [0, R]$ ,  $M(B_r) \leq \frac{3}{2}\omega_m r^m$ .
- (3)  $\Lambda := \|H_M\|_{L^\infty(M)} < \infty$ .

**Proposition 2.3** (Weighted monotonicity formula). *Let  $f : B_R \rightarrow \mathbb{R}$  be a non-negative, convex function such that  $\|\nabla f\|_\infty \leq 1$ . Provided  $0 \in \text{supp } M$ , then for every  $0 < r \leq R$*

$$\frac{1}{\omega_m r^m} \int_{B_r} f \, dM \geq f(0) - C_0 \Lambda (\|f\|_{L^\infty(M)} + r) r$$

*for some  $C_0$  universal.*

**Remark 2.4.** Although this fact will not be used in the following, we point out that the above result holds true provided  $f$  is  $m$ -convex, meaning that the sum of the  $m$  smallest eigenvalues of  $D^2 f$  is non-negative.

**Proof.** We begin with some preliminary computations. Assume<sup>1</sup>  $m > 2$  and let

$$h(x) = \frac{1}{\omega_m m(m-2)} \begin{cases} \frac{m}{2} - \frac{m-2}{2}|x|^2 & \text{if } |x| \leq 1 \\ |x|^{2-m} & \text{if } |x| \geq 1. \end{cases}$$

Note that  $h \in C^{1,1}(\mathbb{R}^d) \cap C^2(\mathbb{R}^d) \setminus \partial B_1$  and

$$(2.3) \quad \nabla h(x) = -\frac{1}{m\omega_m} \begin{cases} x & \text{if } |x| < 1 \\ \frac{x}{|x|^m} & \text{if } |x| \geq 1 \end{cases}$$

as well as

$$(2.4) \quad D^2 h(x) = -\frac{1}{\omega_m} \begin{cases} \frac{1}{m} \mathbb{1} & \text{if } |x| < 1 \\ |x|^{-m} \left( \frac{1}{m} \mathbb{1} - \frac{x \otimes x}{|x|^2} \right) & \text{if } |x| > 1. \end{cases}$$

For any  $0 < r \leq s < R$ , we then let

$$g_{r,s}(x) = r^{2-m} h(x/r) - s^{2-m} h(x/s).$$

Note that  $g_{r,s} \geq 0$  and  $g_{r,s} \equiv 0$  outside  $B_s$ . Straightforward computations give, using (2.3) and (2.4), for any  $S \in \text{Gr}(m, d)$ :

$$\text{div}_S \nabla g_{r,s}(x) = -\frac{\chi_{B_r}}{\omega_m r^m} + \frac{\chi_{B_s}}{\omega_m s^m} + \frac{\chi_{B_s \setminus B_r}}{\omega_m |x|^m} \left( \frac{|Sx|^2}{|x|^2} - 1 \right) \leq -\frac{\chi_{B_r}}{\omega_m r^m} + \frac{\chi_{B_s}}{\omega_m s^m}.$$

<sup>1</sup>In the case  $m = 2$ , a logarithmic term replaces  $|x|^{2-m}$ .

Moreover, since  $f$  is non-negative and convex

$$(2.5) \quad \begin{aligned} \operatorname{div}_S(f\nabla g_{r,s} - g_{r,s}\nabla f) &= f \operatorname{div}_S(\nabla g_{r,s}) - g_{r,s} \operatorname{div}_S(\nabla f) \\ &\leq f \operatorname{div}_S(\nabla g_{r,s}) \\ &\leq -\frac{f\chi_{B_r}}{\omega_m r^m} + \frac{f\chi_{B_s}}{\omega_m s^m}. \end{aligned}$$

Let now

$$I(r) = \frac{1}{\omega_m r^m} \int_{B_r} f \, dM.$$

By (2.2) and (2.5), for any  $0 < r < s \leq R$ , it holds

$$I(s) - I(r) \geq \int \operatorname{div}_{T_x M}(f\nabla g_{r,s} - g_{r,s}\nabla f) \, dM = - \int H \cdot (f\nabla g_{r,s} - g_{r,s}\nabla f) \, dM.$$

Therefore

$$(2.6) \quad \frac{I(s) - I(r)}{s - r} \geq - \int f H \cdot \frac{\nabla g_{r,s}}{s - r} \, dM + \int \frac{g_{r,s}}{s - r} H \cdot \nabla f \, dM =: -A_1 + A_2.$$

By Hölder's inequality, we may estimate

$$|A_1| \leq \Lambda \sup \left| \frac{\nabla g_{r,s}}{s - r} \right| \int_{B_s} f \, dM \leq C \Lambda \sup \left| \frac{\nabla g_{r,s}}{s - r} \right| s^m I(s)$$

and

$$|A_2| \leq \Lambda \sup \left| \frac{g_{r,s}}{s - r} \right| \|\nabla f\|_\infty M(B_s) \leq C \Lambda \sup \left| \frac{g_{r,s}}{s - r} \right| s^m$$

where in the second inequality we have used the fact that  $\|\nabla f\|_\infty \leq 1$  and Item 2 in Assumption 2.2 above. Direct computations give, using (2.3) and (2.4) together with the convexity of  $t \mapsto t^{-m}$ ,

$$\limsup_{s \searrow r} \sup_{x \in B_R} \left| \frac{\nabla g_{r,s}}{s - r} \right| \leq C r^{-m}, \quad \limsup_{s \searrow r} \sup_{x \in B_R} \left| \frac{g_{r,s}}{s - r} \right| \leq C r^{1-m}$$

for some  $C$  depending only on  $m$ . Thus, letting  $s \searrow r$  in (2.6), together with standard approximation arguments (see for example [15, §17]), gives

$$I'(r) \geq -C\Lambda(I(r) + r)$$

in the sense of distributions. We now multiply both sides of the latter inequality by the integrating factor  $F(r) = e^{C\Lambda r}$  and integrate from some  $h > 0$  to  $r$  to obtain

$$I(r) \geq e^{C\Lambda(h-r)} I(h) - C\Lambda \int_h^r e^{C\Lambda(t-r)} t \, dt \geq I(h) - C\Lambda r \|f\|_{L^\infty(M)} - C\Lambda r^2,$$

where in the second inequality we have used the facts that  $e^t \geq 1 + t$  and that, by Item 2 in Assumption 2.2,  $I(h) \leq C\|f\|_{L^\infty(M)}$ . Finally, if  $\Theta^m(M, 0) \geq 1$ , letting  $h \searrow 0$  in the above inequality gives the desired result. Otherwise, if  $0 \in \operatorname{supp} M$  then there is a sequence  $x_j \rightarrow 0$  such that  $\Theta^m(M, x_j) \geq 1$ . The result follows by continuity of  $f$ .  $\square$

**Remark 2.5.** The role of the factor  $3/2$  in Assumption 2.2 (2), as will be clear from the proof of Proposition 2.6 below, is to rule out the possibility that  $M$  consists of two separated sheets, which would clearly violate the conclusion of a Harnack inequality. On the other hand, for what concerns Proposition 2.3, by virtue of the classical monotonicity formula (see, for instance, [15, §17]), (2) in Assumption 2.2 may be replaced by the weaker assumptions

$$M(B_R) \leq E_0 R^m \text{ and } \Lambda \leq \frac{E_0}{R},$$

with the caveat that  $C_0$  in the conclusion of Proposition 2.6 depends on  $E_0$ .

The above result allows us to prove a *partial Harnack inequality*. In order to do so, for  $S \in \text{Gr}(m, d)$ ,  $M$  as in Assumption 2.2, and  $x_0 \in \text{supp } M$  with  $B_r(x_0) \subset B_R$ , we introduce the notation

$$\text{osc}_S(M, B_r(x_0)) = \frac{1}{2} \sup \{ |S^\perp(x - y)| : x, y \in \text{supp } M \cap B_r(x_0) \}$$

to denote the radius of the smallest cylinder of the form  $\{x \in \mathbb{R}^d : |S^\perp(x - y)| \leq h\}$  for some  $y \in \mathbb{R}^d$  that contains  $\text{supp } M \cap B_r(x_0)$ . Whenever  $x_0 = 0$  and the indication of  $M$  and of the  $m$ -plane  $S$  is unnecessary, as it is in the next two results, we write  $\text{osc}(B_r)$  in place of  $\text{osc}_S(M, B_r(x_0))$ .

**Proposition 2.6** (Harnack inequality). *Let  $M$  satisfy Assumption 2.2 and let  $0 \in \text{supp } M$ . There exists a universal constant  $\eta \in (0, 1)$  such that, if*

$$\text{osc}(B_R) \leq \eta R \quad \text{and} \quad \Lambda \leq \frac{\text{osc}(B_R)}{R^2},$$

then

$$\text{osc}(B_{\eta R}) \leq (1 - \eta) \text{osc}(B_R).$$

**Proof.** Let  $\varepsilon = \frac{1}{R} \text{osc}(B_R)$  and  $\Sigma = \text{supp } M$ . We can thus choose  $y_0 \in B_{\varepsilon R}$  such that  $\Sigma \cap B_R \subset \{y : |S^\perp(y - y_0)| \leq \varepsilon R\}$ . Assume, by contradiction, that there are points  $y_1, y_2 \in \Sigma \cap B_{\eta R}$  such that  $|S^\perp(y_1 - y_2)| > 2(1 - \eta)\varepsilon R$ . Denote  $\omega := \frac{S^\perp(y_1 - y_2)}{|S^\perp(y_1 - y_2)|}$  and consider

$$f_1(x) = \left( (x - y_0) \cdot \omega - \frac{\varepsilon R}{2} \right)^+ \quad \text{and} \quad f_2(x) = \left( -(x - y_0) \cdot \omega - \frac{\varepsilon R}{2} \right)^+$$

and let

$$I_i(r) = \frac{1}{\omega_m r^m} \int_{B_r(y_i)} f_i \, dM.$$

Notice that, by assumption,  $\|f_i\|_{L^\infty(M \cap B_R)} \leq \frac{\varepsilon}{2} R$  and  $\Lambda \leq \varepsilon R^{-1}$ . Moreover,  $f_i(y_i) \geq (\frac{1}{2} - 2\eta)\varepsilon R$ . Therefore, by Proposition 2.3, for every  $\theta \in [\eta, 1/2]$  (to be determined later) it holds

$$(2.7) \quad \begin{aligned} I_i(\theta R) &\geq f_i(y_i) - C_0 \Lambda \theta R (\|f_i\|_{L^\infty(M)} + \theta R) \\ &\geq \left( \frac{1}{2} - 2\eta \right) \varepsilon R - C_0 \varepsilon \theta \left( \frac{\varepsilon}{2} + \theta \right) R. \end{aligned}$$

Since, by assumption,  $\varepsilon \leq \eta \leq \theta \leq \frac{1}{2}$  we have  $C_0 \varepsilon \theta \left( \frac{\varepsilon}{2} + \theta \right) \leq 2C_0 \theta^2 \varepsilon \leq C_0 \theta \varepsilon$ . We thus obtain from (2.7)

$$(2.8) \quad I_i(\theta R) \geq \left( \frac{1}{2} - 2\eta - C_0 \theta \right) \varepsilon R \geq \left( \frac{1}{2} - (C_0 + 2)\theta \right) \varepsilon R$$

Note that  $\text{supp } f_1 \cap \text{supp } f_2 = \emptyset$ . Using this fact and the assumptions  $M(B_r) \leq \frac{3}{2} \omega_m r^m$  and  $|y_1|, |y_2| \leq 2\varepsilon R \leq 2\eta R$ , we have

$$\begin{aligned} I_1(\theta R) + I_2(\theta R) &\leq \|f_i\|_\infty \frac{M(B_{\theta R}(y_1) \cup B_{\theta R}(y_2))}{\omega_m (\theta R)^m} \\ &\leq \frac{\varepsilon R}{2} \frac{M(B_{\theta R+2\eta R})}{\omega_m (\theta R)^m} \\ &\leq \frac{3}{4} \left( 1 + \frac{2\eta}{\theta} \right)^m \varepsilon R \leq \frac{3}{4} \left( 1 + C_1 \frac{\eta}{\theta} \right) \varepsilon R \end{aligned}$$

for some  $C_1$  depending only on  $m$ . We now specify our choices of  $\theta$  and  $\eta$ . We first choose  $\theta$  much smaller than  $\min\{C_1^{-1}, (C_0 + 2)^{-1}\}$  such that

$$2\left(\frac{1}{2} - (C_0 + 2)\theta\right) > \frac{3}{4}\left(1 + C_1\theta\right).$$

Fixing  $\eta \leq \theta^2$  we obtain a contradiction from (2.8) and summing for  $i \in \{1, 2\}$ .  $\square$

As a corollary, we have the following

**Proposition 2.7.** *There are positive universal constants  $\beta$  and  $C$  with the following property. Let  $M$  satisfy Assumption 2.2 and  $0 \in \text{supp } M$ . Then, for every  $r \in [C(\text{osc}(B_R) + \Lambda R^2), R]$ , it holds*

$$\text{osc}(B_r) \leq C(\text{osc}(B_R) + \Lambda R^2) \left(\frac{r}{R}\right)^\beta.$$

**Proof.** By rescaling, assume  $R = 1$ . For some  $K$  large to be determined later, let

$$F(r) = \text{osc}(B_r) + K\Lambda r^2.$$

We claim that, if  $\eta$  is the constant determined in Proposition 2.6, then for any  $r$  such that  $\text{osc}(B_r) \leq \eta r$  it holds

$$(2.9) \quad F(\eta r) \leq (1 - \eta)F(r).$$

Indeed, if  $\Lambda \leq \text{osc}(B_r)r^{-2}$ , then Proposition 2.6 yields  $\text{osc}(B_{\eta r}) \leq (1 - \eta)\text{osc}(B_r)$ , hence

$$F(\eta r) \leq (1 - \eta)\text{osc}(B_r) + \eta^2 K\Lambda r^2 \leq (1 - \eta)F(r)$$

since  $\eta$  is much smaller than 1. Otherwise, if  $\text{osc}(B_r) \leq \Lambda r^2$ , then

$$\begin{aligned} F(\eta r) &\leq \text{osc}(B_r) + K\Lambda\eta^2 r^2 \\ &\leq K\Lambda r^2 \left(\frac{1}{K} + \eta^2\right) \\ &\leq (1 - \eta)F(r) \end{aligned}$$

as desired, provided  $K \geq \frac{1}{1 - \eta - \eta^2}$ .

By induction, if  $\eta^k \geq F(1)$ , then

$$\text{osc}(B_{\eta^k}) \leq F(\eta^k) \leq (1 - \eta)^k F(1).$$

Now, for  $F(1) \leq r < 1$ , let  $k \in \mathbb{N} \cup \{0\}$  be such that  $\eta^{k+1} \leq r < \eta^k$ . Then we have

$$\text{osc}(B_r) \leq \text{osc}(B_{\eta^k}) \leq (1 - \eta)^k F(1) \leq \frac{1}{1 - \eta} \eta^{\beta(k+1)} F(1) \leq Cr^\beta F(1)$$

provided  $\beta$  is chosen small so that  $\eta^\beta \geq (1 - \eta)$  and  $C \geq \frac{1}{1 - \eta}$ .  $\square$

**2.2. Allard's theorem.** We now prove the following

**Theorem 2.8** (Improvement of flatness). *For every  $\alpha \in (0, 1)$ , there are positive constants  $c, \varepsilon_0, \eta$  and  $C$  with the following property. Let  $M \in \mathcal{M}_m^\infty(B_R(x_0))$  be such that*

- $x_0 \in \text{supp } M$ ;
- $\Theta^m(M, x) \geq 1$  for  $M$ -almost every  $x$ ;
- for every  $x \in B_{R/2}(x_0)$  and  $0 < r \leq R/2$  it holds  $M(B_r(x)) \leq \frac{3}{2}\omega_m r^m$ ;

and, for some  $\varepsilon \leq \varepsilon_0$  and  $S \in \text{Gr}(m, d)$ ,

$$\text{osc}_S(M, B_R(x_0)) \leq \varepsilon R \quad \text{and} \quad \|H_M\|_{L^\infty(M)} \leq c\varepsilon R^{-1}.$$

Then there is  $T \in \text{Gr}(m, d)$  with  $|S - T| \leq C\varepsilon$  such that

$$(2.10) \quad \text{osc}_T(M, B_{\eta R}(x_0)) \leq \eta^{1+\alpha} \varepsilon R.$$

Before proceeding with the proof, we introduce the following notation. To every  $M \in \mathcal{M}_m(U)$  we associate a rectifiable varifold  $V_M$  such that  $\int \varphi(x, S) dV_M(x, S) = \int \varphi(x, T_x M) dM(x)$  for every  $\varphi \in C_c^0(U \times \text{Gr}(m, d))$ . If  $M \in \mathcal{M}_m^p(U)$  and  $H_M = 0$ , we say that  $V_M$  is *stationary*.

**Proof.** By rescaling and translating, we assume  $R = 1$  and  $x_0 = 0$ . We argue by contradiction and compactness. Assume there are sequences  $\varepsilon_j \searrow 0, c_j \searrow 0$  and  $\{M^j\} \subset \mathcal{M}_m^\infty(B_1)$  such that the assumption of the theorem are satisfied with  $\varepsilon_0$  replaced by  $\varepsilon_j$  and  $c$  replaced by  $c_j$ , for which however (2.10) fails for any choice of  $\eta$  and  $T$ .

Before proceeding, we remark that, by compactness (see, for instance, [15, Theorem 42.7]), there is  $M \in \mathcal{M}_m(B_1)$  such that  $0 \in \text{supp } M$ ,  $\Theta^m(M, x) \geq 1$  for  $M$ -a.e.  $x$ ,  $M(B_1 \setminus S) = 0$ ,  $V_M$  is stationary and  $V_{M^j} \rightarrow V_M$ . Therefore  $M = \Theta^m(M, x) \mathcal{H}^m \llcorner S$  and  $S \ni x \mapsto \Theta^m(M, x)$  has vanishing distributional gradient, thus  $M = \Theta_0 \mathcal{H}^m \llcorner S$  for some  $\Theta_0 \geq 1$ .

Let now  $\Sigma^j = \text{supp } M^j$ ; for every  $x \in \mathbb{R}^d$ , define  $F_j(x) = (Sx, \varepsilon_j^{-1} S^\perp x) \in \mathbb{R}^d$  and let

$$\tilde{\Sigma}^j = F_j(\Sigma^j) \subset B_1^S \times \overline{B_1^{S^\perp}}.$$

Notice that  $\tilde{\Sigma}^j$  are relatively closed subsets of  $B_1^S \times \overline{B_1^{S^\perp}}$  and that they are non-empty, since by assumption  $0 \in \tilde{\Sigma}^j$  for every  $j$ . Therefore there is a relatively closed set  $\tilde{\Sigma}$  in  $B_1^S \times \overline{B_1^{S^\perp}}$  such that, up to subsequences,  $\tilde{\Sigma}^j$  converges to  $\tilde{\Sigma}$  in the Hausdorff distance.

For every  $x' \in B_1^S$ , we let

$$\mathbf{u}(x') = \{y \in S^\perp : (x', y) \in \tilde{\Sigma}\}.$$

First, we show that  $\mathbf{u}(x') \neq \emptyset$  for every  $x' \in B_1^S$ . Indeed, otherwise, there would be  $r > 0$  such that for every  $j$  large enough

$$\tilde{\Sigma}^j \cap (B_r^S(x') \times \overline{B_1^{S^\perp}}) = \emptyset,$$

hence, by compactness and lower semicontinuity of the mass,

$$0 = \liminf_j M^j(B_r^S(x') \times \overline{B_1^{S^\perp}}) \geq M(B_r^S(x') \times \overline{B_1^{S^\perp}}) \geq \omega_m r^m,$$

which is false.

Secondly, we prove that  $\mathbf{u}(x')$  is a singleton for every  $x' \in B_{1/2}^S$ . Indeed, by Proposition 2.7, for every  $j$ , every  $x \in \Sigma^j \cap B_{1/2}$  and every  $r \in (C\varepsilon_j, \frac{1}{2})$ , it holds

$$\text{osc}_S(M^j, B_r(x)) \leq C\varepsilon_j r^\beta.$$

Therefore, by Hausdorff convergence, for every  $x, y \in \tilde{\Sigma} \cap (B_{1/2}^S \times \overline{B_1^{S^\perp}})$  it holds

$$|S^\perp(x - y)| \leq C|S(x - y)|^\beta.$$

In particular,  $\mathbf{u}(x')$  is a singleton; for the rest of the proof, we denote by  $u(x')$  the only element of  $\mathbf{u}(x')$ . Note further that this implies that  $u \in C^{0, \beta}(B_{1/2}^S)$ . By Lemma 2.9 below,  $u$  is harmonic in  $B_{1/4}^S$ . In particular, since  $\sup_{B_{1/4}^S} |u| \leq 1$ , classical elliptic estimates yield

$$\sup_{B_{1/8}^S} (|\nabla u| + |D^2 u|) \leq C$$

for some  $C$  universal. Since  $u(0) = 0$ , we may choose  $\eta$  small depending only on  $C$  and  $\alpha$  so that

$$|u(x') - u(0) - \nabla u(0)x'| \leq C|x'|^2 \leq \frac{\eta^{1+\alpha}}{2}$$

for every  $x' \in B_{2\eta}^S$ . However, this implies that, for every  $j$  large enough and every  $x \in \Sigma^j \cap B_\eta$ , it holds

$$|S^\perp x - \varepsilon_j \nabla u(0)(Sx)| \leq \varepsilon_j \eta^{1+\alpha},$$

which contradicts the assumptions made at the beginning of the proof.  $\square$

**Lemma 2.9.**  *$u$  defined in the proof of Theorem 2.8 is harmonic.*

**Proof.** We argue as in [14, Lemma 2.4]. Let  $h : B_{1/4}^S \rightarrow S^\perp$  be the harmonic function such that  $(h - u)|_{\partial B_{1/4}^S} = 0$ . If  $u \neq h$ , then there is  $0 < \delta < 1/2$  small such that, for all  $j$  large enough, the function

$$G(x) = \frac{1}{2} \left| \frac{S^\perp x}{\varepsilon_j} - h(Sx) \right|^2 + \frac{\delta}{2} |Sx|^2$$

is such that  $G|_{\Sigma^j}$  achieves its maximum at some point  $x_1$  with  $|Sx_1| \leq \frac{1}{4} - \delta$ . We claim that  $j$  can be chosen so large that, for every  $T \in \text{Gr}(m, d)$ , it holds

$$(2.11) \quad \text{div}_T \nabla G(x_1) > \|H_{M^j}\|_{L^\infty} |\nabla G(x_1)|,$$

which would contradict Proposition A.1. In the rest of the proof,  $C$  denotes constants (whose value may change from one expression to the other) which depend only on  $d, m$  and  $\delta$ , but they are independent of  $j$ . We start by noticing that, by standard elliptic estimates,

$$\max\{|\nabla h(Sx_1)|, |D^2 h(Sx_1)|\} \leq C.$$

Therefore  $|\nabla G(x_1)| \leq \frac{C}{\varepsilon_j}$  for  $j$  large enough and  $\|H_{M^j}\|_{L^\infty} |\nabla G(x_1)| \leq C\varepsilon_j \rightarrow 0$  as  $j \rightarrow \infty$ . Thus, in order to prove (2.11), it is sufficient to prove that for  $j$  sufficiently large

$$(2.12) \quad \inf_{T \in \text{Gr}(m, d)} \text{div}_T \nabla G(x_1) \geq \delta.$$

Let  $f(x) = \frac{1}{\varepsilon_j} S^\perp x - h(Sx)$  and let  $\mathbb{1}_S$  denote the orthogonal projection onto  $S$ . Then  $D^2 G(x_1) = D^2 f(x_1) \cdot f(x_1) + \nabla f(x_1) (\nabla f(x_1))^\top + \delta \mathbb{1}_S \geq D^2 f(x_1) \cdot f(x_1) + \delta \mathbb{1}_S =: A$ . In particular, since  $\Delta h = 0$ ,

$$\text{div}_S \nabla G(x_1) \geq \text{trace}_S A = \Delta h(Sx_1) \cdot \left( \frac{S^\perp x_1}{\varepsilon_j} - h(Sx_1) \right) + m\delta = m\delta,$$

where  $\text{trace}_S A = \sum_{i=1}^m A\xi_i \cdot \xi_i$  for any orthonormal basis  $\{\xi_i\}_{i=1}^m$  of  $S$ . Since  $|A| \leq C$ , by continuity it holds  $\text{div}_T \nabla G(x_1) \geq \delta$  for every  $T \in \text{Gr}(m, d)$  such that  $|T - S| \leq \gamma$  for  $\gamma > 0$  sufficiently small. This proves (2.12) in the case  $|T - S| \leq \gamma$ .

On the other hand, if  $|T - S| \geq \gamma$ , then there is a unit vector  $\eta \in T$  such that  $|S^\perp \eta| \geq c_0 \gamma$  for some  $c_0 > 0$  depending only on the dimension. Then

$$\text{div}_T \nabla G(x_1) \geq \text{trace}_T A + |(\nabla f(x_1))^\top \eta|^2.$$

We have  $|\text{trace}_T A| \leq |A| \leq C$  and

$$\left| (\nabla f(x_1))^\top \eta \right| = \left| -\nabla h(Sx_1) S\eta + \frac{1}{\varepsilon_j} S^\perp \eta \right| \geq \frac{1}{\varepsilon_j} |S^\perp \eta| - |\nabla h(Sx_1)| \geq \frac{c_0 \gamma}{\varepsilon_j} - C,$$

thus  $\text{div}_T \nabla G(x_1) \geq \left( \frac{c_0 \gamma}{\varepsilon_j} - C \right)^2 - C \geq \delta$  for  $j$  large enough. This proves (2.12) in the case  $|T - S| \geq \gamma$ .  $\square$

Notice that the conclusion of Theorem 2.8 may be iterated at all scales. In particular, if we let

$$E(M, B_r(x)) = \inf_{S \in \text{Gr}(m, d)} \frac{\text{osc}_S(M, B_r(x))}{r} + C \|H\|_{L^\infty(M \setminus B_r)} r$$

for some  $C$  large, then Theorem 2.8 yields the existence of some universal constants  $\eta$  and  $\varepsilon_0$  such that

$$E(M, B_{\eta R}) \leq \eta^\alpha E(M, B_R)$$

provided  $E(M, B_R) \leq \varepsilon_0$ . A straight-forward induction argument yields, for every  $r > 0$ ,

$$(2.13) \quad \inf_{S \in \text{Gr}(m, d)} \frac{\text{osc}_S(M, B_r)}{r} \leq E(M, B_r) \leq C \left(\frac{r}{R}\right)^\alpha E(M, B_R)$$

for some  $C$  universal, provided  $E(M, B_R) \leq \varepsilon_0$ .

We finally prove Theorem 2.1.

**Proof of Theorem 2.1.** Assume the result is false. Then there exist a sequence  $\delta_j \searrow 0$  and a sequence  $\{M^j\} \subset \mathcal{M}_m^\infty(B_1)$  that satisfy the assumption of Theorem 2.1 with  $\delta_0$  replaced by  $\delta_j$ , for which the conclusion of the theorem does not hold.

We claim that, for  $j$  large enough, the assumptions of Theorem 2.8 are in place for  $R = \frac{1}{2}$  and any  $x_0 \in \text{supp } M^j \cap B_{1/2}$ . Then, the fact that  $M^j$  can be written as the graph of a  $C^{1, \alpha}$  function follows from (2.13) above.

In the direction of proving the above claim, we first remark that, up to subsequences, there exists  $M \in \mathcal{M}_m(B_1)$  such that  $\Theta^m(M, x) \geq 1$   $M$ -almost everywhere,  $V_M$  is stationary and  $M(B_1) \leq \omega_m$ . By [2, Theorem 5.3], we have that  $M = \mathcal{H}^m \llcorner S$  for some  $S \in \text{Gr}(m, d)$ .

Given  $\alpha \in (0, 1)$ , let now  $\varepsilon_0$  be the constant given in Theorem 2.8. Then, for  $j$  large enough,

$$\text{supp } M^j \subset \{y : |S^\perp y| \leq \varepsilon_0/2\}.$$

The only thing left to prove is that

$$(2.14) \quad M^j(B_{1/2}(x)) \leq \left(1 + \frac{1}{4}\right) \frac{\omega_m}{2^m}$$

for every  $j$  large and every  $x \in \text{supp } M^j$ . Given the above inequality, by monotonicity ([15, §17]), provided  $\delta_0$  is smaller than some universal constant, we obtain  $M^j(B_r(x)) \leq \frac{3}{2} \omega_m r^m$  for every  $0 \leq r \leq \frac{1}{2}$ , as required in Theorem 2.8.

We now prove (2.14). If the result is false, then there exist a subsequence  $j_k$  and points  $x_k \in \text{supp } M^{j_k} \cap B_{1/2}$  such that

$$(2.15) \quad M^{j_k}(B_{1/2}(x_k)) \geq \left(1 + \frac{1}{4}\right) \frac{\omega_m}{2^m}.$$

Up to extracting a further subsequence,  $x_k \rightarrow x \in \overline{B_{1/2}}$  and, by monotonicity,  $x \in \text{supp } M$ . Therefore, for any  $\varepsilon > 0$ ,

$$\omega_m \left(\frac{1}{2} + \varepsilon\right)^m = M\left(B_{\frac{1}{2} + \varepsilon}(x)\right) \geq \limsup_k M^{j_k}\left(B_{\frac{1}{2} + \varepsilon}(x)\right) \geq \limsup_k M^{j_k}(B_{1/2}(x_k))$$

contradicting (2.15) and thus concluding the proof.  $\square$

### 3. BRAKKE'S REGULARITY THEOREM

In this section, we show how the arguments in the previous section may be adapted to the case of mean curvature flows. Inspired by [11], we give the following definition.

**Definition 3.1** (Brakke flow with transport term). *We say that a family  $\mathbf{M} = \{M_t\}_{t \in [0, \Omega]}$  of Radon measures on  $U \subset \mathbb{R}^d$  is a Brakke flow with transport term  $v$  in  $U \times [0, \Omega]$  if the following hold true:*

- (1) *for almost every  $t \in [0, \Omega]$ ,  $M_t \in \mathcal{M}_m^2(U)$  and, for  $M_t$ -almost every  $x$ ,  $\Theta^m(M_t, x)$  is a positive integer.*

(2) For every  $W \subset\subset U$ ,

$$\int_0^\Omega \int_W (1 + |v| + |H_{M_t}|^2) dM_t dt < \infty$$

(3) For every non-negative test function  $\phi \in C_c^1(U \times [0, \Omega])$  and for every  $[t_1, t_2] \subset [0, \Omega]$ , it holds

$$(3.1) \quad M_{t_2}(\phi(\cdot, t_2)) - M_{t_1}(\phi(\cdot, t_1)) \leq \int_{t_1}^{t_2} \int \left( \partial_t \phi + (-\phi H + \nabla \phi) \cdot (H + v^\perp) \right) dM_t dt,$$

where  $H$  is the generalized mean curvature vector of  $M_t$ , as defined in (2.2), and  $v^\perp(x) = (T_x M_t)^\perp v(x)$  for  $M_t$ -almost every  $x$ .

For a Brakke flow with transport as above, we define the measure  $M$  on  $U \times [0, \Omega]$  by  $\int \varphi(x, t) dM(x, t) = \int \int \varphi(x, t) dM_t(x) dt$  and the space-time track of  $\mathbf{M}$ :

$$\Sigma_{\mathbf{M}} = \text{Clos} \left( \bigcup_{t \in [0, \Omega]} \text{supp } M_t \times \{t\} \right).$$

**Remark 3.2.** By virtue of [4, Chapter 5], it holds

$$(3.2) \quad H_{M_t}(x) \perp T_x M_t$$

at  $M$ -almost every  $(x, t)$ . This fact is used in Proposition 3.5 below.

**3.1. Decay of oscillations.** In the following, we denote by  $Q_R(x_0, t_0)$  the backwards parabolic cylinder  $B_R(x_0) \times [t_0 - R^2, t_0]$ . The following assumptions will be used in the present subsection.

**Assumption 3.3.**

- (1) There is  $E_1 < \infty$  such that, for every  $(x, t) \in Q_R$  and  $B_r(x) \subset B_R$ , it holds  $M_t(B_r(x)) \leq E_1 r^m$ ;
- (2)  $\Lambda := \|v\|_{L^\infty(M)} < \infty$ .

Before proceeding, we give the following definition:

**Definition 3.4** ( $m$ -dimensional backward heat kernel). *Let  $\phi \in C_c^\infty([0, 1])$  be a cut-off function that we fix hereafter such that  $\phi \equiv 1$  in  $[0, 1/2]$ ,  $|\phi'| \leq 2$  and  $0 \leq \phi \leq 1$  everywhere. For  $R > 0$ , we let  $\Psi_R : \mathbb{R}^d \times (-\infty, 0) \rightarrow \mathbb{R}$  be defined as*

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

A direct calculation yields that there is  $C$  universal such that, for every  $S \in \text{Gr}(m, d)$  and for every  $(x, t) \in \mathbb{R}^d \times (-\infty, 0)$ , it holds

$$(3.3) \quad \left| \partial_t \Psi_R + \text{div}_S \nabla \Psi_R + \frac{|S^\perp \nabla \Psi_R|^2}{\Psi_R} \right| \leq C \frac{\chi_{B_R}}{R^{m+2}}.$$

**Proposition 3.5** (Weighted Huisken's monotonicity formula). *For every  $E_1$ , there is  $C > 0$  with the following property. Let  $\mathbf{M}$  and  $v$  satisfy Assumption 3.3, and let  $f : B_R \rightarrow \mathbb{R}$  be a non-negative, convex function with  $\|\nabla f\|_\infty \leq 1$ . Then, for every  $-R^2 \leq t < 0$ , it holds*

$$(3.4) \quad \int f \Psi_R(\cdot, t) dM_t \geq f(0) - C \left( \Lambda + \frac{\varepsilon}{R^2} + \varepsilon \Lambda^2 \right) (-t),$$

where  $\varepsilon := \|f\|_{L^\infty(M)}$ .

**Proof.** Up to rescalings, we may assume  $R = 1$ ; for brevity, we let  $\widehat{\Psi} := \Psi_1$ . By mollification, we may also assume that  $f \in C^2(\overline{B_1})$ . Notice that, in this case,  $\operatorname{div}_S \nabla f \geq 0$  for any  $S \in \operatorname{Gr}(m, d)$ . By (3.3) and the above inequality, we have

$$\begin{aligned} \partial_t(f\widehat{\Psi}) &\leq f \operatorname{div}_S \nabla \widehat{\Psi} - f \frac{|S^\perp \nabla \widehat{\Psi}|^2}{\widehat{\Psi}} + Cf\chi_{B_1} \\ &\leq \operatorname{div}_S(\widehat{\Psi} \nabla f - f \nabla \widehat{\Psi}) - f \frac{|S^\perp \nabla \widehat{\Psi}|^2}{\widehat{\Psi}} + Cf\chi_{B_1}. \end{aligned}$$

In particular, for almost every  $t$ , it holds

$$(3.5) \quad \int \partial_t(f\widehat{\Psi}) \, dM_t \leq \int \left( -H \cdot (\widehat{\Psi} \nabla f - f \nabla \widehat{\Psi}) - f \frac{|(\nabla \widehat{\Psi})^\perp|^2}{\widehat{\Psi}} \right) \, dM_t + CE_1 \varepsilon,$$

where we have used the facts that  $M_t \in \mathcal{M}_m^2$  for a.e.  $t$  and that  $M_t(B_1) \leq E_1$ .

Let now  $I(s) = \int f\widehat{\Psi}(\cdot, s) \, dM_s$ . We use  $\phi = f\widehat{\Psi}$  as a test function in Definition 3.1. By (3.5), we obtain, for  $-1 \leq t \leq s < 0$ :

$$\begin{aligned} I(s) - I(t) &\leq \int_t^s \int \left( -f \frac{|(\nabla \widehat{\Psi})^\perp|^2}{\widehat{\Psi}} + 2fH \cdot \nabla \widehat{\Psi} - f\widehat{\Psi}|H|^2 \right) \, dM_\tau \, d\tau \\ &\quad + \int_t^s \int \left( v^\perp \cdot (-f\widehat{\Psi}H + \nabla(f\widehat{\Psi})) \right) \, dM_\tau \, d\tau + CE_1 \varepsilon(s-t) \\ &= \int_t^s \int f\widehat{\Psi} \left( - \left| \frac{(\nabla \widehat{\Psi})^\perp}{\widehat{\Psi}} - H \right|^2 + v^\perp \cdot \left( \frac{(\nabla \widehat{\Psi})^\perp}{\widehat{\Psi}} - H \right) \right) \, dM_\tau \, d\tau \\ &\quad + \int_t^s \int \widehat{\Psi} \nabla f \cdot v^\perp \, dM_\tau \, d\tau + CE_1 \varepsilon(s-t), \end{aligned}$$

where (3.2) was used in the above equality. We then use Young's inequality to bound  $v^\perp \cdot \left( \frac{(\nabla \widehat{\Psi})^\perp}{\widehat{\Psi}} - H \right) \leq |v|^2 + \left| \frac{(\nabla \widehat{\Psi})^\perp}{\widehat{\Psi}} - H \right|^2$  and obtain

$$(3.6) \quad I(s) - I(t) \leq \int_t^s \int \left( f\widehat{\Psi}|v|^2 + \widehat{\Psi}|\nabla f||v| \right) \, dM_\tau \, d\tau + CE_1 \varepsilon(s-t).$$

Next, we estimate

$$\int f\widehat{\Psi}|v|^2 \, dM_\tau \leq \Lambda^2 \int f\widehat{\Psi} \, dM_\tau \quad \text{and} \quad \int \widehat{\Psi}|\nabla f||v| \, dM_\tau \leq \Lambda \int \widehat{\Psi} \, dM_\tau \leq C\Lambda E_1,$$

where the latter inequality follows from Item 1 in Assumption 3.3. Therefore, going back to (3.6), dividing both sides by  $s-t$  and letting  $s \searrow t$ , we obtain

$$I'(t) \leq \Lambda^2 I(t) + C(\Lambda + \varepsilon)$$

in the sense of distributions, for some  $C$  depending also on  $E_1$ . Hence for every  $-1 \leq t \leq s < 0$  it

$$I(t) \geq I(s)e^{-\Lambda^2(s-t)} - C(\Lambda + \varepsilon) \int_t^s e^{\Lambda^2(t-\tau)} \, d\tau \geq I(s) - C\Lambda^2 \varepsilon(-t) - C(\Lambda + \varepsilon)(-t)$$

where in the second inequality we used the inequality  $e^t \geq 1+t$  and the fact that  $I(s) \leq CE_1 \varepsilon$ . We conclude the proof by remarking that

$$\lim_{s \searrow 0} \int f(\cdot) \Psi_R(\cdot - x, s) \, dM_{t+s} \geq f(x)$$

at  $M$ -almost every  $(x, t)$ .  $\square$

Similarly to what we did in Section 2, for  $S \in \text{Gr}(m, d)$  and  $X \in \Sigma_{\mathbf{M}}$  we define the quantity

$$\text{osc}_S(\mathbf{M}, Q_r(X)) = \frac{1}{2} \sup \{ |S^\perp(x-y)| : (x, t), (y, s) \in \Sigma_{\mathbf{M}} \cap Q_r(X) \}.$$

**Proposition 3.6** (Harnack inequality). *For every  $E_1$ , there exist  $\eta \in (0, 1)$  with the following property. Let  $\mathbf{M}$  and  $v$  satisfy Assumption 3.3 with  $(0, 0) \in \Sigma_{\mathbf{M}}$ . Moreover, assume that, for every  $t \in [-R^2, 0)$ ,*

$$\int \Psi_R(\cdot, t) \, dM_t \leq \frac{3}{2},$$

that  $\text{osc}(Q_R) \leq \eta R$  and that  $\Lambda R^2 \leq \text{osc}(Q_R)$ . Then

$$\text{osc}(Q_{\eta R}) \leq (1 - \eta) \text{osc}(Q_R).$$

**Proof.** Let  $\varepsilon = \frac{1}{R} \text{osc}(Q_R)$ . We can thus choose  $y_0 \in B_{\varepsilon R}$  such that  $\Sigma_{\mathbf{M}} \cap B_R \subset \{y : |S^\perp(y - y_0)| \leq \varepsilon R\}$ . Assume by contradiction that there are two points  $Y_i = (y_i, s_i) \in \Sigma \cap Q_{\eta R}$ ,  $i = 1, 2$ , such that  $|S^\perp(y_1 - y_2)| \geq 2(1 - \eta)\varepsilon R$ . Denote  $\omega := \frac{S^\perp(y_1 - y_2)}{|S^\perp(y_1 - y_2)|}$  and consider

$$f_1(x) = \left( (x - y_0) \cdot \omega - \frac{\varepsilon R}{2} \right)^+ \quad \text{and} \quad f_2(x) = \left( -(x - y_0) \cdot \omega - \frac{\varepsilon R}{2} \right)^+.$$

Consider  $\eta \leq \theta \leq 1$  and let  $T = -\theta R^2$ . Since the  $f_i$  are convex and  $\|\nabla f_i\|_\infty \leq 1$ , by Proposition 3.5

$$\begin{aligned} \int f_i \Psi_{R/2}(\cdot - y_i, T - s_i) \, dM_T &\geq f_i(y_i) - C \left( \Lambda + \frac{\|f_i\|_\infty}{R^2} + \|f_i\|_\infty \Lambda^2 \right) (s_i - T) \\ &\geq f_i(y_i) - 3C_0 \varepsilon \theta R, \end{aligned}$$

where we have used that  $\|f_i\|_\infty \leq \varepsilon R$  and  $\Lambda \leq \varepsilon R^{-1}$ . By choosing  $\theta$  small so that  $3C_0 \theta \leq \frac{1}{10}$  and remarking that  $f_i(y_i) \geq (\frac{1}{2} - 2\eta)\varepsilon R$ , we obtain

$$(3.7) \quad \int f_i \Psi_{R/2}(\cdot - y_i, T - s_i) \, dM_T \geq \left( \frac{2}{5} - 2\eta \right) \varepsilon R.$$

Next, we bound from above the left-hand side of (3.7): provided  $\eta^2$  is much smaller than  $\theta$  chosen above, it holds  $T - s_i \leq -\frac{\theta}{2} R^2$ , hence there is some constant  $L$  such that

$$|\Psi_{R/2}(x - y_i, T - s_i) - \Psi_{R/2}(x, T)| \leq L (R^{-m-1}|y_i| + R^{-m-2}|s_i|) \leq L\eta R^{-m}$$

for every  $x \in \mathbb{R}^d$ . Therefore

$$(3.8) \quad \begin{aligned} \int f_i \Psi_{R/2}(\cdot - y_i, T - s_i) \, dM_T &\leq \frac{\varepsilon R}{2} \int_{\text{supp } f_i} \Psi_{R/2}(\cdot - y_i, T - s_i) \, dM_T \\ &\leq \frac{\varepsilon R}{2} \left( \int_{\text{supp } f_i} \Psi_{R/2}(\cdot, T) \, dM_T + LE_1 \eta \right) \end{aligned}$$

where we have used the fact that  $\frac{M_T(B_R)}{R^m} \leq E_1$ .

If  $\eta$  is smaller than some universal constant, then  $\text{supp } f_1 \cap \text{supp } f_2 = \emptyset$ , thus summing (3.7) and (3.8) for  $i = 1, 2$ , we obtain

$$\begin{aligned} \int \Psi_{R/2}(\cdot, T) \, dM_T &\geq \left( \int_{\text{supp } f_1} \Psi_{R/2}(\cdot, T) \, dM_T + \int_{\text{supp } f_2} \Psi_{R/2}(\cdot, T) \, dM_T \right) \\ &\geq 2 \frac{2}{\varepsilon R} \left( \frac{2}{5} - 2\eta \right) \varepsilon R - 2LE_1 \eta \\ &= \frac{8}{5} - C\eta \end{aligned}$$

for some  $C$  depending on  $E_1$ . Choosing  $\eta$  smaller, if needed, contradicts the assumption that  $\int \Psi_{R/2}(\cdot, T) dM_T \leq \frac{3}{2}$ , thus concluding the proof.  $\square$

As Proposition 2.6 implies Proposition 2.7, we obtain the following result as a corollary of Proposition 3.6.

**Proposition 3.7** (Decay of oscillations). *For every  $E_1$ , there exist  $C$  and  $\beta$  with the following property. Let  $\mathbf{M}$  be a Brakke flow with transport term  $v$  in  $Q_R$  that satisfies Assumption 3.3 with  $(0, 0) \in \Sigma_{\mathbf{M}}$  and such that, for every  $t \in [-R^2, 0)$ ,*

$$\int \Psi_R(\cdot, t) dM_t \leq \frac{3}{2}.$$

*Then, for every  $r \geq C(\text{osc}(Q_R) + \Lambda R^2)$ , it holds*

$$\text{osc}(Q_r) \leq C(\text{osc}(Q_R) + \Lambda R^2) \left(\frac{r}{R}\right)^\beta.$$

**3.2. Brakke's theorem.** The present subsection is dedicated to the proof of the following version of Brakke's regularity theorem:

**Theorem 3.8** (Interior regularity). *For every  $\alpha \in (0, 1)$  and every  $E_1$ , there are positive and small constants  $\theta$  and  $\delta_0$  with the following property. Let  $\mathbf{M}$  be a Brakke flow with transport term  $v$  in  $Q_1$ . Assume that:*

- $(0, 0) \in \Sigma_{\mathbf{M}}$ ;
- for every  $t \in [-1, 0]$  and every  $B_r(x) \subset B_1$ ,  $M_t(B_r) \leq E_1 r^m$ ;
- $\int_{B_1} \Psi(\cdot, -\theta^2) dM_{-\theta^2} \leq 1 + \delta_0$ ;
- $\|v\|_{L^\infty(M)} \leq \delta_0$ .

*Then  $\Sigma_{\mathbf{M}} \cap Q_{\theta/8}$  is included in the graph of some function  $u \in C^{1,\alpha}(Q_{\theta/8}^m; \mathbb{R}^{d-m})$ .*

In the above statement, by  $u \in C^{1,\alpha}(Q_{\theta/8}^m; \mathbb{R}^{d-m})$  we mean that  $(x, t) \mapsto u(x, t)$  is differentiable with respect to  $x$  and that there is  $C > 0$  such that, for every  $(x', t), (y', s) \in Q_{\theta/8}^m$  it holds

$$|u(x, t) - u(y, s) - \nabla u(x, t) \cdot (y - x)| \leq C(|x - y|^2 + |t - s|)^{\frac{1+\alpha}{2}}.$$

**Remark 3.9.** This result is analogous to the “end-time regularity” proved in [16], although our proof requires the forcing term  $v$  to be a  $L^\infty$  function. We also remark that, differently from the case of minimal varifolds, higher regularity on  $v$  does not straightforwardly yield higher regularity for  $\Sigma_{\mathbf{M}}$ . In this regard, see [17].

Similarly to Theorem 2.1, Theorem 3.8 follows from an *improvement of flatness*, which we state and prove next. We first make some preliminary assumptions:

**Assumption 3.10.**

*For some  $X_0 = (x_0, t_0)$  and  $R > 0$ :*

- (1)  $\mathbf{M}$  is a Brakke flow with transport term  $v$  in  $Q_R(X_0)$ .
- (2)  $X_0 \in \Sigma_{\mathbf{M}}$ .
- (3) For every  $(x, t) \in Q_{R/2}(X_0)$  and every  $-\frac{R^2}{4} \leq \tau < 0$ , it holds

$$\int_{B_{R/2}(x)} \Psi(\cdot - x, \tau) dM_t \leq \frac{3}{2}.$$

**Theorem 3.11** (Improvement of flatness). *For every  $\alpha \in (0, 1)$ , there are universal constants  $\varepsilon_0, \eta$  (small) and  $C$  (large) with the following property. Let  $\mathbf{M}$  satisfy Assumption 3.10. Assume, in addition, that for some  $\varepsilon \leq \varepsilon_0$  and  $S \in \text{Gr}(m, d)$ :*

$$(3.9) \quad \text{osc}_S(\mathbf{M}, Q_R(X_0)) \leq \varepsilon R \quad \text{and} \quad C\|v\|_{L^\infty(M)} \leq \frac{\varepsilon}{R}$$

Then there exists  $T \in \text{Gr}(m, d)$  with  $|T - S| \leq C\varepsilon$  such that

$$(3.10) \quad \text{osc}_T(\mathbf{M}, Q_{\eta R}(X_0)) \leq \eta^{1+\alpha} \varepsilon R.$$

**Proof.** By rescaling and translating, we may assume  $R = 1$  and  $X_0 = (0, 0)$ . We argue by contradiction and compactness: assume there exist a sequence  $\varepsilon_j \searrow 0$  and two sequences  $\{\mathbf{M}^j\}$  and  $\{v^j\}$  such that  $\mathbf{M}^j$  is a Brakke flow with transport  $v^j$  for which the assumptions of the theorem are satisfied with  $\varepsilon$  replaced by  $\varepsilon_j$ . For brevity, we let  $\Sigma^j := \Sigma_{\mathbf{M}^j}$ .

**Step 1: compactness and convergence to a plane.** By Proposition 3.12 below, up to extracting a subsequence (which we do not relabel),  $\{\mathbf{M}^j\}$  converges to a Brakke flow without transport term  $\mathbf{M}^\infty$ . Moreover,  $M_t^\infty(S^c) = 0$  for every  $t$ , hence  $M_t^\infty = \Theta^m(M_t^\infty, \cdot) \mathcal{H}^m \llcorner S$ . Now, for almost every  $t$ ,  $M_t^\infty \in \mathcal{M}_m^2$  and, by Item 3 in Assumption 3.10 and by Huisken's monotonicity,  $\Theta^m(M_t^\infty, x) \leq E_1$  for every  $x$ . In particular, by (2.1), for every  $\varphi \in C_c^\infty(S; S)$  it holds

$$\int_S \Theta^m \text{div} \varphi = \int \text{div}_S \tilde{\varphi} dM_t^\infty \leq C \left( \int |\varphi|^2 dM_t^\infty \right)^{1/2} \leq C \left( E_1 \int_S |\varphi|^2 \right)^{1/2},$$

where  $\tilde{\varphi} \in C_c^\infty(\mathbb{R}^d; \mathbb{R}^d)$  is any extension of  $\varphi$  such that  $\|\tilde{\varphi}\|_\infty \leq \|\varphi\|_\infty$ . Hence  $S \ni x \mapsto \Theta^m(M_t^\infty, x)$  is a locally  $W^{1,2}$  function and, by Definition 3.1, it is integer-valued. Hence, for almost every  $t$ , either  $\Theta^m(M_t^\infty, \cdot) \equiv 0$  or  $\Theta^m(M_t^\infty, \cdot) \geq 1$  for  $\mathcal{H}^m$ -almost every  $x$ . However,  $(0, 0) \in \Sigma_{\mathbf{M}^\infty}$  hence, by (3.1), it cannot be that  $M_t^\infty = 0$  for some  $t < 0$ . Therefore  $M_t^\infty \geq \mathcal{H}^m \llcorner S$  for almost every  $t \in [-1, 0]$ .

**Step 2: Hausdorff convergence.** Let now

$$\tilde{\Sigma}^j = \left\{ \left( Sx, \frac{1}{\varepsilon_j} S^\perp x, t \right) : (x, t) \in \Sigma^j \right\} \subset B_1^S \times B_1^{S^\perp} \times [-1, 0].$$

By Item 2 in Assumption 3.10,  $\tilde{\Sigma}^j \neq \emptyset$ . Therefore, up to subsequences,  $\{\tilde{\Sigma}^j\}$  converges in the Hausdorff distance to some relatively closed set  $\tilde{\Sigma} \subset B_1^S \times B_1^{S^\perp} \times [-1, 0]$ .

**Step 3:  $\tilde{\Sigma}$  is a graph.** Let

$$\mathbf{u}(x', t) = \{y \in B_1^{S^\perp} : (x', y, t) \in \tilde{\Sigma}\}.$$

Arguing as in the proof of Theorem 2.8, by Step 1 we first show that  $\mathbf{u}(x', t) \neq \emptyset$  for every  $(x', t) \in Q_1^S$ . Next, by Proposition 3.7, we conclude that  $\mathbf{u}(x', t) = \{u(x', t)\}$  is a singleton for every  $(x', t) \in Q_{1/2}^S$  and that  $u$  is Hölder continuous.

**Step 4: conclusion.** Arguing as in Lemma 2.9, Proposition A.2 below yields that  $u|_{Q_{1/4}^S}$  is a solution to the heat equation. The desired result follows by the Hausdorff convergence established in Step 2 and by classical Schauder estimates for the heat equation.  $\square$

The following result was used in the proof of Theorem 3.11.

**Proposition 3.12.** *Let  $\{\mathbf{M}^j\}, \{v_j\}$  be two sequences such that, for each  $j$ ,  $\mathbf{M}^j$  is a Brakke flow with transport term  $v_j$  in  $Q_R$ . Assume that, for every  $W \subset\subset B_R$ ,*

$$\sup_{t \in [-R^2, 0]} \sup_{j \in \mathbb{N}} M_t(W) < \infty$$

and that

$$\|v_j\|_{L^\infty(M^j)} \rightarrow 0.$$

Then there exist a subsequence  $\{j_\ell\} \subset \mathbb{N}$  and a Brakke flow  $\mathbf{M}$  without transport term in  $Q_R$  such that, for every  $t \in [-R^2, 0]$ ,  $M_{t_\ell}^{j_\ell} \rightharpoonup M_t$  as Radon measures in  $B_R$ .

**Proof.** The proof, in the case of Brakke flows without transport term, can be found in [10, §7]: the more general case can be proved by straightforward modifications. See also [9, Section 4.2.2].  $\square$

We conclude the present section by sketching the proof of Theorem 3.8.

**Proof of Theorem 3.8.** Argue by contradiction. For  $E_1$  and  $\alpha$  fixed, assume there is a sequence  $\{\mathbf{M}^j\}$  of Brakke flows with transport term  $v_j$  that satisfy the assumptions of Theorem 3.8 with  $\delta_0$  and  $\theta$  replaced by some  $\delta_j \searrow 0$  and  $\theta_j \searrow 0$ , respectively. By Proposition 3.12 above, the rescalings

$$\widetilde{\mathbf{M}}^j = \left\{ \widetilde{M}_t^j = \theta_j^{-m} (\mu_{\theta_j})_{\#} M_{\theta_j^2 t}^j \right\}_{t \in [-\theta_j^{-2}, 0]}$$

(where  $\mu_r(y) = \frac{y}{r}$ ) converge to a Brakke flow  $\widetilde{\mathbf{M}}$  without transport term in  $\mathbb{R}^d \times (-\infty, 0]$ . By Huisken's monotonicity formula,  $\widetilde{\mathbf{M}}$  must be a stationary unit-density plane. Therefore, for  $j$  large enough, the assumptions of Theorem 3.11 are in place for every  $X_0 \in \Sigma_{\mathbf{M}^j} \cap Q_{\theta_j/2}$  and  $R = \theta_j/2$ .  $\square$

#### APPENDIX A. MAXIMUM PRINCIPLES

For the sake of completeness, we prove two maximum principles which were used in the proofs of Theorems 2.8 and 3.11. The proofs of Proposition A.1 and Proposition A.2 are inspired by [20] and [19, Theorem 14.1], respectively. In what follows, for a symmetric matrix  $A$ , we let  $\text{trace}_m A$  denote the sum of the  $m$  smallest eigenvalues of  $A$ .

**Proposition A.1.** *Let  $M \in \mathcal{M}_m^\infty(U)$  be such that  $\Lambda := \|H\|_{L^\infty(M)} < \infty$ . If  $f \in C^2(U)$  is such that  $f|_{\text{supp } M}$  achieves a local maximum at  $x_0 \in \text{supp } M \cap \text{Int } U$ , then*

$$\text{trace}_m D^2 f(x_0) \leq \Lambda |\nabla f(x_0)|.$$

**Proof.** Without loss of generality, we assume that  $x_0 = 0$  and that  $f$  has a global strict maximum at 0. By contradiction, assume that there is  $r > 0$  such that

$$\text{trace}_m D^2 f(x) - \Lambda |\nabla f(x)| \geq \delta$$

for every  $x \in B_r$ . Up to choosing  $r$  smaller and up to adding a constant to  $f$ , we also assume that  $f(0) > 0$  and that  $\{f > 0\} \cap \text{supp } M \subset B_r$ . By mollification, (2.2) and the assumption  $\|H_M\|_\infty \leq \Lambda$  give

$$\begin{aligned} \Lambda \int_{\{f>0\}} f |\nabla f| \, dM &\geq \int_{\{f>0\}} \text{div}_{T_x M} (f \nabla f) \, dM \\ &= \int_{\{f>0\}} \left( |\nabla_{T_x M} f(x)|^2 + f(x) \text{div}_{T_x M} \nabla f(x) \right) \, dM(x) \\ &\geq \int_{\{f>0\}} f \text{div}_{T_x M} \nabla f \, dM \end{aligned}$$

where  $\nabla_{T_x M} f(x)$  is the projection of  $\nabla f(x)$  onto  $T_x M$ . This yields

$$\delta \int_{\{f>0\}} f \, dM \leq \int_{\{f>0\}} f (\text{div}_{T_x M} \nabla f - \Lambda |\nabla f|) \, dM \leq 0,$$

which contradicts the fact that  $0 \in \text{supp } M$ .  $\square$

**Proposition A.2.** *Let  $\mathbf{M}$  be a Brakke flow with transport term  $v$  in  $U \times [-\Omega, 0]$  and assume that  $\Lambda := \|v\|_{L^\infty(M)} < \infty$ . If  $f \in C^2(U \times [-\Omega, 0])$  is such that  $f|_{\Sigma_{\mathbf{M}}}$  has a local maximum at  $X_0 = (x_0, t_0) \in \text{Int } U \times (-\Omega, 0]$ , then*

$$\text{trace}_m D^2 f(X_0) - \partial_t f(X_0) \leq \Lambda |\nabla f(X_0)|.$$

**Proof.** Up to a translation, we assume that  $(x_0, t_0) = (0, 0)$ . Moreover, we assume that  $f|_{\Sigma \cap \{t \leq 0\}}$  has a strict local maximum at  $(0, 0)$ . Assume the result does not hold. In particular, since  $f \in C^2$ , we can choose  $\rho > 0$  and  $\varepsilon > 0$  small so that:

- (1)  $\partial_t f - \text{trace}_m D^2 f + \Lambda |\nabla f| < -\varepsilon < 0$  in  $Q_\rho$ ;
- (2)  $0 < f(0, 0) \leq \frac{2\varepsilon}{\Lambda^2}$  and  $f < 0$  in  $\Sigma \cap \{t \leq 0\} \setminus Q_\rho$ ;

where the second point holds up to adding a constant to  $f$ . We now let  $\varphi(x, t) = (f^+(x, t))^4$ , where  $f^+ = \max\{f, 0\}$ , and we use  $\varphi$  as a test function for (3.1). Since  $\varphi(\cdot, -\rho^2) = 0$  by assumption, we have

$$(A.1) \quad \begin{aligned} 0 &\leq \int \varphi(\cdot, 0) dM_0 - \int \varphi(\cdot, -\rho^2) dM_{-\rho^2} \\ &\leq \int_{-\rho^2}^0 \int (\partial_t \varphi + (-\varphi H + \nabla \varphi) \cdot (H + v^\perp)) dM_t dt. \end{aligned}$$

We now have

$$\int H \cdot \nabla \varphi dM_t = - \int \text{div}_{T_x M_t} \nabla \varphi dM_t$$

for a.e.  $t$ , and we may bound  $-\varphi H \cdot v^\perp \leq \varphi |H|^2 + \varphi |v|^2$ . Therefore (A.1) gives

$$(A.2) \quad 0 \leq \int_{-\rho^2}^0 \int (\partial_t \varphi - \text{div}_{T_x M_t} \nabla \varphi + \varphi \Lambda^2 + |\nabla \varphi| \Lambda) dM_t dt.$$

Straightforward computations give

$$|\nabla \varphi| = 4(f^+)^3 |\nabla f|, \quad \text{div}_{T_x M_t} \nabla \varphi = 4(f^+)^3 \text{div}_{T_x M_t} \nabla f \geq 4(f^+)^3 \text{trace}_m D^2 f$$

and  $\partial_t \varphi = 4(f^+)^3 \partial_t f$ . Therefore (A.2) reads

$$(A.3) \quad 0 \leq \int_{-\rho^2}^0 \int 4(f^+)^3 \left( \partial_t f - \text{trace}_m D^2 f + |\nabla f| \Lambda + \Lambda^2 \frac{f^+}{4} \right) dM_t dt.$$

Since  $f$  has a maximum at 0, it holds  $\Lambda^2 \frac{f^+}{4} \leq \Lambda^2 \frac{f(0,0)}{4} \leq \frac{\varepsilon}{2}$ , hence

$$\partial_t f - \text{trace}_m D^2 f + |\nabla f| \Lambda + \Lambda^2 \frac{f^+}{4} \leq -\varepsilon + \frac{\varepsilon}{2} = -\frac{\varepsilon}{2}$$

in  $Q_\rho \cap \{f \geq 0\}$ . The latter inequality and (A.3) yield  $M(Q_\rho) = 0$ , which contradicts the assumption  $(0, 0) \in \Sigma_{\mathbf{M}}$ .  $\square$

## REFERENCES

- [1] Allard's interior regularity theorem: an invitation to stationary varifolds. *Nonlinear analysis in geometry and applied mathematics*, 2:23–49, 2018.
- [2] William K. Allard. On the First Variation of a Varifold. *The Annals of Mathematics*, 95(3):417, 1972.
- [3] Frederick J. Almgren, Jr. Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure. *The Annals of Mathematics*, 87(2):321, 1968.
- [4] Kenneth A. Brakke. *The motion of a surface by its mean curvature*. Mathematical notes; 20. Princeton University Press, Princeton, 1978.
- [5] Luis A Caffarelli. The obstacle problem revisited. *Journal of Fourier Analysis and Applications*, 4(4):383–402, 1998.
- [6] Luis A. Caffarelli and Antonio Córdoba. An elementary regularity theory of minimal surfaces. *Differential and Integral Equations*, 6(1):1–13, 1993.
- [7] Ennio De Giorgi. Frontiere orientate di misura minima. *Seminario di Matematica della Scuola Normale Superiore di Pisa*, 1960–1961:57, 1961.
- [8] Carlo Gasparetto. Epsilon-regularity for the Brakke flow with boundary. 2022.
- [9] Carlo Gasparetto. *Epsilon-regularity for the mean curvature flow with boundary and transport term: a viscosity approach*. PhD thesis, Scuola Internazionale Superiore di Studi Avanzati, Trieste, 2022.

- [10] Tom Ilmanen. Elliptic regularization and partial regularity for motion by mean curvature. *Memoirs of the American Mathematical Society*, 108(520), 1994.
- [11] Kota Kasai and Yoshihiro Tonegawa. A general regularity theory for weak mean curvature flow. *Calculus of Variations and Partial Differential Equations*, 50(1-2):1–68, 2014.
- [12] Ovidiu Savin. Small Perturbation Solutions for Elliptic Equations. *Communications in Partial Differential Equations*, 32(4):557–578, 2007.
- [13] Ovidiu Savin. Regularity of flat level sets in phase transitions. *Annals of Mathematics*, 169(1):41–78, 2009.
- [14] Ovidiu Savin. Viscosity solutions and the minimal surface system. In *Nonlinear Analysis in Geometry and Applied Mathematics, Part 2*, pages 135–145, Somerville, MA, 2018. International Press of Boston, Inc.
- [15] Leon Simon. *Lectures on geometric measure theory*. Number 3 in Proceedings of the Centre for Mathematical Analysis. Centre for Mathematical Analysis, Australian National University, Canberra, 1984.
- [16] Salvatore Stuvard and Yoshihiro Tonegawa. End-time regularity theorem for brakke flows. 2022.
- [17] Yoshihiro Tonegawa. A second derivative Hölder estimate for weak mean curvature flow. *Advances in Calculus of Variations*, 7(1), 2014.
- [18] Yu Wang. Small perturbation solutions for parabolic equations. *Indiana University Mathematics Journal*, 62(2):671–697, 2013.
- [19] Brian White. Mean curvature flow (math 258) lecture notes. Notes by Otis Chodosh, available online at <http://web.stanford.edu/~ochodosh/MCFnotes.pdf>, 2015.
- [20] Brian White. Controlling area blow-up in minimal or bounded mean curvature varieties. *Journal of Differential Geometry*, 102(3), 2016.

*G. De Philippis*: COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY,  
251 MERCER ST., NEW YORK, NY 10012, USA.  
*Email address*: [guido@cims.nyu.edu](mailto:guido@cims.nyu.edu)

*C. Gasparetto*: DIPARTIMENTO DI MATEMATICA, UNIVERSITÀ DI PISA, LARGO BRUNO PONTECORVO  
5, 56127 PISA, ITALY  
*Email address*: [carlo.gasparetto@dm.unipi.it](mailto:carlo.gasparetto@dm.unipi.it)

*F. Schulze*: DEPARTMENT OF MATHEMATICS, ZEEMAN BUILDING, UNIVERSITY OF WARWICK, GIB-  
BET HILL ROAD, COVENTRY CV4 7AL, UK  
*Email address*: [felix.schulze@warwick.ac.uk](mailto:felix.schulze@warwick.ac.uk)