REGULARITY OF AREA MINIMIZING CURRENTS I: GRADIENT L^p ESTIMATES

CAMILLO DE LELLIS AND EMANUELE SPADARO

Abstract. In a series of papers, including the present one, we give a new, shorter proof of Almgren's partial regularity theorem for area minimizing currents in a Riemannian manifold, with a slight improvement on the regularity assumption for the latter. This note establishes a new a priori estimate on the excess measure of an area minimizing current, together with several statements concerning approximations with Lipschitz multiple valued graphs. Our new a priori estimate is an higher integrability type result, which has a counterpart in the theory of Dir-minimizing multiple valued functions and plays a key role in estimating the accuracy of the Lipschitz approximations.

0. Foreword: a new proof of Almgren's partial regularity

In the present work we continue the investigations started in [14, 17], which together with the forthcoming papers [15, 16] lead to a proof of the following theorem.

Theorem 0.1. *Let $\Sigma \subset \mathbb{R}^{m+n}$ be a C^{3,ε_0} submanifold for some $\varepsilon_0 > 0$ and T an m-dimensional area minimizing integral current in Σ. Then, there is a closed set* $\mathrm{Sing}(T)$ *of Hausdorff dimension at most $m-2$ such that T is a C^{3,ε_0} embedded submanifold in* $\Sigma \setminus (\mathrm{spt}(\partial T) \cup \mathrm{Sing}(T))$.

Theorem 0.1 was first proved by Almgren in his monumental work [3], assuming slightly better regularity on Σ, namely $\Sigma \in C^5$. The improvement itself is therefore not so significant, but our proof, besides being much shorter, introduces new ideas and establishes several new results, which we hope will provide useful tools for further investigations in the area. Indeed, although we still follow Almgren's program and use many of his groundbreaking discoveries, the main steps are achieved in a more efficient way thanks to new estimates and techniques. A striking example is the construction of the so-called center manifold, which is by far the most intricate part of Almgren's work and the least explored, in spite of its importance: in this respect, our construction in [15] is considerably simpler and shorter than [3, Chapter 4], and establishes better results.

Some of our improvements are more transparent, although not substantially simpler, when $\Sigma = \mathbb{R}^{m+n}$ and in a book in preparation [12] we will provide a complete and self-contained account of Theorem 0.1 under such assumption. Moreover, building on our understanding of the various issues involved to the analysis of higher codimension singularities, we plan to tackle Chang's improvement [8], which shows that $\mathrm{Sing}(T)$ consists of isolated points when $m = 2$. His arguments rely on a center manifold construction which does not match exactly the statements of [3] and it is not fully justified, but only briefly sketched in the appendix of [8]. In [18], instead, we give a detailed, simple construction for such center manifold and a complete proof of this refined regularity result.

An alternative route to Chang's result for J-holomorphic currents in symplectic manifolds has been given recently in [27, 28]. The interest in the regularity theory for this class of area minimizing 2-dimensional currents has been generated by the seminal paper of Taubes [32] on the equivalence between Gromov and Seiberg-Witten invariants, where it plays an important role. Moreover, the papers [27, 28] have stimulated a lot of activity in the area, cf., for example, [6, 24, 25, 26]. In [6] Bellettini and Rivière proved that, when T is a special Lagrangian cone in $\mathbb{R}^6$, $\mathrm{Sing}(T)$ consists of finitely many half-lines meeting at the origin. This is, to our knowledge, the only result of its type not covered by the Almgren-Chang works. We believe that the Bellettini-Rivière regularity theorem can be extended to general 3-dimensional area minimizing cones in any space dimension, combining the techniques developed in [14]–[18]. Most of the proofs in [6, 24, 25, 26, 27, 28, 32] take advantage of two specific assumptions, the underlying almost complex structure and the 2-dimensionality of the objects of study. Nonetheless these works have had a profound influence on our research.

0.1. **A blow-up proof: a very brief overview.** In the rest of this foreword we will give a rough outline of the proof of Theorem 0.1, highlighting the contents of this note and the way it merges with its companion papers [15, 16], while comparing them to [3]. Our discussion will be based on a well-known class of examples for which the statement of Theorem 0.1 is optimal, namely singular holomorphic curve of $\mathbb{C}^2$. As it was first observed by Federer (cf. [20, 5.4.19]), the integral currents induced by holomorphic subvarieties of $\mathbb{C}^n$ (with their natural orientation) are area minimizing.

We denote by $D_Q(T)$ the set of points in $\mathrm{spt}(T)\setminus\mathrm{spt}(\partial T)$ where the density of a current T equals the natural number $Q\geq 1$. One first pioneering contribution by Almgren is an elementary, but very clever, generalization of Federer's reduction argument, which has been widely used in several contexts (see [29, Theorem 35.3] and [34]). This argument implies that, if T is area minimizing, then $\mathrm{spt}(T)\setminus(\cup_Q D_Q(T)\cup\mathrm{spt}(\partial T))$ has Hausdorff dimension at most $m-3$. Thus, to prove Theorem 0.1 it suffices to show that the Hausdorff dimension of $\mathrm{Sing}_Q(T) := \mathrm{Sing}(T)\cap D_Q(T)$ is at most $m-2$. Since the "classical" regularity theory ensures that T is a $C^{1,\alpha}$ submanifold in the neighborhood of any point $x\in D_1(T)$, it is natural to argue by induction on Q.

Let us therefore consider the case $Q=2$ and a point $x\in D_2(T)$. By the monotonicity formula, in some neighborhood U of x, $\|T\|$-almost all points have density 1 or 2. If the points of density 1 are a set of $\|T\|$-measure zero, by the classical regularity theory x is a regular point for T. So any $x\in\mathrm{Sing}_2(T)$ must be surrounded by many points of density 1, as it is, for instance, for the complex curve $\{z^2=w^3\}\subset\mathbb{C}^2$ at $x=0$. On the other hand, in such an example 0 is an isolated singularity, whereas, if T were to contradict Theorem 0.1, by standard measure theoretic arguments there would be a point $x\in\mathrm{Sing}_2(T)$ surrounded by many points of density 2. From now on we argue by contradiction and assume that this happens for some area minimizing T at the point $0\in D_2(T)$. Moreover, by known facts in geometric measure theory, we can reduce the contradiction to the case that, for a suitable sequence of radii $r_k \downarrow 0$, the homothetic rescalings of the current T by a factor $1/r_k$ (from now on denoted by T_k) converge to a double copy of an m-dimensional plane, while at the same time $D_2(T_k)$ remains rather large.

It was first recognized by De Giorgi that the convergence of T_k to a *single copy* of a flat plane implies that $\mathrm{spt}(T_k)$ can be well approximated by the graph of Lipschitz functions which are "almost harmonic". However, the example $\{z^2 = w^3\} \subset \mathbb{C}^2$ shows that this is not always the case if the limiting plane has higher multiplicity. Motivated by this fact, Almgren undertook in [3] the strikingly ambitious program of giving a rather complete existence and regularity theory for *multiple valued* functions minimizing a suitable generalization of the Dirichlet energy, called Dir-minimizers. The crowning achievement of this theory is that, except for a closed set of codimension at most 2, Dir-minimizers can be locally decomposed in classical (i.e. single-valued) non-intersecting harmonic sheets (possibly counted with multiplicity). Such "linear theory" is developed in [3, Chapter 2] and revisited in our paper [14]. Moreover, it is complemented by several technical statements linking the multiple valued graphs to the integral currents, a task which is accomplished in [3, Chapter 1] by Almgren and in [17] by us (we refer to the introduction to our previous two papers [14, 17] for more details).

The guiding idea in the contradiction argument is to approximate the currents T_k with Lipschitz 2-*valued* functions and, after a suitable renormalization of their Dirichlet energy, show that they converge to a Dir-minimizer. If the limit inherits a large singular set from the currents T_k, then it contradicts the linear regularity theory. Obviously, this strategy requires suitable approximations of area minimizing currents with multiple valued graphs, accomplished by Almgren in [3, Chapter 3] and by us in the present paper. If one follows our approach, the convergence of these approximations to a Dir-minimizer can be concluded in a rather direct way. However, we cannot expect that such limit inherits the singular set of the current. For example, given the complex curve $\{(z, w) : (z - w^2)^2 = w^5\} \subset \mathbb{C}^2$, any reasonable approximations of homothetic rescalings of this algebraic variety in a neighborhood of the origin converge to a double copy of the classical holomorphic graph $\{(w, w^2) : w \in \mathbb{C}\}$, which has lost the singularity at the origin.

In order to perform the blow-up argument, we then need to "modulate lower order regularities out". This is accomplished by the construction of a center manifold (see [3, Chapter 4] and [15]): such an object is a regular $C^{3,\alpha}$ submanifold which is very close to the average of the sheets of the current at any scale where the latter is "very collapsed". The final blow-up argument is then carried over to a new sequence of 2-valued approximations of T_k, performed on the normal bundles of the center manifolds (see [3, Chapter 5] and [16]). By a delicate unique continuation principle, based on a new monotonicity formula discovered by Almgren, a suitable normalization of the latter approximations does converge to a Dir-minimizer which would be forced to have a large singular set, reaching the desired contradiction. This final step builds upon very delicate computations, which thus require a lot of accuracy in the construction of the center manifold, that in turn needs very good estimates on the approximation results of this note. Thus, unlike the two works [14, 17], which can be considered separately, the present paper and [15, 16] are intimately interconnected.

0.2. **Our contribution; or, what is new.** In their overall structure, our five papers match bijectively the five chapters of [3]. Moreover, it is clear that the ultimate reason for the success of the program is the very same prodigious and celebrated discovery of Almgren: the monotonicity of the frequency function and its astonishing robustness, which enters twice in the plan: at the very beginning, in the linear regularity theory, and at the end, in the convergence of the final approximations (cf. [14, 16]).

So, what is new in our proof? Aside from finer details, which are explained in the introductions to each of our papers, there are some new contributions which come at a higher level. Our investigations started with the idea that the machinery developed in metric analysis and metric geometry in the last 30 years could reduce the complexity of several arguments in Almgren's program. This is, indeed, the case at many levels in the two papers [14, 17] and in this note. Approaching vast parts of Almgren's theory with these tools, we not only get shorter and more transparent proofs, but often also achieve stronger analytic estimates, which give a better starting point for the PDE parts of the program. Moreover, as it often happens when "abstract nonsense" simplifies preexisting mathematical theories, such machinery provides also a better insight to the material of [3], as it highlights the important points in the proofs therein.

However, this alone would not explain the shortness of our papers compared to [3, Chapters 3,4,5]. The other important reason is that we also derive some fundamental, new "hard" estimates. A primary example is the present paper, where the main a priori estimate is a new higher integrability result, which comes from a Gehring-type argument and is inspired by a simple remark in the linear theory (the higher integrability of gradients of Dir-minimizers) which to our knowledge is not observed in Almgren's monograph. Similar instances are present in the papers [15, 16], where some new quantities and guiding principles are introduced (for instance, the "modified frequency" function in [16] and the "splitting-before-tilting" principle in [15], inspired by [26]), which probably lead to the improvement on the regularity assumptions of the ambient manifold Σ. In all these cases we provide more efficient tools compared to [3] and invoke more PDE theory at several levels, drawing connections with fairly classical concepts from other areas of analysis (such as maximal functions, Lipschitz truncations, elliptic systems, Sobolev capacity). Unfortunately we do not understand Almgren's arguments at a sufficiently deep level to draw a fine parallel between our papers [15, 16] and the last two chapters of his book, where the intricacy of the arguments in [3] is almost prohibitive. It remains the fact that our papers are much more accessible, and we hope that in the near future our work will be used to penetrate further in the richness and beauty of Almgren's monograph and to go beyond Theorem 0.1.

Acknowledgments. The first author is deeply indebted to Tristan Rivière who "infected him" with the beauty of the problem. Both authors are also warmly thankful to Giovanni Alberti, Bill Allard, Luigi Ambrosio and Bernd Kirchheim not only for several enlightening conversations, but also for their constant enthusiastic encouragement.

Several other colleagues and friends have contributed with important scientific conversations at some specific stage, for which they will be acknowledged specifically in the various papers. In particular, for this first one we are grateful to Stefano Bianchini, Sergio Conti, Matteo Focardi, Jonas Hirsch and Luca Spolaor for very useful discussions and comments.

This work was carried over several years and the authors wish to thank many institutions where they spent very productive visits, namely: the University of Rome La Sapienza, the Scuola Normale Superiore and the University of Pisa, the Max Planck Institute for Mathematics in the Sciences and the University of Leipzig, the University of Zürich, the SISSA in Trieste, the University of Warwick and, most of all, Princeton University, which hosted the first author during his last sabbatical. We also acknowledge the support of the ERC grant RAM, ERC 306247.

We finally thank the anonymous referee for his/her careful reading and very valuable suggestions, which contributed to a substantial improvement of the initial manuscript.

1. Introduction

1.1. **A priori gradient L^p estimate.** In order to state the main results, we start specifying some assumptions, which will hold throughout the paper. For the notation concerning submanifolds $\Sigma \subset \mathbb{R}^{m+n}$ we refer to [17, Section 1]. With $\mathbf{B}_r(p)$ and $B_r(x)$ we denote, respectively, the open ball with radius r and center p in $\mathbb{R}^{m+n}$ and the open ball with radius r and center x in $\mathbb{R}^m$. $\mathbf{C}_r(x)$ will always denote the cylinder $B_r(x) \times \mathbb{R}^n$ and the point x will be omitted when it is the origin. In fact, by a slight abuse of notation, we will often treat the center x as a point in $\mathbb{R}^{m+n}$, avoiding the correct, but more cumbersome, $(x, 0)$. Let e_i be the unit vectors in the standard basis, π_0 the (oriented) plane $\mathbb{R}^m \times \{0\}$ and $\vec{\pi}_0$ the m-vector $e_1 \wedge \ldots \wedge e_m$ orienting it. We denote by $\mathbf{p}$ and $\mathbf{p}^\perp$ the orthogonal projections onto, respectively, π_0 and its orthogonal complement $\pi_0^\perp$. In some cases we need orthogonal projections onto other planes π and their orthogonal complements $\pi^\perp$, for which we use the notation $\mathbf{p}_\pi$ and $\mathbf{p}_\pi^\perp$. For what concerns integral currents we use the definitions and the notation of [29].

Assumption 1.1. $\Sigma \subset \mathbb{R}^{m+n}$ *is a* C^2 *submanifold of dimension* $m + \bar{n} = m + n - l$, *which is the graph of an entire function* $\Psi : \mathbb{R}^{m+\bar{n}} \to \mathbb{R}^l$ *and satisfies the bounds*

$$\|D\Psi\|_0 \leq c_0 \quad \text{and} \quad \mathbf{A} := \|A_\Sigma\|_0 \leq c_0, \tag{1.1}$$

where c_0 *is a positive (small) dimensional constant.* T *is an integral current of dimension* m *with bounded support contained in* Σ *and which, for some open cylinder* $\mathbf{C}_{4r}(x)$ *(with* $r \leq 1$*) and some positive integer* Q, *satisfies*

$$\mathbf{p}_\sharp T = Q\,[\![B_{4r}(x)]\!] \quad \text{and} \quad \partial T \llcorner \mathbf{C}_{4r}(x) = 0\,. \tag{1.2}$$

If we say that T *is* area minimizing *we then mean that it is area-minimizing in* $\Sigma \cap \mathbf{C}_{4r}(x)$, *namely that* $\mathbf{M}(T) \leq \mathbf{M}(T + \partial S)$ *for any integral* S *with* $\mathrm{spt}(S) \subset \Sigma \cap \mathbf{C}_{4r}(x)$.

Definition 1.2 (Excess measure)**.** For a current T as in Assumption 1.1 we define the *cylindrical excess* $\mathbf{E}(T, \mathbf{C}_{4r}(x))$, the *excess measure* $\mathbf{e}_T$ and its *density* $\mathbf{d}_T$:

$$\mathbf{E}(T, \mathbf{C}_r(x)) := \frac{\|T\|(\mathbf{C}_r(x))}{\omega_m r^m} - Q,$$

$$\mathbf{e}_T(A) := \|T\|(A \times \mathbb{R}^n) - Q\,|A| \qquad \text{for every Borel } \; A \subset B_r(x),$$

$$\mathbf{d}_T(y) := \limsup_{s\to 0} \frac{\mathbf{e}_T(B_s(y))}{\omega_m\, s^m} = \limsup_{s\to 0} \mathbf{E}(T, \mathbf{C}_s(y)),$$

where ω_m is the measure of the m-dimensional unit ball (the subscripts $_T$ will be omitted if clear from the context).

Since T has finite mass, the function $\mathbf{d}$ is naturally an L^1 function. However, we can show the following higher integrability estimate when T is, in addition, area minimizing. We call it a gradient L^p estimate because we will show that $\mathbf{d}$ coincides with the gradient of an appropriate Lipschitz function on a large region.

Theorem 1.3 (Gradient L^p estimate)**.** *There exist constants $p_1 > 1$ and $C, \varepsilon_{10} > 0$ (depending on $m, n, \bar{n}, Q$) with the following property. Let T be as in Assumption 1.1 in the cylinder $\mathbf{C}_4$. If T is area minimizing and $E = \mathbf{E}(T, \mathbf{C}_4) < \varepsilon_{10}$, then*

$$\int_{\{\mathbf{d}\leq 1\}\cap B_2} \mathbf{d}^{p_1} \leq C\, E^{p_1 - 1}\left(E + \mathbf{A}^2\right). \tag{1.3}$$

In the case $Q = 1$ or $\bar{n} = 1$, it follows from the classical regularity theory (essentially due to De Giorgi, cf. [9]) that T is a $C^{1,\alpha}$ submanifold in $\mathbf{C}_2$. However, when $\min\{Q, \bar{n}\} \geq 2$, T is not necessarily regular and Theorem 1.3 gives in fact an *a priori* regularity estimate: in this case (1.3) cannot be improved (except for optimizing the constants p_1, C and ε_{10}). Indeed, for $Q = m = 2$, $\Sigma = \mathbb{R}^4$ and $p_1 = 2$, (1.3) is false no matter how large ε_{10}^{-1} and C are chosen (cf. [10, Section 6.2]).

In order to prove Theorem 1.3 we develop the following tools:

(a) a general scheme to approximate integer rectifiable currents with multiple valued functions, relying heavily on the "metric analysis" of [14] and on a modified "Jerrard–Soner" BV estimate for the slicing of currents (cf. Proposition 2.2);
(b) a simple and robust harmonic approximation of area minimizing currents with multiple valued functions (cf. Theorem 4.2);
(c) the higher integrability of the gradient of Dir-minimizing multiple valued functions (cf. Theorem 5.1 – see also [30] for a different proof and related results).

In turn, Theorem 1.3 will be combined with (a) to achieve a very accurate approximation result for area minimizing current, stated in Theorem 1.4. This theorem and some corollaries of our analysis play a fundamental role in the papers [15, 16] and, as explained in the Foreword, have a counterpart in [3, Chapter 3]. However, our derivation of Theorem 1.4 differs substantially from Almgren's and when we use some of his ideas, as it is for the existence of the almost projection $\boldsymbol{\rho}^\star$ of Section 7, we give independent arguments for the main steps of the proof.

1.2. **Strong approximation of area minimizing currents.** Concerning multiple valued functions we will follow the notation and terminology of [14, 17]. In particular, a Q-valued function is a map f (usually defined over a measurable subset Ω of $\mathbb{R}^m$) taking values in the space $\mathcal{A}_Q(\mathbb{R}^n)$ of unordered Q-tuples of points in $\mathbb{R}^n$, denoted by $\sum_i [\![P_i]\!]$. $\mathcal{A}_Q(\mathbb{R}^n)$ can be equipped with a natural metric $\mathcal{G}$ (cf. [14, Definition 0.2]) and for f measurable there exist measurable functions $f_i : \Omega \to \mathbb{R}^n$ such that $f(x) = \sum_i [\![f_i(x)]\!]$ $\forall x \in \Omega$ (cf. [14, Proposition 0.4]). The functions f_i are not uniquely determined, but in using this notation we assume to have fixed some suitable f_i's. Moreover, if f is Lipschitz, resp. $f \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^n))$ (cf. [14, Definition 0.5]) and Ω is open, then there exist measurable functions $Df_i \in L^\infty$, resp. L^2, such that $\sum_i [\![Df_i(x)]\!]$ is the approximate differential of f (cf. [14, Definition 2.6]) at a.e. x. In fact in this case the f_i's and Df_i's can be chosen so that the first are approximately differentiable a.e. and the second are their approximate differentials in the classical sense (cf. [17, Lemma 1.1]). The Dirichlet energy of f is then $\mathrm{Dir}(f, \Omega) := \int_\Omega |Df|^2$, where $|Df|^2 := \sum_i |Df_i|^2$. Following [17, Definition 1.10], we denote by $\mathbf{G}_f$ the integer rectifiable current, in $\mathbb{R}^{m+n}$, naturally associated to the graph of a Lipschitz Q-valued map $f : \mathbb{R}^m \supset A \to \mathcal{A}_Q(\mathbb{R}^n)$. Moreover, we will use the notation $\mathrm{osc}\,(f)$ for the quantity $\inf_p \sup_x \mathcal{G}(f(x), Q\,[\![p]\!])$.

Theorem 1.4 (Almgren's strong approximation)**.** *There exist constants $C, \gamma_1, \varepsilon_1 > 0$ (depending on $m, n, \bar{n}, Q$) with the following property. Assume that T is area minimizing, satisfies Assumption 1.1 in the cylinder $\mathbf{C}_{4r}(x)$ and $E = \mathbf{E}(T, \mathbf{C}_{4\,r}(x)) < \varepsilon_1$. Then, there is a map $f : B_r(x) \to \mathcal{A}_Q(\mathbb{R}^n)$, with $\mathrm{spt}(f(x)) \subset \Sigma$ for every x, and a closed set $K \subset B_r(x)$ such that*

$$\mathrm{Lip}(f) \leq C E^{\gamma_1}, \tag{1.4}$$

$$\mathbf{G}_f \llcorner (K \times \mathbb{R}^n) = T \llcorner (K \times \mathbb{R}^n) \quad \textit{and} \quad |B_r(x) \setminus K| \leq C\, E^{\gamma_1} \left(E + r^2\, \mathbf{A}^2\right) r^m, \tag{1.5}$$

$$\left| \|T\|(\mathbf{C}_{\sigma\,r}(x)) - Q\,\omega_m\,(\sigma\,r)^m - \tfrac{1}{2} \int_{B_{\sigma\,r}(x)} |Df|^2 \right| \leq C\, E^{\gamma_1} \left(E + r^2\, \mathbf{A}^2\right) r^m \quad \forall\, 0 < \sigma \leq 1. \tag{1.6}$$

If in addition $\mathbf{h}(T, \mathbf{C}_{4r}(x), \pi_0) := \sup\{|\mathbf{p}^\perp(x) - \mathbf{p}^\perp(y)| : x, y \in \mathrm{spt}(T) \cap \mathbf{C}_{4r}(x)\} \leq r$, then

$$\mathrm{osc}\,(f) \leq C\mathbf{h}(T, \mathbf{C}_{4r}(x), \pi_0) + C(E^{1/2} + r\,\mathbf{A})\,r\,. \tag{1.7}$$

The gain of a small power E^{γ_1} in the three estimates (1.4)-(1.6) plays a crucial role in the papers [15, 16]. When $Q = 1$ and $\Sigma = \mathbb{R}^{m+1}$, this approximation theorem was first proved by De Giorgi in [9]. In the generality above it appears in the big regularity paper for the first time (cf. [3, Sections 3.28-3.30]). Its proof is an elementary consequence of Theorem 6.1 and the Lipschitz approximation algorithm mentioned above. In turn Theorem 6.1 will be derived from Theorem 1.3 using a suitable competitor argument. In the case $Q = 1$, the competitor is the convolution of (a first) Lipschitz approximation with a smooth kernel, a classical argument which in fact appears already in De Giorgi's seminal paper [9], although in a slightly different form (cf. [13, Appendix]).

Here we need a similar approach in the framework of multiple valued functions. However, since $\mathcal{A}_Q(\mathbb{R}^n)$ is highly nonlinear, it is not possible to regularize directly by convolution.

We exploit at this point a key idea of Almgren, embedding $\mathcal{A}_Q(\mathbb{R}^n)$ in an Euclidean space and using some suitable "almost projections" $\boldsymbol{\rho}^\star_\delta$. Our proof of the existence of these almost projections is however different from the one given by Almgren in [3, Theorem 1.3] and, indeed, gives better bounds in terms of the relevant parameters (see Proposition 6.2).

1.3. **Harmonic approximation.** A second ingredient which in [15, 16] will play a key role is the harmonic approximation of Theorem 1.6 below (already mentioned in (b) above). In order to state it we need to set some notation about the ambient manifold Σ.

Remark 1.5 (Estimates on Ψ in good Cartesian coordinates)**.** Assume that T is as in Assumption 1.1 in the cylinder $\mathbf{C}_{4r}(x)$. If $E := \mathbf{E}(T, \mathbf{C}_{4r}(x))$ is smaller than a geometric constant, we can assume, without loss of generality, that the function $\Psi : \mathbb{R}^{m+\bar n} \to \mathbb{R}^l$ parameterizing Σ satisfies $\Psi(x) = 0$, $\|D\Psi\|_0 \le C\, E^{1/2} + C\mathbf{A}r$ and $\|D^2\Psi\|_0 \le C\mathbf{A}$. Indeed observe that

$$E = \mathbf{E}(T, \mathbf{C}_{4r}(x)) = \frac{1}{2\,\omega_m\,(4r)^m} \int_{\mathbf{C}_{4r}(x)} |\vec{T}(y) - \vec{\pi}_0|^2\, d\|T\|(y)\,.$$

Thus, we can fix a point $p \in \mathrm{spt}(T) \cap \mathbf{C}_{4r}(x)$ such that $|\vec{T}(p) - \vec{\pi}_0| \le C\, E^{1/2}$. Then, we can find an associated rotation $O \in O(m+n)$ such that $O_\sharp \vec{T}(p) = \vec{\pi}_0$ and $|O| \le C\, E^{1/2}$. It follows that $\pi := O(T_p\Sigma)$ is a $(m+\bar n)$-dimensional plane such that $\pi_0 \subset \pi$ and $\|\pi - T_p\Sigma\| \le CE^{1/2}$. We choose new coordinates so that π_0 remains equal to $\mathbb{R}^m \times \{0\}$ but $\mathbb{R}^{m+\bar n} \times \{0\}$ equals π. Since the excess E is assumed to be sufficiently small, we can write Σ as the graph of a function $\Psi : \pi \to \pi^\perp$. If $(z, \Psi(z)) = p$, then $|D\Psi(z)| \le C\|T_p\Sigma - \mathbb{R}^{m+\bar n} \times \{0\}\| \le CE^{1/2}$. However, $\|D^2\Psi\|_0 \le C\mathbf{A}$ and so $\|D\Psi\|_0 \le CE^{1/2} + C\mathbf{A}r$. Moreover, $\Psi(x) = 0$ is achieved translating the system of reference by a vector orthogonal to $\mathbb{R}^{m+\bar n} \times \{0\}$ and, hence, belonging to $\{0\} \times \mathbb{R}^n$.

From now on, we will often consider Q-valued maps $y \mapsto w(y) \in \mathcal{A}_Q(\mathbb{R}^n) = \mathcal{A}_Q(\mathbb{R}^{\bar n} \times \mathbb{R}^l)$ which take the form $w(y) = \sum_i [\![(u_i(y), \Psi(y, u_i(y)))]\!]$, where $u = \sum_i [\![u_i]\!]$ is evidently a map taking values in $\mathcal{A}_Q(\mathbb{R}^{\bar n})$. For w we will then use the short-hand notation $w = (u, \Psi(y, u))$. We also recall the notation for the average map $\eta : \mathcal{A}_Q(\mathbb{R}^n) \to \mathbb{R}^n$ defined by

$$\mathcal{A}_Q(\mathbb{R}^n) \ni T = \sum_{i=1}^Q [\![P_i]\!] \mapsto \eta(T) := \frac{1}{Q}\sum_{i=1}^Q P_i \in \mathbb{R}^n.$$

Theorem 1.6 (Harmonic approximation)**.** *Let γ_1 be the constant of Theorem 1.4. Then, for every $\bar\eta, \bar\delta > 0$, there is a positive constant $\bar\varepsilon_1$ with the following property. Assume that T is as in Theorem 1.4, $E := \mathbf{E}(T, \mathbf{C}_{4\,r}(x)) < \bar\varepsilon_1$ and $r\mathbf{A} \le E^{1/4+\bar\delta}$. If f is the map in Theorem 1.4 and we fix Cartesian coordinates as in Remark 1.5, then there exists a* Dir*-minimizing function $u : B_r(x) \to \mathcal{A}_Q(\mathbb{R}^{\bar n})$ such that $w := (u, \Psi(y, u))$ satisfies*

$$r^{-2}\int_{B_r(x)} \mathcal{G}(f, w)^2 + \int_{B_r(x)} \left(|Df| - |Dw|\right)^2 + \int_{B_r(x)} |D(\boldsymbol{\eta}\circ f) - D(\boldsymbol{\eta}\circ w)|^2 \le \bar\eta\, E\, r^m\,. \quad (1.8)$$

This theorem is the multi-valued analog of De Giorgi's harmonic approximation (cf. [9]). We prove it via a compactness argument which, although very close in spirit to De Giorgi's original one, is to our knowledge new (even when $n = \bar{n} = 1$). Indeed, it uses neither the monotonicity formula nor a regularization by convolution of the Lipschitz approximation, and we expect it to be useful in different contexts.

1.4. **Persistence of Q-points.** A major ingredient in [16] is the persistence of points of maximal multiplicity in the approximation of Theorem 1.4, when interpreted in a suitable "limiting sense". If the current T has a point of density Q, f must satisfy the following integral bound (even though f might have no values of multiplicity Q).

Theorem 1.7 (Persistence of Q-points)**.** *For every $\hat{\delta}, C^\star > 0$, there is $\bar{s} \in]0, \frac{1}{2}[$ such that, for every $s < \bar{s}$, there exists $\hat{\varepsilon}(s, C^*, \hat{\delta}) > 0$ with the following property. If T is as in Theorem 1.4, $E := \mathbf{E}(T, \mathbf{C}_{4\,r}(x)) < \hat{\varepsilon}$, $r^2\mathbf{A}^2 \leq C^\star E$ and $\Theta(T, (p, q)) = Q$ at some $(p, q) \in \mathbf{C}_{r/2}(x)$, then the approximation f of Theorem 1.4 satisfies*

$$\int_{B_{sr}(p)} \mathcal{G}(f, Q\,[\![\boldsymbol{\eta} \circ f]\!])^2 \leq \hat{\delta} s^m r^{2+m} E\,. \tag{1.9}$$

1.5. **A remark on notation.** Finally we remark that we follow closely the notation of [14, 17], except for a subtle point. We denote by $\boldsymbol{\xi}$ the map in [14, Corollary 2.2], which there was denoted by $\boldsymbol{\xi}_{BW}$, since the symbol $\boldsymbol{\xi}$ was in fact used for the "precursor map"of [14, Theorem 2.1]. So, here $\boldsymbol{\xi} : \mathcal{A}_Q(\mathbb{R}^n) \to \mathbb{R}^{N(Q,n)}$ is an injective function satisfying the following three properties:

(i) $\mathrm{Lip}(\boldsymbol{\xi}) \leq 1$;
(ii) $\mathrm{Lip}(\boldsymbol{\xi}^{-1}|_{\mathcal{Q}}) \leq C(n, Q)$, where $\mathcal{Q} = \boldsymbol{\xi}(\mathcal{A}_Q)$;
(iii) $|Df| = |D(\boldsymbol{\xi} \circ f)|$ almost everywhere for every $f \in W^{1,2}(\Omega, \mathcal{A}_Q)$.

This "improved" $\boldsymbol{\xi}$ was suggested by Brian White and appears for the first time in [8]. The conclusion (iii) above is actually not explicitly stated in [14], but it follows easily: indeed [14, Corollary 2.2] implies the identity $|Df| = |D(\boldsymbol{\xi} \circ f)|$ at every point of differentiability of a Lipschitz map and, hence, almost everywhere. The case of a general $f \in W^{1,2}(\Omega, \mathcal{A}_Q)$ can then be concluded from [14, Proposition 2.5].

We will use the notation C and c for generic positive dimensional constants, which may possibly change from line to line: we will always understand that these constants depends only on the dimensional parameters $m, \bar{n}, n, Q, c_0$ of Assumption 1.1.

2. Lipschitz approximation

To begin with, we develop a robust algorithm to approximate currents T as in Assumption 1.1 with graphs of multiple valued functions. Following the work of Ambrosio and Kirchheim [5], we view the slice map $x \mapsto \langle T, \mathbf{p}, x\rangle$ as a function taking values in the space $\boldsymbol{I}_0(\mathbb{R}^n)$ of 0-dimensional integral currents. A key estimate of Jerrard and Soner (cf. [5, 23]) implies that this map has bounded variation in the metric sense introduced by Ambrosio in [4]. On the other hand, following [14], Q-valued functions can be viewed as Sobolev maps taking values into (a subset of) $\boldsymbol{I}_0(\mathbb{R}^n)$. Thus, finding Lipschitz multiple valued approximations of T can be seen as a particular case of the more general task of finding Lipschitz approximations of BV maps with a fairly general target space.

Definition 2.1 (Maximal function of the excess measure)**.** Given a current T as in Assumption 1.1 we introduce the "non-centered" maximal function of $\mathbf{e}_T$:

$$\mathbf{me}_T(y) := \sup_{y\in B_s(w)\subset B_{4r}(x)} \frac{\mathbf{e}_T(B_s(w))}{\omega_m\, s^m} = \sup_{y\in B_s(w)\subset B_{4r}(x)} \mathbf{E}(T,\mathbf{C}_s(w)).$$

We can now state the main result of the section, which provides the first Lipschitz approximation for rectifiable currents.

Proposition 2.2 (Lipschitz approximation)**.** *There exists a constant $C>0$ with the following property. Let T and Ψ be as in Assumption 1.1 in the cylinder $\mathbf{C}_{4s}(x)$. Set $E=\mathbf{E}(T,\mathbf{C}_{4s}(x))$, let $0<\delta_{11}<1$ be such that $16^m E<\delta_{11}$, and define*

$$K := \big\{\mathbf{me}_T < \delta_{11}\big\}\cap B_{3s}(x)\,.$$

Then, there is $u\in \mathrm{Lip}(B_{3s}(x),\mathcal{A}_Q(\mathbb{R}^n))$ such that $\mathrm{spt}(u(y))\subset\Sigma$ for every $y\in B_{3s}(x)$ and

$$\mathrm{Lip}(u)\leq C\,\big(\delta_{11}^{1/2}+\|D\Psi\|_0\big),\qquad \mathrm{osc}\,(u)\leq C\mathbf{h}(T,\mathbf{C}_{4s}(x),\pi_0)+Cs\|D\Psi\|_0\,,$$

$$\mathbf{G}_u\llcorner(K\times\mathbb{R}^n)=T\llcorner(K\times\mathbb{R}^n),$$

$$|B_r(x)\setminus K|\leq\frac{10^m}{\delta_{11}}\,\mathbf{e}_T\Big(\{\mathbf{me}_T>2^{-m}\delta_{11}\}\cap B_{r+r_0 s}(x)\Big)\quad\forall\, r\leq 3\,s,\tag{2.1}$$

where $r_0=16\,\sqrt[m]{E/\delta_{11}}<1$.

The proof of the proposition is based on a BV estimate which differs from the ones of [5, 23]. Note that we do not assume that T is area minimizing. Indeed, even the assumption (1.2) could be relaxed, but we do not pursue this issue here.

2.1. **The modified Jerrard–Soner estimate.** Recall that each element $S\in \boldsymbol{I}_0(\mathbb{R}^{m+n})$ is simply a finite sum of Dirac deltas, $S=\sum_{i=1}^h w_i\,\delta_{z_i}$, where $h\in\mathbb{N}$, $w_i\in\{-1,1\}$ and the z_i's are (not necessarily distinct) points in $\mathbb{R}^{m+n}$. Let T be a current as in Assumption 1.1 in the cylinder $\mathbf{C}_4$. The slicing map $x\mapsto\langle T,\mathbf{p},x\rangle$ takes values in $\boldsymbol{I}_0(\mathbb{R}^{m+n})$ and is characterized by (cf. [29, Section 28]):

$$\int_{B_4}\langle T,\mathbf{p},x\rangle\,(\varphi)dx = T(\varphi\,dx)\quad\text{for every }\ \varphi\in C^\infty_c(\mathbf{C}_4).\tag{2.2}$$

Moreover $\mathrm{spt}(\langle T,\mathbf{p},x\rangle)\subseteq\mathbf{p}^{-1}(\{x\})$ and therefore $\langle T,\mathbf{p},x\rangle=\sum_i w_i\,\delta_{(x,y_i)}$. The assumption (1.2) guarantees that $\sum_i w_i=Q$ for almost every x. In order to state our BV estimate, we consider the push-forwards of $\langle T,\mathbf{p},x\rangle$ into the vertical directions:

$$T_x := \mathbf{p}^\perp_\sharp\big(\,\langle T,\mathbf{p},x\rangle\,\big)\in\boldsymbol{I}_0(\mathbb{R}^n)\,.\tag{2.3}$$

It follows from (2.2) that the currents T_x are characterized through the identity:

$$\int_{B_4}T_x(\psi)\varphi(x)\,dx = T(\varphi(x)\,\psi(y)\,dx)\quad\text{for every }\ \varphi\in C^\infty_c(B_4),\ \psi\in C^\infty_c(\mathbb{R}^n).\tag{2.4}$$

Proposition 2.3 (BV estimate). *Assume T satisfies Assumption 1.1 in $\mathbf{C}_4$ (i.e. $r = 1$ and $x = 0$ in Assumption 1.1). For every $\psi \in C_c^\infty(\mathbb{R}^n)$, set $\Phi_\psi(x) := T_x(\psi)$. If $\|D\psi\|_\infty \leq 1$, then $\Phi_\psi \in BV(B_4)$ and satisfies*

$$\big(|D\Phi_\psi|(A)\big)^2 \leq 2m^2\, \mathbf{e}_T(A)\, \|T\|(A \times \mathbb{R}^n) \quad \textit{for every Borel set } A \subseteq B_4. \tag{2.5}$$

Note that in the usual Jerrard-Soner estimate the RHS of (2.5) would be $(\|T\|(A\times\mathbb{R}^n))^2$.

Proof. It is enough to prove (2.5) for every open set $A \subseteq B_4$. To this aim, recall that:

$$|D\Phi_\psi|(A) = \sup\left\{\int_A \Phi_\psi(x)\operatorname{div}\varphi(x)\,dx \;:\; \varphi \in C_c^\infty(A,\mathbb{R}^m),\ \|\varphi\|_\infty \leq 1\right\}. \tag{2.6}$$

For any smooth vector field φ we have $(\operatorname{div}\varphi(x))\,dx = d\,\Xi$, where

$$\Xi = \sum_j \varphi_j\, d\hat{x}^j \quad \text{and} \quad d\hat{x}^j = (-1)^{j-1}dx^1\wedge\cdots\wedge dx^{j-1}\wedge dx^{j+1}\wedge\cdots\wedge dx^m.$$

From (2.4) and the assumption $\partial T \llcorner \mathbf{C}_4 = 0$ in (1.2), we conclude that

$$\begin{aligned}\int_A \Phi_\psi(x)\operatorname{div}\varphi(x)\,dx &= \int_{B_4} T_x(\psi)\operatorname{div}\varphi(x)\,dx = T(\psi\operatorname{div}\varphi\,dx)\\ &= T(\psi\, d\,\Xi) = T(d(\psi\,\Xi)) - T(d\psi\wedge\Xi) = -T(d\psi\wedge\Xi)\,.\end{aligned} \tag{2.7}$$

Observe that the m-form $d\psi\wedge\Xi$ has no dx component, since

$$d\psi\wedge\Xi = \sum_{j=1}^m\sum_{i=1}^n \frac{\partial\psi}{dy^i}(y)\,\varphi_j(x)\,dy^i\wedge d\hat{x}^j. \tag{2.8}$$

Write $\vec{T} = \langle\vec{T},\vec{\pi}_0\rangle\,\vec{\pi}_0 + \vec{S}$. Then,

$$(T(d\psi\wedge\Xi))^2 = \Big(\int\langle\vec{S}, d\psi\wedge\Xi\rangle\, d\|T\|\Big)^2 \leq \||d\psi\wedge\Xi|\|_\infty^2\|T\|(A\times\mathbb{R}^n)\int_{A\times\mathbb{R}^n}|\vec{S}|^2\, d\,\|T\|\,,$$

($|\cdot|$ denotes the norms on Λ_m and Λ^m induced by the natural inner products $\langle,\rangle$). Since $|\vec{S}|^2 = 1 - \langle\vec{T},\vec{\pi}_0\rangle^2 \leq 2 - 2\langle\vec{T},\vec{\pi}_0\rangle$, we have

$$\int_{A\times\mathbb{R}^n}|\vec{S}|^2\, d\,\|T\| \leq 2\int_{A\times\mathbb{R}^n}\Big(1 - \langle\vec{T},\vec{\pi}_0\rangle\Big)\, d\,\|T\| = 2\,\mathbf{e}_T(A).$$

Moreover, by (2.8), $\||d\psi\wedge\Xi|\|_\infty \leq m\,\|D\psi\|_\infty\,\|\varphi\|_\infty \leq m$. Summarizing, we get

$$\int_A \Phi_\psi(x)\operatorname{div}\varphi(x)\,dx \leq \big(2m^2\,\mathbf{e}_T(A)\,\|T\|(A\times\mathbb{R}^n)\big)^{1/2} \qquad \forall\varphi\in C_c^\infty(A,\mathbb{R}^m), \|\varphi\|_\infty \leq 1\,.$$

Taking the supremum over φ's we conclude (2.5) through (2.6). □

2.2. **Proof of Proposition 2.2.** Since the statement is invariant under translations and dilations, without loss of generality we assume $x = 0$ and $s = 1$. Consider the slices $T_x := \mathbf{p}^\perp_\sharp \langle T, \mathbf{p}, x\rangle \in \boldsymbol{I}_0(\mathbb{R}^n)$ and recall that $\|T\|(A\times\mathbb{R}^n) \geq \int_A \mathbf{M}(T_x)\,dx$ for every open set A (cf. [29, Lemma 28.5]). Therefore,

$$\mathbf{M}(T_x) \leq \lim_{r\to 0} \frac{\|T\|(\mathbf{C}_r(x))}{\omega_m\, r^m} \leq \mathbf{me}_T(x) + Q \quad \text{for almost every } x.$$

Since $\delta_{11} < 1$, we infer $\mathbf{M}(T_x) < Q+1$ for a.e. $x \in K$. There are, then, Q functions $g_i : K \to \mathbb{R}^n$ such that $T_x = \sum_{i=1}^Q \delta_{g_i(x)}$ for a.e. $x\in K$. Define $g : K \mapsto \mathcal{A}_Q(\mathbb{R}^n)$ as $g := \sum_i [\![g_i]\!]$ and fix $\psi \in C^\infty_c(\mathbb{R}^n)$. Proposition 2.3 gives

$$(|D\Phi_\psi|(B_r(y)))^2 \leq 2m^2\, \mathbf{e}_T(B_r(y))\, \|T\|(\mathbf{C}_r(y)) = 2m^2\, \mathbf{e}_T(B_r(y))\big(Q|B_r(y)| + \mathbf{e}_T(B_r(y))\big)\,.$$

Hence, if we define the maximal function

$$\mathbf{m}|D\Phi_\psi|(x) := \sup_{x\in B_r(y)\subset B_{4r}} \frac{|D\Phi_\psi|(B_r(y))}{|B_r(y)|}\,,$$

we conclude that

$$(\mathbf{m}|D\Phi_\psi|(x))^2 \leq 2m\, \mathbf{me}_T(x)^2 + 2m\, Q\, \mathbf{me}_T(x) \leq C\delta_{11} \quad \text{for every } x\in K.$$

Therefore, the theory of BV functions gives a dimensional constant C such that

$$|\Phi_\psi(x) - \Phi_\psi(y)| \leq C\,\delta_{11}^{1/2}\,|x-y| \qquad \forall x,y\in K \text{ Lebesgue points of } \Phi_\psi, \tag{2.9}$$

(see for instance [19, Section 6.6.2]: although in that reference the authors use the *centered* maximal function, the proof works obviously also in our context). Consider next the Wasserstein distance of exponent 1 on 0-dimensional integral currents S_1, S_2:

$$W_1(S_1,S_2) := \sup\big\{\langle S_1 - S_2, \psi\rangle \,:\, \psi\in C^1(\mathbb{R}^n),\ \|D\psi\|_\infty \leq 1\big\}\,. \tag{2.10}$$

Obviously, when $S_1 = \sum_i [\![S_{1i}]\!]\,, S_2 = \sum_i [\![S_{2i}]\!] \in \mathcal{A}_Q(\mathbb{R}^n)$, the supremum in (2.10) can be taken over a suitable countable subset of $\psi\in C^\infty_c(\mathbb{R}^n)$, chosen independently of the S_i's. Moreover, we have that

$$W_1(S_1,S_2) = \min_{\sigma\in\mathscr{P}_Q}\sum_i |S_{1i} - S_{2\sigma(i)}| \geq \min_{\sigma\in\mathscr{P}_Q}\Big(\sum_i |S_{1i} - S_{2\sigma(i)}|^2\Big)^{1/2} = \mathcal{G}(S_1,S_2). \tag{2.11}$$

So $\mathcal{G}(g(x),g(y)) \leq C\,\delta_{11}^{1/2}\,|x-y|$ for a.e. $x,y\in K$. (The first equality in (2.11) is well-known, but not easy to find in the literature. It can be derived by suitably modifying the arguments of [20, 4.1.12]. Another quick derivation is the following. Consider the set Π of probability measures π on $\mathbb{R}^n\times\mathbb{R}^n$ of the form $\sum_{i,j} c_{ij}\delta_{(S_{1i},S_{2j})}$, where the matrix of coefficients c_{ij} consists of nonnegative entries with $\sum_k c_{kj} = 1$ and $\sum_k c_{ik} = 1$ for every i and j, i.e. it is a doubly stochastic matrix. It then follows from the Kantorovich duality, see for instance [33, Theorem 1.14], that $W_1(S_1,S_2) = \min_{\pi\in\Pi}\int |x-y|\,d\pi(x,y)$. Observe however that $\int |x-y|\,d\pi(x,y)$ is a linear function of the coefficients c_{ij}: the space of such matrices, also called Birkhoff polytope, is a compact convex set and so the minimum is attained on the subset of extremal points. By the classical Birkhoff-von Neumann Theorem this set consists of the permutations matrices (see [7]) and so $\min_{\pi\in\Pi}\int|x-y|\,d\pi = \min_{\sigma\in\mathscr{P}_Q}\sum_i |S_{1i}-S_{2\sigma(i)}|$.)

Next, write $g(x) = \sum_i [\![(h_i(x),\Psi(x,h_i(x)))]\!]$. Obviously $x\mapsto h(x) := \sum_i [\![h_i(x)]\!] \in \mathcal{A}_Q(\mathbb{R}^{\bar n})$ is a Lipschitz map on K with Lipschitz constant $\leq C\,\delta_{11}^{1/2}$. Recalling [14, Theorem 1.7], we can extend it to a map $\bar u \in \mathrm{Lip}(B_3,\mathcal{A}_Q(\mathbb{R}^{\bar n}))$ satisfying $\mathrm{Lip}(\bar u)\leq C\,\delta_{11}^{1/2}$ and $\mathrm{osc}\,(\bar u)\leq C\mathrm{osc}\,(h)$. Set finally $u(x) = \sum_i [\![(\bar u_i(x),\Psi(x,\bar u_i(x)))]\!]$. We start showing the Lipschitz bound. Fix $x_1,x_2\in B_3$ and assume, without loss of generality, that $\mathcal{G}(\bar u(x_1),\bar u(x_2))^2 = \sum_i |\bar u_i(x_1)-\bar u_i(x_2)|^2$. Then

$$\begin{aligned}\mathcal{G}(u(x_1),u(x_2))^2 &\leq \sum_i \left|(\bar u_i(x_1),\Psi(x_1,\bar u_i(x_1))) - (\bar u_i(x_2),\Psi(x_2,\bar u_i(x_2)))\right|^2\\ &\leq 2\sum_i \Big((1+\|D_y\Psi\|_0^2)|\bar u_i(x_1)-\bar u_i(x_2)|^2 + \|D_x\Psi\|_0^2|x_1-x_2|^2\Big)\\ &\leq 2(1+\|D\Psi\|_0^2)\mathcal{G}(\bar u(x_1),\bar u(x_2))^2 + 2\|D\Psi\|_0^2|x_1-x_2|^2\\ &\leq C(\delta_{11}+\|D\Psi\|_0^2)|x_1-x_2|^2\,.\end{aligned}$$

As for the L^∞ bound, let $\eta>0$ be arbitrary and $p\in\mathbb{R}^{\bar n}$ be such that $\mathrm{osc}(\bar u)\leq \sup_{x\in B_3}\mathcal{G}(\bar u(x),Q\,[\![p]\!])+\eta$. Proceeding as above

$$\begin{aligned}\mathrm{osc}(u)^2 &\leq \sup_{x\in B_3}\mathcal{G}(u(x),Q\,[\![(p,\Psi(0,p))]\!])^2\\ &\leq 2\sup_{x\in B_3}\Big((1+\|D\Psi\|_0^2)\mathcal{G}(\bar u(x),Q\,[\![p]\!])^2 + \|D\Psi\|_0^2|x|^2\Big)\\ &\leq 4(1+\|D\Psi\|_0^2)\big(\mathrm{osc}(\bar u)^2+\eta^2\big) + 18\,\|D\Psi\|_0^2.\end{aligned}$$

Since $\mathrm{osc}(h)\leq \mathbf{h}(T,\mathbf{C}_4,\pi_0)$, the estimate on $\mathrm{osc}(u)$ follows letting $\eta\downarrow 0$.

The identity $\mathbf{G}_u\llcorner(K\times\mathbb{R}^n) = T\llcorner(K\times\mathbb{R}^n)$ is a consequence of $u(x)=T_x$ for a.e. $x\in K$. Indeed, recall that both T and $\mathbf{G}_u$ are rectifiable and observe that $\langle\vec T,\vec\pi_0\rangle\neq 0$ $\|T\|$-a.e. on $K\times\mathbb{R}^n$, because $\mathbf{me}_T<\infty$ on K. Similarly, $\langle\vec{\mathbf{G}}_u,\vec\pi_0\rangle\neq 0$ $\|\mathbf{G}_u\|$-a.e. on $K\times\mathbb{R}^n$, by [17, Proposition 1.4]. Thus, $(\mathbf{G}_u-T)\llcorner K\times\mathbb{R}^n = 0$ if and only if $(\mathbf{G}_u-T)\llcorner dx\mathbf{1}_{K\times\mathbb{R}^n}=0$. The latter identity follows from the slicing formula and the property $\langle T,\mathbf{p},x\rangle = \langle\mathbf{G}_u,\mathbf{p},x\rangle = \sum_i\delta_{(x,u_i(x))}$, valid for a.e. $x\in K$.

Finally, for each $x\in B_r\setminus K$ choose a ball $x\in B^x = B_{r(x)}(y(x))\subset B_4$ such that $\mathbf{e}_T(B^x)\geq 2^{-m}\delta_{11}\omega_m r(x)^m$. By the $5r$-Covering theorem, we choose balls $\hat B^i = B_{5r(x_i)}(y(x_i))$ which cover $B_r\setminus K$ and such that the balls B^{x_i} are pairwise disjoint. We then conclude

$$|B_r\setminus K|\leq 10^m\delta_{11}^{-1}\mathbf{e}_T\left(\bigcup_i B^{x_i}\right)\,. \tag{2.12}$$

Fix $y\in B^{x_i}$. Since $B^{x_i}\subset B_4$, we have $2^{-m}\delta_{11}\omega_m r(x_i)^m\leq \mathbf{e}_T(B^{x_i})\leq\mathbf{e}_T(B_4)=4^m\omega_m E$, which implies $2r(x_i)\leq r_0<1$. Thus, $y\in B_{r+r_0}\subset B_4$. By definition of $\mathbf{me}_T$ we obviously have $\mathbf{me}_T(y)\geq 2^{-m}\delta_{11}$. So $\cup_i B^{x_i}\subset B_{r+r_0}\cap\{\mathbf{me}_T>2^{-m}\delta_{11}\}$ and (2.12) implies (2.1).

3. Patching multiple valued graphs

In this section we prove some complementary results to the theory of multiple valued functions as exposed in [14, 17]. In particular, we show here a concentration compactness principle for Q-valued functions, and give an algorithm to construct suitable competitors for the Dirichlet energy, which will be also used in [16]. We first introduce some terminology.

Definition 3.1 (Translating sheets)**.** Let $\Omega \subset \mathbb{R}^m$ be a bounded open set. A sequence of maps $\{h_k\}_{i\in\mathbb{N}} \subset W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^n))$ is called a sequence of *translating sheets* if there are:

(a) integers $J \geq 1$ and $Q_1, \ldots, Q_J \geq 1$ satisfying $\sum_{j=1}^J Q_j = Q$,
(b) vectors $y_k^j \in \mathbb{R}^n$ (for $j \in \{1, \ldots, J\}$ and $k \in \mathbb{N}$) with

$$\lim_k |y_k^j - y_k^i| = +\infty \qquad \forall i \neq j, \tag{3.1}$$

(c) and maps $\zeta^j \in W^{1,2}(\Omega, \mathcal{A}_{Q_j})$ for $j \in \{1, \ldots, J\}$,

such that $h_k = \sum_{j=1}^J [\![\tau_{y_k^j} \circ \zeta^j]\!]$, where for any generic $y \in \mathbb{R}^n$ we denote by $\tau_y : \mathcal{A}_Q(\mathbb{R}^n) \to \mathcal{A}_Q(\mathbb{R}^n)$ the translation map (cp. [14, Section 3.3.3])

$$\mathcal{A}_Q(\mathbb{R}^n) \ni T = \sum_i [\![P_i]\!] \mapsto \tau_y(T) := \sum_i [\![P_i - y]\!] \in \mathcal{A}_Q(\mathbb{R}^n).$$

Remark 3.2. Assume that h_k, Q_j, y_k^j and ζ^k satisfy all the requirements of Definition 3.1 except for (3.1). Up to subsequences and relabellings, assume that $y_k^1 - y_k^2$ converges to a vector $2\bar{y}$. We can replace

- the integers Q_1 and Q_2 with $Q' = Q_1 + Q_2$;
- the vectors y_k^1 and y_2^k with $y_k' = (y_k^1 + y_k^2)/2$;
- the maps ζ^1 and ζ^2 with $\zeta' := [\![\tau_{\bar{y}} \circ \zeta^1]\!] + [\![\tau_{-\bar{y}} \circ \zeta^2]\!]$.

The new collections $Q', Q_3, \ldots, Q_J$, $y_k', y_k^3, \ldots, y_k^J$ and $\zeta', \zeta^3, \ldots, \zeta^J$, and the function $h_k' := [\![\zeta']\!] + \sum_{j=3}^J [\![\zeta^j]\!]$, satisfy again all the requirements of Definition 3.1 except, possibly, for (3.1). Moreover, $\|\mathcal{G}(h_k', h_k)\|_{L^2} \to 0$ and $|Dh_k'| = |Dh_k|$. Obviously, we can iterate this procedure only a finite number of times, obtaining a subsequence of translating sheets $\hat{h}_k$ asymptotic to h_k in the L^2 distance with $|D\hat{h}_k| = |Dh_k|$.

3.1. Concentration compactness.

Translating sheets give a useful device to recover a suitable "compactness statement" for sequences of maps with equi-bounded energy.

Proposition 3.3 (Concentration compactness)**.** *Let $\Omega \subset \mathbb{R}^m$ be a Lipschitz bounded open set and $(g_k)_{k\in\mathbb{N}} \subset W^{1,2}(\Omega, \mathcal{A}_Q)$ a sequence of functions with $\sup_k \int_\Omega |Dg_k|^2 < \infty$. Then, there exist a subsequence (not relabeled) and a sequence of translating sheets h_k such that $\|\mathcal{G}(g_k, h_k)\|_{L^2} \to 0$ and the following inequalities hold for every open $\Omega' \subset \Omega$ and any* *sequence of measurable sets J_k with $|J_k| \to 0$:*

$$\liminf_{k\to+\infty}\left(\int_{\Omega'\setminus J_k}|Dg_k|^2-\int_{\Omega'}|Dh_k|^2\right)\geq 0 \tag{3.2}$$

$$\limsup_{k\to+\infty}\int_\Omega (|Dg_k|-|Dh_k|)^2\leq \limsup_k \int_\Omega \left(|Dg_k|^2-|Dh_k|^2\right)\,. \tag{3.3}$$

Proof. We start proving, by induction on Q, the existence of translating sheets $\{h_k\}$ (and a subsequence) with $\|\mathcal{G}(h_k,g_k)\|_{L^2}\to 0$ and satisfying the following additional property. If J, Q_j, y^j_k and ζ^j are as in Definition 3.1, then there are Q_j valued functions w^j_k such that, after setting $f_k=\sum_j [\![w^j_k]\!]$, we have

$$\|\mathcal{G}(f_k,g_k)\|_{L^2}+|\{g_k\neq f_k\}|\to 0,\quad \|\mathcal{G}(\tau_{-y^j_k}\circ w^j_k,\zeta^j)\|_{L^2}\to 0\quad\text{and}\quad |Df_k|\leq|Dg_k|\,. \tag{3.4}$$

If $Q=1$ the claim with $f_k=g_k$ is an easy corollary of the Poincaré inequality and the compact embedding $W^{1,2}\hookrightarrow L^2$. Assuming that the claim holds for any $Q^*<Q$, we prove it for Q. By the generalized Poincaré inequality [14, Proposition 2.12], there exist points $\bar g_k\in\mathcal{A}_Q(\mathbb{R}^n)$ and a real number M such that

$$\int_\Omega \mathcal{G}(g_k,\bar g_k)^2\leq C\int_\Omega |Dg_k|^2\leq M<\infty\qquad\forall\, k\in\mathbb{N}\,.$$

Recall the separation $s(T)$ and the diameter $d(T)$ of a point $T=\sum_i [\![P_i]\!]$ introduced in [14, Definition 3.4]: $s(T):=\min\left\{|P_i-P_j| : P_i\neq P_j\right\}$ and $d(T):=\max\{|P_i-P_j|\}$. We distinguish between to cases.

Case 1: $\liminf_k d(\bar g_k)<\infty$. After passing to a subsequence, we find $y_k\in\mathbb{R}^n$ such that the functions $\tau_{y_k}\circ g_k$ are equi-bounded in the $W^{1,2}$-metric. By the Sobolev embedding [14, Proposition 2.11], there exists a Q-valued map $\zeta\in W^{1,2}$ such that $\tau_{y_k}\circ g_k\to\zeta$ in $L^2(\Omega)$.

Case 2: $\lim_k d(\bar g_k)=+\infty$. By [14, Lemma 3.8] there are points $S_k\in\mathcal{A}_Q$ such that

$$\beta\, d(\bar g_k)\leq s(S_k)<+\infty\quad\text{and}\quad \mathcal{G}(S_k,\bar g_k)\leq s(S_k)/32,$$

where β is a dimensional constant. Write $S_k=\sum_{i=1}^J \kappa_i [\![P^i_k]\!]$, with $P^i_k\neq P^j_k$ for $i\neq j$. Both J and κ_i may depend on k but they have a finite range: therefore, after extracting a subsequence, we can assume that they do not depend on k. Set next $r_k=\frac{s(S_k)}{16}$ and let ϑ_k be the retraction of $\mathcal{A}_Q(\mathbb{R}^n)$ into $\overline{B_{r_k}(S_k)}$ provided by [14, Lemma 3.7]. Clearly, the functions $\hat f_k=\vartheta_k\circ g_k$ satisfy $|D\hat f_k|\leq|Dg_k|$ and there are κ_i-valued functions z^i_k such that

$$\hat f_k=\sum_{i=1}^J [\![z^i_k]\!]\,,\quad\text{with}\quad \|\mathcal{G}(z^i_k,\kappa_i\,[\![P^i_k]\!])\|_\infty\leq r_k.$$

Since $\kappa_i<Q$, we apply the inductive hypothesis to each sequence $(z^i_k)_k$ and, using Remark 3.2 reach a subsequence (not relabeled) of $\hat f_k$, a sequence of translating sheets h_k and corresponding functions f_k which satisfy (3.4) with $\hat f_k$ replacing g_k.

We next claim that (3.4) holds even for g_k, i.e. that $\lim_k \left(\|\mathcal{G}(f_k, g_k)\|_{L^2} + |\{f_k \neq g_k\}|\right) = 0$. To this aim, recall first that

$$\left\{g_k \neq \hat{f}_k\right\} = \{\mathcal{G}\,(g_k, S_k) > r_k\} \subseteq \{\mathcal{G}\,(g_k, \bar{g}_k) > r_k/2\}\,.$$

Thus,

$$\left|\left\{g_k \neq \hat{f}_k\right\}\right| \leq |\,\{\mathcal{G}\,(g_k, \bar{g}_k) > r_k/2\}\,| \leq \frac{C}{r_k^2}\int_{\left\{\mathcal{G}(g_k,\bar{g}_k)>\frac{r_k}{2}\right\}} \mathcal{G}\,(g_k, \bar{g}_k)^2 \leq \frac{CM}{(d(\bar{g}_k))^2}. \tag{3.5}$$

Since $d(\bar{g}_k) \to +\infty$ and (3.4) holds with $\hat{f}_k$ replacing g_k, we conclude $|\{f_k \neq g_k\}| \to 0$. Next, since $\vartheta_k(\bar{g}_k) = \bar{g}_k$ and $\mathrm{Lip}(\vartheta_k) = 1$, we have $\mathcal{G}(\hat{f}_k, \bar{g}_k) \leq \mathcal{G}(g_k, \bar{g}_k)$. Therefore, by the Sobolev embedding and the Poincaré inequality, for any $p \in]2, 2^*[$, we infer

$$\begin{aligned}
&\int_\Omega \mathcal{G}(\hat{f}_k, g_k)^2 = \int_{\{g_k\neq\hat{f}_k\}} \mathcal{G}(\hat{f}_k, g_k)^2 \leq 2\int_{\{\hat{f}_k\neq g_k\}} \mathcal{G}(\hat{f}_k, \bar{g}_k)^2 + 2\int_{\{\hat{f}_k\neq g_k\}} \mathcal{G}(\bar{g}_k, g_k)^2\\
&\leq 4\int_{\{\hat{f}_k\neq g_k\}} \mathcal{G}(\bar{g}_k, g_k)^2 \leq C\,\|\mathcal{G}\,(g_k, \bar{g}_k)\|_{L^p}^2 \left|\left\{\hat{f}_k \neq g_k\right\}\right|^{1-\frac{2}{p}} \stackrel{(3.5)}{\leq} \frac{CM^{1-2/p}}{d(\bar{g}_k)^{2-4/p}}\int_\Omega |Dg_k|^2.
\end{aligned}$$

Since $d(\bar{g}_k)$ diverges, this shows $\|\mathcal{G}(\hat{f}_k, g_k)\|_{L^2} \to 0$ and by inductive hypothesis that $\|\mathcal{G}(f_k, g_k)\|_{L^2} \to 0$.

We now show that (3.2) and (3.3) are consequences of (3.4). For each j we consider the corresponding embedding $\boldsymbol{\xi}_j : \mathcal{A}_{Q_j}(\mathbb{R}^n) \to \mathbb{R}^{N(Q_j,n)}$ and, by a slight abuse of notation, we drop the j subscript. Then, we conclude that $\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j \to \boldsymbol{\xi}\circ\zeta^j$ in L^2 and $\|D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)\|_{L^2}$ is a bounded sequence, from which

$$D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j) \rightharpoonup D(\boldsymbol{\xi}\circ\zeta^j) \quad \text{in } L^2(\Omega)\,. \tag{3.6}$$

If J_k is a sequence of measurable sets with $|J_k| \downarrow 0$, then $\mathbf{1}_{\Omega'\setminus J_k} \to \mathbf{1}_{\Omega'}$ in $L^2(\Omega)$ and it follows from (3.6) that

$$D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)\mathbf{1}_{\Omega'\setminus J_k} \rightharpoonup D(\boldsymbol{\xi}\circ\zeta^j)\mathbf{1}_{\Omega'} \quad \text{in } L^2(\Omega)\,,$$

and, hence,

$$\mathrm{Dir}(\zeta^j, \Omega') = \int_{\Omega'} |D(\boldsymbol{\xi}\circ\zeta^j)|^2 \leq \liminf_k \int_{\Omega'\setminus J_k} |D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)|^2 = \liminf_k \int_{\Omega'\setminus J_k} |Dw_k^j|^2.$$

Summing over j, we obtain (3.2). As for (3.3), set $J_k := \{g_k \neq f_k\}$. Thus,

$$\begin{aligned}
&\int_{\Omega\setminus J_k} (|Dg_k| - |Dh_k|)^2 \leq \sum_j \int_{\Omega\setminus J_k} (|Dw_k^j| - |D\zeta^j|)^2\\
&= \sum_j \int_{\Omega\setminus J_k} \left(|D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)| - |D(\boldsymbol{\xi}\circ\zeta^j)|\right)^2 \leq \sum_j \int_{\Omega\setminus J_k} |D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j) - D(\boldsymbol{\xi}\circ\zeta^j)|^2\\
&= \sum_j \int_{\Omega\setminus J_k} \Big(|D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)|^2 + |D(\boldsymbol{\xi}\circ\zeta^j)|^2 - 2\,D(\boldsymbol{\xi}\circ\tau_{-y_k^j}\circ w_k^j)\cdot D(\boldsymbol{\xi}\circ\zeta^j)\Big)\,.
\end{aligned} \tag{3.7}$$

Therefore, by (3.6) (and taking into account that $|J_k| \to 0$) one gets

$$\begin{aligned}
&\limsup_{k\to+\infty} \int_{\Omega\setminus J_k} (|Dg_k| - |Dh_k|)^2 \\
&\leq \lim_{k\to+\infty} \sum_j \int_{\Omega\setminus J_k} \Big(|D(\boldsymbol{\xi} \circ \tau_{-y_k^j} \circ w_k^j)|^2 + |D(\boldsymbol{\xi}\circ\zeta^j)|^2 - 2\, D(\boldsymbol{\xi} \circ \tau_{-y_k^j} \circ w_k^j) \cdot D(\boldsymbol{\xi}\circ\zeta^j) \Big) \\
&= \limsup_{k\to+\infty} \int_{\Omega\setminus J_k} \sum_j |D(\boldsymbol{\xi} \circ \tau_{-y_k^j} \circ w_k^j)|^2 - \int_\Omega \sum_j |D(\boldsymbol{\xi}\circ\zeta^j)|^2 \\
&= \limsup_{k\to+\infty} \int_{\Omega\setminus J_k} |Dg_k|^2 - \int_\Omega |Dh_k|^2. \qquad (3.8)
\end{aligned}$$

On the other hand, since $|J_k| \to 0$ we conclude

$$\limsup_{k\to\infty} \int_{J_k} (|Dg_k| - |Dh_k|)^2 = \limsup_{k\to\infty} \int_{J_k} |Dg_k|^2 .$$

Observe that, after passing to a subsequence, we can actually assume that all limsups are in fact limits. Summing (3.8) and the last equation we then conclude (3.3). □

3.2. **Dirichlet competitors.** We consider next a standard procedure to construct competitors for the Dirichlet energy of a sequence of functions with equi-bounded energy.

Proposition 3.4 (Construction of a competitor)**.** *Consider two radii $1 \leq r_0 < r_1 < 4$ and maps $g_k, h_k \in W^{1,2}(B_{r_1}, \mathcal{A}_Q(\mathbb{R}^n))$ such that $\{h_k\}_k$ is a sequence of translating sheets,*

$$\sup_k \mathrm{Dir}(g_k, B_{r_1}) < +\infty \quad \textit{and} \quad \|\mathcal{G}(g_k, h_k)\|_{L^2(B_{r_1}\setminus B_{r_0})} \to 0.$$

For every $\eta > 0$, there exist $r \in]r_0, r_1[$, a subsequence of $\{g_k\}_k$ (not relabeled) and functions $H_k \in W^{1,2}(B_{r_1}, \mathcal{A}_Q(\mathbb{R}^n))$ such that $H_k|_{B_{r_1}\setminus B_r} = g_k|_{B_{r_1}\setminus B_r}$ and $\mathrm{Dir}(H_k, B_{r_1}) \leq \mathrm{Dir}(h_k, B_{r_1}) + \eta$. In addition, there is a dimensional constant C and a constant C^* (depending on η and the two sequences, but not on *k) such that*

$$\mathrm{Lip}(H_k) \leq C^* \,(\mathrm{Lip}(g_k) + 1), \qquad (3.9)$$

$$\|\mathcal{G}(H_k, h_k)\|_{L^2(B_r)} \leq C\mathrm{Dir}(g_k, B_r) + C\mathrm{Dir}(H_k, B_r)\,, \qquad (3.10)$$

$$\|\boldsymbol{\eta}\circ H_k\|_{L^1(B_{r_1})} \leq C^* \,\|\boldsymbol{\eta}\circ g_k\|_{L^1(B_{r_1})} + C\|\boldsymbol{\eta}\circ h_k\|_{L^1(B_{r_1})}\,. \qquad (3.11)$$

In order to prove the proposition, we need to recall the following two lemmas, which are slight variants of [14, Proposition 4.4] and [14, Lemma 2.15].

Lemma 3.5 (Lipschitz approximation)**.** *Let $f \in W^{1,2}(B_r, \mathcal{A}_Q)$. Then, for every $\varepsilon > 0$, there exists $f_\varepsilon \in \mathrm{Lip}(B_r, \mathcal{A}_Q)$ such that*

$$\int_{B_r} \mathcal{G}(f, f_\varepsilon)^2 + \int_{B_r} \big(|Df| - |Df_\varepsilon|\big)^2 + \int_{B_r} \big(|D(\boldsymbol{\eta}\circ f)| - |D(\boldsymbol{\eta}\circ f_\varepsilon)|\big)^2 \leq \varepsilon. \qquad (3.12)$$

If $f|_{\partial B_r} \in W^{1,2}(\partial B_r, \mathcal{A}_Q)$, then f_ε can be chosen to satisfy also

$$\int_{\partial B_r} \mathcal{G}(f, f_\varepsilon)^2 + \int_{\partial B_r} \big(|Df| - |Df_\varepsilon|\big)^2 \leq \varepsilon. \tag{3.13}$$

Proof. By an obvious scaling argument we can assume $r = 1$. We start noticing that (3.12) is a corollary of [14, Proposition 4.4]. On the other hand, if $f|_{\partial B_1} \in W^{1,2}(\partial B_1)$, we extend the map to B_2 by setting $f(x) = f(\frac{x}{|x|})$ if $|x| \geq 1$. We then can apply [14, Proposition 2.5] to find a sequence of Lipschitz maps f_k such that $f_k \to f$ strongly in $W^{1,2}(B_2)$. Given $\delta > 0$, define the maps $f^\delta(x) = f((1+\delta)x)$ and $f^\delta_k(x) = f_k((1+\delta)x)$. Obviously, $f^\delta_k \to f^\delta$ strongly in $W^{1,2}(B_1)$ and $f^\delta \to f$ strongly in $W^{1,2}(B_1)$ as $\delta \downarrow 0$. By a standard Fubini argument, for each j we can find a $\delta_j < \frac{1}{j}$ and a subsequence $\{f_{k,j}\}_k$ such that $f_{k,j}|_{\partial B_{1+\delta_j}} \to f|_{\partial B_{1+\delta_j}}$ (i.e. $f^{\delta_j}_{k,j}|_{\partial B_1} \to f^{\delta_j}|_{\partial B_1} = f|_{\partial B_1}$) strongly in $W^{1,2}(\partial B_{1+\delta_j})$ as $k \uparrow \infty$. By standard diagonal argument we can arrange the subsequences so that $\{f_{k,j}\} \supset \{f_{k,j+1}\}$. Thus, a suitable diagonal sequence $\bar f_j := f^{\delta_j}_{k(j),j}$ has the property that $\bar f_j \to f$ in $W^{1,2}(B_1)$ and $\bar f_j|_{\partial B_1} \to f|_{\partial B_1}$ in $W^{1,2}(\partial B_1)$. □

Lemma 3.6 (Interpolation)**.** *There exists a constant $C_0 = C_0(m,n,Q) > 0$ with the following property. Assume $r \in]1,3[$, $f \in W^{1,2}(B_r, \mathcal{A}_Q)$ and $g \in W^{1,2}(\partial B_r, \mathcal{A}_Q)$ are given maps such that $f|_{\partial B_r} \in W^{1,2}(\partial B_r, \mathcal{A}_Q)$. Then, for every $\varepsilon \in]0, r[$ there exists a function $h \in W^{1,2}(B_r, \mathcal{A}_Q)$ such that $h|_{\partial B_r} = g$ and*

$$\int_{B_r} |Dh|^2 \leq \int_{B_r} |Df|^2 + \varepsilon \int_{\partial B_r} \big(|D_\tau f|^2 + |D_\tau g|^2\big) + \frac{C_0}{\varepsilon} \int_{\partial B_r} \mathcal{G}(f,g)^2\,, \tag{3.14}$$

$$\mathrm{Lip}(h) \leq C_0 \left\{ \mathrm{Lip}(f) + \mathrm{Lip}(g) + \varepsilon^{-1} \sup_{\partial B_r} \mathcal{G}(f,g) \right\}\,, \tag{3.15}$$

$$\int_{B_r} |\boldsymbol{\eta} \circ h| \leq C_0 \int_{\partial B_r} |\boldsymbol{\eta} \circ g| + C_0 \int_{B_r} |\boldsymbol{\eta} \circ f|\,, \tag{3.16}$$

(here D_τ denotes the tangential derivative).

Proof. The first conclusion is an obvious corollary of [14, Lemma 2.15]. It is then straightforward to see that the map constructed in the proof of [14, Lemma 2.15] satisfies also (3.15). As for the final claim, let $\bar g := \sum [\![g_i - \boldsymbol{\eta} \circ g]\!]$, $\bar f := \sum [\![f_i - \boldsymbol{\eta} \circ f]\!]$ and consider the interpolation map $\bar h$ between $\bar f$ and $\bar g$ given by [14, Lemma 2.15]. Set $\hat h = \sum_i [\![\bar h_i - \boldsymbol{\eta} \circ \bar h]\!]$ and observe that $\mathrm{Lip}(\hat h) \leq \mathrm{Lip}(\bar h)$ and $\mathrm{Dir}(\hat h) \leq \mathrm{Dir}(\bar h)$. We apply again [14, Lemma 2.15] in the case $Q = 1$ to $\boldsymbol{\eta} \circ f$ and $\boldsymbol{\eta} \circ g$, and get the interpolation u. It is then easy to check that the map $h := \sum_i [\![\hat h_i + u]\!]$ has all the desired properties. □

Proof of Proposition 3.4. Set for simplicity $A_k := \|\mathcal{G}(g_k, h_k)\|_{L^2(B_{r_1} \setminus B_{r_0})}$ and $B_k := \|\boldsymbol{\eta} \circ g_k\|_{L^1(B_{r_1})}$. If $A_k \equiv 0$, then there is nothing to prove and so we can assume that, for a subsequence, not relabeled, $A_k > 0$. Assuming that for yet another subsequence (not relabeled) $B_k > 0$, we consider the function

$$\psi_k(r) := \int_{\partial B_r} \big(|Dg_k|^2 + |Dh_k|^2\big) + A_k^{-2} \int_{\partial B_r} \mathcal{G}(g_k, h_k)^2 + B_k^{-1} \int_{\partial B_r} |\boldsymbol{\eta} \circ g_k|. \tag{3.17}$$

By assumption $\liminf_k \int_{r_0}^{r_1} \psi_k(r)\,dr < \infty$. So, by Fatou's Lemma, there is $r \in]r_0, r_1[$ and a subsequence, not relabeled, such that $\lim_k \psi_k(r) < \infty$. Thus, for some $M > 0$ we have

$$\int_{\partial B_r} \mathcal{G}(g_k, h_k)^2 \to 0, \tag{3.18}$$

$$\mathrm{Dir}(h_k, \partial B_r) + \mathrm{Dir}(g_k, \partial B_r) \leq M, \tag{3.19}$$

$$\int_{\partial B_r} |\boldsymbol{\eta} \circ g_k| \leq M \, \|\boldsymbol{\eta} \circ g_k\|_{L^1(B_{r_1})}. \tag{3.20}$$

In case $B_k = 0$ for all k large enough, we define ψ_k dropping the last summand in (3.17) and reach the same conclusion.

Let ζ^j be the blocks of the translating sheets h_k as in Definition 3.1. We apply Lemma 3.5 to each ζ^j and find Lipschitz functions ζ^j_η satisfying the conclusion of the lemma with $\bar\varepsilon_1 = \bar\varepsilon_1(\eta, M) > 0$ (which will be chosen later). We also choose a standard radial convolution kernel φ in $\mathbb{R}^m$ and a small parameter $\bar\rho$ (also to be chosen later). Then, set

$$h_{k,\eta} := \sum_{j=1}^{J} [\![\tau_{y_k^j} \circ \zeta^j_\eta]\!] \quad \text{and} \quad \bar h_{k,\eta} := \sum_{i=1}^{Q} [\![(h_{k,\eta})_i - \boldsymbol{\eta} \circ h_{k,\eta} + (\boldsymbol{\eta} \circ h_k) * \varphi_{\bar\rho}]\!],$$

and choose $\bar\rho$ so small that

$$Q^2 \|\boldsymbol{\eta} \circ h_k - (\boldsymbol{\eta} \circ h_k) * \varphi_{\bar\rho}\|^2_{L^2} \leq \bar\varepsilon_1, \tag{3.21}$$

$$\int_{B_r} \left(|D(\boldsymbol{\eta} \circ h_k)|^2 - |D(\boldsymbol{\eta} \circ h_k * \varphi_{\bar\rho})|^2 \right) \leq \bar\varepsilon_1. \tag{3.22}$$

Note that this is possible because, from the fact that h_k is a sequence of translating sheets, it follows that $\eta \circ h_k(x) = F(x) + p_k$ for some $F \in W^{1,2}$ and a sequence of vectors $p_k \in \mathbb{R}^n$. Therefore $(\eta \circ h_k) * \varphi_{\bar\rho} = F * \varphi_{\bar\rho} + p_k$ and $D(\eta \circ h_k) * \varphi_{\bar\rho} = DF * \varphi_{\bar\rho}$, and (3.21) and (3.22) follows if $\bar\rho$ is sufficiently small by the usual convolution estimates. In particular by very rough estimates,

$$\|\mathcal{G}(g_k, \bar h_{k,\eta})\|_{L^2} \overset{(3.21)}{\leq} \|\mathcal{G}(g_k, h_k)\|_{L^2} + 2\|\mathcal{G}(h_k, h_{k,\eta})\|_{L^2} + \bar\varepsilon_1 \leq o(1) + 3\,\bar\varepsilon_1, \tag{3.23}$$

$$\mathrm{Dir}(\bar h_{k,\eta}, \partial B_r) \leq 2\,M + 2\,\bar\varepsilon_1 \tag{3.24}$$

and

$$\begin{aligned}
\mathrm{Dir}(\bar h_{k,\eta}, B_r) &= \sum_i \int_{B_r} \left|D(h_{k,\eta})_i - D(\boldsymbol{\eta} \circ h_{k,\eta}) + D(\boldsymbol{\eta} \circ h_k * \varphi_{\bar\rho})\right|^2 \\
&= \int_{B_r} \left(|Dh_{k,\eta}|^2 - Q|D(\boldsymbol{\eta} \circ h_{k,\eta})|^2 + Q|D(\boldsymbol{\eta} \circ h_k * \varphi_{\bar\rho})|^2 \right) \\
&= \mathrm{Dir}(h_{k,\eta}, B_r) + Q \int_{B_r} \left(|D(\boldsymbol{\eta} \circ h_k)|^2 - |D(\boldsymbol{\eta} \circ h_{k,\eta})|^2 \right) \\
&\quad + Q \int_{B_r} \left(|D(\boldsymbol{\eta} \circ h_k * \varphi_{\bar\rho})|^2 - |D(\boldsymbol{\eta} \circ h_k)|^2 \right)
\end{aligned}$$

$$\stackrel{(3.12),(3.22)}{\leq} \mathrm{Dir}(h_{k,\eta}, B_r) + 2\,Q\,\bar\varepsilon_1. \tag{3.25}$$

We can then apply Lemma 3.6 to $\bar h_{k,\eta}$ and g_k with $\bar\varepsilon_2 = \bar\varepsilon_2(\eta, M) > 0$, and get (up to subsequences) maps H_k satisfying $H_k|_{\partial B_r} = g_k|_{\partial B_r}$ and

$$\begin{aligned}\mathrm{Dir}\,(H_k, B_r) &\leq \mathrm{Dir}\left(\bar h_{k,\eta}, B_r\right) + \bar\varepsilon_2\,\mathrm{Dir}\left(\bar h_{k,\eta}, \partial B_r\right) + \bar\varepsilon_2\,\mathrm{Dir}(g_k, \partial B_r) + \frac{C_0}{\bar\varepsilon_2}\int_{\partial B_r} \mathcal{G}\left(\bar h_{k,\eta}, g_k\right)^2\\ &\leq \mathrm{Dir}(h_k, B_r) + Q\bar\varepsilon_1 + 3\,\bar\varepsilon_2\,(M + \bar\varepsilon_1) + 3\,C_0\,\bar\varepsilon_2^{-1}\bar\varepsilon_1\end{aligned}$$

where in the last line we have used (3.18), (3.19) and (3.23) - (3.25). An appropriate choice of the parameters ε_1 and ε_2 gives the desired bound $\mathrm{Dir}\,(H_k, B_r) \leq \mathrm{Dir}(h_k, B_r) + \eta$.

Observe next that, by construction, $\limsup_k \mathrm{Lip}(\bar h_{k,\eta}) \leq C^*$, for some constant which depends on η and the two sequences, but not on k. Moreover,

$$\|\mathcal{G}(\bar h_{k,\eta}, g_k)\|_{L^\infty(\partial B_r)} \leq \|\mathcal{G}(\bar h_{k,\eta}, g_k)\|_{L^2(\partial B_r)} + C\mathrm{Lip}(g_k) + C\mathrm{Lip}(\bar h_{k,\eta})\,.$$

Thus (3.9) follows from (3.15).

Finally, (3.10) follows from the Poincaré inequality applied to $\mathcal{G}(H_k, g_k)$ (which vanishes identically on ∂B_r), and (3.11) follows from (3.16), because of (3.20) and $\|\boldsymbol{\eta}\circ\bar h_{k,\eta}\|_{L^1(B_r)} = \|(\boldsymbol{\eta}\circ h_k)*\varphi_{\bar\rho}\|_{L^1(B_r)} \leq \|\boldsymbol{\eta}\circ h_k\|_{L^1(B_{r_1})}$ if $\bar\rho$ is also chosen small enough such that $r+\bar\rho < r_1$. □

4. Harmonic approximation

In what follows we will always apply Proposition 2.2 with $\delta_{11} = E^{2\beta}$ and under a certain scaling of $\mathbf{A}$.

Definition 4.1 (E^β-Lipschitz approximation)**.** Let $\beta \in \left(0, \frac{1}{2m}\right)$, T be as in Proposition 2.2 such that $32E^{(1-2\beta)/m} < 1$ and $s\mathbf{A} \leq E^{1/4+\delta}$ for some $\delta > 0$. If the coordinates are fixed as in Remark 1.5, the map u given by Proposition 2.2 for $\delta_{11} = E^{2\beta}$ is then called the *E^β-Lipschitz approximation of T in* $\mathbf{C}_{3s}(x)$ and will be denoted by f.

In this section we prove that, if T is also area minimizing, the corresponding E^β-Lipschitz approximation is close to a Dir-minimizing function w. This comes with an $o(E)$-improvement of the estimates in Proposition 2.2.

Theorem 4.2 (First harmonic approximation)**.** *For every $\eta_1, \delta > 0$ and every $\beta \in (0, \frac{1}{2m})$, there exist constants $\varepsilon_{12}, C_{12} > 0$ with the following property. Let T be as in Assumption 1.1 in $\mathbf{C}_{4s}(x)$ and assume it is area minimizing. If $E = \mathbf{E}(T, \mathbf{C}_{4s}(x)) \leq \varepsilon_{12}$ and $s\mathbf{A} \leq E^{1/4+\delta}$, then the E^β-Lipschitz approximation f in $\mathbf{C}_{3s}(x)$ satisfies*

$$\int_{B_{2s}(x)\setminus K} |Df|^2 \leq \eta_1 E\,\omega_m\,(4\,s)^m = \eta_1\,\mathbf{e}_T(B_{4s}(x)). \tag{4.1}$$

Moreover, if we consider the coordinates of Remark 1.5, there exists a Dir*-minimizing function* $u : B_{2s}(x) \to \mathcal{A}_Q(\mathbb{R}^{\bar n})$ *such that the map* $B_{2s}(x) \ni y \mapsto w = (u, \Psi(y,u))$ *satisfies*

$$s^{-2}\int_{B_{2s}(x)} \mathcal{G}(f,w)^2 + \int_{B_{2s}(x)} \big(|Df| - |Dw|\big)^2 \le \eta_1 E\,\omega_m\,(4\,s)^m = \eta_1\,\mathbf{e}_T(B_{4s}(x))\,, \tag{4.2}$$

$$\int_{B_{2s}(x)} |D(\boldsymbol{\eta}\circ f) - D(\boldsymbol{\eta}\circ w)|^2 \le \eta_1 E\,\omega_m\,(4\,s)^m = \eta_1\,\mathbf{e}_T(B_{4s}(x))\,. \tag{4.3}$$

Remark 4.3 (Isoperimetric inequality)**.** If $S \subset \mathbb{R}^{m+n}$ is an integral current of dimension $m-1$ with $\partial S = 0$, then there is an m-dimensional integral current $R \subset \mathbb{R}^{m+n}$ such that $\partial R = S$ and $\mathbf{M}(R) \le C\mathbf{M}(S)^{m/(m-1)}$, where the constant C is only dimensional (see [29, Theorem 30.1]). It is also well-known that, when $\mathrm{spt}(S) \subset \Sigma$ and Σ is as in Assumption 1.1 the same inequality holds for some $\bar R$ with $\mathrm{spt}(\bar R) \subset \Sigma$ and $\partial \bar R = S$, with a dimensional constant C which depends additionally on the constant c_0. This can be easily seen as follows: let $\mathbf{q} : \mathbb{R}^{m+n} \to \mathbb{R}^{m+\bar n}$ be the orthogonal projection and $\Lambda : \mathbb{R}^{m+n} \to \Sigma$ be the map $\Lambda(p) = (\mathbf{q}(p), \Psi(\mathbf{q}(p)))$. Λ is a global Lipschitz retraction of $\mathbb{R}^{m+n}$ onto Σ which is the identity on Σ: thus we can simply set $\bar R = \Lambda_\sharp R$.

Remark 4.4 (Taylor expansion of the mass)**.** There are dimensional constants $c, C > 0$ such that the following holds. Let $V \subset \mathbb{R}^m$ be a bounded measurable set and let $u : V \to \mathcal{A}_Q(\mathbb{R}^n)$ be a Lipschitz function with $\mathrm{Lip}(u) \le c$. Denote by $\mathbf{G}_u$ the integer rectifiable current associated to the graph of u as in [17, Definition 1.10]. Then, the following Taylor expansion of the mass of $\mathbf{G}_u$ holds:

$$\mathbf{M}(\mathbf{G}_u) = Q\,|V| + \int_V \frac{|Du|^2}{2} + \int_V \sum_i R(Du_i),$$

where $R : \mathbb{R}^{n\times m} \to R$ is a C^1 function satisfying $|R(D)| = |D|^3\,L(D)$ for some positive function L such that $L(0) = 0$ and $\mathrm{Lip}(L) \le C$. This Taylor expansion is proven in [17, Corollary 3.3] (although the corollary is stated for V open, the proof works obviously when V is merely measurable).

Remark 4.5. There exists a dimensional constant $c > 0$ such that, if $E \le c$, then the E^β-Lipschitz approximation satisfies the following estimates:

$$\mathrm{Lip}(f) \le C\,E^\beta, \tag{4.4}$$

$$\int_{B_{3s}(x)} |Df|^2 \le C\,E\,s^m. \tag{4.5}$$

Indeed (4.4) follows from Proposition 2.2, Remark 1.5 and $\|D\Psi\|_0 \le C(E^{1/2} + \mathbf{A}) \le C\,E^\beta$ by the choice of β and the scaling of $\mathbf{A}$. While (4.5) follows from Remark 4.4 since for E sufficiently small

$$\int_{B_{3s}(x)} \sum_i R(Df_i) \le C\,E^{2\beta}\int_{B_{3s}(x)} |Df|^2 < \frac{1}{4}\int_{B_{3s}(x)} |Df|^2,$$

and therefore

$$\begin{aligned}\int_{B_{3s}(x)} |Df|^2 \leq & C\,(\mathbf{M}(\mathbf{G}_f \llcorner \mathbf{C}_{3s}(x)) - Q\,\omega_m\,(3\,s)^m)\\ \leq & C\,(\mathbf{M}(T \llcorner \mathbf{C}_{3s}(x)) - Q\,\omega_m\,(3\,s)^m) + C\,\mathbf{M}(\mathbf{G}_f \llcorner (B_{3s}(x)\setminus K)\times\mathbb{R}^n)\\ \leq & C\,E\,s^m + C\,E^{2\beta}\,|B_{3s}(x)\setminus K| \leq C\,E\,s^m.\end{aligned}$$

(4.5) is therefore a rather simple corollary of the "maximal function truncation" argument employed in Proposition 2.2. Other approximation schemes give instead worse bounds for the Lipschitz constant of the approximating map, cf. for instance [29, Theorem 5.1.1].

Proof of Theorem 4.2. By rescaling and translating, it is not restrictive to assume that $x = 0$ and $s = 1$. Thus, by Remark 1.5 we can assume $\Psi(0) = 0$, $\|D\Psi\|_0 \leq C(E^{1/2} + \mathbf{A})$ and $\|D^2\Psi\|_0 \leq \mathbf{A}$. The proof of (4.1) is by contradiction. Assume there exist a constant $c_1 > 0$, a sequence of currents $(T_k)_{k\in\mathbb{N}}$ satisfying Assumption 1.1 and area minimizing, ambient manifolds Σ_k (parametrized by Ψ_k, with second fundamental forms bounded by $\mathbf{A}_k$) and corresponding E_k^β-Lipschitz approximations $(f_k)_{k\in\mathbb{N}}$ such that

$$E_k := \mathbf{E}(T_k, \mathbf{C}_4) \to 0\,, \quad \mathbf{A}_k \leq E_k^{1/4+\delta} \quad \text{and} \quad \int_{B_2\setminus K_k} |Df_k|^2 \geq c_1\,E_k, \tag{4.6}$$

where $K_k := \{x \in B_3 : \mathbf{me}_{T_k}(x) < E_k^{2\beta}\}$. Set $\Gamma_k := \{x \in B_4 : \mathbf{me}_{T_k}(x) \leq 2^{-m}E_k^{2\beta}\}$ and observe that $\Gamma_k \cap B_3 \subset K_k$. From Proposition 2.2, it follows that

$$\mathrm{Lip}(f_k) \leq C E_k^\beta, \tag{4.7}$$

$$|B_r \setminus K_k| \leq C E_k^{-2\beta}\mathbf{e}_T\big(B_{r+r_0(k)}\setminus\Gamma_k\big) \quad \text{for every } r \leq 3\,, \tag{4.8}$$

where $r_0(k) = 16\,E_k^{(1-2\beta)/m} < \frac{1}{2}$. We also assume

$$\Psi_k(0) = 0 \qquad \text{and} \qquad \|D\Psi_k\|_0 + \|D^2\Psi_k\|_0 \leq C E_k^{1/4+\delta}\,. \tag{4.9}$$

Then, (4.6), (4.7) and (4.8) give

$$c_1\,E_k \leq \int_{B_2\setminus K_k} |Df_k|^2 \leq C\,\mathbf{e}_{T_k}(B_s\setminus\Gamma_k) \quad \forall\, s\in\left[\tfrac{5}{2}, 3\right].$$

Setting $c_2 := c_1/(2C)$, we have $2c_2E_k \leq \mathbf{e}_{T_k}(B_s\setminus\Gamma_k) = \mathbf{e}_{T_k}(B_s) - \mathbf{e}_{T_k}(B_s\cap\Gamma_k)$, thus leading to

$$\mathbf{e}_{T_k}(\Gamma_k\cap B_s) \leq \mathbf{e}_{T_k}(B_s) - 2\,c_2\,E_k. \tag{4.10}$$

Next observe that $\omega_m 4^m E_k = \mathbf{e}_{T_k}(B_4) \geq \mathbf{e}_{T_k}(B_s)$. Therefore, by the Taylor expansion in Remark 4.4, (4.10) and $E_k \downarrow 0$, it follows that, for every $s \in [5/2, 3]$,

$$\begin{aligned}\int_{\Gamma_k\cap B_s}\frac{|Df_k|^2}{2} &\leq (1 + C\,E_k^{2\beta})\,\mathbf{e}_{T_k}(\Gamma_k\cap B_s)\\ &\leq (1 + C\,E_k^{2\beta})\left(\mathbf{e}_{T_k}(B_s) - 2\,c_2\,E_k\right) \leq \mathbf{e}_{T_k}(B_s) - c_2\,E_k.\end{aligned} \tag{4.11}$$

Our aim is to show that (4.11) contradicts the minimizing property of T_k. To construct a competitor we write $f_k(x) = \sum_i [\![f_k^i(x)]\!] \in \mathcal{A}_Q(\mathbb{R}^{\bar n}\times\mathbb{R}^l)$, and denote by $(f_k^i)'(x)$ the first

$\bar n$ components of the points $f_k^i(x)$. This induces a map $f_k' := \sum_i [\![(f_k^i)']\!]$ taking values into $\mathcal{A}_Q(\mathbb{R}^{\bar n})$. Observe that, since $f_k^i(x)$ are indeed points of the manifold Σ_k

$$f_k = \sum_i [\![((f_k^i)'(x), \Psi_k(x,(f_k^i)'(x)))]\!] \,. \tag{4.12}$$

We consider $g_k := {E_k}^{-1/2} f_k'$. Since by Remark 4.5 $\sup_k \mathrm{Dir}(g_k, B_3) < \infty$ and $|B_3 \setminus \Gamma_k| \to 0$, by Proposition 3.3 we can find a subsequence (not relabelled) of translating sheets h_k satisfying (3.2) - (3.3) and $\|\mathcal{G}(g_k,h_k)\|_{L^2(B_3)} \to 0$. In particular, we are in the position to apply Proposition 3.4 to g_k and h_k, with $r_0 = \frac{5}{2}$, $r_1 = 3$ and $\eta = \frac{c_2}{4}$, and find $r \in \left(\frac{5}{2}, 3\right)$ and competitor functions H_k satisfying $H_k|_{B_3\setminus B_r} = g_k|_{B_3\setminus B_r}$,

$$\mathrm{Dir}(H_k, B_r) \le \mathrm{Dir}(h_k, B_r) + \frac{c_2}{4}, \tag{4.13}$$

$$\mathrm{Lip}(H_k) \le C^*\, E_k^{\beta - 1/2} \tag{4.14}$$

$$\|\mathcal{G}(H_k, g_k)\|_{L^2(B_r)} \le C^*\, \mathrm{Dir}(g_k, B_r) + C\, \mathrm{Dir}(H_k, B_r) \le M < \infty. \tag{4.15}$$

Moreover, Proposition 3.3 implies that, for k is large enough,

$$\mathrm{Dir}(h_k, B_r) \le \mathrm{Dir}(g_k, B_r \cap \Gamma_k) + \frac{c_2}{4} \overset{(4.11)}{\le} \frac{\mathbf{e}_{T_k}(B_r)}{E_k} - \frac{3c_2}{4} E_k \,. \tag{4.16}$$

Note that (4.14) follows from (3.9) observing that $E_k^{\beta-1/2} \uparrow \infty$: thus C^* depends on c_2 and the two chosen sequences, but not on k. From now on, although this and similar constants are not dimensional, we will keep denoting them by C, with the understanding that they do not depend on k. Note that, from (4.7) and (4.8), one gets

$$\begin{aligned}\|T_k - \mathbf{G}_{f_k}\|(\mathbf{C}_3) &\le \|T_k\|((B_3 \setminus K_k)\times \mathbb{R}^n) + \|\mathbf{G}_{f_k}\|((B_3 \setminus K_k)\times\mathbb{R}^n)\\ &\le Q\,|B_3\setminus K_k| + E_k + Q\,|B_3 \setminus K_k| + C\,|B_3\setminus K_k|\,\mathrm{Lip}(f_k)\\ &\le E_k + C\,E_k^{1-2\beta} \le C\,E_k^{1-2\beta}. \end{aligned} \tag{4.17}$$

Let (z,y) be coordinates on $\mathbb{R}^m \times \mathbb{R}^n$ and consider the function $\varphi(z,y) = |z|$ and the slice $\langle T_k - \mathbf{G}_{f_k}, \varphi, r\rangle$. Observe that, by the coarea formula and Fatou's Lemma,

$$\int_r^3 \liminf_k E_k^{2\beta-1} \mathbf{M}(\langle T_k - \mathbf{G}_{f_k}, \varphi, s\rangle)\, ds \le \liminf_k E_k^{2\beta-1} \|T_k - \mathbf{G}_{f_k}\|(\mathbf{C}_3) \le C\,.$$

Therefore, for some $\bar r \in (r,3)$ and a subsequence, not relabeled, $\mathbf{M}\big(\,\langle T_k - \mathbf{G}_{f_k}, \varphi, \bar r\rangle\,\big) \le C\,E_k^{1-2\beta}$.

Let now $v_k := {E_k}^{1/2}\, H_k|_{B_{\bar r}}$, $u_k := (v_k, \Psi_k(x, v_k))$ and consider the current $Z_k := \mathbf{G}_{u_k} \llcorner \mathbf{C}_{\bar r}$. Since $u_k|_{\partial B_{\bar r}} = f_k|_{\partial B_{\bar r}}$, one gets $\partial Z_k = \langle \mathbf{G}_{f_k}, \varphi, \bar r\rangle$ and, hence, $\mathbf{M}(\partial(T_k \llcorner \mathbf{C}_{\bar r} - Z_k)) \le C E_k^{1-2\beta}$. We define

$$S_k = T_k \llcorner (\mathbf{C}_4 \setminus \mathbf{C}_{\bar r}) + Z_k + R_k\,. \tag{4.18}$$

where (cp. Remark 4.3) R_k is an integral current supported in Σ_k such that

$$\partial R_k = \partial(T_k \llcorner \mathbf{C}_{\bar r} - Z_k) \quad \text{and} \quad \mathbf{M}(R_k) \le C E_k^{\frac{(1-2\beta)m}{m-1}}\,.$$

S_k is supported in Σ_k and $\partial S_k = \partial(T_k \llcorner \mathbf{C}_4)$. We now show that, since $\beta < \frac{1}{2m}$, for k large enough, the mass of S_k is smaller than that of T_k. To this aim we write

$$\mathrm{Dir}(u_k, B_{\bar r}) - \mathrm{Dir}(f_k, B_{\bar r}\cap\Gamma_k) = \underbrace{\int_{B_{\bar r}} |Dv_k|^2 - \int_{B_{\bar r}\cap\Gamma_k} |Df_k'|^2}_{I_1}$$
$$+ \underbrace{\int_{B_{\bar r}} |D(\Psi_k(x,v_k))|^2 - \int_{B_{\bar r}} |D(\Psi_k(x,f_k'))|^2}_{I_2} + \underbrace{\int_{B_{\bar r}\setminus\Gamma_k} |D(\Psi_k(x,f_k'))|^2}_{I_3}.$$

The first term is estimated by (4.13) and (3.2): recalling that $v_k = E_k^{1/2} H_k$ and $f_k' = E_k^{1/2} g_k$ (but also that the two functions coincide on $B_{\bar r}\setminus B_r$) we achieve $I_1 \leq \frac{c_2}{2} E_k$ for k large enough. For what concerns the second, we proceed as follows. First we write

$$I_2 = \sum_i \int_{B_{\bar r}} (D(\Psi_k(x,u_k(x))_i - D(\Psi_k(x,f_k'(x))_i) : (D(\Psi_k(x,u_k(x))_i + D(\Psi_k(x,f_k'(x))_i).$$

Next, recalling the chain rule [14, Proposition 1.12], we get

$$\begin{aligned}\left|D(\Psi_k(x,u_k(x))_i + D(\Psi_k(x,f_k'(x))_i\right| &\leq C\|D_x\Psi_k\|_0 + C\|D_u\Psi_k\|_0(\mathrm{Lip}(u_k)+\mathrm{Lip}(f_k'))\\ &\stackrel{(4.9)}{\leq} C E_k^{1/4+\delta}.\end{aligned}$$

Using the letter inequality, the chain rule and (4.9), once again we achieve

$$\begin{aligned} I_2 \leq & C\, E_k^{1/4+\delta} \int_{B_{\bar r}} \Big(\sum_i |D_x\Psi_k(x, v_k^i(x)) - D_x\Psi_k(x,(f_k^i)'(x))| \\ & \qquad + \|D_u\Psi_k\|_0\,(|Dv_k| + |Df_k'|) \Big) \\ \leq & C\, E_k^{1/4+\delta} \|D^2\Psi_k\|_0 \int_{B_{\bar r}} \mathcal{G}(v_k, f_k') + C\, E_k^{1/2+2\delta} \int_{B_{\bar r}} (|Dv_k| + |Df_k'|) \\ \leq & C\, E_k^{1/2+2\delta}\, E_k^{1/2} + C\, E_k^{1+2\delta} \leq C E_k^{1+2\delta}. \end{aligned} \tag{4.19}$$

Finally, $I_3 \leq C\|D\Psi_k\|_\infty^2 |B_3\setminus\Gamma_k| \leq C E_k^{1+\beta}$. Thus, for k large enough we achieve $I_2 + I_3 \leq \frac{c_2}{4} E_k$, thereby reaching $\mathrm{Dir}(u_k, B_{\bar r}) - \mathrm{Dir}(f_k, B_{\bar r}\cap\Gamma_k) \leq \frac{3c_2}{4} E_k$. Hence,

$$\begin{aligned} \mathbf{M}(S_k) - \mathbf{M}(T_k) &\leq \mathbf{M}(Z_k) + C\,\mathbf{M}(R_k) - \mathbf{M}(T_k \llcorner \mathbf{C}_{\bar r}) \\ &\leq Q\,|B_{\bar r}| + \int_{B_{\bar r}} \frac{|Du_k|^2}{2} + C\, E_k^{1+2\beta} + C\, E_k^{\frac{(1-2\beta)m}{m-1}} - Q|B_{\bar r}| - \mathbf{e}_{T_k}(B_{\bar r}) \\ &\leq \int_{B_{\bar r}\cap\Gamma_k} \frac{|Df_k|^2}{2} + \frac{3}{4} c_2\, E_k + C\, E_k^{1+2\beta} + C\, E_k^{\frac{(1-2\beta)m}{m-1}} - \mathbf{e}_{T_k}(B_{\bar r}) \\ &\stackrel{(4.11)}{\leq} -\frac{c_2\, E_k}{4} + C\, E_k^{1+\beta} + C\, E_k^{\frac{(1-2\beta)m}{m-1}} < 0, \end{aligned} \tag{4.20}$$

as soon as E_k is small enough. This gives the desired contradiction and proves (4.1).

For what concerns (4.2) and (4.3), we argue similarly. Without loss of generality we assume $x = 0$ and $s = 1$. Hence, we let $(T_k)_k$, $(\Sigma_k)_k$ and $(\Psi_k)_k$ be sequences with vanishing $E_k := \mathbf{E}(T_k, \mathbf{C}_4)$ and satisfying (4.9), but contradicting (4.2) or (4.3). So, being f_k the E_k^β-Lipschitz approximations, we know that, for any sequence of Dir-minimizing functions $\bar{u}_k$ which we might choose, when we set $w_k = (\bar{u}_k, \Psi_k(x, \bar{u}_k))$ we will have

$$\liminf_k \underbrace{E_k^{-1} \int_{B_2} \left(\mathcal{G}(f_k, w_k)^2 + (|Df_k| - |Dw_k|)^2 + |D(\boldsymbol{\eta} \circ f_k - \boldsymbol{\eta} \circ w_k)|^2\right)}_{=:I(k)} > 0\,. \tag{4.21}$$

As in the previous argument we introduce the maps f_k' satisfying (4.12), the normalized functions $g_k = E_k^{-1/2} f_k'$ and, after extraction of a subsequence, the translating sheets h_k satisfying (3.2) - (3.3) and $\|\mathcal{G}(g_k, h_k)\|_{L^2(B_3)} \to 0$. We next claim that

(i) $\lim_k \int_{B_2} |Dg_k|^2 = \int_{B_2} |Dh_{k_0}|^2$, for any k_0 (recall that $\int_{B_2} |Dh_k|^2$ is constant);

(ii) h_k is Dir-minimizing in B_2.

If (i) is false, then there is a positive constant c_2 such that, for any $r \in [5/2, 3]$,

$$\int_{B_r} \frac{|Dh_k|^2}{2} \leq \int_{B_r} \frac{|Dg_k|^2}{2} - c_2 \leq \frac{\mathbf{e}_{T_k}(B_r)}{E_k} - \frac{c_2}{2}, \tag{4.22}$$

provided k large enough (where the last inequality is again an effect of the Taylor expansion of Remark 4.4). We next define the competitor currents S_k as in the argument leading to (4.20): this latter inequality is reached thanks to (4.22), which substitutes (4.11) and (4.16). On the other hand (4.20) contradicts the minimizing property of T_k. If (ii) is false, then h_k is not Dir-minimizing in B_2. This implies that one of the ζ^j in the translating sheets h_k is not Dir-minimizing in B_2. Indeed, in the opposite case, by [14, Theorem 3.9], $\|\mathcal{G}(\zeta^j, Q\,[\![0]\!])\|_{C^0(B_2)} < \infty$ and, since $h_k = \sum_i [\![\tau_{y_k^i} \circ \zeta^i]\!]$ and $|y_k^i - y_k^j| \to \infty$ for $i \neq j$, by the maximum principle of [14, Proposition 3.5], h_k would be Dir-minimizing. Thus, for some ζ^j we can find a competitor $\hat{\zeta}^j$ with less energy in the ball B_2. So the functions $F_k = \sum_j [\![\tau_{y_k^j} \circ \hat{\zeta}^j]\!]$ satisfy, for any $r \in [5/2, 3]$,

$$\int_{B_r} \frac{|DF_k|^2}{2} \leq \int_{B_r} \frac{|Dh_k|^2}{2} - c_2 = \lim_k \int_{B_r} \frac{|Dg_k|^2}{2} - c_2 \leq \frac{\mathbf{e}_T(B_r)}{E_k} - \frac{c_2}{2} \tag{4.23}$$

provided k is large enough (here $c_2 > 0$ is some constant independent of r and k). On the other hand $F_k = h_k$ on $B_3 \setminus B_{5/2}$ and therefore $\|\mathcal{G}(F_k, g_k)\|_{L^2(B_3 \setminus B_{5/2})} \to 0$. We then construct the competitor current S_k of (4.18): this time we use, however, the map F_k in place of h_k to construct H_k via Proposition 3.4 and we reach the contradiction (4.20) using (4.23) in place of (4.11) and (4.16).

We next set $\bar{u}_k := E_k^{1/2} h_k$ and we aim at showing that, for $w_k = (\bar{u}_k, \Psi_k(x, \bar{u}_k))$, $I(k) \to 0$, a contradiction to (4.21). Observe first that, by $\|\mathcal{G}(g_k, h_k)\|_{L^2} \to 0$, we have $D(\boldsymbol{\xi} \circ g_k) - D(\boldsymbol{\xi} \circ h_k) \rightharpoonup 0$ in L^2 (recall the definition of $\boldsymbol{\xi}$ in Section 1.5). On the other hand, recall that $D(\boldsymbol{\xi} \circ h_k)$ is actually a single function, independent of k, because h_k is a sequence of translating sheets. So, (i) and the identities $|D(\boldsymbol{\xi} \circ g_k)| = |Dg_k|$, $|D(\boldsymbol{\xi} \circ h_k)| = |Dh_k|$ imply that $D(\boldsymbol{\xi}\circ g_k)-D(\boldsymbol{\xi}\circ h_k)$ converge strongly to 0 in L^2. If we next set $\hat{h}_k=\sum_i [\![h_k^i-\boldsymbol{\eta}\circ h_k]\!]$ and $\hat{g}_k=\sum_i [\![g_k^i-\boldsymbol{\eta}\circ g_k]\!]$, we obviously have $\|\mathcal{G}(\hat{h}_k,\hat{g}_k)\|_{L^2}+\|\boldsymbol{\eta}\circ h_k-\boldsymbol{\eta}\circ g_k\|_{L^2}\to 0$. Recall however that the Dirichlet energy enjoys the splitting

$$\mathrm{Dir}(g_k)=Q\int |D(\boldsymbol{\eta}\circ g_k)|^2+\mathrm{Dir}(\hat{g}_k)\qquad \mathrm{Dir}(h_k)=Q\int |D(\boldsymbol{\eta}\circ h_k)|^2+\mathrm{Dir}(\hat{h}_k)\,.$$

So (i) implies that the Dirichlet energies of $\boldsymbol{\eta}\circ g_k$ and $\hat{g}_k$ converge, respectively, to those of $\boldsymbol{\eta}\circ h_k$ and $\hat{h}_k$ (which, we recall again, are independent of k because the h_k's are translating sheets). We thus infer that $D(\boldsymbol{\eta}\circ h_k)-D(\boldsymbol{\eta}\circ g_k)$ converges to 0 strongly in L^2.

Coming back to w_k we observe that

$$E_k^{-1}\int_{B_2}\mathcal{G}(w_k,f_k)^2\leq (2+\mathrm{Lip}(D\Psi)^2)E_k^{-1}\int_{B_2}\mathcal{G}(\bar{u}_k,f_k')^2=C\int_{B_2}\mathcal{G}(h_k,g_k)^2\to 0\,. \tag{4.24}$$

So,

$$\begin{aligned}\limsup_k I(k)\leq& 2\limsup_k\int_{B_2}\big(|Dg_k|-|Dh_k|)^2+|D(\boldsymbol{\eta}\circ g_k-\boldsymbol{\eta}\circ h_k)|^2\big)\\ &+C(Q)\limsup_k E_k^{-1}\int_{B_2}\mathcal{G}(D(\Psi(x,f_k')),D(\Psi(x,\bar{u}_k)))^2\\ \leq& C\limsup_k E_k^{-1}\int_{B_2}\mathcal{G}(D(\Psi(x,f_k')),D(\Psi(x,\bar{u}_k)))^2=\limsup_k E_k^{-1}J(k)\,. \end{aligned}\tag{4.25}$$

Recalling the chain rule of [14, Proposition 1.12], we have

$$\begin{aligned}D(\Psi(x,f_k'))(x)&=\sum_i [\![D_x\Psi(x,(f_k^i)'(x))+D_v\Psi(x,(f_k^i)'(x))\cdot D(f_k^i)'(x)]\!]\\ D(\Psi(x,\bar{u}_k))(x)&=\sum_i [\![D_x\Psi(x,\bar{u}_k^i(x))+D_v\Psi(x,\bar{u}_k^i(x))\cdot D\bar{u}_k^i(x)]\!]\,.\end{aligned}$$

So we can estimate

$$J(k)\leq C\mathrm{Lip}(D_x\Psi)^2\int_{B_2}\mathcal{G}(f_k',\bar{u}_k)^2+C\|D\Psi\|_0^2\int_{B_2}(|Df_k'|^2+|D\bar{u}_k|^2)\stackrel{(4.9)}{\leq} CE_k^{3/2+2\delta}\,.$$

We therefore conclude that $E_k^{-1}J(k)\to 0$ and thus $I(k)\to 0$, which contradicts (4.21). □

5. Gradient L^p estimate

In this section we prove Theorem 1.3. The result is a consequence of an higher integrability estimate for the gradient of Dir-minimizing functions, the $o(E)$-improved estimate for the excess measure given in Proposition 5.4 and a very careful "covering and stopping radius" argument (cf. [31] for an exposition in a more elementary context).

5.1. **Higher integrability of the gradient of Dir-minimizers.** Most of the energy of a Dir-minimizer lies where the gradient is relatively small. We prove indeed the following a priori estimate (cf. [30] for a different proof and some improvements).

Theorem 5.1 (Higher integrability of Dir-minimizers)**.** *There exists $p_{10} > 2$ such that, for every $\Omega' \subset\subset \Omega \subset \mathbb{R}^m$ open domains, there is a constant $C > 0$ such that*

$$\|Du\|_{L^{p_{10}}(\Omega')} \leq C\, \|Du\|_{L^2(\Omega)} \quad \textit{for every } \mathrm{Dir}\textit{-minimizing } u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^n)). \tag{5.1}$$

Proof. The statement is a corollary of Proposition 5.2 below and a Gehring type lemma, cf. [21, Proposition 5.1]. □

Proposition 5.2. *Let $\frac{2\,(m-1)}{m} < p_{11} < 2$. Then, there exists $C = C(m,n,Q,p_{11})$ such that, for every $u : \Omega \to \mathcal{A}_Q$ Dir-minimizing, the following holds*

$$\left(\fint_{B_r(x)} |Du|^2\right)^{1/2} \leq C \left(\fint_{B_{2r}(x)} |Du|^{p_{11}}\right)^{1/p_{11}} \quad \forall\, x \in \Omega,\ \forall\, r < \min\left\{1, \mathrm{dist}(x, \partial\Omega)/2\right\}.$$

Proof. Since the estimate is invariant under translations and rescalings, it is enough to prove it for $x = 0$ and $r = 1$. We assume, therefore $\Omega = B_2$. Let $u : \Omega \to \mathcal{A}_Q(\mathbb{R}^n)$ be Dir-minimizing and let $F = \boldsymbol{\xi} \circ u : \Omega \to \mathcal{Q} \subset \mathbb{R}^N$. Denote by $\bar{F} \in \mathbb{R}^N$ the average of F on B_2. By Fubini's theorem and the Poincaré inequality, there exists $s \in [1,2]$ such that

$$\int_{\partial B_s} \left(|F - \bar{F}|^{p_{11}} + |DF|^{p_{11}}\right) \leq C \int_{B_2} \left(|F - \bar{F}|^{p_{11}} + |DF|^{p_{11}}\right) \leq C \|DF\|^{p_{11}}_{L^{p_{11}}(B_2)}.$$

Consider $F|_{\partial B_s}$. Since $\frac{1}{2} > \frac{1}{p_{11}} - \frac{1}{2\,(m-1)}$, we can use the embedding $W^{1,p_{11}}(\partial B_s) \hookrightarrow H^{1/2}(\partial B_s)$ (see, for example, [1]). Hence, we infer that

$$\left\|F - \bar{F}\right\|_{H^{1/2}(\partial B_s)} \;\leq\; C\, \|DF\|_{L^{p_{11}}(B_2)}\,. \tag{5.2}$$

Let $\hat{F}$ be the harmonic extension of $F|_{\partial B_s}$ in B_s. It is well known (one could, for example, use the result in [1] on the half-space together with a partition of unity) that

$$\|D\hat{F}\|_{L^2(B_s)} \leq C(m) \min_{p \in \mathbb{R}^N} \|\hat{F} - p\|_{H^{1/2}(\partial B_s)} \stackrel{(5.2)}{\leq} C\, \|DF\|_{L^{p_{11}}(B_2)}\,. \tag{5.3}$$

Consider the map $\boldsymbol{\rho}$ of [14, Theorem 2.1]. Since $\boldsymbol{\rho} \circ \hat{F}|_{\partial B_s} = u|_{\partial B_s}$ and $\boldsymbol{\rho} \circ \hat{F}$ takes values in $\mathcal{Q}$, by the minimizing property of u and the Lipschitz continuity of $\boldsymbol{\xi}$, $\boldsymbol{\xi}^{-1}$ and $\boldsymbol{\rho}$, we conclude:

$$\left(\int_{B_1} |Du|^2\right)^{1/2} \leq C \left(\int_{B_s} |D\hat{F}|^2\right)^{1/2} \leq C \left(\int_{B_2} |DF|^{p_{11}}\right)^{1/p_{11}} = C \left(\int_{B_2} |Du|^{p_{11}}\right)^{1/p_{11}}. \quad \square$$

Remark 5.3. Proposition 5.2 can be proved in several different ways, which are based on more common test function arguments: cf. the intrinsic proof (i.e. which does not use the biLipschitz embedding $\boldsymbol{\xi}$) in [30] or the usual Caccioppoli's inequality for quasi minima [22, Theorem 6.5].

5.2. **Improved excess estimate.** The higher integrability of the Dir-minimizing functions and the harmonic approximation lead to the following estimate, which we call "weak" since we will improve it in Theorem 6.1.

Proposition 5.4 (Weak excess estimate)**.** *For every $\eta_{10} > 0$, there exists $\varepsilon_{13} > 0$ with the following property. Let T be area minimizing and assume it satisfies Assumption 1.1 in $\mathbf{C}_{4s}(x)$. If $E = \mathbf{E}(T, \mathbf{C}_{4s}(x)) \leq \varepsilon_{13}$, then*

$$\mathbf{e}_T(A) \leq \eta_{10}\, E\, s^m + C\, \mathbf{A}^2\, s^{m+2}, \tag{5.4}$$

for every $A \subset B_s(x)$ Borel with $|A| \leq \varepsilon_{13}|B_s(x)|$.

Proof. Without loss of generality, we can assume $s = 1$ and $x = 0$. We distinguish the two regimes: $E \leq \mathbf{A}^2$ and $\mathbf{A}^2 \leq E$. In the former, clearly $\mathbf{e}_T(A) \leq C\, E \leq C\, \mathbf{A}^2$. In the latter, we let f be the $E^{\frac{1}{4m}}$-Lipschitz approximation of T in $\mathbf{C}_3$ and, arguing as for the proof of Theorem 4.2, we find a radius $r \in (1, 2)$ and a current R such that

$$\partial R = \langle T - \mathbf{G}_f, \varphi, r\rangle \quad \text{and} \quad \mathbf{M}(R) \leq C E^{(1-\frac{1}{2m})\frac{m}{m-1}}.$$

Therefore, by the Taylor expansion in Remark 4.4, we have:

$$\|T\|(\mathbf{C}_r) \leq \mathbf{M}(\mathbf{G}_f \llcorner \mathbf{C}_r + R) \leq \|\mathbf{G}_f\|(\mathbf{C}_r) + C\, E^{\frac{2m-1}{2m-2}} \leq Q\,|B_r| + \int_{B_r} \frac{|Df|^2}{2} + C\, E^{1+\gamma}, \tag{5.5}$$

where $\gamma = \frac{1}{2m}$. On the other hand, using again the Taylor expansion for the part of the current which coincides with the graph of f, we deduce as well that

$$\|T\|((B_r \cap K) \times \mathbb{R}^n) \geq Q\,|B_r \cap K| + \int_{B_r \cap K} \frac{|Df|^2}{2} - C\, E^{1+\gamma}. \tag{5.6}$$

Subtracting (5.6) from (5.5), we deduce

$$\mathbf{e}_T(B_r \setminus K) \leq \int_{B_r \setminus K} \frac{|Df|^2}{2} + CE^{1+\gamma}. \tag{5.7}$$

If ε_{13} is chosen small enough, we infer from (5.7) and (4.1) in Theorem 4.2 that

$$\mathbf{e}_T(B_r \setminus K) \leq \eta\, E + CE^{1+\gamma}, \tag{5.8}$$

for a suitable $\eta > 0$ to be chosen. Let now $A \subset B_1$ be such that $|A| \leq \varepsilon_{13}\, \omega_m$. Combining (5.8) with the Taylor expansion, we have

$$\mathbf{e}_T(A) \leq \mathbf{e}_T(A \setminus K) + \int_A \frac{|Df|^2}{2} + C\, E^{1+\gamma} \leq \int_A \frac{|Df|^2}{2} + \eta\, E + CE^{1+\gamma}. \tag{5.9}$$

If ε_{13} is small enough, we can again apply Theorem 4.2. Using the coordinates of Remark 1.5, there is a Dir-minimizing u such that $|Df|$ is close in L^2 (with an error ηE) to $|Dw|$ with $w = (u, \Psi(x, u))$ and by Remark 4.5 $\mathrm{Dir}(u) \leq CE$. On the other hand $|Dw(x)| \leq$

$(1+\|D\Psi\|_0)|Du|+\|D\Psi\|_0$. Since $\|D\Psi\|_0 \leq C E^{1/2}$, by Theorem 5.1 $\||Dw|\|_{L^{p_{10}}(B_1)} \leq C E^{1/2}$. Therefore,

$$\mathbf{e}_T(A) \stackrel{(4.2)}{\leq} \int_A |Dw|^2 + 3\,\eta\, E + C E^{1+\gamma} \leq C\left(|A|^{1-2/p_{10}} + \eta\right) E + C E^{1+\gamma}. \tag{5.10}$$

Hence, if ε_{13} and η are suitably chosen, (5.4) follows from (5.10). □

5.3. **Proof of Theorem 1.3.** We assume without loss of generality that $E > 0$ and divide the proof into two steps.

Step 1. There exist constants $\gamma \geq 2^m$ and $\varrho > 0$ such that, for every $c \in [1, (\gamma\, E)^{-1}]$ and $s \in [2, 4]$ with $\bar{s} = s + 4\, c^{-1/m} \leq 4$, we have

$$\int_{\{\gamma\, c\, E \leq \mathbf{d} \leq 1\} \cap B_s} \mathbf{d} \leq \gamma^{-\varrho} \int_{\left\{\frac{c\,E}{\gamma} \leq \mathbf{d} \leq 1\right\} \cap B_{\bar{s}}} \mathbf{d} + C\, c^{-2/m}\, \mathbf{A}^2. \tag{5.11}$$

In order to prove it, let N_B be the constant in Besicovich's covering theorem [19, Section 1.5.2] and choose $N \in \mathbb{N}$ so large that $N_B < 2^{N-1}$. Let ε_{13} be as in Proposition 5.4 when we choose $\eta_{10} = 2^{-2m-N}$, and set

$$\gamma = \max\{2^m, \varepsilon_{13}^{-1}\} \quad \text{and} \quad \varrho = \min\left\{-\log_\gamma(N_B/2^{N-1}), \frac{1}{2m}\right\}.$$

Let c and s be any real numbers as above. For almost every $x \in \{\gamma\, c\, E \leq \mathbf{d} \leq 1\} \cap B_s$, there exists r_x such that

$$E(T, \mathbf{C}_{4r_x}(x)) \leq c\, E \qquad \text{and} \qquad E(T, \mathbf{C}_t(x)) \geq c\, E \quad \forall t \in]0, 4\, r_x[. \tag{5.12}$$

Indeed, since $\mathbf{d}(x) = \lim_{r\to 0} E(T, \mathbf{C}_r(x)) \geq \gamma\, c\, E \geq 2^m c\, E$ and

$$E(T, \mathbf{C}_t(x)) = \frac{\mathbf{e}_T(B_t(x))}{\omega_m\, t^m} \leq \frac{4^m\, E}{t^m} \leq c\, E \quad \text{for} \quad t \geq \frac{4}{\sqrt[m]{c}},$$

we just choose $4r_x = \min\{t \leq 4/\sqrt[m]{c} : E(T, \mathbf{C}_t(x)) \leq cE\}$. Note also that $r_x \leq 1/\sqrt[m]{c}$. Consider the current T in $\mathbf{C}_{4r_x}(x)$. Setting $A = \{\gamma\, c\, E \leq \mathbf{d}\} \cap B_{4r_x}(x)$, we have that

$$\mathbf{E}(T, \mathbf{C}_{4r_x}(x)) \leq c\, E \leq \frac{E}{\gamma\, E} \leq \varepsilon_{13} \quad \text{and} \quad |A| \leq \frac{c\, E\, |B_{4r_x}(x)|}{\gamma\, c\, E} \leq \varepsilon_{13} |B_{4r_x}(x)|.$$

Hence, we can apply Proposition 5.4 to $T \llcorner \mathbf{C}_{4r_x}(x)$ to get

$$\begin{aligned}\int_{B_{r_x}(x) \cap \{\gamma\, c\, E \leq \mathbf{d} \leq 1\}} \mathbf{d} &\leq \int_A \mathbf{d} \leq \mathbf{e}_T(A) \leq 2^{-2\,m-N}\, \mathbf{e}_T(B_{4r_x}(x)) + C\, r_x^{m+2} \mathbf{A}^2 \\ &\leq 2^{-2\,m-N}\, (4\, r_x)^m\, \omega_m\, \mathbf{E}(T, \mathbf{C}_{4r_x}(x)) + C\, r_x^{m+2} \mathbf{A}^2 \stackrel{(5.12)}{\leq} 2^{-N}\, \mathbf{e}_T(B_{r_x}(x)) + C\, r_x^{m+2} \mathbf{A}^2. \end{aligned} \tag{5.13}$$

Thus,

$$\begin{aligned}
\mathbf{e}_T(B_{r_x}(x)) &= \int_{B_{r_x}(x)\cap\{\mathbf{d}>1\}} \mathbf{d} + \int_{B_{r_x}(x)\cap\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}} \mathbf{d} + \int_{B_{r_x}(x)\cap\left\{\mathbf{d}<\frac{c\,E}{\gamma}\right\}} \mathbf{d}\\
&\leq \int_A \mathbf{d} + \int_{B_{r_x}(x)\cap\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}} \mathbf{d} + \frac{c\,E}{\gamma}\,\omega_m\, r_x^m\\
&\stackrel{(5.12),(5.13)}{\leq} \left(2^{-N}+\gamma^{-1}\right)\,\mathbf{e}_T(B_{r_x}(x)) + C\,r_x^{m+2}\mathbf{A}^2 + \int_{B_{r_x}(x)\cap\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}} \mathbf{d}.
\end{aligned} \tag{5.14}$$

Therefore, recalling that $\gamma \geq 2^m \geq 4$, from (5.13) and (5.14) we infer:

$$\begin{aligned}
\int_{B_{r_x}(x)\cap\{\gamma\,c\,E\leq\mathbf{d}\leq 1\}} \mathbf{d} &\leq \frac{2^{-N}}{1-2^{-N}-\gamma^{-1}} \int_{B_{r_x}(x)\cap\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}} \mathbf{d} + C\,r_x^{m+2}\mathbf{A}^2\\
&\leq 2^{-N+1}\int_{B_{r_x}(x)\cap\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}} \mathbf{d} + C\,r_x^{m+2}\mathbf{A}^2.
\end{aligned}$$

By Besicovich's covering theorem, we choose N_B families of disjoint balls $\overline{B}_{r_x}(x)$ whose union covers $\{\gamma\,c\,E \leq \mathbf{d} \leq 1\}\cap B_s$ and, since as already noticed $r_x \leq 1/\sqrt[m]{c}$ for every x, we conclude:

$$\int_{\{\gamma\,c\,E\leq\mathbf{d}\leq 1\}\cap B_s} \mathbf{d} \leq N_B\,2^{-N+1}\int_{\left\{\frac{c\,E}{\gamma}\leq\mathbf{d}\leq 1\right\}\cap B_{s+\frac{2}{\sqrt[m]{c}}}} \mathbf{d} + C\,c^{-\frac{2}{m}}\,\mathbf{A}^2,$$

which, for the above defined ϱ, implies (5.11).

Step 2. We iterate (5.11) in order to conclude (1.3). Denote by L the largest integer smaller than $2^{-1}\left((\log_\gamma E^{-1})-1\right)$, $s_L = 2$ and recursively $s_k = s_{k+1} + 2\,\gamma^{-\frac{2k}{m}}$ for $k \in \{L-1,\ldots,1\}$. Notice that, since $\gamma \geq 2^m$, $s_k < 4$ for every k. Thus, we can apply (5.11) with $c = \gamma^{2k}$, $s = s_k$ and $\bar{s} = s_{k-1}$ to conclude

$$\int_{\{\gamma^{2k+1}\,E\leq\mathbf{d}\leq 1\}\cap B_{s_k}} \mathbf{d} \leq \gamma^{-\varrho}\int_{\left\{\gamma^{2k-1}\,E\leq\mathbf{d}\leq 1\right\}\cap B_{s_{k-1}}} \mathbf{d} + C\,\gamma^{-\frac{4\,k}{m}}\mathbf{A}^2 \quad \forall\, k\in\{2,\ldots,L\}\,.$$

In particular, iterating this estimate we get

$$\int_{\{\gamma^{2\,k+1}\,E\leq\mathbf{d}\leq 1\}\cap B_2} \mathbf{d} \leq \gamma^{-(k-1)\,\varrho}\int_{\{\gamma\,E\leq\mathbf{d}\leq 1\}\cap B_{s_1}} \mathbf{d} + C\,\mathbf{A}^2\sum_{\ell=0}^{k-2}\gamma^{-\left(\frac{4\,(k-\ell)}{m}+\ell\,\varrho\right)}. \tag{5.15}$$

Set $A_0 = \{\mathbf{d} < \gamma E\}$, $A_k = \{\gamma^{2k-1} E \le \mathbf{d} < \gamma^{2k+1} E\}$ for $k = 1, \ldots, L$, and $A_{L+1} = \{\gamma^{2L+1} E \le \mathbf{d} \le 1\}$. Since $\cup A_k = \{\mathbf{d} \le 1\}$, for $p_1 < 1 + \frac{\varrho}{2} \le 1 + \frac{1}{m}$, we conclude:

$$
\begin{aligned}
\int_{B_2 \cap \{\mathbf{d} \le 1\}} \mathbf{d}^{p_1} &= \sum_{k=0}^{L+1} \int_{A_k \cap B_2} \mathbf{d}^{p_1} \le \sum_k \gamma^{(2k+1)(p_1-1)} E^{p_1-1} \int_{A_k \cap B_2} \mathbf{d} \\
&\overset{(5.15)}{\le} C \sum_k \gamma^{k(2(p_1-1)-\varrho)} E^{p_1} + C \sum_k \sum_{\ell=0}^{k-2} \gamma^{k(2(p_1-1)-\frac{4}{m})+\ell(\frac{4}{m}-\varrho)} E^{p_1-1} \mathbf{A}^2 \\
&\le C E^{p_1} + C \sum_k \gamma^{k(2(p_1-1)-\varrho)} E^{p_1-1} \mathbf{A}^2 .
\end{aligned}
$$

6. Almgren's approximation theorem

In this section we show how Theorem 1.3 gives a simple proof of the approximation result in Theorem 1.4. The key point is the following theorem.

Theorem 6.1 (Almgren's strong excess estimate)**.** *There are constants $\varepsilon_{11}, \gamma_{11}, C > 0$ (depending on $m, n, \bar{n}, Q$) with the following property. Assume T satisfies Assumption 1.1 in $\mathbf{C}_4$ and is area minimizing. If $E = \mathbf{E}(T, \mathbf{C}_4) < \varepsilon_{11}$, then*

$$
\mathbf{e}_T(A) \le C \left(E^{\gamma_{11}} + |A|^{\gamma_{11}}\right) \left(E + \mathbf{A}^2\right) \quad \textit{for every Borel } A \subset B_{\frac{9}{8}} . \tag{6.1}
$$

This estimate complements (1.3) enabling to control the excess in the region where $\mathbf{d} > 1$. We call it strong Almgren's estimate because a similar formula can be found in the big regularity paper (cf. [3, Sections 3.24-3.26 & 3.30(8)]) and is a strengthened version of Proposition 5.4. To achieve (6.1) we construct a suitable competitor to estimate the size of the set K where the graph of the E^β-Lipschitz approximation f differs from T. Following Almgren, we embed $\mathcal{A}_Q$ in a large Euclidean space, via a biLipschitz embedding $\boldsymbol{\xi}$. We then regularize $\boldsymbol{\xi} \circ f$ by convolution and project it back onto $\mathcal{Q} = \boldsymbol{\xi}(\mathcal{A}_Q)$. To avoid loss of energy we need a rather special "almost projection" $\boldsymbol{\rho}^\star_\delta$.

Proposition 6.2. *For every $\bar{n}, Q \in \mathbb{N} \setminus \{0\}$ there are geometric constants $\delta_0, C > 0$ with the following property. For every $\delta \in]0, \delta_0[$ there is $\boldsymbol{\rho}^\star_\delta : \mathbb{R}^{N(Q,\bar{n})} \to \mathcal{Q} = \boldsymbol{\xi}(\mathcal{A}_Q(\mathbb{R}^{\bar{n}}))$ such that $|\boldsymbol{\rho}^\star_\delta(P) - P| \le C\, \delta^{8^{-\bar{n}Q}}$ for all $P \in \mathcal{Q}$ and, for every $u \in W^{1,2}(\Omega, \mathbb{R}^N)$, the following holds:*

$$
\int |D(\boldsymbol{\rho}^\star_\delta \circ u)|^2 \le \left(1 + C\, \delta^{8^{-\bar{n}Q-1}}\right) \int_{\{\mathrm{dist}(u,\mathcal{Q}) \le \delta^{\bar{n}Q+1}\}} |Du|^2 + C \int_{\{\mathrm{dist}(u,\mathcal{Q}) > \delta^{\bar{n}Q+1}\}} |Du|^2 . \tag{6.2}
$$

The proof of Proposition 6.2 is postponed to the next section. Here we show Theorem 6.1 and hence conclude the Theorems 1.4 and 1.6. Theorem 1.3 enters crucially in the argument when estimating the second summand of (6.2) for the regularization of $\boldsymbol{\xi} \circ f$.

6.1. Regularization by convolution.

Here we construct the competitor.

Proposition 6.3. *Let $\beta_1 \in \left(0, \frac{1}{2m}\right)$ and T be an area minimizing current satisfying Assumption 1.1 in $\mathbf{C}_4$. Let f be its E^{β_1}-Lipschitz approximation. Then, there exist constants*

$\bar\varepsilon_{12}, \gamma_{12}, C > 0$ and a subset of radii $B \subset [9/8, 2]$ with $|B| > 1/2$ with the following properties. If $\mathbf{E}(T, \mathbf{C}_4) \leq \bar\varepsilon_{12}$, for every $\sigma \in B$, there exists a Q-valued function $g \in \mathrm{Lip}(B_\sigma, \mathcal{A}_Q)$ such that

$$g|_{\partial B_\sigma} = f|_{\partial B_\sigma}, \quad \mathrm{Lip}(g) \leq C\, E^{\beta_1}, \quad \mathrm{spt}(g(x)) \subset \Sigma \quad \forall\, x \in B_\sigma,$$

and

$$\int_{B_\sigma} |Dg|^2 \leq \int_{B_\sigma \cap K} |Df|^2 + C\, E^{\gamma_{12}} \left(E + \mathbf{A}^2\right). \tag{6.3}$$

Proof. By Remark 1.5 we assume that $\Psi(0) = 0$, $\|D\Psi\|_0 \leq C(E^{1/2} + \mathbf{A})$ and $\|D^2\Psi\|_0 \leq C\mathbf{A}$. Since $|Df|^2 \leq C\, \mathbf{d}_T \leq C E^{2\beta_1} \leq 1$ on K, by Theorem 1.3 there is $q_1 = 2\,p_1 > 2$ such that

$$\||Df|\|^2_{L^{q_1}(K \cap B_2)} \leq C\, E^{1-1/p_1} (E + \mathbf{A}^2)^{1/p_1} \leq C(E + \mathbf{A}^2). \tag{6.4}$$

Given two (vector-valued) functions h_1 and h_2 and two radii $0 < \bar r < r$, we denote by $\mathrm{lin}(h_1, h_2)$ the linear interpolation in $B_r \setminus \bar B_{\bar r}$ between $h_1|_{\partial B_r}$ and $h_2|_{\partial B_{\bar r}}$. More precisely, if $(\theta, t) \in \mathbb{S}^{m-1} \times [0, \infty)$ are spherical coordinates, then

$$\mathrm{lin}(h_1, h_2)(\theta, t) = \frac{r - t}{r - \bar r}\, h_2(\theta, t) + \frac{t - s}{r - \bar r}\, h_1(\theta, t).$$

Next, let $\delta > 0$ and $\varepsilon > 0$ be two parameters and let $1 < r_1 < r_2 < r_3 < 2$ be three radii, all to be chosen later. To keep the notation simple, we will write $\boldsymbol{\rho}^\star$ in place of $\boldsymbol{\rho}^\star_\delta$. Let $\varphi \in C^\infty_c(B_1)$ be a standard (nonnegative!) mollifier. We also use the notation $f(x) = (f_1(x), f_2(x)) \in \mathcal{A}_Q(\mathbb{R}^{\bar n} \times \mathbb{R}^l)$ meaning that $f(x) = \sum_i [\![(f^i_1(x), f^i_2(x))]\!]$ with $(f^i_1(x), f^i_2(x)) \in \mathbb{R}^{\bar n} \times \mathbb{R}^l$ and the maps f_1 and f_2 are then given by $f_j(x) = \sum_i [\![f^i_j(x)]\!]$. This does not create confusion in "ordering the sheets": since the points $f^i(x)$ belong to Σ we have indeed the relation $f^j_2(x) = \Psi(x, f^j_1(x))$. We moreover set $f' := \boldsymbol{\xi} \circ f_1$. Recall the map $\boldsymbol{\rho}$ of [14, Theorem 2.1] and define:

$$g' := \begin{cases} \sqrt{E}\, \boldsymbol{\rho} \circ \mathrm{lin}\left(\frac{f'}{\sqrt{E}}, \boldsymbol{\rho}^\star\left(\frac{f'}{\sqrt{E}}\right)\right) & \text{in } B_{r_3} \setminus B_{r_2},\\ \sqrt{E}\, \boldsymbol{\rho} \circ \mathrm{lin}\left(\boldsymbol{\rho}^\star\left(\frac{f'}{\sqrt{E}}\right), \boldsymbol{\rho}^\star\left(\frac{f'}{\sqrt{E}} * \varphi_\varepsilon\right)\right) & \text{in } B_{r_2} \setminus B_{r_1},\\ \sqrt{E}\, \boldsymbol{\rho}^\star\left(\frac{f'}{\sqrt{E}} * \varphi_\varepsilon\right) & \text{in } B_{r_1}. \end{cases} \tag{6.5}$$

Finally set $g_1 := \boldsymbol{\xi}^{-1} \circ g'$ and $g := \sum_i [\![(g^i_1, \Psi(x, g^i_1))]\!]$. We claim that, for $\sigma := r_3$ in a suitable set $B \subset [9/8, 2]$ with $|B| > 1/2$, we can choose $r_2 = r_3 - s$ and $r_1 = r_2 - s$ so that g satisfies the conclusion of the proposition. Some computations will be simplified taking into account that our choice of the parameters will imply the following inequalities:

$$\delta^{2\cdot 8^{-\bar n Q}} \leq s, \quad \varepsilon \leq s \quad \text{and} \quad E^{1-2\beta_1} \leq \varepsilon^m. \tag{6.6}$$

We start noticing that clearly $g|_{\partial B_{r_3}} = f|_{\partial B_{r_3}}$. As for the Lipschitz constant, it suffices to estimate the Lipschitz constant of g'. This can be easily done observing that:

$$\begin{cases} \mathrm{Lip}(g') \leq C\, \mathrm{Lip}(f' * \varphi_\varepsilon) \leq C\, \mathrm{Lip}(f') \leq C\, E^{\beta_1} & \text{in } B_{r_1},\\ \mathrm{Lip}(g') \leq C\, \mathrm{Lip}(f') + C\, \frac{\|f' - f' * \varphi_\varepsilon\|_{L^\infty}}{s} \leq C(1 + \frac{\varepsilon}{s})\, \mathrm{Lip}(f') \leq C\, E^{\beta_1} & \text{in } B_{r_2} \setminus B_{r_1},\\ \mathrm{Lip}(g') \leq C\, \mathrm{Lip}(f') + C\, E^{1/2}\, \frac{\delta^{8^{-\bar n Q}}}{s} \leq C\, E^{\beta_1} + C\, E^{1/2} \leq C\, E^{\beta_1} & \text{in } B_{r_3} \setminus B_{r_2}. \end{cases}$$

In the first inequality of the last line we have used that, since $\mathcal{Q}$ is a cone, $E^{-1/2}f'(x) \in \mathcal{Q}$ for every x: therefore $|\boldsymbol{\rho}^\star(f'/E^{1/2}) - f'/E^{1/2}| \leq C\delta^{8^{-\bar{n}Q}}$. We pass now to estimate the Dirichlet energy of g.

Step 1. Energy in $B_{r_3} \setminus B_{r_2}$. By Section 1.5, the energy of the first component g_1 coincides with the (classical!) Dirichlet energy of g'. By Proposition 6.2, $|\boldsymbol{\rho}^\star(P) - P| \leq C\,\delta^{8^{-\bar{n}Q}}$ for all $P \in \mathcal{Q}$. Thus, elementary estimates on the linear interpolation give

$$\int_{B_{r_3}\setminus B_{r_2}} |Dg'|^2 \leq \frac{C\,E}{(r_3-r_2)^2}\int_{B_{r_3}\setminus B_{r_2}} \left|\tfrac{f'}{\sqrt{E}} - \boldsymbol{\rho}^\star\left(\tfrac{f'}{\sqrt{E}}\right)\right|^2 + C\int_{B_{r_3}\setminus B_{r_2}} |Df'|^2$$
$$+ C\int_{B_{r_3}\setminus B_{r_2}} |D(\boldsymbol{\rho}^\star\circ f')|^2 \leq C\int_{B_{r_3}\setminus B_{r_2}} |Df'|^2 + C\,E\,s^{-1}\,\delta^{2\cdot 8^{-\bar{n}Q}}\,. \quad (6.7)$$

As for g_2, we compute $Dg_2^i(x) = D_x\Psi(x, g_1^i(x)) + D_u\Psi(x, g_1^i(x))Dg_1^i(x)$ and so

$$\int_{B_{r_3}\setminus B_{r_2}} |Dg_2|^2 \leq C\,s\,(E + \mathbf{A}^2)\,, \quad (6.8)$$

where we used the estimate $\|Dg_2\|_0 \leq C\,\|D\Psi\|_0 \leq C(E^{1/2} + \mathbf{A})$.

Step 2. Energy in $B_{r_2} \setminus B_{r_1}$. Here, using the same interpolation inequality and a standard estimate on convolutions of $W^{1,2}$ functions, we get

$$\int_{B_{r_2}\setminus B_{r_1}} |Dg'|^2 \leq C\int_{B_{r_2+\varepsilon}\setminus B_{r_1-\varepsilon}} |Df'|^2 + \frac{C}{(r_2-r_1)^2}\int_{B_{r_2}\setminus B_{r_1}} |f' - \varphi_\varepsilon * f'|^2$$
$$\leq C\int_{B_{r_2+\varepsilon}\setminus B_{r_1-\varepsilon}} |Df'|^2 + C\,\varepsilon^2 s^{-2}\int_{B_3} |Df'|^2 \leq C\int_{B_{r_2+\varepsilon}\setminus B_{r_1-\varepsilon}} |Df'|^2 + C\,\varepsilon^2\,E\,s^{-2}\,. \quad (6.9)$$

Similarly, for the second component we have that

$$\int_{B_{r_2+\varepsilon}\setminus B_{r_1-\varepsilon}} |Dg_2|^2 \leq C\,(\mathbf{A}^2 + E)\,s. \quad (6.10)$$

Step 3. Energy in B_{r_1}. Define $Z := \left\{\mathrm{dist}\left(\tfrac{f'}{\sqrt{E}} * \varphi_\varepsilon, \mathcal{Q}\right) > \delta^{\bar{n}Q+1}\right\}$ and use (6.2) to get

$$\int_{B_{r_1}} |Dg'|^2 \leq \left(1 + C\,\delta^{8^{-\bar{n}Q-1}}\right)\int_{B_{r_1}\setminus Z} |D\,(f' * \varphi_\varepsilon)|^2 + C\int_Z |D\,(f' * \varphi_\varepsilon)|^2 =: I_1 + I_2. \quad (6.11)$$

We consider I_1 and I_2 separately. For I_1 we first observe the elementary inequality

$$\|D(f' * \varphi_\varepsilon)\|_{L^2}^2 \leq \||Df'| * \varphi_\varepsilon\|_{L^2}^2 \leq \|(|Df'|\,\mathbf{1}_K) * \varphi_\varepsilon\|_{L^2}^2 + \|(|Df'|\,\mathbf{1}_{K^c}) * \varphi_\varepsilon\|_{L^2}^2$$
$$+ 2\|(|Df'|\,\mathbf{1}_K) * \varphi_\varepsilon\|_{L^2}\|(|Df'|\,\mathbf{1}_{K^c}) * \varphi_\varepsilon\|_{L^2}\,, \quad (6.12)$$

where K^c is the complement of K in B_3. Recalling $r_1 + \varepsilon \leq r_1 + s = r_2$ we estimate the first summand in (6.12) as follows:

$$\|(|Df'|\,\mathbf{1}_K) * \varphi_\varepsilon\|_{L^2(B_{r_1})}^2 \leq \int_{B_{r_1+\varepsilon}} (|Df'|\,\mathbf{1}_K)^2 \leq \int_{B_{r_2}\cap K} |Df'|^2\,. \quad (6.13)$$

To treat the other terms recall that $\mathrm{Lip}(f') \leq C\, E^{\beta_1}$ and $|K^c| \leq C\, E^{1-2\beta_1}$:

$$\|(|Df'|\mathbf{1}_{K^c})*\varphi_\varepsilon\|^2_{L^2(B_{r_1})} \leq CE^{2\beta_1}\|\mathbf{1}_{K^c}*\varphi_\varepsilon\|^2_{L^2} \leq CE^{2\beta_1}\,\|\mathbf{1}_{K^c}\|^2_{L^1}\,\|\varphi_\varepsilon\|^2_{L^2} \leq \frac{CE^{2-2\beta_1}}{\varepsilon^m}. \tag{6.14}$$

Putting (6.13) and (6.14) in (6.12) and recalling $E^{1-2\beta_1} \leq \varepsilon^m$ and $\int |Df'|^2 \leq CE$, we get

$$I_1 \leq \int_{B_{r_2}\cap K} |Df'|^2 + C\,\delta^{8^{-\bar n Q-1}}\,E + C\,\varepsilon^{-m/2}\,E^{3/2-\beta_1}\,. \tag{6.15}$$

For what concerns I_2, first we argue as for I_1, splitting in K and K^c, to deduce that

$$I_2 \leq C\int_Z ((|Df'|\,\mathbf{1}_K)*\varphi_\varepsilon)^2 + C\,\varepsilon^{-m/2}E^{3/2-\beta_1}. \tag{6.16}$$

Then, regarding the first summand in (6.16), we note that

$$|Z|\,\delta^{2\bar n Q+2} \leq \int_{B_{r_1}} \left|\tfrac{f'}{\sqrt{E}}*\varphi_\varepsilon - \tfrac{f'}{\sqrt{E}}\right|^2 \leq C\,\varepsilon^2. \tag{6.17}$$

Since $|Df'| \leq |Df|$ (and recalling that $q_1 = 2p_1 > 2$), we use (6.4) to obtain

$$\begin{aligned}\int_Z ((|Df'|\,\mathbf{1}_K)*\varphi_\varepsilon)^2 &\leq |Z|^{\frac{p_1-1}{p_1}}\|(|Df'|\,\mathbf{1}_K)*\varphi_\varepsilon\|^2_{L^{q_1}} \leq C\,\left(\frac{\varepsilon}{\delta^{\bar n Q+1}}\right)^{\frac{2\,(p_1-1)}{p_1}}\,\||Df'|\|^2_{L^{q_1}(K)}\\ &\leq C\,\left(\frac{\varepsilon}{\delta^{\bar n Q+1}}\right)^{\frac{2\,(p_1-1)}{p_1}}\,(E+\mathbf{A}^2)\,.\end{aligned} \tag{6.18}$$

Gathering all the estimates together, (6.11), (6.15), (6.16) and (6.18) give

$$\int_{B_{r_1}} |Dg'|^2 \leq \int_{B_{r_2}\cap K} |Df'|^2 + C\Big(E\delta^{8^{-\bar n Q-1}} + \frac{E^{3/2-\beta_1}}{\varepsilon^{m/2}} + (E+\mathbf{A}^2)\left(\frac{\varepsilon}{\delta^{\bar n Q+1}}\right)^{\frac{2\,(p_1-1)}{p_1}}\Big). \tag{6.19}$$

On the other hand, for what concerns g_2 we can estimate as follows

$$\begin{aligned}\int_{B_{r_1}} |Dg_2|^2 &= \int_{B_{r_1}} |Df_2|^2 + \sum_i \int_{B_{r_1}} (Dg_2^i - Df_2^i)\cdot(Dg_2^i + Df_2^i)\\ &\leq \int_{B_{r_1}\cap K} |Df_2|^2 + \int_{B_{r_1}\setminus K} |Df_2|^2 + C\left(\mathbf{A} + E^{\frac12}\right)\sum_i \int_{B_{r_1}} |Dg_2^i - Df_2^i|\end{aligned} \tag{6.20}$$

We already observed that $|Df_2| \leq C(\mathbf{A} + E^{1/2})$, leading to the estimate $\int_{K^c} |Df_2|^2 \leq C(\mathbf{A}^2+E)|K^c| \leq C(\mathbf{A}^2+E)E^{1-2\beta_1}$. As for the latter summand we compute

$$\begin{aligned}|Dg_2^i - Df_2^i| \leq& |D_x\Psi(x,g_1^i) - D_x\Psi(x,f_1^i)|\\ &+ |D_u\Psi(x,g_1^i(x))Dg_1^i| + |D_u\Psi(x,f_1^i(x))Df_1^i|\\ \leq& C\mathbf{A}\mathcal{G}(g_1,f_1) + C\left(\mathbf{A}+E^{1/2}\right)E^{\beta_1}\,.\end{aligned}$$

We next estimate $\|\mathcal{G}(g_1,f_1)\|_\infty \leq Cg'' - f'\|_\infty$ and

$$\begin{aligned}\|g'-f'\|_\infty &\leq C\,\sqrt{E}\left(\left\|\boldsymbol{\rho}^*\left(\tfrac{f'}{\sqrt{E}}*\varphi_\varepsilon\right) - \boldsymbol{\rho}^*\left(\tfrac{f'}{\sqrt{E}}\right)\right\|_\infty + \left\|\boldsymbol{\rho}^*\left(\tfrac{f'}{\sqrt{E}}\right) - \tfrac{f'}{\sqrt{E}}\right\|_\infty\right)\\ &\leq C\,\mathrm{Lip}(\boldsymbol{\rho}^*)\,\|f'*\varphi_\varepsilon - f'\|_{L^\infty} + C\,E^{1/2}\delta^{8^{-\bar n Q}} \leq C\,E^{\beta_1}\,\varepsilon + C\,E^{1/2}\delta^{8^{-\bar n Q}} \leq CE^{\beta_1}\,.\end{aligned}$$

We therefore conclude

$$\int_{B_{r_1}} |Dg_2|^2 \le \int_{B_{r_1}\cap K} |Df_2|^2 + C(\mathbf{A}^2+E)E^{\beta_1}\,. \tag{6.21}$$

Final estimate. Since $|Dg|^2 = |Dg'|^2 + |Dg_2|^2$, summing (6.7), (6.8), (6.9), (6.10), (6.19) and (6.21) (and recalling $\varepsilon < s$), we conclude

$$\begin{aligned}\int_{B_{r_3}} |Dg|^2 \le \int_{B_{r_1}\cap K} |Df|^2 + C\int_{B_{r_1+3s}\setminus B_{r_1-s}} |Df'|^2 + C(\mathbf{A}+E^2)(s+E^{\beta_1})\\ + C\,E\left(\delta^{8^{-\bar n Q-1}} + \frac{\varepsilon^2}{s^2} + \frac{\delta^{2\cdot 8^{-\bar n Q}}}{s} + \frac{E^{1/2-\beta_1}}{\varepsilon^{m/2}} + \left(1+\mathbf{A}^2\,E^{-1}\right)\left(\frac{\varepsilon}{\delta^{\bar n Q+1}}\right)^{\frac{2\,(p_1-1)}{p_1}}\right).\end{aligned}$$

We set $\varepsilon = E^a$, $\delta = E^b$ and $s = E^c$, where

$$a = \frac{1-2\,\beta_1}{2\,m}, \quad b = \frac{1-2\,\beta_1}{4\,m\,(\bar n\,Q+1)} \quad \text{and} \quad c = \frac{1-2\,\beta_1}{8^{\bar n Q}\,4m\,(\bar n\,Q+1)}.$$

This choice respects (6.6). Assume E is small enough so that $s \le \frac{1}{16}$. Now, if $C > 0$ is a sufficiently large constant, there is a set $B' \subset [9/8, \frac{29}{16}]$ with $|B'| > 1/2$ such that,

$$\int_{B_{r_1+3s}\setminus B_{r_1-s}} |Df'|^2 \le C\,s\int_{B_2} |Df'|^2 \le C\,E^{1+c} \quad \text{for every } r_1 \in B'.$$

Indeed by integrating in polar coordinates and by Fubini's Theorem we have that

$$\begin{aligned}\int_{\frac{9}{8}}^{\frac{29}{16}} dr \int_{B_{r+3s}\setminus B_{r-s}} |Df'|^2 &= \int_{\frac{9}{8}}^{\frac{29}{16}} dr \int_{r-s}^{r+3s} dt \int_{\partial B_t} |Df'|^2 d\mathcal{H}^{n-1}\\ &\le 4\,s\int_{\frac{9}{8}-s}^{2} dt \int_{\partial B_t} |Df'|^2 d\mathcal{H}^{n-1} \le 4\,s\int_{B_2} |Df'|^2,\end{aligned}$$

from which the conclusion follows for C big enough:

$$\begin{aligned}&\left|\left\{r \in \left[\frac{9}{8}, \frac{29}{16}\right] : \int_{B_{r+3s}\setminus B_{r-s}} |Df'|^2 \ge C\,s\int_{B_2} |Df'|^2\right\}\right|\\ &\le \frac{1}{C\,s\int_{B_2}|Df'|^2}\int_{\frac{9}{8}}^{\frac{29}{16}} dr \int_{B_{r+3s}\setminus B_{r-s}} |Df'|^2 \le \frac{4}{C} < \frac{1}{8}.\end{aligned}$$

For $\sigma = r_3 \in B = 2s + B'$ we then conclude, for some $\gamma(\beta_1, \bar n, n, m, Q) > 0$,

$$\int_{B_\sigma} |Dg|^2 \le \int_{B_\sigma\cap K} |Df|^2 + C E^{\gamma}(E+\mathbf{A}^2)\,. \qquad \square$$

6.2. **Proof of Theorem 6.1.** Choose $\beta_1 = \frac{1}{4m}$ and consider the set $B \subset [9/8, 2]$ given in Proposition 6.3. Using the coarea formula and the isoperimetric inequality (the argument and the map φ are the same in the proof of Theorem 4.2 and that of Proposition 5.4), we find $s \in B$ and an integer rectifiable current R such that

$$\partial R = \langle T - \mathbf{G}_f, \varphi, s\rangle \quad \text{and} \quad \mathbf{M}(R) \leq C E^{\frac{2m-1}{2m-2}}.$$

Since $g|_{\partial B_s} = f|_{\partial B_s}$ and g takes values in Σ, we can use g in place of f in the estimates and, arguing as before (see e.g. the proof of Theorem 5.4), we get, for a suitable $\gamma > 0$:

$$\|T\|(\mathbf{C}_s) \leq Q\,|B_s| + \int_{B_s} \frac{|Dg|^2}{2} + CE^{1+\gamma} \stackrel{(6.3)}{\leq} Q\,|B_s| + \int_{B_s\cap K} \frac{|Df|^2}{2} + CE^{\gamma}(E + \mathbf{A}^2)\,. \tag{6.22}$$

On the other hand, by Taylor's expansion in Remark 4.4,

$$\begin{aligned}\|T\|(\mathbf{C}_s) &= \|T\|((B_s \setminus K)\times\mathbb{R}^n) + \|\mathbf{G}_f\|((B_s\cap K)\times\mathbb{R}^n)\\ &\geq \|T\|((B_s \setminus K)\times\mathbb{R}^n) + Q\,|K\cap B_s| + \int_{K\cap B_s}\frac{|Df|^2}{2} - C\,E^{1+\gamma}.\end{aligned}\tag{6.23}$$

Hence, from (6.22) and (6.23), we get $\mathbf{e}_T(B_s\setminus K) \leq C\,E^{\gamma}\,(E + \mathbf{A}^2)$.

This is enough to conclude the proof. Indeed, let $A \subset B_{9/8}$ be a Borel set. Using the higher integrability of $|Df|$ in K (see (6.4)) and possibly selecting a smaller $\gamma > 0$, we get

$$\begin{aligned}\mathbf{e}_T(A) &\leq \mathbf{e}_T(A\cap K) + \mathbf{e}_T(A\setminus K) \leq \int_{A\cap K}\frac{|Df|^2}{2} + C\,E^{\gamma}\left(E + \mathbf{A}^2\right)\\ &\leq C\,|A\cap K|^{\frac{p_1-1}{p_1}}\left(\int_{A\cap K}|Df|^{q_1}\right)^{2/q_1} + C\,E^{\gamma}\left(E+\mathbf{A}^2\right)\\ &\leq C\,|A|^{\frac{p_1-1}{p_1}}\left(E+\mathbf{A}^2\right) + C\,E^{\gamma}\left(E+\mathbf{A}^2\right).\end{aligned}$$

6.3. **Proofs of Theorems 1.4 and 1.6.** As usual we assume, w.l.o.g., $r = 1$ and $x = 0$. Choose $\beta_{11} < \min\{\frac{1}{2m}, \frac{\gamma_{11}}{2(1+\gamma_{11})}\}$, where γ_{11} is the constant in Theorem 6.1. Let f be the $E^{\beta_{11}}$-Lipschitz approximation of T. Clearly (1.4) follows directly from Proposition 2.2 if $\gamma_1 < \beta_{11}$. Set next $A := \left\{\mathbf{m}\mathbf{e}_T > 2^{-m}E^{2\,\beta_{11}}\right\} \cap B_{9/8}$. By Proposition 2.2, $|A| \leq CE^{1-2\beta_{11}}$. If ε_1 is sufficiently small, apply (2.1) and estimate (6.1) to A to conclude:

$$|B_1\setminus K| \leq C\,E^{-2\,\beta_{11}}\,\mathbf{e}_T\,(A) \leq C\,E^{\gamma_{11}-2\beta_{11}(1+\gamma_{11})}(E+\mathbf{A}^2).$$

By our choice of γ_{11} and β_{11}, this gives (1.5) for some positive γ_1. Finally, set $S = \mathbf{G}_f$. Recalling the strong Almgren estimate (6.1) and the Taylor expansion in Remark 4.4, we conclude: for every $0 < \sigma \leq 1$

$$\begin{aligned}&\left|\|T\|(\mathbf{C}_\sigma) - Q\,\sigma^m\,\omega_m - \int_{B_\sigma}\frac{|Df|^2}{2}\right| \leq \mathbf{e}_T(B_\sigma\setminus K) + \mathbf{e}_S(B_\sigma\setminus K) + \left|\mathbf{e}_S(B_\sigma) - \int_{B_\sigma}\frac{|Df|^2}{2}\right|\\ &\qquad \leq C\,E^{\gamma_{11}}(E+\mathbf{A}^2) + C\,|B_\sigma\setminus K| + C\,\mathrm{Lip}(f)^2\int_{B_\sigma}|Df|^2 \leq C\,E^{\gamma_1}(E+\mathbf{A}^2).\end{aligned}$$

The L^∞ bound follows from Proposition 2.2 recalling that, by Remark 1.5, we can assume $\|D\Psi\|_0 \leq C(E^{1/2} + \mathbf{A})$. Finally, Theorem 1.6 is a special case of Theorem 4.2, since the map f in Theorem 1.4 is the E^{γ_1}-Lipschitz approximation of T.

7. The "almost" projections $\boldsymbol{\rho}^\star_\delta$

In this section we show the existence of the maps $\boldsymbol{\rho}^\star_\delta$ in Proposition 6.2. Compared to the original ones introduced by Almgren, our $\boldsymbol{\rho}^\star_\delta$'s have the advantage of depending on a single parameter. Our proof is different from Almgren's and gives more explicit estimates, relying heavily on the following simple corollary of Kirszbraun's Theorem.

Lemma 7.1. *Let $f : \Omega \subset \mathbb{R}^{N_1} \to C \subset \mathbb{R}^{N_2}$ be a Lipschitz function and assume that C is closed and convex. Then, there is an extension $\hat{f}$ of f to the whole $\mathbb{R}^{N_1}$ which preserves the Lipschitz constant and takes values in C.*

To prove Lemma 7.1 it suffices to take the map $\tilde{f}$ of the classical statement of Kirszbraun's theorem (see [20, Theorem 2.10.43]) which takes values in $\mathbb{R}^{N_2}$ and compose it with the orthogonal projection π_C onto the convex closed set C, which is a 1-Lipschitz map in $\mathbb{R}^{N_2}$.

Proof of Proposition 6.2. The proof consists of four parts: the first one is a detailed description of the set $\mathcal{Q}$, whereas the remaining three give a rather explicit construction in this order:

(1) first we specify $\boldsymbol{\rho}^\star_\delta$ on $\mathcal{Q}$: the resulting map will be called $\boldsymbol{\rho}^\flat$;
(2) then we extend it to a map $\boldsymbol{\rho}^\sharp$ on $\mathcal{Q}_{\delta^{nQ+1}}$, the δ^{nQ+1}-neighborhood of $\mathcal{Q}$; $\boldsymbol{\rho}^\sharp$ will satisfy $\mathrm{Lip}(\boldsymbol{\rho}^\sharp) \leq 1 + C\delta^{8^{-\bar{n}Q-1}}$ and $|\boldsymbol{\rho}^\sharp(P) - P| \leq C\delta^{8^{-\bar{n}Q}}$ for every $p \in \mathcal{Q}$;
(3) we then extend it to all $\mathbb{R}^N$ keeping its Lipschitz constant bounded.

(3) follows easily from (2): we consider $\boldsymbol{\xi}^{-1} \circ \boldsymbol{\rho}^\sharp : \mathcal{Q}_{\delta^{\bar{n}Q+1}} \to \mathcal{A}_Q$ and a Lipschitz extension $h : \mathbb{R}^N \to \mathcal{A}_Q$ of it with $\mathrm{Lip}(h) \leq C$, using [14, Theorem 1.7]. Our map is then $\boldsymbol{\rho}^\star_\delta := \boldsymbol{\xi} \circ h$. Then (6.2) is an easy consequence of (2), (3) and the chain rule.

The description of $\mathcal{Q}$ and the proofs of (1) and (2) are given in the next subsections. $\square$

From now on we use n instead of $\bar{n}$ to simplify the notation.

7.1. Conical simplicial structure of $\mathcal{Q}$.

We first prove that $\mathcal{Q}$ is the union of families $\{\mathcal{F}_i\}_{i=0}^{nQ}$ of sets, the "i-dimensional faces" of $\mathcal{Q}$, with the following properties:

(p1) $\mathcal{Q} = \cup_i \cup_{F \in \mathcal{F}_i} F$;
(p2) $\mathcal{F} := \cup_i \mathcal{F}_i$ is a collection of finitely many disjoint sets;
(p3) each face $F \in \mathcal{F}_i$ is a convex *open* i-dimensional cone, where open means that for every $x \in F$ there exists an i-dimensional disk D with $x \in D \subset F$;
(p4) for each $F \in \mathcal{F}_i$, $\bar{F} \setminus F$ is the union of some elements of $\cup_{j<i} \mathcal{F}_j$.
(p5) for each $i < k \leq nQ$ and for each $F \in \mathcal{F}_i$, there exists $G \in \mathcal{F}_k$ such that $F \subset \bar{G}$.

Remark 7.2. With a small abuse of notation ∂F will denote $\overline{F} \setminus F$ for any $F \in \mathcal{F}$.

So, $\mathcal{F}_0 = \{0\}$; $\mathcal{F}_1$ consists of finitely many half-lines meeting at 0, i.e. of sets of type $l_v = \{\lambda\, v \,:\, \lambda \in]0, +\infty[\}$ for $v \in \mathbb{S}^{N-1}$; $\mathcal{F}_2$ consists of finitely many 2-dimensional "infinite triangles" delimited by pairs of half lines l_{v_1}, $l_{v_2} \in \mathcal{F}_1$ and by $\{0\}$; and so on. To prove this statement, first of all we recall the construction of $\boldsymbol{\xi}$ (see [14, 2.1.2]). After selecting a suitable finite collection of non zero vectors $\{e_k\}_{k=1}^h$ (in general $h > n$), we define the linear map $L : \mathbb{R}^{n\,Q} \to \mathbb{R}^N$ with $N := h\,Q > n\,Q$ given by

$$L(P_1, \ldots, P_Q) := \big(\underbrace{P_1 \cdot e_1, \ldots, P_Q \cdot e_1}_{w^1}, \underbrace{P_1 \cdot e_2, \ldots, P_Q \cdot e_2}_{w^2}, \ldots, \underbrace{P_1 \cdot e_h, \ldots, P_Q \cdot e_h}_{w^h}\big)\,.$$

Then, we consider the map $O : \mathbb{R}^N \to \mathbb{R}^N$ which maps $(w^1 \ldots, w^h)$ into the vector $(v^1, \ldots, v^h)$ where each v^i is obtained from w^i ordering its components in increasing order. Note that the composition $O \circ L : (\mathbb{R}^n)^Q \to \mathbb{R}^N$ is now invariant under the action of the symmetric group $\mathscr{P}_Q$. Therefore, $\boldsymbol{\xi}$ is simply the induced map on $\mathcal{A}_Q = (\mathbb{R}^n)^Q/\mathscr{P}_Q$ and $\mathcal{Q} = \boldsymbol{\xi}(\mathcal{A}_Q) = O(V)$ where $V := L(\mathbb{R}^{n\,Q})$. Moreover, since the vectors e_i's span $\mathbb{R}^n$ (cf. [14, 2.1.2]), the map L is injective and thus V is an nQ-dimensional subspace.

Consider the following equivalence relation $\sim$ on V:

$$(w^1, \ldots, w^h) \sim (z^1, \ldots, z^h) \qquad \text{if} \qquad \begin{cases} w^i_j = w^i_k & \Leftrightarrow z^i_j = z^i_k \\ w^i_j > w^i_k & \Leftrightarrow z^i_j > z^i_k \end{cases} \quad \forall\, i, j, k\,, \tag{7.1}$$

where $w^i = (w^i_1, \ldots, w^i_Q)$ and $z^i = (z^i_1, \ldots, z^i_Q)$: if $w \sim z$, then O rearranges their components with the same permutation. We let $\mathcal{E}$ denote the set of corresponding equivalence classes in V and $\mathcal{C} := \{L^{-1}(E) : E \in \mathcal{E}\}$. The following fact is an obvious consequence of definition (7.1):

$$L(P) \sim L(S) \quad \text{if and only if} \quad L(P_{\pi(1)}, \ldots, P_{\pi(Q)}) \sim L(S_{\pi(1)}, \ldots, S_{\pi(Q)}) \quad \forall\, \pi \in \mathscr{P}_Q\,.$$

Thus, $\pi(C) \in \mathcal{C}$ for every $C \in \mathcal{C}$ and every $\pi \in \mathscr{P}_Q$. Since $\boldsymbol{\xi}$ is injective and is induced by $O \circ L$, it follows that, for every pair E_1, $E_2 \in \mathcal{E}$, either $O(E_1) = O(E_2)$ or $O(E_1) \cap O(E_2) = \emptyset$. Therefore, the family $\mathcal{F} := \{O(E) : E \in \mathcal{E}\}$ is a partition of $\mathcal{Q}$.

Clearly, each $E \in \mathcal{E}$ is a convex cone. Let i be its dimension and D any i-dimensional disk $D \subset E$. Denote by x the center of D and let y be any other point of E. Then, by (7.1), the point $z = y - \varepsilon(x - y) = (1 + \varepsilon)\,y - \varepsilon\,x$ belongs as well to E for any $\varepsilon > 0$ sufficiently small. The convex envelope of $D \cup \{z\}$, which is contained in E, contains in turn an i-dimensional disk centered in y: therefore E is an open convex cone. Since $O|_E$ is a linear injective map, $F = O(E)$ is an open convex cone of dimension i. Therefore, $\mathcal{F}$ satisfies (p1)-(p3).

Next notice that, having fixed $w \in E$, a point z belongs to $\bar{E} \setminus E$ if and only if

(1) $w^i_j \geq w^i_k$ implies $z^i_j \geq z^i_k$ for every i, j and k;
(2) there exists r, s and t such that $w^r_s > w^r_t$ and $z^r_s = z^r_t$.

Thus, if d is the dimension of E, $\partial E := \overline{E} \setminus E$ (cf. Remark 7.2) is the union of some elements of $\cup_{j<d} \mathcal{E}_j$, where with $\mathcal{E}_j$ we denote the j-dimensional elements of $\mathcal{E}$. Observe that, since O is continuous, we must have $\overline{F} \supset O(\overline{E})$. On the other hand, if $x \in \overline{F}$ and $x_k \to x$ is a sequence contained in F, then there is a sequence $\{y_k\} \subset E$ with $O(y_k) = x_k$. By the definition of O the sequence $\{y_k\}$ is bounded and hence, up to subsequence, we can assume that it converges to $y \in \overline{F}$: thus $O(y) = x$ and $O(\overline{E}) = \overline{F}$. On the other hand, for

equivalence classes E_1, E_2 of different dimension we necessarily have $O(E_1) \cap O(E_2) = \emptyset$. Thus $O(\partial E) \cap O(E) = \emptyset$, i.e. $\partial F = O(\partial E)$, which shows ($\mathrm{p}_4$).

For what concerns (p_5) we show first that if $L(P) = z \in E \in \mathcal{E}$ is such that $z^i_j \neq z^i_k$ for all i and for all $j \neq k$, then $O(E) \in \mathcal{F}_{nQ}$. Indeed, if $t < {}^1\!/_4 \min_{i,j\neq k} |z^i_j - z^i_k|$, then $L(P + v) \in E$ for every $v \in B_t(0) \subset \mathbb{R}^{nQ}$, i.e. E is an (nQ)-dimensional convex cone. Therefore it follows that for every $F \in \mathcal{F}_i$ with $i < nQ$ there exists $G \in \mathcal{F}_{nQ}$ such that $F \subset \bar{G}$. To show this claim it is enough to prove that, if $F = O(E)$ and $L(P) = z \in E$, then z is the limit of points $w \in V$ such that $w^i_j \neq w^i_k$ for all i, j, k, which can be easily proved by a simple perturbation argument. Next, we argue inductively on k: knowing that $F \in \mathcal{F}_i$ is contained in $\overline{G}$ for some $G \in \mathcal{F}_k$ with $k > i + 1$, we show that there is $H \in \mathcal{F}_{k-1}$ such that $F \subset \overline{H}$. Observe indeed that $F \subset \partial G = \overline{G} \setminus G$ and that, for dimensional reasons, $\overline{G} \setminus G$ must be contained in the closure of those $H \in \mathcal{F}_{k-1}$ such that $H \subset \overline{G}$. Let $H \in \mathcal{F}_{k-1}$ be such that $F \cap \overline{H} \neq \emptyset$. Consider $E, K \in \mathcal{E}$ such that $F = O(E)$ and $H = O(K)$. Let $x \in E$ such that $O(x) \in F \cap \overline{H}$ and $z \in K$. We then must have that $x^i_k \geq x^i_j$ whenever $z^i_k > z^i_j$ and that $x^i_k = x^i_j$ whenever $z^i_k = z^i_j$. By the very definition of $\sim$, the same property holds even if we replace x with another element $\xi \in E$. Therefore the open segment $]\xi, z[$ must be contained in K, which in turn implies that $\xi \in \overline{K}$. Thus we conclude $F \subset \overline{H}$.

7.2. **Construction of $\boldsymbol{\rho}^\flat$.** The main building block in the construction of $\boldsymbol{\rho}^\flat$ is given by the following lemma.

Lemma 7.3. *For $\tau \in]0, \frac{1}{4}[$ and any $D \in \mathbb{N} \setminus \{0\}$ consider the map $\Phi_\tau : \mathbb{R}^D \to \mathbb{R}^D$ defined by:*

$$\Phi_\tau(x) = \begin{cases} 0 & \text{if } |x| \leq \tau \\ \sqrt{\tau}\,\frac{|x|-\tau}{\sqrt{\tau}-\tau}\,\frac{x}{|x|} & \text{if } \tau \leq |x| \leq \sqrt{\tau} \\ x & \text{if } |x| \geq \sqrt{\tau}. \end{cases}$$

Then $|\Phi_\tau(x) - x| \leq \tau$ and $\mathrm{Lip}(\Phi_\tau) \leq 1 + 2\sqrt{\tau}$.

Proof. The proofs of the two claims are straightforward computations. First $\Phi_\tau(x) = x$ if $|x| \geq \sqrt{\tau}$ and $|\Phi_\tau(x) - x| = |x| \leq \tau$ if $|x| \leq \tau$. For $\tau \leq |x| \leq \sqrt{\tau}$ we compute

$$|\Phi_\tau(x) - x| = \left| \frac{\sqrt{\tau}\,(|x|-\tau)}{\sqrt{\tau}-\tau} - |x| \right| = \tau\,\frac{\sqrt{\tau}-|x|}{\sqrt{\tau}-\tau} \leq \tau.$$

Next we show that $|D\Phi_\tau(x) \cdot v| \leq (1 + 2\sqrt{\tau})|v|$ at any point of differentiability. This inequality obviously imply the claimed Lipschitz constant estimate because Φ_τ is Lipschitz and its domain of definition is a convex set. The inequality is, moreover, obvious when $|x| < \tau$ and $|x| > \sqrt{\tau}$. For $\tau < |x| < \sqrt{\tau}$, we can compute

$$D\Phi_\tau(x) = \frac{1 - \frac{\tau}{|x|}}{1 - \sqrt{\tau}}\mathrm{Id} + \frac{\frac{\tau}{|x|}}{1 - \sqrt{\tau}}\,\frac{x}{|x|} \otimes \frac{x}{|x|}\,.$$

The matrix is symmetric with positive eigenvalues (because $|x| > \tau$) and the maximal eigenvalue is $(1 - \sqrt{\tau})^{-1} \leq 1 + 2\sqrt{\tau}$, thereby proving our claim. □

7.2.1. *Special coordinates, conical sections and separation.* Let S_k be the k-dimensional skeleton of $\mathcal{Q}$, i.e. the union of $F \in \mathcal{F}_k$ and denote by $(S_k)_\sigma$ its σ-neighborhood $\{x : \operatorname{dist}(x, S_k) < \sigma\}$. Incidentally, $(S_k)_\sigma$ contains $(S_i)_\sigma$ for every $i < k$.

Definition 7.4 (Coordinates and conical sections)**.** Fix any face $F \in \mathcal{F}_k$ and introduce Cartesian coordinates $(y, z) \in \mathbb{R}^k \times \mathbb{R}^{N-k}$ in such a way that $F \subset \mathbb{R}^k \times \{0\}$. For a positive constant $\tilde{c}$ consider the cone $\mathscr{C}(F) := \{(y, z) \in \mathcal{Q} : (y, 0) \in F \,, |z| \leq \tilde{c} \operatorname{dist}((y, 0), S_{k-1})\}$. For any $p = (y, 0) \in F$ we set $V_p := (\{y\} \times \mathbb{R}^{N-k}) \cap \mathscr{C}(F)$.

Note that, if $\tilde{c}$ is sufficiently small, we will have the following property

$$\mathscr{C}(F) \cap \mathscr{C}(G) \neq \emptyset \quad \Longrightarrow \quad \text{either } F \subset \overline{G} \text{ or } G \subset \overline{F}.$$

For every constants $a, b > 0$, $k = 1 \ldots, nQ - 1$ and $F \in \mathcal{F}_k$, we fix coordinates as in Definition 7.4 and denote by $F_{a,b}$ the sets

$$F_{a,b} := \big\{(y, z) \,: |z| \leq a, (y, 0) \in F \setminus (S_{k-1})_b\big\}.$$

For the faces $F \in \mathcal{F}_{nQ}$ of maximal dimension and for every $a > 0$, $F_{\star,a}$ denotes the set $F_{\star,a} := F \setminus (S_{nQ-1})_a$. The following lemma is an obvious corollary of the linear simplicial and conical structures of $\mathcal{Q}$.

Lemma 7.5. *There is a constant $\bar{c} > 0$ (independent of a, b below) with the following property. Assume F and G are two distinct k dimensional faces.*

- *If $k = nQ$, $a > 0$, $x \in F_{\star,a}$ and $x' \in G_{\star,a}$, then $|x - x'| \geq \bar{c}a$;*
- *If $k < nQ$, $b/a > \bar{c}^{-1}$, $x \in F_{a,b}$ and $x' \in G_{a,b}$, then $|x - x'| \geq \bar{c}b$.*

Moreover, if $F \in \mathcal{F}_k$, $H \in \mathcal{F}_i$ with $i > k$ and $F \not\subset \partial H$ (cf. Remark 7.2), then $|x - x'| \geq \bar{c}a$ for every $x \in H$ and $x' \in F \setminus (S_{k-1})_a$.

7.2.2. *The domains* $\operatorname{Dom}(f_k)$. Next we choose constants $c_k := \delta^{8^{-nQ+k}}$. If δ is small enough, each family $\{F_{2\sqrt{c_k}, c_{k-1}^2}\}_{F \in \mathcal{F}_k}$ with $k < nQ$ is made by pairwise disjoint sets, which are at least $\bar{c}c_{k-1}^2$ far apart, where $\bar{c}$ is the constant of Lemma 7.5, and it holds $F_{2\sqrt{c_k}, c_{k-1}^2} \subset \mathscr{C}(F) \subset \mathcal{Q}$. We are ready to define the map $\boldsymbol{\rho}^\flat := \boldsymbol{\rho}^\star|_{\mathcal{Q}}$ inductively "from the top to the bottom". More precisely we will define a family of maps $\{f_k\}_{k \in \{0, \ldots, nQ\}}$ on domains $\operatorname{Dom}(f_k) \subset \mathcal{Q}$ starting from f_{nQ} and ending with $f_0 = \boldsymbol{\rho}^\flat$. We first explicitly define $\operatorname{Dom}(f_k) := \mathcal{Q} \setminus (S_{k-1})_{c_{k-1}}$ for $k > 0$ and $\operatorname{Dom}(f_0) = \mathcal{Q}$, and in order to simplify our notation we then agree that $c_{-1} = \delta^{8^{-nQ-1}}$ and $S_{-1} = (S_{-1})_{c_{-1}} = \emptyset$. Note that $\operatorname{Dom}(f_{k+1}) \not\subset \operatorname{Dom}(f_k)$. It is obvious that

$$\operatorname{Dom}(f_k) = \Big(\operatorname{Dom}(f_{k+1}) \cup \bigcup_{F \in \mathcal{F}_k} F_{2\sqrt{c_k}, c_{k-1}^2}\Big) \setminus (S_{k-1})_{c_{k-1}} \,. \tag{7.2}$$

Indeed, if $x \in \operatorname{Dom}(f_k) \setminus \operatorname{Dom}(f_{k+1})$ we then must have $\operatorname{dist}(x, S_k) < c_k$ and $\operatorname{dist}(x, S_{k-1}) \geq c_{k-1}$. Let $q \in S_k$ be such that $|x - q| < c_k$. Since $\operatorname{dist}(x, S_{k-1}) \geq c_{k-1} > c_k$, the point q must necessarily belong to a k-dimensional face F. Fix coordinates as in Definition 7.4. If $x = (y, z)$, we then obviously have $|z| < c_k \leq 2\sqrt{c_k}$. On the other hand $\operatorname{dist}((y, 0), S_{k-1}) \geq \operatorname{dist}(x, S_{k-1}) - |z| \geq c_{k-1} - c_k > c_{k-1}^2$. This shows that $x \in F_{2\sqrt{c_k}, c_{k-1}^2}$.

7.2.3. *The maps f_k.* On $\mathrm{Dom}(f_{nQ})$ we define $f_{nQ} = \mathrm{Id}$ and specify next the procedure to define f_k knowing f_{k+1}. Along the procedure we claim inductively the following.

Assumption 7.6 (Inductive step)**.** *The map f_{k+1} has the following three properties.*

(a_{k+1}) $\mathrm{Lip}(f_{k+1}) \le 1 + C c_{k+1}^{1/2}$ *and* $|f_{k+1}(x) - x| \le C\, c_{k+1}$.

(b_{k+1}) *Consider $i \le k+1$, an i-dimensional face F, the cone $\mathscr{C}(F)$ in Definition 7.4 and the corresponding coordinates. Then, f_{k+1} factorizes on* $\mathrm{Dom}(f_{k+1}) \cap \mathscr{C}(F)$ *as*

$$f_{k+1}(y,z) = (y, h^F_{k+1}(y,z)) \in \mathbb{R}^i \times \mathbb{R}^{N-i}. \tag{7.3}$$

(c_{k+1}) *For every $G \in \mathcal{F}_i$ with $i \ge k+1$, f_{k+1} maps* $\mathrm{Dom}(f_{k+1}) \cap \{x : \mathrm{dist}(x, G) < \delta\}$ *into $\overline{G}$. Moreover the restriction of f_{k+1} to G_{c_i,c_k} is the orthogonal projection onto G.*

The constants involved depend on k but not on the parameter δ and since the process is iterated finitely many times, we will not keep track of such dependence. Note that f_{nQ} satisfies (a_{nQ}), (b_{nQ}) and (c_{nQ}) trivially, because it is the identity map. Given f_{k+1} we next show how to construct f_k. For every $p \in G \in \mathcal{F}_k$ with $p \notin (S_{k-1})_{c^2_{k-1}}$, set coordinates as in Definition 7.4 and consider the cone $W_p := \{(y,z) \in V_p : |z| \le 2\sqrt{c_k}\}$. Let now Φ_τ be the map of Lemma 7.3 with $\tau = 2c_k$. The function f_k is defined in W_p by

$$f_k(x) = f_k(y,z) := (y, h^F_k(y,z)) := \begin{cases} (y,0) & \text{for } |z| \le \tau/2 = c_k, \\ \big(y, \Phi_\tau(h^F_{k+1}(y,z))\big) & \text{otherwise.} \end{cases} \tag{7.4}$$

If $q \in \mathrm{Dom}(f_k)$ does not belong to any W_p as above, then we set $f_{k+1}(q) = f_k(q)$.

Observe that the definition above gives values to f_k on a set which is larger than $\mathrm{Dom}(f_k)$: this will be useful to carry on some of the estimates, but we insist that Assumption 7.6 will only be checked on $\mathrm{Dom}(f_k)$.

7.2.4. *Well-definition and continuity.* Consider a point $q \in \mathrm{Dom}(f_k)$. If q is not contained in $F_{2\sqrt{c_k}, c^2_{k-1}}$ for some k-dimensional face, then by (7.2) it is contained in the domain of f_{k+1} and thus $f_k(q)$ is defined. If q is contained in $F_{2\sqrt{c_k}, c^2_{k-1}}$ for some k-dimensional face, then q belongs to some W_p as above. Let $q = (y,z)$. If $|z| \le c_k$, then $f_k(q)$ is defined; otherwise, since $\mathrm{dist}(q, S_k) \ge c_k$, we infer that $q \in \mathrm{Dom}(f_{k+1})$ and $f_k(q)$ is also defined.

As for the continuity, fix $(y,z) \in W_p \cap \mathrm{Dom}(f_k)$ with $p = (y,0) \in F \in \mathcal{F}_k$. If $|z| = c_k$, then by (a_{k+1}) we have $|h^F_{k+1}(y,z)| \le |z| + C c_{k+1} \le \tau/2 + C\tau^8$. For δ sufficiently small this obviously implies $|h^F_{k+1}(y,z)| \le \tau$ and thus, by the definition of Φ_τ, $\Phi_\tau(h^F_{k+1}(y,z)) = 0$. On the other hand, if $|z| = 2\sqrt{c_k}$, then $|h^F_{k+1}(y,z)| \ge |z| - C c_{k+1} = 2\sqrt{c_k} - C c_k^8 \ge \sqrt{2c_k}$ and thus $\Phi_\tau(h^F_{k+1}(y,z)) = h^F_{k+1}(y,z)$. Therefore under this assumption we have $f_{k+1}(q) = f_k(q)$.

We next check that f_k maps $\mathrm{Dom}(f_k)$ into $\mathcal{Q}$. This is true by induction where f_k coincides with f_{k+1}. Fix therefore a point q in some $W_p \cap \mathrm{Dom}(f_k)$ with $p \in F \in \mathcal{F}_k$ and let G be the i-dimensional face containing q with $i > k$. Then, $f_{k+1}(q)$ belongs to a face $\overline{G}$, by Assumption 7.6. By the estimate in (a_{k+1}) and the assumption (b_{k+1}), the face G must intersect $\mathscr{C}(F)$ and thus $F \subset \bar{G}$. Observe that, by the properties of Φ_τ and by the inductive assumption (b_{k+1}), $f_k(q)$ is mapped in the segment joining $f_{k+1}(q)$ and q and thus must belong to $\overline{G}$.

7.2.5. *The inductive conclusions* (c_k) *and* (b_k). The first claim of (c_k) is simple to prove: as noticed, if a point $q \in \mathrm{Dom}(f_k)$ belongs also to $\mathrm{Dom}(f_{k+1})$, then f_k maps it into the closure of the face containing q. If the point is not contained in $\mathrm{Dom}(f_{k+1})$, then it must be contained in the c_k-neighborhood of some k-dimensional face F and hence it is mapped into F: when this happens F is a portion of the boundary of the face containing q. Next, fix a face $G \in \mathcal{F}_i$. If $i = k$, by the very definition of f_k, we have that the restriction of f_k to $\mathrm{Dom}(f_k) \cap G_{c_k, c_{k-1}}$ is the orthogonal projection onto G. If $i > k$, we actually have that $f_k = f_{k+1}$ on $\mathrm{Dom}(f_k) \setminus (S_k)_{2\sqrt{c_k}} \supset \mathcal{Q} \setminus (S_{i-1})_{c_{k-1}}$.

Fix now an i-dimensional face L with $i \leq k$, consider coordinates $\mathbb{R}^i \times \mathbb{R}^{n-i}$ as in Definition 7.4 and the corresponding $\mathscr{C}(L)$. If $q = (y, 0) \in L$, the condition (b_k) is equivalent to saying that $V_q \cap \mathrm{Dom}(f_k)$ gets mapped into $\{(y, 0)\} \times \mathbb{R}^{N-i}$. Fix a point $\tilde{q} \in V_q$. If $f_{k+1}(\tilde{q}) = f_k(\tilde{q})$ there is nothing to prove. Otherwise it turns out that there is a k-dimensional face F such that $\tilde{q} \in \mathscr{C}(F)$. But then we necessarily have $L \subset \bar{F}$. So, set coordinates $\mathbb{R}^i \times \mathbb{R}^{k-i} \times \mathbb{R}^{n-k}$ so that at the same time $L \subset \mathbb{R}^i \times \{0\} \times \{0\}$ and $F \subset \mathbb{R}^i \times \mathbb{R}^{k-i} \times \{0\}$. Thus, $(y, 0, 0)$ is the coordinate of q and (y, z, w) that of $\tilde{q}$. According to our definition of f_k, $f_k(\tilde{q}) = (y, z, w')$ for some w', which indeed implies the desired claim.

7.2.6. C^0 *estimate.* Observe that, for every x where f_k coincides with f_{k+1}, we have $|f_k(x) - x| \leq C c_{k+1} \leq C c_k^8$. Instead, for any point x where f_k is newly defined, we distinguish the following two cases: either $x = (y, z)$ with $|z| \leq c_k$, in which case $|f_k(x) - x| \leq c_k$; or $x = (y, z)$ with $|z| > c_k$, and then by the estimates of Lemma 7.3 and the triangle inequality we have

$$|f_k(x) - x| \leq |f_{k+1}(x) - f_k(x)| + |f_{k+1}(x) - x| \leq C c_{k+1} + \tau \leq C c_{k+1} + 2 c_k\,.$$

7.2.7. *Lipschitz estimate.* We fix $x, x' \in \mathrm{Dom}(f_k)$ and, apart from the trivial one $f_k(x) = f_{k+1}(x)$ and $f_k(x') = f_{k+1}(x')$, we distinguish three cases.

Case 1: $x, x' \in G_{2\sqrt{c_k}, c_{k-1}^2}$ for some k-dimensional face G. Choosing coordinates as in Definition 7.4, we set $x = (y, z)$ and $x' = (y', z')$. If both $|z|, |z'| \leq \frac{\tau}{2}$, then $|f_k(x) - f_k(x')| = |y - y'| \leq |x - x'|$. If $|z| \geq \frac{\tau}{2}$ and $|z'| \geq \frac{\tau}{2}$, then

$$\begin{aligned}
|f_k(x) - f_k(x')|^2 \leq & |y - y'|^2 + (1 + 2\tau^{1/2})^2 |h^F_{k+1}(y, z) - h^F_{k+1}(y', z')|^2 \\
\leq & (1 + 2\tau^{1/2})^2 \left(|y - y'|^2 + |h^F_{k+1}(y, z) - h^F_{k+1}(y', z')|^2\right) \\
= & (1 + \sqrt{2c_k})^2 |f_{k+1}(x) - f_{k+1}(x')|^2 \leq (1 + \sqrt{2c_k})^2 (1 + C\sqrt{c_{k+1}})^2 |x - x'|^2\,.
\end{aligned}$$

If $|z| \leq \frac{\tau}{2}$ and $|z'| > \frac{\tau}{2}$, let $\tilde{z}$ be the point with $|\tilde{z}| = \frac{\tau}{2}$ on the segment joining z and z', and $\tilde{x} = (y, \tilde{z})$. Observe that $f_k(\tilde{x}) = f_k(x) = (y, 0)$ and that $|\tilde{x} - x'|^2 = |y - y'|^2 + |z' - \tilde{z}|^2 \leq |y - y'|^2 + |z - z'|^2 \leq |x - x'|^2$. On the other hand we have just shown $|f_k(x') - f_k(\tilde{x})| \leq (1 + C c_k^{1/2}) |x' - \tilde{x}|$.

Case 2: $x \in F_{2\sqrt{c_k}, c_{k-1}^2}$, $x' \in G_{2\sqrt{c_k}, c_{k-1}^2}$ for distinct $F, G \in \mathcal{F}_k$. By Lemma 7.5, $|x - x'| \geq \bar{c}\, c_{k-1}^2 \geq \bar{c} c_k^{1/4}$. On the other hand, we also have, by the C^0 estimate,

$$|f_k(x) - f_k(x')| \leq |x - x'| + C c_k \leq (1 + C c_k^{3/4}) |x - x'|\,.$$

Case 3: $x \in G_{2\sqrt{c_k}, c_{k-1}^2}$ **for some** k**-dimensional face** G **and** $f_k(x') = f_{k+1}(x')$**.** Without loss of generality we assume

- $G \in \mathcal{F}_k$;
- $x' \notin G_{2\sqrt{c_k}, c_{k-1}^2}$;
- $x' \in H$ for some face H (of dimension $i > k$).

We have two possibilities.
Case 3a: $G \not\subset \overline{H}$. Consider the closed set $\tilde{G} := G \setminus (S_{k-1})_{c_{k-1}^2}$. By Lemma 7.5 $\mathrm{dist}(x', \tilde{G}) \geq \bar{c}c_{k-1}^2$ and thus $|x - x'| \geq \bar{c}c_{k-1}^2 - 2\sqrt{c_k} \geq \frac{\bar{c}}{2}c_{k-1}^2$. We can therefore argue as in Case 2.
Case 3b: $G \subset \overline{H}$. We then have two possibilities. The first is that $x \in \mathrm{Dom}(f_{k+1})$. Since $f_k(x') = f_{k+1}(x')$, we have $|f_k(x') - x'| \leq Cc_{k+1} = Cc_k^8$. We use the coordinates of Definition 7.4 and (a_{k+1}) to conclude $f_k(x') = f_{k+1}(y', z') = (y'', z'')$ with $|z''| \geq |z'| - C\,c_k^8 \geq 2\sqrt{c_k} - C\,c_k^8 \geq \sqrt{2c_k}$. We can therefore write $f_k(x') = (y'', \Phi_\tau(z''))$ (because $\Phi_\tau(z'') = z''$) and, hence, recalling $f_k(x) = (y, \Phi_\tau(h_{k+1}^F(y, z)))$ and $f_{k+1}(x) = (y, h_{k+1}^F(y, z))$,

$$\begin{aligned}&|f_k(x') - f_k(x)|^2 \leq |y - y''|^2 + (1 + 2\sqrt{\tau})^2 |h_{k+1}^F(y, z) - z''|^2 \\ \leq &(1 + 2\sqrt{\tau})^2 |f_{k+1}(x) - f_{k+1}(x')|^2 .\end{aligned}$$

We therefore conclude $|f_k(x') - f_k(x)| \leq (1 + C\tau^{1/2})|x' - x| \leq (1 + Cc_k^{1/2})|x' - x|$.

The second possibility is that x is not in the domain of definition of f_{k+1}. In that case x is at distance c_k from G and thus $|x - x'| \geq \sqrt{c_k}$. We then conclude that $|f_k(x) - f_k(x')| \leq |x - x'| + C\,c_k \leq (1 + C\sqrt{c_k})|x - x'|$.

7.2.8. *Summary.* After nQ steps, we get a function $f_0 = \boldsymbol{\rho}^\flat : \mathcal{Q} \to \mathcal{Q}$ which satisfies

$$\mathrm{Lip}(\boldsymbol{\rho}^\flat) \leq 1 + C\,\delta^{8^{-nQ-1}} \quad \text{and} \quad |\boldsymbol{\rho}^\flat(x) - x| \leq C\,\delta^{8^{-nQ}}, \tag{7.5}$$

$$\boldsymbol{\rho}^\flat(\{x : \mathrm{dist}(x, F) \leq \delta\}) \subset \overline{F} \quad \text{for every } F \in \mathcal{F}_k, \tag{7.6}$$

$$\boldsymbol{\rho}^\flat : F_{\delta, c_0^{1/8}} \to F \text{ is the orthogonal projection on } F \text{ for every } F \in \mathcal{F}_k. \tag{7.7}$$

7.3. **The extension $\boldsymbol{\rho}^\sharp$ of $\boldsymbol{\rho}^\flat$ to $\mathcal{Q}_{\delta^{nQ+1}}$.** Next we extend the map $\boldsymbol{\rho}^\flat : \mathcal{Q} \to \mathcal{Q}$ to the δ^{nQ+1}-neighborhood of $\mathcal{Q}$, keeping the estimate (7.5). We first observe that, since the number of all the faces is finite, when δ is small enough, there exists a constant $C = C(N)$ with the following property. Consider two distinct faces F and H in $\mathcal{F}_i$. If x, y are two points contained, respectively, in $F_{\delta^{i+1}} \setminus \cup_{j<i} \cup_{G \in \mathcal{F}_j} G_{\delta^{j+1}}$ and $H_{\delta^{i+1}} \setminus \cup_{j<i} \cup_{G \in \mathcal{F}_j} G_{\delta^{j+1}}$, then

$$\mathrm{dist}(x, y) \;\geq\; C\,\delta^i. \tag{7.8}$$

Similarly if $F \in \mathcal{F}_l$ and $H \in \mathcal{F}_i$ with $l < i$ and $F \not\subset \bar{H}$, then for every $x \in F_{\delta^{l+1}}$ and $y \in H_{\delta^{i+1}} \setminus \cup_{j<i} \cup_{G \in \mathcal{F}_j} G_{\delta^{j+1}}$ it holds

$$\mathrm{dist}(x, y) \;\geq\; C\,\delta^i. \tag{7.9}$$

The extension $\boldsymbol{\rho}^\sharp$ is defined inductively, but this time "from the bottom to the top". The first extension g_0 is identically 0 on $B_\delta(0)$ (note that this is feasible because $\boldsymbol{\rho}^\flat \equiv 0$ in

$B_\delta(0) \cap \mathcal{Q}$). Now we come to the inductive step. Suppose we have an extension g_ℓ of $\boldsymbol{\rho}^\flat$, defined on the union of the $\delta^{\ell+1}$-neighborhoods of the ℓ-skeletons S_ℓ, for $\ell \in \{0, \ldots, k\}$, i.e.

$$\mathbf{L}_k \;:=\; \mathcal{Q} \cup B_\delta(0) \cup \bigcup_{\ell=1}^{k} \bigcup_{F \in \mathcal{F}_\ell} F_{\delta^{\ell+1}}\,.$$

Assume inductively that $\mathrm{Lip}(g_k) \leq 1 + C\,\delta^{8^{-nQ-1}}$ and assume that g_k maps any δ^{j+1}-neighborhood of any j-dimensional face into its closure, when $j \leq k$. Then, we define the extension of g_k to $\mathbf{L}_{k+1}$ in the following way. For every face $F \in \mathcal{F}_{k+1}$, we set

$$g_{k+1} := \begin{cases} \boldsymbol{\rho}^\flat & \text{on} \quad \mathcal{Q}\,, \\ g_k & \text{on} \quad (S_k)_{\delta^{k+1}} \cap F_{\delta^{k+2}}\,, \\ \mathbf{p}_F & \text{on} \quad \{x \in \mathbb{R}^N \,:\, \mathbf{p}_F(x) \in F_{\delta,1}\} \cap F_{\delta^{k+2}}\,, \end{cases} \tag{7.10}$$

where $\mathbf{p}_F$ stands for the orthogonal projection on F (recall that by (7.7) $\boldsymbol{\rho}^\flat = \mathbf{p}_F$ on $F \cap F_{\delta^{k+2},1}$). Consider now a face F as above and $U(F)$ the union of all the δ^{j+1}-neighborhoods of the j-dimensional faces which belong to $\overline{F}$. As defined above, g_{k+1} maps a portion of $U(F)$ into $\overline{F}$. We can use Lemma 7.1 to extend g_{k+1} to $U(F)$ keeping the same Lipschitz constant, which we now compute. This constant is obviously smaller than $1 + C\delta^{8^{-nQ-1}}$ on the domain $((S_k)_{\delta^{k+1}} \cap F_{\delta^{k+2}}) \cup F$ by inductive hypothesis. The same constant is 1 on $\{x \in \mathbb{R}^N \,:\, \mathbf{p}_F(x) \in F_{\delta,1}\} \cap F_{\delta^{k+2}}$. Consider now a point $x \in \{x \in \mathbb{R}^N \,:\, \mathbf{p}_F(x) \in F_{\delta,1}\} \cap F_{\delta^{k+2}}$ and a point $y \in F \cup ((S_k)_{\delta^{k+1}} \cap F_{\delta^{k+2}})$. If $y \notin (S_k)_{c_0^{1/8}}$, then necessarily $y \in F$ and we then have

$$|g_{k+1}(x) - g_{k+1}(y)| = |\mathbf{p}_F(x) - y| = |\mathbf{p}_F(x) - \mathbf{p}_F(y)| \leq |x - y|.$$

Otherwise we have $|x - y| \geq 1 - c_0^{1/8} = 1 - \delta^{8^{-nQ-1}}$ and we can write

$$\begin{aligned} |g_{k+1}(x) - g_{k+1}(y)| \leq & |g_{k+1}(x) - y| + C c_0 \leq |x - y| + \delta^{k+2} + C c_0 \\ & \leq \left(1 + \frac{\delta^{k+2} + C c_0}{1 - C c_0^{1/8}}\right) |x - y| \leq (1 + C\delta^{8^{-nQ-1}})|x - y|. \end{aligned}$$

Note that, if $x \in U(F_1) \cap U(F_2)$ for two distinct $F_1, F_2 \in \mathcal{F}_{k+1}$, then $x \in \mathbf{L}_k$. Thus, the map g_{k+1} is continuous. We next bound the global Lipschitz constant of g_{k+1}. Indeed consider points $x \in U(F_1) \setminus U(F_2)$ and $y \in U(F_2) \setminus U(F_1)$ for two distinct $F_i \in \mathcal{F}_{k+1}$. Since by (7.8) and (7.9) $|x - y| \geq C\,\delta^{k+1}$, we easily see that

$$\begin{aligned} |g_{k+1}(x) - g_{k+1}(y)| &\leq |g_{k+1}(x) - g_{k+1}(\mathbf{p}_{F_1}(x))| + |g_{k+1}(\mathbf{p}_{F_1}(x)) - g_{k+1}(\mathbf{p}_{F_2}(y))| \\ &\quad + |g_{k+1}(\mathbf{p}_{F_2}(y)) - g_{k+1}(y)| \\ &\leq 2(1 + C\delta^{8^{-nQ-1}})\,\delta^{k+2} + |\boldsymbol{\rho}^\flat(\mathbf{p}_{F_1}(x)) - \boldsymbol{\rho}^\flat(\mathbf{p}_{F_2}(y))| \\ &\leq 2(1 + C\delta^{8^{-nQ-1}})\,\delta^{k+2} + (1 + C\,\delta^{8^{-nQ-1}})|\mathbf{p}_{F_1}(x) - \mathbf{p}_{F_2}(y)| \\ &\leq 2(1 + C\delta^{8^{-nQ-1}})\,\delta^{k+2} + (1 + C\,\delta^{8^{-nQ-1}})\Big(|x - y| + 2\,\delta^{k+2}\Big) \\ &\leq (1 + C\,\delta^{8^{-nQ-1}})\,|x - y|. \end{aligned}$$

Next, consider the case $x \in \mathcal{Q} \setminus U(F), y \in U(F)$. If $|x-y| \geq \delta^{k+1}$, we can then argue as above and (considering that $g_{k+1}(x) = \boldsymbol{\rho}^\flat(x)$) we bound

$$\begin{aligned}
&|g_{k+1}(x) - g_{k+1}(y)| \leq (1 + C\delta^{8^{-nQ-1}})\,\delta^{k+2} + |\boldsymbol{\rho}^\flat(x) - \boldsymbol{\rho}^\flat(\mathbf{p}_F(y))| \\
&\leq (1 + C\delta^{8^{-nQ-1}})\,\left(\delta^{k+2} + |x - \mathbf{p}_F(y)|\right) \leq (1 + C\delta^{8^{-nQ-1}})\,\left(\delta^{k+2} + |x-y| + \delta^{k+2}\right) \\
&\leq (1 + C\,\delta^{8^{-nQ-1}})\,|x-y|.
\end{aligned}$$

We therefore assume $|x-y| \leq \delta^{k+1}$. Observe also that, if $y \notin \{x \in \mathbb{R}^N : \mathbf{p}_F(x) \in F_{\delta,1}\} \cap F_{\delta^{k+2}}$, then $g_{k+1}(y) = g_k(y)$ and since $g_{k+1}(x) = \boldsymbol{\rho}^\flat(x) = g_k(x)$, we know the Lipschitz bound by inductive assumption. We therefore conclude that $x \in F_{\delta^{k+2}+\delta^{k+1}, 1-\delta^{k+1}}$. Assuming δ_0 small enough, $\delta^{k+2} + \delta^{k+1} \leq \delta$ and $1 - \delta^{k+1} \geq \delta^{8^{-nQ-1}} = c_0^{1/8}$, therefore $x \in F_{\delta, c_0^{1/8}}$. By (7.7) we then have $|g_{k+1}(x) - g_{k+1}(y)| = |\mathbf{p}_F(x) - \mathbf{p}_F(y)| \leq |x-y|$.

Since $\mathcal{Q}$ and the union of the $U(F_i)$ is the domain of definition of g_{k+1}, this shows $\mathrm{Lip}(g_{k+1}) \leq 1 + C\,\delta^{8^{-nQ-1}}$. Note that by construction we also have that $U(F)$ is mapped into $\overline{F}$, which is the other inductive hypothesis.

After making the step above nQ times we arrive to a map g_{nQ} which extends $\boldsymbol{\rho}^\flat$ and is defined in a δ^{nQ+1}-neighborhood of $\mathcal{Q}$. This is the map $\boldsymbol{\rho}^\sharp$.

8. Persistence of Q-points: Proof of Theorem 1.7

Proof of Theorem 1.7. As usual, by scaling and translating we assume $x = 0$ and $r = 1$. According to [14, Theorem 3.9], there are constants $\bar{C}(m,n,Q), \kappa(m,n,Q) > 0$ such that

$$\sup_{x \neq y \in B_{1/2}} \frac{\mathcal{G}(w(x), w(y))}{|y-x|^\kappa} \leq \bar{C}(\mathrm{Dir}(w))^{\frac{1}{2}} \quad \text{for any Dir-minimizer } w : B_1 \to \mathcal{A}_Q(\mathbb{R}^n). \tag{8.1}$$

The final choice of $\bar{s}$ will be specified at the very end, but for the moment we impose $\bar{s} < \frac{1}{4}$. Fix now $s < \bar{s}$ and $C^\star$ as in the statement and assume by contradiction that, no matter how small we choose $\hat{\varepsilon} > 0$, there are a current T and a submanifold Σ as in Theorem 1.4 and a point $(p,q) \in \mathbf{C}_{1/2}$ satisfying:

(a) $E := \mathbf{E}(T, \mathbf{C}_4) < \hat{\varepsilon}$ and $\mathbf{A}^2 \leq C^\star E$;
(b) $\Theta(T, (p,q)) = Q$;
(c) the E^{γ_1}-approximation f (which is the map of Theorem 1.4) violates (1.9), that is

$$\int_{B_s(p)} \mathcal{G}(f, Q\,[\![\boldsymbol{\eta} \circ f]\!])^2 > \hat{\delta} s^m E\,. \tag{8.2}$$

Set $\bar{\delta} = \frac{1}{4}$ and fix $\bar{\eta} > 0$ (whose choice will be specified later). By (a), for a suitably small $\hat{\varepsilon}$ we can apply Theorem 1.6 in the coordinates of Remark 1.5: we let u be the corresponding Dir-minimizer and $w = (u, \Psi(x,u))$. If $\bar{\eta}$ and $\hat{\varepsilon}$ are suitably small, we have

$$\int_{B_s(p)} \mathcal{G}(w, Q\,[\![\boldsymbol{\eta} \circ w]\!])^2 \geq \tfrac{3\hat{\delta}}{4} s^m E\,,$$

and $\sup\left\{\mathrm{Dir}(f), \mathrm{Dir}(w)\right\} \leq CE$ (here we use Remark 4.5). Thus there is $\bar p \in B_s(p)$ with $\mathcal{G}(w(\bar p), Q\,[\![\boldsymbol{\eta}\circ w(\bar p)]\!])^2 \geq \frac{3\hat\delta}{4\omega_m}\,E$ and, by (8.1), we conclude

$$g(x) := \mathcal{G}(w(x), Q\,[\![\boldsymbol{\eta}\circ w(x)]\!]) \geq \left(\tfrac{3\hat\delta}{4\omega_m}E\right)^{1/2} - 2\,(CE)^{1/2}\bar C \bar s^{\kappa} \geq \left(\tfrac{\hat\delta}{2}E\right)^{1/2}, \tag{8.3}$$

where we assume that $\bar s$ is chosen small enough in order to satisfy the last inequality. Setting $h(x) := \mathcal{G}(f(x), Q\,[\![\boldsymbol{\eta}\circ f(x)]\!])$, we recall that we have

$$\int_{B_s(p)} |h-g|^2 \leq C\,\bar\eta E\,.$$

Consider therefore the set $A := \left\{h > \left(\frac{\hat\delta}{4}E\right)^{1/2}\right\}$. If $\bar\eta$ is sufficiently small, we can assume that $|B_s(p)\setminus A| < \frac{1}{8}|B_s|$. Further, define $\bar A := A\cap K$, where K is the set of Theorem 1.4. Assuming $\hat\varepsilon$ is sufficiently small we ensure $|B_s(p)\setminus \bar A| < \frac{1}{4}|B_s|$. Let N be the smallest integer such that $N\frac{\hat\delta E}{64Qs} \geq \frac{s}{2}$. Set $\sigma_i := s - i\frac{\hat\delta E}{64Qs}$ for $i\in\{0,1\ldots,N\}$ and consider, for $i\leq N-1$, the annuli $\mathcal{C}_i := B_{\sigma_i}(p)\setminus B_{\sigma_{i+1}}(p)$. If $\hat\varepsilon$ is sufficiently small, we can assume that $N\geq 2$ and $\sigma_N \geq \frac{s}{4}$. For at least one of these annuli we must have $|\bar A\cap \mathcal{C}_i| \geq \frac{1}{2}|\mathcal{C}_i|$. We then let $\sigma := \sigma_i$ be the corresponding outer radius and we denote by $\mathcal{C}$ the corresponding annulus.

Consider now a point $x\in \mathcal{C}\cap\bar A$ and let T_x be the slice $\langle T, \mathbf{p}, x\rangle$. Since $\bar A\subset K$, for a.e. $x\in\bar A$ we have $T_x = \sum_{i=1}^Q [\![(x, f_i(x))]\!]$. Moreover, there exist i and j such that $|f_i(x) - f_j(x)|^2 \geq \frac{1}{Q}\mathcal{G}(f(x), [\![\boldsymbol{\eta}\circ f(x)]\!])^2 \geq \frac{\hat\delta}{4Q}E$ (recall that $x\in\bar A\subset A$). When $x\in\mathcal{C}$ and the points (x,y) and (x,z) belong both to $\mathbf{B}_\sigma((p,q))$, we must have

$$|y-z|^2 \leq 4\left(\sigma^2 - \left(\sigma - \tfrac{\hat\delta E}{64Qs}\right)^2\right) \leq \tfrac{\sigma\hat\delta E}{8Qs} \leq \tfrac{\hat\delta E}{8Q}\,.$$

Thus, for $x\in\bar A\cap\mathcal{C}$ at least one of the points $(x, f_i(x))$ is not contained in $\mathbf{B}_\sigma((p,q))$. We conclude therefore

$$\begin{aligned}\|T\|(\mathbf{C}_\sigma(p)\setminus \mathbf{B}_\sigma((p,q))) &\geq |\mathcal{C}\cap\bar A| \geq \frac{1}{2}|\mathcal{C}| = \frac{\omega_m}{2}\left(\sigma^m - \left(\sigma - \tfrac{\hat\delta E}{64Qs}\right)^m\right)\\ &\geq \frac{\omega_m}{2}\sigma^m\left(1 - \left(1 - \tfrac{\hat\delta E}{64Qs\sigma}\right)^m\right)\,.\end{aligned} \tag{8.4}$$

Recall that, for τ sufficiently small, $(1-\tau)^m \leq 1 - \frac{m\tau}{2}$. Since $\sigma\geq\frac{s}{4}$, if $\hat\varepsilon$ is chosen sufficiently small we can therefore conclude

$$\|T\|(\mathbf{C}_\sigma(p)\setminus\mathbf{B}_\sigma(p)) \geq \frac{\omega_m\sigma^m\hat\delta E}{256Qs\sigma} \geq \frac{\omega_m}{1024Q}\hat\delta E\sigma^{m-2} = c_0\hat\delta E\sigma^{m-2}\,. \tag{8.5}$$

Next, by Theorem 1.4 and Theorem 1.6,

$$\|T\|(\mathbf{C}_\sigma(p)) \leq Q\omega_m\sigma^m + CE^{1+\gamma_1} + \bar\eta E + \int_{B_\sigma(p)} \frac{|Dw|^2}{2}\,. \tag{8.6}$$

Moreover, as shown in [14, Section 3.3] (cf. [14, Proposition 3.10]), we have

$$\int_{B_\sigma(p)} |Dw|^2 \leq \|D\Psi\|^2\sigma^m + C\int_{B_\sigma(p)} |Du|^2 \leq C(1+C^\star)E\sigma^m + C\mathrm{Dir}(u)\sigma^{m-2+2\kappa}, \quad (8.7)$$

(for some constants κ and C depending only on m, n and Q; in fact the exponent κ is the one of (8.1)). Combining (8.5), (8.6) and (8.7), we conclude

$$\|T\|(\mathbf{B}_\sigma((p,q))) \leq Q\omega_m\sigma^m + (\bar\eta + C(1+C^\star)\sigma^m)E + CE^{1+\gamma_1} + CE\sigma^{m-2+2\kappa} - c_0\sigma^{m-2}\hat\delta E\,. \quad (8.8)$$

Next, by the monotonicity formula, $\rho \mapsto \exp(C\mathbf{A}^2\rho^2)\rho^{-m}\|T\|(\mathbf{B}_\rho((p,q)))$ is a monotone function (indeed, the usual monotonicity formula of the theory of varifolds with bounded mean curvature gives the monotonicity of $\rho \mapsto \exp(C\mathbf{A}\rho)\rho^{-m}\|T\|(\mathbf{B}_\rho((p,q)))$, cf. [29, Theorem 17.6]); the slight improvement needed in this proof follows from minor modifications of the usual argument but, since we have not been able to find a reference, we provide a proof in Lemma A.1 in the appendix). Using $\mathbf{A}^2 \leq C^\star E$, $\Theta(T,(p,q)) = Q$ and the Taylor expansion of the exponential, we conclude

$$\|T\|(\mathbf{B}_\sigma((p,q))) \geq Q\omega_m\sigma^m - CC^\star E\sigma^{m+2}\,. \quad (8.9)$$

Combining (8.8) and (8.9) we conclude

$$C(1+C^\star)\sigma^2 + (\bar\eta + CE_1^\gamma)\sigma^{2-m} + C\sigma^{2\kappa} \geq c_0\hat\delta\,. \quad (8.10)$$

Recalling that $\sigma \leq s < \bar s$, we can, finally, specify $\bar s$: it is chosen so that $C(1+C^\star)\bar s^2 + C\bar s^{2\kappa}$ is smaller than $\frac{c_0}{2}\hat\delta$. Combined with (8.3) this choice of $\bar s$ depends, therefore, only upon $\hat\delta$. (8.10) becomes then

$$(\bar\eta + CE^{\gamma_1})\sigma^{2-m} \geq \tfrac{c_0}{2}\hat\delta\,. \quad (8.11)$$

Next, recall that $\sigma \geq \frac{s}{4}$. We then choose $\hat\varepsilon$ so that $(\bar\eta + C\hat\varepsilon^{\gamma_1})(\frac{s}{4})^{2-m} \leq \frac{c_0}{4}\hat\delta$. This choice is incompatible with (8.11), thereby reaching a contradiction: for this choice of the parameter $\hat\varepsilon$ (which in fact depends only upon $\hat\delta$ and s) the conclusion of the Theorem, i.e. (1.9), must then be valid. □

Appendix A. Monotonicity formula

Lemma A.1. *There is a constant C depending only on m, n and $\bar n$ with the following property. If $\Sigma \subset \mathbb{R}^{m+n}$ is a C^2 $(m+\bar n)$-dimensional submanifold with $\|A_\Sigma\|_\infty \leq \mathbf{A}$ and T an m-dimensional integer-rectifiable current supported in Σ which is stationary in Σ, then for every $\xi \in \Sigma$ the function $\rho \mapsto \exp(C\mathbf{A}^2\rho^2)\rho^{-m}\|T\|(\mathbf{B}_\rho(\xi))$ is monotone on the interval $]0,\bar\rho[$, where $\bar\rho := \min\{\mathrm{dist}(x,\mathrm{spt}(\partial T)), (C\mathbf{A})^{-1}\}$.*

Proof. The argument is a minor variant of the classical proof of the monotonicity formula for varifolds with bounded mean curvature due to Allard (cf. [2]). Here the stronger hypothesis that T is stationary in a C^2-submanifold allows a better estimate of the relevant error term. Without loss of generality assume $\xi = 0$, let $s \in]0,\bar\rho[$ and $\varphi \in C^1_c(]-1,1[)$ with $\varphi \equiv 1$ in a neighborhood of 0. For each $x \in \Sigma$ let $\mathbf{p}_x : \mathbb{R}^{m+n} \to T_x\Sigma$ be the orthogonal projection onto the tangent space to Σ in x and consider the vector field $X_s(x) := \varphi(\frac{|x|}{s})\mathbf{p}_x(x)$. Note that X_s is tangent to Σ and thus $\delta T(X_s) = 0$. In order to compute $\delta T(X_s)$, consider at $\|T\|$-a.e. $x \in \mathrm{spt}(T)$ an orthonormal frame $e_1, \ldots, e_m$ with $e_1 \wedge \ldots \wedge e_m = \vec{T}$. It turns out that

$$\delta T(X_s) = \int \mathrm{div}_{\vec{T}} X_s \, d\|T\| = \int \sum_i \langle D_{e_i} X_s, e_i \rangle \, d\|T\| \, .$$

Next, at any $x \in \Sigma$ let $\nu_1, \ldots, \nu_l$ $(l = n - \bar{n})$ be an orthonormal frame orthogonal to Σ. Since $\mathbf{p}_x(x) = x - \sum_j \langle x, \nu_j \rangle \nu_j$ and $\langle e_i, \nu_j \rangle = 0$, we compute:

$$\mathrm{div}_{\vec{T}} X_s(x) = \underbrace{\sum_i \left[D_{e_i} \left(\varphi \left(\tfrac{|x|}{s} \right) \right) \langle x, e_i \rangle + \varphi \left(\tfrac{|x|}{s} \right) \langle D_{e_i} x, e_i \rangle \right]}_{I} - \underbrace{\varphi \left(\tfrac{|x|}{s} \right) \sum_{i,j} \langle x, \nu_j \rangle \langle D_{e_i} \nu_j, e_i \rangle}_{II} \, .$$

I is the usual expression appearing in the proof of the standard monotonicity formula for stationary varifolds. If we use the notation r for the function $x \mapsto |x|$ and $\nabla^\perp r$ for the orthogonal projection on the orthogonal complement of $\mathrm{Span}\{e_1, \ldots, e_m\}$, we find $I = m\varphi(\frac{r}{s}) + \frac{r}{s}\varphi'(\frac{r}{s})(1 - |\nabla^\perp r|^2)$ (see for instance [11, (2.2)]). In order to bound II, we first observe that $\langle D_{e_i} \nu_j, e_i \rangle = -\langle A(e_i, e_i), \nu_j \rangle$. Next, since $r \leq (C\mathbf{A})^{-1}$, if C is chosen sufficiently large we can assume that the geodesic segment of Σ connecting 0 and x has length $\ell < 2r$. Denote by $\gamma : [0, \ell] \to \Sigma$ a parametrization by arc-length of such a segment. Then,

$$\langle x, \nu_j(x) \rangle = \int_0^\ell \langle \dot{\gamma}(\sigma), \nu_j(\gamma(\ell)) \rangle \, d\sigma = \int_0^\ell \underbrace{\langle \dot{\gamma}(\sigma), [\nu_j(\gamma(\ell)) - \nu_j(\gamma(\sigma))] \rangle}_{g(\sigma)} \, d\sigma \, , \tag{A.1}$$

and observe that

$$|g'(\sigma)| \leq 2 \left| \frac{d}{d\sigma} \dot{\gamma}(\sigma) \right| + \left| \langle \dot{\gamma}(\sigma), \frac{d}{d\sigma} \nu_j(\gamma(\sigma)) \rangle \right| \leq 3 |A(\dot{\gamma}(\sigma), \dot{\gamma}(\sigma))|.$$

Since $g(\ell) = 0$, integrating the latter inequality we conclude $|g(\sigma)| \leq 3\ell\mathbf{A} \leq 6r\mathbf{A}$, which in turn, together with (A.1), gives $|x \cdot \nu_j(x)| \leq 12 r^2 \mathbf{A}$.

Putting all estimates together, we achieve the inequality $|II| \leq C\varphi(\frac{r}{s}) r^2 \mathbf{A}^2$. From here on we can follow the usual strategy leading to the monotonicity formula (cf. [29] or [11, Proof of Theorem 2.1]): letting the test function φ converge from below to the indicator function of $]-1, 1[$, after few manipulations we achieve the inequality

$$\frac{d}{ds} \frac{\|T\|(\mathbf{B}_s)}{s^m} \geq -C\mathbf{A}^2 s \frac{\|T\|(\mathbf{B}_s)}{s^m} \, ,$$

which leads to the desired claim. □

Remark A.2. The proof can be easily extended to varifolds which are stationary in Σ. In fact the argument above can be considerably shortened using directly the Monotonicity Formula of Section 5 in [2].

References

[1] Robert A. Adams. *Sobolev spaces*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

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

[3] Frederick J. Almgren, Jr. *Almgren's big regularity paper*, volume 1 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[4] Luigi Ambrosio. Metric space valued functions of bounded variation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 17(3):439–478, 1990.

[5] Luigi Ambrosio and Bernd Kirchheim. Currents in metric spaces. *Acta Math.*, 185(1):1–80, 2000.

[6] Costante Bellettini and Tristan Rivière. The regularity of special Legendrian integral cycles. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 11(1):61–142, 2012.

[7] Garrett Birkhoff. Three observations on linear algebra. *Univ. Nac. Tucumán. Revista A.*, 5:147–151, 1946.

[8] Sheldon Xu-Dong Chang. Two-dimensional area minimizing integral currents are classical minimal surfaces. *J. Amer. Math. Soc.*, 1(4):699–778, 1988.

[9] Ennio De Giorgi. *Frontiere orientate di misura minima*. Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61. Editrice Tecnico Scientifica, Pisa, 1961.

[10] C. De Lellis. Almgren's Q-valued functions revisited. *Proceedings of the ICM 2010, Hyderabad.*

[11] Camillo De Lellis. Allard's interior regularity theorem: an invitation to stationary varifolds. *Preprint*, 2012.

[12] Camillo De Lellis and Emanuele Spadaro. *The regularity theorey for area-minimizing currents in codimension higher than* 1. In preparation.

[13] Camillo De Lellis and Emanuele Spadaro. Center manifold: a case study. *Discrete Contin. Dyn. Syst.*, 31(4):1249–1272, 2011.

[14] Camillo De Lellis and Emanuele Spadaro. Q-valued functions revisited. *Mem. Amer. Math. Soc.*, 211(991):vi+79, 2011.

[15] Camillo De Lellis and Emanuele Spadaro. Regularity of area-minimizing currents II: center manifold. *Preprint*, 2013.

[16] Camillo De Lellis and Emanuele Spadaro. Regularity of area-minimizing currents III: blow-up. *Preprint*, 2013.

[17] Camillo De Lellis and Emanuele Spadaro. Multiple valued functions and integral currents. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, to appear (2014).

[18] Camillo De Lellis, Emanuele Spadaro, and Luca Spolaor. *Forthcoming*, 2013.

[19] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[20] Herbert Federer. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.

[21] M. Giaquinta and G. Modica. Regularity results for some classes of higher order nonlinear elliptic systems. *J. Reine Angew. Math.*, 311/312:145–169, 1979.

[22] Enrico Giusti. *Direct methods in the calculus of variations*. World Scientific Publishing Co., Inc., River Edge, NJ, 2003.

[23] R. L. Jerrard and H. M. Soner. Functions of bounded higher variation. *Indiana Univ. Math. J.*, 51(3):645–677, 2002.

[24] David Pumberger and Tristan Rivière. Uniqueness of tangent cones for semicalibrated integral 2-cycles. *Duke Math. J.*, 152(3):441–480, 2010.

[25] Tristan Rivière. Approximating J-holomorphic curves by holomorphic ones. *Calc. Var. Partial Differential Equations*, 21(3):273–285, 2004.

[26] Tristan Rivière. A lower-epiperimetric inequality for area-minimizing surfaces. *Comm. Pure Appl. Math.*, 57(12):1673–1685, 2004.

[27] Tristan Rivière and Gang Tian. The singular set of J-holomorphic maps into projective algebraic varieties. *J. Reine Angew. Math.*, 570:47–87, 2004.

[28] Tristan Rivière and Gang Tian. The singular set of 1-1 integral currents. *Ann. of Math. (2)*, 169(3):741–794, 2009.

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

[30] Emanuele Spadaro. Complex varieties and higher integrability of Dir-minimizing Q-valued functions. *Manuscripta Math.*, 132(3-4):415–429, 2010.

[31] Emanuele Spadaro. Nondoubling A^∞ weights. *Adv. Calc. Var.*, 5(3):345–354, 2012.

[32] Clifford Henry Taubes. *Seiberg Witten and Gromov invariants for symplectic 4-manifolds*, volume 2 of *First International Press Lecture Series*. International Press, Somerville, MA, 2000. Edited by Richard Wentworth.

[33] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[34] Brian White. Stratification of minimal surfaces, mean curvature flows, and harmonic maps. *J. Reine Angew. Math.*, 488:1–35, 1997.