

# RECTIFIABILITY OF THE SINGULAR SET OF MULTIPLE VALUED ENERGY MINIMIZING HARMONIC MAPS

JONAS HIRSCH, SALVATORE STUWARD, AND DANIELE VALTORTA

ABSTRACT. In this paper we study the singular set of energy minimizing  $Q$ -valued maps from  $\mathbb{R}^m$  into a smooth compact manifold  $\mathcal{N}$  without boundary. Similarly to what happens in the case of single valued minimizing harmonic maps, we show that this set is always  $(m - 3)$ -rectifiable with uniform Minkowski bounds. Moreover, as opposed to the single valued case, we prove that the target  $\mathcal{N}$  being non-positively curved but not simply connected does not imply continuity of the map.

KEYWORDS:  $Q$ -valued functions, harmonic maps, singular set, rectifiability, Reifenberg theorem, quantitative stratification.

AMS SUBJECT CLASSIFICATION (2010): 49Q20, 58E20.

## 0. INTRODUCTION

Multiple-valued harmonic functions (Dir-minimizers) were originally introduced by Almgren in [Alm00] as first order approximations for the branching singularities of minimal surfaces in codimension higher than one. Roughly speaking, a  $Q$ -valued Dirichlet minimizer is a function which attains  $Q$  different values (counted with multiplicity) for each point in the domain, and minimizes a suitably defined Dirichlet energy with respect to boundary data.

Even though at first sight it might seem that  $Q$ -valued functions are easy generalizations of classical (single valued) functions, there are some crucial differences. For instance, the space of such functions is not linear, in the sense that the sum of two  $Q$ -valued functions is not a well-defined notion. These differences make the study of such objects both more complicated than their classical counterpart and more interesting. For a recent survey on results of this kind, we direct the reader to [DLS11], where the authors revisit Almgren's original regularity theory of Dir-minimizing  $Q$ -valued functions suggesting a more intrinsic approach which has its roots in the new techniques developed in the last two decades to perform analysis on metric spaces.

Several generalizations of the original  $Q$ -valued Dir-minimizing functions have been studied in literature, both in the direction of analyzing multiple-valued functions taking values in more general target spaces than  $\mathbb{R}^n$ , and in the sense of more general functionals to minimize. Here we limit ourselves to mentioning some of these works. In the direction of functionals more general than the Dirichlet energy there are the works [Mat83, DLFS11], as well as the recent work [Stu17a] by one of the authors, where a complete multi-valued theory for the stability operator is studied. The papers [DLGT04, Gob09, BDPW15, BDPG15, Stu17b] focus instead on more general target spaces.

The work [Hir16b] of one of the authors started analyzing  $Q$ -valued harmonic maps into compact Riemannian manifolds, introducing the appropriate definitions and developing the basic continuity theory for such objects. In particular, using a suitably modified version of Federer-Almgren's dimension-reduction argument, [Hir16b] proves that

**Theorem 0.1** ([Hir16b, Theorem 0.1]). *Given a smooth compact Riemannian manifold  $\mathcal{N}$  and a  $Q$ -valued map  $u : \Omega \subset \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  locally minimizing the Dirichlet energy, the singular set of  $u$*

defined as

$$\text{sing}_H(u) := \{x \in \Omega \text{ s.t. } u \text{ is not continuous in a neighborhood of } x\} \quad (0.1)$$

is a closed set of Hausdorff dimension at most  $m - 3$ . Moreover, outside this set the map  $u$  is locally  $C^{0,\alpha}$  continuous, with  $\alpha = \alpha(m, Q) > 0$ .

Note that this is the multi-valued counterpart of the classical Schoen-Uhlenbeck results in [SU82] for the singularities of (single-valued) harmonic maps between manifolds. Also observe that if the target manifold  $\mathcal{N}$  is replaced by Euclidean space  $\mathbb{R}^n$  then it actually holds  $\text{sing}_H(u) = \emptyset$  (cf. [DLS11, Theorem 0.9]). In fact, in the classical regularity theory for  $\mathcal{A}_Q(\mathbb{R}^n)$ -valued energy minimizing maps, the singular set  $\text{sing}(u)$  consists of those points  $x \in \Omega$  having the following property: there is no neighborhood  $\mathcal{U}$  of  $x$  where the map  $u$  admits a decomposition  $u = \sum_{\ell=1}^Q \llbracket u_\ell \rrbracket$  in classical smooth harmonic functions  $u_\ell: \mathcal{U} \rightarrow \mathbb{R}^n$  such that for any  $\ell, \ell' \in \{1, \dots, Q\}$  either  $u_\ell \equiv u_{\ell'}$  in  $\mathcal{U}$  or  $u_\ell(y) \neq u_{\ell'}(y)$  for every  $y \in \mathcal{U}$ . For such a singular set, it holds that  $\dim_{\mathcal{H}}(\text{sing}(u)) \leq m - 2$  (cf. [Alm00] and [DLS11, Theorem 0.11]). The same definition of  $\text{sing}(u)$  can be naturally extended to the case when  $u$  is  $\mathcal{A}_Q(\mathcal{N})$ -valued, by simply replacing the decomposition in harmonic functions with a decomposition in  $\mathcal{N}$ -valued smooth harmonic maps. With this definition, it evidently holds that  $\text{sing}_H(u) \subset \text{sing}(u)$ . Furthermore, the two singular sets coincide if  $Q = 1$ , as in this case Hölder continuity can be upgraded to full smoothness by standard elliptic regularity theory. In contrast, no information is available at the moment on  $\text{sing}(u)$  for  $\mathcal{A}_Q(\mathcal{N})$ -valued energy minimizing maps when  $Q > 1$ .

The goal of this work is, instead, to improve the result in Theorem 0.1, and prove  $(m-3)$ -rectifiability for the singular set  $\text{sing}_H(u)$  along with uniform  $(m-3)$  Minkowski bounds. In particular, we want to show that

**Theorem 0.2.** *Given an energy-minimizing  $Q$ -valued map  $u: B_2(0) \subseteq \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  with energy bounded by  $\Lambda$ , if  $B_r(\text{sing}_H(u)) := \bigcup_{x \in \text{sing}_H(u)} B_r(x)$  then we have*

$$\text{Vol}(B_r(\text{sing}_H(u)) \cap B_1(0)) \leq Cr^3, \quad (0.2)$$

where  $C = C(m, \mathcal{N}, Q, \Lambda)$ . Moreover,  $\text{sing}_H(u)$  is  $(m-3)$ -rectifiable.

In order to prove this result, we are going to apply the techniques developed in [NV17], which roughly speaking rely on a quantitative version of the dimension-reduction argument. However, here we will present an alternative definition of the quantitative stratification used in [NV17] and originally introduced by Cheeger and Naber in [CN13a, CN13b].

More precisely, the quantitative stratification is based on the analysis of symmetries and approximate symmetries of the map  $u$  at different points and scales, and roughly speaking the quantitative stratum  $\mathcal{S}_{\varepsilon,r}^k$  is the set of points  $x$  for which  $u$  on  $B_r(x)$  is  $\varepsilon$ -far away from being homogeneous and invariant with respect to a  $k$ -dimensional subspace. While the notion of closeness employed by [CN13a, CN13b] relies on the  $L^2$  distance of the map  $u$  from some homogeneous and  $k$ -symmetric model map  $h$ , we propose a notion that focuses on the  $L^2$  norm of the *gradient* of  $u$  restricted to arbitrary  $k$ -subspaces. The advantage of the new definition is that it is more tailored towards the available  $\varepsilon$ -regularity theorems. This stratification is introduced in detail in section 3.

In section 8, we also consider the special case of non-positive sectional curvatures in the target  $\mathcal{N}$ . For classical harmonic maps, this assumption implies full-blown continuity of the map  $u$  everywhere. On the other hand, in the case of  $Q$ -valued maps this is true only if  $\mathcal{N}$  is assumed to be also *simply connected*. We will provide a counterexample to show that this assumption is needed. This example is

based on the fact that the graph of a  $Q$ -valued map can have a different topology from the one of its domain, and it shows once more that the properties of  $Q$ -valued maps can be very different from their single valued counterparts.

The plan of the paper is the following: first, we introduce  $Q$ -valued energy minimizers and quickly review the standard properties of these maps. In particular, we study the normalized energy  $\mathcal{E}(x, B_r(x)) := r^{2-m} \int_{B_r(x)} |Du|^2$  and its mollified version, which, although morally similar to the classical one, will prove itself to be more useful in quantitative estimates.

We then move on to the study of different versions of the  $\varepsilon$ -regularity theorem for  $Q$ -valued maps. Soon after, we prove the main estimates on the singular set of  $u$  and its stratification. This result relies on a sharp version of Reifenberg's theorem, which we quote from literature. Finally, we close the paper with the analysis of the case of non-positively curved target manifolds.

#### ACKNOWLEDGMENTS

The authors wish to thank the anonymous referee for carefully reading an earlier version of this manuscript, and for his/her valuable comments and suggestions. The research of J. H. has been supported by the MIUR SIR-grant "Geometric Variational Problems", ID RBSI14RVEZ. S. S. was supported by the ERC-grant RAM "Regularity of Area Minimizing currents", ID 306246. D. V. has been supported by the SNSF grant PZ00P2\_168006.

#### 1. NOTATION AND PRELIMINARIES

Throughout the whole paper, we will denote by  $\Omega$  an open subset of Euclidean space  $\mathbb{R}^m$ ,  $m \geq 2$ , and by  $\mathcal{N}$  a smooth compact Riemannian manifold of dimension  $n$  with empty boundary. Without loss of generality, we regard  $\mathcal{N}$  as an isometrically embedded submanifold of a Euclidean space  $\mathbb{R}^N$ . The symbol  $A$  will denote the second fundamental form of the embedding  $\mathcal{N} \hookrightarrow \mathbb{R}^N$ .

The Euclidean scalar product in  $\mathbb{R}^N$  is denoted by  $\langle \cdot, \cdot \rangle$ . Since the metric on  $\mathcal{N}$  is induced by the flat metric on  $\mathbb{R}^N$ , the same symbol will also be adopted for the scalar product between tangent vectors to  $\mathcal{N}$ . The standard connection in  $\mathbb{R}^m$  is denoted by  $D$ . If  $\{e_i\}_{i=1}^m$  is an orthonormal basis of  $\mathbb{R}^m$ , we will denote by  $D_i$  the directional derivative operator  $D_{e_i}$ .

The open ball with center  $x$  and radius  $r$  in  $\mathbb{R}^m$  is denoted  $B_r(x)$ . If  $1 \leq k \leq m-1$  and  $L \subset \mathbb{R}^m$  is a linear subspace of dimension  $\dim(L) = k$ , then we will denote the disc  $(x + L) \cap B_r(x)$  by  $B_r^L(x)$ , or often with the simpler notation  $B_r^k(x)$ .

**1.1. Multiple valued functions.** Fix an integer  $Q \geq 1$ . We will assume that the reader is familiar with the theory of Almgren's  $Q$ -valued functions, for which we refer to [DLS11]. In what follows, we briefly recall the main definitions and properties we are going to need in the sequel. The space of  $Q$ -points in  $\mathbb{R}^N$  is denoted  $\mathcal{A}_Q(\mathbb{R}^N)$ , and defined by

$$\mathcal{A}_Q(\mathbb{R}^N) := \left\{ T = \sum_{\ell=1}^Q \llbracket p_\ell \rrbracket : \text{each } p_\ell \in \mathbb{R}^N \right\},$$

where  $\llbracket p_\ell \rrbracket$  is the Dirac delta measure centered at  $p_\ell$ . Observe that, by definition, a  $Q$ -point  $T$  is a purely atomic measure of mass  $Q$  in  $\mathbb{R}^N$  which is obtained as the linear combination of Dirac deltas with integer multiplicities. If  $T \in \mathcal{A}_Q(\mathbb{R}^N)$ , the symbol  $\text{spt}(T)$  will denote the *support* of the aforementioned measure. We endow  $\mathcal{A}_Q(\mathbb{R}^N)$  with the structure of complete metric space determined

by the distance  $\mathcal{G}(T_1, T_2)$  given by

$$\mathcal{G}(T_1, T_2)^2 := \min_{\sigma \in \mathcal{P}_Q} \sum_{\ell=1}^Q |p_\ell - q_{\sigma(\ell)}|^2,$$

where  $\mathcal{P}_Q$  denotes the group of permutations of  $\{1, \dots, Q\}$ , and  $T_1 = \sum_{\ell} \llbracket p_\ell \rrbracket$ ,  $T_2 = \sum_{\ell} \llbracket q_\ell \rrbracket$ .

Any map  $f: \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$  will be called a  $Q$ -valued function. It is a simple observation (cf. [DLS11, Proposition 0.4]) that if  $f$  is a measurable  $Q$ -valued function then there are measurable maps  $f_\ell: \Omega \rightarrow \mathbb{R}^N$  for  $\ell = 1, \dots, Q$  such that  $f(x) = \sum_{\ell=1}^Q \llbracket f_\ell(x) \rrbracket$  at a.e.  $x \in \Omega$ . Any choice of  $\{f_\ell\}_{\ell=1}^Q$  as above is called a measurable *selection* for  $f$ .

For  $p \in [1, \infty]$ , the spaces  $L^p(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  consist of those measurable  $f: \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$  for which

$$\begin{aligned} \|f\|_{L^p}^p &:= \int_{\Omega} \mathcal{G}(f(x), Q\llbracket 0 \rrbracket)^p dx < \infty \quad \text{when } 1 \leq p < \infty, \\ \|f\|_{L^\infty} &:= \operatorname{ess\,sup}_{x \in \Omega} \mathcal{G}(f(x), Q\llbracket 0 \rrbracket) < \infty. \end{aligned}$$

For the sake of notational simplicity, we will often set  $|T| := \mathcal{G}(T, Q\llbracket 0 \rrbracket)$  if  $T \in \mathcal{A}_Q(\mathbb{R}^N)$ . We remark that if  $Q > 1$  then  $\mathcal{A}_Q(\mathbb{R}^N)$  is not a linear space: hence, in spite of the notation,  $T \mapsto |T|$  is not a norm.

A map  $f: \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$  belongs to the Sobolev space  $W^{1,p}(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  if there exists  $\psi \in L^p(\Omega)$  such that for every Lipschitz function  $\phi: \mathcal{A}_Q(\mathbb{R}^N) \rightarrow \mathbb{R}$  it holds:

- (i)  $\phi \circ f \in W^{1,p}(\Omega)$ ;
- (ii)  $|D(\phi \circ f)(x)| \leq \operatorname{Lip}(\phi)\psi(x)$  at a.e.  $x \in \Omega$ .

By [DLS11, Proposition 4.2], if  $f \in W^{1,p}(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  then for every  $i \in \{1, \dots, m\}$  there exists a unique  $g_i \in L^p(\Omega)$  such that

- (i)  $|D_i(\mathcal{G}(f, T))| \leq g_i$  a.e. for every  $T \in \mathcal{A}_Q(\mathbb{R}^N)$ ;
- (ii) if  $h_i \in L^p(\Omega)$  is such that  $|D_i(\mathcal{G}(f, T))| \leq h_i$  a.e. for every  $T \in \mathcal{A}_Q(\mathbb{R}^N)$  then  $g_i \leq h_i$  a.e.

We will call the function  $g_i$  the *metric derivative* of  $f$  in the direction  $e_i$ .

As usual,  $W_{loc}^{1,p}(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  consists of those measurable functions which are in  $W^{1,p}(\Omega', \mathcal{A}_Q(\mathbb{R}^N))$  for every  $\Omega' \Subset \Omega$ .

A  $Q$ -valued function  $f: \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$  is *differentiable* at a point  $x \in \Omega$  if there exist  $Q$  linear maps  $\lambda_\ell: \mathbb{R}^m \rightarrow \mathbb{R}^N$  satisfying

- (i)  $\mathcal{G}(f(y), \sum_{\ell} \llbracket f_\ell(x) + \lambda_\ell \cdot (y - x) \rrbracket) = o(|y - x|)$  for  $|y - x| \rightarrow 0$ ;
- (ii)  $\lambda_\ell = \lambda_{\ell'}$  if  $f_\ell(x) = f_{\ell'}(x)$ .

If  $f$  is differentiable at  $x$ , then the  $Q$ -point  $\sum_{\ell=1}^Q \llbracket \lambda_\ell \rrbracket \in \mathcal{A}_Q(\mathbb{R}^{N \times m})$  is the *differential* of  $f$  at  $x$ , and will be denoted  $Df(x)$  or  $Df|_x$ . We will write  $Df_\ell(x)$  for the map  $\lambda_\ell$ , so that  $Df(x) = \sum_{\ell} \llbracket Df_\ell(x) \rrbracket$ , and we establish the notation  $D_\tau f(x) := \sum_{\ell} \llbracket Df_\ell(x) \cdot \tau \rrbracket = \sum_{\ell} \llbracket D_\tau f_\ell(x) \rrbracket \in \mathcal{A}_Q(\mathbb{R}^N)$  for the directional derivative in the direction  $\tau \in \mathbb{R}^m$ . We will also sometimes write  $Df(x) \cdot \tau$  for  $D_\tau f(x)$ , so that  $Df(x) \cdot \tau = \sum_{\ell} \llbracket Df_\ell(x) \cdot \tau \rrbracket$ .

It is a consequence of the Lipschitz approximation theorem for Sobolev  $Q$ -valued functions [DLS11, Proposition 2.5] and of the  $Q$ -valued counterpart of Rademacher's theorem [DLS11, Theorem 1.13] that every Sobolev  $Q$ -valued map is approximately differentiable at a.e.  $x \in \Omega$ . Furthermore, as shown in [DLS11, Proposition 2.17], if  $f \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  then for every  $i \in \{1, \dots, m\}$  it holds

$$g_i^2 = \mathcal{G}(D_i f, Q\llbracket 0 \rrbracket)^2 = |D_i f|^2 \quad \text{a.e. in } \Omega.$$

This makes unambiguous the use of the notation  $|D_i f|$  for the metric derivative  $g_i$ . In particular, for  $f \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^N))$  there is a well-defined notion of (rescaled) *Dirichlet energy* in a ball  $B_r(x) \Subset \Omega$ , given by

$$\mathcal{E}(f, B_r(x)) := r^{2-m} \int_{B_r(x)} |Df(y)|^2 dy = r^{2-m} \int_{B_r(x)} \sum_{i=1}^m |D_i f(y)|^2 dy.$$

**1.2.  $Q$ -valued energy minimizing maps.** Now, set

$$W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N})) := \left\{ u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathbb{R}^N)) : \text{spt}(u(x)) \subset \mathcal{N} \text{ for a.e. } x \in \Omega \right\}. \quad (1.1)$$

**Definition 1.1** (Energy minimizers, cf. [Hir16b, Definition 1.1]). A map  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  is a *local minimizer*, or simply *minimizer*, of the Dirichlet energy if for any  $B_r(x) \Subset \Omega$  the following holds

$$\mathcal{E}(u, B_r(x)) \leq \mathcal{E}(v, B_r(x)) \quad (1.2)$$

for every  $v \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  such that  $v \equiv u$  in a neighborhood of  $\partial B_r(x)$ .

The  $Q$ -valued energy minimizers defined in Definition 1.1 are the multi-valued counterpart of classical energy minimizing harmonic maps. We refer the reader to the beautiful monographs of Simon [Sim96], Moser [Mos05] or Lin-Wang [LW08] for more about classical (single-valued) energy minimizing maps.

As anticipated in the introduction, a partial regularity theory for  $Q$ -valued energy minimizers was developed by the first author in [Hir16b]. For further reference, and for the readers' convenience, let us briefly collect the main results of [Hir16b] which we are going to use in the sequel.

A first important observation is that if  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  is energy minimizing and if  $B_r(x) \Subset \Omega$  then one can test the minimality of  $u$  along suitably chosen families  $u_\varepsilon$  of competitors in order to infer that  $u$  satisfies some integral equations, known as *variational equations*, which turn out to be of fundamental importance for the regularity theory. There are two important kinds of variations that one may consider in this context: the *inner variations* (obtained by perturbing  $u$  by means of right compositions with diffeomorphisms in the *domain*) and the *outer variations* (obtained by perturbing  $u$  by means of left compositions with diffeomorphisms in the *target*).

**Proposition 1.2** (Variational equations, cf. [Hir16b, Equations (2.2) and (2.5)]). *Fix  $B_r(x) \Subset \Omega$  and let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be energy minimizing. Then, for every vector field  $X = (X^1, \dots, X^m) \in C_c^1(B_r(x), \mathbb{R}^m)$  the following inner variation formula holds:*

$$\int_{B_r(x)} \sum_{i,j=1}^m \left( |Du|^2 \delta_{ij} - 2 \sum_{\ell=1}^Q \langle D_i u_\ell, D_j u_\ell \rangle \right) D_i X^j dy = 0. \quad (1.3)$$

*Moreover, for any vector field  $Y \in C^1(B_r(x) \times \mathbb{R}^N, \mathbb{R}^N)$  such that  $Y(y, p) = 0$  for  $y$  in a neighborhood of  $\partial B_r(x)$  we have the following outer variation formula:*

$$\int_{B_r(x)} \sum_{i=1}^m \sum_{\ell=1}^Q (\langle D_i u_\ell, D_i(Y(y, u_\ell)) \rangle + \langle A_{u_\ell}(D_i u_\ell, D_i u_\ell), Y(y, u_\ell) \rangle) dy = 0. \quad (1.4)$$

Recall that in the classical case  $Q = 1$  a map  $u \in W_{loc}^{1,2}(\Omega, \mathcal{N})$  satisfying the identity (1.4) for any  $Y$  is referred to as a *weakly harmonic map*, whereas a map  $u$  for which both (1.3) and (1.4) hold for any choice of  $X$  and  $Y$  is called a *stationary harmonic map*. Analogously, we will call *stationary  $Q$ -harmonic* any map  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  for which both equations (1.3) and (1.4) hold. Of course, by Proposition 1.2 every  $Q$ -valued energy minimizing map  $u$  is stationary  $Q$ -harmonic. On the other

hand, some of the results that we present here hold true under the weaker assumption that  $u$  is stationary  $Q$ -harmonic rather than minimizing, since their proofs are a consequence solely of the variational equations. We will explicitly underline in our statements every time that the result applies also to stationary  $Q$ -harmonic maps.

A first powerful result stemming from the variational equations is the monotonicity of the map  $r \in (0, \text{dist}(x, \partial\Omega)) \mapsto \mathcal{E}(u, B_r(x))$  for every fixed point  $x \in \Omega$ , cf. [Hir16b, Equation (2.6)]. As a consequence, if  $u$  is stationary  $Q$ -harmonic then for every  $x \in \Omega$  the *density*  $\Theta_u(x)$  of  $u$  at  $x$  is well-defined by the formula

$$\Theta_u(x) := \lim_{r \downarrow 0} \mathcal{E}(u, B_r(x)).$$

Multiple-valued energy minimizers also enjoy the following compactness theorem.

**Theorem 1.3** (Compactness, cf. [Hir16b, Lemma 4.1]). *Let  $\{u_h\}_{h=1}^\infty \subset W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be a sequence of  $Q$ -valued minimizing harmonic maps with  $\sup_{h \geq 1} \mathcal{E}(u_h, B_r(x)) < \infty$  for each ball  $B_r(x) \Subset \Omega$ . Then, there is a subsequence  $u_{h_j}$  and a minimizing harmonic map  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  such that*

- (i)  $\lim_{j \rightarrow \infty} \int_\Omega \mathcal{G}(u_{h_j}, u)^2 dy = 0$ ;
- (ii)  $\lim_{j \rightarrow \infty} \mathcal{E}(u_{h_j}, B_r(x)) = \mathcal{E}(u, B_r(x))$  for every ball  $B_r(x) \Subset \Omega$ .

The monotonicity of the rescaled energy at a fixed point  $x_0 \in \Omega$  together with the Compactness Theorem 1.3 allow to conclude the existence of tangent maps. In particular, for every sequence  $r_h$  of radii with  $r_h \downarrow 0$  there exists a subsequence  $r_{h'}$  such that the maps  $T_{x_0, r_{h'}}^u(y) := u(x_0 + r_{h'}y)$  converge in  $L^2$  and locally in energy to a  $Q$ -valued energy minimizing map  $\phi \in W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$ . Any map  $\phi$  arising as a limit of a sequence  $T_{x_0, r_h}^u$  for some sequence  $r_h \downarrow 0$  is called a *tangent map* to  $u$  at  $x_0$  (cf. [Hir16b, Definition 6.1]). Every tangent map  $\phi$  is homogeneous of degree zero with respect to  $0 \in \mathbb{R}^m$ , and thus it satisfies  $\mathcal{E}(\phi, B_\rho(0)) \equiv \Theta_\phi(0) = \Theta_u(x_0)$  for every  $\rho > 0$ . Furthermore, the map  $y \in \mathbb{R}^m \mapsto \Theta_\phi(y)$  attains its maximum at  $y = 0$ . The set of points  $y \in \mathbb{R}^m$  for which  $\Theta_\phi(y) = \Theta_\phi(0)$  is classically called the *spine* of  $\phi$ , and it is denoted  $S(\phi)$ . It turns out that, exactly as in the classical case,  $S(\phi)$  is a linear subspace of  $\mathbb{R}^m$ , and  $\phi$  is invariant with respect to compositions with translations by elements in  $S(\phi)$ , that is  $\phi(x + y) = \phi(x)$  for every  $x \in \mathbb{R}^m$ , for any  $y \in S(\phi)$ . The dimension of  $S(\phi)$  is the number of independent directions along which  $\phi$  is invariant. Now, if  $u$  is Hölder continuous in a neighborhood of  $x_0$  then it is easy to see that  $\Theta_u(x_0) = 0$ , and thus  $u$  admits in  $x_0$  a tangent map  $\phi$  for which  $S(\phi) = \mathbb{R}^m$ , and  $\phi$  is constant. The set  $\text{sing}_H(u)$  of points  $x$  such that  $u$  is not Hölder continuous in a neighborhood can be instead classically stratified according to the number of symmetries that the tangent maps at points in it have. In particular, for  $0 \leq k \leq m - 1$  one defines

$$\mathcal{S}^k(u) := \{x \in \text{sing}_H(u) : \dim S(\phi) \leq k \text{ for every tangent map } \phi \text{ to } u \text{ at } x\}. \quad (1.5)$$

The following  $\varepsilon$ -regularity theorem in the spirit of Schoen-Uhlenbeck [SU82] is the core of [Hir16b] and the key to completing the partial regularity theory.

**Theorem 1.4** ( $Q$ -valued  $\varepsilon$ -regularity, cf. [Hir16b, Lemma 5.2]). *There exist constants  $\varepsilon_0 > 0$ ,  $\alpha > 0$  and  $C > 1$  depending on  $m, \mathcal{N}, Q$  with the property that if  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  is energy minimizing in  $B_{R_0}(x_0)$  with*

$$\mathcal{E}(u, B_{R_0}(x_0)) \leq \varepsilon_0,$$

*then the following energy decay estimate holds:*

$$\mathcal{E}(u, B_r(x)) \leq C \left(\frac{r}{R}\right)^{2\alpha} \mathcal{E}(u, B_R(x)) \quad \forall x \in B_{\frac{R_0}{2}}(x_0), \forall 0 < r \leq R \leq \frac{R_0}{2}.$$

In particular,  $u \in C^{0,\alpha}(B_{\frac{R_0}{2}}(x_0), \mathcal{A}_Q(\mathcal{N}))$ .

By the  $\varepsilon$ -regularity theorem, the set  $\text{sing}_H(u)$  coincides with the set  $\{x : \Theta_u(x) > 0\}$ . Using this information, it is standard to conclude that if  $u$  is minimizing then  $\mathcal{H}^{m-2}(\text{sing}_H(u)) = 0$ . On the other hand, if  $x_0 \in \text{sing}_H(u)$  then for any tangent map  $\phi$  one has  $\Theta_\phi(0) = \Theta_u(x_0) > 0$ , and thus the spine  $S(\phi)$  is a subset of  $\text{sing}_H(\phi)$ . Since  $S(\phi)$  is a linear subspace, and since  $\mathcal{H}^{m-2}(\text{sing}_H(\phi)) = 0$ , we have that  $\dim S(\phi) \leq m - 3$ . Hence,  $\mathcal{S}^{m-1}(u) = \mathcal{S}^{m-2}(u) = \mathcal{S}^{m-3}(u) = \text{sing}_H(u)$ . By a variation of the standard Federer-Almgren dimension-reduction argument, one then concludes that  $\dim_{\mathcal{H}}(\text{sing}_H(u)) = \dim_{\mathcal{H}}(\mathcal{S}^{m-3}(u)) \leq m - 3$ .

In the next sections we will turn our attention to the *quantitative stratification* for  $Q$ -valued minimizing maps, which will allow us to obtain better information on the fine properties of the singular strata  $\mathcal{S}^k(u)$ . Before doing that, we will slightly modify the definition of the rescaled energy: the new ‘‘mollified’’ energy that we are going to introduce in the coming section, will be more useful for quantitative estimates.

## 2. THE MOLLIFIED DIRICHLET ENERGY AND ITS MONOTONICITY

**Definition 2.1** (Mollified energy). Let  $\varphi = \varphi(t)$  be any non-negative function in  $C_c^1([0, 1])$  which is constant in a neighborhood of  $t = 0$ .

Then, for any  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  and for any  $B_r(x) \subset \Omega$  we define the quantity

$$\theta_u(x, r) := r^{2-m} \int \varphi\left(\frac{|x-y|}{r}\right) |Du(y)|^2 dy. \quad (2.1)$$

When the map  $u$  is fixed, we will simply write  $\theta(x, r)$  for the sake of notational simplicity. In what follows, we show that, under suitable assumptions on  $\varphi$ , the function  $r \mapsto \theta(x, r)$  is monotone non-decreasing for fixed  $x$ , and we explicitly compute its derivative.

**Notation 2.2.** For any  $x \in \mathbb{R}^m$ , we shall denote by  $r_x$  the radial unit vector field with respect to  $x$ , defined by

$$r_x(y) := \frac{y-x}{|y-x|} \quad \text{for every } y \in \mathbb{R}^m \setminus \{x\}.$$

**Lemma 2.3.** Let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map, and let  $x \in \Omega$ . For any  $\varphi$  as in Definition 2.1, the following identity holds true for all  $r$  such that  $B_r(x) \subset \Omega$ :

$$\frac{d}{dr} \theta(x, r) = -2r^{2-m} \int \varphi'\left(\frac{|x-y|}{r}\right) \frac{|x-y|}{r^2} |D_{r_x} u(y)|^2 dy. \quad (2.2)$$

In particular, if we let  $\psi = \psi(t)$  denote a primitive function of  $\varphi'(t)t^{m-2}$ , then for  $0 < s < r < \text{dist}(x, \partial\Omega)$  we have:

$$\theta(x, r) - \theta(x, s) = \int \left( \psi\left(\frac{|x-y|}{r}\right) - \psi\left(\frac{|x-y|}{s}\right) \right) |x-y|^{2-m} |D_{r_x} u(y)|^2 dy. \quad (2.3)$$

In case we choose  $\varphi$  to be non-increasing, we have that  $r \mapsto \theta(x, r)$  is non-decreasing; furthermore, if  $-\varphi'(t) \geq (1-t)^+$  then it holds

$$\theta(x, r) - \theta(x, r/2) \geq C \int_{B_{\frac{r}{2}}(x)} \frac{|x-y|}{r^{m-1}} |D_{r_x} u(y)|^2 dy \quad (2.4)$$

for some positive constant  $C = C(m)$ .

*Proof.* The identity (2.2) follows from the inner variation formula, equation (1.3). Indeed, for any fixed  $x \in \Omega$  and  $0 < r < \text{dist}(x, \partial\Omega)$  define the vector field  $X(y) := \varphi\left(\frac{|x-y|}{r}\right)(y-x)$ . If we plug this choice of  $X$  in (1.3), we easily deduce the identity

$$(m-2) \int \varphi\left(\frac{|x-y|}{r}\right) |Du(y)|^2 dy + \int \varphi'\left(\frac{|x-y|}{r}\right) \frac{|x-y|}{r} (|Du(y)|^2 - 2|D_{r_x}u(y)|^2) dy = 0.$$

To conclude, we can differentiate the quantity  $\theta(x, r)$  in  $r$  and obtain the differential identity (2.2).

Now, let  $\psi$  be a primitive function of  $\varphi'(t)t^{m-2}$ . We have

$$\frac{d}{dr}\psi\left(\frac{|x-y|}{r}\right) = -\frac{1}{r}\varphi'\left(\frac{|x-y|}{r}\right)\left(\frac{|x-y|}{r}\right)^{m-1},$$

and thus we can rewrite (2.2) as

$$\frac{d}{dr}\theta(x, r) = 2\frac{d}{dr}\int \psi\left(\frac{|x-y|}{r}\right) |x-y|^{2-m} |D_{r_x}u(y)|^2 dy.$$

Integrating immediately leads to (2.3).

If we choose  $\varphi'(t) \leq 0$ , then (2.2) implies that  $r \mapsto \theta(x, r)$  is non-decreasing. In case  $-\varphi'(t) \geq (1-t)^+$ , we have for  $0 < a \leq \frac{1}{2}$

$$\psi(a) - \psi(2a) = -\int_a^{2a} \varphi'(t)t^{m-2} dt \geq a^{m-1} \left( \frac{2^{m-1}-1}{m-1} - a \frac{2^m-1}{m} \right) \geq C_m a^{m-1}. \quad (2.5)$$

Hence, the estimate (2.4) can be deduced from (2.3) by using the fact that  $\psi$  is non-increasing to estimate

$$\theta(x, r) - \theta(x, r/2) \geq \int_{B_{\frac{r}{2}}(x)} \left( \psi\left(\frac{|x-y|}{r}\right) - \psi\left(\frac{2|x-y|}{r}\right) \right) |x-y|^{2-m} |D_{r_x}u(y)|^2 dy, \quad (2.6)$$

and then using the inequality in (2.5) with  $a = \frac{|x-y|}{r}$  for  $y \in B_{\frac{r}{2}}(x)$ .  $\square$

**Assumption 2.4.** For the rest of the paper, we will assume that  $\varphi$  has been fixed, and that it satisfies the condition  $-\varphi'(t) \geq (1-t)^+$ , so that the inequality (2.4) holds.

### 3. QUANTITATIVE STRATIFICATION

The first step towards the definition of the quantitative singular strata is to introduce the notion of ‘‘model maps’’ having a given number of symmetries. This definition is analogous to [NV17, Definition 1.1].

**Definition 3.1** (*k*-symmetric maps). A map  $h \in W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  is said to be:

- *homogeneous* with respect to  $x \in \mathbb{R}^m$  if

$$h(x + \lambda v) = h(x + v) \quad \text{for all } \lambda > 0, \text{ for every } v \in \mathbb{R}^m,$$

or equivalently if

$$D_{r_x}h = Q[[0]] \quad \text{a.e. in } \mathbb{R}^m.$$

- *k-symmetric* if it is homogeneous with respect to the origin and there exists a linear subspace  $L \subset \mathbb{R}^m$  with  $\dim(L) = k$  along which  $h$  is invariant, that is

$$h(x + v) = h(x) \quad \text{for every } x \in \mathbb{R}^m, \text{ for all } v \in L,$$

or, equivalently, such that

$$D_v h(x) = Q[[0]], \quad \text{for a.e. } x \in \mathbb{R}^m, \text{ for all } v \in L.$$

Observe that if  $h \in W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  is stationary and homogeneous with respect to  $x$  then  $\theta_h(x, s) = \theta_h(x, r)$  for every  $0 < s < r$  by (2.3). Also, if  $h$  is  $k$ -symmetric with invariance subspace  $L$  then the energy of  $h$  in the direction of any  $v \in L$  vanishes. Hence, it is very natural to give the following definition, which is the starting point for introducing the quantitative stratification.

**Definition 3.2.** Given a stationary  $Q$ -harmonic map  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$ , we say that a ball  $B_r(x)$  with  $B_{2r}(x) \subset \Omega$  is  $(k, \varepsilon)$ -symmetric for  $u$  if and only if the following conditions hold:

- (a)  $\theta_u(x, 2r) - \theta_u(x, r) < \varepsilon$ ;
- (b) there exists a linear subspace  $L \subset \mathbb{R}^m$  with  $\dim(L) = k$  such that

$$r^{2-m} \int_{B_r(x)} |D_L u(y)|^2 dy \leq \varepsilon,$$

where

$$\int_{B_r(x)} |D_L u(y)|^2 dy := \int_{B_r(x)} \sum_{i=1}^k |D_{e_i} u(y)|^2 dy,$$

for any orthonormal basis  $\{e_i\}_{i=1}^k$  of  $L$ .

**Remark 3.3.** Observe that the conditions (a) and (b) above are scale-invariant in the following sense. For  $x \in \Omega$  and  $r > 0$  such that  $B_{2r}(x) \subset \Omega$ , consider the blow-up map  $T_{x,r}^u$  given by

$$T_{x,r}^u(y) := u(x + ry).$$

Then,  $B_r(x)$  is  $(k, \varepsilon)$ -symmetric with respect to  $u$  if and only if  $B_1(0)$  is  $(k, \varepsilon)$ -symmetric with respect to  $T_{x,r}^u$ .

**Definition 3.4** (Quantitative stratification). Let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be stationary  $Q$ -harmonic, and let  $\varepsilon, r > 0$  and  $k \in \{0, \dots, m\}$ . We will set

$$\mathcal{S}_{\varepsilon,r}^k(u) := \{x \in \Omega : \text{for no } r \leq s < 1 \text{ is the ball } B_s(x) \text{ } (k+1, \varepsilon)\text{-symmetric with respect to } u\}.$$

It is an immediate consequence of the definition that if  $k' \leq k$ ,  $\varepsilon' \geq \varepsilon$  and  $r' \leq r$  then

$$\mathcal{S}_{\varepsilon',r'}^{k'}(u) \subseteq \mathcal{S}_{\varepsilon,r}^k(u).$$

Hence, we can set

$$\mathcal{S}_{\varepsilon}^k(u) := \bigcap_{r>0} \mathcal{S}_{\varepsilon,r}^k(u),$$

and consider the union

$$\bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^k(u).$$

**Remark 3.5.** Note that from Theorem 1.4 one easily deduces that if  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  is energy minimizing and a ball  $B_r(x)$  is  $(m, \varepsilon_0)$ -symmetric for  $u$ , with the  $\varepsilon_0$  given in there, then  $u$  is Hölder continuous in  $B_{\frac{r}{2}}(x)$ , and thus in particular  $\mathcal{S}_{\varepsilon}^k(u) \cap B_{\frac{r}{2}}(x) = \emptyset$  for every  $k \leq m-1$  and for every  $\varepsilon > 0$ . In fact, we can also conclude that  $(\bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^m(u)) \setminus (\bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^{m-1}(u)) = \Omega \setminus (\bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^{m-1}(u))$  coincides with the set  $\text{reg}_H(u) := \Omega \setminus \text{sing}_H(u)$  of points of Hölder continuity for  $u$ , and  $\text{sing}_H(u) = \bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^{m-1}(u)$ .

In the following proposition, we show that if  $u$  is energy minimizing then the set  $\bigcup_{\varepsilon>0} \mathcal{S}_{\varepsilon}^k(u)$  coincides with the  $k^{\text{th}}$  singular stratum  $\mathcal{S}^k(u)$  as it was defined in (1.5).

**Proposition 3.6.** *Let  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be energy minimizing. Then*

$$\bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^k(u) = \{x : \text{no tangent map to } u \text{ at } x \text{ is } (k+1)\text{-symmetric}\} = \mathcal{S}^k(u).$$

*Proof.* First recall that for any  $x \in \Omega$  there exists at least one tangent map  $\phi \in W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  to  $u$  at  $x$ , and that all tangent maps are energy minimizing and 0-symmetric.

Now, let  $x$  be a point such that there exists a tangent map  $\phi$  to  $u$  at  $x$  which is  $(k+1)$ -symmetric. Then, there is a sequence  $r_j \searrow 0$  of radii such that the corresponding sequence of blow-up maps  $u_j := T_{x,r_j}^u$  satisfies  $\mathcal{G}(u_j, \phi) \rightarrow 0$  in  $L_{loc}^2(\mathbb{R}^m)$  as  $j \rightarrow \infty$  and furthermore

$$\theta_\phi(0, \rho) = \lim_{j \rightarrow \infty} \theta_{u_j}(0, \rho) \quad \forall \rho > 0.$$

In particular, since  $\phi$  is homogeneous with respect to the origin, and thus  $\theta_\phi(0, 2) - \theta_\phi(0, 1) = 0$  by (2.3), for any  $\varepsilon > 0$  there exists  $j_0 = j_0(\varepsilon)$  such that

$$\theta_{u_j}(0, 2) - \theta_{u_j}(0, 1) < \varepsilon \quad \forall j \geq j_0. \quad (3.1)$$

Moreover, since  $\phi$  is  $(k+1)$ -symmetric there exists a linear subspace  $L \subset \mathbb{R}^m$  with  $\dim(L) = k+1$  such that  $D_L \phi = Q[[0]]$  a.e. in  $\mathbb{R}^m$ . Hence, from the convergence of energy for minimizers we deduce that if  $j_0$  is chosen suitably large then also

$$\int_{B_1(0)} |D_L u_j|^2 dy \leq \varepsilon \quad \forall j \geq j_0. \quad (3.2)$$

Together, equations (3.1) and (3.2) imply that  $B_{r_j}(x)$  is  $(k+1, \varepsilon)$ -symmetric for  $u$  if  $j \geq j_0(\varepsilon)$ , and thus  $x \notin \bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^k(u)$ . This proves the first inclusion, namely

$$\bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^k(u) \subseteq \{x \in \Omega : \text{no tangent map to } u \text{ at } x \text{ is } (k+1)\text{-symmetric}\}.$$

In order to prove the other inclusion, assume that  $x \notin \bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^k(u)$ . Then, for every  $j \in \mathbb{N}$  there exist a radius  $r_j > 0$  and a  $(k+1)$ -dimensional linear subspace  $L_j \subset \mathbb{R}^m$  such that if we set  $u_j := T_{x,r_j}^u$  then

$$\theta_{u_j}(0, 2) - \theta_{u_j}(0, 1) < \frac{1}{j} \quad (3.3)$$

and

$$\int_{B_1(0)} |D_{L_j} u_j|^2 dy \leq \frac{1}{j}. \quad (3.4)$$

Modulo a simple right composition of each  $u_j$  with a rotation, we can assume that the invariant subspace is a fixed  $(k+1)$ -dimensional subspace  $L \subset \mathbb{R}^m$ . By the compactness theorem for  $Q$ -valued energy minimizing maps, a subsequence (not relabeled) of the  $u_j$ 's converges in  $L_{loc}^2$  and in energy to an energy minimizing map  $\phi$ . From (3.3) together with (2.4) we deduce that the limit map  $\phi$  is homogeneous with respect to the origin. Furthermore, (3.4) implies that  $\phi$  is invariant along the subspace  $L$ , and thus  $\phi$  is  $(k+1)$ -symmetric. Now, if a subsequence of the  $r_j$ 's converges to 0 then  $\phi$  is by definition a tangent map to  $u$  at  $x$ . If, on the other hand, the  $r_j$ 's are bounded away from 0 then  $u \equiv \phi$  on a ball of positive radius centered at  $x$ , and thus, in particular, all tangent maps to  $u$  at  $x$  coincide with  $\phi$ . In either case, this completes the proof.  $\square$

**Corollary 3.7.** *Let  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be energy minimizing. Then*

$$\left( \bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^{m-1}(u) \right) \setminus \left( \bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^{m-3}(u) \right) = \emptyset.$$

Hence,

$$\text{sing}_H(u) = \bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^{m-3}(u).$$

*Proof.* This is a direct consequence of the Proposition 3.6, since the identity  $\mathcal{S}^{m-1}(u) = \mathcal{S}^{m-3}(u)$  holds for the standard stratification.  $\square$

**Notation 3.8.** From now on, we will simple use the writing  $\mathcal{S}^k(u)$  to denote the union  $\bigcup_{\varepsilon > 0} \mathcal{S}_\varepsilon^k(u)$  of the quantitative strata whenever  $u$  is stationary  $Q$ -harmonic. Following Proposition 3.6, this set coincides with the classical  $\mathcal{S}^k(u)$  when  $u$  is also energy minimizing.

The definition of quantitative stratification that we have proposed differs from the original one introduced by Cheeger and Naber in [CN13a, CN13b] and then used by Naber and Valtorta in [NV17]. Of course, the Cheeger-Naber quantitative stratification can be without any difficulties extended to the  $Q$ -valued context. We recall the definition here, in order to compare it with Definition 3.4.

**Definition 3.9.** Let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$ , and fix  $k \in \{0, \dots, m\}$  and  $\varepsilon > 0$ . A ball  $B_r(x)$  with  $B_{2r}(x) \subset \Omega$  is said to be  $(k, \varepsilon)$ -symmetric for  $u$  in the sense of Cheeger-Naber, or briefly [CN]  $(k, \varepsilon)$ -symmetric, if there exists some  $k$ -symmetric map  $h \in W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  such that

$$\int_{B_r(x)} \mathcal{G}(u(y), h(y-x))^2 dy \leq \varepsilon. \quad (3.5)$$

The definitions of [CN]  $(\varepsilon, r)$ -singular strata and [CN]  $\varepsilon$ -singular strata can be then straightforwardly obtained according to the definition of [CN]  $(k, \varepsilon)$ -symmetry. In particular,  $\mathcal{S}_{[CN]}^k(u) := \bigcup_{\varepsilon > 0} \mathcal{S}_{\varepsilon, [CN]}^k(u)$  classically consists of all points  $x \in \Omega$  having the property that there exists  $\varepsilon > 0$  such that no ball  $B_r(x)$  is [CN]  $(k+1, \varepsilon)$ -symmetric with respect to  $u$ .

The following simple proposition shows that if  $u$  is a minimizing  $Q$ -valued map then Definition 3.2 and Definition 3.9 are equivalent, in the sense that they generate the same stratification.

**Proposition 3.10.** For every  $\varepsilon > 0$  there exists  $\delta = \delta(m, \mathcal{N}, N, Q, \Lambda, \varepsilon) > 0$  such that for any  $Q$ -valued Dirichlet minimizing  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  with  $\mathcal{E}(u, B_{10}(0)) \leq \Lambda$ , and for any  $x, r$  such that  $B_{2r}(x) \subset B_{10}(0)$ :

- (i) if  $B_r(x)$  is  $(k, \delta)$ -symmetric for  $u$ , then it is [CN]  $(k, \varepsilon)$ -symmetric for  $u$ ;
- (ii) if  $B_r(x)$  is [CN]  $(k, \delta)$ -symmetric for  $u$ , then it is  $(k, \varepsilon)$ -symmetric for  $u$ .

*Proof.* Since both the definitions of symmetry are scale-invariant, modulo translations and dilations it suffices to show the validity of the proposition for  $x = 0$  and  $r = 1$ . We start proving the first claim. Assume by contradiction that there exist  $\varepsilon_0 > 0$  and a sequence  $\{u_j\}_{j \in \mathbb{N}}$  of maps as satisfying the assumptions of the proposition and such that  $B_1$  is  $(k, j^{-1})$ -symmetric but with

$$\int_{B_1} \mathcal{G}(u_j(y), h(y))^2 dy > \varepsilon_0 \quad \text{for every } k\text{-symmetric function } h, \text{ for every } j \in \mathbb{N}. \quad (3.6)$$

Modulo rotations, we can assume that the  $k$ -dimensional linear subspace  $L$  such that condition (b) in Definition 3.2 is satisfied is fixed along the sequence: namely, we can assume without loss of generality that

$$\theta_{u_j}(0, 2) - \theta_{u_j}(0, 1) < j^{-1}$$

and that

$$\int_{B_1} |D_L u_j(y)|^2 dy \leq j^{-1}$$

for some fixed  $k$ -dimensional plane  $L \subset \mathbb{R}^m$ . Now, the compactness theorem for  $Q$ -valued energy minimizing maps implies that a subsequence of the  $u_j$ 's (not relabeled) converges in  $L^2(B_{10}(0), \mathcal{A}_Q(\mathbb{R}^N))$  and in energy to a  $Q$ -valued energy minimizing map  $h$  for which

$$\theta_h(0, 2) - \theta_h(0, 1) = 0$$

and

$$D_L h = Q \llbracket 0 \rrbracket \quad \text{a.e. in } B_1.$$

Hence, by (2.4) the map  $h|_{B_1}$  can be extended to a  $k$ -symmetric map (which we still denote by  $h$ ), and the fact that  $\mathcal{G}(u_j, h) \rightarrow 0$  in  $L^2(B_1)$  contradicts (3.6).

For the converse, assume again by contradiction that there exist  $\varepsilon_0$ , a sequence  $\{u_j\}_{j \in \mathbb{N}}$  of maps satisfying the assumptions of the proposition and a sequence  $\{h_j\}_{j \in \mathbb{N}} \subset W_{loc}^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  of  $k$ -symmetric maps such that

$$\int_{B_1} \mathcal{G}(u_j(y), h_j(y))^2 dy \leq j^{-1} \quad (3.7)$$

and such that the ball  $B_1$  is not  $(k, \varepsilon_0)$ -symmetric. Again, after applying suitable rotations we can assume that the invariant subspace for the maps  $h_j$  is a fixed  $k$ -dimensional plane  $L \subset \mathbb{R}^m$ . By compactness, the maps  $u_j$  converge, up to subsequences, to an energy minimizing  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$ . By (3.7), also  $h_j \rightarrow u$  strongly in  $L^2(B_1, \mathcal{A}_Q(\mathcal{N}))$ . Since the space of  $k$ -symmetric maps is  $L^2$ -closed, we deduce that  $u$  is  $k$ -symmetric. Since the  $u_j$ 's converge to  $u$  also in energy, the ball  $B_1$  must be  $(k, \varepsilon_0)$ -symmetric for  $u_j$  if  $j$  is sufficiently large, which is the required contradiction.  $\square$

**Corollary 3.11.** *Let  $u$  be as above. Then, for every  $k \in \{0, \dots, m\}$  one has*

$$\mathcal{S}^k(u) = \mathcal{S}_{[\text{CN}]}^k(u). \quad (3.8)$$

Using more quantitative estimates, the comparison between the two notions of quantitative symmetry can be carried to the case of stationary  $Q$ -harmonic maps.

**Proposition 3.12.** *There exists a constant  $C = C(m, \mathcal{N}, N, Q) > 0$  with the following property. Let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map. If a ball  $B_r(x) \Subset \Omega$  is  $(k, \varepsilon)$ -symmetric for  $u$ , then  $B_{\frac{r}{4}}(x)$  is  $(k, C|\varepsilon \ln(\varepsilon)|)$ -symmetric for  $u$  in the sense of Cheeger-Naber.*

*Proof.* Without loss of generality, we prove the claim for  $x = 0$  and  $r = 1$ . The idea of the proof is to explicitly construct from  $u$  a  $k$ -symmetric map in  $B_{\frac{1}{4}}$ . Modulo a rotation, we can assume that the  $k$ -dimensional plane  $L$  of  $\varepsilon$ -almost symmetry is  $L = \{x_i = 0 : i > k\} = \mathbb{R}^k \times \{0\}$ . For convenience, we will denote the variables of  $\mathbb{R}^k$  with  $y, y'$  and the variables of  $\mathbb{R}^{m-k}$  with  $z, z'$ . The point  $x \in \mathbb{R}^m$  will be therefore given coordinates  $x = (y, z) \in \mathbb{R}^k \times \mathbb{R}^{m-k}$ . With a slight abuse of notation, we will also sometimes regard  $y$  and  $z$  as vectors in  $\mathbb{R}^m$ , thus avoiding the cumbersome, although more correct, writings  $(y, 0)$  and  $(0, z)$ . Finally, when we integrate a function with respect to the variable  $y$  over a ball  $B_r^k \subset \mathbb{R}^k$  we will use the notation  $B_r^y$  as domain of integration (and analogously for the variables  $y', z, z'$ ).

In order to construct the  $k$ -symmetric map, we need to prove two simple inequalities for multiple-valued functions.

*Claim 1:* there exists a constant  $C = C(k, N, Q)$  with the following property. For any function  $f = f(y, z)$  in  $W^{1,2}(B_1^k \times B_1^{m-k}, \mathcal{A}_Q(\mathbb{R}^N))$ , one has

$$\int_{B_1^{y'}} \int_{B_1^y \times B_1^z} \mathcal{G}(f(y, z), f(y', z))^2 \leq C \int_{B_1^y \times B_1^z} |D_L f|^2. \quad (3.9)$$

*Claim 2:* Let  $0 < s_0 < a < 1$ . There exists a constant  $C = C(d, a, Q)$  such that for any  $f \in W^{1,2}(B_1^d, \mathcal{A}_Q(\mathbb{R}^N))$  and every  $a < t \leq 1$  such that  $f|_{\partial B_t^d} \in W^{1,2}(\partial B_t^d, \mathcal{A}_Q(\mathbb{R}^N))$  the following holds:

$$\int_{B_1^d \setminus B_{s_0}^d} \mathcal{G}\left(f(x), f\left(t \frac{x}{|x|}\right)\right)^2 \leq C |\ln(s_0)| \int_{B_1^d \setminus B_{s_0}^d} |Df(x) \cdot x|^2. \quad (3.10)$$

*Proof of Claim 1:* The proof is a consequence of the Poincaré inequality for multiple valued functions, see [DLS11, Proposition 4.9]. Indeed, first observe that for a.e.  $z \in B_1^{m-k}$  the map  $y \mapsto f(y, z)$  is in  $W^{1,2}(B_1^k, \mathcal{A}_Q(\mathbb{R}^N))$ . Hence, by the aforementioned Poincaré inequality, for any such a  $z$  there exists a point  $\bar{f}(z) \in \mathcal{A}_Q(\mathbb{R}^N)$  such that

$$\int_{B_1^y} \mathcal{G}(f(y, z), \bar{f}(z))^2 \leq C \int_{B_1^y} |D_L f(y, z)|^2,$$

where  $C = C(k, N, Q)$ . Hence, by triangle inequality we infer that

$$\begin{aligned} \int_{B_1^y} \int_{B_1^{y'}} \mathcal{G}(f(y, z), f(y', z))^2 &\leq 2\mathcal{H}^k(B_1^k) \left( \int_{B_1^y} \mathcal{G}(f(y, z), \bar{f}(z))^2 + \int_{B_1^{y'}} \mathcal{G}(f(y', z), \bar{f}(z))^2 \right) \\ &\leq C \int_{B_1^y} |D_L f(y, z)|^2. \end{aligned}$$

Integrating now this inequality in  $z \in B_1^{m-k}$  gives (3.9).

*Proof of Claim 2:* First note that for  $\mathcal{H}^{d-1}$ -a.e.  $w \in \partial B_1^d$  the map  $r \mapsto g^w(r) := f(rw)$  is in  $W^{1,2}((0, 1), \mathcal{A}_Q(\mathbb{R}^N))$ . By the  $W^{1,2}$ -selection theorem for multiple-valued functions of one variable (cf. [DLS11, Proposition 1.2]), there exist  $W^{1,2}$  functions  $g_\ell^w: (0, 1) \rightarrow \mathbb{R}^N$  for  $\ell = 1, \dots, Q$  such that  $\left| \frac{d}{dr} g_\ell^w(r) \right| \leq |D_w f(rw)|$  for a.e.  $r \in (0, 1)$ . Now, fix  $t \in (a, 1)$ . Then, by one-dimensional calculus, we have for  $s_0 < s \leq t$  and for every  $\ell \in \{1, \dots, Q\}$  that

$$\begin{aligned} |g_\ell^w(s) - g_\ell^w(t)|^2 &\leq \left( \int_s^t r^{-d-1} dr \right) \left( \int_s^t \left| \frac{d}{dr} g_\ell^w(r) \right|^2 r^{d+1} dr \right) \\ &\leq \frac{s^{-d}}{d} \int_{s_0}^1 |Df(rw) \cdot w|^2 r^{d+1} dr \\ &= \frac{s^{-d}}{d} \int_{s_0}^1 |Df(rw) \cdot rw|^2 r^{d-1} dr. \end{aligned}$$

For  $t \leq s \leq 1$  the same computation holds true interchanging  $t$  and  $s$ : in this case, we estimate  $\frac{t^{-d}}{d} \leq \frac{a^{-d}}{d} s^{-d}$ . Hence in both cases we have

$$|g_\ell^w(s) - g_\ell^w(t)|^2 \leq C s^{-d} \int_{s_0}^1 |Df(rw) \cdot rw|^2 r^{d-1} dr.$$

Summing over  $\ell \in \{1, \dots, Q\}$  and recalling the definition of the metric  $\mathcal{G}$  this produces

$$\mathcal{G}(f(sw), f(tw))^2 \leq C s^{-d} \int_{s_0}^1 |Df(rw) \cdot rw|^2 r^{d-1} dr \quad \text{for every } s \in (s_0, 1),$$

where  $C = C(d, a, Q)$ . Multiply by  $s^{d-1}$  and integrate in  $s$  between  $s_0$  and 1 to obtain

$$\int_{s_0}^1 \mathcal{G}(f(sw), f(tw))^2 s^{d-1} ds \leq C |\ln(s_0)| \int_{s_0}^1 |Df(rw) \cdot rw|^2 r^{d-1} dr.$$

Integrating now in  $w \in \partial B_1$  gives inequality (3.10).

We are now ready to prove the proposition. Let  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map, and assume that  $B_1$  is  $(k, \varepsilon)$ -symmetric for  $u$ . By (b) in Definition 3.2, we can fix  $\frac{1}{4} \leq t \leq \frac{1}{\sqrt{2}}$  such that  $x \in \partial B_t \mapsto u(x)$  is in  $W^{1,2}(\partial B_t, \mathcal{A}_Q(\mathcal{N}))$  and satisfies  $\int_{\partial B_t} |D_L u|^2 \leq C \int_{B_1} |D_L u|^2$ . For a.e.  $y' \in B_t^k$  we have that the map  $z \mapsto v_{y'}(z) := u(y', z)$  is in  $W^{1,2}(B_t^{m-k}, \mathcal{A}_Q(\mathcal{N}))$ . Hence, by the scaled version of (3.10) with  $d = m - k$  we have for any  $0 < s_0 < \frac{1}{4}$  that

$$\int_{B_t^z \setminus B_{s_0}^z} \mathcal{G} \left( v_{y'}(z), v_{y'} \left( t \frac{z}{|z|} \right) \right)^2 \leq C |\ln(s_0)| \int_{B_t^z} |D u(y', z) \cdot z|^2,$$

where  $C = C(m, Q)$ . Integrating this now in  $y' \in B_t^k$  we obtain

$$\int_{B_t^{y'}} \int_{B_t^z \setminus B_{s_0}^z} \mathcal{G} \left( v_{y'}(z), v_{y'} \left( t \frac{z}{|z|} \right) \right)^2 \leq C |\ln s_0| \int_{B_t^k \times B_t^{m-k}} |D u(y, z) \cdot z|^2 \, dy dz.$$

Adding the scaled version of (3.9), since  $B_t^k \times B_t^{m-k} \subset B_1$  we obtain

$$\begin{aligned} \int_{B_t^{y'}} \left( \int_{B_t^y \times B_t^z} \mathcal{G}(u(y, z), v_{y'}(z))^2 + \int_{B_t^z \setminus B_{s_0}^z} \mathcal{G} \left( v_{y'}(z), v_{y'} \left( t \frac{z}{|z|} \right) \right)^2 \right) \\ \leq C \left( \int_{B_1} |D_L u|^2 + |\ln(s_0)| \int_{B_1} |D u(x) \cdot x|^2 \right). \end{aligned}$$

Hence there exists  $y'_0 \in B_t^k$  such that

$$\begin{aligned} \int_{B_t^y \times B_t^z} \mathcal{G}(u(y, z), v_{y'_0}(z))^2 + \int_{B_t^z \setminus B_{s_0}^z} \mathcal{G} \left( v_{y'_0}(z), v_{y'_0} \left( t \frac{z}{|z|} \right) \right)^2 \\ \leq \frac{C}{\mathcal{H}^k(B_t^k)} \left( \int_{B_1} |D_L u|^2 + |\ln(s_0)| \int_{B_1} |D u(x) \cdot x|^2 \right) \\ \leq C(1 + |\ln(s_0)|)\varepsilon, \end{aligned}$$

where in the last inequality we have used that the ball  $B_1$  is, by assumption,  $(k, \varepsilon)$ -symmetric for  $u$  together with (2.4).

Set  $h(x) = h(y, z) := v_{y'_0} \left( t \frac{z}{|z|} \right) \in W^{1,2}(B_t, \mathcal{A}_Q(\mathcal{N}))$ . Note that, by definition,  $h$  is homogeneous with respect to 0. Furthermore,  $h$  is  $k$ -symmetric. An application of the triangle inequality gives

$$\begin{aligned} \int_{B_t^y \times B_t^z} \mathcal{G}(u(y, z), h(x))^2 \leq 2 \int_{B_t^y \times B_t^z} \mathcal{G}(u(y, z), v_{y'_0}(z))^2 + 2 \int_{B_t^y \times (B_t^z \setminus B_{s_0}^z)} \mathcal{G}(v_{y'_0}(z), h(x))^2 \\ + 2 \int_{B_t^y \times B_{s_0}^z} \mathcal{G}(v_{y'_0}(z), h(x))^2. \end{aligned}$$

As we have shown above, the first two integrals can be bounded by  $C(1 + |\ln(s_0)|)\varepsilon$ . As for the last integral, we estimate it brutally by fixing a point  $p \in \mathcal{N}$  and computing

$$\int_{B_t^y \times B_{s_0}^z} \mathcal{G}(v_{y'_0}(z), h(x))^2 \leq 2 \sup_{x \in B_1} \mathcal{G}(u(x), Q[[p]])^2 \mathcal{H}^k(B_1^k) \mathcal{H}^{m-k}(B_{s_0}^{m-k}) \leq CQ \text{diam}(\mathcal{N})^2 s_0^{m-k}$$

Hence choosing  $s_0 = \varepsilon$  proves the proposition, since we get

$$\int_{B_t} \mathcal{G}(u(x), h(x))^2 \leq C|\varepsilon \ln(\varepsilon)|.$$

□

**Corollary 3.13.** *Let  $u \in W_{loc}^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map. Then*

$$\mathcal{S}_{[\text{CN}]}^k(u) \subset \mathcal{S}^k(u).$$

We conclude the section with two propositions about the characterization of the singular set for minimizing and stationary maps. The first one is the following effective version of Corollary 3.7.

**Proposition 3.14.** *There exists  $\varepsilon = \varepsilon(m, \mathcal{N}, N, Q, \Lambda)$  such that for any  $Q$ -valued Dirichlet minimizing  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  with  $\mathcal{E}(u, B_{10}(0)) \leq \Lambda$ , the following holds:*

$$B_1 \cap (\mathcal{S}^{m-1}(u) \setminus \mathcal{S}_\varepsilon^{m-3}(u)) = \emptyset.$$

*Proof.* The proof is by contradiction. Assume, therefore, that for every  $j \in \mathbb{N}$  there exists  $u_j$  satisfying the assumptions of the proposition with a point  $x_j \in B_1 \cap (\mathcal{S}^{m-1}(u_j) \setminus \mathcal{S}_{j^{-1}}^{m-3}(u_j))$ . Since  $x_j \notin \mathcal{S}_{j^{-1}}^{m-3}(u_j)$ , there exists  $0 < r_j < 1$  and a linear subspace  $L_j \subset \mathbb{R}^m$  with  $\dim(L_j) = m - 2$  such that

$$\theta_{u_j}(x_j, 2r_j) - \theta_{u_j}(x_j, r_j) \leq j^{-1}, \quad (3.11)$$

$$r_j^{2-m} \int_{B_{r_j}(x_j)} |D_{L_j} u_j|^2 \leq j^{-1}. \quad (3.12)$$

As usual, without loss of generality we assume that the  $(m - 2)$ -planes of  $j^{-1}$ -almost symmetry are a fixed subspace  $L$  along the sequence. Set  $v_j(y) := u_j(x_j + r_j y)$ , and re-write the equations (3.11) and (3.12) in terms of  $v_j$ :

$$\theta_{v_j}(0, 2) - \theta_{v_j}(0, 1) \leq j^{-1}, \quad (3.13)$$

$$\int_{B_1} |D_L v_j|^2 \leq j^{-1}. \quad (3.14)$$

Now, by an elementary computation it is immediate to see that for every  $\rho \in (0, 8)$  one has

$$\rho^{2-m} \int_{B_\rho} |Dv_j|^2 = (\rho r_j)^{2-m} \int_{B_{\rho r_j}(x_j)} |Du_j|^2 \leq C_m \Lambda.$$

Hence, by the Compactness Theorem 1.3, the sequence  $\{v_j\}_{j \in \mathbb{N}}$  converges up to subsequences in  $L^2(B_8, \mathcal{A}_Q(\mathbb{R}^N))$  and in energy to a  $Q$ -valued energy minimizing map  $v$  for which

$$\theta_v(0, 2) - \theta_v(0, 1) = 0, \quad (3.15)$$

$$\int_{B_1} |D_L v|^2 = 0. \quad (3.16)$$

In particular,  $v|_{B_1}$  can be extended to an  $(m - 2)$ -symmetric energy minimizer. This implies that a fortiori  $0 \in \text{reg}_H(v)$ . Thus,  $0 \notin \mathcal{S}^{m-1}(v_j)$  for  $j$  large, which contradicts the fact that  $x_j \in \mathcal{S}^{m-1}(u_j)$ . □

In the single-valued case  $Q = 1$ , we have the following result on the quantitative stratification for stationary harmonic maps.

**Proposition 3.15.** *There exists  $\varepsilon = \varepsilon(m, \mathcal{N})$  such that for any single-valued stationary harmonic map  $u \in W_{loc}^{1,2}(\Omega, \mathcal{N})$  the following holds:*

$$\mathcal{S}^{m-1}(u) \setminus \mathcal{S}_\varepsilon^{m-2}(u) = \emptyset.$$

*Proof.* Proposition 3.15 is a consequence of the inner variation formula. First we derive a general estimate and show afterwards how it implies the proposition.

Let us consider a single-valued harmonic map  $u$  in  $B_1$  that satisfies the inner variation formula (1.3) with  $Q = 1$ . We fix two non-negative, non-increasing functions  $\psi, \varphi \in C_c^1\left(\left[0, \frac{1}{\sqrt{2}}\right]\right)$  and a  $k$ -dimensional subspace  $L \subset \mathbb{R}^m$ . After a rotation, we may assume that  $L = \{x_i = 0: i = k+1, \dots, m\}$ . To make the notation a bit simpler we will write  $x = (y, z) \in L \times L^\perp$ , and by a slight abuse of notation we shall again consider  $z = (0, z)$  as a vector in  $\mathbb{R}^m$ . Consider the vector field  $X(y, z) := \psi(|y|)\varphi(|z|)z = \psi\varphi z$ . We have  $DX = \psi\varphi P^\perp + \frac{\psi\varphi'}{|z|}z \otimes z + \frac{\psi'\varphi}{|y|}z \otimes y$ , where  $P^\perp$  denotes the orthogonal projection onto  $L^\perp$ . We use this vector field in the inner variation formula (1.3) and obtain

$$0 = \int |Du|^2 \left( (m-k)\psi\varphi + \psi\varphi'|z| \right) - 2 \left( \psi\varphi |D_{L^\perp} u|^2 + \psi\varphi' \frac{1}{|z|} |Du \cdot z|^2 + \psi'\varphi \left\langle Du \cdot z, Du \cdot \frac{y}{|y|} \right\rangle \right).$$

Observe that

$$(m-k)\psi\varphi |Du|^2 - 2\psi\varphi |D_{L^\perp} u|^2 = (m-k-2)\psi\varphi |Du|^2 + 2\psi\varphi |D_L u|^2.$$

Furthermore, we can write  $z = x - y$  to estimate

$$\begin{aligned} |Du \cdot z|^2 &\leq 2|Du \cdot x|^2 + 2|Du \cdot y|^2 \\ \left\langle Du \cdot z, Du \cdot \frac{y}{|y|} \right\rangle &\leq \frac{1}{2}|Du \cdot x|^2 + \frac{1}{2} \left| Du \cdot \frac{y}{|y|} \right|^2. \end{aligned}$$

Combining all together, and recalling that  $\varphi', \psi' \leq 0$ , we obtain the inequality

$$\begin{aligned} &\int - \left( (m-k-2)\psi\varphi + \psi\varphi'|z| \right) |Du|^2 \\ &\leq \int 2\psi\varphi |D_L u|^2 + \int \frac{-4\psi\varphi'}{|z|} \left( |Du \cdot x|^2 + |Du \cdot y|^2 \right) - \psi'\varphi \left( |Du \cdot x|^2 + \left| Du \cdot \frac{y}{|y|} \right|^2 \right). \end{aligned} \quad (3.17)$$

We are ready to prove the proposition. Fix  $\varepsilon > 0$  to be determined later, and suppose by contradiction that there is a point  $x \in \mathcal{S}^{m-1}(u) \setminus \mathcal{S}_\varepsilon^{m-2}(u)$ . Since  $x \notin \mathcal{S}_\varepsilon^{m-2}$ , there exists  $r = r(\varepsilon) > 0$  and an  $(m-1)$ -dimensional subspace  $L = L(\varepsilon)$  such that  $r^{2-m} \int_{B_r(x)} |D_L u|^2 < \varepsilon$  and  $\theta(x, 2r) - \theta(x, r) < \varepsilon$ . By translation and scaling, i.e. passing to  $T_{x,r}^u$ , we may assume that  $x = 0$  and  $r = 1$ . However, for notational convenience, we will still write  $u$  for  $T_{x,r}^u$ . After a further rotation we may assume that  $L = \{x_m = 0\}$ . Now, we have  $B_{\frac{1}{2}} \subset B_{\frac{1}{\sqrt{2}}}^{m-1} \times \left( -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \subset B_1$ . Fix a function  $\eta \in C^1$  with  $\eta' \leq 0$  and  $\eta(t) = 1$  for  $t \leq \frac{1}{2}$ ,  $\eta(t) = 0$  for  $t \geq \frac{1}{\sqrt{2}}$ . Set  $\varphi = \psi := \eta$  in (3.17). Recall that in our situation  $k = m-1$ , and thus  $-(m-k-2) = 1$ . Furthermore, we have  $\psi\varphi \geq \mathbf{1}_{B_{\frac{1}{2}}}$ , and  $\frac{|4\psi\varphi'}{|z|}, |\psi'\varphi|$  are bounded and supported in  $B_1$ . Hence (3.17) reads in our case

$$\int_{B_{\frac{1}{2}}} |Du|^2 \leq C \int_{B_1} \left( |x| |D_r u|^2 + |D_L u|^2 \right),$$

where  $r(x) = r_0(x) = \frac{x}{|x|}$ . By (2.4), we deduce that  $\int_{B_{\frac{1}{2}}} |Du|^2 \leq C\varepsilon$ . If  $\varepsilon > 0$  is chosen sufficient small, i.e.  $C\varepsilon < \varepsilon_0$  where  $\varepsilon_0 = \varepsilon_0(m, \mathcal{N})$  is the threshold in the  $\varepsilon$ -regularity theorem for stationary harmonic maps (cf. [Bet93, RS08]), this allows to infer that  $u$  is Hölder continuous in  $B_{\frac{1}{4}}$ , and hence 0 is a regular point. This contradicts the assumption that  $0 \in \mathcal{S}^{m-1}$ .  $\square$

**Remark 3.16.** Note that the above proposition could be extended (with exactly the same proof) to the case of stationary  $Q$ -harmonic maps if an  $\varepsilon$ -regularity theorem was available in that case.

**Remark 3.17.** A converse to Corollary 3.13 cannot hold in general due to an example of [HLP92]. They show the existence of a stationary harmonic map  $u : B_1 \subset \mathbb{R}^3 \rightarrow \mathbb{S}^2$  that has a constant tangent map at a singularity  $\bar{x}$ . Hence we have

$$\bar{x} \in \mathcal{S}_{[\text{CN}]}^m(u) \setminus \mathcal{S}_{[\text{CN}]}^{m-1}(u).$$

On the other hand, due to the  $\varepsilon$ -regularity theorem in [Bet93], it cannot be that  $\bar{x} \notin \mathcal{S}^{m-1}(u)$ . Nevertheless, if one knows a priori that  $\mathcal{N}$  has the property that there is no smooth harmonic map  $v : \mathbb{S}^2 \rightarrow \mathcal{N}$ , [Lin99, Lemma 3.1, Remark 3.2] imply that every blow-up sequence  $T_{x, r_k}^u$  converges in fact strongly in energy to one of its tangent maps  $T$ . In this case, the proof of Proposition 3.10 applies as well for stationary maps, and hence  $\mathcal{S}_{[\text{CN}]}^k(u) = \mathcal{S}^k(u)$ .

**3.1. Main theorem on the quantitative strata.** Since the relevant terminology has been introduced now, we can finally state the main estimates that we are going to prove on the singular strata.

**Theorem 3.18.** *Given a stationary  $Q$ -valued map  $u : B_2(0) \subseteq \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  with  $\mathcal{E}(u, B_2(0)) \leq \Lambda$ , let  $\mathcal{S}_{\varepsilon, r}^k(u)$  be its quantitative singular strata. Then, if  $B_r(\mathcal{S}_{\varepsilon, r}^k(u)) := \bigcup_{x \in \mathcal{S}_{\varepsilon, r}^k(u)} B_r(x)$ , we have*

$$\text{Vol}\left(B_r\left(\mathcal{S}_{\varepsilon, r}^k(u) \cap B_1(0)\right)\right) \leq C(m, \mathcal{N}, Q, \Lambda, \varepsilon)r^{m-k}. \quad (3.18)$$

Moreover,  $\mathcal{S}_{\varepsilon}^k(u)$  is  $k$ -rectifiable for all  $\varepsilon \geq 0$ .

**Remark 3.19.** This theorem is similar in spirit to [NV17, Theorem 1.3].

Note that this and the  $\varepsilon$ -regularity theorem for  $Q$ -valued energy minimizers immediately imply Theorem 0.2 as a corollary. Indeed, we have

**Corollary 3.20** (Theorem 0.2). *Given an energy minimizing  $Q$ -valued map  $u : B_2(0) \subseteq \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  with energy bounded by  $\Lambda$ , if  $B_r(\text{sing}_H(u)) := \bigcup_{x \in \text{sing}_H(u)} B_r(x)$  then we have*

$$\text{Vol}\left(B_r(\text{sing}_H(u) \cap B_1(0))\right) \leq Cr^3, \quad (3.19)$$

where  $C = C(m, \mathcal{N}, Q, \Lambda)$ . Moreover,  $\text{sing}_H(u)$  is  $(m-3)$ -rectifiable.

*Proof.* By remark 3.5 and proposition 3.14, there exists an  $\varepsilon$  such that

$$\mathcal{S}_{\varepsilon}^{m-3}(u) \cap B_1(0) = \text{sing}_H(u) \cap B_1(0). \quad (3.20)$$

Thus Theorem 3.18 immediately proves the volume estimates and rectifiability for  $\text{sing}_H(u)$ .  $\square$

We postpone the proof of Theorem 3.18 to Section 7, after having discussed a few technical tools needed to complete it.

4. QUANTITATIVE  $\varepsilon$ -REGULARITY THEOREMS

In this section we are going to present the proof of a quantitative version of the  $\varepsilon$ -regularity theorem for  $Q$ -valued minimizers, cf. Theorem 4.3 below, which in turn implies Corollary 4.4, providing sufficient conditions under which the singular set  $\text{sing}_H(u)$  is constrained to live in the tubular neighborhood of an affine subspace of  $\mathbb{R}^m$  of appropriate dimension. We start with the following definition, analogous to [NV17, Definition 4.5].

**Definition 4.1.** Let  $y_0, y_1, \dots, y_k$  be  $(k+1)$  points in  $B_1(0) \subset \mathbb{R}^m$ , and let  $\rho > 0$ . We say that these points  $\rho$ -effectively span a  $k$ -dimensional affine subspace if

$$\text{dist}(y_i, y_0 + \text{span}[y_1 - y_0, \dots, y_{i-1} - y_0]) \geq 2\rho \quad \text{for every } i = 1, \dots, k. \quad (4.1)$$

A set  $F \subset B_1(0)$   $\rho$ -effectively spans a  $k$ -dimensional subspace if there exist points  $\{y_i\}_{i=0}^k \subset F$  which  $\rho$ -effectively span a  $k$ -dimensional subspace.

**Remark 4.2.** It is easy to see that if the points  $\{y_i\}_{i=0}^k$   $\rho$ -effectively span a  $k$ -dimensional affine subspace then for every point

$$x \in y_0 + \text{span}[y_1 - y_0, \dots, y_k - y_0]$$

there exists a unique set of numbers  $\{\alpha_i\}_{i=1}^k$  such that

$$x = y_0 + \sum_{i=1}^k \alpha_i (y_i - y_0), \quad |\alpha_i| \leq C(m, \rho) |x - y_0|.$$

Furthermore, the notion of  $\rho$ -effectively spanning a  $k$ -dimensional affine subspace passes to the limit: if for every  $j \in \mathbb{N}$  the points  $\{y_i^j\}_{i=0}^k$   $\rho$ -effectively span a  $k$ -dimensional subspace and there exist the limits  $y_i := \lim_{j \rightarrow \infty} y_i^j$ , then also the points  $\{y_i\}_{i=0}^k$   $\rho$ -effectively span a  $k$ -dimensional subspace.

We can now state the main theorem of this section.

**Theorem 4.3.** Let  $\varepsilon, \rho > 0$  be fixed. There exists  $\delta, \bar{r} > 0$ , depending on  $m, \rho, \Lambda, \varepsilon$ , with the following property. Let  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map with energy bounded by  $\Lambda$ , let  $r \leq 1$ , and let

$$F := \{y \in B_r(0) : \theta(y, 4r) - \theta(y, 2r) < \delta\}.$$

If  $F$   $(\rho \cdot r)$ -effectively spans a  $k$ -dimensional subspace  $L$ , then

$$\left( \mathcal{S}_{\varepsilon, r\bar{r}}^k(u) \cap B_{\frac{r}{2}}(0) \right) \subset B_{\rho r}(L). \quad (4.2)$$

**Corollary 4.4.** For every  $\rho > 0$ , there exists  $\delta = \delta(m, \mathcal{N}, N, Q, \Lambda, \rho) > 0$  with the following property. Let  $u: B_{10}(0) \subset \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  be a  $W^{1,2}$  map with energy bounded by  $\Lambda$ , and let  $r \leq 1$ .

(i) In case  $u$  is energy minimizing, if there exist  $m-2$  points  $\{y_i\}_{i=0}^{m-3} \subset B_r(0)$  which  $(\rho \cdot r)$ -effectively span an  $(m-3)$ -dimensional affine subspace  $L \subset \mathbb{R}^m$  and such that

$$\theta(y_i, 4r) - \theta(y_i, 2r) < \delta \quad \text{for every } i = 0, \dots, m-3,$$

then

$$\text{sing}_H(u) \cap B_{\frac{r}{2}}(0) \subset B_{\rho r}(L);$$

(ii) in case  $u$  is single-valued and stationary harmonic, if there exist  $m-1$  points  $\{y_i\}_{i=0}^{m-2} \subset B_r(0)$  which  $(\rho \cdot r)$ -effectively span an  $(m-2)$ -dimensional affine subspace  $L \subset \mathbb{R}^m$  and such that

$$\theta(y_i, 4r) - \theta(y_i, 2r) < \delta \quad \text{for every } i = 0, \dots, m-2,$$

then

$$\text{sing}_H(u) \cap B_{\frac{r}{2}}(0) \subset B_{\rho r}(L).$$

In particular, if  $m = 3$  and  $u$  is a  $Q$ -valued energy minimizer, and if  $\theta(y_0, 4r) - \theta(y_0, 2r) < \delta$  then

$$B_{\frac{r}{2}}(y_0) \setminus B_{\rho r}(y_0) \subset \text{reg}_H(u).$$

The same holds if  $m = 2$  and  $u$  is single-valued and stationary.

*Proof of Corollary 4.4.* It follows immediately from Theorem 4.3 and Propositions 3.14 for the minimizing case and 3.15 for the stationary harmonic case.  $\square$

For the proof of Theorem 4.3 we will need the following lemma.

**Lemma 4.5.** *Let  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map, and let  $r \leq 1$ . If  $\{y_i\}_{i=0}^k \subset B_r(0)$  ( $\rho \cdot r$ )-effectively span a  $k$ -dimensional affine subspace  $L \subset \mathbb{R}^m$ , then*

$$r^{-m} \int_{B_r(0)} \left( r^2 |D_{\hat{L}} u(z)|^2 + |D_v u(z)|^2 \right) dz \leq C(m, \rho) \sum_{i=0}^k (\theta(y_i, 4r) - \theta(y_i, 2r)), \quad (4.3)$$

where  $\hat{L}$  is the linear part of  $L$  and  $v$  is the vector field  $v(z) := D\left(\frac{1}{2} \text{dist}^2(z, L)\right)$ .

*Proof.* It is an immediate consequence of (2.4) that there exists a constant  $C = C(m)$  such that

$$r^{-m} \int_{B_r(x)} |Du(z) \cdot (z-x)|^2 dz \leq \int_{B_r(x)} \frac{|z-x|}{r^{m-1}} \left| Du(z) \cdot \frac{z-x}{|z-x|} \right|^2 dz \leq C(m) (\theta(x, 2r) - \theta(x, r)) \quad (4.4)$$

whenever  $B_{2r}(x) \subset B_{10}(0)$ . Now, assume that  $y_0, y_1, \dots, y_k$  are as in the statement, and observe that for every unit vector  $e$  in the linear part  $\hat{L}$  of  $L$  there exists a unique set of numbers  $\{\alpha_i\}_{i=1}^k$  such that

$$e = r^{-1} \sum_{i=1}^k \alpha_i (y_i - y_0), \quad |\alpha_i| \leq C(m, \rho).$$

Hence, we get

$$\begin{aligned} r^{2-m} \int_{B_r(0)} |D_e u(z)|^2 dz &\leq C(m, \rho) r^{-m} \sum_{i=1}^k \int_{B_r(0)} |Du(z) \cdot (y_i - y_0)|^2 dz \\ &\leq C(m, \rho) r^{-m} \sum_{i=0}^k \int_{B_r(0)} |Du(z) \cdot (z - y_i)|^2 dz \\ &\leq C(m, \rho) r^{-m} \sum_{i=0}^k \int_{B_{2r}(y_i)} |Du(z) \cdot (z - y_i)|^2 dz \\ &\stackrel{(4.4)}{\leq} C(m, \rho) \sum_{i=0}^k (\theta(y_i, 4r) - \theta(y_i, 2r)). \end{aligned}$$

Summing over an orthonormal basis  $e_1, \dots, e_k$  of  $\hat{L}$  produces

$$r^{2-m} \int_{B_r(0)} |D_{\hat{L}} u(z)|^2 dz \leq C(m, \rho) \sum_{i=0}^k (\theta(y_i, 4r) - \theta(y_i, 2r)). \quad (4.5)$$

As for the second term, let  $z \in B_r(0)$ , and let  $\pi := \pi_L(z)$  be the orthogonal projection of  $z$  onto  $L$ . Of course,

$$v(z) := D\left(\frac{1}{2}\text{dist}^2(z, L)\right) = z - \pi.$$

On the other hand, we have as usual that

$$\pi = y_0 + \sum_{i=1}^k \alpha_i (y_i - y_0), \quad |\alpha_i| \leq C(m, \rho) |\pi - y_0| \leq C(m, \rho) r,$$

and thus

$$v(z) = z - \left( y_0 + \sum_{i=1}^k \alpha_i (y_i - y_0) \right).$$

Arguing as above, one concludes that also

$$r^{-m} \int_{B_r(0)} |D_v u(z)|^2 dz \leq C(m, \rho) \sum_{i=0}^k (\theta(y_i, 4r) - \theta(y_i, 2r)), \quad (4.6)$$

which together with (4.5) completes the proof of (4.4).  $\square$

For  $k = m - 2$ , the conclusions of the previous lemma can be improved using again the inner variation formula. Although we are not going to use this lemma explicitly in the following pages, we want to record it here for its own sake and because it is a quantitative version of Proposition 3.15.

**Lemma 4.6.** *Let  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map, and let  $r \leq 1$ . If  $\{y_i\}_{i=0}^{m-2} \subset B_r(0)$   $(\rho \cdot r)$ -effectively span an  $(m - 2)$ -dimensional affine subspace  $L \subset \mathbb{R}^m$ , then*

$$r^{-m} \int_{B_r(0)} \left( r^2 |D_{\hat{L}} u|^2 + |D_{\hat{L}^\perp} u|^2 |v|^2 \right) \leq C(m, \rho) \sum_{i=0}^{m-2} (\theta(y_i, 8r) - \theta(y_i, 4r)), \quad (4.7)$$

where  $\hat{L}$  is the linear part of  $L$ ,  $\hat{L}^\perp$  is its orthogonal complement in  $\mathbb{R}^m$  and  $v$  is the vector field  $v(x) := D\left(\frac{1}{2}\text{dist}^2(x, L)\right)$ .

*Proof.* The proof is very similar to the one of Proposition 3.15: also in this case, we will make use of the stationary equation with a suitable choice of the vector field  $X$ . Without loss of generality, we can assume that  $r = 1$ . Furthermore, modulo translations and rotations we can assume that  $L = \{x_i = 0 : i = m - 1, m\}$ . As usual, coordinates on  $L$  and  $L^\perp$  will be denoted by  $y$  and  $z$  respectively, and in order to simplify our notation the vectors  $(y, 0)$  and  $(0, z)$  in  $L \times L^\perp$  will be simply denoted by  $y$  and  $z$ . Observe that under these assumptions one has  $v(x) = z$  for every  $x = (y, z) \in B_1$ . Now, let  $\psi = \psi(y)$  be a cut-off function of the variable  $y \in L$ , with  $\psi \equiv 1$  in  $B_1^{m-2}$ ,  $\text{spt}(\psi) \subset B_2^{m-2}$  and  $|D\psi| \leq 1$ . Let also  $\varphi(t) := \max\{1 - t, 0\}$ , and consider the vector field  $X(y, z) := \psi(y)\varphi(|z|^2)z = \psi\varphi z$ . We can immediately compute  $DX = \psi\varphi P^\perp + \varphi z \otimes D\psi + 2\psi\varphi' z \otimes z$ . With this choice of  $X$ , the inner variation formula (1.3) reads

$$\begin{aligned} 0 &= \int |Du|^2 \left( 2\psi\varphi + 2\psi\varphi'|z|^2 \right) - 2 \left( \psi\varphi |D_{L^\perp} u|^2 + \varphi \langle Du \cdot z, Du \cdot D\psi \rangle + 2\psi\varphi' |Du \cdot z|^2 \right) \\ &= 2 \int \psi\varphi |D_L u|^2 + \psi\varphi' |Du|^2 |z|^2 - \left( \varphi \langle Du \cdot z, Du \cdot D\psi \rangle + 2\psi\varphi' |Du \cdot z|^2 \right). \end{aligned}$$

In particular, since  $\varphi'(|z|^2) = -\chi_{\{|z|\leq 1\}}$ ,  $\psi|_{\{|y|\leq 1\}} \equiv 1$  and  $B_1 \subset B_1^{m-2} \times B_1^2 \subset B_2$ , we immediately deduce

$$\int_{B_1} |Du|^2 |z|^2 \leq C \int_{B_2} (|D_L u|^2 + |Du \cdot z|^2).$$

The estimate (4.7) then follows from Lemma 4.5.  $\square$

We are now ready to prove Theorem 4.3.

*Proof of Theorem 4.3.* Since the statement is scale-invariant, there is no loss of generality in proving it only in the case  $r = 1$ . Let  $\{y_i\}_{i=0}^k \subset F$   $\rho$ -effectively span the  $k$ -dimensional subspace  $L$ , and let  $x$  be any point in  $B_{\frac{1}{2}}(0) \setminus B_\rho(L)$ . The goal is to prove that  $x \notin \mathcal{S}_{\varepsilon, \bar{r}}^k(u)$  for some  $\bar{r} > 0$ , and thus that there exists  $\bar{r} > 0$  and a radius  $r_x \in [\bar{r}, 1)$  such that the ball  $B_{r_x}(x)$  is  $(k+1, \varepsilon)$ -symmetric for  $u$ . Let  $0 < \delta \ll 1$  to be chosen later. Since  $x \in B_{\frac{1}{2}}(0)$ ,  $B_\sigma(x) \subset B_1(0)$  for every  $0 < \sigma < \frac{1}{2}$ . Hence, we deduce from Lemma 4.5 that

$$\int_{B_\sigma(x)} |D_{\hat{L}} u|^2 \leq C(m, \rho) \delta$$

for any such  $\sigma$ . In order to gain another direction along which the energy is small, we let  $v(z) := D\left(\frac{1}{2} \text{dist}^2(z, L)\right)$ , and we set  $e := \frac{v(x)}{|v(x)|}$ . Note that  $|v(x)| = \text{dist}(x, L) \geq \rho$ . Again by Lemma 4.5 and by the monotonicity of the function  $r \mapsto \mathcal{E}(u, B_r(x))$ , we have

$$\begin{aligned} \int_{B_\sigma(x)} |D_e u|^2 &\leq \rho^{-2} \int_{B_\sigma(x)} |Du(z) \cdot v(x)|^2 \\ &\leq 2\rho^{-2} \left( \int_{B_\sigma(x)} |Du(z) \cdot v(z)|^2 + \int_{B_\sigma(x)} |Du(z) \cdot (v(z) - v(x))|^2 \right) \\ &\leq C \int_{B_1(0)} |D_v u|^2 + C\sigma^2 \int_{B_\sigma(x)} |Du|^2 \\ &\leq C\delta + C\Lambda\sigma^m, \end{aligned}$$

where  $C = C(m, \rho)$ . Hence, if  $V := \hat{L} \oplus \text{span}(e)$  then

$$\int_{B_\sigma(x)} |D_V u|^2 \leq C\delta + C\Lambda\sigma^m \quad (4.8)$$

for every  $0 < \sigma < \frac{1}{2}$ . Note that  $\dim(V) = k+1$ .

Fix now  $\varepsilon > 0$ , and let  $\bar{\sigma} = \bar{\sigma}(m, \rho, \Lambda, \varepsilon) < \frac{1}{2}$  be such that  $C\Lambda\bar{\sigma}^2 \leq \frac{\varepsilon}{2}$ . We claim that for any  $0 < \tau \ll 1$  there exists  $\tau\bar{\sigma} \leq r_x < \bar{\sigma}$  such that

$$\theta(x, 2r_x) - \theta(x, r_x) \leq \frac{2c_1(m)\Lambda}{-\log_2(2\tau)}. \quad (4.9)$$

Indeed, otherwise for any integer  $M \in \left(\frac{3}{4} \log_2\left(\frac{1}{2\tau}\right), \log_2\left(\frac{1}{2\tau}\right)\right)$  we would get

$$c_1(m)\Lambda \geq \theta(x, \bar{\sigma}) \geq \sum_{i=0}^M \theta(x, 2^{-i}\bar{\sigma}) - \theta(x, 2^{-(i+1)}\bar{\sigma}) \geq M \frac{2c_1(m)\Lambda}{-\log_2(2\tau)} \geq \frac{3}{2} c_1(m)\Lambda,$$

which is impossible. Hence, if we fix  $\tau = \tau(m, \Lambda, \varepsilon)$  so small that  $\frac{2c_1(m)\Lambda}{-\log_2(2\tau)} \leq \varepsilon$ , the above argument allows to conclude that if we set  $\bar{r} := \tau\bar{\sigma}$  then there is a radius  $r_x \in (\bar{r}, \bar{\sigma})$  such that

$$\theta(x, 2r_x) - \theta(x, r_x) \leq \varepsilon. \quad (4.10)$$

Furthermore, formula (4.8) with  $r_x$  in place of  $\sigma$  implies that

$$r_x^{2-m} \int_{B_{r_x}(x)} |D_V u|^2 \leq C \delta \bar{r}^{2-m} + C \Lambda \bar{\sigma}^2. \quad (4.11)$$

We can finally chose  $\delta = \delta(m, \rho, \Lambda, \varepsilon)$  such that  $C \delta \bar{r}^{2-m} \leq \frac{\varepsilon}{2}$ . From equations (4.10) and (4.11) we infer that  $B_{r_x}(x)$  is  $(k+1, \varepsilon)$ -symmetric for  $u$ .  $\square$

We conclude the section with the following proposition, according to which if the mollified energy is pinched enough at  $(k+1)$  points spanning a  $k$ -plane  $L$ , then it is almost constant along this  $L$ .

**Proposition 4.7.** *Let  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  be a stationary  $Q$ -harmonic map with  $\mathcal{E}(u, B_{10}(0)) \leq \Lambda$ , and let  $r \leq 1$ . Let  $0 < \rho < 1$  and  $\eta > 0$  be fixed. There exists  $\delta_0 = \delta_0(m, \Lambda, \rho, \eta) > 0$  with the following property. Let  $F := \{y \in B_r(0) : \theta(y, 8r) - \theta(y, \rho r) < \delta_0\}$ . If there are points  $y_0, \dots, y_k \in F$  which  $(2\rho \cdot r)$ -effectively span a  $k$ -dimensional affine subspace  $L \subset \mathbb{R}^m$ , then*

$$|\theta(x, \rho r) - \theta(y_j, \rho r)| < \eta \quad \text{for every } x \in L \cap B_r(0), \text{ for every } j \in \{0, \dots, k\}.$$

*Proof.* Due to the scaling invariance of the problem we may without loss of generality assume that  $r = 1$ . We calculate the derivative of  $x \mapsto \theta(x, \rho)$  along a direction  $v$  of the linear part of  $L$ .

$$D_v \theta(x, \rho) = \rho^{2-m} \int \varphi' \left( \frac{|x-y|}{\rho} \right) \frac{(x-y) \cdot v}{\rho |x-y|} |Du|^2(y) dy.$$

If we consider the vector field  $X(y) := \rho^{2-m} \varphi' \left( \frac{|x-y|}{\rho} \right) v$  in the inner variation formula (1.3) we obtain from there

$$\begin{aligned} & \rho^{2-m} \int \varphi' \left( \frac{|x-y|}{\rho} \right) \frac{(x-y) \cdot v}{\rho |x-y|} |Du|^2(y) dy \\ &= -2\rho^{2-m} \int \varphi' \left( \frac{|x-y|}{\rho} \right) D_v u(y) \cdot Du(y) \cdot \frac{(y-x)}{\rho |y-x|} dy \\ &\leq 2 \left( \rho^{-m} \int_{B_\rho(x)} |D_v u|^2 dy \right)^{\frac{1}{2}} \left( \rho^{2-m} \int_{B_\rho(x)} |Du|^2 dy \right)^{\frac{1}{2}}. \end{aligned}$$

Observe that by the monotonicity formula we have  $\rho^{2-m} \int_{B_\rho(x)} |Du|^2 \leq \Lambda$  for all  $x \in B_1$ . Using (4.3) we have that

$$\begin{aligned} & \rho^{-m} \int_{B_\rho(x)} |D_v u|^2 dy \leq \rho^{-m} \int_{B_{1+\rho}(0)} |D_L u|^2 dy \\ &\leq C(m, \rho) \sum_{i=0}^k (\theta(y_i, 4(1+\rho)) - \theta(y_i, 2(1+\rho))) \leq C(m, \rho)(k+1)\delta_0. \end{aligned}$$

Hence we conclude that for all  $x \in B_1$

$$|D_v \theta(x, \rho)| \leq C(m, \rho) ((k+1)\Lambda\delta_0)^{\frac{1}{2}}.$$

Integrating along a path from a point  $y_i$  to any point  $x \in L \cap B_1$  we obtain that

$$|\theta(x, \rho) - \theta(y_i, \rho)| \leq C(m, \rho) ((k+1)\Lambda\delta_0)^{\frac{1}{2}}.$$

Choosing  $\delta_0 > 0$  sufficient small we obtain the proposition.  $\square$

## 5. REIFENBERG THEOREM

This section is dedicated to Reifenberg-type results needed for the proof of the main theorem. The results will only be quoted without proof, and they are in some sense a quantitative generalization of Reifenberg's topological disk theorem (see [Rei60]). Many generalizations of this landmark theorem are available in literature, we limit ourselves to citing [Tor95, DT12] among the various present. Here we will need two versions of this theorem originally proved in [NV17].

Before quoting the theorems, we need the following definition of the so-called Jones'  $\beta_2$  numbers.

**Definition 5.1.** Given a positive Borel measure  $\mu$  defined in  $\mathbb{R}^m$ , for all positive radii  $r > 0$  and dimensions  $k \in \mathbb{N}$ , we define

$$D_\mu^k(x, r) := \min \left\{ \int_{B_r(x)} \frac{\text{dist}^2(y, V)}{r^2} \frac{d\mu(y)}{r^k} : V \subset \mathbb{R}^m \text{ is an affine subspace with } \dim(V) = k \right\}. \quad (5.1)$$

Usually in literature this quantity is referred to as Jones'  $\beta_2$ -number  $\beta_{2,\mu}^k(x, r)^2$ .

$D$  captures in a scale invariant way the distance between the support of  $\mu$  and some  $k$ -dimensional subspace  $V$ . Indeed, the factor  $r^{-2}$  in the distance term makes the integrand scale-invariant, while  $r^{-k}\mu$  takes care of the scaling properties of the measure  $\mu$ .

Here we mention two easy and crucial properties of  $D$ .

**Lemma 5.2** (Bounds on  $D$ ). *Given two measures  $\mu, \mu'$  such that  $\mu' \leq \mu$ , for all  $x, r$  and  $k \in \mathbb{N}$  we can bound*

$$D_{\mu'}^k(x, r) \leq D_\mu^k(x, r). \quad (5.2)$$

Also, for all  $x, y, r$  such that  $|x - y| \leq r$ :

$$D_\mu^k(x, r) \leq 2^{k+2} D_\mu^k(y, 2r). \quad (5.3)$$

*Proof.* The proof follows immediately from the definition.  $\square$

**5.1. Quantitative Reifenberg Theorems.** Assuming a sort of integral Carleson-type condition on the  $D$  numbers, we can obtain uniform scale invariant properties on the measure  $\mu$ . For the reader's convenience, here we quote two key theorems that we are going to use in order to get the final estimates on the singular set of  $Q$ -valued minimizers. The first one is about upper Ahlfors bounds for discrete measures, and is quoted from [NV17, Theorem 3.4]. This theorem is enough for our purposes, but we mention that some generalizations have been obtained in [ENV16]. The second important theorem is about rectifiability properties for general  $\mu$ , and is quoted from [AT15, Theorem 1.1].

**Theorem 5.3.** [NV17, Theorem 3.4] *For some constants  $\delta_R(m)$  and  $C_R(m)$  depending only on the dimension  $m$ , the following holds. Let  $\{B_{r_x/10}(x)\}_{x \in \mathcal{D}} \subseteq B_3(0) \subset \mathbb{R}^m$  be a collection of pairwise disjoint balls with their centers  $x \in B_1(0)$ , and let  $\mu \equiv \sum_{x \in \mathcal{D}} r_x^k \delta_x$  be the associated measure. Assume that for each  $B_r(x) \subseteq B_2$*

$$\int_{B_r(x)} \left( \int_0^r D_\mu^k(y, s) \frac{ds}{s} \right) d\mu(y) < \delta_R^2 r^k. \quad (5.4)$$

*Then, we have the uniform estimate*

$$\sum_{x \in \mathcal{D}} r_x^k < C_R(m). \quad (5.5)$$

Condition (5.4) prescribes some integral Carleson-type control over the quantity  $D(x, r)$ . If the measure  $\mu$  is the  $k$ -dimensional Hausdorff measure restricted to some  $S$ , this bound is enough to guarantee also the rectifiability of  $S$ , as seen in the following theorem. Note that in [AT15] the theorem is presented in a more general form.

**Theorem 5.4** ([AT15, Corollary 1.3]). *Given a Borel measurable subset  $S$  of  $\mathbb{R}^m$ , let  $\mu := \mathcal{H}^k \llcorner S$  be the  $k$ -dimensional Hausdorff measure restricted to  $S$ . The set  $S$  is countably  $k$ -rectifiable if and only if*

$$\int_0^1 D_\mu^k(x, s) \frac{ds}{s} < \infty \quad \text{for } \mu\text{-a.e. } x. \quad (5.6)$$

## 6. BEST APPROXIMATING PLANE

In this section, we record the main technical lemma needed for the final proof of Theorem 3.18. Although several technical points need to be addressed, this lemma contains most of the important estimates in the paper and provides an estimate on the  $D$  numbers using the normalized energy  $\theta(x, r)$ .

The basic ideas behind the estimates in this section are similar to the ones in [NV17, Theorem 7.1], however the new definition of  $(k, \varepsilon)$ -symmetries allows for more quantitative and easier proofs.

For any  $f \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$ , and for all  $B_r(x) \subseteq \Omega$ , we introduce the following quantity

$$\mathcal{P}_f(x, r) := r^{-m} \int_{B_r(x)} |Df(y) \cdot (y - x)|^2 dy. \quad (6.1)$$

Note that in the case  $u$  is an energy minimizing  $Q$ -valued harmonic map we have by (4.4):

$$\mathcal{P}_u(x, r) \leq C(m) [\theta(x, 2r) - \theta(x, r)]. \quad (6.2)$$

However, here we carry out the estimates in a very general setting, and we will exploit this bound only at the very last step in our main proof.

**Theorem 6.1.** *Let  $u \in W^{1,2}(B_2(0), \mathcal{A}_Q(\mathcal{N}))$ , and fix  $\varepsilon > 0$ ,  $0 < r \leq 1$  and some  $x \in B_1(0)$ . Let also  $\mu$  be any positive Radon measure supported on  $B_1(0)$ . Assuming that*

$$\inf \left\{ r^{2-m} \int_{B_r(x)} |D_V u|^2 : V \subset \mathbb{R}^m \text{ linear with } \dim(V) = k + 1 \right\} \geq \varepsilon, \quad (6.3)$$

we conclude

$$D_\mu^k(x, r) \leq \frac{(m-k)(k+1)2^m}{\varepsilon r^k} \int_{B_r(x)} \mathcal{P}_u(y, 2r) d\mu(y). \quad (6.4)$$

**Remark 6.2.** We remark that (6.4) does not change if  $\mu$  is multiplied by a positive constant, thus for convenience for the rest of this section we are going to assume without loss of generality that  $\mu$  is a probability measure. Moreover, we can also assume without loss of generality that  $x = 0$  and  $r = 1$ .

Note that for this theorem we will not exploit any property specific to energy minimizers. For future convenience, we record a simple corollary that rephrases the previous theorem with the language of energy minimizers and quantitative stratification.

**Corollary 6.3.** *For any  $Q$ -valued Dirichlet minimizing  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  with  $\mathcal{E}(u, B_{10}(0)) \leq \Lambda$ , and for all  $\varepsilon > 0$ ,  $0 < r \leq 1$  and  $x \in B_1(0)$ , the following holds. Let  $\mu$  be any positive Radon measure supported on  $B_1(0)$ . Assuming that  $B_r(x)$  is  $(k, \varepsilon)$ -symmetric but NOT  $(k+1, \varepsilon)$ -symmetric, we conclude*

$$D_\mu^k(x, r) \leq \frac{C(m)}{\varepsilon r^k} \int_{B_r(x)} [\theta(y, 4r) - \theta(y, 2r)] d\mu(y). \quad (6.5)$$

*Proof.* The proof follows immediately from the definition of  $(k + 1, \varepsilon)$ -symmetry and the bound in (6.2).  $\square$

**6.1. Properties of the best approximating plane.** For fixed  $k$ , and given any probability measure  $\mu$ , for all  $(x, r)$  we set  $V(x, r)$  to be the  $k$ -dimensional affine subspace minimizing

$$\int_{B_r(x)} \text{dist}^2(y, V) \, d\mu(y), \quad (6.6)$$

so that, in particular,

$$D_\mu^k(x, r) = r^{-(k+2)} \int_{B_r(x)} \text{dist}^2(y, V(x, r)) \, d\mu(y). \quad (6.7)$$

Since in this section we focus on  $x = 0$  and  $r = 1$ , we will in fact mostly consider only the  $k$ -dimensional subspace  $V(0, 1)$ .

First of all, note that necessarily  $V(x, r)$  will pass through the center of mass of  $\mu$  in  $B_r(x)$ , defined as

$$x_m(\mu, x, r) = x_m := \int_{B_r(x)} x \, d\mu(x). \quad (6.8)$$

It will be convenient to phrase some of the estimates needed for theorem 6.1 in terms of a suitable quadratic form on  $\mathbb{R}^m$ , defined as

$$R(w) := \int_{B_1(0)} |\langle x - x_m, w \rangle|^2 \, d\mu(x). \quad (6.9)$$

By standard linear algebra, there exists an orthonormal basis  $\{e_1, \dots, e_m\}$  of eigenvectors for  $R$  with non-negative eigenvalues  $\lambda_1, \dots, \lambda_m$ , which we will take for convenience in decreasing order. Note that by the variational characterization of  $\lambda_k$  we have that

$$e_k \in \operatorname{argmax} \left\{ \int_{B_1(0)} |\langle x - x_m, e \rangle|^2 \, d\mu(x) \text{ s.t. } |e|^2 = 1 \text{ and } \langle e, e_i \rangle = 0 \, \forall i \leq k \right\}, \quad (6.10)$$

$$\lambda_k = \int_{B_1(0)} |\langle x - x_m, e_k \rangle|^2 \, d\mu(x), \quad (6.11)$$

and so

$$D_\mu^k(0, 1) = \int_{B_1(0)} \text{dist}^2(x, V(0, 1)) \, d\mu(x) = \sum_{i=k+1}^m \lambda_i. \quad (6.12)$$

Indeed, by minimality of  $V$ ,  $V(0, 1) = x_m + \operatorname{span}[e_1, \dots, e_k]$ , and thus

$$\int_{B_1(0)} \text{dist}^2(x, V(0, 1)) \, d\mu(x) = \sum_{i=k+1}^m \int_{B_1(0)} |\langle x - x_m, e_i \rangle|^2 \, d\mu(x) = \sum_{i=k+1}^m \lambda_i. \quad (6.13)$$

Using simple geometry, it is possible to prove that for *any* map  $f \in W^{1,2}$  we have the following estimate involving  $\lambda_k$  and  $\mathcal{P}_f$ .

**Lemma 6.4.** *Let  $f = \sum_{\ell=1}^Q \llbracket f_\ell \rrbracket \in W^{1,2}(B_{3r}(x), \mathcal{A}_Q(\mathcal{N}))$ , and let  $\mu$  be a probability measure on  $B_r(x)$ . Then*

$$\lambda_k \int_{B_r(x)} |Df(z) \cdot e_k|^2 \, dz \leq 2^m \int_{B_r(x)} \mathcal{P}_f(y, 2r) \, d\mu(y) \quad \text{for every } k = 1, \dots, m. \quad (6.14)$$

*Proof.* For simplicity, we assume  $x = 0$  and  $r = 1$ . Moreover, note that evidently we can assume  $\lambda_k > 0$ , otherwise there is nothing to prove. Fix some  $z \in B_1(0)$ . By definition of eigenvectors  $e_k$ , we have for every  $\ell \in \{1, \dots, Q\}$  that

$$\int_{B_1(0)} \langle x - x_m, e_k \rangle (Df_\ell(z) \cdot (x - x_m)) \, d\mu(x) = \lambda_k Df_\ell(z) \cdot e_k. \quad (6.15)$$

By definition of center of mass, we can write

$$\int_{B_1(0)} \langle x - x_m, e_k \rangle (z - x_m) \, d\mu(x) = 0, \quad (6.16)$$

and so

$$\lambda_k Df_\ell(z) \cdot e_k = \int_{B_1(0)} \langle x - x_m, e_k \rangle (Df_\ell(z) \cdot (x - z)) \, d\mu(x). \quad (6.17)$$

By Cauchy-Schwartz and by (6.11), we have

$$\lambda_k^2 |Df_\ell(z) \cdot e_k|^2 \leq \lambda_k \int |Df_\ell(z) \cdot (x - z)|^2 \, d\mu(x), \quad (6.18)$$

and thus, summing over  $\ell$ ,

$$\lambda_k |Df(z) \cdot e_k|^2 \leq \int |Df(z) \cdot (x - z)|^2 \, d\mu(x). \quad (6.19)$$

Taking the integral of this inequality in  $B_1(0)$  with respect to the volume measure in  $z$ , we obtain the estimate

$$\begin{aligned} \lambda_k \int_{B_1(0)} |Df(z) \cdot e_k|^2 \, dz &\leq \iint_{B_1(0) \times B_1(0)} |Df(z) \cdot (x - z)|^2 \, dz \, d\mu(x) \\ &\leq \int_{B_1(0)} \int_{B_2(x)} |Df(z) \cdot (x - z)|^2 \, dz \, d\mu(x) \leq 2^m \int_{B_1(0)} \mathcal{P}_f(x, 2) \, d\mu(x). \end{aligned} \quad (6.20)$$

□

From this proposition, the proof of Theorem 6.1 follows as a simple corollary.

*Proof of theorem 6.1.* As before, we assume without loss of generality that  $x = 0$  and  $r = 1$ . Moreover, by (6.12) it is sufficient to prove that

$$\lambda_{k+1} \leq \frac{C(m)}{\varepsilon} \int_{B_1(0)} \mathcal{P}_u(y, 2) \, d\mu(y). \quad (6.21)$$

By the previous lemma, we have

$$\lambda_{k+1} \sum_{j=1}^{k+1} \int_{B_1(0)} |Du \cdot e_j|^2 \leq \sum_{j=1}^{k+1} \lambda_j \int_{B_1(0)} |Du \cdot e_j|^2 \leq C(m) \int_{B_1(0)} \mathcal{P}_u(x, 2) \, d\mu(x). \quad (6.22)$$

By the lower bound in (6.3), we must have

$$\sum_{j=1}^{k+1} \int_{B_1(0)} |Du \cdot e_j|^2 \geq \varepsilon, \quad (6.23)$$

and this concludes the proof. □

## 7. PROOF OF THE MAIN THEOREM VIA COVERING ARGUMENTS

This section is dedicated to the proof of the Theorem 3.18. We split it into two pieces, one containing the uniform Minkowski bounds and one with the rectifiability part. Once the Minkowski bounds are obtained, the rectifiability is almost an immediate corollary.

The Minkowski bounds will be obtained with a covering argument similar to the one in [NV16].

**Proposition 7.1** (Fine covering). *There exist a small constant  $\delta = \delta(m, \mathcal{N}, Q, \Lambda, \varepsilon) > 0$  and  $C_{III}(m)$  such that the following holds. Let  $u$  be a  $Q$ -valued Dirichlet minimizing map  $u \in W^{1,2}(B_{10}(0), \mathcal{A}_Q(\mathcal{N}))$  with  $\mathcal{E}(u, B_{10}(0)) \leq \Lambda$ , let  $\varepsilon > 0$ ,  $p \in B_1(0)$ , and  $0 < r \leq R$ ,  $0 < R \leq 1$  be chosen in an arbitrary fashion. For any subset  $\mathcal{S} \subseteq \mathcal{S}_{\varepsilon, \delta r}^k(u)$ , setting  $E = \sup_{x \in B_{2R}(p) \cap \mathcal{S}} \theta(x, 3R)$ , there exists a covering*

$$\mathcal{S} \cap B_R(p) \subseteq \bigcup_{x \in \mathcal{D}} B_{r_x}(x), \quad \text{with } r_x \geq r \quad \text{and} \quad \sum_{x \in \mathcal{D}} r_x^k \leq 2C_{III}(m)R^k. \quad (7.1)$$

Moreover, for all  $x \in \mathcal{D}$ , either  $r_x = r$ , or for all  $y \in B_{2r_x}(x)$ :

$$\theta(y, 3r_x) \leq E - \delta. \quad (7.2)$$

**Remark 7.2.** This fine covering will be proved as a corollary of the ‘‘intermediate covering’’ stated in proposition 7.3.

**7.1. Proof of the main theorem 3.18 given proposition 7.1.** Before we move to the proof of the proposition, we use it to prove the main theorem. This proof is basically a corollary of the covering proposition 7.1. We will use this proposition inductively to produce a family of coverings of

$$\mathcal{S} := \mathcal{S}_{\varepsilon, \delta r}^k(u) \cap B_1(0) \quad (7.3)$$

indexed by a parameter  $i \in \mathbb{N}$  of the form

$$\mathcal{S} \subseteq \bigcup_{x \in \mathcal{D}^i} B_{r_x}(x), \quad \sum_{x \in \mathcal{D}^i} r_x^k \leq (c(m)C_{III}(m))^i. \quad (7.4)$$

Moreover, if  $E = \sup_{x \in \mathcal{S}_{\varepsilon, \delta r}^k(u) \cap B_2(0)} \theta(x, 3)$ , we have for all  $i$

$$r_x \leq r \quad \text{or} \quad \forall y \in \mathcal{S} \cap B_{2r_x}(x), \quad \theta(y, 3r_x) \leq E - i\delta. \quad (7.5)$$

Evidently, for  $i \geq \lceil E/\delta \rceil + 1$ , the second condition cannot be verified, and so all the radii in the covering are going to be equal to  $r$ . As a consequence, we have the Minkowski bound

$$\text{Vol} \left( B_r \left( \mathcal{S}_{\varepsilon, \delta r}^k(u) \right) \cap B_1(0) \right) \leq (c(m)C_{III}(m))^{\lceil E/\delta \rceil + 1} r^{m-k}. \quad (7.6)$$

Since  $\delta = \delta(m, \varepsilon, \Lambda, Q, \mathcal{N})$ , it is clear that, up to enlarging the constant in the estimate, the same bound holds also for  $\mathcal{S}_{\varepsilon, r}^k(u)$  in the place of  $\mathcal{S}_{\varepsilon, \delta r}^k(u)$ , and this concludes the proof of the Minkowski bounds in (3.18).

In order to produce the covering in (7.4), we will apply inductively the covering proposition 7.1. For  $i = 1$ , we can apply this proposition to  $B_1(0)$  and obtain the desired covering. Inductively, consider all the balls  $\{B_{r_x}(x)\}_{x \in \mathcal{D}^i}$  and apply proposition 7.1 to these balls. For each  $x \in \mathcal{D}^i$ , we obtain a covering of the form

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \mathcal{D}_x} B_{r_y}(y), \quad \sum_{y \in \mathcal{D}_x} r_y^k \leq 2C_{III}(m)r_x^k, \quad (7.7)$$

$$r_y \leq r \quad \text{or} \quad \forall z \in \mathcal{S} \cap B_{2r_y}(y), \quad \theta(z, 3r_y) \leq E - (i+1)\delta. \quad (7.8)$$

Set

$$\mathcal{D}^{i+1} = \bigcup_{x \in \mathcal{D}^i} \mathcal{D}_x, \quad (7.9)$$

and the induction step is completed.

7.1.1. *Proof of the rectifiability of  $\mathcal{S}_\varepsilon^k$ .* As for the rectifiability, this is going to be a corollary of Theorem 5.4, the uniform Minkowski bound (3.18) and the approximation theorem 6.1.

In particular, let  $\mu = \mathcal{H}^k \llcorner \left\{ \mathcal{S}_\varepsilon^k(u) \cap B_1(0) \right\}$ . From (3.18) we deduce that this measure is finite, as

$$\mu(B_1(0)) \leq C(m, \Lambda, \varepsilon).$$

In turn, by scaling this implies that for all  $x \in B_1(0)$  and  $r > 0$

$$\mu(B_r(x)) \leq C(m, \Lambda, \varepsilon)r^k, \quad (7.10)$$

and thus  $\mu$  is Ahlfors upper  $k$ -regular.

We claim that there exists a constant  $c(m, \varepsilon, \Lambda) > 0$  such that for all  $x \in \mathcal{S}_\varepsilon^k$ , all  $r > 0$  and  $(k+1)$ -dimensional subspaces  $V$ :

$$r^{2-m} \int_{B_r(x)} |D_V u|^2 \geq c(m, \varepsilon, \Lambda). \quad (7.11)$$

In order to prove this, fix any  $r \in (0, 1]$ , and consider the sequence of radii  $r_i = r2^{-i}$  for  $i = 0, 1, \dots$ . Since

$$\sum_{i=0}^{\infty} [\theta_u(x, r_i) - \theta_u(x, r_{i+1})] \leq C\Lambda, \quad (7.12)$$

and since all terms in the sum are nonnegative by monotonicity, at most  $C\Lambda\varepsilon^{-1}$  of these terms can be bigger than  $\varepsilon$ . In particular, this implies that for all  $x \in \mathcal{S}_\varepsilon^k$  and  $r > 0$ , there exists  $r'(r, x) \geq 2^{-C\Lambda\varepsilon^{-1}}r$  such that  $\theta(x, r') - \theta(x, r'/2) < \varepsilon$ . By definition of  $\mathcal{S}_\varepsilon^k$  and some simple estimates, this implies that for all  $(k+1)$ -dimensional subspaces  $V$  we have

$$r^{2-m} \int_{B_r(x)} |D_V u|^2 \geq \frac{r^{2-m}}{(r'(r, x))^{2-n}} (r'(r, x))^{2-m} \int_{B_{r'(r, x)}(x)} |D_V u|^2 \geq \frac{r^{2-m}}{(r'(r, x))^{2-m}} \varepsilon \geq c(m, \varepsilon, \Lambda). \quad (7.13)$$

With this lower bound, we can apply the best approximation theorem 6.1, and with a simple change of variables we can write

$$\begin{aligned} \int_{B_1(0)} \int_0^1 D_\mu^k(x, r) \frac{dr}{r} d\mu(x) &\leq C(m, \varepsilon, \Lambda) \int_0^1 \int_{B_1(0)} r^{-k} \int_{B_r(x)} \mathcal{P}_u(y, 2r) d\mu(y) d\mu(x) \frac{dr}{r} \\ &\leq C(m, \varepsilon, \Lambda) \int_0^1 \int_{B_1(0)} \mathcal{P}_u(y, 2r) \left( r^{-k} \int_{B_r(y)} d\mu(x) \right) d\mu(y) \frac{dr}{r} \\ &\leq C(m, \varepsilon, \Lambda) \int_{B_1(0)} \int_0^1 \mathcal{P}_u(y, 2r) \frac{dr}{r} d\mu(x) \\ &\leq C(m, \varepsilon, \Lambda)\Lambda, \end{aligned} \quad (7.14)$$

where the last inequality follows from

$$\begin{aligned} \int_0^1 \mathcal{P}_u(y, 2r) \frac{dr}{r} &\stackrel{(6.2)}{\leq} \int_0^1 [\theta(y, 4r) - \theta(y, 2r)] \frac{dr}{r} = \lim_{t \rightarrow 0} \int_t^1 [\theta(y, 4r) - \theta(y, 2r)] \frac{dr}{r} \\ &= \int_{1/2}^1 \theta(y, 4r) \frac{dr}{r} + \underbrace{\lim_{t \rightarrow 0} \int_t^{1/2} \theta(y, 4r) \frac{dr}{r} - \int_{2t}^1 \theta(y, 2r) \frac{dr}{r}}_{=0} - \lim_{t \rightarrow 0} \int_t^{2t} \theta(y, 2r) \frac{dr}{r} \leq C(m)\Lambda. \end{aligned} \quad (7.15)$$

The rectifiability of  $\mathcal{S}_\varepsilon^k(u)$  is now a consequence of theorem 5.4.

By countable additivity, the rectifiability of  $\mathcal{S}^k(u)$  is a corollary of the rectifiability of  $\mathcal{S}_\varepsilon^k(u)$  for all  $\varepsilon > 0$ .

It is worth remarking that the uniform Ahlfors upper estimates obtained a priori for the measure  $\mu = \mathcal{H}^k \llcorner \{\mathcal{S}_\varepsilon^k(u) \cap B_1(0)\}$  are *essential* to carry out this computation, and actually they are the most difficult part of the estimate. This is why the proof of the rectifiability property is so easy.  $\square$

**7.2. Proof of Proposition 7.1.** Now we turn to the proof of the covering proposition. We split this proof in two pieces by introducing a secondary covering proposition.

**Proposition 7.3** (Intermediate covering). *Under the assumptions of proposition 7.1, for all  $0 < \rho < 1/100$ , there exist  $\delta = \delta(m, \mathcal{N}, Q, \Lambda, \varepsilon, \rho) > 0$  and  $C_{II}(m)$  such that the following is true.*

*There exists a finite covering of  $\mathcal{S} = \mathcal{S}_{\varepsilon, \delta r}^k(u) \cap B_R(p)$  of the form*

$$\mathcal{S} \subseteq \bigcup_{x \in \mathcal{D}} B_{r_x}(x), \quad \text{with } r_x \geq r \text{ and } \sum_{x \in \mathcal{D}} r_x^k \leq C_{II}(m)R^k. \quad (7.16)$$

Moreover, for each  $x \in \mathcal{D}$ , either there exists a  $(k-1)$ -dimensional space  $W_x$  such that

$$F_{x, r_x} \equiv \{y \in \mathcal{S} \cap B_{2r_x}(x) \text{ with } \theta(y, \rho r_x/20) \geq E - \delta\} \subseteq B_{\rho r_x/10}(W_x), \quad (7.17)$$

or  $r_x = r$ .

Assuming this proposition, we prove proposition 7.1. The idea is simple: we consider this second covering, and refine it inductively on each ball with  $r_x \geq r$  and no uniform energy drop, so basically on the sets  $F_{x, r_x}$ . Given the last characterization of this set, we only need to refine the covering around a subspace of a dimension  $m-1$ , which allows us to keep uniform  $m$  dimensional covering estimates on the original set.

*Proof of proposition 7.1 given proposition 7.3.* Let  $0 < \rho < 1$  to be fixed later, and let  $A \in \mathbb{N}$  be the first integer such that  $\rho^A < r$ . Also assume without loss of generality  $p = 0$  and  $R = 1$ .

For all  $i = 1, \dots, A$ , we construct a covering of  $\mathcal{S}$  of the form

$$\mathcal{S} \cap B_1(0) \subseteq \bigcup_{x \in \mathcal{R}_i} B_r(x) \cup \bigcup_{x \in \mathcal{F}_i} B_{r_x}(x) \cup \bigcup_{x \in \mathcal{B}_i} B_{r_x}(x), \quad (7.18)$$

where  $\mathcal{R}_i$  are the balls of radius  $r$  in the covering,  $\mathcal{F}_i$  are the balls where the uniform energy drop condition (7.2) is satisfied, and  $\mathcal{B}_i$  are the bad balls, where none of the two conditions is verified. We want to obtain uniform packing bounds on  $\mathcal{R}_i$  and  $\mathcal{F}_i$ , and exponentially small packing bounds on  $\mathcal{B}_i$ . We will refine our covering only on bad balls by re-applying the second covering lemma on those, and this is why we need smallness on their packing bounds. In detail, we want

$$\sum_{x \in \mathcal{R}_i \cup \mathcal{F}_i} r_x^k \leq C_{III}(m) \left( \sum_{j=0}^i 7^{-j} \right), \quad \sum_{x \in \mathcal{B}_i} r_x^k \leq 7^{-i}. \quad (7.19)$$

7.2.1. *Induction step.* In this step, we consider a generic ball  $B_R(p)$  and we want to obtain a covering of  $\mathcal{S} \cap B_R(p)$  of the form

$$\mathcal{S} \cap B_R(p) \subseteq \bigcup_{x \in \mathcal{R}_p \cup \mathcal{F}_p} B_{r_x}(x) \cup \bigcup_{x \in \mathcal{B}_p} B_{r_x}(x) \quad (7.20)$$

such that  $r_x \geq r$  and for all  $x \in \mathcal{R}_p$ ,  $r_x = r$ . Moreover, we want to obtain the covering estimates

$$\sum_{z \in \mathcal{R}_p \cup \mathcal{F}_p} r_z^k \leq C_{III}(m)R^k \quad \sum_{z \in \mathcal{B}_p} r_z^k \leq 7^{-1}R^k, \quad (7.21)$$

and we also want to make sure that the balls  $\{B_{r_x}(x)\}_{x \in \mathcal{F}_p}$  have a definite energy drop, in particular we will prove that for all  $y \in \mathcal{F}_x$  and  $z \in B_{2 \cdot (\rho r_x / 60)}(y) = B_{2r_y}(y)$ ,

$$\theta(z, 3r_y) < E - \delta. \quad (7.22)$$

In order to do so, we first apply the second covering in Proposition 7.3 to the ball  $B_R(p)$ . We obtain a covering of the form

$$\mathcal{S} \cap B_R(p) \subseteq \bigcup_{x \in \mathcal{D}} B_{r_x}(x), \quad \text{with } r_x \geq r \text{ and } \sum_{x \in \mathcal{D}} r_x^k \leq C_{II}(m)R^k. \quad (7.23)$$

We split  $\mathcal{D}$  into two disjoint sets:  $\mathcal{D} = \mathcal{D}_r \cup \mathcal{D}_+$ , where  $x \in \mathcal{D}_r$  if  $r_x \leq 60\rho^{-1}r$ , and otherwise  $x \in \mathcal{D}_+$ . Observe that if  $x \in \mathcal{D}_+$  then (7.17) is valid.

We refine slightly this covering by recovering each of the balls in  $\mathcal{D}_r$  and  $\mathcal{D}_+$  separately. In particular, for all  $x \in \mathcal{D}_r$ , consider a simple covering of  $B_{r_x}(x)$  by balls of radius  $r$  with number bounded by  $c(m)\rho^{-m}$ , and let  $\mathcal{R}_p$  be the union of all centers in these coverings. Note that if  $r_x = r$ , we can keep  $B_{r_x}(x)$  unchanged.

For all  $x \in \mathcal{D}_+$ , consider a covering of  $B_{r_x}(x)$  made of balls of radius  $\rho r_x / 60 > r$  centered inside this ball and such that the family of balls with half the radius are pairwise disjoint. In particular, let

$$B_{r_x}(x) \subseteq \bigcup_{y \in \mathcal{B}_x} B_{\rho r_x / 60}(y) \cup \bigcup_{y \in \mathcal{F}_x} B_{\rho r_x / 60}(y), \quad (7.24)$$

where

$$F_{x,r_x} \cap \bigcup_{y \in \mathcal{F}_x} B_{2 \cdot (\rho r_x / 60)}(y) = \emptyset, \quad \mathcal{B}_x \subseteq B_{\rho r_x}(W_x). \quad (7.25)$$

Thus, the balls in  $\mathcal{F}_x$  will have a uniform energy drop, in particular we have that for all  $y \in \mathcal{F}_x$  and  $z \in B_{2 \cdot (\rho r_x / 60)}(y) = B_{2r_y}(y)$ ,

$$\theta(z, 3r_y) < E - \delta. \quad (7.26)$$

Moreover, the number of balls in  $\mathcal{B}_x$  is well controlled. Indeed, since  $\mathcal{B}_x \subseteq B_{\rho r_x}(W_x)$ ,  $\{B_{\rho r_x / 120}(y)\}_{y \in \mathcal{B}_x}$  are pairwise disjoint and  $W_x$  is a  $k$ -dimensional subspace, then

$$\#\{\mathcal{F}_x\} \leq c(m)\rho^{-m}, \quad \#\{\mathcal{B}_x\} \leq c(m)\rho^{1-k}. \quad (7.27)$$

Set  $\mathcal{B}_p = \bigcup_{x \in \mathcal{D}_+} \mathcal{B}_x$  and  $\mathcal{F}_p = \bigcup_{x \in \mathcal{D}_+} \mathcal{F}_x$ . We have

$$\sum_{z \in \mathcal{R}_p \cup \mathcal{F}_p} r_z^k \leq c(m)\rho^{-m+k} \sum_{x \in \mathcal{D}} r_x^k \leq c(m)\rho^{-m+k} C_{II}(m)R^k, \quad (7.28)$$

$$\sum_{z \in \mathcal{B}_p} r_z^k \leq c(m)\rho^1 \sum_{x \in \mathcal{D}} r_x^k \leq c(m)\rho C_{II}(m)R^k. \quad (7.29)$$

We choose  $\rho = \rho(m) \leq 1/100$  sufficiently small so that

$$c(m)\rho C_{II}(m) \leq 1/7. \quad (7.30)$$

In this way, we have as desired the estimates

$$\sum_{z \in \mathcal{R}_p \cup \mathcal{F}_p} r_z^k \leq C_{III}(m)R^k \quad \sum_{z \in \mathcal{B}_p} r_z^k \leq 7^{-1}R^k, \quad (7.31)$$

where  $C_{III}(m) = c(m)\rho(m)^{-m+k}C_{II}(m)$ .

**7.2.2. Finishing the proof.** With the induction step, the proof follows easily. For  $i = 1$ , apply the induction step to  $B_1(0)$  and we obtain (7.18) for  $i = 1$  with (7.19). In particular, we obtain the covering given in (7.20).

For generic  $i$ , we have by induction

$$\mathcal{S} \cap B_1(0) \subseteq \bigcup_{x \in \mathcal{R}_i} B_r(x) \cup \bigcup_{x \in \mathcal{F}_i} B_{r_x}(x) \cup \bigcup_{x \in \mathcal{B}_i} B_{r_x}(x). \quad (7.32)$$

Apply the induction step on all the balls  $\{B_{r_x}(x)\}_{x \in \mathcal{B}_i}$  separately, and define

$$\mathcal{R}_{i+1} = \mathcal{R}_i \cup \bigcup_{x \in \mathcal{B}_i} \mathcal{R}_x, \quad \mathcal{F}_{i+1} = \mathcal{F}_i \cup \bigcup_{x \in \mathcal{B}_i} \mathcal{F}_x, \quad \mathcal{B}_{i+1} = \bigcup_{x \in \mathcal{B}_i} \mathcal{B}_x. \quad (7.33)$$

By construction, we have the estimates

$$\sum_{z \in \mathcal{R}_{i+1} \cup \mathcal{F}_{i+1}} r_z^k \leq C_{III}(m) \sum_{s=0}^i 7^{-s} \sum_{z \in \mathcal{B}_{i+1}} r_z^k \leq 7^{-i-1}. \quad (7.34)$$

Note that at the step  $i = A$  all the balls in our covering will either have energy drop (if they are in  $\mathcal{F}_A$ ) or have radius  $= r$  (if they are in  $\mathcal{R}_A$ ). Equation (7.34) for  $i = A$  gives the desired bound on the final covering.  $\square$

Now we turn our attention to the proof of proposition 7.3, which is the last one needed to complete the main theorem.

**7.3. Proof of proposition 7.3.** For convenience, we assume  $p = 0$  and  $R = 1$ . Fix  $\varepsilon, \rho > 0$ , and let  $A$  be such that  $\rho^A \leq r < \rho^{A-1}$ . The proof is based on an inductive covering by balls, where the discrete Reifenberg is applied in order to control the number of these balls.

**7.3.1. Construction of the covering.** We split the inductive covering in two parts: at first we simply construct the covering inductively, and then we prove the packing bounds using Reifenberg's theorem. Specifically, we start by looking for an inductive (for  $i = 0, 1, \dots, A$ ) covering for  $\mathcal{S} = \mathcal{S}_{\varepsilon, \delta r}^k(u) \cap B_R(p)$  of the form

$$\mathcal{S} \subseteq \bigcup_{x \in \mathcal{B}_i} B_{r_x}(x) \cup \bigcup_{x \in \mathcal{G}_i} B_{r_x}(x), \quad (7.35)$$

where the elements of  $\mathcal{B}_i$  are the centers of the *bad balls* in our covering, and  $\mathcal{G}_i$  are the centers of the *good balls*. In particular, we want :

- (1) for all  $i$  and  $x \in \mathcal{B}_i$ ,  $r_x \geq \rho^i$  and there exists a  $(k-1)$ -dimensional subspace  $W_x$  such that
$$F_{x, r_x} \equiv \{y \in \mathcal{S} \cap B_{2r_x}(x) \text{ such that } \theta(y, \rho r_x/20) \geq E - \delta\} \subseteq B_{\rho r_x/10}(W_x), \quad (7.36)$$

where  $\delta > 0$  is fixed, to be determined later;

- (2) for all  $i = 1, \dots, A$  and  $x \in \mathcal{G}_i$ ,  $r_x = \rho^i$  and the set  $F_{x, r_x}$  defined above  $(\rho r_x/20)$ -effectively spans some  $k$ -dimensional affine subspace  $V_x$ ;
- (3) for  $i = A$ , we have the bound

$$\sum_{x \in \mathcal{B}_A \cup \mathcal{G}_A} r_x^k \leq C_{II}(m). \quad (7.37)$$

Moreover, we request some extra properties of the centers of the covering in order to apply the discrete-Reifenberg theorem:

- (4) for all  $i$ , the balls in the collection  $\{B_{r_x/10}(x)\}_{x \in \mathcal{G}_i \cup \mathcal{B}_i}$  are pairwise disjoint;
- (5) for all  $i \geq 1$  and  $x \in \mathcal{G}_i \cup \mathcal{B}_i$ , we have the energy bound

$$\theta(x, r_x) \geq E - \eta \quad \text{for some } \eta > 0 \text{ to be determined}; \quad (7.38)$$

- (6) there exists a constant  $c_L(m, \varepsilon, \Lambda)$  such that for all  $i$ ,  $x \in \mathcal{G}_i \cup \mathcal{B}_i$  and  $s \in [r_x, 1]$ , and for all  $(k+1)$ -dimensional subspaces  $T$ , we have

$$s^{2-m} \int_{B_s(x)} |D_T u|^2 \geq C_T(m, \Lambda, \varepsilon). \quad (7.39)$$

At each induction step, we will refine our covering on the good balls, while leaving the bad balls untouched.

For  $i = 0$ , consider the set  $F_{0,1}$ . If this set does **not**  $\rho/20$ -effectively span something  $k$ -dimensional, then we call  $B_1(0)$  a bad ball, set  $\mathcal{G}_i = \emptyset$  for all  $i$  and  $\{0\} = \mathcal{B}_0 = \mathcal{B}_A$  with  $r_0 = 1$ . This covering immediately satisfies all the properties of proposition 7.3, and we stop our construction here.

In the other case, set  $\mathcal{G}_0 = \{0\}$  with  $r_0 = 1$ , and we proceed with our construction with this as our base case.

**Induction step.** Assuming by induction that all the properties listed above are valid up to the index  $i$ , we want to produce the covering for  $i+1$ . In order to do so, we want to refine our covering on good balls, and leave the previous bad balls intact.

Fix an arbitrary  $x \in \mathcal{G}_i$ , and consider the set  $F_{x, r_x}$ . Since  $B_{\rho^i}(x)$  is a good ball, by definition this set  $[\rho^{i+1}/20]$ -effectively spans a  $k$ -dimensional affine subspace  $V_x$ . By applying theorem 4.3 to the ball  $B_{4\rho^i}(x)$ , we find that there exists a  $\delta(m, \Lambda, \varepsilon, \rho)$  sufficiently small so that

$$\mathcal{S}_{\varepsilon, \delta r}^k(u) \cap B_{2\rho^i}(x) \subset B_{\rho^{i+1}/10}(V_x) \quad (7.40)$$

Consider the set

$$K = \bigcup_{x \in \mathcal{G}_i} (B_{\rho^i}(x) \cap V_x) \setminus \bigcup_{x \in \mathcal{B}_i} B_{4r_x/5}(x). \quad (7.41)$$

Given the inclusion (7.40), and since we have chosen  $\rho \leq 1/100$ , we have

$$\mathcal{S} \setminus \bigcup_{x \in \mathcal{B}_i} B_{r_x}(x) \subseteq B_{\rho^{i+1}/5}(K). \quad (7.42)$$

Indeed, if

Let  $\mathcal{D}_K \subseteq K$  be a maximal subset of points at least  $\rho^{i+1}/5$  apart, so that the balls  $\{B_{\rho^{i+1}/10}(x)\}_{x \in \mathcal{D}_K}$  are pairwise disjoint. Note that these balls are also disjoint from  $\{B_{r_x/3}(x)\}_{x \in \mathcal{B}_i}$  by construction. Moreover, by maximality of the subset  $\mathcal{D}_K$ , we have the inclusion

$$\mathcal{S} \setminus \bigcup_{x \in \mathcal{B}_i} B_{r_x}(x) \subseteq \bigcup_{x \in \mathcal{D}_K} B_{2\rho^{i+1}/5}(x) \quad (7.43)$$

We can discard from this collection all the balls  $B_{2\rho^{i+1}/5}(x)$  that have empty intersection with  $\mathcal{S}$ . Now consider the collection

$$\left\{ B_{\rho^{i+1}}(x) \right\}_{x \in \mathcal{D}_K}, \quad (7.44)$$

and classify these points into good and bad balls according to whether or not (7.36) is satisfied. In particular, if  $F_{x, \rho^{i+1}}(\rho^{i+2}/20)$ -effectively spans a  $k$ -dimensional subspace  $V_x$ , then we say that  $x \in \tilde{\mathcal{G}}_{i+1}$ , and  $x \in \tilde{\mathcal{B}}_{i+1}$  otherwise. We set

$$\mathcal{B}_{i+1} = \mathcal{B}_i \cup \tilde{\mathcal{B}}_{i+1}, \quad \mathcal{G}_{i+1} = \tilde{\mathcal{G}}_{i+1}. \quad (7.45)$$

This takes care of properties 1 and 2 in the induction.

Now fix any  $x \in \mathcal{D}_K$ . By construction, there exists an  $x' \in \mathcal{G}_i$  such that  $x \in V_{x'} \cap B_{r_{x'}}(x')$ . Hence, we can apply proposition 4.7, and prove that for all  $\eta > 0$  there exists a  $\delta(m, \mathcal{N}, Q, \Lambda, \rho, \eta)$  sufficiently small so that

$$\theta(x, \rho^{i+1}/40) \geq E - \eta. \quad (7.46)$$

Moreover, there also exists some  $x' \in \mathcal{S} \cap B_{2\rho^{i+1}/5}(x)$ . Fix any  $s \in [\rho^{i+1}, 1]$ , and for  $A \in \mathbb{N}$ ,  $A \geq \frac{10\Lambda}{\varepsilon}$  consider the sequence of radii  $s_0, \dots, s_A$  where  $s_0 = s/5$  and  $s_k = 2^{-k}s_0/5$ . We claim that there exists  $\bar{i} \in [1, A]$  such that

$$\theta(x, s_i) - \theta(x, s_{i+1}) \leq \varepsilon. \quad (7.47)$$

This is an easy consequence of the monotonicity and the relation

$$\Lambda \geq \sum_{i=1}^A \theta(x', s_i) - \theta(x', s_{i+1}). \quad (7.48)$$

If  $\delta \leq 2^{-10\Lambda/\varepsilon}$ , by definition of  $\mathcal{S} = \mathcal{S}_{\varepsilon, \delta}^k$  we have that for every  $(k+1)$ -dimensional subspace  $T$ :

$$\left(\frac{s_{\bar{i}}}{s}\right)^{2-m} s^{2-m} \int_{B_s(x)} |D_T u|^2 \geq s_{\bar{i}}^{2-m} \int_{B_{s_{\bar{i}}}(x')} |D_T u|^2 \geq \varepsilon. \quad (7.49)$$

Thus we have proved (7.39) with  $C_T(m, \Lambda, \varepsilon) = \left(\frac{s_A}{s}\right)^{m-2} \varepsilon$ .

**i=A** For  $i = A$ , one can use the same construction as above, but with radius  $r$  instead of radius  $\rho^A$ . At this stage, one also does not need to make any distinction between good and bad balls.

At this stage, we define  $\mathcal{D}$  of proposition 7.3 to be

$$\mathcal{D} = \mathcal{B}_A \cup \mathcal{G}_A. \quad (7.50)$$

We are left to prove the packing estimates (7.16).

**7.3.2. Volume estimates.** We will apply the discrete Reifenberg theorem to the measure

$$\mu_{\mathcal{D}} = \sum_{x \in \mathcal{D}} r_x^k \delta_x. \quad (7.51)$$

In order to do so, we need to check that (5.4) is satisfied for this  $\mu$ , and we exploit the best approximation theorem 6.1.

However, as it will be evident later on, we cannot apply this theorem directly. Instead, we will prove the volume estimate with an upwards induction.

7.3.3. *Inductive statement.* For convenience, we define the one-parameter family of measures  $\mu_t$  by setting

$$\mathcal{D}_t = \mathcal{D} \cap \{r_x \leq t\}, \quad \mu_t = \mu_{\mathcal{D}} \llcorner \mathcal{D}_t \leq \mu_{\mathcal{D}}. \quad (7.52)$$

Let  $T$  be such that  $2^{T-1}r < 1/70 \leq 2^T r$ . We will prove by induction on  $j = 0, 1, \dots, T$  that there exists a constant  $C_I(m)$  such that for all  $x \in B_1(0)$  and  $s = 2^j r$ :

$$\mu_s(B_s(x)) \equiv \sum_{y \in \mathcal{D}_s \cap B_s(x)} r_y^k \leq C_I(m) s^k. \quad (7.53)$$

Once this has been proved, with a simple covering argument we can turn the estimates for  $j = T$  into the estimates (7.16), replacing  $C_I(m)$  with  $C_{II}(m) = c(m)C_I(m)$  if necessary.

Base step in the induction,  $j = 0$ . The first step of the induction is easy. Since by construction  $r_x \geq r$  for all  $x \in \mathcal{D}$ , and since the balls  $\{B_{r_x/10}(x)\}_{x \in \mathcal{D}}$  are pairwise disjoint, a standard covering argument shows that for all  $x \in B_1(0)$ ,

$$\mu_r(B_r(x)) \leq C_0(m) r^k. \quad (7.54)$$

7.3.4. *Induction step.* The induction step is divided into two parts: first we are going to prove a weak packing bound for balls of radius  $2^{j+1}r$ . With this estimate, we will be able to apply the discrete Reifenberg theorem, which gives us a uniform scale invariant upper bound for the measure that lets us complete the induction.

*Coarse bounds.* Assuming that the induction step  $j$  is proved, we can easily obtain a rough bound for  $j + 1$ . Indeed, let  $x \in B_1(0)$  be arbitrary, and consider the ball  $B_{2^{j+1}r}(x)$ . By covering this ball with  $c(m)$  balls of half the radius, and using the induction hypothesis, we can estimate

$$\mu_{2^j r}(B_{2^{j+1}r}(x)) \leq c(m)C_I(m)(2^{j+1}r)^k. \quad (7.55)$$

With a similar covering argument, we can estimate the “new contributions” in  $\mu_{2^{j+1}r}$ . To be precise, since  $\{B_{r_x/10}(x)\}_{x \in \mathcal{D}}$  are all pairwise disjoint, we have

$$\bar{\mathcal{D}} = \left\{ y \in \mathcal{D} \cap B_{2^{j+1}r}(x) \text{ with } r_y \in (2^j r, 2^{j+1}r] \right\} \quad \sum_{y \in \bar{\mathcal{D}}} r_y^k \leq C_0(m)(2^{j+1}r)^k. \quad (7.56)$$

Thus, choosing  $C_I(m) \geq C_0(m)$ , we have

$$\mu_{2^{j+1}r}(B_{2^{j+1}r}(x)) \leq c(m)C_I(m)(2^{j+1}r)^k. \quad (7.57)$$

*Refined estimate.* In order to refine this last estimate, we need to apply the discrete Reifenberg 5.3. An essential tool is given by the estimates in corollary 6.3. Fix any  $B_{2^{j+1}r}(z)$  for  $z \in \mathcal{D}$ . For convenience, hereafter we will denote

$$\mu_{2^{j+1}r} \llcorner B_{2^{j+1}r}(x) \equiv \mu. \quad (7.58)$$

Set also for  $y \in \mathcal{D}$ :

$$\mathcal{W}_{\mathcal{D}}(y, s) = \begin{cases} \theta(y, 4s) - \theta(y, 2s) & \text{for } s \geq r_y/10, \\ 0 & \text{for } s < r_y/10. \end{cases} \quad (7.59)$$

By construction, and in particular by the estimates in (7.39) and (7.38) (i.e., by points 5 and 6 in the inductive construction at page 32), for  $\eta \leq \eta(m, \delta, \varepsilon, \Lambda, Q, \mathcal{N})$  sufficiently small we can apply Theorem 6.1 to  $\mu$  and any ball  $B_s(x)$  with  $x \in \mathcal{D}$  and  $s \in [r_x, 1]$ , and obtain

$$D_{\mu}^k(x, s) \leq c(m)\varepsilon^{-1}s^{-k} \int_{B_s(x)} \mathcal{W}_{\mathcal{D}}(y, s) d\mu(y). \quad (7.60)$$

As a corollary of this and (5.3), we can extend this relation for all  $s \in [r_x/10, 1]$  and obtain

$$D_\mu^k(x, s) \leq c(m)\varepsilon^{-1}s^{-k} \int_{B_{10s}(x)} \mathcal{W}_D(y, 10s) d\mu(y). \quad (7.61)$$

Note that this relation is trivially true also for  $s \leq r_x/10$ , because in this case the support of the measure  $\mu$  inside the ball  $B_{r_x/10}(x)$  is an isolated point.

We can use this estimate to prove (5.4) for the measure  $\mu$ . Indeed, fix any  $y \in B_{2^{j+2}r}(x)$ ,  $t \in (0, 2^{j+1}r]$ , and in turn choose any  $s \in [0, t]$ . For these parameters, we can bound:

$$\int_{B_t(y)} D_\mu^k(z, s) d\mu(z) \stackrel{(7.61)}{\leq} c(m)\varepsilon^{-1}s^{-k} \int_{B_t(y)} \left[ \int_{B_{10s}(z)} \mathcal{W}_D(p, 10s) d\mu(p) \right] d\mu(z). \quad (7.62)$$

Considering that

$$\{(p, z) \text{ s.t. } |z - y| \leq t \text{ and } |p - z| \leq 10s\} \subset \{(p, z) \text{ s.t. } |p - y| \leq t + 10s \text{ and } |p - z| \leq 10s\}, \quad (7.63)$$

we can exchange the variables of integration and estimate

$$\begin{aligned} \int_{B_t(y)} D_\mu^k(z, s) d\mu(z) &\leq c(m)\varepsilon^{-1} \int_{B_{11t}(y)} \frac{\mu(B_{10s}(p))}{s^k} \mathcal{W}_D(p, 10s) d\mu(p) \\ &\leq c(m)\varepsilon^{-1} C_I \int_{B_{11t}(y)} \mathcal{W}_D(p, 10s) d\mu(p). \end{aligned} \quad (7.64)$$

Recall that by (7.58),  $\mu(A) = \mu(A \cap B_{2^{j+1}r}(x))$ . Note that the induction hypothesis and the coarse estimates have been used to obtain the last inequality.

By integrating this inequality on  $\int_0^t \frac{ds}{s}$ , we get

$$\int_{B_t(y)} \left( \int_0^t D_\mu^k(z, s) \frac{ds}{s} \right) d\mu(z) \leq c(m)\varepsilon^{-1} C_I \int_{B_{11t}(y)} \left[ \int_0^t \mathcal{W}_D(z, 10s) \frac{ds}{s} \right] d\mu(z). \quad (7.65)$$

Note that for all  $x \in \mathcal{D}$ ,  $\theta(0, 1) - \theta(0, r_x) \leq \eta$ . Thus since  $t \leq 1/70$  we have

$$\int_0^t \mathcal{W}_D(x, 10s) \frac{ds}{s} = \int_{r_x}^t [\theta(x, 40s) - \theta(x, 20s)] \frac{ds}{s} \quad (7.66)$$

$$= \int_{t/2}^t \theta(x, 40s) \frac{ds}{s} + \underbrace{\int_{r_x}^{t/2} \theta(x, 40s) \frac{ds}{s} - \int_{2r_x}^{t/2} \theta(x, 20s) \frac{ds}{s}}_{=0} - \int_{r_x}^{2r_x} \theta(x, 20s) \frac{ds}{s} \quad (7.67)$$

$$= \int_{t/2}^t \left[ \theta(x, 40s) - \theta\left(x, 40 \frac{r_x}{t} s\right) \right] \frac{ds}{s} \leq c\eta. \quad (7.68)$$

This in turn implies

$$\int_{B_t(y)} \left( \int_0^t D_\mu^k(z, s) \frac{ds}{s} \right) d\mu(z) \leq c(m)\varepsilon^{-1} C_I \eta t^k. \quad (7.69)$$

By picking  $\eta$  sufficiently small (in turn: by picking  $\delta(m, \mathcal{N}, Q, \Lambda, \varepsilon, \eta)$  sufficiently small), we can apply the discrete Reifenberg theorem to  $\mu$  and prove that

$$\mu_{2^{j+1}r}(B_{2^{j+1}r}(x)) \leq C_R(m)(2^{j+1}r)^k. \quad (7.70)$$

By picking  $C_{II}(m) = \max\{C_0(m), C_R(m)\}$ , we complete the induction step, and in turn the proof of this proposition.

□

## 8. CONTINUITY IN NON-POSITIVELY CURVED SPACES

This section is devoted to the proof of the following result.

**Theorem 8.1.** *Let  $\mathcal{N}$  be a complete, simply connected manifold all of whose sectional curvatures are non-positive. Then, every minimizing harmonic map  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  satisfies*

$$\text{sing}_H(u) = \emptyset.$$

The proof will be split into two parts. In the first part of the argument we will show a general lemma, Lemma 8.2. Then, in subsection 8.1 we will show how the lemma implies the theorem.

Observe that in the single-valued case  $Q = 1$  the hypothesis that  $\pi_1(\mathcal{N}) = \{0\}$  is not necessary: indeed, in subsection 8.2 we will show how the same result holds when  $Q = 1$  under the weaker assumption that  $\mathcal{N}$  is connected. The proof will follow from the simply connected situation by means of lifting of Lipschitz-continuous functions into covering spaces. The hypothesis that  $\mathcal{N}$  is simply connected, instead, is indispensable when  $Q > 1$ : in subsection 8.3 we will provide an example of a *singular*  $Q$ -valued minimizing harmonic map in a flat target manifold  $\mathcal{N}$ .

**Lemma 8.2.** *Let  $f : \mathcal{N} \rightarrow \mathbb{R}$  be a  $C^2$ -regular function such that  $\nabla^2 f \geq 0$  on  $T_p\mathcal{N}$  for all  $p \in \mathcal{N}$ . Then*

$$f \circ u = \sum_{\ell=1}^Q f(u_\ell) = \text{const.}$$

for any 0-homogeneous energy minimizer  $u : \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$ .

*Proof.* We will split the proof of the lemma into two steps:

Claim 1: for any energy minimizer  $u : \Omega \rightarrow \mathcal{A}_Q(\mathcal{N})$ ,  $\Omega \subset \mathbb{R}^m$  open, we have that  $f \circ u : \Omega \rightarrow \mathbb{R}$  is subharmonic in the sense of distributions i.e.

$$\Delta(f \circ u) \geq 0; \tag{8.1}$$

Claim 2: any 0-homogeneous subharmonic function is constant.

The lemma is an immediate consequence of claim 1 and claim 2.

*Proof of claim 1:* Let  $\hat{f}$  be any extension of  $f$  to  $\mathbb{R}^N$  such that  $\hat{f}$  is  $C^2$  (for instance, take  $\hat{f}(p) := \phi(p)f(\Pi(p))$ , where  $\Pi(p) : \mathbf{U}_\delta(\mathcal{N}) \rightarrow \mathcal{N}$  is the nearest point projection from a  $\delta$ -tubular neighborhood  $\mathbf{U}_\delta(\mathcal{N})$  and  $\phi$  is a non-negative smooth bump function supported in  $\mathbf{U}_\delta(\mathcal{N})$  and constantly equal to 1 in a small neighborhood of  $\mathcal{N}$ ). Observe that for every  $p \in \mathcal{N}$  we have  $\nabla^2 f = (D(D\hat{f})^{T_p})^{T_p}$ , where  $v^{T_p}$  denotes the orthogonal projection of  $v$  onto  $T_p\mathcal{N}$ . In order to deduce the claim, let  $\varphi = \varphi(x) \in C_c^1(\Omega)$  non-negative be given and define the vector field

$$Y(x, p) := \varphi(x)\nabla\hat{f}(p) = \varphi(x)(D\hat{f}(p))^{T_{\Pi(p)}}.$$

The outer variation formula (1.4) provides now

$$\begin{aligned} 0 &= \int_{\Omega} \sum_{i=1}^m \sum_{\ell=1}^Q \left( \langle D_i u_\ell, \nabla\hat{f}(u_\ell) \rangle D_i \varphi + \langle D_i u_\ell, D\nabla\hat{f} \cdot D_i u_\ell \rangle \varphi \right) \\ &= \int_{\Omega} \sum_{i=1}^m \left( D_i(f \circ u) D_i \varphi + \sum_{\ell=1}^Q \nabla^2 f(u_\ell) (D_i u_\ell, D_i u_\ell) \varphi \right). \end{aligned}$$

In the last line we have used that  $D_i u_\ell \in T_{u_\ell} \mathcal{N}$  and so  $\langle D_i u_\ell, D \nabla \hat{f} \cdot D_i u_\ell \rangle = \nabla^2 f(u_\ell)(D_i u_\ell, D_i u_\ell)$ . Since the last term is non-negative we deduce the claim:

$$\int_{\Omega} \langle D(f \circ u), D\varphi \rangle \leq 0 \quad \text{for all } \varphi \in C_c^1(\Omega), \varphi \geq 0.$$

*Proof of claim 2:* Let  $h \in W^{1,2}(\mathbb{R}^m)$  be 0-homogeneous and subharmonic in the sense of distributions i.e.

$$\int \langle Dh, D\varphi \rangle \leq 0 \quad \text{for all } \varphi \in C_c^1(\mathbb{R}^m), \varphi \geq 0. \quad (8.2)$$

Suppose  $h$  is not constant. Then there exists  $a > 0$  such that  $h$  is not constant on the super-level set  $\{x : h(x) \geq -a\}$ , which in turn implies  $(h+a)^+$  is not constant. Take any non-negative  $\eta(t)$ , satisfying  $\eta(t) = 0$  for  $t > R$  (possibly a smooth approximation of  $(R-t)^+$ ), and consider the test function  $\varphi(x) = \eta(|x|^2)(h+a)^+$  in (8.2). Observe that  $D_i \varphi = \eta(|x|^2) D_i(h+a)^+ + \eta'(|x|^2) 2x^i (h+a)^+$ . But  $\sum_i D_i h(x) x^i = 0$  for a.e.  $x$  in  $\mathbb{R}^m$  because  $h$  is 0-homogeneous. Hence we deduce

$$0 \geq \int |D(h+a)^+|^2 \eta(|x|^2).$$

But this contradicts the assumption that  $(h+a)^+$  is not constant.  $\square$

**8.1. Proof of Theorem 8.1.** In this subsection we conclude the proof of Theorem 8.1. Recall that the hypotheses on  $\mathcal{N}$  imply by the Cartan-Hadamard Theorem that  $\exp_p : T_p \mathcal{N} \rightarrow \mathcal{N}$  is a covering map for every  $p \in \mathcal{N}$ . Furthermore, since  $\mathcal{N}$  is assumed to be smooth we have  $\text{dist}_{\mathcal{N}}(q, p) = |\exp_p^{-1}(q)|$ . As a further consequence we deduce that for each  $p$  the map  $q \mapsto d_p^2(q) := \text{dist}_{\mathcal{N}}(q, p)^2$  is smooth. By the second variation formula for length we deduce that  $\nabla^2 d_p^2 \geq 0$ .

*Proof of theorem 8.1.* Again we split the proof in two parts:

Claim 1: every 0-homogeneous and locally minimizing  $u : \mathbb{R}^m \rightarrow \mathcal{A}_Q(\mathcal{N})$  is constant;

Claim 2: claim 1 implies that every locally minimizing map  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  is continuous.

Obviously claim 2 is equivalent to the theorem. Let us first show how claim 2 follows from claim 1: *Proof of claim 2:* Let  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be locally energy minimizing, and suppose by contradiction that  $\text{sing}_H(u) \neq \emptyset$ . Due to the characterization of the Hölder regular set by means of the tangent maps [Hir16b, Lemma 6.1], there is  $y \in \text{sing}_H(u)$  with a non-constant tangent map  $T_y^u$  at  $y$ . But every tangent map is 0-homogeneous and locally minimizing, and thus constant by claim 1. This is the required contradiction.

*Proof of claim 1:* Let  $u \in W^{1,2}(\mathbb{R}^m, \mathcal{A}_Q(\mathcal{N}))$  be any 0-homogeneous locally minimizing map. As a consequence of the previous discussion, for every  $k > 1$  and  $p \in \mathcal{N}$  the function  $q \in \mathcal{N} \mapsto f(q) := (d_p(q)^2)^k$  is  $C^2$  regular and satisfying  $\nabla^2 f \geq 0$  on  $T_q \mathcal{N}$  since  $t \mapsto t^k$  is convex. Hence, we can apply lemma 8.2 and deduce that for all  $p \in \mathcal{N}$ ,  $k > 1$

$$d_p^{2k} \circ u = \sum_{\ell=1}^Q d_p^{2k}(u_\ell) \quad (8.3)$$

is constant. To conclude we need the following small algebraic fact, whose proof we postpone and first show the end of the argument.

**Lemma 8.3.** *Let  $\{a_\ell\}_{\ell=1}^Q, \{b_\ell\}_{\ell=1}^Q$  be two families of non-negative real numbers. Suppose that for some sequence  $k_i \rightarrow \infty$  we have*

$$\sum_{\ell=1}^Q a_\ell^{k_i} = \sum_{\ell=1}^Q b_\ell^{k_i}.$$

*Then,  $\{a_\ell\}_{\ell=1}^Q = \{b_\ell\}_{\ell=1}^Q$ .*

In order to conclude the proof, fix any  $x, y \in \mathbb{R}^m$  and let  $u(x) = \sum_{\ell=1}^Q \llbracket p_\ell \rrbracket, u(y) = \sum_{\ell=1}^Q \llbracket q_\ell \rrbracket$ . For a fixed  $p_j$  we have by (8.3) that for all  $k > 1$

$$\sum_{\ell=1}^Q \text{dist}_{\mathcal{N}}(p_\ell, p_j)^{2k} = \sum_{\ell=1}^Q \text{dist}_{\mathcal{N}}(q_\ell, p_j)^{2k}.$$

But so the lemma 8.3 implies that the number of zeros of the left- and right-hand side are the same. So we conclude that  $\#\{\ell : p_\ell = p_j\} = \#\{\ell : q_\ell = p_j\}$ . Since  $p_j$  was arbitrary we have  $u(x) = u(y)$ , that is  $u$  is constant.

It remains to give the proof of the lemma.

*Proof of lemma 8.3.* This lemma follows by induction on  $Q$ . For  $Q = 1$  the claim is obvious. Suppose the claim is proven for  $Q' < Q$ . We may assume that the families are ordered, i.e.  $a_1 \geq a_2 \geq \dots \geq a_Q$  and  $b_1 \geq b_2 \geq \dots \geq b_Q$ . If  $a_1 = 0$  the claim follows. Hence we may assume  $a_1 > 0$ . The hypothesis implies that for all  $k_i$

$$\sum_{\ell=1}^Q \left(\frac{a_\ell}{a_1}\right)^{k_i} = \sum_{\ell=1}^Q \left(\frac{b_\ell}{a_1}\right)^{k_i}.$$

If we consider the limits for  $k_i \rightarrow \infty$  we deduce that the LHS converges to  $\#\{\ell : a_\ell = a_1\}$ . If  $b_1 > a_1$ , the RHS converges to  $+\infty$ . If  $b_1 < a_1$ , on the other hand, the RHS converges to 0. Hence,  $b_1 = a_1$ . Furthermore the RHS converges therefore to  $\#\{\ell : b_\ell = b_1 = a_1\}$  which must be the same number as for the family  $\{a_\ell\}_{\ell=1}^Q$ . Hence we conclude that the assumption can now be written as

$$\sum_{\ell: a_\ell=a_1} a_1^{k_i} + \sum_{\ell: a_\ell \neq a_1} a_\ell^{k_i} = \sum_{\ell: b_\ell=a_1} b_1^{k_i} + \sum_{\ell: b_\ell \neq a_1} b_\ell^{k_i}.$$

As we have just shown the first sum on the left agrees with the first sum on the right, hence we deduce equality for the second sums for all  $k_i$ . The lemma follows now by induction hypothesis.  $\square$

$\square$

**8.2. The improved result when  $Q = 1$ .** Although it is a known result we want to give a short proof of how the previous implies the following theorem. The important fact to remark is that for the single valued case the topology of the target does not play a role.

**Theorem 8.4.** *Let  $\mathcal{N}$  be a complete, connected manifold all of whose sectional curvatures are non-positive. Then, every locally energy minimizing map  $u \in W^{1,2}(\Omega, \mathcal{N})$  is smooth.*

*Proof.* It is classical that every continuous harmonic map is smooth, hence it is sufficient to prove the continuity of the harmonic map. We will show it by induction on the dimension  $m$  of the base space  $\Omega \subset \mathbb{R}^m$ . In fact, we will proceed similarly to the simply connected situation:

Claim 1: every 0-homogeneous locally energy minimizer  $u : \mathbb{R}^m \rightarrow \mathcal{N}$  is constant;

Claim 2: every locally energy minimizing map  $u \in W^{1,2}(\Omega, \mathcal{N})$  is continuous.

*Proof of claim 1:* Assume claim 1 is proven for  $m' < m$ . In a first step we want to show that the map  $u|_{\mathbb{S}^{m-1}}$  is continuous. For  $m \leq 3$  this holds true since  $\mathcal{H}^{m-2}(\text{sing}(u)) = 0$ , [Sim96, Lemma 1 section 2.10]. Now let  $u : \mathbb{R}^m \rightarrow \mathcal{N}$  be 0-homogeneous and energy minimizing, but suppose by contradiction that when restricted to the sphere  $\mathbb{S}^{m-1}$   $u$  is not continuous, i.e.  $\text{sing}(u) \cap \mathbb{S}^{m-1} \neq \emptyset$ . Hence we can find  $y \in \text{sing}(u) \cap \mathbb{S}^{m-1}$  at which there is a tangent map  $T$  with at least one line of symmetry, i.e. such that for some  $z \in \mathbb{R}^m$  one has  $T(x + \lambda z) = T(x)$  for all  $\lambda \in \mathbb{R}$ , for all  $x$ . But this implies that  $T$  is a locally energy minimizing 0-homogeneous map from  $\mathbb{R}^{m-1}$  to  $\mathcal{N}$ . By induction hypothesis  $T$  must be constant. Hence  $\text{sing}(u) \cap \mathbb{S}^{m-1} = \emptyset$ .

We have thus concluded that  $v := u|_{\mathbb{S}^{m-1}} : \mathbb{S}^{m-1} \rightarrow \mathcal{N}$  is continuous and so smooth. Let  $P : \tilde{\mathcal{N}} \rightarrow \mathcal{N}$  be an isometric covering map e.g. we can take  $P = \exp_p : T_p \mathcal{N} \rightarrow \mathcal{N}$  by the Cartan-Hadamard Theorem. Since  $\mathbb{S}^{m-1}$  is simply connected we have that  $u_*(\pi_1(\mathbb{S}^{m-1})) \subset P_*(\pi_1(\mathbb{R}^n))$  and hence there exists a lift  $\tilde{v} : \mathbb{S}^{m-1} \rightarrow \tilde{\mathcal{N}}$  of  $v$ , that is with  $P \circ \tilde{v} = v$ , compare [Hat02, Proposition 1.33]. The 0-homogeneous extension  $\tilde{u}(x) := \tilde{v}(\frac{x}{|x|})$  must be locally energy minimizing since  $P$  is isometric (indeed, if  $\tilde{w}$  is a local competitor for  $\tilde{u}$  then  $w := P \circ \tilde{w}$  is a local competitor for  $u$ , and  $\int_{\Omega} |Dw|^2 = \int_{\Omega} |D\tilde{w}|^2$ ; hence,  $\tilde{u}$  must be locally minimizing if  $u$  is). But as proven in the simply connected situation every 0-homogeneous locally energy minimizing map  $\tilde{u} : \mathbb{R}^m \rightarrow \tilde{\mathcal{N}}$  is constant, compare claim 1 in subsection 8.1 with  $Q = 1$ . This shows the claim.

*Proof of claim 2:* Assume  $\text{sing}(u) \neq \emptyset$ . Hence we can find  $y \in \text{sing}(u)$  at which there is a non-trivial tangent map  $T$ . But the existence of  $T$  is ruled out by claim 1.  $\square$

**8.3.  $Q$ -valued counterexample.** In this subsection we want to present an example that the continuity fails for  $Q$ -valued functions if the target is not simply connected. Due to the results in subsection 8.1 we already know that the reason must be of topological nature.

**Proposition 8.5.** *There is a 2-valued energy minimizing map  $u$  from  $B^3 \subset \mathbb{R}^3$  into the flat torus  $\mathbb{T}^2 = \mathbb{C}/\mathbb{Z}^2$  with the property that  $u|_{\mathbb{S}^2}$  is Lipschitz continuous,  $\text{sing}_H(u) \in B^3$  and  $\text{sing}_H(u) \neq \emptyset$ .*

*Proof.* The construction of the example proceeds as follows:

- (1) we present an explicit example of a branched covering  $\pi : \mathcal{V} \rightarrow \mathbb{S}^2$ , where  $\mathcal{V}$  is a torus.  $\mathcal{V}$  is constructed as a complex variety in  $\hat{\mathbb{C}} \times \hat{\mathbb{C}}$ ;
- (2) using  $\pi$  we construct a 2-valued, Lipschitz continuous map  $v$  from  $\mathbb{S}^2$  into the flat torus  $\mathbb{T}^2 = \mathbb{C}/\mathbb{Z}^2$  with finite energy;
- (3) let  $u$  be a minimizer of the Dirichlet energy with respect to the boundary datum  $g(x) := v(\frac{x}{|x|})$  in  $B^3 = B_1^3(0)$ . We will show that  $u$  cannot be continuous. This last step is based on the fact that if  $u$  were continuous then  $u_{\#} [\mathbb{S}^2]$  would belong to the trivial homology in the sense of currents, a contradiction.

Let us now present the details to the outlined steps:

*Step 1:* Let  $\hat{\mathbb{C}}$  be the Riemann sphere. We fix two non zero, unequal complex numbers  $a, b$  and define the meromorphic function  $m(z) := z \frac{z-a}{z-b}$ . Consider the complex variety

$$\mathcal{V} := \left\{ (w, z) \in \hat{\mathbb{C}} \times \hat{\mathbb{C}} : w^2 = z \frac{z-a}{z-b} \right\}.$$

Consider the projection  $\pi : \hat{\mathbb{C}} \times \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  onto the second component. Restricted to  $\mathcal{V}$ , we obtain a ramified covering map

$$\pi : \mathcal{V} \rightarrow \hat{\mathbb{C}}.$$

The map  $\pi$  by definition is a two valued covering with ramification points in  $P_1 = (0, 0)$ ,  $P_2 = (0, a)$ ,  $P_3 = (\infty, b)$  and  $P_4 = (\infty, \infty)$ . We claim that  $\pi$  takes the form  $\pi(\zeta) = \zeta^2$  at each of the ramification points  $P_i$ . Furthermore this implies that  $\mathcal{V}$  is smoothly embedded, i.e. does not have any singular points.

Set  $p_1 = 0 = p'_4$ ,  $p_2 = a$ ,  $p'_3 = \frac{1}{b}$  ( $p_3 = b = \frac{1}{p'_3}$ ,  $p_4 = +\infty = \frac{1}{p'_4}$ ).

At  $P_1, P_2$  we have  $m(z) = (z - p_i)h_i(z - p_i)$  with  $h_i$  holomorphic in a neighborhood  $U_i$  of 0 and  $h_i(0) \neq 0$ . We deduce that  $\varphi_i(z) := (z - p_i)h_i(z - p_i) = m(z)$  is locally a holomorphic diffeomorphism between  $p_i + U_i$  and a neighborhood  $V_i \subset \mathbb{C}$  of 0. Now it is straightforward to check that

$$\Phi_i : \zeta \in V_i \mapsto (\zeta, \varphi_i^{-1}(\zeta^2)),$$

is a local parametrization of  $\mathcal{V}$  around  $P_i$ , i.e.  $\varphi_i \circ \pi \circ \Phi_i(\zeta) = \zeta^2$ . Changing  $U_i$  we may assume that  $V_i = \mathbb{D}_{r_i}$  for each  $i = 1, 2$ , where  $\mathbb{D}_r$  is the disc centered at  $0 \in \mathbb{C}$  with radius  $r$ . Furthermore since  $\Phi_i$  is a smooth regular map  $P_i$  is not a singular point of  $\mathcal{V}$ .

To analyze the ramification points  $P_3, P_4$  we use the inversion  $I : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$  with  $I(z) = \frac{1}{z}$ . Observe that  $(w, z) \in \mathcal{V}$  if and only if  $(w' = I(w), z' = I(z))$  is a solution of  $(w')^2 = m'(z')$  with  $m'(z) = I \circ m \circ I = \frac{b}{a} z' \frac{z' - \frac{1}{b}}{z' - \frac{1}{a}}$  or

$$I(\mathcal{V}) = \left\{ (w', z') \in \hat{\mathbb{C}} \times \hat{\mathbb{C}} : w'^2 = \frac{b}{a} z' \frac{z' - \frac{1}{b}}{z' - \frac{1}{a}} \right\}.$$

Now we can argue for  $P_3, P_4$  as for  $P_1, P_2$  interchanging  $p_1, p_2$  with  $p'_4$  and  $p'_3$  (and denote with  $U'_i$ ,  $i = 3, 4$  the related neighborhoods of 0). As a conclusion we can apply the Riemann-Hurwitz formula, and obtain

$$\chi(\mathcal{V}) = -4 + \sum_{i=1}^4 (2 - 1) = 0.$$

Hence  $\mathcal{V}$  is a torus.

*Step 2:* In the following we equip  $\mathcal{V}$  with the pullback metric  $\mathbf{g} := \iota^* \delta$  of its immersion  $\iota : \mathcal{V} \hookrightarrow \hat{\mathbb{C}} \times \hat{\mathbb{C}}$ . Observe that the metric  $\mathbf{g}$  is compatible with the conformal structure considered in step 1. The construction of  $v$  will be done in two steps. First, since  $\pi : \mathcal{V} \rightarrow \hat{\mathbb{C}}$  is a branched conformal covering of degree two there is a natural way to define 2-valued maps with finite energy. These maps are not Lipschitz continuous, in fact only  $C^{0, \frac{1}{2}}$ , but we are able to find a Lipschitz continuous map with similar properties nearby.

Let  $f : \mathcal{V} \rightarrow \mathcal{N}$  be any smooth function from the Riemann surface  $\mathcal{V}$  into a manifold  $\mathcal{N}$ . We define a two valued map  $u = u_f : \hat{\mathbb{C}} \rightarrow \mathcal{A}_2(\mathcal{N})$  using the branched covering map  $\pi : \mathcal{V} \rightarrow \hat{\mathbb{C}}$  as follows

$$u(z) := \sum_{P \in \pi^{-1}(z)} \llbracket f(P) \rrbracket,$$

counting multiplicities i.e.  $u(p_i) = 2 \llbracket f(P_i) \rrbracket$  for  $i = 1, \dots, 4$ .

We claim that  $u \in W^{1,2}(\mathbb{S}^2, \mathcal{A}_2(\mathcal{N}))$  with

$$\int_{\mathbb{S}^2} |\nabla u|^2 = \int_{\mathcal{V}} |\nabla f|^2. \quad (8.4)$$

Let  $\gamma$  be a smooth path connecting  $p_1, p_2, p_3, p_4$ . We obtain a simply connected domain  $\Omega \subset \mathbb{C}$  setting

$$\Omega := \hat{\mathbb{C}} \setminus \left( \bigcup_{i=1,2} (p_i + U_i) \cup \bigcup_{i=3,4} I(p'_i + U'_i) \cup \gamma \right).$$

Hence there exist two holomorphic maps  $\psi_i : \Omega \rightarrow \pi^{-1}(\Omega)$  with  $\psi_1(\Omega) \cup \psi_2(\Omega) = \pi^{-1}(\Omega)$  such that

$$u(z) = \llbracket f \circ \psi_1 \rrbracket + \llbracket f \circ \psi_2 \rrbracket \quad \text{for every } z \in \Omega.$$

Since the Dirichlet energy is conformally invariant (cf. [DLS11, Lemma 3.12]), we have

$$\int_{\Omega} |\nabla u|^2 = \int_{\pi^{-1}(\Omega)} |\nabla f|^2.$$

Now we consider a ramification point, for instance  $P_1$  and the related neighborhood  $p_1 + U_1$ . Using the previously introduced parametrization  $\Phi_1$  we have

$$u \circ \varphi_1^{-1}(\zeta) = \llbracket f \circ \Phi_1(\zeta^{\frac{1}{2}}) \rrbracket + \llbracket f \circ \Phi_1(-\zeta^{\frac{1}{2}}) \rrbracket.$$

The maps  $\zeta \in \mathbb{D}_{r_2} \mapsto \pm \zeta^{\frac{1}{2}}$  both together parametrize  $\mathbb{D}_{r_1}$ . Hence, as before, due to the conformal invariance of Dirichlet energy we obtain

$$\int_{\varphi_1^{-1}(\mathbb{D}_{r_2})} |\nabla u|^2 = \int_{\Phi_1(\mathbb{D}_{r_1})} |\nabla f|^2.$$

Summing up all the pieces and using that  $\mathcal{H}^2(\gamma) = 0$  we obtain (8.4).

By step 1  $\mathcal{V}$  is a smoothly embedded torus in  $\hat{\mathbb{C}} \times \hat{\mathbb{C}}$ ; hence, there exists a smooth diffeomorphism  $F : \mathcal{V} \rightarrow \mathbb{T}^2$ . Apply the above construction with the specific choice  $f = F$  to obtain

$$\tilde{v}(z) := \sum_{P \in \pi^{-1}(z)} \llbracket F(P) \rrbracket \in W^{1,2}(\hat{\mathbb{C}}, \mathcal{A}_2(\mathbb{T}^2)).$$

For the further argument let us note the following. We may fix a point  $z_0 \in \Omega$  and set  $q_i := F \circ \psi_i(z_0)$  for  $i = 1, 2$ . Since  $F$  is a diffeomorphism there is  $\varepsilon > 0$  such that  $|F \circ \psi_1(z_0) - F \circ \psi_2(z_0)| \geq 8\varepsilon$ . Hence we can find a ball  $B_{r_0}(z_0) \subset \Omega$  s.t.  $|F \circ \psi_i(z_0) - F \circ \psi_i(z)| \leq \varepsilon$  for all  $z \in B_{r_0}(z_0)$ ,  $i = 1, 2$ . Also, since  $F$  is a diffeomorphism there is a  $\delta > 0$  s.t.

$$\text{dist}_{\mathbb{T}^2}(F(q), F \circ \psi_i(z_0)) \geq 2\delta \quad \text{for all } q \in \mathcal{V} \text{ s.t. } \pi(q) \notin B_{r_0}(z_0) \text{ and } i = 1, 2.$$

Observe that this implies that

$$\tilde{v}|_{B_{r_0}(z_0)} = \llbracket \tilde{v}_1 \rrbracket + \llbracket \tilde{v}_2 \rrbracket \quad (8.5)$$

with  $\tilde{v}_1, \tilde{v}_2$  Lipschitz continuous and  $|\tilde{v}_1(z) - \tilde{v}_2(z')| \geq 4\varepsilon$  for  $z, z' \in B_{r_0}(z_0)$ . Furthermore we have

$$\text{dist}_{\mathbb{T}^2}(\text{spt}(\tilde{v}(z)), \text{spt}(\tilde{v}(z_0))) \geq 2\delta \quad \text{for all } z \notin B_{r_0}(z_0). \quad (8.6)$$

It remains to show that there is  $v \in \text{Lip}(\hat{\mathbb{C}}, \mathcal{A}_2(\mathbb{T}^2))$  nearby. This will be a consequence of the following approximation lemma:

**Lemma 8.6.** *Given  $w \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N})) \cap C^0(\Omega, \mathcal{A}_Q(\mathcal{N}))$ , for every  $\Omega' \Subset \Omega$  there exists  $w_j \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N})) \cap C^0(\Omega, \mathcal{A}_Q(\mathcal{N}))$  with*

$$\begin{aligned} w_j &\in \text{Lip}(\Omega', \mathcal{A}_Q(\mathcal{N})); & w_j &= w \text{ in a neighborhood of } \partial\Omega \\ \|\mathcal{G}(w_j, w)\|_{L^\infty(\Omega')} &\rightarrow 0; & \int_{\Omega'} |Dw_j|^2 &\rightarrow \int_{\Omega'} |Dw|^2 \text{ as } j \rightarrow \infty. \end{aligned}$$

*In particular if  $w$  is Lipschitz continuous on  $\Omega \setminus \Omega'$ , each  $w_j$  is Lipschitz continuous on the whole set  $\Omega$ .*

Before coming to the proof of this lemma let us present how to conclude. Apply the lemma to the 0-homogeneous extension of  $\tilde{v}$  in the annulus  $\Omega := B_2^3(0) \setminus B_{\frac{1}{4}}^3(0)$  to obtain an approximating sequence  $v_j \in W^{1,2}(B_2^3(0) \setminus B_{\frac{1}{4}}^3(0), \mathcal{A}_2(\mathbb{T}^2)) \cap \text{Lip}(B_{\frac{3}{2}}^3(0) \setminus B_{\frac{1}{2}}^3(0), \mathcal{A}_2(\mathbb{T}^2))$ . Choosing  $j$  sufficiently large we can guarantee that (8.5), (8.6) hold as well for the  $v_j$ . Now fix such  $j$  sufficiently large and set  $v := v_j|_{\hat{C}}$ . The 0-homogeneous extension of  $v$  i.e.  $g(x) := v(\frac{x}{|x|})$  for  $x \in B_1 \subset \mathbb{R}^3$  is an element of  $W^{1,2}(B_1, \mathcal{A}_2(\mathbb{T}^2))$  and Lipschitz continuous outside of 0. Now we may apply the direct method to obtain an energy minimizing map  $u : B_1 \rightarrow \mathcal{A}_2(\mathbb{T}^2)$  with  $u|_{\mathbb{S}^2} = g|_{\mathbb{S}^2}$ , compare [DLS11, Theorem 0.8].

*Proof of Lemma 8.6.* Since  $\mathcal{N} \hookrightarrow \mathbb{R}^N$  smooth isometrically there exists a smooth nearest point projection  $\Pi : \mathbf{U}_\delta(\mathcal{N}) \rightarrow \mathcal{N}$  for some  $\delta > 0$ . Let  $\xi_{BW} : \mathcal{A}_Q(\mathbb{R}^N) \rightarrow \mathbb{R}^M$  be the locally isometric "improved" Almgren/B. White embedding of  $\mathcal{A}_Q(\mathbb{R}^N)$ , cf. [DLS11, Section 2]. We will denote with  $\rho_{BW} : \mathbb{R}^M \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$  the related Lipschitz retraction, satisfying  $\rho_{BW} \circ \xi_{BW} = \text{id}$  on  $\mathcal{A}_Q(\mathbb{R}^N)$ , [DLS11, Corollary 2.2].

Since  $w$  is assumed to be continuous, there exists  $\tilde{w}_j$  with  $\tilde{w}_j \rightarrow \xi_{BW} \circ w$  in  $L^\infty(\Omega, \mathbb{R}^M) \cap W^{1,2}(\Omega, \mathbb{R}^M)$ ,  $\tilde{w}_j \in \text{Lip}(\Omega', \mathbb{R}^M)$  for every  $\Omega' \Subset \Omega$ , and  $\tilde{w}_j = \xi_{BW} \circ w$  in a neighborhood of  $\partial\Omega$ . For instance, one may take  $\tilde{w}_j = (1 - \theta) \xi_{BW} \circ w + \theta \eta_{\varepsilon_j} \star (\xi_{BW} \circ w)$ , for an appropriate cut-off-function  $\theta$  and a sequence of mollifiers  $\eta_{\varepsilon_j}$ .

Since  $\rho_{BW}$  is a Lipschitz-retraction and  $\xi_{BW}$  is a local isometry we conclude that the sequence

$$\hat{w}_j := \rho_{BW} \circ \tilde{w}_j : \Omega \rightarrow \mathcal{A}_Q(\mathbb{R}^N)$$

has the claimed properties up to the fact that  $\hat{w}_j$  does not necessarily take values in  $\mathcal{N}$ . But for sufficient large  $j$  we have  $\mathcal{G}(\hat{w}_j(x), w(x)) < \frac{1}{2}\delta$  for all  $x \in \Omega$  hence

$$w_j(x) := \Pi \circ \hat{w}_j(x) = \sum_{\ell=1}^Q \llbracket \Pi((\hat{w}_j(x))_\ell) \rrbracket$$

is well-defined and has all the claimed properties. It is clearly Lipschitz continuous on  $\Omega'$  since  $\Pi$  is smooth and Lipschitz. The sequence  $w_j$  converges uniformly to  $w$  since  $\Pi$  is the identity on  $\mathcal{N}$  and finally

$$\int_{\Omega} |\nabla w|^2 \leq \liminf_{j \rightarrow \infty} \int_{\Omega} |\nabla w_j|^2 \leq \liminf_{j \rightarrow \infty} \int_{\Omega} \frac{|\nabla \hat{w}_j|^2}{(1 - \text{dist}(\hat{w}_j(x), \mathcal{N})C)^2} = \int_{\Omega} |\nabla w|^2.$$

In the first inequality we used the lower-semicontinuity of the Dirichlet energy, in the second an estimate on the derivative of the nearest point projection  $\Pi$ , compare [Hir16b, Remark 2.1 (iv)].  $\square$

*Step 3:* That  $\text{sing}_H(u) \Subset B^3$  follows from the fact that  $u|_{\mathbb{S}^2}$  is Lipschitz continuous and a boundary regularity result for  $Q$ -valued locally energy minimizing maps, which can be obtained from the analogous result of [Hir16a] for "classical"  $\mathbb{R}^N$ -valued Dir-minimizers modulo slight modifications of the arguments: precisely, this is how to proceed in order to obtain the boundary regularity result [Hir16a, Theorem 0.1] in the manifold valued setting for  $s = 1$ . Only in the proof of Proposition 3.3, one replaces the application of Lemma B.2. to obtain the interpolation  $\varphi(k')$  by the application of the  $Q$ -valued Luckhaus lemma, [Hir16b, Lemma 3.1] to obtain  $\varphi(k')$ . Due to the  $L^\infty$ -bound in the Luckhaus lemma one can apply the nearest point projection  $\Pi : \mathbf{U}_\delta(\mathcal{N}) \rightarrow \mathcal{N}$  and obtain an interpolation function  $\Pi \circ \varphi(k')$  that satisfies the same bounds.

To show that  $\text{sing}_H(u) \neq \emptyset$  the idea is to use the "degree" of  $u|_{\mathbb{S}^2}$  to show that  $u$  cannot be continuous. We will use the notion of "degree" suggested by the theory of Cartesian currents. We

will need the following fact about push-forwards of integral currents by  $Q$ -valued proper Lipschitz continuous functions (see, for instance, [DS15, Section 1] or [Stu17b]): let  $\Omega \subset \mathbb{R}^m$  be open (not necessarily connected) with smooth boundary  $\partial\Omega$ ,  $\Sigma \subset \Omega$  any smooth  $k$ -dimensional surface, and  $f: \Omega \rightarrow \mathcal{A}_Q(\mathcal{N})$  Lipschitz and proper. Then, the following holds:

- $T := f_{\#} \llbracket \Omega \rrbracket$  is an  $m$ -dimensional integer rectifiable current in  $\mathcal{N}$ ,  $S := f_{\#} \llbracket \Sigma \rrbracket$  is a  $k$ -dimensional integer rectifiable current in  $\mathcal{N}$ ;
- it holds  $\partial T = f_{\#} \llbracket \partial\Omega \rrbracket$ .

In case  $\Omega$  is 3-dimensional,  $\Sigma$  and  $\mathcal{N}$  are 2-dimensional without boundary, the constancy theorem for integral currents implies that

- $T = f_{\#} \llbracket \Omega \rrbracket = 0$  since  $T$  is a 3-dimensional current supported in a 2-dimensional manifold;
- $S = f_{\#} \llbracket \Sigma \rrbracket = \theta_{\Sigma} \llbracket \mathcal{N} \rrbracket$  for some  $\theta_{\Sigma} \in \mathbb{Z}$  since  $S$  is a 2-dimensional integer rectifiable current without boundary supported in a 2-dimensional manifold;
- the following identity holds true

$$0 = \partial T = f_{\#} \llbracket \partial\Omega \rrbracket = \sum_{j=1}^J \theta_{\Sigma_j} \llbracket \mathcal{N} \rrbracket \quad (8.7)$$

where  $\Sigma_j$  are the different components of  $\partial\Omega$  i.e.  $\partial\Omega = \bigcup_{j=1}^J \Sigma_j$ .

Now we can conclude *step 3*. Assume by contradiction that  $u$  is continuous. First extend  $u$  to  $B_2$  setting  $u(x) = u(\frac{x}{|x|})$  for  $|x| > 1$ . Apply the approximation lemma 8.6 to  $u$  with  $\Omega = B_{\frac{3}{2}}$  and  $\Omega' = B_1$  to obtain a sequence  $u_j \in W^{1,2}(B_{\frac{3}{2}}, \mathcal{A}_2(\mathbb{T}^2))$  with  $u_j|_{\partial B_{\frac{3}{2}}} = u|_{\partial B_{\frac{3}{2}}}$  for all  $j$ . Since  $u$  is Lipschitz continuous on  $B_2 \setminus \overline{B_1}$  we have that  $u_j \in \text{Lip}(B_{\frac{3}{2}}, \mathcal{A}_2(\mathbb{T}^2))$ . Modifying  $u_j$  slightly we can assume that  $u_j$  is constant in a small ball  $B_r(0)$ . This can be achieved for instance by composing  $u_j$  with a Lipschitz function of the form

$$\psi(x) := \begin{cases} x & \text{for } |x| \geq 2r \\ \frac{|x|-r}{r}x & \text{for } r \leq |x| < 2r \\ 0 & \text{for } |x| < r. \end{cases}$$

Now consider the set  $\Omega = B_{\frac{3}{2}} \setminus B_{\frac{r}{2}}$  with smooth boundary components  $\Sigma_1, \Sigma_2$  given by  $\llbracket \Sigma_1 \rrbracket = \llbracket \partial B_{\frac{3}{2}} \rrbracket$  and  $\llbracket \Sigma_2 \rrbracket = -\llbracket \partial B_{\frac{r}{2}} \rrbracket$  in the sense of currents. Since  $u_j$  is constant on  $B_r$  we have  $(u_j)_{\#} \llbracket \Sigma_2 \rrbracket = 0$  by the very definition of push-forward. The identity (8.7) implies that

$$0 = (u_j)_{\#} \llbracket \Sigma_1 \rrbracket = u_{\#} \llbracket \partial B_{\frac{3}{2}} \rrbracket = u_{\#} \llbracket \partial B_1 \rrbracket. \quad (8.8)$$

We used that  $u_j = u$  on  $\partial B_{\frac{3}{2}}$  for all  $j$  and  $u$  is 0-homogeneous on  $B_2 \setminus B_1$ . Now note that  $u$  satisfies (8.5), (8.6), so that a small region around  $F \circ \psi_1(z_0)$  is covered only once, hence  $u_{\#} \llbracket \partial B_1 \rrbracket \neq 0$ . This contradicts (8.8).  $\square$

**8.4. Example of a "non-classical" tangent map.** In this section we want to observe that tangent maps of  $Q$ -valued locally energy minimizing maps may have different structures than "classical" one-valued tangent maps.

Following the classical scheme we make the following definition:

**Definition 8.7.** Let  $u \in W^{1,2}(\Omega, \mathcal{A}_Q(\mathcal{N}))$  be energy minimizing. A point  $x \in \text{sing}_H(u)$  is called a regular-singular point if for every tangent map  $T$  at  $x$  there are classical one-valued tangent maps  $T_\ell : \mathbb{R}^m \rightarrow \mathcal{N}$ , i.e. 0-homogeneous locally energy minimizing maps, such that

$$T = \sum_{\ell=1}^Q \llbracket T_\ell \rrbracket.$$

It is worth noting that every continuity point of a locally energy minimizing map has the property above, by the identification of regular points by the existence of a constant tangent map, [Hir16b, Lemma 6.1 (iii)].

We will show the following

**Proposition 8.8.** *Let  $u : B_1(0) \subset \mathbb{R}^3 \rightarrow \mathcal{A}_2(\mathbb{T}^2)$  be the energy minimizing map constructed in the previous section. Then,  $\text{sing}_H(u)$  does not contain any regular-singular point.*

*Proof.* It was shown in step 3 of the previous section that  $\text{sing}_H(u) \neq \emptyset$  and  $\text{sing}_H(u) \Subset B^3$ , hence at every point  $x \in \text{sing}_H(u)$  a tangent map exists. Let  $T : \mathbb{R}^3 \rightarrow \mathcal{A}_2(\mathbb{T}^2)$  be an arbitrary tangent map at some  $y \in \text{sing}_H(u)$ . Assume by contradiction that there are "classical" tangent maps  $T_1, T_2 : \mathbb{R}^3 \rightarrow \mathbb{T}^2$  such that

$$T = \llbracket T_1 \rrbracket + \llbracket T_2 \rrbracket.$$

Each  $T_i$  is 0-homogeneous and locally energy minimizing. Since  $\mathbb{T}^2$  is flat each  $T_i$  satisfies the assumptions of claim 1 in the proof of Theorem 8.4, hence  $T_i$  must be constant. But this contradicts that  $T$  is a non-constant tangent map and concludes the proof of the proposition.  $\square$

#### REFERENCES

- [Alm00] F. J. ALMGREN, JR., *Almgren's big regularity paper*, World Scientific Monograph Series in Mathematics **1**, World Scientific Publishing Co., Inc., River Edge, NJ, 2000,  $Q$ -valued functions minimizing Dirichlet's integral and the regularity of area-minimizing rectifiable currents up to codimension 2, With a preface by Jean E. Taylor and Vladimir Scheffer. MR 1777737. Zbl 0985.49001.
- [AT15] J. AZZAM and X. TOLSA, Characterization of  $n$ -rectifiability in terms of Jones' square function: Part II, *Geom. Funct. Anal.* **25** no. 5 (2015), 1371–1412. MR 3426057. Zbl 06521333. <https://doi.org/10.1007/s00039-015-0334-7>.
- [Bet93] F. BETHUEL, On the singular set of stationary harmonic maps, *Manuscripta Math.* **78** no. 4 (1993), 417–443. MR 1208652. <https://doi.org/10.1007/BF02599324>.
- [BDPG15] P. BOUAFIA, T. DE PAUW, and J. GOBLET, Existence of  $p$  harmonic multiple valued maps into a separable Hilbert space, *Ann. Inst. Fourier (Grenoble)* **65** no. 2 (2015), 763–833. MR 3449167. Available at [http://aif.cedram.org/item?id=AIF\\_2015\\_\\_65\\_2\\_763\\_0](http://aif.cedram.org/item?id=AIF_2015__65_2_763_0).
- [BDPW15] P. BOUAFIA, T. DE PAUW, and C. WANG, Multiple valued maps into a separable Hilbert space that almost minimize their  $p$  Dirichlet energy or are squeeze and squash stationary, *Calc. Var. Partial Differential Equations* **54** no. 2 (2015), 2167–2196. MR 3396449. <https://doi.org/10.1007/s00526-015-0861-y>.
- [CN13a] J. CHEEGER and A. NABER, Lower bounds on Ricci curvature and quantitative behavior of singular sets, *Invent. Math.* **191** no. 2 (2013), 321–339. MR 3010378. Zbl 06143195. <https://doi.org/10.1007/s00222-012-0394-3>.
- [CN13b] J. CHEEGER and A. NABER, Quantitative stratification and the regularity of harmonic maps and minimal currents, *Comm. Pure Appl. Math.* **66** no. 6 (2013), 965–990. MR 3043387. Zbl 1269.53063. <https://doi.org/10.1002/cpa.21446>.
- [DT12] G. DAVID and T. TORO, Reifenberg parameterizations for sets with holes, *Mem. Amer. Math. Soc.* **215** no. 1012 (2012), vi+102. MR 2907827. Zbl 1236.28002. <https://doi.org/10.1090/S0065-9266-2011-00629-5>.
- [DLFS11] C. DE LELLIS, M. FOCARDI, and E. N. SPADARO, Lower semicontinuous functionals for Almgren's multiple valued functions, *Ann. Acad. Sci. Fenn. Math.* **36** no. 2 (2011), 393–410. MR 2757522. <https://doi.org/10.5186/aasfm.2011.3626>.

- [DLGT04] C. DE LELLIS, C. R. GRISANTI, and P. TILLI, Regular selections for multiple-valued functions, *Ann. Mat. Pura Appl. (4)* **183** no. 1 (2004), 79–95. MR 2044333. <https://doi.org/10.1007/s10231-003-0081-5>.
- [DS15] C. DE LELLIS and E. SPADARO, Multiple valued functions and integral currents, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **XIV** (2015), 1239–1269.
- [DLS11] C. DE LELLIS and E. N. SPADARO,  $Q$ -valued functions revisited, *Mem. Amer. Math. Soc.* **211** no. 991 (2011), vi+79. MR 2663735. Zbl 1246.49001. <https://doi.org/10.1090/S0065-9266-10-00607-1>.
- [ENV16] N. EDELEN, A. NABER, and D. VALTORTA, Quantitative Reifenberg theorem for measures, *ArXiv:1612.08052* (2016). Available at <https://arxiv.org/abs/1612.08052>.
- [Gob09] J. GOBLET, Lipschitz extension of multiple Banach-valued functions in the sense of Almgren, *Houston J. Math.* **35** no. 1 (2009), 223–231. MR 2491878.
- [HLP92] R. HARDT, F.-H. LIN, and C.-C. POON, Axially symmetric harmonic maps minimizing a relaxed energy, *Comm. Pure Appl. Math.* **45** no. 4 (1992), 417–459. MR 1161539. Available at <https://doi.org/10.1002/cpa.3160450404>.
- [Hat02] A. HATCHER, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR 1867354.
- [Hir16a] J. HIRSCH, Boundary regularity of Dirichlet minimizing  $Q$ -valued functions, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **16** no. 4 (2016), 1353–1407. MR 3616337.
- [Hir16b] J. HIRSCH, Partial Hölder continuity for  $Q$ -valued energy minimizing maps, *Comm. Partial Differential Equations* **41** no. 9 (2016), 1347–1378. MR 3551461. <https://doi.org/10.1080/03605302.2016.1204313>.
- [Lin99] F.-H. LIN, Gradient estimates and blow-up analysis for stationary harmonic maps, *Ann. of Math. (2)* **149** no. 3 (1999), 785–829. MR 1709303. Available at <https://doi.org/10.2307/121073>.
- [LW08] F. LIN and C. WANG, *The analysis of harmonic maps and their heat flows*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2008. MR 2431658. <https://doi.org/10.1142/9789812779533>.
- [Mat83] P. MATTILA, Lower semicontinuity, existence and regularity theorems for elliptic variational integrals of multiple valued functions, *Trans. Amer. Math. Soc.* **280** no. 2 (1983), 589–610. MR 716839. <https://doi.org/10.2307/1999635>.
- [Mos05] R. MOSER, *Partial regularity for harmonic maps and related problems*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. MR 2155901. <https://doi.org/10.1142/9789812701312>.
- [NV16] A. NABER and D. VALTORTA, Stratification for the singular set of approximate harmonic maps, *ArXiv:1611.03008* (2016). Available at <https://arxiv.org/abs/1611.03008>.
- [NV17] A. NABER and D. VALTORTA, Rectifiable-Reifenberg and the regularity of stationary and minimizing harmonic maps, *Ann. of Math. (2)* **185** no. 1 (2017), 131–227. MR 3583353. <https://doi.org/10.4007/annals.2017.185.1.3>.
- [Rei60] E. R. REIFENBERG, Solution of the Plateau Problem for  $m$ -dimensional surfaces of varying topological type, *Acta Math.* **104** (1960), 1–92. MR 0114145. Zbl 0099.08503. Available at <http://link.springer.com/article/10.1007%2F2FBF02547186>.
- [RS08] T. RIVIÈRE and M. STRUWE, Partial regularity for harmonic maps and related problems, *Comm. Pure Appl. Math.* **61** no. 4 (2008), 451–463. MR 2383929. <https://doi.org/10.1002/cpa.20205>.
- [SU82] R. SCHOEN and K. UHLENBECK, A regularity theory for harmonic maps, *J. Differential Geom.* **17** no. 2 (1982), 307–335. MR 664498. Zbl 0521.58021. Available at <http://projecteuclid.org/getRecord?id=euclid.jdg/1214436923>.
- [Sim96] L. SIMON, *Theorems on regularity and singularity of energy minimizing maps*, *Lectures in Mathematics ETH Zürich*, Birkhäuser Verlag, Basel, 1996, Based on lecture notes by Norbert Hungerbühler. MR 1399562. <https://doi.org/10.1007/978-3-0348-9193-6>.
- [Stu17a] S. STUWARD, Multiple valued Jacobi fields, *ArXiv:1701.08753* (2017). Available at <https://arxiv.org/abs/1701.08753>.
- [Stu17b] S. STUWARD, Multiple valued sections of vector bundles: the reparametrization theorem for  $Q$ -valued functions revisited, *ArXiv:1705.00054* (2017). Available at <https://arxiv.org/abs/1705.00054>.
- [Tor95] T. TORO, Geometric conditions and existence of bi-Lipschitz parameterizations, *Duke Math. J.* **77** no. 1 (1995), 193–227. MR 1317632. Zbl 0847.42011. <https://doi.org/10.1215/S0012-7094-95-07708-4>.

S. S. and D. V., UNIVERSITÄT ZÜRICH, WINTERTHURERSTRASSE 190, CH-8057 ZÜRICH, SWITZERLAND

*E-mail address*, S. S.: `salvatore.stuward@math.uzh.ch`

*E-mail address*, D. V.: `daniele.valtorta@math.uzh.ch`