Stratification for the singular set of approximate harmonic maps

Aaron Naber · **Daniele Valtorta**

Abstract The aim of this note is to extend the results in [NV] to the case of approximate harmonic maps. More precisely, we will proved that the singular strata $S^k(u)$ of an approximate harmonic map are k-rectifiable, and we will show effect bounds on the quantitative strata. In the process we will simplify many of the arguments from [NV], and in particular we produce a new main covering lemmas which vastly simplifies the older argument.

Keywords Approximate harmonic maps · Quantitative stratification · Reifenberg theorem

Mathematics Subject Classification (2000) 58E20 · 35J60

Contents

A. Naber
Northwestern University, Evanston Illnois USA
E-mail: anaber@math.northwestern.edu

D. Valtorta
University of Zurich (CH)
E-mail: daniele.valtorta@math.uzh.ch

1 Introduction

In this note, we will work with maps between two compact Riemannian manifolds M and N such that

$$|\sec_{B_2(p)}| \leq K_M, \;\; \text{inj}(B_2(p)) \geq K_M^{-1},$$
$$\partial N = \emptyset, \;\; |\sec_N| \leq K_N, \;\; \text{inj}(N) \geq K_N^{-1}, \;\; \text{diam}(N) \leq K_N, \tag{1.1}$$
$$\dim(M) = m, \dim(N) \leq n.$$

Moreover, we will always assume that N is isometrically embedded into $\mathbb{R}^n$.

A harmonic map $u \in H^1(M, N)$ is a critical point of the Dirichlet energy for fixed boundary values. This map satisfies the Euler-Lagrange equation

$$\Delta u + A(u)(\nabla u, \nabla u) = 0, \tag{1.2}$$

where A is the second fundamental form of $N \subset \mathbb{R}^n$. An important object in the study of harmonic maps is the normalized energy

$$\theta(x, r) = r^{2-m} \int_{B_r(x)} |\nabla u|^2. \tag{1.3}$$

This quantity turns out to be (almost) monotone for *stationary* harmonic maps. Combined with an ϵ-regularity theorem, this immediately yields the partial regularity of such maps.

Loosely speaking, approximate harmonic maps are solutions to

$$\Delta u + A(u)(\nabla u, \nabla u) = f, \tag{1.4}$$

where f in an L^2 function. Such a map is called *stationary* if it also satisfies

$$\nabla^i \left(|\nabla u|^2 g_{ij} - 2 \left\langle \nabla_i u, \nabla_j u \right\rangle \right) + 2 \left\langle \nabla_j u, f \right\rangle = 0. \tag{1.5}$$

As for harmonic maps, one defines the singular set of the map u as the set of points in the domain which are not continuous:

$$\mathcal{S}(u) = \{x \in M \;\; s.t. \;\; \exists r > 0 \;\; s.t. \;\; u|_{B_r(x)} \;\; \text{is continuous} \}^C. \tag{1.6}$$

By elliptic regularity, it is easy to see that around a regular point almost harmonic maps are $C^{0,\alpha}$ continuous, where $\alpha = \alpha(m, K_M, N, p) > 0$, but in general higher regularity depends on the regularity of f.

As for standard harmonic maps, one can define *stationary* almost harmonic maps and prove that the normalized energy $\theta(x, \cdot)$ is almost monotone, in the sense that has bounded variation at every point. Using only this property, and an ϵ-regularity theorem, one shows that $\mathcal{S}(u)$ has zero $m - 2$-Hausdorff measure and that each of these maps has a (possibly non-unique) tangent map at every point.

By looking at the symmetries of these tangents, see Section 1.2 for precision, one can define a stratification $\mathcal{S}^k(u)$ for $\mathcal{S}(u)$ by setting

$$\mathcal{S}^k(u) = \{x \in B_1(0) \;\; s.t. \;\; \text{no tangent map at } x \text{ is k-symmetric} \}. \tag{1.7}$$

The aim of this note is to prove that the results obtained in [NV] for stationary harmonic maps continue to hold in this setting, and in particular the strata $S^k(u)$ are all k-rectifiable. Our full collection of main theorems is listed in Section 3, after some preliminaries are introduced, however we begin by stating the main structure result for the stratification itself:

Theorem 1 *Let $u : B_3(p) \subseteq M \to N$ be a stationary approximately harmonic map, where M and N satisfy (1.1), and $f \in L^p$ for $p > \frac{m}{2}$. Then for each k we have that $S^k(u)$ is k-rectifiable.*

In fact, in Section 3 we will weaken slightly the assumptions on f. Moreover, we will also prove in Section 3 uniform volume bounds for the *quantitative* strata $S^k_{\epsilon,r}$, which will be introduced in the next section.

There are three key elements in the proof of these estimates: an estimate on the β_2-Jones' numbers for the singular strata derived using the monotonicity formula and the L^2-best subspace approximation theorem, a Reifenberg-type theorem which allows us to turn these estimates into volume bounds and rectifiability for the singular strata, and an inductive covering argument which guarantees the applicability of this Reifenberg theorem on the strata. While all the results and their proofs are essentially the same as in [NV], albeit done in slightly more generality, it is worth noticing that the proof inductive covering argument has been simplified here.

1.1 Domains with curvature

For the sake of simplicity, throughout the rest of this note we will assume that $M = \Omega \subseteq \mathbb{R}^m$. Assuming that the domain M is flat simplifies the technicalities involved in the computations, but involves no fundamental changes in comparison to the general case - primarily, one obtains errors from the almost aspect of the monotonicity formula which simply need to be accounted for in a standard manner. Thus, all the results described in this note carry over to the general case, up to minor technical differences. In particular, all the constants involved in the bounds would also depend on the manifold M, and actually just on a lower bound on its injectivity radius and an upper bound on the sectional curvature. Moreover, the normalized energy θ, and its adapted version for the almost harmonic case $\hat{\theta}$, would need to be slightly changed as described in Remark 13.

1.2 Quantitative stratification for general maps

In this section, we introduce in detail the quantitative stratification of the singular set $S(u)$. As mentioned in the introduction, the idea behind the quantitative stratification is to group all points in the domain of the map u according to the number of approximate symmetries of u at some scale, as opposed to the standard stratification which looks at the exact symmetries of the set of tangent maps.

A first version of the quantitative stratification can be found in Almgren's big regularity paper (see [Alm00, section 2.25]), and it was reintroduced in [CN13a, CN13b] to provide the first effective regularity results. In particular, in [CN13a, CN13b] the authors use the quantitative stratification to prove estimates on noncollapsed manifolds with Ricci curvature bounds, and on the singular set of stationary and minimizing harmonic maps. As a corollary of the estimates, in [CN13b] the authors obtain uniform $W^{1,p}$ bounds for minimizing harmonic maps for all $p < 3$. This technique has since been used in [CHN13b], [CHN13a], [CNV15], [FMS15], [BL15] to prove similar results in the areas of mean curvature flow, critical sets of elliptic equations, harmonic map flow, and biharmonic maps.

A significant improvement on these techniques has been obtained in [NV], where the authors use sharp estimates on the Jones' β_2 numbers (called *distortion* in [NV]) in terms of the normalized energy to prove sharp volume bounds and rectifiability for the singular strata of harmonic maps. As mentioned in the introduction, the aim of this note is to extend these results to the case of approximately harmonic maps, but perhaps more importantly to give somewhat simplified proofs of the main arguments of [NV].

In order to define precisely the quantitative stratification, we need to introduce the concept of k-symmetric maps.

Definition 2 Given a map $h \in H^1(\mathbb{R}^m, N)$, we say that

1. h is homogeneous wrt the point p if $h(p + \lambda v) = h(p + v)$ for all $\lambda > 0$ and $v \in \mathbb{R}^m$. Equivalently, $\partial_{r_p} h = 0$, where r_p is the radial direction wrt p,
2. h is k-symmetric if it is homogeneous wrt the origin and it has an invariant k-subspace, i.e., if there exists a linear subspace $V \subseteq \mathbb{R}^m$ of dimension k such that

$$h(x + v) = h(x) \ \ \forall x \in \mathbb{R}^m \ \ \forall v \in V . \tag{1.8}$$

As a notation, we say that h is 0-symmetric iff it is homogeneous wrt the origin. In the definition, we insist that a k-symmetric map be *both homogeneous and k-invariant*.

Example 3 It is very easy to produce examples of these maps by taking maps defined on S^{m-1} and extending them by homogeneity on the whole $\mathbb{R}^m$. Note that necessarily if a map is homogeneous, then it is continuous only if it is constant.

The first nonconstant explicit example is the map $h : \mathbb{R}^m \to S^{m-1}$ given by

$$h(x) = \frac{x}{|x|} , \tag{1.9}$$

which is a H^1 map for $m \geq 3$ (and actually a *minimizing* harmonic map in these cases, see [Lin87, CG89]). We can easily build from this example a k-symmetric map by defining $g : \mathbb{R}^m \times \mathbb{R}^k \to S^{m-1}$ as $g(x, y) = h(x)$.

The quantitative stratification is based on how close a map is to some k-symmetric map at different scales.

Definition 4 Given a map $g \in H^1(\Omega, N)$, we say that $B_r(x) \subset \Omega$ is (k, ϵ)-symmetric for g if there exists some k-symmetric function h such that

$$\fint_{B_r(x)} |g(y) - h(y - x)|^2 \equiv \frac{1}{\omega_m r^m} \int_{B_r(x)} |g(y) - h(y - x)|^2 \le \epsilon. \qquad (1.10)$$

An equivalent definition may be given in terms of the blow-up map $T_{x,r}^g$, defined by

$$T_{x,r}^g(y) = g(x + ry). \qquad (1.11)$$

Indeed, since h is homogeneous by assumption, we have

$$\fint_{B_r(x)} |g(y) - h(y - x)|^2 = \fint_{B_1(0)} \left| T_{x,r}^g(y) - h(y) \right|^2. \qquad (1.12)$$

Given this definition of approximate symmetry, we can define the quantitative stratification by classifying the points in the domain according to how close they look at different scales to something which is k-symmetric.

Definition 5 Given $u \in H^1(\Omega, N)$, $r, \epsilon > 0$ and $k \in \{0, 1, \cdots, m\}$, we define

$$\mathcal{S}_{\epsilon,r}^k(u) \equiv \{x \in \Omega \ \text{ s.t. } \text{ for no } r \le s < 1, \ B_s(x) \text{ is } (k + 1, \epsilon)\text{-symmetric wrt } u\}. \qquad (1.13)$$

Note that these sets have some immediate inclusion properties coming from their definition. In particular, if $k' \le k$, $\epsilon' \ge \epsilon$ and $r' \le r$, we have

$$\mathcal{S}_{\epsilon',r'}^{k'}(u) \subseteq \mathcal{S}_{\epsilon,r}^k(u). \qquad (1.14)$$

Given this, one can construct the sets

$$\mathcal{S}_\epsilon^k(u) = \bigcap_{r > 0} \mathcal{S}_{\epsilon,r}^k(u), \quad \mathcal{S}^k(u) = \bigcup_{\epsilon > 0} \mathcal{S}_\epsilon^k(u). \qquad (1.15)$$

Remark 6 Note that, for an approximate harmonic map, the set $\mathcal{S}^k(u)$ has another characterization:

$$\mathcal{S}^k(u) = \{x \in \Omega \ \text{ s.t. } \text{ no tangent maps at } x \text{ is } k + 1 - \text{symmetric}\}. \qquad (1.16)$$

For a precise statement, see Lemma 26.

Example 7 It is interesting to study these sets with an example. In particular, consider the map $g : \mathbb{R}^3 \to S^2$ given by $g(x) = x/|x|$. Since the map g is 0-homogeneous but not 1-symmetric, it is clear that $0 \in \mathcal{S}_{\epsilon,r}^0$ for all r and $\epsilon < \epsilon_0$ sufficiently small. In particular, ϵ_0 can be taken to be the $L^2(B_1(0))$ distance from g to all 1-symmetric maps h, which is easily seen to be positive.

Now given any $x \in \mathbb{R}^n \setminus \{0\}$, since g is continuous at x, for all $\epsilon > 0$ there exists a radius $r(\epsilon)$ such that

$$\fint_{B_{r|x|}(x)} |g(y) - g(x)|^2 \le \epsilon. \qquad (1.17)$$

Thus every point $x \neq 0$ will eventually be almost 3-symmetric, if we consider a small enough radius. However, for r sufficiently big, $g|_{B_r(x)}$ and $g|_{B_r(0)}$ are close in the L^2 norm. Thus, for $r \geq s(|x|, \epsilon)$ sufficiently large we have that x will belong to $\mathcal{S}^0_{\epsilon,r}$. Summing up, we obtain that

$$\mathcal{S}^0_{\epsilon,r}(g) = \overline{B_{s(r,\epsilon)}(0)} \quad \text{where} \quad s(r,\epsilon) \text{ is increasing in } r \text{ and decreasing in } \epsilon. \tag{1.18}$$

Moreover, for $\epsilon < \epsilon_0$, we have $s(0, \epsilon) = \lim_{r \to 0} s(0, \epsilon) = 0$. Up to different constants, the sets $\mathcal{S}^1_{\epsilon,r}(g)$ and $\mathcal{S}^2_{\epsilon,r}(g)$ also behave in a similar way, while evidently $\mathcal{S}^3_{\epsilon,r}(g) = \mathbb{R}^3$ (but this last statement has clearly no meaning).

2 Approximate harmonic maps: definition and monotonicity

Before moving to approximate harmonic maps, we briefly recall the most important aspects of the regularity theory for harmonic maps.

2.1 Stationary harmonic maps: regularity

Loosely speaking, harmonic maps are critical points of the Dirichlet energy, in particular with respect to compact variations in the target space and in the domain space, in the case of stationary harmonic maps. The definition of these object is standard in literature, here we briefly recall it.

Given a compact C^2 Riemannian manifold N, we define the Sobolev space $H^1(M, N)$ by isometrically embedding N into a Euclidean space $\mathbb{R}^n$, and considering

$$H^1(\Omega, N) = \left\{ f \in H^1(\Omega, \mathbb{R}^n) \ s.t. \text{ for a.e. } x \in M, \ f(x) \in N \right\}. \tag{2.1}$$

Let $\Omega \subset \mathbb{R}^m$ be a domain. A stationary harmonic map $u \in H^1(\Omega, N)$ is a critical point for this Dirichlet energy. In particular, let $v : M \to \mathbb{R}^n$ and $w : \Omega \to \mathbb{R}^m$ be smooth vector fields, both with compact support in Ω. Then if u is a stationary harmonic map we have

$$\frac{d}{dt}\Big|_{t_0} \int_\Omega |\nabla (\Pi_N(u + tv))|^2 = 0 \quad \text{and} \quad \frac{d}{dt}\Big|_{t_0} \int_\Omega |\nabla (u(x + tw))|^2 = 0, \tag{2.2}$$

where Π_N is the nearest point projection from $\mathbb{R}^n$ onto N, a map which is well-defined and C^1 on a small neighborhood of N.

By standard computations (see for example [Xin96, Mos05]), the first condition gives the Euler-Lagrange equation

$$\Delta u + A(u)(\nabla u, \nabla u) = 0. \tag{2.3}$$

This is an equation satisfied by $u : \Omega \to \mathbb{R}^n$ in the weak H^1 sense, and $A(u)$ is the second fundamental form of the embedding $N \hookrightarrow \mathbb{R}^n$ evaluated at $u(x)$. Maps satisfying only this equation are called *weakly harmonic*.

The second condition gives rise to a separate Euler-Lagrange equation. In particular, we see that for all $i = 1, \cdots, m$ we have

$$\mathrm{div}_j \left(|\nabla u|^2 \, \delta_{ij} - 2 \left\langle \nabla_i u, \nabla_j u \right\rangle \right) = 0 \,. \tag{2.4}$$

Also this equation is to be interpreted in the weak sense, in particular we have that for all smooth vector fields with compact support $\xi : \Omega \to \mathbb{R}^m$

$$\int_\Omega \left(|\nabla u|^2 \, \delta_{ij} - 2 \partial_i u \partial_j u \right) \partial_i \xi_j = 0 \,. \tag{2.5}$$

Note that a map $u \in C^2(\Omega, \mathbb{R}^n)$ satisfying (2.3) will automatically satisfy also (2.5). However, in general this is not the case if $u \in H^1$ only. Maps satisfying both (2.3) and (2.5) are called *stationary* harmonic maps.

An important consequence of (2.5) is the monotonicity of the normalized energy θ defined by

$$\theta(x, r) = r^{2-m} \int_{B_r(x)} |\nabla u|^2 \,. \tag{2.6}$$

Lemma 8 *For a stationary harmonic map $u \in H^1(\Omega, \mathbb{R}^n)$, and for almost all x, r such that $B_r(x) \subset \Omega$, we have the monotonicity identity*

$$\theta'(x, r) = 2r^{-m} \int_{\partial B_r(x)} |(y - x) \cdot \nabla u|^2 \, dS(y) \geq 0 \,. \tag{2.7}$$

This in particular implies that $\theta(x, r)$ is a monotone function in r, and it also gives quantitative estimates for the integral of the radial part of the energy. Indeed, as a corollary we obtain that for all $0 \leq s \leq r$:

$$2 \int_{B_r(x) \backslash B_s(x)} \frac{|(y - x) \cdot \nabla u|^2}{|y - x|^m} = \theta(x, r) - \theta(x, s) \,. \tag{2.8}$$

Proof This identity is obtained by plugging in the vector field $\xi_i = \chi_{B_r(x)} \cdot x_i$ in (2.5) (or better, a sequence of smooth approximations of this field).

The last ingredient needed for a very basic regularity theory for stationary harmonic maps is the ϵ-regularity theorem by Bethuel (see [Bet93]).

Theorem 9 (ϵ-**regularity theorem**) *Let $u \in H^1$ be a stationary harmonic map. There exists an $\epsilon = \epsilon(m, N) > 0$ such that $\theta(x, r) \leq \epsilon$ implies that u is smooth on $B_{r/2}(x)$ with*

$$\left| \nabla^k u(y) \right| \leq C(m, k, N) r^{-k} \theta(x, r) \,. \tag{2.9}$$

for all $k = 1, \cdots, \infty$ and $y \in B_{r/2}(x)$.

With these two theorems and a simple covering argument (see [Bet93]), it is easy to see that

$$\mathcal{H}^{n-2}(\mathcal{S}(u)) = 0 \,. \tag{2.10}$$

Note that in general this is not the case for weakly harmonic maps.

2.2 Monotonicity and ϵ-regularity

Following the natural approach of [Mos05], here we introduce approximately harmonic maps. In this note, approximately harmonic maps will always mean *stationary* approximately harmonic maps.

Basically, we say that a map is approximately harmonic if it satisfies equations (2.3) and (2.4) up to a controlled error.

Definition 10 A map $u \in H^1(\Omega, N)$ is said to be approximately harmonic if the following are satisfied in the sense of distributions

$$\Delta(u) + A(u)(\nabla u, \nabla u) = f\,, \tag{W}$$

$$\mathrm{div}_{\mathrm{j}} \left(|\nabla u|^2\, \delta_{ij} - 2 \left\langle \nabla_i u, \nabla_j u \right\rangle \right) + 2 \left\langle \nabla_i u, f \right\rangle = 0\,, \tag{S}$$

for some $f \in L^2(\Omega)$. In order to obtain almost monotonicity for the normalized energy of the solution u, we assume that there exists $\mathcal{F}, \gamma > 0$ such that for all $B_r(x) \subset \Omega$:

$$r^{4-m} \int_{B_r(x)} |f|^2 \le \mathcal{F} r^\gamma\,, \tag{f}$$

where $\gamma > 0$, and $\mathcal{F}$ and γ are independent of x. One can interpret (f) as saying some weighted maximal function of f is uniformly bounded. Although this assumption might sound extremely technical or unnatural, it is easy to see that any $f \in L^p(\Omega)$ with $p > m/2$ satisfies this condition. Indeed, by Hölder inequality we have

$$r^{4-m} \int_{B_r(x)} |f|^2 \le r^{4-m} \left(\int_{B_r(x)} |f|^p \right)^{\frac{2}{p}} (\omega_m r^m)^{\frac{p-2}{p}} \le c(m, p)\, \|f\|^2_{L^p(\Omega)}\, r^{\frac{2}{p}(2p-m)}\,. \tag{2.11}$$

Remark 11 Note that the condition (S) is quite natural, since it would be satisfied automatically by a smooth map u solving (W). However, in general an H^1 solution to (W) will *not* satisfy also (S).

The reason why we insist on (S) is that with this relation we can prove an almost monotonicity formula for approximately harmonic maps, which is essential for the estimates we need. The following lemma is taken from [Mos05, lemma 4.1], however the quantities analyzed are slightly different. For the reader's convenience, we sketch a proof here.

Lemma 12 *For $m \ge 3$, let $u \in H^1(B_3(0), N)$ be an approximately harmonic map satisfying* (W) *and* (S) *with* (f). *Suppose also that*

$$\theta(0, 3) = 3^{2-m} \int_{B_3(0)} |\nabla u|^2 \le \Lambda\,. \tag{2.12}$$

Define the function

$$\hat{\theta}_u(x, r) \equiv \hat{\theta}(x, r) = \theta(x, r) - \frac{2}{m-2} r^{2-m} \int_{B_r(x)} \left\langle (y-x) \cdot \nabla u|_y, f(y) \right\rangle dVol(y) + \frac{1}{(m-2)^2} \int_{B_r(x)} \frac{|f|^2}{|y-x|^{m-4}} dVol(y)\,. \tag{2.13}$$

Then for all $B_s(x) \subseteq B_r(x)$ with $r \le 1$ and $x \in B_1(0)$ we have $\hat{\theta}(x, r) \ge 0$ and

$$0 \le \int_{B_r(x) \backslash B_s(x)} \frac{|(y-x) \cdot \nabla u|^2}{|y-x|^m} \le \hat{\theta}(x, r) - \hat{\theta}(x, s). \tag{2.14}$$

Moreover, for all $a > 0$ we have the bounds

$$(1-a)\hat{\theta}(x, r) - \frac{c(m, \gamma)\mathcal{F}}{a} r^\gamma \le \theta(x, r) \le (1+a)\hat{\theta}(x, r) + \frac{c(m, \gamma)\mathcal{F}}{a} r^\gamma, \tag{2.15}$$

and the uniform bounds

$$\theta(x, r) + \hat{\theta}(x, r) \le c(m)\Lambda + c(m, \gamma)\mathcal{F}r^\gamma. \tag{2.16}$$

Proof For convenience, we assume $x = 0$. First of all, we prove that, under condition (f), the function $\hat{\theta}$ is finite. For $m = 3, 4$, this is obvious since $f \in L^2$, and (f) is not needed. For $m \ge 5$, by integration by parts we have

$$\int_0^r s^{4-m} \int_{\partial B_s(0)} |f|^2 \, d\theta ds = \left[s^{4-m} \int_{B_s(0)} |f|^2 \right]_0^r + \int_0^r s^{3-m} dt \int_{B_t(0)} |f|^2 \le \tag{2.17}$$

$$\le \mathcal{F}r^\gamma + (m-4)\mathcal{F} \int_s^r t^{3-m} t^{m-4+\gamma} dt \le \mathcal{F}\left(1 + \frac{m-4}{\gamma}\right) r^\gamma. \tag{2.18}$$

As for the monotonicity of $\hat{\theta}$, the proof is a simple application of the stationary equation (S). Indeed, let ϕ be any Lipschitz radial cutoff function with $\phi(0) = 1$ and $\phi(r) = 0$, and consider the vector field $\xi(y) = \phi(|y|)\mathbf{y}$. By testing (S) with ξ, we get

$$\int_\Omega \phi\left((m-2)|\nabla u|^2 - 2\langle y \cdot \nabla u, f\rangle\right) = \int_\Omega (-\phi')|y|\left(|\nabla u|^2 - 2|\hat{r} \cdot \nabla u|^2\right) \tag{2.19}$$

where $\hat{r} = |y|^{-1}y$ is the unit norm radial vector. By letting ϕ converge to $\chi_{B_r(0)}$, we prove that for almost all r

$$\int_{B_r(0)} \left((m-2)|\nabla u|^2 - 2\langle y \cdot \nabla u, f\rangle\right) = r \int_{\partial B_r(0)} \left(|\nabla u|^2 - 2|\hat{r} \cdot \nabla u|^2\right). \tag{2.20}$$

The derivative of $\hat{\theta}$ is, at least a.e. in r,

$$\hat{\theta}(0, r)' = \frac{2-m}{r}\theta(0, r) + r^{2-m} \int_{\partial B_r(0)} |\nabla u|^2 + 2r^{2-m} \int_{B_r(0)} \langle y \cdot \nabla u, f\rangle + \tag{2.21}$$

$$-\frac{2}{m-2} r^{2-m} \int_{\partial B_r(0)} \langle y \cdot \nabla u, f\rangle + \frac{1}{(m-2)^2} r^{4-m} \int_{\partial B_r(0)} |f|^2. \tag{2.22}$$

By plugging in (2.20), we obtain for a.e. r

$$\hat{\theta}(0, r)' = r^{-m}\left(2 \int_{\partial B_r(0)} |y \cdot \nabla u|^2 - \frac{2}{m-2} r^2 \int_{\partial B_r(0)} \langle y \cdot \nabla u, f\rangle + \frac{1}{(m-2)^2} r^4 \int_{\partial B_r(0)} |f|^2\right) \ge \tag{2.23}$$

$$\ge r^{-m} \int_{\partial B_r(0)} |y \cdot \nabla u|^2. \tag{2.24}$$

where the last estimate is a simple application of Young's inequality. By integrating this relation over $[s, r]$, we obtain (2.14).

As for (2.15), this follows immediately from Young's inequality applied to the rhs of (2.13) and (2.18). Indeed:

$$\left| \frac{2}{m-2} r^{2-m} \int_{B_r(x)} \left\langle (y-x) \cdot \nabla u|_y, f(y) \right\rangle \right| \le a\ \theta(x, r) + \frac{r^{4-m}}{(m-2)^2\,a} \int_{B_r(x)} |f|^2\ . \qquad (2.25)$$

We conclude by noticing that since $\hat{\theta}$ is monotone in r, and since by (2.18) and (2.25):

$$\hat{\theta}(x, 1) \le 2\theta(x, 1) + \mathcal{F} + c(m, \gamma)\mathcal{F} \le c(m)\Lambda + c(m, \gamma)\mathcal{F}, \qquad (2.26)$$

then the bounds on $\hat{\theta}$ are obvious. The uniform bounds on θ are then a consequence of the previous estimate (2.15).

Remark 13 In case the domain space is a Riemannian manifold M, the definition of $\hat{\theta}$ would need to be changed a little. We refer the reader to [Xin96, section 2.2] for more details on this. We just mention that the changes arise from the fact that the Hessian of the distance function r is a little different than in the Euclidean case. This is related to the derivatives of the radial vector field ξ used in the proof of the monotonicity formula.

For approximate harmonic maps, it is also possible to prove an ϵ-regularity theorem as for stationary harmonic maps. The underlying techniques are basically the same, up to minor technical details. Here we quote the ϵ-regularity theorem in [Mos05, proposition 4.1].

Theorem 14 *Let u solve* (W) *and* (S) *with* (f). *Then there exists* $\epsilon_0, \alpha > 0$ *depending only on* m, N, γ *such that*

$$\theta(x, r) \le \epsilon_0 \quad \implies \quad u \in C^{0,\alpha}(B_{r/2}(x))\ \ with\ \ \|u\|_{C^{0,\alpha}} \le C(m, N, \mathcal{F}, \gamma)\,. \qquad (2.27)$$

Proof The proof is based on a polynomial decay (in r) for the normalized energy $\theta(x, r)$ which is very similar to the proof of the ϵ-regularity theorem for stationary harmonic maps. For the details, we refer the reader to [Mos05, proposition 4.1].

As a corollary, we obtain a similar statement for $\hat{\theta}$.

Corollary 15 *Let u be as above. Then there exists* $\epsilon_0, r_0, \alpha > 0$ *depending only on* m, N, γ *such that*

$$\hat{\theta}(x, r) \le \epsilon_0 \quad and \quad r \le r_0 \quad \implies \quad u \in C^{0,\alpha}(B_{r/2}(x))\ \ with\ \ \|u\|_{C^{0,\alpha}} \le C(m, N, \mathcal{F}, \gamma)\,. \quad (2.28)$$

Proof This corollary follows immediately from the estimates in (2.15) and the previous proposition.

As with stationary harmonic maps, this ϵ-regularity theorem and a simple covering argument imply that

$$\mathcal{H}^{n-2}(\mathcal{S}(u)) = 0\,. \qquad (2.29)$$

Invariance by scale As it is well-known, for stationary harmonic map the normalized energy θ is scale-invariant, in the sense that if we consider the map

$$T^u_{x,r}(y) \equiv u(x + ry) \,, \tag{2.30}$$

then $\theta_T(0, 1) = \theta_u(x, r)$, without any scaling factors. This is an essential property of the normalized energy.

In the case of approximate harmonic maps, the quantity $\hat{\theta}$ satisfies similar properties. However, some scaling factors are inevitably present on the zero order term f. Indeed, the map $T \equiv T^u_{x,r}$ in this case will be an approximate harmonic map satisfying

$$\Delta(T) + A(T)(\nabla T, \nabla T) = \tilde{f} \,, \tag{2.31}$$

$$\mathrm{div}\left(|\nabla T|^2 \, e_i - 2 \left\langle \nabla_{e_i} T, \nabla T \right\rangle\right) + 2 \left\langle \nabla_{e_i} T, \tilde{f} \right\rangle = 0 \,, \tag{2.32}$$

$$\tilde{f}(y) \equiv r^2 f(x + ry) \,. \tag{2.33}$$

Thus, we obtain that $\hat{\theta}_u(x, r) = \hat{\theta}_T(0, 1)$ if we replace f with $\tilde{f}$ in the definition of $\hat{\theta}_T$. Note also that if (f) is satisfied, then for all $B_r(x) \subset B_2(0)$ we have

$$\int_{B_1(0)} \left|\tilde{f}\right|^2 = r^4 \int_{B_1(0)} |f(x + ry)|^2 = r^{4-m} \int_{B_r(x)} |f(y)|^2 \le \mathcal{F}r^\gamma \,. \tag{2.34}$$

2.3 Weak convergence

In this section, we recall a standard result about the convergence of almost harmonic maps. In particular, we want to show that given a sequence of approximate harmonic maps with bounded energy and such that $f_i \to 0$, their weak sub-limit is a weakly harmonic map. This result is an easy adaptation of standard estimates in literature, see for example [TW95, theorem 4] or [Sch84, corollary 2.3].

Proposition 16 *Let u_i solve* (W) *and* (S), *where $f_i \in L^2(B_3(0))$ satisfies* (f) *with $\mathcal{F}$ and γ independent of i. Assume that $\int_{B_3(0)} |\nabla u_i|^2 \le C$ and that $f_i \rightharpoonup^* 0$ in weak L^2. Then there exists a subsequence (which will still be denoted by u_i) such that*

1. *u_i converges in the weak H^1 sense to some u*
2. *there exists a close set Σ with finite $n - 2$ packing content such that the sequence u_i converges strongly in $H^1_{loc}(B_1(0) \setminus \Sigma) \cap C^{0,\alpha/2}_{loc}(B_1(0) \setminus \Sigma)$ to u, which is a smooth map on $B_1(0) \setminus \Sigma$.*
3. *u is weakly-harmonic (but not necessarily stationary harmonic)*
4. *u enjoys the unique continuation property, in the sense that if there exists another weakly harmonic map v such that $u = v$ a.e. on an open set, and v is smooth away from a set of Σ' with finite $n - 2$ packing content, then $u = v$*

Remark 17 Note that even if u_i are stationary harmonic (i.e. if $f_i = 0$ for all i), their limit might not be stationary. An interesting example of this is given in [DLL03].

Remark 18 As a corollary of this theorem, we obtain the following. Consider the measure $|\nabla u_i|^2 \, d\text{Vol}$ on $B_1(0)$. Then in the weak sense of measures we have

$$|\nabla u_i|^2 \, d\text{Vol} \rightharpoonup^* |\nabla u|^2 \, d\text{Vol} + \nu \,, \tag{2.35}$$

where ν is a nonnegative measure by Fatou's lemma, which is supported on the set Σ described in the previous theorem. In particular, the support of ν has finite $n - 2$ packing content. For stationary harmonic maps, this measure is called *defect measure*, and it has been extensively studied in [Lin99].

Before stating this result, we need a technical lemma about weak H^1 convergence in the ϵ-regularity region.

Lemma 19 *Let u_i solve* (W) *and* (S), *where $f_i \in L^2(B_3(0))$ satisfies* (f) *with $\mathcal{F}$ and γ independent of i. Assume that*

1. *$\theta_{u_i}(0, 2) \leq \epsilon_0$, where ϵ_0 is the parameter in theorem 14,*
2. *$u_i \rightharpoonup^* u$ in the weak $H^1(B_3(0))$ sense,*
3. *$f_i \rightharpoonup^* f$ in the weak L^2 sense.*

Then u_i converges to u in the strong $H^1(B_1(0))$ sense and on $B_1(0)$ the map u solves

$$\Delta(u) = A(u)(\nabla u, \nabla u) + f \tag{2.36}$$

in the sense of distributions.

Proof The proof relies on standard techniques, but for the sake of completeness we outline it here. Note that by the ϵ-regularity theorem in theorem 14, we have that $\|u_i\|_{C^{0,\alpha}(B_1(0))} \leq C$, with a uniform bound independent of i. Since N is a compact manifold, we also have that $\|u_i\|_{L^\infty(B_3(0))}$ is uniformly bounded. Thus u_i converges to u in the strong $C^{0,\alpha/2}(B_1(0))$ sense.

First of all, we prove that ∇u_i converges to ∇u in the strong $L^2(B_1(0))$ sense. For this purpose, it is sufficient to show that for all $\phi \in C_C^\infty(B_1(0))$,

$$\int_{B_1(0)} |\nabla(u_i - u)|^2 \, \phi \to 0 \,. \tag{2.37}$$

We can split this integral as

$$\int |\nabla(u_i - u)|^2 \, \phi = \int \langle \nabla(u_i - u), \nabla(u_i - u) \rangle \phi = \int \langle \nabla(u_i - u), \nabla u_i \rangle \phi - \int \langle \nabla(u_i - u), \nabla u \rangle \phi \,. \tag{2.38}$$

By weak convergence, the second integral tends to 0 with i. As for the first, we have

$$-\int \langle \nabla(u_i - u), \nabla u_i \rangle \phi = \int \langle u_i - u, \Delta u_i \rangle \phi + \int \langle u_i - u, \nabla u_i \cdot \nabla \phi \rangle = \tag{2.39}$$

$$= \int \langle u_i - u, A(u_i)(\nabla u_i, \nabla u_i) + f_i \rangle \phi + \int \langle u_i - u, \nabla u_i \cdot \nabla \phi \rangle \,. \tag{2.40}$$

Since $\|u_i - u\|_{L^\infty(B_1(0))} \to 0$, it is easy to see that this integral converges to 0 as well. This completes the proof of the strong H^1 convergence.

As for the equation solved by u, we have for all smooth test functions $\phi \in C_C^\infty(B_1(0))$:

$$-\int \phi \Delta(u) = \int \langle \nabla u, \nabla \phi \rangle = \lim_i \int \langle \nabla u_i, \nabla \phi \rangle = \lim_i \int \phi \Delta(u_i) = \quad (2.41)$$

$$= \lim_i \int \phi \left[A(u_i)(\nabla u_i, \nabla u_i) + f_i \right] = \int \phi f + \lim_i \int \phi A(u_i)(\nabla u_i, \nabla u_i). \quad (2.42)$$

Since A is continuous, and since $u_i \to u$ in $C^{\alpha/2}(B_1(0))$, we have $|A(u_i) - A(u)| \to 0$. Here we interpret $A(\cdot)$ as a continuous function on N which takes $\mathbb{R}^m \times \mathbb{R}^m$ into $\mathbb{R}^m$ in a bilinear way. Moreover, the strong convergence of ∇u_i to ∇u allow us to estimate

$$\int \phi A(u_i)(\nabla u_i, \nabla u_i) = \int \phi \left\{ [A(u_i) - A(u)] (\nabla u_i, \nabla u_i) + A(u)(\nabla u_i, \nabla u_i) - A(u)(\nabla u, \nabla u) + A(u)(\nabla u, \nabla u) \right\} = $$

$$(2.43)$$

$$= \int \phi A(u)(\nabla u, \nabla u) + \|A(u_i) - A(u)\|_\infty \int \phi |\nabla u_i|^2 + \int \phi \left[A(u)(\nabla(u_i - u), \nabla u_i) + A(u)(\nabla u, \nabla(u_i - u)) \right].$$

$$(2.44)$$

The strong convergence of ∇u_i to ∇u implies the thesis.

Now we are in a position to prove the original proposition.

Proof (Proof of Proposition 16) The proof is based on the monotonicity formula for u_i and the ϵ-regularity theorem.

Define the set

$$\Sigma = \bigcap_{r > 0} \left\{ x \in \overline{B_1(0)} \ \ s.t. \ \ \liminf_{i \to \infty} \theta_{u_i}(x, r) \geq \epsilon_0 \right\}, \quad (2.45)$$

where ϵ_0 is taken from theorem 14. It is easy to see that this set is closed. Let $B_{r_j}(x_i)$ be a sequence of disjoint balls contained in $B_3(0)$ such that $x_j \in \Sigma$. For each j, there exists a subsequence of u_i (still denoted with the same indexes for the sake of simplicity) such that

$$\theta_{u_i}(x_j, r_j) \geq \epsilon_0 > 0. \quad (2.46)$$

By a diagonal procedure, it is possible to find a subsequence of u_i such that (2.46) is valid for all i and j. This implies immediately that

$$\int_{B_3(0)} |\nabla u_i|^2 \geq \sum_j r_j^{n-2} \theta_{u_i}(x_j, r_j) \geq \epsilon_0 \sum_j r_j^{n-2}, \quad (2.47)$$

as desired. Note that evidently this uniform packing estimate implies upper estimates on the $n - 2$ Minkowski content and Hausdorff measure.

For all $x \notin \Sigma$, we can apply Lemma 19, and obtain that u is a Hölder continuous function in a neighborhood of x solving $\Delta u = A(u)(\nabla u, \nabla u)$. By standard estimates on continuous harmonic maps (see for example [Mos05, theorem 3.1]), u is smooth in a neighborhood of x.

The last thing to check is that u is globally weakly harmonic, and this is a consequence of the fact that u is smooth and harmonic on Σ^C, and Σ has bounded $n - 2$ packing estimates.

This fact implies that Σ has 2-capacity zero, and in particular for all compact sets $K \Subset B_1(0)$, there exists a sequence of smooth functions ϕ_i such that

$$\phi_i(x) = 1 \text{ for } x \text{ in a neighborhood of } \Sigma \cap K, \quad \phi_i(x) = 0 \text{ if } d(x, \Sigma \cap K) > i^{-1}, \quad \lim_{i \to \infty} \int_{B_3(0)} |\nabla \phi_i|^2 = 0.$$

(2.48)

Since u is smooth on Σ^C, we can write for all $\psi \in C_C^\infty(B_1(0))$ that

$$\int \psi \Delta(u) = \int \psi \phi_i \Delta(u) + \int \psi(1 - \phi_i)\Delta(u) = \int \psi \phi_i \Delta(u) + \int \psi(1 - \phi_i)A(u)(\nabla u, \nabla u).$$

(2.49)

As i converges to ∞, ϕ_i converges to 1 a.e. in $B_3(0)$. Thus by dominated convergence

$$\lim_{i \to \infty} \int \psi(1 - \phi_i)A(u)(\nabla u, \nabla u) = \int \psi A(u)(\nabla u, \nabla u).$$

(2.50)

Moreover, we have

$$\limsup_{i \to \infty} \left| \int \psi \phi_i \Delta(u) \right| \leq \limsup_{i \to \infty} \left[\int \phi_i |\nabla \psi \cdot \nabla u| + \int \psi |\nabla \psi_i \cdot \nabla u| \right] = 0.$$

(2.51)

We are left to prove the unique continuation property. Note that if u and v are smooth, then unique continuation follows because $\Delta(u) = A(u)(\nabla u, \nabla u) = A(u)(\nabla u)[\nabla u]$ can be viewed as a linear equation on u with smooth first order coefficient $A(u)(\nabla u)$.

Since u and v are assumed to be smooth outside the close set $\Sigma \cup \Sigma'$, and this close set is $n - 2$ dimensional and thus non-disconnecting, we easily obtain that $u = v$ a.e. on $(\Sigma \cup \Sigma')^C$, and thus on the whole domain.

3 Main theorems

Now we are in a position to state precisely the main results we intend to obtain. The main theorem we want to prove is

Theorem 20 *Let $u : B_3(0) \subseteq \mathbb{R}^m \to N$ an approximately harmonic mapping solving* (W) *and* (S) *with* (f) *such that $3^{2-m} \int_{B_3(0)} |\nabla u|^2 \leq \Lambda$. Then for each $\epsilon > 0$ there exists $C_\epsilon(m, N, \Lambda, \mathcal{F}, \gamma, \epsilon)$ such that for all $r \in (0, 1]$:*

$$Vol\left(B_r\left(\mathcal{S}_{\epsilon,r}^k(u)\right) \cap B_1(0)\right) \leq C_\epsilon r^{n-k}.$$

(3.1)

As a corollary, we can estimate for all $r \in (0, 1]$:

$$Vol\left(B_r\left(\mathcal{S}_\epsilon^k(u)\right) \cap B_1(0)\right) \leq C_\epsilon r^{n-k},$$

(3.2)

moreover $\mathcal{S}_\epsilon^k$ is k-rectifiable.

As a corollary of this theorem we obtain the rectifiability of the strata.

Theorem 21 *Let $u : B_3(0) \subseteq \mathbb{R}^m \to N$ an approximately harmonic mapping solving* (W) *and* (S) *with* (f). *Then for all k the strata $\mathcal{S}^k(u)$ are k-rectifiable.*

It is worth noticing that simple adaptation of the proofs described here allow to obtain a slightly better result. In particular, we can obtain uniform packing estimates instead of Minkowski estimates. Since the proof of this result requires no additional idea, but would make the exposition more technical and confusing, we state the result here and leave the details to the reader.

Theorem 22 *Let $u : B_3(0) \subseteq \mathbb{R}^m \to N$ an approximately harmonic mapping solving* (W) *and* (S) *with* (f) *such that $3^{2-m} \int_{B_3(0)} |\nabla u|^2 \leq \Lambda$. Let $\{B_{r_x}(x)\}_{x \in \mathcal{C}}$ be a collection of pair-wise disjoint balls with $r_x \in (0, 1]$ and for all x, $x \in \mathcal{S}^k_{\epsilon, r_x}$. Then there exists a constant $C_\epsilon(m, N, \Lambda, \mathcal{F}, \gamma, \epsilon)$ such that*

$$\sum_{x \in \mathcal{C}} r_x^n \leq C_\epsilon. \tag{3.3}$$

In the following sections, we will develop all the techniques needed for the proof of these results. First of all, we will give a quantitative link between 0-symmetry (and higher order symmetries) and the properties $\hat{\theta}$. We will then briefly recall without proof the Reifenberg theorems which we will use, and then we will turn to the L^2-best subspace approximation theorem and the covering arguments needed to complete the proofs.

4 Quantitative ϵ-regularity theorems

The aim of this section is to show that there's a quantitative link between $\hat{\theta}$ and the almost symmetries of the map u. First, we are going to show an adaptation of [CN13b, theorem 3.3]. In particular, we will show that if $\hat{\theta}(x, \cdot)$ is sufficiently *pinched* on two consecutive scales (i.e., if $\hat{\theta}(x, r) - \hat{\theta}(x, r/2)$ is small enough), then $B_r(x)$ will be $(0, \delta)$-symmetric.

We will then turn our attention to higher order symmetries. We will show a very natural sufficient condition for $B_r(x)$ to be (k, δ)-symmetric based on the geometry of the "pinched points" $y \in B_r(x)$ such that $\hat{\theta}(y, r) - \hat{\theta}(y, r/2)$ is small. Note that while a single pinched point is enough to guarantee 0-symmetries, we will ask the set of pinched points to "effectively span" some k-dimensional affine subspace in order to guarantee higher order symmetry.

4.1 Quantitative homogeneity

It is easy to see that if u is stationary harmonic and $\hat{\theta}(0, 1) = \hat{\theta}(0, 1/2)$, then u must be 0-symmetric. This is a direct consequence of (2.7) and the unique continuation property for harmonic maps. By an easy compactness argument, we can see that this characterization

of 0-symmetric map has a rigidity property, in the sense that if $\hat{\theta}(0,1) - \hat{\theta}(0,1/2)$ is small enough, then u is close to a 0-symmetric map. In the case of an approximately harmonic map, this statement remains true up to focusing on a small enough scale, so that the error coming from f becomes small enough.

Remark 23 Throughout this section, we will assume that $u \in H^1(B_3(0), N)$ is an approximately harmonic map satisfying (W) and (S) with (f). Moreover, we will also assume the uniform energy bound $\hat{\theta}(0,3) \le \Lambda$.

Proposition 24 *Let $u \in H^1(B_3(0))$ be a solution to (W) and (S) with (f) and $\hat{\theta}(0,3) \le \Lambda$. Then for all $\delta_1 > 0$, there exist $\delta_2 = \delta_2(m, \Lambda, \gamma, \delta_1) > 0$ such that if $\mathcal{F} \le \delta_2$ and for some $x \in B_1(0)$*

$$\hat{\theta}(x,r) - \hat{\theta}(x,r/2) < \delta_2\,, \tag{4.1}$$

then $B_r(x)$ is $(0, \delta_1)$-symmetric.

Proof Consider by contradiction a sequence of maps u_i and a sequence of balls $B_{r_i}(x_i)$ such that $\hat{\theta}(x_i, r_i) - \hat{\theta}(x_i, r_i/2) < i^{-1}$, $\mathcal{F} < i^{-1}$, but such that the balls $B_{r_i}(x_i)$ are not $(0, \delta_1)$-symmetric. Let $T_i(y) = u(x_i + r_i(y))$ be their blow-up maps, and recall that $\hat{\theta}_{T_i}(0, s) = \hat{\theta}_{u_i}(x_i, sr)$. By (2.16), T_i have uniform H^1 bounds, and so there exists a weakly convergence subsequence, for convenience denoted with the same index. Note that T_i are approximately harmonic maps solving

$$\Delta(T_i) = A(T_i)(\nabla T_i, \nabla T_i) + \tilde{f}_i\,, \quad \tilde{f}_i(y) = r_i^2 f_i(x_i + r_i y)\,. \tag{4.2}$$

By (f), we have $\int_{B_1(0)} \left|\tilde{f}_i\right|^2 = r^{4-m} \int_{B_{r_i}(x_i)} |f|^2 \le i^{-1} r_i^\gamma \to 0$. Thus T is weakly harmonic by Proposition 16 and smooth away from a close set Σ of dimension $n-2$.

Moreover, by (2.14), we have the estimate

$$\int_{B_1(0)\backslash B_{1/2}(0)} |y \cdot \nabla T_i|^2 \le c(m)\left[\hat{\theta}_{T_i}(0,1) - \hat{\theta}_{T_i}(0,1/2)\right] \to 0\,. \tag{4.3}$$

Thus T is radially invariant on $B_1(0) \setminus B_{1/2}(0)$, and by unique continuation it is homogeneous. Indeed, let T' be the homogeneous continuation of T over the whole $B_1(0)$. It is easily seen that both maps are weakly harmonic and smooth away from an $n-2$ dimensional set, thus we can apply point (4) of proposition 16 and prove that $T = T'$. Since T_i converges weakly in H^1 to T, and strongly in the L^2 norm, we have reached a contradiction.

Note that, as a corollary of the proof, we obtain a characterization of all tangent maps for approximate harmonic function.

Corollary 25 *Let u be as above, then all of its tangent maps are weakly harmonic homogeneous maps. In particular, for all $x \in B_1(0)$, and for all sequences $r_i \to 0$, there exists a subsequence r_{i_j} such that $T_j = T^u_{x, r_{i_j}}$ converges in the weak $H^1_{loc}(\mathbb{R}^m, N)$ sense to a 0-symmetric weakly harmonic map which is smooth on an open, dense, connected subset.*

We close this section with the characterization of $\mathcal{S}^k(u)$ promised above.

Lemma 26 *Let u be an approximate harmonic map with* (f). *Then*

$$\mathcal{S}^k(u) = \bigcup_{\epsilon>0} \mathcal{S}^k_\epsilon = \bigcup_{\epsilon>0} \bigcap_{r>0} \mathcal{S}^k_{\epsilon,r} = \{x \in \Omega \ \ s.t. \ \ \text{no tangent maps at } x \text{ is } k+1-\text{symmetric}\} .$$

(4.4)

Proof Let $x \in \Omega$. We now know that for all x, there exists at least one (possibly more) tangent map T, and all tangent maps are 0-symmetric.

Suppose that at $x \in \mathcal{S}^k_\epsilon$ for some $\epsilon > 0$. Then since for all $r > 0$, $T^u_{x,r}$ is at least ϵ-apart in the L^2 sense from all $k+1$-symmetric maps, then also any tangent map $T = \lim_i T^u_{x,r_i}$ must be ϵ-apart from all $k+1$ symmetric maps. Thus

$$\bigcup_{\epsilon>0} \mathcal{S}^k_\epsilon \subseteq \{x \in \Omega \ \ s.t. \ \ \text{no tangent maps at } x \text{ is } k+1-\text{symmetric}\} . \quad (4.5)$$

Now if $x \notin \bigcup_{\epsilon>0} \mathcal{S}^k_\epsilon$, by definition for all $i > 0$, there exists some $r_i > 0$ and some $k+1$-symmetric map T_i such that

$$\int_{B_1(0)} \left| T^u_{x,r_i} - T_i \right|^2 \le i^{-1} . \quad (4.6)$$

By passing to a subsequence $T_i \to T$ which is $k+1$-symmetric, we see also that $T^u_{x,r_i} \rightharpoonup^* T$. If there's a subsequence of r_i converging to 0, then T is by definition a tangent map, and it is also $k+1$-symmetric. If r_i is bounded away from 0, then $u = T$ on a ball of positive size around x, and in particular all tangent maps of u at x are equal to T. In either case, we have proved the claim.

4.2 Quantitative higher order symmetries

In order to have higher order symmetries, one point where $\hat{\theta}$ is pinched is not enough. However, if we have more points where $\hat{\theta}$ is pinched, and these points span in some sense a k-dimensional space, this is enough to guarantee higher order symmetries.

Example 27 As a guiding example, consider again the case of a stationary harmonic map. If for two distinct points x_1, x_2 we have $\theta(x_i, 1) - \theta(x_i, 1/2) = 0$, then the map is homogeneous with respect to x and homogeneous with respect to y, which immediately implies that this map is invariant with respect to the line L joining x and y.

Moreover, note also the following. If we take any $z \notin L$, then at z two distinct directional derivatives of u are null, one in the L direction, and another in the direction joining z and x (or y). Now consider a small enough ball $B_r(z)$. On this ball, two directional derivatives are null, and if r is small enough this ball is almost 0-symmetric by monotonicity of $\theta(z, \cdot)$. Thus this ball will be almost 2-symmetric.

In the next lemmas, we will prove in detail quantitative versions of these observations. Before doing that, let us record the definition of a quantitative version of linear independence, which will be used throughout the rest of this section.

Definition 28 Let $y_0, \cdots, y_k \subset B_1(0) \subset \mathbb{R}^m$. We say that these points ρ-effectively span a k-dimensional affine subspace if for all $i = 1, \cdots, k$,

$$y_i \notin B_{2\rho}(y_0 + \text{span}(y_1 - y_0, \cdots, y_{i-1} - y_0)) \ . \tag{4.7}$$

As an immediate consequence, we obtain that $\{y_i\}$ are effectively linear independent, in the sense that for all

$$x \in \{y_0 + \text{span}(y_1 - y_0, \cdots, y_k - y_0)\} \ , \tag{4.8}$$

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

$$x = y_0 + \sum_{i=1}^k \alpha_i(y_i - y_0), \quad |\alpha_i| \le C(m, \rho) \|x - y_0\| \ . \tag{4.9}$$

Moreover, note that this property is preserved under limits, as opposed to the property of being simply linearly independent. Indeed, if we have a sequence $\left\{y_{i,j}\right\}_{i=0}^k$ such that for all j, $\left\{y_{i,j}\right\}_{i=0}^k$ ρ-effectively span a k-dimensional space, and if $\lim_j y_{i,j} = \bar{y}_i$, then also $\{\bar{y}_i\}_{i=0}^k$ ρ-effectively spans a k-dimensional space.

Definition 29 Given $F \subset B_1(0)$, we say that F ρ-effectively spans a k-dimensional subspace if there exist $\{y_0, \cdots, y_k\} \subseteq F$ that ρ-effectively spans a k-dimensional subspace according to the previous definition.

With this concept in mind, we can extend Proposition 24 to the case where we have $k+1$ distinct points of pinching for $\hat{\theta}$. Note that we will not actually use this proposition, indeed we will need the more refined version of this statement, which is the main focus of Section 6.1. However, we think it is reasonable to record this proposition in order to give the reader a better understanding of the direction that our argument is taking.

Proposition 30 Let $u \in H^1(B_3(0))$ be a solution to (W) and (S) with (f) and $\theta(0, 3) \le \Lambda$. Then for all $\epsilon, \rho > 0$, there exist $\delta(m, \Lambda, \gamma, \epsilon, \rho)$ such that if $\mathcal{F} < \delta$ and for some $\{x_i\}_{i=0}^k \subset B_1(0)$ we have

1. $\{x_i\}$ ρ-effectively spans a k-dimensional affine subspace V,
2. $\hat{\theta}(x_i, r) - \hat{\theta}(x_i, r/2) < \delta$ for all i,

then $B_r(x)$ is (k, ϵ)-symmetric.

Proof The proof is a simple adaptation of Proposition 24. □

The next proposition shows that in case when the set of points with pinching effectively spans a k-dimensional plane V, then $\mathcal{S}_{\epsilon,r}^k$ is contained in a tube around V.

Proposition 31 Let $\rho, \epsilon > 0$ be fixed. There exists $\delta_3(m, \Lambda, N, \gamma, \rho, \epsilon) > 0$ such that the following holds. Let $F = \left\{y \in B_2(0) \ \text{s.t.} \ \hat{\theta}(y, 1) - \hat{\theta}(y, \rho) < \delta_3\right\}$. If $\mathcal{F} \le \delta_3$ and F ρ-effectively spans a k-dimensional subspace V, then

$$\mathcal{S}_{\epsilon, \delta_3}^k \subseteq B_{2\rho}(V) \ . \tag{4.10}$$

We split the proof of this proposition into two parts, the first of which is contained in the following technical lemma.

Lemma 32 *Let u be as above. There exits a $\delta_4(m, \Lambda, N, \gamma, \rho, \epsilon) > 0$ such that if $\mathcal{F} < \delta_4$ and*

$$\int_{B_1(0)} |P \cdot \nabla u|^2 < \delta_4 , \tag{4.11}$$

for some $k + 1$ dimensional linear subspace P, then $\mathcal{S}^k_{\epsilon, \bar{r}} \cap B_{1/2}(0) = \emptyset$, where $\bar{r} = \delta_4^{\frac{1}{2(m-2)}}$.

Remark 33 Notationally we define $|P \cdot \nabla u|^2 \equiv \sum |\nabla_{e_i} u|^2$, where e_i is an orthonormal basis of P.

Proof We want to show by contradiction that for all $x \in B_{1/2}(0)$, there exists an $r_x \in [\bar{r}, 1/2]$ such that $B_{r_x}(x)$ is $(k + 1, \epsilon)$-symmetric. Note that, by monotonicity, for all $x \in B_{1/2}(0)$ there exists an $r_x \in [\bar{r}, 1/2]$ such that

$$\hat{\theta}(x, r_x) - \hat{\theta}(x, r_x/2) < \frac{C_1(m, \gamma)\Lambda}{-\log(\delta_4)} . \tag{4.12}$$

Indeed, otherwise we would have

$$c(m, \gamma)(\Lambda + \delta_4) \stackrel{(2.16)}{\geq} \hat{\theta}(x, 1/2) = \sum_{i=1}^{-\log(\bar{r})+1} \left(\hat{\theta}(x, 2^{-i}) - \hat{\theta}(x, 2^{-i-1}) \right) \stackrel{(2.15)}{\geq} c(m)C_1(m, \gamma)\Lambda , \tag{4.13}$$

which is impossible if we set $C_1(m, \gamma) = 2c(m, \gamma)c(m)^{-1}$. Moreover, note that

$$r_x^{2-n} \int_{B_{r_x}(x)} |P \cdot \nabla u|^2 \leq \delta_4 r_x^{2-n} \leq \delta_4^{1/2} \tag{4.14}$$

by definition of r_x. Thus consider a contradicting sequence u_i with $\delta_{4,i} \to 0$ such that there exists $x_i \in B_{1/2}(0)$ and $r_i \in [\bar{r}, 1]$ such that $B_{r_i}(x_i)$ is not $(k + 1, \epsilon)$-symmetric but such that (4.12) and (4.14) are valid. By a simple rotation, we can assume that the $k + 1$ dimensional subspace P is fixed throughout the sequence. Consider the maps $T_i = T_{x_i, r_i}^{u_i}$. Their weak limit converges to some weakly harmonic T which is 0-symmetric by unique continuation, (4.12) and Proposition 24. Moreover, T is also invariant wrt the $k+1$-dimensional P by (4.14). This clearly is a contradiction.

Now we turn to the proof of the main proposition.

Proof (Proof of Proposition 31) Let $\{y_0, \cdots, y_k\} \subseteq F$ ρ-effectively span the k-dimensional subspace V, and consider any $x \in B_2(0) \setminus B_{2\rho}(V)$. Note that for all $i = 0, \cdots, k$, we have

$$B_\rho(x) \subset B_2(y_i) \setminus B_\rho(y_i) . \tag{4.15}$$

By (2.14), we obtain that for all $i = 0, \cdots, k$:

$$\int_{B_\rho(x)} |(z - y_i) \cdot \nabla u|^2 (z) \leq c(m) \left[\hat{\theta}(y_i, 1) - \hat{\theta}(y_i, \rho) \right] . \tag{4.16}$$

Let w be any norm 1-vector in $\hat{V}$, the linear subspace associated to the affine V. By definition of $\{y_i\}$, there exists $\{\alpha_i\}_{i=1}^k$ such that

$$w = \sum_i \alpha_i(y_i - y_0) \qquad |\alpha_i| \leq c(m, \rho). \tag{4.17}$$

Thus we also have

$$\int_{B_\rho(x)} |w \cdot \nabla u|^2\,(z) \leq C(m, \rho) \sum_i \int_{B_\rho(x)} |[(y_i - z) + (z - y_0)] \cdot \nabla u|^2\,(z) \leq \tag{4.18}$$

$$\leq c(m, \rho) \sum_i \left[\hat{\theta}(y_i, 1) - \hat{\theta}(y_i, \rho) \right] \leq c(m, \rho, \gamma)\delta_3. \tag{4.19}$$

This in particular implies

$$\int_{B_\rho(x)} |V \cdot \nabla u|^2 \leq c(m, \rho, \gamma)\delta_3. \tag{4.20}$$

In order to gain one more independent direction along which the energy is small, set for all $z \in B_1(0) \setminus V$

$$h(z) = \frac{z - \pi_V(z)}{\|z - \pi_V(z)\|}. \tag{4.21}$$

By an argument similar to the one above, we obtain also

$$\int_{B_\rho(x)} |h(z) \cdot \nabla u|^2\,(z) \leq c(m, \rho, \gamma)\delta_3. \tag{4.22}$$

By a geometric argument, it is easy to see that if $d(z, V) \geq \rho$, then $|h(z) - h(x)| \leq C(m)\,|x - z|\,\rho^{-1}$. This implies that for all $r \leq \rho$, we have

$$\frac{1}{2}\int_{B_r(x)} |h(x) \cdot \nabla u|^2\,(z) \leq \int_{B_r(x)} |h(z) \cdot \nabla u|^2\,(z) + \int_{B_r(x)} |h(x) - h(z)|^2\,|\nabla u|^2\,(z) \leq c(m, \rho, \gamma)\delta_3 + c(m, \rho, \gamma, \Lambda)r^m,$$
$$\tag{4.23}$$

where in the last line we have used the uniform bound $\theta(x, r) \leq c(m, \gamma)(\Lambda + 1)$ for all $x \in B_2(0)$ and $r \in [0, 1]$ (see the estimates in (2.16)).

For the $k + 1$-dimensional linear subspace $P = \hat{L} \oplus h(x)$, and for all $r \leq \rho$, we get the estimate

$$r^{2-m}\int_{B_r(x)} |P \cdot \nabla u|^2 \leq c(m, \rho, \gamma)\delta_3 r^{2-m} + C(m, \rho, \gamma, \Lambda)r^2. \tag{4.24}$$

Now we can choose r and subsequently $0 < \delta_3 << \delta_4$ small enough in order to apply Lemma 32 to $B_r(x)$, and we obtain the thesis.

We close this section by observing that if $B_r(x)$ is not $(k+1, \epsilon)$-symmetric, but it is almost 0-symmetric, then necessarily u must have some energy on *any* $k + 1$ distinct directions on $B_r(x)$. Actually, part of this energy must be concentrated on the annulus $B_r(x) \setminus B_{r/2}(x)$.

Lemma 34 *Let u be as above. Then for each $\epsilon > 0$ there exists $\delta_5(m, N, \Lambda, \gamma, \epsilon) > 0$ such that if $\mathcal{F} < \delta_5$, and $B_1(0)$ is $(0, \delta_5)$-symmetric but is not $(k + 1, \epsilon)$-symmetric, then for every $k + 1$ linear subspace P we have*

$$\int_{A_{3/8,1/2}(0)} |P \cdot \nabla u|^2 \geq \delta_5\,, \tag{4.25}$$

where $A_{3/8,1/2}(0) \equiv B_{1/2}(0) \setminus B_{3/8}(0)$.

Proof Also this proof is based on a simple contradiction argument which hinges on the H^1_{weak} compactness given by the uniform energy bounds. Thus let u_i be a contradicting sequence. In particular, let u_i be approximately harmonic maps with $\mathcal{F}_i \leq i^{-1}$ such that $B_1(0)$ is $(0, i^{-1})$-symmetric but not $(k + 1, \epsilon)$-symmetric. Moreover, let P be a $k + 1$-dimensional subspace such that for all i

$$\fint_{A_{3/8,1/2}(0)} |P \cdot \nabla u_i|^2 \leq i^{-1} \to 0\,. \tag{4.26}$$

Note that we can assume that P is not changing with i simply by making a rotation in the domain space. After passing to a subsequence, $u_i \rightharpoonup^* u$ in the weak H^1 sense, with u being a weakly harmonic map. Now (4.26) and the H^1_{weak} convergence guarantee that

$$\int_{A_{3/8,1/2}(0)} |P \cdot \nabla u|^2 = 0\,. \tag{4.27}$$

Moreover, u will be 0-homogeneous by unique continuation and the pinching assumption. Thus, we obtain that

$$\int_{B_1(0)} |P \cdot \nabla u|^2 = 0\,, \tag{4.28}$$

and this implies that u is $k + 1$-symmetric. Since u_i converges strongly to u in $L^2(B_1(0))$, we obtain a contradiction.

4.3 Uniformity of the energy and of the non-symmetry

One moral to be taken from the previous section, in particular from Proposition 30, is that if we have a lot of pinched points which span something k-dimensional, then the map u is almost constant along these k-directions.

Here we prove another two important variations of this general philosophical point. First of all, we will show that in the situation described above, u is almost constant also in some H^1 sense, not just in an L^2 sense. In particular $\hat{\theta}$ remains almost constant on all pinched points.

Lemma 35 *Let $u : B_3(0) \to N$ be a solution to (W) and (S) with (f). Let $\rho > 0$ and $\eta > 0$ be fixed, and assume that for all $y \in B_1(0)$, $\hat{\theta}(y, 1) \leq E$, then there exists $\delta_6 = \delta_6(m, \Lambda, N, \rho, \gamma, \eta)$*

such that if $\mathcal{F} \leq \delta_6$, *the following holds. Let* $F(u, \delta_6) \subseteq \left\{ y \in B_1(0) \quad s.t. \quad \hat{\theta}(y, \rho) > E - \delta_6 \right\}$. *If* F ρ-*effectively spans a* k-*dimensional subspace* L, *then for all* $x \in L \cap B_2(0)$, *we have*

$$\hat{\theta}(x, \rho) > E - \eta. \tag{4.29}$$

Moreover, if $k \geq m - 1$, *then* $E \leq \eta$.

Proof We prove this statement by contradiction. Fix any $\eta > 0$, and let u_i be a contradicting sequence. In particular, u_i will be approximately harmonic maps on $B_3(0)$ with $\theta_{u_i}(y, 1) \leq E$ for all i, and such that $F_i = F(u_i, i^{-1})$ ρ-effectively spans a k-dimensional subspace L. Up to translations and rotations, we can assume that L is fixed for all i.

Moreover, we assume by contradiction that there exists some $x_i \in L$ such that $\hat{\theta}_{u_i}(x_i, \rho) \leq E - \epsilon$. Note that, evidently, we can assume that L has dimension at least 1, otherwise there's nothing to prove.

By weak compactness, we can pass to a subsequence and have $u_i \rightharpoonup^* u$, where u is a weakly harmonic map, and

$$|\nabla u_i|^2 \, d\mathrm{Vol} \rightharpoonup^* \mu \equiv |\nabla u|^2 \, d\mathrm{Vol} + \nu, \tag{4.30}$$

where the convergence is in the weak sense of measures, and ν is the defect measure, which is nonnegative by Fatou's lemma. Set by definition $\theta_\mu(y, r) = r^{2-n}\mu(B_r(y))$, and note that θ_μ is monotone in r for all y fixed. Indeed, let $0 < r_1 < r_2$, and consider that

$$\theta_\mu(y, r_2) - \theta_\mu(y, r_1) = \lim_{i \to \infty} \left[\theta_{u_i}(y, r_2) - \theta_{u_i}(y, r_1) \right] \overset{(2.15)}{=} \lim_{i \to \infty} \left[\hat{\theta}_{u_i}(y, r) - \hat{\theta}_{u_i}(y, r_1) \right] \geq 0 \quad (4.31)$$

Now let $\left\{ y_{i,j} \right\}_{j=0}^{k} \subset F_i$ be a set of points which ρ-effectively spans L. By passing to a subsequence, we can assume that $x_i \to x \in L$ and $\lim_{i \to \infty} y_{i,j} = \bar{y}_j$, where $\left\{ \bar{y}_j \right\}$ ρ-effectively spans L.

For all j and $r > 0$, we have that

$$\left(\frac{r - |y_i - y|}{r} \right)^{n-2} \theta_{u_i}(y_i, r - |y_i - y|) \leq \theta_{u_i}(y, r) \leq \left(\frac{r}{r + |y_i - y|} \right)^{n-2} \theta_{u_i}(y_i, r + |y_i - y|), \quad (4.32)$$

thus by taking the limit on i, we can conclude that for all j:

$$\theta_\mu(y_j, 1) - \theta_\mu(y_j, \rho) = 0. \tag{4.33}$$

By an easy adaptation of [Lin99, lemma 1.7] (in particular, by the proof of point *ii* in this lemma, carried out at pages 797–800), this implies that μ, ν and u are radially invariant on $B_1(y_j) \setminus B_\rho(y_j)$ for all j. Since y_j ρ-effectively span L, as an immediate consequence we have that u, ν and μ are invariant wrt L on the whole $B_2(0)$.

Thus $\theta_\mu(y, \rho) = E$ for all $y \in L$, in particular $\theta_\mu(x, \rho) = E$. Since $x_i \to x$ and $\rho > 0$, we obtain our contradiction by (4.32).

As a last point, if we assume that L has dimension at least $m - 1$, then we know that both ν and u are invariant wrt an $m - 1$ dimensional affine subspace. However, by point 2 in Proposition 16, the support of ν must have finite $m-2$ dimensional Hausdorff measure. Thus,

we see that $v = 0$, which means that u_i converges strongly in H^1 to u, which is an $m - 1$-invariant weakly harmonic map. In other words, u is a weakly harmonic map depending only on 1 variable, and by standard Sobolev embeddings those maps are well-known to be continuous. Since u is also 0-symmetric wrt y_0, it is a constant map. This proves the last claim.

We close this section with the last technical result we need. In particular, we prove that almost symmetry (or lack thereof) is preserved under some suitable pinching condition.

Lemma 36 *Let $u : B_8(0) \to N$ be a solution to (W) and (S) with (f). Let $\rho > 0$ and $\epsilon > 0$ be fixed. There exists $\delta_7 = \delta_7(m, \Lambda, N, \rho, \gamma, \epsilon)$ such that if $\mathcal{F} \leq \delta_7$, then the following holds. If $\hat{\theta}(0, 1) - \hat{\theta}(0, 1/2) < \delta_7$ and there exists a point $y \in B_3(0)$ with*

1. *$\hat{\theta}(y, 1) - \hat{\theta}(y, 1/2) < \delta_7$,*
2. *for some $r \in [\rho, 2]$, $B_r(y)$ is not $(k + 1, \epsilon)$-symmetric*

Then $B_r(0)$ is not $(k + 1, \epsilon/2)$-symmetric.

Remark 37 Note that we do not assume that $|y - 0| \geq \rho$, and thus we cannot obtain in general that u will be invariant on the line joining 0 and y.

Proof Consider again a contradicting sequence. In particular, u_i will be a sequence of maps with $\hat{\theta}(0, 1) - \hat{\theta}(0, 1/2) \leq i^{-1}$, and for some $y_i \in B_3(0)$ we will have $\hat{\theta}(y_i, 2) - \hat{\theta}(y_i, 1/2) \leq i^{-1}$. Moreover, $B_r(y_i)$ is not $(k + 1, \epsilon)$-symmetric, but $B_r(0)$ is $(k + 1, \epsilon/2)$-symmetric.

Thus, for all i, there exists an $h_i : B_{10}(0) \to N$ which is $k + 1$-symmetric such that

$$\fint_{B_r(0)} |u_i - h_i|^2 \leq \epsilon/2 \,. \tag{4.34}$$

By passing to a subsequence if necessary, we obtain $y_i \to y \in \overline{B_3(0)}$, $u_i \rightharpoonup^* u$ in $H^1(B_8(0))$ and thus $u_i \to u$ in the strong $L^2(B_8(0))$ sense, with u weakly harmonic and homogeneous wrt 0 and y. Moreover, $h_i \rightharpoonup^* h$ in L^2, where h is a $k + 1$-symmetric map with

$$\fint_{B_r(0)} |u - h|^2 \leq \limsup_{i \to \infty} \fint_{B_r(0)} |u - h_i|^2 \leq 2\epsilon/3 \,. \tag{4.35}$$

If $y = 0$, then obviously we obtain a contradiction. Indeed, since N is compact, we have

$$\lim_{i \to \infty} \fint_{B_r(y_i)} |u_i - h|^2 = \fint_{B_r(0)} |u - h|^2 \leq 2\epsilon/3 \,. \tag{4.36}$$

In a similar way, if $y \neq 0$, then u is invariant wrt the line joining 0 and y, and we get

$$\lim_{i \to \infty} \left(\fint_{B_r(0)} |u_i(y_i + z) - h(z)|^2 \, dz \right)^{1/2} \leq \lim_{i \to \infty} \left(\fint_{B_r(0)} |u_i(y_i + z) - u(y + z)|^2 \, dz \right)^{1/2} + (2\epsilon/3)^{1/2} = (2\epsilon/3)^{1/2} \,. \tag{4.37}$$

5 Reifenberg theorem

In this section, we quote without proof the appropriate Reifenberg results from [NV] that are needed in this paper. These results have since been extended and improved upon, with simplified proofs, in [ENV16]. Before doing that, we write a brief introduction to these results, and state the necessary definitions.

5.1 Reifenberg theorem in literature

The Reifenberg topological disk theorem, introduced in [Rei60], states that if a subset $S \subset \mathbb{R}^n$ is sufficiently close in the Hausdorff sense at all scales to a k-dimensional plane, then S is $C^{0,\alpha}$-homeomorphic to a disk.

This theorem has been improved during the years, with the objective of obtaining some $C^{0,1}$ information on S. For example, Toro proves in [Tor95] that the correspondence is $C^{0,1}$ under some summability assumption on Jones' β_2 numbers for the set S. More general results along the same line are available in [DS93, DT12].

It is worth mentioning that just by working with β_2-numbers, without the Reifenberg-type techniques, rectifiability results for sets and measures similar to the ones discussed have been obtained very recently in [AT15, Tol15].

In [NV] and [ENV16], effective estimates in the form of upper Ahlfor's regularity and rectifiability are obtained for sets and measures under the appropriate Dini conditions on the β_2-numbers. These results play a key role in the finiteness and structure theorems of this paper. Here we quote these theorems, and refer the reader to [NV] and [ENV16] for their proofs.

5.2 Definitions

Here we define the so-called Jones' β_2 numbers.

Definition 38 Let μ be a nonnegative Radon measure on $B_3(0)$, and fix any $r > 0$ and $k \in \mathbb{N}$. We define the k-dimensional Jones' β number by

$$\beta_{2,\mu}^k(x,r)^2 = \min_{V \subseteq \mathbb{R}^n} \int_{B_r(x)} \frac{d^2(y,V)}{r^2} \frac{d\mu(y)}{r^k}, \tag{5.1}$$

where the minimum is taken over all affine subspaces V of dimension k.

It is clear that β_2 is suitable to quantify how close the support of μ is to a k-dimensional subspace. Note that the scaling factor r^{-2-k} in the definition of μ is chosen to make β_2 "scale invariant" in some sense. Indeed, r^{-2} takes care of the scaling properties of $d(x,V)^2$, and since we expect μ to behave like a k-dimensional measure, $r^{-k}\mu(B_r(x))$ is the right scale invariant quantity to consider. This is the case if, for example, μ if k-Ahlfors regular, in the sense that for all $x \in \text{supp}(\mu) \cap B_1(0)$ and $r \leq 1$

$$C^{-1}r^k \leq \mu(B_r(x)) \leq Cr^k, \tag{5.2}$$

or if μ is the k-dimensional Hausdorff measure on any set S. It is worth recording two basic properties of β.

Lemma 39 β_2^2 *is monotone in* μ, *in the sense that if* $\mu \leq \nu$, *then for all* x, r

$$\beta_{2,\mu}^k(x,r)^2 \leq \beta_{2,\nu}^k(x,r)^2 . \tag{5.3}$$

Moreover, if $|x - y| \leq r$, *then*

$$\beta_{2,\mu}^k(x,r)^2 \leq 2^{k+2}\beta_{2,\mu}^k(y,2r)^2 . \tag{5.4}$$

Proof Both these properties are immediate consequences of the definition of β.

5.3 Generalized Reifenberg Theorems

Now we are ready to state the two versions of the quantitative Reifenberg theorems from [NV] that we will use to prove the uniform volume bounds on $\mathcal{S}_{\epsilon,r}^k$.

Theorem 40 *[NV, theorem 3.4] For some constants* $\delta_R(m)$ *and* $C_R(m)$ *depending only on the dimension* m, *the following holds. Let* $\{B_{r_x/5}(x)\}_{x \in \mathcal{C}} \subseteq B_3(0) \subset \mathbb{R}^m$ *be a collection of pairwise disjoint balls with their centers* $x \in B_1(0)$, *and let* $\mu \equiv \sum_{x \in \mathcal{C}} \omega_k r_x^k \delta_x$ *be the associated measure. Assume that for each* $B_r(x) \subseteq B_2$

$$\int_{B_r(x))} \left(\int_0^r \beta_{2,\mu}^k(y,s)^2 \frac{ds}{s} \right) d\mu(y) < \delta_R^2 r^k . \tag{5.5}$$

Then we have the uniform estimates

$$\sum_{x \in \mathcal{C}} r_x^k < C_R(m) . \tag{5.6}$$

Remark 41 See [ENV16] for a more recent generalization of the above.

This theorem will be used on some carefully chosen discrete approximation of the singular set of u. In order to guarantee the assumption (5.5), we will use the β_2 estimates of Section 6.1. The next result is the rectifiable Reifenberg from [NV], see also [ENV16]:

Theorem 42 *[NV, theorem 3.3] There exists* $\delta_R(m) > 0$ *such that the following holds. Let* $S \subseteq B_3(0) \subseteq \mathbb{R}^n$ *be a* λ^k-*measurable subset, and assume for each* $B_r(x)$ *with* $x \in B_1(0)$ *and* $r \leq 1$ *we have*

$$\int_{S \cap B_r(x)} \left(\int_0^r \beta_{2,\lambda^k|_S}^k(y,s)^2 \frac{ds}{s} \right) d\lambda^k(y) < \delta_R^2 r^k . \tag{5.7}$$

Then S *is* k-*rectifiable such that for each* $x \in S$ *we have* $\lambda^k(B_r(x)) < C_R r^k$.

Remark 43 The basis for the ideas in the above are a technical refinement of a new $W^{1,p}$-Reifenberg which is proved by the authors in [NV]. Since we do not directly need this $W^{1,p}$-Reifenberg we do not state it here.

Remark 44 Note that [AT15, theorem 1.1] proved the above without the Ahlfor's upper bound. See also [ENV16] for a more recent generalization, which both applies to a much more general class of measures, and does so under weaker assumptions.

6 Proof of the main theorems

In order to be in a position to prove the main theorem, we need to obtain two important ingredients. First we will discuss estimates linking the β_2 of a generic measure μ with support contained in $\mathcal{S}^k_{\epsilon,r}(u)$ and the monotone quantity $\hat{\theta}$, and later on we will describe a covering argument that will allow us to split the covering of the set $\mathcal{S}^k_{\epsilon,r}(u)$ into suitable pieces with nice estimates.

6.1 L^2 subspace approximation theorem

The aim of this section is to prove that the β_2 defined in the previous sections can be controlled using the monotone quantities θ and $\hat{\theta}$, and the parameters $\mathcal{F}$ and γ.

In order to ease the notation, we define for $x \in B_1(0)$ and $r > 0$ the quantity

$$W_r(x) \equiv W_{8r,r}(x) \equiv \int_{B_{8r}(x) \setminus B_r(x)} \frac{|(y - x) \cdot \nabla u(y)|^2}{|y - x|^m} d\mathrm{Vol}(y) \geq 0. \qquad (6.1)$$

Note that for an approximate harmonic map, by (2.14) we have the bound

$$W_r(x) \leq \hat{\theta}(x, 8r) - \hat{\theta}(x, r). \qquad (6.2)$$

Note that, at least philosophically, bounds on $W_r(x)$ and $W_r(y)$, for $|x - y| \leq r$, give bounds on the scale-invariant L^2 norm of $(x - y) \cdot \nabla u$ in an annulus around x and y. This is an easy consequence of the fact that for all $z \in \mathbb{R}^m$, the vectors $(z - y)$ and $(z - x)$ always span the vector $(x - y)$. In this section, we will exploit this simple idea and some easy tricks to prove β_2 estimates in a very general setting.

The main estimate in this section is the following. Note that, up to minor technical details, this theorem is similar to [NV, theorem 7.1].

Theorem 45 *Let u be as above, and fix $\epsilon > 0$, $0 < r \leq 1$ and $x \in B_1(0)$. There exists a constant $C_1(m, N, \Lambda, \gamma, \epsilon) > 0$ such that if $\mathcal{F} \leq \delta_5$ and $B_{8r}(x)$ is $(0, \delta_5)$-symmetric but not $(k + 1, \epsilon)$-symmetric, then for any nonnegative finite measure μ on $B_r(x)$ we can estimate*

$$\beta^k_{2,\mu}(x, r)^2 = \inf_V r^{-2-k} \int_{B_r(x)} d^2(x, V) \, d\mu(x) \leq C_1 r^{-k} \int_{B_r(x)} W_r(x) \, d\mu(x), \qquad (6.3)$$

where the inf *is taken over all k-dimensional affine subspaces $V \subseteq \mathbb{R}^m$.*

Remark 46 As it will be clear in the proof, δ_5 here is the same as the one given by Lemma 34. Moreover, $C_1 = c(m)\delta_5^{-1}$.

Remark 47 Note that the quantity on the rhs of this theorem can be easily estimated in terms of $\hat{\theta}$ by (2.14).

The proof of this theorem hinges on some manipulations over the eigenvalues and eigenvectors of the "inertia matrix" associated to every measure μ.

6.1.1 Eigenvalue and eigenvectors of the matrix associated to a measure

Let us consider a probability measure μ with support in $B_1(0)$, and let x_{cm} be its center of mass, i.e.:

$$x_{cm} = x_{cm}(\mu) \equiv \int x \, d\mu(x) \,. \tag{6.4}$$

Consider the bilinear quadratic form $Q(v, w)$ defined by

$$Q(v, w) \equiv \int [(x - x_{cm}) \cdot v] \, [(x - x_{cm}) \cdot w] \, d\mu(x) \,. \tag{6.5}$$

In this section, we study the eigenvalue and eigenvectors of Q and their relations with the β_2 defined above.

Definition 48 Given a probability measure $\mu \in B_1(0)$, we set $\lambda_1(\mu), \cdots, \lambda_m(\mu)$ to be the eigenvalues of $Q(\mu)$ in decreasing order, and $v_1(\mu), \cdots, v_m(\mu)$ to be its eigenvectors. In case one eigenvalue has higher multiplicity, we take any choice of orthonormal eigenvectors inside the eigenspace.

Note that by definition of eigenvectors, we have

$$Q(v_k) = \lambda_k v_k = \int [(x - x_{cm}) \cdot v_k] \, (x - x_{cm}) \, d\mu(x) \tag{6.6}$$

We also have a variational characterization of the eigenvalues given by

$$\lambda_1 = \lambda_1(\mu) \equiv \max_{|v|^2=1} \int |(x - x_{cm}) \cdot v|^2 \, d\mu(x) \,. \tag{6.7}$$

and $v_1 = v_1(\mu)$ is any of the norm 1 vectors obtaining this maximum. By induction, we also have

$$\lambda_{k+1} = \lambda_{k+1}(\mu) \equiv \max \left\{ \int |(x - x_{cm}) \cdot v|^2 \, d\mu(x) \ \ s.t. \ \ |v|^2 = 1 \text{ and } \forall i \le k, \ v \cdot v_i = 0 \right\}, \tag{6.8}$$

and v_{k+1} is a vector obtaining this maximum. Note that, by definition of v_k, the subspace $V_k = x_{cm} + \text{span}\{v_1, \ldots, v_k\}$ is the k-dimensional affine subspace (or one of the subspaces) achieving the minimum in the β_2. In other words

$$\min_{V \subseteq \mathbb{R}^m, \ \dim(V)=k} \int d^2(x, V) \, d\mu(x) = \int d^2(x, V_k) \, d\mu(x) = \lambda_{k+1}(\mu) + \cdots + \lambda_m(\mu) \,. \tag{6.9}$$

Remark 49 Note that evidently V_k must pass through the center of mass of μ. This is an immediate corollary of Jensen's inequality (or Steiner's theorem).

By simple manipulations with λ_k and v_k, we obtain the following important estimate:

Proposition 50 *Let $u : B_9(0) \to N$ be an H^1 map, and let μ be a probability measure on $B_1(0)$ with $\lambda_k(\mu), v_k(\mu)$ defined as in Definition 48. Then there exists $C(m) > 0$ such that*

$$\lambda_k \int_{A_{3,4}(0)} |v_k \cdot \nabla u(z)|^2 \, dVol(z) \leq C(m) \int W_0(x) \, d\mu(x) \,. \tag{6.10}$$

Proof For simplicity, and without essential loss of generality, we assume that $x_{cm}(\mu) = 0$ (otherwise a simple translation will do the trick).

For any $z \in A_{3,4}$ and $k = 1, \cdots, m$, we take the scalar product of (6.6) with $\nabla u(z)$, and obtain

$$\lambda_k \, (v_k \cdot \nabla u(z)) = \int (x \cdot v_k) \, (\nabla u(z) \cdot x) \, d\mu(x) \,, \tag{6.11}$$

By definition of center of mass (see (6.4)), we have for all fixed z:

$$\int x \cdot z \, d\mu(x) = x_{cm} \cdot z = 0 \,. \tag{6.12}$$

Thus we can re-write (6.11) in the form:

$$\lambda_k \, (\nabla u(z) \cdot v_k) = \int (x \cdot v_k) \, [\nabla u(z) \cdot (x - z)] \, d\mu(x) \,. \tag{6.13}$$

A simple application of Hölder inequality tells us that for all fixed z:

$$\lambda_k^2 \, |\nabla u(z) \cdot v_k|^2 \leq \lambda_k \int |\nabla u(z) \cdot (x - z)|^2 \, d\mu(x) \,. \tag{6.14}$$

Note that we can evidently assume $\lambda_k > 0$, otherwise there's nothing to prove. By integrating both sides of the previous inequality on $A_{3,4}(0)$, we get

$$\lambda_k \int_{A_{3,4}(0)} |\nabla u(z) \cdot v_k|^2 \, d\text{Vol}(z) \leq \int \int_{A_{3,4}(0)} |\nabla u(z) \cdot (x - z)|^2 \, d\text{Vol}(z) \, d\mu(x) \leq \tag{6.15}$$

$$\leq \int \int_{A_{3,4}(0)} \frac{|\nabla u(z) \cdot (x - z)|^2}{|x - z|^m} \, |x - z|^m \, d\text{Vol}(z) \, d\mu(x) \leq \tag{6.16}$$

$$\leq C(m) \int \int_{A_{1,8}(x)} \frac{|\nabla u(z) \cdot (x - z)|^2}{|x - z|^m} \, d\text{Vol}(z) \, d\mu(x) \leq C(m) \int W_0(x) \, d\mu(x) \,, \tag{6.17}$$

as desired.

6.1.2 Proof of Theorem 45

We are now in a position to prove Theorem 45. By rescaling (6.3), we can assume for convenience that $\mu(B_1(0)) = 1$. Since we have ordered λ_k to be decreasing in value, and by (6.9), we have

$$\beta_{2,\mu}^k(0, 1)^2 = \lambda_{k+1}(\mu) + \cdots + \lambda_m(\mu) \leq (m - k)\lambda_{k+1} \,. \tag{6.18}$$

By applying Proposition 50 to each $j = 1, \cdots, k + 1$, we obtain

$$\sum_{j=1}^{k+1} \lambda_j \int_{A_{3,4}(0)} |\nabla u(z) \cdot v_j|^2 \, d\text{Vol}(z) \leq (k + 1)C \int W_0(x) \, d\mu(x). \qquad (6.19)$$

Let $V^{k+1} = \text{span}(v_1, \cdots, v_{k+1})$ be the linear part of the best $k + 1$-dimensional subspace of μ. Given that λ_j are decreasing in j (by definition), the last estimate leads to

$$\lambda_{k+1} \int_{A_{3,4}(0)} |V^{k+1} \cdot \nabla u(z)|^2 \, d\text{Vol}(z) = \lambda_{k+1} \sum_{j=1}^{k+1} \int_{A_{3,4}(p)} |\nabla u(z) \cdot v_j|^2 \, d\text{Vol}(z) \leq C \int W_0(x) \, d\mu(x).$$
$$(6.20)$$

By assumption, we know that $B_8(0)$ is $(0, \delta_5)$-symmetric and not $(k + 1, \epsilon)$-symmetric. Thus, by Lemma 34,

$$\int_{A_{3,4}(p)} |\nabla u(z) \cdot V^{k+1}|^2 \, d\text{Vol}(z) \geq 8^{m-2} \delta_5. \qquad (6.21)$$

This allows us to estimate

$$c(m)\delta_5 \lambda_{k+1} \leq \lambda_{k+1} \int_{A_{3,4}(0)} |\nabla u(z) \cdot V^{k+1}|^2 \, d\text{Vol}(z) \leq C \int W_0(x) \, d\mu(x). \qquad (6.22)$$

Since δ_5 is a positive constant depending only on $(m, N, \Lambda, \gamma, \epsilon)$, and by (6.18), we can conclude

$$\beta_{2,\mu}^k(0, 1)^2 \leq C(m, N, \Lambda, \gamma, \epsilon) \int W_0(x) \, d\mu(x) \qquad (6.23)$$

as desired.

$\square$

6.2 Covering argument

In this subsection, we prove the inductive covering argument needed for the main theorem. We split this covering argument into two lemmas: in the first one, we keep refining inductively a covering by balls until all but a controlled amount of points in our balls have some definite drop in $\hat{\theta}$, and in the second one we show that this controlled amount of points without drop is small so that they can be "ignored".

Lemma 51 (Covering Lemma I) *Let $u : B_3(0) \to N$ be an approximately harmonic map satisfying (W) and (S), with the conditions (f). Fix any $\epsilon > 0$, $0 < \rho < \rho(m) \leq 100^{-1}$, and $0 < r < R$, $0 < R \leq 1$ arbitrary, set $E = \sup_{x \in B_{2R}(0) \cap S} \hat{\theta}_1(x)$, and assume the uniform bound $E \leq \Lambda$. There exists $\delta = \delta(m, N, \Lambda, \gamma, \rho, \epsilon) > 0$ and $C_V(m)$ such that the following is true.*

If $\mathcal{F} < \delta$ then for any subset $S \subseteq S_{\epsilon, \delta r}^k$ there exists a finite covering of $S \cap B_R(0)$ such that

$$S \cap B_R(0) \subseteq \bigcup_{x \in \mathcal{C}} B_{r_x}(x) \quad \text{with} \quad r_x \geq r \quad \text{and} \quad \sum_{x \in \mathcal{C}} r_x^k \leq C_V(m)R^k. \qquad (6.24)$$

Moreover, for each $x \in \mathcal{C}$, one of the following is verified

  i) $r_x = r$

  ii) *the set of points* $F_x \equiv \{y \in \mathcal{S} \cap B_{2r_x}(x) \; s.t. \; \theta(y, \rho r_x/10) > E - \delta\}$ *is contained in* $B_{\rho r_x/5}(L_x) \cap$
    $B_{2r_x}(x)$, *where* L_x *is some* $k - 1$ *dimensional affine subspace.*

Remark 52 By the scale-invariance properties of $\hat{\theta}$, it is clear that for simplicity we can assume wlog that $R = 1$.

Remark 53 Note that the set F_x may be empty.

Remark 54 For convenience, and without any loss of generality, we will assume in the proof that r is some (positive) power of ρ, and that ρ is some (negative) power of 2. In particular:

$$r = \rho^{\bar{j}} \quad \text{and} \quad \rho = 2^{-a}, \quad \text{with} \quad a, \bar{j} \in \mathbb{N}. \tag{6.25}$$

6.2.1 Proof of Lemma 51

The idea of the proof is the following. We are going to build inductively on i a covering of the set $\mathcal{S}_{\epsilon,r}^k$ by a family of balls of radius $r_i = \rho^i$. In the inductive step, we will look at each ball of radius r_i and determine if this is a "good" or a "bad" ball according to how many points inside this ball have $\hat{\theta}(y, \rho r_i/10) \geq E - \delta$.

If this set of points "effectively span" some k-dimensional affine subspace V, then we will apply Lemma 31 in order to see that the whole set $\mathcal{S}_{\epsilon,r}^k \cap B_{r_i}(x)$ is contained in small neighborhood of V. Moreover, using Lemma 35, we will see that we can cover the whole neighborhood of V by balls with uniform radius, and this covering will satisfy the assumptions of the discrete Reifenberg theorem. These balls are the good balls.

If this set of points is empty, or it does not "effectively span" something k-dimensional, then we will stop refining our covering, because by definition condition (ii) is verified.

The uniform k-dimensional content estimates will follow from the discrete Reifenberg theorem 40 applied to the natural measure associated with this covering. The β_2 estimates needed to apply the Reifenberg theorem are a consequence of Section 6.1.

6.2.2 Inductive covering: first step

Consider the map $u : B_3(0) \to N$, let $\mathcal{S} \subseteq \mathcal{S}_{\epsilon,\delta r}^k$ be an arbitrary subset and define the set

$$F = \left\{y \in B_2(0) \cap \mathcal{S} \; s.t. \; \hat{\theta}(y, \rho/10) > E - \delta\right\}. \tag{6.26}$$

If there exists a $k - 1$-dimensional subspace L such that $F \subset B_{\rho/5}(L)$, then there's nothing to prove. In this case, we call $B_1(0)$ a *bad* ball.

Otherwise, we say that $B_1(0)$ is a *good* ball. In this second case, let V be a k-dimensional subspace which is $(\rho/10)$-effectively spanned by the set F. Thus by definition there exists $\{y_j\}_{j=0}^k \subset F$ that $(\rho/10)$-effectively span V. For δ sufficiently small, we can apply Lemma 31 to $B_1(0)$, we obtain that

$$\mathcal{S}_{\epsilon,\delta r}^k \cap B_1(0) \subset B_{\rho/5}(V). \tag{6.27}$$

Consider a finite covering of $B_{\rho/5}(V) \cap B_1$ by balls $\left\{B_\rho(x)\right\}_{x \in \mathcal{C}}$ such that

1. $x \in V \cap B_1(0)$
2. if $x \neq y$, then $B_{\rho/5}(x) \cap B_{\rho/5}(y) = \emptyset$

Note that, by Lemma 35, we have for all $x \in \mathcal{C}$:

$$\hat{\theta}(x, \rho/10) \geq E - \eta, \tag{6.28}$$

as long as δ is sufficiently small. Under the same smallness assumption, Lemma 36 implies that for all x we have $x \in S^k_{\rho, \epsilon/2}$. We will need these two properties later on in order to apply the discrete Reifenberg Theorem 40 to the measure associated to our final covering.

This completes the base step of the inductive covering we will be constructing in the next subsection. Now we will consider any of the balls $B_\rho(x)$ in this covering and start over the process.

6.2.3 Inductive step

We will build by induction a sequence of coverings

$$S \subseteq \bigcup_{x \in \mathcal{C}^j} B_{r_x^j}(x) = \bigcup_{x \in \mathcal{C}^j_b} B_{r_x^j}(x) \cup \bigcup_{x \in \mathcal{C}^j_g} B_{r_x^j}(x) \equiv B_{r_x^j}\left(\mathcal{C}^j_b\right) \cup B_{r_x^j}\left(\mathcal{C}^j_g\right), \tag{6.29}$$

where $\mathcal{C}^j_b$ will represent the centers of a collection of "bad balls" and $\mathcal{C}^j_g$ will represent the centers of a collection of "good balls" such that

1. If $x \in \mathcal{C}^j_b$ then $r_x^j \geq \rho^j$ and the set $F_x = \left\{ y \in S \cap B_{2r_x^j}(x) \ \ s.t. \ \hat{\theta}(y, \rho r_x^j/10) \geq E - \delta \right\}$ is contained in some $B_{\rho r_x^j/5}(L_x)$, where L_x is a $k-1$-dimensional affine subspace.
2. If $x \in \mathcal{C}^j_g$ then $r_x^j \equiv \rho^j$ and the set $F_x = \left\{ y \in S \cap B_{2r_x^j}(x) \ \ s.t. \ \hat{\theta}(y, \rho r_x^j/10) \geq E - \delta \right\}$ $(\rho r_x^j/10)$-effectively spans a k-dimensional affine subspace V_x.
3. For all $x \neq y \in \mathcal{C}^j$ we have $B_{r_x/5}(x) \cap B_{r_y/5}(y) = \emptyset$.
4. For all $x \in \mathcal{C}^j$ we have $\hat{\theta}(x, r_x) \geq E - \eta$.
5. For all $x \in \mathcal{C}^j$ and for all $s \in [r_x, 1]$, $B_s(x)$ is not $(k+1, \epsilon/2)$-symmetric.

Suppose that we have this covering for some j, and consider the set

$$R_j = S \setminus \bigcup_{x \in \mathcal{C}^j_b} B_{r_x}(x) = S \setminus B_{r_x}\left(\mathcal{C}^j_b\right). \tag{6.30}$$

Note that by definition this set is contained in $B_{\rho^j}\left(\mathcal{C}^j_g\right)$. For each $x \in \mathcal{C}^j_g$, we know that F_x $[\rho^{j+1}/10]$-effectively spans a k-dimensional subspace V_x. As seen in the first inductive step, by Proposition 31 we have that

$$S^k_{\epsilon, \delta r} \cap B_{2\rho^j}(x) \subset B_{\rho^{j+1}/5}(V_x) \tag{6.31}$$

for all $x \in \mathcal{C}^j_g$ as long as

$$\delta \leq \delta_3(m, \Lambda, N, \gamma, \rho, \epsilon). \tag{6.32}$$

In order to build an open covering of R_j, consider the set

$$A = \bigcup_{x \in \mathcal{C}_g^j} \left(B_{\rho^j}(x) \cap V_x \right) \setminus B_{r_x/2}\left(\mathcal{C}_b^j \right) . \tag{6.33}$$

By (6.30) and (6.31), and since $\rho < 100^{-1}$, we have

$$R_j \subseteq B_{\rho^{j+1}/5}(A) . \tag{6.34}$$

Now first note that by the definition of A and since $\rho \leq 100^{-1}$, all of these balls are disjoint from $B_{r_x/10}\left(\mathcal{C}_b^j \right)$. Moreover, by Lemma 35, if we choose δ sufficiently small, for all $y \in A$ we have

$$\hat{\theta}\left(y, \rho^{j+1}/10\right) \geq E - \eta . \tag{6.35}$$

In particular, we need

$$0 < \delta \leq \delta_6(m, \Lambda, N, \rho, \gamma, \eta) . \tag{6.36}$$

Furthermore, if we choose δ small enough, by Lemma 36, we obtain that for all $s \in [\rho, 1]$ and for all $y \in A$

$$B_s(y) \text{ is not } (k+1, \epsilon/2)\text{-symmetric} . \tag{6.37}$$

In particular, we need

$$0 < \eta \leq \delta_7(m, \Lambda, N, \rho, \gamma, \epsilon) \quad \Longleftarrow \quad 0 < \delta \leq \min\{\delta_7(m, \Lambda, N, \rho, \gamma, \epsilon), \delta_6(m, \Lambda, N, \rho, \gamma, \delta_7)\} . \tag{6.38}$$

Now consider a (finite) Vitali subcovering of this set given by

$$R_j \subseteq \bigcup_{x \in \mathcal{C}^A} B_{\rho^{j+1}}(x) . \tag{6.39}$$

We can classify all the balls in this covering into good and bad according to how spread their set F is. In particular, for all $x \in \mathcal{C}^A$ consider as above the set

$$F_x = \left\{ y \in \mathcal{S} \cap B_{2\rho^{j+1}}(x) \;\; s.t. \;\; \hat{\theta}\left(y, \rho^{j+2}/10\right) \geq E - \delta \right\} . \tag{6.40}$$

If F_x $[\rho^{j+2}/10]$-effectively spans a k dimensional subspace V_x, then we say that $B_{\rho^{j+1}}(x_t)$ is a good ball, and we put $x \in \mathcal{C}_g^A$. Otherwise, we say that $B_{\rho^{j+1}}(x)$ is a bad ball, and we put $x \in \mathcal{C}_b^A$.

We define

$$\mathcal{C}_b^{j+1} = \mathcal{C}_b^j \cup \mathcal{C}_b^A , \quad \mathcal{C}_g^{j+1} = \mathcal{C}_g^A . \tag{6.41}$$

Note that the set of bad balls contains *all* the bad balls encountered at any previous step. On the contrary, good balls get refined at each stage, and at each induction step the previous bad balls disappear from the set $\mathcal{C}_g$.

Now the induction is complete. Indeed, property 1 and 2 are a direct consequence of the definition of $\mathcal{C}_g$ and $\mathcal{C}_b$. Property 3 comes from the definition of A and the Vitali covering lemma. Finally, property 4 is a consequence of (6.35) and property 5 comes from (6.37).

6.2.4 Volume estimates

Now we are in a position to prove the desired volume estimates, and in particular

$$\sum_{x \in \mathcal{C}} r_x^k \le C_V(m), \tag{6.42}$$

where $\mathcal{C} = \mathcal{C}^{\bar{j}}$ for $\bar{j}$ such that $\rho^{\bar{j}} = r$.

We will prove this estimate by an induction on the radius. For convenience, we define the measure

$$\mu = \omega_k \sum_{x \in \mathcal{C}} r_x^k \delta_x. \tag{6.43}$$

Upwards induction For all $t \in (0, 1]$, set $\mathcal{C}_t = \{x \in \mathcal{C} \ \ s.t. \ \ r_x \le t\}$, and define the measure

$$\mu_t \equiv \omega_k \sum_{x \in \mathcal{C}_t} r_x^k \delta_x \le \mu. \tag{6.44}$$

Now we want to prove inductively on $t = r, 2r, 2^2r, 2^3r, \cdots, 1/8$ that for some universal constant $C_R(m)$, for all $x \in B_3(0)$ and $s \ge r$ we have

$$\mu_t(B_t(x)) \equiv \left(\sum_{x \in \mathcal{C} \ \ s.t. \ r_x \le t} \omega_k r_x^k \delta_x \right)(B_t(x)) \le C_R(m)t^k. \tag{6.45}$$

Note that $C_R(m)$ is the constant in Theorem 40. Note also that $\mu_1 = \mu$, so at the last step of the induction we will have recovered an estimate for the whole μ, up to a covering of $B_1(0)$ by balls $B_{1/8}(p_i)$. In other words, we prove (6.42) with

$$C_V(m) = c(m)C_R(m). \tag{6.46}$$

Note that the base step is easily seen to be true for $t = r$. Indeed, at this stage we have

$$\mu_r = \sum_{x \in \mathcal{C}_r} \omega_k r^k \delta_x, \tag{6.47}$$

where all $B_{r/5}(x_i)$ are disjoint. Thus we immediately have $\mu_r(B_r(x)) \le c(m)r^k$.

Now, suppose that we have proven (6.45) for $t \le 2^j r$, we will show that (6.45) holds also for $t = 2^{j+1}r$.

Rough estimate First of all, we note that by a very bad estimate we have for all $x \in B_1(0)$:

$$\mu_{2\bar{r}}(B_{2\bar{r}}(x)) \le c(m)C_R(m)(2\bar{r})^k, \tag{6.48}$$

where for convenience we have set $\bar{r} = 2^j r$. Indeed, we can split $\mu_{2\bar{r}}$ into

$$\mu_{2\bar{r}} = \mu_{\bar{r}} + \tilde{\mu}_{2\bar{r}} \equiv \sum_{x \in \mathcal{C}_{\bar{r}}} \omega_k r_x^k \delta_x + \sum_{x \in \mathcal{C} \ \ s.t. \ r_x \in (\bar{r}, 2\bar{r}]} \omega_k r_x^k \delta_x. \tag{6.49}$$

Take a covering of $B_{2\bar{r}}(x)$ by balls $B_{\bar{r}}(y_i)$ such that $B_{\bar{r}/2}(y_i)$ are disjoint. The number of these balls has a universal bound $c(m)$, and by induction we have

$$\mu_{\bar{r}}(B_{2\bar{r}}(x)) \le \sum_i \mu_{\bar{r}}(B_{\bar{r}}(y_i)) \le c(m)C_R(m)\bar{r}^k . \qquad (6.50)$$

As for the other part of μ, by definition of this measure all the balls $B_{r_x/5}(x)$ are pairwise disjoint, and so we get immediately

$$\tilde{\mu}_{2\bar{r}}(B_{2\bar{r}}(x)) \le c(m)(2\bar{r})^k . \qquad (6.51)$$

Reifenberg estimates We will show inductively that we can apply Theorem 40 to the measures $\mu_{2\bar{r}}$ on each fixed $B_{2\bar{r}}(x)$. For convenience, we set

$$\bar{\mu} = \mu_{2\bar{r}}|_{B_{2\bar{r}}(x)} . \qquad (6.52)$$

Note that for all $x \in \text{supp}(\mu)$, and all $s \in [r_x, 1]$, we have $\hat{\theta}(x, s) - \hat{\theta}(x, s/2) < \eta$ because $\hat{\theta}(x_i, s) \le E$ by monotonicity of $\hat{\theta}$ and by definition of E, and $\hat{\theta}(x_i, s/2) \ge E - \eta$ by condition (4) of our constructed covering. Now we can choose η small enough so that for all $x \in \text{supp}(\mu)$ and $0 < s \le 1$ we have the β_2 estimate

$$\beta_{2,\bar{\mu}}(x, s)^2 \overset{thm.45}{\le} C_1 s^{-k} \int_{B_s(x)} \hat{W}_s(y) \, d\bar{\mu}(y) , \qquad (6.53)$$

where we have set for all $x \in \text{supp}(\mu)$:

$$\hat{W}_s(x) = \begin{cases} W_s(x_i) & \text{if } s > r_x , \\ 0 & \text{if } s \le r_x . \end{cases} \qquad (6.54)$$

Indeed, for $s \le r_x$, $\text{supp}(\mu) \cap B_s(x) = \{x\}$, and there's nothing to prove. If $s \ge r_x$, then for all $y \in B_{r_x}(x)$, $r_y < s$ by construction of μ.

Now for $r_x \le s \le 1/8$, the ball $B_{8s}(x)$ is not $(k + 1, \epsilon/2)$-symmetric by (6.37). Let $\delta_5(m, N, \Lambda, \gamma, \epsilon)$ be the parameter found in Lemma 34 and Theorem 45. By Proposition 24, we can choose a threshold

$$\eta_0(m, N, \Lambda, \gamma, \epsilon) = \delta_1(m, N, \Lambda, \gamma, \epsilon, \delta_5) > 0 \qquad (6.55)$$

such that $\hat{\theta}(x, 8s) - \hat{\theta}(x, 4s) < \eta$ with $\eta \le \eta_0$ implies that $B_{8s}(x_i)$ is $(0, \delta_5)$-symmetric. Thus all the assumptions of Theorem 45 are satisfied, and we have the estimate (6.53) as desired.

Now we can prove that for all $y \in B_{2\bar{r}}(x)$, and $r \le 2\bar{r}$, we have

$$\int_{B_r(y)} \left(\int_0^r \beta_{2,\bar{\mu}}^k(z, s)^2 \frac{ds}{s} \right) d\bar{\mu}(z) < c(m)C_1 C_R^2 \eta r^k . \qquad (6.56)$$

Indeed, by (6.53) we can estimate for all $s \le r$:

$$\int_{B_r(y)} \beta_{2,\bar{\mu}}^k(z, s)^2 \, d\bar{\mu}(z) \le C_1 s^{-k} \int_{B_r(y)} \left[\int_{B_s(z)} \hat{W}_s(t) \, d\bar{\mu}(t) \right] d\bar{\mu}(z) . \qquad (6.57)$$

Now, on $B_s(z)$, either $\bar\mu = \mu_s|_{B_{2\bar r}(x)}$, or there exists an $x \in \text{supp}(\mu) \cap B_s(z)$ with $r_x > s$. Since $z \in \text{supp}(\mu)$ as well, by construction we have $z = x = \text{supp}(\mu) \cap B_s(z)$, and $\hat W_s(z) = 0$. Thus in either case we have

$$\int_{B_r(y)} \beta_{2,\bar\mu}^k(z,s)^2 \, d\bar\mu(z) \le C_1 s^{-k} \int_{B_r(y)\cap B_{2\bar r}(x)} \left[\int_{B_s(z)\cap B_{2\bar r}(x)} \hat W_s(t) \, d\mu_s(t) \right] d\mu_s(z). \qquad (6.58)$$

By induction, and by the rough estimates in (6.48), for all $s \in (0, 2\bar r]$ and $z \in B_1(0)$ we can estimate

$$\mu_s(B_s(z)) \le c(m)C_R s^k. \qquad (6.59)$$

Thus we obtain

$$\int_{B_r(y)} \beta_{2,\bar\mu}^k(z,s)^2 \, d\bar\mu(z) \le c(m)C_1 C_R \int_{B_{r+s}(y)\cap B_{2\bar r}(x)} \hat W_s(z) d\mu_s(z) = c(m)C_1 C_R \int_{B_{r+s}(y)} \hat W_s(z) d\bar\mu(z). \qquad (6.60)$$

This yields

$$\int_{B_r(y)} \left(\int_0^r \beta_{2,\bar\mu}^k(z,s)^2 \, \frac{ds}{s} \right) d\bar\mu(z) \le c(m)C_1 C_R \int_{B_{2r}(y)} \left[\int_0^r \hat W_s(z) \frac{ds}{s} \right] d\bar\mu(z). \qquad (6.61)$$

Note that for all $x \in \text{supp}(\mu)$ and $r \le 2\bar r \le 1/8$, we have

$$\int_0^r \hat W_s(x) \frac{ds}{s} = \int_{r_x}^r \hat W_s(x) \frac{ds}{s} \le \int_{r_x}^{1/8} \hat W_s(x) \frac{ds}{s} \overset{(2.14)}{\le} c\left[\hat\theta(x,1) - \hat\theta(x,r_x) \right] \le c\eta. \qquad (6.62)$$

Thus, using again the induction hypothesis and the rough estimates (6.48), we prove (6.56).

If we choose η small enough, in particular

$$\eta \le \eta_1(m, N, \Lambda, \gamma, \epsilon) = c(m) \frac{\delta_R^2}{C_1 C_R^2}, \qquad (6.63)$$

we can apply Theorem 40 to $\bar\mu$ and obtain (6.45) as wanted.

The only thing left to do is to choose $\delta = \delta(m, N, \Lambda, \gamma, \rho, \epsilon) > 0$ in such a way that (6.35) is satisfied with

$$\eta \le \min\{\eta_0, \eta_1, \delta_7\} \qquad (6.64)$$

and also (6.32) is satisfied. Given (6.55) and (6.63), as noted above this is a simple application of Lemmas 35 and 36. This finishes the proof of Lemma 51.

$\square$

6.2.5 Second covering lemma

By repeating this covering argument over bad balls, we obtain the following

Lemma 55 (Covering Lemma II) *Let* $u : B_3(0) \to N$ *be an approximately harmonic map satisfying* (W) *and* (S)*, with the conditions* (f)*. Fix any* $\epsilon > 0$ *and* $0 < r \le R$*,* $0 < R \le 1$*, set* $E = \sup_{x \in B_{2R}(0) \cap S} \hat{\theta}_1(x)$*, and assume the uniform bound* $E \le \Lambda$*. There exists* $\delta = \delta(m, N, \Lambda, \gamma, \epsilon) > 0$ *and* $C_F(m)$ *such that the following is true.*

If $\mathcal{F} < \delta$*, for any subset* $S \subseteq S^k_{\epsilon, \delta r}$*, there exists a finite covering of* $S \cap B_R(0)$ *such that*

$$S \cap B_R(0) \subseteq \bigcup_{x \in \mathcal{C}} B_{r_x}(x), \quad \text{with} \quad r_x \ge r \quad \text{and} \quad \sum_{x \in \mathcal{C}} r_x^k \le C_F(m) R^k. \quad (6.65)$$

Moreover, for each $x \in \mathcal{C}$*,*

  i) *either* $r_x = r$
  ii) *or we have the following uniform energy drop*

$$\forall y \in B_{r_x}(x) \cap S, \ \hat{\theta}(y, r_x/10) \le E - \delta. \quad (6.66)$$

Remark 56 As for the previous covering lemma, also in this case we can assume for simplicity and wlog that $R = 1$.

Proof We need to refine the covering of the previous lemma. Recall that by lemma 51 we have a covering of $S \cap B_1(0)$ given by

$$S \cap B_1(0) \subseteq \bigcup_{x \in \mathcal{C}} B_r(x) \equiv \bigcup_{x \in \mathcal{C}_r} B_r(x) \cup \bigcup_{x \in \mathcal{C}_+} B_{r_x}(x) \quad \text{with} \quad r_x \ge r \quad \text{and} \quad \sum_{x \in \mathcal{C}_r \cup \mathcal{C}_+} r_x^k \le C_V(m),$$
$$(6.67)$$

where we have set

$$\mathcal{C}_r = \{x \in \mathcal{C} \text{ s.t. } r_x = r\} \quad \text{and} \quad \mathcal{C}_+ = \{x \in \mathcal{C} \text{ s.t. } r_x > r\}, \quad \mathcal{C} = \mathcal{C}_r \cup \mathcal{C}_+. \quad (6.68)$$

We will of course keep $\mathcal{C}_r$ as part of our final covering, while we will refine the covering on each of the balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}_+}$ in an inductive way. By item (ii) of lemma 51, for each $x \in \mathcal{C}_+$ the set $F_x \equiv \{y \in S \cap B_{2r_x}(x) \text{ s.t. } \theta(y, \rho r_i/10) > E - \delta\}$ is close to a $k-1$-dimensional space. Assuming that $F_x = \emptyset$, all we need to do in order to achieve (6.66) is to re-cover $B_{r_x}(x)$ with balls $\{B_{\rho r_x}(y)\}_{y \in \mathcal{C}_x^{(1,f)}}$. These balls are the final covering we are looking for. Evidently, the number of these balls is bounded by a constant $C_f(m, \rho)$.

If $F_x \ne \emptyset$, we need to exploit the fact that we still know $F_x \subseteq B_{\rho r_x/5}(L_x) \cap B_{2r_x}(x)$, where L_x is at most $k - 1$ dimensional. Thus we can cover $B_{r_x}(x) \setminus B_{\rho r_x}(F_x)$ as above, and cover $B_{\rho r_x}(F_x)$ separately by balls $\{B_{\rho r_x}(y)\}_{y \in \mathcal{C}_x^{(1,b)}}$. On these "bad balls", we will not be able to obtain any information on the energy drop over these new balls in the covering. However, their k-dimensional content is small since F_x behaves like a $k - 1$ dimensional set. This will allow us to start over on each of these bad balls separately, and keep a uniform k-dimensional estimate on the content of the final covering. More precisely:

6.2.6 Re-covering of bad balls: Induction

In detail, we will build by induction on i a sequence of coverings of $\mathcal{S} \subseteq \mathcal{S}^k_{\epsilon,\delta r} \cap B_1(0)$ such that

1. For all $i = 1, 2, \cdots$

$$\mathcal{S} \subseteq \bigcup_{x \in \mathcal{C}^{(i,r)}} B_r(x) \cup \bigcup_{x \in \mathcal{C}^{(i,f)}} B_{r_x}(x) \cup \bigcup_{x \in \mathcal{C}^{(i,b)}} B_{r_x}(x) \, . \tag{6.69}$$

2. For all $x \in \mathcal{C}^{(i,r)}$, $r_x = r$. In other words, on these "r-balls" option (i) of our lemma is verified,
3. For all $x \in \mathcal{C}^{(i,f)}$ and all $z \in B_{2r_x}(x)$ we have $\hat{\theta}(z, r_x/10) \le E - \delta$. In other words, on these "final balls" option (ii) of our lemma is verified,
4. for all $x \in \mathcal{C}^{(i,b)}$, $r < r_x \le \rho^i$. On these "bad balls", none of the two stopping options is verified, thus we need to refine our covering here.
5. For some constant $C_F(m)$, we have the estimates

$$\sum_{x \in \mathcal{C}^{(i,r)} \cup \mathcal{C}^{(i,f)}} r_x^k \le C_F(m) \left(\sum_{j=1}^{i} 2^{-j} \right), \quad \sum_{x \in \mathcal{C}^{(i,b)}} r_x^k \le 2^{-i} \, . \tag{6.70}$$

Thus the estimates on r and final balls has uniform bounds, while our estimates on bad balls has exponentially decreasing bounds.

6.2.7 Re-covering of bad balls: First step in the induction

For $i = 1$, consider the covering (6.67) given by the previous lemma. We keep the balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}_r}$ as they are, while for each $x \in \mathcal{C}_+$ consider two coverings of $B_{\rho r_x}(F_x)$ and its complement

$$B_{r_x}(x) \setminus B_{\rho r_x}(F_x) \subseteq \bigcup_{y \in \mathcal{C}^{(1,f)}_x} B_{\rho r_x}(y) \, , \quad B_{r_x}(x) \cap B_{\rho r_x}(F_x) \subseteq \bigcup_{y \in \mathcal{C}^{(1,b)}_x} B_{\rho r_x}(y) \, , \tag{6.71}$$

where $B_{\rho r_x/2}(y)$ are pairwise disjoint in both coverings.

By definition of F_x, for all $y \in \mathcal{C}^{(1,f)}_x$ the energy drop condition (6.66) is satisfied. Moreover we have the trivial estimates

$$\sum_{y \in \mathcal{C}^{(1,f)}_x} (\rho r_x)^k = (\rho r_x)^k \# \left\{ y \in \mathcal{C}^{(1,f)}_x \right\} \le c(m) \rho^{k-m} r_x^k \equiv C_f(m, \rho) r_x^k \, . \tag{6.72}$$

Since the energy drop is verified on these balls, we define $\mathcal{C}^{(1,f)}$ to be the set of final balls at the step $i = 1$ by

$$\mathcal{C}^{(1,f)} = \bigcup_{x \in \mathcal{C}_+} \mathcal{C}^{(1,f)}_x \, . \tag{6.73}$$

For $y \in \mathcal{C}^{(1,b)}_x$, the energy drop condition is not verified. However, since there exists a $k - 1$ dimensional space L_x such that

$$F_x = \left\{ y \in \mathcal{S} \cap B_{2r_x}(x) \ \ s.t. \ \ \hat{\theta}(y, \rho r_x/10) \ge E - \delta \right\} \subseteq B_{\rho r_x/5}(L_x) \, , \tag{6.74}$$

then we can estimate

$$\sum_{y \in \mathcal{C}_x^{(1,b)}} (\rho r_x)^k = \rho^k r_x^k \# \left\{ \mathcal{C}_x^{(1,b)} \right\} \le c(m) \rho^{1-k} \rho^k r_x^k \equiv C_c(m) \rho r_x^k. \tag{6.75}$$

On these balls, we can either have the stopping condition $\rho r_x = r$, or we need to refine the covering further. Thus we define

$$\mathcal{C}^{(1,b)} = \bigcup_{x \in \mathcal{C}_+, \ \rho r_x > r} \mathcal{C}_x^{(1,b)}, \quad \mathcal{C}^{(1,r)} = \mathcal{C}_r \cup \bigcup_{x \in \mathcal{C}_+, \ \rho r_x = r} \mathcal{C}_x^{(1,b)}. \tag{6.76}$$

$\mathcal{C}^{(1,b)}$ represents the set of "bad balls" where we need to refine our covering further.

By this and lemma 51, in particular by the estimates in (6.67), we obtain that

$$\sum_{y \in \mathcal{C}^{(1,b)}} r_y^k \le C_c(m) \rho \sum_{x \in \mathcal{C}_+} r_x^k \le C_V(m) C_c(m) \rho. \tag{6.77}$$

If we choose

$$0 < \rho(m) \le \min \left\{ 100^{-1}, \frac{1}{2} C_V(m)^{-1} \cdot C_c(k)^{-1} \right\}, \tag{6.78}$$

we can rephrase the above estimates as

$$\sum_{y \in \mathcal{C}^{(1,b)}} r_y^k \le \frac{1}{2}. \tag{6.79}$$

If we set

$$C_F(m) = 2 C_V(m) \left(C_f(m, \rho(m)) + C_c(m) \right), \tag{6.80}$$

the estimates on the final and r-balls are

$$\sum_{y \in \mathcal{C}^{(1,r)} \cup \mathcal{C}^{(1,f)}} r_y^k = \# \{\mathcal{C}_r\} r^k + \sum_{x \in \mathcal{C}_+} r_x^k \left(C_f(m, \rho) + C_c(m, \rho) \right) \le C_V(m) \left(C_f(m, \rho) + C_c(m) \right) = \frac{1}{2} C_F(m). \tag{6.81}$$

Note that clearly for all $y \in \mathcal{C}^{(1,b)}$, we have $r_y \le \rho$.

6.2.8 Re-covering of bad balls: Induction step

Suppose that we have obtained our covering for i. It is clear that we need to improve our covering only on the balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}^{(i,b)}}$. In order to do so, we consider each of these balls separately.

Since all the assumptions on lemma 51 are satisfied on each of the $B_{r_x}(x)$, we can apply again this lemma to each $B_{r_x}(x)$, and obtain that for all x there exists a covering

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \hat{\mathcal{C}}_{r,x}} B_r(y) \cup \bigcup_{y \in \hat{\mathcal{C}}_{+,x}} B_{r_y}(y) \quad \text{with} \quad r_y \ge r \quad \text{and} \quad \sum_{y \in \hat{\mathcal{C}}_{r,x} \cup \hat{\mathcal{C}}_{+,x}} r_y^k \le C_V(m) r_x^k. \tag{6.82}$$

Moreover, for each $y \in \hat{\mathcal{C}}_{+,x}$, there exists a $k - 1$ dimensional subspace L_y such that

$$F_y \equiv \left\{ z \in \mathcal{S} \cap B_{2r_y}(y) \;\; s.t. \;\; \theta(z, \rho r_y/10) > E - \delta \right\} \subseteq B_{\rho r_y/5}\left(L_y\right) \cap B_{2r_y}(y) \,. \tag{6.83}$$

By applying exactly the same procedure described in the first step of the induction to each of the balls $\left\{ B_{r_y}(y) \right\}_{y \in \hat{\mathcal{C}}_{+,x}}$, we obtain the new desired covering. In particular, for each $y \in \hat{\mathcal{C}}_{+,x}$ we can find a covering

$$B_{r_y}(y) \setminus B_{\rho r_y}\left(F_y\right) \subseteq \bigcup_{z \in \hat{\mathcal{C}}_y^{(i+1,f)}} B_{\rho r_y}(z) \,, \quad B_{r_y}(y) \cap B_{\rho r_y}\left(F_y\right) \subseteq \bigcup_{z \in \mathcal{C}_y^{(i+1,b)}} B_{\rho r_y}(z) \,, \tag{6.84}$$

where for all $z \in \hat{\mathcal{C}}_y^{(i+1,f)}$ and all $p \in \mathcal{S} \cap B_{2\rho r_y}(z)$, we have $\hat{\theta}(p, \rho r_y) \le E - \delta$, and we have the estimates

$$\sum_{z \in \hat{\mathcal{C}}_y^{(i+1,f)}} (\rho r_y)^k \le C_f(m, \rho) r_y^k \,, \quad \sum_{z \in \mathcal{C}_y^{(i+1,b)}} (\rho r_y)^k \le C_c(m) \rho r_y^k \,. \tag{6.85}$$

The new set $\mathcal{C}^{(i+1,f)}$ is now defined as the previous set of "final balls" $\mathcal{C}^{(i,f)}$ along with the new final balls $\hat{\mathcal{C}}^{(i+1,f)}$ obtained with this covering, thus making

$$\hat{\mathcal{C}}^{(i+1,f)} = \bigcup_{x \in \mathcal{C}^{(i,b)}} \bigcup_{y \in \hat{\mathcal{C}}_{+,x}} \hat{\mathcal{C}}_y^{(i+1,f)} \,, \quad \mathcal{C}^{(i+1,f)} = \mathcal{C}^{(i,f)} \cup \hat{\mathcal{C}}^{(i+1,f)} \,. \tag{6.86}$$

In a similar way for the r-balls, we obtain

$$\hat{\mathcal{C}}^{(i+1,r)} = \bigcup_{x \in \mathcal{C}^{(i,b)}} \left(\hat{\mathcal{C}}_{r,x} \cup \bigcup_{y \in \hat{\mathcal{C}}_{+,x}, \; \rho r_y = r} \mathcal{C}_y^{(i+1,b)} \right) \,, \quad \mathcal{C}^{(i+1,r)} = \mathcal{C}^{(i,r)} \cup \hat{\mathcal{C}}^{(i+1,r)} \,. \tag{6.87}$$

However, evidently the new set of "bad balls" does not contain the bad balls at the previous scale, since those are the ones that were just re-covered. In particular

$$\mathcal{C}^{(i+1,b)} = \bigcup_{x \in \mathcal{C}^{(i,b)}} \bigcup_{y \in \mathcal{C}_{+,x}, \; \rho r_y > r} \mathcal{C}_y^{(i+1,b)} \,. \tag{6.88}$$

The k-dimensional content estimate of our covering are obtained by iterating the estimates obtained in the first step. In detail, by arguing as in (6.77) and (6.79), and by choosing ρ according to (6.78), we obtain

$$\sum_{z \in \mathcal{C}^{(i+1,b)}} r_z^k \le \sum_{x \in \mathcal{C}^{(i,b)}} \frac{1}{2} r_x^k = 2^{-1-i} \,. \tag{6.89}$$

As for final and r-balls, arguing as in (6.81) we can estimate the contribution given by the new r and final balls by

$$\sum_{z \in \hat{\mathcal{C}}^{(i+1,r)} \cup \hat{\mathcal{C}}^{(i+1,f)}} r_z^k \le \left(\frac{1}{2} C_F(m) \right) \sum_{z \in \mathcal{C}^{(i,b)}} r_z^k = 2^{-i-1} C_F(m) \,. \tag{6.90}$$

This yields the desired result (6.70), and in turn concludes the proof of the lemma.

It is worth noticing that at the i-th step of the induction, the radius of the biggest ball in the covering is smaller than ρ^i. Thus eventually $\rho^i \le r$ and this induction will stop in a finite number of steps.

6.2.9 Keeping track of the constants

For the reader's convenience we record here how all the constants involved in the previous two lemmas were chosen.

First of all, note that $\epsilon > 0$ is arbitrary, as well as $r > 0$. However, it is of course important that all the constants here are *independent of r*.

$C_R(m)$ is the constant coming from the Reifenberg theorem 40, and it depends only on m. $C_V(m)$ is fixed in (6.46), and it is just a dimensional constant $c(m)$ (coming from a rough cover of $B_1(0)$ by balls of radius $1/8$) times $C_R(m)$. Thus $C_V(m)$ clearly depends only on m. $C_c(m)$ is fixed in (6.75), and is just another covering constant whose value depends only on m.

The parameter ρ, which was a free parameter in the first covering, is fixed once and for all in (6.78) as a constant depending only on m. For convenience, we can also pick a ρ satisfying (6.25). Once this choice has been fixed, also the constant $C_F(m)$ introduced in (6.80) depends only on m.

The parameter $\eta > 0$ is chosen according to (6.38), (6.55) and (6.63), as explained in (6.64). At last, with this positive value of η fixed, we choose δ in such a way that (6.32), (6.35) and (6.38) are all satisfied.

6.3 Proof of the main theorems

Before proving our main theorems, we provide an argument that justifies the assumption $\mathcal{F} < \delta$ which is present in all of our technical lemmas and covering arguments. The idea is that by condition (f), we can focus on small enough scales r on which the value of $\mathcal{F}$ has "decayed" by a factor r^γ.

6.3.1 First covering by balls of small radius

In all the estimates we need, an important assumption is that the constant $\mathcal{F}$ in (f) is sufficiently small, in other words our estimates apply if u is an approximate harmonic map and the error f is small enough. This assumption is not too restrictive because (f) is better than scale invariant in nature. Indeed, if we restrict ourselves to small enough scales $r \le r_0$, the rescaled maps $T^u_{x,r} : B_3(0) \to N$ are approximate harmonic maps solving (2.31) and (2.32) with (2.33), and the error function $\tilde{f}$ satisfies

$$s^{4-m} \int_{B_y(s)} \left| \tilde{f} \right|^2 \le (\mathcal{F}r_0^\gamma)s^\gamma \tag{6.91}$$

for all $y \in B_1(0)$ and $s \le 1$. Thus, if we choose $r_0(m, \mathcal{F}, \gamma, \delta)$ in such a way that $\mathcal{F}r_0^\gamma \le \delta$, we can guarantee the smallness hypothesis $\mathcal{F} < \delta$ on all smaller scales.

We can cover the original ball $B_1(0)$ with balls $B_{r_0/2}(x_i)$ such that $B_{r_0/4}(x_i)$ are disjoint, and then start over on each of this smaller balls. Evidently, we have

$$\sum_i r_0^k \le C(m)r_0^{k-m} \le C_0(m, \mathcal{F}, \gamma, \delta). \tag{6.92}$$

Since we will pick the parameter δ from the covering Lemma 55, where we have $\delta = \delta(m, N, \Lambda, \gamma, \epsilon)$, and since we are only doing this rough covering once on the first big ball, the final estimates of Theorem 20 are not modified by this. Also the statements about the rectifiability (which is stable under countable unions, let alone finite unions) is not effected by this covering.

Now we are in a position to prove our main theorems. We start with the volume estimates of Theorem 20.

6.3.2 Proof of Theorem 20

This proof is basically a corollary of the covering Lemma 55.

By the argument in Section 6.3.1, we can assume that $\mathcal{F} < \delta$ throughout this proof. Moreover, in a similar spirit, instead of the estimates of (3.1), we will prove the slightly less powerful estimate

$$\mathrm{Vol}\left(B_r\left(\mathcal{S}^k_{\epsilon, \delta r}(u)\right) \cap B_1(0)\right) \leq C'_\epsilon r^{n-k} , \tag{6.93}$$

the difference being the δr in $\mathcal{S}^k_{\epsilon, \delta r}$. As above, since $\delta = \delta(m, N, \Lambda, \gamma, \epsilon)$, this does not affect the final estimate in (3.1), if not by enlarging the constant C'_ϵ to C_ϵ.

Consider the set $\mathcal{S} = \mathcal{S}^k_{\epsilon, \delta r}(u) \cap B_1(0)$. Note that by the monotonicity of $\hat{\theta}$ and the estimates in Lemma 12, we have the uniform bounds

$$\forall x \in B_1(0) , \quad \forall r \in [0, 1], \quad \hat{\theta}(x, r) \leq \Lambda' = c(m)\Lambda + c(m, \gamma)\mathcal{F} . \tag{6.94}$$

Let $E = \sup_{x \in \mathcal{S}} \hat{\theta}(x, 1) \leq \Lambda'$.

6.3.3 Induction on energy upper bounds

Using the covering Lemma 55, we will prove by induction on $i = 0, 1, \cdots, \lfloor \delta^{-1}E \rfloor + 1$ that there exist coverings of $\mathcal{S}$ by balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}^i}$ such that

$$\mathcal{S} \subseteq \bigcup_{c \in \mathcal{C}^i} B_{r_x}(x) , \quad \sum_{x \in \mathcal{C}^i} r_x^k \leq (c(m)C_F(m))^i . \tag{6.95}$$

Moreover, for all i we have

$$r_x \leq r \quad \text{or} \quad \forall y \in \mathcal{S} \cap B_{2r_x}(x) , \quad \hat{\theta}(y, r_x) \leq E - i\delta . \tag{6.96}$$

It is clear that if we pick $i = \lfloor \delta^{-1}E \rfloor + 1$, then the second condition cannot be true anywhere, which means that the first condition must be true, which will complete the construction of the covering.

Now it is clear that the estimate (6.93) is true with

$$C'_\epsilon(m, N, \Lambda, \mathcal{F}, \gamma, \epsilon) = 2^k (c(m)C_F(m))^{\lfloor \delta^{-1}E \rfloor + 1} . \tag{6.97}$$

So the only thing left to do is to prove the properties of this inductive covering in (6.95). Note that this covering is trivial for $i = 0$, since $\mathcal{S} \subseteq B_1(0)$ does the trick at this stage.

By induction, suppose that (6.95) and (6.96) are true for i. Pick any $x \in \mathcal{C}^i$, and consider $B_{r_x}(x)$. By the covering lemma 55 (or better, by an r_x-rescaled version of this lemma), there exists a covering $\hat{\mathcal{C}}_x$ of $\mathcal{S} \cap B_{r_x}(x)$ such that

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \hat{\mathcal{C}}_x} B_{r_y}(y), \quad r_y \le \rho r_x \le \rho^i, \quad \sum_{y \in \hat{\mathcal{C}}_x} r_y^k \le C_F(m) r_x^k. \tag{6.98}$$

Moreover, for all $y \in \hat{\mathcal{C}}_x$, we have

$$\text{either } r_y = r \quad \text{or} \quad \forall z \in \mathcal{S} \cap B_{2r_y}(y), \ \hat{\theta}(z, \rho r_y/10) \le E - i\delta - \delta = E - (i+1)\delta. \tag{6.99}$$

By covering each $B_{r_y}(y)$ again by a minimal set of balls of radius $\rho(m) r_y \le r_y/10$, we obtain a covering $\mathcal{C}_x$ such that

$$\mathcal{S} \cap B_{r_x}(x) \subseteq \bigcup_{y \in \mathcal{C}_x} B_{r_y}(y), \quad r_y \le \rho r_x \le \rho^i, \quad \sum_{y \in \mathcal{C}_x} r_y^k \le c(m) C_F(m) r_x^k. \tag{6.100}$$

Moreover, for all $y \in \mathcal{C}_x$, we have

$$\text{either } r_y \le r \quad \text{or} \quad \forall z \in \mathcal{S} \cap B_{2r_y}(y), \ \hat{\theta}(z, \rho r_y) \le E - i\delta - \delta = E - (i+1)\delta. \tag{6.101}$$

By summing all the contributions coming from balls $\{B_{r_x}(x)\}_{x \in \mathcal{C}^i}$, we obtain

$$\mathcal{C}^{i+1} = \bigcup_{x \in \mathcal{C}^i} \mathcal{C}_x, \quad \sum_{y \in \mathcal{C}^{i+1}} r_y^k = \sum_{x \in \mathcal{C}^i} \left(\sum_{y \in \mathcal{C}_x} r_y^k \right) \le (c(m) C_F(m))^{i+1}, \tag{6.102}$$

as desired.

6.3.4 Proof of Theorem 21 and rectifiability of $\mathcal{S}_\epsilon^k$

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

By the volume estimates in (3.1), we have $\lambda^k\left(\mathcal{S}_\epsilon^k \cap B_1(0)\right) \le C_\epsilon$. By applying the same estimates on any ball $B_r(x)$ with $x \in B_1(0)$ and $r \le 1$, we obtain that

$$\lambda^k\left(\mathcal{S}_\epsilon^k \cap B_r(x)\right) \le C_\epsilon r^k, \tag{6.103}$$

in other words, $\mathcal{S}_\epsilon^k$ is upper-Ahlfors regular.

We will prove that for all measurable subsets $\mathcal{S} \subseteq \mathcal{S}_\epsilon^k \cap B_1(0)$, there exists a k-measurable subset $E \subset \mathcal{S}$ with $\lambda^k(E) \le 7^{-1} \lambda^k(\mathcal{S})$ such that $\mathcal{S} \setminus E$ is k-rectifiable. Since $\mathcal{S}$ is an arbitrary measurable subset, this is enough to prove rectifiability by a standard density argument.

Consider any $\mathcal{S} \subseteq \mathcal{S}_\epsilon^k \cap B_1(0)$. We can assume wlog that $\lambda^k(\mathcal{S}) > 0$, otherwise there is nothing to prove. Consider the function $f(x,r) = \hat{\theta}(x,r) - \hat{\theta}(x,0)$ on $B_1(0)$. This function is monotone nondecreasing in r, uniformly bounded for all $x \in B_1(0)$ and $r \le 1$, and pointwise converging to 0 as $r \to 0$.

Thus, by dominated convergence, for all $\delta > 0$, there exists a radius $\bar{r} > 0$ such that

$$\fint_{\mathcal{S}} f(x, 10\bar{r}) d\lambda^k(x) \le \delta^2. \tag{6.104}$$

Let $E \subset \mathcal{S}$ be a measurable subset with $\lambda^k(E) \le \delta \lambda^k(\mathcal{S})$ and such that $f(x, 10\bar{r}) \le \delta$ for all $x \in F \equiv \mathcal{S} \setminus E$.

Now cover F by a finite number of balls $B_{\bar{r}}(x_i)$ centered on F. We want to show that, if δ is chosen small enough, then on each of these balls we can apply Theorem 42 to $F \cap B_{\bar{r}}(x_i)$ for all i, and thus proving that F is k-rectifiable as desired.

Reifenberg estimates The estimates here are basically equivalent to the estimates carried out in Section 6.2.4. Actually, since we already know that (6.103) holds, we do not even need the upper induction part of that argument. For this reason, we will only sketch the main passages in the estimates.

Fix any i, and consider the set $F \cap B_{\bar{r}}(x_i)$. For convenience, we rescale the ball $B_{\bar{r}}(x_i)$ to $B_1(0)$. With an abuse of notation, we will keep denoting by u, $\hat{\theta}$, $\mathcal{S}_\epsilon^k$ and F also the rescaled objects.

By definition of $F \subset \mathcal{S}_\epsilon^k$, we have that $\hat{\theta}(x, 10) - \hat{\theta}(x, 0) \le \delta$ for all $x \in F$. By an estimate analogous to (6.53), we have for all $x \in F$ and $s \le 1$

$$\beta_{2, \lambda^k|_F}(x, s)^2 \le C_1 s^{-k} \int_{B_s(x)} W_s(y) \, d\lambda^k|_F(y) \tag{6.105}$$

By integrating, and by (6.103), we obtain for all $x \in B_1(0)$ and $s \le r \le 1$:

$$\int_{B_r(x)} \beta_{2, \lambda^k|_F}(z, s)^2 \, d\lambda^k|_F(z) \le C_1 s^{-k} \int_{B_r(y)} \left[\int_{B_s(z)} W_s(t) \, d\lambda^k|_F(t) \right] d\lambda^k|_F(z) \le C_1 C_\epsilon \int_{B_{r+s}(y)} W_s(z) d\lambda^k|_F(z) . \tag{6.106}$$

Integrating again in s, we finally get for all $x \in B_1(0)$ and $r \le 1$:

$$\int_{B_r(x)} \left[\int_0^s \beta_{2, \lambda^k|_F}(z, s)^2 \, \frac{ds}{s} \right] d\lambda^k|_F(z) \le C_1 C_\epsilon \int_{B_{2r}(x)} [\hat{\theta}(x, 8r) - \hat{\theta}(x, 0)] d\lambda^k|_F(z) \le c(m) C_1 C_\epsilon^2 \delta r^k . \tag{6.107}$$

By choosing

$$\delta \le \frac{\delta_R^2}{c(m) C_1 C_\epsilon^2} , \tag{6.108}$$

we can apply Theorem 42 to the set $F \cap B_1(0)$, thus proving that it is k-rectifiable. This concludes the proof.

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** (2015), 1371–1412. MR 3426057.
            Zbl 06521333. http://dx.doi.org/10.1007/s00039-015-0334-7.

Bet93.   F. Bethuel, On the singular set of stationary harmonic maps, *Manuscripta
            Math.* **78** (1993), 417–443. MR 1208652. Zbl 0792.53039. http://dx.doi.
            org/10.1007/BF02599324.

BL15.   C. Breiner and T. Lamm, Quantitative stratification and higher regularity
            for biharmonic maps, *Manuscripta Math.* **148** (2015), 379–398. MR 3414482.
            Zbl 1327.53079. http://dx.doi.org/10.1007/s00229-015-0750-x.
            Available at http://arxiv.org/abs/1410.5640.

CHN13a.   J. Cheeger, R. Haslhofer, and A. Naber, Quantitative stratification and
            the regularity of harmonic map flow, *accepted to Calc. of Var. and PDE* (2013).
            Available at http://arxiv.org/abs/1308.2514.

CHN13b.   _____ , Quantitative stratification and the regularity of mean curvature flow,
            Geom. Funct. Anal. **23** (2013), 828–847. MR 3061773. Zbl 1277.53064.
            http://dx.doi.org/10.1007/s00039-013-0224-9. Available at http://
            arxiv.org/abs/1308.2514.

CN13a.   J. Cheeger and A. Naber, Lower bounds on Ricci curvature and quantitative
            behavior of singular sets, *Invent. Math.* **191** (2013), 321–339. MR 3010378.
            Zbl 1268.53053. http://dx.doi.org/10.1007/s00222-012-0394-3.
            Available at http://arxiv.org/abs/1103.1819.

CN13b.   _____ , Quantitative stratification and the regularity of harmonic maps and
            minimal currents, *Comm. Pure Appl. Math.* **66** (2013), 965–990. MR 3043387.
            Zbl 1269.53063. http://dx.doi.org/10.1002/cpa.21446. Available at
            http://arxiv.org/abs/1107.3097.

CNV15.   J. Cheeger, A. Naber, and D. Valtorta, Critical sets of elliptic equations,
            Comm. Pure Appl. Math. **68** (2015), 173–209. MR 3298662. Zbl 06399723.
            http://dx.doi.org/10.1002/cpa.21518. Available at http://arxiv.
            org/abs/1207.4236.

CG89.   J.-M. Coron and R. Gulliver, Minimizing *p*-harmonic maps into spheres, *J.
            Reine Angew. Math.* **401** (1989), 82–100. MR 1018054. Zbl 0677.58021. http:
            //dx.doi.org/10.1515/crll.1989.401.82.

DS93.   G. David and S. Semmes, *Analysis of and on uniformly rectifiable sets*, *Math-
            ematical Surveys and Monographs* **38**, American Mathematical Society, Prov-
            idence, RI, 1993. MR 1251061. Zbl 0832.42008. http://dx.doi.org/10.
            1090/surv/038.

DT12.   G. David and T. Toro, Reifenberg parameterizations for sets with holes, *Mem.
            Amer. Math. Soc.* **215** (2012), vi+102. MR 2907827. Zbl 1236.28002. http:
            //dx.doi.org/10.1090/S0065-9266-2011-00629-5.

DLL03.   W. Ding, J. Li, and W. Li, Nonstationary weak limit of a stationary harmonic
            map sequence, *Comm. Pure Appl. Math.* **56** (2003), 270–277. MR 1934622.
            Zbl 1034.58012. http://dx.doi.org/10.1002/cpa.10058.

ENV16. N. S. Edelen, A. C. Naber, and D. Valtorta, Quantitative Reifenberg Theorem for Measures, *in preparation* (2016).

FMS15. M. Focardi, A. Marchese, and E. Spadaro, Improved estimate of the singular set of Dir-minimizing Q-valued functions via an abstract regularity result, *J. Funct. Anal.* **268** (2015), 3290–3325. MR 3336726. Zbl 06475693. http://dx.doi.org/10.1016/j.jfa.2015.02.011.

Lin87. F.-H. Lin, A remark on the map $x/|x|$, *C. R. Acad. Sci. Paris Sér. I Math.* **305** (1987), 529–531. MR 916327. Zbl 0652.58022.

Lin99. ———, Gradient estimates and blow-up analysis for stationary harmonic maps, *Ann. of Math. (2)* **149** (1999), 785–829. MR 1709303. Zbl 0949.58017. http://dx.doi.org/10.2307/121073. Available at http://arxiv.org/abs/math/9905214.

Mos05. R. Moser, *Partial regularity for harmonic maps and related problems*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005. MR 2155901. Zbl 1246.58012. http://dx.doi.org/10.1142/9789812701312.

NV. A. Naber and D. Valtorta, Rectifiable-reifenberg and the regularity of stationary and minimizing harmonic maps, *accepted on Annals of Mathematics*. Available at http://annals.math.princeton.edu/articles/10010.

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%2FBF02547186.

Sch84. R. M. Schoen, Analytic aspects of the harmonic map problem, in *Seminar on nonlinear partial differential equations (Berkeley, Calif., 1983)*, *Math. Sci. Res. Inst. Publ.* **2**, Springer, New York, 1984, pp. 321–358. MR 765241. Zbl 0551.58011. http://dx.doi.org/10.1007/978-1-4612-1110-5_17.

Tol15. X. Tolsa, Characterization of n-rectifiability in terms of Jones' square function: part I, *Calc. Var. Partial Differential Equations* **54** (2015), 3643–3665. MR 3426090. Zbl 06544048. http://dx.doi.org/10.1007/s00526-015-0917-z.

Tor95. T. Toro, Geometric conditions and existence of bi-Lipschitz parameterizations, *Duke Math. J.* **77** (1995), 193–227. MR 1317632. Zbl 0847.42011. http://dx.doi.org/10.1215/S0012-7094-95-07708-4.

TW95. T. Toro and C. Wang, Compactness properties of weakly p-harmonic maps into homogeneous spaces, *Indiana Univ. Math. J.* **44** (1995), 87–113. MR 1336433. Zbl 0826.58014. http://dx.doi.org/10.1512/iumj.1995.44.1979.

Xin96. Y. Xin, *Geometry of harmonic maps*, *Progress in Nonlinear Differential Equations and their Applications, 23*, Birkhäuser Boston, Inc., Boston, MA, 1996. MR 1391729 (97h:58052). Zbl 0848.58014. http://dx.doi.org/10.1007/978-1-4612-4084-6.