THE SINGULAR STRUCTURE AND REGULARITY
OF STATIONARY VARIFOLDS

AARON NABER AND DANIELE VALTORTA

ABSTRACT. If one considers an integral varifold $I^m \subseteq M$ with bounded mean curvature, and if $S^k(I) \equiv \{x \in M :$ no tangent cone at x is $k + 1$-symmetric$\}$ is the standard stratification of the singular set, then it is well known that $\dim S^k \leq k$. In complete generality nothing else is known about the singular sets $S^k(I)$. In this paper we prove for a general integral varifold with bounded mean curvature, in particular a stationary varifold, that every stratum $S^k(I)$ is k-rectifiable. In fact, we prove for k-a.e. point $x \in S^k$ that there exists a unique k-plane V^k such that *every* tangent cone at x is of the form $V \times C$ for some cone C.

In the case of minimizing hypersurfaces $I^{n-1} \subseteq M^n$ we can go further. Indeed, we can show that the singular set $S(I)$, which is known to satisfy $\dim S(I) \leq n-8$, is in fact $n-8$ rectifiable with uniformly *finite* $n-8$ measure. An effective version of this allows us to prove that the second fundamental form A has *apriori* estimates in L^7_{weak} on I, an estimate which is sharp as $|A|$ is not in L^7 for the Simons cone. In fact, we prove the much stronger estimate that the regularity scale r_I has L^7_{weak}-estimates.

The above results are in fact just applications of a new class of estimates we prove on the *quantitative* stratifications $S^k_{\epsilon,r}$ and $S^k_{\epsilon} \equiv S^k_{\epsilon,0}$. Roughly, $x \in S^k_{\epsilon} \subseteq I$ if no ball $B_r(x)$ is ϵ-close to being $k + 1$-symmetric. We show that S^k_{ϵ} is k-rectifiable and satisfies the Minkowski estimate $\mathrm{Vol}(B_r \, S^k_{\epsilon}) \leq C_{\epsilon} r^{n-k}$. The proof requires a new L^2-subspace approximation theorem for integral varifolds with bounded mean curvature, and a $W^{1,p}$-Reifenberg type theorem proved by the authors in [NVa].

CONTENTS

1. INTRODUCTION

In this paper we will study integral m-varifolds I^m with bounded mean curvature, and in particular stationary varifolds and area minimizing currents, see Section 2 for an introduction to the basics. In the case of a varifold I with bounded mean curvature we can consider the stratification of I given by

$$S^0(I) \subseteq \cdots \subseteq S^k(I) \subseteq \cdots S^m(I) \subseteq I, \tag{1.1}$$

where the singular sets are defined as

$$S^k(I) \equiv \{x \in M : \text{ no tangent cone at } x \text{ is } k + 1\text{-symmetric}\}, \tag{1.2}$$

see Definition 1.2 for a precise definition and more detailed discussion. An important result of Almgren (see [Alm00, sections 2.25, 2.26]) tells us that in the context of area minimizing currents we have the Hausdorff dimension estimate

$$\dim S^k(I) \leq k. \tag{1.3}$$

With some work, this estimate can be carried over to the context of integral varifolds. Unfortunately, essentially nothing else is understood about the structure of the singular sets in any generality. In the case where I is a codimension 1 area minimizing current more is understood, at least for the top stratum of the singular set. Namely, if I^{n-1} is an area minimizing hypersurface it has been shown in [Sim95] that $S(I) = S^{n-8}(I)$ is rectifiable. This result relies on Simons' work [Sim68], where the author proves that all tangent cones to hypersurfaces in $\mathbb{R}^n$ are hyperplanes if $n \leq 7$.

Moreover, for a minimizing hypersurface, i.e., a minimizing current of codimension 1, one can use the ϵ-regularity theorem of [DG61], (see also [Fed69],[CN13b] and theorem 2.11) to show that the whole singular set $S(I)$ coincides with $S^{n-8}(I)$. As it is well-known, this is not the case for minimizing currents of codimension higher than one, as singular points in this context may also arise as branching points on the top stratum $S^m(I) \setminus S^{m-1}(I)$.

The first goal of this paper is to give improved regularity results for stratification and associated quantitative stratification for varifolds under only a bounded mean curvature assumption. Indeed, we will show for such a varifold that the strata $S^k(I)$ are k-rectifiable for every k. In fact, we will show that for k-a.e. $x \in S^k(I)$ there exists a unique k-dimensional subspace $V \subseteq T_xM$ such that $every$ tangent cone at x is of the form $V^k \times C$ for some cone C. Note that we are not claiming that the cone factor C is necessarily unique, just that the Euclidean factor V^k of the cone is. It is a good question as to whether tangent cones need to be unique in general under only the assumption of bounded mean curvature.

For a varifold $I^m = I^{n-1}$ which is an area minimizing hypersurface these results can be improved. To begin with, we have that the top stratum of the singular set $S(I) = S^{n-8}(I)$ is $n - 8$ rectifiable, and that there are $a\ priori$ bounds on its $n - 8$ Hausdorff measure. That is, if the mass $\mu_I(B_2) \leq \Lambda$ is bounded, then we have the estimate $H^{n-8}(S(I) \cap B_1) \leq C(\Lambda, B_2)$. In fact, we have the stronger Minkowski estimate

$$\text{Vol}\left(B_r\left(S(I)\right) \cap B_1\right), \, r\mu_I\left(B_r\left(S(I)\right) \cap B_1\right) \leq Cr^8, \tag{1.4}$$

where $B_r\left(S(I)\right)$ is the tube of radius r around the singular set. Indeed, we can prove much more effective versions of these estimates. That is, in theorem 1.8 we show that the second fundamental form A of I, and

in fact the regularity scale r_I, have *a priori* bounds in L^7_{weak}. More precisely, we have that

$$\mu_I\left(\left\{|A| > r^{-1}\right\} \cap B_1\right) \le \mu_I\left(B_r\left\{|A| > r^{-1}\right\} \cap B_1\right) \le C(\Lambda, B_2)r^7. \tag{1.5}$$

Let us observe that these estimates are sharp, in that the Simons' cone satisfies $|A| \notin L^7_{loc}(I)$. Let us also point out that this sharpens estimates of [CN13b], where it was proven that for minimizing hypersurfaces that $|A| \in L^p$ for all $p < 7$. We refer the reader to Section 1.2 for the precise and most general statements.

Now the techniques of this paper all center around the notion of the *quantitative* stratification. In fact, it is for the quantitative stratification that the most important results of the paper hold, everything else can be seen to be corollaries of these statements. The quantitative stratification was first introduced in [CN13a], and later used in [CN13b] with the goal of giving effective and L^p estimates on stationary and minimizing currents. It 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. More recently, in [NVa] the authors have used ideas similar to those in this paper to prove structural theorems for the singular sets of stationary harmonic maps.

Before describing the results in this paper on the quantitative stratification, let us give more precise definitions of everything. To begin with, to describe the stratification and quantitative stratification we need to discuss the notion of symmetry associated to an integral varifold with bounded mean curvature. Specifically:

Definition 1.1. For $y \in \mathbb{R}^n$ and $\lambda > 0$, let $\eta_{y,\lambda} : \mathbb{R}^n \to \mathbb{R}^n$ and $\tau_y : \mathbb{R}^n \to \mathbb{R}^n$ be the functions

$$\eta_{y,\lambda}(x) = y + \frac{x - y}{\lambda}, \quad \tau_y(x) = x + y, \tag{1.6}$$

and let $\eta_{y,\lambda\,\#}$ and $\tau_{y\#}$ be their pushforwards. We define the following:

(1) An integral varifold $I^m \subseteq \mathbb{R}^n$ is called k-symmetric if $\eta_{0,\lambda\,\#}I = I \; \forall \; \lambda > 0$ (which means that I is a cone at the origin), and if there exists a k-plane $V^k \subseteq \mathbb{R}^n$ such that for each $y \in V^k$ we have that $\tau_{y\#}I = I$.

(2) Given an integral varifold $I^m \subseteq M$ and $\epsilon > 0$, we say a ball $B_r(x) \subseteq M$ with $r < \text{inj}(x)$ and $x \in I^m$ is (k, ϵ)-symmetric if there exists a k-symmetric integral varifold $\tilde{I}^m \subseteq T_xM$ such that $d(\eta_{0,r\,\#}I^m \cap B_1(0), \tilde{I}^m \cap B_1(0)) < \epsilon$, where we have used the exponential map to identify I as a varifold on T_xM.

Remark 1.1. The distance d may be taken to be the weak distance induced by the Frechet structure of varifold convergence. In the case of minimizing currents, it is equivalent to take d to be the flat distance.

Thus, an integral varifold I^m is k-symmetric if $I = V^k \times C$ for some cone C. A varifold is (k, ϵ)-symmetric on a ball $B_r(x)$ if it is weakly close to a k-symmetric varifold on this ball.

With the notion of symmetry in hand, we can define precisely the quantitative stratification associated to a solution. The idea is to group points together based on the amount of symmetry that balls centered at those points contain. In fact, there are several variants which will play a role for us. Let us introduce them all and briefly discuss them:

Definition 1.2. For an integral varifold I^m with bounded mean curvature and finite density we make the following definitions:

(1) For $\epsilon, r > 0$ we define the kth (ϵ, r)-stratification $S^k_{\epsilon,r}(I)$ by

$$S^k_{\epsilon,r}(I) \equiv \{x \in I \cap B_1 : \text{ for no } r \leq s < 1 \text{ is } B_s(x) \text{ a } (k+1, \epsilon)\text{-symmetric ball}\}. \tag{1.7}$$

(2) For $\epsilon > 0$ we define the kth ϵ-stratification $S^k_\epsilon(I)$ by

$$S^k_\epsilon(I) = \bigcap_{r>0} S^k_{\epsilon,r}(I) \equiv \{x \in I \cap B_1 : \text{ for no } 0 < r < 1 \text{ is } B_r(x) \text{ a } (k+1, \epsilon)\text{-symmetric ball}\}. \tag{1.8}$$

(3) We define the kth-stratification $S^k(I)$ by

$$S^k(I) = \bigcup_{\epsilon>0} S^k_\epsilon(I) = \{x \in I \cap B_1 : \text{ no tangent cone at } x \text{ is } k+1\text{-symmetric}\}. \tag{1.9}$$

Remark 1.2. It is a small but important exercise to check that the standard stratification $S^k(I)$ as defined in (1.1) agrees with the set $\bigcup_{\epsilon>0} S^k_\epsilon(I)$.

Let us discuss in words the meaning of the quantitative stratification, and how it relates to the standard stratification. As discussed at the beginning of the section, the stratification $S^k(I)$ of I is built by separating points of I based on the infinitesimal symmetries of I at those points. The quantitative stratifications $S^k_\epsilon(I)$ and $S^k_{\epsilon,r}(I)$ are, on the other hand, instead built by separating points of I based on how many symmetries exist on balls of definite size around the points. In practice, the quantitative stratification has two advantages over the standard stratification. First, on minimizing hypersurfaces the quantitative stratification allows to prove effective estimates. In particular, in [CN13b] the L^p estimates

$$\fint_{B_1 \cap I} |A|^{7-\delta} < C_\delta \; \forall \delta > 0, \tag{1.10}$$

on minimizing hypersurfaces were proved by exploiting this fact. The second advantage is that the estimates on the quantitative stratification are much stronger than those on the standard stratification. Namely, in [CN13b] the Hausdorff dimension estimate (1.3) on $S^k(I)$ was improved to the Minkowski content estimate

$$\text{Vol}\left(B_r\left(S^k_{\epsilon,r}\right)\right) \leq C_{\epsilon,\delta} r^{n-k-\delta} \; \forall \delta > 0. \tag{1.11}$$

One of the key technical estimates of this paper is that in theorem 1.3 we drop the δ from this estimate and obtain an estimate of the form

$$\text{Vol}\left(B_r\left(S^k_{\epsilon,r}\right)\right) \leq C_\epsilon r^{n-k}. \tag{1.12}$$

From this we are able to conclude in theorem 1.4 an estimate on S^k_ϵ of the form

$$\text{Vol}\left(B_r\left(S^k_\epsilon\right)\right) \leq C_\epsilon r^{n-k}. \tag{1.13}$$

In particular, this estimate allows us to conclude that S^k_ϵ has uniformly finite k-dimensional measure. In fact, the techniques will prove much more for us. They will show us that S^k_ϵ is k-rectifiable, and that for k-a.e. point $x \in S^k_\epsilon$ there is a unique k-plane $V^k \subseteq T_x M$ such that *every* tangent cone at x is k-symmetric with respect to V. By observing that $S^k(I) = \bigcup S^k_\epsilon(I)$, this is what allows us to prove in theorem 1.5 our main results on the classical stratification. This decomposition of S^k into the pieces S^k_ϵ is crucial for the proof.

On the other hand, (1.13), combined with the ϵ-regularity theorems of [Fed69],[CN13b], allow us to conclude in the minimizing hypersurface case both the weak L^7 estimate on $|A|$, and the $m - 7$-finiteness of the singular set of I. Thus we will see that theorems 1.6 and 1.8 are fairly quick consequences of (1.13).

Thus we have seen that (1.12) and (1.13), and more generally theorem 1.3 and theorem 1.4, are the main challenges of the paper. We will give a more complete outline of the proof in Section 1.3, however let us mention for the moment that two of the new ingredients to the proof are a new L^2-subspace approximation theorem for integral varifolds with bounded mean curvature, proved in Section 6, and a $W^{1,p}$-Reifenberg theorem described in Section 3. The L^2-approximation result roughly tells us that the L^2-distance of a measure from being contained in a k-dimensional subspace may be estimated by integrating the volume drop of the integral varifold over the measure. To exploit the estimate we prove a new $W^{1,p}$-Reifenberg type theorem. The classical Reifenberg theorem states that if we have a set S which is L^∞-approximated by an affine k-dimensional subspace at every point and scale, then S is bi-Hölder to a k-dimensional manifold, see theorem 2.18 for a precise statement. It is important for us to improve on this bi-Hölder estimate, at least enough that we are able to control the kth-dimensional measure of the set and prove rectifiability. In particular, we want to improve the C^α-maps to $W^{1,p}$-maps for $p > k$, and we will want to do it using a condition which is integral in nature. More precisely, we will only require a form of summable L^2-closeness of the subset S to the approximating subspaces. We will see in theorem 6.1 that by using the L^2-subspace approximation theorem that the conditions of this new $W^{1,p}$-Reifenberg are in fact controllable for the quantitative stratifications S^k_ϵ.

1.1. **Results for Varifolds with Bounded Mean Curvature.**

We now turn our attention to giving precise statements of the main results of this paper. In this subsection we focus on those concerning the singular structure of integral varifolds with bounded mean curvature. Precisely, let (M^n, g, p) be a Riemannian manifold satisfying

$$|\sec_{B_2(p)}| \le K, \ \mathrm{inj}(B_2(p)) \ge K^{-1}, \tag{1.14}$$

and let I^m be an integral varifold on M with mean curvature bounded by H on $B_2(p)$. That is, for every smooth vector field X on M with compact support in $B_2(p)$ we have the estimate

$$\left|\delta\mu_I(X)\right| \le H \int_I |X|\, dI, \tag{1.15}$$

see Section 2.1 for more on this. Let us begin by discussing our main theorem for the quantitative stratifications $S^k_{\epsilon,r}(I)$:

Theorem 1.3 ((ϵ, r)-Stratification of Integral Varifolds with BMC). *Let I^m be an integral varifold on M^n satisfying the curvature bound (1.14), the mean curvature bound (1.15), and the mass bound $\mu_I(B_2(p)) \le \Lambda$. Then for each $\epsilon > 0$ there exists $C_\epsilon(n, K, H, \Lambda, \epsilon)$ such that*

$$Vol\Big(B_r\big(S^k_{\epsilon,r}(I)\big) \cap B_1(p)\Big) \le C_\epsilon r^{n-k}. \tag{1.16}$$

When we study the quantitative stratification $S_\epsilon^k(I)$ we can refine the above to prove structure theorems on the set itself:

Theorem 1.4 (ϵ-Stratification of Integral Varifolds with BMC). *Let I^m be an integral varifold on M^n satisfying the curvature bound* (1.14)*, the mean curvature bound* (1.15)*, and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. Then for each $\epsilon > 0$ there exists $C_\epsilon(n, K, H, \Lambda, \epsilon)$ such that*

$$Vol\left(B_r(S_\epsilon^k(I)) \cap B_1(p)\right) \leq C_\epsilon r^{n-k} . \tag{1.17}$$

In particular, we have the kth-dimensional Hausdorff measure estimate $\lambda^k(S_\epsilon^k(I)) \leq C_\epsilon$. Further, $S_\epsilon^k(I)$ is k-rectifiable, and for k-a.e. $x \in S_\epsilon^k$ there exists a unique k-plane $V^k \subseteq T_xM$ such that every tangent cone of x is k-symmetric with respect to V^k.

Finally, we end this subsection by stating our main results when it comes to the classical stratification $S^k(I)$. The following may be proved from the previous theorem in only a few lines given the formula $S^k(I) = \bigcup S_\epsilon^k(I)$:

Theorem 1.5 (Stratification of Integral Varifolds with BMC). *Let I^m be an integral varifold on M^n satisfying the curvature bound* (1.14)*, the mean curvature bound* (1.15)*, and the mass bound $\mu_I(B_2(p)) < \infty$. Then for each k we have that $S^k(I)$ is countably k-rectifiable. Further, for k-a.e. $x \in S^k(I)$ there exists a unique k-plane $V^k \subseteq T_xM$ such that every tangent cone of x is k-symmetric with respect to V^k.*

1.2. Results for Minimizing Hypersurfaces. In this section we restrict ourselves to studying codimension one integral currents I^{n-1} which minimize the area functional on $B_2(p)$ with respect to compact variations. That is, if I' is another integral current with $\partial I = \partial I'$ in $B_2(p)$ and $\text{supp}\,(I - I') \Subset B_2(p)$, then $\mu_I(B_2) \leq \mu'_I(B_2)$. One could easily restrict to local minimizers to obtain similar results. Most of the results of this section follow quickly by combining the quantitative stratification results of Section 1.1 with the ϵ-regularity of [Fed69],[CN13b], see Section 2.5 for a review of these points.

Our first estimate is on the singular set $\text{Sing}(I)$ of a minimizing hypersurface. Recall that $\text{Sing}(I)$ is the set of points where I is not smooth.

Theorem 1.6 (Structure of Singular Set). *Let $I^m = I^{n-1}$ be a minimizing integral current on M^n with $\partial I \cap B_2(0) = 0$ satisfying the curvature bound* (1.14) *and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. Then $\text{Sing}(I)$ is $m - 7$-rectifiable and there exists $C(n, K, \Lambda)$ such that*

$$\text{Vol}\left(B_r(\text{Sing}(I)) \cap B_1(p)\right) \leq C r^8 ,$$
$$\mu_I\left(B_r(\text{Sing}(I)) \cap B_1(p)\right) \leq C r^7 . \tag{1.18}$$

In particular, $\lambda^{m-7}(\text{Sing}(I)) \leq C$.

The above can be extended to effective regularity estimates on I. To state the results in full strength let us recall the notion of the regularity scale associated to a function. Namely:

Definition 1.7. Let I^m be an integral varifold. For $x \in I \cap B_1(p)$ we define the regularity scale $r_I(x)$ by

$$r_I(x) \equiv \max \left\{ 0 \le r \le 1 : \sup_{B_r(x)} |A| \le r^{-1} \right\}. \tag{1.19}$$

By definition, $r_I(x) \equiv 0$ if I is not C^2 in a neighborhood of x.

Remark 1.3. The regularity scale is *scale invariant*. That is, if $r \equiv r_I(x)$ and we rescale $B_r(x) \to B_1(x)$, then on the rescaled ball we will have $|A| \le 1$ on $B_1(x)$.

Remark 1.4. We have the easy estimate $|A|(x) \le r_I(x)^{-1}$. However, a lower bound on $r_I(x)$ at a point is in principle *much* stronger than an upper bound on $|A|(x)$.

Remark 1.5. Notice that the regularity scale is a Lipschitz function with $|\nabla r_I| \le 1$, and it is an upper semicontinuous function.

Remark 1.6. If I satisfies an elliptic equation, e.g. is a stationary varifold, then we have estimates of the form

$$\sup_{B_{r_I/2}(x)} |\nabla^k A| \le C_k \, r_I(x)^{-(k+1)}. \tag{1.20}$$

In particular, control on r_I gives control on all higher order derivatives.

Now let us state our main estimates for minimizing hypersurfaces:

Theorem 1.8 (Estimates on Minimizing Hypersurfaces). *Let $I^m = I^{n-1}$ be a minimizing integral current on M^n with $\partial I \cap B_2(0) = 0$ satisfying the curvature bound (1.14) and the mass bound $\mu_I(B_2(p)) \le \Lambda$. Then there exists $C(n, K, \Lambda)$ such that*

$$Vol\Big(\{x \in B_1(p) : |A| > r^{-1}\} \cap B_1(p)\Big) \le Cr^8,$$

$$Vol\Big(\{x \in B_1(p) : r_I(x) < r\} \cap B_1(p)\Big) \le Cr^8,$$

$$\mu_I\Big(\{x \in B_1(p) : r_I(x) < r\} \cap B_1(p)\Big) \le Cr^7. \tag{1.21}$$

In particular, both $|A|$ and r_I^{-1} have bounds in $L^7_{weak}\big(I \cap B_1(p)\big)$, the space of weakly L^7 functions on I wrt the measure μ_I.

1.3. **Outline of Proofs and Techniques.** In this subsection we give a brief outline of the proof of the main theorems. To describe the new ingredients involved it will be helpful to give a comparison to the proofs of previous results in the area, in particular the dimension estimate (1.3) of Federer and the Minkowski and L^p estimates (1.10) of [CN13b].

Indeed, the starting point for the study of singular sets for solutions of geometric equations typically looks the same, that is, one needs a monotone quantity. In the case of integral varifolds I^m with bounded mean curvature, we consider the volume density

$$\theta_r(x) \equiv r^{-m}\mu_I(B_r(x)). \tag{1.22}$$

For simplicity sake let us take $M \equiv \mathbb{R}^n$ and I^m to be stationary in this discussion, which is really of no loss except for some small technical work. Then $\frac{d}{dr}\theta_r \geq 0$, so $\theta_r(x)$ is precisely a monotone quantity, and it is independent of r if and only if I is 0-symmetric, see Section 2.3 for more on this. Interestingly, this is the only information one requires to prove the dimension estimate (1.3). Namely, since $\theta_r(x)$ is monotone and bounded, it must converge as r tends to zero. Therefore, if we consider the sequence of scales $\mathfrak{r}_\alpha = 2^{-\alpha}$ then we have for each x that

$$\lim_{\alpha \to \infty} \left| \theta_{\mathfrak{r}_\alpha}(x) - \theta_{\mathfrak{r}_{\alpha+1}}(x) \right| \to 0. \tag{1.23}$$

From this one can conclude that *every* tangent cone of I is 0-symmetric. This fact combined with some very general dimension reduction arguments originating with Federer from geometric measure theory [Sim83], yield the dimension estimate (1.3).

The improvement in [CN13b] of the Hausdorff dimension estimate (1.3) to the Minkowski content estimate (1.11), and from there the L^p estimate of (1.10), requires exploiting more about the monotone quantity $\theta_r(x)$ than just that it limits as r tends to zero. Indeed, an effective version of (1.23) says that for each $\delta > 0$ there exists $N(\Lambda, \delta) > 0$ such that

$$\left| \theta_{\mathfrak{r}_\alpha}(x) - \theta_{\mathfrak{r}_{\alpha+1}}(x) \right| < \delta \tag{1.24}$$

holds for all except for at most N scales $\alpha \in \{\alpha_1, \ldots, \alpha_N\} \subseteq \mathbb{N}$. These *bad* scales where (1.24) fails may differ from point to point, but the number of such scales is uniformly bounded. This allows one to conclude that, for all but at most $N(\Lambda, \epsilon)$-scales, $B_{\mathfrak{r}_{alpha}}(x)$ is $(0, \epsilon)$-symmetric, see Section 2.4. To exploit this information, a new technique other than dimension reduction was required in [CN13b]. Indeed, in [CN13b] the quantitative 0-symmetry of (1.24) was instead combined with the notion of cone splitting and an energy decomposition in order to conclude the estimates (1.10),(1.11). Since we will use them in this paper, the quantitative 0-symmetry and cone splitting will be reviewed further in Section 2.4.

Now let us begin to discuss the results of this paper. The most challenging aspect of this paper is the proof of the estimates on the quantitative stratifications of theorems 1.3 and 1.4, and so we will focus on these in our outline. Let us first observe that it might be advantageous to replace (1.24) with a version that forces an actual rate of convergence, see for instance [Sim93], [Sim95]. More generally, if one is in a context where an effective version of tangent cone uniqueness can be proved then this may be exploited. In fact, in the context of critical sets of elliptic equations one can follow exactly this approach, see the authors' work [NVb] where versions of theorems 1.3 and 1.4 were first proved in this context. However, in the general context of this paper such an approach fails, as tangent cone uniqueness is not available, and potentially not

correct.

Instead, we will first replace (1.24) with the following relatively simple observation. Namely, for each x there exists $N(\Lambda, \delta)$ and a finite number of scales $\{\alpha_1, \ldots, \alpha_N\} \subseteq \mathbb{N}$ such that

$$\sum_{\alpha_j < \alpha < \alpha_{j+1}} \left| \theta_{r_\alpha}(x) - \theta_{r_{\alpha+1}}(x) \right| < \delta. \tag{1.25}$$

That is, not only does the mass density drop by less than δ between these scales, but the sum of the all the mass density drops is less than δ between these scales.

Unfortunately, exploiting (1.25) in order to prove estimates on the singular set turns out to be *substantially* harder to use than exploiting either (1.23) or even (1.24). In essence, this is because it is not a local assumption in terms of scale, and one needs estimates which can see many scales simultaneously, but which do not require any form of tangent cone uniqueness statements. Accomplishing this requires four new ideas, two of which have been introduced in the last few years in [CN13b],[NVb], and two of which are new to this paper. The first point is to replace the study of the singular set with the study of the quantitative singular set, as introduced in [CN13b] for integral varifolds. It will be clear from the proofs that there is no direct way to apply the ideas of this paper to the singular set itself without decomposing it into these quantitative pieces. The sharp estimates themselves also depend on a new covering argument, first introduced by the authors in [NVb]. We will only briefly discuss the covering in the outline, but this covering argument has the advantage of giving very sharp packing estimates on sets and exploits well the condition (1.25), see Section 7 for more on this in the context of this paper. The disadvantage of the strategy of this covering is that it requires comparing balls of arbitrarily different sizes. In [NVb] this was accomplished by proving an effective tangent cone uniqueness statement in the context of critical sets of elliptic equations. Unfortunately, in the context of this paper it is not clear such a statement exists.

Thus, we arrive at discussing the new ideas supplied in this paper. As discussed, in order to apply the strategy of the covering argument of [NVb] we need to be able to estimate collections of balls of potentially arbitrarily different radii. Accomplishing this requires two ingredients, a rectifiable-Reifenberg type theorem, and a new L^2-best subspace approximation theorem for integral varifolds with bounded mean curvature, which will allow us to apply the rectifiable-Reifenberg. Let us discuss these two ingredients separately.

We begin by briefly discussing the $W^{1,p}$-Reifenberg and rectifiable-Reifenberg theorems, which are stated in Section 3. Recall that the classical Reifenberg theorem, reviewed in Section 2.7, gives criteria under which a set becomes C^α-Hölder equivalent to a ball $B_1(0^k)$ in the Euclidean space $\mathbb{R}^k$. In the context of this paper, it is important to improve on this result so that we have gradient and volume control of our set. Let us remark that there have been many generalizations of the classical Reifenberg theorem in the literature, see for instance [Tor95] [DT12], however those results have hypotheses which are much too strong for the purposes of this paper. Instead, we will follow the approach introduced by the authors in [NVa] to prove rectifiability and Minkowski-type bounds on the singular sets of harmonic maps. In particular, we need to improve the C^α-equivalence to a $W^{1,p}$-equivalence. This is strictly stronger by Sobolev embedding, and if $p > k$ then this results in volume estimates and a rectifiable structure for the set. More generally, we will

require a version of the theorem which allows for more degenerate structural behavior, namely a rectifiable-Reifenberg theorem. In this case, the assumptions will conclude that a set S is rectifiable with volume estimates. Of course, what is key about this result is that the criteria will be checkable for our quantitative stratifications, thus let us discuss this criteria briefly. Roughly, if $S \subseteq B_1(0^n)$ is a subset equipped with the k-dimensional Hausdorff measure λ^k, then let us define the k-dimensional distortion of S by

$$D_S^k(y, s) \equiv s^{-2} \inf_{L^k} s^{-k} \int_{S \cap B_s(x)} d^2(y, L^k) \, d\lambda^k(y), \tag{1.26}$$

where the inf is over all k-dimensional affine subspaces of $\mathbb{R}^n$. That is, D^k measures how far S is from being contained in a k-dimensional subspace. Our rectifiable-Reifenberg then requires this be small on S in an integral sense, more precisely that

$$r^{-k} \int_{S \cap B_r(x)} \sum_{\mathfrak{r}_\alpha \leq r} D^k(y, \mathfrak{r}_\alpha) d\lambda^k(x) < \delta^2 \,. \tag{1.27}$$

For δ sufficiently small, the conclusions of the rectifiable-Reifenberg theorem 3.4 are that the set S is rectifiable with effective bounds on the k-dimensional measure. Let us remark that one cannot possibly conclude better than rectifiable under this assumption, see for instance the Examples of Section 3.3.

Thus, in order to prove the quantitative stratification estimates of theorems 1.3 and 1.4, we will need to verify that the integral conditions (1.27) hold for the quantitative stratifications $S_\epsilon^k(I)$, $S_{\epsilon,r}^k(I)$ on all balls $B_r(x)$. In actuality the proof is more complicated. We will need to apply a discrete version of the rectifiable-Reifenberg, which will allow us to build an iterative covering in Section 7 of the quantitative stratifications, and each of these will satisfy (1.27). This will allow us to keep effective track of all the estimates involved. However, let us for the moment just focus on the main estimates which allows us to turn (1.27) into information about our integral varifolds, without worrying about such details.

Namely, in Section 6 we prove a new and very general approximation theorem for integral varifolds with bounded mean curvature. As always in this outline, let us assume $M \equiv \mathbb{R}^n$ and that I^m is stationary, the general case is no harder. Thus we consider an arbitrary measure μ which is supported on $B_1(0^n)$. We would like to study how closely the support of μ can be approximated by a k-dimensional affine subspace $L^k \subseteq \mathbb{R}^n$, in the appropriate sense, and we would like to estimate this distance by using properties of I. Indeed, if we assume that $B_8(0^n)$ is *not* $(k + 1, \epsilon)$-symmetric with respect to I, then for an arbitrary μ we will prove in theorem 6.1 that

$$\inf_{L^k \subseteq \mathbb{R}^n} \int_{B_1(0)} d^2(x, L^k) \, d\mu \leq C \int_{B_1(0)} |\theta_8(x) - \theta_1(x)| \, d\mu, \tag{1.28}$$

where C will depend on ϵ, the mass of I, and the geometry of M. That is, if I does not have $k + 1$ degrees of symmetry, then how closely the support of an arbitrary measure μ can be approximated by a k-dimensional subspace can be estimated by looking at the mass drop of I along μ. In applications, μ will be the restriction to B_1 of some discrete approximation of the k-dimensional Hausdorff measure on S_ϵ^k, and thus the symmetry assumption on I will hold for all balls centered on the support of μ. Notice that if the ambient space is not $\mathbb{R}^n$ but a smooth manifold, (1.28) has an extra term on the rhs as in (6.1), but this term is easily seen to be

not a problem for the estimates given its exponential decay.

In practice, applying (1.28) to (1.27) is subtle and must be done inductively on scale. Additionally, in order to prove the effective Hausdorff estimates $\lambda^k(S^k_\epsilon \cap B_r) \leq Cr^k$ we will need to use the Covering lemma 7.1 to break up the quantitative stratification into appropriate pieces, and we will apply the estimates to these. This decomposition is based on a covering scheme first introduced by the authors in [NVb]. Thus for the purposes of our outline, let us assume we have already proved the Hausdorff estimate $\lambda^k(S^k_\epsilon \cap B_r) \leq Cr^k$, and use this to be able to apply the rectifiable-Reifenberg in order to conclude the rectifiability of the singular set. This may feel like a cheat, however it turns out the proof of the Hausdorff estimate will itself be proved by a very similar tactic, though will also require an inductive argument on scale and use the discrete rectifiable-Reifenberg of theorem 3.3 in replace of the rectifiable-Reifenberg of theorem 3.4.

Thus let us choose a ball B_r and let $E \equiv \sup_{B_r} \theta_r(y)$. Let us consider the subset $\tilde{S}^k_\epsilon \subseteq S^k_\epsilon \cap B_r$ defined by

$$\tilde{S}^k_\epsilon \equiv \{y \in S^k_\epsilon \cap B_r : \theta_0(y) > E - \eta\}, \tag{1.29}$$

where $\eta = \eta(n, K, H, \Lambda, \epsilon)$ will be chosen appropriately later. We will show now that $\tilde{S}^k_\epsilon$ is rectifiable. Since η is fixed and the ball B_r is arbitrary, the rectifiability of all of S^k_ϵ will follow quickly by an easy covering argument. Thus, let us estimate (1.27) by plugging in (1.28) and the Hausdorff estimate to conclude:

$$r^{-k} \int_{\tilde{S}^k_\epsilon} \sum_{\mathfrak{r}_\alpha \leq r} D^k(x, \mathfrak{r}_\alpha) \, d\lambda^k$$

$$= r^{-k} \int_{\tilde{S}^k_\epsilon} \sum_{\mathfrak{r}_\alpha \leq r} \left(\inf_{L^k} \mathfrak{r}_\alpha^{-2-k} \int_{\tilde{S}^k_\epsilon \cap B_{\mathfrak{r}_\alpha}(x)} d^2(y, L^k) d\lambda^k(y) \right) d\lambda^k(x)$$

$$\leq Cr^{-k} \int_{\tilde{S}^k_\epsilon} \sum_{\mathfrak{r}_\alpha \leq r} \left(\mathfrak{r}_\alpha^{-k} \int_{\tilde{S}^k_\epsilon \cap B_{\mathfrak{r}_\alpha}(x)} |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| d\lambda^k(y) \right) d\lambda^k(x)$$

$$= Cr^{-k} \sum_{\mathfrak{r}_\alpha \leq r} \mathfrak{r}_\alpha^{-k} \int_{\tilde{S}^k_\epsilon} \lambda^k(\tilde{S}^k_\epsilon \cap B_{\mathfrak{r}_\alpha}(y)) |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| \, d\lambda^k(y)$$

$$\leq Cr^{-k} \int_{\tilde{S}^k_\epsilon \cap B_r(x)} \sum_{\mathfrak{r}_\alpha \leq r} |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| \, d\lambda^k(y)$$

$$\leq Cr^{-k} \int_{\tilde{S}^k_\epsilon} |\theta_{8r}(y) - \theta_0(y)| \, d\lambda^k(y)$$

$$\leq Cr^{-k} \lambda_k(\tilde{S}^k_\epsilon) \cdot \eta$$

$$< \delta^2, \tag{1.30}$$

where in the last line we have chosen $\eta = \eta(n, K, H, \Lambda, \epsilon)$ so that the estimate is less than the required δ from the rectifiable-Reifenberg. Notice also that the bound $\sum_{\mathfrak{r}_\alpha \leq r} |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| \leq C(\theta_{16r}(y) - \theta_0(y))$ follows from the monotonicity of θ and the fact that the sum on the lhs can be bounded with telescopic sums. Indeed,

$$\sum_{\mathfrak{r}_\alpha \leq r} |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| = \sum_{i=1}^{3} \sum_{\alpha \bmod(3) = i, \ \mathfrak{r}_\alpha \leq r} |\theta_{8\mathfrak{r}_\alpha}(y) - \theta_{\mathfrak{r}_\alpha}(y)| \leq 3(\theta_{16r}(y) - \theta_0(y)). \tag{1.31}$$

Thus, if r is small enough, we can apply the rectifiable-Reifenberg of theorem 3.4 in order to conclude the rectifiability of the set $\tilde{S}^k_\epsilon$, which in particular proves that S^k_ϵ is itself rectifiable, as claimed.

2. PRELIMINARIES

2.1. **Integral Varifolds and Mean Curvature.** Let us begin by giving a very brief introduction to integral varifolds, the first variation, and their relationship to mean curvature. See [DL12, Sim83] for a more complete and nice introduction. We begin with the definition of an integral varifold:

Definition 2.1. Given a smooth Riemannian manifold (M^n, g), consider the set of pairs (S, θ), where S is a countably m-dimensional rectifiable set in M^n and θ is a positive function locally integrable wrt λ^m_S. We define the equivalence relation $(S, \theta) \sim (S', \theta')$ by asking that $\lambda^m(S \setminus S') + \lambda^m(S' \setminus S) = 0$ and $\theta = \theta'$ λ^m-a.e. on $S \cap S'$. An m-dimensional rectifiable varifold I^m in (M^n, g) is the equivalence class of pairs (S, θ). We say that I is an *integral m-rectifiable varifold*, or simply integral m-varifold, if the function θ takes values in the positive integers λ^m_S-a.e.

Remark 2.1. From now on, with an abuse of notation, we will denote the varifold I simply by one representative in its equivalence class (S, θ).

We associate to I the measure $\mu_I \equiv \theta \, d\lambda^m_S$, where λ^m_S is the m-dimensional Hausdorff measure restricted to S. Then the total mass of I can be denoted

$$\mu_I(M) = \int_M d\mu_I = \int_M \theta d\lambda^m_S \,. \tag{2.1}$$

To define the mean curvature of an integral varifold we begin by recalling the notion of the first variation. Given a smooth vector field X on M with compact support, let ϕ^X_t be the family of diffeomorphisms generated by X, then we can define the first variation of I as the distribution

$$\delta\mu_I(X) \equiv \frac{d}{dt}\Big|_{t=0} |\phi_{t\#}I| = \int_M \operatorname{div}_I X \, d\mu_I \,, \tag{2.2}$$

where $\operatorname{div}_I(X)$ is the divergence of I on the tangent space of I_S, which is well defined a.e. We can now say that I has mean curvature bounded by H (or first variation bounded by H) on an open set $U \subseteq M$ if for every smooth vector field whose support is in U we have that

$$\left|\delta\mu_I(X)\right| \le H \int |X| \, d\mu_I \,. \tag{2.3}$$

We say that I is a stationary integral varifold on U if the mean curvature is bounded by 0, and in particular we have that I is then a critical point of the area functional, with respect to variations in U.

An important compactness theorem for integral m-varifolds due to Allard states that given a sequence of integral varifolds with uniformly bounded mass and mean curvature has a converging subsequence in the same class. Here we state the theorem for the reader's convenience, and refer to [All72] or [Sim83, chapter 8] for the proof and more details on the subject.

Theorem 2.2 (Allard compactness theorem). *Let I_j be a sequence of integral m-varifolds in (M^n, g) with mass bound $\mu_{I_j}(M) \leq \Lambda$ and mean curvature uniformly bounded by H. Then, up to passing to a subsequence, I_j converges in the sense of varifolds to some integral m-varifold I with $\mu_I(M) \leq \Lambda$ and mean curvature bounded by H.*

2.2. **Minimizing integral currents.** Next, let us consider the class of integral currents. These objects arise naturally in the study of minimal surfaces, here we briefly recall their definition and main properties. We refer the reader to [Fed69, Mor00, DL15] for a more complete introduction on currents.

Integral currents arise naturally in the study of the Plateau's problem. The lack of compactness in the family of classical manifolds with respect to their volume makes it natural to introduce a sort of "weak version" of these objects in order to apply classical variational methods to prove the existence of minimizers. In this spirit, De Rham defines the currents as duals of smooth forms in a domain.

Definition 2.3. Let $m \leq n$. Given a domain $\Omega \subseteq M^n$, let $\Lambda^m(\Omega)$ be the space of smooth compactly supported m-forms on Ω with the strong topology. We denote by $\|\lambda\|$ the comass of the form λ, i.e.

$$\|\lambda\|_c = \max \{|\langle \lambda(p), v_1 \wedge \cdots \wedge v_m \rangle| \text{ with } p \in \Omega \text{ and } |v_1 \wedge \cdots v_m| = 1\}, \tag{2.4}$$

where $|v_1 \wedge \cdots v_m|$ is the m-dimensional Hausdorff measure in $T_p(M)$ of the parallelogram determined by the tangent vectors $v_1, \cdots, v_m$.

An m-dimensional current I is a continuous linear functional $I : \Lambda^m \to \mathbb{R}$.

By integration, it is evident that any smooth m-dimensional orientable submanifold can be viewed as a current. The boundary ∂I is defined (when it exists) as the only $m - 1$ dimensional current such that for all $\lambda \in \Lambda^{m-1}$ the integration by parts holds:

$$I(d\lambda) = \partial I(\lambda). \tag{2.5}$$

It is natural to associate to each current I and open set $A \subseteq \Omega$ a mass $\|I\|(A)$ by setting

$$\|I\|(A) = \sup \left\{ \frac{|I(\lambda)|}{\|\lambda\|_c} \text{ s.t. } \lambda \in \Lambda^m, \ \lambda \neq 0, \ \text{supp}(\lambda) \subset A \right\}, \tag{2.6}$$

where $\|\lambda\|_c$ is the comass of λ.

Definition 2.4. We say that an m-dimensional current I on $\Omega \subset \mathbb{R}^{m+n}$ is an *integer current* if $\|I\|(\Omega) < \infty$ and there exists a sequence of C^1 oriented m-dimensional submanifolds $M_i \subset \mathbb{R}^{m+n}$, a sequence of pairwise disjoint closed sets $K_i \subset M_i$ and a sequence of integers k_i such that

$$I(\lambda) = \sum_i k_i \int_{K_i} \lambda \tag{2.7}$$

The current I is said to be an *integral current* if both I and ∂I are integer currents.

It is clear that we can naturally associate a varifold V_I to each integral current I. Note that the mass of the current $\|I\|(A)$ coincides with the mass in the varifold sense $\mu_{V_I}(A)$.

We say that an integral current I is *minimizing* if it minimizes the mass among all other integral currents with the same boundary, in particular

Definition 2.5. Given an integral current I on Ω, we say that I is area-minimizing in Ω if $\|I\|(A) \leq \|J\|(A)$ for all integral currents J such that $\partial I = \partial J$ in Ω and $A \equiv \text{supp}\,(I - J)$ is a compact subset of Ω.

It is worth mentioning that integral currents enjoy an important compactness property with respect to their mass. Indeed, given a sequence I_i of integral currents with $\|I_i\|(\Omega) + \|\partial I_i\|(\Omega) < C < \infty$, there exists a converging subsequence, where the convergence is intended in the weak-$*$ topology with respect to $\Lambda^m(\Omega)$. This compactness property, proved by Federer and Flemming in [FF60], immediately implies the existence of minimizers of the generalized Plateau problem, in the sense that

Theorem 2.6. *Given an $m - 1$ integer rectifiable currents J such that there exists an integral current I with $\partial I = J$ and $\|I\| < \infty$, there exists an integral current I' minimizing the norm $\|I'\|$ of all currents with $\partial I' = J$.*

2.3. **Bounded Mean Curvature and Monotonicity.** For the purposes of this paper, the most important property of an integral varifold with bounded mean curvature is the existence of a monotone quantity at each point. For simplicity let us first consider the case of a stationary integral varifold I^m in $\mathbb{R}^n$. Then for $x \in I$ and $r > 0$ we can consider the density function

$$\theta_r(x) = r^{-m}\mu_I(B_r(x)).\tag{2.8}$$

Then we have that $\theta_r(x)$ is monotone increasing and for $0 < s \leq r$

$$\theta_r(x) - \theta_s(x) = \int_{B_r(x)\backslash\overline{B_s(x)}} d_x^{-m}\langle N_y I, n_x\rangle^2 \, d\mu_I(y)\,,\tag{2.9}$$

where $N_y I$ is the orthogonal complement of the tangent space of I at y, which is defined μ_I-a.e., $d_x(y) = |x-y|$ is the distance function to x, and $n_x(y) = \frac{y-x}{|y-x|}$ is the normal vector field from x. From this it is easy to check that $\theta_r(x)$ is independent of r iff I is 0-symmetric, see [All72, section 5] or [Alm66]. More generally, if $\theta_r(x) = \theta_s(x)$ then I is 0-symmetric on the annulus $A_{s,r}(x)$. In Section 2.4 we will recall a quantitative version of this statement introduced in [CN13b]. Motivated by this, we see that what we are really interested in is the amount the density drops from one scale to the next, and thus we define for $0 < s \leq r$

$$W_{s,r}(x) \equiv \theta_r(x) - \theta_s(x) \geq 0\,.\tag{2.10}$$

Often times we will want to enumerate our choice of scale, so let us define the scales $\mathfrak{r}_\alpha \equiv 2^{-\alpha}$ for $\alpha \geq 0$, and the corresponding mass density drop:

$$W_\alpha(x) \equiv W_{\mathfrak{r}_\alpha,\mathfrak{r}_{\alpha-3}}(x) \equiv \theta_{\mathfrak{r}_{\alpha-3}}(x) - \theta_{\mathfrak{r}_\alpha}(x) \geq 0\,.\tag{2.11}$$

From this one can prove that at every point, every tangent cone is 0-symmetric, which is the starting point for the dimension estimate (1.3) of Federer. In Section 2.4 we discuss quantitative versions of this point, first introduced in [CN13b] and used in this paper as well, and also generalizations which involve higher degrees of symmetry. These points were first used in [CN13b] to prove Minkowski estimates on the quantitative stratification of a stationary varifold. They will also play a role in our arguments, though in a different manner.

In the general case when $M \neq \mathbb{R}^n$ and/or the mean curvature is only bounded, essentially the same statements may be made, however $\theta_r(x)$ is now only *almost* monotone, meaning that $e^{Cr}\theta_r(x)$ is monotone for some constant $C = C(n, K, H)$ which depends only on the geometry of M and the mean curvature bound

H, see for example [DL12, CM11]. In particular, for the almost monotonicity of the normalized volume see [CM11, pag 234], where the authors carry out the computations using the Hessian comparison theorem. Notice also that in the general case the equality in (2.9) still holds up to an error proportional to r

$$\left| \theta_r(x) - \theta_s(x) - \int_{B_r(x)\setminus\overline{B_s(x)}} d_x^{-m} \langle N_y I, n_x \rangle^2 \, d\mu_I(y) \right| \leq Cr. \tag{2.12}$$

2.4. **Quantitative 0-Symmetry and Cone Splitting.** In this subsection we review some of the quantitative symmetry and splitting results of [CN13b], in particular those which will play a role in this paper.

The first result we will discuss acts as an *effective* formulation of the fact that every tangent cone is 0-symmetric. Namely, the quantitative 0-symmetry of [CN13b] says that for each $\epsilon > 0$ and point, that away from a finite number of scales every ball looks $(0, \epsilon)$-symmetric. To be precise, let us consider the radii $r_\alpha \equiv 2^{-\alpha}$, and then the statement is the following:

Theorem 2.7 (Quantitative 0-Symmetry [CN13b]). *Let I^m be an integral varifold on M^n satisfying the curvature bound (1.14), the mean curvature bound (1.15), and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. Then for each $\epsilon > 0$ the following hold:*

(1) *There exists $\delta(n, \Lambda, K, H, \epsilon) > 0$ such that for each $x \in B_1(p)$ and $0 < r \leq r(n, \Lambda, K, H, \epsilon)$, if we have $|\theta_r(x) - \theta_{\delta r}(x)| < \delta$, then $B_r(x)$ is $(0, \epsilon)$-symmetric.*

(2) *For each $x \in B_1(p)$ there exists a finite number of scales $\{\alpha_1, \ldots, \alpha_N\} \subseteq \mathbb{N}$ with $N \leq N(n, \Lambda, K, H, \epsilon)$ such that for $r \notin (r_{\alpha_j}/2, 2r_{\alpha_j})$ we have that $B_r(x)$ is $(0, \epsilon)$-symmetric.*

Remark 2.2. In [CN13b] the result is stated for a stationary varifold, however the verbatim (quick) proof holds just as well for integral varifolds with bounded mean curvature. For the reader's convenience, here we give a sketch of it.

Sketch of the proof. Assume for simplicity that the ambient space is Euclidean, i.e., $K = 0$, and $p = 0$. Consider by contradiction a sequence of varifolds I_i in $B_2(0)$ with uniformly bounded mass and mean curvature, and consider a sequence of balls $B_{r_i}(x_i)$ with $x_i \in B_1(0)$ such that $r_i \leq i^{-1}$ and $\theta(x_i, r_i) - \theta(x_i, i^{-1}r_i) \leq i^{-1}$ but $B_{r_i}(x_i)$ is not $(0, \epsilon)$-symmetric, for some $\epsilon > 0$ fixed.

After rescaling $B_{r_i}(x_i) \to B_1(0)$, we obtain a sequence $\tilde{I}_i$ of varifolds of bounded mass with mean curvature $H_i \to 0$ and such that $\theta(0, 1) - \theta(0, i^{-1}) \leq i^{-1}$. Allard compactness theorem ensures that $\tilde{I}_i$ converges weakly to some $\tilde{I}$ with bounded mass. Given the pinching condition on θ, we obtain that for all $r > 0$:

$$\lim_{i \to \infty} \int_{B_1(0)\setminus B_r(0)} d_x^{-m} \langle N_y \tilde{I}_i, n_x \rangle^2 \, d\mu_{\tilde{I}_i}(y) = 0, \tag{2.13}$$

thus $\tilde{I}$ is a 0-symmetric varifold. Since $\tilde{I}_i$ converges weakly to $\tilde{I}$, we arrived at a contradiction.

A mildly more technical but morally verbatim argument works when the ambient space is not Euclidean but has sectional curvature bounds.

As for the second part, we can assume wlog that the constant δ from the first part is $\delta = 2^{-A}$ for some $A \in \mathbb{N}$. Consider the sequence $\alpha = (0, A, 2A, 3A, ...)$. Since

$$\sum_{k=0}^{\infty} \theta_{2^{-kA}}(x) - \theta_{2^{-(k+1)A}}(x) = \theta_1(x) - \theta_0(x) \tag{2.14}$$

and by monotonicity, we have that at most $N_1(n, \Lambda, K, H, \epsilon)$ values of k fail to satisfy $\theta_{2^{-kA}}(x) - \theta_{2^{-(k+1)A}}(x) < \delta$, and thus at most $N_1(n, \Lambda, K, H, \epsilon)$ balls in the collection $B_{2^{-kA}}(x)$ are not $(0, \epsilon)$-symmetric.

By repeating the same argument for the sequences $\alpha = (i, A + i, 2A + i, 3A + i, ...)$ with $i = 0, \cdots, A - 1$, we obtain the desired estimate with $N(n, \Lambda, K, H, \epsilon) = AN_1$. □

Another technical tool that played an important role in [CN13b] was that of cone splitting. This will be used in this paper when proving the existence of unique tangent planes of symmetry for the singular set, so we will discuss it here. In short, cone splitting is the idea that multiple 0-symmetries add up to give rise to a k-symmetry. To state it precisely let us give a careful definition of the notion of independence of a collection of points:

Definition 2.8. We say a collection of points $\{x_1, \ldots, x_\ell\} \subset \mathbb{R}^n$ is independent if they are linearly independent. We say the collection is τ-independent at x if $d(x_{k+1}, \text{span}\{x_1, \ldots, x_k\}) > \tau$ for each k. If $\{x_1, \ldots, x_\ell\} \subset M$ then we say the collection is τ-independent with respect to x if $d(x, x_j) < \text{inj}(x)$ and the collection is τ-independent when written in exponential coordinates at x.

Now we are in a position to state the effective cone splitting of [CN13b]:

Theorem 2.9 (Cone Splitting [CN13b]). *Let I^m be an integral varifold on M^n satisfying the curvature bound* (1.14)*, the mean curvature bound* (1.15)*, the mass bound $\mu_I(B_3(p)) \leq \Lambda$, and let $\epsilon, \tau > 0$ be fixed. Then there exists $\delta(n, \Lambda, \epsilon, \tau) > 0$ such that if $K + H < \delta$ and $x_1, \ldots, x_k \in B_1(p)$ are such that*

(1) *$B_2(x_j)$ are $(0, \delta)$-symmetric, and $B_2(p)$ is $(0, \delta)$-symmetric*
(2) *$\{x_1, \ldots, x_k\}$ are τ-independent at p,*

then $B_1(p)$ is (k, ϵ)-symmetric.

Remark 2.3. The assumption $K + H < \delta$ is of little consequence, since this just means focusing the estimates on balls of sufficiently small radius after rescaling.

Proof. The proof can be carried out with a simple compactness argument similar to the one used in the proof of theorem 2.7. □

We end with the following, which one can view as a quantitative form of dimension reduction.

Theorem 2.10 (Quantitative Dimension Reduction). *Let I^m be an integral varifold on M^n satisfying the curvature bound* (1.14)*, the mean curvature bound* (1.15)*, and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. Then for each $\epsilon > 0$ there exists $\delta(n, \Lambda, \epsilon), r(n, \Lambda, \epsilon) > 0$ such that if $K + H < \delta$ and $B_2(p)$ is (k, δ)-symmetric with respect to some k-plane V^k, then for each $x \in B_1(p) \setminus B_\epsilon(V^k)$ we have that $B_r(x)$ is $(k + 1, \epsilon)$-symmetric.*

The proof of this theorem is standard and it follows from theorem 2.7 and a contradiction argument.

2.5. ϵ-regularity for Minimizing Hypersurfaces.

In this subsection we quickly review the ϵ-regularity theorem of [CN13b], which itself follows quickly from the difficult work of [Sim68],[Fed69] after an easy contradiction argument. This will be our primary technical tool in upgrading the structural results on varifolds with bounded mean curvature to the regularity results for area minimizing hypersurfaces. The main theorem of this subsection is the following:

Theorem 2.11 (ϵ-Regularity [Sim68],[Fed69],[CN13b]). *Let $I^{n-1} \subseteq B_2$ be a minimizing hypersurface with $\partial I \cap B_2 (0) = 0$ satisfying the bounds (1.14), and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. Then there exists $\epsilon(n, \Lambda) > 0$ such that if $K < \epsilon$ and $B_2(p)$ is $(n - 7, \epsilon)$-symmetric, then*

$$r_I(p) \geq 1 \,.$$

Remark 2.4. The assumption $K < \epsilon$ is of little consequence, since this just means focusing the estimates on balls of sufficiently small radius after rescaling. Note that since I is a minimizing hypersurface, we know that $H = 0$ in this case.

2.6. **Hausdorff, Minkowski, and packing Content.** In this subsection we give a brief review of the notions of Hausdorff, Minkowski, and packing content. We will also use this to recall the definition of Hausdorff measure. The results of this subsection are completely standard, but this gives us an opportunity to introduce some notation for the paper. For a more detailed reference, we refer the reader to [Mat95, Fed69]. Let us begin with the notions of content:

Definition 2.12. Given a set $S \subseteq \mathbb{R}^n$ and $r > 0$ we define the following:

(1) The k-dimensional Hausdorff r-content of S is given by

$$\lambda_r^k(S) \equiv \inf \left\{ \sum \omega_k r_i^k : S \subseteq \bigcup E_i \text{ and } \operatorname{diam}(E_i) \leq 2r_i \leq 2r \right\}. \tag{2.15}$$

(2) The k-dimensional Minkowski r-content of S is given by

$$m_r^k(S) \equiv (2r)^{k-n} \operatorname{Vol}(B_r(S)) \,. \tag{2.16}$$

(3) The k-dimensional packing r-content of S is given by

$$p_r^k(S) \equiv \sup \left\{ \sum \omega_k r_i^k : x_i \in S \text{ and } \{B_{r_i}(x_i)\} \text{ are disjoint and } r_i \leq r \right\}. \tag{2.17}$$

These definitions make sense for any $k \in [0, \infty)$, though in this paper we will be particularly interested in integer valued k. Notice that if S is a compact set then $\lambda_r^k(S), m_r^k(S) < \infty$ for any $r > 0$, and that we always have the relations

$$\lambda_r^k(S) \lesssim m_r^k(S) \lesssim p_r^k(S) \,. \tag{2.18}$$

In particular, bounding the Hausdorff content is less powerful than bounding the Minkowski content, which is itself less powerful than bounding the packing content.

Primarily in this paper we will be mostly interested in content estimates, because these are the most effective estimates. However, since it is classical, let us go ahead and use the Hausdorff content to define a measure. To accomplish this, let us more generally observe that if $r \leq r'$ then $\lambda_r^k(S) \geq \lambda_{r'}^k(S)$. In particular, we can define the limit

$$\lambda_0^k(S) \equiv \lim_{r \to 0} \lambda_r^k(S) = \sup_{r > 0} \lambda_r^k(S) \,.$$

It turns out that λ_0^k is a genuine measure.

Definition 2.13. Given a set $S \subseteq \mathbb{R}^n$ we define its k-dimensional Hausdorff measure by $\lambda^k(S) \equiv \lambda_0^k(S)$.

Similar constructions can be carried out for the Minkowski and packing content. In particular, we can define

$$\overline{m}_0^k(S) \equiv \limsup_{r \to 0} m_r^k(S), \quad \underline{m}_0^k(S) \equiv \liminf_{r \to 0} m_r^k(S), \tag{2.19}$$

$$p_0^k(S) = \lim_{r \to 0} p_r^k(S) = \inf_{r > 0} p_r^k(S). \tag{2.20}$$

Definition 2.14. Given a set $S \subseteq \mathbb{R}^n$ we define its Hausdorff and Minkowski dimension (or box-dimension) by

$$\dim_H S \equiv \inf \left\{ k \geq 0 : \lambda_0^k(S) = 0 \right\},$$
$$\dim_M S \equiv \inf \left\{ k \geq 0 : \overline{m}_0^k(S) = 0 \right\}. \tag{2.21}$$

Remark 2.5. Note that we could define an upper and lower Minkowski dimension by

$$\overline{\dim}_M S \equiv \inf \left\{ k \geq 0 : \overline{m}_0^k(S) = 0 \right\}, \quad \underline{\dim}_M S \equiv \inf \left\{ k \geq 0 : \underline{m}_0^k(S) = 0 \right\}. \tag{2.22}$$

In general, $\underline{\dim}_M S \leq \overline{\dim}_M S$, where the inequality may be strict. However, for the purposes of this paper we will only be interested in the *upper* Minkowski dimension.

As an easy example consider the rationals $\mathbb{Q}^n \subseteq \mathbb{R}^n$. Then it is a worthwhile exercise to check that $\dim_H \mathbb{Q}^n = 0$, while $\dim_M \mathbb{Q}^n = n$.

A very important notion related to measures is the *density* at a point. Although this is standard, for completeness we briefly recall the definition of Hausdorff density, and refer the reader to [Mat95, chapter 6] for more on this subject.

Definition 2.15. Given a set $S \subset \mathbb{R}^n$ which is λ^k-measurable, and $x \in \mathbb{R}^n$, we define the k-dimensional upper and lower density of S at x by

$$\theta^{\star k}(S, x) = \limsup_{r \to 0} \frac{\lambda^k(S \cap B_r(x))}{\omega_k r^k}, \quad \theta_\star^k(S, x) = \liminf_{r \to 0} \frac{\lambda^k(S \cap B_r(x))}{\omega_k r^k}. \tag{2.23}$$

In the following, we will use the fact that for almost any point in a set with finite λ^k-measure, the density is bounded from above and below.

Proposition 2.16 ([Mat95]). *Let* $S \subset \mathbb{R}^n$ *be a set with* $\lambda^k(S) < \infty$. *Then for* k-*a.e.* $x \in S$:

$$2^{-k} \leq \theta^{\star k}(S, x) \leq 1, \tag{2.24}$$

while for k-*a.e.* $x \in \mathbb{R}^n \setminus S$

$$\theta^{\star k}(S, x) = 0. \tag{2.25}$$

2.7. **The Classical Reifenberg theorem.** In this Section we recall the classical Reifenberg theorem, as well as some more recent generalizations. The Reifenberg theorem gives criteria on a closed subset $S \subseteq B_2 \subseteq \mathbb{R}^n$ which determine when $S \cap B_1$ is bi-Hölder to a ball $B_1(0^k)$ in a smaller dimensional Euclidean space. The criteria itself is based on the existence of *good* best approximating subspaces at each scale. We start by recalling the Hausdorff distance.

Definition 2.17. Given two sets $A, B \subseteq \mathbb{R}^n$, we define the Hausdorff distance between these two by

$$d_H(A, B) = \inf \{r \geq 0 \ \ s.t. \ \ A \subset B_r(B) \ \ \text{and} \ \ B \subset B_r(A)\} \ . \tag{2.26}$$

Recall that d_H is a distance on closed sets, meaning that $d_H(A, B) = 0$ implies $\overline{A} = \overline{B}$.

The classical Reifenberg theorem says the following:

Theorem 2.18 (Reifenberg theorem [Rei60, Sim])**.** *For each $0 < \alpha < 1$ and $\epsilon > 0$ there exists $\delta(n, \alpha, \epsilon) > 0$ such that the following holds. Assume $0^n \in S \subseteq B_2 \subseteq \mathbb{R}^n$ is a closed subset, and that for each $x \in S \cap B_1(0)$ and $r \in (0, 1]$ we have*

$$\inf_{L^k} d_H(S \cap B_r(x), L^k \cap B_r(x)) < \delta \, r \, , \tag{2.27}$$

where the $\inf$ is taken over all k-dimensional affine subspaces $L^k \subseteq \mathbb{R}^n$. Then there exists $\phi : B_1(0^k) \to S$ which is a C^α bi-Hölder homeomorphism onto its image with $[\phi]_{C^\alpha}, [\phi^{-1}]_{C^\alpha} < 1 + \epsilon$ and $S \cap B_1 \subseteq \phi(B_1(0^k))$.

Remark 2.6. In fact, one can prove a little more. In particular, under the hypothesis of the previous theorem, there exists a closed subset $S' \subset \mathbb{R}^n$ such that $S' \cap B_1(0) = S \cap B_1(0)$ and which is homeomorphic to a k-dimensional subspace $0^n \in T_0 \subseteq \mathbb{R}^n$ via the C^α bi-Hölder homeomorphism $\phi : T_0 \to S'$. Moreover, $|\phi(x) - x| \leq C(n)\delta$ for all $x \in T_0$ and $\phi(x) = x$ for all $x \in T_0 \setminus B_2(0)$.

One can paraphrase the above to say that if S can be well approximated on every ball by a subspace in the L^∞-sense, then S must be bi-Hölder to a ball in Euclidean space.

Let us also mention that there are several more recent generalizations of the classic Reifenberg theorem. In [Tor95], the author proves a strengthened version of (2.27) that allows one to improve bi-Hölder to bi-Lipschitz. Unfortunately, for the applications of this paper the hypotheses of [Tor95] are much too restrictive. We will require a weaker condition than in [Tor95], which is more integral in nature, see theorem 3.2. In exchange, we will improve the bi-Hölder of the classical Reifenberg to $W^{1,p}$.

We will also need a version of the classical Reifenberg which only assumes that the subset S is contained near a subspace, not conversely that the subspace is also contained near S. In exchange, we will only conclude the set is rectifiable. A result in this direction was first proved in [DT12], but again the hypotheses are too restrictive for the applications of this paper, and additionally there is a topological assumption necessary for the results of [DT12], which is not reasonable in the context in this paper. We will see how to appropriately drop this assumption in theorem 3.4.

3. The $W^{1,p}$-Reifenberg theorem

In this Section we recall some Reifenberg-type theorems first introduced in [NVa]. In [NVa] we focused our attention on proving the rectifiable-Reifenberg of theorem 3.4, and in this section we will focus our attention on proving the discrete Reifenberg of theorem 3.3. The proofs of the two results are very similar.

Some interesting generalizations of Reifenberg's theorem regarding rectifiability are present in literature. Generalizations similar to the ones studied here are obtained in [Tor95] and [DT12]. However, the requirements of the theorems proved in these papers are too stringent to be applicable in our situation. Recently, and using techniques independent from this work, [AT15],[Tol15] proved necessary and sufficient conditions for the rectifiability of a set or of a measure that are closely related to the results of this section. However, these results don't provide any *apriori* control over the volume of the sets in question, which is a key estimate of this section and it is essential in the application to the singular strata of currents.

3.1. **Statement of Main $W^{1,p}$-Reifenberg and rectifiable-Reifenberg Results.** Before turning to the statements of the theorems, let us introduce some definitions in order to keep the statements as clean and intuitive as possible.

Definition 3.1. Let μ be a measure on B_2 with $r > 0$ and $k \in \mathbb{N}$. Then we define the k-dimensional displacement by

$$D_\mu^k(x, r) \equiv \inf_{L^k} r^{-(k+2)} \int_{B_r(x)} d^2(y, L^k) \, d\mu(y) \,, \tag{3.1}$$

if $\mu(B_r(x)) \geq \epsilon_n r^k \equiv (1000n)^{-7n^2} r^k$, and $D_\mu^k(x, r) \equiv 0$ otherwise, where the inf's are taken over all k-dimensional affine subspaces $L^k \subseteq \mathbb{R}^n$. If $S \subseteq B_2$, then we can define its k-displacement $D_S^k(x, r)$ by associating to S the k-dimensional Hausdorff measure λ_S^k restricted to S and setting

$$D_S^k(x, r) \equiv D_{\lambda^k|_S}^k(x, r) \,. \tag{3.2}$$

Sometimes, we will omit the index k and the subscript μ or S when there can be no risk of confusion from the context. In particular, we will often write $D(x, r)$ for $D_\mu^k(x, r)$.

Remark 3.1. One can replace ϵ_n by any smaller lower bound, and the proofs and statements will all continue to hold.

Remark 3.2. Notice that the definitions are scale invariant. In particular, if we rescale $B_r \to B_1$ and let $\tilde{S}$ be the induced set, then $D_S^k(x, r) \to D_{\tilde{S}}^k(x, 1)$.

Remark 3.3. Notice the monotonicity given by the following: If $\mu' \leq \mu$, then $D_{\mu'}^k(x, r) \leq D_\mu^k(x, r)$.

Remark 3.4. It is immediate to see from the definition that, up to dimensional constants, $D_\mu^k(x, r)$ is controlled on both sides by $D_\mu^k(x, r/2)$ and $D_\mu^k(x, 2r)$. In particular, if $\mu(B_r(x)) \geq \gamma_k r^k = \omega_k 40^{-k} r^k >> \epsilon_n r^k$, then for all $y \in B_r(x)$, $D_\mu^k(x, r) \leq 2^{k+2} D_\mu^k(y, 2r)$. As a corollary we have the estimate

$$D_\mu^k(x, r) \leq 2^{k+2} \fint_{B_r(x)} D_\mu^k(y, 2r) d\mu(y) \,. \tag{3.3}$$

Before introducing the results which are really needed for the paper, it is worth mentioning the $W^{1,p}$-Reifenberg theorem obtained in [NVa]. This is a natural generalization of the Reifenberg and gives intuition and motivation for the rest of the statements, which are essentially more complicated versions of it. We do not prove the theorem in this paper, instead we refer the reader to [NVa, theorem 3.2].

Theorem 3.2 ($W^{1,p}$-Reifenberg). [NVa, theorem 3.2] *For each $\epsilon > 0$ and $p \in [1, \infty)$ there exists $\delta(n, \epsilon, p) > 0$ such that the following holds. Let $S \subseteq B_4 \subseteq \mathbb{R}^n$ be a closed subset with $0^n \in S$, and assume for each $x \in S \cap B_1$ and $B_r(x) \subseteq B_4$ that*

$$\inf_{L^k} d_H(S \cap B_r(x), L^k \cap B_r(x)) < \delta r, \tag{3.4}$$

$$\int_{S \cap B_r(x)} \left(\int_0^r D_S^k(y, s) \, \frac{ds}{s} \right) d\lambda^k(y) < \delta^2 r^k. \tag{3.5}$$

Then the following hold:

(1) *there exists a mapping $\phi : \mathbb{R}^k \to \mathbb{R}^n$ which is a $1 + \epsilon$ bi-$W^{1,p}$ map onto its image and such that $S \cap B_1(0^n) = \phi(B_1(0^k)) \cap B_1(0^n)$.*

(2) *$S \cap B_1(0^n)$ is countably k-rectifiable.*

(3) *For each ball $B_r(x) \subseteq B_1$ with $x \in S$ we have*

$$(1 - \epsilon)\omega_k r^k \leq \lambda^k(S \cap B_r(x)) \leq (1 + \epsilon)\omega_k r^k. \tag{3.6}$$

Remark 3.5. Results (2) and (3) both follow from (1). We get (3) by applying the result of (1) to all smaller balls $B_r(x) \subseteq B_1$, since the assumptions of the theorem hold on these balls as well.

Remark 3.6. Note that, for $p > k$, a bi-$W^{1,p}$ map is a bi-C^α map, in particular we see that $\phi(B_1(0^k))$ is homeomorphic to the ball $B_1(0^k)$.

Remark 3.7. As it is easily seen, the requirement that S is closed is essential for this theorem, and in particular for the lower bound on the Hausdorff measure. As an example, consider any set $S \subseteq \mathbb{R}^k$ which is dense but has zero Hausdorff measure. In the following theorems, we will not be concerned with lower bounds on the measure, and we will be able to drop the closed assumption.

We are going to state another generalization of Reifenberg's theorem, more discrete in nature, which will be particularly important in the proof of the main theorems of this paper:

Theorem 3.3 (Discrete Reifenberg). [NVa, theorem 3.4] *There exists $\delta(n) > 0$ and $D(n)$ such that the following holds. Let $\{B_{r_s}(x_s)\}_{s \in S} \subseteq B_2$ be a collection of disjoint balls with $x_s \in B_1(0)$, and let $\mu \equiv \sum_{s \in S} \omega_k r_s^k \delta_{x_s}$ be the associated measure. Assume that for each $B_r(x) \subseteq B_4$ with $\mu(B_r(x)) \geq \gamma_k r^k = \omega_k(r/40)^k$ we have*

$$\int_{B_r(x)} \left(\int_0^r D_\mu^k(y, t) \, \frac{dt}{t} \right) d\mu(y) < \delta^2 r^k. \tag{3.7}$$

Then we have the estimate

$$\sum_{s \in S} r_s^k < D(n). \tag{3.8}$$

Remark 3.8. Instead of (3.7) we may assume the estimate

$$\sum_{\mathfrak{r}_\alpha \leq r/2} \int_{B_r(x)} D_\mu^k(y, \mathfrak{r}_\alpha) \, d\mu(y) < \delta^2 r^k \,. \tag{3.9}$$

In the applications, this will be the more convenient phrasing.

In order to prove rectifiability of the strata, we will also need the following version of Reifenberg's theorem. The proof of this theorem relies on the same ideas as the discrete-Reifenberg, for this reason we do not report it here and we refer the interested reader to [NVa, theorem 3.3].

Theorem 3.4 (Rectifiable-Reifenberg). [NVa, theorem 3.3] *For every $\epsilon > 0$, there exists $\delta(n, \epsilon) > 0$ such that the following holds. Let $S \subseteq B_4 \subseteq \mathbb{R}^n$ be a λ^k-measurable subset, and assume for each $B_r(x) \subseteq B_4$ with $\lambda^k(S \cap B_r(x)) \geq \gamma_k r^k$ that*

$$\int_{S \cap B_r(x)} \left(\int_0^r D_S^k(y, s) \, \frac{ds}{s} \right) d\lambda^k(y) < \delta^2 r^k \,. \tag{3.10}$$

Then the following holds:

(1) *For each ball $B_r(x) \subseteq B_1$ with $x \in S$ we have*

$$\lambda^k(S \cap B_r(x)) \leq (1 + \epsilon)\omega_k r^k \,. \tag{3.11}$$

(2) *$S \cap B_1(0^n)$ is countably k-rectifiable.*

Remark 3.9. Notice that for the statement of the theorem we do not need control over balls which already have small measure. This will be quite convenient for the applications.

Remark 3.10. Instead of (3.10) we may assume the essentially equivalent estimate

$$\sum_{\mathfrak{r}_\alpha \leq r/2} \int_{S \cap B_r(x)} D_S^k(y, \mathfrak{r}_\alpha) \, d\lambda^k(y) < \delta^2 r^k \,. \tag{3.12}$$

In the applications, this will be the more convenient phrasing.

3.2. **Explanatory example.** In order to understand better the idea behind the improvement of the Reifenberg theorem, we use the famous snow-flake as a test case.
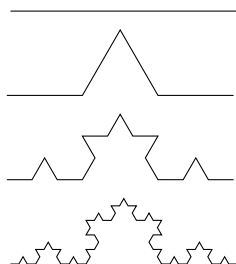
The construction of a snowflake of parameter $\eta > 0$ is well known (see for example [Mat95, section 4.13]). Take the unit segment $[0, 1] \times \{0\} \subseteq \mathbb{R}^2$, and replace the middle part $[1/3, 2/3] \times \{0\}$ with the top part of the isosceles triangle with base $[1/3, 2/3] \times \{0\}$ and of height $\eta \cdot \text{lenght}([1/3, 2/3] \times \{0\})$. In other words, you are replacing the segment $[1/3, 2/3] \times \{0\}$ with the two segments joining $(1/3, 0)$ to $(1/2, \eta/3)$, and $(1/2, \eta/3)$ to $(2/3, 0)$. Then repeat this construction inductively on each of the 4 straight segments in the new set. Here on the left hand side you can see the very classical picture

of the first three steps in the construction of the standard snowflake, with $\eta = \sqrt{3}/2$.

It is clear that the length of the curve at step i is equal to the length at step $i - 1$ times $2/3 + \sqrt{1 + \eta^2}/3$, so the length of the snowflake will be infinity for any $\eta > 0$. This is a simple application of the Pythagorean theorem, and the extra square power on η comes from the fact that at each step we are adding some length η to the curve, but in a direction perpendicular to it.

However, if we replace the fixed parameter η with a variable parameter η_i, we see immediately that the length of the limit curve will be finite if and only if $\sum \eta_i^2 < \infty$.

This suggests that the finiteness of the Hausdorff measure of the set S is related to the summability properties of $D_S^k(x, \mathfrak{r}_\alpha)$ over scales.

Indeed, if we introduce the following L^∞ analogue of the distortion D:

$$\bar{D}_S^k(x, r) \equiv r^{-2} \inf_{L^k} \left[d_H \left(S \cap B_r(x), L^k \cap B_r(x) \right) \right]^2 \tag{3.13}$$

and require that the sum over scales of $D(x, r)$ is small in the sense that

$$\sum_{\alpha=0}^{\infty} \sup_{x \in S \cap B_1(0)} D_S^k(x, \mathfrak{r}_\alpha) < \delta^2, \tag{3.14}$$

then we obtain that S is a bi-Lipschitz image of a k-dimensional disk. This result was proved by Toro in [Tor95].

However, one can also lower these requirements and replace the L^∞ norms with more flexible L^2 norms and integrals, and still obtain finiteness of the k-dimensional Hausdorff measure and rectifiability under the less restrictive assumptions of theorem 3.4. Moreover, by a simple covering argument, it is not necessary to ask control over $d_H \left(S \cap B_r(x), L^k \cap B_r(x) \right)$, but just on $d(x, L^k)$ for $x \in S$. In other words, it is only important to have control over how close S is to some k-dimensional subspace, and not vice-versa. Thus it is not a problem to have "holes" in the set S.

As mentioned above, similar results to the ones stated in this section were proved in [DT12], but still the theorems proved there have slightly stronger assumptions than the ones used here, and on the other hand obtain stronger topological results than the ones needed for this paper.

3.3. Examples.
In this subsection we present a few examples which motivate the sharpness of our results.

3.3.1. *The Simons Cone and Sharp Estimates.*
In order to study the sharpness of the estimates of theorem 1.8 we study some examples. One may use either the Simons cone or the Lawson cone for this analysis. The Simons cone $C \subseteq \mathbb{R}^8$ is a cone over the surface

$$S^3\left(\frac{1}{\sqrt{2}}\right) \times S^3\left(\frac{1}{\sqrt{2}}\right) \subseteq S^7. \tag{3.15}$$

It has been shown [BDGG69] that C is an area minimizing cone. It is easy to check that for $x \in C$ we have that $|A|(x) = \sqrt{6}|x|^{-1}$, where A is the second fundamental form (see [Sim83, remark B.3]). In particular, we get that $|A| \in L^7_{weak}$, but $|A| \notin L^7_{loc}$, showing that the estimates of theorem 1.8 are sharp.

3.3.2. *Rectifiable-Reifenberg Example I.* Let us begin with an easy example, which shows that the rectifiable conclusions of theorem 3.4 is sharp. That is, one cannot hope for better structural results under the hypothesis. Indeed, consider any k-dimensional subspace $V^k \subseteq \mathbb{R}^n$, and let $S \subseteq V^k \cap B_2(0^n)$ be an arbitrary measurable subset. Then clearly $D(x, r) \equiv 0$ for each x and $r > 0$, and thus the hypotheses of theorem 3.4 are satisfied, however S clearly need not be better than rectifiable. In the next example we shall see that S need not even come from a single rectifiable chart, as it does in this example.

3.3.3. *Rectifiable-Reifenberg Example II.* With respect to the conclusions of theorem 3.4 there are two natural questions regarding how sharp they are. First, is it possible to obtain more structure from the set S than rectifiable? In particular, in theorem 3.2 there are topological conclusions about the set, is it possible to make such conclusions in the context of theorem 3.4? In the last example we saw this is not the case. Then a second question is to ask whether we can at least find a single rectifiable chart which covers the whole set S. This example taken from [DT12, counterexample 12.4] shows that the answer to this question is negative as well.

To build our examples let us first consider a unit circle $S^1 \subseteq \mathbb{R}^3$. Let $M^2 \supset S^1$ be a smooth Möbius strip around this circle, and let $S_\epsilon \subseteq M^2 \cap B_\epsilon(S^1) \equiv M_\epsilon^2$ be an arbitrary λ^2-measurable subset of the Möbius strip, contained in a small neighborhood of the S^1. In particular, $Area(S_\epsilon) \leq C\epsilon \to 0$ as $\epsilon \to 0$. It is not hard, though potentially a little tedious, to check that assumptions of theorem 3.4 hold for $\delta \to 0$ as $\epsilon \to 0$.

However, we have learned two points from these example. First, since S_ϵ was an arbitrary measurable subset of a two dimensional manifold, we have that it is 2-rectifiable, however that is the most which may be said of S_ϵ. That is, structurally speaking we cannot hope to say better than 2-rectifiable about the set S_ϵ. More than that, since S_ϵ is a subset of the Möbius strip, we see that even though S_ϵ is rectifiable, we cannot even cover S_ϵ by a single chart from $B_1(0^2)$, as a Möbius strip is not orientable. See [DT12] for more on this.

4. Technical Constructions toward New Reifenberg Results

In this section, we prove some technical lemmas needed for dealing with the relation between best L^2 subspaces. These elementary results will be used in many of the estimates of subsequent sections.

4.1. **Hausdorff distance and subspaces.** We start by recalling some standard facts about affine subspaces in $\mathbb{R}^n$ and Hausdorff distance.

Definition 4.1. Given two linear subspaces $L, V \subseteq \mathbb{R}^n$, we define the Grassmannian distance between these two as

$$d_G(L, V) = d_H(L \cap B_1(0), V \cap B_1(0)) = d_H\left(L \cap \overline{B_1(0)}, V \cap \overline{B_1(0)}\right). \tag{4.1}$$

Note that if $\dim(L) \neq \dim(V)$, then $d_G(L, V) = 1$.

For general subsets in $\mathbb{R}^n$, it is evident that $A \subseteq B_\delta(B)$ does not imply $B \subseteq B_{c\delta}(A)$. However, if A and B are affine spaces with the same dimension, then it is not difficult to see that this property holds. More precisely:

Lemma 4.2. *Let V, W be two k-dimensional affine subspaces in $\mathbb{R}^n$, and suppose that $V \cap B_{1/2}(0) \neq \emptyset$. There exists a constant $c(k,n)$ such that if $V \cap B_1(0) \subseteq B_\delta(W \cap B_1(0))$, then $W \cap B_1(0) \subseteq B_{c\delta}(V \cap B_1(0))$. Thus in particular $d_H(V \cap B_1(0), W \cap B_1(0)) \leq c\delta$.*

Proof. The proof relies on the fact that V and W have the same dimension. Let $x_0 \in V$ be the point of minimal distance from the origin. By assumption, we have that $\|x_0\| \leq 1/2$. Let $x_1, \cdots, x_k \in V \cap \overline{B_1(0)}$ be a sequence of points such that

$$\|x_i - x_0\| = 1/2 \quad \text{and for } i \neq j, \quad \langle x_i - x_0, x_j - x_0 \rangle = 0. \tag{4.2}$$

In other words, $\{x_i - x_0\}_{i=1}^k$ is an affine base for V. Let $\{y_i\}_{i=0}^k \subseteq W \cap \overline{B_1(0)}$ be such that $d(x_i, y_i) \leq \delta$. Then

$$\|y_i - y_0\| \geq 1/2 - 2\delta \quad \text{and for } i \neq j, \quad \left|\langle y_i - y_0, y_j - y_0 \rangle\right| \leq 4\delta + 4\delta^2. \tag{4.3}$$

This implies that for $\delta \leq \delta_0(n)$, $\{y_i - y_0\}_{i=1}^k$ is an affine base for W and for all $y \in W$

$$y = y_0 + \sum_{i=1}^k \alpha_i(y_i - y_0), \quad |\alpha_i| \leq 10 \|y - y_0\|. \tag{4.4}$$

Now let $y \in W \cap \overline{B_1(0)}$ be the point of maximum distance from V, and let π be the projection onto V and $\pi^\perp$ the projection onto $V^\perp$, which is the linear subspace orthogonal to V. Then

$$d(y, V) = d(y, \pi(y)) = \left\|\pi^\perp(y - x_0)\right\| \leq \left\|\pi^\perp(y_0 - x_0)\right\| + \sum_{i=1}^k |\alpha_i| \left\|\pi^\perp(y_i - y_0)\right\| \leq c'(n,k)\delta. \tag{4.5}$$

Since $y \in \overline{B_1(0)}$ and $\|x_0\| \leq 1/2$, by a simple geometric argument $\pi(y) \in V \cap B_{1+c'\delta}(0)$, and thus $d(y, V \cap B_1(0)) \leq 2c'\delta \equiv c\delta$. This proves the claim. $\square$

Next we will see that the Grassmannian distance between two subspaces is enough to control the projections with respect to these planes. In order to do so, we recall a standard estimate.

Lemma 4.3. *Let V, W be linear subspaces of a Hilbert space. Then $d_G(V, W) = d_G(V^\perp, W^\perp)$.*

Proof. We will prove that $d_G(V^\perp, W^\perp) \leq d_G(V, W)$. By symmetry, this is sufficient.

Take $x \in V^\perp$ such that $\|x\| = 1$, and consider that $d(x, W^\perp) = \|\pi_W(x)\|$. Let $z = \pi_W(x)$ and $y = \pi_V(z)$. We want to show that if $d_G(V, W) \leq \epsilon < 1$, then $\|z\| \leq \epsilon$. We can limit our study to the space spanned by x, y, z, and assume wlog that $x = (1, 0, 0)$, $y = (0, b, 0)$ and $z = (a, b, c)$. By orthogonality between z and $z - x$, we have

$$a^2 + b^2 + c^2 + (1 - a)^2 + b^2 + c^2 = 1 \implies a = a^2 + b^2 + c^2, \tag{4.6}$$

and since $z \in W$, we also have $\|z - y\| \leq \epsilon \|z\|$, which implies

$$a^2 + c^2 \leq \epsilon^2 \left(a^2 + b^2 + c^2\right) \implies a^2 + c^2 \leq \frac{\epsilon^2}{1 - \epsilon^2} b^2. \tag{4.7}$$

Since the function $f(x) = x^2/(1 - x^2)$ is monotone increasing for $0 \le x < 1$, we can define $0 \le \alpha < 1$ in such a way that

$$a^2 + c^2 = \frac{\alpha^2}{1 - \alpha^2} b^2 , \quad a = a^2 + b^2 + c^2 = \frac{1}{1 - \alpha^2} b^2 . \tag{4.8}$$

Note that necessarily we will have $\alpha \le \epsilon$. Now we have

$$\frac{1}{(1 - \alpha^2)^2} b^4 = a^2 \le \frac{\alpha^2}{1 - \alpha^2} b^2 \quad \implies \quad b^2 \le \alpha^2 \left(1 - \alpha^2\right) \quad \implies \quad \|z\|^2 = a^2 + b^2 + c^2 \le \alpha^2 \le \epsilon^2 . \tag{4.9}$$

This proves that $V^\perp \cap B_1(0) \subset B_\epsilon(W^\perp)$. In a similar way, one proves the opposite direction. $\square$

As a corollary, we prove that the Grassmannian distance $d_G(V, W)$ is equivalent to the distance given by $\|\pi_V - \pi_W\|$.

Lemma 4.4. *Let V, W be linear subspaces of $\mathbb{R}^n$. Then for every $x \in \mathbb{R}^n$,*

$$\|\pi_V(x) - \pi_W(x)\| \le 2 d_G(V, W) \|x\| . \tag{4.10}$$

In particular, if $x \in W^\perp$, then $\|\pi_V(x)\| \le 2 d_G(V, W) \|x\|$.

Conversely, we have

$$d_G(V, W) \le \sup_{x \in \mathbb{R}^n \setminus \{0\}} \left\{ \frac{\|\pi_V(x) - \pi_W(x)\|}{\|x\|} \right\} . \tag{4.11}$$

Proof. The proof is just a corollary of the previous lemma. Assume wlog that $\|x\| = 1$, and let $x = y + z$ where $y = \pi_V(x)$ and $z = \pi_{V^\perp}(x)$. Then

$$\|\pi_V(x) - \pi_W(x)\| = \|y - \pi_W(y) - \pi_W(z)\| \le \|y - \pi_W(y)\| + \|z - \pi_{W^\perp}(z)\| = d(y, W) + d(z, W^\perp) . \tag{4.12}$$

Since $\|y\|^2 + \|z\|^2 = \|x\|^2 = 1$, by the previous lemma we get the first estimate.

The reverse estimate is an immediate consequence of the definition. $\square$

4.2. Distance between L^2 best planes.

Here we study the distance between best approximating subspaces for our measure μ on different balls. Let us begin by fixing our notation for this subsection, and pointing out the interdependencies of the constants chosen here. Throughout this subsection, our choice of scale $\rho = \rho(n, M) > 0$ is a constant which will eventually be fixed according to lemma 4.7. For applications to future sections, it is sufficient to know that we can take $\rho(n, M) = 10^{-10}(100n)^{-n} M^{-1}$. We also point out that in Section 5, we will fix $M = 40^n$, and so ρ will be a constant depending only on n. In particular, we can use the very coarse estimate

$$\rho = 10^{-10}(100n)^{-3n} . \tag{4.13}$$

We will also introduce a threshold value $\gamma_k = \omega_k 40^{-k}$. The dimensional constant γ_k is chosen simply to be much smaller than any covering errors which will appear.

We will consider a positive Radon measure μ supported on $S \subseteq B_1(0)$, and use $D(x, r) \equiv D_\mu^k(x, r)$ to bound the distances between best L^2 planes at different points and scales. By definition let us denote by $V(x, r)$ a best k-dimensional plane on $B_r(x)$, i.e., a k-dimensional affine subspace minimizing $\int_{B_r(x)} d(x, V)^2 d\mu$. Note that, in general, this subspace may not be unique. We want to prove that, under reasonable hypothesis, the

distance between $V(x, r)$ and $V(y, r')$ is small if $d(x, y) \sim r$ and $r' \sim r$.

In order to achieve this, we will need to understand some minimal properties of μ. First, we need to understand how concentrated μ is on any given ball. For this reason, for some $\rho > 0$ and all $x \in B_1(0)$ we will want to consider the upper mass bound

$$\mu(B_\rho(x)) \leq M\rho^k \quad \forall x \in B_1(0). \tag{4.14}$$

However, an upper bound on the measure is not enough to guarantee best L^2-planes are close, as the following example shows:

Example 4.1. Let V, V' be k-dimensional subspaces, $0 \in V \cap V'$, and set $S = (V \cap B_1(0) \setminus B_{1/10}(0)) \cup S'$, where $S' \subseteq V' \cap B_{1/10}(0)$ and $\mu = \lambda^k|_S$. Then evidently $D(0, 1) \leq \lambda^k(S')$ and $D(0, 1/10) = 0$, independently of V and V'. However, $V(0, 1)$ will be close to V, while $V(0, 1/10) = V'$. Thus, in general, we cannot expect $V(0, 1)$ and $V(0, 1/10)$ to be close if $\mu(B_{1/10}(0))$ is too small.

Thus, in order to prove that the best planes are close, we need to have some definite amount of measure on the set, in such a way that S "effectively spans" a k-dimensional subspace, where by effectively span we mean the following:

Definition 4.5. Given a sequence of points $p_i \in \mathbb{R}^n$, we say that $\{p_i\}_{i=0}^k$ α−effectively span a k-dimensional affine subspace if for all $i = 1, \cdots, k$

$$\|p_i - p_0\| \leq \alpha^{-1}, \quad p_i \notin B_\alpha(p_0 + \text{span}\{p_1 - p_0, \cdots, p_{i-1} - p_0\}). \tag{4.15}$$

Note that this definition is basically an affine version of Definition 2.8. The definition implies that the vectors $p_i - p_0$ are linearly independent in a quantitative way. In particular, we obtain immediately that

Lemma 4.6. *If $\{p_i\}_{i=0}^k$ α-effectively span the k-dimensional affine subspace*

$$V = p_0 + \text{span}\{p_1 - p_0, \cdots, p_k - p_0\},$$

then for all $x \in V$ there exists a unique set $\{\alpha_i\}_{i=1}^k$ such that

$$x = p_0 + \sum_{i=0}^k \alpha_i(p_i - p_0), \quad |\alpha_i| \leq c(n, \alpha) \|x - p_0\|. \tag{4.16}$$

Proof. The proof is quite straightforward. Since $\{p_i - p_0\}_{i=1}^k$ are linearly independent, we can apply the Gram-Schmidt orthonormalization process to obtain an orthonormal basis $e_1, \cdots, e_k$ for the linear space $\text{span}\{p_i - p_0\}_{i=1}^k$. By induction and (4.15), it is easy to check that for all i

$$e_i = \sum_{j=1}^i \alpha'_{ij}(p_j - p_0), \quad |\alpha'_{ij}| \leq c(n, \alpha). \tag{4.17}$$

Now the estimate follows from the fact that for all $x \in V$

$$x = p_0 + \sum_{i=1}^k \langle x - p_0, e_i \rangle e_i. \tag{4.18}$$

$\square$

With these definitions, we are ready to prove that in case μ is not too small, then its support must effectively span something k-dimensional.

Lemma 4.7. *Let* $\gamma_k = \omega_k 40^{-k}$. *There exists a* $\rho_0(n, \gamma_k, M) = \rho_0(n, M)$ *such that if* (4.14) *holds for some* $\rho \leq \rho_0$ *and if* $\mu(B_1(0)) \geq \gamma_k$, *then for every affine subspace* $V \subseteq \mathbb{R}^n$ *of dimension* $\leq k - 1$, *there exists an* $x \in S \cap B_1(0)$ *such that* $B_{10\rho}(x) \cap V = \emptyset$ *and* $\mu\left(B_\rho(x) \cap B_1(0)\right) \geq c(n, \rho) = c(n)\rho^n > 0$.

Proof. Let V be any $k - 1$-dimensional subspace, and consider the set $B_{11\rho}(V)$. Let $B_i = B_\rho(x_i)$ be a sequence of balls that cover the set $B_{11\rho}(V) \cap B_1(0)$ and such that $B_i/2 \equiv B_{\rho/2}(x_i)$ are disjoint and $x_i \in B_{11\rho}(V) \cap B_1(0)$. If N is the number of these balls, then a standard covering argument gives

$$N\omega_n\rho^n/2^n \leq \omega_{k-1}(1+\rho)^{k-1}\omega_{n-k+1}(12\rho)^{n-k+1} \leq 24^n\omega_{k-1}\omega_{n-k+1}\rho^{n-k+1}$$

$$\implies \quad N \leq 48^n \frac{\omega_{k-1}\omega_{n-k+1}}{\omega_n}\rho^{1-k}. \tag{4.19}$$

By (4.14), the measure of the set $B_{11\rho}(V)$ is bounded by

$$\mu(B_{11\rho}(V)) \leq \sum_i \mu(B_i) \leq MN\rho^k \leq 48^n \frac{\omega_{k-1}\omega_{n-k+1}}{\omega_n} M\rho \leq 10^5(50n)^n M\rho = c(n)M\rho. \tag{4.20}$$

where the next-to-last estimate is an extremely rough bound on the constants involved. Thus if

$$\rho \leq 10^{-5}(50n)^{-n}\gamma_k/(4M), \tag{4.21}$$

then $\mu(B_{11\rho}(V)) \leq \gamma_k/4$. In particular, we get that there must be some point of S not in $B_{11\rho}(V)$. More effectively, let us consider the set $S \cap B_1(0) \setminus B_{11\rho}(V)$. This set can be covered by at most $c(n, \rho) = 4^n\rho^{-n}$ balls of radius ρ centered in $x \in S \cap B_1(0) \setminus B_{11\rho}(V)$, and we also see that

$$\mu\left(B_1(0) \setminus B_{11\rho}(V)\right) \geq \frac{3\gamma_k}{4}. \tag{4.22}$$

Thus, there must exist at least one ball of radius ρ centered in x and disjoint from $B_{10\rho}(V)$ such that

$$\mu\left(B_\rho(x) \cap B_1(0)\right) \geq \frac{3\gamma_k}{4}4^{-n}\rho^n \geq c(n)\rho^n. \tag{4.23}$$

$\square$

Now if at two consecutive scales there are some balls on which the measure μ *effectively spans* k-dimensional subspaces, we show that these subspaces have to be close together.

Lemma 4.8. *Let* μ *be a positive Radon measure and assume* $\mu(B_1(0)) \geq \gamma_k$. *Additionally, let* $B_\rho(x) \subset B_1(0)$ *be a ball such that* $\mu(B_\rho(x)) \geq \gamma_k\rho^k$ *and for each* $y \in B_\rho(x)$ *we have* $\mu(B_{\rho^2}(y)) \leq M\rho^{2k}$, *where* $\rho \leq \rho_0$. *Then if* $A = V(0,1) \cap B_\rho(x)$ *and* $B = V(x, \rho) \cap B_\rho(x)$ *are* L^2-*best subspace approximations of* μ *with* $d(x, A) < \rho/2$, *then*

$$d_H(A, B)^2 \leq c(n, \rho, M)\left(D_\mu^k(x, \rho) + D_\mu^k(0, 1)\right). \tag{4.24}$$

Proof. Let us begin by observing that if $c(n, \rho, M) > 4\rho^2\delta^{-1}(n, \rho, M)$, which will be chosen later, then we may assume without loss of generality that

$$D_\mu^k(x, \rho) + D_\mu^k(0, 1) \leq \delta = \delta(n, \rho, M), \tag{4.25}$$

since otherwise (4.24) is trivially satisfied. Moreover, note that $\gamma_k >> \epsilon_n$, so equation (3.1) is valid on $B_\rho(x)$ and on $B_1(0)$.

We will estimate the distance $d_H(A, B)$ by finding $k + 1$ balls $B_{\rho^2}(y_i)$ which have enough mass and effectively span in the appropriate sense $V(x, \rho)$. Given the upper bounds on D_μ^k, we will then be in a position to prove our estimate.

Consider any $B_{\rho^2}(y) \subseteq B_1(0)$ with $\mu\left(B_{\rho^2}(y)\right) > 0$ and let $p(y) \in B_{\rho^2}(y)$ be the center of mass of μ restricted to $B_\rho(x) \cap B_{\rho^2}(y)$. Let also $\pi(p)$ be the orthogonal projection of p onto $V(x, \rho)$. By Jensen's inequality:

$$d(p(y), V(x,\rho))^2 = d(p(y), \pi(p(y)))^2 = d\left(\fint_{B_{\rho^2}(y)} z\, d\mu(z), V(x,\rho)\right)^2 \le \frac{1}{\mu(B_{\rho^2}(y))} \int_{B_{\rho^2}(y)} d(z, V(x,\rho))^2 d\mu(z).$$
(4.26)

Using this estimate and lemma 4.7 (or better its rescaled version applied to $B_\rho(x)$), we want to prove that there exists a sequence of $k + 1$ balls $B_{\rho^2}(y_i)$ with $y_i \in B_\rho(x)$ such that

  (i) $\mu\left(B_\rho(x) \cap B_{\rho^2}(y_i)\right) \ge c(n, \rho, M) > 0$
  (ii) $\{\pi(p(y_i))\}_{i=0}^k \equiv \{\pi_i\}_{i=0}^k$ effectively spans $V(x, \rho)$. In other words for all $i = 1, \cdots, k$, $\pi_i \in V(x, \rho)$ and

$$\pi_i \notin B_{5\rho^2}\left(\pi_0 + \text{span}\,(\pi_1 - \pi_0, \cdots, \pi_{i-1} - \pi_0)\right).$$
(4.27)

We prove this statement by induction on $i = 0, \cdots, k$. For $i = 0$, the statement is trivially true since $\mu(B_\rho(x)) \ge \gamma_k \rho^k$. In order to find y_{i+1}, consider the subspace $V^{(i)} = \pi_0 + \text{span}\,(\pi_1 - \pi_0, \cdots, \pi_i - \pi_0)$. By lemma 4.7 applied to the ball $B_\rho(x)$, there exists some $B_{\rho^2}(y_{i+1})$ such that $\mu\left(B_\rho(x) \cap B_{\rho^2}(y_{i+1})\right) \ge c(n, \rho, M) > 0$, $y_{i+1} \in B_\rho(x)$ and

$$y_{i+1} \notin B_{10\rho^2}\left(\pi_0 + \text{span}\,(\pi_1 - \pi_0, \cdots, \pi_i - \pi_0)\right).$$
(4.28)

By definition of center of mass, it is clear that $d(y_{i+1}, p(y_{i+1})) \le \rho^2$. Moreover, by item (i) and equation (4.26), we get

$$d(p(y_{i+1}), V(x,\rho))^2 \le c \int_{B_\rho(x) \cap B_{\rho^2}(y_{i+1})} d(z, V(x,\rho))^2 d\mu(z) \le c D_\mu^k(x, \rho) \le c\delta.$$
(4.29)

Thus by the triangle inequality we have $d(y_{i+1}, \pi_{i+1}) \le 2\rho^2$ if $\delta \le \delta_0(n, \rho, M)$ is small enough. This implies (4.27). Using similar estimates, we also prove $d(p(y_{i+1}), V(0, 1))^2 \le c' D_\mu^k(0, 1)$ for all $i = -1, 0, \cdots, k - 1$. Thus by the triangle inequality

$$d(\pi_{i+1}, V(0, 1)) \le d(\pi_{i+1}, p(y_{i+1})) + d(p(y_{i+1}), V(0, 1)) \le c(n, \rho, M)\left(D_\mu^k(x, \rho) + D_\mu^k(0, 1)\right)^{1/2}.$$
(4.30)

Now consider any $y \in V(x, \rho)$. By item (ii) and lemma 4.6, there exists a unique set $\{\beta_i\}_{i=1}^k$ such that

$$y = \pi_0 + \sum_{i=1}^k \beta_i(\pi_i - \pi_0), \quad |\beta_i| \le c(n, \rho) \|y - \pi_0\|.$$
(4.31)

Hence for all $y \in V(x, \rho) \cap B_\rho(x)$, we have

$$d(y, V(0, 1)) \leq d(\pi_0, V(0, 1)) + \sum_i |\beta_i| \left[d(\pi_i, V(0, 1)) + d(\pi_0, V(0, 1)) \right] \leq c(n, \rho, M) \left(D_\mu^k(x, \rho) + D_\mu^k(0, 1) \right)^{1/2}.$$

(4.32)

By lemma 4.2, this completes the proof of (4.24). □

4.3. Comparison between L^2 and L^∞ planes.

Given $B_r(x)$, we denote as before by $V(x, r)$ one of the k-dimensional subspace minimizing $\int_{B_r(x)} d(y, V)^2 d\mu$. Suppose that the support of μ satisfies a uniform one-sided Reifenberg condition, i.e. suppose that there exists a k-dimensional plane $L(x, r)$ such that $x \in L(x, r)$ and

$$\text{supp} (\mu) \cap B_r(x) \subseteq B_{\delta r}(L(x, r)).$$

(4.33)

Then, by the same technique used in lemma 4.8, we can prove that

Lemma 4.9. *Let μ be a positive Radon measure with $\mu(B_1(0)) \geq \gamma_k$ and such that for all $B_\rho(y) \subseteq B_1(0)$ we have $\mu(B_\rho(y)) \leq M\rho^k$ and (4.33). Then*

$$d_H(L(0, 1) \cap B_1(0), V(0, 1) \cap B_1(0))^2 \leq c(n, \rho, M) \left(\delta^2 + D_\mu^k(0, 1) \right).$$

(4.34)

4.4. bi-Lipschitz equivalences.

In this subsection, we study a particular class of maps with nice local properties. These maps are a slightly modified version of the maps which are usually exploited to prove Reifenberg's theorem, see for example [Rei60, Tor95, DT12], [Mor66, section 10.5] or [Sim]. The estimates in this section are standard in literature.

We start by defining the functions σ. For some $0 < r \leq 1$, let $\{x_i\}$ be an $r/10$-separated subset of $\mathbb{R}^n$, i.e.,

(i) $d(x_i, x_j) \geq r/10$.

Let also p_i be a points in $\mathbb{R}^n$ with

(ii) $p_i \in B_{10r}(x_i)$

and let V_i be a sequence of k-dimensional linear subspaces.

By standard theory, it is easy to find a locally finite smooth partition of unity $\lambda_i : \mathbb{R}^n \to [0, 1]$ such that

(iii) $\text{supp}(\lambda_i) \subseteq B_{3r}(x_i)$ for all i,
(iv) for all $x \in \bigcup_i B_{2r}(x_i)$, $\sum_i \lambda_i(x) = 1$ and $\sum_i \lambda_i(x') \leq 1$ for all $x' \in \mathbb{R}^n$,
(v) $\sup_i \|\nabla \lambda_i\|_\infty \leq c(n)/r$,
(vi) if we set $1 - \psi(x) = \sum_i \lambda_i(x)$, then ψ is a nonnegative smooth function with $\|\nabla \psi\|_\infty \leq c(n)/r$.

Note that by (iii), and since x_i is r-separated, there exists a constant $c(n)$ such that for all x, $\lambda_i(x) > 0$ for at most $c(n)$ different indexes.

For convenience of notation, set $\pi_V(v)$ to be the orthogonal projection onto the linear subspace V of the free vector v, and set

$$\pi_{p_i, V_i}(x) = p_i + \pi_{V_i}(x - p_i).$$

(4.35)

In other words, π_{p_i,V_i} is the affine projection onto the affine subspace $p_i + V_i$. Recall that π_{V_i} is a linear map, and so the gradients of π_{V_i} and of π_{p_i,V_i} at every point are equal to π_{V_i}.

Definition 4.10. Given $\{x_i, p_i, \lambda_i\}$ satisfying (i) to (vi), and given a family of linear k-dimensional spaces V_i, we define a smooth function $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\sigma(x) = x + \sum_i \lambda_i(x)\pi_{V_i^\perp}(p_i - x) = \psi(x)x + \sum_i \lambda_i(x)\pi_{p_i,V_i}(x) . \tag{4.36}$$

By local finiteness, it is evident that σ is smooth. Moreover, if $\psi(x) = 1$, then $\sigma(x) = x$. It is clear that philosophically σ is a form of "smooth interpolation" between the identity and the projections onto the subspaces V_i. It stands to reason that if V_i are all close together, then this map σ is close to being an orthogonal projection in the region $\bigcup_i B_{2r}(x_i)$.

Lemma 4.11. *Suppose that there exists a k-dimensional linear subspace $V \subseteq \mathbb{R}^n$ and a point $p \in \mathbb{R}^n$ such that for all i*

$$d_G(V_i, V) \leq \delta , \quad d(p_i, p + V) \leq \delta . \tag{4.37}$$

Then the map σ restricted to the set $U = \psi^{-1}(0) = \left(\sum_i \lambda_i\right)^{-1}(1)$ can be written as

$$\sigma(x) = \pi_{p,V}(x) + e(x) , \tag{4.38}$$

and $e(x)$ is a smooth function with

$$\|e\|_\infty + r\|\nabla e\|_\infty \leq c(n)\delta . \tag{4.39}$$

Remark 4.1. Thus, on U we have that σ is the affine projection onto V plus an error which is small in C^1.

Proof. On the set U, we can define

$$e(x) = \sigma(x) - \pi_{p,V}(x) = -\pi_{p,V}(x) + \sum_i \lambda_i(x) \cdot \left(\pi_{p_i,V_i}(x)\right)$$

$$= \sum_i \lambda_i(x) \cdot (p_i - p - \pi_V(p_i - p) + \pi_V(p_i) - \pi_{V_i}(p_i) + \pi_{V_i}(x) - \pi_V(x)) . \tag{4.40}$$

By (4.37) and lemma 4.4, we have the estimates

$$\|p_i - p - \pi_V(p_i - p)\| \leq \delta , \quad \left\|\pi_V(x - p_i) - \pi_{V_i}(x - p_i)\right\| \leq 2\delta \|x - p_i\| \leq 20\delta r . \tag{4.41}$$

This implies

$$\|e\|_{L^\infty(U)} \leq c(n)(1 + 13r)\delta \leq c(n)\delta . \tag{4.42}$$

As for ∇e, we have

$$\nabla e = \sum_i \nabla\lambda_i(x) \cdot (p_i - p - \pi_V(p_i - p) + \pi_V(p_i) - \pi_{V_i}(p_i) + \pi_{V_i}(x) - \pi_V(x)) + \sum_i \lambda_i(x)\nabla\left(\pi_{V_i}(x) - \pi_V(x)\right) . \tag{4.43}$$

The first sum is easily estimated, and since $\langle \nabla(\pi_W)|_x, w \rangle = \pi_W(w)$, we can still apply lemma 4.4 and conclude:

$$\|\nabla e\|_{L^\infty(U)} \leq \frac{c(n)}{r}\delta . \tag{4.44}$$

$\square$

As we have seen, σ is in some sense close to the affine projection to $p + V$. In the next lemma, which is similar in spirit to [Sim, squash lemma], we prove that the image through σ of a graph over V is again a graph over V with nice bounds.

Lemma 4.12 (squash lemma). *Fix $\rho \leq 1$ and some $B_{r/\rho}(y) \subseteq \mathbb{R}^n$, let $I = \{x_i\} \cap B_{5r/\rho}(y)$ be an $r/10$-separated set and define σ as in Definition 4.10. Let also p_i be point such that $p_i \in B_{10r}(x_i)$. Suppose that there exists a k-dimensional subspace V and some $p \in \mathbb{R}^n$ such that $d(y, p + V) \leq \delta r$ and for all i:*

$$d(p_i, p + V) \leq \delta r \quad and \quad d_G(V_i, V) \leq \delta \,. \tag{4.45}$$

Suppose also that there exists a C^1 function $g : V \to V^\perp$ such that $G \subseteq \mathbb{R}^n$ is the graph

$$G = \{p + x + g(x) \ \ for \ \ x \in V\} \cap B_{r/\rho}(y) \,,$$

and $r^{-1} \|g\|_\infty + \|\nabla g\|_\infty \leq \delta'$. There exists a $\delta_0(n) > 0$ sufficiently small such that if $\delta \leq \delta_0 \rho$ and $\delta' \leq 1$, then

(i) *$\forall z \in G$, $r^{-1} |\sigma(z) - z| \leq c(n)(\delta + \delta')\rho^{-1}$, and σ is a C^1 diffeomorphism from G to its image,*

(ii) *the set $\sigma(G)$ is contained in a C^1 graph $\{p + x + \tilde{g}(x), \ \ x \in V\}$ with*

$$r^{-1} \|\tilde{g}\|_\infty + \|\nabla \tilde{g}\|_\infty \leq c(n)(\delta + \delta')\rho^{-1} \,. \tag{4.46}$$

(iii) *moreover, if U' is such that $B_{c(\delta+\delta')\rho^{-1}}(U') \subseteq \psi^{-1}(0)$, then the previous bound is independent of δ', in the sense that*

$$r^{-1} \|\tilde{g}\|_{L^\infty(U' \cap V)} + \|\nabla \tilde{g}\|_{L^\infty(U' \cap V)} \leq c(n)\delta\rho^{-1} \,. \tag{4.47}$$

For example, if $\delta' \leq \delta_0(n)\rho^{-1}$, we can take $U' = \bigcup_i B_{1.5r}(x_i)$.

(iv) *the map σ is a bi-Lipschitz equivalence between G and $\sigma(G)$ with bi-Lipschitz constant $\leq 1 + c(n)(\delta + \delta')^2\rho^{-2}$.*

Proof. For convenience, we fix $r = 1$ and $p = 0$. By notation, given any map $f : \mathbb{R}^n \to \mathbb{R}^m$, $p \in \mathbb{R}^n$ and $w \in T_p(\mathbb{R}^n) = \mathbb{R}^n$, we will denote by $\nabla|_p f[w]$ the gradient of f evaluated at p and applied to the vector w.

Recall that

$$\sigma(x + g(x)) = \psi(z)(x + g(x)) + \sum_{x_i \in I} \lambda_i(z) \left(\pi_{p_i, V_i}(x + g(x)) \right), \quad 1 - \psi(x) = \sum_{x_i \in I} \lambda_i(x), \tag{4.48}$$

where we have set for convenience $z = z(x) = x + g(x)$. Define $h(x)$ by

$$(1 - \psi(z))x + h(x) \equiv \sum_i \lambda_i(z) \left(\pi_{p_i, V_i}(x + g(x)) \right). \tag{4.49}$$

Set also $h^T(x) = \pi_V(h(x))$ and $h^\perp(x) = \pi_{V^\perp}(h(x))$. By projecting the function $\sigma(x + g(x))$ onto V and its orthogonal complement we obtain

$$\sigma(x + g(x)) \equiv \sigma^T(x) + \sigma^\perp(x) \,,$$

$$\sigma^T(x) = x + h^T(x) \,, \quad \sigma^\perp(x) = \psi(z)g(x) + h^\perp(x) \,. \tag{4.50}$$

We claim that if $\delta' \leq 1$, then

$$\left\| h^T(x) \right\| + \left\| \nabla h^T(x) \right\| \leq \frac{c\delta}{\rho} \,, \tag{4.51}$$

where this bound is *independent* of δ' as long as $\delta' \leq 1$. Indeed, for all $x \in V$ we have

$$h^T(x) = \pi_V \left[\sum_i \lambda_i(z) \left(\pi_{p_i, V_i}(x + g(x)) - x \right) \right] = \sum_i \lambda_i(z) \pi_V \left[\left(\pi_{p_i, V_i}(x) - \pi_V(x) \right) + \pi_{V_i}(g(x)) \right] \qquad (4.52)$$

Given (4.45) and lemma 4.4, with computations similar to (4.41), we get $\left\| h^T(x) \right\| \leq c\delta(1 + \rho^{-1}) \leq c\delta\rho^{-1}$. As for the gradient, we get for any vector $w \in V$

$$\nabla h^T|_x[w] = \pi_V \left[\sum_i \nabla \lambda_i|_z \left[w + \nabla g|_x[w] \right] \left(\pi_{p_i, V_i}(x + g(x)) - x \right) + \sum_i \lambda_i(z) \left(\pi_{V_i}(w + \nabla g[w]) - w \right) \right], \quad (4.53)$$

In particular, we obtain

$$\left\| \nabla h^T|_x[w] \right\| \leq \sum_i \left\| \nabla \lambda_i \right\| (1 + \|\nabla g\|) \|w\| \left\| \pi_{p_i, V_i}(x + g(x)) - x \right\| + \sum_i \lambda_i(z) \left(\left\| \pi_{V_i}(w) - w \right\| + \left\| \pi_{V_i}(\nabla g[w]) \right\| \right). \qquad (4.54)$$

For the first term, we can estimate

$$\|\nabla \lambda_i\| \leq c(n), \quad \|\nabla g\| \leq \delta' \leq 1, \quad \left\| \pi_{p_i, V_i}(x + g(x)) - x \right\| \leq \left\| \pi_{p_i, V_i}(x) - x \right\| + \left\| \pi_{V_i}(g(x)) \right\|. \qquad (4.55)$$

Since $x \in V$ with $\|x\| \leq \rho^{-1}$, and $g(x) \in V^{\perp}$, by (4.45) and lemma 4.4 we obtain

$$\left\| \pi_{p_i, V_i}(x + g(x)) - x \right\| \leq c\delta\rho^{-1}. \qquad (4.56)$$

As for the second term, we have

$$\left\| \pi_{V_i}(w) - w \right\| \leq c\delta \|w\|, \quad \left\| \pi_{V_i}(\nabla g[w]) \right\| \leq c\delta\delta' \|w\| \leq c\delta \|w\|. \qquad (4.57)$$

Summing all the contributions, we obtain (4.51) as wanted.

Thus we can apply the inverse function theorem on the function $\sigma^T(x) : V \to V$ and obtain a C^1 inverse Q such that for all $x \in V$, $\|Q(x) - x\| + \|\nabla Q - id\| \leq c(n)\delta\rho^{-1}$, and if $\psi(x + g(x)) = 1$, then $Q(x) = x$. So we can write that for all $x \in V$

$$\sigma(x + g(x)) = \sigma^T(x) + \tilde{g}(\sigma^T(x)) \quad \text{where} \quad \tilde{g}(x) = \sigma^{\perp}(Q(x)) = h^{\perp}(Q(x)) + \psi(z(Q(x))) g(Q(x)). \qquad (4.58)$$

Arguing as above, we see that $h^{\perp}(x)$ is a C^1 function with

$$\left\| h^{\perp}(x) \right\| + \left\| \nabla h^{\perp}(x) \right\| \leq \frac{c\delta}{\rho}, \qquad (4.59)$$

and this bound is independent of δ' (as long as $\delta' \leq 1$).

Thus the function $\tilde{g} : V \to V^{\perp}$ satisfies for all x in its domain

$$\|\tilde{g}(x)\| + \|\nabla \tilde{g}(x)\| \leq c(n)(\delta + \delta')\rho^{-1}, \qquad (4.60)$$

Moreover, for those x such that $\psi(Q(x) + g(Q(x))) = 0$, the estimates on $\tilde{g}$ are independent of δ', in the sense that $\|\tilde{g}(x)\| + \|\nabla \tilde{g}(x)\| \leq c(n)\delta\rho^{-1}$. Note that by the previous bounds we have

$$\|Q(x) + g(Q(x)) - x\| \leq c(\delta + \delta')\rho^{-1}, \qquad (4.61)$$

and so if $B_{c(\delta+\delta')\rho^{-1}}(U') \subset \psi^{-1}(0)$, then for all $x \in U' \cap V$, $\psi(Q(x) + g(Q(x))) = 0$. This proves items (ii), (iii). As for item (i), it is an easy consequence of the estimates in (4.51), (4.59).

Now since both G and $\sigma(G)$ are Lipschitz graphs over V, it is clear that the bi-Lipschitz map induced by π_V would have the right bi-Lipschitz estimate. Since σ is close to π_V, it stands to reason that this property

remains true. In order to check the estimates, we need to be a bit careful about the horizontal displacement of σ.

bi-Lipschitz estimates In order to prove the estimate in (iv), we show that for all $z = x + g(x) \in G$ and for all unit vectors $w \in T_z(G) \subset \mathbb{R}^n$, we have

$$\left| \|\nabla\sigma|_z[w]\|^2 - 1 \right| \le c(\delta + \delta')^2 . \tag{4.62}$$

First of all, note that if $\psi(z) = 1$, then σ is the identity, and there's nothing to prove.

In general, we have that

$$\nabla\sigma|_z[w] = \underbrace{\left(\psi(z)w + \sum_i \lambda_i(z)\pi_{V_i}[w] \right)}_{:=A} + \underbrace{\left(z\nabla\psi[w] + \sum_i \pi_{p_i,V_i}(z)\nabla\lambda_i[w] \right)}_{:=B} . \tag{4.63}$$

Since $\psi(z) + \sum_i \lambda_i(z) = 1$ everywhere by definition, we have

$$\|B\| = \left\| \sum_i (\pi_{p_i,V_i}(z) - z)\nabla\lambda_i[w] \right\| \le c \sup_i \left\{ \left\| \pi_{p_i,V_i}(z) - z \right\| \right\} \le c(\delta + \delta') . \tag{4.64}$$

This last estimate comes from the fact that G is the graph of g over V with $\|g\|_\infty \le \delta'$. Moreover, we can easily improve the estimate for B in the horizontal direction using lemma 4.4. Indeed, since $\pi_{p_i,V}(z) - z = -\pi_{V_i^\perp}(z - p_i)$, we have

$$\|\pi_V B\| = \left\| \sum_i \pi_V \left(\pi_{p_i,V_i}(z) - z \right) \nabla\lambda_i[w] \right\| \le c \sup_i \left\{ \left\| \pi_V \left(\pi_{p_i,V_i}(z) - z \right) \right\| \right\} \tag{4.65}$$

$$\le c \sup_i \left\{ \left\| \pi_V \left(\pi_{V_i^\perp}(x + g(x)) - \pi_{V_i^\perp}(p_i) \right) \right\| \right\} \le c(\delta^2 + \delta'\delta) .$$

As for A, by adapting the proof of lemma 4.11, we get $\|A - \pi_V[w]\| \le c (\delta + \delta')$. Moreover, also in this case we get better estimates for A in the horizontal direction. Indeed, we have

$$\|\pi_V(A) - \pi_V[w]\| = \left\| \psi(z)\pi_V[w] + \sum_i (\lambda_i(z)\pi_V[\pi_{V_i}[w]]) - \pi_V[w] \right\| = \left\| \sum_i \lambda_i(z) (\pi_V[\pi_{V_i}[w] - \pi_V[w]]) \right\| . \tag{4.66}$$

Now let $w = \pi_V[w] + \pi_{V^\perp}[w] = w_V + w_{V^\perp}$. Then we have

$$\|\pi_V(A) - \pi_V[w]\| \le \sum_i \lambda_i(z) \left(\left\| \pi_V[\pi_{V_i}[w_V] - w_V] \right\| + \left\| \pi_V[\pi_{V_i}[w_{V^\perp}]] \right\| \right) \tag{4.67}$$

$$= \sum_i \lambda_i(z) \left(\left\| \pi_V[\pi_{V_i^\perp}[w_V]] \right\| + \left\| \pi_V[\pi_{V_i}[w_{V^\perp}]] \right\| \right) .$$

Since G is the Lipschitz graph of g over V with $\|\nabla g\| \le c\delta'$, then $\|\pi_{V^\perp}[w]\| \le c\delta'$. Then, by lemma 4.4, we have

$$\|\pi_V(A) - \pi_V[w]\| \le c \sum_i \lambda_i(z) \left(\delta^2 + \delta\delta' \right) . \tag{4.68}$$

Summing up, since $\|\pi_V[w]\| \leq \|w\| = 1$, we obtain that

$$\left| \|\nabla\sigma|_z[w]\|^2 - 1 \right| = \left| \|\pi_{V^\perp}\nabla\sigma|_z[w]\|^2 + \|(\pi_V\nabla\sigma|_z[w] - \pi_V[w]) + \pi_V[w]\|^2 - 1 \right|$$

$$\leq c(\delta + \delta')^2 + \left| \|\pi_V[w]\|^2 - 1 \right| = c(\delta + \delta')^2 + \|\pi_{V^\perp}[w]\|^2 \leq c(\delta + \delta')^2. \tag{4.69}$$

$\square$

4.5. Pointwise Estimates on D.

We wish to see in this subsection how (3.7) implies pointwise estimates on D, which will be convenient in the proof of the generalized Reifenberg results. Indeed, the following is an almost immediate consequence of Remark 3.4:

Lemma 4.13. *Assume* $B_{4r}(x) \subseteq B_4(0)$ *satisfies* $\mu(B_r(x)) \geq \gamma_k r^k >> \epsilon_n r^k$ *and* $\int_{B_{2r}(x)} D_\mu^k(y, 2r) \, d\mu(y) < \delta^2 (2r)^k$. *Then there exists* $c(n)$ *such that* $D_\mu^k(x, r) < c\delta^2$. *In particular, if* (3.7) *holds then for every* $B_r(x) \subseteq B_1(0)$ *such that* $\mu(B_r(x)) \geq 4^k \gamma_k r^k$ *we have that* $D_\mu^k(x, r) < c\delta^2$.

Proof. First of all, note that in all of our theorems we just investigate properties of $\mu|_{B_1(0)}$, thus the first assumption is not too restrictive.

Moreover, it is easy to see that for all $y \in B_2(0)$ and $s \in [2r, 4r]$, we have

$$D(y, 2r) \leq 16^{-k-2} D(y, s) = c(k) D(y, s). \tag{4.70}$$

Moreover, notice that if (3.7) holds, then

$$\int_{B_{2r}(x)} D(y, 2r) d\mu(y) \leq \int_{B_{4r}(x)} \left(\int_{2r}^{4r} D_\mu^k(y, t) \, \frac{dt}{t} \right) d\mu(y) < \delta^2 r^k. \tag{4.71}$$

$\square$

5. PROOF OF THEOREM 3.3: THE DISCRETE-REIFENBERG

First of all, note that, by definition of μ, the statement of this theorem is equivalent to

$$\mu(B_1(0)) \leq D(n)/\omega_k. \tag{5.1}$$

In the proof, we will fix the constant $C_1(k) \leq 40^k \omega_k$ and therefore the positive scale $\rho(n, C_1(k)) = \rho(n) < 1$ according to lemma 4.7. For convenience, we will assume that $\rho = 2^q$, $q \in \mathbb{N}$, so that we will be able to use the sum bounds (3.12) more easily.

Reduction to a quantized measure Given a general measure of the form $\mu \equiv \sum_{s \in S} \omega_k r_s^k \delta_{x_s}$, it will be convenient for us to consider a similar measure but with

$$r_s \in \rho(n)^{\mathbb{N}} = \{\rho^n \ s.t. \ n \in \mathbb{N}\}. \tag{5.2}$$

In order to do so, we define $\tilde{r}_s = \max \left\{ t \in \rho(n)^{\mathbb{N}} \ s.t. \ t \leq r_s \right\}$ and

$$\tilde{\mu} = \sum_{s \in S} \omega_k \tilde{r}_s^k \delta_{x_s}. \tag{5.3}$$

It is clear that $\tilde{\mu} \leq \mu \leq \rho(n)^k \tilde{\mu}$, and given the monotonicity of D_μ^k wrt μ explained in remark 3.3, the bound (3.7) is still valid with $\tilde{\mu}$ in place of μ.

We will prove the theorem for $\tilde{\mu}$, and in particular we will prove that

$$\tilde{\mu}(B_1(0)) \leq C_1(k). \tag{5.4}$$

With this bound, it is evident that

$$\mu(B_1(0)) \leq \rho(n)^k \tilde{\mu}(B_1(0)) \leq C_1(k)\rho(n)^k \equiv D(n)/\omega_k. \tag{5.5}$$

In other words, we can assume without essential loss of generality that $r_s \in \rho(n)^{\mathbb{N}}$.

For convenience of notation, we will still denote $\tilde{\mu}$ simply by μ.

Bottom scale In the proof, it will be convenient to assume that $r_s \geq \bar{r} > 0$. It is clear that, by means of a simple limiting argument, this assumption is not restrictive. In particular, fix any positive radius $\bar{r} = r_A = \rho^A$ for some $A \in \mathbb{N}$, and consider the measure $\mu_{\bar{r}} \leq \mu$ defined by

$$\mu_{\bar{r}} = \sum_{s \ \ s.t. \ \ r_s \geq \bar{r}} \omega_k r_s^k \delta_{x_s}. \tag{5.6}$$

Note that this is a finite sum if $\bar{r}$ is positive. By Remark 3.3, we see that $\mu_{\bar{r}}$ satisfies all the hypothesis of this theorem, and since $\mu_{\bar{r}} \nearrow \mu$, if we prove uniform bounds on $\mu_{\bar{r}}(B_1(0))$ which are *independent* of $\bar{r}$, we can conclude the theorem.

For this reason, in the rest of the proof we will assume for simplicity that $r_s \geq \bar{r} = \rho^A > 0$ for all s.

5.1. First induction: upwards.

We are going to prove inductively on $j = A, \cdots, 0$ that for all $x \in B_1(0) \subset \mathbb{R}^n$ and $\mathfrak{r}_j = \rho^j \leq 1$, either $B_{\mathfrak{r}_j}(x)$ is contained in one of the balls $\{B_{r_s}(x_s)\}_{s \in S}$, or we have the bound

$$\mu\left(B_{\mathfrak{r}_j}(x)\right) \leq C_1(k)\mathfrak{r}_j^k. \tag{5.7}$$

Note that, for $j = A$, this bound follows from the definition of the measure μ and the assumption that $r_s \geq \bar{r}$. Note also that this implies $\mathfrak{r}_j \leq 2r_s$ for all $s \in S$ and $x_s \in \overline{B_{\mathfrak{r}_j}(x)}$.

Clearly, we can assume wlog that $\mu\left(B_{\mathfrak{r}_j}(x)\right) \geq \gamma_k \mathfrak{r}_j^k$, otherwise there is nothing to prove. This observation will be essential in order to apply lemma 4.8.

Moreover, as long as we are trying to prove (5.7), we can replace wlog μ with $\mu|_{B_{\mathfrak{r}_j}(x)}$. Indeed, by Remark 3.3, all the hypotheses of theorem 3.3 hold also for any restriction of μ, in particular equation (3.7). Thus, from now on, μ will indicate $\mu|_{B_{\mathfrak{r}_j}(x)}$ and $S = \text{supp}(\mu) \subset \overline{B_{\mathfrak{r}_j}(x)}$.

Remark 5.1. Note that if $\nu = \mu|_{B_{\mathfrak{r}_j}(x)}$, then all the D_ν^k on balls $B_s(y) \supset B_{\mathfrak{r}_j}(x)$ are controlled. Indeed, we have

$$D_\mu^k(x, \mathfrak{r}_j) = D_\nu^k(x, \mathfrak{r}_j) = \mathfrak{r}_j^{-k-2} \int_{B_{\mathfrak{r}_j}(x)} d(z, V(x, \mathfrak{r}_j))^2 d\mu(z) = \mathfrak{r}_j^{-k-2} \int_{B_s(y)} d(z, V(x, \mathfrak{r}_j))^2 d\nu(z) \geq \frac{s^{k+2}}{\mathfrak{r}_j^{k+2}} D_\nu^k(y, s). \tag{5.8}$$

In particular, this implies that if $B_{\mathfrak{r}_j}(x) \subseteq B_s(y)$, then

$$D_\nu^k(y, s) \leq \left(\frac{\mathfrak{r}_j}{s}\right)^{k+2} D_\nu^k(x, \mathfrak{r}_j) = \left(\frac{\mathfrak{r}_j}{s}\right)^{k+2} D_\mu^k(x, \mathfrak{r}_j). \tag{5.9}$$

In turn, as long as $\mu(B_{\mathrm{r}_j}(x)) \leq c(n)\mathrm{r}_j^k$, we also have the bound

$$\int_{\mathbb{R}^n} \left(\int_0^\infty D_\nu^k(z,t)\, \frac{dt}{t} \right) d\nu(z) < c\delta^2 s^k \,. \tag{5.10}$$

5.2. Rough estimate.
Fix some j, and suppose that (5.7) holds on all scales below r_j, i.e., for all $y \in B_1(0)$ and $\bar{r} \leq \mathrm{r}_i \leq \mathrm{r}_j$, $\mu(B_{\mathrm{r}_i}(y)) \leq C_1(k)\mathrm{r}_i^k$.

Let us first observe that we can easily obtain a bad upper bound on $\mu\left(B_{\chi\mathrm{r}_j}(x) \right)$ for any fixed $\chi > 1$. Consider the points in $\{x_s\}_{s\in S} \cap B_{\chi\mathrm{r}_j}(x)$, and divide them into two groups: the ones with $r_s \leq \mathrm{r}_j$ and the ones with $r_s > \mathrm{r}_j$. Note that by (5.2), $r_s > \mathrm{r}_j$ is equivalent to $r_s \geq \mathrm{r}_{j-1}$.

For the first group, cover them by balls $B_{\mathrm{r}_j}(z_i)$ such that $B_{\mathrm{r}_j/2}(z_i)$ are disjoint. Since there can be at most $c(n,\chi)$ balls of this form, and for all of these balls the upper bound (5.7) holds, we have an induced upper bound on the measure of this set.

As for the points with $r_s > \mathrm{r}_j$, by construction there can be only $c(n,\chi)$ many of them, and we also have the bound $r_s \leq 2\chi\mathrm{r}_j$. Summing up the two contributions, we get the very rough estimate

$$\mu\left(B_{\chi\mathrm{r}_j}(x) \right) \leq C_2(n,\chi)\mathrm{r}_j^k \,, \tag{5.11}$$

where $C_2 >> C_1$. Note that, as long as the inductive hypothesis holds, C_2 is independent of j. However, it is clear that successive repetitions of the above estimate will not lead to (5.7).

5.3. Second induction: downwards. Outline of the proof.
Suppose that (5.7) is true for all $x \in B_1(0)$ and $i = j+1, \cdots, A$. Fix $x \in \mathbb{R}^n$, and consider the set $B = B_{\mathrm{r}_j}(x)$. Recall that we always assume that B is not contained in one of the balls $B_{r_s}(x_s)$, otherwise the bound (5.7) might fail for trivial reasons. We are going to build by induction on $i \geq j$ a sequence of smooth maps $\sigma_i : \mathbb{R}^n \to \mathbb{R}^n$ and smooth k-dimensional manifolds T_i which will serve as approximations for the support of μ at scale r_i. Let us outline the inductive procedure now, and introduce all the relevant terminology. Everything described in the remainder of this subsection will be discussed more precisely over the coming pages. To begin with, we will have at the first step that

$$\begin{aligned} \sigma_j &= id, \\ T_j &= V(x, \mathrm{r}_j) \subset \mathbb{R}^n \,, \end{aligned} \tag{5.12}$$

where $V(x, \mathrm{r}_j)$ is one of the k-dimensional affine subspaces which minimizes $\int_{B_{\mathrm{r}_j}(x)} d^2(y,V)\, d\mu$. Thus, the first manifold T_j is a k-dimensional affine subspace which best approximates $B_{\mathrm{r}_j}(x)$. At future steps we can recover T_{i+1} from T_i and σ_{i+1} from the simple relation

$$T_{i+1} = \sigma_{i+1}(T_i) \,. \tag{5.13}$$

We will see that σ_{i+1} is a diffeomorphism when restricted to T_i, and thus each additional submanifold T_{i+1} is also diffeomorphic to $\mathbb{R}^k$. As part of our inductive construction we will build at each stage a Vitali covering of T_i given by

$$B_{\mathrm{r}_i}(T_i) \cap B_{\mathrm{r}_j}(x) \sim \bigcup_{t=j}^{i} \left(\bigcup_{y\in I_b^t} B_{\mathrm{r}_t}(y) \cup \bigcup_{x_s\in I_f^t} B_{r_s}(x_s) \right) \cup \bigcup_{y\in I_g^i} B_{\mathrm{r}_i}(y) \,, \tag{5.14}$$

where I_g, I_b, and I_f represent the *good*, *bad*, and *final* balls in the covering. Final balls are balls belonging to the original covering $B_{r_s}(x_s)$ such that $r_s \in [r_i, r_{i-1})$ (equivalently, by (5.2), $r_s = r_i$), and the other balls in the covering are characterized as good or bad according to how much measure they carry. Good balls are those with large measure, bad balls the ones with small measure. More precisely, we have

$$\mu(B_{r_i}(y)) \geq \gamma_k r_i^k, \quad \text{if} \quad y \in I_g^i,$$
$$\mu(B_{r_i}(y)) < \gamma_k r_i^k, \quad \text{if} \quad y \in I_b^i. \tag{5.15}$$

We will see that, over each good ball $B_{r_i}(y)$ in this covering, T_i can be written as a graph over the best approximating subspace $V(y, r_i)$ with good estimates.

Our goal in these constructions is the proof of (5.7) for the ball $B = B_{r_j}(x)$, and thus we will need to relate the submanifolds T_i, and more importantly the covering (5.14), to the set B. Indeed, this covering of T_i almost covers the set B, at least up to an excess set E_{i-1}. That is,

$$\text{supp}(\mu) \cap B \subseteq E_{i-1} \cup \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{r_t}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(y) \right) \cup \bigcup_{y \in I_g^i} B_{r_i}(y). \tag{5.16}$$

We will see that the set E_{i-1} consists of those points of B which do not satisfy a uniform Reifenberg condition. Thus in order to prove (5.7) we will need to estimate the covering (5.14), as well as the excess set E_{i-1}.

Let us now outline the main properties used in the inductive construction of the mapping $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$, and hence $T_{i+1} = \sigma_{i+1}(T_i)$. As is suggested in (5.14), it is the good balls and not the bad and final balls which are subdivided at further steps of the induction procedure. In order to better understand this construction let us begin by analyzing the good balls $B_{r_i}(y)$ more carefully. On each such ball we may consider the best approximating k-dimensional subspace $V(y, r_i)$. Since $B_{r_i}(y)$ is a good ball, one can check that most of $\text{supp}(\mu) \cap B_{r_i}(y)$ must satisfy a uniform Reifenberg and reside in a small neighborhood of $V(y, r_i)$. We denote those points which don't by $E(y, r_i)$, see (5.36) for the precise definition. Then we can define the next step of the excess set by

$$E_i = E_{i-1} \cup \bigcup_{y \in I_g^i} E(y, r_i). \tag{5.17}$$

Thus our excess set represents all those points which do not lie in an appropriately small neighborhood of the submanifolds T_i. With this in hand we can then find a submanifold $T_i' \subseteq T_i$, which is roughly defined by

$$T_i' \approx T_i \setminus \left(\bigcup_{t=j}^{i+1} \bigcup_{y \in I_b^t} B_{r_t/6}(y) \cup \bigcup_{t=j}^{i+1} \bigcup_{x_s \in I_f^t} B_{r_s/6}(x_s) \right), \tag{5.18}$$

see (5.50) for the precise inductive definition, such that

$$\text{supp}(\mu) \cap B \subseteq E_i \cup \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{r_s}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(y) \right) \cup B_{r_{i+1}/4}(T_i') \equiv R_i \cup \bigcup B_{r_{i+1}/4}(T_i'), \tag{5.19}$$

where R_i represents our remainder term, and consists of those balls and sets which will not be further subdivided at the next stage of the induction.

The basic idea is that if in our induction we find a bad ball or a final ball $B_{\mathfrak{r}}(x)$, we know that the measure carried by this ball is bounded by $C\mathfrak{r}^k$. Because of this upper bound, in order to get the final estimate on μ we do not need to further analyze the measure inside any of these balls. However, we do need to keep track of the measure carried by these balls in successive induction steps. This is why every time we find a bad or final ball, we create a corresponding "hole" in the manifold T_i, and obtain as a result T_i'. By construction, the k-dimensional measure of these holes is comparable to $\mu(B_{\mathfrak{r}}(x))$, and thus the k-dimensional measure of T_i (without holes) already "includes" the μ-measure of the final and bad balls at all bigger scales.

Now in order to finish the inductive step of the construction, we can cover $B_{\mathfrak{r}_{i+1}/4}(T_i')$ by some Vitali set

$$B_{\mathfrak{r}_{i+1}/4}(T_i') \subseteq \bigcup_{y \in I} B_{\mathfrak{r}_{i+1}}(y) \,, \tag{5.20}$$

where $y \in I \subseteq T_i'$. We may then decompose the ball centers

$$I^{i+1} = I_g^{i+1} \cup I_b^{i+1} \cup I_f^{i+1} \,, \tag{5.21}$$

based on (5.15). Now we will use Definition 4.10 and the best approximating subspaces $V(y, \mathfrak{r}_{i+1})$ to build $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$ such that

$$\mathrm{supp}\{\sigma_{i+1} - Id\} \subseteq \bigcup_{y \in I_g^{i+1}} B_{3\mathfrak{r}_{i+1}}(y) \,. \tag{5.22}$$

In order to prove the final bounds, we need to track the measure of the approximating manifolds T_i as i goes to infinity. We can use the local bi-Lipschitz estimates for σ_i at scale $\mathfrak{r}_i$ and integrate them along each manifold T_i to obtain uniform bounds on $\lambda^k(T_i)$ as i goes to infinity. This completes the outline of the inductive construction.

5.4. **First steps in the induction.** In order to make the proof more understandable, we give in detail the proof of the first steps in the downwards induction, which contains most of the necessary ideas to carry out the whole construction.

Fix any $B_{\mathfrak{r}_j}(x)$. Without loss of generality, we assume that

$$\mu(B_{\mathfrak{r}_j}(x)) \geq 2\gamma_k \mathfrak{r}_j^k \,, \tag{5.23}$$

otherwise we clearly have the measure estimate we want to prove.

With this condition, it makes sense to talk about a best L^2 approximating subspace for the support of μ on $B_{\mathfrak{r}_j}(x)$. Denote this subspace by $V(x, \mathfrak{r}_j) \equiv T_j$.

Now, we want to cover the support of μ with balls of radius $9\mathfrak{r}_{j+1}/10$ (roughly one scale smaller) in such a way to have good k-dimensional packing estimates on these balls. The idea is that condition (3.7) will force the measure μ to be *almost* supported in a small tubular neighborhood of T_j, up to small measure. So we split our ball in

$$B_{\mathfrak{r}_j}(x) = \left(B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right)\right) \cup \left(B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right)^C\right) \,. \tag{5.24}$$

The second part in this splitting is in some sense preventing the measure μ to satisfy an L^∞ Reifenberg condition. We call this part *excess set*, in particular

$$E(x, \mathfrak{r}_j) = B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right)^C . \tag{5.25}$$

Although $\mu(E)$ can be positive, it cannot be too big. Indeed, we have the trivial estimate

$$\int_{B_{\mathfrak{r}_j}(x) \setminus E(x, \mathfrak{r}_j)} d(y, V(x, \mathfrak{r}_j))^2 \, d\mu(y) + \mu\left(E(x, \mathfrak{r}_j)\right)(\mathfrak{r}_{j+1}/11)^2 \leq \int_{B_{\mathfrak{r}_j}(x)} d(y, V(x, \mathfrak{r}_j))^2 \, d\mu(y) = \mathfrak{r}_j^{k+2} D_\mu^k(x, \mathfrak{r}_j) \leq c(n, \rho)\mathfrak{r}_j^{k+2}\delta .$$

$$\tag{5.26}$$

Now, almost all of the measure μ must be concentrated in $B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right)$. In order to estimate this part, we build a covering with good overlapping properties of this set by balls of radius $\geq 9\mathfrak{r}_{j+1}/10$ centered on $V(x, \mathfrak{r}_j) \cap B_{\mathfrak{r}_j}(x)$.

This covering is built in the following way. First of all, we consider separately all the balls $\{B_{r_s}(x_s)\}_{s \in S}$ with $r_s \geq \mathfrak{r}_{j+1}$. Recall that by construction $r_s \leq 2\mathfrak{r}_j$ for all $s \in S$ with $x_s \in B_{\mathfrak{r}_j}(x)$, otherwise there's nothing to prove since $B_{\mathfrak{r}_j}(x)$ would be contained in $B_{r_s}(x_s)$ for some s. Note that, all of these balls are pairwise disjoint. We will call these balls *final balls*, and set I_j^f to be the set of centers of these balls.

In most cases, or at least in the most interesting cases, the set of final balls will be empty or very small. We complete this partial covering of $B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right)$ with other balls centered on $V(x, \mathfrak{r}_j)$ of radius $9\mathfrak{r}_{j+1}/10$ in such a way that this covering have a Vitali property.

Thus we obtain

$$B_{\mathfrak{r}_j}(x) \cap B_{\mathfrak{r}_{j+1}/11}\left(V(x, \mathfrak{r}_j)\right) \subset \bigcup_{x_s \in I_j^f} B_{r_s}(x_s) \cup \bigcup_{q \in Q} B_{9\mathfrak{r}_{j+1}/10}\left(x_q\right) . \tag{5.27}$$

Now, we split the set Q according to how much measure is contained in $B_{\mathfrak{r}_{j+1}}\left(x_q\right)$. In particular, if $\mu\left(B_{\mathfrak{r}_{j+1}}\left(x_q\right)\right) \geq \gamma_k \mathfrak{r}_{j+1}^k$, we say that this is a *good ball*, otherwise we say that this is a *bad ball*.

Now we want to build a new best approximating manifold T_{j+1} at this scale. On good balls, we have a best approximating subspace $V(x_q, \mathfrak{r}_{j+1})$, and since these balls carry enough measure, we can apply lemma 4.8 and obtain a quantitative estimate on the distance between $V(x, \mathfrak{r}_j)$ and $V(x_q, \mathfrak{r}_{j+1})$. In turn, this will allow us to apply the construction in the squash lemma 4.12. In particular, we have a smooth map σ defined on $\mathbb{R}^n$, and moreover $\sigma(T_j) \equiv T_{j+1}$ is a diffeomorphism onto its image when restricted to T_j with good quantitative bi-Lipschitz estimates. The details of this construction are carried out in subsection 5.8.

As for bad balls, we don't have to worry too much about those, since they carry really small measure. In particular,

$$\mu\left(B_{\mathfrak{r}_{j+1}}\left(x_q\right)\right) < \lambda^k \left(T_{j+1} \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)\right) . \tag{5.28}$$

Thus, in order to keep track of the measure carried by bad balls, we can simply keep track of the k-dimensional measure of the approximating manifolds T_i. In some sense, every time we hit a bad ball, we can compare its measure μ to the k-dimensional Hausdorff measure of the "hole" $T_{j+1} \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$. Since, as we will prove, for $i \geq j+1$, $T_i \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$ and $T_j \cap B_{\mathfrak{r}_{j+1}/6}\left(x_q\right)$ are substantially equal, by estimating the measure of T_i we also estimate the total measure of all the bad balls.

Moreover, since this estimate covers the measure of the whole bad ball, from this step forward we do not have to worry about $\mu|_{B_{\mathfrak{r}_{j+1}}(x_q)}$ any longer in the induction. We can use a similar argument to track the measure of final balls.

This construction is carried out in Subsection 5.9

Evidently, we cannot hope to apply these considerations also to good balls in order to get the estimates we want, because the measure of good balls is not small (a priori it could be anything). Instead, on the new good balls, we start over the same construction we outlined here (excess set, construction of the new best approximating manifold, and so on) and keep going by induction. The inductive estimates on the k-dimensional measure are carried out in Subsection 5.10.

5.5. **Second induction: details of the construction.** Let us now describe precisely the proof of this inductive construction which will lead to (5.7). For $j \leq i \leq A$, we will define a sequence of approximating manifolds T_i for the support of μ and a sequence of smooth maps σ_i such that

(i) $\sigma_j = id$, $T_j = V(x, \mathfrak{r}_j) \subset \mathbb{R}^n$,

(ii) $T_i = \sigma_i(T_{i-1})$,

(iii) for $i \geq j + 1$ and $y \in T_{i-1}$,

$$d(\sigma_i(y), y) \leq c\delta \mathfrak{r}_i, \tag{5.29}$$

and $\sigma_i|_{T_{i-1}}$ is a diffeomorphism onto T_i,

(iv) for every $y \in T_i$, $T_i \cap B_{2\mathfrak{r}_i}(y)$ is the graph over some k-dimensional affine subspace of a smooth function f satisfying

$$\frac{\|f\|_\infty}{\mathfrak{r}_i} + \|\nabla f\|_\infty \leq c\delta. \tag{5.30}$$

As outlined before, the manifolds T_i will be good approximations of the set S up to some "excess" set of small measure. Moreover, we will also introduce the concept of *good, bad* and *final* balls (whose centers will be in the sets I_g^i, I_b^i and I_f^i), a *remainder set* R_i, and the manifolds $T_i' \subseteq T_i$. Before giving the precise definitions (which are in equations (5.47), (5.41), (5.38) and (5.50) respectively), let us group here all the properties that we will need (and prove) for these objects, so that the reader can always come back to this page to have a clear picture of what are the objectives of the proof.

(v) for every $i \geq j + 1$ and $y \in I_g^i$, $d(y, V(y, \mathfrak{r}_i)) \leq c\delta \mathfrak{r}_i$, the set $T_i \cap B_{1.5\mathfrak{r}_i}(y)$ is the graph over $V(y, \mathfrak{r}_i)$ of a smooth function f satisfying (5.30), where $V(y, r)$ is one of the k-dimensional affine subspaces minimizing $\int_{B_r(y)} d^2(y, V) \, d\lambda^k$,

(vi) for all i, we have the inclusion

$$\mathrm{supp}\,(\mu) \cap B \subseteq B_{\mathfrak{r}_{i+1}/10}\left(T_i'\right) \cup R_i, \tag{5.31}$$

The last two properties needed are the key for the final volume estimates:

(vii) we can estimate

$$\lambda^k(\sigma_i^{-1}(T_i' \cap B_{2\mathfrak{r}_j}(x))) + \left(\#\left(I_b^i\right) + \#\left(I_f^i\right)\right)\omega_k(\mathfrak{r}_i/10)^k \leq \lambda^k(T_{i-1}' \cap B_{2\mathfrak{r}_j}(x)), \tag{5.32}$$

(viii) we can estimate the excess set by

$$\mu(E(y, \mathfrak{r}_i))\mathfrak{r}_{i+1}^2 \leq C(n)\mathfrak{r}_i^{k+2}D_\mu^k(y, 2\mathfrak{r}_i). \tag{5.33}$$

At the first step of our induction, we can assume wlog that $\mu(B_{\mathfrak{r}_j}(x)) \geq 2\gamma_k \mathfrak{r}_j^k$. We set $I_b^j = \emptyset$, $I_g^j = \{\bar{x}\}$, where $\bar{x}$ is the center of mass of $\mu|_{B_{\mathfrak{r}_j}(x)}$, $T_j = V(x, \mathfrak{r}_j)$ and $\sigma_j = id$. We set $E(x, \mathfrak{r}_j)$ to be the excess set defined by (5.25), and $R_j = E(x, \mathfrak{r}_j)$.

Moreover, we also set

$$I_f^j(y) = \left\{ x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s \in [\mathfrak{r}_j, 1) \right\} \overset{(5.2)}{=} \left\{ x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s = \mathfrak{r}_j \right\}, \tag{5.34}$$

where we used the fact that $B_{\mathfrak{r}_j}(x)$ is not contained in any of the balls $\{B_{r_s}(x_s)\}_{s \in S}$ to show the equivalence, and we set

$$T_j' = T_j \setminus \bigcup_{x_s \in I_f^j} B_{\mathfrak{r}_j/6}(x_s) . \tag{5.35}$$

It is clear from these definitions that all the properties (i)-(viii) are satisfied.

Now we proceed by induction assuming that we have defined for all $t \in \{j, \cdots, i\}$ I_g^t, I_b^t, I_f^t, the maps σ_t and the manifolds T_t, T_t'. Moreover, we assume that we also have defined for all $t \in \{j, \cdots, i-1\}$ the excess sets $\{E(y, \mathfrak{r}_t)\}_{y \in I_g^t}$ and the remainder R_{i-1}.

In the induction step, we will first build $\{E(y, \mathfrak{r}_i)\}_{y \in I_g^i}$ and R_i, and then move on to the construction of $I_g^{i+1}, I_b^{i+1}, I_f^{i+1}, \sigma_{i+1}$ and T_{i+1}, T_{i+1}'.

5.6. Excess set. Let us begin by describing the construction of the excess set. We will only be interested here in a ball $B_{\mathfrak{r}_i}(x)$ which is a good ball, in the sense that $\mu(B_{\mathfrak{r}_i}(x)) \geq \gamma_k \mathfrak{r}_i^k$.

Thus define $V(x, r)$ to be (one of) the k-dimensional plane minimizing $\int_{B_r(x)} d(y, V)^2 d\mu$, and define also the excess set to be the set of points which are some definite amount away from the best plane V. Precisely,

$$E(x, \mathfrak{r}_i) = \left(B_{\mathfrak{r}_i}(x) \setminus B_{\mathfrak{r}_{i+1}/11}(V) \right) \cap S . \tag{5.36}$$

The points in $\text{supp}(\mu) \cap E$ are in some sense what prevents the set S from satisfying a uniform one-sided Reifenberg condition at this scale. By construction, all points in E have a uniform lower bound on the distance from V, so that if we assume $\mu(B_{\mathfrak{r}_i}(x)) \geq \gamma_k \mathfrak{r}_i^k$, i.e. $B_{\mathfrak{r}_i}(x)$ is a good ball, then we can estimate

$$\int_{B_{\mathfrak{r}_i}(x) \setminus E(x, \mathfrak{r}_i)} d(y, V(x, \mathfrak{r}_i))^2 \, d\mu(y) + \mu(E(x, \mathfrak{r}_i))(\mathfrak{r}_{i+1}/11)^2 \leq \int_{B_{\mathfrak{r}_i}(x)} d(y, V(x, \mathfrak{r}_i))^2 \, d\mu(y) = \mathfrak{r}_i^{k+2} D_\mu^k(x, \mathfrak{r}_i) . \tag{5.37}$$

5.7. Good, bad and final balls. Inductively, let us define the remainder set to be the union of all the previous bad balls, final balls, and the excess sets:

$$R_i = \bigcup_{t=j}^{i} \left(\bigcup_{y \in I_b^t} B_{\mathfrak{r}_t}(y) \cup \bigcup_{x_s \in I_f^t} B_{r_s}(x_s) \cup \bigcup_{y \in I_g^t} E(y, \mathfrak{r}_t) \right) . \tag{5.38}$$

The set R_i represents everything we want to throw out at the inductive stage of the proof. We will see later in the proof how to estimate this remainder set itself. Note that $R_j = E(x, \mathfrak{r}_j) \cup \left(\bigcup_{x_s \in I_f^j} B_{r_j}(x_s) \right)$.

Now consider the points $x_s \in S$ outside the remainder set, and separate the balls $B_{r_s}(x_s)$ with radius $r_s = \mathfrak{r}_{i+1}$ from the others. Notice that by construction all the points $x_s \in S$ outside the remainder set have $r_s \leq \mathfrak{r}_{i+1}$.

Define for $y \in I_g^i$ the sets

$$I_f^{i+1}(y) = \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s \in [\mathfrak{r}_{i+1}, \mathfrak{r}_i)\} \stackrel{(5.2)}{=} \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s = \mathfrak{r}_{i+1}\}, \quad (5.39)$$

and

$$J^{i+1}(y) = \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s < \mathfrak{r}_{i+1}\} \stackrel{(5.2)}{=} \{x_s \in (S \setminus R_i) \cap B_{\mathfrak{r}_i}(y) \quad s.t. \quad r_s \leq \mathfrak{r}_{i+2}\}. \quad (5.40)$$

From this we can construct the sets

$$I_f^{i+1} = \cup_{y \in I_g^i} I_f^{i+1}(y) \quad \text{and} \quad J^{i+1} = \cup_{y \in I_g^i} J^{i+1}(y). \quad (5.41)$$

Note that by construction and inductive item (v), we have

$$S \setminus R_i = I_f^{i+1} \cup J^{i+1} \subset B_{\mathfrak{r}_{i+1}/10}(T_i'). \quad (5.42)$$

Recall that, as long as we are trying to prove the estimate (5.7), we can assume wlog that $\mu = \mu|_{B_{\mathfrak{r}_j}(x)}$ and $S = \mathrm{supp}(\mu) \subset B_{\mathfrak{r}_j}(x)$.

Let us now consider a covering of (5.42) given by

$$S \setminus R_i \subseteq I_f^{i+1} \cup \bigcup_{z \in I} B_{9\mathfrak{r}_{i+1}/10}(z), \quad (5.43)$$

where $I \subseteq T_i'$, and for any $p \neq q \in I_f^{i+1} \cup I$, $B_{r_p/5}(p) \cap B_{r_q/5}(q) = \emptyset$. Here we denote for convenience $r_p = \mathfrak{r}_{i+1}$ if $p \in I \cup I_f^{i+1}$. Note that this second property is true by definition for $p, q \in I_f^{i+1}$, we only need to complete this partial Vitali covering with other balls of the same size. To be precise, note that by (5.38) and (5.36)

$$(S \setminus R_i) \setminus \bigcup_{z \in I_f^{i+1}} B_{3r_z/5}(z) \subset \left(\bigcup_{y \in I_g^i} (B_{\mathfrak{r}_{i+1}/4}(V(y, \mathfrak{r}_i)) \cap B_{9\mathfrak{r}_i/10}(y)) \right) \cap \left(\bigcup_{t=j}^{i+1} \bigcup_{z \in I_f^t} B_{3r_z/5}(z) \cup \bigcup_{t=j}^{i} \bigcup_{z \in I_b^t} B_{3\mathfrak{r}_t/5}(z) \right)^C. \quad (5.44)$$

Take a finite covering of this last set by balls $\{B_{3\mathfrak{r}_{i+1}/10}(y)\}_{y \in Y}$. Note that we can pick

$$Y \cap \left(\bigcup_{t=j}^{i+1} \bigcup_{z \in I_f^t} B_{3r_z/5}(z) \cup \bigcup_{t=j}^{i} \bigcup_{z \in I_b^t} B_{3\mathfrak{r}_t/5}(z) \right) = \emptyset. \quad (5.45)$$

By item (iv), T_i is locally a Lipschitz graph over some k-dimensional subspace with (5.30), and thus we can choose $Y \subset T_i$.

Consider a Vitali subcovering of this set, denote I the set of centers in this subcovering. Such a subcovering will have the property that the balls $\{B_{3\mathfrak{r}_{i+1}/10}(y)\}_{y \in I}$ will be pairwise disjoint. These balls will also be disjoint from $\bigcup_{z \in I_f^{i+1}} B_{r_z/5}(z)$ by (5.45). The (finite version of) Vitali covering theorem ensures that $\bigcup_{y \in I} B_{9\mathfrak{r}_{i+1}/10}(y)$ will cover the whole set in (5.44).

Now by construction of I_f and the remainder set, all the balls $\{B_{r_s}(x_s)\}_{s \in S}$ with $r_s \geq r_{i+1}$ have already been accounted for. This means that

$$(S \setminus R_i) \setminus \bigcup_{x_s \in I_f^{i+1}} B_{r_{r_{i+1}}}(x_s) \subset \bigcup_{y \in I} B_{9r_{i+1}/10}(y) , \tag{5.46}$$

as desired.

We split the balls with centers in I into two subsets, according to how much measure they carry. In particular, let

$$I_g^{i+1} = \left\{y \in I \ \ s.t. \ \mu\left(B_{r_{i+1}}(y)\right) \geq \gamma_k r_{i+1}^k\right\} , \quad I_b^{i+1} = \left\{y \in I \ \ s.t. \ \mu\left(B_{r_{i+1}}(y)\right) < \gamma_k r_{i+1}^k\right\} . \tag{5.47}$$

5.8. **Map and manifold structure.** Let $\left\{\lambda_s^{i+1}\right\} = \{\lambda_s\}$ be a partition of unity such that for each $y_s \in I_g^{i+1}$

- $\text{supp}(\lambda_s) \subseteq B_{3r_{i+1}}(y_s)$
- for all $z \in \cup_{y_s \in I_g^{i+1}} B_{2r_{i+1}}(y_s)$, $\sum_s \lambda_s(z) = 1$
- $\max_s \|\nabla \lambda_s\|_\infty \leq C(n)/r_{i+1}$.

For every $y_s \in I_g^{i+1}$, let $V(y_s, r_{i+1})$ to be (one of) the k-dimensional subspace that minimizes $\int_{B_{r_{i+1}}(y_s)} d(z, V)^2 d\mu$. By Remark 3.4 we have

$$r_{i+1}^{-k-2} \int_{B_{r_{i+1}}(y_s)} d(z, V(y_s, r_{i+1}))^2 d\mu(z) \leq D_\mu^k(y_s, r_{i+1}) . \tag{5.48}$$

Let $p_s \in B_{r_{i+1}}(y_s)$ be the center of mass of $\mu|_{B_{r_{i+1}}(y_s)}$. It is worth observing that $p_s \in V(y_s, r_{i+1})$.

Define the smooth function $\sigma_{i+1} : \mathbb{R}^n \to \mathbb{R}^n$ as in Definition 4.10, i.e.,

$$\sigma_{i+1}(x) = x + \sum_s \lambda_s^{i+1}(x) \pi_{V(y_s, r_{i+1})^\perp}(p_s - x) . \tag{5.49}$$

With this function, we can define inductively for $i \geq j$ the sets

$$T_j = V(x, r_j), \quad T_j' = T_j \setminus \bigcup_{x_s \in I_f^j} B_{r_s/6}(x_s) \tag{5.50}$$

$$T_{i+1} = \sigma_{i+1}(T_i), \quad T_{i+1}' = \sigma_{i+1}\left(T_i' \setminus \left(\bigcup_{y \in I_b^{i+1}} B_{r_{i+1}/6}(y) \cup \bigcup_{x_s \in I_f^{i+1}} B_{r_s/6}(x_s)\right)\right) . \tag{5.51}$$

Fix any $y \in I_g^{i+1}$, and let $z \in I_g^i$ be such that $B_{9r_{i+1}/10}(y) \cap B_{9r_i/10}(z) \neq \emptyset$. By induction, $T_i \cap B_{9r_{i+1}/10}(y) \subseteq T_i \cap B_{1.5r_i}(z)$ is the graph of a C^1 function over $V(z, r_i)$. Consider the points $\{y_t\}_{t \in T_y} = I_g^{i+1} \cap B_{6r_{i+1}}(y)$. By construction it is easy to see that $d(y_t, V(z, r_i)) \leq r_{i+1}/9$, and so for all $t \in T_y$ we can apply the estimates in lemma 4.8 to the couple $B_{r_{i+1}}(y_t) \subseteq B_{r_i}(z)$ with $M = C_1$ by the first induction. Note that by (5.2) and by construction of good and final balls, for $t \in T$ and $x_s \in B_{r_{i+1}}(y_t)$, $r_s \leq r_{i+2}$. Using condition (3.7) and lemma 4.13, we obtain that for all y_s:

$$r_{i+1}^{-1} d_H\left(V(z, r_i) \cap B_{r_{i+1}}(y_s), V(y_s, r_{i+1}) \cap B_{r_{i+1}}(y_s)\right) \leq c\left(D_\mu^k(y_s, r_{i+1}) + D_\mu^k(z, r_i)\right)^{1/2} \leq c(n, \rho, C_1)\delta . \tag{5.52}$$

This implies that, if $\delta(n, \rho, C_1)$ is small enough, $T_i \cap B_{2r_{i+1}}(y)$ is a graph also over $V(y, r_{i+1})$ satisfying the same estimates as in (5.30), up to a worse constant c. That is, if δ is sufficiently small, we can apply lemma

4.12 and prove induction point (v).

Points (iii) **and** (iv) Points (iii) and (iv) are proved with similar methods. We briefly sketch the proofs of these two points.

Let $y \in T_{i-1}$, and recall the function $\psi_i \equiv 1 - \sum \lambda_s$. If $\psi_i|_{B_{2\mathfrak{r}_i}(y)}$ is identically 1, then $\sigma_i|_{B_{2\mathfrak{r}_i}(y)} = id$, and there is nothing to prove.

Otherwise, there must exist some $z' \in I_g^i \cap B_{5\mathfrak{r}_i}(y)$, and thus there exists a $z \in I_g^{i-1}$ such that $B_{8\mathfrak{r}_i}(y) \subseteq B_{1.5\mathfrak{r}_{i-1}}(z)$. By point (v) in the induction, $T_{i-1} \cap B_{1.5\mathfrak{r}_{i-1}}(z)$ is a Lipschitz graph over $V(z, \mathfrak{r}_{i-1})$. Proceeding as before, by the estimates in lemma 4.8 and lemmas 4.12, 4.13, we obtain that $T_i \cap B_{2\mathfrak{r}_i}(y)$ is also a Lipschitz graph over $V(z, \mathfrak{r}_{i-1})$ with small Lipschitz constant, and that $\|\sigma_i(p) - p\| \le c\delta\mathfrak{r}_i$ for all $p \in T_{i-1}$.

Moreover, $\sigma_i|_{T_{i-1}}$ is locally a diffeomorphism at scale $\mathfrak{r}_i$. From this we see that σ_i is a diffeomorphism on the whole T_{i-1}.

It is worth to remark a subtle point. In order to prove point (iv), we cannot use inductively (iv), we need to use point (v). Indeed, as we have seen, given any $z \in I_g^{i-1}$, then $T_{i-1} \cap B_{1.5\mathfrak{r}_{i-1}}(z)$ is a Lipschitz graph of a function f where $\|\nabla f\| \le c\delta$, and this c is *independent* of the induction step we are considering by (iii) in lemma 4.12. If we tried to iterate directly the bound given by (iv), the constant c would depend on the induction step i, and thus we could not conclude the estimate we want.

5.9. **Properties of the manifolds T_i'.**

Here we want to prove the measure estimate in (5.32). The basic idea is that bad and final balls correspond to holes in the manifold T_i, and each of these holes carries a k-dimensional measure which is proportionate to the measure inside the balls. In particular, let $B_r(y)$ be a bad or a final ball. In the first case, $r = \mathfrak{r}_{i+1}$, while in the second $r \in [\mathfrak{r}_{i+1}, \mathfrak{r}_i)$. In either case, we will see that y must be $\sim \mathfrak{r}_{i+1}$-close to T_i, which is a Lipschitz graph at scale $\mathfrak{r}_i$. This implies that $\mu(B_r(y) \cap T_i) \sim r^k$, and thus we can bound the measure of a bad or final ball with the measure of the hole we have created on T_i.

In detail, point (vi) is an immediate consequence of the definition of R_i.

In order to prove the volume measure estimate, consider that

$$T_i' \setminus \sigma_{i+1}^{-1}(T_{i+1}') \subseteq \left(\bigcup_{y \in I_b^{i+1}} B_{\mathfrak{r}_{i+1}/6}(y) \cup \bigcup_{x_s \in I_f^{i+1}} B_{r_s/6}(x_s) \right). \tag{5.53}$$

Note that the balls in the collection $\{B_{\mathfrak{r}_{i+1}/5}(y)\}_{y \in I_f^{i+1} \cup I_b^{i+1}}$ are pairwise disjoint. Pick any $y \in I_b^{i+1}$, and let $z \in I_g^i$ be such that $y \in B_{\mathfrak{r}_i}(z)$. By definition, $y \in T_i'$ and $\mu(B_{\mathfrak{r}_{i+1}}(y)) < \gamma_k\mathfrak{r}_{i+1}^k < 10^{-k}\omega_k r_{i+1}^k$. Since $T_i \cap B_{2\mathfrak{r}_i}(z)$ is a graph over $V(z, \mathfrak{r}_i)$ with $y \in T_i$, and since $B_{\mathfrak{r}_{i+1}/6}(y) \cap T_i$ is disjoint from the "holes" of T_i', then

$$\lambda^k(T_i' \cap B_{\mathfrak{r}_{i+1}/6}(y)) \ge \omega_k 7^{-k}\mathfrak{r}_{i+1}^k. \tag{5.54}$$

A similar estimate holds for the final balls. The only difference is that if $x_s \in I_f^{i+1}$, then it is not true in general that $x_s \in T_i$. However, since $D_\mu^k(z, \mathfrak{r}_i) \le c\delta$ and $\mu(\{x_s\}) = \omega_k r_{i+1}^k$, we still have that $d(x_s, V(z, \mathfrak{r}_i)) \le$

$c\delta\mathfrak{r}_{i+1} \leq \mathfrak{r}_{i+1}/100$. Given (5.30), we can conclude

$$\lambda^k(T_i' \cap B_{\mathfrak{r}_{i+1}/7}(x_s)) \geq \omega_k 10^{-k} \mathfrak{r}_{i+1}^k \,. \tag{5.55}$$

Now it is evident from the definition of T_{i+1}' that

$$\lambda^k(\sigma_{i+1}^{-1}(T_{i+1}') \cap B_{2\mathfrak{r}_j}(x)) + \#\left(I_b^{i+1} \cup I_f^{i+1}\right)\omega_k(\mathfrak{r}_{i+1}/10)^k \leq \lambda^k(T_i' \cap B_{2\mathfrak{r}_j}(x)) \,. \tag{5.56}$$

5.10. **Volume estimates on the manifold part.** Here we want to prove that for every measurable $\Omega \subseteq T_i$

$$\lambda^k(\sigma_{i+1}(\Omega)) \leq \lambda^k(\Omega) + c(n,\rho,C_1)\int_{B_{\mathfrak{r}_j}(x)} D(p, 2\mathfrak{r}_{i+1})d\mu(p) \,. \tag{5.57}$$

The main applications will be with $\Omega = T_i$ and $\Omega = T_i'$. In order to do that, we need to analyze in a quantitative way the bi-Lipschitz correspondence between T_i and T_{i+1} given by σ_{i+1}.

As we already know, $\sigma_{i+1} = id$ on the complement of the set $G = \cup_{y \in I_g^{i+1}} B_{5\mathfrak{r}_{i+1}}(y)$, so we can concentrate only on this set.

Using the same techniques as before, and in particular by lemmas 4.8 and 4.12, we can prove that for each $y \in I_g^{i+1}$, the set $T_i \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a Lipschitz graph over $V(y, \mathfrak{r}_{i+1})$ with Lipschitz constant bounded by

$$c(n,\rho,C_1)\left(D(y,\mathfrak{r}_{i+1}) + \sum_{z \in I_g^i \cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right)^{1/2} \,. \tag{5.58}$$

In a similar manner, we also have that $T_{i+1} \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a Lipschitz graph over $V(y, \mathfrak{r}_{i+1})$ with Lipschitz constant bounded by

$$c(n,\rho,C_1)\left(\sum_{z \in I_g^{i+1} \cap B_{10\mathfrak{r}_{i+1}}(y)} D(z,\mathfrak{r}_{i+1}) + \sum_{z \in I_g^i \cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right)^{1/2} \,. \tag{5.59}$$

Moreover, by the bi-Lipschitz estimates of lemma 4.12, we also know that σ_{i+1} restricted to $T_i \cap B_{5\mathfrak{r}_{i+1}}(y)$ is a bi-Lipschitz equivalence with bi-Lipschitz constant bounded by

$$L(y, 5\mathfrak{r}_{i+1}) \leq 1 + c\left(\sum_{z \in I_g^{i+1} \cap B_{10\mathfrak{r}_{i+1}}(y)} D(z,\mathfrak{r}_{i+1}) + \sum_{z \in I_g^i \cap B_{5\mathfrak{r}_i}(y)} D(z,\mathfrak{r}_i)\right) \,. \tag{5.60}$$

In order to estimate this upper bound, we use (3.3) and the definition of good balls to write

$$D(z,r) \leq c \fint_{B_r(z)} D(p, 2r)d\mu(p) \leq c(n,\rho,C_1)r^{-k}\int_{B_r(z)} D(p, 2r)d\mu(p) \,. \tag{5.61}$$

Since by construction any point $x \in \mathbb{R}^n$ can be covered by at most $c(n)$ different good balls at different scales, we can bound

$$L(y, 5\mathfrak{r}_{i+1}) \leq 1 + \frac{c(n,\rho,C_1)}{\mathfrak{r}_{i+1}^k}\int_{B_{5\mathfrak{r}_i}(y)} [D(p, 2\mathfrak{r}_{i+1}) + D(p, 2\mathfrak{r}_i)]\,d\mu(p) \,. \tag{5.62}$$

We can also badly estimate

$$D(p, 2\mathfrak{r}_{i+1}) + D(p, 2\mathfrak{r}_i) \leq c(n,\rho)D(p, 2\mathfrak{r}_i) \,. \tag{5.63}$$

Now let P_s be a measurable partition of $\Omega \cap G$ such that for each s, $P_s \subseteq B_{5r_{i+1}}(y_s)$. By summing up the volume contributions of P_s, and since evidently $\lambda^k(P_s) \leq 7^k \omega_k r_{i+1}^k$, we get

$$\lambda^k(\sigma_{i+1}(\Omega)) = \sum_s \lambda^k(\sigma_{i+1}(P_s)) \leq \sum_s \lambda^k(P_s)\left(1 + \frac{c}{r_{i+1}^k}\int_{B_{5r_i}(y_s)} D(p, 2r_i)d\mu(p)\right)$$

$$\leq \lambda^k(\Omega) + c\int_{\bigcup_{y_s \in I_g^{i+1}} B_{5r_i}(y_s)} D(p, 2r_i)d\mu(p)$$

$$\leq \lambda^k(\Omega) + c(n, \rho, C_1)\int_{B_{r_j}(x)} D(p, 2r_i)d\mu(p). \tag{5.64}$$

5.11. **Estimates on the excess set.** In this paragraph, we estimate the total measure of the excess set, which is defined by

$$E_T = \bigcup_{i=j}^{A}\bigcup_{y \in I_g^i} E(y, r_i). \tag{5.65}$$

At each y and at each scale, we have by (5.37) and (3.3)

$$\mu(E(y, r_i)) \leq c(n, \rho)r_i^k D_\mu^k(y, r_i) \leq c(n, \rho)r_i^k \fint_{B_{r_i}(x)} D_\mu^k(p, 2r_i)d\mu(p). \tag{5.66}$$

Since by definition of excess set $B_{r_i}(y)$ must be a good ball, then

$$\mu(E(y, r_i)) \leq c(n, \rho)\int_{B_{r_i}(y)} D_\mu^k(p, 2r_i)d\mu(p). \tag{5.67}$$

Now by construction of the good balls, there exists a constant $c(n)$ such that at each step i, each $x \in \mathbb{R}^n$ belongs to at most $c(n)$ good balls. Thus for each $i \geq j$, we have

$$\sum_{y \in I_g^i} \mu(E(y, r_i)) \leq c(n, \rho)\int_{\bigcup_{y \in I_g^i} B_{r_i}(y)} D_\mu^k(p, 2r_i)d\mu(p) \leq c(n, \rho)\int_{B_{r_j}(x)} D_\mu^k(p, 2r_i)d\mu(p). \tag{5.68}$$

If we sum over all scales, we get

$$\mu(E_T) \leq c(n, \rho)\sum_{i=j}^{A}\int_{B_{r_j}(x)} D_\mu^k(p, 2r_i)d\mu(p). \tag{5.69}$$

Since $\rho = 2^q$, it is clear that

$$\mu(E_T) \leq c(n, \rho)\sum_{i=j}^{A}\int_{B_{r_j}(x)} D_\mu^k\left(p, 2^{1-qi}\right)d\mu(p) \leq c(n, \rho)\delta r_j^k, \tag{5.70}$$

since the sum in the middle is clearly bounded by (3.12).

This estimate is exactly what we want from the excess set.

5.12. **Volume estimates.** By adding (5.57), with $\Omega \equiv \sigma_{i+1}^{-1}(T'_{i+1}) \cap B_{2\mathfrak{r}_j}(x)$, and (5.32), we prove that for all $i = j, \cdots, A + 1, ...$

$$\lambda^k(T'_{i+1} \cap B_{2\mathfrak{r}_j}(x)) + \#\left(I_b^{i+1} \cup I_f^{i+1}\right)\omega_k(\mathfrak{r}_{i+1}/10)^k \leq \lambda^k(T'_i \cap B_{2\mathfrak{r}_j}(x)) + c(n,\rho,C_1)\int_{B_{\mathfrak{r}_j}(x)} D(p, 2\mathfrak{r}_{i+1})d\mu(p).$$

(5.71)

Adding the contributions from all scales, by (3.7) we get

$$\lambda^k(T'_{i+1} \cap B_{2\mathfrak{r}_j}(x)) + \sum_{t=j}^{i+1} \#\left(I_b^{i+1} \cup I_f^{i+1}\right)\omega_k(\mathfrak{r}_t/10)^k$$

$$\leq \lambda^k(T_j \cap B_{2\mathfrak{r}_j}(x)) + c(n,\rho,C_1)\sum_{s=j+1}^{i+1}\int_{B_{\mathfrak{r}_j}(y)\cap B_1(0)} D(p, 2r_s)d\mu(p)$$

$$\leq \lambda^k\left(T_j \cap B_{2\mathfrak{r}_j}(x)\right)[1 + c(n,\rho,C_1)\delta] ,$$

(5.72)

where in the last line we estimated $\lambda^k\left(T_j \cap B_{2\mathfrak{r}_j}(x)\right) \sim \mathfrak{r}_j^k$, since T_j is a k-dimensional subspace, and we bounded the sum using (3.12).

In the same way, we can also bound the measure of T_{i+1} by

$$\lambda^k(T_{i+1} \cap B_{2\mathfrak{r}_j}(x)) \leq \lambda^k\left(T_j \cap B_{2\mathfrak{r}_j}(x)\right)[1 + c(n,\rho,C_1)\delta] .$$

(5.73)

5.13. **Upper estimates for μ.** Since we have assumed $r_s \geq \bar{r} = r_A$ for all $x_s \in S$, we know that for $i = A$ our construction ends, and the whole support S is contained in final and bad balls and excess sets. In other words, the sets $I_g^A = I_b^A = \emptyset$, and

$$\text{supp}(\mu) \cap B = \text{supp}(\mu) \cap B_{\mathfrak{r}_j}(x) \subseteq R_A .$$

(5.74)

This fact and the estimates in (5.70) and (5.72) imply

$$\mu(B) \leq \sum_{t=j}^{A} \#\left(I_b^t\right)\gamma_k\omega_k\mathfrak{r}_t^k + \sum_{t=j}^{A} \#\left(I_f^t\right)\omega_k\mathfrak{r}_t^k + \mu(E_T)$$

$$\leq 10^k\left(\sum_{t=j}^{A} \#\left(I_b^t\right)\omega_k(\mathfrak{r}_t/10)^k + \sum_{t=j}^{A}\sum_{x_s \in I_f^t}\omega_k(r_s/10)^k\right) + \mu(E_T) \leq C_3(k)(1 + c(n,\rho,C_1)\delta)\mathfrak{r}_j^k.$$

(5.75)

In this last estimate, we can fix $C_3(k) = 20^k\omega_k$, and $C_1(k) = 2C_3(k) = 2 \cdot 20^k\omega_k \leq 40^k\omega_k$, and $\rho(n, C_1)$ according to lemma 4.7. Now, it is easy to see that if $\delta(n,\rho,C_1)$ is sufficiently small, then

$$\mu(B) \leq C_1(k)\mathfrak{r}_j^k,$$

(5.76)

which finishes the proof of the downward induction, and hence the actual ball estimate (5.7).

6. L^2-Best Approximation theorems for Stationary varifolds

In this Section we prove the main estimate necessary for us to be able to apply the rectifiable-Reifenberg of theorems 3.3 and 3.4 to the singular sets $S_\epsilon^k(I)$ of the stratification induced by integral varifolds with bounded mean curvature.

Namely, we need to understand how to estimate on a ball $B_r(x)$ the L^2-distance of S_ϵ^k from the best approximating k-dimensional subspace. Here we carry out the computations in the Euclidean setting with mean curvature $H = 0$. The generalization to the manifold case with bounded mean curvature is straightforward.

Theorem 6.1 (L^2-Best Approximation theorem). *Let $I^m \subseteq B_9 \subset \mathbb{R}^n$ be an integral varifold satisfying* (1.14)*, and the mass bound $\mu_I(B_9(p)) \le \Lambda$, and let $\epsilon > 0$. Then there exists $\delta(n, \Lambda, \epsilon)$, $C(n, \Lambda, \epsilon) > 0$ such that if $r \le 1$ and $B_{4r}(p)$ is $(0, \delta)$-symmetric but not $(k + 1, \epsilon)$-symmetric, then for any finite measure μ on $I^m \cap B_r(p)$ we have that*

$$D_\mu(p, r) = r^{-2-k} \inf_{L^k} \int_{B_r(p)} d^2(x, L^k) \, d\mu(x) \le C r^{-k} \int_{B_r(p)} W_{8r,r}(x) \, d\mu(x), \qquad (6.1)$$

where the inf *is taken over all k-dimensional affine subspaces $L^k \subseteq T_p M$.*

6.1. Symmetry and Gradient Bounds.

In this subsection we study integral varifolds I^m with bounded mean curvature which are *not* $(k + 1, \epsilon)$-symmetric on some ball. In particular, we show that this forces for each $k + 1$-subspace V^{k+1} that $|\pi_I^\perp[V](x)| = |\langle N_x I, V \rangle|$ has some definite size in L^2, where $\pi_I^\perp[V](x)$ is the projection of V to the $N_x I$, the orthogonal complement of the tangent space $T_x I$ of I at x. More precisely:

Lemma 6.2. *Let $I^m \subseteq B_9$ be an integral varifold satisfying* (1.14)*, the mean curvature bound* (1.15)*, and the mass bound $\mu_I(B_9(p)) \le \Lambda$. Then for each $\epsilon > 0$ there exists $\delta(n, \Lambda, \epsilon) > 0$ such that if $B_4(p)$ is $(0, \delta)$-symmetric but is not $(k + 1, \epsilon)$-symmetric, then for every $k + 1$-subspace $V^{k+1} \subseteq T_p M$ we have*

$$\int_{A_{3,4}(p)} |\pi_I^\perp[V]|^2 d\mu_I \ge \delta, \qquad (6.2)$$

where $\pi_I^\perp[V](x)$ is the projection of V to $N_x I$, the orthogonal compliment to the tangent space of I at x.

Proof. The proof is by contradiction. So with $n, \Lambda, \epsilon > 0$ fixed let us assume the result fails. Then there exists a sequence I_i of integral varifolds of $B_9(p_i)$, with $B_4(p_i)$ being $(0, \delta_i)$-symmetric but not $(k + 1, \epsilon)$-symmetric, and such that for some subspaces V_i^{k+1} we have that

$$\int_{A_{3,4}(p_i)} |\pi_{I_i}^\perp[V_i]|^2 d\mu_{I_i} \le \delta_i \to 0. \qquad (6.3)$$

After rotation we can assume $V_i^{k+1} = V^{k+1}$, and then after passing to a subsequence we have that

$$I_i \longrightarrow I \subseteq B_9(0^n), \qquad (6.4)$$

in the sense of varifolds (or in the sense of the flat distance if I is an integral current). In particular, we have that I is a stationary varifold and

$$\int_{A_{3,4}(p_i)} |\pi_I^\perp[V]|^2 d\mu_I = 0. \qquad (6.5)$$

On the other hand, the $(0, \delta_i)$-symmetry of the I_i tells us that I is 0-symmetric. Combining these tells us that

$$\int_{B_4(0^n)} |\pi_I^\perp[V]|^2 d\mu_I = 0 \,, \tag{6.6}$$

and hence we have that I is $k + 1$-symmetric. Because the convergence $I_i \to I$ is in the varifold sense, this contradicts that the I_i are not $(k + 1, \epsilon)$-symmetric for i sufficiently large, which proves the lemma.

$\square$

6.2. **Best L^2-Subspace Equations.** In order to prove theorem 6.1 we need to identify which subspace minimizes the L^2-energy, and the properties about this subspace that allow us to estimate the distance. We begin in Section 6.2.1 by studying some very general properties of the second directional moments of a general probability measure $\mu \subseteq B_1(p)$. We will then study in Section 6.2.2 a quantitative notion of almost-symmetry for stationary varifolds.

6.2.1. *Second Directional Moments of a Measure.* Let us consider a probability measure $\mu \subseteq B_1(0^n)$, and let

$$x_{cm}^i = x_{cm}^i(\mu) \equiv \int x^i \, d\mu(x) \,, \tag{6.7}$$

be the center of mass. Let us inductively consider the maximum of the second directional moments of μ. More precisely:

Definition 6.3. Let $\lambda_1 = \lambda_1(\mu) \equiv \max_{|v|^2=1} \int |\langle x - x_{cm}, v \rangle|^2 \, d\mu(x)$ and let $v_1 = v_1(\mu)$ with $|v_1| = 1$ be any vector obtaining this maximum. Now let us define inductively the pair (λ_{k+1}, v_{k+1}) from $v_1, \ldots, v_k$ by

$$\lambda_{k+1} = \lambda_{k+1}(\mu) \equiv \max_{|v|^2=1, \langle v, v_i \rangle=0 \ \forall i \le k} \int |\langle x - x_{cm}, v \rangle|^2 \, d\mu(x) \,, \tag{6.8}$$

where v_{k+1} is any vector obtaining this maximum.

Thus $v_1, \ldots, v_n$ defines an orthonormal basis of $\mathbb{R}^n$, ordered so that they maximize the second directional moments of μ. Let us define the subspaces

$$V^k = V^k(\mu) \equiv x_{cm} + \mathrm{span}\{v_1, \ldots, v_k\} \,. \tag{6.9}$$

The following is a simple but important exercise:

Lemma 6.4. *If μ is a probability measure in $B_1(0^n)$, then for each k the functional*

$$\min_{L^k \subseteq \mathbb{R}^n} \int d^2(x, L^k) \, d\mu(x) \,, \tag{6.10}$$

where the min *is taken over all k-dimensional affine subspaces, attains its minimum at V^k. Further, we have that*

$$\min_{L^k \subseteq \mathbb{R}^n} \int d^2(x, L^k) \, d\mu(x) = \int d^2(x, V^k) \, d\mu(x) = \lambda_{k+1}(\mu) + \cdots + \lambda_n(\mu) \,. \tag{6.11}$$

Note that the best affine subspace V^k will necessarily pass through the center of mass x_{cm}.

Now let us record the following Euler-Lagrange formula, which is also an easy computation:

Lemma 6.5. *If μ is a probability measure in $B_1(0^n)$, then we have that $v_1(\mu), \ldots, v_n(\mu)$ satisfy the Euler-Lagrange equations:*

$$\int \langle x - x_{cm}, v_k \rangle (x - x_{cm})^i \, d\mu(x) = \lambda_k v_k^i, \tag{6.12}$$

where

$$\lambda_k = \int |\langle x - x_{cm}, v_k \rangle|^2 \, d\mu(x). \tag{6.13}$$

6.2.2. *Projections on the tangent spaces.* Our goal now is to study an integral varifold I^m in the directions spanned by $v_1(\mu), \ldots, v_n(\mu)$, associated to a probability measure. The main result of this subsection is the following, which holds for a general integral varifold with bounded mean curvature. We recall that by definition

$$W_\alpha(x) \equiv W_{\mathfrak{r}_\alpha, \mathfrak{r}_{\alpha-3}}(x) \equiv \theta_{\mathfrak{r}_{\alpha-3}}(x) - \theta_{\mathfrak{r}_\alpha}(x) \geq 0, \tag{6.14}$$

where $\mathfrak{r}_\alpha = 2^{-\alpha}$.

Proposition 6.6. *Let $I^m \subseteq B_9$ be an integral varifold with mass bound $\mu_I(B_9(p)) \leq \Lambda$. Let μ be a probability measure on $B_1(p)$ with $\lambda_k(\mu), v_k(\mu)$ defined as in Definition 6.3. Then there exists $\delta(n) > 0$ such that*

$$\lambda_k \int_{A_{3,4}(p)} |\pi_I^\perp[v_k]|^2 \, d\mu_I(z) \leq \delta^{-1} \int W_0(x) \, d\mu(x), \tag{6.15}$$

where $\pi^\perp[v](x)$ is the projection of v to $N_x I$, the orthogonal compliment of the tangent space $T_x I$, which exists a.e.

Proof. Note first that there is no harm in assuming that $x_{cm} \equiv 0$. If not we can easily translate to make this so, in which case we still have that $\operatorname{supp}(\mu) \subseteq B_2$. Additionally, we will simplify the technical aspect of the proof by assuming that $M \equiv \mathbb{R}^n$ and $H = 0$. By working in normal coordinates the proof of the general case is no different except up to some mild technical work.

Now let us fix any $z \in A_{3,4}$ in the support of I and such that $N_z I$ is well defined (note that this second set coincides with the support of I up to a set of μ_I measure zero). Observe that

$$\int \langle x, v_k \rangle \, d\mu(x) = \langle x_{cm}, v_k \rangle = 0. \tag{6.16}$$

Then combined with lemma 6.5 we can inner product both sides of (6.12) by $N_z I$, the normal to the tangent space of I at z, to obtain for each k and $A_{3,4}$:

$$\lambda_k \langle N_z I, v_k \rangle = \int \langle x, v_k \rangle \langle N_z I, x \rangle \, d\mu(x) = \int \langle x, v_k \rangle \langle N_z I, x - z \rangle \, d\mu(x). \tag{6.17}$$

We can then estimate

$$\lambda_k^2 |\langle N_z I, v_k \rangle|^2 \leq \lambda_k \int |\langle N_z I, x - z \rangle|^2 \, d\mu(x). \tag{6.18}$$

Integrating with respect to z on both sides we get the estimate

$$\lambda_k \int_{A_{3,4}} |\langle N_z I, v_k \rangle|^2 \, d\mu_I(z) \leq \int \int_{A_{3,4}} |\langle N_z I, x - z \rangle|^2 \, d\mu_I(z) \, d\mu(x). \tag{6.19}$$

Set for convenience $n_x(z) = (z - x)/|z - x|$, i.e., $n_x(z)$ is the radial vector from x to z. Now for $x \in \text{supp}(\mu)$ we can estimate

$$\int_{A_{3,4}} |\langle N_z I, x - z \rangle|^2 \, d\mu_I(z) = \int_{A_{3,4}} |\langle N_z I, n_x(z) \rangle|^2 |x - z|^{-m} |x - z|^{2+m} \, d\mu_I(z)$$

$$\leq C(n) \int_{A_{3,4}} |\langle N_z I, n_x(z) \rangle|^2 |x - z|^{-m} \, d\mu_I(z)$$

$$\leq C(n) \int_{A_{1,8}(x)} |\langle N_z I, n_x(z) \rangle|^2 |x - z|^{-m} \, d\mu_I(z)$$

$$= C(n) W_0(x). \tag{6.20}$$

Applying this to (6.19) we get the estimate

$$\lambda_k \int_{A_{3,4}} |\langle N_z I, v_k \rangle|^2 \, d\mu_I(z) \leq C(n) \int W_0(x) \, d\mu(x). \tag{6.21}$$

$\square$

6.3. **Proof of theorem 6.1.** Let us now combine the results of this Section in order to prove theorem 6.1. Indeed, let μ be a measure in $B_1(p) \subseteq T_p M$. We can assume that μ is a probability measure without any loss of generality, since both sides of our estimate scale. Let $(\lambda_1(\mu), v_1(\mu)), \ldots, (\lambda_n(\mu), v_n(\mu))$ be the directional second moments as defined in Definition 6.3, with V^k the induced subspaces defined as in (6.9). Using lemma 6.4 we have that

$$\min_{L^k \subseteq \mathbb{R}^n} \int d^2(x, L^k) \, d\mu(x) = \int d^2(x, V^k) \, d\mu(x) = \lambda_{k+1}(\mu) + \cdots + \lambda_n(\mu) \leq (n - k) \lambda_{k+1}(\mu), \tag{6.22}$$

where we have used that $\lambda_j \leq \lambda_i$ for $j \geq i$. Therefore our goal is to estimate λ_{k+1}. To begin with, proposition 6.6 tells us that for each j

$$\lambda_j \int_{A_{3,4}(p)} |\langle N_z I, v_j \rangle|^2 \, d\mu_I(z) \leq C \int W_0(x) \, d\mu(x). \tag{6.23}$$

Let us sum the above for all $j \leq k + 1$ in order to obtain

$$\sum_{j=1}^{k+1} \lambda_j \int_{A_{3,4}(p)} |\langle N_z I, v_j \rangle|^2 \, d\mu_I(z) \leq (k + 1) C \int W_0(x) \, d\mu(x), \tag{6.24}$$

or by using that $\lambda_{k+1} \leq \lambda_j$ for $k + 1 \geq j$ we get

$$\lambda_{k+1} \int_{A_{3,4}(p)} |\langle N_z I, V^{k+1}\rangle|^2 \, d\mu_I(z) = \lambda_{k+1} \sum_{j=1}^{k+1} \int_{A_{3,4}(p)} |\langle N_z I, v_j\rangle|^2 \, d\mu_I(z) \leq C \int W_0(x) \, d\mu(x). \tag{6.25}$$

Now we use that $B_9(p)$ is $(0, \delta)$-symmetric, but *not* $(k + 1, \epsilon)$-symmetric in order to apply lemma 6.2 and conclude that

$$\int_{A_{3,4}(p)} |\langle N_z I, V^{k+1}\rangle|^2 \, d\mu_I(z) \geq \delta. \tag{6.26}$$

Combining this with (6.25) we obtain

$$\delta\lambda_{k+1} \leq \lambda_{k+1} \int_{A_{3,4}(p)} |\langle N_z I, V^{k+1}\rangle|^2 \, d\mu_I(z) \leq C \int W_0(x) \, d\mu(x), \tag{6.27}$$

or that

$$\lambda_{k+1} \leq C(n, \Lambda, \epsilon) \int W_0(x) \, d\mu(x). \tag{6.28}$$

Combining this with (6.22) we have therefore proved the theorem. □

7. THE INDUCTIVE COVERING LEMMA

This Section is dedicated to the basic covering lemma needed for the proof of the main theorems of the paper. The covering scheme is similar in nature to the one introduced by the authors in [NVb, NVa] in order to prove structural theorems on critical and singular sets. Specifically, let us consider an integral varifold I^m with bounded mean curvature, then we will build a covering of the quantitative stratification

$$S_{\epsilon,r}^k(I) \cap B_1(p) \subseteq U_r \cup U_+ = U_r \cup \bigcup B_{r_i}(x_i), \tag{7.1}$$

which satisfies several basic properties. To begin with, the set U_+ is a union of balls satisfying $r_i > r \geq 0$, and should satisfy the packing estimate $\omega_k \sum r_i^k \leq C$. Each ball $B_{r_i}(x_i)$ should have the additional property that there is a definite mass drop of I when compared to $B_2(p)$. To describe the set U_r we should distinguish between the case $r > 0$ and $r \equiv 0$. In the case $r > 0$ we will have that $U_r = \bigcup B_r(x_i^r)$ is a union of r-balls and satisfies the Minkowski estimate $\mathrm{Vol}(U_r) \leq Cr^{n-k}$. In the case when $r \equiv 0$ we will have that U_0 is k-rectifiable with the Hausdorff estimate $\lambda^k(U_0) \leq C$. Let us be more precise:

Lemma 7.1 (Covering lemma). *Let $I^m \subseteq B_2$ be an integral varifold satisfying the bounds* (1.14)*, the mean curvature bound* (1.15)*, the mass bound $\mu_I(B_2(p)) \leq \Lambda$, and assume $K + H < \delta(n, \Lambda, \epsilon)$. Let $E = \sup_{x \in B_1(p)} \theta_1(x)$ with $\epsilon > 0$, $r \geq 0$, and $k \in \mathbb{N}$. Then for all $\eta \leq \eta(n, \Lambda, \epsilon)$ and $0 \leq r \leq 1$, there exists a covering $S_{\epsilon,r/100}^k(I) \cap B_1(p) \subseteq U = U_r \cup U_+$ such that*

  (1) *$U_+ = \bigcup B_{r_i}(x_i)$ with $r_i > r$ and $\sum r_i^k \leq C(n, \Lambda, \epsilon)$.*
  (2) *$\sup_{y \in B_{r_i}(x_i)} \theta_{\eta r_i/2}(y) \leq E - \eta$.*
  (3) *If $r > 0$ then $U_r = \bigcup_1^N B_r(x_i^r)$ with $N \leq C(n)r^{-k}$.*

(4) If $r = 0$ then U_0 is k-rectifiable and satisfies $\mathrm{Vol}(B_s\,(U_0)) \leq C(n)s^{n-k}$ for each $s > 0$.
In particular, $\lambda^k(U_0) \leq C(n)$.

Remark 7.1. The assumption $K + H < \delta$ is of little consequence, since given any K this just means focusing the estimates on balls of sufficiently small radius after rescaling.

Remark 7.2. As in the previous section, throughout this section we will assume that the ambient manifold M is actually Euclidean and I has zero mean curvature, that is $K = H = 0$. The generalization to the general case is straightforward, but it would involve minor details that would anyway add another layer of technicalities.

Remark 7.3. Note that, up to enlarging the constants in the covering by some $C(n)$, the conclusions of this lemma clearly hold also for the set $S^k_{\epsilon,r}$, not just for the slightly smaller set $S^k_{\epsilon,r/100}$. We state this lemma with the factor 10^{-2} just for technical reasons.

To prove the result let us begin by outlining the construction of the covering, we will then spend the rest of this section proving the constructed cover has all the desired properties.

Thus let us consider some $\eta > 0$ fixed, and then define the mass scale for $x \in B_1(p)$ by

$$s_x = s_x^{E,\eta} \equiv \inf \left\{ r \leq t \leq 1 : \sup_{B_s(x) \cap S^k_{\epsilon,r/100}} \theta_{\eta s/2}(y) \geq E - \eta \ \text{ for all } \ t \leq s \leq 1 \right\}. \tag{7.2}$$

Note that the mass scale implicitly depends on many constants. If $r = 0$ let us define the set U_0 by

$$U_0 \equiv \{x \in S^k_{\epsilon,0}(I) \cap B_1(p) : s_x = 0\}, \tag{7.3}$$

while if $r > 0$ let us define

$$U_r \equiv \bigcup B_r(x_i^r), \tag{7.4}$$

where

$$\{x_i^r\} \subseteq \{x \in S^k_{\epsilon,r/100}(I) \cap B_1(p) : s_x = r\}, \tag{7.5}$$

is such that $\{B_{r/8}(x_i^r)\}$ are a maximal pairwise disjoint collection of balls centered on $S^k_{\epsilon,r/100}(I) \cap B_1(p)$.

In order to define the covering $U_+ = \{B_{r_i}(x_i)\}$ let us first consider the covering

$$\{x \in S^k_{\epsilon,r/100}(I) \cap B_1(p) : s_x > r\} \subseteq \bigcup_{s_x > r} B_{s_x/10}(x), \tag{7.6}$$

and choose from it a Vitali subcovering and set

$$U_+ \equiv \bigcup_{i \in J} B_{r_i}(x_i), \tag{7.7}$$

where $r_i \equiv s_{x_i}/2$ and $\{B_{r_i/5}(x_i)\}$ are all disjoint. It is clear U_+ satisfies the version of (2) given by

$$(2') \quad \sup_{y \in B_{2r_i}(x_i)} \theta_{\eta r_i/2}(y) \leq E - \eta. \tag{7.8}$$

It is clear from the construction that we have built a covering

$$S^k_{\epsilon,r/100}(I) \cap B_1(p) \subseteq U_r \cup U_+ \,, \tag{7.9}$$

so we need to study the properties of U_r and U_+.

Once Lemma 7.1 is proved, since $\eta(n, K_N, \Lambda, \epsilon)$ is fixed, we can re-cover in a trivial way all balls $B_{r_i}(x_i)$ with smaller balls of radius ηr_i and obtain for this new covering that

(1) $r^{k-n}\text{Vol}(B_r \, U_r) + \omega_k \sum r_i^k \le C(n, K_N, \Lambda, \epsilon) \, s^k$,

(2) $\sup_{y \in B_{r_i}(x_i)} \theta_{r_i}(y) \le E - \eta$,

as desired.

The outline of this Section is as follows. In Sections 7.1 and 7.2 we will deal with estimating the set U_r. In fact, from a technical standpoint U_r is much easier than U_+ to deal with, and for the rectifiability and local λ^k-finiteness of $S^k_{\epsilon,r/100}$ this is the set which is most important. Indeed, for U_r we will be able to almost directly use the rectifiable-Reifenberg of theorems 3.4 and 3.3, at least when combined with an additional induction argument. However, for the global Minkowski estimates and k-dimensional Hausdorff measure bounds on $S^k_{\epsilon,r/100}$ it is crucial to deal with the set U_+ as well. To do this we first prove a variety of technical lemmas in Section 7.3, these will allow us to exchange U_+ for a more manageable collection of balls without losing much content. Then in Section 7.4 we will be able argue as in the U_r case by applying the rectifiable Reifenberg of theorem 3.3, however we will apply it not to U_+ but to a more carefully chosen covering obtained by exploiting the results of Section 7.3.

7.1. **Estimating U_r in lemma 7.1 for $r > 0$.** Let us begin by altering the collection $U_r = \bigcup B_r(x_i^r)$ slightly. Indeed, by definition of U_r, we have that for each ball $B_r(x_i^r)$ there exists a point $x_i' \in B_r(x_i^r)$ such that $\theta_{\eta r}(x_i') \ge E - \eta$. Then we have the covering

$$U_r \subseteq \bigcup_1^N B_{2r}(x_i') \,, \tag{7.10}$$

and let $U_r' \equiv \bigcup_1^{N'} B_{2r}(x_i')$ be a maximal disjoint subset of these balls. Let us first observe that by a standard covering argument we have $N \le C(n)N'$, and thus to estimate U_r it is enough to estimate N'. Indeed, by the maximality of U_r' we have that for every ball center $x_i^r \in U_r$ there exists a ball center $x_j' \in U_r'$ such that $x_i^r \in B_{4r}(x_j')$ and $\theta_{\eta r}(x_j') \ge E - \eta$. Since the balls $\{B_{r/8}(x_i^r)\}$ are disjoint, we have that

$$N \cdot \omega_n \left(\frac{r}{8}\right)^n \le \text{Vol}(U_r) \le \sum_1^{N'} \text{Vol}(B_{4r}(x_j')) \le N' \cdot \omega_n (4r)^n \,, \tag{7.11}$$

which in particular gives us the desired estimate $N \le 32^n N'$.

Now let us proceed to estimate N'. Thus consider the measure associated to U_r' given by

$$\mu = \sum \omega_k r^k \delta_{x_i'} \,. \tag{7.12}$$

Let us consider the sequence of radii $\mathfrak{r}_\alpha = 2^{-\alpha}$. We now wish to prove that for each $x \in B_1$ and all $r \le \mathfrak{r}_\alpha \le 2^{-7}$ that

$$\mu(B_{\mathfrak{r}_\alpha}(x)) \le D(n)\mathfrak{r}_\alpha^k, \tag{7.13}$$

where $D(n)$ is from theorem 3.3. Let us first observe that once (7.13) has been proved then we have the desired estimate on N'. Indeed, by a simple covering argument we obtain that $\mu(B_1(0)) \le C(n) \cdot D(n)$, which in turn implies

$$N'\omega_n r^k = \sum_1^{N'} \omega_k r^k = \mu(B_1) \le C(n) \cdot D(n), \tag{7.14}$$

and this gives the claimed estimate $N' \le C(n)r^{-k}$.

Thus let us now concentrate on proving (7.13). We will prove it by induction on α. Let α_0 be such that $r \le \mathfrak{r}_{\alpha_0} < 2r$. Let us begin by observing that the result clearly holds for $\mathfrak{r}_{\alpha_0}$. Therefore let us now assume we have proved (7.13) for some $\mathfrak{r}_{\alpha+1}$, and then proceed to prove it for $\mathfrak{r}_\alpha$.

Let us first observe a rough estimate in this direction through a covering argument. Namely, let us consider the radii $\mathfrak{r}_{\alpha+1} \le s \le 4\mathfrak{r}_\alpha$. We can cover $B_s(x)$ by a collection of at most $C(n)$ balls $\{B_{\mathfrak{r}_{\alpha+1}}(y_i)\}$, and thus by the inductive assumption we have for all $x \in B_1$ and $s \le 4\mathfrak{r}_\alpha$ that

$$\mu(B_s(x)) \le \sum \mu(B_{\mathfrak{r}_{\alpha+1}}(y_i)) \le D(n)C(n)\mathfrak{r}_\alpha^k \le D'(n)s^k, \tag{7.15}$$

where of course $D'(n) \gg D(n)$.

Now let us consider a ball $B_s(y) \subseteq B_2$ with $s \le 2\mathfrak{r}_\alpha$. If $\mu(B_s(y)) \le \epsilon_n s^k$ then $D_\mu(y, s) \equiv 0$ by definition, while if $s \le r$ then we also have that $D_\mu(y, s) \equiv 0$, since the support of μ in $B_r(x)$ contains at most one point and thus is precisely contained in a k-dimensional subspace. Thus let us consider the case when $s > r$ and $\mu(B_s(y)) > \epsilon_n s^n$. In this case notice by theorem 2.7 that for all the points $y \in \mathrm{supp}\,(\mu) \cap B_s\,(y)$, the ball $B_{32s}\,(y)$ is $(0, \delta)$-symmetric, where $\delta = \delta(\eta|n, \Lambda) \to 0$ as $\eta \to 0$. Thus for $\eta \le \eta(n, \Lambda)$ we can apply theorem 6.1 to see that

$$D_\mu(y, s) \le C(n, \Lambda, \epsilon)s^{-k} \int_{B_s(y)} W_s(z)\, d\mu(z). \tag{7.16}$$

By applying this to all $r \le t \le s$ we have

$$s^{-k} \int_{B_s(x)} D_\mu(y, t)\, d\mu(y) \le C s^{-k} \int_{B_s(x)} t^{-k} \int_{B_t(y)} W_t(z)\, d\mu(z)\, d\mu(y)$$

$$= C s^{-k} t^{-k} \int_{B_{2s}(x)} \mu(B_t(z))W_t(z)\, d\mu(z)$$

$$\le C s^{-k} \int_{B_{2s}(x)} W_t(z)\, d\mu(z), \tag{7.17}$$

where we have used our rough estimate (7.15) in the last line and the fact that if $|z - y| \leq t$ and $|x - y| \leq s$, then $|z - x| \leq t + s \leq 2s$. Let us now consider the case when $t = \mathfrak{r}_\beta \leq s \leq 2\mathfrak{r}_\alpha$. Then we can sum to obtain:

$$
\sum_{\mathfrak{r}_\beta \leq 2s} s^{-k} \int_{B_s(x)} D_\mu(y, \mathfrak{r}_\beta) \, d\mu(y) \leq C \sum_{r \leq \mathfrak{r}_\beta \leq 2s} s^{-k} \int_{B_{2s}(x)} W_{\mathfrak{r}_\beta}(y) \, d\mu(y)
$$

$$
= C s^{-k} \int_{B_{2s}(x)} \sum_{r \leq \mathfrak{r}_\beta \leq 2s} W_{\mathfrak{r}_\beta}(y) \, d\mu(y)
$$

$$
\leq C \, s^{-k} \int_{B_{2s}(x)} \left| \theta_{8s}(y) - \theta_r(y) \right| d\mu(y) \leq C(n, \Lambda, \epsilon)\eta, \tag{7.18}
$$

where in the last line we have used both the rough estimate (7.15) on $\mu(B_s(x))$ and that in the support of μ we have that $\left| \theta_{8s}(y) - \theta_r(y) \right| \leq \eta$ by the construction of U'_r. Now let us choose $\eta \leq \eta(n, \Lambda, \epsilon)$ such that we have

$$
\sum_{\mathfrak{r}_\beta \leq 2s} s^{-k} \int_{B_{2s}(x)} D_\mu(y, \mathfrak{r}_\beta) \, d\mu(y) \leq \delta^2, \tag{7.19}
$$

where δ is taken from theorem 3.3. Since the estimate (7.19) holds for all $B_s \subseteq B_{2\mathfrak{r}_\alpha}(x)$, we can therefore apply theorem 3.3 to conclude the estimate

$$
\mu(B_{\mathfrak{r}_\alpha}(x)) \leq D(n)\mathfrak{r}_\alpha^k. \tag{7.20}
$$

This finishes the proof of (7.13), and hence the proof of estimate of U_r for $r > 0$ in lemma 7.1. $\qquad \square$

7.2. **Estimating U_0 in lemma 7.1.** Let us begin by proving the Minkowski estimates on U_0. Indeed, observe that for any $r > 0$ that $U_0 \subseteq U_r$, and thus we have the estimate

$$
\mathrm{Vol}(B_r\,(U_0)) \leq \mathrm{Vol}(B_r\,(U_r)) \leq \omega_n(2r)^n \cdot N \leq C(n, \Lambda, \epsilon)r^{n-k}, \tag{7.21}
$$

which proves the Minkowski claim. In particular, we have as a consequence the k-dimensional Hausdorff measure estimate

$$
\lambda^k(U_0) \leq C(n, \Lambda, \epsilon). \tag{7.22}
$$

In fact, let us conclude a slightly stronger estimate, since it will be a convenient technical tool in the remainder of the proof. If $B_s(x)$ is any ball with $x \in B_1$ and $s < \frac{1}{2}$, then by applying the same proof to the rescaled ball $B_s(x) \to B_1(0)$, we can obtain the Hausdorff measure estimate

$$
\lambda^k(U_0 \cap B_s(x)) \leq C s^k. \tag{7.23}
$$

To finish the construction we need to see that U_0 is rectifiable. We will in fact apply theorem 3.4 in order to conclude this. To begin with, let $\mu \equiv \lambda^k\big|_{U_0}$ be the k-dimensional Hausdorff measure, restricted to U_0. Let $B_s(y)$ be a ball with $y \in B_1 \cap \mathrm{supp}\,(\mu)$ and $s < \frac{1}{10}$, now we will now argue in a manner similar to Section 7.1. Thus, if $\mu(B_s(y)) \leq \epsilon_n s^k$ then $D_\mu(y, s) \equiv 0$, and otherwise we then have by theorem 6.1 that for $\eta \leq \eta(n, \Lambda)$

$$
D_\mu(y, s) \leq C(n, \Lambda, \epsilon)s^{-k} \int_{B_s(y)} W_s(z) \, d\mu(z). \tag{7.24}
$$

By applying this to all $t \leq s$ we have

$$s^{-k} \int_{B_s(x)} D_\mu(y,t) \, d\mu(y) \leq C s^{-k} \int_{B_s(x)} t^{-k} \int_{B_t(y)} W_t(z) \, d\mu(z) \, d\mu(y)$$

$$= C s^{-k} t^{-k} \int_{B_{2s}(x)} \mu(B_t(z)) W_t(z) \, d\mu(z)$$

$$\leq C s^{-k} \int_{B_{2s}(x)} W_t(z) \, d\mu(z), \qquad (7.25)$$

where we have used our estimate (7.23) in the last line. Let us now consider the case when $t = r_\beta = 2^{-\beta} \leq s$. Then we can sum to obtain:

$$\sum_{r_\beta \leq s} s^{-k} \int_{B_s(x)} D_\mu(y, r_\beta) \, d\mu(y) \leq C \sum_{r_\beta \leq s} s^{-k} \int_{B_s(x)} W_{r_\beta}(y) \, d\mu(y)$$

$$= C s^{-k} \int_{B_s(x)} \sum_{r_\beta \leq s} W_{r_\beta}(y) \, d\mu(y)$$

$$\leq C \, s^{-k} \int_{B_s(x)} \left| \theta_{8s}(y) - \theta_0(y) \right| d\mu(y) \leq C(n, \Lambda, \epsilon) \eta, \qquad (7.26)$$

where we have used two points in the last line. First, we have used our estimate $\mu(B_s(x)) \leq C s^k$. Second, we have used that by the definition of U_0, for each point in the support of μ we have that $\left| \theta_s(y) - \theta_0(y) \right| \leq \eta$. Now let us choose $\eta \leq \eta(n, \Lambda, \epsilon)$ such that we have

$$\sum_{r_\beta \leq s} s^{-k} \int_{B_s(x)} D_\mu(y, r_\beta) \, d\mu(y) \leq \delta^2, \qquad (7.27)$$

where δ is chosen from theorem 3.4. Thus, by applying theorem 3.4 we see that U_0 is rectifiable, which finishes the proof of lemma 7.1 in the context of U_0. $\qquad \square$

7.3. **Technical Constructions for Estimating U_+.** Estimating the set U_+ is in spirit similar to the estimates obtained in the last subsections on U_r and U_0. However, the estimate on U_+ itself is a bit more delicate, and we cannot directly apply the discrete Reifenberg of theorem 3.3 to this set. Instead, we will need to replace U_+ with a different covering at each stage, which will be more adaptable to theorem 3.3. This subsection is dedicated to proving a handful of technical results which are important in the construction of this new covering.

Throughout this subsection we are always working under the assumptions of lemma 7.1. Let us begin with the following point, which is essentially a consequence of the continuity of the mass:

Lemma 7.2. *For each $\eta' > 0$ there exists $R(n, \Lambda, \eta') > 3$ such that with $\eta \leq \eta(n, \Lambda, \eta')$ we have for each $z \in B_{r_i}(x_i)$ with $B_{\eta R r_i}(z) \subseteq B_2(0)$ the estimate*

$$\theta_{\eta R r_i}(z) > E - \eta'. \qquad (7.28)$$

Proof. The proof relies on a straight forward comparison using the definition and monotonicity of the mass. Namely, let $x, y \in B_{1/2}$ with $s < 1$ and let us denote $d \equiv d(x, y)$. Then we have the estimate

$$\theta_s(y) = s^{-k} \int_{B_s(y)} d\mu_I \le s^{-k} \int_{B_{s+d}(x)} d\mu_I = \left(\frac{s}{s+d}\right)^{-k} \theta_{s+d}(x). \tag{7.29}$$

To apply this in our context, let us note for each x_i in our covering, that by our construction of U_+ there must exist $y_i \in B_{r_i}(x_i)$ such that $\theta_{\eta(R-2)r_i}(y_i) \ge \theta_{\eta r_i}(y_i) \ge E - \eta$. Let us now apply (7.29) to obtain

$$\theta_{\eta R r_i}(z) \ge \left(\frac{\eta R r_i}{\eta(R-2)r_i}\right)^{-k} \theta_{\eta(R-2)r_i}(y_i) \ge \left(\frac{R}{R-2}\right)^{-k} (E - \eta). \tag{7.30}$$

If $R = R(n, \Lambda, \eta') > 0$ and $\eta \le \eta(n, \Lambda, \eta')$, then we obtain from this the claimed estimate. $\qquad\square$

In words, the above lemma is telling us that even though we have no reasonable control over the size of $\theta_{r_i}(x_i)$, after we go up a controlled number of scales we can again assume that the mass is again close to E.

In the last lemma the proof was based on continuity estimates on the mass θ. In the next lemma we wish to show an improved version of this continuity under appearance of symmetry. Precisely:

Lemma 7.3 (Improved Continuity of θ). *Let $I^m \subseteq B_2$ be an integral varifold satisfying the bounds (1.14), the mean curvature bound (1.15), and the mass bound $\mu_I(B_4(p)) \le \Lambda$. Then for $0 < \tau, \beta, \gamma < 1$ there exists $\delta(n, \Lambda, \gamma, \beta, \tau) > 0$ such that if there exists $x_1, \ldots, x_k \in I \cap B_1(p)$ which are τ-independent at x_0 with $|\theta_3(x_j) - \theta_\delta(x_j)| < \delta$, and if V^k is the k-dimensional affine subspace $x_0 + \mathrm{span}\{x_1 - x_0, \ldots, x_k - x_0\}$, then for all $x, y \in I \cap B_1(p) \cap B_\delta(V^k)$ and $10^{-4}\beta \le s \le 1$ we have that $|\theta_s(x) - \theta_s(y)| < \gamma$.*

Proof. The proof is by contradiction. Thus, imagine no such δ exists. Then there exists a sequence of integral varifolds I_i on $B_4(p_i)$ satisfying $I_i(B_4(p_i)) \le \Lambda$ and such that

(1) there exists $x_{i,1}, \ldots, x_{i,k} \in B_1(p_i)$ which are τ-independent at $x_{i,0}$ with $\left|\theta_{\delta_i}^{I_i}(x_{i,j}) - \theta_3^{I_i}(x_{i,j})\right| < \delta_i \to 0$,

however we have that there exists $x_i, y_i \in I \cap B_1(p) \cap B_{\delta_i}(V_i^k)$ and $10^{-4}\beta \le s_i \le 1$ such that $\left|\theta_{s_i}^{I_i}(x_i) - \theta_{s_i}^{I_i}(y_i)\right| \ge \gamma$. Since $\partial I \cap B_2(p) = \emptyset$, the masses are uniformly bounded and the bound on the mean curvature δ_i converges to zero, we may apply Allard's compactness theorem and obtain a converging subsequence (where for convenience we will not change indexes) $I_i \to I$ as well as the collection

$$\begin{aligned} &x_{i,j} \to x_j \in B_1(0^n), \\ &x_i \to x, \, y_i \to y \in V = x_0 + \mathrm{span}\{x_1 - x_0, \ldots, x_k - x_0\}, \\ &s_i \to s. \end{aligned} \tag{7.31}$$

In particular, we have in this limit that

$$\begin{aligned} &\left|\theta_0^I(x_j) - \theta_3^I(x_j)\right| = 0, \\ &\left|\theta_s^I(x) - \theta_s^I(y)\right| \ge \gamma. \end{aligned} \tag{7.32}$$

However, we have by theorem 2.7 and the standard cone splitting of theorem 2.9 we have that I is k-symmetric with respect to the k-plane V on B_1. In particular, I is invariant under translation by elements of V. However, this is a contradiction to (7.32), and therefore we have proved the result.

$\square$

Now the first goal is to partition U_+ into a finite collection, each of which will have a few more manageable properties than U_+ itself. More precisely:

Lemma 7.4. *For each $R \geq 5$ there exists $N(n, R) > 1$ such that we can break up U_+ as a union*

$$U_+ = \bigcup_{a=1}^{N} U_+^a = \bigcup_{a=1}^{N} \bigcup_{i \in J^a} B_{r_i}(x_i) , \tag{7.33}$$

such that each U_+^a has the following property: if $i \in J^a$, then for any other $j \in J^a$ we have that if $x_j \in B_{Rr_i}(x_i)$, then $r_j < R^{-2}r_i$.

Proof. Let us recall that the balls in the collection $\{B_{r_i/5}(x_i)\}$ are pairwise disjoint. In particular, given $R \geq 5$ if we fix a ball $B_{r_i}(x_i)$ then by the usual covering arguments there can be at most $N(n, R)$ ball centers $\{x_j\}_1^N \subset U_+ \cap B_{R^3 r_i}(x_i)$ with the property that $r_j \geq r_i$. Indeed, if $\{x_j\}_1^N$ is such a collection of balls then we get

$$\omega_n (2R^3 r_i)^n = \mathrm{Vol}(B_{2R^3 r_i}(x_i)) \geq \sum_1^N \mathrm{Vol}(B_{r_j/5}(x_j)) \geq N\omega_n (r_i/5)^n , \tag{7.34}$$

which by rearranging gives the estimate $N \leq N(n, R)$ as claimed.

Now we wish to build our decomposition $U_+ = \bigcup_1^N U_+^a$, where N is from the first paragraph. We shall do this inductively, with the property that at each step of the inductive construction we will have that the sets U_+^a will satisfy the desired property. In particular, for every a and $i \in J^a$, if $j \in J^a$ is such that $x_j \in B_{Rr_i}(x_i)$, then $r_j < R^{-2}r_i$.

Begin by letting each J^a be empty. We are going to sort the points $\{x_i\}_{i \in J}$ into the sets J^a one at a time. At each step let $i \in J \setminus \bigcup_{a=1}^{N} J^a$ be an index such that $r_i = \max r_j$, where the max is taken over all indexes in $J \setminus \bigcup_{a=1}^{N} J^a$, i.e., over all indexes which haven't been sorted out yet. Now let us consider the collection of ball centers $\{y_j\}_{j \in J'}$ such that $J' \subset J$, $x_i \in B_{Rr_j}(y_j)$ and $r_i \leq r_j \leq R^2 r_i$. Note that, by construction, either y_j has already been sorted out in some J^a, or $r_j = r_i$. Now evidently $y_j \in B_{R^3 r_i}(x_i)$ for all $j \in J'$ and so by the first paragraph the cardinality of J' is at most $N(n, R)$. In particular, there must be some J^a such that $J^a \cap J' = \emptyset$. Let us assign the index i to this piece of the decomposition, so that $i \in J^a$ and $x_i \in U_a^+$. Clearly the decomposition $\bigcup U_+^a$ still satisfies the inductive hypothesis after the addition of this point, and so this finishes the inductive step of construction. Since at each stage we have chosen x_i to have the maximum radius, this process will continue indefinitely to give the desired decomposition of U_+.

$\square$

Now with a decomposition fixed, let us consider for each $1 \leq a \leq N$ the measures

$$\mu^a \equiv \sum_{i \in J^a} \omega_k r_i^k \delta_{x_i} . \tag{7.35}$$

The following is a crucial point in our construction. It tells us that each ball $B_{10r_i}(x_i)$ either has small μ^a-volume, or the point x_i must have large mass at scale r_i. Precisely:

Lemma 7.5. *Let $\eta', \beta, D > 0$ be fixed. There exists $R = R(n, \Lambda, \eta', D, \beta, \epsilon) > 0$ such that if we consider the decomposition (7.33), and if*

   (1) $\eta \leq \eta(n, \Lambda, \eta', \beta, \epsilon)$, $r_i < 10^{-2}$,
   (2) *we have $\mu^a(B_{10r_i}(x_i)) \geq 2\omega_k r_i^k$,*
   (3) *for all ball centers $y_j \in A_{r_i/5, 10r_i}(x_i) \cap U_+^a$ and for $s = 10^{-2n}D^{-1}\omega_n r_i$, we have that $\mu^a(B_s(y_j)) \leq Ds^k$,*

then we have that $\theta_{\beta r_i/10}(x_i) \geq E - \eta'$.

Proof. Let us begin by choosing η'' sufficiently small, which will be fixed later in the proof, and let us also define $\tau \equiv 10^{-2n}D^{-1}\omega_n$. Let $\delta(n, \Lambda, \beta, \eta', \tau)$ be from lemma 7.3 so that the conclusions hold with $10^{-1}\eta'$. Now throughout we will assume $\eta'' < \delta$. We will also choose $R = R(n, \Lambda, \eta'', D, \epsilon) > \max\{\tau^{-1}, \delta^{-1}, \delta'^{-1}\}$ so that lemma 7.2 is satisfied with η''.

Since it will be useful later, let us first observe that $r_i \geq R^2 r$. Indeed, if not then for each ball center $y_j \in A_{r_i/5, 10r_i}(x_i) \cap U_+^a$ we would have $r \leq r_j < R^{-2}r_i < r$. This tells us that there can be no ball centers in $A_{r_i/5, 10r_i}(x_i) \cap U_+^a$. However, by our volume assumption we have that

$$\mu^a(A_{r_i/5, 10r_i}(x_i)) = \mu^a(B_{10r_i}(x_i)) - \mu^a(B_{r_i/5}(x_i)) \geq 2\omega_k r_i^k - \omega_k r_i^k = \omega_k r_i^k, \tag{7.36}$$

which contradicts this. Therefore we must have that $r_i \geq R^2 r$.

Now our first real claim is that under the assumptions of the lemma there exists ball centers $y_0, \ldots, y_k \in U_+^a \cap A_{r_i/5, 10r_i}(x_i)$ which are τr_i-independent in the sense of Definition 2.8. Indeed, assume this is not the case, then we can find a $k - 1$-plane V^{k-1} such that

$$\{y_i\}_{i \in J^a} \cap A_{r_i/5, 10r_i}(x_i) \subseteq B_{\tau r_i}\left(V^{k-1}\right). \tag{7.37}$$

In particular, by covering $\{y_i\}_{i \in J^a} \cap B_{\tau r_i}(V) \cap B_{10r_i}(x_i)$ by $C(n)\tau^{1-k} < 10^{2n}\tau^{1-k}$ balls of radius τr_i centered on $\{y_i\}_{i \in J^a}$, and using our assumption that $\mu_+^a(B_{\tau r_i}(y_i)) \leq D\tau^k r_i^k$, we are then able to conclude the estimate

$$\mu^a\left(A_{r_i/5, 10r_i}(x_i)\right) \leq \mu^a\left(B_{\tau r_i}(V) \cap B_{10r_i}(x_i)\right) \leq 10^n D\tau r_i^k < \omega_n r_i^k. \tag{7.38}$$

On the other hand, our volume assumption guarantees that

$$\mu^a(A_{r_i/5, 10r_i}(x_i)) = \mu^a(B_{10r_i}(x_i)) - \mu^a(B_{r_i/5}(x_i)) \geq 2\omega_k r_i^k - \omega_k r_i^k \geq \omega_k r_i^k, \tag{7.39}$$

which leads to a contradiction. Therefore there must exist $k + 1$ ball centers $y_0, \ldots, y_k \in A_{r_i/5, 10r_i}(x_i)$ which are τr_i-independent points, as claimed.

Let us now remark on the main consequences of the existence of these $k + 1$ points. Note first that for each y_j we have that $\theta_{\eta R^{-1} r_i}(y_j) > E - \eta''$, since by the construction of U_+^a we have that $r_j \leq R^{-2}r_i$, and therefore we can apply lemma 7.2. Thus we have $k + 1$ points in $B_{10r_i}(x_i)$ which are τr_i-independent, and whose mass densities are η''-pinched. To exploit this, let us first apply lemma 7.3 in order to conclude that

for each $x \in B_{\delta r_i}(V)$ we have

$$\theta_{\beta r_i/10}(x) \geq \theta_{\beta r_i/10}(y_j) - |\theta_{\beta r_i/10}(x) - \theta_{\beta r_i/10}(y_j)| \geq E - \eta'' - 10^{-1}\eta' > E - \eta'. \tag{7.40}$$

In particular, if we assume that x_i is such that $\theta_{\beta r_i/10}(x_i) < E - \eta'$, then we must have that $r_i^{-1}d(x_i, V) \geq \delta = \delta(n, \Lambda, \eta')$.

Therefore, let us now assume $\theta_{\beta r_i/10}(x_i) < E - \eta'$, and thus to prove the lemma we wish to find a contradiction. To accomplish this notice that we have our $k + 1$ points $y_0, \ldots, y_k \in B_{10r_i}$ which are τr_i-independent and for which $|\theta_{20r_i}(y_j) - \theta_{\eta R^{-1}r_i}(y_j)| < \eta''$. Therefore by applying the cone splitting of theorem 2.9 we have for each $\epsilon' > 0$ that if $\eta'' \leq \eta''(n, \Lambda, \epsilon')$ then $B_{10r_i}(x_i)$ is (k, ϵ')-symmetric with respect to the k-plane V^k. However, since $d(x_i, V) > \delta r_i$, we have by theorem 2.10 that if $\epsilon' \leq \epsilon'(n, \Lambda, \epsilon)$ then there exists some $\tau' = \tau'(n, \Lambda, \epsilon)$ such that $B_{\tau'r_i}(x_i)$ is $(k + 1, \epsilon)$-symmetric. However, we can assume after a further increase that $R = R(n, \Lambda, D, \epsilon) > 4\tau'^{-1}$, and thus we have that $\tau'r_i > 4R^{-1}r_i > 4r$. This contradicts that $x_i \in S^k_{\epsilon, r/100}$, and thus we have contradicted that $\theta_{\beta r_i/10}(x_i) < E - \eta'$, which proves the lemma. $\square$

7.4. Estimating U_+ in lemma 7.1.

Now we proceed to finish the proof of lemma 7.1 by estimating the set U_+. First let us pick $D' = D'(n) \equiv 2^{16n}D(n)$, where $D(n)$ is from theorem 3.3. For some η', β fixed depending only on n, Λ, ϵ, we can choose R as in lemma 7.5. It is then enough to estimate each of the sets U_+^a, as there are at most $N = N(n, \Lambda, \epsilon, \eta', \beta)$ pieces to the decomposition. Thus we will fix a set U_+^a and focus on estimating the content of this set. Let us begin by observing that if $r > 0$ then we have the lower bound $r_i \geq r$. Otherwise, let us fix any $r > 0$ and restrict ourselves to the collection of balls in U_+^a with $r_i \geq r$. The estimates we will prove in the end will be independent of r, and thus by letting $r \to 0$ we will obtain estimates on all of U_+^a.

Now let us make the precise statement we will prove in this subsection. Namely, consider any of ball centers $\{x_i\}_{i \in J^a}$ and any radius $2^{-4}r_i \leq \mathfrak{r}_\alpha \leq 2^{-6}$, where $\mathfrak{r}_\alpha = 2^{-\alpha}$. Then we will show that

$$\mu^a(B_{\mathfrak{r}_\alpha}(x_i)) \leq 2^{8n}D(n)\mathfrak{r}_\alpha^k. \tag{7.41}$$

Let us observe that once we have proved (7.41) then we have finished the proof of the Covering lemma, as a simple covering argument gives us the the estimate

$$\sum r_i^k = \mu^a(B_2(x_i)) \leq C(n). \tag{7.42}$$

We prove (7.41) inductively on α. To begin notice that for each x_i if α is the largest integer such that $2^{-4}r_i \leq \mathfrak{r}_\alpha$, then the statement clearly holds, by the definition of the measure μ^a. In fact, we can go further than this. For each x_i, let $r_i' \in [r_i/5, 10r_i]$ be the largest radius such that for all $r_i/5 \leq s < r_i'$ we have

$$\mu^a(B_s(x_i)) \leq 2\omega_k(5s)^k. \tag{7.43}$$

In particular, we certainly have the much weaker estimate $\mu^a(B_s(x_i)) \leq 2^{8n}D(n)s^k$, and hence (7.41) is also satisfied for all $2^{-4}r_i \leq \mathfrak{r}_\alpha \leq r_i'$. Notice that we then also have the estimate

$$\omega_k r_i^k \leq \mu^a(B_{r_i/8}(x_i)) \leq \mu^a(B_{r_i'}(x_i)) \leq 2\omega_k(5r_i')^k. \tag{7.44}$$

Now let us focus on proving the inductive step of (7.41). Namely, assume α is such that for all x_i with $2^{-4}r_i \leq \mathfrak{r}_{\alpha+1} < 2$ we have that (7.41) holds. Then we want to prove that the same estimate holds for $\mathfrak{r}_\alpha$. Let us begin by seeing that a weak version of (7.41) holds. Namely, for any index $i \in J^a$ and any radius $\mathfrak{r}_\alpha \leq s \leq 8\mathfrak{r}_\alpha$, by covering $B_s(x_i)$ by at most 2^{8n} balls $\{B_{\mathfrak{r}_{\alpha+1}}(y_j)\}$ of radius $\mathfrak{r}_{\alpha+1}$ we have the weak estimate

$$\mu^a(B_s(x_i)) \leq \sum \mu^a(B_{\mathfrak{r}_{\alpha+1}}(y_j)) \leq D'(n)s^k, \tag{7.45}$$

where of course $D'(n) >> 2^{8n}D(n)$.

To improve on this, let us fix an $i \in J^a$ and the relative ball center $x_i \in U_+^a$ with $2^{-4}r_i \leq \mathfrak{r}_\alpha$. Now let $\{x_j\}_{j \in J} = \{x_j\}_{j \in J^a} \cap B_{\mathfrak{r}_\alpha}(x_i)$ be the collection of ball centers in $B_{\mathfrak{r}_\alpha}(x_i)$. Notice first that if $r_j' > 2\mathfrak{r}_\alpha$ for any of the ball centers $\{x_j\}$, then we can estimate

$$\mu^a(B_{\mathfrak{r}_\alpha}(x_i)) \leq \mu^a(B_{2\mathfrak{r}_\alpha}(x_j)) \leq 2\omega_k(10\mathfrak{r}_\alpha)^k \leq 2^{8n}D(n)\mathfrak{r}_\alpha^k, \tag{7.46}$$

so that we may fairly assume $r_j' \leq 2\mathfrak{r}_\alpha$ for every $j \in J$. Now for each ball $B_{r_j'}(x_j)$ let us define a new ball $B_{\bar{r}_j}(y_j)$ which is roughly equivalent, but will have some additional useful properties needed to apply the discrete Reifenberg. Namely, for a given ball $B_{r_j'}(x_j)$, let us consider the two options $r_j' < 10r_j$ or $r_j' = 10r_j$. If $r_j' < 10r_j$, then we let $y_j \equiv x_j$ with $\bar{r}_j \equiv r_j'$. In this case we must have that $\mu^a(B_s(x_j)) > 2\omega_k(5s)^k$ for some s arbitrarily close to $\bar{r}_j$, and thus we can apply lemma 7.5 in order to conclude that $\theta_{\beta\bar{r}_j/2}(y_j) \geq E - \eta'$. In the case when $r_j' = 10r_j$ is maximal, let $y_j \in B_{r_j}(x_j)$ be a point such that $\theta_{\eta r_j}(y_j) = E - \eta$, such a point exists by the definition of r_j, and let $\bar{r}_j \equiv 9r_j$. In either case we then have the estimates

$$\theta_{\bar{r}_j/8}(y_j) \geq E - \eta',$$
$$\omega_k 10^{-k}\bar{r}_j^k \leq \mu^a(B_{\bar{r}_j/8}(y_j)) \leq \mu^a(B_{\bar{r}_j}(y_j)) \leq 2\omega_k(5r_j')^k \leq 10^k\omega_k\bar{r}_i^k,$$
$$x_j \in B_{\bar{r}_j/5}(y_j). \tag{7.47}$$

Since $\text{supp}(\mu_a) \cap B_{\mathfrak{r}_\alpha}(x_i) \subseteq \bigcup B_{\bar{r}_j/5}(y_j)$, we can choose a Vitali subcovering of the support such that

$$\text{supp}(\mu_a) \cap B_{\mathfrak{r}_\alpha}(x_i) \subseteq \bigcup B_{\bar{r}_j}(y_j), \tag{7.48}$$

such that $\{B_{\bar{r}_j/5}(y_j)\}$ are disjoint, where we are now being loose on notation and referring to $\{y_j\}_{j \in \bar{J}}$ as the ball centers from this subcovering. Let us now consider the measure

$$\mu' \equiv \sum_{j \in \bar{J}} \omega_k\left(\frac{\bar{r}_j}{10}\right)^k \delta_{y_j}. \tag{7.49}$$

That is, we have associated to the disjoint collection $\{B_{\bar{r}_j/10}(y_j)\}$ the natural measure. Our goal is to prove that

$$\mu'(B_{\mathfrak{r}_\alpha}(x_i)) \leq D(n)\mathfrak{r}_\alpha^k. \tag{7.50}$$

Let us observe that if we prove (7.50) then we are done. Indeed, using (7.47) we can estimate

$$\mu^a(B_{\mathfrak{r}_\alpha}(x_i)) \le \sum \mu^a(B_{\bar{r}_j}(y_j)) \le 10^k \omega_k \sum \bar{r}_j^k = 10^{2k} \mu'(B_{\mathfrak{r}_\alpha}(x_i)) \le 2^{8n} D(n) \mathfrak{r}_\alpha^k \,, \tag{7.51}$$

which would finish the proof of (7.41) and therefore the lemma.

Thus let us concentrate on proving (7.50). We will want to apply the discrete Reifenberg in this case to the measure μ'. Let us begin by proving a weak version of (7.50). Namely, for any ball center y_j from our subcovering and radius $\bar{r}_j < s \le 4\mathfrak{r}_\alpha$ let us consider the set $\{z_\ell\} = \{y_t\}_{t \in \bar{J}} \cap B_s(y_j)$ of ball centers inside $B_s(y_j)$. Since the balls $\{B_{\bar{r}_k/5}(z_\ell)\}$ are disjoint we have that $\bar{r}_k \le 8s$. Using this, (7.45), and (7.47) we can estimate

$$\mu'(B_s(y_j)) = \sum_{z_\ell \in B_s(y_j)} \omega_k 10^{-k} \bar{r}_k^k \le C(n) \sum_{z_\ell \in B_s(y_j)} \mu^a(B_{\bar{r}_k/8}(z_\ell)) \le C(n) \mu^a(B_{2s}(y_j)) \le C(n) s^k \,, \tag{7.52}$$

where of course $C(n) >> 2^{8n} D(n)$.

Now let us finish the proof of (7.50). Thus let us pick a ball center $y_j \in B_{\mathfrak{r}_\alpha}(x_i)$ and a radius $s < 4\mathfrak{r}_\alpha$. If $\mu'(B_s(y_j)) \le \epsilon_n s^k$ then $D_{\mu'}(y_j, s) \equiv 0$ by definition, and if $s \le \bar{r}_j/5$ then $D_{\mu'}(y_j, s) \equiv 0$, since the support of μ' in $B_{\bar{r}_j/5}(y_j)$ contains at most one point and thus is precisely contained in a k-dimensional subspace. In the case when $s > \bar{r}_i/5$ and $\mu(B_s(y_j)) > \epsilon_n s^k$, we want to apply the estimates in theorem 6.1. In order to do so, we first remark that by picking $\beta \le \beta_0(n, \Lambda, \beta')$ sufficiently small, since $\theta_1(y_j) - \theta_{\beta \bar{r}_j}(y_j) \le \eta'$, we can apply theorem 2.7 in order to prove that $B_{4s}(y)$ is $(0, \beta')$-symmetric, with $\beta' = \beta'(n, \Lambda, \epsilon)$, and in turn we obtain by theorem 6.1 that

$$D_{\mu'}(y_j, s) \le C(n, \Lambda, \epsilon) s^{-k} \int_{B_s(y)} W_s(z) \, d\mu'(z) \,. \tag{7.53}$$

Note that $B_s(y_j)$ is not $(k+1, \epsilon)$-symmetric since $y_j \in S^k_{\epsilon, r/100}$ and $s \ge \bar{r}_j/5 \ge r/100$.

By applying this to all $\bar{r}_j/5 < t \le s$ we can estimate

$$s^{-k} \int_{B_s(x)} D_{\mu'}(y, t) \, d\mu'(y) \le C s^{-k} \int_{B_s(x)} t^{-k} \int_{B_t(y)} W_t(z) \, d\mu'(z) \, d\mu'(y)$$

$$= C s^{-k} t^{-k} \int_{B_{2s}(x)} \mu'(B_t(z)) W_t(z) \, d\mu'(z)$$

$$\le C s^{-k} \int_{B_{2s}(x)} W_t(z) \, d\mu'(z) \,, \tag{7.54}$$

where we have used our estimate on $\mu'(B_t(y))$ from (7.52) in the last line. Let us now consider the case when $\bar{r}_j/5 < t = \mathfrak{r}_\beta \le s \le 2\mathfrak{r}_\alpha$. Then we can sum to obtain:

$$\sum_{\mathfrak{r}_\beta \le s} s^{-k} \int_{B_s(x)} D_{\mu'}(y, \mathfrak{r}_\beta) \, d\mu'(y) \le C \sum_{r'_y \le \mathfrak{r}_\beta \le s} s^{-k} \int_{B_{2s}(x)} W_{\mathfrak{r}_\beta}(y) \, d\mu'(y)$$

$$= C s^{-k} \int_{B_{2s}(x)} \sum_{\bar{r}_y \le \mathfrak{r}_\beta \le s} W_{\mathfrak{r}_\beta}(y) \, d\mu'(y)$$

$$\le C s^{-k} \int_{B_{2s}(x)} \left| \theta_{8s}(y) - \theta_{\bar{r}_y}(y) \right| d\mu'(y) \le C(n, \Lambda, \epsilon) \eta' \,, \tag{7.55}$$

where we are using (7.47) in the last line in order to see that $\left|\theta_s(y) - \theta_{\bar{r}_y}(y)\right| \leq \eta'$. Now let us choose $\eta' \leq \eta'(n, \Lambda, \epsilon)$ such that

$$\sum_{\mathfrak{r}_\beta \leq s} s^{-k} \int_{B_{2s}(x)} D_{\mu'}(y, \mathfrak{r}_\beta) \, d\mu'(y) \leq \delta^2 \,, \tag{7.56}$$

where δ is chosen from the discrete rectifiable-Reifenberg of theorem 3.3. Since the estimate (7.56) holds for all $B_{2s} \subseteq B_{2\mathfrak{r}_\alpha}(x)$, we can therefore apply theorem 3.3 to conclude the estimate

$$\mu(B_{\mathfrak{r}_\alpha}(x_i')) \leq D(n)\mathfrak{r}_\alpha^k \,. \tag{7.57}$$

This finishes the proof of (7.47) , and hence the proof of lemma 7.1.   □

8. Proof of Main theorem's for Integral Varifolds with Bounded Mean Curvature

In this section we prove the main theorems of the paper concerning integral varifolds with bounded mean curvature. With the tools of Sections 3, 6, and 7 developed, we will at this stage mainly be applying the covering of lemma 7.1 iteratively to arrive at the estimates. When this is done carefully, we can combine the covering lemma with the cone splitting in order to check that for k-a.e. $x \in S_\epsilon^k$ there exists a unique k-dimensional subspace $V^k \subseteq T_x M$ such that every tangent cone of I at x is k-symmetric with respect to V.

Remark 8.1. For the proofs of the theorems of this section let us make the following remark. For any $\delta > 0$ we can cover $B_1(p)$ by a collection of balls

$$B_1(p) \subseteq \bigcup_1^N B_{M^{-1}\delta}(p_i) \,, \tag{8.1}$$

where $N \leq C(n)M^n\delta^{-n}$. Thus if $\delta = \delta(n, \Lambda, \epsilon)$, $M = K + H$, and we can analyze each such ball, we can conclude from this estimates on all of $B_1(p)$. In particular, by rescaling $B_{M^{-1}\delta}(p_i) \to B_1(p_i)$, we see that we can assume in our analysis that $K + H < \delta$ without any loss of generality. We shall do this throughout this section.

8.1. **Proof of theorem 1.3.** Let $I^m \subseteq B_2$ be an integral varifold satisfying the bounds (1.14), the mean curvature bound (1.15), and the mass bound $\mu_I(B_2(p)) \leq \Lambda$. With $\epsilon, r > 0$ fixed, let us choose $\eta(n, \Lambda, \epsilon) > 0$ and $\delta(n, \Lambda, \epsilon) > 0$ as in lemma 7.1. By Remark 8.1, we see that we can assume that $K + H < \delta$, which we will do for the remainder of the proof.

Now let us begin by first considering an arbitrary ball $B_s(x)$ with $x \in B_1(p)$ and $r < s \leq 1$, potentially quite small. We will use lemma 7.1 in order to build a special covering of $B_s(x)$. Let us define

$$E \equiv \sup_{y \in B_s(x)} \theta_s(y) \,, \tag{8.2}$$

and thus if we apply lemma 7.1 with $\eta(n, K_N, \Lambda, \epsilon)$ fixed to $B_s(x)$, then we can build a covering

$$S_{\epsilon,r}^k \cap B_s(x) \subseteq U_r \cup U_+ = \bigcup B_r(x_i^r) \cup \bigcup B_{r_i}(x_i) \,, \tag{8.3}$$

with $r_i > r$. Let us recall that this covering satisfies the following:

(a) $r^{k-n}\text{Vol}(B_r\, U_r) + \omega_k \sum r_i^k \leq C(n, \Lambda, \epsilon)\, s^k$.

(b) $\sup_{y \in B_{r_i}(x_i)} \theta_{\eta r_i}(y) \leq E - \eta$.

As remarked above, since $\eta(n, K_N, \Lambda, \epsilon)$ is fixed, we can re-cover in a trivial way all balls $B_{r_i}(x_i)$ with smaller balls of radius ηr_i and obtain for this new covering that

(a') $r^{k-n}\text{Vol}(B_r\, U_r) + \omega_k \sum r_i^k \leq C(n, K_N, \Lambda, \epsilon)\, s^k$.

(b') $\sup_{y \in B_{r_i}(x_i)} \theta_{r_i}(y) \leq E - \eta$.

Now that we have built our required covering on an arbitrary ball $B_s(x)$, let us use this iteratively to build our final covering of $S_{\epsilon,r}^k(I)$. First, let us apply it to $B_1(p)$ is order to construct a covering

$$S_{\epsilon,r}^k(I) \subseteq U_r^1 \cup U_+^1 = \bigcup B_r\left(x_i^{r,1}\right) \cup \bigcup B_{r_i^1}\left(x_i^1\right), \tag{8.4}$$

such that

$$r^{k-n}\text{Vol}\left(B_r\left(U_r^1\right)\right) + \omega_k \sum \left(r_i^1\right)^k \leq C(n, K_N, \Lambda, \epsilon), \tag{8.5}$$

and with

$$\sup_{y \in B_{r_i^1}(x_i^1)} \theta_{r_i^1}(y) \leq \Lambda - \eta. \tag{8.6}$$

Now let us tackle the following claim, which is our main iterative step in the proof:

Claim: For each ℓ there exists a constant $C_\ell(\ell, n, \Lambda, \epsilon)$ (recall that $K + H \leq \delta$) and a covering

$$S_{\epsilon,r}^k(I) \subseteq U_r^\ell \cup U_+^\ell = \bigcup B_r\left(x_i^{r,\ell}\right) \cup \bigcup B_{r_i^\ell}\left(x_i^\ell\right), \tag{8.7}$$

with $r_i^\ell > r$, such that the following two properties hold:

$$r^{k-n}\text{Vol}\left(B_r\left(U_r^\ell\right)\right) + \omega_k \sum \left(r_i^\ell\right)^k \leq C_\ell(\ell, n, \Lambda, \epsilon),$$
$$\sup_{y \in B_{r_i^\ell}(x_i^\ell)} \theta_{r_i^\ell}(y) \leq \Lambda - \ell \cdot \eta. \tag{8.8}$$

To prove the claim let us first observe that we have shown this holds for $\ell = 1$. Thus let us assume we have proved the claim for some ℓ, and determine from this how to build the covering for $\ell + 1$ with some constant $C_{\ell+1}(\ell + 1, n, \Lambda, \epsilon)$, which we will estimate explicitly.

Thus with our covering determined at stage ℓ, let us apply the covering of (8.3) to each ball $\left\{B_{r_i^\ell}\left(x_i^\ell\right)\right\}$ in order to obtain a covering

$$S_{\epsilon,r}^k \cap B_{r_i^\ell}(x_i^\ell) \subseteq U_{i,r} \cup U_{i,+} = \bigcup_j B_r\left(x_{i,j}^r\right) \cup \bigcup_j B_{r_{i,j}}\left(x_{i,j}\right), \tag{8.9}$$

such that

$$r^{k-n} \text{Vol}\left(B_r\left(U_{i,r}\right)\right) + \omega_k \sum_j (r_{i,j})^k \leq C(n, \Lambda, \epsilon)(r_i^\ell)^k,$$

$$\sup_{y \in B_{r_{i,j}}(x_{i,j})} \theta_{r_{i,j}}(y) \leq \Lambda - (\ell + 1)\eta. \tag{8.10}$$

Let us consider the sets

$$U_r^{\ell+1} \equiv U_r^\ell \cup \bigcup_i U_{i,r},$$

$$U_+^{\ell+1} \equiv \bigcup_{i,j} B_{r_{i,j}}(x_{i,j}). \tag{8.11}$$

Notice that the second property of (8.8) holds for $\ell + 1$ by the construction, hence we are left analyzing the volume estimate of the first property. Indeed, for this we combine our inductive hypothesis (8.8) for U^ℓ and (8.10) in order to estimate

$$r^{k-n} \text{Vol}\left(B_r\left(U_r^{\ell+1}\right)\right) + \omega_k \sum_{i,j} (r_{i,j})^k \leq r^{k-n} \text{Vol}\left(B_r\left(U_r^\ell\right)\right) + \sum_i \left(r^{k-n} \text{Vol}\left(B_r\left(U_{i,r}\right)\right) + \omega_k \sum_j (r_{i,j})^k\right)$$

$$\leq C_\ell + C \sum_i (r_i^\ell)^k$$

$$\leq C(n, \Lambda, \epsilon) \cdot C_\ell(\ell, n, \Lambda, \epsilon)$$

$$\equiv C_{\ell+1}. \tag{8.12}$$

Hence, we have proved that if the claim holds for some ℓ then the claim holds for $\ell + 1$. Since we have already shown the claim holds for $\ell = 1$, we have therefore proved the claim for all ℓ.

Now we can finish the proof. Indeed, let us take $\ell = \lceil \eta^{-1} \Lambda \rceil = \ell(\eta, \Lambda)$. Then if we apply the Claim to such an ℓ, we must have by the second property of (8.8) that

$$U_+^\ell \equiv \emptyset, \tag{8.13}$$

and therefore we have a covering

$$S_{\epsilon,r}^k \subseteq U_r^\ell = \bigcup_i B_r(x_i). \tag{8.14}$$

But in this case we have by (8.8) that

$$\text{Vol}\left(B_r\left(S_{\epsilon,r}^k(I)\right)\right) \leq \text{Vol}\left(B_r\left(U_r^\ell\right)\right) \leq C(n, \Lambda, \epsilon) r^{n-k}, \tag{8.15}$$

which proves the theorem. $\square$

8.2. **Proof of theorem 1.4.** There are several pieces to theorem 1.4. To begin with, the volume estimate follows easily now that theorem 1.3 has been proved. That is, for each $r > 0$ we have that

$$S^k_\epsilon(I) \subseteq S^k_{\epsilon,r}(I), \tag{8.16}$$

and therefore we have the volume estimate

$$\text{Vol}\left(B_r\left(S^k_\epsilon(I)\right) \cap B_1(p)\right) \le \text{Vol}\left(B_r\left(S^k_{\epsilon,r}(I)\right) \cap B_1(p)\right) \le C(n,K,H,\Lambda,\epsilon)r^{n-k}. \tag{8.17}$$

In particular, this implies the much weaker Hausdorff measure estimate

$$\lambda^k\left(S^k_\epsilon(I) \cap B_1(p)\right) \le C(n,K,H,\Lambda,\epsilon), \tag{8.18}$$

which proves the first part of the theorem.

Let us now focus on the rectifiability of S^k_ϵ. We consider the following claim, which is the $r = 0$ version of the main Claim of theorem 1.3. We will be applying lemma 7.1, which requires $K + H < \delta$. As in the proof of theorem 1.3 we can just assume this without any loss, as we can cover $B_1(p)$ by a controlled number of balls of radius $M^{-1}\delta$, so that after rescaling we can analyze each of these balls with the desired curvature assumption. Thus let us consider the following:

Claim: If $K + H < \delta$, then for each ℓ there exists a covering $S^k_\epsilon(I) \subseteq U^\ell_0 \cup U^\ell_+ = U^\ell_0 \bigcup B_{r^\ell_i}(x^\ell_i)$ such that

(1) $\lambda^k(U^\ell_0) + \omega_k \sum (r^\ell_i)^k \le C_\ell(\ell, n, K, H, \Lambda, \epsilon)$.
(2) U^ℓ_0 is k-rectifiable.
(3) $\sup_{y \in B_{r^\ell_i}(x^\ell_i)} \theta_{r^\ell_i}(y) \le \Lambda - \ell \cdot \eta$

The proof of the Claim follows essentially the same steps as those for the main Claim of theorem 1.3. For base step $\ell = 0$, we consider the decomposition $S^k_\epsilon \subseteq U^0_0 \cup U^0_+$ where $U^0_0 = \emptyset$ and $U^0_+ = B_1(p)$.

Now let us assume we have proved the claim for some ℓ, then we wish to prove the claim for $\ell + 1$. Thus, let us consider the set U^ℓ_+ from the previous covering step given by

$$U^\ell_+ = \bigcup B_{r^\ell_i}(x^\ell_i). \tag{8.19}$$

Now let us apply lemma 7.1 to each of the balls $B_{r^\ell_i}(x^\ell_i)$ in order to write

$$S^k_\epsilon \cap B_{r^\ell_i}(x^\ell_i) \subseteq U_{i,0} \cup U_{i,+} = U_{i,0} \cup \bigcup_j B_{r_{i,j}}(x_{i,j}), \tag{8.20}$$

with the following properties:

(a) $\lambda^k(U_{i,0}) + \omega_k \sum_j r^k_{i,j} \le C(n, K, H, \Lambda, \epsilon, p)(r^\ell_i)^k$,
(b) $\sup_{y \in B_{r_{i,j}}(x_{i,j})} \theta_{r_{i,j}}(y) \le \Lambda - (\ell + 1)\eta$,
(c) $U_{i,0}$ is k-rectifiable.

Now let us define the sets

$$U^{\ell+1}_0 = \bigcup U_{i,0} \cup U^\ell_0,$$
$$U^{\ell+1}_+ = \bigcup_{i,j} B_{r_{i,j}}(x_{i,j}). \tag{8.21}$$

Conditions (2) and (3) from the Claim are clearly satisfied. We need only check condition (1). Using (a) and the inductive hypothesis we can estimate that

$$\lambda^k(U_0^{\ell+1}) + \omega_k \sum_{i,j} (r_{i,j})^k \leq \lambda^k(U_0^\ell) + \sum_i \left(\lambda^k(U_{i,0}) + \omega_k \sum_j (r_{i,j})^k \right),$$

$$\leq C_\ell + C(n, K, H, \Lambda, \epsilon) \sum_i (r_i^\ell)^k$$

$$\leq C(n, K, H, \Lambda, \epsilon) \cdot C_\ell$$

$$\equiv C_{\ell+1} . \tag{8.22}$$

Thus, we have proved the inductive part of the claim, and thus the claim itself.

Let us now finish the proof that $S_\epsilon^k(I)$ is rectifiable. So let us take $\ell = \lceil \eta^{-1}\Lambda \rceil = \ell(\eta, \Lambda)$. Then if we apply the above Claim to ℓ, then by the third property of the Claim we must have that

$$U_+^\ell \equiv \emptyset , \tag{8.23}$$

and therefore we have the covering

$$S_\epsilon^k \subseteq U_0^\ell , \tag{8.24}$$

where U_0^ℓ is k-rectifiable with the volume estimate $\lambda^k(U_0^\ell) \leq C$, which proves that S_ϵ^k is itself rectifiable.

Finally, we prove that for k a.e. $x \in S_\epsilon^k$ there exists a k-dimensional subspace $V_x \subseteq T_x M$ such that *every* tangent cone at x is k-symmetric with respect to V_x. To see this we proceed as follows. For each $\eta > 0$ let us consider the finite decomposition

$$S_\epsilon^k = \bigcup_{\alpha=0}^{\lceil \eta^{-1}\Lambda \rceil} W_{\epsilon,\eta}^{k,\alpha} , \tag{8.25}$$

where by definition we have

$$W_{\epsilon,\eta}^{k,\alpha} \equiv \{ x \in S_\epsilon^k : \theta_0(x) \in [\alpha\eta, (\alpha + 1)\eta) \} . \tag{8.26}$$

Note then that each $W_{\epsilon,\eta}^{k,\alpha}$ is k-rectifiable, and thus there exists a full measure subset $\tilde{W}_{\epsilon,\eta}^{k,\alpha} \subseteq W_{\epsilon,\eta}^{k,\alpha}$ such that for each $x \in \tilde{W}_{\epsilon,\eta}^{k,\alpha}$ the tangent cone of $W_{\epsilon,\eta}^{k,\alpha}$ exists and is a subspace $V_x \subseteq T_x M$.

Now let us consider such an $x \in \tilde{W}_{\epsilon,\eta}^{k,\alpha}$, and let V_x^k be the tangent cone of $W_{\epsilon,\eta}^{k,\alpha}$ at x. For all $r << 1$ sufficiently small we of course have $|\theta_r(x) - \theta_0(x)| < \eta$. Thus, by the monotonicity and continuity of θ we have for all $r << 1$ sufficiently small and all $y \in W_{\epsilon,\eta}^{k,\alpha} \cap B_r(x)$ that $|\theta_r(y) - \theta_0(y)| < 2\eta$. In particular, by theorem 2.7 we have for each $y \in W_{\epsilon,\eta}^{k,\alpha} \cap B_r(x)$ that $B_r(y)$ is $(0, \delta_\eta)$-symmetric, with $\delta_\eta \to 0$ as $\eta \to 0$. Now let us recall the cone splitting of theorem 2.9. Since the tangent cone at x is V_x^k, for all r sufficiently small we can find $k + 1$ points $x_0, \ldots, x_k \in B_r(x) \cap W_{\epsilon,\eta}^{k,\alpha}$ which are $10^{-1}r$-independent, see Definition 2.8, and for which $B_{2r}(x_j)$ are $(0, \delta_\eta)$-symmetric. Thus, by the cone splitting of theorem 2.9 we have that $B_r(x)$ is (k, δ_η)-symmetric with respect to V_x^k for all r sufficiently small, where $\delta_\eta \to 0$ as $\eta \to 0$. In particular, every

tangent cone at x is (k, δ_η)-symmetric with respect to V_x, where $\delta_\eta \to 0$ as $\eta \to 0$.

Now let us consider the sets

$$\tilde{W}^k_{\epsilon,\eta} \equiv \bigcup_\alpha \tilde{W}^{k,\alpha}_{\epsilon,\eta}. \tag{8.27}$$

So $\tilde{W}^k_{\epsilon,\eta} \subseteq S^k_\epsilon$ is a subset of full k-dimensional measure, and for every point $x \in \tilde{W}^k_{\epsilon,\eta}$ we have seen that *every* tangent cone of is (k, δ_η)-symmetric with respect to some $V_x \subseteq T_x M$, where $\delta_\eta \to 0$ as $\eta \to 0$. Finally let us define the set

$$\tilde{S}^k_\epsilon \equiv \bigcap_j \tilde{W}^k_{\epsilon,j^{-1}}. \tag{8.28}$$

This is a countable intersection of full measure sets, and thus $\tilde{S}^k_\epsilon \subseteq S^k_\epsilon$ is a full measure subset. Further, we have for each $x \in \tilde{S}^k_\epsilon$ that *every* tangent cone must be (k, δ)-symmetric with respect to some V_x, for all $\delta > 0$. In particular, every tangent cone at x must be $(k, 0) = k$-symmetric with respect to some V_x. This finishes the proof of the theorem. □

8.3. **Proof of theorem 1.5.** Let us begin by observing the equality

$$S^k(I) = \bigcup_{\epsilon > 0} S^k_\epsilon(I) = \bigcup_{\beta \in \mathbb{N}} S^k_{2^{-\beta}}(I). \tag{8.29}$$

Indeed, if $x \in S^k_\epsilon(I)$, then no tangent cone at x can be $(k + 1, \epsilon/2)$-symmetric, and in particular $k + 1$-symmetric, and thus $x \in S^k(I)$. This shows that $S^k_\epsilon(I) \subseteq S^k(I)$. On the other hand, if $x \in S^k(I)$ then we claim there is some $\epsilon > 0$ for which $x \in S^k_\epsilon(I)$. Indeed, if this is not the case, then there exists $\epsilon_i \to 0$ and $r_i > 0$ such that $B_{r_i}(x)$ is $(k + 1, \epsilon_i)$-symmetric. If $r_i \to 0$ then we can pass to a subsequence to find a tangent cone which is $k + 1$-symmetric, which is a contradiction. On the other hand, if $r_i > r > 0$ then we see that $B_r(x)$ is itself $k + 1$-symmetric, and in particular *every* tangent cone at x is $k + 1$-symmetric. In either case we obtain a contradiction, and thus $x \in S^k_\epsilon(I)$ for some $\epsilon > 0$. Therefore we have proved (8.29).

As a consequence, $S^k(I)$ is a countable union of k-rectifiable sets, and therefore is itself k-rectifiable. On the other hand, theorem 1.4 tells us that for each $\beta \in \mathbb{N}$ there exists a set $\tilde{S}^k_{2^{-\beta}}(I) \subseteq S^k_{2^{-\beta}}(I)$ of full measure such that

$$\tilde{S}^k_{2^{-\beta}} \subseteq \{x : \exists V^k \subseteq T_x M \text{ s.t. every tangent cone at } x \text{ is } k\text{-symmetric wrt } V\}. \tag{8.30}$$

Hence, let us define

$$\tilde{S}^k(I) \equiv \bigcup \tilde{S}^k_{2^{-\beta}}(I). \tag{8.31}$$

Then we still have that $\tilde{S}^k(I)$ has k-full measure in $S^k(I)$, and if $x \in \tilde{S}^k(I)$ then for some β we have that $x \in \tilde{S}^k_{2^{-\beta}}$, which proves that there exists a subspace $V \subseteq T_x M$ such that every tangent cone at x is k-symmetric with respect to V. We have finished the proof of the theorem. □

9. PROOF OF MAIN THEOREM'S FOR MINIMIZING HYPERSURFACES

In this section we prove the main theorems of the paper concerning minimizing hypersurfaces. That is, we finish the proofs of theorem 1.6 and theorem 1.8. In fact, the proofs of these two results are almost identical, though the first relies on theorem 1.4 and the later on theorem 1.3. However, for completeness sake we will include the details of both.

9.1. **Proof of theorem 1.6.** We wish to understand better the size of the singular set $\text{Sing}\left(I^{n-1}\right)$, where $I^{n-1} \subset M^n$, of a minimizing hypersurface. Let us recall that the ϵ-regularity of theorem 2.11 tells us that if I is minimizing, then there exists $\epsilon(n, K, H, \Lambda) > 0$ with the property that if $x \in B_1(p)$ and $0 < r < r(n, K, H, \Lambda)$ is such that $B_{2r}(x)$ is $(n - 7, \epsilon)$-symmetric, then $r_I(x) \geq r$. In particular, x is a smooth point, and we have for $\epsilon(n, K, H, \Lambda) > 0$ that

$$\text{Sing}(I) \cap B_1(p) \subseteq S_\epsilon^{n-8}(I). \tag{9.1}$$

Thus by theorem 1.4 there exists $C(n, K, H, \Lambda) > 0$ such that for each $0 < r < 1$ we have

$$\text{Vol}\left(B_r\left(\text{Sing}(I)\right) \cap B_1(p)\right) \leq \text{Vol}\left(B_r\left(S_\epsilon^{n-8}(I)\right) \cap B_1(p)\right) \leq Cr^8. \tag{9.2}$$

This of course immediately implies, though of course is much stronger than, the Hausdorff measure estimate

$$\lambda^{n-8}(\text{Sing}(I) \cap B_1(p)) \leq C, \tag{9.3}$$

which finishes the proof of the first estimate in (1.18). As a simple corollary of this and the uniform bound $\theta(x, r) \leq c\Lambda$ for all $x \in B_1(p)$ and $r < 1$, we obtain also the second estimate in (1.18). $\qquad\square$

9.2. **Proof of theorem 1.8.** We begin again by considering the ϵ-regularity of theorem 2.11. This tells us that if I is a minimizing hypersurface, then there exists $\epsilon(n, K, H, \Lambda) > 0$ with the property that if $x \in B_1(p)$ and $0 < r < r(n, K, H, \Lambda)$ are such that $B_{2r}(x)$ is $(n - 7, \epsilon)$-symmetric, then $r_I(x) \geq r$. In particular, we have for such ϵ, r that

$$\{x \in B_1(p) : r_I(x) < r\} \subseteq S_{\epsilon,r}^{n-8}(I). \tag{9.4}$$

Thus by theorem 1.4 there exists $C(n, K, H, \Lambda) > 0$ such that for each $0 < r < 1$ we have

$$\text{Vol}(B_r\{x \in B_1(p) : r_I(x) < r\}) \leq \text{Vol}\left(B_r\left(S_{\epsilon,r}^{n-8}(I)\right) \cap B_1(p)\right) \leq Cr^8, \tag{9.5}$$

which proves the second estimate of (1.21). To prove the first we observe that $|A|(x) \leq r_I(x)^{-1}$, while the third is a corollary of the bound $\theta(x, r) \leq c\Lambda$ for all $x \in B_1(p)$ and $r < 1$. This concludes the proof of the theorem. $\qquad\square$

ACKNOWLEDGMENTS

References

[All72] W. K. Allard, On the first variation of a varifold, *Ann. of Math. (2)* **95** (1972), 417–491. MR 0307015. Zbl 0252.49028.

[Alm66] F. J. Almgren, Jr., Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem, *Ann. of Math. (2)* **84** (1966), 277–292. MR 0200816. Zbl 0146.11905.

[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.

[BDGG69] E. Bombieri, E. De Giorgi, and E. Giusti, Minimal cones and the Bernstein problem, *Invent. Math.* **7** (1969), 243–268. MR 0250205. Zbl 0183.25901.

[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.

[CM11] T. H. Colding and W. P. Minicozzi, II, *A course in minimal surfaces*, *Graduate Studies in Mathematics* **121**, American Mathematical Society, Providence, RI, 2011. MR 2780140. Zbl 1242.53007. http://dx.doi.org/10.1090/gsm/121.

[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.

[DG61] E. De Giorgi, *Frontiere orientate di misura minima*, *Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61*, Editrice Tecnico Scientifica, Pisa, 1961. MR 0179651.

[DL12] C. De Lellis, Allard's interior regularity theorem: an invitation to stationary varifolds, Lecture notes, 2012. Available at http://www.math.uzh.ch/fileadmin/user/delellis/publikation/allard_31.pdf.

[DL15] _____, The size of the singular set of area-minimizing currents, *lecture notes* (2015). Available at http://arxiv.org/abs/1506.08118.

[Fed69] H. Federer, *Geometric measure theory*, *Die Grundlehren der mathematischen Wissenschaften, Band 153*, Springer-Verlag New York Inc., New York, 1969. MR 0257325. Zbl 0176.00801.

[FF60] H. Federer and W. H. Fleming, Normal and integral currents, *Ann. of Math. (2)* **72** (1960), 458–520. MR 0123260 (23 #A588). Zbl 0187.31301.

[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.

[Mat95]   P. MATTILA, *Geometry of sets and measures in Euclidean spaces*, *Cambridge Studies in Advanced Mathematics* **44**, Cambridge University Press, Cambridge, 1995, Fractals and rectifiability. MR 1333890. Zbl 0819.28004. http://dx.doi.org/10.1017/CBO9780511623813.

[Mor00]   F. MORGAN, *Geometric measure theory*, third ed., Academic Press, Inc., San Diego, CA, 2000, A beginner's guide. MR 1775760. Zbl 0974.49025.

[Mor66]   C. B. MORREY, JR., *Multiple integrals in the calculus of variations*, *Die Grundlehren der mathematischen Wissenschaften, Band 130*, Springer-Verlag New York, Inc., New York, 1966. MR 0202511. Zbl 0142.38701.

[NVa]   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.

[NVb]   ———— , Volume estimates on the critical sets of solutions to elliptic pdes, *Preprint*. Available at http://arxiv.org/abs/1403.4176.

[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.

[Sim]   L. SIMON, Reifenberg's topological disc theorem. Available at http://www.math.uni-tuebingen.de/ab/analysis/pub/leon/reifenberg/reifenberg.dvi.

[Sim83]   ———— , *Lectures on geometric measure theory*, *Proceedings of the Centre for Mathematical Analysis, Australian National University* **3**, Australian National University Centre for Mathematical Analysis, Canberra, 1983. MR 756417. Zbl 0546.49019.

[Sim93]   ———— , Cylindrical tangent cones and the singular set of minimal submanifolds, *J. Differential Geom.* **38** (1993), 585–652. MR 1243788. Zbl 0819.53029. Available at http://projecteuclid.org/euclid.jdg/1214454484.

[Sim95]   ———— , Rectifiability of the singular sets of multiplicity 1 minimal surfaces and energy minimizing maps, in *Surveys in differential geometry, Vol. II (Cambridge, MA, 1993)*, Int. Press, Cambridge, MA, 1995, pp. 246–305. MR 1375258. Zbl 0874.49033.

[Sim68]   J. SIMONS, Minimal varieties in riemannian manifolds, *Ann. of Math. (2)* **88** (1968), 62–105. MR 0233295. Zbl 0181.49702. Available at http://www.jstor.org/stable/1970556.

[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.