QUANTITATIVE REIFENBERG THEOREM FOR MEASURES

NICK EDELEN, AARON NABER, AND DANIELE VALTORTA

ABSTRACT. We study generalizations of Reifenberg’s Theorem for measures in $\mathbb{R}^n$ under assumptions on the Jones’ $\beta$-numbers, which appropriately measure how close the support is to being contained in a subspace. Our main results, which holds for general measures without density assumptions, give effective measure bounds on $\mu$ away from a closed $k$-rectifiable set with bounded Hausdorff measure. We show examples to see the sharpness of our results. Under further density assumptions one can translate this into a global measure bound and $k$-rectifiable structure for $\mu$. Applications include quantitative Reifenberg theorems on sets and discrete measures, as well as upper Ahlfors’ regularity estimates on measures which satisfy $\beta$-number estimates on all scales.

CONTENTS

1. Introduction 2  
2. Main Theorems 6  
2.1. Main theorems 8  
2.2. Mass Bounds with Density 9  
2.3. $\beta_2$-Estimates on all Scales 10  
2.4. Applications. 12  
2.5. Outline of proof 14  
3. Examples 16  
3.1. Graphs 17  
3.2. Generalize Koch Snowflake 17  
3.3. $\beta$-Control without Density Bounds 18  
3.4. Sharpness without Density Assumptions 19  
4. Preliminaries 20  
4.1. Definitions and notation 20  
4.2. Basic lemmas 22  
4.3. Generalized center of mass 25  
4.4. Tilting and packing control 27  
5. Main construction 29  
6. Good tree 31  
6.1. Inductive hypothesis 35  
6.2. Covering control 35  

Date: February 28, 2017.
The first author was supported by NSF grant DMS-1606492, the second author has been supported by NSF grant DMS-1406259, the third author has been supported by SNSF grant 200021_159403/1.
1. INTRODUCTION

In his famous work on the solution of Plateau’s problem [Rei60], Reifenberg proved that if a closed set $S \subset \mathbb{R}^n$ is sufficiently well approximated by $k$-dimensional planes at all scales and points in the set, then it is $C^{0,\alpha}$-bi-Hölder to a disk. The exponent $\alpha$ can be made arbitrarily close to 1, so long as the approximation is sufficiently close. Specifically, Reifenberg considered a two-sided closeness condition on $S$:

$$\inf_{V} r^{-1}d_H(S \cap B_r(x), V \cap B_r(x)) \leq \delta(n, \alpha), \quad \forall x \in S \cap B_1, r \leq 8,$$

where the infimum is taken over all affine $k$-planes. Given (1.1) Reifenberg concludes that $S \cap B_1^n$ is bi-Hölder to $B_1^k$. Sets satisfying (1.1) are often called Reifenberg flat.

We remark that (1.1) has a strong uniform connotation. First of all, it rules out holes in $S$. Indeed, under (1.1) $S$ is topologically a disk, not just a closed subset of a disk. As a consequence, (1.1) automatically endows $S$ with lower mass bounds, that is, a set $S$ satisfying (1.1) is “lower $k$-Ahlfors regular”. However (1.1) cannot guarantee an upper mass bound (“upper Ahlfors regularity”), since we cannot take $\alpha = 1$ in Reifenberg’s theorem, except in the trivial case $\delta = 0$. See Example 3.4 for more on this.

A second consequence of (1.1) is that the set $S$ has no “excess set”, meaning that at all points and at all scales $S$ is uniformly contained in a small tubular neighborhood around a $k$-dimensional plane. In other words, $S$ cannot have parts of small measure (or even of zero measure) far away from its approximating plane at all points and at all scales.

This paper is motivated by the question: how can we control the mass and structure of $S$ if we allow for holes and excess sets? This question has begun playing a role in various settings, in particular in the analysis of singular sets of nonlinear PDE’s, see [NV17].
Especially in recent years, various papers dealing with this question have appeared in literature. In these papers, a need arose for one-sided notions of closeness with an integral flavor, which in particular allows for holes and excess set. Natural quantities that satisfy these requirements were introduced by Jones in [Jon90]. They are the so-called $\beta$-numbers for a nonnegative Borel measure $\mu$:

\[ \beta_{\mu,p}^k(x,r) := \left( \inf_{V \in k} \frac{1}{r^p} \int_{B_r(x)} r^{-p} d(z,V) d\mu(z) \right)^{1/p} \quad \text{for } 1 \leq p < \infty \]

\[ \beta_{\mu,\infty}^k(x,r) := \inf_{V \in k} (\inf \{ \delta : \text{supp} \mu \cap B_r(x) \subset B_{\delta r}(V) \}) \]

where the infima are taken over all affine $k$-planes $V^k$. See Definition 2.1 for a slight generalization, which will be used in this paper.

Toro [Tor95] and David-Toro [DT12] considered several generalizations of Reifenberg’s Theorem. They showed, for example, that if a closed set $S$ satisfies the uniform Reifenberg (1.1) and

\[ \int_0^1 \beta_{\mathcal{H}^k \cap S,2}^k(x,r)^2 dr/r \leq M \quad \forall x \in S, \]

with $p = 1$ or $\infty$, then the parameterization of Reifenberg’s Theorem is bi-Lipschitz. The reason it suffices to assume the $\beta$-numbers are square-summable is captured in the Koch snowflake of Example 3.4.

David-Toro [DT12] also proved that if a collection of points is in some discrete sense Reifenberg flat, then one can find a bi-Hölder mapping containing all the points. Assuming further a kind of “discrete $\beta$-number” summability condition, one can take the mapping to be bi-Lipschitz. Because [DT12] parameterize the entire net of points by a disk, a topological assumption in the spirit of small-tilting is necessary for their results (see [DT12, counterexample 12.4]; think thin twisted bands attached to an annulus).

If one is not interested in topological information, but only in control over the mass and rectifiability, it is reasonable to think that bounds on the $\beta_p$ ($p < \infty$) will be sufficient. The intuition is that we should only need to parameterize the region of $S$ near the $L^p$-best planes; the “excess set” away from the best-planes already has controlled mass.

The works of Tolsa [Tol15] and Azzam-Tolsa [AT15] verify part of this intuition. They prove that a set $S$ is $k$ rectifiable if and only if

\[ \int_0^1 \beta_{\mathcal{H}^k \cap S,2}^k(x,r)^2 dr/r < \infty \quad (\mathcal{H}^k \cap S) \text{-a.e. } x. \]

In fact their Theorem holds for general Radon measures, assuming $\mu$-a.e. positivity and finiteness of the upper density.

In this paper we shall prove effective bounds for the total measure and mass under a Dini-type assumption on the $\beta_2$, without any Ahlfors-regularity or Reifenberg-flat assumptions. In fact, we shall consider general nonnegative Borel-regular measures, without any ($\sigma$-)finiteness or density assumptions, and prove a combination of measure estimates on $\mu$ and Minkowski estimates on its support. As a corollary we recover the rectifiability theorem of [AT15].

Our complete list of theorems for the paper is rather long and at times technical, since it is important for the applications to consider the estimates in a good deal of generality. We therefore wait until Section 2 to give a complete list of main theorem. In preparation we list here a few corollaries of these main results.
which illustrate what type of behavior one might expect. We begin with the following, which holds for a
general Borel-regular measure without density assumptions:

**Theorem 1.1.** Let \( \mu \) be a nonnegative Borel-regular measure supported in \( B^1_1(0) \), and \( \epsilon \geq 0 \). Suppose

\[
\mu \left\{ z \in B_1(0) : \int_0^2 \beta_{\mu,2}(z,r)^2 \frac{dr}{r} > M \right\} \leq \Gamma. \tag{1.5}
\]

Then there is a closed, \( k \)-rectifiable set \( K \subset B_1(0) \) with \( \mathcal{H}^k(K) < c(n) \), so that we have

\[
|B_r(K)| \leq c(n)r^{n-k} \quad \forall r \leq 1, \quad \text{and} \quad \mu(B_r(x)) \geq \epsilon r^k \quad \forall x \in K, r \leq 1, \tag{1.6}
\]

and

\[
\mu(B_1(0) \setminus K) \leq c(n)(\epsilon + M) + \Gamma. \tag{1.7}
\]

Here \( |A| \) denotes the \( n \)-dimensional Lebesgue measure of \( A \), and \( B_r(A) \) the \( r \)-tubular neighborhood of \( A \). In this general case when one does not assume upper or lower density bounds, effective estimates must be split into two pieces: a \( \mu \)-noncollapsed set \( K \) for which there exists effective Minkowski estimates (in particular \( K \) has finite \( k \)-dimensional Hausdorff measure), and a set \( B_1 \setminus K \) which has uniformly bounded \( \mu \)-measure. See Example 3.7 to see that this setup is the sharp one.

Assuming additionally a bound on the lower-density allows us to prove apriori upper bounds on the measure of \( \mu \).

**Corollary 1.2.** If \( \mu \) is a non-negative Borel-regular measure satisfying the hypothesis (1.5), and with

\[
\Theta^k_\epsilon(\mu, x) \leq b \quad \text{for } \mu\text{-a.e. } x \in B_1, \tag{1.8}
\]

then

\[
\mu(B_1) \leq c(n)(b + M) + \Gamma. \tag{1.9}
\]

A key consequence of Corollary 1.2 is proving upper-Ahlfors-regularity from some kind of \( \beta \)-number condition. In certain Theorems (e.g. [GS]) this allows us to remove the assumption of upper-Ahlfors-regularity. We also bring the readers attention to [Mis16], who has recently and independently obtained a similar result to Theorem 1.3, assuming upper and lower bounds on the upper-density.

**Theorem 1.3.** Suppose \( \mu \) is a nonnegative Borel-regular measure supported in \( B_1 \) with the property

\[
\Theta^k_\epsilon(\mu, x) \leq b \quad \text{for } \mu\text{-a.e. } x \in B_1, \tag{1.10}
\]

and satisfying one of the following conditions:

A) For \( \mu\text{-a.e. } x \in B_1, \)

\[
\int_0^2 \beta_{\mu,2}(x,r)^2 \frac{dr}{r} \leq M. \tag{1.11}
\]

B) For \( \mu\text{-a.e. } x \in B_1, \) and every \( 0 < r \leq 1, \)

\[
\int_{B_r(x)} \int_0^{2r} \beta_{\mu,2}(z,r)^2 \frac{dr}{r} d\mu(z) \leq M^2r^k. \tag{1.12}
\]
C) \( \mu(B_1) < \infty \), and for \( \mu \)-a.e. \( x \in B_1 \), and every \( 0 < r \leq 1 \),

\[
\int_{B_r(x)} \int_0^{2r} \beta_{\mu,2}(z,r) \frac{dr}{r} d\mu(z) \leq M \mu(B_r(x)).
\]  

(1.13)

Then for every \( x \in B_1 \), \( 0 < r \leq 1 \), we have

\[
\mu(B_r(x)) \leq c(n)(b + M)r^k.
\]  

(1.14)

Moreover, if \( \Theta^{*,k}(\mu, x) > 0 \) \( \mu \)-a.e., then \( \mu \) is \( k \)-rectifiable.

When \( \mu = \mathcal{H}^k \llcorner S \), where \( \mathcal{H}^k \) is the \( k \)-dimensional Hausdorff measure, any one of the conditions A)\(-\)C) of Theorem 1.3 implies \( \mu \) is \( \sigma \)-finite, and therefore we obtain as an immediate Corollary:

**Corollary 1.4.** If \( \mu = \mathcal{H}^k \llcorner S \) satisfies any of the conditions A)\(-\)C) of Theorem 1.3, then \( S \) is \( k \)-rectifiable, and for every \( x \in B_1 \), \( 0 < r \leq 1 \), we have

\[
\mathcal{H}^k(S \cap B_r(x)) \leq c(n)(1 + M)r^k.
\]  

(1.15)

The corollary is similar in spirit to [AS, section 3], where the authors use a different notion of \( \beta \)-numbers based on a Choquet integral in terms of Hausdorff content and require lower regularity of the set \( S \).

Summability conditions like (1.5) have arisen in several interesting scenarios. The \( \beta \)-numbers have long been a bridge between rectifiability and boundedness of Calderon-Zygmund operators. Indeed for Ahlfors-David-regular sets \( S \), David-Semmes [DS93] show that uniform rectifiability is equivalent to an \( L^1 \)-bound of the form

\[
\int_{B_r(x) \cap S} \beta_{\mathcal{H}^k \llcorner S,2}(z,s)^2 d\mathcal{H}^k(z) \leq M \mu(B_r(x)) \quad \forall x, r.
\]  

(1.16)

Also, in the past 20 years there has been significant research on boundedness of Calderon-Zygmund operators on sets or measures satisfying only an upper-Ahlfors-regularity condition (e.g. [Dav98], [Tol99], [NTV97]). Very recently [GS] (see also [AT15] Theorem 1.4, and [JNT]) has shown that if a Radon measure \( \mu \) is upper-Ahlfors-regular, and satisfies

\[
\int_{B_r(x) \cap S} \beta_{\mu,2}^k(z,s)^2 \frac{d\mu(B_s(z))}{s} s d\mu(z) \leq M \mu(B_r(x)) \quad \forall x, r,
\]  

(1.17)

then Calderon-Zygmund operators are bounded in \( L^2(\mu) \).

If combine this result with Theorem 1.3, we obtain

**Corollary 1.5 (Theorem 1.3 + [GS]).** Let \( \mu \) be a nonnegative Borel-regular measure with the property that for \( \mu \)-a.e. \( x \in B_1 \), and every \( 0 < r \leq 1 \), we have

\[
\Theta^k_{\mu}(\mu, x) \leq b, \quad \int_{B_r(x)} \int_0^{2r} \beta_{\mu,2}^k(z,r)^2 \frac{dr}{r} d\mu(z) \leq M \mu(B_r(x)).
\]  

(1.18)

Then Calderon-Zygmund operators are bounded in \( L^2(\mu) \).

Another important application can be found in the analysis of singular sets. In [NV17], [NV15] the authors use a similar (though less general) technique to prove uniform mass bounds and rectifiability for the singular set of nonlinear harmonic maps and integral currents with bounded mean curvature.
In fact our Theorems generalize the discrete and rectifiable Reifenberg Theorems of [NV17], which prove mass bounds for discrete measures assuming small \( \beta \)-bounds at all scales. Our Main Theorem 2.6 works for arbitrary Borel-regular measures, relaxes the smallness assumption, allows the user to specify the scale of packing, and most importantly requires only weak \( \beta \)-bounds at a single scale.

Note that there is a growing literature on \( \beta \)-numbers and problems related to rectifiability. For example, [BS15, BS] proves necessary and sufficient condition for rectifiability of one dimensional curves, and [AS] focuses on the traveling salesman problem.

Our proofs are based on a generalized corona decomposition for the support of \( \mu \). Starting from \( B_1(0) \), we cover \( spt \mu \) with smaller and smaller balls, where at each step we decrease the radius of our covering by a factor of \( \rho \). For each ball in the covering, we consider the \( k \)-dimensional plane achieving the minimum in the definition of \( \beta \), i.e., the plane minimizing the \( L^2(\mu) \) norm of the distance. The part of \( spt \mu \) away from this best approximating plane \( V \) will be counted as “excess set”, and the estimates on \( \beta \) imply that this part has uniformly bounded measure. The part close to \( V \) will be covered by balls satisfying a Vitali condition. Each of these balls is classified into “good” or “bad” according to how spread the support of \( \mu \) is inside them. In particular, good balls are the ones where the support of \( \mu \) effectively spans \( V \), while bad balls are those where \( \mu \) is concentrated near some lower dimensional subspace.

On good balls, we will be able to exploit the bounds on the \( \beta \) numbers to obtain nice estimates on the tilting of the best planes from one scale to the next. This will allow us to use a Reifenberg-type construction to create a sequence of smooth manifolds \( T_i \) approximating \( spt \mu \) over good balls. These manifolds will have uniform bi-Lipschitz bounds.

On bad balls, we will lose the ability to refine this Reifenberg-type construction. However, on these balls the support of \( \mu \) is almost contained in a \( k-1 \) dimensional space, making it easy to obtain uniform \( k \)-dimensional estimates on it.

Thus we will be able to decompose \( \mu \) into a piece with controlled measure, and a piece supported on a set with uniform \( k \)-dimensional packing estimates. By assuming an upper bound over the (lower) density of \( \mu \), we will be able to turn the \( k \)-dimensional estimates on the support into mass estimates on \( \mu \) itself. Lower bounds on the (upper) density of \( \mu \) will allow us to conclude the \( k \)-rectifiability of this measure.

Acknowledgments The authors thank Max Engelstein and Benjamin Jaye for leading us to reference [GS], and warmly thank Tatiana Toro for several useful conversations. The first author expresses his deepest gratitude to Leon Simon for his insight and inspiration.

2. MAIN THEOREMS

In this section, we introduce the relevant definitions in order to state in a clean fashion our main theorems. We begin by slightly generalizing the standard definition of \( \beta_2 \), by allowing \( \beta_2 \) to only be zero on balls with small \( \mu \)-measure (as in the results this is all one needs).
Definition 2.1. Let $\mu$ be a Borel measure in $\mathbb{R}^n$, and $\bar{\epsilon}_\beta \geq 0$. The $k$-dimensional truncated $\beta$-number of $\mu$ in $B_r(x)$ is defined to be
\[
\beta_{\mu,2,\bar{\epsilon}_\beta}^k(x,r)^2 = \begin{cases} 
\inf_{V^k} r^{-k-2} \int_{B_r(x)} d(z,V)^2 d\mu(z) & \mu(B_r(x)) > \bar{\epsilon}_\beta r^k \\
0 & \mu(B_r(x)) \leq \bar{\epsilon}_\beta r^k.
\end{cases}
\] (2.1)
where we take the infimum over affine $k$-planes $V^k$. We write $V_{\mu}^k(x,r)$ for any of the affine $k$-planes achieving $\beta_{\mu,2}^k(x,r)$.

Remark 2.2. In all the theorems in this paper, we will assume that $\bar{\epsilon}_\beta$ is smaller than a constant depending on the parameters of the specific theorem. This means that we do not require any control over the measure $\mu$ on balls with small mass. This can be helpful in some applications.

Since this assumption is not common, we remark for the sake of clarity that with $\bar{\epsilon}_\beta = 0$ all the statements of this paper are evidently still valid.

Often we will simply drop subscripts when no confusion arises, for instance we may write $\beta_2$ in place of $\beta_{2,\mu,\bar{\epsilon}_\beta}$ and $V$ in place of $V_{\mu}^k$. We make some elementary remarks on distortion:

Remark 2.3 (Monotonicity of $\beta_2$). Note the following monotonicity type properties:

1. If $B_r(y) \subseteq B_R(x)$, and we know $\mu B_R(x) > \bar{\epsilon}_\beta R^k$, then $\beta_2(y,r)^2 \leq (R/r)^{k+2} \beta_2(x,r)^2$.
2. If $|x-y| < r$, and $\mu B_{2r}(y) > 2^k \epsilon_\beta r^k$, then $\beta_2(x,r)^2 \leq 2^{k+2} \beta_2(y,2r)^2$.
3. Lastly, if $\text{spt}\mu \subseteq B_1$, and $\mu B_1 > \bar{\epsilon}_\beta 2^k$, then for any $x \in B_1$ we have $\beta_2(x,16)^2 \leq c(k) \beta_2(x,2)^2$.

Remark 2.4 (Scale-invariance of $\beta_2$). A trivial but important property to note is the following: if $\mu_{x,r}(A) := r^{-k} \mu(x + rA)$, then $\beta_{\mu_{x,r},2}^k(y,s) = \beta_{\mu,2}^k(x + ry, rs)$.

Now our main theorem will be quite general in nature, something which naturally involves more technical details than would be liked in an introduction, thus let us take a moment to preface it. Besides allowing $\mu$ to be a general Borel-regular measure, it may be in practice that one only has $\beta$-number estimates on part of $\mu$, or only has estimates until a given fixed scale, which may vary from point to point in $\mu$, for instance in the case of discrete measures. Therefore, associated to our estimates we must consider a covering of the measure $\mu$ which tells us where the estimates hold. The following notation gives us a consistent way of dealing with this:

Definition 2.5. Let $C$ be a closed subset of $\mathbb{R}^n$, and $r_x : C \to \mathbb{R}_+$ a radius function, we say $(C, r_x)$ is a covering pair and we decompose $C = C_+ \cup C_0$, with $r_x > 0$ on $C_+$ and $r_x = 0$ on $C_0$. Given a measure $\mu$ on $\mathbb{R}^n$, a covering pair $(C, r_x)$ is called a covering pair for $\mu$ if $C \supseteq \text{spt}\mu$.

If $C' \subseteq C$, we write $C'_+ = C' \cap C_+$ and $C'_0 = C' \cap C_0$. For shorthand we will sometimes write
\[
B_{r_x}(C') := C'_0 \cup \bigcup_{x \in C'_+} B_{r_x}(x).
\] (2.2)

Associated to any $C'$ we have the packing measure
\[
\mu_{C'} := \mathcal{H}^k \cup C'_0 + \sum_{x \in C'_+} r_x^k \delta_x.
\] (2.3)

By $\delta_x$ we mean the Dirac delta centered at $x$. 

We are now ready to state our main theorems. Aside from the principal results, which are stated in a very general fashion, we prove many corollaries under different additional assumptions. For the reader’s convenience, we split these results into different subsections.

2.1. **Main theorems.** Our main technical Theorem is the following, which is stated with a good deal of generality. In particular, besides only assuming our measure \( \mu \) is Borel-regular, without additional density assumptions, we only assume weak summability estimates on the \( \beta \) numbers, as opposed to the more standard assumption of a.e. control or \( L^1 \) control. In subsequent subsections we will add additional hypotheses to obtain improved results, however almost all will follow in rather short order from the following:

**Theorem 2.6 (Main Theorem).** There is a constant \( c(n) \) so that the following holds: Let \( \epsilon \geq 0 \), and \( \mu \) be a non-negative Borel-regular measure in \( \mathbb{R}^n \) supported in \( B_1(0) \), and \( (C, r_x) \) a covering pair for \( \mu \) with \( r_x \leq 1 \). Suppose

\[
\mu \left\{ z \in B_1(0) : \int_{r}^{2r} \beta^2_{\mu,2,\epsilon_{\beta}}(z,r) \frac{dr}{r} > M \right\} \leq \Gamma, \tag{2.4}
\]

and that \( \bar{\epsilon}_{\beta} \leq c(k) \max\{M, \epsilon\} \). Then there is a closed subset \( C' \subset C \) which admits the following properties:

A) **Packing bounds:**

\[
\mathcal{H}^k(C'_0) + \sum_{x \in C'_+} r_x^k \leq c(n). \tag{2.5}
\]

In fact, we have the Minkowski type estimate

\[
r_x^{k-n}|B_r(C')| \leq c(n) \quad \forall 0 < r \leq 1. \tag{2.6}
\]

B) **Measure control away from \( C' \):**

\[
\mu \left( B_1 \setminus B_{r_x}(C') \right) \leq c(n)(\epsilon + M) + \Gamma. \tag{2.7}
\]

C) **A noncollapsing condition on \( C' \):**

\[
\mu(B_r(x)) \geq \max\{M, \epsilon\} r^k \quad \forall x \in C', \quad \forall r_x \leq r \leq 1. \tag{2.8}
\]

D) **Fine-scale packing structure:** \( C'_0 \) is closed, rectifiable, and admits the upper bounds

\[
\mathcal{H}^k(C'_0 \cap B_r(x)) \leq c(n) r^k \quad \forall x \in B_1, 0 < r \leq 1. \tag{2.9}
\]

**Remark 2.7.** Theorem 2.6 has a scale-breaking quality to it: the hypothesis and measure estimates B), C) scale with the measure, but the packing estimate does not. Note also that in this theorem we make no assumption on the density of \( \mu \).

**Remark 2.8.** The \( C'_0 \) is actually a kind of “uniformly rectifiable,” in the sense that whenever \( \mathcal{H}^k(C'_0 \cap B_r(x)) > 2^{-k-1} \omega_k r^k \), there is some bi-Lipschitz \( k \)-manifold \( T \) so that

\[
\mathcal{H}^k(C'_0 \cap T) \geq \frac{1}{2} \mathcal{H}^k(C'_0 \cap B_r(x)). \tag{2.10}
\]

**Remark 2.9.** If (2.4) holds for some \( C \) that not closed, then conclusions A)-D) hold for the closure \( \overline{C'} = C'_+ \cup C'_0 \) (this equality is due to Lemma 8.4).

**Remark 2.10.** In fact one can prove upper-Ahlfors-regularity of the packing measure \( \mu_{\overline{C'}} \), either by using Theorem 2.14 or by analyzing the tree structures more closely.
Let us also remark in the above that though we estimate all of \( \mu \) on the ball, it is necessary to break \( \mu \) into two complimentary pieces and give apriori fundamentally different types of estimates on each piece. Indeed, in (A) we look at the part of \( \mu \) covered by \( \mathcal{C}' \) and prove packing estimates on the support of \( \mu \), which give little control over \( \mu \) itself. On the other hand, we show in (B) that on the remaining part of \( \mu \) a uniform bound on the measure itself. Packing estimates and measure estimates are not related to one another for a general \( \mu \), though we will see in the next subsection how to convert between the two under additional lower or upper density estimates.

2.2. Mass Bounds with Density. The previous theorem works under essentially no assumptions on \( \mu \) other than the weak \( \beta_2 \) bound. As expected, we obtain better control over \( \mu \) by assuming also some control over its \( k \)-dimensional density. In particular, let us now see that assuming either upper or lower density bounds on \( \mu \) allows us to apply Theorem 2.6 in order to conclude mass bounds.

For the applications, in particular for a generalization of the discrete Reifenberg of [NV17], we will again want a version which is with respect to a general covering. In this case it is unnecessary, and indeed unnatural given the purpose of such a covering, to make density assumptions on \( \mu \) below the scale being considered. Thus let us consider the lower and upper density of \( \mu \) with respect to a covering pair \((\mathcal{C}, r_x)\) given by

\[
\Theta^*, k(\mu, \mathcal{C}, x) \equiv \limsup_{r \to r_x} \frac{\mu(B_{r/50}(x))}{(r/50)^k},
\]

\[
\Theta^k(\mu, \mathcal{C}, x) \equiv \liminf_{r \to r_x} \frac{\mu(B_{r/50}(x))}{(r/50)^k}.
\] (2.11)

Of course, if \( \mathcal{C} = \mathcal{C}_0 \) then this is the usual upper and lower density, while if \( \mathcal{C} = \mathcal{C}_+ \) and \( \mu \) is finite then \( \Theta^*, k(\mu, \mathcal{C}, x) = \Theta^k(\mu, \mathcal{C}, x) = (r_x/50)^{-k}\mu(B_{r_x/50}(x)) \).

Recall that if \( \mu = \mathcal{H}^k \downarrow S \), and \( \mathcal{H}^k(S) < \infty \), then for \( \mathcal{H}^k \)-a.e. \( s \in S \):

\[
2^{-k} \leq \Theta^*, k(\mathcal{H}^k \downarrow S, x) \leq 1.
\] (2.12)

Let us now see how to derive global mass bounds under density assumptions on our measure:

**Theorem 2.11** (Mass Bounds under Density). Let \( \mu \) be a non-negative Borel-regular measure in \( B_1 \) and \((\mathcal{C}, r_x)\) a covering pair for \( \mu \) such that \( r_x \leq 1 \). Suppose

\[
\mu \left\{ z \in B_1 : \int_{r_x}^{2r_x} \beta_{\mu,2,\epsilon_\beta}(z,r)^2 \frac{dr}{r} > M \right\} \leq \Gamma,
\] (2.13)

Then we have:

L) If \( \Theta^*, k(\mu, \mathcal{C}, x) \geq a > 0 \) for \( \mu \)-a.e. \( x \), then there exists \( \mathcal{C}' \subseteq \mathcal{C} \) with \( \mu(B_1 \setminus B_{r_x}(\mathcal{C}')) = 0 \) so that

\[
\mathcal{H}^k(\mathcal{C}_0) + \sum_{\mathcal{C}_+'} r_x^k \leq c(n)(\bar{\epsilon}_\beta + M + \Gamma) + c(n).
\] (2.14)

U) If \( \Theta^k(\mu, \mathcal{C}, x) \leq b \) for \( \mu \)-a.e. \( x \in \mathcal{C} \), then

\[
\mu(B_1) \leq c(n)(\bar{\epsilon}_\beta + b + M) + \Gamma.
\] (2.15)
Note that, since (2.13) is only an assumption at one scale, one cannot hope to obtain upper Ahlfors regularity for $\mu$. However, in Theorem 2.14 we will see that if one assumes (2.13) on all points and scales, then one may obtain uniform upper Ahlfors regularity on the measure. We refer to Examples 3.5 and 3.6 to see that the results of Theorem 2.11 are sharp, in that the improvement from Theorem 2.6 to Theorem 2.11 may fail without the corresponding density bounds.

As in the main theorem, our primary applications of the above are either in the case when $C_+ = \emptyset$, so we are controlling $\mu$ on all points and scales, or in the discrete case $C_0 = \emptyset$. Let us record our statements in these cases:

**Corollary 2.12 (Mass Bounds under Density).** Let $\mu$ be a non-negative Borel-regular measure in $B_1$, and $(C, r_x)$ a covering pair for $\mu$ with $C = C_0$ and $r_x \leq 1$. Suppose

$$
\mu \left\{ z \in B_1 : \int_0^2 \beta_{\mu, 2, \epsilon_\beta}^k (z, r)^2 \frac{dr}{r} > M \right\} \leq \Gamma.
$$

Then we have

L) If $\Theta^{*, k}(\mu, x) > a > 0$ for $\mu$-a.e. $x$, then $\exists C' \subseteq \text{spt} \mu$ with $\mu(B_1 \setminus C') = 0$ s.t.

$$
\mathcal{H}^k(C'_0) \leq \frac{c(n)(\epsilon_\beta + M + \Gamma)}{a} + c(n).
$$

U) If $\Theta_{\epsilon_\beta}^k(\mu, x) \leq b$ for $\mu$-a.e. $x$, then $\mu(B_1) \leq c(n)(\epsilon_\beta + b + M) + \Gamma$.

Additionally, in the case of a discrete cover (i.e. $C_0 = \emptyset$) we have the following:

**Corollary 2.13 (Packing Estimates under Density).** Let $\mu$ be a non-negative Borel-regular measure supported in $B_1$, and $(C, r_x)$ a covering pair for $\mu$ with $C = C_+$ and $r_x \leq 1$. Suppose

$$
\mu \left\{ z \in B_1 : \int_{r_x}^{2r_x} \beta_{\mu, 2, \epsilon_\beta}^k (z, r)^2 \frac{dr}{r} > M \right\} \leq M.
$$

Then we have

L) If $\mu(B_{r_x}(x)) \geq ar_x^k$, then $\exists C' \subseteq \text{spt} \mu$ with $\mu(B_1 \setminus B_{r_x}(C')) = 0$ s.t.

$$
\sum_{x \in C_+'} r_x^k \leq \frac{c(n)(\epsilon_\beta + M + \Gamma)}{a} + c(n).
$$

U) If $\mu(B_{r_x}(x)) \leq br_x^k$, then $\mu(B_1) \leq c(n)(\epsilon_\beta + b + M) + \Gamma$.

2.3. $\beta_2$-Estimates on all Scales. In this subsection we study what improvements to the main results are obtained when the $\beta_2$-estimates are assumed to exist on all scales. Under a density bound, either lower or upper, one is able to conclude upper Ahlfors regularity estimates. Our most general result is the following:

**Theorem 2.14 (Upper Ahlfors’s Regularity).** Take $\mu$ a non-negative Borel-regular measure in $B_1$ and $(C, r_x)$ a covering pair for $\mu$ such that $r_x \leq 1$, and for all $x \in C$ and all $r > 0$ we have

$$
\mu \left\{ z \in B_r(x) : \int_{r_x}^{2r} \beta_{\mu, 2, \epsilon_\beta}^k (z, s)^2 \frac{ds}{s} > M \right\} \leq Mr^k,
$$

(2.20)
Then we have

L) If $\Theta^{*,k}(\mu, C, x) \geq a > 0$ for $\mu$-a.e. $x \in C$, then there is a $C' \subset C$ with $\mu(B_1 \setminus B_{r_x}(C')) = 0$, satisfying:

$$H^k(C'_0 \cap B_r(x)) + \sum_{z \in C'_0 \cap B_r(x)} r_z^k \leq \left(\frac{c(n)(\bar{\epsilon}_\beta + M)}{a} + c(n)\right) r^k \quad \forall x \in C', r \geq r_x. \quad (2.21)$$

U) If $\Theta^k_*(\mu, C, x) \leq b$ for $\mu$-a.e. $x \in C$, then there is a $C' \subset C$, with $B_{r_x}(C') \supset C$, and satisfying:

$$\mu(B_r(x)) \leq c(n) (\bar{\epsilon}_\beta + b + M) r^k \quad \forall x \in C', r \geq r_x. \quad (2.22)$$

Despite appearances this theorem is not a straightforward corollary of Theorem 2.11. The issue is $C_+:\text{ in L), we want a single choice of balls which admit packing at all scales. When } C_+ = \emptyset \text{ then the proof becomes much easier.}$

In the spirit of the previous results let us again study the above in the cases when $C_+ = \emptyset$ or in the discrete case when $C_0 = \emptyset$. In the case $C_+ = \emptyset$ we have produced a generalization of the rectifiable Reifenberg from [NV17], while in the discrete case $C_0 = \emptyset$ we have produced a generalization of the discrete Reifenberg from [NV17]. We will analyze this a bit more in the next subsection. Let us begin with the $C_+ = \emptyset$ case:

**Corollary 2.15 (Upper Ahlfors’ Regularity).** Let $\mu$ be a non-negative Borel-regular measure in $B_1$, and $(C, r_x)$ a covering pair for $\mu$ with $C = C_0$ and $r_x \leq 1$. Suppose that for all $x$ and $r > 0$:

$$\mu \left\{ z \in B_r(x) : \int_0^{2r} \beta^{k}_{\mu, 2, \bar{\epsilon}_\beta}(z, s)^2 \frac{ds}{s} > M \right\} \leq Mr^k, \quad (2.23)$$

Then:

L) If $\Theta^{*,k}(\mu, x) \geq a$ for $\mu$-a.e. $x$, then there is a $C' = C'_0 \subset C$ with $\mu(B_1 \setminus C') = 0$, so that

$$H^k(C' \cap B_r(x)) \leq \left(\frac{c(n)(\bar{\epsilon}_\beta + M)}{a} + c(n)\right) r^k \quad \forall x \in B_1, r > 0. \quad (2.24)$$

U) If $\Theta^k_*(\mu, x) \leq b$ for $\mu$-a.e. $x$, then

$$\mu(B_r(x)) \leq c(n) (\bar{\epsilon}_\beta + b + M) r^k \quad \forall x \in B_1, r > 0. \quad (2.25)$$

Finally let us end with the following discrete case when $C_0 = \emptyset$:

**Corollary 2.16 (Upper Ahlfors’ Regularity).** Let $\mu$ be a non-negative Borel-regular measure supported in $B_1$, and $(C, r_x)$ a covering pair for $\mu$ with $C = C_+$ and $r_x \leq 1$. Suppose that for all $x$ and $r > 0$ we have

$$\mu \left\{ z \in B_r(x) : \int_{r_x}^{2r} \beta^{k}_{\mu, 2, \bar{\epsilon}_\beta}(z, s)^2 \frac{ds}{s} > M \right\} \leq Mr^k, \quad (2.26)$$

Then:

L) If $\mu(B_{r_x}(x)) \geq ar^k_x$, then there is a $C' \subset C$ with $\mu(B_1 \setminus B_{r_x}(C')) = 0$, and satisfying

$$\sum_{z \in C'_0 \cap B_r(x)} r_z^k \leq \left(\frac{c(n)(\bar{\epsilon}_\beta + M)}{a} + c(n)\right) r^k \quad \forall x \in C', r \geq r_x. \quad (2.27)$$
2.4. Applications. In this subsection we record a variety of applications of the previous results. In some of these applications the distinction between rectifiability of a set and of a measure is important. Let us be explicit: we say a Borel measure $\mu$ on $\mathbb{R}^n$ is $k$-rectifiable if $\mu \ll \mathcal{H}^k$, and $\mu(\mathbb{R}^n \setminus K) = 0$ for some $k$-rectifiable set $K$.

Our first application is to the structure of measures which satisfy an $L^1$ $\beta$-estimate at some scale. Under the additional hypothesis of lower and upper finiteness on the upper density it was shown in [AT15] that such measures are rectifiable. A decomposition similar in spirit to the one obtained here has been proven in [BS] for the case $k = 1$. Here we prove a sharp generalization of these results which does not require any density assumptions. Let us begin with an effective version:

**Theorem 2.17 (Structure under $\beta_2$-control).** Let $\mu$ be a non-negative Borel-regular measure supported in $B_1$, and $\epsilon > 0$. Suppose that

$$\int_{B_1} \int_0^{2r} \beta_{\mu, 2, 0}(z, r)^2 \frac{dr}{r} d\mu(z) \leq M^2$$

Then we can decompose $\mu = \mu_h + \mu_\ell + \mu_0$, so that:

A) $\mu_h(B_1 \setminus K_h) = 0$ for some $k$-rectifiable set $K_h$, with volume bounds

$$\mathcal{H}^k(K_h \cap B_r(x)) \leq c(n) r^k \quad \forall x \in B_1, 0 < r \leq 1,$$

and such that $\mathcal{H}^k(K_h \cap B_1) \leq c(k) \epsilon^{-1} \mu$ on open sets.

B) $\mu_\ell$ is $k$-rectifiable, with mass bounds $\mu_\ell(B_1) \leq c(n)(\epsilon + M)$.

C) $\mu_0(B_1) \leq c(n)(\epsilon + M)$, and $\Theta^{-k}(\mu, x) = 0$ for $\mu_0$-a.e. $x$.

The $\mu_0$ piece of the decomposition is important in the above, see Example 3.5 for an illustration of this. If we assume $\beta$-control at all scales, we obtain a similar decomposition, which satisfies an upper-Ahlffors-regularity condition on $\mu_\ell, \mu_0$. Notice again we assume nothing on the density.

**Corollary 2.18 ($\beta_2$-control at all scales).** Let $\mu$ be a non-negative Borel-regular measure supported in $B_1$, and $\epsilon > 0$. Suppose that

$$\int_{B_1} \int_0^{2r} \beta_{\mu, 2, 0}(z, r)^2 \frac{dr}{r} d\mu(z) \leq M^2 r^k \quad \forall x \in B_1, 0 < r \leq 1.$$  

Then we can decompose $\mu = \mu_h + \mu_\ell + \mu_0$ in the same manner as Theorem 2.17, but which instead of (2.30) admits the volume measures bounds:

$$\mathcal{H}^k(K_h \cap B_r(x)) \leq c(n)(1 + M/\epsilon) r^k \quad \text{and} \quad (\mu_\ell + \mu_0)(B_r(x)) \leq c(n)(\epsilon + M) r^k$$

for all $x \in B_1$, and $0 < r \leq 1$.

Assuming lower and upper bounds on the upper density, and finiteness of $\int_0^1 \beta_2(x, r)^2 \frac{dr}{r}$ for $\mu$-almost all $x$, it was shown in [AT15] that such measures are rectifiable. Here, as a Corollary to Theorem 2.17 we prove a generalization of this result where we only require lower bounds on the upper density to prove rectifiability of the support of $\mu$, and upper bounds on the lower density to prove rectifiability for the whole measure $\mu$. 

U) If $\mu(B_{r_x}(x)) \leq b r_x^k$, then there is a $C' \subset C$ with $B_{r_x}(C') \supset C$, and satisfying

$$\mu(B_r(x)) \leq c(n)(\epsilon \beta + b + M) r^k \quad \forall x \in C', r \geq r_x.$$  

(2.28)
Theorem 2.19 (compare to [AT15]). Let $\mu$ a non-negative Borel-regular measure in $B_1$ such that for $\mu$-a.e. $x$

$$\int_0^2 \beta_{\mu,2,0}^k (x,s)^2 \frac{ds}{s} < \infty, \quad \text{and} \quad \Theta^{*,k}(\mu, x) > 0.$$  \hfill (2.33)

Then there is a $k$-rectifiable set $K_0$ so that $\mu(B_1 \setminus K_0) = 0$. Moreover, if $\Theta^k(\mu, x) < \infty$ for $\mu$-a.e. $x$, then $\mu$ is $k$-rectifiable.

In another direction, applications of these techniques arise in [NV17], where it was proved that the singular sets of some nonlinear equations are rectifiable with uniform measure bounds. The key connection between singular sets and the type of theory presented here are the rectifiable Reifenberg and discrete Reifenberg theorems of [NV17]. The techniques of the current paper not only reproduce these, but allow us to prove generalizations which are substantially better in several regards. To discuss this let us state our generalized discrete Reifenberg.

Theorem 2.20 (Discrete Reifenberg). Let $\{B_{r_i}(x_i)\} \subseteq B_1$ be a disjoint collection of balls, and consider a general packing measure $\mu \equiv \sum \mu_i r_i^k \delta_{x_i}$. Let us suppose we have the estimate

$$\mu \left\{ x_i \in B_1 : \int_{r_i}^2 \beta_{\mu,2,\tilde{\epsilon}_\beta}^k (x_i, r)^2 \frac{dr}{r} > M \right\} \leq \Gamma.$$ \hfill (2.34)

Then the following hold:

L) If $\mu_i \geq a > 0$, then we have $\sum r_i^k \leq c(n)(\tilde{\epsilon}_\beta + M + \Gamma) + c(n)$.

U) If $\mu_i \leq b$, then we have $\mu(B_1) \leq c(n)(b + \tilde{\epsilon}_\beta + M) + \Gamma$.

The above generalizes the discrete Reifenberg of [NV17] in several ways, for instance it applies to a general class of discrete measures and only requires a bound on the $\beta$-estimates, not smallness. The most fundamental improvement however, is that one only needs to assume the $\beta$-estimates of (2.34) on the first scale in order to conclude the packing estimates. In [NV17], the discrete Reifenberg requires $\beta$-estimates analogous to (2.34) on every point and scale in order to conclude the desired mass bound.

We want to point out that in a recent and interesting work [Mis16] has independently obtained a generalization of discrete Reifenberg, which also only requires boundedness of the $\beta$-estimates (as opposed to smallness), but still requires $L^1$-estimates at all points and scales.

In a similar spirit, let us consider a generalization of the rectifiable Reifenberg of [NV17]. In [NV17], it is shown that if a set $S$ satisfies the correct $\beta$-estimates on every point and scale, then $S$ is rectifiable with uniform mass bounds. In the same vein as our generalized discrete Reifenberg, to obtain the mass bounds we may weaken this assumption substantially by only requiring $\beta$-estimates on the first scale:

Theorem 2.21 (Mass Bounds for Sets). Take $S \subseteq B_1$, and assume we have the estimate

$$\mathcal{H}^k \left\{ x \in S : \int_0^2 \beta_{\mathcal{H}^k,2,\tilde{\epsilon}_\beta}^k (x,r)^2 \frac{dr}{r} > M \right\} \leq \Gamma.$$ \hfill (2.35)

Then have that $\mathcal{H}^k(S) \leq c(n)(1 + \tilde{\epsilon}_\beta + M) + \Gamma$.

Remark 2.22. It is worth noting that the under the above assumption $S$ need not be rectifiable. It follows from Theorem 2.6 that all but some definite amount must be, but to prove all of $S$ is rectifiable requires assuming a true weak-$L^1$ assumption.
2.5. Outline of proof. Let us establish some basic intuition about what information $\beta$-numbers can and cannot give us.

First, it is instructive to point out that the $n$-Lebesgue measure $\mathcal{L}^n$ has summably finite $\beta$-number at every point:

$$\int_0^1 \beta^k_{\mathcal{L}^n,2}(x,r)^2 \frac{dr}{r} \leq c(n).$$

(2.36)

This is calculated in Example (3.5). This says that finiteness or even smallness of the summed $\beta$-number does not by itself give any kind of $k$-dimensional structure. In particular, given a condition like (2.4), without any extra information, we cannot hope to have a $k$-dimensional packing or Hausdorff estimate away from arbitrarily small $\mu$-measure. In other words, the $M$ in the measure estimate (2.7) is necessary.

Moreover, as illustrated in Example (3.6), by approximating the $n$-Lebesgue measure by packed $k$-spheres, even if we are dealing with a nice $k$-dimensional manifold we cannot hope to improve the $k$-dimensional Hausdorff estimates we obtain to stronger Minkowski or packing estimates. The problem in this case is that we have no information about the scale at which things start to look $k$-dimensional. Thus we see there is some subtlety as to what estimates exist, and it is important to split our general estimates into two types: Hausdorff estimates on one piece and measure estimates on the other.

However, let us come to terms with this, and attempt prove a packing and measure estimate as in Theorem 2.6. A basic motivating observation is the that $\beta$-numbers control mass at any fixed distance from the $L^2$-best planes. For example, if we write $V^k$ for the $L^2$-best plane of $B_1(0)$ which realizes $\beta^k_{\mu,2}(0,1)$, then

$$\mu(B_1(0) \setminus B_\rho(V)) \leq \rho^{-2} \beta(0,1)^2$$

(2.37)

In other words, the region of $B_1(0)$ with $\mu$-mass much larger than $\beta(0,1)^2$ is concentrated in the neighborhood $B_\rho(V^k)$.

(2.37) suggests the following naive strategy. Take a Vitali cover of $B_\rho(V^k)$ by balls of size $\rho$, and thereby obtain a $k$-dimensional packing estimate on high-mass region in $B_1(0)$. In each ball $B_\rho(x)$, we can repeat the same process with balls of size $\rho^2$, obtaining a $k$-packing estimate on the high-mass region of $B_\rho(x)$. Then repeat in each $B_{\rho^2}$, and then in each $B_{\rho^3}$, etc...

Unfortunately, no matter how small we fix $\rho$, attempting to iterate this will result in an exponential error, due to double-counting on the ball-overlaps. If one attempts to use disjoint cubes instead of balls, an exponential error arises for a different reason: the $k$-volume of a $k$-plane in the cube varies with orientation.

In order to obtain a finite $k$-dimensional packing estimate, we require either some kind of global memory as we progress in scale, or the knowledge that big mass regions look like a very small portion of a $k$-plane. It turns out these two options give a dichotomy: in any ball, the region with big mass either looks “very $k$-dimensional,” giving us enough tilting control to construct a global memory; or it looks “at most $(k-1)$-dimensional,” giving us very good packing estimates to compensate for double-counting. This dichotomy is captured in Lemma 4.12.

We call balls satisfying the first condition of being “very $k$-dimensional” good, and those satisfying the second condition of being “at most $(k-1)$-dimensional” bad. See Definition 4.3 for a precise formulation. It is worth emphasizing that our notion of a good ball, as compared to similar constructions in previous
results (e.g. [NV17]), is much more general. Our ability to technically deal with such good balls without additional measure assumptions is a key observation for the paper.

Now let us assume we have a condition like
\[
\int_0^1 \beta_{\mu,2}^k(z,r)^2 \frac{dr}{r} \leq \delta(n)^2 M \quad \forall z \in B_1(0),
\]
where \( \delta(n) \) is some small constant. We say a ball \( B_r(x) \) is a stop ball if \( \mu(B_r(x)) \leq M r^k \), so that non-stop balls have big mass relative to \( \delta(n)^2 M \).

We implement two separate stopping-time arguments, one for good balls (Section 6) and one for bad balls (Section 7). In each case we build a tree of balls, which is a successive refinement of coverings by good, bad, and stop balls. We then chain these trees together (Section 8) to obtain our estimate. Let us describe the good and bad tree constructions in a little more detail.

A good tree is built in the following manner. Start at some initial good ball \( B_{r_0}(g_0) \), with \( L^2 \)-best plane \( L_0 \). We define the good/bad/stop balls at the next scale \( r_1 \) to be a Vitali cover of \( B_{r_1/40}(L_0) \). Let us call these \( \{B_{r_1}(g)\}_{g \in G_1}, \{B_{r_1}(b_1)\}_{b \in B_1} \), and \( \{B_{r_1}(s)\}_{s \in S_1} \), respectively.

Now \( B_{r_0}(g_0) \) is covered by \( B_{r_0}(g_0) \setminus B_{r_1/50}(L_0) \), and the good/bad/stop balls at scale \( r_1 \). The set \( B_{r_1}(g_0) \setminus B_{r_1/50}(L_0) \) is called the excess set for \( B_{r_0}(g_0) \).

Write \( L_{1,g} \) for the \( L^2 \)-best plane associated to good ball \( B_{r_1}(g) \). We define the good/bad/stop balls at scale \( r_2 \) to be a Vitali cover of
\[
\left( \bigcup_{g \in G_1} B_{r_1}(g) \cap B_{r_2/40}(L_{1,g}) \right) \setminus \left( \bigcup_{b \in B_1} B_{r_1}(b) \cup \bigcup_{s \in S_1} B_{r_1}(s) \right).
\]

In other words, we cover the neighborhoods of \( L^2 \)-best planes, while avoiding all previous bad/stop balls. Now \( B_{r_0}(g_0) \) is covered by the various excess sets, and the good \( r_2 \)-balls, and the bad/stop balls at scales \( r_1 \) and \( r_2 \).

We proceed in this fashion, inducting into the \( r_1 \)-good balls, and avoiding the bad/stop balls at scales \( r_1, r_1^{-1}, \ldots, r_1 \). We end up with a cover of \( B_{r_0}(g_0) \) by excess sets, bad/stop balls at all scales, and a region inside every good \( r_1 \)-ball collection. This covering is called the good tree.

In the good tree, we have good tilting control. We can perform a construction analogous to [Rei60], [Tor95], [DT12], [NV17], to build at each scale a bi-Lipschitz manifold \( T_i \) that cuts through all \( r_i \)-good balls in a controlled fashion, and all bad/stop balls at scales \( r_i, r_i^{-1}, \ldots, r_1 \). Since the corresponding balls of radius \( r_i/10 \) are pairwise disjoint, and their intersection with \( T_i \) has \( k \)-measure comparable to \( r_i^k \), we obtain the \( k \)-packing estimate in a good tree, and we can bound the measure of stop balls and excess sets in terms of the packing estimate.

Let us elaborate a little on the manifold construction. As in [Rei60], each \( T_i \) is essentially a glorified interpolation of planes, constructed inductively via maps \( \sigma_i \). We set \( T_0 = L_0 \), and then \( T_{i+1} = \sigma_{i+1}(T_i) \). More precisely, if \( \{B_{r_i}(g)\}_{g \in G_i} \) is the collection of good \( r_i \)-balls, and \( \{L_{i,g}\} \) are associated \( L^2 \)-best planes, then \( \sigma_i \) is the following interpolation of projection mappings
\[
\sigma_i(x) = \sum_{g \in G_i} \phi_{i,g}(x)(\proj_{L_{i,g}}(x - X_{i,g}) + X_{i,g}).
\]
Here $X_{i,g}$ is the (generalized) $\mu$-center of mass for $B_{ri}(g)$ (see Definition 4.2), $\text{proj}_{L_{i,g}}$ the linear projection to $L_{i,g}$, and $\phi_{i,g}$ is a partition of unity subordinate to $\{B_{3ri}(g)\}$.

Estimates of Section 4.4 control the tilting of subsequential $L^2$-best planes in terms of $M^{-1/2}\beta$, and consequently in Section 6.3 we obtain scale-invariant $C^1$ bounds on $\sigma_i$ in terms of $M^{-1/2}\beta$ also. These estimates are essentially standard, but we emphasize that our construction requires no upper mass control in the good and bad balls. We may be packing regions with infinite $\mu$-mass.

These various estimates imply that the composition $\sigma_i \circ \cdots \circ \sigma_1$ has a bi-Lipschitz bound at $x$ like

$$\prod_{\ell=0}^{i}(1 + c(n)M^{-1}\beta(x,r_\ell)2) \leq e^{c(n)\delta^2}. \quad (2.41)$$

This gives a uniformly bi-Lipschitz bound on each $T_i$, and shows that $T_i \rightarrow T_\infty$ for some bi-Lipschitz $T_\infty$.

Ultimately, the main Theorem 6.6 of Section 6 says: the good tree decomposes $B_{r_0}(g_0)$ into a region with bounded mass, a collection of bad balls with packing, and a subset of $T_\infty$ (the region inside every good $r_i$-ball collection).

The bad tree is constructed in essentially the same way as the good tree, except with the roles of good and bad balls swapped, and with a “best” $(k - 1)$-plane instead of a best $L^2$ $k$-plane. In bad trees we are covering a neighborhood of a $(k - 1)$-plane, and can therefore naively estimate the $k$-packing of all bad/good/stop balls at across scales. In fact we can make our $k$-packing estimate very small, by choosing the scale-drop very small, which also allows us to not worry about issues like double counting, which is a serious concern for the good ball coverings. The measure of stop balls and excess sets is then controlled by the packing.

The main Theorem 7.4 of Section 7 says: the bad tree decomposes a bad ball $B_{r_0}(b_0)$ into a region with bounded mass, a collection of good balls with small packing, and a region with $H^k$-measure 0 (the region inside every bad $r_i$-ball collection).

We outline how to chain these trees together. In a good tree, away from the bad balls we have packing and measure estimates down to a small scale. We sometimes call the bad balls of a good tree the leaves. Similarly, in a bad tree we have small-scale packing and measure estimates away from the good balls (the bad tree leaves).

To obtain global estimates down to a small scale we first construct an initial good or bad tree at $B_1(0)$, which of course we can assume is not a stop ball (otherwise we are done anyway). At any leaf of this good/bad tree, we build bad/good tree, and thereby obtain global estimates away from smaller leaves. Then in leaves of the secondary trees, we build tertiary trees, and continue this construction inductively.

Each time we build a new family of trees, the trees switch type, and we get estimates on a smaller scale. The type-switching is very important: it means we can always cancel double-counting error with the small bad-tree packing, choosing an appropriate scale drop (equation (8.1)). In fact we end up with global packing on all leaves of all trees (Theorem 8.1), which allows us to concatenate packing and measure estimates from each individual tree.

3. Examples

In this section we gather some examples regarding measures and $\beta$ number estimates. We start with a basic example on the finiteness of the $\beta$ number for graphs in order to help build an intuition.
3.1. Graphs.

Example 3.1. Let $S$ be the graph of a $C^2$ function $f : \mathbb{R}^k \to \mathbb{R}^{n-k}$, so that $S \subset \mathbb{R}^n$. Evidently, one expects $\mathcal{H}^k \llcorner S$ to have nice $\beta^2$ bounds. In order to verify this, we compute the Taylor expansion of $f$ around $(x_0, f(x_0)) \in S$:

$$|f(x) - f(x_0) - \nabla f(x_0) \cdot (x - x_0)| \leq \|\nabla^2 f\|_\infty |x - x_0|^2. \quad (3.1)$$

In particular, using simple geometric considerations, this proves that

$$\beta^k S((x_0, f(x_0)), r)^2 \leq c(k, \|\nabla f\|_\infty) \|\nabla^2 f\|_\infty r. \quad (3.2)$$

In turn, this implies the uniform pointwise bound

$$\int_0^1 \beta^k S((x_0, f(x_0)), r)^2 \frac{dr}{r} \leq c(k, \|\nabla f\|_\infty) \|\nabla^2 f\|_\infty. \quad (3.3)$$

It is easy to see that a similar computation holds for all graphs of $C^{1, \alpha}$ functions.

As opposed to the previous example, we show an easy case where the integral in (3.3) does not converge.

Example 3.2. Let $S$ be the graph of the function $\alpha |x|$ in $\mathbb{R}^2$, where $\alpha \neq 0$. Then clearly $\beta^k S((0,0), r)$ is constant in $r$, and thus $\int_0^1 \beta^k S((0,0), r)^2 \frac{dr}{r} = \infty$. This proves that the estimate in (3.3) cannot hold in a pointwise sense for Lipschitz graphs.

It is worth noticing however that although the pointwise bound of (3.3) does not hold for Lipschitz graphs, an integral estimate follows from [1, theorem 6] (see also [AS, theorem 1.2] and [DS93, Theorem 1.42]). For the sake of completeness, here we report the result (without proof):

**Proposition 3.3.** Let $f : \mathbb{R}^k \to \mathbb{R}^{n-k}$ be a Lipschitz map with Lipschitz constant bounded by $L$, and let $S$ be the graph of $f$. Then for all $x \in S$ and $r \geq 0$ we have

$$\int_{B_r(x)} \left( \int_0^r \beta^k S(x, s)^2 \frac{ds}{s} \right) d\mathcal{H}^k(x) \leq C(k)(1 + L^2)k/2 L^2 r^k. \quad (3.4)$$

3.2. Generalize Koch Snowflake.

Example 3.4. In this example, we recall some properties of the snowflake curve, which will allow us to illustrate why it is reasonable to use summability properties of the square of the beta numbers in order to obtain measure bounds and rectifiability. We start by briefly recalling the construction of the Koch curve.

The construction of a Koch curve (snowflake) of height $\kappa > 0$ is well known (as a reference, see [Mat95, section 4.13]). The basic building block is the following: Given $a, b \in \mathbb{R}^2$ and the induced segment $\ell = [a, b] \subset \mathbb{R}^2$ of length $|\ell|$, split the segment into three equal parts $[a, x_1] \cup [x_1, x_2] \cup [x_2, b]$, and replace the middle segment $[x_1, x_2]$ with a pair of segments $[x_1, z] \cup [z, x_2]$ having the same length and such that $d(z, \ell) = \kappa |\ell|$. To build the Koch snowflake let us consider a sequence of piecewise segments $\gamma_i$ defined as follows. Let $\gamma_0$ be the unit segment $[0, 1] \times \{0\} \subset \mathbb{R}^2$. To construct $\gamma_{i+1}$ from $\gamma_i$ let us apply the building block construction above with parameter $\kappa$ to each segment of $\gamma_i$. Limiting gives the Koch curve $\gamma$. 
It is clear that the length of the curve \( \gamma_i \) is given by \( |\gamma_i| = \left(1 + \left(\sqrt{1 + c\kappa_i^2} - 1\right)/3\right)|\gamma_{i+1}| \), so the length of the snowflake \( \gamma \) will be infinity for any \( \kappa > 0 \). This is a simple application of the Pythagorean theorem, and the extra square power on \( \kappa \) comes from the fact that loosely speaking at each step we are adding some length \( \kappa \) to the curve, but in a direction perpendicular to it.

To build the generalized Koch snowflake let us instead consider a sequence of parameters \( \kappa_i \). Let \( \gamma_0 \) be the unit segment as before, and now to construct \( \gamma_{i+1} \) from \( \gamma_i \) let us apply the building block construction above with parameter \( \kappa_i \) to each segment of \( \gamma_i \). Limiting gives the generalized snowflake \( \gamma \). It is clear that the length of the curve \( \gamma_{i+1} \) is given by \( |\gamma_{i+1}| \sim (1 + c\kappa_i^2)|\gamma_i| \), so the length of the snowflake \( \gamma \) is finite if and only if \( \sum \kappa_i^2 < \infty \).

This suggests that the finiteness of the Hausdorff measure of the set \( S \) is related to the summability properties of squares of its \( \beta \) numbers over scales.

### 3.3. \( \beta \)-Control without Density Bounds

The following examples illustrate that some of the assumptions in the main theorems (and its corollaries) are necessary. In particular, we will see how densities play a role in our results.

**Example 3.5.** This example illustrates how in Theorem 2.17, we cannot get rid of the \( \mu_0 \) part in the decomposition without assuming lower bounds on \( \Theta^{*,k} \). Consider the \( n \)-dimensional Lebesgue measure \( \mu = \mathcal{L}^n \). Then for every \( x, r \), we have

\[
\beta^k(x, r)^2 \leq r^{-k-2} \int_{B_r} 2r^2 d\mathcal{L}^n = 2\omega_n r^{n-k}.
\]

Hence for all \( x \in B_1(0) \) and \( R \leq 2 \):

\[
\int_0^R \beta^k(x, r)^2(x, r) \frac{dr}{r} \leq c(n) \int_0^R r^{n-k-1}dr \leq \frac{c(n)}{n-k} R^{n-k}.
\]  

(3.6)

However \( \mathcal{L}^n \) is trivially not supported on any set of finite \( \mathcal{H}^k \)-measure. This tells us that finiteness (or even smallness) of \( \int \beta^2 dr/r \) by itself does not give any kind of \( k \)-dimensional structure.

In the next example, we show that in the main Theorem 2.6, we cannot obtain uniform packing estimates on the whole support of the measure \( \mu \). In other words, we show that in general point \( (B) \) in theorem 2.6 is necessary.

**Example 3.6.** Fix \( \rho \leq 1 \). Let \( \{x_1\} \) be a maximal \( \rho \)-net in \( B_2 \), and we can assume the \( x_i \) lie on a lattice parallel to coordinate axes. For each \( x_i \), let \( S_i \) be a \( k \)-sphere centered at \( x_i \), of radius \( \rho^{n/k} \). Set \( S = B_1 \cap (\cup_i S_i) \), and let \( \mu = \mathcal{H}^k \cap S \).

We have by construction that

\[
\mu(B_1) \leq c(n), \quad \Theta^k(x, \mu) = 1 \quad \forall x \in \text{spt}\mu.
\]

(3.7)

We claim that for all \( x \in B_1(0) \):

\[
\int_0^2 \beta^k(x, r)^2 \frac{dr}{r} \leq c(n).
\]

(3.8)
Let us prove the Claim. Morally what is happening, is that on scales \( \rho^{n/k} \) we look like a \( k \)-sphere, while on scales \( \rho^{n/k} > \rho^{n/k} \) we look like the Lebesgue measure as in the last example. Fix \( r > 0 \). If \( r \leq \rho^{n/k}/10 \), then \( S \cap B_r(x) \) is the graph of a \( C^2 \) function over some \( k \)-dimensional plane. Thus, according to example 3.1, we deduce

\[
\int_0^{10^{-1} \rho^{n/k}} \beta^k(x,r)^2 \frac{dr}{r} \leq c(n) \quad (3.9)
\]

for every \( x \).

Suppose \( r \geq \rho^{n/k} \). Then as with the Lebesgue measure we can calculate

\[
\beta^k(x,r)^2 \leq c(n)r^{-k-2} \cdot (r/\rho)^k \sum_{\ell=0}^{r/\rho} (\ell \rho + \rho^{n/k})^2 \cdot \rho^n (r/\rho)^{n-k-1} \leq c(n)r^{n-k} \quad (3.10)
\]

Therefore

\[
\int_0^2 \beta^k(x,r)^2 \frac{dr}{r} \leq c(n) \quad \text{and thus} \quad \int_0^2 \beta^k(x,r)^2 \frac{dr}{r} \leq c(n) \quad (3.11)
\]

also. This proves the claim.

So \( \mu \) satisfies all the hypothesis of Theorem 2.6 uniformly in \( \rho \), and has uniformly bounded upper density. However we trivially have that \( S \) fails any kind of packing estimate for \( r \geq \rho \):

\[
|B_\rho(S)| \geq |B_1| \geq c(n) \quad (3.12)
\]

Let’s go further, and even assume that we have cut out a set \( S’ \) of measure \( \mathcal{H}^k(S’) \leq \epsilon \). Then the number of entire \( k \)-spheres \( S_i \) which can be covered by \( S’ \) is at most \( c(n)\epsilon \rho^{-n} \). Therefore

\[
|B_\rho(S \setminus S’)| \geq |B_1 \setminus B_{c(n)\epsilon}| \geq 1 - c(n)\epsilon^n \quad (3.13)
\]

This shows that in general we cannot hope to have reasonable \( k \)-dimensional packing estimates on the support of \( \mu \), not even if we consider the support of \( \mu \) away from a portion with \( \epsilon \) mass.

3.4. **Sharpness without Density Assumptions.** In this subsection we give a brief example both illustrating Theorem 2.6 and showing that the estimates given are sharp.

**Example 3.7.** For \( L^k \subseteq \mathbb{R}^n \) a \( k \)-dimensional subspace with \( \{x_i\} \subseteq L^k \cap B_1 \) a countable dense subset and \( \Lambda_i > 0 \) any collection of positive numbers. Let us consider the measure

\[
\mu = L^n_{|B_1} + \sum_i \Lambda_i \delta_{x_i},
\]

where \( L^n \) is the standard Lebesgue measure and \( \delta_{x_i} \) is a Dirac delta centered at \( x_i \in L^k \cap B_1 \). As in Example 3.5 we have the pointwise \( \beta \)-control

\[
\int_0^2 \beta^k_\mu(x,s)^2 \frac{ds}{s} < C(n) \quad \text{for all} \ x \in B_1, \quad (3.15)
\]

where notice the bound on the right hand side is independent of the \( \Lambda_i \). Clearly, this measure has no upper or lower \( k \)-density bounds at any point. Applying Theorem 2.6, or more appropriately Corollary 2.15, we
obtain the $k$-rectifiable subset $\mathcal{K} = L^k \cap B_1$ with uniform $k$-Minkowski and $k$-Hausdorff estimates such that $\mu(B_1 \setminus \mathcal{K}) < C(n)$, as claimed.

We see in this extreme example where no point has either an upper or lower density bound that we cannot strengthen the estimates of the main theorems.

4. Preliminaries

In this section, we gather the necessary preliminary results on $\beta$-estimates and tilting of the best approximating planes. Our ambient space will always be $\mathbb{R}^n$, equipped with the usual Euclidean metric.

Most of our Theorems are stated in a $k$-dimensional scale-invariant form. When we say “apply Theorem X to the measure $\mu$ at scale $B_r(x)$,” we mean “apply Theorem X to the measure $\mu_{x,r}$,” where

$$\mu_{x,r}(A) := r^{-k} \mu(x + rA).$$

(4.1)

4.1. Definitions and notation. Let us outline some basic notation and definitions used throughout the paper. Given a set $A \subset \mathbb{R}^n$, write

$$B_r(A) = \{x : \text{dist}(x, A) < r\}$$

(4.2)

for the (open) $r$-tubular neighborhood of $A$. We write $|A| = \mathcal{L}^n(A)$ for the $n$-dimensional volume, and $\overline{A}$ for the closure of the set $A$. For ease of notation we will sometimes write $B_r(0)$ as $B_r$. We write $\mathcal{H}^k(A)$ for the $k$-dimensional Hausdorff measure of $A$, and $\mathcal{H}^k_\delta$ for the usual $\delta$-approximation of Hausdorff measure.

Given a measure $\mu$, the upper-/lower-$k$-densities are defined as

$$\Theta^{*,k}(\mu, x) = \limsup_{r \to 0} \frac{\mu(B_r(x))}{r^k}, \quad \Theta^{k}(\mu, x) = \liminf_{r \to 0} \frac{\mu(B_r(x))}{r^k}. \quad (4.3)$$

We write $d_H$ for the usual Hausdorff distance between sets

$$d_H(A_1, A_2) = \inf \{r : A_1 \subset B_r(A_2) \quad \text{and} \quad A_2 \subset B_r(A_1)\}. \quad (4.4)$$

Given two subspaces $L_1, L_2$, we define a Grassmanian type distance

$$d_G(L_1, L_2) = d_H(B_1(0) \cap L_1, B_1(0) \cap L_2). \quad (4.5)$$

If $L_1, L_2$ are affine spaces, we define $d_G$ in terms of the associated subspaces.

Given an affine or linear space $L$, we will write $p_L$ for the linear projection mapping, onto the associated linear subspace. It’s easy to show that

$$d_G(L_1, L_2) \leq |p_{L_1} - p_{L_2}| \leq 2d_G(L_1, L_2), \quad (4.6)$$

where $|p| = \sup_{|v| \leq 1} |p(v)|$ denotes the operator norm. In this paper we will deal exclusively with the operator norm for linear mappings.

Given a function $f$, we define the $C^k$-norm at scale $\rho$ to be

$$|f|_{C^k_\rho} = \sum_{\ell=0}^{k} \sup_{x \in \text{dom}(f)} \rho^{\ell-1} |D^\ell f|, \quad \text{where} \quad \left| D^\ell f \right| = \sup_{|v_1| = \cdots = |v_\ell| = 1} \left| D^\ell f[v_1, \cdots, v_\ell] \right|. \quad (4.7)$$
and $|f|_{C^k_\rho(\Omega)} = |f|_{\Omega}$. This definition is scale-invariant for graph-dilation: if $f_\rho(y) = \rho^{-1} f(\rho y)$, then $|f|_{C^k_\rho} = |f_\rho|_{C^k_1}$. Given an affine $k$-plane $L$, and a function $f : \Omega \subset L \rightarrow L^\perp$,

$$\text{graph}_\Omega(f) = \{x + f(x) : x \in \Omega\}. \quad (4.8)$$

We require the following effective notion of linear-independence for a collection of points.

**Definition 4.1.** We say a collection of points $p_0, \ldots, p_k$ are in $\rho$-general position if, for each $i$, we have $p_i \not\in B_\rho(< p_0, \ldots, p_{i-1} >)$. Here by $< p_0, \ldots, p_{i-1} >$ we mean the affine $(i-1)$-space containing these points.

The $\mu$-center of mass of $B_r(x)$ is defined to be

$$\frac{1}{\mu(B_r(x))} \int_{B_r(x)} z \, d\mu(z). \quad (4.9)$$

Let define here a generalized $\mu$-center of mass, which captures a crucial property of the center of mass even if $\mu(B_r(x)) = \infty$. We beg the reader’s patience in justifying the well-definedness of Definition 4.2 until Proposition 4.10.

**Definition 4.2.** We define the generalized $\mu$-center of mass $X_{x,r}$ of a ball $B_r(x)$ as follows. If $\mu(B_r(x)) < \infty$, let

$$X_{x,r} = \frac{1}{\mu(B_r(x))} \int_{B_r(x)} z \, d\mu(z) \quad (4.10)$$

be the usual center of mass. If $\mu(B_r(x)) = \infty$, we let $X_{x,r}$ be any point in the intersection

$$\overline{B_r(x)} \cap \bigcap \left\{ \text{affine } V^k : \int_{B_r(x)} d(z, V)^2 \, d\mu(z) < \infty \right\}. \quad (4.11)$$

We now define our fundamental notion of good and bad ball. This definition depends on a choice of $\rho$, $m$, and is of course specific to some measure $\mu$. Usually the $\rho$, $m$, $\mu$ will be unambiguous from the context, but if not we will be explicit.

**Definition 4.3.** We say a ball $B_r(x)$ is good if there are points $y_0, \ldots, y_k \in B_r(x)$, such that for each $i$

$$\mu(B_{\rho r}(y_i)) \geq m(\rho r)^k, \quad (4.12)$$

and if $Y_i$ is the generalized $\mu$-center of mass of $B_{\rho r}(y_i)$, then the $\{Y_i\}$ are in $\rho r$-general position. We say $B_r(x)$ is bad if it is not good.

In simple terms, a good ball has large mass which is spread on some affine $k$-plane. The reason we care is because $\beta$-numbers control the distance between centers of mass, and $L^2$-best planes (Proposition 4.11). If we have lots of mass nicely spread out over a $k$-plane, we can effectively control tilting between different $L^2$-best planes.
4.2. **Basic lemmas.** We collect a few basic results used throughout the paper. First, we require the following standard measure-theoretic result.

**Lemma 4.4.** Let $\mu$ be a non-negative Borel-regular measure.

A) Let $A_1 \subset A_2$, and suppose
\[
\Theta^{s,k}(\mu, A_2, x) \geq t \text{ for } \mu\text{-a.e. } x \in A_1.
\]
Then
\[
t \mathcal{H}^k(A_1) \leq \mu(A_2).
\]

B) Suppose $\mu$ is finite, and for some $A$ we know instead that
\[
\Theta^{s,k}(\mu, x) \geq t \text{ for } \mu\text{-a.e. } x \in A,
\]
then
\[
t \mathcal{H}^k(A) \leq \mu(A).
\]

**Proof.** Let us prove A). The proof is given in [Sim83], but for the convenience of the reader we reproduce it here. We can assume $\mu(A_2) < \infty$ and $t > 0$. Take $\delta > 0$, and $0 < \tau < t$.

By assumption, the collection of balls
\[
\{ B_s(x) : x \in A_1, \quad s < \delta, \quad \text{and} \quad \mu(B_s(x) \cap A_2) > \tau \omega_k \rho^k \}
\]
covers $A_1$ finely. Therefore we can choose a disjoint Vitali subcollection $\{ B_s(x_i) \}_i$, so that
\[
A_1 \subset \bigcup_{i=1}^N B_s(x_i) \cup \bigcup_{i=N+1}^\infty B_s(x_i) \quad \forall N.
\]
We have
\[
\tau \sum_i \omega_k s_i \leq \mu(\cup_i B_s(x_i) \cap A_2) \leq \mu(A_2) < \infty.
\]

We then calculate, for any $N$,
\[
\tau \mathcal{H}^k_{5\delta}(A_1) \leq \tau \sum_{i=1}^N \omega_k s_i^k + \tau \sum_{i=N+1}^\infty \omega_k s_i^k \leq \mu(A_2) + o(1) \quad \text{as } N \to \infty.
\]
Taking $N \to \infty$, $\delta \to 0$, then $\tau \to t$ proves part A).

Let us prove part B). Since $\mu$ is finite, we can choose an open $U \supset A$ so that $\mu(U) \leq \mu(A) + \epsilon$. Then taking $A_1 = A$, and $A_2 = U$, we can apply the first part of the Lemma to deduce
\[
t \mathcal{H}^k(A) \leq \mu(A) + \epsilon.
\]
Now take $\epsilon \to 0$. □

Next it is worth pointing out that $\beta_k^k(x, r)$ is lower-semicontinuous in $(x, r)$. This implies in particular that pointwise bounds on the $\beta$-numbers or summed-$\beta$-numbers in a set can always be extended to the closure.
Lemma 4.5. Suppose \( \beta_{\mu,\beta}(x,r)^2 < \infty \). If \( x_i \to x \) and \( r_i \to r > 0 \), then
\[
\beta_{\mu}(x,r)^2 \leq \liminf_i \beta_{\mu}(x_i,r_i)^2.
\] (4.23)

Proof. We can assume \( \beta_{\mu}(x,r) > 0 \), otherwise the statement is vacuous, so in particular \( \mu B_r(x) > \bar{c}_\beta r^K \).

Suppose, towards a contradiction, we have a sequence \( r_i \to r \), and \( x_i \to x \), so that
\[
\liminf_i \beta_{\mu}(x_i,r_i) < \beta_{\mu}(x,r).
\] (4.24)

First of all, since \( 1_{B_r(x)} \leq \liminf_{i \to \infty} 1_{B_{r_i}(x_i)} \), by Fatou’s lemma we have
\[
\liminf_i \mu B_{r_i}(x_i) > \bar{c}_\beta r^K.
\] (4.25)

Moreover, let \( V_i = V(x,r_i) \) and \( V = V(x,r) \). We can assume, by passing to a subsequence, that \( V_i \to W \) for some \( W \). Since
\[
d(z,V_i)^2 1_{B_{r_i}(x_i)} \to d(z,W)^2 1_A \quad \forall z,
\] (4.26)

for some set \( A \supset B_r(x) \), we have again by Fatou that
\[
r^{-k-2} \int_{B_r(x)} d(z,W)^2 d\mu(z) \leq \liminf_i \beta_{\mu}(x_i,r_i) < \beta_{\mu}(x,r),
\] (4.27)

which contradicts our definition of \( V \). \[ \square \]

Remark 4.6. Another observation which will prove useful is that we can replace the Dini-type condition \( \int \beta_{\frac{dr}{r}} \) with a sum over scales. By monotonicity of \( \beta \), we have
\[
\sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq r/2} \beta_{\mu,2^\alpha \tilde{\epsilon}}(x,2^\alpha)^2 \leq 2^{k+3} \int_0^r \beta_{\mu,2^\alpha \tilde{\epsilon}}(x,s)^2 ds/s,
\] (4.28)

and conversely
\[
\int_0^r \beta_{\mu,2^\alpha \tilde{\epsilon}}(x,s)^2 ds/s \leq 2^{k+3} \sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq r} \beta_{\mu,2^\alpha \tilde{\epsilon}}(x,2^\alpha)^2.
\] (4.29)

As discussed in Definition 4.2 we will be using the (generalized) center of mass extensively to control tilting between \( L^2 \)-best planes. We need two preparatory Lemmas, which allow us to control \( k \)-plane Grassmanian distances by considering only a choice of points in general position.

Lemma 4.7. Suppose \( x_0, \ldots, x_k \) in \( B_1 \) are in \( \rho \)-general position, and there is an affine space \( W \) so that \( d(x_i,W) \leq \delta \) for each \( i \). Then \( \langle x_0, \ldots, x_k \rangle \cap B_1 \subset B_{c(n,\rho)\delta}(W) \).

Proof. By Gram-Schmidt orthogonalization, we can write any \( y \in \langle x_0, \ldots, x_k \rangle \cap B_1 \) as
\[
y = x_0 + \sum_{i=1}^k a_i(x_i-x_0),
\] (4.30)

with \( |a_i| \leq c(n,\rho) \). Let \( w_i \in W \) with \( |x_i-w_i| \leq \delta \), then we can estimate
\[
|y_i - w_0 - \sum a_i(w_i - w_0)| \leq |x_0 - w_0| + \sum |a_i| \left(|x_i - w_i| + |x_0 - w_0| \right) \leq c(n,\rho)\delta,
\] (4.31)
as claimed. \[ \square \]
In the following Lemma the fact that both planes have the same dimension is crucial.

**Lemma 4.8.** Let $V, W$ be affine $k$-spaces, such that $V \cap B_2 \subset B_\delta(W)$, and $V \cap B_1 \neq \emptyset$. Then $d_H(V \cap B_2, W \cap B_2) \leq c(n)\delta$.

**Proof.** We can assume $\delta \leq \delta_0(n)$. The assumptions imply $V \cap B_2 \subset B_{c(n)\delta}(W \cap B_2)$. Choose $x_0 \in V$ to minimize $|x_0|$. Now let $x_i \in V$ be chosen so that

$$|x_i - x_0| = 1, \quad <x_i - x_0, x_j - x_0> = 0. \quad (4.32)$$

Let $y_i$ be a point in $W$ with $d(x_i, y_i) \leq \delta$. Then

$$|y_i - y_0| \geq 1 - 2\delta, \quad |<y_i - y_0, y_j - y_0>| \leq 8\delta. \quad (4.33)$$

Hence, for $\delta \leq \delta_0(n)$, the $y_i$ are in $1/2$-general position in $W$. One can see this using the Gram-Schmidt orthogonalization process.

For any vector $w \in W \cap B_2$, since $y_i \in B_{2+c(n)\delta} \cap W$, we have

$$w = y_0 + \sum_i a_i(y_i - y_0), \quad |a_i| \leq c(n), \quad (4.34)$$

and therefore $d(w, V) \leq c(n)\delta$. Since $V \cap B_1 \neq \emptyset$, we have $W \subset B_{c(n)\delta}(V \cap B_2)$.

Our approximating maps $\sigma_i$ in the good tree are locally perturbations of affine projections. We need a standard lemma, due in this formulation to L. Simon, which tracks how graphs change under almost-projections.

**Lemma 4.9** (“Squash Lemma” [Sim]). Let $L, \tilde{L}$ be an affine $k$-planes, with $L \cap B_{3/2} \neq \emptyset$. Let $G \subset \mathbb{R}^n$ be a graph over $L$:

$$G = \text{graph}_{\Omega \subset L} f, \quad |f| + |Df| \leq 1, \quad B_{5/2} \cap L \subset \Omega \subset L. \quad (4.35)$$

Consider the mapping $\Phi : B_3^n \to \mathbb{R}^n$ defined by

$$\Phi(x) = m + p_{\tilde{L}}(x - m) + e(x), \quad (4.36)$$

where $p_{\tilde{L}}$ is the linear projection operator onto $\tilde{L}$.

Suppose

$$d(m, \tilde{L}) \leq \epsilon, \quad d_H(L \cap B_3, \tilde{L} \cap B_3) \leq \epsilon, \quad \text{and} \quad |e| + |De| \leq \epsilon \text{ on } B_{5/2}. \quad (4.37)$$

Then provided $\epsilon \leq \epsilon_0(n)$, we have

$$\Phi(G \cap B_{5/2}) = \text{graph}_{U \subset \tilde{L}} \tilde{f}, \quad |\tilde{f}| + |D\tilde{f}| \leq 8\epsilon, \quad B_2 \cap \tilde{L} \subset U \subset \tilde{L}. \quad (4.38)$$

**Proof.** For ease of notation let us write $\tilde{p}, \tilde{p}^\perp, p$ for the linear projections onto $\tilde{L}, \tilde{L}^\perp, L$, respectively.

Write $L = y + L_0$, and $\tilde{L} = \tilde{y} + \tilde{L}_0$, where $L_0$ and $\tilde{L}_0$ are linear subspaces. Choose ON bases $\{e_i\}_i$ of $L_0$, and $\{\tilde{e}_i\}_i$ of $\tilde{L}_0$, so that $\tilde{p}(e_i) = \lambda_i \tilde{e}_i$. By assumption,

$$|\lambda_i - 1| \leq 2d_G(L, \tilde{L}) \leq 6\epsilon. \quad (4.39)$$
Define \( h : \Omega \to \bar{L} \) by setting
\[
h(x) = \tilde{y} + \tilde{p}(x - \tilde{y}) + \tilde{p}(f(x)) + \tilde{p}(e(x + f(x)))
\] (4.40)
\[
\equiv \tilde{p}^\perp(\tilde{y}) + \tilde{p}(x) + (\tilde{p} - p)(f(x)) + \tilde{p}(e(x + f(x)) + \tilde{p}(e(x + f(x))))
\] (4.41)
so that
\[
\Phi(x + f(x)) = h(x) + \tilde{p}^\perp(e(x + f(x))) + \tilde{p}^\perp(m - \tilde{y}).
\] (4.42)
For \( x \in L \cap B_{5/2} \), we have,
\[
|h(x) - x| \leq |\tilde{y} + \tilde{p}(x - \tilde{y})| + |\tilde{p} - p||f(x)| + |e(x + f(x))|
\] (4.43)
\[
\leq d_H(L \cap B_3, \bar{L} \cap B_3) + 2d_G(L, \bar{L}) + \epsilon
\] (4.44)
\[
\leq 8\epsilon,
\] (4.45)
and similarly,
\[
|Dh(x) - \tilde{p}| \leq |\tilde{p} - p||Df| + |\tilde{p}||De||I + Df|
\] (4.46)
\[
\leq 8\epsilon.
\] (4.47)
Let us identify \( L, \bar{L} \) (and hence \( L_0, \bar{L}_0 \)) with \( \mathbb{R}^k \) via the bases \( \{e_i\}_i, \{\tilde{e}_i\}_i \). Under this identification, we have
\[
|Dh - I| \leq |Dh - \tilde{p}| + |\tilde{p} - I| \leq 20\epsilon \quad \text{on} \quad B_{5/2}^k.
\] (4.48)
Therefore, provided \( \epsilon \leq \epsilon_0(n) \), \( h \) is a diffeomorphism from \( B_{5/2} \cap L \) onto its image \( U \) in \( \bar{L} \), with \( |Dh^{-1}| \leq 2 \). From (4.45) we can say \( U \supset B_2 \cap \bar{L} \).
Define
\[
\tilde{f}(z) = \tilde{p}^\perp(m - \tilde{y}) + \tilde{p}^\perp(e(h^{-1}(z) + f(h^{-1}(z))))
\] (4.49)
By assumption and (4.47), we have for \( z \in U \),
\[
|\tilde{f}| \leq d(m, \bar{L}) + |e| \leq 2\epsilon, \quad |D\tilde{f}| \leq |De||I + Df||Dh^{-1}| \leq 4\epsilon.
\] (4.50)
Since we can write
\[
\Phi(x + f(x)) = h(x) + \tilde{f}(h(x)),
\] (4.51)
this completes the proof of Lemma 4.9. \( \square \)

4.3. Generalized center of mass. We prove well-definedness of the generalized \( \mu \)-center of mass, and the relation it has with \( L^2 \)-best planes.

**Proposition 4.10.** Suppose that \( \mu(B_r(x)) = \infty \). Then there exists \( X \in \overline{B_r(x)} \) such that \( X \in V^k \) for all affine \( k \)-planes \( V^k \) such that \( \int_{\overline{B_r(x)}} d(x, \bar{V})^2d\mu(x) < \infty \).

In particular, the notion of generalized \( \mu \)-center of mass (Definition 4.2) is well-defined.
Proof. Consider the collection \( \{ V_\alpha \}_{\alpha \in A} \) of all \( k \)-affine spaces \( V_\alpha \) such that
\[
\int_{B_r(x)} d(x, V)^2 d\mu(x) < \infty.
\] (4.52)
It’s easy to see that there must be some finite subcollection \( V_0, \ldots, V_{k+1} \) so that
\[
W := \bigcap_{\alpha} V_\alpha = \bigcap_{i=0}^{k+1} V_i.
\] (4.53)
Suppose, towards a contradiction, that \( B_r(x) \cap W = \emptyset \). Then it must follow
\[
\overline{B_r(x)} \cap \bigcap_{i=0}^{k+1} B_\epsilon(V_i) = \emptyset
\] (4.54)
for some \( \epsilon > 0 \), because \( \overline{B_r(x)} \) and each \( V_i \) are closed.
But then, using the definition of \( V_i \), we have
\[
\mu(B_r(x)) = \mu \left( B_r(x) \cap (\mathbb{R}^n \setminus \bigcup_{i=0}^{k+1} B_\epsilon(V_i)) \right)
\]
\[
\leq \sum_{i=0}^{k+1} \mu(B_r(x) \cap (\mathbb{R}^n \setminus B_\epsilon(V_i))
\]
\[
\leq \sum_{i=0}^{k+1} \epsilon^{-2} \int_{B_r(x)} d(z, V_i)^2 d\mu(z)
\]
\[
< \infty,
\] (4.58)
which is a contradiction. In the penultimate inequality we used that \( k < n \). \( \square \)

Recall that we defined in Definition 4.2 the generalized \( \mu \)-center of mass to be any \( X \) as in Proposition 4.10 whenever a ball has infinite \( \mu \)-mass. The essential characteristics of the generalized center of mass is that its distance from best approximating planes is controlled by the \( \beta \)-numbers. More specifically:

**Proposition 4.11.** Suppose \( B_r(y) \subset B_R(x) \), and \( \mu B_R(x) > \bar{c}_\beta R^k \). Let \( Y \) be the generalized center of mass for \( B_r(y) \). Then
\[
\left( d(Y, V(x, R)) \right)^2 \leq \frac{R^{k+2}}{\mu(B_r(y))} \frac{\beta^k}{\mu \cdot \epsilon \beta}(x, R)^2.
\] (4.59)

**Proof.** We can assume \( \beta^k(x, R) < \infty \), otherwise there is nothing to show. By the previous proposition, if \( \mu B_r(y) = \infty \), then \( Y \in V(x, R) \). Otherwise, by Jensen’s inequality, we calculate
\[
d(Y, V(x, R))^2 \leq \frac{1}{\mu(B_r(y))} \int_{B_r(y)} d(z, V(x, R))^2 d\mu(z)
\]
\[
\leq \frac{R^{k+2}}{\mu(B_r(y))} R^{-k-2} \int_{B_R(x)} d(z, V(x, R))^2 d\mu(z).
\] (4.60) \( \square \)
4.4. **Tilting and packing control.** We are ready to prove the original dichotomy described in the Proof Outline: in good balls the $\beta$-numbers control tilting, and in bad balls we have good packing estimates.

Let us first give some intuition about why it is necessary to have mass effectively spread out to say anything the $L^2$-best plane. As a guiding example, consider the measure $\mu_1 = H^{k-1} \cdot L$, where $L$ is a linear $(k - 1)$-dimensional space. In this case, it is clear that $\beta_{\mu_1}^k(x, r) = 0$ for all $x$ and $r$. However, the notion of best approximating $k$-dimensional plane is not even uniquely defined.

We can make this example more precise by slightly modifying the measure $\mu$. Consider

$$\mu_2 = \mu_1 + \delta_{x_1} + \chi \delta_{x_2},$$

where $|x_1| = d(x_1, L) = 1/2$ and $|x_2| = d(x_2, < L, x_1 >) = 1/10$. In other words, $x_1$ is orthogonal to $L$ and $x_2$ is orthogonal to the subspace spanned by $x_1$ and $L$.

It is clear that $V(0, 1/4) < L, x_2 >$, and $\beta_{\mu_2}^k(0, 1/4) = 0$, while we have the simple estimate

$$\beta_{\mu_2}^k(0, 1) \leq \chi/10^2.$$  \hfill (4.62)

Moreover, it is easy to see that $V(0, 1)$ can be made arbitrarily close to $< L, x_1 >$ by choosing $\chi$ small enough. This implies that the distance between $V(0, 1)$ and $V(0, 1/4)$ cannot be uniformly bounded by $\beta(0, 1) + \beta(0, 1/4)$.

The issue in both these examples is that mass is not sufficiently $k$-dimensional at scale $\approx 1$ for $\beta(0, 1)$ and $\beta(0, 1/4)$ to sense the right plane. This is why we can only hope to get tilting control in good balls. In bad balls, mass will be sufficiently $(k - 1)$-dimensional that we have good packing.

We start by showing this straightforward dichotomy:

**Lemma 4.12.** Given a non-negative Borel measure $\mu$, and a choice of $m, \rho$, then at least one of the following occurs:

A) There are points $x_0, \ldots, x_k \in B_1$ such that for each $i$ we have

$$\mu B_\rho(x_i) > m \rho^k, \quad \text{and} \quad X_i \notin B_\rho(< X_0, \ldots, X_{i-1} >),$$

where $X_i$ is the generalized center of mass of $B_\rho(x_i)$.

B) There is an affine $W^{k-1}$ such that

$$\mu(B_1 \setminus B_{2\rho}(W^{k-1})) \leq 2^n \rho^{k-n} m.$$  \hfill (4.64)

In particular, if $B_1$ is bad then (4.64) holds for some $W^{k-1}$.

**Proof.** Let us construct such a collection $x_0, \ldots, x_k$ by induction, and see what happens when it fails. Take $i \geq 0$, and suppose we’ve obtained $x_0, \ldots, x_{i-1}$ satisfying (4.63). Write $W = < X_0, \ldots, X_{i-1} >$.

Suppose we can find an $x_i \in B_1 \setminus B_{2\rho}(W)$ with $\mu B_\rho(x_i) > m \rho^k$. Then we have

$$d(X_i, V) \geq d(x_i, V) - \rho \geq \rho.$$  \hfill (4.65)

And therefore we have constructed an $x_i$.

Otherwise, necessarily

$$y \in B_1 \setminus B_{2\rho}(W) \implies \mu B_\rho(y) \leq m \rho^k.$$  \hfill (4.66)
Take a maximal \( \rho \)-net \( \{ y_j \}_j \) of \( B_1 \setminus B_{2\rho}(W) \). Then the balls \( \{ B_{\rho}(y_j) \}_j \) cover \( B_1 \setminus B_{2\rho}(W) \), and
\[
\# \{ y_j \}_j \leq 2^n \rho^{-n}
\] (4.67)
and \( \mu_{B_{\rho}(y_j)} \leq m \rho^k \) for each \( j \). Since \( \dim V = i \leq k - 1 \), the lemma is established. \( \square \)

An easy consequence of Lemma 4.12 is the bad-ball packing.

**Theorem 4.13.** Let \( B_1 \) be a bad ball, and let \( W^{k-1} \) be the \((k-1)\)-plane of Lemma 4.12. Then

A) If \( \{ y_i \}_i \) is a maximal \( 2\rho/5 \)-net in \( B_{5\rho}(W^{k-1}) \), then
\[
\# \{ y_i \}_i \rho^k \leq c_1(n) \rho.
\] (4.68)

B) There is an \( m_0(n, \rho) > 0 \) so that if \( m = m_0(n, \rho) \epsilon \), then
\[
\mu(B_1 \setminus B_{2\rho}(W^{k-1})) \leq \epsilon/2.
\] (4.69)

C) If we know
\[
\mu(B_{\rho}(y)) \leq \Gamma \rho^k \forall y \in B_1.
\] (4.70)

Then provided \( \rho \leq \rho_0(n, \epsilon, \Gamma) \) and \( m = m_0(n, \rho) \epsilon \) as in part A), we have
\[
\mu(B_1) \leq \epsilon.
\] (4.71)

**Proof.** Part A) follows because
\[
\# \{ y_i \}_i \leq \frac{\omega_{n-k+1} \omega_{k-1} \cdot 30^n}{\omega_n} \rho^{-k+1}.
\] (4.72)
Part B) follows by choosing \( m_0 = \rho^{n-k} 2^{-n}/2 \). In part C), the assumption and parts A), B) imply
\[
\mu(B_1) \leq 2^n \rho^{k-n} m + c_1(n) \Gamma \rho.
\] (4.73)
\( \square \)

Let us now prove tilting control in good balls.

**Theorem 4.14 (Good-ball tilting).** Suppose \( x \in B_7 \), and \( B_{\rho}(x) \) is a good ball, and that \( \mu_{B_8} > \bar{\epsilon}_\beta 8^k \), \( \mu_{B_{8\rho}(x)} > \bar{\epsilon}_\beta (8\rho)^k \). Then we have that
\[
d_H(V(x, 8\rho) \cap B_{8\rho}(x), V(0, 8) \cap B_{8\rho}(x))^2 \leq \frac{c(n, \rho)}{m} \left( \beta^{k}(x, 8\rho)^2 + \beta^{k}_\mu(0, 8)^2 \right).
\] (4.74)
Therefore we get that
\[
d_G(V(x, 8\rho), V(0, 8))^2 \leq \frac{c(n, \rho)}{m} \left( \beta^{k}(x, 8\rho)^2 + \beta^{k}_\mu(0, 8)^2 \right).
\] (4.75)

**Proof.** In the below any \( c \) will depend only on \( n, \rho \). By virtue of being a good ball, we have \( y_0, \ldots, y_k \in B_{\rho}(x) \) so that
\[
\mu_{B_{\rho^2}(y_i)} > m \rho^{2k},
\] (4.76)
and the generalized center-of-masses \( Y_i \) are in \( \rho^2 \)-general position.
By Proposition 4.11 we have
\[
d(Y_i, V(x, 8\rho))^2 \leq \frac{(8\rho)^{k+2}}{m^2 \rho^{2k}} \beta^k_\mu(x, 8\rho)^2 \leq \frac{c}{m} \beta^k_\mu(x, 8\rho)^2,
\]
\[
d(Y_i, V(0, 8))^2 \leq \frac{8^{k+2}}{m^2 \rho^{2k}} \beta^k_\mu(0, 8)^2 \leq \frac{c}{m} \beta^k_\mu(0, 8)^2.
\]
(4.77)

Therefore, letting \( L = \langle Y_0, \ldots, Y_k \rangle \) we have by Lemma 4.7 (at scale \( 8\rho \)) that
\[
L \cap B_{8\rho}(x) \subset B_{cm^{-1/2}\beta^k_\mu(x, 8\rho)}(V(x, 8\rho)),
\]
\[
L \cap B_{8\rho}(x) \subset B_{cm^{-1/2}\beta^k_\mu(0, 8)}(V(0, 8)).
\]
(4.78)

Since \( Y_i \in B_{2\rho}(x) \) we have \( L \cap B_{2\rho}(x) \neq \emptyset \). Therefore, we have by Lemma 4.8 (at scale \( 8\rho \))
\[
d_H(V(x, 8\rho) \cap B_{8\rho}(x), V(0, 8) \cap B_{8\rho}(x))^2
\]
\[
\leq 2d_H(L \cap B_{8\rho}(x), V(x, 8\rho) \cap B_{8\rho}(x))^2 + 2d_H(L \cap B_{8\rho}(x), V(0, 8) \cap B_{8\rho}(x))^2
\]
\[
\leq \frac{c}{m} (\beta^k_\mu(x, 8\rho)^2 + \beta^k_\mu(0, 8)^2),
\]
(4.79)
as claimed.

\[\square\]

5. MAIN CONSTRUCTION

We will use the dichotomy of good/bad balls of Definition 4.3 to obtain good estimates as we drop down in scale. In very simple terms, either we having good tilting control, and can obtain volume bounds by a Reifenberg type process; or, in bad balls, we have very good packing control and can compensate for the naive overlaps by choosing our scale sufficiently small.

In this construction, we fix \( m = m_0(n, \rho)M \) as in Theorem 4.13. The construction has a running dependence on \( \rho \), which for Theorem 5.3 will be fixed as \( \rho = \rho(n) \) in (8.1), but in general may be chosen differently (as in Proposition 10.1). However we can always assume \( \rho \leq 1/20 \). For convenience, we can assume \( \rho = 2^{\alpha_0} \) for some \( \alpha_0 \in \mathbb{Z} \), and we use the following notation.

**Definition 5.1.** We write \( r_i = \rho^i \).

Recall that we are given a covering pair \( (\mathcal{C}, r_x) \). We may refer to the \( \{B_{r_y}(y)\}_{y \in \mathcal{C}^+} \) as the original balls. As explained in the outline, our construction is based on refining inductively on scales approximations of the support of \( \mu \). Our refining needs to stop when we hit some of the “original balls” in the covering \( \mathcal{C}^+ \). Here we define this stopping condition in detail.

**Definition 5.2.** A ball \( B_r(x) \) is a stop ball if either \( \mu B_r(x) \leq Mr^k \), or if there is some original ball \( B_{r_y}(y) \) satisfying \( x \in B_{r_y+2r}(y) \) and \( r < r_y \leq r/\rho \).

Our key technical result is the following. From this Theorem 2.6 follows more or less directly.
Theorem 5.3. There are constants $c_{\text{key}}(n)$, $\delta_0(n)$ so that the following holds: Let $M > 0$, $\mu$ be a non-negative Borel-regular measure in $\mathbb{R}^n$ supported in $B_1$, and $(C, r_x)$ a covering pair for $\mu$ with $r_x \leq 1$. Suppose we know
\begin{equation}
\sum_{\alpha \in \mathbb{Z} : r_x < 2^\alpha \leq 2} \beta_{\mu,2,\beta}(x, 2^n)^2 \leq \delta_0(n)^2 M \quad \text{for } \mu - a.e. x \in C.
\end{equation}

Then provided $\bar{\epsilon}_\beta \leq 2^{-k} \omega_k M/c_{\text{key}}(n)$, there is a closed subset $C' \subset C$ which admits the following properties:

A) Packing bound:
\begin{equation}
\mathcal{H}^k(C'_0) + \sum_{x \in C'_x} r_x^k \leq c_{\text{key}}(n),
\end{equation}

and Minkowski estimates
\begin{equation}
r^{k-n}|B_r(C')| \leq c_{\text{key}}(n) \quad \forall 0 < r \leq 1,
\end{equation}

B) Upper measure estimates:
\begin{equation}
\mu\left(B_1 \setminus B_{4r_x}(C')\right) \leq c_{\text{key}}(n)M,
\end{equation}

C) Lower measure estimates:
\begin{equation}
\mu(B_r(x)) \geq \frac{M}{c_{\text{key}}(n)} r^k \quad \forall 4r_x < r \leq 1, \quad \forall x \in C',
\end{equation}

D) Fine-scale packing structure: $C'_0$ is closed, rectifiable, and admits bounds
\begin{equation}
\mathcal{H}^k(C'_0 \cap B_r(x)) \leq c_{\text{key}}(n) r^k \quad \forall x \in B_1, 0 < r \leq 1.
\end{equation}

Remark 5.4. This Theorem only requires the weaker hypothesis on $C$ that $\mu(\mathbb{R}^n \setminus B_{r_x}(C)) = 0$, except in this case we would require (5.1) to hold at every $x \in C$.

In fact the only real reason we define a covering pair by $C \supset \text{spt}\mu$, instead of $\mu(\mathbb{R}^n \setminus B_{r_x}(C)) = 0$, is to make sense of (weak) $L^1$ conditions as in Theorem 2.6. However, in the presence of $L^\infty$ bounds, like in Theorem 5.3, the two notions are essentially the same: using monotonicity of $\beta$ one can always extend $C$ to cover $\text{spt}\mu$, and still satisfy (5.1) up to a factor of $c(k)$.

Let us demonstrate how Theorem 2.6 follows from Theorem 5.3.

Proof of Theorem 2.6 given 5.3. We can of course assume $c_{\text{key}} = \delta_0^{-2}$, by enlarging $c_{\text{key}}$ or shrinking $\delta_0$ as necessary.

Define the set
\begin{equation}
A = \left\{ z \in B_1 : \int_{r_z}^2 \beta^{k}_{\mu,2,\beta}(z,r)^2 dr/r > M \right\}.
\end{equation}

Since $\mu$ is regular, and by (2.4), there is a Borel set $U \supset A$ with $\mu(U) = \mu(A) \leq \Gamma$.

Define the measure
\begin{equation}
\mu' = \mu_\ast(B_1 \setminus U).
\end{equation}

Then $C$ is trivially still a covering pair for $\mu'$.
By monotonicity of $\beta$ in remark 2.3, and our choice of $\mu'$ (supported in $B_1$), we have
\[ \int_{r_x/4}^{4} \beta_{\mu',2,z,r}^k(z,r)^2 dr / r \leq c(k)M \quad \text{for $\mu'$-a.e. } z \in C, \] (5.9)
and therefore, again by monotonicity of $\beta$,
\[ \sum_{r_x/4 \leq 2^n \leq 2} \beta_{\mu',2,c(k)\rho}^k(z,2^n)^2 \leq c(k)M \quad \text{for $\mu'$-a.e. } z \in C. \] (5.10)
Here of course $c(k) \geq 1$. Note we have used the scale bounds (2.3) to drop the factor of 4 above, which will be convenient.

We can suppose the $\epsilon$ of Theorem 2.6 is equal to $M$, as enlarging either only weakens the assumptions, and leaves the conclusions unchanged.

First suppose $\epsilon = M = 0$, so that necessarily $\mu'$ is supported in some $k$-plane $V$. In this case Theorem 2.6 becomes trivial. Set $C'_0 = C_0 \cap V \cap B_2$, and let $C'_+ \subset V \cap B_2$ be (the centers of) a Vitali subcover of
\[ \{ B_{r_x/5}(x) : x \in C_+ \cap V \cap B_2 \}, \] (5.11)
so that balls $\{ B_{r_x/5}(x) \}_{x \in C'_+}$ are disjoint, and
\[ B_{r_x}(C'_+) \supset B_{r_x/5}(C_+ \cap V \cap B_2) \supset C_+ \cap V \cap B_2. \] (5.12)

Necessarily $B_{r_x}(C') \supset \text{spt } \mu'$, and the packing bounds of conclusion A) follow simply because $C' \subset V$, and $\{ B_{r_x/5}(x) \cap V : x \in C'_+ \}$ are disjoint. The measure estimate B) is immediate, the non-collapsing C) becomes vacuous, and conclusion D) follows since $C'_0 \subset V$.

If $\epsilon = M > 0$, then we can apply Theorem 5.3 to $\mu'$, with $c(k)c_{\text{key}}M$ in place of $M$, and covering pair $(C, r_x/4)$.

6. Good Tree

We define precisely the good tree at $B_1(0)$. As explained in the proof outline of Section 2.5, the good tree is a succession of coverings at finer and finer scales, which decompose $B_1(0)$ into a part of controlled measure, a family of bad balls with packing estimates, and a subset of a Lipschitz manifold.

Suppose $B_1(0)$ is a good ball as in Definition 4.3 for a non-negative Borel-regular measure $\mu$, with parameters $\rho$ and $m = m_0(n, \rho)M$. We take $C$ a covering pair for $\mu$, with the property that
\[ B_2(0) \cap B_{r_y}(y) \neq \emptyset \implies r_y \leq 1 \quad \forall y \in C_+. \] (6.1)
In principle we are assuming $r_y \leq 1$, but for technical reasons it is better for tree-chaining (Section 8) to allow for big balls in $C$ that are far away.

We define inductively families of good balls $\{ B_r(y) \}_{y \in G_i}$, bad balls $\{ B_{r_i}(b) \}_{b \in B_i}$, and stop balls $\{ B_{r_i}(s) \}_{s \in S_i}$. At our initial scale $r_0 = 1$, $B_1(0)$ is the only ball, and it is good by assumption. So $G_0 = \{ 0 \}$, and $B_0 = S_0 = \emptyset$. 
Let us now discuss the remainder and excess sets, which are useful in the technical constructions. In each good ball $B_{r_i}(g)$ we have an associated $L^2$-best $k$-plane $L_{i g}^k := V^k_{\mu}(g, 8r_i)$. For technical reasons we take the plane associated with scale $8r_i$ rather than $r_i$. We define the good ball excess set to be

$$E_{ig} = B_{r_i}(g) \setminus B_{r_{i+1}/50}(L_{i g}),$$

and the total excess

$$E_i = \bigcup_{g \in G_i} E_{ig}.$$

The excess sets will have controlled measure.

Define inductively the remainder set to be the bad and stop balls at all previous scales:

$$R_i = \bigcup_{\ell=0}^{i} (B_{r_{\ell}}(S_{\ell}) \cup B_{r_{\ell}}(B_{\ell}))$$

We now move on to the inductive construction of our coverings. Suppose we have defined the good/bad/stop balls down through scale $r_{i-1}$. We define the $r_i$-balls as follows. Let $J_i$ form a maximal $2r_i/5$-net in $B_1 \cap \bigcup_{g \in G_{i-1}} \left( (B_{r_{i-1}}(g) \cap B_{r_{i}/40}(L_{i-1}, g)) \setminus R_{i-1} \right)$, so the balls $\{B_{r_i}(z)\}_{z \in J_i}$ cover (6.5) and the balls $\{B_{r_i/5}(z)\}_{z \in J_i}$ are disjoint. Now define

$$S_i = \{ z \in J_i : B_{r_i}(z) \text{ is a stop ball} \},$$

$$G_i = \{ z \in J_i : B_{r_i}(z) \text{ is not stop, but is good} \},$$

$$B_i = \{ z \in J_i : B_{r_i}(z) \text{ is not stop, but is bad} \}.$$

Evidently $J_i = S_i \cup G_i \cup B_i$.

Let $p_{ig}$ be the linear projection to $L_{i g}$, and $p_{ig}^\perp$ the linear projection to $L_{i g}^\perp$. For each $i$ define the map

$$\sigma_i(x) = x - \sum_{g \in G_i} \phi_{ig}(x)p_{ig}^\perp(x - X_{ig}).$$

Here $X_{ig}$ is the generalized $\mu$-center of mass of $B_{r_i}(g)$, and $\{\phi_{ig}\}_{g \in G_i}$ is a partition of unity subordinate to $\{B_{3r_i}(g)\}_{g \in G_i}$, which satisfies

$$\text{spt}\phi_{ig} \subset B_{3r_i}(g), \quad \sum_{g \in G_i} \phi_{ig} = 1 \text{ on } B_{5r_i/2}(g), \quad |D\phi_{ig}| \leq c(n)/r_i.$$  

We use maps $\sigma_i$ to build inductively a sequence of manifolds $T_i$. First set $T_0 = L_{0,1}$, and then define

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

This completes the good tree inductive construction.

Let us make some elementary remarks concerning the good tree construction.

Remark 6.1. Every good/bad/stop ball is contained in $B_2$, and for any $i$, the balls

$$\{B_{r_i/5}(g) : g \in G_i\} \cup \bigcup_{\ell=0}^{i} \{B_{r_i/5}(x) : x \in B_\ell \cup S_\ell\}$$

(6.12)
are pairwise disjoint.

Remark 6.2. For each $i$, $\sigma_i$ is the identity outside $B_{3r_i}(G_i)$. In particular, since $\rho \leq 1/20$, we have

$$T_i = T_{i-1} \text{ inside } \bigcup_{\ell=0}^{i-1} B_{4r_i/5}(B_\ell \cup S_\ell). \quad (6.13)$$

Definition 6.3. The construction defined above is said to be the good tree rooted at $B_1(0)$, and may be written as $T = T(B_1)$. Given such a tree $T$ we refer to the collection of all good/bad balls by

$$G(T) := \bigcup_i G_i, \quad B(T) := \bigcup_i B_i. \quad (6.14)$$

In a slight abuse of notation, we let $r_g$ and $r_b$ be the associated radius functions for $G(T)$ and $B(T)$. So for example, if $g \in G_i \subset G(T)$, then $r_g = r_i$.

For every stop ball $B_{r_i}(s)$ we have either $\mu B_{r_i}(s) \leq Mr_i^k$, or $s \in B_{r_y+2r_i}(y_s)$ for some choice of $y_s \in C_+$. Define

$$C_+(T) = \bigcup_{i=0}^{\infty} \bigcup_{s \in S_i, \mu B_{r_i}(s) > Mr_i^k} y_s \quad \text{(6.15)}$$

to be a choice of $y$’s arising in this way (so, at most one $y$ per stop ball). Define

$$C_0(T) = \bigcap_i \overline{B_{r_i}(G_i)} \quad \text{(6.16)}$$

In a good tree, we refer to the collection of all bad balls as the tree leaves.

Let us make some remarks concerning the above definition.

Remark 6.4. We have $C_+(T) \subset C_+$. By (6.33), we have $C_0(T) \subset C_0 \setminus B_{r_x}(C_+)$, and from (6.32), $C(T) \subset B_3$.

Remark 6.5. By construction we have the inclusion

$$\bigcup_{i=0}^{\infty} \bigcup_{s \in S_i, \mu B_{r_i}(s) > Mr_i^k} B_{r_i}(s) \subset \bigcup_{y \in C_+(T)} B_{4r_y}(y), \quad \text{(6.17)}$$

and consequently we also have the lower bound

$$\mu B_{4r_y}(y) > \frac{M}{\rho^k} r_y^k \quad \forall y \in C_+(T). \quad \text{(6.18)}$$

The following is our main result about good trees in this section:

**Theorem 6.6 (Good Tree Structure).** There is a $\delta_1(n, \rho)$ so that if $M > 0$, and

$$\sum_{\alpha \in \mathbb{Z} : r_y < 2^{\alpha} \leq 16} \beta^k_{\mu,\rho}(y, 2^{\alpha})^2 \leq \delta^2 \leq \delta_1(n, \rho)^2 M \quad \text{for } \mu\text{-a.e. } y \in C, \quad \text{(6.19)}$$
then provided \( \epsilon_\beta \leq c(k)M \), we have for any \( i \geq 0 \) the packing estimate

\[
\# \{ G_i \} r^k_i + \sum_{\ell=0}^{i} \# \{ B_\ell \cup S_\ell \} r^k_\ell \leq 50^k,
\]

and for any \( i \geq 0 \) the measure estimate

\[
\mu \left[ B_1 \setminus \left( \bigcup_{g \in G_i} B_{r_g}(g) \cup \bigcup_{b \in B_i} \bigcup_{\ell \leq i} B_{r_\ell}(b) \cup \bigcup_{x \in C(T)} B_{4r_x}(x) \right) \right] \leq c(k)M + c(k, \rho)\delta^2.
\]

Moreover, for each \( i \), the manifold \( T_i \cap B_2(0) \) is a \( e^{cM^{-1}\delta^2} \)-biLipschitz to a \( k \)-dimensional disk, and satisfies

\[
B_{r_i}(G_i) \subset B_{2r_i}(T_i \cap B_1) \subset B_{4r_i}(G_i) \cup R_i
\]

There is a fixed Lipschitz manifold \( T_\infty \), also \( e^{cM^{-1}\delta^2} \)-biLipschitz to a disk, so that \( T_i \to T_\infty \) in \( C^{0,\alpha}(B_1) \) (for any \( \alpha < 1 \)) as \( i \to \infty \). Here \( c = c(n, \rho) \).

**Remark 6.7.** Here, as always in this paper, we mean intrinsically bi-Lipschitz, in the sense that the inverse mapping \( T_\infty \to T_0 \) is Lipschitz with respect to the intrinsic distance in \( T_\infty \).

In the language of Definition 6.3, we have

**Corollary 6.8.** Suppose the hypotheses of Theorem 6.6. Then for the associated good-tree \( T \) we have

A) **Tree-leaf/bad-ball packing:**

\[
\sum_{b \in B(T)} r^k_b \leq 50^k.
\]

B) **Original ball packing:**

\[
\sum_{x \in C_+(T)} r^k_x \leq c(n, \rho) \quad \text{and} \quad r^{k-n} |B_r(C(T))| \leq c(n, \rho) \quad \forall 0 < r \leq 1.
\]

C) **Upper measure estimates:**

\[
\mu \left[ B_1 \setminus \left( \bigcup_{b \in B(T)} B_{r_0}(b) \cup B_{4r_0}(C(T)) \right) \right] \leq c(k)M + c(k, \rho)\delta^2,
\]

D) **Lower measure estimates:**

\[
\mu B_r(y) > \frac{M}{c(n, \rho)} r^k \quad \forall y \in C(T), \quad \forall 4r_y < r \leq 1.
\]

E) **Fine-scale packing structure:** \( C_0(T) \) is a closed subset of \( T_\infty \).

For the duration of this Section we assume the hypothesis of Theorem 6.6, and the good-tree construction. Our main goal of this section will now be to prove Theorem 6.6.
6.1. *Inductive hypothesis.* The key properties we wish of our manifolds and coverings in the tree construction are the following. We shall prove them inductively on the scale $i$. Recall from (6.3) and (6.4) the definitions of the excess and remainder sets. We will want the following inductive control:

1. “covering control”:
   \[ B_1 \subset (\cup_{g \in G_i} B_{r_i}(g)) \cup R_i \cup (\cup_{k \leq i-1} E_k), \]  
   \hspace{1cm} (6.27)

2. “radius control”:
   \[ y \in C \text{ and } B_{2r_i} (G_i \cup B_i) \cap B_{r_y}(y) \neq \emptyset \implies r_y \leq r_i. \]  
   \hspace{1cm} (6.28)

3. “graphicality”:
   \[ \forall i, \forall g \in G_i, \begin{cases} T_i \cap B_{2r_i}(g) = \text{graph}_{\Omega_{ig} \subset L_{ig}} f_{ig} \\ |f_{ig}|_{C_i} \leq \Lambda(n, \rho) M^{-1/2} \delta \\ B_{3r_i/2}(g) \cap L_{ig} \subset \Omega_{ig} \subset L_{ig}. \end{cases} \]  
   \hspace{1cm} (\star_i)

(this last condition is really just to ensure $\Omega_{ig} \neq \emptyset$)

Let us now concentrate on proving these properties.

6.2. *Covering control.* The first two properties are the easiest.

**Lemma 6.9** (covering control). *Conditions “covering control” and “radius control” hold for every $i$. *

**Proof.** We prove the first inclusion (6.27). Trivially this holds at $i = 0$. Suppose (6.27) holds at scale $i - 1$. By construction, for each $g \in G_{i-1}$ we have

\[ B_{r_i/40} (L_{i-1},g) \cap B_{r_{i-1}} (g) \subseteq R_{i-1} \cup B_{r_i} (G_i \cup B_i \cup S_i) \]
\[ = \cup B_{r_i} (G_i) \cup R_i. \]  
   \hspace{1cm} (6.29) \hspace{1cm} (6.30)

And therefore

\[ (\cup_{g \in G_{i-1}} B_{r_{i-1}} (g)) \subseteq \cup_{g \in G_{i}} B_{r_i} (g) \cup R_i \cup E_{i-1}. \]  
   \hspace{1cm} (6.31)

By our inductive hypothesis, we’ve therefore proven (6.27) at scale $i$.

We prove “radius control” (6.28). This holds at scale 0 since we assume $r_y \leq 1$ for any $y \in C_+$ with $B_{r_y}(y) \cap B_2$. Suppose (6.28) holds down to scale $i - 1$, and consider a $g \in G_i$ with $g \in B_{r_y+2r_i}(y)$. We can choose a $g' \in G_{i-1}$, so that $g \in B_{r_{i-1}}(g')$.

Since $2r_i < r_{i-1}$, by our inductive hypothesis $r_y \leq r_{i-1}$. Therefore if $r_y > r_i$, by construction $B_{r_i}(g)$ would be a stop ball, a contradiction. \( \square \)

Some key relations which follows from “radius control” are below:

**Corollary 6.10.** *For any $y \in C_+(T)$, with $r_y \leq r_i$, we have*

\[ B_{r_y}(y) \cap B_{2r_{i-1}} (G_{i-1} \cup B_{i-1}) \neq \emptyset. \]  
   \hspace{1cm} (6.32)

Conversely, if $y \in C_+$ and $r_y > r_i$, then

\[ B_{r_y}(y) \cap B_{2r_i} (G_i \cup B_i) = \emptyset. \]  
   \hspace{1cm} (6.33)
For any $y \in C(T)$, we have

$$B_{\frac{r_y}{2^{p_\epsilon}}} (y) \cap B_{r_x} (x) = \emptyset \quad \forall x \in C \text{ such that } r_x > \rho^{-3} r_y.$$  \hspace{1cm} (6.34)

**Proof.** Relation (6.32) is immediate from the construction, and (6.33) follows from “radius control.” We prove (6.34). If $r_y > 0$ choose $r_i$ so that $\rho r_i \leq r_y < r_i$, otherwise take any $r_i > 0$. By relation (6.32), we deduce

$$B_{\frac{r_y}{2^{p_\epsilon}}} (z) \subset B_{2r_{i-2}} (G_{i-2} \cup B_{i-2}).$$  \hspace{1cm} (6.35)

Using (6.33), we have $B_{r_y} (y) \cap B_{r_x} (x) = \emptyset$ for any $x \in C$ with $r_x \geq r_{i-2}$. This proves (6.34).

We note that Lemma 6.9 also gives control over the $\beta$ numbers centered at $G_i$.

**Corollary 6.11.** For any $g \in G_i$, there is a $y \in C \cap B_{2r_i} (g)$ with $r_y \leq r_i$ and

$$\sum_{\alpha \in \mathbb{Z} : r_y < 2^\alpha \leq 16} \beta_{\mu,2,\epsilon,\beta}^k (y, 2^\alpha)^2 \leq \delta^2.$$  \hspace{1cm} (6.36)

In particular, we have

$$\beta (g, 8r_i)^2 \leq c(k) \beta (y, 11r_i)^2 \leq c(k) \delta^2.$$  \hspace{1cm} (6.37)

**Proof.** This is a direct corollary of the construction and Remark 2.3. In particular.

6.3. **Tilting control in good balls.** We can apply theorem 4.14 to obtain Reifenberg-type control between subsequent good balls’ best planes.

**Lemma 6.12.** Take $g' \in B_{2r_{i-1}} (g)$ for some $g' \in G_i$ and $g \in G_{i-1}$. Then for $c = c(n, \rho)$ we have

$$d_H (L_{ig'} \cap B_{8r_i} (g'), L_{i-1, g} \cap B_{8r_i} (g')) \leq cM^{-1/2} (\beta (g', 8r_i) + \beta (g, 8r_{i-1})) r_i.$$  \hspace{1cm} (6.38)

And hence

$$d_G (L_{ig'}, L_{i-1,g}) \leq cM^{-1/2} (\beta (g', 8r_i) + \beta (g, 8r_{i-1}))$$  \hspace{1cm} (6.39)

$$d (X_{ig'}, L_{i-1,g}) \leq cM^{-1/2} \beta (g, 8r_{i-1}) r_i.$$  \hspace{1cm} (6.40)

Similarly, if $g'' \in G_i$ satisfies $|g'' - g'| < 7r_i$, then

$$d_G (L_{ig'}, L_{ig''}) \leq cM^{-1/2} (\beta (g', 8r_i) + 2\beta (g, 8r_{i-1}) + \beta (g'', 8r_{i-1}))$$  \hspace{1cm} (6.41)

$$d (X_{ig''}, L_{ig'}) \leq cM^{-1/2} \beta (g', 8r_i) r_i.$$  \hspace{1cm} (6.42)

More coarsely, we have the following inequalities:

$$d_G (L_{ig'}, L_{i-1,g}) \leq cM^{-1/2} \delta, \quad d (X_{ig'}, L_{i-1,g}) \leq cM^{-1/2} \delta r_i$$  \hspace{1cm} (6.43)

$$d_G (L_{ig''}, L_{ig'}) \leq cM^{-1/2} \delta, \quad d (X_{ig''}, L_{ig'}) \leq cM^{-1/2} \delta r_i,$$  \hspace{1cm} (6.44)

where $c$ is always a constant $c = c(n, \rho)$. 
Proof. Apply Theorem 4.14 at scale \( r_i \) to deduce first two equations (6.38) and (6.39), recalling our choice of \( m = m_0(n)M \). Equation (6.40) follows by Proposition 4.11. Precisely the same reasoning works with \( g' \) replaced by \( g'' \), so (6.41) follows from the triangle inequality. The estimate in (6.42) is again deduced from Proposition 4.11, since \( B_{r_i}(g'') \subset B_{8r_i}(g') \).

Evidently, the coarse inequalities (6.43) and (6.44) are direct consequences of the previous estimates and the coarse estimate (6.37). \( \square \)

With this control over the tilting of best planes we can prove good \( C^1 \) bounds on our maps \( \sigma_i \) and manifolds \( T_i \). We prove two classes of estimates, following [Sim]. The “coarse estimates” give bounds on \( \sigma_i \) in terms of the graphical bounds of \( T_{i-1} \). From this one can deduce slightly worse graph bounds of \( T_i \) at scale \( i \), however this technique cannot hope to give \( (\ast_i) \) as the bounds would degenerate as \( i \to \infty \).

The “squash estimates” say that provided \( (\ast_{i-1}) \) holds, we get good graphical bounds at scale \( i \) which are independent of \( \Lambda \). This is the key to demonstrating \( (\ast_i) \) inductively. Then by going two scales back, and linking the squash estimates with the coarse estimates, we get bounds on \( \sigma_i \) which are independent of scale (the “refined coarse estimates”).

Lemma 6.13 (“coarse estimates”). Take \( g' \in G_i \cap B_{r_{i-1}}(g) \) for some \( g \in G_{i-1} \). Suppose

\[
\begin{align*}
\begin{cases}
T_{i-1} \cap B_{2r_{i-1}}(g) = \text{graph}_{\Omega_i, g \subset L_{i-1}, g} f \\
|f|_{C^{1}_{i-1}} \leq \gamma.
\end{cases}
\end{align*}
\tag{6.45}
\]

Then

\[
\sup_{B_{3r_i}(g') \cap T_{i-1}} \left( r_i^{-1} |\sigma_i - Id| + |D_{T_{i-1}}\sigma_i - Id| + |D_{T_{i-1}}^\perp \sigma_i - Id||^{1/2} \right)
\leq c(n, \rho)(\gamma + M^{-1/2} \beta(g, 8r_{i-1}))
\leq c_{cs}(n, \rho)(\gamma + M^{-1/2} \delta).
\tag{6.46}
\]

Here \( D_{T_{i-1}} \), \( D_{T_{i-1}}^\perp \) are the projections of \( D_{T_{i-1}} \) to the tangent, normal bundles of \( T_{i-1} \).

Proof. Throughout the proof \( c \) will denote a constant depending only on \( n, \rho \). Recall that

\[
\sigma_i(x) - x = - \sum_{z\in G_i} \phi_{iz}(x) p_{iz}^x(x - X_{iz}),
\tag{6.47}
\]

and consequently for any vector \( V \) we have

\[
D_V(\sigma_i(x) - x) = - \sum_{z\in G_i} (D_V \phi_{iz}(x)) p_{iz}^x(x - X_{iz}) - \sum_{z\in G_i} \phi_{iz}(x) p_{iz}^x(V).
\tag{6.48}
\]

By assumption we know

\[
B_{3r_i}(g') \cap T_{i-1} = \text{graph}_{\Omega_{i-1}, g \subset L_{i-1}, g} f \quad \text{with} \quad |f|_{C^{1}_{i-1}} \leq \gamma.
\tag{6.49}
\]
Now fix an \( x \in B_{3r_i}(g') \cap T_{i-1} \), and write \( x = (\zeta, f(\zeta)) \) for \( \zeta \in L_{i-1, g} \). Suppose \( \phi_{iz}(x) > 0 \), so in particular \( |g' - z| < 6r_i \). We calculate that
\[
|p_{iz}^\perp(x - X_{iz})| \leq |p_{iz}^\perp - p_{iz}^\perp_{1-g}||x - X_{iz}| + |p_{iz}^\perp_{1-g}(x - \zeta)| + |p_{iz}^\perp_{1-g}(\zeta - X_{iz})|
\]
\[
\leq cd_G(L_{iz}, L_{i-1,g})r_i + \gamma r_i + d(X_{iz}, L_{i-1,g})
\]
\[
\leq cM^{-1/2}(\beta(z, 8r_i) + \beta(g, 8r_{i-1}))r_i + \gamma r_{i-1}
\]
(6.50)

There are at most \( c(n) \) \( z \in G_i \) such that \( \phi_{iz}(x) > 0 \). We therefore obtain
\[
|\sigma_i(x) - x| \leq c\gamma r_i + c \left( \beta(g, 8r_{i-1}) + \sum_{z \in G_i \cap B_{6r_i}(g')} \beta(z, 8r_i) \right) r_i
\]
\[
\leq c\gamma r_i + cM^{-1/2}\beta(g, 8r_{i-1})r_i
\]
(6.51)

We bound the gradient term. Fix \( V \) a unit vector in \( T_x T_{i-1} \). The first term of the gradient is bounded in precisely the same way, since \( |D\phi_{ig}| \leq 10/r_i \). The second term can be bounded like
\[
|p_{iz}^\perp(V)| \leq |p_{iz}^\perp - p_{iz}^\perp_{1-g}| + |p_{iz}^\perp_{1-g}(V)|
\]
\[
\leq 2d_G(L_{iz}, L_{i-1,g}) + 2d_G(L_{i-1,g}, T_x T_{i-1})
\]
\[
\leq cM^{-1/2}(\beta(z, 8r_i) + \beta(g, 8r_{i-1})) + c\gamma.
\]
(6.52)

The extra tangent power comes about because, given a unit vector \( W \in T_x T_{i-1}, \) and \( Y \) arbitrary, we have
\[
|W \cdot p_{iz}^\perp(Y)| \leq \left( |p_{iz}^\perp - p_{iz}^\perp_{1-g}| + |p_{iz}^\perp_{1-g}(W)| \right) |p_{iz}^\perp(Y)|
\]
\[
\leq (2d_G(L_{iz}, L_{i-1,g}) + \gamma) |p_{iz}^\perp(Y)|
\]
\[
\leq c \left[ M^{-1/2}\beta(g, 8r_{i-1}) + \gamma \right] |p_{iz}^\perp(Y)|.
\]
(6.53)

**Lemma 6.14** ("squash estimates"). Take \( g' \in G_i \cap B_{r_{i-1}}(g) \) for some \( g \in G_{i-1} \), and suppose \((*)_{i-1}\) holds. Then provided \( \delta \leq \delta_2(n, \rho, \Lambda)M^{1/2} \), we have
\[
T_i \cap B_{2r_i}(g') = \text{graph}_{\Omega_{ig'} \cap L_{ig'}} f
\]
\[
|f|_{C^1} \leq c(n, \rho)M^{-1/2}\beta(g, 8r_{i-1}) \leq c_{eq}(n, \rho)M^{-1/2}\delta
\]
\[
B_{\beta/2}(g') \cap L_{ig'} \subset \Omega_{ig'} \subset L_{ig'}.
\]
(6.54)

**Proof.** As before \( c \) will denote a generic constant depending only on \( n, \rho \). For convenience define \( p = p_{ig} \), \( L = L_{ig} \) and \( X = X_{ig} \). We have
\[
\sigma_i(x) = X + p(x - X) + p^\perp(x - X) - \sum_{z \in G_i} \phi_{iz}(x)p_{iz}^\perp(x - X_{iz})
\]
\[
= X + p(x - X) + e(x).
\]
(6.55)
For \( x \in B_{5r_i/2}(g') \), since \( \sum_{z \in G_i} \phi_{iz}(x) = 1 \), we have

\[
e(x) = \sum_{z \in G_i} \phi_{iz} \left( p^\perp(x - X) - p^\perp_{iz}(x - X_{iz}) \right)
= \sum_{z \in G_i} \phi_{iz} \left( p^\perp(X_{iz} - X) + (p^\perp - p^\perp_{iz})(x - X_{iz}) \right).
\]

(6.56)

If \( \phi_{iz}(x) > 0 \), then \( |g' - z| < 6r_i \), and so by Lemma 6.12 we have

\[
|p^\perp(X_{iz} - X)| \leq cM^{-1/2} \beta(g', 8r_i) r_i.
\]

(6.57)

and

\[
|(p^\perp - p^\perp_{iz})(x - X_{iz})| \leq 2d_G(L, L_{iz})|x - X_{iz}|
\leq cM^{-1/2} \beta(\delta(g, 8r_i - 1)) + \beta(g', 8r_i) + \beta(z, 8r_i) r_i.
\]

(6.58)

Since there are at most \( c(n) \) \( z \in G_i \) for which \( \phi_{iz}(x) > 0 \), we obtain that

\[
|e(x)| \leq cM^{-1/2} \left( \beta(g, 8r_i - 1) + \sum_{z \in B_{6r_i}(g')} \beta(z, 8r_i) \right) r_i \leq cM^{-1/2} \beta(g, 8r_i - 1)
\]

(6.59)

For any unit vector \( V \) we have

\[
D_V e = \sum_{z \in G_i} (D_V \phi_{iz}) \left( p^\perp(X_{iz} - X) - p^\perp_{iz}(x - X_{iz}) \right) - \sum_{z \in G_i} \phi_{iz}(x)(p^\perp - p^\perp_{iz})(V).
\]

(6.60)

Use the same estimates on the first term, with \( |D\phi_{iz}| \leq 10/r_i \). The second term we can estimate as

\[
|(p^\perp - p^\perp_{iz})(V)| \leq 2d_G(L, L_{iz}).
\]

(6.61)

We deduce

\[
|e|_{C^1_1(B_{sr_i/2}(g'))} \leq cM^{-1/2} \beta(g, 8r_i - 1)
\]

(6.62)

By \((*_{i-1})\) we have

\[
B_{3r_i}(g') \cap T_{i-1} = \text{graph}_{\Omega_{i-1,g} \subset L_{i-1,g} f}, \quad |f|_{C^1_{i-1}} \leq \Lambda \delta_2,
\]

(6.63)

for some \( g \in G_{i-1} \) with \( g' \in B_{r_i-1}(g) \), and some \( \Omega_{i-1,g} \neq \emptyset \). From Lemma 6.12, we have

\[
d(X_{ig'}, L_{i-1,g}) \leq c\delta r_i \quad \text{and} \quad d_H(L \cap B_{8r_i}(g'), L_{i-1,g} \cap B_{8r_i}(g')) \leq c\delta r_i.
\]

(6.64)

Therefore, taking \( \delta_2(\Lambda, n, \rho) \) small, we have \( L_{i-1,g} \cap B_{3r_i/2}(g') \neq \emptyset \), and since \( T_{i-1} \) is a topological disc without boundary in \( B_{3r_i}(g') \), we can ensure that \( \Omega_{i-1,g} \supset B_{5/2}(g') \cap L_{i-1,g} \).

Apply the Squash Lemma 4.9 at scale \( r_i \) to deduce

\[
\sigma_i(B_{sr_i/2}(g') \cap T_{i-1}) \cap B_{2r_i}(g') = \text{graph}_{U \subset L f}, \quad |f|_{C^1_{i}} \leq 4|e|_{C^1_{i}} \leq cM^{-1/2} \beta(g, 8r_i - 1),
\]

(6.65)

where \( B_{3r_i/2}(g') \cap L \subset U \subset L \).

By the coarse estimates we know \( |\sigma_i(x) - x| \leq c(1 + \Lambda)\delta_2 r_i \). Thus if \( c(1 + \Lambda)\delta_2 \) is universally small, we can replace \( \sigma_i(B_{sr_i/2}(g') \cap T_{i-1}) \cap B_{2r_i}(g') \) in the above with \( \sigma_i(T_{i-1}) \cap B_{2r_i}(g') \).

\[
\]

By plugging the squash estimates at scale \( i - 1 \) into the coarse estimates at scale \( i \), we obtain
Lemma 6.15 (refined estimates). Suppose (*) holds through $i - 1$. If $i \geq 2$, then given $g' \in G_i$, $g'' \in G_{i-1}$, and $g \in G_{i-2}$ so that $g' \in B_{r_{i-1}}(g'')$ and $g'' \in B_{r_{i-2}}(g)$, then we have the estimate

$$\sup_{B_{3r_i}(g') \cap T_{i-1}} \left( r_i^{-1} |\sigma_i - Id| + |D^{\perp}_{T_{i-1}}(\sigma_i - Id)| + |D^{\perp}_{T_{i-1}}(\sigma_i - Id)|^{1/2} \right)$$

$$\leq c(n, \rho) M^{-1/2} \beta(g, 8r_{i-2})$$

(6.66)

$$\leq c_{rcs}(n, \rho) M^{-1/2} \delta.$$  

(6.67)

As a consequence, we have

$$\sup_{x \in B_{3r_i}(g') \cap T_{i-1}} |D_{T_{i-1}} \sigma_i| \leq 1 + c(n, \rho) M^{-1} \beta(g, 8r_{i-1})^2.$$  

(6.68)

If $i = 1$, then for any $g \in G_1$ we have

$$\sup_{B_{3r_1}(g) \cap T_0} \left( r_1^{-1} |\sigma_1 - Id| + |D^{\perp}_{T_0}(\sigma_1 - Id)| + |D^{\perp}_{T_0}(\sigma_1 - Id)|^{1/2} \right)$$

$$\leq c(n, \rho) M^{-1/2} \beta(0, 8)$$

(6.69)

$$\leq c_{rcs}(n, \rho) M^{-1/2} \delta.$$  

The above estimates suffice to prove “graph control”:

Theorem 6.16. We have $(\ast_i)$ for every $i \geq 0$.

Proof. Take $\Lambda = c_{sq}$, the constant from the squash estimates. Then ensure $\delta_1 \leq \delta_2(n, \rho, c_{sq})$.  

Corollary 6.17. Provided $\delta_1(n, \rho)$ is sufficiently small, $\sigma_i$ is a diffeomorphism onto its image, and $\sigma_i^{-1}$ satisfies the same coarse estimates (Lemma 6.15) as $\sigma_i$.

Proof. Take $g' \in B_{r_{i-1}}(g)$. Hypothesis $(\ast_i)$ implies $T_{i-1} \cap B_{5r_i}(g')$ is $C^1$ close to $L_{i-1,g}$. The coarse estimates 6.15 show the $\sigma_i$ is locally invertible, and $\sigma_i(T_{i-1} B_{3r_i}(g')) \subset B_{5r_i}(g')$.

6.4. Properties of the approximating manifolds. We show that the manifolds $T_i$ are uniformly bi-Lipschitz to $T_0$, and approach in the Hausdorff sense a limit $T_\infty$, sharing the same bi-Lipschitz bound. For $i > j \geq 1$, define the maps

$$\tau_{i,j} = \sigma_i \circ \cdots \circ \sigma_j : T_{j-1} \to T_i.$$  

(6.70)

By Corollary 6.17 each $\tau_{i,j}$ is a diffeomorphism between $T_{j-1}$ and $T_i$.

Lemma 6.18. For every $i \geq j$, we have

$$|\tau_{i,j}(x) - x| \leq c_2 M^{-1/2} \delta r_j,$$  

(6.71)

and

$$|D\tau_{i,j}| \leq e^{c_2 M^{-1/2}}, \quad |D\tau_{i,j}^{-1}| \leq e^{c_2 M^{-1/2}}.$$  

(6.72)

Here $c_2 = c_2(n, \rho)$. 
\textbf{Proof.} The first estimate follows immediately from the coarse estimates: since \(|\sigma_k(x) - x| \leq c_{\text{rcs}} M^{-1/2} \delta r_k\) for every \(x \in T_{k-1}\), we have
\[
|\tau_{i,j}(x) - x| \leq c_{\text{rcs}} M^{-1/2} \sum_{k=j}^{i-1} r_k \leq c_{\text{rcs}} M^{-1/2} \delta r_j \quad \forall x \in T_{i-1}.
\]

In order to prove the gradient estimates, let \(x = x_{j-1} \in T_{j-1}\), and let \(x_k = \tau_{k,j}(x) \in T_k\). By definition of \(\tau_{ij}\), we have
\[
|D_{T_{j-1}} \tau_{ij}(x)| \leq \prod_{k=j}^{i} |D_{T_{k-1}} \sigma_k(x_{k-1})|.
\]
(Here, as usual, the gradient of \(\sigma_k\) is computed on the tangent space of \(T_{k-1}\).)

Let us bound each term in the above product. If \(x_{k-1} \notin B_{3r_k} (G_k)\) then \(|D \sigma_k(x_{k-1})| = 1\) and there is nothing to show. Otherwise, choose \(g_k' \in G_k\) so that \(x_{k-1} \in B_{3r_k} (g_k')\). By Lemma 6.15, there is a \(g_{k-2} \in G_{k-2}\) so that \(g_k' \in B_{r_{k-1} + r_{k-2}} (g_{k-2})\), and
\[
|D_{x_{k-1}} \sigma_k(x_{k-1})| \leq 1 + cM^{-1} \beta(g_{k-2}, 8r_{k-2})^2.
\]

We can assume that \(x_{i-1} \in B_{3r_i} (G_i)\). By assumption and “radius control” (6.28), we can find a \(y \in C \cap B_{2r_i} (g_i')\) with \(r_y < r_i\) and such that (5.1) holds (see also Corollary 6.11). Using the refined squash estimates of Lemma 6.15, we have
\[
|y - g_{k-2}| \leq |y - x_{i-1}| + |x_{i-1} - x_{k-1}| + |x_{k-1} - g_k'| + |g_k' - g_{k-2}|
\leq 4r_i + c\delta r_k + 3r_k + r_{k-1} + r_{k-2}
\leq 2r_{k-2}.
\]
From the monotonicity of \(\beta\) given in Remark 2.3, we deduce that
\[
|D \sigma_k(x_{k-1})| \leq 1 + cM^{-1} \beta(y, 10r_{k-2})^2 \quad \forall k = j, \ldots, i.
\]
Similarly, we have \(|D \sigma_k^{-1}(x_k)| \leq 1 + c\beta(y, 10r_{k-2})^2\) for each \(k\).

Therefore
\[
\log |D_{T_{j-1}} \tau_{ij}(x)| \leq \sum_{k=j}^{i} \log \left(1 + cM^{-1} \beta(y, 10r_{k-2})^2\right)
\leq cM^{-1} \sum_{k=j-2}^{i} \beta(y, 10r_k)^2
\leq cM^{-1} \delta^2.
\]
The bound for \(\tau_{i,j}^{-1}\) follows in precisely the same manner. \(\square\)

The above Lemma shows that the \(\tau_{i,1}\) are uniformly Cauchy in \(C^{\alpha,0}\), and therefore we have a map \(\tau_\infty\) so that \(\tau_{i,1} \to \tau_\infty\) uniformly. Using the gradient bounds above, we deduce \(\tau_\infty\) is bi-Lipschitz wrt the intrinsic distance in \(T_\infty\), with constants
\[
e^{-c_2 M^{-1} \delta^2} |x - y| \leq |\tau_\infty(x) - \tau_\infty(y)| \leq e^{c_2 M^{-1} \delta^2} |x - y|.
\]
Note that this in particular implies that $\tau_\infty$ is a Lipschitz mapping from $\mathbb{R}^k$ to $\mathbb{R}^n$ with Lipschitz constant $e^{c_2M^{-1}\delta^2}$.

Define $T_\infty = \tau_\infty(T_0)$, so $T_\infty$ is bi-Lipschitz to $T_0$. We immediately have that
\[
e^{-ke_2M^{-1}\delta^2} \mathcal{H}^k(T_0 \cap B_3) \leq \mathcal{H}^k(T_\infty \cap B_3) \leq e^{ke_2M^{-1}\delta^2} \mathcal{H}^k(T_0 \cap B_3).
\] (6.84)

6.5. Packing estimate. The construction of the balls and the $T_i$ ensure that for every $i$, the collection of good $r_i$ balls, and bad/stop balls at scales $r_i, r_{i-1}, \ldots, r_1$ form a Vitali covering not only in $\mathbb{R}^n$, but also in the manifold $T_i$. This allows us to obtain the following packing estimates:

Lemma 6.19. We have
\[
\# \{G_i\} r_i^k + \sum_{\ell \leq i} \# \{B_\ell \cup S_\ell\} r_\ell^k \leq 30ke^{c_3M^{-1}\delta^2},
\] (6.85)

provided $\delta_1(n, \rho)$ is sufficiently small. Here $c_3 = c_3(n, \rho)$.

Proof. We first demonstrate that each good/bad/stop $r_i/5$-ball intersects a definite amount of the $T_{i-1}$. More precisely, let $x \in (B_i \cup G_i \cup S_i) \cap B_{r_{i-1}}(g)$ for some $g \in G_{i-1}$. We have $\text{dist}(x, L_{i-1, g}) < r_i/40$ by construction, and hence by graphicality we know
\[
\text{dist}(x, T_{i-1}) \leq \Lambda \delta_1 r_{i-1} + r_i/40 < r_i/30,
\] (6.86)

for $\delta_1$ sufficiently small (depending only on $n, \rho$). Since $B_{r_{i-1}/5}(x) \subset B_{2r_{i-1}}(g)$, and we have graphical control over $T_{i-1}$ in this region, we deduce
\[
\mathcal{H}^k(T_{i-1} \cap B_{r_{i}/5}(x)) \geq \omega_k(r_i/10)^k,
\] (6.87)

again provided $\delta_1(n, \rho)$ is small.

Since the $T_i = T_i$ inside bad/stop balls of larger radii (Remark 6.2), we deduce from above that also any bad/stop $r_\ell/5$-ball cuts through a definite amount of $T_{i-1}$, for any $r_\ell \geq r_i$. The collection of good $r_i$-balls, and the bad/stop balls of scale $r_i, r_{i-1}, \ldots, r_1$, form a Vitali collection (Remark 6.1), and in particular the associated regions $T_{i-1} \cap B_{r_\ell/5}(x_\ell)$ for disjoint subsets of $T_{i-1}$. Combining this with (6.87) gives the volume estimate
\[
\omega_k \# \{G_i\} (r_i/10)^k + \sum_{\ell=1}^i \omega_k \# \{B_\ell \cup S_\ell\} (r_i/10)^k, \tag{6.88}
\]
\[
\leq \mathcal{H}^k(T_{i-1} \cap B_{r_i/5}(G_i)) + \sum_{\ell=1}^i \mathcal{H}^k(T_{i-1} \cap B_{r_\ell/5}(B_\ell \cup S_\ell)), \tag{6.89}
\]
\[
= \mathcal{H}^k(T_{i-1} \cap B_{r_i/5}(G_i)) + \sum_{\ell=1}^i \mathcal{H}^k(T_{i-1} \cap B_{r_\ell/5}(B_\ell \cup S_\ell)), \tag{6.90}
\]
\[
\leq \mathcal{H}^k(T_{i-1} \cap B_3), \tag{6.91}
\]
\[
\leq e^{ke_2M^{-1}\delta^2} \omega_k 30^k. \tag{6.92}
\]

The last inequality follows from (6.84). This proves the Lemma. \qed
6.6. **Bounding excess and Finishing the Proof.** In this section, we bound the measure of the excess set in our construction, and finish the proof of Theorem 6.6.

**Lemma 6.20.** Define

\[ Z_i = B_2 \cap \bigcup_{y \in C \atop r_y \leq r_i} B_{r_y}(y). \tag{6.93} \]

We have for each \( i \)

\[ \sum_{g \in G_i} \mu E_{ig} \leq c(k, \rho) e^{c_3 M^{-1} \delta^2} \int_{T_\infty \cap B_{3r_i}(Z_i)} \beta(z, 10r_i)^2 d\mathcal{H}^k(z). \tag{6.94} \]

Therefore,

\[ \sum_{i} \mu E_i \leq c(k, \rho) e^{2c_3 M^{-1} \delta^2} \delta^2. \tag{6.95} \]

**Proof.** For each \( g \in G_i \), by definition of the \( L^2 \)-best plane we have

\[ \mu E_{ig} \leq c(k, \rho) r_i^k \beta(g, 8r_i)^2. \tag{6.96} \]

As in equation (6.87), we have \( d(g, T_i) \leq r_i/30 \) and \( \mathcal{H}^k(T_i \cap B_{r_i/5}(g)) \geq \omega_k(r_i/10)^k \), provided \( \delta_1(n, \rho) \) is sufficiently small. Therefore we deduce that

\[ \mu E_{ig} \leq c \int_{T_i \cap B_{r_i/5}(g)} \beta(z, 9r_i)^2 d\mathcal{H}^k(z). \tag{6.97} \]

Since the \( r_i/5 \)-balls are disjoint, we can control the sum over \( G_i \) by an integral over \( T_\infty \). Using “radius control” and the fact that any good ball contains some mass, we must have

\[ B_{r_i}(g) \cap Z_i \neq \emptyset \quad \forall g \in G_i. \tag{6.98} \]

We therefore calculate

\[ \sum_{g \in G_i} \mu E_{ig} \leq c \int_{T_i \cap B_{2r_i}(Z_i)} \beta(z, 9r_i)^2 d\mathcal{H}^k(z), \tag{6.99} \]

\[ \leq c e^{c_3 M^{-1} \delta^2} \int_{T_\infty \cap B_{3r_i}(Z_i)} \beta(z, r_i)^2 d\mathcal{H}^k, \tag{6.100} \]

\[ \leq c(k, \rho) e^{c_3 M^{-1} \delta^2} \int_{T_\infty \cap B_{3r_i}(Z_i)} \beta(z, 10r_i)^2 d\mathcal{H}^k. \tag{6.101} \]

In inequalities (6.100), (6.101) we used that

\[ |\tau_\infty^{-1} - Id| \leq c(n, \rho) \delta_1 r_i \leq r_i/2, \tag{6.102} \]

ensuring \( \delta_1(n, \rho) \) is sufficiently small. In (6.100) we also used

\[ |J \tau_\infty^{-1}| \leq e^{c_3 M^{-1} \delta^2}. \tag{6.103} \]

This establishes relation (6.94).
We prove relation (6.95). Suppose \( z \in B_{3r_j}(Z_j) \). Then we can find a \( y \in C \) with \( r_y \leq r_j \), and \( |z - y| < 4r_j \), and for which (5.1) holds. We deduce
\[
\sum_{i \geq j} \beta(z, 10r_i)^2 \leq c(k) \sum_{i \geq j} \beta(y, 16r_i)^2 \leq c(k)\delta^2. \tag{6.104}
\]
Since \( j \) and \( z \) were arbitrary, we have
\[
\sum_i 1_{B_{3r_j}(Z_i)}(z)\beta(z, 10r_i)^2 \leq c(k)\delta^2 \quad \forall z \in B_2. \tag{6.105}
\]
We now use Fubini’s theorem and estimate (6.94) to deduce
\[
\sum_i \mu E_i \leq \sum_i \sum_{g \in \mathcal{G}_i} \mu E_{ig}, \tag{6.106}
\]
\[
\leq c(k, \rho) e^{c_3 M^{-1}\delta^2} \sum_i \int_{T_\infty \cap B_{3r_j}(Z_i)} \beta(z, 10r_i)^2 d\mathcal{H}^k, \tag{6.107}
\]
\[
= c\rho e^{c_3 M^{-1}\delta^2} \int_{T_\infty} \sum_i 1_{B_{3r_j}(Z_i)}(z)\beta(z, 10r_i)^2 d\mathcal{H}^k, \tag{6.108}
\]
\[
\leq c\rho e^{c_3 M^{-1}\delta^2} \mathcal{H}^k(T_\infty \cap B_3), \quad \mathcal{H}^k \tag{6.109}
\]
\[
\leq c(k, \rho)\delta^2 e^{2c_3 M^{-1}\delta^2} 50^k. \tag{6.110}
\]
This completes the proof of Lemma 6.20. \( \square \)

To finish the proof of Theorem 6.6 now ensure that \( \delta_1 \) small enough so that \( 30^k e^{c_3 \delta_1^2} \leq 50^k \). The inclusions (6.22) follow from “covering control” and “graphical control,” ensuring \( \delta_1 \) is sufficiently small. This completes the proof of Theorem 6.6. Corollary 6.8 is simply a restatement of Theorem 6.6.

7. Bad Tree

We define precisely the bad tree construction at \( B_1(0) \). As explained in the proof outline of Section 2.5, the bad tree is essentially similar to the good tree construction, except with good and bad balls swapped, and best \((k - 1)\)-planes instead of best \(k\)-planes. Since we are interested in \(k\)-control this simplifies many of the technical details. The resulting bad tree decomposes \( B_1(0) \) into a part of controlled measure, a family of good balls with packing estimates, and a subset with \( \mathcal{H}^k\)-measure 0.

Suppose \( B_1(0) \) is a bad ball for a non-negative Borel-regular measure \( \mu \), with parameters \( \rho \) and \( m = m_0(n, \rho)M \) as described in Section 5. We take \( C \) a covering pair for \( \mu \), with the property that
\[
B_2(0) \cap B_{r_y}(y) \neq \emptyset \implies r_y \leq 1 \quad \forall y \in C_+. \tag{7.1}
\]
As before, in principle this is the assumption \( r_y \leq 1 \), but for technical reasons it is preferable to state it this way.

We define inductively families of bad balls \( \{B_{r_y}(b)\}_{b \in B_1} \), good balls \( \{B_{r_y}(g)\}_{g \in \mathcal{G}_i} \), and stop balls \( \{B_{r_y}(s)\}_{s \in \mathcal{S}_i} \). At our initial scale \( r_0 = 1 \), \( B_1(0) \) is the only ball, and it is bad by assumption. So \( B_0 = \{0\} \), and \( \mathcal{G}_0 = \mathcal{S}_0 = \emptyset \).
In each bad ball $B_i(b)$, we have an associated $(k-1)$-plane $W_{ib}^{k-1}$ from Lemma 4.12. Just for clarity, $\mu$ may be quite badly approximated by $W_{ib}^{k-1}$ in this context, but this will not be relevant since we have a dimension drop. We define the bad ball excess set to be

$$E_{ib} = B_i(b) \setminus B_{2r_{i+1}}(W_{ib}),$$

(7.2)

and the total excess

$$E_i = \bigcup_{b \in B_i} E_{ib}. $$

(7.3)

The excess sets will have controlled measure.

Define inductively the remainder set to be the good and stop balls at all previous scales:

$$R_i = \bigcup_{\ell = 0}^i (B_{r_i}(S_\ell) \cup B_{r_i}(G_\ell))$$

(7.4)

Suppose we have defined the good/bad/stop balls down through scale $r_{i-1}$. We define the $r_i$-balls as follows. Let $J_i$ form a maximal $2r_i/5$-net in

$$B_1 \cap \left[ \bigcup_{b \in B_{i-1}} (B_{r_{i-1}}(b) \cap B_{5r_i}(W_{i-1,b})) \right] \setminus R_{i-1}. $$

(7.5)

so the balls $\{B_{ri}(z)\}_{z \in J_i}$ cover (7.5), and the balls $\{B_{ri/5}(z)\}_{z \in J_i}$ are disjoint.

Now define

$$S_i = \{z \in J_i : B_{ri}(z) \text{ is a stop ball}\}, $$

(7.6)

$$G_i = \{z \in J_i : B_{ri}(z) \text{ is not stop, but is good}\}, $$

(7.7)

$$B_i = \{z \in J_i : B_{ri}(z) \text{ is not stop, but is bad}\}. $$

(7.8)

Evidently $J_i = S_i \cup G_i \cup B_i$. This completes the bad tree construction.

**Definition 7.1.** The construction defined above is said to be the bad tree rooted at $B_1(0)$, and may be written as $T = T(B_1)$. Given such a tree $T$ we refer to the collection of all good/bad balls by

$$G(T) := \cup_i G_i, \quad B(T) := \cup_i B_i. $$

(7.9)

As before we define the associated radius functions $r_g$, $r_b$ for $G(T)$ and $B(T)$.

For every stop ball $B_{ri}(s)$ we have either $\mu B_{ri}(s) \leq Mr_i^k$, or $s \in B_{ry+2ri}(y_s)$ for some choice of $y_s \in C_+$. Define

$$C_+(T) = \bigcup_{i=0}^{\infty} \bigcup_{s \in S_i, \mu B_{ri}(s) > M r_i^k} y_s$$

(7.10)

to be a choice of $y$’s arising in this way (so, at most one $y$ per stop ball). Define also

$$C_0(T) = \cap_i B_{ri}(B_i). $$

(7.11)

In a bad tree, we refer to the collection of all good balls as the tree leaves.
Remark 7.2. As in the good ball construction, we have

\[ y \in C \text{ and } B_{2r_y}(G_i \cup B_i) \cap B_{r_y}(y) \neq \emptyset \implies r_y \leq r_i. \]  

(7.12)

In particular, Corollary 6.10 and Remark 6.4 are still valid.

Remark 7.3. Remark 6.5 continues to hold for bad trees. In other words, we have the inclusion

\[ \bigcup_{i=0}^{\infty} \bigcup_{\mu B_{r_i}(s) > M r_y k} B_{r_i}(s) \subset \bigcup_{y \in C_+(T)} B_{4 r_y}(y). \]  

(7.13)

And therefore we also have the lower bound

\[ \mu B_{4 r_y}(y) > \frac{M}{\rho^k} r_y^k \quad \forall \ y \in C_+(T). \]  

(7.14)

The following is our main theorem in this section:

**Theorem 7.4 (Bad Tree Structure).** Provided \( c_1(n) \rho \leq 1/2 \), where \( c_1(n) \) as in Theorem 4.13, then for any \( i \geq 0 \), and any \( z \in B_1 \), we have the packing estimate

\[ \sum_{\ell=1}^{i} \# \{ G_\ell \cup B_\ell \cup S_\ell \} \leq 2 c_1(n) \rho, \]  

(7.15)

and the measure estimate

\[ \mu \left[ B_1 \setminus \left( \bigcup_{b \in B_i} B_{r_i}(b) \cup \bigcup_{g \in B_i, \ell \leq i} B_{r_i}(g) \cup \bigcup_{x \in C_+(T), \ r_x \geq r_i} B_{r_x}(x) \right) \right] \leq 3M. \]  

(7.16)

Recall that we have chosen \( m = m_0(n, \rho) M \) as in Theorem 4.13.

Remark 7.5. Note that in comparison to the Good Tree Structure theorem we have both smallness on the estimates, and in (7.15) we have that the sum of all balls on all scales have a small packing estimate, not just the balls in the final covering. This is due to our ability to approximate the support by a subspace of dimension strictly less than \( k \).

In the language of Definition 7.1, Theorem 7.4 says the following.

**Corollary 7.6.** Suppose the hypotheses of Theorem 7.4. Then we have

A) Tree-leaf/good-ball packing, with small constant:

\[ \sum_{g \in \mathcal{G}(T)} r_g^k \leq 2 c_1 \rho. \]  

(7.17)

B) Original ball packing:

\[ \sum_{x \in C_+(T)} r_x^k \leq c(n, \rho), \quad \text{and} \quad r^{k-n} |B_r(C(T))| \leq c(n, \rho) \quad \forall 0 < r \leq 1. \]  

(7.18)
C) **Upper measure estimates:**

\[ \mu \left[ B_1 \setminus \left( \bigcup_{g \in \mathcal{G}(\mathcal{T})} B_{r_g}(g) \cup B_{4r_x}(\mathcal{C}(\mathcal{T})) \right) \right] \leq 3M, \quad (7.19) \]

D) **Lower measure estimates:**

\[ \mu B_r(y) > \frac{M}{c(n, \rho)} r^k \quad \forall y \in \mathcal{C}(\mathcal{T}), \quad \forall 4r_y < r \leq 1. \quad (7.20) \]

E) **Fine-scale packing structure:** \( C_0(\mathcal{T}) \) is closed, with \( \mathcal{H}^k(C_0(\mathcal{T})) = 0 \).

---

**Proof of Theorem 7.4 and Corollary 7.6.** By Theorem 4.13 applied to each scale \( r_\ell, \ell = 0, \ldots, i - 1 \), we have

\[ \sum_{g_i \cup B_i \cup S_i} r_i^k \leq c_1(n) \rho \sum_{B_i} r_{i-1}^k \leq (c_1 \rho)^2 \sum_{B_{i-2}} r_{i-2}^k \leq \ldots \leq (c_1 \rho)^i, \quad (7.21) \]

and therefore we obtain the estimate

\[ \sum_{1 \leq \ell \leq i} \# \{ G_\ell \cup B_\ell \cup S_\ell \} r_\ell^k \leq \sum_{\ell=1}^i (c_1(n) \rho)^\ell \leq 2c_1(n) \rho. \quad (7.22) \]

The last inequality follows since \( c_1 \rho \leq 1/2 \). This proves the first packing estimate of (7.15).

Equation (7.16) follows from the packing estimate (7.15), and our choice of \( m \): in any bad ball \( B_{r_i}(b) \) we have

\[ \mu(E_{ib}) \leq Mr_i^k, \quad (7.23) \]

and by construction we have for any \( i \):

\[ B_1 \subset B_{r_i}(B_i) \cup E_0 \cup \bigcup_{\ell=1}^i (B_{r_i}(G_\ell \cup S_\ell) \cup E_\ell). \quad (7.24) \]

Corollary 7.6 is essentially a restatement of Theorem 7.4. We emphasize that the reason we have small packing is because there are no good balls at scale 0. The estimate \( \mathcal{H}^k(C_0(\mathcal{T})) = 0 \) follows from (7.15) (or (7.21)).

---

**8. Finishing the proof**

We wish to obtain the decomposition of Theorem 5.3, which consists of packing estimates inside the original balls \( (C_+ \) or sets \( (C_0), \) and measure estimates away from these. In any single good or bad tree, this decomposition is achieved away from its leaves, which unfortunately are unavoidable. As described in the proof outline in Section 2.5, we implement the following construction: for argument’s sake suppose \( B_1 \) is good, so we build a good tree downwards from \( B_1 \); at any (necessarily bad) leaf of this tree, we build a bad tree; at any (necessarily good) leaf in any of these bad trees, we build a good tree; repeat, etc...

In this way we can continue refining our decomposition, progressing down in scale either infinitely far, thus hitting some part of \( C_0 \), or until we have no leaves in any tree. By Theorems 6.6 and 7.4 applied to each
tree, it is essentially enough to show we have packing on all leaves among all trees, which is the content of Theorem 8.1.

We make the above idea rigorous in the following construction. Let us fix $\rho(n) < 1/20$ so that

$$2c_1(n)\rho50^k \leq 1/2,$$

where $c_1$ is as in Theorem 4.13. Let us also fix

$$m = m_0(n, \rho) M$$

as in Theorem 4.13, and then $\delta_0 = \delta_1(n, \rho)16^{-k-2}$ as in Theorem 6.6.

Let us inductively define for each $t = 0, 1, 2, \ldots$ a family of balls $\{B_r(f)\}_{f \in F_t}$, where $r_f$ is a radius function which varies with $f \in F_t$. We will see in practice that each ball $B_r(f)$ is a tree leaf arising from some tree construction at a previous stage in the induction. For each $t$, these leaves $F_t$ will be either all good balls or all bad balls, but none of them will be stop balls. Moreover, for $t \geq 1$ each leaf of $F_t$ will be a leaf in some good or bad tree. Let us caution the reader that stage index $t$ is only an upper bound on the scale of the leaves in $F_t$, e.g. if $f \in F_t$ then $r_f \leq r_t$.

At stage $t = 0$, $B_1(0)$ is our only leaf. So $F_0 = \{0\}$, and $r_{f=0} = 1$. Trivially all the leaves of $F_0$ are good or bad balls, since we can assume $B_1(0)$ is not a stop ball.

Suppose we have defined leaves down to stage $t - 1$. The leaves $F_{t-1}$ are (by inductive hypothesis) not stop, and either all good or all bad. Let us assume first they are all good. Then for each $f \in F_{t-1}$, build a good tree $T_{G,f} := T(B_r(f))$, with parameters $\rho$ and $m$ fixed as above. Then we set $F_t$ to be the collection of bad leaves of these good trees:

$$F_t = \bigcup_{f \in F_{t-1}} B(T_{G,f}),$$

and we let the radius function $r_f$ be the associated bad-ball radius function in the trees:

$$r_f|_{F_t} = r_b.$$

Similarly, let us assume all the leaves in $F_{t-1}$ are bad. Then for each $f \in F_{t-1}$, we build a bad tree $T_{B,f} := T(B_r(f))$, with parameters $\rho, m$, and define

$$F_t = \bigcup_{f \in F_{t-1}} G(T_{B,f}), \quad r_f|_{F_t} = r_g.$$  

In either case, we have that the $F_t$ are either all good or all bad (but not stop), and when $t \geq 1$ they arise as leaves in some good or bad tree.

We just need to check the conditions of the good/bad tree constructions are satisfied: the conditions (6.1), (7.1) on the covering pair $C$ are satisfied either by assumption (when $t = 1$, we are building a tree at $B_1$), or by Corollary 6.4 and Remark 7.2, since when $t \geq 1$ every leaf sits on some tree. We are therefore justified in building each good/bad tree.

This completes the tree chaining construction. Our key Theorem is the following leaf-packing estimate:

**Theorem 8.1 (Leaf Packing).** We have the packing bound

$$\sum_{t=0}^{\infty} \sum_{f \in F_t} r_f^k \leq c(k).$$
And if \( f \in \mathcal{F}_t \) then \( r_f \leq \rho^t \).

**Proof.** Suppose, for some fixed \( t \geq 1 \), all the \( \mathcal{F}_t \) are good. Then each \( f \in \mathcal{F}_t \) is a leaf of a bad tree rooted at \( \mathcal{F}_{t-1} \). Therefore, by our choice of \( m \), we can apply Theorem 7.4 (or more precisely, Corollary 7.6 part A) to each bad tree in \( \mathcal{F}_{t-1} \) to obtain
\[
\sum_{f \in \mathcal{F}_t} r_f^k \leq 2c_1\rho \sum_{f \in \mathcal{F}_{t-1}} r_f^k. \quad (8.7)
\]

Suppose now all the \( \mathcal{F}_t \) are bad. Then each leaf in \( \mathcal{F}_t \) lies in some good tree rooted at \( \mathcal{F}_{t-1} \). By our choice of \( \delta_0 \), and because \( \mu \) is supported in \( B_1 \), we can apply Theorem 6.6 (more precisely, Corollary 6.8 part A) to obtain
\[
\sum_{f \in \mathcal{F}_t} r_f^k \leq 50^k \sum_{f \in \mathcal{F}_{t-1}} r_f^k. \quad (8.8)
\]

Since by construction the leaves change type from one stage to the next, we can alternatively iterate estimates (8.7) and (8.8) to obtain a packing bound on the leaves at a given stage:
\[
\sum_{f \in \mathcal{F}_t} r_f^k \leq (2c_1\rho 50^k)^t \sum_{f \in \mathcal{F}_t} r_f^k \leq \cdots \leq (2c_1\rho 50^k)^{t/2} \sum_{f \in \mathcal{F}_t} r_f^k \leq c(k)(2c_1\rho 50^k)^{t/2} \leq c(k)2^{-t/2}. \quad (8.9)
\]

where in the last inequality we used our choice of \( \rho \). The packing estimate (8.6) is now immediate.

The last assertion is a basic consequence of our tree constructions. In any tree rooted at \( B_r(x) \), any leaf has radius \( \leq pr \). Therefore
\[
\max_{f \in \mathcal{F}_t} r_f \leq \rho \max_{f \in \mathcal{F}_{t-1}} r_f \leq \rho^t. \quad (8.10)
\]

\[\square\]

**8.1. Measure and packing decomposition.** We now define the subset \( C' \subset C \), and demonstrate \( C' \) satisfies the required packing and measure bounds of Theorem 5.3. Unfortunately, we cannot use the \( C(T) \)s by themselves to build our decomposition \( C' \). The problem is the following:

In any single tree \( T \), we have measure control only away from \( C(T) \) and the leaves. If in some region \( A \subset B_r^n(0) \) we must alternate the good-/bad-tree construction infinitely-many times, then in every tree this region \( A \) lies inside the leaves. In other words, no single tree will see \( A \) in its measure estimate, and we expect strict inclusion
\[
\bigcup_T C_0(T) \subsetneq C'_0. \quad (8.11)
\]

To capture regions like \( A \), we must consider global collections of balls. (Of course, any \( A \) like this will have \( \mathcal{H}^k \)-measure 0, but could still have very large \( \mu \)-measure.)

To handle this issue, we consider the collection \( Q_1 \) of all good and bad \( r_1 \)-balls, contained in any tree. This “horizontal slice” of the tree-structures will see all the regions with big mass, and makes the global Minkowski bounds of \( C' \) easier to handle. Proving the correct measure and packing bounds for \( Q_1 \) is straightforward consequence of the good-/bad-tree structure Theorems 6.6, 7.4, and the leaf packing Theorem 8.1.
Precisely, \( Q_i \) is defined as follows:

\[
Q_i = \left\{ g \in \mathcal{G}(\mathcal{T}(B_{r_f}(f))) : f \in \bigcup_{t=0}^{i} \mathcal{F}_t, \text{ and } r_g = r_i \right\} \quad (8.12)
\]

\[
\cup \left\{ b \in \mathcal{B}(\mathcal{T}(B_{r_f}(f))) : f \in \bigcup_{t=0}^{i} \mathcal{F}_t, \text{ and } r_b = r_i \right\}. \quad (8.13)
\]

From our tree constructions, it is clear that for each fixed \( A \), we have

\[
Q_A \subset B_{r_i}(Q_i) \quad \forall i \leq A. \quad (8.14)
\]

We now define the subcollection \( C' \). We set \( C'_+ \) to be all big-mass stop balls among all trees (recall Definitions 6.3 and 7.1):

\[
C'_+ = \bigcup_{t=0}^{\infty} \bigcup_{f \in \mathcal{F}_t} C_{i}(\mathcal{T}(B_{r_f}(f))). \quad (8.15)
\]

We let \( C'_0 \) be the region contained in every \( r_i \)-collection of good/bad balls:

\[
C'_0 = \bigcap_{i=0}^{\infty} B_{r_i}(Q_i). \quad (8.16)
\]

Packing in individual trees, combined with the leaf packing (8.6), gives us the following

**Lemma 8.2.** For each \( i \) we have the packing estimate

\[
\#\{Q_i\} r_i^k \leq c(k), \quad (8.17)
\]

and measure estimate

\[
\mu \left[ B_1 \setminus (B_{4r_x}(C'_+) \cup B_{r_i}(Q_i)) \right] \leq c(n)M \quad (8.18)
\]

**Proof.** We first prove the packing estimate. Each good/bad \( r_i \)-ball lives in some good/bad tree, rooted in \( \mathcal{F}_t \) with \( t \leq i \). We can therefore apply good-/bad-tree packing estimates (6.20) and (7.15) to each such tree, and then use leaf-packing (8.6) to deduce

\[
\#\{Q_i\} r_i^k \leq c(k) \sum_{t=0}^{i} \sum_{f \in \mathcal{F}_t} r_f^k \leq c(k). \quad (8.19)
\]

We prove the measure estimate. By construction, the leaves \( \mathcal{F}_{t+1} \) include the leaves of any tree rooted at \( \mathcal{F}_t \). Therefore, given any leaf \( f \in \mathcal{F}_t \), we can apply the good/bad-tree measure estimates (6.21) and (7.16) to deduce

\[
\mu \left[ B_{r_f}(f) \setminus (B_{4r_x}(C'_+) \cup B_{r_i}(Q_i) \cup \bigcup_{f' \in \mathcal{F}_{t+1}} B_{r_{f'}}(f')) \right] \leq c(n)(M + \delta_0^2 M) r_f^k. \quad (8.20)
\]
Since \( Q_i \) contains every leaf of scale \( r_i \), and \( r_f \leq r_i \) when \( f \in \mathcal{F}_i \), we deduce

\[
\mu \left[ B_1 \setminus \left( B_{4r_i}(C') \cup B_{r_i}(Q_i) \right) \right] \\
\leq \sum_{t=0}^{i-1} \sum_{f \in \mathcal{F}_t} \mu \left( B_{r_f}(f) \setminus \left( B_{4r_i}(C'_t) \cup B_{r_i}(Q_i) \cup \bigcup_{f' \in \mathcal{F}_{t+1}} B_{r_f'}(f') \right) \right) \\
\leq c(n)(M + \delta_i^2 M) \sum_{t=0}^{i-1} \sum_{f \in \mathcal{F}_t} r_f^k \\
\leq c(n)M. 
\]

(8.21) (8.22) (8.23) (8.24)

Taking \( i \to \infty \), from our definition of \( C' \) we have that Lemma 8.2 implies Theorem 5.3 part B). We now show \( C' \) admits the correct closure, packing, and non-collapsing properties. We require first a Lemma on the comparability between balls in \( C'_+ \), and balls in \( Q_i \).

**Lemma 8.3.** For any \( y \in C'_+ \), with \( r_y \leq r_i \), we have

\[
B_{r_y}(y) \cap B_{2r_{i-1}}(Q_{i-1}) \neq \emptyset. 
\]

(8.25)

Conversely, if \( y \in C_+ \) and \( r_y > r_i \), then

\[
B_{r_y}(y) \cap B_{2r_i}(Q_i) = \emptyset. 
\]

(8.26)

For any \( y \in C' \), we have

\[
B_{100r_y}(y) \cap B_{r_x}(x) = \emptyset \quad \forall x \in C \text{ such that } r_x > c(n)r_y. 
\]

(8.27)

**Proof.** Given \( y \in C'_+ \), then \( y \in C_+(\mathcal{T}) \) for some good or bad tree. By construction, there must be a stop ball \( B_{r_i}(s) \) in this tree with \( r_i < r_y \leq r_{i-1} \) and \( s \in B_{r_y+2r_i}(y) \). But then \( s \in B_{r_{i-1}}(q) \) for some good or bad ball \( B_{r_{i-1}}(q) \), since trees are never rooted in stop balls. We deduce

\[
|y - q| \leq |y - s| + |s - q| < (r_y + 2r_i) + r_{i-1} < r_y + 2r_{i-1}. 
\]

(8.28)

Conversely, if \( y \in C_+ \), then using “radius control” in any tree (see inductive hypothesis (6.28) and Remark 7.2), we deduce

\[
B_{r_y}(y) \cap B_{2r_i}(Q_i) \neq \emptyset \implies r_y \leq r_i. 
\]

(8.29)

The last statement is a direct consequence of the first two, recalling that \( \rho = \rho(n) < 1/10 \).

**Lemma 8.4.** We have that \( C'_0 \subset C_0, C'_+ \subset C_+ \), both \( C'_0 \) and \( C' \) are closed, and \( C' \) satisfies the packing estimate

\[
\sum_{x \in C'_+} r_x^k \leq c(n), \quad \text{and} \quad |B_r(C')| \leq c(n)r^{n-k} \quad \forall 0 < r \leq 1. 
\]

(8.30)
Proof. Let us first prove the packing estimates (8.30). Applying Corollaries 6.8, 7.6 parts B) to each good/bad tree, and recalling that $\rho = \rho(n)$, we have

\[ \sum_{x \in C'_+} r_x^k \leq c(n) \sum_{t=0}^{\infty} \sum_{j \in F_t} r_j^k \leq c(n). \quad (8.31) \]

Take any $r \in (0, 1]$, and choose $i$ so that $r_i < r \leq r_{i-1}$. From Lemma 8.3, we know

\[ \bigcup_{y \in C'_+ : ry \leq r_i} B_r(y) \subset B_{4r_{i-1}}(Q_{i-1}). \quad (8.32) \]

Therefore

\[ r^{k-n} |B_r(C'_i)| \leq r^{k-n} |B_r(C'_0)| + r^{k-n} |B_r(C'_0)| \]
\[ \leq \omega n \sum_{y \in C'_+ : ry > r_i} r^k + 2r^{k-n} |B_{4r_{i-1}}(Q_{i-1})| \]
\[ \leq \omega n \rho^{-k} \sum_{y \in C'_+ : ry > r_i} r^k + 2r^{k-n} \# \{ Q_{i-1} \} \omega n (4r_{i-1})^n \]
\[ \leq \omega n c(n) + 2n \omega n (4/\rho)^n c(n). \quad (8.36) \]

We prove the other statements. By Lemma 8.3 we have that $C'_0 \subset \text{spt} \mu \setminus B_{r_0}^+ (C_+)$, and therefore $C'_0 \subset C_0$. The inclusion $C'_+ \subset C_+$ is by construction.

$C'_0$ is closed by construction. To prove $C'$ is closed, it will suffice to show that any limit point of $C'_+$ must lie in $C'_0$. Let $\{ y_i \} \subset C'_+$ be a sequence converging to some $y$. By packing estimate (8.30), necessarily $r_{y_i} \to 0$. After relabeling we can assume $r_{y_i} < r_i$.

We deduce that, for any fixed $A$, we have

\[ y_i \in B_{5r_{i-1}}(Q_{i-1}) \subset B_{r_{A+o(1)}}(Q_A) \quad \text{as } i \to \infty. \quad (8.37) \]

This proves that $y \in C'_0$. \hfill \Box

Lemma 8.5. We have the lower bound

\[ \mu B_r(y) > \frac{M}{c(n)} r^k \quad \forall y \in C', \forall 4r_y \leq r \leq 1. \quad (8.38) \]

Proof. Take $y \in C'_+$, and choose $A$ so that $r_{A+1} \leq r_y < r_A$. By Lemma 8.3 we have for every $i \leq A - 1$ the existence of a $q \in Q_i$ so that

\[ B_{4r_i}(y) \supset B_{r_i}(q). \quad (8.39) \]

Using the definition of a good/bad ball we therefore have

\[ \mu B_{4r_i}(y) > M r_i^k \quad \forall i \leq A - 1. \quad (8.40) \]

Since $\mu B_{4r_y}(y) > \frac{M}{\rho^2} r_y^k$ (Remarks 6.5, 7.3), this proves (8.38) with $c = c(n, \rho) = c(n)$.

Given $y \in C'_0$, then by construction the above reasoning works for any $i$. \hfill \Box
8.2. **Structure of $C'_0$.** We have established Theorem 5.3 parts A)-C), and that $C'_0$ is closed. To show upper-Ahlfors-regularity and rectifiability of $C'_0$ (conclusion D) we demonstrate that the measure $\eta = \mathcal{H}^k \mathcal{C}'_0$ satisfies

$$
\sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq 2} \beta^k_{\eta, 2, 0}(z, 2^\alpha)^2 \leq c(n)\delta^2_0 \quad \forall z \in B_1. \tag{8.41}
$$

We then can apply our own Theorem 5.3 parts A), B) to the measure $\eta$ to deduce upper-Ahlfors-regularity, and the Proposition 10.1, which will obviously be proven independently, to deduce rectifiability.

To this end, we first apply Lemma 4.4 with $A_1 = C'_0 \cap U$, $A_2 = U$ to deduce

$$
\eta(U) \leq \frac{c(k)}{M} \mu(U) \quad \forall \text{ open } U. \tag{8.42}
$$

Now by Fubini we deduce that for any $B_r(x)$,

$$
\beta^k_{\eta, 2, 0}(x, r)^2 \leq r^{-k-2} \int_{B_r(x)} d(z, V^k_\mu(x, 2r))^2 d\eta \tag{8.43}
$$

$$
= r^{-k-2} \int_0^r \eta\{z \in B_r(x) : d(z, V^k_\mu(x, 2r))^2 > s\} ds \tag{8.44}
$$

$$
\leq \frac{c(k)}{M} r^{-k-2} \int_0^r \mu\{z \in B_r(x) : d(z, V^k_\mu(x, 2r))^2 > s\} ds \tag{8.45}
$$

$$
\leq \frac{c(n)}{M} \beta^k_{\mu, 2, \epsilon^k}(x, 2r)^2, \tag{8.46}
$$

provided $\mu(B_{2r}(x)) > \bar{\epsilon}_\beta(2r)^k$. But using C), if $\eta(B_r(x)) > 0$, then we must have some $z \in C'_0 \cap B_r(x)$, and therefore

$$
\mu(B_{2r}(x)) \geq \mu(B_r(z)) \geq \frac{M}{c_{\text{key}}} \omega_k r^k \geq \bar{\epsilon}_\beta(2r)^k. \tag{8.47}
$$

We can sum over $\alpha$ to find

$$
\sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq 2} \beta^k_{\eta, 2, 0}(x, 2^\alpha)^2 \leq \frac{c(n)}{M} \sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq 4} \beta^2_{\mu, 2, \epsilon^k}(x, 2^\alpha)^2 \leq \frac{c(n)}{M} c(k)\delta^2_0 M. \tag{8.48}
$$

In the last inequality we used our assumption (5.1), and that $\mu$ is supported in $B_1$. This proves the required estimate (8.41).

We use (5.1) to prove upper-Ahlfors-regularity and rectifiability. First, observe that

$$
2^{-k} \leq \Theta^{*k}(\eta, x) \leq 1 \quad \text{for } \eta-\text{a.e. } x, \tag{8.49}
$$

since $\mathcal{H}^k(C'_0) < \infty$.

To prove upper-Ahlfors-regularity we essentially reprove Corollary 2.15 part U, using Theorem 5.3 in place of Theorem 2.6. Fix any $B_r(x)$. Define the covering pair $(S, s_z)$ by setting

$$
S_+ = \{ z \in B_r(x) \cap C'_0 : \Theta^{*k}(\eta, z) \leq 1 \}, \quad S_0 = B_r(x) \setminus S_+. \tag{8.50}
$$

By (8.49) $\eta(S_0) = 0$, and for each $z \in S_+$, we can choose an $s_z < r$ for which $\eta(B_{4s_z}(z)) \leq 2\omega_k(4s_z)^k$. 

Apply Theorem 5.3 parts A, B) to \( \eta_{B_r}(x) \) at scale \( B_r(x) \), and the covering pair \((\mathcal{S}, s_x)\). Then there is a subset \( \mathcal{S}' \) so that

\[
\eta(B_r(x)) \leq c(n)r^k + \sum_{z \in \mathcal{S}'_u} \eta(B_{4r(z)}(z)) \leq c(n)r^k + c(k) \sum_{z \in \mathcal{S}'_u} s_z^k \leq c(n)r^k. \tag{8.51}
\]

We deduce

\[
\eta(B_r(x)) \equiv \mathcal{H}^k(C'_0 \cap B_r(x)) \leq c(n)r^k \quad \forall x \in B_1, 0 < r < 1. \tag{8.52}
\]

Finally, we apply Proposition 10.1, which will be proved independently, to deduce \( C'_0 \) is rectifiable. In fact Lemma 10.2 shows \( C'_0 \) is a kind of “uniformly rectifiable,” in the sense that whenever \( \mathcal{H}^k(C'_0 \cap B_r(x)) > 2^{-k-1}\omega_k r^k \), there is some bi-Lipschitz \( k \)-manifold \( T \) so that

\[
\mathcal{H}^k(C'_0 \cap T) \geq \frac{1}{2} \mathcal{H}^k(C'_0 \cap B_r(x)). \tag{8.53}
\]

This finishes the proof of Theorem 5.3.

9. Packing/Measure estimates and density

In general we cannot do better than a combination measure and packing estimate. At the one extreme, \( \mu \) could have infinite mass concentrated across a \( k \)-plane. At the other, \( \mu \) may have no \( k \)-dimensional structure whatsoever, as when \( \mu = \mathcal{L}^n \).

However in the presence of density bounds, we can deduce purely Hausdorff or purely measure estimates. Lower bounds on upper density give global Hausdorff bounds, and upper bounds on lower density give global measure bounds. Density enforces some kind of \( k \)-dimensional structure locally.

We focus this section on proving Theorems 2.11, 2.14, and their Corollaries. Theorem 2.11 is a more-or-less direct consequence of Theorem 2.6. However, Theorem 2.14 is more involved. As the proof is vastly simpler, and more instructive, let us give first the special case of Corollary 2.12.

Proof of Corollary 2.12. Part L). Apply Theorem 2.6 with \( \epsilon = \bar{\epsilon}_\beta/c(k) \) to find a closed set \( K_0 \) so that

\[
\mathcal{H}^k(K_0) \leq c(n), \quad \mu(B_1 \setminus K_0) \leq c(n)(\epsilon + M) + \Gamma. \tag{9.1}
\]

Define

\[
K_1 = \{ z \in B_1 \setminus K_0 : \Theta^k(\mu, x) \geq \alpha \}, \tag{9.2}
\]

and then use Lemma 4.4 with \( A_1 = K_1, A_2 = B_1 \setminus K_0 \) to deduce

\[
a \mathcal{H}^k(K_1) \leq \mu(B_1 \setminus K_0) \leq c(n)(\epsilon + M) + \Gamma. \tag{9.3}
\]

Now set \( C'_0 = K_0 \cup K_1 \).

Part U). Let \( \mathcal{S} \) be the covering pair defined by

\[
\mathcal{S}_+ = \{ z \in B_1 : \Theta^k(z, s) \leq b \}, \quad \mathcal{S}_0 = \mathcal{C} \setminus \mathcal{S}_+. \tag{9.4}
\]

For \( x \in \mathcal{S}_+ \), choose \( s_x \leq 1 \) so that \( \mu(B_{s_x}(x)) \leq 2\omega_k bs_x^k \). By construction, \( \mu(\mathcal{S}_0) = 0 \), and \( \mathcal{S} \) is closed.
Apply Theorem 2.6 with \( \epsilon = \bar{\epsilon}_\beta/c(k) \) to the covering pair \((S, s_x)\), to obtain a \( S' \). We then have
\[
\mu(B_1) \leq c(n)(M + \epsilon) + \Gamma + \mu(B_{s_x}(S'_x)) \\
\leq c(n)(M + \epsilon) + \Gamma + c(k)b \sum_{x \in S'_x} s_x^k \\
\leq c(n)(M + \epsilon + b) + \Gamma.
\] (9.5) (9.6) (9.7)

There are a couple complications to proving Theorem 2.14 in general. First, in any given ball \( B_r(x) \), condition (2.20) only gives information on points \( y \) with \( r_y \leq 2r \). We need to find a good “minimal” subset \( U \subset \mathcal{C} \), so that in any ball centered in \( U \) we have \( \beta \)-bounds down to a uniform, comparable scale. This is the content of Lemma 9.3.

Second, to prove part L), we cannot use the approach of Corollary 2.12 as proven above, since this can only give a packing estimate at a single scale. Nor can we apply Theorem 2.11 at every scale, since this will give us a different packing for each scale. We must choose a single collection of balls, and demonstrate packing bounds at all scales.

Our strategy to prove part L), in general terms, is to find some discrete approximation \( \mu_{\text{dis}} \) of \( \mu \) on the regions of big mass, prove \( \beta \)-bounds on \( \mu_{\text{dis}} \) in terms of the \( \beta \)-bounds of \( \mu \), and then use Theorem 2.6 to bound \( \mu_{\text{dis}} \). The main headache is that we can only get and use meaningful information on the \( \beta \)-numbers of \( \mu_{\text{dis}} \) if: 1) the balls associated to \( \mu_{\text{dis}} \) are centered at \( \mu \)-centers-of-mass; 2) the balls have some disjointness control; and 3) the balls have lower \( \mu \)-mass bounds.

These issues necessitate the use of an intermediary packing measure \( \nu \), which controls the original packing measure \( \mu_{\text{dis}} \), but which admits the right properties 1)-3). The following Subsection deals with this process.

9.1. Discretizing a measure. We use the following specialization of packing measure as an intermediary.

**Definition 9.1.** We say \( \nu \) is \((k\text{-dimensional) disjoint packing measure}\) if there is some collection of disjoint balls \( \{B_{r_p}(x_p)\}_p \) so that \( \nu \) has the form
\[
\nu = \sum_p r_p^k \delta_{x_p}.
\] (9.8)

Here, as always, \( \delta_{x_p} \) is the Dirac delta at \( x_p \).

A key utility of a disjoint packing measure for us is the following Theorem, what is essentially a baby version of discrete Reifenberg Theorem 2.20.

**Theorem 9.2.** Let \( \nu \) be a \( k \)-dimensional disjoint packing measure supported in \( B_1 \), and suppose
\[
\nu \left\{ z \in B_1 : \int_0^1 \beta_{\nu,2,\bar{\epsilon}_{\beta}}(z,r)^2 \frac{dr}{r} > M \right\} \leq \Gamma.
\] (9.9)

Then provided \( \bar{\epsilon}_\beta \leq c(k) \max \{M, \epsilon\} \), we have
\[
\nu(B_1) \leq c(n)(1 + \epsilon + M) + \Gamma.
\] (9.10)
Proof. By definition we can find a covering pair \((S, r_x)\) so that \(S_0 = \emptyset\), and
\[
\nu = \sum_{x \in S} r_x^k \delta_x. \tag{9.11}
\]

Trivially \(S\) is a covering pair for \(\nu\).

Apply Theorem 2.6 to obtain a \(S' \subset S\). Using Theorem 2.6 parts A), B), and the disjointness hypothesis, we have
\[
\nu(B_1) \leq c(n)(\epsilon + M) + \Gamma + \nu(B_{r_x}(S')) + \sum_{x \in S'} r_x^k, \tag{9.12}
\]
\[
= c(n)(\epsilon + M) + \Gamma + \sum_{x \in S'} r_x^k, \tag{9.13}
\]
\[
\leq c(n)(1 + \epsilon + M) + \Gamma. \tag{9.14}
\]

Note we see above that in the context of a packing measure, the packing control on \(S'\) automatically gives rise to a measure estimate on \(S'\). \(\square\)

The notation \(U\) introduced below is purely a technical issue. Since we a priori only know \(\beta\)-number information down to scale \(r_x\), in a random ball \(B_{r_x}(x)\) we may not have \(\beta\)-number bounds down to a uniform, comparable scale. It turns out one can always find a good subset \(U(C) \subset C\) with the properties we desire.

**Lemma 9.3.** Given a covering pair \((C, r_x)\) with \(r_x \leq 1\), there is a subset \(U(C, r_x) \subset C\) which admits the following properties:

A) if \(x \in U\), and \(r \geq r_x\), then \(y \in C \cap B_r(x) \implies r_y \leq 2r\).

B) \(B_{5r_x}(U) \supset B_{r_x}(C)\), and \(B_{r_x}(U_+) \supset B_{r_x}(C_+\).

**Remark 9.4.** Property A) trivially holds for any \(y \notin B_{r_x}(C_+)\), and \(r > 0\).

**Proof.** Let us write \(r_\alpha = 2^{-\alpha}\). For each integer \(\alpha \geq 0\), define
\[
W_\alpha = \{ z \in C : r_{\alpha+1} < r_z \leq r_\alpha \}. \tag{9.15}
\]

By assumption \(\cup_\alpha W_\alpha = C_+\).

Let \(K_0 = W_0\), and for each integer \(\alpha \geq 1\), let
\[
K_\alpha = W_\alpha \setminus B_{r_x} (\cup_{\beta \leq \alpha-1} W_\beta). \tag{9.16}
\]

Now let
\[
U_+ = \cup_\alpha K_\alpha, \quad U_0 = C_0 \setminus B_{r_x}(C_+). \tag{9.17}
\]

We claim that if \(x \in U\), and \(r \geq r_x\), then
\[
y \in C \cap B_r(x) \implies r_y \leq 2r. \tag{9.18}
\]

If \(x \in U_0\), then this follows trivially, since \(x \notin B_{r_y}(y)\) when \(r_y > 0\). Otherwise, \(x \in K_\alpha\), and if \(r_y > 2r_x\) then \(y \in W_\beta\) for some \(\beta < \alpha\). But then, by construction \(x \notin B_{r_y}(y)\), and so \(r \geq r_y\). This proves the claim.

We now claim that \(B_{5r_x}(U) \supset B_{r_x}(C)\). Let us prove first that
\[
W_\alpha \subset B_{4r_x} (\cup_{\beta \leq \alpha} K_\beta). \tag{9.19}
\]
Take \( x_1 \in W_\alpha \), and write \( \alpha_1 = \alpha \). If \( x_1 \in K_\alpha \), we are done. Otherwise, by construction, there is an \( x_2 \in W_\alpha \), with \( \alpha_2 < \alpha_1 \), so that \( x_1 \in B_{r_{\alpha_2}}(x_2) \). If \( x_2 \notin K_\alpha \), then we can find an \( x_3 \) as we did \( x_2 \). We can thus build a sequence \( x_1, x_2, \ldots \) satisfying

\[
x_i \in W_{\alpha_i} \setminus K_{\alpha_i}, \quad x_{i-1} \in B_{r_{\alpha_i}}(x_i), \quad \alpha_i < \alpha_{i-1}.
\]  

(9.20)

Since \( W_0 = K_0 \), this must terminate at some \( x_N \in K_{\alpha_N} \), with \( x_{N-1} \in B_{\alpha_N}(x_N) \). We therefore calculate

\[
|x_1 - x_N| < r_{\alpha_2} + \ldots + r_{\alpha_N} \leq 2r_{\alpha_N} \leq 4r_{x_N}.
\]

(9.21)

(9.22)

(9.23)

In other words, \( x_1 \in B_{4r_x}(K_{\alpha_N}) \). This proves (9.19), and in particular

\[
B_r(W_\alpha) \subset B_{5r_x}(\cup_{\beta \leq \alpha} K_\beta).
\]

(9.24)

We deduce \( B_{r_x}(C_+) \subset B_{5r_x}(U_+) \). We then have

\[
C_0 \subset U_0 \cup B_{r_x}(C_+) \subset U_0 \cup B_{5r_x}(U_+).
\]

(9.25)

This proves the second claim.

We can now state the main Lemma. It says that if you have the right disjointness to your packing measure, then \( \beta \)-bounds of \( \mu \) controls the packing measure. Due to the center of mass issues we must use an intermediary \( \nu \).

Lemma 9.5. Let \( \mu \) be a non-negative Borel-regular measure in \( B_1 \), and let \((C, r_x)\) be a covering pair for \( \mu \) with \( C = C_+ \) and \( r_x \leq 1 \). Suppose \( C' := C'_+ \subset U(C) \) is such that the balls \( \{B_{2r_x}(x)\}_{x \in C'_+} \) are disjoint, and

\[
\mu(B_{r_x}(x)) \geq a r_x^k \quad \forall x \in C' = C'_+.
\]

(9.26)

Then there is \( k \)-dimensional disjoint packing measure \( \nu \), supported in \( \overline{B_1} \), so that

A) for any \( x \in C', r \geq r_x \), we have \( \mu_{C'}(B_r(x)) \leq \nu(B_{2r}(x)) \).

B) for any \( x \in B_1, 0 < r < 1, \epsilon > 0 \), we have

\[
\nu \left( z \in B_r(x) : \int_0^{2r} \beta_{\epsilon, 2, \mu, k}^k(z, s)^2 \frac{ds}{s} > \frac{c(k)M}{a} \right) \leq \frac{1}{a} \mu \left( z \in B_{4r}(x) : \int_{50r_x}^{32r} \beta_{\epsilon, 2, \mu, k}^k(z, s)^2 \frac{ds}{s} > M \right).
\]

(9.27)

(9.28)

Proof. Given \( y \in C' \), let \( Y \in \overline{B_{r_y}(y)} \cap \overline{B_1} \) be the generalized \( \mu \)-center of mass of \( B_{r_y}(y) \). Define

\[
\nu = \sum_{y \in C'} r_y^k \delta_{y}.
\]

(9.29)

By assumption the balls \( \{B_{r_y}(Y)\}_{y \in C'} \) are disjoint.
Let us prove conclusion A). Given any \( y \in B_r(x) \cap C' \), by disjointness and our choice of \( r \) we know \( r > r_y \). Since the generalized \( \mu \)-center of mass \( Y \in B_{r_y}(y) \), we deduce

\[
\mu_{C'}(B_r(x)) = \sum_{y \in B_r(x) \cap C'} r^k_y \leq \sum_{Y \in B_{2r}(x) \cap \text{spt } \nu} r^k_y = \nu(B_{2r}(x)).
\]  

(9.30)

We now work towards conclusion B). We claim that for each \( X \in \text{spt } \nu \), and \( r \geq r_x \), we have

\[
\nu(B_r(x)) \leq \frac{1}{\alpha} \mu(B_{4r}(x)) \quad \text{and} \quad \beta^k_{\nu,2,4^k\epsilon/a}(x, r)^2 \leq \frac{c(k)}{\alpha} \beta^k_{\mu,2,\epsilon}(x, 4r)^2.
\]  

(9.31)

Let us prove (9.31). If \( Y \in B_r(x) \cap \text{spt } \nu \), then by disjointness \( r_y < r \). We deduce from hypothesis (9.26)

\[
\nu(B_r(x)) = \sum_{y \in B_r(x) \cap \text{spt } \nu} r^k_y \leq \frac{1}{\alpha} \sum_{y \in B_{2r}(x) \cap C'} \mu(B_{r_y}(y)) \leq \frac{1}{\alpha} \mu(B_{4r}(x)).
\]  

(9.32)

Let \( V = V^k_{\mu}(x, 4r) \). The following calculation (specifically, inequality (9.34)) is the key place we use that \( Y \) is the generalized \( \mu \)-center of mass. This inequality (9.34) follows by Proposition 4.10, and of course we only need to sum over \( y \) with \( \mu(B_{r_y}(y)) < \infty \). We calculate

\[
r^{k+2} \beta^k_{\mu}(x, r)^2 \leq \sum_{Y \in B_r(x) \cap \text{spt } \nu} r^k_y d(Y, V)^2,
\]  

(9.33)

\[
\leq \sum_{y \in B_{2r}(x) \cap C'} r^k_y \frac{d(z, V)^2}{\mu(B_{r_y}(y))} \int_{B_{r_y}(y)} d(z, V)^2 d\mu(z),
\]  

(9.34)

\[
\leq \frac{1}{\alpha} \int_{B_{4r}(x)} d(z, V)^2 d\mu(z),
\]  

(9.35)

\[
= \frac{4^{k+2}}{\alpha} r^{k+2} \beta^k_{\mu,2,\epsilon}(x, 4r)^2,
\]  

(9.36)

provided \( \mu(B_{4r}(x)) > \epsilon(4r)^k \), which certainly holds if \( \nu(B_r(x)) > 4^k \epsilon/\alpha r^k \). This proves (9.31).

Equation (9.31) is a pointwise \( \beta \)-number bound at a single scale. Towards conclusion B), we prove a pointwise inequality on the summed-\( \beta \)-numbers. Let us fix an \( x \in B_1 \), and \( r > 0 \) arbitrary.

First, if \( B_r(x) \cap \text{spt } \nu \) is a singleton, then by definition \( \beta_{\nu}(x, r) = 0 \). Therefore, by disjointness, we can assume \( r > r_y \) for every \( Y \in B_r(x) \cap \text{spt } \nu \). Here of course \( Y \) is the generalized \( \mu \)-center of mass for \( B_{r_y}(y) \), where \( y \in C' \).

Otherwise, given \( z \in B_{r_y}(y) \) with \( y \in C' \), we have by (9.31) and our choice of \( C' \subset U \):

\[
\int_{50r_y}^{32r} \beta^k_{\mu,2,\epsilon}(z, s)^2 ds/s \geq \int_{100r_y}^{32r} \beta^k_{\mu,2,\epsilon}(z, s)^2 ds/s,
\]  

(9.37)

\[
\geq \frac{1}{c(k)} \int_{25r_y}^{8r} \beta^k_{\nu,2,4^k\epsilon/a}(z, s)^2 ds/s,
\]  

(9.38)

\[
\geq \frac{a}{c(k)} \int_0^{2r} \beta^k_{\nu,2,\epsilon}(Y, s)^2 ds/s.
\]  

(9.39)

In particular, suppressing the \( \epsilon \) notation, if \( \int_0^{2r} \beta_{\nu}(Y, s)^2 ds/s > \frac{c(k) M}{\alpha} \), then

\[
\mu \left( z \in B_{r_y}(y) : \int_{50r_y}^{32r} \beta_{\mu}(z, s)^2 ds/s > M \right) \geq \mu(B_{r_y}(y)) \geq \alpha r^{k_y}.
\]  

(9.40)
We therefore calculate,
\[
\nu\left( z \in B_r(x) : \int_0^{2r} \beta_\nu(z, s)^2 ds/s > \frac{c(k)M}{a} \right) \quad (9.41)
\]
\[
= \sum \left\{ r_y^k : Y \in B_r(x) \cap \text{spt}\nu \text{ and } \int_0^{2r} \beta_\nu(Y, s)^2 ds/s > \frac{c(k)M}{a} \right\} \quad (9.42)
\]
\[
\leq \frac{1}{a} \sum_{y \in B_{2r}(x)} \mu\left( z \in B_{r_y}(y) : \int_{50r_y}^{32r} \beta_\mu(z, s)^2 ds/s > M \right) \quad (9.43)
\]
\[
\leq \frac{1}{a} \mu\left( z \in B_{4r}(x) : \int_{50r}^{32r} \beta_\mu(z, s)^2 ds/s > M \right). \quad (9.44)
\]
This completes the proof of Lemma 9.5. □

The main Theorem of this subsection is the following. It follows easily from Lemma 9.5.

**Theorem 9.6.** Let \( \mu \) be a non-negative Borel-regular measure in \( B_1 \), and let \( (C, r_x) \) be a covering pair for \( \mu \) with \( C = C_+ \) and \( r_x \leq 1 \). Suppose \( C' := C'_+ \subset U(C) \) is such that the balls \( \{B_{2r_x}(x)\}_{x \in C'} \) are disjoint, and

\[
\mu(B_{r_x}(x)) \geq ar^k_x \quad \forall x \in C' = C'_+. \quad (9.45)
\]

Take \( x \in C' \) and \( r \geq r_x \). If

\[
\mu\left\{ z \in B_{16r}(x) : \int_{50r_x}^{64r} \beta_{\mu,2,\bar{\epsilon}_\beta}(z, r) dr/r > M \right\} \leq \Gamma r^k, \quad (9.46)
\]

and \( \bar{\epsilon}_\beta \leq c(k) \max\{M, \epsilon\} \), then

\[
\mu_{C'}(B_r(x)) \equiv \sum_{y \in C' \cap B_r(x)} r_y^k \leq \left( c(n) + \frac{c(n)(\epsilon + M) + \Gamma}{a} \right) r^k. \quad (9.47)
\]

**Proof.** Let \( \nu \) be the disjoint discrete packing measure of Lemma 9.5. Whenever (9.46) holds, then by Lemma 9.5 we have

\[
\nu\left( z \in B_r(x) : \int_0^{2r} \beta_{\nu,2,c(k)\bar{\epsilon}_\beta}(z, r)^2 ds/s > \frac{c(k)M}{a} \right) \leq \frac{\Gamma}{a} r^k. \quad (9.48)
\]

Using Theorem 9.2 at scale \( B_r(x) \), provided \( \bar{\epsilon}_\beta \leq c(k) \max\{M, \epsilon\} \), we have

\[
\frac{\nu(B_r(x))}{r^k} \leq c(n) \left( 1 + \frac{\epsilon}{a} + \frac{c(k)M}{a} \right) + \frac{\Gamma}{a}. \quad (9.49)
\]

□

**Remark 9.7.** Let us remind the reader that (9.46) is a weaker assumption than the corresponding estimate without the factor of 50.
9.2. Proofs of Theorems. We are now ready to prove the Theorems involving upper and lower density. Corollaries 2.12, 2.13, 2.15, 2.16 are immediate from Theorems 2.11, 2.14, and we will not explicitly prove them.

Proof of Theorems 2.11, 2.14, part L). Let us first handle the \( C'_+ \). Define the covering pair \((S, s_x)\) by setting

\[
S_+ = \{ x \in C_+ : \Theta^{*,k}(\mu, C, x) \geq a \}, \quad S_0 = \varnothing.
\]

Choose \( s_x \in [r_x/50, r_x/25] \) so that \( \mu(B_{s_x}(x)) \geq 50s_x^k \). Clearly \( S \) is a covering pair for \( \mu \), and by assumption \( \mu(C_+ \setminus S_+) = 0 \).

Define \( C'_+ \) to be (the centers of) a Vitali subcover of

\[
\{ B_{5s_x}(x) : x \in \mathcal{U}(S)_+ \}.
\]

In other words, the balls \( \{ B_{5s_x}(x) \}_{x \in C'_+} \) are disjoint, while

\[
B_{r_x}(C'_+) \supset B_{25s_x}(C'_+) \supset B_{5s_x}(\mathcal{U}(S)_+) \supset S_+
\]

In particular, \( \mu(C_+ \setminus B_{r_x}(C'_+)) = 0 \). Clearly \( C_+ \) is independent of \( \epsilon \).

We demonstrate that \( C'_+ \) admits packing bounds at scale \( B_1 \) (in the case of Theorem 2.11), or at all scales (in the case Theorem 2.14). First observe that \( C'_+ \subset \mathcal{U}(S) \) and \( \mu \) as defined satisfy all hypotheses of Theorem 9.6 except (9.46). Fix also \( \epsilon = \bar{\epsilon}/c(k) \) as required by Theorem 9.6.

Since \( r_x \leq 50s_x \), and \( \mu \) is supported in \( B_1 \) the hypothesis of Theorem 2.11 implies (9.46) holds at scale \( B_1 \), except with \( c(k)M \) in place of \( M \), and \( c(k)\Gamma \) in place of \( \Gamma \). We can therefore apply Theorem 9.6, recalling our choice \( \epsilon = \bar{\epsilon}/c(k) \), to deduce

\[
\sum_{x \in C'_+} r_x^k \leq 50^k \sum_{x \in C'_+} s_x^k \leq c(n) + \frac{c(n)(\bar{\varepsilon} + M) + c(k)\Gamma}{a}.
\]

On the other hand, using again \( r_x \leq 50s_x \), the hypothesis of Theorem 2.14 implies (9.46) holds at any scale \( B_{r_x}(x) \), except with \( \Gamma = M \). In this case we can apply Theorem 9.6 at any \( B_{r_x}(x) \) to deduce

\[
\sum_{x \in B_{r_x}(x) \cap C'_+} s_x^k \leq \left( c(n) + \frac{c(n)(\bar{\varepsilon} + M)}{a} \right)x^k.
\]

This proves the required estimate packing for \( C'_+ \).

We now handle \( C'_0 \). Fix \( \delta > 0 \), and define the covering pair

\[
S_+ = \{ x \in C_0 : \Theta^{*,k}(\mu, C, x) \geq a \}, \quad S_0 = S \setminus S_+.
\]

Choose \( s_x < \delta \) so that \( \mu(B_{s_x}(x)) \geq \frac{\omega}{2} s_x^k \).

Define \( C'_0 = S_+ \cap C_0 \), which is clearly independent of \( \delta \) and \( \epsilon \). By assumption \( \mu(C_0 \setminus S_+) = \mu(C_0 \setminus C'_0) = 0 \).

By the same reasoning as above, we can find a Vitali subcover \( S'_+ \) of

\[
\{ B_{5s_x}(x) : x \in \mathcal{U}(S) \},
\]

so that \( \{ B_{5s_x}(x) \}_{x \in S'_+} \) are disjoint, and

\[
B_{25s_x}(S'_+) \supset S_+
\]
By the same argument as before, taking $\epsilon = \tilde{\epsilon}_\beta/c(k)$, the hypothesis of Theorem 2.11 together with Theorem 9.6 implies a packing estimate
\[
\sum_{x \in S'} r^k_x \leq c(n) + \frac{c(n)(\epsilon + M) + c(k)\Gamma}{a}, \quad (9.58)
\]
which in turn gives
\[
\mathcal{H}^k_\delta(C'_0 \cap B_1) \leq c(n) + \frac{c(n)(\tilde{\epsilon}_\beta + M) + c(k)\Gamma}{a}. \quad (9.59)
\]

While the hypothesis of Theorem 2.14 implies that, for any $z \in B_1$ and $0 < r \leq 1$,
\[
\sum_{x \in S' \cap B_r(z)} r^k_x \leq \left(\frac{c(n) + c(n)(\epsilon + M)}{a}\right) r^k, \quad (9.60)
\]
and therefore
\[
\mathcal{H}^k_\delta(C'_0 \cap B_r(x)) \leq \left(\frac{c(n) + c(n)(\tilde{\epsilon}_\beta + M)}{a}\right) r^k. \quad (9.61)
\]
Since $\delta$ is arbitrary, this proves the required Hausdorff estimate on $C'_0$. \qed

**Proof of Theorems 2.11, 2.14, part U).** Let us first prove Theorem 2.11 part U). Fix $\epsilon = \tilde{\epsilon}_\beta/c(k)$ as required by Theorem 2.6.

Let $(S, s_x)$ be a covering pair defined by
\[
S_+ = \{x \in C : \Theta^k_\mu(C, x) \leq b\}, \quad S_0 = S \setminus S_+ \quad (9.62)
\]
and choosing $s_x \in (0, 1)$ so that $\mu(B_{s_x}(x)) \leq 2b\omega k s_x^k$. By assumption $\mu(B_1 \setminus S_+) = \mu(S_0) = 0$, and $S$ is closed.

Apply Theorem 2.6 and obtain a $S' \subset S$. Trivially we have $\mu(S'_0) = 0$. We have
\[
\mu(B_1) \leq c(n) + \sum_{x \in S'_+} \mu(B_{s_x}(x)) \leq c(n) + \sum_{x \in S'_+} s_x^k \leq c(n)(\tilde{\epsilon}_\beta + M) + \sum_{x \in S'_+} s_x^k \leq c(n)(\tilde{\epsilon}_\beta + M) + \Gamma. \quad (9.63)
\]
This proves Theorem 2.11 part U). We now use this to prove Theorem 2.14. We first observe that, trivially, $(C, r_{x/5})$ is a covering pair for $\mu$, and satisfies
\[
\mu \left\{z \in B_{r_{x}}(x) : \int_{S_{r_{x}/5}} \beta^k_{\mu, 2s} (z, s) \frac{ds}{s} > M \right\} \leq M r^k \quad \forall x \in S, r \geq r_{x}. \quad (9.64)
\]
Let $C' = \mathcal{U}(C, r_{x/5})$ as in Lemma 9.3. Then from our choice of $r_{x/5}$, we have $B_{r_{x}}(C') \supset C$. Now take $x \in C'$ and $r \geq r_{x}$, or take $x \notin B_{r_{x}}(C'$) and $r > 0$. In either case, we have by construction that
\[
y \in B_{r}(x) \cap C \implies r_y \leq 10r. \quad (9.65)
\]
Let $\mu' = \mu \ll B_r(x) \leq \mu$. Then $C \cap B_r(x)$ is a covering pair for $\mu'$, which (by monotonicity of $\beta$) satisfies the requirements of Theorem 2.11 part U) at scale $B_{10r}(x)$. We deduce that
\[ \mu(B_r(x)) = \mu'(B_{10r}(x)) \leq c(n)(\bar{\epsilon}_\beta + M + b)(10r)^k, \] (9.68)
which evidently finishes the proof. \[ \square \]

10. Rectifiability

In this section, we prove the rectifiability properties promised in the main theorems, and in particular Theorem 2.19. All these properties will be obtained as corollaries of the following Proposition, which is the main statement of this section. The main purpose of the following Proposition, in comparison with other results in the literature, is that we sharply weaken the density assumptions required on the measure:

**Proposition 10.1.** Let $\mu$ be a non-negative Borel-regular measure such that for $\mu$-a.e.:
\[ \Theta^k_\nu(\mu, x) \leq b < \infty, \quad \Theta^{*,k}(\mu, x) \geq a > 0. \] (10.1)
Suppose also that for $\mu$-a.e.
\[ \int_0^2 \beta_{\mu,2,\bar{\epsilon}_\beta}(x,r)^2 \frac{dr}{r} \leq M, \] (10.2)
where $\bar{\epsilon}_\beta \leq \min\{a, b\}/c(n, b, M)$. Then $\mu$ is $k$-rectifiable, and in particular we have that the density $\Theta^k(\mu, x)$ exists with $a \leq \Theta^k(\mu, x) \leq b$ $\mu$-a.e.

Before proving this Theorem, we briefly discuss the main ideas behind the proofs in this section.

First, to prove rectifiability we need control over the $\beta$-numbers all the way down to radius 0, not just some positive $r_x$. Second, we need to assume some kind of upper and lower density bounds on $\mu$, see for instance Example 3.7. The lower bound prevents any higher dimensional pieces from appearing, and the upper bound prevents any lower dimensional pieces from appearing. Note, however, we will only assume upper bounds on the lower density of the measure and lower bounds on upper density, which is the weakest density assumption one might hope for.

With these two assumptions we can apply Theorem 2.14 and obtain full-blown upper Ahlfors’ regularity for the measure $\mu$ under consideration. This apriori bound immediately implies that $\mu \ll H^k$, and more importantly can be used as in Theorem 4.13 part C) to guarantee that a bad ball carries very small measure.

With this information, by repeating the construction of a good tree rooted at any $B_r(x)$, we can build a Lipschitz manifold $T_\infty$ such that $\mu(B_r(x) \setminus T_\infty) \leq c\epsilon r^k$.

By itself, this is not enough to prove rectifiability, as it is clear from Example 3.5. The problem is that $\mu B_r(x)$ can be much smaller than the part of the measure covered by $T_\infty$. However, under lower bounds on the density, we can guarantee that $B_r(x)$ (for suitably chosen $x$ and $r$) has at least some mass: $\mu(B_r(x)) \geq ar^k$. If $a > 0$ and $\epsilon$ is sufficiently small, this implies that at least half of the measure of
μ(B_r(x)) is covered by our Lipschitz manifold \( T_\infty \). With a simple inductive argument, this implies the rectifiability of \( \mu \).

We start by showing that, under uniform upper Ahlfors bounds, smallness of the \( \beta \) numbers implies that we can cover a big chunk of the support of \( \mu \) with a Lipschitz manifold.

**Lemma 10.2.** For any \( \epsilon > 0 \), there is a \( \delta_2(n, \Gamma, \epsilon) \) so that the following holds: let \( \mu \) be a non-negative Borel-regular measure supported in \( B_1 \), having uniform upper bounds

\[
\mu(B_r(x)) \leq \Gamma r^k \quad \forall x \in B_1, \forall 0 < r \leq 1, \tag{10.3}
\]

and satisfying

\[
\int_0^2 \beta_{\mu,2,\epsilon_\beta}^k(z,r)2\frac{dr}{r} \leq \delta^2 \quad \text{for} \quad \mu\text{-a.e.} \ z \in B_1, \tag{10.4}
\]

with \( \delta \leq \delta_2 \), and \( \epsilon_\beta \leq \epsilon/c(n, \Gamma) \).

Then there is \( c(n, \epsilon, \Gamma) \delta^2 \)-Lipschitz mapping \( \tau : \mathbb{R}^k \to \mathbb{R}^n \), so that

\[
\mu(B_1 \setminus \tau(B_2^k)) \leq c_4(n)\epsilon + c_5(n, \epsilon, \Gamma)\delta^2. \tag{10.5}
\]

**Proof.** Let us remark that the proof of the following will use heavily the results and notation of Section 6.

First, by monotonicity of \( \beta \) (Remark 2.3), and the fact \( \mu \) is supported in \( B_1 \), we have

\[
\sum_{\alpha \in \mathbb{Z} : 2^\alpha \leq 2} \beta_{\mu,2,\epsilon_\beta}^k(z,2^\alpha)^2 \leq c(k)\delta^2. \tag{10.6}
\]

Choose a scale \( \rho(n, \epsilon, \Gamma) < 1/20 \) as in Theorem 4.13 part C) so that any bad ball \( B_r(x) \) with \( m = m_0(n, \rho)e \) as in Theorem 4.13 part A), must have \( \mu(B_r(x)) \leq er^k \).

Define the covering pair \( C = C_0 = B_1(0) \), with \( r_x \equiv 0 \). With our choice of \( \rho \) above, any ball \( B_r(x) \) must be either stop or good (as per Definition 5.2), but cannot be bad.

We can clearly assume without loss of generality that \( B_1 \) is a good ball, and build a good tree \( \mathcal{T} = \mathcal{T}(B_1) \) rooted in \( B_1 \) following the construction of Section 6.

Let \( T_\infty = \tau_\infty(\mathbb{R}^k) \) be the Lipschitz manifold of Theorem 6.6. By the above reasoning, \( B(\mathcal{T}) = \emptyset \), and by construction, \( C_0(\mathcal{T}) \subset T_\infty \). We deduce

\[
\mu(B_1 \setminus T_\infty) \leq \mu(B_1 \setminus C_0(\mathcal{T})) \leq c(n)\epsilon + c(n, \rho)\delta^2. \tag{10.7}
\]

This proves the Lemma. \( \square \)

In order to prove rectifiability, we need smallness of (10.4) relative to the lower density bound. However, rectifiability is a local property, and so we can afford to drop in scale until (10.4) is verified. This is the content of the following Lemma.

**Lemma 10.3.** Take \( \tau, \chi > 0 \). Suppose \( \mu \) is a non-negative Borel-regular measure supported in \( B_1 \), with upper mass bounds

\[
\mu(B_r(x)) \leq \Gamma r^k \quad \forall x \in B_1, 0 < r < 1, \tag{10.8}
\]
and satisfying

\[ \int_{B_1} \int_0^2 \beta_{\mu,2,\bar{\epsilon}_\beta}(z,r)^2 \frac{dr}{r} d\mu(z) < \infty. \]  \hspace{1cm} (10.9) 

Then for \( \mu \text{-a.e. } x \in B_1 \), there is a scale \( R_x > 0 \) so that

\[ \mu \left\{ z \in B_r(x) : \int_0^{2r} \beta_{\mu,2,\bar{\epsilon}_\beta}(z,s)^2 \frac{ds}{s} > \tau \right\} \leq \chi r^k \ \forall 0 < r < R_x. \]  \hspace{1cm} (10.10) 

**Proof.** Let \( F \) be the set of points \( x \) which fail (10.10). Fix an \( 0 < R < 1/4 \). For \( \mu \text{-a.e. } x \in F \), we can find a scale \( 0 < s_x < R \) for which

\[ s_x^k < \frac{1}{\chi} \mu \left\{ z \in B_{s_x}(x) : \int_0^{2s_x} \beta_{\mu,2,\bar{\epsilon}_\beta}(z,r)^2 \frac{dr}{r} d\mu(z) > \tau \right\}. \]  \hspace{1cm} (10.11) 

Choose a Vitali subcovering \( \{ B_{5s_x}(x_i) \}_i \) of \( \{ B_{5s_x}(x) \}_{x \in F} \), so that the \( s_i \)-balls are disjoint. Then we calculate

\[
\mu(F) \leq \sum_i \mu(B_{5s_x}(x_i)) \\
\leq c(k) \Gamma \sum_i s_i^k \\
\leq c(k) \frac{\Gamma}{\chi} \sum_i \frac{1}{\tau} \int_{B_{s_x}(x_i)} \int_0^{2s_x} \beta_{\mu,2,\bar{\epsilon}_\beta}(z,r)^2 \frac{dr}{r} d\mu(z) \\
\leq c(k) \frac{\Gamma}{\chi \tau} \int_{B_1} \int_0^{2R} \beta_{\mu,2,\bar{\epsilon}_\beta}(z,r)^2 \frac{dr}{r} d\mu(z). 
\]  \hspace{1cm} (10.12) 

By the dominated convergence theorem the RHS tends to 0 as \( R \to 0 \). This shows \( \mu(F) = 0 \). \( \square \)

We are now ready to prove our rectifiable Proposition.

**Proof of Proposition 10.1.** Begin by applying Theorem 2.14 part U) to deduce for our measure the estimate

\[ \mu(B_r(x)) \leq c(n)(b+M)r^k =: \Gamma r^k \ \forall x \in B_1, \ 0 < r \leq 1. \]  \hspace{1cm} (10.13) 

This implies in particular that

\[ \int_{B_1} \int_0^2 \beta_{\mu,2,\bar{\epsilon}_\beta}(z,r)^2 \frac{dr}{r} d\mu(z) < \infty, \]  \hspace{1cm} (10.14) 

and that \( \mu \) is absolutely continuous wrt \( \mathcal{H}^k \), in symbols \( \mu \ll \mathcal{H}^k \).

Choose

\[ \epsilon = \frac{\omega_k a}{c_4(n)16 \cdot 10^k}, \quad \delta^2 = \min \left\{ \frac{\omega_k a}{c_5(n,\epsilon,\Gamma)16 \cdot 10^k}, \delta_2(n,b,M,\epsilon) \right\}, \quad \chi = \frac{\omega_k a}{8 \cdot 10^k}, \quad \tau = \delta^2, \]  \hspace{1cm} (10.15) 

so that

\[ \chi + c_4 \epsilon + c_5 \delta^2 \leq \frac{\omega_k a}{4 \cdot 10^k}. \]  \hspace{1cm} (10.16) 

\[ \text{In some cases (as in Section 8.2) } \mu \text{ may already be upper-Ahlfors-regular, in the sense that } \mu(B_r(x)) \leq \Gamma(n,b,M)r^k \ \forall x \in B_1, \ 0 < r \leq 1, \text{ and one does not need these first few lines of the argument.} \]
We claim the following: let \( C \) be any closed set. Then there is a finite collection \( T_1, \ldots, T_N \) of Lipschitz \( k \)-manifolds so that
\[
\mu(B_1 \setminus (C \cup T_1 \cup \cdots \cup T_N)) \leq \frac{1}{2} \mu(B_1 \setminus C). \tag{10.17}
\]
To be specific, each \( T_i \) is the image of a bi-Lipschitz mapping \( \tau_i : B_{3/2}^k \to \mathbb{R}^n \).

Given this claim, rectifiability of \( \mu \) follows by a straightforward iteration. By taking \( C = \emptyset \), then \( C = T_1 \cup \cdots \cup T_N \), etc. we can inductively construct a nested sequence of closed \( k \)-rectifiable sets \( K_1 \subset K_2 \subset \ldots \) so that
\[
\mu(B_1 \setminus K_i) \leq 2^{-i} \mu(B_1) \leq 2^{-i} \Gamma. \tag{10.18}
\]
Letting \( K = \bigcup_i K_i \), then \( \mu(B_1 \setminus K) = 0 \). We have already established \( \mu \ll \mathcal{H}^k \), and therefore \( \mu \) is \( k \)-rectifiable.

Now we turn to the claim (10.17). By assumption and Lemma 10.3, we can find a measurable set \( A \subset B_1 \setminus (A \cup C) = 0 \) such that for all \( x \in A \), there is a radius \( s_x \) so that
\begin{itemize}
  \item[A)] \( B_{s_x}(x) \subset B_1 \setminus C \),
  \item[B)] \( \mu(B_{s_x/5}(x)) \geq \frac{\omega_{2k}}{2} (s_x/5)^k \),
  \item[C)] and \( \mu \left\{ z \in B_{s_x}(x) : \int_0^{2s_x} \beta_{s_x/5}(z, r)^2 \frac{dr}{r} > \delta^2 \right\} \leq \chi_{s_x} \). \tag{10.19}
\end{itemize}

In particular, by Lemma 10.2 and our choice of constants, there is for every such \( x \) a Lipschitz \( k \)-manifold \( T_x \) so that
\[
\mu(B_{s_x}(x) \setminus T_x) \leq \frac{\omega_k a}{4 \cdot 10^k} s_x^k. \tag{10.20}
\]

Let \( \{B_{s_i}(x_i)\} \) be a Vitali subcover of \( \bigcup_{x \in A} B_{s_x}(x) \), so that the \( B_{s_i/5} \)-balls are disjoint, and write \( T_i = T_{x_i} \). We then calculate
\[
\mu \left( B_1 \setminus (C \cup \bigcup_i T_i) \right) \leq \sum_i \mu(B_{s_i}(x_i) \setminus T_i), \tag{10.21}
\]
\[
\leq \sum_i \frac{\omega_k a}{4 \cdot 10^k} s_i^k, \tag{10.22}
\]
\[
\leq \frac{1}{4} \sum_i \frac{\omega_k a}{2} (s_i/5)^k, \tag{10.23}
\]
\[
\leq \frac{1}{4} \mu(B_1 \setminus C). \tag{10.24}
\]
This proves the Claim, taking \( N \) sufficiently large.

Now we are in a position to obtain Theorem 2.19 as a simple corollary of the previous result.

**Proof of Theorem 2.19.** By assumption there is some \( x \in \text{spt} \mu \), with \( \beta_{\mu,2,0}^k(x, 2) < \infty \). Therefore if \( V = V(x, 2) \), and
\[
U_j = \{ z \in B_1 : \text{dist}(z, V) > 1/j \}, \tag{10.25}
\]
then \( \mu U_j \) is Radon, and satisfies
\[
\Theta^*,k(\mu U_j, x) = \Theta^*,k(\mu, x) \quad \forall x \in U_j. \tag{10.26}
\]

For each integer \( i, j \), define the Borel set
\[
K_{ij} = \left\{ z \in U_j : \int_0^2 \beta_{\mu,2,0}(z,r)^2 dr/r \leq i, \quad \text{and} \quad \Theta^*,k(\mu, z) > 1/i \right\}. \tag{10.27}
\]

By assumption and (10.26), we have \( \mu(U_j \setminus \cup_i K_{ij}) = 0 \) for each \( j \).

Using (10.26) and Lemma 4.4, we have
\[
\mathcal{H}^k K_{ij} \leq i \cdot \mu K_{ij}, \tag{10.28}
\]
and therefore by monotonicity of \( \beta \) we know
\[
\int_0^2 \beta_{\mathcal{H}^k K_{ij},2,0}(z,r)^2 dr/r \leq i^2 \quad \forall z \in K_{ij}. \tag{10.29}
\]

Since \( \mu U_j \) is a finite measure, so is \( \mathcal{H}^k K_{ij} \), and hence we have the standard density estimates
\[
2^{-k} \leq \Theta^*,k(\mathcal{H}^k K_{ij}, x) \leq 1 \quad \text{for } \mathcal{H}^k\text{-a.e. } x \in K_{ij}. \tag{10.30}
\]

Using Proposition 10.1, we deduce \( K_{ij} \) is rectifiable.

Clearly
\[
K_0 = V \cup \bigcup_{ij} K_{ij} \tag{10.31}
\]
is still a rectifiable set, and \( \mu(B_1(0) \setminus K_0) = 0 \). This concludes the first part of the proof.

Under the additional assumption that \( \Theta^*_*(\mu, x) < \infty \) for \( \mu \)-a.e. \( x \), define the Borel sets
\[
W_i = \left\{ z \in B_1 : \int_0^2 \beta_{\mu,2,0}(z,r)^2 dr/r \leq i, \quad \text{and} \quad \Theta^*_*(\mu, z) \leq i \right\}. \tag{10.32}
\]

By assumption, \( \mu(B_1 \setminus \cup_i W_i) = 0 \).

For any \( i \) we trivially have that
\[
\int_0^2 \beta_{\mu,W_i,2,0}(z,r)^2 dr/r \leq i, \quad \text{and} \quad \Theta^*_*(\mu W_i, z) \leq i \quad \forall z \in W_i. \tag{10.33}
\]

Therefore, by Theorem 2.14 part U),
\[
\mu(W_i \cap B_r(x)) \leq c(i,n)r^k \quad \forall x \in B_1, 0 < r \leq 1. \tag{10.34}
\]

Therefore \( \mu W_i \ll \mathcal{H}^k \) for each \( i \), and hence \( \mu \ll \mathcal{H}^k \) also. \( \square \)

11. REMAINING THEOREMS

In this last section, we gather all the proofs of the secondary results that were stated in the introduction and in Section 2.4 “Applications”. The proofs are more-or-less direct corollaries of the results already proven.
11.1. **Proof of the statements in the introduction.** Theorem 1.1 follows from Theorem 2.6 by choosing the trivial covering pair \( C = C_0 = B_1(0) \), and \( r_x \equiv 0 \). In a similar way, Corollary 1.2 is a direct consequence of Corollary 2.12.

**Proof of Theorem 1.3.** Consider the measure \( \mu_{x,r}(A) := r^{-k} \mu(x + rA) \). Depending on whether \( \mu \) satisfies the condition A), B), or C), then \( \mu_{x,r} \) satisfies (respectively):

A)
\[
\int_0^2 \beta_{\mu_{x,r},2,0}(z,r)^2 \frac{dr}{r} \leq M \quad \text{for } \mu_{x,r}\text{-a.e. } z \in B_1,
\]

B) or
\[
\int_{B_1} \int_0^2 \beta_{\mu_{x,r},2,0}(z,r)^2 \frac{dr}{r} d\mu_{x,r}(z) \leq M^2,
\]

C) or
\[
\int_{B_1} \int_0^2 \beta_{\mu_{x,r},2,0}(z,r)^2 \frac{dr}{r} d\mu_{x,r}(z) \leq M \mu_{x,r}(B_1).
\]

In cases A) or B), apply Theorem 2.11 with \( \Gamma = M \), to deduce
\[
\mu_{x,r}(B_1) \leq c(n)(M + b + \epsilon) + M \quad \forall \epsilon > 0.
\]

In case C), apply Theorem 2.11 with \( \Gamma = \frac{1}{2} \mu_{x,r}(B_1) \), to deduce
\[
\mu_{x,r}(B_1) \leq c(n)(2M + b + \epsilon) + \frac{1}{2} \mu_{x,r}(B_1) \quad \forall \epsilon > 0.
\]

Therefore, in all cases, we have \( \mu(B_r(x)) \leq c(n)(M + b)r^k \).

The rectifiability is a direct consequence of Theorem 2.19.

**Proof of Corollary 1.4.** We first note that if \( \mathcal{H}^k(S) < \infty \) (as is automatic in case C), then
\[
2^{-k} \leq \Theta^{*,k}(\mu, x) \leq 1 \quad \text{for } \mathcal{H}^k\text{-a.e. } x,
\]

and can therefore apply Theorem 1.3 directly.

Let us consider cases A) and B). In either case, there is an \( x \in \text{spt}\mu \cap B_1 \) so that
\[
\beta_{\mu,2}(x,2) < \infty.
\]

Trivially, \( S \cap V^k(x,2) \) is \( k \)-rectifiable, and
\[
\mathcal{H}^k(V^k(x,2) \cap S \cap B_r(x)) \leq c(n)r^k \quad \forall x \in B_1, 0 < r \leq 1.
\]

On the other hand, let us define for each \( i \)
\[
S_i = \{ z \in B_1 : d(z, V^k(x,2)) > 1/i \}.
\]

Then \( \mathcal{H}^k(S_i) < \infty \), and by monotonicity of \( \beta \) the measure \( \mathcal{H}^k, S_i \) satisfies either condition A) or B). We deduce from Theorem 1.3 that \( S_i \) is \( k \)-rectifiable, and admits Hausdorff bounds
\[
\mathcal{H}^k(S_i \cap B_r(x)) \leq c(n)(1 + M)r^k \quad \forall x \in B_1, 0 < r \leq 1.
\]

Taking \( i \to \infty \), we deduce the required estimate. \( \square \)
We conclude this section with the proof of Corollary 1.5, which follows immediately from our measure bounds and the result of [GS].

**Proof of Corollary 1.5.** By Theorem 1.3, we obtain that $\mu$ is upper Ahlfors regular. In particular, we have the uniform bounds $\mu(B_r(x))/r^k \leq c(n)(M+b)$ for all $x$ and $0 < r \leq 1$. Therefore condition (1.17) is satisfied, and the conclusion follows by [GS].

11.2. Applications. Most of the work the following Theorems is in decomposing the measure in various ways via Theorem 2.6.

**Proof of Theorem 2.17.** Apply Theorem 2.6 at scale $B_1$, with the trivial covering pair $C \equiv C_0 = B_1$ (so $r_x \equiv 0$), and $\Gamma = M$, to obtain a closed, $k$-rectifiable set $\mathcal{K}_0$ with

$$\mathcal{H}^k(\mathcal{K}_0 \cap B_r(x)) \leq c(n)r^k \forall x \in B_1, 0 < r \leq 1, \quad \text{and} \quad \mu(B_1 \setminus \mathcal{K}_0) \leq c(n)(M + \epsilon).$$

(11.11)

Since $\mathcal{K}_0$ is closed, we know

$$\Theta^x_k(\mu_0(B_1 \setminus \mathcal{K}_0), x) = \Theta^x_k(\mu, x) \quad \text{for every} \ x \in B_1 \setminus \mathcal{K}_0. \quad (11.12)$$

Consider the Borel sets

$$U_+ = \{ x \in B_1 \setminus \mathcal{K}_0 : \Theta^x_k(\mu, x) > 0 \} \quad \text{and} \quad \mathcal{K}_\infty = \{ x \in B_1 \setminus \mathcal{K}_0 : \Theta^x_k(\mu, x) = \infty \}. \quad (11.13)$$

Since $\mu_0(B_1 \setminus \mathcal{K}_0) < \infty$, we know by (11.12) and Lemma 4.4 that $\mathcal{H}^k(\mathcal{K}_\infty) = 0$, and by Radon-Nikodym theorem

$$0 < \Theta^x_k(\mu_0(U_+ \setminus (\mathcal{K}_0 \cup \mathcal{K}_\infty)), x) < \infty \quad \text{for} \ \mu\text{-a.e.} \ x \in U_+ \setminus (\mathcal{K}_0 \cup \mathcal{K}_\infty). \quad (11.14)$$

Define

$$\mu_\ell = \mu_0(U_+ \setminus (\mathcal{K}_0 \cup \mathcal{K}_\infty)). \quad (11.15)$$

By monotonicity of $\beta$,

$$\int_0^2 \beta^k_{\mu,2,0}(z, r)^2 dr/r < \infty \quad \mu_\ell \text{-a.e.} \ z, \quad (11.16)$$

and therefore by Theorem 2.19 $\mu_\ell$ is rectifiable.

Set $\mathcal{K}_h = \mathcal{K}_0 \cup \mathcal{K}_\infty$, and $\mu_h = \mu_0\mathcal{K}_h$. Since $\mathcal{K}_0$ is rectifiable, and $\mathcal{H}^k(\mathcal{K}_\infty) = 0$, then $\mathcal{K}_h$ is rectifiable also. The volume bounds of part A) follow since $\mathcal{H}^k(\mathcal{K}_\infty) = 0$. Using Theorem 2.6 part C) with Lemma 4.4, we have

$$\mathcal{H}^k(\mathcal{K}_h \cap U) = \mathcal{H}^k(\mathcal{K}_0 \cap U) \leq c(k)\epsilon^{-1}\mu(U) \quad \forall \text{open} \ U. \quad (11.17)$$

Last, set $U_0$ to be the points $x \in B_1 \setminus \mathcal{K}_0$ for which $\Theta^x_k(\mu, x) = 0$. By definition,

$$B_1 \subset (\mathcal{K}_0 \cup \mathcal{K}_\infty) \cup U_+ \cup U_0, \quad (11.18)$$

thus we can set $\mu_0 = \mu_0 U_0$, and this completes the decomposition. Since $\mu_0 \leq \mu_0(B_1 \setminus \mathcal{K}_0)$, the mass bound of Theorem 2.17 part C) is immediate. The density assertion in C) is by construction. \hfill \Box

In a similar way, to prove Theorem 2.18, we decompose $\mu$ as in Theorem 2.17 at every scale.
Proof of Theorem 2.18. Let $K_0$, $K_\infty$, $U_+$, and $U_0$ be defined as in the proof of Theorem 2.17. Thus, $K_0$ is closed, rectifiable, and satisfies
\begin{equation}
\mathcal{H}^k(K_0 \cap B_r(x)) \leq c(n) r^k \quad \forall x \in B_1, 0 < r \leq 1, \quad \text{and} \quad \mu(B_1 \setminus K_0) \leq c(n)(\epsilon + M). \tag{11.19}
\end{equation}
Moreover, by Theorem 2.6 part C) and Lemma 4.4,
\begin{equation}
\mathcal{H}^k \setminus K_0 \leq c(k) \epsilon^{-1} \mu \quad \text{on open sets}. \tag{11.20}
\end{equation}
Recall that $\mathcal{H}^k(K_\infty) = 0$.

Define the intermediate measure
\begin{equation}
\mu_1 = \mu_{L}(B_1 \setminus (K_0 \cup K_\infty)). \tag{11.21}
\end{equation}
Then $\mu_1$ is Radon.

Let $\{B_{s_i}(x_i)\}_i$ be a countable collection of balls, so that centers and radii $\{(x_i, s_i)\}_i$ are dense in $B_1 \times (0, 1)$. For each $i$, apply Theorem 2.6 to $\mu_1$ at scale $B_{s_i}(x_i)$, with covering pair $C = C_0 = B_{s_i}(x_i)$ and $\Gamma = M$.

This produces a sequence of closed, $k$-rectifiable sets $K_i$ such that
\begin{equation}
\mu_1(B_{s_i}(x_i) \setminus K_i) \leq c(n)(\epsilon + M)s_i^k \quad \text{and} \quad \mathcal{H}^k \setminus K_i \leq c(k) \epsilon^{-1} \mu_{1L}K_i. \tag{11.22}
\end{equation}

This last statement follows from Theorem 2.6 part C), Lemma 4.4, and finiteness of $\mu_1$. Since each $K_i$ is Borel, we have
\begin{equation}
\mathcal{H}^k \setminus \bigcup_i K_i \leq c(k) \epsilon^{-1} \mu_{1L} \bigcup_i K_i \quad \text{on Borel sets}. \tag{11.23}
\end{equation}

We deduce by monotonicity of $\beta$ that the measure $\eta = \mathcal{H}^k \setminus \bigcup_i K_i$ satisfies
\begin{equation}
\int_{B_r(x)} \int_0^2 \beta^k_{2,0}(z, r)^2 \frac{dr}{r} d\eta(z) \leq c(k) M^2 / \epsilon^2 \quad \forall x \in B_1, 0 < r \leq 1. \tag{11.24}
\end{equation}

Since $\mu_1$ is finite, $\mathcal{H}^k(\bigcup_i K_i) < \infty$, and so $\Theta^{*,k}(\eta, x) \leq 1$ at $\eta$-a.e. $x$. Theorem 2.14 part U) implies that
\begin{equation}
\eta(B_r(x)) \leq c(n)(1 + M/\epsilon) r^k \quad \forall x \in B_1, 0 < r \leq 1. \tag{11.25}
\end{equation}

Set
\begin{equation}
K_h = K_0 \cup K_\infty \cup \bigcup_i K_i, \quad \mu_\ell = \mu_{1L}(U_+ \setminus \bigcup_i K_i), \quad \mu_0 = \mu_{1L}(U_0 \setminus \bigcup_i K_i). \tag{11.26}
\end{equation}

With this definition, $\mu_\ell$ is $k$-rectifiable, and $\Theta^{*,k}(\mu_\ell, x) = 0$ at $\mu_0$-a.e. $x$, as in the proof of Theorem 2.17. Upper-Ahlfors-regularity of $\mu_\ell + \mu_0$ follows by (11.22). \hfill \Box

We conclude this section with the proofs of Theorems 2.20 and 2.21, which are immediate consequences of previous results.

Proof of Theorem 2.20. Define the covering pair $C = C_+ = \{x_i\}_i$ with $r_{x_i} = r_i$. Then Theorem 2.20 is an immediate Corollary of Theorem 2.11. \hfill \Box

Proof of Theorem 2.21. By lower-semi-continuity of $\beta$ (Lemma 4.5) there is no loss in assuming that $S$ is closed. As in the Proof of Corollary 1.4, hypothesis (2.35) implies $S$ is $\sigma$-finite. Therefore $\Theta^{*,k}(\mathcal{H}^k \setminus S, x) \leq 1$ at $\mathcal{H}^k$-a.e. $x \in S$. In order to conclude, simply choose the covering pair $C = C_0 = S$, and apply Theorem 2.11. \hfill \Box
REFERENCES

[AS] J. AZZAM and R. SCHUL, An analyst’s traveling salesman theorem for sets of dimension larger than one, preprint, arXiv 1609.02892. Available at https://arxiv.org/abs/1609.02892.

[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.

[BS] M. BADGER and R. SCHUL, Multiscale analysis of 1-rectifiable measures ii: characterizations, preprint, arXiv 1602.03823. Available at https://arxiv.org/abs/1602.03823.

[BS15] Multiscale analysis of 1-rectifiable measures: necessary conditions, Math. Ann. 361 (2015), 1055–1072. MR 3319560. http://dx.doi.org/10.1007/s00208-014-1104-9. Available at http://dx.doi.org/10.1007/s00208-014-1104-9.

[Dav98] G. DAVID, Uncountifiable \( s \)-sets have vanishing analytic capacity, Rev. Mat. Iberoamericana 14 (1998), 369–479. MR 1654535. Zbl 0913.30012. http://dx.doi.org/10.4171/RMI/242.

[DS93] G. DAVID and S. SEMMES, Analysis of and on uniformly rectifiable sets, Mathematical Surveys and Monographs 38, American Mathematical Society, Providence, RI, 1993. MR 1251061. Zbl 0832.42008. http://dx.doi.org/10.1090/surv/038.

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

[GS] D. GIRELA-SARRIN, Geometric conditions for the \( L^p \)-boundedness of singular integral operators with odd kernels with respect to measures with polynomial growth in \( r^d \), arXiv:1505.07264. Available at https://arxiv.org/abs/1505.07264.

[INT] B. JAYE, F. NAZAROV, and X. TOLSA, The measures with an associated square function operator bounded in \( L^2 \), arXiv:1612.04754. Available at http://dx.doi.org/10.1007/1612.04754.

[Jon90] P. W. JONES, Rectifiable sets and the traveling salesman problem, Invent. Math. 102 (1990), 1–15. MR 1069238. Zbl 0731.30018. http://dx.doi.org/10.1007/BF01233418. Available at http://dx.doi.org/10.1007/BF01233418.

[Mai95] 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.

[Mis16] M. MISKIEWICZ, Discrete reifenberg-type theorem, (2016), arXiv:1612.02461. Available at https://arxiv.org/pdf/1612.02461.pdf.

[NV15] A. NABER and D. VALTORTA, The singular structure and regularity of stationary and minimizing varifolds, (2015), arXiv:1505.03428. Available at https://arxiv.org/abs/1505.03428.

[NV17] Rectifiable-reifenberg and the regularity of stationary and minimizing harmonic maps, Annals of Mathematics 185 (2017), Available at https://arxiv.org/abs/1504.02043.

[NTV97] F. NAZAROV, S. TREIL, and A. VOLBERG, Cauchy integral and Calderon-Zygmund operators on non-homogeneous spaces, Internat. Math. Res. Notices (1997), 703–726. MR 1470373. Zbl 0889.42013. http://dx.doi.org/10.1155/S1073792897000469.

[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/lectures/08092002.pdf.

[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.

[Tol99] X. TOLSA, \( L^2 \)-boundedness of the Cauchy integral operator for continuous measures, Duke Math. J. 98 (1999), 269–304. MR 1695200. Zbl 0945.30032. http://dx.doi.org/10.1215/S0012-7094-99-09808-3.
[Tol15] 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.