

# QUANTITATIVE METRIC DENSITY AND CONNECTIVITY FOR SETS OF POSITIVE MEASURE

GUY C. DAVID AND BRANDON OLIVA

**ABSTRACT.** We show that in doubling, geodesic metric measure spaces, (including, for example, Euclidean space) sets of positive measure have a certain large-scale metric density property. As an application, we prove that a set of positive measure in the unit cube of  $\mathbb{R}^d$  can be decomposed into a controlled number of subsets that are “well-connected” within the original set, along with a “garbage set” of arbitrarily small measure. Our results are *quantitative*, i.e., they provide bounds independent of the particular set under consideration.

## 1. INTRODUCTION

**1.1. Background: qualitative vs. quantitative results in analysis.** Many important theorems in analysis are “qualitative” or “infinitesimal”. They assert that a certain desirable small-scale limiting behavior occurs, but do not provide any guarantee of a fixed scale at which such a behavior occurs up to a given error.

A basic example is Rademacher’s theorem: a Lipschitz function from  $\mathbb{R}^d$  to  $\mathbb{R}$  is differentiable almost everywhere. While this is a wonderful result, neither the theorem nor its typical proofs (see, e.g., [11, 6]) guarantee anything about how far one might have to zoom in at a generic point to find a scale where this Lipschitz function looks like an affine function with at most, say, 1% error.

With completely different arguments, however, such a guarantee can be made. Let us say that a function  $f$  is “ $\epsilon$ -close to affine” on a ball  $B$  of radius  $r$  if there is an affine function  $A$  such that  $\sup |f - A| \leq \epsilon r$  on  $B$ . A result of Dorronsoro [4] implies the following quantitative statement. *For each  $d \in \mathbb{N}$  and  $\epsilon > 0$ , there is an  $r_0 > 0$  such that every 1-Lipschitz  $f: [0, 1]^d \rightarrow \mathbb{R}$  has a ball of radius at least  $r_0$  on which it is  $\epsilon$ -close to being affine.* (See also expositions by Semmes [5, B.29] and Young [13, Theorem 2.4].) In fact, more is true: in some sense “most” balls in  $[0, 1]^d$  have the property that  $f$  is  $\epsilon$ -close to affine on them. Note that the parameter  $r_0$  is a guarantee that can be made *independent of the particular 1-Lipschitz function  $f$* .

There is by now a whole literature in quantitative analysis that is too large to survey here, including both new phenomena and quantitative analogs of classical results. In addition to the paper of Dorronsoro referenced above, some of the foundational texts are by Jones [9] and David–Semmes [3].

**1.2. The Lebesgue density theorem.** Another “qualitative” result in analysis, in the sense described above, is the Lebesgue density theorem. Using  $\lambda$  to denote Lebesgue measure in  $\mathbb{R}^n$ , a version of the theorem is as follows:

**Theorem 1.1** (Lebesgue Density Theorem). *Let  $E \subseteq \mathbb{R}^d$  be measurable. Then at almost every point  $x \in E$ ,*

$$\lim_{r \rightarrow 0} \frac{\lambda(B(x, r) \cap E)}{\lambda(B(x, r))} = 1.$$

In other words, at infinitesimally small scales,  $E$  “fills up” nearly all the measure of a ball. We note that the same statement holds in arbitrary complete, doubling metric measure spaces; see below for the definitions of these terms and [6] for the result in this generality.

---

*Date:* April 19, 2024.

*2020 Mathematics Subject Classification.* 30L99, 28A75.

G. C. David was partially supported by the National Science Foundation under Grant No. DMS-2054004. The main results of the paper form part of the master’s thesis of B. Oliva.

As a naïve quantitative analog of the Lebesgue density theorem, paralleling the quantitative Rademacher theorem mentioned above, one might hope for the following:

$$(1.1) \quad \begin{aligned} &\text{For each } d \in \mathbb{N} \text{ and } \alpha, \epsilon > 0, \text{ there is an } r_0 > 0 \text{ such that if } E \subseteq [0, 1]^d \text{ has } \lambda(E) \geq \alpha, \\ &\text{then there is a ball } B \text{ of radius } r \geq r_0 \text{ such that } \frac{\lambda(B \cap E)}{\lambda(B)} \geq 1 - \epsilon. \end{aligned}$$

This statement turns out to be **false**, however, as we demonstrate with a straightforward counterexample in Section 3.

**1.3. Quantitative metric density.** To give a correct quantitative statement inspired by the Lebesgue density theorem, we view density in a metric rather than measure-theoretic sense.

**Definition 1.2.** We define the *sparsity* of a set  $E$  in a closed ball  $B$  of radius  $r$  by

$$\text{sparse}(E, B) = \frac{\sup\{\text{dist}(x, E \cap B) : x \in B\}}{r}.$$

If  $E \cap B = \emptyset$ , we regard the sparsity  $\text{sparse}(E, B)$  as undefined.

Thus, if  $\text{sparse}(E, B) = 0$ , then  $\overline{E \cap B} = B$ , although of course it may be that the measure of  $E \cap B$  is 0. If  $\text{sparse}(E, B)$  is small, we consider  $E$  to be “very dense” in  $B$ .

In order to formulate the idea that a set has small sparsity at “most” locations and scales, we also require the notion of a “multi-resolution family” of balls in a metric space.

**Definition 1.3.** Let  $(X, d)$  be a metric space. Let  $\{N_k : k = 0, 1, 2, \dots\}$  be a family of  $2^{-k}$ -nets in  $X$  satisfying  $N_0 \subseteq N_1 \subseteq N_2 \subseteq \dots$  (See Definition 2.4.)

Given a constant  $A \geq 1$ , a *multiresolution family* is the collection

$$\mathcal{B} = \{B(x, A \cdot 2^{-k}) : x \in N_k, 0 \leq k < \infty\}$$

Note that if  $A \geq 1$ , then the collection of all balls in  $\mathcal{B}$  of fixed radius  $A \cdot 2^{-k}$  forms a cover of  $X$ .

With these definitions, we can state a “Quantitative Metric Density Theorem”. The terminology employed in the statement (e.g., “doubling”, “geodesic”) will be defined in Section 2.

**Theorem 1.4.** Let  $(X, d, \mu)$  be a complete, geodesic, metric measure space equipped with a multiresolution family  $\mathcal{B}$ . Assume that  $\mu$  is a  $C$ -doubling measure,  $\text{diam } X = 1$ , and  $\mu(X) = 1$ .

Given  $\epsilon > 0$ , there is a  $K$  (depending only on  $\epsilon$ ,  $C$ , and  $A$ ) such that, for any set  $E \subseteq X$ , if we set

$$\mathcal{B}(E, \epsilon) = \{B \in \mathcal{B} : B \cap E \neq \emptyset, \text{sparse}(E, B) \geq \epsilon\}.$$

then

$$(1.2) \quad \sum_{B \in \mathcal{B}(E, \epsilon)} \mu(B) \leq K.$$

Theorem 1.4 says that a set of positive measure must look very dense in “most” balls that intersect it, in a way which is *independent of the particular set under consideration*.

For those unfamiliar with the sort of sum over balls of all scales in (1.2) (sometimes called a “Carleson packing” condition), we include the following small piece of intuition: Suppose we did not include the condition that  $\text{sparse}(E, B) \geq \epsilon$  in the sum in (1.2). In other words, supposed we simply summed  $\mu(B)$  over all balls  $B \in \mathcal{B}$  intersecting  $E$ . In that case, the above sum would diverge (if  $\mu(E) > 0$ ) because

$$\sum_{B \in \mathcal{B}, B \cap E \neq \emptyset} \mu(B) = \sum_{k=0}^{\infty} \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset \\ \text{rad}(B) = A2^{-k}}} \mu(B) \geq \sum_{k=0}^{\infty} \mu(E) = \infty.$$

Therefore, Theorem 1.4 says vaguely that  $E$  has sparsity  $< \epsilon$  in “most” balls in  $\mathcal{B}$  that intersect it.

To relate this back to the discussion above, we note that Theorem 1.4 will easily imply the following:

**Corollary 1.5.** *Let  $(X, d, \mu)$  be a complete, geodesic metric measure space with  $\mu$  a  $C$ -doubling measure. Suppose that  $\text{diam } X = 1$  and  $\mu(X) = 1$ .*

*Given  $\alpha, \epsilon > 0$ , there is an  $r_0 > 0$  (depending only on  $\alpha, \epsilon, C$ ) such that if a Borel set  $E \subset X$  has  $\mu(E) \geq \alpha$ , then there is a ball  $B$  of radius  $r \in [r_0, 1]$  such that*

$$(1.3) \quad E \cap B \neq \emptyset \text{ and } \text{sparse}(E, B) < \epsilon.$$

In other words, we guarantee a large scale (with size independent of  $E$ ) in which  $E$  is very dense.

**Remark 1.6.** The main ideas behind Theorem 1.4 and Corollary 1.5 are not really new, although we believe our presentation of the result has new features. The key intermediate step in the proof (Proposition 4.4) can essentially be found in [3, Section IV.1.2] and [12, Lemma 1.4], although stated with more restrictive assumptions.

Thus, our main purpose in stating and proving Theorem 1.4 is the application to the original Theorem 1.9 below. As a secondary concern, we take the opportunity to provide a general proof of the result in the setting of abstract metric measure spaces.

**Remark 1.7.** The assumptions that  $\text{diam } X = 1$  and  $\mu(X) = 1$  in Theorem 1.4 and Corollary 1.5 are not really restrictive, since one can always normalize to these parameters. See Remark 4.8 for an example.

**1.4. Quantitative connectivity.** Our main new result is about sets of positive measure in the unit cube of  $\mathbb{R}^d$ . We show that such a set can be decomposed, up to a “garbage set” of small measure, into pieces that are “well-connected” within the larger set, in a way we now make precise.

Given  $\eta > 0$ , an  $\eta$ -chain in a metric space is a finite list of points  $(x_0, \dots, x_m)$  such that  $d(x_i, x_{i+1}) < \eta$  for each  $i \in \{0, \dots, m-1\}$ . We say that the chain is “from  $x$  to  $y$ ” if  $x = x_0$  and  $y = x_m$ . The *length* of the  $\eta$ -chain is the sum  $\sum_{i=0}^{m-1} d(x_i, x_{i+1})$ .

**Definition 1.8.** Let  $E \subset [0, 1]^d$ , and let  $F \subset E$ . We say that  $F$  is  $\delta$ -well-connected in  $E$  if for every  $x, y \in F$ , there is a  $\delta|x - y|$ -chain from  $x$  to  $y$  in  $E$  with length at most  $(1 + \delta)|x - y|$ .

Thus, if  $F$  is well-connected in  $E$ , then any pair of points in  $F$  can be joined by a discrete path in  $E$  whose *steps are small* (compared to the distance between the endpoints) and whose *length is almost as small as possible*.

We show that all sets of positive measure have “large” pieces that are well-connected within the original set, in a quantitative way.

**Theorem 1.9.** *Given  $d \in \mathbb{N}$ ,  $\alpha > 0$ , and  $\delta > 0$ , there is an  $M \in \mathbb{N}$  (depending only on  $d, \alpha, \delta$ ) with the following property: If  $E \subset [0, 1]^d$ , then there are sets  $F_1, F_2, \dots, F_M, Z \subset E$  such that:*

- (1)  $E = F_1 \cup F_2 \cup \dots \cup F_M \cup Z$ .
- (2)  $\lambda(Z) < \alpha$ .
- (3) For each  $i \in \{1, 2, \dots, M\}$ ,  $F_i$  is  $\delta$ -well-connected in  $E$ .

The “quantitative” aspect here is the fact that the number  $M$  of well-connected subsets depends only on the given parameters, and *not* on the original set  $E$ .

**Remark 1.10.** The set  $E$  is not assumed to be measurable in Theorem 1.9. However, the set  $Z$  can always be taken to be the intersection of  $E$  with a measurable subset of  $\mathbb{R}^d$  of measure  $< \alpha$ . Thus, if  $E$  is not measurable, then conclusion (2) of the theorem can be interpreted in the sense of Lebesgue outer measure.

The main steps in the proof of Theorem 1.9 are as follows: We apply Theorem 1.4 (suitably reinterpreted for dyadic cubes rather than balls) to show that in “most” dyadic cubes that touch  $E$ , the sparsity of  $E$  is small. (In fact, it is important to use expansions of dyadic cubes, but we elide this for now.) This will imply that most points of  $E$ , outside of a small garbage set  $Z$ , lie in a controlled number of “bad” dyadic cubes in which  $E$  is not very dense.

The next step is a “coding” argument, originally due to Jones [8], to separate  $E \setminus Z$  into a controlled number of sets  $F_i$  with the following property: If two points  $x, y \in E$  are in the same  $F_i$ , then there is a cube containing them of size  $\approx |x - y|$  in which  $E$  is very dense. This coding argument is Lemma 5.5; we take this opportunity to give a different proof than the ones presented in [8] and [2].

Finally, we complete the proof by showing that each set  $F_i$  is well-connected in  $E$ .

**Remark 1.11.** Theorem 1.9 bears some resemblance to the “Checkerboard Theorem” [10, Theorem 1.3] of Jones–Katz–Vargas, but it seems that neither result implies the other. The result of [10] shows that a set  $B$  of positive measure in the unit cube admits a (quantitatively) large subset  $A$  that is “checkerboard connected” within the original set. This means that any pair of points in  $A$  can be joined by a finite sequence of jumps parallel to the coordinate axes whose endpoints lie in  $B$ , and whose total length is controlled.

In Theorem 1.9, the jumps are not necessarily parallel to the coordinate axes (otherwise the length bound in Definition 1.8 would be impossible). However, the jumps in Theorem 1.9 are always *small* compared to the diameter of the path, whereas in [10, Theorem 1.3] this is not required.

The proofs are also different; for instance, [10] makes essential use of maximal function bounds, and we do not.

**Remark 1.12.** It seems likely that a version of Theorem 1.9 is true in the generality of doubling, geodesic metric spaces. One main ingredient, Theorem 1.4, is already stated and proven in this setting.

The main change needed to push the argument for Theorem 1.9 to this setting would be a version of dyadic cubes in abstract metric spaces. Such a construction exists [1, 7], but in the interest of keeping the present paper relatively streamlined we do not attempt to give the most general possible result here.

**1.5. Outline of the paper.** Basic definitions and notations are explained in Section 2. Section 3 contains a counterexample to the quantitative measure-theoretic density statement (1.1). Section 4 contains the proofs of Theorem 1.4 and Corollary 1.5, and Section 5 contains the proof of Theorem 1.9.

## 2. PRELIMINARIES

The section contains some additional definitions, notation, and basic facts used in the paper.

**Definition 2.1.** A *metric measure space* is a triple consisting of a space  $X$ , metric  $d$ , and (outer) measure  $\mu$ , denoted  $(X, d, \mu)$ , with an associated  $\sigma$ -algebra of measurable sets.

We use  $B(x, r)$  to denote the closed ball of radius  $r$  centered at  $x$ , i.e.,  $\{y \in X : d(x, y) \leq r\}$ .

**Definition 2.2.** Let  $(X, d)$  be a metric measure space and  $C > 0$ . A non-zero Borel regular measure  $\mu$  on  $X$  is called *doubling* (or *C-doubling* to emphasize the constant) if it assigns finite measure to every ball and for each ball  $B(x, r) \subset X$  we have

$$\mu(B(x, 2r)) \leq C \cdot \mu(B(x, r)).$$

Of course, Lebesgue measure  $\lambda$  on  $\mathbb{R}^n$  is an example of a doubling measure.

**Definition 2.3.** Let  $(X, d)$  be a metric space. A metric space is said to be a *doubling metric space* if there is an  $N \in \mathbb{N}$  such that for every ball  $B(x, 2r) \subset X$ , there are  $N$  balls  $B(x_k, r)$  such that

$$B(x, 2r) \subseteq \bigcup_{k=1}^N B(x_k, r).$$

In other words, a metric space is a doubling metric space if every ball can be covered by a controlled amount of balls of half its original radius. If  $(X, d, \mu)$  is a metric measure space with  $\mu$  a doubling measure, then  $(X, d)$  is a doubling metric space; see [6, p. 82].

**Definition 2.4.** Let  $(X, d)$  be a metric space. Given  $s > 0$ , an *s-net* is a set  $N \subset X$  that satisfies the following:

- (1) If  $x, y \in N$  and  $x \neq y$ , then  $d(x, y) \geq s$ .

(2) If  $z \in X$ , then there is an  $x \in N$  such that  $d(x, z) < s$ .

We say that a set is *s-separated* if only the first condition above is satisfied.

**Definition 2.5.** Let  $(X, d_X)$ ,  $(Y, d_Y)$  be two metric spaces, and let  $f : X \rightarrow Y$  be a function. We call  $f$  an isometric embedding if for every  $x_1, x_2 \in X$

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2).$$

**Definition 2.6.** Let  $(X, d)$  be a metric space. We say  $X$  is *geodesic* if for every  $x, y \in X$ , there is an isometric embedding  $f : [0, d(x, y)] \rightarrow X$  such that  $f(0) = x$  and  $f(d(x, y)) = y$ . We call this  $f$  a geodesic.

Our basic examples of doubling, geodesic spaces are  $\mathbb{R}^d$  and  $[0, 1]^d$  with the Euclidean metric. In these settings, we will also need the standard dyadic cube decomposition:

**Definition 2.7.** Let  $a_1, a_2, \dots, a_d$  be integers between 0 and  $2^{-k} - 1$ ,  $k \in \mathbb{N} \cup \{0\}$ . We define a *dyadic cube* in  $[0, 1]^d$  by

$$Q = [a_1 \cdot 2^{-k}, (a_1 + 1) \cdot 2^{-k}) \times [a_2 \cdot 2^{-k}, (a_2 + 1) \cdot 2^{-k}) \times \dots \times [a_d \cdot 2^{-k}, (a_d + 1) \cdot 2^{-k}).$$

One can triple the dyadic cube by

$$3Q = [(a_1 - 1) \cdot 2^{-k}, (a_1 + 2) \cdot 2^{-k}) \times [(a_2 - 1) \cdot 2^{-k}, (a_2 + 2) \cdot 2^{-k}) \times \dots \times [(a_d - 1) \cdot 2^{-k}, (a_d + 2) \cdot 2^{-k}).$$

By analogy, we can define  $5Q$ ,  $7Q$ , etc.

If  $d = 1$ , we refer to “dyadic intervals” rather than dyadic cubes.

The side length of a cube  $Q$  is written  $\text{side}(Q)$ . We denote the set of dyadic cubes in  $[0, 1]^d$  of side length  $2^{-k}$  by  $\Delta_k$ , with  $d$  understood from context, and we set  $\Delta = \bigcup_{k=0}^{\infty} \Delta_k$ .

A *child* of a dyadic cube  $Q$  is a dyadic cube  $R \subset Q$  with  $\text{side}(R) = \frac{1}{2}\text{side}(Q)$ .

If  $Q$  is a dyadic cube in  $\mathbb{R}^d$ , then  $3Q$  is the union of  $3^d$  disjoint dyadic cubes of the same side length as  $Q$ . We refer to these as the *cells* of  $3Q$ .

A basic fact about dyadic cubes is that, if  $Q, R \in \Delta$ , then either  $Q \cap R = \emptyset$ ,  $Q \subset R$ , or  $R \subset Q$ .

### 3. COUNTEREXAMPLE TO NAÏVE QUANTITATIVE LEBESGUE DENSITY

Here, we show by a simple 1-dimensional construction that the quantitative measure-theoretic density statement (1.1) cannot hold.

Let  $d = 1$ ,  $\alpha = \frac{1}{2}$ , and  $\epsilon = \frac{1}{10}$ . Let  $r_0 \in (0, 1)$  and  $k \in \mathbb{N}$  be the smallest integer such that  $2^{-k} \leq r_0$ .

Let  $E_k \subseteq [0, 1]$  be defined as:

$$E_k = \left(0, \frac{1}{2^{k+2}}\right) \cup \left(\frac{2}{2^{k+2}}, \frac{3}{2^{k+2}}\right) \cup \dots \cup \left(\frac{2^{k+2}-2}{2^{k+2}}, \frac{2^{k+2}-1}{2^{k+2}}\right).$$

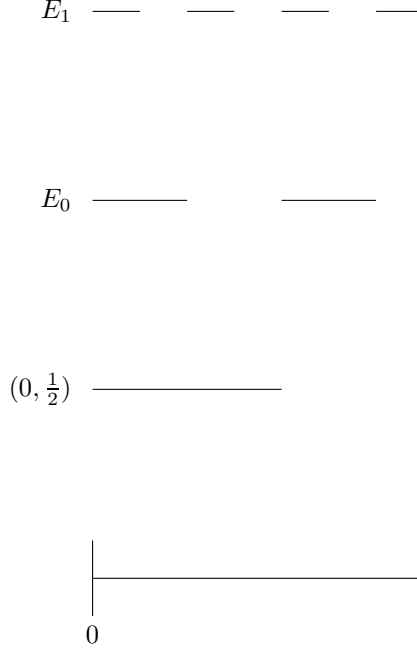
In other words,  $E_k$  is the union of  $2^{k+1}$  intervals each of measure  $2^{-(k+2)}$ , each interval separated by a length of  $2^{-(k+2)}$ , and the union is contained in  $[0, 1]$ . (See Figure 1 below.) Of course,  $\lambda(E_k) = \frac{1}{2} = \alpha$  for all  $k$ .

Note that we have,

$$\lambda(E_k) = 2^{k+1} \cdot 2^{-(k+2)} = \frac{1}{2}.$$

Let  $I$  be a dyadic interval in  $[0, 1]$  such that  $2^{-(k+1)} \leq \lambda(I) \leq 1$ . Then, observe that

$$(3.1) \quad \lambda(E_k \cap I) = \frac{1}{2}\lambda(I).$$

FIGURE 1. The sets  $E_k$  for  $k = 0, 1$ .

Take any  $x \in [0, 1]$  and  $r \in [r_0, 1]$ . We will show that

$$\frac{\lambda(E_k \cap [x - r, x + r])}{\lambda([x - r, x + r])} = \frac{\lambda(E_k \cap B(x, r))}{\lambda(B(x, r))} < 1 - \epsilon = 0.9,$$

thus disproving (1.1).

**Claim 3.1.** *There is a dyadic interval  $I \subseteq [x - r, x + r] \cap [0, 1]$  such that  $\frac{r}{2} \leq \lambda(I) \leq 2r$*

*Proof.* Let  $j \in \mathbb{N}$  be the first integer where  $2^{-j} \leq r$ . Let  $I$  be a dyadic interval in  $[0, 1]$  of length  $2^{-j}$  containing  $x$ . Note that  $I \subset [x - r, x + r]$ , so clearly  $\lambda(I) \leq 2r$ .

If  $2^{-j} < \frac{r}{2}$ , then we have

$$2^{-(j-1)} = 2^{-j+1} < r.$$

This contradicts our choice of  $j$ . Therefore,

$$\frac{r}{2} \leq \lambda(I) = 2^{-j} < r \leq 2r.$$

□

So since  $\frac{r}{2} \leq \lambda(I) \leq 2r$ , observe that

$$2^{-(k+1)} \leq \frac{r_0}{2} \leq \frac{r}{2} \leq \lambda(I) \leq 2r.$$

Therefore by (3.1) above,

$$\lambda(E_k \cap I) = \frac{1}{2} \lambda(I).$$

So,

$$\lambda(E_k \cap [x - r, x + r]) \leq 2r - \frac{1}{2} \lambda(I) \leq 2r - \frac{r}{4} = \frac{7}{4}r.$$

Hence,

$$\frac{\lambda(E_k \cap [x-r, x+r])}{\lambda([x-r, x+r])} \leq \frac{7}{4}r \cdot \frac{1}{2r} = \frac{7}{8} < 1 - \frac{1}{10} = \frac{9}{10}.$$

Thus, statement (1.1) fails in general.

#### 4. QUANTITATIVE METRIC DENSITY

Our goal in this section is to prove Theorems 1.4 and Corollary 1.5.

The following proposition guarantees the existence of nested families of nets, and therefore of multiresolution families, in any metric space. It is well-known, and we omit the standard proof using Zorn's lemma.

**Proposition 4.1.** *Let  $(X, d)$  be a metric space. Then there is a sequence of  $2^{-k}$ -nets  $N_k$  in  $X$  such that  $N_0 \subseteq N_1 \subseteq N_2 \subseteq \dots$*

We also need the following basic consequence of the doubling condition.

**Lemma 4.2.** *Let  $(X, d)$  be a doubling metric space, and let  $N$  be its doubling constant. Let  $\mathcal{B}$  be a multiresolution family with constant  $A > 0$ . Then there is a constant  $M > 0$ , depending only on  $N$  and  $A$ , such that for every  $x \in X$  and  $k \in \{0, 1, 2, \dots\}$ , at most  $M$  balls from  $\mathcal{B}$  of radius  $A2^{-k}$  can contain  $x$ .*

*Proof.* Fix  $x \in X$  and let  $k \in \{0, 1, 2, \dots\}$ . Let  $\{B_1, B_2, B_3, \dots\}$  be the balls of radius  $A \cdot 2^{-k}$  in  $\mathcal{B}$  containing  $x$ , where  $B_i = B(x_i, A \cdot 2^{-k})$  for  $i \in \{1, 2, \dots\}$  and some  $k \in \{1, 2, \dots\}$ . Thus, the points  $x_i$  are  $2^{-k}$ -separated and all contained in  $B(x, A2^{-k})$ .

Iterating the doubling property of  $(X, d)$  tells us that, given  $p \in \mathbb{N}$ , the ball  $B(x, A2^{-k})$  can be covered by at most  $N^p$  balls of radius  $\frac{1}{2^p} \cdot A \cdot 2^{-k}$ .

Therefore,  $B(x, A2^{-k})$  can be covered by at most  $M$  balls of radius smaller than  $\frac{1}{2} \cdot 2^{-k}$  for some  $M > 0$  depending only on  $N$  and  $A$ . Each such ball contains at most one element of the set  $\{x_1, x_2, \dots\}$ , so there at most  $M$  distinct elements in this set.  $\square$

In order to prove Theorem 1.4, we first introduce an intermediate definition.

**Definition 4.3.** Let  $(X, d, \mu)$  be a metric measure space with  $\mu$  Borel regular. Let  $E \subset X$  be a nonempty set. Given a ball  $B$  of radius  $r$ , we define

$$(4.1) \quad d_E(B) = \frac{1}{\mu(B)} \int_B \frac{\text{dist}(x, E)}{r} d\mu(x).$$

This is simply the average distance of a point in  $B$  to the set  $E$ , normalized by the radius of  $B$ . A simple calculation shows that if  $B \cap E \neq \emptyset$ , then  $0 \leq d_E(B) \leq 2$ .

We are now ready to state and prove the following proposition that will help us to prove Theorem 1.4. As noted in the introduction, the main idea of this proof can already be found in [3, Section IV.1.2] and [12, Lemma 1.4].

**Proposition 4.4.** *Let  $(X, d, \mu)$  be a doubling metric measure space equipped with a multiresolution family  $\mathcal{B}$  of constant  $A > 0$ . Assume  $\mu(X) = 1$  and  $\text{diam } X = 1$ .*

*Let  $E \subset X$ . Then*

$$(4.2) \quad \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset}} d_E(B) \cdot \mu(B) \leq C_X,$$

where  $C_X$  is a constant depending only on  $A$  and the doubling constant of  $X$  (and not on  $E$ ).

*Proof.* Note that we have

$$\sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset}} d_E(B) \cdot \mu(B) = \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset}} \int_B \frac{\text{dist}(x, E)}{r} d\mu(x) = \int_X \left( \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset \\ x \in B}} \frac{\text{dist}(x, E)}{r} \right) d\mu(x),$$

where the  $r$  in the summand denotes the radius of the ball  $B \in \mathcal{B}$  being summed over.

We will now bound the integrand. Note that if  $x \in \bar{E}$ , then  $\text{dist}(x, E) = 0$ . So fix  $x \in X$  with  $x \notin \bar{E}$ , and consider

$$(4.3) \quad \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset \\ x \in B}} \frac{\text{dist}(x, E)}{r}.$$

Let  $N$  be the biggest integer such that  $A \cdot 2^{-N} \geq \frac{1}{2} \cdot \text{dist}(x, E)$ . For balls of radius  $r = A \cdot 2^{-N}$ , we have

$$\frac{\text{dist}(x, E)}{r} \leq \frac{2 \cdot A \cdot 2^{-N}}{r} = \frac{2 \cdot A \cdot 2^{-N}}{A \cdot 2^{-N}} = 2.$$

Note that there are no balls in the summation in (4.3) with radius  $A \cdot 2^{-(N+1)}, A \cdot 2^{-(N+2)}, \dots$  because such a ball would have radius  $r \leq A \cdot 2^{-(N+1)} < \frac{1}{2} \text{dist}(x, E)$ , and such a ball could not touch both  $E$  and  $x$ . Recall that, since  $N$  is the biggest integer such that  $A \cdot 2^{-N} \geq \frac{1}{2} \cdot \text{dist}(x, E)$ , we have

$$\text{dist}(x, E) \leq A \cdot 2^{-(N-1)}.$$

If  $j \in \{0, 1, 2, \dots, N\}$  and  $B$  is a ball in the sum (4.3) of radius  $A \cdot 2^{-(N-j)}$ , then the summand

$$\frac{\text{dist}(x, E)}{r} \leq \frac{A \cdot 2^{-(N-1)}}{r} = \frac{A \cdot 2^{-(N-1)}}{A \cdot 2^{-(N-j)}} = \frac{1}{2^{j-1}}.$$

For each fixed radius  $A \cdot 2^{-k}$  ( $k \in \{0, 1, \dots, N\}$ ), Lemma 4.2 tells us that there are at most  $M$  balls containing  $x$  with radius equal to  $A \cdot 2^{-k}$ , where  $M$  depends on  $N$  and  $A$ . Therefore we have,

$$\sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset \\ x \in B}} \frac{\text{dist}(x, E)}{r} \leq M \cdot (2 + 1 + \frac{1}{2} + \frac{1}{4} + \dots) \leq M \cdot \left( \sum_{i=-1}^{\infty} \frac{1}{2^i} \right) = M \cdot (2 + 2) = 4M.$$

Let  $C_X = 4M$ . We now have

$$\int_X \left( \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset \\ x \in B}} \frac{\text{dist}(x, E)}{r} \right) d\mu(x) \leq \int_X C_X d\mu(x) = C_X \cdot \mu(X) = C_X.$$

Therefore

$$\sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset}} d_E(B) \cdot \mu(B) \leq C_X.$$

□

The following lemma connects  $\text{sparse}(E, B)$  to  $d_E(B)$ . There is a minor technical issue that we must overcome, which is that  $\text{sparse}(E, B)$  considers only points in  $E \cap B$ , while  $d_E(B)$  considers points of  $E$  that may lie outside of  $B$ . The geodesic assumption allows us to link the two quantities despite this.



**Lemma 4.5.** *Let  $(X, d, \mu)$  be a complete, geodesic metric measure space, and  $\mu$  a  $C$ -doubling measure. Let  $E \subset X$  and  $B$  be a ball in  $X$ .*

*Given  $\epsilon > 0$ , there is a  $\delta > 0$ , depending only on  $\epsilon$  and  $C$ , such that if  $d_E(B) < \delta$ , then*

$$\text{sparse}(E, B) < \epsilon.$$

*Proof.* We will prove this by contrapositive. Without loss of generality, let  $0 < \epsilon < 1$  be given. Suppose we have a metric measure space  $(X, d, \mu)$  satisfying the assumptions on the lemma, a subset  $E \subset X$ , and a ball  $B = B(p, r)$  such that  $\text{sparse}(E, B) \geq \epsilon$ . We will bound  $d_E(B)$  from below by a constant depending only on  $\epsilon$  and  $C$ .

The assumption on sparsity translates to

$$\sup\{\text{dist}(x, E \cap B) : x \in B\} \geq \epsilon r.$$

Thus, there is a point  $x \in B$  such that  $\text{dist}(x, E \cap B) > \epsilon \frac{3r}{4}$ . The ball  $B(x, \frac{3\epsilon}{4}r)$  therefore contains no points of  $E \cap B$ .

**Claim 4.6.** *There is a point  $y \in B(x, \frac{\epsilon}{4}r)$  such that*

$$B(y, \frac{\epsilon}{4}r) \subseteq B \cap B(x, \frac{3\epsilon}{4}r)$$

*Proof.* If  $d(p, x) < \frac{\epsilon}{4}$ , set  $y = x$ . The conclusion then follows immediately from the triangle inequality, recalling our assumption that  $\epsilon < 1$ .

If  $d(p, x) \geq \frac{\epsilon}{4}$ , we do the following. Let  $\gamma : [0, d(p, x)] \rightarrow X$  be a geodesic from  $x$  to  $p$  such that  $\gamma(0) = x$ ,  $\gamma(d(p, x)) = p$ . Set  $y = \gamma(\frac{\epsilon}{4}r)$ . Then

$$d(y, x) = \frac{\epsilon}{4}r$$

and

$$d(y, p) = d(p, x) - \frac{\epsilon}{4}r.$$

It then follows easily from the triangle inequality that

$$B(y, \frac{\epsilon}{4}r) \subseteq B \cap B(x, \frac{3\epsilon}{4}r)$$

□

**Claim 4.7.**  *$\text{dist}(z, E) \geq \frac{\epsilon}{8}r$  for every  $z \in B(y, \frac{\epsilon}{8}r)$ .*

*Proof.* By the triangle inequality and the previous claim, if  $z \in B(y, \frac{\epsilon}{8}r)$ , then we have

$$B(z, \frac{\epsilon}{8}r) \subseteq B(y, \frac{\epsilon}{4}r) \subseteq B \cap B(x, \frac{3\epsilon}{4}r).$$

The final set in this inclusion contains no points of  $E$ , so  $B(z, \frac{\epsilon}{8}r)$  contains no points of  $E$ , and this proves the claim. □

We now have

$$\begin{aligned} d_E(B) &= \frac{1}{\mu(B)} \int_B \frac{\text{dist}(z, E)}{r} d\mu(z) \geq \frac{1}{\mu(B)} \int_{B(y, \frac{\epsilon}{8}r)} \frac{\text{dist}(z, E)}{r} d\mu(z) \\ &\geq \frac{1}{\mu(B)} \int_{B(y, \frac{\epsilon}{8}r)} \frac{\epsilon r}{8r} d\mu(z) \geq \frac{\epsilon}{8} \cdot \frac{\mu(B(y, \frac{\epsilon}{8}r))}{\mu(B)}. \end{aligned}$$

If  $N$  is the smallest integer such that  $2^N \cdot \frac{\epsilon}{8} \geq 2$ , then  $B \subset B(y, 2^N \cdot \frac{\epsilon}{8}r)$ . Because  $\mu$  is  $C$ -doubling,

$$C^N \cdot \mu(B(y, \frac{\epsilon}{8}r)) \geq \mu(B(y, 2^N \cdot \frac{\epsilon}{8}r)) \geq \mu(B).$$

So we have

$$\frac{\mu(B(y, \frac{\epsilon}{8}r))}{\mu(B)} \geq \frac{1}{C^N}$$

Let  $\delta = \frac{\epsilon}{8} \cdot \frac{1}{C^N}$ , which depends only on  $\epsilon$  and  $C$ . Then

$$d_E(B) \geq \frac{\epsilon}{8} \cdot \frac{\mu(B(y, \frac{\epsilon}{8}r))}{\mu(B)} \geq \frac{\epsilon}{8} \cdot \frac{1}{C^N} = \delta.$$

□

We can now prove Theorem 1.4.

*Proof of Theorem 1.4.* Given  $\epsilon > 0$ , choose  $\delta > 0$  as in Lemma 4.5. This yields that

$$\sum_{B \in \mathcal{B}(E, \epsilon)} \mu(B) \leq \sum_{\substack{B \in \mathcal{B} \\ d_E(B) \geq \delta \\ B \cap E \neq \emptyset}} \mu(B)$$

By Proposition 4.4, we have

$$\sum_{\substack{B \in \mathcal{B} \\ d_E(B) \geq \delta \\ B \cap E \neq \emptyset}} \mu(B) \leq \sum_{\substack{B \in \mathcal{B} \\ B \cap E \neq \emptyset}} \frac{1}{\delta} d_E(B) \cdot \mu(B) \leq \frac{1}{\delta} \cdot C_X.$$

□

**Remark 4.8.** Suppose that  $(Y, d_Y, \mu)$  is a metric measure space with  $\mu$  a  $C$ -doubling measure,  $\mu(Y) = 1$ , but  $\text{diam } Y = D > 1$ . Let  $\mathcal{B}$  be a multiresolution family in  $Y$  consisting of balls of radii  $D \cdot A \cdot 2^{-k}$  for  $k \geq 0$ , and let  $E \subset Y$  be measurable. Consider the metric measure space  $(X, d_X, \mu) = (Y, \frac{1}{D}d, \mu)$ . In the space  $X$ , the balls of  $\mathcal{B}$  have radii  $A \cdot 2^{-k}$  for  $k \geq 0$ . Note in addition that the sparsity of a set in a ball is unaffected by this rescaling, and therefore the notation  $\mathcal{B}(E, \epsilon)$  is unambiguous.

Thus given  $\epsilon > 0$ , there is a  $K(\epsilon, C, A) > 0$  such that if  $E \subset Y$  then

$$\sum_{B \in \mathcal{B}(E, \epsilon)} \mu(B) \leq K(\epsilon, C, A).$$

Thus, a version of Theorem 1.4 with  $\tilde{A} = DA$  holds for metric measure spaces with diameter  $D$  greater than 1.

We now prove Corollary 1.5.

*Proof of Corollary 1.5:* Fix  $A = 1$  in the definition of our multiresolution family. By Theorem 1.4, if  $E \subset X$  then

$$\sum_{B \in \mathcal{B}(E, \epsilon)} \mu(B) \leq K$$

where  $K$  depends only on  $\epsilon$  and on the doubling measure constant  $C$ . Let  $N$  be the first integer such that some ball  $B$  in  $\mathcal{B}$  of radius  $A \cdot 2^{-N}$  touches  $E$  and is not in  $\mathcal{B}(E, \epsilon)$ , that is,  $\text{sparse}(E, B) < \epsilon$ . This implies that all balls with radii  $A \cdot 2^0, A \cdot 2^{-1}, \dots, A \cdot 2^{-(N-1)}$  that touch  $E$  are contained in  $\mathcal{B}(E, \epsilon)$ . Therefore, the above sum is bounded below by the sum of all the measures of these balls. For each  $n \in \{0, 1, \dots, N-1\}$ , the measure of the union of all such balls at a given scale is greater than or equal to the measure of  $E$ . Therefore

$$\mu(E) \cdot N \leq \sum_{B \in \mathcal{B}(E, \epsilon)} \mu(B) \leq K.$$

Note  $\mu(E) \geq \alpha$ , so we have

$$N \cdot \alpha \leq N \cdot \mu(E) \leq K$$

and so

$$N \leq \frac{K}{\alpha}.$$

Thus, there is a ball  $B$  with radius  $A \cdot 2^{-N} \geq A \cdot 2^{-\frac{K}{\alpha}}$  that intersects  $E$  and has  $\text{sparse}(E, B) < \epsilon$ . So we can take  $r = A \cdot 2^{-N}$  and  $r_0 = A \cdot 2^{-\frac{K}{\alpha}}$ .  $\square$

At this point, it may be instructive to return to the example set(s)  $E_k$  in Section 3, which showed that statement (1.1) failed, i.e., that one cannot expect a large ball where a measurable set is measure-theoretically dense.

Corollary 1.5, however, applies perfectly well to these sets. For small values of  $k$ ,  $E_k$  contains large full dyadic intervals, and thus of course has large scales in which it is metrically dense (i.e., has small sparsity). For large values of  $k$ ,  $E_k$  is metrically very dense in the whole interval  $[0, 1]$ , and so again has a large scale where it has small sparsity.

## 5. QUANTITATIVE CONNECTIVITY

In this section, we prove Theorem 1.9. This concerns subsets of  $[0, 1]^d$ , rather than general metric spaces.

If  $E \subseteq [0, 1]^d$  and  $Q$  is a dyadic cube such that  $7Q \cap E \neq \emptyset$ , we set

$$\text{sparse}(E, 7Q) = \frac{1}{\text{side}(7Q)} \sup\{\text{dist}(x, 7Q \cap E) : x \in 7Q\}.$$

We first slightly adjust Theorem 1.4 to apply to dyadic cubes.

**Corollary 5.1.** *Given  $\epsilon > 0$ , there is a  $\tilde{K} > 0$ , depending only on  $\epsilon$  and  $d$ , such that if  $E \subset [0, 1]^d$ , then*

$$\sum_{\substack{Q \in \Delta \\ \text{sparse}(E, 7Q) \geq \epsilon \\ 3Q \cap E \neq \emptyset}} \lambda(Q) \leq \tilde{K}.$$

*Proof.* Fix a multi-resolution family of balls  $\mathcal{B}$  in  $[0, 1]^d$  with scaling factor  $A = 7\sqrt{d}$ .

For each  $Q \in \Delta_k$ , we may choose a ball  $B_Q \in \mathcal{B}$  of radius  $A2^{-k} = 7\sqrt{d}2^{-k} = \text{diam}(7Q)$  such that  $7Q \subseteq B_Q$ .

Now fix a set  $E \subset [0, 1]^d$ .

**Claim 5.2.** *If  $E \cap 7Q \neq \emptyset$ , then  $\text{sparse}(E, 7Q) \leq \frac{A(d+1)}{7} \text{sparse}(E, B_Q)$ .*

*Proof.* Let  $x \in 7Q$  and  $s = \text{sparse}(E, B_Q)$ . Let  $k$  be such that  $Q \in \Delta_k$ , which forces the radius  $r$  of  $B_Q$  to be  $A2^{-k}$ .

To prove the claim, it will suffice to find a point  $y \in E \cap 7Q$  with  $|x - y| \leq (d+1)sr$ , which we now do.

If  $\text{dist}(x, \partial(7Q)) > sr$ , then this is easy: there is a  $y \in E \cap B$  such that  $d(x, y) \leq sr$ , and the assumption on the distance to the boundary forces  $y \in E \cap 7Q$ .

Otherwise,  $\text{dist}(x, \partial(7Q)) \leq sr$ . In that case, we first perturb  $x$  along  $d$  separate line segments parallel to the coordinate axes, each of length at most  $sr$ , until we reach a point  $x' \in 7Q$  with  $\text{dist}(x', \partial(7Q)) > sr$ . By the previous case, there is a  $y \in E \cap 7Q$  such that  $|x' - y| \leq sr$ . It follows that

$$|x - y| \leq sr + dsr = (d+1)sr.$$

Thus,

$$\begin{aligned} \text{sparse}(E, 7Q) &\leq \frac{(d+1)sr}{\text{side}(7Q)} = \frac{(d+1)r \text{sparse}(E, B)}{7 \cdot 2^{-k}} = \frac{(d+1)A \cdot 2^{-k} \text{sparse}(E, B)}{7 \cdot 2^{-k}} \\ &= \frac{A(d+1)}{7} \text{sparse}(E, B). \end{aligned}$$

$\square$

Now, each  $B_Q$  contains the corresponding cube  $7Q$ . In addition, each ball  $B \in \mathcal{B}$  can act as  $B_Q$  for at most  $D$  different cubes  $Q$ , for some  $D$  depending on  $d$ , by a simple volume argument.

Therefore, using the previous claim,

$$\begin{aligned} \sum_{\substack{Q \in \Delta \\ \text{sparse}(E, 7Q) \geq \epsilon \\ 3Q \cap E \neq \emptyset}} \lambda(Q) &\leq \sum_{\substack{Q \in \Delta \\ \text{sparse}(E, 7Q) \geq \epsilon \\ 3Q \cap E \neq \emptyset}} \lambda(B_Q) \\ &\leq D \left( \sum_{\substack{B \in \mathcal{B} \\ \text{sparse}(E, B) \geq \frac{7\epsilon}{A(d+1)} \\ B \cap E \neq \emptyset}} \lambda(B) \right) \end{aligned}$$

By Theorem 1.4 (and Remark 4.8), this sum is controlled by a constant depending only on  $\epsilon$  and  $d$ , which completes the proof.  $\square$

The next result uses Corollary 5.1 to show, roughly speaking, that most points of  $E$  are not contained in too many cubes of large sparsity.

**Proposition 5.3.** *Let  $\epsilon > 0$  be given, and suppose  $E \subset [0, 1]^d$ .*

*For  $N \in \mathbb{N}$ , let  $Z_N$  be the collection of points that are contained in at least  $N$  distinct cubes  $7Q$  such that  $Q \in \Delta$ ,  $3Q \cap E \neq \emptyset$ , and  $\text{sparse}(E, 7Q) \geq \epsilon$ .*

*Then  $\lambda(Z_N) \leq 7^d \frac{\tilde{K}}{N}$ , where  $\tilde{K}$  is the constant from Corollary 5.1.*

*Proof.* For  $N \in \mathbb{N}$  we have

$$Z_N = \{x \in [0, 1]^d : x \text{ is in at least } N \text{ distinct dyadic cubes } 7Q \text{ such that } 3Q \cap E \neq \emptyset \text{ and } \text{sparse}(E, 7Q) \geq \epsilon\}.$$

The set  $Z_N$  is a union of countably many cubes, and therefore Lebesgue measurable. By Corollary 5.1,

$$\begin{aligned} \tilde{K} &\geq \sum_{\substack{Q \in \Delta \\ 3Q \cap E \neq \emptyset \\ \text{sparse}(E, 7Q) \geq \epsilon}} \lambda(Q) \\ &= 7^{-d} \sum_{\substack{Q \in \Delta \\ 3Q \cap E \neq \emptyset \\ \text{sparse}(E, 7Q) \geq \epsilon}} \lambda(7Q) \\ &= 7^{-d} \int_{[0, 1]^d} \left( \sum_{\substack{Q \in \Delta \\ x \in 7Q \\ 3Q \cap E \neq \emptyset \\ \text{sparse}(E, 7Q) \geq \epsilon}} 1 \right) dx \\ &\geq 7^{-d} \int_{Z_N} \left( \sum_{\substack{Q \in \Delta \\ x \in 7Q \\ 3Q \cap E \neq \emptyset \\ \text{sparse}(E, 7Q) \geq \epsilon}} 1 \right) dx \end{aligned}$$

$$\begin{aligned}
&\geq 7^{-d} \int_{Z_N} N d\lambda \\
&= 7^{-d} N \lambda(Z_N).
\end{aligned}$$

The conclusion then follows.  $\square$

As preparation for the “coding” argument below, we need the following simple fact about dyadic cubes.

**Lemma 5.4.** *Let  $Q, R \in \Delta$  such that  $\text{side}(R) \geq \text{side}(Q)$  and  $3Q \cap 3R \neq \emptyset$ . Then  $3Q \subset 7R$ .*

*Proof.* Let  $x \in 3Q \cap 3R \neq \emptyset$ . There is a cell  $R' \subset 3R$  containing  $x$ . Then  $3Q \subset 5R' \subset 7R$ .  $\square$

The following lemma provides useful decompositions of sets for which each point is contained in only a controlled number of cubes from a given family. To our knowledge, the idea appears first (in a different context) in [8, Lemma 2.2] and is also explained in [2, Lemma 8.4].

We take the opportunity here to present the lemma in a fairly general form and also give a proof that is somewhat different than those cited above, and which we find a bit more transparent.

**Lemma 5.5.** *Let  $G \subset [0, 1]^d$ , and  $B \subset \Delta$ . Assume that  $3Q \cap G \neq \emptyset$  for every  $Q \in B$ , and that each  $x \in G$  is contained in at most  $N$  different cubes  $7Q$  for  $Q \in B$ .*

*There is a constant  $M > 0$ , depending only on  $d$  and  $N$ , and sets  $F_1, F_2, \dots, F_M$  with the following properties:*

- (i)  $G = F_1 \cup F_2 \cup \dots \cup F_M$
- (ii) *If  $x, y \in F_i$  are distinct and  $Q$  is a dyadic cube of minimal side length such that  $x, y \in 3Q$ , then  $Q \notin B$ .*

*Proof.* The proof is split into multiple claims for reading purposes.

**Claim 5.6.** *There are disjoint collections  $\mathcal{A}_1, \dots, \mathcal{A}_N \subset B$  such that the following are true:*

- (i)  $B = \mathcal{A}_1 \cup \mathcal{A}_2 \cup \dots \cup \mathcal{A}_N$ .
- (ii) *If  $Q, R \in \mathcal{A}_i$  are distinct for some  $i \in \{1, 2, \dots, N\}$ , then  $3Q \cap 3R = \emptyset$ .*

*Proof.* Order  $B = \{Q_1, Q_2, \dots\}$  by decreasing size, i.e., so that  $\text{side}(Q_i) \geq \text{side}(Q_j)$  if  $i \leq j$ . We will construct the collections  $\mathcal{A}_1, \dots, \mathcal{A}_N$  inductively. Put  $Q_1 \in \mathcal{A}_1$ .

Assume now that  $Q_1, \dots, Q_k$  have been put in various collections among the  $\mathcal{A}_1, \dots, \mathcal{A}_N$ , that the collections are (so far) disjoint from each other, and that if  $Q, R \in \mathcal{A}_i$  then  $3Q \cap 3R = \emptyset$ . We wish to place  $Q_{k+1}$  in some  $\mathcal{A}_i$  with the property that  $3Q_{k+1} \cap 3R = \emptyset$  for all  $R$  currently in  $\mathcal{A}_i$ .

Suppose there is no  $i \in \{1, 2, \dots, N\}$  such that  $3Q_{k+1} \cap 3R = \emptyset$  for every  $R \in \mathcal{A}_i \cap \{Q_1, \dots, Q_k\}$ . Then  $3Q_{k+1}$  intersects distinct cubes  $3R_1, 3R_2, \dots, 3R_N$ , where  $R_i \in \mathcal{A}_i$  and  $\text{side}(R_i) \geq \text{side}(Q_{k+1})$ . By Lemma 5.4,  $3Q_{k+1} \subset 7R_i$  for every  $i \in \{1, 2, \dots, N\}$ . Then there is an  $x \in 3Q_{k+1} \cap G$  such that  $x \in 7Q_{k+1}, 7R_1, \dots, 7R_N$ . This contradicts the assumption that  $x$  can be contained in at most  $N$  different cubes  $7Q$  for  $Q \in B$ .

Therefore, we may place cube  $Q_{k+1}$  in one of the collections  $\mathcal{A}_i$  and maintain the desired properties. It follows by induction that the sets may be constructed as desired.  $\square$

Now let  $\alpha = \{0, 1, \dots, 3^d\}$ . We introduce a correspondence between the elements of  $\alpha \setminus \{0\}$  and the  $3^d$  different cells of a (hence any) tripled dyadic cube  $3Q$ . The precise choice of correspondence does not matter at all; to be concrete, we can order the cells by the “dictionary” order of their centers, so that cell 1 is the cell closest to the origin, cell 2 is its neighbor in the  $x_1$  direction, etc.

We will consider “words  $w$  of length  $N$  from the alphabet  $\alpha$ ,” i.e.,  $w \in \alpha^N$ . For such a  $w$ ,  $w_i \in \alpha$  denotes its  $i$ th letter.

For each  $x \in G$  we define  $\omega(x) \in \alpha^N$  in the following way. Consider  $i \in \{0, \dots, 3^d\}$ . The point  $x$  can be in at most one cube  $3Q$  for  $Q \in \mathcal{A}_i$ . Let us call this cube  $Q_{x,i}$ , if it exists. We set the  $i$ th letter of  $w(x)$  to be

$$\omega(x)_i = \begin{cases} 0 & \text{if } x \notin \bigcup_{Q \in \mathcal{A}_i} 3Q \\ k \in \alpha \setminus \{0\} & \text{if } x \text{ is in the } k\text{th cell of } 3Q_{x,i}. \end{cases}$$

Thus, for each  $x$ , there are at most  $M = (3^d + 1)^N$  possibilities for the word  $\omega(x)$ . Enumerate these possibilities in any order by  $\alpha^N = \{\omega_1, \dots, \omega_M\}$ . Let  $F_j = \{x \in G : \omega(x) = \omega_j\}$  for  $j \in \{1, 2, \dots, M\}$ . It is immediate that  $G = \cup_{j=1}^M F_j$ , so it remains to verify the second conclusion of the lemma.

**Claim 5.7.** *Suppose  $x, y \in F_j$  and  $x \neq y$ . Let  $Q$  be a dyadic cube of minimal side length such that  $x, y \in 3Q$ . Then  $Q \notin B$ .*

*Proof.* Suppose  $Q \in B$ . Then  $Q \in \mathcal{A}_i$  for some  $i \in \{1, 2, \dots, N\}$ . Since  $x, y \in F_j$ ,  $\omega(x) = \omega(y) = \omega_j$ . So  $\omega(x)_i = \omega(y)_i$ . It follows from the definition of  $\omega(x)_i$  and  $\omega(y)_i$  that  $x$  and  $y$  both lie in the same cell  $Q'$  of  $3Q$ .

Let  $R$  be a child of  $Q'$  that contains  $x$ . Then  $x, y \in Q' \subset 3R$ . Since  $R$  is half the size of  $Q$ , this contradicts the fact that  $Q$  is a minimal dyadic cube such that  $x, y \in 3Q$ . Therefore  $Q \notin B$ .  $\square$

Both conclusions of the lemma have now been verified.  $\square$

We will also need the following basic fact.

**Lemma 5.8.** *Let  $Q \in \Delta$  be a minimal dyadic cube such that  $x, y \in 3Q$ . Then  $\text{side}(Q) \leq 2|x - y|$ .*

*Proof.* Let  $Q'$  be a cell of  $3Q$  such that  $x \in Q'$ . Let  $R \subset Q'$  be a child of  $Q'$  such that  $x \in R$ . Because  $Q$  is a minimal cube such that  $x, y \in 3Q$ , it must be that  $y \notin 3R$ . So

$$\frac{1}{2}\text{side}(Q) = \text{side}(R) \leq |x - y|.$$

This implies that  $\text{side}(Q) \leq 2|x - y|$ .  $\square$

We are now ready to prove Theorem 1.9.

*Proof of Theorem 1.9.* Let  $\alpha, \delta > 0$ . Without loss of generality we may assume  $\delta < \frac{1}{2}$ . Let  $P \in \mathbb{N}$  be chosen large enough so that  $\frac{1}{P} \leq \frac{3}{4}\delta$  and  $P > 1 + \frac{1}{\delta}$ . Let  $\epsilon = \frac{\delta}{28P}$  and set  $B_0 = \{Q \in \Delta : 3Q \cap E \neq \emptyset, \text{sparse}(E, 7Q) \geq \epsilon\}$ . For  $N \in \mathbb{N}$  to be chosen momentarily, define  $Z$  by

$$Z = \{x \in E : x \text{ is in at least } N > 0 \text{ different cubes } 7Q \text{ for } Q \in B_0\}.$$

Then  $Z = Z_N \cap E$ , where  $Z_N$  is the (measurable) set defined in Proposition 5.3. Using that proposition, we have that

$$\lambda(Z) \leq 7^d \frac{\tilde{K}}{N}$$

(with the understanding that this denotes Lebesgue outer measure if  $E$  is not measurable).

For  $N$  large enough, depending only on  $\epsilon$  and  $d$ , we have  $\lambda(Z) < \alpha$ . Let  $G = E \setminus Z$ , and  $B = \{Q \in B_0 : 3Q \cap G \neq \emptyset\}$ . Let  $x \in G$ . Then  $x \notin Z$ , so  $x$  is in at most  $N$  different cubes of  $B_0$  and hence in at most  $N$  different cubes of  $B$ . By Lemma 5.5, we can write  $G = F_1 \cup F_2 \cup \dots \cup F_M$  with the property that if  $x, y \in F_i$  then a minimal dyadic cube  $Q$  with  $x, y \in 3Q$  satisfies  $Q \notin B$ . Here  $M = (3^d + 1)^N$ .

At this point, we have written

$$E = G \cup Z = F_1 \cup \dots \cup F_M \cup Z,$$

where  $\lambda(Z) < \alpha$ . To complete the proof, it remains to check that each set  $F_j$  is  $\delta$ -well-connected in  $E$ .

Suppose  $x, y \in F_j$  are distinct for some  $j \in \{1, 2, \dots, M\}$ . Let  $Q$  be a minimal dyadic cube such that  $x, y \in 3Q$ . Then  $Q \notin B$ . Since  $x, y \in G \cap 3Q$ , we have  $G \cap 3Q \neq \emptyset$ , from which it follows that  $Q \notin B_0$ . This implies that  $\text{sparse}(E, 7Q) < \epsilon$ . Note also that  $\text{side}(Q) \leq 2|x - y|$  by the previous lemma.

Let  $\gamma : [0, |x - y|] \rightarrow 3Q$  parametrize the line segment from  $x$  to  $y$ . Let  $x_i = \gamma(\frac{i}{P}|x - y|)$  for  $i \in \{0, 1, 2, \dots, P\}$ . Note that  $x_0 = x$  and  $x_P = y$ . For each  $i \in \{1, 2, \dots, P - 1\}$ , the fact that  $\text{sparse}(E, 7Q) < \epsilon$  allows us to choose  $z_i \in E$  such that  $|z_i - x_i| < \epsilon \cdot \text{side}(7Q)$ . We also set  $z_0 = x$  and  $z_P = y$ , which are in  $F_j \subseteq G \subseteq E$  by assumption. Thus, we have a discrete path  $\{z_0 = x, z_1, z_2, \dots, z_{P-1}, z_P = y\} \subset E$ . Figure 2 illustrates the construction of this path.

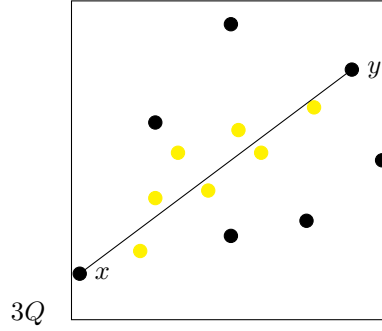


FIGURE 2. An illustration of the chain  $\{z_i\}$ , highlighted in yellow, built by choosing points of  $E$  near the line segment from  $x$  to  $y$ .

We now verify that this path has the desired properties to make  $F_j$   $\delta$ -well-connected. First, we check, using our choices of  $\epsilon$  and  $P$ , that each “step size” is small.

$$\begin{aligned}
 |z_{i+1} - z_i| &\leq |z_{i+1} - x_{i+1}| + |x_{i+1} - x_i| + |x_i - z_i| \\
 &< 2\epsilon \text{side}(7Q) + \frac{1}{P}|x - y| \\
 (5.1) \quad &\leq 28\epsilon|x - y| + \frac{1}{P}|x - y| \\
 &\leq \left(\frac{\delta + 1}{P}\right)|x - y| \\
 &\leq \delta|x - y|.
 \end{aligned}$$

Next, we check that the total length of this path is small. Reusing inequality (5.1) from the previous calculation and our definitions of  $\epsilon$  and  $P$ , we obtain that

$$\begin{aligned}
 \sum_{i=0}^{P-1} |z_{i+1} - z_i| &\leq \sum_{i=0}^{P-1} \left(28\epsilon + \frac{1}{P}\right)|x - y| \\
 &= P \left(28\epsilon + \frac{1}{P}\right)|x - y| \\
 &\leq (\delta + 1)|x - y|
 \end{aligned}$$

Thus, each  $F_j$  is  $\delta$ -well-connected in  $E$ , and this completes the argument.  $\square$

## REFERENCES

- [1] Michael Christ. “A  $T(b)$  theorem with remarks on analytic capacity and the Cauchy integral”. In: *Colloq. Math.* 60/61.2 (1990), pp. 601–628. ISSN: 0010-1354,1730-6302. DOI: 10.4064/cm-60-61-2-601-628. URL: <https://doi.org/10.4064/cm-60-61-2-601-628>.
- [2] Guy David. *Wavelets and Singular Integrals on Curves and Surfaces*. Lecture Notes in Mathematics. Springer-Verlag, 1991. ISBN: ISBN 3-540-53902-6.
- [3] Guy David and Stephen Semmes. *Analysis of and on Uniformly Rectifiable Sets*. Mathematical Surveys and Monographs. American Mathematical Society, 1991.
- [4] José R. Dorronsoro. “A characterization of potential spaces”. In: *Proc. Amer. Math. Soc.* 95.1 (1985), pp. 21–31. ISSN: 0002-9939. DOI: 10.2307/2045567.

- [5] Misha Gromov. *Metric structures for Riemannian and non-Riemannian spaces*. English. Modern Birkhäuser Classics. Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates. Birkhäuser Boston, Inc., Boston, MA, 2007, pp. xx+585. ISBN: 978-0-8176-4582-3; 0-8176-4582-9.
- [6] Juha Heinonen. *Lectures on Analysis on Metric Spaces*. Universitext. Springer-Verlag, New York, 2001, pp. 1–6. ISBN: 0-387-95104-0. DOI: 10.1007/978-1-4613-0131-8. URL: <https://doi.org/10.1007/978-1-4613-0131-8>
- [7] Tuomas Hytönen and Anna Kairema. “Systems of dyadic cubes in a doubling metric space”. In: *Colloq. Math.* 126.1 (2012), pp. 1–33. ISSN: 0010-1354,1730-6302. DOI: 10.4064/cm126-1-1. URL: <https://doi.org/10.4064/cm126-1-1>
- [8] Peter W. Jones. “Lipschitz and bi-Lipschitz Functions”. In: *Revista Matemática Iberoamericana* 4.1 (1988), pp. 115–121.
- [9] Peter W. Jones. “Rectifiable sets and the traveling salesman problem”. In: *Invent. Math.* 102.1 (1990), pp. 1–15. ISSN: 0020-9910,1432-1297. DOI: 10.1007/BF01233418. URL: <https://doi.org/10.1007/BF01233418>.
- [10] Peter W. Jones, Nets Hawk Katz, and Ana Vargas. “Checkerboards, Lipschitz functions and uniform rectifiability”. In: *Revista Matemática Iberoamericana* 13.1 (1997), pp. 189–210. ISSN: 0213-2230. DOI: 10.4171/RMI/219. URL: <https://doi.org/10.4171/RMI/219>.
- [11] P. Mattila. *Geometry of sets and measures in Euclidean spaces*. Vol. 44. Cambridge Studies in Advanced Mathematics. Fractals and rectifiability. Cambridge University Press, Cambridge, 1995, pp. xii+343. ISBN: 0-521-46576-1; 0-521-65595-1. DOI: 10.1017/CB09780511623813. URL: <https://doi-org.proxy.bsu.edu/10.1017/CB09780511623813>
- [12] Raanan Schul. *Big-Pieces-of-Lipschitz-Images Implies a Sufficient Carleson Estimate in a Metric Space*. 2007. arXiv: 0706.2517 [math.MG].
- [13] Robert Young. *Notes on quantitative rectifiability and differentiability*. <https://math.nyu.edu/~ryoung/courses/qdiff/diffNotes.pdf>. Apr. 2020.

DEPARTMENT OF MATHEMATICAL SCIENCES, BALL STATE UNIVERSITY, MUNCIE, IN 47306  
 Email address: `gcdavid@bsu.edu`

DEPARTMENT OF MATHEMATICAL SCIENCES, BALL STATE UNIVERSITY, MUNCIE, IN 47306  
 Email address: `brandon.oliva@bsu.edu`