

On Falconer's distance set problem in the plane

**Larry Guth, Alex Iosevich, Yumeng Ou
& Hong Wang**

Inventiones mathematicae

ISSN 0020-9910

Invent. math.

DOI 10.1007/s00222-019-00917-x



Your article is protected by copyright and all rights are held exclusively by Springer-Verlag GmbH Germany, part of Springer Nature. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".



On Falconer's distance set problem in the plane

Larry Guth¹ · Alex Iosevich² ·
Yumeng Ou³ · Hong Wang¹

Received: 18 September 2018 / Accepted: 18 August 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract If $E \subset \mathbb{R}^2$ is a compact set of Hausdorff dimension greater than $5/4$, we prove that there is a point $x \in E$ so that the set of distances $\{|x - y|\}_{y \in E}$ has positive Lebesgue measure.

1 Introduction

For a set $E \subset \mathbb{R}^d$, define the distance set

$$\Delta(E) = \{|p - p'| : p, p' \in E\}.$$

Falconer's distance problem [11] is about the connection between the Hausdorff dimension of a set E and the size of $\Delta(E)$. Given a compact set E in \mathbb{R}^d , $d \geq 2$, the problem is to understand how large the Hausdorff dimension of E needs to be to ensure that the Lebesgue measure of $\Delta(E)$ is positive. Falconer proved that if $\dim_{\mathcal{H}}(E) > \frac{d+1}{2}$, then $\mathcal{L}(\Delta(E)) > 0$. Using an example based on the integer lattice, he showed for every $s \leq \frac{d}{2}$ there exist sets of Haus-

✉ Yumeng Ou
Yumeng.Ou@baruch.cuny.edu

¹ Department of Mathematics, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

² Department of Mathematics, University of Rochester, 915 Hylan Building, Rochester, NY 14627, USA

³ Department of Mathematics, City University of New York, Baruch College, One Bernard Baruch Way, New York, NY 10010, USA

dorff dimension s for which $\mathcal{L}(\Delta(E)) = 0$. This led to the conjecture that if $\dim_{\mathcal{H}}(E) > \frac{d}{2}$, then the Lebesgue measure of the distance set is positive. This is known as the Falconer Distance Conjecture.

In [36], Wolff proved that if $E \subset \mathbb{R}^2$ is a compact set with Hausdorff dimension greater than $4/3$, then $\Delta(E)$ has positive Lebesgue measure. In this paper, we improve the bound.

Theorem 1.1 *If $E \subset \mathbb{R}^2$ is a compact set with Hausdorff dimension greater than $5/4$, then $\Delta(E)$ has positive Lebesgue measure.*

In higher dimensions, Erdoğan proved in [9] that if $\dim_{\mathcal{H}}(E) > \frac{d}{2} + \frac{1}{3}$, then $\mathcal{L}(\Delta(E)) > 0$. Recently, these estimates were improved for all $d \geq 3$ by Du, Guth, Ou, Wang, Wilson, and Zhang [6]. In dimension 3, they showed that the Falconer conjecture holds when $\dim_{\mathcal{H}}(E) > 9/5$. The estimates for $d \geq 4$ were further improved by Du and Zhang [7]. For large d , they prove that Falconer's conjecture holds when $\dim_{\mathcal{H}}(E) > \frac{d}{2} + \frac{1}{4} + o(1)$. These works brought into play the decoupling theorem of Bourgain and Demeter [3]. This approach will also play a key role in our proof.

Returning to the planar case, there have been a number of important recent results. Orponen [29] proved that if E is a compact Ahlfors-David regular set of dimension $s \geq 1$, then $\Delta(E)$ has packing dimension 1. Note that packing dimension 1 is only slightly weaker than positive measure. This result was striking because in previous work on the problem, there was no evidence that the Ahlfors-David case would be any easier than the general case. This approach was further developed by Keleti and Shmerkin [24]. They proved very strong estimates for sets that are even roughly like Ahlfors-David regular sets. They also proved results about the Hausdorff dimension of $\Delta(E)$. For instance, if E is a compact set with Hausdorff dimension strictly greater than 1, then they proved that the Hausdorff dimension of $\Delta(E)$ is at least .685.... Bourgain [2] had proven that if E has Hausdorff dimension at least 1, then $\Delta(E)$ has Hausdorff dimension at least $1/2 + \delta$ for some $\delta > 0$. The value of δ could be made explicit but it would be very small, and so the .685... is quite striking. We will use one of the key ideas of [29] and [24] in the proof of Theorem 1.1.

There is a variant of the Falconer Distance Problem involving pinned distance sets. For any point x , the pinned distance set $\Delta_x(E)$ is defined by

$$\Delta_x(E) = \{|x - y| : y \in E\}.$$

Peres and Schlag [31] proved that if $E \subset \mathbb{R}^d$, $d \geq 2$ and $\dim_{\mathcal{H}}(E) > \frac{d+1}{2}$, then $\mathcal{L}(\Delta_x(E)) > 0$ for every $x \in E$ except for a set of small Hausdorff dimension. Improvements on the size of the exceptional set were obtained by the second listed author and Liu in [19].

Recently, in [25], Liu showed that if $\dim_{\mathcal{H}}(E) > \frac{d}{2} + \frac{1}{3}$, then $\mathcal{L}(\Delta_x(E)) > 0$ for every $x \in E$ except those in a set of small Hausdorff dimension. Using Liu's method, we are also able to bound the size of pinned distance sets.

Theorem 1.2 *If $E \subset \mathbb{R}^2$ is a compact set with Hausdorff dimension larger than $\frac{5}{4}$, then there is a point $x \in E$ such that its pinned distance set $\Delta_x(E)$ has positive Lebesgue measure.*

1.1 Other norms

The Falconer problem has also been studied for other norms. Suppose that K is a symmetric convex body in \mathbb{R}^d and $\|\cdot\|_K$ is the norm with unit ball K . We let $\Delta_K(E)$ be the set of distances $\|x - y\|_K$ with $x, y \in E$ and we let $\Delta_{K,x}(E)$ be the set of distances $\|x - y\|_K$ with $y \in E$. If K is the cube $[-1, 1]^d$, then $\|\cdot\|_K$ is the l^∞ norm, and it is not difficult to construct a compact set $E \subset \mathbb{R}^d$ with Hausdorff dimension d so that $\Delta_K(E)$ has measure zero. But there are non-trivial results if K is curved. We focus on the case that ∂K is C^∞ smooth and has positive Gaussian curvature. It is plausible that Falconer's conjecture remains true for all such norms, and most previous results on the problem extend to this setting. For instance, Erdoğan's bound extends to this class of norms—cf. Remark 1.6 in [9]. Our method also extends to this class of norms.

Theorem 1.3 *Let K be a symmetric convex body in \mathbb{R}^2 whose boundary ∂K is C^∞ smooth and has strictly positive curvature. Let $E \subset \mathbb{R}^2$ be a compact set whose Hausdorff dimension is larger than $\frac{5}{4}$. Then, there exists a point $x \in E$ so that the pinned distance set*

$$\Delta_{K,x}(E) := \{\|x - y\|_K : y \in E\}$$

has positive Lebesgue measure.

Remark 1.4 One can adapt the proof of Theorems 1.1 and 1.3 to yield the following result. Suppose that the Hausdorff dimension of a compact set $E \subset \mathbb{R}^2$ is equal to $s > 1$ and K is as in Theorem 1.3. Then there exists $x \in E$ such that the upper Minkowski dimension of $\Delta_{x,K}(E)$ is $\geq \frac{4s}{3} - \frac{2}{3}$. Keleti and Shmerkin [24] obtained the lower bound $\frac{1}{4}(1 + s + \sqrt{3s(2-s)})$ in the case of the Euclidean metric. Their estimate is better than ours near $s = 1$, but ours is preferable as s nears $\frac{5}{4}$. The sketch of this argument is given in Appendix where we also discuss the complications of replacing the upper Minkowski dimension by the Hausdorff dimension in the claim above.

Falconer's distance problem can be thought of as a continuous analogue of a combinatorial problem raised by Erdős [10]: given a set P of N points

in \mathbb{R}^d , what is the smallest possible cardinality of $\Delta(P)$. A grid is the best known example in all dimensions. In two dimensions, Guth and Katz [14] proved a lower bound for $|\Delta(P)|$ which nearly matches the grid example (up to a factor of $\log^{1/2} N$). In higher dimensions, there is a larger gap, and the best known result is due to Solymosi and Vu [34]. The Erdős distinct distance problem also makes sense for general norms and much less is known about it. In the planar case, if K is smooth and has strictly positive curvature, the best known bound says that if $|P| = N$, then $|\Delta_K(P)| \gtrsim N^{3/4}$, with stronger estimates established by Garibaldi in special cases [12]. There is a conversion mechanism to go from Falconer-type results to Erdős-type results that was developed by the second author together with Hoffman [16], Laba [18], and Rudnev and Uriarte-Tuero [17]. It gives estimates for point sets that are fairly spread out. Applying the conversion mechanism to Theorem 1.3 we get the following corollary:

Corollary 1.5 *Let K be a symmetric convex body in \mathbb{R}^2 whose boundary ∂K is C^∞ smooth and has strictly positive curvature. Let P be a set of N points in $[0, 1]^2$ so that the distance between any two points is $\gtrsim N^{-1/2}$. Then there exists $x \in P$ such that*

$$|\Delta_{K,x}(P)| \gtrsim N^{\frac{4}{5}}. \quad (1.1)$$

Note that the $N^{-1/2}$ -separation condition in the above essentially says that P is a homogeneous set, i.e. there is one point in each of the $\sim N$ cells in $[0, 1]^2$.

1.2 The main obstacle

The work on the Falconer problem by Wolff [36] and Erdoğan [9] is based on a framework developed by Mattila [26, 27] which connects the original geometric problem to estimates in Fourier analysis. Suppose that E is a compact set with positive α -dimensional Hausdorff measure. Then there is a probability measure μ supported on E with $\mu(B(x, r)) \lesssim r^\alpha$ for every ball $B(x, r)$. The measure μ is called a Frostman measure (cf. [37], Proposition 8.2.). Let $d(x, y) = |x - y|$. Mattila considered the pushforward measure $d_*(\mu \times \mu)$, defined by

$$\int_{\mathbb{R}} \psi(t) d_*(\mu \times \mu) := \int_{E \times E} \psi(|x - y|) d\mu(x) d\mu(y).$$

In particular $d_*(\mu \times \mu)$ is a probability measure supported on $\Delta(E)$. Mattila noted that if $\|d_*(\mu \times \mu)\|_{L^2}^2 = \int d_*(\mu \times \mu)(t)^2 dt$ is finite, then Cauchy–Schwarz forces the Lebesgue measure of $\Delta(E)$ to be positive. Then he described an interesting way to rewrite $\|d_*(\mu \times \mu)\|_{L^2}^2$ in terms of the Fourier

transform of μ . The resulting integral is connected to restriction theory, and Wolff used that connection to prove the bound in [36], building on earlier work by Bourgain [1].

In [25], Liu used a different framework for the Falconer problem which leads to estimates on pinned distance sets. For any x , define $d^x(y) = |x - y|$. He studied the quantity

$$\int \|d_*^x(\mu)\|_{L^2}^2 d\mu(x). \quad (1.2)$$

If this key quantity is finite, then for almost every $x \in E$, $\|d_*^x\mu\|_{L^2}$ is finite, and then a Cauchy-Schwarz argument forces the Lebesgue measure of $\Delta_x(E)$ to be positive. Liu introduced an interesting way to rewrite this quantity in terms of the Fourier transform of μ . It can then be studied using restriction theory, leading to estimates on the pinned distance problem.

In the planar case, there is an obstruction to pushing either one of these methods to dimensions below $4/3$. For every $\alpha < 4/3$, there is a set E of dimension α and a Frostman measure μ on E so that $\|d_*(\mu \times \mu)\|_{L^2}$ is infinite, and also $\|d_*^x(\mu)\|_{L^2}$ is infinite for every $x \in E$. (In three dimensions, it is known [9, Remark 1.5] that $5/3$ is the lowest threshold one can possibly obtain for the Falconer problem using the Mattila integral method, however, we are not aware of similar obstructions in higher dimensions.) This set is a variation on an example from [22]. The set E looks roughly like several parallel train tracks. In the following figure, we show an approximation of the set E at a small scale R^{-1} . The measure μ (approximated at scale R^{-1}) is just the normalized area measure on this set.

The set E is divided among several large $R^{-1/2} \times 1$ rectangles. Within each of these large rectangles, the set E consists of evenly spaced parallel rectangles with dimensions $R^{-1/2} \times R^{-1}$. Each of these smaller rectangles is called a slat. The restriction of E to one of the larger rectangles is called a train track. The spacing between two consecutive slats is controlled by the dimension of E , and it works out to $R^{-\alpha/2}$. If x and y are in the same train track, on roughly opposite ends, and if y is M slats from x , then $|x - y|$ lies in the interval

$$I_M := [MR^{-\alpha/2} - R^{-1}, MR^{-\alpha/2} + R^{-1}].$$

Therefore, $d_*(\mu \times \mu)$ assigns a lot of mass to the union of the intervals I_M . This union is quite small, and even though the mass involved is significantly less than 1, it is still enough to force $\int |d_*(\mu \times \mu)|^2$ to be very large.

There is a similar issue for $d_*^x(\mu)$. If we fix any $x \in E$, and we let T_0 be the large rectangle containing x , then $d_*^x(\mu|_{T_0})$ is mostly concentrated on $\cup I_M$, and this forces $\int |d_*^x\mu|^2$ to be very large. On the other hand, if T is a large rectangle which is far from x , then $d_*^x(\mu|_T)$ is rather evenly distributed—in fact $d_*^x(\mu|_T)$ is close to the pushforward of the uniform measure on T with the

same total mass. So if we graph $d_*^x(\mu)$, it has some peaks along $\cup I_M$ coming from the rectangle T_0 through x , but the bulk of $d_*^x(\mu)$ is spread rather evenly and comes from rectangles T far from x . In particular, the support of $d_*^x(\mu)$ indeed has positive Lebesgue measure.

This example (which will be revisited in more detail in Sect. 6) is the main obstacle to proving the Falconer conjecture for dimensions less than $4/3$. Starting with a general Frostman measure, we separate out a part of it that resembles the train tracks in the example above. Then we estimate the train-track part and the non-train-track part in different ways.

For technical reasons, we consider two subsets $E_1, E_2 \subset E$ separated by distance ~ 1 , and we let μ_1 and μ_2 be Frostman measures on E_1, E_2 . In fact, we will show that for some $x \in E_2$, the distance set $\Delta_x(E_1)$ has positive Lebesgue measure. In the example above, we can imagine that E_1 is the bottom third and E_2 is the top third. We divide μ_1 into two pieces

$$\mu_1 = \mu_{1,good} + \mu_{1,bad},$$

where $\mu_{1,bad}$ is essentially the train-track-like part of μ_1 . We always arrange, however, that $\int \mu_{1,bad} = 0$.

For example, if μ_1 is the normalized area measure on the set E_1 in Fig. 1 above, then $\mu_{1,good}$ would be (approximately) the normalized area measure on the union of the large rectangles. The bad part, $\mu_{1,bad}$, is equal to $\mu_1 - \mu_{1,good}$, so it would be large on the slats and slightly negative on the parts of the large rectangles outside of the slats. If T_0 is the large rectangle containing x , then $d_*^x(\mu_{1,good}|_{T_0})$ would be much more spread out than $d_*^x(\mu_1|_{T_0})$. On the other hand, if T is far from x , then $d_*^x(\mu_{1,good}|_T)$ would be almost the same as $d_*^x(\mu_1|_T)$. All together, the graph of $d_*^x(\mu_{1,good})$ would look like the graph of $d_*^x(\mu_1)$ with the peaks damped out. The pushforward $d_*^x(\mu_{1,good})$ would be quite evenly spread and its L^2 norm would be finite. The graph of $d_*^x(\mu_{1,bad})$ would include the tall thin peaks from $d_*^x(\mu_1)$, and it would be slightly negative between the peaks. Since the thin peaks have small mass, the L^1 norm of $d_*^x(\mu_{1,bad})$ would be small.

To prove Theorem 1.2, we will show that the features of $\mu_{1,good}$ and $\mu_{1,bad}$ that we just observed in the example from Fig. 1 will occur for any set E of dimension greater than $5/4$. There are two main estimates. The first estimate, in Proposition 2.1, says that for most $x \in E_2$, $\|d_*^x(\mu_{1,bad})\|_{L^1}$ is small, and so the L^1 distance between $d_*^x(\mu_1)$ and $d_*^x(\mu_{1,good})$ is small. The bad part, $\mu_{1,bad}$, is made from train track configurations, and that helps us analyze it. Analyzing each individual train track is not difficult. However, unlike in our example above, it could happen that each point lies in many different train tracks going in different directions. To control this type of behavior, we use an estimate of Orponen from [30] which also

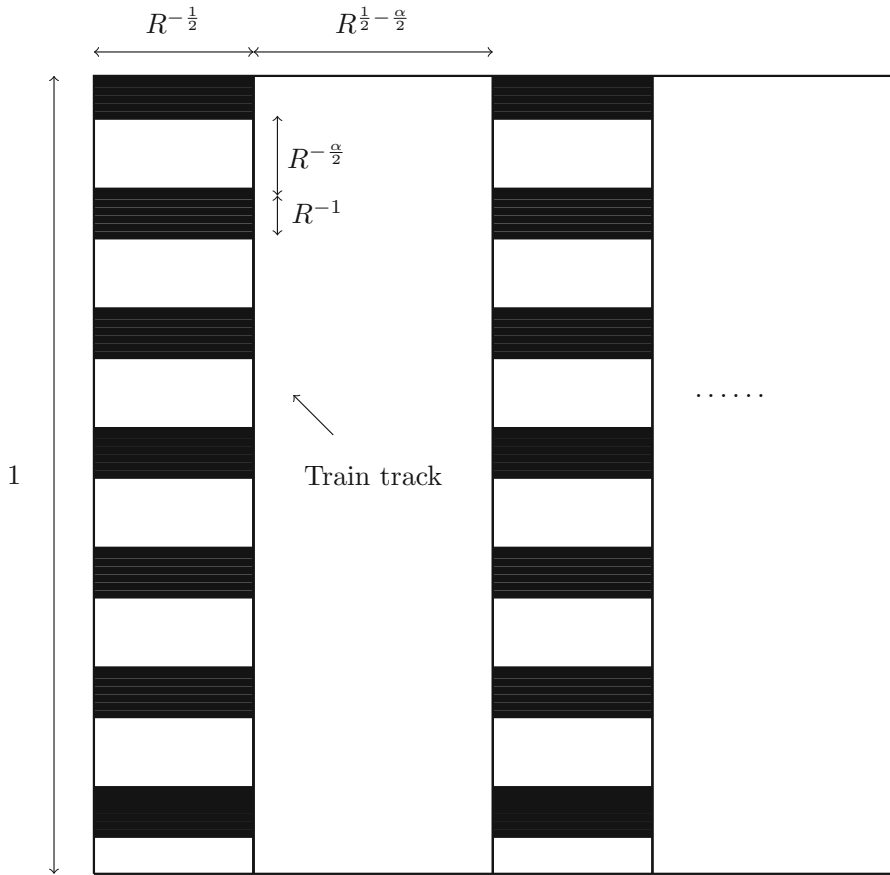


Fig. 1 The train track example

played a key role in Keleti and Shmerkin's work on the Falconer problem [24].

The second estimate says that $d_*^x \mu_{1,good}$ is better behaved in L^2 than $d_*^x \mu_1$. More precisely, Proposition 2.2 says that if $\alpha > 5/4$, then $\int_{E_2} \|d_*^x \mu_{1,good}\|_{L^2}^2$ is finite. The proof of Proposition 2.2 is based on Liu's framework and on decoupling. We will prove and then use a refinement of the decoupling theorem (Theorem 4.2) which is related to the refined Strichartz estimates that appear in [4, 5], and [6]. This refinement of decoupling was proven independently by Xiumin Du and Ruixiang Zhang (personal communication). It may be of independent interest.

Here is an outline of the paper. In Sect. 2, we set up our framework (defining $\mu_{1,good}$ precisely) and outline the main estimates. At that point, we will be able to make some further comments about the proofs of the two main propositions. In Sect. 3, we prove Proposition 2.1. In Sect. 4, we state and prove a refinement

of the decoupling theorem. At that point, we will give some more context about this result. Section 4 does not depend on any previous sections. In Sect. 5, we prove Proposition 2.2 by combining Liu's framework with our decoupling tools. This will finish the proof of Theorem 1.2. In Sect. 6, we present in detail the train track example that we introduced above. In Sect. 7, we adapt our arguments to general metrics $\|\cdot\|_K$, proving Theorem 1.3. In Sect. 8, we prove Corollary 1.5.

2 Setup and outline of the main estimates

Let $E \subset \mathbb{R}^2$ be a compact set with positive α -dimensional Hausdorff measure. Without loss of generality, we can suppose that E is contained in the unit disk. Let E_1 and E_2 be subsets of E with positive α -dimensional Hausdorff measure so that the distance from E_1 to E_2 is $\gtrsim 1$. Each subset E_i admits a measure μ_i with the following two properties:

$$\mu_i \text{ is a probability measure supported on } E_i. \quad (2.1)$$

$$\mu_i(B(x, r)) \lesssim r^\alpha. \quad (2.2)$$

We will explain how to define $\mu_{1,good}$ by removing “train-track like” pieces from μ_1 . Before going into the details, let us explain the features of a train track that motivate our definition of $\mu_{1,good}$. Let μ be the example in Fig. 1 and let T be one of the $R^{-1/2} \times 1$ rectangles containing a train track of the set E . One feature of T is that $\mu(T)$ is large. Because the slats of the train track are perpendicular to the direction of T , the Fourier transform of $\mu|_T$ is concentrated on frequencies that are in the same direction as T . This is a second feature of T . So to build $\mu_{1,good}$, we will first identify rectangles T with large μ measure and call them bad rectangles. Then for each bad rectangle T , we will identify the part of μ with physical support in T and frequency support in the direction of T , and remove that part. Here is the precise definition.

We consider a sequence of scales R_0, R_1, R_2 , etc. Here R_0 is a large number that we will choose later and $R_j = 2^j R_0$. Cover the annulus $R_{j-1} \leq |\omega| \leq R_j$ by rectangular blocks τ with dimensions approximately $R_j^{1/2} \times R_j$. The long direction of each block τ is the radial direction. We choose a partition of unity subordinate to this cover, so that

$$1 = \psi_0 + \sum_{j \geq 1, \tau} \psi_{j, \tau}.$$

Let $\delta > 0$ be a small constant that will be determined later. Note that δ will play a role in the definition of $\mu_{1,good}$ and will be eventually taken sufficiently small in terms of α in Proposition 2.2 below.

For each (j, τ) , cover the unit disk with tubes T of dimensions approximately $R_j^{-1/2+\delta} \times 2$ with the long axis parallel to the long axis of τ . The covering has uniformly bounded overlap, each T intersects at most two other tubes. Let $\mathbb{T}_{j,\tau}$ be the collection of all these tubes, and let η_T be a partition of unity subordinate to this covering, so that for each choice of j and τ , $\sum_{T \in \mathbb{T}_{j,\tau}} \eta_T$ is equal to 1 on the disk of radius 2 and each η_T is smooth.

Define an operator M_T associated to a tube $T \in \mathbb{T}_{j,\tau}$ by

$$M_T f := \eta_T(\psi_{j,\tau} \hat{f})^\vee.$$

Morally, $M_T f$ is the part of f which has Fourier support in τ and physical support in T . We also let $M_0 f := (\psi_0 \hat{f})^\vee$. We denote $\mathbb{T}_j = \cup_\tau \mathbb{T}_{j,\tau}$ and $\mathbb{T} = \cup_{j \geq 1} \mathbb{T}_j$. If f is a function supported on the unit disk, then $f = M_0 f + \sum_{T \in \mathbb{T}} M_T f$, up to a tiny error (see Lemma 3.4 below for a precise statement).

Let $c(\alpha) > 0$ be a large constant to be determined later in Lemma 3.6, and $4T$ denote the concentric tube of four times the radius. We call a tube $T \in \mathbb{T}_{j,\tau}$ *bad* if

$$\mu_2(4T) \geq R_j^{-1/2+c(\alpha)\delta}.$$

To get a sense of what this means, notice that the number of tubes $T \in \mathbb{T}_{j,\tau}$ is $\sim R_j^{1/2-\delta}$. If each tube $T \in \mathbb{T}_{j,\tau}$ contained the same amount of the measure μ_2 , then for each tube we would have $\mu_2(T) \sim R_j^{-1/2+\delta}$. A tube is *bad*, if its neighborhood contains significantly more μ_2 measure than this. A tube is *good* if it is not bad. Now we define $\mu_{1,\text{good}}$ to be the sum of contributions from all the good tubes.

$$\mu_{1,\text{good}} := M_0 \mu_1 + \sum_{T \in \mathbb{T}, T \text{ good}} M_T \mu_1.$$

Note that in our discussion below, only those tubes that intersect both E_1 and E_2 matter, as otherwise they don't make any contribution to the quantities in Propositions 2.1 and 2.2. And according to the definition, $\mu_{1,\text{good}}$ is defined with respect to μ_2 .

We describe a couple of examples to give a sense of how $\mu_{1,\text{good}}$ behaves. If μ_1 is the normalized area measure on the set E in Fig. 1 above, then $\mu_{1,\text{good}}$ would be (approximately) the normalized area measure on the union of the large rectangles. On the other hand, if we took the set E in Fig. 1 above and we changed it by tilting the slats at a 45 degree angle while keeping the large rectangles vertical, then $\mu_{1,\text{good}}$ would be essentially equal to μ_1 . In general, $\mu_{1,\text{good}}$ may not be real-valued, but it is a distribution (in fact, it is a complex measure supported on the disk of radius 2).

Our main theorem (Theorem 1.2) follows from two estimates about the pushforward measures $d_*^x \mu_1$ and $d_*^x \mu_{1,good}$.

Proposition 2.1 *If $\alpha > 1$, and if we choose R_0 large enough, then there is a subset $E'_2 \subset E_2$ so that $\mu_2(E'_2) \geq 1 - \frac{1}{1000}$ and for each $x \in E'_2$,*

$$\|d_*^x(\mu_1) - d_*^x(\mu_{1,good})\|_{L^1} < \frac{1}{1000}.$$

Proposition 2.2 *If $\alpha > 5/4$, then for sufficiently small δ in terms of α ,*

$$\int_{E_2} \|d_*^x(\mu_{1,good})\|_{L^2}^2 d\mu_2(x) < +\infty.$$

In the above, we have slightly abused notation, using $d_*^x(\mu)$, $d_*^x(\mu_{1,good})$ to denote both the pushforward measures and their densities. To be completely rigorous, one would need to define the density as limit of approximate identity, then derive the propositions above uniformly with respect to the limiting process. We omit the details as the process is fairly standard (for example see [25]).

Proof of Theorem 1.2 using Proposition 2.1 and Proposition 2.2 The two propositions tell us that there is a point $x \in E_2$ so that

$$\|d_*^x(\mu_1) - d_*^x(\mu_{1,good})\|_{L^1} < 1/1000, \text{ and} \quad (2.3)$$

$$\|d_*^x(\mu_{1,good})\|_{L^2} < +\infty. \quad (2.4)$$

Since $d_*^x(\mu_1)$ is a probability measure, (2.3) guarantees that

$$\int |d_*^x(\mu_{1,good})| \geq 1 - \frac{1}{1000}.$$

Note that the support of $d_*^x(\mu_1)$ is contained in $\Delta_x(E)$. Therefore

$$\begin{aligned} \int_{\Delta_x(E)} |d_*^x \mu_{1,good}| &= \int |d_*^x(\mu_{1,good})| - \int_{\Delta_x(E)^c} |d_*^x(\mu_{1,good})| \\ &\geq 1 - \frac{1}{1000} - \int |d_*^x(\mu_1) - d_*^x(\mu_{1,good})| \geq 1 - \frac{2}{1000}. \end{aligned}$$

But on the other hand,

$$\int_{\Delta_x(E)} |d_*^x \mu_{1,good}| \leq |\Delta_x(E)|^{1/2} \left(\int |d_*^x \mu_{1,good}|^2 \right)^{1/2}. \quad (2.5)$$

Since (2.4) tells us that $\int |d_*^x \mu_{1,good}|^2$ is finite, it follows that $|\Delta_x(E)|$ is positive. \square

To end this section, let us make some comments about the proofs of Proposition 2.1 and Proposition 2.2. To prove Proposition 2.1, the first observation is that if x is far from T , then removing $M_T \mu_1$ from μ_1 has a negligible effect on the pushforward measure $d_*^x(\mu_1)$. So the difference between $d_*^x \mu_1$ and $d_*^x \mu_{1,good}$ only comes from the bad tubes going through x . Recall that a tube T is bad if its μ_2 measure is too large. In general a point x could lie in many bad tubes, and we need to know that the total μ_1 measure of all these bad tubes is small (for most $x \in E_2$). This follows from Orponen's radial projection theorem from [30]. This theorem plays an important role in Keleti and Shmerkin's work on the Falconer problem [24], which is where we learned about it.

To discuss Proposition 2.2, we first describe the framework from [25]. Let σ_t denote the normalized arc length measure on the circle of radius t (so the total measure is 1). In [25], Liu proved the following remarkable identity: for any function f ,

$$\int_0^\infty |f * \sigma_t(x)|^2 t dt = \int_0^\infty |f * \hat{\sigma}_r(x)|^2 r dr.$$

It follows from this identity that

$$\int_{E_2} \|d_*^x(\mu_{1,good})\|_{L^2}^2 d\mu_2(x) \lesssim \int_0^\infty \left(\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r|^2 d\mu_2(x) \right) r dr.$$

Now the Fourier transform of $\mu_{1,good} * \hat{\sigma}_r$ is supported on the circle of radius r , and studying such functions is the subject of restriction theory. We can decompose $\mu_{1,good} * \hat{\sigma}_r$ as

$$\mu_{1,good} * \hat{\sigma}_r = \sum_{T \text{ good}} M_T \mu_1 * \hat{\sigma}_r.$$

The right-hand side is essentially the wave packet decomposition of $\mu_{1,good} * \hat{\sigma}_r$. This means that $M_T \mu_1 * \hat{\sigma}_r|_{B^2(1)}$ is essentially supported in T and its Fourier transform is essentially supported in an arc of S_r^1 in the direction of T . Since the tubes T are all good, each tube T has a small μ_2 measure, and we will take advantage of this to bound the inner integral $\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r|^2 d\mu_2(x)$.

Since T has a small μ_2 measure, we can immediately get a good estimate for $\int_{E_2} |M_T \mu_1 * \hat{\sigma}_r|^2 d\mu_2$. But to bound $\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r|^2 d\mu_2$, we need to know how the wave packets $M_T \mu_1 * \hat{\sigma}_r$ interact with each other. Is it possible that these wave packets have a lot of positive interference on the set E_2 ? We will

use decoupling theory to control such positive interference. We will discuss this further in Sect. 4.

2.1 Notations and preliminaries

We write $A \lesssim B$ if $A \leq CB$ for some absolute constant C ; $A \sim B$ if $A \lesssim B$ and $B \lesssim A$; $A \lesssim_\epsilon B$ if $A \leq C_\epsilon B$; $A \lesssim_R B$ if $A \leq C_\epsilon R^\epsilon B$ for any $\epsilon > 0$, $R > 1$ (or with N replacing R depending on the context).

For a large parameter R , we use $\text{RapDec}(R)$ to denote those quantities that are bounded by a huge (absolute) negative power of R , i.e. $\text{RapDec}(R) \leq C_N R^{-N}$ for arbitrary $N > 0$. Such quantities are negligible in our argument. We say a function is *essentially supported* in a region if (the appropriate norm of) the tail outside the region is $\text{RapDec}(R)$ for the underlying parameter R .

A typical situation where such terms would appear is the following. Recall that we have defined at the beginning of this section that $M_T f = \eta_T(\psi_{j,\tau} \hat{f})^\vee$, for a tube $T \in \mathbb{T}_{j,\tau}$ that has radius $R_j^{-1/2+\delta}$ and length 2. Hence, $M_T f$ is supported in T because of η_T . In the Fourier space, write $\widehat{M_T f} = \widehat{\eta_T} * (\psi_{j,\tau} \hat{f})$, we claim that $M_T f$ has Fourier support essentially contained in 2τ , where τ has dimension approximately $R_j^{1/2} \times R_j$ and has long axis parallel to the long axis of T . To see this, note that η_T is a smooth bump function supported in T that has height 1, and it has Fourier transform

$$\widehat{\eta_T}(\xi) = \int_T \eta_T(x) e^{ix \cdot \xi} dx = e^{ic(T) \cdot \xi} \int_T \eta_T(x) e^{i(x-c(T)) \cdot \xi} dx,$$

where $c(T)$ denotes the center of T . If ξ is in the dual rectangle T^* of T , i.e. the rectangle centered at the origin with dimension $\sim R_j^{1/2-\delta} \times \frac{1}{2}$ with short axis parallel to the long axis of T , then $|(x-c(T)) \cdot \xi| < \frac{1}{10}$. Hence, there is essentially no cancellation and $|\widehat{\eta_T}(\xi)| \sim |T|$. If instead, ξ is outside a R_j^δ neighborhood of T^* , cancellation plays the primary role and from the smoothness of η_T and integration by parts one has $|\widehat{\eta_T}(\xi)| \lesssim \text{RapDec}(R_j)$. In conclusion, $\widehat{\eta_T}$ is essentially supported in T^* which makes $\widehat{M_T f} = \widehat{\eta_T} * (\psi_{j,\tau} \hat{f})$ essentially supported in 2τ .

3 Proof of Proposition 2.1

We will study the pushforward measures $d_*^X(\mu_1)$ and $d_*^X(\mu_{1,\text{good}})$. Recall by definition that

$$\int \psi(t) d_*^X(\mu) = \int \psi(|x-y|) d\mu(y).$$

In particular, if ψ is the characteristic function of the interval $t_0 \leq t \leq t_0 + \Delta t$, then we see that

$$\int_{t_0}^{t_0+\Delta t} d_*^x(\mu) = \int_{t_0 \leq |x-y| \leq t_0+\Delta t} d\mu.$$

To evaluate $d_*^x(\mu)$ at t , we can take the limit as $\Delta t \rightarrow 0$. If we think of μ as $\mu(y)dy$, then we get

$$d_*^x(\mu)(t) = \int_{S^1(x,t)} \mu(y)dl(y),$$

where $dl(y)$ denotes the arc length measure on the circle $S^1(x, t)$.

To control $\|d_*^x(\mu_1) - d_*^x(\mu_{1,good})\|_{L^1}$ we will start by studying $d_*^x(M_T\mu_1)$ for different T . For a tube $T \in \mathbb{T}$, let $2T$ denote the concentric tube of twice the radius. If $x \notin 2T$, we show that $d_*^x(M_T\mu_1)$ is negligible.

Lemma 3.1 *If $T \in \mathbb{T}_{j,\tau}$, and $x \in E_2$, and $x \notin 2T$, then*

$$\|d_*^x(M_T\mu_1)\|_{L^1} \lesssim \text{RapDec}(R_j).$$

Proof We will prove the stronger estimate that for every t :

$$d_*^x(M_T\mu_1)(t) \lesssim \text{RapDec}(R_j).$$

Recall that

$$d_*^x(M_T\mu_1)(t) = \int_{S^1(x,t)} M_T\mu_1(y)dl(y). \quad (3.1)$$

We also recall that

$$M_T\mu_1 = \eta_T(\psi_{j,\tau}\hat{\mu}_1)^\vee = \eta_T(\psi_{j,\tau}^\vee * \mu_1).$$

Now $\psi_{j,\tau}^\vee$ is concentrated on a $R_j^{-1/2} \times R_j^{-1}$ rectangle centered at 0 and it decays rapidly outside that rectangle. Since $x \in E_2$, the distance from x to the support of μ_1 is $\gtrsim 1$. Therefore, $d_*^x M_T\mu_1(t)$ is tiny unless $t \sim 1$.

To study the case when $t \sim 1$, we expand out $M_T\mu_1$:

$$M_T\mu_1(y) = \eta_T(y)(\psi_{j,\tau}\hat{\mu}_1)^\vee(y) = \eta_T(y) \int e^{2\pi i \omega \cdot y} \psi_{j,\tau}(\omega) \hat{\mu}_1(\omega) d\omega.$$

Since $|\hat{\mu}_1(\omega)| \leq 1$, and $\psi_{j,\tau}(\omega)$ is supported on τ and bounded by 1, it suffices to check that for each $\omega \in \tau$,

$$\int_{S^1(x,t)} \eta_T(y) e^{2\pi i \omega \cdot y} dl(y) \leq \text{RapDec}(R_j). \quad (3.2)$$

We will prove this rapid decay by stationary phase. There are two slightly different cases, depending on whether T intersects $S^1(x, t/2)$ or not. Let us start with the case that T intersects $S^1(x, t/2)$, since this case is a little harder. After a coordinate rotation, we can assume that ω has the form $(0, \omega_2)$ with $\omega_2 \sim R_j$. Recall that a tube $T \in \mathbb{T}_{j,\tau}$ has long axis in the direction of the center of τ . In particular, our tube T must be nearly vertical, up to an angle of $R_j^{-1/2}$. The tube T intersects $S^1(x, t)$ in two arcs, which we deal with one at a time. Each arc is a graph of the form $y_2 = h(y_1)$, where y_1 lies in an interval $I(T)$ of length $\sim R_j^{-1/2+\delta}$. Since T intersects $S^1(x, t/2)$, and the tube T is nearly vertical, the function h and all its derivatives are $\lesssim 1$ on $I(T)$.

The following point is crucial for stationary phase. Since T is within an angle $R_j^{-1/2}$ of vertical, and $x \notin 2T$, then the distance from T to the top or bottom points of the circle is $\gtrsim R_j^{-1/2+\delta}$, and so we get

$$|h'(y_1)| \gtrsim R_j^{-1/2+\delta} \text{ on the interval } I(T).$$

In these coordinates, our integral becomes

$$\int_{I(T)} \eta_T(y_1, h(y_1)) e^{2\pi i \omega_2 h(y_1)} J(y_1) dy_1,$$

where $J(y_1)$ is the Jacobian factor that relates the arclength on the circle to dy_1 . Since h and all its derivatives are $\lesssim 1$ on $I(T)$, the same applies to J . The function η_T is smooth at scale $R_j^{-1/2+\delta}$, and so if we abbreviate $\eta(y_1) := \eta_T(y_1, h(y_1))J(y_1)$, then η obeys

$$|\eta^{(k)}| \lesssim (R_j^{1/2-\delta})^k,$$

and η is supported on $I(T)$. We let $\phi(y_1) = 2\pi\omega_2 h(y_1)$. We now have to bound the following integral:

$$\int_{I(T)} \eta(y_1) e^{i\phi(y_1)} dy_1.$$

This integral can be bounded using stationary phase. The method is essentially the same as in [35], Chapter 8, Proposition 1. Here is a sketch. We note that on $I(T)$,

$$\begin{aligned} |\phi'(y_1)| &= |\omega_2| |h'(y_1)| \gtrsim R_j^{1/2+\delta}, \text{ and} \\ |\phi^{(k)}(y_1)| &= |\omega_2| |h^{(k)}(y_1)| \lesssim R_j. \end{aligned}$$

Next we note that

$$\frac{1}{i\phi'} \frac{d}{dy_1} e^{i\phi(y_1)} = e^{i\phi(y_1)}.$$

We define $D = \frac{1}{i\phi'} \frac{d}{dy_1}$, so our integral becomes $\int_{I(T)} \eta D^N e^{i\phi} dy_1$, where N is an arbitrary integer. Next we expand out $D^N e^{i\phi}$ and we integrate by parts many times so that none of the derivatives actually lands on $e^{i\phi}$. Using our lower bound on $|\phi'|$ and our upper bounds on the higher derivatives of ϕ and the derivatives of η , it follows that our integral is bounded by $C_N R_j^{-2\delta N}$. For instance, if all the derivatives land on η , then we get a bound on $(R_j^{1/2-\delta})^N (R_j^{1/2+\delta})^{-N} \sim R_j^{-2\delta N}$, and this is the worst case. Since N is arbitrary we get the desired bound.

If T does not intersect $S^1(x, t/2)$, then we choose our coordinates differently so that we can still arrange that $|h'|$ is bounded. This time, we rotate so that $\omega = (\omega_1, 0)$, where $\omega_1 \sim R_j$. The tube T intersects $S^1(x, t)$ in one or two arcs, and each arc is a graph of the form $y_2 = h(y_1)$ over an interval $I(T)$, and h and all its derivatives are $\lesssim 1$ on $I(T)$. Our integral now has the form

$$\int_{I(T)} \eta_T(y_1, h(y_1)) J(y_1) e^{2\pi i \omega_1 y_1} dy_1.$$

Since $\omega_1 \sim R_j$, and η_T is smooth on the scale $R_j^{-1/2+\delta}$, this integral can also be bounded by stationary phase (in fact more simply than in the other case). \square

Next we prove a simple bound to cover the case that $x \in 2T$.

Lemma 3.2 *For any $T \in \mathbb{T}_{j,\tau}$ and any function f supported in the unit disk,*

$$\|M_T f\|_{L^1} \lesssim \|f\|_{L^1(2T)} + \text{RapDec}(R_j) \|f\|_{L^1}.$$

Proof Recall that for a tube $T \in \mathbb{T}_{j,\tau}$, we defined M_T by

$$M_T f := \eta_T(\psi_{j,\tau} \hat{f})^\vee = \eta_T(\psi_{j,\tau}^\vee * f).$$

Now $\psi_{j,\tau}^\vee$ is essentially supported in a rectangle of dimensions $R_j^{-1} \times R_j^{-1/2}$ and $\|\psi_{j,\tau}^\vee\|_{L^1} \lesssim 1$. Since the thickness of T is $R_j^{-1/2+\delta}$, we get

$$\int |M_T f| \lesssim \int_T |\psi_{j,\tau}^\vee * f| \lesssim \int_{2T} |f| + \text{RapDec}(R_j) \|f\|_{L^1}.$$

\square

Corollary 3.3 For any point x , and any tube $T \in \mathbb{T}$,

$$\|d_*^x(M_T \mu_1)\|_{L^1} \lesssim \mu_1(2T) + \text{RapDec}(R_j).$$

Next we check carefully that μ_1 is very close to $M_0 \mu_1 + \sum_{T \in \mathbb{T}} M_T \mu_1$.

Lemma 3.4 For any L^1 function f supported in the unit disk

$$\|f - M_0 f - \sum_{T \in \mathbb{T}} M_T f\|_{L^1} \lesssim \text{RapDec}(R_0) \|f\|_{L^1}.$$

Proof Recall that $\{\psi_{j,\tau}\}$ is a partition of unity. We define $M_{j,\tau} f = (\psi_{j,\tau} \hat{f})^\vee$, so that $f = \sum_{j,\tau} M_{j,\tau} f$. It suffices to bound

$$\|M_{j,\tau} f - \sum_{T \in \mathbb{T}_{j,\tau}} M_T f\|_{L^1} \lesssim \text{RapDec}(R_j) \|f\|_{L^1}.$$

The left hand side is

$$\|(1 - \sum_{T \in \mathbb{T}_{j,\tau}} \eta_T)(\psi_{j,\tau}^\vee * f)\|_{L^1}.$$

Now as we noted in the proof of Lemma 3.2, $\psi_{j,\tau}^\vee$ is essentially supported on an $R_j^{-1/2} \times R_j^{-1}$ rectangle. Also, $\sum_{T \in \mathbb{T}_{j,\tau}} \eta_T$ is equal to one on the disk of radius 2 and then decays outside it. Since f is supported in the unit disk, $\psi_{j,\tau}^\vee * f$ is essentially supported in the disk of radius 2, and we get the desired rapid decay. \square

Now we can relate $\|d_*^x(\mu_{1,\text{good}}) - d_*^x(\mu_1)\|_{L^1}$ to the geometry of the bad rectangles. For each point x and each j , we define

$$\text{Bad}_j(x) := \bigcup_{T \in \mathbb{T}_j : x \in 2T \text{ and } T \text{ is bad}} 2T.$$

Lemma 3.5 For any point x in E_2 ,

$$\|d_*^x(\mu_{1,\text{good}}) - d_*^x(\mu_1)\|_{L^1} \lesssim \sum_{j \geq 1} R_j^\delta \mu_1(\text{Bad}_j(x)) + \text{RapDec}(R_0).$$

Proof Recall that $\mu_{1,\text{good}}$ is defined by

$$\mu_{1,\text{good}} := M_0 \mu_1 + \sum_{T \in \mathbb{T}, T \text{ good}} M_T \mu_1.$$

Using Lemma 3.4, we see that

$$\|d_*^x(\mu_{1,good}) - d_*^x(\mu_1)\|_{L^1} \lesssim \sum_j \sum_{T \in \mathbb{T}_j, T \text{ bad}} \|d_*^x(M_T \mu_1)\|_{L^1} + \text{RapDec}(R_0).$$

If $x \in 2T$, then we apply Corollary 3.3, and if $x \notin 2T$, then we apply Lemma 3.1. We get

$$\|d_*^x(\mu_{1,good}) - d_*^x(\mu_1)\|_{L^1} \lesssim \sum_j \sum_{T \in \mathbb{T}_j, x \in 2T, T \text{ bad}} \mu_1(2T) + \text{RapDec}(R_0).$$

Since the distance from E_2 to E_1 is $\gtrsim 1$, each point of E_1 is contained in $2T$ for $\lesssim R_j^\delta$ tubes $T \in \mathbb{T}_j$ with $x \in 2T$. Therefore, the right hand side is

$$\lesssim \sum_j R_j^\delta \mu_1 \left(\bigcup_{T \in \mathbb{T}_j, x \in 2T, T \text{ bad}} 2T \right) + \text{RapDec}(R_0) = \sum_j R_j^\delta \mu_1(\text{Bad}_j(x)) + \text{RapDec}(R_0). \quad \square$$

Next we need to estimate the measure of $\text{Bad}_j(x)$. We will do this using Orponen's radial projection theorem. Before introducing the theorem, we need to set up a little more notation.

$\text{Bad}_j := \{(x_1, x_2) : \text{there is a bad } T \in \mathbb{T}_j \text{ so that } 2T \text{ contains } x_1 \text{ and } x_2\}.$

Notice that $\text{Bad}_j(x)$ is just the set of y so that $(y, x) \in \text{Bad}_j$. Therefore,

$$\mu_1 \times \mu_2(\text{Bad}_j) = \int \mu_1(\text{Bad}_j(x)) d\mu_2(x).$$

Our main estimate about the bad rectangles is

Lemma 3.6 *For each $\alpha > 1$, there exists sufficiently large $c(\alpha) > 0$ (that appears in the definition of bad tubes), so that for each $j \geq 1$,*

$$\mu_1 \times \mu_2(\text{Bad}_j) \lesssim R_j^{-2\delta}.$$

Before turning to the proof, let us use this lemma to finish the proof of Proposition 2.1.

Proof of Proposition 2.1 using Lemma 3.6 We want to find a set $E'_2 \subset E_2$ with $\mu_2(E'_2) \geq 1 - \frac{1}{1000}$ so that for each $x \in E'_2$,

$$\|d_*^x(\mu_{1,good}) - d_*^x(\mu_1)\|_{L^1} \leq \frac{1}{1000}.$$

We recall that

$$\mu_1 \times \mu_2(\text{Bad}_j) = \int \mu_1(\text{Bad}_j(x)) d\mu_2(x).$$

Therefore, we can choose $B_j \subset E_2$ so that $\mu_2(B_j) \leq R_j^{-(1/2)\delta}$ and for all $x \in E_2 \setminus B_j$,

$$\mu_1(\text{Bad}_j(x)) \lesssim R_j^{-(3/2)\delta}.$$

We define $E'_2 = E_2 \setminus \bigcup_{j \geq 1} B_j$. As long as R_0 is sufficiently large (compared to δ and α), we have $\mu_2(E'_2) \geq 1 - \frac{1}{1000}$ as desired. Now for each $x \in E'_2$, we have

$$\begin{aligned} \|d_*^x(\mu_{1,good}) - d_*^x(\mu_1)\|_{L^1} &\lesssim \sum_{j \geq 1} R_j^\delta \mu_1(\text{Bad}_j(x)) + \text{RapDec}(R_0) \\ &\lesssim \sum_{j \geq 1} R_j^{-(1/2)\delta} + \text{RapDec}(R_0) \lesssim R_0^{-(1/2)\delta}. \end{aligned}$$

By choosing R_0 sufficiently large, we get the desired bound. \square

Now we introduce Orponen's radial projection theorem. The statement we use appears as Proposition 3.11 in [24], and it appears as Equation (3.5) in Orponen's paper [30]. Define a radial projection map $P_y : \mathbb{R}^2 \setminus \{y\} \rightarrow S^1$ by

$$P_y(x) = \frac{x - y}{|x - y|}.$$

Theorem 3.7 (Orponen [30]) *For every $\alpha > 1$ there exists $p(\alpha) > 1$ so that the following holds. Suppose that μ_1 and μ_2 are measures on the unit disk with disjoint supports and that for every ball $B(x, r)$, $\mu_i(B(x, r)) \lesssim r^\alpha$. Then*

$$\int \|P_y \mu_2\|_{L^p}^p d\mu_1(y) < +\infty.$$

Proof of Lemma 3.6 Recall that $\text{Bad}_j(y)$ is defined to be

$$\text{Bad}_j(y) := \bigcup_{T \in \mathbb{T}_j: y \in 2T \text{ and } T \text{ is bad}} 2T.$$

In other words, $\text{Bad}_j(y)$ is the set of x so that (y, x) lies in Bad_j . Therefore,

$$\mu_1 \times \mu_2(\text{Bad}_j) = \int \mu_2(\text{Bad}_j(y)) d\mu_1(y).$$

Suppose that $T \in \mathbb{T}_j$ is a bad rectangle and $y \in 2T \cap E_1$. Let $A(T)$ be the arc of the circle whose center corresponds to the direction of the long axis of T and with length $\sim R_j^{-1/2+\delta}$. Since the distance from E_1 to E_2 is $\gtrsim 1$, it follows that $P_y(4T \cap E_2) \subset A(T)$, and so

$$P_y \mu_2(A(T)) \geq \mu_2(4T) \geq R_j^{-1/2+c(\alpha)\delta}. \quad (3.3)$$

So we see that $P_y(\text{Bad}_j(y))$ can be covered by arcs $A(T)$ of length $\sim R_j^{-1/2+\delta}$ which each enjoy (3.3). By the Vitali covering lemma, we can choose a disjoint subset of the arcs $A(T)$ so that $5A(T)$ covers $P_y(\text{Bad}_j(y))$. This implies that the arc length measure of $P_y(\text{Bad}_j(y))$ is bounded by

$$|P_y(\text{Bad}_j(y))| \lesssim R_j^{-(c(\alpha)-1)\delta}.$$

Now we bound

$$\begin{aligned} \mu_1 \times \mu_2(\text{Bad}_j) &= \int \mu_2(\text{Bad}_j(y)) d\mu_1(y) \\ &\leq \int \left(\int_{P_y(\text{Bad}_j(y))} P_y \mu_2 \right) d\mu_1(y). \end{aligned}$$

By Holder's inequality and by choosing $c(\alpha)$ sufficiently large, this is

$$\leq \sup_y |P_y(\text{Bad}_j(y))|^{1-\frac{1}{p}} \int \|P_y \mu_2\|_{L^p} d\mu_1 \lesssim R_j^{-2\delta}.$$

□

4 Refined Strichartz estimates

The proof of Proposition 2.2 will use a refined Strichartz type estimate, which in turn is based on the decoupling theorem of Bourgain-Demeter [3].

Theorem 4.1 ([3]) *Suppose that $S \subset \mathbb{R}^d$ is a strictly convex C^2 hypersurface with Gaussian curvature ~ 1 . Decompose the R^{-1} -neighborhood of S into blocks θ of dimensions $R^{-1/2} \times \dots \times R^{-1/2} \times R^{-1}$. Suppose that \hat{f}_θ is supported in θ and $f = \sum_\theta f_\theta$. Then for any p in the range $2 \leq p \leq \frac{2(d+1)}{d-1}$,*

$$\|f\|_{L^p(B_R)} \lesssim \left(\sum_\theta \|f_\theta\|_{L^p(w_{B_R})}^2 \right)^{1/2}, \quad (4.1)$$

where w_{B_R} is a weight which is ~ 1 on B_R and rapidly decaying.

The decoupling theorem is a remarkably strong and sharp theorem in some situations, for instance if $|f_\theta(x)|$ is roughly constant on B_R for each θ . On the other hand, if the supports of the different f_θ are disjoint from each other, then one trivially gets the stronger inequality $\|f\|_{L^p(B_R)} \leq (\sum_\theta \|f_\theta\|_{L^p(B_R)}^p)^{1/p}$. The idea of refined Strichartz estimates is to use the decoupling theorem where it is strong, but also to take advantage of disjointness when it occurs. The first version of the refined Strichartz inequality appeared in [4], and it was generalized in [5]. We need here a slightly more flexible version of the inequality. The inequality we prove here was discovered independently by Xiumin Du and Ruixiang Zhang (personal communication).

We will state our estimate in terms of wave packets. Here is the setup. Let S and θ be as above. Let \mathbb{T}_θ be a finitely overlapping covering of \mathbb{R}^d by tubes T of length $\sim R^{1+\delta}$ and radius $\sim R^{\frac{1+\delta}{2}}$ with long axis normal to the surface S at θ . We write $\mathbb{T} = \cup_\theta \mathbb{T}_\theta$. Each $T \in \mathbb{T}$ belongs to \mathbb{T}_θ for a single θ , and we let $\theta(T)$ denote this θ . We say that f is microlocalized to $(T, \theta(T))$ if f is essentially supported in $2T$ and \hat{f} is essentially supported in $2\theta(T)$. A function f_T which is microlocalized to $(T, \theta(T))$ is called a wave packet. If $\omega \in \theta(T)$, then f_T morally has the form $f_T \approx a \chi_T e^{2\pi i \omega x}$, where $a \in \mathbb{C}$ and χ_T denotes a smooth bump function on T . Our theorem gives an estimate for the constructive interference between wave packets.

Theorem 4.2 *Let p be in the range $2 \leq p \leq \frac{2(d+1)}{d-1}$. For any $\epsilon > 0$, suppose there exists $0 < \delta \ll \epsilon$ satisfying the following. Let $\mathbb{W} \subset \mathbb{T}$ and suppose that each $T \in \mathbb{W}$ lies in B_R . Let $W = |\mathbb{W}|$. Suppose that $f = \sum_{T \in \mathbb{W}} f_T$, where f_T is microlocalized to $(T, \theta(T))$. Suppose that $\|f_T\|_{L^p}$ is roughly constant among all the $T \in \mathbb{W}$. Let Y be a union of $R^{1/2}$ -cubes in B_R each of which intersects at most M tubes $T \in \mathbb{W}$. Then*

$$\|f\|_{L^p(Y)} \lesssim_\epsilon R^\epsilon \left(\frac{M}{W} \right)^{\frac{1}{2} - \frac{1}{p}} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2 \right)^{1/2}. \quad (4.2)$$

The fraction M/W measures to what extent the wave packets of \mathbb{W} are disjoint from each other. If $M = 1$, then the wave packets are completely disjoint, and the inequality above becomes $\|f\|_{L^p(Y)} \lesssim (\sum_T \|f_T\|_{L^p}^p)^{1/p}$.

Before proving Theorem 4.2, let us explain how it relates to the decoupling theorem (Theorem 4.1). In Theorem 4.1, consider the special case that f_θ is non-zero for N caps θ , and that for each of these caps, $f_\theta = \sum_{T \in \mathbb{T}_\theta} f_T$ is a sum of P non-zero wave packets f_T , and that all these wave packets have the same amplitude. In [3], the general theorem was reduced to this special case by pigeonholing, so it is not really so special. In this case, the decoupling inequality (4.1) can be written in the form

$$\|f\|_{L^p(B_R)} \lesssim \left(\frac{1}{P}\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2\right)^{1/2}. \quad (4.3)$$

Now if Q is any $R^{1/2}$ -square, then it can lie in $\lesssim 1$ tubes T in each direction. Therefore, we have $M \leq N$, and $W = NP$, and so $\frac{M}{W} \leq \frac{1}{P}$. So we see that (4.2) is at least as strong as (4.3), and it is stronger whenever M is much less than N . When M is much less than N , then it means that each cube Q lies in wave packets from only a small fraction of the different caps θ , which means that the supports of the f_θ don't intersect as much as they could. In summary, Theorem 4.2 is like Theorem 4.1, but it gives a stronger estimate when the supports of the f_θ don't intersect too much.

Proof of Theorem 4.2 Without loss of generality, we can assume that

$$\|f\|_{L^p(Q)} \sim \text{constant for all } R^{1/2}\text{-cubes } Q \subset Y. \quad (4.4)$$

To set up the argument, we decompose f as follows. We cover S with larger blocks τ of dimensions $R^{-1/4} \times \dots \times R^{-1/4} \times R^{-1/2}$. For each τ we cover $B^d(R)$ with cylinders \square with radius $R^{3/4}$ and length R , with the long axis perpendicular to τ . Each cylinder \square is associated to one τ , which denote $\tau(\square)$. Then we define

$$\mathbb{W}_\square := \{T \in \mathbb{W} : \theta(T) \subset \tau(\square) \text{ and } T \cap B_R \subset \square\}.$$

We define $f_\square = \sum_{T \in \mathbb{W}_\square} f_T$. We note that \hat{f}_\square is essentially supported in $\tau(\square)$. An $R^{1/2}$ -cube Q lies in one cylinder \square associated to each cap τ . So by applying decoupling at scale $R^{1/2}$, we get

$$\|f\|_{L^p(Q)} \lesssim \left(\sum_{\square} \|f_\square\|_{L^p(Q)}^2\right)^{1/2}. \quad (4.5)$$

(Strictly speaking, we have a weight on the right-hand side. However, if the tail of the weight dominates for some $Q \subset Y$, then we trivially get the conclusion of the theorem. Indeed, because the weight decays rapidly away from Q , the total contribution of the tails from different Q 's is bounded by R^{-N_0} . By choosing N_0 sufficiently large, this is trivially bounded by the right hand side of (4.2). Therefore, we can ignore the tail of the weight.)

The next ingredient is induction on scales. After parabolic rescaling, the decomposition $f_\square = \sum_{T \in \mathbb{W}_\square} f_T$ is equivalent to the setup of the theorem at scale $R^{1/2}$ instead of scale R . So by induction on the radius, we get a version of our main inequality for each function f_\square . It goes as follows:

Write \square as a union of $R^{1/2} \times R^{3/4}$ cylinders running parallel to the long axis of \square . Let $Y_{\square, M'}$ be the union of those cylinders that intersect $\sim M'$ of the tubes $T \in \mathbb{W}_\square$. Then

$$\|f_\square\|_{L^p(Y_{\square, M'})} \lesssim R^{\epsilon/2} \left(\frac{M'}{|\mathbb{W}_\square|} \right)^{\frac{1}{2} - \frac{1}{p}} \left(\sum_{T \in \mathbb{W}_\square} \|f_T\|_{L^p}^2 \right)^{1/2}. \quad (4.6)$$

Indeed, by shrinking the short side of \square by $R^{1/4}$ and the long side of it by $R^{1/2}$, one can turn \square into a cube of side length $R^{1/2}$. Under the same parabolic rescaling, the small cylinders inside \square will become small cubes of side length $R^{1/4}$ and the rescaled tubes T contained in \square will have radius $R^{1/4+\delta}$ and length $R^{1/2}$.

Now we dyadically pigeonhole M' so that

$$\|f\|_{L^p(Q)} \lesssim \left\| \sum_{\square: Q \subset Y_{\square, M'}} f_\square \right\|_{L^p(Q)}$$

for a fraction ≈ 1 of $Q \subset Y$.

We fix this value of M' , and from now on we abbreviate $Y_\square = Y_{\square, M'}$.

Next we dyadically pigeonhole $|\mathbb{W}_\square|$. Let $\mathbb{B}_{W'}$ be the set of \square with $|\mathbb{W}_\square| \sim W'$. We dyadically pigeonhole W' so that

$$\|f\|_{L^p(Q)} \lesssim \left\| \sum_{\square \in \mathbb{B}_{W'}: Q \subset Y_\square} f_\square \right\|_{L^p(Q)}. \quad (4.7)$$

for a fraction ≈ 1 of $Q \subset Y$.

We fix this value of W' and from now on we abbreviate $\mathbb{B} = \mathbb{B}_{W'}$.

We also note that for each $\square \in \mathbb{B}$,

$$\sum_{T \in \mathbb{W}_{\square}} \|f_T\|_{L^p}^2 \sim \frac{W'}{W} \sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2. \quad (4.8)$$

Finally, we dyadically pigeonhole the cubes $Q \subset Y$ according to the number of $\square \in \mathbb{B}$ so that $Q \subset Y_{\square}$. We get a subset $Y' \subset Y$ so that for each cube $Q \subset Y'$, $Q \subset Y_{\square}$ for $\sim M''$ choices of $\square \in \mathbb{B}$, and Q obeys (4.7). Moreover, by dyadic pigeonholing, we have $|Y'| \approx |Y|$. Since each cube $Q \subset Y$ had approximately equal L^p norm, we also get $\|f\|_{L^p(Y')} \approx \|f\|_{L^p(Y)}$.

We also note that

$$M'M'' \leq M.$$

because a cube $Q \subset Y'$ belongs to Y_{\square} for $\sim M''$ different \square , and if $Q \subset Y_{\square}$, then it belongs to T for $\sim M'$ different $T \in \mathbb{W}_{\square}$.

Similarly, we note that

$$W'|\mathbb{B}| \leq W$$

because for each $\square \in \mathbb{B}$, $|\mathbb{W}_{\square}| \sim W'$, and \mathbb{W}_{\square} are disjoint subsets of \mathbb{W} .

Now we are ready to begin our estimate. For each $Q \subset Y'$, we have

$$\|f\|_{L^p(Q)} \lesssim \left\| \sum_{\square \in \mathbb{B}: Q \subset Y_{\square}} f_{\square} \right\|_{L^p(Q)}.$$

Applying decoupling as in (4.5), this is bounded by

$$\lesssim \left(\sum_{\square \in \mathbb{B}: Q \subset Y_{\square}} \|f_{\square}\|_{L^p(Q)}^2 \right)^{1/2}.$$

The number of terms in the sum is $\sim M''$. Applying Hölder, we get

$$\lesssim (M'')^{\frac{1}{2} - \frac{1}{p}} \left(\sum_{\square \in \mathbb{B}: Q \subset Y_{\square}} \|f_{\square}\|_{L^p(Q)}^p \right)^{1/p}.$$

We raise this inequality to the p^{th} power and sum over $Q \subset Y'$ to get

$$\|f\|_{L^p(Y)}^p \lesssim \|f\|_{L^p(Y')}^p \lesssim (M'')^{\frac{p}{2}-1} \sum_{\square \in \mathbb{B}} \|f_{\square}\|_{L^p(Y_{\square})}^p.$$

Now we can use our induction on scales—Eq. (4.6)—which gives

$$\lesssim R^{p\epsilon/2} \left(\frac{M'M''}{W'} \right)^{\frac{p}{2}-1} \sum_{\square \in \mathbb{B}} \left(\sum_{T \in \mathbb{W}_{\square}} \|f_T\|_{L^p}^2 \right)^{p/2}.$$

By (4.8), this is

$$\lesssim R^{p\epsilon/2} \left(\frac{M'M''}{W} \right)^{\frac{p}{2}-1} \frac{|\mathbb{B}|W'}{W} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2 \right)^{p/2}.$$

Since $M'M'' \leq M$ and $|\mathbb{B}|W' \leq W$, we get

$$\lesssim R^{p\epsilon/2} \left(\frac{M}{W} \right)^{\frac{p}{2}-1} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2 \right)^{p/2}.$$

Putting everything together and taking account of \lesssim throughout, we get

$$\|f\|_{L^p(Y)} \lesssim R^{3\epsilon/4} \left(\frac{M}{W} \right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2 \right)^{1/2}.$$

This closes the induction and finishes the proof. \square

One can also apply a rescaling to this theorem. If we rescale in Fourier space by a factor λ , then each $R^{-1/2} \times \cdots \times R^{-1}$ block θ is replaced by a $\lambda R^{-1/2} \times \cdots \times \lambda R^{-1}$ block. There is a corresponding rescaling in physical space so that each $R^{1/2+\delta} \times \cdots \times R$ tube T is replaced by a $\lambda^{-1} R^{1/2+\delta} \times \cdots \times \lambda^{-1} R$ tube T . The case of interest for us is $\lambda = R$.

Corollary 4.3 *Suppose that $S \subset \mathbb{R}^d$ is a strictly convex C^2 hypersurface with Gaussian curvature ~ 1 . For any $\epsilon > 0$, suppose there exists $0 < \delta \ll \epsilon$ satisfying the following. Suppose that the 1-neighborhood of RS is partitioned into $R^{1/2} \times \cdots \times R^{1/2} \times 1$ blocks θ . For each θ , let \mathbb{T}_{θ} be a set of tubes of dimensions $R^{-1/2+\delta} \times 1$ with long axis perpendicular to θ , and let $\mathbb{T} = \cup_{\theta} \mathbb{T}_{\theta}$.*

Let p be in the range $2 \leq p \leq \frac{2(d+1)}{d-1}$. Let $\mathbb{W} \subset \mathbb{T}$ and suppose that each $T \in \mathbb{W}$ lies in the unit ball. Let $W = |\mathbb{W}|$. Suppose that $f = \sum_{T \in \mathbb{W}} f_T$, where

f_T is microlocalized to $(T, \theta(T))$. Suppose that for each $T \in \mathbb{W}$, $\|f_T\|_{L^p}$ is roughly constant. Let Y be a union of $R^{-1/2}$ -cubes in the unit ball each of which intersects at most M tubes $T \in \mathbb{W}$. Then

$$\|f\|_{L^p(Y)} \lesssim_\epsilon R^\epsilon \left(\frac{M}{W}\right)^{\frac{1}{2}-\frac{1}{p}} \left(\sum_{T \in \mathbb{W}} \|f_T\|_{L^p}^2\right)^{1/2}.$$

Corollary 4.3 is the result we will actually use in our estimates about the Falconer problem.

Theorem 4.2 is closely related to the refined Strichartz estimates from [4, 5] and [6], and we record a corollary in a similar form. To set up the statement, we need to set up a little notation. We find it most convenient to work with the case that S is a graph, so suppose S is defined by $\omega_d = \phi(\omega_1, \dots, \omega_{d-1})$, and $(\omega_1, \dots, \omega_{d-1}) \in B^{d-1}(1)$. We assume that ϕ is C^2 and that the eigenvalues of the Hessian $\nabla^2 \phi$ are ~ 1 . Then for a function $g : B^{d-1} \rightarrow \mathbb{C}$, we can define the extension operator by

$$Eg(x) = \int_{B^{d-1}} e^{2\pi i(x_1\omega_1 + \dots + x_{d-1}\omega_{d-1} + x_d\phi)} g(\omega_1, \dots, \omega_{d-1}) d\omega_1, \dots, d\omega_{d-1}. \quad (4.9)$$

We decompose B^{d-1} into finitely overlapping balls θ of radius $\sim R^{-1/2}$, and then we can decompose g as

$$g = \sum_{\theta, v} g_{\theta, v},$$

where

1. $v \in R^{1/2+\delta}\mathbb{Z}^{d-1}$;
2. $g_{\theta, v}$ is supported on θ ;
3. $\hat{g}_{\theta, v}$ is essentially supported on a ball around v of radius $R^{1/2+\delta}$;
4. the functions $g_{\theta, v}$ are approximately orthogonal;
5. $Eg_{\theta, v}$ restricted to B_R is essentially supported on a tube $T_{\theta, v}$ of radius $\sim R^{1/2+\delta}$ and length $\sim R$;
6. if we think of θ as a cap in S , then the long axis of $T_{\theta, v}$ is normal to S . Also $T_{\theta, v}$ intersects the plane $x_d = 0$ at the point $(v, 0)$.

See Section 3 of [13] for background on this wave packet decomposition, including proofs of these standard facts.

Now we are ready to state our refined Strichartz estimate.

Theorem 4.4 *Let E be the extension operator as in (4.9), where ϕ is C^2 and the eigenvalues of the Hessian $\nabla^2 \phi$ are ~ 1 . Suppose that $g : B^{d-1} \rightarrow \mathbb{C}$. For any $\epsilon > 0$, suppose there exists $0 < \delta \ll \epsilon$ satisfying the following. Suppose*

that $g = \sum_{(\theta, v) \in \mathbb{W}} g_{\theta, v}$, where $\|g_{\theta, v}\|_{L^2}$ are comparable for all $(\theta, v) \in \mathbb{W}$. Let $W = |\mathbb{W}|$. Suppose Y is a union of $R^{1/2}$ -cubes in B_R^d which each intersect $\sim M$ of the tubes $T_{\theta, v} \in \mathbb{W}$. Suppose that $p = \frac{2(d+1)}{d-1}$. Then

$$\|Eg\|_{L^p(Y)} \lesssim_\epsilon R^\epsilon \left(\frac{M}{W}\right)^{\frac{1}{2} - \frac{1}{p}} \|g\|_{L^2}.$$

Proof Let η_{B_R} be a bump function associated to the ball of radius R . We define

$$f_{\theta, v} = \eta_{B_R} E g_{\theta, v}.$$

The function $f_{\theta, v}$ is essentially supported in $T_{\theta, v}$ and its Fourier transform is essentially supported in the R^{-1} -neighborhood of θ (viewing θ as a cap in $S \subset \mathbb{R}^d$). Therefore, the functions $f_{\theta, v}$ have the right microlocalization to apply Theorem 4.2. Before doing so, we need to sort them by L^p -norm. We define

$$\mathbb{W}_\lambda := \{(\theta, v) \in \mathbb{W} : \|f_{\theta, v}\|_{L^p} \sim \lambda\}.$$

We define $g_\lambda := \sum_{(\theta, v) \in \mathbb{W}_\lambda} g_{\theta, v}$, and $W_\lambda = |\mathbb{W}_\lambda|$. Now Theorem 4.2 gives

$$\|Eg_\lambda\|_{L^p(Y)} \lesssim \left(\frac{M}{W_\lambda}\right)^{\frac{1}{2} - \frac{1}{p}} \left(\sum_{(\theta, v) \in \mathbb{W}_\lambda} \|Eg_{\theta, v}\|_{L^p(B_R)}^2\right)^{1/2}.$$

Next we note that $\|Eg_{\theta, v}\|_{L^p(B_R)} \lesssim \|g_{\theta, v}\|_{L^2}$. This is a consequence of the Strichartz or Tomas-Stein inequality, but because $Eg_{\theta, v}$ is a single wave packet, there is an even simpler argument:

$$\begin{aligned} \|Eg_{\theta, v}\|_{L^p(B_R)} &\lesssim \|Eg_{\theta, v}\|_{L^p(T_{\theta, v})} \leq |T_{\theta, v}|^{1/p} \|Eg_{\theta, v}\|_{L^\infty} \\ &\leq |T_{\theta, v}|^{1/p} \int_\theta |g_{\theta, v}| \leq |T_{\theta, v}|^{1/p} |\theta|^{1/2} \|g_{\theta, v}\|_{L^2}. \end{aligned}$$

Now $T_{\theta, v}$ has volume $R^{\frac{1}{2}(d-1)+1+\delta(d-1)}$ and θ has volume $R^{-\frac{d-1}{2}}$ and so $|T_{\theta, v}|^{1/p} |\theta|^{1/2} \lesssim R^{c\delta}$, which is negligible since $\delta \ll \epsilon$. Plugging in this bound, we get

$$\|Eg_\lambda\|_{L^p(Y)} \lesssim \left(\frac{M}{W_\lambda}\right)^{\frac{1}{2} - \frac{1}{p}} \left(\sum_{(\theta, v) \in \mathbb{W}_\lambda} \|g_{\theta, v}\|_{L^2}^2\right)^{1/2}.$$

Since all the $\|g_{\theta,v}\|_{L^2}$ are comparable, we get

$$\|Eg_\lambda\|_{L^p(Y)} \lesssim \left(\frac{M}{W_\lambda}\right)^{\frac{1}{2}-\frac{1}{p}} \left(\frac{W_\lambda}{W}\right)^{1/2} \|g\|_{L^2}.$$

We have $W_\lambda \leq W$, and the total power of W_λ on the right-hand side is positive, and so we get the bound

$$\|Eg_\lambda\|_{L^p(Y)} \lesssim \left(\frac{M}{W}\right)^{\frac{1}{2}-\frac{1}{p}} \|g\|_{L^2}.$$

Since this estimate holds for every λ , the theorem is proven. \square

5 Proof of Proposition 2.2

In this section, we prove Proposition 2.2. The proof is based on adding a refined Strichartz estimate (Corollary 4.3) to the framework of [25]. We want to show that if $\alpha > 5/4$, then

$$\int_{E_2} \|d_*^x(\mu_{1,good})\|_{L^2}^2 d\mu_2(x) < +\infty.$$

We follow Liu's approach from [25]. Let σ_t be the normalized arc length measure on the circle of radius t (normalized so that the total measure is 1). Then

$$\|d_*^x(\mu_{1,good})\|_{L^2}^2 = \int_0^\infty |\mu_{1,good} * \sigma_t(x)|^2 t^2 dt.$$

Now we would like to make use of Liu's identity:

Theorem 5.1 ([25]) *For any function $f : \mathbb{R}^2 \rightarrow \mathbb{C}$, and any $x \in \mathbb{R}^2$,*

$$\int_0^\infty |f * \sigma_t(x)|^2 t dt = \int_0^\infty |f * \hat{\sigma}_r(x)|^2 r dr.$$

Notice that on the left-hand side we have $t dt$ instead of $t^2 dt$. If $x \in E_2$, then $\mu_1 * \sigma_t(x) = 0$ unless $t \sim 1$ because E_1 and E_2 are contained in the unit disk and the distance between them is $\gtrsim 1$. Therefore, we can write

$$\int_0^\infty |\mu_1 * \sigma_t(x)|^2 t^2 dt \sim \int_0^\infty |\mu_1 * \sigma_t(x)|^2 t dt.$$

We would like to write the same thing with $\mu_{1,good}$ in place of μ_1 . To justify this, we need to argue that $\mu_{1,good}$ is essentially supported in a small neighborhood of E_1 , which we now check.

Lemma 5.2 *Let A be the complement of the $R_0^{-1/2+\delta}$ -neighborhood of E_1 . Then*

$$\int_A |\mu_{1,good}| = \text{RapDec}(R_0) \text{ and } \max_{x \in A} |\mu_{1,good}(x)| = \text{RapDec}(R_0).$$

Proof By definition,

$$\begin{aligned} \mu_{1,good} &= M_0 \mu_1 + \sum_{j,\tau} \sum_{\substack{T \in \mathbb{T}_{j,\tau} \\ T \text{ good}}} M_T \mu_1 \\ &= \psi_0^\vee * \mu_1 + \sum_{j,\tau} \sum_T \eta_T (\psi_{j,\tau}^\vee * \mu_1). \end{aligned}$$

Now ψ_0^\vee is essentially supported on a ball of radius R_0^{-1} and $\psi_{j,\tau}^\vee$ is essentially supported on a rectangle of dimensions $R_j^{-1/2} \times R_j^{-1}$ centered at the origin. Since μ_1 is supported on E_1 , the result follows. \square

Since $\mu_{1,good}$ is essentially supported in a thin neighborhood of E_1 , we can indeed say that for any $x \in E_2$,

$$\int_0^\infty |\mu_{1,good} * \sigma_t(x)|^2 t^2 dt \lesssim \int_0^\infty |\mu_{1,good} * \sigma_t(x)|^2 t dt.$$

Now we can apply Theorem 5.1 to get

$$\begin{aligned} \int_{E_2} \|d_*^x(\mu_{1,good})\|_{L^2}^2 d\mu_2(x) &\lesssim \int_{E_2} \int_0^\infty |\mu_{1,good} * \hat{\sigma}_r(x)|^2 r dr d\mu_2(x) \\ &= \int_0^\infty \left(\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) \right) r dr. \end{aligned} \quad (5.1)$$

We will use Theorem 4.2 to estimate the inner integral for each r .

Proposition 5.3 *For any $\alpha > 0$, $r > 0$, and δ sufficiently small depending on α, ϵ :*

$$\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) \leq C(R_0) r^{-\frac{\alpha+1}{3} + \epsilon} r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi,$$

where ψ_r is a weight function which is ~ 1 on the annulus $r - 1 \leq |\xi| \leq r + 1$ and decays off of it. To be precise, we could take

$$\psi_r(\xi) = (1 + |r - |\xi||)^{-100}.$$

The conclusion here is very similar to saying

$$\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) \leq C(R_0) r^{-\frac{\alpha+1}{3}+\epsilon} \|\hat{\mu}_1\|_{L^2(d\sigma_r)}^2.$$

For technical reasons, we have the bound in the form above. Before turning to the proof of Proposition 5.3, let us see how it implies Proposition 2.2. Like most previous work on the Falconer problem, the proof uses the idea of the β -dimensional energy of a measure. Recall that this energy is given by

$$I_\beta(\mu) := \int |x - y|^{-\beta} \mu(x) \mu(y).$$

If a measure μ on the unit ball obeys $\mu(B(x, r)) \lesssim r^\alpha$, then $I_\beta(\mu)$ is finite for every $\beta < \alpha$ (cf. Lemma 8.3 of [37]). In particular, $I_\beta(\mu_1) < \infty$ for every $\beta < \alpha$. There is also a Fourier representation for $I_\beta(\mu)$ (cf. Proposition 8.5 of [37]): if μ is a measure on \mathbb{R}^n , then

$$I_\beta(\mu) = c_{n,\beta} \int_{\mathbb{R}^n} |\xi|^{-(n-\beta)} |\hat{\mu}(\xi)|^2 d\xi.$$

Proof of Proposition 2.2 using Proposition 5.3 By (5.1),

$$\int_{E_2} \|d_*^x(\mu_{1,good})\|_{L^2}^2 d\mu_2(x) \lesssim \int_0^\infty \left(\int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) \right) r dr. \quad (5.2)$$

Plugging in Proposition 5.3 to bound the inner integral, we get

$$\begin{aligned} &\lesssim_{R_0} \int_0^\infty \int_{\mathbb{R}^2} r^{-\frac{\alpha+1}{3}+\epsilon} \psi_r(\xi) |\hat{\mu}_1(\xi)|^2 d\xi dr \\ &\lesssim \int_{\mathbb{R}^2} |\xi|^{-\frac{\alpha+1}{3}+\epsilon} |\hat{\mu}_1(\xi)|^2 d\xi \sim I_\beta(\mu_1) \end{aligned}$$

with $\beta = 2 - \frac{\alpha+1}{3} + \epsilon$. We know that $I_\beta(\mu_1) < \infty$ as long as $\beta < \alpha$, so we get the desired bound as long as

$$2 - \frac{\alpha+1}{3} < \alpha.$$

This is equivalent to $\alpha > 5/4$. \square

Proof of Proposition 5.3 Recall that

$$\mu_{1,good} = M_0\mu_1 + \sum_{j \geq 1, \tau} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} M_T\mu_1.$$

When we convolve with $\hat{\sigma}_r$, the only terms that remain are those with Fourier support intersecting the circle of radius r . The interesting case is when $r > 10R_0$. We will return at the end to the case $r < 10R_0$. Assuming $r > 10R_0$, $\mu_{1,good} * \hat{\sigma}_r$, up to a rapidly decaying tail that is negligible for our desired bound, is essentially equal to

$$\sum_{R_j \sim r} \sum_{\tau} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} M_T\mu_1 * \hat{\sigma}_r.$$

Let η_1 be a bump function adapted to the unit ball. We define

$$f_T = \eta_1(M_T\mu_1 * \hat{\sigma}_r).$$

We claim that each f_T is microlocalized in the way we would want to apply Corollary 4.3. If $T \in \mathbb{T}_{j,\tau}$, then we let $\theta(T)$ be the 1-neighborhood of $3\tau \cap S_r^1$. We claim that \hat{f}_T is essentially supported in $\theta(T)$. First we recall that $\widehat{M_T\mu_1}$ is essentially supported in 2τ . Therefore, the Fourier transform of $M_T\mu_1 * \hat{\sigma}_r$ is essentially supported in $2\tau \cap S_r^1$. Finally, the Fourier transform of f_T is essentially supported in the 1-neighborhood of $2\tau \cap S_r^1$, which is contained in θ . Note that θ is a rectangular block of dimensions roughly $r^{1/2} \times 1$.

Next we claim that f_T is essentially supported in $2T$. We know that $M_T\mu_1$ is supported in T . Let $\tilde{\psi}_\tau$ be a smooth bump function which is 1 on 2τ and rapidly decaying. Since the Fourier transform of $M_T\mu_1$ is essentially supported on 2τ , we have $M_T\mu_1 * \hat{\sigma}_r$ is essentially equal to $M_T\mu_1 * (\tilde{\psi}_\tau\sigma_r)^\wedge$. It is standard to check by stationary phase (or see [13, Lemma 3.1]) that $(\tilde{\psi}_\tau\sigma_r)^\wedge$ is bounded by $\text{RapDec}(r)$ on $B^2(1)$ outside of a tube of radius $r^{-1/2+\delta}$ in the direction of τ passing through the origin. So $M_T\mu_1 * (\tilde{\psi}_\tau\sigma_r)^\wedge$ is negligible on $B^2(1) \setminus 2T$. So f_T is essentially supported on $2T$.

We have $\mu_{1,good} * \hat{\sigma}_r$ is essentially equal to $\sum_{T \text{ good}} f_T$. Next we sort the f_T according to their L^p norms.

$$\mathbb{W}_\lambda := \{T : \|f_T\|_{L^p} \sim \lambda\}.$$

$$f_\lambda := \sum_{T \in \mathbb{W}_\lambda} f_T.$$

Since the number of scales λ is $\lesssim \log r$, it suffices to prove the bound

$$\int |f_\lambda(x)|^2 d\mu_2(x) \lesssim r^{-\frac{\alpha+1}{3}+\epsilon} \|\hat{\mu}_1\|_{L^2(d\sigma_r)}^2.$$

Next we divide the unit ball into $r^{-1/2}$ -squares q and sort them. We let

$$\mathcal{Q}_M := \{r^{-1/2} \text{ squares } q : q \text{ intersects } \sim M \text{ tubes } T \in \mathbb{W}_\lambda\}.$$

We let $Y_M = \bigcup_{q \in \mathcal{Q}_M} q$. Since there are only $\sim \log r$ choices of M , it suffices to bound $\int_{Y_M} |f_\lambda|^2 d\mu_2$. Next we bound the measure of Y_M .

Lemma 5.4 *For any M , the $r^{-1/2}$ -neighborhood of Y_M has measure*

$$\mu_2(N_{r^{-1/2}}(Y_M)) \lesssim \frac{|\mathbb{W}_\lambda| r^{-1/2+c(\alpha)\delta}}{M}.$$

Proof This is a double counting argument. Define $\chi(q, T) = 1$ if $q \cap T \neq \emptyset$, 0 otherwise. Let

$$J = \sum_{q \in \mathcal{Q}_M, T \in \mathbb{W}_\lambda} \mu_2(2q) \chi(q, T).$$

On one hand, one has

$$J \leq 100 \sum_{T \in \mathbb{W}_\lambda} \mu_2(4T)$$

because the q 's have only finitely overlapping, hence we over count at most 100 times $\mu_2(2q)$, and if $q \cap T \neq \emptyset$, then $2q \subset 4T$. Since a good tube T satisfies $\mu_2(4T) \lesssim r^{-1/2+c(\alpha)\delta}$, therefore

$$J \leq 100 \sum_{T \in \mathbb{W}_\lambda} \mu_2(4T) \lesssim r^{-1/2+c(\alpha)\delta} |\mathbb{W}_\lambda|.$$

On the other hand, $J \gtrsim \mu_2(N_{r^{-1/2}}(Y_M)) \cdot M$. The desired estimate follows by playing these two bounds against each other. \square

Now we are ready to bound

$$\int_{Y_M} |f_\lambda|^2 d\mu_2.$$

The Fourier support of f_λ is essentially contained in the 1-neighborhood of S_r^1 , and so f_λ is (morally) locally constant at scale $\sim r^{-1}$. Therefore we

can replace $d\mu_2$ by $\mu_2 * \eta_{1/r}$, where $\eta_{1/r}$ is a bump function with integral 1 essentially supported on a ball of radius $1/r$. Then we can use Hölder to bound

$$\int_{Y_M} |f_\lambda|^2 d\mu_2 \lesssim \left(\int_{Y_M} |f_\lambda|^6 \right)^{1/3} \left(\int_{Y_M} |\mu_2 * \eta_{1/r}|^{3/2} \right)^{2/3}.$$

To bound the first factor, we use Corollary 4.3 with $W = |\mathbb{W}_\lambda|$ wave packets and multiplicity M . We get

$$\|f_\lambda\|_{L^6(Y_M)} \lesssim \left(\frac{M}{|\mathbb{W}_\lambda|} \right)^{1/3} \left(\sum_{T \in \mathbb{W}_\lambda} \|f_T\|_{L^6}^2 \right)^{1/2}.$$

By Lemma 5.4, we can bound $M/|\mathbb{W}_\lambda|$ to get

$$\lesssim \left(\frac{r^{-1/2+c(\alpha)\delta}}{\mu_2(N_{r^{-1/2}}(Y_M))} \right)^{1/3} \left(\sum_{T \in \mathbb{W}_\lambda} \|f_T\|_{L^6}^2 \right)^{1/2}.$$

To bound the second factor, we note that μ_2 of a ball of radius r^{-1} is at most $r^{-\alpha}$. Therefore,

$$\|\mu_2 * \eta_{1/r}\|_{L^\infty} \lesssim r^{2-\alpha}.$$

And so

$$\int_{Y_M} |\mu_2 * \eta_{1/r}|^{3/2} \lesssim (r^{2-\alpha})^{1/2} \int_{Y_M} d\mu_2 * \eta_{1/r} \sim r^{1-\frac{\alpha}{2}} \mu_2(N_{r^{-1/2}}(Y_M)).$$

Plugging in these two bounds, we get

$$\begin{aligned} \int_{Y_M} |f_\lambda|^2 d\mu_2 &\lesssim r^{O(\delta)} r^{-1/3} \mu_2(N_{r^{-1/2}}(Y_M))^{-2/3} \\ &\quad \sum_{T \in \mathbb{W}_\lambda} \|f_T\|_{L^6}^2 \cdot r^{\frac{2}{3}-\frac{\alpha}{3}} \mu_2(N_{r^{-1/2}}(Y_M))^{2/3}. \end{aligned}$$

Notice that the powers of $\mu_2(N_{r^{-1/2}}(Y_M))$ cancel, leaving

$$\int_{Y_M} |f_\lambda|^2 d\mu_2 \lesssim r^{O(\delta)} r^{\frac{1-\alpha}{3}} \sum_{T \in \mathbb{W}_\lambda} \|f_T\|_{L^6}^2.$$

Next we record an elementary bound for $\|f_T\|_{L^6}$. Since f_T is essentially supported on T , $\|f_T\|_{L^6} \lesssim |T|^{1/6} \|f_T\|_{L^\infty} \sim r^{-1/12} \|f_T\|_{L^\infty}$. Recall that

$$f_T = \eta_1(M_T \mu_1 * \hat{\sigma}_r) = \eta_1 \int_{S_r^1} \widehat{M_T \mu_1} e^{2\pi i x \cdot \omega} d\sigma_r.$$

Since $\widehat{M_T \mu_1}$ restricted to S_r^1 is essentially supported on $\theta(T)$, we get

$$\|f_T\|_{L^\infty} \lesssim \sigma_r(\theta(T))^{1/2} \|\widehat{M_T \mu_1}\|_{L^2(d\sigma_r)} \sim r^{-1/4} \|\widehat{M_T \mu_1}\|_{L^2(d\sigma_r)}.$$

Therefore

$$\|f_T\|_{L^6} \lesssim r^{-1/3} \|\widehat{M_T \mu_1}\|_{L^2(d\sigma_r)}.$$

Plugging into the last bound, we get

$$\int_{Y_M} |f_\lambda|^2 d\mu_2 \lesssim r^{O(\delta)} r^{\frac{-1-\alpha}{3}} \sum_{T \in \mathbb{W}_\lambda} \|\widehat{M_T \mu_1}\|_{L^2(d\sigma_r)}^2.$$

To finish the proof of Proposition 5.3, it just remains to take δ sufficiently small depending on α so that $O(\delta) < \epsilon$ and to check that

$$\sum_{R_j \sim r} \sum_{\tau} \sum_{T \in \mathbb{T}_{j,\tau}} \int |\widehat{M_T \mu_1}|^2 d\sigma_r \lesssim r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi. \quad (5.3)$$

Morally, we are showing that the $\widehat{M_T \mu_1}$ are approximately orthogonal with respect to $d\sigma_r$ and/or ψ_r . The pieces $\widehat{M_T \mu_1}|_{S_r^1}$ correspond to the wave packet decomposition of $\hat{\mu}_1 * \hat{\sigma}_r$. It's a standard fact that the wave packets in a wave packet decomposition are approximately orthogonal. (For instance, see Section 3 of [13] for related orthogonality arguments.) But because of the direction of (5.3), it takes some extra care to be completely rigorous. In particular, it makes matters easier to put ψ_r instead of $d\sigma_r$ on the right-hand side of (5.3), although we're not sure whether this is necessary. Now we turn to the details.

Recall that ψ_r is a weight function which is ~ 1 on the annulus $r-1 \leq |\xi| \leq r+1$ and then rapidly decaying. Similarly, define $\psi_{j,\tau,r}$ to be a weight function which is roughly 1 on the intersection of τ with the annulus $r-1 \leq |\xi| \leq r+1$ and then rapidly decaying. We recall that if $T \in \mathbb{T}_{j,\tau}$, then $\widehat{M_T \mu_1}$ is rapidly decaying outside of τ . Since $M_T \mu_1$ is supported in $T \subset B^2(1)$, its Fourier transform is morally locally constant on scale 1. Therefore, for any $T \in \mathbb{T}_{j,\tau}$, we have

$$\int |\widehat{M_T \mu_1}|^2 d\sigma_r \lesssim r^{-1} \int |\widehat{M_T \mu_1}|^2 \psi_{j,\tau,r} d\xi,$$

where the r^{-1} comes because σ_r is the normalized arc-length measure on S_r^1 , which is equal to approximately $1/r$ times arc length measure. Next we expand out

$$r^{-1} \int |\widehat{M_T \mu_1}|^2 \psi_{j,\tau,r} d\xi = r^{-1} \int |\hat{\eta}_T * (\psi_{j,\tau} \hat{\mu}_1)|^2 \psi_{j,\tau,r} d\xi.$$

Since $\hat{\eta}_T$ is essentially supported in a rectangle of dimensions $R^{1/2} \times 1$, with the long direction parallel to S_r^1 at points in $\tau \cap S_r^1$, we can bound

$$r^{-1} \int |\hat{\eta}_T * (\psi_{j,\tau} \hat{\mu}_1)|^2 \psi_{j,\tau,r} d\xi \lesssim r^{-1} \int |\hat{\eta}_T * (\tilde{\psi}_{j,\tau,r} \hat{\mu}_1)|^2 d\xi,$$

where $\tilde{\psi}_{j,\tau,r}$ is again rapidly decaying outside of $\tau \cap \{r-1 \leq |\xi| \leq r\}$, but a bit more slowly than $\psi_{j,\tau,r}$. The point of all these adjustments is that we can now apply Plancherel in a clean way:

$$\sum_{T \in \mathbb{T}_{j,\tau}} \int |\widehat{M_T \mu_1}|^2 d\sigma_r \lesssim r^{-1} \sum_{T \in \mathbb{T}_{j,\tau}} \int |\eta_T|^2 |\tilde{\psi}_{j,\tau,r}^\vee * \mu_1|^2 dx.$$

Since any point x lies in $\lesssim 1$ different $T \in \mathbb{T}_{j,\tau}$, the last expression is bounded by

$$\lesssim r^{-1} \int |\tilde{\psi}_{j,\tau,r}^\vee * \mu_1|^2 dx = r^{-1} \int |\tilde{\psi}_{j,\tau,r} \hat{\mu}_1|^2 d\xi.$$

So now

$$\begin{aligned} \sum_{R_j \sim r} \sum_{\tau} \sum_{T \in \mathbb{T}_{j,\tau}} \int |\widehat{M_T \mu_1}|^2 d\sigma_r &\lesssim r^{-1} \sum_{R_j \sim r, \tau} \int |\tilde{\psi}_{j,\tau,r} \hat{\mu}_1|^2 d\xi \\ &= r^{-1} \int \left(\sum_{R_j \sim r, \tau} \tilde{\psi}_{j,\tau,r}^2 \right) |\hat{\mu}_1|^2 d\xi. \end{aligned}$$

The regions where $\tilde{\psi}_{j,\tau,r}$ are ~ 1 tile the annulus $r-1 \leq |\xi| \leq r+1$, with each point lying in $\lesssim 1$ regions. Therefore, $\sum_{R_j \sim r, \tau} \tilde{\psi}_{j,\tau,r}^2$ is ~ 1 on the annulus $r-1 \leq \xi \leq r+1$ and rapidly decaying elsewhere. So a weight function ψ_r satisfying the desired properties can be chosen so that $\sum_{R_j \sim r, \tau} \tilde{\psi}_{j,\tau,r}^2 \lesssim \psi_r$, and we get

$$\sum_T \int |\widehat{M_T \mu_1}|^2 d\sigma_r \lesssim r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi.$$

This gives (5.3) and finishes the proof for the main case $r > 10R_0$.

If $r < 10R_0$, we give a more elementary estimate. It is rather lossy, but the loss can be absorbed into the factor $C(R_0)$. We write

$$\begin{aligned} \int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) &\leq \|\mu_{1,good} * \hat{\sigma}_r\|_\infty^2 \leq \|\widehat{\mu_{1,good}}\|_{L^1(d\sigma_r)}^2 \\ &\leq \|\widehat{\mu_{1,good}}\|_{L^2(d\sigma_r)}^2. \end{aligned}$$

Recall that $\mu_{1,good}$ is the sum of the good $M_T \mu_1$ while μ_1 is the sum of all $M_T \mu_1$. As we discussed above, the $M_T \mu_1$ are approximately orthogonal with respect to ψ_r , and so a similar argument to the one above shows that

$$\|\widehat{\mu_{1,good}}\|_{L^2(d\sigma_r)}^2 \lesssim r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi.$$

Since $r \leq 10R_0$, we get

$$\begin{aligned} \int_{E_2} |\mu_{1,good} * \hat{\sigma}_r(x)|^2 d\mu_2(x) &\lesssim r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi \\ &\lesssim C(R_0) r^{-\frac{\alpha+1}{3}+\epsilon} r^{-1} \int |\hat{\mu}_1|^2 \psi_r d\xi. \end{aligned}$$

□

6 Train track examples

As we mentioned in the introduction, when $\alpha < 4/3$, there are examples of measures where the Mattila integral is infinite, and the related L^2 integral in Liu's framework is also infinite. The relevant sets look like several train tracks. These train track examples are based on the train track example in [22] (page 151). In this section, we discuss these measures and their properties.

Proposition 6.1 *For every $1 \leq \alpha < 4/3$ and every B , there is a probability measure μ on $B^2(1)$ with the following properties:*

1. *For any ball $B(x, r)$, $\mu(B(x, r)) \lesssim r^\alpha$.*
2. *If $d(x, y) := |x - y|$, then*

$$\|d_*(\mu \times \mu)\|_{L^2} > B.$$

3. *If $d^x(y) := |x - y|$, then for every x in the support of μ ,*

$$\|d_*^x(\mu)\|_{L^2} > B.$$

Proof Let R be a large parameter. Let A_R be the set of points $(x_1, x_2) \in [0, 1]^2$ where $0 \leq x_1 \leq R^{-1/2}$ and where, for some integer M ,

$$MR^{-\alpha/2} \leq x_2 \leq MR^{-\alpha/2} + R^{-1}.$$

This set is reminiscent of a train track. The slats of the train track are rectangles with dimensions $R^{-1/2} \times R^{-1}$, and there are $\sim R^{\alpha/2}$ slats evenly spaced inside a vertical rectangle of dimension $R^{-1/2} \times 1$. We form a set E_R by taking the union of $R^{\frac{\alpha-1}{2}}$ train tracks that are evenly spread. To be definite, let us define $A_{R,l}$ to be the translate of A_R by the vector $(R^{-\frac{\alpha-1}{2}}l, 0)$. and then define E_R to be the union of $A_{R,l}$ as l goes from 0 to $R^{\frac{\alpha-1}{2}}$. (There is considerable freedom in how to take the union of the train tracks, and we could make similar examples with non-parallel train tracks also.) Let μ be the normalized area measure on E_R .

First we check that $\mu(B(x, r)) \lesssim r^\alpha$. The number of R^{-1} -boxes in E_R is $R^{\frac{\alpha-1}{2}} R^{\alpha/2} R^{1/2} = R^\alpha$. So we have to check that if $r = AR^{-1}$, then the number of R^{-1} boxes in $E_R \cap B(x, r)$ is $\lesssim A^\alpha$. If $A \leq R^{1/2}$, then $B(x, r) \cap E_R$ is contained inside one train track. The spacing between horizontal slats is $R^{-\alpha/2}$, and so the number of horizontal slats that intersect the ball $B(x, r)$ is at most

$$\max\left(1, \frac{r}{R^{-\alpha/2}}\right) = \max(1, AR^{-1}R^{\alpha/2}).$$

Each horizontal slat intersects $B(x, r)$ in at most $A R^{-1}$ -boxes. So the total number of R^{-1} -boxes in $B(x, r)$ is at most

$$\max(A, A^2 R^{-\frac{2+\alpha}{2}}) = \max(A, A^\alpha A^{2-\alpha} R^{-\frac{2-\alpha}{2}}) \leq A^\alpha,$$

where in the last inequality we used $A \leq R^{1/2}$.

Suppose $A \geq R^{1/2}$. Morally, since the train tracks are spaced evenly, the estimates will be even better than for the case $A = R^{1/2}$. Here are the details. Since the spacing between train tracks is $R^{-\frac{\alpha-1}{2}}$, the number of train tracks that $B(x, r)$ intersects is at most

$$rR^{\frac{\alpha-1}{2}} + 1 = AR^{\frac{\alpha-3}{2}} + 1.$$

Within each train track, the number of slats that $B(x, r)$ intersects is at most

$$rR^{\alpha/2} = AR^{\frac{\alpha-2}{2}}.$$

Each slat contains $R^{1/2}$ R^{-1} -boxes. So the total number of R^{-1} -boxes in $B(x, r)$ is at most

$$A^2 R^{\alpha-2} + A R^{\frac{\alpha-1}{2}} = A^\alpha A^{2-\alpha} R^{\alpha-2} + A^\alpha A^{1-\alpha} R^{\frac{\alpha-1}{2}} \leq A^\alpha,$$

where in the last inequality we used $R^{1/2} \leq A \leq R$.

Next we estimate $\int |d_*(\mu \times \mu)|^2$. The key point is that $d_*(\mu \times \mu)$ assigns a large measure to each interval $I_M = [MR^{-\alpha/2} - 2R^{-1}, MR^{-\alpha/2} + 2R^{-1}]$, where M is an integer with $M \sim R^{\alpha/2}$. Indeed, if x is any point in E_R , and if y lies in the same train track as x , in a horizontal slat which is M steps from the horizontal slat containing x , then $|x - y| \in [MR^{-\alpha/2} - 2R^{-1}, MR^{-\alpha/2} + 2R^{-1}]$. The μ_R measure of a single slat is $R^{-\alpha+1/2}$, because the slat contains $R^{1/2}$ R^{-1} -boxes, which each have μ_R measure $R^{-\alpha}$. Therefore,

$$d_*(\mu \times \mu)(I_M) \gtrsim R^{-\alpha+\frac{1}{2}}.$$

By Cauchy–Schwarz,

$$\int_{I_M} d_*(\mu \times \mu)^2 dt \gtrsim \frac{R^{-2\alpha+1}}{|I_M|} \sim R^{-2\alpha+2}.$$

The number of different I_M is $\sim R^{\alpha/2}$, and so

$$\int d_*(\mu \times \mu)^2 dt \gtrsim R^{-\frac{3}{2}\alpha+2}.$$

If $\alpha < 4/3$, then the power of R is positive, and $\int d_*(\mu \times \mu)^2 dt$ goes to infinity with R .

Finally we estimate $\int |d_*^x(\mu)|^2$. The computation is similar: $d_*^x(\mu)$ assigns a large measure to each interval I_M defined above. In fact, just as above,

$$d_*^x(\mu)(I_M) \gtrsim R^{-\alpha+\frac{1}{2}}.$$

because if y lies in the slat of E_R lying in the same train track as x and M horizontal slats from the horizontal slat containing x , then $d^x(y) = |x - y| \in I_M$, and the μ_R measure of this slat is $R^{-\alpha+1/2}$. Then just as above we get

$$\begin{aligned} \int_{I_M} d_*^x(\mu)^2 dt &\gtrsim \frac{R^{-2\alpha+1}}{|I_M|} \sim R^{-2\alpha+2} \text{ and} \\ \int d_*^x(\mu)^2 dt &\gtrsim R^{-\frac{3}{2}\alpha+2}. \end{aligned}$$

If $\alpha < 4/3$, then the right-hand side tends to infinity as desired. \square

We can also take limits of these examples with different scalings. Suppose that R_j is a sequence of scales that goes to infinity rapidly (say, $R_{j+1} = 2^{R_j}$). Define

$$E_j = \bigcap_{i=1}^j E_{R_i}, \text{ and}$$

$$E = E_\infty = \bigcap_{i=1}^{\infty} E_{R_i}.$$

Define μ_i to be μ_{R_i} restricted to E_i and renormalized, and let $\mu = \mu_\infty$ be a weak limit of the measures μ_i . It is not hard to check that the Hausdorff dimension of E is α , that $\mu(B(x, r)) \lesssim r^\alpha$, and that $\|d_*(\mu \times \mu)\|_{L^2} = +\infty$ and $\|d_*^x(\mu)\|_{L^2} = +\infty$ for each $x \in E$.

7 Generalization to other norms: proof of Theorem 1.3

In this section, we consider the generalization of the Falconer problem where the Euclidean norm is replaced by other norms. We will show that our main theorem generalizes to other norms as long as the unit ball of the norm is strictly convex and smooth.

Theorem 7.1 *Let K be a symmetric convex body in \mathbb{R}^2 whose boundary ∂K is C^∞ smooth and has everywhere positive curvature bounded from above and below. Let $\|\cdot\|_K$ denote the norm with unit ball K . Define the pinned distance set*

$$\Delta_{x,K}(E) := \{\|x - y\|_K\}_{y \in E}.$$

If $E \subset \mathbb{R}^2$ has Hausdorff dimension $> 5/4$, then there exists $x \in E$ so that the pinned distance set $\Delta_{x,K}(E)$ has positive Lebesgue measure.

Many parts of the proof work in the same way, and we will only discuss the required changes. The most interesting new ingredient is a generalization of Liu's identity—Theorem 5.1.

Let us start by discussing the analogue of the train track examples for a general norm. This will help us motivate the right way to decompose μ_1 into pieces $M_T \mu_1$. A train track consists of many parallel slats of dimensions $\sim R^{-1/2} \times R^{-1}$ contained in a larger rectangle of dimensions $\sim R^{-1/2} \times 1$. In the original Euclidean case, the direction of the slats is perpendicular to the direction of the larger rectangle. But to build an interesting example for the norm $\|\cdot\|_K$, the angle of the slats should be dictated by the geometry of K in the following way. Suppose that the long axis of the large rectangle

is parallel to a vector v . By rescaling, we can assume that $v \in \partial K$. Then build a train track where the slats are rectangles with long axis parallel to the tangent vector of K at v . In this case, if x lies in one slat in the train track, and y_1, y_2 lie in the same slat at the opposite side of the train track, then $\|x - y_1\|_K = \|x - y_2\|_K + O(R^{-1})$. From here on, train track examples have the same properties as in the Euclidean case.

The angles of the slats have a nice interpretation in terms of the dual norm K^* . We recall here some standard facts about dual norms. Let $\|\cdot\|_{K^*}$ be the dual norm to $\|\cdot\|_K$, and let K^* be the unit ball of the dual norm, which will also be smooth and strictly convex. Recall that the dual norm is defined by

$$\|\omega\|_{K^*} = \sup_{v \in K} \omega \cdot v.$$

By strict convexity, there is a unique $v \in K$ which achieves the supremum, which we denote by $v(\omega)$. The plane $\omega \cdot v = \omega \cdot v(\omega)$ is tangent to K at $v(\omega)$, and so we see that ω is normal to ∂K at $v(\omega)$. Similarly, for each vector v , there is a unique $\omega(v) \in \partial K^*$ so that $\omega(v) \cdot v = \|v\|_K$. The plane $\{\omega : \omega \cdot v = \omega(v) \cdot v\}$ is tangent to K^* at $\omega(v)$ and so v is normal to ∂K^* at $\omega(v)$. If $\omega \in \partial K^*$ and $v \in \partial K$, then $v \cdot \omega = 1$ if and only if $v = v(\omega)$ if and only if $\omega = \omega(v)$. Therefore $\omega(v(\omega)) = \omega$ and $v(\omega(v)) = v$. Since $\omega(v)$ is normal to ∂K at v , the map $\omega : \partial K \rightarrow \partial K^*$ is essentially the Gauss map. Because the curvature of ∂K is ~ 1 , the map ω is bilipschitz: for $v_1, v_2 \in \partial K$, $|\omega(v_1) - \omega(v_2)| \sim |v_1 - v_2|$. Therefore the map $v : \partial K^* \rightarrow \partial K$ is bilipschitz. This shows that the curvature of K^* is ~ 1 .

We can now generalize the decomposition $\mu_1 = \sum_T M_T \mu_1$ to the case of general norms $\|\cdot\|_K$, where $M_T \mu_1$ is designed to isolate the train track configurations described above. We let R_j and τ and $\psi_{j,\tau}$ be the same as in the Euclidean case. But we redefine the tubes $\mathbb{T}_{j,\tau}$. For each τ , consider $\partial(RK^*)$ where R is chosen so that $\partial(RK^*) \cap \tau$ is non-empty (so $R \sim R_j$). Then we let $\mathbb{T}_{j,\tau}$ be a set of tubes T with long direction perpendicular to $\partial(RK^*) \cap \tau$. In other words, if $\omega \in \tau$, then the direction of the tubes T is $v(\omega)$. As before, the dimensions of the tubes are $R^{-1/2+\delta} \times 1$ and the set of $T \in \mathbb{T}_{j,\tau}$ covers $B^2(2)$. We choose η_T so that $\sum_{T \in \mathbb{T}_{j,\tau}} \eta_T$ is 1 on $B^2(2)$. Then we define, for each $T \in \mathbb{T}_{j,\tau}$,

$$M_T f := \eta_T(\psi_{j,\tau} \hat{f})^\vee.$$

We define good and bad tubes in the same way as in the Euclidean case, and as before we let

$$\mu_{1,good} = M_0 \mu_1 + \sum_{j \geq 1} \sum_{\tau} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} M_T \mu_1.$$

In the Euclidean case, we studied the pushforwards $d_*^x \mu_1$ and $d_*^x \mu_{1,good}$ for $x \in E_2$. In the case of general norms, we will use a small variation of the pushforward measure. We need the small variation because of the way that the generalization of Liu's identity is stated, cf. Lemma 7.5 below.

Recall that σ_t denotes the normalized arc length measure on the (Euclidean) circle of radius t . Then as we saw at the beginning of Sect. 5,

$$d_*^x(fdarea)(t) = t\sigma_t * f(x).$$

Define σ_t^K to be the normalized (Euclidean) arc length measure on $S_K^1(t)$ —the circle of radius t in the norm $\|\cdot\|_K$. Then we define

$$T_{K,x}(fdarea)(t) = t^{1/2}\sigma_t^K * f(x).$$

We note that the support of $T_{K,x}(fdarea)$ is contained in $\Delta_{x,K}(\text{supp}(f))$. In particular, if $x \in E_2$, then the support of $T_{K,x}\mu_1$ is contained in $\Delta_{x,K}(E_1) \subset \Delta_{x,K}(E)$. For comparison, if we let $d_K^x(y) = \|x - y\|_K$, we would have

$$d_{K,*}^x(fdarea)(t) = t(k\sigma_t^K) * f(x),$$

where $k(y)$ is a smooth positive function which only depends on the direction of y —i.e. $k(\lambda y) = k(y)$ for $\lambda \neq 0$. If $x \in E_2$ and $y \in E_1$, then $\|x - y\|_K \sim 1$. Therefore, both $d_{K,*}^x \mu_1$ and $T_{K,x}\mu_1$ are supported in $t \sim 1$, and by comparing the two formulas, we see that $T_{K,x}\mu_1(t) \sim d_{K,*}^x \mu_1(t)$. In particular, this implies that $\int T_{K,x}\mu_1(t)dt \sim 1$.

To prove Theorem 7.1, we have to prove analogues of Propositions 2.1 and 2.2:

Proposition 7.2 *Let K be a symmetric convex body in \mathbb{R}^2 whose boundary ∂K is C^∞ smooth and has everywhere positive curvature bounded from above and below. If $\alpha > 1$, then for sufficiently large R_0 , there is a subset $E'_2 \subset E_2$ so that $\mu_2(E'_2) \geq 1 - \frac{1}{1000}$ and for each $x \in E'_2$,*

$$\|T_{K,x}\mu_1 - T_{K,x}\mu_{1,good}\|_{L^1} < \frac{1}{1000}.$$

Proposition 7.3 *Let K be a symmetric convex body in \mathbb{R}^2 whose boundary ∂K is C^∞ smooth and has everywhere positive curvature bounded from above and below. If $\alpha > 5/4$, then for sufficiently small δ in terms of α ,*

$$\int_{E_2} \|T_{K,x}\mu_{1,good}\|_{L^2}^2 d\mu_2(x) < +\infty.$$

7.1 Proposition 2.1 for general norms

In this section, we discuss the proof of Proposition 7.2 the analogue of Proposition 2.1. We explain what needs to be modified in the proof of Proposition 2.1. The most significant part is the proof of the first lemma, Lemma 3.1. In the context of general norms, the lemma still holds with the same statement, but when we look at the proof we will need to use the way that τ and the direction of $T \in \mathbb{T}_{j,\tau}$ are related to each other.

Lemma 7.4 *If $T \in \mathbb{T}_{j,\tau}$, and $x \in E_2$, and $x \notin 2T$, then*

$$\|T_{K,x}(M_T\mu_1)\|_{L^1} \lesssim \text{RapDec}(R_j).$$

Proof We will prove the stronger estimate that for every t :

$$T_{K,x}(M_T\mu_1)(t) \lesssim \text{RapDec}(R_j).$$

Recall that

$$T_{K,x}M_T\mu_1(t) = t^{1/2} \int_{S_K^1(x,t)} M_T\mu_1(y) d\sigma_t(y), \quad (7.1)$$

where $S_K^1(x, t)$ is the circle around x of radius t in the norm $\|\cdot\|_K$ and σ_t is the normalized arc length measure on it.

We also recall that

$$M_T\mu_1 = \eta_T(\psi_{j,\tau}\hat{\mu}_1)^\vee = \eta_T(\psi_{j,\tau}^\vee * \mu_1).$$

Now $\psi_{j,\tau}^\vee$ is concentrated on a $R_j^{-1/2} \times R_j^{-1}$ rectangle centered at 0 and it decays rapidly outside that rectangle. Since $x \in E_2$, the distance from x to the support of μ_1 is $\gtrsim 1$. Therefore, $T_{K,x}(M_T\mu_1)(t)$ is tiny unless $t \sim 1$.

To study the case when $t \sim 1$, we expand out $M_T\mu_1$:

$$M_T\mu_1(y) = \eta_T(y)(\psi_{j,\tau}\hat{\mu}_1)^\vee(y) = \eta_T(y) \int e^{2\pi i \omega \cdot y} \psi_{j,\tau}(\omega) \hat{\mu}_1(\omega) d\omega.$$

Since $|\hat{\mu}_1(\omega)| \leq 1$, and $\psi_{j,\tau}(\omega)$ is supported on τ and bounded by 1, it suffices to check that for each $\omega \in \tau$,

$$\int_{S_K^1(x,t)} \eta_T(y) e^{2\pi i \omega \cdot y} d\sigma_t(y) \leq \text{RapDec}(R_j). \quad (7.2)$$

We will prove this rapid decay by stationary phase. After a coordinate rotation, we can assume that ω has the form $(0, \omega_2)$ with $\omega_2 \sim R_j$. Let $T_0 \in \mathbb{T}_{j,\tau}$

be the tube that passes through x . The tubes of $\mathbb{T}_{j,\tau}$ have long axis perpendicular to $\partial(RK^*)$ at a point in τ . In other words, the long axis of a tube in $\mathbb{T}_{j,\tau}$ is parallel to $v(\omega)$ for $\omega \in \tau$ (up to angle $R^{-1/2}$). The tube T_0 intersects $S_K^1(x, t)$ in two arcs, and on these arcs, the normal vector to $S_K^1(x, t)$ points in the direction $\omega(v(\omega)) = \omega$, which is vertical. Now T is not T_0 —we know that $x \notin 2T$, and so the distance from T to T_0 is $\gtrsim R^{-1/2+\delta}$. By the strict convexity of K , if $y \in S_K^1(x, t) \cap T$, then the normal vector to $S_K^1(x, t)$ at y makes an angle $\gtrsim R^{-1/2+\delta}$ with the vertical.

The tube T intersects $S_K^1(x, t)$ in one or two arcs. We parametrize each arc as a graph—either $y_2 = h(y_1)$ or $y_1 = h(y_2)$ —over an interval $I(T)$. By choosing one of these two options, we can assume that h and all its derivatives are $\lesssim 1$ on $I(T)$. Let us assume first that $y_2 = h(y_1)$ since this is the more interesting case. Our integral becomes

$$\int_{I(T)} \eta_T(y_1, h(y_1)) e^{2\pi i \omega_2 h(y_1)} J(y_1) dy_1.$$

This is the same integral that appears in the proof of Lemma 3.1. If $y_1 \in I(T)$, then $(y_1, y_2) \in S_K^1(x, t) \cap T$, and so the normal vector to $S_K^1(x, t)$ at y makes an angle $\gtrsim R_j^{-1/2+\delta}$ with the vertical. Therefore, for $y_1 \in I(T)$, $|h'(y_1)| \gtrsim R_j^{-1/2+\delta}$, just like in the proof of Lemma 3.1. We can prove the desired estimate by the same stationary phase argument as in the proof of Lemma 3.1.

If $y_1 = h(y_2)$, then we have a similar but easier integral:

$$\int_{I(T)} \eta_T(h(y_2), y_2) e^{2\pi i \omega_2 y_2} J(y_2) dy_2.$$

This integral is the same as the one appearing at the end of the proof of Lemma 3.1. \square

It is also straightforward to check that if f is supported on the annulus $\{y : \|x - y\|_K \sim 1\}$, then

$$\|T_{K,x} f\|_{L^1} \lesssim \|f\|_{L^1}.$$

The rest of the proof of Proposition 2.1 is unchanged.

7.2 A general curve version of Liu's identity

In the proof of Proposition 2.2, the only ingredient that needs to be adjusted for general norms K is Liu's L^2 identity—Theorem 5.1. The argument in [25] seemingly relies heavily on the rotation invariance of the circle. We give a

different approach to Theorem 5.1, and we show that it extends (modulo a negligible tail term) to more general metrics.

The analogue of Liu's theorem is the following.

Lemma 7.5 *There is a smooth (not necessarily positive) measure σ^{K^*} on $\partial K^* = S_{K^*}^1(1)$ so that the following holds. Define a measure $\sigma_r^{K^*}$ on $S_{K^*}^1(r)$ by setting*

$$\sigma_r^{K^*}(A) = \sigma^{K^*}(A/r).$$

Suppose that $f(y)$ is supported in the annulus $\|x - y\|_K \sim 1$. Then

$$\int |f * \sigma_t^K(x)|^2 t dt \lesssim_\epsilon \int |f * \widehat{\sigma_r^{K^*}}(x)|^2 r dr + O(\|f\|_{\dot{H}^{-\frac{1}{2}+\epsilon}}^2 + \|f\|_{\dot{H}^{-\frac{1}{2}}}^2). \quad (7.3)$$

Proof We abbreviate

$$F(t) = T_{K,x} f(t) = t^{1/2} \sigma_t^K * f(x).$$

Note that the left-hand side of (7.3) is $\int F(t)^2 dt$. Also, because of the support condition on f , $F(t)$ is supported on $t \sim 1$.

We begin with an estimate for $\widehat{\sigma^K}$ which was derived by Herz [15]: When $|\xi| \geq 1$, we have

$$\begin{aligned} \widehat{\sigma^K}(\xi) &= |\xi|^{-\frac{1}{2}} \kappa(\xi)^{-\frac{1}{2}} e^{2\pi i(\|\xi\|_{K^*} - \frac{1}{8})} + |\xi|^{-\frac{1}{2}} \kappa(-\xi)^{-\frac{1}{2}} e^{-2\pi i(\|\xi\|_{K^*} - \frac{1}{8})} \\ &\quad + O(|\xi|^{-\frac{3}{2}}), \end{aligned}$$

where $\kappa(\xi)$ is the Gaussian curvature of ∂K at $v(\xi)$ —the vector with $\xi \cdot v(\xi) = \max_{v \in K} \xi \cdot v$. Note that ∂K is symmetric and so $\kappa(\xi) = \kappa(-\xi)$. Also $\widehat{\sigma_t^K}(\xi) = \widehat{\sigma^K}(t\xi)$. Therefore,

$$\widehat{\sigma_t^K}(\xi) = |t\xi|^{-1/2} \kappa(\xi)^{-\frac{1}{2}} (e^{2\pi i(t\|\xi\|_{K^*} - \frac{1}{8})} + e^{-2\pi i(t\|\xi\|_{K^*} - \frac{1}{8})}) + O(|t\xi|^{-\frac{3}{2}}).$$

This bound holds for $|t\xi| \gtrsim 1$. If $|t\xi| \lesssim 1$, then we have the simpler bound $|\widehat{\sigma_t^K}(\xi)| \lesssim 1$ which gives the same expression with a remainder term of the form $|t\xi|^{-1/2}$ on the right-hand side.

Now we return to $F(t)$. We have

$$F(t) = t^{1/2} \sigma_t^K * f(x) = t^{1/2} \int e^{2\pi i x \cdot \xi} \widehat{\sigma_t^K}(\xi) \hat{f}(\xi) d\xi.$$

Plugging in the formula above for $\widehat{\sigma_t^K}$, we get two main terms and a remainder term—for every $t \sim 1$,

$$F(t) = F_1(t) + F_2(t) + O\left(\int_{|\xi| \geq 1} |\widehat{f}(\xi)| |\xi|^{-\frac{3}{2}} d\xi + \int_{|\xi| \leq 1} |\widehat{f}(\xi)| |\xi|^{-1/2} d\xi\right),$$

where

$$F_1(t) = e^{-i\frac{\pi}{4}} \int e^{2\pi i x \cdot \xi} \widehat{f}(\xi) \kappa(\xi)^{-\frac{1}{2}} |\xi|^{-\frac{1}{2}} e^{2\pi i t \|\xi\|_{K^*}} d\xi.$$

$$F_2(t) = e^{i\frac{\pi}{4}} \int e^{2\pi i x \cdot \xi} \widehat{f}(\xi) \kappa(\xi)^{-\frac{1}{2}} |\xi|^{-\frac{1}{2}} e^{-2\pi i t \|\xi\|_{K^*}} d\xi.$$

Notice that the $t^{1/2}$ in the expression $F(t) = t^{1/2} \sigma_t^K * f(x)$ cancelled the $t^{-1/2}$ in front of $\widehat{\sigma_t^K}(\xi)$. This cancellation is the motivation for the expression $t^{1/2} \sigma_t^K * f$. It leads to a simple formula for \widehat{F}_1 and \widehat{F}_2 , which we can use to estimate $\int |F_1(t)|^2$ and $\int |F_2(t)|^2$.

Before turning to Plancherel, let us mention that the formula for $F_i(t)$ makes sense for all real t .

To find the formula for $\widehat{F}_1(r)$, we will massage the definition of $F_1(t)$ into the form $F_1(t) = \int_0^\infty e^{2\pi i r t} G(r) dr$. Then it will follow that $\widehat{F}_1(r) = G(r)$ (and that \widehat{F}_1 is supported in $[0, \infty)$.) Now we process the formula for F_1 . First, we write $\xi = r\theta$ where $r = \|\xi\|_{K^*}$ and $\theta \in S_{K^*}^1 = \partial K^*$. We can do a change of variables $d\xi = J(\theta) r dr d\theta$, where $d\theta$ is arc length measure on $S_{K^*}^1$ and $J(\theta)$ is smooth and bounded. Then we get

$$F_1(t) = e^{-i\frac{\pi}{4}} \int_0^\infty \int_{S_{K^*}^1} e^{2\pi i x \cdot \xi} \widehat{f}(\xi) \kappa(\theta)^{-\frac{1}{2}} |\xi|^{-\frac{1}{2}} J(\theta) d\theta e^{2\pi i t r} r dr,$$

where $\xi = r\theta$. We rewrote $\kappa(\xi) = \kappa(\theta)$ since $\kappa(\xi)$ only depends on the direction of ξ . Up to another smooth factor $\tilde{J}(\theta)$, we have $|\xi| = r$, and so, after redefining $J(\theta)$, we have

$$F_1(t) = \int_0^\infty \left(e^{-i\frac{\pi}{4}} \int_{S_{K^*}^1} e^{2\pi i x \cdot \xi} \widehat{f}(\xi) \kappa(\theta)^{-\frac{1}{2}} J(\theta) d\theta r^{1/2} \right) e^{2\pi i t r} dr.$$

Therefore,

$$\widehat{F}_1(r) = e^{-i\frac{\pi}{4}} \int_{S_{K^*}^1} e^{2\pi i x \cdot \xi} \widehat{f}(\xi) \kappa(\theta)^{-\frac{1}{2}} J(\theta) d\theta r^{1/2}.$$

Now define $\sigma^{K^*} = \kappa(\theta)^{-1/2} J(\theta) d\theta$, a smooth measure on $S_{K^*}^1$, and we have

$$\hat{F}_1(r) = e^{-i\frac{\pi}{4}} f * \widehat{\sigma_r^{K^*}}(x) r^{1/2}.$$

Now by Plancherel, we get

$$\int |F_1(t)|^2 dt = \int |\hat{F}_1(r)|^2 dr = \int_0^\infty |f * \widehat{\sigma_r^{K^*}}(x)|^2 r dr.$$

A similar bound applies to F_2 .

Finally, we put it all together:

$$\begin{aligned} \int |F(t)|^2 dt &\sim \int_{t \sim 1} |F(t)|^2 dt \lesssim \int |F_1(t)|^2 dt \\ &+ \int |F_2(t)|^2 dt + \text{Remainder}, \end{aligned}$$

where

$$|\text{Remainder}| \lesssim \int_{|\xi| \geq 1} |\hat{f}(\xi)| |\xi|^{-\frac{3}{2}} d\xi + \int_{|\xi| \leq 1} |\hat{f}(\xi)| |\xi|^{-\frac{1}{2}} d\xi.$$

The main two terms are $\lesssim \int_0^\infty |f * \widehat{\sigma_r^{K^*}}(x)|^2 r dr$. The remainder terms are controlled by Cauchy–Schwarz:

$$\begin{aligned} \left(\int_{|\xi| \geq 1} |\hat{f}(\xi)| |\xi|^{-\frac{3}{2}} d\xi \right)^2 &\lesssim \left(\int |\hat{f}(\xi)|^2 |\xi|^{-1+2\epsilon} d\xi \right) \left(\int_{|\xi| \geq 1} |\xi|^{-2-2\epsilon} d\xi \right) \\ &\lesssim \|f\|_{\dot{H}^{-\frac{1}{2}+\epsilon}}^2. \\ \left(\int_{|\xi| \leq 1} |\hat{f}(\xi)| |\xi|^{-\frac{1}{2}} d\xi \right)^2 &\lesssim \int_{|\xi| \leq 1} |\hat{f}|^2 |\xi|^{-1} d\xi \leq \|f\|_{\dot{H}^{-\frac{1}{2}}}^2. \quad \square \end{aligned}$$

We wish to apply this Lemma with $f = \mu_{1,good}$. Our $\mu_{1,good}$ is rapidly decaying outside of a tiny neighborhood of E_1 , and so if $x \in E_2$, $\mu_{1,good}$ is essentially supported in an annulus of the form $\|x - y\|_K \sim 1$. So we can apply Lemma 7.5, and to make use of it we just need to check that the remainder terms are finite: in other words

$$\|\mu_{1,good}\|_{\dot{H}^{-s}} < \infty,$$

for $s = 1/2$ or $1/2 - \epsilon$.

Let us first check that $\|\mu_{1,good}\|_{\dot{H}^{-s}} \lesssim \|\mu_1\|_{\dot{H}^{-s}}$. Indeed,

$$\|\mu_{1,good}\|_{\dot{H}^{-s}}^2 \lesssim \sum_{j,\tau} R_j^{-2s} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} \int |\widehat{M_T \mu_1}|^2.$$

Applying Plancherel,

$$\begin{aligned} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} \int |\widehat{M_T \mu_1}|^2 &= \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} \int |\eta_T|^2 |(\psi_{j,\tau} \hat{\mu}_1)^\vee|^2 \\ &\lesssim \int |(\psi_{j,\tau} \hat{\mu}_1)^\vee|^2 = \int |\psi_{j,\tau}|^2 |\hat{\mu}_1|^2. \end{aligned}$$

Plugging into the above, we see that

$$\|\mu_{1,good}\|_{\dot{H}^{-s}}^2 \lesssim \sum_{j,\tau} R_j^{-2s} \int |\psi_{j,\tau}|^2 |\hat{\mu}_1|^2 \lesssim \int |\xi|^{-2s} |\hat{\mu}_1|^2 = \|\mu_1\|_{\dot{H}^{-s}}^2.$$

The norm $\|\mu_1\|_{\dot{H}^{-s}}$ is related to the dimension α as follows. Recall that the β -dimensional energy of a measure μ is given by

$$I_\beta(\mu) := \int |x - y|^{-\beta} \mu(x) \mu(y).$$

There is also a Fourier representation for $I_\beta(\mu)$ (cf. Proposition 8.5 of [37]): if μ is a measure on \mathbb{R}^n , then

$$I_\beta(\mu) = c_{n,\beta} \int_{\mathbb{R}^n} |\xi|^{-(n-\beta)} |\hat{\mu}(\xi)|^2 d\xi.$$

In particular $\|\mu_1\|_{\dot{H}^{-s}}^2 = \int |\xi|^{-2s} |\hat{\mu}_1|^2 = I_{2-2s}(\mu_1)$. If a measure μ on the unit ball obeys $\mu(B(x, r)) \lesssim r^\alpha$, then $I_\beta(\mu)$ is finite for every $\beta < \alpha$ (cf. Lemma 8.3 of [37]). In particular, $I_\beta(\mu_1) < \infty$ for every $\beta < \alpha$. Therefore, $\|\mu_1\|_{\dot{H}^{-s}} < \infty$ whenever $2 - 2s < \alpha$ or $s > 1 - \alpha/2$. In particular, if $\alpha > 1$, then the remainder terms are controlled and we get

$$\begin{aligned} \|T_{K,x} \mu_{1,good}\|_{L^2}^2 &= \int |\mu_{1,good} * \sigma_t^K(x)|^2 t dt \\ &\lesssim \int_0^\infty |\mu_{1,good} * \widehat{\sigma_r^{K*}}(x)|^2 r dr + O(1). \end{aligned}$$

Integrating with respect to $d\mu_2(x)$, we get

$$\int_{E_2} \|T_{K,x}\mu_{1,\text{good}}\|_{L^2}^2 \lesssim \int_0^\infty \left(\int_{E_2} |\mu_{1,\text{good}} * \widehat{\sigma_r^{K^*}}(x)|^2 d\mu_2(x) \right) r dr + O(1).$$

As in the proof in the Euclidean case, we bound the inner integral using Corollary 4.3. The proof is essentially the same as in the Euclidean case, but when we check that each piece f_T is microlocalized correctly, we have to take into account the angles between the tubes $T \in \mathbb{T}_{j,\tau}$ and the normal vector to ∂K^* in the τ direction. Here are the details.

$$\mu_{1,\text{good}} = M_0\mu_1 + \sum_{j \geq 1, \tau} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} M_T\mu_1.$$

When we convolve with $\widehat{\sigma_r^{K^*}}$, the only terms that remain are those with Fourier support intersecting $S_{K^*}^1(r)$. So $\mu_{1,\text{good}} * \widehat{\sigma_r^{K^*}}$ is essentially equal to

$$\sum_{R_j \sim r} \sum_{\tau} \sum_{T \in \mathbb{T}_{j,\tau}, T \text{ good}} M_T\mu_1 * \widehat{\sigma_r^{K^*}}.$$

Let η_1 be a bump function adapted to the unit ball. We define

$$f_T = \eta_1 \left(M_T\mu_1 * \widehat{\sigma_r^{K^*}} \right).$$

We claim that each f_T is microlocalized in the way we would want to apply Corollary 4.3. If $T \in \mathbb{T}_{j,\tau}$, then we let $\theta(T)$ be the 1-neighborhood of $3\tau \cap S_{K^*}^1(r)$. We claim that $\widehat{f_T}$ is essentially supported in $\theta(T)$. First we recall that $\widehat{M_T\mu_1}$ is essentially supported in 2τ . Therefore, the Fourier transform of $M_T\mu_1 * \widehat{\sigma_r^{K^*}}$ is essentially supported in $2\tau \cap S_{K^*}^1(r)$. Finally, the Fourier transform of f_T is essentially supported in the 1-neighborhood of $2\tau \cap S_{K^*}^1(r)$, which is contained in $\theta(T)$. Note that $\theta(T)$ is a rectangular block of dimensions roughly $r^{1/2} \times 1$.

Next we claim that f_T is essentially supported in $2T$. We know that $M_T\mu_1$ is supported in T . Let $\tilde{\psi}_\tau$ be a smooth bump function which is 1 on 2τ and rapidly decaying. Since the Fourier transform of $M_T\mu_1$ is essentially supported on 2τ , we have $M_T\mu_1 * \widehat{\sigma_r^{K^*}}$ is essentially equal to $M_T\mu_1 * (\tilde{\psi}_\tau \sigma_r^{K^*})^\wedge$. It is standard to check by stationary phase that $(\tilde{\psi}_\tau \sigma_r^{K^*})^\wedge$ is bounded by $\text{RapDec}(r)$ on $B^2(1)$ outside of a tube of radius $r^{-1/2+\delta}$ in the direction which is normal to $S_{K^*}^1(r)$ in τ . By construction, the tube T also goes in this direction. Therefore, $M_T\mu_1 * (\tilde{\psi}_\tau \sigma_r^{K^*})^\wedge$ is negligible on $B^2(1) \setminus 2T$. So f_T is essentially supported on $2T$.

The rest of the proof of Proposition 2.2 is the same as in the Euclidean case. When we apply Theorem 4.2, the surface S that we use is ∂K^* . Since Theorem 4.2 only requires S to be a C^2 hypersurface with all extrinsic curvatures ~ 1 , it applies to ∂K^* .

7.3 Norms with some points of vanishing curvature

Theorem 7.1 applies to norms $\|\cdot\|_K$ where ∂K has strictly positive curvature everywhere. This assumption rules out the l^p norms for all $p \neq 2$. If $1 < p < \infty$, and $p \neq 2$, then there are finitely many points on the boundary of the unit ball where the curvature vanishes. Theorem 7.1 can be generalized to the case when ∂K is smooth and the curvature vanishes at finitely many points by a small extra trick. We first set up E_1 , E_2 , μ_1 , and μ_2 as usual, but then we refine them to avoid the directions where the curvature of K vanishes. Let r_0 be a small radius that we can choose later. Let B_1 be any ball of radius r_0 with $\mu_1(B_1) > 0$, and replace E_1 by $E_1 \cap B_1$. Then cover E_2 with balls of radius r_0 . We call a ball B from this covering bad if there are points $x_1 \in B_1$ and $x_2 \in B$ so that the vector $x_2 - x_1$ is parallel to a vector $v \in \partial K$ where the curvature of ∂K vanishes. The number of bad balls is $\lesssim r_0^{-1}$. Since $\mu_2(B(x, r)) \lesssim r^\alpha$ with $\alpha > 1$, we can find a good ball B_2 with $\mu_2(B_2) > 0$. Now we replace E_2 by $E_2 \cap B_2$. We redefine μ_1 and μ_2 to be supported on our new smaller sets E_1 and E_2 .

If $x_i \in B_i$, then the vector $(x_2 - x_1)/|x_2 - x_1|$ lies in an arc of ∂K of length $\sim r_0$ which avoids all the flat points of ∂K . Now we define \tilde{K} to be a different symmetric convex body so that $\partial \tilde{K}$ includes this arc of ∂K but $\partial \tilde{K}$ is smooth with strictly positive curvature everywhere. We can apply our proof to $\|\cdot\|_{\tilde{K}}$. It gives us a point $x \in E_2$ so that $d_{\tilde{K}}^x(E_1)$ has positive Lebesgue measure. But if $x_1 \in E_1$ and $x \in E_2$, then $\|x - x_1\|_K = \|x - x_1\|_{\tilde{K}}$, and so $d_K^x(E_1)$ has positive Lebesgue measure also.

8 Applications of the main results to the Erdős distance problem for general norms

The purpose of this section is to prove Corollary 1.5 and extend it to a more general collection of point sets. The following definition is due to the second listed author, Rudnev and Uriarte-Tuero [17].

Definition 8.1 Let P be a set of N points contained in $[0, 1]^d$. Define the measure

$$d\mu_P^s(x) = N^{-1} \cdot N^{\frac{d}{s}} \cdot \sum_{p \in P} \chi_B(N^{\frac{1}{s}}(x - p)) dx, \quad (8.1)$$

where χ_B is the indicator function of the ball of radius 1 centered at the origin. We say that P is s -adaptable if there exists C independent of N such that

$$I_s(\mu_P) = \int \int |x - y|^{-s} d\mu_P^s(x) d\mu_P^s(y) \leq C. \quad (8.2)$$

It is not difficult to check that if the points in set P are separated by distance $cN^{-1/s}$, then (8.2) is equivalent to the condition

$$\frac{1}{N^2} \sum_{p \neq p'} |p - p'|^{-s} \leq C. \quad (8.3)$$

In dimension $d = 2$, it is also easy to check that if the distance between any two points of P is $\gtrsim N^{-1/2}$, then (8.3) holds for any $s \in [0, 2)$, and hence P is s -adaptable.

We will prove that if P is s -adaptable, then for some $x \in P$, $|\Delta_{K,x}(P)| \gtrsim N^{4/5}$. As a special case, this implies Corollary 1.5.

Fix $s \in (\frac{5}{4}, 2)$ and define $d\mu_P^s$ as above. Note that the support of $d\mu_P^s$ is $P^{N^{-\frac{1}{s}}}$, the $N^{-\frac{1}{s}}$ -neighborhood of P . Since $I_s(\mu_P^s)$ is uniformly bounded, the proof of Theorem 1.3 implies that there exists $x_0 \in P^{N^{-1/s}}$ so that

$$\mathcal{L}(\Delta_{K,x_0}(P^{N^{-\frac{1}{s}}})) \geq c > 0,$$

where the constant c only depends on the value of C in (8.3).

Let x be a point of P with $|x - x_0| \leq N^{-1/s}$. It follows that for any y , $\|x_0 - y\|_K = \|x - y\|_K + O(N^{-1/s})$. Let $E_{N^{-1/s}}(\Delta_{K,x}(P))$ be the smallest number of $N^{-1/s}$ -intervals needed to cover $\Delta_{K,x}(P)$. We know that $\Delta_{K,x_0}(P^{N^{-1/s}})$ is contained in the $O(N^{-1/s})$ neighborhood of $\Delta_{K,x}(P)$, and so

$$\mathcal{L}(\Delta_{K,x_0}(P^{N^{-\frac{1}{s}}})) \lesssim N^{-\frac{1}{s}} E_{N^{-1/s}}(\Delta_{K,x}(P)).$$

Then our lower bound on $\mathcal{L}(\Delta_{K,x_0}(P^{N^{-\frac{1}{s}}}))$ gives

$$E_{N^{-1/s}}(\Delta_{K,x}(P)) \gtrsim N^{1/s}.$$

In other words, $\Delta_{K,x}(P)$ contains $\gtrsim N^{1/s}$ different distances that are pairwise separated by $\gtrsim N^{-1/s}$. In particular, $|\Delta_{K,x}(P)| \gtrsim N^{1/s}$. Since this holds for every $s > 5/4$, we get $|\Delta_{K,x}(P)| \gtrsim N^{4/5}$ as desired.

Acknowledgements The authors are grateful to the anonymous referees for helpful comments and suggestions. Larry Guth is supported by a Simons Investigator grant. Alex Iosevich is supported in part by the NSA Grant H98230-15-0319. Yumeng Ou is supported in part by NSF-DMS #1854148 (previously #1764454).

9 Appendix: Discussion of the lower bound on the upper Minkowski dimension of $\Delta_{x,K}(E)$ in Remark 1.4

Let ρ be a smooth cut-off function supported in the ball of radius 2 and equal to 1 in the ball of radius 1 centered at the origin. Let $\rho_\delta(x) = \delta^{-d} \rho(\delta^{-1}x)$. Following the argument in (2.5) with $\mu_{1,good}$ replaced by $\mu_{1,good} * \rho_\delta$, we see that the Lebesgue measure of the δ -neighborhood of $\Delta_{x,K}(E)$ is bounded from below by

$$\frac{(1 - \frac{2}{1000})^2}{\int |d_*^x \mu_{1,good} * \rho_\delta|^2}.$$

Following (5.2) with $\mu_{1,good}$ replaced by $\mu_{1,good} * \rho_\delta$, we see that the expression above is bounded from below by $C\delta^{\frac{5}{3} - \frac{4\alpha}{3} + \epsilon}$, hence there exists $x \in E$ such that the upper Minkowski dimension of $\Delta_{x,K}(E)$ is bounded from below by

$$1 - \left(\frac{5}{3} - \frac{4\alpha}{3} \right) = \frac{4}{3}\alpha - \frac{2}{3},$$

as claimed.

It would be interesting to obtain a lower bound on the Hausdorff dimension of $\Delta_{x,K}(E)$. If $\mu_{1,good}$ were positive, it would be sufficient to show that

$$\int_{E_2} I_\gamma d_*^x(\mu_{1,good}) d\mu_2(x) \tag{9.1}$$

is bounded with $\gamma < \frac{4}{3}\alpha - \frac{2}{3}$. This estimate follows from the same argument as in (5.2) above. Unfortunately, in view of the fact that $\mu_{1,good}$ is complex valued, the estimate (9.1) does not appear to be sufficient to draw the desired conclusion.

References

1. Bourgain, J.: Hausdorff dimension and distance sets. *Israel J. Math.* **87**(1–3), 193–201 (1994)
2. Bourgain, J.: On the Erdős-Volkmann and Katz-Tao ring conjectures. *Geom. Funct. Anal.* **13**(2), 334–365 (2003)

3. Bourgain, J., Demeter, C.: The proof of the l^2 decoupling conjecture. *Ann. Math. (2)* **182**(1), 351–389 (2015)
4. Du, X., Guth, L., Li, X.: A sharp Schrödinger maximal estimate in \mathbb{R}^2 . *Ann. Math.* **186**, 607–640 (2017)
5. Du, X., Guth, L., Li, X., Zhang, R.: Pointwise convergence of Schrödinger solutions and multilinear refined Strichartz estimate (2018). [arXiv:1803.01720](https://arxiv.org/abs/1803.01720)
6. Du, X., Guth, L., Ou, Y., Wang, H., Wilson, B., Zhang, R.: Weighted restriction estimates and application to Falconer distance set problem (2018). [arXiv:1802.10186](https://arxiv.org/abs/1802.10186)
7. Du, X., Zhang, R.: Sharp L^2 estimate of Schrödinger maximal function in higher dimensions (2018). [arXiv:1805.02775](https://arxiv.org/abs/1805.02775)
8. Eswarathasan, S., Iosevich, A., Taylor, K.: Fourier integral operators, fractal sets and the regular value theorem. *Adv. Math.* **228**, 2385–2402 (2011)
9. Erdoğan, B.: A bilinear Fourier extension theorem and applications to the distance set problem. *Int. Math. Res. Not.* **2005**(23), 1411–1425 (2005)
10. Erdős, P.: On sets of distances of n points. *Am. Math. Mon.* **53**, 248–250 (1946)
11. Falconer, K.J.: On the Hausdorff dimensions of distance sets. *Mathematika* **32**(2), 206–212 (1985)
12. Garibaldi, J.: Erdős distance problem for convex metrics. Thesis (Ph.D.), University of California, Los Angeles (2004)
13. Guth, L.: Restriction estimates using polynomial partitioning II (2016). [arXiv:1603.04250](https://arxiv.org/abs/1603.04250)
14. Guth, L., Katz, N.H.: On the Erdős distinct distance problem in the plane. *Ann. Math. (2)* **181**(1), 155–190 (2015)
15. Herz, C.S.: Fourier transforms related to convex sets. *Ann. Math. (2)* **75**(1), 81–92 (1962)
16. Hofmann, S., Iosevich, A.: Circular averages and Falconer/Erdős distance conjecture in the plane for random metrics. *Proc. Am. Mat. Soc.* **133**, 133–144 (2005)
17. Iosevich, A., Rudnev, M., Uriarte-Tuero, I.: Theory of dimension for large discrete sets and applications. *Math. Model. Nat. Phenom.* **9**(5), 148–169 (2014)
18. Iosevich, A., Łaba, I.: K -distance sets, Falconer conjecture, and discrete analogs. *Integers Electron. J. Comb. Number Theory* **5** (2005). #A08 (In: *Topics in Combinatorial Number Theory: Proceedings of the Integers Conference 2003 in Honor of Tom Brown, DIMATIA, ITI Series*, vol. 261)
19. Iosevich, A., Liu, B.: Pinned distance problem, slicing measures and local smoothing estimates (2017). [arXiv:1706.09851](https://arxiv.org/abs/1706.09851)
20. Iosevich, A., Rudnev, M.: Distance measures for well-distributed sets. *Discret. Comput. Geom.* **38**, 61–80 (2007)
21. Iosevich, A., Senger, S.: Sharpness of Falconer's $\frac{d+1}{2}$ estimate. *Ann. Acad. Sci. Fenn. Math.* **41**(2), 713–720 (2016)
22. Katz, N., Tao, T.: Some connections between Falconer's distance set conjecture and sets of Furstenberg type. *N. Y. J. Math.* **7**, 149–187 (2001)
23. Katz, N., Tardos, G.: A new entropy inequality for the Erdős distance problem. In: Pach, J. (ed.) *Towards a Theory of Geometric Graphs*, vol. 342, pp. 119–126. Contemporary Mathematics Publication, American Mathematical Society, Providence (2004)
24. Keleti, T., Shmerkin, P.: New bounds on the dimensions of planar distance sets (2018). [arXiv:1801.08745](https://arxiv.org/abs/1801.08745)
25. Liu, B.: An L^2 -identity and pinned distance problem (2018). [arXiv:1802.00350](https://arxiv.org/abs/1802.00350)
26. Mattila, P.: On the Hausdorff dimension and capacities of intersections. *Mathematika* **32**, 213–217 (1985)
27. Mattila, P.: Spherical averages of Fourier transforms of measures with finite energy: dimensions of intersections and distance sets. *Mathematika* **34**, 207–228 (1987)
28. Moser, L.: On the different distances determined by n points. *Am. Math. Mon.* **59**(2), 85–91 (1952)

29. Orponen, T.: On the distance sets of Ahlfors-David regular sets. *Adv. Math.* **307**, 1029–1045 (2017)
30. Orponen, T.: On the dimension and smoothness of radial projections (2017). [arXiv:1710.11053v2](https://arxiv.org/abs/1710.11053v2) (**preprint**)
31. Peres, Y., Schlag, W.: Smoothness of projections, Bernoulli convolutions and the dimension of exceptions. *Duke Math J.* **102**, 193–251 (2000)
32. Shmerkin, P.: On distance sets, box-counting and Ahlfors regular sets. *Discret. Anal.* **9**, 22 (2017)
33. Shmerkin, P.: On the Hausdorff dimension of pinned distance sets (2017). [arXiv:1706.00131](https://arxiv.org/abs/1706.00131)
34. Solymosi, J., Vu, V.: Near optimal bounds for the Erdős distinct distances problem in high dimensions. *Combinatorica* **28**(1), 113–125 (2008)
35. Stein, E.: *Harmonic Analysis*. Princeton University Press, Princeton (1993)
36. Wolff, T.: Decay of circular means of Fourier transforms of measures. *Int. Math. Res. Not.* **1999**(10), 547–567 (1999)
37. Wolff, T.: *Lectures on Harmonic Analysis*. University Lecture Series, vol. 29. American Mathematican Society, Providence (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.