MAXIMAL OPERATORS AND DECOUPLING
FOR $\Lambda(p)$ CANTOR MEASURES

Izabella Łaba

University of British Columbia, Department of Mathematics
Vancouver, B.C., V6T 1Z2, Canada; ilaba@math.ubc.ca

Abstract. For $2 \le p < \infty$, $\alpha' > 2/p$, and $\delta > 0$, we construct Cantor-type measures on $\mathbf{R}$ supported on sets of Hausdorff dimension $\alpha < \alpha'$ for which the associated maximal operator is bounded from $L^p_\delta(\mathbf{R})$ to $L^p(\mathbf{R})$. Maximal theorems for fractal measures on the line were previously obtained by Łaba and Pramanik [17]. The result here is weaker in that we are not able to obtain L^p estimates; on the other hand, our approach allows Cantor measures that are self-similar, have arbitrarily low dimension $\alpha > 0$, and have no Fourier decay. The proof is based on a decoupling inequality similar to that of Łaba and Wang [18].

1. Introduction

1.1. The Cantor set constuction. We define a Cantor set $E \subset \mathbf{R}$, with the associated measure μ supported on it, as follows. Let N be a large positive integer, and let $N_0 \in \mathbf{N}$ with $0 < N_0 < N$. Let Σ be a non-empty collection of subsets S of $[N] := \{0, 1, \ldots, N-1\}$ such that $|S| = N_0$ for all $S \in \Sigma$. Choose $S_1 \in \Sigma$, and let

$$A_1 = 1 + N^{-1}S_1, \quad E_1 = A_1 + [0, N^{-1}].$$

For each $a \in A_1$, choose $S_{2,a} \in \Sigma$ and let

$$A_{2,a} = a + N^{-2}S_{2,a}, \quad A_2 = \bigcup_{a \in A_1} A_{2,a}, \quad E_2 = A_2 + [0, N^{-2}].$$

Continuing by induction, let $k \ge 2$, and suppose that A_j and E_j, $j = 1, 2, \ldots, k$, have been constructed. For every $a \in A_k$, choose $S_{k+1,a} \in \Sigma$, and let

$$A_{k+1,a} = a + N^{-k-1}S_{k+1,a}, \quad A_{k+1} = \bigcup_{a \in A_k} A_{k+1,a}, \quad E_{k+1} = A_{k+1} + [0, N^{-k-1}].$$

This yields a sequence of sets $[1, 2] \supset E_1 \supset E_2 \supset E_3 \supset \ldots$, where each E_j consists of N_0^j intervals of length N^{-j}. For each j, let

$$\mu_j = \frac{1}{|E_j|}\mathbf{1}_{E_j}.$$

We will identify the functions μ_j with the absolutely continuous measures $\mu_j \, dx$. It is easy to see that μ_j converge weakly as $j \to \infty$ to a probability measure μ supported on the Cantor set $E_\infty = \bigcap_{j=1}^\infty E_j$, and that E_∞ has Hausdorff and Minkowski dimensions both equal to $\alpha := \frac{\log N_0}{\log N}$ (so that $N_0 = N^\alpha$). Furthermore, there is a constant $C_\mu > 0$ such that for all $x \in \operatorname{supp}\mu$ we have

$$(1) \qquad C_\mu^{-1} r^\alpha \le \mu((x - r, x + r)) \le C_\mu r^\alpha \quad \forall \, r > 0.$$

We are particularly interested in the self-similar case, with $S_1 = S_{j,a} = S$ for a fixed $S \in \Sigma$ and all j, a in the construction. Then μ is a self-similar measure supported on the set

$$E_\infty = \left\{ x \in [1,2] \colon x = 1 + \sum_{j=1}^\infty x_j N^{-j}, \ x_j \in S \text{ for all } j \in \mathbf{N} \right\},$$

and has similarity dimension α. However, self-similarity is not required for our proof. Our assumptions could be weakened further: for example, the same argument works (with appropriately modified constants) if the assumption that $|S_{j,a}| = N_0$ for all j, a is replaced by the weaker condition $c^{-1} N_0 \le |S_{j,a}| \le c N_0$ for some $c > 0$, as long as $S_{j,a}$ continue to be $\Lambda(p)$-sets. It should also be possible to allow constructions with slowly varying parameters as in [18].

1.2. $\Lambda(p)$ sets. Our Cantor digit set S will be provided by a theorem of Bourgain on $\Lambda(p)$ sets [3]; see also Talagrand [31].

Theorem 1. (Bourgain [3]) *Let $p > 2$. For every $N \in \mathbf{N}$ sufficiently large, there is a set $S \subset [N]$ of size $|S| \ge c_0 N^{2/p}$ such that for any set of coefficients $\{c_a\}_{a \in S}$ we have*

$$(2) \qquad \left\| \sum_{a \in S} c_a e^{2\pi i a x} \right\|_{L^p([0,1])} \le C(p) \left(\sum_{a \in S} |c_a|^2 \right)^{1/2},$$

with c_0 and $C(p)$ independent of N.

A set $S \subset [N]$ satisfying the conclusion (2) of Theorem 1 for all choices of $\{c_a\}_{a \in S}$ is called a $\Lambda(p)$ set. It is well known (see [3]) that the cardinality of any $\Lambda(p)$ set $S \subset [N]$ must be bounded by $c_1 N^{2/p}$, with c_1 depending on the constant $C(p)$ in (2). Throughout the rest of this paper we will always assume that $S \subset [N]$ is a $\Lambda(p)$ set such that

$$(3) \qquad c_0 N^{2/p} \le |S| \le c_1 N^{2/p}$$

with the constants c_0, c_1 independent of N. The existence of $\Lambda(p)$ sets obeying the first inequality in (3) is provided by Bourgain's theorem, and the second inequality is the general upper bound on the size of $\Lambda(p)$ sets as discussed above. For convenience, we will always assume that $S \subset [N-1]$, i.e. $N-1 \notin S$. This can always be arranged by removing $N-1$ from S and adjusting the constants if necessary.

1.3. Main result. We define the maximal operator with respect to a probability measure μ:

$$(4) \qquad \mathcal{M}f(x) := \sup_{t>0} \int |f(x - ty)| \, d\mu(y) = \sup_{t>0} \big[\mathcal{A}_t |f| \big](x), \quad f \in \mathcal{S},$$

where

$$\mathcal{A}_t f := \int f(x - ty) \, d\mu(y) = \int \widehat{f}(\xi) \widehat{\mu}(t\xi) e^{2\pi i x \xi} \, d\xi.$$

Our main result is a bound on $\mathcal{M}$ when μ is a Cantor measure with $\Lambda(p)$ digit sets. We first specify rigorously the class of measures under consideration.

Definition 1. We say that $E_\infty = \bigcap_{j \in \mathbf{N}} E_j \subset \mathbf{R}$ is a $\Lambda(p)$ *Cantor set* if it has been constructed as in Section 1.1, with the additional constraint that all sets S in Σ are $\Lambda(p)$ sets contained in $[N-1]$ and obeying (2) and (3). We will also say that the probability measure μ defined in Section 1.1 and supported on E_∞ is a $\Lambda(p)$ *Cantor measure,*

Bourgain's theorem ensures that if $p \in (2, \infty)$ is given, then for all sufficiently large N we can choose $N_0 = N_0(N)$ for which there exist $\Lambda(p)$ sets $S \subset [N-1]$ satisfying $|S| = N_0$ and obeying the $\Lambda(p)$ assumptions (2) and (3) for some $c_0, c_1, C(p)$ independent of N. We will fix these $c_0, c_1, C(p)$ throughout this paper, assume N to be sufficiently large, and choose N_0 and Σ accordingly. Note that $\alpha = (\log N_0)/(\log N)$ may depend slightly on N, but by (3) we will always have

$$(5) \qquad \frac{2}{p} + \frac{\log c_0}{\log N} \leq \alpha \leq \frac{2}{p} + \frac{\log c_1}{\log N},$$

so that α can be as close to $2/p$ as we wish if N is large enough.

Theorem 2. *Let $2 \leq p < \infty$. Then for any $\alpha' > 2/p$, $\delta > 0$, and for every $\Lambda(p)$ Cantor measure μ with N sufficiently large depending on p and α', we have*

(i) *μ is supported on a $\Lambda(p)$ Cantor set E_∞ of Hausdorff dimension $\alpha < \alpha'$,*

(ii) *the maximal operator $\mathcal{M}$ given by (4) obeys the bound*

$$(6) \qquad \|\mathcal{M}f\|_p \leq C'_{N,\delta} \|f\|_{L^p_\delta(\mathbf{R})}, \quad f \in \mathcal{S},$$

where $L^p_\delta(\mathbf{R})$ is the inhomogeneous Sobolev space with the norm

$$\|f\|_{L^p_\delta} = \left\| (1 - \Delta)^{\delta/2} f \right\|_p = \left\| \left[(1 + |\xi|^2)^{\delta/2} \widehat{f} \right]^\vee \right\|_p.$$

By interpolation with the trivial L^∞ bound, (6) implies the same bound with p replaced by q for any $q \in (p, \infty)$.

1.4. Averaging estimates. We briefly discuss the implications in terms of averaging estimates. Consider first a single-scale averaging operator $f \to \mathcal{A}_1 f = f * \mu$ for a probability measure μ on $\mathbf{R}$. By Young's inequality, we always have the trivial estimate

$$(7) \qquad \|\mathcal{A}_1 f\|_p \leq \|f\|_p, \quad 1 \leq p \leq \infty.$$

If the measure μ satisfies a Fourier decay condition

$$(8) \qquad \widehat{\mu}(\xi) \leq C_\beta (1 + |\xi|)^{-\beta},$$

we can improve this to an L^2-Sobolev estimate by writing

$$(9) \qquad \|\mathcal{A}_1 f\|_{L^2_\beta(\mathbf{R})} = \left\| \left[(1 + |\xi|^2)^{\beta/2} \widehat{\mu}(\xi) \widehat{f}(\xi) \right]^\vee \right\|_2 \lesssim \|f\|_2.$$

It is well known (see e.g. [34]) that if μ is supported on a set of Hausdorff dimension α, then (8) can only hold for $\beta \leq \alpha/2$. We will say that a measure μ is a *Salem measure* if it has optimal Fourier decay except possibly for the endpoint, i.e. (8) holds for all $\beta < \alpha/2$. There are numerous constructions of such measures in the literature; within the framework of our construction of $\Lambda(p)$ Cantor measures, we can ensure that μ is Salem by using the "rotations mod N" technique of [16] as in [18]. In that case, we get (9) for all $\beta < \alpha/2$.

On the other hand, when the measure μ in Theorem 2 is self-similar, it is easy to check that $\widehat{\mu}(N^j) \not\to 0$ as $j \to \infty$, so that an estimate of the form (8) cannot hold with any $\beta > 0$. For such measures, we cannot upgrade (7) to a Sobolev estimate (consider a sequence of functions with Fourier supports in $O(1)$ neighbourhoods of N^j).

Minor modifications of the proof of Theorem 2 yield the following estimates on averages of $\mathcal{A}_t f$ with respect to t.

Theorem 3. *Let $2 \leq p < \infty$, and let μ be a $\Lambda(p)$ Cantor measure as in Theorem 2. Then:*

(i) *for every r with $p < r < \infty$, we have*

$$
(10) \qquad \left\| \|\mathcal{A}_t f(x)\|_{L^r([1,2],dt)} \right\|_{L^p(dx)} \lesssim_N \|f\|_p,
$$

   provided that N is sufficiently large depending on r,

(ii) *for $p = r$, we have the following Sobolev improvement for $\gamma < \alpha/2$:*

$$
(11) \qquad \left\| \left[(1 + |\xi|^2)^{\gamma/2} \widehat{f}(\xi) \widehat{\mu}(t\xi) \right]^{\vee} \right\|_{L^p(\mathbf{R}_x \times [1,2]_t)} \lesssim_N \|f\|_p,
$$

   provided that N is sufficiently large depending on γ. We use $^{\vee}$ to denote the inverse Fourier transform in x only.

The estimates (10) and (11) hold for general $\Lambda(p)$ measures, including the self-similar case when no Fourier decay is available and we cannot do better than (7) for a fixed t.

1.5. A geometric corollary. Let $X, Y \subset \mathbf{R}$ be Lebesgue measurable. Suppose that for some choice of positive numbers $\{t(x)\}_{x \in Y}$ we have

$$
(12) \qquad \bigcup_{x \in Y} (x + t(x) E_\infty) \subset X,
$$

where E_∞ is a $\Lambda(p)$ Cantor set. Let $f = \mathbf{1}_X$. Then $\|f\|_p^p = |X|$, and $\mathcal{M}f = 1$ on Y so that $\|\mathcal{M}f\|_p^p \geq |Y|$. If we knew that $\mathcal{M}$ is bounded on $L^p(\mathbf{R})$, it would follow that $|X| \gtrsim |Y|$; in particular, it would follow that if Y has positive measure, then so does X, We are not able to prove this, but we can prove the following weaker statement.

Corollary 1. *Let $\delta > 0$, and let μ be a $\Lambda(p)$ Cantor measure as in Theorem 2 for some $p \in [2, \infty)$, with N large enough depending on p and δ. Let $X, Y \subset \mathbf{R}$ be sets such that (12) holds for some choice of $\{t(x)\}_{x \in Y}$. Let X_j and Y_j denote the N^{-j}-neighbourhoods of X and Y. Then $|X_j| \gtrsim_N N^{-j\delta p}|Y_j|$, uniformly in j. In particular, we have*

$$
\overline{\dim_M}(X) \geq \overline{\dim_M}(Y) - \delta,
$$

where we use $\overline{\dim_M}$ to denote the upper Minkowski dimension of a set, and the same is true for the lower Minkowski dimension.

To see this, let $f_j = \mathbf{1}_{X_j} * \varphi_j$, where $\varphi_j(x) = N^j \varphi(N^j x)$ for a Schwartz function φ such that $\varphi \geq 0$, $\varphi(x) \geq 1$ on $[-1,1]$, and $\widehat{\varphi}$ is supported in $|\xi| \lesssim 1$. Then $\|f_j\|_{L^p_\delta} \lesssim N^{j\delta}|X_j|^{1/p}$, and $\mathcal{M}f_j \gtrsim 1$ on Y_j, so that $\|\mathcal{M}f_j\|_p \gtrsim |Y_j|^{1/p}$. The conclusion follows from Theorem 2.

1.6. Literature overview. Maximal and averaging operators associated with measures supported on lower-dimensional submanifolds of $\mathbf{R}^d$ have been widely studied in harmonic analysis. A fundamental prototype result in this area is the spherical maximal theorem, due to Stein [29] in dimensions $d \geq 3$ and Bourgain [2] for $d = 2$, which asserts that the maximal operator associated with the Lebesgue measure on the sphere S^{d-1} in $\mathbf{R}^d$ is bounded on $L^p(\mathbf{R}^d)$ for $p > \frac{d}{d-1}$. There is a large body of work on similar estimates under varying conditions on the dimensionality, smoothness and curvature of the underlying manifold, or on the Fourier decay of the measure μ; see e.g. [7] or [30] for a partial overview.

We mention a few prior results that allow fractal measures on $\mathbf{R}^d$ with $d \geq 2$. A theorem of Rubio de Francia [22] provides a maximal estimate for measures μ on $\mathbf{R}^d$

that obey the Fourier decay condition (8) with $\beta > 1/2$. In particular, this allows fractal measures for which (8) holds. However, the result is void when $d = 1$, since measures on $\mathbf{R}$ that are not absolutely continuous can never satisfy (8) with $\beta > 1/2$. In a different direction, Iosevich and Sawyer [13] proved a maximal estimate in the special case of spherically symmetric fractals in dimensions $d \geq 2$. Iosevich, Krause, Sawyer, Taylor and Uriarte-Tuero [12] studied a variant where the averages in $\mathcal{A}_t$ are taken with respect to the spherical Lebesgue measure, but the L^p norms of f and $\mathcal{A}_t f$ are evaluated with respect to fractal measures.

Relatively little is known about maximal estimates for fractal measures in dimension 1. The first such results were proved by the author and Pramanik in [17]. Specifically, for any $0 < \epsilon < \frac{1}{3}$, there is a probability measure $\mu = \mu_\epsilon$ supported on a set $E_\infty \subset [1, 2]$ of Hausdorff dimension $1 - \epsilon$ such that the associated maximal operator $\mathcal{M}$ is bounded on $L^p(\mathbf{R})$ for $p > \frac{1+\epsilon}{1-\epsilon}$ (the best possible range would be $p > 1/(1 - \epsilon)$). Furthermore, in the case corresponding to $\epsilon = 0$, there exists a probability measure μ supported on a set $E_\infty \subset [1, 2]$, of Hausdorff dimension 1 but Lebesgue measure 0, such that $\mathcal{M}$ is bounded on $L^p(\mathbf{R})$ for all $p > 1$. This implies L^p differentiation theorems with the same range of p (answering a question of Aversa and Preiss). Results on $L^p \to L^q$ boundedness of appropriately modified maximal operators are also obtained. The construction in [17] is probabilistic and relies on "correlation conditions" (essentially, estimates on the size of intersections of two or more rescaled and translated copies of the support of μ). It does not produce explicit examples or allow self-similar sets.

Shmerkin and Suomala [27] have told me that they were able to improve this as follows: for $k \in \{2, 3, \ldots\}$, and for p_0 with the dual exponent $p_0' = k$, they construct Cantor measures μ of dimension $\alpha = \frac{1}{p_0} = 1 - \frac{1}{k}$ such that $\mathcal{M}$ is bounded on L^p for the optimal range $p > p_0$. Their proof follows the general scheme of [17], but with improved correlation conditions obtained via the methods of [25], [26].

It turns out to be very difficult to decide whether specific fractal measures can differentiate $L^p(\mathbf{R})$ for sufficiently large but finite p. Mathé (unpublished, see [14]) has reportedly constructed explicit fractal measures on $\mathbf{R}$ that cannot differentiate $L^p(\mathbf{R})$ for any $p < \infty$. The problem remains open for self-similar measures, including the middle-third Cantor measure (this question was already raised by Aversa and Preiss in the 1990s; see [17] for a more thorough discussion of the relevant history). Hochman [11] proved using entropy methods from [10] that if $X \subset \mathbf{R}$ contains a scaled copy of a Cantor set K centered at every point of a set Y of positive Hausdorff dimension, then $\dim_H(X) > \dim_H(K)$; however, the proof of differentiation would require a similar estimate with $\dim_H(K)$ replaced by $\dim_H(Y)$ which can be much larger. Our Corollary 1 is a partial result in that direction.

Our present approach via decoupling is not sufficient to yield L^p boundedness of $\mathcal{M}$ for any $p < \infty$. It is likely that this will require an additional combinatorial argument; we hope to address this in a future paper. On the other hand, Theorem 2 extends the study of maximal operators for Cantor sets on the line in several directions that were not covered in [17], [27]. We can allow α', therefore α, to be arbitrarily small (in [17], we require $\alpha > 2/3$; Shmerkin and Suomala require $\alpha \geq 1/2$). Our construction of $\Lambda(p)$ Cantor sets allows self-similar measures, with $S = S_{j,a}$ the same for all j and a. Furthermore, explicit constructions of $\Lambda(p)$ sets are available in some cases (e.g. Sidon sets for $p = 4$, see [28], [1], [23]), hence we can give explicit examples of measures for which the theorem holds. This also shows that Theorem 2 can

hold for measures on $\mathbf{R}$ without Fourier decay[1]. (In [17], the "correlation condition" imposed on our measures forced them to obey (8) with some $\beta > 0$.)

An unpleasant feature of the problem is that there does not seem to be an easy way to use (8) to obtain further improvements in Theorems 2 and 3, even when μ is Salem. This is in contrast to papers such as [19] or [21], where both decoupling (or square function estimates) and Fourier decay play a role. One issue is that Fourier-analytic proofs of maximal theorems usually require good estimates on the derivatives of $\widehat{\mu}$, which are not available in our case.

1.7. Outline of proof. We will follow a Fourier-analytic approach, developed in [19] and then adapted in [33], [21] to use decoupling instead of square function estimates. Let

$$F_\gamma f(x,t) = \langle D_t \rangle^\gamma \big(\rho(t) \mathcal{A}_t f(x) \big), \quad f \in \mathcal{S},$$

where ρ is a smoothed out characteristic function of $[1,2]$, $D_t = \frac{1}{2\pi i}\frac{\partial}{\partial t}$ and $\langle u \rangle = (1 + |u|^2)^{1/2}$. Suppose that we could prove that

$$(13) \qquad\qquad \|F_\gamma f\|_{L^p(dxdt)} \lesssim \|f\|_{L^p(dx)}, \quad f \in \mathcal{S},$$

for some $\gamma > 1/p$. Then by the Sobolev embedding theorem, we would have

$$\sup_t |\mathcal{A}_t f(x)| \lesssim \|F_\gamma f(x, \cdot)\|_{L^p(dt)};$$

taking the L^p norms in x would then yield a maximal estimate. We will not be able to actually prove (13), so instead we proceed as follows to get a weaker estimate.

By a standard reduction (see Section 2.2, it suffices to consider the single-scale maximal operator $\tilde{\mathcal{M}}$ with the range of t restricted to $[N^{-1}, 1]$. We will be seeking bounds of the form

$$(14) \qquad\qquad \|\tilde{\mathcal{M}}f\|_p \le C_N N^{j\beta}\|f\|_p, \quad j \ge j_0,$$

for some $\beta \in \mathbf{R}$ and for all f with $\widehat{f}$ supported in $|\xi| \sim N^j$. Adding the appropriate cut-offs and then applying the Sobolev embedding argument, we reduce the problem to estimating a Fourier multiplier operator F_j given by

$$\widehat{F_j f}(\xi, s) = \tilde{m}_j(\xi, s)\widehat{f}(\xi), \quad f \in \mathcal{S},$$

with the multiplier m_j supported (up to small errors) on a neighbourhood of the Cantor bush $\mathcal{K}_j = \bigcup_{a \in A_j} \mathcal{K}_{j,a}$, where

$$\mathcal{K}_{j,a} = \big\{ (\xi, s) \in \mathbf{R}^2 : |\xi| \sim N^j, \ |\xi a - s| \le 1 \big\}.$$

Let $F_j f = \sum_{a \in A_j} F_{j,a}f$, where (again, up to small errors) $F_{j,a}f$ is Fourier supported on $\mathcal{K}_{j,a}$. The main ingredient of the proof is the decoupling estimate

$$(15) \qquad\qquad \|F_j f\|_p \lesssim N^{j\epsilon}\Big(\sum_{a \in A_j} \|F_{j,a}f\|_p^2 \Big)^{1/2}$$

for some small $\epsilon > 0$. For each individual $a \in A_j$, we have the estimate

$$\|F_{j,a}f\|_p \lesssim N^{j(\gamma-\alpha)}\|f\|_p,$$

---

[1]In higher dimensions, a related but different phenomenon arises in the work of Keleti, Nagy and Shmerkin [15], Thornton [32], and Olivo and Shmerkin [20] on packing theorems and maximal operators associated with cube skeletons.

which can be proved by writing out $F_{j,a}$ in its integral operator form and using Young's inequality. Plugging this into (15), and summing over $a \in A_j$ with $|A_j| = N^{j\alpha}$, we get

$$(16) \qquad \|F_j f\|_p \lesssim N^{j\epsilon} N^{j(\gamma - \alpha/2)} \|f\|_p.$$

This implies (14) with $\beta = \gamma - \frac{\alpha}{2} + \epsilon$. In order to apply Sobolev's embedding theorem, we must have $\gamma > 1/p$, and recall from (5) that $1/p$ is very close to $\alpha/2$. Thus we will not get (14) with $\beta > 0$ (which would be needed in order to prove that $\tilde{\mathcal{M}}$, and therefore $\mathcal{M}$, is bounded on L^p), but we will be able to arrange for $\beta > 0$ to be arbitrarily small by taking sufficiently large N, which leads to Theorem 2.

The decoupling estimate used in (15) is proved in Section 4. The argument is similar to that of Łaba and Wang [18] for $\Lambda(p)$ Cantor sets on the line, and in fact uses the single-step decoupling inequality from [18, Lemma 5] as a basic building block. The proof in [18] is, in turn, based on iterating a continuous variant of Bourgain's $\Lambda(p)$ estimate in Theorem 1, and on the decoupling techniques from the work of Bourgain and Demeter [4], [5]. The additional geometric observation needed to prove a similar estimate for functions on $\mathbf{R}^2$ with Fourier transforms supported in the Cantor bush is the following: at each step of the iterative construction of the Cantor bush, the $k+1$-level branches contained in a single k-th level branch are close to parallel when restricted to ξ-intervals of length $N^{j-O(1)}$. This allows us to apply the "parallel decoupling" argument (cf. [4, Section 8]) to pass from one-dimensional Cantor sets to a two-dimensional bush.

It is easy to see that, in general, (15) cannot hold with the exponent p replaced by q with $q > p$. Indeed, if that were possible, then we could just consider functions whose Fourier transform is supported on a thin horizontal slice $\{\xi_0 \leq \xi \leq \xi_0 + 1\}$ of the Cantor bush. Then the estimate (15) becomes essentially one-dimensional and any improvement in the exponent would have to correspond to a similar improvement in Bourgain's $\Lambda(p)$ theorem, which is known to be impossible. We also note that Demeter [6] has proved a decoupling estimate for Cantor sets on a parabola; while there is at least a nominal similarity to this paper, his result is based on the curvature of the parabola and does not apply in our setting.

2. Initial reductions

2.1. Notation. We write $[N] = \{0, 1, \ldots, N-1\}$. For $d = 1, 2$, we use $|\cdot|$ to denote the Euclidean norm of a vector in $\mathbf{R}^d$, the cardinality of a finite set, or the d-dimensional Lebesgue measure of a subset of $\mathbf{R}^d$, depending on the context. We will also write $B(x, r) = \{y \in \mathbf{R}^d : |x - y| \leq r\}$. If $b \in \mathbf{R}^d$, $c \in \mathbf{R}$, and $B_1, B_2 \subset \mathbf{R}^d$, we write $b + B_1 = \{b + b_1 : b_1 \in B_1\}$, $cB_1 = \{cb_1 : b_1 \in B_1\}$, and $B_1 + B_2 = \{b_1 + b_2 : b_1 \in B_1, \ b_2 \in B_2\}$.

We use $X \lesssim Y$ to say that $X \leq CY$ for some constant $C > 0$, and $X \sim Y$ to say that $X \lesssim Y$ and $X \gtrsim Y$. The constants such as C, C', etc. and the implicit constants in $\lesssim$ may change from line to line, and may depend on d and p, but are independent of variables or parameters such as x, R, j, k. Whenever a constant depends on N, we will indicate this explicitly by writing $C_N, C(N), X \lesssim_N Y$, etc; all other constants will be independent of N.

A word on how the constants are organized: in our main decoupling inequality (Proposition 3), we lose a factor of the form C^j with C independent of j. We then want to argue that, given $\epsilon > 0$, this can be dominated by $N^{j\epsilon}$, provided that N was chosen large enough depending on ϵ. In order for this to work, it is crucial that the

constant C and all constants leading up to it be independent of N as well. All other parts of the proof are non-iterative and the dependence of the constants there on N is harmless.

For a function $f \colon \mathbf{R} \to \mathbf{C}$, we define its Fourier transform

$$\widehat{f}(\xi) = \int e^{-2\pi i x\xi} f(x)\, dx, \quad \xi \in \mathbf{R},$$

and similarly for a measure μ on $\mathbf{R}$,

$$\widehat{\mu}(\xi) = \int e^{-2\pi i x\xi}\, d\mu(x), \quad \xi \in \mathbf{R}.$$

For functions $f \colon \mathbf{R}_x \times \mathbf{R}_t \to \mathbf{C}$, we reserve ξ and s to denote the Fourier variables dual to x and t respectively, so that $\widehat{f}(\xi, s) = \iint e^{-2\pi i(x\xi + ts)} f(x, t)\, dx\, dt$. We will also sometimes use $\mathcal{F}$ for the Fourier transform, so that $\mathcal{F}f = \widehat{f}$. If the Fourier transform of a function $f(x, t)$ is taken only in one variable, we will indicate this using subscripts, e.g., $\mathcal{F}_{x \to \xi} f$. We will use the notation

$$D_x = \frac{1}{2\pi i} \frac{\partial}{\partial x}$$

so that $D_x e^{2\pi i x\xi} = \xi e^{2\pi i x\xi}$, and similarly for the t variable. We will also write $\langle u \rangle = (1 + |u|^2)^{1/2}$. If $p \in [1, \infty]$, we use p' to denote the dual exponent defined via $\frac{1}{p} + \frac{1}{p'} = 1$.

Throughout the rest of this paper, μ will be a $\Lambda(p)$ Cantor measure as in Theorem 2. We also need additional notation associated with Cantor sets. In the introduction, we defined $A_{k+1,a} = a + N^{-(k+1)} A$ for $a \in A_k$, so that $A_{k+1} = \bigcup_{a \in A_k} A_{k+1,a}$. Let

$$E_{k+1,a} := A_{k+1,a} + [0, N^{-(k+1)}] = E_{k+1} \cap [a, a + N^{-k}]$$

so that $E_{k+1} = \bigcup_{a \in A_k} E_{k+1,a}$. We will also use the decomposition $\mu = \sum_{a \in A_k} \mu_{k,a}$, where

$$\mu_{k,a} = \mu\big|_{a + [0, N^{-k}]}.$$

In the self-similar case, $\mu_{k,a}$ is a similar copy of μ, rescaled to $a + [0, N^{-k}]$ and with total mass $N_0^{-k} = N^{-k\alpha}$.

2.2. Reduction to a single scale.

Lemma 1. *Define the restricted maximal operator*

$$(17) \qquad \widetilde{\mathcal{M}}f(x) := \sup_{N^{-1} \leq t \leq 1} \left| \int f(x - ty)\, d\mu(y) \right| = \sup_{N^{-1} \leq t \leq 1} |\mathcal{A}_t f(x)|, \quad f \in \mathcal{S},$$

with $\mathcal{A}_t f := \int f(x - ty)\, d\mu(y)$ as before. Let $2 \leq p < \infty$. Suppose that for some $j_0 \in \mathbf{N}$ we have the estimate

$$(18) \qquad \|\widetilde{\mathcal{M}}f\|_p \leq C_N N^{j\beta} \|f\|_p, \quad j \geq j_0,$$

for all $f \in \mathcal{S}$ with $\operatorname{supp} \widehat{f} \subset \{N^j \leq |\xi| \leq 2N^{j+1}\}$, with the constant C_N independent of j. Then the full maximal operator $\mathcal{M}$ defined in (4) obeys

$$(19) \qquad \|\mathcal{M}f\|_p \leq C'_N \|f\|_p \quad \text{if } \beta < 0,$$

$$(20) \qquad \|\mathcal{M}f\|_p \leq C'_{N,\epsilon} \|f\|_{L^p_{\beta+\epsilon}} \quad \text{if } \beta > 0.$$

Proof. The argument here is well known (see [2]), but since it is short and we need to keep track of the scaling, we include the proof for completeness. Our presentation follows [24], with the scaling factor 2 replaced by N.

It suffices to prove that the bounds (19), (20) hold with $\mathcal{M}$ replaced by

$$\mathcal{M}_R f(x) := \sup_{0 < t < R} \left| \int f(x - ty) \, d\mu(y) \right|,$$

and with constants independent of R. By scaling, it suffices to consider $R = 1$.

Let $\phi \in C_c^\infty(B(0,2))$ be a function such that $0 \leq \phi \leq 1$ and $\phi(\xi) = 1$ for $|\xi| \leq 1$. Define $\phi_0(\xi) = \phi(N^{-1}\xi)$ and $\phi_j(\xi) = \phi(N^{-j-1}\xi) - \phi(N^{-j}\xi)$ for $j \in \mathbf{N}$. Then $\sum_{j=0}^\infty \phi_j \equiv 1$ and ϕ_j is supported in the region $N^j \leq |\xi| \leq 2N^{j+1}$ for $j > 0$. Define f_j via

$$\widehat{f_j} = \phi_j \widehat{f}$$

so that $f = \sum_{j=0}^\infty f_j$. Then

$$\mathcal{M}_1 f(x) = \sup_{k \geq 0} \sup_{N^{-k-1} \leq t \leq N^{-k}} |\mathcal{A}_t f(x)|$$

$$\leq \sup_{k \geq 0} \sup_{N^{-k-1} \leq t \leq N^{-k}} \left| \mathcal{A}_t \Big[\sum_{j < j_0 + k} f_j \Big](x) \right| + \sup_{k \geq 0} \sup_{N^{-k-1} \leq t \leq N^{-k}} \left| \mathcal{A}_t \Big[\sum_{j \geq j_0 + k} f_j \Big](x) \right|$$

$$=: I_1(x) + I_2(x).$$

The I_1 part is dominated by a constant (depending on N, j_0) multiple of the Hardy–Littlewood maximal operator, therefore bounded on all L^p with $p > 2$. To estimate I_2, we use (18) and scaling. Let $f_{j,k}(x) = N^{-k/p} f_j(N^{-k}x)$, then $\|f_{j,k}\|_p = \|f_j\|_p$ and $\widehat{f_{j,k}}$ is supported in $N^{j-k} \leq |\xi| \leq 2N^{j-k+1}$. We have

$$\mathcal{A}_t f_j(x) = N^{k/p} \big[\mathcal{A}_{N^k t} f_{j,k} \big](N^k x),$$

so that

$$I_2(x) \leq \sup_{k \geq 0} \sum_{j \geq j_0 + k} \sup_{N^{-k-1} \leq t \leq N^{-k}} |\mathcal{A}_t f_j(x)| \leq \sup_{k \geq 0} \sum_{j \geq j_0 + k} N^{k/p} \tilde{\mathcal{M}} f_{j,k}(N^k x)$$

$$\leq \left(\sum_{k \geq 0} \Big[\sum_{j \geq j_0 + k} N^{k/p} \tilde{\mathcal{M}} f_{j,k}(N^k x) \Big]^p \right)^{1/p}.$$

It follows that

$$\|I_2\|_p \leq \left(\int \sum_{k \geq 0} \Big[\sum_{j \geq j_0 + k} N^{k/p} \tilde{\mathcal{M}} f_{j,k}(N^k x) \Big]^p dx \right)^{1/p}$$

$$= \left(\sum_{k \geq 0} \Big\| \sum_{j \geq j_0 + k} N^{k/p} \tilde{\mathcal{M}} f_{j,k}(N^k \cdot) \Big\|_p^p \right)^{1/p}$$

$$\leq \left(\sum_{k \geq 0} \Big[\sum_{j \geq j_0 + k} \|\tilde{\mathcal{M}} f_{j,k}\|_p \Big]^p \right)^{1/p} \lesssim_N \left(\sum_{k \geq 0} \Big[\sum_{j \geq j_0 + k} N^{(j-k)\beta} \|f_j\|_p \Big]^p \right)^{1/p},$$

where at the last step we used (18) and that $\|f_{j,k}\|_p = \|f_j\|_p$. If $\beta < 0$, we use discrete Young's inequality and then Littlewood–Paley to estimate

$$\|I_2\|_p \lesssim_N \Big(\sum_j \|f_j\|_p^p \Big)^{1/p} \lesssim \Big\| \Big(\sum_j |f_j|^2 \Big)^{1/2} \Big\|_p \lesssim \|f\|_p.$$

If on the other hand $\beta > 0$, we have instead

$$\|I_2\|_p \lesssim_N \left(\sum_{k \geq 0} N^{-k\beta p} \left[\sum_{j \geq j_0 + k} N^{-j\epsilon} N^{j(\beta + \epsilon)} \|f_j\|_p \right]^p \right)^{1/p}$$

$$\lesssim_N \left(\sum_{k \geq 0} N^{-k\beta p} \left[\sum_{j \geq j_0 + k} N^{-j\epsilon} \|\langle D_x \rangle^{\beta + \epsilon} f\|_p \right]^p \right)^{1/p}$$

$$\lesssim_{N,\epsilon} \|f\|_{L^p_{\beta + \epsilon}}. \qquad \qquad \square$$

2.3. A multiplier problem. Following [19] (see also [21]), we perform a further reduction as follows. For $\gamma > 0$, we define the operator F_γ, mapping functions $f \in \mathcal{S}(\mathbf{R}_x^d)$ to Schwartz functions on $\mathbf{R}_x^d \times \mathbf{R}_t$:

$$F_\gamma f(x, t) = \langle D_t \rangle^\gamma \big(\rho(t) \mathcal{A}_t f(x) \big),$$

where $\rho \in C_c^\infty(\frac{1}{2N}, 2)$ is a fixed function such that $\rho \geq 0$ and $\rho \equiv 1$ on $[\frac{1}{N}, 1]$. Recall that $D_t = \frac{1}{2\pi i} \frac{\partial}{\partial t}$ and $\langle u \rangle = (1 + |u|^2)^{1/2}$, so that for a function $h(t)$ we have

$$\mathcal{F}(\langle D_t \rangle^\gamma h)(s) = \langle s \rangle^\gamma \widehat{h}(s).$$

Suppose that we can prove that for some function $f \in \mathcal{S}$,

$$(21) \qquad \qquad \|F_\gamma f\|_{L^p(dxdt)} \lesssim K \|f\|_{L^p(dx)},$$

for some γ such that $\gamma p > 1$. Then by the Sobolev embedding theorem, we have

$$\tilde{\mathcal{M}} f(x) \lesssim \sup_{N^{-1} \leq t \leq 1} |\rho(t) \mathcal{A}_t f(x)| \lesssim \|F_\gamma f(x, \cdot)\|_{L^p(dt)},$$

so that

$$(22) \qquad \qquad \|\tilde{\mathcal{M}} f\|_{L^p(dx)} \lesssim \|F f\|_{L^p(dxdt)} \lesssim K \|f\|_{L^p(dx)}.$$

Our strategy will be to prove (21) (therefore (22)) for all $f \in \mathcal{S}$ such that $\mathrm{supp}\, \widehat{f} \subset \{N^j \leq |\xi| \leq 2N^{j+1}\}$, with $K \lesssim_N N^{j\beta}$ uniformly in j, then use Lemma 1 to pass to the unrestricted maximal operator.

We first set up the appropriate band-limited operators. Let ϕ and ϕ_j be the functions defined at the beginning of the proof of Lemma 1, and define σ_j via

$$\widehat{\sigma}_j = \phi_j \widehat{\mu}$$

so that $\sigma_j \in \mathcal{S}$ and $\widehat{\mu} = \sum_{j=0}^\infty \widehat{\sigma}_j$. Let

$$F_{\gamma,j} f(x, t) = \langle D_t \rangle^\gamma \left(\rho(t) \int \widehat{f}(\xi) \widehat{\sigma}_j(t\xi) e^{2\pi i x\xi} \, d\xi \right).$$

Lemma 2. *Let* $2 \leq p < \infty$ *and* $\gamma > 1/p$. *With* $F_{\gamma,j}$ *as above, suppose that we have the estimate*

$$(23) \qquad \qquad \|F_{\gamma,j} f\|_{L^p(dxdt)} \lesssim_N N^{j\beta} \|f\|_{L^p(dx)}, \quad f \in \mathcal{S}.$$

Then $\mathcal{M}$ *obeys the conclusions (19) or (20) of Lemma 1, depending on the sign of* β.

Proof. We prove below in Lemma 3 that $F_{\gamma,j}$ is a Fourier multiplier operator on $\mathbf{R}_{x,t}^2$ with a multiplier supported in $\frac{1}{2}N^j \leq |\xi| \leq 4N^{j+2}$. It follows that for functions f with supp $\widehat{f} \subset \{N^j \leq |\xi| \leq 2N^{j+1}\}$ with $j \geq 3$, we have

$$F_\gamma f = \sum_{k=j-2}^{j+1} F_{\gamma,k} f.$$

Therefore, if (23) holds, then so do (21) and (by the above discussion) (22) with $K = C_N N^{j\beta}$. Hence the assumption (18) holds with $j_0 = 3$, and the conclusion follows from Lemma 1. $\square$

In the sequel, $\gamma > 0$ will be fixed and we will omit it from notation, writing $F_{\gamma,j} = F_j$.

Lemma 3. *We have the Fourier multiplier representation*

(24) $$\widehat{F_j f}(\xi, s) = \tilde{m}_j(\xi, s)\widehat{f}(\xi), \quad f \in \mathcal{S},$$

where $\tilde{m}_j(\xi, s)$ is a Schwartz function in 2 variables, given by

$$\tilde{m}_j(\xi, s) = \langle s \rangle^\gamma \int \sigma_j(y)\widehat{\rho}(\xi y + s)\, dy.$$

and supported in $\frac{1}{2}N^j \leq |\xi| \leq 4N^{j+2}$.

Proof. Let

$$\mathcal{A}_{t,j}f := \int f(x - ty)\sigma_j(y)\, dy = \int \widehat{f}(\xi)\widehat{\sigma}_j(t\xi)e^{2\pi i x\xi}\, d\xi.$$

Then $\rho(t)\mathcal{A}_{t,j}f \in \mathcal{S}_{x,t}$, therefore so does $F_j f$. Taking the partial Fourier transform in t, we get

$$\mathcal{F}_{t \to s}\left(F_j f\right)(x, s) = \langle s \rangle^\gamma \mathcal{F}_{t \to s}\left(\rho(t)\mathcal{A}_{t,j}f\right)(x, s)$$

$$= \langle s \rangle^\gamma \int e^{-2\pi i t s}\rho(t) \int \widehat{f}(\xi)\widehat{\sigma}_j(t\xi)e^{2\pi i x\xi}\, d\xi\, dt.$$

Interchanging the order of integration, we get that

(25) $$\mathcal{F}_{t \to s}\left(F_j f\right)(x, s) = \int e^{2\pi i x\xi}\widehat{f}(\xi)\tilde{m}_j(\xi, s)\, d\xi,$$

where

$$\tilde{m}_j(\xi, s) = \langle s \rangle^\gamma \int e^{-2\pi i t s}\rho(t)\widehat{\sigma}_j(t\xi)\, dt.$$

For $t \in$ supp $\rho \subset [\frac{1}{2N}, 2]$, $\widehat{\sigma}_j(t\xi)$ as a function of ξ is supported in $[t^{-1}N^j, 2t^{-1}N^{j+1}] \subset [\frac{1}{2}N^j, 4N^{j+2}]$. Therefore $\tilde{m}_j$ is a Schwartz function supported in $\frac{1}{2}N^j \leq |\xi| \leq 4N^{j+2}$.

Next, we rewrite $\tilde{m}_j$ as

$$\tilde{m}_j(\xi, s) = \langle s \rangle^\gamma \int e^{-2\pi i t s}\rho(t) \int \sigma_j(y)e^{-2\pi i t\xi y}\, dy\, dt$$

$$= \langle s \rangle^\gamma \int \sigma_j(y)\left[\int \rho(t)e^{-2\pi i t(\xi y + s)}\, dt\right]dy$$

$$= \langle s \rangle^\gamma \int \sigma_j(y)\widehat{\rho}(\xi y + s)\, dy,$$

as claimed. Finally, taking the Fourier transform in x in (25) proves (24) and completes the proof of the lemma. $\square$

3. Localization estimates

Recall that $\mu = \sum_{a \in A_j} \mu_{j,a}$, where $\mu_{j,a} = \mu\big|_{a+[0,N^{-j}]}$. Accordingly, let

$$\widehat{\sigma_{j,a}} = \phi_j \widehat{\mu_{j,a}}.$$

Then $F_j f = \sum_{a \in A_j} F_{j,a} f$, where

(26)
$$\widehat{F_{j,a}f}(\xi, s) = \tilde{m}_{j,a}(\xi, s) \widehat{f}(\xi),$$
$$\tilde{m}_{j,a} = \langle s \rangle^\gamma \int \sigma_{j,a}(y)\, \widehat{\rho}(\xi y + s)\, dy.$$

Recalling the definition of ϕ_j, we see that $\sigma_{j,a} = \mu_{j,a} * N^j \psi(N^j \cdot)$, where $\psi := \mathcal{F}^{-1}(\phi(N^{-1}\xi) - \phi(\xi))$ is a fixed Schwartz function. Moreover, $\mu_{j,a}$ is a rescaling to the interval $a + [0, N^{-j}]$ of a Cantor measure of the same type as μ, with the same N and t, and with total mass $N^{-j\alpha}$. It follows that

$$\sigma_{j,a}(y) = N^{j(1-\alpha)} \sigma_{j,a}^0(N^j(y - a)),$$

where $\sigma_{j,a}^0 \in \mathcal{S}$ have all Schwartz seminorms bounded uniformly in a and j.

We now fix an $a \in A_j$. For that a, we define new coordinates (ξ, τ_a) in the Fourier space:

$$(\xi, \tau_a) := (\xi, s + a\xi).$$

In the rest of this section, a will be fixed and we will supress the dependence of τ_a on a, writing $\tau = \tau_a$, Then

$$\tilde{m}_{j,a}(\xi, s) = N^{-j\alpha} \langle a\xi - \tau \rangle^\gamma \int N^j \sigma_{j,a}^0(N^j(y - a))\, \widehat{\rho}(\xi y - a\xi + \tau)\, dy$$

$$= N^{-j\alpha} \langle a\xi - \tau \rangle^\gamma \int N^j \sigma_{j,a}^0(N^j y)\, \widehat{\rho}(\xi y + \tau)\, dy$$

$$= N^{-j\alpha} \langle a\xi - \tau \rangle^\gamma \int \sigma_{j,a}^0(y)\, \widehat{\rho}(N^{-j}\xi y + \tau)\, dy.$$

Define

(27)
$$m_{j,a}(\xi, \tau) := N^{j(\alpha-\gamma)} \tilde{m}_{j,a}(\xi, s) = \frac{\langle a\xi - \tau \rangle^\gamma}{N^{j\gamma}} \lambda_{j,a}(\xi, \tau),$$

where

$$\lambda_{j,a}(\xi, \tau) = \int \sigma_{j,a}^0(y)\, \widehat{\rho}(N^{-j}\xi y + \tau)\, dy.$$

In addition to changing variables, we also normalized the multipliers to simplify the forthcoming calculations.

We also note the following representations of $F_{j,a}$ as a Fourier integral operator.

Lemma 4. *We have*

(28)
$$F_{j,a}(x, t) = N^{j(\gamma-\alpha)} \int f(y) K_{j,a}(x - y, t)\, dy,$$

where

(29)
$$K_{j,a}(x, t) = \iint e^{2\pi i((x - ta)\xi + t\tau)} m_{j,a}(\xi, \tau)\, d\xi\, d\tau.$$

Proof. Taking the inverse Fourier transform of the first equation in (26) in both x and t, we get that $F_{j,a}(x,t) = \int f(y)\tilde{K}_{j,a}(x-y,t)\,dy$, where

$$\tilde{K}_{j,a}(x,t) = \iint e^{2\pi i(x\xi+ts)}\tilde{m}_{j,a}(\xi,s)\,d\xi\,ds.$$

Substituting $\tilde{m}_{j,a}(\xi,s) = N^{j(\gamma-\alpha)}m_{j,a}(\xi,\tau)$ and changing variables $(\xi,s) \to (\xi,\tau)$ in the integral defining $\tilde{K}_{j,a}$, with Jacobian $\left|\frac{\partial(\xi,\tau)}{\partial(\xi,s)}\right| = 1$, we get (29). $\qquad\square$

Proposition 1. *For any $M \in \mathbf{N}$, we have*

$$(30) \qquad \left|K_{j,a}(x,t)\right| \le C_{M,N}N^j \min\left[\left(1 + N^{2j}|x-ta|^2\right)^{-M},\right.$$

$$(31) \qquad\qquad\qquad \left.\left(1 + |t|^2\right)^{-M}\right]$$

with $C_{M,N}$ independent of j and a.

Proof. We first prove the bound in (30). Let $u = x - ta$. Let also $L_M = (1 + N^{2j}|u|^2)^{-M}(1 + N^{2j}D_\xi^2)^M$, so that $L_M e^{2\pi i u\xi} = e^{2\pi i u\xi}$. Integrating by parts in ξ (with boundary terms 0, since $m_{j,a} \in \mathcal{S}$), we get

$$(1 + N^{2j}|u|^2)^M K_{j,a}(x,t) = (1 + N^{2j}|u|^2)^M \iint e^{2\pi i t\tau}(L_M e^{2\pi i u\xi})m_{j,a}(\xi,\tau)\,d\xi\,d\tau$$

$$= (1 + N^{2j}|u|^2)^M \iint e^{2\pi i(t\tau+u\xi)}(L_M m_{j,a})(\xi,\tau)\,d\xi\,d\tau$$

$$= \iint e^{2\pi i(t\tau+u\xi)}\left((1 + N^{2j}D_\xi^2)^M m_{j,a}\right)(\xi,\tau)\,d\xi\,d\tau.$$

Hence

$$\left|K_{j,a}(x,t)\right| \le (1 + N^{2j}|u|^2)^{-M}I_{M,N}(j),$$

where

$$I_{M,N}(j) = \iint \left|\left((1 + N^{2j}D_\xi^2)^M m_{j,a}\right)(\xi,\tau)\right|\,d\xi\,d\tau$$

$$\le \sum_{0 \le n_1+n_2 \le 2M} \iint (N^j D_\xi)^{n_1}\frac{\langle a\xi - \tau\rangle^\gamma}{N^{j\gamma}} \cdot (N^j D_\xi)^{n_2}\lambda(\xi,\tau)\,d\xi\,d\tau.$$

We need to prove that $I_{M,N} \lesssim_{M,N} N^j$. Recall that $m_{j,a}$ is supported in $\frac{1}{2}N^j \le |\xi| \le 4N^{j+2}$, so that the integration in $I_{M,N}$ is restricted to the same region. Hence it suffices to prove that the integrands are bounded by $C_{M,N}(1 + |\tau|)^{-M}$ with constants independent of j. This follows from the estimates (32) and (33) below.

- We have $\lambda_{j,a}(\xi,\tau) = \int \sigma_{j,a}^0(y)\,\widehat{\rho}(N^{-j}\xi y + \tau)\,dy$, so that for $n_2 \ge 0$

$$(N^j D_\xi)^{n_2}\lambda_{j,a}(\xi,\tau) = \int \sigma_{j,a}^0(y)y^{n_2}(D^{n_2}\widehat{\rho})(N^{-j}\xi y + \tau)\,dy.$$

For a fixed ξ, this integral as a function of τ is a convolution of two Schwartz functions whose Schwartz seminorms are bounded uniformly in $j \in \mathbf{N}$ and ξ in the range $|N^{-j}\xi| \sim_N 1$. Hence for any $M' > 0$ we have

$$(32) \qquad |(N^j D_\xi)^{n_2}\lambda_{j,a}(\xi,\tau) \lesssim_{n_2,N,M'} (1 + |\tau|)^{-M'}.$$

• We claim that for $n_1 \geq 0$ and ξ in the indicated range,

$$(33) \qquad (N^j D_\xi)^{n_1} \frac{\langle a\xi - \tau \rangle^\gamma}{N^{j\gamma}} \lesssim_{n_1, N} \langle \tau \rangle.$$

Indeed, we have

$$(N^j D_\xi)^{m_1} \frac{\langle a\xi - \tau \rangle^\gamma}{N^{j\gamma}} \lesssim \frac{\langle a\xi - \tau \rangle^{\gamma - m_1}}{N^{-j(\gamma - m_1)}}.$$

If $m_1 = 0$, we write

$$\frac{\langle a\xi - \tau \rangle}{N^j} \leq \frac{\langle \tau \rangle}{N^j} + \frac{\langle a\xi \rangle}{N^j} \lesssim \langle \tau \rangle + O(N^2) \lesssim_N \langle \tau \rangle.$$

and the claim follows. If $n_1 \geq 1$, the exponent $\gamma - n_1$ is negative, so we need the estimate

$$(34) \qquad N^j \langle a\xi - \tau \rangle^{-1} \lesssim 1 + |\tau|.$$

If $|\tau| \leq \frac{1}{4} N^j$, then for $|\xi| \geq \frac{1}{2} N^j$ we have

$$N^{-j} \langle a\xi - \tau \rangle \geq N^{-j} |a\xi - \tau| \geq N^{-j} \left(\frac{1}{2} N^j - \frac{1}{4} N^j \right) = \frac{1}{4},$$

so that $N^j \langle \tau - a\xi \rangle^{-1} \lesssim 1$. If on the other hand $|\tau| \geq \frac{1}{4} N^j$, then

$$N^j \langle a\xi - \tau \rangle^{-1} \leq N^j \lesssim |\tau|$$

and the claim again is proved.

The proof of (31) is similar, except that instead of L_M we use $L'_M = (1 + |t|^2)^{-M} (1 + D_\tau^2)^M$ with $L_M e^{2\pi i t \tau} = e^{2\pi i t \tau}$ and integrate by parts in τ. The details are omitted. $\qquad \square$

Corollary 2. For $f \in \mathcal{S}(\mathbf{R})$, we have the estimate

$$(35) \qquad \| F_{j,a} f \|_{L^p(dxdt)} \lesssim_N N^{j(\gamma - \alpha)} \| f \|_{L^p(dx)}.$$

Proof. We write $\| F_{j,a} f \|_{L^p(dxdt)} = \big\| \| F_{j,a} f(\cdot, t) \|_{L^p(dx)} \big\|_{L^p(dt)}$. Writing out $F_{j,a} f$ as in (28), we see that it suffices to prove an estimate of the form

$$\left\| \int f(y) K_{j,a}(x - y, t) \, dy \right\|_{L^p(dx)} \lesssim_N (1 + |t|)^{-2} \| f \|_p.$$

By Young's inequality, it suffices to prove that

$$\int |K_{j,a}(x, t)| \, dx \lesssim_N (1 + |t|)^{-2}.$$

But this is an easy consequence of (30) and (31). $\qquad \square$

In the next proposition, we let $\phi \in C_c^\infty([-2, 2])$ be a function such that $0 \leq \phi \leq 1$ and $\phi(\tau) = 1$ for $|\tau| \leq 1$. This can be the same function that we used to define the cut-offs in $|\xi|$. The small number $\epsilon > 0$ will be fixed later.

Proposition 2. *Let*

$$m_{j,a}^{\mathrm{main}}(\xi, \tau) = m_{j,a}(\xi, \tau) \phi(N^{-j\epsilon} \tau),$$

$$K_{j,a}^{\mathrm{main}}(x, t) = \iint e^{2\pi i((x - at)\xi + t\tau)} m_{j,a}^{\mathrm{main}}(\xi, \tau) \, d\xi \, d\tau,$$

$$F_{j,a}^{\mathrm{main}} f(x, t) = N^{j(\gamma - \alpha)} \int f(y) K_{j,a}^{\mathrm{main}}(x - y, t) \, dy, \quad f \in \mathcal{S}(\mathbf{R}).$$

Then for all $1 \le p \le \infty$ and $M'' \in \mathbf{N}$ we have

$$(36) \qquad \|F_{j,a}f - F_{j,a}^{\mathrm{main}}f\|_{L^p(\mathbf{R}^2)} \lesssim_{N,M''} N^{-j\epsilon M''}\|f\|_p,$$

where the implicit constant may depend on p, ϵ, N, and M'', but not on j.

Remark. In the original (independent of a) Fourier coordinates (ξ, s), we have

$$F_{j,a}^{\mathrm{main}}f = \mathcal{F}^{-1}\left[\tilde{m}_{j,a}^{\mathrm{main}}(\xi, s)\widehat{f}(\xi)\right],$$

where $\tilde{m}_{j,a}^{\mathrm{main}}(\xi, s) = N^{j(\gamma-\alpha)}m_{j,a}^{\mathrm{main}}(\xi, s + a\xi)$ is supported in the set

$$\mathcal{K}_{j,a}^{\epsilon} = \left\{(\xi, s) \in \mathbf{R}^2 : \frac{1}{2}N^j \le |\xi| \le 4N^{j+2}, \ |s + a\xi| \le 2N^{j\epsilon}\right\}.$$

Proof. The remark after the proposition follows immediately upon changing coordinates. We now prove the proposition, We have

$$F_{j,a}f - F_{j,a}^{\mathrm{main}}f(x, t) = N^{j(\gamma-\alpha)}\int f(y)(K_{j,a} - K_{j,a}^{\mathrm{main}})(x - y, t)\, dy.$$

As in Corollary 2, it suffices to prove that

$$\int |(K_{j,a} - K_{j,a}^{\mathrm{main}})(x, t)|\, dx \lesssim_{N,M''} N^{-j\epsilon M''}(1 + |t|)^{-2}.$$

This in turn follows from estimates analogous to (30) and (31), namely

$$(37) \qquad \left|(K_{j,a} - K_{j,a}^{\mathrm{main}})(x, t)\right| \le C_{M,M'',N}N^{-j\epsilon M''}\left(1 + N^{2j}|x - ta|^2 + |t|^2\right)^{-M}.$$

with $C_{M,M'',N}$ independent of j and a. To prove this, we proceed as in the proof of (30) and (31), with the following modifications. We have

$$(K_{j,a} - K_{j,a}^{\mathrm{main}})(x, t) = \iint e^{2\pi i((x-at)\xi + t\tau)}\left[m_{j,a}(\xi, \tau) - m_{j,a}^{\mathrm{main}}(\xi, \tau)\right]d\xi\, d\tau,$$

with

$$m_{j,a}(\xi, \tau) - m_{j,a}^{\mathrm{main}}(\xi, \tau) = \frac{\langle a\xi - \tau\rangle^{\gamma}}{N^{j\gamma}}|\lambda_j(\xi, \tau)|\left(1 - \phi(N^{-j\epsilon}\tau)\right).$$

We integrate by parts as in the proof of (30) and (31), but also use that $1 - \phi(N^{-j\epsilon}\tau)$ is supported in $|\tau| \ge N^{j\epsilon}$, so that we can separate out factors $(1 + |\tau|)^{-M''} \lesssim N^{-j\epsilon M''}$ from the estimate (32) and from the analogous estimate for D_τ before proceeding with the rest of the argument. $\square$

4. Decoupling for the Cantor bush

4.1. Preliminaries. We will need to develop decoupling inequalities for functions with Fourier support contained in a neighbourhood of the Cantor bush $\bigcup_{a \in A_j}\mathcal{K}_{j,a}^{\epsilon}$, with $\mathcal{K}_{j,a}^{\epsilon}$ defined in the remark after Proposition 2. We use parts of the decoupling machinery developed by Bourgain and Demeter [4], [5]. The notation below will follow the conventions of [5], with minor modifications. We will also rely on a 1-dimensional Cantor decoupling inequality proved in [18].

For $L > 0$, an *L-interval* in $\mathbf{R}$ will be an interval of length L with endpoints in $L\mathbf{Z}$. If a coordinate system in $\mathbf{R}^2$ is given, an *L-square* will be a 2-dimensional square of side length L, with vertices in $L\mathbf{Z}^2$ and sides parallel to the coordinate axes. We will often use $L = N^k$ with $k \in \mathbf{Z}$; in that case, any N^k-square Q and any $N^{k'}$-square Q' in the same coordinate system are either nested or disjoint except possibly for an edge or vertex. Unless stated otherwise, we will assume all L-squares to be closed.

Note that the definition above relies on a fixed choice of a coordinate system. In the inductive arguments below, we will use many coordinate systems corresponding to different portions of the Cantor set. We will say that two such coordinate systems are *compatible* if a 1-square in one coordinate system can be covered by $O(1)$ 1-squares in the other coordinate system, and vice versa, with the $O(1)$ constants independent of N, j, k.

We will use local weights in 1 and 2 dimensions. If R is the rectangle $\{(x, t): |x - x_0| \leq r_x, \ |t - t_0| \leq r_t\}$, we define

$$(38) \qquad w_R(x, t) = \left(1 + \sqrt{\left(\frac{x - x_0}{r_x}\right)^2 + \left(\frac{t - t_0}{r_t}\right)^2}\right)^{-100}$$

and, for a locally integrable function $g \colon \mathbf{R}^2 \to \mathbf{C}$,

$$\|g\|_{L^p(w_R)} = \left(\int |g|^p w_R\right)^{1/p}.$$

In dimension 1, if I is the interval $x_0 - r \leq x \leq x_0 + r$, we define

$$w_I(x) = \left(1 + \frac{|x - x_0|}{r}\right)^{-1000},$$

and $\|g\|_{L^p(w_I)}$ is defined similarly. Typically, R and I will be N^k-squares and intervals. We will use Q, R, S for squares and I, J for intervals; this will also indicate whether the associated weight w is taken in 1 or 2 dimensions.

We will use repeatedly the following covering argument (cf. [5, Lemma 4.1]).

Lemma 5. *Let $\mathbf{R}^2 = \bigcup_{Q \in \mathcal{Q}} Q$ be a covering of the plane by L-squares associated with some coordinate system. Then we have the following estimates, with the implicit constants independent of L.*

(a) $\sum_{Q \in \mathcal{Q}} w_Q \sim 1$.
(b) $\min_{x \in Q} w_Q(x) \sim \max_{x \in Q} w_Q(x)$.
(c) *Let $\mathbf{R}^2 = \bigcup_{R \in \mathcal{R}} R$ be a covering of the plane by L'-squares in a possibly different but compatible coordinate system, with $L' \sim L$. Suppose that $\{g_i\}_{i \in \mathcal{I}}$ is a finite family of functions such that for $g = \sum g_i$, and for every $R \in \mathcal{R}$, we have the estimate*

$$\|g\|_{L^p(R)} \leq K \left(\sum_i \|g_i\|_{L^p(w_R)}^2\right)^{1/2}.$$

Then we also have

$$\|g\|_{L^p(w_Q)} \lesssim K \left(\sum_i \|g_i\|_{L^p(w_Q)}^2\right)^{1/2}$$

for all $Q \in \mathcal{Q}$.

Proof. Parts (a) and (b) are clear from the definition of w_Q. We now prove (c). For a given Q, let $c_R = \max_{x \in R} w_Q(x)$. We claim that

$$(39) \qquad w_Q \leq \sum_R c_R \mathbf{1}_R \lesssim \sum_R c_R w_R \lesssim w_Q.$$

The first two inequalities in (39) are clear; we need to verify the last one. Using (b) and then (a), we have

$$\sum_R c_R w_R(x) = \sum_R \max_{y \in R} w_Q(y) \cdot w_R(x) \lesssim \sum_R w_Q(x) w_R(x) \lesssim w_Q(x),$$

as required.

With c_R as above, we write

$$\|g\|^2_{L^p(w_Q)} = \left[\int |g|^p w_Q\right]^{2/p} \lesssim \left[\sum_R c_R \int_R |g|^p\right]^{2/p} = \left[\sum_R c_R \|g\|^p_{L^p(R)}\right]^{2/p}$$

$$\lesssim K^2 \left[\sum_R \left(c_R^{2/p} \sum_i \|g_i\|^2_{L^p(w_R)}\right)^{p/2}\right]^{2/p}$$

$$\lesssim K^2 \sum_i \left[\sum_R \left(c_R^{2/p} \|g_i\|^2_{L^p(w_R)}\right)^{p/2}\right]^{2/p},$$

where at the last step we used Minkowski's inequality in $\ell^{p/2}(\mathcal{R})$. Continuing the calculation and using (39) at the end, we get

$$\|g\|^2_{L^p(w_Q)} \lesssim K^2 \sum_i \left[\sum_R c_R \|g_i\|^p_{L^p(w_R)}\right]^{2/p} = K^2 \sum_i \left[\int \sum_R |g_i|^p c_R w_R\right]^{2/p}$$

$$\lesssim K^2 \sum_i \left[\int |g_i|^p w_Q\right]^{2/p} = K^2 \sum_i \|g_i\|^2_{L^p(w_Q)}$$

as claimed. $\qquad\square$

4.2. Decoupling for Cantor strips. Let $S \subset [N-1]$ be a $\Lambda(p)$ set obeying (2) and (3), with N sufficiently large to be determined later. Our basic tool, borrowed from [18, Lemma 5], is the following single-scale decoupling inequality which follows from Bourgain's $\Lambda(p)$ estimate. We will need a slightly modified version with intervals of length 2 instead of 1. This is easy to arrange using a partition of unity, cf. the remark before Lemma 5 in [18].

Lemma 6. *With S as above, let $E = S + [0,2]$, and let $h\colon \mathbf{R} \to \mathbf{C}$ be a locally integrable function with $\widehat{h}$ supported on E. Let $h = \sum_{a\in S} h_a$, where $\widehat{h}_a$ is supported on $a + [0,2]$. Then for $2 \le p \le p_0$ we have*

$$(40) \qquad \|h\|^2_{L^p(w_I)} \le C_1^2 \sum_{a\in S} \|h_a\|^2_{L^p(w_I)}$$

for any 1-interval I, where C_1 depends on p but not on N or h.

We need to extend the estimate (40) to 2-dimensional product sets consisting of parallel strips corresponding to the Cantor intervals.

Lemma 7. *With $E \subset \mathbf{R}$ defined above and $L_1 < L_2$, let $h\colon \mathbf{R}^2 \to \mathbf{C}$ be a locally integrable function such that $\widehat{h}$ is supported on $E \times [L_1, L_2]$. Assume that $h = \sum_{a\in S} h_a$, where $\widehat{h}_a$ is supported on $[a, a+2] \times [L_1, L_2]$. Then*

$$(41) \qquad \|h\|^2_{L^p(w_Q)} \lesssim C_1^2 \sum_{a\in S} \|h_a\|^2_{L^p(w_Q)}$$

for any 1-square Q.

Proof. We will prove that for any $t_1 < t_2$, and for any 1-interval I, we have

$$(42) \quad \int_{t_1}^{t_2} \int |h(x,t)|^p w_I(x)\, dx\, dt \le C_1^2 \left[\sum_{a\in S} \left(\int_{t_1}^{t_2} \int |h_a(x,t)|^p w_I(x)\, dx\, dt\right)^{2/p}\right]^{1/2}.$$

Then (41) follows from Lemma 5 (c).

Consider the function $h^{(t)}(x) := h(x, t)$ as a function of x, with t fixed. We have

$$h(x, t) = \int \Big[\int_{L_1}^{L_2} \widehat{h}(\xi, s) e^{2\pi i t s} \, ds \Big] e^{2\pi i \xi x} \, d\xi,$$

hence $\widehat{h^{(t)}}(\xi) = \int_{L_1}^{L_2} \widehat{h}(\xi, s) e^{2\pi i t s} \, ds$ is supported on E, and satisfies the assumptions of Lemma 6 with $(h^{(t)})_a(x) = h_a(x, t)$. By (40), we have

(43) $$\|h(x, t)\|_{L^p(w_I(x))}^2 \le C^2 \sum_{a \in S} \|h_a(x, t)\|_{L^p(w_I(x))}^2.$$

Let $H_a(t) = \|h_a(x, t)\|_{L^p(w_I(x))}^2$, then by (43) and Minkowski's inequality we have

$$\int_{t_1}^{t_2} \int |h(x, t)|^p w_I(x) \, dx \, dt = \int_{t_1}^{t_2} \|h(x, t)\|_{L^p(w_I(x))}^p \, dx \, dt$$

$$\le C_1^p \int_{t_1}^{t_2} \Big[\sum_a H_a(t) \Big]^{p/2} \, dx \, dt$$

$$= C_1^p \Big\| \sum_a H_a(t) \Big\|_{L^{p/2}([t_1, t_2])}^{p/2}$$

$$\le C_1^p \Big[\sum_a \|H_a(t)\|_{L^{p/2}([t_1, t_2])} \Big]^{p/2}$$

$$= C_1^p \Big[\sum_a \Big(\int_{t_1}^{t_2} \|h_a(x, t)\|_{L^p(w_I(x))}^p \, dt \Big)^{2/p} \Big]^{p/2},$$

which proves (42). $\qquad \qquad \square$

4.3. Local coordinates adjusted to the Cantor bush. The key geometric observation is that, for each $a \in A_k$, the Cantor branches corresponding to $a' \in A_{k+1,a}$ in the next iteration can be treated as parallel when restricted to segments of somewhat shorter length. More precisely, the corresponding part of the Cantor bush can be covered efficiently by a rescaled and rotated copy of the set $\mathcal{E}$ from Lemma 7.

Lemma 8. *Let $a \sim 1$. Define new coordinate systems (u, v) on $\mathbf{R}_{x,t}^2$ and (η, τ) on $\mathbf{R}_{\xi,s}^2$:*

(44) $$u = \frac{x - at}{1 + a^2}, \quad v = \frac{ax + t}{1 + a^2},$$
$$\eta = \xi - as, \quad \tau = a\xi + s.$$

Then:

(a) *We have $(u, v)^T = \mathbf{A}(x, y)^T$ and $(\eta, \tau)^T = (\mathbf{A}^T)^{-1}(\xi, s)^T$, where $\mathbf{A}$ is an orthogonal matrix. Hence the coordinate systems (u, v) and (η, τ) are orthogonal and dual to each other.*

(b) *Let $0 < \xi_1 < \xi_2$, $0 < \Delta a$, and $S \subset [N - 1]$. Assume that*

(45) $$\xi_2 - \xi_1 \le \frac{\xi_1}{N}.$$

Then the set

$$\mathcal{K} = \mathcal{K}[a, \Delta a, \xi_1, \xi_2] = \bigcup_{b \in S} \mathcal{K}_b,$$

where

$$\mathcal{K}_b := \left\{ (\xi, s) \colon \xi_1 \leq \xi \leq \xi_2, \; -\frac{s}{\xi} \in a + \frac{\Delta a}{N}\big(b + [0,1]\big) \right\}$$

is contained in

$$\mathcal{E} := \left\{ (\eta, \tau) \colon \xi_1(1 + a^2) \leq \eta \leq \xi_2(1 + a^2 + a\Delta a), \; \tau \in -\xi_1\frac{\Delta a}{N}(S + [0,2]) \right\},$$

with the individual branches $\mathcal{K}_b$ contained in the corresponding strips $\mathcal{E}_b := \{(\eta, \tau) \in \mathcal{E} \colon \tau \in -\xi_1\frac{\Delta a}{N}(b + [0,2])\}$ of $\mathcal{E}$.

Proof. Part (a) is easily verified by direct calculation. We now turn to (b). We need to prove the following: for

$$(46) \qquad b = \frac{\Delta a}{N}(b_0 + \Delta b), \quad b_0 \in S, \quad 0 \leq \Delta b \leq 1,$$

the line segment $J_b = \{(\xi, s) \colon \xi_1 \leq \xi \leq \xi_2, \; \frac{s}{\xi} + a + b = 0\}$ is contained in the set

$$(47) \qquad \left\{ (\eta, \tau) \colon \xi_1(1 + a^2) \leq \eta \leq \xi_2(1 + a^2 + a\Delta a), \; \tau \in -\xi_1\frac{\Delta a}{N}(b_0 + [0,2]) \right\}$$

Since the set (47) is convex, it suffices to prove this for the endpoints of J_b. A very short calculation shows that these are given by (η_i, τ_i), $i = 1, 2$, where

$$\eta_i = \xi_i(1 + a^2 + ab), \quad \tau_i = -\xi_i b.$$

For b as in (46), we have $0 \leq b \leq \Delta a$. This clearly implies that η_1, η_2 satisfy the constraint in (47). Next, we have

$$-\tau_1 = \xi_1 b \in \xi_1\frac{\Delta a}{N}(b_0 + [0,1]).$$

Finally, we write

$$-\tau_2 = \xi_2 b = \xi_1 b + (\xi_2 - \xi_1)b \in \xi_1\frac{\Delta a}{N}(b_0 + [0,1]) + (\xi_2 - \xi_1)b,$$

and by (45),

$$0 < (\xi_2 - \xi_1)b \leq \xi_1\frac{\Delta a}{N},$$

which completes the proof. $\qquad\square$

Corollary 3. *Let $\mathcal{K} = \bigcup_{b \in S} \mathcal{K}_b$ be as in Lemma 8. For a locally integrable function $g \colon \mathbf{R}^2_{x,t} \to \mathbf{C}$ with $\widehat{g}$ is supported on $\mathcal{K}$, let $\widehat{g}_b = \mathbf{1}_{\mathcal{K}_b}\widehat{g}$. Then*

$$(48) \qquad \|g\|^2_{L^p(w_Q)} \lesssim \sum_{b \in S} \|g_b\|^2_{L^p(w_Q)}$$

for any L-square Q with $L = N(\xi_1 \Delta a)^{-1}$.

Proof. We change the coordinates as in Lemma 8. The function $h(\eta, \tau) = g(x(\eta, \tau), t(\eta, \tau))$ is Fourier supported in $\mathcal{E}$ and satisfies the assumptions of the lemma, with $h_b(\eta, \tau) = g_b(x(\eta, \tau), t(\eta, \tau))$ for $b \in S$. Note that the set $\mathcal{E}$ is a rescaled and reflected copy of the set $E \times [L_1, L_2]$ (with appropriate L_1, L_2) from Lemma 7. Applying (41) to a scaled copy of h and then undoing the scaling and the coordinate change, we get that (48) holds with Q replaced by any L-square in the (η, τ) coordinates. To pass to L-squares in the (x, t) coordinates, we use Lemma 5. $\qquad\square$

4.4. The inductive argument for the Cantor bush.

Lemma 9. *Let $\mathcal{K}_j^\epsilon = \bigcup_{a \in A_j} \mathcal{K}_{j,a}^\epsilon$, where*

$$\mathcal{K}_{j,a}^\epsilon = \left\{ (\xi, s) \in \mathbf{R}^2 \colon \frac{1}{2} N^j \leq |\xi| \leq 4N^{j+2}, \quad |s + a\xi| \leq 2N^{j\epsilon} \right\}.$$

Then $\mathcal{K}_j^\epsilon$ can be covered by $O(N^{-2+j\epsilon})$ finitely overlapping sets of the form $\tilde{\mathcal{K}}_j = \bigcup_{a \in A_j} \tilde{\mathcal{K}}_{j,a}$, where

$$\tilde{\mathcal{K}}_{j,a} = \tilde{\mathcal{K}}_{j,a}[a_0, \xi_1, \xi_2] = \left\{ (\xi, s) \in \mathbf{R}^2 \colon \xi_1 \leq |\xi| \leq \xi_2, \quad -\frac{s}{\xi} \in a_0 + a + [0, N^{-j}] \right\}$$

with $\frac{1}{4} N^j \leq \xi_1 \leq 4N^{j+2}$, $\frac{\xi_1}{2N} \leq \xi_2 - \xi_1 \leq \frac{\xi_1}{N}$, $a_0 \in \frac{1}{2} N^{-j} \mathbf{Z}$ and $|a_0| = O(N^{-j+j\epsilon})$.

Furthermore, if $G_a \in L^p(\mathbf{R}^2)$ is Fourier-supported in $\mathcal{K}_{j,a}^\epsilon$, then there is a decomposition

(49)
$$G_a = \sum_{i \in \mathcal{I}} g_a^{(i)},$$

where the summation runs over a set $\mathcal{I}$ of cardinality $O(N^{-2+j\epsilon})$, each $g_a^{(i)}$ is Fourier supported in some set $\tilde{\mathcal{K}}_{j,a}[a_0, \xi_1, \xi_2]$ as above, and $\|g_a^{(i)}\|_p \lesssim_N \|G_a\|_p$ with the constant independent of j.

Proof. We first cover the strip $\frac{1}{2} N^j \leq |\xi| \leq 4N^{j+2}$ by $O(N^2)$ strips $\xi_1 \leq |\xi| \leq \xi_2$ as indicated. Suppose now that $(\xi, s) \in \mathcal{K}_{j,a}^\epsilon$ for some $a \in A_j$. Then

$$\left| \frac{s}{\xi} + a \right| \leq \frac{2N^{j\epsilon}}{|\xi|} \lesssim N^{-j+j\epsilon},$$

so that $-\frac{s}{\xi} \in a_0 + a + [0, N^{-j}] \subset a_0 + E_j$ for some a_0 as in the statement of the lemma.

For the second claim, allowing overlaps in the covering of $\mathcal{K}_{j,a}^\epsilon$, we can associate with it a smooth partition of unity $\{\Xi_a^{(i)}\}_{i \in \mathcal{I}}$ such that $\sum_{i \in \mathcal{I}} \Xi_a^{(i)} = 1$ on $\mathcal{K}_{j,a}^\epsilon$ and $\|(\Xi_a^{(i)})^\vee\|_1 = O_N(1)$ uniformly in j and a. Then the functions $g_a^{(i)} = G_a * (\Xi_a^{(i)})^\vee$ satisfy the desired conclusions, with the L^p estimate following from Young's inequality. $\square$

Proposition 3. *Let $\tilde{\mathcal{K}}_j[a_0, \xi_1, \xi_2]$ be as in Lemma 9. For a function $g \colon \mathbf{R}^2 \to \mathbf{C}$ with $\text{supp} \, \widehat{g} \subset \tilde{\mathcal{K}}_j[a_0, \xi_1, \xi_2]$, and for $k = 1, \ldots, j$, write $g = \sum_{a \in A_k} g_{k,a}$ with $\widehat{g_{k,a}}$ supported in the set*

(50)
$$\left\{ (\xi, s) \in \mathbf{R}^2 \colon \xi_1 \leq \xi \leq \xi_2, \quad -\frac{s}{\xi} \in a_0 + a + [0, N^{-k}] \right\}.$$

(Note that this defines $g_{k,a}$ uniquely since the sets above are disjoint for different $a \in A_k$.) Then there is a constant C_2 (independent of N, k, j) such that for any $N^k L$-square Q_k with $L = \xi_1^{-1}$, we have

(51)
$$\left(\sum_{S \in \mathcal{T}(Q_k)} \|g\|_{L^p(w_S)}^p \right)^{1/p} \leq C_2^k \left(\sum_{a \in A_k} \|g_{k,a}\|_{L^p(w_{Q_k})}^2 \right)^{1/2},$$

where $Q_k = \bigcup_{S \in \mathcal{T}(Q_k)} S$ is a tiling of Q_k by L-squares.

Proof. The proof is similar to the proof of Proposition 1 in [18]. We proceed by induction in k, using Corollary 3 at each step. To initialize, we write using the notation from Lemma 8

$$\tilde{\mathcal{K}}_j[a_0, \xi_1, \xi_2] \subset \mathcal{K}[a_0, 1, \xi_1, \xi_2].$$

Applying Corollary 3, we get that

$$(52) \qquad \|g\|^2_{L^p(w_{Q_1})} \leq C_3 \sum_{a \in A_1} \|g_{1,a}\|^2_{L^p(w_{Q_1})},$$

with Q_1 as above and some constant C_3 independent of N, j. Similarly, for $1 \leq k \leq j-1$ and $a \in A_k$, the set (50) is contained in $\mathcal{K}[a_0 + a, N^{-k}, \xi_1, \xi_2]$. Applying Corollary 3 again, we get

$$(53) \qquad \|g_{k,a}\|^2_{L^p(w_{Q_{k+1}})} \leq C_3 \sum_{b \in A_{k+1,a}} \|g_{k+1,b}\|^2_{L^p(w_{Q_{k+1}})}.$$

To put the inductive argument together, we use parallel decoupling. Let C_4 be a constant such that for all $N^k L$ squares Q_k with $k \geq 1$,

$$(54) \qquad \sum_{S \in \mathcal{T}(Q_k)} w_S \leq C_4 w_{Q_k}.$$

By the rapid decay of w, we can choose C_4 independent of N and k.

We first note that (51) for $k = 1$ is provided by (52), with the trivial tiling consisting of a single square. Assume now that we have (51) for some k with $1 \leq k \leq j-1$. Let Q_{k+1} be an $N^{k+1}L$-square, and let $Q_{k+1} = \bigcup_{Q_k \in \mathcal{J}} Q_k$ be a tiling of Q_{k+1} by $N^k L$-squares. By the inductive assumption (51), Minkowski's inequality in $\ell^{p/2}(\mathcal{J})$, a rescaling of (54), and (53), in that order, we have

$$
\begin{aligned}
(55) \quad \sum_{S \in \mathcal{I}(Q_{k+1})} \|g\|^p_{L^p(w_S)} &= \sum_{Q_k \in \mathcal{J}} \sum_{S \in \mathcal{T}(Q_k)} \|g\|^p_{L^p(w_S)} \\
&\leq C_2^{kp} \sum_{Q_k \in \mathcal{J}} \left(\sum_{a \in A_k} \|g_{k,a}\|^2_{L^p(w_{Q_k})} \right)^{p/2} \\
&\leq C_2^{kp} \left[\sum_{a \in A_k} \left(\sum_{Q_k \in \mathcal{J}} \|g_{k,a}\|^p_{L^p(w_{Q_k})} \right)^{2/p} \right]^{p/2} \\
&\leq C_2^{kp} C_4 \left[\sum_{a \in A_k} \|g_{k,a}\|^2_{L^p(w_{Q_{k+1}})} \right]^{p/2} \\
&\leq C_2^{kp} C_4 C_3^p \left[\sum_{a \in A_{k+1}} \|f_{k+1,a}\|^2_{L^p(w_{Q_{k+1}})} \right]^{p/2}.
\end{aligned}
$$

This proves (51) with $C_2 = C_3^{1/2} C_4^{1/p}$. $\qquad\square$

Corollary 4. *Let $g = \sum_{a \in A_j} g_{j,a}$ be as in Proposition 3. Then*

$$(56) \qquad \|g\|_{L^p(\mathbf{R}^2)} \lesssim C_2^j \left(\sum_{a \in A_j} \|g_{j,a}\|^2_{L^p(\mathbf{R}^2)} \right)^{1/2}.$$

Proof. We use the notation from the proof of Proposition 3. Note first that by (51), we have

$$(57) \qquad \|g\|_{L^p(Q_j)} \lesssim \Big(\sum_{S \in \mathcal{T}(Q_j)} \|g\|^p_{L^p(w_S)} \Big)^{1/p} \le C_2^j \Big(\sum_{a \in A_j} \|g_{j,a}\|^2_{L^p(w_{Q_j})} \Big)^{1/2}.$$

Let now Q be an MN^jL-square, and let $Q = \bigcup_{Q_j \in \mathcal{J}} Q_j$ be a tiling of Q by N^jL-squares. Using (57) and Minkowski's inequality in $\ell^{p/2}(\mathcal{J})$ as in (55), we get

$$\|g\|^p_{L^p(Q)} = \sum_{Q_j \in \mathcal{J}} \|g\|^p_{L^p(Q_j)} \lesssim C_2^{jp} \sum_{Q_j \in \mathcal{J}} \Big(\sum_{a \in A_j} \|g_{j,a}\|^2_{L^p(w_{Q_j})} \Big)^{p/2}$$

$$\le C_2^{jp} \Bigg[\sum_{a \in A_j} \Big(\sum_{Q_j \in \mathcal{J}} \|g_{j,a}\|^p_{L^p(w_{Q_j})} \Big)^{2/p} \Bigg]^{p/2} \lesssim C_2^{jp} \Bigg[\sum_{a \in A_j} \|g_{j,a}\|^2_{L^p(\mathbf{R})} \Bigg]^{p/2},$$

where the last step follows from Lemma 5 (a). Since this holds for any Q with the constant independent of M and Q, we have proved (56). □

Corollary 5. *Let $G = \sum_{a \in A_j} G_a$, where each $G_a \in L^p(\mathbf{R}^2)$ is Fourier-supported in $\mathcal{K}^\epsilon_{j,a}$. Then*

$$(58) \qquad \|G\|_{L^p(\mathbf{R}^2)} \lesssim_N C_2^j N^{j\epsilon} \Big(\sum_{a \in A_j} \|G_a\|^2_{L^p(\mathbf{R}^2)} \Big)^{1/2}.$$

Proof. This follows by using the decomposition in Lemma 9 and then applying Corollary 4 to each $g^{(i)} = \sum_{a \in A_j} g^{(i)}_a$. □

5. Proof of Theorems 2 and 3

5.1. Proof of Theorem 2. Given $\delta > 0$, choose $\epsilon > 0$ small enough so that $5\epsilon < \delta$. Let N be sufficiently large so that $\alpha < \alpha'$ and $\frac{1}{p} - \frac{\alpha}{2} < \epsilon$. This is possible by (5). Throughout the proof, we may increase N further as needed without changing the other parameters of construction.

By Lemma 1, it suffices to prove the single scale maximal estimate (18) with $\beta = 4\epsilon$, i.e.

$$(59) \qquad \|\tilde{\mathcal{M}}f\|_p \le C_N N^{4j\epsilon} \|f\|_p, \quad j \ge j_0,$$

for all $f \in \mathcal{S}$ with $\mathrm{supp}\,\widehat{f} \subset \{N^j \le |\xi| \le 2N^{j+1}\}$, with the constant C_N independent of j. By Lemma 2, this will follow if we can prove that the operators F_j with $\gamma = (1/p) + \epsilon$ obey

$$(60) \qquad \|F_j f\|_{L^p(dxdt)} \lesssim_N N^{4j\epsilon} \|f\|_{L^p(dx)}, \quad f \in \mathcal{S}.$$

We start with the decomposition $F_j f = \sum_{a \in A_j} F_{j,a} f$ as in Section 3, and an application of Proposition 2. By (36) with M large enough so that $\epsilon M > 2$, we have

$$(61) \qquad \|F_j f\|_p = \Big\| \sum_{a \in A_j} F_{j,a} f \Big\|_p \le \Big\| \sum_{a \in A_j} F^{\mathrm{main}}_{j,a} f \Big\|_p + O_N(N^{-j}) \|f\|_p$$

where each $F^{\mathrm{main}}_{j,a} f$ is Fourier supported in $\mathcal{K}^\epsilon_{j,a}$. Applying Corollary 5 with $G_a = F^{\mathrm{main}}_{j,a} f$, we get

$$\Big\| \sum_{a \in A_j} F^{\mathrm{main}}_{j,a} f \Big\|_p \lesssim_N N^{2j\epsilon} \Big(\sum_{a \in A_j} \|F^{\mathrm{main}}_{j,a} f\|^2_p \Big)^{1/2}.$$

Using (36) again, and then (35), we conclude that

$$\|F_j f\|_p \lesssim_N N^{2j\epsilon} \Big(\sum_{a \in A_j} \|F_{j,a} f\|_p^2 \Big)^{1/2} + O(N^{-j}) \|f\|_p$$

(62)
$$\lesssim_N N^{2j\epsilon} \Big(\sum_{a \in A_j} N^{j(\gamma - \alpha)} \|f\|_p^2 \Big)^{1/2} + O(N^{-j}) \|f\|_p$$

$$\lesssim_N N^{j(\gamma - \frac{\alpha}{2} + 2\epsilon)} \|f\|_p.$$

We have

$$\gamma - \frac{\alpha}{2} + 2\epsilon = \frac{1}{p} + \epsilon - \frac{\alpha}{2} + 2\epsilon \le 4\epsilon,$$

so that (60) holds as claimed. This ends the proof of the theorem.

5.2. Proof of Theorem 3. This is a minor modification of the above. To prove (10), define F_j as in the proof of Theorem 2, but now use $\gamma = \frac{1}{p} - \frac{1}{r} + \epsilon$ instead. Then the exponent at the end of the analogue of (62) is

$$\gamma - \frac{\alpha}{2} + 2\epsilon = \frac{1}{p} - \frac{1}{r} + \epsilon - \frac{\alpha}{2} + 2\epsilon \le 4\epsilon - \frac{1}{r},$$

which is negative if ϵ is small enough. The estimate (10) follows upon summing up in j and then applying the Sobolev embedding theorem as in Section 2.3.

The proof of (11) is similar, but with the factor $\langle s \rangle^\gamma$ in the definition of F_j replaced by $\langle \xi \rangle^\gamma$. Sobolev's embedding theorem is not needed for this part.

References

[1] BOSE, R. C.: An affine analogue of Singer's theorem. - J. Indian Math. Soc. 6, 1942, 1–15.

[2] BOURGAIN, J.: Averages in the plane over convex curves and maximal operators. - J. Anal. Math. 47, 1986, 69–85.

[3] BOURGAIN, J.: Bounded orthogonal systems and the $\Lambda(p)$-set problem. - Acta Math. 162, 1989, 227–245.

[4] BOURGAIN, J., and C. DEMETER: The proof of the ℓ^2 decoupling conjecture. - Ann. of Math. (2) 182, 2015, 351–389.

[5] BOURGAIN, J., and C. DEMETER: A study guide for the ℓ^2 decoupling theorem. - Chin. Ann. Math. Ser. B 38:1, 2017, 173–200.

[6] DEMETER, C.: A decoupling for Cantor-like sets. - Proc. Amer. Math. Soc. 147, 2019, 1037–1050.

[7] DUOANDIKOETXEA, J.: Fourier analysis. - Grad. Stud. Math. 29, Amer. Math. Soc., 2001.

[8] HAMBROOK, K., and I. ŁABA: On the sharpness of Mockenhaupt's restriction theorem. - Geom. Funct. Anal. 23:4, 2013, 1262–1277.

[9] HAMBROOK, K., and I. ŁABA: Sharpness of the Mockenhaupt–Mitsis–Bak–Seeger restriction theorem in higher dimensions. - Bull. London Math. Soc. 48, 2016, 757–770.

[10] HOCHMAN, M.: On self-similar sets with overlaps and inverse theorems for entropy. - Ann. of Math. (2) 180, 2014, 773–822.

[11] HOCHMAN, M.: Some problems on the boundary of fractal geometry and additive combinatorics. - In: Recent Developments in Fractals and Related Fields, J. Barral and S. Seuret (eds.), Springer, 2017, 129–174.

[12] IOSEVICH, A., B. KRAUSE, E. SAWYER, K. TAYLOR, and I. URIARTE-TUERO: Maximal operators: scales, curvature and the fractal dimension. - Anal. Math. (to appear).

[13] IOSEVICH, A., and E. SAWYER: Three problems motivated by the average decay of the Fourier transform. - In: Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), Contemp. Math. 320, Amer. Math. Soc., Providence, RI, 2003, 205–215.

[14] KELETI, T.: Small union with large set of centers. - In: Recent Developments in Fractals and Related Fields, Conference on Fractals and Related Fields III, Ile de Porquerolles, France, 2015 (to appear).

[15] KELETI, T., D. NAGY, and P. SHMERKIN: Squares and their centers. - J. Anal. Math. 134, 2018, 643–669.

[16] ŁABA, I., and M. PRAMANIK: Arithmetic progressions in sets of fractional dimension. - Geom. Funct. Anal. 19:2, 2009, 429–456.

[17] ŁABA, I., and M. PRAMANIK: Maximal theorems and differentiation theorems for sparse sets. - Duke Math. J. 158, 2011, 347–411.

[18] ŁABA, I., and H. WANG: Decoupling and near-optimal restriction estimates for Cantor sets. - Int. Math. Res. Not. IMRN 9, 2018, 2944–2966.

[19] MOCKENHAUPT, G., A. SEEGER, and C. SOGGE: Wave front sets, local smoothing and Bourgain's circular maximal theorem. - Ann. of Math. 136, 1982, 207–218.

[20] OLIVO, A., and P. SHMERKIN: Maximal operators for cube skeletons. - Ann. Acad. Sci. Fenn. Math. 45, 2020, 467–478.

[21] PRAMANIK, M., and A. SEEGER: Averages over curves in $\mathbf{R}^3$ and associated maximal functions. - Amer. J. Math. 129, 2007, 61–103.

[22] RUBIO DE FRANCIA, J. L.: Maximal functions and Fourier transforms. - Duke Math. J. 53, 1986, 395–404.

[23] RUZSA, I. Z.: Solving a linear equation in a set of integers. I. - Acta Arith. 65, 1993, 259–282.

[24] SCHLAG, W.: $L^p \to L^q$ estimates for the circular maximal function. - Ph.D. thesis, California Institute of Technology, 1996.

[25] SHMERKIN, P., and V. SUOMALA: Spatially independent martingales, intersections, and applications. - Mem. Amer. Math. Soc. 251:1195, 2018.

[26] SHMERKIN, P., and V. SUOMALA: A class of random Cantor measures, with applications. - In: Recent Developments in Fractals and Related Fields, J. Barral and S. Seuret (eds.), Springer, 2017, 233–260.

[27] SHMERKIN, P., and V. SUOMALA: Private communication.

[28] SINGER, J.: A theorem in finite projective geometry and some applications to number theory. - Trans. Amer. Math. Soc. 43, 1938, 377–385.

[29] STEIN, E. M.: Maximal functions. I. Spherical means. - Proc. Nat. Acad. Sci. U.S.A. 73:7, 1976, 2174–2175.

[30] STEIN, E. M.: Harmonic analysis. - Princeton Univ. Press, 1993.

[31] TALAGRAND, M.: Sections of smooth convex bodies via majoring measures. - Acta Math. 175, 1995, 273–300.

[32] THORNTON, R.: Cubes and their centers. - Acta Math. Hungar. 152, 2017, 291–313.

[33] WOLFF, T.: Local smoothing estimates in L^p for large p. - Geom. Funct. Anal. 10, 2000, 1237–1288.

[34] WOLFF, T.: Lectures on harmonic analysis. Edited by I. Łaba and C. Shubin. - Amer. Math. Soc., Providence, R.I., 2003.