MAXIMAL OPERATORS AND DIFFERENTIATION THEOREMS FOR SPARSE SETS

IZABELLA ŁABA AND MALABIKA PRAMANIK

Abstract. We study maximal averages associated with singular measures on $\mathbb{R}$. Our main result is a construction of singular Cantor-type measures supported on sets of Hausdorff dimension $1-\epsilon$, $0 \le \epsilon < \frac{1}{3}$ for which the corresponding maximal operators are bounded on $L^p(\mathbb{R})$ for $p > (1+\epsilon)/(1-\epsilon)$. As a consequence, we are able to answer a question of Aversa and Preiss on density and differentiation theorems in one dimension. Our proof combines probabilistic techniques with the methods developed in multidimensional Euclidean harmonic analysis, in particular there are strong similarities to Bourgain's proof of the circular maximal theorem in two dimensions.

1. Introduction

1.1. **Maximal operators.** Let $\{S_k : k \ge 1\}$ be a decreasing sequence of subsets of $\mathbb{R}$. We define the maximal operator associated with this sequence by

$$\tilde{\mathcal{M}} f(x) := \sup_{r>0, k\ge 1} \frac{1}{|S_k|} \int_{S_k} |f(x+ry)| dy. \tag{1.1}$$

While the definition (1.1) is quite general, we will focus on cases where the sequence $\{S_k\}$ arises from a Cantor-type iteration, so that in particular each S_k is a union of finitely many intervals. We will further assume that $|S_k| \to 0$ as $k \to \infty$.

Under mild conditions on the Cantor iteration process, the densities $\phi_k = \frac{1}{|S_k|}\mathbf{1}_{S_k}$ converge weakly to a probability measure μ supported on the set $S = \bigcap_{k=1}^{\infty} S_k$. We then define the maximal operator with respect to μ:

$$\tilde{\mathfrak{M}} f(x) := \sup_{r>0} \int |f(x+ry)| \, d\mu(y). \tag{1.2}$$

We will be interested in the L^p mapping properties of $\tilde{\mathcal{M}}$. Since $\tilde{\mathfrak{M}}$ is clearly dominated by $\tilde{\mathcal{M}}$, similar estimates will follow for $\tilde{\mathfrak{M}}$ with the same range of exponents.

2000 *Mathematics Subject Classification.* 26A24, 26A99, 28A78, 42B25.

We will also be concerned with $L^p \to L^q$ maximal estimates with $p < q$. For this purpose, it is necessary to define the modified maximal operators

$$\tilde{\mathcal{M}}^a f(x) := \sup_{r>0,\ k\geq 1} r^a \int |f(x+ry)|\phi_k(y)dy \ , \tag{1.3}$$

$$\tilde{\mathfrak{M}}^a f(x) := \sup_{r>0} r^a \int |f(x+ry)|\, d\mu(y) \ , \tag{1.4}$$

where the exponent $a = \frac{1}{p} - \frac{1}{q}$ accounts for the appropriate scaling correction. Note that $\tilde{\mathcal{M}}^0 = \tilde{\mathcal{M}}$ and $\tilde{\mathfrak{M}}^0 = \tilde{\mathfrak{M}}$.

Finally, we will need the restricted maximal operators

$$\mathcal{M}f(x) := \sup_{1<r<2,k\geq 1} \frac{1}{|S_k|} \int_{S_k} |f(x+ry)|dy \ , \tag{1.5}$$

$$\mathfrak{M}f(x) := \sup_{1<r<2} \int |f(x+ry)|\, d\mu(y) \ , \tag{1.6}$$

where the range of the dilation factor r is limited to a single scale. These operators will play a critical role in the proofs of the unrestricted maximal estimates.

1.2. The main results.

Theorem 1.1. *There is a decreasing sequence of sets $S_k \subseteq [1,2]$ with the following properties:*

(a) each S_k is a disjoint union of finitely many intervals,
(b) $|S_k| \searrow 0$ as $k \to \infty$,
(c) the weak-$*$ limit μ of the densities $\mathbf{1}_{S_k}/|S_k|$ exists.
(d) The restricted maximal operators $\mathcal{M}$ and $\mathfrak{M}$ defined in (1.5) and (1.6) are bounded from $L^p[0,1]$ to $L^q(\mathbb{R})$ for any $p, q \in (1,\infty)$, and from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ for any $1 < p \leq q < \infty$.
(e) The unrestricted maximal operators $\tilde{\mathcal{M}}^a$ and $\tilde{\mathfrak{M}}^a$ defined in (1.1) and (1.2) are bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ whenever $1 < p \leq q < \infty$, with $a = \frac{1}{p} - \frac{1}{q}$. In particular, $\tilde{\mathcal{M}}$ and $\tilde{\mathfrak{M}}$ are bounded on $L^p(\mathbb{R})$ for $p > 1$.

As a corollary, we obtain a differentiation theorem for averages on S_k that answers a question of Aversa and Preiss [3] (see §1.3.3 for more details).

Theorem 1.2. *Let $\{S_k : k \geq 1\}$ be the sequence of sets given by Theorem 1.1, with the limiting measure μ. Then for every $f \in L^p(\mathbb{R})$ with $p \in (1,\infty)$ we have*

$$\lim_{r\to 0} \sup_k \left| \frac{1}{r|S_k|} \int_{x+rS_k} f(y)dy - f(x) \right| = 0 \textit{ for a.e. } x \in \mathbb{R}, \textit{ and} \tag{1.7}$$

$$\lim_{r\to 0} \left| \int f(x+ry)d\mu(y) - f(x) \right| = 0 \textit{ for a.e. } x \in \mathbb{R}. \tag{1.8}$$

The limiting set $S = \bigcap_{k=1}^{\infty} S_k$ constructed in our proof of Theorem 1.1 has Hausdorff dimension 1. However, we are also able to prove similar maximal estimates for sequences of sets whose limit has Hausdorff dimension $1-\epsilon$ with $\epsilon > 0$, provided that the range of exponents is adjusted accordingly.

Theorem 1.3. *For any $0 < \epsilon < \frac{1}{3}$, there is a decreasing sequence of sets $S_k \subset [1,2]$ obeying the conditions (a)–(c) of Theorem 1.1 and such that:*

(a) $S = \bigcap_{k=1}^{\infty} S_k$ has Hausdorff dimension $1-\epsilon$,

(b) The restricted maximal operators $\mathcal{M}$ and $\mathfrak{M}$ are bounded from $L^p[0,1]$ to $L^q(\mathbb{R})$ for any p, q such that

$$\frac{1+\epsilon}{1-\epsilon} < p < \infty \text{ and } 1 < q < \frac{1-\epsilon}{2\epsilon} p, \tag{1.9}$$

and from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ for any p, q such that $p \leq q$ and (1.9) holds.

(c) The unrestricted maximal operators $\tilde{\mathcal{M}}^a$ and $\tilde{\mathfrak{M}}^a$ are bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ with $a = \frac{1}{p} - \frac{1}{q}$ for any p, q such that $p \leq q$ and (1.9) holds. In particular, $\tilde{\mathcal{M}}$ and $\tilde{\mathfrak{M}}$ are bounded on $L^p(\mathbb{R})$ for $p > \frac{1+\epsilon}{1-\epsilon}$.

(d) The family of sets $\mathcal{S} = \{rS_k : k \geq 1\}$ and the measure μ differentiate $L^p(\mathbb{R})$ in the sense of (1.7) and (1.8) for all $p > \frac{1+\epsilon}{1-\epsilon}$.

Remarks.

1. It is possible to use the ideas of [24] to modify the construction of the sequence of sets S_k so that, in addition to all the conclusions of Theorems 1.1 and 1.3, the limiting set $S = \bigcap_{k=1}^{\infty} S_k$ is a *Salem set.* See §1.3.2 for the definitions and more details.
2. It may be of greater interest that the correlation condition (4.2) used to prove Theorems 1.1 and 1.3 already implies that S has positive *Fourier dimension*, provided that the ϵ in Theorem 1.3 is small enough ($\epsilon < \frac{1}{5}$ will suffice). We hope to address this issue at length in a subsequent paper.
3. An argument due to David Preiss, included here in Subsection 8.2, shows that Theorem 1.2 (hence also Theorem 1.1(e)) cannot hold with $p = 1$. On the other hand, we do not know whether the range of ϵ or the exponents p, q in Theorem 1.3 is optimal.

1.3. **Motivation.** The motivation for the study of the maximal operators introduced in this article comes from two different directions. On the one hand, our maximal operators provide a one-dimensional analogue of higher dimensional Euclidean phenomena that have been studied extensively in harmonic analysis in the context of hypersurfaces and singular measures on $\mathbb{R}^d$. On the other hand, they arise naturally in the consideration of density and differentiation theorems for averages on sparse sets. We describe these below.

1.3.1. *Analogues of averaging operators over submanifolds of* $\mathbb{R}^d$. There is a vast literature on maximal and averaging operators over families of lower-dimensional submanifolds of $\mathbb{R}^d$. A fundamental and representative result is the *spherical maximal theorem*, due to E.M. Stein [37] for $d \geq 3$ and Bourgain [8] for $d = 2$. We state it here for future reference.

Theorem 1.4 (Stein [37], Bourgain [8]). *Recall the spherical maximal operator in $\mathbb{R}^d$:*

$$\tilde{\mathfrak{M}}_{\mathbb{S}^{d-1}} f(x) = \sup_{r>0} \int_{\mathbb{S}^{d-1}} |f(x+ry)| d\sigma(y), \tag{1.10}$$

where σ is the normalized Lebesgue measure on the unit sphere $\mathbb{S}^{d-1}$. Then $\tilde{\mathfrak{M}}_{\mathbb{S}^{d-1}}$ is bounded on $L^p(\mathbb{R}^d)$ for $p > \frac{d}{d-1}$, and this range of p is optimal.

Many results of this type are known for other classes of manifolds in $\mathbb{R}^d$ obeying appropriate smoothness and curvature conditions. We refer the reader to [39], [10], [29], [30] for an introduction to this area of research and further references.

No similar theory has been developed so far in one dimension. Indeed, it is not clear *a priori* what such a theory might look like, given that the real line has no nontrivial lower-dimensional submanifolds. However, given any $\epsilon > 0$, there are many singular measures on $\mathbb{R}$ supported on sets of Hausdorff dimension $1 - \epsilon$. Viewing ϵ as an analogue of "codimension", it is natural to ask whether by imposing additional structure on these sets that would assume the role of curvature, one might obtain L^p estimates similar to those in Theorem 1.4 for the associated maximal operators and for a range $p > p_\epsilon$, where $p_\epsilon \searrow 1$ as $\epsilon \to 0$. Theorem 1.3 provides an affirmative answer to this question. Theorem 1.1 may be interpreted as the limiting situation as $\epsilon \to 0$ (compare with Theorem 1.4 as $n \to \infty$) where the maximal range $(1, \infty]$ of p is achieved for a single set S of zero Lebesgue measure.

1.3.2. *Maximal averages via Fourier decay estimates.* We now turn to the study of maximal operators $\tilde{\mathfrak{M}}$ defined as in (1.2) with μ obeying appropriate Fourier decay conditions. It turns out that such conditions may often be substituted for the geometric assumptions of §1.3.1 (see e.g. [13], [32] and the references therein). From this perspective, our result may be viewed as an extension of the following result by Rubio de Francia [32]. We write $\widehat{\mu}(\xi) = \int e^{-2\pi i \xi x} d\mu(x)$.

Theorem 1.5 (Rubio de Francia [32]). *Suppose that σ is a compactly supported Borel measure on $\mathbb{R}^d$, $d \geq 1$, such that*

$$|\widehat{\sigma}(\xi)| \leq C(1 + |\xi|)^{-a} \tag{1.11}$$

for some $a > \frac{1}{2}$. Then the maximal operator $\tilde{\mathfrak{M}}_\sigma$, defined as in (1.2) but with μ replaced by σ, is bounded on $L^p(\mathbb{R}^d)$ for $p > (2a+1)/(2a)$.

Theorem 1.5 implies Theorem 1.4 for $d \geq 3$, since then the surface measure σ on the sphere obeys the above assumption with $a = \frac{d-1}{2} > \frac{1}{2}$, but it fails to capture the circular maximal estimate in $\mathbb{R}^2$ for which $a = \frac{1}{2}$ just misses the stated range. We also observe that the range of p in Theorem 1.5 is independent of the dimension d; rather, it is given in terms of the Fourier decay exponent a.

It is not possible for a singular measure σ on $\mathbb{R}$ to obey (1.11) with $a > \frac{1}{2}$ (see [34]). In particular, Theorem 1.5 does not apply in this case. On the other hand, there are many such measures obeying (1.11) with a smaller exponent. Recall that the *Fourier dimension* of a compact set $S \subset \mathbb{R}$ is defined by

$$\dim_{\mathbb{F}}(S) = \sup\{0 \leq \beta \leq 1 : \ \exists \text{ a probability measure } \sigma \text{ supported on } S \\ \text{such that } |\widehat{\sigma}(\xi)| \leq C(1+|\xi|)^{-\beta/2} \text{ for all } \xi \in \mathbb{R}\}.$$

It is well known that $\dim_{\mathbb{F}}(S) \leq \dim_{\mathbb{H}}(S)$ for all compact $S \subset \mathbb{R}$, and that the inequality is often strict [26], [14]. However, there are also many examples of sets with $\dim_{\mathbb{F}}(S) = \dim_{\mathbb{H}}(S)$, see e.g. [34], [23], [5], [6], [22], [24]. Such sets are known as *Salem sets.* It is of interest to ask whether there is an analogue of Theorem 1.5 that might apply to singular measures supported on Salem sets and obeying (1.11), possibly with additional assumptions.

It turns out that the proofs of Theorems 1.1 and 1.3 do not use any Fourier decay conditions of the form (1.11). Instead, the key to the proofs is the correlation condition (4.2). If (1.11) indicates the *linear uniformity* of S (see [24]), then (4.2) may be viewed as analogous to higher-order uniformity conditions in additive combinatorics (cf. [16], [18]). Such conditions are known to be strictly stronger than Fourier-analytic estimates. It is in fact possible to prove that the correlation condition (4.2) implies Fourier decay estimates of the form (1.11); in particular, it follows that the sets we construct must have positive Fourier dimension, at least if the ϵ in Theorem 1.3 is sufficiently small ($\epsilon < \frac{1}{5}$ will do). However, the rate of decay obtained in this manner is far from optimal. In the case of the set S of dimension 1 given by Theorem 1.1, our current methods yield (1.11) for all $a < \frac{1}{8}$, whereas the optimal range would be $a \leq \frac{1}{2}$. Note that the range of p in Theorems 1.1 and 1.3 is better than what would follow from the numerology of Theorem 1.5 with that value of a. We do not know whether it is possible to prove maximal estimates such as those in Theorems 1.1 or 1.3 based solely on Fourier decay with $a < \frac{1}{2}$.

With some additional effort, it is possible to construct sequences of sets S_k obeying all conditions of Theorems 1.1 and 1.3, respectively, such that S is also a Salem set. This can be done (as shown in Section 9.1) by adding the appropriate Fourier-analytic conditions to Theorem 5.1 and proving them along the same lines as in [24, Section 6]. However, the Fourier decay is not actually used in the proofs of any of our theorems.

1.3.3. *Density theorems and differentiation of integrals.* In addition to the considerations above, there are natural questions concerning density and differentiation theorems in one dimension that suggest the directions we pursue here. We do not attempt to survey the vast literature on density theorems and differentiation of integrals (see [7], [12] for more information) and focus only on the specific problems relevant to the present discussion.

The following question was raised and investigated by Preiss [31] and Aversa-Preiss [2], [3]: to what extent can the Lebesgue density theorem be viewed as "canonical" in $\mathbb{R}$, in the sense that any other density theorem that takes into account the affine structure of the reals must follow from the Lebesgue density theorem?

Let us clarify and motivate this statement. Consider a family $\mathcal{S}$ of measurable subsets of $\mathbb{R}$. We will say that $\mathcal{S}$ has the *translational density property* if for every measurable set $E \subset \mathbb{R}$ we have

$$\lim_{S\in\mathcal{S},\mathrm{diam}(S\cup\{0\})\to 0} \frac{|(x+S)\cap E|}{|S|} = 1 \text{ for a.e. } x \in E. \tag{1.12}$$

Here and below, we use $x+S$ to denote the translated set $\{x+y : \ y \in S\}$.

It follows from the Lebesgue density theorem that the collection of intervals $\{(-r,r) : r > 0\}$ has this property. A moment's thought shows that collections such as $\{(0,r) : \ r > 0\}$ or $\{(\frac{r}{2},r) : \ r > 0\}$ also have it, simply because the intervals in question occupy at least a fixed positive proportion of $(-r,r)$.

Consider now the family of intervals $\mathcal{S} = \{I_k\}_{k=1}^{\infty}$, where $I_k = (\frac{k}{(k+1)!}, \frac{1}{k!})$. We have $|I_k| = \frac{1}{(k+1)!}$ and $\mathrm{diam}(I_k \cup \{0\}) = \frac{1}{k!}$, hence the last argument no longer applies. In other words, the Lebesgue density theorem does not imply any density properties of $\mathcal{S}$. Nonetheless, $\mathcal{S}$ does have the translational density property, courtesy of the *hearts density theorem* of Preiss [31] and Aversa-Preiss [2] (see also [11] for an alternative proof).

The collection $\mathcal{S}$ in the last example does not generate an *affine invariant density system*: if we let $I_k = (\frac{k}{(k+1)!}, \frac{1}{k!})$ as before and define $\mathcal{S}' = \{rS_k : r > 0, k \in \mathbb{N}\}$, then (1.12) does not hold with $\mathcal{S}$ replaced by $\mathcal{S}'$. (Note that the limit in (1.12) is now being taken over the two parameters k and r.) In fact, Aversa-Preiss prove in [2] that no sequence of intervals I_k can generate an affine invariant density system unless $\liminf_{k\to\infty} |I_k|/\mathrm{diam}(I_k \cup \{0\}) > 0$, in which case the density property in question follows from the Lebesgue theorem as explained above.

On the other hand, if we drop the requirement that $\mathcal{S}$ be a family of intervals, it is possible for $\mathcal{S}$ to generate an affine invariant density system independently of the Lebesgue density theorem. This was announced by Aversa and Preiss in [2] and proved in [3].

Theorem 1.6 (Aversa-Preiss [2], [3])**.** *There is a sequence $\{S_k\}$ of compact sets of positive measure such that $|S_k| \to 0$ and:*

(a) 0 is a Lebesgue density point for $\mathbb{R} \setminus \bigcup S_k$, and in particular we have

$$\lim_{n\to\infty} \frac{|S_k|}{\mathrm{diam}(S_k \cup \{0\})} = 0;$$

(b) the family $\{rS_k : \ r > 0, k \in \mathbb{N}\}$ has the affine density property.

This essentially settles the matter for density theorems, except that constructing an explicit example of sets S_k as in Theorem 1.6 is still an open problem. (The Aversa-Preiss construction is probabilistic, and so is ours below.) However, the analogous question for L^p differentiation theorems remained unanswered.

We will say that $\mathcal{S}$ *differentiates*[1] $L^p_{\rm loc}(\mathbb{R})$ for some $1 \le p \le \infty$ if for every $f \in L^p_{\rm loc}(\mathbb{R})$ we have

$$\lim_{S\in\mathcal{S},\mathrm{diam}(S\cup\{0\})\to 0} \frac{1}{|S|}\int_{x+S} f(y)dy = f(x) \quad \text{for a.e. } x\in\mathbb{R}. \tag{1.13}$$

For instance, the Lebesgue differentiation theorem states that the collection $\{(-r,r):\ r>0\}$ differentiates $L^1_{\rm loc}(\mathbb{R})$. Note that the differentiation property (1.13) implies the density property (1.12), by letting f range over characteristic functions of measurable sets. There is no reason, though, why the converse implication should automatically hold.

While density theorems (such as Theorem 1.6 or the hearts density theorem mentioned earlier) can often be proved using purely geometrical considerations, differentiation theorems tend to require additional analytic input, usually in the form of maximal estimates. A well-known and representative example is provided by the Hardy-Littlewood maximal theorem [20], [41], which easily implies the Lebesgue differentiation theorem.

Aversa and Preiss conjectured in [3] that their Theorem 1.6 could be strengthened to an L^2 differentiation theorem. Specifically, there should exist a sequence of sets $\{S_k\}$ as in Theorem 1.6 such that the family $\{rS_k : r>0,\ k\in\mathbb{N}\}$ differentiates $L^2(\mathbb{R})$ in the sense of (1.13). Our maximal estimates in Theorem 1.1 imply the Aversa-Preiss conjecture along the lines of the standard Hardy-Littlewood argument. Our Theorem 1.2 is in fact stronger, providing a family of sparse sets which differentiates $L^p(\mathbb{R})$ for all $p>1$. Preiss's argument in Subsection 8.2 shows that this range is optimal.

1.4. **Outline of the proofs.** The intuition behind the construction in Theorems 1.1 and 1.3 is, roughly, that such results might hold if the sets S_k (hence also S) are sufficiently randomly distributed throughout the interval $[1,2]$. Thus the challenge is first to find appropriate pseudorandomness conditions that guarantee the boundedness of our maximal operators, then to actually construct a family of sets obeying such conditions. Our arguments are largely inspired by considerations from multidimensional harmonic analysis, in particular by Bourgain's proof of the circular maximal theorem [8]. The probabilistic construction of S_k is somewhat similar to that in [24, Section 6], but significantly more complicated.

The sets S_k will be constructed by randomizing a Cantor-type iteration whose general features are described in Section 2. The main task is to prove that S_k may be chosen so that the restricted maximal operator $\mathcal{M}$ obeys $L^p \to L^q$

[1] This is a slight abuse of the standard terminology, which would require us to say instead that the family $\{S+x\}_{x\in\mathbb{R}}$ differentiates $L^p_{\rm loc}(\mathbb{R})$.

bounds as indicated in Theorems 1.1 and 1.3. Once such bounds are available, the corresponding estimates on $\tilde{\mathcal{M}}^a$ are obtained through the scaling analysis in Section 7, and the estimates on $\mathfrak{M}$ and $\tilde{\mathfrak{M}}$ follow automatically provided that the limiting measure μ exists. The differentiation theorems (Theorem 1.2 and 1.3 (d)) are deduced in Section 8.

Our analysis of $\mathcal{M}$ begins with several preliminary reductions carried out in Section 3.2. Consider the auxiliary restricted maximal operators

$$\mathcal{M}_k f(x) = \sup_{1<t<2} \left| \int f(x+ty)\sigma_k(y)dy \right| ,$$

where $\sigma_k = \phi_{k+1} - \phi_k$, and ϕ_k is the normalized Lebesgue density on S_k. The bulk of the work is to prove appropriate $L^p \to L^q$ bounds on $\mathcal{M}_k$; this implies the bounds on $\mathcal{M}$ upon summation in k. We further replace each $\mathcal{M}_k$ by its discretized and linearized counterpart Φ_k, the discretization being in the space of affine transformations. By duality and interpolation, the desired L^p estimates on Φ_k will follow from restricted strong-type estimates on the "dual" operator Φ_k^*. These reductions are all well known in the harmonic analysis literature, even though the details are specific to the problem at hand. We will follow the approach of [8], [36], and especially [35] with relatively minor modifications.

The main part of our argument is to prove the required estimates on Φ_k^*. Before we describe it in more detail, we pause for a moment to recall the analogous part of Bourgain's proof of the circular maximal theorem in [8]. In his context, the dual linearized operator Φ_k^* acting on characteristic functions $g = \mathbf{1}_\Omega$ has the form

$$\Phi_k^* g(z) = \int_\Omega \frac{1}{|E_{x,k}|} \mathbf{1}_{E_{x,k}}(z) dx,$$

where each $E_{x,k}$ is an annulus of thickness 2^{-k} and radius r_x centered at x. The main task is to prove that Φ_k^* is bounded on $L^{p'}$ with $1 \le p' < 2$. The L^1 bound is trivial, and the proof would be complete if we could prove a similar bound on L^2. We have

$$\begin{aligned} \|\Phi_k^* g\|_2^2 &= \int \int_{\Omega\times\Omega} \frac{1}{|E_{x,k}|\,|E_{y,k}|} \mathbf{1}_{E_{x,k}}(z) \mathbf{1}_{E_{y,k}}(z) dx\, dy\, dz \\ &= \int_{\Omega\times\Omega} \frac{1}{|E_{x,k}|\,|E_{y,k}|} |E_{x,k} \cap E_{y,k}| dx\, dy\,. \end{aligned} \tag{1.14}$$

If we had

$$|E_{x,k} \cap E_{y,k}| \le C_k |E_{x,k}|\,|E_{y,k}|, \tag{1.15}$$

the needed L^2 bound would follow. Unfortunately, (1.15) need not always hold. Specifically, if the two annuli are "internally tangent" in a clamshell configuration, the area of the intersection on the left side of (1.15) can easily be much larger than $|E_{x,k}|\,|E_{y,k}| \approx 2^{-2k}$.

Bourgain's key observation is that geometric considerations put a strict limit on the size of the set of pairs $(x, y) \in \Omega^2$ for which the associated annuli are internally tangent. The remaining generic (or *transverse*) intersections do have reduced area. This allows him to split the region of integration in two parts. One of them involves only transverse intersections, hence there is a good L^2 bound as described above. The other part covers the internal tangencies; here the L^2 estimates are poor, but on the other hand the L^1 estimates can be improved thanks to the small size of the region. An interpolation argument completes the proof.

Let us now try to apply a similar argument in our setting, with p restricted for now to the range $(2, \infty]$ so that $1 \leq p' < 2$. As in Bourgain's proof, the restricted weak L^2 bounds for Φ_k^* are based on estimates on the size of the double intersections $(x + rS_k) \cap (y + sS_k)$ via the appropriate analogue of (1.14). While we still expect that generic double intersections should be significantly smaller than $|S_k|$, the task of actually estimating them turns out to be quite hard, due to the interplay between the different scales in the Cantor iteration.

To illustrate the problem, we consider the following somewhat simplified setting. Suppose that the k-th iteration S_k of the Cantor set is given. Subdivide each of the intervals of S_k into N_{k+1} subintervals of equal length, and choose $N_{k+1}^{1-\epsilon}$ of them within each interval of S_k. Given the translation and dilation parameters x, y, r, s, what is the size of $(x + rS_{k+1}) \cap (y + sS_{k+1})$?

We write the intersection in question as a union of sets

$$(x + r(I \cap S_{k+1})) \cap (y + s(J \cap S_{k+1})), \tag{1.16}$$

where I and J range over all intervals of S_k. If $I \neq J$, the S_{k+1}-subintervals of I and J were chosen independently, hence (1.16) is expected to consist of about $N_{k+1}^{1-2\epsilon}$ such subintervals. In other words, we expect a substantial gain compared to the size of each of the sets $I \cap S_{k+1}$ and $J \cap S_{k+1}$. On the other hand, this argument does not apply to (1.16) with $I = J$, where we cannot expect to do better than the trivial bound.

Following Bourgain, we will refer to the first type of intersections ((1.16) with $I \neq J$) as *transverse intersections*, and to the second type (with $I = J$) as *internal tangencies*. At each step k of the iteration, a typical intersection of two affine copies of S_k will consist of both transverse intersections and internal tangencies. If there are few internal tangencies, we expect an overall gain as described above. If on the other hand there are many internal tangencies, a geometrical argument shows that both $|x - y|$ and $|r - s|$ must be small relative to the current scale, which in turn restricts the relevant domain of (x, y). As in Bourgain's proof, we are able to combine these two observations to prove the desired maximal bound. To extend our bounds to $1 < p \leq 2$ (hence $2 \leq p' < \infty$), we consider the L^n analogues of (1.14) which involve n-fold intersections of affine copies of S_k.

The precise statement of the intersection bound we need is given by the *transverse correlation condition* (4.2). In Section 4 we formulate the correlation condition and prove that it does indeed guarantee a restricted strong type estimate on Φ_k^*. The correlation condition (4.2) may be viewed as a multiscale analogue of the higher order uniformity conditions in additive combinatorics, see e.g. [16], [18]. It appears to be stronger than the pseudorandomness conditions considered so far in the literature, due to the inclusion of the dilation factor and the interplay between different scales.

The random construction of sets S_k obeying our correlation condition is carried out in Section 5. This part of the proof contains the bulk of the technical work and requires the full strength of our probabilistic machinery. The procedure is based on a Cantor-type iteration as described in Section 2, but now each S_k is randomized subject to appropriate constraints on the parameters. We then use large deviation inequalities (specifically, Bernstein's inequality and Azuma's inequality) to prove that at each step of the construction there is a positive probability that the set S_k has the required properties including (4.2). Finally, in Section 6 we fix the parameters of the random construction and complete the proof of our restricted maximal estimates.

1.5. **Acknowledgement.** We are grateful to Vincenzo Aversa and Nir Lev for bringing the Aversa-Preiss conjecture to our attention, and to Nir Lev for a further introduction to questions regarding density and differentiation theorems for sparse sets.

We thank Andreas Seeger for showing us the argument (together with relevant background [28], [13], [9]) that $\tilde{\mathcal{M}}$ and $\tilde{\mathfrak{M}}$ are bounded on all $L^p(\mathbb{R})$ with $1 < p < 2$ for which the restricted operator $\mathcal{M}$ is L^p-bounded. He also referred us to [19], the techniques of which combined with his proof led to the $L^p - L^q$ generalizations for $\tilde{\mathcal{M}}^a$ and $\tilde{\mathfrak{M}}^a$ stated in Theorems 1.1 (e) and 1.3 (c). We include this argument with his kind permission in Section 7 (see Proposition 7.2).

The proof that singular measures cannot differentiate $L^1(\mathbb{R})$ (see Subsection 8.2) is due to David Preiss. We are indebted to him for allowing us to include his argument in this article.

This project was started at the Fields Institute in Toronto, where both authors were participating in the Winter 2008 program on Harmonic Analysis. Both authors are supported by NSERC Discovery Grants.

2. The general Cantor-type construction

2.1. **Basic construction of the sets $\{S_k\}$.** All the nested sequences of sets $\{S_k : k \geq 1\}$ considered in this paper will be obtained using a Cantor-type construction, whose basic features we now describe. The parameters in the construction are the following:

(a) a nondecreasing sequence of positive integers $\{N_k\colon k \geq 1\}$ with $\delta_k^{-1} = N_1 N_2 \cdots N_k$,

(b) certain sequences $\boldsymbol{\kappa}_k$ and $\boldsymbol{\tau}_{k+1}(\mathbf{i})$ of 0-s and 1-s,

$$\boldsymbol{\kappa}_k = \{\kappa_k(\mathbf{i}) : \mathbf{i} = (i_1, \cdots, i_k), 1 \le i_j \le N_j,\ 1 \le j \le k\},\ \text{and}$$

$$\boldsymbol{\tau}_{k+1}(\mathbf{i}) = \{\tau_{k+1}(\mathbf{i}, j) : 1 \le j \le N_{k+1}\}\ \text{ satisfying}$$

$$\kappa_{k+1}(\bar{\mathbf{i}}) = \kappa_k(\mathbf{i})\tau_{k+1}(\bar{\mathbf{i}}), \quad \text{where } \bar{\mathbf{i}} = (i_1, \cdots, i_{k+1}).$$

Given these quantities, we denote

$$\mathbb{I} = \mathbb{I}_k = \{\mathbf{i} = (i_1, \cdots, i_k) \in \mathbb{Z}^k : 1 \le i_r \le N_r,\ 1 \le r \le k\},$$

and for every multi-index $\mathbf{i} = (i_1, \cdots, i_k) \in \mathbb{I}_k$,

$$\alpha(\mathbf{i}) = \alpha_k(\mathbf{i}) = 1 + \frac{i_1 - 1}{N_1} + \frac{i_2 - 1}{N_1 N_2} + \cdots + \frac{i_k - 1}{N_1 \cdots N_k}, \tag{2.1}$$

$$I_k(\mathbf{i}) = [\alpha(\mathbf{i}), \alpha(\mathbf{i}) + \delta_k], \quad \text{so that} \quad I_k(\mathbf{i}) = \bigcup_{i_{k+1}=1}^{N_{k+1}} I_{k+1}(\bar{\mathbf{i}}). \tag{2.2}$$

The argument k will sometimes be suppressed if it is clear from the context. We also set for $k \ge 1$,

$$M_k = N_1 N_2 \cdots N_k (\text{so that } \delta_k = M_k^{-1}), \quad P_k = \#\{\mathbf{i} : \kappa_k(\mathbf{i}) = 1\}.$$

The construction proceeds as follows. Starting with the interval $[1, 2]$ equipped with the Lebesgue measure, we subdivide it into N_1 intervals $\{I_1(i) : 1 \le i \le N_1\}$ of equal length. We choose the P_1 intervals $I_1(i_1)$ for which $\kappa_1(i_1) = 1$ and assign weight P_1^{-1} to each one. At the second step, we subdivide each of the intervals chosen at the first step into N_2 subintervals of equal length δ_2, and choose from $I_1(i_1)$ the subintervals $\{I_2(\mathbf{i}),\ \mathbf{i} = (i_1, i_2)\}$ such that $\tau_2(\mathbf{i}) = 1$. The total number of chosen subintervals at this stage is therefore P_2, and each one is assigned a weight of P_2^{-1}. We continue to iterate the procedure, selecting at the $(k+1)$-th stage subintervals of the intervals chosen at the k-th step, based on the sequences $\boldsymbol{\tau}_{k+1}(\mathbf{i})$. In summary, the sets S_k are chosen according to the scheme

$$S_0 = [1, 2], \qquad S_k = \bigcup_{\mathbf{i}} \{I_k(\mathbf{i}) : \kappa_k(\mathbf{i}) = 1\}.$$

We will always assume that $|S_k| \searrow 0$, i.e., $P_k \delta_k \to 0$.

2.2. **The Hausdorff dimension of the set S.** We now investigate the Hausdorff dimension of the resulting set $S = \bigcap_{k=1}^{\infty} S_k$ as a function of the parameters of the construction.

Lemma 2.1. *Let* $\dim_{\mathbb{H}}(S)$ *denote the Hausdorff dimension of* S *constructed above. Then*

(a) $\dim_{\mathbb{H}}(S) \le \liminf_{k\to\infty} \log(P_k)/\log(M_k)$.
(b) $\dim_{\mathbb{H}}(S) \ge s_0 := \liminf_{k\to\infty} \log(P_k/N_k)/\log(M_{k-1})$.

Proof. Part (a) follows immediately from Proposition 4.1 in [15]. For the proof of part (b), we follow an approach similar to Example 4.6 in [15]. The goal is to define a measure ν on S such that for any $s < s_0$, there exists a constant $C_s < \infty$ satisfying

$$\nu(J) \leq C_s |J|^s \quad \text{for all intervals } J \subset \mathbb{R}. \tag{2.3}$$

The desired conclusion would then follow from Frostman's lemma (see e.g. Proposition 8.2 in [42]).

In order to define ν, we follow a standard procedure due to Caratheodory (see Chapter 4, [26]). Let $\mathcal{B} = \bigcup \mathcal{B}_k$, where $\mathcal{B}_0 = [1, 2]$ and $\mathcal{B}_k$ for $k \geq 1$ is the family of all basic intervals of S_k, i.e., intervals of the form $\{I_k(\mathbf{i}) : \kappa_k(\mathbf{i}) = 1\}$. For each interval $I \in \mathcal{B}$, we define its weight $w(I)$ to be

$$w([1, 2]) = 1, \qquad w(I) = P_k^{-1} \text{ if } I \in \mathcal{B}_k, \tag{2.4}$$

and a family of outer measures ν_k as follows,

$$\nu_k(F) := \inf \Big\{ \sum_{i=1}^{\infty} w(J_i) : F \subseteq \bigcup_{i=1}^{\infty} J_i, \ |J_i| \leq \delta_k, \ J_i \in \mathcal{B} \Big\} \tag{2.5}$$

for all $F \subseteq S$. It is easy to see that ν_k is monotonic, so we can define ν by

$$\nu(F) = \lim_{k \to \infty} \nu_k(F) = \sup_{k \geq 1} \nu_k(F). \tag{2.6}$$

Then ν is a non-negative regular Borel measure of unit mass on subsets of S (Theorem 4.2, [26]).

To prove (2.3), let J be an interval with $0 < |J| \leq \delta_1$. Given such a J, there is a unique $k = k(J)$ such that $\delta_{k+1} \leq |J| < \delta_k$. The number of basic intervals of S_{k+1} that intersect J is

(i) at most $2N_{k+1}$ since J intersects at most two intervals of S_k, and

(ii) at most $|J|/\delta_{k+1}$, since the basic intervals comprising S_{k+1} are of length δ_{k+1} and have disjoint interiors.

It therefore follows from the definitions (2.4) and (2.5) that

$$\begin{aligned} \nu_{k+1}(J) &\leq P_{k+1}^{-1} \min \left[2N_{k+1}, \frac{|J|}{\delta_{k+1}} \right] \\ &\leq P_{k+1}^{-1} (2N_{k+1})^{1-s} \left(\frac{|J|}{\delta_{k+1}} \right)^s \text{ for all } 0 \leq s \leq 1, \end{aligned}$$

$$\text{i.e.,} \quad \frac{\nu_{k+1}(J)}{|J|^s} \leq \frac{2^{1-s} N_{k+1}^{1-s}}{P_{k+1} \delta_{k+1}^s}.$$

Letting $k \to \infty$ and recalling (2.6), we find that the right hand side of the inequality above is bounded above by a constant provided that $s < s_0$. This completes the proof. □

Remark. In our applications, the sequences $\boldsymbol{\kappa}_k$ of 0-s and 1-s will be chosen according to a random mechanism, to be described in Section 5. We will see in these instances that the upper and the lower bounds given by Lemma 2.1 coincide, providing an exact value of the Hausdorff dimension.

2.3. **A limiting measure.** Although most of our results can be stated purely in terms of the maximal operators $\mathcal{M}$ associated with the sequence of sets $\{S_k : k \geq 1\}$, it is often of interest to know whether the normalized Lebesgue measures $\phi_k = \mathbf{1}_{S_k}/|S_k|$ have a nontrivial weak-$*$ limit μ. In this case, the maximal operator $\mathfrak{M}$ associated with μ is bounded by $\mathcal{M}$. If each interval in S_k contains the same number of subintervals of S_{k+1}, it is easy to see that μ exists and is identical to the measure ν defined in the last subsection. Below we provide a sufficient condition for the existence of the weak-$*$ limiting measure under a slightly weaker assumption that will be verified for certain constructions in the sequel.

Lemma 2.2. *Suppose that the distribution of the chosen subintervals $\{I_{\mathbf{i}}(k)\colon \kappa_{\mathbf{i}}(k) = 1\}$ within S_{k-1} is approximately uniform in the following sense:*

$$\sup_{k':k'\geq k} \sum_{\substack{\mathbf{i} \\ \kappa_k(\mathbf{i})=1}} \left| \int_{I_k(\mathbf{i})} \left[\phi_{k'} - \phi_k\right](x)\, dx \right| \to 0 \quad \text{as} \quad k \to \infty. \tag{2.7}$$

Then there exists a probability measure μ on $[1,2]$ such that $\phi_k \to \mu$ in the weak$-*$ topology, i.e., for all $f \in C[1,2]$

$$\int f\phi_k \to \int f d\mu \quad \text{as} \quad k \to \infty.$$

Proof. It suffices to show that $\lim_{k\to\infty} \int f\phi_k$ exists for all continuous functions f on $[1,2]$, i.e., that the sequence $\{\int f\phi_k : k \geq 1\}$ is Cauchy. Since f is uniformly continuous, given any $\epsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(y)| < \frac{\epsilon}{4} \quad \text{whenever} \quad |x - y| < \delta. \tag{2.8}$$

Fix $K \geq 1$ such that $\delta_K < \delta$ and

$$\sup_{k':k'\geq k} \sum_{\substack{\mathbf{i} \\ \kappa_k(\mathbf{i})=1}} \left| \int_{I_k(\mathbf{i})} \left[\phi_{k'} - \phi_k\right](x)\, dx \right| < \frac{\epsilon}{2||f||_\infty} \text{ for all } k \geq K. \tag{2.9}$$

Let $\{x_k(\mathbf{i}) : \kappa_k(\mathbf{i}) = 1\}$ be a collection of points in $[1,2]$ such that $x_k(\mathbf{i}) \in I_k(\mathbf{i})$. Then for all $k' \geq k \geq K$,

$$\begin{aligned}
&\left|\int f(x)\big(\phi_{k'}(x)-\phi_k(x)\big)\,dx\right| \\
&\quad\le \sum_{\substack{\mathbf{i}\\ \kappa_k(\mathbf{i})=1}} \int_{I_k(\mathbf{i})} \Big|\left[f(x)-f(x_k(\mathbf{i}))\right]\left(\phi_{k'}-\phi_k\right)(x)\Big|\,dx \\
&\qquad + \sum_{\substack{\mathbf{i}\\ \kappa_k(\mathbf{i})=1}} |f(x_k(\mathbf{i}))| \left|\int_{I_k(\mathbf{i})} \left(\phi_{k'}-\phi_k\right)(x)\,dx\right| \\
&\quad\le \frac{\epsilon}{4}\int_{S_k}(\phi_{k'}+\phi_k)(x)\,dx + \|f\|_\infty \sum_{\substack{\mathbf{i}\\ \kappa_k(\mathbf{i})=1}} \left|\int_{I_k(\mathbf{i})} \left(\phi_{k'}-\phi_k\right)(x)\,dx\right| \\
&\quad\le 2\frac{\epsilon}{4}+\frac{\epsilon}{2}=\epsilon,
\end{aligned}$$

where we have used (2.8) and (2.9) at the last two steps. □

2.4. **Internal tangencies and transverse intersections.** An important ingredient in the derivation of the maximal estimates is the behavior of the intersections of a fixed number of affine copies of S_k. Obviously, much of our analysis will depend on the specific structure of $\{S_k\}$, which will be described in detail in Section 5. However, we also need certain general properties of the n-fold intersections of affine copies of sets S_k constructed as in Subsection 2.1. The relevant results of this type are collected in this subsection.

Fix $k \ge 1$, $r, s \in [1,2]$ and points x, y in a fixed compact set, say $[-4,0]$ (the reason for this choice will be made clear in the next section). We will be interested in classifying pairs of multi-indices $(\mathbf{i},\mathbf{j}) \in \mathbb{I}_k^2$ such that

$$(x + rI_k(\mathbf{i})) \cap (y + sI_k(\mathbf{j})) \ne \emptyset. \tag{2.10}$$

We will need to distinguish between the situations where $|\alpha_k(\mathbf{i}) - \alpha_k(\mathbf{j})|$ is "small" or "large". The first case will be referred to as an *internal tangency* and the second as a *transverse intersection*. In view of subsequent applications, we give the precise definitions of these notions for general n-fold intersections of intervals. However, the main ideas are already contained in the case $n = 2$, which we encourage the reader to investigate first.

Definition 2.3. *For integers $k \ge 1, n \ge 2$ and any set*

$$\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \le \ell \le n,\ c_\ell \in [-4,0], r_\ell \in [1,2]\}$$

of n translation-dilation pairs, we define a set $\mathbb{F} = \mathbb{F}[n,k;\mathbf{A}_n]$ and n projection maps $\pi_\ell = \pi_\ell[n,k;\mathbf{A}_n](\mathbf{i}_1,\cdots,\mathbf{i}_n) : \mathbb{F} \to \mathbb{I}_k$ as follows,

$$\mathbb{F} = \Big\{(\mathbf{i}_1,\cdots,\mathbf{i}_n) \in \mathbb{I}_k^n : \bigcap_{\ell=1}^n \big(c_\ell + r_\ell I_k(\mathbf{i}_\ell)\big) \ne \emptyset\Big\}, \tag{2.11}$$

$$\pi_\ell(\mathbf{i}_1,\cdots,\mathbf{i}_n) = \mathbf{i}_\ell.$$

Remarks.

1. We emphasize that $\mathbb{F}$ consists of *all* tuples $(\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{I}_k^n$ such that (2.11) holds, regardless of the actual choice of the sets S_k. Thus $\mathbb{F}$ depends only on the parameters n, k, N_j, and on the choice of $\mathbf{A}_n$.
2. Eventually, our translation and dilation parameters c_ℓ and r_ℓ will be chosen from discrete subsets $\mathcal{C}, \mathcal{R}$ of the respective spaces $[-4, 0]$ and $[1, 2]$. Then the total number of possible collections $\mathbb{F}$ cannot exceed $|\mathcal{C}|^n |\mathcal{R}|^n$, again irrespective of the choice of the sets S_k.

The next lemma is an easy observation concerning the "almost injectivity" of the projections π_ℓ.

Lemma 2.4. *For any $1 \leq \ell \leq n$ and any fixed choice of multi-indices $(\mathbf{i}_{\ell'} : 1 \leq \ell' \leq n, \ell' \neq \ell) \in \mathbb{I}_k^{n-1}$,*

$$\max_{\mathbf{i}_\ell :\ (\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F}} \alpha_k(\mathbf{i}_\ell) - \min_{\mathbf{i}_\ell :\ (\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F}} \alpha_k(\mathbf{i}_\ell) \leq 4\delta_k. \tag{2.12}$$

In particular, for any $1 \leq \ell \leq n$ the map π_ℓ is at most four-to-one, i.e.,

$$\sup_{\mathbf{i}_\ell \in \mathbb{I}_k} \#\big(\pi_\ell^{-1}(\mathbf{i}_\ell)\big) \leq 4. \tag{2.13}$$

Proof. The second part of the lemma follows from the first. The inequality in (2.12) is essentially a fact about two-fold intersections. Fix $\ell' \neq \ell$ and $(\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F}$, so that by definition (2.11)

$$(x_{\ell'} \cap r_{\ell'} I_k(\mathbf{i}_{\ell'})) \cap (x_\ell \cap r_\ell I_k(\mathbf{i}_\ell)) \neq \emptyset.$$

Since $r_\ell, r_{\ell'} \in [1, 2]$ any interval of the form $x_{\ell'} + r_{\ell'} I_k(\mathbf{i}_{\ell'})$ can intersect at most four intervals of the form $x_\ell + r_\ell I_k(\mathbf{i}_\ell)$ and these intervals must necessarily be adjacent. The claim follows. ☐

Corollary 2.5. *There exists a decomposition of $\mathbb{F}$ into at most 4^{n-1} subsets so that all the projection maps π_ℓ restricted to each subset are injective.*

Proof. The proof is an easy induction on n combined with (2.13), and is left to the interested reader. ☐

The lemma above motivates the following definition. Setting $\mathbf{i}_\ell = (\mathbf{i}'_\ell, i_{\ell k}) \in \mathbb{I}_{k-1} \times \{1, 2, \cdots, N_k\}$, we find that each $\mathbb{F} = \mathbb{F}[n, k; \mathbf{A}_n]$ decomposes as

$$\mathbb{F} = \mathbb{F}_{\text{int}} \cup \mathbb{F}_{\text{tr}}, \quad \text{where} \quad \mathbb{F}_{\text{int}} := \bigcup_{1 \leq \ell \neq \ell' \leq n} \mathbb{F}_{\text{int}}(\ell, \ell'), \text{ with}$$

$$\mathbb{F}_{\text{int}}(\ell, \ell') := \{(\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F} : \mathbf{i}'_\ell = \mathbf{i}'_{\ell'},\ |i_{\ell k} - i_{\ell' k}| \leq 4\}, \text{ and}$$

$$\mathbb{F}_{\text{tr}} := \mathbb{F} \setminus \mathbb{F}_{\text{int}}.$$

Note that in view of (2.1),

$$(\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F}_{\text{int}}(\ell, \ell') \quad \text{implies} \quad |\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{i}_{\ell'})| \leq 4\delta_k. \tag{2.14}$$

Definition 2.6. *The collections $\mathbb{F}_{int}$ and $\mathbb{F}_{tr}$, which depend only on n, k, $\{N_j : 1 \leq j \leq k\}$ and $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\}$, are referred to as the classes of internal tangencies and transverse intersections respectively.*

A large number of internal tangencies forces a relation between the translation (and hence dilation) parameters, in a sense made precise by the next lemma. (A similar observation was made by Aversa and Preiss in [3].)

Lemma 2.7. *Suppose $\#(\mathbb{F}_{int}) \geq L$. Then*

$$\min\{|c_\ell - c_{\ell'}| : 1 \leq \ell \neq \ell' \leq n\} \leq \min\left(4, 80n(n-1)/L\right).$$

Proof. Since the translation parameters all lie in $[-4, 0]$, we may assume without loss of generality that $L > 20n(n-1)$. Using the definition of $\mathbb{F}_{\mathrm{int}}$ and pigeonholing we can find indices $\ell \neq \ell'$ such that $\#(\mathbb{F}_{\mathrm{int}}(\ell, \ell')) \geq \frac{2L}{n(n-1)}$. By Lemma 2.4, there exists a further subset $\mathbb{F}^* \subseteq \mathbb{F}_{\mathrm{int}}(\ell, \ell')$ such that

$$\#(\mathbb{F}^*) \geq \frac{1}{4}\#(\mathbb{F}_{\mathrm{int}}(\ell, \ell')) \geq \frac{L}{2n(n-1)}, \text{ and } \pi_\ell\Big|_{\mathbb{F}^*} \text{ is injective.} \tag{2.15}$$

Let $(\mathbf{i}_1, \cdots \mathbf{i}_n), (\mathbf{j}_1, \cdots, \mathbf{j}_n) \in \mathbb{F}$. Since $r_\ell, r_{\ell'} \in [1, 2]$, it follows from the definition (2.11) that

$$\begin{aligned} \big|\big(c_\ell + r_\ell \alpha_k(\mathbf{i}_\ell)\big) - \big(c_{\ell'} + r_{\ell'}\alpha_k(\mathbf{i}_{\ell'})\big)\big| &\leq \max(r_\ell, r_{\ell'})\delta_k \leq 2\delta_k, \\ \text{and similarly } \big|\big(c_\ell + r_\ell \alpha_k(\mathbf{j}_\ell)\big) - \big(c_{\ell'} + r_{\ell'}\alpha_k(\mathbf{j}_{\ell'})\big)\big| &\leq 2\delta_k. \end{aligned} \tag{2.16}$$

If further $(\mathbf{i}_1, \cdots, \mathbf{i}_n), (\mathbf{j}_1, \cdots, \mathbf{j}_n) \in \mathbb{F}_{\mathrm{int}}(\ell, \ell')$, then (2.16) and (2.14) imply that

$$\big|(c_\ell - c_{\ell'}) + (r_\ell - r_{\ell'})\alpha_k(\mathbf{i}_\ell)\big| \leq 2\delta_k + r_{\ell'}|\alpha_k(\mathbf{i}_{\ell'}) - \alpha_k(\mathbf{i}_\ell)| \leq 10\delta_k,$$

$$\big|(c_\ell - c_{\ell'}) + (r_\ell - r_{\ell'})\alpha_k(\mathbf{j}_\ell)\big| \leq 2\delta_k + r_{\ell'}|\alpha_k(\mathbf{j}_{\ell'}) - \alpha_k(\mathbf{j}_\ell)| \leq 10\delta_k.$$

Eliminating $(r_\ell - r_{\ell'})$ from the two inequalities above we find that

$$|c_\ell - c_{\ell'}|\big|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{j}_\ell)\big| \leq 40\delta_k.$$

If we now choose $(\mathbf{i}_1, \cdots, \mathbf{i}_n), (\mathbf{j}_1, \cdots, \mathbf{j}_n) \in \mathbb{F}^*$ so that $\big|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{j}_\ell)\big|$ is maximal in this class, it follows from (2.15) that $|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{j}_\ell)| \geq \frac{L\delta_k}{2n(n-1)}$, from which the desired conclusion follows. □

We end this section by applying these definitions to the intersections of the sets S_k. Fix $k \geq 1$, and suppose that the sets $S_1, \ldots, S_k$ have been chosen. Recalling from Subsection 2.1 that $S_k = \bigcup_{\kappa_k(\mathbf{i})=1} I_k(\mathbf{i})$ and restricting the scale factors $r, s \in [1, 2]$, we find that any intersection of the form $(x + rS_k) \cap (y + sS_k)$ is nonempty if and only if there exists at least one pair of multi-indices $(\mathbf{i}, \mathbf{j})$ such that $\kappa_k(\mathbf{i}) = \kappa_k(\mathbf{j}) = 1$ and (2.10) holds. In general, there may be many such pairs $(\mathbf{i}, \mathbf{j})$. Given two affine copies of S_k with a large intersection, one of two cases must arise: either there will be a strong match, in the sense that the number of internal tangencies will be large, or else all but a few such pairs will be transverse intersections. We will need to treat these two situations differently.

As before, the exact definitions are stated for general n-fold intersections of affine copies of S_k.

Definition 2.8. *Let $\{S_k : k \geq 1\}$ be a sequence of sets constructed as in Subsection 2.1. Given $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\} \subseteq [0,1] \times [1,2]$, the sets $x_\ell + r_\ell S_k$ are said to have L internal tangencies (respectively transverse intersections) if*

$$\#\{(\mathbf{i}_1, \cdots, \mathbf{i}_n) \in \mathbb{F}_{int} \text{ (resp. } \mathbb{F}_{tr}) : \kappa_k(\mathbf{i}_1) = \cdots = \kappa_k(\mathbf{i}_n) = 1\} = L.$$

The total number of intersections among $x_\ell + r_\ell S_k$ is defined to be the sum of the numbers of internal tangencies and transverse intersections.

A large number of internal tangencies among $c_\ell + r_\ell S_k$ implies a lower bound on $\#(\mathbb{F}_{\text{int}})$, which in light of Lemma 2.6 (and regardless of what S_k may be) provides a gain in the form of relative proximity of the translation parameters $\{c_\ell\}$. On the other hand, controlling the transverse intersections will be possible only under certain additional assumptions on S_k. We take up this issue in Sections 4 and 5.

3. Preliminary reductions

We now begin our analysis of the restricted maximal operator $\mathcal{M}$ defined in (1.5). In this section, we decompose $\mathcal{M}$ as a sum of auxiliary restricted maximal operators $\mathcal{M}_k$, each of which is then replaced by a linearized and discretized operator Φ_k. We will subsequently investigate the $L^p \to L^q$ mapping properties of Φ_k when acting on functions supported in a fixed compact set. While these reductions are well known and have been used extensively in the literature, it is not entirely straightforward to adapt them to the specific situation at hand, hence we include them for completeness.

3.1. Spatial restriction.

Lemma 3.1. *Suppose that there are exponents (p, q) with $1 \leq p \leq q < \infty$ and a constant $A > 0$ such that $\mathcal{M}$ as in (1.5) satisfies*

$$\|\mathcal{M}f\|_q \leq A\|f\|_p \quad \text{for all } f \in L^p[0,1]. \tag{3.1}$$

Then the inequality in (3.1) continues to hold for all $f \in L^p(\mathbb{R})$, with the constant A replaced by $4^{\frac{1}{q}}A$.

Proof. It suffices to prove the assertion for functions $f \in L^p(\mathbb{R})$ of arbitrary compact support. Given any such f, we can find an integer R such that $f = \sum_{i=-R}^{R} f_i$, where f_i is supported in $[i, i+1]$. Observe that the support of $\mathcal{M}f_i$ is contained in $[i-4, i]$, which implies

$$\begin{aligned} \|\mathcal{M}f\|_q^q &= \Big\|\mathcal{M}\big(\sum_i f_i\big)\Big\|_q^q \leq \Big\|\sum_i \mathcal{M}f_i\Big\|_q^q \\ &\leq 4\sum_i \|\mathcal{M}f_i\|_q^q \leq 4\sum_{i=-R}^{R} A^q\|f_i\|_p^q \,. \end{aligned} \tag{3.2}$$

In the second line we have used the finitely overlapping supports for $\mathcal{M}f_i$, and then applied (3.1) to each f_i. If $p \le q$, we estimate the last sum in (3.2) by

$$4A^q \sum_{i=-R}^{R} \left(\|f_i\|_p^p\right)^{\frac{q}{p}} \le 4A^q \Big[\sum_{i=-R}^{R} \|f_i\|_p^p\Big]^{\frac{q}{p}} \le 4A^q \|f\|_p^q \;.^* \qquad \square$$

We will henceforth assume that all functions are supported on $[0,1]$, so that $\mathcal{M}$ is supported within the fixed compact set $[-4,0]$.

3.2. **Linearization and discretization.** Define the auxiliary restricted maximal operators

$$\mathcal{M}_k f(x) := \sup_{1<r<2} \Big| \int f(x+ry)\sigma_k(y)dy \Big| \quad \text{where} \quad \sigma_k = \phi_{k+1} - \phi_k. \tag{3.3}$$

Then

$$\mathcal{M}f \le \mathcal{N}f + \sum_{k=1}^{\infty} \mathcal{M}_k|f|,$$

where $\mathcal{N}f(x) = \sup_{1<r<2} \int |f(x+ry)|\phi_1(y)dy$. It is an easy exercise to deduce from Hölder's inequality that $\|\mathcal{N}f\|_q \le 4^{1/q}\|\mathcal{N}f\|_\infty \le 4^{1/q}\|\phi_1\|_{p'}\|f\|_p$ for any $p,q \in [1,\infty]$; the main task is to estimate $\mathcal{M}_k$ with $k \ge 1$. We begin by discretizing each $\mathcal{M}_k$ in the space of affine transformations. Specifically, we decompose the spaces of translations x and dilations r (i.e. the intervals $[-4,0]$ and $[1,2]$) into disjoint intervals $\{Q_i\}$ and $\{R_i\}$ respectively, of length δ_{k+1}^L, where L is an integer to be fixed at the end of this subsection. The centers of Q_i and R_i are denoted by c_i and r_i respectively. Let

$$\mathcal{C} = \{c_i : 1 \le i \le 4\delta_{k+1}^{-L}\}, \quad \mathcal{R} = \{r_i : 1 \le i \le \delta_{k+1}^{-L}\}.$$

Proposition 3.2. *Fix $1 < p < \infty$. Then there is a large integer $L = L(p)$ and a small constant $\eta = \eta(p) > 0$ such that the following conclusions hold:*

(a) For every $f \in C_c[0,1]$, there are measurable functions $c(x)$ and $r(x)$ depending on f and taking values in the discrete sets $\mathcal{C}$ and $\mathcal{R}$ respectively, such that

$$\mathcal{M}_k f(x) \le 4|\Phi_k f(x)| + \mathcal{E}_k f(x), \tag{3.4}$$

where

$$\Phi_k f(x) = \int f(z)V_{k,x}(z)dz, \quad \text{with} \quad V_{k,x}(z) = \sigma_k\left(\frac{z-c(x)}{r(x)}\right).$$

(b) Both $\Phi_k f$ and $\mathcal{E}_k f$ are supported on $[-4,0]$.

(c) For every $q \ge 1$ there is a constant $C_{p,q}$ such that

$$\|\mathcal{E}_k f\|_q \le C_{p,q} 2^{-k\eta}\|f\|_p. \tag{3.5}$$

Proof. Fix a function $f \in C_c[0,1]$. Since f is bounded, so is $\mathcal{M}_k f(x)$. Hence we may choose $x_i \in Q_i$ and $\tilde{r}_i \in [1,2]$ such that for all $x \in Q_i$ we have

$$\begin{aligned}
\mathcal{M}_k f(x) &\le 2 \left| \int f(x_i + \tilde{r}_i y) \sigma_k(y) dy \right| \\
&\le 4 \left| \int f(z) \sigma_k \left(\frac{z - x_i}{\tilde{r}_i} \right) dz \right| \\
&\le 4 \left| \int f(z) \sigma_k \left(\frac{z - c_i}{r_{j(i)}} \right) dz \right| + \mathcal{E}_k f(x),
\end{aligned} \tag{3.6}$$

where

$$\mathcal{E}_k f(x) = 4 \left| \int f(z) \left[\sigma_k \left(\frac{z - x_i}{\tilde{r}_i} \right) - \sigma_k \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right| \tag{3.7}$$

and $r_{j(i)}$ is chosen so that $\tilde{r}_i \in R_{j(i)}$. Note that $|\tilde{r}_i - r_{j(i)}| \le \delta_{k+1}^L$. Thus (3.4) holds with $c(x) = c_i$ and $r(x) = r_{j(i)}$.

Since each $\mathcal{M}_k f$ is supported on $[-4, 0]$, it is obvious from (3.6) and (3.7) that so are $\Phi_k f$ and $\mathcal{E}_k f$. It remains to prove (3.5). For this we observe that

$$\begin{aligned}
|\mathcal{E}_k f(x)| \le & 4 \left| \int f(z) \left[\phi_{k+1} \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_{k+1} \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right| \\
& + 4 \left| \int f(z) \left[\phi_k \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_k \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right|.
\end{aligned} \tag{3.8}$$

By Hölder's inequality, the first term on the right side of (3.8) is bounded by

$$\begin{aligned}
&\|f\|_p \left\| \phi_{k+1} \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_{k+1} \left(\frac{z - c_i}{r_{j(i)}} \right) \right\|_{p'} \\
&\qquad = \frac{1}{P_{k+1} \delta_{k+1}} \|f\|_p \left\| \sum_m (\mathbf{1}_{x_i + \tilde{r}_i I_m^{(k+1)}} - \mathbf{1}_{c_i + r_{j(i)} I_m^{(k+1)}}) \right\|_{p'} \\
&\qquad \le \frac{2^{1/p}}{P_{k+1} \delta_{k+1}} \|f\|_p \left\| \sum_m (\mathbf{1}_{x_i + \tilde{r}_i I_m^{(k+1)}} - \mathbf{1}_{c_i + r_{j(i)} I_m^{(k+1)}}) \right\|_1^{1/p'} \\
&\qquad \le \frac{2^{1/p}}{P_{k+1} \delta_{k+1}} \|f\|_p \cdot \left(\sum_m |(x_i + \tilde{r}_i I_m^{(k+1)}) \triangle (c_i + r_{j(i)} I_m^{(k+1)})| \right)^{1/p'}.
\end{aligned}$$

By Lemma 3.3 below, each symmetric difference $(x_i + \tilde{r}_i I_m^{(k+1)}) \triangle (c_i + r_{j(i)} I_m^{(k+1)})$ has measure bounded by $3\delta_{k+1}^L$. Hence the last expression is bounded by

$$\begin{aligned} \frac{2^{1/p}}{P_{k+1}\delta_{k+1}} \|f\|_p \Big(P_{k+1}\delta_{k+1}^L\Big)^{1/p'} &\leq \frac{2^{1/p}\delta_{k+1}^{(L-1)/p'}}{(P_{k+1}\delta_{k+1})^{1/p}} \|f\|_p \\ &\leq 2^{1/p}\delta_{k+1}^{\frac{L}{p'}-1} \|f\|_p \leq C 2^{-(k+1)\eta} \|f\|_p, \end{aligned}$$

where $\eta = \frac{L}{p'} - 1$ is positive for large enough L whenever $p > 1$. We have used the trivial bounds $P_{k+1} \geq 1$ and $N_k \geq 2$. The second term in (3.8) is bounded similarly, with P_{k+1}, δ_{k+1} replaced by P_k, δ_k. Finally, (3.5) follows from the pointwise bound above and the fact that $\mathcal{E}_k$ are supported on the bounded interval $[-4, 0]$. □

Lemma 3.3. *Let $0 < t < 1$, $\frac{1}{2} < r, s < 2$. Then for any $x, y \in \mathbb{R}$ we have*

$$|[x, x + rt] \triangle [y, y + st]| \leq 3\eta$$

whenever $\eta < t/2$ and $|x - y| < \eta$, $|r - s| < \eta$.

Proof. We may assume without loss of generality that $x \leq y$. Observe first that the two intervals cannot be disjoint, since $y - x < \eta < \frac{t}{2} < rt$. Hence we must have either $x \leq y \leq x + rt \leq y + st$ or $x \leq y \leq y + st \leq x + rt$. In the first case, the symmetric difference has measure $(y - x) + (y + st - x - rt) = 2(y - x) + t(r - s) \leq 3\eta$. In the second case, its measure is $(y - x) + (x + rt - y - st) = (r - s)t \leq \eta$. □

3.3. **The interpolation argument.** We now turn to the question of proving $L^p \to L^q$ bounds for Φ_k. In the next lemma we show how such bounds follow from a restricted strong-type estimate for the "adjoint" operator Φ_k^* given by

$$\Phi_k^* g(z) = \int g(x) V_{k,x}(z)\, dx. \tag{3.9}$$

Although similar interpolation arguments are ubiquitous in the literature, the sequence of steps in the proof is somewhat more complicated than usual, due to the additional challenge of keeping track of the dependence of the operator norm of Φ_k^* on k.

Lemma 3.4. *Let Φ_k^* be as in (3.9) and $q_0 \geq 2$. Suppose that Φ_k^* obeys the restricted strong-type estimate*

$$||\Phi_k^* \mathbf{1}_\Omega||_{q_0} \leq 2^{-k\eta_0} |\Omega|^{\frac{q_0 - 1}{q_0}} \quad \textit{for all sets} \quad \Omega \subseteq [0, 1] \tag{3.10}$$

with some $\eta_0 > 0$. Then for any $p > \frac{q_0}{q_0 - 1}$ there is an $\eta(p) > 0$ such that Φ_k is bounded from $L^p[0, 1]$ to $L^{p(q_0 - 1)}[-4, 0]$ with operator norm bounded by $2^{-k\eta(p)}$.

Proof. The operator Φ_k^* satisfies a trivial $L^1 \to L^1$ bound, with operator norm bounded by a constant independent of k. On one hand, by a standard interpolation theorem for operators satisfying restricted weak-type endpoint bounds

(Chapter 4, Theorem 5.5, [4]), Φ_k^* is bounded from $L^p \to L^q$ for all (p,q) satisfying $p' = q_0/\theta$ and $q' = q_0/(\theta(q_0-1))$, $0 < \theta < 1$, with norm bounded uniformly in k but not necessarily decaying as $k \to \infty$. On the other hand, by Hölder's inequality

$$||\Phi_k^* \mathbf{1}_\Omega||_q \leq ||\Phi_k^* \mathbf{1}_\Omega||_{q_0}^\theta ||\Phi_k^* \mathbf{1}_\Omega||_1^{1-\theta} \leq C 2^{-k\eta_0\theta} |\Omega|^{\frac{1}{p}}.$$

By Theorem 5.3 of [4, Chapter 4]), the last two statements imply that the weak-type (p,q) norm of Φ_k^* is bounded by $C2^{-k\eta_0\theta}$ (possibly with a different constant). Note that $p \leq q$, hence we may apply the Marcinkiewicz interpolation theorem (Theorem 4.13 and Corollary 4.14, Chapter 4, [4]) to two such pairs (p,q) to get the desired strong-type Lebesgue mapping properties on all the intermediate spaces and with the operator norms decaying exponentially in k. The statement for Φ_k follows by duality. □

Combining Lemma 3.4 with Proposition 3.2 and Lemma 3.1, we arrive at the following corollary.

Corollary 3.5. *Assume that (3.10) holds. Then for every* $\frac{q_0}{q_0-1} < p < \infty$, *there is an* $\eta(p) > 0$ *such that*

$$\|\mathcal{M}_k f\|_{(q_0-1)p} \leq 2^{-k\eta(p)} \|f\|_p$$

for all $f \in L^p[0,1]$. *Moreover, the restricted maximal operator* $\mathcal{M}$ *is bounded from* $L^p(\mathbb{R})$ *to* $L^{(q_0-1)p}(\mathbb{R})$.

4. Transverse correlations

We now come to the main part of our proof. The first step, to be accomplished in this section, is to reduce the problem of deriving restricted strong-type $L^{\frac{n}{n-1}} \to L^n$ bounds on Φ_k^* to estimating n-fold correlations between affine copies of S_k with few internal tangencies. The construction of a sequence of sets S_k that will meet the correlation condition in question will be addressed in Section 5. We start by setting up the notation for such n-fold correlations and giving a precise statement of our correlation criterion.

Throughout this section, $n \geq 2$ will be a fixed even integer. We will use $\mathfrak{A} = \mathfrak{A}[n,k,L]$ to denote the finite collection of all n-tuples of translation-dilation pairs that arise from the δ_{k+1}^L discretization procedure in Section 3.2:

$$\mathfrak{A} := \{\mathbf{A}_n : \mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\},\ c_\ell \in \mathcal{C}, r_\ell \in \mathcal{R}\}.$$

In particular, we have $\#(\mathfrak{A}) \leq 4\delta_{k+1}^{-2Ln}$. We will also use $\mathfrak{A}_{\text{tr}}$ to denote the subcollection of those n-tuples which have few internal tangencies:

$$\mathfrak{A}_{\text{tr}} = \{\mathbf{A}_n \in \mathfrak{A} :\ \#(\mathbb{F}_{\text{int}}[n,k;\mathbf{A}_n]) < P_k^{1-\epsilon_0}\},$$

where $\epsilon_0 \in (0,1)$ is a fixed constant (eventually, we will let $\epsilon_0 = \frac{1}{2}$). We write $\mathfrak{A}_{\text{int}} = \mathfrak{A} \setminus \mathfrak{A}_{\text{tr}}$.

Definition 4.1. *Let $\mathbf{A}_n \in \mathfrak{A}$, and let $f_1, \dots, f_n$ be functions on $\mathbb{R}$. We define the n-*fold correlation *of $f_1, \dots, f_n$* according to $\mathbf{A}_n$ *as follows:*

$$\Lambda(\mathbf{A}_n; f_1, \dots, f_n) = \int \prod_{\ell=1}^{n} f_\ell\Big(\frac{z - c_\ell}{r_\ell}\Big)\, dz. \tag{4.1}$$

If $f_1 = \cdots = f_n = f$, we will write $\Lambda(\mathbf{A}_n; f, \dots, f) = \Lambda(\mathbf{A}_n; f)$.
The main result in this section is the following.

Proposition 4.2. *Suppose that for some positive even integer $n \geq 1$ and small constant $\epsilon_0 > 0$, the following transverse correlation condition holds:*

$$\sup_{\mathbf{A}_n \in \mathfrak{A}_{tr}} |\Lambda(\mathbf{A}_n; \sigma_k)| \leq C_0(k, n, \epsilon_0) \tag{4.2}$$

Then the operator Φ_k^* defined in (3.9) satisfies the restricted strong-type estimate

$$\sup_{\Omega \subseteq [0,1]} \frac{\|\Phi_k^* \mathbf{1}_\Omega\|_n}{|\Omega|^{\frac{n-1}{n}}} \leq C \left[\max \left(\frac{2^n n^4 P_k^{\epsilon_0 - 1}}{(P_{k+1}\delta_{k+1})^{n-1}}, C_0(k, n, \epsilon_0) \right) \right]^{\frac{1}{n}}, \tag{4.3}$$

where $C > 0$ is an absolute constant independent of n, k and ϵ_0.

Remarks.

(1) Our goal will be to construct sets S_k for which $C_0(k, n, \epsilon_0)$, and indeed the right hand side of (4.3), decay exponentially in k. It will then follow from Corollary 3.5 that $\mathcal{M}$ is bounded from $L^p(\mathbb{R}) \to L^{(n-1)p}(\mathbb{R})$ for all $p > \frac{n}{n-1}$.

(2) The heuristic reason why (4.2) should hold is that, essentially, σ_k are highly oscillating random functions with $\int \sigma_k = 0$, so that two affine copies of σ_k with generic translation and scaling parameters should be close to orthogonal. In other words, there should be a lot of cancellation in the integral defining $\Lambda(\mathbf{A}_n; \sigma_k)$. The only exception to this is when relatively close correlations between two or more such copies are forced by a large number of internal tangencies.

In the proof of the proposition we will need the following trivial bound (ignoring all cancellation) on $\Lambda(\mathbf{A}_n; \sigma_k)$.

Lemma 4.3. *For all $k \geq 1$ and $\mathbf{A}_n \in \mathfrak{A}$, we have*

$$|\Lambda(\mathbf{A}_n; \sigma_k)| \leq \frac{2^{n+1}}{(P_{k+1}\delta_{k+1})^{n-1}}. \tag{4.4}$$

Proof. Recalling that $\sigma_k = \phi_{k+1} - \phi_k$, and expanding the product in $\Lambda(\mathbf{A}_n; \sigma_k)$, we arrive at the expression

$$|\Lambda(\mathbf{A}_n; \sigma_k)| \leq \sum_{\boldsymbol{\lambda} \in \{0,1\}^n} |\Lambda(\mathbf{A}_n; \phi_{k+\lambda_1}, \dots, \phi_{k+\lambda_n})|, \tag{4.5}$$

where $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_n)$. We treat each summand separately. Suppose first that $\lambda_{\ell_0} = 1$ for some ℓ_0. Since $\phi_{k+1} = (P_{k+1}\delta_{k+1})^{-1}\mathbf{1}_{S_{k+1}}$, we may estimate all factors pointwise by $(P_{k+1}\delta_{k+1})^{-1}$, so that

$$\begin{aligned} |\Lambda(\mathbf{A}_n; \phi_{k+\lambda_1}, \dots, \phi_{k+\lambda_n})| &\leq \frac{1}{(P_{k+1}\delta_{k+1})^n} \int \mathbf{1}_{S_{k+1}}\Big(\frac{z - c_{\ell_0}}{r_{\ell_0}}\Big)\, dz \\ &\leq \frac{2P_{k+1}\delta_{k+1}}{(P_{k+1}\delta_{k+1})^n} = \frac{2}{(P_{k+1}\delta_{k+1})^{n-1}}. \end{aligned} \tag{4.6}$$

If on the other hand $\lambda_\ell = 0$ for all ℓ, we have

$$\begin{aligned} |\Lambda(\mathbf{A}_n; \phi_{k+\lambda_1}, \dots, \phi_{k+\lambda_n})| &\leq \frac{1}{(P_k\delta_k)^n} \int \mathbf{1}_{S_k}\Big(\frac{z - c_1}{r_1}\Big)\, dz \\ &\leq \frac{2P_k\delta_k}{(P_k\delta_k)^n} = \frac{2}{(P_k\delta_k)^{n-1}} \leq \frac{2}{(P_{k+1}\delta_{k+1})^{n-1}}, \end{aligned} \tag{4.7}$$

where the last step uses the fact that the sequence $\{P_k\delta_k\}$ is monotone decreasing. Combining (4.5), (4.6) and (4.7) yields the desired conclusion. □

Proof of Proposition 4.2. For $x_1, x_2, \dots, x_n \in [0,1]^n$, let

$$\mathbf{A}(x_1, \dots, x_n) = \{(c(x_\ell), r(x_\ell)) : 1 \leq \ell \leq n\},$$

where $c(x_\ell), r(x_\ell)$ are chosen as in Section 3.2. Thus $\mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}$. Let $\Omega \subseteq [0,1]$, then

$$\begin{aligned} \|\Phi_k^*\mathbf{1}_\Omega\|_n^n &= \Big\|\int_\Omega V_{k,x}(\cdot)dx\Big\|_n^n \\ &= \int \prod_{j=1}^n \Big[\int_\Omega V_{k,x_j}(z)\, dx_j\Big]\, dz \\ &= \int_{\Omega^n} \Big[\int \prod_{j=1}^n V_{k,x_j}(z)\, dz\Big]\, dx_1 \cdots dx_n \\ &= \int_{\Omega^n} \Lambda(\mathbf{A}(x_1, \dots, x_n); \sigma_k)\, dx_1 \dots dx_n \\ &= \Big[\int_{\Theta_1} + \int_{\Theta_2}\Big] \Lambda(\mathbf{A}(x_1, \dots, x_n); \sigma_k)\, dx_1 \dots dx_n, \end{aligned}$$

where

$$\Theta_1 = \{(x_1, \cdots, x_n) \in \Omega^n : \ \mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}_{\text{int}}\},$$

$$\Theta_2 = \{(x_1, \cdots, x_n) \in \Omega^n : \ \mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}_{\text{tr}}\}.$$

We first estimate the integral on Θ_1. While the high order of internal tangency does not allow a better estimate than (4.4) on the integrand, the domain of the integration is restricted to a small set. Specifically, by Lemma 2.7 we have

$$\begin{aligned}\Theta_1 &\subseteq \bigcup_{1\leq \ell\neq\ell'\leq n}\left\{(x_1,\cdots,x_n)\in\Omega^n : |c(x_\ell)-c(x_{\ell'})|\leq \frac{80n(n-1)}{P_k^{1-\epsilon_0}}\right\}\\ &\subseteq \bigcup_{1\leq \ell\neq\ell'\leq n}\left\{(x_1,\cdots,x_n)\in\Omega^n : |x_\ell-x_{\ell'}|\leq \frac{160n(n-1)}{P_k^{1-\epsilon_0}}\right\},\end{aligned}$$

where we used that $|x_\ell - c(x_\ell)| \leq \delta_{k+1}^L \leq \delta_k^L \leq P_k^{-1+\epsilon_0}$. Combining this with Lemma 4.3 we obtain

$$\begin{aligned}&\int_{\Theta_1}\Lambda(\mathbf{A}(x_1,\dots,x_n);\sigma_k)\,dx_1\dots dx_n\\ &\leq \frac{2^{n+1}}{(P_{k+1}\delta_{k+1})^{n-1}}\sum_{1\leq\ell\neq\ell'\leq n}\int_{\Omega^{n-1}}\left[\int_{|x_\ell-x_{\ell'}|\leq\frac{160n(n-1)}{P_k^{1-\epsilon_0}}}dx_\ell\right]\prod_{j\neq\ell}dx_j \qquad (4.8)\\ &\leq \frac{2^{n+1}160n^2(n-1)^2}{(P_{k+1}\delta_{k+1})^{n-1}P_k^{1-\epsilon_0}}|\Omega|^{n-1}\leq 320\frac{2^n n^4 P_k^{\epsilon_0-1}}{(P_{k+1}\delta_{k+1})^{n-1}}|\Omega|^{n-1}.\end{aligned}$$

On the other hand, the desired estimate on the integral on Θ_2 follows directly from (4.2):

$$\int_{\Theta_2}\Lambda(\mathbf{A}(x_1,\dots,x_n);\sigma_k)\,dx_1\dots dx_n\leq C_0(k,n,\epsilon_0)|\Omega|^n\leq C_0(k,n,\epsilon_0)|\Omega|^{n-1}.$$

By (4.8), the conclusion follows. □

5. The random construction

5.1. **Selection of the sets $\{S_k\}$.** We are now ready to describe the probabilistic construction of the sets $\{S_k\}$ satisfying the transverse correlation condition (4.2) with acceptable constants $C_0(k,n,\epsilon_0)$. The basic procedure is as in Subsection 2.1, with the crucial additional point that the sequences $\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k$ are now randomized.

Here and in the sequel, $\{\epsilon_k : k\geq 1\}$ be a sequence of small constants with $0<\epsilon_k<\frac{1}{2}$, and $\{N_k : k\geq 1\}$ will be a nondecreasing sequence of large constants with N_1 large enough. Specific choices of both sequences will be made in the next section. Let $\mathbf{X}_1=\{X_1(i) : 1\leq i\leq N_1\}$ be a sequence of independent and identically distributed Bernoulli random variables:

$$X_1(i)=\begin{cases}1 & \text{with probability } p_1=N_1^{-\epsilon_1},\\ 0 & \text{with probability } 1-p_1.\end{cases}$$

Each realization of the Bernoulli sequence generates a possible candidate for S_1:

$$S_1 = S_1(\mathbf{X}_1) = \bigcup_{\substack{1 \le i \le N_1 \\ X_1(i)=1}} [\alpha_1(i), \alpha_1(i+1)].$$

In general, at the end of the k-th step, we will have selected a realization of $S_1, S_2, \cdots, S_k$. At step $k+1$, we will consider an iid Bernoulli sequence $\mathbf{Y}_{k+1} = \{Y_{k+1}(\bar{\mathbf{i}}) : \bar{\mathbf{i}} = (\mathbf{i}, i_{k+1}) \in \mathbb{I}_{k+1}\}$ with success probability $p_{k+1} = N_{k+1}^{-\epsilon_{k+1}}$, and set

$$\begin{aligned} \mathbf{X}_{k+1} &= \{X_{k+1}(\bar{\mathbf{i}}) : \bar{\mathbf{i}} \in \mathbb{I}_{k+1}\}, \quad X_{k+1}(\bar{\mathbf{i}}) = X_k(\mathbf{i}) Y_{k+1}(\bar{\mathbf{i}}), \\ P_{k+1} &= P_{k+1}(\mathbf{X}_{k+1}) = \sum_{\mathbf{i}} X_{k+1}(\bar{\mathbf{i}}), \\ Q_{k+1} &= P_k N_{k+1} p_{k+1} = P_k N_{k+1}^{1-\epsilon_{k+1}} \\ S_{k+1} &= S_{k+1}(\mathbf{X}_{k+1}) = \bigcup_{X_{k+1}(\bar{\mathbf{i}})=1} \left[\alpha_{k+1}(\bar{\mathbf{i}}), \alpha_{k+1}(\bar{\mathbf{i}}) + \delta_{k+1}\right]. \end{aligned} \tag{5.1}$$

At step $k+1$, the only random variables are the entries of the sequence $\mathbf{Y}_{k+1}$ (and hence $\mathbf{X}_{k+1}$), the sequence $\mathbf{X}_k$ having already been fixed at the previous step. Thus at step $k+1$, P_{k+1} is a random variable, whereas Q_{k+1} is not.

For every $k \ge 1$, we have a large sample space of possible choices for S_k. The goal of this section is to show that at every stage of the construction a selection can be made that satisfies a specified list of criteria, eventually leading up to (4.2). The main result in this section is the following.

Theorem 5.1. *Let $B > 0$ be an absolute constant, independent of k and n ($B = 10$ will work). Then there exists a sequence of sets $\{S_k\}$ constructed as described above (for some realization of the Bernoulli sequences $\mathbf{X}_1$, $\mathbf{Y}_k$) such that all of the following conditions hold:*

(a) $2^{-k} \prod_{j=1}^{k} N_j^{1-\epsilon_j} \le P_k \le 2^k \prod_{j=1}^{k} N_j^{1-\epsilon_j}$.

(b) $|P_k - Q_k| \le B\sqrt{Q_k}$.

(c) The transverse correlation condition (4.2) holds with $\epsilon_0 = \frac{1}{2}$ and

$$C_0\left(k, n, \frac{1}{2}\right) = 4^{n+2} n!\, B 2^{k(n+\frac{3}{2})} \times \left[\prod_{j=1}^{k} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right] N_{k+1}^{n\epsilon_{k+1}} \left[\ln\left(4^n n!\, B \prod_{j=1}^{k+1} N_j^{2Ln}\right)\right]^{1/2}. \tag{5.2}$$

(d) $$\sup_{\mathbf{i}: X_k(\mathbf{i})=1} \left| \sum_{i_{k+1}=1}^{N_{k+1}} \left(X_{k+1}(\bar{\mathbf{i}}) - p_{k+1}\right) \right| \le \left[8 N_{k+1}^{1-\epsilon_{k+1}} \ln(4BP_k)\right]^{\frac{1}{2}}.$$

Corollary 5.2. *Let $\{S_k\}$ be the sequence of sets given by Theorem 5.1. Then:*
(a) The associated operators Φ_k^* defined in (3.9) satisfy the restricted strong-type estimate

$$\sup_{\Omega\subseteq[0,1]} \frac{\|\Phi_k^* \mathbf{1}_\Omega\|_n}{|\Omega|^{\frac{n-1}{n}}} \leq C(n!\,B)^{1/n} 2^{k(1+\frac{3}{2n})} \left[\prod_{j=1}^{k} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right]^{1/n} N_{k+1}^{\epsilon_{k+1}} \times \left[\ln\left(4^n n!\,B \prod_{j=1}^{k+1} N_j^{2Ln}\right)\right]^{1/2n}, \tag{5.3}$$

where $C > 0$ is an absolute constant independent of n and k.
(b) Assume that the parameters N_k, ϵ_k have been set so that

$$\sup_{k\geq 1} \frac{2^{(5+\gamma)k}\ln(M_k)}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \frac{1}{32}. \tag{5.4}$$

for some $\gamma > 0$. Then we further have

$$\sup_{k':k'\geq k} \sum_{\mathbf{i}:X_k(\mathbf{i})=1} \left|\int_{I_k(\mathbf{i})} (\phi_{k'} - \phi_k)dx\right| \leq \frac{2B}{1-2^{-\gamma/2}} 2^{-k\gamma/2}. \tag{5.5}$$

Consequently, the densities ϕ_k converge weakly to a probability measure μ supported on $S = \bigcap_{k=1}^\infty S_k$.

Proof. Part (a) follows from Proposition 4.2. By Theorem 5.1(a), we have

$$\frac{2^n n^4 P_k^{-1/2}}{(P_{k+1}\delta_{k+1})^{n-1}} \leq 2^{2n-1} n^4 2^{k(n-\frac{1}{2})} \left[\prod_{j=1}^{k} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right] N_{k+1}^{(n-1)\epsilon_{k+1}}.$$

Plugging this together with (5.2) into (4.3), we get (5.3). The inequality (5.5) follows from Theorem 5.1(d); we defer the proof of this to Subsection 5.5. Since (5.5) implies in particular that (2.7) holds, the convergence statement follows from Lemma 2.2. □

The proof of Theorem 5.1 is arranged as follows. Note that parts (a)–(b) concern the set S_k, whereas (c)–(d) are properties of S_{k+1}; accordingly, we will say that S_k obeys (a)–(b) if (a)–(b) hold as stated above, and that S_k obeys (c)–(d) if (c)–(d) hold with k replaced by $k-1$. Fix B as in the statement of the theorem, and choose N_1 sufficiently large relative to B. To initialize, we prove that S_1 obeys (a)–(b) with probability at least $1 - B^{-1}$, in particular there exists a choice of S_1 with these properties. Assume now that we have already chosen $S_1, \ldots, S_k$ obeying (a)–(d) (where (c)–(d) hold vacuously for S_1), and consider the space of all possible choices of S_{k+1}. We will prove in Subsections 5.3–5.5 that each of (a)-(b) and (d) fails to hold for S_{k+1} with probability at most B^{-1}, and the event that (a)-(b) hold but (c) fails has probability at most B^{-1}. Thus there is a probability of at least $1-4B^{-1}$ that S_{k+1} obeys all of (a)–(d). Fix this choice of S_{k+1}, and continue by induction.

We emphasize here that we do not attempt to randomize the entire sequence of steps simultaneously. By the $(k+1)$-th stage of the iteration we have restricted attention to a *deterministic* sequence $\mathbf{X}_k$, with the probabilistic machinery being applied to the random sequence $\mathbf{X}_{k+1}$ conditional on the previously obtained $\mathbf{X}_k$. As a consequence, we ensure the existence of *some* sequence of desirable sets, but (in contrast to e.g. Salem's construction in [34]) we can make no claim as to its frequency of occurrence among all possible iterative constructions subject to the given parameters.

5.2. **Two large deviation inequalities.** In this subsection, we record two large deviation inequalities widely used in probability theory that will play a key role in the sequel. The first one is a version of Bernstein's inequality borrowed from [17]. We will use it here much as we did in [24].

Lemma 5.3 (Bernstein's inequality)**.** *Let $Z_1, \dots, Z_m$ be independent random variables with $|Z_j| \leq 1$, $\mathbb{E}Z_j = 0$ and $\mathbb{E}|Z_j|^2 = \sigma_j^2$. Let $\sum \sigma_j^2 \leq \sigma^2$, and assume that $\sigma^2 \geq 6m\lambda$. Then*

$$\mathbb{P}\Big(\Big|\sum_1^n Z_j\Big| \geq m\lambda\Big) \leq 4e^{-m^2\lambda^2/8\sigma^2}. \tag{5.6}$$

We will also need a similar inequality for random variables which are not independent, but instead are allowed to interact with one another to a limited extent. The exact statement that we need is contained in Lemma 5.4 below. Recall that a sequence $U_1, U_2, \dots$ of random variables is a *martingale* if $\mathbb{E}|U_j| < \infty$ for all j and

$$\mathbb{E}(U_{m+1}|U_1, \dots, U_m) = U_m, \quad m = 1, 2, \dots.$$

Lemma 5.4 (Azuma's inequality, [40] or [1], p. 95)**.** *Suppose that $\{U_k : k = 0, 1, 2, \cdots\}$ is a martingale and $\{c_k : k \geq 0\}$ is a sequence of positive numbers such that $|U_{k+1} - U_k| \leq c_k$ a.s. Then for all integers $m \geq 1$ and all $\lambda \in \mathbb{R}$,*

$$\mathbb{P}(|U_m - U_0| \geq \lambda) \leq 2\exp\left(-\frac{\lambda^2}{2\sum_{k=1}^m c_k^2}\right).$$

5.3. **Proof of Theorem 5.1 (a)-(b).** For $k = 1$, let N_1 be chosen so that $6B \leq N_1^{(1-\epsilon_1)/2}$. By Bernstein's inequality (Lemma 5.3) with $Z_i = X_1(i) - p_1$, $m = N_1$, $\sigma^2 = N_1 p_1 = N_1^{1-\epsilon_1}$ and $\lambda = BN_1^{-(1+\epsilon_1)/2}$, we have

$$\begin{aligned}\mathbb{P}\left(|P_1 - N_1p_1| > BN_1^{\frac{1-\epsilon_1}{2}}\right) &= \mathbb{P}\left(\Big|\sum_{i=1}^{N_1}\big[X_1(i) - p_1\big]\Big| > BN_1^{\frac{1-\epsilon_1}{2}}\right) \\ &\leq 4e^{-\frac{B^2}{8}}.\end{aligned}$$

Since $Q_1 = N_1 p_1 = N_1^{1-\epsilon_1}$, this shows that the inequality in (b) holds for $k = 1$ with probability $\geq 1 - 4e^{-B^2/8}$. Further, for any $\mathbf{X}_1$ that satisfies (b), the estimate

$$\frac{1}{2} N_1^{1-\epsilon_1} \leq N_1^{1-\epsilon_1} \left(1 - B N_1^{-\frac{1-\epsilon_1}{2}}\right) \leq P_1 \leq N_1^{1-\epsilon_1} \left(1 + B N_1^{-\frac{1-\epsilon_1}{2}}\right) \leq 2 N_1^{1-\epsilon_1}$$

holds. Assume now that $\mathbf{X}_k$ has been selected so that (a) and (b) hold for some $k \geq 1$. The random variables

$$Z_{\mathbf{i}} = \frac{1}{N_{k+1}} \sum_{i_{k+1}=1}^{N_{k+1}} \left[Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1} \right] ,$$

indexed by $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$, are iid with mean zero and variance $p_{k+1}(1 - p_{k+1})/N_{k+1}$. Hence Lemma 5.3 applies with $m = P_k$, $\sigma^2 = P_k p_{k+1}/N_{k+1}$, and $\lambda = B\sqrt{p_{k+1}/(P_k N_{k+1})}$, yielding

$$\begin{aligned}
&\mathbb{P}\left(\left| P_{k+1} - Q_{k+1} \right| > B\sqrt{Q_{k+1}} \right) \\
&= \mathbb{P}\left(\left| \sum_{\mathbf{i} \in \mathbb{I}_k} X_k(\mathbf{i}) \sum_{i_{k+1}=1}^{N_{k+1}} \left[Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1} \right] \right| > B\sqrt{Q_{k+1}} \right) \\
&= \mathbb{P}\left(\left| \sum_{X_k(\mathbf{i})=1} Z_{\mathbf{i}} \right| > P_k \lambda \right) \\
&\leq 4 e^{-\frac{B^2}{8}} < B^{-1}.
\end{aligned}$$

Thus with large probability, (b) holds with k replaced by $k + 1$. Further by induction hypothesis (a) and the definition of Q,

$$2^{-k} \prod_{j=1}^{k+1} N_j^{1-\epsilon_j} \leq Q_{k+1} \leq 2^k \prod_{j=1}^{k+1} N_j^{1-\epsilon_j}, \tag{5.7}$$

which in particular implies that $Q_{k+1} \geq 2^{-k} N_1^{(k+1)(1-\epsilon_1)} \geq 4B^2$ if N_1 is chosen sufficiently large. Thus for any $\mathbf{X}_{k+1}$ satisfying (b),

$$\frac{1}{2} \leq 1 - \frac{B}{\sqrt{Q_{k+1}}} \leq \frac{P_{k+1}}{Q_{k+1}} \leq 1 + \frac{B}{\sqrt{Q_{k+1}}} \leq 2,$$

which coupled with (5.7) proves the inductive step for (a).

5.4. **Proof of Theorem 5.1(c).** We now begin the proof of (c), which is substantially more difficult. The strategy of the proof is outlined in §5.4.1 below, the execution of the various steps being relegated to the later parts of this subsection.

5.4.1. *Steps of the proof.* Throughout this section we will assume that S_k has been selected so as to obey Theorem 5.1(a)-(b). We begin by replacing the measure $\sigma_k = \phi_{k+1} - \phi_k$ in (4.2) by $\overline{\sigma}_k$, where

$$\overline{\sigma}_k(z) = \frac{1}{Q_{k+1}\delta_{k+1}} \mathbf{1}_{S_{k+1}}(z) - \frac{1}{P_k\delta_k}\mathbf{1}_{S_k}(z). \tag{5.8}$$

This renders the expression in (4.2) more amenable to the application of the large deviation inequalities from Subsection 5.2, at the expense of a harmless error term that we estimate below.

Lemma 5.5 (Step 1). *Assume that Theorem 5.1(a)–(b) holds at step $k+1$. For any $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \le \ell \le n\} \in \mathfrak{A}$,*

$$|\Lambda(\mathbf{A}_n; \sigma_k)| \le |\Lambda(\mathbf{A}_n; \overline{\sigma}_k)| + 2^{2n+1} B 2^{k(n+\frac{3}{2})} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right]. \tag{5.9}$$

In particular, this means that for any $0 < \epsilon_0 < 1$, (4.2) holds with

$$C_0(k, n, \epsilon_0) = \sup_{A_n \in \mathfrak{A}_{tr}} \left|\Lambda(\mathbf{A}_n; \overline{\sigma}_k)\right| + 2^{2n+1} B 2^{k+\frac{3}{2}} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right]. \tag{5.10}$$

Proposition 5.6 (Step 2). *Suppose that there is a constant $C_1(k, n, \epsilon_0)$ such that for all $\mathbf{A}_n \in \mathfrak{A}_{tr}$ the following estimate holds:*

$$\left|\sum_{\mathbf{I} \in \mathbb{F}_{tr}} \prod_{\ell=1}^{n} X_k(\mathbf{i}_\ell) \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^{n} \Big(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1}\Big) \cdot \left|\bigcap_{\ell=1}^{n} (c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell))\right|\right| \le C_1(k, n, \epsilon_0) \tag{5.11}$$

where $\mathbf{I} = (\mathbf{i}_1, \cdots, \mathbf{i}_n)$, $\bar{\mathbf{i}}_\ell = (\mathbf{i}_\ell, i_{k+1,\ell})$, and $\boldsymbol{\iota} = (i_{k+1,1}, \ldots, i_{k+1,n})$ denotes the n-vector whose entries are the $(k+1)$-th entries of $\bar{\mathbf{i}}_1, \cdots, \bar{\mathbf{i}}_n$ respectively (thus $\boldsymbol{\iota}$ ranges over the set $\{1, 2 \cdots, N_{k+1}\}^n$). Then

$$\begin{aligned} \sup_{\mathbf{A}_n \in \mathfrak{A}_{tr}} |\Lambda(\mathbf{A}_n; \overline{\sigma}_k)| \le & C_1(k, n, \epsilon_0) 2^{kn} \left[\prod_{j=1}^{k+1} N_j^{n\epsilon_j}\right] \\ & + 2^{k(n+1-\epsilon_0)+3} \left[\prod_{j=1}^{k} N_j^{-\epsilon_0+\epsilon_j(n+\epsilon_0-1))}\right] N_{k+1}^{n\epsilon_{k+1}}. \end{aligned} \tag{5.12}$$

Proposition 5.7 (Step 3). *The event that (5.11) holds with*

$$C_1(k, n, \epsilon_0) = 4^n n! \left[\prod_{j=1}^{k} N_j^{-\frac{1+\epsilon_j}{2}}\right] \times \left[\ln\left(4^n n! B \prod_{j=1}^{k+1} N_j^{2Ln}\right)\right]^{\frac{1}{2}} \tag{5.13}$$

has probability at least $1 - B^{-1}$.

Assume for now the claims in steps 1–3.

Conclusion of the proof of Theorem 5.1 (c). Of the three estimates (5.10), (5.12), and (5.13), the first one holds with probability at least $1 - B^{-1}$ (Subsection 5.3), the second one holds always, and the third one holds with probability at least $1 - B^{-1}$ as indicated in the last proposition. Combining these estimates yields that

$$\begin{aligned} &|\Lambda(\mathbf{A}_n; \sigma_k)| \\ &\qquad \leq 2^{2n+1} B 2^{k(n+\frac{3}{2})} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})} \right] \\ &\qquad + 2^{k(n+1-\epsilon_0)+3} \left[\prod_{j=1}^{k} N_j^{-\epsilon_0+\epsilon_j(n+\epsilon_0-1)} \right] N_{k+1}^{n\epsilon_{k+1}} \\ &\qquad + 4^n n!\, 2^{k(n+\frac{1}{2})+\frac{1}{2}} \left[\prod_{j=1}^{k} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})} \right] N_{k+1}^{n\epsilon_{k+1}} \times \left[\ln\left(4^n n!\, B \prod_{j=1}^{k+1} N_j^{2Ln} \right) \right]^{\frac{1}{2}}, \end{aligned}$$

with probability at least $1 - 2B^{-1}$. Plugging in $\epsilon_0 = \frac{1}{2}$, we see after some simple algebra that in this event $|\Lambda(\mathbf{A}_n; \sigma_k)|$ is bounded as indicated in (5.2). □

5.4.2. *Proof of Lemma 5.5.* It suffices to prove (5.9), since (5.10) follows directly from it. We write $\sigma_k = \overline{\sigma}_k + e_k$, where $\overline{\sigma}_k$ is as in (5.8) so that

$$e_k(z) = \left[\frac{1}{P_{k+1}\delta_{k+1}} - \frac{1}{Q_{k+1}\delta_{k+1}} \right] \mathbf{1}_{S_{k+1}}(z),$$

Then

$$\Lambda(\mathbf{A}_n; \sigma_k) = \Lambda(\mathbf{A}_n; \overline{\sigma}_k) + E_k, \text{ where}$$

$$E_k = \sum_{\substack{\boldsymbol{\lambda} \in \{0,1\}^n \\ \lambda_1 + \cdots + \lambda_n \geq 1}} \Lambda(\mathbf{A}_n; u_{\lambda_1}, \ldots, u_{\lambda_n}) \text{ with } u_\lambda = \begin{cases} \overline{\sigma}_k & \text{if } \lambda = 0, \\ e_k & \text{if } \lambda = 1. \end{cases}$$

We need to show that $|E_k|$ is bounded by the quantity in (5.9).

We observe that by the definition of Q_k in (5.1) and Theorem 5.1(a) at step k,

$$\begin{aligned} |\overline{\sigma}_k(z)| &\leq \left[\frac{1}{Q_{k+1}\delta_{k+1}} + \frac{1}{P_k \delta_k} \right] \mathbf{1}_{S_k}(z) = \left[\frac{N_{k+1}^{\epsilon_{k+1}}}{P_k \delta_k} + \frac{1}{P_k \delta_k} \right] \mathbf{1}_{S_k}(z) \\ &\leq 2 \frac{N_{k+1}^{\epsilon_{k+1}}}{P_k \delta_k} \mathbf{1}_{S_k}(z) \leq 2^{k+1} \prod_{j=1}^{k+1} N_j^{\epsilon_j} \mathbf{1}_{S_k}(z), \end{aligned}$$

whereas by Theorem 5.1(b) at step $k+1$,

$$\begin{aligned}|e_k(z)| &\leq \frac{|Q_{k+1}-P_{k+1}|}{P_{k+1}Q_{k+1}\delta_{k+1}}\mathbf{1}_{S_{k+1}}(z) \leq B\frac{1}{P_{k+1}\sqrt{Q_{k+1}}\delta_{k+1}}\mathbf{1}_{S_{k+1}}(z)\\ &\leq B2^{\frac{3k}{2}+1}\left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\frac{3\epsilon_j}{2}}\right]\mathbf{1}_{S_{k+1}}(z).\end{aligned}$$

Therefore for any $\boldsymbol{\lambda} \in \{0,1\}^n$ with $\lambda_1 + \cdots + \lambda_n \geq 1$, there exists an index $1 \leq \ell_0 \leq n$ such that

$$\operatorname{supp}\left[\prod_{\ell=1}^{n} u_{\lambda_\ell}\Big(\frac{\cdot - c_\ell}{r_\ell}\Big)\right] \subseteq c_{\ell_0} + r_{\ell_0}S_{k+1}.$$

Note also that the estimate on $|e_k|$ is better than the estimate on $|\overline{\sigma}_k|$ if $N_j \geq N$ and N has been chosen large enough. Hence

$$\begin{aligned}&|\Lambda(\mathbf{A}_n; u_{\lambda_1}, \ldots, u_{\lambda_n})|\\ &\quad\leq \Big(2^{k+1}\prod_{j=1}^{k+1} N_j^{\epsilon_j}\Big)^{n-1}\Big(B2^{\frac{3k}{2}+1}\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\frac{3\epsilon_j}{2}}\Big)\big|(c_{\ell_0}+r_{\ell_0}S_{k+1})\big|\\ &\quad\leq 2^n B 2^{k(n+\frac{1}{2})}\left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n+\frac{1}{2})}\right]P_{k+1}\delta_{k+1}\\ &\quad\leq 2^{n+1}B2^{k(n+\frac{3}{2})}\left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right].\end{aligned}$$

Since the total number of terms in the sum representing E_k is 2^n-1, the desired conclusion follows. ☐

5.4.3. *Proof of Proposition 5.6.* We need to estimate

$$\Lambda(\mathbf{A}_n; \overline{\sigma}_k) = \int \prod_{\ell=1}^{n} \overline{\sigma}_k\Big(\frac{z-c_\ell}{r_\ell}\Big)dz \tag{5.14}$$

for $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\} \in \mathfrak{A}_{\mathrm{tr}}$. We start by rewriting $\overline{\sigma}_k$ as

$$\begin{aligned}\overline{\sigma}_k(z) &= \frac{1}{Q_{k+1}\delta_{k+1}} \sum_{X_{k+1}(\bar{\mathbf{i}})=1} \mathbf{1}_{I_{k+1}(\bar{\mathbf{i}})}(z) - \frac{1}{P_k\delta_k} \sum_{X_k(\mathbf{i})=1} \mathbf{1}_{I_k(\mathbf{i})}(z) \\ &= \frac{1}{Q_{k+1}\delta_{k+1}} \sum_{X_k(\mathbf{i})=1} \sum_{i_{k+1}=1}^{N_{k+1}} (Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}) \mathbf{1}_{I_{k+1}(\bar{\mathbf{i}})}(z) \\ &= \frac{1}{Q_{k+1}\delta_{k+1}} \sum_{\mathbf{i}\in\mathbb{I}_k} X_k(\mathbf{i}) \sum_{i_{k+1}=1}^{N_{k+1}} (Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}) \mathbf{1}_{I_{k+1}(\bar{\mathbf{i}})}(z).\end{aligned}$$

Hence

$$\prod_{\ell=1}^{n} \overline{\sigma}_k\Big(\frac{z-c_\ell}{r_\ell}\Big) = \frac{1}{(Q_{k+1}\delta_{k+1})^n} \sum_{\mathbf{I}\in\mathbb{I}_k^n} \Bigg[\prod_{\ell=1}^{n} X_k(\mathbf{i}_\ell) \times \sum_{\boldsymbol{\iota}} \Bigg(\prod_{\ell=1}^{n} \big(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1}\big) \Bigg) \mathbf{1}_{\bigcap_{\ell=1}^{n}(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell))}(z) \Bigg] \tag{5.15}$$

where $\mathbf{I}$ and $\boldsymbol{\iota}$ are as in Proposition 5.6. Since

$$\bigcap_{\ell=1}^{n}(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)) \subseteq \bigcap_{\ell=1}^{n}(c_\ell + r_\ell I_k(\mathbf{i}_\ell)),$$

a summand in (5.15) is nonzero only if the n-fold intersection on the right hand side above is nonempty, i.e., only if $\mathbf{I} \in \mathbb{F} = \mathbb{F}[n,k;\mathbf{A}_n]$. Splitting $\mathbb{F}$ further into $\mathbb{F}_{\text{int}}$ and $\mathbb{F}_{\text{tr}}$ as in Subsection 2.4, we find that

$$\begin{aligned}\Lambda(\mathbf{A}_n;\overline{\sigma}_k) &= \frac{1}{(Q_{k+1}\delta_{k+1})^n} \Bigg\{ \sum_{\mathbf{I}\in\mathbb{F}_{\text{int}}} + \sum_{\mathbf{I}\in\mathbb{F}_{\text{tr}}} \Bigg\} \Bigg[\prod_{\ell=1}^{n} X_k(\mathbf{i}_\ell) \\ &\qquad \times \sum_{\boldsymbol{\iota}} \Bigg(\prod_{\ell=1}^{n} \big(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1}\big) \Bigg) \Bigg| \bigcap_{\ell=1}^{n}(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)) \Bigg| \Bigg] \\ &:= \Xi_{\text{int}} + \Xi_{\text{tr}}.\end{aligned}$$

We treat these two sums separately.

Since $\mathbf{A}_n \in \mathfrak{A}_{\rm tr}$, we have $\#(\mathbb{F}_{\rm int}) < P_k^{1-\epsilon_0}$, therefore

$$\begin{aligned}\left|\Xi_{\rm int}\right| &\leq \frac{1}{(Q_{k+1}\delta_{k+1})^n} \sum_{\mathbf{I}\in\mathbb{F}_{\rm int}} \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^{n} \left|Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1}\right| \times \left|\bigcap_{\ell=1}^{n}\big(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)\big)\right| \\ &\leq \frac{1}{(Q_{k+1}\delta_{k+1})^n} \sum_{\mathbf{I}\in\mathbb{F}_{\rm int}} \sum_{\boldsymbol{\iota}} \left|\bigcap_{\ell=1}^{n}\big(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)\big)\right| \\ &\leq \frac{4P_k^{1-\epsilon_0} N_{k+1}\delta_{k+1}}{(Q_{k+1}\delta_{k+1})^n} \\ &\leq 2^{k(n+1-\epsilon_0)+3}\left[\prod_{j=1}^{k} N_j^{-\epsilon_0+\epsilon_j(n+\epsilon_0-1)}\right] N_{k+1}^{n\epsilon_{k+1}},\end{aligned}$$

where at the third step we have used Lemma 5.8 below to estimate the number of non-zero summands in the inner sum on the second line by $4N_{k+1}$.

On the other hand, by (5.11)

$$\left|\Xi_{\rm tr}\right| \leq \frac{C_1(k,n,\epsilon_0)}{(Q_{k+1}\delta_{k+1})^n} \leq C_1(k,n,\epsilon_0) 2^{kn}\left[\prod_{j=1}^{k+1} N_j^{n\epsilon_j}\right].$$

Combining the two estimates, we get (5.12). ☐

Lemma 5.8. *For each fixed* $\mathbf{I} \in \mathbb{I}_k$, *there are at most* $4N_{k+1}$ *distinct choices of* $\boldsymbol{\iota} = (i_{k+1,1}, \cdots, i_{k+1,n})$ *such that*

$$\bigcap_{\ell=1}^{n}\big(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)\big) \neq \emptyset. \tag{5.16}$$

Proof. Suppose that (5.16) holds, then

$$\big(\bar{\mathbf{i}}_1, \cdots, \bar{\mathbf{i}}_n\big) \in \mathbb{F}[n, k+1, \mathbf{A}_n]. \tag{5.17}$$

If $i_{k+1,1}$ is fixed, this fixes $\bar{\mathbf{i}}_1$ and it follows from Lemma 2.4 that the number of possible tuples $(\bar{\mathbf{i}}_2, \cdots, \bar{\mathbf{i}}_n)$ such that (5.17) holds is at most 4. Hence the number of possible choices of $(i_{k+1,2}, \ldots, i_{k+1,n})$ is at most 4. This proves the claim, since there are at most N_{k+1} choices of $i_{k+1,1}$. ☐

5.4.4. *Proof of Proposition 5.7.* The heart of the proof is a convenient re-indexing of the sum in (5.11) that permits the application of Azuma's inequality from Subsection 5.2. The next lemma is a preparatory step for arranging this sum in the desired form. The lemma following it completes the verification of the martingale criterion.

Lemma 5.9. *Fix* $\mathbf{A}_n \in \mathfrak{A}$. *Then there is a decomposition of* $\mathbb{F}_{tr}$ *into at most* $4^{n-1}n!$ *subclasses such that*

(a) For all $1 \leq \ell \leq n$, π_ℓ *is injective on each subclass.*

(b) For each subclass, there is a permutation ρ of $\{1, \dots, n\}$ such that

$$\alpha_k(\mathbf{i}_{\rho(1)}) < \cdots < \alpha_k(\mathbf{i}_{\rho(n)}) \tag{5.18}$$

for all $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_n)$ in the subclass.

Proof. Let $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{F}_{\rm tr}$, then for all $\ell \neq \ell'$ we have

$$|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{i}_{\ell'})| > 4\delta_k. \tag{5.19}$$

Thus for every $\mathbf{I}$, all $\alpha_k(\mathbf{i}_\ell)$, $1 \le \ell \le n$, are distinct, and in particular there is a permutation $\rho = \rho(\mathbf{I})$ such that (5.18) holds for that $\mathbf{I}$. Let $\mathcal{F}_\rho = \{\mathbf{I} : \ \rho(\mathbf{I}) = \rho\}$ for each such permutation. By Corollary 2.5, each $\mathcal{F}_\rho$ can be decomposed further into at most 4^{n-1} subsets on which all the projections π_ℓ are injective. □

By a slight abuse of notation, we will continue to use $\mathcal{F}_\rho$ to denote a subclass of $\mathbb{F}_{\rm tr}$ such that both (i) and (ii) hold for the permutation ρ. In view of Lemma 5.9, it suffices to estimate

$$\left| \sum_{\mathbf{I} \in \mathcal{F}_\rho} \prod_{\ell=1}^n X_k(\mathbf{i}_\ell) \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^n \Big(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1} \Big) \cdot \left| \bigcap_{\ell=1}^n \big(c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell) \big) \right| \right| \tag{5.20}$$

for each such $\mathcal{F}_\rho$.

Observe that by part (a) of Lemma 5.9, the index $\mathbf{I}$ in the outer sum is in fact determined uniquely by $\mathbf{i}_{\rho(n)} = \pi_{\rho(n)}(\mathbf{I})$. In other words, the elements $\{\alpha_k(\mathbf{i}_{\rho(n)}) : \mathbf{I} \in \mathcal{F}_\rho\}$ are all distinct. Furthermore, the only indices that contribute to (5.20) are those with $\prod_{\ell=1}^n X_k(\mathbf{i}_\ell) = 1$. Accordingly, let

$$\mathcal{J} = \left\{ \mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathcal{F}_\rho : \ \prod_{\ell=1}^n X_k(\mathbf{i}_\ell) = 1 \right\},$$

and let us arrange the elements of $\mathcal{J}$ in a sequence $\{\mathbf{I}(j) = (\mathbf{i}_1(j), \dots, \mathbf{i}_n(j)) \colon j = 1, \dots, T\}$ so that

$$\alpha_k(\mathbf{i}_{\rho(n)}(1)) < \cdots < \alpha_k(\mathbf{i}_{\rho(n)}(T)). \tag{5.21}$$

For $1 \le j \le T$, we define

$$W_j = \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^n \Big(Y_{k+1}(\mathbf{i}_\ell(j), i_{k+1,\ell}) - p_{k+1} \Big) \left| \bigcap_{\ell=1}^n \big(c_\ell + r_\ell I_{k+1}(\mathbf{i}_\ell(j), i_{k+1,\ell}) \big) \right|, \tag{5.22}$$

where the summation index $\boldsymbol{\iota} = (i_{k+1,1}, \dots, i_{k+1,n})$ is as in the statement of Proposition 5.6, hence ranges over all vectors in $\{1, \dots, N_{k+1}\}^n$. We also let $W_0 = 0$. Then the sum in (5.20) is simply $W_1 + \cdots + W_T$.

Lemma 5.10. *$\{W_j : \ 0 \le j \le T\}$ is a martingale difference sequence (i.e. the sequence $\{W_1 + \cdots + W_m : \ 1 \le m \le T\}$ is a martingale), with $|W_j| \le 4\delta_k$ for all $1 \le j \le T$.*

Proof. We need to prove that $\mathbb{E}(W_m | W_1, \dots, W_{m-1}) = 0$. It suffices to demonstrate that the random variables $Y_{k+1}(\mathbf{i}_{\rho(n)}(m), \cdot)$ are

(i) independent of all $Y_{k+1}(\mathbf{i}_{\rho(\ell)}(m), \cdot)$ with $\ell < n$,
(ii) independent of all W_j with $j < m$.

Once we have this, the desired conclusion follows by setting $\mathcal{W}$ to be the collection of random variables in (i) and (ii) above, and $\mathcal{W}' = \mathcal{W} \setminus \{W_1, \cdots, W_{m-1}\}$, so that

$$\begin{aligned}\mathbb{E}(W_m | W_1, \cdots W_{m-1}) &= \mathbb{E}_{\mathcal{W}'}\Big[\mathbb{E}\big(W_m \big| \mathcal{W}\big)\Big] \\ &= \mathbb{E}_{\mathcal{W}'}\Big[\sum_{\boldsymbol{\iota}} F_{\boldsymbol{\iota},m}(\mathcal{W}) \mathbb{E}\left(Y_{k+1}(\mathbf{i}_{\rho(n)}(m), i_{k+1,\rho(n)}) - p_{k+1}\right)\Big] \\ &= 0.\end{aligned}$$

Here $\{F_{\boldsymbol{\iota},m}\}$ are measurable functions of $\mathcal{W}$ specified by the expression (5.22) for W_m but whose exact functional forms are unimportant.

By (5.18), we have

$$\alpha_k(\mathbf{i}_{\rho(\ell)}(m)) < \alpha_k(\mathbf{i}_{\rho(n)}(m)), \ \ell < n,$$

which implies immediately the first claim (i). It remains to prove (ii). Observe that W_j depends only on $Y_{k+1}(\mathbf{i}_\ell(j), \cdot)$, $1 \leq \ell \leq n$, hence it suffices to prove that

$$\mathbf{i}_{\rho(n)}(m) \notin \Big\{\mathbf{i}_{\rho(\ell)}(j) : \ 1 \leq \ell \leq n, \ 1 \leq j < m\Big\}.$$

But this follows from

$$\alpha_k(\mathbf{i}_{\rho(\ell)}(j)) \leq \alpha_k(\mathbf{i}_{\rho(n)}(j)) < \alpha_k(\mathbf{i}_{\rho(n)}(m)), \ \ell \leq n, \ j < m$$

where we used (5.18) again and then (5.21).

It remains to prove the almost sure bound on W_j. Indeed, by Lemma 5.8 the number of summands in (5.22) that make a non-zero contribution to W_j is bounded by $4N_{k+1}$. Since the size of each summand is bounded by δ_{k+1}, it follows that $|W_j| \leq 4N_{k+1}\delta_{k+1} = 4\delta_k$, as claimed. □

Conclusion of the proof of Proposition 5.7. In light of Lemma 5.10, we apply Azuma's inequality (Lemma 5.4) to the martingale sequence $U_j = W_1 + \cdots + W_j$, with $c_j = 4\delta_k$ and

$$\lambda = 4\delta_k \sqrt{2P_k} \sqrt{\ln(4^n n! B \delta_{k+1}^{-2Ln})},$$

and obtain

$$\mathbb{P}\big((5.20) > \lambda\big) \leq 2\exp\left(-\frac{\lambda^2}{32\delta_k^2 T}\right) \leq 2\exp\left(-\frac{\lambda^2}{32\delta_k^2 P_k}\right) \leq \frac{\delta_{k+1}^{2Ln}}{4^{n-1}n! B}.$$

Since there are at most $4^{n-1}n!$ classes $\mathcal{F}_\rho$, the probability that (5.20)$> \lambda$ for at least of them is bounded by $B^{-1}\delta_{k+1}^{2Ln}$. Summing over such classes, we see that

$$\mathbb{P}\Big(\text{LHS of } (5.11) > 4^{n-1} n!\, \lambda\Big) \leq \frac{\delta_{k+1}^{2Ln}}{B}.$$

Finally, since $\#(\mathfrak{A}) = \delta_{k+1}^{-2Ln}$, there is a probability of at least $1 - \frac{1}{B}$ that (5.11) holds for every $\mathbf{A} \in \mathfrak{A}_{\rm tr}$ with

$$C_1(k, n, \epsilon_0) = 4^{n-1} n!\, \lambda = 4^n n!\, \delta_k \sqrt{2P_k} \sqrt{\ln(4^n n! B \delta_{k+1}^{-2Ln})}.$$

By Theorem 5.1(a) at step k,

$$C_1(k, n, \epsilon_0) \leq 4^n n!\, 2^{\frac{k+1}{2}} \left[\prod_{j=1}^{k} N_j^{-\frac{1+\epsilon_j}{2}} \right] \times \left[\ln \left(4^n n!\, B \prod_{j=1}^{k+1} N_j^{2Ln} \right) \right]^{\frac{1}{2}}.$$

This completes the proof of the proposition. □

5.5. **Existence of the limiting measure.**

5.5.1. *Proof of Theorem 5.1(d).* Let $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$. Applying Bernstein's inequality to the random variables $X_{k+1}(\bar{\mathbf{i}}) - p_{k+1} = Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}$, with $\sigma^2 = N_{k+1} p_{k+1}$ and $\lambda = (8 p_{k+1} \ln(4BP_k)/N_{k+1})^{\frac{1}{2}}$, we obtain

$$\mathbb{P}\left(\left| \sum_{i_{k+1}=1}^{N_{k+1}} \left[Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1} \right] \right| > N_{k+1} \lambda \right) \leq 4 \exp \left[- \frac{N_{k+1}^2 \lambda^2}{8 N_{k+1} p_{k+1}} \right] = \frac{1}{BP_k}.$$

Since there are P_k-many such choices of $\mathbf{i}$, we find that Theorem 5.1(d) holds with probability at least $1 - \frac{1}{B}$, as claimed.

5.5.2. *Proof of Corollary 5.2(b).*

Lemma 5.11. *Assume that (5.4) and Theorem 5.1(d) hold for all k. Then for all $k \geq 1$, $m \geq 0$ and every $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$,*

$$2^{-m} \left[\prod_{r=1}^{m} N_{k+r}^{1-\epsilon_{k+r}} \right] \leq \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \leq 2^m \left[\prod_{r=1}^{m} N_{k+r}^{1-\epsilon_{k+r}} \right], \tag{5.23}$$

where the sum is taken over all m-dimensional multi-indices $\mathbf{j}$ such that $(\mathbf{i}, \mathbf{j}) \in \mathbb{I}_{k+m}$.

Proof. This follows from Theorem 5.1(d) by induction on m. For $m = 0$, (5.23) holds trivially. Assuming that Theorem 5.1(d) holds for m and summing over $\bar{\mathbf{j}} = (\mathbf{j}, j_{m+1})$, we arrive at the following estimate

$$\left| \sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}}) - \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) N_{k+m+1}^{1-\epsilon_{k+m+1}} \right|$$

$$\leq \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \left[N_{k+m+1}^{1-\epsilon_{k+m+1}} \ln(4BP_{k+m}) \right]^{\frac{1}{2}},$$

so that

$$\left|\frac{\sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i},\bar{\mathbf{j}})}{N_{k+m+1}^{1-\epsilon_{k+m+1}}\sum_{\mathbf{j}} X_{k+m}(\mathbf{i},\mathbf{j})} - 1\right| \leq \sqrt{\frac{\ln(4BP_{k+m})}{N_{k+m+1}^{1-\epsilon_{k+m+1}}}} \leq \frac{1}{2},$$

where at the last step we used (5.4). Thus

$$\frac{1}{2} \leq \frac{\sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i},\bar{\mathbf{j}})}{N_{k+m+1}^{1-\epsilon_{k+m+1}}\sum_{\mathbf{j}} X_{k+m}(\mathbf{i},\mathbf{j})} \leq 2 \quad \text{for all } m \geq 1,$$

which yields the desired result by induction. □

Proof of Corollary 5.2(b). Since

$$\begin{aligned}\sup_{k':k'\geq k} \sum_{\mathbf{i}:X_k(\mathbf{i})=1} \left|\int_{I_k(\mathbf{i})} \big(\phi_{k'} - \phi_k\big)(x)\,dx\right| &\leq \sum_{\mathbf{i}:X_k(\mathbf{i})=1}\sum_{m=0}^{\infty}\left|\int_{I_k(\mathbf{i})} \sigma_{k+m}(x)\,dx\right| \\ &\leq P_k \sup_{\mathbf{i}:X_k(\mathbf{i})=1}\left[\sum_{m=0}^{\infty}\left|\int_{I_k(\mathbf{i})} \sigma_{k+m}(x)\,dx\right|\right],\end{aligned}$$

it suffices to prove that the quantity in the last line is bounded above by the right hand side of (5.5). To this end, we fix an $m \geq 0$ and $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$ and write

$$\begin{aligned}P_k \int_{I_k(\mathbf{i})} \sigma_{k+m}(x)\,dx &= \frac{P_k}{P_{k+m+1}}\sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i},\bar{\mathbf{j}}) - \frac{P_k}{P_{k+m}}\sum_{\mathbf{j}} X_{k+m}(\mathbf{i},\mathbf{j}) \\ &= \Xi_1 + \Xi_2, \text{ where}\end{aligned}$$

$$\Xi_1 := P_k\left[\frac{1}{P_{k+m+1}} - \frac{1}{Q_{k+m+1}}\right]\sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i},\bar{\mathbf{j}}), \text{ and}$$

$$\Xi_2 := \frac{P_k}{Q_{k+m+1}}\sum_{\mathbf{j}} X_{k+m}(\mathbf{i},\mathbf{j})\sum_{j_{m+1}}\big(Y_{k+m+1}(\mathbf{i},\mathbf{j},j_{m+1}) - p_{k+m+1}\big).$$

By Theorem 5.1(a) and (5.23), we have

$$\begin{aligned}|\Xi_1| &\leq P_k\frac{|Q_{k+m+1} - P_{k+m+1}|}{P_{k+m+1}Q_{k+m+1}}\sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i},\bar{\mathbf{j}}) \\ &\leq \frac{BP_k}{P_{k+m+1}\sqrt{Q_{k+m+1}}}2^{m+1}\left[\prod_{j=1}^{m+1} N_{k+j}^{1-\epsilon_{k+j}}\right] \\ &\leq B2^{\frac{5}{2}(k+m+1)}\left[\prod_{j=1}^{k+m+1} N_j^{1-\epsilon_j}\right]^{-\frac{1}{2}}.\end{aligned} \tag{5.24}$$

On the other hand, using both Theorem 5.1(d) and (5.23),

$$
\begin{aligned}
|\Xi_2| &\le \frac{P_k}{Q_{k+m+1}} \sum_{\mathbf{j}} X_{k+m}(\mathbf{i},\mathbf{j}) \Big[8N_{k+m+1}^{1-\epsilon_{k+m+1}} \ln(BP_{k+m})\Big]^{\frac{1}{2}} \\
&\le \frac{P_k}{Q_{k+m+1}} 2^m \left[\prod_{j=1}^{m} N_{k+j}^{1-\epsilon_{k+j}}\right] \times \Big[8N_{k+m+1}^{1-\epsilon_{k+m+1}} \ln(BP_{k+m})\Big]^{\frac{1}{2}} \qquad (5.25)\\
&\le 2^{2(k+m)}\sqrt{8}\,\frac{\big[\ln(BP_{k+m})\big]^{\frac{1}{2}}}{N_{k+m+1}^{(1-\epsilon_{k+m+1})/2}}.
\end{aligned}
$$

Combining (5.24) and (5.25) and using (5.4), we obtain

$$
|\Xi_1| + |\Xi_2| \le 2B2^{\frac{5}{2}(k+m+1)} \frac{\big[\ln(BP_{k+m})\big]^{\frac{1}{2}}}{N_{k+m+1}^{(1-\epsilon_{k+m+1})/2}} \le 2B\cdot 2^{-\frac{(k+m)\gamma}{2}}.
$$

The conclusion (5.5) follows upon summation in m. □

6. The estimates for $\mathcal{M}$ and $\mathfrak{M}$

In this section we prove those parts of Theorems 1.1 and 1.3 that concern the restricted maximal operators with $1 < r < 2$. We will do this by fixing the parameters N_k, ϵ_k of the random construction in Section 5 and showing that the conclusions of the theorems hold for the sets S_k with those choices of parameters. Specifically, the conclusions of Theorem 1.1 will hold for S_k with

$$
N_k = N^{k+1},\ \epsilon_k = \frac{1}{k+1}, \qquad (6.1)
$$

and the conclusions of Theorem 1.3 will hold for S_k with

$$
N_k = N^k,\ \epsilon_k = \epsilon, \qquad (6.2)
$$

where N is a large integer.

Lemma 6.1. *Let N_k, ϵ_k be as above with N sufficiently large. Then:*

(a) the set $S = \bigcap_{k=1}^\infty S_k$ has Hausdorff dimension 1 if (6.1) holds and $1-\epsilon$ if (6.2) holds,

(b) assuming (6.1), (3.10) holds for all $q_0 \ge 2$,

(c) assuming (6.2), (3.10) holds for all $2 \le q_0 < q_\epsilon$, where $q_\epsilon = \frac{\epsilon+1}{2\epsilon}$ as in Theorem 1.3,

(d) assuming either (6.1) or (6.2), (5.4) holds with $\gamma = 1$.

Lemma 6.1 will be proved in Subsections 6.1 and 6.2 for (6.1) and (6.2), respectively.

Assuming the lemma, the proof of the restricted maximal estimates is completed as follows. By parts (b) and (c) of the lemma, (3.10) holds with q_0 as above. It follows by Corollary 3.5 that

$$\|\mathcal{M}_k f\|_{(q_0-1)p} \leq C2^{-k\eta(p)}\|f\|_p,\ p > \frac{q_0}{q_0-1}, \tag{6.3}$$

for the same q_0.

Consider first the case when (6.1) holds. We claim that then

$$\|\mathcal{M}_k f\|_q \leq C2^{-k\eta(p)}\|f\|_p \tag{6.4}$$

for all $p, q \in (1,\infty)$. Indeed, fix p and q, and choose q_0 large enough so that $\frac{q_0}{q_0-1} < p$ and $(q_0-1)p > q$. Since $\mathcal{M}_k f$ is supported on $[-4,0]$, we have

$$\|\mathcal{M}_k f\|_q \leq 5^{\frac{1}{q}-\frac{1}{(p-1)q_0}}\|\mathcal{M}_k f\|_{(q_0-1)p}$$

by Hölder's inequality. Combining this with (6.3), we get (6.4).

Summing up (6.4) in k, we see that $\mathcal{M}$ is bounded from $L^p[0,1]$ to $L^q[-4,0]$ for any $p, q \in (1,\infty)$. By Lemma 3.1, it follows that $\mathcal{M}$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ whenever $1 < p \leq q < \infty$.

Assume now that (6.2) holds instead. We claim that in this case (6.4) holds whenever

$$\frac{1+\epsilon}{1-\epsilon} < p < \infty \quad \text{and} \quad 1 < q < \frac{1-\epsilon}{2\epsilon}\,p. \tag{6.5}$$

Indeed, fix such p and q, then $p' < \frac{1+\epsilon}{2\epsilon} = q_\epsilon$. Choose q_0 so that $p' < q_0 < q_\epsilon$, then (6.3) yields (6.4) with $q = (q_0-1)p$. As in the first case, (6.4) also holds for $q < (q_0-1)p$ by Hölder's inequality. Taking $q_0 \to q_\epsilon$, we get (6.4) for all $p' < q_\epsilon$ and $q < (q_\epsilon - 1)p$, which is equivalent to (6.5). We now sum up (6.4) in k to obtain the boundedness of $\mathcal{M}$ from $L^p[0,1]$ to $L^q[-4,0]$ for p, q as in (6.5). By (3.1), $\mathcal{M}$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ whenever $p \leq q$ and (6.5) holds. Note that the range of p, q is nonempty whenever $\epsilon < 1/3$.

The same conclusions follow automatically for $\mathfrak{M}$, provided that the weak limit μ of ϕ_k exists. But thanks to Lemma 6.1(d), (5.4) holds, hence the existence of μ follows from Theorem 5.1(d) for both (6.1) and (6.2).

6.1. **The 1-dimensional case.** Let N_k, ϵ_k be as in (6.1). Then $M_k = N^{\frac{k(k+3)}{2}}$ and, by Theorem 5.1(a),

$$2^{-k}N^{\frac{k(k+1)}{2}} \leq P_k \leq 2^k N^{\frac{k(k+1)}{2}}.$$

By Lemma 2.1(b),

$$\begin{aligned}\dim_{\mathbb{H}}(S) &\geq \liminf_{k\to\infty} \log(P_k/N_k)/\log(M_{k-1}) \\ &\geq \liminf_{k\to\infty} \frac{\log(2^{-k}N^{\frac{k(k+1)}{2}-(k+1)})}{\log(N^{\frac{(k-1)(k+2)}{2}})} = 1.\end{aligned}$$

Hence S has dimension 1.

To prove Lemma 6.1(b), it suffices to show that for any $q_0 \geq 2$ the right side of (5.3) is bounded by $C(q_0)2^{-\eta k}$ with $\eta = \eta(q_0) > 0$. Suppose first that $q_0 = n$ is an even integer. Plugging our values of N_j and ϵ_j into (5.3), we see after some straightforward but cumbersome algebra that

$$\sup_{\Omega\subseteq[0,1]} \frac{\|\Phi_k^* \mathbf{1}_\Omega\|_n}{|\Omega|^{\frac{n-1}{n}}} \leq C(n!\,B)^{1/n} 2^{k(1+\frac{3}{2n})} N^{-\frac{k^2}{4n}+(1-\frac{5}{4n})k+1} \times \Big[\ln(4^n n!\,B) + (k+1)(k+4)Ln\ln N\Big]^{1/2n},$$

which is bounded by $C(n)2^{-\eta(n)k}$ with $\eta(n) = \frac{1}{4n} > 0$ for all even integers n. The estimate in (b) for all $q_0 \geq 2$ (not necessarily an even integer) follows by interpolation.

Finally, to prove (d) we estimate

$$2^{6k} \frac{\ln(M_k)}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \frac{2^{6k-1}\,k(k+3)\ln N}{N^{k+1}} < \frac{1}{32}$$

for all k, provided that N is large enough.

6.2. **The lower-dimensional case.** Let N_k, ϵ_k be as in (6.2). Then $M_k = N^{\frac{k(k+1)}{2}}$ and by Theorem 5.1(a),

$$2^{-k} N^{\frac{k(k+1)}{2}(1-\epsilon)} \leq P_k \leq 2^k N^{\frac{k(k+1)}{2}(1-\epsilon)}.$$

By Lemma 2.1(a),

$$\dim_{\mathbb{H}}(S) \leq \liminf_{k\to\infty} \log(P_k)/\log(M_k) \leq \liminf_{k\to\infty} \frac{\log(2^k N^{\frac{k(k+1)}{2}(1-\epsilon)})}{\log(N^{\frac{k(k+1)}{2}})} = 1-\epsilon,$$

whereas by Lemma 2.1(b),

$$\dim_{\mathbb{H}}(S) \geq \liminf_{k\to\infty} \log(P_k/N_k)/\log(M_{k-1}) \geq \liminf_{k\to\infty} \frac{\log(2^{-k} N^{\frac{k(k+1)}{2}(1-\epsilon)-k})}{\log(N^{\frac{k(k-1)}{2}})} = 1-\epsilon.$$

Hence S has dimension $1-\epsilon$.

Next, we verify Lemma 6.1(c). Plugging (6.2) into (5.3), we see after some more algebra that

$$\sup_{\Omega\subseteq[0,1]} \frac{\|\Phi_k^* \mathbf{1}_\Omega\|_n}{|\Omega|^{\frac{n-1}{n}}} \leq C(n!\,B)^{1/n} 2^{k(1+\frac{3}{2n})} N^{\frac{k(k+1)}{2n}(-\frac{1}{2}+\epsilon(n-\frac{1}{2}))+(k+1)\epsilon} \times \Big[\ln(4^n n!\,B) + (k+1)(k+2)Ln\ln N\Big]^{1/2n}.$$

This is majorized by $C(n)2^{-\eta(n)k}$ with $\eta(n) = \frac{1+\epsilon}{2n} - \epsilon$. Note that $\eta(n) > 0$ if and only if $\epsilon(n - \frac{1}{2}) < \frac{1}{2}$, i.e.

$$\epsilon < \frac{1}{2n-1}, \text{ or } n < \frac{1}{2} + \frac{1}{2\epsilon} = q_\epsilon. \tag{6.6}$$

Let $n_1 = n_1(\epsilon)$ be the largest even integer such that (6.6) holds, and let $n_2 = n_1 + 2$. Interpolating between the estimates for n_1 and n_2, we get that

$$\sup_{\Omega \subseteq [0,1]} \frac{\|\Phi_k^* \mathbf{1}_\Omega\|_{q_0}}{|\Omega|^{\frac{q_0-1}{q_0}}} \leq C(q_0) 2^{-\eta(q_0)k}$$

with $\eta(q_0) > 0$ for all $q_0 < q_\epsilon$.

For part (d), we check as before that

$$2^{6k} \frac{\ln(M_k)}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \frac{2^{6k+1}\, k(k+1) \ln N}{N^{(k+1)(1-\epsilon)}} \leq \frac{1}{32}$$

for all k, if N was chosen large enough. This proves (d) and establishes the existence of μ.

7. Extension to the unrestricted operator

It remains to prove the statements for the unrestricted maximal operators $\tilde{\mathcal{M}}^a$ and $\tilde{\mathfrak{M}}^a$ claimed in Theorem 1.1 (e) and Theorem 1.3(c). Obtaining bounds for global maximal operators using known bounds for single-scale ones is a common theme in the harmonic analysis literature, often involving interpolation and scaling. In this section we present these arguments with the necessary modifications for our problem. The proof naturally splits into two cases $q \geq 2$ and $q < 2$, which are handled in Propositions 7.1 and 7.2 respectively. The former follows an approach closely related to [8], [35]. The proof for $p = q < 2$ is due to Andreas Seeger, who also indicated to us prior work in this direction [28], [9]. Proposition 7.2 combines his argument with interpolation techniques used in a similar setting in [19].

We remark that the scaling arguments below are quite general and apply to any sequence S_k as described in Section 2 subject to the bounds on $\mathcal{M}_k$ and (in Lemma 7.4) the subexponential growth of N_k. In other words, we will not be invoking the probabilistic arguments of Section 5.

Recall the definitions (1.1), (1.2), (1.3), (1.4) and (3.3) of $\tilde{\mathcal{M}}$, $\tilde{\mathfrak{M}}$, $\tilde{\mathcal{M}}^a$, $\tilde{\mathfrak{M}}^a$ and $\mathcal{M}_k$ respectively. Denote by $A_r[k]$ the averaging operator associated to ϕ_k:

$$A_r[k]f(x) = \int f(x+ry)\phi_k(y)\, dy, \qquad \text{where} \qquad \phi_k = \frac{1}{|S_k|}\mathbf{1}_{S_k}. \tag{7.1}$$

The main results in this section are the following.

Proposition 7.1. *Fix two exponents p, q satisfying $1 < p \le q < \infty$, $q \ge 2$. Assume that for some $C > 0$ and $\eta_0 > 0$ we have the estimate*

$$\|\mathcal{M}_k f\|_q \le C 2^{-\eta_0 k} \|f\|_p \tag{7.2}$$

for all f supported on $[0,1]$. Assume furthermore that N_k have been chosen as in (6.1) or (6.2). Then $\tilde{\mathcal{M}}^a$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$, with $a = \frac{1}{p} - \frac{1}{q}$.

Proposition 7.2. *Suppose that there exists $\epsilon \in [0, \frac{1}{3})$ such that (7.2) holds for all functions f supported in $[0,1]$ and all exponents (p,q) satisfying*

$$1 < p \le q \le 2, \quad \frac{1+\epsilon}{1-\epsilon} < p < \infty, \quad 1 < q < \frac{1-\epsilon}{2\epsilon} p. \tag{7.3}$$

Then $\tilde{\mathcal{M}}^a$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ for all such (p,q), with $a = \frac{1}{p} - \frac{1}{q}$.

Remark. Despite the formal similarity, it is worth noting the distinction between the statements of the two propositions. In Proposition 7.1, the assumption (7.2) is for a *fixed* (p,q), and the conclusion is the estimate for the global operator with the same (p,q). In contrast, for Proposition 7.2, the hypothesis (7.2) is for *all* (p,q) in the domain (7.3).

Conclusion of the proofs of Theorems 1.1 (e) and 1.3 (c). Assuming the two propositions, the unrestricted maximal bounds are proved as follows. It suffices to prove the bounds on $\tilde{\mathcal{M}}^a$. Suppose first that we are in the one-dimensional case (6.1). Then (6.4) asserts that the hypotheses of both Propositions 7.1 and 7.2 hold (the latter with $\epsilon = 0$), hence so do the conclusions. In the lower-dimensional case (6.2), the same argument shows that $\tilde{\mathcal{M}}^a$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ whenever p, q obey (6.5) with $p \le q$. □

7.1. **Scaling arguments.** The proofs of both Propositions 7.1 and 7.2 use the Haar decomposition of a function f and the relation between the averaging operators $A_r[k]$ for various scales of the dilation parameter r. We record the necessary facts in the following sequence of lemmas. Following [8], we denote by $\mathcal{D}_s$ the σ-algebra generated by dyadic intervals of length 2^{-s}, and by $\mathbb{E}_s$ the corresponding conditional expectation operators, i.e., $\mathbb{E}_s(f) = \mathbb{E}(f|\mathcal{D}_s)$. We also set

$$\Delta_s f = \mathbb{E}_{s+1}(f) - \mathbb{E}_s(f). \tag{7.4}$$

Lemma 7.3. *Let $1 < p, q < \infty$ and $\eta_0 > 0$. Suppose that (7.2) holds for all f supported on $[0,1]$. Then there exists $\eta > 0$ such that*

$$\|\mathcal{M} f\|_q \le C 2^{-\eta\sqrt{s}} \|f\|_p \tag{7.5}$$

for all functions $f \in L^p(\mathbb{R})$ satisfying $\mathbb{E}_s(f) = 0$, $s \ge 0$.

Proof. For any f supported in $[0,1]$,

$$|A_r[k](f)| \le \mathcal{N} f(x) + \sum_{m=0}^{k-1} \mathcal{M}_m |f|, \quad \text{hence} \quad \mathcal{M} f \le \mathcal{N} f + \sum_{k=0}^{\infty} \mathcal{M}_k |f|,$$

where $\mathcal{N}$ was defined at the beginning of Subsection 3.2. Therefore

$$\|\mathcal{M}f\|_q \le \sum_{k=0}^{\infty} \|\mathcal{M}_k f\|_q. \tag{7.6}$$

The right side is clearly summable by (7.2). To obtain decay as required in (7.5), we will use the assumption that $\mathbb{E}_s f = 0$ to improve the estimate on the terms with $k \le k_0$, where k_0 will be determined shortly. We have

$$\begin{aligned}\mathcal{M}_k f(x) &= \sup_{1<r<2} \Big| \int f(x+ry)(\phi_{k+1}(y) - \phi_k(y))dy \Big| \\ &\le \Big| \int f(x+ry)\phi_k(y)dy \Big| + \Big| \int f(x+ry)\phi_{k+1}(y)dy \Big|.\end{aligned}$$

Suppose that $2^{-s} < \delta_{k+1}$, and consider the term with ϕ_k first. Each of the δ_k-intervals $\{I_k(\mathbf{i}) : \kappa_k(\mathbf{i}) = 1\}$ in the support of ϕ_k can be written as a union of some number of dyadic 2^{-s}-intervals together with two intervals $J_1(\mathbf{i}, s)$ and $J_2(\mathbf{i}, s)$ of length at most 2^{-s}, one at each end of $I_k(\mathbf{i})$. Since f integrates to 0 on each dyadic interval, the only non-zero contribution comes from the intervals $J_j(\mathbf{i}, s)$. By Hölder's inequality, we see that

$$\begin{aligned}\Big| \int f(x+ry)\phi_k(y)\,dy \Big| &= \frac{1}{P_k\delta_k} \Big| \sum_{j=1}^{2} \sum_{\kappa_k(\mathbf{i})=1} \int_{J_j(\mathbf{i},s)} f(ry)dy \Big| \\ &\le \frac{1}{P_k\delta_k} \|f\|_p (2P_k \cdot 2^{-s})^{\frac{1}{p'}} \\ &= \frac{1}{P_k^{1/p}\delta_k} 2^{1-\frac{s}{p'}} \|f\|_p \le M_k 2^{1-\frac{s}{p'}} \|f\|_p \,.\end{aligned}$$

The term with ϕ_{k+1} is estimated similarly. Taking the L^q norm of the left side and using the fixed compact support of $\mathcal{M}_k f$, we see that

$$\|\mathcal{M}_k f\|_q \le M_k 2^{1-\frac{s}{p'}} \|f\|_p \le C 2^{-\frac{s}{p'}} N^{k(k+3)/2} \|f\|_p,$$

where we used (6.1) and (6.2) at the last step. Let $k_0 \approx c\sqrt{s}$ with a small enough constant, then for all $k \le k_0$ we have $N^{k(k+3)/2} 2^{-s/(2p')} \le C$, so that

$$\|\mathcal{M}_k f\|_q \le C 2^{-s/(2p')} \|f\|_p \quad \text{for } k \le k_0.$$

We now use this along with (7.2) to estimate the right side of (7.6):

$$\begin{aligned}\sum_{k=1}^{\infty} \|\mathcal{M}_k f\|_q &= \sum_{k=1}^{k_0} \|\mathcal{M}_k f\|_q + \sum_{k>k_0} \|\mathcal{M}_k f\|_q \\ &\leq C k_0 2^{-s/(2p')} \|f\|_p + C \sum_{k>k_0} 2^{-\eta_0 k} \|f\|_p \\ &\leq C 2^{-s/(4p')} \|f\|_p + C 2^{-c\eta_0\sqrt{s}} \|f\|_p \leq C 2^{-\eta\sqrt{s}} \|f\|_p,\end{aligned}$$

as claimed in (7.5). This proves the result for functions f supported in [0,1]. The extension to a general f is achieved by a "disjointness of support" argument identical to the one given in Lemma 3.1 and is left to the reader. □

We will also need the following rescaled version of (7.5).

Lemma 7.4. *Suppose that (7.5) holds for all functions $f \in L^p(\mathbb{R})$ satisfying $\mathbb{E}_s(f) = 0$ for some $s \geq 0$. Then for any $m \in \mathbb{Z}$ and all $f \in L^p(\mathbb{R})$,*

$$\Big\| \sup_{\substack{k\geq 1 \\ 1\leq r2^m \leq 2}} |A_r[k](\Delta_{s+m} f)| \Big\|_q \leq C \cdot 2^{ma-\eta\sqrt{s}} \|\Delta_{s+m} f\|_p. \tag{7.7}$$

Here $\Delta_s f$ is as in (7.4).

Proof. Let $u = r2^m$, so that $1 \leq u \leq 2$. We have

$$\begin{aligned}A_r[k]f(x) &= \int f(x+ry)\phi_k(y)dy \\ &= \int f(x+2^{-m}uy)\phi_k(y)dy \\ &= \int f(2^{-m}(2^m x + uy))\phi_k(y)dy \\ &= A_u[k](f^{(m)})(2^m x),\end{aligned} \tag{7.8}$$

where $f^{(m)}(\cdot) = f(2^{-m}\cdot)$. Note also that $(\Delta_{s+m} f)^{(m)} = \Delta_{s+m}(2^{-m}\cdot)$ is constant on dyadic 2^{-s}-intervals, i.e. $\mathbb{E}_s((\Delta_{s+m} f)^{(m)}) = 0$. By (7.5), we have

$$\begin{aligned}\Big\| \sup_{\substack{k\geq 1 \\ 1\leq r2^m \leq 2}} |A_r[k](\Delta_{s+m} f)| \Big\|_q &= \Big\| \sup_{\substack{k\geq 1 \\ 1\leq u \leq 2}} \big| \big(A_u[k](\Delta_{s+m} f)^{(m)}\big)(2^m \cdot) \big| \Big\|_q \\ &= 2^{-m/q} \|\mathcal{M}(\Delta_{s+m} f)^{(m)}\|_q \\ &\leq C 2^{-m/q}\, 2^{-\eta\sqrt{s}} \big\|(\Delta_{s+m} f)^{(m)}\big\|_p \\ &= C 2^{-m/q}\, 2^{-\eta\sqrt{s}}\, 2^{m/p} \|\Delta_{s+m} f\|_p \\ &= C 2^{ma}\, 2^{-\eta\sqrt{s}} \|\Delta_{s+m} f\|_p\ .\end{aligned}$$

□

Finally, we need a technical lemma.

Lemma 7.5. *Given any $0 \le a < 1$, there is a constant $C = C(a)$ such that for any $m \in \mathbb{Z}$ and all $f \in L^p(\mathbb{R})$,*

$$\sup_{\substack{k\ge 1\\ 1\le r2^m\le 2}} r^a\big|A_r[k]\mathbb{E}_m f(x)\big| \le Cf^*(x), \text{ where}$$

$$f^*(x) := \sup_{r>0} r^{a-1}\int_{|y|\le r}|f(x-y)|dy.$$

The mapping $f \mapsto f^*$ is bounded from $L^p(\mathbb{R}) \to L^q(\mathbb{R})$ for all $1 < p \le q \le \infty$ for which $a = \frac{1}{p} - \frac{1}{q}$.

Proof. Since $S_k \subseteq [1,2]$ and $r \le 2^{-m+1}$, the set $x + rS_k$ is contained in an interval J centered at x of length 2^{-m+3}. Observe that J can be covered by at most 10 dyadic 2^{-m}-intervals J_i. On each J_i, we have $\mathbb{E}_m(f) \equiv \lambda_i$, where λ_i is the average of f on J_i. Since $A_r[k]\mathbb{E}_m f(x)$ is a convex linear combination of the λ_i-s, it suffices to prove that $r^a\lambda_i \le f^*(x)$. But this follows from

$$r^a\lambda_i = \frac{r^a}{|J_i|}\int_{J_i}|f(y)|dy \le \frac{10r^a}{|J|}\int_{J'}|f(y)|dy \le \frac{C}{|J'|^{1-a}}\int_{J'}|f(y)|dy \le f^*(x),$$

where J' is an interval of length $2|J|$ centered at x so that $J \subset \bigcup J_i \subset J'$.

If $p = q$, then $a = 0$ and f^* is simply the Hardy-Littlewood maximal function of f, which is bounded on all L^p for $p > 1$. If on the other hand $1 < p < q \le \infty$, then $0 < a < 1$ and

$$f^*(x) = \sup_{r>0} r^{a-1}\int_{|x-z|\le r}|f(z)|dz \le \int_{-\infty}^{\infty}\frac{|f(z)|}{|x-z|^{1-a}}dz.$$

Since $f \in L^p(\mathbb{R})$ and $|z|^{a-1}$ is in weak $L^{\frac{1}{1-a}}(\mathbb{R})$, it follows by Young's inequality that the mapping $f \to f^*$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ with $1+\frac{1}{q} = \frac{1}{p}+(1-a)$, as claimed. □

7.2. **Proof of Proposition 7.1.** Given $m \in \mathbb{Z}$ such that $2^{-m} \le r \le 2^{-m+1}$, we write $f = \mathbb{E}_m(f) + \sum_{s\ge m}\Delta_s(f)$, where $\Delta_s(f)$ is defined as in (7.4). Therefore

$$A_r[k](f) = A_r[k](\mathbb{E}_m f) + \sum_{s\ge m} A_r[k](\Delta_s f), \tag{7.9}$$

so that

$$\tilde{\mathcal{M}}^a f \le \sup_{m\in\mathbb{Z}}\sup_{\substack{k\ge 1\\ 1\le r2^m\le 2}} r^a\left[|A_r[k]\mathbb{E}_m(f)(x)| + \Big|\sum_{s\ge m}A_r[k](\Delta_s f)\Big|\right]. \tag{7.10}$$

The first term is bounded from $L^p \to L^q$ by Lemma 7.5. Turning our attention to the second term of (7.10), it suffices to prove that

$$\left\|\sup_{m\in\mathbb{Z}}\sup_{\substack{k\ge 1\\ 1\le r2^m\le 2}} 2^{-ma}\Big|\sum_{s\ge m}A_r[k](\Delta_s f)\Big|\right\|_q \le C\|f\|_p\ . \tag{7.11}$$

We write

$$\sup_{m\in\mathbb{Z}} \sup_{\substack{k\geq 1\\ 1\leq r2^m\leq 2}} 2^{-ma}\Big|\sum_{s\geq m} A_r[k](\Delta_s f)\Big| \leq \Big[\sum_{m\in\mathbb{Z}} 2^{-maq} \sup_{\substack{k\geq 1\\ 1\leq r2^m\leq 2}} \Big|\sum_{s\geq m} A_r[k](\Delta_s f)\Big|^q\Big]^{\frac{1}{q}}$$
$$\leq \Big[\sum_{m\in\mathbb{Z}} 2^{-maq}\Big(\sum_{s\geq m} \sup_{\substack{k\geq 1\\ 1\leq r2^m\leq 2}} |A_r[k](\Delta_s f)|\Big)^q\Big]^{\frac{1}{q}}$$

Taking the L^q-norms of both sides, then using Lemma 7.4 (whose hypothesis in turn is true by Lemma 7.3), we see that the left side of (7.11) is bounded by

$$\Big(\sum_{m\in\mathbb{Z}} 2^{-maq}\Big\|\sum_{s\geq m} \sup_{\substack{k\geq 1\\ 1\leq r2^m\leq 2}} |A_r[k](\Delta_s f)|\Big\|_q^q\Big)^{\frac{1}{q}}$$
$$\leq \Big(\sum_{m\in\mathbb{Z}} 2^{-maq}\Big[\sum_{s\geq m}\Big\| \sup_{\substack{k\geq 1\\ 1\leq r2^m\leq 2}} |A_r[k](\Delta_s f)|\Big\|_q\Big]^q\Big)^{\frac{1}{q}}$$
$$\leq C\Big(\sum_{m\in\mathbb{Z}}\Big[\sum_{s\geq m} 2^{-\eta\sqrt{s-m}}\,\|\Delta_s f\|_p\Big]^q\Big)^{\frac{1}{q}}.$$

The last line is the ℓ^q-norm of the convolution of the discrete functions $\mathbf{1}_{m\geq 0}2^{-\eta\sqrt{m}}$ and $\|\Delta_m f\|_p$. Applying Young's inequality with $s = \max(p,2)$ and $\frac{1}{s}+\frac{1}{r} = 1+\frac{1}{q}$, we bound it by

$$\Big(\sum_{m\geq 0} 2^{-\eta\sqrt{m}r}\Big)^{\frac{1}{r}}\Big(\sum_{m\in\mathbb{Z}}\|\Delta_m f\|_p^s\Big)^{\frac{1}{s}} \leq C\Big(\sum_m \|\Delta_m f\|_p^s\Big)^{\frac{1}{s}}. \tag{7.12}$$

It remains to show that

$$\Big(\sum_m \|\Delta_m f\|_p^s\Big)^{\frac{1}{s}} \leq C\|f\|_p. \tag{7.13}$$

Suppose first that $p\geq 2$, so that $s = p$. Then the claim is trivial for $p=\infty$, and for $p=2$ it follows from the orthogonality of $\Delta_m f$. By interpolation, this implies (7.13) for all $p\in[2,\infty)$. Assume next that $1<p<2$, so that $s=2$. Then

$$\Big(\sum_{m\in\mathbb{Z}}\|\Delta_m f\|_p^2\Big)^{\frac{1}{2}} \leq \Big\|\Big(\sum_{m\in\mathbb{Z}}|\Delta_m f|^2\Big)^{\frac{1}{2}}\Big\|_p \leq C_p||f||_p,$$

where the first step follows from the generalized Minkowski inequality and the second from Littlewood-Paley theory. This proves the claim (7.13).

□

7.3. **Proof of Proposition 7.2.** As indicated in the remark following Proposition 7.2, the conclusion is immediate from Proposition 7.1 if $q = 2$. Fix $\epsilon \in [0, \frac{1}{3})$ and exponents (p, q), $q < 2$ satisfying (7.3). We denote by $C(p, q; R)$ the norm of the linear operator

$$f \mapsto \big\{2^{-ma} A_{r2^{-m}}[k]f : -R \le m \le R,\ 1 \le r \le 2,\ k \ge 1\big\}, \tag{7.14}$$

mapping $L^p(\mathbb{R})$ to $L^q(\ell^\infty_m L^\infty_r \ell^\infty_k)$, where $a = \frac{1}{p} - \frac{1}{q}$. In other words, $C(p, q; R)$ is the best constant such that the following inequality holds for all f:

$$\big\| \sup_{-R\le m\le R} \sup_{1\le r\le 2} \sup_{k\ge 1} 2^{-ma} \big|A_{r2^{-m}}[k]f\big|\big\|_q \le C(p, q; R)\|f\|_p. \tag{7.15}$$

We first ensure that $C(p, q; R)$ is well-defined. The hypothesis (7.2) implies (3.1) after summing in k, hence the inequality in (3.1) continues to hold for all $f \in L^p(\mathbb{R})$ by Lemma 3.1. By the scaling argument in (7.8), this implies that for every fixed $m \in \mathbb{Z}$,

$$\big\|2^{-ma} \sup_{1\le r\le 2} \sup_{k\ge 1} \big|A_{r2^{-m}}[k]f\big|\big\|_q \le \|\mathcal{M}\|_{p\to q}\|f\|_p, \quad f \in L^p(\mathbb{R}). \tag{7.16}$$

Thus we already have the trivial bound $C(p, q; R) \le R\|\mathcal{M}\|_{p\to q}$. Our goal is to show that for each p, q in the indicated range, $C(p, q; R)$ is bounded uniformly in R:

$$C(p, q; R) = O_{p,q}(1). \tag{7.17}$$

This would imply the conclusion of the proposition, since the left hand side of (7.15) converges as $R \to \infty$ to a limit that is bounded above and below by positive constant multiples of $\|\tilde{\mathcal{M}}^a f\|_q$. The convergence is justified by the monotone convergence theorem, which applies because the the operators $A_r[k]$ are non-negative and the functions f can be chosen to be non-negative.

In order to prove (7.17) we fix two other auxiliary exponents (p_1, q_1) and (p_2, q_2) obeying (7.3), such that $p_1 < p < p_2$, $q_2 = 2$, and the points $\{(\frac{1}{p}, \frac{1}{q}), (\frac{1}{p_1}, \frac{1}{q_1}), (\frac{1}{p_2}, \frac{1}{2})\}$ are collinear. The following lemma provides an essential interpolation ingredient of the proof.

Lemma 7.6. *Given any sequence of functions* $\{g_m : -R \le m \le R\}$, *define*

$$\mathcal{T}_a(\{g_m\}) = \big\{2^{-ma} A_{r2^{-m}}[k]g_m : -R \le m \le R,\ 1 \le r \le 2,\ k \ge 1\big\}.$$

(a) For any (p, q) *obeying (7.3), the operator* $\mathcal{T}_{a(p,q)} : L^p_x \ell^\infty_m \to L^q_x \ell^\infty_m L^\infty_r \ell^\infty_k$ *has norm bounded by* $C(p, q; R)$, *with* $a(p, q) = \frac{1}{p} - \frac{1}{q}$.

(b) For any (p_1, q_1) *obeying (7.3), there is a constant* $K_1 = K_1(p_1, q_1)$ *independent of* R *such that the operator* $\mathcal{T}_{a(p_1,q_1)} : L^{p_1}_x \ell^{p_1}_m \to L^{q_1}_x \ell^{p_1}_m L^\infty_r \ell^\infty_k$ *is bounded with norm* $\le K_1$.

(c) If $p_1 < p$, the norm of the operator $\mathcal{T}_{a(p_3,q_3)} : L^{p_3}_x \ell^2_m \to L^{q_3}_x \ell^2_m L^\infty_r \ell^\infty_k$ is bounded by $K_1^{\frac{p_1}{2}} C(p,q;R)^{1-\frac{p_1}{2}}$; i.e.,

$$\Big\| \Big(\sum_{m=-R}^{R} \big[2^{-ma} \sup_{1\le r\le 2} \sup_{k\ge 1} \big| A_{r2^{-m}}[k] g_m \big| \big]^2 \Big)^{\frac12} \Big\|_{q_3} \le K_1^{\frac{p_1}{2}} C(p,q;R)^{1-\frac{p_1}{2}} \Big\| \Big(\sum_m |g_m|^2 \Big)^{\frac12} \Big\|_{p_3}. \tag{7.18}$$

Here $\frac{1}{p_3} = \frac{1}{p} + \frac{p_1}{2}(\frac{1}{p_1} - \frac{1}{p})$, and $(\frac{1}{p_i}, \frac{1}{q_i})$, $i = 1,2,3$ are collinear.

Proof. Part (a) is a consequence of the non-negativity of $A_r[k]$ combined with (7.15):

$$\Big\| \sup_{-R\le m\le R} \sup_{1\le r\le 2} \sup_{k\ge 1} 2^{-ma} \big| A_{r2^{-m}}[k] g_m \big| \Big\|_q \le \Big\| \sup_{-R\le m\le R} \sup_{1\le r\le 2} \sup_{k\ge 1} 2^{-ma} \big| A_{r2^{-m}}[k] \big| \sup_j g_j \big| \big| \Big\|_q \le C(p,q;R) \big\| \sup_m |g_m| \big\|_p.$$

For part (b), by the triangle inequality in L^{q_1/p_1} applied to functions $|G_m|^{p_1}$ we have

$$\Big\| \Big(\sum_m |G_m|^{p_1} \Big)^{\frac{1}{p_1}} \Big\|_{q_1} \le \Big(\sum_m \| G_m \|_{q_1}^{p_1} \Big)^{\frac{1}{p_1}} \quad \text{since } p_1 \le q_1.$$

Using this with $G_m = 2^{-ma} \sup_{1\le r\le 2} \sup_{k\ge 1} A_{r2^{-m}}[k] g_m$, we find

$$\begin{aligned} \Big\| \Big(\sum_m \Big| 2^{-ma} \sup_{1\le r\le 2} \sup_{k\ge 1} \big| A_{r2^{-m}}[k] g_m \big| \Big|^{p_1} \Big)^{\frac{1}{p_1}} \Big\|_{q_1} &\le \Big(\sum_m \Big\| 2^{-ma} \sup_{1\le r\le 2} \sup_{k\ge 1} A_{r2^{-m}}[k] g_m \Big\|_{q_1}^{p_1} \Big)^{\frac{1}{p_1}} \\ &\le \| \mathcal{M} \|_{p_1 \to q_1} \Big(\sum_m \| g_m \|_{p_1}^{p_1} \Big)^{\frac{1}{p_1}} \\ &= K_1 \Big\| \big(\sum_m |g_m|^{p_1} \big)^{\frac{1}{p_1}} \Big\|_{p_1}, \end{aligned}$$

where we have used (7.16) with (p_1, q_1) at the second step. This gives the conclusion with $K_1 = \| \mathcal{M} \|_{p_1 \to q_1}$. Part (c) now follows by complex interpolation of the family of operators $\mathcal{T}_a$ between the spaces in parts (a) and (b). The interpolation works because $p_1 < 2$, so that ℓ^2 is intermediate between ℓ^{p_1} and ℓ^∞. □

Conclusion of the proof of Proposition 7.2. In order to prove (7.17), we start again with the Haar decomposition of the function f, so that (7.9) holds. Thus

$$\begin{aligned}
&\sup_{-R\le m\le R}\ \sup_{1\le r\le 2}\ \sup_{k\ge 1} 2^{-ma}\big|A_{r2^{-m}}[k]f\big| \\
&\qquad \le \sup_{-R\le m\le R}\ \sup_{1\le r\le 2}\ \sup_{k\ge 1} 2^{-ma}\big|A_{r2^{-m}}[k]\mathbb{E}_m f\big| \\
&\qquad\quad + \sum_{s\ge 1} \sup_{-R\le m\le R}\ \sup_{1\le r\le 2}\ \sup_{k\ge 1} 2^{-ma}\big|A_{r2^{-m}}[k]\big(\Delta_{s+m}f\big)\big|.
\end{aligned} \tag{7.19}$$

As before, the first term on the right is bounded pointwise by f^*, and therefore bounded from $L^p \to L^q$ with norm independent of R by Lemma 7.5. We estimate the L^q norms of the summands in (7.19) as follows. On one hand, (7.18) with $g_m = \Delta_{s+m}f$ implies

$$\begin{aligned}
&\Big\|\Big(\sum_{m=-R}^{R}\big[2^{-ma}\sup_{1\le r\le 2}\ \sup_{k\ge 1}\big|A_{r2^{-m}}[k](\Delta_{s+m}f)\big|\big]^2\Big)^{\frac12}\Big\|_{q_3} \\
&\qquad \le K_1^{\frac{p_1}{2}}C(p_1,q_1;R)^{1-\frac{p_1}{2}}\Big\|\Big(\sum_m |\Delta_{s+m}f|^2\Big)^{\frac12}\Big\|_{p_3} \\
&\qquad \le K_1^{\frac{p_1}{2}}C(p_1,q_1;R)^{1-\frac{p_1}{2}}\|f\|_{p_3},
\end{aligned} \tag{7.20}$$

where the last step is a consequence of the Littlewood-Paley inequality. On the other hand, for all $s \ge 1$,

$$\begin{aligned}
&\Big\|\Big(\sum_{m=-R}^{R}\big[2^{-ma}\sup_{1\le r\le 2}\ \sup_{k\ge 1}\big|A_{r2^{-m}}[k](\Delta_{s+m}f)\big|\big]^2\Big)^{\frac12}\Big\|_{2} \\
&\qquad = \Big(\sum_m \Big\|2^{-ma}\sup_{1\le r\le 2}\ \sup_{k\ge 1}\big|A_{r2^{-m}}(\Delta_{s+m}f)\big|\Big\|_2^2\Big)^{\frac12} \\
&\qquad \le C2^{-\eta\sqrt{s}}\Big[\sum_m \|\Delta_{s+m}f\|_{p_2}^2\Big]^{\frac12} \\
&\qquad \le C2^{-\eta\sqrt{s}}\Big\|\Big(\sum_m |\Delta_{m+s}f|^2\Big)^{\frac12}\Big\|_{p_2} \\
&\qquad \le C2^{-\eta\sqrt{s}}\|f\|_{p_2}.
\end{aligned} \tag{7.21}$$

Here η is a positive constant (independent of m) whose existence is guaranteed by Lemma 7.4. The third step above uses the generalized Minkowski inequality (since $p_2 \le 2$) and the fourth follows from Littlewood-Paley theory.

Since $p_3 < p < p_2$ and $\{(\frac{1}{p_3}, \frac{1}{q_3}), (\frac{1}{p}, \frac{1}{q}), (\frac{1}{p_2}, \frac{1}{q_2})\}$ are collinear, we can interpolate between (7.20) and (7.21) to obtain $0 < \theta < 1$ such that

$$\begin{aligned} &\Big\| \sup_{-R \leq m \leq R} \sup_{1 \leq r \leq 2} \sup_{k \geq 1} 2^{-ma} \big| A_{r2^{-m}}\big(\Delta_{s+m} f\big)\big| \Big\|_q \\ &\qquad \leq \Big\| \Big(\sum_{m=-R}^{R} \big[2^{-ma} \sup_{1 \leq r \leq 2} \sup_{k \geq 1} \big| A_{r2^{-m}}[k] \Delta_{s+m} f \big| \big]^2 \Big)^{\frac{1}{2}} \Big\|_q \\ &\qquad \leq \big(K_1^{\frac{p_1}{2}} C(p,q;R)^{1-\frac{p_1}{2}} \big)^{\theta} \big(C 2^{-\eta\sqrt{s}} \big)^{1-\theta} \|f\|_p. \end{aligned}$$

The right hand side is summable in s. In summary, we have obtained the following estimate for the L^q norm of the left hand side of (7.19): there is a large constant K and $0 < \rho < 1$ such that

$$\Big\| \sup_{-R \leq m \leq R} \sup_{1 \leq r \leq 2} \sup_{k \geq 1} 2^{-ma} \big| A_{r2^{-m}}[k] f \big| \Big\|_q \leq K(1 + C(p,q;R)^{\rho}) \|f\|_p.$$

In view of the definition (7.15) of $C(p,q;R)$, we obtain $C(p,q;R) \leq C(1 + C(p,q;R)^{\rho})$. But this implies that $C(p,q;R)$ is bounded above by a constant depending only on K, p, q, but not on R, which is the desired conclusion (7.17). □

8. Differentiation results

8.1. **Proof of Theorem 1.2 and Theorem 1.3(d).** Assume that $\{S_k\}$ is a sequence of sets for which the maximal operator $\tilde{\mathcal{M}}$ is bounded on $L^p(\mathbb{R})$ for some $p \in (1, \infty)$. We claim that in this case $\{rS_k\}$ differentiates L^p in the sense that (1.7) holds.

Let $f \in L^p[0,1]$. We need to prove that

$$\lim_{r \to 0} \sup_{k \geq 1} |A_r[k] f(x) - f(x)| = 0 \tag{8.1}$$

for almost all x, where the averages $A_r[k]$ are defined as in (7.1). In other words, it suffices to show that for any $\lambda > 0$

$$\Big| \Big\{ x : \lim_{r \to 0} \sup_{k \geq 1} \big| A_r[k] f(x) - f(x) \big| > \lambda \Big\} \Big| = 0. \tag{8.2}$$

To this end, fix $t > 0$ and a continuous function f_t on $[0,1]$ such that $\|f - f_t\|_p < \epsilon$. Since (8.1) holds for all x for continuous functions,

$$\begin{aligned}
&\Big|\Big\{x : \limsup_{r\to 0}\sup_{k\geq 1}\big|A_r[k]f(x) - f(x)\big| > \lambda\Big\}\Big| \\
&\qquad = \Big|\Big\{x : \limsup_{r\to 0}\sup_{k\geq 1}\Big|A_r[k](f-f_t)(x) - (f-f_t)(x)\Big| > \lambda\Big\}\Big| \\
&\qquad \leq \Big|\Big\{x : \tilde{\mathcal{M}}(f-f_t)(x) > \frac{\lambda}{2}\Big\}\Big| + \Big|\Big\{x : |f-f_t|(x) > \frac{\lambda}{2}\Big\}\Big| \\
&\qquad \leq \frac{2^p\|\tilde{\mathcal{M}}(f-f_t)\|_p^p}{\lambda^p} + \frac{2^p\|f-f_t\|_p^p}{\lambda^p} \leq C_p\lambda^{-p}t^p,
\end{aligned}$$

where the last step uses the boundedness of $\tilde{\mathcal{M}}$ on L^p. Since t was arbitrary, (8.2) and hence (8.1) are proved.

The proof of (1.8) is similar, except that we use the bounds on the maximal operator $\tilde{\mathfrak{M}}$ instead of $\tilde{\mathcal{M}}$. The details are left to the interested reader.

8.2. **The L^1 case.** The following proposition, due to David Preiss (private communication), shows that (1.8) cannot hold for all $f \in L^1(\mathbb{R})$ if μ is a probability measure singular with respect to Lebesgue.

Proposition 8.1. *Suppose that μ is a probability measure on $\mathbb{R}$ such that its restriction to $\mathbb{R}\setminus\{0\}$ is not absolutely continuous with respect to the Lebesgue measure. Then there is a function $f \in L^1(\mathbb{R})$ such that for every $x \in \mathbb{R}$ the set*

$$Z_x = \Big\{r \in (0,\infty) : \int f(x+ry)d\mu(y) = \infty\Big\}$$

is dense in $(0,\infty)$.

Proof. We may choose an $x_0 \neq 0$ such that $\mu(x_0 - r, x_0 + r)/(2r) \to \infty$ as $r \searrow 0$ (see [33, Theorem 7.15]). In particular, there is a $\rho_0 > 0$ and a continuous function $\eta : (0,\infty) \to [0,\infty)$ such that $\eta(r) \to \infty$ as $r \searrow 0$, $\eta \equiv 0$ on $[\rho_0,\infty)$, η is strictly decreasing on $(0,\rho_0)$, and

$$\frac{\mu(x_0 - r, x_0 + r)}{2r} \geq \eta(r) \text{ for all } r \in (0,\rho_0). \tag{8.3}$$

Let $g \in L^1(0,\infty)$ be a continuous, nonnegative and strictly decreasing function such that

$$\int_0^\infty g(y)\eta(\lambda g(y))dy = \infty \text{ for any } \lambda > 0.$$

(For a construction of such a function, see Subsection 9.2.) Let $h(x) = g^{-1}(|x|)$ for $0 < |x| < \rho_0$ and $h(x) = 0$ for $|x| > \rho_0$. Define

$$f(x) = \sum_{j=1}^\infty 2^{-j}h(x - x_j),$$

where the sequence $\{x_j\}_{j=1}^{\infty}$ is dense in $\mathbb{R}$. Then $f \in L^1(\mathbb{R})$, since

$$\int_{\mathbb{R}} f(x)dx = \int_{\mathbb{R}} h(x)dx = \int_0^\infty |\{x :\ h(x) > t\}|dt = 2\int_0^\infty g(t)dt < \infty.$$

We must prove that for any $x \in \mathbb{R}$ and $a < b$, the interval (a, b) contains a point of Z_x. Indeed, by the density of $\{x_j\}$ there is a $j \geq 1$ such that $r := (x_j - x)/x_0 \in (a, b)$. Then

$$\begin{aligned}
\int h(x - x_j + ry)d\mu(y) &= \int h(r(y - x_0))d\mu(y) \\
&= \int_0^\infty \mu(\{y :\ h(r(y - x_0)) > t\})dt \\
&= \int_0^\infty \mu\Big(x_0 - \frac{g(t)}{r}, x_0 + \frac{g(t)}{r}\Big)dt \\
&\geq \int_0^\infty \frac{g(t)}{r}\eta\Big(\frac{g(t)}{r}\Big)dt = \infty,
\end{aligned} \tag{8.4}$$

so that

$$\int f(x + ry)d\mu(y) \geq \sum_{j=1}^{\infty} 2^{-j}\int h(x - x_j + ry)d\mu(y) = \infty \tag{8.5}$$

as required. □

Remark. The above argument can be adapted to show that (1.7) cannot hold for all $f \in L^1(\mathbb{R})$ if $\{E_k\}$ is a decreasing sequence of subsets of $[1, 2]$ (or any other interval separated from zero) with $|E_k| \to 0$. Namely, fix any such sequence $\{E_k\}$ and let $\phi_k = \mathbf{1}_{E_k}/|E_k|$ as before. Then there is a subsequence $\{\phi_{j_k}\}_{k=1}^{\infty}$ converging weakly to a probability measure μ supported on a set E of measure 0. Without loss of generality we may assume that $j_k = k$. Let also μ_k be the absolutely continuous measure with density ϕ_k. We claim that

$$\lim_{k\to\infty} \int f(x + ry)d\mu_k(y)dy = \infty \text{ for all } r \in Z_x,\ x \in \mathbb{R}. \tag{8.6}$$

To prove (8.6), we first ask the reader to verify that (8.3) implies the following statement: for every $\rho_1 > 0$ there is a $K = K(\rho_1)$ such that

$$\frac{\mu_k(x_0 - \rho, x_0 + \rho)}{2\rho} \geq \frac{1}{4}\eta(\rho) \text{ for all } \rho_1 < \rho < \rho_0,\ k > K(\rho_1).$$

With g, h, f as above, we then have as in (8.4)

$$\begin{aligned}
\int h(x - x_j + ry)d\mu_k(y) &= \int_0^\infty \mu_k\Big(x_0 - \frac{g(t)}{r}, x_0 + \frac{g(t)}{r}\Big)dt \\
&\geq \int_0^R \frac{g(t)}{2r}\eta\Big(\frac{g(t)}{r}\Big)dt,
\end{aligned}$$

for any $R > 0$, provided that $k > K(g(R)/r)$. Since the last integral can be made arbitrarily large as $R \to \infty$, (8.6) follows as in (8.5).

9. Appendix

9.1. **Fourier analytic estimates.** We now discuss the Fourier analytic estimates for μ, as indicated in Remark 1 following Theorems 1.1 - 1.3. Our arguments are very similar to those in Section 6 of [24], hence we only give an outline of the proof and leave the details to the reader. The main result is the following.

Proposition 9.1. *Assume that (5.4) holds. Then there exists a sequence of sets $\{S_k : k \geq 1\}$ that satisfies, in addition to the conditions (a)-(d) of Theorem 5.1, the following estimate: for all $\xi \in \mathbb{R}$,*

$$|\widehat{\sigma}_k(\xi)| \leq B \min\left[1, \frac{M_{k+1}}{|\xi|}\right] \left[\frac{2^{k+3}\ln(BM_{k+1})}{\prod_{j=1}^{k+1} N_j^{1-\epsilon_j}}\right]^{\frac{1}{2}}. \tag{9.1}$$

Proof. It will suffice to prove (9.1) for $\xi \in \mathbb{Z}$, since the more general statement then follows by standard arguments (see e.g. Lemma 9.A.4 in [42]). Setting

$$\mathfrak{S}_k(\xi) = \sum_{\mathbf{i}\in\mathbb{I}_k} X_k(\mathbf{i}) e^{-2\pi i\xi\alpha_k(\mathbf{i})},$$

we observe after a brief calculation that

$$\begin{aligned}
\widehat{\sigma}_k(\xi) &= \frac{1-e^{-2\pi i\xi\delta_{k+1}}}{2\pi i\xi\delta_{k+1}} P_{k+1}^{-1}\mathfrak{S}_{k+1}(\xi) - \frac{1-e^{-2\pi i\xi\delta_k}}{2\pi i\xi\delta_k} P_k^{-1}\mathfrak{S}_k(\xi) \\
&= \frac{1-e^{-2\pi i\xi\delta_{k+1}}}{2\pi i\xi\delta_{k+1}} \left[\Xi_1(\xi) + \Xi_2(\xi)\right], \text{ where} \\
\Xi_1(\xi) &= \left[P_{k+1}^{-1} - Q_{k+1}^{-1}\right] \mathfrak{S}_{k+1}(\xi), \text{ and} \\
\Xi_2(\xi) &= Q_{k+1}^{-1} \sum_{\mathbf{i}\in\mathbb{I}_k} X_k(\mathbf{i}) \sum_{i_{k+1}} \big(Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}\big) e^{-2\pi i\xi\alpha_{k+1}(\bar{\mathbf{i}})}.
\end{aligned}$$

Since $|(1-e^{-2\pi i\xi\delta_{k+1}})/(2\pi i\xi\delta_{k+1})| \leq C\min(1, M_{k+1}/|\xi|)$ with an absolute constant C, and both Ξ_1 and Ξ_2 are M_{k+1}-periodic, it suffices to show that

$$|\Xi_1(\xi)| + |\Xi_2(\xi)| \leq B\left[\frac{2^{k+3}\ln(BM_{k+1})}{\prod_{j=1}^{k+1} N_j^{1-\epsilon_j}}\right]^{\frac{1}{2}} \text{ for } \xi \in \{1, 2, \cdots, M_{k+1}\}.$$

For Ξ_1, this follows even without the logarithmic term from parts (a) and (b) of Theorem 5.1 and the trivial bound $|\mathfrak{S}_{k+1}(\xi)| \leq P_{k+1}$. For Ξ_2, this is a consequence of Bernstein's inequality (Lemma 5.3) applied to the random variables

$$Z_{\mathbf{i}} = \frac{1}{N_{k+1}} \sum_{i_{k+1}=1}^{N_{k+1}} \Big(Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}\Big) e^{-2\pi i\xi\alpha_{k+1}(\bar{\mathbf{i}})}$$

where $\mathbf{i}$ ranges over the indices with $X_k(\mathbf{i}) = 1$, $m = P_k$, $\sigma^2 = P_k p_{k+1}/N_{k+1}$ and $\lambda = \sqrt{8p_{k+1}\ln(BM_{k+1})/(P_k N_{k+1})}$. We omit the details. □

Lemma 9.2. *Assume that S_k have been chosen as in Proposition 9.1, with N_k, ϵ_k given by either (6.1) or (6.2). Then in addition to all conclusions of Lemma 6.1, the limiting measure μ satisfies*

$$|\widehat{\mu}(\xi)| \leq C_\alpha |\xi|^{-\frac{\beta}{2}+\alpha} \quad \textit{for all } \alpha > 0, \tag{9.2}$$

where $\beta = 1$ if (6.1) holds and $\beta = 1 - \epsilon$ if (6.2) holds.

Proof. Assume first that (6.1) holds. We then ask the reader to verify that

$$\begin{aligned} |\widehat{\mu}(\xi)| &\leq \sum_{k=0}^{\infty} \left|\widehat{\sigma}_k(\xi)\right| \\ &\leq B \sum_{k=0}^{\infty} \min\left[1, \frac{N^{\frac{(k+1)(k+4)}{2}}}{|\xi|}\right] \left[\frac{2^{k+3}\left(\ln B + \frac{(k+1)(k+4)}{2}\ln N\right)}{N^{\frac{(k+1)(k+2)}{2}}}\right]^{\frac{1}{2}} \\ &\leq C|\xi|^{-\frac{1}{2}} C^{\sqrt{\ln|\xi|}}, \end{aligned}$$

which implies (9.2) with $\beta = 1$. The proof for (6.2) is similar. □

9.2. **A claim in Subsection 8.2.** In this subsection we describe the construction of the function g used in the proof of Proposition 8.1.

Lemma 9.3. *Given any sequence $N_j \to \infty$, there exist positive constants $\{\gamma_j\}$ such that*

$$\sum_j \gamma_j < \infty, \quad \textit{and} \quad \sum_j \gamma_{j+k} N_j = \infty \textit{ for all } k \geq 0.$$

Proof. We pick a fast-growing subsequence $\{n_j\}$ of the integers such that $n_0 = 0$,

$$\sum_j N_{n_j}^{-1} < \infty, \quad \text{and} \quad n_{j+1} - n_j \nearrow \infty \text{ as } j \to \infty. \tag{9.3}$$

Any positive integer can be written uniquely in the form $n_j + \ell$ for some $0 \leq \ell < n_{j+1} - n_j$, and we set $\gamma_{n_j+\ell} = 2^{-\ell} N_{n_j}^{-1}$. Then

$$\sum_j \gamma_j = \sum_j \sum_{\ell < n_{j+1}-n_j} \gamma_{n_j+\ell} < \sum_{j=1}^{\infty} \frac{1}{N_{n_j}} \sum_{l=0}^{\infty} 2^{-\ell} \leq 2 \sum_{j=1}^{\infty} N_{n_j}^{-1} < \infty,$$

where the last step follows from the first condition in (9.3). On the other hand, given any $k \geq 0$ there exists by the second condition in (9.3) an integer j_0 such that $k < n_{j+1} - n_j$ for all $j \geq j_0$. This implies that

$$\sum_j \gamma_{j+k} N_j \geq \sum_{j \geq j_0} \gamma_{n_j+k} N_{n_j} = 2^{-k} \sum_{j \geq j_0} \frac{1}{N_{n_j}} N_{n_j} = \infty.$$

Thus both claims are proved. □

Lemma 9.4. *Let* $\eta : (0,\infty) \to [0,\infty)$ *be a continuous function such that*

$$\eta \equiv 0 \text{ on } [\rho_0,\infty), \quad \eta \text{ strictly decreasing on } (0,\rho_0], \text{ and } \eta(r) \to \infty \text{ as } r \searrow 0$$

for some $\rho_0 > 0$. *Then there exists a continuous, non-negative and strictly decreasing function* $g \in L^1(0,\infty)$ *such that*

$$\int_0^\infty g(y)\eta(\lambda g(y))\, dy = \infty \text{ for any } \lambda > 0.$$

Proof. Set $\beta_0 = 0$. For a sequence of positive numbers $\{\beta_j : j \geq 1\}$ soon to be specified, we will define g as follows:

$$g(\beta_0) = \rho_0,\ g(\beta_1) = \rho_0 2^{-1},\ g(\beta_1+\beta_2) = \rho_0 2^{-2},\ \cdots, g(\beta_1+\cdots+\beta_j) = \rho_0 2^{-j},$$

and g is linear in the interval $[\beta_0+\cdots+\beta_j, \beta_0+\cdots+\beta_j+\beta_{j+1}]$ subject to the above constraints. Thus g is a piecewise linear, continuous, non-negative and strictly decreasing function, for which

$$\begin{aligned}
\int_0^\infty g(y)\, dy &= \sum_{j=0}^\infty \int g(y)\mathbf{1}_{[\rho_0 2^{-(j+1)},\rho_0 2^{-j})}(g(y))\, dy \\
&\leq \sum_{j=0}^\infty \rho_0 2^{-j} \left|\{y : \rho_0 2^{-(j+1)} \leq g(y) < \rho_0 2^{-j}\}\right| \qquad (9.4)\\
&= \rho_0 \sum_{j=0}^\infty 2^{-j}\beta_{j+1} = 2\rho_0 \sum_{j=0}^\infty 2^{-(j+1)}\beta_{j+1}.
\end{aligned}$$

Also, set $N_j = \eta(\rho_0 2^{-j})$, so that $N_j \nearrow \infty$ as $j \to \infty$. Then given any $\lambda > 0$, there is an integer $k \geq 0$ such that $\lambda \leq 2^k$, and a similar calculation yields

$$\begin{aligned}
\int_0^\infty g(y)\eta(\lambda g(y))\, dy &= \sum_{j\geq 0} \int g(y)\eta(\lambda g(y))\mathbf{1}_{[\rho_0 2^{-(j+1)},\rho_0 2^{-j})}(g(y))\, dy \\
&\geq \rho_0 \sum_{j\geq 0} 2^{-(j+1)}\eta(\rho_0 2^{k-j})\beta_{j+1} \qquad (9.5)\\
&= \rho_0 \sum_{j=k}^\infty 2^{-(j+1)}\beta_{j+1} N_{j-k} \\
&= \rho_0 \sum_{j=0}^\infty 2^{-(j+k+1)}\beta_{j+k+1} N_j.
\end{aligned}$$

Here we used the monotonicity of η at the second step. Now choose the numbers β_j such that $2^{-j}\beta_j = \gamma_j$, where the constants γ_j are as specified by Lemma 9.3. Then the infinite sum on the rightmost side of (9.4) converges, while the one on the rightmost side of (9.5) diverges for all $k \geq 0$. This completes the proof. □

References

1. N. Alon and J.H. Spencer, *The Probabilistic Method*, Wiley Interscience series in discrete mathematics and optimization, second edition, 2000.
2. V. Aversa and D. Preiss, *Hearts density theorems*, Real Analysis Exchange **13(1)** (1987), 28–32.
3. V. Aversa and D. Preiss, *Sistemi di derivazione invarianti per affinita*, preprint, Complesso Universitario di Monte S. Angelo, Napoli, 1995.
4. C. Bennett and R. Sharpley, *Interpolation of operators*, Academic Press 1988.
5. C. Bluhm, *Random recursive construction of Salem sets*, Ark. Mat. **34** (1996), 51–63.
6. C. Bluhm, *On a theorem of Kaufman: Cantor-type construction of linear fractal Salem sets*, Ark. Mat. **36** (1998), 307–316.
7. A. M. Bruckner, *Differentiation of integrals*, Amer. Math. Monthly **78**(9) (1971), Part 2: pp. i–iii+1–51.
8. J. Bourgain, *Averages in the plane over convex curves and maximal operators*, J. Analyse Math. **47** (1986), 69–85.
9. M. Christ, *Endpoint bounds for singular fractional integral operators*, preprint, unpublished (1988).
10. M. Christ, A. Nagel, E.M. Stein, S. Wainger, *Singular and maximal Radon transforms: Analysis and geometry*, Ann. Math. **150** (1999), 489–577.
11. M. Csörnyei, *Density theorems revisited*, Acta Sci. Math. (Szeged), **64** (1998), 59–65.
12. M. de Guzman, *Differentiation of integrals in* $\mathbb{R}^n$, Lecture Notes in Math., no. 481, Springer-Verlag 1975.
13. J. Duoandikoetxea and J. Rubio de Francia, *Maximal and singular integral operators via Fourier transform estimates*, Invent. Math. **84(3)**, 541–561.
14. K. Falconer, *The geometry of fractal sets*, Cambridge Univ. Press 1985.
15. K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons Ltd., 1995.
16. W.T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465–588.
17. B. Green, *Arithmetic progressions in sumsets*, Geom. Funct. Anal. **12** (2002), 584–597.
18. B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), 481–547.
19. A. Greenleaf, A. Seeger and S. Wainger, *On* X*-ray transforms for rigid line complexes and integrals over curves in* $\mathbb{R}^4$, Proc. Amer. Math. Soc **127** (1999), 3533-3545.
20. G. Hardy, J. Littlewood, *A maximal theorem with function-theoretic applications*, Acta Math. **54** (1930), 81–116.
21. C. A. Hayes and C. Y. Pauc, *Derivation and martingales*, Springer-Verlag 1970.
22. J.P. Kahane, *Some Random Series of Functions*, Cambridge Univ. Press, 1985.
23. L. Kaufman, *On the theorem of Jarnik and Besicovitch*, Acta Arith. **39** (1981), 265–267.
24. I. Łaba and M. Pramanik, *Arithmetic progressions in sets of fractional Hausdorff dimension*, to appear in Geom. Funct. Anal. (2009).
25. J.M. Marstrand, *Packing circles in the plane*, Proc. London Math. Soc. **55** (1987), 37–58.
26. P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, 1995.
27. G. Mockenhaupt, *Salem sets and restriction properties of Fourier transforms*, Geom. Funct. Anal. **10** (2000), 1579–1587.
28. A. Nagel, E.M. Stein and J. Wright, *Differentiation in lacunary directions*, Proc. Nat. Acad. Sci. U.S.A. **75(3)** (1978), 1060–1062.
29. D.H. Phong, E.M. Stein, *Hilbert integrals, singular integrals, and Radon transforms I.* Acta Math. **157** (1986), 99–157.

30. D.H. Phong, E.M. Stein, *Hilbert integrals, singular integrals, and Radon transforms I.* Invent. Math. **86** (1986), 75–113.
31. D. Preiss, lectures given in Ravello (1985), not published.
32. J.L. Rubio de Francia, *Maximal functions and Fourier transforms*, Duke Math. J. **53** (1986), 395–404.
33. W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, Inc., 1987.
34. R. Salem, *On singular monotonic functions whose spectrum has a given Hausdorff dimension*, Ark. Mat. **1** (1950), 353–365.
35. W. Schlag, *$L^p \to L^q$ estimates for the circular maximal function*, in: *Topics in Analysis and Its Applications. Selected Theses*, R. Coifman, ed., World Scientific Publishing Co., River Edge, NJ, 2000.
36. W. Schlag, *A geometric proof of Bourgain's circular maximal theorem*, Duke. J. Math. **93** (1998), 505–533.
37. E.M. Stein, *Maximal functions: Spherical means*, Proc. Nat. Acad. Sci. U.S.A. **73** (1976), 2174–2175.
38. E.M. Stein, *Oscillatory integrals in Fourier analysis*, in *Beijing Lectures in Harmonic Analysis* (E.M. Stein, ed.), Ann. Math. Study # 112, Princeton Univ. Press, 1986, 307–355.
39. E.M. Stein, *Harmonic Analysis*, Princeton Univ. Press, Princeton, 1993.
40. J.A. Wellner, *Notes on exponential bounds via Azuma's inequality*, preprint, 2004.
41. N. Wiener, *The ergodic theorem*, Duke Math. J. **5** (1939), 1–18.
42. T. Wolff, *Lectures on Harmonic Analysis*, I. Łaba and C. Shubin, eds., Amer. Math. Soc., Providence, R.I. (2003).

Izabella Łaba and Malabika Pramanik
Department of Mathematics
University of British Columbia
Vancouver, B.C. V6T 1Z2, Canada
E-mail address: ilaba@math.ubc.ca
E-mail address: malabika@math.ubc.ca