

Maximal operators and differentiation theorems for sparse sets

Izabella Łaba and Malabika Pramanik

May 30, 2009

Abstract

We study maximal averages associated with singular measures on \mathbb{R} . Our main result is a construction of singular Cantor-type measures for which the corresponding maximal operators are bounded on L^p with $p \geq 2$. We also prove the L^p boundedness of the restricted (single-scale) maximal operator for all $p > 1$, as well as $L^p \rightarrow L^q$ bounds with appropriate scaling. 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 the proof of Bourgain's circular maximal theorem in two dimensions.

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

1 Introduction

1.1 Maximal operators

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

$$\tilde{\mathcal{M}}f(x) := \sup_{t>0, k \geq 1} \frac{1}{|S_k|} \int_{S_k} |f(x + ty)| dy. \quad (1.1)$$

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 $|S_k| \rightarrow 0$ as $k \rightarrow \infty$.

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 :

$$\tilde{\mathfrak{M}}f(x) := \sup_{t>0} \int |f(x+ty)| d\mu(y). \quad (1.2)$$

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.

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

$$\tilde{\mathcal{M}}^a f(x) := \sup_{t>0, n \geq 1} t^a \int |f(x+ty)| \phi_n(y) dy, \quad (1.3)$$

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

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

$$\mathcal{M}f(x) := \sup_{1 < t < 2, n \geq 1} \frac{1}{|S_k|} \int_{S_k} |f(x+ty)| dy, \quad (1.5)$$

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

where the range of t is limited to a single scale. These operators will play a critical role in the proofs of the unrestricted maximal estimates. Furthermore, we are able to prove maximal estimates for \mathcal{M} and \mathfrak{M} with a better range of exponents; this is likely be of interest from the point of view of possible applications in harmonic analysis, where restricted maximal estimates are often sufficient.

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 $k \rightarrow \infty$,*
- (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 $1 < p \leq q < \infty$ and $q \geq 2$, 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 $p \geq 2$.*

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. Then for every $f \in L^p(\mathbb{R})$ with $p \in [2, \infty)$ we have*

$$\limsup_{r \rightarrow 0} \sup_k \left| \frac{1}{r|S_k|} \int_{x+rS_k} f(y)dy - f(x) \right| = 0 \text{ for a.e. } x \in \mathbb{R}. \quad (1.7)$$

The limiting set $S = \bigcup_{k=1}^{\infty} S_k$ 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 $\{S_k\}$ contained in $[1, 2]$ such that S_k obey conditions (a)–(c) of Theorem 1.1 and:*

- (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

$$\frac{1 + \epsilon}{1 - \epsilon} < p < \infty \text{ and } 1 < q < \frac{1 - \epsilon}{2\epsilon}p, \quad (1.8)$$

and from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ for any p, q such that $p \leq q$ and (1.8) holds.

(c) The unrestricted maximal operators $\tilde{\mathcal{M}}^a$ and $\tilde{\mathfrak{M}}^a$ are bounded from $L^p[0, 1]$ to $L^q(\mathbb{R})$ for any p, q such that (1.8) holds, $p \leq q$ and $q \geq 2$, 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 $p \geq 2$.

(d) The family of sets $\mathcal{S} = \{rS_k : k \geq 1\}$ differentiates $L^p(\mathbb{R})$ in the sense of (1.7) for all $p \geq 2$.

Remarks.

1. The statements of Theorems 1.1 and 1.3 do not mention any Fourier-analytic estimates on μ . In fact it is possible to use the ideas of [21] to construct, with some additional effort, sequences of sets S_k such that, in addition to the existing conclusions of Theorems 1.1 and 1.3, the limiting set $S = \bigcap_{k=1}^{\infty} S_k$ is a *Salem set*.
2. What is possibly of greater interest is 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.

See §1.3.2 for the definitions and more details.

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 [32] for $d \geq 3$ and Bourgain [8] for $d = 2$. We state it here for future reference.

Theorem 1.4. (Stein [32], Bourgain [8]) *Recall the spherical maximal operator in \mathbb{R}^d :*

$$\tilde{\mathfrak{M}}_{\mathbb{S}^{d-1}} f(x) = \sup_{t>0} \int_{\mathbb{S}^{d-1}} |f(x + ty)| d\sigma(y), \quad (1.9)$$

where σ is the normalized Lebesgue measure on the unit sphere \mathbb{S}^{d-1} . Then

$$\|\tilde{\mathfrak{M}}_{\mathbb{S}^{d-1}} f(x)\|_{L^p(\mathbb{R}^d)} \leq C \|f\|_{L^p(\mathbb{R}^d)}, \quad p > \frac{d}{d-1}, \quad (1.10)$$

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 [34], [9], [25], [26] 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 $\epsilon \rightarrow 0$. Theorem 1.3 provides an affirmative answer to this question in the context of the restricted operator \mathfrak{M} . Theorem 1.1 may be viewed as the limiting situation as $\epsilon \rightarrow 0$ (compare with Theorem 1.4 as $n \rightarrow \infty$) 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. The

L^p boundedness of maximal averages in this context has long been a productive avenue of research. Our result may be viewed as an extension of the following result by Rubio de Francia [28]. We write $\widehat{\mu}(\xi) = \int e^{-2\pi i \xi x} d\mu(x)$.

Theorem 1.5. (*Rubio de Francia [28]*) *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 \mathfrak{M}_σ , 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 $a = (d-1)/2 > \frac{1}{2}$, but it fails to capture the circular maximal estimate ($d = 2$) 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 in dimension 1 to obey the assumptions of Theorem 1.5 as stated above. Indeed, it is well-known (see [29]) that

$$\sup \left\{ a > 0 \mid \begin{array}{l} \text{there exists a finite Borel measure } \mu \text{ on } \mathbb{R} \text{ that is} \\ \text{singular with respect to Lebesgue and for which} \\ \sup_{\xi \in \mathbb{R}} |\widehat{\mu}(\xi)(1 + |\xi|)^a| < \infty \end{array} \right\} = \frac{1}{2}.$$

In particular, a singular measure on \mathbb{R} cannot obey (1.11) with $a > \frac{1}{2}$. However, 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

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

It is well known ([23], [12]) that

$$\dim_{\mathbb{F}}(S) \geq \dim_{\mathbb{H}}(S) \text{ for all compact } S \subset \mathbb{R}, \tag{1.12}$$

and that the inequality in (1.12) is often sharp. Whenever $\dim_{\mathbb{F}}(S) = \dim_{\mathbb{H}}(S)$, we say that $S \subset \mathbb{R}$ is a *Salem set*. There are many constructions of Salem sets in the literature, including [29], [20], [5], [6], [19], [21]. It is

therefore of interest to ask whether there is an analogue of Theorem 1.5 that might apply to Salem sets on \mathbb{R} .

Our Theorems 1.1 and 1.3 are related to Theorem 1.5 as follows. 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 [21]), then (4.2) may be viewed as analogous to higher-order uniformity conditions in additive combinatorics (cf. [14], [16]). 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 $a < \frac{1}{4}$ (the optimal range would be $a \leq \frac{1}{2}$).

Alternatively, it is possible with some additional effort 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 by adding the appropriate Fourier-analytic conditions to Theorem 5.1 and proving them along the same lines as in [21, Section 6]. Since the Fourier decay is not actually used in the proofs of any of our theorems, we omit the details.

Finally, note that our maximal estimates on \mathfrak{M} hold for all $p \in [2, \infty]$, which is better than $(\frac{2a+1}{2a}, \infty]$ when $a \leq \frac{1}{2}$. It is known that the range $p > (2a+1)/(2a)$ in Theorem 1.5 is sharp, at least for $a > \frac{1}{2}$, if no additional hypotheses are made beyond the Fourier decay of μ . Our Theorems 1.1 and 1.3 show that the range of p can be improved if the Fourier decay assumption is replaced by the stronger correlation condition (4.2). 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 $a < \frac{1}{2}$.

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 that we pursue in this article. We do not attempt to survey the vast literature on density theorems and differentiation of integrals (see [7], [11] for more information) and focus only on the specific problems relevant

to the present discussion.

The following question was raised and investigated by Preiss [27] 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

$$\lim_{S \in \mathcal{S}, \text{diam}(S \cup \{0\}) \rightarrow 0} \frac{|(x + S) \cap E|}{|S|} = 1 \text{ for a.e. } x \in E. \quad (1.13)$$

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 $\mathcal{S} = \{I_n\}_{n=1}^\infty$, where $I_n = (\frac{n}{(n+1)!}, \frac{1}{n!})$. We have $|I_n| = \frac{1}{(n+1)!}$ and $\text{diam}(I_n \cup \{0\}) = \frac{1}{n!}$, 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 [27] and Aversa-Preiss [2] (see also [10] 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 (1.13) does not hold with \mathcal{S} replaced by \mathcal{S}' . (Note that the limit in (1.13) is now being taken over the two parameters n and r .) In fact, Aversa-Preiss prove in [2] that no sequence of *intervals* I_k can generate an affine invariant density system unless $\liminf_{k \rightarrow \infty} |I_k|/\text{diam}(I_k \cup \{0\}) > 0$, 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 $|S_k| \rightarrow 0$ and:*

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

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

(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 $L^p_{\text{loc}}(\mathbb{R})$ for some $1 \leq p \leq \infty$ if for every $f \in L^p_{\text{loc}}(\mathbb{R})$ we have

$$\lim_{S \in \mathcal{S}, \text{diam}(S \cup \{0\}) \rightarrow 0} \frac{1}{|S|} \int_{x+S} f(y) dy = f(x) \text{ for a.e. } x \in \mathbb{R}. \quad (1.14)$$

For instance, the Lebesgue differentiation theorem states that the collection $\{(-r, r) : r > 0\}$ differentiates $L^1_{\text{loc}}(\mathbb{R})$. Note that the differentiation property (1.14) implies the density property (1.13), 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 estimates of Hardy-Littlewood [17] and Wiener [36], which easily imply 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 (1.14). This is equivalent to our Theorem 1.2: although we state it for all $p \geq 2$, (1.14) is a local property

¹This 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 $L^p_{\text{loc}}(\mathbb{R})$.

and $L_{\text{loc}}^p \subset L_{\text{loc}}^2$ for $p > 2$. Our maximal estimates in Theorem 1.1 provide an analogue of the Hardy-Littlewood estimates for sparse sets and imply the Aversa-Preiss conjecture along the lines of the standard Hardy-Littlewood argument.

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 $[1, 2]$. Accordingly, the proof is divided into two parts. In the first part (Sections 2–4 and 7), we formulate the needed deterministic pseudorandomness conditions on S_k and prove that they indeed guarantee the boundedness of our maximal operators as required by Theorems 1.1 and 1.3. Our arguments here are largely inspired by considerations from multidimensional harmonic analysis, and in particular Bourgain’s proof of the circular maximal theorem [8]. In the second part (Sections 5 and 6), we give a probabilistic construction of sets S_k that obey these conditions. The construction here is somewhat similar to that in [21, Section 6], but significantly more complicated.

Our main task is to prove appropriate $L^p \rightarrow L^q$ bounds on 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 . Once such bounds are obtained, summation over k and interpolation yields similar bounds on the restricted maximal operator \mathcal{M} , and the bounds on $\tilde{\mathcal{M}}^a$ are obtained through the scaling analysis in Section 7. A further preliminary reduction is to replace \mathcal{M}_k by its discretized counterparts, the discretization being in the space of affine transformations. All the reductions discussed so far are standard in the harmonic analysis literature, even though the exact details are specific to the problem at hand. We will follow the approach of [8], [31], and especially [30], with relatively minor modifications.

We now pause for a moment to recall the main ideas of Bourgain’s proof of the circular maximal theorem in [8]. Having performed the appropriate preliminary reductions, Bourgain considers discretized maximal operators of

the form

$$\Phi_k f(x) = \frac{1}{|E_{x,k}|} \int_{E_{x,k}} f(z) dz.$$

Here each $E_{x,k}$ is an annulus of thickness 2^{-k} and radius r_x centered at x , with r_x taking values in a discrete subset of $[1, 2]$. We wish to prove that Φ_k is bounded on L^p for $p > 2$, or equivalently that its dual

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

is bounded on $L^{p'}$. In fact it suffices to prove this for characteristic functions, since then the general case follows by interpolation. Let $g = \mathbf{1}_\Omega$, then

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

This is trivially bounded on L^1 , and the proof would be complete if we could prove a similar bound on L^2 . We have

$$\begin{aligned} \|\Phi_k^* g\|_2^2 &= \int \left| \int_\Omega \frac{1}{|E_{x,k}|} \mathbf{1}_{E_{x,k}}(z) dx \right|^2 dx \\ &= \int \int_{\Omega \times \Omega} \frac{1}{|E_{x,k}| |E_{y,k}|} \mathbf{1}_{E_{x,k}}(z) \mathbf{1}_{E_{y,k}}(z) dx dy dz \\ &= \int_{\Omega \times \Omega} \frac{1}{|E_{x,k}| |E_{y,k}|} |E_{x,k} \cap E_{y,k}| dx dy. \end{aligned} \tag{1.15}$$

If we had

$$|E_{x,k} \cap E_{y,k}| \leq C_k |E_{x,k}| |E_{y,k}|, \tag{1.16}$$

the needed L^2 bound would follow. Unfortunately, (1.16) need not always hold. Specifically, if the two annuli are ‘‘internally tangent’’ in a clamshell configuration, one inside the other with

$$|x - y| = |r_x - r_y| \tag{1.17}$$

up to $2^{-k+k\epsilon}$ -errors, the area of the intersection on the left side of (1.16) can easily be much larger than $|E_{x,k}| |E_{y,k}| \approx 2^{-2k}$.

Bourgain’s key observation is that (1.17), together with additional geometric information specific to circles, restricts those $(x, y) \in \Omega^2$ for which internal tangencies are possible to a rather small set. Once the internal

tangencies are eliminated from consideration, a generic (or *transverse*) intersection of two annuli does have reduced area. This allows Bourgain to split the region of integration into 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 (1.17). An interpolation argument completes the proof.

Let us now try to apply a similar argument in our setting, with p restricted for the moment 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 (1.15). While we still expect that generic double intersections should be significantly smaller than $|E_k|$, 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 $(x + rS_{k+1}) \cap (y + sS_{k+1})$?

Consider first an intersection of the form

$$(x + r(I \cap S_{k+1})) \cap (y + s(J \cap S_{k+1})), \quad (1.18)$$

where I and J are two different intervals of S_k . In this case, the S_{k+1} -subintervals of I and J were chosen independently, hence the intersection 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 $J \cap S_{k+1}$.

This argument does not apply directly to intersections of the form (1.18) with $I = J$. However, if x and y are sufficiently far apart, the intersection of $x + r(I \cap S_{k+1})$ and $y + s(J \cap S_{k+1})$ will consist mostly of intersections of distinct $(k+1)$ -level intervals of S_{k+1} . We may then apply the same argument as above, but at the level $k+1$ instead of k , to obtain the desired gain. The procedure may be iterated as needed.

Following Bourgain, we will refer to the first type of intersections ((1.18) 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 have a gain as described above. If on the other hand there are many internal tangencies, that leads to conditions analogous to (1.17) above. In fact, it implies the stronger statement that both $|x - y|$ and $|t - s|$ must be small, 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 (1.15) which involve n -fold intersections of affine copies of S_k .

The exact statement of the intersection bound that we need is given by the transverse correlation condition (4.2) in Section 4. The correlation condition may be viewed as a multiscale analogue of the higher order uniformity conditions in additive combinatorics, see e.g. [14], [16]. We hope to explore this connection further in a subsequent paper.

It remains to construct a sequence of sets S_k which satisfies our correlation condition. This is accomplished in Sections 5 and 6. The construction is given by a randomized Cantor-type iteration. Having already chosen S_k at the k -th step, at the next step of the iteration we subdivide each of its intervals and let each such subinterval be in S_{k+1} independently with a fixed probability p_{k+1} . We then use large deviation inequalities (specifically, Bernstein's inequality and Azuma's inequality) to prove that there is a positive probability that S_{k+1} obeys the correlation condition.

The paper is arranged as follows. In Section 2, we describe the general class of Cantor-type constructions used in this paper, introduce the notation to describe the n -fold intersections of affine copies of S_k , and collect some basic facts regarding such intersections. The preliminary reductions of discretization, linearization and interpolation are carried out in Section 3.1; by the end of that section the question of proving L^p estimates for \mathcal{M} is reduced to proving restricted weak type $L^{p'}$ estimates for the linearized and discretized operators Φ_k^* . In Section 4 we introduce the correlation condition (4.2) and deduce from it the restricted weak type estimates just mentioned. The actual probabilistic construction of the sets S_k and the proof of (4.2), which contain the technical bulk of the paper, are given in Section 5. In Section 6, we fix the parameters in the random construction and complete the proof of those parts of Theorems 1.1 and 1.3 which concern the restricted maximal operators with the scaling factor $1 < t < 2$. We extend these bounds to the unrestricted maximal operators in Section 7. Finally, in Section 8 we prove Theorem 1.2.

1.5 Acknowledgement

We are grateful to Vincenzo Aversa and Nir Lev for bringing the conjectured Theorem 1.2 to our attention, and to Nir Lev for a further introduction to questions regarding density and differentiation theorems for sparse sets.

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 the 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 sequence of positive integers $\{N_k : k \geq 1\}$,
- (b) certain sequences κ_k and $\tau_{k+1}(\mathbf{i})$ of 0-s and 1-s,

$$\begin{aligned} \kappa_k &= \{\kappa_k(\mathbf{i}) : \mathbf{i} = (i_1, \dots, i_k), 1 \leq i_j \leq N_j, 1 \leq j \leq k\}, \text{ and} \\ \tau_{k+1}(\mathbf{i}) &= \{\tau_{k+1}(\mathbf{i}, j) : 1 \leq j \leq N_{k+1}\} \text{ satisfying} \\ \kappa_{k+1}(\bar{\mathbf{i}}) &= \kappa_k(\mathbf{i})\tau_{k+1}(\bar{\mathbf{i}}), \text{ where } \bar{\mathbf{i}} = (i_1, \dots, i_{k+1}). \end{aligned}$$

Given these quantities, we denote

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

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

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

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

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

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

The construction proceeds as follows. Starting with the interval $[1, 2]$ equipped with the Lebesgue measure, we subdivide it into N_1 intervals $\{I_1(i) : 1 \leq i \leq N_1\}$ of equal length. We choose the P_1 intervals $I_1(i)$ for which $\kappa_1(r) = 1$ 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 $\tau_{k+1}(\mathbf{i})$. In summary, the sets S_k are chosen according to the scheme

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

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

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) $\dim_{\mathbb{H}}(S) \leq \liminf_{k \rightarrow \infty} \log(P_k) / \log(M_k)$.
- (b) $\dim_{\mathbb{H}}(S) \geq s_0 := \liminf_{k \rightarrow \infty} \log(P_k / N_k) / \log(M_{k-1})$.

Proof. The proof of part (a) follows immediately from Proposition 4.1 in [13]. For the proof of part (b), we follow an approach similar to Example 4.6 in [13]. 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

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

The desired conclusion would then follow from Frostman's lemma (see e.g. Proposition 8.2 in [37]).

In order to define ν , we follow a standard procedure due to Caratheodory (see Chapter 4, [23]). 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

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

and a family of outer measures ν_k as follows,

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

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

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

Then ν is a non-negative regular Borel measure of unit mass on subsets of S (Theorem 4.2, [23]).

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

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

Letting $k \rightarrow \infty$ 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. \square

Remark: In our applications, the sequences κ_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 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 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 $\{I_{\mathbf{i}}(k) : \kappa_{\mathbf{i}}(k) = 1\}$ within S_{k-1} is approximately uniform in the following sense:*

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

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

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

Proof. It suffices to show that $\lim_{k \rightarrow \infty} \int f \phi_k$ 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

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

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

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

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 $k' \geq k \geq K$,

$$\left| \int f(x) (\phi_{k'}(x) - \phi_k(x)) dx \right|$$

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

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

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 $k \geq 1$, $x, y \in [0, 1]$, and $r, s \in [1, 2]$. We will be interested in classifying pairs of multi-indices $(\mathbf{i}, \mathbf{j}) \in \mathbb{I}_k^2$ such that

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

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 $k \geq 1, n \geq 2$ and any set

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

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

$$\mathbb{F} = \left\{ (\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{I}_k^n : \bigcap_{\ell=1}^n (c_\ell + r_\ell I_k(\mathbf{i}_\ell)) \neq \emptyset \right\}, \quad (2.11)$$

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

Remarks:

1. We emphasize that \mathbb{F} consists of *all* tuples $(\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{I}_k^n$ 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 x_ℓ and r_ℓ will be chosen from discrete subsets \mathcal{C}, \mathcal{R} of the respective spaces $[0, 1]$ 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 choice of multi-indices $(\mathbf{i}_{\ell'} : 1 \leq \ell' \leq n, \ell' \neq \ell) \in \mathbb{I}_k^{n-1}$,*

$$\max\{\alpha_k(\mathbf{i}_\ell) : (\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{F}\} - \min\{\alpha_k(\mathbf{i}_\ell) : (\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{F}\} \leq 4\delta_k. \quad (2.12)$$

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

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

Proof. The second part of the lemma follows from the first. The inequality in (2.12) is essentially a fact about two-fold intersections. Let us fix $\ell' \neq \ell$ and $(\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{F}$, 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. \square

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 is an injection.*

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

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

$$\begin{aligned} \mathbb{F} &= \mathbb{F}_{\text{int}} \cup \mathbb{F}_{\text{tr}}, \quad \text{where} \quad \mathbb{F}_{\text{int}} := \bigcup_{1 \leq \ell \neq \ell' \leq n} \mathbb{F}_{\text{int}}(\ell, \ell'), \quad \text{with} \\ \mathbb{F}_{\text{int}}(\ell, \ell') &:= \{(\mathbf{i}_1, \dots, \mathbf{i}_n) \in \mathbb{F} : \mathbf{i}'_\ell = \mathbf{i}'_{\ell'}, |i_{\ell k} - i_{\ell' k}| \leq 4\}, \quad \text{and} \\ \mathbb{F}_{\text{tr}} &:= \mathbb{F} \setminus \mathbb{F}_{\text{int}}. \end{aligned}$$

Note that in view of (2.1),

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

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

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

Proof. Since the translation parameters all lie in $[0, 1]$, we may assume without loss of generality that $L > 80n(n-1)$. Using the definition of \mathbb{F}_{int} and pigeonholing we can find indices $\ell \neq \ell'$ such that $\#(\mathbb{F}_{\text{int}}(\ell, \ell')) \geq \frac{2L}{n(n-1)}$. By Lemma 2.4, there exists a further subset $\mathbb{F}^* \subseteq \mathbb{F}_{\text{int}}(\ell, \ell')$ such that

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

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

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

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

$$\begin{aligned} |(c_\ell - c_{\ell'}) + (r_\ell - r_{\ell'})\alpha_k(\mathbf{i}_\ell)| &\leq 2\delta_k + r_{\ell'}|\alpha_k(\mathbf{i}_{\ell'}) - \alpha_k(\mathbf{i}_\ell)| \leq 10\delta_k, \\ |(c_\ell - c_{\ell'}) + (r_\ell - r_{\ell'})\alpha_k(\mathbf{j}_\ell)| &\leq 2\delta_k + r_{\ell'}|\alpha_k(\mathbf{j}_{\ell'}) - \alpha_k(\mathbf{j}_\ell)| \leq 10\delta_k. \end{aligned}$$

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

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

If we now choose $(\mathbf{i}_1, \dots, \mathbf{i}_n), (\mathbf{j}_1, \dots, \mathbf{j}_n) \in \mathbb{F}^*$ so that $|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{j}_\ell)|$ is maximally separated 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. \square

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, \dots, 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*

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

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 $x_\ell + r_\ell S_k$ 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 $\{x_\ell\}$. 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

3.1 Linearization and discretization

In order to analyse the restricted maximal operator \mathcal{M} defined in (1.5), we define the auxiliary restricted maximal operators

$$\mathcal{M}_k f(x) := \sup_{1 < t < 2} \left| \int f(x + ty) \sigma_k(y) dy \right| \quad \text{where} \quad \sigma_k = \phi_{k+1} - \phi_k, \quad (3.1)$$

so that

$$\mathcal{M}f \leq \sum_{k=1}^{\infty} \mathcal{M}_k f.$$

In this section, we replace each \mathcal{M}_k by a linearized and discretized operator Φ_k whose $L^p \rightarrow L^q$ mapping properties we will subsequently investigate. These reduction steps are well known and have been extensively used in the literature. However, minor variations are needed to adapt to the specific situation at hand, and we include these for completeness.

The discretization will be in the space of affine transformations. Observe that if f is supported in the interval $[R, R + 1]$, then the support of $\mathcal{M}f$ is contained in $[R - 4, R + 1]$. Decomposing a function $f \in L^p(\mathbb{R})$ into functions f_i of disjoint support, and exploiting the essential disjointness of $\mathcal{M}f_i$, it is easy to deduce that (see also the proof of Corollary 7.3)

$$\mathcal{M} : L^p([0, 1]) \rightarrow L^q([-4, 1]) \quad \text{implies} \quad \mathcal{M} : L^p(\mathbb{R}) \rightarrow L^q(\mathbb{R}) \quad \text{for } p \leq q. \quad (3.2)$$

We will henceforth assume that all functions are supported on $[0, 1]$, so that the maximal functions are also supported within a fixed compact set. We decompose the space of translations x (namely the interval $[-4, 1]$) and the space of dilations r 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 $\mathcal{C} = \{c_i : 1 \leq i \leq \delta_{k+1}^{-L}\}$ and $\mathcal{R} = \{r_i : 1 \leq i \leq \delta_{k+1}^{-L}\}$ denote the discretized spaces of translation and dilation parameters.

Proposition 3.1. *Fix $1 < p < \infty$. Then there is a large integer L 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 \mathcal{C} and \mathcal{R} , respectively, such that*

$$\mathcal{M}_k f(x) \leq 4|\Phi_k f(x)| + \mathcal{E}_k f(x), \quad (3.3)$$

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 $[-4, 4]$.

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

$$\|\mathcal{E}_k f\|_q \leq C_{p,q} 2^{-k\eta} \|f\|_p. \quad (3.4)$$

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

$$\begin{aligned} \mathcal{M}_k f(x) &\leq 2 \left| \int f(x_i + \tilde{r}_i y) \sigma_k(y) dy \right| \\ &\leq 4 \left| \int f(z) \sigma_k \left(\frac{z - x_i}{\tilde{r}_i} \right) dz \right| \\ &\leq 4 \left| \int f(z) \sigma_k \left(\frac{z - c_i}{r_{j(i)}} \right) dz \right| + \mathcal{E}_k f(x), \end{aligned} \quad (3.5)$$

where

$$\mathcal{E}_k f(x) = 4 \left| \int f(z) \left[\sigma_k \left(\frac{z - x_i}{\tilde{r}_i} \right) - \sigma_k \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right| \quad (3.6)$$

and $r_{j(i)}$ is chosen so that $\tilde{r}_i \in R_{j(i)}$. Note that $|\tilde{r}_i - r_{j(i)}| \leq \delta_{k+1}^L$. Thus (3.3) holds with $c(x) = c_i$ and $r(x) = r_{j(i)}$.

It is obvious from (3.5) and (3.6) that $\Phi_k f$ and $\mathcal{E}_k f$ are supported in $[-4, 4]$. It remains to prove (3.4). For this we observe that

$$\begin{aligned} |\mathcal{E}_k f(x)| &\leq 4 \left| \int f(z) \left[\phi_{k+1} \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_{k+1} \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right| \\ &\quad + 4 \left| \int f(z) \left[\phi_k \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_k \left(\frac{z - c_i}{r_{j(i)}} \right) \right] dz \right|. \end{aligned} \quad (3.7)$$

By Hölder's inequality, the first term on the right side of (3.7) is bounded by

$$\|f\|_p \left\| \phi_{k+1} \left(\frac{z - x_i}{\tilde{r}_i} \right) - \phi_{k+1} \left(\frac{z - c_i}{r_{j(i)}} \right) \right\|_{p'}$$

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

By Lemma 3.2 below, each symmetric difference $(x_i + \tilde{r}_i I_m^{(k+1)}) \Delta (c_i + r_{j(i)} I_m^{(k+1)})$ has measure bounded by $3\delta_{k+1}^L$. Hence the last expression is bounded by

$$\begin{aligned}
\frac{2^{1/p}}{P_{k+1}\delta_{k+1}} \|f\|_p \left(P_{k+1}\delta_{k+1}^L \right)^{1/p'} &\leq \frac{2^{1/p}\delta_{k+1}^{(L-1)/p'}}{(P_{k+1}\delta_{k+1})^{1/p}} \|f\|_p \\
&\leq 2^{1/p}\delta_{k+1}^{\frac{L}{p'}-1} \|f\|_p \leq C2^{-(k+1)\eta} \|f\|_p,
\end{aligned}$$

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 (3.7) is bounded similarly, with P_{k+1}, δ_{k+1} replaced by P_k, δ_k . Finally, (3.4) follows from the pointwise bound above and the fact that \mathcal{E}_k are supported on the bounded interval $[-4, 4]$. \square

Lemma 3.2. *Let $0 < t < 1$, $\frac{1}{2} < r, s < 2$. Then for any $x, y \in \mathbb{R}$ we have*

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

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 $(y - x) + (x + rt - y - st) = (r - s)t \leq \eta$. \square

3.2 The interpolation argument

In view of Proposition 3.1, we need to obtain appropriate $L^p \rightarrow L^q$ bounds for Φ_k . Equivalently by duality, it suffices to obtain $L^{q'} \rightarrow L^{p'}$ bounds for

the “adjoint” operator Φ_k^* given by

$$\Phi_k^*g(z) = \int g(x)V_{k,x}(z) dx. \quad (3.8)$$

This turns out to be a consequence of a weaker bound, via an interpolation argument which we explain in the next lemma. Such arguments are ubiquitous in harmonic analysis, but our specific situation necessitates a careful application of these standard techniques, since we need to keep track of the size (in k) of the operator norm.

Lemma 3.3. *Let Φ_k^* be the operator defined in (3.8) and $q_0 \geq 2$. Suppose that Φ_k^* acting on characteristic functions of measurable sets, satisfies an $L^{\frac{q_0}{q_0-1}}[0, 1] \rightarrow L^{q_0}[-4, 1]$ estimate with an exponentially decaying bound: specifically, there is a constant $\eta_0 > 0$ such that*

$$\|\Phi_k^*\mathbf{1}_\Omega\|_{q_0} \leq 2^{-k\eta_0}|\Omega|^{\frac{q_0-1}{q_0}} \quad \text{for all sets } \Omega \subseteq [0, 1]. \quad (3.9)$$

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

Note that the hypothesis (3.9) is a restricted strong-type estimate.

Proof. The operator Φ_k^* satisfies a trivial $L^1 \rightarrow L^1$ 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 $L^p \rightarrow L^q$ 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 $k \rightarrow \infty$. On the other hand, by Hölder’s inequality,

$$\|\Phi_k^*\mathbf{1}_\Omega\|_q \leq \|\Phi_k^*\mathbf{1}_\Omega\|_{q_0}^\theta \|\Phi_k^*\mathbf{1}_\Omega\|_1^{1-\theta} \leq C2^{-k\eta_0\theta}|\Omega|^{\frac{1}{p}}.$$

The last line implies (by Theorem 5.3, Chapter 4, [4]) that the weak-type (p, q) norm of Φ_k^* continues to satisfy the same exponentially small bound. 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 norm decaying exponentially in k . The statement for Φ_k follows by duality. \square

Combining Lemma 3.3 with Proposition 3.1, we arrive at the following corollary.

Corollary 3.4. *Suppose that Φ_k^* satisfies the restricted strong-type estimate (3.9). Then for every $p > \frac{q_0}{q_0-1}$, there exists a positive constant $\eta(p)$ such that*

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

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

4 Transverse correlations

In this section, we translate the problem of deriving restricted strong-type $L^{\frac{n}{n-1}} \rightarrow L^n$ estimates for Φ_k^* to a question involving n -fold correlations between affine copies of S_k with few internal tangencies. We start by setting up the notation for such n -fold correlations and giving a precise statement of our correlation criterion. The construction of a sequence of sets S_k that will meet our criterion will be addressed in Section 5.

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 3.1:

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

In particular, we have $\#\mathfrak{A} \leq \delta_{k+1}^{-2Ln}$. We will also use \mathfrak{A}_{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:*

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

If $f_1 = \dots = f_n = f$, 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:*

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

Then the operator Φ_k^* defined in (3.8) satisfies the restricted strong-type estimate

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

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

Remarks:

1. While we have not yet specified any size estimate for $C_0(k, n, \epsilon_0)$, in view of Proposition 3.1 we will of course be interested in sets S_k for which $C_0(k, n, \epsilon_0)$ is summable in k , and in fact for which the right hand side of (4.3) has exponential decay. It would then follow from Corollary 3.4 that \mathcal{M} is bounded from $L^p(\mathbb{R}) \rightarrow L^{(n-1)p}(\mathbb{R})$ 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*

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

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

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

where $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_n)$. We treat each summand separately. Suppose first that $\lambda_{\ell_0} = 1$ for some ℓ_0 . Recalling that $\phi_{k+1} = (P_{k+1}\delta_{k+1})^{-1}\mathbf{1}_{S_{k+1}}$, and estimating all factors pointwise by $(P_{k+1}\delta_{k+1})^{-1}$, we see that in this case

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

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

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

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. \square

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 3.1. Thus $\mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}$. Let $\Omega \subseteq [0, 1]^n$, then

$$\begin{aligned} \|\Phi_k^* \mathbf{1}_\Omega\|_n^n &= \left\| \int_\Omega V_{k,x}(\cdot) dx \right\|_n^n \\ &= \int \prod_{j=1}^n \left[\int_\Omega V_{k,x_j}(z) dx_j \right] dz \\ &= \int_{\Omega^n} \left[\int \prod_{j=1}^n V_{k,x_j}(z) dz \right] dx_1 \cdots dx_n \end{aligned}$$

$$\begin{aligned}
&= \int_{\Omega^n} \Lambda(\mathbf{A}(x_1, \dots, x_n); \sigma_k) dx_1 \dots dx_n \\
&= \left[\int_{\Theta_1} + \int_{\Theta_2} \right] \Lambda(\mathbf{A}(x_1, \dots, x_n); \sigma_k) dx_1 \dots dx_n,
\end{aligned}$$

where

$$\begin{aligned}
\Theta_1 &= \{(x_1, \dots, x_n) \in \Omega^n : \mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}_{\text{int}}\}, \\
\Theta_2 &= \{(x_1, \dots, x_n) \in \Omega^n : \mathbf{A}(x_1, \dots, x_n) \in \mathfrak{A}_{\text{tr}}\}.
\end{aligned}$$

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, the discretization in Section 3.1 combined with Lemma 4.3 shows that

$$\begin{aligned}
\Theta_1 &\subseteq \left\{ (x_1, \dots, x_n) \in \Omega^n : \min_{\ell \neq \ell'} |c(x_\ell) - c(x_{\ell'})| \leq \frac{80n(n-1)}{P_k^{1-\epsilon_0}} \right\} \\
&\subseteq \bigcup_{1 \leq \ell \neq \ell' \leq n} \left\{ (x_1, \dots, x_n) \in \Omega^n : \min_{\ell \neq \ell'} |x_\ell - x_{\ell'}| \leq \frac{80n(n-1)}{P_k^{1-\epsilon_0}} + 2\delta_{k+1}^L \right\} \\
&\subseteq \bigcup_{1 \leq \ell \neq \ell' \leq n} \left\{ (x_1, \dots, x_n) \in \Omega^n : |x_\ell - x_{\ell'}| \leq \frac{160n(n-1)}{P_k^{1-\epsilon_0}} \right\}.
\end{aligned}$$

Combining this with Lemma 2.7 we obtain

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

We now turn to the integral on Θ_2 , where the desired gain will follow from the hypothesis (4.2). Namely, for $(x_1, \dots, x_n) \in \Theta_2$ we have

$$|\Lambda(\mathbf{A}(x_1, \dots, x_n); \sigma_k)| \leq C_0(k, n, \epsilon_0),$$

hence

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

Combining this with (4.8) yields the conclusion. □

5 The random construction

5.1 Selection of the sets $\{S_k\}$

We are now ready to describe the probabilistic mechanism and the selection procedure that gives rise to the sets $\{S_k\}$ satisfying the transverse correlation condition (4.2) with acceptable constants $C_0(k, n, \epsilon_0)$. The construction is as in Subsection 2.1, with the crucial additional point that the sequences $\boldsymbol{\kappa}_k, \boldsymbol{\tau}_k$ are now randomized. The parameters of this construction will be a sequence $\{N_k : k \geq 1\}$ of large constants and a sequence $\{\epsilon_k : k \geq 1\}$ of small constants.

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 :

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

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

$N_{k+1}^{-\epsilon_{k+1}}$, and set

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

At the $(k+1)$ -th stage, 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 stage. Thus at the $(k+1)$ -th step, P_{k+1} is a random variable, whereas Q_{k+1} is not.

Let us take this opportunity to clarify the exact probabilistic nature of the construction of the set $S = \bigcap_{k=1}^{\infty} S_k$. Unlike Salem's construction of Salem sets (see [29]), where S satisfying a certain criterion is shown to exist almost surely in a measure space consisting of a large family of Cantor-type constructions, in our case the randomness is limited to individual stages of the process. More precisely, 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 can make no claim as to its frequency of occurrence among all possible iterative constructions subject to the given parameters.

For every $k \geq 1$, we have a large sample space of possible choices for S_k . The goal of this section is to show that at every stage 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. *There exists a sequence of sets $\{S_k\}$ constructed as in Subsection 2.1 (i.e., there exists a nondecreasing sequence of large integers $\{N_k\}$, a nonincreasing sequence of small positive numbers $\{\epsilon_k\}$ and a realization of the Bernoulli sequences \mathbf{Y}_k) satisfying the following properties. There is an absolute constant $B > 0$ (independent of k and n) such that*

- (a) $2^{-k} \prod_{j=1}^k N_j^{1-\epsilon_j} \leq P_k(\mathbf{X}_k) \leq 2^k \prod_{j=1}^k N_j^{1-\epsilon_j}$.
- (b) $|P_k(\mathbf{X}_k) - Q_k| \leq B\sqrt{Q_k}$.

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

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

Corollary 5.2. *Let $\{S_k\}$ be a sequence of sets given by Theorem 5.1. Then the associated operator Φ_k^* defined in (3.8) satisfies the restricted weak-type estimate*

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

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

Proof. This 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). □

5.2 Two large deviation inequalities from probability theory

In this subsection, we record two large deviation inequalities widely used in probability theory that will play a key role in the sequel. We state them here without proof but with appropriate references. The first one is a version of Bernstein's inequality borrowed from [15]. We will use it here much as we did in [21].

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*

$$\mathbb{P}\left(\left|\sum_1^n Z_j\right| \geq m\lambda\right) \leq 4e^{-m^2\lambda^2/8\sigma^2}. \quad (5.4)$$

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, [35] or [1], p. 95). *Suppose that $\{U_k : k = 0, 1, 2, \dots\}$ 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)

The proof is by induction. We fix a large absolute constant B whose exact value will be specified in the sequel (in fact $B = 10$ will work), and choose N_1 sufficiently large relative to B . For $k \geq 1$, we show that the probability of any single item on the criteria list being false is $< B^{-1}$. The probability of the intersection of the desired events is therefore positive, allowing us to pick a realization of the Bernoulli sequence that satisfies all the properties simultaneously.

We start by verifying (a) and (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

$$\begin{aligned} \mathbb{P}(|P_1(\mathbf{X}_1) - N_1 p_1| > BN_1^{\frac{1-\epsilon_1}{2}}) &= \mathbb{P}\left(\left|\sum_{i=1}^{N_1} [X_1(i) - p_1]\right| > BN_1^{\frac{1-\epsilon_1}{2}}\right) \\ &\leq 4e^{-\frac{B^2}{8}}. \end{aligned}$$

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

$$\frac{1}{2}N_1^{1-\epsilon_1} \leq N_1^{1-\epsilon_1}(1 - BN_1^{-\frac{1-\epsilon_1}{2}}) \leq P_1 \leq N_1^{1-\epsilon_1}(1 + BN_1^{-\frac{1-\epsilon_1}{2}}) \leq 2N_1^{1-\epsilon_1}$$

holds. Let us assume then that (a) and (b) hold for some $k \geq 1$. Recalling that \mathbf{X}_k is a deterministic sequence at step $(k + 1)$, we find that for every $\mathbf{i} \in \mathbb{I}_k$ such that $X_k(\mathbf{i}) = 1$, the random variables

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

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

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

Thus with large probability, (b) holds with k replaced by $k + 1$. Further by induction hypotheses,

$$2^{-k} \prod_{j=1}^{k+1} N_j^{1-\epsilon_j} \leq Q_{k+1} = P_k N_{k+1}^{1-\epsilon_{k+1}} \leq 2^k \prod_{j=1}^{k+1} N_j^{1-\epsilon_j}, \quad (5.5)$$

which in particular means 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),

$$\frac{1}{2} \leq \left(1 - \frac{B}{\sqrt{Q_{k+1}}}\right) \leq \frac{P_{k+1}(\mathbf{X}_{k+1})}{Q_{k+1}} \leq Q_{k+1} \left(1 + \frac{B}{\sqrt{Q_{k+1}}}\right) \leq 2,$$

which coupled with (5.5) 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

The first step is to replace the measure $\sigma_k = \phi_{k+1} - \phi_k$ in (4.2) by $\bar{\sigma}_k$, where

$$\bar{\sigma}_k(z) = \frac{1}{Q_{k+1}\delta_{k+1}} \mathbf{1}_{S_{k+1}}(z) - \frac{1}{P_k\delta_k} \mathbf{1}_{S_k}(z). \quad (5.6)$$

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). *For any $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\} \in \mathfrak{A}$,*

$$|\Lambda(\mathbf{A}_n; \sigma_k)| \leq |\Lambda(\mathbf{A}_n; \bar{\sigma}_k)| + 2^{2n+1} B 2^{k(n+\frac{3}{2})} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2} + \epsilon_j(n-\frac{1}{2})} \right]. \quad (5.7)$$

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

$$C_0(k, n, \epsilon_0) \leq \sup_{\mathbf{A}_n \in \mathfrak{A}_{tr}} |\Lambda(\mathbf{A}_n; \bar{\sigma}_k)| + 2^{2n+1} B 2^{k+\frac{3}{2}} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2} + \epsilon_j(n-\frac{1}{2})} \right]. \quad (5.8)$$

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:*

$$\left| \sum_{\mathbf{I} \in \mathbb{F}_{tr}} \prod_{\ell=1}^n X_k(\mathbf{i}_\ell) \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^n \left(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1} \right) \cdot \left| \prod_{\ell=1}^n (c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)) \right| \right| \leq C_1(k, n, \epsilon_0) \quad (5.9)$$

where $\mathbf{I} = (\mathbf{i}_1, \dots, \mathbf{i}_n)$, and $\boldsymbol{\iota} = (i_{k+1,1}, \dots, i_{k+1,n})$ denotes the n -vector whose entries are the $(k+1)$ -th entries of $\bar{\mathbf{i}}_1, \dots, \bar{\mathbf{i}}_n$ respectively (thus $\boldsymbol{\iota}$ ranges over

the set $\{1, 2, \dots, N_{k+1}\}^n$. Then

$$\begin{aligned} \sup_{\mathbf{A}_n \in \mathfrak{A}_{tr}} |\Lambda(\mathbf{A}_n; \bar{\sigma}_k)| &\leq C_1(k, n, \epsilon_0) 2^{kn} \left[\prod_{j=1}^{k+1} N_j^{n\epsilon_j} \right] \\ &+ 2^{k(n+1-\epsilon_0)+3} \left[\prod_{j=1}^k N_j^{-\epsilon_0+\epsilon_j(n+\epsilon_0-1)} \right] N_{k+1}^{n\epsilon_{k+1}}. \end{aligned} \quad (5.10)$$

Proposition 5.7 (Step 3). Let $C_1(k, n, \epsilon_0)$ be the smallest constant satisfying (5.10) in Proposition 5.6. Then

$$C_1(k, n, \epsilon_0) \leq 4^n n! \left[\prod_{j=1}^k N_j^{-\frac{1+\epsilon_j}{2}} \right] \times \left[\ln(4^n n! B \prod_{j=1}^{k+1} N_j^{2Ln}) \right]^{\frac{1}{2}}. \quad (5.11)$$

Assume for now the claims in steps 1–3.

Conclusion of the proof of Theorem 5.1 (c). Combining (5.8), (5.10), and (5.11), we see that

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

We now plug in $\epsilon_0 = \frac{1}{2}$. After some simple algebra, we see that

$$\begin{aligned} |\Lambda(\mathbf{A}_n; \sigma_k)| &\leq 4^{n+2} n! B 2^{k(n+\frac{3}{2})} \\ &\times \left[\prod_{j=1}^k N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})} \right] N_{k+1}^{n\epsilon_{k+1}} \left[\ln(4^n n! B \prod_{j=1}^{k+1} N_j^{2Ln}) \right]^{\frac{1}{2}}, \end{aligned} \quad (5.12)$$

as required. \square

5.4.2 Proof of Lemma 5.5

It suffices to prove (5.7), since (5.8) follows directly from it and the definition of $C_0(k, n, \epsilon_0)$ in (4.2). We write $\sigma_k = \bar{\sigma}_k + e_k$, where $\bar{\sigma}_k$ is as in (5.6) so that

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

Then

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

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

We need to show that E_k is bounded by the quantity in (5.7).

We observe that by the definition of Q_k in (5.1) and part (a) of Theorem 5.1

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

whereas by part (b) of the same theorem

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

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

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

in which case

$$|\Lambda(\mathbf{A}_n; u_{\lambda_1}, \dots, u_{\lambda_n})|$$

$$\begin{aligned}
&\leq \left(2^{k+1} \prod_{j=1}^{k+1} N_j^{\epsilon_j}\right)^{n-1} \left(B2^{\frac{3k}{2}+1} \prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\frac{3\epsilon_j}{2}}\right) |(c_{\ell_0} + r_{\ell_0} S_{k+1})| \\
&\leq 2^n B2^{k(n+\frac{1}{2})} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n+\frac{1}{2})}\right] P_{k+1} \delta_{k+1} \\
&\leq 2^{n+1} B2^{k(n+\frac{3}{2})} \left[\prod_{j=1}^{k+1} N_j^{-\frac{1}{2}+\epsilon_j(n-\frac{1}{2})}\right].
\end{aligned}$$

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

5.4.3 Proof of Proposition 5.6

We need to estimate

$$\Lambda(\mathbf{A}_n; \bar{\sigma}_k) = \int \prod_{\ell=1}^n \bar{\sigma}_k\left(\frac{z - c_\ell}{r_\ell}\right) dz \quad (5.13)$$

for $\mathbf{A}_n = \{(c_\ell, r_\ell) : 1 \leq \ell \leq n\} \in \mathfrak{A}_{\text{tr}}$.

Recalling our notation $\bar{\mathbf{i}} = (\mathbf{i}, i_{k+1}) \in \mathbb{I}_{k+1}$, let us rewrite $\bar{\sigma}_k$ as

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

where the penultimate step follows from the definition of Q_{k+1} in (5.1). We will use this representation to express $\Lambda(\mathbf{A}_n; \bar{\sigma}_k)$ explicitly in terms of the random variables $\mathbf{Y}_{k+1}(\bar{\mathbf{i}})$. We have

$$\begin{aligned} \prod_{\ell=1}^n \bar{\sigma}_k \left(\frac{z - c_\ell}{r_\ell} \right) &= \frac{1}{(Q_{k+1} \delta_{k+1})^n} \sum_{\mathbf{I} \in \mathbb{I}_k^n} \left[\prod_{\ell=1}^n X_k(\mathbf{i}_\ell) \right. \\ &\quad \left. \times \sum_{\boldsymbol{\iota}} \left(\prod_{\ell=1}^n (Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1}) \right) \mathbf{1}_{\bigcap_{\ell=1}^n (c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell))}(z) \right] \end{aligned} \quad (5.14)$$

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 (5.14) 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}_{int} and \mathbb{F}_{tr} as in Subsection 2.4, we find that

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

We treat these two sums separately.

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

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

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 (5.9)

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

Combining the two estimates, we get (5.10). \square

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

$$\bigcap_{\ell=1}^n (c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)) \neq \emptyset. \quad (5.15)$$

Proof. Suppose that (5.15) holds, then

$$(\bar{\mathbf{i}}_1, \dots, \bar{\mathbf{i}}_n) \in \mathbb{F}[n, k+1, \mathbf{A}_n]. \quad (5.16)$$

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 vectors in (5.16) is at most 4, hence the number of possible choices of $(i_{k+1,2}, \dots, i_{k+1,n})$ is at most 4. This proves the claim, since there are at most N_{k+1} choices of $i_{k+1,1}$. \square

5.4.4 Proof of Proposition 5.7

The heart of the proof is a convenient re-indexing of the sum in (5.9) 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*

$$\alpha_k(\mathbf{i}_{\rho(1)}) < \dots < \alpha_k(\mathbf{i}_{\rho(n)}) \quad (5.17)$$

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

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

$$|\alpha_k(\mathbf{i}_\ell) - \alpha_k(\mathbf{i}_{\ell'})| > 4\delta_k. \quad (5.18)$$

Thus for every \mathbf{I} , all $\alpha_k(\mathbf{i}_\ell)$, $1 \leq \ell \leq n$, are distinct, and in particular there is a permutation $\rho = \rho(\mathbf{I})$ such that (5.17) 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}_ρ can be decomposed further into at most 4^{n-1} subsets on which all the projections π_ℓ are injective. \square

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

$$\left| \sum_{\mathbf{I} \in \mathcal{F}_\rho} \prod_{\ell=1}^n X_k(\mathbf{i}_\ell) \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^n \left(Y_{k+1}(\bar{\mathbf{i}}_\ell) - p_{k+1} \right) \cdot \left| \bigcap_{\ell=1}^n (c_\ell + r_\ell I_{k+1}(\bar{\mathbf{i}}_\ell)) \right| \right| \quad (5.19)$$

for each such \mathcal{F}_ρ .

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 (5.19) 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 $\{\mathbf{I}(j) = (\mathbf{i}_1(j), \dots, \mathbf{i}_n(j)) : j = 1, \dots, T\}$ so that

$$\alpha_k(\mathbf{i}_{\rho(n)}(1)) < \dots < \alpha_k(\mathbf{i}_{\rho(n)}(T)). \quad (5.20)$$

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

$$W_j = \sum_{\boldsymbol{\iota}} \prod_{\ell=1}^n \left(Y_{k+1}(\mathbf{i}_\ell(j), i_{k+1,\ell}) - p_{k+1} \right) \left| \bigcap_{\ell=1}^n (c_\ell + r_\ell I_{k+1}(\mathbf{i}_\ell(j), i_{k+1,\ell})) \right|, \quad (5.21)$$

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 $W_0 = 0$.

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

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 $\mathcal{W}' = \mathcal{W} \setminus \{W_1, \dots, W_{m-1}\}$, so that

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

Here $\{F_{\iota, m}\}$ are measurable functions of \mathcal{W} specified by the expression (5.21) for W_m but whose exact functional forms are unimportant.

By (5.17), we have

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

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

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

But this follows from

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

where we used (5.17) again and then (5.20).

It remains to prove the almost sure bound on W_j . Indeed, by Lemma 5.8 the number of summands in (5.21) 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 $|W_j| \leq 4N_{k+1}\delta_{k+1} = \delta_k$, as claimed. \square

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 $U_j = W_1 + \dots + W_j$, 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

$$\begin{aligned} \mathbb{P}\left((5.19) > \lambda\right) &\leq 2 \exp\left(-\frac{\lambda^2}{32\delta_k^2 T}\right) \\ &\leq 2 \exp\left(-\frac{\lambda^2}{32\delta_k^2 P_k}\right) \leq \frac{\delta_{k+1}^{2Ln}}{4^{n-1} n! B}. \end{aligned} \tag{5.22}$$

Since there are at most $4^{n-1} n!$ classes \mathcal{F}_ρ , the probability that (5.22) holds for at least of them is bounded by $B^{-1} \delta_{k+1}^{2Ln}$. Summing over such classes, we see that

$$\mathbb{P}\left(\text{LHS of (5.9)} > 4^{n-1} n! \lambda\right) \leq \frac{\delta_{k+1}^{2Ln}}{B}.$$

Finally, since $\#\mathfrak{A} = \delta_{k+1}^{-2Ln}$, there is a probability of at least $1 - \frac{1}{B}$ that (5.9) holds for every $\mathbf{A} \in \mathfrak{A}_{\text{tr}}$ with

$$\begin{aligned} C_1(k, n, \epsilon_0) &= 4^{n-1} n! \lambda \\ &= 4^n n! \delta_k \sqrt{2P_k} \sqrt{\ln(4^n n! B \delta_{k+1}^{-2Ln})} \\ &\leq 4^n n! 2^{\frac{k+1}{2}} \left[\prod_{j=1}^k N_j^{-\frac{1+\epsilon_j}{2}} \right] \times \left[\ln(4^n n! B \prod_{j=1}^{k+1} N_j^{2Ln}) \right]^{\frac{1}{2}}. \end{aligned}$$

This completes the proof of the proposition. \square

5.5 Existence of a limiting measure

We now turn to the question of the existence of a weak-* limit μ of the normalized Lebesgue densities $\phi_k = \mathbf{1}_{S_k} / |S_k|$, where the S_k -s have been constructed according to the random mechanism described earlier in this section. It turns out that it is relatively simple to choose a realization of the Bernoulli sequence \mathbf{Y}_k at every stage so that the condition (2.7) in Lemma 2.2 can be verified. We now proceed to do so. The two main results in this subsection are the following.

Proposition 5.11. *Assume that the sequence $\ln(P_k)N_{k+1}^{\epsilon_{k+1}-1}$ is bounded by a small constant $\eta \ll 1$, say*

$$\sup_{k \geq 1} \frac{\ln(P_k)}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \eta < \frac{1}{16}. \quad (5.23)$$

Then it is possible to choose a sequence of sets $\{S_k\}$ and an absolute constant B which satisfy, in addition to the properties listed in Theorem 5.1, the two conditions below:

(a) *For every $k \geq 1$,*

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

(b) *For all $k \geq 1$, $m \geq 0$ and every $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$,*

$$2^{-m} \left[\prod_{r=1}^m N_{k+r} p_{k+r} \right] P_k \leq \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \leq 2^m \left[\prod_{r=1}^m N_{k+r} p_{k+r} \right] P_k, \quad (5.24)$$

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

Proposition 5.12. *If the sets $\{S_k\}$ satisfy the two properties stated in Proposition 5.11, then*

$$\sup_{k': k' \geq k} \sum_{\mathbf{i}: X_k(\mathbf{i})=1} \left| \int_{I_k(\mathbf{i})} (\phi_{k'} - \phi_k) dx \right| \leq 2B \sum_{m=0}^{\infty} 2^{\frac{5}{2}(k+m+1)} \frac{[\ln(BP_{k+m})]^{\frac{1}{2}}}{N_{k+m+1}^{\frac{1-\epsilon_{k+m+1}}{2}}}. \quad (5.25)$$

In particular, if the quantity on the right hand side above tends to zero as $k \rightarrow \infty$, then the hypothesis (2.7) of Lemma 2.2, and hence the conclusion hold.

5.5.1 Proof of Proposition 5.11

Proof. We first discuss the proof of part (a). Since $X_{k+1}(\bar{\mathbf{i}}) = X_k(\mathbf{i})Y_{k+1}(\bar{\mathbf{i}}) = Y_{k+1}(\bar{\mathbf{i}})$ for any $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$, we obtain the identity

$$\sum_{i_{k+1}=1}^{N_{k+1}} [X_{k+1}(\bar{\mathbf{i}}) - p_{k+1}] = \sum_{i_{k+1}=1}^{N_{k+1}} [Y_{k+1}(\bar{\mathbf{i}}) - p_{k+1}].$$

The summands on the right hand side are independent random variables, so by Bernstein's inequality with $\sigma^2 = N_{k+1}p_{k+1}$ and $\lambda = (8p_{k+1} \ln(4BP_k)/N_{k+1})^{\frac{1}{2}}$

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

Taking the supremum of the sum above over the P_k -many choices of indices $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$, we find that the statement in part (a) holds with probability $1 - \frac{1}{B}$.

The proof of part (b) follows from (a) and induction on m . For $m = 0$ the inequalities in (5.24) can be replaced by equalities, and reduce to the definition of P_k . Assuming that (5.24) holds for m and setting $\bar{\mathbf{j}} = (\mathbf{j}, j_{m+1})$, we arrive at the following estimate

$$\begin{aligned} & \left| \sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}}) - \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) N_{k+m+1} p_{k+m+1} \right| \\ &= \left| \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \sum_{j_{m+1}} [Y_{k+m+1}(\mathbf{i}, \mathbf{j}, j_{m+1}) - p_{k+m+1}] \right| \\ &\leq \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \left[N_{k+m+1} p_{k+m+1} \ln(4BP_{k+m}) \right]^{\frac{1}{2}}, \end{aligned}$$

so that

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

By (5.23) and choosing N_1 sufficiently large relative to B we can ensure that $\frac{\ln(4BP_k)}{N_{k+1} p_{k+1}} \leq \frac{1}{4}$. Thus

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

which combined with the induction hypothesis yields the desired result. \square

5.5.2 Proof of Proposition 5.12

Proof. Since

$$\sup_{k': k' \geq k} \sum_{\mathbf{i}: X_k(\mathbf{i})=1} \left| \int_{I_k(\mathbf{i})} (\phi_{k'} - \phi_k)(x) dx \right|$$

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

it suffices to prove that the quantity in the last line is bounded above by the right hand side of (5.25). To this end, we fix an $m \geq 0$ and $\mathbf{i} \in \mathbb{I}_k$ with $X_k(\mathbf{i}) = 1$ and write

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

$$\begin{aligned}
\Xi_1 &:= \frac{P_k}{\delta_{k+m+1}} \left[\frac{1}{P_{k+m+1}} - \frac{1}{Q_{k+m+1}} \right] \int_{I_k(\mathbf{i})} \sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}}) \mathbf{1}_{I_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}})}(x) dx \\
&= P_k \left[\frac{1}{P_{k+m+1}} - \frac{1}{Q_{k+m+1}} \right] \sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}}), \text{ and} \\
\Xi_2 &:= \frac{P_k}{Q_{k+m+1}\delta_{k+m+1}} \int_{I_k(\mathbf{i})} \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \sum_{j_{m+1}} (Y_{k+m+1}(\mathbf{i}, \mathbf{j}, j_{m+1}) - p_{k+m+1}) \\
&\quad \times \mathbf{1}_{I_{k+m+1}(\mathbf{i}, \mathbf{j}, j_{m+1})}(x) dx \\
&= \frac{P_k}{Q_{k+m+1}} \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \sum_{j_{m+1}} (Y_{k+m+1}(\mathbf{i}, \mathbf{j}, j_{m+1}) - p_{k+m+1}).
\end{aligned}$$

In order to estimate Ξ_1 , we invoke part (a) of Theorem 5.1 and part (b) of Proposition 5.11 to conclude that

$$\begin{aligned}
|\Xi_1| &= P_k \left| \frac{Q_{k+m+1} - P_{k+m+1}}{P_{k+m+1} Q_{k+m+1}} \right| \sum_{\bar{\mathbf{j}}} X_{k+m+1}(\mathbf{i}, \bar{\mathbf{j}}) \\
&\leq \frac{BP_k}{P_{k+m+1} \sqrt{Q_{k+m+1}}} 2^{m+1} \left[\prod_{j=1}^{m+1} N_{k+j} p_{k+j} \right] \\
&\leq B 2^{\frac{5}{2}(k+m+1)} \left[\prod_{j=1}^{k+m+1} N_j^{1-\epsilon_j} \right]^{-\frac{1}{2}}.
\end{aligned} \tag{5.26}$$

On the other hand using both parts of Proposition 5.11,

$$\begin{aligned}
|\Xi_2| &= \frac{P_k}{Q_{k+m+1}} \left| \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \sum_{j_{m+1}} (Y_{k+m+1}(\mathbf{i}, \mathbf{j}, j_{m+1}) - p_{k+m+1}) \right| \\
&\leq \frac{P_k}{Q_{k+m+1}} \sum_{\mathbf{j}} X_{k+m}(\mathbf{i}, \mathbf{j}) \left[8N_{k+m+1} p_{k+m+1} \ln(BP_{k+m}) \right]^{\frac{1}{2}} \\
&\leq \frac{P_k}{Q_{k+m+1}} 2^m \left[\prod_{j=1}^m N_{k+j} p_{k+j} \right] \times \left[8N_{k+m+1} p_{k+m+1} \ln(BP_{k+m}) \right]^{\frac{1}{2}} \\
&\leq 2^{2(k+m)} \sqrt{8} \frac{[\ln(BP_{k+m})]^{\frac{1}{2}}}{[N_{k+m+1} p_{k+m+1}]^{\frac{1}{2}}} = 2^{2(k+m)} \sqrt{8} \frac{[\ln(BP_{k+m})]^{\frac{1}{2}}}{N_{k+m+1}^{\frac{1-\epsilon_{k+m+1}}{2}}}.
\end{aligned} \tag{5.27}$$

Combining the estimates in (5.26) and (5.27) we obtain

$$\begin{aligned}
|\Xi_1| + |\Xi_2| &\leq B 2^{\frac{5}{2}(k+m+1)} \left[\prod_{j=1}^{k+m+1} N_j^{1-\epsilon_j} \right]^{-\frac{1}{2}} + 2^{2(k+m)} \sqrt{8} \frac{[\ln(BP_{k+m})]^{\frac{1}{2}}}{N_{k+m+1}^{\frac{1-\epsilon_{k+m+1}}{2}}} \\
&\leq 2B 2^{\frac{5}{2}(k+m+1)} \frac{[\ln(BP_{k+m})]^{\frac{1}{2}}}{N_{k+m+1}^{\frac{1-\epsilon_{k+m+1}}{2}}}.
\end{aligned}$$

□

6 The restricted maximal estimates

In this section we prove those parts of Theorems 1.1 and 1.3 that concern the restricted maximal operators with $1 < t < 2$. We will do this by fixing

the parameters N_j, ϵ_j 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

$$N_j = N^j, \epsilon_j = \frac{1}{j}, \quad (6.1)$$

and the conclusions of Theorem 1.3 will hold for S_k with

$$N_j = N^j, \epsilon_j = \epsilon, \quad (6.2)$$

where N is a large integer.

Lemma 6.1. *Let N_j, ϵ_j be as above. 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), (3.9) holds for all $q_0 \geq 2$,*
- (c) *assuming (6.2), (3.9) holds for all $2 \leq q_0 < q_\epsilon$, where $q_\epsilon = \frac{\epsilon+1}{2\epsilon}$ as in Theorem 1.3,*
- (d) *assuming either (6.1) or (6.2), (5.23) holds.*

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, (3.9) holds with q_0 as above. It follows by Corollary 3.4 that

$$\|\mathcal{M}_k f\|_{(q_0-1)p} \leq C 2^{-k\eta(p)} \|f\|_p, \quad p > \frac{q_0}{q_0-1}, \quad (6.3)$$

for the same q_0 .

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

$$\|\mathcal{M}_k f\|_q \leq C 2^{-k\eta(p)} \|f\|_p \quad (6.4)$$

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 $[-4, 1]$, 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 $L^q[-4, 1]$ for any $p, q \in (1, \infty)$. By (3.2), 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

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

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 $q_0 \rightarrow q_\epsilon$, 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 $L^q[-4, 1]$ for p, q as in (6.5). By (3.2), \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), (5.23) holds, hence the existence of μ follows from Proposition 5.11.

6.1 The 1-dimensional case

Let N_j, ϵ_j be as in (6.1). Then $M_k = N^{\frac{k(k+1)}{2}}$ 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),

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

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 simple but somewhat cumbersome algebra that

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

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.

It remains to prove (d), i.e. the estimate (5.23). We have

$$\frac{\ln P_k}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \frac{k \ln 2 + \frac{k(k-1)}{2} \ln N}{N^k} < \frac{1}{20}$$

uniformly in k , provided that N has been chosen large enough. This proves the claim.

6.2 The lower-dimensional case

Let N_j, ϵ_j be as in (6.2). Then $M_k = N^{\frac{k(k+1)}{2}}$ as before, 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),

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

whereas by Lemma 2.1(b),

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

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

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

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.

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

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 $q_0 < q_\epsilon$.

Finally, we prove (5.23). We have

$$\frac{\ln P_k}{N_{k+1}^{1-\epsilon_{k+1}}} \leq \frac{k \ln 2 + \frac{k(k+1)}{2}(1-\epsilon) \ln N}{N^{(k+1)(1-\epsilon)}} \leq \frac{1}{20}$$

for all k , if N was chosen large enough.

7 Extension to the unrestricted operator

Recall that

$$\phi_k(x) = \frac{1}{P_k \delta_k} \sum_{\substack{\mathbf{i} \\ \kappa_{\mathbf{i}}(k)=1}} \mathbf{1}_{I_{\mathbf{i}}(k)}.$$

and denote by $A_t[k]$ the averaging operator associated to ϕ_k :

$$A_t[k]f(x) = \int f(x+ty)\phi_k(y) dy. \quad (7.1)$$

Recall also that

$$\mathcal{M}_k f(x) = \sup_{1 < r < 2} \left| \int f(x + ry) \sigma_k(y) dy \right|, \quad (7.2)$$

where $\sigma_k(x) = \phi_{k+1}(x) - \phi_k(x)$. The purpose of this subsection is to justify the following statement.

Proposition 7.1. *Let $1 < p \leq q < \infty$, $q \geq 2$. Let*

$$\tilde{\mathcal{M}}^a f = \sup_{\substack{n \geq 1 \\ r > 0}} r^a \int |f(x + ry)| \phi_n(y) dy,$$

where $a = \frac{1}{p} - \frac{1}{q}$. Assume that for all f supported on $[0, 1]$ we have

$$\|\mathcal{M}_k f\|_q \leq C 2^{-\eta_0 k} \|f\|_p \quad (7.3)$$

for some $C > 0$ and $\eta_0 > 0$. Then $\tilde{\mathcal{M}}^a$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$.

Assuming the proposition, 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 (7.3) holds for all $p, q \in (1, \infty)$. By Proposition 7.1, $\tilde{\mathcal{M}}^a$ is bounded from $L^p(\mathbb{R})$ to $L^q(\mathbb{R})$ whenever $1 < p \leq q < \infty$ and $q \geq 2$, as claimed in Theorem 1.1. 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), $p \leq q$, and $q \geq 2$.

Proposition 7.1 will be proved in a sequence of lemmas. Following [8], we denote by \mathcal{D}_k the σ -algebra generated by dyadic intervals of length 2^{-k} , and by \mathbb{E}_k the corresponding conditional expectation operators, i.e., $\mathbb{E}_k(f) = \mathbb{E}(f | \mathcal{D}_k)$.

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

$$\left\| \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](f)| \right\|_q \leq C 2^{-\eta_s} \|f\|_p \quad (7.4)$$

for some $\eta > 0$ and all functions f supported on $[0, 1]$ satisfying $\mathbb{E}_s(f) = 0$.

Proof. Let f be as in the lemma. We have

$$|A_t[n](f)| \leq \sum_{k=n}^{\infty} |\mathcal{M}_k f(x)|,$$

hence

$$\left\| \sup_{\substack{n \geq 1 \\ 1 \leq t \leq 2}} |(A_t[n](f))| \right\|_q \leq \sum_{k=1}^{\infty} \|\mathcal{M}_k f\|_q. \quad (7.5)$$

The right side is clearly summable by (7.3). To obtain decay as required in (7.4), we will use the assumption that $\mathbb{E}_s f = 0$ to improve the estimate on the terms with $k \leq k_0$, where k_0 will be determined shortly. We have

$$\begin{aligned} \mathcal{M}_k f(x) &= \left| \int f(x+ty)(\phi_{k+1}(y) - \phi_k(y))dy \right| \\ &\leq \left| \int f(x+ty)\phi_k(y)dy \right| + \left| \int f(x+ty)\phi_{k+1}(y)dy \right|. \end{aligned}$$

Suppose that $2^{-s} < \delta_{k+1}$. Each of the δ_k -intervals in the support of ϕ_k can be written as a union of some number of dyadic 2^{-s} -intervals and at most 2 additional intervals at both ends, each of length less than 2^{-s} . Using also Hölder's inequality, we see that

$$\begin{aligned} \mathcal{M}_k f(x) &\approx \frac{1}{P_k \delta_k} \int_{O(P_k) \text{ intervals of length } < 2^{-s}} |f(ty)|dy \\ &\leq \frac{1}{P_k \delta_k} \|f\|_p (P_k \cdot 2^{-s})^{1/p'} \\ &= \frac{1}{P_k^{1/p} \delta_k} 2^{-s/p'} \|f\|_p. \end{aligned}$$

The term with ϕ_{k+1} is estimated similarly. Taking the L^q norm of the left side and using that f is supported on a fixed bounded interval, we see that

$$\|\mathcal{M}_k f\|_q \leq \frac{C}{P_{k+1}^{1/p} \delta_{k+1}} 2^{-s/p'} \|f\|_p.$$

Let $k_0 \approx cs$ with a small enough constant. Then for $k \leq k_0$ we have

$$\frac{1}{P_{k+1}^{1/p} \delta_{k+1}} 2^{-s/2p'} < 1,$$

so that

$$\|\mathcal{M}_k f\|_q \leq C 2^{-s/2p'} \|f\|_p .$$

We now use this along with (7.3) to estimate the right side of (7.5):

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

as claimed in (7.4) □

Corollary 7.3. *Suppose that the assumptions of Lemma 7.2 hold, and let $R \geq 1$. Then*

$$\left\| \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](f)| \right\|_q \leq C 2^{-\eta s} R^{\max(0, \frac{1}{q} - \frac{1}{p})} \|f\|_p \quad (7.6)$$

for some $\eta > 0$ and all functions f supported on $[0, R]$ such that $\mathbb{E}_s(f) = 0$.

Note in particular that if $p \leq q$, then the bound in (7.6) is independent of R , hence it also holds for all $f \in L^p(\mathbb{R})$ without the compact support assumption.

Proof. We may assume that R is integer. Write $f = \sum_{i=1}^R f^{(i)}$, where $f^{(i)}$ is supported on $[i-1, i]$. Then

$$\begin{aligned} \left\| \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](f)| \right\|_q^q &= \left\| \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](\sum_i f^{(i)})| \right\|_q^q \\ &\leq \left\| \sum_i \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](f^{(i)})| \right\|_q^q \\ &\leq C \sum_i \left\| \sup_{\substack{n \geq 1 \\ 1 < t < 2}} |A_t[n](f^{(i)})| \right\|_q^q \\ &\leq C \sum_{i=1}^R 2^{-\eta s q} \|f^{(i)}\|_p^q . \end{aligned} \quad (7.7)$$

We used that $A_t[n](f^{(i)})$ have finitely overlapping support, then applied Lemma 7.2 to each $f^{(i)}$.

If $p \leq q$, we estimate the last sum in (7.7) by

$$2^{-\eta sq} \sum_{i=1}^R (\|f^{(i)}\|_p)^{q/p} \leq 2^{-\eta sq} \left(\sum_{i=1}^R \|f^{(i)}\|_p \right)^{q/p} \leq 2^{-\eta sq} \|f\|_p^q.$$

If on the other hand $p > q$, we use Hölder's inequality to bound the last sum in (7.7) by

$$2^{-\eta sq} \left(\sum_{i=1}^R \|f^{(i)}\|_p \right)^{q/p} R^{1-\frac{q}{p}} = 2^{-\eta sq} \|f\|_p^q R^{1-\frac{q}{p}}.$$

□

Proof of Proposition 7.1. Given $m \in \mathbb{Z}$ such that $2^{-m} \leq t \leq 2^{-m+1}$, we write

$$f = \mathbb{E}_m(f) + \sum_{k \geq m} \Delta_k(f), \quad \text{where } \Delta_k(f) = \mathbb{E}_{k+1}(f) - \mathbb{E}_k(f).$$

Therefore

$$A_t[n](f) = A_t[n](\mathbb{E}_m(f)) + \sum_{k \geq m} A_t[n](\Delta_k(f)). \quad (7.8)$$

We claim that

$$\sup_{\substack{n \geq 1 \\ t > 0}} t^a |A_t[n]\mathbb{E}_m(f)(x)| \leq C f^*(x), \quad (7.9)$$

where

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

Indeed, since $S_k \subseteq [1, 2]$ and $t \leq 2^{-m+1}$, the set $x + tS_k$ 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 $A_t[n]\mathbb{E}_m f(x)$ is a convex linear combination of the λ_i -s, it suffices to prove that $t^a \lambda_i \leq f^*(x)$. But this follows from

$$t^a \lambda_i = \frac{t^a}{|J_i|} \int_{J_i} |f(y)| dy \leq \frac{10t^a}{|J|} \int_{J'} |f(y)| dy \leq \frac{C}{|J'|^{1-a}} \int_{J'} |f(y)| dy \leq f^*(x),$$

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 $1 < p < q < \infty$, then $0 < a < 1$ and

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

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 $f \rightarrow f^*$ is bounded from L^p to L^q with $1 + \frac{1}{q} = \frac{1}{p} + (1-a)$, as claimed.

Turning our attention to the sum in the second term of (7.8), it suffices to prove that

$$\left\| \sup_{m \in \mathbb{Z}} \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} 2^{-ma} \left| \sum_{k \geq m} A_t[n](\Delta_k(f)) \right| \right\|_q \leq C \|f\|_p. \quad (7.10)$$

We write

$$\begin{aligned} & \sup_{m \in \mathbb{Z}} \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} 2^{-ma} \left| \sum_{k \geq m} A_t[n](\Delta_k(f)) \right| \\ & \leq \left[\sum_{m \in \mathbb{Z}} 2^{-maq} \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} \left| \sum_{k \geq m} A_t[n](\Delta_k(f)) \right|^q \right]^{1/q} \\ & \leq \left[\sum_{m \in \mathbb{Z}} 2^{-maq} \left(\sum_{k \geq m} \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} |A_t[n](\Delta_k(f))| \right)^q \right]^{1/q} \end{aligned}$$

Taking the L^q -norms of both sides, then using Lemma 7.4, we see that the

left side of (7.10) is bounded by

$$\begin{aligned}
& \left(\sum_{m \in \mathbb{Z}} 2^{-maq} \left\| \sum_{k \geq m} \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} |A_t[n](\Delta_k(f))| \right\|_q^q \right)^{1/q} \\
& \leq \left(\sum_{m \in \mathbb{Z}} 2^{-maq} \left[\sum_{k \geq m} \left\| \sup_{\substack{n \geq 1 \\ 1 \leq t 2^m \leq 2}} |A_t[n](\Delta_k(f))| \right\|_q^q \right] \right)^{1/q} \\
& \leq C \left(\sum_{m \in \mathbb{Z}} \left[\sum_{m \geq k} 2^{-\eta(k-m)} \|\Delta_k(f)\|_p \right]^q \right)^{1/q}.
\end{aligned}$$

The last line is the ℓ^q -norm of the convolution of the discrete functions $\mathbf{1}_{m \geq 0} 2^{-\eta m}$ 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

$$\left(\sum_m 2^{-\eta m r} \right)^{1/r} \left(\sum_m \|\Delta_m(f)\|_p^s \right)^{1/s} \leq C \left(\sum_m \|\Delta_m(f)\|_p^s \right)^{1/s}. \quad (7.11)$$

It remains to show that

$$\left(\sum_m \|\Delta_m(f)\|_p^s \right)^{1/s} \leq C \|f\|_p. \quad (7.12)$$

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 (7.12) for all $p \in [2, \infty)$.

Assume now that $1 < p < 2$, so that $s = 2$. Then

$$\begin{aligned}
\left(\sum_m \|\Delta_m(f)\|_p^2 \right)^{1/2} &= \left(\sum_m \left(\int (|\Delta_m(f)|^2)^{p/2} \right)^{2/p} \right)^{1/2} \\
&\leq \left(\sum_m \int (|\Delta_m(f)|^2)^{p/2} \right)^{1/p} \\
&\leq \left(\int \left(\sum_m |\Delta_m(f)|^2 \right)^{p/2} \right)^{1/p}.
\end{aligned}$$

The claim (7.12) now follows from the Littlewood-Paley inequality. \square

Lemma 7.4. *Suppose that (7.4) holds for all functions f satisfying $\mathbb{E}_s(f) = 0$, and define $\Delta_k f$ as above. Then*

$$\left\| \sup_{\substack{n \geq 1 \\ 1 \leq t2^m \leq 2}} |A_t[n](\Delta_{s+m}f)| \right\|_q \leq C \cdot 2^{ma-\eta s} \|\Delta_{s+m}f\|_p. \quad (7.13)$$

Proof. Let $u = t2^m$, so that $1 \leq u \leq 2$. We have

$$\begin{aligned} A_t[n]f(x) &= \int f(x+ty)\phi_n(y)dy \\ &= \int f(x+2^{-m}uy)\phi_n(y)dy \\ &= \int f(2^{-m}(2^m x + uy))\phi_n(y)dy \\ &= A_u[n](f_m)(2^m x), \end{aligned} \quad (7.14)$$

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 (7.4), we have

$$\begin{aligned} \left\| \sup_{\substack{n \geq 1 \\ 1 \leq t2^m \leq 2}} |A_t[n](\Delta_{s+m}(f))| \right\|_q &= \left\| \sup_{\substack{n \geq 1 \\ 1 \leq u \leq 2}} |(A_u[n](\Delta_{s+m}(f))_m)(2^m\cdot)| \right\|_q \\ &= 2^{-m/q} \left\| \sup_{\substack{n \geq 1 \\ 1 \leq u \leq 2}} |A_u[n](\Delta_{s+m}(f))_m| \right\|_q \\ &\leq C 2^{-m/q} 2^{-\eta s} \|(\Delta_{s+m}(f))_m\|_p \\ &= C 2^{-m/q} 2^{-\eta s} 2^{m/p} \|\Delta_{s+m}(f)\|_p \\ &= C 2^{ma} 2^{-\eta s} \|\Delta_{s+m}(f)\|_p. \end{aligned}$$

\square

8 Proof of Theorem 1.6

Let $f \in L^2[0, 1]$. We need to prove that

$$\limsup_{t \rightarrow 0} \sup_{k \geq 1} |A_t[k]f(x) - f(x)| = 0$$

for almost all x , where the averages $A_t[k](x)$ are defined as in (7.1). It suffices to show that

$$\limsup_{t \rightarrow 0} \sup_{k \geq 1} \|A_t[k]f - f\|_2 = 0. \quad (8.1)$$

Given $\epsilon > 0$, we may choose a continuous function f_ϵ on $[0, 1]$ such that $\|f - f_\epsilon\|_2 \leq \epsilon$. Furthermore, since f_ϵ is uniformly continuous on $[0, 1]$, there is a function $\eta(t)$ with $\eta(t) \rightarrow 0$ as $t \rightarrow 0$ such that

$$|f_\epsilon(y) - f_\epsilon(y')| \leq \eta(t) \text{ for all } y, y' \text{ such that } |y - y'| \leq \epsilon. \quad (8.2)$$

We have

$$\begin{aligned} \|A_t[k]f - f\|_2 &\leq \|A_t[k](f - f_\epsilon)\|_2 + \|A_t[k]f_\epsilon - f_\epsilon\|_2 + \|f_\epsilon - f\|_2 \\ &\leq \|\tilde{\mathcal{M}}(f - f_\epsilon)\|_2 + \eta(t) + \epsilon \\ &\leq \epsilon \|\tilde{\mathcal{M}}\|_{L^2 \rightarrow L^2} + \eta(t) + \epsilon. \end{aligned}$$

By Theorem 1.1, $\|\tilde{\mathcal{M}}\|_{L^2 \rightarrow L^2}$ is finite. Thus if t is small enough so that $\eta(t) < \epsilon$, we have

$$\|A_t[k]f - f\|_2 \leq (\|\tilde{\mathcal{M}}\|_{L^2 \rightarrow L^2} + 2)\epsilon.$$

Since $\epsilon > 0$ was arbitrary, it follows that the limit in (8.1) is 0, as claimed.

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, A. Nagel, E.M. Stein, S. Wainger, *Singular and maximal Radon transforms: Analysis and geometry*, Ann. Math. **150** (1999), 489–577.
- [10] M. Csörnyei, *Density theorems revisited*, Acta Sci. Math. (Szeged), **64** (1998), pp. 59–65.
- [11] M. de Guzman, *Differentiation of integrals in \mathbb{R}^n* , Lecture Notes in Math., no. 481, Springer-Verlag 1975.
- [12] K. Falconer, *The geometry of fractal sets*, Cambridge Univ. Press 1985.
- [13] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, John Wiley & Sons Ltd., 1995.
- [14] W.T. Gowers, *A new proof of Szemerédi’s theorem*, Geom. Funct. Anal. **11** (2001), 465–588.
- [15] B. Green, *Arithmetic progressions in sumsets*, Geom. Funct. Anal. **12** (2002), 584–597.
- [16] B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), 481–547.

- [17] G. Hardy, J. Littlewood, *A maximal theorem with function-theoretic applications*, Acta Math. **54** (1930), 81–116.
- [18] C. A. Hayes and C. Y. Pauc, *Derivation and martingales*, Springer-Verlag 1970.
- [19] J.P. Kahane, *Some Random Series of Functions*, Cambridge Univ. Press, 1985.
- [20] L. Kaufman, *On the theorem of Jarnik and Besicovitch*, Acta Arith. **39** (1981), 265–267.
- [21] I. Laba and M. Pramanik, *Arithmetic progressions in sets of fractional Hausdorff dimension*, to appear in Geom. Funct. Anal. (2009).
- [22] J.M. Marstrand, *Packing circles in the plane*, Proc. London Math. Soc. **55** (1987), 37–58.
- [23] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge Studies in Advanced Mathematics, vol. 44, Cambridge University Press, 1995.
- [24] G. Mockenhaupt, *Salem sets and restriction properties of Fourier transforms*, Geom. Funct. Anal. **10** (2000), 1579–1587.
- [25] D.H. Phong, E.M. Stein, *Hilbert integrals, singular integrals, and Radon transforms I*. Acta Math. **157** (1986), 99–157.
- [26] D.H. Phong, E.M. Stein, *Hilbert integrals, singular integrals, and Radon transforms —I*. Invent. Math. **86** (1986), 75–113.
- [27] D. Preiss, lectures given in Ravello (1985), not published.
- [28] J.L. Rubio de Francia, *Maximal functions and Fourier transforms*, Duke Math. J. **53** (1986), 395–404.
- [29] R. Salem, *On singular monotonic functions whose spectrum has a given Hausdorff dimension*, Ark. Mat. **1** (1950), 353–365.
- [30] W. Schlag, *$L^p \rightarrow L^q$ 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.

- [31] W. Schlag, *A geometric proof of Bourgain's circular maximal theorem*, Duke. J. Math. **93** (1998), 505–533.
- [32] E.M. Stein, *Maximal functions: Spherical means*, Proc. Nat. Acad. Sci. U.S.A. **73** (1976), 2174–2175.
- [33] 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, pp. 307-355.
- [34] E.M. Stein, *Harmonic Analysis*, Princeton Univ. Press, Princeton, 1993.
- [35] J.A. Wellner, *Notes on exponential bounds via Azuma's inequality*, preprint, 2004.
- [36] N. Wiener, *The ergodic theorem*, Duke Math. J. **5** (1939), 1–18.
- [37] T. Wolff, *Lectures on Harmonic Analysis*, I. Łaba and C. Shubin, eds., Amer. Math. Soc., Providence, R.I. (2003).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA,
 VANCOUVER, B.C. V6T 1Z2, CANADA
ilaba@math.ubc.ca, malabika@math.ubc.ca