

*Pacific
Journal of
Mathematics*

**MULTILINEAR SINGULAR OPERATORS
WITH FRACTIONAL RANK**

CIPRIAN DEMETER, MALABIKA PRAMANIK AND CHRISTOPH THIELE

MULTILINEAR SINGULAR OPERATORS WITH FRACTIONAL RANK

CIPRIAN DEMETER, MALABIKA PRAMANIK AND CHRISTOPH THIELE

We prove bounds for multilinear operators on \mathbb{R}^d given by multipliers which are singular along a k -dimensional subspace. The new case of interest is when the rank k/d is not an integer. We also investigate connections with the concept of true complexity from additive combinatorics.

1. Introduction

Let $n \geq 3$ and $d \geq 1$. Consider a multiplier $M(\vec{\xi}^{(1)}, \dots, \vec{\xi}^{(n)})$ on the vector space

$$\Gamma := \{\vec{\xi} := (\vec{\xi}^{(1)}, \dots, \vec{\xi}^{(n)}) \in (\mathbb{R}^d)^n : \sum_{i=1}^n \vec{\xi}^{(i)} = \vec{0}\}.$$

This gives rise to the multilinear operator

$$\begin{aligned} T(F_1, \dots, F_{n-1})^{\wedge(-\vec{\xi}^{(n)})} \\ = \int \delta(\vec{\xi}^{(1)} + \dots + \vec{\xi}^{(n)}) M(\vec{\xi}) \widehat{F}_1(\vec{\xi}^{(1)}) \dots \widehat{F}_{n-1}(\vec{\xi}^{(n-1)}) d\xi^{(1)} \dots d\xi^{(n-1)}. \end{aligned}$$

on $n-1$ functions on \mathbb{R}^d .

Theorem 1.1. *Let Γ' be a generic linear subspace of Γ of dimension $k \geq 0$, and assume that $0 \leq k/d < n/2$. Assume the multiplier $M : \Gamma \rightarrow \mathbb{R}$ satisfies*

$$|\partial^\alpha M(\vec{\xi})| \lesssim \text{dist}(\vec{\xi}, \Gamma')^{-|\alpha|},$$

for all partial derivatives up to some finite order. Then

$$T : L^{p_1} \times \dots \times L^{p_{n-1}} \rightarrow L^{p'_n},$$

whenever $2 < p_i \leq \infty$ for each $1 \leq i \leq n$ and

$$(1) \quad 1/p_1 + \dots + 1/p_n = 1,$$

where p'_n is the conjugate exponent to p_n .

MSC2000: 42B20.

Keywords: multilinear singular integral operator, fractional rank, true complexity.

Demeter acknowledges support by a Sloan Research Fellowship and from NSF grant DMS-0556389. Pramanik acknowledges support from NSERC Discovery grant 341763. Thiele acknowledges support from NSF grant DMS-0701302.

The generic character of Γ' in [Theorem 1.1](#) is understood with respect to the Lebesgue measure (for example). In fact, we will need Γ' to satisfy some precise nondegeneracy conditions, and they are generically satisfied. To give the reader a grasp on what these conditions amount to, we will describe them now in the case $d = 2$. The case $d > 3$ involves very similar considerations and will be described in [Section 4C](#).

There will be two sets of nondegeneracy requirements. The first one is that Γ' can be parametrized by any k of the canonical variables. We will work a lot with the following parametrization. For $1 \leq i \leq n$, let $G_i : \mathbb{R}^k \rightarrow \mathbb{R}^d$ be the linear functions such that

$$\begin{aligned} G_1(\xi_1, \dots, \xi_k) &= (\xi_1, \dots, \xi_d), \\ G_2(\xi_1, \dots, \xi_k) &= (\xi_{d+1}, \dots, \xi_{2d}), \\ &\vdots \\ G_{[k/d]+1}(\xi_1, \dots, \xi_k) &= (\xi_{d[k/d]+1}, \dots, \xi_k, \dots), \end{aligned}$$

where the last entries of $G_{[k/d]+1}$ (the ones after the ξ_k entry) and the entries of the remaining G_i for $[k/d] + 1 \leq i \leq n$ are uniquely determined by the requirement that the function $G_1 \times \dots \times G_n$ maps into Γ' .

Let m be the smallest integer that is greater than or equal to k/d , that is, $[k/d]$. We will use the notation $\vec{\xi}^{(i)} := (\xi_1^{(i)}, \dots, \xi_k^{(i)}) \in \mathbb{R}^k$. Let $i_1, \dots, i_k \in \{1, \dots, n\}$ be pairwise distinct indices. Consider the following system of (k vector, or equivalently $2k$ scalar) linear equations in $2k$ variables $\xi_1^{(1)}, \dots, \xi_k^{(2)} \in \mathbb{R}$, and coefficients $v_j \in \mathbb{R}^2$:

$$(2) \quad \left\{ \begin{array}{l} G_{i_1}(\vec{\xi}^{(1)}) - G_{i_1}(\vec{\xi}^{(2)}) = v_1, \\ G_{i_2}(\vec{\xi}^{(1)}) = v_2, \\ \vdots \\ G_{i_m}(\vec{\xi}^{(1)}) = v_m, \\ G_{i_{m+1}}(\vec{\xi}^{(2)}) = v_{m+1}, \\ \vdots \\ G_{i_k}(\vec{\xi}^{(2)}) = v_k. \end{array} \right.$$

When $d = 2$, [Theorem 1.1](#) has the following precise formulation.

Theorem 1.2. *Let $d = 2$ and let Γ' be a linear subspace of Γ of dimension $k \geq 0$. Assume Γ' is the graph over every k of the canonical variables. Moreover, assume that the system (2) has a unique solution (for each choice of v_i) for each pairwise distinct $i_1, \dots, i_k \in \{1, \dots, n\}$. If the remaining hypotheses from [Theorem 1.1](#) are satisfied, then its conclusion will hold.*

It is not hard to see that the requirements in [Theorem 1.2](#) are satisfied for a generic Γ' . The assumption on the compatibility of the system (2) is one of many (similar in style) that work for our approach, and are guaranteed to hold generically. Various other possible alternative assumptions will become apparent from our later analysis. The minimal nondegeneracy conditions that are needed for [Theorem 1.1](#) (or even for our approach) are probably very hard to find, and beyond the goal of this paper. We point out however that if Γ' is degenerate in the sense that it fails to be the graph over *some particular* choice of k canonical variables, then the analysis of the operator T complicates to a significant extent. This has been observed and investigated in [[Demeter and Thiele 2010](#)] in the case $d = 2$.

[Theorem 1.1](#) was proved in [[Muscalu et al. 2002](#)] in the case $d = 1$, so our theorem is only new in the case $d \geq 2$. There, the theorem is proved under just the first nondegeneracy assumption, that Γ' is the graph over any k of the canonical variables, and the result there is proved for a larger class of indices p_i . To simplify our exposition, we choose to prove our theorem in the *locally* L^2 case $p_i > 2$.

A key parameter for our analysis is m , introduced earlier. We will refer to k/d as the *rank* of the operator. When this rank is an integer (and thus equal to m), or more generally, when $m < n/2$, [Theorem 1.1](#) will follow by a fairly straightforward adaptation of the argument in [[Muscalu et al. 2002](#)] to the d -dimensional setting. The novelty here is that k/d can be fractional and sufficiently close to $n/2$ to allow for the “bad” case $m \geq n/2$. A first new case of interest where our theorem is applicable is when $d = 2$, $k = 3$ and $n = 4$.

A simplified version of our approach also gives an alternative proof to the result in [[Muscalu et al. 2002](#)] (the $d = 1$ case), at least in the case $p_i > 2$ (see [Section 4D](#)). Our proof and theirs share much of the analytic part of the argument. Theirs however is structured around an induction on k that is not available in the fractional rank case. We eliminate the induction from the argument, and treat all k in a similar fashion. This new type of approach will involve rather delicate combinatorics.

[Theorem 1.1](#) also has a kernel formulation:

Theorem 1.3. *Let $K : \mathbb{R}^{d(n-1)-k} \rightarrow \mathbb{R}$ be a Calderón–Zygmund kernel, and let $l_i : \mathbb{R}^{d(n-1)-k} \rightarrow \mathbb{R}^d$ be $n - 1$ generic linear forms. Let p_i be as in [Theorem 1.1](#), and assume $0 \leq k/d < n/2$. For Schwartz functions $F_1, \dots, F_n : \mathbb{R}^d \rightarrow \mathbb{R}$ define the multilinear operator*

$$T(F_1, \dots, F_{n-1})(\vec{x}) := \int_{\mathbb{R}^{d(n-1)-k}} \prod_{i=1}^{n-1} F_i(\vec{x} + l_i(\vec{t})) K(\vec{t}) d\vec{t} \quad \text{for } \vec{x} \in \mathbb{R}^d.$$

Then T extends to a bounded operator

$$T : L^{p_1} \times \dots \times L^{p_{n-1}} \rightarrow L^{p'_n}.$$

In the case $k > 0$, the operators we investigate will typically have some modulation invariance. As a consequence, proving their boundedness will involve time-frequency analysis similar to that of the proof of the bilinear Hilbert transform [Lacey and Thiele 1997; 1999].

The assumption $k/d < n/2$ is crucial to our analysis. It can be shown in particular to guarantee that T has no symmetries of higher order (that is, quadratic symmetries). On the other hand, even in the one-dimensional case, the quadratic symmetries may be¹ present when $k \geq n/2$. Perhaps the most famous example with $k/d = n/2$ is the trilinear Hilbert transform ($n = 4$, $k = 2$ and $d = 1$)

$$(3) \quad T(F_1, F_2, F_3)(x) = \int_{\mathbb{R}} F_1(x+t)F_2(x-t)F_3(x+2t) \frac{dt}{t}.$$

Another important example with $k/d = n/2$ is the degenerate two-dimensional bilinear Hilbert transform ($n = 3$, $k = 3$ and $d = 2$)

$$(4) \quad T(F_1, F_2)(x, y) = \int_{\mathbb{R}} F_1(x+t, y)F_2(x, y+t) \frac{dt}{t},$$

where the quadratic symmetries are not singled out, but rather part of an infinite group of symmetries generated by degeneracies. See [Demeter and Thiele 2010] for details. The current techniques seem not enough to address the case $k/d = n/2$, where it is likely that some form of “quadratic Fourier analysis” will play a role. A quick single scale heuristics is provided in Section 4A. However, Theorem 1.1 shows that they can address the case of a rank k/d arbitrarily close to (and smaller than) $n/2$. This paper grew partly as an attempt to get more light on these issues.

A second motivation comes from connections with additive combinatorics, in particular with the issue of *true complexity* of a system of linear equations. Our analysis makes the point that nondegenerate systems characterized by $k/d < n/2$ have true complexity 1, in the language from [Gowers and Wolf 2010]. These things are described in Section 4A. The single scale heuristics provided there sheds a lot of light on the difficulties we encounter in the multiscale context, and we encourage the reader to go over that section first.

To investigate the boundedness properties of the operator in Theorem 1.1, it will be convenient to work with the dualized form defined by

$$\begin{aligned} \Lambda(F_1, \dots, F_n) &:= \int_{\mathbb{R}^d} T(F_1, \dots, F_{n-1})(\vec{x})F_n(\vec{x}) d\vec{x} \\ &= \int \delta(\vec{\xi}^{(1)} + \dots + \vec{\xi}^{(n)})M(\vec{\xi})\widehat{F}_1(\vec{\xi}^{(1)}) \dots \widehat{F}_n(\vec{\xi}^{(n)}) d\vec{\xi}. \end{aligned}$$

¹These symmetries are however not *guaranteed* to exist. In [Gowers and Wolf 2010] there are examples in the case $d = 1$ and $k = n/2$ that do not have any quadratic or higher order symmetries.

We will show that

$$|\Lambda(F_1, \dots, F_n)| \lesssim \prod_{i=1}^n \|F_i\|_{p_i}.$$

In the next section we will discretize the form and convert the problem to the boundedness of a model sum operator. We will use wave packets and multidimensional boxes (called tiles) to serve as their Heisenberg boxes. The tiles are then organized into trees, and eventually into certain products of trees, called vector trees. Most of the argument is then devoted to estimating the counting function associated with these vector trees. This is the main new contribution of our paper, and makes the object of [Section 4](#).

We will assume $m \geq 2$. The case $m = 0$ (that is, $k = 0$) is entirely classical and goes back to the work of Coifman and Meyer. No modulation symmetries are present in this case. The case $m = 1$ can be addressed by the argument in [[Muscalu et al. 2002](#)], by crudely majorizing the rank k/d by m . Indeed, since $1 = m < 3/2 \leq n/2$, we could treat the operator as if it had rank $m \in \mathbb{Z}$. Alternatively, one can apply the argument from [Section 4D](#) here.

2. Discretization

From now on, the notation $|\cdot|$ will refer to the cardinality of a finite set, the length of an interval or the volume of a multidimensional interval, depending on the context. The side length of a cube R will be denoted with $l(R)$. The discretization procedure in this section is very similar to the one from [[Muscalu et al. 2002](#)]. We omit most of the details.

We will work with the constants

$$1 \ll C_0 \ll C_1 \ll C_2 \ll C_3 \ll C_4,$$

whose values will not be specified explicitly, but will rather be clear from the context. The constant C_0 will be chosen first. It will be large enough depending on Γ' , n and d . Then C_1 will be chosen large enough, depending on Γ' but also on the choice of C_0 . Then C_2 , C_3 and C_4 are chosen in this order, sufficiently large compared to their predecessor in the sequence. C_4 will be an integral power of 2. No upper bounds will be forced upon C_i in terms of C_{i-1} , so when some C_i is selected, it can be chosen as large as desired. The fact that these constants will depend on Γ' will be reflected in the fact that the bounds in [Theorem 2.1](#) below also depend on Γ' . This dependence will be ignored.

Let $\bar{\Omega}$ be a finite set of nd -dimensional cubes $\bar{\omega} = \bar{\omega}_1 \times \dots \times \bar{\omega}_n$, where each $\bar{\omega}_i$ is a d -dimensional cube. These cubes are a sparse enough subset of a Whitney decomposing the frequency space $\mathbb{R}^{nd} \setminus \Gamma'$. They will serve to localize various

pieces of the multiplier M . For each $1 \leq i \leq n$ define the projections

$$\bar{\mathbf{O}}_i := \{\bar{\omega}_i : \bar{\omega} \in \bar{\mathbf{O}}\}.$$

These sets will satisfy the following properties:

(i) Separation in scale: For each $\bar{\omega}_1 \times \cdots \times \bar{\omega}_n \in \bar{\mathbf{O}}$, there is a $l \in \mathbb{Z}$ such that

$$|\bar{\omega}_1| = \cdots = |\bar{\omega}_n| = (C_4)^l.$$

(ii) Separation in distance: If $\bar{\omega}_i \neq \bar{\omega}_{i'} \in \bar{\mathbf{O}}_i$ and $|\bar{\omega}_i| = |\bar{\omega}_{i'}|$, then

$$\text{dist}(\bar{\omega}_i, \bar{\omega}_{i'}) \geq C_4 |\bar{\omega}_i|.$$

(iii) Whitney property: For each $\bar{\omega} \in \bar{\mathbf{O}}$, we have

$$10^{-1} C_0 \text{diam}(\bar{\omega}) \leq \text{dist}(\bar{\omega}, \Gamma') \leq 10 C_0 \text{diam}(\bar{\omega}).$$

(iv) Rank m : Any m of the n components of some $\bar{\omega} \in \bar{\mathbf{O}}$ determine uniquely the remaining $n - m$ components.

Let \mathcal{D} be the set of all dyadic cubes in \mathbb{R}^d . Let ϕ be a smooth function whose Fourier transform is adapted² to the cube $[-1/2, 1/2]^d$. For each $R \in \mathcal{D}$ with center $(c_1(R), \dots, c_d(R))$ and each $\bar{\omega}_i \in \bar{\mathbf{O}}_i$ such that $|R||\bar{\omega}_i| = 1$, we define the L^2 normalized wave-packet

$$\phi_{R \times \bar{\omega}_i}(x_1, \dots, x_d) = \frac{1}{|R|^{1/2}} \phi\left(\frac{x_1 - c_1(R)}{l(R)}, \dots, \frac{x_d - c_d(R)}{l(R)}\right) e^{i(c_1(\bar{\omega}_i)x_1 + \cdots + c_d(\bar{\omega}_i)x_d)}.$$

By using a standard discretization procedure like in [Muscalu et al. 2002], involving a decomposition for the multiplier M adapted to the set $\bar{\mathbf{O}}$, and then a Gabor basis decomposition for each F_i , Theorem 1.1 will follow from the discretized version below:

Theorem 2.1. *Let $F_i \in L^{p_i}(\mathbb{R}^d)$ with $2 < p_i \leq \infty$ as in (1). Let $\bar{\mathbf{O}}$ be any finite set satisfying (i)–(iv). Then for generic Γ' ,*

$$\sum_{\substack{R \in \mathcal{D}, \bar{\omega} \in \bar{\mathbf{O}} \\ |R||\bar{\omega}_i|=1}} |R|^{1-n/2} \prod_{i=1}^n |\langle F_i, \phi_{R \times \bar{\omega}_i} \rangle| \lesssim \prod_{i=1}^n \|F_i\|_{p_i}.$$

Moreover, the implicit constant in the above inequality is independent of F_i and of the set $\bar{\mathbf{O}}$.

We will fix $\bar{\mathbf{O}}$ for the remaining part of the paper. The following rank properties will be consequences of (i)–(iv) above (see [Muscalu et al. 2002, Section 6] for details) and of the requirement that Γ' is the graph over any k of the canonical variables:

²That means supported in $[-1/2, 1/2]^d$ and with the first few derivatives bounded by one.

- (v) **Overlapping:** If for some $\bar{\omega}, \bar{\omega}' \in \bar{\mathbf{O}}$ and for some $A \subset \{1, 2, \dots, n\}$ with $|A| = m$ we have $\bar{\omega}_i \subset 3\bar{\omega}'_i$ for each $i \in A$, then $\bar{\omega}_j \subset C_1\bar{\omega}'_j$ for each $1 \leq j \leq n$.
- (vi) **Two lacunary indices:** If for some $\bar{\omega}, \bar{\omega}' \in \bar{\mathbf{O}}$ with the additional property that $\text{diam}(\bar{\omega}) < \text{diam}(\bar{\omega}')$ and for some $A \subset \{1, 2, \dots, n\}$ with $|A| = m$, we have $\bar{\omega}_i \subseteq 3\bar{\omega}'_i$ for each $i \in A$, then there exist $1 \leq i_1 \neq i_2 \leq n$ such that $\bar{\omega}_j \not\subseteq 3\bar{\omega}'_j$ for each $j \in \{i_1, i_2\}$.

Definition 2.2 [Grafakos and Li 2004]. A set \mathcal{G} of intervals in \mathbb{R}^d is called a *central grid* if

- (G1) $R, R' \in \mathcal{G}$ and $R \cap R' \neq \emptyset$ implies $R \subseteq R'$ or $R' \subseteq R$,
- (G2) $R, R' \in \mathcal{G}$ and $R \subseteq R'$ implies $C_2R \subseteq R'$,
- (G3) $R, R' \in \mathcal{G}$ and $2|R| < |R'|$ implies $C_2|R| < |R'|$,
- (G4) $R, R' \in \mathcal{G}$ and $|R| < |R'| \leq 2|R|$ implies $\text{dist}(R, R') \geq C_2|R'|$.

It turns out that each sufficiently sparse set of cubes can be turned into a central grid. See the considerations following [Grafakos and Li 2004, Definition 1] for details.

Lemma 2.3 (centralization). *Let \mathcal{G}_0 be a set of d -dimensional cubes such that*

- (1) $\bar{R}, \bar{R}' \in \mathcal{G}_0$ and $|\bar{R}| < |\bar{R}'|$ implies $|\bar{R}| < C_3|\bar{R}'|$, and
- (2) $\bar{R}, \bar{R}' \in \mathcal{G}_0$ and $|\bar{R}| = |\bar{R}'|$ implies $\text{dist}(\bar{R}, \bar{R}') \geq C_3|\bar{R}|$.

Then for each $\bar{R} \in \mathcal{G}_0$, there is a d -dimensional interval R (not necessarily a cube) such that $\bar{R} \subset R \subset 2\bar{R}$ and such that the set $\mathcal{G} := \{R : \bar{R} \in \mathcal{G}_0\}$ is a central grid.

Note that both the set $\bar{\mathbf{O}}_i$ and the set $2C_1\bar{\mathbf{O}}_i := \{2C_1\bar{\omega}_i : \bar{\omega}_i \in \bar{\mathbf{O}}_i\}$ satisfy the requirements in Lemma 2.3 (if for example $C_3 \ll C_4 - 2C_1$). By applying Lemma 2.3 to each set, we associate to each $\bar{\omega}_i \in \bar{\mathbf{O}}_i$ two intervals ω_i and $\tilde{\omega}_i$ such that $\bar{\omega}_i \subseteq \omega_i \subseteq 2\bar{\omega}_i$, $2C_1\bar{\omega}_i \subseteq \tilde{\omega}_i \subseteq 4C_1\bar{\omega}_i$, and such that the set

$$\mathbf{O}_i := \{\omega_i : \bar{\omega}_i \in \bar{\mathbf{O}}_i\}$$

is a grid, while the set

$$\tilde{\mathbf{O}}_i := \{\tilde{\omega}_i : \bar{\omega}_i \in \bar{\mathbf{O}}_i\}$$

is a central grid. Note that we will need a stronger assumption on the second set. For each $\bar{\omega} = \bar{\omega}_1 \times \dots \times \bar{\omega}_n$, we will use the notation $\omega = \omega_1 \times \dots \times \omega_n$ and $\tilde{\omega} = \tilde{\omega}_1 \times \dots \times \tilde{\omega}_n$, and these two new intervals will form the sets \mathbf{O} and $\tilde{\mathbf{O}}$, respectively.

We hereafter abandon the set $\bar{\mathbf{O}}$ and only refer to the sets \mathbf{O} and $\tilde{\mathbf{O}}$.

Note that the side lengths of each interval $\omega \in \mathbf{O}$ are within a factor of two from each other. Denote by $\mathcal{R}(\omega_i)$ the set of all dyadic cubes R in \mathbb{R}^d such that $|R||\bar{\omega}_i| = 1$. Since $\bar{\omega}$ is a cube, the sets $\mathcal{R}(\omega_i)$ are all the same for $1 \leq i \leq n$. For

each $1 \leq i \leq n$, let \mathbf{S}_i denote the set of all $s_i := R_{s_i} \times \omega_{s_i}$, with ω_i ranging through \mathbf{O}_i and R_{s_i} ranging through $\mathcal{R}(\omega_i)$. The set \mathbf{S} will consist of all $s = R_s \times \omega_s$, with $\omega_s := \omega_{s_1} \times \cdots \times \omega_{s_n}$ ranging through \mathbf{O} and R_s ranging through (any of the) $\mathcal{R}(\omega_{s_i})$.

An element $s = R_s \times \omega_{s_1} \times \cdots \times \omega_{s_n} \in \mathbf{S}$ will be referred to as a multitile, while its components $s_i := R_{s_i} \times \omega_{s_i} \in \mathbf{S}_i$ will be referred to as tiles. The intervals R_s and ω_s will be referred to as the spatial and frequency components of s .

We introduce some relations of order, which are very similar to the ones in [Muscalu et al. 2002, Definition 6.1].

Definition 2.4. Let $s_i, s'_i \in \mathbf{S}_i$ be two tiles. We write

- $s_i \leq s'_i$ if $R_{s_i} \subseteq R_{s'_i}$ and $\omega_{s'_i} \subseteq \omega_{s_i}$,
- $s_i \lesssim s'_i$ if $R_{s_i} \subseteq R_{s'_i}$ and $\tilde{\omega}_{s'_i} \subseteq \tilde{\omega}_{s_i}$,
- $s_i \lesssim' s'_i$ if $s_i \lesssim s'_i$ but $s_i \not\leq s'_i$.

From the rank properties (iv)–(vi) and the grid structure of \mathbf{O} and $\tilde{\mathbf{O}}$ we get:

- (r1) Rank m : Any m of the n frequency components of some $s \in \mathbf{S}$ determine uniquely the remaining $n - m$ components.
- (r2) Overlapping: If for some $s, s' \in \mathbf{S}$ and for some $A \subset \{1, 2, \dots, n\}$ with $|A| = m$ we have $s_i \leq s'_i$ for each $i \in A$, then $s_j \lesssim s'_j$ for each $1 \leq j \leq n$.
- (r3) Two lacunary indices: If for some $s, s' \in \mathbf{S}$ with $|I_{s'}| < |I_s|$ and for some $A \subset \{1, 2, \dots, n\}$ with $|A| = m$ we have $s_i \leq s'_i$ for each $i \in A$, then there exist $1 \leq i_1 \neq i_2 \leq n$ such that $s_j \lesssim' s'_j$ for each $j \in \{i_1, i_2\}$.

Properties (r2) and (r3) are all about the frequency components of multitiles, the spatial components do not play any role.

We also record for future reference the grid properties satisfied by the multitiles:

- (r4) The set $\{\omega_{s_i} : s_i \in \mathbf{S}_i\}$ is a grid.³
- (r5) The set $\{\tilde{\omega}_{s_i} : s_i \in \mathbf{S}_i\}$ is a central grid.

It is also clear (due to (iii)) that if C_0 is sufficiently large, then

- (r6) for each $s \in \mathbf{S}$, we have $C_0^2 \omega_s \cap \Gamma' \neq \emptyset$.

For each tile $s_i = R_{s_i} \times \omega_{s_i}$, we will have a wave packet associated with it, namely

$$\phi_{s_i} := \phi_{R_{s_i} \times \bar{\omega}_{s_i}},$$

where $\bar{\omega}_{s_i}$ is the cube in $\tilde{\mathbf{O}}_i$ that generates ω_{s_i} via the procedure described earlier. Note that the Fourier transform of ϕ_{s_i} is supported in ω_{s_i} , while spatially, ϕ_{s_i} is a bump function quasilocalized near the cube R_{s_i} .

With this notation, Theorem 2.1 can be rephrased as follows:

³While we can arrange that this set is a *central* grid, too, we will not need this strong assumption.

Theorem 2.5. Let $F_i \in L^{p_i}(\mathbb{R}^d)$ with $2 < p_i \leq \infty$ as in (1). Let \mathbf{S} be a set of multitiles satisfying (r1)–(r6). Then for generic Γ'

$$\sum_{s \in \mathbf{S}} |R_s|^{1-n/2} \prod_{i=1}^n |\langle F_i, \phi_{s_i} \rangle| \lesssim \prod_{i=1}^n \|F_i\|_{p_i}.$$

The implicit constant in this inequality is independent of F_i and of the set \mathbf{S} .

The genericity of Γ' will imply an additional rank property (r7), which we choose not to state explicitly at this point, but which will become clear later (for example, see Section 4B).

3. Trees

To prove Theorem 2.5 we will organize each set \mathbf{S}_i into smaller structures, called trees.

Definition 3.1. Let $R_{\mathbf{T}}$ be a dyadic cube in \mathbb{R}^d and let $\zeta_{\mathbf{T}} \in \mathbb{R}^d$. A set of tiles $\mathbf{T} \subset \mathbf{S}_i$ is called an i -tree (also sometimes referred to as tree, when the index i is either not important or when it is clear from the context) with top $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$ if $R_{s_i} \subseteq R_{\mathbf{T}}$ and $\zeta_{\mathbf{T}} \in \tilde{\omega}_{s_i}$ for each $s_i \in \mathbf{T}$.

In case $\zeta_{\mathbf{T}} \in \omega_{s_i}$ for each $s_i \in \mathbf{T}$, the tree \mathbf{T} is called i -overlapping. If $\zeta_{\mathbf{T}} \in \tilde{\omega}_{s_i} \setminus \omega_{s_i}$ for each $s_i \in \mathbf{T}$, the tree will be called i -lacunary.

A tree consisting of a single tile is both lacunary and overlapping (these are actually the only examples of such trees). In general, a tree must not necessarily be overlapping or lacunary. However, each tree can be split as the disjoint union of an overlapping tree and a lacunary tree.

Trees will be used to construct similar structures consisting of multitiles, called vector trees.

Definition 3.2. Let $R_{\tilde{\mathbf{T}}}$ be a dyadic cube in \mathbb{R}^d and let $\zeta_{\tilde{\mathbf{T}}} = (\zeta_{\tilde{\mathbf{T}},1}, \dots, \zeta_{\tilde{\mathbf{T}},n}) \in \mathbb{R}^{dn}$. A set of multitiles $\tilde{\mathbf{T}} \subset \mathbf{S}$ is called a vector tree with top $(R_{\tilde{\mathbf{T}}}, \zeta_{\tilde{\mathbf{T}}})$ if for each $1 \leq i \leq n$, the projection $\mathbf{T}_i := \{s_i : s \in \tilde{\mathbf{T}}\}$ is an i -tree with top $(R_{\tilde{\mathbf{T}}}, \zeta_{\tilde{\mathbf{T}},i})$.

The vector tree $\tilde{\mathbf{T}}$ is called i -overlapping if its projection \mathbf{T}_i is an i -overlapping tree. If this is the case, we call the index i overlapping. Similarly, the vector tree $\tilde{\mathbf{T}}$ is called i -lacunary, if \mathbf{T}_i is an i -lacunary tree. If this is the case, we call the index i lacunary.

Remark 3.3. The rank property (r3) implies that each vector tree has at least two lacunary indices.

Definition 3.4. Let $\mathbf{P} \subseteq \mathbf{S}_i$ be a set of tiles. Its i -size is defined as

$$\text{size}_i(\mathbf{P}) := \sup_{\mathbf{T} \subseteq \mathbf{P}} \left(\frac{1}{|R_{\mathbf{T}}|} \sum_{s_i \in \mathbf{T}} |\langle F_i, \phi_{s_i} \rangle|^2 \right)^{1/2},$$

where the supremum is taken over all lacunary i -trees $\mathbf{T} \subseteq \mathbf{P}$ with tops $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$.

The definition of the size of a set \mathbf{P} depends on the choice of the function F_i . We choose not to index the size by F_i , since F_i will later be fixed.

The next lemma shows that the size is dominated by the Hardy–Littlewood maximal function (see again [Muscalu et al. 2002] for details)

$$M_2 F(\vec{x}) := \sup_{\substack{\vec{x} \in R \\ R \text{ cube in } \mathbb{R}^d}} \left(\frac{1}{|R|} \int_R |F|^2(\vec{y}) d\vec{y} \right)^{1/2}.$$

Lemma 3.5. *Let \mathbf{T} be a lacunary i -tree with top $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$. Then*

$$\left(\frac{1}{|R_{\mathbf{T}}|} \sum_{s_i \in \mathbf{T}} |\langle F_i, \phi_{s_i} \rangle|^2 \right)^{1/2} \lesssim \inf_{\vec{x} \in R_{\mathbf{T}}} M_2 F_i(\vec{x}).$$

As a consequence, $\text{size}_i(\mathbf{P}) \lesssim \sup_{s_i \in \mathbf{P}} \inf_{\vec{x} \in R_{s_i}} M_2 F_i(\vec{x})$.

We will use the following estimate for vector trees:

Lemma 3.6. *Let $\vec{\mathbf{T}}$ be a vector tree with top $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$. Then*

$$\sum_{s \in \mathbf{T}} |R_s|^{1-n/2} \prod_{i=1}^n |\langle F_i, \phi_{s_i} \rangle| \leq |R_{\mathbf{T}}| \prod_{i=1}^n \text{size}_i(\mathbf{T}_i).$$

Proof. Use Hölder’s inequality with an l^2 estimate for two lacunary indices and with an l^∞ estimate for the remaining $n - 2$ indices. □

For a set \mathcal{F} of trees or vector trees, we will use the notation

$$N_{\mathcal{F}}(\vec{x}) := \sum_{\mathbf{T} \in \mathcal{F}} 1_{R_{\mathbf{T}}}(\vec{x}).$$

Also, we will denote by $\|N_{\mathcal{F}}\|_{\text{BMO}}$ the dyadic BMO norm of $N_{\mathcal{F}}$.

We next show how to split a set \mathbf{P} of tiles into sets of trees with controlled size. First, we show how to cut the size in half.

Lemma 3.7. *Let $\mathbf{P} \subseteq \mathbf{S}_i$ be a set of tiles. There is a set \mathcal{F} of disjoint (as sets of tiles) i -trees in \mathbf{P} such that*

$$(5) \quad R_{\mathbf{T}} \subset \{ \vec{x} : \inf_{\vec{x} \in R_{\mathbf{T}}} M_2 F_i(\vec{x}) \gtrsim \text{size}_i(\mathbf{P}) \} \quad \text{for } \mathbf{T} \in \mathcal{F},$$

$$(6) \quad \sum_{\mathbf{T} \in \mathcal{F}} |R_{\mathbf{T}}| \lesssim (\text{size}_i(\mathbf{P}))^{-2} \|F_i\|_2^2,$$

$$(7) \quad \|N_{\mathcal{F}}\|_{\text{BMO}} \lesssim (\text{size}_i(\mathbf{P}))^{-2} \left(\sup_{\mathbf{T} \in \mathcal{F}} \inf_{\vec{x} \in R_{\mathbf{T}}} M_2 F_i(\vec{x}) \right)^2,$$

$$(8) \quad \|N_{\mathcal{F}}\|_q \lesssim (\text{size}_i(\mathbf{P}))^{-2} \left(\sup_{\mathbf{T} \in \mathcal{F}} \inf_{\vec{x} \in R_{\mathbf{T}}} M_2 F_i(\vec{x}) \right)^2 (\text{size}_i(\mathbf{P}))^{p_i} \|F_i\|_{p_i}^{1/q},$$

where

$$1 \leq q < \infty \quad \text{and} \quad \text{size}_i(\mathbf{P} \setminus \bigcup_{\mathbf{T} \in \mathcal{F}} \mathbf{T}) < \frac{1}{2} \text{size}_i(\mathbf{P}).$$

Proof. The proof is very standard, so we only say a few words about it; see for example [Muscalu et al. 2002, Lemma 7.7] or [Muscalu et al. 2004, Lemma 6.6]. The trees in \mathcal{F} are selected in 2^d stages. We need first a definition. Let $1 \leq l \leq d$. A lacunary i -tree with top $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$ is said to be a $(+, l)$ tree if

$$(9) \quad (\zeta_{\mathbf{T}})_l > c_l(\omega_{s_i}) \quad \text{for each } s_i \in \mathbf{T},$$

where $c_l(\omega)$ denotes the l -th component of the center $c(\omega)$ of ω . Similarly, the tree is said to be a $(-, l)$ tree if the inequality is reversed in (9). It is easy to see that each lacunary tree \mathbf{T} is the disjoint union of at most 2^d lacunary subtrees having the same top as \mathbf{T} , each of which is either a $(+, l)$ tree or a $(-, l)$ tree for some $1 \leq l \leq d$. By pigeonholing, it follows that if we eliminate from \mathbf{P} all such trees having size_i at least $(1/(2 \times 2^{d/2}))\text{size}_i(\mathbf{P})$, the remaining set of tiles will have size less than $(1/2)\text{size}_i(\mathbf{P})$.

In a typical stage of the selection⁴ one selects lacunary trees \mathbf{T} with tops $(R_{\mathbf{T}}, \zeta_{\mathbf{T}})$ that are (say) $(+, l)$ trees for a fixed l , and that satisfy

$$(10) \quad \frac{1}{|R_{\mathbf{T}}|} \sum_{s_i \in \mathbf{T}} |\langle F_i, \phi_{s_i} \rangle|^2 \geq \frac{1}{4 \times 2^d} (\text{size}_i(\mathbf{P}))^2.$$

One always aims for the tree that has the minimal value for $(\zeta_{\mathbf{T}})_l$ among all the trees that qualify to be selected at that moment. After such a tree is selected, this tree is added to the set $\mathcal{F}_1(+, l)$ and its tiles are eliminated from \mathbf{P} . The remaining tiles that satisfy $\zeta_{\mathbf{T}} \in \tilde{\omega}_{s_i}$ and $R_{s_i} \subseteq R_{\mathbf{T}}$ will form a tree \mathbf{T}^{sat} , having the same top as \mathbf{T} . Add \mathbf{T}^{sat} to the set $\mathcal{F}_2(+, l)$ and eliminate its tiles from \mathbf{P} . Then the cycle repeats, that is, one starts searching again for a $(+, l)$ tree satisfying (10). When no such tree is available, the $(+, l)$ stage of the selection process ends. One adds all the trees from $\mathcal{F}_1(+, l)$ and from $\mathcal{F}_2(+, l)$ to the set \mathcal{F} . At this point one goes to the next stage of the selection process.

The key observation is that in each stage, the trees from the set $\mathcal{F}_1(+, l)$ have the following property:⁵ If $\mathbf{T}, \mathbf{T}' \in \mathcal{F}_1(+, l)$, $s_i \in \mathbf{T}$, $s'_i \in \mathbf{T}'$ and $\omega_{s_i} \subsetneq \omega_{s'_i}$, then $R_{s'_i} \cap R_{\mathbf{T}} = \emptyset$.

Let us briefly see why this property holds. The fact that $\omega_{s_i} \subsetneq \omega_{s'_i}$, the separation in scales, (r4), and the definition of $(+, l)$ lacunary trees implies that $(\zeta_{\mathbf{T}'})_l > (\zeta_{\mathbf{T}})_l$. This in turn implies that the tree \mathbf{T} was selected earlier than \mathbf{T}' . The separation in scales and the fact that $C_1 \ll C_2$ easily imply that $\zeta_{\mathbf{T}} \in \tilde{\omega}_{s'_i}$. If $R_{s'_i}$ and $R_{\mathbf{T}}$ intersected (and this can only mean that $R_{s'_i} \subseteq R_{\mathbf{T}}$), this together with $\zeta_{\mathbf{T}} \in \tilde{\omega}_{s'_i}$ would imply

⁴The order of the stages does not really matter.

⁵In the literature, this property is referred to as “strong disjointness”.

that right after \mathbf{T} was selected, s'_i would have qualified to be part of the tree \mathbf{T}^{sat} , and thus it would have been eliminated from \mathbf{P} before the selection of \mathbf{T}' began. The contradiction is immediate; thus we conclude that $R_{s'_i} \cap R_{\mathbf{T}} = \emptyset$.

What the property we just proved means is that any two trees from $\mathcal{F}_1(+, l)$ interact very weakly, in that tiles from different trees either have disjoint frequency components or strongly separated spatial components (so that the associated wave packets have little interaction). Using standard arguments, one could argue that this together with (10) implies the Bessel-type inequality

$$\sum_{\mathbf{T} \in \mathcal{F}_1(+, l)} |R_{\mathbf{T}}| \lesssim (\text{size}_i(\mathbf{P}))^{-2} \|F_i\|_2^2.$$

We can clearly replace $\mathcal{F}_1(+, l)$ by $\mathcal{F}_2(+, l)$ in this inequality.

By combining the contribution from all 2^d stages, we get (6). Also, (7) will follow similarly, by a standard localization argument. (8) follows from (5) and (7), via an application of the John–Nirenberg inequality (since $p_i > 2$). Finally, (5) is a consequence of Lemma 3.5. \square

By iterating Lemma 3.7 we get this:

Proposition 3.8. *Let $\mathbf{P}_i \subseteq \mathbf{S}_i$ be a set of tiles. Then one has the decomposition*

$$\mathbf{P}_i := \bigcup_{2^{-k} \leq \text{size}_i(\mathbf{P}_i)} \mathbf{P}_i^{(k)}, \quad \text{where } \text{size}_i(\mathbf{P}_i^{(k)}) \leq 2^{-k+1}$$

and each $\mathbf{P}_i^{(k)}$ is the (disjoint) union of a family $\mathcal{F}_i^{(k)}$ of trees such that

$$\|N_{\mathcal{F}_i^{(k)}}\|_q \lesssim 2^{2k} \left(\sup_{\mathbf{T} \in \mathcal{F}_i^{(k)}} \inf_{\vec{x} \in R_{\mathbf{T}}} M_2 F_i(\vec{x}) \right)^2 (2^{k_i p_i} \|F_i\|_{p_i}^{p_i})^{1/q} \quad \text{for } 1 \leq q < \infty.$$

This proposition gives us good control over the number of trees corresponding to each component $i \in \{1, \dots, n\}$. In the next section we will combine trees from each component to create vector trees and will relate their counting function to the ones of the individual trees.

4. Counting trees

Recall that m is the smallest integer greater than or equal to k/d , and that we have assumed that $m \geq 2$.

Throughout this section we will assume that we have a set of multitiles $\mathbf{P} \subset \mathbf{S}$. We also assume that each projection \mathbf{P}_i consists of a family \mathcal{F}_i of disjoint i -trees

$$\mathbf{P}_i = \bigcup_{\mathbf{T} \in \mathcal{F}_i} \mathbf{T}.$$

Our goal is to split the set \mathbf{P} into a family \mathcal{F} of vector trees with good pointwise control over the counting function $N_{\mathcal{F}}(\vec{x})$ of their tops in terms of the individual counting functions $N_{\mathcal{F}_i}(\vec{x})$ (see (20) and (29) below). To provide a better understanding of what we have to prove, we first give a single scale heuristics.

4A. Single scale heuristics. Assume we are in the particular case where each \mathbf{T} in each family \mathcal{F}_i consists of just one tile, of the form $[0, 1]^d \times \omega$. Thus, the set \mathcal{F}_i will consist of a family of pairwise disjoint tiles of scale 1. Assume that for each i we know the cardinality $|\mathcal{F}_i|$. The question in this case is, subject to axioms (r1)–(r6) and the genericity of Γ' , to estimate the cardinality $|\mathcal{F}|$, where \mathcal{F} consists of all the multitiles s having the property that $s_i \in \mathcal{F}_i$ for each i . Note first that (r1) immediately implies that

$$(11) \quad |\mathcal{F}| \leq \prod_{i=1}^m |\mathcal{F}_i|,$$

with the same being true for each choice of m indices. This leads to

$$|\mathcal{F}| \leq \prod_{i=1}^n |\mathcal{F}_i|^{m/n}.$$

This estimate is only satisfactory when $m < n/2$, for reasons that will become clear in the proof of Proposition 4.1; however, it will be of no use when $m \geq n/2$. The good news is that (11) is only sharp when $k/d = m$, in which case we also know that $m < n/2$. If $k/d < m$, the inequality can be improved, and one has to use the additional rank property (r7) guaranteed by the genericity of Γ' .

For simplicity, let us see this in the case $n = 4$, $d = 2$ and $k = 3$. Then the property (r7) will state that for each pairwise distinct $i_1, i_2, i_3 \in \{1, 2, 3, 4\}$, and for every two multitiles s, s' with $s_1 = s'_1$, knowledge of both s_2 and s'_3 will uniquely determine both s and s' . See Section 4B. Using this and applying Lemma 4.2 as in the next section, we get the improved inequality

$$|\mathcal{F}| \leq \prod_{i=1}^3 |\mathcal{F}_i|^{1/2},$$

and thus, after permuting indices, we get

$$|\mathcal{F}| \leq \prod_{i=1}^4 |\mathcal{F}_i|^{3/8}.$$

In general, one gets

$$(12) \quad |\mathcal{F}| \lesssim \prod_{i=1}^n |\mathcal{F}_i|^{(k/d)/n},$$

which is the optimal inequality. The important fact is that all exponents on the right hand side are less than $1/2$.

We have this consequence of (12):

Proposition 4.1. *Assume as before that $k/d < n/2$. Then, for generic linear forms $l_i : \mathbb{R}^{d(n-1)-k} \rightarrow \mathbb{R}^d$ with $1 \leq i \leq n-1$, we have*

$$\left| \int_{\vec{x} \in [0,1]^d} \int_{\vec{t} \in [0,1]^{d(n-1)-k}} \prod_{i=1}^{n-1} F_i(\vec{x} \oplus l_i(\vec{t})) F_n(\vec{x}) d\vec{t} d\vec{x} \right| \lesssim \prod_{i=1}^n \|\widehat{F}_i\|_\infty^{1-(2k/d)/n}$$

whenever $\|F_i\|_\infty = O(1)$.

Here \oplus denotes addition modulo 1. This shows that the single scale operator relevant to our problem is controlled by the Fourier transform, and thus it provides a heuristics for the fact that Fourier analysis will be able to address the multiscale version of the problem (that is, Theorem 2.5). In contrast to this, we mention that neither the operator in (3) nor that in (4) satisfy any similar bounds.

In the language of Gowers and Wolf [2010], Proposition 4.1 is saying that the system of equations associated with the linear forms l_i has *true complexity* 1.

It will become clear from the argument presented in the following sections that the precise conditions on Γ' needed in Proposition 4.1 (needed to guarantee (12)) amount to the following two requirements. To formulate them, we use the same notation that we have used so far. More exactly, Γ' is the linear subspace of $(\mathbb{R}^d)^n$ defined by

$$\left\{ (\vec{\eta}^{(1)}, \dots, \vec{\eta}^{(n)}) : \sum_{i=1}^{n-1} \vec{\eta}^{(i)} \cdot (\vec{x} + l_i(\vec{t})) + \eta^{(n)} \cdot \vec{x} = \mathbf{0} \in \mathbb{R}[\vec{x}, \vec{t}] \right\}.$$

We ask Γ' to be k -dimensional and parametrizable over every k canonical coordinates. We also ask that the following system of k \mathbb{R}^d -valued equations in d unknowns $\vec{\xi}^{(i)} \in \mathbb{R}^k$ is compatible for each distinct $i_1, \dots, i_m \in \{1, \dots, n\}$:

$$\left\{ \begin{array}{ll} G_{i_1}(\vec{\xi}^{(j)}) - G_{i_1}(\vec{\xi}^{(j+1)}) = v_j & \text{for } 1 \leq j \leq d-1, \\ G_{i_l}(\vec{\xi}^{(j)}) = v_{j,i_l} & \text{for } 1 \leq j \leq d, 2 \leq l \leq m-1, \\ G_{i_m}(\vec{\xi}^{(j)}) = v_{j,i_m} & \text{for } 1 \leq j \leq k+1-d(m-1). \end{array} \right.$$

Proof of Proposition 4.1. By discretizing as in the previous sections (and keeping the notation from there), we are reduced to proving that

$$\sum_{s \in \mathbf{S}} \prod_{i=1}^n |\langle F_i, \phi_{s_i} \rangle| \lesssim \prod_{i=1}^n \|\widehat{F}_i\|_\infty^{1-(2k/d)/n},$$

where \mathbf{S} consists of multitiles of scale 1. Note that $|\langle F_i, \phi_{s_i} \rangle|$ is roughly a Fourier coefficient of F_i . For each $2^{-n_i} \lesssim \|\widehat{F}_i\|_\infty$ let \mathcal{F}_{i,n_i} be the set of all tiles s_i with

$$|\langle F_i, \phi_{s_i} \rangle| \sim 2^{-n_i}.$$

Note that since $\|F_i\|_2 = O(1)$, we have $|\mathcal{F}_{i,n_i}| \lesssim 2^{2n_i}$ (Bessel's inequality). If we use (12), it follows that the cardinality of the set $\mathcal{F}(\vec{n})$ consisting of all multitiles $s \in \mathbf{S}$ such that $s_i \in \mathcal{F}_{i,n_i}$ for each i will be $O(\prod_{i=1}^n 2^{(2kn_i/d)/n})$.

The sum above is then bounded by

$$\sum_{\vec{n}: 2^{-n_i} \lesssim \|\widehat{F}_i\|_\infty} 2^{-\sum_i n_i} \prod_{i=1}^n 2^{(2kn_i/d)/n} \lesssim \prod_{i=1}^n \|\widehat{F}_i\|_\infty^{1-(2k/d)/n}. \quad \square$$

We make two remarks about Proposition 4.1, and in general about the applicability of our Fourier analytic techniques. The first remark shows that some nondegeneracy is in general needed. Take for example

$$\begin{aligned} \int_{(t,s) \in \mathbf{T}^2} \int_{(x,y) \in \mathbf{T}^2} F_1(x+t, y) F_2(x, y+s) F_3(x, y) dt ds dx dy \\ = \int_{(x,y) \in \mathbf{T}^2} \mathcal{F}_x(F_1)(0, y) \mathcal{F}_y(F_2)(x, 0) F_3(x, y) dx dy. \end{aligned}$$

One can check that here $k = d = 2$, and thus $k/d < n/2$. However, it can be easily seen that this form cannot be bounded by a power of (and in general, by no reasonable function of) $\|\widehat{F}_i\|_\infty$. The explanation is that Γ' is degenerate.

The second remark will show that with the current techniques, the requirement $k/d < n/2$ cannot be relaxed. We illustrate this in the case $d = 1$ with the single scale version of (3):

$$\int_{x \in [0,1]} \int_{t \in [0,1]} F_1(x \oplus t) F_2(x \ominus t) F_3(x \oplus 2t) F_4(x) dx dt.$$

Application of the Fourier inversion formula for each function shows that this form is essentially (up to some constants)

$$\sum_{\substack{(n_1, n_2, n_3, n_4) \\ n_1 - n_2 + 2n_3 = 0 \\ n_1 + n_2 + n_3 + n_4 = 0}} \prod_{i=1}^4 \widehat{F}_i(n_i).$$

Our approach relies on using the triangle inequality to bound this sum by

$$(13) \quad \sum_{\substack{(n_1, n_2, n_3, n_4) \\ n_1 - n_2 + 2n_3 = 0 \\ n_1 + n_2 + n_3 + n_4 = 0}} \prod_{i=1}^4 |\widehat{F}_i(n_i)|.$$

Choose $F_1(x) = F_2(x) = F_3(x) = e^{2\pi i N x^2}$ and $F_4(x) = e^{-2\pi i N x^2}$ for large N . One can check that when $i \in \{1, 2, 3\}$, we have $|\widehat{F}_i(n)| \sim N^{-1/2}$ for $N/100 \leq n \leq N$, and that $|\widehat{F}_4(n)| \sim N^{-1/2}$ for $-N \leq n \leq -N/100$. It is easy to see that the term in (13) is $\gtrsim 1$. Since one can check that also $\|\widehat{F}_i\|_\infty \lesssim N^{-1/2}$ for each i , the expression in (13) is not $O(\|\widehat{F}_i\|_\infty)$. Let \mathcal{F}_i refer to the set of n with $|\widehat{F}_i(n)| \sim N^{-1/2}$, and let \mathcal{F} be the set of quadruples (n_1, n_2, n_3, n_4) as in (13), such that $n_i \in \mathcal{F}_i$ for each i . We get

$$|\mathcal{F}| \sim \prod_{i=1}^4 |\mathcal{F}_i|^{1/2},$$

and this sharp inequality becomes inefficient for any application since the exponent is $1/2$.

A moment's reflection shows we can get the same outcome whenever $k = n/2$. Of course, our approach fails to address the case $k/d = n/2$, because of the use of the triangle inequality in (13). It is likely that the correct approach in this case (and in general, whenever $k/d \geq n/2$) is to analyze appropriate quadratic Fourier coefficients.

To address the multiscale case, we will have to count vector trees, rather than just multitiles. To make the exposition more transparent, we will exemplify our approach in the case $d = 2$. The general case $d > 2$ will follow via a similar argument and is briefly sketched afterward.

4B. The case $d = 2$. We start this section by proving that, under the particular assumptions on Γ' from Theorem 1.2, the multitiles in \mathbf{S} satisfy the following additional rank property:

(r7) For all distinct indices $i_1, \dots, i_k \subset \{1, 2, \dots, n\}$ and for all (not necessarily pairwise distinct) multitiles $s, s', p, p' \in \mathbf{S}$, we have

$$\text{dist}(\omega_s, \omega_p), \text{dist}(\omega_{s'}, \omega_{p'}) \leq C_0^5 D_{\max},$$

where

$$\begin{aligned} D_{\max} := & \max\{\text{dist}(\omega_{s_{i_1}}, \omega_{s'_{i_1}}), \text{dist}(\omega_{p_{i_1}}, \omega_{p'_{i_1}})\} \\ & + \max\{\text{dist}(\omega_{s_{i_l}}, \omega_{p_{i_l}}), 2 \leq l \leq m\} + \max\{\text{dist}(\omega_{s'_{i_l}}, \omega_{p'_{i_l}}), m+1 \leq l \leq k\} \\ (14) \quad & + \max\{\text{diam}(\omega_s), \text{diam}(\omega_{s'}), \text{diam}(\omega_p), \text{diam}(\omega_{p'})\}. \end{aligned}$$

It will be important that we can control both $\text{dist}(\omega_s, \omega_p)$ and $\text{dist}(\omega_{s'}, \omega_{p'})$ by a value at most C_0^5 times larger than D_{\max} . Property (r7) will be used twice in the future, in conjunction with a choice of constants such that $C_0 \ll C_1 \ll C_2$.

To see why (r7) holds, let

$$\gamma_s \in C_0^2 \omega_s \cap \Gamma', \quad \gamma_{s'} \in C_0^2 \omega_{s'} \cap \Gamma', \quad \gamma_p \in C_0^2 \omega_p \cap \Gamma' \quad \text{and} \quad \gamma_{p'} \in C_0^2 \omega_{p'} \cap \Gamma'$$

(by (r6)). To simplify notation, assume $i_l = l$. Note that

$$(15) \quad \text{dist}((\gamma_s)_1, (\gamma_{s'})_1), \text{dist}((\gamma_p)_1, (\gamma_{p'})_1), \text{dist}((\gamma_s)_l, (\gamma_p)_l), \text{dist}((\gamma_{s'})_l, (\gamma_{p'})_l) \leq C_0^3 D_{\max}.$$

We will use the notation $\vec{\zeta}^{(i)} := (\zeta_1^{(i)}, \dots, \zeta_k^{(i)}) \in \mathbb{R}^k$. Recall the functions G_i from the introduction. Consider the following system of (k vector, or equivalently $2k$ scalar) linear equations in $2k$ variables $\zeta_1^{(1)}, \dots, \zeta_k^{(2)} \in \mathbb{R}$, and coefficients $v_j \in \mathbb{R}^2$:

$$\left\{ \begin{array}{l} G_1(\vec{\zeta}^{(1)}) - G_1(\vec{\zeta}^{(2)}) = v_1, \\ G_2(\vec{\zeta}^{(1)}) = v_2, \\ \vdots \\ G_m(\vec{\zeta}^{(1)}) = v_m, \\ G_{m+1}(\vec{\zeta}^{(2)}) = v_{m+1}, \\ \vdots \\ G_k(\vec{\zeta}^{(2)}) = v_k \end{array} \right.$$

By hypothesis, the system has a unique solution for each v_j . Let now $\vec{\zeta}^{(1)}, \vec{\zeta}^{(2)} \in \mathbb{R}^k$ consist of the first k entries of $\gamma_s - \gamma_p$ and $\gamma_{s'} - \gamma_{p'}$, respectively, and let v_j be the values corresponding to this choice in the system above. Note that

$$|v_j| \leq 2C_0^3 D_{\max} \quad \text{for each } j \in \{1, \dots, k\}$$

by (15). It will follow that the solution $(\vec{\zeta}^{(1)}, \vec{\zeta}^{(2)})$ will be bounded in norm by the norm of (v_1, \dots, v_k) times a constant that only depends on the coefficients of the linear forms G_1, \dots, G_k . Thus, if C_0 is chosen large enough compared to these coefficients, we get that $\text{dist}(\gamma_s, \gamma_p), \text{dist}(\gamma_{s'}, \gamma_{p'}) \leq C_0^4 D_{\max}$. Now (r7) is immediate.

There will be two stages, each of which will generate some vector trees. In each stage, before we construct the vector trees we will have to carefully reshuffle the sets \mathbf{P}_i .

Let us describe the first stage of the construction. We will first aim at separating the trees in each family \mathcal{F}_i , and to achieve this we will employ a trick first used by Fefferman [1973]. Fix $1 \leq i \leq n$. For each $l \geq 0$, let

$$\mathbf{P}_i(l) := \{s_i \in \mathbf{P}_i : 2^l \leq |\{\mathbf{T} \in \mathcal{F}_i : \zeta_{\mathbf{T}} \in \tilde{\omega}_{s_i}, R_{s_i} \subseteq R_{\mathbf{T}}\}| < 2^{l+1}\}.$$

Note that $(\mathbf{P}_i(l))_{l \geq 0}$ forms a partition of \mathbf{P}_i . Next, we organize each $\mathbf{P}_i(l)$ into i -trees with top tiles. More precisely, consider the set of all tiles $\mathbf{P}_i^*(l) \subseteq \mathbf{P}_i(l)$ that are maximal with respect to the order relation \lesssim .

We observe a few things. First, the tiles in $\mathbf{P}_i^*(l)$ are pairwise not comparable under \lesssim . Second, for each tile $s_i \in \mathbf{P}_i(l)$ there is a *unique* tile $s_i^* \in \mathbf{P}_i^*(l)$ such that $s_i \lesssim s_i^*$. To see the uniqueness part, assume by contradiction that $s_i \lesssim s_i^*$ and $s_i \lesssim s_i^{**}$ for some $s_i^* \neq s_i^{**} \in \mathbf{P}_i^*(l)$. Then $R_{s_i^*} \cap R_{s_i^{**}} \neq \emptyset$, which forces $\tilde{\omega}_{s_i^*} \cap \tilde{\omega}_{s_i^{**}} = \emptyset$. This together with the fact that $s_i^*, s_i^{**} \in \mathbf{P}_i(l)$ will imply that

$$|\{\mathbf{T} \in \mathcal{F}_i : \zeta_{\mathbf{T}} \in \tilde{\omega}_{s_i}, R_{s_i} \subseteq R_{\mathbf{T}}\}| \geq 2 \times 2^l,$$

which forces the contradiction $s_i \in \bigcup_{l' \geq l+1} \mathbf{P}_i(l')$.

Now, for each $s_i^* \in \mathbf{P}_i^*(l)$ we form the i -tree with top $(R_{s_i^*}, c(\omega_{s_i^*}))$ consisting of all tiles in $\mathbf{P}_i(l)$ which are $\lesssim s_i^*$. We have just seen that these trees partition $\mathbf{P}_i(l)$ and that tiles in distinct trees are not comparable under \lesssim . Call the set of these trees $\mathcal{F}_i(l)$. It is easy to see that

$$(16) \quad N_{\mathcal{F}_i(l)}(\vec{x}) \leq N_{\mathcal{F}_i}(\vec{x}) \quad \text{for } \vec{x} \in \mathbb{R}^2.$$

We now use these trees to build our first generation of vector trees. For a moment we will abuse notation and for two $s, s' \in \mathbf{P}$ we will write $s \lesssim s'$ if $s_i \lesssim s'_i$ for each $1 \leq i \leq n$. Note that the extension of \lesssim from tiles to the multitiles in \mathbf{P} remains a relation of order.

For each $\vec{l} := (l_1, \dots, l_n)$ with $l_1, \dots, l_n \geq 0$ denote

$$\mathbf{P}(\vec{l}) := \{s \in \mathbf{P} : s_i \in \mathbf{P}_i(l_i) \text{ for each } 1 \leq i \leq n\}.$$

The selection process goes as follows. Fix \vec{l} . Find a maximal (with respect to \lesssim) $s \in \mathbf{P}(\vec{l})$; then construct the vector tree $\vec{\mathbf{T}}(s)$ with top $(R_s, c(\omega_s))$ consisting of all $s' \in \mathbf{P}(\vec{l})$ such that $s' \lesssim s$. Then eliminate all multitiles in $\vec{\mathbf{T}}(s)$ from $\mathbf{P}(\vec{l})$, and restart the selection, with this new value for $\mathbf{P}(\vec{l})$. When no such vector tree remains to be selected, the value of $\mathbf{P}(\vec{l})$ will be \emptyset . Denote by $\mathcal{F}(\vec{l})$ the family of the vector trees selected at this stage, and by $\mathcal{F}^{**}(\vec{l}) \subset \mathcal{F}(\vec{l})$ those vector trees that consist of at least two multitiles (that is, in addition to their top, they also contain a multitile with a scale distinct from the scale of the top). The vector trees in $\mathcal{F}(\vec{l}) \setminus \mathcal{F}^{**}(\vec{l})$ will be reshuffled later, so we will ignore them for the moment.

We will first show how to control the counting function $N_{\mathcal{F}^{**}(\vec{l})}$ in terms of each $N_{\mathcal{F}_i(l_i)}$. For each $\vec{\mathbf{T}} \in \mathcal{F}(\vec{l})$ we denote by \mathbf{T}_i the projection of $\vec{\mathbf{T}}$ onto \mathbf{S}_i (and this is an i -tree). For each $\vec{x} \in \mathbb{R}^d$, denote by $\mathcal{F}^{**}(\vec{l}, \vec{x})$ the set of all vector trees $\vec{\mathbf{T}} \in \mathcal{F}^{**}(\vec{l})$ such that $\vec{x} \in R_{\vec{\mathbf{T}}}$. A similar definition holds for $\mathcal{F}_i(l_i, \vec{x})$.

Let us first make two easy observations. On the one hand, note that for each $\vec{\mathbf{T}} \in \mathcal{F}(\vec{l})$, all the tiles of \mathbf{T}_i are contained in a unique tree from the family $\mathcal{F}_i(l_i)$ (this follows from an earlier observation, and from the fact that the tiles in \mathbf{T}_i are pairwise comparable under \lesssim). We will refer to this tree as the i -th *standard projection* of $\vec{\mathbf{T}}$. On the other hand, for a fixed \vec{x} and some $\vec{\mathbf{T}} \in \mathcal{F}^{**}(\vec{l}, \vec{x})$, if for each $1 \leq i \leq n$ we know the i -th standard projection of $\vec{\mathbf{T}}$, we will also automatically

know $\vec{\mathbf{T}}$ (this follows from the maximality involved in the selection of $\vec{\mathbf{T}}$). We will see in [Proposition 4.4](#) that more is true, namely that knowledge of just m of the standard projections suffices to determine $\vec{\mathbf{T}}$.

Lemma 4.2 [[Katz and Tao 1999](#)]. *Let X and A be finite sets and let $g : X \rightarrow A$ be a function. Then*

$$|\{(x_1, x_2) \in X^2 : g(x_1) = g(x_2)\}| \geq |X|^2/|A|.$$

We will apply this lemma with $X = \mathcal{F}^{**}(\vec{l}, \vec{x})$, $A = \mathcal{F}_1(l_1, \vec{x})$ and $g(\vec{\mathbf{T}}) := f_1(\vec{\mathbf{T}})$ the first standard projection of $\vec{\mathbf{T}}$ (for later use, we extend this definition to all f_i , $1 \leq i \leq n$). We get that

$$(17) \quad |\{(\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2) \in X^2 : f_1(\vec{\mathbf{T}}_1) = f_1(\vec{\mathbf{T}}_2)\}| \geq (N_{\mathcal{F}^{**}(\vec{l})}(\vec{x}))^2/N_{\mathcal{F}_1(l_1)}(\vec{x}).$$

Next, we will estimate from above the size of the set

$$\mathcal{H} := \{(\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2) \in X^2 : f_1(\vec{\mathbf{T}}_1) = f_1(\vec{\mathbf{T}}_2)\}.$$

In particular, we will show that the function

$$H : \mathcal{H} \rightarrow \prod_{j=2}^k \mathcal{F}_j(l_j, \vec{x}),$$

$$(\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2) \mapsto (f_2(\vec{\mathbf{T}}_1), \dots, f_m(\vec{\mathbf{T}}_1), f_{m+1}(\vec{\mathbf{T}}_2), \dots, f_k(\vec{\mathbf{T}}_2))$$

is injective (recall that $k/d < n/2$, thus $k < n$, and the functions f_i will make sense for each $1 \leq i \leq k$). This fact combined with (17) will lead to the desired estimate

$$(18) \quad N_{\mathcal{F}^{**}(\vec{l})}(\vec{x}) \leq \left(\prod_{j=1}^k N_{\mathcal{F}_j(l_j)}(\vec{x}) \right)^{1/2}.$$

Note that the sum of the exponents on the right hand side equals the rank $k/2$ (and this is the best one can do). It will be crucial that this number is less than $n/2$. Similarly, in the case of general d , one can arrange things such that the sum of the exponents will be⁶ less than $n/2$. This is explained in the next section.

By using this, (16) and the fact that for each $\vec{x} \in \mathbb{R}^2$,

$$(19) \quad N_{\mathcal{F}_i(l_i)}(\vec{x}) = 0 \quad \text{if } 2^i > N_{\mathcal{F}_i}(\vec{x}),$$

⁶It seems likely that one can achieve an inequality where the sum of the exponents is k/d . However, we will content ourselves with a sum barely smaller than $n/2$.

we find that⁷ for each $\epsilon > 0$

$$N_{\mathcal{F}^{**}}(\vec{x}) \lesssim_{\epsilon} \left(\prod_{j=1}^k N_{\mathcal{F}_j}(\vec{x}) \right)^{1/2+\epsilon}.$$

Here and in the following $\mathcal{F}^{**} := \bigcup_{\vec{l}} \mathcal{F}^{**}(\vec{l})$. Also, because there was nothing special about working with indices $1, \dots, k$, we can permute them and get similar inequalities. Combining this with the fact that $k < n$, we get

$$(20) \quad N_{\mathcal{F}^{**}}(\vec{x}) \lesssim \left(\prod_{j=1}^n N_{\mathcal{F}_j}(\vec{x}) \right)^{\delta},$$

for some $\delta < 1/2$, depending only on k and n . The precise value of δ will not be important, what matters for later applications is that it is strictly smaller than $1/2$.

Proposition 4.3. *The function H is injective.*

Proof. Assume that $(\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2), (\vec{\mathbf{U}}_1, \vec{\mathbf{U}}_2) \in \mathcal{H}$ have the same value under H . We will prove that $(\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2) = (\vec{\mathbf{U}}_1, \vec{\mathbf{U}}_2)$. Let s, s', p, p' be the top multitiles of the vector trees $\vec{\mathbf{T}}_1, \vec{\mathbf{T}}_2, \vec{\mathbf{U}}_1, \vec{\mathbf{U}}_2$ and let D_{\max} be as in (14). Our hypothesis will easily imply that

$$(21) \quad D_{\max} \leq C_0 \max\{\text{diam}(\tilde{\omega}_s), \text{diam}(\tilde{\omega}_{s'}), \text{diam}(\tilde{\omega}_p), \text{diam}(\tilde{\omega}_{p'})\},$$

and by (r7) we get

$$(22) \quad \text{dist}(\omega_s, \omega_p), \text{dist}(\omega_{s'}, \omega_{p'}) \\ \leq C_0^6 \max\{\text{diam}(\tilde{\omega}_s), \text{diam}(\tilde{\omega}_{s'}), \text{diam}(\tilde{\omega}_p), \text{diam}(\tilde{\omega}_{p'})\}.$$

We choose one of the four vector trees for which its top multitile has the largest scale of the frequency component. To fix notation, we can assume without any loss of generality that this vector tree is $\vec{\mathbf{T}}_1$. Since $\vec{\mathbf{T}}_1$ has at least two multitiles with distinct scales, we can find $t \in \vec{\mathbf{T}}_1$ such that $\text{diam}(\tilde{\omega}_t) > \text{diam}(\tilde{\omega}_s)$. This observation combined with (22), (r5), and the fact that $C_0, C_1 \ll C_2$ implies that $\tilde{\omega}_p \subset \tilde{\omega}_t$. Since $R_t \subset R_s$ and since $R_s \subset R_p$ (this being a consequence of the facts that $\vec{x} \in R_s \cap R_p$ and that the scale of R_p is larger than the scale of R_s), it follows that $t \lesssim p$. But we also know that $t \lesssim s$, and thus $\vec{\mathbf{T}}_1$ and $\vec{\mathbf{U}}_1$ will share all standard projections. This forces $\vec{\mathbf{T}}_1 = \vec{\mathbf{U}}_1$. This will in turn imply that $\vec{\mathbf{T}}_2$ and $\vec{\mathbf{U}}_2$ share at least m standard projections (corresponding to the indices $j \in \{1, m+1, m+2, \dots, k\}$; recall that $2m-1 \leq k$). The fact that $\vec{\mathbf{T}}_2 = \vec{\mathbf{U}}_2$ will follow from the following proposition, which is somewhat reminiscent of the rank property (iv). \square

Proposition 4.4. *Let $\vec{\mathbf{T}} \in \mathcal{F}^{**}(\vec{l}, \vec{x})$. If we know the value of $f_i(\vec{\mathbf{T}})$ for m of the n values of i , then we know $\vec{\mathbf{T}}$.*

⁷The extra ϵ exponent hides a logarithmic gain.

Proof. Let $\vec{\mathbf{T}}, \vec{\mathbf{U}} \in \mathcal{F}^{**}(\vec{l}, \vec{x})$ such that $f_i(\vec{\mathbf{T}}) = f_i(\vec{\mathbf{U}})$ for each $i \in \{1, \dots, m\}$. Follow exactly the same approach as in the proof of [Proposition 4.3](#), this time applied to the pairs $(\vec{\mathbf{T}}, \vec{\mathbf{T}})$ and $(\vec{\mathbf{T}}, \vec{\mathbf{U}})$. \square

The proof of [\(18\)](#) is now complete, and the first stage of our construction ends.

In the second stage, we will have to deal with the sets of vector trees $\mathcal{F}^*(\vec{l}) := \mathcal{F}(\vec{l}) \setminus \mathcal{F}^{**}(\vec{l})$ each of which consists of just one multitile. We will denote by $\bar{\mathbf{P}}(\vec{l})$ the set of all these multitiles. The additional key property that these multitiles will obey is

$$(23) \quad s \neq s' \in \bar{\mathbf{P}}(\vec{l}) \quad \text{implies} \quad s, s' \text{ are not comparable under } \lesssim.$$

This is a consequence of the maximality involved in the selection of the vector trees from the previous stage, and will turn out to be crucial in proving the analog of [Proposition 4.4](#); see [Proposition 4.5](#) below.

The plan is to reshuffle $\bar{\mathbf{P}}(\vec{l})$ into convenient vector trees. To achieve this, we will first split each $\bar{\mathbf{P}}_i(\vec{l})$ into overlapping trees (interestingly, from now on, the lacunary trees will not play any role). For each $r \in \mathbb{Z}$ with $2^{-r} \leq \text{size}_i(\bar{\mathbf{P}}_i(\vec{l}))$, let

$$(24) \quad \bar{\mathbf{P}}_i^{(r)}(\vec{l}) := \{s_i \in \bar{\mathbf{P}}_i(\vec{l}) : 2^{-r} \leq |R_{s_i}|^{-1/2} |\langle F_i, \phi_{s_i} \rangle| < 2^{-r+1}\}.$$

Note that

$$\bar{\mathbf{P}}_i(\vec{l}) = \bigcup_{2^{-r} \leq \text{size}_i(\bar{\mathbf{P}}_i(\vec{l}))} \bar{\mathbf{P}}_i^{(r)}(\vec{l}).$$

We next use a greedy selection algorithm as before to split each $\bar{\mathbf{P}}_i^{(r)}(\vec{l})$ into a set $\mathcal{F}_i(r, \vec{l})$ of disjoint i -overlapping trees with top tiles pairwise not comparable with respect to \leq . This implies via a classical TT^* argument (see [[Muscalu et al. 2002](#), Corollary 7.6] for example) that

$$(25) \quad \|N_{\mathcal{F}_i(r, \vec{l})}(x)\|_{\text{BMO}} \lesssim 2^{2r} \left(\sup_{s_i \in \bar{\mathbf{P}}_i^{(r)}(\vec{l})} \inf_{\vec{x} \in R_{s_i}} M_2 F_i(\vec{x}) \right)^2.$$

Also, [\(24\)](#) will imply that $N_{\mathcal{F}_i(r, \vec{l})}$ is supported in the set $\{\vec{x} : M_2 F_i(\vec{x}) \gtrsim 2^{-r}\}$. Combining this with [\(25\)](#) and the John–Nirenberg inequality we get

$$(26) \quad \|N_{\mathcal{F}_i(r, \vec{l})}(x)\|_t \lesssim 2^{2r} \left(\sup_{s_i \in \bar{\mathbf{P}}_i^{(r)}(\vec{l})} \inf_{\vec{x} \in R_{s_i}} M_2 F_i(\vec{x}) \right)^2 (2^{p_i r} \|F_i\|_{p_i}^{p_i})^{1/t} \text{ for } 1 \leq t < \infty.$$

We then apply Fefferman's trick again, as in the first stage of our construction, this time however with respect to $\mathcal{F}_i(r, \vec{l})$ and \leq (rather than \lesssim).

For each $q \geq 0$, let

$$\bar{\mathbf{P}}_i^{(r)}(\vec{l}, q) := \{s_i \in \bar{\mathbf{P}}_i^{(r)}(\vec{l}) : 2^q \leq |\{\mathbf{T} \in \mathcal{F}_i(r, \vec{l}) : \zeta_{\mathbf{T}} \in \omega_{s_i}, R_{s_i} \subseteq R_{\mathbf{T}}\}| < 2^{q+1}\}.$$

Note that $(\bar{\mathbf{P}}_i^{(r)}(\vec{l}, q))_{q \geq 0}$ forms a partition of $\bar{\mathbf{P}}_i^{(r)}(\vec{l})$. Next, we organize each $\bar{\mathbf{P}}_i^{(r)}(\vec{l}, q)$ into i -overlapping trees with top tiles. More precisely, consider the set $\bar{\mathbf{P}}_i^{(r,*)}(\vec{l}, q) \subset \bar{\mathbf{P}}_i^{(r)}(\vec{l}, q)$ of all tiles maximal with respect to the order relation \leq .

It follows as before that the tiles in $\bar{\mathbf{P}}_i^{(r,*)}(\vec{l}, q)$ are pairwise not comparable under \leq and that for each tile $s_i \in \bar{\mathbf{P}}_i^{(r)}(\vec{l}, q)$ there is a unique tile $s_i^* \in \bar{\mathbf{P}}_i^{(r,*)}(\vec{l}, q)$ such that $s_i \leq s_i^*$.

For each $s_i^* \in \bar{\mathbf{P}}_i^{(r,*)}(\vec{l}, q)$ we form the i -overlapping tree with top $(R_{s_i^*}, c(\omega_{s_i^*}))$ consisting of all tiles in $\bar{\mathbf{P}}_i^{(r)}(\vec{l}, q)$ that are $\leq s_i^*$. As before, these trees partition $\bar{\mathbf{P}}_i^{(r)}(\vec{l}, q)$ and tiles in distinct trees are not comparable under \leq . Call the set of these trees $\mathcal{F}_i(r, \vec{l}, q)$, and note that

$$(27) \quad N_{\mathcal{F}_i(r, \vec{l}, q)}(\vec{x}) \leq N_{\mathcal{F}_i(r, \vec{l})}(\vec{x}) \quad \text{for } \vec{x} \in \mathbb{R}^2.$$

For all vectors \vec{q} and \vec{r} , let $\mathcal{F}(\vec{r}, \vec{l}, \vec{q})$ consist of all the multitiles s such that $s_i \in \bar{\mathbf{P}}_i^{(r_i)}(\vec{l}, q_i)$ for each i . This consideration is motivated by (23), which implies that each vector tree in $\bar{\mathbf{P}}(\vec{l})$ can have only one multitile. We will prove that

$$(28) \quad N_{\mathcal{F}(\vec{r}, \vec{l}, \vec{q})}(\vec{x}) \leq \left(\prod_{i=1}^k N_{\mathcal{F}_i(r_i, \vec{l}, q_i)}(\vec{x}) \right)^{1/2} \quad \text{for } \vec{x} \in \mathbb{R}^2.$$

Using this, (27), and that $N_{\mathcal{F}_i(r_i, \vec{l}, q_i)}(\vec{x}) = 0$ for each $\vec{x} \in \mathbb{R}^2$ if $2^{q_i} > N_{\mathcal{F}_i(r_i, \vec{l})}(\vec{x})$, we find that for each $\epsilon > 0$,

$$N_{\mathcal{F}(\vec{r}, \vec{l})}(\vec{x}) \lesssim_{\epsilon} \left(\prod_{j=1}^k N_{\mathcal{F}_j(r_j, \vec{l})}(\vec{x}) \right)^{1/2+\epsilon} \quad \text{for } \vec{x} \in \mathbb{R}^2.$$

Here and in the following, $\mathcal{F}(\vec{r}, \vec{l}) := \bigcup_{\vec{q}} \mathcal{F}(\vec{r}, \vec{l}, \vec{q})$. Again, by permuting indices and since $k < n$, we get

$$(29) \quad N_{\mathcal{F}(\vec{r}, \vec{l})}(\vec{x}) \lesssim \left(\prod_{i=1}^n N_{\mathcal{F}_i(r_i, \vec{l})}(\vec{x}) \right)^{\delta} \quad \text{for } \vec{x} \in \mathbb{R}^2 \text{ and some } \delta < 1/2.$$

We mention that (29) will later be used in conjunction with (26) and with the upper bound (19) on the l_i .

We next prove (28). The argument follows exactly the same scheme as in the previous stage of the decomposition: We denote by $X := \mathcal{F}(\vec{r}, \vec{l}, \vec{q}, \vec{x})$ the set of all vector trees $\vec{\mathbf{T}} \in \mathcal{F}(\vec{r}, \vec{l}, \vec{q})$ with $\vec{x} \in R_{\vec{\mathbf{T}}}$. For a vector tree $\vec{\mathbf{T}}$ we will now denote by $f_i(\vec{\mathbf{T}})$ the tree from the set $\mathcal{F}(r_i, \vec{l}, q_i, \vec{x})$ that contains the tiles s_i , for $s \in \vec{\mathbf{T}}$. It will follow as in the previous stage that these projections are well defined. Define

\mathcal{H} and H as before; more precisely,

$$H : \mathcal{H} \rightarrow \prod_{j=2}^k \mathcal{F}_j(r_j, \vec{l}, q_j, \vec{x}).$$

Proposition 4.5. *The function H is injective.*

Proof. The proof of the injectivity of H follows the same lines as the one of [Proposition 4.3](#), with only one key difference. Assume that

$$(\vec{\mathbf{T}}_1 := \{s\}, \vec{\mathbf{T}}_2 := \{s'\}), (\vec{\mathbf{U}}_1 := \{p\}, \vec{\mathbf{U}}_2 := \{p'\}) \in \mathcal{H}$$

have the same value under H . We will prove that $(s, s') = (p, p')$. Let D_{\max} be as in [\(14\)](#). Our hypothesis will easily imply that

$$D_{\max} \leq C_0 \max\{\text{diam}(\omega_s), \text{diam}(\omega_{s'}), \text{diam}(\omega_p), \text{diam}(\omega_{p'})\}.$$

This estimate is stronger than the one in [\(21\)](#), in that on the right hand side here we have the diameters of the cubes ω , rather than those of the cubes $\tilde{\omega}$. This is due to the fact that the standard projections now reflect positioning with respect to \leq , rather than \lesssim . By [\(r7\)](#) we get

$$(30) \quad \text{dist}(\omega_s, \omega_p), \text{dist}(\omega_{s'}, \omega_{p'}) \\ \leq C_0^6 \max\{\text{diam}(\omega_s), \text{diam}(\omega_{s'}), \text{diam}(\omega_p), \text{diam}(\omega_{p'})\}.$$

We choose one of the four vector trees whose frequency component has the largest scale. To fix notation, we can assume without any loss of generality that this vector tree is s . Observe that [\(30\)](#) shows that $\tilde{\omega}_s$ and $\tilde{\omega}_p$ must intersect if $C_0 \ll C_1$. But then [\(r5\)](#) implies that $\tilde{\omega}_p \subseteq \tilde{\omega}_s$. This combined with the fact that $\vec{x} \in R_s \cap R_p \neq \emptyset$ implies that $s \lesssim p$. From [\(23\)](#) we immediately get that $s = p$. This in turn implies that s' and p' share at least m standard projections. This is equivalent to saying that $s'_i \leq p'_i$ (or vice versa), for at least m values of i . From [\(r2\)](#) we get that $s' \lesssim p'$ (or vice versa). A final invocation of [\(23\)](#) concludes that $s' = p'$. \square

4C. The case $d > 2$. The case of arbitrary d follows by considerations very similar to the ones involved in the case $d = 2$. The rank properties (very much in the spirit of [\(r7\)](#)) that will be needed throughout the proof will not be stated explicitly this time, but will rather become clear from the nondegeneracy assumptions on Γ' that will be stated in each case. We briefly sketch the details.

Recall we are under the assumption $m \geq 2$. We can in addition assume that $m \geq n/2$ (and thus $k/d < m$). This is because if $m < n/2$, then, as explained earlier, one can crudely treat the operator as having integral rank m , and apply the methods from [\[Muscalu et al. 2002\]](#) (or alternatively, the approach described in [Section 4D](#)).

We will need the general case of this combinatorial lemma:

Lemma 4.6 [Katz and Tao 1999]. *Let X and A_1, \dots, A_{d-1} be finite sets and for each $1 \leq i \leq d-1$, let $g_i : X \rightarrow A_i$ be a function. Then*

$$|\{(x_1, \dots, x_d) \in X^d : g_i(x_i) = g_i(x_{i+1}) \text{ for all } 1 \leq i \leq d-1\}| \geq \frac{|X|^d}{\prod_{i=1}^{d-1} |A_i|}.$$

The two stages of the construction are the same as in the case $d = 2$, but the choice of the function H involves some modifications. As in the case $d = 2$, H will have the same formula in both stages of the reshuffling process, and proving its injectivity will involve very similar ideas. Thus, to fix notation, we only sketch the argument corresponding to the first stage. Take $X = \mathcal{F}^{**}(\vec{l}, \vec{x})$. Note that $m \leq n+1$ (since $m \leq (n+1)/2$ and $n \geq 3$), so the sets $\mathcal{F}_j(l_j, \vec{x})$ are defined for each $1 \leq j \leq m+1$. We distinguish three separate cases and will address each of them below.

Case 1: n is even. It follows that $m \leq n/2$. Without loss of generality we can assume that

$$(31) \quad \max\{|\mathcal{F}_j(l_j, \vec{x})| : 1 \leq j \leq m\} = |\mathcal{F}_1(l_1, \vec{x})| \leq \min\{|\mathcal{F}_j(l_j, \vec{x})| : m+1 \leq j \leq n\}.$$

Apply [Lemma 4.6](#) with $A_j = \mathcal{F}_1(l_1, \vec{x})$ and $g_j = f_1$. Also, define

$$\mathcal{H} := \{(\vec{\mathbf{T}}_1, \dots, \vec{\mathbf{T}}_d) \in X^d : g_j(\vec{\mathbf{T}}_j) = g_j(\vec{\mathbf{T}}_{j+1}), 1 \leq j \leq d-1\},$$

$$H : \mathcal{H} \rightarrow \prod_j B_j,$$

$$(\vec{\mathbf{T}}_1, \dots, \vec{\mathbf{T}}_d) \mapsto (f_2(\vec{\mathbf{T}}_1), f_3(\vec{\mathbf{T}}_1), \dots, f_m(\vec{\mathbf{T}}_1), f_2(\vec{\mathbf{T}}_2), f_3(\vec{\mathbf{T}}_2), \dots, f_m(\vec{\mathbf{T}}_2), \dots, f_2(\vec{\mathbf{T}}_d), f_3(\vec{\mathbf{T}}_d), \dots, f_m(\vec{\mathbf{T}}_d)).$$

where each B_j equals one of the sets $\mathcal{F}_i(l_i, \vec{x})$. We briefly comment on this construction. It is one of many one can do, and while the nondegeneracy requirements to make a particular choice of H injective will depend on H , they are achieved for generic Γ' . For example, we chose to assign entries of the form f_2, \dots, f_m to each of the trees $\vec{\mathbf{T}}_j$, but we could also have chosen to alternate between these entries and the entries f_{m+1}, \dots, f_{2m-1} , as we did in the case $d = 2$. The only restrictions are that the entries corresponding to each tree $\vec{\mathbf{T}}_j$ are pairwise distinct (otherwise redundancy occurs), and that the entries for $\vec{\mathbf{T}}_j$ are also distinct from $f_1(\vec{\mathbf{T}}_j)$.

Assume for the moment that H is injective. Combining this with [Lemma 4.6](#) we get that

$$|\mathcal{F}^{**}(\vec{l}, \vec{x})| \leq \left(|\mathcal{F}_1(l_1, \vec{x})|^{d-1} \prod_{j=2}^m |\mathcal{F}_j(l_j, \vec{x})| \right)^{1/d}.$$

Due to our assumption (31), one can easily check that this implies

$$(32) \quad |\mathcal{F}^{**}(\vec{l}, \vec{x})| \leq \left(\prod_{j=1}^n |\mathcal{F}_j(l_j, \vec{x})| \right)^\delta \quad \text{for some } \delta < 1/2,$$

which as explained earlier, is the desired inequality.

Let us now see why H is injective. Assume that $(\vec{T}_1, \dots, \vec{T}_d), (\vec{U}_1, \dots, \vec{U}_d) \in \mathcal{H}$ have the same value under H . The nondegeneracy condition that we need is that the system

$$\left\{ \begin{array}{l} G_1(\vec{\xi}^{(1)}) - G_1(\vec{\xi}^{(2)}) = v_1, \\ G_1(\vec{\xi}^{(2)}) - G_1(\vec{\xi}^{(3)}) = v_2, \\ \vdots \\ G_1(\vec{\xi}^{(d-1)}) - G_1(\vec{\xi}^{(d)}) = v_{d-1}, \\ G_2(\vec{\xi}^{(1)}) = v_{1,2}, \\ \vdots \\ G_m(\vec{\xi}^{(d)}) = v_{d,m} \end{array} \right.$$

has *at most one solution* for each choice of targets v_1, \dots, v_{d-1} and $v_{i,j}$ with $1 \leq i \leq d$ and $2 \leq j \leq m$. Note that there are $dm - 1$ \mathbb{R}^d -valued equations in d unknowns $\vec{\xi}^{(j)} \in \mathbb{R}^k$, and that we have $(dm - 1)d \geq dk$, since $m > k/d$. Thus this system is always overdetermined, and we require that the matrix associated with it has maximum rank dk . It is not too hard to check that this is achieved for a generic choice of Γ' . Actually, our choice of H is in such a way that the system consisting of the first k vector-valued equations above will generically give rise to a compatible system.

The injectivity of H now follows as in the case $d = 2$. First, there must be some i_0 such that either \vec{T}_{i_0} or \vec{U}_{i_0} contains a multitile with the frequency scale larger than or equal to the scales of all the multitiles from the trees \vec{T}_i, \vec{U}_i . As before, we get that $\vec{T}_{i_0} = \vec{U}_{i_0}$. The equality $\vec{T}_i = \vec{U}_i$ for the remaining indices i will follow from a “domino effect”. It first follows for the neighboring indices $i = i_0 - 1$ and/or $i = i_0 + 1$, using the fact that

$$f_1(\vec{T}_i) = f_1(\vec{T}_{i_0}) = f_1(\vec{U}_{i_0}) = f_1(\vec{U}_i) \quad \text{and} \quad f_j(\vec{T}_i) = f_j(\vec{U}_i)$$

(for $j \in \{2, 3, \dots, m\}$), and using Proposition 4.4. The domino effect continues until all indices are covered.

Case 2: n is odd. Since we have assumed that $m \geq n/2$, we have $m = (n + 1)/2$. In particular, whenever $i_1 \notin \{i_2, \dots, i_{2m-1}\}$ with $\{i_2, \dots, i_m\}$ pairwise distinct and

$\{i_{m+1}, \dots, i_{2m-1}\}$ pairwise distinct, the system

$$(33) \quad \left\{ \begin{array}{l} G_{i_1}(\vec{\xi}^{(1)}) - G_{i_1}(\vec{\xi}^{(2)}) = v_1, \\ G_{i_2}(\vec{\xi}^{(1)}) = v_{1,2}, \\ \vdots \\ G_{i_m}(\vec{\xi}^{(1)}) = v_{1,m}, \\ G_{i_{m+1}}(\vec{\xi}^{(2)}) = v_{2,m+1}, \\ \vdots \\ G_{i_{2m-1}}(\vec{\xi}^{(2)}) = v_{2,2m-1} \end{array} \right.$$

of $n = 2m - 1$ \mathbb{R}^d -valued equations in 2 unknowns $\vec{\xi}^{(1)}, \vec{\xi}^{(2)} \in \mathbb{R}^k$ will have at most one solution for a generic Γ' . This is one of the two nondegeneracy conditions that will be needed in both of the following two subcases. We will refer to the above system as a *two-scheme*.

Case 2a: n is odd and d is odd. Without loss of generality we can assume that

$$(34) \quad \max\{|\mathcal{F}_j(l_j, \vec{x})| : 1 \leq j \leq m + 1\} = |\mathcal{F}_2(l_2, \vec{x})| \\ \leq \min\{|\mathcal{F}_j(l_j, \vec{x})| : m + 2 \leq j \leq n\}.$$

Apply [Lemma 4.6](#) with $A_j = \mathcal{F}_1(l_1, \vec{x})$ and $g_j = f_1$ if j is odd, $A_j = \mathcal{F}_2(l_2, \vec{x})$ and $g_j = f_2$ if j is even. Also, define

$$\mathcal{H} := \{(\vec{\mathbf{T}}_1, \dots, \vec{\mathbf{T}}_d) \in X^d : g_j(\vec{\mathbf{T}}_j) = g_j(\vec{\mathbf{T}}_{j+1}), 1 \leq j \leq d - 1\}, \\ H : \mathcal{H} \rightarrow \prod_j B_j, \\ (\vec{\mathbf{T}}_1, \dots, \vec{\mathbf{T}}_d) \mapsto := (f_3(\vec{\mathbf{T}}_1), f_4(\vec{\mathbf{T}}_1), \dots, f_m(\vec{\mathbf{T}}_1), f_3(\vec{\mathbf{T}}_2), f_4(\vec{\mathbf{T}}_2), \dots, f_m(\vec{\mathbf{T}}_2), \\ \dots, f_3(\vec{\mathbf{T}}_d), f_4(\vec{\mathbf{T}}_d), \dots, f_m(\vec{\mathbf{T}}_d), \\ f_{m+1}(\vec{\mathbf{T}}_1), f_{m+1}(\vec{\mathbf{T}}_3), \dots, f_{m+1}(\vec{\mathbf{T}}_d)),$$

where each B_j equals one of the sets $\mathcal{F}_i(l_i, \vec{x})$. Note that H has $(m-2)d + (d+1)/2$ entries. If $m = 2$, then the first $(m-2)d$ entries are not present. Note also that the last line above contains $(d+1)/2$ entries of the form $f_{m+1}(\vec{\mathbf{T}}_i)$, for all possible odd indices i . Here is why we cannot use more than $(d+1)/2$ such entries. While more entries would certainly reinforce the injectivity of H , the application of the injectivity (combined with [Lemma 4.6](#)) would be inefficient, in that it would not lead to [\(32\)](#). On the other hand, our choice for H combined with the assumption [\(34\)](#) is easily seen to guarantee [\(32\)](#).

We chose to assign entries $f_{m+1}(\vec{\mathbf{T}}_i)$ to the odd indices i (as opposed to, say, the first $(d+1)/2$ indices) to allow for the domino effect, as explained below.

and \vec{U}_{i_0+1} share $m - 1$ standard projections, \vec{T}_{i_0+2} and \vec{U}_{i_0+2} share $m - 1$ standard projections, and also, recall that $f_2(\vec{T}_{i_0+1}) = f_2(\vec{T}_{i_0+2})$ and $f_2(\vec{U}_{i_0+1}) = f_2(\vec{U}_{i_0+2})$. The analysis of this two-scheme (essentially, a repeat of the argument from the case that $d = 2$) will imply that $\vec{T}_i = \vec{U}_i$ for $i \in \{i_0 + 1, i_0 + 2\}$.

If we allow combinations of these scenarios, it is easy to see that the domino effect will eventually prove that $\vec{T}_i = \vec{U}_i$ for all i .

Case 2b: n is odd and d is even. This is the most delicate case. We will use the same construction as in the previous subcase, but with $d + 1$ vector trees rather than d . More precisely, without loss of generality we can assume that

$$\max\{|\mathcal{F}_j(l_j, \vec{x})| : 1 \leq j \leq m + 1\} = |\mathcal{F}_2(l_2, \vec{x})| \leq \min\{|\mathcal{F}_j(l_j, \vec{x})| : m + 2 \leq j \leq n\}.$$

Apply [Lemma 4.6](#) (this time for $d + 1$ sets) with $A_j = \mathcal{F}_1(l_1, \vec{x})$ and $g_j = f_1$ if j is odd, $A_j = \mathcal{F}_2(l_2, \vec{x})$ and $g_j = f_2$ if j is even. Here $j \in \{1, \dots, d + 1\}$. Also, define

$$\begin{aligned} \mathcal{H} &:= \{(\vec{T}_1, \dots, \vec{T}_{d+1}) \in X^{d+1} : g_j(\vec{T}_j) = g_j(\vec{T}_{j+1}), 1 \leq j \leq d\}, \\ H : \mathcal{H} &\rightarrow \prod_j B_j, \\ (\vec{T}_1, \dots, \vec{T}_{d+1}) &\mapsto (f_3(\vec{T}_1), f_4(\vec{T}_1), \dots, f_m(\vec{T}_1), f_3(\vec{T}_2), f_4(\vec{T}_2), \dots, f_m(\vec{T}_2), \\ &\quad \dots, f_3(\vec{T}_{d+1}), f_4(\vec{T}_{d+1}), \dots, f_m(\vec{T}_{d+1}), \\ &\quad f_{m+1}(\vec{T}_1), f_{m+1}(\vec{T}_3), \dots, f_{m+1}(\vec{T}_{d+1})), \end{aligned}$$

where each B_j equals one of the sets $\mathcal{F}_i(l_i, \vec{x})$. In addition to requiring that the two-schemes [\(33\)](#) have at most one solution, we will also need that the system (a copy of [\(35\)](#) with $d \mapsto d + 1$)

$$\left\{ \begin{aligned} G_1(\vec{\xi}^{(1)}) - G_1(\vec{\xi}^{(2)}) &= v_1, \\ G_2(\vec{\xi}^{(2)}) - G_2(\vec{\xi}^{(3)}) &= v_2, \\ &\vdots \\ G_2(\vec{\xi}^{(d-1)}) - G_2(\vec{\xi}^{(d)}) &= v_{d-1}, \\ G_l(\vec{\xi}^{(j)}) &= v_{l,3} \quad \text{for } 3 \leq l \leq m, 1 \leq j \leq d + 1, \\ G_{m+1}(\vec{\xi}^{(1)}) &= v_{1,m+1}, \\ G_{m+1}(\vec{\xi}^{(3)}) &= v_{3,m+1}, \\ &\vdots \\ G_{m+1}(\vec{\xi}^{(d+1)}) &= v_{d+1,m+1} \end{aligned} \right.$$

has at most one solution. There are $(d + 1)(m - 1/2) - 1/2$ \mathbb{R}^d -valued equations in $d + 1$ variables from \mathbb{R}^k . The fact that the system is overdetermined, since

$$\left((d + 1)\left(m - \frac{1}{2}\right) - \frac{1}{2}\right)d \geq (d + 1)k,$$

is a consequence of the fact that $k/d < m - 1/2$.

The argument will then run as in the previous subcase. We leave details to the interested reader.

4D. The case $d = 1$. As advertised earlier, a simplified version of the combinatorial argument in the case $d = 2$ can also handle the case $d = 1$, and more generally, the case of arbitrary d and $m < n/2$. We present the argument in this generality, and thus assume $m < n/2$, rather than $d = 2$. We note again that the argument here reproves the main theorem in [Muscalu et al. 2002], in the locally L^2 case, without any appeal to induction.

We perform the same two stage decomposition, exactly as in the case $d = 2$. This time however the function H will have a simpler form. More precisely, in the first stage one considers functions of the form

$$H : \mathcal{F}^{**}(\vec{l}, \vec{x}) \rightarrow \prod_{j=1}^m \mathcal{F}_j(l_j, \vec{x}), \quad (\vec{\mathbf{T}}) \mapsto (f_1(\vec{\mathbf{T}}), \dots, f_m(\vec{\mathbf{T}})),$$

with an identical construction (up to notation) for the second stage. The injectivity of H will follow from Proposition 4.4 in the first stage, and from (r2) in the second stage.

The injectivity of H will in turn imply the desired estimates (20) and (29), since $m < n/2$.

Lemma 4.2 and Proposition 4.3 are no longer needed. Thus we need no further nondegeneracy conditions on Γ' , other than the one from [Muscalu et al. 2002] that Γ' is parametrizable over any k canonical variables.

5. Proof of Theorem 2.5

Assume $d = 2$. The argument for $d > 2$ would follow with no essential modification.

For each $\mathbf{P} \subset \mathbf{S}$ we will use the notation

$$\Lambda_{\mathbf{P}}(\vec{x})(F_1, \dots, F_n) := \sum_{s \in \mathbf{P}} |R_s|^{-n/2} \prod_{i=1}^n |\langle F_i, \phi_{s_i} \rangle| 1_{R_s}(\vec{x}).$$

By invoking interpolation and the dilation invariance of our operator, it suffices to prove that for each $2 < p_i \leq \infty$ with $1/p_1 + \dots + 1/p_n = 1/p$ and each $\|F_i\|_{p_i} = 1$,

$$(36) \quad \left| \{\vec{x} : \Lambda_{\mathbf{S}}(F_1, \dots, F_n)(\vec{x}) \gtrsim 1\} \right| \lesssim 1.$$

For the remaining part of the argument, the functions F_i will be fixed as above, and all sizes will be computed with respect to them.

Consider the exceptional set $E = \bigcup_{i=1}^n \{\vec{x} : M_2 F_i(\vec{x}) \geq 1\}$, and note that $|E| \lesssim 1$.

It now suffices to restrict attention in (36) to the set (which for simplicity will also be denoted with \mathbf{S}) of multitiles s that have the property that $R_s \cap E^c \neq \emptyset$.

Lemma 3.5 will now imply that $\text{size}_i(\mathbf{S}_i) \lesssim 1$.

Apply now Proposition 3.8 to the sets \mathbf{S}_i . We get that

$$(37) \quad \mathbf{S}_i := \bigcup_{2^{-k} \lesssim 1} \mathbf{S}_i^{(k)}, \quad \text{where } \text{size}_i(\mathbf{S}_i^{(k)}) \leq 2^{-k+1}$$

and each $\mathbf{S}_i^{(k)}$ is the (disjoint) union of a family $\mathcal{F}_i^{(k)}$ of trees such that

$$(38) \quad \|N_{\mathcal{F}_i^{(k)}}\|_q \lesssim 2^{2k} 2^{kp_i/q} \quad \text{for } 1 \leq q < \infty.$$

An immediate consequence of (38) (choose q large enough) is that

$$|\{\vec{x} : N_{\mathcal{F}_i^{(k)}}(\vec{x}) > 2^{4k}\}| \lesssim 2^{-10k}.$$

By eliminating another exceptional set of measure $O(1)$, it thus suffices to further restrict attention in (36) to those \vec{x} that satisfy

$$(39) \quad N_{\mathcal{F}_i^{(k)}}(\vec{x}) \leq 2^{4k} \quad \text{for each } i \text{ and each } 2^{-k} \lesssim 1.$$

We fix some k_i for each i , and denote by $\mathbf{S}(\vec{k})$ the set of all multitiles s with $s_i \in \mathbf{S}_i^{(k_i)}$ for each i . We follow the procedure described in the case $d = 2$, applied to $\mathbf{P} := \mathbf{S}(\vec{k})$ and $\mathcal{F}_i := \mathcal{F}_i^{(k_i)}$. The set $\mathbf{S}(\vec{k})$ will be the union of three families of vector trees: $\mathcal{F}_{\vec{k}}^{***}$ from the first stage of the construction and

$$(40) \quad \begin{aligned} \mathcal{F}_{\vec{k}}^* &:= \bigcup_{\vec{l}: 1 \leq 2^{l_i} \leq 2^{4k_i}} \bigcup_{\vec{r}: 2^{-r_i} \leq 2^{-k_i+1}} \mathcal{F}_{\vec{k}}(\vec{r}, \vec{l}), \\ \mathcal{F}_{\vec{k}}^- &:= \bigcup_{\vec{l}: 2^{l_i} > 2^{4k_i}} \bigcup_{\vec{r}: 2^{-r_i} \leq 2^{-k_i+1}} \mathcal{F}_{\vec{k}}(\vec{r}, \vec{l}) \end{aligned}$$

from the second stage of the construction. Due (19) and (39), the family $\mathcal{F}_{\vec{k}}^-$ can be ignored, since it will not contribute to $\Lambda_{\mathbf{S}(\vec{k})}$.

Let t be a sufficiently large number. To evaluate $\|N_{\mathcal{F}_{\vec{k}}^{***}}\|_t$, we will invoke (20), (38) and Hölder's inequality, that

$$\|N_{\mathcal{F}_{\vec{k}}^{***}}\|_t \lesssim \prod_{i=1}^n 2^{k_i(2\delta + p_i/(nt))}.$$

Combining this with [Lemma 3.6](#) and the estimate $\text{size}_i(\mathbf{S}_i^{(k)}) \leq 2^{-k+1}$ in [\(37\)](#) for the size, we get

$$\|\Lambda_{\mathcal{F}_k^{**}}\|_t \lesssim \prod_{i=1}^n 2^{k_i(2\delta+p_i/(nt)-1)}.$$

Since $\delta < 1/2$, it follows that

$$\left| \{\vec{x} : \Lambda_{\cup_{k:2^{-k} \lesssim 1} \mathcal{F}_k^{**}}(F_1, \dots, F_n)(\vec{x}) \gtrsim 1\} \right| \lesssim 1.$$

It now remains to evaluate the contribution coming from the vector trees in [\(40\)](#). Fix \vec{r} and \vec{l} . As before, by using [\(26\)](#), [\(29\)](#) and Hölder’s inequality, we get

$$\|N_{\mathcal{F}_k(\vec{r}, \vec{l})}\|_t \lesssim \prod_{i=1}^n 2^{r_i(2\delta+p_i/(nt))}.$$

Combining this with [Lemma 3.6](#) and estimate [\(24\)](#) for the size, we get

$$\|\Lambda_{\mathcal{F}_k(\vec{r}, \vec{l})}\|_t \lesssim \prod_{i=1}^n 2^{r_i(2\delta+p_i/(nt)-1)}.$$

Summing this first over $2^{-r_i} \leq 2^{-k_i+1}$, then over $0 \leq l_i \leq 4k_i$ and finally over $2^{-k_i} \lesssim 1$, we get

$$\left| \{\vec{x} : \Lambda_{\cup_k \mathcal{F}_k^{**}}(F_1, \dots, F_n)(\vec{x}) \gtrsim 1\} \right| \lesssim 1.$$

This finishes the argument.

Acknowledgments

We thank Tamara Kucherenko and Camil Muscalu for helpful discussions.

References

[Demeter and Thiele 2010] C. Demeter and C. Thiele, “On the two dimensional bilinear Hilbert transform”, *Amer. J. Math.* **132**:1 (2010), 201–256. [Zbl 05673321](#)

[Fefferman 1973] C. Fefferman, “Pointwise convergence of Fourier series”, *Ann. of Math.* (2) **98** (1973), 551–571. [MR 49 #5676](#) [Zbl 0268.42009](#)

[Gowers and Wolf 2010] W. T. Gowers and J. Wolf, “The true complexity of a system of linear equations”, *Proc. Lond. Math. Soc.* (3) **100**:1 (2010), 155–176. [MR 2578471](#) [Zbl 05658776](#)

[Grafakos and Li 2004] L. Grafakos and X. Li, “Uniform bounds for the bilinear Hilbert transforms. I”, *Ann. of Math.* (2) **159**:3 (2004), 889–933. [MR 2006e:42011](#) [Zbl 1071.44004](#)

[Katz and Tao 1999] N. H. Katz and T. Tao, “Bounds on arithmetic projections, and applications to the Kakeya conjecture”, *Math. Res. Lett.* **6** (1999), 625–630. [MR 2000m:28006](#) [Zbl 0980.42013](#)

[Lacey and Thiele 1997] M. Lacey and C. Thiele, “ L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$ ”, *Ann. of Math.* (2) **146**:3 (1997), 693–724. [MR 99b:42014](#) [Zbl 0914.46034](#)

- [Lacey and Thiele 1999] M. Lacey and C. Thiele, “On Calderón’s conjecture”, *Ann. of Math. (2)* **149**:2 (1999), 475–496. [MR 2000d:42003](#) [Zbl 0934.42012](#)
- [Muscalu et al. 2002] C. Muscalu, T. Tao, and C. Thiele, “Multi-linear operators given by singular multipliers”, *J. Amer. Math. Soc.* **15**:2 (2002), 469–496. [MR 2003b:42017](#) [Zbl 0994.42015](#)
- [Muscalu et al. 2004] C. Muscalu, T. Tao, and C. Thiele, “ L^p estimates for the biest, II: The Fourier case”, *Math. Ann.* **329**:3 (2004), 427–461. [MR 2005k:42054](#) [Zbl 1073.42010](#)

Received April 8, 2009.

CIPRIAN DEMETER
DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
831 EAST 3RD ST.
BLOOMINGTON IN 47405
UNITED STATES
demeterc@indiana.edu

MALABIKA PRAMANIK
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, BC V6T 1Z2
CANADA
malabika@math.ubc.ca

CHRISTOPH THIELE
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES CA 90095-1555
UNITED STATES
thiele@math.ucla.edu

PACIFIC JOURNAL OF MATHEMATICS

<http://www.pjmath.org>

Founded in 1951 by

E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

EDITORS

V. S. Varadarajan (Managing Editor)
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
pacific@math.ucla.edu

Vyjayanthi Chari
Department of Mathematics
University of California
Riverside, CA 92521-0135
chari@math.ucr.edu

Robert Finn
Department of Mathematics
Stanford University
Stanford, CA 94305-2125
finn@math.stanford.edu

Kefeng Liu
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
liu@math.ucla.edu

Darren Long
Department of Mathematics
University of California
Santa Barbara, CA 93106-3080
long@math.ucsb.edu

Jiang-Hua Lu
Department of Mathematics
The University of Hong Kong
Pokfulam Rd., Hong Kong
jhlu@maths.hku.hk

Alexander Merkurjev
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
merkurev@math.ucla.edu

Sorin Popa
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
popa@math.ucla.edu

Jie Qing
Department of Mathematics
University of California
Santa Cruz, CA 95064
qing@cats.ucsc.edu

Jonathan Rogawski
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
jonr@math.ucla.edu

PRODUCTION

pacific@math.berkeley.edu

Paulo Ney de Souza, Production Manager

Silvio Levy, Senior Production Editor

Matthew Cargo, Production Editor

SUPPORTING INSTITUTIONS

ACADEMIA SINICA, TAIPEI
CALIFORNIA INST. OF TECHNOLOGY
INST. DE MATEMÁTICA PURA E APLICADA
KEIO UNIVERSITY
MATH. SCIENCES RESEARCH INSTITUTE
NEW MEXICO STATE UNIV.
OREGON STATE UNIV.

STANFORD UNIVERSITY
UNIV. OF BRITISH COLUMBIA
UNIV. OF CALIFORNIA, BERKELEY
UNIV. OF CALIFORNIA, DAVIS
UNIV. OF CALIFORNIA, LOS ANGELES
UNIV. OF CALIFORNIA, RIVERSIDE
UNIV. OF CALIFORNIA, SAN DIEGO
UNIV. OF CALIF., SANTA BARBARA

UNIV. OF CALIF., SANTA CRUZ
UNIV. OF MONTANA
UNIV. OF OREGON
UNIV. OF SOUTHERN CALIFORNIA
UNIV. OF UTAH
UNIV. OF WASHINGTON
WASHINGTON STATE UNIVERSITY

These supporting institutions contribute to the cost of publication of this Journal, but they are not owners or publishers and have no responsibility for its contents or policies.

See inside back cover or www.pjmath.org for submission instructions.

The subscription price for 2010 is US \$420/year for the electronic version, and \$485/year for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. Prior back issues are obtainable from Periodicals Service Company, 11 Main Street, Germantown, NY 12526-5635. The Pacific Journal of Mathematics is indexed by [Mathematical Reviews](#), [Zentralblatt MATH](#), [PASCAL CNRS Index](#), [Referativnyi Zhurnal](#), [Current Mathematical Publications](#) and the [Science Citation Index](#).

The Pacific Journal of Mathematics (ISSN 0030-8730) at the University of California, c/o Department of Mathematics, 969 Evans Hall, Berkeley, CA 94720-3840, is published monthly except July and August. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer-review and production is managed by EditFLOW™ from Mathematical Sciences Publishers.

PUBLISHED BY PACIFIC JOURNAL OF MATHEMATICS

at the University of California, Berkeley 94720-3840

A NON-PROFIT CORPORATION

Typeset in L^AT_EX

Copyright ©2010 by Pacific Journal of Mathematics

PACIFIC JOURNAL OF MATHEMATICS

Volume 246 No. 2 June 2010

Geometric structures associated to a contact metric (κ, μ) -space	257
BENIAMINO CAPPELLETTI MONTANO and LUIGIA DI TERLIZZI	
Multilinear singular operators with fractional rank	293
CIPRIAN DEMETER, MALABIKA PRAMANIK and CHRISTOPH THIELE	
A new proof of Reifenberg's topological disc theorem	325
GUANGHAO HONG and LIHE WANG	
Global classical solutions to hyperbolic geometric flow on Riemann surfaces	333
FAGUI LIU and YUANZHANG ZHANG	
An extension of the quintuple product identity and its applications	345
ZHI-GUO LIU	
A generalization of the Pontryagin–Hill theorems to projective modules over Prüfer domains	391
JORGE MACÍAS-DÍAZ	
Elliptic pseudodifferential equations and Sobolev spaces over p -adic fields	407
J. J. RODRÍGUEZ-VEGA and W. A. ZÚÑIGA-GALINDO	
Absolutely isolated singularities of holomorphic maps of \mathbb{C}^n tangent to the identity	421
FENG RONG	
Pullbacks of Eisenstein series from $\mathrm{GU}(3, 3)$ and critical L -values for $\mathrm{GSp}(4) \times \mathrm{GL}(2)$	435
ABHISHEK SAHA	
Isomorphism invariants of restricted enveloping algebras	487
HAMID USEFI	
Spacelike S-Willmore spheres in Lorentzian space forms	495
PENG WANG	