

Lecture notes on Fourier Decoupling

Pavel Zorin-Kranich

January 24, 2020

University of Bonn, summer term 2019

This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

Contents

0	Introduction	3
0.1	Motivation	3
0.2	Space localization	3
0.3	Frequency localization	4
1	Multilinear Kakeya	4
2	Multilinear restriction	8
3	Decoupling for the paraboloid	10
3.1	Optimality of the decoupling inequality	12
3.2	Basic properties of the decoupling constant	12
3.2.1	Parabolic scaling	12
3.2.2	Larger Fourier support	13
3.2.3	Interpolation	13
3.3	Localization	14
3.3.1	Power weights	14
3.3.2	Reverse Hölder inequality	16
3.4	Linear versus multilinear decoupling	16
3.4.1	Transversality	16
3.4.2	Notation	17
3.4.3	Lower dimensional decoupling	18
3.4.4	Bourgain–Guth argument	18
3.5	Bourgain–Demeter iteration	20
3.5.1	L^2 orthogonality	20
3.5.2	Ball inflation	21
3.5.3	Proof of Theorem 3.1 for $p = 2\frac{d+1}{d}$	22
3.5.4	Proof of Theorem 3.1 for $p = 2\frac{d+2}{d}$	23
4	Local smoothing for the wave equation	25
4.1	Decoupling for the truncated cone	25
4.2	Estimates for the pieces of a solution	26
4.3	Examples and the local smoothing conjecture	29
5	Cone square function	30

6	Reverse square function estimate for the cone in \mathbb{R}^{2+1}	33
6.1	Lorentz scaling	34
6.2	Multilinear reverse square function estimate	34
6.2.1	Transversality	34
6.2.2	Bourgain–Guth argument	35
6.3	The $L^3(\mathbb{R}^{2+1}, \ell^2)$ estimate	36
6.3.1	Multilinear restriction	37
6.3.2	Orthogonality	37
6.4	The $L^p(\mathbb{R}^{2+1}, \ell^2)$ estimate	37
6.4.1	Multilinear restriction	37
6.4.2	Orthogonality	38
6.4.3	Scaling	39
6.4.4	Bootstrapping	39
7	Weighted Strichartz estimate for the free Schrödinger equation	40
7.1	Notation	40
7.2	Weighted estimate	40
7.3	Consequences for fractal measures	43
7.4	Lower bounds	43
7.4.1	Modulation of initial data	44
7.4.2	Many wave packets	44
7.4.3	One wave packet	44
7.4.4	Tensor products	45
8	Brascamp–Lieb inequalities	45
8.1	Geometric case	47
8.2	Gaussian-extremizable case	50
8.3	Necessity of BCCT condition	51
8.4	Sufficiency of BCCT condition for simple data	51
9	Decoupling for the moment curve	54
9.1	Vinogradov mean value	54
9.2	Statement of the main result	55
9.2.1	Consequences for Vinogradov mean values	55
9.2.2	Affine scaling	55
9.3	Transversality	56
9.3.1	Projections onto higher order tangential spaces	56
9.3.2	Verification of BCCT condition	57
9.4	Bourgain–Guth argument	58
9.5	Induction on scales	59
9.5.1	Linear decoupling	60
9.5.2	Bourgain–Guth argument	60
9.5.3	Ball inflation	60
9.5.4	Lower degree decoupling	61
9.5.5	Wrapping up the induction	61
A	Estimates for Weyl sums	63
B	Maximal operators along separated directions in \mathbb{R}^2	63
C	Córdoba sector square function on \mathbb{R}^2	67
	References	70

0 Introduction

In this course we will look at some aspects of Fourier restriction theory that are primarily motivated by applications to partial differential equations. Here “Fourier restriction” will mean that we consider functions whose Fourier support is concentrated near a submanifold. In order to motivate this topic we will present in this section a context in which such functions arise.

0.1 Motivation

Consider the initial value problem for the free Schrödinger equation on \mathbb{R}^d :

$$2\pi i \partial_t \psi = \Delta_x \psi, \quad \psi(x, 0) = g(x).$$

This is a linear PDE with constant coefficients, so we can represent the solution in terms of the Fourier transform. Taking the Fourier transform of the equation in the x variable we obtain

$$2\pi i \partial_t \mathcal{F}_x \psi = (2\pi i \xi)^2 \mathcal{F}_x \psi(\xi, t).$$

For each fixed $\xi \in \mathbb{R}^d$ this is an ordinary differential equation, and the solution is given by

$$\mathcal{F}_x \psi(\xi, t) = e(|\xi|^2 t) \mathcal{F}_x \psi(\xi, 0).$$

Here and later

$$e(t) := e^{2\pi i t}.$$

Taking the inverse Fourier transform in the x variable we obtain

$$\psi(x, t) = \int e(|\xi|^2 t + \xi x) \hat{g}(\xi) \, d\xi.$$

The solution ψ can be viewed as a tempered distribution on $\mathbb{R}^d \times \mathbb{R}$, and the above formula shows that its Fourier transform (simultaneously in x and t) variables is supported on a paraboloid. Similarly, the solution of the wave equation has Fourier support on a cone.

This provides the motivation for our study of functions with restricted Fourier support. The abstract results that we will see will eventually lead to a priori estimates for solutions of Schrödinger and wave equations.

0.2 Space localization

Let f be a function on \mathbb{R}^n whose Fourier transform is supported on a submanifold A . Suppose that we want to study the function f locally, say on a ball $B(x, R)$ of radius R . It would be unwise to truncate it to the ball, since this operation would completely destroy any control on the Fourier transform. One can instead use a smooth truncation. Let φ be a Schwartz function such that $\varphi \geq \mathbf{1}_{B(0,1)}$ and such that $\widehat{\varphi}$ has compact Fourier support. Then the function $\varphi_R^x := \varphi(R^{-1}(\cdot - x))$ is ≥ 1 on $B(0, R)$, while its Fourier support is contained in a ball of radius $O(R^{-1})$. It follows that

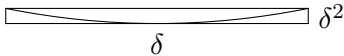
$$\widehat{\varphi_R^x f} = \widehat{\varphi_R^x} * \widehat{f}$$

is supported in a $O(R^{-1})$ -neighborhood of the support of f .

This space localization trick is one of the main reasons why we will have to consider functions with Fourier support near a submanifold.

0.3 Frequency localization

We will study functions with restricted Fourier support by decomposing them into pieces with even smaller Fourier support. Let us see how this such a function might look like. Let f be a function on \mathbb{R}^2 such that \widehat{f} is supported on a piece of parabola of length $\delta \leq 1$ not far from 0.



This piece of parabola is contained in a rectangular box of size $\delta \times \delta^2$. Let ψ be a smooth function adapted to this box that identically equals 1 on the box, so that $\widehat{f} = \widehat{f}\psi$. Then $f = f * \check{\psi}$, and $\check{\psi}$ is a smooth function adapted to the polar rectangle of our rectangular box, that is, a rectangle with same orientation and size $\delta^{-1} \times \delta^{-2}$. Thus, on a heuristical level, the function f is constant at scale δ^{-1} in the tangential direction of the parabola and constant at scale δ^{-2} in the orthogonal direction.

In order to most fully make use of this heuristic it makes sense to study the function f at spatial scale δ^{-2} . This is compatible with the space localization procedure outlined previously: the space localization at scale δ^{-2} leaves the Fourier support near our box of size $\delta \times \delta^2$ (more precisely, the side lengths only have to be enlarged by a constant factor).

So our plan will be to study solutions of the Schrödinger equation at scale R by decomposing them into pieces with Fourier support of size $R^{-1/2}$.

1 Multilinear Keakeya

By the discussion in the introduction, functions with Fourier support in boxes near a parabola (left picture) are morally constant on polar boxes (right picture). We will call boxes “polar” if they have the same orientation and the product of lengths of corresponding sides is 1 (this is morally a special case of “polar sets” in convex analysis).



Let us provisionally call two arcs of the parabola *transverse* if they are separated by more than their combined length. As one sees in the picture, in dimension 2 the interaction between two functions with transverse Fourier supports is relatively simple since they are basically constant in two different directions. In particular, L^p norms of their products can be morally computed by Fubini’s theorem in terms of L^p norms of the individual functions. This is a big improvement over Hölder’s inequality that needs L^{2p} norms of the individual functions. In this section we will see how transversality works in higher dimensions.

Theorem 1.1 (Loomis–Whitney). *Let $n \geq 2$ and denote by $\pi_j : \mathbb{R}^n \rightarrow \mathbb{R}^{n-1}$ the linear map that forgets the j^{th} coordinate:*

$$\pi_j(x_1, \dots, x_n) := (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n).$$

Suppose that $f_j : \mathbb{R}^{n-1} \rightarrow [0, \infty]$ are (measurable) functions. Then the following integral inequality holds.

$$\int_{\mathbb{R}^n} \prod_{j=1}^n f_j(\pi_j(x))^{1/n} dx \leq \prod_{j=1}^n \|f_j\|_{L^1(\mathbb{R}^{n-1})}^{1/n}. \tag{1.1}$$

Proof. By induction on n . For $n = 2$ the claim (1.1) is a direct consequence of Fubini's theorem. Suppose that the claim (1.1) is known for some $n \geq 2$, we will now show it with n replaced by $n + 1$. By Hölder's inequality, the inductive hypothesis, and again Hölder's inequality we obtain

$$\begin{aligned}
& \int_{\mathbb{R}^{n+1}} \prod_{j=1}^{n+1} f_j(\pi_j(x))^{1/n} dx \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}^n} \prod_{j=1}^n f_j(\pi_j(x', x_{n+1}))^{1/n} \cdot f_{n+1}(x')^{1/n} dx' dx_{n+1} \\
&\leq \|f_{n+1}\|_1^{1/n} \int_{\mathbb{R}} \left(\int_{\mathbb{R}^n} \left(\prod_{j=1}^n f_j(\pi_j(x', x_{n+1}))^{1/n} \right)^{n/(n-1)} dx' \right)^{(n-1)/n} dx_{n+1} \\
&\leq \|f_{n+1}\|_1^{1/n} \int_{\mathbb{R}} \left(\prod_{j=1}^n \|f_j(\cdot, x_{n+1})\|_{L^1(\mathbb{R}^{n-1})}^{1/(n-1)} \right)^{(n-1)/n} dx_{n+1} \\
&\leq \|f_{n+1}\|_1^{1/n} \prod_{j=1}^n \left(\int_{\mathbb{R}} \|f_j(\cdot, x_{n+1})\|_{L^1(\mathbb{R}^{n-1})} dx_{n+1} \right)^{1/n} \\
&= \|f_{n+1}\|_1^{1/n} \prod_{j=1}^n \|f_j\|_{L^1(\mathbb{R}^n)}^{1/n}
\end{aligned}$$

as claimed. \square

We will also use a variant in which the functions are constant in not necessarily orthogonal directions.

Corollary 1.2 (Affine invariant Loomis–Whitney). *Let $N_1, \dots, N_n \in \mathbb{R}^n$ be unit vectors and $f_j : N_j^\perp \rightarrow [0, \infty]$ be measurable functions. Then*

$$\int_{\mathbb{R}^n} \prod_{j=1}^n (f_j \circ \pi_j)^{\frac{1}{n-1}} \leq |N_1 \wedge \dots \wedge N_n|^{-\frac{1}{n-1}} \prod_{j=1}^n \|f_j\|_{L^1(N_j^\perp)}^{\frac{1}{n-1}}, \quad (1.2)$$

where π_j denotes the orthogonal projection onto N_j^\perp .

We recall that for column vectors $N_1, \dots, N_n \in \mathbb{R}^n$ we have

$$|N_1 \wedge \dots \wedge N_n| = |\det(N_1 \dots N_n)|.$$

We will not be using any other properties of wedge products, so this identity can be seen as an abbreviation.

Proof. Let A be the linear map such that $A(e_j) = N_j$, where e_1, \dots, e_n is the standard basis of \mathbb{R}^n . By a change of variables and Theorem 1.1 we have

$$\int_{\mathbb{R}^n} \prod_{j=1}^n (f_j \circ \pi_j)^{\frac{1}{n-1}} = |\det A| \int_{\mathbb{R}^n} \prod_{j=1}^n (f_j \circ \pi_j \circ A)^{\frac{1}{n-1}} \leq |\det A| \prod_{j=1}^n \|f_j \circ \pi_j \circ A\|_{L^1(e_j^\perp)}^{\frac{1}{n-1}}.$$

It remains to observe that $\pi_j \circ A$, viewed as a map $e_j^\perp \rightarrow N_j^\perp$, also has determinant $\det A$, so the above equals

$$= |\det A| \prod_{j=1}^n \left(|\det A|^{-1} \|f_j\|_{L^1(N_j^\perp)} \right)^{\frac{1}{n-1}} = |\det A|^{-\frac{1}{n-1}} \prod_{j=1}^n \|f_j\|_{L^1(N_j^\perp)}^{\frac{1}{n-1}}. \quad \square$$

The multilinear Kakeya inequality [BCT06] is a version of Loomis–Whitney in which instead of n single directions we have n families of directions. The argument presented below (from [Gut15]) incurs a loss in the passage to a family of directions, and this loss restricts the kind of functions we can consider. Instead of arbitrary functions $f \circ \pi$ we will use functions that are morally constant at some given scale r , and estimate the integral at a larger scale R .

An r -tube is an infinite cylinder of radius r , or more precisely the r -neighborhood of its *central line*, an affine subspace of dimension 1. The *direction* of a tube is a unit vector parallel to its central line (there are two such vectors, the choice is not important).

Theorem 1.3 (Multilinear Kakeya). *For every $n \in \mathbb{N}$ and $\varepsilon > 0$ there exist $C_1, C_2 < \infty$ such that the following holds. Let $0 < \nu < 1$ and $S_j \subset \mathbb{S}^{n-1}$ be such that for any vectors $v_j \in S_j$,*

$$|v_1 \wedge \cdots \wedge v_n| \geq \nu.$$

Let $r > 0$ and for each $j = 1, \dots, n$ let \mathbf{T}_j be a collection of tubes of radius r in \mathbb{R}^n whose directions lie in S_j . Then for any ball B of radius $R \geq r$ and any positive numbers w_T we have

$$\int_B \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} w_T \mathbf{1}_T \right)^{\frac{1}{n-1}} \leq C_1 \nu^{-C_2} (R/r)^\varepsilon r^n \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} w_T \right)^{\frac{1}{n-1}} \quad (1.3)$$

The power of ν in (1.3) can be taken the same as in the Loomis–Whitney inequality, and the loss $(R/r)^\varepsilon$ can be removed. However, this requires a completely different proof [Gut10] (see also [Zor18] for a more general result).

Proof. By monotone convergence we may assume that the collections \mathbf{T}_j are finite. By continuity we may assume that all coefficients w_T are rational. By scaling we may assume that they are all integers. Allowing repetition of tubes in \mathbf{T}_j we may assume $w_T = 1$ for all T .

Let $\delta \sim \nu^{C_3}$ be a dyadic number with $C_3 = C_3(n, \varepsilon)$ to be chosen later. Replacing ν by $\nu/2$, say, and partitioning each S_j into $O(\delta^{n-1})$ subsets of diameter $\ll \delta$ we may assume that each S_j is a ball or radius $\ll \delta$ centered at some $\tilde{e}_j \in S^{n-1}$. This loses a factor ν^{-C} in (1.3).

With these reductions, for dyadic numbers $r \leq R$ let $\mathcal{K}(R, r)$ be the smallest constant such that the inequality

$$\int_Q \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_T \right)^{\frac{1}{n-1}} \leq \mathcal{K}(R, r) \prod_{j=1}^n |\mathbf{T}_j|^{\frac{1}{n-1}}$$

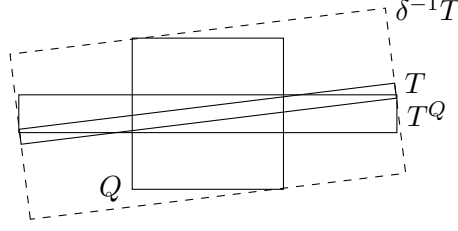
holds for any collections \mathbf{T}_j of tubes of radius r and any cube Q_R of side length R . By scaling $\mathcal{K}(R, r) = r^n \mathcal{K}(R/r, 1)$. Estimating $\sum_{T \in \mathbf{T}_j} \mathbf{1}_T \leq |\mathbf{T}_j|$ it becomes clear that $\mathcal{K}(R, 1) \lesssim R^n$. It suffices to show that in fact $\mathcal{K}(R, 1) \lesssim R^\varepsilon$. We will obtain this estimate by going from scale R to scale 1 in many steps and using the Loomis–Whitney inequality at each step.

Let \mathbf{T}_j be collections of 1-tubes with directions δ -close to \tilde{e}_j . We partition Q_R into cubes Q of side length $\sim \delta^{-1}$. For each such cube Q let $\mathbf{T}_j(Q)$ be the set of tubes in \mathbf{T}_j that intersect Q . For each $T \in \mathbf{T}_j(Q)$, the intersection $T \cap Q$ is contained in a slightly thicker 2-tube T^Q with direction \tilde{e}_j . Therefore,

$$\int_Q \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_T \right)^{\frac{1}{n-1}} = \int_Q \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j(Q)} \mathbf{1}_T \right)^{\frac{1}{n-1}} \leq \int_Q \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_{T^Q} \right)^{\frac{1}{n-1}}.$$

This last integral involves only \tilde{e}_j -parallel tubes of radius 2, and by the Loomis–Whitney inequality it is

$$\leq \nu^{-1/(n-1)} C_n \prod_{j=1}^n |\mathbf{T}_j(Q)|^{\frac{1}{n-1}}.$$



For each $T \in \mathbf{T}_j(Q)$ we have $\delta^{-1}T \supset Q$, where $\delta^{-1}T$ is the δ^{-1} -tube with the same central line as T . Therefore,

$$\prod_{j=1}^n |\mathbf{T}_j(Q)|^{\frac{1}{n-1}} \leq C_n |Q|^{-1} \int_Q \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_{\delta^{-1}T} \right)^{\frac{1}{n-1}}.$$

Summing over the cubes Q we get

$$\int_{Q_R} \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_T \right)^{\frac{1}{n-1}} \leq \nu^{-1/(n-1)} C_n \delta^n \int_{Q_R} \prod_{j=1}^n \left(\sum_{T \in \mathbf{T}_j} \mathbf{1}_{\delta^{-1}T} \right)^{\frac{1}{n-1}}.$$

By definition of \mathcal{K} the right-hand side is bounded by

$$\nu^{-1/(n-1)} C_n \delta^n \mathcal{K}(R, \delta^{-1}) \prod_{j=1}^n |\mathbf{T}_j|^{\frac{1}{n-1}},$$

so

$$\mathcal{K}(R, 1) \leq \nu^{-1/(n-1)} C_n \delta^n \mathcal{K}(R, \delta^{-1}) = \nu^{-1/(n-1)} C_n \mathcal{K}(\delta R, 1).$$

Iterating this inequality $\log R / \log \delta^{-1}$ times and using the trivial estimate for \mathcal{K} at the end we obtain

$$\mathcal{K}(R, 1) \leq C'_n (\nu^{-1/(n-1)} C_n)^{\log R / \log \delta^{-1}} \sim R^{c \log \nu^{-1} / \log \delta^{-1}} \sim R^{c/C''}.$$

Choosing $C''' = C'''(n, \varepsilon)$ sufficiently large finishes the proof. \square

It is inefficient to use the affine-invariant Loomis–Whitney inequality in each step of the iteration; making a change of variables that maps \tilde{e}_j to the standard unit vectors e_j at the beginning of the iteration would result in a better dependence on ν . The above proof of Theorem 1.3 however has the advantage to easily generalize to more general Brascamp–Lieb data, which we will look at later.

Finally, we record what the multilinear Kakeya inequality in Theorem 1.3 tells when applied to general functions.

Corollary 1.4. *Let $S_j \subset \mathbb{S}^{n-1}$ be finite subsets as in Theorem 1.3. Let $0 < r < R < \infty$ and for each $v \in S_j$ let \mathbf{T}_v be a partition \mathbb{R}^n into rectangular boxes of size $r \times \cdots \times r \times R$ whose long axis point in the direction of v . Then for any measurable functions f_v on \mathbb{R}^n and any ball B_R of radius R we have*

$$\int_{B_R} \prod_{j=1}^n \left| \sum_{v \in S_j} f_v \right|^{1/(n-1)} \lesssim_{\varepsilon, \nu} (R/r)^\varepsilon r^n \prod_{j=1}^n \left(\sum_{v \in S_j} \sum_{T \in \mathbf{T}_v: T \cap B_R \neq \emptyset} \sup_{x \in T} |f_v(x)| \right)^{1/(n-1)}$$

Proof. Dominate

$$\mathbf{1}_{B_R} |f_v| \leq \sum_{T \in \mathbf{T}_v: T \cap B_R \neq \emptyset} \mathbf{1}_T \sup_{x \in T} |f_v(x)|$$

and apply (1.3) with $w_T = \sup_{x \in T} |f_v(x)|$. \square

2 Multilinear restriction

The abbreviation $A \lesssim B$ means $A \leq CB$ with some constant C that typically does not depend on functions f , but does depend on dimension and Lebesgue exponents. Dependence on some specific parameters will be indicated by a subscript: \lesssim_ε .

Theorem 2.1 ([BCT06]). *For every $n \in \mathbb{N}$, $\varepsilon > 0$, and compact C^2 hypersurfaces $A_1^0, \dots, A_n^0 \subset \mathbb{R}^n$ there exist $C_1, C_2 < \infty$ such that the following holds.*

Denote by $N_j(x)$ the unit normal vector to A_j at x . Let $A_j \subset A_j^0$, $j = 1, \dots, n$, be compact subsets and assume that for some $0 < \nu < 1$ and all $x_j \in A_j$ we have

$$|N_1(x_1) \wedge \dots \wedge N_n(x_n)| > \nu.$$

For $R > 1$ denote by A_j^R the R^{-1} -neighborhood of A_j . Then for any functions f_j with $\text{supp } \widehat{f_j} \subset A_j^R$ we have

$$\left\| \prod_{j=1}^n \sup_{z \in B(y,1)} |f_j(z)| \right\|_{L_y^{2/(n-1)}(B(0,R))} \leq C_1 \nu^{-C_2} R^{-n/2+\varepsilon} \prod_{j=1}^n \|f_j\|_2. \quad (2.1)$$

The left-hand side of (2.1) is morally equal to

$$\left\| \prod_{j=1}^n |f_j| \right\|_{L^{2/(n-1)}(B(0,R))},$$

since due to the compact Fourier support the functions $|f_j|$ are constant at scale 1. However, in some applications it is convenient to use the formally larger expression in (2.1).

If the surfaces A_j were completely flat, then the Fourier supports A_j^R would be contained in boxes of size $O(1) \times \dots \times O(1) \times O(R^{-1})$, so by the uncertainty principle the functions f_j would be essentially constant on boxes of sizes $1 \times \dots \times 1 \times R$. The estimate (2.1) would then essentially follow from the Loomis–Whitney inequality applied to the functions $|f_j|^2$.

The C^2 hypothesis ensures that the surfaces do in fact look flat, but only at scale $R^{-1/2}$, in the sense that an $R^{-1/2}$ -ball inside a C^2 surface fits into a box of dimensions $R^{-1/2} \times \dots \times R^{-1/2} \times R^{-1}$. We will have to use an induction on scales argument to take advantage of this.

It is conjectured that ν should enter (2.1) with the same exponent as in the Loomis–Whitney inequality. This is currently only known for $n = 2$ (where it is easy) and $n = 3$, see [Ram18]. For one of our applications (maximal estimates for the Schrödinger equation) it will be important that the constant in (2.1) only grows with a power of ν .

Proof of Theorem 2.1. Let $\text{MIRs}(R, \nu)$ be the smallest constant for which the estimate (2.1) holds. By translation invariance the same estimate holds for any ball of radius R , not necessarily centered at 0. For functions f_j with compact Fourier support we have the Bernstein inequality $\|f_j\|_\infty \lesssim \|f_j\|_2$, and it follows that

$$\text{MIRs}(R, \nu) \lesssim R^C. \quad (2.2)$$

Here and later in this proof the implicit constants do not depend on R, ν . We will use induction on scales to obtain a better bound.

Let φ be a Schwartz function on \mathbb{R}^n with compact Fourier support such that $|\varphi| \geq \mathbf{1}_{B(0,2)}$. For each $R \geq 1$ and $x \in \mathbb{R}^n$ let $\varphi_{R^{1/2}}^x(y) := \varphi(R^{-1/2}(y - x))$. Then

$\text{supp } \widehat{\varphi_{R^{1/2}}^x} \subset B(0, CR^{-1/2})$, and it follows that $\text{supp } \widehat{\varphi_{R^{1/2}}^x f_j} \subset A_j^{CR^{1/2}}$. Hence

$$\begin{aligned} \left\| \prod_{j=1}^n \sup_{z \in B(y,1)} |f_j| \right\|_{L_y^{2/(n-1)}(B(x,R^{1/2}))} &\leq \left\| \prod_{j=1}^n \sup_{z \in B(y,1)} |\varphi_{R^{1/2}}^x f_j| \right\|_{L_y^{2/(n-1)}(B(x,R^{1/2}))} \\ &\leq \text{MIRs}(R^{1/2}, \nu) \prod_{j=1}^n \|\varphi_{R^{1/2}}^x f_j\|_2. \end{aligned}$$

Averaging this inequality over $x \in B(0, R)$ we obtain

$$LHS(2.1) \lesssim \text{MIRs}(R^{1/2}, \nu) \left(R^{-n/2} \int_{B(0,R)} \prod_{j=1}^n \|\varphi_{R^{1/2}}^x f_j\|_2^{2/(n-1)} dx \right)^{(n-1)/2}. \quad (2.3)$$

Now for each $1 \leq j \leq n$ we split $A_j = \cup_a A_{j,a}$ into disks of radius $R^{-1/2}$ and make an orthogonal splitting $f_j = \sum_a f_{j,a}$ such that $\text{supp } \widehat{f_{j,a}} \subset A_{j,a}^R$. Since A_j is a C^2 surface, the set $A_{j,a}^R$ is contained in a box $U_{j,a}$ of size $O(R^{-1}) \times O(R^{-1/2}) \times \dots \times O(R^{-1/2})$ whose short side points in the normal direction $v_{j,a}$ at some (arbitrary) point on $A_{j,a}$. Also, for each j the supports of the functions $\widehat{\varphi_{R^{1/2}}^x f_{j,a}}$ have bounded overlap, and by almost orthogonality we obtain

$$(2.3) \lesssim R^{\alpha/2+\varepsilon} \left(R^{-n/2} \int_{B(0,R)} \left(\prod_{j=1}^n \sum_a \|\varphi_{R^{1/2}}^x f_{j,a}\|_2^2 \right)^{1/(n-1)} dx \right)^{(n-1)/2}. \quad (2.4)$$

Let

$$F_{j,a}(x) := \|\varphi_{R^{1/2}}^x f_{j,a}\|_2^2.$$

By the uncertainty principle we expect $F_{j,a}$ to be approximately constant on tubes of size $R \times R^{1/2} \times \dots \times R^{1/2}$ whose long side points in the direction $v_{j,a}$. By the multilinear Kakeya inequality (Corollary 1.4) we estimate

$$\begin{aligned} (2.4) &= \text{MIRs}(R^{1/2}, \nu) \left(R^{-n/2} \int_{B(0,R)} \left(\prod_{j=1}^n \sum_a F_{j,a}(x) \right)^{1/(n-1)} dx \right)^{(n-1)/2} \\ &\leq \text{MIRs}(R^{1/2}, \nu) C_1' \nu^{-C_2'} R^\varepsilon \prod_{j=1}^n \left(\sum_a \sum_{T \in \mathbf{T}_{j,a}} \sup_{x \in T} F_{j,a}(x) \right)^{1/2}, \end{aligned}$$

where $\mathbf{T}_{j,a} = \mathbf{T}_{v_{j,a}}$ is as in Corollary 1.4 with $(R, R^{1/2})$ in place of (R, r) . Here C_1' and C_2' depend on $\varepsilon > 0$.

Now we make the uncertainty principle precise. Fix some Schwartz function with $\mathbf{1}_{B(0,1)} \leq \widehat{\psi} \leq \mathbf{1}_{B(0,2)}$. For each j, a let $\psi_{j,a}$ be its dilation and modulation such that $\widehat{\psi_{j,a}}$ is adapted to $U_{j,a}$, so that $f_{j,a} = \widehat{f_{j,a}} * \psi_{j,a}$. Then in particular by Hölder's inequality

$$\begin{aligned} |f_{j,a}(y)|^2 &= |f_{j,a} * \psi_{j,a}(y)|^2 = \left| \int f_{j,a}(y-z) \psi_{j,a}(z) dz \right|^2 \\ &\leq \left(\int |f_{j,a}(y-z)|^2 |\psi_{j,a}(z)| dz \right) \left(\int |\psi_{j,a}(z)| dz \right) \\ &\lesssim \int |f_{j,a}(y-z)|^2 |\psi_{j,a}(z)| dz = (|f_{j,a}|^2 * |\psi_{j,a}|)(x). \end{aligned}$$

Therefore for every $x \in T \in \mathbf{T}_{j,a}$ we have

$$\begin{aligned} F_{j,a}(x) &\lesssim \int \left(\int |f_{j,a}(y-z)|^2 |\psi_{j,a}(z)| dz \right) |\varphi_{R^{1/2}}^x|^2(y) dy \\ &= \iint |f_{j,a}|^2(y) |\psi_{j,a}(z)| |\varphi_{R^{1/2}}^x|^2(y+z) dz dy \\ &\lesssim R^{-1/2} \int |f_{j,a}|^2(y) \chi_T(y) dy, \end{aligned}$$

where χ_T is a smooth version of the characteristic function. Inserting this above we obtain

$$\begin{aligned}
(2.4) &\lesssim \text{MIRs}(R^{1/2}, \nu) C'_1 \nu^{-C'_2} R^\varepsilon \left(\sum_a \sum_{T \in \mathbf{T}_{j,a}} R^{-1/2} \int |f_{j,a}|^2(y) \chi_T(y) dy \right)^{1/2} \\
&\lesssim \text{MIRs}(R^{1/2}, \nu) C'_1 \nu^{-C'_2} R^{\varepsilon-n/4} \prod_{j=1}^n \left(\sum_a \|f_{j,a}\|_2^2 \right)^{1/2} \\
&= \text{MIRs}(R^{1/2}, \nu) C'_1 \nu^{-C'_2} R^{\varepsilon-n/4} \prod_{j=1}^n \|f_j\|_2.
\end{aligned}$$

Taking a supremum over all possible f_j we obtain

$$\text{MIRs}(R, \nu) \leq \text{MIRs}(R^{1/2}, \nu) C''_1 \nu^{-C'_2} R^{\varepsilon-n/4}.$$

Fix $\varepsilon > 0$. If $R^\varepsilon < C''_1 \nu^{-C'_2}$ we use the trivial bound (2.2). Otherwise we can estimate

$$\text{MIRs}(R, \nu) \leq \text{MIRs}(R^{1/2}, \nu) R^{2\varepsilon-n/4}.$$

Iterating this bound until we arrive at the trivial bound we will obtain

$$\text{MIRs}(R, \nu) \lesssim \nu^{-C} R^{4\varepsilon-n/2}. \quad \square$$

The multilinear restriction inequality is also frequently stated in the following form.

Corollary 2.2. *In the setting of Theorem 2.1 let σ_j denote the surface measure (($n-1$)-dimensional Hausdorff measure) on A_j . Let also $f_j \in L^2(A_j, \sigma_j)$. Then for every $\varepsilon > 0$ and every $R > 1$ we have*

$$\left\| \prod_{j=1}^n \sup_{z \in B(y,1)} |\widehat{f_j \sigma_j}(z)| \right\|_{L_y^{2/(n-1)}(B(0,R))} \lesssim_{\nu, \varepsilon} R^\varepsilon \prod_{j=1}^n \|f_j\|_{L^2(\sigma_j)}. \quad (2.5)$$

Proof. Let $\varphi_R(y) = \varphi(y/R)$, where φ is a Schwartz function with $|\varphi| \geq \mathbf{1}_{B(0,2)}$ and $\text{supp } \widehat{\varphi} \subset B(0,1)$. Then

$$\text{LHS}(2.5) \leq \left\| \prod_{j=1}^n \sup_{z \in B(y,1)} |\varphi_R \widehat{f_j \sigma_j}(z)| \right\|_{L_y^{2/(n-1)}(B(0,R))} \lesssim R^{-n/2+\varepsilon} \prod_{j=1}^n \|\varphi_R \widehat{f_j \sigma_j}\|_2,$$

since

$$\text{supp } \widehat{\varphi_R f_j \sigma_j} \subset \text{supp } \widehat{\varphi_R} + \text{supp}(f_j \sigma_j) \subset A_j^R.$$

Moreover,

$$\|\varphi_R \widehat{f_j \sigma_j}\|_2 = \|\widehat{\varphi_R} * (f_j \sigma_j)\|_{L^2(\mathbb{R}^n)} \lesssim R^{1/2} \|f_j\|_{L^2(\sigma_j)}.$$

The latter inequality can be seen by subdividing \mathbb{R}^n into cubes with side length R^{-1} . \square

3 Decoupling for the paraboloid

The Fourier decoupling inequality is a kind of L^p orthogonality statement for functions with Fourier support close to the paraboloid $\mathbb{P} = \{(\xi, |\xi|^2) \mid \xi \in \mathbb{R}^d\}$. Some of its far-reaching applications will be discussed later on. It was originally proved in [BD15]. Our presentation is based on the simplified proofs in [BD17] and [GZ19].

For inductive purposes it is convenient to formulate the Fourier support condition in a way that is invariant under affine transformations that preserve the paraboloid. We will consider functions with Fourier support inside parallelepipeds adapted to

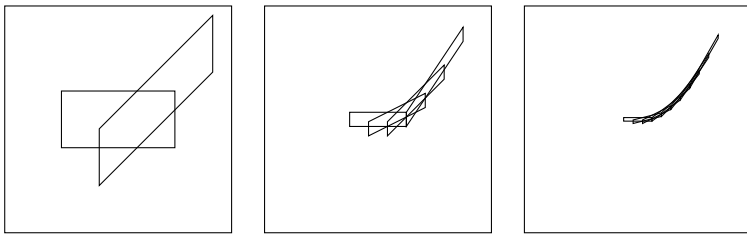


Figure 1: Fourier support parallelepipeds at scales $2^{-1}, 2^{-2}, 2^{-3}$

the paraboloid \mathbb{P} . Figure 3 shows the parallelepipeds of scales $2^{-1}, 2^{-2}$, and 2^{-3} , respectively, inside the parallelepiped of scale 2^0 (in the case $d = 1$). The picture is self-similar: inside each parallelepiped of a given scale there are 2^d parallelepipeds, and any parallelepiped can be mapped onto any other by an affine transformation that preserves the collection of all parallelepipeds.

We proceed with a more formal description. A *dyadic cube* (of side length δ) is a cube of the form $\delta(a + [0, 1]^d)$ with $a \in \mathbb{Z}^d$ and δ a power of 2. We will denote by $\mathcal{P}(Q, \delta)$ the partition of a dyadic cube Q into dyadic cubes with side length δ (δ must be smaller than the side length of Q for this to make sense). We omit Q from the notation $\mathcal{P}(Q, \delta)$ if $Q = [0, 1]^d$.

For $\theta = a + \delta[0, 1]^d \in \mathcal{P}(\delta)$ we will denote by f_θ an arbitrary function of the form $M_\theta f$, where $f \in L^p(\mathbb{R}^{d+1})$ with $\text{supp } \widehat{f} \subset [-2, 2]^{d+1}$ and

$$M_\theta f(x, y) = e(a \cdot x + |a|^2 y)(f \circ L_\theta)(x, y), \quad x \in \mathbb{R}^d, y \in \mathbb{R}, \quad (3.1)$$

where L_θ is the linear transformation

$$L_\theta = \begin{pmatrix} \delta I_d & 0 \\ 0 & \delta^2 \end{pmatrix} \begin{pmatrix} I_d & (2a_1, \dots, 2a_d)^T \\ 0 & 1 \end{pmatrix}.$$

and I_d denotes the identity $d \times d$ matrix. Equivalently, f_θ denotes an arbitrary L^p function with the Fourier support condition

$$\text{supp } \widehat{f}_\theta \subseteq (a, |a|^2) + L_\theta^*([-2, 2]^{d+n}). \quad (3.2)$$

Roughly speaking, $\text{supp } \widehat{f}_\theta$ is contained in a box of size $\delta \times \dots \times \delta \times \delta^2$ that contains the part of paraboloid over θ .

For $1 \leq p, q \leq \infty$ and $\delta > 0$, let the $\ell^q L^p$ decoupling constant $\text{Dec}_d^{p,q}(\delta)$ be the smallest constant such that the inequality

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^p} \leq \text{Dec}_d^{p,q}(\delta) \left(\sum_{\theta} \|f_\theta\|_{L^p}^q \right)^{1/q} \quad (3.3)$$

holds for any functions f_θ as above. The main decoupling estimate is the following.

Theorem 3.1 ([BD15]). *Let $d \geq 1$ and $2 \leq p \leq \infty$. Then for every $\varepsilon > 0$ we have*

$$\text{Dec}_d^{p,2}(\delta) \lesssim_\varepsilon \begin{cases} \delta^{-\varepsilon}, & 2 \leq p \leq 2(d+2)/d, \\ \delta^{-d/2+(d+2)/p-\varepsilon}, & 2(d+2)/d < p \leq \infty. \end{cases} \quad (3.4)$$

Remark. This is a comment on numerology, or the exponents in (3.4). For $p = 2$ and $p = \infty$ the inequality (3.4) holds even with $\varepsilon = 0$. For $p = 2$ this follows from Plancherel's theorem, and for $p = \infty$ from the triangle inequality. By a complex interpolation argument that is explained in Section 3.2.3 it will suffice to consider the endpoint $p = 2(d+2)/d$ to obtain the inequality (3.4) for all $2 \leq p \leq \infty$.

3.1 Optimality of the decoupling inequality

In this section we show that the estimates in Theorem 3.1 are essentially optimal.

Example. Consider first $f_\theta = M_\theta f$, where f is a fixed Schwartz function with $f(0) = 1$ and \hat{f} supported in the unit cube. Then by scaling $\|f_\theta\|_p \sim \delta^{-(d+2)/p}$. On the other hand, $\Re f_\theta \gtrsim 1$ on a fixed neighborhood of 0. Hence

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_p \gtrsim |\mathcal{P}(\delta)| \sim \delta^{-d},$$

and it follows that

$$\text{Dec}_d^{p,q}(\delta) \gtrsim \delta^{-d + \frac{d}{q} + \frac{d+2}{p}}.$$

For $q = 2$ and $p > 2(d+2)/d$ this is the exponent in (3.4) (up to the ε loss).

Example. Consider next $f_\theta(x, y) = \eta(\delta^2(x, y))e(a \cdot x + |a|^2 y)$, where η is a Schwartz function with $\hat{\eta}$ supported in the unit cube and $a \in \theta$. Then $\|f_\theta\|_p \sim \delta^{-2(d+1)/p}$ and by Hölder's inequality with $2 \leq p \leq \infty$ and orthogonality

$$\begin{aligned} \delta^{-2(d+1)(\frac{1}{2} - \frac{1}{p})} \left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_p &\sim \|\eta(\delta^2 \cdot)\|_{\frac{1}{1/2 - 1/p}} \left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_p \geq \left\| \sum_{\theta \in \mathcal{P}(\delta)} \eta(\delta^2 \cdot) f_\theta \right\|_2 \\ &\gtrsim \left(\sum_{\theta \in \mathcal{P}(\delta)} \|\eta(\delta^2 \cdot) f_\theta\|_2^2 \right)^{1/2} \sim \delta^{-d/2} \delta^{-2(d+1)/2}. \end{aligned}$$

It follows that

$$\text{Dec}_d^{p,q}(\delta) \gtrsim \delta^{\frac{d}{q} - \frac{d}{2}}$$

for $2 \leq p \leq \infty$, and in the case $q = 2$ and $2 \leq p \leq 2(d+2)/d$ this is exactly the exponent in (3.4) (up to the ε loss).

Example. Finally, consider again the functions in the first example and translate them so that they become essentially disjointly supported. Then

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^p} \approx \left(\sum_{\theta \in \mathcal{P}(\delta)} \|f_\theta\|_{L^p}^p \right)^{1/p} \approx |\mathcal{P}(\delta)|^{1/p - 1/q} \left(\sum_{\theta \in \mathcal{P}(\delta)} \|f_\theta\|_{L^p}^q \right)^{1/q},$$

so

$$\text{Dec}_d^{p,q}(\delta) \gtrsim \delta^{d/q - d/p}.$$

This shows that we cannot expect any estimates for $p < q = 2$ other than those that follow by interpolation between orthogonality at $p = 2$ and Minkowski's inequality at $p = 1$.

Remark. It is known from [Bou93, p. 118] that the ε loss in (3.4) cannot be completely removed in general. On the other hand, the precise dependence on δ in (3.4) has not been quantified except in the case $d = 1$ [Li17].

3.2 Basic properties of the decoupling constant

3.2.1 Parabolic scaling

We use functions of the form (3.1) in order to make explicit a scaling invariance of the decoupling inequality. If $\delta_0, \delta_1 \leq 1$ are powers of 2 and $\theta_0 \in \mathcal{P}(\delta_0)$, then there is a natural bijection between $\mathcal{P}(\delta_1)$ and $\mathcal{P}(\theta_0, \delta_0 \delta_1)$ given by the composition of translation and scaling that maps $[0, 1]^d$ to θ . Moreover, if θ is mapped to θ' by this

bijection, then $M_{\theta'} = M_{\theta_0} \circ M_{\theta}$, so we can write $f_{\theta'} = M_{\theta_0} \tilde{f}_{\theta}$. Hence

$$\begin{aligned}
\left\| \sum_{\theta' \in \mathcal{P}(\theta_0, \delta_0 \delta_1)} f_{\theta'} \right\|_p &= \left\| M_{\theta_0} \sum_{\theta \in \mathcal{P}(\delta_1)} \tilde{f}_{\theta} \right\|_p \\
&= \delta_0^{-(d+2)/p} \left\| \sum_{\theta \in \mathcal{P}(\delta_1)} \tilde{f}_{\theta} \right\|_p \\
&\leq \delta_0^{-(d+2)/p} \text{Dec}_d^{p,q}(\delta_1) \left(\sum_{\theta \in \mathcal{P}(\delta_1)} \|\tilde{f}_{\theta}\|_p^q \right)^{1/q} \\
&= \text{Dec}_d^{p,q}(\delta_1) \left(\sum_{\theta \in \mathcal{P}(\delta_1)} \|M_{\theta_0} \tilde{f}_{\theta}\|_p^q \right)^{1/q} \\
&= \text{Dec}_d^{p,q}(\delta_1) \left(\sum_{\theta' \in \mathcal{P}(\theta_0, \delta_0 \delta_1)} \|f_{\theta'}\|_p^q \right)^{1/q}.
\end{aligned}$$

3.2.2 Larger Fourier support

The choice of the Fourier support condition $\text{supp } \hat{f} \subset [-2, 2]^d$ in $f_{\theta} = M_{\theta} f$ is not particularly important. In particular, for functions of the form $f_{\theta} = M_{\theta} f$ with $\text{supp } \hat{f} \subset [-C, C]^d$ we obtain

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_{\theta} \right\|_{L^p} \lesssim \text{Dec}_d^{p,q}(C\delta) \left(\sum_{\theta} \|f_{\theta}\|_{L^p}^q \right)^{1/q}. \quad (3.5)$$

This is because such f_{θ} satisfy (3.2) with a larger dyadic cube (of scale $\leq C\delta$).

3.2.3 Interpolation

Let $1 \leq p_0, p_1, q_0, q_1 \leq \infty$ and $0 < \eta < 1$. Define p_{η}, q_{η} by

$$\frac{1}{p_{\eta}} = \frac{1-\eta}{p_0} + \frac{\eta}{p_1}, \quad \frac{1}{q_{\eta}} = \frac{1-\eta}{q_0} + \frac{\eta}{q_1}.$$

Then we have

$$\text{Dec}_d^{p_{\eta}, q_{\eta}}(\delta) \lesssim \text{Dec}_d^{p_0, q_0}(4\delta)^{1-\eta} \text{Dec}_d^{p_1, q_1}(4\delta)^{\eta}. \quad (3.6)$$

This would follow from complex interpolation if we could disregard the Fourier support restrictions, which is of course impossible. In order to apply a standard complex interpolation result we have to reformulate the decoupling inequality (3.3) as an estimate for a linear operator on an $\ell^q L^p$ space. To this end let ψ be a smooth function on \mathbb{R}^{d+1} with $\mathbf{1}_{[-2,2]^{d+1}} \leq \psi \leq \mathbf{1}_{[-4,4]^{d+1}}$. Define the multiplier operator $\widehat{Tf} := \psi \hat{f}$. Since this is a convolution operator with kernel $\check{\psi} \in L^1(\mathbb{R}^{d+1})$, it is bounded on any L^p space with $1 \leq p \leq \infty$. It follows that the operators $T_{\theta} := M_{\theta} T M_{\theta}^{-1}$ are bounded uniformly for all dyadic cubes θ . On the other hand, for an arbitrary function $g \in L^p(\mathbb{R}^{d+1})$ the function $T_{\theta} g$ satisfies the support condition in Section 3.2.2. It follows that for arbitrary functions $g_{\theta} \in L^p(\mathbb{R}^{d+1})$ we have

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} T_{\theta} g_{\theta} \right\|_p \lesssim \text{Dec}_d^{p,q}(4\delta) \left(\sum_{\theta} \|T_{\theta} g_{\theta}\|_p^q \right)^{1/q} \lesssim \text{Dec}_d^{p,q}(4\delta) \left(\sum_{\theta} \|g_{\theta}\|_p^q \right)^{1/q},$$

where we have used (3.5) and the uniform boundedness of the operators T_{θ} . This is now an $\ell^q(\mathcal{P}(\delta), L^p(\mathbb{R}^{d+1})) \rightarrow L^p(\mathbb{R}^{d+1})$ bound for the linear operator $(g_{\theta})_{\theta} \mapsto \sum_{\theta} T_{\theta} g_{\theta}$. By complex interpolation we obtain

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} T_{\theta} g_{\theta} \right\|_{p_{\eta}} \lesssim \text{Dec}_d^{p_0, q_0}(4\delta)^{1-\eta} \text{Dec}_d^{p_1, q_1}(4\delta)^{\eta} \left(\sum_{\theta} \|g_{\theta}\|_{p_{\eta}}^{q_{\eta}} \right)^{1/q_{\eta}}.$$

On the other hand, for f_{θ} satisfying the standing Fourier support condition we have $T_{\theta} f_{\theta} = f_{\theta}$, and this implies (3.6).

3.3 Localization

The freedom to enlarge the Fourier support in Section 3.2.2 can be used to localize decoupling inequalities. Let η be a positive Schwartz function on \mathbb{R}^{d+1} such that $\text{supp } \hat{\eta} \subset B(0, c)$ and $\eta \geq 1$ on $B(0, 1)$. Let $B = B(x, \delta^{-2}) \subset \mathbb{R}^{d+1}$ be a ball of radius δ^{-2} centered at x and $\eta_B := \eta(\delta^2(\cdot - x))$. Then the functions $f_\theta \eta_B$ are as in Section 3.2.2, so we obtain

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \eta_B \right\|_p \lesssim \text{Dec}_d^{p,q}(C\delta) \left(\sum_{\theta} \|f_\theta\|_p^q \right)^{1/q}. \quad (3.7)$$

It is sometimes convenient to use other kinds of weights. For a ball $B = B(c_B, r_B) \subset \mathbb{R}^{d+1}$ and $E > d + 1$, define an associated weight

$$w_{B,E}(x) := \left(1 + \frac{|x - c_B|}{r_B} \right)^{-E}. \quad (3.8)$$

The exponent E will not be important and will be usually omitted from the notation.

From the estimate (3.7) and using $\mathbf{1}_B \leq \eta_B \lesssim w_B$ we immediately obtain

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^p(B)} \lesssim \text{Dec}_d^{p,q}(C\delta) \left(\sum_{\theta} \|f_\theta\|_{L^p(w_B)}^q \right)^{1/q}$$

for balls B of radius R^2 . It is inconvenient that the weights on the left-hand and the right-hand side differ. This can be remedied by the averaging argument in Lemma 3.2 below, and the following estimate can be obtained for $q \leq p$.

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^p(w_B)} \lesssim \text{Dec}_d^{p,q}(\delta) \left(\sum_{\theta} \|f_\theta\|_{L^p(w_B)}^q \right)^{1/q}. \quad (3.9)$$

Similarly we can localize the rescaled decoupling inequality in Section 3.2.1 to balls of size $(\delta_0 \delta_1)^{-2}$.

3.3.1 Power weights

A key property of the weights (3.8) is the inequality

$$\mathbf{1}_B \lesssim \sum_{B' \in \mathcal{B}(B,R)} w_{B'} \lesssim w_B \quad (3.10)$$

that holds for all balls $B \subset \mathbb{R}^n$ and all $0 < R$ that are smaller than the radius of B . Here and later $\mathcal{B}(B, R)$ denotes a boundedly overlapping covering of a set B by balls of radius R . The implicit constants in (3.10) do not depend on B and R .

The following result allows to deduce inequalities for $L^p(w_B)$ norms from inequalities for $L^p(\mathbf{1}_B)$ norms. It is necessitated by the fact that inequalities converse to (3.10) do not hold.

Lemma 3.2 ([BD17, Lemma 4.1]). *Let \mathcal{W} be the collection of all weights, that is, positive, integrable functions on \mathbb{R}^n . Fix $R > 0$ and $E > n$. Let $O_1, O_2 : \mathcal{W} \rightarrow [0, \infty]$ be any functions with the following properties.*

1. $O_1(\mathbf{1}_B) \leq O_2(w_{B,E})$ for all balls $B \subset \mathbb{R}^n$ with radius R
2. $O_1(\alpha u + \beta v) \leq \alpha O_1(u) + \beta O_1(v)$, for each $u, v \in \mathcal{W}$ and $\alpha, \beta > 0$
3. $O_2(\alpha u + \beta v) \geq \alpha O_2(u) + \beta O_2(v)$, for each $u, v \in \mathcal{W}$ and $\alpha, \beta > 0$
4. If $u \leq v$ then $O_i(u) \leq O_i(v)$.
5. If $(u_j)_j \subset \mathcal{W}$ is a monotonically increasing sequence with $u_j \rightarrow u \in \mathcal{W}$ pointwise almost everywhere, then $O_1(u) = \lim_j O_1(u_j)$.

Then for each ball $B \subset \mathbb{R}^n$ with radius R we have

$$O_1(w_{B,E}) \lesssim_{n,E} O_2(w_{B,E})$$

The implicit constant depends only on n and E .

Proof. Let $\mathcal{B} := \mathcal{B}(\mathbb{R}^n, R)$. Note that

$$w_B(x) \leq C \sum_{B' \in \mathcal{B}} w_{B(c_{B'})} \mathbf{1}_{B'}(x)$$

and that

$$\sum_{B' \in \mathcal{B}} w_{B'}(x) w_{B(c_{B'})} \leq C w_B(x)$$

for a sufficiently large constant $C = C(n, E) > 0$. Hence

$$O_1(w_B) \leq \sup_{B' \subset \mathcal{B} \text{ finite}} O_1\left(C \sum_{B' \in B'} w_{B(c_{B'})} \mathbf{1}_{B'}\right) \quad \text{by (5)}$$

$$\leq \sup_{B' \subset \mathcal{B} \text{ finite}} C \sum_{B' \in B'} w_{B(c_{B'})} O_1(\mathbf{1}_{B'}) \quad \text{by (2)}$$

$$\leq C \sup_{B' \subset \mathcal{B} \text{ finite}} \sum_{B' \in B} w_{B(c_{B'})} O_2(w_{B'}) \quad \text{by (1)}$$

$$\leq C^2 \sup_{B' \subset \mathcal{B} \text{ finite}} O_2\left(C^{-1} \sum_{B' \in \mathcal{B}} w_{B'} w_{B(c_{B'})}\right) \quad \text{by (3)}$$

$$\leq C^2 O_2(w_B). \quad \square$$

Remark. Lemma 3.2 will be usually applied with functionals of the form

$$O_1(v) := \|f\|_{L^p(v)}^p \quad (3.11)$$

$$O_2(v) := A \left(\sum_i \|f_i\|_{L^p(v)}^q \right)^{\frac{p}{q}}, \quad (3.12)$$

where $1 \leq q \leq p$. It is clear that conditions (2) and (4) hold for these choices. The condition (3) follows from the reverse Minkowski inequality in $\ell_{\frac{p}{q}}$:

$$\begin{aligned} O_2(v+w) &= A \left(\sum_i \left(\int |f_i|^p v + \int |f_i|^p w \right)^{q/p} \right)^{p/q} \\ &\geq A \left(\sum_i \left(\int |f_i|^p v \right)^{q/p} \right)^{p/q} + A \left(\sum_i \left(\int |f_i|^p w \right)^{q/p} \right)^{p/q} = O_2(v) + O_2(w) \end{aligned}$$

Remark. We recall the proof of the reverse Minkowski inequality, which is basically identical to the proof of the direct Minkowski inequality. Let $0 < r \leq 1$ and let $f : X \times Y \rightarrow [0, \infty]$ be a function such that $0 < \|\int_Y f\|_{L^r(X)} < \infty$. Then by reverse Hölder inequality we have

$$\begin{aligned} \int_X \left| \int_Y f(x, y) dy \right|^r dx &= \int_Y \int_X f(x, y') \left| \int_Y f(x, y) dy \right|^{r-1} dx dy' \\ &\geq \int_Y \left(\int_X f(x, y')^r dx \right)^{1/r} \left(\int_X \left| \int_Y f(x, y) dy \right|^r dx \right)^{1-1/r} dy'. \end{aligned}$$

The reverse Hölder inequality is just the usual Hölder inequality in which one of the terms is brought to the other side. Rearranging we obtain

$$\left(\int_X \left| \int_Y f(x, y) dy \right|^r dx \right)^{1/r} \geq \int_Y \left(\int_X f(x, y')^r dx \right)^{1/r} dy'.$$

3.3.2 Reverse Hölder inequality

We close this section with the following reverse Hölder inequality.

Corollary 3.3 (cf. [BD17, Corollary 4.1]). *For each $1 \leq t \leq p < \infty$, each $E > n$, each $R > 0$ and $\delta > 0$ with $R\delta \geq 1$, each function $f : \mathbb{R}^n \rightarrow \mathbb{C}$ with $\text{diam}(\text{supp } \hat{f}) \lesssim \delta$, and each ball $B \subset \mathbb{R}^n$ with radius R we have*

$$\|f\|_{L^p(w_{B,E})} \lesssim (R\delta)^{n/t-n/p} \|f\|_{L^t(w_{B,\frac{E}{p}})}, \quad (3.13)$$

with the implicit constant independent of R , δ , B , and f .

Here and later we denote normalized L^p norms by

$$\|f\|_{L^p(B)} := |B|^{-1/p} \|f\|_{L^p(B)}, \quad \|f\|_{L^p(w_B)} := |B|^{-1/p} \|f\|_{L^p(w_B)}. \quad (3.14)$$

Proof. Let η be a positive Schwartz function on \mathbb{R}^n with $\mathbf{1}_{B(0,1)} \leq \eta$ and such that $\text{supp}(\hat{\eta}) \subset B(0,1)$. We can thus write

$$\|f\|_{L^p(B)} \leq \|\eta_B f\|_{L^p(\mathbb{R}^n)},$$

where η_B is an appropriate L^∞ -scaling and translation of η . Let θ be a Schwartz function on \mathbb{R}^n such that $\hat{\theta} = 1$ for $|\theta| \leq 10$. Since

$$\text{diam}(\text{supp } \widehat{\eta_B f}) \leq \text{diam}(\text{supp } \widehat{\eta_B}) + \text{diam}(\text{supp } \hat{f}) \lesssim 1/R + \delta \lesssim \delta,$$

we have that

$$\eta_B f = (\eta_B f) * \theta_B,$$

where θ_B is an appropriate L^1 -scaling and modulation of θ . By Young's convolution inequality with exponents

$$\frac{1}{p} = \frac{1}{t} + \frac{1}{r} - 1 = \frac{1}{t} - \frac{1}{r'}$$

we can write

$$\|\eta_B f\|_{L^p(\mathbb{R}^n)} \leq \|\eta_B f\|_{L^t(\mathbb{R}^n)} \|\theta_B\|_{L^r(\mathbb{R}^n)} \lesssim \delta^{n/r'} \|f\|_{L^t(\eta_B^t)}.$$

Rearranging this inequality and estimating $\eta_B \lesssim w_{B,E}^{1/p}$ we obtain

$$|B|^{-1/p} \|f\|_{L^p(B)} \lesssim_{n,E} (R\delta)^{n/r'} |B|^{-1/t} \|f\|_{L^t(w_{B,E}^{t/p})}$$

for any $E > 0$. Now we can apply Lemma 3.2 with

$$\begin{aligned} O_1(v) &:= R^{-n} \int |f|^p v, \\ O_2(v) &:= A(R\delta)^{n/t-n/p} R^{-np/t} \left(\int |f|^t v^{t/p} \right)^{p/t}. \end{aligned} \quad \square$$

3.4 Linear versus multilinear decoupling

3.4.1 Transversality

One of the main advantages of the decoupling inequality (3.3) it can be reduced to the corresponding multilinear inequality involving transverse pieces of the paraboloid. This is the objective of this Section 3.4.1. We fix $d \geq 1$, and if $d > 1$ we assume that Theorem 3.1 is already known with d replaced by $d - 1$. All constants are allowed to depend on d and p , but not on other parameters unless indicated.

Recall that in the multilinear restriction theorem we called subsets S_1, \dots, S_{d+1} of a d -dimensional hypersurface in \mathbb{R}^{d+1} ν -transverse if $|N_1 \wedge \dots \wedge N_{d+1}| > \nu$ for any unit normal vectors N_j to S_j . The same notion of transversality will be used here. However, we also have to understand what it means in terms of the parametrization of the paraboloid \mathbb{P} as the graph of $\Phi(\xi) = |\xi|^2$ on the unit cube $[0, 1]^d$. It is easy to see that $(-2\xi, 1) \in \mathbb{R}^d \times \mathbb{R}$ is a normal vector to \mathbb{P} at the point $(\xi, \Phi(\xi))$. Indeed,

$$(-2\xi, 1) \cdot \partial_j(\xi, \Phi(\xi)) = (-2\xi, 1) \cdot (e_j, 2\xi_j) = 0.$$

For ξ in the unit cube the length of the vector $(-2\xi, 1)$ is ~ 1 , so the pieces of the graph of Φ over $\alpha_1, \dots, \alpha_{d+1} \subseteq [0, 1]^d$ are ν -transverse iff for any $\xi_j \in \alpha_j$ we have

$$\begin{aligned} \nu < \left| \binom{-2\xi_1}{1} \wedge \dots \wedge \binom{-2\xi_{d+1}}{1} \right| &= \left| \binom{2\xi_{d+1} - 2\xi_1}{0} \wedge \dots \wedge \binom{2\xi_{d+1} - 2\xi_d}{0} \binom{-2\xi_{d+1}}{1} \right| \\ &= |2\xi_{d+1} - 2\xi_1 \wedge \dots \wedge 2\xi_{d+1} - 2\xi_d|. \end{aligned}$$

In this case we also call $\alpha_1, \dots, \alpha_{d+1} \subset [0, 1]^d$ ν -transverse.

For illustrative purposes we note that the quantity in the last display is comparable to the volume of the convex hull of ξ_1, \dots, ξ_{d+1} .

3.4.2 Notation

We will use the following notation for L^p norms and ℓ^q norms:

$$L^p_{x \in X} F(x) := \left(\int_{x \in X} |F(x)|^p \right)^{1/p}, \quad \ell^q_{\theta \in \mathcal{P}(\delta)} F(\theta) := \left(\int_{\theta \in \mathcal{P}(\delta)} |F(\theta)|^q \right)^{1/q}.$$

We will also use averaged L^p norms

$$\|f\|_{L^p(B)} := (|B|^{-1} \int_B |f|^p)^{1/p}, \quad \|f\|_{L^p(w_B)} := (|B|^{-1} \int |f|^p w_B)^{1/p}$$

Given f_θ , $\theta \in \mathcal{P}(\delta)$, we write here and later

$$f := \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \quad \text{and} \quad f_\alpha := \sum_{\theta \in \mathcal{P}(\alpha, \delta)} f_\theta$$

for dyadic cubes α of scale $\geq \delta$. For positive numbers A_1, \dots, A_{d+1} , we abbreviate

$$\prod\!\!\!\prod A_i := \left(\prod_{i=1}^{d+1} A_i \right)^{1/(d+1)}.$$

For $K \in 4^{\mathbb{N}}$, a positive number $\nu_K > 0$ that depends on d, K and will be specified in the proof of Proposition 9.6, and $0 < \delta < K^{-1}$ the *multilinear decoupling constant* $\text{MulDec}_d^{p,2}(\delta, K)$ is the smallest constant such that the inequality

$$L^p_{x \in \mathbb{R}^{d+1}} \prod\!\!\!\prod \|f_{\alpha_i}\|_{L^p(B(x, K))} \leq \text{MulDec}_d^{p,2}(\delta, K) \prod\!\!\!\prod \ell^2_{\theta \in \mathcal{P}(\alpha_i, \delta)} \|f_\theta\|_{L^p(\mathbb{R}^{d+1})} \quad (3.15)$$

holds for every ν_K -transverse tuple $\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1})$.

The left-hand side of (3.15) should be considered morally equivalent to $\|\prod\!\!\!\prod |f_{R_i}|\|_p$, since by the uncertainty principle the functions f_{R_i} are morally constant at scale K . However, as noticed in [BD17], inductive arguments are substantially simpler with an additional average.

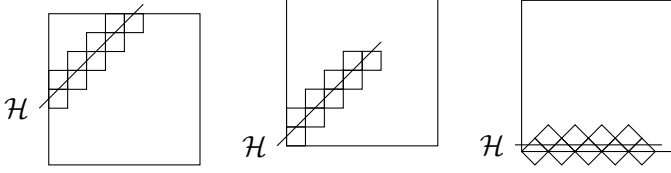
3.4.3 Lower dimensional decoupling

Lemma 3.4. *Let $2 \leq p \leq \infty$, $\mathcal{H} \subset \mathbb{R}^d$ be an affine hyperplane and $\sigma \in 2^{-\mathbb{N}}$. Then*

$$\left\| \sum_{\substack{\beta \in \mathcal{P}(\sigma), \\ 2\beta \cap \mathcal{H} \neq \emptyset}} f_\beta \right\|_p \lesssim \text{Dec}_{d-1}^{p,2}(\sigma) \ell_{\beta \in \mathcal{P}(\sigma)}^2 \|f_\beta\|_p. \quad (3.16)$$

Proof. For $d = 1$ the claim is trivial since the number of summands is bounded uniformly in σ , so we assume $d \geq 2$.

By an affine transformation we may assume that $\text{dist}(\mathcal{H}, 0) \lesssim \sigma$. Then by rotation in \mathbb{R}^d we may assume that \mathcal{H} is the hyperplane $\{\xi_d = \text{const}\}$ (to do that we may have to split β 's into finitely many collections and replace them by slightly larger cubes after the rotation).



Let $\beta \in \mathcal{P}(\sigma)$ be $C\sigma$ -close to \mathcal{H} and write $\beta = \beta' \times \beta_d$ with $\beta' \subset \mathbb{R}^{d-1}$ and $\beta_d \subset \mathbb{R}$. Denote by $\square(\beta)$ the box over β as in (3.2). The crucial geometric observation is that $\square(\beta)$ is contained in a $C\sigma^2$ -neighborhood of $\square(\beta') \times \beta_d$. Hence by (3.5) for each fixed $x_d \in \mathbb{R}$ we can apply the $(d-1)$ -dimensional case of (3.3) to the functions $f_\beta(\cdot, x_d, \cdot)$ of d variables:

$$\left\| \sum_{\beta} f_\beta(\cdot, x_d, \cdot) \right\|_{L^p(\mathbb{R}^d)} \lesssim \text{Dec}_{d-1}^{p,2}(\sigma) \ell_\beta^2 \|f_\beta(\cdot, x_d, \cdot)\|_{L^p(\mathbb{R}^d)}.$$

By Minkowski's inequality it follows that

$$\begin{aligned} \left\| \sum_{\beta} f_\beta \right\|_{L^p(\mathbb{R}^{d+1})} &= L_{x_d \in \mathbb{R}}^p \left\| \sum_{\beta} f_\beta(\cdot, x_d, \cdot) \right\|_{L^p(\mathbb{R}^d)} \lesssim \text{Dec}_{d-1}^{p,2}(\sigma) L_{x_d \in \mathbb{R}}^p \ell_\beta^2 \|f_\beta(\cdot, x_d, \cdot)\|_{L^p(\mathbb{R}^d)} \\ &\leq \text{Dec}_{d-1}^{p,2}(\sigma) \ell_\beta^2 L_{x_d \in \mathbb{R}}^p \|f_\beta(\cdot, x_d, \cdot)\|_{L^p(\mathbb{R}^d)} = \text{Dec}_{d-1}^{p,2}(\sigma) \ell_\beta^2 \|f_\beta\|_{L^p(\mathbb{R}^{d+1})}. \quad \square \end{aligned}$$

3.4.4 Bourgain–Guth argument

We let $d \geq 1$ and assume that Theorem 3.1 holds with d replaced by $d-1$. In the case $d = 1$ this hypothesis is vacuous.

From Hölder's inequality, it follows that

$$\text{MulDec}_d^{p,2}(\delta, K) \lesssim \text{Dec}_d^{p,2}(\delta). \quad (3.17)$$

The Bourgain–Guth argument shows that the converse inequality also holds up to some lower-dimensional terms. To be precise, we will prove

Proposition 3.5. *Let $2 \leq p \leq 2(d+1)/(d-1)$. Then for each $\varepsilon > 0$ there exists K such that*

$$\text{Dec}_d^{p,2}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon} + \delta^{-\varepsilon} \max_{\delta \leq \delta' \leq 1; \delta' \text{ dyadic}} \text{MulDec}_d^{p,2}(\delta', K). \quad (3.18)$$

Proposition 3.5 is proved by iterating the following result $O\left(\frac{\log \delta}{\log K}\right)$ many times after choosing K large enough depending on ε so that $C_\varepsilon \leq K^\varepsilon$.

Proposition 3.6. *Let $2 \leq p \leq 2(d+1)/(d-1)$ and $\varepsilon > 0$. Then for every K that is a power of 4 and $0 < \delta < 1/K$ we have*

$$\text{Dec}_d^{p,2}(\delta) \leq C_\varepsilon K^\varepsilon \text{Dec}_d^{p,2}(\delta K^{1/2}) + C_K \text{MulDec}_d^{p,2}(\delta, K). \quad (3.19)$$

Proof of Proposition 9.6. Fix functions $f_\theta, \theta \in \mathcal{P}(\delta)$. Let $B \subset \mathbb{R}^{d+1}$ be a ball of radius K and

$$S_B := \left(\sum_{\alpha \in \mathcal{P}(K^{-1})} \|f_\alpha\|_{L^p(B)}^2 \right)^{1/2}.$$

We distinguish two cases. The first case is that there exists an affine hyperplane \mathcal{H}_B of \mathbb{R}^d such that

$$\left\| \sum_{\beta \in \mathcal{P}(K^{-1/2}): 2\beta \cap \mathcal{H}_B = \emptyset} f_\beta \right\|_{L^p(B)} \leq S_B. \quad (3.20)$$

In this case we use Lemma 3.4 (for $d > 1$; for $d = 1$ triangle inequality suffices since there are only boundedly many summands in this case) and a simple localization argument as in Section 3.3 to obtain

$$\begin{aligned} \left\| \sum_{\substack{\beta \in \mathcal{P}(K^{-1/2}) \\ 2\beta \cap \mathcal{H}_B \neq \emptyset}} f_\beta \right\|_{L^p(B)} &\lesssim_\varepsilon K^\varepsilon \left(\sum_{\substack{\beta \in \mathcal{P}(K^{-1/2}) \\ 2\beta \cap \mathcal{H}_B \neq \emptyset}} \|f_\beta\|_{L^p(w_B)}^2 \right)^{1/2} \\ &\leq K^\varepsilon \left(\sum_{\beta \in \mathcal{P}(K^{-1/2})} \|f_\beta\|_{L^p(w_B)}^2 \right)^{1/2}. \end{aligned}$$

If (3.20) fails, then for every proper affine hyperplane \mathcal{H} of \mathbb{R}^d there is a dyadic cube $\alpha \in \mathcal{P}(K^{-1})$ such that α is at least $K^{-1/2}$ away from \mathcal{H} and $\|f_\alpha\|_{L^p(B)} \geq c_K S_B$. We can therefore inductively choose such $\alpha_1, \dots, \alpha_{d+1}$ in such a way that α_k is $K^{1/2}$ away from some affine hyperplane passing through $\alpha_1, \dots, \alpha_{k-1}$. In particular the collection $\alpha_1, \dots, \alpha_{d+1}$ is ν_K -transverse for some $\nu_K > 0$ depending only on d and K and

$$\|f\|_{L^p(B)} \leq C_K S_B \leq C_K \prod \|f_{\alpha_i}\|_{L^p(B)},$$

where C_K are constants depending only on d, K . Hence in both cases we obtain

$$\begin{aligned} \left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^p(B)} &\leq \left(\sum_{\alpha \in \mathcal{P}(K^{-1})} \|f_\alpha\|_{L^p(B)}^2 \right)^{1/2} + C_\varepsilon K^\varepsilon \left(\sum_{\beta \in \mathcal{P}(K^{-1/2})} \|f_\beta\|_{L^p(w_B)}^2 \right)^{1/2} \\ &\quad + C_K \sum_{\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1})} \prod \|f_{\alpha_i}\|_{L^p(B)}, \end{aligned}$$

where the latter sum runs over all ν_K -transverse tuples. Replacing all $L^p(B)$ norms by averaged versions $L^p(B)$ and integrating this inequality over all K -balls $B \subset \mathbb{R}^{d+1}$ we obtain

$$\begin{aligned} \|f\|_{L^p(\mathbb{R}^{d+1})} &= L_{x \in \mathbb{R}^{d+1}}^p \|f\|_{L^p(B(x, K))} \\ &\leq L_{x \in \mathbb{R}^{d+1}}^p \left(\sum_{\alpha \in \mathcal{P}(K^{-1})} \|f_\alpha\|_{L^p(B)}^2 \right)^{1/2} \end{aligned} \quad (3.21)$$

$$+ C_\varepsilon K^\varepsilon L_{x \in \mathbb{R}^{d+1}}^p \left(\sum_{\beta \in \mathcal{P}(K^{-1/2})} \|f_\beta\|_{L^p(w_{B(x, K)})}^2 \right)^{1/2} \quad (3.22)$$

$$+ C_K \sum_{\substack{\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} L_{x \in \mathbb{R}^{d+1}}^p \prod \|f_{\alpha_i}\|_{L^p(B(x, K))}, \quad (3.23)$$

In the term (3.21), by Minkowski's inequality and scaling we obtain

$$\begin{aligned}
(3.21) &\leq \left(\sum_{\alpha \in \mathcal{P}(K^{-1})} (I_{x \in \mathbb{R}^{d+1}}^p \|f_\alpha\|_{L^p(B)})^2 \right)^{1/2} \\
&= \left(\sum_{\alpha \in \mathcal{P}(K^{-1})} \|f_\alpha\|_{L^p(\mathbb{R}^{d+1})}^2 \right)^{1/2} \\
&\leq \text{Dec}_d^{p,2}(\delta K) \left(\sum_{\alpha \in \mathcal{P}(\delta)} \|f_\alpha\|_{L^p(\mathbb{R}^{d+1})}^2 \right)^{1/2}.
\end{aligned}$$

The same argument is also applied to (3.22). Note that by scaling $\text{Dec}_d^{p,2}(\delta K) \leq \text{Dec}_d^{p,2}(\delta K^{1/2})$, so the estimate for (3.21) can be absorbed in the estimate for (3.22).

In the last term (3.23) by definition of the multilinear decoupling constant (3.15) we have

$$\begin{aligned}
(3.23) &\leq C_K \text{MulDec}_d^{p,2}(\delta, K) \sum_{\substack{\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} \prod \left(\sum_{\theta \in \mathcal{P}(\alpha_i, \delta)} \|f_\theta\|_{L^p(\mathbb{R}^{d+1})}^2 \right)^{1/2} \\
&\leq C_K \text{MulDec}_d^{p,2}(\delta, K) \left(\sum_{\theta \in \mathcal{P}(\delta)} \|f_\theta\|_{L^p(\mathbb{R}^{d+1})}^2 \right)^{1/2},
\end{aligned}$$

since $\mathcal{P}(\alpha_i, \delta) \subset \mathcal{P}(\delta)$ and there are only C_K choices of $\alpha_1, \dots, \alpha_{d+1}$. \square

3.5 Bourgain–Demeter iteration

We will use two different moves to estimate the left-hand side of (3.15):

1. L^2 orthogonality. This move allows to pull ℓ_τ^2 norms out of the inner L^p norm. This only works for $p = 2$ and at an appropriate spatial scale given by the uncertainty principle.
2. Multilinear Kakeya. This move allows to increase the radius of integration in the inner L^p norm so that we have a chance of applying L^2 orthogonality again.

These moves only work for specific combinations of Lebesgue exponents and scales.

3.5.1 L^2 orthogonality

For every $0 < \delta < 1$ and every ball $B \subset \mathbb{R}^{d+1}$ of radius δ^{-1} we have

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^2(w_B)} \lesssim \ell_{\theta \in \mathcal{P}(\delta)}^2 \|f_\theta\|_{L^2(w_B)}. \quad (3.24)$$

It is important that this estimate holds already on balls of radius δ^{-1} given by the uncertainty principle.

To see (3.24) let η be a bump function adapted to the ball B with $|\eta| \sim 1$ on B . Then

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^2(B)} \lesssim \left\| \sum_{\theta \in \mathcal{P}(\delta)} \eta f_\theta \right\|_{L^2(\mathbb{R}^{d+1})} \lesssim \ell_{\theta \in \mathcal{P}(\delta)}^2 \|\eta f_\theta\|_{L^2(\mathbb{R}^{d+1})} \lesssim \ell_{\theta \in \mathcal{P}(\delta)}^2 \|f_\theta\|_{L^2(w_B)}$$

by Plancherel's theorem, since the Fourier supports $\widehat{\eta f_\theta}$ have bounded overlap (in fact, their projections onto \mathbb{R}^d already have bounded overlap). The estimate (3.24) now follows from Lemma 3.2.

3.5.2 Ball inflation

Lemma 3.7 (Ball inflation, $\ell^t L^t$ version). *Let $0 < \rho \leq K^{-1}$. Let $\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1})$ be a ν -transverse collection of cubes. Let $B \subset \mathbb{R}^{d+1}$ be a ball of radius ρ^{-2} . Then for each $1 \leq t < \infty$ and $\varepsilon > 0$ we have*

$$L_{x \in B}^{\frac{d+1}{d}t} \prod \ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^t \|f_{\tau_i}\|_{L^t(w_{B(x, 1/\rho)})} \lesssim_{\nu, \varepsilon} \rho^{-\varepsilon} \prod \ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^t \|f_{\tau_i}\|_{L^t(w_B)}. \quad (3.25)$$

Proof. For each $\tau \in \mathcal{P}(\rho)$ with center c_τ we cover \mathbb{R}^{d+1} by a family \mathbf{T}_τ of pairwise disjoint boxes (tubes) with d short sides of length ρ^{-1} and 1 long side of length ρ^{-2} pointing in the normal direction $N_{\mathbb{P}}(c_\tau)$. Let $\mathbf{T}_\tau(B) := \{T \in \mathbf{T}_\tau \mid T \cap B \neq \emptyset\}$ and let $T_\tau(x) \in \mathbf{T}_\tau$ denote the unique tube containing x . By the multilinear Keakeya inequality (Corollary 1.4) we obtain

$$\begin{aligned} & LHS(3.25)^{(d+1)t/d} \\ &= |B|^{-1} \int_{x \in B} \prod_{i=1}^{d+1} \left(\sum_{\tau \in \mathcal{P}(\alpha_i, \rho)} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t \right)^{1/d} \\ &\lesssim_{\varepsilon, \nu} |B|^{-1} \rho^{-\varepsilon} \rho^{-(d+1)} \prod_{i=1}^{d+1} \left(\sum_{\tau \in \mathcal{P}(\alpha_i, \rho)} \sum_{T \in \mathbf{T}_\tau(B)} \sup_{x \in T} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t \right)^{1/d} \\ &\sim \rho^{-\varepsilon} \prod_{i=1}^{d+1} \left(\sum_{\tau \in \mathcal{P}(\alpha_i, \rho)} \rho^d \sum_{T \in \mathbf{T}_\tau(B)} \sup_{x \in T} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t \right)^{1/d}. \end{aligned}$$

To conclude the proof it remains to verify that

$$\rho^d \sum_{T \in \mathbf{T}_\tau(B)} \sup_{x \in T} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t \lesssim \|f_\tau\|_{L^t(w_B)}^t.$$

This inequality is morally true because each f_τ is constant at scale $1/\rho$ and the number of summands is of the order ρ^{-d} , so the left-hand side is a normalized L^t norm. We make this explanation precise. Let ψ_τ be an L^1 normalized adapted bump function such that $f_\tau = f_\tau * \psi_\tau$. Then

$$\begin{aligned} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t &\sim \rho^{d+1} \int |f_\tau * \psi_\tau|^t(u) w_{B(x, 1/\rho)}(u) du \\ &\leq \rho^{d+1} \|\psi_\tau\|_{L^1}^{t-1} \int (|f_\tau|^t * |\psi_\tau|)(u) w_{B(x, 1/\rho)}(u) du \\ &\lesssim \rho^{d+1} \int |f_\tau|^t(u) (|\psi_\tau| * w_{B(x, 1/\rho)})(u) du. \end{aligned}$$

For $x \in T \in \mathbf{T}_\tau$ we have $\rho^{d+1} |\psi_\tau| * w_{B(y, \rho^{-1})} \lesssim \tilde{w}_T$, where \tilde{w}_T is an L^1 normalized bump function adapted to T , hence $\|\tilde{w}_T\|_\infty \sim \rho^{d+2}$, so

$$\begin{aligned} \rho^d \sum_{T \in \mathbf{T}_\tau(B)} \sup_{x \in T} \|f_\tau\|_{L^t(w_{B(x, 1/\rho)})}^t &\lesssim \rho^d \sum_{T \in \mathbf{T}_\tau(B)} \int |f_\tau|^t \tilde{w}_T \\ &\lesssim \rho^{2d+2} \int |f_\tau|^t w_B \sim \|f_\tau\|_{L^t(w_B)}^t. \quad \square \end{aligned}$$

Corollary 3.8 (Ball inflation, $\ell^q L^t$ version). *In the setting of Lemma 9.8 let $1 \leq q \leq t < \infty$. Then*

$$L_{x \in B}^{\frac{d+1}{d}t} \prod \ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^q \|f_{\tau_i}\|_{L^t(w_{B(x, 1/\rho)})} \lesssim_{\nu, \varepsilon} \rho^{-\varepsilon} \prod \ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^q \|f_{\tau_i}\|_{L^t(w_B)}. \quad (3.26)$$

Proof. Let

$$\mathcal{P}_{i,l} := \left\{ \tau \in \mathcal{P}(\alpha_i, \rho) \mid 2^{-l-1} < \frac{\|f_\tau\|_{\mathcal{E}^t(w_B)}}{\max_{\tau' \in \mathcal{P}(\alpha_i, \rho)} \|f_{\tau'}\|_{\mathcal{E}^t(w_B)}} \leq 2^{-l} \right\}$$

and partition

$$\mathcal{P}(\alpha_i, \rho) = \left(\bigcup_{l=0}^{\tilde{l}} \mathcal{P}_{i,l} \right) \cup \mathcal{P}_{i,\text{rest}}$$

with $\tilde{l} \sim \log \rho^{-1}$. Then

$$\ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^q \|f_{\tau_i}\|_{\mathcal{E}^t(w_{B(x,1/\rho)})} \leq \sum_{l_i \in \{0, \dots, \tilde{l}\} \cup \{\text{rest}\}} \ell_{\tau_i \in \mathcal{P}_{i,l_i}}^q \|f_{\tau_i}\|_{\mathcal{E}^t(w_{B(x,1/\rho)})},$$

and since there are at most $C \log \rho^{-1}$ summands we may restrict the ℓ^q norm on the left-hand side of (3.25) to $\tau_i \in \mathcal{P}_{i,l_i}$ for some choice of $l_i \in \{0, \dots, \tilde{l}\} \cup \{\text{rest}\}$. Since $q \leq t$, by Hölder's inequality and Lemma 9.8 we obtain

$$\begin{aligned} \mathcal{E}_{x \in B}^{\frac{d+1}{d}t} \prod \ell_{\tau_i \in \mathcal{P}_{i,l_i}}^q \|f_{\tau_i}\|_{\mathcal{E}^t(w_{B(x,1/\rho)})} \\ \leq \left(\prod |\mathcal{P}_{i,l_i}|^{1/q-1/t} \right) \mathcal{E}_{x \in B}^{\frac{d+1}{d}t} \prod \ell_{\tau_i \in \mathcal{P}_{i,l_i}}^t \|f_{\tau_i}\|_{\mathcal{E}^t(w_{B(x,1/\rho)})} \\ \lesssim_{\nu, \varepsilon} \rho^{-\varepsilon} \left(\prod |\mathcal{P}_{i,l_i}|^{1/q-1/t} \right) \prod \ell_{\tau_i \in \mathcal{P}_{i,l_i}}^t \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)}. \end{aligned}$$

It remains to show that

$$|\mathcal{P}_{i,l}|^{1/q-1/t} \ell_{\tau_i \in \mathcal{P}_{i,l}}^t \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)} \lesssim \ell_{\tau_i \in \mathcal{P}(\alpha_i, \rho)}^q \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)}$$

for every $l \in \{0, \dots, \tilde{l}\} \cup \{\text{rest}\}$. For $l \in \{0, \dots, \tilde{l}\}$ we have in fact

$$|\mathcal{P}_{i,l}|^{1/q-1/t} \ell_{\tau_i \in \mathcal{P}_{i,l}}^t \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)} \sim |\mathcal{P}_{i,l}|^{1/q} \max_{\tau_i \in \mathcal{P}_{i,l}} \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)} \sim \ell_{\tau_i \in \mathcal{P}_{i,l}}^q \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)},$$

while for $l = \text{rest}$ we have

$$|\mathcal{P}_{i,\text{rest}}|^{1/q-1/t} \ell_{\tau_i \in \mathcal{P}_{i,\text{rest}}}^t \|f_{\tau_i}\|_{\mathcal{E}^t(w_B)} \lesssim |\mathcal{P}(\alpha_i, \rho)|^{1/q} 2^{-\tilde{l}} \max_{\tau' \in \mathcal{P}(\alpha_i, \rho)} \|f_{\tau'}\|_{\mathcal{E}^t(w_B)},$$

and the claim follows provided that \tilde{l} is a sufficiently large multiple of $\log \rho^{-1}$. \square

3.5.3 Proof of Theorem 3.1 for $p = 2 \frac{d+1}{d}$

The proof of Theorem 3.1 in the range $2 \leq p \leq 2 \frac{d+1}{d}$ is easier than in the general case and also historically it appeared earlier [Bou13]. By interpolation it suffices to consider $p = 2 \frac{d+1}{d}$.

In view of Proposition 3.5 it suffices to estimate the multilinear decoupling constant. Throughout this section let $\alpha_1, \dots, \alpha_{d+1} \in \mathcal{P}(K^{-1})$ be ν_K -transverse cubes. For $\rho \in 2^{-\mathbb{N}}$ we define the quantity

$$A(\rho) := L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{\mathcal{E}^2(w_{B(x,1/\rho)})}. \quad (3.27)$$

We caution the reader that the quantities denoted by A in [BD17] would correspond to our A with L_x^p replaced by $\mathcal{E}_{x \in B}^p$ for a large ball B .

Proposition 3.9. *We have for $p = 2 \frac{d+1}{d}$ and $\varepsilon > 0$*

$$A(\rho) \lesssim_{K, \varepsilon} \rho^{-\varepsilon} A(\rho^2) \quad (3.28)$$

Proof. Using ball inflation (Lemma 9.8) with $t = 2$ we obtain

$$\begin{aligned} A(\rho) &= L_x^p E_{\tilde{x} \in B(x, 1/\rho^2)}^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^2(w_{B(\tilde{x}, 1/\rho)})} \\ &\lesssim_{K, \varepsilon} \rho^{-\varepsilon} L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^2(w_{B(x, 1/\rho^2)})} \end{aligned}$$

By L^2 orthogonality we have

$$\|f_\tau\|_{L^2(w_{B(x, 1/\rho^2)})} \lesssim \ell_{\tau' \in \mathcal{P}(\tau, \rho^2)}^2 \|f_{\tau'}\|_{L^2(w_{B(x, 1/\rho^2)})}, \quad (3.29)$$

and the conclusion follows. \square

Now we relate the quantity A with the multilinear decoupling inequality. Let $\rho \leq (2K)^{-1}$. Then

$$\begin{aligned} LHS(3.15) &= L_x^p E_{\tilde{x} \in B(x, 1/(2\rho))}^p \prod \|f_{\alpha_i}\|_{L^p(B(\tilde{x}, K))} \\ &\text{by Hölder} \leq L_x^p \prod E_{\tilde{x} \in B(x, 1/(2\rho))}^p \|f_{\alpha_i}\|_{L^p(B(\tilde{x}, K))} \\ &\lesssim L_x^p \prod \|f_{\alpha_i}\|_{L^p(B(x, 1/\rho))} \\ &\text{by Minkowski} \leq L_x^p \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \rho)}^1 \|f_\theta\|_{L^p(B(x, 1/\rho))} \\ &\text{by Hölder} \leq \rho^{-d/2} L_x^p \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\theta\|_{L^p(B(x, 1/\rho))} \\ &\text{by reverse Hölder} \lesssim \rho^{-d/2} L_x^p \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\theta\|_{L^2(w_{B(x, 1/\rho)})} \\ &= \rho^{-d/2} A(\rho) \end{aligned} \quad (3.30)$$

On the other hand, for any ρ by Hölder's and Minkowski's inequalities we obtain

$$\begin{aligned} A(\rho) &\leq \prod L_x^p \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^p(w_{B(x, 1/\rho)})} \leq \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 L_x^p \|f_\tau\|_{L^p(w_{B(x, 1/\rho)})} \\ &\sim \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^p(\mathbb{R}^{d+1})}. \end{aligned} \quad (3.31)$$

Using (3.30), iterating (3.28) m times, and using (3.31) at the end we obtain for $\delta = \rho^{2^m}$ the estimate

$$LHS(3.15) \lesssim_{K, m, \varepsilon} \rho^{-d/2} \rho^{-\varepsilon} \dots \rho^{-2^{m-1}\varepsilon} \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\theta\|_{L^p(\mathbb{R}^{d+1})},$$

so that

$$\text{MulDec}_d^{p, 2}(\delta, K) \lesssim_{K, m, \varepsilon} \delta^{-d2^{-m+1}-\varepsilon}.$$

Choosing m large enough this gives $\text{MulDec}_d^{p, 2}(\delta, K) \lesssim_{K, \varepsilon} \delta^{-\varepsilon}$, and this suffices to conclude the proof of Theorem 3.1.

3.5.4 Proof of Theorem 3.1 for $p = 2\frac{d+2}{d}$

For $p > 2\frac{d+1}{d}$ we will have to apply ball inflation with some $t > 2$. This necessitates some additional branching of the estimates since orthogonality is only available in L^2 .

Proposition 3.10. *We have for $p = 2\frac{d+2}{d}$ and $\varepsilon > 0$*

$$A(\rho) \lesssim_{K, \varepsilon} \rho^{-\varepsilon} A(\rho^2)^{1/2} \left(\text{Dec}_d^{p, 2}(\delta/\rho) \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\theta\|_p \right)^{1/2} \quad (3.32)$$

Proof. Using ball inflation (Corollary 3.8) with $q = 2$ and $t = pd/(d+1)$ we obtain

$$\begin{aligned} A(\rho) &= L_x^p E_{\tilde{x} \in B(x, 1/\rho^2)}^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^2(w_{B(\tilde{x}, 1/\rho)})} \\ &\text{by Hölder} \lesssim L_x^p E_{\tilde{x} \in B(x, 1/\rho^2)}^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^t(w_{B(\tilde{x}, 1/\rho)})} \\ &\text{by Cor. 3.8} \lesssim_{K, \varepsilon} \rho^{-\varepsilon} L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^t(w_{B(x, 1/\rho^2)})} \\ &\text{by Hölder} \leq \rho^{-\varepsilon} \left(L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^2(w_{B(x, 1/\rho^2)})} \right)^{1/2} \\ &\quad \cdot \left(L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{L^p(w_{B(x, 1/\rho^2)})} \right)^{1/2}. \end{aligned}$$

Here we used $1/t = (1/2)(1/2) + (1/2)(1/p)$. By L^2 orthogonality (3.29) the first bracket is $\lesssim A(\rho^2)$. In the second bracket we estimate

$$\begin{aligned}
& L_x^p \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{\mathcal{L}^p(w_{B(x, 1/\rho^2)})} \\
& \text{by Hölder} \leq \prod L_x^p \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_{\mathcal{L}^p(w_{B(x, 1/\rho^2)})} \\
& \text{by Minkowski} \leq \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 L_x^p \|f_\tau\|_{\mathcal{L}^p(w_{B(x, 1/\rho^2)})} \\
& \lesssim \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 \|f_\tau\|_p \\
& \text{by scaling} \lesssim \prod \ell_{\tau \in \mathcal{P}(\alpha_i, \rho)}^2 (\text{Dec}_d^{p,2}(\delta/\rho) \ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 \|f_\theta\|_p) \\
& = \text{Dec}_d^{p,2}(\delta/\rho) \prod \ell_{\theta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\theta\|_p. \quad \square
\end{aligned}$$

Lemma 3.11. *Let $p = 2^{\frac{d+2}{d}}$ and suppose that*

$$\text{Dec}_d^{p,2}(\delta) \leq C\delta^{-\eta} \quad (3.33)$$

for some $0 < \eta < \infty$ and some $0 < C < \infty$. Then for every K we have

$$\text{MulDec}^p(\delta, K) \lesssim_{K, \eta, C} \delta^{-\eta(1-2^{-d/\eta-1})}.$$

Proof. Choose ν_K -transverse $\alpha_1, \dots, \alpha_M \in \mathcal{P}(1/K)$. Choose functions f_θ with

$$\ell_{\theta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\theta\|_p = 1.$$

Let $m \in \mathbb{N}$ be chosen later. Using (3.30), iterating (3.32) m times, and using (3.31) at the end we obtain for $\delta = \rho^{2^m}$ the estimate

$$\begin{aligned}
LHS (3.15) & \lesssim_{m, K, \varepsilon} \rho^{-d/2} \prod_{l=0}^{m-1} \left(\rho^{-2^l \varepsilon} \text{Dec}(\delta/\rho^{2^l})^{1/2} \right)^{2^{-l}} \\
& \lesssim \rho^{-d/2} \prod_{l=0}^{m-1} (\delta/\rho^{2^l})^{-2^{-l-1}\eta} \rho^{-\varepsilon} \\
& = \rho^{-d/2-m\varepsilon} \prod_{l=0}^{m-1} \rho^{\eta/2} \delta^{-2^{-l-1}\eta} \\
& = \rho^{-d/2-m\varepsilon+m\eta/2} \delta^{-(1-2^{-m})\eta}.
\end{aligned}$$

If $m\eta > d$ and ε is small enough, then this is $\leq \delta^{-(1-2^{-m})\eta}$, and the claim follows. \square

Now we conclude the proof of Theorem 3.1 for $p = 2^{\frac{d+2}{d}}$. It is easy to see that

$$\text{Dec}_d^{p,2}(\delta) \lesssim \delta^{-\eta} \quad (3.34)$$

for some $\eta < \infty$. Suppose that (3.34) holds for some $\eta > 0$. We will decrease η . Substituting the conclusion of Lemma 3.11 into the conclusion of Proposition 3.5 gives

$$\text{Dec}_d^{p,2}(\delta) \lesssim_\varepsilon \delta^{-\varepsilon} + \delta^{-\varepsilon-(1-2^{-d/\eta-1})\eta} \quad (3.35)$$

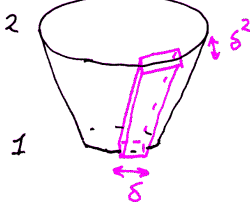
for any $\varepsilon > 0$. Iterating the passage from (3.34) to (3.35) we can make η arbitrarily close to 0.

4 Local smoothing for the wave equation

4.1 Decoupling for the truncated cone

Solutions of the wave equation on \mathbb{R}^d have Fourier support on a cone. For this reason a decoupling theorem for functions with Fourier support near the truncated cone $\mathcal{C} := \{(\xi, |\xi|) \mid \xi \in \mathbb{R}^d, 1 \leq |\xi| \leq 2\}$ will be useful.

Theorem 4.1 ([BD15, Section 8]). *Let $d \geq 2$, $2 \leq p \leq \infty$, let Θ be a boundedly overlapping collection of slabs of size $1 \times \delta \times \dots \times \delta \times \delta^2$ adapted to \mathcal{C} as in the picture.*



Then for any collection of functions f_θ , $\theta \in \Theta$, with $\text{supp } \widehat{f}_\theta \subset \theta$ and $\varepsilon > 0$ we have

$$\left\| \sum_{\theta \in \Theta} f_\theta \right\|_p \lesssim_\varepsilon \delta^{-\varepsilon - \eta_p} \ell_{\theta \in \Theta}^2 \|f_\theta\|_p, \quad (4.1)$$

where

$$\eta_p = \begin{cases} 0, & 2 \leq p \leq 2(d+1)/(d-1), \\ (d-1)/2 - (d+1)/p, & 2(d+1)/(d-1) < p \leq \infty. \end{cases}$$

Proof. By interpolation, which works similarly to the case of the paraboloid, it suffices to consider the critical exponent $p = 2(d+1)/(d-1)$. Alternatively one could treat all exponents directly.

The main idea is that thin sectors of the cone project to something very close to parabolas along the flat direction of the cone. Hence we should be able to apply the decoupling theorem for the parabola in dimension $d-1$ fiberwise. However, the projections are not sufficiently close to the parabola to apply decoupling at scale δ right away. Instead we have to set up an induction on scales argument.

Let M be a positive integer. For $1 \leq m \leq M$ let Θ_m be a boundedly overlapping covering of the thinned cone

$$\mathcal{C}' := \{(\xi, |\xi|) \mid \xi \in \mathbb{R}^d, |\xi_1 - 1| \leq \delta^{2/M}, |\xi_2| \leq \delta^{1/M}, \dots, |\xi_d| \leq \delta^{1/M}\}$$

by adapted slabs of size $\sim \delta^{2/M} \times \delta^{m/M} \times \dots \times \delta^{m/M} \times \delta^{2m/M}$. Partition

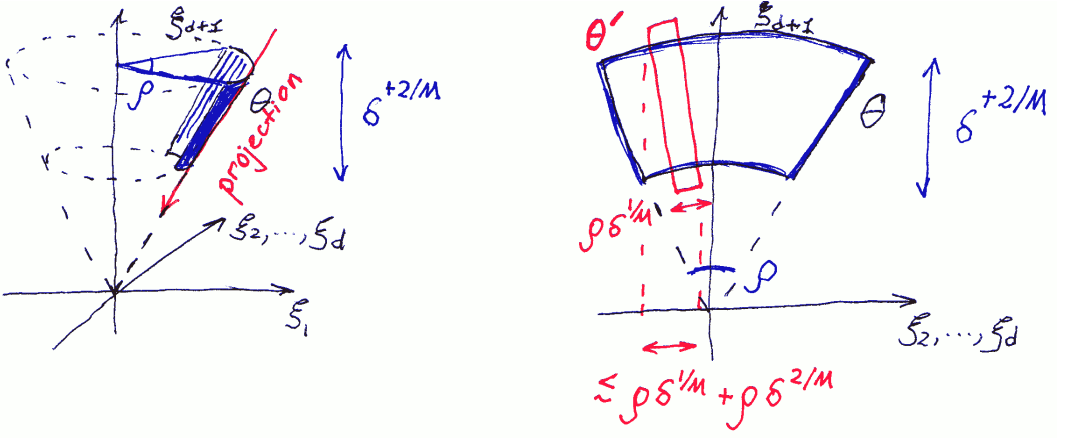
$$\Theta_{m+1} = \cup_{\theta \in \Theta_m} \Theta_{m+1}(\theta)$$

in such a way that for $\theta' \in \Theta_{m+1}(\theta)$ we have $\theta' \cap \theta \neq \emptyset$. Let $f_{\theta'}$, $\theta' \in \Theta_M$ be a collection of functions with $\text{supp } \widehat{f}_{\theta'} \subset \theta'$ and define

$$f_\theta := \sum_{\theta' \in \Theta_{m+1}(\theta)} f_{\theta'}$$

for $m = M-1, \dots, 1$ and $\theta \in \Theta_m$. We claim that for each $1 \leq m < M$ and each $\theta \in \Theta_m$ we have

$$\|f_\theta\|_p \lesssim_{\varepsilon, M} \delta^{-\varepsilon} \ell_{\theta' \in \Theta_{m+1}(\theta)}^2 \|f_{\theta'}\|_p. \quad (4.2)$$



Iterating the claim (4.2) we obtain

$$\left\| \sum_{\theta' \in \Theta_M} f_{\theta'} \right\|_p \leq \sum_{\theta \in \Theta_1} \|f_{\theta}\|_p \lesssim \delta^{-\frac{d-1}{2M}} \ell_{\theta \in \Theta_1}^2 \|f_{\theta}\|_p \lesssim_{\varepsilon, M} \delta^{-\frac{d-1}{2M} - \varepsilon} \ell_{\theta' \in \Theta_M}^2 \|f_{\theta'}\|_p.$$

Since we can partition the cone \mathcal{C} in $O(\delta^{-2/M})$ dilates of the cone \mathcal{C}' and partition functions f_{θ} accordingly, the conclusion (4.1) follows with ε replaced by $C/M + \varepsilon$. Since M was arbitrary we obtain (4.1).

It remains to show the claim (4.2). By rotation invariance we may assume that θ is adapted to the cone segment

$$\mathcal{C}_{\rho} := \{(\xi, |\xi|) \mid \xi \in \mathbb{R}^d, |\xi_1 - 1| \leq \delta^{2/M}, |\xi_2| \leq \rho, \dots, |\xi_d| \leq \rho\}$$

with $\rho = \delta^{m/M}$. Under the projection map

$$(\xi_1, \dots, \xi_d, \xi_{d+1}) \mapsto (\xi_2, \dots, \xi_d, \xi_{d+1} - \xi_1)$$

the parametrization of the cone \mathcal{C}_{ρ} is mapped to

$$(\xi_2, \dots, \xi_d, \sqrt{\xi_1^2 + \dots + \xi_d^2} - \xi_1).$$

Now notice that

$$\begin{aligned} \left| \sqrt{\xi_1^2 + \dots + \xi_d^2} - \xi_1 - \frac{\xi_2^2 + \dots + \xi_d^2}{2} \right| &= \left| \frac{\sqrt{\xi_1^2 + \dots + \xi_d^2} - \xi_1}{\sqrt{\xi_1^2 + \dots + \xi_d^2} + \xi_1} - \frac{\xi_2^2 + \dots + \xi_d^2}{2} \right| \\ &\lesssim \rho^2 (\sqrt{\xi_1^2 + \dots + \xi_d^2} + \xi_1 - 2) \lesssim \rho^2 |\xi_1 - 1| + \rho^4 \lesssim \rho^2 \delta^{2/M}, \end{aligned}$$

so that an $O(\delta^2)$ -neighborhood of \mathcal{C}_{ρ} , and hence θ , is projected to an $O(\rho^2 \delta^{2\nu})$ -neighborhood of a paraboloid of dimension $d - 1$. Moreover, the projections of slabs $\theta' \in \Theta_{m+1}(\theta)$ to the coordinates (ξ_2, \dots, ξ_d) are boundedly overlapping sets of diameter $O(\rho \delta^{1/M})$. Thus (4.2) follows by a fiberwise application of the (rescaled) decoupling theorem for the paraboloid at scale $\rho \delta^{1/M}$. Here it becomes clear why we had to restrict ξ_1 to an interval of length $\delta^{2/M}$: otherwise we could only decouple at scale ρ , and this would not bring us forward in the induction on scales. \square

4.2 Estimates for the pieces of a solution

Let $u : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{C}$ be the solution of the initial value problem

$$\partial_t^2 u = \Delta u, \quad u|_{t=0} = f, \quad \partial_t u|_{t=0} = 0. \quad (4.3)$$

Suppose that \widehat{f} is supported in the annulus $\{\xi \in \mathbb{R}^d \mid 2^n \leq |\xi| \leq 2^{n+1}\}$ with $n \geq 0$. Let also χ be a fixed Schwartz function on \mathbb{R} with compact Fourier support. Then the

function $u(x, t)\chi(t)$ has Fourier support in an $O(1)$ -neighborhood of the truncated cone over the annulus. Hence by Theorem 4.1 and scaling

$$\|u(x, t)\chi(t)\|_{L^p(\mathbb{R}^{d+1})} \lesssim_\varepsilon 2^{(\eta_p + \varepsilon)n/2} \ell_{\theta \in \Theta}^2 \|u_\theta\|_p, \quad (4.4)$$

where Θ is a boundedly overlapping covering of the neighborhood of truncated cone by slabs of size $2^n \times 2^{n/2} \times \cdots \times 2^{n/2} \times 1$ and $u(x, t)\chi(t) = \sum_\theta u_\theta(x, t)$ is a subordinate partition.

Now we estimate the right-hand side of (4.4).

Lemma 4.2. *We have $\|u_\theta\|_\infty \lesssim \|f\|_\infty$.*

Proof. Let A_θ be a smooth cutoff to a sector of angular width $2^{-n/2}$ in \mathbb{R}^d in the direction of θ and $\hat{\eta}$ a function with Fourier support in an annulus such that $\hat{\eta}(\xi) = 1$ if $1 \leq |\xi| \leq 2$. Then

$$u_\theta(x, t) = \chi(t) \int K_\theta(x, t, y) f(y) dy,$$

where

$$K_\theta(x, t, y) = \int_{\mathbb{R}^d} A_\theta(\xi) \hat{\eta}(2^{-n}\xi) e((x - y) \cdot \xi + t|\xi|) d\xi.$$

We claim

$$\|K_\theta(\cdot, t, y)\|_1 \lesssim 1 + |t|^C \quad (4.5)$$

uniformly in n, θ, y . This is proved by an integration by parts argument that first appeared in [SSS91, p. 240]. By translation invariance it suffices to consider $y = 0$, and by a change of variable

$$\begin{aligned} LHS(4.5) &= \left\| 2^{nd} \int_{\mathbb{R}^d} A(\xi) \hat{\eta}(\xi) e(2^n x \cdot \xi + 2^n t |\xi|) d\xi \right\|_{L_x^1} \\ &= \left\| \int_{\mathbb{R}^d} A(\xi) \hat{\eta}(\xi) e(x \cdot \xi + 2^n t |\xi|) d\xi \right\|_{L_x^1}, \end{aligned}$$

where we used homogeneity of $A = A_\theta$. By rotation invariance we may assume that the support of A is centered at $(1, 0, \dots, 0)$. Now we write

$$\int_{\mathbb{R}^d} A(\xi) \hat{\eta}(\xi) e(x \cdot \xi + 2^n t |\xi|) d\xi = \int_{\mathbb{R}^d} \underbrace{A(\xi) \hat{\eta}(\xi) e(2^n t (|\xi| - \xi_1))}_{=: b(\xi)} e(x \cdot \xi + 2^n t \xi_1) d\xi.$$

Consider the self-adjoint differential operator

$$L = (1 - \partial_1^2 - 2^{-n} \partial_2^2 - \cdots - 2^{-n} \partial_d^2).$$

We claim that

$$\|L^N b\|_\infty \lesssim_N 1 + |t|^{2N}. \quad (4.6)$$

Assuming this claim we can write

$$e(x \cdot \xi) = (1 - (2\pi i)^2 |x_1|^2 - (2\pi i)^2 2^{-n} (|x_2|^2 + \cdots + |x_d|^2))^{-N} L^N e(x \cdot \xi + 2^n t \xi_1),$$

and by integration by parts we obtain

$$\begin{aligned} &\left| \int_{\mathbb{R}^d} b(\xi) e(x \cdot \xi) d\xi \right| \\ &\sim (1 + |x_1|^2 + 2^{-n} (|x_2|^2 + \cdots + |x_d|^2))^{-N} \left| \int_{\mathbb{R}^d} b(\xi) L^N e(x \cdot \xi) d\xi \right| \\ &= (1 + |x_1|^2 + 2^{-n} (|x_2|^2 + \cdots + |x_d|^2))^{-N} \left| \int_{\mathbb{R}^d} (L^N b)(\xi) e(x \cdot \xi) d\xi \right| \\ &\leq (1 + |x_1|^2 + 2^{-n} (|x_2|^2 + \cdots + |x_d|^2))^{-N} \|\text{supp } b\| \|L^N b\|_\infty \\ &\lesssim_N (1 + |x_1|^2 + 2^{-n} (|x_2|^2 + \cdots + |x_d|^2))^{-N} 2^{-(d-1)n/2} (1 + |t|^{2N}). \end{aligned}$$

The estimate (4.5) now follows upon taking N large enough.

It remains to show (4.6). To this end it suffices to show

$$|\partial_1^{N_1}(2^{-n/2}\partial_2)^{N_2}\dots(2^{-n/2}\partial_d)^{N_d}\tilde{b}| \lesssim_{N_1,\dots,N_d} 1 + |t|^{N_2+\dots+N_d} \text{ on } \text{supp } b \quad (4.7)$$

for $\tilde{b} = A, \eta, e(2^{nt}(|\xi| - \xi_1))$.

Since $|\xi|, |\xi|^{-1}$ are smooth functions of ξ on $\text{supp } \beta$ we obtain (4.7) for $\tilde{b} = |\xi|$ and $\tilde{b} = |\xi|^{-1}$. Since η is a smooth function of $|\xi|$ we obtain (4.7) for $\tilde{b} = \eta$. The function A can be written as $A(\xi) = \tilde{A}(\xi_2/|\xi|, \dots, \xi_d/|\xi|)$, where $|\nabla^M \tilde{A}| \lesssim 2^{Mn/2}$. By the chain rule each time when ∂_1 hits \tilde{A} we get a factor $\xi_j \xi_1 / |\xi|^3$ with $j \neq 1$. If no $2^{-n/2}\partial_j$ hits this factor, it compensates the derivative of \tilde{A} that we gained because $|\xi_j| \lesssim 2^{-n/2}$. If some $2^{-n/2}\partial_j$ hits this factor, then the derivative of \tilde{A} is compensated by the $2^{-n/2}$ factor from the derivative.¹ Hence we obtain (4.7) for $\tilde{b} = A$.

The function $\tilde{b}(\xi) = e(2^{nt}(|\xi| - \xi_1))$ is the composition of $r(\xi) = |\xi| - \xi_1$ and a function \tilde{B} whose M -th derivative is bounded by $(2^{nt})^M$. Notice first that

$$\partial_1 r(\xi) = (\xi_1 - |\xi|)/|\xi| = -r(\xi)/|\xi|,$$

so any higher derivative $\partial_1^{N_1} r$ is again r times some polynomial in $|\xi|$ and $|\xi|^{-1}$. Also, $|r(\xi)| \lesssim |\xi_2|^2 + \dots + |\xi_d|^2 \lesssim 2^{-n}$ compensates the factor 2^n from the derivative of \tilde{B} . Notice next that for $j \neq 1$ we have

$$\partial_1^{N_1}(2^{-n/2}\partial_j)r(\xi) = \partial_1^{N_1}2^{-n/2}\xi_j/|\xi| = 2^{-n/2}\xi_j(\partial_1^{N_1}|\xi|^{-1}) = O(n^{-1}),$$

and this is enough to compensate the the factor 2^n from the derivative of \tilde{B} . Finally, if at least two derivatives $(2^{-n/2}\partial_j)$ with $j \neq 1$ hit r , then they together bring a factor 2^{-n} that is enough to compensate the the factor 2^n from the derivative of \tilde{B} . Hence we obtain (4.7) in the last remaining case. \square

Corollary 4.3 ([Wol00, Lemma 6.1]). *For $2 \leq p \leq \infty$ we have $\ell_\theta^p \|u_\theta\|_p \lesssim \|f\|_p$.*

Proof. By Lemma 4.2 and complex interpolation² it remain to consider $p = 2$. But in this case we can first decompose f into pieces with Fourier support in sectors, which only requires Plancherel, and use the fact that the wave equation preserves the L^2 norm of each piece.³ \square

By (4.4), Hölder's inequality, and Corollary 4.3 we obtain

$$\begin{aligned} \|u(x, t)\chi(t)\|_{L^p(\mathbb{R}^{d+1})} &\lesssim_\varepsilon 2^{(\eta_p+\varepsilon)n/2} \ell_{\theta \in \Theta}^2 \|u_\theta\|_p \\ &\lesssim 2^{(\eta_p+(d-1)(1/2-1/p)+\varepsilon)n/2} \ell_{\theta \in \Theta}^p \|u_\theta\|_p \\ &\lesssim 2^{(\eta_p+(d-1)(1/2-1/p)+\varepsilon)n/2} \|f\|_p. \end{aligned}$$

Using this estimate for $n \geq 1$ and a trivial estimate for functions with Fourier support in the unit ball we get

Theorem 4.4 (Local smoothing). *For $\alpha > (\eta_p + (d-1)(1/2 - 1/p))/2$ we have*

$$\|u(x, t)\chi(t)\|_{L^p(\mathbb{R}^{d+1})} \lesssim \|f\|_{W^{p,\alpha}}. \quad (4.8)$$

In particular for $p \geq 2(d+1)/(d-1)$ Theorem 4.4 holds in the range $\alpha > \frac{d-1}{2} - \frac{d}{p}$.

¹To make this argument precise one should use a multivariate Faá di Bruno's formula.

²the Fourier support restriction to the annulus has to be replaced by a smooth Fourier cutoff to get estimates on full L^p spaces

³The argument in Lemma 4.2 actually works in every L^p space, so one could also apply that to each piece.

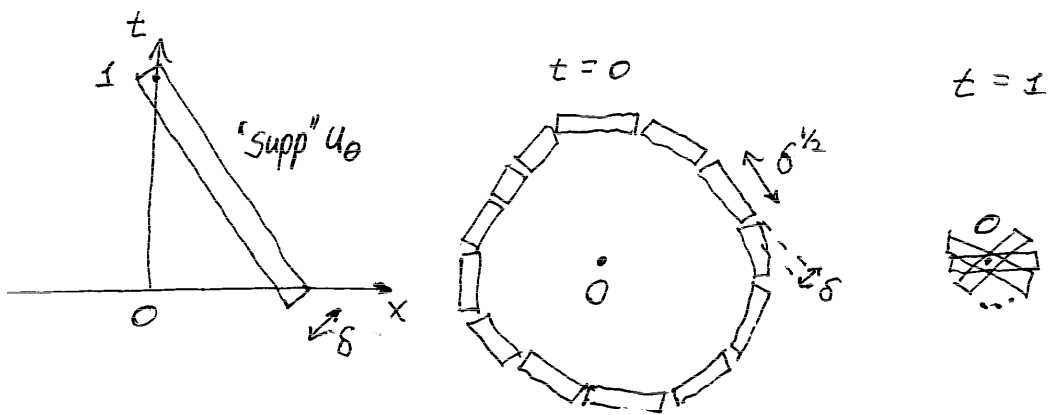


Figure 2: Focusing example

4.3 Examples and the local smoothing conjecture

Let $0 < \delta < 1$. Let θ be a sector of the truncated cone $\{(\xi, |\xi|) \mid |\xi| \sim \delta^{-1}\}$ with opening angle $\sim \delta^{1/2}$. Then θ is contained in a box of dimensions $1 \times \delta^{1/2} \times \dots \times \delta^{1/2} \times \delta$. Take a bump function on θ , multiply with the surface measure on the cone, and denote the Fourier transform of the result by u_θ . Then u_θ is a bump function adapted to a box of dimensions $1 \times \delta^{-1/2} \times \dots \times \delta^{-1/2} \times \delta^{-1}$. It decays rapidly in the directions of long sides and at with power $-1/2$ in the direction of the short side (by stationary phase). Hence $u_\theta \in L^p$ for $p > 2$. We may normalize it in L^∞ and shift it without changing the Fourier support. In particular we make sure that $u_\theta \sim 1$ on a δ -neighborhood of $(0, \dots, 0, 1) \in \mathbb{R}^{d+1}$ and $u_\theta \sim 1$ on a δ -neighborhood of a $\delta^{-1/2}$ -arc of the unit sphere in \mathbb{R}^d for $t = 1$.

Then, summing over a disjoint collection of θ 's of cardinality $\approx \delta^{-(d-1)/2}$ we obtain

$$\left\| \sum_{\theta} u_{\theta}(\cdot, t) \right\|_{L^p(\mathbb{R}^d)} \gtrsim \delta^{-(d-1)/2} \delta^{d/p}$$

for $|t - 1| \lesssim \delta$. On the other hand,

$$\left\| \sum_{\theta} u_{\theta}(\cdot, 0) \right\|_{W^{\alpha,p}} \sim \delta^{1/p-\alpha}.$$

Hence the fixed time estimate

$$\|u(x, t)\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{W^{p,\alpha}}$$

can only hold if $-(d-1)/2 + d/p \geq 1/p - \alpha$, or in other words $\alpha \geq \frac{d-1}{2} - \frac{d-1}{p}$. The same example shows that the local smoothing estimate (4.8) can only hold for $\alpha \geq \frac{d-1}{2} - \frac{d}{p}$. Since by Theorem 4.4 the local smoothing estimate (4.8) does hold in this range for sufficiently large p , it is better than what can be obtained from estimates for fixed times. However, despite the name “local smoothing”, the solution can still be less smooth than the initial data. It seems that the name “local smoothing” comes from the Schrödinger equation, for which it is indeed the case that, on average over time, the solution is smoother than the initial data, see references in [Rog08].

For a single θ the norm $\|u_{\theta}(\cdot, t)\|_{L^p(\mathbb{R}^d)}$ is morally constant for $|t| \lesssim 1$. Hence the local smoothing estimate (4.8) can only hold for $\alpha \geq 0$. It is conjectured that these two examples are essentially the worst.

Conjecture 4.5 ([Sog91]). *For $d \geq 2$ the estimate (4.8) holds for $\alpha > \max(\frac{d-1}{2} - \frac{d}{p}, 0)$.*

Conjecture 4.5 is not known in any dimension. For $d \geq 3$ the best partial result is Theorem 4.4 (and an endpoint estimate [HNS11] for large p). For $d = 2$ two more results are known based on multilinear restriction [LV12; Lee16]. A possible route to Conjecture 4.5 is via the chain of inequalities

$$\|u\|_{L^p(\mathbb{R}^{d+1})} \lesssim \|\ell_\theta^2 u_\theta\|_{L^p(\mathbb{R}^{d+1})} \lesssim \|\ell_\theta^2 f_\theta\|_{L^p(\mathbb{R}^d)} \lesssim \|f\|_{L^p(\mathbb{R}^d)} \quad (4.9)$$

for the critical exponent $p = \frac{2d}{d-1}$, where \lesssim means that the implicit constant grows is $\lesssim_\varepsilon \delta^{-\varepsilon}$ for every $\varepsilon > 0$. The first inequality in this chain would be stronger than the decoupling theorem (Theorem 4.1) since $L^p \ell^2 \leq \ell^2 L^p$. The partial results concern this inequality. The other two inequalities are in fact known for $d = 2$, $p = 4$.

Last inequality in (4.9) is proved in [Cór82] (same argument is repeated in [Cór83]) using a square function estimate from [Cór81] and the estimate for the maximal function with N directions in \mathbb{R}^d .

5 Cone square function

In this section we show a square function estimate of the form $L^4(\mathbb{R}^2) \rightarrow L^4(\mathbb{R}^3, \ell^2)$ for solutions of the wave equation that is effectively due to [MSS92].

Lemma 5.1 ([MSS92, (1.8)]). *Let $d = 2$. Let*

$$T_\theta f(x, t) := \chi(t) \int_{\mathbb{R}^2} K_\theta(x, t, y) f(y) dy.$$

Then for any functions $f_{\theta, l}$ we have

$$\left\| \left(\sum_{\theta, l} |T_{\theta, l} f_{\theta, l}|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^3)} \lesssim (n+1)^{5/4} \left\| \left(\sum_{\theta, l} |f_{\theta, l}|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^2)}. \quad (5.1)$$

In [MSS92, (1.8)] this result is stated for functions $f_{\theta, l}$ of a special form, but the proof works for arbitrary functions.

Proof. We use the classical fact that vector-valued inequalities follow from maximal inequalities. By duality for some function $g \in L^2(\mathbb{R}^3)$ with $\|g\|_2 = 1$ we have

$$LHS(5.1)^2 = \int_{\mathbb{R}^3} \sum_{\theta, l} |T_{\theta, l} f_{\theta, l}|^2 g.$$

Since the kernels $\tilde{K}_\theta(x, t, \cdot) = K_\theta(x, t, \cdot)\chi(t)$ are uniformly integrable and by Hölder's inequality this is

$$\begin{aligned} &= \int_{\mathbb{R}^3} \sum_{\theta, l} \left| \int_{\mathbb{R}^2} \tilde{K}_\theta(x, t, y) f_{\theta, l}(y) dy \right|^2 g(x, t) dx dt \\ &\lesssim \int_{\mathbb{R}^3} \sum_{\theta, l} \int_{\mathbb{R}^2} |\tilde{K}_\theta(x, t, y)| |f_{\theta, l}(y)|^2 dy |g(x, t)| dx dt \\ &= \int_{\mathbb{R}^2} \sum_{\theta, l} |f_{\theta, l}(y)|^2 \left(\int_{\mathbb{R}^3} |\tilde{K}_\theta(x, t, y)| |g(x, t)| dx dt \right) dy \\ &\leq \int_{\mathbb{R}^2} \sum_{\theta, l} |f_{\theta, l}(y)|^2 \left(\sup_{\theta} \int_{\mathbb{R}^3} |\tilde{K}_\theta(x, t, y)| |g(x, t)| dx dt \right) dy \\ &\leq \left\| \sum_{\theta, l} |f_{\theta, l}(y)|^2 \right\|_{L_y^2(\mathbb{R}^2)} \left\| \sup_{\theta} \int_{\mathbb{R}^3} |\tilde{K}_\theta(x, t, y)| |g(x, t)| dx dt \right\|_{L_y^2(\mathbb{R}^2)}. \end{aligned}$$

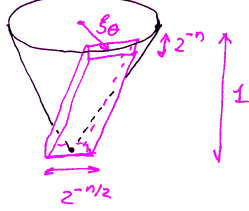
Hence it remains to show the maximal inequality

$$\left\| \sup_{\theta} \int_{\mathbb{R}^3} |\tilde{K}_{\theta}(x, t)| |g(x, t)| dx dt \right\|_{L_y^2(\mathbb{R}^2)} \lesssim (n+1)^{5/2} \|g\|_{L^2(\mathbb{R}^3)}.$$

Let ξ_{θ} be a unit vector in the direction of θ . When estimating K_{θ} we have actually proved

$$|\tilde{K}_{\theta}(x, t, y)| \lesssim_N 2^{3n/2} (1+2^n |\langle y-x, \xi_{\theta} \rangle + t|)^{-N} (1+2^{n/2} |y-x - \langle y-x, \xi_{\theta} \rangle \xi_{\theta}|)^{-N} (1+|t|)^{-N}.$$

So the claimed maximal inequality will follow by scaling from a maximal inequality involving averaging over slabs of dimensions $\sim 2^{-n} \times 2^{-n/2} \times 1$ as in the picture.



We can write the average over such slab as an average over a rectangle of size $2^{-n} \times 2^{-n/2}$ in \mathbb{R}^2 of averages over tubes of length 1, diameter 2^{-n} , and slope 1. Hence our maximal operator is bounded by the composition of a maximal operator with 2^n separated directions in \mathbb{R}^2 and a maximal operator from \mathbb{R}^3 to \mathbb{R}^2 involving tubes as in Lemma 5.2. \square

Lemma 5.2 ([MSS92, Lemma 1.4]).

$$\left\| \sup_{|\xi|=1} \left| \int_{-1}^1 \int_{B(0,1)} g(y - \delta x - \xi t, t) dx dt \right| \right\|_{L_y^2(\mathbb{R}^2)} \lesssim (\log_* \delta)^{3/2} \|g\|_{L^2(\mathbb{R}^3)}. \quad (5.2)$$

Lemma 5.2 extends, up to a worse exponent of $\log_* \delta$, (the basic version of) the maximal estimate for separated directions in \mathbb{R}^2 by setting $g(x, t) = g_0(x) \mathbf{1}_{[-1,1]}(t)$. However, no proof of Lemma 5.2 based on covering arguments seems to be available. The proof uses a Sobolev embedding, TT^* , and a stationary phase decomposition.

Proof. We replace ball averages by smooth averages

$$A_{\alpha} g(y) := \int_{-1}^1 \int_{\mathbb{R}^2} g(y - x + (\cos \alpha, \sin \alpha)t, t) \delta^{-2} \hat{a}(\delta^{-1}x) dx dt,$$

where $\hat{a} \geq 0$ and a has compact support. Let \tilde{g} denote the Fourier transform of g in the x variable. Then

$$A_{\alpha} g(y) = \int_{-1}^1 \int_{\mathbb{R}^2} e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) \tilde{g}(\xi, t) a(\delta \xi) d\xi dt.$$

We decompose $a(\delta \cdot) = a_0 + \sum_{\lambda} a_{\lambda}$, where a_0 is a bump function supported on the unit ball and each a_{λ} is a bump function supported on a dyadic annulus of radius $\sim \lambda$, so that there are $\sim \log_* \delta$ different λ 's. The contribution of a_0 can be easily estimated by going back to the spatial formulation.

Hence it suffices to show that for

$$A_{\alpha}^{\lambda} g(y) := \int_{-1}^1 \int_{\mathbb{R}^2} e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) \tilde{g}(\xi, t) a_{\lambda}(\xi) d\xi dt$$

we have the estimates

$$\left\| \sup_{\alpha} |A_{\alpha}^{\lambda} g| \right\|_{L^2(\mathbb{R}^2)} \lesssim (\log_* \lambda) \|g\|_{L^2(\mathbb{R}^3)}.$$

Indeed, summing these estimates we would obtain

$$\sum_{\lambda} (\log_* \lambda) \|P_{\lambda} g\|_{L^2(\mathbb{R}^3)} \lesssim (\log_* \delta)^{3/2} \left(\sum_{\lambda} \|P_{\lambda} g\|_{L^2(\mathbb{R}^3)}^2 \right)^{1/2} \lesssim (\log_* \delta)^{3/2} \|P_{\lambda} g\|_{L^2(\mathbb{R}^3)}$$

where P_{λ} is the Fourier projection to the support of a_{λ} . In order to show the claim we decompose further

$$a_{\lambda} = a_{\lambda,0} + \sum_{0 < k \lesssim \log_* \lambda} a_{\lambda,k},$$

where

$$a_{\lambda,k}(\xi, \alpha) = a_{\lambda}(\xi) \beta(2^{-k} \lambda^{1/2} \langle (-\sin \alpha, \cos \alpha), \frac{\xi}{|\xi|} \rangle)$$

and β is a smooth function supported on $\pm[1/2, 2]$ such that $\sum_{k \in \mathbb{Z}} \beta(2^{-k} s) = 0$ for $s \neq 0$. Let

$$A_{\alpha}^{\lambda,k} g(y) := \int_{-1}^1 \int_{\mathbb{R}^2} e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) \tilde{g}(\xi, t) a_{\lambda,k}(\xi, \alpha) d\xi dt.$$

It suffices to show

$$\| \sup_{\alpha} |A_{\alpha}^{\lambda,k} g| \|_{L^2(\mathbb{R}^2)} \lesssim \|g\|_{L^2(\mathbb{R}^3)}.$$

An estimate for fixed α is relatively easy. To proceed we use the inequality

$$\begin{aligned} \sup_{\alpha \in [0, 2\pi]} |F(\alpha)^2 - F(0)^2| &= \sup_{\alpha \in [0, 2\pi]} \left| 2 \int_{\alpha'=0}^{\alpha} F(\alpha') F'(\alpha') d\alpha' \right| \\ &\leq 2 \int_{\alpha=0}^{2\pi} |F(\alpha)| |F'(\alpha)| d\alpha \leq 2 \left(\int_{\alpha=0}^{2\pi} |F(\alpha)|^2 d\alpha \right)^{1/2} \left(\int_{\alpha=0}^{2\pi} |F'(\alpha)|^2 d\alpha \right)^{1/2}. \end{aligned}$$

Applying this inequality with $F(\alpha) = A_{\alpha}^{\lambda,k} g(y)$ we see that it suffices to show

$$\left(\int_{\mathbb{R}^2} \int_{\alpha=0}^{2\pi} |A_{\alpha}^{\lambda,k} g(y)|^2 d\alpha dy \right) \left(\int_{\mathbb{R}^2} \int_{\alpha=0}^{2\pi} |\partial_{\alpha} A_{\alpha}^{\lambda,k} g(y)|^2 d\alpha dy \right) \lesssim \|g\|_{L^2(\mathbb{R}^3)}^4. \quad (5.3)$$

Consider the first bracket. Expanding the square we obtain

$$\begin{aligned} &\int_{\mathbb{R}^2} \int_{\alpha=0}^{2\pi} \int e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle - \langle y + (\cos \alpha, \sin \alpha)t', \xi' \rangle) \\ &\quad \cdot \tilde{g}(\xi, t) a_{\lambda,k}(\xi, \alpha) \overline{\tilde{g}(\xi', t') a_{\lambda,k}(\xi', \alpha)} d\xi d\xi' dt dt' d\alpha dy \\ &\quad = \int e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) \\ &\quad \cdot \tilde{g}(\xi, t) a_{\lambda,k}(\xi, \alpha) \mathcal{F}(\overline{\tilde{g}(\cdot, t') a_{\lambda,k}(\cdot, \alpha)})(y + (\cos \alpha, \sin \alpha)t') d\xi dt dt' d\alpha dy \\ &\quad = \int e((t - t') \langle (\cos \alpha, \sin \alpha), \xi \rangle) \\ &\quad \cdot \tilde{g}(\xi, t) a_{\lambda,k}(\xi, \alpha) \overline{\tilde{g}(\xi, t') a_{\lambda,k}(\delta \xi)} d\xi dt dt' d\alpha \\ &\quad = \int H^{\lambda,k}(t, t', \xi) \tilde{g}(\xi, t) \overline{\tilde{g}(\xi, t')} d\xi dt dt', \end{aligned}$$

where

$$H^{\lambda,k}(t, t', \xi) = \int_{\alpha=0}^{2\pi} e((t - t') \langle (\cos \alpha, \sin \alpha), \xi \rangle) |a_{\lambda,k}(\xi, \alpha)|^2 d\alpha.$$

We claim that this function satisfies

$$|H^{\lambda,k}(t, t', \xi)| \lesssim_N \lambda^{-1/2} 2^k (1 + 2^{2k} |t - t'|)^{-N}, \quad |\xi| \sim \lambda, |t| \leq 1, |t'| \leq 1. \quad (5.4)$$

For $k > 0$ this is an oscillatory integral estimate. Indeed, $a_{\lambda,k}(\xi, \cdot)$ is supported on a set of measure $\sim 2^k \lambda^{-1/2}$ and its m -th derivative is bounded by $(2^{-k} \lambda^{1/2})^m$. So after

the change of variable $\alpha = \alpha' 2^k \lambda^{-1/2}$ the amplitude becomes a smooth function with uniform bounds on the support and all derivatives. The phase becomes

$$(t - t') |\xi| \langle (\cos(2^k \lambda^{-1/2} \alpha'), \sin(2^k \lambda^{-1/2} \alpha')), \frac{\xi}{|\xi|} \rangle.$$

By construction the first derivative of the phase is

$$\begin{aligned} (t - t') |\xi| 2^k \lambda^{-1/2} \langle (-\sin(2^k \lambda^{-1/2} \alpha'), \cos(2^k \lambda^{-1/2} \alpha')), \frac{\xi}{|\xi|} \rangle \\ \sim (t - t') |\xi| 2^k \lambda^{-1/2} (2^k \lambda^{-1/2}) \sim 2^{2k} (t - t') \end{aligned}$$

on the support of the phase. On the other hand, for $m \geq 2$ the m -th derivative of the phase is bounded by

$$(t - t') |\xi| (2^k \lambda^{-1/2})^m \lesssim (t - t') |\xi| (2^k \lambda^{-1/2})^2 = 2^{2k} (t - t').$$

Hence (5.4) follows by partial integration. Using (5.4) we obtain

$$\begin{aligned} \int |H^{\lambda, k}(t, t', \xi)| |\tilde{g}(\xi, t) \tilde{g}(\xi, t')| d\xi dt dt' \\ \lesssim_N \lambda^{-1/2} 2^k \int (1 + 2^{2k} |t - t'|)^N |\tilde{g}(\xi, t) \tilde{g}(\xi, t')| d\xi dt dt' \\ \lesssim \lambda^{-1/2} 2^k 2^{-2k} \|g\|_2^2 = \lambda^{-1/2} 2^{-k} \|g\|_2^2 \end{aligned}$$

provided that $N > 1$. The estimate for the term involving ∂_α is similar with the following changes: when in

$$\partial_\alpha \left(e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) a_{\lambda, k}(\xi, \alpha) \right)$$

the derivative falls on $a_{\lambda, k}$ we lose an additional factor $2^{-k} \lambda^{1/2}$. If on the other hand the derivative falls on the exponential we get

$$t |\xi| 2^k \lambda^{-1/2} e(\langle y + (\cos \alpha, \sin \alpha)t, \xi \rangle) \underbrace{2^{-k} \lambda^{1/2} \langle (-\sin \alpha, \cos \alpha), \frac{\xi}{|\xi|} \rangle}_{\text{underlined}} a_{\lambda, k}(\xi, \alpha).$$

The underlined function can be absorbed in $a_{\lambda, k}$ by changing the function β , and we lose a factor $t |\xi| 2^k \lambda^{-1/2} \lesssim 2^k \lambda^{1/2}$. Since there are two derivatives ∂_α we actually lose $(2^k \lambda^{1/2})^2$, and we obtain

$$\int_{\mathbb{R}^2} \int_{\alpha=0}^{2\pi} |\partial_\alpha A_\alpha^{\lambda, k} g(y)|^2 d\alpha dy \lesssim \lambda^{-1/2} 2^{-k} (2^k \lambda^{1/2})^2 \|g\|_2^2.$$

Multiplying this with the previous estimate gives (5.3) for $k > 0$.

Finally, the $k = 0$ case of the estimate (5.4) is even easier because we do not claim any decay in $|t - t'|$, so no partial integration is needed. \square

6 Reverse square function estimate for the cone in \mathbb{R}^{2+1}

For a dyadic number $\delta \in 2^{-\mathbb{Z}}$ we denote by $\mathcal{P}(\delta)$ a boundedly overlapping covering of the truncated cone \mathcal{C} in \mathbb{R}^{2+1} by adapted $1 \times \delta^{1/2} \times \delta$ -slabs (TODO: change to $1 \times \delta \times \delta^2$). One can think of them as being associated to arcs of angular length $2\pi\delta$ on the circle, but such dyadic structure is not preserved by affine scaling that we will use. We fix partitions

$$\mathcal{P}(\delta/2) = \cup_{\theta \in \mathcal{P}(\delta)} \mathcal{P}(\theta, \delta/2)$$

such that $\theta' \in \mathcal{P}(\theta, \delta/2)$ only if $\theta' \cap \theta \neq \emptyset$. Then for any $\delta' < \delta$ and $\theta \in \mathcal{P}(\delta)$ we can define $\mathcal{P}(\theta, \delta')$ in such a way that for $\delta'' < \delta' < \delta$ we have

$$\mathcal{P}(\theta, \delta'') = \cup_{\theta' \in \mathcal{P}(\theta, \delta')} \mathcal{P}(\theta', \delta'').$$

Given f_θ for $\theta \in \mathcal{P}(\delta_0)$ with $\text{supp } \widehat{f}_\theta \subseteq \theta$ for some small δ_0 we define for $\delta_0 < \delta \leq 1$ and $\tau \in \mathcal{P}(\delta)$

$$f_\tau := \sum_{\theta \in \mathcal{P}(\tau, \delta_0)} f_\theta, \quad f := \sum_{\theta \in \mathcal{P}(\delta_0)} f_\theta.$$

Let us denote by $\text{RvSq}^p(\delta)$ (for ‘‘reverse square function estimate’’) the smallest constant such that the estimate

$$\|f\|_p \leq \text{RvSq}^p(\delta) \|\ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta\|_p \quad (6.1)$$

holds for any f_θ as above.

6.1 Lorentz scaling

The cone has a scaling symmetry similar to that of the parabola. Let $0 < \gamma < 1/100$ be a small number and $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be the linear transformation given by

$$T(1, 0, 1) = (1, 0, 1), \quad T(-1, 0, 1) = \gamma^2(-1, 0, 1), \quad T(0, 1, 0) = \gamma(0, 1, 0).$$

Then in the standard coordinates T is given by the matrix

$$\begin{pmatrix} \frac{1+\gamma^2}{2} & 0 & \frac{1-\gamma^2}{2} \\ 0 & \gamma & 0 \\ \frac{1-\gamma^2}{2} & 0 & \frac{1+\gamma^2}{2} \end{pmatrix},$$

and one can verify that it preserves the cone given by the equation $\xi_1^2 + \xi_2^2 = \xi_3^2$. Also, it maps the δ/γ^2 -neighborhood of the truncated cone sector

$$\{(\xi, |\xi|) \mid \xi \in \mathbb{R}^2, 1/2 \leq |\xi| \leq 4, \xi_1 > 1/2\}$$

onto the δ -neighborhood of a sector of the truncated cone \mathcal{C} of width $\sim \gamma$ near $(1, 0, 1)$.

Let $\tau \in \mathcal{P}(\gamma^2)$ be the slab pointing in the direction $(1, 0, 1)$ and $\theta \in \mathcal{P}(\tau, \delta)$. Then $T^{-1}\theta$ is a slab of scale δ/γ^2 adapted to the cone. We apply (6.1) to the functions $f_\theta \circ T^{-t}$. Since $\widehat{f_\theta \circ T^{-t}} = |\det T| \widehat{f_\theta} \circ T$ we obtain

$$\|f_\tau\|_p \lesssim \text{RvSq}^p(\delta/\gamma^2) \|\ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\theta\|_p. \quad (6.2)$$

By rotation invariance this estimate continues to hold for arbitrary $\tau \in \mathcal{P}(\gamma^2)$. Abusing notation we will write

$$S_\delta f := \ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta, \quad S_\delta f_\tau := \ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\theta.$$

6.2 Multilinear reverse square function estimate

6.2.1 Transversality

Let $\mathcal{C}_\tau := \mathcal{C} \cap 5\tau$ be the piece of the truncated cone close to $\tau \in \mathcal{P}(\gamma)$. Denote by $N(\xi)$ the unit normal vector to \mathcal{C} at ξ (say, inward). We call $\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)$ ν -transverse if for every $\xi_i \in \mathcal{C}_{\tau_i}$ we have

$$|N(\xi_1) \wedge N(\xi_2) \wedge N(\xi_3)| \geq \nu, \quad (6.3)$$

see Figure 3. A key geometric property of the cone \mathcal{C} is that if $\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)$ are mutually separated by a distance $d > 10\gamma^{1/2}$, then they are cd^3 -transverse. To see

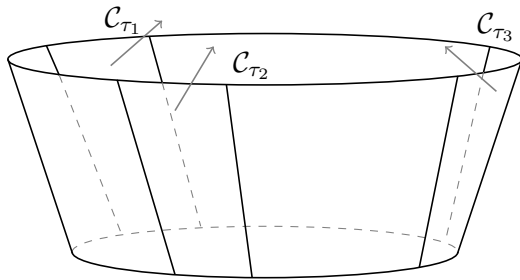


Figure 3: Transverse sectors on the cone (copied from [Lee16]).

this notice that for $\xi \in \mathbb{R}^2$ we have $N((\xi, |\xi|)) = (-\xi/|\xi|, 1)$, so we wedge product in (6.3) is proportional to the area of a certain triangle by the same computation as for the paraboloid in \mathbb{R}^{3+1} . This triangle has side lengths $\gtrsim d$ and vertices on the unit circle, so its area is $\gtrsim d^3$.

We will denote by $\text{MIRvSq}^p(\delta, \gamma)$ (for “multilinear reverse square function estimate”) the smallest constant such that the estimate

$$\left\| \prod_{i=1}^3 \|f_{\tau_i}\| \right\|_p \leq \text{MIRvSq}^p(\delta, \gamma) \prod_{i=1}^3 \|S_\delta f_{\tau_i}\|_p \quad (6.4)$$

holds for all $10\gamma^{1/2}$ -separated triples $\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)$. By Hölder’s inequality we have $\text{MIRvSq}^p(\delta, \gamma) \leq \text{RvSq}^p(\delta)$. We will now show a converse estimate.

6.2.2 Bourgain–Guth argument

The statement below is a simplified version of [LV12, (23)].

Lemma 6.1. *Let $0 < \gamma < 1$. Then for any $x \in \mathbb{R}^3$,*

$$|f(x)| \lesssim \max \left(\max_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)|, \gamma^{-1} \max_{\substack{\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma): \\ \text{dist}(\tau_i, \tau_j) \geq 10\gamma, i \neq j}} \prod_{i=1}^3 |f_{\tau_i}(x)|^{1/3} \right). \quad (6.5)$$

Proof. If

$$|f(x)| \leq C \max_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)|,$$

then we are done. Otherwise we have

$$\sum_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)| \geq |f(x)| \geq C \max_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)|,$$

and since $\#\mathcal{P}(\gamma) \sim \gamma^{-1/2}$ it follows that

$$\#\{\tau \in \mathcal{P}(\gamma) \mid |f_\tau(x)| \geq c\gamma^{1/2} \max_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)|\} \geq C$$

for a small absolute constant $c > 0$ and a large absolute constant C . If the latter constant is large enough, then we can choose 10γ -separated τ_1, τ_2, τ_3 from this set, and with this choice we have

$$|f(x)| \leq \sum_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)| \lesssim \gamma^{-1/2} \max_{\tau \in \mathcal{P}(\gamma)} |f_\tau(x)| \lesssim \gamma^{-1} \prod_{i=1}^3 |f_{\tau_i}(x)|^{1/3}. \quad \square$$

Remark. A slightly more careful argument shows that γ^{-1} can be replaced by $\gamma^{-1/2}$ in (6.5).

Using Lemma 6.1 we can establish the following relation between the linear and trilinear square function estimates.

Proposition 6.2. *Let $p \geq 2$. Then for every $0 < \gamma < 1$ we have*

$$\text{RvSq}^p(\delta) \lesssim \text{RvSq}^p(\delta/\gamma^2) + \gamma^{-2-3/(2p)} \text{MIRvSq}^p(\delta, \gamma) \quad (6.6)$$

with the implicit constant independent of γ .

Proof. Fix γ and f_θ . Raising the estimate (6.5) to the power p , replacing maximum by a sum, and integrating over \mathbb{R}^{2+1} we obtain

$$\|f\|_p^p \lesssim \sum_{\tau \in \mathcal{P}(\gamma)} \|f_\tau\|_p^p + \gamma^{-2p} \sum_{\substack{\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma): \\ \text{dist}(\tau_i, \tau_j) \geq 10\gamma, i \neq j}} \left\| \prod_{i=1}^3 |f_{\tau_i}|^{1/3} \right\|_p^p. \quad (6.7)$$

Consider the first sum on the right-hand side of (6.7). By (6.2) we have

$$\|f_\tau\|_p \lesssim \text{RvSq}^p(\delta/\gamma^2) \|\ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\tau\|_p. \quad (6.8)$$

For $p \geq 2$ we have

$$\begin{aligned} \ell_{\tau \in \mathcal{P}(\gamma)}^p \|\ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\theta\|_p &= \|\ell_{\tau \in \mathcal{P}(\gamma)}^p \ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\theta\|_p \\ &\leq \|\ell_{\tau \in \mathcal{P}(\gamma)}^2 \ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 f_\theta\|_p \\ &= \|\ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta\|_p. \end{aligned}$$

Consider the trilinear term on the right-hand side of (6.7). By definition (6.4) we have

$$\begin{aligned} \sum_{\substack{\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma): \\ \text{dist}(\tau_i, \tau_j) \geq 10\gamma, i \neq j}} \left\| \prod_{i=1}^3 |f_{\tau_i}| \right\|_p^p &\leq \sum_{\substack{\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma): \\ \text{dist}(\tau_i, \tau_j) \geq 10\gamma, i \neq j}} \text{MIRvSq}^p(\delta, \gamma)^p \prod_{i=1}^3 \|\ell_{\theta \in \mathcal{P}(\tau_i, \delta)}^2 f_\theta\|_p^p \\ &\leq \sum_{\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)} \text{MIRvSq}^p(\delta, \gamma)^p \prod_{i=1}^3 \|\ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta\|_p^p \\ &\lesssim \gamma^{-3/2} \text{MIRvSq}^p(\delta, \gamma)^p \|\ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta\|_p^p. \end{aligned}$$

Substituting these estimates in (6.7) we obtain

$$\|f\|_p \lesssim \left(\text{RvSq}^p(\delta/\gamma^2) + \gamma^{-2-3/(2p)} \text{MIRvSq}^p(\delta, \gamma) \right) \|\ell_{\theta \in \mathcal{P}(\delta)}^2 f_\theta\|_p.$$

Since f_θ were arbitrary this gives the claim. \square

Corollary 6.3. *Let $2 \leq p < \infty$ and $\alpha > 0$. If $\text{MIRvSq}^p(\delta, \gamma) \lesssim_\gamma \delta^{-\alpha}$ holds for all $0 < \gamma < 1$, then $\text{RvSq}^p(\delta) \lesssim_\varepsilon \delta^{-\alpha-\varepsilon}$ holds for all $\varepsilon > 0$.*

The proof is the same as for decoupling: choose γ small and iterate (6.6).

6.3 The $L^3(\mathbb{R}^{2+1}, \ell^2)$ estimate

In this section we prove the following result.

Theorem 6.4 ([LV12]). *For every $\varepsilon > 0$ we have*

$$\text{RvSq}^3(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}. \quad (6.9)$$

Theorem 6.4 implies the sharp local smoothing estimate for the wave equation on $L^3(\mathbb{R}^2)$.

6.3.1 Multilinear restriction

Let $\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)$ be 10γ -separated. By multilinear restriction on a ball $B \subset \mathbb{R}^3$ of radius $\delta^{-1/2}$ we have

$$\|\prod f_{\tau_i}\|_{L^3(B)} \lesssim_{\gamma, \varepsilon} \delta^{1/4-\varepsilon} \prod \|f_{\tau_i} \psi_B\|_2, \quad (6.10)$$

where ψ_B is a bump function adapted to B with Fourier support in $B(0, \delta^{1/2})$, since

$$\text{supp } \widehat{f_{\tau_i} \psi_B} \subseteq \text{supp } \widehat{f_{\tau_i}} + \text{supp } \widehat{\psi_B}$$

is contained in the $\delta^{1/2}$ -neighborhood of the cone.

Taking the ℓ_B^3 norm over a finitely overlapping covering of \mathbb{R}^3 by balls of radius $\delta^{-1/2}$ and applying Hölder's inequality we obtain

$$\|\prod f_{\tau_i}\|_3 \sim \ell_B^3 \|\prod f_{\tau_i}\|_{L^3(B)} \lesssim_{\gamma, \varepsilon} \delta^{1/4-\varepsilon/2} \prod \ell_B^3 \|f_{\tau_i} \psi_B\|_2. \quad (6.11)$$

6.3.2 Orthogonality

Since the $\delta^{1/2}$ -neighborhoods of $\theta \in \mathcal{P}(\delta)$ have bounded overlap, by almost orthogonality we have

$$\|f_{\tau_i} \psi_B\|_2 \lesssim \|\ell_{\theta \in \mathcal{P}(\tau_i, \delta)}^2 (f_{\theta} \psi_B)\|_2 \lesssim \|\ell_{\theta \in \mathcal{P}(\tau_i, \delta)}^2 f_{\theta}\|_{L^2(w_B)}.$$

By this estimate and Hölder's inequality,

$$\begin{aligned} \ell_B^3 \|f_{\tau_i} \psi_B\|_2 &\lesssim \ell_B^3 \|S_{\delta} f_{\tau_i}\|_{L^2(w_B)} \\ &\lesssim \delta^{-\frac{3}{2}(\frac{1}{2}-\frac{1}{3})} \ell_B^3 \|S_{\delta} f_{\tau_i}\|_{L^3(w_B)} \lesssim \delta^{-\frac{1}{4}} \|S_{\delta} f_{\tau_i}\|_{L^3(\mathbb{R}^3)}. \end{aligned} \quad (6.12)$$

Substituting this in (6.11) and taking the supremum over f_{θ} we obtain

$$\text{MIRvSq}^3(\delta, \gamma) \lesssim_{\gamma, \varepsilon} \delta^{-\varepsilon}.$$

Theorem 6.4 follows by Corollary 6.3.

6.4 The $L^p(\mathbb{R}^{2+1}, \ell^2)$ estimate

Theorem 6.5 (cf. [Lee16]). *Let $3 \leq p \leq 6$ and $\alpha = \frac{(p-2)(p-3)}{2p^2}$. Then for every $\varepsilon > 0$ we have*

$$\text{RvSq}^p(\delta) \lesssim_{\varepsilon} \delta^{-\alpha-\varepsilon}. \quad (6.13)$$

In [Lee16] this is proved for $p = 4$ and we follow the argument given there. It is somewhat surprising that the resulting exponent α is not an affine function of $1/p$. In particular, for intermediate exponents this is better than what interpolation between cases $p = 3, 4, 6$ would give. For $3 < p < 4$ this gives a small improvement over [BHS18, Figure 7]. The local smoothing conjecture for the wave equation in \mathbb{R}^{2+1} would follow from (6.13) with $p = 4$ and $\alpha = 0$.

6.4.1 Multilinear restriction

Let again $\tau_1, \tau_2, \tau_3 \in \mathcal{P}(\gamma)$ be 10γ -separated and consider a ball $B \subset \mathbb{R}^3$ of radius $\delta^{-\kappa}$. By multilinear restriction we have

$$\|\prod f_{\tau_i}\|_{L^3(B)} \lesssim_{\varepsilon} \delta^{\kappa/2-\varepsilon} \prod \|f_{\tau_i}\|_{L^2(\mathbb{R}^3)}$$

for all f_{τ_i} with Fourier support in δ^{κ} -neighborhoods of \mathcal{C}_{τ_i} . By orthogonality

$$\|f\|_2 \lesssim \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{\kappa})}^2 \|f_{\tau}\|_2.$$

Thus, we have

$$\|\prod f_{\tau_i}\|_{L^3(B)} \lesssim_\varepsilon \delta^{\kappa/2-\varepsilon} \prod \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau\|_2.$$

By Hölder's inequality and decoupling we have

$$\|\prod f_{\tau_i}\|_{L^6(B)} \leq \prod \|f_{\tau_i}\|_{L^6(B)} \lesssim_\varepsilon \delta^{-\varepsilon} \prod \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau\|_6.$$

The last two displays hold for arbitrary functions with $\text{supp } \widehat{f}_\tau \subseteq \tau$ for $\tau \in \mathcal{P}(\delta^\kappa)$, not necessarily of the form $f_\tau = \sum_{\theta \in \mathcal{P}(\tau, \delta)} f_\theta$. Define θ, q, β by

$$\frac{1}{p} = \frac{\theta}{3} + \frac{1-\theta}{6}, \quad \frac{1}{q} = \frac{\theta}{2} + \frac{1-\theta}{6}, \quad \frac{1}{q} = \frac{\beta}{2} + \frac{1-\beta}{p}.$$

Then by complex interpolation

$$\|\prod f_{\tau_i}\|_{L^p(B)} \lesssim_\varepsilon \delta^{\theta\kappa/2-\varepsilon} \prod \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau\|_q.$$

Replacing each f_τ by $f_\tau \psi_B$, where ψ_B is an L^∞ normalized bump function adapted to B with Fourier support in $B(0, \delta^\kappa)$ we obtain

$$\begin{aligned} \|\prod f_{\tau_i}\|_{L^p(B)} &\lesssim \|\prod (f_{\tau_i} \psi_B)\|_{L^p(B)} \\ &\lesssim_\varepsilon \delta^{\theta\kappa/2-\varepsilon} \prod \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau \psi_B\|_q \end{aligned}$$

By Hölder's inequality this implies

$$\|\prod f_{\tau_i}\|_{L^p(B)} \lesssim_\varepsilon \delta^{\theta\kappa/2-\varepsilon} \prod \left(\ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau \psi_B\|_p \right)^{1-\beta} \left(\ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau \psi_B\|_2 \right)^\beta. \quad (6.14)$$

By Hölder's inequality we obtain

$$\begin{aligned} \|\prod f_{\tau_i}\|_{L^p(\mathbb{R}^3)} &\sim \ell_B^p \|\prod f_{\tau_i}\|_{L^p(B)} \lesssim_\varepsilon \delta^{\theta\kappa/2-\varepsilon} \prod \left(\ell_B^p \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau \psi_B\|_p \right)^{1-\beta} \\ &\quad \cdot \left(\ell_B^p \ell_{\tau \in \mathcal{P}(\tau_i, \delta^\kappa)}^2 \|f_\tau \psi_B\|_2 \right)^\beta. \end{aligned} \quad (6.15)$$

Now we specify to $\kappa = 1/2$ and functions of the form $f_\tau = \sum_{\theta \in \mathcal{P}(\tau, \delta)} f_\theta$.

6.4.2 Orthogonality

For each $\theta \in \mathcal{P}(\delta)$ the Fourier support of $f_\theta \psi_B$ is contained in the $\delta^{1/2}$ -neighborhood of θ , and these neighborhoods have bounded overlap (here we need $\kappa \geq 1/2$). Hence

$$\begin{aligned} \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|f_\tau \psi_B\|_2 &= \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \left\| \sum_{\theta \in \mathcal{P}(\tau, \delta)} f_\theta \psi_B \right\|_2 \\ &\lesssim \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \ell_{\theta \in \mathcal{P}(\tau, \delta)}^2 \|f_\theta \psi_B\|_2 = \ell_{\theta \in \mathcal{P}(\tau_i, \delta)}^2 \|f_\theta \psi_B\|_2 \lesssim \|S_\delta f_{\tau_i}\|_{L^2(w_B)}. \end{aligned}$$

By this estimate and Hölder's inequality

$$\begin{aligned} \ell_B^p \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|f_\tau \psi_B\|_2 &\lesssim \ell_B^p \|S_\delta f_{\tau_i}\|_{L^2(w_B)} \\ &\lesssim \delta^{-\frac{3}{2}(\frac{1}{2}-\frac{1}{p})} \ell_B^p \|S_\delta f_{\tau_i}\|_{L^p(w_B)} \lesssim \delta^{-\frac{3}{2}(\frac{1}{2}-\frac{1}{p})} \|S_\delta f_{\tau_i}\|_p. \end{aligned} \quad (6.16)$$

6.4.3 Scaling

Let $\alpha \geq 0$ be the smallest number such that (6.13) holds for all $\varepsilon > 0$. By Minkowski's inequality

$$\ell_B^p \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|f_\tau \psi_B\|_p \lesssim \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \ell_B^p \|f_\tau\|_{L^p(w_B)} \sim \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|f_\tau\|_p.$$

By Lorentz scaling (6.2) we have

$$\|f_\tau\|_p \lesssim_\varepsilon \delta^{-\alpha/2-\varepsilon} \|S_\delta f_\tau\|_p.$$

Inserting this and (6.16) into (6.15) we obtain

$$\|\prod f_{\tau_i}\|_{L^p(\mathbb{R}^3)} \lesssim_\varepsilon \delta^{\theta/4-\varepsilon} \prod \left(\delta^{-\alpha/2} \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|S_\delta f_\tau\|_p \right)^{1-\beta} \cdot \left(\delta^{-\frac{3}{2}(\frac{1}{2}-\frac{1}{p})} \|S_\delta f_{\tau_i}\|_p \right)^\beta.$$

Since

$$\beta(1/2 - 1/p) = (1/q - 1/p) = \theta/2 - \theta/3 = \theta/6$$

it follows that

$$\|\prod f_{\tau_i}\|_{L^p(\mathbb{R}^3)} \lesssim_\varepsilon \delta^{-\alpha(1-\beta)/2-\varepsilon} \prod \left(\ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|S_\delta f_\tau\|_p \right)^{1-\beta} \cdot \|S_\delta f_{\tau_i}\|_p^\beta.$$

By doing the bootstrapping argument more carefully we should be able to obtain a hybrid between decoupling and square function inequalities starting from here. However, we want everything to be in terms of square functions, so we estimate

$$\begin{aligned} \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 \|S_\delta f_\tau\|_p &\lesssim \delta^{-\frac{1}{4}(\frac{1}{2}-\frac{1}{p})} \ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^p \|S_\delta f_\tau\|_p \\ &= \delta^{-\frac{1}{4}(\frac{1}{2}-\frac{1}{p})} \|\ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^p S_\delta f_\tau\|_p \\ &\leq \delta^{-\frac{1}{4}(\frac{1}{2}-\frac{1}{p})} \|\ell_{\tau \in \mathcal{P}(\tau_i, \delta^{1/2})}^2 S_\delta f_\tau\|_p \\ &= \delta^{-\frac{1}{4}(\frac{1}{2}-\frac{1}{p})} \|S_\delta f_{\tau_i}\|_p. \end{aligned}$$

6.4.4 Bootstrapping

So far we proved

$$\|\prod f_{\tau_i}\|_p \lesssim_\varepsilon \delta^{-(1-\beta)(\alpha/2 + \frac{1}{4}(\frac{1}{2}-\frac{1}{p}))-\varepsilon} \prod \|S_\delta f_{\tau_i}\|_p.$$

Since $\varepsilon > 0$ was arbitrary and by Corollary 6.3, we have

$$\alpha \leq (1-\beta) \left(\alpha/2 + \frac{1}{4} \left(\frac{1}{2} - \frac{1}{p} \right) \right),$$

so

$$\alpha(1+\beta)/2 \leq (1-\beta) \frac{1}{4} \left(\frac{1}{2} - \frac{1}{p} \right).$$

Further, $\beta(1/2 - 1/p) = \theta/6 = 1/p - 1/6$, so $\beta = (1/p - 1/6)/(1/2 - 1/p)$ and

$$\begin{aligned} \alpha &\leq (1-\beta) \frac{1}{2} \left(\frac{1}{2} - \frac{1}{p} \right) / (1+\beta) \\ &= \frac{1}{2} (1/2 - 1/p) \left((1/2 - 1/p) - (1/p - 1/6) \right) / \left((1/2 - 1/p) + (1/p - 1/6) \right) \\ &= \frac{1}{2} (1/2 - 1/p) (2/3 - 2/p) / (1/3) = \frac{(p-2)(p-3)}{2p^2}. \end{aligned}$$

This completes the proof.

7 Weighted Strichartz estimate for the free Schrödinger equation

In this section we prove a weighted Strichartz estimate from [DZ18]. That article actually proves a more elaborate estimate using also some results from [Du+18] as an input, but the more elaborate version does not give better results for fractal weights.

7.1 Notation

We will prove a priori estimates for solutions $u = e^{it\Delta}f$ of the free Schrödinger equation with initial data $f \in L^2(\mathbb{R}^n)$. The estimate will be on a ball of spatial scale R . For a small δ we consider small balls of scale K^2 , where $K = R^\delta$ will be the parameter in the Bourgain–Guth argument. By parabolic rescaling we will get a bunch of tubes of different sizes summarized below.

Name	Symbol	Size before scaling	Size after scaling
Big	\mathfrak{B}	R	
Small	\mathfrak{S}	K^2	
Big at previous scale	\mathfrak{B}'	$R/K \times R$	$R/K^2 = R_1$
Small at previous scale	\mathfrak{S}'	$KK_1^2 \times K^2K_1^2$	K_1^2
Outer	\mathfrak{D}	$R/r \times R$	
Inner	\mathfrak{J}	$K^2r^{1-4\delta} \times K^2r^{2-4\delta}$	
Outer at previous scale	\mathfrak{D}'	$R/(Kr) \times R$	$R/(K^2r) \times R/K^2$
Inner at previous scale	\mathfrak{J}'	$KK_1^2r^{1-4\delta} \times K^2K_1^2r^{2-4\delta}$	$K_1^2r^{1-4\delta} \times K_1^2r^{2-4\delta}$

The following possible inclusions will be relevant: $\mathfrak{B} \supset \mathfrak{B}' \supset \mathfrak{S}' \supset \mathfrak{S}$ (for the latter inclusion it is important that $\delta \leq 1/4$).

7.2 Weighted estimate

Proposition 7.1. *Let $n \geq 1$ and E, E' suitably large depending on n . Let $p = \frac{2(n+1)}{n-1}$ ($p = \infty$ when $n = 1$). Fix $\delta \in (0, 1/4)$.*

Let $2 \leq t \leq 2(n+1)/n$, $p_{\text{mr},n} = 2(n+1)/n$, $p_{\text{dec},n-1} = 2(n+1)/(n-1)$,

$$\frac{1}{t_1} = \frac{1}{t} - \frac{1}{p_{\text{mr},n}}, \quad \frac{1}{t_2} = \frac{1}{t} - \frac{1}{p_{\text{dec},n-1}}, \quad \frac{1}{t_3} = \frac{n+2}{p_{\text{dec},n-1}} - \frac{n}{2} = -\frac{1}{n+1}.$$

For $R \geq 1$ let $\text{FrStr}(R)$ denote the smallest constant such that the following holds with $K = R^\delta$.

For every non-negative weight function $(\mu_{\mathfrak{S}})$ and every function f with $\text{supp } \hat{f} \subset B^n(0, 1)$ we have

$$\ell_{\mathfrak{S}\mathfrak{C}\mathfrak{B}}^t \mu_{\mathfrak{S}} \|u\|_{L^p(w_{\mathfrak{S},E})} \leq \text{FrStr}(R) W R^{-1/2} \|u\|_{L^2(w_{\mathfrak{B},E'})}. \quad (7.1)$$

with

$$W := \sup_{1 \leq r \leq R^{1/2}} r^{1/t_3} \sup_{\mathfrak{D}} \ell_{\mathfrak{J}\mathfrak{C}\mathfrak{D}}^{t_1} \ell_{\mathfrak{S}\mathfrak{C}\mathfrak{J}}^{t_2} \mu_{\mathfrak{S}}.$$

Then

$$\text{FrStr}(R) \leq C_\varepsilon K^{C_\varepsilon} R^\varepsilon + C_{\varepsilon'} K^{\varepsilon'} \text{FrStr}(R/K^2).$$

Corollary 7.2. *If δ is sufficiently small depending on ε , then*

$$\ell_{\mathfrak{S}\mathfrak{C}\mathfrak{B}}^t \mu_{\mathfrak{S}} \|u\|_{L^p(\mathfrak{S})} \lesssim_\varepsilon W R^\varepsilon \|f\|_{L^2(\mathbb{R}^n)}. \quad (7.2)$$

In the remaining part of this section we prove Proposition 7.1.

Bourgain–Guth argument Denote by $\mathcal{P}(K^{-1})$ the set of K^{-1} -cubes inside the unit ball in the frequency space \mathbb{R}^n . Using a smooth partition of unity we decompose $f = \sum_{\tau \in \mathcal{P}(K^{-1})} f_\tau$ and $u_\tau = e^{it\Delta} f_\tau$ accordingly. Given a K^2 -cube $\mathfrak{S} \subset \mathbb{R}^{n+1}$, we define its *significant set*

$$\mathcal{S}(\mathfrak{S}) := \left\{ \tau \in \mathcal{P}(K^{-1}) \mid \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \geq \frac{\|u\|_{L^p(w_{\mathfrak{S},E})}}{100(\#\mathcal{P}(K^{-1}))} \right\}.$$

We say a K^2 -cube \mathfrak{S} is *narrow* if there is an $(n-1)$ -dimensional subspace $V \subset \mathbb{R}^n$ such that all cubes in $\mathcal{S}(\mathfrak{S})$ are $100/K$ -close to V , and *broad* otherwise. Then for any broad cube \mathfrak{S} there exist ν -transverse cubes $\tau_1, \dots, \tau_{n+1} \in \mathcal{S}(\mathfrak{S})$ with $\nu \gtrsim K^{-n}$.

Localized multilinear restriction Let $\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1})$ be ν -transverse. Then from multilinear restriction we can deduce

$$\ell_{\mathfrak{S} \subset \mathfrak{B}}^{2(n+1)/n} \prod \|\tilde{u}_{\tau_j}\|_{L^p(\mathfrak{S})} \lesssim_\varepsilon K^C \nu^{-C_\varepsilon} R^{-1/2+\varepsilon} \prod \|\tilde{u}_{\tau_j}\|_{L^2(\mathbb{R}^{n+1})}$$

for any functions \tilde{u}_τ supported in the R^{-1} -neighborhood of the paraboloid over τ . In particular this holds for $\tilde{u}_\tau = u_\tau \psi_{\mathfrak{B}}$ for some bump function of scale R adapted to the ball \mathfrak{B} whose Fourier transform is supported in an R^{-1} -neighborhood of 0, and it follows that

$$\begin{aligned} \ell_{\mathfrak{S} \subset \mathfrak{B}}^{2(n+1)/n} \prod \|u_{\tau_j}\|_{L^p(\mathfrak{S})} &\lesssim_{\varepsilon, E''} K^C \nu^{-C_\varepsilon} R^{-1/2+\varepsilon} \prod \|u_{\tau_j}\|_{L^2(w_{\mathfrak{B}, E''})} \\ &\lesssim_{E'', E'''} K^C \nu^{-C_\varepsilon} R^{-1/2+\varepsilon} \|u\|_{L^2(w_{\mathfrak{B}, E'''})}, \end{aligned}$$

where we use that u_τ is a convolution of u with an L^1 normalized smooth kernel. By an averaging procedure similar to that in the proof of decoupling this implies

$$\ell_{\mathfrak{S} \subset \mathfrak{B}}^{2(n+1)/n} \prod \|u_{\tau_j}\|_{L^p(w_{\mathfrak{S}, E})} \lesssim_{\varepsilon, E, E'} K^C \nu^{-C_\varepsilon} R^{-1/2+\varepsilon} \|u\|_{L^2(w_{\mathfrak{B}, E'})} \quad (7.3)$$

for some suitable E, E' .

Broad case If \mathfrak{S} is broad, then for some ν -transverse $\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1})$ depending on \mathfrak{S} we have

$$\|u\|_{L^p(w_{\mathfrak{S}, E})} \lesssim K^C \prod_{j=1}^{n+1} \|u_{\tau_j}\|_{L^p(w_{\mathfrak{S}, E})}$$

so for the sum over broad \mathfrak{S} by (7.3) we have

$$\begin{aligned} \ell_{\mathfrak{S} \subset \mathfrak{B}}^t \mu_{\mathfrak{S}} \|u\|_{L^p(w_{\mathfrak{S}, E})} &\lesssim K^C \ell_{\mathfrak{S} \subset \mathfrak{B}}^t \mu_{\mathfrak{S}} \sup_{\substack{\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} \prod \|u_{\tau_j}\|_{L^p(w_{\mathfrak{S}, E})} \\ &\leq K^C W_{\mathfrak{B}} \ell_{\mathfrak{S} \subset \mathfrak{B}}^{2(n+1)/n} \sup_{\substack{\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} \prod \|u_{\tau_j}\|_{L^p(w_{\mathfrak{S}, E})} \\ &\leq K^C W_{\mathfrak{B}} \sum_{\substack{\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} \ell_{\mathfrak{S} \subset \mathfrak{B}}^{2(n+1)/n} \prod \|u_{\tau_j}\|_{L^p(w_{\mathfrak{S}, E})} \\ &\lesssim_\varepsilon \nu^{-C_\varepsilon} K^C R^\varepsilon W_{\mathfrak{B}} \sum_{\substack{\tau_1, \dots, \tau_{n+1} \in \mathcal{P}(K^{-1}) \\ \nu\text{-transverse}}} R^{-1/2} \|u\|_{L^2(w_{\mathfrak{B}, E'})} \\ &\lesssim K^{C_\varepsilon} R^\varepsilon W_{\mathfrak{B}} R^{-1/2} \|u\|_{L^2(w_{\mathfrak{B}, E'})}, \end{aligned}$$

where $W_{\mathfrak{B}} = \ell_{\mathfrak{S} \subset \mathfrak{B}}^{t_1} \mu_{\mathfrak{S}} \leq W$. This finishes the estimate for the contribution of the broad cubes.

Narrow case If the cube \mathfrak{S} is narrow, then by lower-dimensional decoupling we have

$$\begin{aligned} \|u\|_{L^p(w_{\mathfrak{S},E})} &\sim \left\| \sum_{\tau \in \mathcal{S}(\mathfrak{S})} u_\tau \right\|_{L^p(w_{\mathfrak{S},E})} \\ &\lesssim_\varepsilon K^\varepsilon \ell_{\tau \in \mathcal{S}(\mathfrak{S})}^2 \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \\ &\leq K^\varepsilon \ell_\tau^2 \|u_\tau\|_{L^p(w_{\mathfrak{S},E})}. \end{aligned}$$

Summing over narrow \mathfrak{S} it remains to estimate

$$\begin{aligned} \ell_{\mathfrak{S} \subset \mathfrak{B}}^t \ell_\tau^2 \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} &\leq \ell_\tau^2 \ell_{\mathfrak{S} \subset \mathfrak{B}}^t \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \\ &\lesssim \ell_\tau^2 \ell_{\mathfrak{B}' : \tau}^t \ell_{\mathfrak{S} \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \\ &\lesssim \ell_\tau^2 \ell_{\mathfrak{B}' : \tau}^2 \ell_{\mathfrak{S} \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})}, \end{aligned}$$

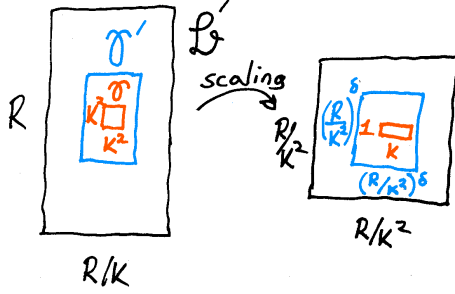
where for each τ we sum over tubes \mathfrak{B}' polar to τ . Fixing τ and \mathfrak{B}' polar to τ we obtain

$$\begin{aligned} \ell_{\mathfrak{S} \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} &\lesssim \ell_{\mathfrak{S}' \subset \mathfrak{B}'}^t \ell_{\mathfrak{S} \subset \mathfrak{S}'}^t \mu_{\mathfrak{S}} \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \\ &\leq \ell_{\mathfrak{S}' \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}'} \ell_{\mathfrak{S} \subset \mathfrak{S}'}^p \|u_\tau\|_{L^p(w_{\mathfrak{S},E})} \\ &\lesssim \ell_{\mathfrak{S}' \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}'} \|u_\tau\|_{L^p(w_{\mathfrak{S}',E})}, \end{aligned} \tag{7.4}$$

where

$$\mu_{\mathfrak{S}'} := \ell_{\mathfrak{S} \subset \mathfrak{S}'}^{t_2} \mu_{\mathfrak{S}}.$$

Recall parabolic scaling L_τ and scaling/modulation operators M_τ that we used in the proof of the decoupling inequality.



We have

$$\begin{aligned} (7.4) &= K^{(n+2)/p} \ell_{\mathfrak{S}' \subset \mathfrak{B}'}^t \mu_{\mathfrak{S}'} \|M_\tau^{-1} u_\tau\|_{L^p(w_{L_\tau \mathfrak{S}', E})} \\ &\leq K^{(n+2)/p} \text{FrStr}(R/K^2) W'(R/K^2)^{-1/2} \|M_\tau^{-1} u_\tau\|_{L^2(w_{L_\tau \mathfrak{B}', E'})} \\ &= K^{(n+2)/p} \text{FrStr}(R/K^2) W'(R/K^2)^{-1/2} K^{-(n+2)/2} \|u_\tau\|_{L^2(w_{\mathfrak{B}', E'})} \\ &= K^{-1/(n+1)} \text{FrStr}(R/K^2) W' R^{-1/2} \|u_\tau\|_{L^2(w_{\mathfrak{B}', E'})} \end{aligned}$$

where W' is computed using the new weights $\mu_{\mathfrak{S}'}$. Since

$$\ell_\tau^2 \ell_{\mathfrak{B}' : \tau}^2 \|u_\tau\|_{L^2(w_{\mathfrak{B}', E'})} \lesssim \ell_\tau^2 \|u_\tau\|_{L^2(w_{\mathfrak{B}, E'})} \lesssim \|u\|_{L^2(w_{\mathfrak{B}, E'})},$$

where we used a localized square function estimate, it remains to observe

$$K^{-1/(n+1)} W' \lesssim W. \tag{7.5}$$

Indeed, every r in the definition of W' corresponds to Kr in the definition of W . This finishes the estimate for the narrow cubes and completes the proof of Proposition 7.1.

7.3 Consequences for fractal measures

Now we evaluate W in the case when $(\mu_{\mathfrak{S}})$ is the characteristic function of an α -dimensional set. Specifically, we assume that $\mu_{\mathfrak{S}} \in \{0, 1\}$ and for any $K^2 \leq r \leq R$ and x we have

$$\sum_{\mathfrak{S} \subseteq B(x, r)} \mu_{\mathfrak{S}} \lesssim r^\alpha. \quad (7.6)$$

Note that $t_2 < t_1$, and it follows that

$$\begin{aligned} W &= \sup_{1 \leq r \leq R^{1/2}} r^{-1/(n+1)} \sup_{\mathfrak{D}} \left(\sum_{\mathfrak{J} \subseteq \mathfrak{D}} \left(\sum_{\mathfrak{S} \subseteq \mathfrak{J}} \mu_{\mathfrak{S}}^{t_2} \right)^{t_1/t_2} \right)^{1/t_1} \\ &\leq \sup_{1 \leq r \leq R^{1/2}} r^{-1/(n+1)} \sup_{\mathfrak{D}} \left(\sum_{\mathfrak{J} \subseteq \mathfrak{D}} \left(\sum_{\mathfrak{S} \subseteq \mathfrak{J}} \mu_{\mathfrak{S}}^{t_2} \right) \cdot \sup_{\mathfrak{J} \subseteq \mathfrak{D}} \left(\sum_{\mathfrak{S} \subseteq \mathfrak{J}} \mu_{\mathfrak{S}}^{t_2} \right)^{t_1/t_2 - 1} \right)^{1/t_1} \\ &\leq \sup_{1 \leq r \leq R^{1/2}} r^{-1/(n+1)} \sup_{\mathfrak{D}} \left(\sum_{\mathfrak{S} \subseteq \mathfrak{D}} \mu_{\mathfrak{S}}^{t_2} \right)^{1/t_1} \cdot \sup_{\mathfrak{J} \subseteq \mathfrak{D}} \left(\sum_{\mathfrak{S} \subseteq \mathfrak{J}} \mu_{\mathfrak{S}}^{t_2} \right)^{1/t_2 - 1/t_1} \end{aligned}$$

Under the hypothesis (7.6) with $\alpha \geq 1$ this is

$$\begin{aligned} &\lesssim \sup_{1 \leq r \leq R^{1/2}} r^{1/t_3} \left(r(R/r)^\alpha \right)^{1/t_1} \cdot (r r^\alpha)^{1/t_2 - 1/t_1} \\ &= R^{\alpha/t_1} \sup_{1 \leq r \leq R^{1/2}} r^{1/t_2 + 1/t_3} r^{\alpha(1/t_2 - 2/t_1)} = R^{\alpha/t_1} \sup_{1 \leq r \leq R^{1/2}} r^{1/t - 1/2} r^{\alpha(1/2 - 1/t)} \\ &= R^{\alpha/t_1} \sup_{1 \leq r \leq R^{1/2}} r^{(\alpha - 1)(1/2 - 1/t)} = R^{\alpha/t_1} R^{(\alpha - 1)(1/2 - 1/t)/2}, \end{aligned}$$

assuming $\alpha \geq 1$. By Corollary 7.2 with $t = 2$ and the above estimate we obtain

$$\boxed{\ell_{\mathfrak{S} \subseteq \mathfrak{B}}^2 \mu_{\mathfrak{S}} \|u\|_{L^p(\mathfrak{S})} \lesssim R^{\alpha/(2(n+1))} \|f\|_2.} \quad (7.7)$$

Up to some uncertainty principle considerations this is the sharp L^2 estimate for the Schrödinger maximal function on sets of dimension α proved in [DZ18, Theorem 2.2].

With $\alpha = n = 1$ and $t = 4$ one can also recover, up to the endpoint, the $L^4(\mathbb{R})$ estimate for the Schrödinger maximal function [KPV91].

One can rescale estimates on a ball of radius R for functions with Fourier support in the unit ball to estimates on a ball of radius 1 for functions with Fourier support in a ball of radius R . Summing these estimates one can formulate these estimates also on H^s spaces, see [Lee06].

7.4 Lower bounds

Applying (7.7) with $\alpha = n$ and using parabolic scaling we see that

$$\left\| \sup_{0 < t < 1/R} |e^{it\Delta} f| \right\|_{L^2(B^n(0,1))} \lesssim R^{\frac{n}{2(n+1)}} \|f\|_{L^2(\mathbb{R}^n)} \quad (7.8)$$

for any function f with $\text{supp } \hat{f} \subset B(0, R)$. In this section we will show that this L^2 bound for the Schrödinger maximal function is optimal up to the R^ε loss. To this end we use special solutions considered in [Bar+07] and [LR17b]. The lower bound that we will show was first proved in [Bou16], but we follow the argument in [LR17a]. Lower bounds for $L^2 \rightarrow L^p$ estimates using the same construction were obtained in [Du+19].

7.4.1 Modulation of initial data

Suppose that $u = e^{it\Delta}f$, or more precisely

$$u(x, t) = \int \widehat{f}(\xi)e(x\xi + t|\xi|^2) d\xi.$$

Modulate f by a frequency η , so that $f_\eta(x) = f(x)e(x\eta)$, and let $u_\eta = e^{it\Delta}f_\eta$, so that

$$\begin{aligned} u_\eta(x, t) &= \int \widehat{f}(\xi - \eta)e(x\xi + t|\xi|^2) d\xi \\ &= \int \widehat{f}(\xi)e(x\xi + x\eta + t|\xi + \eta|^2) d\xi \\ &= e(x\eta + t|\eta|^2) \int \widehat{f}(\xi)e(x\xi + t|\xi|^2 + 2t\xi\eta) d\xi \\ &= e(x\eta + t|\eta|^2)u(x + 2t\eta, t). \end{aligned}$$

So we see that a modulation of the initial data by η causes the solution to travel with speed -2η . We will construct a solution that is large on some set for many times t , and then shift it around using this observation.

7.4.2 Many wave packets

We work in dimension 1. Fix $0 < \sigma < 1$. Let f be given by

$$\widehat{f_{\text{many}}}(\xi) = \widehat{f}(\xi) = \sum_{|l| \leq 2R^\sigma} \varphi(\xi - R^{1-\sigma}l),$$

where φ is a smooth positive function with $\varphi(0) = 1$ supported on a small ρ -neighborhood of the origin. Then

$$\begin{aligned} e^{i(t/R)\Delta}f(x) &= \int \widehat{f}(\xi)e(x\xi + (t/R)|\xi|^2) d\xi \\ &= \sum_{|l| \leq 2R^\sigma} \int \varphi(\xi - R^{1-\sigma}l)e(x\xi + (t/R)|\xi|^2) d\xi \\ &= \sum_{|l| \leq 2R^\sigma} \int \varphi(\xi)e(x(\xi + R^{1-\sigma}l) + (t/R)|\xi + R^{1-\sigma}l|^2) d\xi \end{aligned}$$

Assume $x = x_0 + \tilde{x}$ with $x_0 \in R^{\sigma-1}\mathbb{Z}$, $|x_0| \leq 2$, and $|\tilde{x}| \leq R^{-1}/100$. Assume $t \in R^{2\sigma-1}\mathbb{Z}$ and $|t| \leq 1$. Then

$$\begin{aligned} e^{i(t/R)\Delta}f(x) &= \sum_{|l| \leq 2R^\sigma} \int \varphi(\xi)e((x_0 + \tilde{x})(\xi + R^{1-\sigma}l) + (t/R)|\xi + R^{1-\sigma}l|^2) d\xi \\ &= \sum_{|l| \leq 2R^\sigma} \int \varphi(\xi)e(x_0\xi + \tilde{x}(\xi + R^{1-\sigma}l) + (t/R)|\xi|^2 + (t/R)2\xi R^{1-\sigma}l) d\xi \\ &\sim R^\sigma \end{aligned}$$

since the phase is small.

7.4.3 One wave packet

We still work in dimension 1. Let f be given by

$$\widehat{f_{\text{one}}}(\xi) = \widehat{f}(\xi) = \varphi(R^{-1/2}\xi).$$

Then, assuming $|x| \leq R^{-1/2}$ and $|t| \leq 1$, we obtain

$$e^{i(t/R)\Delta}f = \int \widehat{f}(\xi)e(x\xi + (t/R)|\xi|^2) d\xi \sim R^{1/2}$$

since the phase is small.

7.4.4 Tensor products

If $u_j = e^{it\Delta} f_j$, then

$$u_1(x_1, t) \cdots u_n(x_n, t) = e^{it\Delta} (f_1 \otimes \cdots \otimes f_n)(x_1, \dots, x_n).$$

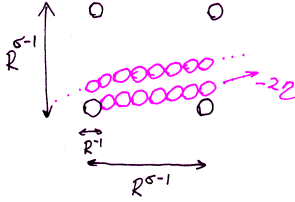
We combine the previous examples by taking a tensor product:

$$f = f_{\text{one}} \otimes f_{\text{many}} \otimes \cdots \otimes f_{\text{many}}.$$

Then the solution $u = e^{i(t/R)\Delta} f$ is $\sim R^{1/2+(n-1)\sigma}$ on the set

$$(-R^{1/2}, R^{1/2}) \times \left((R^{\sigma-1}\mathbb{Z} \cap (-2, 2)) + (-R^{-1}/100, R^{-1}/100) \right)^{n-1} \times \left(R^{2\sigma-1}\mathbb{Z} \cap (0, 1) \right). \quad (7.9)$$

We modulate f by the frequency $\eta = -\frac{1}{2}(R, R^{-2\sigma}, R^{-3\sigma}, \dots, R^{-n\sigma})$. The absolute value of the new solution is then given by $|u(x + 2\eta t, t)|$. We subdivide time into intervals of length $R^{-1/2}$. The shifts of (7.9) by times from non-adjacent time intervals are then essentially disjoint in the first coordinate. The other coordinates of η are chosen in such a way that increasing the time by $R^{2\sigma-1}$ moves the 2nd coordinate by R^{-1} . After R^σ time steps the 2nd coordinate is shifted by $R^{\sigma-1}$, and the 3rd coordinate by R^{-1} . After $R^{(n-1)\sigma}$ time steps the shifts of $R^{-1}/100$ -cubes then cover a positive proportion of the fundamental domain modulo $R^{\sigma-1}$ in coordinates x_2, \dots, x_n .



Choosing $\sigma = \frac{1}{2(n+1)}$ we see that $R^{(n-1)\sigma}$ time steps of length $R^{2\sigma-1}$ add up to a time interval of length $R^{-1/2}$. Moreover, the total shift over the time interval $(0, 1)$ in coordinates x_2, \dots, x_n is bounded by $R^{-2\sigma} \ll 1$. Hence the new solution is $\sim R^{1/2+(n-1)\sigma}$ on a set of measure ~ 1 . Also, $\|f\|_2 \sim R^{(1/2+(n-1)\sigma)/2} \sim R^{n/(2(n+1))}$. Hence if (7.8) holds with exponent s in place of $\frac{n}{2(n+1)}$ we obtain

$$R^{1/2+(n-1)\sigma} \lesssim R^s R^{n/(2(n+1))},$$

so

$$s \geq n/(2(n+1)).$$

Remark. In [Du+19] the same kind of example with

$$f = f_{\text{one}}^{\otimes(n-m+1)} \otimes f_{\text{many}}^{\otimes(m-1)}, \quad 1 \leq m \leq n,$$

is used.

8 Brascamp–Lieb inequalities

We call a finite-dimensional real Hilbert space H with the Lebesgue measure a *Euclidean space*. For an integer $m \geq 0$ an m -*transformation* is a triple

$$\mathbf{B} := (H, (H_j)_{1 \leq j \leq m}, (B_j)_{1 \leq j \leq m})$$

where H, H_1, \dots, H_m are Euclidean spaces and for each j , $B_j : H \rightarrow H_j$ is a linear transformation. We only consider m -transformations that are *non-degenerate* in

the sense that all B_j are surjective, $\dim H_j \neq 0$, and $\bigcap_{j=1}^m \ker(B_j) = \{0\}$. An m -exponent is an m -tuple $\mathbf{p} = (p_j)_{1 \leq j \leq m} \in (0, \infty)^m$ of non-negative real numbers (one can allow zero entries, but this leads to some case distinctions and is not needed in our application). A *Brascamp–Lieb (BL) datum* is a pair (\mathbf{B}, \mathbf{p}) , where \mathbf{B} is an m -transformation and \mathbf{p} is an m -exponent for some integer $m \geq 0$. For a BL datum (\mathbf{B}, \mathbf{p}) and an m -tuple $\mathbf{f} = (f_j)_{1 \leq j \leq m}$ of nonnegative measurable functions $f_j : H_j \rightarrow [0, \infty)$ such that $0 < \int_{H_j} f_j < \infty$ we define

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}; \mathbf{f}) := \frac{\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j}}{\prod_{j=1}^m (\int_{H_j} f_j)^{p_j}}.$$

The *Brascamp–Lieb constant* is the defined as

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) := \sup_{\mathbf{f}} \mathrm{BL}(\mathbf{B}, \mathbf{p}; \mathbf{f}).$$

Equivalently, $\mathrm{BL}(\mathbf{B}, \mathbf{p})$ is the smallest constant for which the m -linear *Brascamp–Lieb inequality*

$$\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j} \leq \mathrm{BL}(\mathbf{B}, \mathbf{p}) \prod_{j=1}^m (\int_{H_j} f_j)^{p_j} \quad (8.1)$$

holds for nonnegative measurable functions $f_j : H_j \rightarrow [0, \infty)$.

Example. The following inequalities are examples of BL inequalities. We refer to the introduction of [Ben+08] for a discussion of these examples.

1. Hölder’s inequality: $B_j = \mathrm{id}_H$ for all j ,
2. Loomis–Whitney inequality: $H = \mathbb{R}^m$, $B_j(x_1, \dots, x_m) = (x_1, \dots, x_{j-1}, x_j, \dots, x_m)$,
3. Young’s convolution inequality: $H = \mathbb{R}^2$, $m = 3$, $B_1(x, y) = x$, $B_2(x, y) = y$, $B_3(x, y) = x + y$.

A central role in the theory of BL inequalities is played by gaussian inputs. A *gaussian input* is an m -tuple $\mathbf{A} = (A_j)_{j=1}^m$ of (strictly) positive definite operators $A_j : H_j \rightarrow H_j$. Let $\mathbf{f}_{\mathbf{A}} = (f_j)_j$ with $f_j(x) = \exp(-\pi \langle A_j x, x \rangle)$ the associated gaussian functions. Then we can compute

$$\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A}) := \mathrm{BL}(\mathbf{B}, \mathbf{p}; \mathbf{f}_{\mathbf{A}}) = \left(\frac{\prod_{j=1}^m (\det_{H_j} A_j)^{p_j}}{\det_H (\sum_{j=1}^m p_j B_j^* A_j B_j)} \right)^{1/2}. \quad (8.2)$$

We define the *gaussian BL constant* by

$$\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) := \sup \{ \mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A}) : \mathbf{A} \text{ is a gaussian input for } (\mathbf{B}, \mathbf{p}) \}. \quad (8.3)$$

Clearly, $\mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) \leq \mathrm{BL}(\mathbf{B}, \mathbf{p})$. Surprisingly, in fact equality holds.

Theorem 8.1 ([Lie90]). *For any BL datum (\mathbf{B}, \mathbf{p}) , we have*

$$\mathrm{BL}(\mathbf{B}, \mathbf{p}) = \mathrm{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}).$$

In the special case of Young’s convolution inequality Theorem 8.1 goes back to [Bec75] for special \mathbf{p} . In the *rank one* case ($\dim H_j = 1$ for all j) Theorem 8.1 for general \mathbf{p} was proved in [BL76b] using rearrangement inequalities in [BLL74]. The proof given by Lieb in [Lie90] has the advantage that it also applies to complex phases, and covers e.g. the Fourier transform. An alternative proof using transportation of mass was given by Barthe [Bar98]. Another alternative approach using heat flow was given by Carlen, Lieb, and Loss [CLL04] in the rank one case and by Bennett, Carbery, Christ, and Tao in [Ben+08].

We follow the latter approach [Ben+08]. We will only need and prove Theorem 8.1 for a special class of BL data called *gaussian extremizable*.

The finiteness of gaussian BL constants is not easy to verify directly. A necessary and sufficient combinatorial condition was found by Bennett, Carbery, Christ, and Tao.

Theorem 8.2 ([Ben+08]). *Let (\mathbf{B}, \mathbf{p}) be a BL datum. Then $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) < \infty$ if and only if we have the scaling condition*

$$\dim(H) = \sum_j p_j \dim(H_j) \tag{8.4}$$

and the dimension condition

$$\dim(V) \leq \sum_{j=1}^m p_j \dim(B_j V) \text{ for every subspace } V \subseteq H. \tag{8.5}$$

We will call (8.5) the *BCCT condition*.

Remark. The conditions (8.4) and (8.5) imply in particular that \mathbf{B} is non-degenerate, as can be seen by testing on $V := H$ and on $V := \bigcap_{j=1}^m \ker(B_j)$.

In the rank one case an equivalent characterization was given in [Bar98].

We will only prove Theorem 8.2 for *simple* data.

Definition 8.3. A BL datum (\mathbf{B}, \mathbf{p}) is called *simple* if the scaling condition (8.4) holds and the BCCT condition (8.5) holds with strict inequality for all subspaces $V \subset H$ with $0 < \dim V < \dim H$.

8.1 Geometric case

Definition 8.4 (Geometric Brascamp–Lieb data). A Brascamp–Lieb datum (\mathbf{B}, \mathbf{p}) is said to be *geometric* if we have

$$B_j B_j^* = \text{id}_{H_j} \text{ for every } j \text{ and} \tag{8.6}$$

$$\sum_{j=1}^m p_j B_j^* B_j = \text{id}_H. \tag{8.7}$$

Example. Hölder’s and Loomis–Whitney inequalities are geometric BL.

Remark. Taking traces of the hypothesis (8.7) we obtain (8.4). One can also verify that (8.5) holds in this case by multiplying (8.7) by the orthogonal projection onto V and taking traces.

For geometric data we will prove the BL inequality using a heat flow.

Proposition 8.5 ([Bal89; Bar98]). *Let (\mathbf{B}, \mathbf{p}) be a geometric Brascamp–Lieb datum. Then*

$$\text{BL}(\mathbf{B}, \mathbf{p}) = \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = 1. \tag{8.8}$$

Proof. Considering the gaussian input $\mathbf{A} = (\text{id}_{H_j})_j$ we see that

$$1 = \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A}) \leq \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}).$$

By 8.1 it therefore suffices to show that for any non-negative measurable functions $f_j : H_j \rightarrow [0, \infty)$ we have

$$\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j} \leq \prod_{j=1}^m \left(\int_{H_j} f_j \right)^{p_j}. \tag{8.9}$$

By Fatou's lemma we may assume that the f_j are smooth and compactly supported. Now let $u_j : \mathbb{R}^+ \times H \rightarrow \mathbb{R}^+$ be the solution to the heat equation

$$\partial_t u_j(t, x) = \Delta_H u_j(t, x) \quad (8.10)$$

$$u_j(0, x) = f_j \circ B_j(x) \quad (8.11)$$

where $\Delta_H := \operatorname{div} \nabla$ is the usual Laplacian on H . More explicitly, we have

$$\begin{aligned} u_j(t, x) &= \frac{1}{(4\pi t)^{\dim(H)/2}} \int_H e^{-\|x-y\|_H^2/4t} f_j(B_j y) \, dy \\ &= \frac{1}{(4\pi t)^{\dim(H_j)/2}} \int_{H_j} e^{-\|B_j x - z\|_{H_j}^2/4t} f_j(z) \, dz. \end{aligned} \quad (8.12)$$

Let

$$Q(t) := \int_H \prod_{j=1}^m u_j^{p_j}(t, x) \, dx.$$

It suffices to show the following three inequalities:

$$\int_H \prod_{j=1}^m (f_j \circ B_j)^{p_j} \leq \limsup_{t \rightarrow 0^+} Q(t), \quad (8.13)$$

$$\limsup_{t \rightarrow 0^+} Q(t) \leq \liminf_{t \rightarrow \infty} Q(t), \quad \text{and} \quad (8.14)$$

$$\liminf_{t \rightarrow \infty} Q(t) \leq \prod_{j=1}^m \left(\int_{H_j} f_j \right)^{p_j}. \quad (8.15)$$

The inequality (8.13) follows from Fatou's lemma.

Let

$$\vec{v}_j := -\frac{\nabla u_j}{u_j}, \quad \vec{v} := \sum_{j=1}^m p_j \vec{v}_j \quad (8.16)$$

We claim that for every $0 < t < \infty$ we have the pointwise inequality

$$\partial_t \left(\prod_{j=1}^m u_j^{p_j} \right) \geq -\operatorname{div} \left(\vec{v} \prod_{j=1}^m u_j^{p_j} \right). \quad (8.17)$$

Assuming (8.17) we obtain

$$\partial_t Q(t) \geq - \int_H \operatorname{div} \left(\vec{v} \prod_{j=1}^m u_j^{p_j} \right) = 0.$$

The latter equality follows from Gauss's theorem since $\vec{v} \prod_{j=1}^m u_j^{p_j}$ is rapidly decreasing in space. Indeed, \vec{v} grows at most polynomially in space and each u_j is rapidly decreasing in the direction of $(\ker B_j)^\perp$. Since \mathbf{B} is non-degenerate, the weighted product of u_j 's is rapidly decreasing in all directions. Thus we have reduced (8.14) to (8.17).

Now we show (8.17). Expanding derivatives on both sides (8.17) becomes

$$\left(\prod_{j=1}^m u_j^{p_j} \right) \left(\sum_{j=1}^m \frac{p_j \partial_t u_j}{u_j} \right) \geq -(\operatorname{div} \vec{v}) \left(\prod_{j=1}^m u_j^{p_j} \right) - \left\langle \vec{v}, \left(\prod_{j=1}^m u_j^{p_j} \right) \left(\sum_{j=1}^m \frac{p_j \nabla u_j}{u_j} \right) \right\rangle.$$

Canceling the product of $u_j^{p_j}$ that is strictly positive for all $t > 0$ we see that this is equivalent to

$$\sum_{j=1}^m \frac{p_j \partial_t u_j}{u_j} \geq -(\operatorname{div} \vec{v}) - \left\langle \vec{v}, \sum_{j=1}^m \frac{p_j \nabla u_j}{u_j} \right\rangle.$$

Inserting the definition (8.16) we calculate

$$\operatorname{div} \vec{v} = - \sum_j p_j \frac{\Delta u_j}{u_j} + \sum_j p_j \frac{\langle \nabla u_j, \nabla u_j \rangle}{u_j^2}.$$

Using the equation (8.10) we see that (8.17) is equivalent to

$$\sum_j p_j \frac{\langle \nabla u_j, \nabla u_j \rangle}{u_j^2} + \left\langle \vec{v}, \sum_{j=1}^m p_j \frac{\nabla u_j}{u_j} \right\rangle \geq 0,$$

which we write in the form

$$\sum_j p_j \langle \vec{v}_j - \vec{v}, \vec{v}_j \rangle \geq 0.$$

Since $u_j(x)$ depends only on $B_j x$ and using (8.6) we see that $\vec{v}_j = B_j^* B_j \vec{v}_j$, so our claim is equivalent to

$$\sum_j p_j \langle B_j^* B_j (\vec{v}_j - \vec{v}), \vec{v}_j \rangle \geq 0.$$

From (8.7) we have

$$\sum_{j=1}^m p_j B_j^* B_j (\vec{v} - \vec{v}_j) = \vec{v} - \sum_{j=1}^m p_j B_j^* B_j \vec{v}_j = \vec{v} - \sum_{j=1}^m p_j \vec{v}_j = 0$$

and hence the claim is equivalent to

$$\sum_j p_j \langle B_j^* B_j (\vec{v}_j - \vec{v}), (\vec{v}_j - \vec{v}) \rangle \geq 0.$$

This is clearly true, so the claim (8.17) holds. This finishes the proof of (8.14).

It remains to show (8.15). Using (8.12) we can write

$$Q(t) = \frac{1}{(4\pi t)^{\sum_{j=1}^m p_j \dim(H_j)/2}} \int_H \prod_{j=1}^m \left(\int_{H_j} e^{-\|B_j x - z\|_{H_j}^2 / 4t} f_j(z) dz \right)^{p_j} dx.$$

Making the change of variables $x = t^{1/2} w$ we obtain

$$Q(t) = \frac{1}{(4\pi)^{\dim(H)/2}} \int_H \prod_{j=1}^m \left(\int_{H_j} e^{-\|B_j w - t^{-1/2} z\|_{H_j}^2 / 4} f_j(z) dz \right)^{p_j} dw.$$

Since the f_j are rapidly decreasing and $\bigcap_{j=1}^m \ker(B_j) = \{0\}$, we may then use dominated convergence to conclude

$$\begin{aligned} \liminf_{t \rightarrow \infty} Q(t) &= \frac{1}{(4\pi)^{\dim(H)/2}} \int_H \prod_{j=1}^m \left(\int_{H_j} e^{-\|B_j w\|_{H_j}^2 / 4} f_j(z) dz \right)^{p_j} dw \\ &= \frac{1}{(4\pi)^{\dim(H)/2}} \prod_{j=1}^m \left(\int_{H_j} f_j \right)^{p_j} \int_H e^{-\langle \sum_{j=1}^m p_j B_j^* B_j w, w \rangle_H / 4} dw. \end{aligned}$$

Using (8.7), the claim (8.15) follows. \square

8.2 Gaussian-extremizable case

Definition 8.6. A Brascamp–Lieb datum (\mathbf{B}, \mathbf{p}) is said to be *gaussian-extremizable* if there exists a gaussian input \mathbf{A} for which $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) = \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})$.

In this section we prove Theorem 8.1 in the gaussian-extremizable case.

Proposition 8.7. *Let (\mathbf{B}, \mathbf{p}) be a gaussian extremizable Brascamp–Lieb datum. Then there exist invertible linear transformations C on H and C_j on H_j such that the BL datum $(\mathbf{B}', \mathbf{p})$ with $\mathbf{B}'_j = C_j B_j C$ is geometric.*

Proof. Let $\mathbf{A} = (A_j)_{1 \leq j \leq m}$ be a gaussian input for (\mathbf{B}, \mathbf{p}) that is a local extremizer for $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})$. Let $M : H \rightarrow H$ be given by $M := \sum_{j=1}^m p_j B_j^* A_j B_j$. Since the A_j are positive definite and \mathbf{B} is non-degenerate, the operator M is also positive definite and in particular invertible. We claim that

$$A_j^{-1} = B_j M^{-1} B_j^* \text{ for all } 1 \leq j \leq m. \quad (8.18)$$

Taking logarithms in (8.2), we see that $\mathbf{A} > 0$ is a local maximizer for the quantity

$$\sum_{j=1}^m p_j \log \det_{H_j} A_j - \log \det_H \sum_{j=1}^m p_j B_j^* A_j B_j. \quad (8.19)$$

A Taylor expansion shows that for an invertible matrix A and any matrix Q we have

$$\frac{d}{d\varepsilon} \log \det(A + \varepsilon Q)|_{\varepsilon=0} = \frac{d}{d\varepsilon} \log \det(1 + A^{-1} \varepsilon Q)|_{\varepsilon=0} = \text{tr}(A^{-1} Q).$$

Differentiating the quantity (8.19) in the variable A_j in the direction of a self-adjoint transformation $Q_j : H_j \rightarrow H_j$ we obtain

$$p_j \text{tr}_{H_j}(A_j^{-1} Q_j) - \text{tr}_H(p_j M^{-1} B_j^* Q_j B_j) = 0.$$

We rearrange this using the cyclic property of the trace as

$$\text{tr}_{H_j}((A_j^{-1} - B_j M^{-1} B_j^*) Q_j) = 0.$$

Since Q_j was an arbitrary self-adjoint transformation, and $A_j^{-1} - B_j M^{-1} B_j^*$ is also self-adjoint, we conclude (8.18).

Let $B'_j := C_j B_j C$ with $C := M^{-1/2}$ and $C_j := A_j^{-1/2}$. It remains to show that $(\mathbf{B}', \mathbf{p})$ is a geometric BL datum. From (8.18) we have $B'_j (B'_j)^* = \text{id}_{H_j}$, and from definition of M we have

$$\sum_{j=1}^m p_j (B'_j)^* B'_j = \sum_{j=1}^m M^{-1/2} p_j B_j^* A_j B_j M^{-1/2} = M^{-1/2} M M^{-1/2} = \text{id}_H. \quad \square$$

Corollary 8.8. *Let (\mathbf{B}, \mathbf{p}) be a gaussian extremizable Brascamp–Lieb datum. Then*

$$\text{BL}(\mathbf{B}, \mathbf{p}) = \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}).$$

Proof. Both the BL constant and the gaussian BL constant are multiplied by the same Jacobian factor when B_j are replaced by $C_j B_j C$. Since they coincide for geometric data by Proposition 8.5, they also coincide in the setting of Proposition 8.7. \square

8.3 Necessity of BCCT condition

Lemma 8.9. *Let (\mathbf{B}, \mathbf{p}) be Brascamp–Lieb datum such that $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) < \infty$. Then the scaling condition (8.4) and the BCCT condition (8.5) hold.*

Proof. Let $\lambda > 0$ be arbitrary. Applying (8.3) with the gaussian input $(\lambda \text{id}_{H_j})_{1 \leq j \leq m}$ we see that

$$\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})^2 \geq \lambda^{\sum_{j=1}^m p_j \dim(H_j) - \dim(H)} / \det\left(\sum_{j=1}^m p_j B_j^* B_j\right).$$

Since λ is arbitrary we obtain (8.4).

Next, let V be any subspace in H , and let $\lambda > 0$. Let $\mathbf{A}^{(\lambda)} = (A_j)_{1 \leq j \leq m}$ be the gaussian input $A_j := \lambda \text{id}_{B_j V} + \text{id}_{(B_j V)^\perp}$. Then $\det_{H_j}(A_j) = \lambda^{\dim(B_j V)}$. Also, we see that $\sum_{j=1}^m p_j B_j^* A_j B_j$ is bounded uniformly in $\lambda \leq 1$, and when restricted to V decays linearly in λ . Thus $\det_H(\sum_{j=1}^m B_j^* A_j B_j) \lesssim \lambda^{\dim(V)}$. Inserting this in (8.2) we obtain

$$\infty > \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})^2 \geq \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A}^{(\lambda)})^2 \gtrsim \frac{\prod_{j=1}^m \lambda^{p_j \dim(B_j V)}}{\lambda^{\dim V}}.$$

Since this holds for all $0 < \lambda \leq 1$ we obtain (8.5). \square

8.4 Sufficiency of BCCT condition for simple data

In this section we prove Theorem 8.2 for simple data, and in fact a more precise statement.

Lemma 8.10. *Let (\mathbf{B}, \mathbf{p}) be a non-degenerate Brascamp–Lieb datum such that (8.5) holds. Then there exists a real number $c > 0$, such that for every orthonormal basis e_1, \dots, e_n of H there exist sets $I_j \subseteq \{1, \dots, n\}$ for each $1 \leq j \leq m$ with $|I_j| = \dim(H_j)$ such that*

$$\sum_{j=1}^m p_j |I_j \cap \{k+1, \dots, n\}| \geq n - k \text{ for all } 0 < k < n. \quad (8.20)$$

and

$$\left\| \bigwedge_{i \in I_j} B_j e_i \right\|_{H_j} \geq c \text{ for all } 1 \leq j \leq m. \quad (8.21)$$

If furthermore (\mathbf{B}, \mathbf{p}) is simple, then we can choose I_j so that the inequality (8.20) is strict.

Proof. We claim that it suffices to show a weaker statement, namely that for every orthonormal basis (\tilde{e}_i) there exist I_j as above with the conclusion by the weaker statement

$$\bigwedge_{i \in I_j} B_j \tilde{e}_i \neq 0 \text{ for all } 1 \leq j \leq m.$$

Then by continuity we have (8.21) for (e_i) in a small neighborhood of (\tilde{e}_i) with the same sets I_j . Since the space of all orthonormal bases is compact this implies the conclusion.

Hence it suffices to make the vectors $(B_j e_i)_{i \in I_j}$ linearly independent for each j . We select the I_j by a backwards greedy algorithm:

$$I_j := \{i \mid B_j e_i \notin \text{lin}\{B_j e_{i'} \mid i < i' \leq n\}\}.$$

Then by backward induction on i we see that $\{B_j e_{i'} \mid i \leq i' \leq n, i' \in I_j\}$ is a basis of the subspace $\text{lin}\{B_j e_{i'} \mid i \leq i' \leq n\}$. Since the B_j are surjective it follows that $|I_j| =$

$\dim(H_j)$. To prove (8.20), we apply the hypothesis (8.5) with $V = \text{lin}\{e_{k+1}, \dots, e_n\}$, to obtain

$$\begin{aligned} n - k &= \dim(V) \leq \sum_j p_j \dim(B_j V) \\ &= \sum_j p_j |I_j \cap \{k+1, \dots, n\}|. \end{aligned}$$

If the datum is simple, then the above inequality is strict for $0 < k < n$. \square

Proposition 8.11. *Let $(\tilde{\mathbf{B}}, \mathbf{p})$ be a Brascamp–Lieb datum such that (8.4) and (8.5) hold. Then there exists a neighborhood \mathcal{B} of $\tilde{\mathbf{B}}$ such that $\sup_{\mathbf{B} \in \mathcal{B}} \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}) < \infty$. Furthermore, if $(\tilde{\mathbf{B}}, \mathbf{p})$ is simple, then (\mathbf{B}, \mathbf{p}) is gaussian-extremizable for every $\mathbf{B} \in \mathcal{B}$.*

Proposition 8.11 combined with Corollary 8.8 in particular shows the sufficiency of (8.4) and (8.5) in Theorem 8.2 in the case of simple data. In fact it provides the more precise conclusion that the gaussian BL constant is uniformly bounded on a neighborhood of the original m -transformation. This was first made explicit in [Ben+18b]. More precise result on regularity of $\text{BL}(\mathbf{B}, \mathbf{p})$ as a function of \mathbf{B} appear in [Val11; Ben+17; Ben+18a].

Proof. The left-hand side of (8.21) is a Lipschitz function of \mathbf{B} , so applying Lemma 8.10 to the datum $(\tilde{\mathbf{B}}, \mathbf{p})$ and replacing c by a slightly smaller number $c(\tilde{\mathbf{B}})$ we may assume that (8.21) continues to hold for all \mathbf{B} in a small neighborhood \mathcal{B} of $\tilde{\mathbf{B}}$.

Let $\mathbf{A} = (A_j)_{1 \leq j \leq m}$ be a gaussian input and $M := \sum_j p_j B_j^* A_j B_j$. We want to estimate

$$\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})^2 = \frac{\prod_{j=1}^m (\det A_j)^{p_j}}{\det M}.$$

The transformation M is self-adjoint; since \mathbf{B} is non-degenerate and $p_j > 0$, we also see that it is positive definite. Thus by choosing an appropriate orthonormal basis $\{e_1, \dots, e_n\} \subset H$ we may assume that $M = \text{diag}(\lambda_1, \dots, \lambda_n)$ for some $\lambda_1 \geq \dots \geq \lambda_n > 0$.

Applying Lemma 8.10, we can find $I_j \subseteq \{1, \dots, n\}$ for each $1 \leq j \leq m$ of cardinality $|I_j| = \dim(H_j)$ obeying (8.20) and (8.21). For each $i \in I_j$, we have

$$\langle A_j B_j e_i, B_j e_i \rangle_{H_j} = \langle e_i, B_j^* A_j B_j e_i \rangle_H \leq \frac{1}{p_j} \langle e_i, M e_i \rangle_H = \lambda_i / p_j.$$

Since A_j is positive definite and by the Cauchy–Schwarz inequality this implies

$$|\langle A_j B_j e_i, B_j e_{i'} \rangle_{H_j}| \leq (\lambda_i \lambda_{i'})^{1/2} / p_j.$$

It follows that

$$|\det(\langle A_j B_j e_i, B_j e_{i'} \rangle_{H_j})_{i, i' \in I_j}| \leq \sum_{\sigma \in S_{I_j}} \prod_{i \in I_j} (\lambda_i \lambda_{\sigma(i)})^{1/2} / p_j = (n_j!) p_j^{-n_j} \prod_i \lambda_i.$$

On the other hand, from (8.21) we see that $(B_j e_i)_{i \in I_j}$ is a basis of H_j with a lower bound on the degeneracy. We thus conclude that

$$\det(A_j) \leq \left\| \bigwedge_{i \in I_j} B_j e_i \right\|_H^{-2} |\det(\langle A_j B_j e_i, B_j e_{i'} \rangle_{H_j})_{i, j}| \leq C_{\mathbf{n}, \mathbf{p}} c(\tilde{\mathbf{B}})^{-2} \prod_{i \in I_j} \lambda_i.$$

Thus

$$\prod_{j=1}^m (\det A_j)^{p_j} \lesssim \prod_{i=1}^n \lambda_i^{\sum_{j=1}^m p_j |I_j \cap \{i\}|}.$$

with an implicit constant depending only on $\mathbf{n}, \mathbf{p}, \tilde{\mathbf{B}}$, but not on \mathbf{B} . We can telescope the right-hand side (using (8.4)) and obtain

$$\prod_{j=1}^m (\det A_j)^{p_j} \lesssim \lambda_1^n \prod_{1 \leq k \leq n-1} (\lambda_{k+1}/\lambda_k)^{\sum_{j=1}^m p_j |I_j \cap \{k+1, \dots, n\}|}.$$

Applying (8.20) we conclude

$$\prod_{j=1}^m (\det A_j)^{p_j} \lesssim \lambda_1^n \prod_{1 \leq k \leq n-1} (\lambda_{k+1}/\lambda_k)^{n-k}$$

which by reversing the telescoping becomes

$$\prod_{j=1}^m (\det A_j)^{p_j} \lesssim \lambda_1 \dots \lambda_k = \det(M),$$

so that $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})^2 \leq C(\mathbf{n}, \mathbf{p}, \tilde{\mathbf{B}})$. Since \mathbf{A} was arbitrary, by definition (8.2) we conclude that $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p})^2 \leq C(\mathbf{n}, \mathbf{p}, \tilde{\mathbf{B}})$.

Now suppose that $(\tilde{\mathbf{B}}, \mathbf{p})$ is simple. Then we have strict inequality in (8.20). Then it follows that

$$\sum_{j=1}^m p_j |I_j \cap \{k+1, \dots, n\}| \geq n - k + c' \text{ for all } 0 < k < n$$

with some strictly positive $c' > 0$ independent of \mathbf{B} since the cardinalities on the left-hand side can only assume finitely many values. We may thus refine the above analysis and conclude that

$$\prod_{j=1}^m (\det A_j)^{p_j} \leq C(\tilde{\mathbf{B}}) \det(M) \prod_{1 \leq k \leq n-1} (\lambda_{k+1}/\lambda_k)^{c'}.$$

This shows that $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A}) \rightarrow 0$ whenever $\lambda_n/\lambda_1 \rightarrow 0$. Thus to evaluate the supremum it suffices to do so in the region $\lambda_1 \leq C\lambda_n$. By the scaling hypothesis (8.4) we may normalize $\lambda_n = 1$. This means that M is now bounded above and below, which by surjectivity of B_j implies that A_j is also bounded. We may now also assume that A_j is bounded from below since otherwise $\text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})$ will be small. We have thus localized each the A_j to a compact set, and hence by continuity we see that an extremizer of $\mathbf{A} \mapsto \text{BL}_{\mathbf{g}}(\mathbf{B}, \mathbf{p}; \mathbf{A})$ exists. Thus (\mathbf{B}, \mathbf{p}) is gaussian-extremizable as desired. \square

Corollary 8.12 (Locally uniform BL inequality). *Let $(\tilde{\mathbf{B}}, \mathbf{p})$ be a simple Brascamp–Lieb datum such that (8.4) and (8.5) hold. Then there exists a neighborhood \mathcal{B} of $\tilde{\mathbf{B}}$ such that*

$$\sup_{\mathbf{B} \in \mathcal{B}} \text{BL}(\mathbf{B}, \mathbf{p}) < \infty.$$

Proof. By Proposition 8.10 we know that the gaussian BL constant is bounded uniformly on a neighborhood \mathcal{B} and each datum (\mathbf{B}, \mathbf{p}) with $\mathbf{B} \in \mathcal{B}$ is gaussian extremizable. By Corollary 8.8 the gaussian BL constants coincide with the full BL constants on this neighborhood. \square

The argument in [Ben+08] for non-simple data is based on a decomposition of non-simple data as a kind of semidirect product of simple data. We will be content with the simple case.

9 Decoupling for the moment curve

9.1 Vinogradov mean value

It is of interest in number theory to estimate Weyl sums such as

$$\sum_{x=1}^X e(\xi_1 x^1 + \cdots + \xi_k x^k).$$

Vinogradov's method deduces bounds for pointwise values (for fixed $\vec{\xi} = (\xi_1, \dots, \xi_k)$) from bounds for averages (*mean values*) over all $\vec{\xi}$. Specifically, it needs estimates on the moments

$$\int_{\vec{\xi} \in [0,1]^k} \left| \sum_{x=1}^X e(\xi_1 x^1 + \cdots + \xi_k x^k) \right|^p d\vec{\xi}. \quad (9.1)$$

When the exponent $p = 2s$ is an even integer, the expression (9.1) can be interpreted as counting the solutions of a system of equations. Indeed, the $2s$ -power of the absolute value can be expanded, and each summand contributes 1 to the integral if its frequency vanishes and 0 otherwise. The frequency vanishes iff the following system of equations holds.

$$\begin{aligned} x_1 + \cdots + x_s &= x_{s+1} + \cdots + x_{2s} \\ &\vdots \\ x_1^k + \cdots + x_s^k &= x_{s+1}^k + \cdots + x_{2s}^k. \end{aligned} \quad (9.2)$$

The number $J_{s,k}(X)$ of integer solutions to the system of equations (9.2) with all entries in the interval $[1, X]$ is certainly at least X^s (considering the ‘‘diagonal’’ solutions $x_{s+j} = x_j$). Another estimate comes from a more elaborate counting argument. For $Y = (Y_1, \dots, Y_k)$ let $J_{s,k}(X, Y)$ be the number of integer solutions with all entries in $[1, X]$ of the system of equations

$$\begin{aligned} x_1 + \cdots + x_s &= Y_1 \\ &\vdots \\ x_1^k + \cdots + x_s^k &= Y_k. \end{aligned}$$

Then

$$X^s = \sum_{Y_1=0}^{sX} \cdots \sum_{Y_k=0}^{sX^k} J_{s,k}(X, Y).$$

On the other hand,

$$J_{s,k}(X) = \sum_{Y_1=0}^{sX} \cdots \sum_{Y_k=0}^{sX^k} (J_{s,k}(X, Y))^2.$$

By the Cauchy–Schwarz inequality it follows that

$$\begin{aligned} X^s &= \sum_{Y_1=0}^{sX} \cdots \sum_{Y_k=0}^{sX^k} J_{s,k}(X, Y) \leq \left(\sum_{Y_1=0}^{sX} \cdots \sum_{Y_k=0}^{sX^k} J_{s,k}(X, Y)^2 \right)^{1/2} \left(\sum_{Y_1=0}^{sX} \cdots \sum_{Y_k=0}^{sX^k} 1 \right)^{1/2} \\ &= J_{s,k}(X)^{1/2} s^{k/2} X^{(1+\cdots+k)/2}. \end{aligned}$$

Combining this with the estimate for the number of diagonal solutions we obtain

$$J_{s,k}(X) = (9.1) \geq \max(X^s, s^{-k} X^{2s - \frac{k(k+1)}{2}}).$$

It turns out that this lower bound is essentially sharp, as proved in [BDG16] and [Woo19]. We will present the decoupling proof from [BDG16] with some simplifications coming from later works [BDG17; GZa; GZb; GZ19]. The two lower bounds that we obtained coincide when $p = 2s = k(k + 1)$, so this “critical” exponent can be expected to play a special role. We will indeed consider only $p = k(k + 1)$, since sharp results for other p ’s can be obtained by interpolation.

Remark. The reduction to a unique critical exponent is special to the one-dimensional situation, in higher dimensions there may be many critical exponents. In fact it becomes preferable not to single out any exponents.

9.2 Statement of the main result

We will formulate a decoupling theorem for functions with Fourier support close to the unit moment curve $\{(\xi, \xi^2, \dots, \xi^k) \mid \xi \in [0, 1]\}$. There is only one parameter, so cubes in the decoupling theorem for the paraboloid become just intervals. We continue to denote by $\mathcal{P}(Q, \delta)$ the partition of an interval Q into dyadic intervals of length δ . We omit Q if $Q = [0, 1]$. We denote by f_θ a function whose Fourier support is adapted to the image of θ in the moment curve (this will be made precise when we describe the affine scaling procedure). We denote by $\text{Dec}_k(\delta)$ the smallest constant in the inequality

$$\left\| \sum_{\theta \in \mathcal{P}(\delta)} f_\theta \right\|_{L^{k(k+1)}(\mathbb{R}^k)} \leq \text{Dec}_k(\delta) \ell_{\theta \in \mathcal{P}(\delta)}^2 \|f_\theta\|_{L^{k(k+1)}(\mathbb{R}^k)}. \quad (9.3)$$

We do not include the exponents $k(k + 1)$ and 2 in the notation since these will be the only exponents that we consider.

Theorem 9.1. *For every $k \geq 1$ and $\varepsilon > 0$ we have $\text{Dec}_k(\delta) \lesssim_\varepsilon \delta^{-\varepsilon}$.*

Theorem 9.1 will be proved by induction on k . The case $k = 1$ is just L^2 orthogonality (in fact we have also already proved the case $k = 2$, since this is the case of the one-dimensional paraboloid). We will henceforth assume that $k \geq 2$ and Theorem 9.1 is already known for smaller values of k .

The main technical difficulties in the case of the moment curve are different from the case of the paraboloid. The Bourgain–Guth argument is much easier, since all lower-dimensional contributions are zero-dimensional. On the other hand, the treatment of multilinear terms becomes more sophisticated.

9.2.1 Consequences for Vinogradov mean values

By the usual procedure the inequality (9.3) can be localized to balls of radius $\gtrsim \delta^{-k}$. Then we choose $\delta \sim X^{-1}$ and let each \widehat{f}_θ be supported in a point on the moment curve over a rational number with denominator X . The right-hand side can then be easily computed since $|f_\theta| \equiv \text{const}$. On the other hand, by periodicity the left-hand side coincides up to scaling with (9.1). After scaling this gives the estimate

$$\int_{\vec{\xi} \in [0, 1]^k} \left| \sum_{x=1}^X e(\xi_1 x^1 + \dots + \xi_k x^k) \right|^{k(k+1)} d\vec{\xi} \lesssim_\varepsilon X^{k(k+1)/2+\varepsilon}.$$

9.2.2 Affine scaling

Let $\theta \in \mathcal{P}(\sigma,)$ with left endpoint $c = c(\theta)$. Consider the affine transformation

$$(L_\theta(x))_j = \sigma^j \sum_{i: 0 \leq i \leq j} \binom{i}{j} (c)^{j-i} x_i, \quad 1 \leq j \leq k, \quad (9.4)$$

where we set $x_0 = 1$. This transformation preserves the moment curve and maps the point 0 to the image of c in the moment curve. The support condition that we impose on our functions is that $\text{supp } \widehat{f}_\theta$ is contained in the image of a fixed ball centered at the origin under L_θ . Then for $\theta_0 \in \mathcal{P}(\delta_0)$ we immediately obtain the rescaled decoupling inequality

$$\| \sum_{\theta' \in \mathcal{P}(\theta_0, \delta_0 \delta_1)} f_{\theta'} \|_{L^{k(k+1)}(\mathbb{R}^k)} \leq \text{Dec}_k(\delta_1) \ell_{\theta' \in \mathcal{P}(\theta, \delta_0 \delta_1)}^2 \|f_{\theta'}\|_{L^{k(k+1)}(\mathbb{R}^k)}.$$

9.3 Transversality

The functions f_θ have Fourier support in boxes of size $\delta \times \delta^2 \times \dots \times \delta^k$. Hence they are morally constant on boxes of size $\delta^{-1} \times \delta^{-2} \times \dots \times \delta^{-k}$. We need a description of orientation of these boxes. To this end we will use the higher order tangent spaces. More precisely, we use the l -th order tangent spaces

$$V^{(l)}(t) := \text{lin}\{\partial^j \Phi(t) \mid 1 \leq j \leq l\} \subseteq \mathbb{R}^k, \quad t \in [0, 1], \quad (9.5)$$

where $\Phi(t) = (t^\gamma)_{1 \leq \gamma \leq k}$ parametrizes the moment curve. If $t \in \theta$, then the function f_θ is morally constant at scale δ^{-l-1} in the orthogonal direction to $V^{(l)}(t)$.

In order to make use of Kakeya–Brascamp–Lieb inequalities we have to verify that the spaces $V^{(l)}(t)$ are transverse when we consider sufficiently widely spaced t 's. Due to the lack of an explicit description of BL constants, transversality in this case means that the BCCT condition for finiteness of BL constants is satisfied. In contrast to the paraboloid case it is not a priori clear how many different t 's one would have to consider to achieve such transversality.

9.3.1 Projections onto higher order tangential spaces

In the case of the moment curve we get lucky and it turns out that for any subspace $V \subseteq \mathbb{R}^k$ the projection onto $V^{(l)}(t)$ almost always has maximal possible dimension.

Theorem 9.2. *For each $k \geq 1$ there exists $M_{0,k}$ such that for every subspace $V \subseteq \mathbb{R}^k$ and every $1 \leq l \leq k$ we have*

$$\dim \pi_{V^{(l)}(t)} V = \min(l, \dim V) \quad (9.6)$$

for all but at most $M_{0,k}$ values of $t \in [0, 1]$.

Proof. Decreasing the dimension of V or l if necessary we may assume without loss of generality $\dim V = l$. Fix a basis (v_1, \dots, v_l) of V such that $a_h := \max\{j \mid v_{h,j} \neq 0\}$ is strictly decreasing in $h = 1, \dots, l$. The dimension on LHS(9.6) equals the rank of the $l \times l$ matrix

$$\mathcal{M}_V^{(l)}(t) := (v_1, \dots, v_l)^T \times (\partial^j \Phi(t))_{1 \leq j \leq l} \quad (9.7)$$

over \mathbb{R} . We claim that the determinant of this matrix is a non-zero polynomial in t . This will suffice to establish the claim since the degree of this polynomial is bounded by some $M_{0,k}$, so it will have at most $M_{0,k}$ zeros. If t is not a zero of this polynomial, then $\mathcal{M}_V^{(l)}(t)$ has real rank l .

To simplify notation write $f_v(x) := \sum_{i=1}^l v_i x^i$. Then

$$\mathcal{M}_V^{(l)}(t) = (\partial^j f_{v_h})_{1 \leq j, h \leq l}.$$

Note that $\deg f_h = a_h$ is strictly decreasing in h . Applying column and row transformations over the field of rational functions $\mathbb{R}(x)$ we obtain the matrix

$$(x^{-a_h} x^j \partial^j f_{v_h})_{1 \leq j, h \leq l}.$$

All entries of this matrix are linear combinations of non-positive powers of x . Hence the constant term in its determinant equals the determinant of constant terms, which is given by

$$\det((a_h \cdots (a_h - j + 1))_{1 \leq j, h \leq l}).$$

Since all v_{h, a_h} are non-zero this determinant is non-zero iff

$$\det(a_h \cdots (a_h - j + 1))_{1 \leq j, h \leq l}$$

is non-zero. In the j -th row all entries are polynomials in a_h of degree j . By row operations one can bring this determinant in the form

$$\det(a_h^j)_{1 \leq j, h \leq l}.$$

But this is $a_1 \cdots a_h$ times a Vandermonde determinant, so non-zero. \square

9.3.2 Verification of BCCT condition

Corollary 9.3. *For every k there exists M such that for any distinct $t_1, \dots, t_M \in [0, 1]$ and any $1 \leq l < k$ we have*

$$BL((V^{(l)}(t_j))_{j=1}^M) < \infty$$

and this BL datum is simple.

It is not necessary to ensure simplicity to proceed with the proof, but we only proved the BCCT criterion for finiteness of BL constants in the simple case.

Proof. By the BCCT condition the BL constant is finite if for every subspace $V \subseteq \mathbb{R}^k$ we have

$$\dim V \leq \frac{k}{lM} \sum_{m=1}^M \dim \pi_{V^{(l)}(t_j)} V \quad (\text{BCCT})$$

with equality for $V = \mathbb{R}^k$. We have actually only proved this in the simple case when the inequality is strict for $0 < \dim V < k$, and we will be able to put ourselves in this situation. The equality in the case $\dim V = k$ is easy to see.

Assume now $\dim V < k$. By Theorem 9.2 we have

$$\dim \pi_{V^{(l)}(t_j)} V = \min(l, \dim V)$$

for all but at most $M_{0,k}$ many t 's. Hence

$$\begin{aligned} RHS(BCCT) &\geq \frac{k(M - M_{0,k})}{lM} \min(l, \dim V) = \frac{M - M_{0,k}}{M} \min(k, \frac{k}{l} \dim V) \\ &\geq \frac{M - M_{0,k}}{M} \frac{k}{k-1} \dim V, \end{aligned}$$

where we used $\dim V \leq k-1$ in the last step. So it suffices to choose M large enough so that $\frac{M - M_{0,k}}{M} \frac{k}{k-1} > 1$. \square

Corollary 9.4. *For every K there exists $\nu = \nu_K$ such that any K^{-1} -separated $\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})$ are ν -transverse in the sense that for every $1 \leq l < k$ and any $x_j \in \alpha_j$ we have*

$$BL((V^{(l)}(x_j))_{j=1}^M) \leq \nu^{-1}.$$

Proof. We have already seen that the BL constants are finite. The uniform upper bound follows from the fact that the BL constants are locally bounded on the set where they are finite and compactness. \square

Remark. The above compactness argument is ineffective. It would be desirable to replace it by an explicit estimate for BL constants.

9.4 Bourgain–Guth argument

We will work with M -linear expressions with M given by Corollary 9.4. We denote $\prod_i A_i := \prod_i A_i := \prod_{i=1}^M A_i^{1/M}$ and $p := k(k+1)$.

For a positive integer K and $0 < \delta < K^{-1}$ we denote by $\text{MulDec}_k(\delta, K)$ the smallest constant such that the inequality

$$L^p_{x \in \mathbb{R}^k} \prod_i \|f_{\alpha_i}\|_{L^p(B(x, K))} \leq \text{MulDec}_k(\delta, K) \prod_{\theta \in \mathcal{P}(\alpha_i, \delta)} \ell^2_{\theta} \|f_{\theta}\|_{L^p(\mathbb{R}^k)} \quad (9.8)$$

holds for every ν_K -transverse tuple $\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})$.

Theorem 9.5. *For each $\varepsilon > 0$ there exists $K \geq 1$ such that for all $0 < \delta < 1$ we have*

$$\text{Dec}(\delta) \lesssim \delta^{-\varepsilon} + \delta^{-\varepsilon} \max_{\delta \leq \delta' \leq 1} (\delta/\delta')^{-\varepsilon} \text{MulDec}(\delta', K). \quad (9.9)$$

Theorem 9.5 is obtained by iterating Corollary 9.7 that is a rescaled version of the following Proposition 9.6. This iteration goes back to [BG11].

Proposition 9.6. *For every $0 < \delta < K^{-1}$ we have*

$$\|f\|_p \lesssim \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_p + K^{M+1} \text{MulDec}(\delta, K) \ell^2_{\theta \in \mathcal{P}(\delta)} \|f_{\theta}\|_p \quad (9.10)$$

Proof of Proposition 9.6. Let $B \subset \mathbb{R}^k$ be a ball of radius K and

$$S_B := \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(B)}.$$

If $\|f\|_{L^p(B)} \geq 2MS_B$, then there exist at least $2M$ cubes $\alpha \in \mathcal{P}(K^{-1})$ with

$$\|f_{\alpha}\|_{L^p(B')} \geq (2K)^{-1} \|f\|_{L^p(B)}.$$

So we can choose M cubes $\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})$ that are K^{-1} -separated and

$$\|f\|_{L^p(B)} \leq K \prod_i \|f_{\alpha_i}\|_{L^p(B)}.$$

Hence in any case

$$\|f\|_{L^p(B)} \lesssim \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(B)} + K \sum_{\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})} \prod_i \|f_{\alpha_i}\|_{L^p(B)},$$

where the sum runs over all K^{-1} -separated tuples. Integrating this inequality over B we obtain

$$\begin{aligned} \|f\|_{L^p(\mathbb{R}^k)} &= L^p_{x \in \mathbb{R}^k} \|f\|_{L^p(B(x, K))} \\ &\leq CL^p_{x \in \mathbb{R}^k} \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(B(x, K))} + KL^p_{x \in \mathbb{R}^k} \sum_{\substack{\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1}) \\ K^{-1}\text{-separated}}} \prod_i \|f_{\alpha_i}\|_{L^p(B(x, K))} \\ &\leq C \ell^2_{\alpha \in \mathcal{P}(K^{-1})} L^p_{x \in \mathbb{R}^k} \|f_{\alpha}\|_{L^p(B(x, K))} + K \sum_{\substack{\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1}) \\ K^{-1}\text{-separated}}} L^p_{x \in \mathbb{R}^k} \prod_i \|f_{\alpha_i}\|_{L^p(B(x, K))} \\ &\leq C \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(\mathbb{R}^k)} + K \sum_{\substack{\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1}) \\ K^{-1}\text{-separated}}} \text{Dec}_k(\delta, K) \prod_{\theta \in \mathcal{P}(\alpha_i, \delta)} \ell^2_{\theta} \|f_{\theta}\|_{L^p(\mathbb{R}^k)} \\ &\leq C \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(\mathbb{R}^k)} + K \sum_{\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})} \text{Dec}_k(\delta, K) \prod_{\theta \in \mathcal{P}(\delta)} \ell^2_{\theta} \|f_{\theta}\|_{L^p(\mathbb{R}^k)} \\ &\leq C \ell^2_{\alpha \in \mathcal{P}(K^{-1})} \|f_{\alpha}\|_{L^p(\mathbb{R}^k)} + K^{M+1} \text{Dec}_k(\delta, K) \ell^2_{\theta \in \mathcal{P}(\delta)} \|f_{\theta}\|_{L^p(\mathbb{R}^k)}. \quad \square \end{aligned}$$

Corollary 9.7. For $0 < \delta < K^{-1}$ we have

$$\text{Dec}(\delta) \leq \max\left(C\text{Dec}(K\delta), K^{M+1}\text{MulDec}(\delta, K)\right). \quad (9.11)$$

Proof of Theorem 9.5. Choose $K \in 2^{k\mathbb{N}}$ so large that the C on the right-hand side of (9.11) is bounded by K^ε . For $\delta < 1/K$ iterate the inequality (9.11) $\lfloor \frac{\log \delta}{\log K} \rfloor$ times and use a trivial estimate for Dec at the end. \square

From Theorem 9.5 it follows that if for some $\eta \geq 0$, all $K \in 2^{\mathbb{N}}$, and all $0 < \delta < 1$ we have

$$\text{MulDec}(\delta, K) \lesssim_K \delta^{-\eta}, \quad (9.12)$$

then we obtain

$$\text{Dec}(\delta) \lesssim_\varepsilon \delta^{-\eta-\varepsilon} \quad (9.13)$$

for every $\varepsilon > 0$. Let η be the smallest exponent in the decoupling inequality.

9.5 Induction on scales

We fix K^{-1} -separated intervals $\alpha_1, \dots, \alpha_M \in \mathcal{P}(K^{-1})$ and functions f_θ . We write

$$n_l = l, \quad \mathcal{K}_l = 1 + \dots + l = \frac{l(l+1)}{2}, \quad p = k(k+1).$$

For $1 \leq l \leq k$ define

$$p_l := p \frac{\mathcal{K}_l}{\mathcal{K}_k} = l(l+1),$$

$$t_l := p \frac{n_l}{n_k} = l(k+1).$$

Here p_l is the sharp decoupling exponent for the l -th moment curve and t_l is an exponent in the BL inequality that we will use.

Define α_l and β_l by

$$\frac{1}{\frac{n_l}{n_k}} = \frac{\alpha_l}{\frac{n_{l+1}}{n_k}} + \frac{1 - \alpha_l}{\frac{\mathcal{K}_l}{\mathcal{K}_k}}, \quad 1 \leq l < k, \quad (9.14)$$

$$\frac{1}{\frac{\mathcal{K}_l}{\mathcal{K}_k}} = \frac{1 - \beta_l}{\frac{\mathcal{K}_{l-1}}{\mathcal{K}_k}} + \frac{\beta_l}{\frac{n_l}{n_k}}, \quad 1 < l < k, \quad (9.15)$$

and $\beta_1 := 1$.

For $2 \leq t \leq p$, $0 < b < 1$, and $b < s$ let

$$A_t(b, s) := L_x^p \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta^b)}^2 \|f_\beta\|_{\mathcal{L}^t(w_{B(x, \delta^{-s})})}. \quad (9.16)$$

The induction on scales argument will involve the quantities

$$A_{t(l)}(b) := A_{t_l}(b, lb),$$

$$A_{p(l)}(b) := A_{p_l}(b, (l+1)b).$$

Here $t(l)$ and $p(l)$ are formal expressions and can be read “of type t with degree l ” and “of type p with degree l ”. For $0 < b < 1$ and $* = t(l), p(l)$ let

$$a_*(b) := \inf\{a \mid A_*(b) \lesssim_{a,K} \delta^{-a} \text{RHS}(9.8) \text{ for all } K\}.$$

9.5.1 Linear decoupling

We can use Hölder to eliminate all multilinearity and use the linear decoupling estimate. For $1 \leq t \leq p$ and $1 \leq l \leq k$ this gives the bound

$$\begin{aligned}
A_t(b, s) &= L_x^p \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta^b)}^2 \|f_\beta\|_{E^t(w_{B(x, \delta^{-lb})})} \\
&\leq \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta^b)}^2 L_x^p \|f_\beta\|_{E^p(w_{B(x, \delta^{-lb})})} \\
&= \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta^b)}^2 \|f_\beta\|_{L^p(\mathbb{R}^k)} \\
&\leq \text{Dec}(\delta^{1-b}) \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\beta\|_p \\
&\lesssim_\varepsilon \delta^{-\eta(1-b)-\varepsilon} \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta)}^2 \|f_\beta\|_p.
\end{aligned} \tag{9.17}$$

This shows

$$a_*(b) \leq \eta(1-b). \tag{9.18}$$

9.5.2 Bourgain–Guth argument

First we estimate the left-hand side of (9.8) by the quantities involved in the iterative procedure. For $1 \leq l \leq k$, $1 \leq t \leq p$, $0 < b \leq s$, and δ sufficiently small so that $\delta^{-s} \geq K$ we have

$$\begin{aligned}
LHS(9.8) &= L_{x \in \mathbb{R}^k}^p \prod \|f_{\alpha_i}\|_{E^p(B(x, K))} \\
&\lesssim L_{x \in \mathbb{R}^k}^p \prod \|f_{\alpha_i}\|_{E^p(B(x, \delta^{-s}))} \\
&\leq L_{x \in \mathbb{R}^k}^p \prod \sum_{\beta \in \mathcal{P}(\alpha_i, \delta^b)} \|f_\beta\|_{E^p(B(x, \delta^{-s}))} \\
&\lesssim \delta^{-b/2-(s-b)k(1/t-1/p)} L_{x \in \mathbb{R}^k}^p \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \delta^b)}^2 \|f_\beta\|_{E^t(w_{B(x, \delta^{-s})})} \\
&\leq \delta^{-Cb} A_*(b).
\end{aligned} \tag{9.19}$$

Here we have used the reverse Hölder inequality (Corollary 3.3) to estimate the E^p norm by the E^t with some loss. This shows

$$\eta \leq Cb + a_*(b). \tag{9.20}$$

9.5.3 Ball inflation

Similar to the paraboloid case we obtain the following results from Keakeya–Brascamp–Lieb inequalities.

Lemma 9.8 (Ball inflation). *Let $1 \leq l < k$, $1 \leq t < \infty$. Let $\rho \leq 1/K$ and let $B \subset \mathbb{R}^k$ be a ball of radius $\rho^{-(l+1)}$. Then we have*

$$L_{x \in B}^{t \frac{k}{l}} \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \rho)}^t \|f_\beta\|_{E^t(w_{B(x, \rho^{-l})})} \lesssim_{\nu, \varepsilon} \rho^{-\varepsilon} \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \rho)}^t \|f_\beta\|_{E^t(w_B)}. \tag{9.21}$$

Corollary 9.9 (Ball inflation). *Let $1 \leq l < k$, $1 \leq q \leq t < \infty$. Let $\rho \leq 1/K$ and let $B \subset \mathbb{R}^k$ be a ball of radius $\rho^{-(l+1)}$. Then we have*

$$L_{x \in B}^{t \frac{k}{l}} \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \rho)}^q \|f_\beta\|_{E^t(w_{B(x, \rho^{-l})})} \lesssim_{\nu, \varepsilon} \rho^{-\varepsilon} \prod \ell_{\beta \in \mathcal{P}(\alpha_i, \rho)}^q \|f_\beta\|_{E^t(w_B)}. \tag{9.22}$$

Note that by (9.14) for $1 \leq l < k$ we have

$$\frac{1}{t_l} = \frac{\alpha_l}{t_{l+1}} + \frac{1 - \alpha_l}{p_l}. \tag{9.23}$$

For $1 \leq l < k$ by Corollary 9.9 and Hölder's inequality together with (9.23) we obtain

$$\begin{aligned}
A_{t(l)}(b) &= A_{t_l}(b, lb) \\
&\lesssim_{\varepsilon, K} \delta^{-b\varepsilon} A_{t_l}(b, (l+1)b) \\
&\lesssim \delta^{-b\varepsilon} A_{t_{l+1}}(b, (l+1)b)^{\alpha_l} A_{p_l}(b, (l+1)b)^{1-\alpha_l} \\
&= \delta^{-b\varepsilon} A_{t(l+1)}(b)^{\alpha_l} A_{p(l)}(b)^{1-\alpha_l}.
\end{aligned} \tag{9.24}$$

This implies

$$a_{t(l)}(b) \leq \alpha_l a_{t(l+1)}(b) + (1 - \alpha_l) a_{p(l)}(b). \tag{9.25}$$

9.5.4 Lower degree decoupling

By (9.15) for $1 < l < k$ we have

$$\frac{1}{p_l} = \frac{1 - \beta_l}{p_{l-1}} + \frac{\beta_l}{t_l}. \tag{9.26}$$

For $1 \leq l < k$ by the localized version of the decoupling inequality for the l -th moment curve and Hölder's inequality with (9.26) we obtain

$$\begin{aligned}
A_{p(l)}(b) &= A_{p_l}(b, (l+1)b) \\
&= L_x^p \prod_{\beta \in \mathcal{P}(\alpha_i, \delta^b)} \ell_{\beta}^2 \|f_{\beta}\|_{\mathcal{E}^{p_l}(w_{B(x, \delta^{-(l+1)b})})} \\
&\lesssim_{\varepsilon} \delta^{-\varepsilon b/l} L_x^p \prod_{\beta \in \mathcal{P}(\alpha_i, \delta^{(l+1)b/l})} \ell_{\beta}^2 \|f_{\beta}\|_{\mathcal{E}^{p_l}(w_{B(x, \delta^{-(l+1)b})})} \\
&\leq \delta^{-\varepsilon b/l} A_{t(l)}\left(\frac{(l+1)b}{l}\right)^{\beta_l} A_{p(l-1)}\left(\frac{(l+1)b}{l}\right)^{1-\beta_l}.
\end{aligned} \tag{9.27}$$

Note that this also holds for $l = 1$ because $p_1 \leq t_1$. This implies

$$a_{p(l)}(b) \leq \beta_l a_{t(l)}((l+1)b/l) + (1 - \beta_l) a_{p(l-1)}((l+1)b/l) \tag{9.28}$$

for $0 < b < l/(l+1)$.

9.5.5 Wrapping up the induction

Proposition 9.10. *The inequalities (9.18), (9.20), (9.25), and (9.28) imply $\eta \leq 0$.*

Proof. We eliminate the dependence on b by setting⁴

$$\tilde{a}_* := \liminf_{b \rightarrow 0} \frac{\eta - a_*(b)}{b}.$$

The hypotheses then imply

$$\tilde{a}_* \geq \eta.$$

$$\tilde{a}_* \leq C.$$

$$\tilde{a}_{t(l)} \geq \alpha_l \tilde{a}_{t(l+1)} + (1 - \alpha_l) \tilde{a}_{p(l)}.$$

$$\tilde{a}_{p(l)}(b) \geq \frac{l+1}{l} (\beta_l \tilde{a}_{t(l)} + (1 - \beta_l) \tilde{a}_{p(l-1)}).$$

⁴This definition is from Tao's blog post

<https://terrytao.wordpress.com/2019/06/14/abstracting-induction-on-scales-arguments/>

Using (9.28) let us verify the last of these inequalities, which is also the least obvious:

$$\begin{aligned}
\tilde{a}_{p(l)}(b) &= \liminf_{b \rightarrow 0} \frac{\eta - a_{p(l)}(b)}{b} \\
&\geq \liminf_{b \rightarrow 0} \frac{\eta - \beta_l a_{t(l)}((l+1)b/l) - (1 - \beta_l) a_{p(l-1)}((l+1)b/l)}{b} \\
&\geq \frac{l+1}{l} \beta_l \liminf_{b \rightarrow 0} \frac{\eta - a_{t(l)}((l+1)b/l)}{(l+1)b/l} + \frac{l+1}{l} (1 - \beta_l) \liminf_{b \rightarrow 0} \frac{\eta - a_{p(l-1)}((l+1)b/l)}{(l+1)b/l} \\
&= \frac{l+1}{l} (\beta_l \tilde{a}_{t(l)} + (1 - \beta_l) \tilde{a}_{p(l-1)}).
\end{aligned}$$

Let

$$W := \sum_{l=1}^{k-1} (l \tilde{a}_{p(l)} + 2l \tilde{a}_{t(l)}).$$

Applying all our inequalities we get

$$\begin{aligned}
W &\geq \sum_{l=1}^{k-1} (l+1) (\beta_l \tilde{a}_{t(l)} + (1 - \beta_l) \tilde{a}_{p(l-1)}) + \sum_{l=1}^{k-1} 2l (\alpha_l \tilde{a}_{t(l+1)} + (1 - \alpha_l) \tilde{a}_{p(l)}) \\
&= 2(k-1) \alpha_{k-1} \tilde{a}_{t(k)} + \sum_{l=1}^{k-1} ((l+1) \beta_l + 2(l-1) \alpha_{l-1}) \tilde{a}_{t(l)} \\
&\quad + \sum_{l=1}^{k-1} ((l+2)(1 - \beta_{l+1}) + 2l(1 - \alpha_l)) \tilde{a}_{p(l)},
\end{aligned}$$

where by convention $\alpha_0 = 0$, $\beta_k = 1$. Now from (9.14) and (9.15) we compute

$$\begin{aligned}
\frac{1}{l/k} &= \frac{\alpha_l}{(l+1)/k} + \frac{1 - \alpha_l}{l(l+1)/(k(k+1))}, \quad \frac{1}{l(l+1)/(k(k+1))} = \frac{1 - \beta_l}{(l-1)l/(k(k+1))} + \frac{\beta_l}{l/k}. \\
(l+1)/(k+1) &= \frac{l}{k+1} \alpha_l + (1 - \alpha_l), \quad \frac{1}{(l+1)/(k+1)} = \frac{1 - \beta_l}{(l-1)/(k+1)} + \beta_l. \\
\alpha_l &= \frac{1 - (l+1)/(k+1)}{1 - l/(k+1)}, \quad \beta_l = \frac{(k+1)/(l+1) - (k+1)/(l-1)}{1 - (k+1)/(l-1)} \\
\alpha_l &= \frac{k-l}{k+1-l}, \quad \beta_l = \frac{(k+1)(l-1) - (l+1)(k+1)}{((l-1) - (k+1))(l+1)} = \frac{2(k+1)}{(l+1)(k-l+2)}
\end{aligned}$$

For $1 < l < k$ we get

$$\begin{aligned}
(l+1) \beta_l + 2(l-1) \alpha_{l-1} &= \frac{2(k+1)}{k-l+2} + 2(l-1) \frac{k-l+1}{k-l+2} \\
&= 2l + \frac{2(k+1)}{k-l+2} - \frac{k-l+1}{k-l+2} - 2l \frac{1}{k-l+2} = 2l.
\end{aligned}$$

For $l = 1$ this is easier. For $1 \leq l \leq k-2$ we get

$$(l+2)(1 - \beta_{l+1}) + 2l(1 - \alpha_l) = (l+2) - \frac{2(k+1)}{k-l+1} + 2l \frac{1}{k-l+1} = l.$$

For $l = k-1$ this is again easier. Hence we get

$$W \geq W + 2(k-1) \alpha_{k-1} \tilde{a}_{t(k)},$$

so $\eta \leq \tilde{a}_{t(k)} \leq 0$. □

Remark. The coefficients used to define W are the Perron–Frobenius eigenvector of the matrix of coefficients of the inequalities for \tilde{a}_* 's.

A Estimates for Weyl sums

[Vau97, Theorem 5.2] Let

$$f_k(\vec{\alpha}; X) := \sum_{x=1}^X e(\alpha_1 x + \cdots + \alpha_k x^k).$$

Theorem A.1. *Let k be an integer with $k > 2$, and let $\vec{\alpha} \in \mathbb{R}^k$. Suppose that there exists a natural number j with $2 \leq j \leq k$ such that, for some $a \in \mathbb{Z}$ and $q \in \mathbb{N}$ with $(a, q) = 1$, one has $|\alpha_j - a/q| \leq q^{-2}$ and $q \leq X^j$. Then one has*

$$f_k(\vec{\alpha}; X) \lesssim X^{1+\varepsilon} (q^{-1} + X^{-1} + qX^{-j})^{1/(k(k-1))}$$

With notation from [Vau97, Section 5.2]

Lemma A.2. *For $2 \leq j \leq k$ and integers $1 \leq x, y \leq N$ we have*

$$\|(j \cdots k)\alpha_j(x - y)\| \lesssim \sum_{h=j-1}^{k-1} \|\gamma_h(x) - \gamma_h(y)\| N^{h-j+1}$$

Proof. By downward induction on j . Suppose that the claim is already known for all larger j . Expand

$$\begin{aligned} \gamma_{j-1}(x) - \gamma_{j-1}(y) &= \sum_{h=j-1}^k \alpha_h \binom{h}{j-1} ((-x)^{h-j+1} - (-y)^{h-j+1}) \\ &= \sum_{h=j}^k (-1)^{h-j+1} \alpha_h \binom{h}{j-1} (x-y)(x^{h-j} + x^{h-j-1}y + \cdots + y^{h-j}) \end{aligned}$$

Hence

$$\begin{aligned} \|(j \cdots k)\alpha_j(x - y)\| &= \|((j+1) \cdots k)\alpha_j \binom{j}{j-1} (x - y)\| \\ &\leq ((j+1) \cdots k) \|\gamma_{j-1}(x) - \gamma_{j-1}(y)\| \\ &\quad + \sum_{h=j+1}^k \|((j+1) \cdots k)\alpha_h \binom{h}{j-1} (x-y)(x^{h-j} + \cdots + y^{h-j})\| \\ &\lesssim \|\gamma_{j-1}(x) - \gamma_{j-1}(y)\| + \sum_{h=j+1}^k \|(h \cdots k)\alpha_h(x-y)\| (x^{h-j} + \cdots + y^{h-j}) \\ &\lesssim \|\gamma_{j-1}(x) - \gamma_{j-1}(y)\| + \sum_{h=j+1}^k N^{h-j} \sum_{h'=h-1}^{k-1} \|\gamma_{h'}(x) - \gamma_{h'}(y)\| N^{h'-h+1} \\ &= \|\gamma_{j-1}(x) - \gamma_{j-1}(y)\| + \sum_{h=j+1}^k \sum_{h'=h-1}^{k-1} \|\gamma_{h'}(x) - \gamma_{h'}(y)\| N^{h'-j+1} \\ &\sim \sum_{h'=j-1}^{k-1} \|\gamma_{h'}(x) - \gamma_{h'}(y)\| N^{h'-j+1}. \quad \square \end{aligned}$$

B Maximal operators along separated directions in \mathbb{R}^2

Keakeya maximal operators on \mathbb{R}^2 were first estimated in [Cor77] (for a nice proof of a stronger result see [BO10]). Keakeya maximal operators are controlled by maximal operators along separated directions, and these were first estimated in [Cór77]. In this section we prove the sharp estimate for such operators obtained in [Str78].

Lemma B.1. *Let $N \geq 2$ be an integer, Ω a $1/N$ -separated set of directions in \mathbb{R}^2 and let \mathcal{R} a set of rectangles in \mathbb{R}^2 for which one of the sides points in a direction in Ω . Then there exists a subset $\mathcal{G} \subset \mathcal{R}$ such that*

$$\int_{\mathbb{R}^2} \left(\sum_{R \in \mathcal{G}} \mathbf{1}_R(x) \right)^2 dx \leq C \left| \bigcup_{R \in \mathcal{G}} R \right|, \quad (\text{B.1})$$

$$\left| \bigcup_{R \in \mathcal{R}} R \right| \leq C(\log N) \left| \bigcup_{R \in \mathcal{G}} R \right|. \quad (\text{B.2})$$

The separation hypothesis on Ω implies that $|\Omega| \lesssim N$. The separation hypothesis was replaced by this cardinality hypothesis in [Kat99] (an alternative proof appeared in [ASV03]). We will only use separated Ω 's.

Proof. Without loss of generality we may assume that \mathcal{R} is finite. Partitioning \mathcal{R} into boundedly many pieces and using rotation invariance we may assume that the long sides of rectangles in \mathcal{R} point in the directions in Ω and that these directions are all within $1/10$ of some fixed direction.

We choose $\mathcal{G} = \{R_1, R_2, \dots\}$ in the following way. Suppose that R_1, \dots, R_{n-1} are already chosen. Then let R_n be a rectangle in \mathcal{R} with the longest long side subject to the condition

$$\sum_{j=1}^{n-1} |R_j \cap R_n| \leq \frac{1}{2} |R_n| \quad (\text{B.3})$$

if such R_n exists, or stop the construction otherwise. By (B.3) we have

$$|R_n \setminus \bigcup_{j < n} R_j| \geq |R_n| - \sum_{j=1}^{n-1} |R_j \cap R_n| \geq \frac{1}{2} |R_n|,$$

and it follows that

$$\begin{aligned} \int_{\mathbb{R}^2} \left(\sum_n \mathbf{1}_{R_n} \right)^2 &\leq 2 \sum_{j \leq n} |R_j \cap R_n| \leq 3 \sum_n |R_n| \\ &\leq 6 \sum_n |R_n \setminus \bigcup_{j < n} R_j| = 6 \left| \bigcup_n R_n \right|. \end{aligned} \quad (\text{B.4})$$

This shows (B.1). By the weak type (1, 1) inequality for the Hardy–Littlewood maximal operator M the remaining inequality (B.2) will follow from

$$\bigcup_{R \in \mathcal{R} \setminus \mathcal{G}} R \subset \{M(\sum_n \mathbf{1}_{\tilde{R}_n}) > c/\log N\}, \quad (\text{B.5})$$

where \tilde{R}_n is the rectangle with the same center, width, and orientation as R_n , and length multiplied by 4.

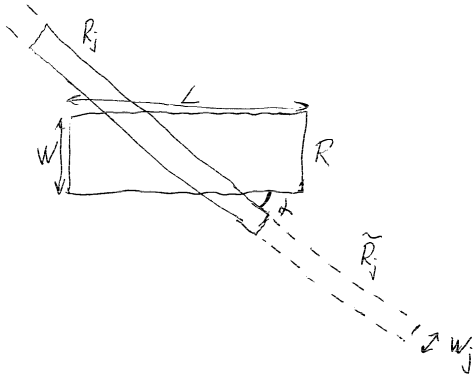
If $R \in \mathcal{R} \setminus \mathcal{G}$, then (B.3) fails for R for all large n . Let n be the smallest index for which (B.3) fails. Then

$$\sum_{j=1}^{n-1} |R_j \cap R| \geq \frac{1}{2} |R|$$

and the long side of R_j is longer than the long side of R for all $j < n$, since otherwise R_j could not have been selected. Denote by θ_j the angle between long sides of R and R_j and by w_j the width of R_j .

Let L be the length of R , W its width, and $0 \leq \alpha \leq 1/10$. Let $s := 8 \max(W, \alpha L)$. Suppose that $\alpha \leq \theta_j \leq 2\alpha$. We claim that for every cube Q with side length s and center in R we have

$$\frac{|\tilde{R}_j \cap Q|}{|Q|} \gtrsim \frac{|R_j \cap R|}{|R|}.$$



If $w_j \geq s$, then in fact both sides are ~ 1 , so assume $w_j < s$. Then

$$|\tilde{R}_j \cap Q| \gtrsim sw_j,$$

and also

$$|R_j \cap R| \lesssim \min(Lw_j, Ww_j/\alpha).$$

Hence

$$\frac{|R_j \cap R|}{|R|} \lesssim w_j \min(L, W/\alpha)/(LW) = w_j/\max(W, \alpha L) \sim w_j/s \lesssim \frac{|\tilde{R}_j \cap Q|}{|Q|}$$

as claimed.

By the pigeon principle there exists $\alpha \in \{0, 2^0 \frac{2\pi}{N}, 2^1 \frac{2\pi}{N}, \dots\}$ such that for the set $J = \{j < n \mid \alpha \leq \theta_j \leq 2\alpha\}$ we have

$$\sum_{j \in J} |R_j \cap R| \gtrsim (\log N)^{-1} |R|.$$

Considering the cube Q with side length s centered at any $x \in R$ we obtain

$$\begin{aligned} M\left(\sum_n \mathbf{1}_{\tilde{R}_n}\right)(x) &\geq M\left(\sum_{j \in J} \mathbf{1}_{\tilde{R}_j}\right)(x) \\ &\geq |Q|^{-1} \sum_{j \in J} |\tilde{R}_j \cap Q| \\ &\geq \sum_{j \in J} |R|^{-1} |R_j \cap R| \\ &\gtrsim (\log N)^{-1}. \end{aligned} \quad \square$$

Theorem B.2 ([Str78]). *Let $N \geq 2$ be an integer and Ω a $1/N$ -separated set of directions in \mathbb{R}^2 . Let M_Ω be the maximal operator associated to averaging over rectangles one of whose sides points in a direction in Ω . Then*

$$\|M_\Omega f\|_{2,\infty} \lesssim (\log N)^{1/2} \|f\|_2. \quad (\text{B.6})$$

Proof. Let \mathcal{R} be the set of all rectangles R with sides pointing in directions in Ω for which $|R|^{-1} \int_R |f| > \lambda$. Let $\mathcal{G} \subset \mathcal{R}$ be as in Lemma B.1. Then

$$\lambda^2 \left| \bigcup_{R \in \mathcal{G}} R \right|^2 \leq \lambda^2 \left(\sum_R |R| \right)^2 < \left(\sum_R \int_R |f| \right)^2 \leq \|f\|_2^2 \sum_{R \in \mathcal{G}} \mathbf{1}_R \|2\|_2^2 \lesssim \|f\|_2^2 \left| \bigcup_{R \in \mathcal{G}} R \right|,$$

and it follows that

$$|\{M_\Omega f > \lambda\}| = \left| \bigcup_{R \in \mathcal{R}} R \right| \lesssim (\log N) \left| \bigcup_{R \in \mathcal{G}} R \right| \lesssim (\log N) \lambda^{-2} \|f\|_2^2. \quad \square$$

We will need the following interpolation argument. It was already used in [Str78], but the proof was not written down, so several later papers prove versions of this result [CHS91; Kat99; Dem10].

Lemma B.3. *Let $N \geq 2$, $1 \leq p < 2$, and T be a quasisubadditive operator such that*

$$\|Tf\|_{p,\infty} \leq N\|f\|_p, \quad \|Tf\|_\infty \leq N\|f\|_\infty, \quad \|Tf\|_{2,\infty} \leq \|f\|_2.$$

Then

$$\|Tf\|_2 \lesssim (\log N)^{1/2} \|f\|_2.$$

Proof. The basic idea is to use the layer cake representation

$$\|Tf\|_2^2 \sim \int_0^\infty \lambda |\{Tf > \lambda\}| d\lambda,$$

for each λ decompose

$$f = f\mathbf{1}_{|f| < N^\alpha \lambda} + f\mathbf{1}_{N^\alpha \lambda \leq |f| \leq N^\beta \lambda} + f\mathbf{1}_{N^\beta \lambda < |f|},$$

use the three different estimates for these pieces and optimize α, β . The rest is an exercise. \square

The various versions of Lemma B.3 in [CHS91; Kat99; Dem10] have a common formulation in terms of real interpolation spaces. We record it here using notation and results from [BL76a]. This general version will not be used in this course.

Lemma B.4. *Let \bar{A}, \bar{B} be compatible couples of Banach spaces, $0 < \theta < 1$, and $1 \leq s \leq q \leq \infty$. Let $N \geq 2$ and let T be a (quasisub-)additive operator such that*

$$\|Ta\|_{B_0} \leq N\|a\|_{A_0}, \quad \|Ta\|_{B_1} \leq N\|a\|_{A_1}, \quad \|Ta\|_{B_{\theta,\infty}} \leq \|a\|_{A_{\theta,s}}.$$

Then

$$\|Ta\|_{B_{\theta,q}} \lesssim (\log N)^{1/s} \|a\|_{A_{\theta,q}}. \quad (\text{B.7})$$

Interpolating between (B.7) with $q = \infty$ and the $A_{\theta,s} \rightarrow B_{\theta,\infty}$ estimate in the hypothesis using the reiteration theorem we further obtain

$$\|Ta\|_{B_{\theta,\infty}} \lesssim (\log N)^{1/s-1/r} \|a\|_{A_{\theta,r}}, \quad s \leq r \leq \infty. \quad (\text{B.8})$$

This allows to recover [CHS91, Proposition 5(i)].

Proof. By the fundamental lemma of interpolation theory there exists a decomposition

$$a = \sum_{j \in \mathbb{Z}} a_j \quad \text{with } J(2^j, a_j) \lesssim K(2^j, a).$$

By subadditivity of the K -functional we split

$$K(2^j, Ta) \lesssim K(2^j, T(\sum_{k < -A} a_{j+k})) + K(2^j, T(\sum_{|k| \leq A} a_{j+k})) + K(2^j, T(\sum_{k > A} a_{j+k})).$$

Now we estimate the individual terms.

$$\begin{aligned} K(2^j, T(\sum_{k < -A} a_{j+k})) &\leq \|T(\sum_{k < -A} a_{j+k})\|_{B_0} \leq N \sum_{k < -A} \|a_{j+k}\|_{A_0} \\ &\leq N \sum_{k < -A} \|a_{j+k}\|_{A_0} \leq N \sum_{k < -A} J(2^{j+k}, a_{j+k}) \leq N \sum_{k < -A} K(2^{j+k}, a) \end{aligned}$$

$$\begin{aligned}
K(2^j, T(\sum_{|k|\leq A} a_{j+k})) &\leq 2^{j\theta} \|T(\sum_{|k|\leq A} a_{j+k})\|_{B_{\theta,\infty}} \leq 2^{j\theta} \|\sum_{|k|\leq A} a_{j+k}\|_{A_{\theta,s}} \\
&\lesssim 2^{j\theta} \left(\sum_{|k|\leq A} (2^{-(j+k)\theta} J(2^{j+k}, a_{j+k}))^s \right)^{1/s} \lesssim 2^{j\theta} \left(\sum_{|k|\leq A} (2^{-(j+k)\theta} K(2^{j+k}, a))^s \right)^{1/s}.
\end{aligned}$$

$$\begin{aligned}
K(2^j, T(\sum_{k>A} a_{j+k})) &\leq 2^j \|T(\sum_{k>A} a_{j+k})\|_{B_1} \leq 2^j N \|\sum_{k>A} a_{j+k}\|_{A_1} \\
&\leq 2^j N \sum_{k>A} \|a_{j+k}\|_{A_1} \leq 2^j N \sum_{k>A} 2^{-(j+k)} J(2^{j+k}, a_{j+k}) \leq N \sum_{k>A} 2^{-k} K(2^{j+k}, a).
\end{aligned}$$

$$\begin{aligned}
\|Ta\|_{B_{\theta,q}} &\sim \left(\sum_{j\in\mathbb{Z}} (2^{-j\theta} K(2^j, Ta))^q \right)^{1/q} \\
&\lesssim \left(\sum_{j\in\mathbb{Z}} \left(\sum_{k<-A} 2^{k\theta} N 2^{-(j+k)\theta} K(2^{j+k}, a) + \left(\sum_{|k|\leq A} (2^{-(j+k)\theta} K(2^{j+k}, a))^s \right)^{1/s} \right)^q \right)^{1/q} \\
&\quad + \sum_{k>A} 2^{-k(1-\theta)} N 2^{-(j+k)\theta} K(2^{j+k}, a) \Big)^{1/q} \\
&\leq \sum_{k<-A} 2^{k\theta} N \left(\sum_{j\in\mathbb{Z}} \left(2^{-(j+k)\theta} K(2^{j+k}, a) \right)^q \right)^{1/q} \\
&\quad + \left(\sum_{j\in\mathbb{Z}} \left(\left(\sum_{|k|\leq A} (2^{-(j+k)\theta} K(2^{j+k}, a))^s \right)^{1/s} \right)^q \right)^{1/q} \\
&\quad + \sum_{k>A} 2^{-k(1-\theta)} N \left(\sum_{j\in\mathbb{Z}} \left(2^{-(j+k)\theta} K(2^{j+k}, a) \right)^q \right)^{1/q} \\
&\lesssim \sum_{k<-A} 2^{k\theta} N \|a\|_{A_{\theta,q}} + (2A+1)^{1/s} \|a\|_{A_{\theta,q}} + \sum_{k>A} 2^{-k(1-\theta)} N \|a\|_{A_{\theta,q}}.
\end{aligned}$$

Choosing $A \sim \log N$ we obtain the claim. \square

Corollary B.5. *In the setting of Theorem B.2 we have*

$$\|M_\Omega f\|_2 \lesssim (\log N) \|f\|_2.$$

Proof. Apply Lemma B.4 to the operator $(\log N)^{-1} M_\Omega$ with estimate $O(N)$ for the $L^p \rightarrow L^p$ norm, $1 < p < 2$, estimate $O(1)$ for the $L^\infty \rightarrow L^\infty$ norm, and estimate (B.6) for the $L^2 \rightarrow L^{2,\infty}$ norm. \square

C Córdoba sector square function on \mathbb{R}^2

We begin with a square function inequality for Fourier multipliers adapted to a collection of *congruent* finitely overlapping rectangles of fixed orientation in \mathbb{R}^n . In [Rub85] this result is attributed to Carleson (in dimension $n = 1$, although the higher-dimensional case does not present additional difficulties). There is also an alternative proof by Córdoba [Cór81] that I have not looked up.

There are also bounds for square functions associated to *arbitrary* finitely overlapping rectangles of fixed orientation. In dimension $n = 1$ it is due to Rubio de Francia [Rub85], and in higher dimensions to Lacey [Lac07]. We will not need such general results.

Lemma C.1. *For every $n \in \mathbb{N}$ and $N > 0$ there exists $M < \infty$ such that the following holds. Let Ξ be a 1-separated set in \mathbb{R}^n and for each $\xi \in \Xi$ let φ_ξ be a Schwartz function with M -th Schwartz norm bounded by 1. Then*

$$\sum_{\xi \in \Xi} \left| \int_{\mathbb{R}^n} \varphi_\xi(x) f(x) e(x \cdot \xi) dx \right|^2 \lesssim_{n,N} \int_{\mathbb{R}^n} |f(x)|^2 (1 + |x|)^{-N} dx. \quad (\text{C.1})$$

Proof. By Plancherel's theorem it is easy to obtain the estimate (C.1) with $N = 0$. To pass to larger N split the domain of f into a ball of radius 1 and dyadic shells of radii $R \in 2^{\mathbb{N}}$. For a dyadic shell of radius R let χ_R be a smooth approximation of its characteristic function. Then, for $C = C(M, N)$ large enough,

$$\|\chi_R \varphi\|_{M\text{-Schwartz}} \lesssim R^{-10N} \|\chi_R \varphi\|_{(M+C)\text{-Schwartz}}.$$

Hence we can sum up the contributions of the dyadic shells. \square

Corollary C.2. *Let \mathcal{Q} be a boundedly overlapping collection of unit cubes in \mathbb{R}^n . For each $Q \in \mathcal{Q}$ let m_Q be a bump function adapted to Q (with sufficiently many derivatives depending on n). Then for $2 \leq p \leq \infty$ we have*

$$\left\| \left(\sum_{Q \in \mathcal{Q}} |\mathcal{F}^{-1}(m_Q \hat{f})|^2 \right)^{1/2} \right\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^p(\mathbb{R}^n)}.$$

By scaling and rotation invariance we can replace \mathcal{Q} by a collection of rectangles of any fixed size with sides parallel to some coordinate axes.

Proof. Let $w \in L^{(p/2)'}(\mathbb{R}^n)$. Applying Lemma C.1 to the shifted function $f(x - \cdot)$ we obtain

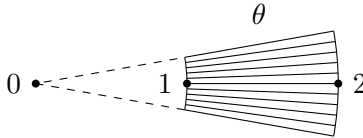
$$\sum_{Q \in \mathcal{Q}} |\mathcal{F}^{-1}(m_Q \hat{f})(x)|^2 \lesssim \int_{\mathbb{R}^n} |f(x - y)|^2 (1 + |y|)^{-N} dy$$

for some $N > n$ provided that we assume sufficiently many derivative bounds on m_Q 's. Multiplying this inequality by $w(x)$ and integrating in x we obtain

$$\begin{aligned} \int_{\mathbb{R}^n} \sum_{Q \in \mathcal{Q}} |\mathcal{F}^{-1}(m_Q \hat{f})(x)|^2 w(x) dx &\lesssim \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x - y)|^2 (1 + |y|)^{-N} dy w(x) dx \\ &= \int_{\mathbb{R}^n} |f(x)|^2 (w * (1 + |\cdot|)^{-N})(x) dx \\ &\leq \|f\|_p^2 \|w * (1 + |\cdot|)^{-N}\|_{(p/2)'} \lesssim \|f\|_p^2 \|w\|_{(p/2)'}. \end{aligned}$$

Taking the supremum over all w with $\|w\|_{(p/2)'} = 1$ and using duality we obtain the claim. \square

Remark. Square function inequalities for *arbitrary* collections of cubes [Rub85; Lac07] hold only in the range $2 \leq p < \infty$. A heuristic explanation for this smaller range is that instead of the convolution of w with a fixed function we would see convolutions with varying functions, so we would need to estimate some maximal function of w , which is not possible for $p = \infty$ since in this case $(p/2)' = 1$. We emphasize however that the above proof is inadequate for more general collections of rectangles.



Proposition C.3. *Let θ be a collection of disjoint truncated sectors of angle $\sim 1/N$ as in the figure. For each θ let P_θ be a smooth Fourier multiplier operator adapted to θ . Then*

$$L^4 \ell_\theta^2 |P_\theta * f| \lesssim (\log N)^{1/4} L^4 f. \quad (\text{C.2})$$

Proof. Expanding the 4-th power of the left-hand side of (C.2) we obtain

$$\int_{\mathbb{R}^2} \sum_{\theta, \theta'} |P_\theta f|^2 |P_{\theta'} f|^2.$$

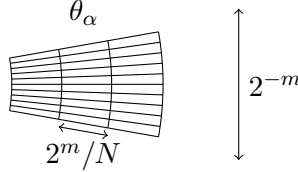
We split the summation

$$\sum_{m=0}^{\log_2 N} \sum_{\theta, \theta': \text{dist}(\theta, \theta') \sim 2^{-m}} \int_{\mathbb{R}^2} |P_\theta f|^2 |P_{\theta'} f|^2.$$

It suffices to obtain a uniform (in N) bound for each m -summand. All estimates below will be uniform in N .

For the diagonal term $\theta = \theta'$ we obtain an $L^4 \rightarrow L^4 \ell^4$ estimate by interpolation between the $L^2 \rightarrow L^2 \ell^2$ estimate which follows from Plancherel's theorem and the trivial $L^\infty \rightarrow L^\infty \ell^\infty$ estimate.

Now fix m and partition each θ in pieces θ_α , $\alpha = 1, \dots, CN/2^m$, of size $1/N \times 2^m/N$. Split $P_\theta = \sum_\alpha P_{\theta_\alpha}$ accordingly (each piece is a smooth multiplier).



For θ, θ' with angular distance $\sim 2^{-m}$ the sumsets $\theta_\alpha + \theta'_\beta$ have bounded overlap as α, β vary. Hence

$$\begin{aligned} \sum_{\theta, \theta': \text{dist}(\theta, \theta') \sim 2^{-m}} \int_{\mathbb{R}^2} |P_\theta f|^2 |P_{\theta'} f|^2 &= \sum_{\theta, \theta': \text{dist}(\theta, \theta') \sim 2^{-m}} \int_{\mathbb{R}^2} \left| \sum_{\alpha, \beta} P_{\theta_\alpha} f P_{\theta'_\beta} f \right|^2 \\ &\lesssim \sum_{\theta, \theta': \text{dist}(\theta, \theta') \sim 2^{-m}} \sum_{\alpha, \beta} \int_{\mathbb{R}^2} |P_{\theta_\alpha} f P_{\theta'_\beta} f|^2. \end{aligned}$$

We split the sum over θ into pieces with $\sim N/2^m$ consecutive summands. For each piece denote by Θ the set of all θ and θ' that give nontrivial contribution, then each Θ contains $\sim N/2^m$ consecutive summands and Θ 's have bounded overlap. We estimate the previous display by

$$\leq \sum_{\Theta} \sum_{\theta, \theta' \in \Theta} \sum_{\alpha, \beta} \int_{\mathbb{R}^2} |P_{\theta_\alpha} f P_{\theta'_\beta} f|^2 = \sum_{\Theta} \int_{\mathbb{R}^2} \left(\sum_{\theta \in \Theta} \sum_{\alpha} |P_{\theta_\alpha} f|^2 \right)^2,$$

By the already mentioned $L^4 \rightarrow L^4 \ell^4$ bound it suffices to consider each Θ individually. Now for fixed Θ all sectors θ_α , $\theta \in \Theta$, are comparable to rectangular boxes of size $\sim 1/N \times 2^m/N$ of a *fixed orientation*. To see this assume without loss of generality that the sectors in Θ are close to horizontal. Then rescale them by a factor 2^{-m} in the horizontal direction. After rescaling the slopes of all sectors θ_α are still $\lesssim 1$, and both their width and height become $\sim 1/N$. Hence we can apply a suitably rescaled version of Corollary C.2. \square

Remark. The truncation of the sectors can be removed as shown in [Cór82; Cór83] at the cost of another power of $\log N$. It is also possible to replace the smooth Fourier multipliers by sharp Fourier cutoffs. To this end one can use the so-called Meyer vector-valued inequality

$$L^4 \ell_{j,l}^2 |T_j f_{j,l}| \lesssim (\log N)^C L^4 \ell_{j,l}^2 |f_{j,l}|, \quad (\text{C.3})$$

where T_j , $j = 1, \dots, N$, are Hilbert transforms along $1/N$ -separated directions in \mathbb{R}^2 . The inequality (C.3) was first proved in [Cór77] using a weighted inequality for the Hilbert transform and bounds for a directional maximal operator in \mathbb{R}^2 . Since the separation hypothesis for directional maximal operators was later removed by Katz [Kat99], this approach also yields (C.3) with an arbitrary set of N directions.

In [DP19, Section 8.2] it is worked out that the weighted approach, when paired with sharp weighted estimates for the Hilbert transform, gives (C.3) with bound $(\log N)^1$.

The article [DP19] also provides an alternative approach to estimates such as (C.2) and (C.3).

References

- [ASV03] A. Alfonseca, F. Soria, and A. Vargas. “An almost-orthogonality principle in L^2 for directional maximal functions”. In: *Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001)*. Vol. 320. Contemp. Math. Amer. Math. Soc., Providence, RI, 2003, pp. 1–7. MR: 1979927 (cit. on p. 64).
- [Bal89] K. Ball. “Volumes of sections of cubes and related problems”. In: *Geometric aspects of functional analysis (1987–88)*. Vol. 1376. Lecture Notes in Math. Springer, Berlin, 1989, pp. 251–260. MR: 1008726 (cit. on p. 47).
- [Bar+07] J. A. Barceló, J. M. Bennett, A. Carbery, A. Ruiz, and M. C. Vilela. “Some special solutions of the Schrödinger equation”. In: *Indiana Univ. Math. J.* 56.4 (2007), pp. 1581–1593. arXiv: math/0604448. MR: 2354692 (cit. on p. 43).
- [Bar98] F. Barthe. “On a reverse form of the Brascamp-Lieb inequality”. In: *Invent. Math.* 134.2 (1998), pp. 335–361. arXiv: math/9705210. MR: 1650312(99i:26021) (cit. on pp. 46, 47).
- [BCT06] J. Bennett, A. Carbery, and T. Tao. “On the multilinear restriction and Kakeya conjectures”. In: *Acta Math.* 196.2 (2006), pp. 261–302. arXiv: math/0509262. MR: 2275834 (cit. on pp. 6, 8).
- [BD15] J. Bourgain and C. Demeter. “The proof of the l^2 decoupling conjecture”. In: *Ann. of Math. (2)* 182.1 (2015), pp. 351–389. arXiv: 1403.5335 [math.CA]. MR: 3374964 (cit. on pp. 10, 11, 25).
- [BD17] J. Bourgain and C. Demeter. “A study guide for the l^2 decoupling theorem”. In: *Chin. Ann. Math. Ser. B* 38.1 (2017), pp. 173–200. arXiv: 1604.06032 [math.CA]. MR: 3592159 (cit. on pp. 10, 14, 16, 17, 22).
- [BDG16] J. Bourgain, C. Demeter, and L. Guth. “Proof of the main conjecture in Vinogradov’s mean value theorem for degrees higher than three”. In: *Ann. of Math. (2)* 184.2 (2016), pp. 633–682. arXiv: 1512.01565 [math.NT]. MR: 3548534 (cit. on p. 55).
- [BDG17] J. Bourgain, C. Demeter, and S. Guo. “Sharp bounds for the cubic Parsell-Vinogradov system in two dimensions”. In: *Adv. Math.* 320 (2017), pp. 827–875. arXiv: 1608.06346 [math.NT]. MR: 3709122 (cit. on p. 55).
- [Bec75] W. Beckner. “Inequalities in Fourier analysis”. In: *Ann. of Math. (2)* 102.1 (1975), pp. 159–182. MR: 0385456 (cit. on p. 46).
- [Ben+08] J. Bennett, A. Carbery, M. Christ, and T. Tao. “The Brascamp-Lieb inequalities: finiteness, structure and extremals”. In: *Geom. Funct. Anal.* 17.5 (2008), pp. 1343–1415. arXiv: math/0505065. MR: 2377493 (cit. on pp. 46, 47, 53).
- [Ben+17] J. Bennett, N. Bez, M. G. Cowling, and T. C. Flock. “Behaviour of the Brascamp-Lieb constant”. In: *Bull. Lond. Math. Soc.* 49.3 (2017), pp. 512–518. arXiv: 1605.08603 [math.CA]. MR: 3723636 (cit. on p. 52).
- [Ben+18a] J. Bennett, N. Bez, S. Buschenhenke, M. Cowling, and T. C. Flock. “On the nonlinear Brascamp-Lieb inequality”. Preprint. 2018. arXiv: 1811.11052 [math.CA] (cit. on p. 52).
- [Ben+18b] J. Bennett, N. Bez, T. C. Flock, and S. Lee. “Stability of the Brascamp-Lieb constant and applications”. In: *Amer. J. Math.* 140.2 (2018), pp. 543–569. arXiv: 1508.07502 [math.CA]. MR: 3783217 (cit. on p. 52).
- [BG11] J. Bourgain and L. Guth. “Bounds on oscillatory integral operators based on multilinear estimates”. In: *Geom. Funct. Anal.* 21.6 (2011), pp. 1239–1295. arXiv: 1012.3760 [math.CA]. MR: 2860188 (cit. on p. 58).
- [BHS18] D. Beltran, J. Hickman, and C. D. Sogge. “Sharp local smoothing estimates for Fourier integral operators”. Preprint. Dec. 2018. arXiv: 1812.11616 [math.AP] (cit. on p. 37).
- [BL76a] J. Bergh and J. Löfström. *Interpolation spaces. An introduction*. Grundlehren der Mathematischen Wissenschaften, No. 223. Berlin: Springer-Verlag, 1976, pp. x+207. MR: 0482275(58#2349) (cit. on p. 66).

- [BL76b] H. J. Brascamp and E. H. Lieb. “Best constants in Young’s inequality, its converse, and its generalization to more than three functions”. In: *Advances in Math.* 20.2 (1976), pp. 151–173. MR: 0412366(54#492) (cit. on p. 46).
- [BLL74] H. J. Brascamp, E. H. Lieb, and J. M. Luttinger. “A general rearrangement inequality for multiple integrals”. In: *J. Functional Analysis* 17 (1974), pp. 227–237. MR: 0346109(49#10835) (cit. on p. 46).
- [BO10] J. A. Barrionuevo and L. S. Oliveira. “A note on maximal averages in the plane”. In: *Proc. Amer. Math. Soc.* 138.1 (2010), pp. 309–313. arXiv: 0810.0911 [math.CA]. MR: 2550196 (cit. on p. 63).
- [Bou13] J. Bourgain. “Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces”. In: *Israel J. Math.* 193.1 (2013), pp. 441–458. arXiv: 1107.1129 [math.CA]. MR: 3038558 (cit. on p. 22).
- [Bou16] J. Bourgain. “A note on the Schrödinger maximal function”. In: *J. Anal. Math.* 130 (2016), pp. 393–396. arXiv: 1609.05744 [math.NT]. MR: 3574661 (cit. on p. 43).
- [Bou93] J. Bourgain. “Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I. Schrödinger equations”. In: *Geom. Funct. Anal.* 3.2 (1993), pp. 107–156. MR: 1209299 (cit. on p. 12).
- [CHS91] A. Carbery, E. Hernández, and F. Soria. “Estimates for the Keakeya maximal operator on radial functions in \mathbf{R}^n ”. In: *Harmonic analysis (Sendai, 1990)*. ICM-90 Satell. Conf. Proc. Springer, Tokyo, 1991, pp. 41–50. MR: 1261427 (cit. on p. 66).
- [CLL04] E. A. Carlen, E. H. Lieb, and M. Loss. “A sharp analog of Young’s inequality on S^N and related entropy inequalities”. In: *J. Geom. Anal.* 14.3 (2004), pp. 487–520. arXiv: math/0408030. MR: 2077162 (cit. on p. 46).
- [Cor77] A. Córdoba. “The Keakeya maximal function and the spherical summation multipliers”. In: *Amer. J. Math.* 99.1 (1977), pp. 1–22. MR: 0447949 (cit. on p. 63).
- [Cór77] A. Córdoba. “The multiplier problem for the polygon”. In: *Ann. of Math. (2)* 105.3 (1977), pp. 581–588. MR: 0438022 (cit. on pp. 63, 69).
- [Cór81] A. Córdoba. “Some remarks on the Littlewood-Paley theory”. In: *Rend. Circ. Mat. Palermo (2)* suppl. 1 (1981), pp. 75–80. MR: 639467 (cit. on pp. 30, 67).
- [Cór82] A. Córdoba. “Geometric Fourier analysis”. In: *Ann. Inst. Fourier (Grenoble)* 32.3 (1982), pp. vii, 215–226. MR: 688026 (cit. on pp. 30, 69).
- [Cór83] A. Córdoba. “Vector valued inequalities for multipliers”. In: *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*. Wadsworth Math. Ser. Wadsworth, Belmont, CA, 1983, pp. 295–305. MR: 730074 (cit. on pp. 30, 69).
- [Dem10] C. Demeter. “Singular integrals along N directions in \mathbb{R}^2 ”. In: *Proc. Amer. Math. Soc.* 138.12 (2010), pp. 4433–4442. arXiv: 1001.1998 [math.CA]. MR: 2680067 (cit. on p. 66).
- [DP19] F. Di Plinio and I. Parissis. “Directional square functions and a sharp Meyer lemma”. Preprint. Feb. 2019. arXiv: 1902.03644 [math.CA] (cit. on p. 70).
- [Du+18] X. Du, L. Guth, X. Li, and R. Zhang. “Pointwise convergence of Schrödinger solutions and multilinear refined Strichartz estimates”. In: *Forum Math. Sigma* 6 (2018), e14, 18. arXiv: 1803.01720 [math.CA]. MR: 3842310 (cit. on p. 40).
- [Du+19] X. Du, J. Kim, H. Wang, and R. Zhang. “Lower bounds for estimates of the Schrödinger maximal function”. Preprint. Feb. 2019. arXiv: 1902.01430 [math.CA] (cit. on pp. 43, 45).
- [DZ18] X. Du and R. Zhang. “Sharp L^2 estimate of Schrödinger maximal function in higher dimensions”. Preprint. May 2018. arXiv: 1805.02775 [math.CA] (cit. on pp. 40, 43).
- [Gut10] L. Guth. “The endpoint case of the Bennett-Carbery-Tao multilinear Keakeya conjecture”. In: *Acta Math.* 205.2 (2010), pp. 263–286. arXiv: 0811.2251 [math.CA]. MR: 2746348 (cit. on p. 6).
- [Gut15] L. Guth. “A short proof of the multilinear Keakeya inequality”. In: *Math. Proc. Cambridge Philos. Soc.* 158.1 (2015), pp. 147–153. arXiv: 1409.4683 [math.AP]. MR: 3300318 (cit. on p. 6).
- [GZa] S. Guo and R. Zhang. “On integer solutions of Parsell-Vinogradov systems”. In: *Invent. Math.* 218.1 (), pp. 1–81. arXiv: 1804.02488. MR: 3994585 (cit. on p. 55).
- [GZb] S. Guo and P. Zorin-Kranich. “Decoupling for moment manifolds associated to Arkhipov-Chubarikov-Karatsuba systems”. In: *Adv. Math.* 360 (). arXiv: 1811.02207 [math.NT]. MR: 4031117 (cit. on p. 55).

- [GZ19] S. Guo and P. Zorin-Kranich. *Decoupling for certain quadratic surfaces of low co-dimensions*. 2019. arXiv: 1902.03450 [math.CA] (cit. on pp. 10, 55).
- [HNS11] Y. Heo, F. Nazarov, and A. Seeger. “Radial Fourier multipliers in high dimensions”. In: *Acta Math.* 206.1 (2011), pp. 55–92. arXiv: 0902.0943 [math.CA]. MR: 2784663 (cit. on p. 30).
- [Kat99] N. H. Katz. “Maximal operators over arbitrary sets of directions”. In: *Duke Math. J.* 97.1 (1999), pp. 67–79. MR: 1681088 (cit. on pp. 64, 66, 69).
- [KPV91] C. E. Kenig, G. Ponce, and L. Vega. “Oscillatory integrals and regularity of dispersive equations”. In: *Indiana Univ. Math. J.* 40.1 (1991), pp. 33–69. MR: 1101221 (cit. on p. 43).
- [Lac07] M. T. Lacey. *Issues related to Rubio de Francia’s Littlewood-Paley inequality*. Vol. 2. New York Journal of Mathematics. NYJM Monographs. State University of New York, University at Albany, Albany, NY, 2007, p. 36. arXiv: math/0306417. MR: 2293255 (cit. on pp. 67, 68).
- [Lee06] S. Lee. “On pointwise convergence of the solutions to Schrödinger equations in \mathbb{R}^{2n} ”. In: *Int. Math. Res. Not.* (2006), Art. ID 32597, 21. arXiv: 0711.0717 [math.CA]. MR: 2264734 (cit. on p. 43).
- [Lee16] J. Lee. “A trilinear approach to square function and local smoothing estimates for the wave operator”. Preprint. July 2016. arXiv: 1607.08426 [math.CA] (cit. on pp. 30, 35, 37).
- [Li17] Z. K. Li. “Effective l^2 decoupling for the parabola”. Preprint. 2017. arXiv: 1711.01202 [math.CA] (cit. on p. 12).
- [Lie90] E. H. Lieb. “Gaussian kernels have only Gaussian maximizers”. In: *Invent. Math.* 102.1 (1990), pp. 179–208. MR: 1069246 (cit. on p. 46).
- [LR17a] R. Lucà and K. Rogers. “A note on pointwise convergence for the Schrödinger equation”. Preprint. Mar. 2017. arXiv: 1703.01360 [math.CA] (cit. on p. 43).
- [LR17b] R. Lucà and K. M. Rogers. “Coherence on fractals versus pointwise convergence for the Schrödinger equation”. In: *Comm. Math. Phys.* 351.1 (2017), pp. 341–359. MR: 3613507 (cit. on p. 43).
- [LV12] S. Lee and A. Vargas. “On the cone multiplier in \mathbb{R}^3 ”. In: *J. Funct. Anal.* 263.4 (2012), pp. 925–940. MR: 2927399 (cit. on pp. 30, 35, 36).
- [MSS92] G. Mockenhaupt, A. Seeger, and C. D. Sogge. “Wave front sets, local smoothing and Bourgain’s circular maximal theorem”. In: *Ann. of Math. (2)* 136.1 (1992), pp. 207–218. MR: 1173929 (cit. on pp. 30, 31).
- [Ram18] J. Ramos. “A trilinear restriction estimate with sharp dependence on transversality”. In: *Amer. J. Math.* 140.6 (2018), pp. 1443–1469. arXiv: 1601.05750 [math.CA]. MR: 3884636 (cit. on p. 8).
- [Rog08] K. M. Rogers. “A local smoothing estimate for the Schrödinger equation”. In: *Adv. Math.* 219.6 (2008), pp. 2105–2122. MR: 2456277 (cit. on p. 29).
- [Rub85] J. L. Rubio de Francia. “A Littlewood-Paley inequality for arbitrary intervals”. In: *Rev. Mat. Iberoamericana* 1.2 (1985), pp. 1–14. MR: 850681 (87j:42057) (cit. on pp. 67, 68).
- [Sog91] C. D. Sogge. “Propagation of singularities and maximal functions in the plane”. In: *Invent. Math.* 104.2 (1991), pp. 349–376. MR: 1098614 (cit. on p. 29).
- [SSS91] A. Seeger, C. D. Sogge, and E. M. Stein. “Regularity properties of Fourier integral operators”. In: *Ann. of Math. (2)* 134.2 (1991), pp. 231–251. MR: 1127475 (cit. on p. 27).
- [Str78] J.-O. Strömberg. “Maximal functions associated to rectangles with uniformly distributed directions”. In: *Ann. Math. (2)* 107.2 (1978), pp. 399–402. MR: 0481883 (cit. on pp. 63, 65, 66).
- [Val11] S. I. Valdimarsson. “Geometric Brascamp-Lieb has the optimal best constant”. In: *J. Geom. Anal.* 21.4 (2011), pp. 1036–1043. MR: 2836590 (cit. on p. 52).
- [Vau97] R. C. Vaughan. *The Hardy-Littlewood method*. Second. Vol. 125. Cambridge Tracts in Mathematics. Cambridge: Cambridge University Press, 1997, pp. xiv+232. MR: 1435742(98a:11133) (cit. on p. 63).
- [Wol00] T. Wolff. “Local smoothing type estimates on L^p for large p ”. In: *Geom. Funct. Anal.* 10.5 (2000), pp. 1237–1288. MR: 1800068 (cit. on p. 28).
- [Woo19] T. D. Wooley. “Nested efficient congruencing and relatives of Vinogradov’s mean value theorem”. In: *Proc. Lond. Math. Soc. (3)* 118.4 (2019), pp. 942–1016. arXiv: 1708.01220 [math.NT]. MR: 3938716 (cit. on p. 55).

[Zor18] P. Zorin-Kranich. “Kekeya-Brascamp-Lieb inequalities”. In: *Collect. Math.* (2018). To appear. arXiv: 1807.09604 [math.CA] (cit. on p. 6).