LINEAR AND BILINEAR RESTRICTION ESTIMATES FOR THE FOURIER TRANSFORM

BY

FARUK TEMUR

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Doctoral Committee:

Professor Richard S. Laugesen, Chair
Professor Joseph Rosenblatt, Contingent Chair
Associate Professor M. Burak Erdoğan, Director of Research
Associate Professor Xiaochun Li

Abstract

This thesis is concerned with the restriction theory of the Fourier transform. We prove two restriction estimates for the Fourier transform. The first is a bilinear estimate for the light cone when the exponents are on a critical line. This extends results proven by Wolff, Tao and Lee-Vargas. The second result is a linear restriction estimate for surfaces with positive Gaussian curvature that improves over estimates proven by Bourgain and Guth, and gives the best known exponents for the well-known restriction conjecture for dimensions that are multiples of three.

To İ. Yorulmaz and M. Yılanlı

Acknowledgments

I would like to express my gratitude to my supervisor M. Burak Erdoğan for introducing me to harmonic analysis, and for many discussions that helped develop my understanding of the area. I would thank for financial support he extended through NSF grant DMS-0900865 that provided partial support while conducting the research that resulted in this thesis. I would also thank to him, and to Xiaochun Li and Nikolaos Tzirakis for their help with my applications for positions after graduation.

I would thank to my family for their emotional support, and to two of my friends, İ. Yorulmaz and M. Yılanlı, to whom this work is dedicated.

I am grateful to members of my thesis committee, R. Laugesen, J. Rosenblatt, and Xiaochun Li for their careful reading of my thesis.

Table of Contents

Chapter 1

Introduction

1.1 Overview

The Fourier transform of $f \in L^1(\mathbb{R}^n)$ is defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}^n} f(x) e^{-2\pi i x \cdot \xi} dx.$$

The Plancherel theorem, which is valid a priori for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$,

$$\|f\|_2 = \|\widehat{f}\|_2,$$

allows us to extend the Fourier transform to $L^2(\mathbb{R}^n)$ functions. Since $f \in L^p(\mathbb{R}^n),\ 1 < p < 2$, can be decomposed into two functions $f_1 \in L^1(\mathbb{R}^n),\ f_2 \in L^2(\mathbb{R}^n)$ the Fourier transform can also be defined for such functions.

We note that if $f \in L^1(\mathbb{R}^n)$, then the Fourier transform $\widehat{f}$ is continuous, so can be meaningfully restricted to any measure zero set, while for $p = 2$, $\widehat{f}$ can be an arbitrary function in $L^2(\mathbb{R}^n)$, and thus is not well defined on measure zero sets. The restriction problem is the question of determining $1 \leq p < 2$ for which the Fourier transform $\widehat{f}$ of any $L^p(\mathbb{R}^n)$ function f, can be restricted to hypersurfaces. It is easy to see that restriction to flat surfaces is possible only for $p = 1$. For example, for a compactly supported non-negative function ϕ defined on $\mathbb{R}^{n-1}$, the Fourier transform of

$$f(x) := \frac{\phi(x_2, \dots, x_n)}{1 + |x_1|}$$

is infinite on every point of the hyperplane $\{\xi \in \mathbb{R}^n : \xi_1 = 0\}$. However it was observed by Stein in 1967 that for curved hypersurfaces restriction for some $p > 1$ is possible, and determining the exact range of p depending on the curved surface is of interest.

Modern formulation of the restriction problem is as follows: let $S \in \mathbb{R}^n$ be a smooth compact hypersurface, and let σ be a measure on this surface with essentially bounded Radon-Nikodym derivative with respect to the surface measure. If

$$\|\widehat{f}\|_{L^q(S,d\sigma)} \leq C_{p,q,S,\sigma} \|f\|_{L^p(\mathbb{R}^n)} \tag{1.1}$$

holds for all functions of the Schwartz class $\mathcal{S}(\mathbb{R}^n)$, then we say that the linear restriction inequality $R_{S,d\sigma}(p \to q)$ holds. The following dual formulation is often used to study the problem. Let $f \in L^1(d\sigma)$ be a function on S, and define the Fourier transform $\widehat{fd\sigma}$ by

$$\widehat{fd\sigma}(\xi) := \int_S f(x) e^{-2\pi i x \cdot \xi} d\sigma(x).$$

If the inequality

$$\|\widehat{fd\sigma}\|_{L^q(\mathbb{R}^n)} \leq C_{p,q,S,\sigma} \|f\|_{L^p(S,d\sigma)} \tag{1.2}$$

holds for all smooth functions on S, then we say that the adjoint linear restriction inequality $R^*_{S,d\sigma}(p \to q)$ holds. This problem is related to some other well-known conjectures in harmonic analysis and partial differential equations, such as the Bochner-Riesz conjecture, the Kakeya conjecture, and the local smoothing conjecture. It has been studied particularly intensively for compact subsets of the paraboloid

$$\{(x_1, \dots, x_n) \in \mathbb{R}^n : x_n = x_1^2 + \dots + x_{n-1}^2\}, \tag{1.3}$$

the sphere

$$\{x \in \mathbb{R}^n : |x| = 1\}, \tag{1.4}$$

and the light cone

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n : x_n = (x_1^2 + \ldots + x_{n-1}^2)^{1/2}\}. \tag{1.5}$$

The first two of these are representatives of elliptic surfaces, that is surfaces of positive Gaussian curvature, and most of the time what is true for them can be generalized to elliptic surfaces. The last is an example of a surface with one vanishing principal curvature. The light cone, of course, is flat for $n = 2$, so it is considered only for $\mathbb{R}^n,\ n \geq 3$.

There are two basic counterexamples that put limitations on the range of exponents for which $R^*_{S,d\sigma}(p \to q)$ holds. The first comes from the decay rate at infinity of the Fourier transform of surface measures. For surface measures of elliptic surfaces this decay is like $(1 + |x|)^{\frac{1-n}{2}}$, while for the light cone it is like $(1 + |x|)^{\frac{2-n}{2}}$. Therefore taking $f \equiv 1$ on the surface, we see that $q > \frac{2n}{n-1}$ is needed to make $|\widehat{fd\sigma}|^q$ integrable for the case of elliptic surfaces. For the light cone the range is $q > \frac{2n}{n-2}$.

The second counterexample is called Knapp example. We describe this for elliptic surfaces, for the light cone it is only slightly different. We take a cap c_ϵ of radius comparable to ϵ on the surface and a function $f \equiv 1$ on this cap. The ellipticity of the surface implies that c_ϵ can be enclosed in a disk of radius comparable to ϵ and thickness ϵ^2. Since c_ϵ has a surface area comparable to ϵ^{n-1}, the magnitude of $\widehat{fd\sigma}$ is comparable to ϵ^{n-1} on the tube dual to the disk, that is, a tube centered at the origin with radius ϵ^{-1} and length ϵ^{-2} oriented in the direction of the unit normal of the surface at a point belonging to the cap. Then the left hand side of (1.2) satisfies $\gtrsim \epsilon^{-\frac{n+1}{q}} \cdot \epsilon^{n-1}$, while the right hand side is comparable to $\epsilon^{\frac{n-1}{p}}$. This gives the necessary condition

$$\frac{n+1}{q} \leq (n-1)(1 - \frac{1}{p}). \tag{1.6}$$

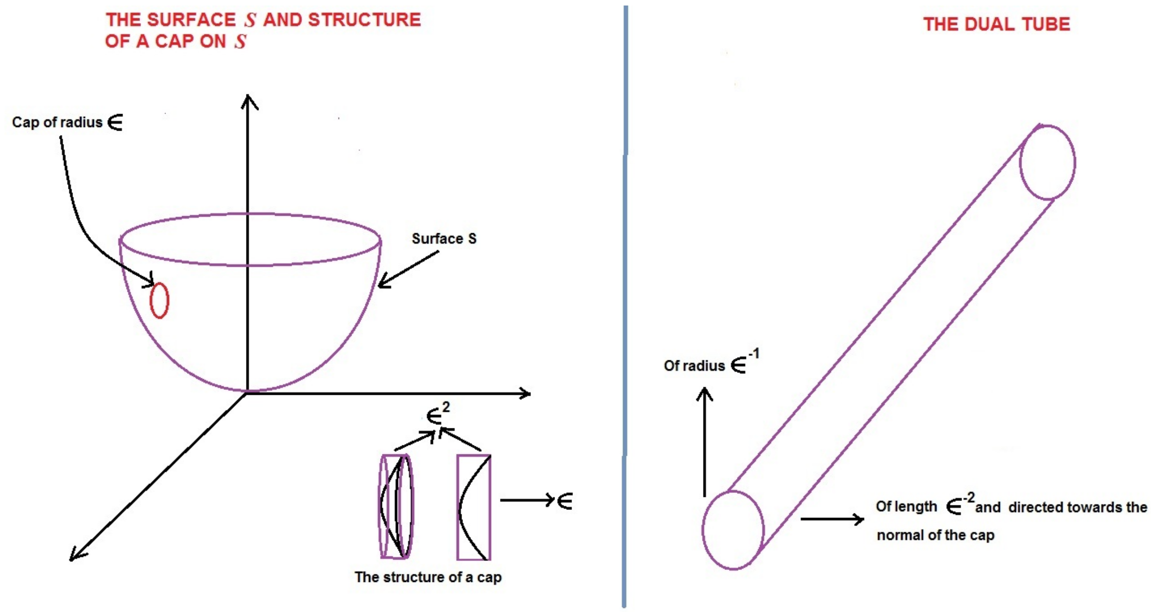

Figure 1.1: The Knapp example for the paraboloid

For the light cone we lengthen the cap in the null direction so that it has length comparable to 1. Then this will be enclosed in a box with dimensions comparable to

$$\underbrace{\epsilon \times \ldots \times \epsilon}_{n-2 \text{ times}} \times \epsilon^2 \times 1$$

with length 1 in the null direction, ϵ^2 in the normal direction, and ϵ in all other directions. The box dual to this has dimensions comparable to

$$\underbrace{\epsilon^{-1} \times \ldots \times \epsilon^{-1}}_{n-2 \text{ times}} \times \epsilon^{-2} \times 1.$$

Therefore, a similar calculation gives this time

$$\frac{n}{q} \leq (n-2)(1-\frac{1}{p}). \tag{1.7}$$

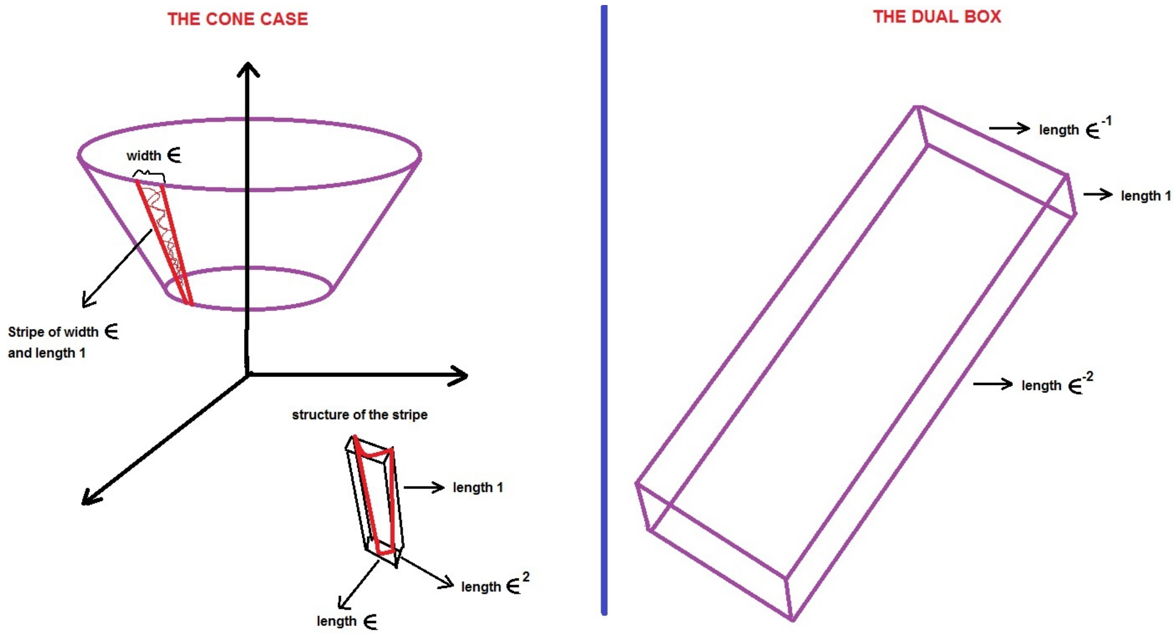

Figure 1.2: The Knapp example for the light cone

The restriction problem $R^*_{S,d\sigma}(p \to q)$ is particularly important for two values of p. In the case $p = 2$, the inequalities $R^*_{S,d\sigma}(p \to q)$ are known as the Strichartz estimates and they have many applications in partial differential equations. The case $p = \infty$ is known to imply all other cases for S a sphere, and also serves as a good measure of progress for other surfaces; see [30]. We will also concentrate on these two cases. For $p = 2$, the necessary conditions stated above give $q \geq \frac{2n+2}{n-1}$ for elliptic surfaces, and $q \geq \frac{2n}{n-2}$ for the light cone. That these conditions are also sufficient follows from the work of Tomas and Stein; see [32]. When $p = \infty$, the necessary conditions coming from the Knapp examples are vacuous, and the problem is known as the restriction conjecture. Notice that since for a set of finite measure $\|\cdot\|_2$ norm of a function f can be dominated by its $\|\cdot\|_\infty$ norm, the resolution of $p = 2$ case already implies a partial result for the case $p = \infty$. This is far from the conjectured range of q, and as opposed to the $p = 2$ case this conjecture is still wide open for elliptic surfaces in dimensions three and higher, and for the light cone in dimensions five and higher. However apart from these already resolved cases there has been significant progress, on which we will further elaborate in section 1.2.

Before ending this section we will talk about bilinear and multilinear restriction estimates that were utilized, apart from other applications, to obtain linear restriction estimates. We first define the concept of transversality of hypersurfaces that will be needed to describe these estimates. Let S_i, $1 \leq i \leq m$ be hypersurfaces in $\mathbb{R}^n$ for $m \leq n$. Let $\xi_i \in S_i$ be arbitrary points on these subsets, and ξ_i' be unit normal vectors at these points. If we have

$$|\xi_1' \wedge \ldots \wedge \xi_m'| > c$$

for a positive constant c regardless of choice of the points $\xi_i \in S_i$ we will call the surfaces S_i transversal.

Consider an estimate of the following type for two compact subsets S_1, S_2 of non-empty interior of S, with S either the light cone or an elliptic surface:

$$\|\widehat{f_1 d\sigma_1} \widehat{f_2 d\sigma_2}\|_{L^q(\mathbb{R}^n)} \leq C_{p,q,S_i,\sigma_i} \|f_1\|_{L^2(S_1,d\sigma_1)} \|f_2\|_{L^2(S_2,d\sigma_2)}. \tag{1.8}$$

Here σ_1, σ_2 are the surface measures of S_1, S_2. This type of restriction estimates are called bilinear estimates. If we choose S_1, S_2 to be the same, this inequality is the same as (1.2), and the inequality holds if and only if $q \geq \frac{n+1}{n-1}$. However, if these subsets are transversal, we realize that the Knapp examples constructed above gives exponents lower than $\frac{n+1}{n-1}$, and this raises the possibility of actually proving this type of estimates for q lower than $\frac{n+1}{n-1}$. Such an estimate was first observed by Carleson and Sjölin, who proved that (1.8) holds for $q \geq 2$ regardless of the dimension n if the surfaces S_1, S_2 are transversal, see [7]. Further, for such surfaces S_1, S_2 Klainerman and Machedon conjectured that $q \geq \frac{n+2}{n}$ is necessary and sufficient. The condition on q comes from a modification of the Knapp examples, that we now describe for elliptic surfaces.

If we take just one cap from each surface of size ϵ and take $f_i \equiv 1$, $i = 1, 2$, the functions $|\widehat{f_1 d\sigma}|, |\widehat{f_2 d\sigma}|$ are comparable to ϵ^{n-1} on corresponding dual tubes. Therefore, the place where we can guarantee that their product is comparable to ϵ^{2n-2} is the intersection of these tubes. Since the tubes are transversal the intersection has measure comparable to only ϵ^{-n}. One can further improve this example by taking more than one cap from each surface. After taking one cap from each side, take about ϵ^{-1} further caps from each side with normals lying in the plane defined by the normals of the two caps we first took. Take f_i to be sum of nonnegative bump functions supported on these caps. The expressions $\widehat{f_i d\sigma_i}$ will be comparable to ϵ^{n-1} on two ends of bushes comprised by dual tubes. If we modulate f_2 so that one end of each bush intersect, $|\widehat{f_1 d\sigma}\widehat{f_2 d\sigma}|$ will be comparable to ϵ^{2n-2} on a box shaped region of dimensions

$$\epsilon^{-2} \times \epsilon^{-2} \times \epsilon^{-1} \times \ldots \times \epsilon^{-1},$$

and this gives the condition $q \geq \frac{n+2}{n}$ as conjectured by Klainerman and Machedon. For the light cone the example is simpler. The caps will be aligned in the null direction, and by appropriately modulating the bump functions dual tubes from both surfaces will comprise a box of size comparable to the one for the elliptic case. We thus get exactly the same exponent.

Bilinear restriction estimates suggest that higher levels of transversality might lead to even better exponents, and encourage the study of multilinear restriction estimates. In 2006 Bennett, Carbery and Tao, [2], proved the following result in this direction. Fix $2 \leq m \leq n$. Let $S_i \in \mathbb{R}^n,\ 1 \leq i \leq m$ be smooth, compact hypersurfaces whose parametrizations $\Sigma_i : U_i \to S_i$ from subsets U_i of $\mathbb{R}^{n-1}$ satisfy the smoothness condition

$$\|\Sigma_i\|_{C^2(U_i)} \leq A.$$

Also assume that they are transverse, that is, for any points, $x_i \in S_i$, unit normals x_i' satisfy

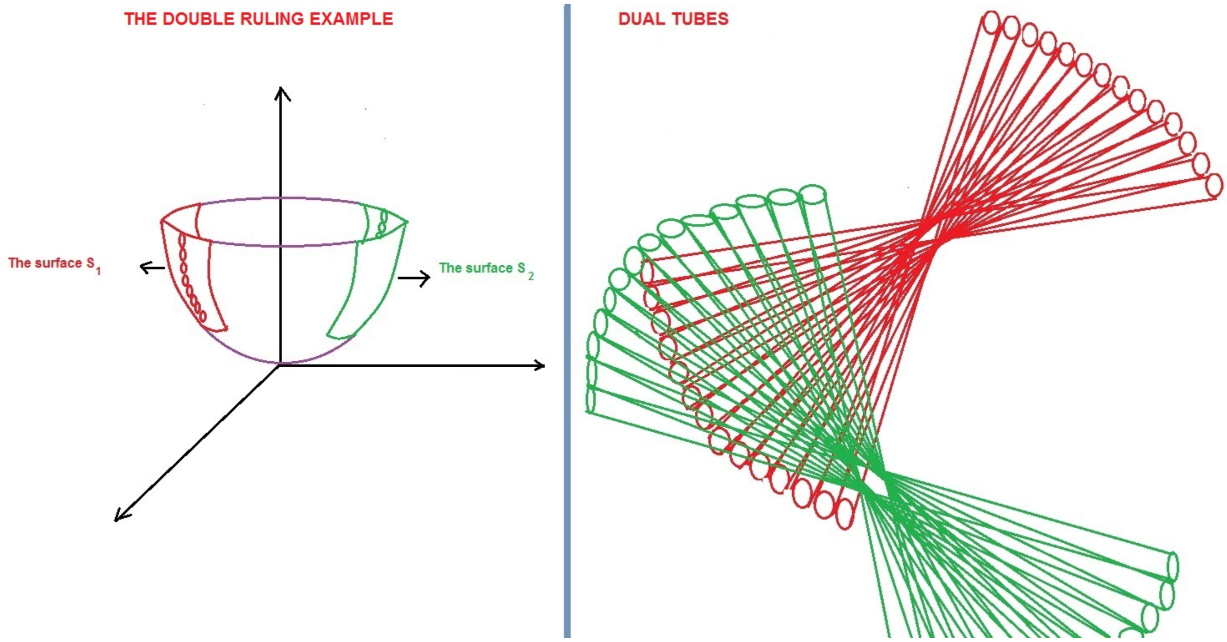

Figure 1.3: The double ruling example for the paraboloid

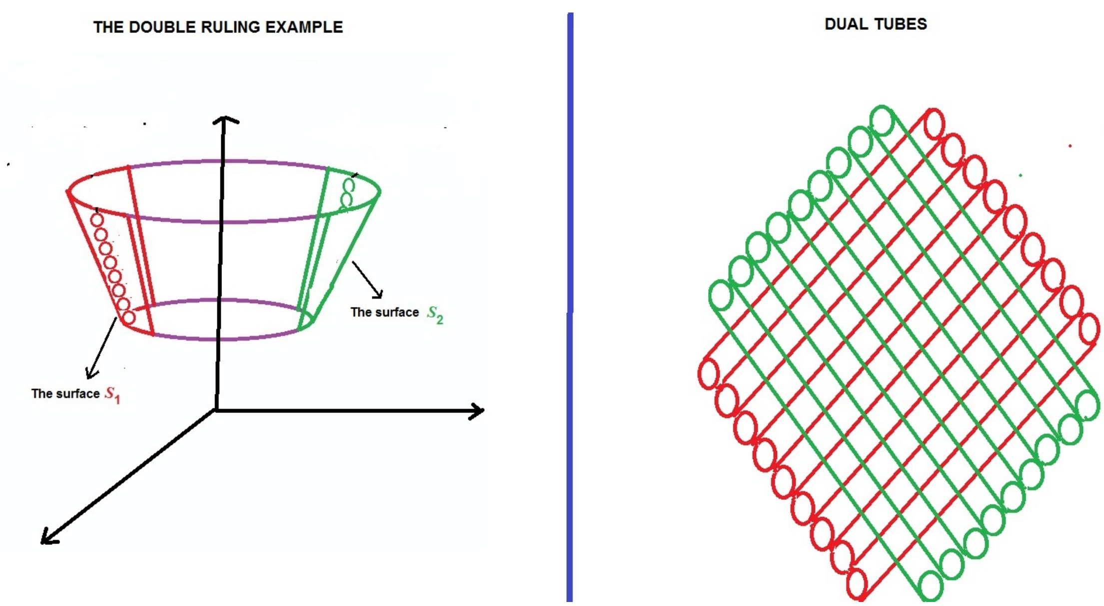

Figure 1.4: The double ruling example for the light cone

$|x_1' \wedge \ldots \wedge x_m'| \geq B$. We then have

$$\| \prod_{i=1}^{m} \widehat{f_i d\sigma_i} \|_{L^q(\mathbb{R}^n)} \leq C_{q,S_i,\sigma_i} \prod_{i=1}^{m} \|f_i\|_{L^2(S_i)} \tag{1.9}$$

for $q > \frac{2}{m-1}$. We remark that we imposed no curvature condition, and without such an assumption this is sharp except for the endpoint. Below we describe the counterexample.

Let S_i be a compact rectangular subset of a hyperplane with unit normal

$$(0, \ldots, \underbrace{1}_{ith\ place}, \ldots, 0), \quad 1 \leq i \leq m$$

with ϵ sidelength in the first m directions except ith, and 1 on the last $n - m$ directions. Thus if we consider $\mathbb{R}^n = \mathbb{R}^m \times \mathbb{R}^{n-m}$, in directions normal to $\mathbb{R}^m$ it has length 1. Let $f_i \equiv 1$ on S_i. Then $| \prod_{i=1}^{m} \widehat{f_i d\sigma_i}|$ is comparable to $\epsilon^{m \cdot (m-1)}$ on a box with sidelength comparable to ϵ^{-1} in directions lying in $\mathbb{R}^m$, and to 1 for those normal to $\mathbb{R}^m$. So the left hand side of (1.9) is comparable to $\epsilon^{m^2 - m - m/q}$ and the right is to $\epsilon^{(m^2-m)/2}$. This implies that the theorem above is optimal except the endpoint.

1.2 History of the Progress

In this section we review the history of progress on problems introduced in the previous section. We start with the restriction conjecture $R^*_{S,d\sigma}(\infty \to q), \quad q \geq \frac{2n}{n-1}$, for S the sphere and the paraboloid. Many of these results were proved in the more general context of the elliptic surfaces or can be extended in a straightforward way to them. In two dimensions the conjecture asserts that $q > 4$, and this was settled by Fefferman and Stein in 1970; see [10]. In dimension three, the conjectured range is $q > 3$, and as we mentioned the Tomas-Stein theorem implies the conjecture for $q \geq 4$. Then in 1991, Bourgain lowered this to $q > 4 - 2/15$; see [3]. This is particularly important, as 4 is a very special exponent for which L^2 based methods are applicable. Wolff, in [37], further improved this to $4 - 2/11$.

Tao, Vargas and Vega, in [25], introduced the use of bilinear estimates of the type (1.8), and proved that $q > 4 - 2/9$ suffices. Tao and Vargas further lowered this to $4 - 2/7$; see [26]. In [29], Tao proved $q > 10/3$, this represents the best range one can obtain using L^2 based bilinear estimates. Finally, using multilinear estimates, Bourgain and Guth, [6], obtained the range $q > 33/10$. In higher dimensions the Tomas-Stein theorem gives $q \geq (2n+2)/(n-1)$. Bourgain, [3], showed that $q \geq (2n+2)/(n-1) - \epsilon_n$ is possible. Wolff used his improved Kakeya estimates to lower this to $(2n^2+n+6)/(n^2+n-1)$; see [37]. In [29] Tao improved this to $(2n+2)/n$. Finally, in [6], Bourgain and Guth improved this to

$$\begin{aligned} q &> \frac{8n+6}{4n-3} \quad \text{if} \quad n \equiv 0 \mod 3 \\ q &> \frac{2n+1}{n-1} \quad \text{if} \quad n \equiv 1 \mod 3 \\ q &> \frac{4n+4}{2n-1} \quad \text{if} \quad n \equiv 2 \mod 3. \end{aligned} \tag{1.10}$$

For the cone case the conjecture is, of course, meaningful for $n \geq 3$, as for $n = 2$ the cone is flat. First progress in this case is the well known work, [24], of Strichartz, which gives $q \geq 2n/(n-2)$. For $n = 3$ the conjecture was settled fully by Barcelo in [1], and for $n = 4$ by Wolff in [34]. In the same work Wolff also showed $q > \frac{2n+4}{n}$ for $n \geq 5$. This represents the best exponent up to date.

We now review the progress on the Klainerman-Machedon conjecture on the L^2 based bilinear restriction estimates. This problem, as opposed to the linear restriction conjecture, is now fully resolved. As we have mentioned the first theorem of the type (1.8) is the Carleson-Sjölin theorem, [7], which proves that $q \geq 2$ suffices for any dimension $n \geq 2$. This holds without any curvature assumption, as in the Bennett-Carbery-Tao result. Also it is sharp for $n = 2$. For $n \geq 3$, of course, we have a better range for elliptic hypersurfaces and the light cone. For the elliptic surfaces $q \geq (n+1)/(n-1)$ follows from the work of Tomas and Stein. For $n > 3$ this improves upon the Carleson-Sjölin work, but for $n = 3$ gives the same exponent. In this case Tao, Vargas and Vega, [25], proved that $q \geq 2 - 5/69$. Later Tao and Vargas, [26], further reduced this to $q \geq 2-2/17$. In 2003, Tao proved the conjecture in all dimensions, except for the endpoint; see [29]. The endpoint is recently proved by Jungjin Lee in [14].

For the light cone the exponent given by Strichartz's work, [24], is $q \geq n/(n-2)$. For $n=3$, Bourgain, [5], improved this to $q \geq 2-13/2048$. Tao and Vargas, [26], further lowered this to $2-8/121$. Finally Wolff proved the conjecture without the endpoint in [34]. In this work he invented the methods that enabled progress also in the elliptic case. The endpoint was proved by Tao in [28].

1.3 Statement of Results

In this section we will introduce the main results of this thesis. Our first result is an estimate of the type (1.8). But to state it we need some further notation.

We define the mixed Lebesgue norms as follows

$$\|f\|_{L^rL^q} := \Big(\int \big(\int |f(x_1,\ldots,x_n)|^q dx_1 \ldots dx_{n-1}\big)^{\frac{r}{q}} dx_n \Big)^{\frac{1}{r}}.$$

Inequalities concerning these norms are useful in the study of dispersive PDE. This is because of the fact that the Fourier transforms of solutions of the Schrödinger equation and the wave equation are distributions supported respectively on the paraboloid and the light cone. This close relation between the restriction estimates and PDE provides incentive for extending these estimates to the mixed Lebesgue norms, treating time and space variables separately. Our first estimate is a result in this direction.

Theorem 1 *Let S_1, S_2 be compact transverse subsets of the light cone, and let σ_1, σ_2 be measures on these surfaces that has essentially bounded Radon-Nikodym derivatives with respect to the canonical measures of S_1, S_2. Let $f_i \in L^2(S_i, d\sigma_i)$, $i = 1, 2$. Then for $1 \leq$*

$q, r \le 2$ with $\frac{1}{r} \le \frac{n}{2}(1-\frac{1}{p})$ and $\frac{1}{q} \le \min(1, \frac{n}{4})$ we have

$$\|\widehat{f_1 d\sigma_1}\widehat{f_2 d\sigma_2}\|_{L^r L^q} \le C_{q,r,S_i,\sigma_i}\|f_1\|_{L^2(S_1,d\sigma_1)}\|f_2\|_{L^2(S_2,d\sigma_2)}. \tag{1.11}$$

Two counterexamples that give conditions $\frac{1}{r} \le \frac{n}{2}(1-\frac{1}{p})$, $\frac{1}{q} \le \min(1, \frac{n}{4})$ will be described in the next chapter. The work of Lee and Vargas, proved this for $\frac{1}{r} < \frac{n}{2}(1-\frac{1}{p})$. Thus our result extends this to the endline.

Our second result is on the restriction conjecture.

Theorem 2 *Let $k \ge 2$ be an integer, and let $n = 3k$. Let $S \subset \mathbb{R}^n$ be a compact subset of an elliptic surface, and $f \in L^\infty(S, d\sigma)$, where σ is a measure on the surface with essentially bounded Radon-Nikodym derivative with respect to the canonical measure of the surface. Then we have*

$$\|\widehat{fd\sigma}\|_{L^q} \le C_{q,S,\sigma}\|f\|_\infty$$

for

$$\begin{aligned} &q > \frac{18}{7} - \frac{2}{735} && \text{if} \quad k = 2 \\ &q > \frac{8k+2}{4k-1} - \frac{(5k-4)(2k+1)}{(4k^2-k)(28k+7)(4k^2-k)} && \text{if} \quad k > 2. \end{aligned}$$

Here observe that minus terms represent improvements over Bourgain-Guth exponents.

Chapter 2

Endline bilinear restriction estimates for the lightcone

This chapter will be dedicated to the proof of Theorem 1.

2.1 Counterexamples and Simplifications

This first section of the proof will describe the counterexamples that show sharpness of the result. We will also reformulate and simplify the result, introduce some notation that we will use in the rest of the proof, and give an overview of the proof.

We start with the counterexamples. The condition $\frac{1}{r} \leq \frac{n}{2}(1-\frac{1}{p})$ comes actually from the modification of the Knapp example described in the first section of the first chapter. If we calculate the left hand side of (1.11) with this example we will obtain this condition. The condition $\frac{1}{q} \leq \min(1, \frac{n}{4})$, on the other hand, needs a new construction. First notice that since we already assume $q \geq 1$, this condition makes sense only for $n = 3$, and it thus suffice to describe it in that case. Let S_1 be a strip of size $\epsilon \times 1$ with length 1 being in the null direction, on the light cone. Consider the intersection of the light cone with a slab of thickness ϵ^2 orthogonal to a normal of S_1 that is at an appropriate distance from the origin. We define S_2 to be the part of the intersection that lies right in front of S_1 for an angular width of $\pi/8$. Let $f_i \equiv 1$ on S_i, $i = 1, 2$. We see that $|\widehat{f_1 d\sigma_1}|$ is comparable to ϵ on the box dual to S_1 centered at the origin , and we have $|\widehat{f_2 d\sigma_2}|$ comparable to ϵ^2 in a box that lies inside that dual, with dimensions comparable to $1 \times 1 \times \epsilon^{-2}$. Thus the left hand side of the inequality is comparable to $\epsilon^{3-2/r}$ while the right is to $\epsilon^{3/2}$. Comparing these gives the desired result.

We now reformulate and simplify Theorem 1. Let S_1, S_2 be given as in the theorem and let U_1, U_2 be the projections of these onto $\mathbb{R}^{n-1} \times \{0\}$ hyperplane. Let f_i be square integrable functions defined respectively on U_i. Define

$$T_{U_i} f_i(\xi) := \int_{U_1} f_i(x) e^{-2\pi i(x \cdot \bar{\xi} + |x| \cdot \xi_n)} dx$$

for $i = 1, 2$ and $\xi = (\bar{\xi}, \xi_n) \in \mathbb{R}^n$. Then proving the inequality

$$\|T_{U_1} f_1 T_{U_2} f_2\|_{L^q L^r} \leq C_{q,r,U_1,U_2} \|f_1\|_{L^2(U_1)} \|f_2\|_{L^2(U_2)} \tag{2.1}$$

for (q, r) satisfying the conditions of Theorem 1, will clearly imply that theorem. To simplify the geometric picture we will take

$$U_1 := \{x \in \mathbb{R}^{n-1} : \angle(x, (1, 0, \dots, 0)) \leq \pi/8,\ 1 \leq x_{n-1} \leq 2\},$$
$$U_2 := \{x \in \mathbb{R}^{n-1} : \angle(x, (-1, 0, \dots, 0)) \leq \pi/8,\ 1 \leq x_{n-1} \leq 2\}.$$

The same proof, with different constants, works for other domains as well. To prove our estimate we will need the following larger domains:

$$U_1' := \{x \in \mathbb{R}^{n-1} : \angle(x, (1, 0, \dots, 0)) \leq \pi/4,\ 1/2 \leq x_{n-1} \leq 4\},$$
$$U_2' := \{x \in \mathbb{R}^{n-1} : \angle(x, (-1, 0, \dots, 0)) \leq \pi/4,\ 1/2 \leq x_{n-1} \leq 4\}.$$

We introduce the modified versions of T_{U_i} as well. Let a_i be smooth, real-valued functions with supports contained in U_i' which are identically 1 on U_i. Then

$$T'_{U_i'} f_i(\xi) := \int a_i(x) f_i(x) e^{-2\pi i(x \cdot \bar{\xi} + |x| \cdot \xi_n)} dx.$$

Notice that if f_i is supported inside U_i, then $T'_{U_i'}$ is no different than T_{U_i}. Henceforth we will use T_i to denote $T'_{U'_i}$.

We can also simplify the conditions on exponents. Since for the non-endline cases this result was proved in [19], proving it for the endline is enough. To further simplify, we will introduce a property of the operators T_i, that is called conservation of mass in PDE literature. We have the following lemma

Lemma 1 *Let $U \in \mathbb{R}^{n-1}$, and let $f \in L^2(U)$. Let T be an operator defined by*

$$Tf(\xi) = \int_U f(x)e^{-2\pi i(x\cdot\overline{\xi}+\phi(x)\xi_n)}dx.$$

for a smooth function ϕ on U. Then we have

$$\int |Tf(\overline{\xi},\xi_n)|^2 d\overline{\xi} = \|f\|_2^2$$

Proof. The proof uses the Plancherel theorem. We have

$$\begin{aligned}\int |Tf(\overline{\xi},\xi_n)|^2 d\overline{\xi} &= \int \Big| \int_U f(x)e^{-2\pi i(x\cdot\overline{\xi}+\phi(x)\cdot\xi_n)}dx\Big|^2 d\overline{\xi} \\ &= \int \Big| \int_U \big(f(x)e^{-2\pi i\phi(x)\cdot\xi_n}\big)e^{-2\pi i x\cdot\overline{\xi}}dx\Big|^2 d\overline{\xi}.\end{aligned} \tag{2.2}$$

This last integral is actually the Fourier transform of the function $f(x)e^{-2\pi i\phi(x)\cdot\xi_n}$. Thus using Plancherel's theorem we have

$$= \|f_(x)\widehat{e^{-2\pi i\phi(x)\cdot\xi_n}}\|_2^2 = \|f(x)e^{-2\pi i\phi(x)\cdot\xi_n}\|_2^2 = \|f\|_2^2.$$

■

We shall utilize this lemma together with the following simple application of Cauchy-

Schwarz inequality

$$
\begin{aligned}
\|T_1f_1T_2f_2\|_{L^\infty L^1} &= \sup_{\xi_n}\int |T_1f_1(\overline{\xi},\xi_n)T_2f_2(\overline{\xi},\xi_n)|d\overline{\xi} \\
&\le \sup_{\xi_n}\Big(\int |T_1f_1(\overline{\xi},\xi_n)|^2 d\overline{\xi}\Big)^{1/2}\Big(\int |T_2f_2(\overline{\xi},\xi_n)|^2 d\overline{\xi}\Big)^{1/2}.
\end{aligned}
\tag{2.3}
$$

Thus one concludes

$$
\|T_1f_1T_2f_2\|_{L^\infty L^1} \le \|f_1\|_2\|f_2\|_2. \tag{2.4}
$$

If Theorem 1 is correct for a pair of (q,r) with $\frac{1}{q}<\min(1,\frac{n}{4})$, $\frac{1}{q}=\frac{n}{2}(1-\frac{1}{r})$, interpolating with what we have above gives Theorem 1 for all points on the line between $(1,\infty)$ and (q,r). Hence it is enough to prove our theorem for (q,r) with q arbitrarily close to $\min(1,\frac{n}{4})$. Thus for each $n\ge 3$ we fix

$$
0<\epsilon<\frac{1}{100n}, \tag{2.5}
$$

and let $1\le q,r\le\infty$ be such that

$$
\frac{1}{q}=\min(1,\frac{n}{4})-\epsilon, \qquad \frac{1}{q}=\frac{n}{2}(1-\frac{1}{r}). \tag{2.6}
$$

Our requirement on ϵ will be of use below.

Finally we also note that it is enough to prove our theorem for $\|f_1\|_2=\|f_2\|_2=1$. We will use this especially in Section 2.4.

Before proceeding further with simplifications we introduce some fixed constants and notation that will be utilized. We will use N to denote the integer $2^{n^{10}}$, and C_0 for $2^{\lfloor N/\epsilon\rfloor^{100}}$. Thus $N,1/\epsilon$ are much smaller than C_0. We also introduce a still larger constant $C_1:=2^{{C_0}^{10}}$. As before C will denote various constants, and dependencies of it will be written as subscripts when necessary. The size of C will in general be rather small such as $10n$, and it will never exceed $2^{N^{15}}$. As N,ϵ depend only on the dimension and (q,r), when C depends on these we will not keep writing this dependency. The choice of C,C_0,C_1 is so that $C<C_0<C_1$, and

each of these is much larger than the constants smaller than it. We introduce the notation $A \lesssim B$ to mean $A \leq CB$ for C as described above, and $A \approx B$ for $A \lesssim B$ and $B \lesssim A$.

Standard method of exploiting transversality for the light cone is to divide functions $f_i,\ i = 1, 2$ into strips so that their Fourier transforms will concentrate on transversal tubes. But this decomposition has to be done with smooth cutoffs which enlarge support of the function concerned. In our case this may take support of f_i beyond U_i. We thus have introduced U_i', and will introduce a quantity that we will call *margin* for functions f_i supported in U_i':

$$margin(f_i) := \mathrm{dist}(supp(f_i), \partial U_i). \tag{2.7}$$

Thus margin is the distance of the support of the function from the boundary of the region U_i. We also introduce the related quantity of angular dispersion of a function f_i as the diameter of the set

$$\{x/|x| : x \in supp(f_i)\}. \tag{2.8}$$

With these quantities at hand, we introduce the following local form of what we want to prove. We note that localization is a standard method of proving restriction estimates, see, for example, [28, 29, 34, 6].

Definition 1 *Let $R \geq C_0 2^{C_1/2}$, and let Q be a cube of sidelength R. We define $A(R)$ to be the best constant for which*

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q)} \leq A(R) \|f_1\|_2 \|f_2\|_2 \tag{2.9}$$

holds for all $f_i \in L^2(U_i)$ with margin requirement

$$margin(f_i) \geq 1/C_0 - (1/R)^{1/N}. \tag{2.10}$$

That $A(R)$ is finite is clear: we have $\|T_i f_i\|_\infty \lesssim \|f_i\|_2$ with constant depending only on the volume of U_i', and volume of Q is R^n, thus

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q)} \leq C R^{\frac{n-1}{r}+\frac{1}{q}} \|f_1\|_2 \|f_2\|_2.$$

So we have $A(R) \leq C R^{\frac{n-1}{r}+\frac{1}{q}}$. Further we have the following property: if $R/C \leq R' \leq R$, then $A(R) \lesssim A(R')$. This holds since any function satisfying the margin requirement for scale R also satisfies it for scale R', thus dividing a cube of size R into cubes of size at most R' we obtain this inequality. Similarly if $R \leq R' \leq CR$ we have $A(R) \lesssim A(R')$. This follows since for any function on U_i, by decomposing U_i, and using Lorentz transforms, dilations and rotations we obtain a constant number C of functions satisfying appropriate margin requirements, and these operations can enlarge the domain of integration by at most another constant factor C. Thus we also see that if we can prove

$$A(R) \lesssim 2^{CC_1}, \tag{2.11}$$

then we have

$$\|T_{U_1} f_1 T_{U_2} f_2\|_{L^q L^r(Q)} \leq C 2^{CC_1} \|f_1\|_2 \|f_2\|_2$$

for any cube Q of any radius, and any choice of functions $f_i \in L^2(U_i)$. This clearly, from the monotone convergence theorem, implies Theorem 1.

The reason we define $A(R)$ for not all functions in $f_i \in L^2(U_i)$, but for those satisfying a margin condition is that, as localization will enlarge the support of f_i, the localized form of f_i may have support exceeding U_i. Then we would not be able to use $A(R)$ to bound this localized form. However, if we impose a margin requirement support of the localized form will still be inside U_i. Then, as for smaller R the margin requirement relaxes, we may use $A(R/C)$ to bound these localized forms. The margin requirement (2.10) above will be called the strict margin requirement, we will also use in coming sections the following relaxed margin requirement

$$margin(f_i) \geq 1/C_0 - 3(1/R)^{1/N}. \tag{2.12}$$

Both of these margin requirements become more strict as R grows, that is for larger R fewer functions satisfy them. Thus we introduce the following variation of $A(R)$

$$\overline{A}(R) := \sup_{C_0 2^{C_1/2} \leq r \leq R} A(r)$$

that will be of use in the proof.

After introducing the relevant notation and concepts, we note one last simple observation before proceeding to give an overview of the proof. As noted above we have from the conservation of mass property the inequality

$$\int |T_i f_i(\overline{\xi}, \xi_n)|^2 d\overline{\xi} = \|f_i\|_2^2.$$

By simply integrating in the last variable we get

$$\|T_i f_i\|_{L^2(Q)} \leq R^{1/2} \|f_i\|_2. \tag{2.13}$$

Thus from the Hölder inequality we have

$$\|T_1 f_1 T_2 f_2\|_{L^1(Q)} \leq R \|f_1\|_2 \|f_2\|_2. \tag{2.14}$$

This is what we will refer to as our trivial L^1 estimate. Note that if we can prove an estimate of this form for L^2 norms with an appropriate negative power of R, by the Hölder inequality and interpolation we can control $A(R)$. Although such an estimate is clearly impossible, we can still exploit this idea partially. Using the idea of wave-packet decomposition, we write Tf_i as sum of functions each localized on a sub-cube of Q. For such a localized function an appropriate estimate is possible on other cubes. On the cube to which a function is localized, we use an inductive hypothesis. Thus, together with interpolation, we are able to control $A(R)$ with sum of $A(R/C)$ with a constant. Now for a non-endline exponent this method gives a negative power of R instead of this constant, which implies desired boundedness for $A(R)$. This is the induction on scales technique of Wolff with which he solved the Klainerman-Machedon conjecture for the light cone except for the endpoint in [34]. But for the endline this arguments alone is not sufficient.

It turns out that the key point is the concept of concentration. To make this concept more clear we define the notion of a disk centered at a point $\xi = (\bar{\xi}, \xi_n)$ of radius r as the intersection of the plane $\mathbb{R}^{n-1} \times \{\xi_n\}$ with the closed ball of radius r centered at ξ. We denote such a disk by $D(\xi, r)$. By concentration we mean having for a radius r much smaller than R

$$\int_{D(\xi,r)} |T_i f_i(\bar{\xi}, \xi_n)|^2 d\bar{\xi} \approx \|f_i\|_2^2$$

for both $i = 1, 2$. If this happens, the concentration persists for a cube Q' of radius r centered at ξ due to the geometry of the cone, - all vectors lying on the cone has an angle $\pi/4$ with the plane $\mathbb{R}^{n-1} \times \{0\}$, that is they have constant velocity, thus the concentration cannot disperse faster than a fixed rate-. Due to transversality we must have very fast decay of $|T_1 f_1 T_2 f_2|$ outside of this box: recall that the transversality assumption forces $T_i f_i$ to concentrate on tubes in different directions. Thus we may regard the quantity $|T_1 f_1 T_2 f_2|$ as essentially zero outside of Q'. Since our trivial L^1 estimate (2.14) depends on the side-length of the cube Q, having a much smaller support improves this estimate, which is sufficient to obtain the desired result for functions $f_i,\ i = 1, 2$ with concentration. If the concentration phenomenon does not occur it is natural to expect $A(R/C)$ to be less than a fraction of $A(R)$. In that case proving that $A(R)$ is less than the sum of $A(R/C)$ with a constant suffices, since this directly implies boundedness of $A(R)$. How to obtain such an expression is just what we mentioned in the previous paragraph.

Executing this vision requires much work. Here we give a brief outline of the rest of chapter 2. In the next section we will give technical definitions, and some preparatory lemmas. The section 2.3 will use these definitions and lemmas to decompose our functions to pieces localized to sub-cubes, and prove properties that these pieces satisfy. The section 2.4 will use these properties as well as the concept of concentration to conclude the proof.

2.2 Some Definitions and Preparatory Lemmas

We start with definitions related to localization to sub-cubes. Let Q be a cube with side-length R. We will use $K_j(Q)$ to denote cubes that arise when we partition Q into subcubes of sidelength $2^{-j}R$. Thus the cardinality of this collection is 2^{jn}. By (c,k) interior $I^{c,k}(Q)$ of Q we will mean the set

$$I^{c,k}(Q) := \bigcup_{q \in K_k(Q)} (1-c)q.$$

Here $(1-c)q$ is a cube with the same center as q that has $1-c$ times the side-length of q. Of course, $0 < c < 1$ and k is a positive integer.

We go on with definitions concerning localization of mass described in the section above. We let η be a fixed non-negative Schwarz function of total mass 1 on $\mathbb{R}^{n-1}$ with Fourier transform supported on the unit disk. We can construct such a function using a real, even compactly supported smooth function ϕ. Because the function is even, $\widehat{\phi}$ is also real. Thus if we let $\widehat{\eta} = \phi * \phi$, then we have $\eta(x) = |\widehat{\phi}|^2(-x)$. Thus we have non-negativity for η and compact support for $\widehat{\eta}$. If we multiply ϕ by an appropriate constant we also obtain the condition that the total mass of η is 1. Using this we define η_r for $r > 0$ by $\eta_r(x) := r^{-n}\eta_0(x/r)$, where $x \in \mathbb{R}^{n-1}$. We define for $i = 1, 2$ the cone subsets

$$C^i(\lambda) := \{(\overline{\lambda} - r\omega, \lambda_n + r) \in \mathbb{R}^n : r \in \mathbb{R},\ \omega \in U_i',\ |\omega| = 1\},$$

and r neighborhoods of these subsets, $C^i(\lambda, r)$. For a given disk $D(\lambda, r)$ localization operators

$P_D, 1-P_D$ are defined as follows: let $f_i \in L^2(U_i')$ and $\widehat{\beta}_r := \chi_D * \eta_{r^{1-1/N}}$. Then

$$P_D f_i(\xi) := T_i(\beta_r * (e^{-2\pi i x\cdot\lambda_n} f_i(x)))(\overline{\xi}, \xi_n - \lambda_n),$$

$$(1-P_D)f_i(\xi) := T_i(e^{-2\pi i x\cdot\lambda_n} f_i(x) - \beta_r * (e^{-ix\cdot\lambda_n} f_i(x)))(\overline{\xi}, \xi_n - \lambda_n),$$

Thus we actually have defined two different operators, one for T_1 and another for T_2. However the reader will easily discern which one is used from the subscript of function f_i. This operator basically localizes the image of a function f_i under the operator T_i to $C^i(\lambda, r)$ while ensuring that this localization is still image of a function under T_i. Thus for each f_i, there are two function $g_i, h_i \in L^2(U_i)$ with $f_i = g_i + h_i$ and

$$P_D f_i(\xi) = T_i g_i(\xi), \quad (1-P_D)f_i(\xi) = T_i h_i(\xi).$$

Furthermore,

$$\|P_D f_i\|_{L^2(\mathbb{R}^{n-1}\times\{0\})} = \|\beta_r * (e^{-2\pi i x\cdot\lambda_n} f_i(x))\|_2 = \|g_i\|_2$$

$$\|(1-P_D)f_i\|_{L^2(\mathbb{R}^{n-1}\times\{0\})} = \|e^{-2\pi i x\cdot\lambda_n} f_i(x) - \beta_r * (e^{-2\pi i x\cdot\lambda_n} f_i(x))\|_2 = \|h_i\|_2$$

In what follows we will, by a small abuse of notation, we use $\|P_D f_i\|_2$ to denote $\|P_D f_i\|_{L^2(\mathbb{R}^{n-1}\times\{0\})}$; and $\|(1-P_D)f_i\|_2$ to denote $\|(1-P_D)f_i\|_{L^2(\mathbb{R}^{n-1}\times\{0\})}$. We now prove a result that make rigorous the concentration properties of this operator. The smooth cutoff for the disk $D(\lambda, r)$

$$\widetilde{\chi}_D(x) := \big(1 + \frac{|x-\overline{\lambda}|}{r}\big)^{N^{-10}}$$

will be useful to make rigorous how strong these concentration properties are.

Lemma 2 *Let $r \geq C_0$, and let $D = D(\lambda, r)$. Let $f_i \in L^2(U_i)$ be functions with $margin(f_i) \geq$*

$C_0 r^{-1+1/N}$. *Then we have the following estimates for* $P_D f_i$

$$margin(P_D f_i) \geq margin(f_i) - Cr^{-1+1/N} \tag{2.15}$$

$$\|\widetilde{\chi}_D^{-N} P_D f_i\|_{L^2(D_+^{ext})} \lesssim r^{-N^2} \|f_i\|_2 \tag{2.16}$$

$$\|(1 - P_D) f_i\|_{L^2(D_-)} \lesssim r^{-N} \|f_i\|_2 \tag{2.17}$$

$$\|P_D f_i\|_2^2 \leq \|f_i\|_{L^2(D_+)}^2 + Cr^{-N} \|f_i\|_2^2 \tag{2.18}$$

$$\|(1 - P_D) f_i\|_2^2 \leq \|f_i\|_{L^2(D_-^{ext})}^2 + Cr^{-N} \|f_i\|_2^2 \tag{2.19}$$

$$\|P_D f_i\|_2,\ \|(1 - P_D) f_i\|_2 \leq \|f_i\|_2 \tag{2.20}$$

where the notation D^{ext} *denotes the rest of the plane* $\mathbb{R}^{n-1} \times \{\lambda_n\}$ *on which* D *lies, and* $D^{\pm} := D(\lambda, r(1 \pm r^{-1/2N}))$.

Proof. The margin estimate follows from the radius of support of β_r, which can be computed from very elementary properties of the Fourier transform. The radius of support of this function is $Cr^{1/N-1}$, thus support of a function convolved with it can enlarge at most this amount. For the other properties we observe that

$$P_D f_i(\lambda) := T_i(\beta_r * (e^{-2\pi i x \cdot \lambda_n} f_i(x)))(\overline{\lambda}, 0) = \widehat{\beta}_r(\overline{\lambda}) \widehat{e^{-2\pi i x \cdot \lambda_n} f_i}(\overline{\lambda}),$$

$$(1 - P_D) f_i(\lambda) := T_i(e^{-2\pi i x \cdot \lambda_n} f_i(x) - \beta_r * (e^{-ix \cdot \lambda_n} f_i(x)))(\overline{\xi}, 0) = (1 - \widehat{\beta}_r(\overline{\lambda})) \widehat{e^{-2\pi i x \cdot \lambda_n} f_i}(\overline{\lambda}).$$

Therefore investigating the behaviour of $\widehat{\beta_r} = \chi_D * \eta_{r^{-1+1/N}}$ will be helpful. Since $\eta_{r^{1-1/N}}$ is a Schwartz function, it has decay faster than any polynomial rate, this combined with the radius of the disk D^+ implies that if $x \in D_+^{ext}$

$$\widetilde{\chi}_D^{-N}(x)(\chi_D * \eta_{r^{1-1/N}})(x) \lesssim r^{-N^2}.$$

This gives (2.16), (2.18). Similarly the size of D^- combined with the fast decay of $\eta_{r^{1-1/N}}$ shows

$$\chi_D * \eta_{r^{1-1/N}}(x) \geq 1 - Cr^{-N},$$

which gives (2.17), (2.19). Finally it is easy to see that

$$0 \leq \chi_D * \eta_{r^{1-1/N}}(x) \leq 1$$

and this proves (2.20). ■

After this investigation of localization of the operator P_D to disks, we turn to localization to cubes. To this end we will look at the operators T_i more closely. We have

$$\begin{aligned} T_i f_i(\xi) = a_i(x) \widehat{f_i(x) e^{-2\pi i |x| \cdot \xi_n}}(\overline{\xi}) &= -\widehat{f}_i * \int_{\mathbb{R}^{n-1}} a_i(x) e^{-2\pi i (x \cdot \overline{\xi} + |x| \cdot \xi_n)} dx \\ &= -T_i f_i(\overline{\xi}, 0) * \int_{\mathbb{R}^{n-1}} a_i(x) e^{-2\pi i (x \cdot \overline{\xi} + |x| \cdot \xi_n)} dx. \end{aligned}$$

We thus define

$$K_{\xi_n}(\overline{\xi}) := -\int_{\mathbb{R}^{n-1}} a_i(x) e^{-2\pi i (x \cdot \overline{\xi} + |x| \cdot \xi_n)} dx.$$

Thus we have

$$T_i f_i(\xi) = (T_i f_i(\cdot, 0) * K_{\xi_n}(\cdot))(\overline{\xi}).$$

Similarly we can replace the zero in the nth coordinate by any λ_n

$$T_i f_i(\xi) = (T_i f_i(\cdot, \lambda_n) * K_{\lambda_n - \xi_n}(\cdot))(\overline{\xi}).$$

We have, from standard non-stationary phase methods, the following strong decay estimate for this kernel:

$$|K_{\xi_n}(\overline{\xi})| \lesssim \left(1 + \operatorname{dist}(\xi, C^i(0))\right)^{-N^{10}}, \tag{2.21}$$

see [28] equation (33), or [15] Lemma 11.

Another important property of T_i is

$$|T_i f_i| \leq \|f_i\|_{L^1} \lesssim \|f_i\|_{L^2}.$$

This clearly follows from the definition of T_i and compactness of the support of f_i. From this we will deduce two important conclusions. First, combining this with the conservation of mass property that was introduced above

$$\|f_i\|_{L^2} = \|T_i f_i(\overline{\xi}, \xi_n)\|_{L^2(\mathbb{R}^{n-1}\times\xi_n)},$$

we can write

$$|T_i f_i| \lesssim \|T_i f_i(\overline{\xi}, \xi_n)\|_{L^2(\mathbb{R}^{n-1}\times\xi_n)}. \tag{2.22}$$

Second combining this with the dominated convergence theorem we see that $T_i f_i$ is continuous on $\mathbb{R}^n$. Now we state our lemma about localization to cubes.

Lemma 3 *Let $D(\xi_D, r)$ be a disk with radius $r \geq 2^{C_0}$. Let $f_i \in L^2(U_i)$, $i = 1,2$ be two functions with f_1 satisfying*

$$margin(f_1) \geq C_0 r^{-1+1/N}.$$

Then for $1 \leq q < r \leq 2$ and $r \lesssim R$ we have

$$\|((1-P_D)f_1)T_2 f_2\|_{L^q L^r(Q(\xi_D, C^{-1}r))} \lesssim r^{C-N}\|f_1\|_2\|f_2\|_2, \tag{2.23}$$

and

$$\|(P_D f_1)T_2 f_2\|_{L^q L^r(Q(\xi_Q,R))\setminus C^1(\xi_D, Cr+R^{1/N}))} \lesssim R^{C-N}\|f_1\|_2\|f_2\|_2 \tag{2.24}$$

whenever $|\xi_n - \xi_{D_n}| \leq C_0 R$.

Before the proof we note that an analogue of this with the places of f_1, f_2 changed also holds, and the proof below works in that case as well with trivial changes.

Proof. Note that the region of integration for the left hand side of (2.23) has measure bounded by a power of r, and as the right hand side shows we can lose such powers. This same argument also holds for (2.24). Thus using Hölder's inequality, it suffices to prove these statements for $p = q = 2$. First we prove (2.23). It clearly suffices to prove an appropriate L^∞ estimate for $(1 - P_D)f_1$ and an appropriate L^2 estimate on $T_2 f_2$. The estimate on $T_2 f_2$ that we need is (2.13). As for the L^∞ estimate, note that

$$\begin{aligned}(1 - P_D)f_1(\xi) &= T_1(e^{-2\pi i x\cdot\xi_{Dn}} f_1(x) - \beta_r * (e^{-2\pi i x\cdot\xi_{Dn}} f_1(x)))(\overline{\xi}, \xi_n - \xi_{Dn}) \\ &= T_1(e^{-2\pi i x\cdot\xi_{Dn}} f_1(x) - \beta_r * (e^{-2\pi i x\cdot\xi_{Dn}} f_1(x)))(\cdot, 0) * K_{\xi_n - \xi_{Dn}}(\cdot)(\overline{\xi}) \\ &= (1 - P_D)f_1(\cdot, \xi_{Dn}) * K_{\xi_n - \xi_{Dn}}(\cdot)(\overline{\xi}) \\ &= \int (1 - P_D)f_1(\overline{\xi} - \overline{\lambda}, \xi_{Dn}) K_{\xi_n - \xi_{Dn}}(\overline{\lambda}) d\overline{\lambda}.\end{aligned} \tag{2.25}$$

Thus using Hölder's inequality we can write

$$|(1 - P_D)f_1(\xi)|^2 \leq \int_{\mathbb{R}^{n-1}} |(1 - P_D)f_1(\overline{\xi} - \overline{\lambda}, \xi_{Dn})|^2 |K_{\xi_n - \xi_{Dn}}(\overline{\lambda})| d\overline{\lambda} \cdot \int_{\mathbb{R}^{n-1}} |K_{\xi_n - \xi_{Dn}}(\overline{\lambda})| d\overline{\lambda}.$$

As we have $|\xi_n - \xi_{Dn}| < C^{-1} r$ for ξ in our cube, the second integral can be bounded by r^C for some appropriate C. For the first integral we partition the domain of integration as union of $D(0, r/3)$ and its complement $\mathbb{R}^{n-1} \setminus D(0, r/3)$. On the first of these $|K_{\xi_n - \xi_{Dn}}(\overline{\lambda})| \lesssim 1$, thus we have

$$\int_{D(0,r/3)} |(1 - P_D)f_1(\overline{\xi} - \overline{\lambda}, \xi_{Dn})|^2 |K_{\xi_n - \xi_{Dn}}(\overline{\lambda})| d\overline{\lambda} \lesssim \|(1 - P_D)f_1\|^2_{L^2(D(\xi_D, r/2))}.$$

But the disk $D(\xi_D, r/2))$ lies inside D^- thus we can use (2.17) to bound it. To deal with the second integral we first note from (2.20) that

$$|(1 - P_D)f_1| \lesssim \|T_1 f_1(\cdot, \xi_{Dn})\|_{L^2(\mathbb{R}^{n-1} \times \xi_{Dn})} \approx \|f_i\|_2$$

Hence the second integral satisfies

$$\lesssim \|f_1\|_2^2 \int_{\mathbb{R}^{n-1}\setminus D(0,r/3)} |K_{\xi_n-\xi_{D_n}}(\overline{\lambda})|d\overline{\lambda}$$

Since $|\xi_n - \xi_{D_n}| < C^{-1}r$, from here decay properties of the kernel gives

$$\lesssim r^{-N}\|f_1\|_2^2.$$

Combining these results gives (2.23).

To prove (2.24) we follow a similar path: we prove an L^∞ estimate on $P_D f_1$, and use (2.13) for $T_2 f_2$. As in (2.25) we can write

$$P_D f_1(\xi) = \int P_D f_1(\overline{\xi} - \overline{\lambda}, \xi_{D_n}) K_{\xi_n-\xi_{D_n}}(\overline{\lambda})d\overline{\lambda},$$

and use the Hölder inequality and the decay estimate (2.21) to write

$$|P_D f_1(\xi)|^2 \lesssim R^C \int_{\mathbb{R}^{n-1}} |P_D f_1(\overline{\xi} - \overline{\lambda}, \xi_{D_n})|^2 |K_{\xi_n-\xi_{D_n}}(\overline{\lambda})|d\overline{\lambda}$$

We divide the domain of integration into $C^1((0,\xi_{D_n}), r + R^{1/2N}) \cap (\mathbb{R}^{n-1} \times \{\xi_n\})$ and its complement $(\mathbb{R}^{n-1} \setminus (C^1((0,\xi_{D_n}), r + R^{1/2N}) \cap ((\mathbb{R}^{n-1} \times \{\xi_n\})$. On the second domain we use the fact that $|P_D f_1| \lesssim \|f_1\|_2$, and the fast decay of the kernel $K_{\xi_n-\xi_{D_n}}$. For the first domain we observe that since $\overline{\xi} \notin C^1(\xi_D, Cr + R^{1/N}) \cap (\mathbb{R}^{n-1} \times \{\xi_n\})$, we must have $\text{dist}(\overline{\xi} - \overline{\lambda}, D) > 2r + R^{1/N}/2$. Combining this with the fact that $|K_{\xi_n-\xi_{D_n}}(\overline{\lambda})| \lesssim 1$ allows us write for this integral

$$\lesssim \|P_D f_1\|_{L^2(D^{ext}(\xi_D, 2r+R^{1/N}/2))}$$

From here the decay of smooth cutoff $\eta_{r^{1-1/N}}$ used for defining P_D gives the desired result.

■

For a cube Q of side-length R, we know that

$$\|T_1 f_1\|_{L^2(Q)} \lesssim R^{1/2}\|f_1\|_2$$

from conservation of mass property. If we want to replace the cube with the cone neighborhood $C^2(\xi_C, R)$, then simply using the conservation of mass property and integrating in the nth coordinate is not enough, since domain of integration in the nth variable has infinite length. But the transversality of normals of $C^1(0)$, and $C^2(\xi_C)$ can be used to essentially limit this length to R, and thus prove the same estimate for $C^2(\xi_C, R)$. Heuristically this can be seen by considering a partition of f_1 into small caps of radius R^{-1}, dual tubes of which will be oriented along the normals of $C^2(\xi_C)$. Owing to transversality each such tube, although of length R^2, will have an intersection with $C^2(\xi_C, R)$ that has length R in the nth coordinate. Although our proof below will not follow exactly such a route, this makes it clear why the result holds.

Lemma 4 *Let $R \gg 1$, $f_1 \in L^2(U_1)$, and let $\xi_C \in \mathbb{R}^n$ be an arbitrary point. Then we have*

$$\|T_1 f_1\|_{L^2(C^2(\xi_C,R))} \lesssim R^{1/2}\|f_1\|_2.$$

Proof. Before begining the proof in earnest we note that by modulating f_1 appropriately, we can take ξ_C to be the origin in $\mathbb{R}^n$. We will use duality type arguments in the proof. Let T denote the operator $\chi_{C^2(\xi_C,R)} T_1$. Then our lemma is equivalent to proving that

$$\|T f_1\|_{L^2(\mathbb{R}^n)} \lesssim R^{1/2}\|f_1\|_2.$$

This operator maps functions on $\mathbb{R}^{n-1}$ to functions on $\mathbb{R}^{n-1}$. Its adjoint T^* is defined for functions g_1 on $\mathbb{R}^n$ via the following formula

$$T^* g_1(x) = a_1(x) \int_{\mathbb{R}} \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x, \xi_n) e^{2\pi i |x| \cdot \xi_n} d\xi_n$$

where $\widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\xi_n)$ denotes $g_1 \cdot \chi_{C^2(0,R)}(\bar{\xi},\xi_n)$ Fourier transformed in $\bar{\xi}$ variable. As is well known $\|T\| = \|T^*\|$, hence it suffices to prove

$$\|T^* g_1\|_{L^2(\mathbb{R}^{n-1})} = \|a_1(x) \int_{\mathbb{R}} \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\xi_n) e^{2\pi i |x| \cdot \xi_n} d\xi_n \|_{L^2(\mathbb{R}^{n-1})} \lesssim R^{1/2} \|g_1\|_{L^2(\mathbb{R}^n)}.$$

By squaring we can write the first term as

$$\|T^* g_1\|^2_{L^2(\mathbb{R}^{n-1})} = \langle T^* g_1, T^* g_1 \rangle_{L^2(\mathbb{R}^{n-1})}.$$

Writing $T^* g_1$ terms explicitly we have

$$\left\langle a_1(x) \int \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\xi_n) e^{2\pi i |x| \cdot \xi_n} d\xi_n,\ a_1(x) \int \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\lambda_n) e^{2\pi i |x| \cdot \lambda_n} d\lambda_n \right\rangle_{L^2(\mathbb{R}^{n-1})}$$

which is equal to

$$\int_{\mathbb{R}^{n-1}} \int_{\mathbb{R}} \int_{\mathbb{R}} a_1^2(x) \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\xi_n) \overline{\widehat{g_1 \cdot \chi_{C^2(0,R)}}}(-x,\lambda_n) e^{2\pi i |x| \cdot (\xi_n - \lambda_n)} d\xi_n d\lambda_n dx.$$

We change the order of integration to obtain

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \langle a_1(x) \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\xi_n) e^{2\pi i |x| \cdot \xi_n}, a_1(x) \widehat{g_1 \cdot \chi_{C^2(0,R)}}(-x,\lambda_n) e^{2\pi i |x| \cdot \lambda_n} \rangle_{L^2(\mathbb{R}^{n-1})} d\xi_n d\lambda_n.$$

Consider the inner product inside the integral for fixed ξ_n, λ_n. In this case we may regard

$$a_1(x) \widehat{g_1 \cdot \chi_{C^2(0,R)}}(x,\xi_n) e^{2\pi i |x| \cdot \xi_n}$$

as the image of $g_1(\bar{\xi},\xi_n)$ under an operator T_{ξ_n} defined on $L^2(\mathbb{R}^{n-1})$. The same view, of course, can be taken towards the other term in the inner product. The dual for such an operator, $T^*_{\xi_n}$, can be calculated from the formula defining adjoint operators on Hilbert spaces and properties of the Fourier transform to be

$$T^*_{\xi_n} h(x) = \chi_{C^2(0,R)}(x,\xi_n) a_1 \cdot \widehat{h \cdot e^{-2\pi i |\cdot| \xi_n}}(x).$$

Again using the equality that defines adjointness on Hilbert spaces we can write our integral above as

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \left\langle g_1, T^*_{\xi_n} T_{\lambda_n} g_1 \right\rangle_{L^2(\mathbb{R}^{n-1})} d\xi_n d\lambda_n. \tag{2.26}$$

The following statement implies our result

$$\left| \left\langle g_1, T^*_{\xi_n} T_{\lambda_n} g_1 \right\rangle_{L^2(\mathbb{R}^{n-1})} \right| \lesssim (1 + |\xi_n - \lambda_n|/R)^{-N} \|g_1(\cdot, \xi_n)\|_2 \|g_1(\cdot, \lambda_n)\|_2. \tag{2.27}$$

We will first see how to arrive at our result from this. Our integral (2.26) above satisfies

$$\lesssim \int_{\mathbb{R}} \int_{\mathbb{R}} (1 + |\xi_n - \lambda_n|/R)^{-N} \|g_1(\cdot, \xi_n)\|_2 \|g_1(\cdot, \lambda_n)\|_2 d\xi_n d\lambda_n.$$

We define a function ω on $\mathbb{R}$ with $\omega(s) = (1 + |s|/R)^{-N}$. Then we can write the integral above as

$$= \int_{\mathbb{R}} \omega * \|g_1\|_{L^2(\mathbb{R}^{n-1} \times \{\xi_n\})}(\lambda_n) \|g_1(\cdot, \lambda_n)\|_2 d\lambda_n.$$

To this we can apply the Hölder inequality to obtain

$$\leq \Big(\int_{\mathbb{R}} \left[\omega * \|g_1\|_{L^2(\mathbb{R}^{n-1} \times \{\xi_n\})}(\lambda_n) \right]^2 d\lambda_n \Big)^{1/2} \cdot \|g_1\|_{L^2(\mathbb{R}^n)}$$

We apply Young's inequality to the first term to obtain

$$\lesssim \|\omega\|_1 \|g_1\|^2_{L^2(\mathbb{R}^n)},$$

and this, as a simple calculation shows, satisfy

$$\lesssim R\|g_1\|^2_{L^2(\mathbb{R}^n)}.$$

Thus it remains to show (2.27). For $|\xi_n - \lambda_n| \lesssim R$ this amounts to showing that

$$|\left\langle g_1, T^*_{\xi_n} T_{\lambda_n} g_1 \right\rangle_{L^2(\mathbb{R}^{n-1})}| \lesssim \|g_1(\cdot, \xi_n)\|_2 \|g_1(\cdot, \lambda_n)\|_2.$$

As both T_{λ_n}, $T^*_{\xi_n}$ are merely combinations of modulations, multiplication by characteristic functions and the Fourier transform they can easily seen to be bounded. Thus applying the Cauchy-Schwarz inequality to the left hand side and then using boundedness of these operators we obtain the desired result. Thus we may assume $|\xi_n - \lambda_n| \gg R$. In this case we first look at our operator $T^*_{\xi_n} T_{\lambda_n}$ closely:

$$T^*_{\xi_n} T_{\lambda_n} g_1(\overline{\xi}) = \chi_{C^2(0,R)}(\overline{\xi}, \xi_n)\Big[- g_1 \cdot \chi_{C^2(0,R)}(\overline{\theta}, \lambda_n) * \int a_1^2(x) \cdot e^{-2\pi i(|x|(\xi_n - \lambda_n) + x\cdot\overline{\theta})} dx\Big](\overline{\xi}).$$

This follows directly from definitions of $T^*_{\xi_n}$, T_{λ_n}. The integral term is of course very similar to the kernel we introduced before, so again we define

$$\widetilde{K}_{\xi_n - \lambda_n}(\overline{\theta}) = - \int a_1^2(x) \cdot e^{-2\pi i(|x|(\xi_n - \lambda_n) + x\cdot\overline{\theta})} dx.$$

This kernel, by the same reasoning, satisfies a decay estimate very similar to (2.21),

$$\widetilde{K}_{\lambda_n - \xi_n}(\overline{\theta}) \lesssim (1 + dist((\overline{\theta}, \xi_n - \lambda_n), C^1(0)))^{-N^{10}}, \tag{2.28}$$

and we have

$$T^*_{\xi_n} T_{\lambda_n} g_1(\overline{\xi}) = \chi_{C^2(0,R)}(\overline{\xi}, \xi_n)\Big[g_1 \cdot \chi_{C^2(0,R)}(\overline{\theta}, \lambda_n) * \widetilde{K}_{\xi_n - \lambda_n}(\overline{\theta})\Big](\overline{\xi})$$
$$= \chi_{C^2(0,R)}(\overline{\xi}, \xi_n) \int_{\mathbb{R}^{n-1}} g_1 \cdot \chi_{C^2(0,R)}(\overline{\theta}, \lambda_n) \widetilde{K}_{\xi_n - \lambda_n}(\overline{\xi} - \overline{\theta}) d\overline{\theta}.$$

Thus

$$|T^*_{\xi_n} T_{\lambda_n} g_1(\overline{\xi})| \leq \|g_1(\cdot, \lambda_n)\|_2 \Big[\int_{\mathbb{R}^{n-1}} \chi_{C^2(0,R)}(\overline{\xi}, \xi_n) \chi_{C^2(0,R)}(\overline{\theta}, \lambda_n) |\widetilde{K}_{\xi_n - \lambda_n}(\overline{\xi} - \overline{\theta})|^2 d\overline{\theta}\Big]^{\frac{1}{2}}.$$

We will now use this information to estimate the inner product. We start with the Cauchy-Schwarz inequality

$$|\left\langle g_1, T^*_{\xi_n} T_{\lambda_n} g_1 \right\rangle_{L^2(\mathbb{R}^{n-1})}| \lesssim \|g_1(\cdot, \xi_n)\|_2 \|T^*_{\xi_n} T_{\lambda_n} g_1\|_2.$$

Then we use our estimate above on $|T^*_{\xi_n} T_{\lambda_n} g_1|$

$$\leq \|g_1(\cdot, \xi_n)\|_2 \|g_1(\cdot, \lambda_n)\|_2 \Big[\int_{\mathbb{R}^{n-1}} \chi_{C^2(0,R)}(\overline{\xi}, \xi_n) \chi_{C^2(0,R)}(\overline{\theta}, \lambda_n) |\widetilde{K}_{\xi_n - \lambda_n}(\overline{\xi} - \overline{\theta})|^2 d\overline{\xi} d\overline{\theta}\Big]^{\frac{1}{2}}.$$

From transversality of $C^1(0,R)$, $C^2(0,R)$, and the estimate (2.28) this last term satisfies

$$\lesssim (1 + |\xi_n - \lambda_n|/R)^{-N}$$

and this finishes our proof.

■

We next give a simple corollary of this lemma which, in certain situations, will be used as a substitute for our simple L^1 estimate (2.14).

Corollary 1 *Let $f_i \in L^2(U_i)$, $i = 1, 2$. Let $R \gg r \gg 1$, $\xi \in \mathbb{R}^n$ and let Q be cube of* *side-length R. Then*

$$\|T_1 f_1 T_2 f_2\|_{L^1(C^3(\xi,r)\cap Q)} \lesssim r^{1/2} R^{1/2} \|f_1\|_2 \|f_2\|_2.$$

Proof. This will easily follow from Lemma 3. We have

$$\|T_1 f_1 T_2 f_2\|_{L^1(C^3(\xi,r)\cap Q)} \leq \|T_1 f_1 T_2 f_2\|_{L^1(C^1(\xi,r)\cap Q)} + \|T_1 f_1 T_2 f_2\|_{L^1(C^2(\xi,r)\cap Q)}.$$

We handle the first summand on the right, the second term is handled the same way. Using the Hölder inequality

$$\|T_1 f_1 T_2 f_2\|_{L^1(C^1(\xi,r)\cap Q)} \leq \|T_1 f_1\|_{L^1(Q)} \|T_2 f_2\|_{L^1(C^1(\xi,r))}$$

From the analogue of Lemma 4 for functions in $L^2(U_2)$, and (2.14) we have

$$\lesssim r^{1/2} R^{1/2} \|f_1\|_2 \|f_2\|_2,$$

and thus we are done.

■

The following lemma represents the gain that can be obtained from having small angular dispersion for at least one of the functions f_1, f_2 in an L^2 estimate. As this suggests in the coming chapters we will partition our functions into small pieces of small angular dispersion to obtain such a gain. This lemma will be of fundamental use to us in this process.

Lemma 5 *Let $r \geq C_0$. Let $f_1 \in L^2(U_1)$ with angular dispersion at most C/r and $f_2 \in L^2(U_2)$. Suppose f_1, f_2 satisfy the margin requirement $margin(f_i) \geq 1/3C_0$, $i = 1, 2$. Let ϕ be a compactly supported function in $B(0,1)$ with Fourier transform non-vanishing on*

$B(0, C)$. *Let* $\phi_R(y) = R^n \phi(Ry)$. *Then for all cubes* Q *of side-length* R *we have*

$$\|T_1 f_1 T_2 f_2\|_{L^2(Q)} \lesssim R^{-n/2} \|T_1 f_1 \widehat{\phi_R}\|_2 \|T_2 f_2 \widehat{\phi_R}\|_2.$$

Proof. We first note that $\widehat{\phi_R}(\xi) = \widehat{\phi}(\xi/R)$. We have

$$\|T_1 f_1 T_2 f_2\|_{L^2(Q)} \lesssim \|T_1 f_1 \widehat{\phi_R} T_2 f_2 \widehat{\phi_R}\|_2.$$

We can of course write

$$T_i f_i(\xi) = \int_{U_i} f(x) e^{-2\pi i x\cdot\xi} dx = \int_{S_i} g(y) e^{-2\pi i y\cdot\xi} d\sigma_i(y) = \widehat{f_i d\sigma_i}(\xi)$$

for S_i cone neighborhoods as defined before, $g(x, |x|) := f(x)$, and $d\sigma$ pullback of the Lebesgue measure on U_i. Thus

$$\|T_1 f_1 \widehat{\phi_R} T_2 f_2 \widehat{\phi_R}\|_2 = \|\widehat{f_1 d\sigma_1}\widehat{\phi_R}\widehat{f_2 d\sigma_2}\widehat{\phi_R}\|_2 = \|f_1 \widehat{d\sigma_1 * \phi_R} f_2 \widehat{d\sigma_2 * \phi_R}\|_2$$

We use the Plancherel theorem to obtain

$$\|(f_1 d\sigma_1 * \widehat{\phi_R) * (}f_2 d\sigma_2 * \phi_R)\|_2 = \|(f_1 d\sigma_1 * \phi_R) * (f_2 d\sigma_2 * \phi_R)\|_2$$

We will, instead of L^2, will prove estimates on L^1 and L^∞ norms and use interpolation. We have from Young's inequality

$$\|(f_1 d\sigma_1 * \phi_R) * (f_2 d\sigma_2 * \phi_R)\|_1 \lesssim \|f_1 d\sigma_1 * \phi_R\|_1 \|f_2 d\sigma_2 * \phi_R\|_1$$

For the L^∞ case we have

$$|(f_1 d\sigma_1 * \phi_R) * (f_2 d\sigma_2 * \phi_R)(y)| \leq \int \big|(f_1 d\sigma_1 * \phi_R)(y - z)(f_2 d\sigma_2 * \phi_R)(z)\big| dz$$

Expressions $f_i d\sigma_i * \phi_R$ are functions supported inside C/R neighborhoods of supports of measures $f_i d\sigma_i$. For any y measure of the set of z for which integrand on the right handside satisfies CR^{-n}. Thus we have

$$\lesssim R^{-n}\|f_1 d\sigma_1 * \phi_R\|_\infty \|f_2 d\sigma_2 * \phi_R\|_\infty.$$

Now interpolation gives

$$\|(f_1 d\sigma_1 * \phi_R) * (f_2 d\sigma_2 * \phi_R)\|_2 \lesssim R^{-n/2}\|f_1 d\sigma_1 * \phi_R\|_2 \|f_2 d\sigma_2 * \phi_R\|_2$$

Applying the Plancherel theorem to the right hand side gives our result. ■

This lemma also have a global version. Its proof is essentially same as the one we presented above, and can be found in [28].

Lemma 6 *Let $r \geq C_0$. Let $f_1 \in L^2(U_1)$ with angular dispersion at most C/r, and $f_2 \in L^2(U_2)$. Suppose f_1, f_2 satisfy the margin requirement $margin(f_i) \geq 1/3C_0,\ i = 1, 2$. Then we have*

$$\|T_1 f_1 T_2 f_2\|_2 \lesssim R^{\frac{2-n}{2}}\|f_1\|_2 \|f_2\|_2.$$

We now prove a technical lemma of a very general nature. It will be frequently used throughout the rest of Chapter 2.

Lemma 7 *Let Q_R be a cube with center ξ_{Q_R} and side-length R. Let $0 < c \leq 2^{-C}$, and $f \in L^\infty(C_0 Q)$. Then we have a cube Q' with side-length CR contained in $C^2 Q$ satisfying*

$$\|f\|_{L^q L^r(Q)} \leq (1 + Cc)\|f\|_{L^q L^r(I^{c,C_0}(Q'))}.$$

Proof. We will prove this for the L^1 norm and then extend it to our case via duality arguments. We let $Q(\xi, R)$ denote a cube centered at ξ with side-length R. If for an integrable function g we can prove the following

$$\|g\|_{L^1(Q)} \leq \frac{1}{|Q_R|}\int_{Q_R}(1+Cc)\|g\|_{L^1(I^{c,C_0}(Q'(\xi,CR))\cap Q)}d\xi, \tag{2.29}$$

there has to be a cube satisfying the desired property for L^1 norm. By Fubini's theorem we have

$$\begin{aligned}\int_{Q_R}\|g\|_{L^1(I^{c,C_0}(Q(\xi,CR))\cap Q_R)}d\xi &= \int_{Q_R}\int_{Q_R}|g(\zeta)|\chi_{I^{c,C_0}(Q(\xi,CR))\cap Q_R}(\zeta)d\xi d\zeta \\ &= \int_{Q_R}|g(\zeta)|\Big(\int_{Q_R}\chi_{I^{c,C_0}(Q(\xi,CR))\cap Q_R}(\zeta)d\xi\Big)d\zeta \\ &= \int_{Q_R}|g(\zeta)||I^{c,C_0}(Q(\zeta,CR))\cap Q_R|d\zeta.\end{aligned}$$

We also have the measure estimate

$$|Q(\zeta,CR)\setminus I^{c,C_0}(Q(\zeta,CR))| \leq c|Q(\zeta,CR)|$$

which leads to

$$|Q_R| \leq (1+Cc)|I^{c,C_0}(Q(\zeta,CR))\cap Q_R|$$

for any $\zeta \in Q_R$. This combined with what we obtained from Fubini's theorem gives (2.29). So we have the L^1 case. To go to our case, first observe that it is enough to prove it for $\|f\|_{L_t^qL_x^p(Q)} = 1$. By duality there exists a function h with properties $\|h\|_{L^{q'}L^{r'}(Q)} = 1$, and

$$\int_Q |f(\xi)|h(\xi)d\xi = 1.$$

But from our result for L^1 functions we have

$$1 = \Big| \int_Q |f(\xi)| h(\xi) d\xi \Big| \leq \|fh\|_{L^1(Q)}$$
$$\leq (1 + Cc)\|fh\|_{L^1(I^{c,C_0}(Q'))}.$$

Applying the Hölder's inequality to the last term gives our lemma.

■

2.3 Main Proposition

This section will provide the main proposition that we will use to conclude the proof. However proving this main propostion is rather long and will be divided into three steps. The first is the wave packet decomposition, which will show that it is possible to partition the function f_i into pieces supported on thin strips such that, not only both pieces themselves but also their images under T_i are well localized. After this the second part will introduce a partitioning of f_i into pieces whose images under T_i localize on sub-cubes. This is the partitioning that we mentioned in the summary of the proof, and this will be achieved using the wave packet decomposition. The third part will give some further information about the decomposition to sub-cubes, and turn the information coming from the first two parts into the exact form that will be applied in the last section. Apart from these we will give an application of the wave packet decomposition that will be of use later. To do all these we will need some new notation that will be introduced along the way.

We start this section with the definition that will make precise the notion of concentration.

Definition 2 *Let $r > 0$, and let $Q \subset \mathbb{R}^n$ be a cube of side-length R centered at a point ξ_Q. Let $f_{1,\alpha} \in L^2(U_1)$, $\alpha \in A$ and $f_{2,\beta} \in L^2(U_2)$, $\beta \in B$ with A, B finite index sets. Then the* *concentration* $E_{r,Q}(f_{1,\alpha},\alpha\in A; f_{2,\beta},\beta\in B)$ *is defined as*

$$\max\Big\{\frac{1}{2}\Big(\sum_{\alpha\in A}\|f_{1,\alpha}\|_2^2\cdot\sum_{\beta\in B}\|f_{1,\beta}\|_2^2\Big)^{\frac{1}{2}},\ \sup_D\Big(\sum_{\alpha\in A}\|T_1f_{1,\alpha}\|^2_{L^2(D)}\cdot\sum_{\beta\in B}\|T_2f_{2,\beta}\|^2_{L^2(D)}\Big)^{\frac{1}{2}}\Big\},$$

where $D=D(\xi_D,r)$ *are disks with* $|\xi_{Dn}-\xi_{Q_n}|\le R/2$*.*

We now prepare for the wave packet decomposition with some definitions. We let $\mathcal{Y}_\rangle$ be a maximal $1/r$ separated subset of unit length vectors lying in U_i', and L the lattice $c^{-2}r\mathbb{Z}^{n-1}$. We will denote by $\mathbf{T}_i$ collections of tubes where a tube is defined as

$$T_i:=\{\xi\in\mathbb{R}^n:|\xi-(\overline{\xi}_{T_i}+\xi_n\omega_{T_i})|\le r\}$$

with $\overline{\xi}_{T_i}\in L,\omega_{T_i}\in\mathcal{Y}_i$ denoting the unique point of the lattice L lying inside the tube and the direction of the tube. Apart from the characteristic function of such a tube we will need the following smooth cutoff

$$\widetilde{\chi}_{T_i}(\xi):=\widetilde{\chi}_{D(\overline{\xi}_{T_i}-\xi_n\omega_{T_i},r)}(\overline{\xi}),$$

namely at every ξ_n, we are taking the previously defined smooth cutoff of the disk given by $\mathbb{R}^{n-1}\times\{\xi_n\}\cap T_i$. With these definitions at hand we state our wave packet decomposition lemma, in which we will describe the process of decomposition, and prove certain properties that the wave packets emerging satisfy. The way we decompose f_i will make clear the margin and angular dispersion properties, and property (2.30) clear. The two others following these will follow from well localization under T_i, while the last two will exploit not only this localization but also localization of pieces themselves.

Lemma 8 *Let* $f_i\in L^2(U_i)$*. We can find a decomposition of* f_i *into wave packets* ϕ_{T_i} *with following properties.*

• *The sum of wave packets give the function itself*

$$f_i = \sum_{\boldsymbol{T}_i} \phi_{T_i}. \tag{2.30}$$

• *Angular dispersion for any* ϕ_{T_i} *is at most* $CR^{-1/2}$ *and margin is at least* $margin(f_i) - CR^{-1/2}$.

• *The wave packet* $T_i\phi_{T_i}$ *is well concentrated on the tube* T_i

$$\|\phi_{T_i}\|_2 \lesssim c^{-C} r \|\widetilde{\chi}_{T_i}(\cdot, \xi_n) T_i f_i(\cdot, \xi_n)\|_{L^2(\mathbb{R}^{n-1}\times\{\xi_n\})} \tag{2.31}$$

if $|\xi_n| \lesssim R$.

• *Due to well concentration on tube* T_i *the function* $T_i\phi_{T_i}$ *is very small on domains distant from the tube: let* Q *be a cube centered at* ξ_Q *with radius* $r_Q \lesssim R$. *If* $|\xi_{Q_n}| \lesssim R$ *and* $dist(T_i, Q) \geq C_0 R$ *then*

$$\|T_i\phi_{T_i}\|_{L^\infty(Q)} \lesssim dist(T, Q)^{C-N} \|f_i\|_2. \tag{2.32}$$

• *The strong localization of* $T_i\phi_{T_i}$ *to* T_i *and the conservation of mass property allows one to obtain*

$$\sum_{\boldsymbol{T}_i} \sup_{q \in K_J(Q)} \widetilde{\chi}_{T_i}^{-3}(\xi_q) \|T_i\phi_{T_i}\|_{L^2(Cq)} \lesssim c^{-C} r \|f_i\|_2^2 \tag{2.33}$$

where ξ_q *stands for the center of* q.

• *Let, for each* T_i, m_{q_0,T_i} *be constants with* q_0 *running through a finite index, and with*

$$\sum_{q_0} m_{q_0,T_i} = 1.$$

we have

$$\Big(\sum_{q_0} \| \sum_{T_i \in \boldsymbol{T}_i} m_{q_0,T_i} \phi_{T_i}\|_2^2\Big)^{1/2} \leq (1 + Cc)\|f_i\|_2. \tag{2.34}$$

Proof. We partition the set U_i' using $\mathcal{Y}_i$

$$\Sigma = \bigcup_{\omega \in \mathcal{Y}_i} Y_\omega$$

with Y_ω denoting those points with direction closer to $\omega \in \mathcal{Y}_i$ than to the other elements of $\mathcal{Y}_i$. Let $\mathcal{G}$ be the set $r^{-1}\mathbb{Z}$. Using this set we further partition each Y_ω as

$$Y_\omega = \bigcup_{\gamma \in \mathcal{G}} Y_{\omega,\gamma}$$

where $Y_{\omega,\gamma}$ being those points closer to γ than any other point of $\mathcal{G}$. We take H to be set of translations given by elements of a disk centered at the origin with radius C/r, and $d\Omega$ to be a smooth compactly supported probability measure in the interior of this disk. Then we define, for $\Omega \in \mathcal{H},\ \omega \in \mathcal{Y},\ \gamma \in \mathcal{G}$ and $f \in L^2(U_i)$, the operators

$$P_{\Omega,\omega,\gamma} f(x) := \chi_{\Omega(Y_{\omega,\gamma})}(\frac{x}{|x|}) f(x).$$

Then of course for any $\Omega \in \mathcal{H}$

$$f_i(x) = \sum_{\omega \in \mathcal{Y}_i, \gamma \in \mathcal{G}} P_{\Omega,\omega,\gamma} f_i(x). \tag{2.35}$$

This decomposition of f_i uses sharp cutoffs, which cannnot guarantee good localization of images under the Fourier transform, thus we will average using the measure $d\Omega$. Then we will sum over these averages to obtain smooth cutoffs for strips, and finally apply smooth cutoffs to the images under the Fourier transform. To this end recall the function η defined in section 2.3. We define the smooth cutoffs $\eta^{\overline{\xi}}$ by

$$\widehat{\eta^{\overline{\xi}}}(\overline{\zeta}) := \widehat{\eta}(\frac{c^2}{r}(\overline{\zeta} - \overline{\xi})).$$

Then from the Poisson summation formula we have

$$\sum_{\bar{\xi}\in L} \widehat{\eta^{\bar{\xi}}} = 1. \tag{2.36}$$

We define the averaging operator

$$\mathcal{A}_{\omega,\gamma} f_i(x) = \int P_{\Omega,\omega,\gamma} f_i(x) d\Omega$$

and sum of averages over strips

$$\mathcal{A}_{\omega} f_i(x) = \sum_{\gamma} \mathcal{A}_{\omega,\gamma} f_i(x).$$

Using these sums and the smooth cutoffs introduced above we define

$$\phi_{T_i} := \eta^{\bar{\xi}_{T_i}} * \mathcal{A}_{\omega_{T_i}} f_i.$$

We first verify that from the sum of these ϕ_{T_i} we recover f_i.

$$\sum_{T_i} \phi_{T_i} = \sum_{\bar{\xi}_{T_i}} \sum_{\omega_{T_i}} \eta^{\bar{\xi}_{T_i}} * \mathcal{A}_{\omega_{T_i}} f_i = \sum_{\bar{\xi}_{T_i}} \eta^{\bar{\xi}_{T_i}} * \Big(\sum_{\omega_{T_i}} \sum_{\gamma\in\mathcal{G}} \mathcal{A}_{\omega_{T_i},\gamma} f_i \Big).$$

The inner integral gives from (2.35)

$$= \sum_{\omega_{T_i},\gamma} \int P_{\Omega,\omega_{T_i}} f_i(x) d\Omega = f_i(x).$$

Turning back we have

$$= \sum_{\bar{\xi}_{T_i}} \eta^{\bar{\xi}_{T_i}} * f_i = f_i * \Big(\sum_{\bar{\xi}_{T_i}} \eta^{\bar{\xi}_{T_i}} \Big).$$

At this point considering (2.36) we see that the Fourier transform of the last expression is equal to $\widehat{f_i}$. Thus we obtain what we desired:

$$\sum_{T_i} \phi_{T_i} = f_i.$$

We now turn to the angular dispersion and margin properties of ϕ_{T_i}. The angular dispersion of $P_{\Omega,\omega,\gamma} f$ is, of course, at most C/R. Averaging over $\Omega \in G$ can increase the support at most by another C/R. Since we are summing over a single ω to obtain $\mathcal{A}_\omega$ the dispersion remains bounded by C/R. From its definition the support of $\eta^{\bar{\xi}_{T_i}}$ is contained in a disk centered at the origin with radius $\lesssim c^2/R$, so convolving with this can only lead to an increase C/R in the angular dispersion. Thus the angular dispersion of ϕ_T is at most C/R. As we see the increase over support of f_i can be at most C/R, hence we also have our margin requirement.

For the well concentration property (2.31), we remark that obtaining this requires some technical work. We first note that

$$T_i \phi_{T_i}(\bar{\xi}, 0) = \widehat{\phi_{T_i}}(\bar{\xi}, 0) = \widehat{\eta^{\bar{\xi}_{T_i}}}(\bar{\xi}) \cdot \widehat{\mathcal{A}_{\omega_{T_i}} f_i}(\bar{\xi})$$

and since,

$$\begin{aligned} \widehat{\mathcal{A}_{\omega_{T_i}} f_i}(\bar{\xi}) = \int \Big(\sum_\gamma \int P_{\Omega,\omega_{T_i},\gamma} f_i(x) d\Omega \Big) e^{-2\pi i x \cdot \bar{\xi}} dx &= \sum_\gamma \int \Big(\int P_{\Omega,\omega_{T_i},\gamma} f_i(x) e^{-2\pi i x \cdot \bar{\xi}} dx \Big) d\Omega \\ &= \sum_\gamma \int \widehat{P_{\Omega,\omega_{T_i}}} f_i(\bar{\xi}) d\Omega, \end{aligned}$$

using the fact that η is a Schwarz function, and the scaling we used to obtain $\eta^{\bar{\xi}_{T_i}}$ we may write

$$
\begin{aligned}
|T_i\phi_{T_i}(\overline{\xi},0)| &\leq \big|\widehat{\eta^{\overline{\xi}_{T_i}}}(\overline{\xi})\big|\big|\sum_\gamma \int \widehat{P_{\Omega,\omega_{T_i}}f_i}(\overline{\xi})d\Omega\big| \\
&\lesssim c^{-C}\widetilde{\chi}^4_{T_i}(\overline{\xi},0)\big|\sum_\gamma \int \widehat{P_{\Omega,\omega_{T_i}}f_i}(\overline{\xi})d\Omega\big|.
\end{aligned}
$$

Hence

$$
\|\phi_{T_i}\|_2 = \|T_i\phi_{T_i}(\cdot,0)\|_2 \lesssim c^{-C}\|\widetilde{\chi}^4_{T_i}(\bar{\cdot},0)\cdot\sum_\gamma \int \widehat{P_{\Omega,\omega_{T_i}}f_i}(\cdot)d\Omega\|_2.
$$

On the right hand side we wish to move along the tube and replace 0 with some appropriate ξ_n. To this end we will prove a statement that will also be of use in proving (2.33). If $f_i \in L^2(U_i')$ is a function of angular dispersion at most C/r centered around vector ω, we have

$$
\|T_if_i\|_{L^2(D(\xi_D,r))} \lesssim C_0^C\|\widetilde{\chi}^5_{D(\xi,r)}T_if_i\|_{L^2(D(\xi,r))} \tag{2.37}
$$

for ξ with $\overline{\xi} = \overline{\xi}_D - (\xi_n - \xi_{D_n})\omega$, and $|\xi_n - \xi_{D_n}| \lesssim R$. We can write, as discussed before

$$
Tf_i(\overline{\xi},\xi_n) = Tf_i(\cdot,\xi_{D_n}) * K^\omega_{\xi_n-\xi_{D_n}}(\cdot)(\overline{\xi}),
$$

with

$$
K^\omega_{\lambda_n}(\overline{\lambda}) = -\int a_\omega(x)e^{-2\pi i(x\cdot\overline{\lambda}+|x|\lambda_n)}dx
$$

where a_ω is a bump function adapted to the support of f_i. We have the strong decay estimate

$$
|K^\omega_{\lambda_n}(\overline{\lambda})| \lesssim C_0^C r^{2-n}(1+|\overline{\lambda}-\lambda_n\omega|/r)^{-N^5}(1+|(\overline{\lambda}-\lambda_n\omega)\cdot\omega|)^{-N^5}
$$

which follows from (2.21), and the nature of the Lorentz scaling that takes a thin strip on the light cone to a wider one. Then an appropriate decomposition around the disk $D(\xi,r)$ gives the desired estimate; see [28], (103). With (2.37) at hand we proceed. We consider the lattice $r\mathbb{Z}^{n-1}$, and a natural partition of $\mathbb{R}^{n-1}$ into disks of radius $2nr$ that one obtains from this lattice. We let $\mathcal{D}$ denote these disks and $D_k(\overline{\xi}_D, 0, 2nr) \in \mathcal{D}$ its elements. Denoting

$$g(x) = \sum_{\gamma} \int P_{\Omega,\omega_{T_i}} f_i(x) d\Omega$$

we have,

$$\begin{aligned}
\|\widetilde{\chi}^4_{T_i}(\bar{\cdot}, 0) T_i g_i(\cdot, 0)\|_2^2 &\leq \sum_{D_k \in \mathcal{D}} c_k^8 \|T_i g_i(\cdot, 0)\|^2_{L^2(D_k)} \\
&\lesssim c^{-C} \sum_{D_k \in \mathcal{D}} c_k^8 \|\widetilde{\chi}^5_{D'_k(\overline{\xi}_D - \xi_n \omega_{T_i}, \xi_n, 2nr)}(\cdot) T_i g_i(\cdot, \xi_n)\|_2^2
\end{aligned}$$

with c_k constants coming from breaking $\widetilde{\chi}_{T_i}$ into disks D_k. Then

$$\lesssim c^{-C} \|\big(\sum_{D_k \in \mathcal{D}} c_k^4 \widetilde{\chi}^5_{D'_k(\overline{\xi}_D - \xi_n \omega_{T_i}, \xi_n, 2nr)}(\cdot)\big) T_i g_i(\cdot, \xi_n)\|_2^2.$$

Observe that

$$\sum_{D_k \in \mathcal{D}} c_k^4 \widetilde{\chi}^5_{D'_k(\overline{\xi}_D - \xi_n \omega_{T_i}, \xi_n, 2nr)}(\cdot) \lesssim \widetilde{\chi}^4_{T_i}(\bar{\cdot}, \xi_n).$$

Hence

$$\lesssim c^{-C} \|\widetilde{\chi}^4_{T_i}(\bar{\cdot}, \xi_n) T_i g_i(\cdot, \xi_n)\|_2^2.$$

As the next step we wish to replace $T_i g_i$ with $T_i f_i$ on the right hand side. To this end observe that if we let

$$S_{\omega_{T_i},\gamma}(x) = \int \chi_{\Omega(Y_{\omega_{T_i}},\gamma)}(x) d\Omega$$

then this S_{T_i} is a compactly supported, smooth function, and

$$\mathcal{A}_{\omega_{T_i}} f(x) = f_i(x) \cdot \sum_{\gamma} S_{\omega_{T_i},\gamma}(x) = g_i(x).$$

Then

$$T_i g_i(\overline{\xi},\xi_n) = \widehat{g_i(x)e^{-2\pi i|x|\cdot\xi_n}}(\overline{\xi}) = -\big[\widehat{f_i(x)e^{-2\pi i|x|\cdot\xi_n}} * \sum_\gamma \widehat{S_{\omega_{T_i},\gamma}}\big](\overline{\xi})$$
$$= -\big[T_i f_i(\cdot,\xi_n) * \sum_\gamma \widehat{S_{\omega_{T_i},\gamma}(\cdot)}\big](\overline{\xi})$$

The Fourier transform $\widehat{S_{\omega_{T_i},\gamma}}$ is, of course, a Schwarz function. We can thus write, using the cutoffs $\widetilde{\chi}_D$ introduced before, and from the properties of the support of $S_{\omega_{T_i}}$,

$$|S_{\omega_{T_i},\gamma}(\overline{\xi})| \lesssim r^{2-n}\widetilde{\chi}_D^3(\overline{\xi}).$$

Then we have

$$\begin{aligned}
\|\widetilde{\chi}_{T_i}^4(\cdot,\xi_n)\cdot T_i g_i(\cdot,\xi_n)\|_2 &= \|\widetilde{\chi}_{T_i}^4(\cdot,\xi_n)\sum_\gamma \big[\widehat{S_{\omega_{T_i},\gamma}}(\overline{\xi}) * T_i f_i(\overline{\xi},\xi_n)\big](\cdot)\|_2 \\
&\le \sum_\gamma \|\widetilde{\chi}_{T_i}^4(\cdot,\xi_n)\big[\widehat{S_{\omega_{T_i},\gamma}}(\overline{\xi}) * T_i f_i(\overline{\xi},\xi_n)\big](\cdot)\|_2 \\
&= \sum_\gamma \|\widetilde{\chi}_{T_i}^4(\cdot,\xi_n)\int \widehat{S_{\omega_{T_i},\gamma}}(\cdot-\overline{\xi})T_i f_i(\overline{\xi},\xi_n)d\overline{\xi}\ \|_2 \\
&\lesssim r^{2-n}\|\widetilde{\chi}_{T_i}^4(\cdot,\xi_n)\int \widetilde{\chi}_D^3(\cdot-\overline{\xi})|T_i f_i(\overline{\xi},\xi_n)|d\overline{\xi}\ \|_2 \\
&\lesssim r^{2-n}\Big(\int \widetilde{\chi}_{T_i}^2(\overline{\zeta},\xi_n)\big(\int \widetilde{\chi}_{T_i}^3(\overline{\zeta},\xi_n)\widetilde{\chi}_D^3(\overline{\zeta}-\overline{\xi})|T_i f_i(\overline{\xi},\xi_n)|d\overline{\xi}\big)^2 d\overline{\zeta}\Big)^{1/2}
\end{aligned}$$

Simply from the definition of functions $\widetilde{\chi}_{T_i},\widetilde{\chi}_D$ we have

$$\widetilde{\chi}_{T_i}^3(\overline{\zeta},\xi_n)\widetilde{\chi}_D^3(\overline{\zeta}-\overline{\xi}) \lesssim \widetilde{\chi}_{T_i}^3(\overline{\xi},\xi_n)$$

Thus we have

$$
\begin{aligned}
&\lesssim r^{2-n}\Big(\int \widetilde{\chi}_{T_i}^2(\overline{\zeta},\xi_n)d\overline{\zeta}\Big)^{1/2}\Big(\int \widetilde{\chi}_{T_i}^3(\overline{\xi},\xi_n)|T_if_i(\overline{\xi},\xi_n)|d\overline{\xi}\Big)\\
&\lesssim r^{2-n+\frac{n-1}{2}}\Big(\int \widetilde{\chi}_{T_i}^3(\overline{\xi},\xi_n)|T_if_i(\overline{\xi},\xi_n)|d\overline{\xi}\Big)\\
&\lesssim r^{2-n+\frac{n-1}{2}}\|\widetilde{\chi}_{T_i}(\cdot,\xi_n)\|_2\|\widetilde{\chi}_{T_i}(\cdot,\xi_n)T_if_i(\cdot,\xi_n)\|_2\\
&\lesssim r\|\widetilde{\chi}_{T_i}(\cdot,\xi_n)T_if_i(\cdot,\xi_n)\|_2.
\end{aligned}
$$

Combining all of these together we obtain (2.31).

We now prove (2.32). This is simpler compared to (2.31). Consider the disk $D(\overline{\xi_Q},0,\frac{C_0}{C}R)\subset\mathbb{R}^{n-1}\times\{0\}$. We have

$$
\begin{aligned}
|T_i\phi_{T_i}(\overline{\xi},\xi_n)| = |T_i\phi_{T_i}(\cdot,0)*K_{\xi_n}(\cdot)(\overline{\xi})| &= |\int T_i\phi_{T_i}(\overline{\lambda},0)K_{\xi_n}(\overline{\xi}-\overline{\lambda})d\overline{\lambda}|\\
&\le |\int_D T_i\phi_{T_i}(\overline{\lambda},0)K_{\xi_n}(\overline{\xi}-\overline{\lambda})d\overline{\lambda}|\\
&+ |\int_{D^{ext}} T_i\phi_{T_i}(\overline{\lambda},0)K_{\xi_n}(\overline{\xi}-\overline{\lambda})d\overline{\lambda}|\\
&= I + II.
\end{aligned}
$$

We evaluate I:

$$
I \le \|T_i\phi_{T_i}(\cdot,0)\|_{L^2(D)}\|K_{\xi_n}(\overline{\xi}-\cdot)\|_{L^2(D)}
$$

Using the trivial estimate $\|K_{\xi_n}\|_\infty \le C$ and the size of radius of D we have

$$
\lesssim R^{5n}\|T_i\phi_{T_i}(\cdot,0)\|_{L^2(D)}.
$$

Then from the simple estimates

$$
\|\widehat{\eta^{\overline{\xi}_{T_i}}}\|_{L^2(D)} \lesssim R^{-N},\quad \|\widehat{\mathcal{A}_{\omega_{T_i}}f}\|_{L^2(D)} \lesssim \|f_i\|_2
$$

we obtain an acceptable bound for I. For II we have

$$II \leq \|T_i\phi_{T_i}(\cdot,0)\|_{L^2(D^{ext})}\|K_{\xi_n}(\overline{\xi}-\cdot)\|_{L^2(D^{ext})}$$

Since the distance of $\overline{\xi}$ to D^{ext} is greater than $\frac{C_0}{C}R$, we have

$$\|K_{\xi_n}(\overline{\xi}-\cdot)\|_{L^2(D^{ext})} \lesssim R^{-N}.$$

The first term clearly satisfies

$$\|T_i\phi_{T_i}(\cdot,0)\|_{L^2(D^{ext})} \lesssim \|f_i\|_2.$$

Thus we obtain (2.32).

To prove (2.33) we utilize (2.37). Fix a cube q centered at ξ_q. We consider $D(\overline{\xi}_q,\xi_n,Cr)$ with $|\xi_n-\xi_{q_n}| \leq Cr$. Then we have

$$\|T_i\phi_{T_i}(\cdot,\xi_n)\|_{L^2(D(\overline{\xi}_q,\xi_n,Cr))} \lesssim c^{-C}\|\widetilde{\chi}^5_{D(\overline{\xi}'_q,0,Cr)}T_i\phi_{T_i}(\cdot,0)\|_2$$

with $\overline{\xi}'_q = \overline{\xi}_q - \xi_{q_n}\omega_{T_i}$, since ξ_n changes an amount at most Cr. Thus squaring and integrating for ξ_n we have

$$\|T_i\phi_{T_i}(\cdot,\xi_n)\|^2_{L^2(Cq)} \lesssim c^{-C}r\|\widetilde{\chi}^5_{D(\overline{\xi}'_q,0,Cr)}T_i\phi_{T_i}(\cdot,0)\|^2_2.$$

So the left hand side of (2.33) can be majorized by

$$c^{-C}r\sum_{T_i\in\mathbf{T}_i}\sup_{q\in K_J(Q)}\widetilde{\chi}^{-3}_{T_i}(\xi_q)\|\widetilde{\chi}^5_{D(\overline{\xi}'_q,0,Cr)}T_i\phi_{T_i}(\cdot,0)\|^2_2.$$

which satisfies,

$$\lesssim c^{-C}r\sum_{T_i\in\mathbf{T}_i}\sup_{q\in K_J(Q)}\widetilde{\chi}^{-3}_{T_i}(\xi_q)\cdot\|\widetilde{\chi}^3_{D(\overline{\xi}'_q,0,Cr)}(\cdot)\widetilde{\chi}^4_{T_i}(\cdot,0)\sum_{\gamma}\int\widehat{P_{\Omega,\omega_{T_i}}f_i}(\cdot)d\Omega\|^2_2.$$

Observe the simple fact

$$\widetilde{\chi}_{D(\bar{\xi}'_q,0,Cr)}(\bar{\xi})\widetilde{\chi}_{T_i}(\bar{\xi},0) \lesssim \widetilde{\chi}_{T_i}(\xi_q),$$

Thus we are left with

$$\lesssim c^{-C} r \sum_{T_i \in \mathbf{T}_i} \|\widetilde{\chi}_{T_i}(\cdot,0)| \sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\cdot) d\Omega|\|_2^2.$$

But this summation satisfies

$$\begin{aligned}
= \int \sum_{T_i \in \mathbf{T}_i} \widetilde{\chi}^2_{T_i}(\bar{\xi},0)|\sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\bar{\xi}) d\Omega|^2 d\bar{\xi} &= \int \sum_{\omega_{T_i}} \sum_{\bar{\xi}_{T_i}} \widetilde{\chi}^2_{\bar{\xi}_{T_i}}(\bar{\xi},0)|\sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\bar{\xi}) d\Omega|^2 d\bar{\xi} \\
&= \int \sum_{\bar{\xi}_{T_i}} \widetilde{\chi}^2_{\bar{\xi}_{T_i}}(\bar{\xi},0) \sum_{\omega_{T_i}} |\sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\bar{\xi}) d\Omega|^2 d\bar{\xi} \\
&\lesssim \int \sum_{\omega_{T_i}} |\sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\bar{\xi}) d\Omega|^2 d\bar{\xi} \\
&\lesssim \sum_{\omega_{T_i}} \|\sum_{\gamma} \int \widehat{P_{\Omega,\omega_{T_i}} f}_i(\cdot) d\Omega\|_2^2 \\
&\lesssim \sum_{\omega_{T_i}} \|\sum_{\gamma} \int P_{\Omega,\omega_{T_i}} f_i(\cdot) d\Omega\|_2^2.
\end{aligned}$$

From here we use that supports of only a finite number of $P_{\Omega,\omega_{T_i}} f_i$ can intersect to obtain

$$\lesssim \|f_i\|_2^2.$$

We finally prove (2.34). We expand the left hand side as

$$\Big(\sum_{q_0} \|\sum_{T_i \in \mathbf{T}_i} m_{q_0,T_i} \eta^{\bar{\xi}_{T_i}} * \mathcal{A}_{\omega_{T_i}} f_i\|_2^2\Big)^{1/2} = \Big(\sum_{q_0} \|\int \sum_{\bar{\xi} \in L} \sum_{\omega \in \mathcal{Y}_i} \sum_{\gamma} m_{q_0,\omega,\bar{\xi}} \widehat{\eta^{\bar{\xi}}} \widehat{P_{\Omega,\omega,\gamma} f}_i(\bar{\xi}) d\Omega\|_2^2\Big)^{1/2}$$

We apply the Minkowski integral inequality to bring the L^2 norm inside:

$$= \Big(\sum_{q_0} (\int \| \sum_{\bar{\xi} \in L} \sum_{\omega \in \mathcal{Y}_i} \sum_{\gamma} m_{q_0,\omega,\bar{\xi}} \widehat{\eta^{\bar{\xi}}} \widehat{P_{\Omega,\omega,\gamma}} f_i(\bar{\xi}) \|_2 d\Omega)^2 \Big)^{1/2}.$$

We apply once more the Minkowski integral inequality to change the order of summation with that of integration

$$= \int \Big(\sum_{q_0} \| \sum_{\bar{\xi} \in L} \sum_{\omega \in \mathcal{Y}_i} \sum_{\gamma} m_{q_0,\omega,\bar{\xi}} \widehat{\eta^{\bar{\xi}}} \widehat{P_{\Omega,\omega,\gamma}} f_i(\bar{\xi}) \|_2^2 \Big)^{1/2} d\Omega.$$

At this stage we define the set W_ω for each $\omega \in \mathcal{Y}_i$ as the set of $x \in Y_\omega$ with a distance at least Cc^2/r to $Y_{\omega'}$ for $\omega' \neq \omega$. Let

$$W = \bigcup_{\omega} W_\omega, \quad W^c = \bigcup_{\omega} Y_\omega \setminus W_\omega.$$

We define

$$P_{\Omega,W_\omega} f_i(x) := \chi_{\Omega(W_\omega)}(x) f_i(x), \quad P_{\Omega,W} f_i(x) := \chi_{\Omega(W)}(x) f_i(x)$$

$$(1 - P_{\Omega,W}) f_i(x) = (1 - \chi_{\Omega(W)})(x) f_i(x).$$

Thus we have

$$\sum_{\omega \in \mathcal{Y}_i} \sum_{\gamma \in \mathcal{G}} P_{\Omega,\omega,\gamma} f_i(x) = f_i(x) = P_{\Omega,W} f_i(x) + (1 - P_{\Omega,W}) f_i(x),$$

and from linearity of the Fourier transform

$$\sum_{\omega \in \mathcal{Y}_i} \sum_{\gamma \in \mathcal{G}} \widehat{P_{\Omega,\omega,\gamma}} f_i(\xi) = \widehat{P_{\Omega,W}} f_i(\xi) + (1 - \widehat{P_{\Omega,W}}) f_i(\xi).$$

We use this decomposition, and the triangle inequality, first for the norm inside, and then for the outer summation, to obtain

$$= \int \Big(\sum_{q_0} \| \sum_{\overline{\xi}\in L} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} \widehat{P_{\Omega,W} f_i}(\overline{\xi}) \|_2^2 \Big)^{\frac{1}{2}} d\Omega + \int \Big(\sum_{q_0} \| \sum_{\overline{\xi}\in L} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} (1 - \widehat{P_{\Omega,W}}) f_i \|_2^2 \Big)^{\frac{1}{2}} d\Omega.$$

For the first term we use the fact that the sets W_ω are at least Cc^2/r separated, so supports of $\eta^{\overline{\xi}} * P_{\Omega,W_\omega} f_i$ are disjoint. Using this and the Plancherel theorem we have

$$= \int \Big(\sum_{q_0} \sum_{\omega\in\mathcal{Y}_i} \| \sum_{\overline{\xi}\in L} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} \widehat{P_{\Omega,W_\omega} f_i} \|_2^2 \Big)^{\frac{1}{2}} d\Omega.$$

We rearrange this as

$$= \int \Big(\sum_{\omega\in\mathcal{Y}_i} \int |\widehat{P_{\Omega,W_\omega} f_i}(\overline{\zeta})|^2 \sum_{q_0} \big(\sum_{\overline{\xi}\in L} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} \big)^2 d\overline{\zeta} \Big)^{\frac{1}{2}} d\Omega$$

$$= \int \Big(\sum_{\omega\in\mathcal{Y}_i} \int |\widehat{P_{\Omega,W_\omega} f_i}(\overline{\zeta})|^2 \big(\sum_{\overline{\xi}\in L} \sum_{q_0} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} \big)^2 d\overline{\zeta} \Big)^{\frac{1}{2}} d\Omega.$$

Summing first over q_0, then over $\overline{\xi}$, and then using disjointness of supports gives

$$= \int \Big(\sum_{\omega\in\mathcal{Y}_i} \int |\widehat{P_{\Omega,W_\omega} f_i}(\overline{\zeta})|^2 d\overline{\zeta} \Big)^{\frac{1}{2}} d\Omega = \int \Big(\sum_{\omega\in\mathcal{Y}_i} \| P_{\Omega,W_\omega} f_i \|_2^2 \Big)^{\frac{1}{2}} d\Omega = \int \| P_{\Omega,W} f_i \|_2 d\Omega$$
$$\leq \| f_i \|_2.$$

We now turn the second term. This time supports are not disjoint, but they are almost disjoint in the sense that only a constant C of them can intersect. Thus by grouping appropriately we obtain

$$\lesssim \int \Big(\sum_{q_0} \sum_{\omega\in\mathcal{Y}_i} \| \sum_{\overline{\xi}\in L} m_{q_0,\omega,\overline{\xi}} \widehat{\eta^{\overline{\xi}}} (1 - \widehat{P_{\Omega,W_\omega}}) f_i \|_2^2 \Big)^{\frac{1}{2}} d\Omega.$$

From this point on exactly the same process reduces this to

$$\lesssim \int \|(1-P_{\Omega,W})f_i\|_2 d\Omega.$$

Applying the Cauchy-Schwarz inequality gives

$$\lesssim \Big(\int \|(1-P_{\Omega,W})f_i\|_2^2 d\Omega\Big)^{1/2} = \Big(\int |f_i(x)|^2\Big(\int |\chi_{\Omega,W^c}(x)| d\Omega\ dx\Big)^{1/2}.$$

The inner integral satisfies $\lesssim c^2$, so we have

$$\leq Cc\|f_i\|_2.$$

We thus obtain (2.34).

■

Thus we complete the first stage in the proof of the main proposition. The following lemma comprises the second stage. It uses the wave packet decomposition to divide f_i into functions whose images under T_i are well localized to subcubes. The lemma proves three facts about these functions. The first is about their margins and follows directly from Lemma 8. The second is about total mass of these functions, and again easily follows from Lemma 8. Th final one, the well-approximation property that will be stated below follows as a consequence of localization, and is the most important part of the lemma.

Lemma 9 *Let $R \geq C_0 2^{C_1}$, and $0 < c \leq 2^{-C_0}$. Let $Q \subset \mathbb{R}^n$ be a cube of side-length R centered at ξ_Q. Let $f_1 \in L^2(U_1)$ with margin requirement $margin(f_i) \gtrsim R^{-1/2}$, and let $f_2 \in L^2(U_2)$. Then for every $q_0 \in K_{C_0}(Q)$ we have a function $F_1^{(q_0)}$ depending on f_1, f_2, q_0 satisfying the following conditions:*

- *The margin requirement*

$$margin(F_1^{(q_0)}) \geq margin(f_1) - CR^{-1/2}.$$

• *Requirement on the total mass*

$$\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2 \leq (1 + Cc)\|f_1\|_2^2. \tag{2.38}$$

• *The function Φ_1 defined by*

$$\Phi_1 := \sum_{q_0 \in K_{C_0}(Q)} |F_1^{(q_0)}| \chi_{q_0}$$

approximates $T_1 f_1$ well

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^2(I^{c, C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4}} \|f_1\|_2 \|f_2\|_2. \tag{2.39}$$

Proof. We first note that if c is extremely small, such as $0 \leq c \leq R^{-\frac{1}{10n}}$ we may just set $F_1^{(q_0)} = 0$ for each q_0, and obtain all properties trivially. Thus we may assume $c > R^{-\frac{1}{10n}}$. As mentioned we will utilize Lemma 8 greatly. We define $F_1^{q_0}$ by

$$F_1^{(q_0)} := \sum_{T_1 \in \mathbf{T}_1} \frac{m_{q_0, T_1}}{m_{T_1}} \phi_{T_1}$$

with

$$m_{q_0, T_1} := \|T_2 f_2 \widetilde{\chi}_{T_1}\|_{L^2(q_0)}^2 + R^{-10n} \|f_2\|_2$$

and

$$m_{T_1} := \sum_{T_1 \in \mathbf{T}_1} m_{q_0, T_1} = \|T_2 f_2 \widetilde{\chi}_{T_1}\|_{L^2(Q)}^2 + R^{-10n} 2^{nC_0} \|f_2\|_2.$$

In the definition of m_{q_0, T_1} the second term is so small as to be dealt with as an error term, and has the function of ensuring that these coefficients are non-zero. As the first term makes clear if $T_2 f_2$ concentrates on a cube q_0 that lies on T_i, we take a larger coefficient for the part of ϕ_{T_1} given to that cube to ensure that (2.39) holds. With this definition we of course have

$$f_1 = \sum_{q_0 \in K_{C_0}(Q)} F_1^{(q_0)}.$$

Lemma 8 makes it clear that the margin requirement, and the requirement on the total mass hold.

Thus it remains to prove (2.39). We have

$$\begin{aligned} |(T_1 f_1 - \Phi_1) T_2 f_2| &= |\big(\sum_{q_0 \in K_{C_0}(Q)} F_1^{(q_0)} - \sum_{q_0 \in K_{C_0}(Q)} F_1^{(q_0)} \chi_{q_0} \big) T_2 f_2| \\ &= \sum_{q_0 \in K_{C_0}(Q)} |F_1^{(q_0)} T_2 f_2| (1 - \chi_{q_0}), \end{aligned} \tag{2.40}$$

thus from our assumption that $R^{\frac{1}{10n}} < c \leq 2^{C_0}$, and the triangle inequality it is sufficient to show that

$$\|T_1 F_1^{(q_0)} T_2 f_2\|_{L^2(I^{c,C_0}(Q) \setminus q_0)} \lesssim c^{-C} R^{\frac{2-n}{4}} \|f_1\|_2 \|f_2\|_2 \tag{2.41}$$

for each $q_0 \in K_{C_0}(Q)$. To prove this we further decompose the cube Q into cubes $q \in K_J(Q)$, where J is an integer such that $2^J \leq R^{1/2} < 2^{J+1}$. Then it suffices to prove that

$$\sum_{q \in K_J(Q): dist(q,q_0) \gtrsim C_0^{-1} cR} \|T_1 F_1^{(q_0)} T_2 f_2\|_{L^2(q)}^2 \lesssim c^{-C} R^{\frac{2-n}{2}} \|f_1\|_2^2 \|f_2\|_2^2. \tag{2.42}$$

Take an indivudual term from this sum, we have from the definition of $F_1^{(q_0)}$ and the triangle inequality

$$\|T_1 F_1^{(q_0)} T_2 f_2\|_{L^2(q)} \leq \sum_{T_1 \in \mathbf{T}_1} \frac{m_{q_0,T_1}}{m_{T_1}} \|T_1 \phi_{T_1} T_2 f_2\|_{L^2(q)}. \tag{2.43}$$

We will first remove from this sum tubes that contribute insignificant amounts: consider tubes that does not intersect $C_0 Q$. By (2.32) for such tubes we have

$$\|T_1 \phi_{T_1} T_2 f_2\|_{L^2(q)} \leq \|T_1 \phi_{T_1}\|_{L^\infty(q)} \|T_2 f_2\|_{L^2(q)} \lesssim \text{dist}(q,T)^{C-N} \|f_1\|_2 \cdot r^{1/2} \|f_2\|_2.$$

The extremely fast decay depending on the distance to the cube allows us to bound the total contribution of these tubes by $CR^{C-N}\|f_1\|_2\|f_2\|_2$. Thus it remains to estimate the tubes that do intersect the cube. In this case it is important to remember that there is at most R^C such tubes. The Lemma 5 gives for one such tube

$$\|T_1\phi_{T_1}T_2f_2\|_{L^2(q)} \lesssim r^{-\frac{n}{2}}\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2\|T_2f_2\widehat{\phi}_r^q\|_2$$

with $\widehat{\phi}_r^q$ are as described in Lemma 5. Putting together the information we gathered on individual terms of the (2.42) we have

$$\|T_1F_1^{(q_0)}T_2f_2\|_{L^2(q)} \lesssim r^{-\frac{n}{2}}\|T_2f_2\widehat{\phi}_r^q\|_2\sum_{T_1\in\mathbf{T}_1}\frac{m_{q_0,T_1}}{m_{T_1}}\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2 + R^{C-N}\|f_1\|_2\|f_2\|_2. \tag{2.44}$$

Since there is at most R^C cubes q, we see that proving (2.42) reduces to proving

$$\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\|T_2f_2\widehat{\phi}_r^q\|_2^2\Big(\sum_{T_1\in\mathbf{T}_1}\frac{m_{q_0,T_1}}{m_{T_1}}\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2\Big)^2 \lesssim c^{-C}r^2\|f_1\|_2^2\|f_2\|_2^2. \tag{2.45}$$

At this point we have a gain of $r^{-\frac{n}{2}}$ arising from narrowness of supports of functions ϕ_{T_1}. The rest of the proof will use a geometric argument to estimate the sum in paranthesis to preserve this gain. We will make use of the fact that the expression $\|T_1\phi_{T_1}\|_{L^2(Cq)}$ is extremely small if the tube T_1 does not intersect Cq. Thus in a sense we can write

$$\sum_{T_1\in\mathbf{T}_1}\frac{m_{q_0,T_1}}{m_{T_1}}\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2 \approx \sum_{T_1\in\mathbf{T}_1}\frac{m_{q_0,T_1}}{m_{T_1}}\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2\chi_{T_1}(\xi_q)$$

where ξ_q stands for the center of q. From here using the Cauchy-Schwarz inequality we sum the terms

$$m_{q_0,T_1}\chi_{T_1}(\xi_q)$$

separately, and these terms are nonzero only for T_1 passing through a cube Cq. But such tubes form a cone neighborhood, which allows us to use Lemma 4 to obtain a good estimate. We now make these ideas rigorous. First observe the trivial inequality

$$\frac{m_{q_0,T_1}}{m_{T_1}} \leq \frac{m_{q_0,T_1}^{1/2}}{m_{T_1}^{1/2}}.$$

This together with inserting the smooth cutoff we defined for tubes to the summation gives

$$\sum_{T_1\in\mathbf{T}_1} \frac{m_{q_0,T_1}}{m_{T_1}} \|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2 \leq \sum_{T_1\in\mathbf{T}_1} \frac{m_{q_0,T_1}^{1/2}}{m_{T_1}^{1/2}} \widetilde{\chi}_{T_1}^{1/2}(\xi_q)\widetilde{\chi}_{T_1}^{-1/2}(\xi_q)\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2.$$

We now use the Cauchy-Schwarz inequality to do separation we mentioned:

$$\leq \sum_{T_1\in\mathbf{T}_1} \frac{\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2^2}{m_{T_1}} \widetilde{\chi}_{T_1}^{-1}(\xi_q) \sum_{T_1\in\mathbf{T}_1} m_{q_0,T_1}\widetilde{\chi}_{T_1}(\xi_q).$$

Observe that

$$\sum_{T_1\in\mathbf{T}_1} m_{q_0,T_1}\widetilde{\chi}_{T_1}(\xi_q) = \sum_{T_1\in\mathbf{T}_1} \|T_2 f_2\widetilde{\chi}_{T_1}\|_{L^2(q_0)}^2\widetilde{\chi}_{T_1}(\xi_q) + \sum_{T_1\in\mathbf{T}_1} R^{-10n}\|f_2\|_2\widetilde{\chi}_{T_1}(\xi_q)$$

$$\leq \|T_2 f_2\Big(\sum_{T_1\in\mathbf{T}_1} \widetilde{\chi}_{T_1}(\xi_q)\widetilde{\chi}_{T_1}^2\Big)^{1/2}\widetilde{\chi}_{q_0}\|_2^2 + R^{-10n}\|f_2\|_2 \sum_{T_1\in\mathbf{T}_1} \widetilde{\chi}_{T_1}(\xi_q)$$

Using very fast decay of $\widetilde{\chi}_{T_1}$ and the fact that the number of tubes T_1 passing through a point ξ_q is bounded by $CR^{(n-2)/2}$, we can bound the second sum by $R^{-5n}\|f_2\|_2$. For the first sum let

$$\chi := \Big(\sum_{T_1\in\mathbf{T}_1} \widetilde{\chi}_{T_1}(\xi_q)\widetilde{\chi}_{T_1}^2\Big)^{1/2}\widetilde{\chi}_{q_0}.$$

and observe that this function represents the part of smooth cutoffs of tubes passing through the point ξ lying inside the cube q_0. But the sum of such smooth cutoffs will be bounded by an appropriate multiple of a smooth cutoff of a cone neighborhood. Since $dist(q_0,\xi_q) \gtrsim C_0^{-1}cR$, at most c^{-C} of tubes passing through ξ_q can intersect in q_0. Since smooth cutoffs $\widetilde{\chi}_{T_1}$ decay very fast the contribution coming from tails can be bounded by another c^{-C}. Thus we have

$$\chi(\xi) \lesssim c^{-C}\big(1 + \frac{dist(\xi, C^1(\xi_q))}{r}\big)^{-10n}.$$

We will perform a dyadic decomposition into shells to exploit this decay. Let $C^1_k(\xi_q)$ be the set of all ξ with $2^{k-1} \leq \mathrm{dist}(\xi, C^1(\xi_q)) < 2^k$ for k a natural number. Then

$$\|T_2 f_2 \chi\|_2^2 = \|T_2 f_2 \chi\|^2_{L^2(C^1(\xi_q, r))} + \sum_{k \in \mathbb{N}} \|T_2 f_2 \chi\|^2_{L^2(C^1_k(\xi_q))}.$$

To the first term on the right hand side we apply Lemma 4 to bound it by

$$\lesssim c^{-C} r \|f_2\|_2^2.$$

For terms in the summation we will apply the same Lemma as well. We have

$$\begin{aligned} \|T_2 f_2 \chi\|^2_{L^2(C^1_k(\xi_q))} &\lesssim c^{-C} 2^{10n(k-1)} \|T_2 f_2\|^2_{L^2(C^1_k(\xi_q))} \\ &\lesssim c^{-C} 2^{10n(k-1)} \|T_2 f_2\|^2_{L^2(C^1(\xi_q, 2^k r))} \\ &\lesssim c^{-C} 2^{-10n(k-1)} 2^k r \|f_2\|_2^2 \end{aligned}$$

So from the exponential decay of the coefficients

$$\|T_2 f_2 \chi\|_2^2 \lesssim c^{-C} r \|f_2\|_2^2 + c^{-C} r \|f_2\|_2^2 \sum_{k \in \mathbb{N}} 2^{-10n(k-1)} 2^k \lesssim c^{-C} r \|f_2\|_2^2.$$

We insert this information back to (2.45) to see that it reduces to proving

$$\sum_{q \in K_J(Q) : dist(q, q_0) \gtrsim C_0^{-1} cR} \|T_2 f_2 \widehat{\phi}^q_r\|_2^2 \sum_{T_1 \in \mathbf{T}_1} \frac{\|T_1 \phi_{T_1} \widehat{\phi}^q_r\|_2^2}{m_{T_1} \widetilde{\chi}_{T_1}(\xi_q)} \lesssim c^{-C} r^2 \|f_1\|_2^2.$$

We can easily rearrange the left-hand side as

$$\sum_{T_1\in\mathbf{T}_1}\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\widetilde{\chi}_{T_1}^{-3}(\xi_q)\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2^2\frac{\|T_2f_2\widehat{\phi}_r^q\|_2^2\widetilde{\chi}_{T_1}^2(\xi_q)}{m_{T_1}}$$

which can be bounded by

$$\sum_{T_1\in\mathbf{T}_1}\sup_{\substack{q\in K_J(Q)\\ dist(q,q_0)\gtrsim C_0^{-1}cR}}\left(\widetilde{\chi}_{T_1}^{-3}(\xi_q)\|T_1\phi_{T_1}\widehat{\phi}_r^q\|_2^2\right)\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\frac{\|T_2f_2\widehat{\phi}_r^q\|_2^2\widetilde{\chi}_{T_1}^2(\xi_q)}{m_{T_1}}. \tag{2.46}$$

Observe that

$$\begin{aligned}\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\|T_2f_2\widehat{\phi}_r^q\|_2^2\widetilde{\chi}_{T_1}^2(\xi_q)\leq \|T_2f_2&\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\widehat{\phi}_r^q\widetilde{\chi}_{T_1}(\xi_q)\|_{L^2(Q)}^2\\ +\|T_2f_2&\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\widehat{\phi}_r^q\widetilde{\chi}_{T_1}(\xi_q)\|_{L^2(\mathbb{R}^n\setminus Q)}^2.\end{aligned}$$

We treat these two terms differently. For the first one note that

$$|\widehat{\phi}_r^q(\xi)|^{1/2}\widetilde{\chi}_{T_1}(\xi_q)\lesssim\widetilde{\chi}_{T_1}(\xi)$$

and

$$\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}|\widehat{\phi}_r^q|^{1/2}\lesssim 1$$

from the Schwarz decay of $\widehat{\phi}_r^q$. So the first term satisfies

$$\|T_2f_2\sum_{q\in K_J(Q):dist(q,q_0)\gtrsim C_0^{-1}cR}\widehat{\phi}_r^q\widetilde{\chi}_{T_1}(\xi_q)\|_{L^2(Q)}^2\lesssim\|T_2f_2\widetilde{\chi}_{T_1}\|_{L^2(Q)}^2.$$

As for the second term from the very fast decay of $\widehat{\phi}_r^q$, and the distance of small cubes q to $\mathbb{R}^n\setminus Q$ we have

$$\lesssim R^{-N}\|f_2\|_2.$$

Hence the inner sum is always bounded by some constant C, and thus (2.46) satisfies

$$\sum_{T_1 \in \mathbf{T}_1} \sup_{\substack{q \in K_J(Q) \\ dist(q,q_0) \gtrsim C_0^{-1} cR}} \left(\widetilde{\chi}_{T_1}^{-3}(\xi_q) \| T_1 \phi_{T_1} \widehat{\phi}_r^q \|_{L^2}^2 \right).$$

To this last term we can apply (2.33). We recall that our T_i intersects $C_0 Q$, thus the number of such T_i is bounded by R^C, and $\widetilde{\chi}_{T_1}(\xi_q)$ cannot be very small. This way we majorize this term by

$$\lesssim \sum_{T_1 \in \mathbf{T}_1} \sup_{q \in K_J(Q)} \left(\widetilde{\chi}_{T_1}^{-3}(\xi_q) \| T_1 \phi_{T_1} \widehat{\phi}_r^q \|_{L^2(Q)}^2 \right) + R^{-N} \| f_1 \|_2.$$

The summation here in turn satisfies

$$\sum_{T_1 \in \mathbf{T}_1} \sup_{q \in K_J(Q)} \left(\widetilde{\chi}_{T_1}^{-3}(\xi_q) \| T_1 \phi_{T_1} \|_{L^2(q)}^2 \right).$$

So from (2.33) we are done.

■

Finally we arrived at the third part, in which we state and prove our main proposition. The proposition obtains functions localized to subcubes from f_1, f_2, and then goes on to prove a margin requirement, a total mass requirement, a persistence of concentration property and two different well approximation properties for them. All of these, with the exception of persistence of concentration property which we start from scratch, follows from the previous lemma and some additional arguments.

Proposition 1 *Let $R \geq C_0 2^{C_1}$, and $0 < c \leq 2^{-C_0}$. Let $Q \subset \mathbb{R}^n$ be a cube of side-length CR centered at ξ_Q. Let $f_i \in L^2(U_i)$, $i = 1, 2$ be functions with the normalization $\|f_1\|_2 = \|f_2\|_2 = 1$, and the margin requirement $margin(f_i) \gtrsim 1/C_0 - 2R^{-1/N}$. Then for every $q_0 \in K_{C_0}(Q)$ we have functions $F_1^{(q_0)}$, $F_1^{(q_0)}$ both depending on f_1, f_2, q_0, Q satisfying the following conditions:*

- *The margin requirement*

$$margin(F_1^{(q_0)}), margin(F_2^{(q_0)}) \geq 1/C_0 - 3R^{-1/N}. \tag{2.47}$$

- *Requirement on the total mass*

$$\sum_{q_0 \in K_{C_0}(Q)} \|F_i^{(q_0)}\|_2^2 \leq (1 + Cc), \quad i = 1, 2. \tag{2.48}$$

- *If the pair f_1, f_2 enjoys good concentration properties, these persist for $F_1^{(q_0)}, F_2^{(q_0)}$,*

$$E_{\rho(1-C_0\rho^{-1/3N}),C_0Q}(F_1^{(q_0)}, q_0 \in K_{C_0}(Q); F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) \leq E_{\rho,C_0Q}(f_1, f_2) + Cc \tag{2.49}$$

for $R^{1/2+3/N} \lesssim \rho \lesssim R^2$.

- *The product of functions Φ_1, Φ_2 defined by*

$$\Phi_i := \sum_{q_0 \in K_{C_0}(Q)} |F_i^{(q_0)}| \chi_{q_0}, \quad i = 1, 2$$

approximates $T_1 f_1 T_2 f_2$ well

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q))} \leq \|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q))} + c^{-C}. \tag{2.50}$$

- *For appropriate cone neighborhoods the following improved version holds*

$$\begin{aligned} \|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q)) \cap C^3(\xi_C, r))} \leq& \|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q)) \cap C^3(\xi_C, r)} \\ &+ c^{-C}(1 + R/r)^{-\epsilon/4}. \end{aligned} \tag{2.51}$$

Proof. We define the functions $F_1^{(q_0)}$ just as in Lemma 9. The requirement on the total mass, then, follows from the definition. We have the desired margin estimates on these from the following computation that makes use of margin properties shown in Lemma 9

$$margin(F_1^{(q_0)}) \geq margin(f_1) - CR^{-1/2} \geq 1/C_0 - 2R^{-1/N} - CR^{-1/2}$$
$$\geq 1/C_0 - 3(1/R)^{1/N}.$$

Now we define $F_2^{(q_0)}$, and this involves a twist. Let ϕ_{T_2}, $T_2 \in \mathbf{T}_2$ denote wave packets coming from decomposition of f_2 according to Lemma 8. Then

$$F_2^{(q_0)} := \sum_{T_2 \in \mathbf{T}_2} \frac{m_{q_0,T_2}}{m_{T_2}} \phi_{T_2}$$

with

$$m_{q_0,T_2} := \sum_{q_1 \in K_{C_0}(Q)} \|T_1 F_1^{(q_1)} \widetilde{\chi}_{T_2}\|^2_{L^2(q_0)} + R^{-10n} \sum_{q_1 \in K_{C_0}(Q)} R^{-10n} \|F_1^{(q_1)}\|_2^2$$

and

$$m_{T_1} := \sum_{q_0 \in K_{C_0}(Q)} m_{q_0,T_1} = \sum_{q_1 \in K_{C_0}(Q)} \|T_1 F_1^{(q_1)} \widetilde{\chi}_{T_1}\|^2_{L^2(Q)} + R^{-10n} 2^{nC_0} \sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_1)}\|_2^2.$$

Note that main diffence in the definition of $F_2^{q_0}$ is in coefficients and not wave packets used, thus margin and total mass estimates will follow similarly. From Lemma 8 the tubes ϕ_{T_2} satisfy the margin requirement

$$margin(\phi_{T_2}) \geq margin(f_2) - CR^{-1/2},$$

thus $F_2^{q_0}$ satisfy the margin requirement by the same calculation above. Total mass requirement directly follows from (2.34).

So we proceed to prove (2.49). We fix $\rho \gtrsim R^{1/2+3/N}$, and pick γ with $\gamma(1 - C_0\gamma^{-1/2N}) = \rho(1 - C_0\rho^{-1/3N})$. It is clear that such a γ exists: we may think of the left-hand side as a function, and apply the intermediate value theorem. This γ also satisfies $\gamma \gtrsim R^{1/2+1/N}$, and $\gamma \leq \rho(1 - C_0\rho^{-1/2N})$. The first of these can easily be shown: $\gamma \geq \rho(1 - C_0\rho^{-1/3N}) \geq$

$\rho/2 \gtrsim R^{1/2+1/N}$. For the second, we need to observe that $\rho > \gamma$, for $\rho \le \gamma$ would mean $\gamma^{-1/2N} \le \rho^{-1/2N} < \rho^{-1/3N}$. This would clearly contradict with the equality we used to define γ. Since ρ, γ are extremely large numbers, we have

$$\rho^{1-1/2N} + \gamma^{1-1/2N} < 2\rho^{1-1/2N} < \rho^{1-1/3N}.$$

We already have from the equality defining γ

$$\gamma = \rho - C_0\rho^{1-1/3N} + C_0\gamma^{1-1/2N}.$$

Combining these two gives the desired result. This new quantity γ and its properties will be used below to divide the proof of (2.49) into two similar stages. We first will prove

$$E_{\rho(1-C_0\rho^{-1/3N}),C_0Q}(F_1^{(q_0)}, q_0 \in K_{C_0}(Q); F_2^{(q_0)}, q_0 \in K_{C_0}(Q))$$

$$= E_{\gamma(1-C_0\gamma^{-1/2N}),C_0Q}(F_1^{(q_0)}, q_0 \in K_{C_0}(Q); F_2^{(q_0)}, q_0 \in K_{C_0}(Q))$$

$$\le E_{\gamma,C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) + Cc,$$

and in the second stage

$$\begin{aligned} E_{\gamma,C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) &\le E_{\rho(1-C_0\rho^{-1/2N}),C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) \\ &\le E_{\rho,C_0Q}(f_1; f_2) + Cc. \end{aligned}$$

We start the first part. Let $D = D(\xi, \gamma(1 - C_0\gamma^{-1/2N}))$ be a disk with $|\xi_n| \le C_0R/2$. If we have

$$E_{\gamma(1-C_0\gamma^{-1/2N}),C_0Q}(F_1^{(q_0)}, q_0 \in K_{C_0}(Q); F_2^{(q_0)}, q_0 \in K_{C_0}(Q))$$

$$= \frac{1}{2}\Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2\Big)^{1/2}\Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2\Big)^{1/2},$$

then by (2.48)

$$\leq (1+Cc)\frac{1}{2}\|f_1\|_2\Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_2^{(q_0)}\|_2^2\Big)^{1/2} \leq E_{\gamma, C_0 Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) + Cc.$$

Thus it suffices to prove that

$$\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)}\|_{L^2(D)}^2\Big)^{1/2}\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)}\|_{L^2(D)}^2\Big)^{1/2}$$
$$\leq E_{\gamma, C_0 Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) + Cc.$$

We consider two more disks $D' := D(\xi, \gamma(1 - \frac{C_0}{2}\gamma^{-1/2N}))$, and $D'' := D(\xi, \gamma)$ that are slightly larger. We then partition f_1 into a part localized on D' and a part localized outside:

$$f_1 = P_{D'} f_1 + (1 - P_{D'}) f_1.$$

We will denote $P_{D'} f_1$ by f_1', and $(1 - P_{D'}) f_1$ by f_1''. By the linearity of decomposition into wave packets we can find for each $q_0 \in K_{C_0}(Q)$ functions $F_1'^{(q_0)}, F_1''^{(q_0)}$ satisfying $F_1'^{(q_0)} + F_1''^{(q_0)} = F_1^{(q_0)}$. We have

$$\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1'^{(q_0)}\|_{L^2(D)}^2\Big)^{1/2} \leq \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1'^{(q_0)}\|_2^2\Big)^{1/2} \leq (1+Cc)\|f_1'\|_2.$$

From Lemma 2 this last term satisfies

$$\leq (1+Cc)\|T_1 f_1\|_{L^2(D'')} + \rho^{C-N}\|f_1\|_2.$$

Now if we can prove that

$$\Big(\sum_{q_0\in K_{C_0}(Q)} \|T_1F_1''^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} \lesssim \rho^{C-N}\|f_1\|_2, \tag{2.52}$$

then from Hölder's inequality and convexity we will have

$$\Big(\sum_{q_0\in K_{C_0}(Q)} \|T_1F_1^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} \le \Big(\sum_{q_0\in K_{C_0}(Q)} \big(\|T_1F_1'^{(q_0)}\|_{L^2(D)} + \|T_1F_1''^{(q_0)}\|_{L^2(D)}\big)^2\Big)^{1/2}$$

$$\le \Big(\sum_{q_0\in K_{C_0}(Q)} \|T_1F_1'^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} + \Big(\sum_{q_0\in K_{C_0}(Q)} \|T_1F_1''^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2}$$

$$\le (1+Cc)\|T_1f_1\|_{L^2(D'')} + \rho^{C-N}\|f_1\|_2 \le \|T_1f_1\|_{L^2(D'')} + Cc.$$

Multiplying everything by

$$\Big(\sum_{q_0\in K_{C_0}(Q)} \|T_1F_1^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2},$$

and recalling Definition 2 we obtain what we desired. So it remains to prove (2.52). Since number of cubes in $K_{C_0}(Q)$ is much smaller than ρ^C it is enough to show

$$\|T_1F_1''^{(q_0)}\|_{L^2(D)} \lesssim \rho^{C-N}\|f_1\|_2.$$

We should have

$$F_1''^{(q_0)} = \sum_{T_1\in \mathbf{T}_1} \frac{m_{q_0,T_1}}{m_{T_1}}\phi''_{T_1},$$

So we can use the Hölder inequality to consider tubes in separate groups. First consider tubes T_i with $\mathrm{dist}(T_i, D) \ge \rho^3$. From the fast decay property (2.32) contribution of these tubes is acceptable. The number tubes with $dist(T_i, D) < \rho^3$ is less than ρ^C, so it is enough to prove

$$\|T_1\phi''_{T_i}\|_{L^2(D)} \lesssim \rho^{C-N}\|f_1\|_2$$

individually for each tube. We further divide these tubes into two groups. Assume we have $\mathrm{dist}(T_i, D) < R^{1/2+1/100N}$. We have from (2.31)

$$\|T_1\phi''_{T_i}\|_{L^2(D)} \le \|\phi''_{T_i}\|_2 \lesssim c^{-C}r\|\widetilde{\chi}_{T_1}(\cdot,\xi_n)T_1f''_1(\cdot,\xi_n)\|_{L^2(\mathbb{R}^{n-1}\times\{\xi_n\})}$$

$$\lesssim c^{-C}r\|\widetilde{\chi}_{T_1}(\cdot,\xi_n)T_1f''_1(\cdot,\xi_n)\|_{L^2(D'_-)} + c^{-C}r\|\widetilde{\chi}_{T_1}(\cdot,\xi_n)T_1f''_1(\cdot,\xi_n)\|_{L^2(D'^{ext}_-)}$$

Our condition ensures that $dist(T_1\cap\mathbb{R}^{n-1}\times\{\xi_n\}, D'^{ext}_-) \ge \frac{C_0}{5}\rho^{1-1/N}$, and we have $\rho^{1-1/N}/R^{1/2} \ge \rho^{1/2N}$. Thus using (2.17), and the decay of the smooth cutoff we have

$$\lesssim c^{-C}r\|T_1f''_1(\cdot,\xi_n)\|_{L^2(D'_-)} + c^{-C}r\|\widetilde{\chi}_{T_1}(\cdot,\xi_n)\|_{L^\infty(D'^{ext}_-)}\|T_1f''_1(\cdot,\xi_n)\|_{L^2(D'^{ext}_-)}$$

$$\lesssim \rho^{C-N}\|f_1\|_2.$$

Now consider the case $\mathrm{dist}(T_i, D) \ge R^{1/2+1/100N}$. If Q_D is a cube centered at ξ with sidelength CR then $\mathrm{dist}(Q_D, T) \gtrsim R^{1/2+1/100N}$. Thus from Lemma 8, (2.32) we have

$$\|T_1\phi''_{T_1}\|_{L^2(D)} \lesssim \mathrm{dist}(T, Q_D)^{C-N}\|f_1\|_2 \lesssim r^{C-N}\|f_1\|_2.$$

This completes the first stage of the proof.

We now turn to the second stage, which will proceed in a very similar way. We have, from monotonicity of the definition of concentration

$$E_{\gamma,C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) \le E_{\rho(1-C_0\rho^{-1/2N}),C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)).$$

Now if we have

$$E_{\rho(1-C_0\rho^{-1/2N}),C_0Q}(f_1; F_2^{(q_0)}, q_0 \in K_{C_0}(Q)) = \frac{1}{2}\|f_1\|_2\Big(\sum_{q_0\in K_{C_0}(Q)}\|F_2^{(q_0)}\|_2^2\Big)^{1/2},$$

then by (2.48)

$$\leq (1 + Cc)\frac{1}{2}\|f_1\|_2\|f_2\|_2 \leq E_{\rho,C_0Q}(f_1; f_2) + Cc.$$

This time let $D := D(\xi, \rho - C_0 \cdot \rho^{-1/2N})$ with $|\xi_n| \leq C_0R/2$. Thus it suffices to prove that

$$\|f\|_{L^2(D)}\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_2F_2^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} \leq E_{\gamma,C_0Q}(f_1; f_2) + Cc.$$

We again consider two more disks $D' := D(\xi, \rho(1 - \frac{C_0}{2} \cdot \rho^{-1/2N}))$, and $D'' := D(\xi, \rho)$ that are slightly larger. We this time partition f_2 into a part localized on D' and a part localized outside:

$$f_2 = P_{D'}f_2 + (1 - P_{D'})f_2.$$

We will denote $P_{D'}f_2$ by f_2', and $(1-P_{D'})f_2$ by f_2''. By the linearity of decomposition into wave packets we can find for each $q_0 \in K_{C_0}(Q)$ functions $F_2'^{(q_0)}, F_2''^{(q_0)}$ satisfying $F_2'^{(q_0)} + F_1''^{(q_0)} = F_2^{(q_0)}$. We have

$$\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_2F_2'^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} \leq \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_2'^{(q_0)}\|^2_2\Big)^{1/2} \leq (1 + Cc)\|f_2'\|_2.$$

From Lemma 2 this last term satisfies

$$\leq (1 + Cc)\|T_2f_2\|_{L^2(D'')} + \rho^{C-N}\|f_2\|_2.$$

This time if we can prove that

$$\Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_2F_2''^{(q_0)}\|^2_{L^2(D)}\Big)^{1/2} \lesssim \rho^{C-N}\|f_1\|_2,$$

from the Hölder inequality and convexity we will have

$$\leq \|T_2f_2\|_{L^2(D'')} + Cc.$$

Then multiplying everything by

$$\|T_1 f_1\|_{L^2(D)}^2$$

and recalling Definition 2 we can conclude the seceond stage. So it remains to prove (2.3). But this follows from exactly the same arguments as before: we first observe that it is sufficient to prove this on individual cubes q_0, and then by the Hölder inequality consider the tubes in three separate groups, each of which follow from arguments already described. Thus we are done.

We turn to well approximation properties. First we prove (2.50). From (2.39) we have

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^2(I^{c,C_0}(Q))} \lesssim c^{-C} R^{(2-n)/4}. \tag{2.53}$$

We have from (2.29) the estimate

$$\|T_1 f_1 T_2 f_2\|_{L^1(I^{c,C_0}(Q))} \lesssim R,$$

and from the definition of Φ_1 and the Cauchy-Schwarz inequality

$$\begin{aligned}
\|\Phi_1 T_2 f_2\|_{L^1(I^{c,C_0}(Q))} &= \|\sum_{q_0 \in K_{C_0}(Q)} |T_1 F_1^{(q_0)}| \chi_{q_0} T_2 f_2\|_{L^1(I^{c,C_0}(Q))} \\
&\leq \sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)} T_2 f_2\|_{L^1(q_0)} \\
&\leq \sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)}\|_{L^2(q_0)} \|T_2 f_2\|_{L^2(q_0)} \\
&\leq \Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)}\|_{L^2(q_0)}^2\Big)^{1/2} \Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_2 f_2\|_{L^2(q_0)}^2\Big)^{1/2}.
\end{aligned} \tag{2.54}$$

Now applying (2.13) and (2.48),

$$\leq \Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)}\|_{L^2(Q)}^2 \Big)^{1/2} \|T_2 f_2\|_{L^2(Q)} \lesssim R^{1/2} \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2 \Big)^{1/2} \cdot R^{1/2} \|f_2\|_2$$
$$\lesssim R. \tag{2.55}$$

From these two L^1 estimates with triangle inequality we obtain

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^1(I^{c,C_0}(Q))} \lesssim R. \tag{2.56}$$

Using the Hölder inequality one obtains from (2.53)

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^q L^2(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4} + \frac{2-q}{2q}}. \tag{2.57}$$

We wish to obtain a $L^q L^1$ inequality to interpolate with this last one. To this end observe that by the Cauchy-Schwarz inequality

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^\infty L^1(I^{c,C_0}(Q))} \leq \|T_1 f_1 - \Phi_1\|_{L^\infty L^2(I^{c,C_0}(Q))} \|T_2 f_2\|_{L^\infty L^2(I^{c,C_0}(Q))}$$

From the triangle inequality

$$\lesssim \big(\|T_1 f_1\|_{L^\infty L^2(I^{c,C_0}(Q))} + \|\Phi_1\|_{L^\infty L^2(I^{c,C_0}(Q))} \big) \|T_2 f_2\|_{L^\infty L^2(I^{c,C_0}(Q))}$$

Now with the conservation of mass we clearly have

$$\|T_1 f_1\|_{L^\infty L^2(I^{c,C_0}(Q))}, \|T_2 f_2\|_{L^\infty L^2(I^{c,C_0}(Q))} \lesssim 1.$$

The same is true for the remaining term:

$$\begin{aligned}
\|\Phi_1\|_{L^\infty L^2(I^{c,C_0}(Q))} &\le \sup_{\xi_n} \| \sum_{q_0 \in K_{C_0}(Q)} |T_1 F_1^{(q_0)}| \chi_{q_0} \|_{L^2(\mathbb{R}^{n-1}\times\{\xi_n\})} \\
&= \sup_{\xi_n} \Big(\sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)} \chi_{q_0}\|^2_{L^2(\mathbb{R}^{n-1}\times\{\xi_n\})} \Big)^{1/2} \\
&\lesssim \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2 \Big)^{1/2} \\
&\lesssim 1
\end{aligned}$$

So we have

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^\infty L^1(I^{c,C_0}(Q))} \lesssim 1.$$

Interpolating this with (2.56) we obtain

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^q L^1(I^{c,C_0}(Q))} \lesssim R^{1/q}. \tag{2.58}$$

Interpolating this with (2.59) we obtain

$$\|(T_1 f_1 - \Phi_1) T_2 f_2\|_{L^q L^p(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4}\frac{2-q}{2q}+\frac{1}{q}\frac{2-p}{p}}. \tag{2.59}$$

We now run a similar process to estimate

$$\|\Phi_1 (T_2 f_2 - \Phi_2)\|_{L^q L^p(I^{c,C_0}(Q))}.$$

We wish to prove that

$$\|\Phi_1 (T_2 f_2 - \Phi_2)\|_{L^2(I^{c,C_0}(Q))} \lesssim c^{-C} R^{(2-n)/4}. \tag{2.60}$$

We have

$$\|\Phi_1(T_2f_2-\Phi_2)\|_{L^2(I^{c,C_0}(Q))} = \|(\sum_{q_0\in K_{C_0}(Q)} |T_1F_1^{(q_0)}|\chi_{q_0})(T_2f_2-\Phi_2)\|_{L^2(I^{c,C_0}(Q))}$$
$$= \Big(\sum_{q_0\in K_{C_0}(Q)} \|(T_2f_2-\Phi_2)T_1F_1^{(q_0)}\|^2_{L^2(I^{c,C_0}(Q))}\Big)^{1/2}.$$

Since there are only C_0^C cubes in $K_{C^0}(Q)$ it suffices to prove (2.60) for an individual summand.

$$\|(T_2f_2-\Phi_2)T_1F_1^{(q_0)}\|_{L^2(I^{c,C_0}(Q))} \lesssim c^{-C}R^{(2-n)/4}.$$

Just as in (2.40) we can write

$$\|(T_2f_2-\Phi_2)T_1F_1^{(q_0)}\|_{L^2(I^{c,C_0}(Q))} \le \|\sum_{q_1\in K_{C_0}(Q)} |T_1F_1^{(q_0)}T_2f_2^{(q_1)}|(1-\chi_{q_1})\|_{L^2(I^{c,C_0}(Q))} \tag{2.61}$$
$$\le \sum_{q_1\in K_{C_0}(Q)} \|T_1F_1^{(q_0)}T_2f_2^{(q_1)}\|_{L^2(I^{c,C_0}(Q)\setminus q_1)}. \tag{2.62}$$

Again since the number of cubes is small it suffices to prove

$$\|T_1F_1^{(q_0)}T_2f_2^{(q_1)}\|_{L^2(I^{c,C_0}(Q)\setminus q_1)} \lesssim c^{-C}R^{(2-n)/4}.$$

We are now in a position very similar to (2.41). From here exactly the same arguments bring us to the analogue of (2.45):

$$\sum_{q\in K_J(Q):dist(q,q_1)\gtrsim C_0^{-1}cR} \|T_1F_1^{(q_0)}\widehat{\phi}_r^q\|_2^2\Big(\sum_{T_2\in\mathbf{T}_2}\frac{m_{q_0,T_2}}{m_{T_2}}\|T_2\phi_{T_2}\widehat{\phi}_r^q\|_2\Big)^2 \lesssim c^{-C}r^2. \tag{2.63}$$

Using the Cauchy-Schwarz inequality exactly in the same way as before

$$\Big(\sum_{T_2\in\mathbf{T}_2}\frac{m_{q_0,T_2}}{m_{T_2}}\|T_2\phi_{T_2}\widehat{\phi}_r^q\|_2\Big)^2 \le \sum_{T_2\in\mathbf{T}_2}\frac{\|T_2\phi_{T_2}\widehat{\phi}_r^q\|_2^2}{m_{T_2}}\widetilde{\chi}_{T_2}^{-1}(\xi_q)\cdot\sum_{T_2\in\mathbf{T}_2} m_{q_0,T_2}\widetilde{\chi}_{T_2}(\xi_q).$$

Observe that

$$\sum_{T_2\in\mathbf{T}_2} m_{q_0,T_2}\widetilde{\chi}_{T_2}(\xi_q) = \sum_{T_2\in\mathbf{T}_2}\sum_{q_2\in K_{C_0}(Q)} \|T_1F_1^{(q_2)}\widetilde{\chi}_{T_2}\|^2_{L^2(q_0)}\widetilde{\chi}_{T_2}(\xi_q)$$
$$+\sum_{q_2\in K_{C_0}(Q)}\|F_1^{(q_2)}\|_2^2\cdot\sum_{T_2\in\mathbf{T}_2} R^{-10n}\widetilde{\chi}_{T_2}(\xi_q).$$

Defining this time

$$\chi := \Big(\sum_{T_2\in\mathbf{T}_2}\widetilde{\chi}_{T_2}(\xi_q)\widetilde{\chi}^2_{T_1}\Big)^{1/2}\widetilde{\chi}_{q_0},$$

and changing order of summation we have

$$\leq \sum_{q_2\in K_{C_0}(Q)}\|T_1F_1^{(q_2)}\chi\|^2_{L^2(q_0)} + \sum_{q_2\in K_{C_0}(Q)}\|F_1^{(q_2)}\|_2^2\cdot R^{-10n}\sum_{T_2\in\mathbf{T}_2}\widetilde{\chi}_{T_1}(\xi_q).$$

The second term, from tube counting and (2.48), is bounded by R^{-5n}. Each term in the first sum , from the same arguments as before, satisfy $\lesssim c^{-C}r\|F_1^{(q_2)}\|_2^2$. Summing and using (2.48) allows one to bound the first sum by $\lesssim c^{-C}r$. Inserting this information back to (2.63) to see that it reduces to proving

$$\sum_{q\in K_J(Q)}\|T_1F_1^{(q_0)}\widehat{\phi}^q_r\|_2^2\sum_{T_1\in\mathbf{T}_2}\frac{\|T_2\phi_{T_2}\widehat{\phi}^q_r\|_2^2}{m_{T_2}\widetilde{\chi}_{T_2}(\xi_q)}\lesssim c^{-C}r.$$

This can be bounded by

$$\sum_{T_2\in\mathbf{T}_2}\sup_{q\in K_J(Q)}\big(\widetilde{\chi}^{-3}_{T_2}(\xi_q)\|T_2\phi_{T_2}\widehat{\phi}^q_r\|_2^2\big)\sum_{q\in K_J(Q)}\frac{\|T_1F_1^{(q_0)}\widehat{\phi}^q_r\|_2^2\widetilde{\chi}^2_{T_2}(\xi_q)}{m_{T_2}}. \tag{2.64}$$

The inner sum, from the same arguments as in the proof of (2.39), satisfies

$$\lesssim \frac{1}{m_{T_2}}\sum_{q\in K_J(Q)}\|T_1F_1^{(q_0)}\widehat{\phi}^q_r\|^2_{L^2(Cq)}\widetilde{\chi}^2_{T_2}(\xi_q)\lesssim 1.$$

Hence (2.64) satisfies

$$\sum_{T_2 \in \mathbf{T}_2} \sup_{q \in K_J(Q)} \left(\widetilde{\chi}_{T_2}^{-3}(\xi_q) \| T_2 \phi_{T_2} \widehat{\phi}_r^q \|_2^2 \right).$$

which, in turn, by (2.33) satisfies

$$\lesssim c^{-C} r.$$

Thus we finally have

$$\|\Phi_1(T_2 f_2 - \Phi_2)\|_{L^2(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4}}. \tag{2.65}$$

From this one gets with the Hölder inequality

$$\|\Phi_1(T_2 f_2 - \Phi_2)\|_{L^q L^2(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4} + \frac{2-q}{2q}}. \tag{2.66}$$

For the L^1 estimates on $\Phi_1 \Phi_2$ and $\Phi_1 T f_2$ the process described in (2.54) suffices. Then by triangle inequality we can obtain the appropriate L^1 estimate on $\Phi_1(T_2 f_2 - \Phi_2)$. An appropriate $L^\infty L^1$ estimate on $\Phi_1(T_2 f_2 - \Phi_2)$ can easily be deduced from repeating the same arguments as above. Interpolation then gives

$$\|\Phi_1(T_2 f_2 - \Phi_2)\|_{L^q L^1(I^{c,C_0}(Q))} \lesssim R^{\frac{1}{q}}.$$

Interpolating this with (2.66) gives

$$\|\Phi_1(T_2 f_2 - \Phi_2)\|_{L^q L^r(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4} \frac{2-q}{2q} + \frac{1}{q} \frac{2-p}{p}}.$$

Finally, this inequality together with (2.59) and the triagnle inequality gives

$$\|T_1 f_1 T_2 f_2 - \Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q))} \lesssim c^{-C} R^{\frac{2-n}{4} \frac{2-q}{2q} + \frac{1}{q} \frac{2-p}{p}}.$$

Notice that for our fixed pair of exponents (q, r) we have

$$\frac{2-n}{4}\frac{2-q}{2q} + \frac{1}{q}\frac{2-p}{p} = 0$$

so from triangle inequality we obtain (2.50).

Finally we prove (2.51). Arguments very similar to those used above will suffice to prove this. First let

$$\Omega := I^{c,C_0}(Q)) \cap C^3(\xi_C, r))$$

and observe that if $r \geq R$ the desired result follows from (2.50). So we will assume $r < R$. Main difference of what follows from what is above will be in L^1 based estimates. Indeed we have from what we proved above

$$\|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^qL^2(\Omega)} \leq \|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^qL^2(I^{c,C_0}(Q))} \lesssim c^{-C}R^{\frac{2-n}{4}+\frac{2-q}{2q}}. \tag{2.67}$$

In the L^1 case however we use Lemma 4 to obtain

$$\|T_1f_1T_2f_2\|_{L^1(\Omega)} \lesssim r^{1/2}R^{1/2} = (r/R)^{1/2}R$$

and

$$\begin{aligned}
\|\Phi_1\Phi_2\|_{L^1(\Omega)} &= \|\sum_{q_0\in K_{C_0}(Q)} |T_1F_1^{(q_0)}|\chi_{q_0} \cdot \sum_{q_1\in K_{C_0}(Q)} |T_2F_2^{(q_1)}|\chi_{q_1}\|_{L^1(\Omega)} \\
&= \|\sum_{q_0\in K_{C_0}(Q)} |T_1F_1^{(q_0)}T_2F_2^{(q_0)}|\chi_{q_0}\|_{L^1(\Omega)} \\
&= \sum_{q_0\in K_{C_0}(Q)} \|T_1F_1^{(q_0)}T_2F_2^{(q_0)}\|_{L^1(\Omega)} \\
&\lesssim (r/R)^{1/2}R \sum_{q_0\in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2\|F_2^{(q_0)}\|_2 \\
&\lesssim (r/R)^{1/2}R\Big(\sum_{q_0\in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2\Big)^{1/2} \cdot \Big(\sum_{q_1\in K_{C_0}(Q)} \|F_2^{(q_1)}\|_2^2\Big)^{1/2} \\
&\lesssim (r/R)^{1/2}R.
\end{aligned}$$

Then triangle inequality yields

$$\|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^1(\Omega)} \lesssim (r/R)^{1/2}R. \tag{2.68}$$

For the $L^\infty L^1$ estimate, however, we do not need any major changes:

$$\|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^\infty L^1(\Omega)} \le \|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^\infty L^1(I^{c,C_0}(Q))} \lesssim 1.$$

Interpolating this with the L^1 estimate gives

$$\|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^qL^1(\Omega)} \lesssim (r/R)^{1/2q}R^{1/q}.$$

Finally interpolating this with (2.67) gives

$$\|T_1f_1T_2f_2 - \Phi_1\Phi_2\|_{L^qL^p(\Omega)} \lesssim c^{-C}(r/R)^{\frac{1}{2q}\frac{2-p}{p}}R^{\frac{2-n}{4}\frac{2-q}{2q}+\frac{1}{q}\frac{2-p}{p}}.$$

Our particular choice of the exponents allow us to write

$$\|T_1 f_1 T_2 f_2 - \Phi_1 \Phi_2\|_{L^q L^p(\Omega)} \lesssim c^{-C} (r/R)^{\epsilon/4},$$

and thus from the triangle inequality we obtain the desired result.

■

The following is an application of the wave packet decomposition that will be useful in the next section. It actually is a simpler version of the well-approximation property (2.51), that removes part of a function lying on a certain domain directly, rather than subtracting a sophisticated well-approximating function. As such its proof is easier.

Lemma 10 *Let* $R \geq 2^{NC_1}$, $2^{NC_1/2} \leq r \leq R^{1/2+4/N}$ *and let* $D = D(\xi_D, C_0^{1/2} r)$. *Let* $f_i \in L^2(U_i)$, $i = 1, 2$ *with margin*$(f_i) \geq 1/3C_0$ *for both functions. Then we have*

$$\|(P_D f_1) T_2 f_2\|_{L^q L^r(Q(\xi_D, 2R) \setminus Q(\xi_D, R))} \lesssim R^{-1/C} \|f_1\|_2 \|f_2\|_2,$$

$$\|T_1 f_1 (P_D f_2)\|_{L^q L^r(Q(\xi_D, 2R) \setminus Q(\xi_D, R))} \lesssim R^{-1/C} \|f_1\|_2 \|f_2\|_2.$$

Proof. As both of these estimates are proven similarly, we will only prove the first one. We may assume $\|f_1\|_2 = \|f_2\|_2 = 1$ by simply normalizing. Also we may set $\xi_D = 0$. Thus the statement we will prove is

$$\|(P_D f_1) T_2 f_2\|_{L^q L^r(Q(0,2R) \setminus Q(0,R))} \lesssim R^{-1/C},$$

We will obtain this, using the Hölder inequality and interpolation, from the L^1 estimate

$$\|(P_D f_1) T_2 f_2\|_{L^1(Q(0,2R) \setminus Q(0,R))} \lesssim R^{C/N} R^{3/4} \tag{2.69}$$

and the L^2 estimate

$$\|(P_D f_1) T_2 f_2\|_{L^2(Q(0,2R)\setminus Q(0,R))} \lesssim R^{C/N} R^{\frac{2-n}{4}}. \tag{2.70}$$

We first handle (2.69). This is relatively easy. We have

$$\begin{aligned}\|(P_D f_1) T_2 f_2\|_{L^1(Q(0,2R)\setminus Q(0,R))} &\le \|(P_D f_1) T_2 f_2\|_{L^1((Q(0,2R)\setminus C^1(0,R^{1/2+C/N}))} \\ &\quad + \|(P_D f_1) T_2 f_2\|_{L^1(C^1(0,R^{1/2+C/N}))}\end{aligned}$$

The first term's contribution satisfies $\lesssim R^{C-N}$ by Lemma 3 and the Hölder inequality. The contribution of the second term, by Lemma 4, and conservation of mass property satisfies $\lesssim R^{3/4+C/N}$. Thus we have (2.69).

For (2.70) we decompose $P_D f_1$ using our wave packet decomposition lemma, choosing $c = 2^{-C_0}$. We have

$$P_D f_1 = \sum_{T_1 \in \mathbf{T}_1} T_1 \phi_{T_1}$$

from Lemma 8. We use the Hölder inequality to obtain

$$\|(P_D f_1) T_2 f_2\|_{L^2(Q(0,2R)\setminus Q(0,R))} \le \sum_{T_1 \in \mathbf{T}_1} \|T_1 \phi_{T_1} T_2 f_2\|_{L^2(Q(0,2R)\setminus Q(0,R))}.$$

For tubes with $\mathrm{dist}(T,0) \gtrsim C_0 R$ we write

$$\|T_1 \phi_{T_1} T_2 f_2\|_{L^2(Q(0,2R)\setminus Q(0,R))} \le \|T_1 \phi_{T_1}\|_{L^\infty(Q(0,2R))} \|T_2 f_2\|_{L^2(Q(0,2R))}$$

For the first factor on the right hand side we utilize (2.32), and for the second conservation of mass to obtain

$$\lesssim R^C \mathrm{dist}(T,0)^{-N}.$$

Thus total contribution of such tubes satisfies

$$\lesssim R^{C-N}.$$

We turn to tubes with $R^{1/2+1/N} \lesssim \text{dist}(T,0) \lesssim C_0 R$. We have, from (2.31), for such tubes

$$\|\phi_{T_1}\|_2 \lesssim c^{-C} R^{1/2} \|\widetilde{\chi}_{T_1}(\cdot,0) P_D f_1(\cdot,0)\|_2 \lesssim R^{C-N}.$$

Since number of such tubes is at most R^C, their total contribution is acceptable as well. We finally deal with tubes satisfying $\text{dist}(T,0) \lesssim R^{1/2+1/N}$. Let $\mathbf{T}_1'$ denote these tubes. Clearly, cardinality of $\mathbf{T}_1'$ satisfies $\lesssim R^C$. We want to estimate

$$\|\sum_{T_1 \in \mathbf{T}_1'} T_1 \phi_{T_1} T_2 f_2\|_{L^2(Q(0,2R)\setminus Q(0,R))}.$$

Let $R^{N^{-1/2}} T_1$ denote the tube with same central line but with radius multiplied by $R^{N^{-1/2}}$. We have, from (2.32)

$$\|T_1 \phi_{T_1}(1 - \chi_{R^{N^{-1/2}} T_1})\|_{L^\infty(Q(0,2R)} \lesssim R^{C-N^{1/2}}.$$

Thus

$$\begin{aligned} &\|\sum_{T_1 \in \mathbf{T}_1'} T_1 \phi_{T_1} T_2 f_2 (1 - \chi_{R^{N^{-1/2}} T_1})\|_{L^2(Q(0,2R)\setminus Q(0,R))} \\ &\le \sum_{T_1 \in \mathbf{T}_1'} \|T_1 \phi_{T_1} T_2 f_2 (1 - \chi_{R^{N^{-1/2}} T_1})\|_{L^2(Q(0,2R))}. \end{aligned}$$

This, then, satisfies

$$\lesssim \sum_{T_1 \in \mathbf{T}_1'} \|T_1 \phi_{T_1}(1 - \chi_{R^{N^{-1/2}} T_1})\|_{L^\infty(Q(0,2R))} \|T_2 f_2\|_{L^2(Q(0,2R))} \lesssim R^{C-N^{1/2}}.$$

Thus it remains to estimate

$$\| \sum_{T_1 \in \mathbf{T}_1'} T_1 \phi_{T_1} T_2 f_2 \chi_{R^{N-1/2} T_1} \|_{L^2(Q(0,2R) \setminus Q(0,R))}.$$

We observe the geometric fact that at most $R^{C/N}$ of tubes $T_1 \in \mathbf{T}_1'$ can intersect inside $Q(0,2R) \setminus Q(0,R)$. Thus using almost orthogonality it is enough to show that

$$\Big(\sum_{T_1 \in \mathbf{T}_1'} \| T_1 \phi_{T_1} T_2 f_2 \|_2^2 \Big)^{1/2} \lesssim R^{\frac{2-n}{4}}.$$

For an individual term we have from Lemma 6

$$\| T_1 \phi_{T_1} T_2 f_2 \|_2 \lesssim R^{\frac{2-n}{4}} \| \phi_{T_1} \|_2 \| f_2 \|_2 \lesssim R^{\frac{2-n}{4}} \| \phi_{T_1} \|_2.$$

Thus

$$\Big(\sum_{T_1 \in \mathbf{T}_1'} \| T_1 \phi_{T_1} T_2 f_2 \|_2^2 \Big)^{1/2} \lesssim R^{\frac{2-n}{4}} \Big(\sum_{T_1 \in \mathbf{T}_1'} \| \phi_{T_1} \|_2^2 \Big)^{1/2}.$$

From (2.34), and (2.20) we have

$$\Big(\sum_{T_1 \in \mathbf{T}_1'} \| \phi_{T_1} \|_2^2 \Big)^{1/2} \lesssim (1 + Cc) \| P_D f_1 \|_2 \leq (1 + Cc) \| f_1 \|_2 \lesssim 1.$$

Thus we are done.

$\blacksquare$

2.4 Conclusion of the Proof

In this section we conclude the proof of Theorem 1. The following is the first step in this direction. It only partially uses the power of the machinery we developed, and will be used to bound $A(R)$ for relatively small R. But we will use again some of the same ideas in its proof.

Proposition 2 *Let $R \geq 2C_0 2^{C_1}$ and $0 < c \leq 2^{C_0}$. Let $f_i \in L^2(U_i),\ i = 1,2$ be functions satisfying the normalization $\|f_i\|_2 = 1,\ i = 1,2$ and the relaxed margin requirement (2.12). Then we have for any cube Q_R with side-length R the inequality*

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R)} \leq (1 + Cc)\overline{A}(R/2)\|f_1\|_2\|f_2\|_2 + c^{-C}. \tag{2.71}$$

Proof. From Lemma 7 it is possible to find a cube Q of sidelength CR inside C^2R with

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R)} \leq (1 + Cc)\,\|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q))}\,.$$

From the well approximation property (2.50) in Proposition 1 we can find $\Phi_i,\ i = 1,2$ with

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q))} \leq (1 + Cc)\,\|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q))} + c^{-C}. \tag{2.72}$$

From the definitions of Φ_i and the triangle inequality we must have

$$\|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q))} \leq \sum_{q_0 \in K_{C_0}(Q)} \|T_1 F_1^{(q_0)} T_2 F_2^{(q_0)}\|_{L^q L^r(q_0)}.$$

Then from definition of the best constants $A(\cdot)$, and from (2.47) of Proposition 1 we have

$$\|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q))} \leq A(2^{-C_0}CR) \sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2 \|F_2^{(q_0)}\|_2.$$

Applying the Cauchy-Schwarz inequality on the right-hand side yields

$$\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2 \|F_2^{(q_0)}\|_2 \leq \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_1^{(q_0)}\|_2^2\Big)^{1/2} \Big(\sum_{q_0 \in K_{C_0}(Q)} \|F_2^{(q_0)}\|_2^2\Big)^{1/2}.$$

From here the requirement on total mass property (2.48) proven in Proposition 1 together with the definition of $\overline{A}(\cdot)$ gives the desired result. ■

This proposition of course gives

$$A(R) \leq (1 + Cc)\overline{A}(R/2)\|f_1\|_2\|f_2\|_2 + c^{-C}.$$

We observe that if $R \geq 4C_0R^{C_1}$, since we have $\overline{A}(\cdot)$ on the right hand side, we can bound $A(R')$ for all $R/2 \leq R' \leq R$ by

$$(1 + Cc)\overline{A}(R/2)\|f_1\|_2\|f_2\|_2 + c^{-C}.$$

Since all $A(R')$ with $R' \leq R/2$ are also bounded by this quantity, we then can conclude that $\overline{A}(R)$ is also bounded by it. Thus it is possible to iterate this proposition. So let $2C_0R^{C_1} \leq R \leq C_0R^{NC_1}$ and let $c = (1/R)^{1/N}$. Let J be a natural number with the property $2C_0R^{C_1} \leq 2^{-J}R < 4C_0R^{C_1}$. We iterate the proposition to obtain

$$\overline{A}(R) \leq \prod_{i=0}^{J-1}(1 + C(2^j/R)^{1/N})\overline{A}(R/2^J) + 2JR^{C/N}.$$

Here of course $J \leq \log_2 R - C_1 \leq R^{1/N}$, so $2J(R)^{C/N} \lesssim R^{C/N}$. As for the first term we first note a simple inequality that will be utilized to estimate it: $\ln 1 + x \leq x$ whenever $x \geq 0$. Now we have

$$\ln\Big(\prod_{i=0}^{J-1} 1 + C(2^j/R)^{1/N}\Big) \leq \sum_{i=0}^{J-1} 1 + C(2^j/R)^{1/N} \leq \frac{C}{R^{1/N}}\sum_{i=0}^{J-1} 2^{j/N}.$$

This last sum can be bounded by

$$\frac{2^{J/N} - 1}{2^{1/N} - 1}.$$

But we have from our inequality $\ln 1 + x \leq x$

$$1 + \frac{1}{N \log_2 e} \leq e^{1/N \log_2 e} = 2^{1/N},$$

so

$$2^{1/N} - 1 \geq 1/2N.$$

Thus the last sum is bounded by $2^{J/N} 2N \leq 2^{\log_2 R^{1/N}} 2N \leq 2NR^{1/N}$. So the product itself should be bounded by $2^{CN} \leq R^{1/N}$. To term $\overline{A}(R/2^J)$ we can apply the trivial inequality $A(R) \leq CR^{n-1/r+1/q}$ that was stated before to get $\overline{A}(R/2^J) \lesssim R^{CC_1}$. Thus we finally get for $R \leq C_0 R^{NC_1}$

$$\overline{A}(R) \lesssim 2^{CC_1}. \tag{2.73}$$

The rest of the proof will deal with larger R. To this end we will introduce a variant of $A(\cdot)$ that is sensitive to concentration and can be related to $A(\cdot)$. This version will allow us to use the improved estimate for cone neighborhoods (2.51).

Definition 3 *Let $R \geq 2^{NC_1/2}$ and $r, r' > 0$. Then we define $A(R, r, r')$ as the best constant satisying*

$$\|f_1 f_2\|_{L_t^q L_x^p(Q_R \cap C^3(\xi_C, r'))} \leq A(R, r, r')(\|f_1\|_2 \|f_2\|_2)^{1/q} E_{r, C_0 Q_R}(f_1, f_2)^{1/q'}$$

for all cubes Q_R of side-length R, all $\xi_C \in \boldsymbol{R}^n$, and all functions $f_i \in L^2(U_i)$ obeying the the margin requirement (2.12)

Before proceeding to relating this newly defined quantity to $A(\cdot)$ we introduce a technical lemma that is of very general nature, and may be helpful whenever one is dealing with mixed norms. It will be used in several propositions below.

Lemma 11 *Let $f_i, 1 \leq i \leq k$ be a finite collection of complex valued functions with $f_i \in$*

$L^qL^r(\boldsymbol{R}^n)$ *for all* i. *If we have* $q < r$ *and supports of these functions mutually disjoint, then*

$$\|\sum_{i=1}^{k} f_i\|_{L^qL^r}^q \leq \sum_{j=1^k} \|f_j\|_{L^qL^r}^q.$$

Proof. We will first exploit the fact that the supports are disjoint:

$$\int (\int |\sum_{i=1}^{k} f_i(\overline{x}, x_n)|^r d\overline{x})^{q/r} dx_n = \int (\sum_{i=1}^{k} \int |f_i(\overline{x}, x_n)|^p dx)^{q/r} dx_n.$$

Now we exploit concavity

$$\leq \sum_{i=1}^{k} \int (\int |f_i(\overline{x}, x_n)|^r d\overline{x})^{q/r} dx_n,$$

and thus we are done.

$\blacksquare$

We now exploit non-concentration to bound $A(\cdot)$ by the variant $A(\cdot, \cdot, \cdot)$ with some gain. It is natural to expect a result such as we state below in the absence of concentration, since in this case the term $E_{\cdot,\cdot}$ will be small, and thus $A(\cdot, \cdot, \cdot)$ large. When there is concentration, the trick is to reduce the cube size through Lemma 10 to such a scale that the size of disk on which concentration occurs is no longer much smaller than the cube size. We now execute this plan.

Proposition 3 *Suppose* $R \geq 2^{NC_1}$. *Then the following inequality holds*

$$A(R) \leq (1 - C_0^{-C}) \sup_{\substack{2^{NC_1} \leq \widetilde{R} \leq R \\ \widetilde{R}^{1/2+4/N} \leq r}} A(\widetilde{R}, r, C_0(1+r)) + 2^{CC_1}.$$

Proof. It is clear that if we can show

$$\|T_1f_1T_2f_2\|_{L^qL^r(Q_R)} \leq (1 - C_0^{-C}) \sup_{\substack{2^{NC_1} \leq \widetilde{R} \leq R \\ \widetilde{R}^{1/2+4/N} \leq r}} A(\widetilde{R}, r, C_0(1+r)) + 2^{CC_1}$$

for all cubes Q_R of side-length R centered at some point $\xi_Q \in \mathbb{R}^n$, and for all $f_i \in L^2(U_i)$ satisfying the normalization $\|f_i\|_2 = 1$ and the strict margin requirement (2.10), from the definition of $A(\cdot)$ our result follows. We can clearly assume that $A(R) \leq 2^{CC_1}$. Furthermore we can assume that $\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R)} \approx A(R)$. We let $0 < \delta < 1/4$ be a small number to be chosen later. We define r to be the supremum of all radii $r \geq 2^{NC_1(1/2+4/N)}$ such that

$$E_{r,C_0 Q_R}(f_1, f_2) \leq 1 - \delta$$

or $r = 2^{NC_1(1/2+4/N)}$ if there is no such radius. Since δ is small, there is a disk $D := D(\xi_D, r)$ satisfying

$$min(\|\phi\|_{L^2(D)}, \|\psi\|_{L^2(D)}) \geq 1 - 2\delta,$$

and $|\xi_{D_n} - \xi_{Q_n}| \leq C_0 R/2$. Such a disk must exist, or else we would have $E_{r,C_0 Q_R}(f_1, f_2) \leq 1 - 2\delta$, conflicting the continuity of $T_i f_i$. Also define the disk $D' = C_0^{1/2} D$, and set $\Omega = Q_R \cap C^3(\xi_D, C_0 r)$. We decompose the functions f_i into two parts: $f_i = P_{D'} f_i + (1 - P_{D'}) f_i$.

We investigate the problem in two cases: $r > R^{1/2+4/N}$ and $r \leq R^{1/2+4/N}$. We handle the first case now. Since $D'^{ext}_{-} \subset D^{ext}$ we have from (2.19) of Lemma 2, and our choice of D, D' the inequalities

$$\|(1 - P_{D'}) f_1\|_2^2, \ \|(1 - P_{D'}) f_2\|_2^2 \lesssim \delta + C_0^{-C}.$$

Hence we must have

$$\|(1 - P_{D'}) f_1 (1 - P_{D'}) f_2\|_{L^q L^r(Q_R)} \lesssim (\delta + C_0^{-C}) A(R). \tag{2.74}$$

On the other hand from (2.24) in Lemma 3 we have

$$\|P_{D'} f_1 T_2 f_2\|_{L^q L^r(Q_R \setminus \Omega)}, \|(1 - P_{D'}) f_1 P_{D'} f_2\|_{L^q L^r(Q_R \setminus \Omega)} \lesssim C_0^{-C}.$$

Thus from the triangle inequality and our assumption $\|T_1f_1T_2f_2\|_{L^qL^r(Q_R)} \approx A(R)$ we have

$$\|T_1f_1T_2f_2\|_{L^qL^r(Q_R\setminus\Omega)} \lesssim (\delta + C_0^{-C})A(R) \lesssim (\delta + C_0^{-C})\|T_1f_1T_2f_2\|_{L^qL^r(Q_R)}. \tag{2.75}$$

Now we will apply Lemma 11. This point of the proof does not work for the uppermost endpoint $(n+1/n-1, 1)$ for $n \geq 4$, and is the reason why we cede it.

$$\begin{aligned}\|T_1f_1T_2f_2\|^q_{L^qL^r(Q_R)} &\leq \|T_1f_1T_2f_2\|^q_{L^qL^r(Q_R\setminus\Omega)} + \|T_1f_1T_2f_2\|^q_{L^qL^r(\Omega)} \\ &\leq C(\delta + C_0^{-C})^q\|T_1f_1T_2f_2\|^q_{L^qL^r(Q_R)} + \|T_1f_1T_2f_2\|^q_{L^qL^r(\Omega)}.\end{aligned}$$

Thus,

$$\|T_1f_1T_2f_2\|_{L^qL^r(\Omega)} \geq (1 - C(\delta + C_0^{-C})^q)^{1/q}\,\|T_1f_1T_2f_2\|_{L^qL^r(Q_R)}\,.$$

But we also have from our choice of r

$$\|T_1f_1T_2f_2\|_{L^qL^r(\Omega)} \leq A(R, r, C_0(1+r))(1-\delta)^{1/q'}.$$

Here setting $\delta = C_0^{-C}$ and showing

$$(1-\delta^q)(1-\delta^q)^{1/q} \geq (1-\delta)^{1/q'} \tag{2.76}$$

will suffice. To see this last inequality remember that we set $1/q = \min(1, \frac{n}{4}) - \epsilon$ and $1/r = 1 - \frac{2}{n}\frac{1}{q}$ where $0 < \epsilon < \frac{1}{10n}$. So we have $1/q \leq 1-\epsilon$, $q \geq \frac{1}{1-\epsilon} \geq 1+\epsilon$, $1 < q < 2$. When $\delta = 0$ both sides of the above inequality are the same, then both decrease for small positive δ. So we shall compare values of derivatives of them for such δ to find the one that decreases faster. On the left the derivative is

$$-(1+\frac{1}{q})q(1-\delta^q)^{1/q}\delta^{q-1} \geq -2q\delta^{q-1} \geq -4\delta^{\epsilon}$$

while on the right we have

$$-\frac{1}{q'}(1-\delta)^{\frac{1}{q'}-1} \leq -\epsilon\frac{1}{(1-\delta)^{1/q}} \leq -\epsilon.$$

For our choice of δ clearly the former is greater.

We now turn to the case where $r \leq R^{1/2+4/N}$. We define $\widetilde{R} := r^{\frac{1}{1/2+4/N}}$. So $2^{NC_1} \leq \widetilde{R} \leq R$ and $r \geq \widetilde{R}^{1/2+4/N}$. We will proceed similarly: just as we obtained upper and lower bounds on $\|T_1 f_1 T_2 f_2\|_{L^q L^r(\Omega)}$ and compared them, this time we will obtain upper and lower bounds on $\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q(\xi_D,\widetilde{R})\cap\Omega)}$ and compare them. If we have $\widetilde{R} > 2^{NC_1}$ then

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q(\xi_D,\widetilde{R})\cap\Omega)} \leq A(R,r,C_0(1+r))(1-\delta)^{1/q'}.$$

If on the other hand $\widetilde{R} = 2^{NC_1}$ from (2.73) we have

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q(\xi_D,\widetilde{R})\cap\Omega)} \leq 2^{CC_1}.$$

These two constitute our upper bound. For the lower bound observe that from the same arguments that we used to obtain (2.75) we can obtain

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q(\xi_D,\widetilde{R})\setminus\Omega)} \lesssim (\delta + C_0^{-C})A(R) \lesssim (\delta + C_0^{-C})\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R)}.$$

So if we can show that

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R\setminus Q(\xi_D,\widetilde{R}))} \lesssim (\delta + C_0^{-C})A(R)$$

using the same arguments as above will give the upper bound, and then the whole result. This inequality follows from the estimate (2.74), and

$$\|(P_{D'}f_1)T_2f_2\|_{L^qL^r(Q_R\setminus Q(\xi_D;\widetilde{R}))} \lesssim (\delta + C_0^{-C})A(R),$$
$$\|(1-P_{D'}f_1)P_{D'}f_2\|_{L^qL^r(Q_R\setminus Q(\xi_D;\widetilde{R}))} \lesssim (\delta + C_0^{-C})A(R).$$

These two estimates can be obtained from Lemma 10 by a dyadic decomposition. We decompose the domain of integration $Q_R \setminus Q(\xi_D;\widetilde{R})$ dyadically into pieces

$$Q(\xi_D, 2\widetilde{R}) \setminus Q(\xi_D, \widetilde{R}), \quad Q(\xi_D, 4\widetilde{R}) \setminus Q(\xi_D, 2\widetilde{R}), \quad Q(\xi_D, 8\widetilde{R}) \setminus Q(\xi_D, 4\widetilde{R}) \ldots$$

. We will have coefficients $(2^k\widetilde{R}) - 1/C$, which means summing over terms $2^{-k/C}$, and this sum clearly converges. Therefore simply summing with the triangle inequality together with our assumption $A(R) \geq 2^{CC_1}$ gives the desired result.

∎

Thus we only need to bound the variant $A(\cdot,\cdot,\cdot)$ by $A(\cdot)$ appropriately. This will be divided into two cases according to absence or presence of concentration. First we suppose the absence of concentration.

Proposition 4 *Let R, r, r', c be positive quantities satisfying $R \geq 2^{NC_1/2}$, $r \geq C_0^C R$, $r' > 0$, and $0 < c \leq 2^{-C_0}$. Then we have the following inequality*

$$A(R, r, r') \leq (1 + Cc)\overline{A}(R) + c^{-C}.$$

Proof. Let $f_i \in L^2(U_i)$, $i = 1, 2$ be functions that satisfy the normalization of mass $\|f_1\|_2 = \|f_2\|_2 = 1$, and the strict margin requirement (2.10). If we can prove that

$$\|T_1f_1T_2f_2\|_{L^qL^r(Q_R)} \leq E_{r,C_0Q_R}(f_1, f_2)^{1/q'}(1 + Cc)\overline{A}(R) + c^{-C}$$

for an arbitrary cube Q_R of side-length R, we will have for arbitrary $\xi_C \in \mathbb{R}^n$

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R \cap C^3(\xi, r'))} \leq \left[(1+Cc)\overline{A}(R) + c^{-C}\right] E_{r,C_0 Q_R}(f_1, f_2)^{1/q'} (\|f_1\|_2 \|f_2\|_2)^{1/q},$$

and thus from the definition of $A(R, r, r')$ we must have the desired inequality. Let $D := D(\xi_Q, r/2)$ where ξ_Q denotes the center of Q. We perform a decomposition with respect to this disk for both functions: $T_i f_i = P_D f_i + (1 - P_D) f_i$. From the definition of the operator P_D we have $T_i g_i = P_D f_i$ with $g_i(x) = \widehat{P_D f_i}(-x, 0)$. These g_i obey the relaxed margin requirement (2.12) and furthermore by (2.18) in Lemma 1

$$\|g_1\|_2 \|g_2\|_2 = \|P_D f_1\|_2 \|P_D f_2\|_2 \leq E_{r,C_0 Q_R}(f_1, f_2) + C R^{C - N/2}.$$

Thus we can apply Proposition 3 to get

$$\|(P_D f_1)(P_D f_2)\|_{L^q L^r(Q_R)} \leq (1 + Cc)(E_{r,C_0 Q_R}(\phi, \psi) + C R^{C-N/2})\overline{A}(R) + c^{-C}.$$

Since we have $\overline{A}(R) \lesssim R^C$, we can absorb the term $C R^{C-N/2} \overline{A}(R)$ into c^{-C}. Thus we will be done if we can show

$$\|((1 - P_D)\phi)\psi\|_{L^q L^r(Q_R)}, \|(P_D \phi)(1 - P_D)\psi\|_{L^q L^r(Q_R)} \leq c^{-C}.$$

But these follow from Lemma 3 and its analogue with the roles of f_1, f_2 reversed.

∎

Thus remains the concentrated case for which we again utilize the idea of passing to smaller scales until concentration disappears. Once concentration vanishes one can again apply the previous propostion. So in effect the following proposition allows us to reduce the concentrated case to the non-concentrated.

Proposition 5 *Let $R \geq C_0 2^{N C_1/2}$, and let $C_0^C R \geq r > R^{1/2+3/N}$. Then for any $0 < c \leq$*

2^{-C_0} *we have*

$$A(R,r,r') \leq (1+Cc)A(R/C_0, r(1-Cr^{-1/3N}), r') + c^{-C}(1+\frac{R}{r'})^{-\epsilon/4}.$$

Proof. As above, from the definition of $A(R,r,r')$ it suffices to prove that given a cube Q_R of sidelength R, a point $\xi_C \in \mathbb{R}^n$, and functions $f_i \in L^2(U_i)$ functions obeying the strict margin requirement (2.10) and the normalization of mass $\|f_i\|_2 = 1$ we have

$$\begin{aligned} \|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R \cap C^3(\xi_C; r'))} &\leq (1+Cc)A(R/C_0, \widetilde{r}, r')E_{r,C_0 Q_R}(f_1, f_2)^{1/q'} \\ &\quad + c^{-C}(1+\frac{R}{r})^{-\epsilon/4} \end{aligned} \tag{2.77}$$

with $\widetilde{r} = r(1-Cr^{-1/3N})$. We now perform a series of reductions. By applying Lemma 7 to $T_1 f_1 T_2 f_2 \chi_{C^3(\xi_C;r')}$ we can find a cube Q of sidelength CR lying inside $C^2 Q_R$ with the property

$$\|T_1 f_1 T_2 f_2\|_{L^q L^r(Q_R \cap C^3(\xi_C; r'))} \leq (1+Cc) \|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q) \cap C^3(\xi_C; r'))}.$$

Thus showing (2.78) with $\|T_1 f_1 T_2 f_2\|_{L^q L^r(I^{c,C_0}(Q) \cap C^3(\xi_C; r'))}$ on the left-hand side suffices. By applying (2.51) of our main proposition we see that it is sufficient to prove

$$\begin{aligned} \|\Phi_1 \Phi_2\|_{L^q L^r(I^{c,C_0}(Q) \cap C^3(\xi_C; r'))} &\leq (1+Cc)A(R/C_0, \widetilde{r}, r')E_{r,C_0 Q_R}(f_1, f_2)^{1/q'} \\ &\quad + c^{-C}(1+\frac{R}{r})^{-\epsilon/4}. \end{aligned} \tag{2.78}$$

Then since $1/2 \leq E_{r,C_0 Q_R}(f_1, f_2)^{1/q'} \leq 1$ it is enough to prove that left-hand side satisfies

$$\leq (1+Cc)A(R/C_0, \widetilde{r}, r')E_{\widetilde{r},C_0 Q_R}(F_1^{(q_0)}, q_0 \in K_{C_0}(Q); F_2^{(q_0)}, q_0 \in K_{C_0}(Q))^{1/q'}. \tag{2.79}$$

Now we apply Lemma 11 to see that

$$\|\Phi_1\Phi_2\|_{L^qL^r(I^{c,C_0}(Q)\cap C^3(\xi_C,r'))} \leq \sum_{q_0\in K_{C_0}(Q)} \|T_1F_1^{(q_0)}T_2F_2^{(q_0)}\|^q_{L^qL^r(q_0\cap C^3(\xi_C,r'))}.$$

Thus it suffices to show that the right-hand side is less than or equal to expression in (2.79). To see this we consider individual terms in the summation. By definition of the quantity $A(\cdot,\cdot,\cdot)$ we have

$$\|T_1F_1^{(q_0)}T_2F_2^{(q_0)}\|^q_{L^qL^r(q_0\cap C^3(\xi_C,r'))} \leq A(R/C_0,\widetilde{r},r')^q\|F_1^{(q_0)}\|_2\|F_2^{(q_0)}\|_2E_{\widetilde{r},C_0Q}(F_1^{(q_0)},F_2^{(q_0)})^{q/q'}.$$

Thus summing these and bearing in mind the fact $E_{\widetilde{r},C_0Q}(F_1^{(q_0)},F_2^{(q_0)}) \leq E_{\widetilde{r},C_0Q_R}(F_1^{(q_0)},q_0 \in K_{C_0}(Q);F_2^{(q_0)},q_0\in K_{C_0}(Q))$ we can bound our summation by

$$A(R/C_0,\widetilde{r},r')^qE_{\widetilde{r},C_0Q_R}(F_1^{(q_0)},q_0\in K_{C_0}(Q);F_2^{(q_0)},q_0\in K_{C_0}(Q))^{q/q'}\sum_{q_0\in K_{C_0}(Q)}\|F_1^{(q_0)}\|_2\|F_2^{(q_0)}\|_2.$$

At this point using the Cauchy-Schwarz inequality and the Proposition (2.48) yields

$$\begin{aligned}\sum_{q_0\in K_{C_0}(Q)}\|F_1^{(q_0)}\|_2\|F_2^{(q_0)}\|_2 &\leq \Big(\sum_{q_0\in K_{C_0}(Q)}\|F_1^{(q_0)}\|_2^2\Big)^{1/2}\Big(\sum_{q_0\in K_{C_0}(Q)}\|F_2^{(q_0)}\|_2^2\Big)^{1/2}\\ &\leq (1+Cc)\|f_1\|_2\|f_2\|_2.\end{aligned}$$

Thus we are done. ∎

The following corollary will allow us to finally conclude our proof by showing that $A(\cdot,\cdot,\cdot)$ can appropriately be related to $A(\cdot)$ regardless of concentration, by handling the non-concentrated case with Proposition 4 and reducing the concentrated case to the non-concentrated one with Proposition 5.

Corollary 2 *Let $R \geq 2^{NC_1}$, and let $r \geq R^{1/2+4/N}$. Then for any $0 < c \leq 2^{-C_0}$ we have*

$$A(R, r, C_0(1+r)) \leq (1 + Cc)\overline{A}(R) + c^{-C}.$$

Proof. If $r \geq C_0^C R$ the result follows from Proposition 4, thus we assume $r < C_0^C R$. We let J denote the least integer with $r/2 \geq C_0^{-J} C_0^C R$. We have $J \lesssim \log r$ since $r \geq R^{1/2+4/N}$. Define $r := r_0 > r_1 > \ldots > r_J$ recursively by $r_{j+1} = r_j(1 - Cr_j^{-1/3N})$. This sequence decreases very slowly, each time r decreases at most by $Cr^{1-1/3N}$, and has only about $\log r$ terms, so

$$r - r_J \lesssim Cr^{1-1/3N} \cdot \log r,$$

which means $r/2 \leq r_J < r$. We define $c_j := C_0^{-1} c C_0^{(j-J)\epsilon/8C}$ for $0 < j \leq J$. We iterate Proposition 4 with these values to obtain

$$A(R, r, C_0(1+r)) \leq \prod_{j=1}^{J}(1+Cc_j)\Big[(1+Cc)A(R/C_0^J, r_J, C_0(1+r)) + \sum_{j=1}^{J} c_j^{-C}(1+\frac{R}{C_0^j(1+r)})^{-\epsilon/4}\Big].$$

We use same techniques introduced at the beginning of this section to estimate the product:

$$\ln \prod_{j=1}^{J}(1 + Cc_j) = \sum_{j=1}^{J} \ln 1 + Cc_j \leq C \sum_{j=1}^{J} c_j.$$

Writing out explicitly one has for this last sum

$$\begin{aligned} C_0^{-1} c C_0^{-J\epsilon/8C} \sum_{j=1}^{J} C_0^{j\epsilon/8C} &= C_0^{-1} c C_0^{-J\epsilon/8C} \frac{C_0^{(J+1)\epsilon/8C} - 1}{C_0^{\epsilon/8C} - 1} \\ &\leq C_0^{-1} c C_0^{-J\epsilon/8C} C_0^{(J+1)\epsilon/8C} \\ &\leq c. \end{aligned}$$

Hence

$$\prod_{j=1}^{J}(1+Cc_j) \leq e^c \leq 1+3c.$$

It remains to estimate the sum inside the paranthesis

$$\begin{aligned}\sum_{j=1}^{J} c_j^{-C}(1+\frac{R}{C_0^j(1+r)})^{-\epsilon/4} &= C_0^C c^{-C}\sum_{j=1}^{J}(C_0^{(j-J)/2}+\frac{C_0^{(j-J)/2}R}{C_0^j(1+r)})^{-\epsilon/4} \\ &\leq C_0^C c^{-C}\sum_{j=1}^{J}(C_0^{(j-J)/2}+\frac{C_0^{-J/2}R}{1+r})^{-\epsilon/4}.\end{aligned}$$

Now observe that $R/(1+r) \geq C_0^{J-2-C}$, thus

$$\leq C_0^C c^{-C}\sum_{j=1}^{J}(C_0^{\frac{J}{2}-2-C})^{-\epsilon/4} \leq C_0^C c^{-C} J C_0^{-J\epsilon/8} \leq c^{-C}.$$

Thus we finally have

$$A(R,r,C_0(1+r)) \leq (1+Cc)A(R/C_0^J, r_J, C_0(1+r)) + c^{-C}.$$

At this point one can apply Proposition 4 to $A(R/C_0^J, r_J, C_0(1+r)$ to obtain the desired result.

■

We can now conclude the proof of Theorem 1 by showing that for all $R > 0$

$$A(R) \lesssim 2^{CC_1}. \tag{2.80}$$

For $R \leq 2^{NC_1}$ we already proved this. For $R > 2^{NC_1}$, set $c = 2^{-C_1}$ and combine Proposition 3 with the corollary above to obtain

$$A(R) \leq (1-C_0^{-C})\overline{A}(R) + 2^{CC_1}.$$

Since we also have (2.80), we can improve this to

$$\overline{A}(R) \leq (1 - C_0^{-C})\overline{A}(R) + 2^{CC_1}.$$

Thus for $R > 2^{NC_1}$ too we have

$$\overline{A}(R) \lesssim 2^{CC_1}.$$

Thus we are done.

Chapter 3

Linear Restriction Estimates for Elliptic surfaces

In this chapter we will prove Theorem 2.

3.1 Preliminaries

In this first section we will reformulate and simplify the result, briefly describe the proof and develop some technical tools for the proof. We first start with reformulation.

We reformulate the theorem in a way very similar to reformulation of Theorem 1 in Chapter 2. We first define an elliptic surface by a parametrization. Let U be a small disk around the origin lying on $\mathbb{R}^{n-1} \times \{0\}$. We can parametrize an elliptic surface S on this neighborhood using a phase function $\phi(x)$ with

$$\phi(x) = \langle Ax, x \rangle + \mathcal{O}(|x|^3)$$

where A is a positive definite matrix. We define

$$Tf(\xi) := \int_U f(x) e^{-2\pi i (x \cdot \bar{\xi} + \phi(x)\xi_n)} dx.$$

Then proving that for every $f \in L^\infty(U)$

$$\|Tf\|_{L^q} \leq C_q \|f\|_\infty$$

whenever q satisfies conditions stated on Theorem 2 will imply Theorem 2. We however note that given a phase ϕ as described above we can always write

$$Tf(\xi) = \int_U f(x) e^{-2\pi i(x\cdot\bar{\xi}+\phi(x)\xi_n)} dx = \int_S g(y) e^{-2\pi i y\cdot\xi} d\sigma(y)$$

with, in the second integral, $y \in \mathbb{R}^n$, $g(\overline{y}, \phi(y)) = f(\overline{y})$, S the surface that the phase ϕ gives, and σ an appropriate measure on the surface -pullback of the Lebesgue measure on U.- Thus we can use both formulations interchangeably.

We further reformulate the problem through localization. Localizing the estimate is a standard approach in proving restriction estimates for the Fourier transform, and our proof in Chapter 2 was an instance of this. From standard ϵ-removal arguments, see [6], to prove the Theorem 2 it suffices to prove that for

$$\begin{aligned} q &= \frac{18}{7} - \frac{2}{735} && \text{if } \; k = 2 \\ q &= \frac{8k+2}{4k-1} - \frac{(5k-4)(2k+1)}{(4k^2-k)(28k+7)(4k^2-k)} && \text{if } \; k > 2. \end{aligned} \tag{3.1}$$

with $n = 3k$, and for ϵ arbitrarily small, we have

$$\|Tf\|_{L^q(B(0,R))} \leq C_{q,\epsilon} R^\epsilon \|f\|_\infty \tag{3.2}$$

for every radius $R > 0$. We further note that it is enough to prove this estimate for $\|f\|_\infty \leq 1$. Thus we introduce $A_q(R)$ as the best constant satisfying

$$\|Tf\|_{L^q(B(0,R))} \leq A_q(R)$$

for $\|f\|_\infty \leq 1$. Since we trivially have

$$\|\widehat{fd\sigma}\|_{L^q(B(0,R))} \lesssim R^{\frac{n}{q}}$$

following from

$$|Tf(\xi)| = \Big|\int_U f(x)e^{-2\pi i(x\cdot\bar{\xi}+\phi(x)\xi_n)}dx\Big| \le C_U$$

this constant is well defined. As a result it suffices to prove, for q satisfying (3.1), and for any $\epsilon > 0$

$$A_q(R) \le C_{q,\epsilon}R^{\epsilon}. \tag{3.3}$$

The proof will use the multilinear estimates of Bennett, Carbery and Tao, [2], described in Chapter 1, the method of rescaling introduced in [25], and the conservation of mass property described in Lemma 1, Chapter 2. Since $\|\cdot\|_2$ norm of a compactly supported function can be bounded by its $\|\cdot\|_\infty$ norm, Bennett-Carbery-Tao estimates are useful. We will observe that we can use these estimates to essentially reduce the dimension of the problem, which makes the method of rescaling work better. This will be achieved through decomposition of the surface into small caps. We will give more detailed heuristics about the proof as we proceed. As we will prove Theorem 2 by proving the local estimates (3.3), this explanation suffices to motivate introduction of the following local form of Bennett, Carbery and Tao result, also proven in [2]. We formulate it using parametrizations. Let $S_i \in \mathbb{R}^n,\ 1 \le i \le m$ be smooth, compact hypersurfaces whose parametrizations $\phi_i : U_i \to S_i$ from subsets U_i of $\mathbb{R}^{n-1}$ satisfy smoothness condition

$$\|\phi_i\|_{C^2(U_i)} \le A,$$

and that are transverse, i.e, for any points $y_i \in S_i$ unit normals y_i' satisfy $|y_1' \wedge \ldots \wedge y_m'| \ge B$. Then for operators T_i associated to ϕ_i we have the localized version

$$\Big\|\prod_{i=1}^{m} T_i f_i\Big\|_{L^q(B(0,R))} \le C_{q,A,B,\epsilon}R^{\epsilon}\prod_{i=1}^{m}\|f_i\|_{L^2(U_i)} \tag{3.4}$$

for $q = \frac{2}{m-1}$ and every $\epsilon > 0$.

Finally, to be used in the proof of one of our lemmas below, we introduce a further variant of the Bennett-Carbery-Tao result, that deals with thin neighborhoods of surfaces, and that was too proven in [2]. Let $S_i \in \mathbb{R}^n$, $1 \le i \le m$ be surfaces just as in the previous paragraph, and let Y_i^R be C/R neighborhoods of these surfaces for $R \gg 1$. Then we have for functions $f_j \in L^2(Y_i^R)$ and every $\epsilon > 0$

$$\| \prod_{i=1}^{m} \widehat{f} \|_{L^q(B(0,R))} \le C_{q,A,B,\epsilon} R^{\epsilon - n/2} \prod_{i=1}^{m} \|f_i\|_{L^2(Y_i^R)} \tag{3.5}$$

for $q = \frac{2}{m-1}$.

We will now prove some lemmas, that will be used in sections 3.2 and 3.3. The first of these will mainly be used to make rigorous heuristic information coming from the uncertainty principle. We will describe how to use the well-known Bernstein inequality for exponents lower than 1.

Lemma 12 *Let $0 < q \le p \le \infty$, and let $f \in L^p(\mathbb{R}^n)$ be a function whose Fourier transform $\widehat{f}$ is supported in a box of measure E. Then*

$$\|f\|_p \lesssim E^{\frac{1}{q} - \frac{1}{p}} \|f\|_q$$

Proof. First of all we note that we may assume $E = 1$, since all other cases will follow from scaling. If $p > 1$ we may use the Bernstein inequality to lower it to 1, so it suffices to prove it for $p \le 1$. First assume that $p = 1$. Then we have

$$\|f\|_1 = \int |f| \le \|f\|_\infty^{1-q} \int |f|^q.$$

We apply the Bernstein inequality to $\|f\|_\infty$:

$$\|f\|_\infty \lesssim \|f\|_1.$$

So we have

$$\|f\|_1 \lesssim \|f\|_1^{1-q} \int |f|^q = \|f\|_1^{1-q} \|f\|_q^q.$$

If $\|f\|_1$ is zero then what we want to prove is trivial, so assuming that it is not zero we divide both sides by $\|f\|_1^{1-q}$ to obtain

$$\|f\|_1^q \lesssim \|f\|_q^q.$$

From which we obtain

$$\|f\|_1 \lesssim \|f\|_q.$$

For $p < q$ we run the same argument:

$$\|f\|_p^p = \int |f|^p \leq \|f\|_\infty^{p-q} \int |f|^q.$$

Then we have

$$\|f\|_\infty \lesssim \|f\|_1 \lesssim \|f\|_p.$$

From here with the same arguments we conclude

$$\|f\|_p \lesssim \|f\|_q.$$

∎

We now proceed to our next lemma. In what follows S will be a surface with properties described in this section, and for $x \in S$ a point, x' will be used to denote the unit normal vector at this point. We will need to use the Gauss map $\Gamma : S^{n-1} \to S$ taking the unit normal y' of a point to itself y. We let $0 < c \ll 1$ denote various small constants. The notation $\fint_E$ will be used to denote average over a set E.

Lemma 13 *Let $S_i \subset S$, $1 \leq i \leq n$ be small caps with $|y_1' \wedge \ldots \wedge y_n'| > c$ for all $y_i \in S_i$, and let $D_i \subset S_i$ be discrete sets with at least $1/M$ separation for $M \gg 1$. Then for any bounded* *function a on S and any ball $B_M \subset \mathbb{R}^n$ of radius M we have*

$$\fint_{B_M} \prod_{i=1}^{n} |\sum_{y\in D_i} a(y)e^{-2\pi i y\cdot\xi}|^{2/n-1} d\xi \ll M^{\epsilon} \prod_{i=1}^{n} (\sum_{y\in D_i} |a(y)|^2)^{1/n-1}.$$

Proof. Our basic strategy is to pass from these sums to integrals over neighborhoods of S_i, and then to apply (3.5). We may clearly assume B_M to be centered at the origin, for otherwise we may incorporate the oscillation coming from being away from the origin into the function a. We introduce the measures $\mu_i = \sum_{y\in D_i} \delta_y$ and define the functions $f_i(z) = a(z)\chi_{D_i}(z)$. Let ϕ be a Schwartz function supported in $B(0, C)$ whose Fourier transform is non-negative on the unit ball. Let $\phi_M(\xi) = M^n\phi(M\xi)$. Then

$$\begin{aligned}
\fint_{B_M} \prod_{i=1}^{n} |\sum_{y\in D_i} a(y)e^{-2\pi i y\cdot\xi}\Big|^{\frac{2}{n-1}} d\xi &= \fint_{B_M} \prod_{i=1}^{n} |\int_{S_i} f_i(z)e^{-2\pi i z\cdot\xi} d\mu_i(z)\Big|^{\frac{2}{n-1}} d\xi \\
&= \fint_{B_M} \prod_{i=1}^{n} |\widehat{f_i d\mu_i}|^{\frac{2}{n-1}} d\xi \\
&\lesssim \fint_{B_M} \prod_{i=1}^{n} |\widehat{f_i d\mu_i}\widehat{\phi_M}|^{\frac{2}{n-1}} d\xi \\
&= \fint_{B_M} \prod_{i=1}^{n} |\widehat{f_i d\mu_i * \phi_M}|^{\frac{2}{n-1}} d\xi.
\end{aligned}$$

To this we can apply Bennett-Carbery-Tao result since $f_i d\mu_i * \phi_M$ are functions in C/M neighbourhood of S_i.

$$\lesssim M^{-n} M^{-\frac{n}{n-1}+\epsilon} \prod_{i=1}^{n} \|f_i d\mu_i * \phi_M\|_2^{\frac{2}{n-1}}.$$

We can calculate individual factors in this product using the fact that supports of at most

C of $a\chi_{D_i}\delta_y * \phi_M$ can intersect as y ranges over D_i. Thus we have

$$\lesssim M^{-n-\frac{n}{n-1}+\frac{n^2}{n-1}+\epsilon}\prod_{i=1}^{n}(\sum_{y\in D_i}|a(y)|^2)^{\frac{1}{n-1}} \lesssim M^{\epsilon}\prod_{i=1}^{n}(\sum_{y\in D_i}|a(y)|^2)^{\frac{1}{n-1}}.$$

■

This last lemma will only be used to prove the next lemma, which will be the one used in the rest of the proof. The idea of the first lemma was to pass from discrete sums to integrals. In the second lemma we will pass from integrals to sums to utilize the first lemma.

Lemma 14 *Let $V \subset \mathbb{R}^n$ be a subspace of dimension m, where $2 \leq m \leq n$. Let $P_1 \dots P_m \in S$ be points with $P_i' \in V$, and $|P_1' \wedge \dots \wedge P_m'| > c$. Let $U_1, \dots, U_m \subset S$ be small neighborhoods of $P_1, \dots, P_m$. Let $M \gg 1$, and $D_i \subset U_i$ be sets of $1/M$ separated points y that obey the condition dist$(y', V) < c/M$ for a small c. Let ϕ be a Schwartz function supported in $B(0, c)$ with Fourier transform non-negative on the unit ball, and let $\phi_M(y) = M^n\phi(My)$. Then for any $f_i \in L^\infty(U_i)$ we have*

$$\fint_{B_M}\prod_{i=1}^{m}|\sum_{y\in D_i}\int_{|z-y|<\frac{c}{M}} f_i(z)e^{-2\pi i z\cdot\xi}d\sigma(z)|^{\frac{2}{m-1}}d\xi \tag{3.6}$$

$$\ll M^{-\frac{m\cdot n}{m-1}+\epsilon}\prod_{i=1}^{m}(\sum_{y\in D_i}\|\widehat{\phi}_M\int_{|z-y|<\frac{c}{M}} f_i(z)e^{-2\pi i z\cdot\xi}d\sigma(z)\|_2^2)^{\frac{1}{m-1}}. \tag{3.7}$$

Proof. We may assume that B_M is centered at the origin just as above. By a simple rotation in the space where the surface S lives we may also assume that $V = \text{span}\{e_1, e_2, \dots, e_m\}$. Let $\hat{V}$ denote the image of $V \cap S^{n-1}$ under the Gauss map. It is possible, by partitioning D_i into subsets and modifying functions f_i if necessary, to assume $y \in \hat{V}$ for all y. Also we may assume $f_i(z) = 0$ if $|z - y| \geq c/M$ for all $y \in D_i$. Let ψ be a non-negative Schwartz function that is identical to 1 in $B(0, 10c)$, and supported in $B(0, 20c)$ whose Fourier transform is nonnegative on $B(0,1)$. We define $\psi_M(y) = M^n\psi(My)$. Then

$$\begin{aligned}
\fint_{B_M} \prod_{i=1}^{m} |\sum_{y\in D_i} \int_{|z-y|<\frac{c}{M}} f_i(z)e^{-2\pi iz\cdot\xi} d\sigma_i(z)|^{\frac{2}{m-1}} d\xi &= \fint_{B_M} \prod_{i=1}^{m} |\int_{S_i} f_i(z)e^{-2\pi iz\cdot\xi} d\sigma_i(z)|^{\frac{2}{m-1}} d\xi \\
&= \fint_{B_M} \prod_{i=1}^{m} |\widehat{f_i d\sigma}(\xi)|^{\frac{2}{m-1}} d\xi \\
&\lesssim \fint_{B_M} \prod_{i=1}^{m} |\widehat{f_i d\sigma}\widehat{\phi_M}(\xi)\widehat{\psi_M}(\xi)|^{\frac{2}{m-1}} d\xi \\
&= \fint_{B_M} \prod_{i=1}^{m} |\widehat{f_i d\sigma * \phi_M} * \psi_M(\xi)|^{\frac{2}{m-1}} d\xi.
\end{aligned}$$

We define $g_i = f_i d\sigma * \phi_M * \psi_M$. We can break $f_i d\sigma$ into pieces $f_{i,y}d\sigma$ by restricting it to the points z with $|z-y| < c/M$. Let $f_{i,y,0}d\sigma$ denote translations of these pieces to the origin -here we are translating both the function and the measure to the translation of the cap, on which these are supported, to the origin-. Then we define $g_{i,y} = f_{i,y}d\sigma * \phi_M * \psi_M$ and $g_{i,y,0} = f_{i,y,0}d\sigma * \phi_M * \psi_M$. We have

$$g_i = f_i d\sigma * \phi_M * \psi_M = \sum_{y\in D_i} f_{i,y}d\sigma * \phi_M * \psi_M = \sum_{y\in D_i} g_{i,y},$$

and

$$g_{i,y,0}(z) = g_{i,y}(z+y).$$

Using these we can write

$$
\begin{aligned}
|\widehat{g_i}(\xi)| = |\int g_i(z)e^{-2\pi i z\cdot\xi}dz| &= |\sum_{y\in D_i}\int_{|z-y|<\frac{15c}{M}} g_{i,y}(z)e^{-2\pi i z\cdot\xi}dz| \\
&= |\sum_{y\in D_i} e^{-2\pi i y\cdot\xi}\int_{|u|<\frac{15c}{M}} g_{i,y}(u+y)e^{-2\pi i u\cdot\xi}du| \\
&= |\sum_{y\in D_i} e^{-2\pi i y\cdot\xi}\int_{|u|<\frac{15c}{M}} g_{i,y,0}(u)e^{-2\pi i u\cdot\xi}du| \\
&= |\int_{|u|<\frac{15c}{M}}\big(\sum_{y\in D_i} g_{i,y,0}(u)e^{-2\pi i y\cdot\xi}\big)e^{-2\pi i u\cdot\xi}du| \\
&\le \int_{|u|<\frac{15c}{M}}|\sum_{y\in D_i} g_{i,y,0}(u)e^{-2\pi i y\cdot\xi}|du \\
&\lesssim M^{-n}\sup_{u:|u|\le\frac{15c}{M}}|\sum_{y\in D_i} g_{i,y,0}(u)e^{-2\pi i y\cdot\xi}|.
\end{aligned}
$$

Observe that the domain of u is compact, and the function concerned is continuous. Thus the supremum is attained at some u_i. Hence turning back we have to estimate

$$
M^{-\frac{2mn}{m-1}}\fint_{B_M}\prod_{i=1}^{m}|\sum_{y\in D_i} g_{i,y,0}(u_i)e^{-2\pi i y\cdot\xi}|^{\frac{2}{m-1}}d\xi.
$$

Writing $y=(v,w)\in\mathbb{R}^m\times\mathbb{R}^{n-m}$ and $\xi=(\alpha,\beta)\in\mathbb{R}^m\times\mathbb{R}^{n-m}$ we can bound this by

$$
\lesssim M^{-\frac{2mn}{m-1}}\max_{w\in B_M^{n-m}}\fint_{B_M^m}\prod_{i=1}^{m}|\sum_{y\in D_i} g_{i,y,0,\omega}(u_i)e^{-2\pi i v\cdot\alpha}|^{\frac{2}{m-1}}d\alpha
$$

where $g_{i,y,0,\omega}(u_i)=g_{i,y,0}(u_i)e^{-2\pi i w\cdot\beta}$. Since S_i has positive definite second fundamental form the projection to first m variables of $\hat{V}$ too has the same property. Thus for fixed w we apply Lemma 13 to obtain

$$
\lesssim M^{-\frac{2mn}{m-1}+\epsilon}\Big(\prod_{i=1}^{m}\sum_{y\in D_i}|g_{i,y,0}(u_i)|^2\Big)^{\frac{1}{m-1}}.
$$

We use Young's inequality to write

$$|g_{i,y}(u_i)| \lesssim \|f_{i,y,0}d\sigma * \phi_M\|_1 \|\psi_M\|_\infty \lesssim M^n \|f_{i,y,0}d\sigma * \phi_M\|_1.$$

On the other hand whenever $u \in B(0,c)$ we have

$$|f_{i,y,0}d\sigma * \phi_M| * \psi_M(u) \geq M^n \|f_{i,y,0}d\sigma * \phi_M\|_1.$$

Utilizing this information and again using Young's inequality we can write

$$\begin{aligned} |g_{i,y}(u_i)| \lesssim M^n \| |f_{i,y,0}d\sigma * \phi_M| * \phi_{\frac{M}{10}}\|_1 &\lesssim M^{\frac{n}{2}} \| |f_{i,y,0}d\sigma * \phi_M| * \psi_M\|_2 \\ &\lesssim M^{\frac{n}{2}} \|f_{i,y,0}d\sigma * \phi_M\|_2 \|\psi_M\|_1 \\ &\lesssim M^{\frac{n}{2}} \|f_{i,y,0}d\sigma * \phi_M\|_2. \end{aligned}$$

We insert this back to get

$$\lesssim M^{-\frac{m\cdot n}{m-1}+\epsilon} \prod_{i=1}^{m} (\sum_{y \in D_i} \|f_{i,y,0}d\sigma * \phi_M\|_2^2)^{\frac{1}{m-1}} \lesssim M^{-\frac{m\cdot n}{m-1}+\epsilon} \prod_{i=1}^{m} (\sum_{y \in D_i} \|f_{i,y}d\sigma * \phi_M\|_2^2)^{\frac{1}{m-1}}.$$

From the Plancherel theorem we have

$$\lesssim M^{-\frac{m\cdot n}{m-1}+\epsilon} \prod_{i=1}^{m} (\sum_{y \in D_i} \|\widehat{f_{i,y}d\sigma}\widehat{\phi_M}\|_2^2)^{\frac{1}{m-1}}$$

$$\lesssim M^{-\frac{m\cdot n}{m-1}+\epsilon} \prod_{i=1}^{m} (\sum_{y \in D_i} \|\widehat{\phi}_M \int_{|z-y|<\frac{c}{M}} f_i(z) e^{-2\pi i \xi \cdot z} d\sigma(z)\|_2^2)^{\frac{1}{m-1}}.$$

■

Our next lemma follows basically from change of variables for integrals. In [25] it was observed that this could be used to prove restriction estimates.

Lemma 15 *Let $S_\rho \subset S$ be a cap of radius ρ with ρ much smaller than the radius of the* *surface S. Then*

$$\Big\| \int_{S_\rho} f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big\|_{L^q(B_R)} \lesssim \rho^{n-1-(n+1)/p} A_q(R).$$

Proof. We can parametrize the surface S, through affine coordinate changes, on a neighborhood U of the origin as

$$\big(x_1, \ldots, x_{n-1}, |x|^2 + \mathcal{O}(|x|^3)\big)$$

Then integral becomes

$$\int_{U_\rho} f(x) e^{-2\pi i (x\cdot\bar{\xi} + \psi(x)\xi_n)} dx$$

with $U_\rho \subset U$ is a cap of radius comparable to ρ and centered at a point v, and $\psi(x) = |x|^2 + \mathcal{O}(|x|^3)$. Our first step is to take the cap U_ρ to the origin:

$$\begin{aligned} \Big\| \int_{U_\rho} f(x) e^{-2\pi i (x\cdot\bar{\xi} + \psi(x)\xi_n)} dx \Big\|^q_{L^q(B_R)} &= \Big\| \int_{U_\rho - v} f(x+v) e^{-2\pi i ((x+v)\cdot\bar{\xi} + \psi(x+v)\xi_n)} dx \Big\|^q_{L^q(B_R)} \\ &= \int_{B_R} \Big| \int_{U_\rho - v} f(x+v) e^{-2\pi i (x\cdot\bar{\xi} + \psi(x+v)\xi_n)} dx \Big|^q d\xi. \end{aligned}$$

By applying coordinate changes of the form $\xi_i' = \xi_i + \xi_n(2v_i + \ldots)$ for $1 \leq i < n$, and $\xi_n' = \xi_n(1 + P(v))$ with $P(v)$ standing for various powers of v, we obtain a new phase $\psi'(x) = |x|^2 + \mathcal{O}(|x|^3)$ and

$$\lesssim \int_{B_{CR}} \Big| \int_{U_\rho - v} f(x+v) e^{-2\pi i (x\cdot\bar{\xi} + \psi'(x)\xi_n)} dx \Big|^q d\xi.$$

Now we change variables as $x' = \rho^{-1}x$, $\xi_i' = \rho\xi_i$ for $1 \leq i < n$, and $\xi_n' = \rho^2 \xi_n$ to obtain

$$\lesssim \rho^{q(n-1)-(n+1)} \int_{B_{\rho CR}} \Big| \int_U f(\rho x + v) e^{-2\pi i (x\cdot\bar{\xi} + (|x|^2 + \rho\mathcal{O}(|x|^3))\xi_n)} dx \Big|^q d\xi.$$

Here U is a domain centered at the origin with radius comparable to 1, $f(\rho x + v)$ a function defined on U with L^∞ norm not exceeding 1, and $|x|^2 + \rho\mathcal{O}(|x|^3)$ is another elliptic phase.

Therefore we have

$$\lesssim \rho^{q(n-1)-(n+1)}(A_q(\rho R))^q \lesssim \rho^{q(n-1)-(n+1)}(A_q(R))^q.$$

$\blacksquare$

3.2 The Bourgain-Guth Argument

We are now ready to describe arguments used by Bourgain and Guth. The argument we use to prove Theorem 2 is in essence an iteration of this argument. We introduce the the constants

$$K_2 \ll K_3 \ll \ldots \ll K_{n-1} \ll K_n \ll R^{\epsilon},$$

with

$$K_m = R^{\epsilon^{10(1+n-m)}}, \qquad 2 \leq m \leq n.$$

We decompose our surface S into small caps S^n_α of size $\frac{1}{K_n}$, and we let $y_\alpha \in S^n_\alpha$. Note that for the projection U of S onto $\mathbb{R}^{n-1} \times \{0\}$ this induces a decomposition into caps U^n_α of size roughly $1/K_n$. In accordance,

$$\int_S f(y)e^{-2\pi i y\cdot\xi}\sigma(dy) = \sum_\alpha e^{-2\pi i y^n_\alpha\cdot\xi}\int_{S^n_\alpha} f(y)e^{-2\pi i(y-y^n_\alpha)\cdot\xi}d\sigma(y).$$

Here we define the operators

$$T^n_\alpha f(\xi) = \int_{S^n_\alpha} f(y)e^{-2\pi i(y-y^n_\alpha)\cdot\xi}d\sigma(y)$$

which could also be defined using U^n_α and a parametrization of S. Thus we have

$$\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y) = \sum_\alpha e^{-2\pi i y^n_\alpha\cdot\xi}T^n_\alpha f(\xi)$$

Now since $|y-y_\alpha^n|\lesssim 1/K_n$ the function $T_\alpha^n f$ is heuristically constant on balls of size comparable to K_n. The aim of the following arguments is to make this rigorous. We will obtain functions c_α^n that dominate $T_\alpha^n f$, and that are approximately constant on balls of this size. We take a Schwartz class function η on $\mathbb{R}^n$, with $\widehat{\eta}(x)=1$ on $B(0,1)$ and $\widehat{\eta}(x)=0$ outside $B(0,2)$. We let $\eta_r(x)=\frac{1}{r^n}\eta(\frac{x}{r})$. Note that if we define $g(y)=f(y+y_\alpha^n)$ we have

$$T_\alpha^n f(\xi)=\int_{S_\alpha^n-y_\alpha^n} g(y)e^{-2\pi i y\cdot\xi}d\sigma(y)$$

that is $T_\alpha^n f$ is the Fourier transform of a surface measure on a cap living in a ball of size $1/K_n$ around the origin. Thus we have the equality

$$T_\alpha^n f=T_\alpha^n f*\eta_{K_n}.$$

We need to obtain our functions c_α^n in a careful way, for we want certain results such as multilinear theory of Bennett-Carbery and Tao to be applicable to them. The following process has this aim. For a fixed ξ we use Lemma 12 to obtain

$$\begin{aligned}|T_\alpha^n f(\xi)|&\le\int|T_\alpha^n f(\xi-\zeta)\eta_{K_n}(\zeta)|d\zeta\\&\lesssim\Big(\int|T_\alpha^n f(\xi-\zeta)\eta_{K_n}(\zeta)|^{\frac{1}{n}}dy\Big)^n K_n^{n-n^2}.\end{aligned}$$

If we take the constant term inside the integral we get

$$\lesssim\Big(\int|T_\alpha^n f(\xi-\zeta)|^{1/n}\frac{1}{K_n^n}|\eta(\frac{\zeta}{K_n})|^{1/n}d\zeta\Big)^n.$$

This has the form, which we want our function c_α^n to have, but it does not have the property of being constant on balls of desired size. To also have this property we define

$$\theta_{K_m}(\zeta)=\frac{1}{K_m^n}\Big(1+\frac{|\zeta|}{K_m}\Big)^{-(10n)^{2n-m}\epsilon^{-n-10}},\qquad 2\le m\le n.$$

Then we set

$$c_\alpha^n(\xi) := \Big(\int |T_\alpha^n f(\xi-\zeta)|^{1/n}\zeta_{K_n}(\zeta)d\zeta\Big)^n.$$

We have $c_\alpha^n(\xi_1) \approx c_\alpha^n(\xi_2)$ whenever $|x_1 - x_2| < K_n$. Moreover the growth of $c_\alpha^n(\xi)$ is bounded depending on decay of θ_{K_n}. Thus convolving c_α^n with functions decaying significantly faster than θ_{K_n} will leave its value at a point approximately the same. This is a fact that will be used in this section as well as in section 3.3. Since η is a Schwarz function

$$|T_\alpha^n f(\xi)| \lesssim c_\alpha^n(\xi).$$

Hence we have

$$\Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \lesssim \sum_\alpha c_\alpha^n(\xi).$$

With this appropriate decomposition at hand, we proceed to the argument proper. We let B_R denote the ball centered at the origin with radius R, and $B(a,M)$ the ball centered at a with radius M. We fix a point $\xi \in \mathbb{R}^n$. We have two possibilities.

1.1. There exist $\alpha_1, \ldots, \alpha_n$ with $|y_1' \wedge \ldots \wedge y_n'| > c(K_n)$ for any $y_i \in U_{\alpha_i}^n$ and

$$c_{\alpha_i}^n(\xi) > K_n^{-n} \max_\alpha c_\alpha^n(\xi).$$

Here $c(K_n)$ stands for some negative power of K_n. We can choose the same $\alpha_1, \ldots, \alpha_n$ for all x in a ball of radius K_n since $c_\alpha^n(x)$ are approximately constant on balls of this size.

1.2. There exist an $(n-1)$-dimensional subspace V_{n-1} such that if $\mathrm{dist}(S_\alpha^n, \widehat{V}_{n-1}) \gtrsim 1/K_n$, then

$$c_\alpha^n(\xi) \le K_n^{-n} \max_\alpha c_\alpha^n(\xi).$$

Here $\widehat{V}$ stands for the image of $V \cap S^{n-1}$ under the Gauss map. Owing to the fact that $c_\alpha^n(\xi)$ are essentially constant on balls of radius K_n, on such balls we can take the linear subspace V_{n-1} to be the same. Note that with appropriate choice of $c(K_n)$ these two cases are complementary, that is one is the negation of the other.

First we assume **1.1** holds. Since the number of caps is comparable to K_n^{n-1}

$$\Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \lesssim K_n^{n-1}\max_\alpha c_\alpha^n(\xi) \lesssim K_n^{2n-1}\big(\prod_{i=1}^n c_{\alpha_i}^n(\xi)\big)^{1/n}$$

Thus letting $B_{1.1}$ denote $x\in B_R$ satisfying **1.1** we have

$$\int_{B_{1.1}}\Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \lesssim K_n^{(2n-1)q}\sum_{\alpha_1,\dots,\alpha_n}\int_{B_R}\big(\prod_{i=1}^n c_{\alpha_i}^n(\xi)\big)^{\frac{q}{n}}d\xi$$

where the sum on the right hand side is over all possible caps satisfying transversality property stated in the case **1.1**. Hence by the definition of $c_{\alpha_i}^n(\xi)$ and the Hölder inequality, noting that $\theta_{K_n}(\zeta_i)d\zeta_i$ has total mass $\lesssim 1$, we have

$$\lesssim K_n^{(2n-1)q}\sum_{\alpha_1,\dots,\alpha_n}\int_{B_R}\big(\prod_{i=1}^n\int|T_{\alpha_i}^n f(\xi-\zeta_i)|^{q/n}\theta_{K_n}(\zeta_i)d\zeta_i\big)d\xi$$

$$= K_n^{(2n-1)q}\sum_{\alpha_1,\dots,\alpha_n}\int_{B_R}\big(\int\prod_{i=1}^n|T_{\alpha_i}^n f(\xi-\zeta_i)|^{q/n}\theta_{K_n}(\zeta_i)d\zeta_1\dots d\zeta_n\big)d\xi.$$

Here observe how the definition of $c_\alpha^n(\xi)$ allowed us to bring the exponent q inside the integral. Now, by Fubini's theorem

$$\lesssim K_n^{(2n-1)q}\sum_{\alpha_1,\dots,\alpha_n}\int\big(\int_{B_R}\prod_{i=1}^n|T_{\alpha_i}^n f(\xi-\zeta_i)|^{q/n}dx\big)\prod_{i=1}^n\theta_{K_n}(\zeta_i)d\zeta_1\dots d\zeta_n \tag{3.8}$$

Assuming $q\ge\frac{2n}{n-1}$ we see that the inner integral has a form suitable for applying (1.9). Thus an individual integral in the sum over $\alpha_1,\dots,\alpha_n$ satisfies

$$\lesssim R^{\epsilon^5},$$

and hence (3.8) satisfies

$$\lesssim K_n^{10n^2} R^{\epsilon} \lesssim R^{\epsilon^3}.$$

The exponent $\frac{2n}{n-1}$ is the best possible according to the restriction conjecture, thus in this n-linear case we encounter no loss. However when we apply this same idea to lower levels of multilinearity we will suffer significant losses, and these losses will increase as level of multilinearity decreases. For example in trilinear case we must have $q \geq 3$, and with this it is not possible to improve best known exponents for the restriction conjecutre. Thus one must also have another method that gets better as multilinearity disappears. This is the method of rescaling, and its use will be illustrated soon.

We now assume that the case **1.2** holds. We use the fact that the number of caps is comparable to K_n^{n-1} to obtain

$$\begin{aligned}
\Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| &\lesssim \Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \\
&+\Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\gtrsim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \\
&\lesssim \Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| + \frac{1}{K_n}\max_{\alpha}\, c_\alpha^n(\xi) \\
&= I + II.
\end{aligned}$$

Letting $B_{1.2}$ denote those $\xi \in B_R$ for which 1.2 holds, we obtain

$$\begin{aligned}
\int_{B_{1.2}}\Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \lesssim &\int_{B_{1.2}}\Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \\
&+\frac{1}{K_n^p}\int_{B_R}\big(\max_{\alpha}\, c_\alpha^n(\xi)\big)^q d\xi
\end{aligned}$$

The second term here is an error term, and it is easy to evaluate. We have

$$\begin{aligned}\frac{1}{K_n^q}\int_{B_R}(\max_\alpha c_\alpha^n(\xi))^q d\xi &\leq \frac{1}{K_n^q}\sum_\alpha \int_{B_R}(c_\alpha^n(\xi))^q d\xi \\ &= \frac{1}{K_n^q}\sum_\alpha \|c_\alpha^n\|_{L^q(B_R)}^q\end{aligned}$$

where the summation is over all caps of size $1/K_n$. Writing out $c_\alpha^n(\xi)$ explicitly and using the Hölder inequality we see that we can apply Lemma 15:

$$\begin{aligned}\|c_\alpha^n\|_{L^q(B_R)}^q &\lesssim \int_{B_R}\Big(\int |T_\alpha^n f(\xi-\zeta)|^q \theta_{K_n}(\zeta)d\zeta\Big)d\xi && (3.9)\\ &\lesssim \int\Big(\int_{B_R}|T_\alpha^n f(\xi-\zeta)|^q d\xi\Big)\theta_{K_n}(\zeta)d\zeta && (3.10)\\ &\lesssim K_n^{n+1-q(n-1)}(A_q(R/K_n))^q. && (3.11)\end{aligned}$$

Thus the contribution of II is bounded by $K_n^{n(2-q)}(A_q(R/K_n))^q$. In an inductive argument trying to bound $A_q(R)$ this is harmless since we will always have $q > \frac{2n}{n-1}$ and this makes the exponent $n(2-q)$ negative.

To evaluate I, we decompose S into caps S_α^{n-1} of radius $1/K_{n-1}$. Then we follow a similar process:

$$\begin{aligned}\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y) &= \sum_\alpha \int_{U_\alpha^{n-1}\cap\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\\ &=: \sum_\alpha e^{-2\pi i y_\alpha^{n-1}\cdot\xi}\widetilde{T}_\alpha^{n-1}f(\xi).\end{aligned}$$

We also define

$$T_\alpha^{n-1}f(\xi) = \int_{S_\alpha^{n-1}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y).$$

We note the difference between $\widetilde{T}$ and T, while the first is defined on intersection of caps with a strip, the second is defined on full caps. We define $\eta_{K_{n-1}}$ just as η_{K_n} was defined, and also recall the definition of θ_{n-1}. Then we write

$$\widetilde{c}_\alpha^{n-1}(\xi) := \Big(\int |\widetilde{T}_\alpha^{n-1} f(\xi - \zeta)|^{1/n-1} \zeta_{K_{n-1}}(\zeta) d\zeta \Big)^{n-1}.$$

$$c_\alpha^{n-1}(\xi) := \Big(\int |T_\alpha^{n-1} f(\xi - \zeta)|^{1/n-1} \theta_{K_{n-1}}(\zeta) d\zeta \Big)^{n-1}.$$

The functions c_α^{n-1} will be needed in the next step of our process. From the same arguments as in the definition of c_α^n we observe that

$$|\widetilde{T}_\alpha^{n-1} f(\xi)| \lesssim \widetilde{c}_\alpha^{n-1}(x), \quad |T_\alpha^{n-1} f(\xi)| \lesssim c_\alpha^{n-1}(\xi),$$

and that both $\widetilde{c}_\alpha^{n-1}, c_\alpha^{n-1}$ are approximately constant on balls of size K_{n-1}. We fix a point $\xi \in B_{1.2}$. We have two distinct and complementary cases:

2.1. There exist $\alpha_1, \ldots, \alpha_{n-1}$ with $|\xi_1' \wedge \ldots \wedge \xi_{n-1}'| > c(K_{n-1})$ for all $\xi_i \in U_{\alpha_i}^{n-1}$, and

$$|\widetilde{c}_{\alpha_i}^{n-1}(\xi)| > K_{n-1}^{-(n-1)} \max_\alpha |\widetilde{c}_\alpha^{n-1}(\xi)|.$$

We can choose $\alpha_1, \ldots, \alpha_{n-1}$ the same for all x in a ball of size K_{n-1} since c_α^{n-1} are essentially constant on balls .

2.2. There exist an $(n-2)$-dimensional subspace V_{n-2} which can be chosen a subspace of V_{n-1} satisfying

$$|\widetilde{c}_\alpha^{n-1}(\xi)| \leq K_{n-1}^{-(n-1)} \max_\alpha |\widetilde{c}_\alpha^{n-1}(\xi)|$$

if $\mathrm{dist}(S_\alpha^{n-1}, \hat{V}_{n-2}) \gtrsim 1/K_{n-1}$. The linear subspace V_{n-2} can be chosen the same for all x in a ball of size K_{n-2}.

We assume that **2.1** holds. We will use both the multilinear theory, and the method of rescaling and compare the exponents the give to decide which one is more advantageous.

Since we have at most K_{n-1}^{n-2} functions $\widetilde{c}_{\alpha}^{n-1}$ we can write

$$\Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \lesssim K_{n-1}^{2n-3}\Big(\prod_{i=1}^{n-1}\widetilde{c}_{\alpha_i}^{n-1}(\xi)\Big)^{1/n-1}.$$

If $q\geq\frac{2n-2}{n-2}$, we can proceed to use the multilinear theory just as in the case **1.1**. So

$$\int_{B_{2.1}}\Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi$$

$$\lesssim C(K_{n-1})\int_{B_R}\Big(\prod_{i=1}^{n-1}\widetilde{c}_{\alpha_i}^{n-1}(\xi)\Big)^{q/n-1}d\xi$$

where we, of course, choose $\widetilde{c}_{\alpha_i}^{n-1}$ depending on ξ. This, then satisfies

$$\lesssim C(K_{n-1})\int_{B_R}\Big(\prod_{i=1}^{n-1}\sum_{\alpha_i}c_{\alpha_i}^{n}(\xi)\Big)^{q/n-1}d\xi$$

$$\lesssim C(K_n)\sum_{\alpha_1,\dots,\alpha_{n-1}}\int_{B_R}\Big(\prod_{i=1}^{n-1}c_{\alpha_i}^{n}(\xi)\Big)^{q/n-1}d\xi$$

and from multilinear theory of Bennett-Carbery and Tao we obtain

$$\lesssim C(K_n)R^{\epsilon^5}\lesssim R^{\epsilon^3}.$$

This time we assume $q<\frac{2n-2}{n-2}$. In this case we have

$$\begin{aligned}&\fint_{B(a,K_n)}\Big|\int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i x\cdot\xi}d\sigma(y)\Big|^q d\xi\\ &\lesssim K_{n-1}^{(2n-3)q}\sum_{\alpha_1,\dots,\alpha_{n-1}}\fint_{B(a,K_n)}\Big(\prod_{i=1}^{n-1}\widetilde{c}_{\alpha_i}^{n-1}(\xi)\Big)^{q/n-1}d\xi\end{aligned}\tag{3.12}$$

with the subspace V_{n-1} being the same for all $x\in B(a,K_n)$. The sum is over those $\alpha_1,\dots,\alpha_{n-1}$ satisfying the transversality condition of case **2.1**. The choice of $\alpha_1,\dots,\alpha_{n-1}$

is the same only for balls of size K_{n-1}, but as the subspace remains the same, caps S_α^{n-1} are always chosen from those intersecting the set

$$\{y : \text{dist}(y, \hat{V}_{n-1}) \lesssim \frac{1}{K_n}\}.$$

Although, as we stated, this time we will apply the method of rescaling, we also will exploit multilinearity partially. This partial exploitation is necessary, and without it improvement of best know exponents would not be possible. We take an individual integral from (3.12) above. As $q < \frac{2n-2}{n-2}$ by the Hölder inequality

$$\fint_{B(a,K_n)} \Big(\prod_{i=1}^{n-1} \widetilde{c}_{\alpha_i}^{n-1}(\xi)\Big)^{\frac{q}{n-1}} d\xi \lesssim \Big(\fint_{B(a,K_n)} \Big(\prod_{i=1}^{n-1} \widetilde{c}_{\alpha_i}^{n-1}(\xi)\Big)^{\frac{2}{n-2}} d\xi\Big)^{\frac{q(n-2)}{2(n-1)}}. \tag{3.13}$$

Writing out $\widetilde{c}_{\alpha_i}^{n-1}(\xi)$ explicitly,

$$\lesssim \Big(\fint_{B(a,K_n)} \Big(\prod_{i=1}^{n-1} \int |\widetilde{T}_\alpha^{n-1} f(\xi - \zeta_i)|^{\frac{1}{n-1}} \theta_{K_{n-1}}(\zeta_i) d\zeta_i\Big)^{\frac{2(n-1)}{n-2}} d\xi\Big)^{\frac{q(n-2)}{2(n-1)}}.$$

We use first the Hölder inequality then Fubini's theorem to obtain

$$\lesssim \Big(\fint_{B(a,K_n)} \Big(\prod_{i=1}^{n-1} \int |\widetilde{T}_\alpha^{n-1} f(\xi - \zeta_i)|^{\frac{2}{n-2}} \theta_{K_{n-1}}(\zeta_i) d\zeta_i\Big) d\xi\Big)^{\frac{q(n-2)}{2(n-1)}}$$

$$\lesssim \Big(\int \Big(\fint_{B(a,K_n)} \prod_{i=1}^{n-1} |\widetilde{T}_\alpha^{n-1} f(\xi - \zeta_i)|^{\frac{2}{n-2}} d\xi\Big) \prod_{i=1}^{n-1} \theta_{K_{n-1}}(\zeta_i) d\zeta_1 \dots d\zeta_{n-1}\Big)^{\frac{q(n-2)}{2(n-1)}}.$$

At this point observe that since ϕ used in Lemma 14 is a Schwarz function, we see that it is possible to apply Lemma 14 to the inner integral and get the result

$$\lesssim K_n^\epsilon \Big(\int \prod_{i=1}^{n-1} \Big(\sum_{\alpha_i} (c_{\alpha_i}^n)^2 (\xi - \zeta_i) d\xi\Big)^{\frac{1}{n-2}} \prod_{i=1}^{n-1} \theta_{K_{n-1}}(\zeta_i) d\zeta_1 \dots d\zeta_{n-1}\Big)^{\frac{q(n-2)}{2(n-1)}}$$

where the summation is over all α_i with $S^n_{\alpha_i} \cap \hat{V}_{n-1} \neq \emptyset$. We rewrite this as

$$K_n^\epsilon \Big(\int \big(\sum_{\alpha_i} (c^n_{\alpha_i})^2 (\xi - \zeta_i) d\xi \big)^{\frac{1}{n-2}} \theta_{K_{n-1}}(\zeta) d\zeta \Big)^{\frac{q(n-2)}{2}}$$

$$\lesssim K_n^\epsilon \Big(\int \big(\sum_{\alpha_i} (c^n_{\alpha_i})^2 (\xi - \zeta) d\xi \big) \theta_{K_{n-1}}(\zeta) d\zeta \Big)^{\frac{q}{2}}$$

As $\theta_{K_{n-1}}$ decays faster than θ_{K_n} the convolution with θ cannot increase the value of c^n_α at a point. Using this and the Hölder inequality

$$\lesssim K_n^\epsilon \sum_{\alpha_i} (c^n_{\alpha_i})^2(\xi) \big)^{\frac{q}{2}} \lesssim K_n^{\epsilon + (n-2)(\frac{q}{2}-1)} \sum_{\alpha_i} (c^n_{\alpha_i})^q (\xi).$$

As we have exploited the fact of α being restricted, we now allow our summation to be over all α. We thus write

$$\fint_{B(a,K_n)} \big(\prod_{i=1}^{n-1} \widetilde{c}^{n-1}_{\alpha_i}(\xi) \big)^{\frac{q}{n-1}} d\xi \lesssim K_n^{\epsilon + (n-2)(\frac{q}{2}-1)} \sum_{\alpha_i} (c^n_{\alpha_i} f(\xi))^q.$$

Integrating both sides over $B(0, R)$ gives

$$\int_{B_R} \big(\prod_{i=1}^{n-1} \widetilde{c}^{n-1}_{\alpha_i}(\xi) \big)^{\frac{q}{n-1}} d\xi \lesssim K_n^{\epsilon + (n-2)(\frac{q}{2}-1)} \int_{B_R} \sum_{\alpha_i} (c^n_{\alpha_i} f(\xi))^q d\xi \tag{3.14}$$

To the individual c^n_α we apply the method of rescaling much the same way as in (3.9) to obtain

$$\|c^n_\alpha\|^q_{L^q(B_R)} \lesssim K_n^{n+1-q(n-1)} (A_q(R/K_n))^q.$$

and thus

$$\int_{B_R} \big(\prod_{i=1}^{n-1} \widetilde{c}^{n-1}_{\alpha_i}(\xi) \big)^{\frac{q}{n-1}} d\xi \lesssim K_n^{\epsilon + (n-2)(\frac{q}{2}-1) + 2n - q(n-1)} (A_q(R))^q.$$

Now if we return to (3.12) and integrate over all $a \in B_{2.1}$ we will obtain

$$\int_{B_{2.1}} \big| \int_{\{y:\mathrm{dist}(y,\hat{V}_{n-1})\lesssim\frac{1}{K_n}\}} f(y)e^{-2\pi i x\cdot\xi}d\sigma(y)\big|^q d\xi \lesssim K_{n-1}^{(2n-3)q} \sum_{\alpha_1,\ldots,\alpha_{n-1}} \int_{B_R} \big(\prod_{i=1}^{n-1} \widetilde{c}_{\alpha_i}^{n-1}(\xi)\big)^{q/n-1} d\xi.$$

We already have estimated a single integral on the right hand side. Summing over all such integrals we obtain

$$\lesssim K_{n-1}^{10n^2} K_n^{\epsilon+(n-2)(\frac{q}{2}-1)+2n-q(n-1)} (A_q(R))^q.$$

It is clear that such a term is acceptable in our inductive argument if

$$(n-2)(\frac{q}{2}-1)+2n-q(n-1) < 0$$

and if ϵ is chosen appropriately . This gives the condition $q > \frac{2n+4}{n}$. Combining this with the exponent coming from the multilinear method we see that we should have

$$q > \min(\frac{2(n-1)}{n-2}, \frac{2(n+2)}{n}).$$

Now how the rest of the argument proceeds is more or less clear. We already have obtained two disjoint sets $B_{1.1}, B_{2.1}$. We will similarly obtain sets $B_{3.1}, \ldots, B_{n-1.1}, B_{n-1.2}$ satisfying

$$B_R = B_{n-1.2} \cup \Big(\bigcup_{j=1}^{n-1} B_{n-1.1}\Big)$$

with all sets on the right hand side disjoint. Then we will obtain

$$\begin{aligned}\int_{B_R} \big| \int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\big|^q d\xi \leq & \sum_{j=1}^{n-1} \int_{B_{j.1}} \big| \int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\big|^q d\xi \\ & + \int_{B_{n-1.2}} \big| \int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\big|^q d\xi,\end{aligned} \tag{3.15}$$

and evaluate these terms. Although how to achieve this all is rather similar to what we already did above, we will further describe how to obtain the set $B_{3.1}$ and how to evaluate integrals on $B_{3.1}, B_{n-1.2}$ in (3.22) to make illuminate certain points.

Having handled ξ satisfying **2.1**, we assume ξ satisfies **2.2**. Notice that we already have

$$B_R = B_{1.1} \cup B_{2.1} \cup B_{2.2}.$$

where sets on the right hand side disjoint. For $\xi \in B_{2.2}$ we have

$$\begin{aligned} \Big|\int_S f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big| \leq \Big| &\int_{\{y:\mathrm{dist}(y,\widehat{V}_{n-2})\lesssim \frac{1}{K_{n-1}}\}} f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big| \\ + \Big| &\int_{\{y:\mathrm{dist}(y,\widehat{V}_{n-2})\gtrsim \frac{1}{K_{n-1}}\}} f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big|, \end{aligned} \tag{3.16}$$

and the set $\{y : \mathrm{dist}(y, \widehat{V}_{n-2}) \gtrsim \frac{1}{K_{n-1}}\}$ can be written as a union of disjoint sets

$$E_1(\xi) := \{y : \mathrm{dist}(y, \widehat{V}_{n-1}) \lesssim \frac{1}{K_n},\ \mathrm{dist}(y, \widehat{V}_{n-2}) \gtrsim \frac{1}{K_{n-1}}\},$$

$$E_2(\xi) := \{y : \mathrm{dist}(y, \widehat{V}_{n-1}) \gtrsim \frac{1}{K_n},\ \mathrm{dist}(y, \widehat{V}_{n-2}) \gtrsim \frac{1}{K_{n-1}}\},$$

as these sets depend on ξ. So

$$\begin{aligned} \Big|\int_{\{y:\mathrm{dist}(y,\widehat{V}_{n-2})\gtrsim \frac{1}{K_{n-1}}\}} f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big| \lesssim \Big| &\int_{E_1(\xi)} f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big| \\ + \Big| &\int_{E_2(\xi)} f(y)e^{-2\pi i y\cdot\xi} d\sigma(y)\Big|. \end{aligned} \tag{3.17}$$

We now use a $1/K_{n-2}$ partitioning of S to partition $\{y : \mathrm{dist}(y, \widehat{V}_{n-2}) \lesssim \frac{1}{K_{n-1}}\}$, and define operators $T_\alpha^{n-2}, \widetilde{T}_\alpha^{n-2}$ and functions $c_\alpha^{n-2}, \widetilde{c}_\alpha^{n-2}$ just as before. Then we obtain two subcases **3.1**, **3.2** and a partitioning of $B_{2.2}$ into $B_{3.1}$, $B_{3.2}$ analogously. Having obtained $B_{3.1}$ as desired, we now estimate the integral on it in (3.22). We of course have

$$\int_{B_{3.1}} \Big|\int_S f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \lesssim \int_{B_{3.1}} \Big|\int_{\{y:\mathrm{dist}(y,\widehat{V}_{n-2})\lesssim\frac{1}{K_{n-1}}\}} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi$$
$$+ \int_{B_{3.1}} \Big|\int_{E_1(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \tag{3.18}$$
$$+ \int_{B_{3.1}} \Big|\int_{E_2(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi.$$

Since ξ satisfies **1.2** and **2.2** we have

$$\Big|\int_{E_1(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \lesssim \sum_\alpha |\widetilde{T}_\alpha^{n-1}(\xi)| \lesssim \sum_\alpha \widetilde{c}^{n-1}(\xi)$$

where α are such that $S_\alpha^{n-1}\cap E_1(\xi)\neq\emptyset$. Thus this last sum satisfies

$$\lesssim K_{n-1}^{n-2}\cdot\frac{1}{K_{n-1}^{n-1}}\max_\alpha \widetilde{c}_\alpha^{n-1}(\xi).$$

Then

$$\int_{B_{3.1}} \Big|\int_{E_1(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \lesssim \frac{1}{K_{n-1}^q}\sum_\alpha\int_{B_{3.1}}\big(\widetilde{c}_\alpha^{n-1}(\xi)\big)^q d\xi$$
$$\lesssim \frac{1}{K_{n-1}^q}\sum_\alpha \|\widetilde{c}_\alpha^{n-1}\|_{L^q(B_R)}^q \tag{3.19}$$

where the last two sums are over all α. Here one uses scaling as before -although this time caps leading to $\widetilde{c}_\alpha^{n-1}$ have a peculiar shape, this really does not matter, we can multiply the function on the cap by cap's characteristic function, and pass to any larger cap.- We see that for any $q>2$ we get

$$\int_{B_{3.1}} \Big|\int_{E_1(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big|^q d\xi \lesssim \frac{1}{K_{n-1}^q}(A_q(R/K_{n-1}))^q \lesssim \frac{1}{K_{n-1}^q}(A_q(R))^q.$$

Similarly

$$\Big|\int_{E_2(\xi)} f(y)e^{-2\pi i y\cdot\xi}d\sigma(y)\Big| \lesssim \sum_\alpha |T_\alpha^{n-1}(\xi)| \lesssim \sum_\alpha c_\alpha^{n-1}(\xi)$$

with α satisfying $E_2(\xi) \cap S_\alpha^n \neq \emptyset$. Then the last sum satisfies

$$\lesssim K_n^{n-1} \cdot \frac{1}{K_n^n} \max_\alpha c_\alpha^n(\xi).$$

So

$$\begin{aligned}\int_{B_{3.1}} \Big| \int_{E_2(\xi)} f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi &\lesssim \frac{1}{K_n^q} \sum_\alpha \int_{B_{3.1}} \big(c_\alpha^n(\xi)\big)^q d\xi \\ &\lesssim \frac{1}{K_n^q} \sum_\alpha \|c_\alpha^n\|_{L^q(B_R)}^q\end{aligned} \tag{3.20}$$

with last two sums being over all α. Then rescaling gives for $q > 2$

$$\int_{B_{3.1}} \Big| \int_{E_2(\xi)} f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi \lesssim \frac{1}{K_n^q} (A_q(R/K_n))^q \lesssim \frac{1}{K_n^q} (A_q(R))^q.$$

Thus remains the term

$$\int_{B_{3.1}} \Big| \int_{\{y : \mathrm{dist}(y, \widehat{V}_{n-2}) \lesssim \frac{1}{K_{n-1}}\}} f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi$$

which we estimate just as in case **2.1** to obtain.

$$q \geq \min(2\frac{n-2}{n-3},\ 2 + \frac{4}{n+1}).$$

This process runs in the same vein until we obtain cases **n-1.1**, **n-1.2**, and thus the sets $B_{n-1.1}, B_{n-1.2}$. There is no difference in handling the integral on $B_{n-1.1}$, but there is a small difference for the integral on $B_{n-1.2}$. In this case we do not proceed to obtain further subcases, and write

$$\begin{aligned}\int_{B_{3.1}} \Big| \int_S f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi \lesssim &\int_{B_{3.1}} \Big| \int_{\{y : \mathrm{dist}(y, \widehat{V}_1) \lesssim \frac{1}{K_2}\}} f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi \\ &+ \int_{B_{3.1}} \Big| \int_{\{y : \mathrm{dist}(y, \widehat{V}_1) \gtrsim \frac{1}{K_2}\}} f(y) e^{-2\pi i y \cdot \xi} d\sigma(y) \Big|^q d\xi.\end{aligned}$$

Now, firstly, the set $\{y : \mathrm{dist}(y, \widehat{V}_1) \lesssim \frac{1}{K_2}\}$ is neighborhood of a point on S, and there is no multilinearity to exploit. We thus proceed directly to scaling for the first term, and obtain the bound

$$\int_{B_{3.1}} \Big| \int_{\{y:\mathrm{dist}(y,\widehat{V}_1)\lesssim \frac{1}{K_2}\}} f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi \lesssim K_{n-2}^{2n+q(1-n)+\epsilon} (A_q(R))^q$$

and this gives the exponent $q > \frac{2n}{n-1}$. The second integral can be handled easily by dividing the set $\{y : \mathrm{dist}(y, \widehat{V}_1) \gtrsim \frac{1}{K_2}\}$ appropriately into subset as we did above. Ultimately the second integral can be bounded by a term

$$C \frac{1}{K_2^q} (A_q(R))^q$$

for any $q > 2$, and such a term is harmless for induction.

Thus we have estimated all terms in (3.22). For the integral on $B_{1.1}$ and $B_{n-1.2}$ we have the condition $q > \frac{2n}{n-1}$. For any integral on $B_{j.1}$ with $2 \leq j \leq n-1$ we have

$$q > \min\left(2\frac{n-j+1}{n-j},\ 2 + \frac{4}{n+j-2}\right).$$

So we see that for higher levels of multilinearity, multilinear method is better, and for lower levels rescaling gives better exponents. At about $j = 2n/3$ the method of rescaling becomes better , but the exact value of j depends on the value of l in the equation $n = 3k + l$. Also the exponent we get when we combine all these conditions on q is taken at about this value of j. We can easily calculate that if $n = 3k$ then the final condition on the exponent q is

$$q > \frac{8n+6}{4n-3} \tag{3.21}$$

and this exponent comes from the term $j = k + 1$, from the method of rescaling. For

$n = 3k + 1$ this is

$$q > \frac{2n+1}{n-1}$$

and comes from $j = k + 1$ using the multilinearity. For $n = 3k + 2$ the exponent is

$$q > \frac{4n+4}{2n-1}$$

and this comes from both $j = k+1$ by multilinearity, and $j = k+2$ by rescaling. In the next section we will limit ourselves to case $n = 3k$ and improve the exponent given by (3.21). We explain the reason why we are able to do this in this case and not the other two. As written above, the condition (3.21) comes from the term $j = k + 1$, namely for $j = k + 1$

$$\frac{8n+6}{4n-3} = \min(2\frac{n-j+1}{n-j},\ 2 + \frac{4}{n+j-2}) = 2 + \frac{4}{n+j-2}$$

The term $2(n - j + 1)/(n - j)$ coming from multilinear estimates is strictly larger than this minimum that comes from rescaling. This will be used for a pigeonholing argument that allows finer estimates.

3.3 Iteration of the Bourgain-Guth Argument

From now on we restrict ourselves to those n with $n = 3k,\ k \in \mathbb{N}$. We recall the expression

$$\begin{aligned} \int_{B_R} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi \leq & \sum_{j=1}^{n-1} \int_{B_{j.1}} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi \\ & + \int_{B_{n-1.2}} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi. \end{aligned} \tag{3.22}$$

We estimated these integrals one by one to obtain our final result. We saw that the worst exponent comes from the integral on $B_{k+1.1}$, that this exponent is obtained via the method of rescaling, and that the integral on $B_{k.1}$ is the last one for which the better exponent comes from the multilinear method. We first observe that if one have $m+1$ level of multilinearity then one also have m level of multilinearity. So a careful consideration of our section 3.2 reveals that we can run the same argument starting with $2k+1$ level of multilinearity instead of n level, and obtain the same result. In this case we partition B_R into

$$B_R = B_{n-1.2} \cup \Big(\bigcup_{j=k}^{n-1} B_{j.1} \Big).$$

Measure of a set here coming from the same level of multilinearity as in the partition before may be different, but this is not of any importance, as we have never used this information. So instead of (3.22) we have

$$\begin{aligned} \int_{B_R} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi \leq & \sum_{j=k}^{n-1} \int_{B_{j.1}} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi \\ & + \int_{B_{n-1.2}} \Big| \int_S f(y) e^{-2\pi i y\cdot\xi} d\sigma(y) \Big|^q d\xi. \end{aligned} \tag{3.23}$$

This reduction will simplify the iteration process that will be described below so significantly that it will allow calculation of improvement for dimensions $n = 3r$ for $r > 2$.

Let $1 \leq m \leq 2k+1$ denote the level of linearity , and so we may define $B_m = B_{j.1}$ for $m + j = n + 1$, and $B_m = B_{n-1.2}$ for $m = 1$. We observe from section 3.2 that for $\xi \in B_m,\ 2 \leq m \leq 2k$ we have

$$|Tf(\xi)| \lesssim C(K_m) \Big(\prod_{i=1}^{m} \widetilde{c}^{\,m}_{\alpha_i}(\xi) \Big)^{1/m-1} + \sum_{i=m+1}^{2k} \frac{1}{K_i} \max_{\alpha} \widetilde{c}^{\,i}_{\alpha}(\xi) + \frac{1}{K_{2k+1}} \max_{\alpha} c^{2k+1}_{\alpha}(\xi),$$

if $m = 1$,

$$|Tf(\xi)| \lesssim \big(\widetilde{c}^{\,2}_{\alpha_i}(\xi) \big)^{1/m-1} + \sum_{i=m+1}^{2k} \frac{1}{K_i} \max_{\alpha} \widetilde{c}^{\,i}_{\alpha}(\xi) + \frac{1}{K_{2k+1}} \max_{\alpha} c^{2k+1}_{\alpha}(\xi),$$

while if $\xi \in B_{m.1},\ m = 2k+1$ we have

$$|Tf(\xi)| \lesssim C(K_m)\Big(\prod_{i=1}^{m} \widetilde{c}^{m}_{\alpha_i}(\xi)\Big)^{1/m}.$$

In all cases, of course, $c^m_{\alpha_i}$ terms are appropriately chosen. Thus for any $\xi \in B_R$ we have

$$\begin{aligned}|Tf(\xi)| \lesssim \sum_{m=2}^{2k} C(K_m)\Big(\prod_{i=1}^{m} \widetilde{c}^{m}_{\alpha_i}(\xi)\Big)^{1/m} + C(K_{2k+1})\Big(\prod_{i=1}^{2k+1} \widetilde{c}^{2k+1}_{\alpha_i}(\xi)\Big)^{1/2k+1} \\ + \widetilde{c}^{2}_{\alpha}(\xi) + \sum_{i=2}^{2k} \frac{1}{K_i}\max_{\alpha} \widetilde{c}^{i}_{\alpha}(\xi) + \frac{1}{K_{2k+1}}\max_{\alpha} c^{2k+1}_{\alpha}(\xi).\end{aligned} \tag{3.24}$$

We can replace terms on the second line by $c_\alpha(\xi)$, denoting maximum of terms on that line. Thus we have

$$|Tf(\xi)| \lesssim \sum_{m=2}^{2k} C(K_m)\Big(\prod_{i=1}^{m} \widetilde{c}^{m}_{\alpha_i}(\xi)\Big)^{1/m} + C(K_{2k+1})\Big(\prod_{i=1}^{2k+1} \widetilde{c}^{2k+1}_{\alpha_i}(\xi)\Big)^{1/2k+1} + c_\alpha(\xi). \tag{3.25}$$

Our aim is to iterate this decomposition for smaller caps that give rise to last two terms on the right-hand side, until we reach the scale $R^{-1/2}$, which allows one to use more direct methods like conservation of energy. But before this we need to make some replacements.

We can obtain a function ϕ_m for $2 \le m \le 2k$ that is approximately constant on balls of radius 1 and that satisfies

$$\prod_{i=1}^{m} (\widetilde{c}^{m}_{\alpha_i}(\xi))^{1/m} = \phi_m \cdot \big(\sum_{\alpha} (c^{m+1}_{\alpha}(\xi))^2\big)^{1/2}. \tag{3.26}$$

To obtain such a function we simply divide the left hand side by the right hand side when the right-hand side is not zero, when it is zero we just set $\phi_m = 0$. Since both sides are approximately constant on much larger balls, the condition of ϕ_m being approximately constant on balls of radius 1 is satisfied. These functions satisfy the following additional property:

$$\Big(\fint_{B(a,K_{m+1})}(\phi_m)^{\frac{2m}{m-1}}\Big)^{\frac{m-1}{2m}} \lesssim K_{m+1}^{\epsilon^5}. \tag{3.27}$$

To see this observe that the process we used to bound the left hand side of (3.13) already gives

$$\Big[\fint_{B(a,K_{m+1})}\prod_{i=1}^{m}(\widetilde{c}_{\alpha_i}^{m}(\xi))^{2/m-1}dx\Big]^{\frac{m-1}{2m}} \lesssim K_{m+1}^{\epsilon^5}\Big(\sum_{\alpha}(c_\alpha^{m+1}(\xi))^2\Big)^{1/2}.$$

From which (3.27) follows. Since $m \leq 2k$, simply using the Hölder inequality gives

$$\Big(\fint_{B(a,K_{m+1})}(\phi_m)^{\frac{4k}{2k-1}}\Big)^{\frac{2k-1}{4k}} \lesssim K_{m+1}^{\epsilon^5}.$$

Inserting this replacement back to (3.25) to obtain

$$\begin{aligned}|Tf(\xi)| \lesssim C(K_{2k+1})\Big(\prod_{i=1}^{2k+1}c_{\alpha_i}^{2k+1}(\xi)\Big)^{1/2k+1} &+ \sum_{m=2}^{2k}K_{m+1}^{\epsilon^5}\phi_m\cdot\Big(\sum_{\alpha}(c_\alpha^{m+1}(\xi))^2\Big)^{1/2} \\ &+ c_\alpha(\xi).\end{aligned} \tag{3.28}$$

The iteration will be performed on the second and third term in the sum above. We will illustrate this for a single c_α^{m+1} from the second term on the right-hand side, for any other term in (3.28), or any other term that will arise as we decompose, the process is the same. This iteration is an entirely elementary, but rather cumbersome process. To simplify we will makes some modifications to (3.28). To this end we introduce

$$a_\alpha^m(\xi) := \int |T_\alpha^m f(\xi-\zeta)|\theta_{K_m}(\zeta)d\zeta,$$

for $2 \leq m \leq n$. Using the Hölder inequality we have

$$c_\alpha^m(\xi) \lesssim a_\alpha^m(\xi).$$

So from (3.28) we pass to

$$|Tf(\xi)| \lesssim C(K_{2k+1})\Big(\prod_{i=1}^{2k+1} a_{\alpha_i}^{2k+1}(\xi)\Big)^{1/2k+1} + \sum_{m=2}^{2k} K_{m+1}^{\epsilon^5}\phi_m \cdot \big(\sum_{\alpha}(a_\alpha^{m+1}(\xi))^2\big)^{1/2} \\ + a_\alpha(\xi). \tag{3.29}$$

We introduce another variant of c_α^m that averages over tubes rather than balls. We will use parabolic rescaling as we iterate our decomposition, and the nature of which dictates use of tubes. Recall the function η that we used in section 3.2 to define c_α^m. For a cap S_α of radius $1/K_m$ rescale this function to make it concentrated on a tube centered at the origin oriented along the normal e_α of the cap S_α. Also define the smooth cutoff

$$\beta_{K_m,\alpha}(\xi) := \Big(1 + \frac{|\xi \cdot e_\alpha|}{K_m^2} + \frac{|\xi - \xi \cdot e_\alpha|}{K_m}\Big)^{-(10n)^{2n-m}\epsilon^{-n-10}}.$$

Then using these we define

$$b_\alpha^m(\xi) := \int |T_\alpha^m f(\xi - \zeta)| \beta_{K_m,\alpha}(\zeta) d\zeta,$$

and obtain from slower decay of β_{K_m} than θ_{K_m}

$$|T_\alpha^m f(\xi)| \lesssim c_\alpha^m(\xi) \lesssim a_\alpha^m(\xi) \lesssim b_\alpha^m(\xi). \tag{3.30}$$

Thus we can write

$$|Tf(\xi)| \lesssim C(K_{2k+1})\Big(\prod_{i=1}^{2k+1} b_{\alpha_i}^{2k+1}(\xi)\Big)^{1/2k+1} + \sum_{m=2}^{2k} K_{m+1}^{\epsilon^5}\phi_m \cdot \big(\sum_{\alpha}(b_\alpha^{m+1}(\xi))^2\big)^{1/2} \\ + b_\alpha(\xi). \tag{3.31}$$

Thus instead of decomposing c_α^{m+1} we may decompose b_α^{m+1}. Since b_α^{m+1} are approximately constant on balls of size $1/K_{m+1}$ we can write

$$b_\alpha^{m+1}(\xi) \approx \int_{B(\xi,1)} b_\alpha^{m+1}(\zeta)d\zeta = \int_{B(\xi,1)} \int |T_\alpha^{m+1} f(\zeta-\eta)|\beta_{K_{m+1},\alpha}(\eta)d\eta d\zeta. \tag{3.32}$$

We expand the term T_α^{m+1} using the projection $U_\alpha = U_\alpha(x_\alpha, 1/K_{m+1})$ of the cap S_α, and a parametrization of S_α on this projection

$$= \int_{B(\xi,1)} \int | \int_{U_\alpha} f(x)e^{-2\pi i[x\cdot(\overline{\zeta}-\overline{\eta})+(|x|^2+\mathcal{O}(|x|^3))(\zeta_n-\eta_n)]}dx|\beta_{K_{m+1},\alpha}(\eta)d\eta d\zeta.$$

From this point onwards we apply change of variables repeatedly. We first translate the cap to the origin

$$\lesssim \int_{B(\xi,1)} \int | \int_{U_\alpha - x_\alpha} f(x+x_\alpha)e^{-2\pi i\Phi(x,\zeta,\eta)}dx|\beta_{K_{m+1},\alpha}(\eta)d\eta d\zeta$$

with

$$\Phi(x,\zeta,\eta) = x\cdot(\overline{\zeta}-\overline{\eta}+(2x_\alpha+\ldots)(\zeta_n-\eta_n)) + (|x|^2(1+\ldots)+\mathcal{O}(|x|^3))(\zeta_n-\eta_n).$$

We now apply appropriate change of variables for two outer integrals to obtain

$$\lesssim \int_{B(\xi',C)} \int | \int_{U_\alpha - x_\alpha} f(x+x_\alpha)e^{-2\pi i[x\cdot(\overline{\zeta}-\overline{\eta})+(|x|^2+\mathcal{O}(|x|^3))(\zeta_n-\eta_n)]}dx|\beta_{K_{m+1},0}(\eta)d\eta d\zeta$$

where $\beta_{K_{m+1},0}$ represents the smooth cutoff for the cap we obtained by translating U_α to the origin. We first apply parabolic rescaling to x with factor $1/K_{m+1}$, then to ζ,η variables with K_{m+1}. We obtain

$$\lesssim K_{m+1}^{n+3} \int_B \int | \int_U f(\frac{x}{K_{m+1}}+x_\alpha)e^{-2\pi i\Psi(x,\zeta,\eta)}dx|\beta_{K_{m+1},0}(K_{m+1}\overline{\eta}, K_{m+1}^2\eta_n)d\eta d\zeta$$

where U a domain of size comparable to 1,

$$\Psi(x,\zeta,\eta) = x\cdot(\overline{\zeta}-\overline{\eta}) + (|x|^2 + \frac{1}{K_{m+1}}\mathcal{O}(|x|^3))(\zeta_n - \eta_n),$$

and

$$B = \{\zeta : |\overline{\zeta}-\overline{\xi}'| \lesssim \frac{1}{K_{m+1}}, \quad |\zeta_n - \xi_n'| \lesssim \frac{1}{K_{m+1}^2}.\}$$

At this point we can apply the decomposition (3.29) to obtain

$$\begin{aligned}
&\lesssim K_{m+1}^{n+3}C(K_{2k+1})\int_B\int\Big(\prod_{i=1}^{2k+1} a_{\tau_i}^{2k+1}(\zeta-\eta)\Big)^{\frac{1}{2k+1}}\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n)d\eta d\zeta \\
&+ K_{m+1}^{n+3}\sum_{l=2}^{2k}K_{l+1}^{\epsilon}\int_B\int\phi_l\cdot\big(\sum_\tau(a_\tau^{l+1}(\zeta-\eta))^2\big)^{1/2}\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n)d\eta d\zeta \\
&+ K_{m+1}^{n+3}\int_B\int a_\tau(\zeta-\eta)\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n)d\eta d\zeta
\end{aligned}$$

All three terms are handled similarly, so we will only demonstrate in detail how to handle the second one, which is the most complicated of all three. It is clear that understanding the term

$$\int_B\int\phi_l\cdot\big(\sum_\tau(a_\tau^{l+1}(\zeta-\eta))^2\big)^{1/2}\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n)d\eta d\zeta$$

is sufficient for this. This term can be written as

$$= \int_B\int\phi_l\cdot\big(\sum_\tau(a_\tau^{l+1}(\zeta-\eta)\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n))^2\big)^{1/2}d\eta d\zeta. \tag{3.33}$$

We have

$$a_\tau^{l+1}(\zeta-\eta) = \int|T_\tau^{l+1}g(\zeta-\eta-\nu)|\theta_{K_{l+1}}(\nu)d\nu$$

with $g(x) = f(x/K_{l+1}+x_\alpha)$. Due to nature of parabolic scaling we have

$$\beta_{K_{m+1},0}(K_{m+1}\overline{\eta},K_{m+1}^2\eta_n) = \frac{1}{K_{m+1}^{n+1}}(1+|\eta|)^{-(10n)^{2n-m-1}\epsilon^{-n-10}} \le \frac{1}{K_{m+1}^{n+1}}(1+|\eta|)^{-(10n)^n\epsilon^{-n-10}},$$

and

$$\theta_{K_{l+1}}(\nu) \leq \frac{1}{K_{l+1}^n}(1+\frac{|\nu|}{K_{l+1}})^{-(10n)^n\epsilon^{-n-10}-5n}.$$

We write

$$a_\tau^{l+1}(\zeta-\eta)\beta'_{K_{m+1},0}(\eta) = \int |T_\tau^{l+1}g(\zeta-\eta-\nu)|\theta_{K_{l+1}}(\nu)\beta_{K_{m+1},0}(K_{m+1}\overline{\eta}, K_{m+1}^2\eta_n)d\nu.$$

We insert what we obtained above to the right-hand side

$$\leq (1+|\eta|)^{-5n}\int |T_\tau^{l+1}g(\zeta-\eta-\nu)|\frac{1}{K_{m+1}^{n+1}K_{l+1}^n}(1+\frac{|\eta+\nu|}{K_{l+1}})^{-(10n)^n\epsilon^{-n-10}+5n}d\nu.$$

For fixed η we perform a change of variables $\mu=\eta+\nu$ we obtain

$$\leq \frac{(1+|\eta|)^{-5n}}{K_{m+1}^{n+1}}\int |T_\tau^{l+1}g(\zeta-\mu)|\frac{1}{K_{l+1}^n}(1+\frac{|\mu|}{K_{l+1}})^{-(10n)^n\epsilon^{-n-10}+5n}d\mu.$$

We insert this back to (3.33) to obtain

$$\lesssim \int_B \Big(\sum_\tau \big(\int |T_\tau^{l+1}g(\zeta-\mu)|\frac{1}{K_{l+1}^nK_{m+1}^{n+1}}(1+\frac{|\mu|}{K_{l+1}})^p d\mu\big)^2\Big)^{\frac{1}{2}}d\zeta \cdot \fint_B\int (1+|\eta|)^{-5n}\phi_l d\eta d\zeta$$

with $p=-(10n)^n\epsilon^{-n-10}+5n$. We will handle these two factors separately. In the first term we have a smooth cutoff for a ball of radius K_{l+1}. By reducing decay a little we can pass to cutoff for a tube of radius K_{l+1} and length K_{l+1}^2

$$\Big(\sum_\tau\int_B\big(\int |T_\tau^{l+1}g(\zeta-\mu)|\frac{1}{K_{l+1}^{n+1}K_{m+1}^{n+1}}(1+\frac{|\mu-\mu\cdot e_\tau|}{K_{l+1}}+\frac{|\mu\cdot e_\tau|}{K_{l+1}^2})^{-(10n)^n\epsilon^{-n-10}+10n}d\mu\big)^2\Big)^{1/2}d\zeta$$

where e_τ is the normal of the cap S_τ^{l+1}. Let $\beta'_{K_{l+1},\tau}$ be the cutoff in this last statement. Then the integral can be written as

$$\lesssim \frac{1}{K_{m+1}^{n+1}}\int_B\int |\int_{U_\tau} f(\frac{x}{K_{m+1}}+x_\alpha)e^{-2\pi i[x\cdot(\overline{\xi}-\overline{\mu})+(|x|^2+\frac{1}{K_{m+1}}\mathcal{O}(|x|^3))(\xi_n-\mu_n)]}dx|\beta'_{K_{l+1},\tau}(\mu)d\mu d\zeta$$

We now will return to our initial domain U_α. To this end we apply parabolic rescaling to x by K_{m+1}, and to ξ, μ by $1/K_{m+1}$. Thus we obtain

$$\lesssim K_{m+1}^{-n-3} \int_{B(\xi',C)} \int |\int_{U_{\alpha_\tau}-x_\alpha} f(x+x_\alpha) e^{-2\pi i[x\cdot(\overline{\zeta}-\overline{\mu})+(|x|^2+\mathcal{O}(|x|^3))(\zeta_n-\mu_n)]} dx| \beta'_{K_{l+1}K_{m+1},0_\tau}(\mu) d\mu d\zeta$$

with $U_{\alpha_\tau} - x_\alpha$ denoting the cap arising from U_τ after scaling, and

$$\beta'_{K_{l+1}K_{m+1},0_\tau}(\mu) = \frac{1}{K_{m+1}^{n+1}} \beta'_{K_{l+1},\tau}(\overline{\mu}/K_{m+1}, \mu_n/K_{m+1}^2).$$

We have $U_{\alpha_\tau} - x_\alpha$ a cap inside $U_\alpha - x_\alpha$, and after translating it will be inside U_α. We are now ready to translate to x_α through change of variables first in ξ, μ, and then in x to obtain

$$\lesssim \int_{B(\xi,1)} \int |\int_{U_{\alpha_\tau}} f(x) e^{-2\pi i[x\cdot(\overline{\zeta}-\overline{\mu})+(|x|^2+\mathcal{O}(|x|^3))(\zeta_n-\mu_n)]} dx| \beta'_{K_{l+1}K_{m+1},\alpha_\tau}(\mu) d\mu d\zeta$$

$$\lesssim \int |\int_{U_{\alpha_\tau}} f(x) e^{-2\pi i[x\cdot(\overline{\xi}-\overline{\mu})+(|x|^2+\mathcal{O}(|x|^3))(\xi_n-\mu_n)]} dx| \beta'_{K_{l+1}K_{m+1},\alpha_\tau}(\mu) d\mu$$

with $\beta'_{K_{l+1}K_{m+1},\alpha_\tau}$ is the function $\beta'_{K_{l+1}K_{m+1},0_\tau}$ rotated to have direction of the normal of the cap U_{α_τ}. This function is a natural smooth tubular cutoff for U_{α_τ}. Thus we obtained a term that is similar to the one we started with, except that this time the cap is of smaller size, and the decay of smooth cutoff is a little slower. But as we started with very fast decay, even after reaching to scale $R^{-1/2}$ we will have enough decay. Thus this final term is acceptable for us. But, of course, we also need to show that the second factor

$$⨍_B \int (1+|\eta|)^{-5n} \phi_l(\zeta-\eta) d\eta d\zeta$$

is also acceptable. We perform change of variables for both variables, we apply parabolic rescaling to both η and ζ by $1/K_{m+1}$:

$$\lesssim \frac{1}{K_{m+1}^{n+1}} \int_{B(\xi',C)} \int (1+\frac{|\overline{\eta}|}{K_{m+1}}+\frac{|\eta_n|}{K_{m+1}^2})^{-5n}\phi_l(\frac{\overline{\zeta}-\overline{\eta}}{K_{m+1}},\frac{\zeta_n-\eta_n}{K_{m+1}^2})d\eta d\zeta.$$

Both functions inside integrals are approximately constant on tubular domains oriented along the vector $(0,\ldots,0,1)$. Applying the change of variables that we used above for these variables after translating caps makes them constant on tubes oriented along e_α, the normal of the cap S_α. Letting $\beta''_{K_{m+1},\alpha}$ denote the first function, and ϕ'_l the second, our integral becomes

$$\lesssim \frac{1}{K_{m+1}^{n+1}} \int_{B(\xi,1)} \int \beta''_{K_{m+1},\alpha}(\eta)\phi'_l(\zeta-\eta)d\eta d\zeta.$$

Inside integral is approximately constant over the domain $B(\xi,1)$, thus

$$\lesssim \frac{1}{K_{m+1}^{n+1}} \int \beta''_{K_{m+1},\alpha}(\eta)\phi'_l(\xi-\eta)d\eta.$$

We define

$$\phi_{\alpha_l}(\xi) := \frac{1}{K_{m+1}^{n+1}} \int \beta''_{K_{m+1},\alpha}(\eta)\phi'_l(\xi-\eta)d\eta.$$

This function satisfies two properties that are required for the iteration process. Firstly it is approximately constant on tubes oriented along e_α with radius K_{m+1} and length K_{m+1}^2. Secondly if B is a box centered at ξ with thickness $K_{m+1}K_{l+1}$ and length $K_{m+1}^2K_{l+1}^2$ along e_{α_τ}, then

$$⨍_B \phi_{\alpha_l}(\xi)d\xi \lesssim K_{l+1}^{\epsilon^5}.$$

To see this we will use the definition of ϕ_{α_l}. We have

$$⨍_B \phi_{\alpha_l}(\xi)d\xi = \frac{1}{K_{m+1}^{n+1}} ⨍_B \int \beta''_{K_{m+1},\alpha}(\eta)\phi'_l(\xi-\eta)d\eta d\xi.$$

Applying the inverse of translation we used above to both η, ξ we get

$$\approx \frac{1}{K_{m+1}^{n+1}} \fint_{B'} \int (1 + \frac{|\overline{\eta}|}{K_{m+1}} + \frac{|\eta_n|}{K_{m+1}^2})^{-5n} \phi_l(\frac{\overline{\xi} - \overline{\eta}}{K_{m+1}}, \frac{\xi_n - \eta_n}{K_{m+1}^2}) d\eta d\xi$$

where B' is now a box of the same size as B, oriented along the normal of the translation of the cap α_τ. We apply parabolic scaling to both terms to obtain

$$\approx \fint_{B''} \int (1 + |\eta|)^{-5n} \phi_l(\xi - \eta) d\eta d\xi = \int (1 + |\eta|)^{-5n} \big(\fint_{B''} \phi_l(\xi - \eta) d\xi \big) d\eta$$

where B'' is a box of radius K_{m+1} and length K_{m+1}^2 oriented along the normal of the cap S_τ that emerges from the parabolic scaling. That the inner integral satisfies $\lesssim K_{l+1}^{\epsilon^5}$ is clear. Thus we have the desired property.

Thus we see that second factor is acceptable to us as well. Hence we have repeated the decomposition succesfully. Further iteration of this decompostion process is much the same as this one. However we will need to deal with the functions ϕ_{α_m} arising at each step. As we iterate we will get a product of these function, and we need to understand this product. To investigate this phenomenon, let ϕ_{τ_m} and ϕ_{ρ_l} be such functions arising consecutively, with ρ denoting the cap τ_m. So ϕ_{τ_m} is constant on boxes τ' dual to the cap τ, while ϕ_{ρ_l} is constant on boxes ρ' dual to the cap ρ. Since ρ arise from the decomposition of τ, if τ is of size δ, ρ must have size δ/K_{m+1} . Let B be a box dual to ρ_l. Since this can be covered with a small number of $K_{l+1}\rho'$ boxes, we may assume that it is a $K_{l+1}\rho'$ box We can decompose this box into ρ' boxes B_α; since ϕ_{ρ_l} is comparable to a constant on a box B_α we have

$$\int_B \phi_{\tau_m}^{4k/2k-1} \phi_{\rho_l}^{4k/2k-1} \lesssim \sum_\alpha \phi_{\rho_l}^{4k/2k-1} \Big|_{B_\alpha} \int_{B_\alpha} \phi_{\tau_k}^{4k/2k-1}$$

with

$$\phi_{\rho_l}^{4k/2k-1} \Big|_{B_\alpha}$$

denoting the constant to which this function is comparable on B_α. Since ρ denotes the cap τ_m, we have

$$\fint_{B_\alpha} \phi_{\tau_m}^{4k/2k-1} \lesssim K_{m+1}^{\epsilon^5}.$$

Hence we have

$$\begin{aligned}\sum_\alpha \phi_{\rho_l}^{4k/2k-1}\Big|_{B_\alpha} \int_{B_\alpha} \phi_{\tau_m}^{4k/2k-1} &\lesssim \sum_\alpha \int_{B_\alpha} \phi_{\rho_l}^{4k/2k-1} \fint_{B_\alpha} \phi_{\tau_m}^{4k/2k-1} \\ &\lesssim K_{m+1}^{\epsilon^5} \sum_\alpha \int_{B_\alpha} \phi_{\rho_l}^{4k/2k-1} \\ &\lesssim K_{m+1}^{\epsilon^5} \int_B \phi_{\rho_l}^{4k/2k-1} \\ &\lesssim K_{m+1}^{\epsilon^5} K_{l+1}^{\epsilon^5} |B|.\end{aligned}$$

So if B is a box dual to ρ_l, or can be covered with a small number of these boxes we have

$$\fint_B \phi_{\tau_m}^{4k/2k-1} \phi_{\rho_l}^{4k/2k-1} \lesssim K_{m+1}^{\epsilon^5} K_{l+1}^{\epsilon^5}.$$

From here we also understand that the constants of type $K_{m+1}^{\epsilon^5}$ incurred during the decomposition are always inversely proportionate to the size of caps τ_m emerging, that is a smaller cap incurs a larger constant. This means as we iterate our decomposition to reach size $R^{-1/2}$ we will not encounter excess growth of constants, for dividing into smaller caps means a smaller number of steps to reach $R^{-1/2}$, while dividing into larger caps incurs smaller constants.

After these explanations we state the final situation after iteration until the scale $R^{-1/2}$.

$$\begin{aligned}|Tf| \lesssim R^{\epsilon^3} \max_{R^{-1/2}<\delta<1} \max_{E_\delta} \Big[\sum_{\tau \in E_\delta} \big(\phi_\tau \big(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1}\big)^{1/2k+1}\big)^2\Big]^{1/2} \\ + R^{\epsilon^3} \max_{E_\gamma} \Big[\sum_{\rho \in E_\gamma} (\phi_\rho b_\rho)^2\Big]^{1/2}\end{aligned} \tag{3.34}$$

1. In the first term E_δ is a collection of δ caps that has cardinality $\lesssim \delta^{-2k+1}$, while in the second E_γ are caps of size $\frac{R^{-1/2}}{K_n} \leq \gamma \leq R^{-1/2}$.

2. In the first term $\tau_i \subset \tau$ are caps of size δ/K_{2k+1}.

3. We have for B a τ' box, K a ρ' box

$$\fint_B \phi_\tau^{4k/2k-1} \lesssim R^{\epsilon^4}, \quad \fint_K \phi_\rho^{4k/2k-1} \lesssim R^{\epsilon^4}.$$

4. The functions $b_{\tau_i}^{2k+1}$ have the form

$$b_{\tau_i}^{2k+1}(\xi) = \big(\int |T_{\tau_i}^{2k+1} f(\xi - \zeta)|^{1/2k+1} \beta_{\delta/K_{2k+1},\tau_i}(\zeta) d\zeta \big)^{2k+1}$$

where we have, for a unit normal e_{τ_i} of the cap τ_i,

$$\beta_{K_{2k+1}/\delta,\tau_i}(\xi) = \frac{1}{K_{2k+1}^{n+1}\delta^{n+1}} \big(1 + \frac{|\xi - \xi \cdot e_{\tau_i}|}{K_{2k+1}/\delta} + \frac{|\xi \cdot e_{\tau_i}|}{K_{2k+1}^2/\delta^2}\big)^{-(10n)^5}.$$

Similarly

$$b_\rho(\xi) = \int |T_\rho f(\xi - \zeta)| \beta_{\gamma,\rho}(\zeta) d\zeta, \quad \beta_{\gamma,\rho}(\xi) = \frac{1}{\gamma^{n+1}} \big(1 + \frac{|\xi - \xi \cdot e_\rho|}{\gamma} + \frac{|\xi \cdot e_\rho|}{\gamma^2}\big)^{-(10n)^5}.$$

We will deal with the terms in (3.34) separately. We start with the first term.

3.3.1 Estimates on the multilinear term

We will estimate this term first in the space L^q, with $q = \frac{4k+2}{2k}$, which is the sharp exponent given for $2k+1$ level multilinearity by the multilinear method -which is better than the one given by rescaling.- Then we will estimate it in $L^{\frac{8k+2}{4k-1}}$ which is the final exponent that our work in section 3.2 gives. We note that $\frac{4k+2}{2k} < \frac{8k+2}{4k-1}$. We will then interpolate between these exponents to improve $\frac{8k+2}{4k-1}$.

We pick an arbitrary acceptable E_δ, and consider

$$\lesssim \Big[\sum_{\tau\in E_\delta} \big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^2\Big]^{1/2}.$$

The terms $b_{\tau_i}^{2k+1}$ come from small caps inside a larger one τ. By rescaling this larger cap to scale 1 we can obtain some gain. We will illustrate this first.

$$\int_{B_R}(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{2/2k} = \int_{B_R}\Big(\prod_{i=1}^{2k+1}\int |T_{\tau_i}f(\xi-\zeta)|^{1/2k+1}_{\delta/K_{2k+1}}\beta_{K_{2k+1}/\delta,\tau_i}(\zeta)d\zeta\Big)^{4k+2/2k}d\xi$$

From this by the Hölder inequality

$$\lesssim \int_{B_R}\Big(\prod_{i=1}^{2k+1}\int |T_{\tau_i}f(\xi-\zeta)|^{1/k}\beta_{K_{2k+1}\delta,\tau_i}(\zeta)d\zeta\Big)d\xi$$

$$= \int_{B_R}\Big(\int\prod_{i=1}^{2k+1} |T_{\tau_i}f(\xi-\zeta_i)|^{1/k}\beta_{K_{2k+1}\delta,\tau_i}(\zeta_i)d\zeta_1\dots d\zeta_{2k+1}\Big)d\xi.$$

Then using Fubini's theorem

$$= \int\Big(\int_{B_R}\prod_{i=1}^{2k+1} |T_{\tau_i}f_{\zeta_i}(\xi-\zeta_i)|^{1/k}d\xi\Big)\prod_{i=1}^{2k+1}\beta_{K_{2k+1}\delta,\tau_i}(\zeta_i)d\zeta_1\dots d\zeta_{2k+1}.$$

Here f_{ζ_i} are modulations of f. We rescale the inner integral to obtain functions $|g_{\zeta_i}|\le 1$ and caps $\alpha_1,\cdots,\alpha_{2k+1}$ with size $1/K_{2k+1}$, and the linear independence condition. To these multilinear estimates apply and we obtain

$$\lesssim \delta^{3k-\frac{1}{k}}R^{\epsilon^5}$$

for the inner integral for any choice of ζ_i. Since outer terms are just smooth cutoffs used to average over certain domains the overall term too satisfies

$$\lesssim \delta^{3k-\frac{1}{k}} R^{\epsilon^5}.$$

With this at hand we proceed to our estimates. First will come simpler and coarser estimates, and the we will refine these using pigeonholing arguments. These coarser estimates will not be directly utilized for the final result, but we will illustrate the arguments that enable the finer estimates in proving them.

$$\begin{aligned}\Big[\sum_{\tau\in E_\delta}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^2\Big]^{1/2} &\le |E_\delta|^{\frac{1}{4k+2}}\Big[\sum_{\tau\in E_\delta}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^{2k+1/k}\Big]^{k/2k+1}\\ &\lesssim \delta^{\frac{-2k+1}{4k+2}}\Big[\sum_{\tau\in E_\delta}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^{2k+1}\Big]^{k/2k+1}\end{aligned}$$

At this point τ ranges over a full partition into δ- caps of S, and thus does not depend on any particular choice of E_δ. Let B stand for τ' boxes. Since $b_{\tau_i}^{2k+1}$ are constant on τ_i' boxes, we can write

$$\begin{aligned}\int_{B_R}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^{\frac{4k+2}{2k}} &\lesssim \sum_B(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{\frac{1}{k}}\Big|_B(\int_B \phi_\tau^{4k+2/2k})\\ &\lesssim \sum_B(\int_B(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{\frac{1}{k}})(\fint_B \phi_\tau^{4k+2/2k})\\ &\lesssim R^{\epsilon^4}\int_{B_R}(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{\frac{1}{k}}\\ &\lesssim R^{\epsilon^3}\delta^{3k-\frac{1}{k}}.\end{aligned} \tag{3.35}$$

Since a full partition of S into δ caps has cardinality $\approx \delta^{-3k+1}$, we have

$$\|\Big[\sum_{\tau\in E_\delta}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^2\Big]^{1/2}\|_{L^{\frac{4k+2}{2k}}(B_R)} \lesssim R^{\epsilon^3}\delta^{\frac{-2k+1}{4k+2}+\frac{2k-2}{4k+2}} = R^{\epsilon^3}\delta^{\frac{-1}{4k+2}}.$$

Exactly same steps together with the fact $b_{\tau_i}^{2k+1} \lesssim \delta^{n-1}$ gives

$$\|\Big[\sum_{\tau\in E_\delta}\big(\phi_\tau(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1})^{1/2k+1}\big)^2\Big]^{1/2}\|_{L^{\frac{8k+2}{4k-1}}(B_R)} \lesssim R^{\epsilon^3}.$$

We now proceed to finer estimates. To this end we will perform pigeonholing arguments. We note that use of pigeonholing is very standard in non-endpoint restriction estimates; see for example [28],[6]. We first set

$$g_\tau = \big(\prod_{i=1}^{2k+1} b_{\tau_i}^{2k+1}\big)^{1/2k+1}.$$

Let $0<\lambda<1$ be a dyadic number, and $g_{\tau,\lambda} = g_\tau \chi_{\{g_\tau \approx \lambda\delta^{n-1}\}}$. Then we have

$$\int_{B_R} g_{\tau,\lambda}^{\frac{8k+2}{4k-1}} < (\lambda\delta^{n-1})^{\frac{8k+2}{4k-1}-\frac{2k+1}{k}} \int_{B_R} g_{\tau,\lambda}^{\frac{2k+1}{k}} \lesssim R^{\epsilon^5}\lambda^{\frac{1}{k(4k-1)}}\delta^{\frac{3k-1}{k(4k-1)}+3k-\frac{1}{k}}.$$

We use this and the techniques introduced for coarser estimates above to obtain

$$\Big[\int_{B_R}\big(\sum_{\tau\in E_\delta}(\phi_\tau g_{\tau,\lambda})^2\big)^{\frac{4k+1}{4k-1}}\Big]^{\frac{4k-1}{8k+2}} \lesssim R^{\epsilon^3}\lambda^{\frac{1}{k(8k+2)}}.$$

We do one more pigeonholing. Let $1\le\mu<\infty$ be dyadic, and decompose

$$\phi_\tau = \sum_\mu \phi_{\tau,\mu}$$

with

$$\phi_{\tau,\mu} = \phi_\tau\chi_{\{\phi_\tau\sim\mu\}}, \qquad \phi_{\tau,1} = \phi_\tau\chi_{\{\phi_\tau\le 1\}}.$$

Then we have on τ' boxes B

$$\fint_B \phi_{\tau,\mu}^{\frac{8k+2}{4k-1}} \le \mu^{-\frac{2}{(2k-1)(4k-1)}}\fint_B \phi_{\tau,\mu}^{\frac{4k}{2k-1}} \lesssim R^{\epsilon^4}\mu^{-\frac{2}{(2k-1)(4k-1)}}.$$

Using this and previous techniques

$$\Big[\int_{B_R}\big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\big)^{\frac{4k+1}{4k-1}}\Big]^{\frac{4k-1}{8k+2}}\lesssim R^{\epsilon^3}\lambda^{\frac{1}{k(8k+2)}}\mu^{-\frac{1}{(2k-1)(4k+1)}}. \tag{3.36}$$

We will estimate the left hand side above using a different method. Clearly

$$\big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\big)^{1/2}\leq\mu\big(\sum_{\tau}g^2_{\tau,\lambda}\big)^{1/2}.$$

Here τ ranges over a full partition of the surface S into caps of size δ. We will write the right hand side as convolutions of measures with tubes, to which using convexity we will apply the Kakeya maximal function estimates. We know that separation between directions of caps τ and τ_i is small, and that $b^{2k+1}_{\tau_i}$ are constant on boxes of size larger than the size of τ'. Thus

$$(b^{2k+1}_{\tau_i})^{1/2k+1}\lesssim(b^{2k+1}_{\tau_{2k+1}})^{1/2k+1}*(\delta^{3k+1}\chi_{\tau'}) \tag{3.37}$$

with $\chi_{\tau'}$ denoting the characteristic function of the tube τ'. Hence for $\xi\in B_R$ we have

$$\begin{aligned}g_\tau(\xi)&\lesssim\int\big(\prod_{i=1}^{2k+1}(b^{2k+1}_{\tau_i})^{1/2k+1}*\delta^{3k+1}\chi_{\tau'}\big)(\zeta)(\delta^{3k+1}\chi_{\tau'})(\xi-\zeta)d\zeta\\&=\int_{B_{3R}}\omega(\zeta)(\delta^{3k+1}\chi_{\tau'})(\xi-\zeta)d\zeta\end{aligned}$$

From (3.37) and the definition of $g_{\tau,\lambda}$ we can write

$$g^2_{\tau,\lambda}\lesssim\omega^2\chi_{\{\omega\gtrsim n^{-15}\lambda\delta^{n-1}\}}.$$

The extra n^{-15} on the right hand side above allows us to write

$$\begin{aligned} g^2_{\tau,\lambda}(\xi) &\lesssim \delta^{3k+1} \int_{B_{3R}} (\omega^2 \chi_{\{\omega \gtrsim n^{-15}\lambda\delta^{n-1}\}})(\zeta)\chi_{\tau'}(\xi - \zeta)d\zeta \\ &= \delta^{3k+1} \int (\omega^2 \chi_{\{\omega \gtrsim n^{-15}\lambda\delta^{n-1}\}}\chi_{B_{3R}})(\zeta)\chi_{\tau'}(\xi - \zeta)d\zeta. \end{aligned}$$

We want to replace the measure

$$(\omega^2 \chi_{\{\omega \gtrsim n^{-15}\lambda\delta^{n-1}\}}\chi_{B_{3R}})(\zeta)d\zeta$$

with a constant multiple of a probability measure with support contained in B_{5R}, from which using convexity we will pass to the Kakeya maximal function. This can be done by estimating the total mass of this measure. By Chebyshev's inequality

$$\int \omega^2 \chi_{\{\omega \gtrsim n^{-15}\lambda\delta^{n-1}\}}\chi_{B_{3R}}(\xi)d\xi \lesssim (\frac{1}{\lambda\delta^{n-1}})^{1/k} \int_{B_{3R}} \omega^{\frac{4k+2}{2k}}(\xi)d\xi$$

which can be expanded as

$$\begin{aligned} &\lesssim (\frac{1}{\lambda\delta^{n-1}})^{1/k} \int \Big(\int_{B_{3R}} (\prod_{i=1}^{2k+1} b^{2k+1}_{\tau_i}(\xi - \zeta_i))^{1/k} d\xi \Big) \Big(\prod_{i=1}^{2k+1} (\delta^{3k+1}\chi_{\tau'})(\zeta_i) \Big) d\zeta_1 \dots d\zeta_{2k+1} \\ &\lesssim R^{\epsilon^4} \lambda^{-\frac{1}{k}} \delta^{3k-3}. \end{aligned}$$

Having estimated the total mass of the measure, we can replace it by this quantity multiplied by a probability measure supported in B_{5R}. Thus

$$g^2_{\tau,\lambda} \lesssim R^{\epsilon^4} \delta^{6k-2} \lambda^{-1/k} \int \chi_{\tau'}(\xi - \zeta) d\mu_\tau(\zeta)$$

with $d\mu_\tau$ being the probability measure. From this we have

$$\|\big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\big)^{1/2}\|_{L^{\frac{8k+2}{4k-1}}(B_R)} \lesssim R^{\epsilon^3}\mu\delta^{3k-1}\lambda^{-1/2k}\|\sum_\tau\big(\int\chi_{\tau'}(\xi-\zeta_\tau)d\mu(\zeta_\tau)\big)^2\|^{1/2}_{\frac{4k+1}{4k-1}}.$$

Using the Hölder inequality we have

$$\lesssim R^{\epsilon^3}\mu\delta^{3k-1}\lambda^{-1/2k}\|\sum_\tau\int\chi_{\tau'}(\xi-\zeta_\tau)d\mu(\zeta_\tau)\|^{1/2}_{\frac{4k+1}{4k-1}}.$$

Let's enumerate the range of τ from 1 to l. We observe that

$$\sum_\tau\int\chi_{\tau'}(\xi-\zeta_\tau)d\mu(\zeta_\tau) = \int\Big(\sum_\tau\chi_{\tau'}(\xi-\zeta_\tau)\Big)d\mu(\zeta_1)\dots d\mu(\zeta_l).$$

Thus from the Hölder inequality we have

$$\lesssim R^{\epsilon^3}\mu\delta^{3k-1}\lambda^{-1/2k}\Big(\int\|\sum_\tau\chi_{\tau'}(\xi-\zeta_\tau)\|_{\frac{4k+1}{4k-1}}d\mu(\zeta_1)\dots d\mu(\zeta_l)\Big)^{1/2}.$$

If we can obtain an estimate for

$$\|\sum_\tau\chi_{\tau'}(\xi-\zeta_\tau)\|_{\frac{4k+1}{4k-1}}$$

independent of the choice of ζ_τ, we can use the fact that the measures concerned are probability measures. We will obtain such an estimate via the Kakeya maximal function estimates. Here there are two different estimates, one is due to Wolff and works better for dimensions $n \leq 8$, the other by Katz-Tao and is better when $n \geq 9$. We will thus apply both estimates. Wolff's estimate for δ-separated δ-tubes T, that is, tubes of radius δ and length 1, gives

$$\|\sum_T\chi_T\|_{\frac{n+2}{n}} \lesssim \delta^{-\frac{3k-2}{3k+2}-},$$

with the minus after the exponent denoting an arbitrarily small deduction from it. Of course the constant hidden in $\gtrsim$ notation depends on this deduction. We clearly have the trivial estimate

$$\|\sum_T \chi_T\|_{L^1} \lesssim 1.$$

Interpolating these two estimates yields

$$\|\sum_T \chi_T\|_{\frac{4k+1}{4k-1}} \lesssim \delta^{-\frac{3k-2}{4k+1}-}.$$

On the other hand Katz-Tao estimate gives

$$\|\sum_T \chi_T\|_{\frac{4n+3}{4n-4}} \lesssim \delta^{-\frac{3k-1}{4k+1}-}.$$

Interpolating this with L^1 estimate gives

$$\|\sum_T \chi_T\|_{\frac{4k+1}{4k-1}} \lesssim \delta^{-\frac{18k-6}{28k+7}-}.$$

We rescale these estimates according to the size of our tubes, to obtain

$$\|\sum_T \chi_T\|_{\frac{4k+1}{4k-1}} \lesssim \delta^{\frac{-24k^2+6k}{4k+1}+\frac{-3k+2}{4k+1}-}$$

from the Wolff estimate and,

$$\|\sum_T \chi_T\|_{\frac{4k+1}{4k-1}} \lesssim \delta^{\frac{-24k^2+6k}{4k+1}+\frac{6-18k}{28k+7}-}$$

from the Katz-Tao estimate. We now use these estimates, and the fact that the measures above are probability measures. From the Wolff estimate we get

$$\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{8k+2}{4k-1}}(B_R)} \lesssim R^{\epsilon^3}\mu\delta^{\frac{k}{8k+2}}\lambda^{-1/2k} \tag{3.38}$$

and from the Katz-Tao estimate we get

$$\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{8k+2}{4k-1}}(B_R)} \lesssim R^{\epsilon^3}\mu\delta^{\frac{5k-4}{28k+7}}\lambda^{-1/2k}. \tag{3.39}$$

For $k=2$, (3.38) is better, and in this case we can write

$$\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} \lesssim R^{\epsilon^3}\mu\delta^{\frac{1}{9}}\lambda^{-1/4} \lesssim R^{\epsilon^3}\mu\delta^{\frac{1}{9}}\lambda^{-3/4}, \tag{3.40}$$

and we combine this with previous estimates to get

$$\begin{aligned}\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} &\lesssim R^{\epsilon^3}\min(\mu\delta^{\frac{1}{9}}\lambda^{-3/4}, \lambda^{1/36}\mu^{-1/27}) \\ &\lesssim R^{\epsilon^3}\delta^{\frac{1}{9\cdot 28}}.\end{aligned}$$

Let $\lambda_0 \approx R^{-100n}$, $\mu \approx R^{100n}$, be two dyadic numbers. Then we may write

$$\begin{aligned}\|\Big(\sum_{\tau\in E_\delta}(\phi_\tau g_\tau)^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} \le &\sum_{\lambda\le\lambda_0}\sum_{\mu}\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} \\ &+\sum_{\lambda>\lambda_0}\sum_{\mu>\mu_0}\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} \\ &+\sum_{\lambda>\lambda_0}\sum_{\mu\le\mu_0}\|\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)}\end{aligned}$$

The first term in the sum above can be shown to satisfy $\lesssim R^{\epsilon^3}\lambda_0$ using (3.36), while the second term is bounded by $\lesssim R^{\epsilon^3}\mu_0^{-\frac{1}{27}}$. Clearly both of these terms satisfy $\lesssim R^{\epsilon^2}\delta^{\frac{1}{9\cdot 28}}$. The last term only has about $C\log_2 R$ terms, thus it, too, is dominated by the same expression. Hence

$$\|\Big(\sum_{\tau\in E_\delta}(\phi_\tau g_\tau)^2\Big)^{1/2}\|_{L^{\frac{18}{7}}(B_R)} \lesssim R^{\epsilon^2}\delta^{\frac{1}{9\cdot 28}}.$$

Since we also have

$$\|\Big(\sum_{\tau\in E_\delta}(\phi_\tau g_\tau)^2\Big)^{1/2}\|_{L^{\frac{5}{2}}(B_R)} \lesssim R^{\epsilon^2}\delta^{\frac{-1}{10}}.$$

Interpolation yields an improvement of

$$\frac{5}{1764}.$$

We turn to $k \geq 3$ case. In this case (3.39) is better, and

$$\begin{aligned}\|\max_{E_\delta}\Big(\sum_{\tau\in E_\delta}(\phi_{\tau,\mu}g_{\tau,\lambda})^2\Big)^{1/2}\|_{L^{\frac{8k+2}{4k-1}}(B_R)} &\lesssim R^{\epsilon^3}\min(\mu\delta^{\frac{5k-4}{28k+7}}\lambda^{-1/2k}, \lambda^{\frac{1}{k(8k+2)}}\mu^{-\frac{1}{(2k-1)(4k+1)}})\\ &\lesssim R^{\epsilon^3}\min(\mu\delta^{\frac{5k-4}{28k+7}}\lambda^{-\frac{2k-1}{2k}}, \lambda^{\frac{1}{k(8k+2)}}\mu^{-\frac{1}{(2k-1)(4k+1)}})\\ &\lesssim R^{\epsilon^3}\delta^{\frac{5k-4}{(28k+7)(8k^2-2k)}}.\end{aligned}$$

Then from the same arguments, the improvement improvement is

$$\frac{(5k-4)(2k+1)}{(4k^2-k)(28k+7)(4k^2-k)}.$$

3.3.2 Estimates on the Linear Term

In this term the scale is about $R^{-1/2}$, and this permits use of the more direct method of conservation of mass presented in Lemma 1. We will combine this method with the methods already introduced above.

$$\|\Big[\sum_{\tau\in E}(\phi_\rho b_\rho)^2\Big]^{1/2}\|_{L^{\frac{4k+2}{2k}}(B_R)} \lesssim R^{\frac{2k-1}{2(4k++2)}}\Big[\sum_{\rho}\int_{B_R}(\phi_\rho b_\rho)^{\frac{4k+2}{2k}}\Big]^{\frac{2k}{4k+2}}.$$

In the last term ρ ranges over a partition into caps of size γ of all of S. We use the technique in (3.35) to obtain

$$\lesssim R^{\frac{2k-1}{2(4k++2)}+\epsilon^4}\Big[\sum_\rho \int_{B_R}(b_\rho)^{\frac{4k+2}{2k}}\Big]^{\frac{2k}{4k+2}}.$$

Due to size of our caps we have $b_\rho \lesssim R^{-3k-1/2+\epsilon^5}$. Using this we have

$$\lesssim R^{\frac{2k-1}{2(4k++2)}+\frac{1-3k}{4k+2}+\epsilon^3}\Big[\sum_\rho \int_{B_R}(b_\rho)^2\Big]^{\frac{2k}{4k+2}}.$$

We now use the conservation of mass property. To this end we expand

$$\begin{aligned}\int_{B_R}(b_\rho(\xi))^2 d\xi &\lesssim \int_{B_R}\Big(\int |T_\rho f(\xi-\zeta)|\beta_{\gamma,\rho}(\zeta)d\zeta\Big)^2 d\xi\\ &\lesssim \int \Big(\int_{B_R} |T_\rho f(\xi-\zeta)|^2 d\xi\Big)\beta_{\gamma,\rho}(\zeta)d\zeta.\end{aligned}$$

To the inner integral we apply conservation of mass. We have

$$\int_{B_R}|T_\rho f(\xi-\zeta)|^2 d\xi \lesssim R^{\frac{3-3k}{2}+\epsilon^5}.$$

Thus what we have

$$\Big[\sum_\rho \int_{B_R}(b_\rho(\xi))^2 d\xi\Big]^{\frac{2k}{4k+2}} \lesssim R^{\frac{2k}{4k+2}+\epsilon^5}.$$

Hence we finally get

$$\lesssim R^{\frac{1}{2(4k+2)}+\epsilon^2}.$$

This time we will obtain estimates in the space $L^{\frac{8k+2}{4k-1}}$ first without using the Kakeya maximal function estimate and then using it. We decompose ϕ_ρ into $\phi_{\rho,\mu}$ exactly as before.

Then

$$
\begin{aligned}
\| \Big[\sum_{\rho \in E} (\phi_{\rho,\mu} b_\rho)^2 \Big]^{1/2} \|_{L^{\frac{8k+2}{4k-1}}(B_R)} &\le R^{\frac{2k-1}{8k+2}} \Big(\sum_\rho \int_{B_R} (\phi_{\rho,\mu} b_\rho)^{\frac{8k+2}{4k-1}} \Big)^{\frac{4k-1}{8k+2}} \\
&\lesssim R^{\frac{2k-1}{8k+2}+\epsilon^4} \mu^{-\frac{1}{(2k-1)(4k+1)}} \Big(\sum_\rho \int_{B_R} (b_\rho)^{\frac{8k+2}{4k-1}} \Big)^{\frac{4k-1}{8k+2}} \\
&\lesssim R^{\frac{2k-1}{8k+2}+\frac{1-3k}{4k+1}+\epsilon^3} \mu^{-\frac{1}{(2k-1)(4k+1)}} \Big(\sum_\rho \int_{B_R} (b_\rho)^2 \Big)^{\frac{4k-1}{8k+2}} \\
&\lesssim R^{\epsilon^2} \mu^{-\frac{1}{(2k-1)(4k+1)}}.
\end{aligned}
$$

Now we will employ the Kakeya maximal function bounds. We will follow exactly the same procedure as in the multilinear case. Thus we first write

$$
b_\rho \lesssim b_\rho * \gamma^{3k+1} \chi_{\rho'}.
$$

Expanding the right hand side and using the Hölder inequality, we have for $\xi \in B_R$

$$
\begin{aligned}
(b_\rho)^2(\xi) \lesssim \gamma^{3k+1} \int (b_\rho)^2(\zeta) \chi_{\rho'}(\xi - \zeta) d\zeta &= \gamma^{3k+1} \int_{B_{2R}} (b_\rho)^2(\zeta) \chi_{\rho'}(\xi - \zeta) d\zeta \\
&= \gamma^{3k+1} \int (b_\rho)^2(\zeta) \chi_{B_{3R}}(\zeta) \chi_{\rho'}(\xi - \zeta) d\zeta.
\end{aligned}
$$

We consider $(b_\rho)^2(\zeta)\chi_{B_{3R}}(\zeta)d\zeta$ as a measure, and we want to replace it with a probability measure supported in B_{5R}. Thus we wish to estimate its total mass. From conservation of mass we have

$$
\int_{B_{3R}} (b_\rho)^2(\xi) d\xi \lesssim R^{\frac{3-3k}{2}+\epsilon^5}.
$$

Hence

$$
(b_\rho)^2(\xi) \lesssim R^{-3k+1+\epsilon^4} \int \chi_{\rho'}(\xi - \zeta) d\mu_\rho(\zeta).
$$

Thus

$$
\begin{aligned}
\Big\| \Big[\sum_{\rho \in E} (\phi_{\rho,\mu} b_\rho)^2 \Big]^{1/2} \Big\|_{L^{\frac{8k+2}{4k-1}}(B_R)} &\leq \mu \Big\| \Big[\sum_{\rho} (b_\rho)^2 \Big]^{1/2} \Big\|_{L^{\frac{8k+2}{4k-1}}(B_R)} \\
&\lesssim R^{\frac{1-3k}{2}+\epsilon^4} \mu \sup_{\zeta_\rho} \Big\| \sum_{\rho} \chi_{\rho'}(\cdot - \zeta_\rho) \Big\|^{1/2}_{\frac{4k+1}{4k-1}} \\
&\lesssim \min\Big(R^{\frac{1-3k}{2}+\frac{1}{4}\frac{24k^2-3k-2}{4k+1}+\epsilon^3}, R^{\frac{1-3k}{2}+\frac{1}{4}\left(\frac{24k^2-6k}{4k+1}+\frac{6-18k}{28k+7}\right)+\epsilon^3}\Big)\mu
\end{aligned}
$$

with different powers of R coming from different Kakeya estimates introduced before. Now for $k=2$ the first is better, and comparison with the other method, and interpolation shows that improvement is again

$$\frac{5}{1764}.$$

For $k \geq 3$ the second is better, and we still have the same improvement:

$$\frac{(5k-4)(2k+1)}{(4k^2-k)(28k+7)(4k^2-k)}.$$

As from both linear and multilinear cases we obtain the same improvements we have Theorem 2.

Chapter 4

References

[1] B. Barcelo, On the restriction of the Fourier transform to a conical surface, Trans. Amer. Math. Soc. **292** (1985), no. 1, 321-333.

[2] J.Bennett, A. Carbery, T. Tao, On the multilinear restriction and Kakeya conjectures, Acta Math., **196** (2006), no.2, 261-302.

[3] J. Bourgain, Besicovitch type maximal operators and applications to Fourier Analysis, Geom. and Funct. Anal., **22** (1991), no.2, 147-187.

[4] J. Bourgain, Estimates for cone multipliers, Operator Theory: Advances and Applications, **77** (1995), 41-60.

[5] J. Bourgain, Harmonic analysis and combinatorics: How much may they contribute to each other?, Mathematics: Frontiers and perspectives, IMU/ Amer. Math. Society, (2000) 13-32.

[6] J. Bourgain, L. Guth, Bounds on oscillatory integral operators based on multilinear estimates, to appear Geom. Funct. Anal.

[7] L. Carleson and P.Sjölin, Oscillatory integrals and a multiplier problem for the disc, Studia Math. **44** (1972), 287-299.

[8] B.Erdoğan, On Falconer's distance set conjecture, Rev. Mat. Iberoamericana **22** (2006), no.2, 649-662.

[9] B.Erdoğan, A bilinear Fourier extension theorem and applications to the distance set problem, Internat. Math. Res. Notices **23** (2005), no.23, 1411-1425,

[10] C. Fefferman, Inequalities for strongly singular convolution operators, Acta Math., **124** (1970), 9-36.

[11] D. Foschi and S. Klainerman, Bilinear space-time estimates for homogeneous wave equations, Ann. Sci. École Norm. Sup., (4) **33** (2000), no. 2, 211-274.

[12] S. Klainerman, M. Machedon, Remark on Strichartz-type inequalities. With appendices by Jean Bourgain and Daniel Tataru, Internat. Math. Res. Notices, (1996), no.5, 201-220.

[13] I. Laba, From harmonic analysis to arithmetic combinatorics, Bull. of Amer. Math. Soc., **45** (2007), no.1, 77-115.

[14] J. Lee, Endpoint bilinear restriction estimate for the paraboloid, preprint

[15] J. Lee, Endpoint estimates for bilinear oscillatory integral operators related to restriction to the cone, Transactions of Amer. Math. Soc., **363** (2011), no.2, 763-800.

[16] S. Lee, Bilinear restriction estimates for surfaces with curvatures of different signs, Trans. Amer. Math. Soc., **358** (2006), no. 8, 3511-3533.

[17] S. Lee, Improved bounds for Bochner-Riesz and maximal Bochner-Riesz operators, Duke Math. J., **122** (2004), no. 1, 205-232.

[18] S. Lee, K. M. Rogers and A. Vargas, Sharp null form estimates for the equation in $\mathbf{R}^{3+1}$, Internat. Math. Res. Notices, (2008), 18pp.

[19] S. Lee and A. Vargas, Sharp null form estimates for the wave equation, American Journal of Mathematics, **130** (2008), no.5, 1279-1326.

[20] G. Mockenhaupt, A note on the cone multiplier, Proc. AMS, **117** (1993), no. 1, 145-152.

[21] S. Shao, Sharp linear and bilinear restriction estimates for paraboloids in the cylindrically symmetrical case, Rev. Mat. Iberoamericana, **25** (2009), no.2, 1127-1168.

[22] P. Sjölin, Regularity of solutions to Schrödinger equations, Duke Math. J., **55** (1987), 699-715.

[23] E.M. Stein, Some problems in harmonic analysis, Harmonic analysis in Eucliden spaces(Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, 3-20.

[24] R.S. Strichartz, Restriction of Fourier transform to quadratic surfaces and Decay of solutions of wave equations, Duke Math. j., **44** (1977), no.3, 705-774.

[25] T. Tao, A. Vargas, L. Vega, A bilinear approach to the restriction and Kakeya conjectures, J. Amer. Math. Soc., **11** (1998), no.4, 967-1000.

[26] T. Tao, A. Vargas, A bilinear approach to cone multipliers I. Restriction Estimates, Geom. Funct. Anal., **10** (2000), no.1, 185-215.

[27] T. Tao, A. Vargas, A bilinear approach to cone multipliers II. Applications, Geom. Funct. Anal., **10** (2000), no.1, 216-258.

[28] T. Tao, Endpoint bilinear restriction theorem for the cone, and some sharp null form estimates, Math. Z., **238** (2001), no.2, 215-268.

[29] T. Tao, A sharp bilinear restriction estimate for paraboloids, Geom. Funct. Anal., **13** (2003), no.6, 1359-1384.

[30] T. Tao, Recent progress on the restriction conjecture, IAS/ Park City Mathematics Series, (2003).

[31] D. Tataru, Null form estimates for second order hyperbolic operators with rough coefficients, Harmonic analysis at Mount Holyoke, Contemp. Math., **320** (2001), 383-409.

[32] P.Tomas, A restriction theorem for the Fourier transform. Bull. Amer. Math. Soc., **81** (1975), no. 477-478.

[33] A. Vargas, Restriction theorems for a surface with negative curvature, Math. Z., **249** (2005), no.1, 97-111.

[34] T. Wolff, A sharp bilinear cone restriction estimate, Ann. of Math., **153** (2001), no.3, 661-698.

[35] T. Wolff, Recent work connected with Kakeya problem, Prospects in Mathematics, 129-162 (1999).

[36] T. Wolff, Local smoothing estimates in L^p for large p, Geom. Funct. Anal., **10** (2000), no.5, 1237-1288.

[37] T. Wolff, An Improved bound for Kakeya type maximal functions, Revista Math. Iberoamericana, **11** (1995), no.3, 651-674.