LINEAR AND BILINEAR RESTRICTION TO CERTAIN
ROTATIONALLY SYMMETRIC HYPERSURFACES

BETSY STOVALL

ABSTRACT. Conditional on Fourier restriction estimates for elliptic hypersurfaces, we prove optimal restriction estimates for polynomial hypersurfaces of revolution for which the defining polynomial has non-negative coefficients. In particular, we obtain uniform—depending only on the dimension and polynomial degree—estimates for restriction with affine surface measure, slightly beyond the bilinear range. The main step in the proof of our linear result is an (unconditional) bilinear adjoint restriction estimate for pieces at different scales.

1. Introduction

Recently, there has been considerable interest (e.g. [1, 7, 8, 9, 10, 15, 19, 21]) in extending the restriction problem to degenerate hypersurfaces, that is, hypersurfaces for which one or more of the principal curvatures is allowed to vanish to some finite (or infinite) order. It has been known for a number of years that if the hypersurface is equipped with Euclidean surface measure, the exponent pairs for which restriction phenomena are possible must depend on the ‘type,’ or order of vanishing of the curvatures. Affine surface measure, however, is conjectured to mitigate the effects of such degeneracies and allow for restriction theorems that are uniform over large classes of hypersurfaces. We verify this conjecture for a class of rotationally symmetric hypersurfaces by proving that the elliptic restriction conjecture implies the restriction conjecture with affine surface measure.

Consider the hypersurface
$$\Gamma = \{(G(\xi), \xi) : \xi \in U \subseteq \mathbb{R}^d\}.$$ We say that $$\Gamma$$ (or $$G$$) is elliptic with parameters $$A, N,$$ and $$1 > \epsilon_0 > 0$$ if $$U$$ is a subset of the unit ball $$B,$$ $$\|\nabla G\|_{C^N(B)} \leq A,$$ and the eigenvalues of $$D^2 G(x)$$ lie in $$(\epsilon_0, \epsilon_0^{-1})$$ for all $$x \in U.$$ The restriction conjecture for elliptic hypersurfaces is the statement that for all pairs $$(p, q)$$ satisfying the (restriction) admissibility condition
$$\frac{2(d+1)}{d+2} \leq q \leq \infty, \quad q = \frac{dp'}{d+2},$$ (1.1)
there exists $$N = N_p$$ such that for all parameters $$A, \epsilon_0,$$ and all elliptic phases $$\Phi$$ with parameters $$A, N, \epsilon_0,$$
$$\left( \int_B |\widehat{f}(\Phi(\xi), \xi)|^q \, d\xi \right)^{\frac{1}{q}} \lesssim \|f\|_{L^p_{\mu}(\mathbb{R}^{1+d})} \quad f \in \mathcal{S}(\mathbb{R}^{1+d}),$$ (1.2)
where $$\mathcal{S}$$ denotes the Schwartz class and the implicit constant is allowed to depend on $$p, A, \epsilon_0.$$ We let $$\mathcal{R}(p \to q)$$ denote the statement that the restriction conjecture
2010 MSC 42
for elliptic hypersurfaces is valid for the exponents $p, q$. We note that our definition of elliptic differs slightly from that in [26], but by a well-known argument (a partition of unity coupled with affine transformations), the corresponding restriction conjectures are easily seen to be equivalent.

In the notation above, affine surface measure on $\Gamma$ is the pushforward by $\xi \mapsto (G(\xi), \xi)$ of

$$
\Lambda_G(\xi) \, d\xi := |\det D^2 G(\xi)|^{\frac{1}{p'}} \, d\xi;
$$

more geometrically, for $\omega \in \Gamma$, it equals $|\kappa(\omega)|^{\frac{1}{p'}} \, d\sigma(\omega)$, where $\kappa$ is the Gaussian curvature and $d\sigma$ is Lebesgue measure on $\Gamma$ [18]. Since this measure gives little weight to the 'bad' flat regions of $\Gamma$, it is natural to ask whether it is possible to prove restriction estimates of the form

$$
\| \hat{f}(G(\xi), \xi) \|_{L^p_{t,x}(\mathbb{R}^{d+1})} \lesssim \| f \|_{L^q_{t,x}(\mathbb{R}^{1+d})},
$$

for $(p, q)$ satisfying the admissibility condition (1.1) and with the implicit constant uniform over $G$ in some reasonably large class. Oscillation is a well-known enemy of restriction estimates—consider, for instance Sjölin’s counter-example $(t, \sin(t^k) e^{-1/t})$ [21]—so it is natural to consider the affine restriction problem for $G$ a polynomial of bounded degree.

Here we specialize somewhat more. Let $P$ be an even polynomial on $\mathbb{R}$ with non-negative coefficients, and let

$$
S_P = \{(P(|\xi|), \xi) : \xi \in \mathbb{R}^d\}.
$$

The following is our main result.

**Theorem 1.1.** Assume that the restriction conjecture $R(p_0 \to q_0)$ holds for some admissible pair of exponents. Then for every restriction admissible pair $(p, q)$ with $p < p_0$, if $P : \mathbb{R} \to \mathbb{R}$ is an even polynomial of degree $N$ with non-negative coefficients, the restriction estimate

$$
\left( \int |\hat{f}(P(|\xi|), \xi)|^q \Lambda_P(\xi) \, d\xi \right)^{\frac{1}{q}} \leq C \| f \|_{L^p_{t,x}(\mathbb{R}^{1+d})}
$$

holds for all $f \in L^p_{t,x}(\mathbb{R}^{1+d})$. The constant $C$ depends only on $p, p_0, d$, the degree of $P$, and the constants in (1.2).

In particular, the restriction estimate (1.3) holds in the bilinear range $p < \frac{2(d+3)}{d+5}$. As pointed out in [17, Section 5.2], the recent Bourgain–Guth [6] and Guth [14] theorems and the bilinear-to-linear method of [26] establish $R(p \to q)$ (and hence Theorem 1.1) in a slightly better (but awkward-to-state) range. (This argument gives us a better range in the monomial case because the existence of an almost transitive group action allows the use of the Maurey–Nikishin–Pisier factorization theorem [20] as in [5].) For $d \geq 3$, analogues of Theorem 1.1 were previously known only in the Stein–Tomas range [9, 10] (these results cover somewhat more general hypersurfaces).

We will primarily focus on restriction with affine surface measure along the scaling line $q = \frac{d+2}{2d+2}$ because this gives essentially the strongest possible estimates for such hypersurfaces. However, in the last section, we will show how to deduce local (i.e. for compact pieces of the hypersurface) estimates from results off the
scaling line \( q = \frac{dp'}{d+2} \) (such as the Bourgain–Guth theorem [6]), as well as sharp unweighted estimates.

It should be possible to relax the hypotheses on \( P \) substantially. Evenness guarantees smoothness of \( S_P \) and the vanishing of the linear term. Neither smoothness at zero nor rotational symmetry are essential for our proof, and variants will be discussed in the last section. The positivity of the coefficients and vanishing of the linear term, however, reflect geometric considerations that do play an important role. Most obviously, the hypothesis that the coefficients are nonnegative rules out negatively curved hypersurfaces, for which no sharp restriction estimates are known beyond the Stein–Tomas range [27, 16]. More subtly, since the linear term vanishes and the coefficients are positive, we can rescale dyadic annuli in \( S_P \) to uniformly elliptic hypersurfaces. That being said, in the last section, we will give a global, but non-uniform result for polynomials \( P \) with \( P''(t) > 0 \) for all \( t > 0. \)

**Sketch of proof.** By duality, \( R(p \to q) \) is equivalent to the adjoint restriction conjecture, which we denote by \( R^*(q' \to p') \). The adjoint restriction operator is also known as the extension operator, and we will say that an exponent pair \((p, q)\) is (extension) admissible if \((q', p')\) is restriction admissible, i.e. if

\[
\frac{2(d+1)}{d} < q \leq \infty, \quad q = \frac{(d+2)p'}{d}.
\]

It will generally be clear from the context whether an ‘admissible’ pair is restriction or extension admissible.

Our goal is to prove that for any admissible \((p_0, q_0)\), \( R^*(p_0 \to q_0) \) implies that the extension operator

\[
E_P f(t, x) = \int e^{itP(|\xi|)+ix\xi} f(\xi) d\xi
\]

satisfies

\[
\|\Lambda_P(\nabla)^{1/p'} E_P f\|_{L_{t,x}^q} \lesssim \|f\|_{L_{\xi}^p}, \quad f \in S(\mathbb{R}^d),
\]

for all admissible \((p, q)\) with \( p < p_0 \), with implicit constants depending on \( d, p, \) and the degree of \( P \).

We will proceed along the following lines. Given a polynomial \( P(t) = a_1 t^2 + \cdots + a_N t^{2N} \) with the \( a_i \) non-negative, we may decompose \( \mathbb{R} \) as a union of intervals, \( \mathbb{R} = \bigcup_{j=1}^{J} I_j \), such that on \( I_j \), \( P \) behaves like the monomial \( a_j t^{2j} \), plus a controllable error. By the triangle inequality, it suffices to prove a uniform restriction estimate for each annular hypersurface \( \{(P(|\xi|, \xi)) : |\xi| \in I_j\} \). By affine invariance of (1.3), we may assume that \( a_j = 1 \). The essential difficulty is then encapsulated by the problem of proving restriction estimates for degenerate hypersurfaces of the form \( \{(|\xi|^2, \xi)\} \), for \( j > 1 \).

By rescaling, the restriction problem on \( \{(|\xi|^2, \xi) : |\xi| \sim 2^k\} \) is equivalent to restriction to \( \{(|\xi|^2, \xi) : |\xi| \sim 1\} \). The latter is (after a partition of unity) elliptic, so we can apply our hypothesis \( R(p_0 \to q_0) \), which implies \( R(p \to q) \) by interpolation. This leaves us to control the interaction between the dyadic annuli and then sum up the dyadic pieces. The former we do by means of a bilinear restriction estimate for transverse hypersurfaces whose curvatures are at different scales; after that the summation is almost elementary.
Prior results. As mentioned earlier, the natural conjectural form of Theorem 1.1 is for arbitrary polynomial hypersurfaces. This is known if $d = 2$ [21]. In fact, a uniform restriction result is known for polynomial curves with affine arclength measure in all dimensions ([23] and the references therein).

For hypersurfaces of dimension two or more, matters seem significantly more complicated. Carbery–Kenig–Ziesler have proved uniform restriction theorems with affine surface measure in 1 + 2 dimensions for $q \leq 2$ for rotationally symmetric hypersurfaces satisfying rather weak conditions on their derivatives [10] (cf. [19]) and for arbitrary homogeneous polynomials [9]. Ikromov–M"{u}ller [15] have proved the sharp unweighted $L^2$ restriction estimates for hypersurfaces in $\mathbb{R}^3$ expressed in adapted coordinates.

Beyond the Stein–Tomas range, very little is known about restriction to degenerate hypersurfaces. While working on this project, the author learned of independent work of Buschenhenke–M"{u}ller–Vargas, which gives the first sharp restriction estimates for any degenerate hypersurfaces (of dimension at least two) beyond the Stein–Tomas range. (This work has appeared in the Ph.D. thesis of Buschenhenke, [7]; a version will be submitted for publication as [8].) In this, the authors establish a Fourier restriction theorem for convex finite type surfaces in $\mathbb{R}^3$ of the form $\{(\phi_1(\xi_1) + \phi_2(\xi_2), \xi) : \xi \in \mathbb{R}^2\}$. Both the form of the result and the methods are different (though there are some coincidental similarities in the proofs of the bilinear results). In particular, the authors use the measure $d\xi$ (rather than the affine surface measure) and directly prove the corresponding scaling critical estimates, which necessarily depend on the $\phi_j$, in the bilinear range, without the use of the square function.

Notation. For two nonnegative quantities $A$ and $B$, the notation $A \lesssim B$ will be used to mean $A \leq CB$ for some constant $C$ that depends only on the dimension, degree of $P$ (or on the ellipticity parameters for more general results), and exponents $p, q, p_0$, unless otherwise stated. We will write $A \sim B$ to mean $A \lesssim B$ and $B \lesssim A$, and $A = O(B)$ to mean $|A| \lesssim |B|$. We will define the notation $A \lesssim B$ later on (at the beginning of the proof of Lemma 4.2 and at the end of Section 5) since its meaning will change. The spatial Fourier transform, which acts on functions on $\mathbb{R}^d$, will be denoted by $f \mapsto \hat{f}$ and its inverse by $g \mapsto \check{g}$. The spacetime Fourier transform, which acts on functions on $\mathbb{R}^{1+d}$, is denoted by $\mathcal{F}$. To simplify exponents, we will consistently ignore the fact that $2\pi \neq 1$.

Acknowledgements. This work was supported by NSF grant DMS-1266336. The author would like to thank Shuanglin Shao, Keith Rogers, Stefan Buschenhenke, Detlef M"{u}ller, and Ana Vargas for enlightening conversations along the way.

2. Bilinear restriction I: Statement of result

We state our bilinear restriction result in the $C^\infty$, rather than polynomial, setting. Let $c_0 > 0$ and let $N$ be sufficiently small and large, respectively, dimensional constants. Let $1 > c_0 > 0$, let $A < c_0$, and let $g_1, g_2 \in C^\infty(B(0, c_0))$ be elliptic phases (as defined in the previous section), which also satisfy the transversality condition $|\nabla g_1(0)| \lesssim |\nabla g_2(0)| \sim 1$; thus $|\nabla g_1| \lesssim |\nabla g_2|$ throughout $B(0, c_0)$. The reader may find it helpful to keep the model case $g_1(\xi) = |\xi|^2$, $g_2(\xi) = |\xi - e_1|^2$ in mind.
Fix \( J > 2 \) and a pair of integers \( k_1 > k_2 \). Define phase functions
\[
h_j(\xi) := 2^{-Jk_1}g_j(2^{k_1}\xi), \quad \xi \in B(0, c_0 2^{-k_1}),
\]
surfaces
\[
S_j := \{(h_j(\xi), \xi) : \xi \in B(0, c_0 2^{-k_1})\},
\]
and extension operators
\[
E_j f(t, x) := \int_{\{||\xi|| < c_0 2^{-k_1}\}} e^{i(t+L_1^{2k_1}(\xi))} (h_j(\xi), \xi) f(\xi) d\xi, \quad j = 1, 2.
\]

For simplicity, we state our bilinear result when \( k_2 = 0 \); the general case may be obtained by scaling.

**Theorem 2.1.** For \( C = C_{d, c_0, J} \) sufficiently large, and all \( k_1 \geq C, k_2 = 0, \delta > 0, \) and \( 2 \leq q > \frac{d+3}{d+1}, \)
\[
\|E_1 f_1 E_2 f_2\|_{L_1^q z} \leq \delta q 2^{k_1(J-2)(\frac{d}{2} - \frac{1}{2} + \delta)\|f_1\|_{L_1^2} \|f_2\|_{L_1^2}, \quad f_1, f_2 \in L_1^q z.
\]
The implicit constant is allowed to depend on \( \delta, q, \) as well as \( d, A, \varepsilon_0, J, \) but not on the phases \( g_1, g_2. \)

**Remarks:** For \( 2 \leq q \leq \infty, \) it is easy to prove this result without the exponential term; combining this with the theorem, we obtain the full range of estimates of the form \( L^2 \times L^2 \to L^q, \) excepting possibly the endpoint \( q = \frac{d+3}{d+1}. \) We have not explored the optimal power of \( 2^{k_1}. \) In fact, the power given here is certainly not optimal since we do not use the small size of \( S_1. \) On the other hand, our argument also works, with some modifications, when \( h_1 \) is replaced by \( 2^{-(J-2)k_1}g_1. \)

In \( 1+2 \) dimensions, bilinear adjoint restriction results have been proved in much greater generality by Buschenhenke–Müller–Vargas in [8]. It is also the author’s understanding that they have independently obtained the above high-dimensional result using their methods (personal communication).

The following scaling critical bilinear restriction result will be used in the proof of the linear restriction theorem.

**Corollary 2.2.** Assume that \( \mathcal{R}^*(p_0 \to q_0) \) holds for some admissible pair \((p_0, q_0)\) with \( p_0 > 2, \) and assume that \( N \) is large enough to satisfy both the hypotheses of the elliptic restriction theorem \( \mathcal{R}^*(p_0 \to q_0) \) and of Theorem 2.1. Then for all admissible pairs \((p, q)\) with \( 2 < p < p_0, \) and any integers \( k_1, k_2, \) we have the bilinear extension estimate
\[
\|(2^{-k_1} \frac{J-2d}{d} E_1 f_1)(2^{-k_2} \frac{J-2d}{d} E_2 f_2)\|_{L_1^2 z} \lesssim 2^{-\delta_p |k_1 - k_2|} \|f_1\|_{L_1^p} \|f_2\|_{L_1^p}, \quad (2.1)
\]
for some \( \delta_p > 0 \) depending on \( p, J. \) The implicit constant depends on \( p, A, \varepsilon_0, J. \)

**Proof of Corollary 2.2.** By considering the special case \( q = \frac{d+2}{d} \) of the bilinear theorem and rescaling, we obtain the bilinear Stein–Tomas inequality
\[
2^{\frac{(J-2d)(k_1+k_2)}{2(d+2)}} \|E_1 f_1 E_2 f_2\|_{L_1^\frac{4d}{d+2}} \lesssim 2^{-c_d |k_1 - k_2|} \|f_1\|_{L_1^p} \|f_2\|_{L_1^p}, \quad (2.2)
\]
for some \( c_d > 0. \)

Supposing that \( \mathcal{R}^*(p_0 \to q_0) \) holds for some admissible pair, by rescaling we see that
\[
\|2^{-k_j} \frac{J-2d}{d} E_j f\|_{L_1^p} \lesssim \|f\|_{L_1^{p_0}}, \quad j = 1, 2.
\]
Thus by Cauchy–Schwarz,
\[
\|2^{-k_1} \frac{x}{2} \hat{f}_1 2^{-k_2} \frac{x}{2} \hat{f}_2\|_{L_{t,x}^\infty} \lesssim \|f_1\|_{L_{t,x}^{2p}} \|f_2\|_{L_{t,x}^{2q}},
\]
for any pair \(k_1, k_2\). By interpolation with (2.2), we obtain the corollary. \(\square\)

**Remark:** In any dimension, an \(L^p_t \times L^q_t \rightarrow L^r_{t,x}\) estimate is easily proved by a well known argument using Plancherel, a change of variables, transversality (not curvature) of the hypersurfaces, and the support sizes. In dimension 1+2, this yields the improved bilinear Stein–Tomas estimate (2.2) directly, giving the corollary without the need for the bilinear machinery. In higher dimensions, this does not quite work; we would want
\[
|\int_{\{|\xi|<2^{-k_1}\}} f_1(\xi_1) f_2(\xi_2) f_3(h_1(\xi_1) + h_2(\xi_2), \xi_1 + \xi_2) d\xi_1 d\xi_2| \lesssim 2^{(k_1+k_2)c_d(J-2)} \|f_1\|_{L^p_{\xi}} \|f_2\|_{L^p_{\xi}} \|f_3\|_{L^q_{\xi}},
\]
for \(c_d > 0\) is sufficiently small. Unlike the \(d = 2\) case, however, the corresponding estimate for flat but transverse hypersurfaces is false, so curvature must play some role. (The full range of exponents in the flat case is given in [2, 3].)

**Notation.** We use \(\mathcal{R}^*(p \times p \rightarrow q)\) as shorthand for the statement that inequality (2.1) holds for extension operators \(\mathcal{E}_1, \mathcal{E}_2\) as described in this section.

### 3. Proof of the Linear Result

This section will be devoted to a proof of Theorem 1.1, using Corollary 2.2 from the previous section.

For the remainder of the section, we assume that the adjoint restriction conjecture \(\mathcal{R}^*(p_0 \rightarrow q_0)\) holds for some (extension) admissible pair \((p_0, q_0)\). We may assume that \(p_0 > 2\).

Fix an admissible pair \((p, q)\) with \(p < p_0\). By interpolation with the trivial \(L^1 \rightarrow L^\infty\) bound, \(\mathcal{R}^*(p_0 \rightarrow q_0)\) implies that \(\mathcal{R}^*(p \rightarrow q)\) holds for all admissible pairs \((p, q)\) with \(p \leq p_0\).

Write \(P(t) = a_0 + a_1 t^2 + \cdots + a_N t^{2N}\), with the \(a_i\) nonnegative. We may assume that \(a_0 = 0\). By duality, it suffices to prove that
\[
\|\Lambda_P(\nabla) \mathcal{R}^* E_P f\|_{L^q_{t,x}} \lesssim \|f\|_{L^p} \quad (3.1)
\]
for all \(f \in L^p_q\) and admissible \((p, q)\) with \(p < p_0\), where the implicit constant depends on \(p, N\). Here \(\Lambda_P(\nabla)\) denotes the Fourier multiplication operator with symbol \(\Lambda_P(\xi)\).

#### 3.1. Initial decomposition

We begin by decomposing \((0, \infty)\) as a union of intervals on which \(P\) is essentially monomial-like.

Define
\[
J_j := \{t \in (0, \infty) : a_j t^{2j} = \max_{1 \leq i \leq N} a_i t^{2i}\}.
\]
Then the \(J_j\) are consecutive intervals, intersecting only at their boundaries, and \((0, \infty) = \bigcup_{j=1}^{N} J_j\). By the triangle inequality, it suffices to prove (3.1) for \(f\) supported on a single annulus \(\{ |\xi| \in J_j \}\). The low frequency case is easy.
Lemma 3.1. Let $B_1 = \{0\} \cup \{\{\xi\} \in J_1\}$. We have the estimate
\[ \|\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{B_1} f\|_{L^q_{1,x}} \lesssim \|f\|_{L^q_x}, \quad (3.2) \]
with uniform implicit constants.

Proof. By rescaling, we may assume that $B_1$ equals $B$, the unit ball. By applying an affine transformation, we may assume that $a_1 = 1$. Then by the definition of $J_1$, $a_j \leq 1$, $2 \leq j \leq N$, so $\Lambda P \sim 1$ on $B$ and (3.2) just follows from our assumption that $\mathcal{R}^*(p_0 \to q_0)$ (and hence $\mathcal{R}^*(p \to q)$) holds. \qed

3.2. Dyadic decomposition. Fix an integer $j \geq 2$. By applying an affine transformation, we may assume that $a_j = 1$.

Let $I_k := J_j \cap [2^{-k-1}, 2^{-k}]$. Assume that $I_k \neq \emptyset$. We will assume that $I_k = [2^{-k-1}, 2^{-k}]$. (For simplicity we ignore intervals containing the endpoints of the $J_j$; they may be treated similarly, and there are only a bounded number of them anyway.) Let $A_k := \{\xi : |\xi| \in I_k\}$. Consider the phase
\[ g_k(\xi) := 2^{2jk}g(2^{-k}|\xi|) = |\xi|^{2j} + \sum_{i \neq j} a_i 2^{2(j-i)k}|\xi|^i, \quad \xi \in A_0. \]
Since $2^{-k} \in J_j$, we have $2^{2(j-i)k}a_i \leq 1$, $i \neq j$, so $g_k$ is elliptic (with the parameters $A_i, \varepsilon$ depending only on the degree of $P$). Thus by our hypothesis that $\mathcal{R}^*(p \to q)$ holds for admissible $(p, q)$ with $p \leq p_0$, and rescaling, we have the following.

Lemma 3.2. For any $k \in \mathbb{Z}$,
\[ \|\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_k} f\|_{L^q_{1,x}} \lesssim \|f\|_{L^q_x}. \]

3.3. Almost orthogonality. The next lemma establishes a decay estimate for the interaction between annular pieces at different scales.

Lemma 3.3. For any integers $k_1, k_2$ such that $I_{k_i} \cap J_j \neq \emptyset$ for $i = 1, 2$ and some $j \geq 2$,
\[ \|\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_{k_1}} f_1 \Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_{k_2}} f_2\|_{L^q_{1,x}} \lesssim 2^{-k_1-k_2} \|f_1\|_{L^q_x} \|f_2\|_{L^q_x}. \quad (3.3) \]

Proof. We know that the $g_k$ are elliptic with uniform parameters; let these be denoted by $A, \varepsilon$. Since $|\nabla g_k| \sim 1$ on $A_0$, we may decompose $A_0$ as a finite union of balls of radius $c_0$, with $c_0$ sufficiently small that $A_{C_0} \ll \varepsilon$ (as was required for Theorem 2.1). Then (3.3) follows from Corollary 2.2 and the triangle inequality. \qed

Summation. Now we put the pieces together. Using boundedness of the Littlewood– Paley square function, Minkowski’s inequality and the fact that $q \leq 4$, Lemma 3.3, the fact that $\delta > 0$, and finally the fact that $q > p$, we have for any $j \geq 2$ that
\[ \|\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{\{\xi \in J_j\}} f\|_{L^q_{1,x}} \lesssim \left( \sum_k \|\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_k} f\|^q \right)^{\frac{2}{q}} \lesssim \sum_{k_1 \leq k_2} \int \left( |\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_{k_1}} f| |\Lambda P(\nabla)^{\frac{3}{2}} \mathcal{E}_P \chi_{A_{k_2}} f| \right)^{\frac{2}{q}} \, dx \, dt \]
\[ \lesssim \sum_{k_1 \leq k_2} \sum_{k_1 \leq k_2} 2^{-\frac{2}{q}[k_1-k_2]} \|\chi_{A_{k_1}} f\|_{L^q_x} \|\chi_{A_{k_2}} f\|_{L^q_x} \lesssim \sum_k \|\chi_{A_k} f\|_{L^q_x} \lesssim \|f\|_{L^q_x}. \]

This completes the proof of Theorem 1.1, modulo the proof of Theorem 2.1. \qed
We note that related applications of square functions (albeit more complex ones) have also appeared in the work \cite{9, 10} of Carbery–Kenig–Ziesler.

We will give the proof of Theorem 2.1 over the next 6 sections.

4. Preliminary reductions

After making an invertible affine transformation of the frequency space $\mathbb{R}^{1+d}$, we may assume that $\nabla g_1(0) = 0$, that $D^2 g_1(0) = I_d$ (the identity), that $\nabla g_2(0) = e_1$ (the first coordinate vector), and that $D^2 g_2(0)$ is positive definite with eigenvalues comparable to 1. We recall that the hypersurface $S_1$ is at scale $2^{-k_1}$, with $k_1 \geq C_{d,J,\delta} \gg 1$, while $S_2$ is at scale 1.

For $R \geq 1$, let $Q_R$ denote the set

$$Q_R = \{(t, x) \in \mathbb{R}^{1+d} : \frac{1}{2} 2^{k_1(J-2)} R \leq t \leq 2^{k_1(J-2)} R, \ |x| \leq R\}.$$  

The main step in the proof of our bilinear restriction theorem is the following local estimate.

**Proposition 4.1.** For every $\delta > 0$ and $R \geq 1$,

$$\|E_1 f_1 E_2 f_2\|_{L_{t,x}^{d+3} (Q_R)} \lesssim 2^{k_1 \delta} R^d 2^{\frac{1}{d+3}} 2^{(J-2)(d-1)} \|f_1\|_{L_{t,x}^{d}} \|f_2\|_{L_{t,x}^{d+1}}, \quad f_2, f_2 \in L^2. \quad (4.1)$$

The implicit constant is allowed to depend on $d, J, \delta$, but not on $R$ or $k_1$.

The remainder of this section will be devoted to a proof of the sufficiency of Proposition 4.1. By interpolation with the easy estimate

$$\|E_1 f_1 E_2 f_2\|_{L_{t,x}^{d}} \approx \|f_1\|_{L_{t,x}^{d}} \|f_2\|_{L_{t,x}^{d}}, \quad (4.2)$$

it suffices to prove the following “epsilon removal” lemma.

**Lemma 4.2.** Assuming Proposition 4.1, for any $\delta > 0$ and $\frac{d+2}{d} > q > \frac{d+3}{d+1}$,

$$\|E_1 f_1 E_2 f_2\|_{L_{t,x}^{d}} \lesssim 2^{k_1 \delta} 2^{\frac{1}{d+3}} 2^{(J-2)(d-1)} \|f_1\|_{L_{t,x}^{d}} \|f_2\|_{L_{t,x}^{d+1}}. \quad (4.3)$$

**Proof.** For the remainder of the section, we will use the notation $A \lesssim B$ if $A \lesssim \delta 2^{k_1 \delta} B$ for each $\delta > 0$.

Fix a nonnegative $\phi \in C_c^{\infty} (\mathbb{R}^{1+d})$ with $\phi \equiv 1$ on $\{|(t,x)| \leq 1\}$ and $\sum_{m \in \mathbb{Z}^{d+1}} \phi(-m) \sim 1$.

We will actually prove that if the local estimate (4.1) holds for slightly expanded surfaces,

$$S_j = \{(h_j(\xi), \xi) : |\xi| < 6c_0 2^{-k_1}\},$$

and corresponding $E_j$, then the bilinear restriction estimate $R^* (2 \times 2 \rightarrow q)$ holds for all $q > \frac{d+3}{d+1}$, but for simplicity we will gloss over the fact that $6 \neq 1$ by using the same notation for these expanded $S_j, E_j$.

Some adjustments are needed to account for the degeneracy of $S_2$, but the argument is essentially that of \cite{25, 4}. Where the argument is identical, we will be brief.

By interpolation with the $L^2$ estimate, it suffices to prove the weak type estimates

$$\|E_1 f_1 E_2 f_2 > \lambda\| \lesssim 2^{\frac{1}{d+3}} 2^{(J-2)(d-1)} \|f_1\|_{L_{t,x}^{d}} \|f_2\|_{L_{t,x}^{d+1}} \lambda^{-q}, \quad \frac{d+2}{d} > q > \frac{d+3}{d+1}.$$  

This in turn may be reduced to proving that

$$\|\chi_E E_1 f_1 E_2 f_2\|_{L_{t,x}^{d+3}} \lesssim 2^{\frac{1}{d+3}} 2^{(J-2)(d-1)} |E|^{\frac{1}{d}} \|f_1\|_{L_{t,x}^{d}} \|f_2\|_{L_{t,x}^{d+1}}, \quad (4.4)$$
for all Borel sets $E$.

Fix $f_1$. By duality (4.4) would follow from
\[
\|E^*_E(\chi_E E_1 f_1 F_2)\|_{L^2} \lesssim 2^{-\frac{k_1(J-2)(d-1)}{2(d+3)}} |E|^\frac{1}{d} \|f_1\|_{L^2} \|F_2\|_{L^\infty},
\]

By Plancherel,
\[
\|E^*_E(\chi_E E_1 f_1 F_2)\|_{L^2}^2 = \langle (\chi_E E_1 f_1 F_2) * E_2, \chi_E E_1 f_1 F_2 \rangle.
\]

Let
\[
R_2 := \max \{1, 2^{k_1(J-2)/d} (\frac{d+4}{4d+1}) |E|^{1/4} (\frac{d+4}{4d+1} - \frac{1}{d}) \},
\]

and define functions
\[
\phi_{2,R_2}(t,x) = \phi(\frac{t}{R_2^2}, \frac{x}{R_2^2}), \quad \phi_{2,R_2}^c = 1 - \phi_{2,R_2} \\
\psi_{2,R_2} = \phi_{2,R_2} E_2 1, \quad \psi_{2,R_2}^c = \phi_{2,R_2}^c E_2 1.
\]

By stationary phase,
\[
\|\psi_{2,R_2}^c\|_{L^\infty} \lesssim R_2^{-\frac{d}{2}}.
\]

By Hölder and Stein–Tomas (rescaled),
\[
\|\chi_E E_1 f_1 F_2\|_{L^1} \lesssim 2^{\frac{k_1(J-2)d}{2(d+3)}} |E|^{\frac{2}{2d+1}} \|f_1\|_{L^2} \|F_2\|_{L^\infty}.
\]

Hence
\[
\langle (\chi_E E_1 f_1 F_2) * \psi_{2,R_2}^c, \chi_E E_1 f_1 F_2 \rangle \lesssim 2^{\frac{k_1(J-2)d}{2(d+3)}} R_2^{-\frac{d}{2}} |E|^{\frac{2}{2d+1}} \|f_1\|_{L^2}^2 \|F_2\|_{L^\infty}^2.
\]

Plugging in the value of $R_2$, we see that this is acceptable, so we turn to the main term.

Let $\mu_2$ denote surface measure on $S_2$; then (using rapid decay of $\psi_{2,R_2}$),
\[
\mathcal{F}(\psi_{2,R_2}) = \mathcal{F}(\phi_{2,R_2}) * \mu_2 \lesssim \sum_{j=0}^{\infty} 2^{-Mj} R_2 S_{2,2^j R_2^{-1}},
\]

where $M$ is sufficiently large for later purposes and
\[
S_{2,2^j R_2^{-1}} := \begin{cases} 
\{(\tau, \xi) : |\xi| < 2c_0, |\tau - h_2(\xi)| < 2^j R_2^{-1} \}, & 2^j \ll R_2 \\
\{(\tau, \xi) : |(\tau, \xi)| < 2^j R_2^{-1} \}, & 2^j \geq R_2.
\end{cases}
\]

Using this and Plancherel,
\[
\langle (\chi_E E_1 f_1 F_2) * \psi_{2,R_2}, \chi_E E_1 f_1 F_2 \rangle = \|\mathcal{F}(\chi_E E_1 f_1 F_2)\|_{L^2_{t,\xi}(\mathcal{F}(\psi_{2,R_2}))}^2 
\lesssim \sum_{j=0}^{\infty} 2^{-Mj} R_2 \|\mathcal{F}(\chi_E E_1 f_1 F_2)\|_{L^2_{t,\xi}(S_{2,2^j R_2^{-1}})}^2.
\]

By a simple covering argument (and translation invariance of our inequality), it suffices to consider the $j = 0$ case. We want
\[
\|\mathcal{F}(\chi_E E_1 f_1 F_2)\|_{L^2_{t,\xi}(S_{2,R_2^{-1}})} \lesssim 2^{\frac{k_1(J-2)(d-1)}{2(d+3)}} R_2^{-\frac{1}{2}} |E|^{\frac{1}{d}} \|f_1\|_{L^2} \|F_2\|_{L^\infty}.
\]

By Plancherel and duality, this is equivalent to
\[
\|\chi_E E_1 f_2 \hat{E}_2 \hat{f}_2\|_{L^1_{t,\xi}} \lesssim 2^{\frac{k_1(J-2)(d-1)}{2(d+3)}} R_2^{-\frac{1}{2}} |E|^{\frac{1}{d}} \|f_1\|_{L^2} \|\hat{f}_2\|_{L^2_{\xi}},
\]

where
\[
\hat{E}_2 \hat{f}_2 = \mathcal{F}^* (\chi_{S_{2,R_2}} \hat{f}_2).
\]
Now fix $\tilde{f}_2 \in L^2$. By duality and Plancherel, (4.5) is equivalent to
\[
\langle (\chi_E F_1 \tilde{E}_2 f_2) * \chi_{E^c} 1, \chi_E F_1 \tilde{E}_2 f_2 \rangle \lesssim 2^{k_1(J-2)\frac{(d-1)}{d+3}} R_2^{-1} |E|^{\frac{3}{d}} \|F_1\|_{L_{\infty}^\infty} \|f_2\|_{L_{\tau, \xi}^2}^2.
\]
Let
\[
R_1 := \max\{1, 2^{-\frac{2k_1(J-2)\frac{(d-1)}{d+3}}{d+3}} |E|^{\frac{d+3}{d+4} - \frac{1}{d+1}} \},
\]
and define
\[
\phi_{1, R_1}(t, x) = \phi\left( \frac{t-t_0}{(\frac{d-1}{2})(d+3)} R_1^{\frac{d+3}{d+4}} \right), \quad \phi_{1, R_1}^c = 1 - \phi_{1, R_1},
\]
\[
\psi_{1, R_1} = \phi_{1, R_1} \chi_{E^c}, \quad \psi_{1, R_1}^c = \phi_{1, R_1}^c \chi_{E^c}.
\]
Using stationary phase\(^{2}\), Hölder, and Stein–Tomas as before,
\[
\langle (\chi_E F_1 \tilde{E}_2 f_2) * \psi_{1, R_1}^c, \chi_E F_1 \tilde{E}_2 f_2 \rangle \lesssim R_1^{-\frac{d+3}{d+2}} |E|^{\frac{d+3}{d+2}} \|F_1\|_{L_{\infty}^\infty} \|f_2\|_{L_{\tau, \xi}^2}^2,
\]
which is acceptable.

We compute
\[
\mathcal{F}(\psi_{1, R_1}) = \mathcal{F}(\phi_{1, R_1}) * \mu_1 \lesssim 2^{k_1(J-2)} R_1 \sum_{j=0}^{\infty} 2^{-Mj} \chi_{S_1, 2^j R_{1-1}},
\]
where
\[
S_1, 2^j R_{1-1} := \begin{cases} \{ (\tau, \xi) : |\xi| < 2c_0 2^{-k_1}, |\tau - h_1(\xi)| < 2^{J-1} 2^{-k_1(J-2)} R_{1-1} \}, & 2^j \ll R_1, \\ \{ (\tau, \xi) : |\xi| < 2^j R_1, |\tau| < 2^{J-1} 2^{-k_1(J-2)} R_{1-1} \}, & 2^j \gg R_1. \end{cases}
\]
Thus to estimate the main term, it suffices to show that
\[
\|\mathcal{F}(\chi_E F_1 \tilde{E}_2 f_2)\|_{L_{\tau, \xi}^\infty(S_{1, R_1})} \lesssim 2^{k_1(J-2)\left(-\frac{d+3}{d+4} + \frac{1}{d+3} - \frac{1}{d+1}\right)} (R_1 R_2)^{-\frac{1}{2}} |E|^{\frac{3}{d}} \|F_1\|_{L_{\infty}^\infty} \|\hat{f}_2\|_{L_{\tau, \xi}^2},
\]
or equivalently,
\[
\|\chi_E \tilde{E}_2 f_2\|_{L_{\tau, \xi}^1(S_{1, R_1})} \lesssim 2^{k_1(J-2)\left(-\frac{d+3}{d+4} + \frac{1}{d+3} - \frac{1}{d+1}\right)} (R_1 R_2)^{-\frac{1}{2}} |E|^{\frac{3}{d}} \|\hat{f}_1\|_{L_{\tau, \xi}^2} \|\hat{f}_2\|_{L_{\tau, \xi}^2},
\]
where
\[
\tilde{E}_2 f_1 = \mathcal{F}^*(\chi_{S_1, R_1} f_1).
\]
By Hölder and the definition of $R_1, R_2$, this would follow from
\[
\|\tilde{E}_2 f_2\|_{L_{\tau, \xi}^\infty(S_{1, R_1})} \lesssim 4^{k_1, \delta (R_1 R_2)^{\delta} 2^{k_1(J-2)\left(-\frac{d+3}{d+4} + \frac{1}{d+3} - \frac{1}{d+1}\right)} (R_1 R_2)^{-\frac{1}{2}} \|\hat{f}_1\|_{L_{\tau, \xi}^2} \|\hat{f}_2\|_{L_{\tau, \xi}^2}, \quad \delta > 0.
\]

In proving (4.7), we may assume that $\text{supp} \hat{f}_j \subseteq S_{j, R_j}$, $j = 1, 2$. To avoid a proliferation of tildes, we will let $\tilde{E}_j f := \mathcal{F}^*(\chi_{S_{j, 3R_j}} f)$. Let $\varphi$ be a smooth non-negative function with $\sum_{m \in \mathbb{Z}^{d+1}} \varphi(-m) = 1$ and $\varphi$ supported in $\{(\tau, \xi) : |(\tau, \xi)| \leq 1\}$. For $(t_0, x_0) \in \mathbb{R}^{1+d}$, define
\[
\varphi(t_0, x_0) = \varphi\left( \frac{t-t_0}{(\frac{d-1}{2})(d+3)} R_1 R_2 \right), \quad \varphi(t_0, x_0) = \varphi\left( \frac{t-t_0}{(\frac{d-1}{2})(d+3)} R_1 R_2 \right).
\]
Then
\[
\sum_{(t_0, x_0)} \varphi(t_0, x_0)^2 \sim \sum_{(t_0, x_0)} \varphi(t_0, x_0)^2 \sim \sum_{(t_0, x_0)} \varphi(t_0, x_0)^2 \sim 1.
\]

\(^{2}\)In fact, a better stationary phase estimate is possible, but we use the one that also works when $h_1$ is replaced by $2^{-(J-2)k_1L_1}$; similarly for (4.6).
where the sum is taken over \((t_0, x_0) \in (2^{k_1(J - 2)} R_1 \mathbb{Z}) \times (R_2 \mathbb{Z})^d\).

Using the triangle inequality and (4.8), our assumptions on the supports of \(\phi, \hat{\varphi}\), the local restriction estimate with Fubini, and finally Cauchy–Schwarz, Plancherel, and (4.8) again,

\[
\|\tilde{E}_1 \tilde{f}_1 \tilde{E}_2 f_2\|_{L^{d+3}_{\tau,x}(R^{d+1})} \lesssim \sum_{(t_0, x_0)} \|\phi_{R_1, R_2}((\varphi_{R_1, R_2}) \tilde{E}_1 \tilde{f}_1)(\varphi_{R_1, R_2}) \tilde{E}_2 f_2\|_{L^{d+3}_{\tau,x}(R^{d+1})} \\
\lesssim \sum_{(t_0, x_0)} \|\tilde{E}_1(\varphi_{R_1, R_2} \ast \tilde{f}_1) \tilde{E}_2(\varphi_{R_1, R_2} \ast \tilde{f}_2)\|_{L^{d+3}_{\tau,x}(Q_{R_2})} \\
\lesssim \delta \sum_{(t_0, x_0)} 2^{k_1 \delta} R_2^{\delta} 2^{k_1(J - 2)} R_1 R_2 \frac{1}{2} 2^{k_1(J - 2)(d - 1)} \\
\times \|\mathcal{F}(\varphi_{R_1, R_2} \ast \tilde{f}_1)\|_{L^2_{\tau,x}} \|\mathcal{F}(\varphi_{R_1, R_2} \ast \tilde{f}_2)\|_{L^2_{\tau,x}} \\
\lesssim \delta 2^{k_1 \delta} R_2^{\delta} 2^{k_1(J - 2)} R_1 R_2 \frac{1}{2} 2^{k_1(J - 2)(d - 1)} \|\tilde{f}_1\|_{L^2_{\tau,x}} \|\tilde{f}_2\|_{L^2_{\tau,x}},
\]

which is what we wanted. This completes the proof. \(\square\)

5. Induction

Let \(R^*(2 \times 2 \to \frac{d+3}{d+1}; \delta, \alpha)\) denote the statement that the local estimate

\[
\|\mathcal{E}_1 f_1 \mathcal{E}_2 f_2\|_{L^{d+3}_{\tau,x}(Q_R)} \lesssim \delta, \alpha, 2^{k_1 \delta} R_2^{\delta} 2^{k_1(J - 2)(d - 1)} \|f_1\|_{L^2_{\tau}} \|f_2\|_{L^2_{\tau}}, \tag{5.1}
\]

holds for all \(R \geq 1\) and \(f_1, f_2 \in L^2_{\tau}\).

**Lemma 5.1.** For all \(\varepsilon > 0\), \(R(2 \times 2 \to \frac{d+3}{d+1}; \varepsilon, \frac{d^2 - 1}{2(d^2 + 3)} + \frac{1}{2})\) holds.

Assuming the lemma, Proposition 4.1 would follow from

\[
R^*(2 \times 2 \to \frac{d+3}{d+1}; \delta, \alpha) \implies R^*(2 \times 2 \to \frac{d+3}{d+1}; \delta + C \varepsilon, \max\{(1 - \varepsilon)\alpha, C \varepsilon\} + C \varepsilon'),
\]

\[
(5.2)
\]

for all \(\alpha > 0\) and \(1 > \delta, \varepsilon, \varepsilon' > 0\). We will prove (5.2) in Sections 6–9, using Wolff’s induction on scales argument from [28] (more precisely, a variant of Tao’s adaptation in [24]). We turn now to the proof of Lemma 5.1.

**Proof of Lemma 5.1.** Let \(\delta > 0\). Assume henceforth that \(R \gtrsim 2^{k_1(J - 2)}\) and that \(\|f_1\|_{L^2_\tau} = \|f_2\|_{L^2_\tau} = 1\). It suffices to prove

\[
\|\mathcal{E}_1 f_1 \mathcal{E}_2 f_2\|_{L^{d+3}_{\tau,x}(Q_R)} \lesssim 2^{k_1(J - 2)} \frac{d^2 - 1}{2(d^2 + 3)} R^\frac{d^2 - 1}{2(d^2 + 3)},
\]

which would establish \(R^*(2 \times 2 \to \frac{d+3}{d+1}; \delta, \frac{d^2 - 1}{2(d^2 + 3)}, \frac{d^2 - 1}{2(d^2 + 3)})\).

By Hölder’s inequality,

\[
\|\mathcal{E}_1 f_1 \mathcal{E}_2 f_2\|_{L^{d+3}_{\tau,x}(Q_R)} \lesssim |Q_R|^{\frac{d+3}{d+1} - \frac{1}{2}} \|\mathcal{E}_1 f_1\|_{L^2_{\tau,x}} \|\mathcal{E}_2 f_2\|_{L^2_{\tau,x}} \\
\lesssim 2^{k_1(J - 2)} \frac{d^2 - 1}{2(d^2 + 3)} R^\frac{d^2 - 1}{2(d^2 + 3)} \|f_1\|_{L^2_\tau} \|\mathcal{E}_2 f_2\|_{L^2_{\tau,x}(Q_R)},
\]

so it suffices to show that

\[
\|\mathcal{E}_2 f_2\|_{L^2_{\tau,x}(Q_R)} \lesssim R^\frac{1}{2}, \tag{5.3}
\]
When \( k_1 = 0 \), (5.3) just follows from Hölder’s inequality (in the time direction) and Plancherel. Inspired by this, we decompose

\[
Q_R = \bigcup_{j=0}^{2^{k_1(j-2)}} Q'_j,
\]

where

\[
Q'_j = Q_j \cap Q_R, \quad \text{and} \quad Q_j = \{(t, x) : R_j \leq t \leq R(j + 1), |x| \leq R\}.
\]

The idea of the proof of (5.3) is that on \( Q'_j \), \( \mathcal{E}_2 f_2 \) is well-approximated by a function \( f_2^{(j)} \) whose extension is spatially localized at time \( R_j \). Moreover, for \( j \neq k \), these pieces are essentially orthogonal.

To make this heuristic rigorous, fix a smooth, non-negative function \( \phi \) with \( \phi \equiv 1 \) on \( \{|\xi| < 2\} \) and \( \phi \equiv 0 \) off \( \{|\xi| < 3\} \). For \( j \in \mathbb{Z} \), define

\[
f^{(j)}_2(\xi) := e^{-iR j h_2(\xi)} \phi(\frac{\xi}{c_0}) |\phi(\frac{\xi}{c_0})| \mathcal{E}_2 f_2(R_j, x)(\xi),
\]

where the inner Fourier transform is taken with respect to the \( x \) variable.

**Lemma 5.2.** For \( (t, x) \in Q'_j \) and \( M \geq 0 \),

\[
|\mathcal{E}_2 f_2(t, x) - \mathcal{E}_2 f^{(j)}_2(t, x)| \lesssim_M R^{-M}. \tag{5.4}
\]

**Lemma 5.3.**

\[
\sum_{j=0}^{2^{k_1(j-2)}} \|f^{(j)}_2\|_{L^2_x}^2 \lesssim \|f_2\|_{L^2_x}^2. \tag{5.5}
\]

We postpone the proofs of Lemmas 5.2 and 5.3 while we complete the proof of (5.3). Choosing \( M \) sufficiently large depending on \( \delta \), and using (5.4) together with Hölder’s inequality, Plancherel, and finally (5.5),

\[
\|\mathcal{E}_2 f_2\|_{L^2_{t,x}(Q_R)} \lesssim 1 + \left( \sum_j \|\mathcal{E}_2 f^{(j)}_2\|_{L^2_{t,x}(Q'_j)}^2 \right)^{\frac{1}{2}} \lesssim 1 + \left( \sum_{j=0}^{2^{k_1(j-2)}} R \|f^{(j)}_2\|_{L^2_x}^2 \right)^{\frac{1}{2}} \lesssim R^{\frac{1}{2}},
\]

and (5.3) (and hence Lemma 5.1) is proved.

**Proof of Lemma 5.2.** Because \( \text{supp } f_2 \subseteq \{|\xi| < c_0\} \),

\[
\mathcal{E}_2 f_2(t, x) = \iiint e^{i(t-R_j x-y)(h_2(\eta), \eta)} \phi(\frac{\eta}{c_0}) e^{i(R_j y)(h_2(\xi), \xi)} \phi(\frac{\xi}{c_0}) f_2(\xi) \, d\xi \, dy \, d\eta.
\]

Thus

\[
\mathcal{E}_2 f_2(t, x) - \mathcal{E}_2 f^{(j)}_2(t, x) = \int P(t, x; \xi) e^{iR j h_2(\xi)} \phi(\frac{\xi}{c_0}) f_2(\xi) \, d\xi,
\]

where

\[
P(t, x; \xi) = \int e^{i(t-R_j x-y)(h_2(\eta), \eta)} \phi(\frac{\eta}{c_0}) \, d\eta e^{iy \xi (1 - \phi(\frac{\eta}{c_0}))} \, dy. \tag{5.7}
\]

For \( (t, x) \in Q'_j \) and \( |y| > CR \), \( |t - R_j| \leq R \) and \( |x - y| \geq |y| - R \), so

\[
|\nabla_{\eta}(t-R_j x-y)(h_2(\eta), \eta)| = |(t-R_j) \nabla h_2(\eta) + (x-y)| \gtrsim |y|,
\]

so integrating by parts in the inner integral of (5.7),

\[
|P(t, x; \xi)| \lesssim_M \int (1 + |y|)^{-(M+d)} (1 - \phi(\frac{\eta}{c_0})) \, dy \lesssim R^{-M}.
\]
Inserting this in (5.6) and using Hölder (and \( \|f_2\|_{L^2_\xi} \sim 1 \)) gives
\[
|\mathcal{E}_2 f_2(t, x) - \mathcal{E}_2 f_2^{(j)}(t, x)| \lesssim_M R^{-M} \|f_2\|_{L^2_\xi} \lesssim R^{-M}.
\]

**Proof of Lemma 5.3.** Define
\[
T_j f(\xi) = e^{-iR_k h_2(\xi)} \hat{f}(\xi) \phi(\frac{\xi}{c_0}) \mathcal{E}_2(\phi(\frac{\eta}{c_0})f(\eta))(R_j, \cdot)\hat{\eta}(\xi).
\]
Then by the support condition on \( f \), that is, \( \|f\|_{L^2_{k\delta}} \leq 1 \), we compute
\[
\|T_k T_j f\|_{L^2_\xi} = \|T_k T_j f\|_{L^2_\xi} \lesssim R^{-M}(1 + |k - j|)^{-M} \|f\|_{L^2_\xi}, \quad f \in L^2;
\]
by Plancherel, this is also valid for \( |k - j| \lesssim 1 \).

By (5.8) and Cotlar–Stein,
\[
\|f\|_{L^2_\xi}^2 \gtrsim \sum_{j=1}^{2^{k_1(J-2)}} \|T_j f\|_{L^2_\xi}^2 = \sum_{j=1}^{2^{k_1(J-2)}} \|T_j f\|_{L^2_\xi}^2 + \sum_{1 \leq j \neq k \leq 2^{k_1(J-2)}} (T_j f, T_k f)
\]
\[
\geq \sum_{1 \leq j \neq k \leq 2^{k_1(J-2)}} \|T_j f\|_{L^2_\xi}^2 - C_M \sum_{j=1}^{2^{k_1(J-2)}} \sum_{k \neq j} (j - k)^{-M} \|f\|_{L^2_\xi}^2
\]
\[
\geq \sum_{1 \leq j \neq k \leq 2^{k_1(J-2)}} \|T_j f\|_{L^2_\xi}^2 - C_M 2^{k_1(J-2)} R^{-M} \|f\|_{L^2_\xi}^2.
\]
Using our lower bound \( R \gtrsim 2^{k_1(J-2)/\delta} \), we obtain (5.5). \( \square \)

**Notation.** We recycle notation, and will say for the remainder of the article that \( A \lesssim_B \) if \( A \lesssim 2^{k_1} R^\varepsilon B \) for all \( \varepsilon > 0 \).

Thus we want to show that
\[
\|\mathcal{E}_1 f_1 E_2 f_2\|_{L^2_{\xi, x}(Q_{\alpha})} \lesssim (\mathcal{R}(1-\varepsilon)^{\alpha} + R^{C_\varepsilon}) 2^{k_1(J-2)/(\delta + 1)} \|f_1\|_{L^2_\xi}^2 \|f_2\|_{L^2_\xi},
\]
and we assume (for the remainder of the argument) that \( \mathcal{R}(2 \rightarrow 2 \rightarrow \frac{d+3}{\delta+1}; \delta, \alpha) \) holds, that \( R \gtrsim 2^{k_1(J-2)/\delta} \), and \( \|f_1\|_{L^2_\xi} = \|f_2\|_{L^2_\xi} = 1 \).
6. Wave packet decomposition

We recall that
\[ Q_R := \{(t,x) : \frac{1}{2}2^{k_1(J-2)} R \leq t \leq 2^{k_1(J-2)} R, \ |x| \leq R \} . \]
Define
\[ X_j := R^{\frac{1}{2}} \mathbb{Z}^d, \quad \Xi_j := (\mathbb{R}^{-\frac{1}{2}} \mathbb{Z}^d) \cap B(0, 4c_0 2^{-k_j}), \quad V_j := \nabla h_j(\Xi_j), \quad j = 1,2 . \]
For \( j = 1,2 \), an \( S_j \)-tube is a set of the form
\[ T_j = \{(t,x) : |x-x_j(T_j)| + tv_j(T_j) < R^\frac{1}{2} \} , \]
where \( x_j(T_j) \in X_1 \) and \( v_j(T_j) \in V_j \).

**Proposition 6.1.** There exist coefficients \((c_{T_j})\) and wave packets \((\phi_{T_j})\), indexed in those \( S_j \)-tubes \( T_j \) satisfying \( \text{dist}(T_j, Q_R) \lesssim R \), such that for any \( M > 0 \),
\[ \|E_1 f_1 E_2 f_2 \|_{L^2, \mathcal{X}^+_{k_1,\alpha}(Q_R)} \lesssim \sum_{T_1} c_{T_1} \|\phi_{T_1}^j \|_{L^2, \mathcal{X}^+_{k_1,\alpha}(Q_R)} + O(1) . \]  
Furthermore, the following hold for each \( j = 1,2 \) and every tube \( T_j \) appearing in the sum:
\[ \|c_{T_j}\| \lesssim 1 \]  
\[ \phi_{T_j} = E_j \phi_{T_j}(0, \cdot) \]  
\[ \text{supp} \phi_{T_j}(0, \cdot) \subseteq \{ |\xi - \xi_j(T_j)| \lesssim R^{1/2} \} \]  
\[ |\phi_{T_j}(t,x)| \lesssim R^{-\frac{1}{4}} \left(1 + \frac{|x-x_j(T_j)| + tv_j(T_j)}{R^{1/2}} \right)^{-M} , \quad (t,x) \in Q_R \]  
\[ \|\sum_{T_j} c_{T_j} \phi_{T_j}(t, \cdot)\|_{L^2} \lesssim \|c_{T_j}\|_{L^2} , \quad \text{for all } (c_{T_j}) \in L^2, \ t \in \mathbb{R} . \]

The proof of this proposition will occupy the remainder of the section.

We begin with the decomposition of \( E_2 f_2 \). Heuristically, an \( S_2 \)-wave packet is concentrated on a tube that is transverse to the long axis of \( Q_R \), so on \( Q_R \) it should be concentrated on a tube of diameter \( R^{\frac{1}{2}} \) and length \( R \). Unfortunately, this heuristic neglects the role of dispersion, which means that we cannot simply decompose the “initial data” \( E_2 f_2(0, \cdot) \) into pieces with Fourier support on \( R^{\frac{1}{2}} \) balls and spatial concentration on \( R^{\frac{3}{2}} \) balls, and then propagate that decomposition forward. Instead, we will apply Tao’s elliptic wave packet decomposition [24] to \( E_2 f_2^{(j)} \) on \( Q'_j \). For the convenience of the reader, we give the precise statement we need and a proof.

**Lemma 6.2 ([24]).** For each \( 0 \leq j \leq 2^{k_1(J-2)} \), there exist coefficients \((c_{T_2}^{(j)})\) and wave packets \((\phi_{T_2}^{(j)})\), indexed in those tubes \( T_2 \) with \( \text{dist}(T_2, Q'_j) \lesssim R \), that satisfy (6.3-6.6), with the superscripts \((j)\) inserted, as well as
\[ \|c_{T_2}^{(j)}\|_{L^2} \lesssim \|f_2^{(j)}\|_{L^2} , \]  
\[ E_2 f_2^{(j)}(t,x) = \sum c_{T_2}^{(j)} \phi_{T_2}^{(j)} + O(R^{-M}) , \quad M > 0 , \quad (t,x) \in Q'_j . \]

**Proof.** We begin by decomposing \( f_2^{(j)} \).

Fix a partition of unity \( \mathbf{1} = \sum_{\xi_2 \in \Xi_2} \varphi_{\xi_2} \), where \( \varphi_{\xi_2}(\xi) = \varphi(R^{1/2}(\xi - \xi_2)) \) for some smooth compactly supported function \( \varphi \). We will show how to decompose
the decomposition of $\mathcal{E}_2 f_2^{(0)}$ may be obtained by summing over $\xi_2 \in \Xi_2$ and using the near-orthogonality of these localized pieces. Let $f_{\xi_2}^{(0)} := \varphi_{\xi_2} f^{(0)}$.

By the Poisson summation formula, we can write $1 \equiv \sum_{x_2 \in X_2} (\psi_{x_2})^2$, where $\psi_{x_2}(x) = \psi(R^{-1/2}(x - x_2))$ for some nonnegative $\psi$ with $\hat{\psi}$ smooth and compactly supported.

Let $T_2$ be the tube with parameters $x_2, \xi_2$, and define
$$c_{T_2} = \|\psi_{2\xi_2} f_{\xi_2}^{(0)}\|_{L^2_{\xi}},$$
$$\phi_{T_2} = c_{T_2}^{-1} \mathcal{E}_2 (\hat{\psi}_{x_2} * \hat{\psi}_{x_2} * f_{\xi_2}^{(0)}).$$

The identity (6.3) and support condition (6.4) are immediate. The $\ell^2$ estimate (6.7) is just Plancherel:
$$\| (c_{T_2}) \|_{\ell^2_{\xi_2}} = \| f_{\xi_2}^{(0)} \|_{L^2_{\xi_2}}.$$

As we will see, standard (non)stationary phase arguments imply the decay estimate
$$|\phi_{T_2}(t, x)| \lesssim_M R^{-d/2} \left(1 + \frac{|x - x(T_2) + t\nu(T_2)|}{R^2}\right)^{-M},$$
if $|t| \lesssim R$ or $|x - x_j| \lesssim R$. (6.9)

Assuming (6.9) for a moment, (6.5) and (6.8) are immediate. By (6.3), it suffices to verify (6.6) for $t = 0$, and for that we apply (6.9) and Schur’s test.

Now for (6.9). We write
$$c_{T_2} \phi_{T_2}(t, x) = R^{d/2} \int \hat{\psi}_{x_2} * f_{\xi_2}^{(0)}(\eta) e^{i x_2 \eta} \int e^{i(t h_2(\xi) + (x - x_2) \xi)} \hat{\psi}(R^{1/2}(\xi - \eta)) d\eta. \tag{6.10}$$

Let
$$K_{T_2}(t, x, \eta) := R^{d/2} \int e^{i(t h_2(\xi) + (x - x_2) \xi)} \hat{\psi}(R^{1/2}(\xi - \eta)) d\xi,$$
and assume that $\eta \in \text{supp} \hat{\psi}_{x_2} * f_{\xi_2}^{(0)}$. Thus for $\xi$ in the region of integration for $K_{T_2}(t, x, \eta)$, $|\xi - \xi_2| \leq |\xi - \eta| + |\eta - \xi_2| \lesssim R^{1/2}$.

From Hölder we obtain $|K_{T_2}(t, x, \eta)| \lesssim 1$. Now assume that
$$|x - x_2 + t\nabla h_2(\xi_2)| \gg R^{1/2}. \tag{6.11}$$

If $|t| \lesssim R$, $t\nabla h_2(\xi) = t\nabla h_2(\xi_2) + O(R^{1/2})$; if $|t| \gg R$ but $|x - x_2| \lesssim R$,
$$|x - x_2 + t\nabla h_2(\xi)| \sim |t\nabla h_2(\xi)| \gg R,$$

since $|\nabla h_2(\xi)| \sim 1$ throughout the region of integration. In either case,
$$|x - x_2 + t\nabla h_2(\xi)| \sim |x - x_2 + t\nabla h_2(\xi_2)|,$$

and integrating by parts $M$ times gives
$$|K_{T_2}(t, x, \eta)| \lesssim (1 + \frac{|x - x(T_2) + t\nu(T_2)|}{R^2})^{-M}. $$

Inserting this into (6.10) and applying Hölder,
$$|c_{T_2} \phi_{T_2}(t, x)| \lesssim R^{-d/2} \|\hat{\psi}_{x_2} * f_{\xi_2}^{(0)}\|_{L^2_{\xi_2}},$$
if $|t| \lesssim R$ or $|x - x_2| \lesssim R$,

which is (6.9).

Given $0 \leq j \leq 2^k (J - 2)$, we can apply exactly the above argument to write
$$\mathcal{E}_2 [e^{i R^j h_2(\xi) f_2^{(j)}}](t, x) = \sum_{T_2} c_{T_2} \phi_{T_2} + O(R^{-M}), \text{ on } Q'_0.$$
Now we translate. Our constants are the same: \( c_{T_2}^{(j)} := c_{T_1} \), but our wave packets are shifted: \( \tilde{\phi}_{T_2}^{(j)}(t,x) := \phi_{T_2}(t-Rj,x) \). Thus \( \tilde{\phi}_{T_2}^{(j)} \) is associated to a tube with parameters \( x_j-Rj \) and \( \xi_2 \), where \( x_2, \xi_2 \) are the parameters for \( \phi_{T_2} \). The conclusions claimed in the lemma are then immediate from those obtained in the case \( j = 0 \), and we are done. \( \square \)

Now let \( \Lambda \subseteq \{0,1,\ldots,2^{k_1(j-2)}\} \) be a C-separated set for some sufficiently large \( C \). Applying the decomposition in Lemma 6.2 to each of the functions \( f_{2}^{(j)} \), and then using the estimate in Lemma 5.2, together with the assumption \( R \gtrsim 2^{d_1} \), we obtain

\[
\mathcal{E}_2 f_2(t,x) = \sum_{T_2} c_{T_2} \phi_{T_2} + O(R^{-M}), \quad (t,x) \in \bigcup_{j \in \Lambda} Q'_j, \tag{6.12}
\]

where the tubes appearing in the sum all lie within a distance \( O(R) \) of one of the \( Q'_j \) with \( j \in \Lambda \). The conclusions (6.3-6.5) follow immediately from Lemma 6.2. Inequality (6.2) just follows from (6.7) and Lemma 5.3, and finally, (6.6) is just a consequence of the corresponding conclusion (with superscripts \( j \) inserted) in Lemma 6.2.

Now we turn to the wave packet decomposition of \( \mathcal{E}_1 f_1 \), which is essentially a rescaling of the elliptic case.

**Lemma 6.3.** There exist coefficients \( (c_{T_1}) \) and wave packets \( (\phi_{T_1}) \), indexed in those tubes with \( \text{dist}(T_1, Q_R) \lesssim R \) and satisfying (6.2-6.6), as well as

\[
\mathcal{E}_1 f_1(t,x) = \sum_{T_1} c_{T_1} \phi_{T_1} + O(R^{-M}), \quad (t,x) \in Q_R. \tag{6.13}
\]

**Proof.** We change variables to write

\[
\tilde{\mathcal{E}}_1 f_1(t,x) = 2^{-\frac{d_1}{2}} \tilde{\mathcal{E}}_1 \tilde{f}_1(2^{-d_1/2} t, 2^{-k_1/2} x),
\]

where \( \tilde{\mathcal{E}}_1 \) is extension from the (elliptic) hypersurface \( \{(g_1(\xi), \xi) : |\xi| < c_0\} \) and \( \tilde{\phi}_1(\xi) := 2^{-\frac{d_1}{2}} \phi_1(2^{-d_1/2} \xi) \). Now apply the same decomposition procedure as the one used for \( \tilde{\mathcal{E}}_2 f_2^{(0)} \) in Lemma 6.2. We note that the long-time decay (6.9) for \( |x-x_2| \lesssim R \) was due to the transversality \( |\nabla h_2| \sim 1 \); \( \nabla g_1 \) is not transverse, but we do not need the long-time decay anyway. Decomposing into packets of frequency width \( 2^{k_1} R^{-1/2} \),

\[
\tilde{\mathcal{E}}_1 \tilde{f}_1(t,x) = \sum_{\tilde{T}_1} c_{\tilde{T}_1} \phi_{\tilde{T}_1},
\]

where the sum is taken over tubes of width \( 2^{-k_1} R^{1/2} \) with base points \( \tilde{x}_1(\tilde{T}_1) \) in \( 2^{-k_1} R^{1/2} \mathbb{Z}^d \) and directions \( \tilde{v}_1(\tilde{T}_1) = \nabla g_1(\tilde{x}_1(\tilde{T}_1)) \) in \( \nabla g_1(2^{k_1} R^{-1/2} \mathbb{Z}^d \cap B(0, 4c_0)) \). (The formula (6.8) was not exact only because we just summed over tubes close to \( Q'_0 \).) Furthermore, we have the estimates

\[
\|\phi_{\tilde{T}_1}(\cdot)\|_{L^2_{\tilde{T}_1}} \lesssim \|\tilde{\phi}_1\|_{L^2_1} = 1 \tag{6.14}
\]

\[
\tilde{\phi}_{\tilde{T}_1} = \tilde{\mathcal{E}}_1 \phi_{\tilde{T}_1}(0, \cdot) \tag{6.15}
\]

\[
\text{supp} \tilde{\phi}_{\tilde{T}_1}(0, \cdot) \subseteq \{ |\xi - \tilde{x}_1(\tilde{T}_1)| \lesssim 2^{k_1} R^{-1/2} \} \tag{6.16}
\]

\[
|\phi_{\tilde{T}_1}(t,x)| \lesssim (2^{k_1} R^{-1/2}) \frac{d}{2} (1 + \frac{|x-\tilde{x}_1(\tilde{T}_1)|}{2^{k_1} R^{1/2}})^{-M}, \quad |(t,x)| \lesssim 2^{-2k_1} R \tag{6.17}
\]
Proof of Proposition 6.1. Given a

Inequalities (6.2-6.6) follow immediately from (6.14-6.18), and the estimate (6.13) (wherein the sum is taken over tubes within $O(R)$ of $Q_R$) follows from the decay estimate.

From here, the proof of Proposition 6.1 is quick.

Proof of Proposition 6.1. Given a $C$-separated subset $\Lambda \subseteq \{0, \ldots, 2^{k_1(j-2)}\}$, let $(c_{T_2}^A)$, $(\phi_{T_2}^A)$ denote the coefficients and wave packets appearing in (6.12). Then

$$|E_{1} f_1 E_{2} f_2(t, x)| \leq \sum_{\Lambda} |T_{T_2} c_{T_2}^A \phi_{T_2}^A| + O(R^{-M}), \quad (t, x) \in Q_R,$$

where the sum is taken over a disjoint collection of $C$ such $\Lambda$’s. Combining this with the wave packet decomposition in Lemma 6.3 and the fact that $|E_{j} f_j| \lesssim 1$ (because $||f_j||_{L^2} \lesssim ||f_j||_{L^2} = 1$)

$$|E_{1} f_1 E_{2} f_2(t, x)| \leq \sum_{\Lambda} |T_{T_1} c_{T_1} \phi_{T_1} T_{T_2} c_{T_2}^A \phi_{T_2}^A| + O(R^{-M}).$$

The estimate (6.1) follows from Hölder, the triangle inequality, the pigeonhole principle, which lets us pick a single $\Lambda$, and $R \gtrsim 2^{k_1}$. The properties (6.2-6.6) have already been established, so we are done.

7. The local and global terms

The wave packet decomposition allows for a number of reductions. First, it suffices to show

$$\left\| \left( \sum_{T_1} c_{T_1} \phi_{T_1} \right) \left( \sum_{T_2} c_{T_2}^A \phi_{T_2}^A \right) \right\|_{L^\infty_{t,x}}(Q_R) \lesssim 2^{k_1 \delta} (R^{(1-\epsilon)\alpha} + R^{C\epsilon}) \frac{2^{k_1(j-2)(d-1)}}{2^{(d+3)}} , \quad \text{(7.1)}$$

whenever the sums are taken over $S_1$-tubes $T_j$ with $\text{dist}(T_j, Q_R) \lesssim R$, $||c_{T_j}||_{L^2_{t,x}} \lesssim 1$, and the wave packets are as described in Proposition 6.1. We only sum over $O(R^{\frac{d}{2}})$ $S_1$-tubes and $O(2^{k_1(j-2)} R^{\frac{d}{2}})$ $S_2$-tubes, so we may assume that for each $T_j$ in the sum, $|c_{T_j}| \gtrsim R^{-c_d} 2^{-k_1(j-2)} c_d$. This leaves $O(k_1 \log R)$ possible dyadic values for $c_{T_j}$ and by pigeonholing, it suffices to prove

$$\left\| \left( \sum_{T_1 \in T_j} \phi_{T_1} \right) \left( \sum_{T_2 \in T_5} \phi_{T_2}^A \right) \right\|_{L^\infty_{t,x}}(Q_R) \lesssim 2^{k_1 \delta} (R^{(1-\epsilon)\alpha} + R^{C\epsilon}) \frac{2^{k_1(j-2)(d-1)}}{2^{(d+3)}} (\#T_j \#T_2)^{\frac{1}{2}} , \quad \text{(7.2)}$$

whenever each $T_j$ is a collection of $S_j$-tubes $T_j$ with $\text{dist}(T_j, Q_R) \lesssim R$.

We decompose $Q_R = \bigcup_{B \in \mathcal{B}} B$, where $\mathcal{B}$ is a collection of finitely overlapping translates of $R^{-\epsilon} Q_R$. We also make a second, finer decomposition $Q_R = \bigcup_{q \in \mathcal{Q}} q$, where $\mathcal{Q}$ is a collection of finitely overlapping $R^{1/2}$ balls. For $q \in \mathcal{Q}$, define

$$T_j(q) = \{ T_j \in T_j : T_j \cap R^q \neq \emptyset \}.$$
Given dyadic values $1 \leq \mu_1, \mu_2, \lambda_1, \lambda_2 \leq 2^{k_1(J-2)R^{2(1+d)}}$, define

$$Q(\mu_1, \mu_2) = \{ q \in Q : \frac{1}{2} \mu_j \leq \# T_j(q) \leq \mu_j, j = 1, 2 \},$$

$$T_j(\lambda_j, \mu_1, \mu_2) = \{ T_j \in \mathcal{T}_j : \frac{1}{2} \lambda_j \leq \# \{ q \in Q(\mu_1, \mu_2) : T_j \in \mathcal{T}_j(q) \} \leq \lambda_j \} \quad \text{(7.3)}$$

$$B_j(T_j, \lambda_1, \mu_1, \mu_2) = \arg \max_{B \in \mathcal{B}} \# \{ q \in Q(\mu_1, \mu_2) : T_j \in \mathcal{T}_j(q) \text{ and } q \cap B \neq \emptyset \} \quad \text{(7.5)}$$

If $B \in \mathcal{B}$ and $T_j \in \mathcal{T}_j$, say $T_j \sim_{\lambda_j, \mu_1, \mu_2} B$ if $B \subseteq C B_j(T_j, \lambda_j, \mu_1, \mu_2)$ and say $T_j \sim B$ if $T_j \sim_{\lambda_j, \mu_1, \mu_2} B$ for some $\lambda_j, \mu_1, \mu_2$. (Here $C$ is sufficiently large for the proof of Lemmas 9.1 and 9.2 in Section 9.) Finally, given $B$, let $\mathcal{T}^{-}_j(B) = \{ T_j \in \mathcal{T}_j : T_j \sim B \}$, $\mathcal{T}^\perp_j(B) = \mathcal{T}_j \setminus \mathcal{T}^{-}_j(B)$.

By the triangle inequality,

$$\| \left( \sum_{T_{1,j} \in \mathcal{T}_1} \phi_{T_{1,j}} \right) \left( \sum_{T_{2,j} \in \mathcal{T}_2} \phi_{T_{2,j}} \right) \|_{\ell^\perp_{L^q(x)}} \leq \sum_{B \in \mathcal{B}} \left\| \left( \sum_{T_{1,j} \in \mathcal{T}_1^-(B)} \phi_{T_{1,j}} \right) \left( \sum_{T_{2,j} \in \mathcal{T}_2^-(B)} \phi_{T_{2,j}} \right) \right\|_{\ell^\perp_{L^q(x)}}(B)$$

$$\quad + \sum_{B \in \mathcal{B}} \left\| \left( \sum_{T_{1,j} \in \mathcal{T}_1^-(B)} \phi_{T_{1,j}} \right) \left( \sum_{T_{2,j} \in \mathcal{T}_2^+(B)} \phi_{T_{2,j}} \right) \right\|_{\ell^\perp_{L^q(x)}}(B)$$

$$\quad + \sum_{B \in \mathcal{B}} \left\| \left( \sum_{T_{1,j} \in \mathcal{T}_1^+(B)} \phi_{T_{1,j}} \right) \left( \sum_{T_{2,j} \in \mathcal{T}_2^+-\mathcal{T}_2^-(B)} \phi_{T_{2,j}} \right) \right\|_{\ell^\perp_{L^q(x)}}(B).$$

As in [24], we will think of the first as the “local term,” and the last three as “global.”

The local term may be bounded easily using the induction hypothesis and the fact that there are only $O(\log R)$ possible dyadic values of $\lambda_1, \lambda_2, \mu_1, \mu_2$:  

$$\sum_{B \in \mathcal{B}} \left\| \left( \sum_{T_{1,j} \in \mathcal{T}_1^-(B)} \phi_{T_{1,j}} \right) \left( \sum_{T_{2,j} \in \mathcal{T}_2^-(B)} \phi_{T_{2,j}} \right) \right\|_{\ell^\perp_{L^q(x)}}(B)$$

$$\quad \leq \sum_{B \in \mathcal{B}} 2^{k_1 \delta R^{1-\varepsilon} \alpha 2^{-k_1(J-2)(d-1) \frac{1}{2(\varepsilon-1)}} (\# T_1^- \# T_2^-)(B)^{\frac{1}{2}}}$$

$$\quad \leq 2^{k_1 \delta R^{1-\varepsilon} \alpha 2^{-k_1(J-2)(d-1) \frac{1}{2(\varepsilon-1)}} (\sum_{B \in \mathcal{B}} \# T_1^-)^{\frac{1}{2}} (\sum_{B \in \mathcal{B}} \# T_2^-)^{\frac{1}{2}}}$$

$$\quad = 2^{k_1 \delta R^{1-\varepsilon} \alpha 2^{-k_1(J-2)(d-1) \frac{1}{2(\varepsilon-1)}} (\sum_{T_{1,j} \in \mathcal{T}_1^-(B) : B \sim T_j} 1)^{\frac{1}{2}} (\sum_{T_{2,j} \in \mathcal{T}_2^-(B) : B \sim T_j} 1)^{\frac{1}{2}}}$$

$$\quad \leq 2^{k_1 \delta R^{1-\varepsilon} \alpha 2^{-k_1(J-2)(d-1) \frac{1}{2(\varepsilon-1)}} (\# T_1 \# T_2)^{\frac{1}{2}}}.$$

It remains to control the global terms.

8. Reduction to Two Combinatorial Estimates

It suffices to show that for each $B \in \mathcal{B}$,

$$\left\| \sum_{T_{1,j} \in \mathcal{T}_1^-} \sum_{T_{2,j} \in \mathcal{T}_2^-} \phi_{T_{1,j}} \phi_{T_{2,j}} \right\|_{\ell^\perp_{L^q(x)}} \leq 2^{k_1 \delta R C \alpha 2^{-k_1(J-2)(d-1) \frac{1}{2(\varepsilon-1)}} (\# T_1 \# T_2)^{\frac{1}{2}}}, \quad \text{(8.1)}$$

in each of the cases $\mathcal{T}_1^-(B) = \mathcal{T}_1^\perp(B)$ and $\mathcal{T}_2^-(B) \subseteq \mathcal{T}_2^\perp(B)$, $\mathcal{T}_1^-(B) \subseteq \mathcal{T}_1$ and $\mathcal{T}_2^-(B) = \mathcal{T}_2^\perp(B)$. The arguments for the different cases will only diverge in the proofs of the combinatorial estimates. For convenience, we will use the notation $\mathcal{T}_j^\perp(B)$ (with
Proof. We begin by estimating the contributions from the $S_j$-tubes separately in the cases $j = 1, 2$.

By Hölder’s inequality, (6.3), and (6.6),

$$
\| \sum_{T_1 \in T_1(B)} \phi_{T_1} \|_{L^2_{t,x}(B)} \lesssim 2^{k_1/j-2} \| R^{1/2} \sum_{T_1 \in T_1(B)} \phi_{T_1}(0) \|_{L^2_{t,x}(B)} \lesssim 2^{k_1/j-2} R^{1/2}(\#T_1)^{1/2}.
$$

(8.3)

Write $B = \bigcup_{j=0}^{2^{k_1/j-2}} B_j$, where each $B_j$ is an $R^{1/\varepsilon}$ cube, and for each $j$, let $T_j(B_j)$ denote the set of tubes $T_j \in T_j(B)$ for which dist$(T_j, B_j) \lesssim R^{-1/\varepsilon}$. Note that each tube is in $T_j(B_j)$ for $O(1)$ values of $j$ by transversality of $S_2$. Using the decay estimate (6.5), Hölder and the fact that $R \gtrsim 2^{k_1/j-2}$, (6.3), (6.6), and the near disjointness of the sets $T_j(B_j)$,

$$
\begin{align*}
\| \sum_{T_2 \in T_2(B)} \phi_{T_2} \|_{L^2_{t,x}(B)} & = \sum_{j=0}^{2^{k_1/j-2}} \| \sum_{T_2 \in T_2(B_j)} \phi_{T_2} + O(R^{-M}) \|_{L^2_{t,x}(B_j)}^2 \\
& \lesssim 1 + R \sum_{j=0}^{2^{k_1/j-2}} \| \sum_{T_2 \in T_2(B_j)} \phi_{T_1}(0) \|_{L^2_{t,x}(B_j)} \lesssim R \sum_{j=0}^{2^{k_1/j-2}} \#T_j(B_j) \lesssim R \#T_j(B).
\end{align*}
$$

(8.4)

Finally, (8.2) just follows from (8.3), (8.4), and Cauchy–Schwarz. □

By interpolation, (8.1) will then follow from the estimate

$$
\| \sum_{T_j(B)} \phi_{T_1} \|_{L^2_{t,x}(B)} \lesssim 2^{k_1/j} R^{d+1/2} \#T_j \#T_j^{1/2}.
$$

(8.5)

We decompose

$$
\| \sum_{T_j(B)} \phi_{T_1} \phi_{T_2} \|_{L^2_{t,x}(B)} \leq \sum_{q \in 2B} \| \sum_{T_j(B)} \phi_{T_1} \phi_{T_2} \|_{L^2_{t,x}(q)}.
$$

(8.6)

By the decay estimate, if $T_j \notin T_j(q)$ (i.e. $T_j \cap R^q = \emptyset$), $|\phi_{T_j}| \lesssim R^{-M}$ on $q$, for arbitrarily large $M$, so the contribution from any tubes not in $T_j(q)$ is negligible.

By this and pigeonholing, it suffices to prove that

$$
\sum_{q \in 2B} \| \sum_{T_j(q)} \phi_{T_1} \phi_{T_2} \|_{L^2_{t,x}(q)} \lesssim R^{C\varepsilon - \frac{d+1}{2}} \#T_j \#T_j,
$$

(8.7)

where

$$
T_j'(q) = T_j'(B) \cap T_j(q) \cap T_j(\lambda_j, \mu_1, \mu_2),
$$

(8.8)
and $1 \leq \mu_1, \mu_2, \lambda_1, \lambda_2 \lesssim R^{100d}$ are arbitrary dyadic values, which will remain fixed for the remainder of the section.

Given $\xi_1 \in B(0, 2^{-k_1+1}c_0)$ and $\xi_2' \in B(0, 2c_0)$, or $\xi_1' \in B(0, 2^{-k_1+1}c_0)$ and $\xi_2 \in B(0, 2c_0)$ (respectively), the functions

\[
\begin{align*}
\xi_1' &\mapsto (h_1(\xi_1') + h_2(\xi_1' + \xi_2 - \xi_1)) - (h_1(\xi_1') + h_2(\xi_2')) \\
\xi_2' &\mapsto (h_1(\xi_1' + \xi_2 - \xi_2') + h_2(\xi_2')) - (h_1(\xi_1') + h_2(\xi_2'))
\end{align*}
\]

have gradients comparable to 1, so the hypersurfaces

\[
\begin{align*}
\pi_1(\xi_1, \xi_2') &= \{\xi_1' \in B(0, 2c_0) : h_1(\xi_1) + h_2(\xi_1' + \xi_2 - \xi_1) = h_1(\xi_1') + h_2(\xi_2')\}, \\
\pi_2(\xi_2, \xi_1') &= \{\xi_2' \in B(0, 2c_0) : h_1(\xi_1' + \xi_2 - \xi_2') + h_2(\xi_2) = h_1(\xi_1') + h_2(\xi_2')\},
\end{align*}
\]

are smoothly embedded.

Given $\xi_1, \xi_1' \in \Xi_1$ and $\xi_2, \xi_2' \in \Xi_2$, define collections

\[
\begin{align*}
T_1'(q, \xi_1, \xi_2') &= \{T_1' \in T_1' (q) : \text{dist}(\xi(T_1'), \pi_1(\xi_1, \xi_2')) \lesssim R^{C\varepsilon^{-1/2}}\}, \\
T_2'(q, \xi_2, \xi_1') &= \{T_2' \in T_2'(q) : \text{dist}(\xi(T_2'), \pi_1(\xi_2, \xi_1')) \lesssim R^{C\varepsilon^{-1/2}}\},
\end{align*}
\]

and quantities

\[
\begin{align*}
\nu_1(q) &= \sup_{\xi_1 \in \Xi_1, \xi_2' \in \Xi_2} \#T_1'(q, \xi_1, \xi_2'), \\
\nu_2(q) &= \sup_{\xi_2 \in \Xi_2, \xi_1' \in \Xi_1} \#T_2'(q, \xi_2, \xi_1').
\end{align*}
\]

Lemma 8.2. For any $q \in Q(\mu_1, \mu_2)$, and $j = 1, 2$,

\[
\| \sum_{T_1'(q)} \sum_{T_2'(q)} \phi_{T_1', T_2'} \|^2_{L^2_{T_1', T_2'}(q)} \lesssim R^{C\varepsilon} R^{-(d-1)/2} \nu_j(q) \#T_j(q) \#T_j'(q) \tag{8.12}
\]

Proof. We give the proof when $j = 1$. By simple arithmetic,

\[
\begin{align*}
\| \sum_{T_1'(q)} \sum_{T_2'(q)} \phi_{T_1', T_2'} \|^2_{L^2_{T_1', T_2'}(q)} &\leq \sum_{T_1, T_1' \in T_1'(q)} \sum_{T_2, T_2' \in T_2'(q)} \langle \phi_{T_1', T_2}, \phi_{T_1, T_2'} \rangle.
\end{align*}
\]

By Plancherel, $\langle \phi_{T_1', T_2}, \phi_{T_1, T_2'} \rangle$ equals zero unless

\[
\begin{align*}
\xi(T_1) + \xi(T_2) &= \xi(T_1') + \xi(T_2') + O(R^{-1/2}) \\
(h_1(\xi(T_1)) + h_2(\xi(T_2))) &= h_1(\xi(T_1')) + h_2(\xi(T_2')) + O(R^{-1/2}),
\end{align*}
\]

i.e. unless \(\text{dist}(T_1', \pi_1(\xi_1(T_1), \xi_2(T_2))), \text{dist}(T_2', \pi_2(\xi_2(T_2), \xi_1(T_1'))) \lesssim R^{-1/2}\).

By Plancherel, a simple change of variables using transversality of the surfaces $S_1, S_2$, and the small frequency support of the $\phi_j(0)$,

\[
\| \phi_{T_1', T_2'} \|_{L^2_{T_1', T_2'}} = \| \mathcal{F}(\phi_{T_1'}) * \mathcal{F}(\phi_{T_2'}) \|_{L^2_{T_1', T_2'}} \lesssim R^{-(d-1)/4} \| \phi_{T_1(0)} \|_{L^2} \| \phi_{T_2(0)} \|_{L^2} \sim R^{-(d-1)/4}.
\]

We claim that, given $T_1, T_1', T_2, T_2'$, (8.13) can hold for at most $O(R^{C\varepsilon})$ tubes $T_2'$ in $T_2'(q)$. Indeed, the second map in (8.9) has gradient comparable to 1, so the equations (8.13) essentially determine $\xi_2(T_2')$; when combined with $q$, this direction determines $T_2'$.

Putting these observations together,

\[
\begin{align*}
\| \sum_{T_1'(q)} \sum_{T_2'(q)} \phi_{T_1', T_2'} \|^2_{L^2_{T_1', T_2'}(q)} &\lesssim \sum_{T_1 \in T_1'(q)} \sum_{T_2' \in T_2'(q)} \sum_{T_1 \in T_1'(q)} \sum_{T_2 \in T_2'(q)} R^{C\varepsilon} \| \phi_{T_1(0)} \|_{L^2} \| \phi_{T_2(0)} \|_{L^2} \lesssim R^{C\varepsilon} \#T_1'(q) \#T_2'(q) \nu_1(q).
\end{align*}
\]
Thus it suffices to show that the left side of this inequality is bounded ($\lesssim$).

\[ \| \sum_{T'_1(q)}^{T_1(q)} \sum_{k'} \phi_{T_1} \phi_{T_2} \|^2 \lesssim R^{C_\varepsilon-d/2} \# T'_1(q) \not\lesssim T'_2(q) \nu_2(q) \]

is exactly the same. \hfill \Box

It remains to control the sum on $q$ of the right side of (8.12). We will show that if $T'_1 = T'_1(q)$, then

\[ \sum_{Q(\mu,B)} \# T'_1(q) \# T'_2(q) \nu_1(q) \lesssim R^{C_\varepsilon} \# T'_1 \# T'_2 \quad (8.14) \]

and that if $T'_2 = T'_2(q)$, then

\[ \sum_{Q(\mu,B)} \# T'_1(q) \# T'_2(q) \nu_1(q) \lesssim R^{C_\varepsilon} \# T'_1 \# T'_2, \quad (8.15) \]

where $T'_j(q)$ is as in (8.8) and $Q(\mu, B) = \{ q \in Q(\mu_1, \mu_2) : q \subseteq 2B \}$. These are our combinatorial estimates.

9. Proofs of the combinatorial estimates

This section will be devoted to the proofs of the combinatorial estimates (8.14) and (8.15). There are some differences in the proofs due to the differing geometries of the intersections of $S_j$'s with $Q_R$ for $j = 1, 2$, but the two inequalities are more similar than not. We will begin with (8.14) and indicate the changes necessary for (8.15). The argument is adapted from that of [24], so we will be somewhat brief.

Recalling the role played by $\mu$ from (7.3), using Fubini, and then recalling the role of $\lambda$ from (7.4) and (8.8),

\[ \sum_{q \in Q(\mu, B)} \# T'_1(q) \# T'_2(q) \nu_1(q) \lesssim \mu_2 \nu_1 \sum_{\tilde{T}_1 \in T_1(q)} \# \{ q \in Q(\mu, B) : T_1 \subseteq T_1(B) \} \lesssim \mu_2 \nu_1 \nu_1 \# T'_1(q). \]

Thus (8.14) will be proven if we can show that for an arbitrary (henceforth fixed) $q_0 \in Q(\mu, B)$ and arbitrary (also fixed) $\xi_1 \in \Xi_1, \xi'_1 \in \Xi_2$, \begin{equation}
\# T'_1(q_0, \xi_1, \xi'_1) \lesssim 2^{k_1 \varepsilon} R^{C \varepsilon} \# T_2 \mu_2 \lambda_1. \quad (9.1)\end{equation}

If $T_1 \in T'_1(q_0, \xi_1, \xi'_1)$, $B \not\subseteq CB_1(T_1, \lambda_1, \mu_1, \mu_2)$, so

\[ \# \{ q \in Q(\mu) : T_1 \cap R^q \neq \emptyset, \ q \cap \frac{1}{2}CB = \emptyset \} \gtrsim R^{-C \varepsilon} \lambda_1. \]

Furthermore, if $q \in Q(\mu)$, $\# \{ T_2 \in T_2 : T_2 \cap R^q \neq \emptyset \} \gtrsim \mu_2$, so

\[ \# \{ (q, T_1, T_2) \in Q(\mu) \times T'_1(q_0, \xi_1, \xi'_1) \times T_2 : T_1 \cap R^q \neq \emptyset, \ T_2 \cap R^q \neq \emptyset, \ q \cap \frac{1}{2}CB = \emptyset \} \gtrsim R^{-C \varepsilon} \lambda_1 \mu_2 \# T'_1(q_0, \xi_1, \xi'_1). \]

Thus it suffices to show that the left side of this inequality is bounded ($\lesssim$) by $2^{k_1 \varepsilon} R^{C \varepsilon} \# T_2$. This will follow from the next lemma.

Lemma 9.1. If $T_2 \in T_2$,

\[ \# \{ (q, T_1) \in Q \times T'_1(q_0, \xi_1, \xi'_1) : T_1 \cap R^q \neq \emptyset, \ T_2 \cap R^q \neq \emptyset, \ q \cap \frac{1}{2}CB = \emptyset \} \lesssim 2^{k_1 \varepsilon} R^{C \varepsilon}. \]
Proof. Let \((t_0, x_0)\) and \((t, x)\) denote the centers of \(q_0\) and \(q\), respectively. Suppose that the pair \((q, T_1)\) is in the set above. Since \(T_1 \cap R^c q_0, T_1 \cap R^c q \neq \emptyset\),

\[ x - x_0 = (t - t_0)v_1(T_1) + O(R^{1/2+\varepsilon}), \]

which implies that \(|x - x_0| \lesssim 2^{-k_1(J-2)}|t - t_0| + O(R^{1/2+\varepsilon}).\) On the other hand, \(q_0 \leq 2B\) and \(q \cap \frac{C}{2}B = \emptyset\) together imply that \(|t - t_0| \gtrsim 2^{k_1(J-2)}R^{1-\varepsilon}\) or \(|x - x_0| \gtrsim R^{1-\varepsilon}\); by the preceding observation, the former must hold. This implies two things.

First, \((t, x)\) must lie within \(O(R^{1/2+\varepsilon})\) of the hypersurface \(\Gamma + (t_0, x_0)\), where

\[ \Gamma = \Gamma (\xi_1, \xi'_2) = \{(t, x) : t \gtrsim 2^{k_1(J-2)}R^{1-\varepsilon}, x = t\nabla h_1(\xi_1)\text{ for some } \xi'_1 \in \pi_1(\xi_1, \xi'_2)\}. \]

We will show that \(\Gamma\) is transverse to directions in \(V_2\). Assuming this for a moment, our tube \(T_2\) intersects \(\Gamma\) in a ball of radius \(R^{1/2}\) and thus picks out \(O(R^{C\varepsilon})\) cubes \(q\).

Second, \(v_1(T_1) = \frac{x-x_0}{t-t_0} + O(2^{-k_1(J-2)}R^{-1/2+C\varepsilon})\), so given \(q\), there are at most \(O(R^{C\varepsilon})\) possible choices for \(T_1\).

The proof of the lemma will be complete once we verify the transversality. By ellipticity of \(g_1\), \(\nabla h_1\) is an invertible function. Unwinding the definitions,

\[ \Gamma = \{(t, x) : h_2(\xi_1 + (\nabla h_1)^{-1}(\xi) - \xi'_2) - h_1((\nabla h_1)^{-1}(\pi)) = h_1(\xi_1) - h_2(\xi'_2)\}. \]

Thus (undoing the scalings), the normal at \((t, x)\) is parallel to

\[
\begin{align*}
-2^{k_1(J-1)}\nabla g_2(\eta_2)(D^2 g_1(\eta')^{-1})\nabla g_1(\eta') + 2^{-2k_1(J-2)}\nabla g_1(\eta'_1)(D^2 g_1(\eta')^{-1})\nabla g_1(\eta'_1),
\end{align*}
\]

where \(\xi = \nabla h_1(\xi'_1) = 2^{-k_1(J-1)}\nabla g_1(\eta'_1)\) and \(\eta' = \xi = \xi_1 + \xi'_1 - \xi'_2\) and \(|\eta_1|, |\eta_2| < c_0\). Recalling that \(D^2 g_1\) is close to the identity and \(\nabla g_2\) is close to \(e_1\), we see that this normal makes a large angle with any \((1, -v_2(T_2))\), so we have the transversality we want.

This completes the proof of (8.14). Now we turn to (8.15). Simply changing subscripts in the earlier argument, we can reduce matters to proving the following.

**Lemma 9.2.** If \(T_1 \in T_1\),

\[ \# \{(q, T_2) \in Q \times T_2^c(q_0, \xi_1, \xi'_2) : T_2 \cap R^c q \neq \emptyset, T_1 \cap R^c q \neq \emptyset, q \cap \frac{C}{2}B = \emptyset\} \lesssim 2^{k_1\delta} R^{C\varepsilon}. \]

**Proof.** As before, let \((t_0, x_0), (t, x)\) denote the centers of \(q_0, q\). This time, if \((q, T_2)\) is in the above set, \(T_2 \cap R^c q_0, T_2 \cap R^c q \neq \emptyset\), which implies that \(|t - t_0| \lesssim R\). Thus since \(q_0 \leq 2B\) and \(q \cap \frac{C}{2}B = \emptyset\), \(|x - x_0| \gtrsim R^{1-\varepsilon}\). Now \(x - x_0 = (t - t_0)v_2(T_2) + O(R^{1/2+\varepsilon})\), and since \(|v_2(T_2)| \lesssim 1, |t - t_0| \gtrsim R^{1-\varepsilon}\) as well.

Now we know that \((t, x)\) must lie within \(O(R^{1/2+\varepsilon})\) of the hypersurface

\[ \{(t, x) : |t - t_0| \gtrsim R^{1-\varepsilon}, (x - x_0) = (t - t_0)v_2(T_2)\text{, for some } \xi'_2 \in \pi_2(\xi_2, \xi'_1)\}. \]

It is similar (but slightly simpler) to show that this hypersurface is transverse to directions in \(V_1\) (such directions are nearly vertical), so \(T_1\) intersects it in a ball of radius \(R^{1/2}\), picking out \(O(R^{C\varepsilon})\) cubes \(q\).

Between the estimate \(v_2(T_2) = \frac{t-t_0}{t-t_0} + O(R^{-1/2+C\varepsilon})\) and the fact that \(T_2\) intersects \(R^c q\), there are only \(O(R^{C\varepsilon})\) possibilities for \(T_2\) as well, so we are done. \(\square\)
The same argument gives bounds for restriction to the graph of $a_1|\xi|^{k_1} + \cdots + a_n|\xi|^{k_n}$, for any coefficients $a_1, \ldots, a_n > 0$ and real powers $2 \leq k_1 < \cdots < k_n$; the coefficients however will depend on the $k_i$, not just on $k_n$.

Let $P(t) = a_2 t^2 + \cdots + a_n t^n$, and assume that $P''(t) > 0$ for all $t > 0$. Let $n_{\min}$ and $n_{\max}$ be the degrees of the lowest and highest (respectively) terms of $P$; their coefficients, $a_{n_{\min}}$ and $a_{n_{\max}}$, must be positive. Let $I_{\min} = \{ t \geq 0 : a_{n_{\min}} t^{n_{\min}} \geq \max_i |a_i t^i| \}$, $I_{\max} = \{ t \geq 0 : a_{n_{\max}} t^{n_{\max}} \geq \max_i |a_i t^i| \}$, and $I_{\text{med}} = [0, \infty) \setminus (I_{\min} \cup I_{\max})$. Then $I_{\min}$ contains all points sufficiently small and $I_{\max}$ all points sufficiently large, so $\{ (P(|\xi|), |\xi|) : |\xi| \in I_{\text{med}} \}$ is compact and elliptic. The methods of the preceding sections apply on $\{ (P(|\xi|), |\xi|) : |\xi| \in I_{\bullet} \}$ for $\bullet = \min, \max$, and we can obtain a non-uniform version of Theorem 1.1 for restriction to the graph of $P(|\xi|)$. Arguing similarly (but only separating out the low frequencies), we may prove such a nonuniform theorem for hypersurfaces of the form

$$\{(\phi(|\xi|), |\xi|) : |\xi| \leq R \},$$

whenever $\phi$ is smooth, $\phi'(0) = 0$, $\phi$ is finite type at 0, and $\phi''(t) > 0$ for $t > 0$. It would be nice to know more uniform versions of these results.

As a corollary of Theorem 1.1, we can obtain an unweighted result, which is necessarily nonuniform.

**Corollary 10.1.** Let $P$ be a polynomial on $\mathbb{R}$ with $P'(0) = 0$ and $P''(t) > 0$ for all $t > 0$. Let $n_{\min}$ denote the lowest nonzero power of $t$ appearing in $P$ and $n_{\max}$ the greatest. Then, conditional on the restriction conjecture $\mathcal{R}(p_0 \to q_0)$ for the admissible pair $(p_0, q_0)$,

$$\| \hat{f}(P(|\xi|), \xi) \|_{L^r(\xi)} \lesssim \| f \|_p,$$  \hspace{1cm} (10.1)

provided $1 \leq p < p_0$ and either $r \geq p$ and $\frac{dp'}{n_{\max} + d} \leq r \leq \frac{dp'}{n_{\min} + d}$, or $r < p$ and $\frac{dp'}{n_{\max} + d} < r < \frac{dp'}{n_{\min} + d}$. The implicit constant depends on $d, p, P$.

For a given value of $p$ the range of $r$ in the corollary is sharp. In particular, the full conjectured range of unweighted bounds would follow from a resolution of the restriction conjecture. We note that in certain cases, some of the exponents $r$ covered in the corollary may be less than 1.

The proof of the corollary uses an argument dating back at least to Drury–Marshall in [12] and some simple observations.

**Proof.** We give the proof when $n_{\min} < n_{\max}$. In the monomial case $n_{\min} = n_{\max}$, the argument is similar but simpler.

By Theorem 1.1 (or the extension mentioned above),

$$\| \hat{f}(P(|\xi|), \xi) \Lambda_P(\xi)^{-\frac{dp'}{d + 2p'}} \|_{L^r} \lesssim \| f \|_p, \hspace{1cm} 1 \leq p < p_0, \hspace{0.5cm} q = \frac{dp'}{d + 2p'}.$$  \hspace{1cm} (10.2)

Let $I_{\min}, I_{\text{med}}, I_{\max}$ be the intervals defined just before the statement of the corollary, and let $A_{\bullet} = I_{\bullet}$ for $\bullet = \min, \text{med}, \max$.

Since $|A_{\text{med}}| < \infty$ and $\Lambda_P(\xi) \sim 1$ on $A_{\text{med}},$

$$\| \hat{f}(P(|\xi|), \xi) \|_{L^r(A_{\text{med}})} \lesssim \| f \|_p, \hspace{1cm} 0 < r \leq \frac{dp'}{d + 2p'}.$$  \hspace{1cm}

This includes the range in the corollary, so it suffices to control the low and high frequency parts.
For \( \xi \in A_\ast \), \( \Lambda_P(\xi) \sim |\xi|^{\frac{(n-2)d}{p'q}} \). Thus by (10.2) and the Lorentz space version of Hölder’s inequality ([22]),
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,q}(A_\ast)} \lesssim \| f \|_p, \quad 0 < r \leq q = \frac{dp'}{d+2}, \quad 1 \leq p < p_0.
\]
Performing Marcinkiewicz interpolation along segments with \( \frac{2d+q}{r} = \frac{d}{r} \) equal to a constant,
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,q}(A_\ast)} \lesssim \| f \|_p, \quad 0 < r < \frac{dp'}{d+2}, \quad 1 < p < p_0. \tag{10.3}
\]
Now we turn to the low frequency part. By (10.3),
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_{\min})} \lesssim \| f \|_p, \quad r = \frac{dp'}{n_{\min}+d}.
\]
When \( r \geq p \), the left side bounds the \( L^r(A_{\min}) \) norm, which in turn bounds the \( L^\nu(A_{\min}) \) norm for all \( s \leq r \), since \( |A_{\min}| < \infty \). If \( r < p \), we set \( A_k = \{|\xi| \sim 2^k\} \) and let \( q = \frac{dp'}{d+2} \). Then \( q > r \), so by Hölder’s inequality and (10.2),
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_{\min})} \lesssim 2^k \| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_k \cap A_\nu)} \lesssim 2^k \| f \|_p. \tag{10.4}
\]
For \( r < \frac{dp'}{n_{\min}+d} \), we can sum over those \( k \) such that \( A_k \cap A_{\min} \neq \emptyset \), obtaining
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_{\min})} \lesssim \| f \|_p.
\]
Now we turn to the high frequency terms. Since \( |\xi| \gtrsim 1 \) on \( A_{\max} \), (10.3) implies that
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_{\max})} \lesssim \| f \|_p, \quad \frac{dp'}{n_{\max}+d} \leq r \leq \frac{dp'}{d+2}, \quad 1 \leq p < p_0.
\]
If \( r \geq p \), the left side of this inequality bounds the \( L^r \) norm and we are done. If \( r < p \), we argue exactly as in (10.4) to obtain
\[
\| \hat{f}(P(|\xi|), \xi) \|_{L^{r,p}(A_k)} \lesssim 2^k \| f \|_p,
\]
which is summable over large \( k \) for \( r > \frac{dp'}{n_{\max}+q} \). \( \square \)

The sharpness of the corollary in the case \( r \geq p \) is known, and for \( r < p \), it has a similar proof to the analogous result in [23].

We close with the essentially trivial deduction of uniform local estimates from elliptic restriction theorems off the scaling line. Our motivations are two-fold. First, this allows us to obtain bounds in the Bourgain–Guth range ([6]). Second, in the negatively curved case, no scaling-critical estimates are known beyond Stein–Tomas ([27, 16]), so these arguments may be helpful in a consideration of more general hypersurfaces.

**Proposition 10.2.** Assume that \( \mathcal{R}^*(p \to q) \) holds for some \( q \) greater than the maximum of \( \frac{2(d+1)p}{d} \) and \( \frac{2(d+2)}{d} \). Then for all bounded sets \( K \subseteq \mathbb{R}^d \) and even polynomials \( P \) with non-negative coefficients,
\[
\| \Lambda_P(\nabla)^{1/p'} \mathcal{E}_P(\chi_K f) \|_{L^p} \lesssim \| \xi \|^\alpha \| f \|_{L^p}, \quad p' := \frac{dq}{d+2}, \quad \alpha < \frac{d}{p} - \frac{d}{p'}.
\]
The implicit constant depends on \( K \), \( \alpha \), and the degree of \( P \).
Proof. We may assume that $K = B(0, R)$ for some $R > 0$. Choose intervals $J_j$ as in Section 3: so that $P(t) \sim a_j t^{2j}$ on $J_j$. It suffices to prove uniform estimates over each annulus $A_j := \{ \xi \in K : \| \xi \| \in J_j \}$. For $k \in \mathbb{Z}$, $2^{-k} \leq 2R$, let

$$A_{jk} = \{ \xi \in A_j : 2^{-k-1} \leq \| \xi \| \leq 2^{-k} \}.$$

Rescaling $\mathcal{R}^*(p \to q)$,

$$\| A_P(\nabla)^{1/p^*} \mathcal{E}_P(\chi_{A_{jk}} f) \|_q \lesssim 2^{-kd/\hat{p}} \frac{(d+2)k}{q} \frac{dk}{p} \| f \chi_{A_{jk}} \|_p$$

$$\sim 2^{k(\alpha - (\frac{d}{p} - \frac{d}{\hat{p}}))} \| \xi \| \| \chi_{A_{jk}} f \|_{L_p}.$$

The right side is clearly summable, with bounds depending on $\alpha, R$. \qed

There is the question of the endpoint $\alpha = \frac{d}{p} - \frac{d}{\hat{p}}$. When $q \geq p$, it is possible to deduce, using the methods of this article, conditional results, but the exponents are typically worse than those in Proposition 10.2. If $q < p$, the endpoint is false. This can be seen by considering functions of the form $f = \sum 2^{kd/\hat{p}} e^{ix \cdot \xi} f_k$, with the $x_k$ sufficiently widely separated, $\text{supp } f_k \subseteq \{ 2^{-k-1} < \| \xi \| < 2^{-k} \}$, and the $f_k$ quasi-extremal in the sense that

$$\| A_P(\nabla)^{1/p^*} \mathcal{E}_P f_k \|_q \gtrsim \| \nabla \|^{\alpha} f_k \|_p \sim 2^{-kd/\hat{p}}.$$

By way of comparison, a scaling critical adjoint restriction theorem for elliptic hypersurfaces, $\mathcal{R}^*(p_0 \to q_0)$, would imply (by Hölder and Theorem 1.1) that for any compact $K \subseteq \mathbb{R}^d$, $q > q_0$, $p \geq \hat{p} := (\frac{d+2}{d})q_0$, and $\alpha < \frac{d}{p} - \frac{d}{\hat{p}}$,

$$\| A_P(\nabla)^{1/p} \mathcal{E}_P f \|_q \lesssim \| \xi \| \| \chi \| \| f \|_p, \quad f \in L^q(\mathbb{R}^d) \quad \text{supp } f \subseteq K. \quad (10.5)$$

If we instead use the Lorentz space version of Hölder’s inequality and argue as in the proof of Corollary 10.1, we would have (10.5) for all $q > q_0$, $q \geq p \geq \hat{p}$ and $\alpha \leq \frac{d}{p} - \frac{d}{\hat{p}}$. In both cases, the implicit constants in (10.5) depend on $q, p, K, \alpha, q_0$, and the degree of $P$.

References


