BILINEAR RESTRICTION ESTIMATES FOR SURFACES WITH CURVATURES OF DIFFERENT SIGNS

SANGHYUK LEE

ABSTRACT. Recently, the sharp $L^2$-bilinear (adjoint) restriction estimates for the cone and the paraboloid were established by Wolff and Tao, respectively. Their results rely on the fact that for the cone and the paraboloid, the nonzero principal curvatures have the same sign. We generalize those bilinear restriction estimates to surfaces with curvatures of different signs.

1. Introduction and the statement of results

Let $S$ be a smooth compact hypersurface with boundary in $\mathbb{R}^{n+1}$, $n \geq 2$, with Lebesgue measure $d\sigma$. The Fourier restriction problem for $S$ is to determine $(p, q)$ for which the (linear) adjoint restriction estimate

$$\|\hat{f}d\sigma\|_q \leq C\|f\|_{L^p(d\sigma)}$$

holds for all $f \in C^\infty_c(S)$. Although many works were devoted to this problem, it remains widely open. However, some significant progress has been recently made in the restriction estimate for the cone [21], the paraboloid and the sphere [13]. These results were obtained by studying a bilinear version of their (adjoint) restriction operators. In this note we aim to generalize the known bilinear restriction estimates ([13, 21]) to more general surfaces.

Let $\phi_1, \phi_2$ be smooth functions on $[-1, 1]^n$. For $i = 1, 2$, let us define extension (adjoint of restriction) operators by

$$E_i f(x, t) = \int_{V_i} e^{2\pi i (x \cdot \xi - t\phi_i(\xi))} f(\xi) d\xi, \quad (x, t) \in \mathbb{R}^n \times \mathbb{R},$$

where $V_1, V_2$ are subcubes in $[-1, 1]^n$. Possibly, $V_1, V_2$ can be the same sets. The bilinear approach to the restriction problem is an attempt to obtain the estimate of the form

$$\|E_1 f \cdot E_2 g\|_{L^q(\mathbb{R}^{n+1})} \leq C\|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^p(\mathbb{R}^n)}.$$

Obviously, if $\phi_1 = \phi_2$ and $V_1 = V_2$, it is equivalent to the linear adjoint restriction estimate. The advantage of the bilinear estimate is that if one imposes some additional conditions (e.g. transversality) on the surfaces $S_i = \{(\xi, -\phi_i(\xi)) : \xi \in V_i\}$, $i = 1, 2$, then the bilinear restriction estimate may have the wider range $(p, q)$ of boundedness than that of the linear one. As shown in [17], [21], for some specific
surfaces (e.g. the sphere or the cone) it is possible to deduce the corresponding linear estimate from bilinear one. So, the bilinear restriction estimate (1.2) can be thought of as a generalization of the linear restriction estimate. For more and related materials, we refer the reader to [13] [14] [15] [16] [17].

The problem is relatively easier on $L^2$, where Plancherel’s theorem can be used freely. Letting $p = 2$, one may try to obtain the best possible $q$ for which

\[ \| E_1 f \cdot E_2 g \|_{L^q(\mathbb{R}^{n+1})} \leq C \| f \|_{L^2(\mathbb{R}^n)} \| g \|_{L^2(\mathbb{R}^n)} \]

holds. This type of estimate was used not only for the restriction problem but it also has applications for a variety of related problems (see [14] [16] and references therein, and also [4] [7] [8] [9]). The estimate (1.3) was first formulated by Bourgain [4] with $\phi_1 = \phi_2 = |\xi|$, $n = 2$ and separation condition (1.4), and he showed it for some $p > 2 - \epsilon$ for some $\epsilon > 0$. Klainerman and Machedon conjectured that when $\phi_1 = \phi_2 = |\xi|$ or $\phi_1 = \phi_2 = |\xi|^2$, (1.3) holds for $q \geq \frac{n+3}{n+1}$ under the condition

\[ \text{dist}(\nabla \phi_1(V_1), \nabla \phi_2(V_2)) \sim 1. \]

Here $A \sim B$ ($A, B > 0$) means $C^{-1} A \leq B \leq CA$ for some $C > 0$. In [13] [17], a systematic study on this problem was carried out and some partial results were obtained. For $\phi_1 = \phi_2 = |\xi|$, the conjecture was later settled by Wolff [21] ($q > \frac{n+3}{n+1}$) and Tao [12] ($q = \frac{n+3}{n+1}$). Recently, Tao [13] obtained the sharp estimate for $\phi_1 = \phi_2 = |\xi|^2$ except for the endpoint estimate ($q = \frac{n+3}{n+1}$). His result also includes the positively curved surfaces, namely, the elliptic surfaces all of which principal curvatures are positive (see [17]).

Both of the results in [13] [21] rely on the fact that for the cone and the paraboloid, the nonzero principal curvatures have the same sign. We try to generalize the bilinear restriction estimates for the cone and the paraboloid to surfaces with curvatures of different signs. The possibility of this kind of generalization was already indicated in [13]. Let us denote by $H\phi$ the Hessian matrix of $\phi$. Our first result is the following.

**Theorem 1.1.** Suppose for $\xi \in [-1, 1]^n$,

\[ \det H\phi_i(\xi) \neq 0, \quad i = 1, 2. \]

Additionally, suppose for all $\xi, \xi', \xi'' \in V_1$ and $\zeta, \zeta', \zeta'' \in V_2$,

\[ |(H\phi_1)^{-1}(\xi'')(\nabla \phi_1(\xi) - \nabla \phi_2(\zeta), \nabla \phi_1(\xi') - \nabla \phi_2(\zeta'))| \geq c > 0, \]

(1.6)

\[ |(H\phi_2)^{-1}(\xi'')(\nabla \phi_1(\xi) - \nabla \phi_2(\zeta), \nabla \phi_1(\xi') - \nabla \phi_2(\zeta'))| \geq c > 0. \]

Then (1.3) holds for $q > \frac{n+3}{n+1}$.

One can see (1.3) is no longer valid for $q < \frac{n+3}{n+1}$ by using the squashed cap functions in [17]. If $\phi_1$ and $\phi_2$ are elliptic functions, condition (1.4) implies (1.6) and (1.7) provided $V_1$ and $V_2$ are sufficiently small since the eigenvalues of $H\phi_i$ have the same sign. So, Theorem 1.1 contains Tao’s result for paraboloid and elliptic surfaces.

The conditions (1.6), (1.7) are related to the rotational curvature conditions of the functions $\Phi^1_2$, $\Phi^2_2$ given by

\[ \Phi^1_2(y, w) = \phi_1(x) + \phi_2(y) - \phi_1(x + y - w) - \phi_2(w), \]

\[ \Phi^2_2(x, z) = \phi_1(x) + \phi_2(z + w - x) - \phi_1(z) - \phi_2(w). \]
Since we are assuming $\det H\phi_i(\xi) \neq 0$, (1.6) and (1.7) are equivalent to
\[
\det \begin{pmatrix}
0 & \partial_y \Phi^T(y, w) \\
\partial_y \Phi^T(y, w) & 0
\end{pmatrix} \neq 0, \quad \det \begin{pmatrix}
0 & \partial_z \Phi^T(x, z) \\
\partial_z \Phi^T(x, z) & 0
\end{pmatrix} \neq 0,
\]
respectively, if $x \in V_1, y \in V_2, z \in V_1, w \in V_2, x + y - w \in V_1$ and $z + w - x \in V_2$. It is easy to see, using the fact that for a nonsingular $n \times n$ matrix $M$ and $v, u \in \mathbb{R}^n$,
\[
\det \begin{pmatrix}
0 & u^T \\
v & M
\end{pmatrix} = -(\det M)(M^{-1}v, u).
\]
Hence, $\Phi^T_1, \Phi^T_2$ satisfy the rotational curvature condition. Therefore, following the same lines of argument in [15, 17], one can see
\[
\|E_1 f \cdot E_2 g\|_2 \leq C\|f\|_{4(n+1)/n} \|g\|_{4(n+1)/n}
\]
which is corresponding to Theorem 2.3 of [17] where elliptic $\phi_1, \phi_2$ were considered.

When the eigenvalues of $H\phi_1$ have different signs, to get (1.8) for $q \geq \frac{n+3}{n+1}$, it is insufficient to impose (1.3) simply. If (1.6) and (1.7) are replaced by the weaker (1.3), then (1.3) fails for $\frac{n+2}{n} > q$. More precisely, under condition (1.3), (1.2) is valid only for $p, q$ satisfying
\[
n \left(1 - \frac{1}{p}\right) \geq \frac{n+2}{2q}.
\]
This is also the necessary condition for the (linear) adjoint restriction $L^p - L^{2q}$ estimate for the surfaces with nonvanishing Gaussian curvature in $\mathbb{R}^{n+1}$. Under the separation condition (1.4) (dist $(V_1, V_2) \sim 1$), the bilinear restriction estimate for hyperboloid has no better boundedness than the linear estimate.

For $n \geq 2$, we consider
\[
\phi_1(\xi) = \phi_2(\xi) = (\xi_1^2 + \cdots + \xi_{k-1}^2 - \xi_k^2 - \cdots - \xi_{n-2}^2 + 2\xi_{n-1}\xi_n) = \frac{1}{2}(M\xi, \xi),
\]
where $M$ is the symmetric matrix which makes the last equality hold. Note that $H\phi_1$ and $H\phi_2$ have $k$ positive and $n - k$ negative eigenvalues. In this case the conditions (1.6) and (1.7) read as follows: For all $\xi, \xi' \in V_1$ and $\zeta, \zeta' \in V_2$,
\[
\langle (M\xi - \zeta, \xi' - \zeta') \rangle \geq c > 0.
\]
Also, (1.3) is equivalent to dist $(V_1, V_2) \sim 1$. Let $f, g$ be the characteristic functions of the balls $B(-e_{n-1}, 1/4), B(e_{n-1}, 1/4) \subset \mathbb{R}^n$, respectively. Then $\|E_1 f E_2 g\|_q \geq c > 0$ because $(E_1 f E_2 g)(0) \neq 0$. For $\lambda \gg 1$ set
\[
f_\lambda = f(\lambda^{1/2} \xi_1, \ldots, \lambda^{1/2} \xi_{n-2}, \xi_{n-1}, \lambda \xi_n)
\]
and similarly define $g_\lambda$. Observe that the supports of $f_\lambda, g_\lambda$ are in small neighborhoods of $-e_{n-1}, e_{n-1}$, respectively. Obviously, (1.4) is satisfied but (1.6), (1.7) (namely (1.3b)) are not valid. Then by re-scaling $\xi \rightarrow (\lambda^{-1/2} \xi_1, \ldots, \lambda^{-1/2} \xi_{n-2}, \xi_{n-1}, \lambda^{-1} \xi_n)$,
\[
E_1 f_\lambda E_2 g_\lambda(x, t) = \lambda^{-n}(E_1 f E_2 g)(\lambda^{-1/2} x', x_{n-1}, \lambda^{-1} x_n, \lambda^{-1} t),
\]
where $x' = (x_1, \ldots, x_{n-2})$. Suppose (1.2) holds; then we have
\[
C\lambda^{-\frac{n}{2}} \|f\|_p \|g\|_q \geq \|E_1 f_\lambda E_2 g_\lambda\|_q = \lambda^{-n} \lambda^{\frac{n+2}{2q}} \|E_1 f E_2 g\|_q.
\]
Letting $\lambda \rightarrow \infty$ we obtain (1.8).
Theorem [14] can be used to improve the known restriction estimate for hyperboloid in \( \mathbb{R}^3 \). Let us set

\[ \mathcal{H} = \{ (\xi_1, \xi_2, \xi_3^2 - \xi_3^2) : \xi \in [-1, 1]^2 \} \]

**Corollary 1.2.** Let \( d\sigma_{\mathcal{H}} \) be the Lebesgue measure on \( \mathcal{H} \). Then for \( 4 > q > \frac{10}{3} \) and \( \frac{2}{q} < (1 - \frac{1}{p}) \), there is a constant \( C \) such that

\[ \| \widehat{f} d\sigma_{\mathcal{H}} \|_{L^p(\mathbb{R}^3)} \leq C \| f \|_{L^p(d\sigma_{\mathcal{H}})}. \]

The sharp estimate for \( q \geq 4 \left( \frac{2}{q} = (1 - \frac{1}{p}) \right) \) is due to Stein [10]. Unfortunately the argument in [17] using the bilinear estimate (1.3) to derive a linear estimate does not seem to be directly applicable for hyperboloids in higher dimensions \( n \geq 3 \). To get a linear estimate in the higher dimensions, one may need to prove a stronger estimate than Theorem [14]. However, we do not know at this point whether it is possible to get a linear estimate from a bilinear estimate such as Theorem [11].

Now we consider the bilinear restriction estimates for some conic surfaces. Let \( N \) be a nonsingular \((n - 1) \times (n - 1)\) symmetric matrix and let \( V_1, V_2 \) be subcubes of \([-1, 1]^n\). For \( i = 1, 2 \), we define extension operators by

\[ E_i f(x, t) = \int_{V_1} \int_{V_2} e^{2\pi i (x' \cdot \eta + xn \cdot \rho - t(\eta, N\eta/\rho))} f(\eta, \rho) d\eta d\rho, \]

where \( (x', x_n, t) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} \). These can be viewed as the adjoint of Fourier restriction to the conic surfaces

\[ \Gamma_i(N) = \{ (\eta, \rho, \tau) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} : \tau = -(\eta, N\eta/\rho), \eta \in V_i, 1 \leq \rho \leq 2 \}, i = 1, 2. \]

By a linear transform on \( \eta, N \) may always be assumed to be a matrix having nonzero entries 1, -1 only on its diagonal. When \( N \) is the identity matrix, \( \Gamma_i(N) \) is a subset of the light cone. For this the bilinear estimate (1.12) under the condition (1.4) was obtained by Wolff [21]. But for other conic type surfaces, especially with curvature of different signs, similar results were unknown. For this we have the following.

**Theorem 1.3.** Let \( N \) be a diagonal matrix having 1, -1 only on its diagonal. For \( i = 1, 2 \), set

\[ \Theta_i = \{ (\eta/\rho) : (\eta, \rho) \in V_i \times [1, 2] \}. \]

Suppose for all \((\theta_1, \theta_2)\), \((\theta'_1, \theta'_2)\) in \( \Theta_1 \times \Theta_2 \),

\[ ||(\theta_1 - \theta_2, N(\theta'_1 - \theta'_2))|| \sim 1. \]

Then, for \( q > \frac{n+3}{n+1} \), there is a constant \( C \) such that

\[ \| E_1 f \hat{E}_2 g \|_q \leq C \| f \|_2 \| g \|_2. \]

If \( N = id \), condition (1.11) is equivalent to (1.4) (to say, \( (\Theta_1, \Theta_2) \sim 1 \)) provided \( \Theta_1, \Theta_2 \) are small enough. Using the similar argument as before, one can see that if (1.11) is replaced by (1.4), then (1.12) holds only for \( \frac{n+3}{n+1} \leq q \) whenever the eigenvalues of \( N \) have different signs. It is possible to replace \( \langle \eta, N\eta/\rho \rangle \) in (1.10) by a more general function \( \rho h(\eta/\rho) \) with some condition on \( h \), which are similar to (1.0), (1.7) but more complicated. The argument in [12] seems likely to give the endpoint estimate \( q = \frac{n+3}{n+1} \).
As an application of Theorem 1.3 we obtain the almost optimal restriction estimate for a conic surface with negative curvature in $\mathbb{R}^4$. Let us set
\[ C = \{(\eta, \rho, (\eta_1^2 - \eta_2^2)/\rho) \in \mathbb{R}^2 \times \mathbb{R} : \eta \in [-1, 1]^2, 1 \leq \rho \leq 2\}. \]
The following can be compared with the restriction estimate for the cone obtained by Wolff [21].

**Corollary 1.4.** Let $d\sigma_C$ be the Lebesgue measure on $C$. Then, for $4 > q > 3$ and $\frac{2}{q} < (1 - \frac{1}{p})$, there is a constant $C$ such that
\[ ||\hat{f}d\sigma_C||_{L^p(\mathbb{R}^4)} \leq C||f||_{L^p(d\sigma_C)}. \]

Therefore, this solves the restriction problem for $C$ except for the endpoint estimates at the critical line $2/q = (1 - 2/p)$, $4 > q > 3$. The sharp estimate for $q \geq 4$ ($2/q = (1 - 2/p)$) was obtained by Greenleaf [6]. As observed in [21] (Corollary 1), a little strong estimate is possible if one use the mixed norm $|| \cdot ||_{L^p(d\eta)(L^2(d\rho))}$ instead of $|| \cdot ||_{L^p(d\sigma)}$.

After this paper was written, the author was informed that A. Vargas [18] independently obtained Corollary 1.2 establishing (1.3) for $q > \frac{5}{3}$ with $\phi_1 = \phi_2 = \xi_1 \xi_2$ on $\mathbb{R}^3$. Also, the author thanks T. Tao for personal communications about this.

**2. Proof of Theorem 1.1**

The proofs of both Theorems 1.1 and Theorem 1.3 are based on Tao’s proof of the bilinear restriction estimates for paraboloids (12), which is a variant of the induction on scales argument due to Wolff [21].

By the $\epsilon-$ removal argument in [15] it is sufficient for the proof of Theorem 1.1 to show that for any $\alpha > 0$, there is a constant $C = C(\alpha)$ such that for $R \gg 1$,
\[ ||E_1f_1 \cdot E_2f_2||_{L^{n+4}(Q_R)} \leq CR^n||f_1||_2||f_2||_2, \tag{2.1} \]
where $Q_R = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : |t|, |x| \leq R\}$.

A bound with large $\alpha$ is easy to obtain using Hölder’s inequality and $||E_2f||_{\infty} \leq ||f||_2$. The main idea of the induction on scales argument is to establish an iterative estimate which enables us to suppress the exponent $\alpha$ as small as possible.

**Proposition 2.1.** Suppose (2.1) holds for some $\alpha > 0$. Then, for all $0 < \delta, \epsilon \ll 1$, there is constant $C$, independent of $R$, such that for $R \gg 1$,
\[ ||E_1f_1 \cdot E_2f_2||_{L^\frac{n+4}{n+2}(Q_R)} \leq CR^{\max(\alpha(1-\delta), c\delta)+\epsilon}||f_1||_2||f_2||_2 \tag{2.2} \]
holds for some constant $c$, independent of $R, \delta, \epsilon$.

This gives (2.1). Indeed, by choosing $\delta = \alpha/(\alpha + c)$, we have $\max(\alpha(1-\delta), c\delta) = c\alpha/(\alpha + c)$. Repeated uses of Proposition 2.1 produce a sequence of exponents $\{\alpha_j\}$ for which (2.1) is valid with $\alpha = \alpha_j$. This is given by the recursive relation
\[ \alpha_{j+1} = c\alpha_j/(\alpha_j + c) + \epsilon, \quad \alpha_0 = C, \]
where $C$ is a large constant. Then it is easy to see that $\{\alpha_j\}$ monotonically converges to $(\epsilon + \sqrt{\epsilon^2 + 4c})/2$. Since $\epsilon$ can be chosen arbitrarily small, we get (2.1) for all $\alpha > 0$.

The last of this section is devoted to showing Proposition 2.1 Fixing $R \gg 1$, we show (2.1) implies (2.2). In the following, by $C, c$ we denote positive constants which may vary from line to line.
First, we decompose $E_i f_i$ into a sum of wave packets which was introduced in [13] (also, see [21]). The wave packets have good localization property in both Fourier transform side and $(x, t)$-space. The wave packet decomposition at scale $R$ makes the support of functions be expanded by $O(R^{-1/2})$ (see Lemma 2.2). So, we need to consider a little bit larger sets than $V_1, V_2$. For $CR^{-1/2} < \epsilon \ll 1$, let us set
\[ \tilde{V}_i = V_i + O(\epsilon), \]
where for $d > 0$ and $A \subset \mathbb{R}^n, A + O(d)$ denotes the set $\{ x \in \mathbb{R}^n : \text{dist} \ (x, A) \leq Cd \}$. By continuity, we may assume conditions (1.5), (1.6) and (1.7) are still valid if $V_1, V_2$ are replaced by $\tilde{V}_1, \tilde{V}_2$, respectively, and $\epsilon$ is sufficiently small.

2.1. Wave packet decomposition at scale $R$. For a fixed $R \gg 1$, we define the space grid $\mathcal{Y}$ by
\[ \mathcal{Y} = R^{1/2} \mathbb{Z}^n \]
and the frequency grids $\mathcal{V}_1, \mathcal{V}_2$, respectively, by setting
\[ \mathcal{V}_i = R^{-1/2} \mathbb{Z}^n \cap \tilde{V}_i, \quad i = 1, 2. \]
Let us set
\[ W_i = \{(y, v) : (y, v) \in \mathcal{Y} \times \mathcal{V}_i\}. \]
For each $w_i = (y_i, v_i) \in W_i$, we define $(R^{1/2})^n \times R$ tube $T_{w_i}$ by
\[ T_{w_i} = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : |t| \leq R, \quad |x - (y_i + t\nabla \phi_i(v_i))| \leq R^{1/2}\}. \]
Obviously $T_{y_i, v_i}$ meets $(y_i, 0)$ and its major direction is parallel to $(\nabla \phi_i(v_i), 1) \in \mathbb{R}^n \times \mathbb{R}$. The following is a simple modification of Lemma 4.1 in [13].

Let $f_1, f_2$ be supported in $\mathcal{V}_1, \mathcal{V}_2$ respectively. If $|t| \leq R$, we can write $E_i f_i$ as
\[ E_i f_i(x, t) = \sum_{w_i \in W_i} C_{w_i} p_{w_i}(x, t), \quad (x, t) \in \mathbb{R}^n \times \mathbb{R}, \]
such that $p_{w_i} = E_i(p_{w_i}(-, 0))$ and $C_{w_i}, p_{w_i}$ satisfy the following:

(P1) For $i = 1, 2$, $\sum_{w_i \in W_i} |C_{w_i}|^2 \leq C \|f_i\|_2^2$.

(P2) If $w_i = (y_i, v_i)$, $\text{supp} \ p_{w_i}(-, t) \subset \{ \xi : \xi = v_i + O(R^{-1/2}) \}$.

(P3) If $\text{dist}((-x, t), T_{y_i, v_i}) \geq R^{\delta + 1/2}$, then $|p_{y_i, v_i}(x, t)| \leq C R^{-100n}$. More precisely,
\[ |p_{y_i, v_i}(x, t)| \leq C R^{-n/4} \left( 1 + \frac{|x - (y_i + t\nabla \phi_i(v_i))|}{R^{1/2}} \right)^{-N} \text{ for any } N. \]

(P4) For any $S \subset W_i$, $\sum_{w_i \in S} \|p_{w_i}(-, t)\|^2 \leq C \# S$.

Proof of Lemma 2.2 Using the Poisson summation formula, one can find an $\eta$ satisfying $\text{supp} \ \tilde{\eta} \subset B(0, 1)$ and $\sum_{k \in \mathbb{Z}^n} \eta(\cdot - k) = 1$. Here $B(a, r) \subset \mathbb{R}^n$ denotes the ball centered at $a$ with radius $r$. Let $\psi \in C_0^\infty(B(0, 1))$ with $\sum_{k \in \mathbb{Z}^n} \psi(\cdot - k) = 1$. For $y \in \mathcal{Y}$ and $v_i \in \mathcal{V}_i$, let us set
\[ \eta_y(x) = \eta(\frac{x + y}{R^{1/2}}), \quad \psi_{y_i}(\xi) = \psi(R^{1/2}(\xi - v_i)). \]
Note that \( f_i = \sum_{v_i \in V_i, y \in Y} \mathcal{F}^{-1}(\hat{\psi}_i \hat{f}_i \eta_y) \) since \( f_i \) is supported in \( V_i \) and \( \mathcal{F}^{-1}(\hat{\psi}_i \hat{f}_i \eta_y) \) is supported in the \((CR^{-1})\)-neighborhood of \( v_i \). Here \( \mathcal{F}^{-1} \) denotes the inverse Fourier transform. Then it follows that
\[
E_i f_i(x, t) = \sum_{v \in V_i, y \in Y} q_{y,v_i}(x, t),
\]

where
\[
q_{y,v_i}(x, t) = \int e^{2\pi i(x \cdot \xi - t \phi_i(\xi))} \mathcal{F}^{-1}(\hat{\psi}_i \hat{f}_i \eta_y) \xi d\xi.
\]
(\text{2.3})

Now we show

**Lemma 2.3.** If \( |t| \leq R \), for any \( N \)
\[
|q_{y,v_i}(x, t)| \leq C M(\psi_i, f_i)(-y) \left( 1 + \frac{|x - (y + t\nabla \phi_i(v_i))|}{R^{1/2}} \right)^{-N},
\]
where \( M \) denotes the Hardy-Littlewood maximal function.

**Proof of Lemma 2.3.** Since \( \mathcal{F}^{-1}(\hat{\psi}_i \hat{f}_i \eta_y) \) is supported in \( v_i + O(R^{-1/2}) \), we can find a smooth cutoff function \( \tilde{\psi} \) so that \( \tilde{\psi}_{v_i} = 1 \) on the support of \( \mathcal{F}^{-1}(\hat{\psi}_i \hat{f}_i \eta_y) \). (Here \( \tilde{\psi}_{v_i} \) is defined from \( \tilde{\psi} \) by the same way as \( \psi_{v_i} \).)

For simplicity we set \( F_{v_i} = \tilde{\psi}_{v_i} \tilde{f}_i \). By putting \( \tilde{\psi}_{v_i} \) into the integral in (2.3), translation and re-scaling, we have
\[
q_{y,v_i}(x, t) = R^{-n/2} \int K(x + z, t) \eta\left(\frac{z + y}{R^{1/2}}\right) F_{v_i}(z) dz,
\]

where
\[
K(x, t) = \int e^{2\pi i(R^{-1/2} x \cdot \xi + x \cdot v_i - t \phi_i(R^{-1/2} \xi + v_i))} \tilde{\psi}(\xi) d\xi.
\]

Note that \( R^{-1/2} t(\nabla \phi_i(R^{-1/2} \xi + v_i) - \nabla \phi_i(v_i)) = O(1) \) if \( |t| \leq R \). From this, if \( |t| \leq R \) and \( |x - t \phi_i(v_i)| \geq CR^{1/2} \),
\[
R^{-1/2}|x - t \nabla \phi_i(v_i)| \leq C|\nabla \xi_R(\xi)| \leq C|\nabla \xi_R(\xi)| \leq C|x - t \nabla \phi_i(v_i)|.
\]

By integration by parts, we see that if \( |t| \leq R \),
\[
|K(x, t)| \leq C \left( 1 + \frac{|x - t \nabla \phi_i(v_i)|}{R^{1/2}} \right)^{-N}
\]
for any \( N \). We translate \( z \to z - y \) and set \( a = x - (y + t \nabla \phi_i(v_i)) \). Then
\[
|q_{y,v_i}(x, t)| \leq CR^{-n/2} \int \left( 1 + \frac{|a + z|}{R^{1/2}} \right)^{-N} |\eta\left(\frac{z}{R^{1/2}}\right) F_{v_i}(z - y)| dz.
\]

Therefore it is sufficient to show that for \( \lambda \gg 1 \)
\[
(2.4) \quad I = \lambda^{-n} \int \left( 1 + \frac{|a + z|}{\lambda} \right)^{-N} |\eta\left(\frac{z}{\lambda}\right) F(z)| dz \leq C(1 + |a|/\lambda)^{-N} M F(0)
\]
for any \( N \). If \( |a| \leq \lambda \), there is nothing to prove. We may assume \( |a| \gg \lambda \). We divide the integral \( I \) into \( \{I_j\} \) so that \( I = \sum_0^\infty I_j \) and
\[
\sum_{j=0}^\infty I_j \leq C \lambda^{-n} \left( \int_{|z| \leq \lambda/2} + \sum_{j=1}^\infty \int_{|z| \simeq \lambda^j} \right) \left( 1 + \frac{|a + z|}{\lambda} \right)^{-N} |\eta\left(\frac{z}{\lambda}\right) F(z)| dz.
\]
Then, if $|a| \sim \lambda^{2j}$, by the rapid decay of $\gamma$, $
abla_j \leq C(|a|/\lambda)^{-M}MF(0)$ for any $M$. If $|a| \not\sim \lambda^{2j}$, $
abla_j \leq C2^{-Mj}(|a|/\lambda)^{-N}MF(0)$ for any $M$ because $C|a + z| \geq |a|$ if $|a| \not\sim |z|$. Summation with respect to $j$ gives (2.4). \hfill \Box

Now we turn to the proof of Lemma 2.2. For $w_i = (y_i, v_i) \in W_i$, let us set

$$C_{y_i, v_i} = R^{j/4}M(\psi_{y_i,v_i})(-y_i), \quad p_{y_i,v_i} = q_{y_i,v_i}/C_{y_i, v_i}.$$ 

Then, $p_{w_i} = E_i(p_{w_i, (\cdot, 0)})$ is obvious from (2.3) and, using Lemma 2.3, we see (P3). By the support property of $\psi_{v_i}$, $\eta_{y_i}$, it is easy to see

$$\text{supp } \hat{\phi}_{y_i,v_i}(\cdot,t) \subset \{\xi : \xi = v_i + O(R^{-1/2})\}$$

for each $t, y_i$. So we get (P2). By Plancherel’s theorem we have

$$\left\| \sum_{v_i \in V_i, y_i \in Y} p_{y_i,v_i}(\cdot,t) \right\|_2 \leq C \left( \sum_{v_i \in V_i} \left\| \sum_{y_i \in Y} p_{y_i,v_i}(\cdot,t) \right\|_2^2 \right)^{1/2}.$$ 

Using this and $\|p_{y_i,v_i}(\cdot,t)\|_2 \leq C$, one can easily see (P4) since for fixed $v_i$, $\{p_{y_i,v_i}(\cdot,t)\}_{y_i \in Y}$ are essentially disjoint (see (P3)).

Since $\psi_{v_i,f_i}$ is supported in a ball of radius about $R^{-1/2}$, if $|x - x'| \leq CR^{1/2}$, $M(\psi_{v_i,f_i})(x) \sim M(\psi_{v_i,f_i})(x')$ (see (2.3)). This can be shown using a bump function adapted to the ball where $\psi_{v_i,f_i}$ is supported. Hence,

$$\sum_{w_i \in W_i} |C_{w_i}|^2 \leq C \sum_{v_i} \int |M(\psi_{v_i,f_i})(y)|^2 dy \leq C \sum_{v_i} \int |\psi_{v_i,f_i}|^2 dy \leq C \int |f_i(y)|^2 dy.$$ 

The second inequality is from the Hardy-Littlewood maximal theorem and the third from Plancherel’s theorem. So we get (P1). \hfill \Box

2.2. Reduction. We normalize $\|f_1\|_2 = \|f_2\|_2 = 1$ and fix a small $\delta > 0$. Then, using Lemma 2.2, we have for $i = 1, 2$,

$$E_i f_i = \sum_{w_i \in W_i} C_{w_i} p_{w_i}.$$ 

For Proposition 2.4 it is sufficient to show that (2.1) implies

$$\sum_{w_1 \in W_1} \sum_{w_2 \in W_2} C_{w_1}C_{w_2}p_{w_1}p_{w_2} \|_{L^{\alpha+\delta}((Q_R))} \lesssim (R^{\alpha(1-\delta)} + R^\delta).$$ 

Here, $A \lesssim B$ means there is a constant $C$ such that $A \leq C, R^{\epsilon}B$ for any $\epsilon > 0$, $R > 1$. For each cube $Q$, we denote by $CQ$ the cube which has the same center as $Q$ and side length $C$-times as long as that of $Q$.

Since $p_{w_i}$ is essentially supported on the tubes $T_{w_i}$ (see (P3) in Lemma 2.2), we may always assume that if $w_i \in W_i$, $i = 1, 2$, then for some large $C > 0$,

$$T_{w_i} \cap CQ_R \neq \emptyset, \quad T_{w_2} \cap CQ_R \neq \emptyset$$

because the contribution from the others is $O(R^{-100n})$. Indeed, since $|C_{w_1}|, |C_{w_1}| \leq C$ by (P1), using (P3) in Lemma 2.2 it is easy to see that for some large $C > 0$,

$$\sum_{T_{w_1} \cap CQ_R = \emptyset \text{ or } T_{w_2} \cap CQ_R = \emptyset} C_{w_1}C_{w_2}p_{w_1}p_{w_2} \|_{L^{\alpha+\delta}((Q_R))} \lesssim R^{-100n}.$$ 

Discarding these, from (2.6) we see that both the numbers of the remaining $w_1, w_2$ (also the tubes) are $O(R^{50n})$. Then one can easily see from a rough estimate that the contribution from the terms with $|C_{w_1}| \leq R^{-200n}$ or $|C_{w_2}| \leq R^{-200n}$ are
O(R^{-30n}). Hence we may also assume \( C \geq |C_{w_1}|, |C_{w_2}| \geq R^{-200n} \). Therefore, by the pigeonhole principle, there are dyadic numbers \( R^{-200n} \leq \mu_1, \mu_2 \leq C \) such that

\[
\text{LHS of (2.5)} \leq C(\log R)^2 \sum_{\mu_1 \leq |C_{w_1}| < 2\mu_1, \mu_2 \leq |C_{w_2}| < 2\mu_2} C_{w_1}C_{w_2}p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)}.
\]

This kind of pigeonhole argument which fixes an involved quantity at the expense of \( C(\log R)^c \) in the bound will be used several times.

Since \( (\#\{w_i : |C_{w_i}| \sim \mu_i\})^\frac{2}{3} \leq C/\mu_i \) from (P1), for (2.5) it is sufficient to show that for any subset \( W_1 \subset W_2 \subset W_2 \),

\[
(2.7) \quad \| \sum_{w_1 \in W_1} \sum_{w_2 \in W_2} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \leq C(\log R)^c (\#W_1 \#W_2)^{1/2}.
\]

In fact, since \( |C_{w_1}|/\mu_1, |C_{w_2}|/\mu_2 \sim 1 \), we may absorb harmless constants \( C_{w_1}/\mu_1, C_{w_2}/\mu_2 \) into \( p_{w_1}, p_{w_2} \), respectively, because this does not affect the properties (P1)- (P4) in Lemma 2.2.

Now we decompose \( Q_R \) into essentially disjoint cubes \( b \) of side length \( R^{1-\delta} \) and denote by \( B \) the collection of these cubes. By triangle inequality

\[
\| \sum_{w_1 \in W_1} \sum_{w_2 \in W_2} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \leq C \sum_{b \in B} \| \sum_{w_1 \in W_1} \sum_{w_2 \in W_2} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}.
\]

Suppose given a relation \( \approx \) between \( w_i, i = 1, 2 \) and \( b \), which will be defined in the next subsection. We break the right-hand side of the above into two parts so that

\[
(2.8) \quad \| \sum_{w_1 \in W_1} \sum_{w_2 \in W_2} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \leq \| \sum_{w_1 \approx b, w_2 \approx b} \sum_{w_1 \approx b} \sum_{w_2 \approx b} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}
\]

Roughly, \( w_1 \approx b \) means \( T_{w_1} \) is one of highly concentrating tubes on \( b \). The division in (2.8) gives the high concentration part and the low concentration part according to the overlapping degree of tubes on \( b \). As to be seen later, the low concentration part can be directly handled by utilizing orthogonality among wave packets and the geometry of concentrating tubes. However it is hard to get estimate for the high concentration part, as one might expect. Instead, the induction assumption (2.1) is to be used for \( \| \sum_{w_1 \approx b} \sum_{w_2 \approx b} p_{w_1}p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)} \). This gives \( O(R^\alpha(1-\delta)) \) bound because \( b \) are cubes of side length \( R^{1-\delta} \). Since there are many \( b \), the number of \( w_1, w_2 \approx b \) should be controlled to get the required bound.

2.3. The relation \( \approx \) between \( w_i \in W_i \) and \( b \). Now we divide \( Q_R \) into essentially disjoint cubes \( q \) of side length \( R^{1/2} \). Let us denote by \( Q \) this collection of \( q \). For \( q \in Q \), let us define

\[
W_i(q) = \{w_i \in W_i : T_{w_i} \cap R^\delta q \neq \emptyset\}.
\]

For dyadic numbers \( 1 \leq \mu_1, \mu_2 \leq R^{100n} \), we set

\[
Q(\mu_1, \mu_2) = \{q \in Q : \mu_1 \leq \#W_1(q) < 2\mu_1, \mu_2 \leq \#W_2(q) < 2\mu_2\}.
\]

For \( w_i \in W_i \) and dyadic numbers \( 1 \leq \mu_1, \mu_2 \leq R^{100n} \), let us set

\[
\lambda(w_1, \mu_1, \mu_2) = \#\{q \in Q(\mu_1, \mu_2) : R^\delta q \cap T_{w_1} \neq \emptyset\}
\]

and for dyadic numbers \( 1 \leq \lambda \leq R^{100n} \),

\[
W_i(\lambda, \mu_1, \mu_2) = \{w_i \in W_i : \lambda \leq \lambda(w_1, \mu_1, \mu_2) < 2\lambda\}.
\]
For each dyadic number $1 \leq \lambda, \mu_1, \mu_2 \leq R^{100} n$ and $w_i \in W_i(\lambda, \mu_1, \mu_2)$, we define $b(w_i, \lambda, \mu_1, \mu_2) \in B$ to be a ball which maximizes the quantity

$$
\# \{ q \in Q(\mu_1, \mu_2) : R^q q \cap T_{w_i} \neq \emptyset, q \cap b(w_i, \lambda, \mu_1, \mu_2) \neq \emptyset \}.
$$

Possibly there may be many candidates for $b(w_i, \lambda, \mu_1, \mu_2)$. Then we simply choose one among them. Since $\# B \sim R^{(n+1)\delta}$, it follows that

$$
(2.9) \quad \# \{ q \in Q(\mu_1, \mu_2) : R^q q \cap T_{w_i} \neq \emptyset, q \cap b(w_i, \lambda, \mu_1, \mu_2) \neq \emptyset \} \gtrsim \lambda R^{-(n+1}\delta).
$$

We define a relation $\approx$ between $w_i \in W_i(\lambda, \mu_1, \mu_2)$ and $b \in B$ by

$$
w_i \approx_{\lambda, \mu_1, \mu_2} b \text{ if } b \cap 10b(w_i, \lambda, \mu_1, \mu_2) \neq \emptyset.
$$

From this we define a relation $\approx$ between $w_i \in W_i$ and $b$ by saying

$$w_i \approx b \text{ if } w_i \approx_{\lambda, \mu_1, \mu_2} b \text{ for some dyadic } 1 \leq \lambda, \mu_1, \mu_2 \leq R^{100} n.$$

Clearly for each $w_i$, there are $O((\log R)^3)$ balls $b$ in $B$ for which $w_i \approx b$ since there are $O((\log R)^3)$ dyadic triples $(\lambda, \mu_1, \mu_2)$. So it follows that for all $w_i \in W_i$,

$$
(2.10) \quad \# \{ b \in B : w_i \approx b \} \lesssim 1.
$$

Since $\sum_{w_i \approx b} p_{w_i} = E_i(\{\sum_{w_i \approx b} p_{w_i} \cdot \mathbf{1}_b \})$ from Lemma 2.2 and $b$ is a $R^{1-\delta}$-cube, the hypothesis (2.1), (P4) in Lemma 2.2 and Plancherel’s theorem give

$$
\sum_{b \in B} \| \sum_{w_i \approx b} p_{w_i} p_{w_2} \|_{L^{\frac{n+1}{n}}} (b) \lesssim CR^{\alpha(1-\delta)} \sum_{b \in B} \# \{ w_i : w_i \approx b \}^{\frac{3}{2}} \# \{ w_2 : w_2 \approx b \}^{\frac{1}{2}}.
$$

By Schwarz’s inequality, changing the order of summation and (2.10), we see

$$
\sum_{b \in B} \| \sum_{w_i \approx b} p_{w_i} p_{w_2} \|_{L^{\frac{n+1}{n}}} (b)
\leq CR^{\alpha(1-\delta)} \left( \sum_{w_i \in W_1} \# \{ b \in B : w_i \approx b \} \right)^{1/2} \left( \sum_{w_2 \in W_2} \# \{ b \in B : w_2 \approx b \} \right)^{1/2}
\lesssim R^{\alpha(1-\delta)} (\# W_1)^{1/2} (\# W_2)^{1/2}.
$$

2.4. $R^{1/2}$-scale decomposition. For (2.7) we need to show

$$
\sum_{b \in B} \| \sum_{w_i \approx b \text{ or } w_2 \approx b} p_{w_i} p_{w_2} \|_{L^{\frac{n+1}{n}}} (b) \lesssim R^{\delta} (\# W_1)^{1/2} (\# W_2)^{1/2}.
$$

Since $\# B \sim R^{(n+1)\delta}$, it suffices to show that for all $b \in B$,

$$
(2.11) \quad \| \sum_{w_i \approx b \text{ or } w_2 \approx b} p_{w_i} p_{w_2} \|_{L^{\frac{n+1}{n}}} (b) \lesssim R^{\delta} (\# W_1)^{1/2} (\# W_2)^{1/2}.
$$

Here, we are assuming the sum is always taken over some subset of $W_1 \times W_2$. Since $p_{w_i} = E_i(\{ p_{w_i} \cdot \mathbf{1}_b \})$, using the local restriction estimate $\| E_i f \|_{L^2(Q_R)} \leq CR^{\frac{1}{2}} \| f \|_2$ (see 17), Schwarz’s inequality and (P4) in Lemma 2.2, we obtain

$$
\| \sum_{w_i \approx b \text{ or } w_2 \approx b} p_{w_i} p_{w_2} \|_{L^1(b)} \lesssim R^{1-\delta} (\# W_1)^{1/2} (\# W_2)^{1/2}.
$$

Hence, in view of interpolation it suffices for (2.11) to show

$$
(2.12) \quad \| \sum_{w_i \approx b \text{ or } w_2 \approx b} p_{w_i} p_{w_2} \|_{L^2(b)} \lesssim R^{\delta} R^{-(n-1)/4} (\# W_1)^{1/2} (\# W_2)^{1/2}.
$$
Since $L^2$-norm is taken in the left-hand side of (2.12), the orthogonality among $p_{w_1}, p_{w_2}$ can be effectively exploited in their Fourier transforms via Plancherel’s theorem. For this, we decompose $b$ into $R^{1/2}$-cubes $q \in \mathcal{Q}$. Obviously, (2.12) follows from

$$
\sum_{q \in \mathcal{Q} : q \subset 2b} || \sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1}p_{w_2} ||_{L^2(q)}^2 \lesssim R^{c \delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.
$$

Now, we make several obvious reductions by pigeonholing. Since $|p_{w_i}| \leq CR^{-100n}$ on $q$ if $T_{w_i} \cap R^q = \emptyset$ by (P3), in the left-hand side of (2.13) we may replace $\sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b}$ by $\sum_{w_1 \in \mathcal{W}_1(q), w_1 \not\approx b} \sum_{w_2 \in \mathcal{W}_2(q), w_2 \not\approx b}$ discarding some $O(R^{-30n})$ terms. Since all the $q$ appearing in (2.13) are contained in $\bigcup_{\mu_1, \mu_2} \mathcal{Q}(\mu_1, \mu_2)$, by pigeonholing on dyadic numbers $1 \leq \mu_{1,2} \leq R^{100n}$, the $\mathcal{Q}$ in the outer sum of (2.13) can also be replaced by $\mathcal{Q}(\mu_1, \mu_2)$ for some $\mu_{1,2}$. Therefore, the matters are reduced to showing

$$
\sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} || \sum_{w_1 \in \mathcal{W}_1(q), w_1 \not\approx b} \sum_{w_2 \in \mathcal{W}_2(q), w_2 \not\approx b} p_{w_1}p_{w_2} ||_{L^2(q)}^2 \lesssim R^{c \delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.
$$

Clearly, $\bigcup_{q \in \mathcal{Q}(\mu_1, \mu_2)} \mathcal{W}_i(q) = \bigcup_{i=1,2} \mathcal{W}_i(\lambda_1, \mu_1, \mu_2), i = 1, 2$. Hence, by pigeonholing over dyadic numbers $1 \leq \lambda_1, \lambda_2 \leq R^{100n}$, we may replace $\mathcal{W}_1(q), \mathcal{W}_2(q)$ in the inner sum of (2.14) by $\mathcal{W}_1(q) \cap \mathcal{W}_2(q)$ in the inner sum of (2.14) by $\mathcal{W}_1(q) \cap \mathcal{W}_2(q)$, respectively. To simplify the notation, we set

$$
\mathcal{W}_i^{\not\approx b}(q, \mu_1, \mu_2, \lambda_i) = \{ w_i \in \mathcal{W}_i(q) \cap \mathcal{W}_i(\lambda_i, \mu_1, \mu_2) : w_i \not\approx b \}.
$$

Let $U_i \subset \mathcal{W}_i$ and define $U_i(q)$ by

$$
U_i(q) = \{ w_i \in U_i : T_{w_i} \cap R^q \neq \emptyset \}.
$$

Breaking $\sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b}$ into $\sum_{w_1 \not\approx b, w_2 \not\approx b} + \sum_{w_1 \not\approx b, w_2 \approx b} + \sum_{w_1 \not\approx b, w_2 \approx b}$ by symmetry it is sufficient for (2.14) to show that for any $U_2 \subset \mathcal{W}_2$ and dyadic numbers $1 \leq \lambda_1, \mu_1, \mu_2 \leq R^{100n}$,

$$
\sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} || \sum_{w_1 \in U_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)} \sum_{w_2 \in U_2(q)} p_{w_1}p_{w_2} ||_{L^2(q)}^2 \lesssim R^{c \delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.
$$

2.5. Orthogonality among wave packets. In this subsection we work on the Fourier transform side and we utilize the frequency localization property of the wave packet decomposition (see (P2) in Lemma 2.2).

For $\xi_1 \in \tilde{V}_1$, $\xi_2' \in \tilde{V}_2$, let us define a function $\Phi_{\xi_1, \xi_2'} : \tilde{V}_1 \to \mathbb{R}$ by

$$
\Phi_{\xi_1, \xi_2'}(\xi_1') = \phi_1(\xi_1') + \phi_2(\xi_1' + \xi_2' - \xi_1) - \phi_1(\xi_1') - \phi_2(\xi_2').
$$

We also define a set $\Pi_{\xi_1, \xi_2'}$ by

$$
\Pi_{\xi_1, \xi_2'} = \{ \xi_1' \in \tilde{V}_1 : \Phi_{\xi_1, \xi_2'}(\xi_1') = 0, \xi_1' + \xi_2' - \xi_1 \in \tilde{V}_2 \}.\]
From (1.6) and (1.7) which imply (1.4), we see
\begin{equation}
|\nabla \Phi_{\xi_1, \xi_2}(\xi_1')| = |\nabla \Phi_{2}(\xi_1' + \xi_2' - \xi_1) - \nabla \Phi_{1}(\xi_1')| \geq c > 0
\end{equation}
provided \( \xi_1' + \xi_2' - \xi_1 \in \tilde{V}_2 \). So, dividing \( \tilde{V}_1, \tilde{V}_2 \) into sufficiently small cubes, we may assume the sets \( \Pi_{\xi_1, \xi_2}' \) are smooth hyper-surfaces for all \( \xi_1 \in \tilde{V}_1, \xi_2' \in \tilde{V}_2 \) because \( E_1fE_{2g} \) is written as a finite sum of operators satisfying this assumption.

For \( \mathcal{U}_1 \subset \mathcal{W}_1 \), we define
\begin{equation}
\mathcal{N}(\mathcal{U}_1) = \sup_{\xi_1 \in \mathcal{V}_1, \xi_2 \in \mathcal{V}_2} \#\{v_1' \in \mathcal{U}_1 : v_1' \in \Pi_{\xi_1, \xi_2} + O(R^{-1/2})\}.
\end{equation}

Lemma 2.4. For \( q \in Q \) and \( \mathcal{U}_i \subset \mathcal{W}_i(q), i = 1, 2 \),
\begin{equation}
\| \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2 \in \mathcal{U}_2} p_{w_1}p_{w_2} \|_{L^2}^2 \lesssim R^q R^{-(n-1)/2} \mathcal{N}(\mathcal{U}_1) \mathcal{N}(\mathcal{U}_1) \#\mathcal{U}_1 \#\mathcal{U}_2.
\end{equation}

Proof of Lemma 2.4. We write
\begin{equation}
\| \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2 \in \mathcal{U}_2} p_{w_1}p_{w_2} \|_{L^2}^2 = \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2' \in \mathcal{U}_2} \sum_{w_1' \in \mathcal{U}_1} I_{w_1, w_2', w_1'}
\end{equation}
where
\begin{equation}
I_{w_1, w_2', w_1'} = \sum_{w_2 \in \mathcal{U}_2} \langle p_{w_1}p_{w_2}, p_{w_1'}p_{w_2'} \rangle.
\end{equation}
By recalling \( w = (y, v) \) and using (P2), we see \( \langle p_{w_1}p_{w_2}, p_{w_1'}p_{w_2'} \rangle = 0 \) unless \( v_1, v_1' \in V_1 + O(R^{-1/2}), v_2, v_2' \in V_2 + O(R^{-1/2}) \),
\begin{equation}
v_1 + v_2 = v_1' + v_2' + O(R^{-1/2}) \quad \text{and} \quad \phi_1(v_1) + \phi_2(v_2) = \phi_1(v_1') + \phi_2(v_2') + O(R^{-1/2})
\end{equation}
because \( \tilde{p}_{w_1} \) are supported in the \( O(R^{-1/2}) \)-neighborhood of the point \((v_1, \phi(v_1))\).

From (2.20), we see \( \langle p_{w_1}p_{w_2}, p_{w_1'}p_{w_2'} \rangle = 0 \) if \( v_1' + v_2' - v_1 \notin \tilde{V}_2 \). So, we may always assume \( v_1' + v_2' - v_1 \in \tilde{V}_2 \). If \( I_{w_1, w_2', w_1'} \neq 0 \) for some fixed \( w_1, w_2' \), then both (2.20) and (2.21) should be satisfied by some \( v_2 \in V_2 \). Hence, if \( I_{w_1, w_2', w_1'} \neq 0 \) for fixed \( w_1 \) and \( w_2' \), then \( v_1' \) satisfies
\begin{equation}
\phi_1(v_1) + \phi_2(v_2) = \phi_1(v_1') + \phi_2(v_2') + O(R^{-1/2})
\end{equation}
Therefore, from (2.13), it is obvious that \( v_1' \) satisfying the above is contained in \( \Pi_{v_1, v_1'} + O(R^{-1/2}) \).

Therefore, the left-hand side of (2.11) is equal to
\begin{equation}
\sum_{w_1 \in \mathcal{U}_1} \sum_{w_2' \in \mathcal{U}_2} \sum_{\{w_1' \in \mathcal{U}_1 : v_1' \in \Pi_{v_1, v_1'} + O(R^{-1/2})\}} \left( \sum_{\{w_2 \in \mathcal{U}_2 : v_2 = v_1' + v_2' - v_1 + O(R^{-1/2})\}} \langle p_{w_1}p_{w_2}, p_{w_1'}p_{w_2'} \rangle \right)
\end{equation}
If \( w_1, w_2', w_1' \) are given, then there are at most \( O(1) \)-\( v_2 \) because these are determined by (2.20). Since all the tubes \( T_{v_2} \) \( (T_{w_1} \text{ resp.)} \) are passing through \( R^q \), there are at most \( O(R^q) \)-\( v_2 \) \( (w_1 \text{ resp.)} \) if \( v_2 \) \( (v_1 \text{ resp.)} \) is determined because \( y_2 \in Y \) \( (y_1 \text{ resp.)} \) are \( R^{1/2} \)-separated. Using (P3) and the transversality between the tubes \( T_{w_1} \) and
Since \( w \leq R \) (by (1.24)), it is easy to see that \( |\langle p_{w_1}, p_{w_2}, p_{w_1}p_{w_2} \rangle| \lesssim R^{-(n-1)/2} \). Therefore, we see that for fixed \( w_1, w_2, w_1', \)
\[
\sum_{\{w_2 \in U_2 : \text{w_2 = w_1'} + v_1' - v_1 + O(R^{-1/2})\}} |\langle p_{w_1}, p_{w_2}, p_{w_1}p_{w_2} \rangle| \lesssim R^\delta R^{-(n-1)/2}.
\]
From this and (2.22) we conclude the proof of Lemma 2.4.

2.6. Proving Theorem 1.1. By showing (2.16), we prove Theorem 1.1 using the following which will be shown in the next subsection.

Lemma 2.5 (Combinatorial estimates). For dyadic numbers \( 1 \leq \mu_1, \mu_2, \lambda_1 \leq R^{100n} \) and \( q \in Q(\mu_1, \mu_2) \), \( q \subset 2b \),
\[
\mathcal{N}(W_1^{gb}(q, \mu_1, \mu_2, \lambda_1)) \lesssim R^\delta \frac{\# W_2}{\lambda_1 \mu_2}.
\]

Using Lemma 2.4, we see
\[
\sum_{q \in Q(\mu_1, \mu_2), q \subset 2b} \sum_{w_1 \in W_1^{gb}(q, \mu_1, \mu_2, \lambda_1)} \sum_{w_2 \in U_2(q)} P_{w_1} P_{w_2} \|P_{w_1} P_{w_2}\|_{L^2(q)}^2 \lesssim R^\delta R^{\frac{1}{4}} \sum_{q \in Q(\mu_1, \mu_2), q \subset 2b} \mathcal{N}(W_1^{gb}(q, \mu_1, \mu_2, \lambda_1)) \# W_1^{gb}(q, \mu_1, \mu_2, \lambda_1) \# U_2(q).
\]

Using Lemma 2.5 for (2.16) it suffices to show
\[
\frac{1}{\lambda_1} \sum_{q \in Q(\mu_1, \mu_2), q \subset 2b} \# W_1^{gb}(q, \mu_1, \mu_2, \lambda_1) \# U_2(q) \lesssim \# W_1 \# W_2.
\]

Since \( q \in Q(\mu_1, \mu_2) \) and \( U_2 \subset W_2 \), \#\( U_2(q) \) \( \leq \mu_2 \). So we need to show
\[
\sum_{q \in Q(\mu_1, \mu_2), q \subset 2b} \# W_1^{gb}(q, \mu_1, \mu_2, \lambda_1) \lesssim \# W_1.
\]

Recalling (2.15) and changing the order of summation,
\[
\sum_{q \in Q(\mu_1, \mu_2), q \subset 2b} \# W_1^{gb}(q, \mu_1, \mu_2, \lambda_1)
\leq \sum_{w_1 \in W_1(\lambda_1, \mu_1, \mu_2)} \# \{q \in Q(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \not= \emptyset, w_1 \not= b\}
\leq \sum_{w_1 \in W_1(\lambda_1, \mu_1, \mu_2)} \# \{q \in Q(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \not= \emptyset\}.
\]

Since \( w_1 \in W_1(\lambda_1, \mu_1, \mu_2) \), \#\( \{q \in Q(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \not= \emptyset\} \leq \lambda_1 \). Therefore, we get (2.23) and hence prove Theorem 1.1.

2.7. Proof of Lemma 2.5. We need to show that for dyadic numbers \( 1 \leq \mu_1, \mu_2, \lambda_1 \leq R^{100n} \), \( \xi_1 \in \tilde{V}_1, \xi_2' \in \tilde{V}_2 \), and \( q_0 \in Q(\mu_1, \mu_2) \), \( q_0 \subset 2b \),
\[
\# \{w_1' \in W_1^{gb}(q_0, \mu_1, \mu_2, \lambda_1) : v_1' \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\} \lesssim R^\delta \frac{\# W_2}{\lambda_1 \mu_2}.
\]

For simplicity we set
\[
W_1^{gb}(\Pi_{\xi_1, \xi_2'}) = \{w_1' \in W_1^{gb}(q_0, \mu_1, \mu_2, \lambda_1) : v_1' \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\}.
\]
Let $w_1 \in W^{p,b}_{1}(\Pi_{1,2})$. Then $T_{w_1} \cap R^q g_0 \neq \emptyset$ and $b \cap 10b(w_1, \lambda_1, \mu_1, \mu_2) = \emptyset$. Since $g_0 \subset 2b$, dist $(q_0, 2b(w_1, \lambda_1, \mu_1, \mu_2)) \geq R^{1-\delta}$. So, by (2.9) we see

\[(2.25) \#\{q \in Q(\mu_1, \mu_2) : R^q \cap T_{w_1} \neq \emptyset, \text{dist} (q_0, q) \geq R^{1-\delta} \} \geq \lambda_1 R^{-(n+1)\delta}.\]

By the definition of $Q(\mu_1, \mu_2)$, $\#W_2(q) \sim \mu_2$ for each $q \in Q(\mu_1, \mu_2)$. From (2.25) and summation in $w_1, w_2$, we get

\[(2.26) \#\{(q, w_1, w_2) \in Q(\mu_1, \mu_2) \times W^{p,b}_{1}(\Pi_{1,2}) \times W_2 : R^q \cap T_{w_1} \neq \emptyset, \text{dist} (q_0, q) \geq R^{1-\delta}, R^q \cap T_{w_2} \neq \emptyset \} \geq R^{-(n+1)\delta} \lambda_1 \#W^{p,b}_{1}(\Pi_{1,2}) \mu_2.\]

Now we try to obtain an upper bound for the left-hand side of the above to obtain (2.24) by a simple arithmetic manipulation. It will be done by the following.

**Lemma 2.6.** For each $w_2 \in W_2$, set

\[S = \{(q, w_1) \in Q(\mu_1, \mu_2) \times W^{p,b}_{1}(\Pi_{1,2}) : R^q \cap T_{w_1} \neq \emptyset, \text{dist} (q_0, q) \geq R^{1-\delta}, R^q \cap T_{w_2} \neq \emptyset \}.\]

Then, $\#S = O(R^c)$ for some $c > 0$, independent of $R$.

Using Lemma 2.6 the left-hand side of (2.20) is bounded by $R^{c} \#W_2$. Therefore,

\[\lambda_1 \#W^{p,b}_{1}(\Pi_{1,2}) \mu_2 \lesssim R^{c} \#W_2.\]

From this (2.24) follows. It remains to show Lemma 2.6. For this we use the following elementary lemma.

**Lemma 2.7.** Let $\Pi$ be a smooth compact hypersurface with boundary in $B(0, 1) \subset \mathbb{R}^n$ and let $\mathbb{P}_p(\Pi) \subset \mathbb{R}^n$ be the hyperplane tangent to $\Pi$ at $p$. For $\delta' \leq \lambda \leq \mu \leq 1$, $0 < \epsilon \ll 1$, set

\[C(\Pi, \lambda, \mu) = \{(s, u, 1) \in \mathbb{R}^n \times \mathbb{R} : \lambda \leq s \leq \mu, u \in \Pi \},\]

\[C(\Pi, \lambda, \mu, \delta) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \text{dist} ((x, t), C(\Pi, \lambda, \mu)) \leq \delta \}.\]

For $y, v \in \mathbb{R}^n$, let $l^y_v \subset \mathbb{R}^{n+1}$ be the line in the direction $(v, 1)$ passing through $(y, 0)$, that is, $l^y_v = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : x = y + vt \}$. And for $0 < \delta \ll 1$, let $T^\omega_y(\delta) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \text{dist} (l^y_v, (x, t)) \leq \delta \}$. Suppose for some $v \in B(0, 1)$,

\[(2.27) \text{dist} (\mathbb{P}_p(\Pi), v) \sim 1 \text{ for all } p \in \Pi.\]

Then for any $y \in \mathbb{R}^n$, $T^\omega_y(\delta) \cap C(\Pi, \lambda, 1, \delta) \subset B(y_0, C\delta/\lambda)$ for some $y_0 \in \mathbb{R}^{n+1}$.

**Proof.** Since $C(\Pi, \lambda, \delta) \subset \bigcup_{k, 1 \geq 2^k \geq \lambda} C(\Pi, 2^{k-1}, 2^k, \delta)$, by re-scaling it is sufficient to show

\[(2.28) T^\omega_y(\delta) \cap C(\Pi, 1/2, 1, \delta) \subset B(y_0, C\delta)\]

for some $y_0 \in \mathbb{R}^{n+1}$. For simplicity we set $C(\Pi) = C(\Pi, 1/2, 1)$. We claim for all $q \in C(\Pi)$,

\[(2.29) \text{dist} (T_q(C(\Pi)), (v, 1)) \sim 1,\]

where $T_q(C(\Pi)) \subset \mathbb{R}^n \times \mathbb{R}$ is the tangent space of $C(\Pi)$ at $q$, namely, $(\mathbb{P}_q(C(\Pi)) - q)$. This means $l^y_v$ intersects $C(\Pi)$ transversally if it meets $C(\Pi)$. Hence, it is easy to see (2.28). This proves Lemma 2.7.

It remains to show our claim. Suppose for some $q = (sp, s) \in C(\Pi)$, $s \in [1/2, 1]$,

\[\text{dist} (T_q(C(\Pi)), (v, 1)) \leq \varepsilon\]
with small $\varepsilon > 0$ (much smaller than the implicit lower bound in (2.29)). Note that

$$T_{q}(C(\Pi)) = \text{span} \{(u,0) : u \in T_{p}(\Pi)\} + \text{span} \{(p,1)\}.$$ 

Hence, $|\langle v,1 \rangle - (c_{1}u_{1} + \cdots + c_{n}u_{n} + c_{n+1}p, c_{n+1})| \leq \varepsilon$ for some $c_{1}, \ldots, c_{n+1} \in \mathbb{R}$ and $u_{1}, \ldots, u_{n} \in T_{p}(\Pi)$. Since $|c_{n+1} - 1| \leq \varepsilon$, $|v - c_{1}u_{1} - \cdots - c_{n}u_{n} - p| \leq \varepsilon$. So dist $(P_{p}(\Pi), v) \leq 2\varepsilon$. This contradicts (2.27). □

**Proof of Lemma 2.7.** To begin with, we may assume the sets $\tilde{V}_{1}, \tilde{V}_{2}$ are small enough so that $\nabla \phi_{1}, \nabla \phi_{2}$ are diffeomorphisms on $\tilde{V}_{1}, \tilde{V}_{2}$, respectively, since $H\phi_{1}, H\phi_{2} \neq 0$.

Note that the major directions $(\nabla \phi_{1}(v), 1)$ of tubes $T_{w_{1}}$ with $w_{1} = (y_{1}, v_{1}) \in W^{2b}_{1}(\Pi_{\xi_{1}^{2}})$ are contained in the set

$$\{(\nabla \phi_{1}(v), 1) \in \mathbb{R}^{n} \times \mathbb{R} : v \in \Pi_{\xi_{1}^{2}, \xi_{2}^{2}} + O(R^{-1/2})\}.$$ 

Since $\xi \rightarrow \nabla \phi_{1}(\xi)$ is a diffeomorphism, $\nabla \phi_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}})$ is also a smooth surface. Let us define a conic surface $C_{\xi_{1}}^{\xi_{2}}$ by

$$C_{\xi_{1}}^{\xi_{2}} = \{(su, s) \in \mathbb{R}^{n} \times \mathbb{R} : u \in \nabla \phi_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}}), |s| \leq 2R\}.$$ 

Since $T_{w_{1}}$ with $w_{1} \in W^{2b}_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}})$ meets $R^{\delta}q_{0}$,

$$\bigcup_{w_{1} \in W^{2b}_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}})} T_{w_{1}} \subset C_{\xi_{1}}^{\xi_{2}}(R^{1/2+\delta}) + q_{0}$$

where $C_{\xi_{1}}^{\xi_{2}}(A) = C_{\xi_{1}}^{\xi_{2}} + O(A)$.

If $(q, w_{1}) \in S$, then $R^{\delta}q$ meets $T_{w_{1}}$. So, $q$ is contained in $C_{\xi_{1}}^{\xi_{2}}(R^{1/2+\delta}) + q_{0}$ if $(q, w_{1}) \in S$ for some $w_{1}$. Since dist $(q_{0}, q) \geq R^{1-\delta}$ if $(q, w_{1}) \in S$, it follows that

$$q \subset C_{\xi_{1}, \xi_{2}^{2}}(R^{1/2+\delta}, R^{1-\delta}, R, q_{0})$$

whenever $(q, w_{1}) \in S$ for some $w_{1}$ where

$$C_{\xi_{1}, \xi_{2}^{2}}(R^{1/2+\delta}, R^{1-\delta}, R, q_{0}) = C_{\xi_{1}}^{\xi_{2}}(R^{1/2+\delta}) \cap \{(x, t) : R^{1-\delta} \leq |t| \leq R\} + q_{0}.$$ 

Furthermore, $R^{\delta}q \cap T_{w_{2}} \neq \emptyset$ if $(q, w_{1}) \in S$. Therefore, we see

$$\bigcup_{q : (q, w_{1}) \in S} q \subset R^{\delta}T_{w_{2}} \cap C_{\xi_{1}}^{\xi_{2}^{2}}(R^{1/2+\delta}, R^{1-\delta}, R, q_{0})$$

where

$$R^{\delta}T_{w_{2}} = \{(x, t) : |t| \leq R_{1}, |x - (y_{2} + t\nabla \phi_{1}(v_{2}))| \leq CR^{1/2+\delta}\}.$$ 

Now we claim that

$$R^{\delta}T_{w_{2}} \cap C_{\xi_{1}}^{\xi_{2}^{2}}(R^{1/2+\delta}, R^{1-\delta}, R, q_{0}) \subset B(y_{0}, R^{1/2+\delta})$$

for some $c > 0$, $y_{0} \in \mathbb{R}^{n+1}$. By rescaling it is sufficient to show that for $\delta^{c} \leq \lambda \leq 1$, $0 < \epsilon \ll 1$,

$$T_{y_{2}}^{\nabla \phi_{2}(v_{2})} (\delta) \cap C(\nabla \phi_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}}), \lambda, 1, \delta) \subset B(y_{0}, C\delta/\lambda)$$

for some $y_{0}$ where $w_{2} = (y_{2}, v_{2}) \in W_{2}$.

The normal vector of the surface $\nabla \phi_{1}(\Pi_{\xi_{1}^{2}, \xi_{2}^{2}})$ at $\nabla \phi_{1}(\xi_{1}^{2})$ is parallel to

$$N_{\xi_{1}}(\xi_{2}^{2}) = (\phi_{1})'_{\xi_{1}}(\xi_{2}^{2})^{-1} \nabla \phi_{2}(\xi_{1}^{2} + \xi_{2}^{2} - \xi_{1}) - \nabla \phi_{1}(\xi_{1}^{2})]$$
with \( \xi' + \xi'' - \xi_1 \in \bar{V}_2 \) because \( \nabla \phi_1(\Pi_{\xi_1,\xi_2}) \) is contained in
\[
\{ \eta \in \nabla \phi_1(\bar{V}_1) : \phi_1(\xi_1) + \phi_2(\nabla \phi_1^{-1}(\eta) + \xi' - \xi_1) - \phi_1(\nabla \phi_1^{-1}(\eta)) - \phi_2(\xi'' - \xi_1) = 0 \}.
\]
Hence, the tangent plane \( P_{\xi_1,\xi_2}(\xi_1') \) of \( \nabla \phi_1(\Pi_{\xi_1,\xi_2}) \) at \( \nabla \phi_1(\xi_1) \) is given by the relation \( \langle N_{\xi_1,\xi_2}(\xi_1'), x - \nabla \phi_1(\xi_1) \rangle = 0 \). So, \( \text{dist}(P_{\xi_1,\xi_2}(\xi_1'), y) \approx |\langle N_{\xi_1,\xi_2}(\xi_1'), y - \nabla \phi_1(\xi_1') \rangle| \) since \( |N_{\xi_1,\xi_2}(\xi_1')| \sim 1 \). Therefore, (1.6) implies for all \( \xi_1, \xi_1' \in \bar{V}_1, \xi_2, \xi_2' \in \bar{V}_2 \)
\[
\text{dist}(P_{\xi_1,\xi_2}(\xi_1'), \nabla \phi_2(\xi_2')) \sim 1.
\]
In particular, \( \text{dist}(P_{\xi_1,\xi_2}(\xi_1'), \nabla \phi_2(v_2)) \sim 1 \) for all \( \xi_1' \in \Pi_{\xi_1,\xi_2} \). Using Lemma 2.7 we see (2.33) and hence (2.32).

Since \( q \in Q \) are \( R^{1/2} \)-separated cubes, from (2.32) and (2.33) it follows that
\[
\# \{ q \in Q : (q, w_1) \in S \text{ for some } w_1 \} = O(R^{\delta}).
\]
Since \( \text{dist}(q, q_0) \geq R^{1-\delta} \) if \( (q, w_1) \in S, \) for each fixed \( q \) the number of possible \( T_{w_1} \)
passing through both \( R^\delta q \) and \( R^\delta q_0 \) is also \( O(R^{\delta}) \), because the major directions
\( (\nabla \phi_1(v), 1), v \in V_1 \) of tubes are separated by \( \sim R^{-1/2} \). This and (2.34) give \#S = O(R^{\delta}).

3. PROOF OF THEOREM 1.3

For the proof of Theorem 1.3 we do not rely an orthogonality lemma such as
Lemma 2.1 in [21], which was crucial to prove the sharp bilinear restriction estimate
for the cone. Instead, we prove Theorem 1.3 by making some additional observation
for conic surfaces upon the same lines of argument as in the previous section.

The arguments in Section 2 (the proof of Theorem 1.1) can be repeated with
\[
\phi_1 = \phi_2 = \langle \eta, N\eta/\rho \rangle, \quad V_i = \bar{V}_i \times [1, 2], \quad i = 1, 2,
\]
which define the extension operators \( \mathcal{E}_1, \mathcal{E}_2 \). From these we define \( V_i, W_i, T_{w_i} \)
by following the same procedure as in the previous section. We keep the same
notations \( W_i(q), \mathcal{Q}(\mu_1, \mu_2), W_i^{\text{trib}}(q_0, \mu_1, \mu_2, \lambda_1), \ldots \), etc.

Since \( \nabla_{\eta,\rho} \phi_1 = \nabla_{\eta,\rho} \phi_2 = (2N\eta/\rho, -\langle \eta, N\eta/\rho \rangle^{2}) \), the major directions \( (\nabla \phi_1, 1) \)
the tubes \( T_{w_i}, w_i \in W_i \) are contained in the set
\[
\{ (2N\theta, -\langle \theta, N\theta \rangle, 1) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} : \theta \in \mathcal{O}_\xi \}.
\]
The transversality between tubes \( T_{w_1} \) and \( T_{w_2}, w_1 \in W_1, w_2 \in W_2 \) is easy to see from (1.11) which implies \( \text{dist}(\mathcal{O}_1, \mathcal{O}_2) \sim 1 \). The set \( \Pi_{\xi_1,\xi_2} \) given by (2.17) is also
a smooth \( n-1 \)-dimensional hypersurface because \( \text{dist}(\nabla \phi_1(V_1), \nabla \phi_2(V_2)) \sim 1 \) (see
(2.18)). Without difficulty one can see that all the arguments in Section 2 work
except Lemma 2.6. Therefore the proof of Theorem 1.3 is to be completed if one
shows the following combinatorial estimate (see Lemma 2.6).

\textbf{Lemma 3.1.} Let \( W_i^{\text{trib}}(\Pi_{\xi_1,\xi_2}) \) be defined by the same way as in Section 2 with
\( \phi_1 = \phi_2 = \langle \eta, N\eta/\rho \rangle, V_i = \bar{V}_i \times [1, 2], i = 1, 2 \). For each \( w_i \in W_2 \), set
\[
S = \{ (q, w_1) \in \mathcal{Q}(\mu_1, \mu_2) \times W_i^{\text{trib}}(\Pi_{\xi_1,\xi_2}) : \}
\[
R^\delta q \cap T_{w_1} \neq \emptyset, \text{dist}(q_0, q) \geq R^{1-\delta}, R^\delta q \cap T_{w_2} \neq \emptyset \}.
\]
Then, \( \#S \lesssim O(R^{\delta}) \).
Unlike the previous case, $\nabla \phi_1$ and $\nabla \phi_2$ are no longer diffeomorphisms. So, the argument for the proof of Lemma 2.6 is not directly applicable. To prove Lemma 3.1, we need to modify the proof of Lemma 2.6.

**Proof of Lemma 3.1.** As before, we claim for some $c > 0$,

\[(3.2) \quad \#\{q \in Q : (q, w_1) \in S \text{ for some } w_1\} = O(R^{\delta}).\]

We define a conic surface $C_1$ by

$C_1 = \{s(2N\theta, -(\theta, N\theta), 1) \in \mathbb{R}^{n-1} \times \mathbb{R} : \theta \in \Theta_1, |s| \leq 2R\}.$

Since $T_{w_1}$ meets with $R^\delta q_0$ if $(q, w_1) \in S$, from (3.1) it follows that

\[\bigcup_{w_1 \in W^{\delta}(\Pi_1, c_2)} T_{w_1} \subset C_1(R^{1/2+\delta}) + q_0\]

where $C_1(A) = C_1 + O(A)$. If $(q, w_1) \in S$, $R^\delta q \cap T_{w_2} \neq \emptyset$ and $\text{dist} \ (q, q_0) \geq R^{1-\delta}$.

Hence,

\[(3.3) \quad \bigcup_{q(w_2) \in S \text{ for some } w_1} q \subset R^\delta T_{w_2} \cap C_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0),\]

where $R^\delta T_{w_2}$ is given by (2.21) and

$C_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0) = C_1(R^{1/2+\delta}) \cap \{(x, t) : R^{1-\delta} \leq |t| \leq R\} + q_0.$

Let us set

$\Pi = \{(2N\theta, -\theta \cdot N\theta) : \theta \in \Theta_1\}.$

Then the normal vector $N_p(\Pi)$ of $\Pi$ at $p = (2N\theta_1, -\theta_1 \cdot N\theta_1) \in \Pi$ is parallel to $(\theta_1, 1)$ since $\Pi$ is parameterized as $\rho + \eta/2 \cdot N\eta/2 = 0$. (Here we use $N^2 = \text{id}$.) So, the tangent plane $P_p(\Pi)$ is given by

$\{(s, t) \in \mathbb{R}^{n-1} \times \mathbb{R} : ((\theta_1, 1), (s - 2N\theta_1, t + \theta_1 \cdot N\theta_1)) = 0\}.$

From (3.1), $\nabla \phi_2(v_2) = (2N\theta_2, -\theta_2 \cdot N\theta_2)$ for some $\theta_2 \in \Theta_2$. Therefore,

\[(3.4) \quad \text{dist} \ (P_p(\Pi), \nabla \phi_2(v_2)) \sim |(\theta_1, 1) \cdot ((2N\theta_2 - 2N\theta_1), -\theta_2 \cdot N\theta_2 + \theta_1 \cdot N\theta_1)|
\]

\[\quad = |(\theta_1 - \theta_2) \cdot N(\theta_1 - \theta_2)| \sim 1\]

for all $\theta_1 \in \Theta_1$ by the hypothesis (1.11). By the same argument in the proof of Lemma 2.6 (rescaling and applying Lemma 2.7 to $T_{y_2}^{\nabla \phi_2(v_2)}(\delta), C_1(\Pi, \lambda, 1, \delta)$ as before), we see

\[R^\delta T_{w_2} \cap C_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0) \subset B(y_0, R^{1/2+\delta})\]

for some $c > 0, y_0$. Therefore, we get (3.2) from (3.3) and the above because $q$ are $R^{1/2}$-separated.

Now we claim that for each fixed $q$ there are $O(R^{\delta})$-wises for which $(q, w_1) \in S$. This and (3.2) prove Lemma 3.1. It remains to show the claim.

Fix a $q$ and let $c(q), c(q_0)$ denote the the centers of cubes $q, q_0$, respectively. Let $u \in \mathbb{R}^n$ be given by

\[(u, 1)/|c(q_0) - c(q)| - c(q)\]
If \((q, w_1) \in S\) with \(w_1 = (y_1, v_1)\), then \(\text{dist } (q_0, q) \geq R^{1-\delta}\) and \(T_{w_1}\) passes through both \(R^d q_0\) and \(R^d q\). So, the major direction \((\phi_1(v_1), 1)\) of \(T_{w_1}\) is contained in the \(O(R^{-1+2+\delta})\)-neighborhood of \((u, 1)\). Therefore if \((q, (y_1, v_1)) \in S\), then
\[
\nabla \phi_1 (v_1) = u + O(R^{-1+2+\delta}).
\]
However, there may be possibly more than \(O(R^{\delta})\)-v1’s satisfying (3.5) since \(\nabla \phi_1\) is not a diffeomorphism.

To get around this, we make three simple observations: (i) if \((q, w_1) \in S\) with \(w_1 = (y_1, v_1)\), then \(v_1\) is contained in \(\Pi_{\xi_1, \xi_2} (R^{-1/2}) = \Pi_{\xi_1, \xi_2} + O(R^{-1/2})\) (see (2.24)), (ii) \(\nabla \phi_1(\eta, \rho) = \nabla \phi_1(\eta', \rho')\) if and only if \(\eta/\rho = \eta'/\rho'\) and (iii) \(V_1\) and \(V_2\) are assumed to be as small as we like since \(E_1\) are written as finite sum of operators with small \(V_1, V_2\).

Let \(\bar{\theta} \in \Theta_1\) be the point satisfying \(2N\bar{\theta}, -(\bar{\theta}, N\bar{\theta}) = u\). Then the \(v_1\)’s satisfying (3.5) are contained in
\[
L(\bar{\theta}, R^{-1+2+\delta}) = \{ \rho(\bar{\theta}, 1) \in \tilde{V}_1 : 1 \leq \rho \leq 2 \} + O(R^{-1+2+\delta})
\]
since (ii) means \(\nabla \phi_1\) fails to be one to one only along the line \(\rho(\bar{\theta}, 1), 1 \leq \rho \leq 2\), for \(\bar{\theta} \in \Theta_1\). If \((\bar{\theta}, 1)\) is transversal to \(\Pi_{\xi_1, \xi_2}\) (namely, the angle \((\bar{\theta}, 1)\) between all the tangent plane of \(\Pi_{\xi_1, \xi_2}\) is bounded away from zero), then \(L(\bar{\theta}, R^{-1+2+\delta}) \cap \Pi_{\xi_1, \xi_2} (R^{-1/2})\) is contained in a ball of radius \(O(R^{-1+2+\delta})\). Therefore, from (i) and (iii), we see that if all the null direction \{\((\theta_1, 1)\)\}_{\theta_1 \in \Theta_1}\) for \(\nabla \phi_1\) is transversal to \(\Pi_{\xi_1, \xi_2}\), there are \(O(R^{\delta})\)-v1 satisfying (3.5) since \(v_1 \in V_1\) are \((R^{-1/2})\)-separated.

We recall from (2.17) that the normal vector of \(\Pi_{\xi_1, \xi_2}\) at \((\eta_1, \rho_1)\) with \(\eta_1 = \theta_1 \rho_1, \theta_1 \in \Theta_1\) is parallel to \(N_{\theta_1, \theta_2} = \nabla \phi_1 (\eta_1, \rho_1) - \nabla \phi_2 (\eta_2, \rho_2) = (2N\theta_1 - 2N\theta_2, -\theta_1 \cdot N\theta_1 + \theta_2 \cdot N\theta_2)\) for some \(\theta_2 \in \Theta_2\) (see (2.18)) where \((\eta_2, \rho_2) \in V_2\) and \(\eta_2 = \theta_2 \rho_2\). Then the null direction \((\theta'_1, 1), \theta'_1 \in \Theta_1\) at \((\eta, \rho), \eta = \theta' \rho\) is transversal to \(\Pi_{\xi_1, \xi_2}\) if and only if
\[
|N_{\theta_1, \theta_2} : (\theta'_1, 1)| \sim 1
\]
because \(|N_{\theta_1, \theta_2}| \sim 1\). Since \(V_1\) and \(V_2\) are assumed to be as small as pleased, we may also assume \(\Theta_1, \Theta_2\) to be small enough. So, \(|N_{\theta_1, \theta_2} : (\theta'_1, 1)| \sim |N_{\theta_1, \theta_2} : (\theta_1, 1)| = |(\theta_1 - \theta_2) \cdot N(\theta_1 - \theta_2)|\). By the hypothesis (1.11), (3.6) is valid for all \(\theta_1, \theta'_1 \in \Theta_1, \theta_2 \in \Theta_2\). This shows the transversality between the null direction \{\((\theta_1, 1)\)\}_{\theta_1 \in \Theta_1}\) and \(\Pi_{\xi_1, \xi_2}\) for all \(\xi_1 \in \tilde{V}_1, \xi_2^2 \in \tilde{V}_2\).

Therefore, from the discussion in the above we see that there are at most \(O(R^{\delta})\)-v1 such that \((q, y_1, v_1) \in S\) for some \(y_1\). Once the direction \(v_1\) is determined, obviously there are at most \(O(R^{\delta})\)-y1 because \(T_{w_1}\) passes through \(R^d q_0\). This proves our claim.

4. APPLICATION TO RESTRICTION ESTIMATES

In this section we prove Corollary (1.2) and Corollary (1.4). These will be derived from Theorems (1.1) and (1.3) by adapting the argument in [17].

4.1. Proof of Corollary (1.2) By making the change of variables \((\eta_1 - \eta_2, \eta_1 + \eta_2) \rightarrow (\xi_1, \xi_2)\), instead of \(\int f d\sigma_\mathcal{H}\), we may consider
\[
Ef(x, t) = \int_{[-2, 2]^2} e^{2\pi i (x \cdot \xi + 4\xi \cdot \xi)} f(\xi) d\xi,
\]
To obtain linear estimate from Theorem 1.1 we begin by stating some bilinear restriction estimates for $E$. From Theorem 1.1 and interpolation (with the trivial estimates, e.g. $\|Ef Eg\|_\infty \leq C\|f\|_1\|g\|_1$) one can easily see the following:

Let $V_1, V_2$ be subcubes of $[-2,2]^2$. Suppose for $(a,b) \in V_1 \times V_2$,

\begin{equation}
\|(a_1 - b_1)(a_2 - b_2)\| \sim 1
\end{equation}

(see (1.9)). Then if $supp f \subset V_1$ and $supp g \subset V_2$, there is a constant $C$ such that

\begin{equation}
\|Ef Eg\|_{L^p([\mathbb{R}^2 \times \mathbb{R}])} \leq C\|f\|_{L^p([\mathbb{R}^2])}\|g\|_{L^p([\mathbb{R}^2])}
\end{equation}

provided $(1/p, 1/q)$ is contained in the interior of the quadrangle $Q$ with vertices $(0,0), (0,3/5), (1/2,3/5), (1,0)$.

**Proposition 4.1.** Let $V_1, V_2$ be subcubes of $[-2,2]^2$. If $dist (V_1, V_2) \sim 1$, supp $f \subset V_1$ and supp $g \subset V_2$, then $\|Ef Eg\|_p \leq C\|f\|_p\|g\|_p$ provided $q > 5/3$ and $1/p + 1/q < 1$.

To prove Corollary 1.2 we need to remove the separation condition $dist (V_1, V_2) \sim 1$ from the above. It can be done by the argument in 17 without modification. However, readers have no difficulty in obtaining Corollary 1.2 by making use of (4.14) and following the argument in the proof of Proposition 4.1 below.

**Proof of Proposition 4.1.** Decomposing both $V_1$ and $V_2$ into small subcubes, we may assume that $V_1 = Q(c_1, \epsilon)$ and $V_2 = Q(c_2, \epsilon)$ with $|c_1 - c_2| \sim 1$, $0 < \epsilon < 1$, where $Q(c,r)$ is the cube centered at $c$ with side length $r$. If $(c_1 - c_2)/|c_1 - c_2|$ is away from $\pm c_1$ and $\pm c_2$, then Proposition 4.1 follows from Theorem 1.1 because (1.11) is satisfied. So we may also assume $(c_1 - c_2)/|c_1 - c_2|$ is close to $\pm c_1$ or $\pm c_2$. Hence by harmless affine transformation it is sufficient for Proposition 4.1 to consider the case $V_1 = Q(-c_1, \epsilon)$, $V_2 = Q(c_1, \epsilon)$.

Let $I = [-2,2]$. We use the decomposition in 17. For each $j \geq 1$, we dyadically decompose $I$ into dyadic subintervals $I_k^j$ of side length $2^{-(j+1)}$. We say $I_k^j \approx I_k^{j'}$ to mean that $I_k^j, I_k^{j'}$ are not adjacent but have adjacent parent intervals of length $2^{-j}$. So, $dist (I_k^j, I_k^{j'}) \sim 2^{-j}$ if $I_k^j \approx I_k^{j'}$. By a Whitney decomposition of $I \times I$ away from its diagonal $D$, we have

\begin{equation}
I \times I \setminus D = \bigcup_{j \geq 1} \bigcup_{I_k^j \approx I_k^{j'}} I_k^j \times I_k^{j'}.
\end{equation}

Let us set

\[ f_k^j(\xi) = \chi_{I_k^j}(\xi_2)f(\xi), \quad g_k^j(\xi) = \chi_{I_k^j}(\xi_2)g(\xi). \]

Since $\sum_{j \geq 1} \sum_{I_k^j \approx I_k^{j'}} \chi_{I_k^j} \chi_{I_k^{j'}} = 1$ almost everywhere in $I \times I$,

\begin{equation}
Ef(x)Eg(x) = \sum_{j \geq 1} B_j(f, g)(x),
\end{equation}

where

\[ B_j(f, g)(x) = \sum_{I_k^j \approx I_k^{j'}} Ef_k^j(x)Eg_k^j(x). \]

If $I_k^j \approx I_k^{j'}$, the Fourier support of $Ef_k^jEg_k^{j'}$ is contained in the rectangle

\[ \{(\xi_1, \xi_2, \tau) : |\xi_2 - c_k^j| \leq 2^j, |\xi_1| \leq C, |\tau| \leq C\}. \]
where \( c_k^j \) is the center of \( I_k^j \). Hence one can easily see that for fixed \( j \) the Fourier supports of \( \{Ef_k^jEg_k^j\}_{k,k':I_k^j\approx I_k'^j} \) are contained in boundedly (at most \( 2^j \)) overlapping rectangles. Using Lemma 4.2 below, we have for \( 1 \leq q \leq \infty \),

\[
\|B_j(f,g)\|_q \leq C \left( \sum_{k,k':I_k^j\approx I_k'^j} \|Ef_k^jEg_k^j\|_{q^*}^{q^*} \right)^{q^*}.
\]

**Lemma 4.2** ([13], Lemma 7.1). Let \( \{R_k\} \) be a collection of rectangles in frequency space such that the dilates \( \{2R_k\} \) are essentially disjoint, and suppose that \( \{F_k\} \) are a collection of functions whose Fourier supports are contained in \( \{R_k\} \). Then for \( 1 \leq p \leq \infty \) we have

\[
(\sum_k \|F_k\|_p^{p^*})^{1/p^*} \lesssim \|F\|_p \lesssim (\sum_k \|F_k\|_p^{p^*})^{1/p^*},
\]

where \( p_\ast = \min(p, p') \) and \( p^\ast = \max(p, p') \).

Now we show that if \( I_k^j \approx I_k'^j \) and \( (1/p, 1/q) \) is contained in the interior of \( Q \),

\[
\|Ef_k^jEg_k^j\|_q \leq C 2^{-2j+2^k + 2^j q} \|f_k^j\|_p \|g_k^j\|_p.
\]

By translation we may assume \( I_k^j = [-2^{-j}, -2^{-j+1}] \) and \( I_k'^j = [2^{-j-1}, 2^{-j}] \), or \( I_k^j = [2^{-j-1}, 2^{-j}] \) and \( I_k'^j = [-2^{-j}, -2^{-j+1}] \). It suffices to consider the first case because the arguments for both cases are similar.

Re-scaling \( (\xi_1, \xi_2) \to (\xi_1, 2^{-j} \xi_2) \) gives

\[
[Ef_k^jEg_k^j](x, t) = 2^{-2j}[E(f_k^j)_jE(g_k^j)_j](x_1, 2^{-j}x_2, 2^{-j}t),
\]

where \( (f_k^j)_j = f_k^j(\xi_1, 2^{-j} \xi_2) \) and \( (g_k^j)_j = g_k^j(\xi_1, 2^{-j} \xi_2) \). Then \( (f_k^j)_j \) and \( (g_k^j)_j \) are supported in \( Q(-1, 1, \frac{1}{2}) \), \( Q((1, 1), \frac{1}{2}) \), respectively, because \( f_k^j \), \( g_k^j \), are supported in \([1-\epsilon, 1+\epsilon] \times [-2^{-j}, -2^{-j+1}]\), \([1-\epsilon, 1+\epsilon] \times [2^{-j-1}, 2^{-j}]\), respectively. Since (4.4) holds for all \( (a, b) \in Q(-1, 1, \frac{1}{2}) \times Q((1, 1), \frac{1}{2}) \), (4.2) and rescaling give (4.6).

Now putting (4.6) into (4.5), we have

\[
\|B_j(f,g)\|_q \leq C 2^{-2j+2^k + 2^j q} \left( \sum_{I_k^j \approx I_k'^j} \|f_k^j\|_p \|g_k^j\|_p \right)^{1/q^*}.
\]

Since the number of \( I_k^j \) with \( I_k^j \approx I_k'^j \) is \( O(1) \), the left-hand side of the above is bounded by \( C 2^{-2j+2^k + 2^j q} \left( \sum_k \|f_k^j\|_p^{2q^*} \right)^{1/2q^*} \left( \sum_k \|g_k^j\|_p^{2q^*} \right)^{1/2q^*} \). Since \( I_k^j \) are disjoint dyadic intervals, \( \sum_k \|f_k^j\|_p^{2q^*} \) is contained in the interior of \( Q \) and \( 2q^* \geq p \). Hence we get

\[
\|B_j(f,g)\|_q \leq C 2^{-2j+2^k + 2^j q} \|f\|_p \|g\|_p
\]

provided \( (1/p, 1/q) \) is contained in the interior of \( Q \) and \( 2q^* \geq p \). Since \( f, g \) are supported in \([-2, 2]^2 \), using Hölder’s inequality, from (4.8) one can easily see that

\[
\|B_j(f,g)\|_q \leq C 2^{-j} \|f\|_p \|g\|_p
\]

for some \( \epsilon > 0 \) if \( q > 5/3 \) and \( 1/p + 1/q < 1 \). Therefore, by (4.4) summation in \( j \) completes the proof. \( \square \)
4.2. Proof of Corollary 4.4 As before, instead of \( \int fd\sigma_C \) it suffices to consider
\[
\mathcal{E}f(x, t) = \int_1^2 \int_{[-2, 2]^2} e^{2\pi i (x' - x) + t\xi_1 / \rho} f(\xi, \rho) d\xi d\rho, \quad x' \in \mathbb{R}^2,
\]
by a linear transformation. Using Theorem 1.3 it is easy to see the following:

Let \( \overline{V}_1, \overline{V}_2 \) be subcubes of \([-1, 1]^2 \). Suppose (4.1) holds for all \( \alpha \in \Theta_1 \) and \( \beta \in \Theta_2 \) (see Theorem 1.3 for the definitions of \( \Theta_1, \Theta_2 \)). Then if \( f \subset \overline{V}_1 \times [1, 2] \) and \( g \subset \overline{V}_2 \times [1, 2] \), there is a constant \( C \) such that
\[
(4.9) \quad \| \mathcal{E}f \mathcal{E}g \|_{L^p(\mathbb{R}^3 \times \mathbb{R})} \leq C \| f \|_{L^p(\mathbb{R}^3)} \| g \|_{L^p(\mathbb{R}^3)}
\]
provided \((1/p, 1/q)\) is contained in the interior of the quadrangle \( Q_0 \) with vertices \((0, 0), (1, 0) (1/2, 2/3), (0, 2/3)\).

Proposition 4.3. Let \( \overline{V}_1, \overline{V}_2 \) be subcubes of \([-1, 1]^2 \). If \( \dist (\Theta_1, \Theta_2) \sim 1 \), then for \( q > \frac{3}{2} \) and \( 1/p + 1/q < 1 \), \( \| \mathcal{E}_1 f \mathcal{E}_2 g \|_q \leq C \| f \|_p \| g \|_p \).

Proof of Proposition 4.3. Decomposing both \( \overline{V}_1 \times [1, 2] \) and \( \overline{V}_2 \times [1, 2] \) into small subcubes, we may assume \( \Theta_1 = Q(c_1, \epsilon) \) and \( \Theta_2 = Q(c_2, \epsilon) \) with \( |c_1 - c_2| \sim 1, 0 < \epsilon \ll 1 \). As we did in the proof of Proposition 4.1, by Theorem 1.3 we may also assume \( |c_1 - c_2|/|c_1 - c_2| \) is close to \( \pm \epsilon_1 \) or \( \pm \epsilon_2 \).

Observe that the change of variables \( (\xi, \rho) \rightarrow (\xi + c_\rho, \rho) \) for \( \mathcal{E}f \) gives
\[
(4.10) \quad \mathcal{E}f(x, t) = \int_1^2 \int \frac{e^{2\pi i (L(x, t) + t\xi_2 / \rho)} \chi_{\Theta_1} \left( \frac{\xi}{\rho} \right) f(\xi + \rho, \rho) d\xi d\rho},
\]
where \( L(x, t) = (x_1 + tc_2, x_2 + tc_1, x_3 + x_1c_1 + x_2c_2 + tc_1c_2) \). Since \( \det L = 1 \), we can move \( \Theta_1, \Theta_2 \) to \( \Theta_1 - c_1, \Theta_2 - c_2 \), respectively, without affecting the estimate (4.9).

Hence, by an linear transform it is sufficient consider the case \( \Theta_1 = Q(-c_1, \epsilon), \Theta_2 = Q(c_1, \epsilon) \).

Using the decomposition (4.3), we set \( f_{k_1}(\xi, \rho) = \chi_{I_{k_1}^c}(\xi_2 / \rho) f(\xi, \rho), g_{k_1}(\xi, \rho) = \chi_{I_{k_1}^c}(\xi_2 / \rho) g(\xi, \rho) \). Then, it follows that
\[
\mathcal{E}f(x) \mathcal{E}g(x) = \sum_{j \geq 1} \mathcal{B}_j(f, g)(x),
\]
where \( \mathcal{B}_j(f, g)(x) = \sum_{k, k_{1}: I_{k_1}^c \approx I_{k_1}^c} \mathcal{E}_{I_{k_1}^c} f_k^j(x) \mathcal{E}_{I_{k_1}^c} g_k^j(x) \). Hence it is sufficient to show that if \( q > \frac{3}{2} \) and \( 1/p + 1/q < 1 \), for some \( \epsilon > 0 \),
\[
(4.11) \quad \| \mathcal{B}_j(f, g) \|_q \leq C \epsilon^{\epsilon} \| f \|_p \| g \|_q.
\]

If \( I_{k_1} \approx I_{k_1} \), the Fourier support of \( \mathcal{E}_{I_{k_1}^c} f_k^j \mathcal{E}_{I_{k_1}^c} g_k^j \) is contained in the set
\[
\{(\xi_1, \xi_2, \rho, \tau) : |\xi_2 - c_0 \rho| \leq 2 \epsilon^{3-j}, 1 \leq \rho < 2, |\xi_1| \leq C, |\tau| \leq C \},
\]
where \( c_0 \) is the center of the smallest interval containing both \( I_{k_1}^c, I_{k_1}^c \). So we see that the Fourier supports of \( \{\mathcal{E}_{I_{k_1}^c} f_k^j \mathcal{E}_{I_{k_1}^c} g_k^j \} \) are contained in boundedly (at most \( 2^6 \)) overlapping rectangles. Hence, using Lemma 4.2 we have for \( 1 \leq q \leq \infty \),
\[
(4.12) \quad \| \mathcal{B}_j(f, g) \|_q \leq C \left( \sum_{I_{k_1}^c \approx I_{k_1}^c} \| \mathcal{E}_{I_{k_1}^c} f_k^j \mathcal{E}_{I_{k_1}^c} g_k^j \|_q^2 \right)^{1/q}.
\]
Now we claim that if \( I_k^1 \approx I_k^2 \) and \((1/p, 1/q)\) is contained in the interior of \( Q_0 \),
\[
\| \mathcal{E} f_k \mathcal{E} g_k' \|_q \leq C 2^{-2j+\frac{3j}{q}+\frac{2j}{p}} \| f_k \|_p \| g_k' \|_p.
\]
(4.13)

Once this is established, by the same argument used before (see (4.7)), (4.12) and (4.13) give \( \| \mathcal{B}_j(f, g) \|_q \leq C 2^{-2j+\frac{3j}{q}+\frac{2j}{p}} \| f \|_p \| g \|_p \) provided \((1/p, 1/q)\) is contained in the interior of \( Q_0 \) and \( 2q_\ast \geq p \). From this, using Hölder’s inequality, one can easily see (4.11).

Finally we show (4.13). By the linear transformation \( \xi_2 \to \xi_2 + c_0 \rho \) for both \( \mathcal{E} f_k \) and \( \mathcal{E} g_k' \) (see (4.10)), we may assume \( I_k^1 = [-2^{-j}, -2^{-j-1}] \) and \( I_k^2 = [2^{-j-1}, 2^{-j}] \), or \( I_k^1 = [2^{-j-1}, 2^{-j}] \) and \( I_k^2 = [-2^{-j}, -2^{-j-1}] \). It suffices to consider the first case. By re-scaling \((\xi_1, \xi_2) \to (\xi_1, 2^{-j} \xi_2)\) we see that
\[
\mathcal{E} f_k \mathcal{E} g_k'(x, t) = 2^{-2j} \mathcal{E} (f_k(x), g_k'(t))(x_1, 2^{-j} x_2, x_3, 2^{-j} t),
\]
where \((f_k(x), g_k'(t))(\xi_1, \xi_2, \rho) = f_k(\xi_1, 2^{-j} \xi_2, \rho)\) and \((g_k'(t))(\xi_1, \xi_2, \rho) = g_k'(\xi_1, 2^{-j} \xi_2, \rho)\). Then \((f_k(x)), (g_k'(t))\) are supported in
\[
\begin{align*}
\mathcal{V}_1 &= \{ (\xi, \rho) : \xi/\rho \in [-1-\epsilon, -1+\epsilon] \times [-1, -1/2], 1 \leq \rho \leq 2 \}, \\
\mathcal{V}_2 &= \{ (\xi, \rho) : \xi/\rho \in [-1-\epsilon, 1+\epsilon] \times [1/2, 1], 1 \leq \rho \leq 2 \},
\end{align*}
\]
respectively. So, (4.11) holds for all \( a \in \Theta_1 \) and \( b \in \Theta_2 \). Therefore by (4.11) and rescaling we get (4.13).

To finish the proof of Corollary 1.4 we need to remove the separation condition \( \text{dist} (\Theta_1, \Theta_2) \sim 1 \) from Proposition 1.3. As before, we decompose \([-2, 2]^2 \times [-2, 2]^2\) away from its diagonal \( D \) (see (4.3)) by a Whitney decomposition. Ignoring some harmless measure zero set, we have
\[
[-2, 2]^2 \times [-2, 2]^2 \setminus D = \bigcup_{j \geq 1} Q^j_k \times Q^j_k',
\]
where \( Q^j_k \) are dyadic squares of side length \( 2^{-j+1} \) and \( Q^j_k \approx Q^j_k' \) means that \( Q^j_k \), \( Q^j_k' \) are not adjacent but have adjacent parent squares of side length \( 2^{-j} \).

Set
\[
\tilde{\mathcal{B}}_j(f, g)(x) = \sum_{k, k': Q^j_k \approx Q^j_k'} \mathcal{E} [f^j_k(x)] \mathcal{E} [g^j_k'](x),
\]
where \([f^j_k(x)](\xi, \rho) = \chi_{Q^j_k}(\xi/\rho) f(\xi, \rho)\). Since \( \mathcal{E} f \mathcal{E} f = \sum_{j \geq 1} \tilde{\mathcal{B}}_j(f, f) \) from (4.14), it follows that \( \| E f \|_q^2 \leq C \sum_{j \geq 1} \| \tilde{\mathcal{B}}_j(f, f) \|_{q/2} \). Hence, it is sufficient to show that if \( q > 3 \) and \( 1/p + 2/q < 1 \), for some \( \epsilon > 0 \)
\[
\| \tilde{\mathcal{B}}_j(f, g) \|_{q/2} \leq C 2^{-\epsilon j} \| f \|_p \| g \|_p.
\]
(4.15)

By repeating the argument used in the proofs of Proposition 1.1, 1.3 and using the orthogonality among \( \{ \mathcal{E} [f^j_k] \mathcal{E} [g^j_k'] \} \) \( Q^j_k \approx Q^j_k' \) (see (4.12)), for (4.15) it is sufficient to show that if \( Q^j_k \approx Q^j_k' \), for \( q > \frac{3}{2} \) and \( 1/p + 1/q < 1 \),
\[
\| \mathcal{E} [f^j_k] \mathcal{E} [g^j_k'] \|_q \leq C 2^{-4j+\frac{3j}{q}+\frac{2j}{p}} \| f_k \|_p \| g_k' \|_p.
\]
(4.16)
By linear transformation $Q^j_k, Q'^j_k$, can be assumed to be contained in $Q(0, 2^{1-\delta})$ (see (4.10)). Re-scaling $\xi \to 2^{-j} \xi$, the squares $Q^j_k, Q'^j_k$ with $Q^j_k \approx Q'^j_k$ are moved to cubes in $[-2, 2]^2$ with distance $\sim 1$. Therefore by Proposition 4.3 (following the same lines of argument as in the proof of Proposition 4.3), we get (4.16).

References


