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

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

holds for all $f \in C_c^\infty(S)$. 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 ϕ_1, ϕ_2 be smooth functions on $[-1, 1]^n$. For $i = 1, 2$, let us define extension (adjoint of restriction) operators by

$$(1.1) \qquad E_i f(x, t) = \int_{V_i} e^{2\pi i(x \cdot \xi - t\phi_i(\xi))} f(\xi) d\xi, \quad (x, t) \in \mathbb{R}^n \times \mathbb{R},$$

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

$$(1.2) \qquad \|E_1 f \cdot E_2 g\|_{L^q(\mathbb{R}^{n+1})} \leq C\|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^p(\mathbb{R}^n)}.$$

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

$$(1.3) \qquad \|E_1 f \cdot E_2 g\|_{L^q(\mathbb{R}^{n+1})} \le C\|f\|_{L^2(\mathbb{R}^n)}\|g\|_{L^2(\mathbb{R}^n)}$$

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 $q \ge \frac{n+3}{n+1}$ under the condition

$$(1.4) \qquad \operatorname{dist}\left(\nabla\phi_1(V_1), \nabla\phi_2(V_2)\right) \sim 1.$$

Here $A \sim B$ $(A, B > 0)$ means $C^{-1}A \le B \le CA$ for some $C > 0$. In [15, 17], a systematical 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 ϕ. Our first result is the following.

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

$$(1.5) \qquad \det H\phi_i(\xi) \ne 0, \quad i = 1, 2.$$

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

$$(1.6) \qquad |\langle (H\phi_1)^{-1}(\xi'')(\nabla\phi_1(\xi) - \nabla\phi_2(\zeta)), \nabla\phi_1(\xi') - \nabla\phi_2(\zeta'))\rangle| \ge c > 0,$$

$$(1.7) \qquad |\langle (H\phi_2)^{-1}(\zeta'')(\nabla\phi_1(\xi) - \nabla\phi_2(\zeta)), \nabla\phi_1(\xi') - \nabla\phi_2(\zeta'))\rangle| \ge c > 0.$$

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 ϕ_1 and ϕ_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 Φ_1^x, Φ_2^w given by

$$\Phi_1^x(y, w) = \phi_1(x) + \phi_2(y) - \phi_1(x + y - w) - \phi_2(w),$$
$$\Phi_2^w(x, z) = \phi_1(x) + \phi_2(z + w - x) - \phi_1(z) - \phi_2(w).$$

Since we are assuming $\det H\phi_i(\xi) \neq 0$, (1.6) and (1.7) are equivalent to

$$\det \begin{pmatrix} 0 & \partial_w \Phi_1^x(y,w) \\ \partial_y \Phi_1^x(y,w) & \partial_{wy}^2 \Phi_1^x(y,w) \end{pmatrix} \neq 0, \quad \det \begin{pmatrix} 0 & \partial_z \Phi_2^w(x,z) \\ \partial_x^w \Phi_2(x,z) & \partial_{zx}^2 \Phi_2^w(x,z) \end{pmatrix} \neq 0,$$

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 $v, u \in \mathbb{R}^n$,

$$\det \begin{pmatrix} 0 & u^t \\ v & M \end{pmatrix} = -(\det M)\langle M^{-1}v, u\rangle.$$

Hence, Φ_1^x, Φ_2^w satisfy the rotational curvature condition. Therefore, following the same lines of argument in [15, 17], one can see

$$\|E_1 f \cdot E_2 g\|_2 \leq C\|f\|_{\frac{4(n+1)}{3n+1}} \|g\|_{\frac{4(n+1)}{3n+1}}$$

which is corresponding to Theorem 2.3 of [17] where elliptic ϕ_1, ϕ_2 were considered.

When the eigenvalues of $H\phi_i$ have different signs, to get (1.3) for $q \geq \frac{n+3}{n+1}$, it is insufficient to impose (1.4) simply. If (1.6) and (1.7) are replaced by the weaker (1.4), then (1.3) fails for $\frac{n+2}{n} > q$. More precisely, under condition (1.4), (1.2) is valid only for p, q satisfying

$$(1.8) \qquad\qquad n\left(1 - \frac{1}{p}\right) \geq \frac{n+2}{2q}.$$

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.

To see (1.8), we consider

$$\phi_1(\xi) = \phi_2(\xi) = (\xi_1^2 + \cdots + \xi_{k-1}^2 - \xi_k^2 - \cdots - \xi_{n-2}^2 + 2\xi_{n-1}\xi_n) = \frac{1}{2}\langle M\xi, \xi\rangle,$$

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 $\zeta, \zeta' \in V_2$,

$$(1.9) \qquad\qquad |\langle M(\xi - \zeta), \xi' - \zeta'\rangle| \geq c > 0.$$

Also, (1.4) 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_λ. Observe that the supports of f_λ, g_λ are in small neighborhoods of $-e_{n-1}$, e_{n-1}, respectively. Obviously, (1.4) is satisfied but (1.6), (1.7) (namely (1.9)) 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

$$C\lambda^{-\frac{n}{p}}\|f\|_p\|g\|_p \geq \|E_1 f_\lambda E_2 g_\lambda\|_q = \lambda^{-n}\lambda^{\frac{n+2}{2q}}\|E_1 f E_2 g\|_q.$$

Letting $\lambda \rightarrow \infty$ we obtain (1.8).

Theorem 1.1 can be used to improve the known restriction estimate for hyperboloid in $\mathbb{R}^3$. Let us set

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

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*

$$\|\widehat{f d\sigma_\mathcal{H}}\|_{L^q(\mathbb{R}^3)} \le C\|f\|_{L^p(d\sigma_\mathcal{H})}.$$

The sharp estimate for $q \ge 4$ ($2/q = (1-1/p)$) 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 $n \ge 3$. To get a linear estimate in the higher dimensions, one may need to prove a stronger estimate than Theorem 1.1. However, we do not know at this point whether it is possible to get a linear estimate from a bilinear estimate such as Theorem 1.1.

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 $\overline{V}_1, \overline{V}_2$ be subcubes of $[-1, 1]^{n-1}$. For $i = 1, 2$, we define extension operators by

$$(1.10) \qquad \mathcal{E}_i f(x, t) = \int_1^2 \int_{\overline{V}_i} e^{2\pi i(x' \cdot \eta + x_n \rho - t\langle \eta, N\eta/\rho\rangle)} f(\eta, \rho) d\eta d\rho,$$

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

$$\Gamma_i(N) = \{(\eta, \rho, \tau) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} : \tau = -\langle \eta, N\eta/\rho\rangle), \eta \in \overline{V}_i, 1 \le \rho \le 2\}, i = 1, 2.$$

By a linear transform on η, 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*

$$\Theta_i = \{(\eta/\rho) : (\eta, \rho) \in \overline{V}_i \times [1, 2]\}.$$

Suppose for all $(\theta_1, \theta_2), (\theta_1', \theta_2') \in \Theta_1 \times \Theta_2$,

$$(1.11) \qquad\qquad |\langle \theta_1 - \theta_2, N(\theta_1' - \theta_2')\rangle| \sim 1.$$

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

$$(1.12) \qquad\qquad \|\mathcal{E}_1 f \mathcal{E}_2 g\|_q \le C\|f\|_2\|g\|_2.$$

If $N = id$, condition (1.11) is equivalent to (1.4) (to say, dist $(\Theta_1, \Theta_2) \sim 1$) provided Θ_1, Θ_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 $\frac{n+1}{n-1} \le q$ 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 (1.6), (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

$$\mathcal{C} = \{(\eta, \rho, (\eta_1^2 - \eta_2^2)/\rho) \in \mathbb{R}^2 \times \mathbb{R} \times \mathbb{R} : \eta \in [-1,1]^2, 1 \le \rho \le 2\}.$$

The following can be compared with the restriction estimate for the cone obtained by Wolff [21].

Corollary 1.4. *Let $d\sigma_{\mathcal{C}}$ be the Leqesgue measure on $\mathcal{C}$. Then, for $4 > q > 3$ and $\frac{2}{q} < (1 - \frac{1}{p})$, there is a constant C such that*

$$\|\widehat{f d\sigma_{\mathcal{C}}}\|_{L^q(\mathbb{R}^4)} \le C\|f\|_{L^p(d\sigma_{\mathcal{C}})}.$$

Therefore, this solves the restriction problem for $\mathcal{C}$ except for the endpoint estimates at the critical line $2/q = (1 - 2/p)$, $4 > q > 3$. The sharp estimate for $q \ge 4$ ($2/q = (1 - 2/p)$) 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 $\| \cdot \|_{L^p(d_{\mathcal{C}})}$.

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 Theorem 1.1 and Theorem 1.3 are based on Tao's proof of the bilinear restriction estimates for paraboloids ([13]), 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 $R \gg 1$,

$$(2.1) \qquad \|E_1 f_1 \cdot E_2 f_2\|_{L^{\frac{n+3}{n+1}}(Q_R)} \le CR^\alpha \|f_1\|_2 \|f_2\|_2,$$

where $Q_R = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : |t|, |x_i| \le R\}$.

A bound with large α is easy to obtain using Hölder's inequality and $\|E_i f\|_\infty \le \|f\|_2$. The main idea of the *induction on scales* argument is to establish an iterative estimate which enables us to suppress the exponent α 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 $R \gg 1$,*

$$(2.2) \qquad \|E_1 f_1 \cdot E_2 f_2\|_{L^{\frac{n+3}{n+1}}(Q_R)} \le CR^{\max(\alpha(1-\delta), c\delta) + \epsilon} \|f_1\|_2 \|f_2\|_2$$

holds for some constant c, independent of R, δ, ϵ.

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 $(\epsilon + \sqrt{\epsilon^2 + 4c\epsilon})/2$. Since ϵ 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

$$\widetilde{V}_i = V_i + O(\epsilon),$$

where for $d > 0$ and $A \subset \mathbb{R}^n$, $A + O(d)$ denotes the set $\{x \in \mathbb{R}^n : \text{dist } (x, A) < Cd\}$. By continuity, we may assume conditions (1.5), (1.6) and (1.7) are still valid if V_1, V_2 are replaced by $\widetilde{V}_1, \widetilde{V}_2$, respectively, and ϵ 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

$$\mathcal{V}_i = R^{-1/2}\mathbb{Z}^n \cap \widetilde{V}_i, \quad i = 1, 2.$$

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].

Lemma 2.2 (Wave packet decomposition). *Suppose f_1, f_2 are supported in V_1, 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 $p_{w_i} = E_i(\widehat{p_{w_i}(\cdot, 0)})$ and C_{w_i}, p_{w_i} satisfy the following:

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

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

(P3) *If $\text{dist } ((x, t), T_{y_i, v_i}) \geq R^{\delta + 1/2}$, then $|p_{y_i, v_i}(x, t)| \leq CR^{-100n}$. More precisely,*

$$|p_{y_i, v_i}(x, t)| \leq CR^{-n/4}\left(1 + \frac{|x - (y_i + t\nabla\phi_i(v_i))|}{R^{1/2}}\right)^{-N} \text{ for any } N.$$

(P4) *For any $S \subset W_i$, $\|\sum_{w_i \in S} p_{w_i}(\cdot, t)\|_2^2 \leq C\#S$.*

Proof of Lemma 2.2. Using the Poisson summation formula, one can find an η satisfying $\text{supp } \widehat{\eta} \subset B(0, 1)$ 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

$$\eta_y(x) = \eta(\frac{x + y}{R^{1/2}}), \quad \psi_{v_i}(\xi) = \psi(R^{1/2}(\xi - v_i)).$$

Note that $f_i = \sum_{v_i \in \mathcal{V}_i, y \in \mathcal{Y}} \mathcal{F}^{-1}(\widehat{\psi_{v_i} f_i} \eta_y)$ since f_i is supported in V_i and $\mathcal{F}^{-1}(\widehat{\psi_{v_i} f_i} \eta_y)$ is supported in the (CR^{-1})-neighborhood of v_i. Here $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Then it follows that

$$E_i f_i(x,t) = \sum_{v \in \mathcal{V}_i, y \in \mathcal{Y}} q_{y,v_i}(x,t),$$

where

(2.3) $$q_{y,v_i}(x,t) = \int e^{2\pi i(x \cdot \xi - t\phi_i(\xi))} \mathcal{F}^{-1}(\widehat{\psi_{v_i} f_i} \eta_y) d\xi.$$

Now we show

Lemma 2.3. *If $|t| \leq R$, for any N*

$$|q_{y,v_i}(x,t)| \leq CM(\widehat{\psi_{v_i} f_i})(-y) \left(1 + \frac{|x - (y + t\nabla\phi_i(v_i))|}{R^{1/2}}\right)^{-N},$$

where M denotes the Hardy-Littlewood maximal function.

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

For simplicity we set $F_{v_i} = \widehat{\psi_{v_i} f_i}$. By putting $\widetilde{\psi}_{v_i}$ into the integral in (2.3), translation and re-scaling, we have

$$q_{y,v_i}(x,t) = R^{-n/2} \int K(x+z,t)\eta(\frac{z+y}{R^{1/2}})F_{v_i}(z)dz,$$

where

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

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 $|x - t\phi_i(v_i)| \geq CR^{1/2}$,

$$R^{-1/2}|x - t\nabla\phi_i(v_i)| \leq C|\nabla_\xi(R^{-1/2}x \cdot \xi + x \cdot v_i - t\phi_i(R^{-1/2}\xi + v_i)|.$$

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

$$|q_{y,v_i}(x,t)| \leq CR^{-n/2} \int \left(1 + \frac{|a+z|}{R^{1/2}}\right)^{-N} |\eta(\frac{z}{R^{1/2}})F_{v_i}(z-y)|dz.$$

Therefore it is sufficient to show that for $\lambda \gg 1$

(2.4) $$I = \lambda^{-n} \int \left(1 + \frac{|a+z|}{\lambda}\right)^{-N} |\eta(\frac{z}{\lambda})F(z)|dz \leq C(1 + |a|/\lambda)^{-N} MF(0)$$

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

$$\sum_0^\infty I_j \leq C\lambda^{-n} \left(\int_{|z| \leq \lambda/2} + \sum_{j=1}^\infty \int_{|z| \sim \lambda 2^j}\right) \left(1 + \frac{|a+z|}{\lambda}\right)^{-N} |\eta(\frac{z}{\lambda})F(z)|dz.$$

Then, if $|a| \sim \lambda 2^j$, by the rapid decay of η, $I_j \leq C(|a|/\lambda)^{-M} MF(0)$ for any M. If $|a| \not\sim \lambda 2^j$, $I_j \leq C2^{-Mj}(|a|/\lambda)^{-N} MF(0)$ for any M because $C|a+z| \geq |a|$ if $|a| \not\sim |z|$. Summation with respect to j gives (2.4).   □

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

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

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

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

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

$$\| \sum_{v_i \in \mathcal{V}_i, y_i \in \mathcal{Y}} p_{y_i, v_i}(\cdot, t) \|_2 \leq C(\sum_{v \in \mathcal{V}_i} \| \sum_{y_i \in \mathcal{Y}} p_{y_i, v_i}(\cdot, t) \|_2^2)^{1/2}.$$

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

Since $\psi_{v_i} f_i$ is supported in a ball of radius about $R^{-1/2}$, if $|x - x'| \leq CR^{1/2}$, $M(\widehat{\psi_{v_i} f_i})(x) \sim M(\widehat{\psi_{v_i} f_i})(x')$ (see [13]). This can be shown using a bump function adapted to the ball where $\psi_{v_i} f_i$ is supported. Hence,

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

The second inequality is from the Hardy-Littlewood maximal theorem and the third from Plancherel's theorem. So we get (P1).   □

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 2.1 it is sufficient to show that (2.1) implies

$$(2.5) \qquad \| \sum_{w_1 \in W_1} \sum_{w_2 \in W_2} C_{w_1} C_{w_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \lesssim (R^{\alpha(1-\delta)} + R^{c\delta}).$$

Here, $A \lesssim B$ means there is a constant C_ϵ such that $A \leq C_\epsilon R^\epsilon B$ for any $\epsilon > 0$, $R \gg 1$. 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 $C > 0$,

$$(2.6) \qquad\qquad T_{w_1} \cap CQ_R \neq \emptyset, \quad T_{w_2} \cap CQ_R \neq \emptyset$$

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 $C > 0$,

$$\| \sum_{T_{w_1} \cap CQ_R = \emptyset \text{ or } T_{w_2} \cap CQ_R = \emptyset} C_{w_1} C_{w_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \lesssim R^{-100n}.$$

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

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

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 $(\#\{w_i : |C_{w_i}| \sim \mu_i\})^{\frac{1}{2}} \leq C/\mu_i$ from (P1), for (2.5) it is sufficient to show that for any subset $\mathcal{W}_1 \subset W_1$, $\mathcal{W}_2 \subset W_2$,

$$(2.7) \qquad \| \sum_{w_1 \in \mathcal{W}_1} \sum_{w_2 \in \mathcal{W}_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \lesssim C(R^{\alpha(1-\delta)} + R^{c\delta})(\#\mathcal{W}_1 \#\mathcal{W}_2)^{1/2}.$$

In fact, since $|C_{w_1}|/\mu_1, |C_{w_2}|/\mu_2 \sim 1$, we may absorb harmless constants C_{w_1}/μ_1, C_{w_2}/μ_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 $\mathcal{B}$ the collection of these cubes. By triangle inequality

$$\| \sum_{w_1 \in \mathcal{W}_1} \sum_{w_2 \in \mathcal{W}_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \leq C \sum_{b \in \mathcal{B}} \| \sum_{w_1 \in \mathcal{W}_1} \sum_{w_2 \in \mathcal{W}_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}.$$

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

$$(2.8) \qquad \| \sum_{w_1 \in \mathcal{W}_1} \sum_{w_2 \in \mathcal{W}_2} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(Q_R)} \leq \sum_{b \in \mathcal{B}} \| \sum_{w_1 \approx b} \sum_{w_2 \approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}$$
$$+ \sum_{b \in \mathcal{B}} \| \sum_{w_1 \not\approx b, \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}.$$

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 $w_i \in \mathcal{W}_i$ and b.** Now we divide Q_R into essentially disjoint cubes q of side length $R^{1/2}$. Let us denote by $\mathcal{Q}$ this collection of q. For $q \in \mathcal{Q}$, let us define

$$\mathcal{W}_i(q) = \{w_i \in \mathcal{W}_i : T_{w_i} \cap R^\delta q \neq \emptyset\}.$$

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

$$\mathcal{Q}(\mu_1, \mu_2) = \{q \in \mathcal{Q} : \mu_1 \leq \#\mathcal{W}_1(q) < 2\mu_1, \quad \mu_2 \leq \#\mathcal{W}_2(q) < 2\mu_2\}.$$

For $w_i \in \mathcal{W}_i$ and dyadic numbers $1 \leq \mu_1, \mu_2 \leq R^{100n}$, let us set

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

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

$$\mathcal{W}_i(\lambda, \mu_1, \mu_2) = \{w_i \in \mathcal{W}_i : \lambda \leq \lambda(w_i, \mu_1, \mu_2) < 2\lambda\}.$$

For each dyadic number $1 \leq \lambda, \mu_1, \mu_2 \leq R^{100n}$ and $w_i \in \mathcal{W}_i(\lambda, \mu_1, \mu_2)$, we define $b(w_i, \lambda, \mu_1, \mu_2) \in \mathcal{B}$ to be a ball which maximizes the quantity

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

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

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

We define a relation $\approx_{\lambda, \mu_1, \mu_2}$ between $w_i \in \mathcal{W}_i(\lambda, \mu_1, \mu_2)$ and $b \in \mathcal{B}$ 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 $w_i \in \mathcal{W}_i$ and b by saying

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

Clearly for each w_i, there are $O((\log R)^3)$ balls b in $\mathcal{B}$ for which $w_i \approx b$ since there are $O((\log R)^3)$ dyadic triples (λ, μ_1, μ_2). So it follows that for all $w_i \in \mathcal{W}_i$,

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

Since $\sum_{w_i \approx b} p_{w_i} = E_i(\sum_{w_i \approx b} \widehat{p_{w_i}(\cdot, 0)})$ 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

$$\sum_{b \in \mathcal{B}} \| \sum_{w_1 \approx b} \sum_{w_2 \approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)} \leq C R^{\alpha(1-\delta)} \sum_{b \in \mathcal{B}} \#\{w_1 : w_1 \approx b\}^{\frac{1}{2}} \#\{w_2 : w_2 \approx b\}^{\frac{1}{2}}.$$

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

$$\sum_{b \in \mathcal{B}} \| \sum_{w_1 \approx b} \sum_{w_2 \approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)}$$

$$\leq C R^{\alpha(1-\delta)} \left(\sum_{w_1 \in \mathcal{W}_1} \#\{b \in \mathcal{B} : w_1 \approx b\} \right)^{1/2} \left(\sum_{w_2 \in \mathcal{W}_2} \#\{b \in \mathcal{B} : w_2 \approx b\} \right)^{1/2}$$

$$\lesssim R^{\alpha(1-\delta)} (\#\mathcal{W}_1)^{1/2} (\#\mathcal{W}_2)^{1/2}.$$

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

$$\sum_{b \in \mathcal{B}} \| \sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)} \lesssim R^{c\delta} (\#\mathcal{W}_1)^{1/2} (\#\mathcal{W}_2)^{1/2}.$$

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

$$(2.11) \qquad \| \sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^{\frac{n+3}{n+1}}(b)} \lesssim R^{c\delta} (\#\mathcal{W}_1)^{1/2} (\#\mathcal{W}_2)^{1/2}.$$

Here, we are assuming the sum is always taken over some subset of $\mathcal{W}_1 \times \mathcal{W}_2$. Since $p_{w_i} = E_i(\widehat{p_{w_i}(\cdot, 0)})$, using the local restriction estimate $\|E_i f\|_{L^2(Q_R)} \leq C R^{\frac{1}{2}} \|f\|_2$ (see [17]), Schwarz's inequality and (P4) in Lemma 2.2, we obtain

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

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

$$(2.12) \qquad \| \sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^2(b)} \lesssim R^{c\delta} R^{-(n-1)/4} (\#\mathcal{W}_1)^{1/2} (\#\mathcal{W}_2)^{1/2}.$$

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

$$(2.13) \qquad \sum_{q \in \mathcal{Q}, q \subset 2b} \| \sum_{w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^2(q)}^2 \lesssim R^{c\delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.$$

Now, we make several obvious reductions by pigeonholing. Since $|p_{w_i}| \leq C R^{-100n}$ on q if $T_{w_i} \cap R^\delta q = \emptyset$ 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 $1 \leq \mu_1, \mu_2 \leq R^{100n}$, the $\mathcal{Q}$ in the outer sum of (2.13) can also be replaced by $\mathcal{Q}(\mu_1, \mu_2)$ for some μ_1, μ_2. Therefore, the matters are reduced to showing

$$(2.14) \qquad \sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \| \sum_{(w_1, w_2) \in \mathcal{W}_1(q) \times \mathcal{W}_2(q), w_1 \not\approx b \text{ or } w_2 \not\approx b} p_{w_1} p_{w_2} \|_{L^2(q)}^2$$
$$\lesssim R^{c\delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.$$

Clearly, $\bigcup_{q \in \mathcal{Q}(\mu_1, \mu_2)} \mathcal{W}_i(q) = \bigcup_\lambda \mathcal{W}_i(\lambda, \mu_1, \mu_2)$, $i = 1, 2$. 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 $\mathcal{W}_1(q) \cap \mathcal{W}_1(\lambda_1, \mu_1, \mu_2)$, $\mathcal{W}_2(q) \cap \mathcal{W}_2(\lambda_2, \mu_1, \mu_2)$, respectively. To simplify the notation, we set

$$(2.15) \qquad \mathcal{W}_i^{\not\approx b}(q, \mu_1, \mu_2, \lambda_i) = \{w_i \in \mathcal{W}_i(q) \cap \mathcal{W}_i(\lambda_i, \mu_1, \mu_2) : w_i \not\approx b\}.$$

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

$$\mathcal{U}_i(q) = \{w_i \subset \mathcal{U}_i : T_{w_i} \cap R^\delta q \neq \emptyset\}.$$

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

$$(2.16) \qquad \sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \| \sum_{w_1 \in \mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)} \sum_{w_2 \in \mathcal{U}_2(q)} p_{w_1} p_{w_2} \|_{L^2(q)}^2$$
$$\lesssim R^{c\delta} R^{-(n-1)/2} \# \mathcal{W}_1 \# \mathcal{W}_2.$$

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

$$(2.17) \qquad \Pi_{\xi_1, \xi_2'} = \{\xi_1' \in \widetilde{V}_1 : \Phi_{\xi_1, \xi_2'}(\xi_1') = 0, \xi_1' + \xi_2' - \xi_1 \in \widetilde{V}_2\}.$$

From (1.6) and (1.7) which imply (1.4), we see

$$(2.18) \qquad |\nabla \Phi_{\xi_1, \xi_2'}(\xi_1')| = |\nabla \phi_2(\xi_1' + \xi_2' - \xi_1) - \nabla \phi_1(\xi_1')| \geq c > 0$$

provided $\xi_1' + \xi_2' - \xi_1 \in \widetilde{V}_2$. So, dividing $\widetilde{V}_1$, $\widetilde{V}_2$ into sufficiently small cubes, we may assume the sets $\Pi_{\xi_1, \xi_2'}$ are smooth hyper-surfaces for all $\xi_1 \in \widetilde{V}_1$, $\xi_2' \in \widehat{V}_2$ because $E_1 f E_2 g$ is written as a finite sum of operators satisfying this assumption.

For $\mathcal{U}_1 \subset \mathcal{W}_1$, we define

$$\mathcal{N}(\mathcal{U}_1) = \sup_{\xi_1 \in \widetilde{V}_1, \xi_2' \in \widetilde{V}_2} \#\{w_1' \in \mathcal{U}_1 : v_1' \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\}.$$

Lemma 2.4. *For $q \in \mathcal{Q}$ and $\mathcal{U}_i \subset \mathcal{W}_i(q)$, $i = 1, 2$,*

$$\| \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2 \in \mathcal{U}_2} p_{w_1} p_{w_2} \|_{L^2}^2 \lesssim R^{c\delta} R^{-(n-1)/2} \mathcal{N}(\mathcal{U}_1) \#\mathcal{U}_1 \#\mathcal{U}_2.$$

Proof of Lemma 2.4. We write

$$(2.19) \qquad \| \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2 \in \mathcal{U}_2} p_{w_1} p_{w_2} \|_{L^2}^2 = \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2' \in \mathcal{U}_2} \sum_{w_1' \in \mathcal{U}_1} I_{w_1, w_2', w_1'}$$

where

$$I_{w_1, w_2', w_1'} = \sum_{w_2 \in \mathcal{U}_2} \langle p_{w_1} p_{w_2}, p_{w_1'} p_{w_2'} \rangle.$$

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 $v_1, v_1' \in V_1 + O(R^{-1/2})$, $v_2, v_2' \in V_2 + O(R^{-1/2})$,

$$(2.20) \qquad\qquad\qquad v_1 + v_2 = v_1' + v_2' + O(R^{-1/2}) \text{ and}$$

$$(2.21) \qquad\qquad \phi_1(v_1) + \phi_2(v_2) = \phi_1(v_1') + \phi_2(v_2') + O(R^{-1/2})$$

because $\widehat{p_{w_i}}$ are supported in the $O(R^{-1/2})$-neighborhood of the point $(v_i, \phi(v_i))$.

From (2.20) we see $\langle p_{w_1} p_{w_2}, p_{w_1'} p_{w_2'} \rangle = 0$ if $v_1' + v_2' - v_1 \notin \widetilde{V}_2$. So we may always assume $v_1' + v_2' - v_1 \in \widetilde{V}_2$. 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

$$\phi_1(v_1) + \phi_2(v_1' + v_2' - v_1) = \phi_1(v_1') + \phi_2(v_2') + O(R^{-1/2}).$$

Therefore, from (2.18) it is obvious that v_1' satisfying the above is contained in $\Pi_{v_1, v_2'} + O(R^{-1/2})$.

Therefore, the left-hand side of (2.19) is equal to

$$(2.22) \qquad \sum_{w_1 \in \mathcal{U}_1} \sum_{w_2' \in \mathcal{U}_2} \sum_{\{w_1' \in \mathcal{U}_1 : v_1' \in \Pi_{v_1, v_2'} + O(R^{-1/2})\}}$$

$$\left(\sum_{\{w_2 \in \mathcal{U}_2 : v_2 = v_1' + v_2' - v_1 + O(R^{-1/2})\}} \langle p_{w_1} p_{w_2}, p_{w_1'} p_{w_2'} \rangle \right).$$

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_{w_2} (T_{w_1} resp.) are passing through $R^\delta q$, there are at most $O(R^{c\delta})$-w_2 (w_1 resp.) if v_2 (v_1 resp.) is determined because $y_2 \in \mathcal{Y}$ (y_1 resp.) are $R^{1/2}$-separated. Using (P3) and the transversality between the tubes T_{w_1} and

T_{w_2} (by (1.4)), 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',

$$\sum_{\{w_2 \in \mathcal{U}_2 : v_2 = v_1' + v_2' - v_1 + O(R^{-1/2})\}} |\langle p_{w_1} p_{w_2}, p_{w_1'} p_{w_2'} \rangle| \lesssim R^{c\delta} R^{-(n-1)/2}.$$

From this and (2.22) we conclude the proof of Lemma 2.4. $\qquad\square$

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 $1 \le \mu_1, \mu_2, \lambda_1 \le R^{100n}$ and $q \in \mathcal{Q}(\mu_1, \mu_2)$, $q \subset 2b$,*

$$\mathcal{N}(\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)) \lesssim R^{c\delta} \frac{\#\mathcal{W}_2}{\lambda_1 \mu_2}.$$

Using Lemma 2.4, we see

$$\sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \Big\| \sum_{w_1 \in \mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)} \sum_{w_2 \in \mathcal{U}_2(q)} p_{w_1} p_{w_2} \Big\|_{L^2(q)}^2$$

$$\lesssim R^{c\delta} R^{-\frac{n-1}{2}} \sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \mathcal{N}(\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)) \#\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1) \#\mathcal{U}_2(q).$$

Using Lemma 2.5, for (2.16) it suffices to show

$$\frac{\#\mathcal{W}_2}{\lambda_1 \mu_2} \sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \#\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1) \#\mathcal{U}_2(q) \lesssim \#\mathcal{W}_1 \#\mathcal{W}_2.$$

Since $q \in Q(\mu_1, \mu_2)$ and $\mathcal{U}_2 \subset \mathcal{W}_2$, $\#\mathcal{U}_2(q) \le \mu_2$. So we need to show

$$(2.23) \qquad \frac{1}{\lambda_1} \sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \#\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1) \lesssim \#\mathcal{W}_1.$$

Recalling (2.15) and changing the order of summation,

$$\sum_{q \in \mathcal{Q}(\mu_1, \mu_2), q \subset 2b} \#\mathcal{W}_1^{\not\approx b}(q, \mu_1, \mu_2, \lambda_1)$$

$$\le \sum_{w_1 \in \mathcal{W}_1(\lambda_1, \mu_1, \mu_2)} \#\{q \in \mathcal{Q}(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \ne \emptyset, w_1 \not\approx b\}$$

$$\le \sum_{w_1 \in \mathcal{W}_1(\lambda_1, \mu_1, \mu_2)} \#\{q \in \mathcal{Q}(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \ne \emptyset\}.$$

Since $w_1 \in \mathcal{W}_1(\lambda_1, \mu_1, \mu_2)$, $\#\{q \in \mathcal{Q}(\mu_1, \mu_2) : T_{w_1} \cap R^\delta q \ne \emptyset\} \le \lambda_1$. 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 $1 \le \mu_1, \mu_2, \lambda_1 \le R^{100n}$, $\xi_1 \in \widetilde{V}_1, \xi_2' \in \widetilde{V}_2$, and $q_0 \in \mathcal{Q}(\mu_1, \mu_2)$, $q_0 \subset 2b$,

$$(2.24) \qquad \#\{w_1' \in \mathcal{W}_1^{\not\approx b}(q_0, \mu_1, \mu_2, \lambda_1) : v_1' \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\} \lesssim R^{c\delta} \frac{\#\mathcal{W}_2}{\lambda_1 \mu_2}.$$

For simplicity we set

$$\mathcal{W}_1^{\not\approx b}(\Pi_{\xi_1, \xi_2'}) = \{w_1' \in \mathcal{W}_1^{\not\approx b}(q_0, \mu_1, \mu_2, \lambda_1) : v_1' \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\}.$$

Let $w_1 \in \mathcal{W}_1^{\not\gtrsim b}(\Pi_{\xi_1, \xi_2'})$. Then $T_{w_1} \cap R^\delta q_0 \neq \emptyset$ and $b \cap 10b(w_1, \lambda_1, \mu_1, \mu_2) = \emptyset$. Since $q_0 \subset 2b$, dist $(q_0, 2b(w_1, \lambda_1, \mu_1, \mu_2)) \geq R^{1-\delta}$. So, by (2.9) we see

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

By the definition of $\mathcal{Q}(\mu_1, \mu_2)$, $\#\mathcal{W}_2(q) \sim \mu_2$ for each $q \in \mathcal{Q}(\mu_1, \mu_2)$. From (2.25) and summation in w_1, w_2, we get

$$(2.26) \quad \#\{(q, w_1, w_2) \in Q(\mu_1, \mu_2) \times \mathcal{W}_1^{\not\gtrsim b}(\Pi_{\xi_1, \xi_2'}) \times \mathcal{W}_2 : R^\delta q \cap T_{w_1} \neq \emptyset,$$
$$\text{dist }(q_0, q) \geq R^{1-\delta}, R^\delta q \cap T_{w_2} \neq \emptyset\} \gtrsim R^{-c\delta} \lambda_1 \#\mathcal{W}_1^{\not\gtrsim b}(\Pi_{\xi_1, \xi_2'})\mu_2.$$

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 $w_2 \in \mathcal{W}_2$, set*

$$S = \{(q, w_1) \in Q(\mu_1, \mu_2) \times \mathcal{W}_1^{\not\gtrsim b}(\Pi_{\xi_1, \xi_2'}) : R^\delta q \cap T_{w_1} \neq \emptyset,$$
$$\text{dist }(q_0, q) \geq R^{1-\delta}, R^\delta q \cap T_{w_2} \neq \emptyset\}.$$

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

Using Lemma 2.6, the left-hand side of (2.26) is bounded by $R^{c\delta}\#\mathcal{W}_2$. Therefore,

$$\lambda_1 \#\mathcal{W}_1^{\not\gtrsim b}(\Pi_{\xi_1, \xi_2'})\mu_2 \lesssim R^{c\delta}\#\mathcal{W}_2.$$

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

Lemma 2.7. *Let Π 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 Π at p. For $\delta^\epsilon \leq \lambda \leq \mu \leq 1$, $0 < \epsilon \ll 1$, set*

$$\mathcal{C}(\Pi, \lambda, \mu) = \{s(u, 1) \in \mathbb{R}^n \times \mathbb{R} : \lambda \leq s \leq \mu, u \in \Pi\},$$
$$\mathcal{C}(\Pi, \lambda, \mu, \delta) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \text{dist }((x, t), \mathcal{C}(\Pi, \lambda, \mu)) \leq \delta\}.$$

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 $T_y^v(\delta) = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} : \text{dist }(l_y^v, (x, t)) \leq \delta\}$. Suppose for some $v \in B(0, 1)$,

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

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

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

$$(2.28) \qquad\qquad T_y^v(\delta) \cap \mathcal{C}(\Pi, 1/2, 1, \delta) \subset B(y_0, C\delta)$$

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

$$(2.29) \qquad\qquad \text{dist }(\mathcal{T}_q(\mathcal{C}(\Pi)), (v, 1)) \sim 1,$$

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

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

$$\text{dist }(\mathcal{T}_q(\mathcal{C}(\Pi)), (v, 1)) \leq \varepsilon$$

with small $\varepsilon > 0$ (much smaller than the implicit lower bound in (2.29)). Note that

$$\mathcal{T}_q(\mathcal{C}(\Pi)) = \text{span}(\{(u,0) : u \in \mathcal{T}_p(\Pi)\}) \oplus \text{span}(\{(p,1)\}).$$

Hence, $|(v,1) - (c_1 u_1 + \cdots + c_n u_n + c_{n+1} p, c_{n+1})| \leq \varepsilon$ for some $c_1, \ldots, c_{n+1} \in \mathbb{R}$ and $u_1, \ldots, u_n \in \mathcal{T}_p(\Pi)$. Since $|c_{n+1} - 1| \leq \varepsilon$, $|v - c_1 u_1 - \cdots - c_n u_n - p| \leq \varepsilon$. So dist $(\mathbb{P}_p(\Pi), v) \leq 2\varepsilon$. This contradicts (2.27). $\qquad\square$

Proof of Lemma 2.6. To begin with, we may assume the sets $\widetilde{V}_1$, $\widetilde{V}_2$ are small enough so that $\nabla\phi_1$, $\nabla\phi_2$ are diffeomorphisms on $\widetilde{V}_1$, $\widetilde{V}_2$, 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 $w_1 = (y_1, v_1) \in \mathcal{W}_1^{\nleq b}(\Pi_{\xi_1, \xi_2'})$ are contained in the set

$$\{(\nabla\phi_1(v), 1) \in \mathbb{R}^n \times \mathbb{R} : v \in \Pi_{\xi_1, \xi_2'} + O(R^{-1/2})\}.$$

Since $\xi \to \nabla\phi_1(\xi)$ is a diffeomorphism, $\nabla\phi_1(\Pi_{\xi_1, \xi_2'})$ is also a smooth surface. Let us define a conic surface $\mathcal{C}_{\xi_1, \xi_2'}$ by

$$\mathcal{C}_{\xi_1, \xi_2'} = \{(su, s) \in \mathbb{R}^n \times \mathbb{R} : u \in \nabla\phi_1(\Pi_{\xi_1, \xi_2'}), |s| \leq 2R\}.$$

Since T_{w_1} with $w_1 \in \mathcal{W}_1^{\nleq b}(\Pi_{\xi_1, \xi_2'})$ meets $R^\delta q_0$,

$$\bigcup_{w_1 \in \mathcal{W}_1^{\nleq b}(\Pi_{\xi_1, \xi_2'})} T_{w_1} \subset \mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}) + q_0$$

where $C_{\xi_1, \xi_2'}(A) = \mathcal{C}_{\xi_1, \xi_2'} + O(A)$.

If $(q, w_1) \in S$, then $R^\delta q$ meets T_{w_1}. So, q is contained in $\mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}) + q_0$ 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

$$q \subset \mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}, R^{1-\delta}, R, q_0)$$

whenever $(q, w_1) \in S$ for some w_1 where

$$\mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}, R^{1-\delta}, R, q_0) = \mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}) \cap \{(x,t) : R^{1-\delta} \leq |t| \leq R\} + q_0.$$

Furthermore, $R^\delta q \cap T_{w_2} \neq \emptyset$ if $(q, w_1) \in S$. Therefore, we see

$$(2.30) \qquad \bigcup_{q:(q,w_1)\in S \text{ for some } w_1} q \subset R^\delta T_{w_2} \cap \mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}, R^{1-\delta}, R, q_0)$$

where

$$(2.31) \qquad R^\delta T_{w_2} = \{(x,t) : |t| \leq R, |x - (y_2 + t\nabla\phi_i(v_2))| \leq CR^{1/2+\delta}\}.$$

Now we claim that

$$(2.32) \qquad R^\delta T_{w_2} \cap \mathcal{C}_{\xi_1, \xi_2'}(R^{1/2+\delta}, R^{1-\delta}, R, q_0) \subset B(y_0, R^{1/2+c\delta})$$

for some $c > 0$, $y_0 \in \mathbb{R}^{n+1}$. By rescaling it is sufficient to show that for $\delta^\epsilon \leq \lambda \leq 1$, $0 < \epsilon \ll 1$,

$$(2.33) \qquad T_{y_2}^{\nabla\phi_2(v_2)}(\delta) \cap \mathcal{C}(\nabla\phi_1(\Pi_{\xi_1, \xi_2'}), \lambda, 1, \delta) \subset B(y_0, C\delta/\lambda)$$

for some y_0 where $w_2 = (y_2, v_2) \in \mathcal{W}_2$.

The normal vector of the surface $\nabla\phi_1(\Pi_{\xi_1, \xi_2'})$ at $\nabla\phi_1(\xi_1')$ is parallel to

$$N_{\xi_1, \xi_2'}(\xi_1') = (\phi_1)_{\xi\xi}''|_{\xi=\xi_1'}^{-1} [\nabla\phi_2(\xi_1' + \xi_2' - \xi_1) - \nabla\phi_1(\xi_1')]$$

with $\xi_1' + \xi_2' - \xi_1 \in \widetilde{V}_2$ because $\nabla \phi_1(\Pi_{\xi_1, \xi_2'})$ is contained in

$$\{\eta \in \nabla \phi_1(\widetilde{V}_1) : \phi_1(\xi_1) + \phi_2(\nabla \phi_1^{-1}(\eta) + \xi_2' - \xi_1) - \phi_1(\nabla \phi_1^{-1}(\eta)) - \phi_2(\xi_2') = 0\}.$$

Hence, the tangent plane $\mathbb{P}_{\xi_1, \xi_2'}(\xi_1')$ 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, dist $(\mathbb{P}_{\xi_1, \xi_2'}(\xi_1'), y) \sim |\langle N_{\xi_1, \xi_2'}(\xi_1'),$ $y - \nabla \phi_1(\xi_1') \rangle|$ since $|N_{\xi_1, \xi_2'}(\xi_1')| \sim 1$. Therefore, (1.6) implies for all $\xi_1, \xi_1' \in \widetilde{V}_1$, $\xi_2, \xi_2' \in \widetilde{V}_2$,

$$\text{dist } (\mathbb{P}_{\xi_1, \xi_2'}(\xi_1'), \nabla \phi_2(\xi_2')) \sim 1.$$

In particular, dist $(\mathbb{P}_{\xi_1, \xi_2'}(\xi_1'), \nabla \phi_2(v_2)) \sim 1$ for all $\xi_1' \in \Pi_{\xi_1, \xi_2'}$. Using Lemma 2.7, we see (2.33) and hence (2.32).

Since $q \in \mathcal{Q}$ are $R^{1/2}$-separated cubes, from (2.32) and (2.30) it follows that

$$(2.34) \qquad \#\{q \in \mathcal{Q} : (q, w_1) \in S \text{ for some } w_1\} = O(R^{c\delta}).$$

Since 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 $O(R^{c\delta})$, because the major directions $(\nabla \phi_1(v), 1), v \in \mathcal{V}_1$ of tubes are separated by $\sim R^{-1/2}$. This and (2.34) give $\#S = O(R^{c\delta})$. $\qquad \square$

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

$$\phi_1 = \phi_2 = \langle \eta, N\eta/\rho \rangle, \quad V_i = \overline{V}_i \times [1, 2], \quad i = 1, 2,$$

which define the extension operators $\mathcal{E}_1, \mathcal{E}_2$. From these we define $\mathcal{V}_i, \mathcal{W}_i, T_{w_i}$ by following the same procedure as in the previous section. We keep the same notations $\mathcal{W}_i(q), \mathcal{Q}(\mu_1, \mu_2), \mathcal{W}_1^{\not\approx b}(q_0, \mu_1, \mu_2, \lambda_1), \dots$, etc.

Since $\nabla_{\eta, \rho} \phi_1 = \nabla_{\eta, \rho} \phi_2 = (2N\eta/\rho, -\langle \eta, N\eta \rangle/\rho^2)$, the major directions $(\nabla \phi_i, 1)$ of the tubes $T_{w_i}, w_i \in \mathcal{W}_i$ are contained in the set

$$(3.1) \qquad \{(2N\theta, -\langle \theta, N\theta \rangle, 1) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} : \theta \in \Theta_i\}.$$

The transversality between tubes T_{w_1} and $T_{w_2}, w_1 \in \mathcal{W}_1, w_2 \in \mathcal{W}_2$ is easy to see from (1.11) which implies dist $(\Theta_1, \Theta_2) \sim 1$. The set $\Pi_{\xi_1, \xi_2'}$ given by (2.17) is also a smooth $n-1$-dimensional hypersurface because 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 2.5).

Lemma 3.1. *Let* $\mathcal{W}_1^{\not\approx b}(\Pi_{\xi_1, \xi_2'})$ *be defined by the same way as in Section* 2 *with* $\phi_1 = \phi_2 = \langle \eta, N\eta/\rho \rangle, V_i = \overline{V}_i \times [1, 2], i = 1, 2.$ *For each* $w_2 \in \mathcal{W}_2$, *set*

$$S = \{(q, w_1) \in \mathcal{Q}(\mu_1, \mu_2) \times \mathcal{W}_1^{\not\approx b}(\Pi_{\xi_1, \xi_2'}) :$$
$$R^\delta q \cap T_{w_1} \neq \emptyset, \text{dist } (q_0, q) \geq R^{1-\delta}, R^\delta q \cap T_{w_2} \neq \emptyset\}.$$

Then, $\#S \lesssim O(R^{c\delta}).$

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 $c > 0$,

$$(3.2) \qquad \#\{q \in \mathcal{Q} : (q, w_1) \in S \text{ for some } w_1\} = O(R^{c\delta}).$$

We define a conic surface $\mathcal{C}_1$ by

$$\mathcal{C}_1 = \{s(2N\theta, -\langle \theta, N\theta \rangle, 1) \in \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R} : \theta \in \Theta_1, |s| \le 2R\}.$$

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

$$\bigcup_{w_1 \in \mathcal{W}_1^{\not\approx b}(\Pi_{\xi_1, \xi_2'})} T_{w_1} \subset \mathcal{C}_1(R^{1/2+\delta}) + q_0$$

where $\mathcal{C}_1(A) = C_1 + O(A)$. If $(q, w_1) \in S$, $R^\delta q \cap T_{w_2} \ne \emptyset$ and dist $(q, q_0) \ge R^{1-\delta}$. Hence,

$$(3.3) \qquad \bigcup_{q:(q,w_1)\in S \text{ for some } w_1} q \subset R^\delta T_{w_2} \cap \mathcal{C}_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0),$$

where $R^\delta T_{w_2}$ is given by (2.31) and

$$\mathcal{C}_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0) = \mathcal{C}_1(R^{1/2+\delta}) \cap \{(x, t) : R^{1-\delta} \le |t| \le R\} + q_0.$$

Let us set

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

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

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

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,

$$(3.4) \quad \text{dist}\,(\mathbb{P}_p(\Pi), \nabla \phi_2(v_2)) \sim |(\theta_1, 1) \cdot ((2N\theta_2 - 2N\theta_1), -\theta_2 \cdot N\theta_2 + \theta_1 \cdot N\theta_1)|$$
$$= |(\theta_1 - \theta_2) \cdot N(\theta_1 - \theta_2)| \sim 1$$

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

$$R^\delta T_{w_2} \cap \mathcal{C}_1(R^{1/2+\delta}, R^{1-\delta}, R, q_0) \subset B(y_0, R^{1/2+c\delta})$$

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 $O(R^{c\delta})$-w_1s 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

$$(u, 1) // \frac{c(q_0) - c(q)}{|c(q_0) - c(q)|}.$$

If $(q, w_1) \in S$ with $w_1 = (y_1, v_1)$, then dist $(q_0, q) \geq R^{1-\delta}$ and T_{w_1} passes through both $R^\delta q_0$ and $R^\delta q$. So, the major direction $(\phi_1(v_1), 1)$ of T_{w_1} is contained in the $O(R^{-1/2+c\delta})$-neighborhood of $(u, 1)$. Therefore if $(q, (y_1, v_1)) \in S$, then

$$(3.5) \qquad\qquad \nabla\phi_1(v_1) = u + O(R^{-1/2+c\delta}).$$

However, there may be possibly more than $O(R^{c\delta})$-v_1'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 $\mathcal{E}_i$ are written as finite sum of operators with small V_1, V_2.

Let $\widetilde{\theta} \in \Theta_1$ be the point satisfying $(2N\widetilde{\theta}, -\langle\widetilde{\theta}, N\widetilde{\theta}\rangle) = u$. Then the v_1's satisfying (3.5) are contained in

$$L(\widetilde{\theta}, R^{-1/2+c\delta}) = \{\rho(\widetilde{\theta}, 1) \in \widetilde{V}_1 : 1 \leq \rho \leq 2\} + O(R^{-1/2+c\delta})$$

since (ii) means $\nabla\phi_1$ fails to be one to one only along the line $\rho(\theta_1, 1)$, $1 \leq \rho \leq 2$, for $\theta_1 \in \Theta_1$. If $(\widetilde{\theta}_1, 1)$ is transversal to $\Pi_{\xi_1, \xi_2'}$ (namely, the angle $(\widetilde{\theta}_1, 1)$ between all the tangent plane of $\Pi_{\xi_1, \xi_2'}$ is bounded away from zero), then $L(\widetilde{\theta}, R^{-1/2+c\delta}) \cap \Pi_{\xi_1, \xi_2'}(R^{-1/2})$ is contained in a ball of radius $O(R^{-1/2+c\delta})$. 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 $O(R^{c\delta})$-v_1 satisfying (3.5) since $v_1 \in \mathcal{V}_1$ are $(R^{-1/2})$-separated.

We recall from (2.17) that the normal vector of $\Pi_{\xi_1, \xi_2'}$ at (η_1, ρ_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 $(\eta, \rho), \eta = \theta_1'\rho$ is transversal to $\Pi_{\xi_1, \xi_2'}$ if and only if

$$(3.6) \qquad\qquad |N_{\theta_1, \theta_2} \cdot (\theta_1', 1)| \sim 1$$

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 Θ_1, Θ_2 to be small enough. So, $|N_{\theta_1, \theta_2} \cdot (\theta_1', 1)| \sim |N_{\theta_1, \theta 2} \cdot (\theta_1, 1)| = |(\theta_1 - \theta_2) \cdot N(\theta_1 - \theta_2)|$. 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 transverality between the null direction $\{(\theta_1, 1)\}_{\theta_1 \in \Theta_1}$ and $\Pi_{\xi_1, \xi_2'}$ for all $\xi_1 \in \widetilde{V}_1, \xi_2' \in \widetilde{V}_2$.

Therefore, from the discussion in the above we see that there are at most $O(R^{c\delta})$-v_1 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 $O(R^{c\delta})$-y_1 because T_{w_1} passes through $R^\delta q_0$. This proves our claim. $\qquad\qquad\square$

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 $(\eta_1 - \eta_2, \eta_1 + \eta_2) \to (\xi_1, \xi_2)$, instead of $\widehat{f d\sigma_{\mathcal{H}}}$ we may consider

$$Ef(x, t) = \int_{[-2,2]^2} e^{2\pi i(x \cdot \xi + t\xi_1\xi_2)} f(\xi)d\xi.$$

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. $\|EfEg\|_\infty \leq C\|f\|_1\|g\|_1$) one can easily see the following:

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

(4.1) $$|(a_1 - b_1)(a_2 - b_2)| \sim 1$$

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

(4.2) $$\|EfEg\|_{L^q(\mathbb{R}^2 \times \mathbb{R})} \leq C\|f\|_{L^p(\mathbb{R}^2)}\|g\|_{L^p(\mathbb{R}^2)}$$

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 $\|EfEg\|_p \leq C\|f\|_p\|g\|_p$ 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 $|c_1 - c_2| \sim 1$, $0 < \epsilon \ll 1$, 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 $\pm e_1$ and $\pm e_2$, then Proposition 4.1 follows from Theorem 1.1 because (4.1) is satisfied. So we may also assume $(c_1 - c_2)/|c_1 - c_2|$ is close to $\pm e_1$ or $\pm e_2$. Hence by harmless affine transformation it is sufficient for Proposition 4.1 to consider the case $V_1 = Q(-e_1, \epsilon)$, $V_2 = Q(e_1, \epsilon)$.

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 $2^{-(j+1)}$ in the usual way. 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

(4.3) $$I \times I \setminus D = \bigcup_{j \geq 1} \bigcup_{I_k^j \approx I_{k'}^j} I_k^j \times I_{k'}^j.$$

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 $I \times I$,

(4.4) $$Ef(x)Eg(x) = \sum_{j \geq 1} B_j(f,g)(x),$$

where

$$B_j(f,g)(x) = \sum_{k,k':I_k^j \approx I_{k'}^j} Ef_k^j(x)Eg_{k'}^j(x).$$

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

$$\{(\xi_1, \xi_2, \tau) : |\xi_2 - c_k^j| \leq 2^{3-j}, |\xi_1| \leq C, |\tau| \leq C\},$$

where c_k^j is the center of I_k^j. Hence one can easily see that for fixed j the Fourier supports of $\{Ef_k^j Eg_{k'}^j\}_{k,k':I_k^j \approx I_{k'}^j}$ are contained in boundedly (at most 2^4) overlapping rectangles. Using Lemma 4.2 below, we have for $1 \leq q \leq \infty$,

$$(4.5) \qquad \|B_j(f,g)\|_q \leq C \left(\sum_{k,k':I_k^j \approx I_{k'}^j} \|Ef_k^j Eg_{k'}^j\|_q^{q_*} \right)^{q_*}.$$

Lemma 4.2 ([16], 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*

$$\left(\sum_k \|F_k\|_p^{p^*}\right)^{1/p^*} \lesssim \|\sum_k F_k\|_p \lesssim \left(\sum_k \|F_k\|_p^{p_*}\right)^{1/p_*},$$

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

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

$$(4.6) \qquad \|Ef_k^j Eg_{k'}^j\|_q \leq C 2^{-2j+\frac{2j}{p}+\frac{2j}{q}} \|f_k^j\|_p \|g_{k'}^j\|_p.$$

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

$$[Ef_k^j Eg_{k'}^j](x,t) = 2^{-2j}[E(f_k^j)_j E(g_{k'}^j)_j](x_1, 2^{-j}x_2, 2^{-j}t),$$

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 $(f_k^j)_j$, $(g_{k'}^j)_j$ 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 (4.1) 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

$$(4.7) \qquad \|B_j(f,g)\|_q \leq C 2^{-2j+\frac{2j}{p}+\frac{2j}{q}} \left(\sum_{I_k^j \approx I_{k'}^j} \|f_k^j\|_p^{q_*} \|g_{k'}^j\|_p^{q_*} \right)^{1/q_*}.$$

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 $C2^{-2j+\frac{2j}{p}+\frac{2j}{q}} (\sum_k \|f_k^j\|_p^{2q_*})^{1/2q_*} (\sum_k \|g_k^j\|_p^{q_*})^{1/2q_*}$. Since I_k^j are disjoint dyadic intervals, $(\sum_k \|f_k^j\|_p^{2q_*})^{1/2q_*} \leq C\|f\|_p$ if $2q_* \geq p$. Hence we get

$$(4.8) \qquad \|B_j(f,g)\|_q \leq C 2^{-2j+\frac{2j}{p}+\frac{2j}{q}} \|f\|_p \|g\|_p$$

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 $\|B_j(f,g)\|_q \leq C 2^{-\epsilon j} \|f\|_p \|g\|_p$ for some $\epsilon > 0$ if $q > 5/3$ and $1/p + 1/q < 1$. Therefore, by (4.4) summation in j completes the proof. $\qquad \square$

4.2. **Proof of Corollary 1.4.** As before, instead of $\widehat{fd\sigma_C}$ it suffices to consider

$$\mathcal{E}f(x,t) = \int_1^2 \int_{[-2,2]^2} e^{2\pi i(x'\cdot\xi + x_3\rho + t\xi_1\xi_2/\rho)} f(\xi,\rho) d\xi d\rho, \quad x' \in \mathbb{R}^2,$$

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 $a \in \Theta_1$ and $b \in \Theta_2$ (see Theorem 1.3 for the definitions of Θ_1, Θ_2). Then if supp $f \subset \overline{V}_1 \times [1,2]$ and supp $g \subset \overline{V}_2 \times [1,2]$, there is a constant C such that

(4.9) $$\|\mathcal{E}f\mathcal{E}g\|_{L^q(\mathbb{R}^3\times\mathbb{R})} \le C\|f\|_{L^p(\mathbb{R}^3)}\|g\|_{L^p(\mathbb{R}^3)}$$

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 $1/p + 1/q < 1$, $\|\mathcal{E}_1 f \mathcal{E}_2 g\|_q \le C\|f\|_p\|g\|_p$.*

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 $\pm e_1$ or $\pm e_2$.

Observe that the change of variables $(\xi, \rho) \to (\xi + c\rho, \rho)$ for $\mathcal{E}f$ gives

(4.10) $$\mathcal{E}f(x,t) = \int_1^2 \int e^{2\pi i(L(x,t)\cdot(\xi_1,\xi_2,\rho) + t\xi_1\xi_2/\rho)} \chi_{\Theta_1}(\frac{\xi}{\rho} + c) f(\xi + c\rho, \rho) d\xi d\rho,$$

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 Θ_1, Θ_2 to $\Theta_1 - c, \Theta_2 - c$, respectively, without affecting the estimate (4.9). Hence, by an linear transform it is sufficient consider the case $\Theta_1 = Q(-e_1, \epsilon)$, $\Theta_2 = Q(e_1, \epsilon)$.

Using the decomposition (4.3), we set $f_k^j(\xi,\rho) = \chi_{I_k^j}(\xi_2/\rho)f(\xi,\rho)$, $g_k^j(\xi,\rho) = \chi_{I_k^j}(\xi_2/\rho)g(\xi,\rho)$. Then, it follows that

$$\mathcal{E}f(x)\mathcal{E}g(x) = \sum_{j\ge 1} \mathcal{B}_j(f,g)(x),$$

where $\mathcal{B}_j(f,g)(x) = \sum_{k,k':I_k^j\approx I_{k'}^j} \mathcal{E}f_k^j(x)\mathcal{E}g_{k'}^j(x)$. Hence it is sufficient to show that if $q > \frac{3}{2}$ and $1/p + 1/q < 1$, for some $\epsilon > 0$,

(4.11) $$\|\mathcal{B}_j(f,g)\|_q \le C2^{-\epsilon j}\|f\|_p\|g\|_q.$$

If $I_k^j \approx I_{k'}^j$, the Fourier support of $\mathcal{E}f_k^j\mathcal{E}g_{k'}^j$ is contained in the set

$$\{(\xi_1, \xi_2, \rho, \tau) : |\xi_2 - c_0\rho| \le 2^{3-j}, 1 \le \rho < 2, |\xi_1| \le C, |\tau| \le C\},$$

where c_0 is the center of the smallest interval containing both I_k^j, $I_{k'}^j$. So we see that the Fourier supports of $\{\mathcal{E}f_k^j\mathcal{E}g_{k'}^j\}_{I_k^j\approx I_{k'}^j}$ are contained in boundedly (at most 2^6) overlapping rectangles. Hence, using Lemma 4.2, we have for $1 \le q \le \infty$,

(4.12) $$\|\mathcal{B}_j(f,g)\|_q \le C \left(\sum_{I_k^j\approx I_{k'}^j} \|\mathcal{E}f_k^j\mathcal{E}g_{k'}^j\|_q^{q_*} \right)^{1/q_*}.$$

Now we claim that if $I_k^j \approx I_{k'}^j$ and $(1/p, 1/q)$ is contained in the interior of Q_0,

$$(4.13) \qquad \|\mathcal{E}f_k^j \mathcal{E}g_{k'}^j\|_q \leq C 2^{-2j+\frac{2j}{p}+\frac{2j}{q}} \|f_k^j\|_p \|g_{k'}^j\|_p.$$

Once this is established, by the same argument used before (see (4.7)), (4.12) and (4.13) give $\|\mathcal{B}_j(f,g)\|_q \leq C 2^{-2j+\frac{2j}{p}+\frac{2j}{q}} \|f\|_p \|g\|_p$ provided $(1/p, 1/q)$ is contained in the interior of Q_0 and $2q_* \geq p$. 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 $\mathcal{E}f_k^j$ and $\mathcal{E}g_{k'}^j$ (see (4.10)), 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. By re-scaling $(\xi_1, \xi_2) \to (\xi_1, 2^{-j}\xi_2)$ we see that

$$[\mathcal{E}f_k^j \mathcal{E}g_{k'}^j](x, t) = 2^{-2j}[\mathcal{E}(f_k^j)_j \mathcal{E}(g_{k'}^j)_j](x_1, 2^{-j}x_2, x_3, 2^{-j}t),$$

where $(f_k^j)_j(\xi_1, \xi_2, \rho) = f_k^j(\xi_1, 2^{-j}\xi_2, \rho)$ and $(g_{k'}^j)_j(\xi_1, \xi_2, \rho) = g_{k'}^j(\xi_1, 2^{-j}\xi_2, \rho)$. Then $(f_k^j)_j$, $(g_{k'}^j)_j$ are supported in

$$\overline{V}_1 = \{(\xi, \rho) : \xi/\rho \in [-1-\epsilon, -1+\epsilon] \times [-1, -1/2], 1 \leq \rho \leq 2\},$$
$$\overline{V}_2 = \{(\xi, \rho) : \xi/\rho \in [1-\epsilon, 1+\epsilon] \times [1/2, 1], 1 \leq \rho \leq 2\},$$

respectively. So, (4.1) holds for all $a \in \Theta_1$ and $b \in \Theta_2$. Therefore by (4.9) and rescaling we get (4.13).   □

To finish the proof of Corollary 1.4, we need to remove the separation condition dist $(\Theta_1, \Theta_2) \sim 1$ from Proposition 4.3. 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

$$(4.14) \qquad [-2, 2]^2 \times [-2, 2]^2 \setminus D = \bigcup_{j \geq 1} \bigcup_{Q_k^j \approx Q_{k'}^j} Q_k^j \times Q_{k'}^j,$$

where Q_k^j are dyadic squares of side length $2^{-(j+1)}$ and $Q_k^j \approx Q_{k'}^j$ means that Q_k^j, $Q_{k'}^j$ are not adjacent but have adjacent parent squares of side length 2^{-j}.

Set

$$\widetilde{\mathcal{B}}_j(f, g)(x) = \sum_{k,k':Q_k^j \approx Q_{k'}^j} \mathcal{E}[f_k^j](x)\mathcal{E}[g_{k'}^j](x),$$

where $[f_k^j](\xi, \rho) = \chi_{Q_k^j}(\xi/\rho)f(\xi, \rho)$. Since $\mathcal{E}f\mathcal{E}f = \sum_{j \geq 1} \widetilde{\mathcal{B}}_j(f, f)$ from (4.14), it follows that $\|Ef\|_q^2 \leq C \sum_{j \geq 1} \|\widetilde{\mathcal{B}}_j(f, f)\|_{q/2}$. Hence, it is sufficient to show that if $q > 3$ and $1/p + 2/q < 1$, for some $\epsilon > 0$

$$(4.15) \qquad \|\widetilde{\mathcal{B}}_j(f, g)\|_{q/2} \leq C 2^{-\epsilon j} \|f\|_p \|g\|_p.$$

By repeating the argument used in the proofs of Proposition 4.1, 4.3 and using the orthogonality among $\{\mathcal{E}[f_k^j]\mathcal{E}[g_{k'}^j]\}_{Q_k^j \approx Q_{k'}^j}$ (see (4.12)), for (4.15) it is sufficient to show that if $Q_k^j \approx Q_{k'}^j$, for $q > \frac{3}{2}$ and $1/p + 1/q < 1$,

$$(4.16) \qquad \|\mathcal{E}[f_k^j]\mathcal{E}[g_{k'}^j]\|_q \leq C 2^{-4j+\frac{4j}{p}+\frac{4j}{q}} \|[f_k^j]\|_p \|[g_{k'}^j]\|_p.$$

By linear transformation Q_k^j, $Q_{k'}^j$ can be assumed to be contained in $Q(0, 2^{1-j})$ (see (4.10)). Re-scaling $\xi \to 2^{-j}\xi$, the squares $Q_k^j, Q_{k'}^j$ with $Q_k^j \approx Q_{k'}^j$ are moved to cubes in $[-2, 2]^2$ with distance ~ 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

[1] B. Barcelo, *On the restriction of the Fourier transform to a conical surface*, Trans. Amer. Math. Soc. **292** (1985), 321-333. MR0805965 (86k:42023)

[2] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal. **1** (1991), 147-187. MR1097257 (92g:42010)

[3] J. Bourgain, *On the restriction and multiplier problems in $\mathbb{R}^3$*, in *Geometric aspects of functional analysis-seminar* 1989-90, Lecture Notes in Math. vol. 1469, Springer-Berlin, (1991), 179-191. MR1122623 (92m:42017)

[4] J. Bourgain, *Estimates for cone multipliers,* Operator Theory: Advances and Applications **77** (1995), 41-60. MR1353448 (96m:42022)

[5] A. Carbery, *Restriction implies Bochner-Riesz for paraboloids*, Math. Proc. Camb. Phil. Soc. **111** (1992), 525-529. MR1151328 (93b:42024)

[6] A. Greenleaf, *Principal curvature and harmonic analysis*, Indiana Univ. Math. J. **30** (1981), 519-537. MR0620265 (84i:42030)

[7] S. Lee, *Some sharp bounds for the cone multiplier of negative order in $\mathbb{R}^3$*, Bull. London Math. Soc. **35** (2003), 373-390. MR1960948 (2004c:42027)

[8] S. Lee, *Endpoint estimates for the circular maximal function*, Proc. Amer. Math. Soc. **131** (2003), 1433-1442. MR1949873 (2003k:42035)

[9] S. Lee, *Improved bounds for Bochner-Riesz and maximal Bochner-Riesz operators*, Duke Math. Journ. **122** (2004), 105-235. MR2046812

[10] E. Stein, *Oscillatory integrals in Fourier analysis*, in *Beijing lectures in harmonic analysis*, Ann. of Math. Study # 112, Princeton Univ. Press, Princeton (1986), 307-355. MR0864375 (88g:42022)

[11] E. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton University Press, Princeton (1993). MR1232192 (95c:42002)

[12] T. Tao, *Endpoint bilinear restriction theorems for the cone and some sharp null from estimates*, Math. Z. **238** (2001), 215-268. MR1865417 (2003a:42010)

[13] T. Tao, *A Sharp bilinear restriction estimate for paraboloids*, Geom. Funct. Anal. **13** (2003), 1359-1384. MR2033842 (2004m:47111)

[14] T. Tao, *Recent progress on the restriction conjecture*, Park City Proceedings, to appear.

[15] T. Tao and A. Vargas, *A bilinear approach to cone multipliers I. Resriction estimates*, Geom. Funct. Anal. **10** (2000), 185-215. MR1748920 (2002e:42012)

[16] T. Tao and A. Vargas, *A bilinear approach to cone multipliers. II. Application*, Geom. Funct. Anal. **10** (2000), 216-258. MR1748921 (2002e:42013)

[17] T. Tao, A. Vargas and L. Vega, *A bilinear approach to the restriction and Kakeya conjecture*, J. Amer. Math. Soc. **11** (1998), 967-1000. MR1625056 (99f:42026)

[18] A. Vargas, *Restriction theorems for a surface with negative curvature*, Math. Z. **249** (2005), 97-111. MR2106972 (2005f:42029)

[19] T. Wolff, *Recent work connected with the Kakeya problem*, Prospects in mathmatics, (Princeton. N.J., 1996), 129-162, Amer. Math. Soc. Providence, RI, 1999. MR1660476 (2000d:42010)

[20] T. Wolff, *Local smoothing type estimates on L^p for large p*, Geom. Funct. Anal. **10** (2000), 1237-1288. MR1800068 (2001k:42030)

[21] T. Wolff, *A sharp cone restriction estimate*, Annals of Math. **153** (2001), 661-698. MR1836285 (2002j:42019)

DEPARTMENT OF MATHEMATICS, POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY, POHANG 790-784, KOREA

E-mail address: sanghyuk@postech.ac.kr

Current address: Department of Mathematics, University of Wisconsin-Madison, Madison, Wisconsin 53706-1388

E-mail address: slee@math.wisc.edu