Singular spherical maximal operators on a class of degenerate two-step nilpotent Lie groups

Naijia Liu · Lixin Yan

Received: 14 September 2022 / Accepted: 25 March 2023 / Published online: 24 April 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
Let \( G \cong \mathbb{R}^d \ltimes \mathbb{R} \) be a finite-dimensional two-step nilpotent group with the group multiplication \((x, u) \cdot (y, v) \to (x + y, u + v + x^T J y)\) where \( J \) is a skew-symmetric matrix satisfying a degeneracy condition with \( 2 \leq \text{rank} \ J < d \). Consider the maximal function defined by
\[
M f (x, u) = \sup_{t > 0} \left| \int_{\Sigma} f (x - ty, u - tx^T J y) d\mu (y) \right|
\]
where \( \Sigma \) is a smooth convex hypersurface and \( d\mu \) is a compactly supported smooth density on \( \Sigma \) such that the Gaussian curvature of \( \Sigma \) is nonvanishing on \( \text{supp} d\mu \). In this paper we prove that when \( d \geq 4 \), the maximal operator \( M \) is bounded on \( L^p (G) \) for the range \( (d - 1)/(d - 2) < p \leq \infty \).

Keywords  Singular spherical maximal operator · Degenerate two step nilpotent Lie groups · Gaussian curvature · Oscillatory integrals

Mathematics Subject Classification  42B25 · 22E30 · 43A80

1 Introduction

Let \( G \) be a finite-dimensional two-step nilpotent group which we may identify with its Lie algebra \( \mathfrak{g} \) by the exponential map. We assume that \( \mathfrak{g} \) splits as a direct sum \( \mathfrak{g} = \mathfrak{w} \oplus \mathfrak{z} \) so that
\[
[\mathfrak{w}, \mathfrak{w}] \subset \mathfrak{z}, \quad [\mathfrak{w}, \mathfrak{z}] = \{0\},
\]
and that \( \dim(\mathfrak{w}) = d \), \( \dim(\mathfrak{z}) = 1 \). In exponential coordinates \( (x, u) \in \mathbb{R}^d \times \mathbb{R} \), the group multiplication is given by
\[
(x, u) \cdot (y, v) \defeq (x + y, u + v + x^T J y),
\]
where $J$ is a skew-symmetric matrix acting on $\mathbb{R}^d$ (i.e. $J^T = -J$). The most prominent examples are the Heisenberg groups $\mathbb{H}^n$ which arise when $d = 2n$ and $J = \frac{1}{2} J_{2n}$ with

$$J_{2n} := \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$$

(1.1)
is the standard symplectic matrix on $\mathbb{R}^{2n}$, see [9]. There is a natural dilation structure relative to $\mathfrak{w}$ and $\mathfrak{z}$, namely for $X \in \mathfrak{w}$ and $U \in \mathfrak{z}$ we consider the dilations

$$\delta_t : (X, U) \to (tX, t^2 U).$$

With the identification of the Lie algebra with the group, $\delta_t$ becomes an automorphism of the group.

Let $\Sigma$ be a smooth convex hypersurface in $\mathfrak{w}$ and let $\mu$ be a compactly supported smooth density on $\Sigma$. Throughout we shall make the following

**Assumption** The Gaussian curvature of $\Sigma$ does not vanish on the support of $\mu$.

Define the dilate $\mu_t$ by

$$\langle \mu_t, f \rangle := \int_{\Sigma} f(tx, 0) d\mu(x).$$

(1.2)

Consider the maximal operator

$$\mathfrak{M} f(x, u) := \sup_{t > 0} \left| f * \mu_t(x, u) \right|,$$

(1.3)

where the convolution is given by

$$f * \mu_t(x, u) := \int_{\Sigma} f(x - ty, u - tx^T J y) d\mu(y).$$

(1.4)

The operator $\mathfrak{M}$ can be understood as an analogue of the classical (Euclidean) spherical maximal function of Stein [16–18] on $\mathbb{R}^n$ (see also Bourgain [2]). In the case of the sphere $\Sigma = S^{2n-1} = \{x : |x| = 1\}$ on the noncentral part of the Heisenberg group $\mathbb{H}^n$ and $\mu$ denotes the normalised surface measure on the sphere $S^{2n-1}$, the maximal function (1.3) was studied by Nevo and Thangavelu in [10], and they obtained $L^p(\mathbb{H}^n)$ boundedness for $d = 2n \geq 4$ and $p > (d - 1)/(d - 2)$ by using spectral methods. Later, improving the results in [10], Narayanan and Thangavelu [11] obtained the optimal range $p > d/(d - 1)$ by modifying the argument in [10] and combining it with Stein’s square function method. In [9], Müller and Seeger independently proved this optimal range for $p$ by using Fourier integral operators in a study which concerns more general surfaces $\Sigma$ with a nonvanishing rotational curvature in “non-degenerate” two-step nilpotent groups including all $\mathbb{H}$-type groups. This topic has attracted a lot of attention in the last decades, and has been a very active research topic in harmonic analysis; see, for example [1, 4, 6, 8, 9, 11–15] and the references therein for further details.

Throughout this paper, we assume that $G$ satisfies the degenerate hypothesis, i.e., for every nonzero linear functional $\omega \in \mathfrak{z}^*$, the rank $\kappa$ of the bilinear form $J_\omega : (X, Y) \to \omega([X, Y])$, which maps $\mathfrak{w} \times \mathfrak{w}$ to $\mathbb{R}$, satisfies $0 < \kappa < d$. Note in exponential coordinates the degenerate hypothesis is equivalent with

$$0 < \kappa = \text{rank } J < d,$$

and $d$ may be odd. The skew symmetry of $J_\omega$ implies that rank $J$ is even, and so we can assume that $2 \leq \kappa < d$. In this paper, we prove the following result.
Theorem 1.1 Suppose $d \geq 4$ and $p > (d - 1)/(d - 2)$. Then the maximal operator $\mathfrak{M}$ extends to a bounded operator on $L^p(G)$ and there exists a constant $C > 0$ such that

$$\|\mathfrak{M}f\|_{L^p(G)} \leq C\|f\|_{L^p(G)}.$$

As a consequence, \( \lim_{t \to 0} f \ast \mu_t(x, u) = cf(x, u) \) a.e. for all $f \in L^p(G)$, with $c = \int d\mu$.

Remark 1.2 (i) The range $p > d/(d - 1)$ is necessary in Theorem 1.1 and it can be seen by testing $\mathfrak{M}$ on the function $f$ given by $f(y, v) = |y|^{-d}(\log |y|)^{-1}\chi(y, v)$ with a suitable nonnegative cutoff function $\chi$, which equals to 1 if $|\langle y, v \rangle| \leq \frac{1}{2}$ and $\text{supp} \chi \subseteq B(0, 1)$ (see [17, Chapter XI] and Sect. 2). Our method gives a positive result only in the range $p > (d - 1)/(d - 2)$, while the range $d/(d - 1) < p \leq (d - 1)/(d - 2)$ for $d \geq 4$ remains open.

(ii) Our method recovers the results of [9, 11] in the partial range $p > (d - 1)/(d - 2)$ when $J$ is invertible. Also, higher values of $\kappa$ ($= \text{rank } J$) do not yield a better range of $p$ in Theorem 1.1.

(iii) Theorem 1.1 can be generalized to the maximal operators associated to surfaces of codimension $m + 1$ with $m > 1$ (see [9]). The result should remain true for $p > (d - 2)/(d - 3)$ whenever $\text{rank } J < d$ and $d \geq 5$.

Our approach is inspired by the work of Müller–Seeger [9]. Like [9], we use square functions and almost orthogonality to boil down the problem of obtaining $L^2$ and weak $(1,1)$ type estimates for a family of dyadic Littlewood–Paley projections of Fourier integral operators associated to $\mathfrak{M}$. The proof of weak $(1,1)$ type estimate proceeds in a standard way. However, for $L^2$ estimate, we employ a different strategy. The corresponding estimate in [9] was established by proving that the canonical relations of the corresponding oscillatory integral operators exhibit a two-sided fold singularity, and then applying the results from [5] (see also [3]) providing sharp $L^2$ estimates in this situation. In this paper, we do not deal with these types of singularities. Instead, we divide our analysis into different cases based on the relative sizes $|\lambda_1|$ and $|\lambda_2|$ respectively of the two frequency variables $\sigma$ and $\tau$ (see the two cases in Lemmas 4.3 and 5.2), and carefully analyze the rank of mixed Hessian of the phase function associated with the relevant oscillatory integral, using Lemma 3.1, to establish Hörmander-type $L^2$ estimates. These $L^2$ oscillatory integral estimates are good enough to establish $L^p$ boundedness of the maximal operator $\mathfrak{M}$ on $L^p(G)$ for $p > (d - 1)/(d - 2)$ with $p \geq 4$.

This article is organized as follows. In Sect. 2, we introduce some notations and preliminary lemmas. We reduce our main theorem to Theorem 2.1. In Sect. 3, we prove a useful lemma, which plays an important role in the proof of Theorem 2.1. The proof of Theorem 2.1 will be given in Sects. 4 and 5 by reducing the estimates for the averages to estimates for oscillatory integral operators.

Throughout, the letters “$c$” and “$C$” will denote (possibly different) constants that may depend only on $d$, $p$, $J$, $\Sigma$ and $\mu$.

2 Preliminaries

The Fourier transform is defined for $f \in \mathcal{S}(\mathbb{R}^n)$ by

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} e^{-ix\xi} f(x) dx$$

and this definition is then extended to $L^2(\mathbb{R}^n)$ and to bounded measure in the usual way. The inverse Fourier transform is denoted by $\mathcal{F}^{-1}(f) = \hat{f}$.
We recall the definition of convolution
\[ f \ast g(x, u) = \int_{\mathbb{R}^{d+1}} f(y, v)g((y, v)^{-1}(x, u))dydv \]
\[ = \int_{\mathbb{R}^{d+1}} f(x - y, u - v - x^T J y)g(y, v)dydv, \]
where \( J \) is skew-symmetric \( d \times d \) matrix acting on \( \mathbb{R}^d \) with \( 2 \leq \text{rank} \ J < d \) by our degenerate hypothesis. By a rotation we can suppose
\[ J = \begin{pmatrix} J_\kappa & 0 \\ 0 & 0 \end{pmatrix} \]  
(2.1)
for a skew-symmetric \( \kappa \times \kappa \) matrix \( J_\kappa \) satisfies rank \( J_\kappa = \kappa \in [2, d) \).

Recall that \( \Sigma \) is a smooth convex hypersurface in \( \mathfrak{w} \) and let \( \mu \) be a compactly supported smooth density on \( \Sigma \). The Gaussian curvature of \( \Sigma \) does not vanish on the support of \( \mu \). In the following, we use notation
\[ x = (x_1, x'), \quad x' = (x_2, \ldots, x_d) \in \mathbb{R}^{d-1}; \]
and also
\[ x = (x'', x_d), \quad x'' = (x_1, \ldots, x_{d-1}) \in \mathbb{R}^{d-1}. \]

By localizations, we only consider that the projection of \( \Sigma \) to \( \mathfrak{w} \) are given by the following two cases since other cases can be studied similarly:

1. \( x_1 = \Gamma_1(x'), x' = (x_2, \ldots, x_d) \) with \( \Gamma_1 \in C^\infty(\mathbb{R}^{d-1}) \) so that \( \mu \) is supported in a small neighborhood of \( (\Gamma_1(x'_0), x'_0) \) for some \( x'_0 \in \mathbb{R}^{d-1}; \)
2. \( x_d = \Gamma_2(x''), x'' = (x_1, \ldots, x_{d-1}) \) with \( \Gamma_2 \in C^\infty(\mathbb{R}^{d-1}) \) so that \( \mu \) is supported in a small neighborhood of \( (x''_0, \Gamma_2(x''_0)) \) for some \( x''_0 \in \mathbb{R}^{d-1} \).

Using the Fourier inversion formula for Dirac measures we may write
\[ \mu_1(x, u) = \chi_{\mu_1}(x, u) \int_{\mathbb{R} \times \mathbb{R}} e^{i(\sigma(x_1 - \Gamma_1(x')) + \tau u)}d\sigma d\tau \]  
(2.2)
and
\[ \mu_2(x, u) = \chi_{\mu_2}(x, u) \int_{\mathbb{R} \times \mathbb{R}} e^{i(\sigma(x_d - \Gamma_2(x'')) + \tau u)}d\sigma d\tau \]  
(2.3)
in the above cases (1) and (2), respectively. Here, \( \chi_{\mu_i}(i = 1, 2) \) is a smooth compactly supported function and the integral convergence in the sense of oscillatory integrals (thus in the sense of distributions). Then Theorem 1.1 follows from the following result:

**Theorem 2.1** Let \( \mu_1 \) and \( \mu_2 \) be given as above. Then we have

(i) Suppose \( d \geq 4 \) and \( p > (d - 1)/(d - 2) \). We have
\[ \left\| \sup_{t > 0} |f \ast (\mu_1)_t| \right\|_{L^p(\mathbb{R}^{d+1})} \leq C \| f \|_{L^p(\mathbb{R}^{d+1})}; \]  
(2.4)

(ii) Suppose \( d \geq 3 \). We have
\[ \left\| \sup_{t > 0} |f \ast (\mu_2)_t| \right\|_{L^p(\mathbb{R}^{d+1})} \leq C \| f \|_{L^p(\mathbb{R}^{d+1})}; \]  
(2.5)
holds if and only if and \( p > d/(d - 1) \).
The proof of (i) and (ii) of Theorem 2.1 will be given in Sects. 4 and 5, respectively. As pointed out in the introduction, \( p > \frac{d}{d-1} \) is necessary in the above theorem and it can be seen by testing \( \mathcal{M} \) on the function \( f \) given by \( f(y, v) = |y|^{1-d}(\log |y|)^{-1} \chi(y, v) \) with a suitable cutoff function \( \chi \), which equals to 1 if \( |(y, v)| \leq \frac{1}{2} \) and \( \text{supp} \chi \subseteq B(0, 1) \). Let \( \mu \) be the induced Lebesgue measure on the sphere \( S^{d-1} \). Then we have

\[
|f * \mu_t(x, u)| \geq -\int_{|y|=1, |x-ty| \leq \frac{|x|}{10}} |x-ty|^{1-d}(\log |x-ty|)^{-1} d\mu(y)
\]

\[
\simeq -\int_0^{\frac{|x|}{10}} r^{1-d}(\log r)^{-1} dr = \infty
\]

if \( |(x, u)| \leq (10(\|J\|+1))^{-1} \) and \( t = |x|, d \geq 4 \). So \( \mathcal{M} f = \sup_{t>0} |f * \sigma_t| \notin L^p(\mathbb{R}^{d+1}) \) for all \( 1 \leq p < \infty \). However, \( f \in L^p(\mathbb{R}^{d+1}) \) for \( 1 \leq p \leq \frac{d}{d-1} \), so \( \mathcal{M} \) is unbounded on \( L^p(\mathbb{R}^{d+1}) \) for \( 1 \leq p \leq d/(d-1) \).

In the end of this section, we state the following two lemmas, which will be useful in the proof of Theorem 2.1.

**Lemma 2.2** Suppose that

\[
\sup_{s \in [1,2]} \left( \sum_{n \in \mathbb{Z}} \| F_n(\cdot, s) \|_2^2 \right)^{1/2} \leq A_1
\]

and

\[
\sup_{s \in [1,2]} \left( \sum_{n \in \mathbb{Z}} \left\| \frac{d}{ds} F_n(\cdot, s) \right\|_2^2 \right)^{1/2} \leq A_2.
\]

Then

\[
\left\| \sup_n \sup_{s \in [1,2]} |F_n(\cdot, s)| \right\|_2 \leq C \left( A_1 + \sqrt{A_1 A_2} \right).
\]

**Proof** For the proof, see [9, Lemma 3.1], [17, p. 499]. \( \square \)

We now state an almost orthogonality lemma, the Cotlar–Stein Lemma.

**Lemma 2.3** Suppose \( 0 < \epsilon < 1, A \leq B/2 \) and let \( \{T_n\}_{n=1}^\infty \) be a sequence of bounded operators on a Hilbert space \( H \) so that the operator norms satisfy

\[
\|T_n\| \leq A
\]

and

\[
\|T_n T_{n'}^*\| \leq B^2 2^{-\epsilon |n-n'|}.
\]

Then for all \( f \in H \)

\[
\left( \sum_{n=1}^\infty \| T_n f \|^2 \right)^{1/2} \leq CA \sqrt{\epsilon^{-1} \log(B/A) \|f\|}.
\]

**Proof** For the proof, see [9, Lemma 3.2]. \( \square \)
3 A useful lemma

Let \( m, n \in [1, d] \). For \( m \times n \) matrix \( A \), we define its norm \( \|A\| \) as

\[
\|A\| = \sup_{X \in \mathbb{R}^n, \|X\| = 1} |AX|,
\]

where \( X = (x_1, \ldots, x_n) \) and \( |X| = \left( \sum_{i=1}^n x_i^2 \right)^{1/2} \). We define \( P_n, E_n \) as the \( d \times n \) matrices:

\[
P_n = \begin{pmatrix} 0 & I_n \end{pmatrix} \quad \text{and} \quad E_n = \begin{pmatrix} I_n \\ 0 \end{pmatrix}
\]

(3.1)

and \( P_n^T, E_n^T \) as their transposes.

In this section, we first prove the following lemma, which plays an essential role in the proof of (i) of Theorem 2.1.

**Lemma 3.1** (i) Let \( J \) be a \( \kappa \times \kappa \) skew-symmetric matrix with rank \( J = \kappa \geq 2 \). For every \( \kappa \times \kappa \) semi-positive definite matrix \( G \) and \( M > 0 \), there exists a constant \( c_0 = c_0(J, G, M) > 0 \) such that

\[
\inf_{|a| \leq M, \|X\| = 1} \| (aG + J)X \| \geq c_0;
\]

(3.2)

(ii) Let \( J \) be a \( d \times d \) skew-symmetric matrix satisfying \( \|Je_1\| > 0 \). For every \( (d-1) \times (d-1) \) positive definite matrix \( Q \) and \( \delta > 0 \), there exists a constant \( c_1 = c_1(J, Q, \delta) > 0 \) such that

\[
\inf_{|a| \geq \delta} \inf_{Y \in \mathbb{R}^d, \|Y\| = 1} \| (P_{d-1} Q P_{d-1}^T + aJ)Y \| \geq c_1.
\]

(3.3)

**Proof** Since \( G \) is a \( \kappa \times \kappa \) semi-positive definite matrix, we can write \( G = B^T B \) for some matrix \( B \). First, let us show that

\[
\inf_{|a| \geq \delta, \|X\| = 1} \| (G + aJ)X \| \geq c_2 := \min \left\{ \frac{\delta J^{-1} - 1}{2}, \frac{\delta^2 J^{-1} - 1}{4} \right\}.
\]

(3.4)

We argue (3.4) by contradiction. Assume that there exist some \( |a| \geq \delta \) and \( \|X\| = 1 \) such that \( \| (G + aJ)X \| < c_2 \). Since \( J \) is a skew-symmetric matrix, we have \( X^T JX = 0 \). Hence,

\[
\|BX\|^2 = |X^T GX| = |X^T (G + aJ)X| \leq \|(G + aJ)X\| < c_2.
\]

From this, we see that \( \|GX\| \leq \|B^T\| \|BX\| < \|B^T\| \sqrt{c_2} \). By assumption, it tells us that

\[
\|JX\| \leq |a|^{-1} \left( \|(G + aJ)X\| + \|GX\| \right) < \delta^{-1} (c_2 + \|B^T\| \sqrt{c_2}) \leq \|J^{-1}\|^{-1}.
\]

However, \( \det J \neq 0 \) and \( 1 = |X| = \|J^{-1} JX\| \leq \|J^{-1}\| \|JX\| \) and so \( \|JX\| \geq \|J^{-1}\|^{-1} \). This will be our desired contradiction.

We now apply (3.4) to prove (3.2). By (3.4),

\[
\|(aG + J)X\| = |a| \|(G + a^{-1} J)X\| \geq |a| c_2 \geq \frac{c_2 \|J^{-1}\|^{-1}}{c_2 + \|G\|}.
\]

whenever \( \|J^{-1}\|^{-1}/(c_2 + \|G\|) \leq |a| \leq M \). On the other hand, we use estimate \( \|aGX\| \leq |a| \|G\| \) to get that for \( |a| \leq \|J^{-1}\|^{-1}/(c_2 + \|G\|) \),

\[
\|(aG + J)X\| \geq \|JX\| - \|aGX\| \geq \|J^{-1}\|^{-1} - |a| \|G\| \geq \frac{c_2 \|J^{-1}\|^{-1}}{c_2 + \|G\|}.
\]
thereby concluding the proof of (3.2) with $c_0 := c_2 \|J^{-1}\|/(c_2 + \|G\|)$.

Next let us show (3.3). Since $\|Je_1\| > 0$, we have $\|J\| > 0$. Let $Y \in \mathbb{R}^d$ and $|Y| = 1$. If $|Y \pm e_1| \leq \frac{\|Je_1\|}{2\|Q\|}$, we have
\[
\|JY\| \geq \|Je_1\| - \|J(Y \pm e_1)\| \geq \|Je_1\| - \|J\| \cdot |Y \pm e_1| \geq \frac{\|Je_1\|}{2}.
\]
If $|Y \pm e_1| \leq \frac{\|Je_1\|}{4\|Q\|}$, we have
\[
\|P_{d-1} Q P_{d-1}^T Y\| = \|P_{d-1} Q P_{d-1}^T (Y \pm e_1)\| \leq \|P_{d-1} Q P_{d-1}^T\| \cdot |Y \pm e_1| = \|Q\| \cdot |Y \pm e_1| \leq \frac{\|Je_1\|}{4}.
\]

Let $\epsilon := \min \left\{ \frac{\|Je_1\|}{2\|J\|}, \frac{\|Je_1\|}{4\|Q\|} \right\} > 0$. If $|Y \pm e_1| \leq \epsilon$, we have
\[
\left\| (P_{d-1} Q P_{d-1}^T + aJ)Y \right\| \geq \delta \|JY\| - \|(P_{d-1} Q P_{d-1}^T)Y\| \geq \frac{\delta \|Je_1\|}{4} \tag{3.5}
\]
for all $|a| \geq \delta$. If $|Y \pm e_1| \geq \epsilon$, we have $|Y'| \geq C\epsilon$ because $|Y| = 1$, which, together with $J^T = -J$, tells us that for all $a \in \mathbb{R}$,
\[
\left\| (P_{d-1} Q P_{d-1}^T + aJ)Y \right\| \geq \|((P_{d-1} Q P_{d-1}^T + aJ)Y, Y)\| = \|((P_{d-1} Q P_{d-1}^T, Y)Y\|\geq C Q \epsilon^2. \tag{3.6}
\]

From (3.5) and (3.6), we complete the proof with
\[
c := \min \left\{ \frac{\delta \|Je_1\|}{4}, C Q \epsilon^2 \right\} = \min \left\{ \frac{\delta \|Je_1\|}{4}, C Q \frac{\|Je_1\|^2}{4\|J\|^2}, C Q \frac{\delta^2 \|Je_1\|^2}{16\|Q\|^2} \right\}.
\]
The proof is complete. \qed

4 Proof of (i) of Theorem 2.1

In this section, we will show that (i) of Theorem 2.1. That is, for $d \geq 4$, the maximal operator $\sup_{r>0} |f \ast (\mu_1)(x, u)|$ is bounded on $L^p(\mathbb{R}^{d+1})$ for $p > (d - 1)/(d - 2)$ where $\mu_1$ is given as in (2.2), i.e.,
\[
\mu_1(x, u) = \chi_{\mu_1}(x, u) \int_{\mathbb{R} \times \mathbb{R}} e^{i(s(x_1 - \Gamma_1(x')) + \tau u)} d\sigma d\tau
\]
with $x' = (x_2, \ldots, x_d)$, $\Gamma_1 \in C^\infty(\mathbb{R}^{d-1})$ and $\mu_1$ is supported in a small neighborhood of $(\Gamma_1(x'_0, x_0))$ for some $x'_0 \in \mathbb{R}^{d-1}$.

Let $\varphi \in C_0^\infty(\mathbb{R})$ be an even smooth function such that $\varphi = 1$ if $|s| \leq 1$ and such that the support of $\varphi$ is contained in $(-2, 2)$. Let $\phi(s) = \varphi(s) - \varphi(2s)$ and let, for $k \geq 1$, $\phi_k(s) := \phi(2^{-k}s)$. Let $\phi_0(s) := \varphi(s)$. Then
\[
\text{supp } \phi_0 \subseteq \{s : |s| \leq 2\}, \quad \text{and } \text{supp } \phi_k \subseteq \{s : 2^{k-1} \leq |s| \leq 2^{k+1}\}
\]
for $k \geq 1$, and we have
\[
1 = \sum_{k=0}^\infty \phi_k(s), \quad \forall s \in \mathbb{R}. \tag{4.1}
\]
Since $\mu_1$ has a sufficiently small support, we can choose a smooth nonnegative function $\chi_1$ defined on $\mathbb{R}^d$, which is equal to 1 on $\text{supp} \mu_1$ and also has a small support. By the curvature hypothesis and the convexity of $\Sigma$, we can further assume $\partial_\nu^2 \Gamma_1(x')$ is positive definite for all $x \in \text{supp} \chi_1$ such that
\[
\inf_{x \in \text{supp} \chi_1} \left| \det \left( \partial_\nu^2 \Gamma_1(x') \right) \right| \geq c_H^{(1)}
\] (4.2) for some $c_H^{(1)} > 0$. Define for $j, k \geq 0$,
\[
K_{j,k}^i(x, u) = \chi_1(x) \int_{\mathbb{R} \times \mathbb{R}} e^{i \sigma(x_1 - \Gamma_1(x'))} e^{i \tau u} \phi_j(\sigma) \phi_k(\tau) d\sigma d\tau
\]
\[
= \chi_1(x) \mathcal{F}^{-1}(\phi_j)(x_1 - \Gamma_1(x')) \mathcal{F}^{-1}(\phi_k)(u)
\] (4.3) such that
\[
\mu_1(x, u) = \chi_1(x) \chi_{\mu_1}(x, u) \delta(x_1 - \Gamma_1(x')) \delta(u) = \chi_{\mu_1}(x, u) \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} K_{j,k}^i(x, u).
\]
For $t > 0$, define the dilates
\[
K_{j,k}^i(t^{-1}x, t^{-2}u) = t^{-(d+2)} K_{j,k}(t^{-1}x, t^{-2}u).
\] (4.4)
These, together with (2.2), tell us that
\[
|f \ast (\mu_1)_t| \leq C |f| \ast \left( \chi_1(x) \delta(x_1 - \Gamma_1(x')) \delta(u) \right)_t = C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} |f| \ast K_{j,k}^i.
\] (4.5)
For $j = k = 0$, it can be verified that $|K_{0,0}^1(x, u)| \leq C_N (1 + |x| + |u|)^{-N}$ for any $N > 0$, and so $\sup_{t > 0} |f \ast K_{i}^0,0 \ast |$ is controlled by the appropriate variant of the Hardy–Littlewood maximal function and therefore (see [17]) we have the inequality
\[
\left\| \sup_{t > 0} |f \ast K_{i}^0,0 \ast | \right\|_{L_p(\mathbb{R}^{d+1})} \leq C \|f\|_{L_p(\mathbb{R}^{d+1})}, \quad 1 < p \leq \infty.
\] (4.6)
Consider $j + k \geq 1$. By definition of $K_{j,k}^i$, we see that if $k = 0$ and $j \geq 1$, then it follows by integration by parts,
\[
\left| \int_{\mathbb{R}^{d+1}} K_{j,0}^i(x, u) dx du \right| \leq C_N 2^{-Nj}.
\] (4.7)
If $k > 0$, then
\[
\int_{\mathbb{R}^{d+1}} K_{j,k}^i(x, u) dx du = \int_{\mathbb{R}^{d+1}} \chi_1(x) \mathcal{F}(\phi_j)(x_1 - \Gamma_1(x')) \mathcal{F}(\phi_k)(u) dx du = 0.
\]
Choose a nonnegative function $b \in C^\infty_c(\mathbb{R}^{d+1})$ such that $\int_{\mathbb{R}^{d+1}} b(x, u) dx du = 1$. Define
\[
\mathbb{K}_{j,k}^i(x, u) = K_{j,k}^i(x, u) - \gamma_{j,k} b(x, u).
\] (4.8)
where
\[
\gamma_{j,k} = \begin{cases}
\int_{\mathbb{R}^{d+1}} K_{j,k}^i(x, u) dx du, & k = 0, \\
0, & k > 0.
\end{cases}
\] (4.9)
and by (4.7), $|\gamma_{j,k}| \leq C_N 2^{-Nj}$ for $k = 0$, and $|\gamma_{j,k}| = 0$ for $k > 0$. Since the maximal operator generated by the kernel $b$ is bounded by the nonisotropic Hardy–Littlewood maximal operator, we see that for $1 < p \leq \infty$,

$$
|\gamma_{j,0}| \sup_{t > 0} |f * b_t|_{L^p(\mathbb{R}^{d+1})} \leq C_N 2^{-Nj} \|f\|_{L^p(\mathbb{R}^{d+1})}.
$$

(4.10)

Finally, we have that

$$
\int_{\mathbb{R}^{d+1}} \mathbb{K}^{j,k}(x, u) dx du = 0.
$$

(4.11)

### 4.1 Square functions and almost orthogonality

**Lemma 4.1** For all $j + k \geq 1$, $\theta \in [0, 1]$, we have

$$
\left\| \sup_{t > 0} |f * \mathbb{K}^j_k| \right\|_{L^2(\mathbb{R}^{d+1})} \leq C(j + k) 2^{j} C_j(k) \|f\|_{L^2(\mathbb{R}^{d+1})},
$$

where

$$
C_j(k)(\theta) := \begin{cases} 2^{-(1-d/2)\theta} 2^{j-\theta k}, & \text{if } M 2^k \geq 2^j, \\
2^{-d/2} j, & \text{if } M 2^k \leq 2^j
\end{cases}
$$

(4.12)

for some sufficiently large constant $M \geq 1$.

**Proof** By (4.8), (4.9) and (4.10), it suffices to show

$$
\left\| \sup_{t > 0} |f * \mathbb{K}^j_k| \right\|_{L^2(\mathbb{R}^{d+1})} \leq C(j + k) 2^{j} C_j(k) \|f\|_{L^2(\mathbb{R}^{d+1})}.
$$

We may write

$$
\sup_{t > 0} |f * \mathbb{K}^j_k| = \sup_{n \in \mathbb{Z}} \sup_{t \in [1, 2]} |f * \mathbb{K}^{j,k}_{2^n t}|.
$$

By Lemma 2.2 with $F_n(x, u, t) = f * \mathbb{K}^{j,k}_{2^n t}(x, u)$, we see that Lemma 4.1 follows from the following estimates which are uniform in $t \in [1, 2]$:

$$
\left( \sum_{n \in \mathbb{Z}} \left\| f * \mathbb{K}^{j,k}_{2^n t} \right\|_{L^2(\mathbb{R}^{d+1})}^2 \right)^{1/2} \leq C(j + k) C_j(k) \|f\|_{L^2(\mathbb{R}^{d+1})},
$$

(4.13)

$$
\left( \sum_{n \in \mathbb{Z}} \left\| s \frac{d}{ds} \mathbb{K}^{j,k}_{2^n t} \right\|_{L^2(\mathbb{R}^{d+1})}^2 \right)^{1/2} \leq C(j + k) 2^j C_j(k) \|f\|_{L^2(\mathbb{R}^{d+1})},
$$

(4.14)

Next we apply an almost orthogonality Lemma 2.3 to prove (4.13) and (4.14). One sees that the inequality (4.13) follows from the following estimates (4.15) and (4.16) if we apply a scaling argument and Lemma 2.3 with $A = 2^{-(1-d/2)\theta} 2^{j-\theta k}$ and $B = 2^{2^j+2k}$:

$$
\|f * \mathbb{K}^{j,k}_{2^n t}\|_{L^2(\mathbb{R}^{d+1})} \leq CC_j(k) \|f\|_{L^2(\mathbb{R}^{d+1})}, \forall \theta \in [0, 1]
$$

(4.15)

and

$$
\|f * (\mathbb{K}^{j,k}_{2^n t})^* * \mathbb{K}^{j,k}_{2^n t}\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4^j+2k} 2^{-|n' - n|} \|f\|_{L^2(\mathbb{R}^{d+1})}
$$

(4.16)
first for \( n \leq n' \) and then by taking adjoints also for \( n > n' \). The inequality (4.14) follows from the following estimates (4.17) and (4.18) if we apply Lemma 2.3 with \( A = 2^j C_{j,k}(\theta) \) and \( B = 2^{3(j+k)} \):

\[
\left\| f * \left[ s \frac{d}{ds} \mathbb{K}_s^{j,k} \right]_{s=2^n t} \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^j C_{j,k}(\theta) \| f \|_{L^2(\mathbb{R}^{d+1})}, \tag{4.17}
\]

and

\[
\left\| f * \left[ \left( s \frac{d}{ds} \mathbb{K}_s^{j,k} \right) \right]_{s=2^n t} \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{6j+6k} 2^{-|n-n'|} \| f \|_{L^2(\mathbb{R}^{d+1})}, \tag{4.18}
\]

which are uniform in \( t \in [1, 2] \). The proof of (4.15), (4.16), (4.17) and (4.18) will be given in Sects. 4.1.2 and 4.1.3, which are obtained by using the cancellation of the kernels of \( \mathbb{K}_s^{j,k} \) and \( s \frac{d}{ds} \mathbb{K}_s^{j,k} \) to show almost orthogonality properties for these operators and certain estimates for oscillatory integrals to establish the decay estimates. \( \square \)

4.1.1 Reduction to oscillatory integral operators

The following lemma is due to Hörmander [7]. Here we prove it again to show the upper bound of the operator norm is irrelevant to the derivative of order 2 of the phase function.

Lemma 4.2 Define

\[ T_\lambda f(x) = \int e^{i\lambda \Phi(x,y)} a(x,y) f(y) dy \]

with \( \Phi(x,y) \in C^\infty(\mathbb{R}^d \times \mathbb{R}^d) \) and \( a(x,y) \in C^\infty_c(\mathbb{R}^d \times \mathbb{R}^d) \). Suppose

\[
\inf_{|X|=1, (x,y) \in \text{supp } a} |\partial^2_{x,y} \Phi(x,y) X| \geq c > 0 \tag{4.19}
\]

for some constant \( c > 0 \) and for each \( |\alpha| + |\beta| \geq 3 \),

\[
\sup_{(x,y) \in \text{supp } a} |\partial^\alpha_x \partial^\beta_y \Phi(x,y)| < \infty. \tag{4.20}
\]

Then there exists a constant \( C > 0 \) independent of \( \lambda \) such that

\[
\| T_\lambda f \|_{L^2(\mathbb{R}^d)} \leq C (1 + |\lambda|)^{-\frac{d}{2}} \| f \|_{L^2(\mathbb{R}^d)}. \tag{4.21}
\]

Proof If \( |\lambda| \leq 1 \), (4.21) can be obtained by Hölder’s inequality since \( a(x,y) \) is a compactly supported function. For all \( |\lambda| \geq 1 \), it suffices (by \( T^*T \) method) to show that

\[
\| T_\lambda^* T_\lambda f \|_{L^2(\mathbb{R}^d)} \leq C |\lambda|^{-d} \| f \|_{L^2(\mathbb{R}^d)}. \tag{4.22}
\]

We may write

\[
T_\lambda^* T_\lambda f(x) = \int_{\mathbb{R}^d} K_\lambda(x,z) f(z) dz,
\]

where

\[
K_\lambda(x,z) = \int_{\mathbb{R}^d} e^{-i\lambda (\Phi(y,x) - \Phi(y,z))} \overline{a}(y,x) a(y,z) dy.
\]
By a partition of unity and Schur’s lemma, the proof of (4.22) reduces to show if \(|x - z|\) is small, then
\[
|K_\lambda(x, z)| \leq C_N (1 + |\lambda||x - z|)^{-N} \tag{4.23}
\]
for a sufficiently large integer \(N \geq 1\). Notice that
\[
|\nabla_y \Phi(y, x) - \nabla_y \Phi(y, z) - \partial_{y,x}^2 \Phi(y, x)(x - z)| \leq C|x - z|^2
\]
for some sufficiently large \(C > 0\). Then by the triangle inequality and (4.19),
\[
|\nabla_y \Phi(y, x) - \nabla_y \Phi(y, z)| \geq c|x - z| - C|x - z|^2 \geq \frac{c}{2}|x - z|
\]
if \(|x - z| \leq c/2C\). On another hand, it follows by (4.20) that for all \(|\alpha| \geq 1\),
\[
|\partial_\alpha^\alpha (\nabla_y \Phi(y, x) - \nabla_y \Phi(y, z))| \leq C_\alpha|x - z|.
\]
By integration by parts, if \(|x - z| \leq c/2C\), then
\[
|K_\lambda(x, z)| \leq C_N |\lambda|^{-N}|x - z|^{-N}
\]
for a sufficiently large integer \(N \geq 1\). Together with the trivial estimate \(|K_\lambda(x, z)| \leq C\), we finish the proof of (4.23) and thus (4.22) follows readily. This completes the proof of Lemma 4.2. \(\square\)

Now we begin to verify estimates (4.15), (4.16), (4.17) and (4.18). To do it, we will reduce them to estimates for oscillatory integral operators. Recall that \(E_{d-1}\) and \(P_{d-1}\) are two \(d \times (d - 1)\) matrices given as in (3.1). One can write
\[
f \ast K^{j,k}(x, u) = \int_{\mathbb{R}^{d+1}} K^{j,k}(x - y, u - v + x^T J_y) f(y, v) dy dv,
\]
where
\[
K^{j,k}(x - y, u - v + x^T J_y) = \chi_1(x - y) \int_{\mathbb{R} \times \mathbb{R}} e^{i(\sigma(x_1 - y_1 - \Gamma_1(x' - y')) + \tau(u - v + x^T J_y))} 
\phi_j(\sigma) \phi_k(\tau) d\sigma d\tau.
\]
We then apply a Fourier transform on \(\mathbb{R}\), in the \(u\) variable of \(f \ast K^{j,k}(x, u)\), and a change of variable in \(\sigma\) to obtain
\[
\mathcal{F}_u(f \ast K^{j,k})(x, \lambda_2) = 2^j \phi_k(\lambda_2) \int_{\mathbb{R}^{d+1}} e^{i2^j(\sigma(x_1 - y_1 - \Gamma_1(x' - y')) - i\lambda_2 x^T J_y)} \chi_1(x - y) \phi(\sigma) 
\mathcal{F}_v^{-1} f(y, \lambda_2) dy d\sigma. \tag{4.24}
\]
Define for \(|\lambda_1| > 1\) and \(\lambda_2 \in \mathbb{R}\),
\[
T_{\lambda_1, \lambda_2} g(x) = \lambda_1 \int_{\mathbb{R}^{d+1}} \chi_1(x - y) e^{i\lambda_1 \sigma(x_1 - y_1 - \Gamma_1(x' - y'))} e^{-i\lambda_2 x^T J_y} \phi(\sigma) g(y) dy d\sigma.
\]
If \(|\lambda_1| \leq 1\), we should replace \(\phi\) by \(\varphi\). Then we have the following result.

**Lemma 4.3** For all \(\theta \in [0, 1]\), there exists a constant \(C > 0\) such that
\[
\|T_{\lambda_1, \lambda_2} g\|_{L^2(\mathbb{R}^d)} \leq \begin{cases} 
C(1 + |\lambda_1|)^{-\frac{(1 - \theta)(d - 2)}{2}} (1 + |\lambda_2|)^{-\theta} \|g\|_{L^2(\mathbb{R}^d)} & \text{if } 2M|\lambda_2| \geq |\lambda_1|; \\
C(1 + |\lambda_1|)^{-\frac{d - 2}{2}} \|g\|_{L^2(\mathbb{R}^d)} & \text{if } M|\lambda_2| \leq 2|\lambda_1| \end{cases} \tag{4.25}
\]
for some sufficiently large \(M \geq 1\).
Proof If $|\lambda_1| \leq 1$, we have

$$T_{\lambda_1, \lambda_2}g(x) = \lambda_1 \int_{\mathbb{R}^{d+1}} \chi_1(x-y)e^{-i\lambda_1 \Psi_{\lambda_1, \lambda_2}(x,y)} \phi(\sigma)g(y)d\sigma,$$

where $\Psi_{\lambda_1, \lambda_2}(x,y) = x^T J y - \lambda_2^{-1} \lambda_1 \sigma(x_1 - y_1 - \Gamma_1(x' - y'))$. If $|\lambda_2| \leq 1$, from Young’s inequality we have

$$\|T_{\lambda_1, \lambda_2}g\|_{L^2(\mathbb{R}^d)} \leq C \|\chi_1 \ast |g|\|_{L^2(\mathbb{R}^d)} \leq C\|g\|_{L^2(\mathbb{R}^d)}.$$

If $|\lambda_2| > 1$, we have

$$\partial^2_{(x_1, \ldots, x_k), (y_1, \ldots, y_k)} \Psi_{\lambda_1, \lambda_2}(x, y) X \geq c_{\text{low}} > 0 \text{ for all } |X| = 1, X \in \mathbb{R}^K.$$

Then (4.25) can be proven by applying Lemma 4.2 in the first $\kappa$ ($\kappa \geq 2$) variables. It remains to show (4.25) if $|\lambda_1| > 1$. Let us first prove (4.25) for $2M|\lambda_2| \geq |\lambda_1|$ where $M \geq 1$ is a constant chosen later. In this case, we write

$$T_{\lambda_1, \lambda_2}g(x) = \lambda_1 \int_{\mathbb{R}^{d+1}} \chi_1(x-y)e^{i\lambda_1 \Psi_{\lambda_1, \lambda_2}(x,y)} \phi(\sigma)g(y)d\sigma,$$

where the phase function $\Psi_{\lambda_1, \lambda_2}$ is given by

$$\Psi_{\lambda_1, \lambda_2}(x,y) = \sigma(x_1 - y_1 - \Gamma_1(x' - y')) - \lambda_1^{-1} \lambda_2 x^T J y.$$

We have

$$\partial^2_{x, y} \Psi_{\lambda_1, \lambda_2}(x,y) = \sigma P_{d-1} \Gamma_1(x' - y') - \lambda_1^{-1} \lambda_2 J,$$

where $P_{d-1}$ is given in (3.1). We apply (3.3) of Lemma 3.1 by taking $Q = \partial^2_{x'} \Gamma_1(x' - y')$ and $a = -\sigma^{-1} \lambda_1^{-1} \lambda_2$ to see that there exists a $c_1 > 0$ such that

$$|\partial^2_{x, y} \Psi_{\lambda_1, \lambda_2}(x, y) X| = |(P_{d-1} \partial^2_{x'} \Gamma_1(x' - y') - \sigma^{d-1} \lambda_1^{-1} \lambda_2 J)X| \geq c_1 > 0,$$

for all $|X| = 1, x - y \in \supp \chi_1$ and $\sigma \in \supp \phi$. Notice that

$$|\partial^\alpha_{x, y} \Psi_{\lambda_1, \lambda_2}(x, y)| = |-\sigma \partial^\alpha_{x, y} \Gamma_1(x' - y')| \leq C_\alpha \text{ for all } |\alpha| \geq 3.$$

Define

$$T^\sigma_{\lambda_1, \lambda_2}g(x) = \lambda_1 \int_{\mathbb{R}^{d+1}} \chi_1(x-y)e^{i\lambda_1 \Psi_{\lambda_1, \lambda_2}(x,y)} g(y)d\sigma.$$

By Fubini’s theorem and Hölder’s inequality, we have

$$|T_{\lambda_1, \lambda_2}g(x)| = \left| \int T^\sigma_{\lambda_1, \lambda_2}g(x)\phi(\sigma)d\sigma \right| \leq C \left( \int |T^\sigma_{\lambda_1, \lambda_2}g(x)|^2 \phi(\sigma)d\sigma \right)^{1/2}.$$

These, together with Lemma 4.2, tell us that

$$\|\chi_1(x)T_{\lambda_1, \lambda_2}g(x)\|_{L^2(\mathbb{R}^d)} \leq C \left( \int \|\chi_1(x)T^\sigma_{\lambda_1, \lambda_2}g(x)\|^2_{L^2(\mathbb{R}^d)} \phi(\sigma)d\sigma \right)^{1/2} \leq C|\lambda_1| \cdot (1 + |\lambda_1|)^{-d/2} \left( \int \|g(y)\|^2_{L^2(\mathbb{R}^d)} \phi(\sigma)d\sigma \right)^{1/2}.$$
By applying Lemma 4.2 and use the method in [17, Chapter XI, p. 500]. Choose a function 

\[ \tilde{\psi}_{\lambda_1, \lambda_2}(x, y) = \lambda_1 \lambda_2^{-1} \sigma(x_1 - y_1 - \Gamma_1(x' - y')) - x^T J y. \]

Then we have 

\[ |\partial_{\lambda_1}^{\alpha} \tilde{\psi}_{\lambda_1, \lambda_2}(x, y)| \leq C_{\alpha} \quad \text{for all } |\alpha| \geq 2 \]

and equal to 1 on 

\[ \mathbb{R}^{\kappa}. \]

We apply (3.2) of Lemma 3.1 with 

\[ g = E_{\kappa} P_{d-1}^{2} \Gamma_{1}(x' - y') P_{d-1}^{T} E_{\kappa} \]

and 

\[ a = -\sigma \lambda_1 \lambda_2^{-1}, \]

to obtain 

\[ |\partial_{(x_1, \ldots, x_k), (y_1, \ldots, y_k)}^{2} \tilde{\psi}_{\lambda_1, \lambda_2}(x, y) X| = |(aG + J_K)X| \geq c_0 > 0 \]

for all 

\[ X \in \mathbb{R}^{\kappa}. \]

By applying Lemma 4.2 in the first \( \kappa (\kappa \geq 2) \) variables and Fubini’s theorem, we have 

\[ \| \chi(x) T_{\lambda_1, \lambda_2} g(x) \|_{L^2(\mathbb{R}^{d})} \leq C |\lambda_1| \cdot (1 + |\lambda_2|)^{-\kappa/2} \| g \|_{L^2(\mathbb{R}^{d})}, \]

(4.27)

By (4.26) and (4.27), for all \( \theta \in [0, 1] \), we have 

\[ \| \chi(x) T_{\lambda_1, \lambda_2} g(x) \|_{L^2(\mathbb{R}^{d})} \leq C (1 + |\lambda_1|)^{-(d+2-\theta)} (1 + |\lambda_2|)^{-\theta} \| g \|_{L^2(\mathbb{R}^{d})}. \]

By a similar translation invariance argument as in [17, p. 236], we have 

\[ \| T_{\lambda_1, \lambda_2} g \|_{L^2(\mathbb{R}^{d})} \leq C (1 + |\lambda_1|)^{-(d+2-\theta)} (1 + |\lambda_2|)^{-\theta} \| g \|_{L^2(\mathbb{R}^{d})}. \]

Next we consider the case \( M|\lambda_2| \leq 2|\lambda_1| \). If \( |\lambda_1| \leq 1 \), we have 

\[ |T_{\lambda_1, \lambda_2} g(x) | \leq C \chi \ast |g|(x). \]

This implies that 

\[ \| T_{\lambda_1, \lambda_2} g \|_{L^2(\mathbb{R}^{d})} \leq C \| \chi \|_{L^1(\mathbb{R})} \| g \|_{L^2(\mathbb{R}^{d})}. \]

Next we consider \( |\lambda_1| \geq 1 \) and use the method in [17, Chapter XI, p. 500]. Choose a function \( \psi \in C_{c}^{\infty}(\mathbb{R}) \) with 

\[ \text{supp } \psi \subseteq [1/2, 2] \] and equal to 1 on \( [\frac{3}{4}, \frac{5}{4}] \). Define 

\[ \varphi_{\lambda_1, \lambda_2}(x, \eta) = \lambda_1 \psi(\eta) \int \chi(x - y) e^{i \lambda_2 \tilde{\psi}_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) \phi(\sigma) g(y) dy d\sigma}, \]

where 

\[ \tilde{\psi}_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) = \sigma \eta (x_1 - y_1 - \Gamma_1(x' - y')) - \lambda_1^{-1} \lambda_2 x^T J y. \]

Observe that 

\[ \partial_{(x, \eta), (y, \sigma)} \psi_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \sigma \eta \tilde{\psi}_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) & \eta & 0 \\ 0 & \sigma \eta \tilde{\psi}_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \]

which implies 

\[ \det \partial_{(x, \eta), (y, \sigma)} \psi_{\lambda_1, \lambda_2}(x, \eta, y, \sigma) = (\sigma \eta)^d \det \partial_{x} \Gamma_1 + O(\lambda_1^{-1} \lambda_2) \]

and 

\[ |\partial_{x, \eta, y, \sigma} \psi(x, \eta, y, \sigma)| \leq C_{\alpha} \quad \text{for all } |\alpha| \geq 2. \]
Then by (4.2), for all \( x - y \in \text{supp} \chi_1, \sigma \in \text{supp} \phi \) and \( \eta \in \text{supp} \psi \), we have

\[
| \det \frac{\partial^2}{\partial (x, \eta, y, \sigma)} \Psi_{\lambda_1, \lambda_2}(x, \eta, y, \sigma)| \geq c_H^{(1)}(\sigma \eta)^{d - C\lambda_1^{-1} \lambda_2 \geq c_H^{(1)} 4^{-d} - 2CM^{-1} \geq \frac{c_H^{(1)}}{2} 4^{-d},
\]

where we choose \( M := 4^{d+2} C[c_H^{(1)}]^{-1} + 1 \) and \( c_H^{(1)} \) is the one defined in (4.2). Thus it follows from Lemma 4.2 that

\[
\| \chi_1(x) \partial_{x, \eta} \chi_1, \lambda_2 g(x, \eta) \|_{L^2(\mathbb{R}^{d+1})} \leq C \| \lambda_1 \| (1 + |\lambda_1|)^{-\frac{d+1}{2}} \| \phi(\sigma) g(y) \|_{L^2(\mathbb{R}^{d+1})}
\]

\[
\leq C (1 + |\lambda_1|)^{-\frac{d+1}{2}} \| g \|_{L^2(\mathbb{R}^d)}.
\]

By a translation invariance argument again,

\[
\| \partial_{\eta} \chi_{1, \lambda_2} g(x, \eta) \|_{L^2(\mathbb{R}^{d+1})} \leq C (1 + |\lambda_1|)^{-\frac{d+1}{2}} \| g \|_{L^2(\mathbb{R}^d)}.
\]

(4.28)

Since

\[
\frac{d}{d \eta} \chi_{1, \lambda_2} g(x, \eta) = -\lambda_1 \psi(\eta) \int \chi_1(x - y) e^{i\eta \cdot \psi(x,y,\sigma)} \frac{d}{d \sigma} \phi(\sigma) g(y) dy d\sigma
\]

\[
+ \lambda_1 \phi'(\eta) \int \chi_1(x - y) e^{i\eta \cdot \psi(x,y,\sigma)} \phi(\sigma) g(y) dy d\sigma,
\]

it is seen that

\[
\left\| \frac{d}{d \eta} \chi_{1, \lambda_2} g(x, \eta) \right\|_{L^2(\mathbb{R}^{d+1})} \leq C (1 + |\lambda_1|)^{-\frac{d+1}{2}} \| g \|_{L^2(\mathbb{R}^d)}.
\]

(4.29)

By H"older’s inequality, we have

\[
\sup_{\eta \in [1,2]} | \partial_{\eta} \chi_{1, \lambda_2} g(x, \eta) |^2 \leq 2 \left( \int_1^2 | \partial_{\eta} \chi_{1, \lambda_2} g(x, \eta) |^2 d\eta \right)^{1/2} \left( \int_1^2 | \frac{d}{d \eta} \partial_{\eta} \chi_{1, \lambda_2} g(x, \eta) |^2 d\eta \right)^{1/2}
\]

\[
+ | \partial_{\eta} \chi_{1, \lambda_2} g(x, 1) |^2,
\]

which, together with (4.28), (4.29) and H"older’s inequality in \( x \), implies us that

\[
\| T_{\lambda_1, \lambda_2} g \|_{L^2(\mathbb{R}^d)} = \| \partial_{\eta} \chi_{1, \lambda_2} g(x, 1) \|_{L^2(\mathbb{R}^d)} \leq C (1 + |\lambda_1|)^{-\frac{d+1}{2}} \| g \|_{L^2(\mathbb{R}^d)}.
\]

This finishes the proof of (4.25).

\[\square\]

4.1.2 Proof of (4.15) and (4.16)

**Proof of (4.15)** If \( M \lambda_2 \geq 2^j \), we have \( 2M |\lambda_2| \geq 2^j \) for all \( \lambda_2 \in \text{supp} \phi_k \). By Plancherel’s theorem, (4.24), Fubini’s theorem and Lemma 4.3, we have

\[
\| f * K^{j,k} \|_{L^2(\mathbb{R}^{d+1})} = \| \mathcal{F}_u (f * K^{j,k}) \|_{L^2(\mathbb{R}^{d+1})}
\]

\[
= \| T_{2^j, \lambda_2} (\mathcal{F}_u f (\cdot, \lambda_2)) (x) \|_{L^2(\mathbb{R}^2) L^2_2(\mathbb{R}^d)}
\]

\[
\leq C 2^{-\left(1 - \frac{d+2}{2} \right) j 2^{-\theta_0} } \| f \|_{L^2(\mathbb{R}^{d+1})}
\]

for all \( \theta \in [0, 1] \). If \( M \lambda_2 \leq 2^j \), we have \( M |\lambda_2| \leq 2^{j+1} \) for all \( \lambda_2 \in \text{supp} \phi_k \). By the above argument, we have

\[
\| f * K^{j,k} \|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{-\frac{d+1}{2} j} \| f \|_{L^2(\mathbb{R}^{d+1})}.
\]

\[\square\]
These, together with (4.8), yields that (4.15) holds. □

**Proof of (4.16)** By a scaling argument, the proof of (4.16) reduces to show for all $n \leq 0$,

$$
\left\| f * (\mathbb{K}^{j,k}_n)^* * \mathbb{K}^{j,k}_n \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4j+4k} 2^n \| f \|_{L^2(\mathbb{R}^{d+1})} \quad (4.30)
$$

and

$$
\left\| f * (\mathbb{K}^{j,k}_n)^* * \mathbb{K}^{j,k}_n \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4j+4k} 2^n \| f \|_{L^2(\mathbb{R}^{d+1})}, \quad (4.31)
$$

where $\mathbb{K}^{j,k}_n$ is defined in (4.8).

We start with the proof of (4.30). We use integration by parts to see that for all $|\alpha| \leq 1$ and $N > 0$, there exists a constant $C = C(\alpha, N) > 0$ such that

$$
|\partial^\alpha_{x,u} \mathbb{K}^{j,k}_n(x, u)| + |\partial^\alpha_{x,u} (\mathbb{K}^{j,k}_n)^*(x, u)| \leq C 2^{2j+2k} \chi_1(x) \left(1 + |u|\right)^{-N} . \quad (4.32)
$$

Note that $|x^T J y| \leq c \|J\| \cdot |y| \leq c^2 \|J\| \|x\| \leq c$. These, in combination with the cancellating property (4.11) of the kernel of $\mathbb{K}^{j,k}_n$, yield

$$
|\left(\mathbb{K}^{j,k}_n \ast \mathbb{K}^{j,k}_n\right)(x, u)| = \left| \int_{\mathbb{R}^{d+1}} \left((\mathbb{K}^{j,k}_n)^*(x - y, u - v - x^T J y)\mathbb{K}^{j,k}_n(y, v)\right) dy dv \right|
$$

$$
= \left| \int_{\mathbb{R}^{d+1}} \left((\mathbb{K}^{j,k}_n)^*(x - y, u - v - x^T J y) - (\mathbb{K}^{j,k}_n)^*(x, u)\right)\mathbb{K}^{j,k}_n(y, v)\right| dy dv
$$

$$
\leq C \int_0^1 \int_{\mathbb{R}^{d+1}} 2^{2j+2k} \left(1 + |x - \theta y| + |u - \theta v|\right)^{-N} \left(|y| + |v|\right)\left|\mathbb{K}^{j,k}_n(y, v)\right| dy dv d\theta.
$$

This gives that for $n \leq 0$,

$$
\int_{\mathbb{R}^{d+1}} \left|\left(\mathbb{K}^{j,k}_n \ast \mathbb{K}^{j,k}_n\right)(x, u)\right| dx du \leq C 2^{2j+2k} \int_{\mathbb{R}^{d+1}} \left(|y| + |v|\right)\left|\mathbb{K}^{j,k}_n(y, v)\right| dy dv
$$

$$
\leq C 2^{4j+4k} 2^n .
$$

By the Schur lemma, (4.30) follows readily.

Next we prove (4.31). We use the cancellating property of $(\mathbb{K}^{j,k})^*$ and a similar argument as above to show that

$$
\int_{\mathbb{R}^{d+1}} \left|\left(\mathbb{K}^{j,k}_n \ast \mathbb{K}^{j,k}_n\right)(x, u)\right| dx du \leq C 2^{4j+4k} 2^n .
$$

This, together with the Schur lemma, gives the desired estimate (4.31). □

### 4.1.3 Proof of (4.17) and (4.18)

**Proof of (4.17)** It suffices to show

$$
\left\| f \ast \left[ s \frac{d}{ds} K^{j,k}_s \right]_{s=2^n t} \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^j C_j, k(\theta) \| f \|_{L^2(\mathbb{R}^{d+1})}.
$$

By (4.3) and change of variables in $\sigma, \tau$, we have

$$
K^{j,k}_s(x, u) = s^{-d+1} \chi_1(s^{-1} x) \int_{\mathbb{R} \times \mathbb{R}} e^{i \sigma (x_1 - s^{-1} x'_1)} e^{i \tau u} \phi_j(s \sigma) \phi_k(s^2 \tau) d\sigma d\tau .
$$
Observe that
\[
\frac{d}{ds} K_s^{j,k} = (K^{(1)})^{j,k}_s + (K^{(2)})^{j,k}_s + 2^j (K^{(3)})^{j,k}_s
\]  
(4.33) where
\[
(K^{(1)})^{j,k}_s = \left[(-d+1)s^{-d+1} \chi_1(s^{-1}x) - s^{-1} x \chi_1'(s^{-1}x)\right] s^{-d+1} \int \int_{\mathbb{R} \times \mathbb{R}} e^{i \sigma (x_1 - s \Gamma_1(s^{-1}x'))} e^{i \tau u} \phi_j(s \sigma) \phi_k(s \tau) d \sigma d \tau,
\]
\[
(K^{(2)})^{j,k}_s = s^{-d+1} \chi_1(s^{-1}x) \int \int_{\mathbb{R} \times \mathbb{R}} e^{i \sigma (x_1 - s \Gamma_1(s^{-1}x'))} e^{i \tau u} \left[2^{-j} s \sigma \phi_j(s \sigma) \phi_k(s \tau) + 2^{-k+1} s^2 \tau \phi_j(s \sigma) \phi_k'(s \tau)\right] d \sigma d \tau,
\]
\[
(K^{(3)})^{j,k}_s = i s^{-1} x \Gamma_1'(s^{-1}x') - i \Gamma_1(s^{-1}x')] s^{-d+1} \chi_1(s^{-1}x) \int \int_{\mathbb{R} \times \mathbb{R}} e^{i \sigma (x_1 - s \Gamma_1(s^{-1}x'))} e^{i \tau u} 2^{-j} s \sigma \phi_j(s \sigma) \phi_k(s \tau) d \sigma d \tau.
\]

By Plancherel’s theorem, Fubini’s theorem and Lemma 4.3, an argument as in the proof of (4.15) yields that for all \( i = 1, 2, 3 \),
\[
\| f \ast (K^{(i)})^{j,k} \|_{L^2(\mathbb{R}^{d+1})} \leq C C_{j,k}(\theta) \| f \|_{L^2(\mathbb{R}^{d+1})}, \quad \forall \theta \in [0,1],
\]
which, together with a scaling argument, implies that
\[
\sup_{t \in [1,2]} \| f \ast \left( s \frac{d}{ds} K_s^{j,k} \right) \|_{L^2(\mathbb{R}^{d+1})} \leq 2^j \sum_{i=1}^{6} \sup_{t \in [1,2]} \| f \ast (K^{(i)})^{j,k}_{2^it} \|_{L^2(\mathbb{R}^{d+1})} \leq C 2^j C_{j,k}(\theta) \| f \|_{L^2(\mathbb{R}^{d+1})}.
\]

This completes the proof of (4.17). \( \square \)

**Proof of (4.18)** The proof is similar as (4.16) and thus we only sketch it. By (4.11), \( s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \) and \( (s \frac{d}{ds} \| \ast \|_{K_s^{j,k}})^{*} \) also have integral zero. By (4.33) and (4.8), an argument as in (4.32), we see that for all \( |\alpha| \leq 1 \) and \( N > 0 \), there exists a constant \( C = C(\alpha, N) > 0 \) such that
\[
\left| \partial_x^{\alpha_1} \partial_u^{\alpha_2} \left( s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right)'(x,u) \right| + \left| \partial_x^{\alpha_1} \partial_u^{\alpha_2} \left( s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right)''(x,u) \right| \leq C 2^{3j+3k}s^{-d-2-|\alpha_1|-2|\alpha_2|} \chi_1(s^{-1}x) \left( 1 + s^{-2}|u| \right)^{-N},
\]
where \( \alpha = (\alpha_1, \alpha_2) \). It suffices to show for all \( n, n' \in \mathbb{Z} \),
\[
\int_{\mathbb{R}^{d+1}} \left| \left( s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right)'_{s=2^{n'}}' \right|^* \left( s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right)'_{s=2^n} (x,u) dxdudu \leq C 2^{6j+6k} 2^{-|n'-n|} \| f \|_{L^2(\mathbb{R}^{d+1})},
\]
(4.34) follows. By the cancelation property of \( \left[ s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right]_{s=2^n} \) if \( n \geq n' \), \( \left[ s \frac{d}{ds} \| \ast \|_{K_s^{j,k}} \right]_{s=2^n} \) if \( n' \geq n \) and mean value theorem, (4.34) follows. \( \square \)
4.2 Weak type \((1, 1)\)-estimate

In this section, we will prove the following result.

**Lemma 4.4** Let \(j, k \geq 0\). For all \(\alpha > 0\), we have

\[
\left| \{ (x, u) : \sup_{t > 0} |f \ast K^{j,k}_t(x, u)| > \alpha \} \right| \leq C (j + k) 2^j \alpha^{-1} ||f||_{L^1(\mathbb{R}^{d+1})}.
\]

**Proof** Let \(\rho(x, u) := |x| + |u|^{1/2}\). For all \((x, u), (y, v) \in \mathbb{R}^{d+1}\), we have

\[
\rho((x, u)(y, v)) \leq c_f (\rho(x, u) + \rho(y, v)), \quad \text{with} \quad c_f = \frac{3}{2} \| J \|^{1/2} + 1.
\]

From this, we see that if \(\rho(x, u) \geq 2c_f \rho(y, v)\), then

\[
\rho((y, v)^{-1}(x, u)) \geq (2c_f)^{-1} \rho(x, u).
\] (4.35)

Since we have already shown the \(L^2\) bounds for the maximal function, from the Calderón–Zygmund theory it suffices to verify the following Hörmander-type condition:

\[
\int_{\rho(x, u) \geq 2c_f \rho(y, v), s > 0} \sup_{\rho(x, u) \geq 2c_f 2^{-n} s} \left| K^{j,k}_s((y, v)^{-1}(x, u)) - K^{j,k}_s(x, u) \right| dxdu \leq C (j + k) 2^j.
\]

By a scaling argument, it suffices to show for all \(\rho(y, v) \leq 2^{-n} \delta\) and \(\delta > 0\),

\[
\sum_{n \in \mathbb{Z}} \int_{\rho(x, u) \geq 2c_f 2^{-n} \delta} \sup_{s \in [1, 2]} \left| K^{j,k}_s((y, v)^{-1}(x, u)) - K^{j,k}_s(x, u) \right| dxdu \leq C (j + k) 2^j.
\] (4.36)

To show (4.36), we first note that from (4.3), we have that for any \(j, k \geq 0\),

\[
|K^{j,k}_s(x, u)| \leq C 2^j 2^k (1 + |x| + |u|^{1/2})^{-N/2} (1 + 2^j |x_1 - \Gamma_1(x')|)^{-N/2} (1 + 2^{k/2} |u|^{1/2})^{-N/2},
\] (4.37)

\[
\sup_{s \in [1, 2]} |\nabla_{x,u} K^{j,k}_s(x, u)| \leq C 2^{j/2} 2^{2k} (1 + |x| + |u|^{1/2})^{-N/2},
\] (4.38)

for large \(N \geq 1\); and by (4.33),

\[
\left| \frac{d}{ds} K^{j,k}_s(x, u) \right| \leq C 2^j 2^{2k} (1 + |x| + |u|^{1/2})^{-N/2} \left( 1 + 2^j |s^{-1}x_1 - \Gamma_1(s^{-1}x')| \right)^{-N/2} (1 + 2^{k/2} |u|^{1/2})^{-N/2}.
\] (4.39)

uniformly in \(s \in [1, 2]\). Now we consider two cases: \(2^{-n} \delta \geq 1\) and \(2^{-n} \delta \leq 1\).

**Case 1:** \(2^{-n} \delta \geq 1\).

In this case, by a triangle inequality, a change of variable, Newton–Leibniz formula, (4.35), (4.37) and (4.39), we see that for all \(\rho(y, v) \leq 2^{-n} \delta\),

\[
\int_{\rho(x, u) \geq 2c_f 2^{-n} \delta} \sup_{s \in [1, 2]} \left| K^{j,k}_s((y, v)^{-1}(x, u)) - K^{j,k}_s(x, u) \right| dxdu \\
\leq \int_{\rho(x, u) \geq 2c_f 2^{-n} \delta} \sup_{s \in [1, 2]} \left( |K^{j,k}_s((y, v)^{-1}(x, u))| + |K^{j,k}_s(x, u)| \right) dxdu \\
\leq 2 \int_{\rho(x, u) \geq 2^{-n} \delta} \sup_{s \in [1, 2]} \left| K^{j,k}_s(x, u) \right| dxdu
\]
\[
\begin{align*}
\leq & \ 2 \int_{\rho(x,u) \geq 2^{-n}\delta} \left| \frac{d}{ds} K_s^{j,k}(x,u) \right| dx du + 2 \int_{\rho(x,u) \geq 2^{-n}\delta} \left| K_{j,k}(x,u) \right| dx du \\
\leq & \ C(2^{-n}\delta)^{-N/2} 2^j,
\end{align*}
\]

where we use \(2c_J \geq 1\). Then we have

\[
\sum_{2^{-n}\delta \geq 1} \int_{\rho(x,u) \geq 2c_J 2^{-n}\delta} \sup_{s \in [1,2]} \left| K_s^{j,k}(y,v)^{-1}(x,u) - K_s^{j,k}(x,u) \right| dx du \leq C 2^j.
\]

**Case 2:** \(2^{-n}\delta \leq 1\).

In this case, we will show for all \(\rho(y,v) \leq 2^{-n}\delta\),

\[
\sum_{2^{-n}\delta \leq 1} \int_{\rho(x,u) \geq 2c_J 2^{-n}\delta} \sup_{s \in [1,2]} \left| K_s^{j,k}(y,v)^{-1}(x,u) - K_s^{j,k}(x,u) \right| dx du \leq C(j + k)2^j.
\]

(4.41)

By the support condition of \(K^{j,k}\), we can suppose \(|x - y| \leq C\|y\|, v\|v\| \leq 1\), we have \(|y| \leq 1\) and \(|v| \leq 1\). As a result, one must have \(|x| \leq C + 1\) and

\[
|y, v + y^T Jx| \leq C \rho(y, v).
\]

(4.42)

Let \(\theta \in [0, 1]\). By (4.35) and \(\rho(x,u) \geq 2c_J \rho(y,v) \geq 2c_J \rho(\theta y, \theta v)\), we have

\[
\rho((\theta y, \theta v)^{-1}(x,u)) \geq (2c_J)^{-1} \rho(x,u).
\]

This, together with Newton–Leibniz formula, (4.38) and (4.42), tells us that

\[
\begin{align*}
& \sup_{s \in [1,2]} \left| K_s^{j,k}(y,v)^{-1}(x,u) - K_s^{j,k}(x,u) \right| \\
\leq & \ \sup_{s \in [1,2]} \int_0^1 \left| (\nabla_{x,u} K_s^{j,k})(((\theta y, \theta v)^{-1}(x,u)) \right| d\theta \left| (y,v + y^T Jx) \right| \\
\leq & \ C 2^{j+2k} \rho(y,v) \int_0^1 \left( 1 + \rho((\theta y, \theta v)^{-1}(x,u)) \right)^{-N/2} d\theta \\
\leq & \ C 2^{j+2k} 2^{-n}\delta \left( 1 + \rho(x,u) \right)^{-N/2}.
\end{align*}
\]

After integrating in \(x, u\), we have
On another hand, we apply (4.37), (4.39) and a similar argument as in the proof of (4.40) to obtain

\[
\sum_{2^{-n}\delta \leq 2^{-2j-2k}} \int_{\rho(x,u) \geq 2e_j 2^{-n}\delta} \sup_{s \in [1,2]} |K^j_k((y,v)^{-1}(x,u)) - K_{s,k}^j(x,u)| \, dx \, du \leq C \sum_{2^{-n}\delta \leq 2^{-2j-2k}} 2^{j+2k} 2^{-n}\delta \leq C. \tag{4.43}
\]

where in the last inequality we use the fact that the summation has only \(O(j+k)\) terms. (4.41) follows from (4.43) and (4.44). Hence (4.36) is proved, and then the proof of Lemma 4.4 is completed.

\[\square\]

4.3 Proof of (i) of Theorem 2.1

Proof of (i) of Theorem 2.1 We first consider \(1 < p \leq 2\). By (4.5), we have

\[
\sup_{t > 0} |f \ast (\mu_1)_t| \leq C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sup_{t > 0} \|f \ast K^j_k\|.
\]

From (4.6), it suffices to consider that \(j + k \geq 1\). Note if \(d \geq 3\),

\[
C_{j,k}(\theta) \leq 2^{-(1-\frac{d}{2}\theta)\frac{d-2}{2}} j^{2-\theta k} \quad \text{for all } j, k \geq 0, \theta \in [0, 1],
\]

where \(C_{j,k}(\theta)\) is the constant defined in (4.12). We use Lemma 4.1 to show that for all \(j + k \geq 1, \theta \in [0, 1]\),

\[
\left\| \sup_{t > 0} |f \ast K^j_k| \right\|_{L^2(\mathbb{R}^{d+1})} \leq C (j + k) 2^{j} 2^{-(1-\frac{d}{2}\theta)\frac{d-2}{2}} j^{2-\theta k} \|f\|_{L^2(\mathbb{R}^{d+1})}.
\]

Lemma 4.4 tells us that

\[
\left\| \sup_{t > 0} |f \ast K^j_k| \right\|_{L^1(\mathbb{R}^{d+1})} \leq C (j + k) 2^{j} \|f\|_{L^1(\mathbb{R}^{d+1})}.
\]

By interpolation, for all \(\theta \in (0, 1)\) we have

\[
\left\| \sup_{t > 0} |f \ast K^j_k| \right\|_{L^p(\mathbb{R}^{d+1})} \leq C (j + k) 2^{\left(\frac{2-\frac{d}{p}}{p} - (2-\frac{d}{2})(1-\frac{d}{2}\theta)\frac{d-2}{2} j^{2-\theta k} \left(\frac{2-(d-2)}{d-2}\theta\right) j^{2-\theta k} \|f\|_{L^p(\mathbb{R}^{d+1})}.
\]

\[\square\]
The summation in $k$ converges automatically and the condition
\[
\frac{1}{p} - (d-2)\left(1 - \frac{1}{p}\right) + \left(1 - \frac{1}{p}\right) \frac{d(d-2)}{2} \theta < 0 \quad \text{for some } \theta \in (0, 1)
\]
makes sure that the summation in $j$ converges. It holds if and only if $\frac{1}{p} - (d-2)(1 - \frac{1}{p}) < 0$.

As a result, the summation in $j, k$ converges whenever $(d-1)/(d-2) < p$ for $d \geq 4$ and $\theta$ is sufficiently small. By interpolation with a trivial $L^\infty$ estimate, we obtain the desired $L^p$ estimate in (i) of Theorem 2.1. □

5 Proof of (ii) of Theorem 2.1

In this section, we will show that (ii) of Theorem 2.1. That is, for $d \geq 3$, the maximal operator $\sup_{t > 0} |f \ast (\mu_2)_t(x, u)|$ is bounded on $L^p(\mathbb{R}^{d+1})$ for $p > d/(d-1)$ where $\mu_2$ is given as in (2.3), i.e.,
\[
\mu_2(x, u) = \chi_{\mu_2}(x, u) \int_{\mathbb{R} \times \mathbb{R}} e^{i(\sigma(x_d - \Gamma_2(x'')) + \tau \mu)} \, d\sigma \, d\tau
\]
with $x_d = \Gamma_2(x'')$, $x'' = (x_1, \ldots, x_{d-1})$ with $\Gamma_2 \in C^\infty(\mathbb{R}^{d-1})$ and $\mu_2$ is supported in a small neighborhood of $(x''_0, \Gamma_2(x''_0))$ for some $x''_0 \in \mathbb{R}^{d-1}$. In the following we give a sketch of the proof since the proof is similar to that of (i) in Theorem 2.1.

Since $\mu_2$ has a sufficiently small support, we can choose a smooth nonnegative function $\chi_2$ defined on $\mathbb{R}^{d-1}$, which is equal to 1 on the projection of supp $\mu_2$ to $x''$-space and also has a small support. By the curvature hypothesis and the convexity of $\Sigma$, we can further assume $\partial^2_{x''} \Gamma_2(x'')$ is positive definite for all $x'' \in \text{supp } \chi_2$ such that
\[
\inf_{x'' \in \text{supp } \chi_2} \det \left( \partial^2_{x''} \Gamma_2(x'') \right) \geq c_H^{(2)}
\]
for some $c_H^{(2)} > 0$. Let $\{\phi_k\}_{k=0}^\infty$ be given as in (4.1). Define for $j, k \geq 0$,
\[
K^{j,k}(x, u) = \chi_2(x'') \int_{\mathbb{R} \times \mathbb{R}} e^{i\sigma(x_d - \Gamma_2(x''))} e^{i\tau \mu} \phi_j(\sigma) \phi_k(\tau) \, d\sigma \, d\tau.
\]
An argument as in (4.5) yields that
\[
|f \ast (\mu_2)_t| \leq C \sum_{j=0}^\infty \sum_{k=0}^\infty |f| \ast K^{j,k}.
\]

Consider $j + k \geq 1$. We have that
\[
\int_{\mathbb{R}^{d+1}} K^{j,k}(x, u) \, dx \, du = \int_{\mathbb{R}^{d+1}} \chi_2(x'') \mathcal{F}(\phi_j)(x_d - \Gamma_2(x'')) \mathcal{F}(\phi_k)(u) \, dx \, du
\]
\[
= \int_{\mathbb{R}^{d+1}} \chi_2(x'') \mathcal{F}(\phi_j)(x_d) \mathcal{F}(\phi_k)(u) \, dx \, du = 0,
\]
since $\int_{\mathbb{R}} \mathcal{F}(\phi_l)(s) \, ds = 0$ if $l > 0$. □

Springer
5.1 Square functions and almost orthogonality

Lemma 5.1 For all \( j + k \geq 1 \) and \( \theta \in [0, 1] \), we have

\[
\left\| \sup_{t > 0} \left| f * K_{t}^{j,k} \right| \right\|_{L^2(\mathbb{R}^{d+1})} \leq C(j + k) 2^{-(d-\frac{3}{2}-\frac{d-1}{2}\theta)j} 2^{-\theta k} \| f \|_{L^2(\mathbb{R}^{d+1})},
\]

(5.5)

Proof We apply Lemma 2.2 to see that (5.5) follows from the following estimates which are uniform in \( n \) and \( \theta \in [0, 1] \):

\[
\left( \sum_{n \in \mathbb{Z}} \left\| f * K_{2^n t}^{j,k} \right\|_{L^2(\mathbb{R}^{d+1})}^{2} \right)^{1/2} \leq C(j + k) 2^{-(1-\theta)\frac{d+1}{2}} 2^{-\theta k} \| f \|_{L^2(\mathbb{R}^{d+1})},
\]

(5.6)

and

\[
\left( \sum_{n \in \mathbb{Z}} \left\| f * \left( \sum_{k=1}^{n} K_{s}^{j,k} \right)^{2} \right\|_{L^2(\mathbb{R}^{d+1})}^{2} \right)^{1/2} \leq C(j + k) 2^{-\theta k} \| f \|_{L^2(\mathbb{R}^{d+1})},
\]

(5.7)

By Lemma 2.3, (5.6) follows from the following estimates (5.8) and (5.9):

\[
\| f * K_{2^n}^{j,k} \|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{-(1-\theta)\frac{d+1}{2}} 2^{-\theta k} \| f \|_{L^2(\mathbb{R}^{d+1})}, \quad \forall \theta \in [0, 1]
\]

(5.8)

and

\[
\| f * (K_{2^n}^{j,k})^{*} K_{2^n}^{j,k} \|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4j+4k} 2^{-|n-n'|} \| f \|_{L^2(\mathbb{R}^{d+1})}
\]

(5.9)

first for \( n \leq n' \) and then by taking adjoints also for \( n > n' \). Again by Lemma 2.3, (5.7) follows from the following estimates (5.10) and (5.11):

\[
\left\| f * \left( \sum_{k=1}^{n} K_{s}^{j,k} \right)^{2} \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{-(d-\frac{3}{2}-\frac{d-1}{2}\theta)j} 2^{-\theta k} \| f \|_{L^2(\mathbb{R}^{d+1})},
\]

(5.10)

and

\[
\left\| f * \left( \sum_{k=1}^{n} K_{s}^{j,k} \right)^{2} \right\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{6j+6k} 2^{-|n-n'|} \| f \|_{L^2(\mathbb{R}^{d+1})}
\]

(5.11)

The proof of (5.8), (5.9), (5.10) and (5.11) will be given in Sects. 5.1.2 and 5.1.3. \( \square \)

5.1.1 Reduction to oscillatory integral operators

To show estimates (5.8), (5.9), (5.10) and (5.11), we will reduce them to estimates for oscillatory integral operators. Recall that \( E_{d-1} \) is a \( d \times (d - 1) \) matrix and \( F_{\kappa} \) is a \( (d - 1) \times \kappa \) matrix with \( \kappa \in [2, d) \), which are given by

\[
E_{d-1} = \begin{pmatrix} I_{d-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad F_{\kappa} = \begin{pmatrix} I_{\kappa} & 0 \\ 0 & 0 \end{pmatrix}.
\]

(5.12)

respectively, and we denote \( E_{d-1}^{T}, F_{\kappa}^{T} \) as their transposes. We write

\[
f * K^{j,k}(x, u) = \int_{\mathbb{R}^{d+1}} K^{j,k}(x - y, u - v + x''^{T} E_{d-1}^{T} E_{d-1} y'') f(y, v) dy dv.
\]
where
\[ K^{j,k}(x-y, u-v + x''^T E_{d-1}^T J E_{d-1} y'') = \chi_2(x'' - y'') \]
\[ \int_{\mathbb{R}^2} e^{i(\sigma(x_y - \Gamma_2(x'' - y'')) + \tau(u-v + x''^T E_{d-1}^T J E_{d-1} y''))} \phi_j(\sigma) \phi_k(\tau) d\sigma d\tau. \]

We then apply a Fourier transform on \( \mathbb{R}^2 \), in the \((x_y, u)\) variables of \( f \ast K^{j,k}(x, u) \), and use the fact that \( K^{j,k}(x, u) = \chi_2(x'') \mathcal{F}^{-1}(\phi_j)(x_y - \Gamma_2(x'')) \mathcal{F}^{-1}(\phi_k)(u) \) to obtain
\[ \mathcal{F}_d \mathcal{F}_u (f \ast K^{j,k})(x'', \lambda_1, \lambda_2) = \phi_j(\lambda_1) \phi_k(\lambda_2) \int_{\mathbb{R}^{d-1}} e^{-i(\lambda_1 \Gamma_2(x'' - y'') - \lambda_2 x''^T E_{d-1}^T J E_{d-1} y'')} \mathcal{F}_d \mathcal{F}_u f(y'') dy''. \]

We have the following result.

Lemma 5.2 Define
\[ T_{\lambda_1, \lambda_2} g(x'') = \int_{\mathbb{R}^{d-1}} e^{-i(\lambda_1 \Gamma_2(x'' - y'') - \lambda_2 x''^T E_{d-1}^T J E_{d-1} y'')} \chi_2(x'' - y'') g(y'') dy''. \]
Then there exists a constant \( C > 0 \) depending on \( \text{supp} \chi_2 \) and \( C^N \) norms of \( \chi_2 \) and \( \Gamma_2 \) on \( \text{supp} \chi_2 \) for some sufficiently large \( N \) such that for all \( \theta \in [0, 1] \),
\[ \| T_{\lambda_1, \lambda_2} g \|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_1|)^{-\frac{d-1}{\theta}} (1 + |\lambda_2|)^{-\theta} \| g \|_{L^2(\mathbb{R}^{d-1})}. \]

Proof One can rewrite
\[ T_{\lambda_1, \lambda_2} g(x'') = \int_{\mathbb{R}^{d-1}} e^{-i\Psi_{\lambda_1, \lambda_2}(x'', y'')} \chi_2(x'' - y'') g(y'') dy'', \]
where the phase function \( \Psi_{\lambda_1, \lambda_2} \) is given by
\[ \Psi_{\lambda_1, \lambda_2}(x'', y'') = \Gamma_2(x'' - y'') - \lambda_1^{-1} \lambda_2 x''^T E_{d-1}^T J E_{d-1} y''. \]

We have \( \partial^2_{x''} \Gamma_2 \Psi_{\lambda_1, \lambda_2}(x'', y'') = -\partial^2_{x''} \Gamma_2(x'' - y'') - \lambda_1^{-1} \lambda_2 E_{d-1}^T J E_{d-1} \). Since \( \partial^2_{x''} \Gamma_2(x'' - y'') \) is a positive definite matrix and \( E_{d-1}^T J E_{d-1} \) is skew-symmetric, it follows that for all \( |X| = 1 \),
\[ |\partial^2_{x''} \Gamma_2 \Psi_{\lambda_1, \lambda_2}(x'', y'') X| \geq |(\partial^2_{x''} \Gamma_2(x'' - y'') X + \lambda_1^{-1} \lambda_2 E_{d-1}^T J E_{d-1} X, X)| \]
\[ = |(\partial^2_{x''} \Gamma_2(x'' - y'') X, X)| \geq c \]
for some \( c > 0 \). By (5.14), we have
\[ |\partial^2_{x''} \Gamma_2 \Psi_{\lambda_1, \lambda_2}(x'', y'')| = |\partial^2_{x''} \Gamma_2(x'' - y'')| \leq C \alpha \]
for all \( x'' - y'' \in \text{supp} \chi_2 \) and \( |\alpha| \geq 3 \). Then by Lemma 4.2,
\[ \| \chi_2 T_{\lambda_1, \lambda_2} g \|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_1|)^{-\frac{d-1}{\theta}} \| g \|_{L^2(\mathbb{R}^{d-1})}. \]

By \( \det J_k \neq 0 \), we have
\[ \inf_{|X|=1} |J_k X| \geq c_k \text{ and } \sup_{|X|=1} |J_k F \partial^2_{x''} \Gamma_2(x'' - y'') F_k X| \leq C \Gamma_{2,k}, \text{ if } x'' - y'' \in \text{supp} \chi_2. \]

Take \( M = 2C \Gamma_{2,k}/c_k \). Now we consider two cases: \( M |\lambda_1| \geq |\lambda_2| \) and \( M |\lambda_1| \leq |\lambda_2| \).
Case 1: $M |\lambda_1| \geq |\lambda_2|$. In this case, we use (5.15) to obtain
$$\|\chi_2 T_{\lambda_1, \lambda_2} g\|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_1|)^{-(1-\theta)^{-\frac{d-1}{2}} + \frac{1}{2}} (1 + |\lambda_2|)^{-\theta} \|g\|_{L^2(\mathbb{R}^{d-1})}. \quad (5.17)$$

Case 2: $M |\lambda_1| \leq |\lambda_2|$. In this case, we rewrite
$$T_{\lambda_1, \lambda_2} g(x'') = \int_{\mathbb{R}^{d-1}} e^{-ik_2 \tilde{\psi}_{\lambda_1, \lambda_2}(x'', y'')} \chi_2(x'' - y'') g(y'') dy'',$$
where $\tilde{\psi}_{\lambda_1, \lambda_2}(x'', y'') = \lambda_1 \lambda_2^{-1} \Gamma_2(x'' - y'') - x'' E_{d-1}^T J E_{d-1} y''$ and thus
$$\partial_{x'', y''}^2 \tilde{\psi}_{\lambda_1, \lambda_2}(x'', y'') = -\lambda_1 \lambda_2^{-1} \partial_x^2 \Gamma_2(x'' - y'') - E_{d-1}^T J E_{d-1}.$$

By a triangle inequality and (5.16),
$$|\partial_{(x_1, \ldots, x_k), (y_1, \ldots, y_k)}^2 \tilde{\psi}_{\lambda_1, \lambda_2}(x'', y'') X| = | - \lambda_1 \lambda_2^{-1} F_k^T \partial_x^2 \Gamma_2(x'' - y'') F_k X - J_k X|$$
$$\geq |J_k X| - M^{-1} |F_k^T \partial_x^2 \Gamma_2(x'' - y'') F_k X|$$
$$\geq c_{j_k} - M^{-1} c_{j_k} c_{j_k} \geq c_{j_k} / 2$$

for all $|X| = 1$. Notice that for all $|\alpha| \geq 2$, $|\partial_{x'', y''}^\alpha \tilde{\psi}_{\lambda_1, \lambda_2}(x'', y'')| \leq C_{\alpha}$. We then apply Lemma 4.2 in the first $k (\kappa \geq 2)$ variables to get
$$\|\chi_2 T_{\lambda_1, \lambda_2} g\|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_2|)^{-\kappa/2} \|g\|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_2|)^{-1} \|g\|_{L^2(\mathbb{R}^{d-1})}.$$

This, together with (5.15), tells us that for all $M |\lambda_1| \leq |\lambda_2|$,
$$\|\chi_2 T_{\lambda_1, \lambda_2} g\|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_1|)^{-(1-\theta)^{-\frac{d-1}{2}} + \frac{1}{2}} (1 + |\lambda_2|)^{-\theta} \|g\|_{L^2(\mathbb{R}^{d-1})}. \quad (5.18)$$

Therefore, it follows from (5.17) and (5.18) that for all $\lambda_1, \lambda_2 \in \mathbb{R}$,
$$\|\chi_2 T_{\lambda_1, \lambda_2} g\|_{L^2(\mathbb{R}^{d-1})} \leq C (1 + |\lambda_1|)^{-(1-\theta)^{-\frac{d-1}{2}} + \frac{1}{2}} (1 + |\lambda_2|)^{-\theta} \|g\|_{L^2(\mathbb{R}^{d-1})}.$$

Then by a translation invariance argument, we complete the proof of Lemma 5.2. □

5.1.2 Proof of (5.8) and (5.9)

Proof of (5.8) By (5.13) and Lemma 5.2, we have
$$\|f * K^{j,k}\|_{L^2(\mathbb{R}^{d+1})} = \|\mathcal{F}_d \mathcal{F}_u (f * K^{j,k})\|_{L^2(\mathbb{R}^{d+1})} = \|T_{\lambda_1, \lambda_2} (\mathcal{F}_d \mathcal{F}_u f (., \lambda_1, \lambda_2))(x'') \|_{L^2_{\lambda_1, \lambda_2}(\mathbb{R}^d)L^2_{x''}(\mathbb{R}^{d-1})} \leq C 2^{-(1-\theta)^{-\frac{d-1}{2}} + \frac{1}{2}} j 2^{-\theta k} \|\mathcal{F}_d \mathcal{F}_u f (x'', \lambda_1, \lambda_2)\|_{L^2_{\lambda_1, \lambda_2}(\mathbb{R}^d)L^2_{x''}(\mathbb{R}^{d-1})} \leq C 2^{-(1-\theta)^{-\frac{d-1}{2}} + \frac{1}{2}} j 2^{-\theta k} \|f\|_{L^2(\mathbb{R}^{d+1})} \quad \text{for all } j, k \geq 0 \text{ and } \theta \in [0, 1]. \square$$

Proof of (5.9) By a scaling argument, the proof of (5.9) reduces to show for all $n \leq 0$,
$$\|f * (K^{j,k} * K^{j,k})\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4j+4k} 2^n \|f\|_{L^2(\mathbb{R}^{d+1})} \quad (5.19)$$
and
$$\|f * (K^{j,k} * K^{j,k})\|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{4j+4k} 2^n \|f\|_{L^2(\mathbb{R}^{d+1})}. \quad (5.20)$$
Let us prove estimate (5.19). We use integration by parts to see that for all \( |\alpha| \leq 1 \) and \( N > 0 \), there exists a constant \( C = C(\alpha, N) > 0 \) such that

\[
|\partial^\alpha_{x,u} K^{j,k}(x, u)| + |\partial^\alpha_{x,u} (K^{j,k})^*(x, u)| \leq C 2^{2j+2k} \chi_2(x'')(1 + |x_d| + |u|)^{-N}.
\] (5.21)

Note that \(|x^T J y| = |x''^T E_{d-1}^T J E_{d-1} y''| \leq c \|J\| \cdot |y''| \leq c^2 \|J\| \text{ if } |x''|, |y''| \leq c \). These, in combination with the cancellation property (5.4) of the kernel of \( K^{j,k} \), yield

\[
|K^{j,k} * K^{j,k}_2(x, u)| \leq C \int_0^1 \int_{\mathbb{R}^{d+1}} 2^{2j+2k} \left(1 + |x - \theta y| + |u - \theta v|\right)^{-N} \left(|y| + |v|\right) |
\]

This gives that for \( n \leq 0 \),

\[
\int_{\mathbb{R}^{d+1}} |(K^{j,k})^* K^{j,k}_2(x, u)| dx du \leq C 2^{2j+2k} \int_{\mathbb{R}^{d+1}} (|y| + |v|)|K^{j,k}_2(y, v)| dy dv
\]

By the Schur lemma, (5.19) follows readily.

For (5.20), we use the cancellation property of the kernel of \( (K^{j,k}_2)^* \) and a similar argument as above to show that

\[
\int_{\mathbb{R}^{d+1}} |(K^{j,k}_2)^* K^{j,k}(x, u)| dx du \leq C 2^{4j+4k} 2^n.
\]

By the Schur lemma again, (5.20) follows readily. \( \square \)

### 5.1.3 Proof of (5.10) and (5.11)

**Proof of (5.10)** By (5.2) and change of variables in \( \sigma, \tau \), we have

\[
K^{j,k}_s(x, u) = s^{-d+1} \chi_2(s^{-1}x'') \int_{\mathbb{R} \times \mathbb{R}} e^{i\sigma(x_d-s\Gamma_2(s^{-1}x''))} e^{i\tau u} \phi_j(s\sigma) \phi_k(s^2\tau) d\sigma d\tau. \tag{5.22}
\]

Note that \( s \frac{d}{ds} K^{j,k}_s(x, u) \) can be written as the finite sum of following terms,

\[
s \frac{d}{ds} K^{j,k}_s(x, u) = (K^{(1)}_s)^{j,k} + (K^{(2)}_s)^{j,k} + 2j (K^{(3)}_s)^{j,k}
\] (5.23)

where

\[
(K^{(1)}_s)^{j,k} = \left( (-d+1) \chi_2(s^{-1}x'') \right) s^{-d+1} \int_{\mathbb{R} \times \mathbb{R}} e^{i\sigma(x_d-s\Gamma_2(s^{-1}x''))} e^{i\tau u} \phi_j(s\sigma) \phi_k(s^2\tau) d\sigma d\tau
\]

\[
(K^{(2)}_s)^{j,k} = s^{-d+1} \chi_2(s^{-1}x'') \int_{\mathbb{R} \times \mathbb{R}} e^{i\sigma(x_d-s\Gamma_2(s^{-1}x''))} e^{i\tau u} \phi_j(s\sigma) \phi_k(s^2\tau) d\sigma d\tau
\]

\[
(K^{(3)}_s)^{j,k} = \left[ -s^{-d+1} \Gamma_2(s^{-1}x'') + is^{-d+1} \Gamma'_2(s^{-1}x'') \right] \chi_2(s^{-1}x'') \int_{\mathbb{R} \times \mathbb{R}} e^{i\sigma(x_d-s\Gamma_2(s^{-1}x''))} e^{i\tau u} 2^{-j} s\sigma \phi_j(s\sigma) \phi_k(s^2\tau) d\sigma d\tau.
\]
The method as in the proof of (5.8) yields that for all \( i = 1, 2, 3 \), we have
\[
\| f * (K^{(i)})^{j,k} \|_{L^2(\mathbb{R}^{d+1})} \leq C 2^{-(1-\theta) \frac{d-1}{2} + j 2^{-\theta k}} \| f \|_{L^2(\mathbb{R}^{d+1})}, \quad \forall \theta \in [0, 1].
\]
These, together with a scaling argument, imply
\[
\sup_{t \in [1,2]} \left\| f * \left[ \frac{d}{ds} K^j_k \right] \right\|_{L^2(\mathbb{R}^{d+1})} \leq 2^j \sum_{i=1}^{3} \sup_{t \in [1,2]} \| f * (K^{(i)})^{j,k} \|_{L^2(\mathbb{R}^{d+1})}
\]
\[
\leq C 2^j 2^{-(1-\theta) \frac{d-1}{2} + j 2^{-\theta k}} \| f \|_{L^2(\mathbb{R}^{d+1})}
\]
\[
= C 2^{-(\frac{d-3}{2} - \frac{d-1}{2}) \theta - \theta k} \| f \|_{L^2(\mathbb{R}^{d+1})}.
\]
Estimate (5.10) follows readily. \( \square \)

**Proof of (5.11)** By (5.23), we can follow the proof of (4.16) to obtain
\[
\left\| f \left( \left[ \frac{d}{ds} K^j_k \right] \right)^* \right\|_{L^2(\mathbb{R}^{d+1})} \leq 2^j \sum_{i,i'=1}^{3} \left\| f * \left( (K^{(i)})^{j,k}_2 \right)^* \left( (K^{(i')})^{j,k}_2 \right)^* \right\|_{L^2(\mathbb{R}^{d+1})}
\]
\[
\leq C 2^{6j+6k} 2^{-|n-n'|} \| f \|_{L^2(\mathbb{R}^{d+1})}.
\]
This completes the proof of (5.11). \( \square \)

### 5.2 Weak type \((1, 1)\)-estimate

For \( p = 1 \), we have the following result.

**Lemma 5.3** Let \( j, k \geq 0 \). For all \( \alpha > 0 \), we have
\[
\left| \{ (x, u) : \sup_{t>0} |f * K^j_k (x, u)| > \alpha \} \right| \leq C (j + k) 2^j \alpha^{-1} \| f \|_{L^1(\mathbb{R}^{d+1})}.
\]

**Proof** The proof of Lemma 5.3 is essentially similar to that of Lemma 4.4. Let \( \rho(x, u) := |x| + |u|^{1/2} \) and \( c_J = \frac{3}{2} \| J \|^{1/2} + 1 \). Like Lemma 4.4, it suffices to show for all \( \rho(y, v) \leq 2^{-n} \delta \) and \( \delta > 0 \),
\[
\sum_{n \in \mathbb{Z}} \int_{\rho(x, u) \geq 2c_J 2^{-n} \delta} \sup_{s \in [1,2]} \left| K^j_k ((y, v)^{-1}(x, u)) - K^j_k (x, u) \right| dx du \leq C (j + k) 2^j.
\]

(5.24)

To show (5.24), we first note that from (5.2), we have that for any \( j, k \geq 0 \),
\[
\left| K^j_k (x, u) \right| \leq C 2^{j+k} (1 + |x| + |u|^{1/2})^{-N/2}
\]
\[
\left( 1 + 2^j |s^{-1} x_d - \Gamma_2 (x'') \right)^{-N/2} \left( 1 + 2^{k/2} |u|^{1/2} \right)^{-N/2},
\]
\[
(5.25)
\]
\[
\sup_{s \in [1,2]} \left| \nabla_{x,u} K^j_k (x, u) \right| \leq C 2^{2j+2k} (1 + |x| + |u|^{1/2})^{-N/2}
\]
\[
(5.26)
\]
for large $N \geq 1$; and by (5.23),
\[
\left| \frac{d}{ds} K^{j,k}_s(x,u) \right| \leq C 2^{j/2} 2^k (1 + |x| + |u|^{1/2})^{-N/2} \left( 1 + 2^j |s^{-1} x_d - \Gamma_2 (s^{-1} x'')| \right)^{-N/2} \left( 1 + 2^{k/2} |u|^{1/2} \right)^{-N/2}.
\]

uniformly in $s \in [1, 2]$. 
If $\rho(y, v) \leq 1$, by the compact $x''$-support of $K^{j,k}$, we can suppose $|x''| \leq C + 1$ and
\[
|(y, v + y^T J x)| = |(y, v + y''^T E^T_{d-1} J E_{d-1} x'')| \leq C \rho(y, v).
\]

5.3 Proof of (ii) of Theorem 2.1

Proof of (ii) of Theorem 2.1 First we consider $1 < p \leq 2$. By (5.3), we have
\[
\sup_{t > 0} |f \ast (\mu_2)_t| \leq C \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \sup_{t > 0} \| f \ast K^{j,k}_t \|.
\]

Because the term $j = k = 0$ can be controlled by the appropriate variant of the Hardy–Littlewood maximal function, it suffices to consider that $j + k \geq 1$. By interpolation, it follows from Lemmas 5.1 and 5.3 that for all $\theta \in [0, 1]$,
\[
\left\| \sup_{t > 0} \| f \ast K^{j,k}_t \| \right\|_{L^p(\mathbb{R}^{d+1})} \leq C (j + k) 2^{-(d/2 - d/2 - 1)\theta/(2 - \theta)} 2^{-\theta/2} 2^{(3/2 - 1)j} \| f \|_{L^p(\mathbb{R}^{d+1})}.
\]
\[
= C (j + k) 2^{(5/2 - 1) - (d - 2)(1 - \theta)(1 - \theta/2)} 2^{(d - 1)(1 - \theta/2)} \| f \|_{L^p(\mathbb{R}^{d+1})}.
\]

The summation in $k$ converges automatically and the condition
\[
\left[ \left( \frac{2}{p} - 1 \right) - (d - 2) \left( 1 - \frac{1}{p} \right) \right] + (d - 1) \left( 1 - \frac{1}{p} \right) \theta < 0 \text{ for some } \theta \in (0, 1)
\]
makes sure that the summation in $j$ converges. It holds if and only if $(\frac{2}{p} - 1) - (d - 2)(1 - \frac{1}{p}) < 0$. As a result, the summation in $j, k$ converges whenever $d/(d - 1) < p \leq 2$ with $d \geq 3$ and $\theta$ is sufficiently small. By interpolation with a trivial $L^\infty$ estimate, we obtain estimate (ii) of Theorem 2.1. \hfill \Box

Acknowledgements We would like to thank the referee for a careful reading of this manuscript and providing many helpful comments and suggestions. The authors were supported by National Key R&D Program of China 2022YFA1005700. N.J. Liu was supported by China Postdoctoral Science Foundation (No. 2022M723673). L.X. Yan was supported by the NNSF of China, Grant (Nos. 11871480 and 12171489). The authors thank Shaoming Guo for helpful suggestions and discussions.

Data availability No associated data used.

Declarations

Conflict of interest There is no conflict of interest.
References

1. Anderson, T.C., Cladek, L., Pramanik, M., Seeger, A.: Spherical means on the Heisenberg group: stability of a maximal function estimate. J. Anal. Math. 145(1), 1–28 (2021)
2. Bourgain, J.: Averages in the plane over convex curves and maximal operators. J. Anal. Math. 47, 69–85 (1986)
3. Cuccagna, S.: $L^2$ estimates for averaging operators along curves with two-sided $k$-fold singularities. Duke Math. J. 89, 203–216 (1997)
4. Cowling, M.: On Littlewood–Paley–Stein theory. Rend. Circ. Mat. Palermo (2) 1, 21–55 (1981)
5. Greenleaf, A., Seeger, A.: On oscillatory integral operators with folding canonical relations. Stud. Math. 132, 125–139 (1999)
6. Fischer, V.: The spherical maximal function on the free two-step nilpotent Lie group. Math. Scand. 99, 99–118 (2006)
7. Hörmander, L.: Oscillatory integrals and multipliers on $FL^p$. Ark. Mat. 11, 1–11 (1973)
8. Mockenhaupt, G., Seeger, A., Sogge, C.D.: Wave front sets, local smoothing and Bourgain’s circular maximal theorem. Ann. Math. 136, 207–218 (1992)
9. Müller, D., Seeger, A.: Singular spherical maximal operators on a class of two step nilpotent Lie groups. Isr. J. Math. 141, 315–340 (2004)
10. Nevo, A., Thangavelu, S.: Pointwise ergodic theorems for radial averages on the Heisenberg group. Adv. Math. 127, 307–334 (1997)
11. Narayanan, E.K., Thangavelu, S.: An optimal theorem for the spherical maximal operator on the Heisenberg group. Isr. J. Math. 144, 211–219 (2004)
12. Pramanik, M., Seeger, A.: $L^p$ regularity of averages over curves and bounds for associated maximal operators. Am. J. Math. 129, 61–103 (2007)
13. Pramanik, M., Seeger, A.: $L^p$–Sobolev estimates for a class of integral operators with folding canonical relations. J. Geom. Anal. 31, 6725–6765 (2021)
14. Schlag, W.: A geometric proof of the circular maximal theorem. Duke Math. J. 93, 505–533 (1998)
15. Schlag, W., Sogge, C.D.: Local smoothing estimates related to the circular maximal theorem. Math. Res. Lett. 4, 1–15 (1997)
16. Stein, E.M.: Maximal functions. I. Spherical means. Proc. Natl. Acad. Sci. USA 73, 2174–2175 (1976)
17. Stein, E.M.: Harmonic Analysis: Real Variable Methods, Orthogonality and Oscillatory Integrals. Princeton Univ. Press, Princeton (1993)
18. Stein, E.M., Wainger, S.: Problems in harmonic analysis related to curvature. Bull. Am. Math. Soc. 84, 1239–1295 (1978)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.