Abstract

We prove a variable coefficient version of the square function estimate of Guth–Wang–Zhang. By a classical argument of Mockenhaupt–Seeger–Sogge, it implies the full range of sharp local smoothing estimates for 2 + 1-dimensional Fourier integral operators satisfying the cinematic curvature condition. In particular, the local smoothing conjecture for wave equations on compact Riemannian surfaces is settled.

MSC 2020

42B37 (primary), 35L05, 35S30 (secondary)
1 INTRODUCTION

1.1 Local smoothing conjectures

Suppose that \((M, g)\) is a \(d\)-dimensional smooth Riemannian manifold without boundary, and \(u(x, t)\) is a solution to the Cauchy problem

\[
\begin{aligned}
\left\{ \begin{array}{l}
(\partial_t^2 - \Delta_g) u(x, t) = 0, \ (x, t) \in M \times \mathbb{R} \\
u(0, x) = u_0(x), \ \partial_t u(x, 0) = u_1(x)
\end{array} \right.,
\end{aligned}
\]

where \(\Delta_g\) denotes the associated Beltrami–Laplacian operator.

When the underlying manifold is Euclidean, that is, \(M = \mathbb{R}^d\) with flat metric, Sogge [22] conjectured that for \(p \geq p_d = \frac{2d}{d-1}\), the space-time norm \(\|u\|_{L^p(\mathbb{R}^d \times [1, 2])}\) is bounded by suitable Sobolev norms of the initial data, which represents a gain in regularity compared to fixed time estimates.

**Conjecture 1.1** (Local smoothing in \(\mathbb{R}^{d+1}\)). Suppose \(M = \mathbb{R}^d\) equipped with flat metric, \(u\) is a solution to the Cauchy problem (1.1), and \(s_p = \frac{d-1}{2} - \frac{d-1}{p}\). Then for any \(p \geq p_d = \frac{2d}{d-1}\), \(\sigma < \frac{1}{p}\), we have

\[
\|u\|_{L^p(\mathbb{R}^d \times [1, 2])} \leq C_\sigma \left( \|u_0\|_{L^p_{s_p}(\mathbb{R}^d)} + \|u_1\|_{L^p_{s_p-1}(\mathbb{R}^d)} \right).
\]

Here \(s_p\) represents the order of regularity from the fixed time estimate; thus, Conjecture 1.1 asserts that integrating in time provides a \(\sigma < 1/p\) order of regularity gain for \(p \geq p_d\). Conjecture 1.1, if true, would be a remarkable result that would imply several major open problems in modern harmonic analysis, including the Bochner–Riesz mean, Fourier restriction, and Kakeya–Nikodym conjectures, all of which are wide open in dimensions \(d \geq 3\). See, for example, [25] for a detailed discussion. In a recent work of Guth–Wang–Zhang [12], Conjecture 1.1 was verified for \(d = 2\).

The situation for a general compact Riemannian manifold \((M, g)\) is quite different. It was shown by Minicozzi and Sogge [19] (see also [23]) that there exist compact manifolds \((M, g)\) that exhibit a certain Kakeya compression phenomenon that forbids favorable Kakeya-type estimates. Therefore, for such manifolds, local smoothing fails for all \(\sigma < 1/p\) if

\[
p < \tilde{p}_{d,+} := \begin{cases} 
\frac{2(3d + 1)}{3d - 3}, & \text{if } d \text{ is odd,} \\
\frac{2(3d + 2)}{3d - 2}, & \text{if } d \text{ is even.}
\end{cases}
\]

This result naturally leads to the following conjecture.
Conjecture 1.2 (Local smoothing for compact manifolds). Let \((M, g)\) be a compact Riemannian manifold with \(u\) being a solution to the Cauchy problem (1.1). For any \(p \geqslant \tilde{p}_d, \sigma < \frac{1}{p}\), we have
\[
\|u\|_{L^p(\mathbb{R}^d \times [1,2])} \leq C_\sigma (\|u_0\|_{L^p_{\tilde{p}_d-\sigma}(\mathbb{R}^d)} + \|u_1\|_{L^p_{\tilde{p}_d-1-\sigma}(\mathbb{R}^d)}).
\] (1.4)

Conjecture 1.2 is open in all dimensions \(d \geqslant 2\). A classical way of attacking Conjecture 1.2 is to utilize a local parametrix to the Cauchy problem (1.1) in terms of Fourier integral operators\(^\dagger\) (FIOs). In particular, it is known that
\[
u(x, t) = F_0 u_0(x, t) + F_1 u_1(x, t),
\] (1.5)

where \(F_j \in I^{j-\frac{1}{2}}(M \times \mathbb{R}, M; \mathcal{C})\) are FIOs with canonical relation \(\mathcal{C}\) satisfying the cinematic curvature condition. In local coordinates, an FIO mapping functions on \(\mathbb{R}^d\) to ones on \(\mathbb{R}^{d+1}\) can be written as a finite sum of operators of the form
\[
F f(z) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\phi(z, \eta)} a(z, \eta) \hat{f}(\eta) d\eta, \quad z = (x, t) \in \mathbb{R}^{d+1}, \eta \in \mathbb{R}^d,
\] (1.6)

where \(a \in S^0(\mathbb{R}^{d+1} \times \mathbb{R}^d)\) is a symbol with compact support in \((x, t)\), and \(\phi \in C^\infty(\mathbb{R}^{d+1} \times (\mathbb{R}^d \setminus \{0\}))\) is a phase function homogeneous of degree 1 in \(\eta\). We say that \(F\) satisfies the (local) cinematic curvature condition if the following conditions are satisfied.

- **Nondegeneracy condition.**
  \[
  \text{rank } \partial^2_{x\eta} \phi(x, t, \eta) = d, \text{ for all } (z, \eta) = (x, t, \eta) \in \text{supp } a.
  \] (1.7)

- **Cone condition.** Consider the Gauss map \(G : \text{supp } a \rightarrow S^d\) by \(G(z, \eta) := \frac{G_0(z, \eta)}{|G_0(z, \eta)|}\) where
  \[
  G_0(z, \eta) := \bigwedge_{j=1}^d \partial_{\eta_j} \partial_z \phi(z, \eta).
  \]

The curvature condition
\[
\text{rank } \partial^2_{\eta\eta} \langle \partial_z \phi(z, \eta), G(z, \eta_0) \rangle |_{\eta = \eta_0} = d - 1
\] (1.8)

holds for all \((z, \eta_0) \in \text{supp } a\).

Although we introduced the concept of FIOs in relation to Conjecture 1.2, the numerology for their local smoothing conjectures is yet again different. It is shown in [2] that examples constructed by Bourgain [3, 4] can be used to identify FIOs satisfying the cinematic curvature condition for which local smoothing fails for all \(\sigma < 1/p\), if
\[
p < \tilde{p}_d := \begin{cases} 
2(d + 1) & \text{if } d \text{ is odd}, \\
2(d + 2) & \text{if } d \text{ is even}.
\end{cases}
\] (1.9)

The local smoothing conjecture for FIOs is the following.

\(^\dagger\) See, for example, [13] for a thorough treatment for the basic theory of FIOs.
**Conjecture 1.3** (Local smoothing for FIOs). Let $F$ be an FIO with symbol of order $\mu$ satisfying the cinematic curvature condition. For any $p \geq \tilde{p}_d$, $\sigma < \frac{1}{p}$, we have

$$
\|Ff\|_{L^p(\mathbb{R}^{d+1})} \leq C_\sigma \|f\|_{L^p_{\mu+sp-\sigma}(\mathbb{R}^d)}. \tag{1.10}
$$

We remark that Conjecture 1.2 essentially corresponds to the subcase of Conjecture 1.3 when the matrix in (1.8) is positive definite. Recently, Conjecture 1.3 has been confirmed by Beltran–Hickman–Sogge [2] for odd spatial dimensions $d$. Nevertheless, Conjecture 1.3 is open for all even values of $d \geq 2$.

### 1.2 Main results

The purpose of this article is to confirm Conjecture 1.3 for $d = 2$.

**Theorem 1.4** (Local smoothing for FIOs in $2+1$ dimensions). Let $F$ be an FIO with symbol of order $\mu$ satisfying the cinematic curvature condition. For any $p \geq \tilde{p}_2 = 4$, $\sigma < \frac{1}{p}$, we have

$$
\|Ff\|_{L^p(\mathbb{R}^{2+1})} \leq C_\sigma \|f\|_{L^p_{\mu+sp-\sigma}(\mathbb{R}^2)}. \tag{1.11}
$$

We remark that $p_d \leq \tilde{p}_+, d \leq \tilde{p}_d$, while $p_d = \tilde{p}_+, d = \tilde{p}_d$ if and only if $d = 2$. Therefore, as a direct consequence of Theorem 1.4, we also confirm Conjecture 1.2 for $d = 2$.

**Corollary 1.5** (Local smoothing for compact surfaces). Let $(M, g)$ be a compact Riemannian surface and $u$ be a solution to the Cauchy problem (1.1). For any $p \geq \tilde{p}_{2,+} = 4$, $\sigma < \frac{1}{p}$, we have

$$
\|u\|_{L^p(M \times [1,2])} \leq C_\sigma (\|u_0\|_{L^p_{\mu+sp-\sigma}(M)} + \|u_1\|_{L^p_{sp-1-\sigma}(M)}). \tag{1.12}
$$

Let us take a moment to briefly survey the history of this problem. The study of Conjecture 1.1 in $2+1$ dimensions dates back to [22] where it was first introduced. In the same article, Sogge also established the first local smoothing estimate with a nontrivial gain of regularity at the endpoint $p = 4$. This result was improved by Mockenhaupt, Seeger, and Sogge [20], while no estimate with a sharp gain of regularity was established. Further improvements can be found in [16, 24, 27]. The seminal paper of Wolff [26] introduced an important tool now known as decoupling inequality for the cone, and used it to obtain sharp local smoothing estimates for $p \geq 74$. The methods of Wolff were further developed and extended by several authors [5, 10, 11, 15]. In [6], Bourgain and Demeter proved a sharp decoupling inequality for the cone for all dimensions $d \geq 2$, which naturally leads to local smoothing estimates in $\mathbb{R}^{d+1}$. In particular, when $d = 2$, their results imply sharp local smoothing estimates for all $p \geq 6$. On the other hand, a certain type of square function estimate is shown to imply local smoothing results by [20]. Following this route, Lee and Vargas [18] proved sharp local smoothing estimates for $p = 3$ by proving a sharp square function estimate using multilinear restriction estimates. As mentioned earlier, the recent work of Guth, Wang, and Zhang [12] settled Conjecture 1.1 for $d = 2$ completely, in which the authors established a sharp $L^4$ square function estimate.
Conjecture 1.2 and 1.3 were studied in parallel to the developments of Conjecture 1.1 [17, 21]. As mentioned above, Beltran, Hickman, and Sogge [2] established Conjecture 1.3 for odd spatial dimensions by proving a variable coefficient analog of Bourgain–Demeter decoupling inequality. In their proof, the authors exploited the multiplicativity of the decoupling constants and reduced the problem to proving a decoupling inequality at small scales via an induction on scales argument. At small enough scales, such a decoupling inequality follows directly from its Euclidean counterpart. More recent partial results can be found in [8, 9]. See the excellent survey article by Beltran, Hickman, and Sogge [1] for a detailed discussion of all three conjectures.

As discussed in [12], the sharp decoupling inequality for the cone does not imply a full range of sharp local smoothing estimates for Euclidean wave equations. This is one of the reasons why the authors resort to proving sharp square function estimates instead. Similarly, the variable coefficient decoupling inequalities of Beltran–Hickman–Sogge [2] do not fully resolve Conjecture 1.3 in even spatial dimensions. In this paper, Theorem 1.4 will be proven as a consequence of a variable coefficient generalization of the square function estimates of Guth–Wang–Zhang, which we shall describe in the next subsection.

### 1.3 Square function estimates for FIOs

The decomposition associated with our square function estimates is classical and the same as that in [2, 21]. Consider an FIO $F$ of the form (1.6) satisfying the cinematic curvature condition (i.e., (1.7) and (1.8)) with a symbol of order 0 and $d = 2$. Fix $\lambda \gg 1$. By a standard Littlewood–Paley decomposition, one may reduce Theorem 1.4 to proving that for any $\varepsilon > 0$,

$$
\| F^\lambda f \|_{L^4(\mathbb{R}^3)} \leq C \varepsilon \lambda^{1/4 + \varepsilon} \| f \|_{L^4(\mathbb{R}^2)},
$$

(1.13)

where $F^\lambda$ is defined as

$$
F^\lambda f(x, t) := \int e^{i \phi^\lambda(x, t, \eta)} a^\lambda(x, t, \eta) \hat{f}(\eta) \, d\eta,
$$

(1.14)

and

$$
\phi^\lambda(x, t, \eta) := \lambda \phi(x/\lambda, t/\lambda, \eta), \quad a^\lambda(x, t, \eta) := a(x/\lambda, t/\lambda, \eta).
$$

(1.15)

Here and throughout, we use the convention that $(x, t)$ are the physical variables and $\eta$ are the frequency variables. We may assume that the physical support of $a(x, t, \eta) \in C^\infty_c(\mathbb{R}^3 \times \mathbb{R}^2)$ is small, and its frequency support is near $e_2$, the basis vector $(0, 1) \in \mathbb{R}^2$.

Let $1 \leq R \leq \lambda$. We decompose the $\eta$-support of $a^\lambda$ angularly by cutting $S^1$ into $\approx R^{1/2}$ many arcs $\vartheta$, each $\vartheta$ spreading an angle $\approx R^{-1/2}$. Let $\{\chi_{\vartheta}\}_{\vartheta} \subset C^\infty(\mathbb{R}^2 \setminus 0)$ be a family of smooth cutoff functions associated with the decomposition in the angular direction, each of which is homogeneous of degree 0, such that $\{\chi_{\vartheta}\}_{\vartheta}$ forms a partition of unity on $\mathbb{R}^2 \setminus 0$. So, we have

$$
\sum_{\vartheta} \chi_{\vartheta}(\eta) \equiv 1, \quad \forall \eta \in \mathbb{R}^2 \setminus 0.
$$

(1.16)

Define

$$
F^\lambda_{\vartheta} f(x, t) := \int_{\mathbb{R}^2} e^{i \phi^\lambda(x, t, \eta)} a^\lambda_{\vartheta}(x, t, \eta) \hat{f}(\eta) \, d\eta,
$$

(1.17)
where the amplitude reads
\[
\alpha_\varnothing^\lambda(x, t, \eta) := a(x/\lambda, t/\lambda, \eta) \chi_\varnothing(\eta).
\] (1.18)

Thus,
\[
P_\varnothing^\lambda f = \sum_\varnothing P_\varnothing^\lambda f.
\]

Now we are ready to state our square function estimates for FIOs.

**Theorem 1.6** (Square function estimates for FIOs). Suppose that \( P_\varnothing^\lambda \) is an operator satisfying the assumptions above. Take \( R = \lambda \). Then for any \( \varepsilon > 0, N \in \mathbb{N} \), there exists constants \( C_\varepsilon, C_N > 0 \) such that
\[
\| P_\varnothing^\lambda f \|_{L^4(\mathbb{R}^3)} \leq C_\varepsilon \lambda^{\varepsilon} \left( \sum_\varnothing |P_\varnothing^\lambda f|^2 \right)^{1/2}_{L^4(\mathbb{R}^3)} + C_N \lambda^{-N} \| f \|_{L^4(\mathbb{R}^2)}.
\] (1.19)

**Remark 1.7.** Theorem 1.6 should be viewed as a variable coefficient generalization of [12, Theorem 1.1], which covers the phase function \( \phi(x, t, \xi) = x \cdot \xi + t |\xi| \). Nevertheless, it is not hard to see that a slight modification of their argument implies the same estimate for phase functions of the form \( \phi(x, t, \xi) = x \cdot \xi + tp(\xi) \), for any function \( p \in C^\infty(\mathbb{R}^2 \setminus \{0\}) \) homogeneous of degree 1 satisfying \( \text{rank } \partial^2 (\xi p(\xi)) = 1 \) on the support of the amplitude. In fact, one can see that the only place where the properties of a circular cone are used is [12, Lemma 4.2]. We include a proof of this lemma for the general cone case in the Appendix. In our case, the function \( p \) will be allowed to smoothly depend on \( x \) and \( t \), and thus, the cone in the phase space could also vary smoothly from point to point.

It is standard to deduce Theorem 1.4 from Theorem 1.6 and a sharp Nikodym maximal function estimate, as demonstrated in [21]; for reader’s convenience, we include a proof in the Appendix. It would be interesting to see if one can prove a higher dimensional analog of Theorem 1.6. Nevertheless, for \( d \geq 3 \), it is unclear whether we should expect the same conjectured threshold, \( \frac{2d}{d-1} \), as in the Euclidean version of this conjecture. Indeed, even though the critical exponents in Conjectures 1.1, 1.2, and 1.3 are all different for \( d \geq 3 \), there seems to be no argument that directly translates the Kakeya compression phenomena for FIOs to counterexample of square function estimates in this setting. It would be interesting to see if such examples can be constructed.

### 1.4 Strategy of proof

It is shown in [2] (see also [14]) that a variable coefficient decoupling inequality can be deduced from its translation-invariant counterpart by reducing it to a small-scale version through a “parabolic rescaling and induction on scales” argument. In [12], the proof of the sharp square function estimate for the cone in \( \mathbb{R}^{2+1} \) is also built upon an induction on scales argument, which utilizes a parabolic rescaling type argument. Thus, it might be tempting to think that one should be able to prove Theorem 1.6 by reducing matters to small scales and then use the square function estimate of Guth–Wang–Zhang as a black box. Unfortunately, there are significant differences between these two induction on scales arguments that seem to forbid one from doing so.
In [2], the authors considered “decoupling constants” $D(R, \lambda)$ with $\lambda$ being the scale associated with $F^\lambda$. Their main result, a variable coefficient decoupling inequality, corresponds to the estimate $D(\lambda^{1-\varepsilon}, \lambda) \leq C\lambda^\varepsilon$. The authors showed that at a tiny scale $K \ll \lambda^{\frac{1}{2} - \varepsilon}$, one could bound $D(K, \lambda)$ favorably by invoking Bourgain–Demeter and a stability lemma. In addition, the authors used a parabolic rescaling argument to bound $D(R, \lambda)$ by $D(K, \lambda) \cdot D(R/K, \lambda/K)$, which allowed them to induct on $R$. This idea was later used in [14], while the argument is completely independent.

In earlier works (see, e.g., [16, 18]), induction on scales arguments was used to prove square function estimates. However, it often seems inefficient to do an induction with respect to a single physical scale. Indeed, Guth, Wang, and Zhang formulated stronger square function estimates [12, Theorem 1.3], which work better with induction on scales. In their new scheme, two different scales $r \leq R$ are involved in the induction, and the “square function constant” $S(r, R)$ depends on both $r$ and $R$. The desired square function estimate corresponds to the inequality $S(1, R) \leq CR^\varepsilon$. The authors then induct on the quotient $\frac{R}{r}$ rather than on $R$ itself. This formulation is one of the key novel contributions of their work. To be more specific, there are three crucial inputs fed into the induction. The first one is a square function estimate for the parabola, which is used to control $S(r, K)$ for some large constant $K = K(\varepsilon)$ (for a $K^{-1}$-chunk of the cone). The second one is a Kakeya-type estimate, which controls $S(r, R)$ for $r \geq \sqrt{R}$. The last ingredient, Lorentz rescaling, can be used to bound $S(r, R)$ essentially by $S(r, r') \cdot \sup_{s \in [0,1]} S(sr', sR)$ with an intermediate scale $r'$. This allows one to induct on the ratio $R/r$. There are two cases for each step of the induction. If $r \leq K^{\frac{1}{2}}$, one goes through $r' = K$ and uses the square function estimate for parabola; if $K^{\frac{1}{2}} \leq r \leq R^{\frac{1}{2}}$, one goes through $r' = r^2$ and applies the Kakeya-type estimate to produce a gain.

Our strategy of proving Theorem 1.6 is inspired by all previous work [2, 12, 14]. We shall consider square function constants of the form $S(r, R, \lambda)$, with $\lambda$ being the ambient scale of the FIO. There are also three inputs in our proof. The first one is a small-scale estimate. In fact, if $r \leq R \leq \lambda^{\frac{1}{2} - \varepsilon}$, the problem looks almost identical to the translation-invariant case at such a small scale, and therefore, we can invoke Guth–Wang–Zhang. The second one is a variable coefficient version of the Kakeya-type estimate, which allows us to control $S(r, R, \lambda)$ for $r \geq R^{\frac{1}{2} + \varepsilon}$. In our case, many key tools available in Euclidean harmonic analysis are absent. We hence develop a new strategy in which we play with multiple different coordinate systems and take advantage of all of them. The authors believe that this strategy is novel and interesting in its own right. With these two estimates in hand, we can bound $S(r, R, \lambda)$ favorably if both $r, R$ are larger or smaller compared to $\lambda^{\frac{1}{2} - \varepsilon}$. Now we are left with the case when $r \ll \lambda^{\frac{1}{2} - \varepsilon} \ll R$. We then use a variable coefficient version of the Lorentz rescaling lemma to go through the intermediate scale $\lambda^{\frac{1}{2} - \varepsilon}$ and close the induction.

Our paper is organized as follows. In Section 2, we introduce basic reductions and state our main estimate Theorem 2.5. In Section 3, we present an induction on scale argument and reduce the proof of Theorem 2.5 to the proof of three lemmas. In Section 4, we collect some geometric facts tied to an important change of variable map that will become handy in our proof of the lemmas. In the last three sections, we prove our lemmas.

### 1.5 Notation

For any arc $\Theta \subset S^1$, we use $d(\Theta)$ to denote its aperture and $\eta_\Theta \in S^1$ to denote its center. For each aperture $s > 0$, we fix a collection $S_s$ of $10s^{-1}$ arcs to form a covering of $S^1$ with bounded overlaps.
For $R \geq 1$, $B_R(z)$ denotes a ball centered at $z$ with radius $R$. We abbreviate $B_R = B_R(z)$ when the center is not of particular interest. We use $\omega_{B_R}$ to denote a Schwartz weight function that is adapted to $B_R = B_R(z_0)$ so that $\omega_{B_R} \geq 1$ on $B_R$ and decays rapidly away from it.

For nonnegative quantities $A$ and $B$, we shall write $A \preceq B$, if there is an absolute constant $C > 0$ such that $A \leq CB$. We shall write $A \approx B$, if $\frac{1}{C}B \leq A \leq CB$ for some absolute constant $C > 1$. Suppose that $A(R)$ and $B(R)$ are now functions of taking nonnegative values. We write $A(R) \preceq \epsilon B(R)$, if there exists a constant $C_\epsilon$ depending on $\epsilon$, such that for all $R \geq 1$, we have $A(R) \leq C_\epsilon R^\epsilon B(R)$.

We shall use $\text{RapDec}(R)$ to denote a term that is rapidly decaying in $R > 1$, that is, for any $N \in \mathbb{N}$, there exists a constant $C_N$ such that $|\text{RapDec}(R)| \leq C_N R^{-N}$.

For a vector-valued function $F = (f_1)$, we denote its Jacobian matrix by $\partial_x F = \frac{\partial F}{\partial x} = (\frac{\partial f_1}{\partial x^j})_j$. In particular, if $F$ is real-valued, $\nabla_x F = \partial_x F$ is a row vector.

We shall say that a function is essentially supported in a set $S$ if the sup norm and $L^1$ norm of this function outside of $S$ are $\text{RapDec}(R)$ with respect to the underlying parameter $R$.

We use $\langle x, y \rangle$ to denote the inner product of vectors. When there is no confusion, we also write $x \cdot y := \langle x, y \rangle$.

## 2 SETUP AND MAIN ESTIMATES

Given $\lambda \gg 1$, we consider $F^\lambda f = \sum_\partial F^\lambda_\partial f$ as defined in (1.17). For simplicity, we write

$$F := F^\lambda f, \quad F_\partial := F^\lambda_\partial f.$$ 

Our goal is to establish the estimate (1.19), which says

$$\|F\|_{L^4(\mathbb{R}^3)} \leq C_\epsilon \lambda^\epsilon \left\| \left( \sum_\partial |F_\partial|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^3)} + C_N \lambda^{-N} \|F\|_{L^4(\mathbb{R}^3)}.$$ 

The reader should keep in mind that our $F$ and $F_\partial$ do not enjoy the Fourier support assumptions in [12]. In fact, they can be categorized by their microlocal support in the following sense.

**Definition 2.1.** We say a smooth function $F : \mathbb{B}_\lambda \to \mathbb{R}$ has microlocal support essentially contained in the $\lambda^{-1}$-neighborhood of a set $S \subset \mathbb{B}_\lambda \times \mathbb{R}^3$, if for any $0 < \epsilon < 1$, $1 \leq R \leq \lambda^{1-\epsilon}$, and any test function $\psi^\lambda(z, \xi) := \psi(z/\lambda, \xi)$, which equals 1 on the $R^{-1}$-neighborhood of $S$, we have

$$\frac{1}{(2\pi)^3} \int \int_{\mathbb{R}^3 \times \mathbb{R}^3} e^{i(w-z) \cdot \xi} (1 - \psi^\lambda(z, \xi)) F(z) \, dz \, d\xi = \text{RapDec}(R)\|F\|_1.$$ 

In addition, if $F$ has microlocal support essentially contained in the $\lambda^{-1}$-neighborhood of $S$, we shall say that the microlocal support of $F$ at $z$ is essentially contained in the $\lambda^{-1}$-neighborhood of the set

$$S_z := \{ \xi : (z, \xi) \in S \}.$$ 

Now, for $F = F^\lambda f$, the relevant integral is the following.

$$\frac{1}{(2\pi)^3} \int \int_{\mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^2} e^{i[(w-z) \cdot \xi + \psi^\lambda(z, \eta)]} (1 - \psi^\lambda(z, \xi)) a^\lambda(z, \eta) \hat{f}(\eta) \, dz \, d\xi \, d\eta.$$
A simple integration by parts argument in $z$ variables yields that $F = F^A f$ has microlocal support essentially contained in the $\lambda^{-1}$ neighborhood of the set

$$\{(z, \xi) : \xi = \nabla_z \phi^A(z, \eta), (z, \eta) \in \text{supp } a^A(\cdot, \cdot)\}. \quad (2.2)$$

Then the microlocal support of $F$ at $z$ is essentially contained in a $\lambda^{-1}$ neighborhood of the cone

$$C_z := \{\xi : \xi = \nabla_z \phi(z, \eta), \eta \in \text{supp } a(z, \cdot)\}. \quad (2.3)$$

Similarly, the microlocal support of $F_\theta$ at $z$ is essentially contained in the $\lambda^{-1}$-neighborhood of the set

$$C^0_z := \{\xi : \xi = \nabla_z \phi(z, \eta), \eta \in \text{supp } a^A_z(z, \cdot)\}. \quad (2.4)$$

For fixed $z$, the $R^{-1}$ neighborhood of $C^0_z$ is approximately a $1 \times R^{-1} \times R^{-1/2}$ plank tangent to the cone $C_z$. The long direction of this plank is given by $\nabla_z \phi^A$.

To prove (2.1), we shall prove a stronger square function estimate that requires further decomposing the square function with respect to various intermediate scales. In this section, we define a key change of variable process associated with each $\theta$, and then, use it to describe the precise square function estimate we aim to prove. This change of variable process should be viewed as a variable coefficient version of the Lorentz transformation. That is, given $\eta \in S^1$, after this process, the light cone in frequency space would be tangent to the horizontal coordinate plane along the direction $\eta^\theta$. See Figure 2 in Section 7 for a special case.

### 2.1 Change of variables

Let $\theta \subset S^1$ be a cap with center $\eta_\theta \in S^1$. We would like to change the variables from $z = (x, t)$ to new variables $(u, t)$, such that

$$\nabla_\eta \phi(z, \eta_\theta) = u.$$  

More precisely, suppose that $\nabla_\eta \phi(x_0, t_0, \eta_0) = u_0$ holds. By the condition (1.7),

$$\det \left( \frac{\partial (\nabla_\eta \phi, t, \eta)}{\partial (x, t, \eta)} \right)(x_0, t_0, \eta_0) \neq 0.$$  

Thus, by the implicit function theorem, there exists a local diffeomorphism $(u, t) \mapsto (x_\eta(u, t), t)$ between a neighborhood of $(u_0, t_0)$ and a neighborhood of $(x_0, t_0)$ such that

$$\nabla_\eta \phi((x_\eta(u, t), t), \eta) = u, \quad (2.5)$$

for each $\eta$ in a neighborhood of $\eta_0$. Since our result is local, the original domain of $(x, t, \eta)$ can be decomposed into finitely many such pieces, and it suffices to work with one of them. For later use, we can also assume that this neighborhood of $\eta_0$ is small enough so that $|\eta - \eta_0|^2 \leq c|\eta - \eta_0|$, where $c$ is a small but fixed constant.
We claim that one can further assume \((x_0, t_0) = (u_0, t_0) = 0\). To see this, write \(P f(x + x_0, t + t_0)\) as
\[
\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\phi(x + x_0, t + t_0, \eta)} a(x + x_0, t + t_0, \eta) \hat{f}(\eta) d\eta = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i(\phi(x + x_0, t + t_0, \eta) - \nabla_\eta \phi(x_0, t_0, \eta_0) \cdot \eta)} a(x + x_0, t + t_0, \eta) \hat{f}(\eta) d\eta,
\]
where \(\hat{f}\) has Fourier transform
\[
\hat{f}(\eta) = e^{i\nabla_\eta \phi(x_0, t_0, \eta_0) \cdot \eta} \hat{f}(\eta).
\]
Since translation does not change the \(L^p\)-norm of \(P f\), and \(\|\hat{f}\|_{L^4} \approx \|f\|_{L^4}\) by the multiplier theorem, we conclude that for our use, it is equivalent to replace the given phase function \(\phi\) by
\[
\phi(x, t, \eta) := \phi(x + x_0, t + t_0, \eta) - \nabla_\eta \phi(x_0, t_0, \eta_0) \cdot \eta,
\]
which satisfies all required conditions to be a phase function and \(\nabla_\eta \phi(0, \eta_0) = 0\), as desired.

As a summary, for every \(\eta\) in a neighborhood of \(\eta_0\), \(x_\eta(u, t)\) is smooth in \(\eta, u, t\), defined in the unit ball in \(\mathbb{R}^d\), of range contained in the unit ball in \(\mathbb{R}^d\), and \((2.5)\) holds.

In our proof below, we only consider \(\eta = \eta_0\), the center of dyadic caps \(\theta\), so for convenience, we denote \(x_\theta := x_{\eta_0}\). Then we define \(x_\lambda^\dagger(u, t) := \lambda x_\theta(u/\lambda, t/\lambda)\), which has domain \(B_\lambda\) in \(\mathbb{R}^d\) and range contained in \(B_\lambda\) in \(\mathbb{R}^d\). By rescaling \((2.5)\), we have
\[
\nabla_\eta \phi^\lambda((x_\lambda^\dagger(u, t), t), \eta_0) = u. \tag{2.6}
\]
By our assumption, \(\phi\) is homogeneous in \(\eta\) of degree 1. It follows that
\[
\phi^\lambda = \langle \nabla_\eta \phi^\lambda, \eta \rangle,
\]
which implies, together with \((2.5)\),
\[
\phi^\lambda(z, \eta_\theta) = \langle u, \eta_\theta \rangle.
\]
This means, if we change variables from \(z\) to \((u, t)\) and denote
\[
\tilde{\phi}^\lambda_\theta(u, t, \eta) := \phi^\lambda(x_\theta^\dagger(u, t), t, \eta), \tag{2.7}
\]
we have, for every \(\eta \parallel \eta_\theta\),
\[
\nabla_{u, t} \tilde{\phi}^\lambda_\theta(u, t, \eta) = (\eta, 0). \tag{2.8}
\]
On the other hand, direct computation on \((2.7)\) gives
\[
\nabla_{u, t} \tilde{\phi}^\lambda_\theta = \nabla_z \phi^\lambda \cdot \frac{\partial(x_\theta^\dagger, t)}{\partial(u, t)}. \tag{2.9}
\]
Recall that at each \( z \in B_\lambda \), the microlocal support of \( F_\theta \) at \( z \) is essentially contained in the \( \lambda^{-1} \)-neighborhood of the plank \( C_z^\phi \) as in (2.4) with its long direction given by \( \nabla_z \phi^\perp \). By comparing (2.8) and (2.9), it follows that under this change of variables, the resulting rectangular box is contained in the \( R^{-1} \)-neighborhood of the horizontal coordinate plane with normal vector \((0,0,1)\). Moreover, the flat direction of the cone is mapped to \((\eta,0)\). To be more precise, a simple integration by parts argument yields the following lemma.

**Lemma 2.2.** If \( 0 < \varepsilon \ll 1 \) and \( 1 \leq R \leq \lambda^{1-\varepsilon} \), let

\[
\tilde{F}_\theta(u, t) := F_\theta(x_\lambda^\lambda(u, t), t).
\]

then its Fourier transform, \( \hat{\tilde{F}}_\theta(\xi) \), is essentially supported in the \( R^{-1} \) neighborhood of the set

\[
\mathcal{N}_\theta := \{(\xi', \xi_3) : |\xi'| \approx 1, |\xi_3| \leq CR^{-1}, \text{dist}(\xi'/|\xi'|, \eta_\theta) \leq CR^{-1/2}\}.
\]

**Proof.** Recall \( F_\theta = F^\lambda \hat{f}_\theta \), then the Fourier transform of \( \tilde{F}_\theta \) reads

\[
\int e^{-i(u, t) \cdot \xi + i\lambda \phi(x_\lambda^\lambda(u, t), t, \eta)} a_\theta(x_\lambda^\lambda(u, t), t, \eta) \, du \, dt \hat{f}(\eta) \, d\eta.
\]

Define \( \mathcal{L}(u, t, \xi, \eta) := -\lambda(u, t) \cdot \xi + \lambda \phi(x_\theta(u, t), t, \eta) \). Define the operator \( D \) as follows:

\[
D := -i \nabla_{u, t} \mathcal{L}(u, t, \xi, \eta) |\nabla_{u, t} \mathcal{L}(u, t, \xi, \eta)|^{-2} \cdot \nabla_{u, t}.
\]

Therefore, for each integer \( N \), after rescaling in \( u, t \), the above integral can be rewritten as

\[
\lambda^3 \int (D)^N e^{-i\lambda(u, t) \cdot \xi + i\lambda \phi(x_\theta(u, t), t, \eta)} a_\theta(x_\theta(u, t), t, \eta) \, du \, dt \hat{f}(\eta) \, d\eta.
\]

Note that \( \nabla_{u, t} \mathcal{L}(u, t, \xi, \eta) = -\lambda \xi + \lambda \partial_{u, t} \phi(x_\theta(u, t), t, \eta) \). Thus, by the above discussion, we know that the set

\[
\{ \partial_{u, t} \phi(x_\theta(u, t), t, \eta) : \eta \in \text{supp } a_\theta(x_\theta(u, t), t, \cdot) \}
\]

is contained in \( \mathcal{N}_\theta \). By choosing \( C \) sufficiently large and noting that \( R \leq \lambda^{1-\varepsilon} \), we see that integration by parts will yield a rapidly decaying term \( \text{RapDec}(\lambda) \).

We remark that this change of variables \( (x, t) = (x_\theta(u, t), t) \) is invertible, namely for each cap \( \theta \) we have \( (u, t) = (u_\theta(x, t), t) \) with

\[
\frac{\partial(u, t)}{\partial(x, t)} = \left( \frac{\partial(x, t)}{\partial(u, t)} \right)^{-1}.
\]

Then, \( \frac{\partial(u, t)}{\partial(x, t)} \) sends the rectangle of direction \((\eta,0)\) on the horizontal coordinate plane to a rectangle tangent to the cone \( C_z \) whose direction coincides with the flat direction of the cone.
2.2 Quantitative conditions

In the proof of our main estimate, we run an induction on scales argument with respect to a class of phase functions satisfying the cinematic curvature condition (1.7) and (1.8). To facilitate the induction process, we need some quantitative upper bounds on the derivatives of the phase and amplitude functions. Let \( \varepsilon_0 > 0 \) be a small number. We first assume that the support of the amplitude function \( a(x, t, \eta) \) is contained in \( B_{\varepsilon_0}(0) \times B_{\varepsilon_0}(e_2) \), i.e,

\[
C_1 : \text{supp}_{x, t, \eta} a(x, t, \eta) \subset B_{\varepsilon_0}(0) \times B_{\varepsilon_0}(e_2).
\]

Let \( L \) be a sufficiently large number depending on the choice of \( \varepsilon_0 \) and \( C_{par} > 0 \) be a fixed constant.

\[
C_2 : \|\partial_\eta^\alpha \partial_{x, t}^\beta \phi(x, t, \eta)\|_{L^\infty(B_{\varepsilon_0}(0) \times B_{\varepsilon_0}(e_2))} \leq C_{par, \eta}(\alpha, \beta) \in \mathbb{N}^2 \times \mathbb{N}^3,
\]

\(|\alpha| \leq L, |\beta| \leq L\). The nonvanishing eigenvalues of

\[
\partial_\eta^2 \langle \partial_z \phi(z, \eta), G(z, \eta_0) \rangle|_{\eta=\eta_0}
\]

fall into \([1/2, 2]\), for \((z, \eta_0) \in \text{supp } a(\cdot, \cdot)\).

For the amplitude function, we assume

\[
C_3 : \|\partial_\eta^\alpha \partial_{x, t}^\beta a(x, t, \eta)\|_{L^\infty(B_{\varepsilon_0}(0) \times B_{\varepsilon_0}(e_2))} \leq C_{par, \eta}(\alpha, \beta) \in \mathbb{N}^2 \times \mathbb{N}^3,
\]

\(|\alpha| \leq L, |\beta| \leq L\).

2.3 Main estimates

In the following part, let \( s \) denote a dyadic number. Consider a cap \( \tau \subset S^1 \) with center \( \eta_\tau \) and \( d(\tau) = s, R^{-1/2} \leq s \leq 1 \). Let \( V_\tau \) be the \( Rs^2 \times Rs \)-rectangle centered at the origin of the plane, whose short edge is parallel to \( \eta_\tau \). Define

\[
U_\tau := \{z = (x_\tau^\lambda(u, t), t) : u \in V_\tau, |t| \leq R\}.
\]

We would like to say some words about the notation \( x_\tau \). By our change of variables in Section 2.1, one can see that \( x_\tau = \tilde{x}_\tau \), where \( \tilde{x}_\tau \) denotes the central \( R^{-1/2} \)-cap of \( \tau \). This is because they have the same center, namely, \( \eta_\tau = \eta_\tilde{\tau} \).

Under this notation, one can tile \((-R, R)^2\) by a collection of translated copies of \( V_\tau \) with finite overlaps. Let us call this collection \( \mathcal{V}_\tau \). Therefore,

\[
\bigcup_{V \in \mathcal{V}_\tau} V \supset (-R, R)^2.
\]

Correspondingly, define

\[
U := \{(x_\tau^\lambda(u, t), t) : u \in V, |t| \leq R, V \in \mathcal{V}_\tau\}.
\]

An equivalent yet more explicit definition of \( U \) can be found in Section 7. Let \( U_\tau \) be the collection of all such \( U \).
Lemma 2.3. The collection $U_\tau$ is finitely overlapping.

Proof. Since the rectangular boxes $\{V \times (-R,R)\}_{V \in \mathcal{V}_\tau}$ are finitely overlapping and each $U$ is obtained from some $V \times (-R,R)$ through a diffeomorphism $(u, t) \mapsto (x^1(u, t), t)$, provided that the size of support $a$ is sufficiently small. Thus, $U_\tau$ is also a covering of $(-R,R)^{2+1}$ with bounded overlaps. \qed

For a given $U \in U_\tau$, define
\begin{equation}
S_U F(z) := \begin{cases} 
\left( \sum_{\theta \in S_{R^{-1/2}} : \theta \subset \tau} |F_\theta|^2(z) \right)^{1/2}, & \text{if } z \in U, \\
0, & \text{otherwise.}
\end{cases}
\tag{2.11}
\end{equation}

We remark that, as pointed out in [12], the size of $U$ can be read off of the index $s$ of the sum and the diameter of the physical ball $R$. We shall prove the following analog of Theorem 1.5 in Guth–Wang–Zhang at the physical scale $R = \lambda$.

**Theorem 2.4.**

$$
\| F \|_{L^4(B_\lambda)}^4 \lesssim \varepsilon \sum_{\lambda^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \sum_{U \in U_\tau} |U|^{-1} \| S_U F \|_{L^2(w_{B_\lambda})}^4 + \text{RapDec}(\lambda) \| f \|_{L^2(B_\lambda)}^4,
$$

Here $s$ denotes a dyadic number ranging from $\lambda^{-1/2}$ to 1.

Let us first see how Theorem 2.4 implies Theorem 1.6. The proof is almost identical to that in [12].

**Theorem 2.4 implies Theorem 1.6.** Given $\tau$ and $U \in U_\tau$, by (2.11), we have
\begin{equation}
\| S_U F \|_{L^2(w_{B_\lambda})}^2 \approx \int_U \sum_{\theta \in S_{R^{-1/2}} : \theta \subset \tau} |F_\theta|^2.
\end{equation}

If we invoke the Cauchy–Schwarz inequality, it follows that
\begin{equation}
|U|^{-1} \| S_U F \|_{L^2(w_{B_\lambda})}^4 \leq \left( \sum_{\theta \in S_{R^{-1/2}} : \theta \subset \tau} |F_\theta|^2 \right)^2.
\end{equation}

Thus Theorem 2.4 implies that, modulo a RapDec$(\lambda) \| f \|_{L^2(B_\lambda)}^4$ error, we have
\begin{equation}
\| F \|_{L^4(B_\lambda)}^4 \leq C \varepsilon \sum_{\lambda^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \sum_{U \in U_\tau} |U|^{-1} \| S_U F \|_{L^2(w_{B_\lambda})}^4
\end{equation}
\begin{equation}
\leq C \varepsilon \sum_{\lambda^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \left( \sum_{\theta \in S_{R^{-1/2}} : \theta \subset \tau} |F_\theta|^2 \right)^{1/2} \left( \sum_{\theta \in S_{R^{-1/2}} : \theta \subset \tau} |F_\theta|^2 \right)^{1/2} \| f \|_{L^2(B_\lambda)}^4.
\end{equation}
\[
\leq C \lambda^\varepsilon \left\| \left( \sum_{\substack{\theta \in \mathcal{S}_{1/2} \setminus 1 \cap \partial \subset \tau}} |F_\theta|^2 \right)^\frac{1}{2} \right\|_{L^4(B_\lambda)}^4.
\]

Here, we used the fact that the sum over \( s \) is dyadic. \( \square \)

Theorem 2.4 is convenient in proving square function estimates. However, as discussed in the Introduction, we shall prove a slightly stronger version that works better with multiscale analysis and induction on scales.

Note that when \( U = B_r \), it is nothing but \( s = 1, R = r \) in the above definition, namely,

\[
S_{B_r} F = \begin{cases} 
\left( \sum_{\theta \in \mathcal{S}_{1/2}} |F_\theta|^2(z) \right)^\frac{1}{2} \quad , & \text{if } z \in B_r, \\
0, & \text{otherwise}.
\end{cases}
\]

Let \( S(r, R, \lambda) \) denote the smallest constant so that for any \( N \in \mathbb{N}, B_R \subset B_\lambda \), there exists a constant \( C_N > 0 \) such that

\[
\sum_{B_r \subset B_R} |B_r|^{-1} \| S_{B_r} F \|_{L^4(B_r)}^4 \leq S(r, R, \lambda) \sum_{R^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \sum_{U \in U'_{\tau}} |U|^{-1} \| S_{U} F \|_{L^4(w_B R)}^4 + C_N \lambda^{-N} \| f \|_{L^2}^4.
\]

(2.12)

It should be understood that \( S(r, R, \lambda) \) is uniform across all \( \phi \) and \( a \) satisfying the cinematic curvature condition and \( C_1, C_2, C_3 \). Then, to prove Theorem 2.4, it suffices to prove the following.

**Theorem 2.5.** For every \( \varepsilon > 0 \), there exists \( C_\varepsilon > 0 \) such that

\[
S(1, \lambda^{1-\varepsilon}, \lambda) \leq C_\varepsilon \lambda^\varepsilon.
\]

Theorem 2.5 implies Theorem 2.4. Take \( R = \lambda^{1-\varepsilon/100}, r = 1 \), then by Theorem 2.5, we have \( S(1, R, \lambda) \leq C_\varepsilon \lambda^\varepsilon/100 \). Note that

\[
\| F \|_{L^4(B_\lambda)}^4 \leq \left( \frac{\lambda^{7/30}}{100} \right) \sup_{B_R \subset B_\lambda} \| F \|_{L^4(B_R)}^4.
\]

Let \( B_R' \subset B_\lambda \) be a ball such that \( \| F \|_{L^4(B_R')}^4 \geq \frac{1}{2} \sup_{B_R \subset B_\lambda} \| F \|_{L^4(B_R)}^4 \), then we have

\[
\| F \|_{L^4(B_\lambda)}^4 \leq C_\varepsilon \lambda^{3\varepsilon/100} \| F \|_{L^4(B_R')}^4.
\]

Noticing that,

\[
\| F \|_{L^4(B_R')}^4 \leq \sum_{B_1 \subset B_R'} \| F \|_{L^4(B_1)}^4 \leq \sum_{B_1 \subset B_R'} |B_1|^{-1} \| S_{B_1} F \|_{L^2(B_1)}^4
\]

\[
\leq S(1, R, \lambda) \sum_{\lambda^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \sum_{U \in U'_{\tau}} |U|^{-1} \| S_{U'} F \|_{L^4(w_B \lambda)}^4 + C_{N+3} R^3 \lambda^{-N-3} \| f \|_{L^2(B_\lambda)}^4
\]

\[
\leq C_\varepsilon \lambda^\varepsilon \sum_{\lambda^{-1/2} \leq s \leq 1} \sum_{\tau \in \mathcal{S}_s} \sum_{U' \in U'} |U'|^{-1} \| S_{U'} F \|_{L^4(w_B \lambda)}^4 + C'_{N} \lambda^{-N} \| f \|_{L^2(B_\lambda)}^4.
\]

(2.13)
Here, we used the following facts: first, for \( U' \supset U \), where the dimension of \( U \) and \( U' \) are \( R \times Rs \times Rs^2 \) and \( \lambda \times \lambda s \times \lambda s^2 \), respectively, we have

\[
\|S_{U}F\|_{L^2(w_{B_R})}^2 \leq (\lambda/R)^{\frac{1}{2}}\|S_{U'}F\|_{L^2(w_{B_s})}^2 = \lambda^2/200\|S_{U'}F\|_{L^2(w_{B_{\lambda s^2}})}^2,
\]

Secondly, \(|U'|/|U| \approx (\lambda/R)^3 = \lambda^3/100\). Finally, the number of \( U \) that is contained in \( U' \) is bounded by the same number \( \lambda^3/100\). \( \square \)

### 3 OUTLINE OF THE PROOF

The rest of this paper is devoted to the proof of Theorem 2.5. We need three lemmas, the proofs of which will be given in Section 5-7.

**Lemma 3.1** (Small-scale estimate). For any \( \mu, \delta > 0 \), \( 1 \leq r \leq R \leq \lambda^{\frac{1}{2} - \mu} \), there exists a constant \( C_{\mu, \delta} \) such that

\[
S(r, R, \lambda) \leq C_{\mu, \delta} R^\delta (R/r)^\delta.
\]

Lemma 3.1 should be compared to the stability lemma in [2] and the small-scale decoupling inequality in [14]. The idea is that when the physical scales \((r, R)\) are small compared to \( \lambda^2 \), the desired square function estimate follows from its Euclidean counterpart rather directly. Indeed, in [12], the authors considered each \( K^{-1} \)-chunk of the cone for \( K = K(\varepsilon) \) and the associated square function constant \( S_K(r, R) \). They proved that \( S_K(r, R) \leq C_\delta (R/r)^\delta \). Summing over all these chunks gives the Euclidean version of Lemma 3.1.

**Lemma 3.2** (Lorentz rescaling). For any \( 1 \leq r_1 \leq r_2 \leq r_3 \leq \lambda^{1-\varepsilon} \), we have

\[
S(r_1, r_3, \lambda) \leq \log r_2 \cdot S(r_1, r_2, \lambda) \max_{r_2^{-1/2} \leq s \leq 1} S(s^2 r_2, s^2 r_3, s^2 \lambda).
\]

Lemma 3.2 is the variable coefficient analog of the Lorentz rescaling lemma in [12]. The proof follows the framework developed in [14], which is inspired by ideas from [2]. In [2] and [14], the authors used decoupling inequality at small scales and a rescaling lemma for decoupling constants to prove variable coefficient analogs. As discussed in the introduction, estimate (2.12) has two distinct scales \((r, R)\) involved. Thus, unlike previous work, Lemmas 3.1 and 3.2 are not sufficient to close the induction. We also need the following variable coefficient version of the Kakeya-type estimate in [12].

**Lemma 3.3** (Kakeya-type inequality). For any \( \nu, \varepsilon > 0 \), \( \lambda^{1/\nu} \leq r^2 \leq \lambda \), there exists a constant \( C_{\varepsilon, \nu} \) such that

\[
S(r_1, r^2, \lambda) \leq C_{\varepsilon, \nu} r^\nu, \text{ for all } r_1 \in [r^{1+\nu}, r^2].
\]

As in [12], we will induct on the ratio \( R/r \). However, in our case, an additional parameter \( \lambda \) is present. As we shall see in the proof, the relation between \( r, R \) and an intermediate scale \( \lambda^{1/2-\mu} \)
will be the key to our proof. Indeed, we shall use Lemma 3.3 to handle the case when both \( r \) and \( R \) are large compared to \( \lambda^{1/2-\mu} \). On the other hand, the case when \( r \) and \( R \) are both small compared to \( \lambda^{1/2-\mu} \) will be covered by Lemma 3.1. Lastly, when \( r \) is small and \( R \) is large, we shall invoke Lemma 3.2 to go through an intermediate scale to close the induction.

**Proof of Theorem 2.5 via three lemmas.** It suffices to prove that

\[
S(r, R, \lambda) \leq C\varepsilon R^{\varepsilon/2}(R/r)^{\varepsilon/2},
\]

for all \( 1 \leq r \leq R \leq \lambda^{1-\varepsilon} \). Note that \( S(r, R, \lambda) \leq 1 \) trivially holds for all \( (r, R, \lambda) \) with \( \frac{R}{r} = 1 \). This will serve as our base case.

**Induction hypothesis.** For \( 1 \leq r \leq R \leq \lambda^{1-\varepsilon} \), we assume that

\[
S(r', R', \lambda') \leq C\varepsilon (R')^{\varepsilon/2}(R'/r')^{\varepsilon/2}
\]

for all \( (r', R', \lambda') \) satisfying \( \lambda' \leq \lambda \) and \( \frac{R'}{r'} \leq \frac{R}{r} \).

We choose \( \mu = \varepsilon^2/4 \), and consider intermediate scales \( R_{\mu} = \lambda^{1-\mu} \) and \( R_{2\mu} = \lambda^{1-2\mu} \). Lemma 3.1 indicates that (3.1) holds if \( R \leq R_{\mu} \). Now for \( R_{\mu} \leq R \leq R_{2\mu}^2 = \lambda^{1-\varepsilon} \), if we also have that \( r \geq R_{2\mu} \), then \( \sqrt{R^{1+\varepsilon}} \leq r \leq R \). Applying Lemma 3.3 with \( \nu = \varepsilon \) gives

\[
S(r, R, \lambda) \leq C\varepsilon R^{\varepsilon/2},
\]

and thus, (3.1) follows from Lemma 3.3 directly in this case.

Finally, we assume that \( r < R_{2\mu} \), then \( \frac{R}{R_{2\mu}} \leq \frac{R}{2r} \) for sufficiently large \( \lambda \). If we invoke Lemma 3.2 with \( r_1 = r, r_2 = R_{\mu}, r_3 = R \), we have

\[
S(r, R, \lambda) \leq \log \lambda \cdot S(r, R_{\mu}, \lambda) \max_{R_{\mu}^{-1/2} \leq s \leq 1} S(s^2 R_{\mu}, s^2 R, s^2 \lambda).
\]

Using Lemma 3.1 and our induction hypothesis, we have

\[
S(r, R, \lambda) \leq \log \lambda \cdot C_{\mu, \delta} R_{\mu}^\delta (R_{\mu}/r)^{\delta-\varepsilon/2} C\varepsilon R^{\varepsilon/2}(R/R_{\mu})^{\varepsilon/2}.
\]

If we choose \( \delta(\varepsilon) \) to be small enough so that \( \log \lambda \cdot C_{\mu, \delta} R_{\mu}^\delta (R_{\mu}/r)^{\delta-\varepsilon/2} \leq 1 \), then (3.1) holds. The induction is closed.

\[ \square \]

4 | GEOMETRY IN CHANGE OF VARIABLES

In Section 2.1, we introduced the change of variables between \((x, t)\) and \((u, t)\). In fact, there is a coordinate system \((x_\theta(u, t), t)\) for each \( \eta_\theta \in S^1 \). These changes of variables are crucial for our analysis, especially in the proof of Lemma 3.3.

Each coordinate system has its own advantages. Under the original coordinates \((x, t)\), the frequency lies near a cone, so we can play tricks with the cone geometry. This does not work under the coordinates \((u, t)\), since these coordinates depend on the choice of \( \eta_\theta \). On the other hand,
under coordinates $(u, t)$ associated with $\theta \subset \tau$, the corresponding set

$$U_\tau := \{(u, t) : (x^{1/2}_\tau(u, t), t) \in U_\tau\} = V_\tau \times (-R, R)$$

is a $1 \times R^{1/2} \times R$-plank, with the property that $U_\tau - U_\tau$ is essentially itself. This, in general, fails on $U_\tau$, which may be curved.

We would like to take advantage of all the different coordinate systems. In this section, we discuss the connection between their geometries carefully. Recall from (2.9) that at each point $(x, t)$, the change from microlocal frequency to the Fourier frequency is given by the matrix

$$\begin{vmatrix} \frac{\partial (u^{1/2}_\theta(x, t), t)}{\partial (x, t)} \end{vmatrix}.$$ 

See Figure 1.

For any dyadic number $R^{-1/2} \leq \sigma \leq 1$, denote $\Theta_{\sigma, \theta}$ the rectangular box of dimensions $\sigma^2 \times \sigma R^{-1/2} \times R^{-1}$, centered at the original, whose $\sigma^2$-edge has direction $(\theta, 0) \in S^2$, $\sigma R^{-1/2}$-edge has direction $(\theta^\perp, 0)$, and $R^{-1}$-edge has direction $(0,0,1)$. By Lemma 2.2 and basic properties of Fourier transforms of products, these planks “on the ground” are where the frequency of $|F_\theta|^2$ lies. Let $\tau \subset S^1$ be a $\sigma^{-1} R^{-1/2}$-cap. We shall define $\Theta_{\sigma, \tau}$ similarly as a box of the same dimensions with orientation associated with $\tau$. The size of $\tau$ is chosen so that $\Theta_{\sigma, \theta} \subset 4 \Theta_{\sigma, \tau}$ for any $\theta \subset \tau$. Therefore, it makes no difference to replace $\Theta_{\sigma, \theta}$ by $\Theta_{\sigma, \tau}$.

Before stating our geometric lemmas on change of variables, we would like to remind the reader of the following notation: $x_\tau = x_{\theta_\tau}$, where $\theta_\tau$ denotes the central $R^{-1/2}$-cap of $\tau$. This is because they have the same center, namely, $\eta_\tau = \eta_{\theta_\tau}$. Similarly, $u_\tau = u_{\theta_\tau}$.

Our first geometric observation is, when $\sigma, \tau, (x, t), \omega \in \Theta_{\sigma, \tau}$ are fixed, the corresponding frequency under the coordinates $(x, t)$ is also essentially fixed when $\theta$ runs over $\tau$.

**Lemma 4.1.** Fix $(x, t), \sigma, \tau$ and $\omega \in \Theta_{\sigma, \tau}$. Then for any $\theta_1, \theta_2 \subset \tau$,

$$\left| \frac{\partial (u^1_{\theta_1}(x, t), t)}{\partial (x, t)} \cdot \omega - \frac{\partial (u^1_{\theta_2}(x, t), t)}{\partial (x, t)} \cdot \omega \right| \lesssim R^{-1}.$$
Proof. Recall from Section 2.1 that for any $\eta \parallel \eta_0$, we have

$$\nabla_\eta \phi(x, t, \eta) = u_\eta(x, t), \phi(x, t, \eta) = \langle u_\eta(x, t), \eta \rangle.$$ 

Take $\nabla_\eta$ on both sides of the second equation. Together with the first equation, it follows that

$$u_\eta = \nabla_\eta \phi = u_\eta + \eta \cdot \partial_\eta u_\eta,$$

which implies for all $(x, t)$,

$$\eta \cdot \partial_\eta u_\eta = 0. \quad (4.2)$$

As $\Theta_{\sigma, \tau}$ is $R^{-1}$-thick, without loss of generality, we may assume $\omega \in \{\xi : \xi_3 = 0\}$. Recall that $\eta_{\tau} \in S^1$ denotes the center of the arc $\tau$. We write $\omega$ as the sum of two vectors whose directions are $\eta_{\tau}, \eta_{\perp},$ say

$$\omega = \omega_l + \omega_s,$$

with $\omega_l \parallel \eta_{\tau}, |\omega_l| \leq \sigma^2,$ and $\omega_s \perp \eta_{\tau}, |\omega_s| \leq \sigma R^{-1/2}$.

Since $\omega_l \cdot \partial_\eta u_{\tau} = \eta_{\tau} \cdot \partial_\eta u_{\tau} = 0$ for all $(x, t)$, it follows that

$$\omega_l \cdot \partial_{x_1} \partial_\eta u_{\tau} = \omega_l \cdot \partial_{x_2} \partial_\eta u_{\tau} = \omega_l \cdot \partial_t \partial_\eta u_{\tau} = 0,$$

and therefore,

$$\left| \langle \omega_l, 0 \rangle \cdot \left( \frac{\partial(u_{\partial_1}^\Lambda(x, t), t)}{\partial(x, t)} - \frac{\partial(u_{\partial_2}^\Lambda(x, t), t)}{\partial(x, t)} \right) \right| \leq [d(\tau)]^2 \cdot |\omega_l| \leq R^{-1}.$$

For $\omega_s$, one simply has

$$\left| \langle \omega_s, 0 \rangle \cdot \left( \frac{\partial(u_{\partial_1}^\Lambda(x, t), t)}{\partial(x, t)} - \frac{\partial(u_{\partial_2}^\Lambda(x, t), t)}{\partial(x, t)} \right) \right| \leq d(\tau) \cdot |\omega_s| \leq R^{-1}. \square$$

The same holds when the direction of change of variables is reversed.

Lemma 4.2. Fix $(u, t), \sigma$ and $\omega \in \Theta_{\sigma, \tau} \cdot \frac{\partial(u_\theta^\Lambda(x, t), t)}{\partial(x, t)}$, for some $\theta \subset \tau$. Then for any $\theta_1, \theta_2 \subset \tau$,

$$\left| \omega \cdot \frac{\partial(x_\theta^\Lambda(u, t), t)}{\partial(u, t)} - \omega \cdot \frac{\partial(x_\theta^\Lambda(u, t), t)}{\partial(u, t)} \right| \leq R^{-1}.$$

Proof. The proof still relies on (4.2), as in the previous lemma. By the previous lemma, we may assume that $\eta_{\tau}$ is the center of the arc $\tau$. Also, as above, we may decompose a vector in $\Theta_{\sigma, \tau}$ into three parts, and only the direction parallel to $(\eta_{\tau}, 0)$ deserves our attention. Therefore, we may
assume

\[ \omega = (\tilde{\omega}_1, 0) \cdot \frac{\partial (u^\lambda_i(x, t), t)}{\partial(x, t)}, \]

with \( \tilde{\omega}_i \parallel \eta_{\tau} \), \( |\tilde{\omega}_i| \leq \sigma^2 \).

By the definition of our change of variables,

\[ \frac{\partial (u^\lambda_i, t)}{\partial(x, t)} (x^\lambda_i(u, t), t) \cdot \frac{\partial (x^\lambda_i, t)}{\partial(u, t)} (u, t) = \text{Id}. \tag{4.3} \]

Take \( \partial_{\eta_k} \), \( k = 1, 2 \) on both sides of (4.3). We obtain

\[ \frac{\partial (u^\lambda_i, t)}{\partial(x, t)} \cdot \left( \frac{\partial (x^\lambda_i, t)}{\partial(u, t)} \right) + A_k \cdot \frac{\partial (x^\lambda_i, t)}{\partial(u, t)} = 0, \tag{4.4} \]

where the matrix \( A_k \) is given by

\[ A_k = \frac{\partial (\partial_{x_1} u^\lambda_i, t)}{\partial(x, t)} \cdot \partial_{\eta_k} (x^\lambda_i)_1 + \frac{\partial (\partial_{x_2} u^\lambda_i, t)}{\partial(x, t)} \cdot \partial_{\eta_k} (x^\lambda_i)_2, \]

with \((x^\lambda_i)_1, (x^\lambda_i)_2) = x^\lambda_i\).

Multiply (4.4) by \((\tilde{\omega}_i, 0)\) from the left. Then the second term becomes

\[ (\tilde{\omega}_i, 0) \cdot \left( \frac{\partial (\partial_{x_1} u^\lambda_i, t)}{\partial(x, t)} \cdot \partial_{\eta_k} (x^\lambda_i)_1 + \frac{\partial (\partial_{x_2} u^\lambda_i, t)}{\partial(x, t)} \cdot \partial_{\eta_k} (x^\lambda_i)_2 \right) \cdot \frac{\partial (x^\lambda_i, t)}{\partial(u, t)}. \tag{4.5} \]

By taking \( \partial_{x_i}, \partial_{x_j}, \partial_{x_i} \partial_t \) on both sides of (4.2) and noting the assumption \( \tilde{\omega}_i \parallel \eta_{\tau} \), one can see that (4.5), in fact, vanishes.

We just proved that, after multiplying (4.4) by \((\tilde{\omega}_i, 0)\) from the left, the second term vanishes. Then the remaining first term gives

\[ 0 = (\tilde{\omega}_i, 0) \cdot \frac{\partial (u^\lambda_i, t)}{\partial(x, t)} \cdot \left( \frac{\partial (x^\lambda_i, t)}{\partial(u, t)} \right) = \omega \cdot \left( \frac{\partial (x^\lambda_i, t)}{\partial(u, t)} \right), \]

for every \( \theta_1, \theta_2 \subset \tau \), as desired. \[ \Box \]

We also claim that, when \( (x, t) \) is fixed, \((u^\lambda_i(x, t), t)\) lies in the same \( U \in \tilde{U}_{\tau} \) for all \( \theta \subset \tau \). As \( t \) is fixed, it suffices to consider \( u^\lambda_{i_1} - u^\lambda_{i_2} \). Recall from Section 2.3 that \( V_{\tau} \) in the plane is the \( \sigma^{-2} \times \sigma^{-1} R^{1/2} \)-rectangle centered at the origin with the short edge parallel to \( \eta_{\tau} \), and both \( U_{\tau} \) defined
Lemma 4.3. Fix \((x, t) \in B_R\). Then for any \(\theta_1, \theta_2 \subset \tau\),
\[
    u^4_{\partial_1}(x, t) - u^4_{\partial_2}(x, t) \in 4V_\tau.
\]

Proof. Without loss of generality, we assume that \(B_R\) is centered at the origin and \(\phi(0, 0, \eta) = 0\) for any \(\eta\), so \(D^2_\eta \phi\) always vanishes when \((x, t) = (0, 0)\). Also by (2.6), with \(\theta = \eta/|\eta|\), we have
\[
    \nabla_\eta \phi^\lambda(x, t, \eta) = u^\lambda_\theta(x, t).
\]
Therefore, all derivatives of \(u^\lambda_\theta\) in \(\eta\) vanish when \((x, t) = (0, 0)\). In particular, this implies
\[
    |D^2_\eta u^\lambda_\theta(x, t)| = |D^2_\eta u^\lambda_\theta(x, t) - D^2_\eta u^\lambda_\theta(0, 0)| \leq R \|D^2_\eta \nabla_{x,t} u^\lambda_\theta\|_{L^\infty} \lesssim R.
\]

Then, along the long edge direction of \(V_\tau\),
\[
    |\langle u^\lambda_{\partial_1} - u^\lambda_{\partial_2}, \tau^\perp \rangle| \leq |u^\lambda_{\partial_1} - u^\lambda_{\partial_2}| \lesssim \|\nabla_\eta u^\lambda_\tau\|_{L^\infty} \cdot d(\tau) \lesssim \sigma^{-1} R^{1/2};
\]
for the short edge direction of \(V_\tau\), which is \(\eta_\tau\), we invoke (4.2) and obtain
\[
    |\langle u^\lambda_{\partial_1} - u^\lambda_{\partial_2}, \tau \rangle| \lesssim \|\nabla_\eta \nabla_\eta u^\lambda_\tau\|_{L^\infty} \cdot [d(\tau)]^2 \lesssim \sigma^{-2},
\]
as desired. \(\square\)

5 | SMALL SCALES

In this section, we prove Lemma 3.1 by exploiting a phenomenon known as a “stability lemma” in the literature. The authors of [2] used it to deduce small-scale decoupling estimates in the proof of Conjecture 1.3 for odd \(d\). They proved a precise stability lemma using Taylor expansions, which allowed them to approximate \(F^\lambda\) by an extension operator at tiny scales. Later, inspired by [2], authors in [14] developed an independent argument to prove a microlocal decoupling inequality. The main idea is that at tiny scales \(R \ll \lambda^{1/2}\), the collection of functions \(\{w_B F_\theta : d(\theta) = R^{-1/2}\}\) essentially satisfies the Fourier support assumption in [12], and thus, enjoys the square function estimates proved there. Indeed, it is not hard to see that at such a tiny scale, the varying conic surface in phase space remains within an \(R^{-1}\)-neighborhood of a fixed cone, and thus, the associated tubes \(U\) that are curved at a larger scale look straight at scale \(R\). In this section, we follow the framework developed in [14] to check the Fourier support of restricted \(F_\theta\), which seems more illustrative for the readers.

Proof of Lemma 3.1. Denote the \(R^{-1}\)-neighborhood of \(C_z\) in (2.3) and \(C_z^0\) in (2.4) by
\[
    C_{z,R} := \{\xi : \text{dist}(\xi, \partial_z \phi^4(z, \eta)) \leq R^{-1}, \eta \in \text{supp } a^4(z, \cdot)\},
\]

in (2.10) and \(U_\tau\) defined in (4.1) are associated with the same \(V_\tau\). Then our claim is reduced to the following lemma.
and
\[
C^\vartheta_{z,R} := \{ \xi : \text{dist}(\xi, \partial_z \phi^\lambda(z, \eta)) \leq R^{-1}, \eta \in \text{supp} \alpha^\lambda_\vartheta(z, \cdot) \}. \tag{5.2}
\]

Fix $\mu > 0$. We shall see that when $R$ is smaller than $\lambda^{1/2-\mu}$, Lemma 3.1 can be reduced to the Euclidean case, then the result of Guth–Wang–Zhang [12] applies directly. Fix $\bar{z} \in B_\lambda$ and consider $z \in B_{\lambda^{1/2-R}}(\bar{z})$, $R < \lambda^{1/2-\mu}$.

**Lemma 5.1.** There exists an absolute constant $C > 0$ such that: for any $z \in B_{\lambda^{1/2-R}}(\bar{z})$, the set $C_{z,R}^\vartheta C_{\bar{z},R}^\vartheta$ are contained in a $CR^{-1}$-neighborhood of $C_{z,R}^\vartheta C_{\bar{z},R}^\vartheta$ respectively.

**Proof.** Recall $|\det(\partial_{zz}^2 \phi^\lambda)| \lesssim 1$, for $C_{z,R}$, it follows from
\[
|\nabla_z \phi^\lambda(z, \eta) - \nabla_z \phi^\lambda(\bar{z}, \eta)| \lesssim |z - \bar{z}| / \lambda \lesssim \lambda^{\mu/2} R / \lambda \lesssim \lambda^{-1/2-\mu+\mu/2} \lesssim R^{-1}.
\]
Similarly, we obtain the desired result for $C_{\bar{z},R}^\vartheta$.

For $1 \leq R \leq \lambda^{1/2-\mu}$, let us consider the Fourier support of the restriction of $F, F_\vartheta$ to the ball $B_R(\bar{z})$ using an appropriate weight function.

**Lemma 5.2.** Let $1 \leq R \leq \lambda^{1/2-\mu}$. The Fourier support of $w_{B_{\lambda^{1/3-R}}(\bar{z})} F$ and $w_{B_{\lambda^{1/3-R}}(\bar{z})} F_\vartheta$ are essentially contained in the $CR^{-1}$-neighborhood of $C_{z,R}^\vartheta C_{\bar{z},R}^\vartheta$, respectively.

**Proof.** Recall $F = F^\lambda f$. The Fourier transform of $w_{B_{\lambda^{1/3-R}}(\bar{z})} F$ is
\[
\iint e^{-iz \cdot \xi + i\phi^\lambda(z, \eta)} \alpha^\lambda(z, \eta) w_{B_{\lambda^{1/3-R}}(\bar{z})}(\xi) \hat{f}(\eta) \, dz \, d\eta.
\]
If $z \notin B_{\lambda^{1/2-R}}(\bar{z})$, then
\[
|w_{B_{\lambda^{1/3-R}}(\bar{z})}(\xi)| = \text{RapDec}(\lambda).
\]
Let $\Psi_{B_{\lambda^{1/2-R}}(\bar{z})}$ be a smooth function with compact support in $B_{\lambda^{1/2-R}}(\bar{z})$. We have, modulo a RapDec($\lambda$)$\|f\|_{L_1}$ error,
\[
\iint e^{-iz \cdot \xi + i\phi^\lambda(z, \eta)} \alpha^\lambda(z, \eta) w_{B_{\lambda^{1/3-R}}(\bar{z})}(z) \hat{f}(\eta) \, dz \, d\eta
\]
\[
= \iint e^{-iz \cdot \xi + i\phi^\lambda(z, \eta)} \alpha^\lambda(z, \eta) w_{B_{\lambda^{1/3-R}}(\bar{z})}(z) \Psi_{B_{\lambda^{1/2-R}}(\bar{z})}(z) \hat{f}(\eta) \, dz \, d\eta. \tag{5.3}
\]
By integrating by parts with respect to $z$, one can check that if $\xi \notin \bigcup_{z \in B_{\lambda^{1/2-R}}(\bar{z})} C_{z,R}$, the integral is rapidly decaying in terms of $\lambda$. We then obtain the desired result by Lemma 5.1. The corresponding statement regarding $w_{B_{\lambda^{1/3-R}}(\bar{z})} F_\vartheta$ can be obtained similarly. □
To sum up, we have reduced estimating $S(r, R, \lambda)$ for $R \leq \lambda^{1/2-\mu}$ to proving corresponding square function estimate associated with the (fixed) cone $C_{\lambda}$. To be more precise, if $R \leq \lambda^{1/2-\mu}$, by Lemma 5.2, it follows that

$$w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f = \chi_{C_{\lambda,R}}(D)(w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f) + \text{RapDec}(\lambda) \|f\|_{L^1},$$

and

$$w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f = \sum_\theta \chi_{C_{\lambda,R}^\theta}(D)(w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f_\theta) + \text{RapDec}(\lambda) \|f\|_{L^1}$$

$$= \sum_\theta \chi_{C_{\lambda,R}^\theta}(D)(w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f) + \text{RapDec}(\lambda) \|f\|_{L^1},$$

(5.4)

where $\chi_{C_{\lambda,R}}(D)$, $\chi_{C_{\lambda,R}^\theta}(D)$ are two Fourier multipliers, and $\chi_E$ denotes the characteristic function of $E$. Therefore, modulo a term rapidly decaying in $\lambda$, we have

$$\chi_{C_{\lambda,R}}(D)(w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f) = \sum_\theta \chi_{C_{\lambda,R}^\theta}(D)(w_{B_{\lambda^{-1/2}R}}(\xi) T^\lambda f).$$

Now, we may apply a version of [12, Proposition 3.4]† with respect to the cone $C_{\lambda}$ (see Remark 1.7) to finish the proof of Lemma 3.1. □

6 A MICROLOCAL KAKEYA-TYPE ESTIMATE

In this section, we prove Lemma 3.3. We first reduce it to the following Kakeya-type lemma, which is a microlocal analog of [12, Lemma 1.4]. The key difference from that in [12] is that we exploit different coordinate systems and take advantage of all of them. We have discussed the necessity of changing variables in Section 4.

Lemma 6.1. Let $G = \sum_{d(\theta) = R^{-1/2}} |F_{\theta}|^2$. Then, for any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that

$$\int_{B_R} |G|^2 \leq C_\varepsilon R^\varepsilon \sum_{R^{-1/2} \leq \xi \leq 1} \sum_{\tau \in \mathbb{S}_1} \sum_{U \in U_{\tau}} |U|^{-1} \|S_{U} F\|_{L^2(w_{B_R})}^4.$$

Lemma 6.1 implies Lemma 3.3. Let $\psi_{B_R} \in C_0^\infty$ be a nonnegative function that equals 1 on $B_R$ and 0 outside $2B_R$. By the definition of $S_{B_{\tau_1} F}$,

$$\|S_{B_{\tau_1}} F\|_{L^2(B_{\tau_1})}^2 \leq \int \psi_{B_{\tau_1}} \cdot \sum_{\tau \in \mathbb{S}_{1/2}} |F_{\tau}|^2.$$

† Indeed, Proposition 3.4 in [12] is an estimate for $S_\varepsilon(r, R)$, while what we need is the same estimate but for $S(r, R)$. Nonetheless, this estimate is a simple corollary of Proposition 3.4, and its proof is nearly identical to the proof of Theorem 1.3 there.
Since $r_1 \leq r^2$, every $r_1^{-1/2}$-cap $\tau$ can be decomposed into a union of $r^{-1}$-caps. This means that one can write

$$F_{\tau} = \sum_{\partial \subset \tau, \partial \in S_{r^{-1}}} F_{\theta}$$

and

$$\int \psi_{B_{r_1}} \cdot \sum_{\tau \in S_{r_1^{-1/2}}} |F_{\tau}|^2 = \int \psi_{B_{r_1}} \cdot \sum_{\tau \in S_{r_1^{-1/2}}} \left| \sum_{\partial \subset \tau, \partial \in S_{r^{-1}}} F_{\theta} \right|^2$$

$$= \sum_{\tau \in S_{r_1^{-1/2}}} \sum_{\partial_1 \subset \tau, \partial_2 \subset \tau} \int \psi_{B_{r_1}} \cdot F_{\theta_1} \tilde{F}_{\theta_2}.$$

We shall show that we only need to consider adjacent caps $\theta_1, \theta_2$.

Recall from (1.17) that

$$F_{\theta}(z) = R_{\theta}^{\lambda} f(z) = \int_{\mathbb{R}^2} e^{i\phi(\lambda, \eta)} a_{\theta}^{\lambda}(z, \eta) \hat{f}(\eta) d\eta,$$

so the integral

$$\int \psi_{B_{r_1}} \cdot F_{\theta_1} F_{\theta_2}$$

$$= \iint \left( \int e^{i(\phi^{\lambda}(z, \eta_1) - \phi^{\lambda}(z, \eta_2))} a_{\theta_1}^{\lambda}(z, \eta_1) a_{\theta_2}^{\lambda}(z, \eta_2) \psi_{B_{r_1}}(z) d\eta_1 d\eta_2 \right) f(\eta_1) \overline{f(\eta_2)} d\eta_1 d\eta_2.$$

To show that it has rapid decay when $\theta_1, \theta_2$ are $r^{-1}$-separated, it suffices to show the kernel

$$\int e^{i(\phi^{\lambda}(z, \eta_1) - \phi^{\lambda}(z, \eta_2))} a_{\theta_1}^{\lambda}(z, \eta_1) a_{\theta_2}^{\lambda}(z, \eta_2) \psi_{B_{r_1}}(z) dz$$

has rapid decay. This is Hörmander's argument. More precisely, the phase function in this kernel satisfies

$$|\nabla_x (\phi^{\lambda}(x, t, \eta_1) - \phi^{\lambda}(x, t, \eta_2))| = |\partial_{\eta_1}^{\lambda} \phi^{\lambda} \cdot (\eta_1 - \eta_2)| + O(|\eta_1 - \eta_2|^2).$$

By condition (1.7), the first term is $\approx |\eta_1 - \eta_2|$, where the implicit constant only depends on $\phi$, and thus, is independent of $\lambda$. We have discussed in Section 2.1 that one can assume the $\eta$-support of the amplitude is small enough so that $|\eta_1 - \eta_2|^2 \ll |\eta_1 - \eta_2|$. Therefore, the $O(|\eta_1 - \eta_2|^2)$ term can be ignored and

$$|\nabla_x (\phi^{\lambda}(x, t, \eta_1) - \phi^{\lambda}(x, t, \eta_2))| \geq |\eta_1 - \eta_2|.$$
which is \( \preceq r^{-1} \) if \( \theta_1, \theta_2 \) are \( r^{-1} \)-separated. Then, by integration by parts,

\[
\left| \int \psi_{B_{r_1}} \cdot F_{\partial_1} \tilde{F}_{\partial_2} \right| = \text{RapDec}(r_1/r) \| f \|_{L^2}^2 = \text{RapDec}(\lambda) \| f \|_{L^2}^2,
\]

where the last equality follows from the assumption \( r_1 \geq r^{1+\upsilon} \geq \lambda \frac{1}{200} \) in Lemma 3.3.

With this local \( L^2 \)-orthogonality, it follows that

\[
\int \psi_{B_{r_1}} \cdot \sum_{\partial \in \mathbb{S}_{r-1}, \theta \in \mathbb{S}_{r-1}} F_{\partial} \left| \sum_{\theta \subset \tau, \theta \in \mathbb{S}_{r-1}} |F_{\theta}|^2 \right| \lesssim \int \psi_{B_{r_1}} \cdot \sum_{\partial \in \mathbb{S}_{r-1}} |F_{\partial}|^2 + \text{RapDec}(\lambda) \| f \|_{L^2}^2.
\]

Then, we take sum in \( \tau \in \mathbb{S}_{r_1/2}^{-1} \) to obtain that, up to a \( \text{RapDec}(\lambda) \) error,

\[
\| S_{B_{r_1}} F \|_{L^2(B_{r_1})}^2 \lesssim \sum_{\tau \in \mathbb{S}_{r_1/2}^{-1}} \int \psi_{B_{r_1}} \cdot \sum_{\partial \in \mathbb{S}_{r-1}, \theta \in \mathbb{S}_{r-1}} |F_{\partial}|^2 = \int \psi_{B_{r_1}} \cdot \sum_{\partial \in \mathbb{S}_{r-1}} |F_{\partial}|^2 = \int \psi_{B_{r_1}} \cdot G.
\]

Finally, by Cauchy–Schwarz, modulo a \( \text{RapDec}(\lambda) \) term,

\[
\sum_{B_{r_1} \subseteq B_r} |B_{r_1}|^{-1} \| S_{B_{r_1}} F \|_{L^2(B_{r_1})}^4 \lesssim \sum_{B_{r_1} \subseteq B_r} \int_{B_{r_1}} |G|^2 = \int_{B_r} |G|^2.
\]

Hence, Lemma 3.3 follows from Lemma 6.1.

Now it remains to prove Lemma 6.1.

Recall that in Section 2.3, for each \( \theta \), there is an associated change of variables \( (x, t) = (x_\theta(u, t), t) \), and we have defined \( \tilde{F}_\theta(u, t) = F_\theta(x_\theta^*(u, t), t) \). Denote its inverse as \( (u, t) = (u_\theta(x, t), t) \) and \( u_\theta^*(\cdot) := \lambda u_\theta(\cdot / \lambda) \). By Fourier inversion,

\[
|F_\theta(x, t)|^2 = |\tilde{F}_\theta(u_\theta^*(x, t), t)|^2 = \int e^{2\pi i (u_\theta^*(x, t)) \cdot \tilde{\omega}} (|\tilde{F}_\theta|)^2 \left( (\tilde{\omega})^\wedge \right) d\tilde{\omega}.
\]

Then

\[
\int_{B_r} |G|^2 = \int_{B_r} \left( \sum_{\theta} \int e^{2\pi i (u_\theta^*(x, t)) \cdot \tilde{\omega}} (|F_\theta|)^2 \left( (\tilde{\omega})^\wedge \right) d\tilde{\omega} \right)^2 dx dt.
\] (6.1)

By Lemma 2.2, the Fourier support of \( F_\theta \) lies in \( \mathcal{N}_\theta \). By basic properties of Fourier transforms of products, one can see that \( \tilde{\omega} \) in (6.1) lies in

\[
\mathcal{N}_\theta + (-\mathcal{N}_\theta),
\]

which is contained in a \( 1 \times R^{-1/2} \times R^{-1} \)-plank \( \Theta_\theta \) centered at the origin, with its 1-edge parallel to \((\eta_\theta, 0)\) and \( R^{-1/2} \)-edge parallel to \((\eta_\theta^\perp, 0)\).
Then we decompose $Θ_{\varnothing}$ as in Section 4. For any dyadic number $R^{-1/2} \leq \varsigma \leq 1$, denote $Θ_{\varsigma,\partial} \subset Θ_{\varnothing}$ as the plank of dimensions $\varsigma^2 \times \varsigma R^{-1/2} \times R^{-1}$, centered at the original of direction, with its $\varsigma^2$-edge parallel to $(\eta_{\varnothing}, 0)$ and $\varsigma R^{-1/2}$-edge parallel to $(\eta_{\perp \varnothing}, 0)$. Let $ϕ_{\Theta_{\varsigma,\partial}} \in C^\infty_0$ that equals 1 on $Θ_{\varsigma,\partial}$ and 0 outside $Θ_{2\varsigma,\partial}$. Then define

$$Ω_{\varsigma,\varnothing} := Θ_{\varsigma,\varnothing} \setminus Θ_{\varsigma/2,\varnothing}, \quad Ψ_{Ω_{\varsigma,\varnothing}} := ϕ_{\Theta_{\varsigma,\varnothing}} - ϕ_{\Theta_{\varsigma/2,\varnothing}}, \quad \text{if } 2R^{-1/2} \leq \varsigma < 1,$$

and

$$Ω_{R^{-1/2},\varnothing} := Θ_{R^{-1/2},\varnothing} \sim B R^{-1}, \quad Ψ_{Ω_{R^{-1/2},\varnothing}} := ϕ_{Ω_{R^{-1/2},\varnothing}}.$$

Now we can decompose the domain of $\tilde{ω}$ in (6.1) to obtain

$$\int_{B_R} \left( \sum_{\varsigma} \sum_{\partial} \int e^{2\pi i (u^2_{\partial}(x,t), t) \cdot \tilde{ω}} \left| \tilde{F}_{\partial} \right|^2 \right)^2 (\tilde{ω}) \psi_{Ω_{\varsigma,\varnothing}} (\tilde{ω}) \, d\tilde{ω} \, dx \, dt$$

$$\quad = \int \left( \chi_{B_R} \cdot \sum_{\varsigma} \sum_{\partial} \int e^{2\pi i (u^2_{\partial}(x,t), t) \cdot \tilde{ω}} \left| \tilde{F}_{\partial} \right|^2 \right)^2 (\tilde{ω}) \psi_{Ω_{\varsigma,\varnothing}} (\tilde{ω}) \, d\tilde{ω} \, dx \, dt. \quad (6.2)$$

By Plancherel in $(x,t)$, it equals

$$\int \left( \int_{B_R} \sum_{\partial} \sum_{\varsigma} \int e^{2\pi i (u^2_{\partial}(x,t), t) \cdot \tilde{ω}} \left| \tilde{F}_{\partial} \right|^2 \right)^2 (\tilde{ω}) \psi_{Ω_{\varsigma,\varnothing}} (\tilde{ω}) \, d\tilde{ω} \, dx \, dt \, d\omega \, d\omega' \quad (6.3)$$

Notice that, by integration by parts in $(x,t)$, it suffices to consider $(\partial, (x,t), (\omega, \tilde{ω}))$ satisfying

$$\left| \frac{\partial (u^2_{\partial}(x,t), t)}{\partial (x,t)} \cdot \omega - \tilde{ω} \right| \lesssim R^{-1}. \quad (6.4)$$

We choose a smooth cutoff function $χ_\partial (x,t, \omega, \tilde{ω})$ that equals 1 for all $(x,t, \omega, \tilde{ω})$ satisfying (6.4). For each fixed $\varsigma$, divide the circle into $\varsigma^{-1}R^{-1/2}$-arcs $\tau$. When $(x,t), \varsigma, \tau, \omega$ are fixed and $\partial \subset \tau$, (6.4) and Lemma 4.2 imply that $\tilde{ω}$ essentially lies in a ball of radius $R^{-1}$. Therefore, there is no extra cost to replace $χ_\partial$ by $χ_{\partial, \tau}$, where $\partial, \tau$ again denotes the central $R^{-1/2}$-cap of $\tau$. For convenience, we denote $χ_\tau := χ_{\partial, \tau}$. Then the integral (6.3) becomes

$$\int \left( \int_{B_R} \sum_{\sigma} \sum_{\tau \in S_{\varsigma^{-1}R^{-1/2}}} \int_{Ω_{\sigma,\tau}} \left( \int_{Ω_{\sigma,\tau}} \sum_{\partial \subset \tau} \chi_\tau \, d\tilde{ω} \right) \, dx \, dt \right)^2 \, d\omega. \quad (6.5)$$

Now we apply Cauchy–Schwarz to (6.5). More precisely, it is bounded above by an integral on $\omega$, with integrand the product of

$$\int_{B_R} \sum_{\sigma} \sum_{\tau \in S_{\varsigma^{-1}R^{-1/2}}} \int_{Ω_{\sigma,\tau}} \chi_\tau \, d\tilde{ω} \, dx \, dt \quad (6.6)$$
\[
\int_{B_R} \sum_{\sigma} \sum_{\tau} \int_{\Omega_{\sigma,\tau}} \chi_{\tau} \left| \sum_{\theta \in S_{R^{-1/2} : \partial \subset \tau}} e^{2\pi i ((u^\beta(x,t),t) - \tilde{\omega} - (x,t) - \tilde{\omega}) (|F_{\theta}|^2)^\wedge(\tilde{\omega})} \right|^2 d\tilde{\omega} \, dx \, dt. \quad (6.7)
\]

For the first factor (6.6), we have seen that when \( \omega, (x, t), \sigma, \tau \) are fixed, (6.4) and Lemma 4.2 imply that \( \tilde{\omega} \) essentially lies in a ball of radius \( R^{-1} \). Therefore (6.6) is
\[
\lesssim_\varepsilon R^{-3} \sum_{\sigma} \int_{B_R} \sum_{\tau} \chi^\circ_{\tau}(x,t,\omega) \, dx \, dt,
\]
where
\[
\chi^\circ_{\tau}(x,t,\omega) = \sup_{\tilde{\omega} \in \Omega_{\sigma,\tau}} \chi_{\tau}(x,t,\omega,\tilde{\omega}).
\]

By (6.4), one can see that \( \chi^\circ_{\tau} \) vanishes unless \( \omega \) essentially lies in the \( R^{-1} \)-neighborhood of
\[
\Omega_{\sigma,x,t} := \bigcup_{\partial} \Omega_{\sigma,\partial} \frac{\partial(u^\beta(x,t),t)}{\partial_\tau(x,t)}.
\]

Here comes the key observation from Guth–Wang–Zhang (see [12, Lemma 4.2]). Let \( N_{R^{-1}}(\Omega_{\sigma,x,t}) \) denote the \( R^{-1} \) neighborhood of \( \Omega_{\sigma,x,t} \). When \( \sigma, (x,t), \omega \in N_{R^{-1}}(\Omega_{\sigma,x,t}) \) are fixed, there are only finitely many \( \tau \) with \( \chi^\circ_{\tau} \) nonvanishing. This is proved in [12] for the standard cone \((\xi, |\xi|) \in \mathbb{R}^3 \), while as we explained right after Theorem 1.6, it remains valid on a general cone with its curvature bounded from above and below. In particular, it applies to our cones for each fixed \((x,t)\), thanks to our cinematic curvature assumption. For completeness, we provide proof of a generalization of [12, Lemma 4.2] to a general cone in the Appendix (Lemma A.2).

The discussion above implies that (6.8) is bounded above by
\[
R^{-3} \sum_{\sigma} \int_{B_R} dx \, dt \lesssim \log R.
\]

We just showed that the factor (6.6) is \( \lesssim_\varepsilon 1 \). It remains to consider the integral on \( \omega \) with integrand (6.7), which equals
\[
\int \int_{B_R} \sum_{\sigma} \sum_{\tau \in S_{-1R^{-1/2}}} \int_{\Omega_{\sigma,\tau}} \chi_{\tau} \left| \sum_{\theta \in S_{R^{-1/2} : \partial \subset \tau}} e^{2\pi i ((u^\beta(x,t),t) - \tilde{\omega} - (x,t) - \tilde{\omega}) (|F_{\theta}|^2)^\wedge(\tilde{\omega})} \right|^2 d\tilde{\omega} \, dx \, dt \, d\omega
\]

When \((x, t), \sigma, \tau, \tilde{\omega}\) are fixed, (6.4) and Lemma 4.1 imply that \( \chi_{\tau} \) vanishes unless \( \omega \) essentially lies in a ball of radius \( R^{-1} \). Therefore, by integrating \( d\omega \), it is bounded above by
\[
R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau \in S_{-1R^{-1/2}}} \int_{\Omega_{\sigma,\tau}} \left| \sum_{\theta \in S_{R^{-1/2} : \partial \subset \tau}} e^{2\pi i ((u^\beta(x,t),t) - \tilde{\omega} - (x,t) - \tilde{\omega}) (|F_{\theta}|^2)^\wedge(\tilde{\omega})} \right|^2 d\tilde{\omega} \, dx \, dt.
\]
Fix \((x, t)\) and apply Plancherel. It becomes

\[
R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau} \int_{\partial \in \mathbb{S}_{R-1/2}} \sum_{\theta \subset \tau} \left| \int_{\mathbb{R}^3} e^{2\pi i ((\lambda^2(x,t),t)+(u_1,t_1))} \hat{F}_\theta(u_2,t_2) \right|^2 du_1 dt_1 dx dt
\]

\[
= R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau} \int_{\partial \in \mathbb{S}_{R-1/2}} \sum_{\theta \subset \tau} \int K_\theta(x,t,u_1,t_1,u_2,t_2) \left| \hat{F}_\theta(u_2,t_2) \right|^2 du_2 dt_2 \] dx dt,
\]

where the kernel \(K_\theta\) is given by

\[
K_\theta(x,t,u_1,t_1,u_2,t_2) = \int e^{2\pi i ((\lambda^2(x,t),t)+(u_1,t_1)-(u_2,t_2))} \psi_{\Theta,\sigma,\tau}(\hat{\omega}) d\hat{\omega}.
\]

Notice that \(\Theta,\sigma,\tau\) and \(\tilde{U}\) defined in (4.1) are dual rectangular boxes to each other. Therefore,

\[
\|K_\theta\|_{L^\infty} \lesssim |\Theta,\sigma,\tau| = |\tilde{U}|^{-1} \approx |U|^{-1},
\]

and by integration by parts \(K_\theta\) has rapid decay unless essentially

\[
(u^j_\theta(x,t),t) + (u_1,t_1) - (u_2,t_2) \in \Theta,\sigma,\tau = \tilde{U}.
\]

We cover the domain of \((u_1,t_1)\) in (6.9) by a collection of finitely overlapping rectangular boxes \(\tilde{U}_\tau := \{U\}\), where each \(U\) is a translation copy of \(\tilde{U}_\tau\). By (6.12) and Lemma 4.3, it follows that when \((x,t)\) is fixed and \((u_1,t_1)\) lies in some \(U \subset \tilde{U}_\tau\), we only need to consider \((u_2,t_2)\) contained in some translation of \(U\), denoted by \(U_{x,t}\). We would like to point out that

(i) \(U_{x,t}\) only depends on \(U\), not on specific point \((u_1,t_1) \in U\).
(ii) As \(U\) runs over all rectangular boxes parallel to \(\tilde{U}_\tau\), so does \(U_{x,t}\).

Therefore, with the estimate (6.11), we have (6.9)

\[
\lesssim R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau} \sum_{U \subset \tilde{U}_\tau} \int_{\mathbb{R}^3} \left| \sum_{\partial \in \mathbb{S}_{R-1/2}} \sum_{\theta \subset \tau} \int_{U_{x,t}} \left| \tilde{U}^{-1} \hat{F}_\theta(u_2,t_2) \right|^2 du_2 dt_2 \right|^2 dx dt
\]

\[
\lesssim R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau} \sum_{U \subset \tilde{U}_\tau} \left| \tilde{U}^{-1} \left( \int_{U_{x,t}} \sum_{\theta \in \mathbb{S}_{R-1/2}} \left| \hat{F}_\theta \right|^2 \right)^2 \right| dx dt
\]

\[
= R^{-3} \int_{B_R} \sum_{\sigma} \sum_{\tau} \sum_{U \subset \tilde{U}_\tau} \left| \tilde{U}^{-1} \left( \int_{\tilde{U}} \sum_{\theta \in \mathbb{S}_{R-1/2}} \left| \hat{F}_\theta \right|^2 \right)^2 \right| dx dt
\]

\[
\approx \sum_{\sigma} \sum_{\tau} \sum_{U \subset \tilde{U}_\tau} \left| \tilde{U}^{-1} \left( \int_{\tilde{U}} \sum_{\theta \in \mathbb{S}_{R-1/2}} \left| \hat{F}_\theta \right|^2 \right)^2 \right|
\]

where the second line follows from (i), the third line follows from (ii), and the last line follows because nothing in the integrand depends on \((x,t)\).
Finally, inside the square, we change variables back from $F_\theta(u, t)$ to $F_\theta(x, t)$. Since $|U| \approx |U|$, this quantity is approximately equal to

$$\sum_\sigma \sum_\tau \sum_{U \in U_\tau} |U|^{-1} \left( \int_{\rho \in \mathbb{S}_{\tau}} \sum_{\partial \subset \tau} |F_\theta|^2 \right)^2,$$

as desired.

7 | LORENTZ RESCALING

In this section, we establish the Lorentz rescaling lemma, which plays a crucial role in the induction on scales argument. Lorentz rescaling for the cone is the analogous counterpart of parabolic rescaling for the parabola, and both of them connect estimates from different scales. To proceed, we first formulate an equivalent yet more quantitative definition of wave envelopes.

7.1 | An equivalent description of $U$

Let $0 < s \leq 1$ be a dyadic number and fix an arc $\tau \subset S^1$ with center $\eta_\tau \in S^1$ and aperture $d(\tau) = s$. There exists an orthogonal matrix $A_\tau$ such that $A_\tau e_2 = \eta_\tau$. For each $m = (m_1, m_2) \in \mathbb{Z}^2$, let $m_\tau = A_\tau (s^{-1} m_1, m_2) \in \mathbb{R}^2$. We define $B_{\tau, m}$ by

$$B_{\tau, m} := \{ y = (y_1, y_2) \in \mathbb{R}^2 : |\Pi_{\eta_\tau} (y - m_\tau)| \leq C, |\Pi_{\eta_\perp} (y - m_\tau)| \leq Cs^{-1} \}.$$

where $\Pi_{\eta_\tau}, \Pi_{\eta_\perp}$ denote the projection onto $\eta_\tau$ and $\eta_\perp$, respectively. It is then straightforward to check that the whole space $\mathbb{R}^2$ is covered by $\{B_{\tau, m}\}_m$ with finite overlaps. Define $U_{\tau, m}$ to be

$$\{ z : |\Pi_{\eta_\tau} (\partial_{\eta} \phi^4(z, \eta_\tau) - R^2 m_\tau)| \leq CRs^2, \quad |\Pi_{\eta_\perp} (\partial_{\eta} \phi^4(z, \eta_\tau) - R^2 m_\tau)| \leq CRs, |t| \leq R \}.$$  \hspace{1cm} (7.1)

One may check that $U_{\tau, m}$ gives a curved slab of length $R$ and a rectangular cross-section of size $Rs^2 \times Rs$.

We claim that the definition of the wave envelope in (7.1) is equivalent to that in (2.10). Indeed, by changing variables

$$x \mapsto x_\tau^\lambda (u, t), t \mapsto t,$$

the curved tube $U_{\tau, 0}$ is mapped to

$$\{(u, t) : |t| \leq R, |\Pi_{\eta_\tau} (u)| \leq CRs^2, |\Pi_{\eta_\perp} (u)| \leq CRs \},$$

which corresponds to $V_\tau \times (-R, R)$. Each $U_{\tau, m}$ then corresponds to some $V \in \mathcal{V}_\tau$. Therefore, our claim is true.

We cover $B_R$ by $\{U_{\tau, m}\}_m$ so that

$$B_R \subset \bigcup_m U_{\tau, m},$$

with finite overlaps.
We consider $S(r, R, \lambda)$ as defined in (2.12). Using the equivalent definition of the curved tube $U_{\tau}$, we have, up to a rapid decay in $\lambda$,

$$
\sum_{B_r \subseteq B_R} |B_r|^{-1} \left\| \left( \sum_{\theta \in S_{r^{-1/2}}} |F_{\theta}^\lambda f|^2 \right) \right\|_{L^2(B_r)}^{1/2} \leq S(r, R, \lambda) \sum_{R^{-1/2} \leq \varepsilon \leq 1} \sum_{\tau \in S_\varepsilon} \sum_{m} |U_{\tau, m}|^{-1} \left\| \left( \sum_{\delta \in \tau} |F_{\delta}^\lambda f|^2 \right) \right\|_{L^2(U_{\tau, m})}^{1/2}.
$$

7.2 | Lorentz rescaling

With the above ingredients, we are ready to establish the Lorentz rescaling lemma. First, we take $\phi(x, t, \eta) = x \cdot \eta + t|\eta|$ as an example to illustrate the idea. In this case, the hypersurface associated with $x \cdot \eta + t|\eta|$ is the standard cone $(\eta, |\eta|)$. To proceed, we apply the (Euclidean) Lorentz transformation to the cone first, which will put us in a better position to perform a parabolic rescaling. By the change of variables,

$$
\eta_1 \mapsto \eta_1, \quad \eta_2 \mapsto \frac{1}{\sqrt{2}}(\eta_3 - \eta_2), \quad \eta_3 \mapsto \frac{1}{\sqrt{2}}(\eta_3 + \eta_2),
$$

the cone $(\eta, |\eta|)$ is transformed to the tilted cone $\eta_3 = \eta_1^2/2\eta_2$, the formula of which is invariant under parabolic rescaling in $\eta_2$ and $\eta_3$. Moreover, this cone is tangent to the $\eta_1\eta_2$-plane. See Figure 2.

For our case, the phase function $\phi(x, t, \eta)$ satisfies the cinematic curvature condition, which may not be translation invariant. A significant difference is that the associated cone in phase space is given by $(z, \partial_z \phi(z, \eta))$, which varies smoothly depending on $z$. Therefore, the Euclidean Lorentz transformation does not work directly. To overcome this difficulty, we perform a change of variables with respect to the $(x, t)$-variables as displayed in Section 2.1.

Denote

$$
\tilde{\phi}(u, t, \eta) := \phi(x_c(u, t), t, \eta).
$$

One can easily check that $\tilde{\phi}(u, t, \eta)$ also satisfies the cinematic curvature condition. By the discussion in Section 2.1, the normal direction of the cone $\partial_{u, \lambda}\tilde{\phi}(u, t, \eta)$ is $(0, 0, 1)$ at $\eta = \eta_{\tau}$. This manifests
that the above change of variable plays a similar role as the Euclidean Lorentz transformation, because both of which tilt the cone so that it is tangent to the horizontal plane.

Now we restate the Lorentz rescaling lemma as follows.

**Lemma 7.1.** Let $r_1 \leq r_2 \leq r_3 \leq R \leq \lambda^{1-\varepsilon}$, then

$$S(r_1, r_3, \lambda) \leq \log r_2 S(r_1, r_2, \lambda) \max_{r_2^{-1/2} \leq s \leq 1} S(s_2 r_2, s^2 r_3, s^2 \lambda).$$

**Proof.** By the definition of $S(r_1, r_2, \lambda)$, we obtain

$$\sum_{B_{r_1} \subset R} |B_{r_1}|^{-1/2} \left\| \left( \sum_{\delta \in S_{-1/2}} |F^\delta_{\theta} f|^2 \right)^{1/2} \right\|_{L^2(B_{r_1})}^4 \leq S(r_1, r_2, \lambda) \sum_{r_2^{-1/2} \leq s \leq 1} \sum_{\tau \in S_{s}} \sum_{m} |U_{\tau, m}|^{-1/2} \left\| \left( \sum_{\delta \subset \tau : \delta \in S_{-1/2}} |F^\delta_{\theta} f|^2 \right)^{1/2} \right\|_{L^2(U_{\tau, m})}^4 + \text{RapDec}(\lambda)\|f\|_{L^2}^4.$$  \hspace{1cm} (7.3)

To complete the proof, it suffices to show that for a fixed $\tau$ with aperture $s$, we have

$$\sum_{m} |U_{\tau, m}|^{-1/2} \left\| \left( \sum_{\delta \subset \tau : \delta \in S_{-1/2}} |F^\delta_{\theta} f|^2 \right)^{1/2} \right\|_{L^2(U_{\tau, m})}^4 \leq S(s_2 r_2, s^2 r_3, s^2 \lambda) \sum_{r_3^{-1/2} \leq s' \leq 1} \sum_{\tau' \subset \tau \in S_{s'}} \sum_{m} |U_{\tau', m}|^{-1/2} \left\| \left( \sum_{\delta \subset \tau' : \delta \in S_{-1/2}} |F^\delta_{\theta} f|^2 \right)^{1/2} \right\|_{L^2(U_{\tau', m})}^4 + \text{RapDec}(\lambda)\|f\|_{L^2}^4.$$  \hspace{1cm} (7.4)

In fact, if we plug (7.4) into (7.3) and notice that the number of dyadic numbers located in $(r_2^{-1/2}, 1)$ is $\log r_2$, the lemma follows.

Now we prove (7.4). For a given $\tau$, let $A_\tau$ be an orthogonal matrix such that $A_\tau e_2 = \eta_\tau$. We shall perform the following change of variables

$$x \mapsto x^\lambda_\tau(A_\tau u, t), \quad t \mapsto t, \quad \eta \mapsto A_\tau \eta.$$

Define

$$\tilde{F}_{\theta}^\lambda f(u, t) : = \int e^{i\tilde{\phi}_\lambda(u, t, \eta)} \tilde{a}_{\theta}^\lambda(u, t, \eta) \hat{f}(\eta) d\eta,$$

$$\tilde{\phi}_\lambda(u, t, \eta) : = \phi^\lambda(x^\lambda_\tau(A_\tau u, t), A_\tau \eta), \quad \tilde{a}_{\theta}^\lambda(u, t, \eta) = a_{\theta}^\lambda(x^\lambda_\tau(A_\tau u, t), t, A_\tau \eta).$$
Under the new coordinates, the curved tube $U_{\tau,m}$ is straightened and given by 

$$U_{\tau,m} := \{(u,t) : |t| \leq r_2, |u_2 - r_2 s^2 m_2| \leq r_2 s^2, |u_1 - r_2 s m_1| \leq r_2 s\},$$  

(7.5)

and the left-hand side of (7.4) is transformed to 

$$\sum_m |U_{\tau,m}|^{-1} \left\| \sum_{\theta \in S_{r_2^{-1/2}}} |\tilde{f}_{\theta} L^{s} f|^2 \right\|_{L^2(U_{\tau,m})}^{1/2} \right\|^4.$$  

(7.6)

Next, we perform parabolic rescaling 

$$\eta_1 \mapsto s \eta_1, \eta_2 \mapsto \eta_2, u_1 \mapsto s^{-1} u_1, u_2 \mapsto u_2, t \mapsto s^{-2} t.$$  

Define $\tilde{\phi}$ and $\tilde{a}$ to be 

$$\tilde{\phi}(u,t,\eta) := s^{-2} \tilde{\phi}(s u_1, s^2 u_2, t, s \eta_1, \eta_2),$$

$$\tilde{a}_{\theta}(u,t,\eta) := \tilde{a}_{\theta}(s u_1, s^2 u_2, t, s \eta_1, \eta_2).$$

Correspondingly, by (7.5), $U_{\tau,m}$ is transformed to a cube $Q_m$ of side-length $r_2 s^2$, namely, 

$$Q_m = \{(u,t), |t| \leq r_2 s^2 : |u_2 - r_2 s^2 m_2| \leq r_2 s^2, |u_1 - r_2 s^2 m_1| \leq r_2 s^2\}.$$  

Then, (7.6) becomes 

$$\sum_m |Q_m|^{-1} \left\| \sum_{\theta \in S_{r_2^{-1/2}}} |\tilde{f}_{\theta} L^{s} f|^2 \right\|_{L^2(Q_m)}.$$  

where $\tilde{f}_{\theta} L^{s} f$ is defined by 

$$\tilde{f}_{\theta} L^{s} f(u,t) := \int e^{i \tilde{\phi}_{\theta}(u,t,\eta)} \tilde{a}_{\theta}(u,t,\eta) \tilde{f}(\eta) \, d\eta \quad \text{with} \quad \tilde{f}(\eta) := f(s \eta_1, \eta_2).$$

We claim that the phase function $\tilde{\phi}$ satisfies our assumptions on the phase functions. Assuming this, by the definition of $S(s^2 r_2, s^2 r_3, s^2 \lambda)$, we have 

$$\sum_m |Q_m|^{-1} \left\| \sum_{\theta \in S_{r_2^{-1/2}}} |\tilde{f}_{\theta} L^{s} f|^2 \right\|_{L^2(Q_m)}^{1/2} \leq S(s^2 r_2, s^2 r_3, s^2 \lambda) \sum_{r_2^{-1/2} \leq s \leq s} \sum_{\tau' \in S_{s'}} \sum_{m} |U_{\tau',m}|^{-1}.$$  

$$\text{RapDec}(\lambda) \left\| f \right\|_{L^2}^{4} + \sum_{\theta' \in S_{r_3^{-1/2}}} |\tilde{f}_{\theta'} L^{s} f|^2 \right\|_{L^2(U_{\tau,m})}^{1/2} \right\|^4.$$  

(7.7)
Note that $U_{r',m}$ is defined by

$$
U_{r',m} := \{(u, t), |t| \leq r_3 s^2 : |\Pi_{\eta_{r'}} (\partial_{\eta} \tilde{\phi}_{x}^s(u, t, \eta_{r'}) - r_3 s^2 m_{r'})| \leq r_3 (s')^2, \\
|\Pi_{\eta_{r'}} (\partial_{\eta} \tilde{\phi}_{x}^s(u, t, \eta_{r'}) - r_3 s^2 m_{r'})| \leq r_3 s s',
$$

with $\eta_{r'} \in S^1$ satisfying

$$
\eta_{r'} \parallel A_{r'} \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} \eta_{r'},
$$

and $m_{r'} := A_{r'}(s/s' m_1, m_2)$. After changing back to the original variables, we obtain (7.4).

To conclude this section, it remains to check that $\tilde{\phi}_s$ satisfies our assumptions on the phase and the amplitude functions, that is, $\tilde{\phi}_s$ satisfies the quantitative conditions $C_2$ and the cinematic curvature condition and the conditions $C_1, C_3$ on the amplitude functions.

It is easy to verify the conditions $C_1, C_3$. Indeed, one can employ the following change of variables:

$$
(u, t) \mapsto (u, t)/A, \eta \mapsto \eta/A,
$$

where $A > 0$ is a sufficiently large absolute constant, and then decompose the enlarged domain into a finite number of smaller ones. Thus, it suffices to verify the condition $C_2$.

For the new phase function $\tilde{\phi}_s$, by homogeneity, we have

$$
s^{-2} \tilde{\phi}(su_1, s^2 u_2, t, s \eta_1, \eta_2) = s^{-2} \eta_2 \tilde{\phi}(su_1, s^2 u_2, t, s \eta_1/\eta_2, 1).
$$

Using Taylor’s formula, we get

$$
\tilde{\phi}(su_1, s^2 u_2, t, s \eta_1/\eta_2, 1) = \tilde{\phi}(su_1, s^2 u_2, t, e_2) + s \partial_{\eta_1} \tilde{\phi}(su_1, s^2 u_2, t, e_2) \eta_1/\eta_2 \\
+ 1/2 s^2 \partial_{\eta_1}^2 \tilde{\phi}(su_1, s^2 u_2, t, e_2) (\eta_1/\eta_2)^2 + O(s^3(\eta_1/\eta_2)^3).
$$

By Euler’s formula, we obtain

$$
\tilde{\phi}(su_1, s^2 u_2, t, e_2) = \partial_{\eta_2} \tilde{\phi}(su_1, s^2 u_2, t, e_2).
$$

Taking (7.10) and (7.9) into (7.8) and invoking (2.6), we have

$$
\tilde{\phi}_s = s^{-2} \tilde{\phi}(su_1, s^2 u_2, t, s \eta_1, \eta_2) = u \cdot \eta + \frac{1}{2} \partial_{\eta_1}^2 \tilde{\phi}(su_1, s^2 u_2, t, e_2) \eta_1^2/\eta_2 + O(s(\eta_1^3/\eta_2^3)),
$$

where the implicit constant in the big $O$ only depends on a finite number of derivatives of $\phi$. Therefore, the condition $C_2$ can be ensured.

Finally, let us check that $\tilde{\phi}_s$ satisfies the cinematic curvature condition.

$$
\tilde{\phi}_s(u, t, \eta) = s^{-2} \phi(x_\tau(A_\tau(su_1, s^2 u_2), t), t, A_\tau(s \eta_1, \eta_2)).
$$
By the homogeneity of $\phi$, it is clear that $\tilde{\phi}_s(u, t, \eta)$ is homogeneous of degree 1 with respect to $\eta$. Since $A_\tau e_2 = \eta_\tau$, we have

$$s^{-2} \partial_\eta \tilde{\phi}_s(u, t, e_2) = s^{-2} \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} A_\tau \partial_\eta \phi(x_\tau(A_\tau(su_1, s^2u_2), t), t, \eta_\tau)$$

$$= s^{-2} \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix} A_\tau A_\tau \begin{pmatrix} s & 0 \\ 0 & s^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \quad (7.11)$$

Therefore,

$$s^{-2} \partial_\eta \partial_u \tilde{\phi}_s(u, t, e_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and then, the nondegeneracy condition (1.7) follows from the continuity of the phase function and our assumption of smallness on the support of $a(x, t, \eta)$.

Now we verify the curvature condition (1.8) of the phase function $\tilde{\phi}_s$. The two tangent vectors of the hypersurface $\{\partial_{(u,t)} \tilde{\phi}_s(u, t, \eta) : \eta \in \text{supp } a^s(z, \cdot)\}$ at $e_2$ are

$$\partial_\eta \partial_{(u,t)} \tilde{\phi}_s(u, t, e_2), \partial_\eta \partial_{(u,t)} \tilde{\phi}_s(u, t, e_2).$$

Recall from (7.11) that we have normalized the normal vector to the hypersurface $\{\partial_{(u,t)} \tilde{\phi}_s(u, t, \eta) : \eta \in \text{supp } a^s(z, \cdot)\}$ at $e_2$ to be $e_3$. A simple calculation gives that

$$\partial^2_{\eta \eta} \langle \partial_{(u,t)} \tilde{\phi}_s(u, t, \eta), e_3 \rangle|_{\eta=e_2} = \partial^2_{\eta \eta} \partial_t \tilde{\phi}_s(u, t, \eta)|_{\eta=e_2}. \quad (7.12)$$

Since

$$\partial^2_{\eta \eta} \partial_t \tilde{\phi}_s(u, t, \eta)|_{\eta=e_2} = B^t \partial^2_{\eta \eta} \partial_t \phi(x_\tau(A_\tau(su_1, s^2u_2), t), t, \eta_\tau)B,$$

where

$$B = A_\tau \begin{pmatrix} s & 0 \\ 0 & 1 \end{pmatrix}.$$

Also, note that

$$\text{rank } \partial^2_{\eta \eta} \langle \partial_{(u,t)} \phi(x_\tau(u, t), t, \eta_\tau), e_3 \rangle = \text{rank } \partial^2_{\eta \eta} \partial_t \phi(x_\tau(u, t), t, \eta_\tau) = 1. \quad (7.13)$$

Finally, by (7.13),(7.12), and the nondegeneracy of $B$, we have

$$\text{rank } \partial^2_{\eta \eta} \partial_t \tilde{\phi}_s(u, t, \eta)|_{\eta=e_2} = 1.$$

The proof of the claim is complete. □

**APPENDIX**

**A.1 | Theorem 1.6 implies local smoothing estimate**

For the reader’s convenience, we provide details on how to get a local smoothing estimate from Theorem 1.6 via an argument from [21].
Let $\vartheta \subset S^1$ be an arc of aperture $\lambda^{-1/2}$. By a standard Littlewood–Paley decomposition, Conjecture 1.3 can be deduced from (1.13), which says
\[
\|F^\lambda f\|_{L^4(\mathbb{R}^3)} \leq C_\varepsilon \lambda^{1/4+\varepsilon} \|f\|_{L^4(\mathbb{R}^2)}.
\]

Assuming the square function estimate
\[
\left\| \sum_\vartheta |F^\lambda_\vartheta f|^2 \right\|_{L^4(\mathbb{R}^3)} \leq \varepsilon \lambda^{1/4+\varepsilon} \|f\|_{L^4(\mathbb{R}^2)}.
\]

one can see that to establish (1.10), it suffices to prove
\[
\left(\sum_\vartheta |F^\lambda_\vartheta f|^2\right)^{1/2} \leq \varepsilon \lambda^{1/4+\varepsilon} \|f\|_{L^4(\mathbb{R}^2)}.
\]

To this end, we use $K^\lambda_\vartheta$ to denote the kernel of the operator $F^\lambda_\vartheta$, that is,
\[
K^\lambda_\vartheta(x, t, y) := \int_{\mathbb{R}^2} e^{i(\phi^\lambda_\vartheta(x, t, \eta) - y \cdot \eta)} \chi_\vartheta(\eta) \, d\eta.
\]

By integration by parts in $\eta$, we have for each $N \in \mathbb{N}$
\[
|K^\lambda_\vartheta(x, t, y)| \lesssim_N \lambda^{-1/2} \left( 1 + |\Pi_{\eta^\vartheta}(\partial_\eta^\vartheta \phi^\lambda(z, \eta^\vartheta) - y) + \lambda^{-1/2} |\Pi_{\eta^\vartheta}(\partial_\eta^\vartheta \phi^\lambda(z, \eta^\vartheta) - y)| \right)^{-N}. \tag{A.1}
\]

By Riesz’s representation theorem, we have
\[
\left\| \left( \sum_\vartheta |F^\lambda_\vartheta f|^2 \right)^{1/2} \right\|_{L^4(\mathbb{R}^3)}^2 = \sup_{\|g\|_{L^2(\mathbb{R}^3)} = 1} \int_{\mathbb{R}^3} \sum_\vartheta |F^\lambda_\vartheta f(x, t)|^2 g(x, t) \, dx \, dt.
\]

Then, by Hölder’s inequality,
\[
\int_{\mathbb{R}^3} \sum_\vartheta |F^\lambda_\vartheta f(x, t)|^2 g(x, t) \, dx \, dt
\]
\[
= \int_{\mathbb{R}^3} \sum_\vartheta \left| \int_{\mathbb{R}^2} K^\lambda_\vartheta(x, t, y) f_\vartheta(y) \, dy \right|^2 g(x, t) \, dx \, dt
\]
\[
\lesssim \int \sum_\vartheta \left( \int_{\mathbb{R}^2} |K^\lambda_\vartheta(x, t, y)| \, dy \right) \left( \int |K^\lambda_\vartheta(x, t, y)|^2 |f_\vartheta(y)|^2 \, dy \right) |g(x, t)| \, dx \, dt
\]
\[
\lesssim \left( \sup_{\vartheta, x, t} \int_{\mathbb{R}^2} |K^\lambda_\vartheta(x, t, y)| \, dy \right) \int_{\mathbb{R}^2} \left( \sup_\vartheta \int |K^\lambda_\vartheta(x, t, y)||g(x, t)| \, dx \, dt \right) \sum_\vartheta |f_\vartheta|^2 \, dy
\]
\[
\lesssim \int_{\mathbb{R}^2} \left( \sum_\vartheta |f_\vartheta(y)|^2 \right) \left( \sup_\vartheta \int_{\mathbb{R}^3} |K^\lambda_\vartheta(x, t, y)g(x, t)| \, dx \, dt \right) \, dy
\]
\[
\lesssim \left\| \sum_\vartheta |f_\vartheta|^2 \right\|_{L^2(\mathbb{R}^2)} \left\| \sup_\vartheta \int_{\mathbb{R}^3} |K^\lambda_\vartheta(x, t, y)g(x, t)| \, dx \, dt \right\|_{L^2(\mathbb{R}^2)}.
\]
where, in the second to last line, we use (A.1) to obtain
\[ \sup_{\theta, x, t} \int |K_{\theta}^\lambda(x, t, y)| \, dy \lesssim 1. \]

Then the following two inequalities complete the proof.

**Square function inequality**
\[ \left\| \left( \sum_{\theta} |f_\theta|^2 \right)^{\frac{1}{2}} \right\|_{L^p(\mathbb{R}^2)} \lesssim \lambda^\epsilon \|f\|_{L^p(\mathbb{R}^2)}, \quad 2 \leq p \leq 4. \] (A.2)

**Nikodym maximal function**
\[ \left\| \sup_{\theta} \int_{\mathbb{R}^3} |K_{\theta}^\lambda(x, t, y)g(x, t)| \, dx \, dt \right\|_{L^2(\mathbb{R}^2)} \lesssim \lambda^{1/2+\epsilon}. \] (A.3)

(A.2) can be found in [7]. It remains to check (A.3). One can see from (A.1) that when \( \theta, y \) are fixed, the essential support of \( K_{\theta}^\lambda \) is contained in a curved tube \( T_{\theta, y} \) defined by
\[ T_{\theta, y} := \{(x, t) : |t| \leq \lambda, \Pi_{\eta}(\partial_{\eta} \phi^{\lambda}(x, t, \eta_\theta) - y) \leq 1, |\Pi_{\eta^\perp}(\partial_{\eta} \phi^{\lambda}(x, t, \eta_\theta) - y) | \leq \lambda^{1/2}\}. \]

Thus, we may view (A.3) as a Nikodym-type maximal function estimate, in the sense that
\[ \left( \int_{\mathbb{R}^2} \sup_{\theta} \left| \frac{1}{|T_{\theta, y}|} \int_{T_{\theta, y}} |g(x, t)| \, dx \, dt \right|^2 \, dy \right)^{\frac{1}{2}} \lesssim \epsilon \lambda^{-1/2+\epsilon} \|g\|_{L^2(\mathbb{R}^3)}. \] (A.4)

Equation (A.4), in turn, is a direct consequence of the following Nikodym maximal function estimate from [21].

**Theorem A.1 ([21]).** Assume \( 1 \leq R \leq \lambda \) and \( \phi \) satisfies the cinematic curvature condition. Let tube \( T_{\theta, y} \) be defined as follows:
\[ T_{\theta, y} := \{(x, t) \in B_R(0) \times [0, R] : |(\partial_{\eta} \phi^{\lambda}(x, t, \eta_\theta) - y)| \leq 1\}, \]
then
\[ \left( \int_{\mathbb{R}^2} \sup_{\theta} \left| \frac{1}{|T_{\theta, y}|} \int_{T_{\theta, y}} |g(x, t)| \, dx \, dt \right|^2 \, dy \right)^{\frac{1}{2}} \lesssim \epsilon R^{-1/2+\epsilon} \|g\|_{L^2(\mathbb{R}^3)}. \] (A.5)

It is not hard to see that (A.5) implies (A.4). Indeed, one may decompose \( T_{\theta, y} \) into a collection of finitely overlapping tubes \( \{T_{\theta, y'}\}_{y'} \) such that
\[ T_{\theta, y} \subset \bigcup_{y'} T_{\theta, y'}. \]

Therefore,
\[ \frac{1}{|T_{\theta, y}|} \int_{T_{\theta, y}} |g(x, t)| \, dx \, dt \leq \operatorname{Avg}_{y'} \frac{1}{|T_{\theta, y'}|} \int_{T_{\theta, y'}} |g(x, t)| \, dx \, dt, \] (A.6)
where $\text{Avg}_n$ is defined by

$$\text{Avg}_n a_n := \frac{1}{\#\{n\}} \sum_n a_n.$$  

By Minkowski’s inequality, (A.4) follows from (A.5).

Finally, we remark that the higher dimensional analog of the above Nikodym maximal function estimate (A.5) is an interesting open problem. Nevertheless, as was shown in [2] ([19] for manifolds), in higher dimensions, we, in general, do not have favorable variable coefficient Nikodym maximal function estimates, and therefore going from square function estimates to local smoothing bounds is more subtle in higher dimensions.

A.2 | Finite overlapping of the planks for a general cone

Let $f : [0, 1] \to \mathbb{R}$ be a smooth function with $|f'| \leqslant C_1$ and $9/10 \leqslant f'' \leqslant 11/10$, where $C_1 > 0$ is an absolute constant. We shall use $f$ to model (a local piece of) a general cone with its curvature bounded above and below. Define the cone

$$C(s, t) := \{(st, sf(t), s) : 0 \leqslant s \leqslant 1, 0 \leqslant t \leqslant 1\}.$$  

Let

$$C(t) := \frac{(t, f(t), 1)}{\sqrt{1 + t^2 + f(t)^2}}, \quad T(t) := \frac{(1, f'(t), 0)}{\sqrt{1 + [f'(t)]^2}}, \quad N(t) := \frac{C(t) \times N(t)}{|C(t) \times N(t)|}.$$  

Let $r^{-1} < \sigma \leqslant 1$ and $0 \leqslant i \leqslant \frac{1}{2} \sigma r$ with $i \in \mathbb{N}$. Define $t_i := (i + \frac{1}{2})\sigma^{-1} r^{-1}$. Define the planks

$$\Theta(\sigma, t_i) := \{\xi \in \mathbb{R}^3 : |C(t_i) \cdot \xi| \leqslant \sigma^2, |T(t_i) \cdot \xi| \leqslant r^{-1} \sigma, |N(t_i) \cdot \xi| \leqslant r^{-2}\}.$$  

Let

$$\Theta_\sigma := \bigcup_i \Theta(\sigma, t_i).$$  

Now we are ready to state a generalization of [12, Lemma 4.2] for cone $C$.

Lemma A.2. If $\xi \in \Theta_\sigma \setminus \Theta_{\sigma/2}$, then there exist at most $C_0$ planks that are contained in $\Theta_\sigma$ overlapping at $\xi$, where $C_0$ is an absolute constant that only depends on $f$.

Proof. It is easy to verify the case when $\xi$ is in the annulus

$$A_\sigma := \left\{ \xi : \frac{\sigma^2}{4} \leqslant |\xi| \leqslant \sigma^2 \right\},$$

the planks $\{\Theta(\sigma, t_i) \cap A_\sigma\}$ are finitely overlapping. It remains to consider the case when $|\xi| \leqslant \frac{\sigma^2}{4}$.

If $|\xi| \leqslant \frac{\sigma^2}{4}$ and $\xi \in \Theta_\sigma \setminus \Theta_{\sigma/2}$, then $\xi \in \bigcup_i \Theta(\sigma, t_i)$, where

$$\bar{\Theta}(\sigma, t_i) := \{\xi \in \mathbb{R}^3 : |C(t_i) \cdot \xi| \leqslant \frac{\sigma^2}{4}, cr^{-1} \sigma \leqslant |T(t_i) \cdot \xi| \leqslant r^{-1} \sigma, |N(t_i) \cdot \xi| \leqslant r^{-2}\},$$

and $c$ is a fixed small constant.
Fix a point $\xi = (\xi_1, \xi_2, s_0) \in \bigcup_i \overline{\Theta}(\sigma, t_i)$ on the plane $\xi_3 = s_0 \approx \sigma^2$. By the strict convexity of the curve $C(s_0, t)$, we know that there are at most two lines $\ell_j$, $j = 1, 2$ that are tangent to $C(s_0, t)$ at $C(s_0, \tau_j^\xi)$ for some $\tau_j^\xi \in [0, 1]$, $j = 1, 2$. Without loss of generality, by possibly extend the domain of $f$, we may assume that there are exactly two such lines.

It now suffices to show that if a line $\ell_k$ that is tangent to $C(s_0, t)$ at $C(s_0, t_{ik})$ and it passes through $B_{r-2}(\xi)$, then $|t_{ik} - \tau_j^\xi| \leq C\sigma^{-1}r^{-1}$ or $|t_{ik} - \tau_2^\xi| \leq C\sigma^{-1}r^{-1}$ for some absolute constant $C$. In fact, it is elementary to check that if $|\xi - \hat{\xi}| \leq r^{-2}$, then $|\tau_j^\xi - \tau_j^\hat{\xi}| \leq C\sigma^{-1}r^{-1}$ for an absolute constant $C$ that only depends on the cone. See Figure A.1.

**ACKNOWLEDGEMENTS**

This project was supported by the National key R&D program of China No. 2022YFA1005700 and 2022YFA1007200. C. Gao was supported by Chinese Postdoc Foundation Grant No. 8206300279. B. Liu was supported by SUSTech start-up counterpart Y01286235. C. Miao was partially supported by NSF China grant No. 11831004. Y. Xi was partially supported by NSF China grant No. 12171424 and the Fundamental Research Funds for the Central Universities 2021QNA3001. The authors would like to thank Ruixiang Zhang for suggesting this problem. The authors would like to thank Larry Guth and Hong Wang for some helpful comments. The authors would like to thank the anonymous referees for their valuable comments and suggestions, which have greatly improved the presentation of this paper.

**JOURNAL INFORMATION**

The *Proceedings of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

**REFERENCES**

1. D. Beltran, J. Hickman, and C. D. Sogge, *Sharp local smoothing estimates for Fourier integral operators*, Proceedings Geometric Aspects of Harmonic Analys, Springer International Publishing, pp. 29–105.
2. D. Beltran, J. Hickman, and C. D. Sogge, Variable coefficient Wolff-type inequalities and sharp local smoothing estimates for wave equations on manifolds, Anal. PDE 13 (2020), no. 2, 403–433.
3. J. Bourgain, $L^p$-estimates for oscillatory integrals in several variables, Geom. Funct. Anal. 1 (1991), no. 4, 321–374.
4. J. Bourgain, Some new estimates on oscillatory integrals, Essays on Fourier analysis in honor of Elias M. Stein (Princeton, NJ, 1991), vol. 42, Princeton Math. Ser., Princeton University Press, Princeton, NJ, 1995, pp. 83–112.
5. J. Bourgain, Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces, Israel J. Math. 193 (2013), no. 1, 441–458.
6. J. Bourgain and C. Demeter, The proof of the $l^2$ decoupling conjecture, Ann. Math. (2) 182 (2015), no. 1, 351–389.
7. A. Córdoba, Geometric Fourier analysis, Ann. Inst. Fourier (Grenoble) 32 (1982), no. 3, vii, 215–226.
8. C. Gao, C. Miao, and J. Yang, Improved variable coefficient square functions and local smoothing of Fourier integral operators.
9. C. Gao, C. Miao, and J. Yang, Function inequality for a class of Fourier integral operators satisfying cinematic curvature conditions, Forum Math, 32 (2020), no. 6, 1375–1394.
10. G. Garrigós and A. Seeger, On plate decompositions of cone multipliers, Proc. Edinb. Math. Soc. (2) 52 (2009), no. 3, 631–651.
11. G. Garrigós and A. Seeger, A mixed norm variant of Wolff's inequality for paraboloids, Harmonic analysis and partial differential equations, vol. 505, Contemp. Math., Amer. Math. Soc., Providence, RI, 2010, pp. 179–197.
12. L. Guth, H. Wang, and R. Zhang, A sharp square function estimate for the cone in $\mathbb{R}^3$, Ann. of Math. (2) 192 (2020), no. 2, 551–581.
13. L. Hörmander, Fourier integral operators. I, Acta Math. 127 (1971), 79–183.
14. A. Iosevich, B. Liu, and Y. Xi, Microlocal decoupling inequalities and the distance problem on Riemannian manifolds, Amer. J. Math. 144 (2022), no. 6, 1601–1639.
15. I. Łaba and T. Wolff, A local smoothing estimate in higher dimensions, vol. 88, 2002, pp. 149–171. Dedicated to the memory of Tom Wolff.
16. J. Lee, A trilinear approach to square function and local smoothing estimates for the wave operator, arXiv:1607.08426.
17. S. Lee and A. Seeger, Lebesgue space estimates for a class of Fourier integral operators associated with wave propagation, Math. Nachr. 286 (2013), no. 7, 743–755.
18. S. Lee and A. Vargas, On the cone multiplier in $\mathbb{R}^3$, J. Funct. Anal. 263 (2012), no. 4, 925–940.
19. W. P. Minicozzi II and C. D. Sogge, Negative results for Nikodym maximal functions and related oscillatory integrals in curved space, Math. Res. Lett. 4 (1997), no. 2–3, 221–237.
20. G. Mockenhaupt, A. Seeger, and C. D. Sogge, Wave front sets, local smoothing and Bourgain's circular maximal theorem, Ann. Math. (2) 136 (1992), no. 1, 207–218.
21. G. Mockenhaupt, A. Seeger, and C. D. Sogge, Local smoothing of Fourier integral operators and Carleson-Sjölin estimates, J. Amer. Math. Soc. 6 (1993), no. 1, 65–130.
22. C. D. Sogge, Propagation of singularities and maximal functions in the plane, Invent. Math. 104 (1991), no. 2, 349–376.
23. C. D. Sogge, Y. Xi, and H. Xu, On instability of the Nikodym maximal function bounds over Riemannian manifolds, J.Geom. Anal. 28 (2018), no. 3, 2886–2901.
24. T. Tao and A. Vargas, A bilinear approach to cone multipliers. II. Applications, Geom. Funct. Anal. 10 (2000), no. 1, 216–258.
25. T. Tao, The Bochner-Riesz conjecture implies the restriction conjecture, Duke Math. J. 96 (1999), no. 2, 363–375.
26. T. Wolff, Local smoothing type estimates on $L^p$ for large $p$, Geom. Funct. Anal. 10 (2000), no. 5, 1237–1288.
27. T. Wolff, A sharp bilinear cone restriction estimate, Ann. of Math. (2) 153 (2001), no. 3, 661–698.