STUDIA MATHEMATICA 230 (2) (2015)

An improved maximal inequality for 2D fractional order Schrödinger operators

by

CHANGXING MIAO (Beijing), JIANWEI YANG (Beijing)
and JIQIANG ZHENG (Nice)

Abstract. The local maximal operator for the Schrödinger operators of order $\alpha > 1$ is shown to be bounded from $H^s(\mathbb{R}^2)$ to L^2 for any $s > 3/8$. This improves the previous result of Sjölin on the regularity of solutions to fractional order Schrödinger equations. Our method is inspired by Bourgain's argument in the case of $\alpha = 2$. The extension from $\alpha = 2$ to general $\alpha > 1$ faces three essential obstacles: the lack of Lee's reduction lemma, the absence of the algebraic structure of the symbol and the inapplicable Galilean transformation in the deduction of the main theorem. We get around these difficulties by establishing a new reduction lemma and analyzing all the possibilities in using the separation of the segments to obtain the analogous bilinear L^2-estimates. To compensate for the absence of Galilean invariance, we resort to Taylor's expansion for the phase function. The Bourgain–Guth inequality (2011) is also generalized to dominate the solution of fractional order Schrödinger equations.

Contents

2010 *Mathematics Subject Classification*: 42B25, 35Q41.
Key words and phrases: local maximal inequality, multilinear restriction estimate, induction on scales, localization argument, oscillatory integral operator.
Received 20 January 2014; revised 8 December 2015.
Published online 27 January 2016.

DOI: 10.4064/sm8190-12-2015

1. Introduction and the main result. For $\alpha > 1$, we define the αth order Schrödinger evolution operator by

$$U(t)f(x) := (2\pi)^{-d/2} \int_{\mathbb{R}^d} e^{i[x\cdot\xi + t|\xi|^\alpha]} \widehat{f}(\xi)\, d\xi,$$

and consider the local maximal inequality

$$\Big\| \sup_{0<t<1} |U(t)f| \Big\|_{L^2(B(0,1))} \le C_{\alpha,s,d} \|f\|_{H^s(\mathbb{R}^d)}, \qquad s \in \mathbb{R}, \tag{1.1}$$

where $H^s(\mathbb{R}^d)$ is the usual inhomogeneous Sobolev space defined via the Fourier transform, and $B(0,1)$ is the unit ball centered at the origin. As a consequence of (1.1), we have the following pointwise convergence (from a standard process of approximation and Fatou's lemma):

$$\lim_{t\to 0} U(t)f(x) = f(x), \qquad \text{a.e. } x,\ \forall f \in H^s(\mathbb{R}^d).$$

If $s > d/2$, we obtain (1.1) immediately from Sobolev's embedding. Thus, it is natural to ask what the minimal s is to ensure (1.1).

First of all, let us briefly review the known results about (1.1) in the case when $\alpha = 2$. This problem was raised by Carleson [Cr], who proved $s \ge 1/4$ in the 1D case. This result was shown to be optimal by Dahlberg and Kenig [DK]. The higher dimensional cases of (1.1) were established independently by Sjölin [S] and Vega [V] with $s > 1/2$. In particular, the result was strengthened to $s = 1/2$ by Sjölin [S] when $d = 2$. Meanwhile, Vega [V] demonstrated that (1.1) fails in any dimension if $s < 1/4$. It is then conjectured that $s \ge 1/4$ is sufficient for all dimensions.

A breakthrough was achieved by Bourgain [B1, B2], who showed that (1.1) holds with $\alpha = 2$ for some $s < 1/2$ when $d = 2$. His work was continued and improved by many authors, including Moyua, Vargas and Vega [MVV], Tao and Vargas [TV1, TV2], Lee [L] and Shao [Sh]; the best result hitherto is $s > 3/8$ due to Lee [L].

Previous to [B3], the results for $d \ge 3$ had $s > 1/2$, and $s \ge 1/4$ was still believed to be the correct condition for (1.1) in every dimension. The study of this problem stagnated for several years until the recent work [B3], where the 1/2-barrier was broken for *all* dimensions. More surprisingly, Bourgain also discovered some counterexamples to the widely believed 1/4-threshold. Specifically, he showed that $s \ge 1/2 - 1/d$ is necessary for (1.1) if $d \ge 4$. These examples originated essentially from an observation on arithmetical progressions.

Now, let us turn to the fractional order case. Sjölin [S] proved (1.1) with $s = 1/2$, $d = 2$ for *all* $\alpha > 1$. His proof involves a TT^* argument, which reduces the problem to a dispersive estimate of a specific oscillatory integral. After localizing the integral to high and low frequencies, the author employed a classical result by Miyachi [M] to treat the high frequency part. The other part is estimated by means of the inequality

$$\int_{\mathbb{R}^2} |y|^{-1}(1+|x-y|^4)^{-1}\,dy \le c|x|^{-1},$$

where the decay rate on the right hand side cannot be improved. A crucial fact which Sjölin's proof relied heavily on is that the factor $|t(x)-t(y)|^{1/\alpha}$ can be canceled at the end of the computation exactly for $s = 1/2$. This is usually referred to as the Kolmogorov–Seliverstov–Plessner method (see [Cr] and [S] for more details). For these reasons, it seems difficult to pursue Sjölin's original approach to improve this result. In this paper, we prove

Theorem 1.1. *If $d = 2$ and $\alpha > 1$, then* (1.1) *is valid for all $s > 3/8$.*

Remark 1.2. As noted in [B3], one may modify the method to treat a general multiplier operator $\Phi(D)$ having the property that for some constants $C, c > 0$ and all multi-indices γ,

$$|\partial^\gamma \Phi(\xi)| \le C|\xi|^{2-|\gamma|}, \qquad |\nabla\Phi(\xi)| \ge c\,|\xi|.$$

However, this does not concern the fractional order case.

As a consequence, we get some improvement on the higher dimensional results by using the scheme of induction on dimensions formulated in [B3].

Corollary 1.3. *For $d \ge 3$ and $\alpha > 1$, there exists a θ_d such that* (1.1) *is valid for all $s > \theta_d$ with*

$$\theta_d = \frac{1}{2} - \sigma\left(\frac{1}{2} - \theta_{d-1}\right)$$

for some $\sigma \in (0, 1/2)$. In particular, $\theta_d < 1/2$ for every $d \ge 2$ since $\theta_2 < 1/2$.

Remark 1.4. This improves [S, Theorem 2] in higher dimensions. Noting that the induction argument in [B3] is independent of the order α, we may apply it verbatim to obtain Corollary 1.3.

As in [B3], the proof is based on the multilinear restriction theorem of [BCT]. To achieve this, an important observation introduced by Bourgain and Guth [BG] is that up to an R^ε factor and a well behaved remainder, one can successfully control the free solution of the Schrödinger equation with a sum of triple products fulfilling the transversality condition for which the multilinear restriction estimate can be used. Roughly speaking, one gains structure by losing R^ε, but this is acceptable if we do not intend to solve the end-point problem. These triple products, which we will call type I terms in

Section 5, are generated by iteration with respect to scales. As a result, they are used to collect the contributions obtained at different scales. In this sense, this is also reminiscent of Wolff's induction on scale argument in [W] ([1]). In this paper, we call this robust device Bourgain–Guth's inequality.

Let us take this opportunity to try to moderately clarify several points in Bourgain's argument, keeping the notation of [BG] and [B3]. Of course, a complete clarification of Bourgain's treatment will be far beyond our reach. Instead, we focus only on the points which are directly relevant to this paper.

First, a crucial input is the Bourgain–Guth inequality for oscillatory integrals from [BG]. It collects the contributions of the transversal triple products from all dyadic scales between $R^{-1/2}$ and 1, so that we can use the multilinear restriction theorem of [BCT] to evaluate the contributions at each scale. Since we are dealing with dyadic scales in $(R^{-1/2}, 1)$, we can safely consider items from all scales by taking an ℓ^2 sum, losing at most a factor of $\log R$. To obtain this inequality, Bourgain and Guth [BG] tactically used a "local constant trick" according to the following principle. By writing the oscillatory integral as trigonometric sums $T_\alpha f(x)$ with variable coefficients, one may regard $T_\alpha f(x)$ as constant on each ball of radius K thanks to the uncertainty principle, where these "constants" certainly depend on the position of the ball. This heuristic point is justified by convolving $T_\alpha f$ with suitable bump functions. However, further manipulations, especially the iterative process, are awkward to write out explicitly. Instead, one prefers a formal calculation for brevity and clarity. Based on this observation, one may insert/extract the factor $T_\alpha f(x)$ into/from an integral over a ball of radius K, or more generally over a tile of suitable shape and size. All this can be justified by invoking the uncertainty principle.

This simple and important observation is very efficient in simplifying various explicit calculations, so that the Bourgain–Guth inequality can be established in [BG] by iteration.

Let us say more about the establishment of Bourgain–Guth's inequality before turning to the argument for the Schrödinger maximal function. The brilliant novelty in [BG], which we will follow in Section 5, is embodied in the way of using Bonnet–Carbery–Tao's multiplier restriction theorem. The idea might be roughly described as follows: after writing $Tf(x)$ as a variable coefficient trigonometric sum, one may estimate it for each $x \in B_R$ in three different manners, where only a small portion of the members in $\{T_\alpha f(x)\}_\alpha$ would dominate the behavior of $Tf(x)$. As can be seen in [BG] and Section 5, these members correspond respectively to three different scenarios which cover all the possibilities for a particular $x \in B_R(0)$. According

([1]) It is thus interesting to consider how to combine these two important ideas together to improve the argument in this work.

to [BG], they are called *non-coplanar interaction, non-transverse interaction* and *transverse coplanar interaction.* We refer to Section 5 for more details about this classification.

Now we turn to Bourgain's treatment of the Schrödinger maximal function. The idea is that by using Bourgain–Guth's inequality, one is reduced to controlling each item in the ℓ^2 summation with the desired bound. To achieve this, one tiles $\mathbb{R}^3$ with translates of polar sets of the cap τ, which contains a triple of transversal subcaps τ_1, τ_2, τ_3. This provides a decomposition of $B_R \subset \mathbb{R}^3$. Invoking the local constancy principle, one may raise and lower the moment exponents on each tile so that the trilinear restriction of [BCT] can be used. During this calculation, Galilean transform is employed to shift the center of the square where the frequency is localized to the center. Although we have compressed Bourgain's argument into as few words as possible, it is far more difficult and subtle in concrete manipulations as in [B3, BG]. We confine ourselves to this brief description of Bourgain's approach and turn to our situation below.

To use the strategy of [B3], we need to retrieve the Bourgain–Guth inequality for general $\alpha > 1$. Although this inequality is invented in [BG] for $\alpha = 2$, it is rather non-trivial to generalize it to $\alpha > 1$, as will be seen in Section 5. One of the obstructions is the absence of the algebraic structure of $|\xi|^\alpha$ when α is not an integer. This fact leads to the differences of our argument from [B3] and [BG] in almost every aspect, especially in the proof of the bilinear L^2-estimate in Subsection 5.2 where we introduce a new argument.

Besides the reestablishment of Bourgain–Guth's inequality, we need a fractional order version of Lee's reduction lemma of [L] for general $\alpha > 1$. In Section 4, we establish this result using a different method. This extends the result of [L] to a more general setting. We will use the method of stationary phase in the spirit of [Sh]. However, to justify the proof, we involve a localization argument which eliminates Schwartz tails by losing R^ε. To be more precise, we separate the Poisson summation into relatively large and small scales, where either the rapid decrease of Schwartz functions or the stationary phase argument can be used to handle the error terms. This principle is also used in the proof of the main theorem. The essence of this argument is exploiting the orthogonality in "phase space" via stationary phase and the Poisson summation formula. In doing so, one only needs to afford an R^ε loss, but one may sum the pieces that are well-estimated efficiently. See Sections 3 and 4 for more details.

To end this section, let us say a few words about the potential of Bourgain–Guth's approach to oscillatory integrals. In harmonic analysis, one of the most important principles is that structures are favorable conditions to help us use deep results. For instance, Whitney's decomposition was employed in [TV1, TV2, TVV] to generate the transversality conditions for the use of bilinear estimates. On the other hand, the proof of Bourgain and Guth's inequality enlightened a new approach to generating structures by means of logical classification, i.e. exploiting the intrinsic structures implicitly involved in the summation of large numbers of elements creatively using logical division. The idea is fairly new and the argument is really a *tour de force*, bringing in ideas and techniques from combinatorics, as will be seen in Section 5. We believe this approach is very promising to get improvements on the open questions in classical harmonic analysis. In particular, the result in this paper might be improved further by refining this method.

This paper is organized as follows. In Section 2, we introduce some preliminaries and basic lemmas. In Section 3, we prove the main result. Section 4 is devoted to the proof of Lemma 2.11, and Section 5 to the proof of Lemma 2.4.

2. Preliminaries. This section includes the list of the frequently used notation, the statement of a crucial lemma which plays the key role in deducing the main result, as well as the primary reduction for the proof of Theorem 1.1.

2.1. Notation. Throughout this paper the following notation will be used:

- ⋄ $\Omega = [-1/2, 1/2] \times [-1/2, 1/2]$.
- ⋄ $[r]$ is the greatest integer not exceeding r.
- ⋄ If Ω is a subset of $\mathbb{R}^d$, we define $\Omega^c = \mathbb{R}^d \setminus \Omega$.
- ⋄ χ_Ω denotes the characteristic function of $\Omega \subset \mathbb{R}^d$.
- ⋄ If ξ is a vector in $\mathbb{R}^d$, we define $\overline{\xi} = \xi/|\xi|$.
- ⋄ $\mathcal{I} = \{\xi \in \mathbb{R}^d \mid 1/2 \le |\xi| \le 2\}$ and, except in Lemma 2.11 below, we always assume $d = 2$.
- ⋄ For f a measurable function and $a \in \mathbb{R}^d$, we define $\tau_a f(x) = f(x-a)$.
- ⋄ $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz class on $\mathbb{R}^d$, and $\mathcal{S}'(\mathbb{R}^d)$ the space of tempered distributions.
- ⋄ $\mathcal{F}_{x\to\xi} f$ and $\widehat{f}(\xi)$ denote the Fourier transform of a tempered distribution $f(x)$ and
$$\widehat{f}(\xi) = (2\pi)^{-d/2} \int_{\mathbb{R}^d} f(x) e^{-ix\cdot\xi}\, dx.$$
- ⋄ $B(a, K)$ or $B_{a,K}$ is the ball in $\mathbb{R}^d$ centered at a of radius K.
- ⋄ If B is a convex body in $\mathbb{R}^d$ and $\lambda > 0$, we use λB to denote the convex set having the same center with B but enlarged in size by λ.

- ⋄ C stands for a constant which might be different from line to line and $c \ll C$ means c is far less than C. This is clear in the context.
- ⋄ $A \lesssim B$ means $A \le CB$ for some constant C, and $A \simeq B$ means both $A \lesssim B$ and $B \lesssim A$.
- ⋄ $A \lesssim_{\eta,\zeta,\dots} B$ means that there is a constant $C = C(\eta, \zeta, \dots)$ such that $A \le CB$.
- ⋄ $\zeta = \mathcal{O}_\alpha(\eta)$ means $\zeta \lesssim_\alpha \eta$.

2.2. Caps, tiles and the Bourgain–Guth inequality. Now we introduce some terminology. Let $R \gg 2^{5\alpha} > 1$ and $1/\sqrt{R} < \delta < 1$. Partition $\mathbb{R}^2$ into $\bigcup_\tau \Omega_\tau$ where Ω_τ is a $\delta \times \delta$-square centered at $\xi_\tau \in \delta\mathbb{Z}^2$ with edges parallel to the coordinate axes. Let $\vec{n}_\tau$ be the exterior unit normal to the immersed surface $(\xi, |\xi|^\alpha)$ at the point $(\xi_\tau, |\xi_\tau|^\alpha)$. We define

$$\Pi_\tau^\delta = (\xi_\tau, |\xi_\tau|^\alpha) + \{z \in \mathbb{R}^3 \mid |\langle z, \vec{n}_\tau\rangle| \le \delta^\alpha\},$$
$$\mathcal{C}_\tau = \Pi_\tau^\delta \cap \{z \in \mathbb{R}^3 \mid z = (z_1, z_2, z_3),\ (z_1, z_2) \in \Omega_\tau\}.$$

Obviously, $\mathcal{C}_\tau$ is a parallelepiped with dimensions $\sim \delta,\ \delta,\ \delta^\alpha$.

Definition 2.1. The parallelepiped $\mathcal{C}_\tau$ is called a δ-*cap* associated to Ω_τ.

Definition 2.2. The *polar set* of $\mathcal{C}_\tau$ is defined as

$$\mathcal{C}_\tau^* = \{z \in \mathbb{R}^3 \mid |\langle z, w\rangle| \le 1,\ \forall w \in \mathcal{C}_\tau - (\xi_\tau, |\xi_\tau|^\alpha)\}.$$

It is easy to see that $\mathcal{C}_\tau^*$ is essentially a $1/\delta \times 1/\delta \times 1/\delta^\alpha$-rectangle centered at the origin, with the longest side in the direction of $\vec{n}_\tau$. Moreover, we may tile $\mathbb{R}^3$ with translations of $\mathcal{C}_\tau^*$. This decomposes $\mathbb{R}^3$ naturally into the union of essentially disjoint $\mathcal{C}_\tau^*$-boxes. We call this decomposition a *tiling of* $\mathbb{R}^3$ *with* $\mathcal{C}_\tau^*$*-boxes.*

Define an oscillatory integral by

$$Tf(x) = \int_{\mathcal{I}} e^{i[x_1\xi_1 + x_2\xi_2 + x_3|\xi|^\alpha]} \widehat{f}(\xi)\, d\xi,$$

where $x = (x_1, x_2, x_3) \in \mathbb{R}^3$ and $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$. Setting $x' = (x_1, x_2)$ and regarding x_3 as the temporal variable $t = x_3$, we have

$$U(t)f(x') = Tf(x). \tag{2.1}$$

Remark 2.3. The general d-dimensional counterpart of (2.1) is defined in the same way, with x_{d+1} in place of x_3, and (x_1, x_2) and (ξ_1, ξ_2) replaced by $(x_1, \dots, x_d)$ and $(\xi_1, \dots, \xi_d)$. This will only be used in Lemma 2.11 below, which is proved for general dimensions.

Now we can state Bourgain–Guth's inequality which will be used to control the oscillatory integral $Tf(x)$ in terms of $\{Tf_\tau\}_\tau$, where $\widehat{f_\tau}$ is supported in a much smaller region $\Omega_\tau \subset \mathcal{I}$.

Lemma 2.4. *If* $\operatorname{supp}\widehat{f} \subset \mathcal{I}$ *and* $1 \ll K \ll R$, *then for any* $\varepsilon > 0$ *we have the following estimate on the cylinder* $B(0,R) \times [0,R] \subset \mathbb{R}^{2+1}$:

$$|Tf(x)| \lesssim R^{\varepsilon} \max_{1/K \geq \delta > 1/\sqrt{R}} \max_{\mathcal{E}_\delta} \Big[\sum_{\Omega_\tau \in \mathcal{E}_\delta} \Big(\psi_\tau \prod_{j=1}^{3} |Tf_{\tau_j}|^{1/3} \Big)^2 \Big]^{1/2} \tag{2.2}$$

$$+ R^{\varepsilon} \max_{\mathcal{E}_{1/\sqrt{R}}} \Big[\sum_{\Omega_\tau \in \mathcal{E}_{1/\sqrt{R}}} (\psi_\tau |Tf_\tau|)^2 \Big]^{1/2}, \tag{2.3}$$

where $\widehat{f}_{\tau_j}$ *is supported in* Ω_{τ_j} *for* $j = 1,2,3$, *and*

- $\mathcal{E}_\delta$ *consists of at most* $(1/\delta)^{1+\varepsilon}$ *disjoint* $\delta \times \delta$*-squares* Ω_τ;
- $\{\Omega_{\tau_j}\}_{j=1}^3$ *is a triple of non-collinear* $\delta/K \times \delta/K$*-squares inside* Ω_τ;
- *for each* τ, ψ_τ *is a non-negative function on* $\mathbb{R}^3$ *which is constant on unit cubes centered at points of* $\mathbb{Z}^3$ *and satisfies*

$$\frac{1}{|B|} \int_B \psi_\tau(x)^4 \, dx \lesssim R^{\varepsilon} \tag{2.4}$$

for all B *in a tiling of* $\mathbb{R}^3$ *with* $\mathcal{C}_\tau^*$*-boxes.*

Remark 2.5. A triplet $(\Omega_{\tau_1}, \Omega_{\tau_2}, \Omega_{\tau_3})$ in Ω_τ is said to be *non-collinear* if

$$|\xi_{\tau_1} - \xi_{\tau_2}| \geq |\xi_{\tau_1} - \xi_{\tau_3}| \geq \operatorname{dist}(\xi_{\tau_3}, \ell(\xi_{\tau_1}, \xi_{\tau_2})) > 10^3 \alpha 2^{\alpha}/K, \tag{2.5}$$

where ξ_j is the center of Ω_{τ_j} for $j = 1,2,3$ and $\ell(\xi_{\tau_1}, \xi_{\tau_2})$ is the line through ξ_{τ_1} and ξ_{τ_2}. Consequently, the caps $\mathcal{C}_{\tau_1}, \mathcal{C}_{\tau_2}, \mathcal{C}_{\tau_3}$ are transversal, that is, the exterior normal vectors to these three caps are linearly independent, uniformly with respect to the variables belonging to Ω_{τ_j} for $j = 1,2,3$. We refer to [BCT] for the precise description of the transversality condition (see also Section 3). This condition is required for the multilinear restriction estimate established in [BCT], and frequently used in [B3, BG].

Remark 2.6. This lemma is established in the spirit of Bourgain and Guth, but it differs from [BG] in two respects. First, the non-collinearity condition is reformulated in (2.5) to handle general $\alpha > 1$. Second, the scales of the caps and dual caps depend on α already. The absence of the algebraic structure of the symbol $|\xi|^\alpha$ for general $\alpha > 1$ will lead to some difficulties in the deduction of (2.2)–(2.3) as well as the application of this inequality in the proof of Theorem 1.1. These obstacles make our argument more complicated than in [B3].

Remark 2.7. If $\widehat{f}_\tau$ is supported in a square Ω_τ of size δ, $|Tf_\tau(x)|$ can be regarded essentially as a constant on each $\mathcal{C}_\tau^*$-box. We call this the *local constancy property*, indicated in [BG].

This inequality allows us to gain the transversality condition in each term of the summation (2.2) by losing only an R^{ε} factor. This is favorable especially in proving some non-endpoint estimates. We also point out that the precise cardinality of $\mathcal{E}_\delta$ will not be used in the proof of Theorem 1.1. We will prove Lemma 2.4 in Section 5.

2.3. A primary reduction of the problem. By Littlewood–Paley's theory, Sobolev embedding and Hölder's inequality, Theorem 1.1 amounts to showing that for any $\varepsilon > 0$, there is a $C_{\alpha,\varepsilon}$ such that

$$\Big\| \sup_{0<x_3<1} |Tf(\cdot, x_3)| \Big\|_{L^2(B(0,1))} \le C_{\alpha,\varepsilon} R^{3/8+\varepsilon} \|f\|_2 \tag{2.6}$$

for $\widehat{f}(\xi)$ supported in $\{\xi \in \mathbb{R}^2 \mid R/2 \le |\xi| \le 2R\}$ with R large enough.

After rescaling, (2.6) reduces to

$$\Big\| \sup_{0<x_3<R^{\alpha}} |Tf(\cdot, x_3)| \Big\|_{L^2(B(0,R))} \le C_{\alpha,\varepsilon} R^{3/8+\varepsilon} \|f\|_2, \qquad \operatorname{supp} \widehat{f} \subset \mathcal{I}. \tag{2.7}$$

When $\alpha = 2$, it was observed by Lee [L] that to get (2.7), it suffices to prove it with the supremum taken only for $0 < x_3 < R$. This simplifies the problem significantly so that the result of $s > 3/8$ can be deduced for $d = 2$. This reduction is also necessary for the argument in [B3]. We extend this result to *all* $\alpha > 1$ by proving the following lemma.

Lemma 2.8. *Suppose that for any $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that*

$$\Big\| \sup_{0<x_3<R} |Tf(\cdot, x_3)| \Big\|_{L^2(B(0,R))} \le C_{\varepsilon} R^{3/8+\varepsilon} \|f\|_2 \tag{2.8}$$

for R sufficiently large and $\operatorname{supp} \widehat{f} \subset \{\xi \in \mathbb{R}^2 \mid 3/8 \le |\xi| \le 17/8\}$. *Then* (2.7) *holds.*

Remark 2.9. Intuitively, one might expect that the x_3-interval over which the supremum in (2.8) is taken should be $(0, R^{\alpha/2})$. Although this can be deduced easily by modifying our argument slightly, we will lose more derivatives in Theorem 1.1 if we use (2.8) with $0 < x_3 < R^{\alpha/2}$. The loss of derivatives forces the s in (1.1) to rely heavily on α, and this will confine α to a small range in order to improve Sjölin's result. However, our result can be strengthened so that s is independent of α thanks to Lemma 2.8. We point out that the global maximal inequality is α-dependent. See [R] for details.

Remark 2.10. Heuristically, the idea behind this lemma can be expressed in terms of propagation speed. If the frequency of the initial data f is localized in $\mathcal{I}$, then the propagation speed of $U(t)f$ can be morally regarded as finite. Suppose R is large enough so that f is mainly concentrated in $B(0, R)$. If one waits at a position in $B(0, R)$ for the maximal amplitude of the solution to occur during the time period $0 < t < R^{\alpha}$, then by the finite speed of propagation, this maximal amplitude can be expected to happen before time R. This heuristic intuition is justified by the following lemma.

LEMMA 2.11. *Let* $\operatorname{supp}\widehat{f}(\xi)\subset\mathcal{I}$ *and* $j=0,1,\dots,[R^{\alpha-1}]$, $t_j=jR$. *Set* $I_j=[t_j,t_{j+1})$ *for* $j<[R^{\alpha-1}]$ *and* $I_{[R^{\alpha-1}]}=[t_{[R^{\alpha-1}]},R^\alpha)$. *Denote* $x=(x',x_{d+1})$ *where* $x'=(x_1,\dots,x_d)$ *and take* $\varphi\in C_0^\infty(B(0,2R))$ *such that* $\varphi(x')=1$ *on* $B(0,R)$. *Then for any* $\varepsilon>0$, *there is a* $C_{\alpha,\varepsilon}>0$ *and a family* $\{f_j\}_j$ *of functions satisfying*

$$\operatorname{supp}\widehat{f_j}\subset\{\xi\in\mathbb{R}^d\mid 1/2-1/R\le|\xi|\le 2+1/R\}=:\mathcal{I}_{1/R},$$

such that for $x_{d+1}\in I_j$,

$$\varphi(x')Tf(x)=\varphi(x')\chi_{I_j}(x_{d+1})Tf_j(x',x_{d+1}-t_j)+\mathcal{O}_{\alpha,\varepsilon}(R^{-99d}\|f\|_2), \tag{2.9}$$

or equivalently, viewing $x_{d+1}=t\in I_j$,

$$\varphi(x')U(t)f(x')=\varphi(x')\chi_{I_j}(t)U(t-t_j)f_j(x')+\mathcal{O}_{\alpha,\varepsilon}(R^{-99d}\|f\|_2). \tag{2.10}$$

Moreover, there exists a positive constant $c_d>0$ *such that*

$$\Big\|\Big(\sum_{j=0}^{[R^{\alpha-1}]}|f_j|^2\Big)^{1/2}\Big\|_2\le C_{\alpha,\varepsilon}R^{\varepsilon c_d}\|f\|_2. \tag{2.11}$$

To prove this lemma, we introduce a localization argument which allows us to regard a Schwartz function with compact Fourier frequencies as a smooth cut-off function by losing R^ε. That is why we have to lose $R^{\varepsilon c_d}$ in (2.11), but this is suitable for our purposes.

REMARK 2.12. In our proof, f_j is constructed by localizing $Tf(x,t_j)$ with Schwartz functions. This leads to a $1/R$-enlargement of $\mathcal{I}$ in the frequency space, but this does not affect the use of this lemma.

We end this section by showing that Lemma 2.8 follows from Lemma 2.11.

Proof of Lemma 2.8. In view of (2.10), for $d=2$ we have

$$\varphi(x')|Tf(x',x_3)|\lesssim_{\alpha,\varepsilon}\varphi(x')\sum_{j=0}^{[R^{\alpha-1}]}|\chi_{I_j}(x_3)Tf_j(x',x_3-t_j)|+R^{-198}\|f\|_2.$$

Choosing R large enough and neglecting $R^{-198}\|f\|_2$, we obtain

$$\sup_{0<x_3<R^\alpha}|\varphi(x')Tf(x)|^2\lesssim_{\alpha,\varepsilon}\sum_{j=0}^{[R^{\alpha-1}]}\sup_{0<x_3-t_j<R}|\varphi(x')Tf_j(x',x_3-t_j)|^2.$$

Integrating both sides of the above inequality on $B(0,R)$, we may estimate the left side of (2.7) by

$$\Big(\sum_{j=0}^{[R^{\alpha-1}]}\Big\|\sup_{0<x_3-t_j<R}|\varphi(x')Tf_j(x',x_3-t_j)|\Big\|_2^2\Big)^{1/2}.$$

Using (2.8) and (2.11), we obtain

$$(2.7)\lesssim_{\alpha,\varepsilon} R^{3/8+\varepsilon}\|f\|_2. \ \blacksquare$$

Remark 2.13. In proving (2.8), we always fix an $\varepsilon>0$ first and then take R large, which may depend on ε, α and $\|f\|_2$. This allows us to eliminate as many error terms as possible by repeatedly using the localization argument.

Remark 2.14. In fact, the original proof of [L, Lemma 2.1] for the classical Schrödinger equation works for the generalized case as well. Moreover, the proof of [LR, Lemma 2.1] can be used to get rid of the ε loss.

3. Proof of the main result. Now we are in a position to prove Theorem 1.1. For any fixed $\varepsilon>0$, we normalize $\|f\|_2=1$. In light of Lemmas 2.4 and 2.8, (2.8) amounts to obtaining the following two estimates for R large enough:

$$\sum_{\substack{1/\sqrt{R}<\delta<1/K\\ \delta\,\text{dyadic}}}\Big[\sum_{\Omega_\tau:\,\delta\times\delta}\Big\|\psi_\tau\prod_{j=1}^{3}|Tf_{\tau_j}|^{1/3}\Big\|^2_{L^2(|x'|<R)L^\infty(|x_3|<R)}\Big]^{1/2}\lesssim R^{3/8+\varepsilon}, \tag{3.1}$$

$$\Big[\sum_{\Omega_\tau:\,1/\sqrt{R}\times1/\sqrt{R}}\big\|\psi_\tau|Tf_\tau|\big\|^2_{L^2(|x'|<R)L^\infty(|x_3|<R)}\Big]^{1/2}\lesssim R^{3/8+\varepsilon}, \tag{3.2}$$

where $x'=(x_1,x_2)$ and $\Omega_\tau:\delta\times\delta$ refers to the partition of $\mathcal{I}_{1/8}$ into the union of $\delta\times\delta$-squares.

By orthogonality, it suffices to prove

$$\int_{|x'|<R}\sup_{|x_3|<R}\Big|\psi_\tau\prod_{j=1}^{3}|Tf_{\tau_j}|^{1/3}\Big|^2(x',x_3)\,dx'\lesssim_\varepsilon R^{3/4+2\varepsilon}\|f_\tau\|_2^2, \tag{3.3}$$

$$\int_{|x'|<R}\sup_{|x_3|<R}(\psi_\tau|Tf_\tau|)^2(x',x_3)\,dx'\lesssim_\varepsilon R^{3/4+2\varepsilon}\|f_\tau\|_2^2, \tag{3.4}$$

where f_τ is defined as $\widehat{f_\tau}=\widehat{f}\chi_{\Omega_\tau}$.

3.1. The proof of (3.3). For brevity, we set $G_{\tau_1,\tau_2,\tau_3}=\prod_{j=1}^{3}|Tf_{\tau_j}|^{1/3}$ in (3.3), where (τ_1,τ_2,τ_3) corresponds to the squares $(\Omega_{\tau_1},\Omega_{\tau_2},\Omega_{\tau_3})$ with the properties of Lemma 2.4. Let ξ^τ be the center of Ω_τ. Then we may assume $\xi^\tau=(0,|\xi^\tau|)$ due to the invariance of (3.3) under orthogonal transformations. Let $\mathcal{C}_\tau$ be the δ-cap associated to Ω_τ, and $\mathcal{C}^*_\tau$ be the polar set of $\mathcal{C}_\tau$.

After tiling $B(0,R)\times[0,R]\subset\mathbb{R}^3$ with $\mathcal{C}^*_\tau$-boxes, we have

$$B(0,R)\times[0,R]\subset\bigcup_{j,k}B_{j,k}$$

where $B_{j,k}$ is a $\mathcal{C}^*_\tau$-box labeled by j and k, with j corresponding to the horizontal translation and k to the vertical one (see Figure 1). Adopting the notation of [B3], we denote the projection of each $B_{j,k}$ to the (x_1,x_2)-variables by $I_j=\pi_{x'}(B_{j,k})$. Let $\mathcal{P}_{x'}$ be the plane through the point $(x',0)$ and perpendicular to the x_2-direction. We define $J_k^{x'}=\pi_{x_3}(B_{j,k}\cap\mathcal{P}_{x'})$. Then $|J_k^{x'}|\sim 1/\delta$ for all x' and k . For I_j, it is easy to see that the length of the side in the direction of x_1 is approximately $1/\delta$, and the side in the x_2-direction has length $1/\delta^\alpha$.

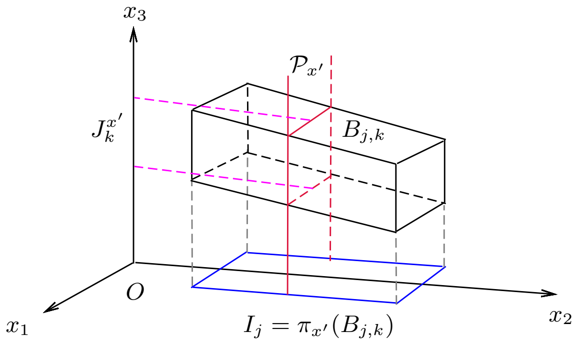

Fig. 1. The $\mathcal{C}^*_\tau$-box $B_{j,k}$

By Hölder's inequality and $\ell^3\hookrightarrow\ell^\infty$, we have

$$\begin{aligned}(3.5)\qquad &\sum_j\|\psi_\tau G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2_{x'}(I_j)L^\infty(|x_3|<R)}\\ &\le\delta^{-(1+\alpha)/3}\sum_j\|\psi_\tau G_{\tau_1,\tau_2,\tau_3}\|^2_{L^3_{x'}(I_j)L^\infty(|x_3|<R)}\\ &\le\delta^{-(1+\alpha)/3}\sum_j\max_k\|\psi_\tau G_{\tau_1,\tau_2,\tau_3}\|^2_{L^3_{x'}(I_j)L^\infty_{x_3}(J_k^{x'})}\\ &\le\delta^{-(1+\alpha)/3}\sum_j\Big[\sum_k\|\psi_\tau G_{\tau_1,\tau_2,\tau_3}\|^3_{L^3_{x'}L^\infty_{x_3}(B_{j,k})}\Big]^{2/3}.\end{aligned}$$

Using (2.4) and Remark 2.7, we obtain

$$\begin{aligned}&\|\psi_\tau G_{\tau_1,\tau_2,\tau_3}\|_{L^3_{x'}L^\infty_{x_3}(B_{j,k})}\\ &\lesssim\|\psi_\tau\|_{L^3_{x'}L^\infty_{x_3}(B_{j,k})}\bigg(\frac{1}{|I_j|}\int_{I_j}\bigg(\frac{1}{|J_k^{x'}|}\int_{J_k^{x'}}|G_{\tau_1,\tau_2,\tau_3}(x',x_3)|^3\,dx_3\bigg)^{2/3}dx'\bigg)^{1/2}\\ &\lesssim|I_j|^{-1/2}|J_k|^{-1/3}\|\psi_\tau\|_{L^3_{x'}L^4_{x_3}(B_{j,k})}\|G_{\tau_1,\tau_2,\tau_3}\|_{L^2_{x'}L^3_{x_3}(B_{j,k})}\end{aligned}$$

$$
\begin{aligned}
&\lesssim |I_j|^{1/3-1/4-1/2}|J_k|^{-1/3}\|\psi_\tau\|_{L^4(B_{j,k})}\|G_{\tau_1,\tau_2,\tau_3}\|_{L^2_{x'},L^3_{x_3}(B_{j,k})}\\
&\lesssim \delta^{-(1+\alpha)(1/3-3/4)+1/3}R^{\varepsilon}|B_{j,k}|^{1/4}\|G_{\tau_1,\tau_2,\tau_3}\|_{L^2_{x'},L^3_{x_3}(B_{j,k})}\\
&\lesssim R^{\varepsilon}\delta^{1/4+\alpha/6}\|G_{\tau_1,\tau_2,\tau_3}\|_{L^2_{x'},L^3_{x_3}(B_{j,k})},
\end{aligned}
$$

where, in the second inequality, we have used the fact that ψ_τ is constant on unit cubes. This allows us to control the L^∞ norm of ψ_τ with respect to the x_3-variable on $J_k^{x'}$ by the L^4 norm.

Plugging this into (3.5), by Minkowski's inequality we get

$$
\begin{aligned}
|(3.5)| &\lesssim \delta^{-(1+\alpha)/3}\sum_j\Big[\sum_k R^{3\varepsilon}\delta^{3/4+\alpha/2}\|G_{\tau_1,\tau_2,\tau_3}\|^3_{L^2_{x'},L^3_{x_3}(B_{j,k})}\Big]^{2/3}\\
&\lesssim R^{2\varepsilon}\delta^{1/6}\sum_j\int_{I_j}\Big(\sum_k\int_{J_k^{x'}}|G_{\tau_1,\tau_2,\tau_3}|^3(x',x_3)\,dx_3\Big)^{2/3}dx'\\
&\lesssim R^{2\varepsilon}\delta^{1/6}\|G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2(|x'|<R)L^3(|x_3|<R)}. \qquad (3.6)
\end{aligned}
$$

REMARK 3.1. In (3.6), α is absent from the exponent of δ and R. However, the role of α is encoded in the estimation of

$$
\|G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2(|x'|<R)L^3(|x_3|<R)}. \tag{3.7}
$$

To handle this expression, Bourgain used a Galilean transformation to shift the center of the domain for the integral Tf_ν to the origin when $\alpha=2$. We cannot directly use this due to the absence of the algebraic structure of $|\xi|^\alpha$ for general $\alpha>1$. To adapt his strategy, we get around this obstacle by using Taylor's expansion. We also use a localization argument as in the proof of Lemma 2.11.

It remains to evaluate (3.7). Let us introduce some notation. Define

$$
\mathcal{T}_{\delta,\tau}h(x_1,x_2,x_3)=\int_{\mathbb{R}^2}e^{i[x_1\eta_1+x_2\eta_2+x_3\Phi(\xi^\tau,\lambda_0,\alpha,\delta,\eta)]}\chi(\eta)\widehat{h}(\eta)\,d\eta,
$$

where χ is a smooth function adapted to the unit square Ω and

$$
\Phi(\xi^\tau,\lambda_0,\alpha,\delta,\eta)=\frac{\alpha}{2}|\xi^\tau|^{\alpha-2}(\eta_1^2+(\alpha-1)\eta_2^2)+\Theta(\eta)\delta|\eta|^3,
$$

$$
\Theta(\eta)=\frac{\alpha(\alpha-2)}{6}|\xi^\tau+\delta\eta\lambda_0|^{\alpha-3}\big[3\langle\overline{\xi^\tau+\delta\eta\lambda_0},\overline{\eta}\rangle+(\alpha-4)\langle\overline{\xi^\tau+\delta\eta\lambda_0},\overline{\eta}\rangle^3\big],
$$

with $\lambda_0\in(0,1)$.

Let $\widehat{g^\delta_{\tau_\nu}}(\eta)=\delta\widehat{f_{\tau_\nu}}(\xi^\tau+\delta\eta)$ and consider Taylor's expansion of $|\xi^\tau+\delta\eta|^\alpha$ at ξ^τ up to the third order. For some $\lambda_0\in(0,1)$ we have

$$
|Tf_{\tau_\nu}(x_1,x_2,x_3)|=\delta\big|\mathcal{T}_{\delta,\tau}(g^\delta_{\tau_\nu})\big(\delta x_1,\delta(x_2+x_3\alpha|\xi^\tau|^{\alpha-1}),\delta^2x_3\big)\big|.
$$

Using Hölder's inequality in x_2 and making the change of variables

$$(x_1, x_2, x_3) \to (\delta^{-1} y_1, \delta^{-1} y_2, \delta^{-2} y_3),$$

we get

$$\begin{aligned}
&\|G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2(|x'|<R)L^3(|x_3|<R)} \\
&\quad \lesssim \int\limits_{|x'|<R} \Big(\int\limits_{|x_3|<R} \prod_{\nu=1}^{3} \delta \big| \mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})\big(\delta x_1, \delta(x_2 + \alpha x_3 |\xi^{\tau}|^{\alpha-1}), \delta^2 x_3\big) \big| \, dx_3 \Big)^{2/3} dx' \\
&\quad \lesssim \delta^2 R^{1/3} \int\limits_{|x_1|<2R} \Big(\iint\limits_{\substack{|x_3|<R \\ |x_2|<\alpha 2^{\alpha} R}} \prod_{\nu=1}^{3} |\mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})(\delta x_1, \delta x_2, \delta^2 x_3)| \, dx_2 \, dx_3 \Big)^{2/3} dx_1 \\
&\quad \lesssim \delta^2 R^{1/3} \int\limits_{|y_1|<2\delta R} \Big(\iint\limits_{\substack{|y_3|<\delta^2 R \\ |y_2|<\alpha 2^{\alpha} \delta R}} \prod_{\nu=1}^{3} |\mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})(y_1, y_2, y_3)| \, \frac{dy_2}{\delta} \, \frac{dy_3}{\delta^2} \Big)^{2/3} \frac{dy_1}{\delta}.
\end{aligned}$$

Partitioning the range of y_1 into consecutive intervals I_μ as

$$(-2\delta R, 2\delta R) = \bigcup_{\mu} I_\mu, \qquad |I_\mu| = \delta^2 R,$$

we get

$$\begin{aligned}
&\|G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2(|x'|<R)L^3(|x_3|<R)} \\
&\qquad \lesssim R^{1/3} \delta^{-1} \sum_{\mu} \int_{I_\mu} \Big\| \prod_{\nu=1}^{3} \mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})(y_1, \cdot, \cdot) \Big\|^{2/3}_{L^1(|y_2|<\alpha 2^{\alpha} \delta R;\, |y_3|<\delta^2 R)} dy_1.
\end{aligned}$$

Applying Hölder's inequality with respect to y_1 on each I_μ and then decomposing the interval for y_2 similarly as

$$(-2\delta R, 2\delta R) = \bigcup_{\mu'} I_{\mu'}, \qquad |I_{\mu'}| = \delta^2 R,$$

we obtain

$$\begin{aligned}
(3.8) \qquad &\|G_{\tau_1,\tau_2,\tau_3}\|^2_{L^2(|x'|<R)L^3(|x_3|<R)} \\
&\lesssim R^{2/3} \delta^{-1/3} \sum_{\mu} \Big(\sum_{\mu'} \int\limits_{y_2 \in I_{\mu'}} dy_2 \iint\limits_{\substack{y_1 \in I_\mu \\ |y_3|<\delta^2 R}} \prod_{\nu=1}^{3} |\mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})(y_1, y_2, y_3)| \, dy_1 \, dy_3 \Big)^{2/3} \\
&\lesssim R^{2/3} \delta^{-1/3} \sum_{\mu,\mu'} \Big\| \prod_{\nu=1}^{3} \mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu}) \Big\|^{2/3}_{L^1(Q_{\mu,\mu'} \times [-\delta^2 R, \delta^2 R])},
\end{aligned}$$

where $Q_{\mu,\mu'} = I_\mu \times I_{\mu'}$ is a $\delta^2 R \times \delta^2 R$-square.

To evaluate

$$\Big\|\prod_{\nu=1}^{3}\mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})\Big\|_{L^1(Q_{\mu,\mu'}\times[\delta^2R,\delta^2R])},$$

we need to introduce a localization argument based on Poisson summation with respect to the (y_1,y_2)-variables.

Denote the center of $Q_{\mu,\mu'}$ by $y'_{\mu,\mu'}$, which belongs to $\delta^2R\mathbb{Z}^2=:\mathcal{Z}$. Choose a Schwartz function $\beta\geq 0$ such that $\operatorname{supp}\widehat{\beta}\subset B(0,1/2)\subset\mathbb{R}^2$ and $\widehat{\beta}(0)=1$. For all $z\in\mathbb{R}^2$ we have

$$\delta^{2\varepsilon}\sum_{y'_{\mu,\mu'}\in\mathcal{Z}}\beta\Big(\frac{y'_{\mu,\mu'}-z}{\delta^{2-\varepsilon}R}\Big)=1. \tag{3.9}$$

Fix $y'_{\mu_0,\mu'_0}\in\mathcal{Z}$ and define

$$K(y',y_3,z)=\int_{\mathbb{R}^2_\eta}e^{i[\langle y'-z,\eta\rangle+y_3\Phi(\xi^\tau,\lambda_0,\alpha,\delta,\eta)]}\chi_{Q_{\mu_0,\mu'_0}}(y')\chi(\eta)\,d\eta.$$

Then

$$|\chi_{Q_{\mu_0,\mu'_0}}(y')\mathcal{T}_{\delta,\tau}(g^{\delta}_{\tau_\nu})(y',y_3)|\lesssim F_1+F_2, \tag{3.10}$$

where

$$F_1=\delta^{2\varepsilon}\Big|\int K(y',y_3,z)\sum_{\substack{y'_{\mu,\mu'}\in\mathcal{Z}\\ |y'_{\mu,\mu'}-y'_{\mu_0,\mu'_0}|\leq\delta^2R^{1+\varepsilon}}}\beta\Big(\frac{y'_{\mu,\mu'}-z}{\delta^{2-\varepsilon}R}\Big)g^{\delta}_{\tau_\nu}(z)\,dz\Big|,$$

$$F_2=\delta^{2\varepsilon}\Big|\int K(y',y_3,z)\sum_{\substack{y'_{\mu,\mu'}\in\mathcal{Z}\\ |y'_{\mu,\mu'}-y'_{\mu_0,\mu'_0}|>\delta^2R^{1+\varepsilon}}}\beta\Big(\frac{y'_{\mu,\mu'}-z}{\delta^{2-\varepsilon}R}\Big)g^{\delta}_{\tau_\nu}(z)\,dz\Big|.$$

First,

$$F_2\leq\mathfrak{F}_{2,1}+\mathfrak{F}_{2,2},$$

where

$$\mathfrak{F}_{2,1}=\delta^{2\varepsilon}\int_{\substack{|z-y'_{\mu_0,\mu'_0}|\\ \leq\alpha 2^{4\alpha}\delta^2R^{1+\varepsilon_1}}}|K(y,z)|\sum_{\substack{z_{\mu,\mu'}\in\mathbb{Z}^2\\ |z_{\mu,\mu'}-y'_{\mu_0,\mu'_0}/\delta^2R|>R^{\varepsilon}}}\beta\Big(\delta^{\varepsilon}\Big(z_{\mu,\mu'}-\frac{z}{\delta^2R}\Big)\Big)|g^{\delta}_{\tau_\nu}(z)|\,dz,$$

$$\mathfrak{F}_{2,2}=\delta^{2\varepsilon}\int_{\substack{|z-y'_{\mu_0,\mu'_0}|\\ >\alpha 2^{4\alpha}\delta^2R^{1+\varepsilon_1}}}|K(y,z)|\sum_{\substack{z_{\mu,\mu'}\in\mathbb{Z}^2\\ |z_{\mu,\mu'}-y'_{\mu_0,\mu'_0}/\delta^2R|>R^{\varepsilon}}}\beta\Big(\delta^{\varepsilon}\Big(z_{\mu,\mu'}-\frac{z}{\delta^2R}\Big)\Big)|g^{\delta}_{\tau_\nu}(z)|\,dz,$$

with $z_{\mu,\mu'}=y'_{\mu,\mu'}\delta^{-2}R^{-1}$ and $\varepsilon_1=0.01\varepsilon$.

Since R can be chosen so large that $R^\varepsilon \gg \alpha 2^{4\alpha} > 1$, in $\mathfrak{F}_{2,1}$ we have

$$\left| z_{\mu,\mu'} - \frac{z}{\delta^2 R} \right| \geq \left| z_{\mu,\mu'} - \frac{y'_{\mu_0,\mu_0'}}{\delta^2 R} \right| - \left| \frac{z}{\delta^2 R} - \frac{y'_{\mu_0,\mu_0'}}{\delta^2 R} \right| \geq \frac{R^{0.9\varepsilon}}{2}.$$

Hence

$$\sum_{|z_{\mu,\mu'} - y'_{\mu_0,\mu_0'}/\delta^2 R| > R^\varepsilon} \delta^{2\varepsilon} \beta\left(\delta^\varepsilon \left(z_{\mu,\mu'} - \frac{z}{\delta^2 R} \right) \right) \tag{3.11}$$

is bounded by

$$\delta^{2\varepsilon} \int_{|z| > R^{0.9\varepsilon}/2} \beta(\delta^\varepsilon z)\, dz \lesssim_N \int_{|z| > 0.5(\delta R^{0.9})^\varepsilon} (1 + |z|)^{-N}\, dz.$$

Noting that $\delta > R^{-1/2}$, for suitably large N depending on ε we have

$$(3.11) \lesssim_\varepsilon R^{-2000}.$$

By Cauchy–Schwarz's inequality in z and the boundedness of $\|K(y,\cdot)\|_2$,

$$\mathfrak{F}_{2,1} \lesssim_\varepsilon R^{-2000} \int \left| K(y', y_3, z) g^\delta_{\tau_\nu}(z) \right| dz \lesssim_\varepsilon R^{-2000} \|g^\delta_{\tau_\nu}\|_2.$$

To estimate $\mathfrak{F}_{2,2}$, in view of (3.9) we write

$$\mathfrak{F}_{2,2} \leq \int_{|z - y'_{\mu_0,\mu_0'}| > \alpha 2^{4\alpha} \delta^2 R^{1+\varepsilon_1}} |K(y', y_3, z)| \cdot |g^\delta_{\tau_\nu}(z)|\, dz. \tag{3.12}$$

Since y' is restricted to the $2\delta^2 R$-neighborhood of $y'_{\mu_0,\mu_0'}$, we have

$$\begin{aligned} &|y' - z| - |y_3 \nabla_\eta \Phi(\xi^\tau, \lambda_0, \alpha, \delta\eta)| \\ &\qquad \geq |z - y'_{\mu_0,\mu_0'}| - |y' - y'_{\mu_0,\mu_0'}| - |y_3 \nabla_\eta \Phi(\xi^\tau, \lambda_0, \alpha, \delta\eta)| \\ &\qquad \gtrsim \alpha 2^{4\alpha} \delta^2 R^{1+\varepsilon_1} - 2\delta^2 R - \alpha 2^{3\alpha} \delta^2 R \gtrsim \alpha 2^{\alpha-1} \delta^2 R^{1+\varepsilon_1}. \end{aligned}$$

By introducing the differential operator

$$\mathfrak{D} = \frac{y' - z + y_3 \nabla_\eta \Phi}{|y' - z + y_3 \nabla_\eta \Phi|^2} \cdot \nabla_\eta,$$

we may estimate $K(y', y_3, z)$ in $\mathfrak{F}_{2,2}$ using integration by parts to get

$$K(y', y_3, z) \lesssim_N |y' - z|^{-N}.$$

Inserting this into (3.12) and using Cauchy–Schwarz, for suitable $N = N(\varepsilon_1)$ we have

$$\mathfrak{F}_{2,2} \lesssim_{\alpha,\varepsilon} R^{-2000} \|g^\delta_{\tau_\nu}\|_2.$$

Consequently, the contribution of F_2 to (3.10) is negligible.

Now, let us evaluate the contribution of F_1. For brevity, we denote

$$B_{\mu_0,\mu_0'}(z) = \sum_{\substack{y'_{\mu,\mu'}\in\mathcal{Z} \\ |y'_{\mu,\mu'}-y'_{\mu_0,\mu_0'}|\le\delta^2R^{1+\varepsilon}}} \beta\Big(\frac{y'_{\mu,\mu'}-z}{\delta^{2-\varepsilon}R}\Big).$$

From the definition of $\widehat{g^\delta_{\tau_\nu}}$, we have

$$\operatorname{supp}\widehat{B_{\mu_0,\mu_0'}g^\delta_{\tau_\nu}} \subset \frac{1}{\delta}(\Omega_{\tau_\nu}-\xi^\tau)+\mathcal{O}(1/R^{\varepsilon/2}).$$

Note that we may choose $K \ll R^{\varepsilon/2}$, and non-collinearity still holds for the supports of $\{\widehat{B_{\mu_0,\mu_0'}g^\delta_{\tau_\nu}}\}_{\nu=1}^3$. The main contribution to

$$\Big\|\prod_{\nu=1}^3 \mathcal{T}_{\delta,\tau}(g^\delta_{\tau_\nu})\Big\|_{L^1(Q_{\mu,\mu'}\times[-\delta^2R,\delta^2R])}$$

comes from

$$\delta^{2\varepsilon}\Big\|\prod_{\nu=1}^3 \mathcal{T}_{\delta,\tau}(B_{\mu,\mu'}g^\delta_{\tau_\nu})\Big\|_{L^1(Q_{\mu,\mu'}\times[-\delta^2R,\delta^2R])}. \tag{3.13}$$

For the $\varepsilon>0$ at the beginning of this section, we claim that

$$\Big\|\prod_{\nu=1}^3 \mathcal{T}_{\delta,\tau}(B_{\mu,\mu'}g^\delta_{\tau_\nu})\Big\|_{L^1(Q_{\mu,\mu'}\times[-\delta^2R,\delta^2R])} \lesssim_\varepsilon R^\varepsilon\prod_{\nu=1}^3\|B_{\mu,\mu'}g^\delta_{\tau_\nu}\|_{L^2}. \tag{3.14}$$

We postpone the proof of (3.14) to Subsection 3.3. At present, we show how (3.14) implies (3.3). Using (3.8), (3.14) and Hölder's inequality, we obtain

$$(3.6)\lesssim R^{2/3+3\varepsilon}\delta^{1/6-1/3}\prod_{\nu=1}^3\Big(\sum_{\mu,\mu'}\|B_{\mu,\mu'}g^\delta_{\tau_\nu}\|_{L^2}^2\Big)^{1/3}.$$

From the definition of $B_{\mu,\mu'}$, by Cauchy–Schwarz we have

$$\begin{aligned}\sum_{\mu,\mu'}\int|B_{\mu,\mu'}(z)g^\delta_{\tau_\nu}(z)|^2\,dz \tag{3.15}\\ &\le R^{2\varepsilon}\sum_{\mu,\mu'}\sum_{\substack{y_0'\in\mathcal{Z}\\ |y_0'-y'_{\mu,\mu'}|\le\delta^2R^{1+\varepsilon}}}\int_{\mathbb{R}^2}\Big|\beta\Big(\frac{y_0'-z}{\delta^{2-\varepsilon}R}\Big)g^\delta_{\tau_\nu}(z)\Big|^2\,dz\\ &\le R^{2\varepsilon}\sum_{y_0'\in\mathcal{Z}} r_{y_0'}\int_{\mathbb{R}^2}\Big|\beta\Big(\frac{y_0'-z}{\delta^{2-\varepsilon}R}\Big)g^\delta_{\tau_\nu}(z)\Big|^2\,dz,\end{aligned}$$

where

$$r_{y_0'} = \#\{y'_{\mu,\mu'}\in\mathcal{Z}\mid |y_0'-y'_{\mu,\mu'}|\le\delta^2R^{1+\varepsilon}\}\lesssim R^{2\varepsilon}.$$

Invoking $1/\delta < \sqrt{R}$, we get

$$(3.15) \lesssim R^{4\varepsilon}\delta^{-4\varepsilon}\int\Big[\delta^{2\varepsilon}\sum_{y_0'\in\mathcal{Z}}\beta\Big(\frac{y_0'-z}{\delta^{2-\varepsilon}R}\Big)|g^{\delta}_{\tau_\nu}(z)|\Big]^2\,dz$$
$$\lesssim \Big(\frac{R}{\delta}\Big)^{4\varepsilon}\|g^{\delta}_{\tau_\nu}\|_2^2 \lesssim R^{6\varepsilon}\|f_\tau\|_2^2.$$

As a consequence, we have

$$(3.6) \lesssim_\varepsilon R^{3/4+9\varepsilon}\|f_\tau\|_2^2.$$

This implies (3.3) since $\varepsilon > 0$ can be taken arbitrarily small.

3.2. The proof of (3.4). Letting $\delta = 1/\sqrt{R}$, we adopt the same argument as in Subsection 3.1 to obtain (3.6) with Tf_τ in place of G_{τ_1,τ_2,τ_3} so that

$$\int_{|x'|<R}\sup_{|x_3|<R}(\psi_\tau|Tf_\tau|)^2(x',x_3)\,dx' \lesssim_\varepsilon R^\varepsilon\delta^{1/6}\|Tf_\tau\|^2_{L^2(|x'|<R)L^3(|x_3|<R)},$$

where Ω_τ is a $1/\sqrt{R}\times 1/\sqrt{R}$-square.

Denote $\widehat{g^{\delta}_{\tau}}(\eta) = \delta\widehat{f_\tau}(\xi^\tau+\delta\eta)$. The previous argument yields

$$\int_{|x'|<R}\sup_{|x_3|<R}(\psi_\tau|Tf_\tau|)^2(x',x_3)\,dx' \lesssim_\varepsilon R^{1/12+\varepsilon}\sum_{\mu,\mu'}\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|^2_{L^3(Q_{\mu,\mu'}\times[-1,1])},$$

where $Q_{\mu,\mu'}$ is a square with unit length. Invoking the definition of $\mathcal{T}_{\delta,\tau}(g^\delta_\tau)$, we have

$$\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|_{L^\infty(Q_{\mu,\mu'}\times[-1,1])} \le \|g^\delta_\tau\|_2 \le \|f_\tau\|_2.$$

By Hölder's inequality and Plancherel's theorem, we obtain

$$\sum_{\mu,\mu'}\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|^2_{L^3(Q_{\mu,\mu'}\times[-1,1])} \lesssim \|f_\tau\|_2^{2/3}\sum_{\mu,\mu'}\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|^{4/3}_{L^2(Q_{\mu,\mu'}\times[-1,1])}$$
$$\lesssim \|f_\tau\|_2^{2/3}\Big(\sum_{\mu,\mu'}1\Big)^{1/3}\Big(\sum_{\mu,\mu'}\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|^2_{L^2(Q_{\mu,\mu'}\times[-1,1])}\Big)^{2/3}$$
$$\lesssim R^{2/3}\|f_\tau\|_2^{2/3}\|\mathcal{T}_{\delta,\tau}(g^\delta_\tau)\|^{4/3}_{L^\infty(|x_3|<1)L^2(|x'|\le R)} \lesssim R^{2/3}\|f_\tau\|_2^2.$$

Therefore (3.4) follows. Collecting (3.3) and (3.4), we conclude that (3.14) implies (2.8). We shall prove (3.14) in the next subsection.

3.3. The proof of (3.14). To prove (3.14), we need the multilinear restriction theorem of [BCT]. Since a special form of this theorem is sufficient for our purposes, we formulate it only in this form.

Now we introduce some basic assumptions. Let $U \subset \mathbb{R}^{d-1}_\eta$ be a compact neighborhood of the origin and $\Sigma : U \to \mathbb{R}^d$ be a smooth parametrization of a $(d-1)$-hypersurface of $\mathbb{R}^d$. For $U_\nu \subset U$ and g_ν supported in $U_\nu \subset \mathbb{R}^{d-1}$ with $1 \le \nu \le d$, assume that there is a constant $\mu > 0$ such that

$$\det\bigl(X(\eta^{(1)}), \dots, X(\eta^{(d)})\bigr) > \mu \tag{3.16}$$

for all $\eta^{(1)} \in U_1, \dots, \eta^{(d)} \in U_d$, where

$$X(\eta) = \bigwedge_{k=1}^{d-1} \frac{\partial}{\partial \eta_k} \varSigma(\eta), \qquad \eta = (\eta_1, \dots, \eta_{d-1}).$$

Assume also that there is a constant $A \ge 0$ such that

$$\|\varSigma\|_{C^2(U_\nu)} \le A \quad \text{ for all } 1 \le \nu \le d. \tag{3.17}$$

For each $1 \le \nu \le d$ and $g_\nu \in L^p(U_\nu)$, $p \ge 1$, define

$$\mathcal{S} g_\nu(x) = \int_{U_\nu} e^{ix\cdot\varSigma(\eta)} g_\nu(\eta)\, d\eta.$$

THEOREM 3.2. *Under the assumption of* (3.16) *and* (3.17), *for each* $\varepsilon > 0$, $q \ge 2d/(d-1)$ *and* $p' \le (d-1)q/d$, *there is a constant* $C > 0$, *depending only on* A, ε, p, q, d, μ, *such that*

$$\Bigl\|\prod_{\nu=1}^{d} \mathcal{S} g_\nu\Bigr\|_{L^{q/d}(B(0,R))} \le C_\varepsilon R^\varepsilon \prod_{\nu=1}^{d} \|g_\nu\|_{L^p(U_\nu)} \tag{3.18}$$

for all $g_1, \dots, g_d \in L^p(\mathbb{R}^{d-1})$ *and all* $R \ge 1$.

REMARK 3.3. We shall use (3.18) below with $d = q = 3$ and $p = 2$.

The proof of (3.14) amounts to showing

$$\Bigl\|\prod_{\nu=1}^{3} |\mathcal{T}_{\delta,\tau}(g^\delta_{\tau_\nu})|\Bigr\|_{L^1(B(0,\lambda))} \le C_\varepsilon \lambda^\varepsilon \prod_{\nu=1}^{3} \|g^\delta_{\tau_\nu}\|_{L^2}, \qquad \forall \lambda > 0. \tag{3.19}$$

To prove (3.19), we use (3.18) with

$$\begin{aligned} &\mathcal{S} = \mathcal{T}_{\delta,\tau},\\ &U_\nu = \delta^{-1}(\varOmega_{\tau_\nu} - \xi^\tau) + \mathcal{O}(1/R^{\varepsilon/2}), \qquad \nu = 1,2,3,\\ &g_\nu = \widehat{g^\delta_{\tau_\nu}} \end{aligned}$$

and

$$\varSigma : (\eta_1, \eta_2) \to \bigl(\eta_1, \eta_2, \varPhi(\xi^\tau, \lambda_0, \alpha, \delta, \eta)\bigr).$$

If $\varSigma$ satisfies (3.16) and (3.17), then (3.19) follows immediately. Since the smoothness condition (3.17) is clear from the definition of $\varPhi$, we only need to show the transversality condition (3.16).

A simple calculation yields

$$\begin{cases} \partial_{\eta_1} \varSigma = \bigl(1, 0, \partial_{\eta_1} \varPhi(\xi^\tau, \lambda_0, \alpha, \delta, \eta)\bigr),\\ \partial_{\eta_2} \varSigma = \bigl(0, 1, \partial_{\eta_2} \varPhi(\xi^\tau, \lambda_0, \alpha, \delta, \eta)\bigr), \end{cases}$$

where

$$\begin{cases} \partial_{\eta_1}\Phi = \alpha|\xi^\tau|^{\alpha-2}\eta_1 + \delta\partial_{\eta_1}(\Theta(\eta)|\eta|^3), \\ \partial_{\eta_2}\Phi = \alpha(\alpha-1)|\xi^\tau|^{\alpha-2}\eta_2 + \delta\partial_{\eta_2}(\Theta(\eta)|\eta|^3). \end{cases}$$

Since

$$\Theta(\eta) = \frac{\alpha(\alpha-2)}{6}|\xi^\tau+\delta\eta\lambda_0|^{\alpha-3}\big[3\langle\overline{\xi^\tau+\delta\eta\lambda_0},\overline{\eta}\rangle + (\alpha-4)\langle\overline{\xi^\tau+\delta\eta\lambda_0},\overline{\eta}\rangle^3\big],$$

we have

$$\nabla_\eta(\Theta(\eta)|\eta|^3) = \mathcal{O}_\alpha(1).$$

This along with $\delta \le 1/K \ll 1$ allows us to write

$$X(\eta) = \big(\alpha(\alpha-1)|\xi^\tau|^{\alpha-2}\eta_1, \alpha|\xi^\tau|^{\alpha-2}\eta_2, -1\big) + \mathcal{O}_\alpha(1)(1/K, 1/K, 0),$$

hence

$$\begin{aligned} \det\big(X(\eta^{\tau_1}), X(\eta^{\tau_2}), X(\eta^{\tau_3})\big) \\ = \alpha^2(\alpha-1)|\xi^\tau|^{2(\alpha-2)}\det\begin{pmatrix} -1 & -1 & -1 \\ \eta_1^{\tau_1} & \eta_1^{\tau_2} & \eta_1^{\tau_3} \\ \eta_2^{\tau_1} & \eta_2^{\tau_2} & \eta_2^{\tau_3} \end{pmatrix} + \mathcal{O}_\alpha(1/K). \end{aligned} \tag{3.20}$$

In view of the non-collinearity of $\Omega_{\tau_1}, \Omega_{\tau_2}, \Omega_{\tau_3}$, the area of the triangle formed by η^{τ_1}, η^{τ_2} and η^{τ_3} is uniformly bounded away from zero, or equivalently, there is a $C > 0$ such that

$$\left|\det\begin{pmatrix} -1 & -1 & -1 \\ \eta_1^{\tau_1} & \eta_1^{\tau_2} & \eta_1^{\tau_3} \\ \eta_2^{\tau_1} & \eta_2^{\tau_2} & \eta_2^{\tau_3} \end{pmatrix}\right| \ge C$$

for all $\eta^{\tau_\nu} \in U_\nu$, $\nu = 1, 2, 3$. Therefore, we can reorder the columns in (3.20) to ensure

$$\det\begin{pmatrix} -1 & -1 & -1 \\ \eta_1^{\tau_1} & \eta_1^{\tau_2} & \eta_1^{\tau_3} \\ \eta_2^{\tau_1} & \eta_2^{\tau_2} & \eta_2^{\tau_3} \end{pmatrix} \ge C > 0.$$

Next, we take K large enough so that

$$(3.20) \ge \alpha^2(\alpha-1)|\xi^\tau|^{\alpha-2}C/2 > 0.$$

Consequently, we have (3.19), and this completes the proof of Theorem 1.1.

4. Proof of Lemma 2.11. Take $\eta \in \mathcal{S}(\mathbb{R}^d)$ such that $\eta \ge 0$ and $\widehat{\eta}$ is supported in $B(0,1/2)$ with $\widehat{\eta}(0) = 1$. Denoting $\mathfrak{X} = R\,\mathbb{Z}^d$, we have Poisson's summation formula

$$\sum_{x^\sigma\in\mathfrak{X}} \eta\Big(\frac{x-x^\sigma}{R}\Big) = 1, \qquad \forall x \in \mathbb{R}^d.$$

We adopt the notation of Lemma 2.11. First, noting that

$$U(t)f(x') = U(t-t_j)U(t_j)f(x'),$$

we may write, for $x_{d+1} \in (0, R^\alpha)$,

$$\psi(x')Tf(x) = \sum_{j=0}^{[R^{\alpha-1}]} \int_{\mathbb{R}^d} \chi_{I_j}(x_{d+1}) K(x,y;t_j) Tf(y,t_j)\,dy, \tag{4.1}$$

with

$$K(x,y;t_j) = \int_{\mathbb{R}^d} e^{i[(x'-y)\cdot\xi + (x_{d+1}-t_j)|\xi|^\alpha]} \psi(x')\chi(\xi)\,d\xi,$$

where $\chi \in C_c^\infty(\mathbb{R}^d)$ with $\chi(\xi) \equiv 1$ for $\xi \in \mathcal{I}_{1/R}$. Without loss of generality, we may assume x_{d+1} belongs to I_j for some $j \in \{0, \dots, [R^{\alpha-1}]\}$. Using Poisson's summation formula, we have

$$(4.1) = J_1 + J_2,$$

where

$$J_1 = \sum_{\substack{y^\sigma \in \mathfrak{X} \\ |y^\sigma| \geq 10R^{1+\varepsilon}}} \int_{\mathbb{R}^d} K(x,y;t_j)\eta\Big(\frac{y-y^\sigma}{R}\Big) Tf(y,t_j)\,dy,$$

$$J_2 = \sum_{\substack{y^\sigma \in \mathfrak{X} \\ |y^\sigma| < 10R^{1+\varepsilon}}} \int_{\mathbb{R}^d} K(x,y;t_j)\eta\Big(\frac{y-y^\sigma}{R}\Big) Tf(y,t_j)\,dy.$$

4.1. The estimation of J_1. We have

$$J_1 \leq J_{1,1} + J_{1,2},$$

where

$$J_{1,1} := \int_{|y| \leq \alpha 2^{\alpha+2}R} |K(x,y;t_j)| \sum_{y^\sigma \in \mathfrak{X},\, |y^\sigma| \geq 10R^{1+\varepsilon}} \eta\Big(\frac{y-y^\sigma}{R}\Big) |Tf(y,t_j)|\,dy,$$

$$J_{1,2} := \int_{|y| > \alpha 2^{\alpha+2}R} |K(x,y;t_j)|\,|Tf(y,t_j)|\,dy.$$

We show J_1 is negligible by estimating the contributions of $J_{1,1}$ and $J_{1,2}$ separately.

- **Estimation of $J_{1,1}$.** It is easy to see that

$$|J_{1,1}| \leq \int_{|y| \leq \alpha 2^{\alpha+2}R} \sum_{z \in \mathbb{Z}^d,\, |z| \geq 10R^\varepsilon} \eta\Big(\frac{y}{R} - z\Big) |K(x,y;t_j)|\,|Tf(y,t_j)|\,dy.$$

Since $R^\varepsilon \gg \alpha 2^{\alpha+1} > 1$, $|y| \le \alpha 2^{\alpha+2} R$ and $\eta \in \mathcal{S}(\mathbb{R}^d)$, we have

$$\sum_{z\in\mathbb{Z}^d,\,|z|\ge 10R^\varepsilon} \eta\Big(\frac{y}{R} - z\Big) \lesssim_N \int_{|x|>5R^\varepsilon} \Big(1 + \Big|\frac{y}{R} - x\Big|\Big)^{-N} dx \lesssim_N R^{-\varepsilon(N-d)}.$$

Choosing $N \approx 100d/\varepsilon + d$, we obtain

$$\Big|\sum_{|z|\ge 10R^\varepsilon} \eta\Big(\frac{y}{R} - z\Big)\Big| \lesssim_\varepsilon R^{-100d}.$$

By Cauchy–Schwarz and (2.1),

$$|J_{1,1}| \lesssim_\varepsilon R^{-100d} \|\mathcal{F}_{y\to\xi} K(x,\cdot,t_j)\|_{L^2_\xi} \|Tf(\cdot,t_j)\|_{L^2_y} \lesssim_\varepsilon R^{-100d}\|f\|_2.$$

Thus $J_{1,1}$ is negligible.

- **Estimation of $J_{1,2}$.** First, the phase function of $K(x,y;t_j)$ reads

$$\Phi(x,y,\xi,t_j) := (x'-y)\cdot\xi + (x_{d+1} - t_j)|\xi|^\alpha.$$

Thus the critical points of $\Phi(x,y,\xi,t_j)$ occur only when

$$y = x' + (x_{d+1} - t_j)\alpha|\xi|^{\alpha-2}\xi.$$

Since $R^\varepsilon \gg \alpha 2^{\alpha+1} > 1$ and

$$|x'| \le 2R, \quad 0 < x_{d+1} - t_j < R, \quad |y| \ge \alpha 2^{\alpha+2} R,$$

by the triangle inequality we have

$$|\nabla_\xi \Phi| \ge |y| - |x'| - \alpha 2^{\alpha-1}|x_{d+1} - t_j| \ge |y| - \alpha 2^{\alpha+1} R \ge R.$$

Using integration by parts, we may estimate $K(x,y;t_j)$ in $J_{1,2}$ by

$$|K(x,y;t_j)| \lesssim_\alpha \psi(x')(|y| - \alpha 2^{\alpha+1}R)^{-100d}.$$

As a consequence,

$$|J_{1,2}| \lesssim_\alpha \psi(x') R^{-99d}\|f\|_2.$$

Hence this term is also negligible.

4.2. The estimation of J_2. Rewrite J_2 as

$$J_2 = \psi(x')\int e^{i[x'\cdot\xi + (x_{d+1}-t_j)|\xi|^\alpha]}\chi(\xi)\widehat{f_j}(\xi)\,d\xi = \psi(x')Tf_j(x', x_{d+1} - t_j),$$

where

$$f_j(y) = \sum_{y^\sigma\in\mathfrak{X},\,|y^\sigma|<10R^{1+\varepsilon}} \eta\Big(\frac{y - y^\sigma}{R}\Big) Tf(y,t_j). \tag{4.2}$$

This gives the first term on the right side of (2.10). It suffices to show that the f_j's defined by (4.2) satisfy (2.11).

To do this, we first perform some reductions. Let $\{\xi^{(k)}\}_k$ be a family of maximal $R^{1-\alpha}$-separated points of $\mathcal{I}_{1/R}$ and cover $\mathcal{I}_{1/R}$ with essentially disjoint balls $B(\xi^{(k)}, R^{1-\alpha})$. This covering admits a partition of unity

$$\sum_k \varphi_k(\xi) = 1,$$

where φ_k is a smooth function supported in $B(\xi^{(k)}, R^{1-\alpha})$. Hence $f = \sum_k f_{(k)}$ and $f_j = \sum_k f_{j,(k)}$ for $j \in \{0, \dots, [R^{\alpha-1}]\}$, where

$$\widehat{f_{(k)}}(\xi) = \widehat{f}(\xi)\varphi_k(\xi) \quad \text{and} \quad \widehat{f_{j,(k)}}(\xi) = \widehat{f_j}(\xi)\varphi_k(\xi)$$

are all supported in $B(\xi^{(k)}, R^{1-\alpha})$. By Plancherel's theorem and almost orthogonality, it suffices to find some $c_d > 0$ such that

$$\sum_{j=0}^{[R^{\alpha-1}]} \|f_{j,(k)}\|_2^2 \le C_\varepsilon R^{\varepsilon c_d} \|f_{(k)}\|_2^2, \tag{4.3}$$

with $C_\varepsilon > 0$ independent of k.

Without loss of generality, we only deal with the case $k = 0$ and suppress the subscript k in $f_{(k)}$ and $f_{j,(k)}$ for brevity. As a result, we may assume $\operatorname{supp} \widehat{f} \subset B(\xi^{(0)}, R^{1-\alpha})$ in the following argument and normalize $\|f\|_2 = 1$. By Cauchy–Schwarz, we have

$$|f_j(y)|^2 \lesssim R^{2d\varepsilon} \sum_{y^\sigma \in \mathfrak{X},\, |y^\sigma| < 10R^{1+\varepsilon}} \eta\left(\frac{y - y^\sigma}{R}\right)^2 |Tf(y, t_j)|^2.$$

Integrating with respect to y and summing over j, we obtain

$$\sum_j \|f_j\|_2^2 \lesssim R^{2d\varepsilon} \sum_j \sum_{y^\sigma \in \mathfrak{X},\, |y^\sigma| < 10R^{1+\varepsilon}} \int \eta\left(\frac{y - y^\sigma}{R}\right)^2 |Tf(y, t_j)|^2 \, dy. \tag{4.4}$$

Invoking the definition of $Tf(y, t_j)$, we can write

$$Tf(y, t_j) = I_1 + I_2,$$

where

$$I_1 := \int_{\Omega_{y,j}^c} \mathfrak{K}(y, z; t_j) f(z)\, dz, \qquad I_2 := \int_{\Omega_{y,j}} \mathfrak{K}(y, z; t_j) f(z)\, dz,$$

$$\mathfrak{K}(y, z; t_j) := \int e^{i[(y-z)\cdot\xi + t_j|\xi|^\alpha]} \chi(\xi)\, d\xi,$$

$$\Omega_{y,j} := \big\{ z \in \mathbb{R}^d \;\big|\; \big|z - y - \alpha t_j |\xi^{(0)}|^{\alpha-2} \xi^{(0)}\big| < \alpha 2^{\alpha+2} R \big\}.$$

Thus, it suffices to evaluate the contributions of I_1 and I_2 to (4.4).

• **The contribution of I_1.** Since $|\xi - \xi^{(0)}| \le R^{1-\alpha}$ and $|t_j| \le R^\alpha$, we have

$$\big|\nabla_\xi [(y - z)\cdot\xi + t_j|\xi|^\alpha]\big| \ge |z - (y + \alpha t_j |\xi^{(0)}|^{\alpha-2}\xi^{(0)})| - \alpha 2^\alpha R \ge \alpha 2^{\alpha+1} R.$$

This allows us to use integration by parts to evaluate

$$|\mathfrak{K}(y,z;t_j)| \lesssim_N \big|z-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|^{-N}.$$

Choosing N large enough, we see that the contribution of I_1 to (4.4) is bounded by

$$\begin{aligned}
&R^{2\varepsilon d}\sum_j \sum_{y^\sigma\in\mathfrak{X},\,|y^\sigma|\le 10R^{1+\varepsilon}} \int_{\mathbb{R}^d} \eta\Big(\frac{y-y^\sigma}{R}\Big)^2 |I_1|^2\,dy\\
&\quad\lesssim_N R^{3\varepsilon d}\sum_j \sup_{y^\sigma}\int_{\mathbb{R}^d}\eta\Big(\frac{y-y^\sigma}{R}\Big)^2 \int_{\Omega_{y,j}^c}\big|z-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|^{-2N}\,dz\,dy\\
&\quad\lesssim_N R^{3\varepsilon d-2N+2d+\alpha-1}\lesssim_\varepsilon R^{-200d}.
\end{aligned}$$

• **The contribution of I_2.** We use Poisson's summation formula with respect to z-variable in I_2 to get

$$\sum_j \sum_{y^\sigma\in\mathfrak{X},\,|y^\sigma|\le 10R^{1+\varepsilon}} \int_{\mathbb{R}^d} \eta\Big(\frac{y-y^\sigma}{R}\Big)^2 \Big|\int_{\Omega_{y,j}} \mathfrak{K}(y,z;t_j)f(z)\,dz\Big|^2\,dy \le L_1+L_2,$$

where

$$L_1=\sum_j \sum_{\substack{y^\sigma\in\mathfrak{X}\\ |y^\sigma|\le 10R^{1+\varepsilon}}} \int_{\mathbb{R}^d}\eta\Big(\frac{y-y^\sigma}{R}\Big)^2\Big|\sum_{\substack{z_0\in\mathfrak{X}\\ z_0\notin\mathfrak{A}(y)}}\int_{\Omega_{y,j}}\eta\Big(\frac{z-z_0}{R}\Big)\mathfrak{K}(y,z;t_j)f(z)\,dz\Big|^2\,dy,$$

$$L_2=\sum_j \sum_{\substack{y^\sigma\in\mathfrak{X}\\ |y^\sigma|\le 10R^{1+\varepsilon}}} \int_{\mathbb{R}^d}\eta\Big(\frac{y-y^\sigma}{R}\Big)^2\Big|\sum_{\substack{z_0\in\mathfrak{X}\\ z_0\in\mathfrak{A}(y)}}\int_{\Omega_{y,j}}\eta\Big(\frac{z-z_0}{R}\Big)\mathfrak{K}(y,z;t_j)f(z)\,dz\Big|^2\,dy,$$

$$\mathfrak{A}(y)=\big\{z_0\in\mathfrak{X}\ \big|\ \big|z_0-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|\le 10R^{1+\varepsilon}\big\}.$$

Now, we show L_1 is also negligible. In fact, since

$$\big|z_0-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|>10R^{1+\varepsilon}$$

and

$$\big|z-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|<\alpha 2^{\alpha+2}R,$$

by taking R so large that $R^\varepsilon\gg\alpha 2^{\alpha+2}$ we have

$$|z-z_0|\ge\frac12\big|z_0-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|.$$

Under the above conditions, we obtain

$$\begin{aligned}
\sum_{z_0\in\mathfrak{X}\setminus\mathfrak{A}(y)}\eta\Big(\frac{z-z_0}{R}\Big) &\lesssim_N R^N\sum_{z_0\in\mathfrak{X}\setminus\mathfrak{A}(y)}\big|z_0-(y+\alpha t_j|\xi^{(0)}|^{\alpha-2}\xi^{(0)})\big|^{-N}\\
&\lesssim_N R^{-(N-d)(1+\varepsilon)}.
\end{aligned}$$

Choosing

$$N \approx \frac{1}{2\varepsilon}(200d + \alpha + 1),$$

and using Cauchy–Schwarz as before, we get

$$\begin{aligned} L_1 &\lesssim_N R^{-2\varepsilon N + 2d(1+\varepsilon)} \\ &\qquad \times \sum_j \sum_{|y^\sigma| \le 10R^{1+\varepsilon}} \int_{\mathbb{R}^d} \eta\Big(\frac{y-y^\sigma}{R}\Big)^2 \Big(\int_{\Omega_{y,j}} |\mathfrak{K}(y,z;t_j) f(z)|\, dz \Big)^2 dy \\ &\lesssim_N R^{-2\varepsilon N + 5d + \alpha - 1} \lesssim_\varepsilon R^{-100d}. \end{aligned}$$

Thus the contribution of L_1 is negligible.

Next, we turn to the evaluation of L_2. This term contains the non-trivial contribution to (4.4). First, applying Cauchy–Schwarz's inequality to the summation with respect to z_0, we have

$$L_2 \lesssim R^{\varepsilon d}(H_1 + H_2),$$

where for $\gamma = 1, 2$,

$$\begin{aligned} H_\gamma = \sum_j \sum_{|y^\sigma| \le 10R^{1+\varepsilon}} \sum_{z_0 \in \mathfrak{A}(y)} \int_{\mathbb{R}^d} \eta\Big(\frac{y-y^\sigma}{R}\Big)^2 \\ \times \Big| \int_{\Omega^\gamma_{y,j,z_0}} \eta\Big(\frac{z-z_0}{R}\Big) \mathfrak{K}(y,z;t_j) f(z)\, dz \Big|^2 dy \end{aligned}$$

with

$$\Omega^1_{y,j,z_0} = \Omega_{y,j} \cap B(z_0, R^{1+\varepsilon}), \qquad \Omega^2_{y,j,z_0} = \Omega_{y,j} \setminus B(z_0, R^{1+\varepsilon}).$$

- **The evaluation of H_1.** Since

$$|t_j - t_{j+1}| = R, \qquad j = 0, \dots, [R^{\alpha-1}] - 1,$$

at most R^ε of the $\Omega_{y,j}$'s intersect $B(z_0, R^{1+\varepsilon})$. Denote by χ_1 the characteristic function of $B(z_0, R^{1+\varepsilon})$. Applying Plancherel's theorem to H_1 yields

$$\begin{aligned} H_1 &\lesssim R^\varepsilon \max_j \sum_{z_0 \in \mathfrak{X}} \int_{\mathbb{R}^d} \Big| \iint e^{i[(y-z)\cdot\xi + t_j|\xi|^\alpha]} \chi_{\Omega_{y,j}} \eta\Big(\frac{z-z_0}{R}\Big) \chi_1(z) f(z)\, dz\, d\xi \Big|^2 dy \\ &\lesssim R^\varepsilon \max_j \sum_{z_0 \in \mathfrak{X}} \int_{\mathbb{R}^d} \Big| \int e^{i(y\cdot\xi + t_j|\xi|^\alpha)} \mathcal{F}_{z\to\xi}\Big(\chi_{\Omega_{y,j}} \eta\Big(\frac{\cdot - z_0}{R}\Big) \chi_1 f(\cdot)\Big)(\xi)\, d\xi \Big|^2 dy \\ &\lesssim R^\varepsilon \sum_{z_0 \in \mathcal{X}} \int_{\mathbb{R}^d} \Big[\eta\Big(\frac{z-z_0}{R}\Big) \chi_1(z) |f(z)| \Big]^2 dz \lesssim R^\varepsilon. \end{aligned}$$

• **The evaluation of H_2.** Since $\eta \in \mathcal{S}(\mathbb{R}^d)$, we have

$$H_2 \lesssim_N R^{2\varepsilon d} \sup_{y^\sigma} \sum_j \int_{\mathbb{R}^d} \eta\Big(\frac{y-y^\sigma}{R}\Big) \sup_{z_0} \bigg| \int_{|z-z_0|\geq R^{1+\varepsilon}} \bigg|\frac{z-z_0}{R}\bigg|^{-N} |f(z)|\,dz\bigg|^2 dy$$
$$\lesssim_N R^{-\varepsilon(N-d)+d+\alpha-1} \lesssim_\varepsilon R^{\varepsilon d}.$$

Collecting all the estimations of I_1, L_1 and H_1, we eventually get (2.11), and this completes the proof of Lemma 2.11.

Remark 4.1. After finishing this work, we were informed by Professor Sanghyuk Lee that Lemma 2.11 can be deduced without losing R^ε by adapting the argument for the temporal localization Lemma 2.1 in [CLV]. We have however decided to include our method since it exhibits different techniques which are interesting in their own right.

5. Proof of Lemma 2.4. The proof is divided into three parts. First, we establish an auxiliary result (Lemma 5.1). Second, we deduce an inductive formula with respect to different scales by exploiting the self-similarity of Lemma 5.1. Finally, we iterate this inductive formula to get Lemma 2.4.

5.1. An auxiliary lemma. Let us begin with an outline of the main steps. First, we partition the support of $\widehat{f}$ into the union of $1/K \times 1/K$-squares with $K \ll R$. Then we rewrite $Tf(x)$ as a superposition of solutions of the linear Schrödinger equation, where each initial datum is frequency-localized in one of these squares. The oscillatory integral $Tf(x)$ can be transformed into an exponential sum, where the fluctuations of the coefficients on every box $Q_{a,K}$ of size $K^{1-\varepsilon} \times K^{1-\varepsilon} \times K^{1-\varepsilon}$ are so slight that they can be viewed essentially as constant on each such box.

Next, we partition $B(0,R) \times [0,R] \subset \mathbb{R}^3$ into the union of disjoint $Q_{a,K}$'s and estimate the exponential sum on each $Q_{a,K}$. In doing so, we encounter three possibilities. For the first one, we will have the transversality condition so that the multilinear restriction theorem of [BCT] can be applied. When the transversality fails, we consider the other two possibilities. For this part, we use more information from geometric structures along with Córdoba's square function estimates [Co]. Now, let us turn to details.

We partition $\mathcal{I}$ into the union of disjoint $1/K \times 1/K$-squares Ω_ν, centered at ξ_ν,

$$\mathcal{I} \subset \bigcup_\nu \Omega_\nu.$$

Then, we rewrite $Tf(x)$ as an exponential sum

$$Tf(x) = \sum_\nu T_\nu f(x) e^{i\phi(x,\xi_\nu)},$$

where $\phi(x,\xi) = x_1\xi_1 + x_2\xi_2 + x_3|\xi|^\alpha$ and

$$T_\nu f(x) = \int\limits_{\Omega_\nu} e^{i[\phi(x,\xi)-\phi(x,\xi_\nu)]}\widehat{f}(\xi)\,d\xi.$$

• **The local constancy of $T_\nu f(x)$.** From a direct computation, $\widehat{T_\nu f}(y)$ is supported in

$$\{y \in \mathbb{R}^3 \mid y = (y_1, y_2, y_3),\ |y_j| \le 1/K,\ j = 1,2,3\}.$$

If we take a smooth radial function $\hat{\eta}(\omega)$ such that $\hat{\eta}(\omega) = 1$ for $|\omega| < 2$ and $\hat{\eta}(\omega) = 0$ for $|\omega| > 4$, $\omega \in \mathbb{R}^3$, then

$$\widehat{T_\nu f}(\omega) = \widehat{T_\nu f}(\omega)\widehat{\eta_K}(\omega)$$

for $\eta_K(x) = K^{-3}\eta(K^{-1}x)$. Consequently, $T_\nu f = T_\nu f * \eta_K$.

Let $Q_a = Q_{a,K}$ be a $K^{1-\varepsilon} \times K^{1-\varepsilon} \times K^{1-\varepsilon}$-box centered at $a \in K^{1-\varepsilon}\mathbb{Z}^3$. Then

$$B(0,R) \times [0,R] \subset \bigcup_a Q_a,$$

where the union is taken over all a such that $Q_a \cap (B(0,R) \times [0,R]) \neq \emptyset$. Denote by χ_a the characteristic function of Q_a. For $x \in Q_a$, making the change of variables $x = \tilde{x} + a$ with $\tilde{x} \in Q_{0,K}$, we have

$$\begin{aligned} |T_\nu f(x)| &= \Big|\int \tau_{-a}(T_\nu f)(z)\eta_K(\tilde{x} - z)\,dz\Big| \\ &\le \int |\tau_{-a}(T_\nu f)(z)| \sup_{\tilde{x}\in Q_{0,K}} |\eta_K(\tilde{x} - z)|\,dz =: c_{a,\nu}. \end{aligned}$$

Thus, we associate to any Q_a the sequence $\{c_{a,\nu}\}_\nu$.

Noting that

$$|\nabla_x\phi(x,\xi) - \nabla_x\phi(x,\xi_\nu)| \le 1/K, \quad \forall \xi \in \Omega_\nu,$$

one easily derives the uniform estimate

$$|T_\nu f(x') - T_\nu f(x'')| \le \|f_\nu\|_2/K, \quad \forall |x' - x''| < K.$$

In particular, $|T_\nu f(x)|$ deviates from $c_{a,\nu}$ by only $\|f_\nu\|_2/K$ whenever $x \in Q_a$.

The local constancy trick can also be regarded as an extension of the Shannon–Nyquist sampling theorem of [T]. From

$$T_\nu f(x) = \int T_\nu f(x - Ky)\eta(y)\,dy,$$

and since $\int \eta = 1$, we see that $T_\nu f(x)$ is essentially an average on the ball $B(x,K)$ up to Schwartz tails. From the uncertainty principle, $|T_\nu f(x)|$ is essentially constant on the boxes $Q_{a,K}$. We refer to [T] for the standard exposition of this issue. Thus, whenever $x \in Q_a$, we may regard $T_\nu f(x)$ as $c_{\nu,a}$.

In the preparations, we set $K^{1-\varepsilon}$ to be the side length of $Q_{a,K}$ where ε is necessary for a technical reason so that the local constancy holds. However, to simplify the notation, we will suppress this small ε in the following. Keeping this in mind, we next classify the Q_a's into three categories.

• **The classification of $\{Q_a\}_a$.** Let $\mathcal{A}$ consist of all a associated to the boxes Q_a as above. We will write $\mathcal{A}$ as the union of $\mathcal{A}_j$ for $j = 1, 2, 3$ with $\mathcal{A}_j \subset \mathcal{A}$ defined as follows.

Let $c_a^* = \max_\nu c_{a,\nu}$ and $\xi_{\nu_a^*}$ be the center of the square Ω_{ν^*} associated to c_a^*. We define $\mathcal{A}_1 \subset \mathcal{A}$ so that $a \in \mathcal{A}_1$ if and only if there exist $\nu_1, \nu_2, \nu_3 \in \{1, \ldots, \sim K^2\}$ with

$$\min\{c_{a,\nu_1}, c_{a,\nu_2}, c_{a,\nu_3}\} > K^{-4} c_a^*,$$

and $\xi_{\nu_1}, \xi_{\nu_2}, \xi_{\nu_3}$ are non-collinear in the sense that

$$|\xi_{\nu_1} - \xi_{\nu_2}| \geq |\xi_{\nu_1} - \xi_{\nu_3}| \geq \operatorname{dist}(\xi_{\nu_3}, \ell(\nu_1, \nu_2)) > 10^3 \alpha 2^\alpha / K, \tag{5.1}$$

where $\ell(\nu_1, \nu_2)$ is the straight line through ξ_{ν_1}, ξ_{ν_2}.

Next, we take $1 \ll K_1 \ll K \ll R$ and define $\mathcal{A}_2 \subset \mathcal{A}$ so that $a \in \mathcal{A}_2$ if and only if

$$|\xi_\nu - \xi_{\nu_a^*}| > 4/K_1 \;\Rightarrow\; c_{a,\nu} \leq K^{-4} c_a^*. \tag{5.2}$$

Let $\mathcal{A}_3 = (\mathcal{A} \setminus \mathcal{A}_1) \cap (\mathcal{A} \setminus \mathcal{A}_2)$. Then $\mathcal{A} \subset \mathcal{A}_1 \cup \mathcal{A}_2 \cup \mathcal{A}_3$.

We claim that if $a \in \mathcal{A}_3$, then there exists a ν_a^{**} such that $c_{a,\nu_a^{**}} > K^{-4} c_a^*$ and

$$|\xi_{\nu_a^{**}} - \xi_{\nu_a^*}| > 4/K_1, \tag{5.3}$$

and moreover

$$\operatorname{dist}(\xi_\nu, \ell(\nu^*, \nu^{**})) > 10^3 \alpha 2^\alpha K_1 / K \;\Rightarrow\; c_{a,\nu} \leq K^{-4} c_a^*. \tag{5.4}$$

Thus, $\{Q_a\}_a$ is classified according to a being in $\mathcal{A}_1$, $\mathcal{A}_2$ or $\mathcal{A}_3$.

Since the first part of the claim is clear from the definition of $\mathcal{A}_3$, it suffices to show (5.4). Suppose there is an $a \in \mathcal{A}_3$ for which (5.4) fails; then there is a ν such that $c_{a,\nu} > K^{-4} c_a^*$ but

$$\operatorname{dist}(\xi_\nu, \ell(\nu_a^*, \nu_a^{**})) > 10^3 \alpha 2^\alpha K_1 / K. \tag{5.5}$$

We claim (5.5) implies that ξ_ν, $\xi_{\nu_a^*}$ and $\xi_{\nu_a^{**}}$ satisfy (5.1). Hence $a \in \mathcal{A}_1$, which contradicts $\mathcal{A}_1 \cap \mathcal{A}_3 = \emptyset$.

To prove the claim, we consider the following two alternatives due to symmetry:

Case 1: $\min\{|\xi_\nu - \xi_{\nu_a^{**}}|, |\xi_\nu - \xi_{\nu_a^*}|\} \leq |\xi_{\nu_a^*} - \xi_{\nu_a^{**}}|$,

Case 2: $\min\{|\xi_\nu - \xi_{\nu_a^{**}}|, |\xi_\nu - \xi_{\nu_a^*}|\} > |\xi_{\nu_a^*} - \xi_{\nu_a^{**}}|$.

For Case 1, assuming $|\xi_\nu - \xi_{\nu_a^*}| \le |\xi_\nu - \xi_{\nu_a^{**}}|$ without loss of generality, we deduce (5.1) immediately from

$$\begin{aligned}|\xi_{\nu_a^*} - \xi_{\nu_a^{**}}| \ge |\xi_\nu - \xi_{\nu_a^*}| &\ge \operatorname{dist}\big(\xi_\nu, \ell(\nu_a^*, \nu_a^{**})\big)\\ &> 10^3\alpha 2^\alpha K_1/K > 10^3\alpha 2^\alpha/K.\end{aligned}$$

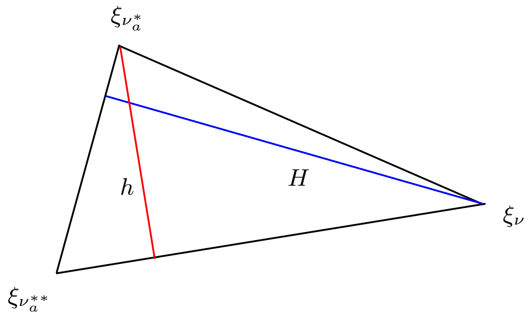

Fig. 2. The triangle $\triangle\xi_\nu\xi_{\nu_a^*}\xi_{\nu_a^{**}}$

For Case 2, we may assume $|\xi_{\nu_a^*} - \xi_{\nu_a^{**}}| < |\xi_\nu - \xi_{\nu_a^*}| \le |\xi_\nu - \xi_{\nu_a^{**}}|$ by symmetry and consider the triangle $\triangle\xi_\nu\xi_{\nu_a^*}\xi_{\nu_a^{**}}$ (Figure 2), with

$$H = \operatorname{dist}(\xi_\nu, \ell(\nu_a^*, \nu_a^{**})), \qquad h = \operatorname{dist}(\xi_{\nu_a^*}, \ell(\nu, \nu_a^{**})).$$

Considering the measure of $\triangle\xi_\nu\xi_{\nu_a^*}\xi_{\nu_a^{**}}$, from (5.3) we get

$$h = \frac{|\xi_{\nu_a^*} - \xi_{\nu_a^{**}}|}{|\xi_{\nu_a^{**}} - \xi_\nu|} H \ge \frac{4}{K_1} \cdot \frac{10^3 K_1}{K} \alpha 2^\alpha \cdot \frac{1}{4} > 10^3 \frac{\alpha 2^\alpha}{K},$$

where we have used $|\xi_{\nu_a^{**}} - \xi_\nu| \le 4$, and hence (5.1) follows.

• **An auxiliary lemma.** The following lemma exhibits once more the spirit of Bourgain–Guth's method, namely the failure of non-coplanar interactions implies small Fourier supports with possible additional separation structures.

Lemma 5.1. *Let $B(0,R) \times [0,R] \subset \bigcup_a Q_a$ be as before. On each Q_a we have*

$$|Tf(x)| \lesssim K^8 \max_{\substack{\nu_1,\nu_2,\nu_2\\ \textit{non-collinear}}} \Big(\prod_{j=1}^{3} |T_{\nu_j} f(x)|\Big)^{1/3} \tag{5.6}$$

$$+ \max_\mu \Big| \int_{\tilde\Omega_\mu} e^{i\phi(x,\xi)} \hat f(\xi)\, d\xi \Big| \tag{5.7}$$

$$+ K_1^2 \max_{\substack{\mathcal{L}_1,\mathcal{L}_2\subset\mathcal{L}_a\\ \operatorname{dist}(\mathcal{L}_1,\mathcal{L}_2)>1/K_1}} \prod_{j=1}^{2} \Big| \sum_{\Omega_\nu\subset\mathcal{L}_j} e^{i\phi(x,\xi_\nu)} T_\nu f(x)\Big|^{1/2} \tag{5.8}$$

$$+ K_1^3 K^{-1/2} \Big(\sum_{\Omega_\nu\subset\mathcal{L}_a} |T_\nu f(x)|^2 \Big)^{1/2}, \tag{5.9}$$

where $\tilde{\Omega}_\mu$ is a $1/K_1 \times 1/K_1$-square centered at $\xi_\mu \in \mathcal{I}$, and $\mathcal{L}_a$ is the $(\alpha 2^\alpha 10^3 K_1/K)$-neighborhood of the line $\ell(\nu_a^*, \nu_a^{**}) =: \ell_a^*$. The two separated portions $\mathcal{L}_1, \mathcal{L}_2$ of $\mathcal{L}_a$ are obtained by intersecting $\mathcal{L}_a$ with some $\tilde{\Omega}_{\mu_1}$ and $\tilde{\Omega}_{\mu_2}$ respectively.

Proof. For $x \in Q_a$, we estimate $|Tf(x)|$ in different ways depending on whether a is in $\mathcal{A}_1$, $\mathcal{A}_2$ or $\mathcal{A}_3$. If $a \in \mathcal{A}_1$, from the definition of $\mathcal{A}_1$ we have

$$\begin{aligned} |Tf(x)|^3 &\le (K^2 c_a^*)^3 \le K^6 K^{12} c_{a,\nu_1} c_{a,\nu_2} c_{a,\nu_3} \\ &\le K^{18} \sum_{\substack{\nu_1,\nu_2,\nu_3 \\ \text{non-collinear}}} \prod_{j=1}^{3} |T_{\nu_j} f(x)| \\ &\le K^{24} \max_{\substack{\nu_1,\nu_2,\nu_3 \\ \text{non-collinear}}} \prod_{j=1}^{3} |T_{\nu_j} f(x)|. \end{aligned}$$

If $a \in \mathcal{A}_2$, we use (5.2) for $\mathcal{A}_2$ to estimate $Tf(x)$. For brevity, we define $\|\xi\| = \max\{|\xi_1|, |\xi_2|\}$ for $\xi \in \mathbb{R}^2$ and let

$$\Omega_* = \bigcup \{\Omega_\nu \mid \|\xi_\nu - \xi_{\nu_a^*}\| \le 10/K_1\}.$$

Then Ω_* is a $10/K_1 \times 10/K_1$-square with center $\xi_{\nu_a^*}$. We may write

$$\begin{aligned} |Tf(x)| &\le \Big| \sum_{\|\xi_\nu - \xi_{\nu_a^*}\| \le 10/K_1} e^{i\phi(x,\xi_\nu)} T_\nu f(x) \Big| + \Big| \sum_{\|\xi_\nu - \xi_{\nu_a^*}\| > 10/K_1} e^{i\phi(x,\xi_\nu)} T_\nu f(x) \Big| \\ &\le \Big| \sum_{\|\xi_\nu - \xi_{\nu_a^*}\| \le 10/K_1} \int_{\Omega_\nu} e^{i\phi(x,\xi)} \hat{f}(\xi)\, d\xi \Big| + \sum_{|\xi_\nu - \xi_{\nu_a^*}| > 4/K_1} |T_\nu f(x)| \\ &\le 100 \max_{\tilde{\Omega}_\mu \subset \Omega_{\nu_a^*}} \Big| \int_{\tilde{\Omega}_\mu} e^{i\phi(x,\xi)} \hat{f}(\xi)\, d\xi \Big| + K^2 K^{-4} c_a^* \\ &\lesssim (5.7) + (5.9). \end{aligned}$$

If $a \in \mathcal{A}_3$, we write

$$Tf(x) = \mathfrak{D}_1 + \mathfrak{D}_2,$$

where

$$\mathfrak{D}_j = \chi_a(x) \int_{D_j} e^{i\phi(x,\xi)} \widehat{f}(\xi)\, d\xi, \qquad j = 1, 2,$$

with

$$\begin{aligned} D_1 &= \{\xi \mid \operatorname{dist}(\xi, \ell_a^*) > \alpha 2^\alpha 10^4 K_1/K\}, \\ D_2 &= \{\xi \mid \operatorname{dist}(\xi, \ell_a^*) \le \alpha 2^\alpha 10^4 K_1/K\}. \end{aligned}$$

From (5.4), we have

$$|\mathfrak{D}_1| \lesssim \sum_{\xi_\nu:\,\mathrm{dist}(\xi_\nu,\ell_a^*)>\alpha 2^\alpha 10^3 K_1/K} |e^{i\phi(x,\xi_\nu)}T_\nu f(x)| \le K^2K^{-4}c_a^* \le (5.9).$$

To evaluate $\mathfrak{D}_2$, we assume without loss of generality

$$\operatorname{supp}\widehat{f} \subset \{\xi \in \mathbb{R}^2 \mid \mathrm{dist}(\xi,\ell_a^*) \le \alpha 2^\alpha 10^4 K_1/K\}.$$

Let $\{\tilde{\Omega}_\mu\}_\mu$ be a family of disjoint $1/K_1 \times 1/K_1$-squares such that $\mathcal{I} \subset \bigcup_\mu \tilde{\Omega}_\mu$. For any $x \in Q_a$, we define

$$c_\mu(x) = \int_{\tilde{\Omega}_\mu} e^{i\phi(x,\xi)}\hat{f}(\xi)\,d\xi.$$

Let

$$\mathcal{H}_a = \Big\{x \in Q_a \;\Big|\; \|Tf(x)| \le 10^8 \max_\mu |c_\mu(x)|\Big\}.$$

Then

$$|\mathfrak{D}_2| \le |Tf(x)|\chi_{\mathcal{H}_a}(x) + |Tf(x)|\chi_{Q_a\setminus\mathcal{H}_a}(x),$$

where the first term is bounded by (5.7). To handle the second term, we observe that $x \in Q_a \setminus \mathcal{H}_a$ implies

$$|c_\mu(x)| \le 10^{-8}|Tf(x)|, \qquad \forall\mu. \tag{5.10}$$

Set

$$\mathcal{J}(x) = \Big\{\mu \;\Big|\; \frac{10^{-1}}{K_1^2}|Tf(x)| \le |c_\mu(x)| \le 10^{-8}|Tf(x)|,\, x \in Q_a \setminus \mathcal{H}_a\Big\}.$$

We have $\#\mathcal{J}(x) \ge 10^7$ for all $x \in Q_a \setminus \mathcal{H}_a$. Indeed, suppose $\#\mathcal{J}(x_0) < 10^7$ for some x_0; then

$$|Tf(x_0)| \le \sum_{\mu\in\mathcal{J}(x_0)} |c_\mu(x_0)| + \sum_{\mu\notin\mathcal{J}(x_0)} |c_\mu(x_0)|. \tag{5.11}$$

Because of (5.10), we can bound the right side of (5.11) by

$$10^7 \times 10^{-8}|Tf(x_0)| + K_1^2 \cdot \frac{10^{-1}}{K_1^2}|Tf(x_0)| < \frac{1}{5}|Tf(x_0)|,$$

which is impossible. Noting that the centers of $\{\tilde{\Omega}_\mu\}_\mu$ are $1/K_1$-separated, we can choose $\mu_1, \mu_2 \in \mathcal{J}(x)$ (x-dependent) such that

$$\mathrm{dist}(\tilde{\Omega}_{\mu_1}, \tilde{\Omega}_{\mu_2}) \ge 10^4/K_1$$

and

$$|Tf(x)| \le 10K_1^2 \min\{|c_{\mu_1}(x)|, |c_{\mu_2}(x)|\}, \qquad x \in Q_a \setminus \mathcal{H}_a.$$

It follows that

$$|Tf(x)|\chi_{Q_a\setminus\mathcal{H}_a}(x) \le 10K_1^2 \prod_{j=1}^2 \Big| \int_{\tilde{\Omega}_{\mu_j}} e^{i\phi(x,\xi)}\widehat{f}(\xi)\,d\xi\Big|^{1/2}, \tag{5.12}$$

where μ_1 and μ_2 might depend on x. Now in view of (5.2),

$$|c_{\mu_j}(x)| \le \Big| \sum_{\substack{\Omega_\nu \subset \tilde{\Omega}_{\mu_j} \\ \operatorname{dist}(\xi_\nu, \ell_a^*) \le \alpha 2^\alpha 10^3 K_1/K}} \int_{\Omega_\nu} e^{i\phi(x,\xi)} \widehat{f}(\xi)\, d\xi \Big| + K^2 K^{-4} c_a^*. \tag{5.13}$$

Thus

$$\begin{aligned} &|Tf(x)|\chi_{Q_a \setminus \mathcal{H}_a}(x) \\ &\le 10K_1^2 \Big[\prod_{j=1}^{2} \Big| \sum_{\Omega_\nu \subset \mathcal{L}_j} e^{i\phi(x,\xi_\nu)} T_\nu f(x) \Big|^{1/2} + \Big(2K^{-2} c_a^* \sum_{\Omega_\nu \subset \mathcal{L}_a} |T_\nu f(x)| \Big)^{1/2} + K^{-2} c_a^* \Big], \end{aligned} \tag{5.14}$$

where $\mathcal{L}_j = \tilde{\Omega}_{\mu_j} \cap \mathcal{L}_a$ with $\mathcal{L}_a$ the $\alpha 2^\alpha 10^3 K_1/K$-neighborhood of ℓ_a^*. Noting that

$$c_a^* \le \Big(\sum_{\Omega_\nu \subset \mathcal{L}_a} |T_\nu f(x)|^2 \Big)^{1/2}$$

and

$$\sum_{\Omega_\nu \subset \mathcal{L}} |T_\nu f(x)| \lesssim_\alpha K_1 K^{1/2} \Big(\sum_{\Omega_\nu \subset \mathcal{L}_a} |T_\nu f(x)|^2 \Big)^{1/2},$$

the last two terms in (5.14) are bounded by

$$K_1^3 K^{-1/2} \Big(\sum_{\Omega_\nu \subset \mathcal{L}_a} |T_\nu f(x)|^2 \Big)^{1/2}.$$

Therefore,

$$|\mathfrak{D}_2| \lesssim (5.7) + (5.8) + (5.9).$$

This completes the proof of Lemma 5.1. ∎

5.2. A self-similar iterative formula. The formula we deduce in this part is the engine for the iterative process to prove the fractional order Bourgain–Guth inequality. Let Ω_τ be a δ-square centered at ξ_τ, and let

$$\begin{aligned} \Omega_\nu^\tau &= \{\xi \in \mathbb{R}^2 \mid \|\xi - (\xi_\tau + d\xi_\nu)\| < \delta/K\}, \\ \tilde{\Omega}_\mu^\tau &= \{\xi \in \mathbb{R}^2 \mid \|\xi - (\xi_\tau + d\xi_\mu)\| < \delta/K_1\}. \end{aligned}$$

Denote $\widehat{f}_{\tau,\nu} = \widehat{f} \cdot \chi_{\Omega_\nu^\tau}$. We prove the following iterative formula from scale δ to δ/K for all $a \in \mathcal{A}$:

$$|T_\tau f|(x) \lesssim K^8 \max_{\substack{\nu_1,\nu_2,\nu_3 \\ \text{non-collinear}}} \prod_{j=1}^{3} \|Tf_{\tau,\nu_j}(x)|^{1/3} \tag{5.15}$$

$$+ \psi_\tau(x)\Big(\sum_{\Omega_\nu^\tau \subset \mathcal{L}_\tau} |Tf_{\tau,\nu}(x)|^2\Big)^{1/2} \tag{5.16}$$

$$+ \max_\mu \Big| \int_{\tilde{\Omega}_\mu^\tau \cap \Omega_\tau} e^{i[\xi\cdot x' + x_3|\xi|^\alpha]} \hat{f}(\xi)\, d\xi \Big| \tag{5.17}$$

for all $x \in \mathcal{T}_a^*$, where

$$\mathcal{T}_a^* := \{x \in \mathbb{R}^3 \mid |x_1 - a_1| < K/\delta,\ |x_2 - a_2| < K/\delta,\ |x_3 - a_3| < K/\delta^\alpha\},$$

$\mathcal{L}_\tau$ is the δ/K-neighborhood of a line segment, and ψ_τ is a function satisfying

$$\Big(\frac{1}{|B|}\int_B \psi_\tau(x)^4\, dx\Big)^{1/4} \lesssim K_1^{2\alpha}, \tag{5.18}$$

where B is a $K\mathcal{C}_\tau^*$-box centered at a.

To deduce (5.15)–(5.18), we need the following estimate, which is a standard square function estimate going back to Córdoba [Co]. The crucial L^4-estimate is used in [BG] to tackle the worst scenario by exploiting the separation of the line segments in which the frequencies are localized. This part of frequencies corresponds to the terms of main contributions.

Lemma 5.2. *For any $a \in \mathcal{A}$ and all $x \in Q_a$,*

$$\Big(\frac{1}{|Q_a|}\int_{Q_a} |(5.8) + (5.9)|^4\, dx\Big)^{1/4} \lesssim K_1^{2\alpha}\Big(\sum_{\Omega_\nu \subset \mathcal{L}} |T_\nu f(x)|^2\Big)^{1/2}, \tag{5.19}$$

where the implicit constant is independent of a.

Remark 5.3. This observation is crucial for the iteration in the next subsection. To prove (5.19), we rely heavily on the $10^4/K_1$-separation of the segments $\mathcal{L}_1, \mathcal{L}_2$. Since we are dealing with a fractional order symbol, the proof is more intricate than that in [BG], where the algebraic structure simplifies the proof significantly.

Proof of Lemma 5.2. We need to estimate the L^4-average of (5.8) and (5.9) over Q_a. Since on every Q_a, $|T_\nu f(x)|$ can be viewed as a constant, we immediately get

$$\Big(\frac{1}{|Q_a|}\int_{Q_a} (5.9)^4\, dx\Big)^{1/4} \lesssim K_1^3\Big(\sum_{\Omega_\nu \subset \mathcal{L}} |T_\nu f(x)|^2\Big)^{1/2}.$$

Next, we estimate (5.8). First, we have

$$
\begin{aligned}
(5.20)\qquad & \int_{Q_a} \max_{\substack{\mathcal{L}_1,\mathcal{L}_2\subset\mathcal{L}_a\\ \operatorname{dist}(\mathcal{L}_1,\mathcal{L}_2)\geq 10^4/K_1}} \prod_{j=1}^{2}\Big|\sum_{\Omega_\nu\subset\mathcal{L}_j} e^{i\phi(x,\xi_\nu)}T_\nu f(x)\Big|^2\,dx \\
&\leq \sum_{\substack{\mathcal{L}_1,\mathcal{L}_2\subset\mathcal{L}_a\\ \operatorname{dist}(\mathcal{L}_1,\mathcal{L}_2)\geq 10^4/K_1}} \int_{Q_a\cap\mathfrak{S}_{\mathcal{L}_1,\mathcal{L}_2}} \prod_{j=1}^{2}\Big|\sum_{\Omega_\nu\subset\mathcal{L}_j} e^{i\phi(x,\xi_\nu)}T_\nu f\Big|^2\,dx
\end{aligned}
$$

where

$$
\mathfrak{S}_{\mathcal{L}_1,\mathcal{L}_2}=\Big\{x\;\Big|\;(5.8)\leq 2\Big|\sum_{\Omega_\nu\subset\mathcal{L}_1} e^{i\phi(x,\xi_\nu)}T_\nu f(x)\Big|^{1/2}\Big|\sum_{\Omega_\nu\subset\mathcal{L}_2} e^{i\phi(x,\xi_\nu)}T_\nu f(x)\Big|^{1/2}\Big\},
$$

and the summation in (5.20) is taken over all pairs of $10^4/K_1$-separated subsegments of $\mathcal{L}_a$. Since there are at most K_1^2 such pairs, it suffices to estimate each term in the summation. Take a Schwartz function $\rho\geq 0$ such that $\widehat{\rho}$ is compactly supported and $\rho_a(x):=\rho((x-a)/K)=1$ for all $x\in Q_a$,

$$
\begin{aligned}
(5.21)\qquad & \int_{Q_a\cap\mathfrak{S}_{\mathcal{L}_1,\mathcal{L}_2}} \prod_{j=1}^{2}\Big|\sum_{\Omega_\nu\subset\mathcal{L}_j} e^{i\phi(x,\xi_\nu)}T_\nu f(x)\Big|^2\,dx \\
&\leq \int_{\mathbb{R}^3}\prod_{j=1}^{2}\Big|\rho_a(x)\sum_{\Omega_\nu\subset\mathcal{L}_j} e^{i\phi(x,\xi_\nu)}T_\nu f(x)\Big|^2\,dx \\
&\leq \sum_{\substack{\Omega_{\nu_1},\Omega_{\nu_2}\subset\mathcal{L}_1\\ \Omega_{\nu_1'},\Omega_{\nu_2'}\subset\mathcal{L}_2}} \Big|\int_{\mathbb{R}^3}[\rho_a T_{\nu_1}f\cdot\overline{\rho_a T_{\nu_2}f\cdot\rho_a T_{\nu_1'}f}\cdot\rho_a T_{\nu_2'}f](x)e^{i\Psi(x,\xi_{\nu_1},\xi_{\nu_2},\xi_{\nu_1'},\xi_{\nu_2'})}\,dx\Big|,
\end{aligned}
$$

where

$$
\Psi(x,\xi_{\nu_1},\xi_{\nu_2},\xi_{\nu_1'},\xi_{\nu_2'})=\phi(x,\xi_{\nu_1})-\phi(x,\xi_{\nu_2})-\phi(x,\xi_{\nu_1'})+\phi(x,\xi_{\nu_2'}).
$$

Considering the support of the function

$$
\widehat{\rho_a}*\widehat{T_{\nu_1}f}*\widehat{\rho_a}*\widehat{\overline{T_{\nu_2}f}}*\widehat{\rho_a}*\widehat{\overline{T_{\nu_1'}f}}*\widehat{\rho_a}*\widehat{T_{\nu_2'}f},
$$

we may restrict the summation to those quadruples $\nu_1,\nu_2,\nu_1',\nu_2'$ such that

$$
(5.22)\qquad \begin{cases} |\xi_{\nu_1}-\xi_{\nu_2}-\xi_{\nu_1'}+\xi_{\nu_2'}|\lesssim 1/K,\\ \big||\xi_{\nu_1}|^\alpha-|\xi_{\nu_2}|^\alpha-|\xi_{\nu_1'}|^\alpha+|\xi_{\nu_2'}|^\alpha\big|\lesssim 1/K.\end{cases}
$$

Denote $\ell_a^*=\mathbb{R}v+b$, where $v,b\in\mathbb{R}^2$ with $|v|=1$, $|b|\leq 4$ and $v\perp b$. Since ξ_{ν_1}, ξ_{ν_2}, $\xi_{\nu_1'}$ and $\xi_{\nu_2'}$ lie in the $\alpha 2^\alpha 10^3K_1/K$-neighborhood of ℓ_a^*, there are t_1, t_2, t_1' and t_2' such that for $j=1,2$, we have

$$
|\xi_{\nu_j}-t_jv-b|\leq\alpha 2^\alpha 10^3K_1/K,\qquad |\xi_{\nu_j'}-t_j'v-b|\leq\alpha 2^\alpha 10^3K_1/K.
$$

In view of (5.22) and the $10^4/K_1$-separation of $\mathcal{L}_1$ and $\mathcal{L}_2$, we have

$$|t_1 - t_2| < 2/K_1, \quad |t_1' - t_2'| < 2/K_1, \quad |t_1 - t_1'| > 10^4/K_1, \tag{5.23}$$

$$|t_1 - t_2 - t_1' + t_2'| \lesssim K_1/K, \tag{5.24}$$

$$|\varphi(t_1) - \varphi(t_2) - \varphi(t_1') + \varphi(t_2')| \lesssim K_1/K, \tag{5.25}$$

where $\varphi(t) := (t^2 + |b|^2)^{\alpha/2}$.

We claim that (5.23)–(5.25) imply

$$|t_1 - t_2| \lesssim K_1^\alpha/K, \quad |t_1' - t_2'| \lesssim K_1^\alpha/K. \tag{5.26}$$

As a consequence, we may deduce that

$$|\xi_{\nu_1} - \xi_{\nu_2}| \lesssim K_1^\alpha/K, \quad |\xi_{\nu_1'} - \xi_{\nu_2'}| \lesssim K_1^\alpha/K,$$

which implies

$$(5.21) \lesssim K_1^{4\alpha} \sum_{\Omega_\nu \subset \mathcal{L}_1,\, \Omega_{\nu'} \subset \mathcal{L}_2} \int_{Q_a} |Tf_\nu|^2 |Tf_{\nu'}|^2(x)\, dx.$$

Summing up (5.21) with respect to pairs of $(\mathcal{L}_1, \mathcal{L}_2)$ we obtain, from the local constancy of the $|T_\nu f|(x)$'s,

$$\begin{aligned}\int_{Q_a} (5.8)^4\, dx &\lesssim K_1^{4\alpha+2} \int_{Q_a} \Big(\sum_{\Omega_\nu \subset \mathcal{L}} |Tf_\nu(x)|^2\Big)^2 dx \\ &\lesssim K_1^{4\alpha+2} |Q_a| \Big(\sum_{\Omega_\nu \subset \mathcal{L}} |Tf_\nu(x)|^2\Big)^2, \quad \forall x \in Q_a.\end{aligned}$$

This proves (5.19).

To prove the claim (5.26), we need to consider the following two possibilities:

$$t_1 < t_2 < t_1' < t_2' \quad \text{and} \quad t_1 < t_2 < t_2' < t_1'.$$

In view of (5.24), the second possibility implies

$$|t_1 - t_2| + |t_1' - t_2'| \lesssim K_1/K.$$

Then (5.26) follows immediately. It is thus sufficient to handle the first possibility. To do this, we consider the following three cases.

- $t_1 < t_2 < 0 < t_1' < t_2'$. For this case, from (5.25) we have

$$|\varphi(t_1) - \varphi(t_2)| + |\varphi(t_1') - \varphi(t_2')| \lesssim K_1/K. \tag{5.27}$$

By the triangle inequality and (5.23), we have either $t_1' > 10^3/K_1$ or $t_2 < -10^3/K_1$. We only handle the case when $t_1' > 10^3/K_1$ since the other case is exactly the same. We apply the mean value theorem to $\varphi(t)$ to get a t_* with $t_1' < t_* < t_2'$ such that

$$\frac{K_1}{K} \gtrsim |\varphi(t_1') - \varphi(t_2')| \gtrsim \alpha |t_*|^{\alpha-1} |t_1' - t_2'| \gtrsim \frac{\alpha}{K_1^{\alpha-1}} |t_1' - t_2'|.$$

Hence

$$|t_1' - t_2'| \lesssim_\alpha K_1^\alpha / K.$$

By the triangle inequality again and (5.24), we obtain

$$|t_1 - t_2| \lesssim_\alpha K_1^\alpha / K.$$

• $t_1 < 0 < t_2 < t_1' < t_2'$. First, we always have $t' > 10^3/K_1$ in this case. If $t_1 < -t_2$, we also get (5.27), so (5.26) follows immediately as above. For $-t_2 \le t_1$, from (5.25) we have

$$\big||\varphi(t_1) - \varphi(t_2)| - |\varphi(t_1') - \varphi(t_2')|\big| \lesssim K_1/K. \tag{5.28}$$

Suppose $|t_1 - t_2| \gg K_1^\alpha/K$; then $|t_1' - t_2'| \gg K_1^\alpha/K$ by (5.23). Moreover, we have

$$\frac{1}{2}|t_1' - t_2'| \le |t_1 - t_2| \le 2|t_1' - t_2'|. \tag{5.29}$$

By the mean value theorem, there are $t_* \in (t_1, t_2)$ and $t_*' \in (t_1', t_2')$ with $t_*' > \max\{10^3/K_1, t_* + 10^3/K_1\}$ such that

$$\begin{aligned} \frac{K_1}{K} &\gtrsim |(\varphi(t_2') - \varphi(t_1')) - (\varphi(t_2) - \varphi(t_1))| \\ &\gg \alpha \frac{K_1^\alpha}{K}(|t_*'|^{\alpha-1} - |t_*|^{\alpha-1}) \gg \alpha \frac{K_1}{K}, \end{aligned}$$

which is impossible since $\alpha > 1$ and $K_1 \gg 1$.

• $0 < t_1 < t_2 < t_1' < t_2'$. This can be reduced to the above two cases, and thus the proof of (5.19) is complete. ■

Here, another observation made by Bourgain and Guth [BG] is that as a consequence of (5.19), for $x \in Q_a$ one can define $\psi(x)$ by writing

$$(5.8) + (5.9) = \psi(x)\Big(\sum_{\Omega_\nu \subset \mathcal{L}} |Tf_\nu(x)|^2\Big)^{1/2}.$$

Clearly ψ is non-negative and

$$\bigg(\frac{1}{|Q_a|}\int_{Q_a} \psi(x)^4\,dx\bigg)^{1/4} \lesssim K_1^{2\alpha}. \tag{5.30}$$

To see this is possible, one only needs to apply the local constancy of $T_\nu f(x)$'s so that $\psi(x)$ can be defined on each ball of radius K due to (5.19). Then we glue all the pieces of $\psi(x)$ on the balls together. By an averaging argument and the local constancy of $(5.8) + (5.9)$ and of the functions $T_\nu f(x)$ on each $Q_{a,K}$ box, we may assume $\psi(x)$ is constant on unit cubes centered at lattice points.

Remark 5.4. By writing (5.8) + (5.9) as a product of an appropriate ψ and a square function, we may iterate this part step by step in the subsequent context to generate the items having transversality structures corresponding to all the dyadic scales. This is one of the brilliant ideas due to Bourgain and Guth, which is applied in [B3] and [BG] as a substitution of Wolff's induction on scale technique.

Substituting (5.8) + (5.9) in Lemma 5.1 for

$$\psi(x)\Big(\sum_{\Omega_\nu\subset\mathcal{L}}|Tf_\nu(x)|^2\Big)^{1/2},$$

we obtain

$$|Tf(x)| \lesssim K^8 \max_{\substack{\nu_1,\nu_2,\nu_2\\ \text{non-collinear}}}\Big(\prod_{j=1}^{3}|T_{\nu_j}f(x)|\Big)^{1/3} \tag{5.31}$$

$$+\,\psi(x)\Big(\sum_{\Omega_\nu\subset\mathcal{L}}|Tf_\nu(x)|^2\Big)^{1/2} \tag{5.32}$$

$$+\max_\mu\Big|\int_{\tilde\Omega_\mu} e^{i\phi(x,\xi)}\hat f(\xi)\,d\xi\Big|. \tag{5.33}$$

Now, we are ready to prove (5.15)–(5.18). Observe that $Tf(x)$ is controlled in terms of Tf_ν's with $\widehat{f_\nu}$ supported in a square of size $1/K$, whereas $\widehat{f}$ is supported in a region of size 1. Thus it is natural to scale each $\widehat{f_\nu}$ to be a function $\widehat{g}_\nu$ such that $\operatorname{supp}\widehat{g}_\nu$ is of size 1. After applying (5.31)–(5.33) to each Tg_ν, we rescale the estimates on Tg_ν back to the original size $1/K$. More generally, this process can be carried out with Tf_τ in place of Tf on the left side of (5.31), where $\widehat{f_\tau}$ is supported in a square of size δ.

Proof of (5.15)–(5.18). Let $\widehat{f_\tau}=\widehat{f}|_{\Omega_\tau}$ and $x\in\mathcal{T}_a^*=a+Q_{0,K}^\delta$, with

$$Q_{0,K}^\delta=\{x\mid(\delta x_1,\delta x_2,\delta^\alpha x_3)\in Q_{0,K}\}.$$

Making the change of variables

$$x'=a'+\tilde x'/\delta,\qquad x_3=a_3+\tilde x_3/\delta^\alpha,\qquad \xi=\xi_\tau+\delta\eta, \tag{5.34}$$

where ξ_τ is the center of Ω_τ, and $\tilde x\in Q_{0,K}$, we have

$$\begin{aligned}\chi_{\mathcal{T}_a^*}(x)Tf_\tau(x)&=\chi_{Q_{0,K}}(\tilde x)\int_\Omega e^{i[\tilde x'\cdot\eta+\tilde x_3|\eta|^\alpha]}\widehat{g_a^{\tau,\delta}}(\eta)\,d\eta\\&=\chi_{Q_{0,K}}(\tilde x)T(g_a^{\tau,\delta})(\tilde x),\end{aligned} \tag{5.35}$$

with

$$\widehat{g_a^{\tau,\delta}}(\eta)$$
$$=e^{i[\delta a'\cdot\eta+a_3\delta^\alpha|\eta|^\alpha+(a'+\tilde x'/\delta)\cdot\xi_\tau+(a_3+\tilde x_3/\delta^\alpha)(|\xi_\tau+\delta\eta|^\alpha-|\delta\eta|^\alpha)]}\delta^2\widehat{f_\tau}(\xi_\tau+\delta\eta)\chi_\Omega(\eta).$$

Now that $\widehat{g_a^{\tau,\delta}}$ is supported in a square of size 1, we can apply (5.31)–(5.33) to (5.35) with $\tilde{x} \in Q_{0,K}$ to obtain

$$|T(g_a^{\tau,\delta})(\tilde{x})| \lesssim K^8 \max_{\substack{\nu_1,\nu_2,\nu_3 \\ \text{non-collinear}}} \prod_{j=1}^{3} |T_{\nu_j}(g_a^{\tau,\delta})(\tilde{x})|^{1/3} \tag{5.36}$$

$$+ \psi(\tilde{x})\Big(\sum_{\Omega_\nu \subset \mathcal{L}} |T_\nu(g_a^{\tau,\delta})(\tilde{x})|^2\Big)^{1/2} \tag{5.37}$$

$$+ \max_{\mu} \Big| \int_{\widetilde{\Omega}_\mu} e^{i\phi(\tilde{x},\xi)} \widehat{g_a^{\tau,\delta}}(\eta)\, d\eta \Big|. \tag{5.38}$$

Rescaling the $\widehat{g_a^{\tau,\delta}}$ in (5.36)–(5.38) back to $\widehat{f_\tau}$ by using (5.34) and setting $\eta = (\zeta - \xi_\tau)/\delta$ we obtain

$$\begin{aligned} |\chi_{Q_{0,K}}(\tilde{x}) T_\nu(g_a^{\tau,\delta})(\tilde{x})| &= \Big| \int_{\|\eta-\xi_\nu\|<1/K} e^{i[\tilde{x}'\cdot\eta + \tilde{x}_3|\eta|^\alpha]} \widehat{g_a^{\tau,\delta}}(\eta)\, d\eta \Big| \\ &= \Big| \int_{\|\zeta-(\xi_\tau+d\xi_\nu)\|<\delta/K} e^{i[\zeta\cdot(a'+\tilde{x}/\delta)+(a_3+\tilde{x}_3/\delta^\alpha)|\zeta|^\alpha]} \hat{f}_\tau(\zeta)\, d\zeta \Big| \\ &= \Big| \int_{\|\zeta-(\xi_\tau+d\xi_\nu)\|<\delta/K} e^{i[\zeta\cdot x' + x_3|\zeta|^\alpha]} \hat{f}_\tau(\zeta)\, d\zeta \Big|, \qquad x \in \mathcal{T}_a^*. \end{aligned} \tag{5.39}$$

From (5.35), (5.36) and (5.39), we get (5.15)–(5.17) on $\mathcal{T}_a^*$.

Since

$$\psi_\tau(x) = \psi\big(\delta(x_1 - a_1), \delta(x_2 - a_2), \delta^\alpha(x_3 - a_3)\big), \qquad \forall x \in \mathcal{T}_a^*,$$

we obtain (5.18) from (5.30). ∎

5.3. Iteration and the end of proof. This part follows closely the ideas of Bourgain and Guth [BG]; however, we provide more explicit calculations during the iteration process so that this marvelous idea can be grasped even by novice readers. It is hoped that this robust machine will be upgraded, so that further improvements seem possible in this area.

Let $1 \ll K_1 \ll K \ll R$. From Lemma 5.1, we have

$$|Tf(x)| \lesssim K^8 \max_{\substack{\Omega_{\tau_1},\Omega_{\tau_2},\Omega_{\tau_3}:\,1/K\text{-cubes} \\ \text{non-collinear}}} \prod_{j=1}^{3} |T_{\tau_j} f(x)|^{1/3} \tag{5.40}$$

$$+ \psi(x)\Big[\sum_{\substack{\Omega_\tau \subset \mathcal{L} \\ \Omega_\tau:\,1/K\text{-cubes}}} |T_\tau f(x)|^2\Big]^{1/2} \tag{5.41}$$

$$+ \max_{\tilde{\Omega}_{\tilde{\tau}}:\,1/K_1\text{-cubes}} |T_{\tilde{\tau}} f(x)|, \tag{5.42}$$

where ψ is approximately constant on unit boxes and obeys

$$\Big(\frac{1}{|Q_a|}\int_{Q_a}\psi(x)^4\,dx\Big)^{1/4}\lesssim K_1^{2\alpha}$$

for any Q_a.

Noting that (5.40) involves a triple product of $|T_{\tau_j}f|^{1/3}$ with $j=1,2,3$, we say this term is *of type I.* The term (5.41) is a product of a suitable function ψ and an ℓ^2 sum of $\{T_\tau f\}_\tau$, and we call it a term *of type II.* The term (5.42) is an ℓ^∞ norm of $\{T_{\tilde\tau}f\}_{\tilde\tau}$, and we call it *of type III.* In each step of the iteration below, we will encounter plenty of terms belonging to type I, II and III from the previous step. These are called *newborn* terms. We add the newborn terms of type I to the type I terms of the previous generations and keep on iterating all the newborn terms of type II and III to get the next generation of type I, II and III terms. This is the iterating mechanism.

To be more precise, we apply (5.15)–(5.17) to $T_\tau f$ in (5.41) with $\delta=1/K$, and to $T_{\tilde\tau}f$ in (5.42) with $\delta=1/K_1$. This is exactly the first step of the iteration. After this, we obtain terms of type I, II and III generated by (5.41) and (5.42). In each type of the terms, the supports of $\widehat{f_\tau}$'s could be of the scales like

$$\frac{1}{K^2},\qquad\frac{1}{KK_1}\quad\text{or}\quad\frac{1}{K_1^2}.$$

Adding the newborn terms of type I to the previous type I terms, we repeat the same argument as in the first step to all the terms of type II and type III to get the second generation. This process is continued with newborn terms of type I added to the previous type I terms until the scale of the support of $\widehat{f_\tau}$ in the terms of type II and type III becomes $1/\sqrt{R}$. Finally, we obtain a collection of type I terms at different scales and a remainder consisting of type II and III terms at scale $1/\sqrt{R}$, which is controlled by (2.3). This yields (2.2) and (2.3).

Now, we present the explicit computation for the first step. Applying (5.15)–(5.17) to (5.41) with $\delta=1/K$, by Minkowski's inequality we have

$$\psi(x)\Big(\sum_{\Omega_\tau\subset\mathcal{L}}|T_\tau f(x)|^2\Big)^{1/2}$$

$$\lesssim K^8\Big[\sum_{\Omega_\tau\subset\mathcal{L}}\Big(\max_{\substack{\Omega_\tau\supset\Omega_{\tau_1^{(1)}},\Omega_{\tau_2^{(1)}},\Omega_{\tau_3^{(1)}}(x)\\ \text{non-collinear }1/K^2\text{-squares}}}\psi\prod_{j=1}^{3}|T_{\tau_j^{(1)}}f|^{1/3}\Big)^2\Big]^{1/2}(x)\tag{5.43}$$

$$+\Big[\sum_{\Omega_\tau\subset\mathcal{L}}\ \sum_{\substack{\mathcal{L}^{(1)}\supset\Omega_{\tau^{(1)}}\\ 1/K^2\text{-squares}}}(\psi\psi_\tau|T_{\tau^{(1)}}f|)^2\Big]^{1/2}(x) \tag{5.44}$$

$$+\Big(\sum_{\Omega_\tau\subset\mathcal{L}}\psi^2\max_{\substack{\tilde\Omega_{\tilde\tau^{(1)}}\subset\Omega_\tau\\ 1/(K_1K)\text{-squares}}}|T_{\tilde\tau^{(1)}}f|^2\Big)^{1/2}(x), \tag{5.45}$$

where the superscript in $\tau^{(k)}$ denotes the kth step of the iteration.

Denoting $\psi_{:\tau^{(1)}:}=\psi\psi_\tau$, we need to verify that for any $\varepsilon>0$,

$$\frac{1}{|\mathcal{C}_\tau^*|}\int_{\mathcal{C}_\tau^*}\psi(x)^4\,dx\lesssim R^\varepsilon, \tag{5.46}$$

$$\frac{1}{|\mathcal{C}_{\tau^{(1)}}^*|}\int_{\mathcal{C}_{\tau^{(1)}}^*}\psi_{:\tau^{(1)}:}(x)^4\,dx\lesssim R^\varepsilon. \tag{5.47}$$

To get (5.46), we use the boxes Q_a to subdivide $\mathcal{C}_\tau^*$ so that

$$\mathcal{C}_\tau^*\subset\bigcup_a Q_a\subset 2\mathcal{C}_\tau^*.$$

Then (5.30) gives

$$\begin{aligned}\frac{1}{|\mathcal{C}_\tau^*|}\int_{\mathcal{C}_\tau^*}\psi(x)^4\,dx&\lesssim\frac{1}{|\bigcup_a Q_a|}\int_{\bigcup Q_a}\psi(x)^4\,dx\\&\lesssim\max_{\mathcal{C}_\tau^*\subset Q_a\subset 2\mathcal{C}_\tau^*}\frac{1}{|Q_a|}\int_{Q_a}\psi(x)^4\,dx\lesssim K_1^{8\alpha}\ll R^\varepsilon.\end{aligned}$$

To verify (5.47), we note that $\mathcal{C}_{\tau^{(1)}}^*$ is a $K^2\times K^2\times K^{2\alpha}$-box in the direction of the normal vector of the surface at $\xi_{\tau^{(1)}}$. It follows that $\mathcal{C}_{\tau^{(1)}}^*$ can be covered as follows:

$$\mathcal{C}_{\tau^{(1)}}^*\subset\bigcup_\tau B_\tau^*\subset 2\mathcal{C}_{\tau^{(1)}}^*$$

where B_τ^* is a $K\mathcal{C}_\tau^*$-box. Then

$$\frac{1}{|\mathcal{C}_{\tau^{(1)}}^*|}\int_{\mathcal{C}_{\tau^{(1)}}^*}\psi^4\psi_\tau^4(x)\,dx\lesssim\max_{B_\tau^*\subset 2\mathcal{C}_{\tau^{(1)}}^*}\frac{1}{|B_\tau^*|}\int_{B_\tau^*}\psi^4\psi_\tau^4(x)\,dx. \tag{5.48}$$

To estimate the right side, we let $\{\mathcal{B}_\rho\}_\rho$ be a collection of essentially disjoint $\mathcal{C}_\tau^*$-boxes such that

$$B_\tau^*\subset\bigcup_\rho\mathcal{B}_\rho\subset 2B_\tau^*.$$

Since ψ_τ is approximately constant on each $\mathcal{B}_\rho$, we have

$$(5.49)\qquad \int_{B_\tau^*} \psi(x)^4\psi_\tau(x)^4\,dx \lesssim \sum_\rho \Big[\int_{\mathcal{B}_\rho} \psi(x)^4\,dx\Big](\psi_\tau|_{\mathcal{B}_\rho})^4, \\ \lesssim K_1^{8\alpha}\sum_\rho (\psi_\tau|_{\mathcal{B}_\rho})^4|\mathcal{B}_\rho| \lesssim K_1^{8\alpha}\int_{B_\tau^*}\psi_\tau(x)^4\,dx \\ \lesssim K_1^{16\alpha}|\mathcal{C}_\tau^*| \ll R^\varepsilon|\mathcal{C}_\tau^*|,$$

where we have used (5.18), (5.46) and the fact that $\mathcal{B}_\rho$ is a $\mathcal{C}_\tau^*$-box. This along with (5.48) proves (5.47).

Next, we apply (5.15)–(5.17) to (5.42) with $\delta = 1/K_1$, and obtain

$$(5.50)\qquad |(5.42)| \lesssim \max_{\tilde{\Omega}_\tau:\,1/K_1\text{-squares}} K^8 \max_{\substack{\tilde{\Omega}_{\tilde{\tau}}\supset\Omega_{\tilde{\tau}_1^{(1)}},\Omega_{\tilde{\tau}_2^{(1)}},\Omega_{\tilde{\tau}_3^{(1)}}\\ \text{non-collinear } 1/(K_1K)\text{-squares}}} \prod_{j=1}^{3}|T_{\tilde{\tau}_j^{(1)}}f|^{1/3}$$

$$(5.51)\qquad +\max_{\tilde{\Omega}_\tau:\,1/K_1\text{-squares}}\Bigg(\sum_{\substack{\Omega_{\tilde{\tau}^{(1)}}\subset\mathcal{L}^{(1)}\cap\tilde{\Omega}_{\tilde{\tau}}\\ 1/(K_1K)\text{-squares}}}\psi_{\tilde{\tau}}^2|T_{\tilde{\tau}^{(1)}}f|^2\Bigg)^{1/2}$$

$$(5.52)\qquad +\max_{\tilde{\Omega}_\tau:\,1/K_1\text{-squares}}\ \max_{\substack{\tilde{\Omega}_\tau\supset\tilde{\Omega}_{\tau^{(1)}}\\ \tilde{\Omega}_{\tau^{(1)}}:\,1/K_1^2\text{-squares}}}|T_{\tilde{\tau}^{(1)}}f|.$$

We also have estimates on the L^4-average akin to (5.46) and (5.47).

Remark 5.5. After the first step, if we already have

$$K^2 \sim KK_1 \sim \sqrt{R}$$

then Lemma 2.4 is proved. However, this is not the case since $1 \ll K_1 \ll K \ll R$. Therefore we have to use the above argument recursively to deduce Lemma 2.4.

Noting that the scale for type I terms at the kth generation is $K^{-m_1}K_1^{-m_2}$ with $m_1, m_2 \in \mathbb{Z}$, $m_1+m_2 = k+1$, $m_1, m_2 \geq 0$, we find the type I terms of the kth stage from the previous $(k-1)$th stage are dominated by a k-fold sum

$$\Bigg(\sum_{\Omega_\tau\subset\mathcal{L}}\sum_{\Omega_{\tau^{(1)}}\subset\mathcal{L}^{(1)}}\cdots\sum_{\Omega_{\tau^{(k)}}\subset\mathcal{L}^{(k)}}\max_{\substack{\Omega_{\tau_j^{(k+1)}}\subset\Omega_{\tau^{(k)}},\,j=1,2,3\\ \Omega_{\tau_j^{(k+1)}}:\,1/K^{k+1}\text{-squares}\\ \text{non-collinear}}}\psi^2_{:\tau^{(k)}:}\Bigg(\prod_{j=1}^{3}|T_{\tau_j^{(k+1)}}f|\Bigg)^{2/3}\Bigg)^{1/2} + \sum_{\text{mixed scales}},$$

where the summation over mixed scales represents the cases when $m_2 \geq 1$. For brevity, we only write out the case when $m_2 = 0$ explicitly; the other cases are similar. Noticing that in each fold of the sum, there are at most

KK_1 terms involved, the above expression can be controlled by

$$C(K)\max_{\mathcal{E}_{(1/K)^k}}\Big[\sum_{\Omega_\tau^{(k)}\in\mathcal{E}_{(1/K)^{k(1+\varepsilon)}}}\Big(\psi_{:\tau(k):}\prod_{j=1}^{3}|T_{\tau_j^{(k+1)}}f|^{1/3}\Big)^2\Big]^{1/2}+\sum_{\text{mixed scales}},$$

where we have adopted the notation of Lemma 2.4.

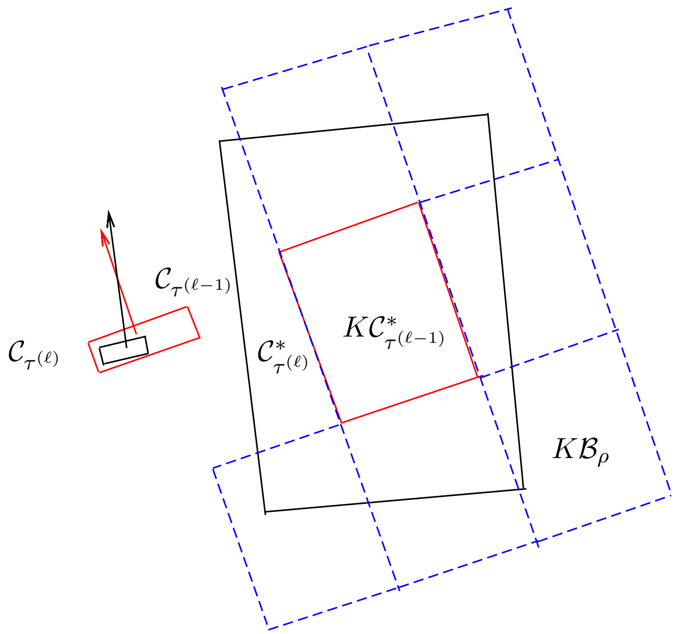

Fig. 3. The boxes $\mathcal{C}^*_{\tau(\ell)}$ and $K\mathcal{C}^*_{\tau(\ell-1)}$

It remains to show

$$\frac{1}{|\mathcal{C}^*_{\tau(k)}|}\int_{\mathcal{C}^*_{\tau(k)}}\psi_{:\tau(k):}(x)^4\,dx\lesssim R^\varepsilon\quad\text{ for any }\varepsilon>0.$$

To prove this, we use induction on k. We observe that the L^4-average of $\psi_{:\tau(1):}$ over $\mathcal{C}^*_{\tau(1)}$ is bounded by $K_1^{4\alpha}$ in the first step. Assuming the $(\ell-1)$th stage, we already have

$$\Big(\frac{1}{|\mathcal{C}^*_{\tau(\ell-1)}|}\int_{\mathcal{C}^*_{\tau(\ell-1)}}\psi^4_{:\tau(\ell-1):}\Big)^{1/4}\lesssim K_1^{2\ell\alpha}. \tag{5.53}$$

Since $\psi_{:\tau(\ell):}=\psi_{:\tau(\ell-1):}\psi_{\tau(\ell)}$, we need to evaluate

$$\frac{1}{|\mathcal{C}^*_{\tau(\ell)}|}\int_{\mathcal{C}^*_{\tau(\ell)}}\psi^4_{:\tau(\ell-1):}\psi^4_{\tau(\ell)}(x)\,dx. \tag{5.54}$$

Because the angle between the two normal vectors of $\mathcal{C}_{\tau(\ell-1)}$ at $\xi_{\tau(\ell-1)}$ and $\mathcal{C}_{\tau(\ell)}$ at $\xi_{\tau(\ell)}$ is also controlled by $1/K$ (see Figure 3), we may construct a cover of $\mathcal{C}^*_{\tau(\ell)}$ by $K\mathcal{C}^*_{\tau(\ell-1)}$-boxes as follows. Denote by $\{\mathcal{B}_\rho\}_\rho$ a collection of $\mathcal{C}^*_{\tau(\ell-1)}$-boxes such that (see Figure 3)

$$\mathcal{C}^*_{\tau(\ell)} \subset \bigcup_\rho K\mathcal{B}_\rho \subset 2\,\mathcal{C}^*_{\tau(\ell)}.$$

On account of this, we can estimate

$$(5.54) \lesssim \max_{\mathcal{B}_\rho} \frac{1}{|K\mathcal{B}_\rho|} \int_{K\mathcal{B}_\rho} \psi^4_{:\tau(\ell-1):}\psi^4_{\tau(\ell)}(x)\,dx.$$

By the induction hypothesis, we have

$$\begin{aligned}\int_{K\mathcal{B}_\rho} \psi_{:\tau(\ell-1):}(x)^4\psi_{\tau(\ell)}(x)^4\,dx &\lesssim \sum_\rho (\psi_{\tau(\ell)}|_{\mathcal{B}_\rho})^4 \int_{\mathcal{B}_\rho} \psi_{:\tau(\ell-1):}(x)^4\,dx \\ &\lesssim K_1^{8\alpha\ell} \sum_\rho (\psi_{\tau(\ell)}|_{\mathcal{B}_\rho})^4 |\mathcal{B}_\rho| \\ &\lesssim K_1^{8\alpha\ell} \int_{K\mathcal{B}_\rho} \psi_{\tau(\ell)}(x)^4\,dx \lesssim K_1^{8\alpha(\ell+1)}|K\mathcal{B}_\rho|,\end{aligned}$$

where in the last estimate we have used

$$\frac{1}{|K\mathcal{C}^*_{\tau(\ell-1)}|}\int_{K\mathcal{C}^*_{\tau(\ell-1)}} \psi_{\tau(\ell)}(x)^4\,dx \lesssim K_1^{8\alpha}.$$

We denote

$$\delta = K^{-(\ell+1)},$$

and assume at the ℓth stage

$$1/\sqrt{R} < \delta.$$

Noting that $K_1 \ll K$ and

$$\ell + 1 = \log(1/\delta)/\log K,$$

we have

$$K_1^{8\alpha(\ell+1)} < R^{\frac{\log K_1^{4\alpha}}{\log K}} \ll R^\varepsilon, \qquad \forall \varepsilon > 0.$$

If δ ranges over all dyadic numbers between $R^{-1/2}$ and K^{-1}, we see the contribution from all the type I terms is bounded by (2.2). The contributions from (5.45), (5.51) and (5.52) to (2.2) can be evaluated in a similar manner.

When the scale reaches $1/\sqrt{R}$, the remainder term is bounded by (2.3). Finally, we lose an R^ε-factor by taking the maximum in (2.2) and (2.3) with respect to dyadic $\delta \in (R^{-1/2}, 1/K)$. This completes the proof of Lemma 2.4.

Acknowledgements. We thank Professor Sanghyuk Lee for indicating the recent work [CLV]. The authors were supported by the NSF of China under grants No. 11231006 and 11371059.

References

[BCT] J. Bennett, T. Carbery and T. Tao, *On the multilinear restriction and Kakeya conjectures*, Acta Math. 196 (2006), 261–302.

[B1] J. Bourgain, *A remark on Schrödinger operators*, Israel J. Math. 77 (1992), 1–16.

[B2] J. Bourgain, *Some new estimates on oscillatory integrals*, in: Essays on Fourier Analysis in Honor of Elias M. Stein (Princeton, NJ, 1991), Princeton Math. Ser. 42, Princeton Univ. Press, 1995, 83–112.

[B3] J. Bourgain, *On the Schrödinger maximal function in higher dimensions*, Proc. Steklov Inst. Math. 280 (2013), 46–60.

[BG] J. Bourgain and L. Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. 21 (2011), 1239–1295.

[Cr] L. Carleson, *Some analytic problems related to statistical mechanics*, in: Euclidean Harmonic Analysis, Lecture Notes in Math. 779, Springer, Berlin, 1979, 5–45.

[CLV] C. Cho, S. Lee and A. Vargas, *Problems on pointwise convergence of solutions to the Schrödinger equation*, J. Fourier Anal. Appl. 18 (2012), 972–994.

[Co] A. Córdoba, *Geometric Fourier analysis*, Ann. Inst. Fourier 32 (1982), no. 3, 215–226.

[DK] B. Dahlberg and C. E. Kenig, *A note on the almost everywhere behavior of solutions to the Schrödinger equations*, in: Harmonic Analysis, Lecture Notes in Math. 908, Springer, Berlin, 1982, 205–209.

[L] S. Lee, *On pointwise convergence of the solutions to Schrödinger equations in* $\mathbb{R}^2$, Int. Math. Res. Notices 2006, art. ID 32597, 21 pp.

[LR] S. Lee and K. M. Rogers, *The Schrödinger equation along curves and the quantum harmonic oscillator*, Adv. Math. 229 (2012), 1359–1379.

[M] A. Miyachi, *On some singular Fourier multipliers*, J. Fac. Sci. Univ. Tokyo Sec. IA Math. 28 (1981), 267–315.

[MVV] A. Moyua, A. Vargas and L. Vega, *Schrödinger maximal function and restriction properties of the Fourier transform*, Int. Math. Res. Notices 1996, 793–815.

[R] K. Rogers, *A local smoothing estimate for the Schrödinger equation*, Adv. Math. 219 (2008), 2105–2122.

[Sh] S. Shao, *On localization of the Schrödinger maximal operator*, arXiv:1006.2787v1 (2010).

[S] P. Sjölin, *Regularity of solutions to the Schrödinger equation*, Duke Math. J. 55 (1987), 699–715.

[T] T. Tao, *Time-frequency harmonic analysis*, http://www.math.ucla.edu/~tao/254a.1.01w/blurb.html.

[TV1] T. Tao and A. Vargas, *A bilinear approach to cone multipliers. I. Restriction estimates*, Geom. Funct. Anal. 10 (2000), 185–215.

[TV2] T. Tao and A. Vargas, *A bilinear approach to cone multipliers. II. Applications*, Geom. Funct. Anal. 10 (2000), 216–258.

[TVV] T. Tao, A. Vargas and L. Vega, *A bilinear approach to the restriction and Kakeya conjectures*, J. Amer. Math. Soc. 11 (1998), 967–1000.

[V] L. Vega, *Schrödinger equations: pointwise convergence to the initial data*, Proc. Amer. Math. Soc. 102 (1988), 874–878.
[W] T. Wolff, *A sharp bilinear cone restriction estimate*, Ann. of Math. (2) 153 (2001), 661–698.

Changxing Miao
Institute of Applied Physics
and Computational Mathematics
100088 Beijing, China
E-mail: miao_changxing@iapcm.ac.cn

Jianwei Yang
Beijing International Center
for Mathematical Research
Peking University
100871 Beijing, China
E-mail: geewey_young@pku.edu.cn

Jiqiang Zheng
Université Nice Sophia-Antipolis
06108 Nice Cedex 02, France
E-mail: zhengjiqiang@gmail.com
zheng@unice.fr