SHARP $L^2$ ESTIMATE OF SCHRÖDINGER MAXIMAL FUNCTION IN HIGHER DIMENSIONS

XIUMIN DU AND RUIXIANG ZHANG

Dedicated to Jean Bourgain with admiration

Abstract. We show that, for $n \geq 3$, $\lim_{t \to 0} e^{it\Delta} f(x) = f(x)$ holds almost everywhere for all $f \in H^s(\mathbb{R}^n)$ provided that $s > \frac{n}{2(n+1)}$. Due to a counterexample by Bourgain, up to the endpoint, this result is sharp and fully resolves a problem raised by Carleson. Our main theorem is a fractal $L^2$ restriction estimate, which also gives improved results on the size of divergence set of Schrödinger solutions, the Falconer distance set problem and the spherical average Fourier decay rates of fractal measures. The key ingredients of the proof include multilinear Kakeya estimates, decoupling and induction on scales.

1. Introduction

The solution to the free Schrödinger equation

$$
\begin{cases}
    iu_t - \Delta u = 0, & (x, t) \in \mathbb{R}^n \times \mathbb{R} \\
    u(x, 0) = f(x), & x \in \mathbb{R}^n
\end{cases}
$$

is given by

$$
e^{it\Delta} f(x) = (2\pi)^{-n} \int e^{i(x \cdot \xi + t|\xi|^2)} \hat{f}(\xi) \, d\xi.
$$

In [8], Carleson proposed the problem of identifying the optimal $s$ for which $\lim_{t \to 0} e^{it\Delta} f(x) = f(x)$ almost everywhere whenever $f \in H^s(\mathbb{R}^n)$, and proved convergence for $s \geq \frac{1}{4}$ when $n = 1$. Dahlberg and Kenig [10] then showed that this result is sharp. The higher dimensional case has since been studied by several authors [7, 9, 28, 30, 2, 26, 29, 20, 3, 23, 11, 4, 24, 12, 13]. In particular, almost everywhere convergence holds if $s > \frac{1}{2} - \frac{1}{4n}$ when $n = 2$ due to Lee [20] and $n \geq 2$ due to Bourgain [3]). Recently Bourgain [4] gave counterexamples showing that convergence can fail if $s < \frac{n}{2(n+1)}$. Since then, Guth, Li and the first author [12] improved the sufficient condition when $n = 2$ to the almost sharp range $s > \frac{1}{4}$. In higher dimensions ($n \geq 3$), Guth, Li and the authors [13] proved the convergence for $s > \frac{n+1}{2(n+2)}$.

In this article, we establish the following theorem, which is sharp up to the endpoint.

Theorem 1.1. Let $n \geq 1$. For every $f \in H^s(\mathbb{R}^n)$ with $s > \frac{n}{2(n+1)}$, $\lim_{t \to 0} e^{it\Delta} f(x) = f(x)$ almost everywhere.

We use $B^m(x, r)$ to represent a ball centered at $x$ with radius $r$ in $\mathbb{R}^m$. By a standard smooth approximation argument, Theorem 1.1 is a consequence of the following estimate of Schrödinger maximal function:
Remark 1.4. Our results above are new in the case (1.4), argument using wave packet decomposition. Following two problems are still open when previous sharp $L^p$ estimates of Schrödinger maximal function are not as strong as the previous sharp $L^p$ estimates:

\begin{equation}
\left\| \sup_{0<t\leq 1} \left| e^{it\Delta} f \right| \right\|_{L^2(B^n(0,1))} \leq C_s \|f\|_{H^s(R^n)}.
\end{equation}

Via a localization argument, Littlewood-Paley decomposition and parabolic rescaling, Theorem 1.2 is reduced to the following theorem which we will prove in this paper:

**Theorem 1.3.** Let $n \geq 1$. For any $\epsilon > 0$, there exists a constant $C_\epsilon$ such that

\begin{equation}
\left\| \sup_{0<t\leq R} \left| e^{it\Delta} f \right| \right\|_{L^2(B^n(0,R))} \leq C_\epsilon R^{\frac{n}{2} + \epsilon} \|f\|_2
\end{equation}

holds for all $R \geq 1$ and all $f$ with $\text{supp} \hat{f} \subset A(1) = \{ \xi \in R^n : |\xi| \sim 1 \}$.

**Remark 1.4.** Our results above are new in the case $n \geq 3$. When $n = 1, 2$, despite the fact that we recover the almost sharp results of pointwise convergence problem, our sharp $L^2$ estimates of Schrödinger maximal function are not as strong as the previous sharp $L^p$ estimates:

\begin{equation}
\left\| \sup_{t>0} \left| e^{it\Delta} f \right| \right\|_{L^4(R)} \leq C \|f\|_{H^{1/4}(R)} , \quad [19, \text{Kenig-Ponce-Vega}],
\end{equation}

and

\begin{equation}
\left\| \sup_{0<t\leq 1} \left| e^{it\Delta} f \right| \right\|_{L^3(R^2)} \leq C_s \|f\|_{H^s(R^2)}, \forall s > \frac{1}{3}, \quad [12, \text{D.-Guth-Li}].
\end{equation}

From the same examples as what one has for the restriction conjecture, it seems reasonable to have a similar conjecture in higher dimensions - it would be interesting to ask if the following holds for general $n$:

\begin{equation}
\left\| \sup_{0<t\leq 1} \left| e^{it\Delta} f \right| \right\|_{L^{\frac{2(n+1)}{n}}(R^n)} \leq C \|f\|_{H^\infty(R^n)}.
\end{equation}

From (1.4) and (1.5) we see that (1.6) is true for $n = 1$, and is true up to the endpoint for $n = 2$. However, the estimate (1.6) fails in higher dimensions. During the MRC week in June 2018, Jongchon Kim, Hong Wang and the authors worked together on the following local estimates:

\begin{equation}
\left\| \sup_{0<t\leq 1} \left| e^{it\Delta} f \right| \right\|_{L^p(B^n(0,1))} \leq C_s \|f\|_{H^s(R^n)}, \forall s > \frac{n}{2(n+1)},
\end{equation}

and by looking at Bourgain’s counterexample [3] in every intermediate dimension, the working group observed that (1.7) fails if $p > p_0 := 2 + \frac{4}{(n-1)(n+2)}$. Note that $2(n+1) > p_0$ when $n \geq 3$ and henceforth (1.6) fails. To our best knowledge, the following two problems are still open when $n \geq 3$: determine the optimal $p = p(n)$ for which we can have (1.7) and identify the optimal $s = s(n,p)$ for which (1.7) with $p > 2$ fixed holds. Our working group at MRC had some partial results and planned to deal with these problems in an upcoming paper.

*The global $L^3$ estimate (1.5) follows easily from the local $L^3$ estimate in [12], via a localization argument using wave packet decomposition.
Our main result is the following fractal $L^2$ restriction estimate, from which Theorem 1.3 follows.

**Theorem 1.5.** Let $n \geq 1$. For any $\varepsilon > 0$, there exists constant $C_\varepsilon$ such that the following holds for all $R \geq 1$ and all $f$ with supp $\widehat{f} \subset B^n(0,1)$. Suppose that $X = \bigcup_k B_k$ is a union of lattice unit cubes in $B^{n+1}(0,R)$ and each lattice $R^{1/2}$-cube intersecting $X$ contains $\sim \lambda$ many unit cubes in $X$. Let $1 \leq \alpha \leq n+1$ and $\gamma$ be given by

$$\gamma := \max_{B^{n+1}(x',r) \subset B^{n+1}(0,R)} \frac{\# \{B_k : B_k \subset B(x',r) \}}{r^\alpha}.$$  

Then

$$\|e^{it\Delta} f\|_{L^2(X)} \leq C_\varepsilon \gamma^{\frac{n+1}{n+\alpha}} R^{\frac{n}{2(n+1)} + \varepsilon} \|f\|_2.$$  

Note that in Theorem 1.5, $\lambda \leq \gamma R^{\alpha/2}$. As a direct result of Theorem 1.5, there holds a slightly weaker fractal $L^2$ restriction estimate. It has a relatively simpler statement:

**Corollary 1.6.** Let $n \geq 1$. For any $\varepsilon > 0$, there exists constant $C_\varepsilon$ such that the following holds for all $R \geq 1$ and all $f$ with supp $\widehat{f} \subset B^n(0,1)$. Suppose that $X = \bigcup_k B_k$ is a union of lattice unit cubes in $B^{n+1}(0,R)$. Let $1 \leq \alpha \leq n+1$ and $\gamma$ be given by

$$\gamma := \max_{B^{n+1}(x',r) \subset B^{n+1}(0,R)} \frac{\# \{B_k : B_k \subset B(x',r) \}}{r^\alpha}.$$  

Then

$$\|e^{it\Delta} f\|_{L^2(X)} \leq C_\varepsilon \gamma^{\frac{n+1}{n+\alpha}} R^{\frac{n}{2(n+1)} + \varepsilon} \|f\|_2.$$  

We will see that Corollary 1.6 is sufficient to derive the sharp $L^2$ estimate of Schrödinger maximal function (Theorem 1.3) and all other applications in Section 2. This corollary can also be proved directly by a slightly simpler argument. The case $n = 1$ of Corollary 1.6 can be recovered using the ingredients in Wolff’s paper [31]. See Subsection 3.3 for a discussion.

Nevertheless, Theorem 1.5 has two advantages compared to Corollary 1.6. Firstly, it gives us a better $L^2$ restriction estimate if the set $X$ of unit cubes is fairly sparse. Secondly, it tells us some geometric information about a set $X$ of unit cubes when $\|e^{it\Delta} f\|_{L^2(X)}$ is comparable to $\|e^{it\Delta} f\|_{L^2(B(0,R))}$. For example, taking $\alpha = n+1$ (hence $\gamma \lesssim 1$) we have:

**Corollary 1.7.** Let $n \geq 1$. Suppose that $X = \bigcup_k B_k$ is a union of lattice unit cubes in $B^{n+1}(0,R)$ and each lattice $R^{1/2}$-cube intersecting $X$ contains $\sim \lambda$ many unit cubes in $X$. Suppose there is a function $f$ with supp $\widehat{f} \subset B^n(0,1)$ and $\|f\|_2 \neq 0$ such that $\|e^{it\Delta} f\|_{L^2(X)} \gtrsim R^{1/2} \|f\|_2$. Then $\lambda \gtrsim R^{n/2}$.  

As a remark, the scale $R^{1/2}$ in Corollary 1.7 is the largest one can have. Indeed, with the assumption of the corollary, the unit cubes in $X$ do not have to almost fill $R^3$-cubes completely for $\beta > 1/2$. One can see this from the Knapp example where we only have one wave packet.
To prove our main result - Theorem 1.5 - we will use a broad-narrow analysis, which has similar spirit as the techniques in the work of Bourgain-Guth [6], Bourgain [3], Bourgain-Demeter [5] and Guth [18].

In the broad case, we can exploit the transversality and apply the multilinear refined Strichartz estimate, which is a result obtained by Guth, Li and the authors in [13] (see [12, 14, 13] for applications of refined Strichartz estimate). In the narrow case, we use the $l^2$ decoupling theorem of Bourgain-Demeter [5] in a lower dimension and perform induction on scales. The way we do induction has its roots in the proof of the linear refined Strichartz estimate, due to Guth, Li and the first author (essentially proved in [12], see [13] for the statement in general setting).

Our method is related to Bourgain’s in [3]. In [3], Bourgain had a similar broad-narrow analysis (we have here the size of the small ball being $K^2$ instead of $K$ as in [3] for a technical issue similar to what one has in [5, 18]). He then applied multilinear restriction to control the broad part in the sharp range $s > 2 n (n+1)$ (except the endpoint). He speculated from this that the above range of $s$ might be sharp (see the end of the introduction in [4]). In [3] the narrow part was handled following the general approach from [6], which gives non-sharp estimates. Historically, one could view the present non-endpoint solution to Carleson’s problem as building on [3], providing a subtler way of handling the narrow part and proving Corollary 1.6.

For the stronger Theorem 1.5 and Corollary 1.7, one needs a different ingredient, namely the multilinear refined Strichartz in [13], to handle the broad part.

In Section 2 we show how Corollary 1.6 and Theorem 1.3 follow from Theorem 1.5 and we also present applications of Theorem 1.5 to other problems - bounding the size of divergence set of Schrödinger solutions (Theorem 2.4), the Falconer distance set problem (Theorem 2.6 and 2.7) and the spherical average Fourier decay rates of fractal measures (Theorem 2.8). We prove Theorem 1.5 in Section 3.

**Notation.** We write $A \lesssim B$ if $A \leq CB$ for some absolute constant $C$, $A \sim B$ if $A \lesssim B$ and $B \lesssim A$; $A \ll B$ if $A$ is much less than $B$; $A \preccurlyeq B$ if $A \leq C_\varepsilon R^\varepsilon B$ for any $\varepsilon > 0, R > 1$. Sometimes we also write $A \lesssim B$ if $A < C \varepsilon B$ for some constant $C_\varepsilon$ depending on $\varepsilon$ (when the dependence on $\varepsilon$ is unimportant).

By an $r$-ball (cube) we mean a ball (cube) of radius (side length) $r$. An $r \times \cdots \times r \times L$-tube (box) means a tube (box) with radius (short sides length) $r$ and length $L$. For a set $S$, $\#S$ denotes its cardinality.

**Acknowledgements.** The authors would like to thank Larry Guth and Xiaochun Li for several discussions. They also thank Larry Guth for making some historical remarks, as well as sharing his lecture notes on decoupling online, from which they got a lot inspiration. The second author would like to thank Jean Bourgain and Zihua Guo who introduced the problem to him. The authors are also indebted to Daniel Eceizabarrena and Luis Vega for a discussion on the history of Schrödinger maximal estimate in dimension $1 + 1$.

The work of X.D. is supported by the National Science Foundation under Grant No. 1638352 and the Shing-Shen Chern Fund. The work of R.Z. is supported by the National Science Foundation under Grant No. 1638352 and the James D. Wolfensohn Fund. The authors are grateful to the IAS for providing them wonderful research environment.
2. Applications of Theorem 1.5

2.1. Sharp $L^2$ estimate of Schrödinger maximal function. In this subsection, we show how Corollary 1.6 and Theorem 1.3 follow from Theorem 1.5 via dyadic pigeonholing argument and locally constant property.

**Proof of (Theorem 1.5 $\implies$ Corollary 1.6).** Given $X = \bigcup B_k$, a union of lattice unit cubes in $B^{n+1}(0, R)$ satisfying the assumptions of Corollary 1.6, we sort the lattice $R^{1/2}$-cubes in $\mathbb{R}^{n+3}$ intersecting $X$ by the number $\lambda$ of unit cubes $B_k$ contained in it. Since $1 \leq \lambda \leq R^{O(1)}$, there are only $O(\log R)$ choices for dyadic number $\lambda$. So we can choose a dyadic number $\lambda$ and a subset $\mathcal{B}_\lambda$ of $\{B_k\}$ such that for each unit cube $B$ in $\mathcal{B}_\lambda$, the lattice $R^{1/2}$-cube containing it contains $\sim \lambda$ many unit cubes from $\mathcal{B}_\lambda$ and

$$
\|e^{it\Delta}f\|_{L^2(X)} \lesssim \|e^{it\Delta}f\|_{L^2(\bigcup_{B \in \mathcal{B}_\lambda} B)}.
$$

By applying Theorem 1.5 to $\|e^{it\Delta}f\|_{L^2(\bigcup_{B \in \mathcal{B}_\lambda} B)}$, we get

$$
\|e^{it\Delta}f\|_{L^2(X)} \lesssim \gamma^{(n+1)(n+2)} R^{(n+1)(n+2)} \|f\|_2,
$$

and (1.11) follows from the fact that $1 \leq \gamma R^{\alpha/2}$.

**Proof of (the case $\alpha = n$ of Corollary 1.6 $\implies$ Theorem 1.3).** We aim to show that

$$
\left\| \sup_{0 < t < R} |e^{it\Delta}f| \right\|_{L^2(B^n(0, R))} \lesssim R^{\frac{n}{n+\alpha}} \|f\|_2
$$

holds for all $R \geq 1$ and all $f$ with Fourier support in $A(1) := \{\xi \in \mathbb{R}^n : |\xi| \sim 1\}$.

By viewing $|e^{it\Delta}f(x)|$ essentially as constant on unit balls, we can find a set $X$ described as follows: $X$ is a union of unit balls in $B^n(0, R) \times [0, R]$ satisfying that each vertical thin tube of dimensions $1 \times \cdots \times 1 \times R$ contains exactly one unit ball in $X$, and

$$
\left\| \sup_{0 < t < R} |e^{it\Delta}f| \right\|_{L^2(B^n(0, R))} \lesssim \|e^{it\Delta}f\|_{L^2(X)}.
$$

The desired estimate follows by applying Corollary 1.6 to $\|e^{it\Delta}f\|_{L^2(X)}$ with $\alpha = n$ and $\gamma \lesssim 1$.

2.2. Other applications. By formalizing the locally constant property, from Corollary 1.6 we derive some weighted $L^2$ estimates - Theorem 2.2 and 2.3, which in turn have applications on several problems described below.

**Definition 2.1.** Let $\alpha \in (0, d]$. We say that $\mu$ is an $\alpha$-dimensional measure in $\mathbb{R}^d$ if it is a probability measure supported in the unit ball $B^d(0, 1)$ and satisfies that

$$
\mu(B(x, r)) \leq C \mu r^\alpha, \quad \forall r > 0, \quad \forall x \in \mathbb{R}^d.
$$

Denote $d\mu(x) := R^\alpha d\mu(x/R)$.

**Theorem 2.2.** Let $n \geq 1$, $\alpha \in (0, n]$ and $\mu$ be an $\alpha$-dimensional measure in $\mathbb{R}^n$. Then

$$
\left\| \sup_{0 < t < R} |e^{it\Delta}f| \right\|_{L^2(B^n(0, R); d\mu(x))} \lesssim R^{\frac{n}{n+\alpha}} \|f\|_2,
$$

We refer the readers to [6, Sections 2-5] for the standard formalism of this locally constant property.
whenever \( R \geq 1 \) and \( f \) has Fourier support in \( B^n(0,1) \).

**Theorem 2.3.** Let \( n \geq 1, \alpha \in (0, n+1] \) and \( \mu \) be an \( \alpha \)-dimensional measure in \( \mathbb{R}^{n+1} \). Then

\[
\|e^{it\Delta}f\|_{L^2(B^{n+1}(0,R);d\mu_R(x,t))} \lesssim R^{\frac{n}{2}\alpha+n} \|f\|_2,
\]

whenever \( R \geq 1 \) and \( f \) has Fourier support in \( B^n(0,1) \).

Theorems 2.2 and 2.3 are new when \( n \geq 2, \alpha \in (2,3) \) or \( n \geq 3, \alpha \in (\frac{n+1}{2}, n+1) \) (c.f. [14, 22]). We defer the proof of these weighted \( L^2 \) estimates to the end of this subsection. Let’s first see their applications. We omit history and various previous results on the following three problems and refer the readers to [13, 14, 22] and the references therein.

**I. Hausdorff dimension of divergence set of Schrödinger solutions**

A natural refinement of Carleson’s problem was initiated by Sjögren and Sjölin [27]: determine the size of divergence set, in particular, consider

\[
\alpha_n(s) := \sup_{f \in H^s(\mathbb{R}^n)} \dim \{ x \in \mathbb{R}^n : \lim_{t \to 0} e^{it\Delta}f(x) \neq f(x) \},
\]

where \( \dim \) stands for the Hausdorff dimension.

The following theorem is a direct result of Theorem 2.2 (c.f. [13, 22]). When \( n = 2 \), it recovers the corresponding result derived from the sharp \( L^3 \) estimate of Schrödinger maximal function in D.-Guth-Li [12]. When \( n \geq 3 \), it improves the previous best known result in D.-Guth-Li-Z. [13].

**Theorem 2.4.** Let \( n \geq 2 \). Then

\[
\alpha_n(s) \leq n + 1 - \frac{2(n+1)s}{n}, \quad \frac{n}{2(n+1)} < s < \frac{n}{4}.
\]

**II. Falconer distance set problem**

Let \( E \subset \mathbb{R}^d \) be a compact subset, its distance set \( \Delta(E) \) is defined by

\[
\Delta(E) := \{|x-y| : x, y \in E\}.
\]

**Conjecture 2.5** (Falconer [15]). Let \( d \geq 2 \) and \( E \subset \mathbb{R}^d \) be a compact set. Then

\[
\dim(E) > \frac{d}{2} \Rightarrow |\Delta(E)| > 0.
\]

Here \( |\cdot| \) denotes the Lebesgue measure and \( \dim(\cdot) \) is the Hausdorff dimension.

Following a scheme due to Mattila (c.f. [14] Proposition 2.3), Theorem 2.3 implies the following result towards Falconer’s conjecture. When \( d = 2,3 \), this recovers the previous best known results of Wolff (\( d=2 \), [31]) and D.-Guth-Ou-Wang-Wilson-Z. (\( d=3 \), [14]), via a different approach. In the case \( d \geq 4 \), this improves the previous best known result in [14]:

**Theorem 2.6.** Let \( d \geq 2 \) and \( E \subset \mathbb{R}^d \) be a compact set with

\[
\dim(E) \geq \frac{d^2}{2d-1} = \frac{d}{2} + \frac{1}{4} + \frac{1}{8d-4}.
\]

Then \( |\Delta(E)| > 0 \).

By applying a very recent work of Liu [21, Theorem 1.4], Theorem 2.3 also implies the following result for the pinned distance set problem, with the same threshold:
Theorem 2.7. Let $d \geq 2$ and $E \subset \mathbb{R}^d$ be a compact set with
\[
\dim(E) > \frac{d^2}{2d-1} = \frac{d}{2} + \frac{1}{4} + \frac{1}{8d-4}.
\]
Then there exists $x \in E$ such that its pinned distance set
\[
\Delta_x(E) := \{|x - y| : y \in E\}
\]
has positive Lebesgue measure.

(III) Spherical average Fourier decay rates of fractal measures

Let $\beta_d(\alpha)$ denote the supremum of the numbers $\beta$ for which
\[
\|\hat{\mu}(R)\|_{L^2(B^d(0, R))} \leq C_{\alpha, \mu} R^{-\beta}
\]
whenever $R > 1$ and $\mu$ is an $\alpha$-dimensional measure in $\mathbb{R}^d$. The problem of identifying the precise value of $\beta_d(\alpha)$ was proposed by Mattila [25].

A lower bound of $\beta_d(\alpha)$ as in Theorem 2.8 follows from Theorem 2.3 (c.f. [14, Remark 2.5]). When $d = 2$, this recovers the sharp result of Wolff [31]. When $d = 3$ and $\alpha \in (\frac{3}{4}, 2]$, this recovers the previous best known result of D.-Guth-Ou-Wang-Wilson-Z. [14]. In the case $d = 3, \alpha \in (2, 3)$ or $d \geq 4, \alpha \in (d/2, d)$, this improves the previous best known result in [14].

Theorem 2.8. Let $d \geq 2$ and $\alpha \in (\frac{d}{2}, d)$. Then
\[
\beta_d(\alpha) \geq \frac{(d-1)\alpha}{d}.
\]

The proofs of Theorems 2.2 and 2.3 are entirely similar and we only do the proof of the former here, which is slightly more involved.

Proof of Theorem 2.2 Denote $e^{it\Delta}f(x)$ by $Ef(x, t)$, and $(x, t)$ by $\hat{x}$. Since supp $\hat{f} \subseteq B^n(0, 1)$, we have supp $\hat{E}f \subseteq B^{n+1}(0, 1)$. Thus there exists a Schwartz bump function $\psi$ on $\mathbb{R}^{n+1}$ (we require $\hat{\psi} \equiv 1$ on $B^{n+1}(0, 100)$) such that $(Ef)^2 = (Ef)^2 * \psi$.

The function max$_{\|\hat{y} - \hat{x}\| \leq 10000}$ $|\psi(\hat{y})|$ is rapidly decaying. We call it $\psi_1(\hat{x})$. Note also that any $(x, t)$ in $\mathbb{R}^{n+1}$ belongs to a unique integral lattice cube whose center we denote by $\tilde{m} = (m, m_{n+1}) = (m_1, \ldots, m_{n+1}) = \tilde{m}(x, t)$.

Then we have
\[
\left\| \sup_{0 < t < R} |e^{it\Delta}f| \right\|_{L^2(B^n(0, R); d\mu_R)}^2 = \int_{B^n(0, R)} \sup_{0 < t < R} |Ef(x, t)|^2 d\mu_R(x)
\leq \int_{B^n(0, R)} \sup_{0 < t < R} (|Ef|^2 * |\psi_1|)(x, t) d\mu_R(x)
\leq \sum_{m = (m_1, \ldots, m_n) \in \mathbb{Z}^n, \ |m_1| \leq R} \left( \int_{|x-m| \leq 10} d\mu_R(x) \right) \cdot \sup_{m_{n+1} \in \mathbb{Z}} (|Ef|^2 * |\psi_1|)(m, m_{n+1}).
\]
We denote that for each \( m \in \mathbb{Z}^n \), the supremum of \((|Ef|^2|\psi_1|)(m, m_{n+1})\) over all integers \( 0 \leq m_{n+1} \leq R \) is attained at \( m_{n+1} = b(m) \). Also we assume \(|f||e = 1\) so \(|e^{i\Delta f}|\) is uniformly bounded pointwisely. For each \( m \in \mathbb{Z}^n \) we define

\[
\nu_m := \int_{|x-m| \leq 10} d\mu_R(x) \lesssim 1.
\]

By (2.8), we have

\[
\left\| \sup_{0 < t < R} |e^{i\Delta f}| \right\|_{L^2(B^n(0, R); d\mu_R)}^2 \lesssim \sum_{\nu \text{ dyadic}} \sum_{m \in \mathbb{Z}^n, |m| \leq R} \nu \cdot (|Ef|^2|\psi_1|)(m, b(m)) + R^{-89n}.
\]

For each dyadic \( \nu \), denote \( A_{\nu} = \{ m \in \mathbb{Z}^n : |m| \leq R, \nu_m \sim \nu \} \). Performing a dyadic pigeonholing over \( \nu \) we see that there exists a dyadic \( \nu \in [R^{-100n}, 1] \) such that for any small \( \varepsilon > 0 \),

\[
\left\| \sup_{0 < t < R} |e^{i\Delta f}| \right\|_{L^2(B^n(0, R); d\mu_R)}^2 \lesssim \sum_{m \in A_{\nu}} \nu \cdot (|Ef|^2|\psi_1|)(m, b(m)) + R^{-89n}
\]

\[
\lesssim \nu \cdot \left( \int_{B^{n+1}((m, b(m)), R^\alpha)} |Ef|^2 \right) + R^{-89n}
\]

\[
\lesssim \nu \cdot \int_{B^{n+1}((m, b(m)), R^\alpha)} |Ef|^2 + R^{-89n}.
\]

Consider the set \( X_{\nu} = \bigcup_{m \in A_{\nu}} B^{n+1}((m, b(m)), R^\alpha) \). It is a union of a bunch of distinct \( R^\varepsilon \)-balls and is in turn a union of unit balls. These balls’ projection onto the \((x_1, \ldots, x_n)\)-plane are essentially disjoint (a point can be covered \( \lesssim \varepsilon \) times). For every \( r > R^{2\varepsilon} \) by the definition of \( A_{\nu} \), the intersection of \( X_{\nu} \) and any \( r \)-ball can be contained in no more than \( R^{100n}r^{-1}r^{-\alpha} \) disjoint \( R^\varepsilon \)-balls. Hence we can apply Corollary 1.6 to \( X_{\nu} \) with \( \gamma \lesssim R^{100n}r^{-1}r^{-\alpha} \) and \( \alpha \). With (2.10) this gives

\[
\left\| \sup_{0 < t < R} |e^{i\Delta f}| \right\|_{L^2(B^n(0, R); d\mu_R)}^2 \lesssim \nu^{n-1} R^\alpha R^{\alpha \varepsilon} ||f||_2^2 \lesssim R^{\alpha \varepsilon} ||f||_2^2.
\]

This concludes the proof. \( \square \)

3. **Main inductive proposition and proof of Theorem 1.5**

To prove Theorem 1.5, we will use a broad-narrow analysis which involves inductions. To make everything work we introduce another parameter \( K \) and state the theorem in a slightly different way. This is our main inductive proposition:

**Proposition 3.1.** Let \( n \geq 1 \). For any \( 0 < \varepsilon < 1/100 \), there exist constants \( C_{\varepsilon} \) and \( 0 < \delta = \delta(\varepsilon) \ll \varepsilon \) (e.g. \( \delta = \varepsilon^{100} \)) such that the following holds for all \( R \geq 1 \) and all \( f \) with \( \text{supp} \hat{f} \subset B^n(0, 1) \). Let \( p = \frac{2(n+1)}{n-1} \) \((p = \infty \text{ when } n = 1)\). Suppose that
\[ Y = \bigcup_{k=1}^{M} B_k \] is a union of lattice \( K^2 \)-cubes in \( B^{n+1}(0, R) \) and each lattice \( R^{1/2} \)-cube intersecting \( Y \) contains \( \sim \lambda \) many \( K^2 \)-cubes in \( Y \), where \( K = R^8 \). Suppose that
\[ \|e^{it\Delta}f\|_{L^p(B_k)} \] is dyadically a constant in \( k = 1, 2, \cdots, M \).

Let \( 1 \leq \alpha \leq n + 1 \) and \( \gamma \) be given by
\[ \gamma := \max_{B^{n+1}(x', r) \subseteq B^{n+1}(0, R)} \frac{\# \{ B_k : B_k \subset B(x', r) \}}{r^\alpha}. \]

Then
\[ \|e^{it\Delta}f\|_{L^p(Y)} \leq C_M^{-1} \gamma^{(n+1)/2} \lambda^{(n+1)/(n+2)} R^{(n+1)/(n+2)} \|f\|_2. \]

Theorem 1.5 follows from Proposition 3.1 by a dyadic pigeonholing argument:

**Proof of (Proposition 3.1) \Rightarrow Theorem 1.5.** Given \( X = \bigcup_k B_k \), a union of lattice unit cubes satisfying the assumptions of Theorem 1.5 we sort these unit cubes \( B_k \) according to the value of \( \|e^{it\Delta}f\|_{L^p(B_k)} \). Assuming \( \|f\|_2 = 1 \), there are only \( O(\log R) \) significant dyadic choices for this value. Therefore we can choose \( X' \subset X \), a union of unit cubes \( B \), such that
\[ \|e^{it\Delta}f\|_{L^p(B)} \] is dyadically a constant in \( B \) from \( X' \) and
\[ \|e^{it\Delta}f\|_{L^2(X)} \lesssim \|e^{it\Delta}f\|_{L^2(X')}. \]

Let \( M \) be the total number of unit cubes \( B \) in \( X' \). In view of \( |e^{it\Delta}f| \) being essentially constant on unit balls, the estimate (1.9) is equivalent to
\[ \|e^{it\Delta}f\|_{L^p(X')} \lesssim M^{1 - \frac{1}{n+1}} \gamma^{(n+1)/2} \lambda^{(n+1)/(n+2)} R^{(n+1)/(n+2)} \|f\|_2, \]
where \( p = \frac{2(n+1)}{n-1} \), and \( \gamma, \lambda \) are as in the assumptions of Theorem 1.5.

We further sort the unit cubes \( B \) in \( X' \) as follows:

1. Let \( \beta \) be a dyadic number, and \( B_\beta \) a sub-collection of the unit cubes in \( X' \) such that for each \( B \in B_\beta \), the lattice \( K^2 \)-cube \( \tilde{B} \) containing \( B \) satisfies
   \[ \|e^{it\Delta}f\|_{L^p(\tilde{B})} \sim \beta. \]

   Denote the collection of relevant \( K^2 \)-cubes by \( \tilde{B}_\beta \).

2. Fix \( \beta \). Let \( \lambda' \) be a dyadic number and \( B_{\beta, \lambda'} \) a sub-collection of \( B_\beta \) such that for each \( B \in B_{\beta, \lambda'} \), the lattice \( R^{1/2} \)-cube \( Q \) containing \( B \) contains \( \sim \lambda' \) many \( K^2 \)-cubes from \( \tilde{B}_\beta \). Denote the collection of relevant \( K^2 \)-cubes by \( \tilde{B}_{\beta, \lambda'} \).

Since there are only \( O(\log R) \) many significant choices for each dyadic number \( \beta, \lambda' \), we can choose some \( \beta \) and \( \lambda' \) so that \( \#B_{\beta, \lambda'} \gtrsim M \). Then it follows easily by definition that
\[ M' := \#\tilde{B}_{\beta, \lambda'} \gtrsim M, \quad \lambda' \leq \lambda, \]
and
\[ \gamma' := \max_{B^{n+1}(x', r) \subseteq B^{n+1}(0, R)} \frac{\# \{ \tilde{B} \in \tilde{B}_{\beta, \lambda'} : \tilde{B} \subset B(x', r) \}}{r^\alpha} \leq \gamma. \]
Applying Proposition 3.1 to \( \|e^{it\Delta}f\|_{L^p(Y)} \) with \( Y = \bigcup_{\hat{B} \in \mathcal{B}_{p',\lambda'}} \hat{B} \) and parameters \( M', \gamma', \lambda' \), we get
\[
\|e^{it\Delta}f\|_{L^p(X)} \lesssim \|e^{it\Delta}f\|_{L^p(Y)} \lesssim M^{-\frac{1}{2\tau}\gamma\frac{n(n+2)}{(n+1)(n+2)} \lambda\frac{n(n+2)}{(n+1)(n+2)} R_{\tau}} \|f\|_2,
\]
as desired.

The rest of this section is devoted to a proof of Proposition 3.1. Note that when the radius \( R \lesssim 1 \), the estimate (3.2) is trivial. So we can assume that \( R \) is sufficiently large compared to any constant depending on \( \varepsilon \). We will induct on radius \( R \) in our proof.

In the proof, we will sometimes have paragraphs starting with Intuition. We hope that those could help the readers understand what we do next.

**Intuition.** For our union \( Y \) of \( K^2 \)-cubes, we want to use decoupling theory on each \( K^2 \)-cube. This will relate the whole \( e^{it\Delta}f \) to extended functions \( \hat{e}^{it\Delta}f \), from various \( 1/K \)-caps \( \tau \) in the frequency space. Instead of doing decoupling in dimension \( n+1 \), we are going to do a broad-narrow analysis following Bourgain-Guth [6], Bourgain [3], Bourgain-Demeter [5] and Guth [15]: for each \( K^2 \)-cube, one of the two has to happen:

(i) It is broad in the sense that some contributing caps are transversal and form an \((n+1)\)-linear structure. In this case the function is controlled by multilinear estimates which are usually strong enough.

(ii) It is narrow (i.e. not broad). In this case all the contributing caps have normal directions close to a hyperplane, which enables us to use decoupling in dimension \( n \).

Either way we get better estimates than a direct \((n+1)\)-dimensional decoupling. We control the broad part directly, and do an induction on the narrow part. Our induction has its roots in the proof of the refined Strichartz estimate in [12] [13].

Throughout this section we fix \( p = \frac{2(n+1)}{n-1} \). In the frequency space we decompose \( B^n(0,1) \) into disjoint \( K^{-1} \)-cubes \( \tau \). Denote the set of \( K^{-1} \)-cubes \( \tau \) by \( S \). For function \( f \) with \( \text{supp} \hat{f} \subset B^n(0,1) \) we have \( f = \sum \tau f_\tau \), where \( f_\tau \) is \( f \) restricted to \( \tau \). Given a \( K^2 \)-cube \( B \), we define its **significant** set
\[
S(B) := \left\{ \tau \in S : \|e^{it\Delta}f_\tau\|_{L^p(B)} \geq \frac{1}{100(#S)} \|e^{it\Delta}f\|_{L^p(B)} \right\}.
\]
Note that
\[
\| \sum_{\tau \in S(B)} e^{it\Delta}f_\tau \|_{L^p(B)} \sim \|e^{it\Delta}f\|_{L^p(B)}.
\]
We say a \( K^2 \)-cube \( B \) is **narrow** if there is an \( n \)-dimensional subspace \( V \) such that for all \( \tau \in S(B) \)
\[
\text{Angle}(G(\tau), V) \leq \frac{1}{100nK},
\]
where \( G(\tau) \subset S^n \) is a spherical cap of radius \( \sim K^{-1} \) given by
\[
G(\tau) := \left\{ \frac{(-2\xi, 1)}{|(-2\xi, 1)|} : \xi \in \tau \right\},
\]
and \( \text{Angle}(G(\tau), V) \) denotes the smallest angle between any non-zero vector \( v \in V \) and \( v' \in G(\tau) \). Otherwise we say the \( K^2 \)-cube \( B \) is **broad**. It follows from this
definition that for any broad $B$, there exist $\tau_1, \cdots, \tau_{n+1} \in S(B)$ such that for any $v_j \in G(\tau_j)$

$$|v_1 \wedge v_2 \wedge \cdots \wedge v_{n+1}| \gtrsim K^{-n}. \tag{3.4}$$

Denote the union of broad $K^2$-cubes $B_k$ in $Y$ by $Y_{\text{broad}}$, and the union of narrow $K^2$-cubes $B_k$ in $Y$ by $Y_{\text{narrow}}$. We call it the broad case if $Y_{\text{broad}}$ contains $\geq M/2$ many $K^2$-cubes, the narrow case otherwise. We will deal with the broad case in Subsection 3.1 using multilinear refined Strichartz estimate from [13]. And we handle the narrow case in Subsection 3.2 by an inductive argument via the Bourgain-Demeter $l^2$ decoupling theorem [5] and induction on scales.

### 3.1. Broad case

Recall that $K = R^d$. A key tool we are using in the broad case is the following multilinear refined Strichartz estimate from [13], which is proved using $l^2$ decoupling, induction on scales and multilinear Kakeya estimates (see [11, 16]).

**Theorem 3.2** (c.f. Theorem 4.2 in [13]). Let $q = \frac{2(n+2)}{n}$. Let $f$ be a function with Fourier support in $B^n(0,1)$. Suppose that $\tau_1, \cdots, \tau_{n+1} \in S$ and (3.4) holds for any $v_j \in G(\tau_j)$. Suppose that $Q_1, Q_2, \cdots, Q_N$ are lattice $R^{1/2}$-cubes in $B_R^{n+1}$, so that

$$\|e^{it\Delta}f_{\tau_i}\|_{L^q(Q_j)} \text{ is dyadically a constant in $j$, for each $i = 1, 2, \cdots, n+1$.}$$

Let $Y$ denote $\bigcup_{j=1}^N Q_j$. Then for any $\epsilon > 0$,

$$\left\| \prod_{j=1}^{n+1} |e^{it\Delta}f_{\tau_j}|^{\frac{1}{n+1}} \right\|_{L^q(Y)} \leq C\epsilon R^n N^{\frac{-n}{(n+1)(n+\epsilon)}} \|f\|_2. \tag{3.5}$$

In the broad case, there are $\sim M$ many broad $K^2$-cubes $B$. Denote the collection of $(n+1)$-tuple of transverse caps by $\Gamma$:

$$\Gamma := \{ \hat{\tau} = (\tau_1, \cdots, \tau_{n+1}) : \tau_j \in S \text{ and } (3.4) \text{ holds for any } v_j \in G(\tau_j) \}. \tag{3.4}$$

Then for each broad $B$,

$$\|e^{it\Delta}f\|_{L^p(B)}^p \leq K^{O(1)} \prod_{j=1}^{n+1} \left( \int_B |e^{it\Delta}f_{\tau_j}|^p \right)^{\frac{1}{n+1}}, \tag{3.6}$$

for some $\hat{\tau} = (\tau_1, \cdots, \tau_{n+1}) \in \Gamma$. In order to exploit the transversality, we want to bound the above geometric average of integrals by an integral of geometric average up to a loss of $K^{O(1)}$. We can do this by random translation and locally constant property. Given a $K^2$-cube $B$, denote its center by $x_B$. We break $B$ into finitely overlapping balls of the form $B(x_B + v, 2)$, where $v \in B(0,K^2) \cap \mathbb{Z}^{n+1}$. For each $\tau_j$, we can view $|e^{it\Delta}f_{\tau_j}|$ essentially as constant on each $B(x_B + v, 2)$. Choose $v_j \in B(0,K^2) \cap \mathbb{Z}^{n+1}$ such that $\|e^{it\Delta}f_{\tau_j}\|_{L^\infty(B)}$ is obtained in $B(x_B + v_j, 2)$. Denote $v_j = (x_j, t_j)$ and define $f_{\tau_j, v_j}$ by

$$\hat{f}_{\tau_j, v_j}(\xi) := \hat{f}_{\tau_j}(\xi) e^{i(x_j \cdot \xi + t_j |\xi|^2)}.$$

Then $e^{it\Delta}f_{\tau_j, v_j}(x) = e^{i(x_j t_j)} e^{i(x_j \cdot \xi + t_j |\xi|^2)}$ and $|e^{it\Delta}f_{\tau_j, v_j}(x)|$ attains $\|e^{it\Delta}f_{\tau_j}\|_{L^\infty(B)}$ in $B(x_B, 2)$. Therefore

$$\int_B |e^{it\Delta}f_{\tau_j}|^p \leq K^{O(1)} \int_{B(x_B, 2)} |e^{it\Delta}f_{\tau_j, v_j}|^p. \tag{3.7}$$
Now for each broad $B$, we find some $\tilde{\tau} = (\tau_1, \cdots, \tau_{n+1}) \in \Gamma$ and $\tilde{\nu} = (\nu_1, \cdots, \nu_{n+1})$ such that

$$
\left\| e^{it\Delta} f \right\|_{L^p(B)} \leq K^{O(1)} \prod_{j=1}^{n+1} \left( \int_{B(x_B, 2)} |e^{it\Delta} f_{\tau_j, \nu_j}|^p \right)^{\frac{1}{p+\tilde{\nu}}} 
$$

(3.8)

Since there are only $K^{O(1)}$ choices for $\tilde{\tau}$ and $\tilde{\nu}$. We can choose some $\tilde{\tau}$ and $\tilde{\nu}$ such that (3.8) holds for $\geq K^{-C} M$ broad balls $B$. From now on, fix $\tilde{\tau}$ and $\tilde{\nu}$, and let $f_j$ denote $f_{\tau_j, \nu_j}$. Next we further sort the collection $B$ of remaining broad balls as follows:

1. For a dyadic number $A$, let $B_A$ be a sub-collection of $B$ in which for each $B$ we have

$$
\left\| \prod_{j=1}^{n+1} |e^{it\Delta} f_j|^\frac{1}{p+\tilde{\nu}} \right\|_{L^\infty(B(x_B, 2))} \sim A .
$$

2. Fix $A$, for dyadic numbers $\tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}$, let $B_{A, \tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}}$ be a sub-collection of $B_A$ in which for each $B$, the $R^{1/2}$-cube $Q$ containing $B$ contains $\sim \tilde{\lambda}$ cubes from $B_A$ and

$$
\left\| e^{it\Delta} f_j \right\|_{L^q(Q)} \sim \tilde{t}_j, \quad j = 1, 2, \cdots, n+1 .
$$

Here $q = \frac{2(n+2)}{n}$.

We can assume that $\|f\|_2 = 1$. Then all the above dyadic numbers making significant contributions can be assumed to be between $R^{-C}$ and $R^C$ for a large constant $C$. Therefore, there exist some dyadic numbers $A, \tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}$ such that $B_{A, \tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}}$ contains $\geq K^{-C} M$ many cubes $B$. Fix a choice of $A, \tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}$ and denote $B_{A, \tilde{\lambda}, \tilde{t}_1, \cdots, \tilde{t}_{n+1}}$ by $B$ for convenience (a mild abuse of notation). Then, in the broad case, it follows from (3.8) and our choice of $A$ that

$$
\left\| e^{it\Delta} f \right\|_{L^p(Y)} \leq K^{O(1)} \left\| \prod_{j=1}^{n+1} |e^{it\Delta} f_j|^\frac{1}{p+\tilde{\nu}} \right\|_{L^p(\cup_{B \in B} B(x_B, 2))} 
$$

(3.9)

\begin{align*}
&\leq K^{O(1)} M^{\tilde{\lambda} - \frac{1}{q}} \left\| \prod_{j=1}^{n+1} |e^{it\Delta} f_j|^\frac{1}{p+\tilde{\nu}} \right\|_{L^q(\cup_{B \in B} B(x_B, 2))} \\
&\leq K^{O(1)} M^{\tilde{\lambda} - \frac{1}{(n+1)(n+2)}} \left\| \prod_{j=1}^{n+1} |e^{it\Delta} f_j|^\frac{1}{p+\tilde{\nu}} \right\|_{L^q(\cup_{Q \in Q} Q)} ,
\end{align*}

where $Q$ is the collection of relevant $R^{1/2}$-cubes $Q$ when we define $B$. Note that

$$
(#Q)\lambda \geq (#Q)\tilde{\lambda} \sim #B \geq K^{-C} M ,
$$

so

$$
\tilde{N} := #Q \geq K^{-C} \frac{M}{\lambda} .
$$

(3.10)
Applying Theorem 3.2 we get
\[
\left\| \prod_{j=1}^{n+1} e^{it\Delta} f_j \right\|_{L^q(\bigcup_{Q \in Q} Q)} \leq K^{O(1)} \left( \frac{M}{\alpha} \right)^{-\frac{2}{(n+1)(n+2)}} \|f\|_2,
\]
and therefore by (3.9),
\[
\|e^{it\Delta} f\|_{L^p(Y)} \leq K^{O(1)} M^{-\frac{1}{n+1}} \lambda^{(n+1)(n+2)} \|f\|_2.
\]
Note that
\[
M^{-\frac{1}{n+1}} \lambda^{(n+1)(n+2)} \leq K^{O(1)} M^{-\frac{1}{n+1}} \lambda^{(n+1)(n+2)} R^{(n+1)(n+2)}
\]
holds if and only if \( M \leq K^{O(1)} \gamma^2 R^a \). Indeed, by definition (3.1) of \( \gamma \), we have \( M \leq \gamma R^a \) and \( \gamma \geq K^{-2a} \). So the broad case is done.

3.2. Narrow case. For each narrow ball, we have the following lemma which is a consequence of \( L^2 \) decoupling theorem in dimension \( n \) and Minkowski’s inequality. This argument is essentially contained in Bourgain-Demeter’s proof of the \( L^2 \) decoupling conjecture and we omit the details (see the proof of Proposition 5.5 in [3]).

Lemma 3.3. Suppose that \( B \) is a narrow \( K^2 \)-cube in \( \mathbb{R}^{n+1} \). Then for any \( \varepsilon > 0 \),
\[
\|e^{it\Delta} f\|_{L^p(B)} \leq C_\varepsilon K^{\varepsilon} \left( \sum_{\tau \in S} \|e^{it\Delta} f_\tau\|_{L^p(\omega_B)}^2 \right)^{1/2},
\]
here \( p = \frac{2(n+1)}{n-1} \), \( S \) denotes the set of \( K^{-1} \)-cubes which tile \( B^n(0,1) \), and \( \omega_B \) is a weight function which is essentially a characteristic function on \( B \).

For each \( \tau \in S \), we will deal with \( e^{it\Delta} f_\tau \) by parabolic rescaling and induction on radius. In order to do so, we need to further decompose \( f \) in physical space and perform dyadic pigeonholing several times to get the right picture for our inductive hypothesis at scale \( R_1 := R/K^2 \) after rescaling.

Intuition. For each \( 1/K \)-cap \( \tau \), all wave packets associated with \( f_\tau \) through a given point have to lie in a common box that has one side length \( R \) and other side lengths \( R/K \). Every single box of this type will become an \( R/K^2 \)-ball if we perform a parabolic rescaling to transform \( \tau \) into the standard \( 1 \)-cap. We want to use the inductive hypothesis for radius \( R/K^2 \) in an efficient way. A bunch of dyadic pigeonholing steps will be needed.

First, we break the physical ball \( B^n(0,R) \) into \( R/K \)-cubes \( D \). For each pair \((\tau,D)\), let \( f_{\tau,D} \) be the function formed by cutting off \( f \) on the cube \( D \) (with a Schwartz tail) in physical space and the cube \( \tau \) in Fourier space. Note that \( e^{it\Delta} f_{\tau,D} \), restricted to \( B^{n+1}(R) \), is essentially supported on an \( R/K \times \cdots \times R/K \times R \)-box where we denote by \( \square_{\tau,D} \). The box \( \square_{\tau,D} \) is in the direction given by \((-2c(\tau),1)\) and intersects \( t = 0 \) at the cube \( D \), where \( c(\tau) \) is the center of \( \tau \). For a fixed \( \tau \), the different boxes \( \square_{\tau,D} \) tile \( B^{n+1}(0,R) \). In particular, for each \( \tau \), a given

\[\text{In reality, our boxes will have edge length slightly larger, say being larger by } K^{100} \text{ times. See e.g. the wave packet decomposition theorem in [17]. This would not hurt us in any way and we omit this technicality for reading convenience.}\]
$K^2$-cube $B$ lies in exactly one box $\square_{r,D}$. We write $f = \sum_{\Box} f_{\Box}$ for abbreviation. By Lemma 3.3 for each narrow $K^2$-cube $B$,

$$\|e^{it\Delta} f\|_{L^p(B)} \lesssim K^2 \left( \sum_{\Box} \|e^{it\Delta} f\|_{L^p(\omega_B)}^2 \right)^{1/2}.$$  \hspace{1cm} (3.11)

Figure 1. Tubes of different scales in the $\square$

Next, we perform a dyadic pigeonholing to get our inductive hypothesis for each $f_{\Box}$. Denote

$$R_1 := R/K^2 = R^{1-2\delta}, \quad K_1 := R_1^{\delta} = R^{\delta-2\delta^2}.$$ 

Tile $\square$ by $KK_1^2 \times \cdots \times KK_1^2 \times K^2 K_1^2$-tubes $S$, and also tile $\square$ by $R^{1/2} \times \cdots \times R^{1/2} \times K K_1^{1/2}$-tubes $S'$ (all running parallel to the long axis of $\square$). To understand these scales, see Figure 2 for the change in physical space (3.20) during the process of parabolic rescaling. In particular, after rescaling the $\square$ becomes an $R_1$-cube, the tubes $S'$ and $S$ become lattice $R_1^{1/2}$-cubes and $K_1^2$-cubes respectively.

We apply the following to regroup tubes $S$ and $S'$ inside each $\square$:

1. Sort those tubes $S$ which intersect $Y$ according to the value $\|e^{it\Delta} f_{\square}\|_{L^p(S)}$ and the number of narrow $K^2$-cubes contained in it. For dyadic numbers $\eta, \beta_1$, we use $S_{\square, \eta, \beta_1}$ to stand for the collection of tubes $S \subset \square$ which each satisfy that $S$ contains $\sim \eta$ narrow $K^2$-cubes in $Y_{\text{narrow}}$ and $\|e^{it\Delta} f_{\square}\|_{L^p(S)} \sim \beta_1$.

2. For fixed $\eta, \beta_1$, we sort the tubes $S' \subset \square$ according to the number of tubes $S \in S_{\square, \eta, \beta_1}$ contained in it. For dyadic number $\lambda_1$, let $S_{\square, \eta, \beta_1, \lambda_1}$ be the sub-collection of $S_{\square, \eta, \beta_1}$ such that for each $S \in S_{\square, \eta, \beta_1, \lambda_1}$, the tube $S'$ containing $S$ contains $\sim \lambda_1$ tubes from $S_{\square, \eta, \beta_1}$.

3. For fixed $\eta, \beta_1, \lambda_1$, we sort the boxes $\square$ according to the value $\|f_{\square}\|_2$, the number $\#S_{\square, \eta, \beta_1, \lambda_1}$ and the value $\gamma_1$ defined below. For dyadic numbers
\(\beta_2, M_1, \gamma_1, \eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1\) denote the collection of boxes \(\square\) which each satisfy that
\[
\|f\|_2 \sim \beta_2, \quad \#S_{\mathcal{D}, \eta, \beta_1, \lambda_1} \sim M_1
\]
and
\[
(3.12) \quad \max_{T, r \subset \mathcal{T}, r \geq K^{-1}} \#\{S \in S_{\mathcal{D}, \eta, \beta_1, \lambda_1} : S \subset T\} \sim \gamma_1,
\]
where \(T_r\) are \(Kr \times \cdots \times Kr \times K^2r\)-tubes in \(\square\) running parallel to the long axis of \(\square\).

Let \(Y_{\mathcal{D}, \eta, \beta_1, \lambda_1}\) be the union of the tubes \(S\) in \(S_{\mathcal{D}, \eta, \beta_1, \lambda_1}\), and \(X_{Y_{\mathcal{D}, \eta, \beta_1, \lambda_1}}\) the corresponding characteristic function. Then on \(Y_{\text{narrow}}\) we can write
\[
e^{it\Delta} f = \sum_{\eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1} \left( \sum_{\square \subset B_{\mathcal{D}, \eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1}} e^{it\Delta} f_{\square} \cdot X_{Y_{\mathcal{D}, \eta, \beta_1, \lambda_1}} \right) + \text{RapDec}(R).
\]

Here \(\text{RapDec}(R)\) means some term that is smaller than a huge negative power of \(R\). As before it will not hurt us in any way. We will neglect this term in the sequel. Again to make the statement really rigorous one needs to increase the side lengths of \(\square\) by a tiny power of \(R\), say \(R^{\log R} \sim K^d\). As before we do not present this technicality for reading convenience.

In particular, on each narrow \(B\) we have
\[
e^{it\Delta} f = \sum_{\eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1} \left( \sum_{B \subset Y_{\mathcal{D}, \eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1}} e^{it\Delta} f_B \right).
\]

Without loss of generality, we assume that \(\|f\|_2 = 1\). Then we can further assume that the dyadic numbers above are in reasonable ranges, say
\[
1 \leq \eta \leq K^{O(1)}, \quad R^{-C} \leq \beta_1 \leq K^{O(1)}, \quad 1 \leq \lambda_1 \leq R^{O(1)}
\]
and
\[
R^{-C} \leq \beta_2 \leq 1, \quad 1 \leq M_1 \leq R^{O(1)}, \quad K^{-2n} \leq \gamma_1 \leq R^{O(1)},
\]
where \(C\) is a large constant such that the contributions from those \(\beta_1\) and \(\beta_2\) less than \(R^{-C}\) are negligible. Therefore, there are only \(O(\log R)\) significant choices for each dyadic number. Because of (3.11) and (3.13), by pigeonholing, we can choose \(\eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1\) so that
\[
(3.14) \quad \|e^{it\Delta} f\|_{L^p(B)} \lesssim (\log R)^6 K^{-1} \left( \sum_{\square \subset B_{Y_{\mathcal{D}, \eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1}} \|e^{it\Delta} f_{\square}\|_{L^2(\omega_B)}^2 \right)^{1/2}
\]
holds for a fraction \(\gtrsim (\log R)^{-6}\) of all narrow \(K^2\)-cubes \(B\).

We fix \(\eta, \beta_1, \lambda_1, \beta_2, M_1, \gamma_1\) for the rest of the proof. Let \(Y_{\mathcal{D}}\) and \(B\) stand for the abbreviations of \(Y_{\mathcal{D}, \eta, \beta_1, \lambda_1}\) and \(B_{\mathcal{D}, \eta, \beta_1, \beta_2, M_1, \gamma_1}\) respectively. Finally we sort the
narrow balls \( B \) satisfying \( \text{[3.14]} \) by \( \#\{ \square \in \mathbb{B} : B \subset Y_\square \} \). Let \( Y' \subset Y_{\text{narrow}} \) be a union of narrow \( K^2 \)-cubes \( B \) which each obey

\[
\| e^{it\Delta} f \|_{L^p(B)} \lesssim (\log R)^6 K^{\frac{e^4}{7}} \left( \sum_{\square \in \mathbb{B} : B \subset Y_\square} \| e^{it\Delta} f \|_{L^p(\omega_B)}^2 \right)^{1/2}
\]

and

\[
\#\{ \square \in \mathbb{B} : B \subset Y_\square \} \sim \mu
\]

for some dyadic number \( 1 \leq \mu \leq K^{O(1)} \), moreover the number of \( K^2 \)-cubes \( B \) in \( Y' \) is \( \gtrsim (\log R)^{-7} M \).

Now we are done with dyadic pigeonholing argument and let us put all these together. By our assumption that \( \| e^{i\tau \Delta} f \|_{L^p(B_\tau)} \) is essentially constant in \( k = 1, 2, \cdots, M \), in the narrow case we have

\[
\| e^{it\Delta} f \|_{L^p(Y')}^p \lesssim (\log R)^7 \sum_{B \subset Y'} \| e^{it\Delta} f \|_{L^p(B)}^p.
\]

For each \( B \subset Y' \), it follows from \( \text{[3.15]}, \text{[3.16]} \) and Hölder’s inequality that

\[
\| e^{it\Delta} f \|_{L^p(B)}^p \lesssim (\log R)^{6p} K^{e^4 p \mu^{\frac{1}{2}} - 1} \sum_{\square \in \mathbb{B} : B \subset Y_\square} \| e^{it\Delta} f \|_{L^p(\omega_B)}^p.
\]

Putting \( \text{[3.17]} \) and \( \text{[3.18]} \) together and as before omitting the rapidly decaying tails,

\[
\| e^{it\Delta} f \|_{L^p(Y')} \lesssim (\log R)^{13} K^{e^4 \mu^{\frac{1}{2}} + \frac{7}{15}} \left( \sum_{\square \in \mathbb{B}} \| e^{it\Delta} f \|_{L^p(Y_\square)}^p \right)^{1/p}.
\]

Next, we apply parabolic rescaling and induction on radius to each \( \| e^{i\tau \Delta} f \|_{L^p(Y_\square)} \).

For each \( 1/K \)-cube \( \tau = \tau_\square \) in \( B^n(0, 1) \), we write \( \xi = \xi_0 + K^{-1} \zeta \in \tau \), where \( \xi_0 \) is the center of \( \tau \), then

\[
| e^{it\Delta} f_\square(x) | = K^{-n/2} | e^{i\Delta} g(\tilde{x}) |
\]

for some function \( g \) with Fourier support in the unit cube and \( \| g \|_2 = \| f_\square \|_2 \), where the new coordinates \( (\tilde{x}, \tilde{t}) \) are related to the old coordinates \( (x, t) \) by

\[
\begin{align*}
\tilde{x} &= K^{-1} x + 2t K^{-1} \xi_0, \\
\tilde{t} &= K^{-2} t.
\end{align*}
\]

For simplicity, denote the above relation by \( (\tilde{x}, \tilde{t}) = F(x, t) \). Therefore

\[
\| e^{it\Delta} f_\square(x) \|_{L^p(Y_\square)} = K^{\frac{n+2}{p} - \frac{2}{7}} \| e^{i\Delta} g(\tilde{x}) \|_{L^p(\tilde{Y})} = K^{-\frac{1}{p\tau + 1}} \| e^{i\Delta} g(\tilde{x}) \|_{L^p(\tilde{Y})},
\]

where \( \tilde{Y} \) is the image of \( Y_\square \) under the new coordinates.

Note that we can apply our inductive hypothesis \( \text{[3.2]} \) at scale \( R_1 = R/K^2 \) to \( \| e^{i\Delta} g(\tilde{x}) \|_{L^p(\tilde{Y})} \) with new parameters \( M_1, \gamma_1, \lambda_1, R_1 \). More precisely, \( \tilde{Y} = F(Y_\square) \) consists of \( \sim M_1 \) distinct \( K^2_\tau \)-cubes \( F(S) \) in an \( R_1 \)-ball \( F(\square) \), and the \( K^2_\tau \)-cubes \( F(S) \) are organized into \( R_1^{1/2} \)-cubes \( F(S') \) such that each cube \( F(S') \) contains \( \sim \lambda_1 \) cubes \( F(S) \). Moreover, \( \| e^{i\Delta} g(\tilde{x}) \|_{L^p(F(S))} \) is dyadically a constant in \( S \subset Y_\square \). By our choice of \( \gamma_1 \), we have

\[
\max_{B^{n+1}(x', r) \subset F(\square)} \# \{ F(S) : F(S) \subset B(x', r) \} \sim r^{\alpha} \sim \gamma_1.
\]
Henceforth, by (3.21) and inductive hypothesis (3.2) at scale $R_1$ we have
\begin{equation}
\| e^{it\Delta} f_\square(x) \|_{L^p(Y_\square)} \leq K^{-\frac{1}{n+2}} M_1^{-\frac{1}{n+2}} \gamma_1^{-\frac{1}{n+1}} \nu_1^{-\frac{n+1}{(n+2)(n+2)}} \left( \frac{R}{K^2} \right)^{\frac{n+1}{(n+2)(n+2)} + \varepsilon} \| f_\square \|_2.
\end{equation}

From (3.19) and (3.22) we obtain
\begin{equation}
\| e^{it\Delta} f \|_{L^p(Y)} \leq K^{2\alpha} \left( \frac{\mu}{\#B} \right)^{\frac{1}{n+2}} K^{-\frac{1}{n+2}} M_1^{-\frac{1}{n+2}} \gamma_1^{-\frac{1}{n+1}} \nu_1^{-\frac{n+1}{(n+2)(n+2)}} \left( \frac{R}{K^2} \right)^{\frac{n+1}{(n+2)(n+2)} + \varepsilon} \left( \sum_{\square \in B} \| f_\square \|_2^p \right)^{1/p} \| f \|_2,
\end{equation}

where the last inequality follows from orthogonality $\sum_{\square \in B} |f_\square|^2 \lesssim \|f\|_2^2$ and the assumption that $\|f_\square\|_2 \sim \text{constant in } \square \in B$.

\textbf{Intuition.} To finish our inductive argument, we have to relate the old and new parameters. Our setup allows ourselves to do this in a nice way: Given $M_1, \lambda_1$ and $Y_\square$, if $\eta$ is small, i.e. each $S$ contains very few narrow $K^2$-cubes, then $M$ is relatively small; if $\eta$ is large, i.e. each $S$ contains a lot of narrow $K^2$-cubes, then $\lambda$ and $\gamma$ are relatively large. Both make the RHS of what we want to prove reasonably large. This is the reason why one could believe the numerology will work out.

Consider the cardinality of the set $\{(\square, B) : \square \in B, B \subset Y_\square \cap Y'\}$. By our choice of $\mu$ as in (3.16), there is a lower bound

$$\# \{(\square, B) : \square \in B, B \subset Y_\square \cap Y'\} \gtrsim (\log R)^{-7} M \mu.$$ 

On the other hand, by our choices of $M_1$ and $\eta$, for each $\square \in B$, $Y_\square$ contains $\sim M_1$ tubes $S$ and each $S$ contains $\sim \eta$ narrow cubes in $\gamma$-narrow, so

$$\# \{(\square, B) : \square \in B, B \subset Y_\square \cap Y'\} \lesssim (\# B) M_1 \eta.$$ 

Therefore, we get
\begin{equation}
\frac{\mu}{\#B} \lesssim (\log R)^7 M_1 \eta.
\end{equation}

Next by our choices of $\gamma_1$ as in (3.12) and $\eta$,$$
\gamma_1 \cdot \eta \sim \max_{T_r \subset \square \cdot r \geq K^2} \frac{\# \{ S : S \subset Y_\square \cap T_r \} \cdot \# \{ B : B \subset S \cap \gamma \text{ narrow for any fixed } S \subset Y_\square \}}{r^\alpha} \leq \frac{K \gamma(Kr)^\alpha}{r^\alpha} = K \gamma^\alpha.
\end{equation}

where the last inequality follows from the definition (3.1) of $\gamma$ and the fact that we can cover a $K r \times \cdots \times K r \times K^2 r$-tube $T_r$ by $\sim K$ finitely overlapping $Kr$-balls. Hence,
\begin{equation}
\eta \lesssim \frac{\gamma K^{\alpha+1}}{\gamma_1}.
\end{equation}
Finally we relate $\lambda_1$ and $\lambda$ by considering the number of narrow $K$-balls in each relevant $R^{1/2} \times \cdots \times R^{1/2} \times KR^{1/2}$-tube $S'$. Recall that each relevant $S'$ contains $\sim \lambda_1$ tubes $S$ in $Y_\Omega$ and each such $S$ contains $\sim n$ narrow balls. On the other hand, we can cover $S'$ by $\sim K$ finitely overlapping $R^{1/2}$-balls and by assumption each $R^{1/2}$-ball contains $\lesssim \lambda$ many $K$-cubes in $Y$. Thus it follows that

$$\lambda_1 \lesssim \frac{K \lambda}{\eta}.$$ (3.26)

By inserting (3.24) and (3.26) into (3.23),

$$\|e^{it\Delta} f\|_{L^p(Y)} \lesssim \frac{K^{3\epsilon^4}}{K^{2\epsilon}} \left( \frac{\eta}{K^{\alpha+1}} \right) \left( \frac{\eta}{K^{\alpha+1}} \right)^{\frac{2}{(n+1)(n+2)}} M^{-\frac{1}{n+1}} \lambda^{\frac{n}{(n+1)(n+2)}} \sigma R^{\frac{\alpha}{(n+1)(n+2)}} \|f\|_2 \lesssim \frac{K^{3\epsilon^4}}{K^{2\epsilon}} M^{-\frac{1}{n+1}} \lambda^{\frac{n}{(n+1)(n+2)}} \sigma R^{\frac{\alpha}{(n+1)(n+2)}} \|f\|_2 ,$$

where the last inequality follows from (3.25). Since $K = R^\delta$ and $R$ can be assumed to be sufficiently large compared to any constant depending on $\varepsilon$, we have $\frac{K^{3\epsilon^4}}{K^{2\epsilon}} \ll 1$ and the induction closes for the narrow case. This completes the proof of Proposition 3.1.

3.3. Remark. In Section [2] we have seen that Corollary [1,6] is a direct result of Theorem [1,5] and they are equally useful in applications to the sharp $L^2$ estimate of Schrödinger maximal function. We can also prove Corollary [1,6] from scratch using a similar argument as in this section, which is slightly easier in two aspects compared to that of Theorem [1,5]. First, in the broad case, it is sufficient to use multilinear restriction estimates and not necessary to invoke multilinear refined Strichartz. Secondly, because there is one parameter less, the dyadic pigeonholing argument in the narrow case would be slightly reduced, for example, see Figure 2 for tubes of different scales in the $\square$ under the setting of Corollary [1,6].

**Figure 2.** Tubes of different scales in the $\square$ (in inductive argument for Corollary [1,6])
In fact, an adaptation of some arguments in the work [31] of Wolff on the Falconer distance set problem in dimension 2 can already imply Corollary 1.6 when \( n = 1 \). In the special case \( n = 1 \), the broad versus narrow dichotomy becomes the one on bilinear versus linear. To handle the linear part, the idea of induction on scales and splitting the ball into rectangular boxes “□” of size \( R \times R/K \) in our proof already existed in Wolff’s paper. We thank Hong Wang for pointing this out to us and sharing some notes on it.

References

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

[2] J. Bourgain, Some new estimates on oscillatory integrals, Essays on Fourier Analysis in Honor of Elias M. Stein (Princeton, NJ, 1991), Princeton Math. Ser., vol. 42, Princeton University Press, New Jersey, 1995, pp. 83-112.

[3] J. Bourgain, On the Schrödinger maximal function in higher dimension, Proceedings of the Steklov Institute of Math. 2013, vol. 280, pp. 46-60. (2012).

[4] J. Bourgain, A note on the Schrödinger maximal function, J. Anal. Math. 130 (2016), 393-396.

[5] J. Bourgain and C. Demeter, The proof of the \( L^2 \) decoupling conjecture, Ann. of Math. (2) 182 (2015), no. 1, 351-389.

[6] J. Bourgain and L. Guth, Bounds on oscillatory integral operators based on multilinear estimates, Geom. Funct. Anal. 21 (2011), no. 6, 1239-1295.

[7] A. Carbery, Radial Fourier multipliers and associated maximal functions, In Recent Progress in Fourier Analysis (El Escorial, 1983), vol. 111 of North-Holland Math. Stud. North-Holland, Amsterdam, 1985, pp. 49-56.

[8] L. Carleson, Some analytic problems related to statistical mechanics, Euclidean Harmonic Analysis (Proc. Sem., Univ. Maryland, College Park, Md, 1979), Lecture Notes in Math.779, pp. 5-45.

[9] M. Cowling, Pointwise behavior of solutions to Schrödinger equations, In Harmonic Analysis (Cortona, 1982), vol. 992 of Lecture Notes in Math. Springer, Berlin, 1983, pp. 83-90.

[10] B.E.J. Dahlberg and C.E. Kenig, A note on the almost everywhere behavior of solutions to the Schrödinger equation, Harmonic Analysis (Minneapolis, Minn, 1981), Lecture Notes in Math. 908, pp.205-209.

[11] C. Demeter and S. Guo, Schrödinger maximal function estimates via the pseudoconformal transformation, arXiv:1608.07640

[12] X. Du, L. Guth and X. Li, A sharp Schrödinger maximal estimate in \( \mathbb{R}^2 \), Annals of Math 186 (2017), 607-640.

[13] X. Du, L. Guth, X. Li and R. Zhang, Pointwise convergence of Schrödinger solutions and multilinear refined Strichartz estimate, preprint (2018), arXiv:1803.01720

[14] X. Du, L. Guth, Y. Ou, H. Wang, B. Wilson and R. Zhang, Weighted restriction estimates and application to Falconer distance set problem, preprint (2018), arXiv:1802.10186

[15] K. J. Falconer, On the Hausdorff dimensions of distance sets, Mathematika 32 (1985), no. 2, 206-212 (1986).

[16] L. Guth, A short proof of the multilinear Kakeya inequality, Math. Proc. Cambridge Philos. Soc. 158 (2015), no. 1, 147-153.

[17] L. Guth, A restriction estimate using polynomial partitioning, J. Amer. Math. Soc. 29(2) (2016), 371-413.

[18] L. Guth, Restriction estimates using polynomial partitioning II, preprint (2016), arXiv:1603.01425

[19] C.E. Kenig, G. Ponce and L. Vega, Oscillatory integrals and regularity of dispersive equations, Indiana Univ. Math. J. 40 (1991), no. 1, 33-69.

[20] S. Lee, On pointwise convergence of the solutions to Schrödinger equations in \( \mathbb{R}^2 \), Intern. Math. Research Notices. 2006, 32597, 1-21 (2006).

[21] B. Liu, An \( L^2 \)-identity and pinned distance problem, preprint (2018), arXiv:1802.00350

[22] R. Lucà and K. Rogers, Average decay for the Fourier transform of measures with applications, J. Eur. Math. Soc. (2016, to appear)
[23] R. Lucá and K. Rogers, Coherence on fractals versus pointwise convergence for the Schrödinger equation, Comm. Math. Phys. 351 (2017), no. 1, 341-359.
[24] R. Lucá and K. Rogers, A note on pointwise convergence for the Schrödinger equation, Math. Proc. Camb. Phil. Soc. (2017)
[25] P. Mattila, Hausdorff dimension, projections, and the Fourier transform., Publ. Mat. 48 (2004), no. 1, 3-48.
[26] A. Moyua, A. Vargas, and L. Vega, Schrödinger maximal function and restriction properties of the Fourier transform, International Math. Research Notices 1996 (16), 793-815 (1996).
[27] P. Sjögren and P. Sjölin, Convergence properties for the time-dependent Schrödinger equation, Ann. Acad. Sci. Fenn. 14 (1989), no. 1, 13-25.
[28] P. Sjölin, Regularity of solutions to the Schrödinger equation, Duke Math. Journal 55(3), 699-715 (1987).
[29] T. Tao and A. Vargas, A bilinear approach to cone multipliers. II. Applications, Geometric and Functional Analysis 10 (1), 185-215 (2000).
[30] L. Vega, Schrödinger equations: pointwise convergence to the initial data, Proceedings of the American Mathematical Society 102 (4), 874-878 (1988).
[31] T. Wolff, Decay of circular means of Fourier transforms of measures, Int. Math. Res. Not. 10 (1999), 547-567.

Institute for Advanced Study, Princeton, NJ
E-mail address: xdu@math.ias.edu

Institute for Advanced Study, Princeton, NJ
E-mail address: rzhang@math.ias.edu