A SHARP SCHRÖDINGER MAXIMAL ESTIMATE IN $\mathbb{R}^2$

XIUMIN DU, LARRY GUTH, AND XIAOCHUN LI

Abstract. We show that $\lim_{t\to 0} e^{it\Delta} f(x) = f(x)$ almost everywhere for all $f \in H^s(\mathbb{R}^2)$ provided that $s > 1/3$. This result is sharp up to the endpoint. The proof uses polynomial partitioning and decoupling.

1. Introduction

The solution to the free Schrödinger equation

$$\begin{cases} iu_t - \Delta u = 0, & (x,t) \in \mathbb{R}^n \times \mathbb{R} \\ u(x,0) = f(x), & x \in \mathbb{R}^n \end{cases} \tag{1.1}$$

is given by

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

We consider the following problem posed by Carleson in [5]: determine the optimal s for which $\lim_{t\to 0} e^{it\Delta} f(x) = f(x)$ almost everywhere whenever $f \in H^s(\mathbb{R}^n)$. Our main result is the following:

Theorem 1.1. *For every $f \in H^s(\mathbb{R}^2)$ with $s > 1/3$, $\lim_{t\to 0} e^{it\Delta} f(x) = f(x)$ almost everywhere.*

Recently, Bourgain [3] gave examples showing that such convergence can fail for any $s < 1/3$, and so Theorem 1.1 is sharp up to the endpoint.

This problem originates from Carleson [5], who proved convergence for $s \geq 1/4$ when $n = 1$. Dahlberg and Kenig [6] showed that the convergence does not hold for $s < 1/4$ in any dimension. Sjölin [18] and Vega [20] proved independently the convergence for $s > 1/2$ in all dimensions. The sufficient condition for pointwise convergence was improved by Bourgain [1], Moyua-Vargas-Vega [16], and Tao-Vargas [19]. The best known sufficient condition in dimension $n = 2$ was $s > 3/8$, due to Lee [13] using Tao-Wolff's bilinear restriction method. In general dimension $n \geq 2$, Bourgain [2] showed the convergence for $s > 1/2 - 1/(4n)$, using multilinear methods. When $n = 2$, this approach gives a different proof of Lee's result for $s > 3/8$.

For many years, it had seemed plausible that convergence actually holds for $s > 1/4$ in every dimension. Only in 2012, Bourgain [2] gave a counterexample showing that this is false in sufficiently high dimensions. Improved counterexamples were given by Lucá-Rogers [14] [15] and Demeter-Guo [7]. Very recently, in [3], Bourgain gave counterexamples showing that convergence can fail if $s < \frac{n}{2(n+1)}$. In particular, for $n = 2$, convergence can fail if $s < 1/3$.

We will follow the standard approach by bounding the associated maximal function. We use $B^n(c, r)$ to represent a ball centered at c with radius r in $\mathbb{R}^n$, and

use χ_E to denote the characteristic function of any measurable set E. For brevity, $B(c,r)$ represents $B^2(c,r)$, a ball in $\mathbb{R}^2$.

Theorem 1.2. *For any $s > 1/3$, the following bound holds: for any function $f \in H^s(\mathbb{R}^2)$,*

$$\left\| \sup_{0<t\leq 1} |e^{it\Delta} f| \right\|_{L^3(B(0,1))} \leq C_s \|f\|_{H^s(\mathbb{R}^2)}. \tag{1.2}$$

If the support of $\widehat{f}$ lies in $A(R) = \{\xi \in \mathbb{R}^2 : |\xi| \sim R\}$, then Theorem 1.2 boils down to the bound

$$\left\| \sup_{0<t\leq 1} |e^{it\Delta} f| \right\|_{L^3(B(0,1))} \leq C_\epsilon R^{1/3+\epsilon} \|f\|_{L^2}. \tag{1.3}$$

After parabolic rescaling, this bound reduces to the following estimate for functions f with $\widehat{f}$ supported in $A(1)$.

Theorem 1.3. *For any $\epsilon > 0$, there exists a constant C_ϵ such that*

$$\left\| \sup_{0<t\leq R} |e^{it\Delta} f| \right\|_{L^3(B(0,R))} \leq C_\epsilon R^\epsilon \|f\|_2 \tag{1.4}$$

holds for all $R \geq 1$ and all f with $\mathrm{supp}\widehat{f} \subset A(1) = \{\xi \in \mathbb{R}^2 : |\xi| \sim 1\}$.

Here is an outline of the proof of Theorem 1.3. The proof uses polynomial partitioning. This technique was introduced by Nets Katz and the second author in [8], where it was applied to incidence geometry. In [9] and [10], the second author applied this technique to restriction estimates in Fourier analysis. Polynomial partitioning is a divide and conquer technique. We begin by finding a polynomial whose zero set divides some object of interest into equal pieces. For instance, in [8], it was proven that for any finite volume set $E \subset \mathbb{R}^3$ and any degree $D \geq 1$, there is a polynomial P of degree at most D so that $\mathbb{R}^3 \setminus Z(P)$ is a union of $\sim D^3$ disjoint open sets O_i, and the volumes $|O_i \cap E|$ are all equal. Hence for any i, $|E| \lesssim D^3 |O_i \cap E|$. In our paper, we choose the polynomial P to behave well with respect to the $L^p_x L^q_t$ norm of $e^{it\Delta} f$. For any $p \leq q < \infty$ and any degree $D \geq 1$, we show that there is a polynomial P of degree at most D so that $\mathbb{R}^3 \setminus Z(P)$ is a union of $\sim D^3$ disjoint open sets O_i, and for any i,

$$\|e^{it\Delta} f\|^p_{L^p_x L^q_t (B(0,R)\times[0,R])} \lesssim D^3 \|\chi_{O_i} e^{it\Delta} f\|^p_{L^p_x L^q_t (B(0,R)\times[0,R])}. \tag{1.5}$$

(To prove Theorem 1.3, we will use q finite but very large and p close to 3. The degree D will be a tiny power of R, so D is large compared to 1, but very small compared to R.)

Breaking spacetime into cells O_i is useful because of the way it interacts with the wave packet decomposition of $e^{it\Delta} f$, which we now recall. We decompose f into pieces that are localized in both physical space and frequency space. We tile the physical space $B(0,R)$ with $R^{1/2}$-cubes ν, and we tile the frequency space $B(0,1)$ with $R^{-1/2}$-cubes θ. Then we decompose f as $f = \sum_{\theta,\nu} f_{\theta,\nu}$, where $f_{\theta,\nu}$ is essentially supported on ν in physical space and essentially supported on θ in frequency space. Each function $e^{it\Delta} f_{\theta,\nu}$ is called a wave packet. The restriction of $e^{it\Delta} f_{\theta,\nu}$ to the domain $B(0,R) \times [0,R]$ is essentially supported on a tube $T_{\theta,\nu}$ of

radius $R^{1/2}$ and length R. This tube intersects the time slice $\{t = 0\}$ at ν, and the direction of the tube depends on θ.

A key fact in the applications of polynomial partitioning in combinatorics is that a line can enter at most $D + 1$ of the cells O_i. To see this, we note that the polynomial P can vanish at most D times along a line, unless it vanishes on the whole line, and so a line can cross $Z(P)$ at most D times. A wave packet $e^{it\Delta} f_{\theta,\nu}$ is supported on a tube $T_{\theta,\nu}$ of radius $R^{1/2}$. This tube can potentially enter many or even all the cells O_i, but it cannot penetrate deeply into very many cells. We define W to be the $R^{1/2}$-neighborhood of $Z(P)$ in $B(0, R) \times [0, R]$, and we define O_i' to be $O_i \setminus W$. Now the central line of $T_{\theta,\nu}$ can enter at most $D+1$ of the original cells O_i, and so the tube $T_{\theta,\nu}$ can enter at most $D + 1$ of the smaller cells O_i'. In other words, each wave packet $e^{it\Delta} f_{\theta,\nu}$ is essentially supported on the union of W and $D + 1$ cells O_i'.

We can use induction to study $e^{it\Delta} f$ on each smaller cell O_i'. To study $e^{it\Delta} f$ on a cell O_i', we only need to take account of those wave packets that intersect O_i'. Therefore, we define f_i to be the sum of $f_{\theta,\nu}$ over those pairs (θ, ν) for which $T_{\theta,\nu}$ enters O_i'. On the cell O_i', $e^{it\Delta} f$ is essentially equal to $e^{it\Delta} f_i$. We can control the L^2 norms of the f_i by using the fact that $f_{\theta,\nu}$ are (approximately) orthogonal and the fact that each tube $T_{\theta,\nu}$ enters $\lesssim D$ smaller cells O_i'. In particular, we will prove that

$$\sum_i \|f_i\|_2^2 \lesssim D\|f\|_2^2.$$

We can now use induction to control $e^{it\Delta} f$ on each cell O_i'. In this way, we get good control of the contribution to $\|e^{it\Delta} f\|_{L_x^p L_t^q(B(0,R)\times[0,R])}$ coming from the union of all smaller cells O_i'. It remains to control the contribution coming from W.

The most difficult scenario is the following: $e^{it\Delta} f$ is a sum of wave packets $e^{it\Delta} f_{\theta,\nu}$ for which the tubes $T_{\theta,\nu}$ are all contained in W. The polynomial partitioning method allows us to reduce the original problem to this special scenario. This scenario indeed occurs in Bourgain's example in [3]. Let us take a moment to describe this example.

In the example from [3], the zero set $Z(P)$ can be taken to be a plane $t = x_1$. The set W is a planar slab of thickness $R^{1/2}$. The solution $e^{it\Delta} f$ is essentially supported in W. On the plane $t = x_1$, $e^{it\Delta} f$ is a solution of the Schrödinger equation in 1 + 1 dimensions. In other words, we can choose coordinates (y, s) on this plane and an initial data g so that $e^{is\Delta} g$ is essentially equal to $e^{it\Delta} f$ on the plane. Also, $|e^{it\Delta} f(x_1, x_2)|$ is approximately constant as we vary x_1 within the slab W. The initial data is chosen so that $|e^{is\Delta} g(y)|$ is large on a set X of $\sim R^{3/2}$ unit squares in $[0, R] \times [0, R]$. It follows that $|e^{it\Delta} f(x)|$ is large on a set of $\sim R^{3/2}$ 3-dimensional rectangles of dimensions $R^{1/2} \times 1 \times 1$ in $B(0, R) \times [0, R]$. Moreover, the projections of these rectangles are roughly disjoint, and so they cover a positive proportion of $B(0, R)$. Therefore $\sup_{0<t<R} |e^{it\Delta} f(x)|$ is large on a positive proportion of $B(0, R)$.

In this construction, the set X needs to be fairly sparse because the projections of the $R^{1/2} \times 1 \times 1$ rectangles need to be disjoint in $B(0, R)$. In particular, there can be at most $R^{1/2}$ unit squares of X in any $R^{1/2}$-ball in $[0, R] \times [0, R]$. In the example of [3], $|e^{is\Delta} g| \sim R^{-5/12}\|g\|_{L^2([0,R])}$ on the set X. During our proof, we will need to show that this quantity $R^{-5/12}\|g\|_{L^2}$ could not be any larger. In rough

terms, we need to show that a solution $e^{is\Delta}g$ cannot focus too much on a set X which is sparse and spread out.

We will prove such bounds using the l^2 decoupling theorem of Bourgain and Demeter [4]. We think of these bounds as refinements of the Strichartz inequality. Here is one such estimate:

Theorem 1.4. *Suppose that $g : \mathbb{R} \to \mathbb{C}$ has frequency supported in $B^1(0,1)$. Suppose that $Q_1, Q_2, ...$ are lattice $R^{1/2}$-cubes in $[0,R]^2$, so that*

$$\|e^{it\Delta}g\|_{L^6(Q_j)} \text{ is essentially constant in } j.$$

Suppose that these cubes are arranged in horizontal strips of the form $\mathbb{R} \times \{t_0, t_0 + R^{1/2}\}$, and that each strip contains $\sim \sigma$ cubes Q_j. Let Y denote $\bigcup_j Q_j$. Then for any $\epsilon > 0$,

$$\|e^{it\Delta}g\|_{L^6(Y)} \le C_\epsilon R^\epsilon \sigma^{-1/3} \|g\|_{L^2}.$$

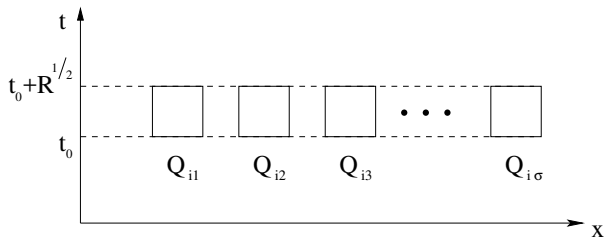

FIGURE 1. $\sim \sigma$ many cubes in a horizontal strip

The Strichartz inequality says that $\|e^{it\Delta}g\|_{L^6([0,R]^2)} \lesssim \|g\|_{L^2}$. Theorem 1.4 says that we get a stronger estimate when the solution $e^{it\Delta}g$ is spread out in space. To get a sense of what the theorem says, consider the following example. Suppose that $e^{it\Delta}g$ is a sum of σ wave packets supported on disjoint $R^{1/2} \times R$ rectangles. We can take Y to be the union of these rectangles. By scaling, we can suppose that $|e^{it\Delta}g| \sim 1$ on these σ rectangles and negligibly small elsewhere, and then a direct calculation shows that $\|e^{it\Delta}g\|_{L^6(Y)} \sim \|e^{it\Delta}g\|_{L^6([0,R]^2)} \sim \sigma^{-1/3}\|g\|_{L^2([0,R])}$. So Theorem 1.4 roughly says that if $e^{it\Delta}g$ is "as spread out as" σ disjoint wave packets, then its L^6 norm cannot be much bigger than the L^6 norm of σ disjoint wave packets.

This theorem helps us to control the size of $e^{it\Delta}g$ on a sparse, spread out set X as above. Suppose that the function $e^{it\Delta}g$ is evenly spread out on $[0,R]^2$ in the sense that $\|e^{it\Delta}g\|_{L^6(Q)}$ is roughly constant among all $R^{1/2}$-boxes $Q \subset [0,R]^2$. In this case, we can take $\sigma = R^{1/2}$ in Theorem 1.4, which gives

$$\|e^{it\Delta}g\|_{L^6([0,R]^2)} \lesssim R^{-1/6+\epsilon}\|g\|_{L^2}.$$

In the example from [3], X contains $\sim R^{1/2}$ unit squares in each $R^{1/2}$-box of $[0,R]^2$, and each of these boxes indeed has a roughly equal value of $\|e^{it\Delta}g\|_{L^6(Q)}$. If $|e^{it\Delta}g| \sim H$ on the set X, then Theorem 1.4 gives

$$H|X|^{1/6} \lesssim \|e^{it\Delta}g\|_{L^6([0,R]^2)} \lesssim R^{-1/6+\epsilon}\|g\|_{L^2}.$$

Since $|X| \sim R^{3/2}$, we get the bound $H \lesssim R^{-5/12+\epsilon}\|g\|_{L^2}$. This upper bound matches the behavior of the example from [3] up to a factor R^ϵ.

Theorem 1.4 lets us deal with the case that $Z(P)$ is a plane. We need to deal with the more general case that $Z(P)$ is a possibly curved surface of degree at most D. We prove a more general version of Theorem 1.4, Theorem 7.1, which covers the case of wave packets concentrated into a curved surface.

Acknowledgements. *The second author is supported by a Simons Investigator grant.*

2. Main inductive theorem

Here we state a slightly more complicated theorem which will imply all the theorems in the introduction. Our proof uses induction, and we need the slightly more complicated formulation to make all the inductions work. First of all, the polynomial partitioning involves a topological argument, and the topological argument does not work well with the sup appearing in our maximal function. Therefore, we replace the norm $L^p_x L^\infty_t$ with the norm $L^p_x L^q_t$ for q very large. Another technical issue has to do with parabolic rescaling. Suppose that $\widehat{f}$ is supported in a smaller ball $B(\xi_0, M^{-1}) \subset B(0,1)$. In this situation, one can often apply parabolic rescaling to reduce the problem at hand to a problem on a smaller ball in physical space. However, the change of coordinates in such a parabolic rescaling does not interact well with mixed norms of the form $L^p_x L^q_t$. Therefore, we instead do induction on the size of the ball $B(\xi_0, M^{-1})$, proving slightly stronger bounds when the ball is small. Taking account of these small issues, we formulate our result in the following way:

Theorem 2.1. *For $p > 3$, for any $\epsilon > 0$, there exists a constant C_ϵ such that for any $q > 1/\epsilon^4$,*

$$\left\|e^{it\Delta} f\right\|_{L^p_x L^q_t(B(0,R)\times[0,R])} \leq C_{p,\epsilon} M^{-\epsilon^2} R^\epsilon \|f\|_2 \tag{2.6}$$

holds for all $R \geq 1$, any $\xi_0 \in B^2(0,1)$, any $M \geq 1$ and all f with $\mathrm{supp}\widehat{f} \subset B^2(\xi_0, M^{-1})$.

Let us quickly explain how Theorem 2.1 implies the theorems in the introduction. We note that by the dominated convergence theorem we have

$$\left\| \sup_{0<t\leq R} |e^{it\Delta} f| \right\|_{L^p(B(0,R))} = \lim_{q\to\infty} \left\|e^{it\Delta} f\right\|_{L^p_x L^q_t(B(0,R)\times[0,R])},$$

for any L^2-function f with compact Fourier support or any Schwartz function f. Therefore, Theorem 2.1 implies that for any $R \geq 1$ and any f with the support of $\widehat{f} \subset B(0,1)$, and for any $p > 3$, we have

$$\left\|e^{it\Delta} f\right\|_{L^p_x L^\infty_t(B(0,R)\times[0,R])} \leq C_{p,\epsilon} R^\epsilon \|f\|_2 \tag{2.7}$$

So far we assume $p > 3$. But it is straightforward to prove a bound of the form

$$\|e^{it\Delta} f\|_{L^2_x L^\infty_t(B(0,R)\times[0,R])} \leq R^{O(1)} \|f\|_2.$$

Combining these bounds using Hölder's inequality, we see that Equation (2.7) holds for $p = 3$ as well. This establishes Theorem 1.3.

We write $A \lessapprox B$ if $A \leq C_\epsilon R^\epsilon B$ for any $\epsilon > 0$. Suppose now that $\widehat{g}$ is supported in $A(R)$. To prove Theorem 1.2, we want to show that

$$\left\| \sup_{0<t\leq 1} |e^{it\Delta} g| \right\|_{L^3(B(0,1))} \lessapprox R^{1/3} \|g\|_{L^2}. \tag{2.8}$$

After parabolic rescaling, we are led to a function f with $\widehat{f}$ supported in $A(1)$, and we need to show the bound

$$\left\| \sup_{0<t\leq R^2} |e^{it\Delta} f| \right\|_{L^3(B(0,R))} \lessapprox \|f\|_{L^2}.$$

But applying Theorem 1.3 with R^2 in place of R gives:

$$\left\| \sup_{0<t\leq R^2} |e^{it\Delta} f| \right\|_{L^3(B(0,R^2))} \lessapprox \|f\|_{L^2}.$$

This implies Equation (2.8). Now, given $s > 1/3$ and $f \in H^s(\mathbb{R}^2)$, we decompose f in a Littlewood-Paley decomposition: $f = \sum_{k\geq 0} f_k$ where $\widehat{f_0}$ is supported in $B(0,1)$ and $\widehat{f_k}$ is supported in $A(2^k)$ for $k \geq 1$. We have $\|f_k\|_{L^2} \lesssim 2^{-ks}\|f\|_{H^s}$. Applying (2.8) to each f_k and using the triangle inequality, we get Theorem 1.2.

Theorem 1.2 implies Theorem 1.1 by a standard smooth approximation argument, which we briefly recall. If f is Schwartz, then it is well-known that $e^{it\Delta} f(x) \to f(x)$ uniformly in x. Schwartz functions are dense in H^s, and so we can write $f = g + h$ where g is Schwartz and $\|h\|_{H^s} < \epsilon^{100}$. Since g is Schwartz, we can find a time $t_\epsilon > 0$ so that $|e^{it\Delta} g(x) - g(x)| < \epsilon$ for all x and all $0 \leq t \leq t_\epsilon$. On the other hand, by the maximal estimate in Theorem 1.2, $|e^{it\Delta} h(x)| < \epsilon$ for all $0 \leq t \leq 1$ and all x in $B(0,1) \setminus X_\epsilon$, where $|X_\epsilon| < \epsilon$. Taking a sequence of $\epsilon \to 0$ exponentially fast, and doing a little measure theory, it follows that $e^{it\Delta} f(x) \to f(x)$ for almost every $x \in B(0,1)$. The same applies to any other ball, and we see that $e^{it\Delta} f(x) \to f(x)$ for almost every $x \in \mathbb{R}^2$.

We also remark that the local bound (1.2) from Theorem 1.2 can be used to derive immediately a global estimate in $L^3(\mathbb{R}^2)$ for the maximal function $\sup_{0<t\leq 1} |e^{it\Delta} f|$, following from Theorem 10 in [17]. We are indebted to K. Rogers for pointing this out to us.

In the rest of the paper, we prove Theorem 2.1. In Section 3, we review polynomial partitioning, and in Section 4, we review wave packet decomposition. Then we begin the proof of Theorem 2.1 in Section 5.

3. Polynomial Partitioning

First we state a variation of the ham-sandwich theorem, which introduces a polynomial P in the polynomial ring $\mathbb{R}[x,t]$ such that the variety $Z(P) = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : P(x,t) = 0\}$ bisects every member in a collection of some quantities. It relies on Borsuk-Ulam Theorem, which asserts that *if $F : \mathbb{S}^N \to \mathbb{R}^N$ is a continuous function, where $\mathbb{S}^N$ is the N-dimensional unit sphere, then there exists a point $v \in \mathbb{S}^N$ with $F(v) = F(-v)$.*

Lemma 3.1. *If $W_1, W_2, \cdots, W_N \in L^1_x L^r_t(\mathbb{R}^n \times \mathbb{R}), 1 \le r < \infty$, then there exists a non-zero polynomial P on $\mathbb{R}^n \times \mathbb{R}$ of degree $\le c_n N^{1/(n+1)}$ such that for each W_j,*

$$\big\|\chi_{\{P>0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} = \big\|\chi_{\{P<0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})}.$$

Proof. Let V be the vector space of polynomials on $\mathbb{R}^n \times \mathbb{R}$ of degree at most D, then

$$\mathrm{Dim} V = \binom{D+n+1}{n+1} \sim_n D^{n+1}.$$

So we can choose $D \sim N^{1/(n+1)}$ such that $\mathrm{Dim} V \ge N+1$, and without loss of generality we can assume $\mathrm{Dim} V = N+1$ and identify V with $\mathbb{R}^{N+1}$. We define a function G as follows:

$$\begin{aligned} \mathbb{S}^N \subseteq V\backslash\{0\} &\xrightarrow{G} \mathbb{R}^N \\ P &\mapsto \{G_j(P)\}_{j=1}^N, \end{aligned}$$

where

$$G_j(P) := \big\|\chi_{\{P>0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} - \big\|\chi_{\{P<0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})},$$

it is obvious that $G(-P) = -G(P)$. Assume that the function G is continuous, then Borsuk-Ulam Theorem tells us that there exists $P \in \mathbb{S}^N \subseteq V\backslash\{0\}$ with $G(P) = G(-P)$, hence $G(P) = 0$, and P obeys the conclusion of Lemma 3.1. It remains to check the continuity of the functions G_j on $V\backslash\{0\}$.

Suppose that $P_k \to P$ in $V\backslash\{0\}$. Note that

$$|G_j(P_k) - G_j(P)| \le 2\big\|\chi_{\{P_k P \le 0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})},$$

while $P_k \to P$ implies that

$$\bigcap_{k_0} \bigcup_{k \ge k_0} \{(x,t) : P_k(x,t)\cdot P(x,t) \le 0\} \subseteq P^{-1}(0).$$

By the dominated convergence theorem,

$$\lim_{k_0\to\infty} \big\|\chi_{\cup_{k\ge k_0}\{P_k P\le 0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} = \big\|\chi_{\{P^{-1}(0)\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} = 0.$$

This proves that $\lim_{k\to\infty} |G_j(P_k) - G_j(P)| = 0$, showing that G_j is continuous on $V\backslash\{0\}$. □

By applying Lemma 3.1 repeatedly, we get the following polynomial partitioning result:

Theorem 3.2. *If $W \in L^1_x L^r_t(\mathbb{R}^n \times \mathbb{R})\backslash\{0\}, 1 \le r < \infty$, then for each D there exists a non-zero polynomial P of degree at most D such that $(\mathbb{R}^n \times \mathbb{R})\backslash Z(P)$ is a union of $\sim_n D^{n+1}$ disjoint open sets O_i and for each i we have*

$$\big\|W\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} \le c_n D^{n+1} \big\|\chi_{O_i} W\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})}.$$

Proof. By Lemma 3.1, we obtain a polynomial P_1 of degree $\lesssim 1$ such that

$$\big\|\chi_{\{P_1>0\}} W\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} = \big\|\chi_{\{P_1<0\}} W\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})}.$$

Next, we let $W_+ := \chi_{\{P_1>0\}} W$ and $W_- := \chi_{\{P_1<0\}} W$, and by Lemma 3.1 again we obtain a polynomial P_2 of degree $\lesssim 2^{1/(n+1)}$ such that

$$\big\|\chi_{\{P_2>0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})} = \big\|\chi_{\{P_2<0\}} W_j\big\|_{L^1_x L^r_t(\mathbb{R}^n\times\mathbb{R})},$$

for $j = +, -$. Continuing inductively, we construct polynomials $P_1, P_2, \cdots, P_s$. Let $P := \prod_{k=1}^{s} P_k$. The sign conditions of the polynomials cut $(\mathbb{R}^n \times \mathbb{R})\backslash Z(P)$ into 2^s cells O_i, and by construction and triangle inequality we have that, for each i,

$$\|W\|_{L_x^1 L_t^r(\mathbb{R}^n \times \mathbb{R})} \leq 2^s \|\chi_{O_i} W\|_{L_x^1 L_t^r(\mathbb{R}^n \times \mathbb{R})}.$$

By construction, $\deg P_k \lesssim 2^{(k-1)/(n+1)}$, therefore $\deg P \leq c_n 2^{s/(n+1)}$. We can choose s such that $c_n 2^{s/(n+1)} \in [D/2, D]$, then $\deg P \leq D$ and the number of cells $2^s \sim_n D^{n+1}$. □

Definition 3.3. *We say that a polynomial P is non-singular if $\nabla P(z) \neq 0$ for each point z in $Z(P)$.*

It is well-known that non-singular polynomials are dense in the space of all polynomials, cf. Lemma 1.5 in [9]. Following from the density of non-singular polynomials and the proof of Theorem 3.2, we can assume that the polynomial in the partitioning theorem enjoys nice geometric properties.

Theorem 3.4. *If $W \in L_x^1 L_t^r(\mathbb{R}^n \times \mathbb{R})\backslash\{0\}, 1 \leq r < \infty$, then for each D there exists a non-zero polynomial P of degree at most D such that $(\mathbb{R}^n \times \mathbb{R})\backslash Z(P)$ is a union of $\sim_n D^{n+1}$ disjoint open sets O_i and for each i we have*

$$\|W\|_{L_x^1 L_t^r(\mathbb{R}^n \times \mathbb{R})} \leq c_n D^{n+1} \|\chi_{O_i} W\|_{L_x^1 L_t^r(\mathbb{R}^n \times \mathbb{R})}.$$

Moreover, the polynomial P is a product of distinct non-singular polynomials.

4. Wave Packet Decomposition

We focus on the dimension $n = 2$ in the rest of the paper.

A (dyadic) rectangle in $\mathbb{R}^2$ is a product of (dyadic) intervals with respect to given coordinate axes of $\mathbb{R}^2$. A rectangle $\theta = \prod_{j=1}^{2} \theta_j$ in frequency space and a rectangle $\nu = \prod_{j=1}^{2} \nu_j$ in physical space are said to be dual if $|\theta_j||\nu_j| = 1$ for $j = 1, 2$. We say that (θ, ν) is a tile if it is a pair of *dual* (dyadic) rectangles. The dyadic condition is not essential in our decomposition.

Let φ be a Schwartz function from $\mathbb{R}$ to $\mathbb{R}$ whose Fourier transform is non-negative, supported in a small interval, of radius κ (κ is a fixed small constant), about the origin in $\mathbb{R}$, and identically 1 on another smaller interval around the origin. For a (dyadic) rectangular box $\theta = \prod_{j=1}^{2} \theta_j$, set

$$\widehat{\varphi_\theta}(\xi_1, \xi_2) = \prod_{j=1}^{2} \frac{1}{|\theta_j|^{1/2}} \widehat{\varphi}\left(\frac{\xi_j - c(\theta_j)}{|\theta_j|}\right). \tag{4.1}$$

Here $c(\theta_j)$ is the center of the interval θ_j and hence $c(\theta) = (c(\theta_1), c(\theta_2))$ is the center of the rectangle θ. We also note that $\|\varphi_\theta\|_{L^2} \sim 1$. We let $c(\nu)$ denote the center of ν. For a tile (θ, ν) and $x \in \mathbb{R}^2$, we define

$$\widehat{\varphi_{\theta,\nu}}(\xi) = e^{2\pi i c(\nu)\cdot\xi} \widehat{\varphi_\theta}(\xi). \tag{4.2}$$

We say that two tiles (θ, ν) and (θ', ν') have the same dimensions if $|\theta_j| = |\theta'_j|$ for all j, which then implies that $|\nu_j| = |\nu'_j|$ for all j. Let $\mathbf{T}$ be a collection of all tiles with fixed dimensions and coordinate axes. Then for any Schwartz function f from $\mathbb{R}^2$ to $\mathbb{R}$, we have the following representation

$$f(x) = c_\kappa \sum_{(\theta,\nu)\in\mathbf{T}} f_{\theta,\nu} := c_\kappa \sum_{(\theta,\nu)\in\mathbf{T}} \langle f, \varphi_{\theta,\nu}\rangle \varphi_{\theta,\nu}(x), \tag{4.3}$$

where c_κ is an absolute constant. This representation can be proved directly (see [11]) or by employing inductively the one-dimensional result in [12].

We will only use tiles (θ,ν) where θ is an $R^{-\frac{1}{2}}$-cube in frequency space and ν is an $R^{\frac{1}{2}}$-cube in physical space. Indeed, let θ be an $R^{-\frac{1}{2}}$-cube (or ball) in $B(0,1)\subset\mathbb{R}^2$. Let $\mathbf{T}_\theta$ be a collection of all tiles (θ',ν) such that ν's are $R^{\frac{1}{2}}$-cubes and $\theta'=\theta$. Then for any Schwartz function f with $\mathrm{supp}\widehat{f}\subset B(0,1)$, we have

$$f(x)=c_\kappa\sum_{\theta}\sum_{(\theta',\nu)\in\mathbf{T}_\theta}\langle f,\varphi_{\theta',\nu}\rangle\varphi_{\theta',\nu}(x)\,. \tag{4.4}$$

Here θ's range over all possible cubes in $\mathrm{supp}\widehat{f}$. We use $\mathbf{T}$ to denote $\cup_\theta\mathbf{T}_\theta$. It is clear that

$$\sum_{(\theta,\nu)\in\mathbf{T}}\big|\langle f,\varphi_{\theta,\nu}\rangle\big|^2\sim\|f\|_2^2\,. \tag{4.5}$$

We set

$$\psi_{\theta,\nu}(x,t)=e^{it\Delta}\varphi_{\theta,\nu}(x)\,. \tag{4.6}$$

From (4.4), we end up with the following representation for $e^{it\Delta}f$:

$$e^{it\Delta}f(x)=c_\kappa\sum_{(\theta,\nu)\in\mathbf{T}}e^{it\Delta}f_{\theta,\nu}(x)=c_\kappa\sum_{(\theta,\nu)\in\mathbf{T}}\langle f,\varphi_{\theta,\nu}\rangle\psi_{\theta,\nu}(x,t)\,. \tag{4.7}$$

We shall analyze the localization of $\psi_{\theta,\nu}$ in the physical and frequency space.

On the domain $B(0,R)\times[0,R]$, the function $\psi_{\theta,\nu}$ is essentially supported on a tube $T_{\theta,\nu}$ defined as follows. Let

$$T_{\theta,\nu}:=\{(x,t)\in\mathbb{R}^2\times\mathbb{R}\,:\,0\le t\le R,|x-c(\nu)+2tc(\theta)|\le R^{1/2+\delta}\}\,, \tag{4.8}$$

where $\delta=\epsilon^2$ is a small positive parameter. We see that $T_{\theta,\nu}$ is a tube of length R, of radius $R^{1/2+\delta}$, in the direction $G_0(\theta)=(-2c(\theta),1)$, and intersecting $\{t=0\}$ at an $R^{1/2+\delta}$-ball centered at $c(\nu)$. In order to see this, let ψ be a Schwartz function with Fourier transform supported in $[-1,1]$ and $2\psi(t)\ge\chi_{[0,1]}(t)$. Here $\chi_{[0,1]}$ is the characteristic function on $[0,1]$. On $B(0,R)\times[0,R]$, we have $|\psi_{\theta,\nu}|\le 2|\psi^*_{\theta,\nu}|$, where

$$\psi^*_{\theta,\nu}(x,t)=\psi_{\theta,\nu}(x,t)\psi\big(\frac{t}{R}\big)\,. \tag{4.9}$$

From the definitions of $e^{it\Delta}$ and $\psi_{\theta,\nu}$, it is easy to check that, by integration by parts, $\psi^*_{\theta,\nu}$ is essentially supported in the tube $T_{\theta,\nu}$. More precisely, we have

$$|\psi^*_{\theta,\nu}(x,t)|\le\frac{1}{\sqrt{R}}\chi^*_{T_{\theta,\nu}}(x,t)\,, \tag{4.10}$$

where $\chi^*_{T_{\theta,\nu}}$ denotes a bump function satisfying that $\chi^*_{T_{\theta,\nu}}=1$ on $\{(x,t)\in\mathbb{R}^2\times\mathbb{R}:0\le t\le R,|x-c(\nu)+2tc(\theta)|\le\sqrt{R}\}$, and $\chi^*_{T_{\theta,\nu}}=O(R^{-1000})$ outside $T_{\theta,\nu}$. We can essentially treat $\chi^*_{T_{\theta,\nu}}$ as $\chi_{T_{\theta,\nu}}$, the indicator function on the tube $T_{\theta,\nu}$.

On the other hand, the Fourier transform of $\psi^*_{\theta,\nu}$ enjoys

$$\widehat{\psi^*_{\theta,\nu}}(\xi_1,\xi_2,\xi_3)=R\widehat{\varphi_{\theta,\nu}}(\xi_1,\xi_2)\widehat{\psi}\left(\frac{\xi_3-(\xi_1^2+\xi_2^2)}{1/R}\right)\,. \tag{4.11}$$

Hence $\widehat{\psi^*_{\theta,\nu}}$ is supported in the $\frac{1}{R}$-neighborhood of the parabolic cap over θ, that is,

$$\operatorname{supp} \widehat{\psi^*_{\theta,\nu}} \subseteq \big\{(\xi_1,\xi_2,\xi_3) : (\xi_1,\xi_2) \in \theta, |\xi_3 - (\xi_1^2+\xi_2^2)| \le \frac{1}{R}\big\}. \tag{4.12}$$

We denote this $\frac{1}{R}$-neighborhood of the parabolic cap over θ by θ^*. In the rest of the paper, we can assume that the function $\psi_{\theta,\nu}$ is essentially localized in $T_{\theta,\nu}$ in physical space, and θ^* in frequency space.

5. Cell contributions

The rest of the paper is devoted to a proof of Theorem 2.1, using polynomial partitioning. Recall that the functions f in Theorem 2.1 are Fourier supported in $B(\xi_0, M^{-1}) \subset \mathbb{R}^2$ with arbitrary $\xi_0 \in B(0,1)$ and $M \ge 1$. Also $p > 3$ and $q > \epsilon^{-4}$. The function f can be assumed to be a Schwartz function since the collection of all Schwartz functions is dense in L^2. We need to prove the bound (2.6):

$$\left\|e^{it\Delta} f\right\|_{L^p_x L^q_t (B(0,R)\times[0,R])} \le C_{p,\epsilon} M^{-\epsilon^2} R^{\epsilon} \|f\|_2 .$$

The proof of Theorem 2.1 is by induction on the radius R in physical space and the radius $1/M$ in frequency space. First we cover the bases of the induction. If $M \ge R^{10}$, then we bound $|e^{it\Delta} f(x)|$ by $M^{-1}\|f\|_2$ and Theorem 2.1 is trivial. If $R^{1/2-O(\delta)} < M < R^{10}$, then all associated wave packets are in the same direction, and by a direct computation we can bound the left-hand side of (2.6) by $R^{(3-p)/(2p)+O(\delta)}\|f\|_2$, from which Theorem 2.1 follows immediately. Therefore we can assume that $M \ll \sqrt{R}$. We can assume that R is sufficiently large, otherwise Theorem 2.1 is trivial. This covers the base of the induction. Now we turn to the inductive step. By induction, we can assume that Theorem 2.1 holds for physical radii less than $R/2$ or for physical radius R and frequency radius less than $\frac{1}{2M}$.

Let $\mathbf{B}^*_R$ denote the set $B(0,R)\times[0,R]$.

We pick a degree $D = R^{\epsilon^4}$, and apply polynomial partitioning with this degree to the function $\chi_{\mathbf{B}^*_R}|e^{it\Delta} f(x)|^p$. By Theorem 3.4 with $r = q/p$, there exists a non-zero polynomial P of degree at most D such that $(\mathbb{R}^2\times\mathbb{R})\backslash Z(P)$ is a union of $\sim D^3$ disjoint open sets O_i and for each i we have

$$\left\|e^{it\Delta} f(x)\right\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)} \le cD^3 \left\|\chi_{O_i} e^{it\Delta} f(x)\right\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)}. \tag{5.1}$$

Moreover, the polynomial P is a product of distinct non-singular polynomials.

We define

$$W := N_{R^{1/2+\delta}} Z(P) \cap \mathbf{B}^*_R , \tag{5.2}$$

where $\delta = \epsilon^2$ and $N_{R^{1/2+\delta}} Z(P)$ stands for the $R^{1/2+\delta}$-neighborhood of the variety $Z(P)$ in $\mathbb{R}^3$. We have the wave packet decomposition for $e^{it\Delta} f$ as in (4.7). For each cell O_i, we set

$$O_i' := [O_i \cap \mathbf{B}^*_R]\backslash W \text{ and } \mathbf{T}_i := \{(\theta,\nu) \in \mathbf{T} : T_{\theta,\nu} \cap O_i' \ne \emptyset\}. \tag{5.3}$$

Here $T_{\theta,\nu}$ is the tube associated to each tile (θ,ν), as defined in (4.8). For each function f we define

$$f_i := \sum_{(\theta,\nu)\in\mathbf{T}_i} f_{\theta,\nu} . \tag{5.4}$$

From (4.10), it follows that on each cell O_i',

$$e^{it\Delta}f(x) \sim e^{it\Delta}f_i(x)\,. \tag{5.5}$$

By the fundamental theorem of algebra, we have a simple yet important geometric observation:

Lemma 5.1. *For each tile $(\theta,\nu)\in\mathbf{T}$, the number of cells O_i' that intersect the tube $T_{\theta,\nu}$ is $\le D+1$.*

Proof. If $T_{\theta,\nu}$ intersects O_i', then the central line of $T_{\theta,\nu}$ must enter O_i. On the other hand, a line can cross the variety $Z(P)$ at most D times, hence can enter at most $D+1$ cells O_i. □

By triangle inequality, we dominate $\big\|e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)}$ by

$$\sum_i \big\|\chi_{O_i'}e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} + \big\|\chi_W e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)}\,. \tag{5.6}$$

We call the first term in (5.6) the cellular term, and the second the wall term. Using induction we will see that the desired bound (2.6) holds unless the wall term makes a significant contribution. In particular, we will show that (2.6) holds unless

$$\big\|e^{it\Delta}f\big\|_{L^p_xL^q_t(\mathbf{B}^*_R)} \lesssim R^{O(\epsilon^4)}\big\|\chi_W e^{it\Delta}f\big\|_{L^p_xL^q_t(\mathbf{B}^*_R)}. \tag{5.7}$$

Define

$$\mathcal{I} = \left\{ i\,:\, \big\|e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} \le 10cD^3\big\|\chi_{O_i'}e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} \right\}, \tag{5.8}$$

where c is the constant from (5.1). By triangle inequality and (5.1), for each $i\in\mathcal{I}^c$, we have

$$\begin{aligned}\big\|e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} &\le \frac{10}{9}cD^3\big\|\chi_{O_i\cap W}e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)}\\ &\lesssim R^{3\epsilon^4}\big\|\chi_W e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)}.\end{aligned}$$

So if $\mathcal{I}^c$ is non-empty, then (5.7) holds. For the moment, we are considering the case where (5.7) does not hold, and so every index i is in $\mathcal{I}$, and hence $|\mathcal{I}|\sim D^3$.

In addition, by Lemma 5.1,

$$\sum_i \|f_i\|_2^2 \lesssim (D+1)\sum_{\theta,\nu}\|f_{\theta,\nu}\|_2^2 \lesssim D\|f\|_2^2\,. \tag{5.9}$$

Henceforth, by pigeonhole principle, there exists $i\in\mathcal{I}$ such that

$$\|f_i\|_2^2 \lesssim D^{-2}\|f\|_2^2\,. \tag{5.10}$$

Now we use induction: we apply (2.6) to this special f_i at radius $\frac{R}{2}$. We can cover $B(0,R)\times[0,R]$ by $O(1)$ cylinders with dimensions $B(0,R/2)\times[0,R/2]$. Therefore, we get the bound

$$\begin{aligned}&\big\|e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} \lesssim D^3\big\|\chi_{O_i'}e^{it\Delta}f(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)} \lesssim D^3\big\|e^{it\Delta}f_i(x)\big\|^p_{L^p_xL^q_t(\mathbf{B}^*_R)}\\ \lesssim& D^3\left[C_{p,\epsilon}M^{-\epsilon^2}R^\epsilon\|f_i\|_2\right]^p \lesssim D^{3-p}\left[C_{p,\epsilon}M^{-\epsilon^2}R^\epsilon\|f\|_2\right]^p\,.\end{aligned}$$

Recall that $D = R^{\epsilon^4}$, and we can assume R is very large (compared to p). Since $p > 3$ we have $D^{3-p} \ll 1$. Therefore, we see that induction closes (unless (5.7) holds).

It remains to prove the desired bounds when (5.7) holds – when the wall term is almost as big as the whole.

6. Contribution from the wall: transverse and tangent terms

From Section 5, it remains to estimate the wall contribution, the second term in (5.6). To deal with the contribution from the wall W, we break B_R^* into $\sim R^{3\delta}$ balls B_j of radius $R^{1-\delta}$. (Recall from the last section that δ is defined to be ϵ^2.)

For any tile $(\theta, \nu) \in \mathbf{T}$, we say that $T_{\theta,\nu}$ is *tangent* to the wall W in a given ball B_j if it satisfies that $T_{\theta,\nu} \cap B_j \cap W \neq \emptyset$ and

$$\text{Angle}(G_0(\theta), T_z[Z(P)]) \leq R^{-1/2+2\delta} \tag{6.1}$$

for any non-singular point $z \in 10T_{\theta,\nu} \cap 2B_j \cap Z(P)$. Recall that $G_0(\theta) = (-2c(\theta), 1)$ is the direction of the tube $T_{\theta,\nu}$. Here $T_z[Z(P)]$ stands for the tangent space to the variety $Z(P)$ at the point z, and by a non-singular point we mean a point z in $Z(P)$ with $\nabla P(z) \neq 0$. Since P is a product of distinct non-singular polynomials, the non-singular points are dense in $Z(P)$. We note that if $T_{\theta,\nu}$ is tangent to W in B_j, then $T_{\theta,\nu} \cap B_j$ is contained in the $R^{1/2+\delta}$-neighborhood of $Z(P) \cap 2B_j$.

We say that $T_{\theta,\nu}$ is *transverse* to the wall W in the ball B_j if it enjoys that $T_{\theta,\nu} \cap B_j \cap W \neq \emptyset$ and

$$\text{Angle}(G_0(\theta), T_z[Z(P)]) > R^{-1/2+2\delta} \tag{6.2}$$

for some non-singular point $z \in 10T_{\theta,\nu} \cap 2B_j \cap Z(P)$.

Let $\mathbf{T}_{j,\text{tang}}$ represent the collection of all tiles $(\theta, \nu) \in \mathbf{T}$ such that $T_{\theta,\nu}$'s are tangent to the wall W in B_j, and $\mathbf{T}_{j,\text{trans}}$ denote the collection of all tiles $(\theta, \nu) \in \mathbf{T}$ such that $T_{\theta,\nu}$'s are transverse to the wall W in B_j.

We define $f_{j,\text{tang}} := \sum_{(\theta,\nu)\in\mathbf{T}_{j,\text{tang}}} f_{\theta,\nu}$ and $f_{j,\text{trans}} := \sum_{(\theta,\nu)\in\mathbf{T}_{j,\text{trans}}} f_{\theta,\nu}$. Then on $B_j \cap W$, we have

$$e^{it\Delta} f(x) \sim e^{it\Delta} f_{j,\text{tang}}(x) + e^{it\Delta} f_{j,\text{trans}}(x)\,. \tag{6.3}$$

The following Lemma is about how a tube crosses a variety transversely, which was proved by the second author in [9]. It says that $T_{\theta,\nu}$ crosses the wall W transversely in at most $R^{O(\epsilon^4)}$ many balls B_j.

Lemma 6.1. *(Lemma 3.5 in [9]) For each tile $(\theta, \nu) \in \mathbf{T}$, the number of $R^{1-\delta}$-balls B_j for which $(\theta, \nu) \in \mathbf{T}_{j,\text{trans}}$ is at most Poly$(D) = R^{O(\epsilon^4)}$.*

For points $(x, t) \in B_j \cap W$, we could break up $e^{it\Delta} f(x)$ into a transverse term and a tangent term. However, when we analyze the tangent contribution in subsequent sections, we will need to use a bilinear structure. So we do a more refined decomposition: we break $e^{it\Delta} f(x)$ into a linear transverse term and a bilinear tangent term.

We decompose $B(\xi_0, M^{-1}) \subset \mathbb{R}^2$, the Fourier support of function f, into balls τ of radius $1/(KM)$. Here $K = K(\epsilon)$ is a large parameter. We write $f = \sum_\tau f_\tau$, where supp $\widehat{f_\tau} \subseteq \tau$.

We let $B_\epsilon := \{(x,t) \in B(0,R)\times[0,R] \,:\, \exists\, \tau \text{ s.t. } |e^{it\Delta} f_\tau(x)| > K^{-\epsilon^4}|e^{it\Delta}f(x)|\}$. We will show by induction on the radius $(1/M)$ in frequency space that the contribution from B_ϵ is acceptable. In fact, by the definition of B_ϵ,

$$\left\|\chi_{B_\epsilon} e^{it\Delta} f(x)\right\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)} \le K^{\epsilon^4 p} \sum_\tau \left\|e^{it\Delta} f_\tau(x)\right\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)}.$$

By applying (2.6) in Theorem 2.1 the right-hand side is bounded by

$$\begin{aligned}
&\lesssim K^{\epsilon^4 p} \sum_\tau \left[C_\epsilon (KM)^{-\epsilon^2} R^\epsilon \|f_\tau\|_2\right]^p \\
&\le K^{(\epsilon^4-\epsilon^2)p} \left[C_\epsilon M^{-\epsilon^2} R^\epsilon \|f\|_2\right]^p
\end{aligned}$$

We choose $K = K(\epsilon)$ large so that $K^{(\epsilon^4-\epsilon^2)} \ll 1$, which yields by induction that the term involving B_ϵ plays an unimportant role.

For points (x,t) not in B_ϵ, we have the following decomposition into a transverse term and a bilinear tangent term.

Lemma 6.2. *For each point $(x,t) \in B_j \cap W$ satisfying $\max_\tau |e^{it\Delta} f_\tau(x)| \le K^{-\epsilon^4}|e^{it\Delta}f(x)|$, there exists a sub-collection I of the collection of all possible $1/(KM)$ balls τ, such that*

$$|e^{it\Delta} f(x)| \lesssim |e^{it\Delta} f_{I,j,\text{trans}}(x)| + K^{10}\text{Bil}(e^{it\Delta} f_{j,\text{tang}}(x)), \tag{6.4}$$

where

$$f_{I,j,\text{trans}}(x) := \sum_{\tau\in I} f_{\tau,j,\text{trans}}(x),$$

and the bilinear tangent term is given by

$$\text{Bil}(e^{it\Delta} f_{j,\text{tang}}(x)) := \max_{\substack{\tau_1,\tau_2 \\ dist(\tau_1,\tau_2)\ge 1/(KM)}} |e^{it\Delta} f_{\tau_1,j,\text{tang}}(x)|^{1/2} |e^{it\Delta} f_{\tau_2,j,\text{tang}}(x)|^{1/2}.$$

Proof. Let I be defined by $I := \{\tau \,:\, |e^{it\Delta} f_{\tau,j,\text{tang}}(x)| \le K^{-10}|e^{it\Delta}f(x)|\}$. Then clearly

$$I^c = \{\tau \,:\, |e^{it\Delta} f_{\tau,j,\text{tang}}(x)| > K^{-10}|e^{it\Delta}f(x)|\}.$$

If there exist $\tau_1, \tau_2 \in I^c$ with $\text{dist}(\tau_1,\tau_2) \ge 1/(KM)$, then $|e^{it\Delta}f(x)| \lesssim K^{10}\text{Bil}(e^{it\Delta} f_{j,\text{tang}}(x))$. Otherwise, the number of balls τ in I^c is $O(1)$, and

$$\sum_{\tau\in I^c} |e^{it\Delta} f_\tau(x)| \le CK^{-\epsilon^4}|e^{it\Delta}f(x)| \le \frac{1}{10}|e^{it\Delta}f(x)|.$$

Hence, by the fact that $f = \sum_\tau f_\tau$ and the definition of I,

$$\begin{aligned}
\frac{9}{10}|e^{it\Delta}f(x)| &\le |\sum_{\tau\in I} e^{it\Delta} f_\tau(x)| \\
&\lesssim |e^{it\Delta} f_{I,j,\text{tang}}(x)| + |e^{it\Delta} f_{I,j,\text{trans}}(x)| \\
&\le CK^{-8}|e^{it\Delta}f(x)| + |e^{it\Delta} f_{I,j,\text{trans}}(x)|,
\end{aligned}$$

which implies that $|e^{it\Delta}f(x)| \lesssim |e^{it\Delta} f_{I,j,\text{trans}}(x)|$. □

By Lemma 6.2 we can now estimate the wall contribution in (5.6) by

$$\begin{aligned}
&\big\|\chi_W\, e^{it\Delta} f(x)\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)} \\
\lesssim &\big\|\chi_{B_\epsilon} e^{it\Delta} f(x)\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)} \qquad (6.5)\\
&+\sum_j \big\| \max_I \chi_{B_j\cap W}\, |e^{it\Delta} f_{I,j,\mathrm{trans}}(x)|\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)} \qquad (6.6)\\
&+K^{10p}\sum_j \big\|\chi_{B_j\cap W}\mathrm{Bil}(e^{it\Delta} f_{j,\mathrm{tang}}(x))\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)}\,. \qquad (6.7)
\end{aligned}$$

As we explained above, the first term (6.5) obeys an acceptable bound by induction on M. We now estimate the linear transverse term (6.6). The term (6.6) is dominated by

$$\sum_j \sum_{I\subseteq\mathcal{T}} \big\|\chi_{B_j\cap W}\, e^{it\Delta} f_{I,j,\mathrm{trans}}(x)\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)}\,, \qquad (6.8)$$

where $\mathcal{T}$ is the collection of all possible $1/(KM)$-balls in $B(\xi_0, 1/M)$, and the sum is taken over all subsets of $\mathcal{T}$. Since there are at most 2^{K^2} I's, we apply (2.6) in Theorem 2.1 with radius $R^{1-\delta}$ to obtain

$$(6.8) \le \sum_j 2^{K^2}\left[C_\epsilon M^{-\epsilon^2} R^{(1-\delta)\epsilon}\|f_{j,\mathrm{trans}}\|_2\right]^p\,, \qquad (6.9)$$

which is bounded by, using Lemma 6.1,

$$2^{K^2} R^{O(\epsilon^4)-\delta\epsilon p}\left[C_\epsilon M^{-\epsilon^2} R^{\epsilon}\|f\|_2\right]^p\,. \qquad (6.10)$$

Since $\delta=\epsilon^2$, it is clear that $2^{K^2}R^{O(\epsilon^4)-\delta\epsilon p} < 1/100$ and so the induction on the transverse term closes.

It remains to estimate the bilinear tangent term (6.7). We state the result on the bilinear maximal estimate in this section, and prove it in Section 8.

Proposition 6.3. *For $p>3$, the following maximal estimate of the bilinear tangent term holds, uniformly in M:*

$$\left(\int_{B(0,R)} \sup_{t:(x,t)\in W\cap B_j} \big|\mathrm{Bil}(e^{it\Delta} f_{j,\mathrm{tang}}(x))\big|^p dx\right)^{1/p} \le\ C_\epsilon R^{\epsilon/2}\|f\|_2\,. \qquad (6.11)$$

Given Proposition 6.3, we estimate the bilinear tangent term (6.7) as follows, for any $q>1/\epsilon^4$,

$$\begin{aligned}
&\big\|\chi_{B_j\cap W}\mathrm{Bil}(e^{it\Delta} f_{j,\mathrm{tang}}(x))\big\|^p_{L^p_x L^q_t(\mathbf{B}^*_R)}\\
\le &R^{p/q}\int_{B(0,R)} \sup_{t:(x,t)\in W\cap B_j} \big|\mathrm{Bil}(e^{it\Delta} f_{j,\mathrm{tang}}(x))\big|^p dx\\
\le &R^{O(\delta)+\epsilon p/2}\|f\|_2^p\,.
\end{aligned}$$

Hence Theorem 2.1 follows from Proposition 6.3 and the inductions.

7. Variations on the Strichartz inequality using decoupling

In this section we obtain both linear and bilinear local refinements of the Strichartz inequality, via the Bourgain-Demeter l^2-decoupling theorem [4]. In Section 8 we will use the bilinear refinement to prove the bilinear maximal estimate in Proposition 6.3.

For the bilinear tangent term in Proposition 6.3, all wave packets are tangent to a variety. Suppose that $Z = Z(P)$ where P is a product of non-singular polynomials. For any tile $(\theta, \nu) \in \mathbf{T}$, we say that $T_{\theta,\nu}$ is $ER^{-1/2}$-*tangent* to Z if

$$T_{\theta,\nu} \subset N_{ER^{1/2}} Z \cap \mathbf{B}_R^*, \text{and}$$

$$\text{Angle}(G_0(\theta), T_z[Z(P)]) \leq ER^{-1/2} \tag{7.1}$$

for any non-singular point $z \in N_{2ER^{1/2}}(T_{\theta,\nu}) \cap 2\mathbf{B}_R^* \cap Z$.

Let

$$\mathbf{T}_Z(E) := \{(\theta, \nu) \,|\, T_{\theta,\nu} \text{ is } ER^{-1/2}\text{-tangent to } Z\},$$

and we say that f is concentrated in wave packets from $\mathbf{T}_Z(E)$ if

$$\sum_{(\theta,\nu) \notin \mathbf{T}_Z(E)} \|f_{\theta,\nu}\|_2 \leq \text{RapDec}(R)\|f\|_2.$$

Since the radius of $T_{\theta,\nu}$ is $R^{1/2+\delta}$, R^δ is the smallest interesting value of E.

In this section, we establish the following local refinements of the Strichartz estimates.

Theorem 7.1. *Suppose that f has Fourier support in $B^2(0,1)$, and is concentrated in wave packets from $\mathbf{T}_Z(E)$, where $Z = Z(P)$ and P is a product of distinct non-singular polynomials. Suppose that $Q_1, Q_2, ...$ are lattice $R^{1/2}$-cubes in $B^3(R)$, so that*

$$\|e^{it\Delta} f\|_{L^6(Q_j)} \text{ is essentially constant in } j.$$

Suppose that these cubes are arranged in horizontal strips of the form $\mathbb{R} \times \mathbb{R} \times \{t_0, t_0 + R^{1/2}\}$, and that each strip contains $\sim \sigma$ cubes Q_j. Let Y denote $\bigcup_j Q_j$. Then

$$\|e^{it\Delta} f\|_{L^6(Y)} \lessapprox E^{O(1)} R^{-1/6} \sigma^{-1/3} \|f\|_{L^2}. \tag{7.2}$$

To get some intuition, we consider a special case of Theorem 7.1, in which the variety Z is naturally replaced by a 2-plane V, and $E \approx 1$. In the planar case, all wave packets are contained in the $\approx R^{1/2}$-neighborhood of V, and the absolute value $|e^{it\Delta} f(x)|$ is essentially constant along a certain direction which is roughly normal to V. Note that $e^{it\Delta} f(x)|_V$ is a Schrödinger solution in dimension 2. Denote $e^{it\Delta} f(x)|_V$ by $e^{ir\Delta} h(y)$ for some function h with Fourier support in $B^1(1)$, where (y, r) are coordinates of V. Hence the conclusion in Theorem 7.1 can be rephrased in terms of h. Indeed, observe that

$$\|e^{it\Delta} f(x)\|_{L^6(Y)}^6 \sim R^{1/2} \|e^{ir\Delta} h(y)\|_{L^6(Y \cap V)}^6,$$

$$\|f\|_2^2 \sim R^{-1} \|e^{it\Delta} f\|_{L^2(B^3(R))}^2 \sim R^{-1} R^{1/2} \|e^{ir\Delta} h\|_{L^2(B_R^3 \cap V)}^2 \sim R^{1/2} \|h\|_2^2.$$

Therefore the estimate (7.2) is equivalent to

$$\|e^{ir\Delta} h\|_{L^6(Y \cap V)} \lessapprox \sigma^{-1/3} \|h\|_{L^2}. \tag{7.3}$$

It follows from the Strichartz inequality that $\|e^{ir\Delta}h\|_{L^6(Y\cap V)} \lesssim \|h\|_{L^2}$. We get an improvement when σ is large. The condition that σ is large forces the solution $e^{it\Delta}f$ to be spread out in space, and we will exploit this spreading out to get our improvement.

Moreover, Theorem 7.1 has the following bilinear refinement.

Theorem 7.2. *For functions f_1 and f_2 with separated Fourier supports in $B^2(0,1)$, separated by ~ 1, suppose that f_1 and f_2 are concentrated in wave packets from $\mathbf{T}_Z(E)$, where $Z = Z(P)$ and P is a product of distinct non-singular polynomials. Suppose that $Q_1, Q_2, \cdots, Q_N$ are lattice $R^{1/2}$-cubes in $B^3(R)$, so that for each i,*

$$\|e^{it\Delta}f_i\|_{L^6(Q_j)} \text{ is essentially constant in } j.$$

Let Y denote $\bigcup_{j=1}^N Q_j$. Then

$$\left\| |e^{it\Delta}f_1 e^{it\Delta}f_2|^{1/2} \right\|_{L^6(Y)} \lessapprox E^{O(1)} R^{-1/6} N^{-1/6} \|f_1\|_{L^2}^{1/2} \|f_2\|_{L^2}^{1/2}.$$

7.1. **Proof of Theorem 7.1.** The proof uses the Bourgain-Demeter l^2-decoupling theorem, together with induction on the radius and parabolic rescaling. First we recall the decoupling result of Bourgain and Demeter in [4].

Theorem 7.3 (Bourgain-Demeter)**.** *Suppose that the R^{-1}-neighborhood of the unit parabola in $\mathbb{R}^2$ is divided into $R^{1/2}$ disjoint rectangular boxes τ, each with dimensions $R^{-1/2} \times R^{-1}$. Suppose $\widehat{F}_\tau$ is supported in τ and $F = \sum_\tau F_\tau$. Then*

$$\|F\|_{L^6(\mathbb{R}^2)} \lessapprox \left(\sum_\tau \|F_\tau\|_{L^6(\mathbb{R}^2)}^2 \right)^{1/2}.$$

If $E \geq R^{1/4}$ (or any fixed power of R), then the estimate (7.2) is trivial because of the factor $E^{O(1)}$. So we assume that $E \leq R^{1/4}$.

To set up the argument, we decompose f as follows. We break the unit ball $B^2(1)$ in frequency space into small balls τ of radius $R^{-1/4}$, and divide the physical space ball $B^2(R)$ into balls B of radius $R^{3/4}$. For each pair (τ, B), we let $f_{\square_{\tau,B}}$ be the function formed by cutting off f on the ball B (with a Schwartz tail) in physical space and the ball τ in Fourier space. We note that $e^{it\Delta}f_{\square_{\tau,B}}$, restricted to $B^3(R)$, is essentially supported on an $R^{3/4} \times R^{3/4} \times R$-box, which we denote by $\square_{\tau,B}$ (compare the discussion in Section 4). The box $\square_{\tau,B}$ is in the direction given by $(-2c(\tau), 1)$ and intersects $t = 0$ at a disk centered at $(c(B), 0)$, where $c(\tau)$ and $c(B)$ are the centers of τ and B respectively. For a fixed τ, the different boxes $\square_{\tau,B}$ tile $B^3(R)$. In particular, for each τ, a given cube Q_j lies in exactly one box $\square_{\tau,B}$.

Since f is concentrated in wave packets from $\mathbf{T}_Z(E)$, we only need to consider those $R^{1/2}$-cubes Q_j that are contained in the $ER^{1/2}$-neighborhood of Z. For each such $R^{1/2}$-cube Q_j, we will see that the wave packets that pass through Q_j are nearly coplanar. Because of this, we will be able to apply the 2-dimensional decoupling theorem to study $e^{it\Delta}f$ on Q_j:

Lemma 7.4. *Suppose that f has Fourier support in $B^2(0,1)$ and is concentrated in wave packets from $\mathbf{T}_Z(E)$, where $E \leq R^{1/4}$ and $Z = Z(P)$ is a finite union of non-singular varieties. Suppose that an $R^{1/2}$-cube Q is in $N_{ER^{1/2}}(Z)$. Then we have the decoupling bound*

$$\|e^{it\Delta}f\|_{L^6(Q)} \lessapprox \left(\sum_{\square} \|e^{it\Delta}f_{\square}\|^2_{L^6(10Q)}\right)^{1/2} + R^{-1000}\|f\|_{L^2}. \tag{7.4}$$

Remark: The $R^{-1000}\|f\|_{L^2}$ is a negligibly small term which covers minor contributions coming from the tails of the Fourier transforms of smooth functions. We will neglect this term in the sequel.

Proof. Observe that $Q \subset N_{ER^{1/2}}Z$ implies that there exists a non-singular point $z_0 \in Z \cap N_{ER^{1/2}}Q$. Thus for each wave packet $T_{\theta,\nu}$ that intersects Q, we have $z_0 \in Z \cap N_{2ER^{1/2}}(T_{\theta,\nu})$. By the definition of $\mathbf{T}_Z(E)$ we get the angle bound

$$\text{Angle}(G_0(\theta), T_{z_0}[Z(P)]) \leq ER^{-1/2}\,. \tag{7.5}$$

We recall from Section 4 that $G_0(\theta) = (-2c(\theta), 1)$. Suppose that $T_{z_0}Z$ is the plane given by $a_1x_1 + a_2x_2 + bt = 0$, with $a_1^2 + a_2^2 + b^2 = 1$. The angle condition above restricts the location of θ as follows:

$$|-2a\cdot c(\theta) + b| \lesssim ER^{-1/2}. \tag{7.6}$$

We note that each tube $T_{\theta,\nu}$ makes an angle $\gtrsim 1$ with the plane $t = 0$, because $\theta \subset B(0,1)$. We can assume that there are some tubes $T_{\theta,\nu}$ tangent to $T_{z_0}Z$, and so $|a| \gtrsim 1$. Therefore, (7.6) confines θ to a strip of width $\sim ER^{-1/2}$ inside of $B(0,1)$. We denote this strip by $S \subset B(0,1)$.

Let $\mathbf{T}_{Z,Q}(E)$ be the set of (θ,ν) in $\mathbf{T}_Z(E)$ for which each $T_{\theta,\nu}$ intersects Q. For each (θ,ν) in $\mathbf{T}_{Z,Q}(E)$, θ obeys (7.6), and so $\theta \subset S$. Let η be a smooth bump function which approximates χ_Q. We note that $\eta e^{it\Delta}f$ is essentially equal to

$$\sum_{(\theta,\nu)\in\mathbf{T}_{Z,Q}(E)} \eta e^{it\Delta}f_{\theta,\nu}.$$

Therefore, the Fourier transform of the localized solution $\eta e^{it\Delta}f$ is essentially supported in

$$S^* := \{(\xi_1,\xi_2,\xi_3) : (\xi_1,\xi_2) \in S \text{ and } |\xi_3 - \xi_1^2 - \xi_2^2| \lesssim R^{-1/2}\}. \tag{7.7}$$

(The contribution of the not essential parts is covered by the negligible term $R^{-1000}\|f\|_{L^2}$ in the statement of the Lemma.)

After a rotation in the (x_1,x_2)-plane we can suppose that the strip S is defined by

$$a_1 \leq \xi_1 \leq a_1 + ER^{-1/2},$$

for some $a_1 \in [-1,1]$. We note that at each point $(\xi_1,\xi_2) \in S$,

$$\partial_1\left(\xi_1^2 + \xi_2^2\right) = 2a_1 + O(ER^{-1/2}). \tag{7.8}$$

Let v be the vector

$$v = (1, 0, 2a_1).$$

Let Π be a 2-plane perpendicular to v. Because $E \leq R^{1/4}$, we claim that the projection of S^* onto Π lies in the $\sim R^{-1/2}$-neighborhood of a parabola. We can see this as follows. Let

$$S^*_{core} := \{(\xi_1, \xi_2, \xi_3) : \xi_1 = a_1, |\xi_2| \leq 1, \xi_3 = \xi_1^2 + \xi_2^2\}.$$

The set S^*_{core} is a parabola, and its projection onto Π is also a parabola. We claim that the projection of S^* to Π lies in the $\sim R^{-1/2}$-neighborhood of this parabola. If $(\xi_1, \xi_2, \xi_3) \in S^*$, then (7.8) tells us that

$$(\xi_1^2 + \xi_2^2) = a_1^2 + \xi_2^2 + 2a_1(\xi_1 - a_1) + O(ER^{-1/2} \cdot |\xi_1 - a_1|).$$

Therefore,

$$(\xi_1, \xi_2, \xi_3) = (a_1, \xi_2, a_1^2 + \xi_2^2) + (\xi_1 - a_1)v + O(ER^{-1/2}|\xi_1 - a_1| + R^{-1/2}).$$

The first term on the right-hand side lies is S^*_{core}. Since Π is perpendicular to v, the projection to Π kills the second term on the right-hand side. So the distance from the projection of ξ to the projection of S^*_{core} is at most

$$ER^{-1/2}|\xi_1 - a_1| + R^{-1/2} \lesssim E^2R^{-1} + R^{-1/2} \sim R^{-1/2}.$$

Therefore, if we restrict $\eta e^{it\Delta} f$ to Π, the resulting 2-dimensional function has Fourier support in the $\sim R^{-1/2}$-neighborhood of a parabola.

We consider the decomposition $f = \sum_{(\tau,B):\square_{\tau,B}\cap Q\neq\emptyset} f_{\square_{\tau,B}}$. If $e^{it\Delta} f_{\square_{\tau,B}}$ contributes to $\|e^{it\Delta} f\|_{L^6(Q)}$, there must be a wave packet $T_{\theta,\nu}$ that intersects the $R^{1/2}$-cube Q with $\theta \subset \tau$, and so $\tau \cap S$ must be non-empty. Also, for a given τ, there is only one B so that $\square_{\tau,B} \cap Q$ is non-empty. Also, the Fourier support of $\eta e^{it\Delta} f_{\square_{\tau,B}}$ lies in $S^* \cap (\tau \times \mathbb{R})$, by the same argument we used above for $\eta e^{it\Delta} f$. The projection onto Π of $S^* \cap (\tau \times \mathbb{R})$ is an $R^{-1/4} \times R^{-1/2}$ rectangular box. The union of these boxes over all τ intersecting S is the $R^{-1/2}$-neighborhood of a parabola. Therefore, we have the hypotheses to apply the 2-dimensional decoupling theorem, Theorem 7.3, which gives:

$$\|\eta e^{it\Delta} f\|_{L^6(\Pi)} \lessapprox \left(\sum_{\square} \|\eta e^{it\Delta} f_{\square}\|^2_{L^6(\Pi)} \right)^{1/2}.$$

Now we integrate in the direction perpendicular to Π and apply Fubini and Minkowski to get

$$\|\eta e^{it\Delta} f\|_{L^6(\mathbb{R}^3)} \lessapprox \left(\sum_{\square} \|\eta e^{it\Delta} f_{\square}\|^2_{L^6(\mathbb{R}^3)} \right)^{1/2}.$$

This implies the desired conclusion. □

Next, by induction on the radius R, we will show that each function $f_\square$ obeys a version of Theorem 7.1. Here is the statement. Suppose that $S_1, S_2, ...$ are $R^{1/2} \times R^{1/2} \times R^{3/4}$-tubes in $\square$ (running parallel to the long axis of $\square$), and that

$$\|e^{it\Delta} f_{\square}\|_{L^6(S_j)} \text{ is essentially constant in } j.$$

Suppose that these tubes are arranged into $R^{3/4}$-strips running parallel to the short axes of $\square$ and that each such strip contains $\sim \sigma_\square$ tubes S_j. Let $Y_\square$ denote $\cup_j S_j$. Then

$$\|e^{it\Delta} f_\square\|_{L^6(Y_\square)} \lessapprox E^{O(1)} R^{-1/12} R^{-1/12} \sigma_\square^{-1/3} \|f_\square\|_{L^2}. \tag{7.9}$$

This inequality follows by doing a parabolic rescaling and then using Theorem 7.1 at scale $R^{1/2}$, which we can assume holds by induction on R. We write down the details of this parabolic rescaling, and in particular we will check that the tangent-to-variety condition is preserved under parabolic rescaling. For each $R^{-1/4}$-ball τ in $B^2(1)$, we write $\xi = \xi_0 + R^{-1/4}\zeta \in \tau$, then

$$|e^{it\Delta} f_\tau(x)| = R^{-1/4} |e^{i\tilde{t}\Delta} g(\tilde{x})|$$

for some function g with Fourier support in $B^2(1)$ and $\|g\|_2 = \|f_\tau\|_2$, where the new coordinates $(\tilde{x}, \tilde{t})$ are related to the old coordinates (x, t) by

$$\begin{cases} \tilde{x} = R^{-1/4} x + 2tR^{-1/4}\xi_0 \,, \\ \tilde{t} = R^{-1/2} t \,. \end{cases} \tag{7.10}$$

Therefore

$$\|e^{it\Delta} f_\square(x)\|_{L^6(Y_\square)} = R^{-1/12} \|e^{i\tilde{t}\Delta} g(\tilde{x})\|_{L^6(\tilde{Y})},$$

where $\tilde{Y}$ is the image of $Y_\square$ under the new coordinates. Note that $\tilde{Y}$ is a union of $R^{1/4}$-cubes inside an $R^{1/2}$-cube. These $R^{1/4}$-cubes are arranged in $R^{1/4}$-horizontal strips, and each strip contains $\sim \sigma_\square$ $R^{1/4}$-cubes. Moreover, by the relation (7.10), we see that each wave packet T, at scale R, of dimensions $R^{1/2+\delta} \times R^{1/2+\delta} \times R$ in the old coordinates is mapped to a corresponding wave packet $\tilde{T}$, at scale $R^{1/2}$, of dimensions $R^{1/4+\delta} \times R^{1/4+\delta} \times R^{1/2}$ in the new coordinates. The variety $Z(P)$ corresponds to a new variety $Z(Q)$, given by the relation $Q(\tilde{x}, \tilde{t}) = Q(R^{-1/4}x + 2tR^{-1/4}\xi_0, R^{-1/2}t) = P(x, t)$. We claim that, under the above correspondence, if the wave packet T at scale R is $ER^{-1/2}$-tangent to $Z(P)$, then the wave packet $\tilde{T}$ at scale $R^{1/2}$ is $ER^{-1/4}$-tangent to $Z(Q)$ in the new coordinates.

By the relation (7.10), the distance condition $T \subset N_{ER^{1/2}} Z(P)$ implies that $\tilde{T} \subset N_{ER^{1/4}} Z(Q)$. Given the direction $(-2\xi, 1)$ of T, the angle condition

$$\text{Angle}((-2\xi, 1), T_{z_0}[Z(P)]) \le ER^{-1/2}$$

is equivalent to

$$\frac{|(-2\xi, 1) \cdot (P_x(x_0, t_0), P_t(x_0, t_0))|}{|(P_x(x_0, t_0), P_t(x_0, t_0))|} \lesssim ER^{-1/2}, \tag{7.11}$$

where $z_0 = (x_0, t_0)$. Note that the direction of the corresponding wave packet $\tilde{T}$ is given by $(-2\zeta, 1)$, where ξ and ζ are related by $\xi = \xi_0 + R^{-1/4}\zeta$. Let $\tilde{z}_0 = (\tilde{x}_0, \tilde{t}_0)$ denote the point corresponding to z_0. Using the relations

$$P_x = R^{-1/4} Q_{\tilde{x}}, \quad P_t = 2R^{-1/4}\xi_0 \cdot Q_{\tilde{x}} + R^{-1/2} Q_{\tilde{t}} \,,$$

after some computation, (7.11) yields that

$$\frac{|(-2\zeta, 1) \cdot (Q_{\tilde{x}}(\tilde{x}_0, \tilde{t}_0), Q_{\tilde{t}}(\tilde{x}_0, \tilde{t}_0))|}{|(Q_{\tilde{x}}(\tilde{x}_0, \tilde{t}_0), Q_{\tilde{t}}(\tilde{x}_0, \tilde{t}_0))|} \lesssim ER^{-1/4},$$

which implies that

$$\text{Angle}((-2\zeta, 1), \tilde{T}_{\tilde{z}_0}[Z(Q)]) \le ER^{-1/4}.$$

Therefore the tangent-to-variety condition is preserved under parabolic rescaling and the induction on radius is justified.

We have now established inequality (7.9). To apply this inequality, we need to identify a good choice of $Y_\square$. We do this by some dyadic pigeonholing. For each $\square$, we apply the following algorithm to regroup tubes in $\square$.

(1) We sort those $R^{1/2}\times R^{1/2}\times R^{3/4}$-tubes S's contained in the box $\square$ according to the order of magnitude of $\|e^{it\Delta}f_\square\|_{L^6(S)}$, which we denote λ. For each dyadic number λ, we use $\mathbb{S}_\lambda$ to stand for the collection of tubes $S\subset\square$ with $\|e^{it\Delta}f_\square\|_{L^6(S)}\sim\lambda$.

(2) For each λ, we sort the tubes $S\in\mathbb{S}_\lambda$ by looking at the number of such tubes in an $R^{3/4}$-strip. For any dyadic number η, we let $\mathbb{S}_{\lambda,\eta}$ be the set of tubes $S\in\mathbb{S}_\lambda$ so that the number of tubes of $\mathbb{S}_\lambda$ in the $R^{3/4}$-strip containing S is $\sim\eta$.

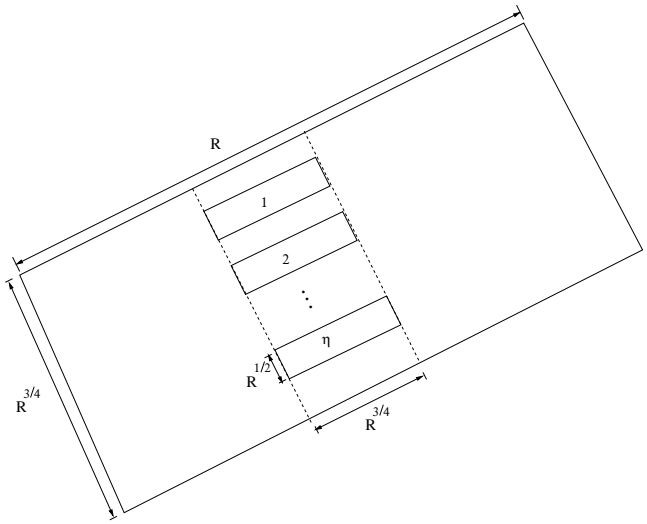

FIGURE 2. Tubes in a given strip in the $\square$

Let $Y_{\square,\lambda,\eta}$ be the union of the tubes in $\mathbb{S}_{\lambda,\eta}$. Then we represent

$$e^{it\Delta}f=\sum_{\lambda,\eta}\left(\sum_{\square}e^{it\Delta}f_\square\cdot\chi_{Y_{\square,\lambda,\eta}}\right).$$

Note that $\|e^{it\Delta}f_\square\|_{L^6(S)}\le R^{O(1)}\|f\|_2$, for each tube S as above and the number of $\square$'s does not exceed $R^{O(1)}$. We see that the contribution from those λ's with $\lambda\le R^{-C}\|f\|_2$ is at most $R^{-C/2}\|f\|_2$. Here the constant C can be selected to be sufficiently large so that $R^{-C/2}\|f\|_2$ is negligible. So without loss of generality, we can assume that the terms with small λ contribute insignificantly to $\|e^{it\Delta}f\|_{L^6(Q_j)}$ for every Q_j. Therefore there are only $O(\log R)$ significant choices for each of λ,η. By pigeonholing, we can choose λ,η so that

$$\|e^{it\Delta}f\|_{L^6(Q_j)}\lesssim(\log R)^2\big\|\sum_{\square}e^{it\Delta}f_\square\cdot\chi_{Y_{\square,\lambda,\eta}}\big\|_{L^6(Q_j)}\tag{7.12}$$

holds for a fraction ≈ 1 of all cubes Q_j in Y. We need this uniform choice of (λ, η), which is independent of Q_j, because later we will sum over all Q_j and arrive at $\|e^{it\Delta} f_\square\|_{L^6(Y_{\square,\lambda,\eta})}$.

We fix λ and η for the rest of the proof. Let $Y_\square$ stand for the abbreviation of $Y_{\square,\lambda,\eta}$. We note that $Y_\square$ obeys the hypotheses for our inductive estimate (7.9), with $\sigma_\square$ being the value of η that we have fixed.

The following geometric estimate will play a crucial role in our proof. Each set $Y_\square$ contains $\lesssim \sigma_\square$ tubes in each strip parallel to the short axes of $\square$. Since the angle between the short axes of $\square$ and the x-axes is bounded away from $\pi/2$, it follows that $Y_\square$ contains $\lesssim \sigma_\square$ cubes Q_j in any $R^{1/2}$-horizontal row. Therefore,

$$|Y_\square \cap Y| \lesssim \frac{\sigma_\square}{\sigma} |Y|. \tag{7.13}$$

Next we sort the the boxes $\square$ according to the dyadic size of $\|f_\square\|_{L^2}$. We can restrict matters to $\lesssim \log R$ choices of this dyadic size, and so we can choose a set of $\square$'s, $\mathbb{B}$, so that $\|f_\square\|_{L^2}$ is essentially constant for $\square \in \mathbb{B}$ and

$$\|e^{it\Delta} f\|_{L^6(Q_j)} \lessapprox \| \sum_{\square \in \mathbb{B}} e^{it\Delta} f_\square \cdot \chi_{Y_\square} \|_{L^6(Q_j)} \tag{7.14}$$

for a fraction ≈ 1 of cubes Q_j in Y.

Finally we sort the cubes $Q_j \subset Y$ according to the number of $Y_\square$ that contain them. We let $Y' \subset Y$ be a set of cubes Q_j which obey (7.14) and which each lies in $\sim \mu$ of the sets $\{Y_\square\}_{\square \in \mathbb{B}}$. Because (7.14) holds for a large fraction of cubes, and because there are only dyadically many choices of μ, $|Y'| \approx |Y|$. By the equation (7.13), we see that

$$|Y_\square \cap Y'| \le |Y_\square \cap Y| \lessapprox \frac{\sigma_\square}{\sigma} |Y| \approx \frac{\sigma_\square}{\sigma} |Y'|.$$

Therefore, the multiplicity μ is bounded by

$$\mu \lessapprox \frac{\sigma_\square}{\sigma} |\mathbb{B}|. \tag{7.15}$$

We now are ready to combine all our ingredients and finish our proof. For each $Q_j \subset Y'$, we have

$$\|e^{it\Delta} f\|_{L^6(Q_j)} \lessapprox \left\| \sum_{\square \in \mathbb{B}} e^{it\Delta} f_\square \cdot \chi_{Y_\square} \right\|_{L^6(Q_j)}.$$

Now we apply Lemma 7.4 to the function $\sum_{\square \in \mathbb{B}, Q_j \in Y_\square} f_\square$ to bound the right hand side by

$$\lessapprox \left(\sum_{\square \in \mathbb{B}, Q_j \subset Y_\square} \left\| e^{it\Delta} f_\square \right\|^2_{L^6(Q_j)} \right)^{1/2}.$$

Since the number of $Y_\square$ containing Q_j is $\sim \mu$, we can apply Hölder to get

$$\left\| \sum_{\square \in \mathbb{B}} e^{it\Delta} f_\square \cdot \chi_{Y_\square} \right\|_{L^6(Q_j)} \lessapprox \mu^{1/3} \left(\sum_{\square \in \mathbb{B}, Q_j \subset Y_\square} \left\| e^{it\Delta} f_\square \right\|^6_{L^6(Q_j)} \right)^{1/6}.$$

Now we raise to the sixth power and sum over $Q_j \subset Y'$ to get

$$\left\|e^{it\Delta}f\right\|^6_{L^6(Y')} \lessapprox \mu^2 \sum_{\square\in\mathbb{B}} \left\|e^{it\Delta}f_\square\right\|^6_{L^6(Y_\square)}.$$

Since $|Y'| \gtrapprox |Y|$, and since each cube $Q_j \subset Y$ makes an equal contribution to $\|e^{it\Delta}f\|_{L^6(Y)}$, we see that $\|e^{it\Delta}f\|_{L^6(Y)} \approx \|e^{it\Delta}f\|_{L^6(Y')}$ and so

$$\left\|e^{it\Delta}f\right\|^6_{L^6(Y)} \lessapprox \mu^2 \sum_{\square\in\mathbb{B}} \left\|e^{it\Delta}f_\square\right\|^6_{L^6(Y_\square)}.$$

By a parabolic rescaling, Figure 2 becomes Figure 3. Henceforth, applying our

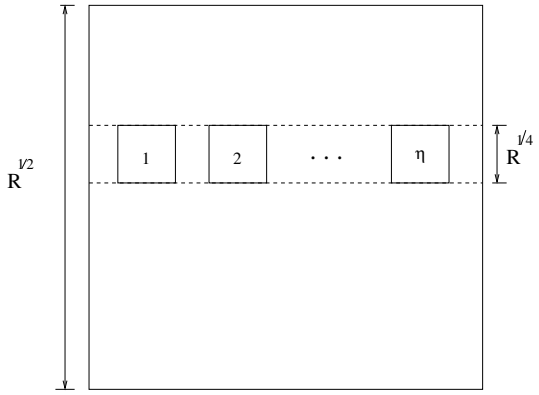

FIGURE 3. Cubes in a given strip in an $R^{1/2}$-cube

inductive hypothesis (7.9) at scale $R^{1/2}$ to the right-hand side, we see that

$$\left\|e^{it\Delta}f\right\|^6_{L^6(Y)} \lessapprox E^{O(1)}R^{-1}\mu^2\sigma_\square^{-2}\sum_{\square\in\mathbb{B}}\|f_\square\|^6_{L^2}. \tag{7.16}$$

Plugging in our bound for μ in (7.15), this is bounded by

$$\lesssim E^{O(1)}R^{-1}\sigma^{-2}|\mathbb{B}|^2\sum_{\square\in\mathbb{B}}\|f_\square\|^6_{L^2}.$$

Now since $\|f_\square\|_{L^2}$ is essentially constant among all $\square \in \mathbb{B}$, the last expression is

$$\sim E^{O(1)}R^{-1}\sigma^{-2}(\sum_{\square\in\mathbb{B}}\|f_\square\|^2_{L^2})^3 \le E^{O(1)}R^{-1}\sigma^{-2}\|f\|^6_{L^2}.$$

Taking the sixth root, we obtain our desired bound:

$$\|e^{it\Delta}f\|_{L^6(Y)} \lessapprox E^{O(1)}R^{-1/6}\sigma^{-1/3}\|f\|_{L^2}.$$

This closes the induction on radius and completes the proof.

7.2. **Proof of Theorem 7.2.** It can be proved by the method used in the proof of Theorem 7.1. By Hölder,

$$\left\|\left|e^{it\Delta}f_1e^{it\Delta}f_2\right|^{1/2}\right\|_{L^6(Y)} \le \prod_{i=1}^2\left\|e^{it\Delta}f_i\right\|^{1/2}_{L^6(Y)}.$$

For each i, we process $\|e^{it\Delta}f_i\|_{L^6(Y)}$ following the proof of Theorem 7.1. We decompose $f_i = \sum_\square f_{i,\square}$, and we follow the proof of Theorem 7.1. We define $Y_{i,\square}$ by

dyadic pigeonholing, so that $Y_{i,\square}$ is arranged in several $R^{3/4}$-strips (running parallel to the short axes of $\square$) with $\sim \sigma_{i,\square}$ $R^{1/2} \times R^{1/2} \times R^{3/4}$-tubes in each strip. When we use dyadic pigeonholing to pick a subset of cubes $Q_j \subset Y$, we pigeonhole for f_1 and f_2 simultaneously, and so we pick out a set of cubes that works well for both functions. Following the argument up to Equation (7.14), we see that for a fraction ≈ 1 of cubes Q_j,

$$\|e^{it\Delta} f_i\|_{L^6(Q_j)} \lessapprox \| \sum_{\square \in \mathbb{B}_i} e^{it\Delta} f_{i,\square} \cdot \chi_{Y_{i,\square}} \|_{L^6(Q_j)} \text{ for } i = 1, 2. \tag{7.17}$$

Similarly, we sort the cubes $Q_j \subset Y$ according to the number of $Y_{i,\square}$ that contain them. We let $Y' \subset Y$ be a set of cubes Q_j which obey (7.17) and which each lies in $\sim \mu_1$ of the sets $\{Y_{1,\square}\}_{\square \in \mathbb{B}_1}$ and $\sim \mu_2$ of the sets $\{Y_{2,\square}\}_{\square \in \mathbb{B}_2}$. Because (7.14) holds for a large fraction of cubes, and because there are only dyadically many choices of μ_1, μ_2, $|Y'| \approx |Y|$. Following the proof of Theorem 7.1 further, up to Equation (7.16), we see that for each i,

$$\left\| e^{it\Delta} f_i \right\|_{L^6(Y)} \lessapprox E^{O(1)} R^{-1/6} \left[\mu_i^2 \sigma_{i,\square}^{-2} \sum_{\square \in \mathbb{B}_i} \|f_{i,\square}\|_{L^2}^6 \right]^{1/6}. \tag{7.18}$$

Finally, we give a geometric estimate for μ_1 and μ_2 that takes advantage of the bilinear structure. If $\square_1 \in \mathbb{B}_1$ and $\square_2 \in \mathbb{B}_2$, then the angle between their long axes is ~ 1. Therefore, their intersection is contained in a ball of radius $\sim R^{3/4}$, and so $Y_{\square_1} \cap Y_{\square_2}$ contains $\lesssim \sigma_{1,\square}\sigma_{2,\square}$ different $R^{1/2}$-balls (see Figure 4). For each of the

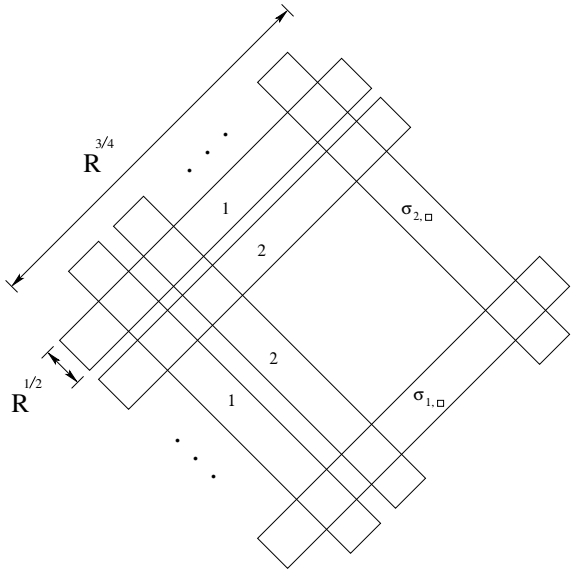

FIGURE 4. at most $O(\sigma_{1,\square}\sigma_{2,\square})$ cubes created by two transversal families of rectangular boxes

$\approx N$ cubes Q_j in Y', for each i, the cube Q_j lies in $\sim \mu_i$ of the sets $\{Y_{\square_i}\}_{\square_i \in \mathbb{B}_i}$.

Therefore,

$$N\prod_{i=1}^{2}\mu_i \lessapprox \prod_{i=1}^{2}\sigma_{i,\square}|\mathbb{B}_i|. \tag{7.19}$$

Starting with (7.18) and inserting this estimate, we see that

$$\prod_{i=1}^{2}\left\|e^{it\Delta}f_i\right\|_{L^6(Y)}^{1/2} \lessapprox E^{O(1)}R^{-1/6}\prod_{i=1}^{2}\left[\mu_i^2\sigma_{i,\square}^{-2}\sum_{\square\in\mathbb{B}_i}\|f_{i,\square}\|_{L^2}^6\right]^{\frac{1}{6}\cdot\frac{1}{2}}$$

$$\lessapprox E^{O(1)}R^{-1/6}\prod_{i=1}^{2}\left[N^{-1}|\mathbb{B}_i|^2\sum_{\square\in\mathbb{B}_i}\|f_{i,\square}\|_{L^2}^6\right]^{\frac{1}{6}\cdot\frac{1}{2}} \lesssim E^{O(1)}R^{-1/6}N^{-1/6}\prod_{i=1}^{2}\|f_i\|_{L^2}^{1/2},$$

as desired.

8. Bilinear maximal estimate with small separation

In this section, using Theorem 7.2 and parabolic rescaling, we prove the following proposition, which implies Proposition 6.3.

Proposition 8.1. *Suppose that $\xi_0 \in B^2(0,1)$ and that f_i have Fourier supports in $B(\xi_0, 1/M)$ for some $M \geq 1$. Also suppose that the Fourier supports of f_i are separated by at least $1/(KM)$, where $K = K(\epsilon)$ is a large constant. Suppose that each f_i is concentrated in wave packets from $\mathbf{T}_Z(E)$, where $E \geq R^\delta$ and $Z = Z(P)$ and P is a product of distinct non-singular polynomials. Then*

$$\left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^3_x(B_R)L^\infty_t(0,R)} \lessapprox E^{O(1)}\|f_1\|_{L^2}^{1/2}\|f_2\|_{L^2}^{1/2}. \tag{8.1}$$

Proof. We can assume $M \ll R^{1/2}$, otherwise all wave packets were in the same direction and a direct computation would give us the desired result.

Since f is concentrated in wave packets from $\mathbf{T}_Z(E)$, we decompose $N_{ER^{1/2}}Z$ into balls Q of radius $R^{1/2}$. Let η be a smooth bump function approximating χ_Q. As we saw in the proof of Lemma 7.4, in Equation (7.7), the Fourier support of each function $\eta e^{it\Delta}f_i$ is essentially supported on

$$S^* := \{(\xi_1,\xi_2,\xi_3) : (\xi_1,\xi_2) \in S \text{ and } |\xi_3 - \xi_1^2 - \xi_2^2| \lesssim R^{-1/2}\},$$

where $S \subset B(0,1)$ is a strip of width $ER^{-1/2}$. Since the Fourier support of each f_i is also contained in $B(\xi_0, 1/M)$, the Fourier support of $\eta e^{it\Delta}f_i$ is also essentially contained in $B(\xi_0, \frac{2}{M}) \times \mathbb{R}$. The intersection of S^* with the cylinder $B(\xi_0, \frac{2}{M}) \times \mathbb{R}$ is contained in a rectangle of dimensions $\sim ER^{-1/2} \times 1/M \times 1/M$. We denote this rectangle by $A^*(Q)$. Since the Fourier support of each $\eta e^{it\Delta}f_i$ is contained in $A^*(Q)$, $|\eta e^{it\Delta}f_i|$ is morally constant on dual rectangles with dimensions $M \times M \times E^{-1}R^{1/2}$. We tile Q with such dual rectangles, which we denote $A_k(Q)$. The projection of each dual rectangle $A_k(Q)$ to the x-plane is an $M \times E^{-1}R^{1/2}$-rectangle.

Suppose that $\sup_{0<t<R}|e^{it\Delta}f_1 e^{it\Delta}f_2|^{1/2} \sim H$ on a set $U \subset B(0,R)$. It suffices for us to prove the bound

$$H|U|^{1/3} \lessapprox E^{O(1)}\|f_1\|_{L^2}^{1/2}\|f_2\|_{L^2}^{1/2}. \tag{8.2}$$

We will bound $|U|$ using the rectangles $A_k(Q)$. For the time being, let us suppose that $|\eta e^{it\Delta}f_i|$ is roughly constant on each $A_k(Q)$. This is not quite rigorous, but

useful for intuition. On the next page, we will come back to this point and give a rigorous argument.

There must be a collection of dual rectangles $A_k(Q_j)$ whose projections cover U and so that $|e^{it\Delta}f_1 e^{it\Delta}f_2|^{1/2} \sim H$ on each dual rectangle. We let X denote the union of these dual rectangles. Each $M \times M \times E^{-1}R^{1/2}$ rectangle $A_k(Q_j) \subset X$ has a projection with area $ME^{-1}R^{1/2}$, and since these projections cover U, we have the bound

$$|U| \lesssim M^{-1}|X|. \tag{8.3}$$

We can also assume that no two rectangles $A_k(Q_j) \subset X$ have essentially the same projection. This implies that X contains $\lesssim E^{O(1)}R^{1/2}M^{-1}$ rectangles $A_k(Q)$ in each cube Q. So for each cube Q, we get the bound

$$|X \cap Q| \lesssim E^{O(1)}MR. \tag{8.4}$$

We consider the $R^{1/2}$-cubes Q in $B^2(R) \times [0, R]$ that intersect X. We sort these $R^{1/2}$-cubes Q according to the dyadic value of $\left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Q)}$. We can choose a set of of $R^{1/2}$-cubes Q_j, $j = 1, 2, \cdots, N$, so that

$$\left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Q_j)} \text{ is essentially constant in } j, \tag{8.5}$$

and $|X| \lessapprox |X \cap Y|$, where $Y := \bigcup_{j=1}^N Q_j$. Using the locally constant property that $|e^{it\Delta}f_1 e^{it\Delta}f_2|^{1/2} \sim H$ on each rectangle $A_k(Q_j) \subset X$, we see that

$$H|X|^{1/6} \lessapprox E^{O(1)} \left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Y)}. \tag{8.6}$$

Since $|X \cap Q_j| \lesssim E^{O(1)}MR$ for each cube Q_j, $j = 1, ...N$, we see that $|X| \lessapprox |X \cap Y| \lesssim E^{O(1)}MNR$. Therefore,

$$H|X|^{1/3} \lessapprox E^{O(1)}M^{1/6}N^{1/6}R^{1/6} \left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Y)}. \tag{8.7}$$

Finally, since $|U| \lesssim M^{-1}|X|$, we have

$$H|U|^{1/3} \lessapprox E^{O(1)}M^{-1/6}N^{1/6}R^{1/6} \left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Y)}. \tag{8.8}$$

Therefore, our desired bound (8.2) follows from a generalization of Theorem 7.2, which we now state.

Proposition 8.2. *Suppose that f_1 and f_2 are as in Proposition 8.1. Suppose that $Q_1, Q_2, \cdots, Q_N$ are lattice $R^{1/2}$-cubes in $B^3(R)$ so that*

$$\left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Q_j)} \text{ is essentially constant in } j. \tag{8.9}$$

Let Y denote $\bigcup_{j=1}^N Q_j$. Then

$$\left\||e^{it\Delta}f_1|^{1/2}|e^{it\Delta}f_2|^{1/2}\right\|_{L^6(Y)} \lessapprox E^{O(1)}M^{1/6}N^{-1/6}R^{-1/6}\|f_1\|_{L^2}^{1/2}\|f_2\|_{L^2}^{1/2}. \tag{8.10}$$

If $M = 1$, then f_1 and f_2 have Fourier supports separated by ~ 1, and we can apply Theorem 7.2. We first find $Y' \subset Y$ with $|Y'| \approx |Y|$ so that for each i, $\|e^{it\Delta} f_i\|_{L^6(Q_j)}$ is essentially constant among all $Q_j \subset Y'$. Then we apply Theorem 7.2 to Y' to get (8.10):

$$\left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y)} \approx$$

$$\approx \left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y')} \lessapprox E^{O(1)} N^{-1/6} R^{-1/6} \|f_1\|_{L^2}^{1/2} \|f_2\|_{L^2}^{1/2}.$$

For larger M, the Fourier supports of f_1 and f_2 are only separated by $\sim 1/M$, and so we will need to apply parabolic rescaling before we can use Theorem 7.2.

Before we do this parabolic rescaling and prove Proposition 8.2, let us return to the issue of $|e^{it\Delta} f_i|$ being morally roughly constant on each rectangle $A_k(Q)$. We used the locally constant property to justify (8.6) above. We can rigorously prove (8.6) as follows. We mentioned above that each function $\eta_Q e^{it\Delta} f_i$ has Fourier transform essentially supported in a rectangle $A^*(Q)$ of dimensions $\sim ER^{-1/2} \times M^{-1} \times M^{-1}$. So the Fourier transform of their product, $g := \eta_Q^2 e^{it\Delta} f_1 e^{it\Delta} f_2$, is essentially supported in a rectangle with the same orientation and roughly the same dimensions. If $\hat{\psi}$ is designed to be identically 1 on this rectangle, then $g * \psi$ is essentially equal to g. We can choose such a ψ where $|\psi|$ is a rapidly-decaying approximation of $|A_k(Q_j)|^{-1} \chi_{A_k(Q_j)}$. Therefore, we see that

$$\sup_{A_k(Q)} |e^{it\Delta} f_1 e^{it\Delta} f_2| \lesssim R^{O(\delta)} \frac{\int_{R^\delta A_k(Q)} |e^{it\Delta} f_1 e^{it\Delta} f_2|}{|A_k(Q_j)|} + R^{-1000} \|f_1\|_{L^2} \|f_2\|_{L^2}, \tag{8.11}$$

where the second term accounts for the tail of ψ. Since $E \geq R^\delta$, we can assume that $R^\delta A_k(Q) \subset Q$.

We let X be a union of rectangles $A_k(Q_j)$ which each obeys

$$H \lesssim \sup_{A_k(Q_j)} |e^{it\Delta} f_1 e^{it\Delta} f_2|^{1/2}.$$

We can arrange that the projections of $10A_k(Q_j)$ cover U and also that any two rectangles $A_k(Q_j)$ in X have essentially different projections. Because of this covering, we still have $|U| \lesssim M^{-1}|X|$. Now if $H \lesssim R^{-100} \|f_1\|_{L^2}^{1/2} \|f_2\|_{L^2}^{1/2}$, then (8.2) follows trivially. Therefore, (8.11) tells us that for each $A_k(Q_j) \subset X$:

$$\int_{R^\delta A_k(Q)} |e^{it\Delta} f_1 e^{it\Delta} f_2| \gtrsim R^{-O(\delta)} |A_k(Q_j)| H^2.$$

We define Y just as above, and this inequality lets us rigorously justify (8.6):

$$H|X|^{1/6} \approx H|X \cap Y|^{1/6} \lessapprox E^{O(1)} \left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y)}.$$

It only remains to prove Proposition 8.2.

Proof. For function f with Fourier support in $B(\xi_0, 1/M)$, by parabolic rescaling, we have

$$\|e^{it\Delta} f(x)\|_{L^p(B^3(R))} \sim M^{\frac{4}{p}-1} \|e^{ir\Delta} \tilde{f}(y)\|_{L^p(B_{R/M} \times I_{R/M^2})}, \tag{8.12}$$

where $\tilde{f}$ has Fourier support in $B^2(0,1)$, $\|\tilde{f}\|_2 = \|f\|_2$, the new coordinates (y,r) and old coordinates (x,t) are related by

$$\begin{cases} y = x/M + 2t\xi_0/M, \\ r = t/M^2, \end{cases}$$

and $B_{R/M} \times I_{R/M^2}$ is a box of dimensions $\sim \frac{R}{M} \times \frac{R}{M} \times \frac{R}{M^2}$, which is the range for (y,r) under the change of variables as above. By (8.12), we have

$$\left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y)} \sim M^{-1/3} \left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y})}, \tag{8.13}$$

where $\tilde{f}_1, \tilde{f}_2$ have $1/K$-separated Fourier supports in $B^2(0,1)$, and $\tilde{Y}$ is a union of N $\frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M^2}$-boxes in $B_{R/M} \times I_{R/M^2}$, in correspondence to Y under the change of variables as above.

To use Theorem 7.2 to estimate $\left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y})}$, we decompose $B_{R/M} \times I_{R/M^2}$ as a union of $\frac{R}{M^2}$-balls $Q_{k,R/M^2}$, and inside each $Q_{k,R/M^2}$ we consider the $\sqrt{R}/M$-cubes $Q^{(k)}$ that intersect $\tilde{Y}$. First, we sort the balls $Q_{k,R/M^2}$ according to the dyadic values $\|e^{ir\Delta} \tilde{f}_i\|_{L^2(Q_{k,R/M^2})}$, $i = 1,2$. Then inside each $Q_{k,R/M^2}$ we sort the cubes $Q^{(k)}$ according to the dyadic values $\|e^{ir\Delta} \tilde{f}_i\|_{L^6(Q^{(k)})}$, $i = 1,2$. We can choose balls $Q_{k,R/M^2}$, $k = 1,2,\cdots,\bar{W}$, and inside each $Q_{k,R/M^2}$ we can choose a set of $\sqrt{R}/M$-cubes $Q_j^{(k)}$, $j = 1,2,\cdots,N_k$, so that

$$\approx N \text{ boxes in } \tilde{Y} \text{ are contained in } \bigcup_{k=1}^{\bar{W}} \tilde{Y}_k, \tag{8.14}$$

where $\tilde{Y}_k := \bigcup_{j=1}^{N_k} Q_j^{(k)}$, and the following conditions hold:

- (a). For each $i = 1,2$, $\|e^{ir\Delta} \tilde{f}_i\|_{L^2(Q_{k,R/M^2})}$ is essentially constant in $k = 1,\cdots,\bar{W}$.
- (b). For each $k = 1,\cdots,\bar{W}$, for each $i = 1,2$, $\|e^{ir\Delta} \tilde{f}_i\|_{L^6(Q_j^{(k)})}$ is essentially constant in $j = 1,\cdots,N_k$.
- (c). $\left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y}_k)}$ is essentially constant in $k = 1,\cdots,\bar{W}$.

Now by (8.9), (8.14) and the condition (c) as above, for each $1 \le k \le \bar{W}$ we have

$$\left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y})} \lessapprox \bar{W}^{\frac{1}{6}} \left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y}_k)}.$$

Since tangent-to-variety condition is preserved under parabolic rescaling, we can apply Theorem 7.2 to bound $\left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y}_k)}$ by

$$\lessapprox E^{O(1)} \left(\frac{R}{M^2}\right)^{-1/6} N_k^{-1/6} \left(\frac{R}{M^2}\right)^{-1/2} \prod_{i=1}^{2} \left\| e^{ir\Delta} \tilde{f}_i \right\|_{L^2(Q_{k,R/M^2})}^{1/2}.$$

By the condition (a) as above and parabolic rescaling (8.12), we have

$$\prod_{i=1}^{2} \left\| e^{ir\Delta} \tilde{f}_i \right\|_{L^2(Q_{k,R/M^2})}^{1/2} \lesssim \bar{W}^{-1/2} \prod_{i=1}^{2} \|e^{ir\Delta} \tilde{f}_i\|_{L^2(B_{R/M} \times I_{R/M^2})}^{1/2}$$

$$\sim \bar{W}^{-1/2} M^{-1} \prod_{i=1}^{2} \|e^{it\Delta} f_i\|^{1/2}_{L^2(B^3(R))} \lesssim \bar{W}^{-1/2} M^{-1} R^{1/2} \prod_{i=1}^{2} \|f_i\|_2^{1/2}.$$

Combining (8.13) and the above estimates for $\left\| |e^{ir\Delta} \tilde{f}_1|^{1/2} |e^{ir\Delta} \tilde{f}_2|^{1/2} \right\|_{L^6(\tilde{Y})}$, we get

$$\left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y)} \lessapprox E^{O(1)} \bar{W}^{-1/3} N_k^{-1/6} R^{-1/6} \prod_{i=1}^{2} \|f_i\|_2^{1/2}.$$

The above estimate holds for $\bar{W}$ indexes k's. For each k, there are N_k $\frac{\sqrt{R}}{M}$-cubes in $\tilde{Y}_k$, each $\frac{\sqrt{R}}{M}$-cube contains at most M $\frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M^2}$-boxes in $\tilde{Y}$, and there are $\approx N$ $\frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M} \times \frac{\sqrt{R}}{M^2}$-boxes in $\tilde{Y}$ that are contained in $\bigcup_{k=1}^{\bar{W}} \tilde{Y}_k$. By pigeonholing there is an index k satisfying

$$N \lessapprox N_k \bar{W} M.$$

Therefore

$$\left\| |e^{it\Delta} f_1|^{1/2} |e^{it\Delta} f_2|^{1/2} \right\|_{L^6(Y)} \lessapprox E^{O(1)} \bar{W}^{-1/6} N^{-1/6} M^{1/6} R^{-1/6} \prod_{i=1}^{2} \|f_i\|_2^{1/2}. \tag{8.15}$$

Since $\bar{W} \geq 1$, this completes the proof of Proposition 8.2. □

This finishes the proof of Proposition 8.1. □

Finally, to prove Proposition 6.3, we apply Proposition 8.1 to $f_{j,tang}$ on each ball B_j. We expand $f_{j,tang}$ into wave packets at the scale $\rho = R^{1-\delta}$ on the ball B_j. Because of the definition of $f_{j,tang}$, each wave packet will lie in the $R^{1/2+\delta}$-neighborhood of Z and the angles between the wave packets and the tangent space of Z will be bounded by $R^{-1/2+2\delta}$. For a detailed description of the wave packet decomposition of $f_{j,tang}$ on a smaller ball, see Section 7 of [10]. We define E so that $\rho^{1/2} E = R^{1/2+\delta}$. Since $\rho = R^{1-\delta}$, we get $E = R^{(3/2)\delta}$, and so $E\rho^{-1/2} = R^{-1/2+2\delta}$. Each new wave packet lies in the $E\rho^{1/2}$-neighborhood of Z, and the angles between the wave packets and the tangent space of Z are bounded by $E\rho^{-1/2}$. Therefore, the new wave packets are concentrated in $\mathbf{T}_Z(E)$. Now since $E^{O(1)} = R^{O(\delta)}$, the bound from Proposition 8.1 implies Proposition 6.3.

References

[1] 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.

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

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

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

[5] 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.

[6] 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.

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

[8] L. Guth and N. Katz *On the Erdős distinct distance problem in the plane*, Annals of Math 181 (2015), 155-190.
[9] L. Guth, *A restriction estimate using polynomial partitioning*, J. Amer. Math. Soc. 29 (2016), no. 2, 371-413.
[10] L. Guth, *Restriction estimates using polynomial partitioning II*, arXiv:1603.04250
[11] M. Lacey and X. Li, *On a Conjecture of E.M. Stein of the Hilbert Transform on Vector Fields*, Memoirs of the AMS **205** (2010), no. 965.
[12] M. T. Lacey and C. M. Thiele, *L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$,* Ann. Math. **146** (1997), 693-724.
[13] S. Lee, *On pointwise convergence of the solutions to Schrödinger equations in $\mathbb{R}^2$*, International Math. Research Notices. 2006, 32597, 1-21 (2006).
[14] R. Lucá and K. Rogers, An improved necessary condition for the Schrödinger maximal estimate, arXiv: 1506.05325v1.
[15] R. Lucá and K. Rogers, Coherence on fractals versus pointwise convergence for the Schrödinger equation, Comm. Math. Phys. 351 (2017), no. 1, 341-359.
[16] 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).
[17] K. M. Rogers, A local smoothing estimate for the Schrödinger equation, Adv. in Math., 219 (2008), 2105-2122.
[18] P. Sjölin, *Regularity of solutions to the Schrödinger equation*, Duke Math. Journal 55(3), 699-715 (1987).
[19] T. Tao and A. Vargas, *A bilinear approach to cone multipliers.II. Applications*, Geometric and Functional Analysis 10 (1), 185-215 (2000).
[20] L. Vega, *Schrödinger equations: pointwise convergence to the initial data*, Proceedings of the American Mathematical Society 102 (4), 874-878 (1988).

Mathematics Department, University of Illinois at Urbana-Champaign, Urbana IL
E-mail address: xdu7@illinois.edu

Department of Mathematics, MIT, Cambridge MA
E-mail address: lguth@math.mit.edu

Mathematics Department, University of Illinois at Urbana-Champaign, Urbana IL
E-mail address: xcli@math.uiuc.edu