LINEAR PROFILE DECOMPOSITIONS FOR A FAMILY OF FOURTH ORDER SCHRÖDINGER EQUATIONS

JIN-CHENG JIANG, SHUANGLIN SHAO, AND BETSY STOVALL

Abstract. We establish linear profile decompositions for the fourth order linear Schrödinger equation and for certain fourth order perturbations of the linear Schrödinger equation, in dimensions greater than or equal to two. We apply these results to prove dichotomy results on the existence of extremizers for the associated Stein–Tomas/Strichartz inequalities; along the way, we also obtain lower bounds for the norms of these operators.

1. Introduction

We consider the family of fourth order Schrödinger equations
\[
\begin{cases}
  iu_t + \Delta^2 u - \mu \Delta u = 0, & \mu \geq 0, \\
  u(0) = u_0 \in L^2_x(\mathbb{R}^d)
\end{cases}
\]  
(1.1)

where \( u : \mathbb{R} \times \mathbb{R}^d \to \mathbb{C}, \ d \geq 2. \) For \( \mu \neq 0, \) this is the free form of the nonlinear Schrödinger equation with a fourth order perturbation; this equation was introduced by Karpman [9] (see also [12, 10, 11, 13]) to study the effects of higher order dispersion in the propagation of solitary waves in plasmas.

The main result of this article (Theorem 3.1) is a linear profile decomposition for the equations given by (1.1). The theorem roughly states that, after passing to a subsequence, an \( L^2_x \)-bounded sequence of initial data may be decomposed as sum of asymptotically orthogonal pieces that are compact modulo symmetries, plus an error term with arbitrarily small dispersion. We then use this result to obtain dichotomy results on the existence of extremizers to the inequalities
\[
\| (\mu + |\nabla|^2)^{\frac{d+2}{2(d+2)}} e^{it(\Delta^2 - \mu \Delta)} f \|_{L^2_{t,x}} \leq C_{d,\mu} \| f \|_{L^2_x},
\]
in the spirit of [27]. Our results generalize those of [8], wherein the analogous theorems were proved in the one-dimensional case. Though the results we obtain are similar, we encounter new challenges in higher dimensions. One reason for this is that the propagator \( e^{it(\Delta^2 - \mu \Delta)} \) may be viewed as a Fourier extension operator, and the analysis of such operators seems to be much more difficult in dimensions greater than or equal to two. Further, comparatively little is known about Fourier restriction to surfaces where the curvature degenerates.

We are additionally motivated by recent applications of linear profile decompositions to the study of other dispersive equations, including wave [1, 16], Schrödinger [2, 22, 4, 15, 19, 21], KdV [27, 17], and Klein–Gordon [18].

Finally, much of the argument seems amenable to an extension to more general perturbations of the Schrödinger equation, for instance with \( \Delta^2 \) replaced by \( |\nabla|^\alpha \)

Date: August 13, 2013.
for $\alpha > 1$, but the authors have not investigated the extent to which this argument would need to be changed.

We now turn to a brief outline of the proof. For the remainder of this article, we let $S_\mu(t)$ denote the data-to-solution map,

$$S_\mu(t) = e^{it(\Delta^2 - \mu \Delta)}$$

and let $D_\mu$ denote the differential operator

$$D_\mu = \sqrt{\mu + |\nabla|^2}.$$

Section 2 is devoted to a proof of a refinement (Proposition 2.2) of the Stein–Tomas/Strichartz inequality

$$\| D_\mu^{d/2} S_\mu(t) f \|_{L_2^{2d+4} L_2^2} \lesssim \| f \|_{L_2^2}. \quad (1.2)$$

Roughly, this result states that if the left side of this inequality is greater than a constant times the right side, the function $f$ must contain a nontrivial wave packet that is concentrated on a small ‘cap’ on the Fourier side. This result is stronger than the annular refinement obtained in [5, Proposition 2.3] (and necessary for our finer-scale decomposition), and its proof strongly relies on Tao’s bilinear restriction theorem for elliptic hypersurfaces from [29]. One challenge that we face in proving the refined Strichartz inequality for these fourth order equations (compared with wave, Schrödinger, or Klein–Gordon) is the lack of scale invariance when $\mu \neq 0$, coupled with the absence of any natural analogue of the Lorentz or Galilei boosts. (The phase shifts $S_\mu(t)f \mapsto S_\mu(t)e^{it\alpha}f$ will provide a rough stand-in.)

Once we have obtained this refinement, we turn in Section 3 to the proof of the linear profile decomposition. This result, Theorem 3.1, follows by a familiar inductive argument. If $(u_n)$ is an $L_2^2$-bounded sequence such that $\| D_\mu^{d/2} S_\mu(t) u_n \|_{L_2^{2d+4} L_2^2}$ does not tend to 0, by the refined Strichartz estimate, there exists a sequence $\{g_n^1\}$ of pseudo-symmetries (true symmetries, i.e. spacetime translations, composed with scalings and phase shifts) such that the sequence $(g_n^1)^{-1} u_n$ has a nonzero weak limit $\phi^1 \in L_2^2$. The first profile is $\phi^1_n = g_n^1 \phi^1$, and we then repeat the argument on the sequence $u_n - \phi^1_n$. In fact, the refined Strichartz estimate gives a quantitative lower bound on the $L_2^2$ norm of $\phi^1$, and it is this that allows us to eventually show that for large $l$, the error terms $w_l^1$ are negligible. Having established the linear profile decomposition, for our application, we need to prove that the pieces $S_\mu(t)\phi^1_n$ and $S_\mu(t)\phi^k_n$ are asymptotically orthogonal for $j \neq k$. This is the content of Proposition 3.2.

Finally, in Section 4, we apply the linear profile decomposition to prove Theorems 4.1 and 4.2, which give lower bounds for the operator norms and dichotomy results on the existence of extremizers to (1.2) when $\mu = 0$ or 1 (by scaling, this extends to the general case). Very roughly, because of the asymptotic orthogonality of the profiles in the decomposition, after passing to a subsequence, an extremizing sequence to (1.2) must contain a single profile. After passing to a subsequence, there are three possibilities: compactness, convergence to a free Schrödinger wave, or convergence to a free fourth order Schrödinger wave. In the first case, extremizers exist. In the latter two, extremizers may fail to exist, but these cases give us the desired lower bounds on the operator norms.
Acknowledgements. The research of the first author was supported by National Science Council Grant NSC100-2115-M-007-009-MY2. The second author was supported by NSF DMS-1160981. The third author was supported by NSF DMS-0902667 and 1266336.

2. The refined Strichartz inequality

Fix $d \geq 2$ and $\mu \geq 0$. For the remainder of the article, let $\phi : \mathbb{R}^d \to [0, 1]$ be a smooth, radial, decreasing bump function with $\phi \equiv 1$ on $\{ |\xi| \leq 1 \}$ and $\phi \equiv 0$ on $\{ |\xi| \geq 2 \}$.

In this section, we prove the refined Strichartz inequality, which is crucial to our profile decomposition. Before stating it, we need a little notation.

Definition 2.1. A $\mu$-cap is a ball $\kappa = \{ \xi : |\xi - \xi_0| < r \}$, for some $\xi_0 \in \mathbb{R}^d$ and $r > 0$ satisfying $8r \leq |\xi_0| + \sqrt{\mu}$.

Given a cap $\kappa$ with center $\xi_0$ and radius $r$, we define an associated cutoff $\phi_\kappa(\xi) = \phi(\frac{\xi - \xi_0}{r})$.

Given $f$ we denote by $f_\kappa$ the function whose Fourier transform is given by $\hat{f}_\kappa = \phi_\kappa \hat{f}$.

With this notation in place, our refined Strichartz inequality is the following.

Proposition 2.2 (Refined Strichartz). Let $q = \frac{2(d^2 + 3d + 1)}{d^2}$ and $\theta = \frac{2}{(d+2)q}$. If $f \in L^2_x(\mathbb{R}^d)$, then

\[ \| S_\mu(t) D^{\frac{3q}{4} - \frac{1}{4}} f \|_{L^q_{t,x}} \leq \left( \sup_\kappa |\kappa|^{\frac{d+2}{4q} - \frac{1}{4}} \| S_\mu(t) D^{\frac{3}{2}} f_\kappa \|_{L^q_{t,x}} \right)^{\theta} \| f \|_{L^2_x}^{1-\theta}. \] (2.1)

Here the supremum is taken over all $\mu$-caps $\kappa$ as in Definition 2.1, and the implicit constant depends only on $d$.

Propositions of this kind have appeared in many places in the literature, and the outline we follow is a familiar one (cf. [2, 20]). The main new ingredient here is the parameter $\mu$. When $\mu = 0$, the graph of the function $\mu|\xi|^2 + |\xi|^4$ has vanishing curvature at 0, while in the case $\mu > 0$, the curvature never vanishes, but the scaling symmetry is broken. To deal with these issues, we begin by proving a refined Strichartz estimate associated to the decoupling of dyadic frequency annuli. By this scaling, it suffices to establish a refinement of the Stein–Tomas inequality for the Fourier extension operator associated to the surfaces $\{ (|\xi|^2 + |\xi|^4, \xi) : |\xi| \lesssim 1 \}$ and $\{ (\varepsilon|\xi|^2 + |\xi|^4, \xi) : |\xi| \sim 1 \}$, where $0 \leq \varepsilon \lesssim 1$. These surfaces are uniformly well-curved; indeed, they are elliptic on small (but uniform) frequency scales. We can thus apply Tao’s bilinear restriction theorem, and use methods developed in the context of the linear profile decomposition for Schrödinger to obtain refined estimates on these frequency localized regions. Finally, it is a simple matter to undo the scaling and glue the pieces back together, thereby obtaining (2.1).

We begin by noting that it suffices to prove Proposition 2.2 when $\mu = 0, 1$. Indeed, if $\mu > 0$, a simple computation shows that

\[ S_\mu(t) f = [S_1(\mu^2 t) f(\frac{\cdot}{\sqrt{\mu}})](\sqrt{\mu} \cdot), \quad \text{and} \quad D_\mu f = [\sqrt{\mu} D_1 f(\frac{\cdot}{\sqrt{\mu}})](\sqrt{\mu} \cdot). \]
Furthermore, if \( \kappa = \{ \xi : |\xi - \xi_0| < r \} \) is a 1-cap, \( \sqrt{\mu} \kappa = \{ \xi : |\xi - \sqrt{\mu} \xi_0| < \sqrt{\mu} r \} \) is a \( \mu \)-cap. Thus Proposition 2.2 for any \( \mu > 0 \) follows from Proposition 2.2 in the case \( \mu = 1 \) by scaling. For the remainder of this section, we will consider only the cases \( \mu = 0, 1 \).

We will employ two different Littlewood–Paley decompositions, depending on whether \( \mu = 0 \) or \( \mu = 1 \). Define

\[
\psi_N^0(\xi) = \phi(\frac{\xi}{N}), \quad \psi_N^1(\xi) = \phi(\frac{\xi}{N}) - \phi(\frac{2\xi}{N}), \quad \text{for } N \in \mathbb{Z}^+,
\]

\[
\psi_N^{\mu}(\xi) = \phi(\frac{\xi}{N}) - \phi(\frac{\xi}{2N}), \quad \text{for } N \in \mathbb{Z}^+.
\]

Regardless of the value of \( \mu \), the \( \psi_N^{\mu} \) form a partition of unity. We define the Littlewood–Paley projections \( P_N^\mu \) by \( \hat{f}_N^\mu = \hat{P}_N^\mu f = \psi_N^\mu \hat{f} \), \( \mu = 0, 1 \).

**Lemma 2.3.** If \( f \in L^2_t \), then

\[
\|S_\mu(t)D_\mu^{\frac{1}{2}} f\|_{L^{2(d+2)}_t L^{\frac{1}{2}}(\mathbb{R}^d)} \lesssim \sup_N \|S_\mu(t)D_\mu^{\frac{1}{2}} f_N^\mu\|_{L^{2(d+2)}_t L^{\frac{1}{2}}(\mathbb{R}^d)} \| f\|_{L^2}^\mu, \tag{2.2}
\]

for \( \mu = 0, 1 \). Here, the supremum is taken over frequencies \( N \in \{0\} \cup 2\mathbb{N} \) or \( N \in 2\mathbb{Z} \), depending on the value of \( \mu \).

In the proof of Lemma 2.3, we will use the following Strichartz estimates, which may be proved using the methods of stationary phase together with the main theorem of [14]. See [24] for further details.

**Proposition 2.4.** Let \( \mu \geq 0 \). For all \((q,r)\) satisfying \( 2 \leq q, r \leq \infty \), \((q,r) \neq (2,\infty)\), and \( \frac{2}{q} + \frac{d}{r} = \frac{d}{2} \),

\[
\|D_\mu^{\frac{1}{2}} S_\mu(t) f\|_{L^q_t L^r(\mathbb{R} \times \mathbb{R}^d)} \lesssim \| f\|_{L^2} \tag{2.3}
\]

The implicit constant may be taken to depend only on \( d, q, r \) and, in particular, may be chosen independently of \( \mu \).

**Proof.** Let \( d = 2 \). By the Littlewood–Paley square function estimate, simple arithmetic, Hölder’s inequality, and the Strichartz inequality, for \( \mu = 0, 1 \),

\[
\|S_\mu(t)D_\mu f\|_{L^{2d}_t L^2} \sim \iint (\sum_M |S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu|^2) (\sum_N |S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu|^2) \, dx \, dt
\]

\[
\sim \sum_{M \geq N} \|(S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu)(S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu))\|_{L^2_t}^2
\]

\[
\lesssim \sum_{M \geq N} \|S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu\|_{L^1_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu\|_{L^1_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu\|_{L^1_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu\|_{L^1_t}
\]

\[
\times \|S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu\|_{L^2_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu\|_{L^2_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_M^\mu\|_{L^2_t} \|S_\mu(t)D_\mu^\frac{1}{2} f_N^\mu\|_{L^2_t}.
\]

If \( \mu = 0 \), this implies by Plancherel that

\[
\|S_\mu(t)D_\mu^\frac{1}{2} f\|_{L^2_t}^2 \lesssim \sup_K \|S_\mu(t)D_\mu^\frac{1}{2} f_K^\mu\|_{L^2_t}^2 \sum_{M \geq N} M^{-\frac{1}{2}} N^{\frac{1}{2}} \|f_M^0\|_{L^2} \|f_N^0\|_{L^2},
\]

while if \( \mu = 1 \),

\[
\|S_\mu(t)D_\mu^\frac{1}{2} f\|_{L^2_t}^2 \lesssim \sup_K \|S_\mu(t)D_\mu^\frac{1}{2} f_K^1\|_{L^2_t}^2 \sum_{M \geq N} \langle M \rangle^{-\frac{1}{2}} \|f_M^0\|_{L^2} \|f_N^1\|_{L^2}.
\]
In either case, (2.2) follows from an application of Schur's test.

If $d > 2$, then by the square function estimate, the triangle inequality, Hölder's inequality, and the endpoint Strichartz inequality, for $\mu = 0, 1$,

$$
\|S_\mu(t)D_{\mu}^{\frac{d}{2}}f\|_{L^2_{t,x}} \lesssim \int \int (\sum_M |S_\mu(t)D_{\mu}^{\frac{d}{2}}f|^2_M) \frac{dt}{t^2} dx
$$

$$
\times (\sum_N |S_\mu(t)D_{\mu}^{\frac{d}{2}}f_N|^2) \frac{dx}{x^2} dt
$$

$$
\lesssim \sum_{M \geq N} \| (S_\mu(t)D_{\mu}^{\frac{d}{2}}f_M) (S_\mu(t)D_{\mu}^{\frac{d}{2}}f_N) \|_{L^2_{t,x}}^{\frac{d}{2}+2}
$$

$$
\lesssim \sum_{M \geq N} \| S_\mu(t)D_{\mu}^{\frac{d}{2}}f_M \|_{L^2_{t,x}} \| S_\mu(t)D_{\mu}^{\frac{d}{2}}f_M \|_{L^2_{t,x}}^{\frac{d}{2}}
$$

$$
\times \| S_\mu(t)D_{\mu}^{\frac{d}{2}}f_M \|_{L^2_{t,x}} \| S_\mu(t)D_{\mu}^{\frac{d}{2}}f_M \|_{L^2_{t,x}}^{\frac{d}{2}}
$$

$$
\times \sup_K \| S_\mu(t)D_{\mu}^{\frac{d}{2}}f_K \|_{L^2_{t,x}}^{\frac{d}{2}}
$$

$$
\times \sum_{M \geq N} \| f_M \|_{L^2_{t,x}} \| D_{\mu}^{\frac{d}{2}}f_M \|_{L^2_{t,x}} \| f_N \|_{L^2_{t,x}} \| D_{\mu}^{\frac{d}{2}}f_N \|_{L^2_{t,x}}.
$$

By Plancherel, if $\mu = 0$,

$$
\|S_\mu(t)D_{\mu}^{\frac{d}{2}}f\|_{L^2_{t,x}} \lesssim \sup_K \|S_\mu(t)D_{\mu}^{\frac{d}{2}}f_K\|_{L^2_{t,x}}^{\frac{d}{2}}
$$

$$
\times \sum_{M \geq N} M^{-\frac{4}{d(d+2)}} \| f_M \|_{L^2_{t,x}} \| f_N \|_{L^2_{t,x}},
$$

while if $\mu = 1$,

$$
\|S_\mu(t)D_{\mu}^{\frac{d}{2}}f\|_{L^2_{t,x}} \lesssim \sup_K \|S_\mu(t)D_{\mu}^{\frac{d}{2}}f_K\|_{L^2_{t,x}}^{\frac{d}{2}}
$$

$$
\times \sum_{M \geq N} \langle M \rangle^{-\frac{4}{d(d+2)}} \| f_M \|_{L^2_{t,x}} \| f_N \|_{L^2_{t,x}}.
$$

As in the $d = 2$ case, (2.1) follows by Schur's test. 

We now rescale one more time, to frequencies $|\xi| \sim 1$. We claim that the proposition follows from the next lemma.

**Lemma 2.5.** Let $q = \frac{2(d^3+3d+1)}{d^2}$. For $f \in L^2_d(\mathbb{R}^d)$, at scale $1$ we have

$$
\|S_\varepsilon(t)P_\varepsilon f\|_{L^2_{t,x}} \lesssim (\sup_K |\xi|^{\frac{d+2}{d+1} - \frac{d}{d+2}} \| S_\varepsilon(t)\hat{f}_\kappa \|_{L^2_{t,x}}^{\frac{1}{d+1}} \| f \|_{L^2_{t,x}}^{\frac{d+1}{d} \varepsilon + \frac{d}{d+1}}, \quad 0 \leq \varepsilon \leq 1,
$$

(2.4)

where the supremum is taken over caps $\kappa = \{ \xi : |\xi - \xi_0| < r \}$, with $\frac{1}{2} \leq |\xi_0| \leq 2$ and $r < \frac{1}{10}$. Additionally, at scale $0$,

$$
\|S_1(t)P_1 f\|_{L^2_{t,x}} \lesssim (\sup_K |\xi|^{\frac{d+2}{d+1} - \frac{d}{d+2}} \| S_1(t)\hat{f}_\kappa \|_{L^2_{t,x}}^{\frac{1}{d+1}} \| f \|_{L^2_{t,x}} \frac{d+1}{d}) \varepsilon + \frac{d}{d+1} \varepsilon + \frac{d}{d+1} \varepsilon,
$$

(2.5)

where the supremum is taken over caps $\kappa = \{ \xi : |\xi - \xi_0| < r \}$, with $|\xi_0| \leq 1$ and $r < \frac{1}{10}$. 

Assuming the lemma for the moment, we complete the proof of Proposition 2.2.

Proof of Proposition 2.2. By scaling, (2.4) implies that

$$N^{\frac{2d}{N}} \|S_{N\xi}(t) P^0_N f\|_{L^{2d-2}_{t,x}} \lesssim \left( \sup_\kappa |\kappa\|^ {\frac{4d-2}{32}} \frac{1}{N} N\|S_{N\xi}(t) f_{\kappa}\|_{L^2_{t,x}} \right)^{\frac{2d}{N}} \|f\|_{L^2_x},$$

(2.6)

for $0 \leq \varepsilon \lesssim 1$ and $N \in \mathbb{N}$, where the supremum is taken over caps $\kappa = \{\xi - \xi_0 < r\}$, where $\frac{N}{4} \leq |\xi_0| \leq 4N$ and $r < \frac{N}{32}$. We note that if $\mu = 0$ or $\mu = 1$ and $N \geq 1$, these are $\mu$-caps.

In the case $\mu = 0$, the refined Strichartz estimate (2.1) just follows from (2.2), (2.6) (with $\varepsilon = 0$), and Bernstein’s inequality.

In the case $\mu = 1$, we first apply (2.2). Then we use Bernstein’s inequality together with (2.5) for the scale $0$ term and (2.6) with $\varepsilon = \frac{1}{N}$ for the scale $N$ term.

(Note that if $N \in 2\mathbb{N}$, $P^0_N = P^1_N$.)

The remainder of the section will be devoted to the proof of Lemma 2.5. We will give the details for (2.4); (2.5) is similar, but slightly simpler because the parameter $\varepsilon$ is not present.

We recall that $P^0 f = \psi_0 f$, where $\psi_0$ is supported on $\{\frac{1}{2} \leq |\xi| \leq 2\}$. There exists a finite decomposition

$$\psi_0 = \sum_{j=1}^{C_d} \psi_{1,j},$$

(2.7)

with $\psi_{1,j} = \psi_0 \phi_j$, for smooth bump functions $\phi_j : \mathbb{R}^d \to [0, 1]$ whose supports have diameter equal to a small dimensional constant $c_d$, which will be determined in a moment. Crucially, the (large) constant $C_d$ may be chosen to depend on $c_d$ and $d$; more precisely, $C_d \lesssim_d c_d^{-d}$.

By (2.7), for any $f \in L^2_{t,x}$,

$$P^0 f = \sum_{j=1}^{C_d} f_{1,j}, \text{ where } f_{1,j} = \psi_{1,j} f, \quad i = 1, 2.$$  

(2.8)

By (2.8) and the triangle inequality,

$$\|S_{\varepsilon}(t) f_{1,j}\|_{L^{2d-2}_{t,x}} \leq \sum_{j=1}^{C_d} \|S_{\varepsilon}(t) f_{1,j}\|_{L^{2d-2}_{t,x}} \leq C_d \max_j \|S_{\varepsilon}(t) f_{1,j}\|_{L^{2d-2}_{t,x}}.$$ 

(2.9)

Fix $\frac{1}{4} < |\xi_0| < 4$ and define

$$\sigma_\varepsilon,\xi_0 = \{(|\varepsilon|\xi)^2 + |\xi|, \xi : |\xi - \xi_0| \leq 3c_d\}.$$  

Let $T_{\varepsilon,\xi_0}$ denote the translation

$$T_{\varepsilon,\xi_0}(\tau, \xi) = (\tau - \varepsilon|\xi_0|^2 - |\xi_0|^4, \xi - \xi_0)$$

and let $R_{\varepsilon,\xi_0} \in SO(d + 1)$ be a rotation mapping $\sigma_\varepsilon,\xi_0$ into standard position, that is,

$$R_{\varepsilon,\xi_0}(1, -2\varepsilon \xi_0 - 4|\xi_0|^4 \xi_0) = (\sqrt{1 + 4\varepsilon^2|\xi_0|^2 + 16\varepsilon|\xi_0|^4} + 16|\xi_0|^6, 0).$$

(2.10)

We may uniquely specify $R_{\varepsilon,\xi_0}$ by also requiring that

$$R_{\varepsilon,\xi_0}(2|\varepsilon|\xi_0^2 + 4|\xi_0|^4, \xi_0) = (0, \xi_0 \sqrt{1 + 4\varepsilon^2|\xi_0|^2 + 16\varepsilon|\xi_0|^4} + 16|\xi_0|^6)$$

(2.11)

and that $R_{\varepsilon,\xi_0}$ leave vectors perpendicular to $(1, 0)$ and $(0, \xi_0)$ invariant.
For $c_d$ sufficiently small, $R_{\varepsilon, \xi_0} T_{\varepsilon, \xi_0}(\sigma_{\varepsilon, \xi_0})$ may be written as a graph,

$$R_{\varepsilon, \xi_0} T_{\varepsilon, \xi_0}(\sigma_{\varepsilon, \xi_0}) \subset \{ (\Phi_{\varepsilon, \xi_0}(\xi), \xi) : |\xi| \leq b_d \}. $$

Furthermore, $\Phi_{\varepsilon, \xi_0}$ is smooth (actually analytic), uniformly in $0 \leq \varepsilon \leq 1$, and satisfies

- $\Phi_{\varepsilon, \xi_0}(0) = 0, \nabla \Phi_{\varepsilon, \xi_0}(0) = 0,$
- $|\Phi_{\varepsilon, \xi_0}|_{C^N([|\xi| < b_d])} \leq B_d$, where $N = N_d$ may be taken arbitrarily large,
- $H_{\varepsilon, \xi_0} = (\partial_i \partial_j \Phi_{\varepsilon, \xi_0})_{1 \leq i, j \leq d}$ is uniformly elliptic: $0 < \lambda_d \| \xi \|_{H_{\varepsilon, \xi_0}}$.

The last observation above follows from the fact that $\sigma_{\varepsilon, \xi_0}$ has positive Gaussian curvature, which is uniformly bounded below for $0 \leq \varepsilon \leq 1$.

Let $D_{\varepsilon, \xi_0}(\tau, \xi) = (\tau, (H_{\varepsilon, \xi_0}(0))^{\frac{1}{2}} \xi)$. Then

$$D_{\varepsilon, \xi_0} R_{\varepsilon, \xi_0} T_{\varepsilon, \xi_0}(\sigma_{\varepsilon, \xi_0}) \subset \{ (\frac{1}{2}|\xi|^2 + f_{\varepsilon, \xi_0}(\xi), \xi) : |\xi| \leq a_d \},$$

where $f_{\varepsilon, \xi_0}$ is analytic and satisfies

- $f_{\varepsilon, \xi_0}(0) = 0, \nabla f_{\varepsilon, \xi_0}(0) = 0, D^2 f_{\varepsilon, \xi_0}(0) = 0$
- $|f_{\varepsilon, \xi_0}|_{C^N([|\xi| < a_d])} \leq A_d$, uniformly in $0 \leq \varepsilon \leq 1$.

Let $\delta_d > 0$ be a small constant and $N$ be a large integer, which will be determined in a moment. We observe that decreasing $c_d$ has no affect on $f_{\varepsilon, \xi_0}$, but simply allows us to decrease the size of $a_d$. By Taylor’s theorem,

$$a_d^{-3} \|f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} + a_d^{-2} \|D f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} + a_d^{-1} \|D^2 f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} \lesssim \|D^3 f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} \lesssim 1,$$

so taking $a_d$ sufficiently small, we may assume that

$$a_d^{-2} \|f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} + a_d^{-1} \|D f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} + \|D^2 f_{\varepsilon, \xi_0}|_{C^0([|\xi| < a_d])} < \delta_d. \tag{2.12}$$

We may further assume that

$$a_d \|f_{\varepsilon, \xi_0}|_{C^N([|\xi| < a_d])} < \delta_d. \tag{2.13}$$

Let $B_{a_d}(\tau, \xi) = (a_d^{-2} \tau, a_d^{-1} \xi)$ and define $A_{\varepsilon, \xi_0} = B_{a_d} D_{\varepsilon, \xi_0} R_{\varepsilon, \xi_0} T_{\varepsilon, \xi_0}$. Then

$$A_{\varepsilon, \xi_0}(\sigma_{\varepsilon, \xi_0}) \subset \{ (\frac{1}{2}|\xi|^2 + g_{\varepsilon, \xi_0}(\xi), \xi) : |\xi| \leq 1 \},$$

where $g_{\varepsilon, \xi_0}$ is analytic and satisfies $g_{\varepsilon, \xi_0}(\xi) = a_d^{-2} f_{\varepsilon, \xi_0}(a_d \xi)$. By (2.12) and (2.13),

$$\|g_{\varepsilon, \xi_0}|_{C^N([|\xi| < 1])} < \delta_d.$$

As explained Section 9 of [29], the proof of the bilinear restriction theorem for paraboloids also establishes the following bilinear restriction result for bounded elliptic hypersurfaces.

**Theorem 2.6** ([29]). *There exists $\delta_d > 0$ and an integer $N$ such that if $\Phi(\xi) = \frac{1}{2}|\xi|^2 + g(\xi)$ with $g(\xi)$ smooth and $\|g\|_{C^N([|\xi| \leq 1])} < \delta_d$, then the operator $E_\Phi$ defined by

$$E_\Phi f(t, x) = \int_{\{|\xi| \leq 1\}} e^{i(t,x)(\Phi(\xi)\xi)} \hat{f}(\xi) d\xi$$

satisfies the following. Let $\rho_1, \rho_2 \subset \{|\xi| \leq 1\}$ satisfy

$$\text{diam}(\rho_1) \sim \text{diam}(\rho_2) \sim \text{dist}(\rho_1, \rho_2) \sim 1.$$*
If \( f_1, f_2 \) are \( L^2 \) functions whose Fourier transforms are supported on \( \rho_1, \rho_2 \), respectively, then
\[
\left\| \mathcal{F}_\Phi f_1 \right\|_{L^p_{t,x}} \lesssim \left\| f_1 \right\|_{L^2_{t,x}} \left\| f_2 \right\|_{L^2_{t,x}}, \quad p > \frac{d+3}{d+1},
\]
where the implicit constant depends only on \( d, p \).

Taking advantage of the symmetries of the Fourier transform, we have the following.

**Corollary 2.7.** For \( c_d \) sufficiently small, if \( \kappa_1 \) and \( \kappa_2 \) are two caps contained in \( \{ \frac{1}{3} < |\xi| < \frac{2}{3} \} \) satisfying
\[
\text{diam}(\kappa_1) \sim \text{diam}(\kappa_2) \sim \text{dist}(\kappa_1, \kappa_2) \sim r < c_d,
\]
then for any \( f \in L^2_{t,x} \) and \( p > \frac{d+3}{d+1} \),
\[
\left\| (S_\tau(t)f_{\kappa_1})(S_\tau(t)f_{\kappa_2}) \right\|_{L^p_{t,x}} \lesssim_d \| f_{\kappa_1} \|_{L^2_{t,x}} \| f_{\kappa_2} \|_{L^2_{t,x}}, \quad p > \frac{d+3}{d+1}.
\]

**Proof of Corollary 2.7 from Theorem 2.6.** Define \( \sigma_1 \) and \( \sigma_2 \) by
\[
\sigma_j = \{ (\varepsilon|\xi|^2 + |\xi|^4, \xi) : \xi \in \kappa_j \}.
\]
Choose \( \xi_0 \) such that \( \frac{1}{3} < |\xi_0| < \frac{2}{3} \) and \( \kappa_1, \kappa_2 \subset \{ |\xi - \xi_0| \leq 3r \} \). Let \( A_{\varepsilon,\xi_0} \) be the map defined above. Then
\[
A_{\varepsilon,\xi_0}(\sigma_j) \subset \{ (\frac{1}{2}|\xi|^2 + g_{\varepsilon,\xi_0}(\xi), \xi) : |\xi| < 1 \}.
\]
Let \( \pi \) denote projection onto the hyperplane \( \{ \tau = 0 \} \), and define \( F_{\varepsilon,\xi_0}(\xi) = \pi(A_{\varepsilon,\xi_0}(\varepsilon|\xi|^2 + |\xi|^4, \xi)) \). Then \( F_{\varepsilon,\xi_0} \) is a diffeomorphism, uniformly in \( 0 \leq \varepsilon \leq 1 \), and
\[
\| F_{\varepsilon,\xi_0} \|_{C^1}\{ (|\xi| < c_d) \} + \| F_{\varepsilon,\xi_0}^{-1} \|_{C^1}\{ (|\xi| < c_d) \} \lesssim_d 1. \tag{2.14}
\]
Furthermore, for \( j = 1, 2 \),
\[
\text{diam}(F_{\varepsilon,\xi_0}(\kappa_j)) \sim \text{dist}(F_{\varepsilon,\xi_0}(\kappa_1), F_{\varepsilon,\xi_0}(\kappa_2)) \sim r.
\]
Let \( \rho_j = r^{-1}F_{\varepsilon,\xi_0}(\kappa_j), j = 1, 2 \), let \( \Phi(\xi) = \frac{1}{2}|\xi|^2 + r^{-2}g_{\varepsilon,\xi_0}(r\xi) \), and let \( B_\tau(\tau, \xi) = (r^{-2}\tau, r^{-1}\xi) \).

For \( j = 1, 2 \) and \((t, x) \in \mathbb{R}^{1+d}\), by the definition of \( F_{\varepsilon,\xi_0} \), the change of variables formula, the definition of \( \Phi \), and the change of variables formula again,
\[
S_\tau(t)f_{\kappa_j}(t, x) = \int_{\mathbb{R}^d} e^{i(t, x) \cdot (\frac{1}{2}|\xi|^2 + |\xi|^4, \xi) F_{\varepsilon,\xi_0}(\xi)} d\xi
\]
\[
= \int_{\mathbb{R}^d} \exp\{i(t, x) \cdot A_{\varepsilon,\xi_0}^{-1}(\frac{1}{2}|F_{\varepsilon,\xi_0}(\xi)|^2 + g_{\varepsilon,\xi_0}(F_{\varepsilon,\xi_0}(\xi)), F_{\varepsilon,\xi_0}(\xi))\} f_{\kappa_j} \circ F_{\varepsilon,\xi_0}^{-1} |\det DF_{\varepsilon,\xi_0}(\eta)| d\eta
\]
\[
= \int_{\mathbb{R}^d} \exp\{i[A_{\varepsilon,\xi_0}^{-1}]^T(t, x)(\frac{1}{2}|\eta|^2 + g_{\varepsilon,\xi_0}(\eta), \eta)\} f_{\kappa_j} \circ F_{\varepsilon,\xi_0}^{-1} |\det DF_{\varepsilon,\xi_0}(\eta)| d\eta
\]
\[
= \int_{\mathbb{R}^d} \exp\{i[A_{\varepsilon,\xi_0}^{-1}]^T(t, x)(r^2\Phi(r^{-1}\eta), \eta)\} f_{\kappa_j} \circ F_{\varepsilon,\xi_0}^{-1} (r^{-1}\eta) |\det DF_{\varepsilon,\xi_0}(r\eta)| d\eta.
\]

Let \( g_j \) be the function whose Fourier transform is given by
\[
\hat{g}_j(\eta) = f_{\kappa_j} \circ F_{\varepsilon,\xi_0}^{-1} (r\eta) |\det DF_{\varepsilon,\xi_0}(r\eta)|.
\]
Then \( g_j \) is supported on \( \rho_j \),
\[
S_\tau(t)f_{\kappa_j}(t, x) = r^d \mathcal{F}_\Phi g_j(B_r^{-1}(A_{\varepsilon,\xi_0}^{-1})^T(t, x)), \tag{2.15}
\]
and by Plancherel, the change of variables formula, and (2.14)
\[ \|g_j\|_{L^2} \sim r^{-\frac{d}{2}} \|f_{\kappa_j}\| \det DF_{\kappa, \xi_0}^{-1} \|_{L^2} \sim r^{-\frac{d}{2}} \|f_{\kappa_j}\|_{L^2}. \] (2.16)

For any \( p > \frac{d+3}{d+1} \), by (2.15), a change of variables and the fact that \( \det A_{\kappa, \xi_0} \sim 1 \), Theorem 2.6, and (2.16),
\[ \|S(t)f_{\kappa_j}(S(t)f_{\kappa_j})\|_{L^p_t} \]
\[ = r^{2d}\| (\mathcal{E}_\Phi g_1) \circ B_{\kappa_j}^{-1}(A_{\kappa_j, \xi_0}^{-1})^T ((\mathcal{E}_\Phi g_2) \circ B_{\kappa_j}^{-1}(A_{\kappa_j, \xi_0}^{-1})^T) \|_{L^p_t} \]
\[ \sim r^{2d-\frac{d+2}{p}} \| (\mathcal{E}_\Phi g_1)(\mathcal{E}_\Phi g_2) \|_{L^p_t} \lesssim r^{2d-\frac{d+2}{p}} \|g_1\|_{L^2} \|g_2\|_{L^2} \]
\[ \sim r^{d-\frac{d+2}{p}} \|f_{\kappa_j}\|_{L^2} \|f_{\kappa_j}\|_{L^2}. \]

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

From the corollary, the proof of the following follows in what is by now a standard manner.

**Lemma 2.8.** Let \( q = \frac{2(d+3d+1)}{2d} \). If \( f \in L^2_x \), then
\[ \|S(t)f_{1,j}\|_{L^q_{t,x}} \lesssim ( \sup_\kappa |(\frac{d}{d+2})^{-\frac{d}{d+2}} \|S(t)(f_{1,j})_\kappa\|_{L^q_{t,x}} |)^{\frac{d+1}{2d}} \|f_{1,j}\|_{L^2_x}. \] (2.17)

where the supremum may be taken over caps \( \kappa \) with
\[ \kappa \subset \{ \frac{1}{2} \leq |\xi| \leq 4 \} \quad \text{and} \quad \text{diam}(\kappa) < 0.1. \] (2.18)

This lemma may be proved in the same way, *mutatis mutandis*, as Proposition 4.23 of [20], so we omit the details.

Since \( |\psi_{1,j}| \leq 1 \), by Plancherel, \( \|f_{1,j}\|_{L^2} \leq \|f\|_{L^2} \). Furthermore, \( S(t)(f_{1,j})_\kappa = (S(t)f_{1,j})_\kappa \). By construction, \( \|\psi_{1,j}\|_{L^1} \leq 1 \), so by Young’s convolution inequality (in the \( x \)-variable), \( \|S(t)(f_{1,j})_\kappa\|_{L^1_{t,x}} \lesssim \|S(t)f_{1,j}\|_{L^1_{t,x}} \). So finally, by combining (2.9) and (2.17), we obtain (2.4). The inequality (2.5) follows in a similar manner (but is simpler because there is no \( \varepsilon \)). Therefore Lemma 2.5 is proved, and this completes the proof of Proposition 2.2 and the section.

**3. Linear Profile Decomposition**

**Theorem 3.1** (Linear profile decomposition). Let \( (u_n)_{n \geq 1} \) be a bounded sequence of \( L^2_x \) functions. After passing to a subsequence, the following hold. For each \( j \geq 1 \), there exist \( \phi^j \in L^2_x \), a sequence of parameters \( (h_n^j, \xi_n^j, x_n^j, t_n^j) \in (0, \infty) \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \), and a sequence of \( L^2_x \) errors \( (w_n^j)_{n \geq 1} \) such that for each \( l \geq 1 \),
\[ u_n = \sum_{j=1}^l S_{\mu}(-t_n^j)g_n^j[e^{i(\xi_n^j, \xi_n^j)} \phi^j] + w_n^j, \] (3.1)

where \( g_n^j(\phi) := \frac{1}{h_n^j} \phi(\frac{x-x_n^j}{h_n^j}) \), and for each \( j \), either \( |h_n^j, \xi_n^j| \to \infty \) or \( \xi_n^j \equiv 0 \). The errors satisfy
\[ \lim_{l \to \infty} \limsup_{n \to \infty} \|D_{\mu}^{\frac{d+2}{d+1}} S_{\mu}(t)w_n^j\|_{L^2_{t,x} \mathbb{R}^d} = 0. \] (3.2)
For \( j \neq k \), \((h_n^j, \xi_n^j, x_n^j, t_n^j)_{n \geq 1}\) and \((h_n^k, \xi_n^k, x_n^k, t_n^k)_{n \geq 1}\) are pairwise orthogonal in the sense that

\[
\lim_{n \to \infty} \left( \frac{h_n^j + h_n^k}{h_n^j h_n^k} + \frac{|h_n^j \xi_n^j - h_n^k \xi_n^k|}{h_n^j h_n^k} + \frac{|t_n^j - t_n^k|}{(h_n^j)^4} + \frac{|t_n^j - t_n^k|((\mu + 2|\xi_n^j|^2)}{(h_n^j)^2} + \frac{|x_n^j - x_n^k + 2(t_n^j - t_n^k)(2|\xi_n^j|^2 + \mu)|}{h_n^j} \right) = \infty. \tag{3.3}
\]

Furthermore, if \( \lim_{n \to \infty} |h_n^j \xi_n^j - h_n^k \xi_n^k| \neq \infty \), then \( h_n^j \xi_n^j = h_n^k \xi_n^k \) for all \( n \), and if \( \lim_{n \to \infty} (h_n^j + h_n^k) \neq \infty \), then \( h_n^j = h_n^k \) for all \( n \). Finally, for each \( l \geq 1 \),

\[
\lim_{n \to \infty} \left( \|u_n\|_{L^2_t}^2 - \left( \sum_{j=1}^l \|\phi^j\|_{L^2_t}^2 + \|w_n^l\|_{L^2_t}^2 \right) \right) = 0. \tag{3.4}
\]

The orthogonality condition of parameters implies the following.

**Proposition 3.2 (Orthogonality of profiles).** For \( j \neq k \), along the subsequence satisfying (3.3),

\[
\lim_{n \to \infty} \left\| \left( D_{\mu}^{\frac{d}{d+2}} S_{\mu}(t) S_{\mu}(-t_n^j) g_n^j [e^{i(\cdot)h_n^j \xi_n^j \phi^j}] \right) \times \left( D_{\mu}^{\frac{d}{d+2}} S_{\mu}(t) S_{\mu}(-t_n^k) g_n^k [e^{i(\cdot)h_n^k \xi_n^k \phi^k}] \right) \right\|_{L^2_{t,x}}^{\frac{2(d+2)}{2(d+2)}} = 0. \tag{3.5}
\]

Thus by (3.2),

\[
\lim_{l \to \infty} \sup_{n \to \infty} \left\| D_{\mu}^{\frac{d}{d+2}} S_{\mu}(t) u_n \right\|_{L^2_{t,x}} \leq \sum_{j=1}^l \left\| D_{\mu}^{\frac{d}{d+2}} S_{\mu}(t-t_n^j) g_n^j [e^{i(\cdot)h_n^j \xi_n^j \phi^j}] \right\|_{L^2_{t,x}}^{\frac{2(d+2)}{2(d+2)}} = 0. \tag{3.6}
\]

We begin with the linear profile decomposition.

**Proof of Theorem 3.1.** The proof follows a familiar outline (cf. [1, 4, 8]). We construct the linear profile decomposition inductively. At each stage we will pass to a further subsequence, but to avoid a proliferation of sub- and superscripts, we will denote each subsequence by \((u_n)\). For each \( n, j \), let \( B_n^j \) denote the \( L^2_t \) isometry

\[
B_n^j = S_{\mu}(-t_n^j) g_n^j e^{i(\cdot)h_n^j \xi_n^j}.
\]

Set \( w_n^0 = u_n \) and assume that for some \( l \geq 0 \) we have found a subsequence of \((u_n), L^2_t \) functions \( \phi^j \), and sequences \((h_n^j, \xi_n^j, x_n^j, t_n^j), 1 \leq j \leq l \) such that for all \( 1 \leq k \leq l \), we have the following: (3.4) holds (with \( k \) in place of \( l \)); (3.3) holds for all \( k \neq j \in \{1, \ldots, l\} \); either \( |h_n^j \xi_n^k| \to \infty \) or \( \xi_n^k \equiv 0 \);

\[
u_n = \sum_{j=1}^k B_n^j \phi^j + w_n^k; \tag{3.7}
\]

\[
(B_n^k)^{-1} w_n^{k-1} \rightharpoonup \phi^k, \text{ weakly in } L^2_t. \tag{3.8}
\]

Define

\[
A_l^j = \limsup_{n \to \infty} \|u_n^j\|_{L^2_t}, \quad e^j = \limsup_{n \to \infty} \|D_{\mu}^{\frac{d}{d+2}} S_{\mu}(t) u_n^j\|_{L^2_{t,x}}^{\frac{2(d+2)}{2(d+2)}}.
\]
By the Strichartz inequality, $\varepsilon^l \lesssim A^l$. If $\varepsilon^l = 0$, then we are done, so we may assume that $\varepsilon^l > 0$. Passing to a subsequence, we may assume that $\|w_n^l\|_{L^2_x} \sim A^l$ and $\|D_{t,x}^{\alpha+} S_\mu(t) w^l_n\|_{L^2_{t,x}} \sim \varepsilon^l$, for all $n$.

By Proposition 2.2 and a little arithmetic, there exists a sequence $(\kappa_n)$ of $\mu$-caps such that for all sufficiently large $n$, 

$$A^l \left( \frac{\varepsilon^l}{A^l} \right)^{\frac{2}{d+2}} \lesssim |\kappa_n|^{-\frac{d+2}{d}} \|S_\mu(t) D_{t,x}^{\frac{d}{2}} (w^l_n) \kappa_n\|_{L^2_{t,x}}. \tag{3.9}$$

By Hölder’s inequality, Bernstein’s inequality (since $(w^l_n)_{\kappa_n}$ is frequency localized), and Young’s convolution inequality (since $\|\hat{\phi}_n\|_{L^1} \sim 1$),

$$\|S_\mu(t) D_{t,x}^{\frac{d}{2}} (w^l_n) \kappa_n\|_{L^2_{t,x}} \lesssim \|S_\mu(t) D_{t,x}^{\frac{d}{2}} (w^l_n) \kappa_n\|_{L^2_{t,x}} \leq (A^l)^{\frac{2}{d+2}} \|S_\mu(t) (w^l_n) \kappa_n\|_{L^2_{t,x}}^{\frac{1}{2(d+2)}}. \tag{3.10}$$

Combining this with (3.9) and the fact that $(w^l_n)_{\kappa_n}$ is smooth (since it has compact Fourier support), there exist parameters $(x_n, t_n)$ such that

$$A^l \left( \frac{\varepsilon^l}{A^l} \right)^{\frac{2}{d+2}} \lesssim |\kappa_n|^{-\frac{d+2}{d}} \|S_\mu(t_n) (w^l_n) \kappa_n(x_n)|. \tag{3.10}$$

Write $\kappa_n = \{\xi : |\xi - \xi_n| < h_n^{-1}\}$, and set

$$(h_n^{-1}, \xi_n^{-1}, x_n^{-1}, t_n^{-1}) = (h_n, \xi_n, x_n, t_n).$$

The sequence

$$(B_{n+1}^l)^{-1} w^l_n = e^{-ix_n \xi_n} g_n^{-1} S_\mu(t_n) w^l_n$$

is bounded in $L^2_x$, so after passing to a subsequence, we may extract a weak limit; say 

$$(B_{n+1}^l)^{-1} w^l_n \rightharpoonup \phi^{l+1}, \text{ weakly in } L^2_x. \tag{3.11}$$

Thus we have

$$\lim_{n \to \infty} (\|w^l_n\|_{L^2_x}^2 - \|w^l_n - B_{n+1}^l \phi^{l+1}\|_{L^2_x}^2 - \|\phi^{l+1}\|_{L^2_x}^2) = \lim_{n \to \infty} (2((B_{n+1}^l)^{-1} w^l_n, \phi^{l+1}) - 2\|\phi^{l+1}\|_{L^2_x}^2) = 0, \tag{3.12}$$

which implies that (3.4) holds with $l$ replaced by $l + 1$ and 

$$w^{l+1} = w^l_n - B_{n+1}^l \phi^{l+1}. \tag{3.13}$$
Thus by the Strichartz inequality, after iterating, we obtain (3.2). (Indeed for (3.2) to fail, $\varepsilon^l$ must stay large, but in this case (3.13) implies that $A^l$ must decrease to 0, a contradiction to the Strichartz inequality.)

By changing $\phi^{l+1}$ if necessary, we may assume that either $|h_n^{l+1}\xi_n| \to \infty$ or $\xi_n \equiv 0$. Indeed, if $|h_n^{l+1}\xi_n| \not\to \infty$, after passing to a subsequence, $h_n^{l+1}\xi_n \to \xi_0 \in \mathbb{R}^d$, and so we may replace $\phi^{l+1}$ with $e^{it\xi_0}g^{l+1}$ and $\xi_n^{l+1}$ with 0. Similar arguments justify the assertions that for each $k < l+1$, either $\left(\frac{h_n^k}{h_n^{l+1}} + \frac{h_n^{l+1}}{h_n^k}\right) \to \infty$ or $h_n^k \equiv h_n^{l+1}$ and that either $|h_n^k\xi_n - h_n^{l+1}\xi_n| \to \infty$ or $h_n^k \equiv h_n^{l+1}$. 

Finally, we turn to (3.3). The crux of the argument will be the following.

**Lemma 3.3.** If the limit in (3.3) is infinite, $(B_n^k)^{-1}B_n^l \to 0$ in the weak operator topology on $\mathcal{B}(L^2_x)$. Otherwise, after passing to a subsequence, there exists an $L^2_x$ isometry $B_n^k$ such that $(B_n^k)^{-1}B_n^l \to B_n^k$ in the strong operator topology on $\mathcal{B}(L^2_x)$.

**Proof.** Let $B_n^k = (B_n^k)^{-1}B_n^l$. It suffices to prove that if the limit in (3.3) is infinite, 

$$\lim_{n \to \infty} \langle B_n^k \phi, \psi \rangle = 0,$$

for all Schwartz functions $\phi, \psi$ with compact frequency support, and that otherwise after passing to a subsequence, $B_n^k \phi \to B^k \phi$ in $L^2_x$ for all $\phi \in L^2_x$.

A simple computation shows that 

$$\|B_n^k \phi\|_{L^2_x} \leq \left(\frac{h_n^k}{h_n^l}\right)^{\frac{d}{2}} \|\hat{\phi}\|_{L^1_x}, \quad \|(B_n^k)^{-1} \psi\|_{L^2_x} \leq \left(\frac{h_n^l}{h_n^k}\right)^{\frac{d}{2}} \|\hat{\psi}\|_{L^1_x}.$$ 

By Hölder’s inequality and Plancherel, 

$$|\langle B_n^k \phi, \psi \rangle| \lesssim \min\left\{\left(\frac{h_n^k}{h_n^l}\right)^{\frac{d}{2}} \|\hat{\phi}\|_{L^1_x} \|\hat{\psi}\|_{L^2_x}, \left(\frac{h_n^l}{h_n^k}\right)^{\frac{d}{2}} \|\hat{\psi}\|_{L^1_x} \|\hat{\phi}\|_{L^2_x}\right\},$$

and since $\phi, \psi$ are Schwartz, if $(\frac{h_n^k}{h_n^l} + \frac{h_n^l}{h_n^k}) \to \infty$, the right hand side of the above inequality tends to 0. Thus we may henceforth assume that $h_n^k \equiv h_n^k \equiv h_n$.

Now assume that $|h_n^k\xi_n - h_n\xi_n| \to \infty$. By assumption, $\hat{\phi}, \hat{\psi}$ have compact support; say $\text{Supp} \hat{\phi}, \text{Supp} \hat{\psi} \subset \{||\xi|| \leq R\}$. Then 

$$\text{Supp} \hat{B_n^k} \phi \subset \{||\xi - \xi_n^k|| < h_n^{-1}R\}, \quad \text{Supp} \hat{B_n^k} \psi \subset \{||\xi - \xi_n^k|| < h_n^{-1}R\}.$$ 

For sufficiently large $n$, these sets are disjoint, so 

$$\langle B_n^k \phi, \psi \rangle = \langle \hat{B_n^k} \phi, \hat{B_n^k} \psi \rangle \to 0.$$ 

Thus we may assume that $\xi_n^k \equiv \xi_n \equiv \xi_n$.

With these assumptions in place, we compute 

$$B_n^k \phi = \omega_n^k T_n^k S_n^k R_n^k P_n^k \phi,$$

where 

$$\omega_n^k = \exp\left[i(\xi_n(x_n^k - x_n^k) + (t_n^k - t_n^l))(||\xi_n||^2 + \mu||\xi_n||^2)\right],$$

$$T_n^k \psi(x) = \psi(x + x_n^k - x_n^l + t_n^k - t_n^l + 4|\xi_n|^2 \xi_n^2 + 2\mu \xi_n^2) \hat{\psi}(\xi) d\xi,$$

$$S_n^k \psi = \int \exp\left[i(x \xi + \frac{x_n^k - t_n^l}{h_n^k})(2||\xi||^2|\xi_n|^2 + 4\xi_n^2 + 2\mu||\xi||^2)\right] \hat{\psi}(\xi) d\xi,$$

$$R_n^k \psi(x) = \int \exp\left[i(x \xi + \frac{4(t_n^k - t_n^l)}{h_n^k})(2||\xi||^2|\xi_n|^2)\right] \hat{\psi}(\xi) d\xi,$$

$$P_n^k \psi = \exp\left[4\xi_n^2 \frac{(t_n^k - t_n^l)}{h_n^k} \Delta^2\right] \psi.$$
After passing to a subsequence, \( \omega^{kj} \to \omega \in S^1 \), so \( B^{kj}_n \phi - \omega T^{kj}_n S^{kj}_n R^{kj}_n P^{kj}_n \phi \to 0 \) in \( L^2_\omega \). Thus it suffices to show that the conclusions of the lemma hold for \( T^{kj}_n S^{kj}_n R^{kj}_n P^{kj}_n \phi \) instead of \( B^{kj}_n \).

We write
\[
T^{kj}_n S^{kj}_n R^{kj}_n P^{kj}_n \phi(x) = \int e^{i \Phi^{kj}_n(x, \xi)} \hat{\phi}(\xi) \, d\xi,
\]
where
\[
\Phi^{kj}_n(x, \xi) = \xi \cdot \left( x + \frac{x_n^k - x_j^j + (t_n^k - t_j^j)(4|\xi_n|^2 + 2\mu \xi_n)}{h_n} \right) + \frac{(t_n^k - t_j^j)|\xi|^4}{h_n^4} + \frac{(t_n^k - t_j^j)(2|\xi_n|^2|\xi|^2 + 4(\xi_n \cdot \xi)^2 + \mu|\xi|^2)}{h_n^2} + \frac{4(t_n^k - t_j^j)|\xi|^2 |\xi_n \cdot \xi|}{h_n^3}.
\]

Since \( \hat{\Phi^{kj}_n}(x, \xi) = \frac{24(t_n^k - t_j^j)}{h_n^3} \), by the method of stationary phase ([28, Chapter VIII.2.2]),
\[
|T^{kj}_n S^{kj}_n R^{kj}_n P^{kj}_n \phi(x)| \lesssim (1 + \frac{|t_n^k - t_j^j|}{h_n^3})^{-\frac{3}{4}}.
\]
Thus if \( \frac{|t_n^k - t_j^j|}{h_n^3} \to \infty \), \( T^{kj}_n S^{kj}_n R^{kj}_n P^{kj}_n \phi \to 0 \), weakly in \( L^2_\omega \). Thus we may assume that \( \frac{|t_n^k - t_j^j|}{h_n^3} \not\to \infty \). Passing to a subsequence, \( \frac{|t_n^k - t_j^j|}{h_n^3} \to s^{kj} \), so
\[
\|P^{kj}_n \phi - e^{is^{kj} \Delta^2} \phi\|_{L^2_\omega} \to 0,
\]
for every \( \phi \in L^2_\omega \). Thus it suffices to show that the conclusions of the lemma hold for \( T^{kj}_n S^{kj}_n R^{kj}_n \phi \) instead of \( B^{kj}_n \).

Passing to a subsequence, either \( \xi_n \equiv 0 \) or \( \frac{\xi_n}{|\xi_n|} \to \xi_0 \in S^{d-1} \). The case when \( \xi_n \equiv 0 \) is much easier, so we assume henceforth that \( \frac{\xi_n}{|\xi_n|} \to \xi_0 \in S^{d-1} \). Passing to a further subsequence, \( 0 < \xi_n \cdot \xi_0 \sim |\xi_n| \) for all \( n \). We write
\[
T^{kj}_n S^{kj}_n R^{kj}_n \phi(x) = \int e^{i \Phi^{kj}_n(x, \xi)} \hat{\psi}(x) \, d\xi,
\]
where now we set
\[
\Phi^{kj}_n(x, \xi) = \xi \cdot \left( x + \frac{x_n^k - x_j^j + (t_n^k - t_j^j)(4|\xi_n|^2 + 2\mu \xi_n)}{h_n} \right) + \frac{(t_n^k - t_j^j)(2|\xi_n|^2|\xi|^2 + 4(\xi_n \cdot \xi)^2 + \mu|\xi|^2)}{h_n^2} + \frac{4(t_n^k - t_j^j)|\xi|^2 |\xi_n \cdot \xi|}{h_n^3}.
\]
Since
\[
|\xi_n \cdot \nabla|^{3/2} \Phi^{kj}_n(x, \xi)| = \frac{24(t_n^k - t_j^j)|\xi_n||\xi_n \cdot \xi|}{|\xi_n|^{3/2}} \lesssim \frac{|t_n^k - t_j^j||\xi_n|}{h_n^3},
\]
if \( \frac{|t_n^k - t_j^j||\xi_n|}{h_n^3} \to \infty \), by stationary phase (as above), \( \|T^{kj}_n S^{kj}_n R^{kj}_n \phi\|_{L^\infty_\omega} \to 0 \), so \( T^{kj}_n S^{kj}_n R^{kj}_n \phi \to 0 \), weakly in \( L^2_\omega \). Otherwise, as above, after passing to a subsequence, \( R^{kj}_n \phi \to R^{kj}_n \phi \) in \( L^2_\omega \), for some unitary operator \( R^{kj}_n \). Thus it suffices to consider \( T^{kj}_n S^{kj}_n \).

Similar arguments show that \( T^{kj}_n S^{kj}_n \phi \to 0 \), weakly in \( L^2_\omega \) if \( \frac{|t_n^k - t_j^j|(2|\xi_n|^2 + \mu)}{h_n^3} \to \infty \), so we may assume that this term is bounded, and after passing to a subsequence, \( S^{kj}_n \phi \to S^{kj}_n \phi \) in \( L^2_\omega \). This reduces matters to proving that the conclusions of the lemma hold for \( T^{kj}_n \), and since \( T^{kj}_n \) is just a translation, elementary arguments show that if \( \frac{|x_n^k - x_j^j + (t_n^k - t_j^j)(4|\xi_n|^2 + 2\mu \xi_n)}{h_n^3} \to \infty \), then \( T^{kj}_n \phi \to 0 \), weakly in \( L^2_\omega \).
if 

\[ \frac{x^k_r - x^{l_r}_j + (t^k_n - t^{l_n}_j)(4|\xi_j|^2 + 2|\xi_n|)}{h_n} \rightarrow y^{kj} \in \mathbb{R}^d, \]

then \( T^{kj}_n \phi \rightarrow \phi(y^{kj}) \) in \( L^2_x \). This completes the proof of the lemma. \( \square \)

Now we complete the proof of the linear profile decomposition by showing that (3.3) holds for all \( 1 \leq k < j = l + 1 \). Suppose (3.3) failed for some \( 1 \leq k < j = l + 1 \). Then

\[
0 = \operatorname{wk-lim}_{n \to \infty} (B^n_k)^{-1} w^{k-1}_n - \phi^k = \operatorname{wk-lim}_{n \to \infty} (B^n_k)^{-1} w^{k}_n
\]

\[
= \operatorname{wk-lim}_{n \to \infty} (B^n_k)^{-1} [b^{k+1}_n \phi^{k+1} + \cdots + b^l_n \phi^l + w^l_n].
\]

We have assumed that (3.3) holds for all \( k < j \leq l \), so by Lemma 3.3,

\[
\operatorname{wk-lim}_{n \to \infty} (B^n_k)^{-1} [b^{k+1}_n \phi^{k+1} + \cdots + b^l_n \phi^l] = 0,
\]

which implies that

\[
\operatorname{wk-lim}_{n \to \infty} (B^n_k)^{-1} w^l_n = 0.
\]

On the other hand, by Lemma 3.3, there exists a unitary operator \( B^{l+1,k} \) such that after passing to a subsequence, \( (B^{l+1,k}_n)^{-1} B^{l+1}_n \to B^{l+1,k} \) in the strong operator topology on \( L^2_x \); thus for any test function \( \psi \),

\[
\langle (B^{l+1,k})^{-1} \phi^{l+1}, \psi \rangle = \langle \phi^{l+1}, B^{l+1,k} \psi \rangle = \lim_{n \to \infty} \langle B^{l+1}_n w^l_n, B^{l+1,k} \psi \rangle
\]

\[
= \lim_{n \to \infty} \langle B^{l+1}_n w^l_n, (B^{l+1}_n)^{-1} B^n_k \psi \rangle = \lim_{n \to \infty} \langle B^n_k w^l_n, \psi \rangle = 0.
\]

Since \( B^{l+1,k} \) is unitary and \( \phi^{l+1} \neq 0 \), this is a contradiction. Thus (3.3) must hold for all \( 1 \leq k < j \leq l + 1 \), and this completes the proof of Theorem 3.1. \( \square \)

Next we prove Proposition 3.2.

**Proof of Proposition 3.2.** Since

\[
\psi \mapsto D_{\mu}^{\frac{d}{\alpha}} S_{\mu}(t) S_{\mu}(-t^l_n) g^l_n[e^{t^{(\cdot)}_{\alpha}} h^l_n, \xi_n, \xi_n \psi]
\]

is a bounded linear operator from \( L^2_x \to L^2_{t,x} \), with operator norm bounded by a constant independent of \( \mu, h^l_n, \xi_n, x_n^l, t^l_n \), by standard approximation arguments, it suffices to prove (3.5) for \( \phi^l \) and \( \phi^k \) lying in some dense subclass of \( L^2_x \). We will assume henceforth that they are Schwartz functions whose Fourier transforms are supported on a compact set that does not contain 0.

Our proof will use the following pointwise estimates.

**Lemma 3.4.** Fix \( \mu \geq 0 \) and let \( \psi \) be a Schwartz function with compact frequency support that does not contain 0. There exists an \( L^2_{t,x} \) function \( v = v_\psi \), depending only on \( \psi \), such that

\[
|D_{\mu}^{\frac{d}{\alpha}} S_{\mu}(t) \psi(x)| \leq (1 + \mu)^{\frac{d}{\alpha}} v((1 + \mu)t, x),
\]

and such that if \( |a| \gg \max \{ |\xi| : \xi \in \operatorname{Supp} \hat{\psi} \} \),

\[
|D_{\mu}^{\frac{d}{\alpha}} S_{\mu}(t) e^{i(\cdot)a} \psi(x)| \leq (2|a|^2 + \mu)^{\frac{d}{\alpha}} v((2|a|^2 + \mu)t, x + (4|a|^2 + 2\mu)at)
\]

(3.14)
Proof of Lemma 3.4. We give the details for the second case, when $|a| \gg \max\{|\xi| : \xi \in \text{Supp } \hat{\psi}\}$. The case when $a = 0$ is similar, but a little simpler.

Consider the function

$$w_{a,\mu}(\xi) = \left(\frac{\sqrt{\mu + |\xi + a|^2}}{\sqrt{\mu + 2|a|^2}}\right)^{\frac{|a|}{2}}.$$

Then $w_{a,\mu}$ and all of its derivatives are bounded on the support of $\hat{\psi}$, uniformly in $a$ and $\mu$. This follows from a simple induction argument and the fact that

$$c_\psi \leq |\xi + a| \leq \sqrt{\mu + |\xi + a|^2} \leq \sqrt{\mu + |a|^2} + |\xi| \leq 2\sqrt{\mu + 2|a|^2},$$

for all $\xi \in \text{Supp } \hat{\psi}$.

Fix $x$ and define $y = x + t(4|a|^2a + 2\mu a)$. Then

$$|D_{\mu}^{\frac{|a|}{2}} S_\mu(t)e^{i\langle \cdot \rangle a}\hat{\psi}(x)|$$

\begin{align*}
&= (\mu + 2|a|^2)^{\frac{|a|}{2}} \int_{\mathbb{R}^d} e^{i(x\xi + t(|\xi|^4 + |\mu|\xi^2)}w_{a,\mu}(\xi - a)\hat{\psi}(\xi - a) \, d\xi \\
&= (\mu + 2|a|^2)^{\frac{|a|}{2}} \int_{\mathbb{R}^d} e^{i(x\xi + t(|\xi|^4 + 4|\xi|^2|a|^2 + 4(\xi a)^2 + 4\mu \xi^2))}
\times w_{a,\mu}(\xi)\hat{\psi}(\xi) \, d\xi.
\end{align*}

(3.16)

Define $\Phi = \Phi_{t, y, a, \mu}$ by

$$\Phi(\xi) = y\xi + t(|\xi|^4 + 4|\xi|^2|a|^2 + 2|\xi|^2a + 4(\xi a)^2 + \mu |\xi|^2).$$

We compute the gradient and Hessian of $\Phi$:

$$\nabla \Phi(\xi) = y(4|\xi|^2\xi + 8(\xi a)\xi + 4|\xi|^2a + 4|a|^2\xi + 8(\xi a)a + 2\mu \xi)$$

(3.17)

$$H_\Phi = (\Phi_{ij}(\xi)) = t(8\xi_\xi\xi_j + 4|\xi|^2\delta_{ij} + 8\xi_\xi a_j + 8(\xi a)\delta_{ij} + 8a_j \xi_j
+ 4|a|^2\delta_{ij} + 8a_j a_j + 2\mu \delta_{ij}),$$

where $\delta_{ij}$ equals 1 if $i = j$ and 0 otherwise.

We now prove the estimate (3.15) by standard techniques from harmonic analysis (see [28, Chapter 8]). By Hölder’s inequality and (3.16),

$$|D_{\mu}^{\frac{|a|}{2}} S_\mu(t)e^{i\langle \cdot \rangle a}\hat{\psi}(x)| \leq (\mu + 2|a|^2)^{\frac{|a|}{2}} \|w_{a,\mu}\| L^\infty \|\hat{\psi}\| L^1 \lesssim (\mu + 2|a|^2)^{\frac{|a|}{2}},$$

(3.19)

for all $(t, x) \in \mathbb{R}^{1+d}$.

On the support of $\hat{\psi}$, we have

$$|4|\xi|^2\xi + 8(\xi a)\xi + 4|\xi|^2a + 4|a|^2\xi + 8(\xi a)a + 2\mu \xi| \lesssim \mu + 2|a|^2,$$

so if $(\mu + 2|a|^2)|t| \ll |y|, |\nabla \Phi(\xi)| \gtrsim |y|$ throughout the support of $\hat{\psi}$. Therefore for any $N \geq 1$, integrating by parts $N$ times in the right side of (3.16),

$$|D_{\mu}^{\frac{|a|}{2}} S_\mu(t)e^{i\langle \cdot \rangle a}\hat{\psi}(x)| \leq C_{\psi, N}(\mu + 2|a|^2)^{\frac{|a|}{2}} \left(\frac{1}{1 + |\mu|}\right)^N,$$

(3.20)

whenever $(\mu + 2|a|^2)|t| \ll |y|$.

If $(\mu + 2|a|^2)|t| \gtrsim |y|$, then $\nabla \Phi$ may vanish on the support of $\hat{\psi}$, so we examine the Hessian of $\Phi$. Since $|a| \gg \langle \xi \rangle$ for all $\xi \in \text{Supp } \hat{\psi}$, $t^{-1}H_\Phi(\xi) \gtrsim c_\psi(\mu + 2|a|^2)$ (i.e. $t^{-1}H_\Phi - c_\psi(\mu + 2|a|^2)$ is a positive matrix). Thus by stationary phase

$$|D_{\mu}^{\frac{|a|}{2}} S_\mu(t)e^{i\langle \cdot \rangle a}\hat{\psi}(x)| \leq C_{\psi}(\mu + 2|a|^2)^{\frac{|a|}{2}} \left(\frac{1}{1 + |\mu + 2|a|^2||t|}\right)^{-\frac{1}{2}},$$

(3.21)
wherever \((\mu + 2|a|^2)|t| \gtrsim |y|\).

Combining (3.19), (3.20), and (3.21), and recalling the definition of \(y\), (3.14) holds with

\[
v(s, y) = C\psi \left[ \frac{1}{1 + |y|} \right]^2 \chi_{|y| \gg |s|} + \left( \frac{1}{1 + |x|} \right)^2 \chi_{|y| \leq |s|},
\]

and it is easy to check that \(v \in L^{2(d+2)}_{s,y}\). This completes the proof of Lemma 3.4. □

Now we return to estimating the quantity in (3.5), where we continue to assume that \(\phi^k, \phi^j\) are Schwartz functions with compact frequency supports that do not contain zero. We will use two families of \(L^{2(d+2)}_{t,x}\) isometries:

\[
G^j_n u(t, x) = (h^j_n)^{- \frac{d(d+4)}{2(d+2)} + 2} u(t - t^j_n, x - x^j_n h^j_n) \quad \text{if } a = 0
\]

\[
L_{a, \mu} u(t, x) = \begin{cases}
(1 + \mu)(x - x^j_n \xi_n) \mu(t, x), & \text{if } a = 0 \\
2(a|a|^2 + \mu)(x - x^j_n \xi_n) \mu(t, x) + (4|a|^2 + 2\mu)at, & \text{if } a \neq 0.
\end{cases}
\]

If we consider \(a^j_n = h^j_n \xi_n\), after passing to a subsequence, either \(a^j_n = 0\) for all \(n\), or \(|a^j_n| \gg \max \{(|\xi| : \xi \in \text{Supp} \phi^j)\} \) for all \(n\). In either case, we can apply Lemma 3.4 once we move the translation/scaling isometries across the differential operators:

\[
|D^a_{\mu} S_{\mu}(t - t^j_n) g^j_n [e^{i(\xi_n \xi_n^j \phi^j)](x)}] = |G^j_n D^a_{\mu} S_{\mu}(t)[e^{i(\xi_n \xi_n^j \phi^j)](x)}] \\
\leq G^j_n L_{a^j_n, \mu^j_n} v^j(t, x),
\]

and similarly,

\[
|D^a_{\mu} S_{\mu}(t - t^j_n) g^k_n [e^{i(\xi_n \xi_n^j \phi^k)](x)}] \leq G^k_n L_{a^j_n, \mu^j_n} v^k(t, x),
\]

where \(v^j \in L^{2(d+2)}_{t,x}\), \(a^j_n = h^j_n \xi_n\), \(\mu^j_n = (h^j_n)^2 \mu\), and similarly with \(j\) replaced by \(k\).

The proof of the proposition will thus be complete once we prove the following.

**Lemma 3.5.** If \(v^j, v^k\) are any \(L^{2(d+2)}_{t,x}\) functions and the orthogonality condition (3.3) holds, then

\[
\lim_{n \to \infty} \|[G^j_n L_{a^j_n, \mu^j_n} v^j][G^k_n L_{a^j_n, \mu^j_n} v^k]\|_{L^{d+2}_{t,x}} = 0. \tag{3.22}
\]

**Proof of Lemma 3.5.** Since the \(G^j_n\) and \(L_{a, \mu}\) operators are uniformly bounded on \(L^{d+2}_{t,x}\), it suffices to prove this for \(v^j\) and \(v^k\) lying in some dense subclass of \(L^{d+2}_{t,x}\). We assume henceforth that they are compactly supported Schwartz functions; say \(\text{Supp} v^j, \text{Supp} v^k \subseteq \{|x| \leq R\}\).

Passing to a subsequence, we may assume that each of the summands in (3.3) has a limit. (This passage is harmless because to prove (3.22), it suffices to prove that every subsequence has a further subsequence along which the limit is zero.)
We first consider the case when $\frac{h^k_n}{h_n} \to \infty$. Using the $L_{t,x}^{2(\frac{d+2}{d})}$ isometry properties and Hölder’s inequality,
\[
\left\| [G_n^j L_{a_n^j,\mu_n^j} v^j] [G_n^k L_{a_n^k,\mu_n^k} v^k] \right\|_{L_{t,x}^{\frac{d+2}{d}}}
= \left\| [v^j] [L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v^k] \right\|_{L_{t,x}^{\frac{d+2}{d}}}
\leq \left\| [v^j] \right\|_{L_{t,x}^{2(\frac{d+2}{d})}} \left( \text{Supp}(L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v^k) \right)
\times \left\| L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v^k \right\|_{L_{t,x}^{2(\frac{d+2}{d})}}.
\tag{3.23}
\]

By (3.23), this implies (3.22). By a similar argument, (3.22) also holds if $h_n \to \infty$.

The exact computation of $L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v^k$ is elementary but tedious; however, it is painless to verify that
\[
L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v^k(t, x) = c_j^k v^k(d_j^k t - s_j^k, h_j^k n x - y_j^k(t)),
\]
for positive constants $c_j^k, d_j^k, s_j^k$, real numbers $h_j^k$, and functions $y_j^k : \mathbb{R} \to \mathbb{R}^d$. Fix $t$. If
\[
(t, x) \in \text{Supp}(L_{a_n^j,\mu_n^j}^{-1} (G_n^j)^{-1} G_n^k L_{a_n^k,\mu_n^k} v_k),
\]
then $x - \frac{h_j^k}{h_n^k} y_j^k(t) \leq \frac{h_j^k}{h_n^k} R$. By Hölder,
\[
\left\| v^j(t, \cdot) \right\|_{L_{t,x}^{2(\frac{d+2}{d})}} \leq \left\| v^j \right\|_{L_{t,x}^{\infty}}.
\]
Integrating the above estimate with respect to $t$ and recalling that $\frac{h_j^k}{h_n^k} \to \infty$,
\[
\left\| v^j \right\|_{L_{t,x}^{2(\frac{d+2}{d})}} \leq \left\| v^j \right\|_{L_{t,x}^{\infty}}.
\]
By (3.23), this implies (3.22). By a similar argument, (3.22) also holds if $\frac{h_k^j}{h_n^k} \to \infty$.

Henceforth, we may assume that $h_n^k \equiv h_n^k \equiv h_n$, so $\mu_n^j \equiv \mu_n^k \equiv \mu_n$ as well. Using a change of variables, we may now remove the dilations from the $G_n^j$:
\[
\left\| [G_n^j L_{a_n^j,\mu_n^j} v^j] [G_n^k L_{a_n^k,\mu_n^k} v^k] \right\|_{L_{t,x}^{\frac{d+2}{d}}} = \left\| [L_{a_n^j,\mu_n^j} v^j] [G_n^k L_{a_n^k,\mu_n^k} v^k] \right\|_{L_{t,x}^{\frac{d+2}{d}}},
\tag{3.24}
\]
where $G_n^j v(t, x) = v(t - \frac{h_n^j - h_n^j}{h_n^k}, x - \frac{x_n^j - x_n^j}{h_n^k})$.

We now break into three cases: $a_n^j \equiv a_n^k \equiv 0$; $a_n^j \equiv 0$ and $|a_n^k| \to \infty$; and $|a_n^j|, |a_n^k| \to \infty$.

We deal with the easiest of the three cases, $a_n^j \equiv a_n^k \equiv 0$, first. In this case,
\[
L_{a_n^j,\mu_n^j} v^j(t, x) = (1 + \mu_n) \frac{d}{d-t} v^j((1 + \mu_n) t, x)
\]
\[
G_n^j L_{a_n^j,\mu_n^j} v^j(t, x) = (1 + \mu_n) \frac{d}{d-t} v^k((1 + \mu_n) t - \frac{h_n^j - h_n^j}{h_n^k}, x - \frac{x_n^j - x_n^j}{h_n^k}).
\]
By our assumptions on the various parameters, including the assumption that the orthogonality condition (3.3) holds, either $\frac{v^j - v^k}{h_n^j} \to \infty$ or $\frac{x_n^j - x_n^k}{h_n^j} \to \infty$. In either case, $\text{Supp}(L_{a_n^j,\mu_n^j} v^j) \cap \text{Supp}(G_n^j L_{a_n^k,\mu_n^k} v^k)$ is empty for sufficiently large $n$, so the
right hand side of (3.24) is eventually zero. By the identity (3.24), this establishes (3.22).

We turn now to the case when \( a_n \equiv 0, \ |a_n^k| \to \infty \). (By symmetry, this argument also covers the case when the roles of \( j \) and \( k \) are reversed.) Arguing similarly to (3.23),

\[
\|L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^j\|_{L_t^2} \lesssim \|v_j\|_{L_t^2} (\text{Supp} \ L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^k).
\]

We compute:

\[
L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^k(t, x) = \left( \frac{2|a_n|^2 + \mu_n}{1 + \mu_n} \right)^{\frac{d}{2(d+2)}} v^k \left( \frac{2|a_n|^2 + \mu_n}{1 + \mu_n} \right) \left( t - \frac{(t_n^k - \ell_n^j)(1 + \mu_n)}{\overline{a_n}} \right),
\]

\[
x - \frac{x_n^k - x_j}{\overline{a_n}} + \frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} a_n \left( t - \frac{(t_n^k - \ell_n^j)(1 + \mu_n)}{\overline{a_n}} \right)
\]

where

\[
s_n^{jk} = \frac{(t_n^k - \ell_n^j)(1 + \mu_n)}{\overline{a_n}},
\]

\[
y_n^{jk} = \frac{x_n^k - x_j}{\overline{a_n}} + \frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} a_n s_n^{jk}.
\]

Fix \( x \). If \( (t, x) \in (\text{Supp} \ v^j) \cap (\text{Supp} \ L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^k) \), it satisfies \( |x| \leq R \) and \( |x + \frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} a_n t - y_n^{jk}| \leq 2R \). Therefore

\[
|\frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} a_n t - y_n^{jk}| \leq 2R,
\]

so recalling that \( |a_n^k| \gtrsim 1 \) for all \( n \),

\[
|t - r_n^{jk}| \leq \frac{2R(1 + \mu_n)}{\frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} |a_n|} \lesssim \frac{R}{|a_n|},
\]

where

\[
r_n^{jk} = \frac{(1 + \mu_n)y_n^{jk} - a_n^k}{\frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} |a_n|}.
\]

Integrating in \( t \),

\[
\|v^j(\cdot, x)\|_{L_t^\infty} \left( \left\{ |t + r_n^{jk}| \gtrsim \frac{R}{|a_n|} \right\} \right) \lesssim \left( \|v^j\|_{L_t^\infty} \right) \frac{R}{|a_n|}.
\]

Integrating this in \( x \),

\[
\|v^j\|_{L_t^\infty} \left( \frac{d}{2(d+2)} \right) L_t^{\frac{d}{2}} (\text{Supp} \ L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^k) \lesssim \|v^j\|_{L_t^\infty} R^d \frac{R}{|a_n|} \to 0.
\]

This implies (3.22), and completes the case when \( a_n^j \equiv 0, \ |a_n^k| \to \infty \). We compute

\[
L_{a_n,k}^{-1} G_n^{jk} L_{a_n,k} v^k(t, x) = \left( c_n^{jk} \right)^{\frac{d}{2(d+2)}} v^k \left( c_n^{jk} \left( t - s_n^{jk} \right), x - y_n^{jk} - b_n^{jk} t \right),
\]

where

\[
c_n^{jk} = \frac{2|a_n|^2 + \mu_n}{2|a_n|^2 + \mu_n},
\]

\[
s_n^{jk} = \frac{(t_n^k - \ell_n^j)(2|a_n|^2 + \mu_n)}{\overline{a_n}}
\]

\[
y_n^{jk} = \frac{x_n^k - x_j}{\overline{a_n}} + \frac{4|a_n|^2 + 2\mu_n}{1 + \mu_n} \left( t_n^k - \ell_n^j \right).
\]
Suppose that $|a_n^j - a_n^k| \to \infty$. We claim that $|b_{n}^{jk}| \to \infty$. Indeed,

$$
|b_{n}^{jk}| \geq b_{n}^{jk} \cdot \frac{a_n^k - a_n^j}{|a_n^k - a_n^j|} \\
= \left\{ (2|a_n^j|^2 + \mu_n) |a_n^k - a_n^j| \right\}^{-1} \left\{ 5[a_n^j \cdot (a_n^k - a_n^j)]^2 + 6|a_n^k - a_n^j|^2 a_n^j \cdot (a_n^k - a_n^j) \\
+ 2|a_n^k - a_n^j|^4 + |a_n^k - a_n^j|^2 (|a_n^j|^2 + \mu_n) \right\} \\
\geq \frac{|a_n^k - a_n^j|^2 (|a_n^j|^2 + \mu_n)}{2|a_n^j|^2 + \mu_n} |a_n^k - a_n^j| \geq \frac{1}{2} |a_n^k - a_n^j|,
$$

where for the second inequality, we used Cauchy–Schwartz and the elementary inequality

$$
5(ab)^2 + 2b^4 \geq \left( \frac{3}{2} \right) (ab)^2 + 2b^4 \geq 6|ab|^3 b,
$$

for all $a, b \in \mathbb{R}$. Since $|b_{n}^{jk}| \to \infty$, arguing exactly as we did in the case $a_n^j \equiv 0, |a_n^k| \to \infty$, we can establish (3.22).

Henceforth, we may assume that $a_n^j \equiv a_n^k$. Thus $a_n^j \equiv 1$ and $b_{n}^{jk} \equiv 0$. If $|x^j| \to \infty$ or $|y^j| \to \infty$, the (compact) supports of $\psi$ and $L_{a_n^j, \mu_n}^{-1} G_n^{jk} L_{a_n^k, \mu_n}^{jk}$ are disjoint for sufficiently large $n$, so (3.22) holds. Since $|x^j| \geq \frac{|x^j - y^j|}{h_n}$, this completes the proof in the case $|x^j|, |y^j| \to \infty$. \hfill \Box

4. Application: Dichotomy result on the existence of extremizers

As an application of the profile decomposition in Theorem 3.1, we establish lower bounds for the operator norms and a dichotomy result on existence of extremizers. Similar results have previously appeared in [8]. We begin by defining

$$
A_{\mu} := \sup_{\phi \in L^2_{x, s}, \phi \neq 0} \| \phi \|_{L^2_{x, s}}^{-1} \| D_{\mu}^{\frac{d+1}{2}} S_{\mu}(\cdot) \phi \|_{L^2_{t, x} \times \mathbb{R}^d} \quad \mu \geq 0, \quad (4.1)
$$

and

$$
B := \sup_{\psi \in L^2_{x, s}, \psi \neq 0} \| \psi \|_{L^2_{x, s}}^{\frac{1}{2}} \| e^{it\Delta} \psi \|_{L^2_{t, x} \times \mathbb{R}^d}. \quad (4.2)
$$

These are finite by the Strichartz inequalities for the fourth order Schrödinger and Schrödinger equations.

We say that a function $\phi$ is an extremizer for $A_{\mu}$ (resp. $B$) if $\| \phi \|_{L^2_{x, s}} \neq 0$ and $\phi$ maximizes the ratio in (4.1) (resp. (4.2)). A sequence $\{ \phi_n \}$ is an extremizing sequence for $A_{\mu}$ if

$$
A_{\mu} = \lim_{n \to \infty} \| \phi_n \|_{L^2_{x, s}}^{-1} \| D_{\mu}^{\frac{d+1}{2}} S_{\mu}(\cdot) \phi_n \|_{L^2_{t, x} \times \mathbb{R}^d};
$$

$\{ \phi_n \}$ is $L^2_{x, s}$-normalized if $\| \phi_n \|_{L^2_{x, s}} = 1$ for all $n$.

Extremizers are known to exist for $B$ ([6, 7] for $d = 1, 2$, [26] for $d \geq 3$). In dimensions 1 and 2, it is known in addition that the extremizers are Gaussian functions, modulo symmetries of the Schrödinger equation.
\textbf{Theorem 4.1.} The operator norms $A_0$ and $B$ satisfy:

$$A_0 \geq 3^{-\frac{1}{d+2}} 2^{-\frac{d}{d+2}} B.$$  

(4.3)

If the inequality is strict in (4.3), then extremizers exist for $A_0$. If extremizers do not exist and $\{\phi_n\}$ is an $L_2^d$-normalized extremizing sequence for $A_0$, there exist a sequence of parameters $(h_n, \xi_n, x_n, t_n)$ with $|h_n\xi_n| \rightarrow \infty$ and an extremizer $\psi$ for $B$ such that after passing to a subsequence,

$$\lim_{n \rightarrow \infty} \|\phi_n - S_0(t_n)g_n(e^{i(t)h_n\xi_n}\psi \circ \ell_{h_n\xi_n}^{-1})\|_{L_2^d} = 0,$$  

(4.4)

where for $a \in \mathbb{R}^d$, $\ell_a$ denotes the transformation

$$\ell_a(\xi) = \sqrt{6} \text{proj}_a(\xi) + \sqrt{2}(\xi - \text{proj}_a(\xi)).$$  

(4.5)

Conversely, if equality holds in (4.3), any sequence $\{\phi_n\}$ satisfying (4.4) for a sequence of parameters $(h_n, \xi_n, x_n, t_n)_{n \geq 1}$ with $|h_n\xi_n| \rightarrow \infty$ and an extremizer $\psi$ for $B$ is an extremizing sequence for $A_0$.

If $\mu > 0$, then by scaling, $A_{\mu} = A_1$; scaling also gives a natural correspondence between extremizing sequences (and, if they exist, extremizers) for $A_{\mu}$ and those for $A_1$. Thus we only state the dichotomy result in the case $\mu = 1$.

\textbf{Theorem 4.2.} The operator norms satisfy $A_1 \geq \max\{A_0, B\}$, and if this inequality is strict, extremizers exist for $A_1$. If extremizers do not exist and $\{\phi_n\}$ is an $L_2^d$-normalized extremizing sequence for $A_1$, then there exist a sequence of parameters $(h_n, \xi_n, x_n, t_n)$ and a function $\phi \in L_2^d$ such that after passing to a subsequence,

$$\lim_{n \rightarrow \infty} \|\phi_n - S_1(t_n)g_n e^{i(t)h_n\xi_n}\psi\|_{L_2^d} = 0,$$  

(4.6)

Moreover, in this case, one of the following occurs: either $A_1 = B$, $h_n \rightarrow \infty$, $|\xi_n| \rightarrow 0$, and $\psi$ is an extremizer for $B$, or $A_1 = A_0$, $h_n \rightarrow 0$, $\xi_n \equiv 0$, and $\psi$ is an extremizer for $A_0$.

The analogue of the final conclusion of Theorem 4.1 for $A_1$ is the following. If $A_1 = B$, $h_n \rightarrow \infty$, $|\xi_n| \rightarrow 0$, $\psi$ is an extremizer for $B$, and $\phi_n$ satisfies (4.6), then $\phi_n$ is an extremizing sequence for $A_1$. If $A_1 = A_0$, then $A_0 \geq B$, so the inequality is strict in (4.3), which implies that extremizers exist for $A_0$. Furthermore, in this case, if $h_n \rightarrow 0$, $\xi_n \equiv 0$, $\psi$ is an extremizer for $A_0$, and $\phi_n$ satisfies (4.6), then $\phi_n$ is an extremizing sequence for $A_1$.

The proofs of these theorems will rely on the linear profile decomposition and the following lemmas.

\textbf{Lemma 4.3.} Let $\phi \in L_2^d(\mathbb{R}^d)$. If $(a_n)$ is a sequence in $\mathbb{R}^d$ with $|a_n| \rightarrow \infty$, then

$$\lim_{n \rightarrow \infty} \|\nabla |\nabla |^\frac{d}{2}\phi_n e^{it\Delta} e^{i(a_n)\phi} - a_n|\nabla |^\frac{d}{2}\phi_n e^{i(x\cdot a_n + |a_n|)\phi} e^{-i|a_n|^2t\Delta} (\phi \circ \ell_{a_n})(\ell_{a_n}^{-1}(x + 4t|a_n|^2a_n))\|_{L_{t,x}^{2(4+2)}} = 0,$$  

(4.7)

with $\ell_a$ as in (4.5).

\textbf{Lemma 4.4.} Let $\phi \in L_2^d(\mathbb{R}^d)$. Let $(h_n, \xi_n)$ be a sequence in $(0, \infty) \times \mathbb{R}^d$. Define $g_n \phi(x) = h_n^{-\frac{d}{2}} \phi(\frac{x}{h_n})$. If $h_n \rightarrow 0$ and $\xi_n \equiv 0$, then

$$\lim_{n \rightarrow \infty} \|\nabla |\nabla |^\frac{d}{2} S_1(t) g_n e^{i(t)h_n\xi_n \phi} - \nabla |\nabla |^\frac{d}{2} g_n \phi\|_{L_{t,x}^{2(4+2)}} = 0.$$  

(4.8)
If \( h_n \to \infty \) and \( \xi_n \equiv 0 \),
\[
\lim_{n \to \infty} \|(\nabla)^{\frac{d}{2}} S_1(t)g_n e^{i(t)h_n \xi_n} \phi - e^{-it\Delta} g_n \phi\|_{L_{t,x}^{2(d+2)}} = 0.
\] (4.9)

If \( |\xi_n| \lesssim 1 \) for all \( n \) and \( |h_n \xi_n| \to \infty \),
\[
\lim_{n \to \infty} \|(\nabla)^{\frac{d}{2}} S_1(t)g_n e^{i(t)h_n \xi_n} \phi - (\xi_n)^{\frac{d}{2}} e^{i(x\xi_n + t|\xi_n|^2)} e^{-it\Delta}
\]
\[
\times [g_n(\phi \circ \ell_n)](\ell_n^{-1}(x + 4t|\xi_n|^2 \xi_n + 2t \xi_n))\|_{L_{t,x}^{2(d+2)}} = 0,
\] (4.10)
where \( \ell_n(\xi) = \sqrt{6|\xi_n|^2 + 1 (\text{proj}_{\xi_n} \xi) + \sqrt{2|\xi_n|^2 + 1 (\xi - \text{proj}_{\xi_n} \xi})} \).

If \( |\xi_n| \to \infty \) and \( |h_n \xi_n| \to \infty \),
\[
\lim_{n \to \infty} \|(\nabla)^{\frac{d}{2}} S_1(t)g_n e^{i(t)h_n \xi_n} \phi - (\xi_n)^{\frac{d}{2}} e^{i(x\xi_n + t|\xi_n|^2)} e^{-i\frac{|\xi_n|^2}{h_n} t\Delta}
\]
\[
\times [g_n(\phi \circ \ell_n)](\ell_n^{-1}(x + 4t|\xi_n|^2 \xi_n + 2t \xi_n))\|_{L_{t,x}^{2(d+2)}} = 0,
\] (4.11)
with \( \ell_n = \ell_{\xi_n} \), using the notation from (4.5).

The lemmas will be proved at the end of this section. Before proceeding to their proofs, we show how Theorem 4.1 can be proved from Lemma 4.3 and indicate how to adapt this proof for Theorem 4.2.

\textbf{Proof of Theorem 4.1.} Let \( (h_n, \xi_n, x_n, t_n) \) be a sequence of parameters such that \( |h_n \xi_n| \to \infty \) and let \( \psi \) be an extremizer for \( B \). Assume that the sequence \( \{\phi_n\} \subset L_2^2 \) satisfies
\[
\lim_{n \to \infty} \|\phi_n - e^{it_n \Delta^2} g_n(\phi(\xi_n) \psi \circ \ell_{h_n \xi_n}^{-1})\|_{L_2^2} = 0.
\]

By the Strichartz inequality (2.3), changes of variables, Lemma 4.4, and the assumption that \( \psi \) is an extremizer,
\[
A_0 \geq \lim_{n \to \infty} \|\phi_n\|_{L_2^2}^{-1} \|\nabla e^{it_n \Delta^2} \phi_n\|_{L_{t,x}^{2(d+2)}}
\]
\[
= \lim_{n \to \infty} \|e^{it_n \Delta^2} g_n(\xi_n \psi \circ \ell_{\xi_n}^{-1})\|_{L_2^2}^{-1}
\]
\[
\times \|\nabla e^{i(t_n + t)\Delta^2} g_n(\xi_n \psi \circ \ell_{\xi_n}^{-1})\|_{L_{t,x}^{2(d+2)}}
\]
\[
= 3^{-\frac{1}{2}} 2^{-\frac{d}{2}} \|\phi\|_{L_2^2}^{-1} \lim_{n \to \infty} \|\nabla e^{it_n \Delta^2} g_n(\xi_n \psi \circ \ell_{\xi_n}^{-1})\|_{L_{t,x}^{2(d+2)}}
\]
\[
= 3^{-\frac{1}{2}} 2^{-\frac{d}{2}} \|\phi\|_{L_2^2}^{-1} \lim_{n \to \infty} \|\nabla g_n(\xi_n \psi \circ \ell_{\xi_n}^{-1})\|_{L_{t,x}^{2(d+2)}}
\]
\[
= 3^{-1} 2^{-d} \|\phi\|_{L_2^2}^{-1} \|\psi\|_{L_2^2}^{-1} \|e^{-it_n \Delta} \psi\|_{L_{t,x}^{2(d+2)}} = 3^{-1} 2^{-d} \|\phi\|_{L_2^2}^{-1} \|\psi\|_{L_2^2}^{-1} \|e^{-it_n \Delta} \psi\|_{L_{t,x}^{2(d+2)}}
\] (4.12)

This verifies (4.3). Conversely, if equality holds in (4.3), then it holds everywhere in the computation above, establishing the final conclusion of the theorem.

In the other direction, let \( \{\phi_n\} \) be an \( L_2^2 \)-normalized extremizing sequence for \( A_0 \). By Theorem 3.1, there exist sequences \( \{\phi_j\}_{j \geq 1} \), \( \{w_n^j\}_{j \geq 1, n \geq 1} \), and parameters \( (h_j^t, \xi_j, t_n^j)_{j \geq 1, n \geq 1} \) such that for each \( j \), \( |h_j^t \xi_j| \to \infty \) or \( h_j^t \xi_j \equiv 0 \) and such that after passing to a subsequence,
\[
\phi_n = \sum_{j=1}^J e^{it_n^j \Delta^2} g_n(\xi_n \psi \circ \ell_n^j) + w_n^j,
\]
where (3.2), (3.4), and (3.6) hold.

Therefore,
\[
A_0^{2(d+2)} - \lim_{n \to \infty} \frac{2(d+2)}{\sum_{j=1}^{2(d+2)}} \left\| \nabla \left[ \frac{d}{d^2} e^{it\Delta} \phi_n \right] \right\|_{L^2_{t,x}}^{2(d+2)} \\
\leq \lim_{l \to \infty} \lim_{n \to \infty} \sum_{1 \leq j \leq t} \left\| \nabla \left[ \frac{d}{d^2} e^{i(t+t_n)\Delta} \phi_n \right] \right\|_{L^2_{t,x}}^{2(d+2)} \\
\leq \lim_{l \to \infty} \lim_{n \to \infty} \sum_{1 \leq j \leq t} A_0^{2(d+2)} \left\| \nabla \left[ \frac{d}{d^2} e^{i(t+t_n)\Delta} \phi_n \right] \right\|_{L^2_{t,x}}^{2(d+2)} \\
\leq \sum_{j=1}^{2(d+2)} \left( \sup \left\| \phi_j \right\|_{L^2_{t,x}} \right) \sum_{j} \left\| \phi_j \right\|_{L^2_{t,x}}^{\frac{1}{2}}.
\]

By (3.4), the right hand side is strictly less than the left hand side (a contradiction) unless there exists \(j\) such that \(\left\| \phi_j \right\|_{L^2_{t,x}} = 1\). In this case, there is only one profile and the error terms tend to zero in \(L^2_{t,x}\):
\[
\phi_n = e^{it_n\Delta} g_n(e^{i(h_n\xi_n)} \phi_n) + w_n, \quad \lim_{n \to \infty} \left\| w_n \right\|_{L^2_{x}} = 0. \tag{4.13}
\]

If \(h_n\xi_n = 0\), then since
\[
\left\| e^{it_n\Delta} g_n(\phi) \right\|_{L^2_{x}} = \left\| \phi \right\|_{L^2_{x}},
\]
\[
\left\| \nabla \frac{d}{d^2} e^{i(t_n+\xi_n)\Delta} g_n(\phi) \right\|_{L^2_{t,x}} = \left\| \nabla \frac{d}{d^2} e^{it\Delta} \phi \right\|_{L^2_{t,x}},
\]
\(\phi\) is an extremizer for \(A_0\). Thus if \(A_0\) does not have an extremizer, every \(L^2_{t,x}\)-normalized extremizing sequence must satisfy (after passing to a subsequence)
\[
\left\| \phi_n - e^{it_n\Delta} g_n(e^{i(h_n\xi_n)} \phi) \right\|_{L^2_{x}} \to 0, \tag{4.14}
\]
for some function \(\phi \in L^2_{t,x}\) and parameters \((h_n, \xi_n, t_n, x_n)\) with \(|h_n\xi_n| \to \infty\). By the essentially the same computation as (4.12), this implies that
\[
A_0 \leq 3^{-\frac{d+1}{2} - \frac{d+1}{2}} B,
\]
and hence that equality holds in (4.3).

Thus it remains to show that if \(\{\phi_n\}\) is an \(L^2_{t,x}\)-normalized extremizing sequence for \(A_0\), (4.3) holds with equality, and (4.14) holds for some \(\phi\) and \((h_n, t_n, x_n, \xi_n)\) with \(|h_n\xi_n| \to \infty\), then (4.4) holds with \(\psi\) an extremizer for \(B\).

Passing to a further subsequence, there exists \(\omega \in S^{d-1}\) such that \(h_n\xi_n \to \omega\). Let \(\psi = \phi \circ \ell_{\omega}\). For \(a \neq 0\), \(\ell_{\omega}\) depends only on \(\frac{\omega}{|\omega|}\), so \(\phi - \psi \circ \ell_{-1}^{h_n\xi_n} \to 0\) in \(L^2_{t,x}\).

Therefore
\[
\left\| e^{it_n\Delta} g_n(e^{i(h_n\xi_n)} \phi) - e^{it_n\Delta} g_n(e^{i(h_n\xi_n)} (\psi \circ \ell_{-1}^{h_n\xi_n})) \right\|_{L^2_{x}} \to 0,
\]
which implies by (4.13) that
\[
\left\| \phi_n - e^{it_n\Delta} g_n(e^{i(h_n\xi_n)} (\psi \circ \ell_{-1}^{h_n\xi_n})) \right\|_{L^2_{x}} \to 0.
\]
That \(\psi\) is an extremizer for \(A_0\) follows from the same computations as in (4.12). This completes the proof of Theorem 4.1. \(\Box\)
Adapting the argument for Theorem 4.2. There are two relatively minor differences in the proof of Theorem 4.2. First, $A_1$ must be compared to two operator norms, $A_0$ and $B$. To obtain the estimate $A_1 \geq B$, we simply take an extremizer for $B$ and use (4.9), arguing similarly to (4.12). Given $\varepsilon > 0$, we can show that $A_1 \geq (1 - \varepsilon)A_0$ by selecting an $L^2_\varepsilon$-normalized function $\psi$ satisfying $\| \nabla \frac{\partial^2}{\partial x^2} e^{i t \Delta} \psi \| \geq (1 - \varepsilon)A_0$ and using (4.8); letting $\varepsilon \to 0$, we see that $A_1 \geq A_0$.

Second, we must rule out the case in which (4.6) holds for some $\psi \in L^2_\varepsilon$ and some sequence of parameters $(h_n, \xi_n, t_n, x_n)$ with $\xi_n \neq 0$. Passing to a subsequence and using the fact that spacetime translations do not affect any of the relevant operator norms, it suffices to consider the cases when $(t_n, x_n) \equiv (0, 0), [h_n \xi_n] \to \infty$, and either $\xi_n \to \xi_0 \neq 0$ or $|\xi_n| \to \infty$. If $|\xi_n| \to \infty$, we apply (4.11) and compute

$$
\| \xi_n \|_{L^2_\varepsilon} e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} [g_n(\phi \circ \ell_n)](\xi_n^{-1} (x + 2 t h_n^2 |\xi_n|^2 + 2 t)) \|_{L^2_{t,x}}^{\frac{d+2}{2(d+2)}} \leq 3^{-\frac{d}{2(d+2)}} 2^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon} < B,
$$

provided $\xi_n \neq 0$, $\ell_n = \xi \xi_n$ is as in (4.5), and $\| \phi \|_{L^2_\varepsilon} \leq 1$. If $\xi_n \to \xi_0 \neq 0$ and $\| \phi \|_{L^2_\varepsilon} \leq 1$, we use (4.10) and compute

$$
\limsup_{n \to \infty} \| \xi_n \|_{L^2_\varepsilon} e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} \leq \limsup_{n \to \infty} (|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (6|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (2|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon} < B,
$$

so this case can be ruled out as well. \hfill \square

Finally, we prove the lemmas.

Proof of Lemmas 4.3 and 4.4. We begin by observing that by the change of variables formula and the Strichartz inequality for Schrödinger,

$$
\| [a_n] e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} (\phi \circ \ell_n)(\xi_n^{-1} (x + 2 t |a_n|^2 a_n)) \|_{L^2_{t,x}}^{\frac{d}{2(d+2)}} = 3^{-\frac{d}{2(d+2)}} 2^{-\frac{d}{2(d+2)}} \| e^{-i t \Delta} \phi \circ \ell_n \|_{L^2_{t,x}}^{\frac{d}{2(d+2)}} \leq 3^{-\frac{d}{2(d+2)}} 2^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon}.
$$

Similar computations give:

$$
\| \nabla \frac{\partial^2}{\partial x^2} e^{i t \Delta} g_n \phi \|_{L^2_{t,x}}^{\frac{d}{2(d+2)}} \leq A_0 \| \phi \|_{L^2_\varepsilon}, \quad \| e^{-i t \Delta} g_n \phi \|_{L^2_{t,x}}^{\frac{d}{2(d+2)}} = B \| \phi \|_{L^2_\varepsilon},
$$

$$
\| [\xi_n] e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} \leq (|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (6|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (2|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon},
$$

$$
\| [\xi_n] e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} \leq (|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (6|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (2|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon},
$$

$$
\| \xi_n \|_{L^2_\varepsilon} e^{i (x \xi_n + t h_n^2 |\xi_n|^4 + t |\xi_n|^2)} e^{-i |\xi_n|^2 t \Delta} \leq (|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (6|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} (2|\xi_n|^2 + 1)^{-\frac{d}{2(d+2)}} B \| \phi \|_{L^2_\varepsilon}.
$$
Since the function \(e^{i\xi_n} \phi\) when \(|\xi| \sim a_n\) is in some dense subset of \(L_x^2\), so by the fundamental theorem of calculus,

\[
\left| \frac{a_n}{|\xi|} \right| - 1 \sim |a_n| - \frac{|\xi|}{\frac{a_n}{|\xi|}} \leq |a_n| - \frac{|\xi|}{\frac{a_n}{|\xi|}} a_n - \frac{|\xi|}{\frac{a_n}{|\xi|}} R.
\]

In addition, by the Strichartz inequality (2.3) for 4th order Schrödinger, the operator

\[
\phi \mapsto D_{\mu}^{\frac{a_n}{|\xi|}} S_\mu(t) g_n(e^{i\xi_n} \phi)
\]

is also uniformly bounded from \(L_x^2\) to \(L_{t,x}^{2(d+2)}\), so it suffices to prove the lemmas when \(\phi\) is in some dense subset of \(L_x^2\). Thus we may assume that \(\phi\) is a Schwartz function with compact frequency support that does not contain 0:

\[
\text{Supp} \hat{\phi} \subseteq \{R^{-1} \leq |\xi| \leq R\}.
\]

Under the hypotheses of Lemma 4.3, we assume \(|a_n| \geq 2R\) and \(|\xi - a_n| \leq R\). Then \(|\xi| \sim |a_n|\), so by the fundamental theorem of calculus,

\[
\left| \frac{a_n}{|\xi|} \right| - 1 \sim |a_n| - \frac{|\xi|}{\frac{a_n}{|\xi|}} \leq |a_n| - |\xi| - a_n| - \frac{|\xi|}{\frac{a_n}{|\xi|}} \frac{a_n}{|\xi|} R.
\]

Therefore, since \(D_{\mu}^{\frac{a_n}{|\xi|}} S_\mu(t)\) is a bounded operator from \(L_x^2\) to \(L_{t,x}^{2(d+2)}\), (4.7) would follow from

\[
|a_n| \frac{d}{d\xi} \left|e^{i\xi_n} \phi\right|_2 \left[\left|e^{i\xi_n} \phi\right|_2 - e^{i(x-a_n t) |a_n|^2} e^{-it\Delta} (\phi \circ \ell_{a_n}) (e_n^{-1}(x + 4t|a_n|^2 a_n))\right] \to 0.
\]

Changing variables in \(t\), the left hand side of (4.15) equals

\[
\left|e^{i\xi_n} \phi\right|_2 - e^{i(x-a_n t) |a_n|^2} e^{-it\Delta} (\phi \circ \ell_{a_n}) (e_n^{-1}(x + 4t|a_n|^2 a_n)) \to 0.
\]

Next, we compute

\[
e^{i|x-a_n|^2 |a_n|^2} e^{i\xi_n} \phi(x) = \int e^{i(x+\xi_n \xi)} e^{i\xi_n \xi} \phi\left(\frac{x - a_n}{\xi}\right) d\xi,
\]

\[
= \int e^{i(x+\xi_n \xi)} e^{i\xi_n \xi} e^{i\xi_n \xi} \phi\left(\frac{x - a_n}{\xi}\right) d\xi
\]

\[
= e^{i(x-a_n t) |a_n|^2} \int e^{i(x+4t|a_n|^2)} e^{it\xi_n} e^{it\xi_n} \phi\left(\frac{x - a_n}{\xi}\right) d\xi.
\]

Thus (4.15) would follow from

\[
\left|e^{i\xi_n} \phi\right|_2 - e^{i(x-a_n t) |a_n|^2} e^{-it\Delta} (\phi \circ \ell_{a_n}) (e_n^{-1}(x)) \left|e^{i(x-a_n t) |a_n|^2} e^{-it\Delta} (\phi \circ \ell_{a_n}) (e_n^{-1}(x)) \right|_2 \to 0.
\]

Since

\[
(e^{-it\Delta} (\phi \circ \ell_{a_n})) \circ \ell_{a_n}^{-1} = e^{it\xi_n (\xi_n)} \phi,
\]

(4.17)
the limit in (4.16) equals zero if and only if
\[
\lim_{n \to \infty} \|e^{it(\frac{\Delta^2}{|a_n|^2} + \frac{4i\Delta \nabla \phi_n}{|a_n|^2})} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} = 0, \quad \text{if } |a_n| \to \infty. \quad (4.18)
\]

Similar computations show that (4.8), (4.9), (4.10), and (4.11) would (respectively) follow from
\[
\lim_{n \to \infty} \|e^{it\Delta^2} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} = 0, \quad \text{if } h_n \to 0 \quad (4.19)
\]
\[
\lim_{n \to \infty} \|e^{it\Delta^2} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} = 0, \quad \text{if } h_n \to \infty \quad (4.20)
\]
\[
\lim_{n \to \infty} \|e^{it(\frac{\Delta^2}{|a_n|^2} + \frac{4i\Delta \nabla \phi_n}{|a_n|^2})} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} = 0, \quad \text{if } |\xi_n| \lesssim 1, h_n \to \infty \quad (4.21)
\]
\[
\lim_{n \to \infty} \|e^{it(\frac{\Delta^2}{|a_n|^2} + \frac{4i\Delta \nabla \phi_n}{|a_n|^2})} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} = 0, \quad \text{if } |\xi_n|, |h_n\xi_n| \to \infty. \quad (4.22)
\]

Since \(\hat{\phi}\) is smooth with compact support,
\[
\|e^{it(\frac{\Delta^2}{|a_n|^2} + \frac{4i\Delta \nabla \phi_n}{|a_n|^2})} - 1\|_{\mathcal{L}_{t,x}^{2(\frac{d+2}{2})}} \to 0
\]
pointwise in \(t,x\), so (4.18) will follow from the dominated convergence theorem if we show that there exists a function that dominates each term in the sequence.

Let
\[
\Phi_n(t,x,\xi) = t(\frac{|\xi|^4}{|a_n|^2} + \frac{4|\xi|^2a_n}{|a_n|^2} + \frac{\mu|\xi|^2}{|a_n|^2} + |\ell_{a_n}(\xi)|^2) + x\xi
\]
\[
\Psi_n(t,x,\xi) = t|\ell_{a_n}(\xi)|^2 + x\xi.
\]

Then the left hand side of (4.18) is the \(L_{t,x}^{2(\frac{d+2}{2})}\) norm of
\[
(t,x) \mapsto \int (e^{i\Phi_n(t,x,\xi)} - e^{i\Psi_n(t,x,\xi)})\hat{\phi}(\xi) \, d\xi.
\]

Since \(\hat{\phi}\) is smooth with compact support, this quantity is uniformly bounded. The gradients of the phases are
\[
\nabla \Phi_n(t,x,\xi) = t\left(\frac{4|\xi|^2\xi}{|a_n|^2} + \frac{8\xi a_n \xi}{|a_n|^2} + \frac{4|\xi|^2 a_n}{|a_n|^2} + \frac{2\mu \xi^2}{|a_n|^2} + 2\ell_{a_n} \circ \ell_{a_n}(\xi) + x\right)
\]
\[
\nabla \Psi_n(t,x,\xi) = 2\ell_{a_n} \circ \ell_{a_n}(\xi) + x,
\]
and for \(|a_n| \gg (1 + R)^2\) and \(|x| > 100Rt\), these are nonvanishing for \(\xi \in \text{Supp} \hat{\phi} \subset \{|\xi| \leq R\}\). In particular,
\[
|\nabla \Phi_n(t,x,\xi)|, |\nabla \Psi_n(t,x,\xi)| \gtrsim |x|, \quad |x| > 100Rt. \quad (4.23)
\]

Furthermore, the Hessian matrices of the phases satisfy
\[
D^2_{\xi} \Phi_n(t,x,\xi), D^2_{\xi} \Psi_n(t,x,\xi) = tO_{\xi,a_n}(\frac{R^2}{|a_n|^2}) + 2\ell_{a_n}^2,
\]
where \(O_{\xi,a_n}(\frac{R^2}{|a_n|^2})\) is a matrix whose coefficients are uniformly bounded by \(\frac{R^2}{|a_n|^2}\), and we are identifying the transformation \(\ell_{a_n}\) with its matrix. Since \(2\ell_{a_n}^2\) is uniformly positive definite (indeed, its eigenvalues are 12, 4, ... , 4), for \(|a_n|\) sufficiently
large (depending on $R$), the critical points of the phases must be nondegenerate. Thus by (4.23), (4.24), and the principle of stationary phase (cf. [28, Ch. VIII]),

$$\left| \int \left[ e^{i\Phi_n(t,x,\xi)} - e^{i\Psi_n(t,x,\xi)} \right] \phi(\xi) \, d\xi \right| \lesssim \left( \frac{1}{1+|t|} \right)^N \chi_{\{|x| \gg t\}} + \left( \frac{1}{1+|t|} \right)^{\frac{d}{2}} \chi_{\{|x| \lesssim t\}} \lesssim \left( \frac{1}{1+|t|+|x|} \right)^{\frac{d}{2}}. $$

The right hand side of the above inequality is in $L^2_{t,x} \chi_{\{|x| \gg t\}}$, so (4.18) does indeed follow by the dominated convergence theorem.

The limits in (4.19), (4.20), (4.21), and (4.22) may be verified in a similar manner. (For (4.19) and (4.21), one also uses that $0 \notin \text{Supp} \hat{\phi}$.) This completes the proof of Lemmas 4.3 and 4.4. $\square$

References


