SHARP BOUNDS ON SPECTRAL CLUSTERS FOR LIPSCHITZ METRICS

HERBERT KOCH, HART F. SMITH, AND DANIEL TATARU

ABSTRACT. We establish L^p bounds on L^2 normalized spectral clusters for self-adjoint elliptic Dirichlet forms with Lipschitz coefficients. In two dimensions we obtain best possible bounds for all $2 \le p \le \infty$, up to logarithmic losses for $6 < p \le 8$. In higher dimensions we obtain best possible bounds for a limited range of p.

1. INTRODUCTION

Let M be a compact, 2-dimensional manifold without boundary, on which we fix a smooth volume form dx. Let g be a section of positive definite symmetric quadratic forms on $T^*(M)$, and let ρ be a strictly positive function on M.

Consider the eigenfunction problem

$$div\,(\mathrm{g}\, d\phi) + \lambda^2 \rho \phi = 0\,,$$

where div, which maps vector fields to functions, is the dual of d under dx. This setup includes as a special case eigenfunctions of the Laplace-Beltrami operator. We refer to the real parameter λ as the frequency of ϕ, and take $\lambda \ge 0$. Under the condition that g and ρ are measurable and bounded from above and below, there exists a complete orthonormal basis $\{\phi_j\}_{j=1}^{\infty}$ of eigenfunctions for $L^2(M, \rho\, dx)$, with frequencies satisfying $\lambda_j \to \infty$.

In this paper we prove the following, where for convenience of the statement we take $\lambda \ge 2$. Except for the factor $(\log \lambda)^\sigma$ in (1.2), these bounds are the best possible for general Lipschitz g and ρ, in terms of the growth in λ, by examples of [8].

Theorem 1.1. *Suppose that* $\mathrm{g}, \rho \in \mathrm{Lip}(M)$. *Assume that the frequencies of* u *are contained in the interval* $[\lambda, \lambda+1]$, *so that*

$$u = \sum_{j:\lambda_j \in [\lambda,\lambda+1]} c_j\, \phi_j\,.$$

Then

$$\|u\|_{L^p(M)} \le C_p\, \lambda^{\frac12 - \frac2p} \|u\|_{L^2(M)}\,, \quad 8 < p \le \infty\,, \tag{1.1}$$

and

$$\|u\|_{L^p(M)} \le C \lambda^{\frac23(\frac12 - \frac1p)} (\log \lambda)^\sigma \|u\|_{L^2(M)}\,, \quad 6 \le p \le 8\,, \tag{1.2}$$

where $\sigma = \frac32$ *for* $p = 8$, *and* $\sigma = 0$ *for* $p = 6$.

The authors were supported in part by NSF grant DMS-0654415.

For $2 \le p \le 6$, the bound (1.2) was established, without logarithmic loss, in [7],

$$\|u\|_{L^p(M)} \le C\lambda^{\frac{2}{3}(\frac{1}{2}-\frac{1}{p})}\|u\|_{L^2(M)}\,, \quad 2 \le p \le 6\,.$$

To put the above estimates in context, we recall the previously known results. In case g and ρ are C^∞, Sogge [10] established the following bounds in general dimensions $d \ge 2$,

$$\begin{aligned} &\|u\|_{L^p(M)} \le C\,\lambda^{d(\frac{1}{2}-\frac{1}{p})-\frac{1}{2}}\,\|f\|_{L^2(M)}\,, \qquad \tfrac{2(d+1)}{d-1} \le p \le \infty\,, \\ &\|u\|_{L^p(M)} \le C\,\lambda^{\frac{d-1}{2}(\frac{1}{2}-\frac{1}{p})}\,\|f\|_{L^2(M)}\,, \qquad 2 \le p \le \tfrac{2(d+1)}{d-1}\,, \end{aligned} \tag{1.3}$$

which are best possible at all p for unit width spectral clusters. Semiclassical generalizations were obtained by Koch-Tataru-Zworski in [5].

The estimates (1.3) were extended to $C^{1,1}$ coefficients in [6]. On the other hand, the examples of [8] show that for small p they fail for coefficients of lower Hölder regularity. In particular, for Lipschitz coefficients the following would in general be best possible

$$\begin{aligned} &\|u\|_{L^p(M)} \le C\,\lambda^{d(\frac{1}{2}-\frac{1}{p})-\frac{1}{2}}\,\|f\|_{L^2(M)}\,, \qquad \tfrac{2(d+2)}{d-1} \le p \le \infty\,, \\ &\|u\|_{L^p(M)} \le C\,\lambda^{\frac{2(d-1)}{3}(\frac{1}{2}-\frac{1}{p})}\,\|f\|_{L^2(M)}\,, \qquad 2 \le p \le \tfrac{2(d+2)}{d-1}\,. \end{aligned} \tag{1.4}$$

Except for the factors of $\log\lambda$ for $6 < p \le 8$, we have thus obtained the sharp results in two dimensions. We also establish in this paper partial results for dimensions $d \ge 3$. Precisely, in the last section of this paper we prove the first estimate in (1.4) for $\frac{6d-2}{d-1} < p \le \infty$. The second estimate in (1.4) was established in [7] on the range $2 \le p \le \frac{2(d+1)}{d-1}$. That proof proceeded by establishing the no-loss estimate (1.3) on sets of diameter $\lambda^{-\frac{1}{3}}$, using methods analogous to the $\lambda^{-\frac{1}{3}}$ time scale representation of u in this paper. This scaling had been used in [11] to prove Strichartz estimates with loss for wave equations with Lipschitz coefficients, and the examples showing optimality of the estimates were constructed in [9].

For most of this paper we focus for simplicity on the case of two dimensions, where our results are strongest. We start in section 2 with the reduction of Theorem 1.1 to two key propositions, one involving short time dispersive estimates and the other long time energy overlap bounds. The short time estimates, Proposition 2.5, are established in sections 4 and 6, through a combination of Strichartz and bilinear estimates. The energy overlap bounds, Proposition 2.4, are established in sections 3 and 5, and depend on energy propagation estimates for Lipschitz metrics. The weak localization of these energy propagation estimates is the main reason for the limitation on our results in higher dimensions, which we establish in section 7.

2. The argument

In this section, we reduce the proof of Theorem 1.1 to the two key results of this paper, Propositions 2.4 and 2.5. Since the estimate (1.2) follows by interpolation from the case $p = 8$ and the known case $p = 6$, we consider $8 \le p \le \infty$. To avoid unnecessary repetition, we focus on the case $p = 8$ in the following steps. At the end of this section we then show how to deduce the estimate (1.1) for $p > 8$ by a simple interpolation argument.

The spectral localization of u, integration by parts, elliptic regularity, and the equation, yield the following over M

$$\lambda^{-1}\| div\,(\mathrm{g}\,du) + \lambda^2 \rho u\|_{L^2} + \lambda^{-1}\|du\|_{L^2} + \lambda^{-2}\|d^2 u\|_{L^2} + \lambda\|u\|_{H^{-1}} \lesssim \|u\|_{L^2}\,. \tag{2.1}$$

By choosing a partition of unity subordinate to suitable local coordinates, for $p=8$ we are then reduced to the following.

Theorem 2.1. *Suppose that* g *and* ρ *are globally defined on* $\mathbb{R}^2$*, with*

$$\|\mathrm{g}^{ij} - \delta^{ij}\|_{Lip(\mathbb{R}^2)} + \|\rho - 1\|_{Lip(\mathbb{R}^2)} \le c_0 \ll 1\,.$$

Then the following estimate holds for functions u *supported in the unit cube of* $\mathbb{R}^2$*,*

$$\|u\|_{L^8(\mathbb{R}^2)} \lesssim \lambda^{\frac14}(\log\lambda)^{\frac32}\bigl(\|u\|_{L^2(\mathbb{R}^2)} + \lambda^{-1}\|(div\,(\mathrm{g}\,du) + \lambda^2\rho u\|_{L^2(\mathbb{R}^2)}\bigr)\,. \tag{2.2}$$

Step 1: *Reduction to a frequency localized first order problem.* In proving Theorem 2.1 we may replace the function g by g_λ, where g_λ is obtained by smoothly truncating $\hat{\mathrm{g}}(\xi)$ to $|\xi| \le c\lambda$, c some fixed small constant. Since

$$\|\mathrm{g} - \mathrm{g}_\lambda\|_{L^\infty(\mathbb{R}^2)} \lesssim \lambda^{-1}\,, \qquad \|\nabla(\mathrm{g} - \mathrm{g}_\lambda)\|_{L^\infty(\mathbb{R}^2)} \lesssim 1\,,$$

the right hand side of (2.2) is comparable to the same quantity after this replacement. Similarly we replace ρ by ρ_λ.

By the Coifman-Meyer commutator theorem [2] (see also [12, Prop. 3.6.B]), the commutator of g_λ or ρ_λ with a multiplier $\Gamma(D)$ of type S^0 maps $H^s \to H^{s+1}$ for $-1 \le s \le 0$, so we may take a conic partition of unity to reduce matters to establishing (2.2) with u replaced by $\Gamma(D)u$, with $\Gamma(\xi)$ supported where $|\xi_2| \le c\,\xi_1$. This step loses compact support of u, but we may still take the L^8 norm over the unit cube. Finally, arguments as in [7, Corollary 5] reduce matters to considering $\hat u(\xi)$ supported where $|\xi| \approx \lambda$.

We now label $x_1 = t$, and $x_2 = x$, and let (τ,ξ) be the dual variables to (t,x). Thus, with c above small, $\hat u(\tau,\xi)$ is supported where $\{|\xi| \le \frac12\lambda\,, \tau \approx \lambda\}$. For $|\xi| \le \frac34\lambda$ we can factor

$$-\mathrm{g}_\lambda(t,x)\cdot(\tau,\xi)^2 + \lambda^2\rho_\lambda(t,x) = -\mathrm{g}_\lambda^{00}(t,x)\bigl(\tau + \tilde a(t,x,\xi,\lambda)\bigr)\bigl(\tau - a(t,x,\xi,\lambda)\bigr)\,,$$

where $\tilde a\,, a > 0$, and both belong to $\lambda C^1 S_{\lambda,\lambda}$ according to the following definition.

Definition 2.2. *We say* $b(t,x,\xi,\lambda) \in S_{\lambda,\lambda^\delta}$ *if, for all multi-indices* α,β*,*

$$|\partial^\alpha_{t,x}\partial^\beta_\xi b(t,x,\xi,\lambda)| \le C_{\alpha,\beta}\lambda^{-|\beta| + \delta|\alpha|}\,.$$

We say $b(t,x,\xi,\lambda) \in C^1 S_{\lambda,\lambda^\delta}$ *if, in addition, for* $|\alpha| \ge 1$ *the stronger estimate holds*

$$|\partial^\alpha_{t,x}\partial^\beta_\xi b(t,x,\xi,\lambda)| \le C_{\alpha,\beta}\lambda^{-|\beta| + \delta(|\alpha|-1)}\,.$$

We write $b \in \lambda^m S_{\lambda,\lambda^\delta}$ *to indicate* $\lambda^{-m} b \in S_{\lambda,\lambda^\delta}$*.*

In our applications either $\delta = 1$ or $\delta = \frac23$, and we suppress the explicit λ when writing b. The symbols $b \in S_{\lambda,\lambda}$ are simply a bounded family of $S^0_{0,0}(\mathbb{R}^2,\mathbb{R})$ symbols rescaled by $(t,x,\xi) \to (\lambda t, \lambda x, \lambda^{-1}\xi)$, so that $L^2(\mathbb{R})$ boundedness of $b(t,x,D)$, as well as the Weyl quantization $b^w(t,x,D)$, follows (with uniform bounds in t) by [3]. For symbols in $C^1 S_{\lambda,\lambda^\delta}$, the asymptotic laws for composition and adjoint hold to first order. In particular, if $a \in C^1 S_{\lambda,\lambda^\delta}$ then

$$a^w(t,x,D) = a(t,x,D) + r(t,x,D)\,, \quad r \in \lambda^{-1} S_{\lambda,\lambda^\delta}\,,$$

and if $a \in S_{\lambda,\lambda^\delta}$, $b \in C^1 S_{\lambda,\lambda^\delta}$, then

$$a(t,x,D)b(t,x,D) = (ab)(t,x,D) + r(t,x,D)\,, \quad r \in \lambda^{-1} S_{\lambda,\lambda^\delta}\,.$$

Let $a_\lambda(t,x,\xi)$ be obtained by smoothly truncating the (t,x)-Fourier transform of $a(\cdot\,,\cdot\,,\xi,\lambda)$ to frequencies less than $c\lambda$. Then since $a \in \lambda C^1 S_{\lambda,\lambda}$,

$$a(t,x,\xi,\lambda) - a_\lambda(t,x,\xi) \in S_{\lambda,\lambda}\,.$$

The symbol $a_\lambda(t,x,\xi)$ inherits from $a(t,x,\xi,\lambda)$ the estimates, for $|\xi| \le \frac{3}{4}\lambda$,

$$a_\lambda(t,x,\xi) \approx \lambda\,, \qquad \partial_\xi^2 a_\lambda(t,x,\xi) \approx -\lambda^{-1}\,. \tag{2.3}$$

We now extend a_λ globally in ξ so that $a_\lambda(t,x,\xi) = \lambda$ for $|\xi| \ge (\frac{3}{4} - c)\lambda$, and a_λ is unchanged for $|\xi| \le \frac{5}{8}\lambda$, while maintaining the frequency localization in the (t,x) variables. In particular, (2.3) holds for $|\xi| \le \frac{5}{8}\lambda$. A similar observation holds for $\tilde a$.

By the above, we have the factorization over $|\xi| \le \frac{5}{8}\lambda$,

$$div\, \mathrm{g}_\lambda\, d + \lambda^2 \rho_\lambda = \mathrm{g}_\lambda^{00}\big(D_t + \tilde a_\lambda^w(t,x,D)\big)\big(D_t - a_\lambda^w(t,x,D)\big) + r(t,x,D)\,,$$

where $r \in \lambda S_{\lambda,\lambda}$. Since $a_\lambda^w(t,x,D)u$ is supported where $\tau \approx \lambda$, and $D_t + a_\lambda(t,x,D)$ admits a parametrix in $\lambda^{-1} S_{\lambda,\lambda}$ there, we have thus reduced Theorem 2.1 to establishing the following estimate over $(t,x) \in [0,1] \times \mathbb{R}$,

$$\|u\|_{L^8([0,1]\times\mathbb{R})} \lesssim \lambda^{\frac14} (\log\lambda)^{\frac32} \big(\|u\|_{L^2([0,1]\times\mathbb{R})} + \|(D_t - a_\lambda^w(t,x,D))u\|_{L^2([0,1]\times\mathbb{R})}\big)\,.$$

We denote by $S(t,s)$ the evolution operators for $a_\lambda^w(t,x,D)$, which are unitary on $L^2(\mathbb{R})$. Precisely, $u(t,x) = S(t,t_0)f$ satisfies

$$\big(D_t - a_\lambda^w(t,x,D)\big)u = 0\,, \qquad u(t_0,\cdot\,) = f\,. \tag{2.4}$$

If $\hat f$ is supported in $|\xi| \le \frac34\lambda$, then so is $\hat u(t,\cdot\,)$, since $a_\lambda = \lambda$ for $|\xi| \ge (\frac34 - c)\lambda$, and a_λ is spectrally localized in x to the $c\lambda$ ball. By the Duhamel formula, it then suffices to prove that

$$\|u\|_{L^8([0,1]\times\mathbb{R})} \lesssim \lambda^{\frac14}(\log\lambda)^{\frac32} \|u_0\|_{L^2(\mathbb{R})}\,, \qquad u = S(t,0)u_0\,,$$

with $\widehat{u}_0$ supported in $|\xi| \le \frac34\lambda$, and $a_\lambda(t,x,\xi)$ as above.

Henceforth, we will take $\|u_0\|_{L^2(\mathbb{R})} = 1$.

Step 2: *Decomposition in a wave packet frame on $\lambda^{-\frac13}$ time slices.* Let $a_{\lambda^{2/3}}(t,x,\xi)$ be obtained by smoothly truncating the (t,x)-Fourier transform of $a_\lambda(\cdot\,,\cdot\,,\xi)$, or equivalently that of $a(t,x,\xi,\lambda)$, to frequencies less than $c\lambda^{\frac23}$. Then $a_{\lambda^{2/3}} \in \lambda S_{\lambda,\lambda^{2/3}}$, and $a_{\lambda^{2/3}}$ also satisfies (2.3), for $|\xi| \le \frac58\lambda$ in case of the second estimate in (2.3).

We divide the time interval $[0,1]$ into subintervals of length $\lambda^{-\frac13}$, and thus write $[0,1]\times\mathbb{R}$ as a union of slabs $[l\lambda^{\frac13}, (l+1)\lambda^{\frac13}]\times\mathbb{R}$. Within each such slab we will consider an expansion of u in terms of homogeneous solutions for $D_t - a^w_{\lambda^{2/3}}(t,x,D)$, We refer to the homogeneous solutions on each $\lambda^{-\frac13}$ time interval as *tube solutions*, since they will be highly localized to a collection of tubes T.

The collection of tubes is indexed by triples of integers $T = (l,m,n)$, with $0 \le l \le \lambda^{\frac13}$ referencing the slab $[l\lambda^{\frac13}, (l+1)\lambda^{\frac13}]\times\mathbb{R}$. We will describe the construction for the slab $[0,\lambda^{-\frac13}]\times\mathbb{R}$; the tube solutions supported on the other slabs are obtained in an identical manner. Thus, T is here identified with a pair $(m,n) \in \mathbb{Z}^2$.

We start with a $\lambda^{\frac{2}{3}}$-scaled Gabor frame on $\mathbb{R}$, with compact frequency support. That is, we select a Schwartz function ϕ, with $\hat\phi$ supported in $|\xi|\le\frac{9}{8}$, such that, with $x_T=\frac12\lambda^{-\frac23}m$ and $\xi_T=\lambda^{\frac23}n$, the space-frequency translates

$$\phi_T(x)=\lambda^{\frac13}e^{ix\xi_T}\phi(\lambda^{\frac23}(x-x_T))$$

form a tight frame, in that for all $f\in L^2(\mathbb{R})$,

$$f=\sum_{m,n}c_T\phi_T\,,\qquad c_T=\int\overline{\phi_T(x)}f(x)\,dx\,,$$

from which it follows that

$$\|f\|^2_{L^2(\mathbb{R})}\approx\sum_T|c_T|^2\,.$$

Since the function f in our application will be frequency restricted, the index ξ_T will only run over $|n|\le\lambda^{\frac13}$, by the compact support condition on $\hat\phi$.

The frame is not orthogonal, so it is not necessarily true, for arbitrary coefficients b_T, that $\|\sum b_T\phi_T\|^2_{L^2(\mathbb{R})}\approx\sum|b_T|^2$. However, by the bounded overlap condition on $\widehat{\phi_T}$, this does hold if the sum is over a collection of T for which the corresponding ξ_T are distinct,

$$\Big\|\sum_{T\in\Lambda}b_T\phi_T\Big\|^2_{L^2(\mathbb{R})}\approx\sum_{T\in\Lambda}|b_T|^2\quad\text{if}\quad\xi_T\ne\xi_{T'}\quad\text{when}\quad T\ne T'\in\Lambda\,.\tag{2.5}$$

Let v_T denote the solution to

$$\big(D_t-a^w_{\lambda^{2/3}}(t,x,D)\big)v_T=0\,,\qquad v_T(0,\cdot\,)=\phi_T\,.$$

We define $x_T(t)$ by

$$x_T(t)=x_T-t\,\partial_\xi a_{\lambda^{2/3}}(0,x_T,\xi_T)\,.$$

By Theorem 5.5, for $t\in[0,\lambda^{\frac13}]$ the function $v_T(t,\cdot)$ is a $\lambda^{\frac23}$-scaled Schwartz function with frequency center ξ_T, and spatial center $x_T(t)$, where the envelope function satisfies uniform Schwartz bounds over t. Thus, v_T is localized, to infinite order in x, to the following tube, which we also refer to as T,

$$T=\{(t,x):|x-x_T(t)|\le\lambda^{-\frac23}\,,t\in[0,\lambda^{-\frac13}]\}\,.$$

Since the $a^w_{\lambda^{2/3}}(t,x,D)$ flow is unitary, for each t the functions $\{v_T(t,\cdot)\}$ form a tight frame on $L^2(\mathbb{R})$. On each $\lambda^{-\frac13}$ time slab we can thus expand

$$u(t,x)=\sum_T c_T(t)\,v_T(t,x)\,,\qquad c_T(t)=\int\overline{v_T(t,x)}\,u(t,x)\,dx\,.$$

Differentiating the equation, we see that

$$c_T'(t)=i\int\overline{v_T(t,x)}\,\big(a^w_\lambda(t,x,D)-a^w_{\lambda^{2/3}}(t,x,D)\big)u(t,x)\,dx\,.$$

Since $a_\lambda(t,x,\xi)-a_{\lambda^{2/3}}(t,x,\xi)\in\lambda^{\frac13}S_{\lambda,\lambda}$, and $\|u_0\|_{L^2}=1$, we then have uniformly for $t\in[0,\lambda^{-\frac13}]$,

$$\sum_{m,n}|c_T(t)|^2\lesssim1\,,\qquad\sum_{m,n}|c_T'(t)|^2\lesssim\lambda^{\frac23}\,,$$

which together imply the following bounds, that will then hold uniformly on each $\lambda^{-\frac{1}{3}}$ time slab,

$$\sum_{T:l=l_0} \|c_T\|_{L^\infty}^2 + \lambda^{-\frac{1}{3}} \|c_T'\|_{L^2}^2 \lesssim 1\,. \tag{2.6}$$

We apply this expansion separately to the solution u on each $\lambda^{-\frac{1}{3}}$ time slab, and obtain the full tube decomposition $u = \sum_T c_T v_T$, where if $T = (l,m,n)$, the functions c_T and v_T are supported by $t \in I_T \equiv [l\lambda^{-\frac{1}{3}}, (l+1)\lambda^{-\frac{1}{3}}]$.

Step 3: *Interval decomposition according to packet size.* Here, given a coefficient c_T, we partition the time interval I_T into smaller dyadic subintervals where the coefficient c_T is essentially constant. This is done according to the following lemma.

Lemma 2.3. *Let $c : I \to \mathbb{C}$ with*

$$\|c\|_{L^\infty(I)}^2 + |I| \cdot \|c'\|_{L^2(I)}^2 = B\,.$$

Given $\epsilon > 0$, there is a partition of I into dyadic sub-intervals I_j, for each of which either

$$\|c\|_{L^\infty(I_j)} \geq 2|I_j|^{\frac{1}{2}} \|c'\|_{L^2(I_j)} \quad \text{or} \quad \|c\|_{L^\infty(I_j)} < \epsilon\,. \tag{2.7}$$

Independent of ϵ, the following bound holds

$$\sum_j |I_j|^{-1} \|c\|_{L^\infty(I_j)}^2 \leq 16B|I|^{-1}\,.$$

Proof. If the condition (2.7) holds on I then no partition is needed. Otherwise we divide the interval in half and retest. The condition automatically holds if $|I_j| \leq \epsilon^2 B^{-1}|I|/4$. The bound holds by comparing the sum to the L^2 norm of c' over the parent intervals, which have overlap at most 2. $\square$

We will take ϵ to be $\lambda^{-\frac{1}{3}}$. For each T, this gives a finite partition of its corresponding time interval I_T into dyadic subintervals,

$$I_T = \bigcup_j I_{T,j}$$

so that (2.7) holds for c_T in each subinterval,

$$\|c_T\|_{L^\infty(I_{T,j})} \geq 2|I_{T,j}|^{\frac{1}{2}} \|c_T'\|_{L^2(I_{T,j})} \quad \text{or} \quad \|c_T\|_{L^\infty(I_{T,j})} < \lambda^{-\frac{1}{3}}\,, \tag{2.8}$$

and we have the square summability relation

$$\sum_j \lambda^{-\frac{1}{3}} |I_{T,j}|^{-1} \|c_T\|_{L^\infty(I_{T,j})}^2 \lesssim \|c_T\|_{L^\infty}^2 + \lambda^{-\frac{1}{3}} \|c_T'\|_{L^2}^2\,. \tag{2.9}$$

We introduce the notation

$$c_{T,j} = 1_{T,j} c_T\,, \qquad c_{T,j}' = 1_{T,j} c_T'\,.$$

Using these interval decompositions, we partition the function u on $[0,1] \times \mathbb{R}$ into a dyadically indexed sum

$$u = \sum_{a \leq 1} \sum_{k \geq 0} u_{a,k} + u_\epsilon\,, \tag{2.10}$$

where the index a runs over dyadic values between $\epsilon = \lambda^{-\frac{1}{3}}$ and 1,

$$u_{a,k} = \sum_{(T,j)\in\mathcal{T}_{a,k}} c_{T,j} v_T \,,$$

and where

$$\mathcal{T}_{a,k} = \{(T,j) : |I_{T,j}| = 2^{-k}\lambda^{-\frac{1}{3}},\ \|c_T\|_{L^\infty(I_{T,j})} \in (a, 2a]\,\}\,.$$

We call the functions $u_{a,k}$ above (a,k)-packets, and note that by (2.8)

$$\frac{a}{4} \le |c_T(t)| \le a\,, \quad t \in I_{T,j}\,.$$

We will separately bound in L^8 each of the functions $u_{a,k}$. Since there are at most $\lambda^{\frac{1}{3}}$ tubes T over any point, and $|v_T| \lesssim \lambda^{\frac{1}{3}}$, we see that $\|u_\epsilon\|_{L^\infty} \lesssim \lambda^{\frac{1}{3}}$. On the other hand, since the decomposition (2.10) is almost orthogonal, we have $\|u_\epsilon\|_{L^2} \lesssim 1$, and hence $\|u_\epsilon\|_{L^8} \lesssim \lambda^{\frac{1}{4}}$, as desired. For $p > 8$ the bounds on u_ϵ are even better than needed.

We note here that, by (2.6) and (2.9),

$$\sum_{(T,j)\in\mathcal{T}_{a,k}} \|c_{T,j}\|_{L^2}^2 \lesssim 2^{-2k}\,,$$

hence $\|u_{a,k}\|_{L^2([0,1]\times\mathbb{R})} \lesssim 2^{-k}$.

Step 4: *Localization weights and bushes.* To measure the size of each packet v_T we introduce a bump function in $I_T \times \mathbb{R}$, namely

$$\chi_T(t,x) = 1_{I_T}(t)\left(1 + \lambda^{\frac{2}{3}}|x - x_T(t)|\right)^{-2}.$$

To measure the local density of (a,k)-packets we introduce the function

$$\chi_{a,k} = \sum_{(T,j)\in\mathcal{T}_{a,k}} 1_{I_{T,j}} \chi_T\,.$$

We note that

$$|v_T| \lesssim \lambda^{\frac{1}{3}} \chi_T\,,$$

therefore we have the straightforward pointwise bound

$$|u_{a,k}| \lesssim \lambda^{\frac{1}{3}} a\, \chi_{a,k}\,.$$

This suffices in the low density region

$$A_{a,k,0} = \{\chi_{a,k} \le 1\}\,,$$

as interpolating the above pointwise bound with the above estimate $\|u_{a,k}\|_{L^2} \lesssim 2^{-k}$, we obtain

$$\|u_{a,k}\|_{L^8(A_{a,k,0})} \lesssim \lambda^{\frac{1}{4}} a^{\frac{3}{4}} 2^{-\frac{k}{4}}\,.$$

We may sum over $k \ge 0$ and $a \le 1$ to obtain the desired L^8 bound without log factors. For $p > 8$ the resulting bound is even better than needed.

To obtain bounds over sets where $\chi_{a,k}$ is large, we need to consider how the solution u behaves on regions larger than a single $\lambda^{-\frac{1}{3}}$ slab, in addition to more precise bounds within each slab.

Step 5: *Concentration scales and bushes.* Here we introduce a final parameter $m \ge 1$ which measures the dyadic size of the packet density. Precisely, we consider the sets

$$A_{a,k,m} = \{(t,x) \in [0,1]\times\mathbb{R} : 2^{m-1} < \chi_{a,k}(t,x) \le 2^m\}\,.$$

The points in $A_{a,k,m}$ are called (a,k,m)-bush centers, since as shown in the next section they correspond to the intersection at time t of about 2^m of the packets comprising $u_{a,k}$.

We remark that by fixed-time L^2 bounds on u, and tube overlap considerations, the parameter m must satisfy

$$2^m a^2 \lesssim 1\,, \qquad 2^m \lesssim \lambda^{\frac{1}{3}}\,. \tag{2.11}$$

Our goal will be to measure $\|u_{a,k}\|_{L^8(A_{a,k,m})}$. Some heuristic considerations are appropriate here.

We first note that a collection of 2^m tubes that overlap at a common time t, which we call a 2^m-bush, can retain full overlap for time $\delta t = 2^{-m}\lambda^{-\frac{1}{3}}$. For this to happen the tubes in the bush must have close angles. If the bush is more spread out then the coherence time decreases.

On the other hand, with favourable geometry a focused 2^m-bush may come back together after time $\delta t = 2^m\lambda^{-\frac{1}{3}}$. This indicates that beyond this scale our only available tool is summation with respect to the number of such time intervals.

In effect, the worst case scenario for 2^m-bushes is when there are exactly 2^m tubes which come together on each $\delta t = 2^m\lambda^{-\frac{1}{3}}$ time interval. If the total number of tubes is larger than 2^m then we expect an improvement on the L^8 norm.

Given the above considerations, we decompose the unit time interval $[0,1]$ into a collection $\mathcal{I}_m$ of intervals of size $\delta t = 2^{-m}\lambda^{-\frac{1}{3}}$; such intervals are then dyadic subintervals of the decomposition into $\lambda^{-\frac{1}{3}}$ time slices made in step 2. The proof of Theorem 2.1 is concluded using the following two propositions. The first one counts how many of these slices may contain (a,k,m)-bushes.

Proposition 2.4. *There are at most $\approx \lambda^{\frac{1}{3}} 2^{-3m} a^{-4} \big\langle \log(2^m a^2)\big\rangle^3$ intervals $I \in \mathcal{I}_m$ which intersect $A_{a,k,m}$.*

The second one estimates $\|u_{a,k}\|_{L^8(A_{a,k,m})}$ on a single $2^{-m}\lambda^{-\frac{1}{3}}$ time slice.

Proposition 2.5. *For each interval $I \in \mathcal{I}_m$, we have*

$$\|u_{a,k}\|_{L^8(A_{a,k,m}\cap I\times\mathbb{R})} \lesssim \lambda^{\frac{5}{24}} 2^{\frac{3m}{8}} a^{\frac{1}{2}} 2^{-\frac{k}{4}}\,. \tag{2.12}$$

Combining the two propositions we obtain

$$\|u_{a,k}\|_{L^8(A_{a,k,m})} \lesssim \lambda^{\frac{1}{4}} \big\langle \log(2^m a^2)\big\rangle^{\frac{3}{8}}\, 2^{-\frac{k}{4}}\,.$$

The sets $A_{a,k,m}$ are disjoint, and $\langle \log(2^m a^2)\rangle \lesssim \log\lambda$. Since there are at most $\log\lambda$ values of m, we obtain

$$\|u_{a,k}\|_{L^8([0,1]\times\mathbb{R})} \lesssim \lambda^{\frac{1}{4}} (\log\lambda)^{\frac{1}{2}} 2^{-\frac{k}{4}}\,.$$

We may sum over $k \geq 0$ without additional loss, and there are at most $\log\lambda$ distinct values of a, which yields the desired conclusion (2.2).

For $p > 8$, we interpolate (2.12) with $|u_{a,k}| \lesssim \lambda^{\frac{1}{3}} 2^m a$ to obtain

$$\|u_{a,k}\|^p_{L^p(A_{a,k,m}\cap I\times\mathbb{R})} \lesssim \lambda^{\frac{p}{3}-1}\, 2^{m(p-5)} a^{p-4}\, 2^{-2k}\,,$$

and summing over intervals yields

$$\|u_{a,k}\|^p_{L^p(A_{a,k,m})} \lesssim \lambda^{\frac{p-2}{3}}\, 2^{m(p-8)} a^{p-8}\, 2^{-2k} \big\langle \log(2^m a^2)\big\rangle^3$$
$$= \lambda^{\frac{p}{2}-2} (2^{\frac{m}{2}} a)^{p-8}\, (\lambda^{-\frac{1}{6}} 2^{\frac{m}{2}})^{p-8}\, 2^{-2k}\, \big\langle \log(2^m a^2)\big\rangle^3 .$$

By (2.11), the quantity a takes on dyadic values less than $2^{-\frac{m}{2}}$, whereas 2^m takes on dyadic values less than $\lambda^{\frac{1}{3}}$. We may thus sum over a, k, m to obtain the desired bound

$$\|u\|^p_{L^p([0,1]\times\mathbb{R})} \lesssim \lambda^{\frac{p}{2}-2} ,$$

which together with the estimate for $p = 8$ concludes the proof of Theorem 1.1. □

3. Bush counting

In this section we prove Proposition 2.4. There are $2^m \lambda^{\frac{1}{3}}$ intervals in $\mathcal{I}_m$, so the bound is trivial unless $a \geq 2^{-m}$. It suffices to prove, for ϵ a fixed small number, that if among $\epsilon\, 2^{3m} a^2 \langle \log(2^m a^2)\rangle^{-1}$ consecutive slices in $\mathcal{I}_m$ there are M slices that intersect $A_{a,k,m}$, then

$$M \lesssim (2^m a^2)^{-1} \langle \log(2^m a^2)\rangle^2 . \tag{3.1}$$

Heuristically we would like to say that a point in $A_{a,k,m}$ corresponds to 2^m packets through a point. To make this precise, we need to take into account the tails in the bump functions χ_T.

Consider a point (t, x) in a $2^{-m}\lambda^{-\frac{1}{3}}$ slice $I \times \mathbb{R}$, such that $\chi_{a,k}(t, x) \geq 2^m$. For each $y \in \mathbb{R}$, we denote by $N(y)$ the number of tubes T in the definition of $\chi_{a,k}$ which are centered near y at time t, i.e.

$$N(y) = \#\big\{(T, j) \in \mathcal{T}_{a,k} \,:\, t \in I_{T,j} \;\text{ and }\; |x_T(t) - y| \leq \lambda^{-\frac{2}{3}}\big\} .$$

Then

$$\chi_{a,k}(t, x) \lesssim \lambda^{\frac{2}{3}} \int \big(1 + \lambda^{\frac{2}{3}} |x - y|\big)^{-2} N(y)\, dy .$$

Hence there must exist some point y such that $N(y) \gtrsim 2^m$. Thus, we can find a point y and $\gtrsim 2^m$ indices $(T, j) \in \mathcal{T}_{a,k}$ for which $|x_T(t) - y| \leq \lambda^{-\frac{2}{3}}$ and $t \in I_{T,j}$. Since there are at most 5 values of T with the same ξ_T for which $|x_T(t) - y| \leq \lambda^{-\frac{2}{3}}$, we may select a subset of $\approx 2^m$ packets which have distinct values of ξ_T. We call this an (a, k, m)-*bush* centered at (t, y). For simplicity, we assume the bush contains exactly 2^m terms.

Consider a collection $\{B_n\}_{n=1}^M$ of M distinct (a, k, m)-bushes, centered at (t_n, x_n), with

$$\epsilon\, \lambda^{-\frac{1}{3}} 2^{2m} a^2 \langle \log(2^m a^2)\rangle^{-1} \geq |t_n - t_{n'}| \geq \lambda^{-\frac{1}{3}} 2^{-m} \quad \text{when} \quad n \neq n' .$$

Denote by $\{v_{n,l}\}_{l=1,2^m}$ the collection of 2^m terms v_T comprising B_n. For each n we define the bounded projection operators P_n on $L^2(\mathbb{R})$ at time t_n by

$$P_n f = 2^{-m} a^2 \left(\sum_l \overline{c_{n,l}}(t_n)^{-1} v_{n,l}(t_n, \cdot\,) \right) \left(\sum_l c_{n,l}(t_n)^{-1} \langle v_{n,l}(t_n, \cdot\,), f\rangle \right)$$

where we recall that $|c_{n,l}(t_n)| \approx a$, so that $\|P_n f\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}$. Applying these projections to our solution u to (2.4) at time t_n, recalling that $\langle v_{n,l}(t,\cdot\,), u(t,\cdot\,)\rangle = c_{n,l}(t)$, we obtain

$$P_n u = a^2 \left(\sum_l \overline{c_{n,l}}(t_n)^{-1}\, v_{n,l}(t_n,\cdot\,) \right)$$

therefore

$$\|P_n u\|^2_{L^2(\mathbb{R})} \approx 2^m a^2\,, \tag{3.2}$$

where we are using (2.5) and the fact that the $v_{n,l}$ have distinct ξ_T values. Note that the L^2 norm is invariant under the $a^w_{\lambda^{2/3}}$ flow.

If these projectors were orthogonal with respect to the flow of (2.4), that is

$$P_{n'} S(t_{n'}, t_n) P_n = 0\,,$$

then we would obtain

$$1 = \|u_0\|^2_{L^2(\mathbb{R})} \gtrsim \sum_n \|P_n u\|^2_{L^2(\mathbb{R})} \approx M 2^m a^2\,,$$

and (3.1) would be trivial. This is too much to hope for. Instead, we will prove that the operators P_n satisfy an almost orthogonality relation:

Lemma 3.1. *Let $\alpha = \max\bigl(\lambda^{-\frac13}|t_{n'} - t_n|^{-1}, \lambda^{\frac13}|t_{n'} - t_n|\bigr)$. Then the operators P_n satisfy*

$$\|P_{n'} S(t_{n'}, t_n) P_n\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \lesssim 2^{-m}\alpha\,\langle \log(2^{-m}\alpha)\rangle\,. \tag{3.3}$$

We postpone the proof of Lemma 3.1 to the end of section 5. This estimate is not strong enough to allow us to use Cotlar's lemma. However, we can prove a weaker result, namely that for any solution u to (2.4) we have

$$\sum_n \|P_n u\|_{L^2(\mathbb{R})} \lesssim C^{\frac12} \|u_0\|_{L^2(\mathbb{R})}\,, \qquad C = M + \sum_{n,n'} 2^{-m}\alpha\,\langle \log(2^{-m}\alpha)\rangle\,. \tag{3.4}$$

Indeed, by duality (3.4) is equivalent to

$$\Bigl\| \sum_n S(0,t_n) P_n f_n \Bigr\|_{L^2(\mathbb{R})} \lesssim C^{\frac12} \sup_n \|f_n\|_{L^2(\mathbb{R})}\,.$$

Using (3.3) we have

$$\begin{aligned}\Bigl\| \sum_n S(0,t_n) P_n f_n \Bigr\|^2_{L^2(\mathbb{R})} &= \sum_{n,n'} \langle f_{n'}, P_{n'} S(t_{n'},t_n) P_n f_n\rangle \\ &\lesssim \Bigl(M + \sum_{n\neq n'} 2^{-m}\alpha\,\langle \log(2^{-m}\alpha)\rangle \Bigr) \sup_n \|f_n\|^2_{L^2(\mathbb{R})}\,,\end{aligned}$$

establishing (3.4).

Comparing (3.2) and (3.4) applied to u, it follows that

$$2^{\frac m2} a M \lesssim C^{\frac12}\,,$$

or, in expanded form,

$$2^m a^2 M^2 \lesssim M + \sum_{n\neq n'} 2^{-m}\alpha\,\langle \log(2^{-m}\alpha)\rangle\,. \tag{3.5}$$

This will be the source of our bound in (3.1) for M. If

$$2^m a^2 M^2 \lesssim M$$

then we are done, so we consider the summation term.

First consider the sum over terms where $\lambda^{\frac{1}{3}}|t_{n'} - t_n| \geq 1$, for which we have

$$\sum_{n\neq n'} 2^{-m}\lambda^{\frac{1}{3}}|t_{n'} - t_n| \left\langle \log(2^{-m}\lambda^{\frac{1}{3}}|t_{n'} - t_n|\,) \right\rangle \lesssim \epsilon\,|\log\epsilon|\,2^m a^2 M^2\,,$$

where we use that $r\langle\log r\rangle$ is an increasing function, and that (recall (2.11))

$$2^{-m}\lambda^{\frac{1}{3}}|t_{n'} - t_n| \leq \epsilon\,2^m a^2 \langle\log(2^m a^2)\rangle^{-1} \ll 1\,.$$

Taking ϵ small we can thus absorb these terms into the left hand side of (3.5).

To conclude the proof, we consider the sum over $\lambda^{\frac{1}{3}}|t_{n'} - t_n| \leq 1$. By the $2^{-m}\lambda^{-\frac{1}{3}}$ separation of the M points t_n, we have

$$\sum_{n\neq n'} 2^{-m}\lambda^{-\frac{1}{3}}|t_{n'} - t_n|^{-1} \left\langle \log(2^m\lambda^{\frac{1}{3}}|t_{n'} - t_n|) \right\rangle \lesssim M(\log M)^2\,.$$

We conclude that

$$2^m a^2 M \lesssim (\log M)^2\,,$$

hence that

$$M \lesssim (2^m a^2)^{-1}\left\langle\log(2^m a^2)\right\rangle^2.$$

4. Short time bounds

In this section we prove Proposition 2.5. We recall the bound we need,

$$\|u_{a,k}\|_{L^8(A_{a,k,m}\cap I\times\mathbb{R})} \lesssim \lambda^{\frac{5}{24}} 2^{\frac{3}{8}m} a^{\frac{1}{2}} 2^{-\frac{k}{4}}$$

where $|I| = 2^{-m}\lambda^{-\frac{1}{3}}$. Here $u_{a,k}$ on $I\times\mathbb{R}$ has the form

$$1_I(t)\cdot u_{a,k} = \sum_{(T,j)\in\mathcal{T}_{a.k}\,:\,I_{T,j}\cap I\neq\emptyset} 1_I c_{T,j} v_T\,,$$

where we recall that $c_{T,j} = 1_{I_{T,j}} c_T$, $c'_{T,j} = 1_{I_{T,j}} c'_T$. Also, $|I_{T,j}| = 2^{-k}\lambda^{-\frac{1}{3}}$, and

$$|c_{T,j}| \approx a\,, \qquad \|c'_{T,j}\|_{L^2} \lesssim \lambda^{\frac{1}{6}} 2^{\frac{k}{2}} a\,.$$

Note that if $k \leq m$, then $I_{T,j} \supseteq I$ for each term in the sum, whereas if $k > m$, then $I_{T,j}$ is a dyadic subinterval of I, and there may be multiple terms associated to a tube T.

We let N denote the number of terms in the sum for $u_{a,k}$, and note that, by (2.6) and (2.9), we have

$$N a^2 \lesssim 2^{-k}\,. \tag{4.1}$$

Using this bound, and dividing $u_{a,k}$ by a, we then need establish the following.

Lemma 4.1. *Let $\mathcal{T}$ be a collection of N distinct pairs (T,j), and $I_{T,j}$ corresponding intervals of length $2^{-k}\lambda^{-\frac{1}{3}}$ which intersect the interval I of length $2^{-m}\lambda^{-\frac{1}{3}}$. Assume that*

$$\|c_{T,j}\|_{L^\infty} + 2^{-\frac{k}{2}}\lambda^{-\frac{1}{6}}\|c'_{T,j}\|_{L^2} \leq 1\,.$$

Then with $v = \sum_{\mathcal{T}} c_{T,j} v_T$*, the following holds*

$$\|v\|_{L^8(A_m \cap I \times \mathbb{R})} \lesssim \lambda^{\frac{5}{24}} N^{\frac{1}{4}} 2^{\frac{3}{8}m}\,,$$

where

$$A_m = \{\chi \approx 2^m\}\,, \qquad \chi = \sum_{T \in \mathcal{T}} \chi_T\,.$$

To start the proof, we first show that we can dispense with the high angle interactions. We want to establish

$$\|v^2\|_{L^4(A_m \cap I \times \mathbb{R})} \lesssim \lambda^{\frac{5}{12}} N^{\frac{1}{2}} 2^{\frac{3}{4}m}\,.$$

We express v^2 using a bilinear angular decomposition. Fixing some reference angle θ we can write

$$v^2 = \sum_l \pm v_l^2 \;\; + \sum_{\angle(T,S) \geq \theta} c_{T,j} v_T \cdot c_{S,k} v_S$$

where $v_l = \sum_{\xi_T \in K_l} c_{T,j} v_T$ consists of the terms for which ξ_T lies in an interval K_l of length $\approx \lambda\theta$, where the K_l have overlap at most 3. The second sum is over a subset of $\mathcal{T} \times \mathcal{T}$ subject to the condition $\angle(T,S) = \lambda^{-1}|\xi_T - \xi_S| \geq \theta$.

For the second term we have a bilinear L^2 estimate,

Lemma 4.2. *The following bilinear* L^2 *bound holds,*

$$\Big\| \sum_{\angle(T,S) \geq \theta} c_{T,j} v_T \cdot c_{S,k} v_S \Big\|_{L^2(I \times \mathbb{R})} \lesssim \theta^{-\frac{1}{2}} \sum_{T \in \mathcal{T}} \Big(\|c_{T,j}\|_{L^\infty}^2 + 2^{-k} \lambda^{-\frac{1}{3}} \|c'_{T,j}\|_{L^2}^2 \Big)\,. \tag{4.2}$$

For this estimate there is no restriction on the number of tubes, nor do we require equal size of the c_T. The integral can furthermore be taken over the interval of length $\lambda^{-\frac{1}{3}}$ containing I; the short time condition is needed only for the small angle interactions.

We prove (4.2) in section 6. In our case, each term in the sum on the right is bounded by 1. On the other hand, we have $|v_T| \lesssim \lambda^{\frac{1}{3}} \chi_T$ therefore

$$\Big| \sum_{\angle(T,S) \geq \theta} c_{T,j} v_T \cdot c_{S,k} v_S \Big| \lesssim \lambda^{\frac{2}{3}} \chi^2\,,$$

which yields

$$\Big\| \sum_{\angle(T,S) \geq \theta} c_{T,j} v_T \cdot c_{S,k} v_S \Big\|_{L^\infty(A_m)} \lesssim \lambda^{\frac{2}{3}} 2^{2m}$$

Interpolating the L^2 and the L^∞ bounds we obtain

$$\Big\| \sum_{\angle(T,S) \geq \theta} c_{T,j} v_T \cdot c_{S,k} v_S \Big\|_{L^4(A_m)} \lesssim \lambda^{\frac{1}{3}} N^{\frac{1}{2}} 2^m \theta^{-\frac{1}{4}}\,.$$

This is what we need for the high angle component provided that

$$\lambda^{\frac{1}{3}} N^{\frac{1}{2}} 2^m \theta^{-\frac{1}{4}} = \lambda^{\frac{5}{12}} N^{\frac{1}{2}} 2^{\frac{3}{4}m}\,,$$

or equivalently

$$\theta = \lambda^{-\frac{1}{3}} 2^m\,.$$

Hence it suffices to restrict ourselves to the terms v_l, where $\xi_T \in K_l$, an interval of width $\delta\xi = \lambda^{\frac{2}{3}} 2^m$ centered on ξ_l. Let N_l denote the number of terms (T, j) in v_l, so that $\sum_l N_l \le 3N$. We will prove that

$$\big\|(1 + \lambda^{-\frac{4}{3}} 2^{-2m} |D - 2\xi_l|^2) v_l^2\big\|_{L^3(I\times\mathbb{R})} \lesssim \lambda^{\frac{1}{3}} N_l^{\frac{2}{3}} 2^{\frac{m}{3}} \,. \tag{4.3}$$

Let $Q_l(D) = (1 + \lambda^{-\frac{4}{3}} 2^{-2m} |D - 2\xi_l|^2)^{-1}$. We observe that, for $w_l \in \mathcal{S}(\mathbb{R})$,

$$\Big\|\sum_l Q_l w_l\Big\|_{L^3(\mathbb{R})} \lesssim \Big(\sum_l \|w_l\|_{L^3(\mathbb{R})}^{\frac{3}{2}}\Big)^{\frac{2}{3}} ,$$

which follows by interpolating the bounds

$$\Big\|\sum_l Q_l w_l\Big\|_{L^\infty(\mathbb{R})} \lesssim \sum_l \|w_l\|_{L^\infty(\mathbb{R})} \,, \qquad \Big\|\sum_l Q_l w_l\Big\|_{L^2(\mathbb{R})} \lesssim \Big(\sum_l \|w_l\|_{L^2(\mathbb{R})}^2\Big)^{\frac{1}{2}} .$$

The second follows by the finite overlap condition, the first since Q_l is convolution with respect to an L^1 function. Applying this to (4.3) yields

$$\Big\|\sum_l v_l^2\Big\|_{L^3(I\times\mathbb{R})} \lesssim \lambda^{\frac{1}{3}} N^{\frac{2}{3}} 2^{\frac{m}{3}} \,.$$

Interpolating with the L^∞ bounds as above yields the desired bound

$$\Big\|\sum_l v_l^2\Big\|_{L^4(A_m\cap I\times\mathbb{R})} \lesssim \lambda^{\frac{5}{12}} N^{\frac{1}{2}} 2^{\frac{3}{4}m} \,.$$

By Leibniz' rule and Hölder's inequality, (4.3) follows from showing, for $n \le 2$, that

$$\big\|\big(\lambda^{-\frac{2}{3}} 2^{-m}(D - \xi_l)\big)^n v_l\big\|_{L^6(I\times\mathbb{R})} \lesssim \lambda^{\frac{1}{6}} N_l^{\frac{1}{3}} 2^{\frac{m}{6}} \,. \tag{4.4}$$

We first note the following bound on v_l in L^6 over the entire $2^{-\min(k,m)}\lambda^{-\frac{1}{3}}$ time slice $I^* \times \mathbb{R}$ on which the v_l are supported:

$$\big\|\big(\lambda^{-\frac{2}{3}} 2^{-m}(D - \xi_l)\big)^n v_l\big\|_{L^6(I^*\times\mathbb{R})} \lesssim \lambda^{\frac{1}{6}} N_l^{\frac{1}{2}} \,. \tag{4.5}$$

To establish this, it suffices by the generalized Minkowski inequality to establish it on a time interval J of length $2^{-k}\lambda^{-\frac{1}{3}}$, with N_l replaced by the number N_J of (T, j) for which $I_{T,j} = J$. If $k \le m$, then there is only one interval J to consider, whereas $k > m$ means J is a dyadic subdivision of I. If t_0 is the left endpoint of J, then we have the initial data bound

$$\|\big(\lambda^{-\frac{2}{3}} 2^{-m}(D - \xi_l)\big)^n v_l(t_0)\|_{L^2(\mathbb{R})} \lesssim \Big(\sum |c_{T,j}(t_0)|^2\Big)^{\frac{1}{2}} \lesssim N_J^{\frac{1}{2}} ,$$

and for the inhomogeneous term we have

$$\|\big(\lambda^{-\frac{2}{3}} 2^{-m}(D - \xi_l)\big)^n \big(D_t - a^w_{\lambda^{2/3}}(t, x, D)\big) v_l\|_{L^1_t L^2_x(J\times\mathbb{R})} \lesssim |J|^{\frac{1}{2}} \Big(\sum \|c'_{T,j}\|_{L^2(J)}^2\Big)^{\frac{1}{2}} \lesssim N_J^{\frac{1}{2}} \,.$$

The result then holds by the weighted Strichartz estimates, Theorem 5.4.

To obtain the gain in the norm over the slice $I \times \mathbb{R}$, we make a further decomposition $v_l = \sum_B v_B$ into "bushes". This is made by decomposing the x-axis into disjoint intervals of radius $\lambda^{-\frac{2}{3}}$, indexed by B, with center x_B, and letting v_B denote the sum of the $c_{T,j} v_T$ in v_l whose center x_T at time t_0 satisfies $|x_T - x_B| \le \lambda^{-\frac{2}{3}}$.

For simplicity, we take $t_0 = 0$. Let $x_B(t)$ denote the bicharacteristic curve passing through (x_B, ξ_l). Then $|x_T - x_B| \lesssim \lambda^{-\frac{2}{3}}$, provided T is part of v_B. By Theorem 5.4 we thus have the weighted Strichartz estimates,

$$\left\| \left(1 + \lambda^{\frac{4}{3}} |x - x_B(t)|^2\right) \left(\lambda^{-\frac{2}{3}} 2^{-m} (D - \xi_l)\right)^n v_B \right\|_{L^6(I \times \mathbb{R})} \lesssim \lambda^{\frac{1}{6}} N_B^{\frac{1}{2}} ,$$

where N_B is the number of terms in v_B. We may sum over B to obtain

$$\left\| \left(\lambda^{-\frac{2}{3}} 2^{-m} (D - \xi_l)\right)^n v_l \right\|_{L^6(I \times \mathbb{R})} \lesssim \lambda^{\frac{1}{6}} \Big(\sum_B N_B^3 \Big)^{\frac{1}{6}} \leq \lambda^{\frac{1}{6}} N_l^{\frac{1}{6}} 2^{\frac{m}{3}} , \tag{4.6}$$

where at the last step we used $N_B \leq 2^m$, and $\sum_B N_B = N_l$. Combining (4.6) with (4.5) yields (4.4). □

5. Wave packet propagation

In this section we establish the basic properties of the wave packet solutions v_T on the $\lambda^{-\frac{1}{3}}$ time scale, and prove weighted Strichartz estimates. In addition, we give the proof of Lemma 3.1. Throughout this section, we let $A = a^w_{\lambda^{2/3}}(t, x, D)$, and let u solve $(D_t - A)u = 0$, $u(0, \cdot\,) = u_0$. We assume $u_0 \in \mathcal{S}$, so that all derivatives of u are rapidly decreasing in x. Throughout, I is an interval with left hand endpoint 0 and $|I| \leq \lambda^{-\frac{1}{3}}$.

Lemma 5.1. *For any $m, n \geq 0$ and $\xi_0 \in \mathbb{R}$,*

$$\sum_{j=0}^{n} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \| (D - \xi_0)^j u \|_{L^\infty L^2(I \times \mathbb{R})} \leq C_n \sum_{j=0}^{n} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \| (D - \xi_0)^j u_0 \|_{L^2(\mathbb{R})} .$$

Proof. We use induction on n. The case $n = 0$ follows by self-adjointness of A, so we assume the result holds for $n - 1$. We may write the commutator

$$\lambda^{-\frac{2}{3}} [(D - \xi_0), A] = \lambda^{\frac{1}{3}} b^w(t, x, D) , \quad b \in S_{\lambda, \lambda^{2/3}} ,$$

whereas commuting with $\lambda^{-\frac{2}{3}}(D - \xi_0)$ preserves the set of Weyl-pseudodifferential operators with symbol in $S_{\lambda, \lambda^{2/3}}$. Hence, we may write

$$(D_t - A) \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^n (D - \xi_0)^n u = \lambda^{\frac{1}{3}} \sum_{j=0}^{n-1} b^w(t, x, D) \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j (D - \xi_0)^j u ,$$

where $b \in S_{\lambda, \lambda^{2/3}}$ may vary with j. The proof follows by L^2 boundedness of $b^w(t, x, D)$ and the Duhamel formula, since $|I| \leq \lambda^{-\frac{1}{3}}$. □

A similar proof, using the fact that

$$\lambda^{\frac{1}{3}} [x, A] = \lambda^{\frac{1}{3}} b^w(t, x, D) , \quad b \in S_{\lambda, \lambda^{2/3}} ,$$

and that commuting with $\lambda^{\frac{1}{3}} x$ preserves $S_{\lambda, \lambda^{2/3}}$, yields the following.

Corollary 5.2. *For any $l, m, n \geq 0$, and all $x_0, \xi_0 \in \mathbb{R}$,*

$$\sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(x-x_0)^k (D-\xi_0)^j u\|_{L^\infty L^2(I\times\mathbb{R})}$$
$$\leq C_{n,l} \sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(x-x_0)^k (D-\xi_0)^j u_0\|_{L^2(\mathbb{R})}\,.$$

To obtain weighted localization in x at the $\lambda^{-\frac{2}{3}}$ scale we need to evolve the spatial center of u along the bicharacteristic flow. Additionally, we must work on a time interval I so that the spread of bicharacteristics due to the spread of frequency support is less than $\lambda^{-\frac{2}{3}}$.

Lemma 5.3. *Let $x_0(t) = x_0 - t\,\partial_\xi a_{\lambda^{2/3}}(0, x_0, \xi_0)$, and suppose that $|I| \leq 2^{-m}\lambda^{-\frac{1}{3}}$. Then for $l \leq n$, and general m, n, x_0, ξ_0,*

$$\sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{2k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(x-x_0(t))^k (D-\xi_0)^j u\|_{L^\infty L^2(I\times\mathbb{R})}$$
$$\leq C_n \sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{2k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(x-x_0)^k (D-\xi_0)^j u_0\|_{L^2(\mathbb{R})}\,.$$

Proof. We write

$$i\,[x-x_0(t), D_t - A] = (\partial_\xi a_{\lambda^{2/3}})^w(t,x,D) - \partial_\xi a_{\lambda^{2/3}}(0,x_0,\xi_0)\,,$$

and taking a Taylor expansion write

$$\lambda^{\frac{2}{3}}\left(\partial_\xi a_{\lambda^{2/3}}(t,x,\xi) - \partial_\xi a_{\lambda^{2/3}}(0,x_0,\xi_0)\right) =$$
$$\lambda^{\frac{1}{3}} 2^m \Big(b_1(t,x,\xi)\lambda^{\frac{1}{3}}(x-x_0) + b_2(t,x,\xi)\lambda^{\frac{1}{3}} t + b_3(t,x,\xi)\lambda^{-\frac{2}{3}}2^{-m}(\xi-\xi_0)\Big)\,,$$

with $b_j \in S_{\lambda,\lambda^{2/3}}$, where we use $2^m \geq 1$. Additionally, commuting with $\lambda^{\frac{2}{3}}x$ preserves the class of $b^w(t,x,D)$ with $b \in S_{\lambda,\lambda^{2/3}}$. The proof now proceeds along the lines of the proof of Lemma 5.1, using that $|I| \leq \lambda^{-\frac{1}{3}}2^{-m}$. □

We remark that the proof of Lemma 5.3 in fact shows that one may bound

$$\sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{2k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(D_t - A)(x-x_0(t))^k (D-\xi_0)^j u\|_{L^1 L^2(I\times\mathbb{R})}$$
$$\leq C_n \sum_{j=0}^{n} \sum_{k=0}^{l} \lambda^{\frac{2k}{3}} \left(\lambda^{-\frac{2}{3}} 2^{-m}\right)^j \|(x-x_0)^k (D-\xi_0)^j u_0\|_{L^2(\mathbb{R})}\,,$$

provided $l \leq n$ and $|I| \leq 2^{-m}\lambda^{-\frac{1}{3}}$, or $|I| \leq \lambda^{-\frac{1}{3}}$ in case $l = 0$. We thus can prove weighted Strichartz estimates as an easy corollary of the unweighted version. We state the result for $p = q = 6$, but it holds for all allowable values of (p,q) for which the unweighted version holds.

Theorem 5.4. *Let* $x_0(t) = x_0 - t\,\partial_\xi a_{\lambda^{2/3}}(0,x_0,\xi_0)\,,$ *and suppose that* $|I| \le 2^{-m}\lambda^{-\frac13}$. *Then for* $l \le n$, *and general* m, n, x_0, ξ_0,

$$\sum_{j=0}^{n}\sum_{k=0}^{l}\lambda^{\frac{2k}{3}}\bigl(\lambda^{-\frac23}2^{-m}\bigr)^j\|(x-x_0(t))^k(D-\xi_0)^j u\|_{L^6(I\times\mathbb{R})}$$
$$\le C_n\,\lambda^{\frac16}\sum_{j=0}^{n}\sum_{k=0}^{l}\lambda^{\frac{2k}{3}}\bigl(\lambda^{-\frac23}2^{-m}\bigr)^j\|(x-x_0)^k(D-\xi_0)^j u_0\|_{L^2(\mathbb{R})}\,.$$

If $l = 0$, *then the result holds for* $|I| \le \lambda^{-\frac13}$.

Proof. By the above remarks, the result follows by the Duhamel theorem from the case $n = l = 0$. That case, in turn, follows from Theorem 2.5 of [4]. An alternate proof is contained in [1]. That paper dealt with $\lambda^{-1}\Delta_{\mathrm{g}}$ instead of A, but the analysis is similar for A as above. $\square$

If we take $m = 0$, then Lemma 5.3 applies to the evolution of a $\lambda^{-\frac23}$ packet. Precisely,

Theorem 5.5. *Suppose that* ϕ *is a Schwartz function, and* $\phi_T = \lambda^{\frac13}e^{ix\xi_T}\phi(\lambda^{\frac23}(x-x_T))$. *Let* v_T *satisfy*

$$\bigl(D_t - A\bigr)v_T = 0\,,\qquad v_T(0,\cdot\,) = \phi_T\,.$$

Then with

$$x_T(t) = x_T - t\,\partial_\xi a_{\lambda^{2/3}}(0,x_T,\xi_T)\,,$$

for $t \in [0,\lambda^{\frac13}]$ *one can write*

$$v_T(t,\cdot) = \lambda^{\frac13}e^{ix\xi_T}\psi_T\bigl(t,\lambda^{\frac23}(x-x_T(t))\bigr)\,,$$

where $\{\psi_T(t,\cdot)\}_{t\in I}$ *is a bounded family of Schwartz functions on* $\mathbb{R}$, *with all Schwartz norms uniformly bounded over* T, $t \in I$, *and* $\lambda \ge 1$.

We conclude this section with the proof of Lemma 3.1. Let P_0 denote the bounded projection on $L^2(\mathbb{R})$ defined by $P_0 f = 2^{-m}\langle w_0, f\rangle w_0$, where w_0 is a sum of 2^m normalized packets centered at x_0 with disjoint frequency centers; that is,

$$w_0 = \sum_{l=1}^{2^m}\lambda^{\frac13}e^{ix\xi_l}\psi_l\bigl(\lambda^{\frac23}(x-x_0)\bigr)\,,$$

where the ψ_l are a bounded collection of Schwartz functions, and the ξ_l are distinct points on the $\lambda^{\frac23}$-spaced lattice in $\mathbb{R}$, with $|\xi_l| \le \frac12\lambda$. We take P_1 to be similarly defined where x_0 is replaced by x_1, possibly with a different set of ξ_l and ψ_l. We need to prove

$$\|P_1 S(t_1,t_0)P_0\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} \lesssim 2^{-m}\alpha\,\langle\log(2^{-m}\alpha)\rangle\,, \tag{5.1}$$

where $\alpha = \max\bigl(\lambda^{-\frac13}|t_1-t_0|^{-1},\lambda^{\frac13}|t_1-t_0|\bigr)$.

The proof of (5.1) requires control of the solution u over times greater than $\lambda^{-\frac13}$, which we will express in the form of weighted energy estimates. Heuristically, for $t \le \lambda^{-\frac13}$ one can localize energy flow at the symplectic $\lambda^{\frac23}$ scale. For $t \ge \lambda^{-\frac13}$, energy flow cannot be localized finer than the uncertainty in the Hamiltonian flow, where ξ is determined only within λt, with resulting uncertainty in x. In the notation below, our weighted energy estimate localizes ξ to within $\delta\lambda$, and x to within δ^2. The linear growth of the weights reflects the Lipschitz regularity of $a_\lambda(t,x,\xi)$.

In the following, we let

$$\delta = \begin{cases} \lambda^{-\frac{1}{3}}\,, & |t_0-t_1| \le \lambda^{-\frac{1}{3}}\,, \\ |t_0-t_1|\,, & |t_0-t_1| \ge \lambda^{-\frac{1}{3}}\,. \end{cases}$$

Let $q_j(t,x,D)$, $j=0,1$, denote the symbol

$$q_j(t,x,\xi) = \delta^{-2}\big(x - x_j + (t-t_j)\,\partial_\xi a_\lambda(t_j,x_j,\xi)\big)\,,$$

and set $Q_j(t) = q_j(t,x,D)$. We will prove that

$$\|Q_0(t_1)S(t_1,t_0)f\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})} + \|Q_0(t_0)f\|_{L^2(\mathbb{R})} + \delta^{-1}\|(x-x_0)f\|_{L^2(\mathbb{R})}\,. \tag{5.2}$$

Assuming (5.2) for the moment, we consider the constant coefficient symbols

$$m_0(\xi) = \langle q_0(t_1,x_1,\xi)\rangle\,, \qquad m_1(\xi) = \langle q_1(t_0,x_0,\xi)\rangle\,.$$

We will use (5.2) to prove that

$$\big\|m_0(D)^{\frac{1}{2}}\langle\delta^{-2}(x-x_1)\rangle^{-2}S(t_1,t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2}m_1(D)^{\frac{1}{2}}f\big\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}\,, \tag{5.3}$$

with bounds uniform over the various parameters. We then factor

$$\begin{aligned} P_1S(t_1,t_0)P_0 = P_1\langle\delta^{-2}(x-x_1)\rangle^2 m_0(D)^{-\frac{1}{2}}m_0(D)^{\frac{1}{2}}\langle\delta^{-2}(x-x_1)\rangle^{-2} \\ S(t_1,t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2}m_1(D)^{\frac{1}{2}}m_1(D)^{-\frac{1}{2}}\langle\delta^{-2}(x-x_0)\rangle^2P_0\,, \end{aligned}$$

which reduces (5.1) to showing that

$$\big\|P_1\langle\delta^{-2}(x-x_1)\rangle^2m_0(D)^{-\frac{1}{2}}f\big\|^2_{L^2(\mathbb{R})} \lesssim 2^{-m}\alpha\,\langle\log(2^{-m}\alpha)\rangle\|f\|^2_{L^2(\mathbb{R})}\,, \tag{5.4}$$

where we use symmetry and adjoints to conclude that the rightmost factor above satisfies the same bounds. Since $\|w_1\|_{L^2(\mathbb{R})} \approx 2^{\frac{m}{2}}$, (5.4) is implied by

$$\big\|m_0(D)^{-\frac{1}{2}}\langle\delta^{-2}(x-x_1)\rangle^2w_1\|^2_{L^2(\mathbb{R})} \lesssim \alpha\,\langle\log(2^{-m}\alpha)\rangle\,. \tag{5.5}$$

To prove (5.5), note that since $\delta \ge \lambda^{-\frac{1}{3}}$, and the packets in w_1 are centered at x_1, the function $\langle\delta^{-2}(x-x_1)\rangle^2w_1$ is of the same form as w_1. The Fourier transform of w_1 is a sum of 2^m L^2-normalized Schwartz functions, concentrated on the $\lambda^{\frac{2}{3}}$-scale about the distinct ξ_l, and the left hand side of (5.5) can thus be compared to

$$\lambda^{-\frac{2}{3}}\int\Big|\sum_l m_0(\xi)^{-\frac{1}{2}}\big(1+\lambda^{-\frac{2}{3}}|\xi-\xi_l|\big)^{-N}\Big|^2\,d\xi \lesssim \sum_l m_0(\xi_l)^{-1}\,.$$

By (2.3),

$$\partial_\xi a_\lambda(t_0,x_0,\xi_l) - \partial_\xi a_\lambda(t_0,x_0,\xi_{l'}) \approx \lambda^{-1}(\xi_{l'}-\xi_l)\,.$$

Since $\lambda^{\frac{1}{3}}\delta^2|t_1-t_0|^{-1} = \alpha \ge 1$, the estimate (5.5) then follows by comparison to the worst case sum

$$\sum_{j=0}^{2^m}(1+\alpha^{-1}j)^{-1} \lesssim \alpha\big(1+|\log(2^m\alpha^{-1})|\,\big)\,.$$

To see that equation (5.3) is a consequence of (5.2), we observe that (5.3) follows by interpolation from showing, with uniform bounds,

$$\begin{aligned} &\left\| m_0(D)\langle\delta^{-2}(x-x_1)\rangle^{-2} S(t_1,t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2} f\right\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}\,, \\ &\left\| \langle\delta^{-2}(x-x_1)\rangle^{-2} S(t_1,t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2} m_1(D) f\right\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}\,. \end{aligned} \tag{5.6}$$

The second line follows from the first by symmetry and adjoints, so we prove the first estimate. We first note that

$$\|m_0(D)\langle\delta^{-2}(x-x_1)\rangle^{-2} f\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})} + \|Q_0(t_1)\langle\delta^{-2}(x-x_1)\rangle^{-2} f\|_{L^2(\mathbb{R})}\,.$$

The commutator of $Q_0(t_1)$ and $\langle\delta^{-2}(x-x_1)\rangle^{-2}$ is bounded; this uses the fact that $\partial_\xi^2 a_\lambda(t,x,\xi) \in \lambda^{-1} S_{\lambda,\lambda}$, that $|t_1-t_0|\le\delta$, and that $\delta^{-3}\le\lambda$. Thus, (5.6) reduces to showing that

$$\|Q_0(t_1)S(t_1,t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2} f\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})}\,,$$

which follows from (5.2), since the purely spatial weight $Q_0(t_0)\langle\delta^{-2}(x-x_0)\rangle^{-2}$ is bounded, and $\delta\le 1$.

To establish (5.2), we calculate

$$\partial_t\|Q_0(t)S(t,t_0)f\|^2_{L^2(\mathbb{R})} = 2\,\mathrm{Re}\,\big\langle(\partial_t Q_0 + i[Q_0, a^w_\lambda(t,x,D)])S(t,t_0)f, Q_0(t)S(t,t_0)f\big\rangle\,,$$

so since $|t_1-t_0|\le\delta$ it suffices to show that

$$\|(\partial_t Q_0 + i[Q_0, a^w_\lambda(t,x,D)])S(t,t_0)f\|_{L^2(\mathbb{R})} \lesssim \delta^{-1}\|f\|_{L^2(\mathbb{R})} + \delta^{-2}\|(x-x_0)f\|_{L^2(\mathbb{R})}\,.$$

The operator $\partial_t Q_0 + i[Q_0, a^w_\lambda(t,x,D)]$ is equal to

$$\delta^{-2}\big(\partial_\xi a_\lambda(t_0,x_0,D) - \partial_\xi a^w_\lambda(t,x,D)\big) + i\delta^{-2}(t-t_0)\big[\partial_\xi a_\lambda(t_0,x_0,D), a^w_\lambda(t,x,D)\big]\,.$$

The commutator term is bounded on L^2 since $a_\lambda\in\lambda C^1 S_{\lambda,\lambda}$, and $\partial_\xi^2 a_\lambda(t_0,x_0,\xi)\in\lambda^{-1}C^1 S_{\lambda,\lambda}$. Since $|t-t_0|\le\delta$, the second term is thus bounded by δ^{-1}. The L^2 norm of the first term is bounded by $\delta^{-1}+\delta^{-2}(x-x_0)$, so we have to bound

$$\delta^{-1}\|(x-x_0)S(t,t_0)f\|_{L^2(\mathbb{R})} \lesssim \|f\|_{L^2(\mathbb{R})} + \delta^{-1}\|(x-x_0)f\|_{L^2(\mathbb{R})}\,.$$

This follows by Corollary 5.2, since $\delta^{-1}\le\lambda^{\frac13}$. □

6. Bilinear L^2 estimates

We prove here the bilinear estimate Lemma 4.2. For this section, we will let

$$\chi_T(t,x) = \big(1+\lambda^{\frac23}|x-x_T(t)|\big)^{-N}$$

for some suitably large but fixed N, and use the fact that $|v_T|\lesssim\lambda^{\frac13}\chi_T$, by Theorem 5.5. Also in this section we let

$$a(t,x,\xi) = a_{\lambda^{2/3}}(t,x,\xi)\,,\quad a_\xi(t,x,\xi) = \partial_\xi a_{\lambda^{2/3}}(t,x,\xi)\,.$$

We first reduce Lemma 4.2 to the case that the c_T are constants, that is, to the following lemma.

Lemma 6.1. *Suppose that $b_T, d_S \in \mathbb{C}$. Then, for any subset Λ of tube pairs (T,S) satisfying $\angle(T,S) \geq \theta$, the following bilinear L^2 bound holds,*

$$\Big\| \sum_{(T,S)\in\Lambda} b_T v_T \cdot d_S v_S \Big\|_{L^2} \lesssim \theta^{-\frac12}\Big(\sum_T |b_T|^2\Big)^{\frac12}\Big(\sum_S |d_S|^2\Big)^{\frac12}. \tag{6.1}$$

The norm is taken over the common $\lambda^{-\frac13}$ time slice in which the tubes lie, and there is no restriction on the number of terms.

To make the reduction of Lemma 4.2 to Lemma 6.1, we note that it suffices by Minkowski to establish (4.2) on an interval J of size $2^{-k}\lambda^{-\frac13}$, including only the $c_{T,j}$ for which $I_{T,j} = J$. On such an interval we can write

$$c_{T,j}(t) = c_{T,j}(t_0) + \int_0^t c'_{T,j}(s)\,ds\,,$$

where t_0 is the left endpoint of J. We can thus bound

$$\Big|\sum c_{T,j} v_T \cdot c_{S,k} v_S\Big| \leq \Big|\sum c_{T,j}(t_0) v_T \cdot c_{S,k}(t_0) v_S\Big| + \int_J \Big|\sum c'_{T,j}(r) v_T \cdot c_{S,k}(t_0) v_S\Big|\,dr$$
$$+ \int_J \Big|\sum c_{T,j}(t_0) v_T \cdot c'_{S,k}(s) v_S\Big|\,ds + \int_{J\times J}\Big|\sum c'_{T,j}(r) v_T \cdot c'_{S,k}(s) v_S\Big|\,dr\,ds\,.$$

Bringing the integral out of the L^2 norm and applying (6.1) together with the Schwartz inequality yields the desired bound

$$\Big\| \sum_{\angle(T,S)\geq\theta} c_{T,j} v_T \cdot c_{S,k} v_S \Big\|_{L^2(J\times\mathbb{R})} \lesssim \theta^{-\frac12} \sum_{I_{T,j}=J} \Big(\|c_{T,j}\|^2_{L^\infty} + 2^{-k}\lambda^{-\frac13}\|c'_{T,j}\|^2_{L^2}\Big)\,.$$

We now turn to the proof of Lemma 6.1. One estimate we will use is the following. Suppose that the tubes T (respectively S) all point in the same direction, that is, ξ_T is the same for all T, and ξ_S is the same for all S, where $|\xi_T - \xi_S| \geq \lambda\theta$. Then

$$\int \Big(\sum_{T,S} |b_T|\chi_T \cdot |d_S|\chi_S\Big)^2 dt\,dx \lesssim \lambda^{-\frac43}\theta^{-1}\Big(\sum_T |b_T|^2\Big)\Big(\sum_S |d_S|^2\Big)\,. \tag{6.2}$$

This follows since different tubes T (respectively S) are disjoint, and the intersection of any pair of tubes T and S is a $\lambda^{-\frac23}$ interval in x times a $\theta^{-1}\lambda^{-\frac23}$ interval in t about the center of the intersection. Precisely, one can make a change of variables of Jacobian $\lambda^{\frac43}\theta$ to reduce matters to $\lambda = 1$ and $\theta = \frac{\pi}{2}$, where the result is elementary. We remark that since the terms are positive, this holds even if the sum over T and S on the left includes just a subset of the collection of all T and S.

Consider now an integral

$$\int v_T\, v_S\, \overline{v_{T'}}\,\overline{v_{S'}}\,dt\,dx\,.$$

Recalling that $\lambda^{-\frac23}\xi_T \in \mathbb{Z}$, we will relabel

$$\xi_T = \xi_{m+j}\,,\quad \xi_S = \xi_{m-j}\,,\quad \xi_{T'} = \xi_{n+i}\,,\quad \xi_{S'} = \xi_{n-i}\,,$$

corresponding to $\xi_T = \lambda^{\frac23}(m+j)$, etc. Here, m, n, i, and j may take on half-integer values; for simplicity we assume $m \neq n$ and $i \neq j$, and i, j nonzero. The cases of equality are simpler in what follows. By symmetry we may take $j > i \geq 1$. Assuming that $\angle(T,S) \geq \theta$ implies $j \geq \lambda^{\frac{1}{3}}\theta$, and similarly $\angle(T',S') \geq \theta$ implies $i \geq \lambda^{\frac{1}{3}}\theta$.

We introduce the quantities

$$\begin{aligned} v_{(4)} &= v_T\, v_S\, \overline{v_{T'}}\, \overline{v_{S'}}\,, \\ a_{(4)} &= a(t,x,\xi_{m+j}) + a(t,x,\xi_{m-j}) - a(t,x,\xi_{n+i}) - a(t,x,\xi_{n-i})\,, \\ a_{(4')} &= a(t,x,\xi_{m+j}) + a(t,x,\xi_{m-j}) - a(t,x,\xi_{m+i}) - a(t,x,\xi_{m-i})\,. \end{aligned}$$

Since $a_{\xi\xi} \approx \lambda^{-1}$, we have simultaneous upper and lower bounds for $a_{(4')}$,

$$a_{(4')} = \lambda^{\frac{2}{3}} \int_i^j a_\xi(t,x,\xi_{m+s}) - a_\xi(t,x,\xi_{m-s})\, ds \approx \lambda^{\frac{1}{3}} \int_i^j s\, ds \approx \lambda^{\frac{1}{3}}(j^2 - i^2)\,. \tag{6.3}$$

We may similarly use $|\partial^\alpha_{t,x} a_{\xi\xi}| \lesssim \lambda^{-1+\frac{2}{3}(|\alpha|-1)}$ for nonzero α to deduce

$$\bigl|\partial^\alpha_{t,x} a_{(4')}\bigr| \lesssim \lambda^{\frac{1}{3}+\frac{2}{3}(|\alpha|-1)}(j^2 - i^2)\,, \qquad |\alpha| \geq 1\,. \tag{6.4}$$

To control the difference of $a_{(4)}$ and $a_{(4')}$, we introduce the quantity

$$\begin{aligned} r_{(4)} &= a_{(4)} - a_{(4')} - 2(\xi_m - \xi_n) a_\xi(t,x,\xi_m) \\ &= \sum_{\pm} a(t,x,\xi_{m\pm i}) - a(t,x,\xi_{n\pm i}) - (\xi_{m\pm i} - \xi_{n\pm i})\, a_\xi(t,x,\xi_m)\,, \end{aligned}$$

which by a Taylor expansion in ξ is seen to satisfy

$$\bigl|\partial^\alpha_{t,x} r_{(4)}\bigr| \lesssim \lambda^{\frac{1}{3}+\frac{2}{3}\max(0,|\alpha|-1)}(i + |m-n|)\,|m-n|\,.$$

Using a Taylor expansion of $a(t,x,\xi)$ about $\xi = \xi_T$, we can write

$$a^w(t,x,D)v_T = \bigl[a(t,x,\xi_T) + a_\xi(t,x,\xi_T)(D - \xi_T) + r^w(t,x,D)\bigr]v_T\,,$$

where $\lambda^{-\frac{1}{3}} r^w(t,x,D)$ applied to a $\lambda^{\frac{2}{3}}$-scaled Schwartz function with frequency center ξ_T, such as v_T, yields a Schwartz function of comparable norm, uniformly over λ.

We then write

$$\begin{aligned} a_\xi(t,x,\xi_T)(D - \xi_T)v_T = \bigl[a_\xi(t,x,\xi_T) - a_\xi(t,x,\xi_m)\bigr](D - \xi_T)v_T& \\ + a_\xi(t,x,\xi_m)(D - \xi_T)v_T\,.& \end{aligned}$$

The same expansions hold with ξ_T replaced by ξ_S, $\xi_{T'}$, and $\xi_{S'}$.

Replacing ξ_T by any of $\xi_{m\pm j}$ or $\xi_{n\pm i}$, the function $a_\xi(t,x,\xi_T) - a_\xi(t,x,\xi_m)$ satisfies

$$\bigl|\partial^\alpha_{t,x}\bigl(a_\xi(t,x,\xi_T) - a_\xi(t,x,\xi_m)\bigr)\bigr| \lesssim (i + j + |m-n|)\lambda^{-\frac{1}{3}+\frac{2}{3}\max(0,|\alpha|-1)}\,.$$

Let $L = D_t - a_\xi(t,x,\xi_m)D$, where $D = D_x$ as always. Writing $D_t v_T = a^w(t,x,D)v_T$, and using the above expansion for the latter, we see that $(L - a_{(4')})v_{(4)}$ can be written as a sum of 5 terms,

$$\begin{aligned} (L - a_{(4')})v_{(4)} = (L_T v_T)\, v_S\, \overline{v_{T'}}\, \overline{v_{S'}} + v_T\,(L_S v_S)\, \overline{v_{T'}}\, \overline{v_{S'}} - v_T\, v_S\, \overline{(L_{T'} v_{T'})}\, \overline{v_{S'}}& \\ - v_T\, v_S\, \overline{v_{T'}}\, \overline{(L_{S'} v_{S'})} + r_{(4)} v_{(4)}\,,& \end{aligned}$$

where we wrote $\xi_T + \xi_S - \xi_{T'} - \xi_{S'} = 2(\xi_m - \xi_n)$, and where

$$\begin{aligned} L_T v_T &= \Big(a^w(t,x,D) - a(t,x,\xi_T) - a_\xi(t,x,\xi_m)(D-\xi_T)\Big) v_T \\ &= \Big(\big[a_\xi(t,x,\xi_T) - a_\xi(t,x,\xi_m)\big](D-\xi_T) + r^w(t,x,D)\Big) v_T\,. \end{aligned}$$

In the expressions for L_S, $L_{T'}$, and $L_{S'}$, T is respectively replaced by S, T', and S', but the ξ_m is the same for each. The L_T's thus depend on all 4 subscripts, but this is fine since the below analysis is applied separately to each term.

Since $(D-\xi_T)$ applied to v_T counts as $\lambda^{\frac23}$, then $|L_T v_T| \lesssim \lambda^{\frac13}(i+j+|m-n|)\lambda^{\frac13}\chi_T$. Indeed, $L_T v_T$ can be written, at each fixed time t, as $\lambda^{\frac13}(i+j+|m-n|)$ times a Schwartz function of the same scale and phase space center as v_T.

Consequently,

$$\big|(L - a_{(4')})v_{(4)}\big| \lesssim \lambda^{\frac13}(i+j+|m-n|)\,\langle m-n\rangle\,\lambda^{\frac43}\chi_{(4)}\,,$$

where $\chi_{(4)}$ is the product of the corresponding χ_T.

We also need the estimate

$$\big|(L - a_{(4')})^2 v_{(4)}\big| \lesssim \lambda^{\frac23}(i+j+|m-n|)^2\langle m-n\rangle^2\lambda^{\frac43}\chi_{(4)}\,.$$

Following the above arguments, we can write $(L-a_{(4')})^2 v_{(4)}$ as a sum of 16 terms like

$$(L_T^2 v_T)\,v_S\,\overline{v_{T'}}\,\overline{v_{S'}} + (L_T v_T)\,(L_S v_S)\,\overline{v_{T'}}\,\overline{v_{S'}} - (L_T v_T)\,v_S\,\overline{(L_{T'}v_{T'})}\,\overline{v_{S'}} - \cdots$$

plus 4 commutator terms

$$\big(\,\big[D_t - a^w(t,x,D)\,,L_T\big]v_T\big)\,v_S\,\overline{v_{T'}}\,\overline{v_{S'}} + v_T\,\big(\,\big[D_t - a^w(t,x,D)\,,L_S\big]v_S\big)\,\overline{v_{T'}}\,\overline{v_{S'}} - \cdots$$

plus remainder terms

$$(Lr_{(4)})\,v_{(4)} + r_{(4)}(L - a_{(4')})v_{(4)}\,.$$

Except for the commutator terms, the desired bounds follow by the same estimates as for $(L-a_{(4')})v_{(4)}$. The commutator terms depend on the fact that the commutator of two symbols in $C^1 S_{\lambda,\lambda^{2/3}}$ is in $\lambda^{-1}S_{\lambda,\lambda^{2/3}}$.

We let ϕ be a smooth cutoff to a $\lambda^{-\frac13}$ time interval in t. We can then write

$$\begin{aligned} \int v_{(4)}\,\phi\,dt\,dx = \int \frac{\big(L-a_{(4')}\big)^2 v_{(4)}}{a_{(4')}^2}\,\phi\,dt\,dx + \int \frac{\big(\phi(L'a_{(4')}) - a_{(4')}(L'\phi)\big)\,v_{(4)}}{a_{(4')}^2}\,dt\,dx \\ - \int \frac{\big(2\phi(L'a_{(4')}) - a_{(4')}(L'\phi)\big)\big(L-a_{(4')}\big)\,v_{(4)}}{a_{(4')}^3}\,dt\,dx\,, \end{aligned}$$

where we have integrated by parts in the last two terms, and L' is the transpose of L. We assume here that the v_T are extended to a slightly larger time interval to allow integration by parts in t.

By the above estimates the integrand of the first term on the right is bounded by

$$\frac{(i+j+|m-n|)^2\langle m-n\rangle^2\lambda^{\frac43}\chi_{(4)}}{(j^2-i^2)^2} \le \frac{\langle m-n\rangle^4}{\langle\, j-i\,\rangle^2}\,\lambda^{\frac43}\chi_{(4)}\,.$$

By (6.3) and (6.4), we have $(L'a_{(4')}) \lesssim a_{(4')}$. The integrands of the last two terms on the right hand side are then respectively dominated by

$$\frac{\lambda^{\frac{4}{3}}\chi_{(4)}}{(j^2-i^2)} + \frac{(i+j+|m-n|)\langle m-n\rangle \lambda^{\frac{4}{3}}\chi_{(4)}}{(j^2-i^2)^2} \le \frac{\langle m-n\rangle^2}{\langle j-i\rangle^2}\,\lambda^{\frac{4}{3}}\chi_{(4)}\,.$$

Consequently, we have shown that we can write $\int v_{(4)}\phi\,dt\,dx = \int w_{(4)}\,dt\,dx$, where

$$|w_{(4)}| \lesssim \frac{\langle m-n\rangle^4}{\langle j-i\rangle^2}\,\lambda^{\frac{4}{3}}\chi_{(4)}\,.$$

Given a collection Λ of pairs of tubes (T,S), we let

$$b_n = \Big(\sum_{\xi_T=\xi_n}|b_T|^2\Big)^{\frac{1}{2}}, \qquad d_n = \Big(\sum_{\xi_S=\xi_n}|d_S|^2\Big)^{\frac{1}{2}},$$

where the sum is over all T, respectively S, in the collection that have frequency center ξ_n. Then

$$\begin{aligned}\Big\|\phi\sum_{(T,S)\in\Lambda} b_T v_T\cdot d_S v_S\Big\|_{L^2}^2 &= \sum_{(T,S)\in\Lambda}\ \sum_{(T',S')\in\Lambda} b_T\,d_S\,\overline{b_{T'}}\,\overline{d_{S'}}\int v_{(4)}\phi\,dt\,dx\\ &\le \sum_{m,j}\sum_{n,i}\ \sum_{(T,S)\in\Lambda_{m,j}}\ \sum_{(T',S')\in\Lambda_{n,i}}\big|\,b_T\,d_S\,\overline{b_{T'}}\,\overline{d_{S'}}\,\big|\int w_{(4)}\,dt\,dx\,,\end{aligned}$$

where $\Lambda_{m,j}\subseteq\Lambda$ consists of the pairs $(T,S)\in\Lambda$ such that $\xi_T=\xi_{m+j}$, and $\xi_S=\xi_{m-j}$.

By the above this is dominated by

$$\begin{aligned}&\lambda^{\frac{4}{3}}\sum_{m,n,i,j}\frac{\langle m-n\rangle^4}{\langle\, i-j\,\rangle^2}\int\Big(\sum_{(T,S)\in\Lambda_{m,j}}|\,b_T\,d_S|\,\chi_T\chi_S\Big)\Big(\sum_{(T',S')\in\Lambda_{n,i}}|\,b_{T'}\,d_{S'}|\,\chi_{T'}\chi_{S'}\Big)\,dt\,dx\\ &\qquad\lesssim \theta^{-1}\sum_{m,n,i,j}\frac{\langle m-n\rangle^2}{\langle i-j\rangle^2}b_{m+j}\,d_{m-j}\,b_{n+i}\,d_{n-i}\,,\end{aligned}$$

where we used the Cauchy-Schwartz inequality and (6.2).

We next show that we may write $\int v_{(4)}\phi\,dt\,dx = \int w_{(4)}\,dt\,dx$, where

$$|w_{(4)}| \lesssim \frac{1}{\langle m-n\rangle^{18}}\,\lambda^{\frac{4}{3}}\chi_{(4)}\,. \tag{6.5}$$

Since

$$\min\Big(\frac{\langle m-n\rangle^4}{\langle\, i-j\,\rangle^2}\,,\ \frac{1}{\langle m-n\rangle^{18}}\Big) \le \frac{1}{\langle\, i-j\,\rangle^{\frac{3}{2}}\langle m-n\rangle^{\frac{3}{2}}}\,,$$

this will establish that

$$\Big\|\phi\sum_{(T,S)\in\Lambda} b_T v_T\cdot d_S v_S\Big\|_{L^2}^2 \lesssim \theta^{-1}\sum_{m,n,i,j}\frac{1}{\langle\, i-j\,\rangle^{\frac{3}{2}}\langle m-n\rangle^{\frac{3}{2}}}b_{m+j}\,d_{m-j}\,b_{n+i}\,d_{n-i}\,.$$

By Schur's lemma, this is in turn bounded by

$$\theta^{-1}\Big(\sum_{m,j}b_{m+j}^2\,d_{m-j}^2\Big)^{\frac{1}{2}}\Big(\sum_{n,i}b_{n+i}^2\,d_{n-i}^2\Big)^{\frac{1}{2}} \le \theta^{-1}\Big(\sum_T|b_T|^2\Big)\Big(\sum_S|d_S|^2\Big)\,,$$

where the last sum is over all T,S that occur in Λ.

To prove (6.5), we write

$$2(m-n)\int v_{(4)}\phi\,dt\,dx = \int \lambda^{-\frac{2}{3}}(\xi_T+\xi_S-\xi_{T'}-\xi_{S'})\,v_{(4)}\phi\,dt\,dx = \int w_{(4)}\phi\,dt\,dx\,,$$

where

$$w_{(4)} = (\lambda^{-\frac{2}{3}}(D-\xi_T)\,v_T)\,v_S\,\overline{v_{T'}}\,\overline{v_{S'}} + v_T\,(\lambda^{-\frac{2}{3}}(D-\xi_S)\,v_S)\,\overline{v_{T'}}\,\overline{v_{S'}} - \cdots$$

We repeat this process, and use that $|\lambda^{-\frac{2k}{3}}(D-\xi_T)^k v_T| \lesssim \lambda^{\frac{1}{3}}\chi_T$. □

7. Results for dimension $d \geq 3$

In this section we work on a compact d-dimensional manifold M without boundary. We consider spectral clusters for $\mathrm{g},\rho \in \mathrm{Lip(M)}$ exactly as in Theorem 1.1. We will apply the methods and results of the previous sections to prove the following, which establishes the conjectured result (1.4) for a partial range of p.

Theorem 7.1. *Let u be a spectral cluster on M, where M is of dimension $d \geq 3$. Then*

$$\|u\|_{L^p(M)} \leq C_p\,\lambda^{d(\frac{1}{2}-\frac{1}{p})-\frac{1}{2}}\|u\|_{L^2(M)}\,, \qquad \tfrac{6d-2}{d-1} < p \leq \infty\,. \tag{7.1}$$

The following partial range result for the other estimate in (1.4),

$$\|u\|_{L^p(M)} \leq C\,\lambda^{\frac{2(d-1)}{3}(\frac{1}{2}-\frac{1}{p})}\,\|f\|_{L^2(M)}\,, \qquad 2 \leq p \leq \tfrac{2(d+1)}{d-1}\,,$$

was established in [7], as was the $p=\infty$ case of (7.1).

The proof of Theorem 7.1 follows the same steps as Theorem 1.1, and so we focus below on the modifications necessary in each step.

Step 1: *Reduction to a frequency localized first order problem.* Care must be taken in the frequency localization step to handle the high-frequency terms, since Sobolev embedding as used in the $d=2$ case is not sufficient to establish the desired result for large p in high dimensions. In particular, the analogue of Theorem 2.1 does not hold for $p=\frac{6d-2}{d-1}$. Instead, we use the following estimate from [7], valid for Lipschitz g,ρ,

$$\|u\|_{L^\infty(M)} \lesssim \lambda^{\frac{d-1}{2}}\|u\|_{L^2(M)}\,. \tag{7.2}$$

We remark that this estimate used the strict spectral localization of u and intrinsic Sobolev embedding on M to deduce it from results for smaller p.

By (7.2), if ϕ is a bump function supported in a local coordinate patch, and

$$\phi u = (\phi u)_{<\lambda} + (\phi u)_\lambda + (\phi u)_{>\lambda}$$

is the decomposition of ϕu into terms with local-coordinate frequencies respectively less than $c\lambda$, comparable to λ, and greater than $c^{-1}\lambda$, then each term in the decomposition has L^∞ norm bounded by $\lambda^{\frac{d-1}{2}}\|u\|_{L^2(M)}$. The proof of [7, Corollary 5], together with (2.1) and Sobolev embedding on $\mathbb{R}^d$, yields

$$\|(\phi u)_{<\lambda}\|_{L^{\frac{2d}{d-2}}} + \|(\phi u)_{>\lambda}\|_{L^{\frac{2d}{d-2}}} \lesssim \|u\|_{L^2(M)}\,.$$

Interpolation with (7.2) then yields even better bounds than those of Theorem 7.1 for these terms.

Hence we are reduced to bounding $\|(\phi u)_\lambda\|_{L^p}$. With $a^w_\lambda(t,x,D)$ and $S(t,t_0)$ defined as they are for $d=2$, where x and ξ are now of dimension $d-1$, we then reduce matters as before to establishing

$$\|u\|_{L^p([0,1]\times\mathbb{R}^{d-1})} \le C_p\,\lambda^{d(\frac12-\frac1p)-\frac12}\|u_0\|_{L^2(\mathbb{R}^{d-1})}\,,\quad u=S(t,0)u_0\,,\quad \tfrac{6d-2}{d-1}<p\le\infty\,,$$

with $\widehat{u}_0$ supported in $|\xi|\le\frac34\lambda$. As before we will take $\|u_0\|_{L^2(\mathbb{R}^{d-1})}=1$.

The expansion of u in terms of tube solutions v_T on each $\lambda^{-\frac13}$ time slab and the definition of $A_{a,k,m}$ bushes then proceeds for $d\ge3$ the same as for $d=2$, but where we take $\epsilon=\lambda^{-\frac{d-1}{3}}$ as the lower bound for a in the sum $u=\sum u_{a,k}$ in order to trivially obtain the desired bounds for u_ϵ.

In d-dimensions, a 2^m-bush has angular spread at least $2^{\frac{m}{d-1}}\lambda^{-\frac13}$, and so can retain full overlap for time $\delta t=2^{-\frac{m}{d-1}}\lambda^{-\frac13}$. Thus, in dimension d we decompose the unit time interval into a collection $\mathcal{I}_m$ of intervals of size $\delta t=2^{-\lfloor\frac{m}{d-1}\rfloor}\lambda^{-\frac13}$; such intervals are then dyadic subintervals of the decomposition into $\lambda^{-\frac13}$ time slabs.

The proof of Theorem 7.1 will then be concluded using the following two propositions.

Proposition 7.2. *There are at most $\lambda^{\frac13}2^{-\frac{m}{d-1}}(2^ma^2)^{-2}\langle\log(2^ma^2)\rangle$ intervals $I\in\mathcal{I}_m$ which intersect $A_{a,k,m}$.*

Proposition 7.3. *For each interval $I\in\mathcal{I}_m$, we have*

$$\|u_{a,k}\|^p_{L^p(A_{a,k,m}\cap I\times\mathbb{R}^{d-1})}\lesssim\lambda\big(\lambda^{\frac{d-1}{3}}2^ma\big)^{p-p_d}2^{-\frac{kp_d}{2}}\,,\qquad p\ge p_d=\tfrac{2(d+1)}{d-1}\,.$$

Indeed, combining the two propositions we obtain

$$\begin{aligned}\|u_{a,k}\|^p_{L^p(A_{a,k,m})}&\lesssim\lambda\big(\lambda^{\frac{d-1}{3}}2^ma)^{p-p_d}\lambda^{\frac13}2^{-\frac{m}{d-1}}(2^ma^2)^{-2}\langle\log(2^ma^2)\rangle2^{-\frac{kp_d}{2}}\\&\lesssim\lambda^{\frac43+\frac{d-1}{3}(p-p_d)}2^{m(\frac{p-p_d}{2}-\frac{1}{d-1})}(2^ma^2)^{\frac{p-p_d}{2}-2}\langle\log(2^ma^2)\rangle2^{-\frac{kp_d}{2}}\,.\end{aligned}$$

Recall that 2^m and a both take on dyadic values such that

$$2^ma^2\lesssim1\,,\qquad 2^m\lesssim\lambda^{\frac{d-1}{3}}\,.$$

When the exponent of 2^ma^2 is positive,

$$p>p_d+4=\tfrac{6d-2}{d-1}\,,$$

we may sum over m, a and k to obtain

$$\|u\|^p_{L^p}\le C_p^p\,\lambda^{1+\frac{d-1}{2}(p-p_d)}\,,$$

giving the desired result

$$\|u\|_{L^p}\le C_p\,\lambda^{\frac{d-1}{2}-\frac{d}{p}}\,.$$

By (7.2) the constant C_p remains bounded as $p\to\infty$, but may diverge as $p\to\frac{6d-2}{d-1}$. On the other hand C_p is bounded by a power of $\log\lambda$ for $p=\frac{6d-2}{d-1}$, since there are only $\approx\log\lambda$ terms in each index.

Proof of Proposition 7.2. There are $\approx 2^{\frac{m}{d-1}}\lambda^{\frac{1}{3}}$ intervals in $\mathcal{I}_m$, so we may assume that $a^2 \gg 2^{-m(1+\frac{1}{d-1})}$. It suffices to prove, for ϵ a fixed small number, that if among $\epsilon\, 2^{m(1+\frac{2}{d-1})}a^2$ consecutive slices in $\mathcal{I}_m$ there are M slices that intersect $A_{a,k,m}$, then

$$M \lesssim (2^m a^2)^{-1}\langle \log(2^m a^2)\rangle\,. \tag{7.3}$$

Consider a collection $\{B_n\}_{n=1}^M$ of M distinct (a,k,m)-bushes, centered at (t_n,x_n), such that

$$\epsilon\,\lambda^{-\frac{1}{3}}2^{m(1+\frac{1}{d-1})}a^2 \geq |t_n - t_{n'}| \geq \lambda^{-\frac{1}{3}}2^{-\frac{m}{d-1}} \quad \text{when} \quad n \neq n'\,.$$

Denote by $\{v_{n,l}\}_{l=1,2^m}$ the collection of 2^m packets in B_n. As in the proof of Proposition 2.4, for each n we define the bounded projection operators P_n on $L^2(\mathbb{R}^{d-1})$ at time t_n by

$$P_n f = 2^{-m}a^2\left(\sum_l \overline{c_{n,l}}(t_n)^{-1}v_{n,l}(t_n,\cdot\,)\right)\left(\sum_l c_{n,l}(t_n)^{-1}\langle v_{n,l}(t_n,\cdot\,), f\rangle\right)$$

where we recall that $|c_{n,l}(t_n)| \approx a$, so that

$$\|P_n u(t_n,\cdot\,)\|^2_{L^2(\mathbb{R}^{d-1})} \approx 2^m a^2\,. \tag{7.4}$$

As with Proposition 2.4, the key estimate is the following analogue of Lemma 3.1,

Lemma 7.4. *Let* $\alpha = \max\bigl(\lambda^{-\frac{1}{3}}|t_{n'}-t_n|^{-1}, \lambda^{\frac{1}{3}}|t_{n'}-t_n|\bigr)$*. Then the operators* P_n *satisfy*

$$\|P_{n'}S(t_{n'},t_n)P_n\|_{L^2(\mathbb{R}^{d-1})\to L^2(\mathbb{R}^{d-1})} \lesssim 2^{-\frac{m}{d-1}}\alpha\,. \tag{7.5}$$

Proof. We follow the proof of Lemma 3.1 at the end of Section 5. The same steps follow, where the q_j are vector valued if $d \geq 3$. The analogue of (5.5) to be proven is

$$\bigl\|m_0(D)^{-\frac{1}{2}}\langle \delta^{-2}(x-x_1)\rangle^2 w_1\|^2_{L^2(\mathbb{R})} \lesssim 2^{m(1-\frac{1}{d-1})}\alpha\,,$$

which is established by comparison to the worst case sum, where $j \in \mathbb{Z}^{d-1}$,

$$\sum_{|j|\leq 2^{\frac{m}{d-1}}}(1+\alpha^{-1}|j|)^{-1} \lesssim \alpha\int_0^{2^{\frac{m}{d-1}}} r^{d-3}\,dr \lesssim 2^{m(1-\frac{1}{d-1})}\alpha\,,$$

where we used that $\alpha \geq 1$ to handle the $j=0$ term. $\square$

As before, Lemma 7.4 leads to the bound

$$\sum_n \|P_n u\|_{L^2(\mathbb{R}^{d-1})} \lesssim C^{\frac{1}{2}}\|u_0\|_{L^2(\mathbb{R})}\,, \qquad C = M + \sum_{n,n'} 2^{-\frac{m}{d-1}}\alpha\,. \tag{7.6}$$

Comparing (7.4) and (7.6) applied to u, it follows that

$$2^m a^2 M^2 \lesssim M + \sum_{n\neq n'} 2^{-\frac{m}{d-1}}\alpha\,. \tag{7.7}$$

The bound (7.3) is trivial if $2^m a^2 M^2 \lesssim M$, so we consider the summation term. For the sum over terms where $\lambda^{\frac{1}{3}}|t_{n'}-t_n| \geq 1$ we have

$$\sum_{n\neq n'} 2^{-\frac{m}{d-1}}\lambda^{\frac{1}{3}}|t_{n'}-t_n| \lesssim \epsilon\, 2^m a^2 M^2\,.$$

Taking ϵ small we can thus absorb these terms into the left hand side of (7.7).

To conclude, we may assume then that

$$2^m a^2 M^2 \lesssim 2^{-\frac{m}{d-1}} \lambda^{-\frac{1}{3}} \sum_{n \neq n'} |t_{n'} - t_n|^{-1} .$$

We use the $2^{-\frac{m}{d-1}} \lambda^{-\frac{1}{3}}$ separation of the t_n's to bound

$$\sum_{n \neq n'} |t_{n'} - t_n|^{-1} \lesssim 2^{\frac{m}{d-1}} \lambda^{\frac{1}{3}} M \log M ,$$

thus

$$2^m a^2 M \lesssim \log M ,$$

or

$$M \lesssim (2^m a^2)^{-1} \langle \log(2^m a^2) \rangle .$$

□

Proof of Proposition 7.3. We estimate $\|u_{a,k}\|_{L^p(A_{a,k,m})}$ on a single slice $I \times \mathbb{R}^{d-1}$, where $I \in \mathcal{I}_m$. By (4.1), if there are N terms in the sum for $u_{a,k}$, then $Na^2 \lesssim 2^{-k}$, so by orthogonality the total energy of $u_{a,k}$ is $\lesssim 2^{-\frac{k}{2}}$. We then have the Strichartz estimates

$$\|u_{a,k}\|_{L^{p_d}(I \times \mathbb{R}^{d-1})} \lesssim \lambda^{\frac{1}{p_d}} 2^{-\frac{k}{2}} , \qquad p_d = \tfrac{2(d+1)}{d-1} .$$

If $k > \frac{m}{d-1}$, this is proven on each $2^{-k} \lambda^{-\frac{1}{3}}$ dyadic subinterval of I then summed.

We interpolate this with the L^∞ bound

$$\|u_{a,k}\|_{L^\infty(A_{a,k,m} \cap I \times \mathbb{R}^{d-1})} \lesssim \lambda^{\frac{d-1}{3}} 2^m a ,$$

to obtain

$$\|u_{a,k}\|^p_{L^p(A_{a,k,m} \cap I \times \mathbb{R}^{d-1})} \lesssim \lambda (\lambda^{\frac{d-1}{3}} 2^m a)^{p-p_d} 2^{-\frac{kp_d}{2}} . \quad \square$$

References

[1] M. Blair, H. Smith, and C. Sogge, *On Strichartz estimates for Schrödinger operators in compact manifolds with boundary*, Proc. Amer. Math. Soc., **136** (2008), 247–256.

[2] R. Coifman and Y. Meyer, Commutateurs d'integrales singulieres et operateurs multilineaires, Ann. Inst. Fourier Grenoble **28** (1978), 177–202.

[3] L. Hörmander, *The Weyl calculus of pseudo-differential operators*, Comm. Pure Appl. Math. **32** (1979), 359–443.

[4] H. Koch and D. Tataru, *Dispersive estimates for principally normal operators*, Comm. Pure Appl. Math. **58** (2005), 217–284.

[5] H. Koch, D. Tataru and M. Zworski, *Semiclassical L^p estimates*, Ann. Henri Poincaré **8** (2007), no. 5, 885–916.

[6] H. Smith, *Spectral cluster estimates for $C^{1,1}$ metrics*, Amer. Jour. Math. **128** (2006), 1069–1103.

[7] ———, *Sharp $L^2 \to L^q$ bounds on spectral projectors for low regularity metrics*, Math. Res. Lett. **13** (2006), 965–972.

[8] H. Smith and C. Sogge, *On Strichartz and eigenfunction estimates for low regularity metrics*, Math. Res. Lett. **1** (1994), 729–737.

[9] H. Smith and D. Tataru, *Sharp counterexamples for Strichartz estimates for low regularity metrics*, Math. Res. Lett. **9** (2002), 199–204.

[10] C. Sogge, *Concerning the L^p norm of spectral clusters for second order elliptic operators on compact manifolds*, J. Funct. Anal. **77** (1988), 123–134.

[11] D. Tataru, *Strichartz estimates for operators with nonsmooth coefficients III*, J. Amer. Math. Soc. **15** (2002), 419–442.

[12] M. Taylor, Pseudodifferential Operators and Nonlinear PDE, Progress in Mathematics, vol. 100, Birkhäuser, Boston, 1991.

Mathematisches Institute, University of Bonn, Beringstrasse 1, 53115 Bonn, Germany

E-mail address: koch@math.uni-bonn.de

Department of Mathematics, University of Washington, Seattle, WA 98195

E-mail address: hart@math.washington.edu

Department of Mathematics, University of California, Berkeley, CA 94720

E-mail address: tataru@math.berkeley.edu