



Logarithmic improvements in L^p bounds for eigenfunctions at the critical exponent in the presence of nonpositive curvature

Matthew D. Blair¹ · Christopher D. Sogge²

Received: 23 June 2017 / Accepted: 20 March 2019 / Published online: 29 March 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract We consider the problem of proving L^p bounds for eigenfunctions of the Laplacian in the high frequency limit in the presence of nonpositive curvature and more generally, manifolds without conjugate points. In particular, we prove estimates at the “critical exponent” $p_c = \frac{2(d+1)}{d-1}$, where a spectrum of scenarios for phase space concentration must be ruled out. Our work establishes a gain of an inverse power of the logarithm of the frequency in the bounds relative to the classical L^p bounds of the second author.

1 Introduction

Let (M, g) be a boundaryless, compact, connected Riemannian manifold with $d = \dim(M) \geq 2$ and Δ_g the associated negative Laplace–Beltrami operator. The spectrum of $-\Delta_g$ is discrete and we let e_λ denote any L^2 -normalized eigenfunction

The first author was supported in part by the National Science Foundation Grants DMS-1301717 and DMS-1565436, and the second by the National Science Foundation Grant DMS-1665373.

✉ Matthew D. Blair
blair@math.unm.edu

Christopher D. Sogge
sogge@jhu.edu

¹ Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131, USA

² Department of Mathematics, Johns Hopkins University, Baltimore, MD 21093, USA

$$(\Delta_g + \lambda^2)e_\lambda = 0, \quad \|e_\lambda\|_{L^2(M)} = 1. \quad (1.1)$$

Here $L^p(M)$ is the space of L^p functions with respect to Riemannian measure dV_g . The frequency λ thus parameterizes the eigenvalues of

$$P := \sqrt{-\Delta_g}.$$

We are concerned with L^p bounds on eigenfunctions e_λ in the high frequency limit $\lambda \rightarrow \infty$ and more generally, “spectral clusters”, meaning sums of eigenfunctions in the range of a spectral projector $\mathbf{1}_{[\lambda, \lambda+h(\lambda)]}(P)$, the operator which projects a function onto all the eigenspaces of P whose corresponding eigenvalue lies in a band of width $h(\lambda)$ to the right of λ . In [25], the second author showed that for $2 < p \leq \infty$, with $h(\lambda) \equiv 1$

$$\|\mathbf{1}_{[\lambda, \lambda+1]}(P)\|_{L^2(M) \rightarrow L^p(M)} \lesssim \lambda^{\delta(p, d)}, \quad \lambda \geq 1, \quad (1.2)$$

$$\delta(p, d) = \begin{cases} \frac{d-1}{2} - \frac{d}{p}, & p_c \leq p \leq \infty, \\ \frac{d-1}{2} \left(\frac{1}{2} - \frac{1}{p} \right), & 2 \leq p \leq p_c, \end{cases} \quad p_c := \frac{2(d+1)}{d-1}. \quad (1.3)$$

Note that $\delta(p_c, d) = 1/p_c$. The case $p = \infty$ here can be seen as a consequence of classical pointwise Weyl laws. One of the key contributions of [25] was to treat these bounds at the “critical” exponent p_c , so that interpolation yields the remaining cases. This gives a discrete analog of the Stein-Tomas Fourier restriction bound for the sphere [26, p.135] (or more precisely the adjoint bound). Given (1.2), any eigenfunction as in (1.1) satisfies

$$\|e_\lambda\|_{L^p(M)} \lesssim \lambda^{\delta(p, d)}, \quad \lambda \geq 1. \quad (1.4)$$

As observed in [24], the exponent $\delta(p, d)$ in (1.4) cannot be improved when (M, g) is the round sphere. The zonal harmonics provide a sequence of eigenfunctions saturating the bound when $p_c \leq p \leq \infty$ and the highest weight spherical harmonics saturate it when $2 < p \leq p_c$. This is not surprising since the spectrum of P in this setting is nearly arithmetic, meaning the projector in (1.2) is essentially the same as projection onto an eigenspace. However, the geometries for which the corresponding eigenfunctions saturate (1.4) are in some sense exceptional, and if it does occur then the geodesic flow expects to have similar dynamics to that of the sphere. Well known classical Gaussian beam constructions show that when (M, g) has a stable elliptic orbit, then there are highly accurate approximate eigenfunctions that saturate (1.2) when $2 < p \leq p_c$. The works [30, 31, 33] characterize geometries which saturate (1.4) when $p = \infty$, showing, for instance, that in the real analytic case, this will only occur if the unit speed geodesics emanating from a point $x \in M$

loop back to a point at a common time. These features are absent from several Riemannian manifolds of interest such as manifolds without conjugate points.

When (M, g) has nonpositive sectional curvatures, it is known that

$$\|\mathbf{1}_{[\lambda, \lambda + (\log \lambda)^{-1}]}(P)\|_{L^2(M) \rightarrow L^p(M)} \lesssim_p \frac{\lambda^{\delta(p,d)}}{(\log \lambda)^{\sigma(p,d)}}, \quad p \neq p_c, \quad (1.5)$$

for some exponent $\sigma(p, d) > 0$. For $p_c < p \leq \infty$, a work of Hassell and Tacy [19] shows that one can take $\sigma(p, d) = \frac{1}{2}$ here,¹ though the implicit constant tends to infinity as $p \searrow p_c$. Their work draws from a classical work of Berard [1] on the remainder in the pointwise Weyl law (which already implies the $p = \infty$ case). When $2 < p < p_c$, the bounds (1.5) result from the authors' works [9, 10], but the exponents obtained satisfy $\lim_{p \rightarrow p_c^-} \sigma(p, d) = 0$, again leaving the critical $p = p_c$ case open. In particular, [9] shows that the left hand side of (1.5) is dominated by so-called “Kakeya-Nikodym” averages, which bound the mass of these spectral clusters within shrinking tubular neighborhoods about a geodesic segment. The work [10] then shows that these averages are then seen to exhibit a logarithmic gain in the presence of nonpositive curvature (cf. (5.6) below).

The two strategies outlined here are therefore very effective towards obtaining a logarithmic gain in the $L^2(M) \rightarrow L^p(M)$ bounds on the projector in (1.5) when $p \neq p_c$: they either rule out mass concentration similar to the zonal harmonics, yielding improvements for $p_c < p \leq \infty$, or concentration similar to the highest weight spherical harmonics, yielding improvements for $2 < p < p_c$. However, by themselves they do not seem to give an effective strategy for obtaining a logarithmic gain at the critical exponent p_c .

We also remark that a work of Hezari and Rivière [18] uses quantum ergodicity methods to show a logarithmic gain in the L^p bounds on eigenfunctions in the presence of negative curvature for all $2 < p \leq \infty$, but this is only for a full density subsequence of them. In the present work, we are interested in bounds satisfied by the full sequence.

A breakthrough on this critical problem came from the second author in [29], who demonstrated a gain of an inverse power of $\log \log \lambda$ in the $L^2(M) \rightarrow L^{p_c}(M)$ bounds on this spectral projector. The strategy there was to instead consider bounds on the projector in weak- L^p spaces, which in turn yields strong L^p bounds after interpolation with Lorentz space bounds of Bak and Seeger [6]. The weak bounds were then treated by methods analogous to Bourgain's approach to Fourier restriction to the sphere in [4, §6]. We outline the strategy of [29] in Sect. 2 below.

¹ The works [1, 3] show this also holds if (M, g) merely lacks conjugate points.

In the present work, we show that the $\log \log \lambda$ gain exhibited in [29] can be improved to a $\log \lambda$ gain. This is significant as the latter essentially corresponds to the largest time scale over which the frequency localized wave kernel is currently understood in the setting of nonpositive curvature, closely related to considerations involving the “Ehrenfest time” in quantum mechanics. In what follows, ρ is an even, real valued function satisfying

$$\rho \in \mathcal{S}(\mathbb{R}), \quad \rho(0) = 1, \quad \text{supp}(\widehat{\rho}) \subset \{|t| \in (1/4, 1/2)\}.$$

We also assume that $|\rho(t)| \leq 1$ for every $t \in \mathbb{R}$ so that for any $\tau, \lambda > 0$

$$\|\rho(\tau(\lambda - P))\|_{L^2(M) \rightarrow L^2(M)} \leq 1. \quad (1.6)$$

Throughout the work, we let $c_0 > 0$ be a sufficiently small but fixed constant and define $T = T(\lambda)$ by

$$T := c_0 \log \lambda. \quad (1.7)$$

Theorem 1.1 *Suppose (M, g) has nonpositive sectional curvatures. There exists $\varepsilon_0 > 0$ such that*

$$\|\rho(T(\lambda - P))\|_{L^2(M) \rightarrow L^{p_c}(M)} \lesssim \frac{\lambda^{1/p_c}}{(\log \lambda)^{\varepsilon_0}}, \quad p_c = \frac{2(d+1)}{d-1}. \quad (1.8)$$

Consequently,

$$\|\mathbf{1}_{[\lambda, \lambda + (\log \lambda)^{-1}]}(P)\|_{L^2(M) \rightarrow L^{p_c}(M)} \lesssim \frac{\lambda^{1/p_c}}{(\log \lambda)^{\varepsilon_0}}. \quad (1.9)$$

and in particular, any eigenfunction as in (1.1) satisfies

$$\|e_\lambda\|_{L^{p_c}(M)} = \mathcal{O}(\lambda^{1/p_c} (\log \lambda)^{-\varepsilon_0}).$$

That (1.9) follows from (1.8) is standard. Indeed, taking c_0 sufficiently small, $\rho(T(\lambda - P))$ is invertible on the range of $\mathbf{1}_{[\lambda, \lambda + (\log \lambda)^{-1}]}(P)$ with inverse uniformly bounded on $L^2(M)$. We are thus focused on proving (1.8).

Remark 1.2 The argument shows that in fact

$$\varepsilon_0 = \begin{cases} \frac{4}{3(d+1)^3}, & d \geq 4, \\ \frac{1}{48} - & d = 3 \\ \frac{1}{72}, & d = 2, \end{cases} \quad (1.10)$$

where in the $d = 3$ case the minus sign means that the exponent can be taken strictly less than but arbitrarily close to $1/48$. As noted in Remark 2.4 below, this exponent can be improved when $d = 2, 3$ if it is assumed that (M, g) has strictly negative sectional curvatures, leading to an exponent of $\varepsilon_0 = \frac{1}{36}$ when $d = 2$ and a removal of the loss when $d = 3$ so that $\varepsilon_0 = \frac{1}{48}$.

To gain an appreciation as to why treating the case of “critical” exponents is subtle, it is helpful to consider the analog of (1.2) for the constant coefficient Laplacian on \mathbb{R}^d , which amounts to considering the Fourier multiplier onto frequencies $\{\xi \in \mathbb{R}^d : \lambda \leq |\xi| \leq \lambda + 1\}$. Correspondingly, take $f_{\lambda, \theta} \in L^2(\mathbb{R}^d)$, $\|f_{\lambda, \theta}\|_{L^2} = 1$, to be defined as the function whose Fourier transform is the characteristic function of the set

$$\left\{ \xi \in \mathbb{R}^d : \lambda \leq |\xi| \leq \lambda + 1, \left| \frac{\xi}{|\xi|} - (1, 0, \dots, 0) \right| \leq \theta \right\}, \quad (1.11)$$

multiplied by an L^2 -normalization constant $c_{\lambda, \theta} \approx \lambda^{-\frac{d-1}{2}} \theta^{-\frac{d-1}{2}}$. When $\theta = \lambda^{-1/2}$, a linearization of the phase function of the Fourier integral $\int e^{ix \cdot \xi} \hat{f}_{\lambda, \lambda^{-1/2}}(\xi) d\xi$ shows that $|f_{\lambda, \lambda^{-1/2}}(x)| \gtrsim \lambda^{\frac{d-1}{4}}$ on the set

$$\left\{ x = (x_1, x') \in \mathbb{R} \times \mathbb{R}^{d-1} : |x'| \leq \varepsilon \lambda^{-1/2}, |x_1| \leq \varepsilon \right\}, \quad (1.12)$$

for some $\varepsilon > 0$. It is then easily verified that $\|f_{\lambda, \lambda^{-1/2}}\|_{L^p} \gtrsim \lambda^{\frac{d-1}{2}(\frac{1}{2} - \frac{1}{p})}$, resulting in a function analogous to the “Knapp example” from Fourier restriction theory. On the other hand, when $\theta = 1$, stationary phase suggests

$$|f_{\lambda, 1}(x)| \approx \lambda^{\frac{d-1}{2}} (1 + \lambda|x|)^{-\frac{d-1}{2}}, \quad |x_1| \gtrsim |x'|, \quad (1.13)$$

for sufficiently many x so that $\|f_{\lambda, 1}\|_{L^p} \gtrsim \lambda^{\frac{d-1}{2} - \frac{d}{p}}$ when $p > \frac{2d}{d-1}$. This yields families of functions which saturate the exponent in (1.2) when $2 < p \leq p_c$ and $p_c \leq p \leq \infty$ respectively. However, by carefully splitting into oscillatory regions where stationary phase can be applied and θ -dependent non-oscillatory regions similar to (1.12), it can be seen that at $p = p_c$, $\|f_{\lambda, \theta}\|_{L^{p_c}} \gtrsim \lambda^{1/p_c}$ for any $\theta \in [\lambda^{-1/2}, 1]$, hence its designation as the “critical” exponent. These computations were carried out rigorously in [34].

Analogous constructions can be carried out for suitable approximations to $1_{[\lambda, \lambda+1]}(P)$ on any (M, g) , only now the x_1 axis is replaced by a geodesic segment and Riemannian distance replaces Euclidean (see [26, Ch.5]). Moreover, localization analogous to that in (1.11) can be achieved by pseudodifferential operators (PDOs). These considerations demonstrate that in order to show (1.9) at the critical exponent, one must rule out a spectrum of scenarios for phase

space concentration: simply disproving either maximal mass concentration in $\lambda^{-1/2}$ tubular neighborhoods or decay akin to (1.13) as in previous works is not enough by itself. We shall see that the method in [29] is effective in proving nonconcentration for $\theta \geq \lambda^{-1/2+\varepsilon}$ for any fixed $\varepsilon \in (0, 1/2)$. A key idea in the present work to accomplish this for microlocalized modes corresponding to the remaining cases $\theta \in (\lambda^{-1/2}, \lambda^{-1/2+\varepsilon})$.

Unlike [29], the present work does not rely on the known bounds (1.5) when $2 < p < p_c$. The bounds in Theorem 1.1 can be interpolated with the $p = 2$ case to show L^p bounds for this range of p . As noted above, the exponent $\sigma(p, d)$ vanishes as $p \nearrow p_c$ so the interpolation yields an improved exponent for p interval to the left of p_c , but not all values of $2 < p < p_c$.

Outline of the work

In Sect. 2, we review the method introduced in [29]. We then show how to generate improvements on this approach for modes microlocalized to a conic sector about a fixed covector field, analogous to the angular localization in (1.11). This adapts the approach in [10].

The third section then details a proof by contradiction for our main result, Theorem 1.1. The arguments here are partially inspired by strategies in nonlinear PDE, particularly dispersive ones, which seek to characterize the phase space concentration of solutions which develop a singularity, then disprove the possibility of such concentration. While the present work does not develop an explicit “profile decomposition” for spectral clusters, akin to those which are common for nonlinear Schrödinger equations, the approach here is reminiscent of works in that vein such as [5, 11]. In Sect. 3.1, we review the local structure of spectral multipliers which roughly project onto frequency bands of width 1 and then define an almost orthogonal decomposition adapted to these operators which achieves the microlocalization considered in Sect. 2. This culminates in the statement of Theorem 3.2, which bounds the weak- L^{p_c} quasi-norms of such spectral multipliers by the mass of the elements in the decomposition. The contradiction is then finalized in Sect. 3.2. The proof of Theorem 1.1 thus relies in a crucial way on Theorem 3.2 and the improvements from Sect. 2, in particular Corollary 2.2. Together these are the central developments in the present work.

The fourth section sets the stage for bilinear estimates on approximate projections onto bands of width 1, which will yield the proof of Theorem 3.2. We then need to show how the elements of our decomposition behave under these approximate projections, which is done in Sect. 4.2. The bilinear estimates can then be concluded. The final subsection Sect. 4.3 then shows that products of the members of the decomposition obey an almost orthogonality principle in L^r spaces, a crucial lemma in the proof of the bilinear bounds.

The fifth and final section then considers results for geometric hypotheses on (M, g) weaker than nonpositive curvature.

Semiclassical analysis

This work uses a modest amount of semiclassical analysis, though instead of using the notation h commonly used in this practice, we use $\lambda = h^{-1}$ as the frequency parameter. The primary use is to quantize various compactly supported pseudodifferential symbols $q_\lambda(x, \xi)$ so that $Q_\lambda = \text{Op}(q_\lambda)$ is the operator with Schwartz kernel

$$Q_\lambda(x, y) = \frac{\lambda^d}{(2\pi)^d} \int e^{i\lambda(x-y)\cdot\xi} q_\lambda(x, \xi) d\xi \quad (\text{standard quantization}). \quad (1.14)$$

In the present work, one will be able to view these operations as the result of taking a classical symbol, compactly supported where $|\xi| \approx \lambda$ with uniform estimates in $S_{1,0}^0$, $S_{7/8,1/8}^0$, and applying the rescaling $\xi \mapsto \lambda\xi$. Such a rescaling yields symbols in the classes S_0 , $S_{1/8}$ respectively in the sense of [38, §4.4] The semiclassical Fourier transform is thus defined consistently by $\mathcal{F}_\lambda(f)(\xi) = \hat{f}(\lambda\xi)$ with inverse $\mathcal{F}_\lambda^{-1}(f)(x) = \lambda^d \check{f}(\lambda x)$ where \hat{f} , \check{f} are the classical Fourier transform and its inverse respectively. The use of semiclassical quantization makes for a convenient use of stationary phase.

Notation We take the common convention that $A \lesssim B$ means that $A \leq CB$ for some large constant C which depends only on (M, g) and in particular is uniform in λ and possibly other parameters except when they are given in the subscript of \lesssim . Similarly, $A \ll B$ means that $A \leq cB$ for some small uniform constant c . The notation $A \approx B$ means that $A \lesssim B$ and $B \lesssim A$. Certain variables may be reassigned when the analysis in a given section is independent of prior sections.

Throughout, ρ_λ abbreviates the operator $\rho(T(\lambda - P))$ in (1.8), where T is as in (1.7). We will also use “local” projectors σ_λ defined by $\rho(\tilde{c}_0(\lambda - P))$ for some fixed, but sufficiently small constant \tilde{c}_0 (much less than the injectivity radius of (M, g)). When these operators are restricted to some sequence of $\lambda_k \rightarrow \infty$, we abbreviate ρ_{λ_k} , σ_{λ_k} as ρ_k , σ_k respectively. Finally, we use $\Theta_g(x, y)$ to denote the Riemannian distance between two points x, y on M .

2 Review of [29] and improved weak bounds for microlocalized modes

2.1 Review of [29]

We review the arguments of the second author in [29] used to prove (1.8) with $\log \lambda$ replaced by $\log \log \lambda$. We begin by recalling weak- L^p and Lorentz spaces

on (M, g) with respect to Riemannian measure. The weak- L^p functions are the measurable functions for which the following quasi-norm is finite

$$\|f\|_{L^{p,\infty}(M)} = \sup_{\alpha > 0} \alpha |\{x \in M : |f(x)| > \alpha\}|^{\frac{1}{p}},$$

where the bars are used denote the Riemannian measure. The well-known Chebyshev inequality shows that functions in $L^p(M)$ are also in weak- L^p with $\|f\|_{L^{p,\infty}(M)} \leq \|f\|_{L^p(M)}$. More generally, the Lorentz spaces are a family of interpolation spaces which include both $L^p(M)$ and weak- L^p . They are defined by first considering the distribution function for measurable functions as

$$d_f(\alpha) = |\{x \in M : |f(x)| > \alpha\}|$$

then defining $L^{p,q}(M)$ as being the measurable functions for which the following quasi-norm is finite

$$\|f\|_{L^{p,q}(M)} := p^{\frac{1}{q}} \left(\int_0^\infty [d_f(s)^{1/p} s]^q \frac{ds}{s} \right)^{\frac{1}{q}}, \quad 0 < q < \infty.$$

Lorentz spaces are often equivalently defined using the decreasing rearrangement of f . A well known identity from measure theory shows $\|f\|_{L^{p,p}(M)} = \|f\|_{L^p(M)}$. As suggested by the notation above, when $q = \infty$ the Lorentz space $L^{p,\infty}(M)$ is just the weak- L^p functions.

As observed in [29, §4], an interpolation in Lorentz spaces yields (recalling $\rho_\lambda := \rho(T(\lambda - P))$)

$$\|\rho_\lambda\|_{L^2(M) \rightarrow L^{p_c}(M)} \lesssim \|\rho_\lambda\|_{L^2(M) \rightarrow L^{p_c,\infty}(M)}^{1 - \frac{2}{p_c}} \|\rho_\lambda\|_{L^2(M) \rightarrow L^{p_c,2}(M)}^{\frac{2}{p_c}}. \quad (2.1)$$

In [6, Corollary 1.3], Bak and Seeger showed $\|\rho_\lambda\|_{L^2 \rightarrow L^{p_c,2}} = \mathcal{O}(\lambda^{\frac{1}{p_c}})$. Consequently, it suffices to obtain weak L^{p_c} bounds on ρ_λ .

We consider a slightly more general setting for the weak bounds, considering instead weak bounds for $Q_\lambda \circ \rho_\lambda$ where Q_λ is either the identity or a semiclassical pseudodifferential operator as in (1.14) corresponding to a compactly supported symbol $q_\lambda \in S_{1/8}$ in that $|\partial^\alpha q| \lesssim_\alpha \lambda^{1/8}$. Note that [29] only considers the case where Q_λ is the identity.

Fix a unit vector $f \in L^2(M)$, then consider for $\alpha > 0$ and some coordinate system $\Omega \subset M$ the sets A_α defined by

$$A_\alpha := \{x \in \Omega : |((Q_\lambda \circ \rho_\lambda)f)(x)| > \alpha\}, \quad \|f\|_{L^2(M)} = 1. \quad (2.2)$$

Denoting the Riemannian measure of this set as $|A_\alpha|$, we seek a bound

$$\alpha|A_\alpha|^{\frac{1}{p_c}} \lesssim \lambda^{\frac{1}{p_c}} (\log \lambda)^{-\varepsilon_1}, \quad \varepsilon_1 := \frac{\varepsilon_0 p_c}{p_c - 2} = \frac{d+1}{2} \varepsilon_0. \quad (2.3)$$

We begin by restricting attention to the case

$$\lambda^{\frac{d-1}{4}} (\log \lambda)^{-\frac{1}{2}} \lesssim \alpha. \quad (2.4)$$

We now set

$$r := \lambda \alpha^{-\frac{4}{d-1}} (\log \lambda)^{-\frac{2}{d-1}} \quad \text{so that} \quad (\lambda r^{-1})^{\frac{d-1}{2}} = \alpha^2 \log \lambda. \quad (2.5)$$

Given (2.4), $r \ll 1$. At the cost of replacing A_α by a set of proportional measure, we may write $A_\alpha = \bigcup_j A_{\alpha,j}$ where $d(A_{\alpha,j}, A_{\alpha,k}) > C_0 r$ in Euclidean distance for some $C_0 > 0$ sufficiently large when $j \neq k$. To see this, cover the original set A_α by a lattice of nonoverlapping cubes of sidelength r . Then partition the cubes in this cover into $\mathcal{O}(1)$ subcollections such that the centers of the cubes in each subcollection are separated by a distance of at least $4^d C_0 r$. By the pigeonhole principle, the intersection of at least one subcollection in the partition with A_α must have measure comparable to A_α . We may thus replace A_α by its intersection with this subcollection of cubes.

Now let $\mathbf{1}_A$ denote the characteristic function of A , and $a_j = \mathbf{1}_{A_j} \psi_\lambda$ where ψ_λ is defined as

$$\psi_\lambda(x) = \begin{cases} \frac{((Q_\lambda \circ \rho_\lambda)f)(x)}{|((Q_\lambda \circ \rho_\lambda)f)(x)|}, & ((Q_\lambda \circ \rho_\lambda)f)(x) \neq 0, \\ 1, & ((Q_\lambda \circ \rho_\lambda)f)(x) = 0. \end{cases}$$

Since ρ_λ is self-adjoint and $\|f\|_{L^2(M)} = 1$,

$$\alpha|A_\alpha| \leq \left| \int ((Q_\lambda \circ \rho_\lambda)f) \overline{\psi_\lambda \mathbf{1}_{A_\alpha}} \right| \leq \left(\int \left| \sum_j (\rho_\lambda \circ Q_\lambda^*) a_j \right|^2 \right)^{\frac{1}{2}}.$$

This now yields (with $\rho_\lambda^2 = \rho_\lambda \circ \rho_\lambda$)

$$\alpha^2 |A_\alpha|^2 \leq \sum_j \int |(\rho_\lambda \circ Q_\lambda^*) a_j|^2 + \sum_{j \neq k} \int (Q_\lambda \circ \rho_\lambda^2 \circ Q_\lambda^*) a_j \overline{a_k} =: I + II. \quad (2.6)$$

We now consider the consequences of (2.6) when Q_λ is the identity and when this is a semiclassical PDO with $q_\lambda \in S_{1/8}$ separately.

2.1.1 Consequences of (2.6) when Q_λ is the identity

We further review the arguments in [29], assuming Q_λ is the identity. The arguments in [1, 19] used to prove (1.5) when $p = \infty$ also show that

$$\|\rho_\lambda\|_{L^2(M) \rightarrow L^\infty(M)} \lesssim \lambda^{\frac{d-1}{2}} (\log \lambda)^{-1/2}.$$

In fact, this is a consequence of (2.10) below and duality. Hence A_α as defined in (2.2) is vacuous unless $\alpha \lesssim \lambda^{\frac{d-1}{2}} (\log \lambda)^{-1/2}$, meaning we only need to consider cases where $r \gtrsim \lambda^{-1}$.

In [29, (30)], it is shown that ρ_λ satisfies local L^2 bounds over balls $B(x, r)$ when $\lambda^{-1} \lesssim r \leq \text{inj}(M)$

$$\|\rho_\lambda\|_{L^2(M) \rightarrow L^2(B(x, r))} = \|\rho_\lambda\|_{L^2(B(x, r)) \rightarrow L^2(M)} \lesssim r^{\frac{1}{2}}, \quad (2.7)$$

where the implicit constant is independent of x and the equality holds since ρ_λ is self-adjoint. Hence

$$I \lesssim r \sum_j \int |a_j|^2 \lesssim r |A_\alpha| = \lambda \alpha^{-\frac{4}{d-1}} (\log \lambda)^{-\frac{2}{d-1}} |A_\alpha|. \quad (2.8)$$

Moreover, with $K(w, z)$ denoting the integral kernel of ρ_λ^2

$$\begin{aligned} II &\lesssim \left(\sup_{j \neq k} \sup_{(w, z) \in A_{\alpha, j} \times A_{\alpha, k}} |K(w, z)| \right) \sum_{j \neq k} \|a_j\|_{L^1} \|a_k\|_{L^1} \\ &\lesssim \left(\sup_{j \neq k} \sup_{(w, z) \in A_{\alpha, j} \times A_{\alpha, k}} |K(w, z)| \right) |A_\alpha|^2. \end{aligned} \quad (2.9)$$

Lemma 3.3 in [29] then appeals to results of Bérard [1] to observe that there exists $C = C(M, g)$ sufficiently large such that

$$|K(w, z)| \leq \frac{C}{T} \left(\frac{\lambda}{\lambda^{-1} + \Theta_g(w, z)} \right)^{\frac{d-1}{2}} + C \lambda^{\frac{d-1}{2}} \exp(CT). \quad (2.10)$$

Recalling (1.7), we then have that

$$\begin{aligned} II &\lesssim \left(C_0^{-\frac{d-1}{2}} (\log \lambda)^{-1} (\lambda r^{-1})^{\frac{d-1}{2}} + \lambda^{C c_0 + \frac{d-1}{2}} \right) |A_\alpha|^2 \\ &\lesssim \left(C_0^{-\frac{d-1}{2}} \alpha^2 + \lambda^{C c_0 + \frac{d-1}{2}} \right) |A_\alpha|^2 \end{aligned} \quad (2.11)$$

Given (2.6), (2.8), and (2.11) we then have desirable bounds on $|A_\alpha|$ when $\alpha \geq \lambda^{\frac{d-1}{4}+\varepsilon}$ where ε can be made small by choosing c_0 much smaller and C_0 large. However, the smaller we wish to take ε , the smaller we must take c_0 , which does have to be uniform in the proof. In [29], this is remedied by taking $T = c_0 \log \log \lambda$ and appealing to the results in [9, 10], to handle smaller values of α . This in turn only yields a gain of a power of $(\log \log \lambda)^{-1}$ in the final estimates. In the present work, we assume c_0 is small enough so that the argument outlined here yields

$$\alpha |A_\alpha|^{\frac{1}{pc}} \lesssim \lambda^{\frac{1}{pc}} (\log \lambda)^{-\frac{1}{d+1}} \quad \text{for } \lambda^{\frac{d-1}{4}+\frac{1}{8}} \leq \alpha, \quad (2.12)$$

so that the crucial matter is to treat the cases $\alpha < \lambda^{\frac{d-1}{4}+\frac{1}{8}}$. The choice of $\varepsilon = \frac{1}{8}$ is not crucial, but a convenient choice for the sake of concreteness as it does influence other parameters throughout the work. We stress that in the remainder of this work, (2.12) is only applied to the case $Q_\lambda = I$.

2.1.2 Consequences of (2.6) when Q_λ is a semiclassical PDO

We now reconsider the bounds on I and II just established in Sect. 2.1.1 but with Q_λ now a semiclassical PDO with symbol in $S_{1/8}$. We would like for (2.7) to yield

$$I \lesssim r \sum_j \int |Q_\lambda^* a_j|^2 \lesssim r |A_\alpha| = \lambda \alpha^{-\frac{4}{d-1}} (\log \lambda)^{-\frac{2}{d-1}} |A_\alpha|. \quad (2.13)$$

However, the kernel of Q_λ^* is only rapidly decreasing outside a $\lambda^{-7/8}$ neighborhood of the diagonal and hence this estimate does not follow at scales finer than $r \leq \lambda^{-7/8}$. But given (2.12), we will only need to bound I when $\alpha < \lambda^{\frac{d-1}{4}+\frac{1}{8}}$, meaning that $r > \lambda^{-\frac{1}{2(d-1)}} (\log \lambda)^{-\frac{2}{d-1}}$, which always determines a much coarser scale of at least $r \gg \lambda^{-3/4}$. Hence in these cases, $|Q_\lambda^* a_j(x)| = \mathcal{O}(\lambda^{-N})$ for any N outside a cube of sidelength $\approx r$, so the local estimates in (2.7) do indeed yield (2.13).

Turning to the bounds on II in (2.9), we now consider the effect of replacing the kernel $K(w, z)$ of ρ_λ^2 there by the kernel of $Q_\lambda \circ \rho_\lambda^2 \circ Q_\lambda^*$ as indicated by (2.6). In the next subsection, we will show that for suitable choices of Q_λ , the corresponding kernel $K(w, z)$ satisfies

$$|K(w, z)| \lesssim \frac{1}{T} \left(\frac{\lambda}{\lambda^{-1} + \Theta_g(w, z)} \right)^{\frac{d-1}{2}} + c(\lambda) \lambda^{\frac{d-1}{2}},$$

for some $c(\lambda) \searrow 0$ at least as fast as an inverse power of $\log \lambda$ but no faster than $(\log \lambda)^{-1}$ (so that (2.4) is ultimately respected in this argument). Hence (2.11) can be improved to read

$$II \lesssim \left(C_0^{-\frac{d-1}{2}} \alpha^2 + \lambda^{\frac{d-1}{2}} c(\lambda) \right) |A_\alpha|^2. \quad (2.14)$$

Taking C_0 sufficiently large, we obtain an improvement on (2.12):

$$\alpha |A_\alpha|^{\frac{1}{pc}} \lesssim \lambda^{\frac{1}{pc}} (\log \lambda)^{-\frac{1}{d+1}} \quad \text{for } \lambda^{\frac{d-1}{4}} c(\lambda)^{\frac{1}{2}} \lesssim \alpha \leq \lambda^{\frac{d-1}{4} + \frac{1}{8}}. \quad (2.15)$$

2.2 Improved weak estimates for microlocalized modes

Consider any local coordinate chart Ω on M . Suppose $q_\lambda(x, \xi)$ is a semiclassical symbol such that for some unit covector field $\omega(x)$, $|\omega(x)|_{g(x)} = 1$ (with $g(x)$ the ‘‘cometric’’, the inner product on the T^*M induced by the metric),

$$\begin{aligned} \text{supp}(q_\lambda) \subset & \left\{ (x, \xi) \in T^* \Omega : |\xi/|\xi|_{g(x)} - \omega(x)|_{g(x)} \lesssim \lambda^{-1/8}, |\xi| \approx 1 \right\}, \\ & \left| \langle \omega(x), d_\xi \rangle^j \partial_{x, \xi}^\beta q_\lambda(x, \xi) \right| \lesssim_{\beta, j} \lambda^{|\beta|/8}. \end{aligned} \quad (2.16)$$

The symbol q_λ thus lies in the subcritical class² $S_{1/8}$ (as in [38, §4.4]). If one sets $Q_\lambda := \text{Op}(q_\lambda)$ as in (1.14), we show the following improvement on (2.10) of the kernel of the composition $Q_\lambda \circ \rho_\lambda^2 \circ Q_\lambda^*$:

Theorem 2.1 *Let $K(w, z)$ denote the kernel of $Q_\lambda \circ \rho_\lambda^2 \circ Q_\lambda^*$. We then have*

$$\begin{aligned} |K(w, z)| & \lesssim \frac{1}{T} \left(\frac{\lambda}{\Theta_g(w, z)} \right)^{\frac{d-1}{2}} + c(\lambda) \lambda^{\frac{d-1}{2}}, \\ c(\lambda) & = \begin{cases} (\log \lambda)^{-1/2}, & \text{if } d = 2, \\ (\log \lambda)^{-1} \log \log \lambda, & \text{if } d = 3, \\ (\log \lambda)^{-1}, & \text{if } d \geq 4. \end{cases} \end{aligned} \quad (2.17)$$

where the implicit constants can be taken independent of λ and depend only on finitely many of the derivative bounds in (2.16).

Corollary 2.2 *Let $Q_\lambda, c(\lambda)$ be as in Theorem 2.1, A_α as in (2.2). Then*

$$\alpha |A_\alpha|^{\frac{1}{pc}} \lesssim \lambda^{\frac{1}{pc}} c(\lambda)^{\frac{1}{d+1}}, \quad 0 < \alpha \leq \lambda^{\frac{d-1}{4} + \frac{1}{8}}. \quad (2.18)$$

² Again, the choice of 1/8 is not crucial here, only a convenient one.

Proof of Corollary 2.2 Given (2.15), it suffices to assume $\alpha \lesssim \lambda^{\frac{d-1}{4}} c(\lambda)^{\frac{1}{2}}$. But since $\|Q_\lambda \circ \rho_\lambda\|_{L^2(M) \rightarrow L^2(M)} \lesssim 1$ uniformly, we have $\alpha|A_\alpha|^{\frac{1}{2}} \lesssim 1$, hence

$$\alpha|A_\alpha|^{\frac{1}{pc}} = \alpha^{1-\frac{2}{pc}} \left(\alpha|A_\alpha|^{\frac{1}{2}}\right)^{\frac{2}{pc}} \lesssim \lambda^{\frac{1}{pc}} c(\lambda)^{\frac{1}{d+1}},$$

by the upper bound on α . \square

2.2.1 Consequences of the Hadamard parametrix and the proof of Theorem 2.1

Since $\hat{\rho}^2 = \hat{\rho} * \hat{\rho}$ is supported in $[-1, 1]$, the key to (2.17) is to bound the following integral by the *second term* on the right hand side of (2.17):

$$\frac{1}{2\pi T} \int_{-T}^T (1-\beta)(t) \hat{\rho}^2(t/T) e^{i\lambda t} (Q_\lambda \circ \cos(tP) \circ Q_\lambda^*)(w, z) dt. \quad (2.19)$$

where β is of sufficiently small compact support and identically one in a neighborhood of 0. Indeed, without the factor of $1 - \beta$ in the integrand, this is the kernel of $Q_\lambda \circ \rho_\lambda^2 \circ Q_\lambda^*$, up to negligible errors, by Euler's formula. It is a classical result of Hörmander [16] that if one replaces $1 - \beta$ by β here, the resulting kernel is bounded by the first term on the right in (2.17).

Since (M, g) does not have conjugate points, the kernel of $\cos(tP)$ can be analyzed by lifting to the universal cover (\tilde{M}, \tilde{g}) where \tilde{g} is defined by pulling the metric tensor g back via the covering map. Fix a fundamental domain $D \subset \tilde{M}$ and let \tilde{w}, \tilde{z} denote the unique points in D which project onto w, z in M via the covering map. Recall that the classical Cartan-Hadamard theorem ensures that \tilde{M} is diffeomorphic to \mathbb{R}^d via the exponential map at any point. Here we take global geodesic coordinates on \tilde{M} via the exponential map at \tilde{w} . We also assume that the geodesic in \tilde{M} from \tilde{w} with initial covector $\omega(\tilde{w})$ lies along the first coordinate axis and let $\tilde{\gamma}(t) = (t, 0, \dots, 0)$ denote this unit speed geodesic.

If $\tilde{P} = \sqrt{-\Delta_{\tilde{g}}}$, with $\Delta_{\tilde{g}}$ the Laplacian on (\tilde{M}, \tilde{g}) , we have

$$\cos(tP)(w, z) = \sum_{\alpha \in \Gamma} \cos(t\tilde{P})(\tilde{w}, \alpha(\tilde{z}))$$

where Γ denotes the group of deck transformations which preserve the covering map.³ Note that by finite speed of propagation, we may restrict attention to

³ The proof of Theorem 2.1 is more or less independent of the other sections, so we temporarily reassign α as indexing Γ in the interest of consistency with prior works.

the $\alpha \in B(\tilde{w}, T)$. For $\tilde{x} \in D$ and $\tilde{y} \in \mathbb{R}^d$, we first concern ourselves with

$$V(\tilde{x}, \tilde{y}) := \frac{1}{2\pi T} \int_{-T}^T (1 - \beta)(t) \hat{\rho}^2(t/T) e^{i\lambda t} \cos(t \tilde{P})(\tilde{x}, \tilde{y}) dt.$$

If we extend the kernel of Q_λ^* to be periodic with respect to $\alpha \in \Gamma$, we have (with $d\tilde{x}, d\tilde{y}$ implicitly the Riemannian measure with respect to \tilde{g})

$$\begin{aligned} (2.19) &= \sum_{\alpha \in \Gamma} U_\alpha(\tilde{w}, \tilde{z}), \\ U_\alpha(\tilde{w}, \tilde{z}) &:= \int_{\alpha(D)} \int_D Q_\lambda(\tilde{w}, \tilde{x}) V(\tilde{x}, \tilde{y}) Q_\lambda^*(\tilde{y}, \alpha^{-1}(\tilde{z})) d\tilde{x} d\tilde{y}. \end{aligned} \quad (2.20)$$

Using the Hadamard parametrix for the wave equation on $(\mathbb{R}^d, \tilde{g})$ and stationary phase (see for example, [8, Lemma 5.1], [10, §3], [12, Lemma 3.1]), it is known that

$$V(\tilde{x}, \tilde{y}) := \frac{\lambda^{\frac{d-1}{2}}}{T \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})^{\frac{d-1}{2}}} \sum_{\pm} e^{\pm i\lambda \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})} a_{\lambda, \pm}(\tilde{x}, \tilde{y}) + R_\lambda(\tilde{x}, \tilde{y}). \quad (2.21)$$

Here $a_{\lambda, \pm}, R_\lambda$ vanish for $\Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) \geq T$ by finite speed of propagation and $a_{\lambda, \pm}$ also vanishes if $\Theta_{\tilde{g}}(\tilde{x}, \tilde{y})$ is sufficiently small since β vanishes in a neighborhood of the origin. The remainder can be taken so that $|R_\lambda(\tilde{x}, \tilde{y})| \lesssim \lambda^{-2}$. Moreover, $a_{\lambda, \pm}$ can be written as

$$a_{\lambda, \pm}(\tilde{x}, \tilde{y}) = \vartheta(\tilde{x}, \tilde{y}) a_{\lambda, \pm, 1}(\Theta_{\tilde{g}}(\tilde{x}, \tilde{y})) + a_{\lambda, \pm, 2}(\tilde{x}, \tilde{y}), \quad (2.22)$$

where $|\partial_r^j a_{\lambda, \pm, 1}(r)| \lesssim_j r^{-j}$ and there exists C_d so that for $0 < |\beta| < 16d$,

$$|\partial_{\tilde{x}, \tilde{y}}^\beta \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})|, \lambda^2 |\partial_{\tilde{x}, \tilde{y}}^\beta a_{\lambda, \pm, 2}(\tilde{x}, \tilde{y})|, |\partial_{\tilde{x}, \tilde{y}}^\beta \vartheta(\tilde{x}, \tilde{y})| \lesssim \exp(C_d \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})). \quad (2.23)$$

The function $\vartheta(\tilde{x}, \tilde{y})$ is the leading coefficient in the Hadamard parametrix. It is characterized by the property that $dV_g = \vartheta^{-2}(\tilde{x}, \tilde{y}) d\mathcal{L}$ in normal coordinates at \tilde{x} , with \mathcal{L} denoting Lebesgue measure on \mathbb{R}^d . Since (\tilde{M}, \tilde{g}) has nonpositive sectional curvatures, it is observed in [32] that ϑ is uniformly bounded as a consequence of the Günther comparison theorem. Moreover, if the curvatures are strictly negative and bounded above by $-\kappa^2$, the same theorem implies $\vartheta(\tilde{x}, \tilde{y}) \lesssim \exp(-\frac{\kappa(d-1)}{2} \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}))$.

Given the properties of the support of $a_{\pm, \lambda}$ and R_λ , there are at most $\mathcal{O}(e^{CT})$ nonzero terms in the sum (2.20) as a consequence of lattice point counting arguments. As observed above, $|R_\lambda(\tilde{x}, \tilde{y})| \lesssim \lambda^{-2}$ and hence by Sobolev embedding and L^2 bounds on Q_λ , we may restrict attention to the sum over \pm in (2.21).

We next observe that in our global coordinate system, we may assume that up to acceptable $\mathcal{O}(\lambda^{-2})$ error, the kernel of Q_λ is of the form

$$\frac{\lambda^d}{(2\pi)^d} \int e^{i\lambda(\tilde{w}-\tilde{x})\cdot\eta} q_\lambda(\tilde{w}, \tilde{x}, \eta) d\eta,$$

$$\text{supp}(q_\lambda) \subset \left\{ (\tilde{w}, \eta) \in T^*D, \tilde{x} \in D : |\eta/|\eta| - (1, 0, \dots, 0)| \lesssim \lambda^{-\frac{1}{8}}, |\eta| \approx 1 \right\}.$$

Here we have used a compound symbol, deviating slightly from (1.14) to ensure the kernel is supported in $D \times D$. We may assume the same for the support of the symbol q_λ^* of the adjoint. Restricting attention to the main term in (2.21), $U_\alpha(\tilde{w}, \tilde{z})$ is a sum over \pm

$$\begin{aligned} & \frac{\lambda^{\frac{5d-1}{2}}}{(2\pi)^{2d} T} \int e^{i\lambda\varphi_\pm(\tilde{w}, \tilde{x}, \tilde{y}, \tilde{z}, \eta, \zeta)} q_\lambda(\tilde{w}, \tilde{x}, \eta) a_{\pm, \lambda}(\tilde{x}, \tilde{y}) \\ & \quad \times q_\lambda^*(\tilde{y}, \alpha^{-1}(\tilde{z}), \eta) d\tilde{x} d\tilde{y} d\eta d\zeta, \end{aligned} \quad (2.24)$$

$$\varphi_\pm(\tilde{w}, \tilde{x}, \tilde{y}, \tilde{z}, \eta, \zeta) := (\tilde{w} - \tilde{x}) \cdot \eta \pm \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) + (\tilde{y} - \alpha^{-1}(\tilde{z})) \cdot \zeta,$$

where as before the domain of integration is $(\tilde{x}, \tilde{y}) \in D \times \alpha(D)$.

Applying stationary phase to (2.24) shows that for any $\alpha \in B(w, T)$,

$$|U_\alpha(\tilde{w}, \tilde{z})| \lesssim \frac{\lambda^{\frac{d-1}{2}}}{T} (\vartheta(\tilde{w}, \alpha(\tilde{z})) + \lambda^{-2}) \left(1 + \Theta_g(\tilde{w}, \alpha(\tilde{z}))\right)^{-\frac{d-1}{2}}. \quad (2.25)$$

The main idea in the proof of (2.17) is that one can improve upon this bound when $\alpha(D)$ is outside a tubular neighborhood of \tilde{y} . The proof is similar to that in [10] where the authors made use of the following consequence of the Toponogov triangle comparison theorem (see [10, Proposition 2.1] for further details).

Lemma 2.3 *Suppose $(\mathbb{R}^d, \tilde{g})$ is the cover of (M, g) given by the exponential map at w and that its sectional curvatures are bounded below by -1 . Given $T \gg 1$ and $\theta \ll 1$, let $C(\theta; T)$ denote the set of points in the metric ball of radius T about w such that the geodesic through the point and w forms an angle less than θ with \tilde{y} . Fix R sufficiently large. Then if*

$$\mathcal{T}_R := \{\tilde{x} \in \mathbb{R}^d : \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) \leq R\},$$

we have $C(\theta_T; T) \subset \mathcal{T}_R$ if $\sin(\frac{\theta_T}{2}) = \frac{\sinh(R/2)}{\sinh T}$.

Note that we may assume the sectional curvatures of (M, g) and (\tilde{M}, \tilde{g}) are bounded below by -1 by rescaling the metric in the outset of the proof.

Fix $R = 100 \cdot \text{diam}(D)$. Given the lemma, we take c_0 in (1.7) so that

$$C(\lambda^{-1/16}; c_0 \log \lambda) = C(\lambda^{-1/16}; T) \subset \mathcal{T}_R, \text{ and hence}$$

$$|\pm d_{\tilde{w}} \Theta_{\tilde{g}}(\tilde{w}, \tilde{y}) - (1, 0, \dots, 0)| \gtrsim \lambda^{-\frac{1}{16}}, \quad \tilde{y} \notin \mathcal{T}_R. \quad (2.26)$$

2.2.2 Proof of Theorem 2.1

As in [10], set

$$\Gamma_{\mathcal{T}_R} := \{\alpha \in \Gamma : \alpha(D) \cap \mathcal{T}_R \neq \emptyset\}.$$

The arguments on p. 202 in that work then show that the cardinality of $\{\alpha \in \Gamma_{\mathcal{T}_R} : \Theta_{\tilde{g}}(\tilde{w}, \alpha(\tilde{z})) \in [2^k, 2^{k+1}]\}$ is $\mathcal{O}(2^k)$. Therefore given (2.25),

$$\sum_{\alpha \in \Gamma_{\mathcal{T}_R}} |U_\alpha(\tilde{w}, \tilde{z})| \lesssim \frac{\lambda^{\frac{d-1}{2}}}{T} \sum_{0 \leq k \lesssim \log_2 \lambda} 2^k 2^{-k \frac{d-1}{2}} \lesssim c(\lambda) \lambda^{\frac{d-1}{2}}. \quad (2.27)$$

Indeed, so geometric summation shows the inequality.

We are now left to show that

$$|U_\alpha(\tilde{w}, \tilde{z})| \lesssim 1, \quad \text{for } \alpha \notin \Gamma_{\mathcal{T}_R}. \quad (2.28)$$

Indeed, if this holds, then given (1.7) we have for some uniform constant C ,

$$\sum_{\alpha \notin \Gamma_{\mathcal{T}_R}} |U_\alpha(\tilde{w}, \tilde{z})| \lesssim e^{CT} \lesssim \lambda^{C c_0} \lesssim c(\lambda) \lambda^{\frac{d-1}{2}},$$

since we take c_0 sufficiently small.

Next observe that with φ_\pm as in (2.24)

$$d_{\tilde{x}} \varphi_\pm = \pm d_{\tilde{x}} \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) - \eta,$$

$$d_\eta \varphi_\pm = \tilde{w} - \tilde{x}, \quad d_\zeta \varphi_\pm = \tilde{y} - \alpha^{-1}(\tilde{z}).$$

Now recall (2.23) and the constant C_d there. If we take c_0 small so that $\lambda^{C_d c_0} \ll \lambda^{1/16}$, integration by parts in (2.24) yields

$$|U_\alpha(\tilde{w}, \tilde{z})| \lesssim \sup_{\tilde{x}, \tilde{y}, \eta, \pm} \lambda^{\frac{5d-1}{2}} \left(1 + \lambda^{\frac{7}{8}} |\pm d_{\tilde{x}} \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) - \eta| \right.$$

$$\left. + \lambda^{\frac{7}{8}} |\tilde{y} - \alpha^{-1}(\tilde{z})| + \lambda^{\frac{7}{8}} |\tilde{w} - \tilde{x}| \right)^{-8d},$$

where the supremum is over all points inside the support of the amplitude. However, there exists C such that

$$|d_{\tilde{x}}\Theta_{\tilde{g}}(\tilde{x}, \tilde{y}) - d_{\tilde{w}}\Theta_{\tilde{g}}(\tilde{w}, \alpha^{-1}(\tilde{z}))| \lesssim e^{CT} (|\tilde{w} - \tilde{x}| + |\tilde{y} - \alpha^{-1}(\tilde{z})|),$$

so taking $c_0 < \frac{1}{16C}$ in (1.7), the constant on the right is $\lambda^{C_{c_0}} \ll \lambda^{1/16}$, hence

$$|U_\alpha(\tilde{w}, \tilde{z})| \lesssim \sup_{\eta, \pm} \lambda^{\frac{5d-1}{2}} \left(1 + \lambda^{\frac{3}{4}} |\pm d_{\tilde{w}}\Theta_{\tilde{g}}(\tilde{w}, \alpha^{-1}(\tilde{z})) - \eta| \right)^{-8d},$$

But since $|\eta - (1, 0, \dots, 0)| \lesssim \lambda^{-1/8}$, and $\alpha \notin \Gamma_{T_R}$, as a consequence of (2.26) the second factor is $\mathcal{O}(\lambda^{-3d})$ which is stronger than (2.28).

Remark 2.4 When the curvatures of (M, g) are strictly negative, one can take $c(\lambda) = (\log \lambda)^{-1}$ in Theorem 2.1 and its corollary in any dimension, leading to an improvement in the exponent ε_0 in Remark 1.2 when $d = 2, 3$ via the argument in Sect. 3. As observed above, ϑ decays exponentially in $\Theta_{\tilde{g}}$ in this case, and hence the sum in (2.27) is $\mathcal{O}(\lambda^{\frac{d-1}{2}} / \log \lambda)$ for any $d \geq 2$.

3 The proof by contradiction

To obtain a contradiction to Theorem 1.1, suppose there exists a sequence of triples $\{(f_k, \lambda_k, B_k)\}_{k=1}^\infty$ such that $\|f_k\|_{L^2(M)} = 1$, $B_k, \lambda_k \rightarrow \infty$ such that

$$0 < \frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \|\rho_k f_k\|_{L^{p_c, \infty}(M)}, \quad \varepsilon_1 = \frac{\varepsilon_0 p_c}{p_c - 2} = \frac{d+1}{2} \varepsilon_0, \quad (3.1)$$

where ε_0 is in our main L^{p_c} estimate in Theorem 1.1 (cf. Remark 1.2) and as before, $\rho_k = \rho_{\lambda_k}$. Indeed, if we had

$$\limsup_{\lambda \rightarrow \infty} \lambda^{-1/p_c} (\log \lambda)^{\varepsilon_0} \|\rho_\lambda\|_{L^2(M) \rightarrow L^{p_c}(M)} = \infty,$$

then a similar inequality holds with different values of $B_k \rightarrow \infty$, a strong L^p bound replacing this weak one, and the larger log-exponent ε_1 replaced by ε_0 . But then the Lorentz interpolation argument (2.1) yields (3.1).

Taking ε_0 small enough so that $\varepsilon_1 \leq \frac{1}{d+1}$, given the consequence (2.12) of the results in [29], we may assume for each k , there is $\alpha_k > 0$ such that

$$\frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k |\{x \in M : |\rho_k f(x)| > \alpha_k\}|^{\frac{1}{p_c}}, \quad \alpha_k \leq \lambda_k^{\frac{d-1}{4} + \frac{1}{8}}. \quad (3.2)$$

In order to take advantage of the improved microlocalized bounds in Theorem 2.1 and Corollary 2.2, we will appeal to methods emanating from the Fourier restriction problem and their relatives. In particular, we want to control the $L^{p_c, \infty}$ quasi-norm of the $\rho_k f$ by the L^{p_c} and L^2 norm of expressions such as $Q_\lambda \rho_k f$ with Q_λ being the pseudodifferential cutoff function as in Theorem 2.1 (though the notation will change slightly below). While the operator ρ_λ is still too poorly understood to apply such classical methods, we can instead use local operators $\sigma_k = \rho(\tilde{c}_0(\lambda_k - P))$ (as in the notation section) in order to achieve this. This is in the same spirit of the authors' previous work, and that of others, where the local operators are treated in a way that make them amenable to global analysis.

The main idea is that $(I - \sigma_k) \circ \rho_k$ is an acceptable error term. Indeed, since $(1 - \rho)(0) = 0$, we have

$$|(1 - \rho)(\tilde{c}_0(\lambda - \tau))\rho(T(\lambda - \tau))| \lesssim T^{-1}(1 + T|\lambda - \tau|)^{-N},$$

and hence the classical $L^2 \rightarrow L^{p_c}$ bounds (1.2) for spectral projectors $\mathbf{1}_{[l, l+1]}(P)$ imply

$$\|(I - \sigma_k) \circ \rho_k f_k\|_{L^{p_c}(M)} \lesssim \lambda^{1/p_c} (\log \lambda)^{-1}. \quad (3.3)$$

Since $\varepsilon_1 < 1$, we may assume that (3.2) holds with $\sigma_k \rho_k f_k$ replacing $\rho_k f_k$ (the former abbreviating $(\sigma_k \circ \rho_k) f_k$).

Now take a finite partition of unity subordinate to an open cover of a suitable family of coordinate domains. By the pigeonhole principle, we may assume that at the cost of shrinking the B_k and α_k by a uniform factor and passing to a subsequence of the triples indexed by k there is a bump function ψ supported in a coordinate chart $\Omega \subset \mathbb{R}^d$ centered at the origin for which

$$\frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k |\{x \in \Omega : |\psi(x)(\sigma_k \rho_k f_k)(x)| > \alpha_k\}|^{\frac{1}{p_c}}.$$

After another harmless shrinking of B_k, α_k , we may also assume that the measure here is the usual Lebesgue measure in coordinates instead of Riemannian measure. By a second application of the pigeonhole principle, we may assume that there exists a Fourier multiplier $m \in S_{1,0}^0$ truncating to a conic sector of small aperture about a fixed vector such that

$$\frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k |\{x \in \Omega : |(m(D)\psi \sigma_k \rho_k f_k)(x)| > \alpha_k\}|^{\frac{1}{p_c}}. \quad (3.4)$$

After a possible rotation of coordinates, we may further assume that the fixed vector is $(1, 0, \dots, 0)$, that is,

$$\text{supp}(m) \subset \{\xi : |\xi/|\xi| - (1, 0, \dots, 0)| \ll 1\}.$$

3.1 Analysis of σ_λ

We may assume that in the coordinate chart Ω , $g^{ij}(0) = \delta^{ij}$ and that for some $\epsilon > 0$ sufficiently small

$$\Omega = [-\epsilon, \epsilon]^d. \quad (3.5)$$

We now recall the method for computing the kernel of $\sigma_\lambda = \rho(\tilde{c}_0(\lambda - P))$ from [26, Ch. 5]. There it is observed that σ_λ can be realized as an operator valued integral involving the wave kernel e^{-itP}

$$\sigma_\lambda = \frac{1}{2\pi\tilde{c}_0} \int_{-\tilde{c}_0}^{\tilde{c}_0} e^{it\lambda} e^{-itP} \widehat{\rho}(t/\tilde{c}_0) dt.$$

Using a Lax parametrix, it is well known that for $|t| \leq \tilde{c}_0$ there exists a phase function $\varphi(t, x, \xi)$ and an amplitude $v(t, x, \xi)$ such that the Schwartz kernel of $m(D)\psi e^{-itP}$ is given by an oscillatory integral

$$\left(m(D)\psi e^{-itP} \right)(x, y) = 2\pi\tilde{c}_0 \int e^{i(\varphi(t, x, \xi) - y \cdot \xi)} v(t, x, \xi) \tilde{\psi}(y) d\xi + \text{error},$$

where the error is smoothing to a sufficient order and hence can be neglected in what follows. Here $\tilde{\psi}$ is a bump function of slightly larger support and we may assume $v(t, \cdot, \xi)$, $\tilde{\psi}$ are supported in the same coordinate chart Ω as above. Moreover, we may take

$$\text{supp}(v(t, x, \cdot)) \subset \{\xi : |\xi/|\xi| - (1, 0, \dots, 0)| \ll 1\},$$

for some conic sector of slightly larger aperture than the one containing $\text{supp}(m)$ (cf. (3.4)). Up to negligible error, the kernel of $m(D)\psi\sigma_\lambda$ is

$$\int_{-\tilde{c}_0}^{\tilde{c}_0} \int e^{i(\lambda t + \varphi(t, x, \xi) - y \cdot \xi)} \widehat{\rho}(t/\tilde{c}_0) v(t, x, \xi) d\xi dt \cdot \tilde{\psi}(y).$$

An integration by parts in t shows that the contribution of the region where $|\xi| \ll \lambda$ or $|\xi| \gg \lambda$ to this integral is $\mathcal{O}(\lambda^{-N})$ for any N and hence negligible. Hence we may assume that $v(t, x, \cdot)$ is further supported where $|\xi| \approx \lambda$. Rescaling $\xi \mapsto \lambda\xi$, we are reduced to considering a semiclassical Fourier

integral operator $\tilde{\sigma}_\lambda$ given by integration against the kernel

$$\tilde{\sigma}_\lambda(x, y) := \lambda^d \int_{-\tilde{c}_0}^{\tilde{c}_0} \int e^{i\lambda(t+\varphi(t, x, \xi) - y \cdot \xi)} \widehat{\rho}(t/\tilde{c}_0) v(t, x, \xi) d\xi dt \cdot \tilde{\psi}(y) \quad (3.6)$$

where now $v(t, x, \cdot)$ is supported where $|\xi| \approx 1$ and in the same conic region as before. Therefore in what follows, we may assume that any function on which $\tilde{\sigma}_\lambda$ operates has its semiclassical Fourier transform supported in this region. Note that the operator $\tilde{\sigma}_\lambda$ is $m(D)\psi\sigma_\lambda$ up to negligible error.

We pause to remark that [26, Lemma 5.1.3] uses stationary phase on (3.6) to show that $\tilde{\sigma}_\lambda$ is an oscillatory integral operator with Carleson-Sjölin phase (see also the $2^j \approx 1$ case of Lemma 4.3 below). As observed there, the L^p theory for such operators due to Hörmander and Stein then yield the following linear estimates on $\tilde{\sigma}_\lambda$, which in turn imply (1.2):

$$\|\tilde{\sigma}_\lambda\|_{L^2 \rightarrow L^{p_c}} \lesssim \lambda^{\frac{1}{p_c}}. \quad (3.7)$$

We now want to decompose the identity into a family of pseudodifferential operators which have the effect of localizing a function in phase space in a fashion similar to Fourier multipliers defined by the characteristic functions in (1.11). However, this requires care as the operators must in some sense be invariant under the geodesic flow. We achieve this by fixing a hyperplane, namely the $x_1 = 0$ plane, then localizing the momenta so that it is within a $\lambda^{-1/8}$ neighborhood of a fixed vector as it passes through this hyperplane. In the construction, it is convenient to use the trivialization $T^*\Omega \cong \Omega \times \mathbb{R}^d$ to define a family of constant covector fields along the hyperplane which serve as the centers of these neighborhoods (constant in the sense that their expression in the coordinate frame is independent of position).

3.1.1 Analysis of the geodesic flow

In preparation for the decomposition, we study χ_t , which we denote as the flow on $T^*\Omega$ generated by the Hamiltonian vector field of $p(x, \xi) = |\xi|_{g(x)}$. Hence $\chi_t(x, \xi)$ is the time t value of the integral curve of the Hamiltonian vector field of p with initial data (x, ξ) . Recall that the phase function φ in the construction above satisfies

$$\chi_t(d_\xi \varphi(t, x, \xi), \xi) = (x, d_x \varphi(t, x, \xi)). \quad (3.8)$$

For initial data in the cosphere bundle $S^*\Omega$ defined by

$$S^*\Omega := \{(x, \xi) \in T^*\Omega : |\xi|_{g(x)} = 1\},$$

the integral curves of p coincide with geodesics of (M, g) as curves in the cotangent bundle. We write $x = (x_1, x')$ so that in particular $(0, x')$ gives coordinates on the $x_1 = 0$ hyperplane. Consider the restriction of this flow to a neighborhood of origin in the hyperplane $x_1 = 0$ and ξ in a conic neighborhood of $(1, 0, \dots, 0)$ in $S_x^*\Omega$, the cosphere space at x . Assuming that ϵ in (3.5), and \tilde{c}_0 is sufficiently small, we have for $|t| \leq \tilde{c}_0$, the mapping

$$(t, x', \eta) \mapsto \chi_t(0, x', \eta)$$

generates a diffeomorphism from the neighborhood to a conic neighborhood of $(1, 0, \dots, 0)$ in the cosphere bundle $S^*\Omega$. Indeed, recalling our assumption that $g^{ij}(0) = \delta^{ij}$, the derivative of this mapping at $(0, 0, (1, 0, \dots, 0))$ is the identity. Denote the inverse as

$$(\iota(x, \omega), \Phi(x, \omega), \Psi(x, \omega)) \in (-\tilde{c}_0, \tilde{c}_0) \times \{y_1 = 0\} \times S_{\Phi(x, \omega)}^*\Omega. \quad (3.9)$$

Equivalently, these functions can be described in terms of the minimizing unit speed geodesic passing through (x, ω) : this geodesic passes through the $y_1 = 0$ plane at the point $y' = \Phi(x, \omega)$, the covector at this intersection point is given by $\Psi(x, \omega)$, and $\iota(x, \omega) = \Theta_g(x, \Phi(x, \omega))$.

We note that we may further assume that for any $x \in \Omega$, $\omega \mapsto \Psi(x, \omega)$ is an invertible mapping, and if $\eta \mapsto \omega(x, \eta)$ denotes the inverse, then $\omega(x, \eta)$ is the unit covector along the geodesic through x whose covector at the intersection point with $y_1 = 0$ is η .

3.1.2 The almost orthogonal decomposition

Now let v index a collection of vectors in a neighborhood of $(1, 0, \dots, 0)$ on \mathbb{S}^{d-1} separated by a distance of at least $\frac{1}{2}\lambda^{-1/8}$. Define a corresponding partition of unity $\beta_v(\xi)$ such that $\text{supp}(\beta_v)$ is contained in a spherical cap of diameter $2\lambda^{-1/8}$ about v and $\sum_v \beta_v(\xi) = 1$ for $\xi \in \mathbb{S}^{d-1}$. Then extend $\beta_v(\xi)$ to all of $\mathbb{R}^d \setminus \{0\}$, so that it is homogeneous of degree zero. Now define

$$q_v(x, \xi) = \tilde{\psi}(x)\beta_v(\Psi(x, \xi/|\xi|_{g(x)}))\tilde{\beta}(|\xi|_{g(x)}),$$

where $\tilde{\beta}$ is a bump function such that $\tilde{\psi}(x)v(t, x, \xi) = \tilde{\psi}(x)\tilde{\beta}(|\xi|_{g(x)})v(t, x, \xi)$ is supported where $|\xi| \approx 1$ and in a slightly larger conic region than $v(t, x, \cdot)$. Moreover, we take $\tilde{\psi}$ to be a bump function supported in Ω and identically one on $\tilde{\psi}$. This bump function means that $q_v(x, \xi)$ is not invariant under χ_t , but we can assume that \tilde{c}_0 and the support is chosen suitably so that $q_v(\chi_t(x, \xi)) = q_v(x, \xi)$ when $\tilde{\psi}(x)v(t, x, \xi) \neq 0$ and $|t| \leq \tilde{c}_0$.

The function q_v thus defines a semiclassical symbol in the class $S_{1/8}$. It is of the form considered in Theorem 2.1 where the unit covector field $\omega(x) =$

$\omega(x, v)$ is that of the minimizing geodesic passing through x such that its intersection with $y_1 = 0$ has the covector $v/|v|_{g(x)}$. We define $Q_v = \text{Op}(q_v)$ as in (1.14) and hence up to error which is $\mathcal{O}(\lambda^{-N})$ in L^2 for some N sufficiently large

$$\tilde{\sigma}_\lambda h = \sum_v \tilde{\sigma}_\lambda Q_v h. \quad (3.10)$$

Moreover, the selection of the indices v ensures that there exists a constant C_d such that for any fixed v

$$\#\{\tilde{v} : \text{supp}(q_v) \cap \text{supp}(q_{\tilde{v}}) \neq \emptyset\} \leq C_d. \quad (3.11)$$

In this work we will exploit the almost orthogonality of the decomposition (3.10) at the level of L^2 and also for products of these members in L^r for $1 \leq r \leq \infty$. We begin by considering the former; the more general theory will be considered in Lemma 4.2 and is adapted to $\tilde{\sigma}_\lambda$.

We first observe that by appealing to the FBI transform as in [38, Theorem 13.3], we have for any symbol⁴ $q \in S_{1/8}$

$$\|\text{Op}(q)\|_{L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)} \leq \|q\|_\infty + \mathcal{O}(\lambda^{-\frac{3}{4}}). \quad (3.12)$$

Since we may restrict attention to sufficiently large λ , we have that for any subcollection F of the v and an arbitrary sequence $\epsilon_v = \pm 1$,

$$\left\| \sum_{v \in F} \epsilon_v Q_v \right\|_{L^2 \rightarrow L^2}, \quad \left\| I - \sum_{v \in F} Q_v \right\|_{L^2 \rightarrow L^2}, \quad \|Q_v\|_{L^2 \rightarrow L^2} \leq 2. \quad (3.13)$$

Proposition 3.1 *Suppose $h \in L^2(\mathbb{R}^d)$ and the semiclassical Fourier transform of h is supported where $\tilde{\beta} = 1$. Then for λ sufficiently large,*

$$(4C_d)^{-1} \|h\|_{L^2}^2 \leq \sum_v \|Q_v h\|_{L^2}^2 \leq 4\|h\|_{L^2}^2. \quad (3.14)$$

Moreover, if F is any subcollection of the v ,

$$\sum_{v \in F} \|Q_v h\|_{L^2}^2 + \left\| h - \sum_{v \in F} Q_v h \right\|_{L^2}^2 \leq 4\|h\|_{L^2}^2. \quad (3.15)$$

⁴ This theorem can be applied to the rescaled symbol $q(\lambda^{-1/8}(x, \xi))$, which yields the decay rate of $\lambda^{-3/4}$ for the error term upon return to the original coordinates. Since we are working in a subcritical symbol class, the distinction between Weyl quantization and the standard one in (1.14) is inconsequential given change of quantization formulae.

Proof We begin with the first inequality in (3.14). The symbolic calculus means that if $\text{supp}(q_\nu) \cap \text{supp}(q_{\tilde{\nu}}) = \emptyset$, then

$$\|Q_{\tilde{\nu}}^* \circ Q_\nu h\|_{L^2 \rightarrow L^2} \lesssim_N \lambda^{-N}. \quad (3.16)$$

Taking $N > \frac{d-1}{4}$ here we see that for λ sufficiently large

$$\|h\|_{L^2}^2 \leq 2 \sum \{ \langle Q_\nu h, Q_{\tilde{\nu}} h \rangle_{L^2} : \text{supp}(q_\nu) \cap \text{supp}(q_{\tilde{\nu}}) \neq \emptyset \}.$$

Indeed, (3.16) ensures that the contribution of the remaining terms is negligible. An application of Cauchy–Schwarz now shows that the right hand side here is in turn bounded by $4C_d \sum_\nu \|Q_\nu h\|_{L^2}^2$, where C_d is as in (3.11).

Turning to the second inequality in (3.14), this is essentially a consequence of (3.13) and the fact that the constant in Khintchine’s inequality can be taken to be 1 when $p = 2$. More directly, consider the usual family of Rademacher functions $r_k(t) = \text{sgn}(\sin(2^k \pi t))$, $k = 1, 2, 3, \dots$, which are known to form an orthonormal sequence in $L^2([0, 1])$. Hence for any injection $\nu \mapsto k(\nu)$, (3.13) gives that

$$\sum_\nu \|Q_\nu h\|_{L^2}^2 = \int_0^1 \int_{\mathbb{R}^d} \left| \sum_\nu r_{k(\nu)}(t) Q_\nu h(x) \right|^2 dx dt \leq 4 \|h\|_{L^2}^2.$$

To see (3.15), now consider an injection $\nu \mapsto k(\nu)$ defined on F such that $k(\nu) \neq 1$ for all $\nu \in F$. The triangle inequality implies that for all t ,

$$\left| \sum_{\nu \in F} r_{k(\nu)}(t) q_\nu(x, \xi) + r_1(t) \left(1 - \sum_{\nu \in F} q_\nu(x, \xi) \right) \right| \leq 1.$$

Proceeding similarly, we now have

$$\begin{aligned} & \sum_{\nu \in F} \|Q_\nu h\|_{L^2}^2 + \left\| h - \sum_{\nu \in F} Q_\nu h \right\|_{L^2}^2 \\ &= \int_0^1 \int_{\mathbb{R}^d} \left| \sum_{\nu \in F} r_{k(\nu)}(t) Q_\nu h(x) + r_1(t) (I - \sum_{\nu \in F} Q_\nu) h(x) \right|^2 dx dt \\ &\leq 4 \|h\|_{L^2}^2. \end{aligned}$$

□

Note that (3.14) implies the crude bound

$$\|h\|_{L^2} \lesssim \lambda^{\frac{d-1}{16}} \sup_v \|Q_v h\|_{L^2}. \quad (3.17)$$

In Sect. 4, we will prove the following theorem:

Theorem 3.2 *Suppose h is supported in Ω with $\|h\|_{L^2(\Omega)} \leq 4$. Assume further that h satisfies (3.14), (3.17), and (3.10) (without error term). Let $\tilde{\sigma}_\lambda$ be as in (3.6) and define*

$$\tilde{A}_\alpha = \{x \in \Omega : |\tilde{\sigma}_\lambda h(x)| > \alpha\}.$$

Let $\delta_d = 2/(d+1)$ when $d \geq 3$ and $\delta_2 = 1/3$ when $d = 2$. Then

$$\sup \left\{ \alpha |\tilde{A}_\alpha|^{\frac{1}{p_c}} : \alpha \in (0, \lambda^{\frac{d-1}{4} + \frac{1}{8}}] \right\} \lesssim \lambda^{\frac{1}{p_c}} \left(\max_v \|Q_v h\|_{L^2}^{\delta_d} \right) + \lambda^{\frac{1}{p_c}}. \quad (3.18)$$

Here the second term on the right means λ is raised to some given power which is strictly less than $1/p_c$.

As we shall see in (4.8) below, the assumption $\alpha \in (0, \lambda^{\frac{d-1}{4} + \frac{1}{8}}]$ will allow us to exploit gains in bilinear estimates that correspond to the “subcritical” range of L^q spaces with $2 < q < p_c$. Indeed, (4.8) is a subtle but crucial observation in the present work, showing that weak bounds can be combined with known bilinear estimates to avoid the impediments presented by localizing the momenta in scales as fine as $\lambda^{-1/2}$ (a necessary technical difficulty in [9]). The second term in the right side of (3.18) corresponds to the gain in the bilinear estimates corresponding to angular separation larger than $\lambda^{-1/8}$ and the above assumptions on α , while the first term in the right side of (3.18) corresponds to the contribution to the bilinear estimate for near-diagonal terms corresponding to separation smaller than $\lambda^{-1/8}$. As noted above, there is nothing special about the power $1/8$. Any number between 0 and $1/2$ should work after adjusting the power of α correspondingly.

3.2 Finalizing the contradiction

Recall from (3.4), we have sequences $B_k, \lambda_k \rightarrow \infty$ and corresponding α_k satisfying

$$\frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k |\{|\tilde{\sigma}_k \rho_k f_k| > \alpha_k\}|^{\frac{1}{p_c}}, \quad 0 < \alpha_k \leq \lambda_k^{\frac{d-1}{4} + \frac{1}{8}}.$$

Here we use the same notation convention $\tilde{\sigma}_k = \tilde{\sigma}_{\lambda_k}$, and it is understood that the set on the right in the inequality is $\{x \in \Omega : |(\tilde{\sigma}_k \rho_k f_k)(x)| > \alpha_k\}$.

Recall that the semiclassical wave front set of the kernel of $\tilde{\sigma}_k$ is contained in $\Omega \times \{|\xi| \approx 1\}$ given the localization of the symbol $v(t, x, \xi)$ above. We therefore make a slight abuse of notation and assume that $\rho_k f_k$ satisfies the assumptions on h in Theorem 3.2, including having support in Ω , though strictly speaking this only applies to a microlocalization of this function. In particular we assume $\rho_k f_k$ satisfies (3.14), (3.17), and (3.10) (at the cost of shrinking the B_k and α_k one last time). By (3.15), for any set of $\{Q_{v_l}\}_{l=1}^L$ with $v_l \neq v_j$ when $j \neq l$,

$$\left\| \rho_k f_k - \sum_{l=1}^L Q_{v_l} \rho_k f_k \right\|_{L^2(\Omega)}^2 \leq 4 \|\rho_k f_k\|_{L^2(\Omega)}^2 - \sum_{l=1}^L \|Q_{v_l} \rho_k f_k\|_{L^2(\Omega)}^2 \quad (3.19)$$

Let $C(2C_d)^{-\delta_d}$ exceed the implicit constant in (3.18), where C_d is defined in (3.11). Take $N_k \in \mathbb{N}$ such that

$$\frac{N_k}{2} \leq 4 \left(\frac{2C(\log \lambda_k)^{\varepsilon_1}}{B_k} \right)^{2/\delta_d} < N_k.$$

If the middle expression is strictly less than $1/2$, take $N_k = 1$. We note for future use that in either case, we have

$$N_k = o((\log \lambda_k)^{2\varepsilon_1/\delta_d}).$$

We claim there exists a selection of distinct $Q_{v_1} \rho_k f_k, \dots, Q_{v_{N_k}} \rho_k f_k$, with Q_v as in Theorem 3.2, which satisfies for k sufficiently large

$$\left(\frac{B_k}{2C(\log \lambda_k)^{\varepsilon_1}} \right)^{1/\delta_d} \leq \|Q_{v_l} \rho_k f_k\|_{L^2(\Omega)}, \text{ for any } l = 1, \dots, N_k, \quad (3.20)$$

$$\frac{B_k \lambda_k^{1/p_c}}{2(\log \lambda_k)^{\varepsilon_1}} \leq \alpha_k \left\| \left\{ \tilde{\sigma}_k \left(\rho_k f_k - \sum_{l=1}^L Q_{v_l} \rho_k f_k \right) \right| > \frac{\alpha_k}{2} \right\|^{1/p_c}, \quad (3.21)$$

for any $L = 1, \dots, N_k$,

and in the latter case, $\tilde{\sigma}_k$ acts on the function in parentheses. We now show how to derive a contradiction assuming these two hold. Recall that the integral kernel of $\tilde{\sigma}_k$ is supported in $\Omega \times \Omega$. Hence (3.21) and the classical L^{p_c} bounds of the second author on $\tilde{\sigma}_k$ in (3.7) gives

$$\frac{B_k \lambda_k^{1/p_c}}{2(\log \lambda_k)^{\varepsilon_1}} \lesssim \lambda_k^{1/p_c} \left\| \rho_k f_k - \sum_{l=1}^{N_k} Q_{v_l} \rho_k f_k \right\|_{L^2(\Omega)}.$$

We now multiply by λ_k^{-1/p_c} , square both sides, and apply (3.19) and (3.20) to obtain

$$\begin{aligned} \left(\frac{B_k}{2(\log \lambda_k)^{\varepsilon_1}} \right)^2 &\lesssim 4\|\rho_k f_k\|_{L^2(\Omega)}^2 - \sum_{l=1}^{N_k} \|Q_{v_l} \rho_k f_k\|_{L^2(\Omega)}^2 \\ &\leq 4 - N_k \left(\frac{B_k}{2C(\log \lambda_k)^{\varepsilon_1}} \right)^{2/\delta_d}. \end{aligned}$$

Here we have used that $\|f_k\|_{L^2} = 1$ and our assumption $\|\rho\|_{L^\infty} \leq 1$ (cf. (1.6)). Since N_k is selected so that the right hand side is negative, we obtain a contradiction.

To see how to construct $Q_{v_1} \rho_k f_k, \dots, Q_{v_{N_k}} \rho_k f_k$, we proceed inductively. For any $L = 1, \dots, N_k - 1$, we show how to select the successive function in the collection given the previously chosen $Q_{v_1} \rho_k f_k, \dots, Q_{v_L} \rho_k f_k$ which satisfy (3.20), (3.21). The initial selection of Q_{v_1} is essentially the same, simply take $h_2 = 0$ in the following argument. Denote

$$h_1 = \rho_k f_k - \sum_{l=1}^L Q_{v_l} \rho_k f_k = \sum_{v \neq v_l} Q_{v_l} \rho_k f_k, \quad h_2 = \sum_{l=1}^L Q_{v_l} \rho_k f_k,$$

where the second expression for h_1 is a sum over all v distinct from the v_1, \dots, v_L . Then, by our assumptions

$$\frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k \left| \left\{ |\tilde{\sigma}_k h_1| > \frac{\alpha_k}{2} \right\} \right|^{\frac{1}{p_c}} + \alpha_k \left| \left\{ |\tilde{\sigma}_k h_2| > \frac{\alpha_k}{2} \right\} \right|^{\frac{1}{p_c}}. \quad (3.22)$$

Our first main claim is that we can use Corollary 2.2 to see that

$$\alpha_k \left| \left\{ |\tilde{\sigma}_k h_2| > \frac{\alpha_k}{2} \right\} \right|^{\frac{1}{p_c}} = o\left(\lambda_k^{1/p_c} (\log \lambda_k)^{-\varepsilon_1}\right). \quad (3.23)$$

We initially observe the following $L^2 \rightarrow L^{p_c}$ “commutator bounds”

$$\|\tilde{\sigma}_\lambda Q_v - Q_v \tilde{\sigma}_\lambda\|_{L^2(\Omega) \rightarrow L^{p_c}(\Omega)} \lesssim \lambda^{-\frac{1}{4}}. \quad (3.24)$$

Morally, this is Sobolev embedding and Egorov’s theorem combined with the invariance of q_v under the flow χ_t . However, we give a direct proof below that will be shown after the related Lemma 4.3. Assuming (3.24), we use properties

of the distribution function and Chebyshev's inequality to get

$$\alpha_k^{p_c} \left| \left\{ |\tilde{\sigma}_k h_2| > \frac{\alpha_k}{2} \right\} \right| \lesssim \sum_{l=1}^L \alpha_k^{p_c} \left| \left\{ |Q_{v_l} \tilde{\sigma}_k \rho_k f_k| > \frac{\alpha_k}{4L} \right\} \right| + L \lambda^{-\frac{p_c}{4}}.$$

Since the $1/p_c$ power of the second term on the right is much stronger than the bounds in (3.23), we are left with estimating the first term on the right hand side.

We next observe that for any $p \in [1, \infty]$, $\|Q_{v_l}\|_{L^p \rightarrow L^p} \lesssim 1$, which will allow us to eliminate $\tilde{\sigma}_k$ and apply Corollary 2.2. Indeed, given that q_v satisfies (2.16) with $\omega(x) = \omega(x, v)$, integration by parts in the expression for the integral kernel $Q_{v_l}(x, y)$ yields the pointwise bounds

$$|Q_{v_l}(x, y)| \lesssim \lambda^{1+\frac{7}{8}(d-1)} (1 + \lambda |\omega(x, v) \cdot (x - y)| + \lambda^{\frac{7}{8}} |x - y|)^{-(d+1)}.$$

Hence the uniform bounds on L^p follow from the generalized Young's inequality. By Chebyshev's inequality, (3.24), and (3.3), we now have

$$\alpha_k^{p_c} \left| \left\{ |Q_{v_l} \tilde{\sigma}_k \rho_k f_k| > \frac{\alpha}{4L} \right\} \right| \lesssim \alpha_k^{p_c} \left| \left\{ |Q_{v_l} \rho_k f_k| > \frac{\alpha}{8L} \right\} \right| + \lambda (\log \lambda)^{-p_c},$$

and as before the last term on the right is of the desired size in (3.23).

We may now use that Corollary 2.2 yields the following bound,

$$\sum_{l=1}^L \alpha_k^{p_c} \left| \left\{ |Q_{v_l} \rho_k f_k| > \frac{\alpha}{8L} \right\} \right| \lesssim L^{1+p_c} \lambda_k c(\lambda_k)^{\frac{p_c}{d+1}} \leq \lambda_k N_k^{1+p_c} c(\lambda_k)^{\frac{p_c}{d+1}}. \quad (3.25)$$

To see that (3.23) now follows, take p_c -th roots of both sides here and recall that $N_k = o((\log \lambda_k)^{2\varepsilon_1/\delta_d})$. The condition on the exponent in Remark 1.2 and the relation $\varepsilon_1 = \frac{(d+1)\varepsilon_0}{2}$ then implies

$$\begin{cases} \frac{2\varepsilon_1}{\delta_2} \left(\frac{p_c+1}{p_c} \right) - \frac{1}{2\cdot 3} = -\varepsilon_1, & \text{if } d = 2, \\ \frac{2\varepsilon_1}{\delta_3} \left(\frac{p_c+1}{p_c} \right) - \frac{1}{4} < -\varepsilon_1, & \text{if } d = 3, \\ \frac{2\varepsilon_1}{\delta_d} \left(\frac{p_c+1}{p_c} \right) - \frac{1}{d+1} = -\varepsilon_1, & \text{if } d \geq 4. \end{cases} \quad (3.26)$$

We show the details behind this when $d \geq 4$ so that $c(\lambda) = (\log \lambda)^{-1}$, and note that the other cases are verified similarly. Given the prior observation on the size of N_k , the p_c -th root of the right hand side of (3.25) is

$$\mathcal{O} \left(\lambda_k^{\frac{1}{p_c}} N_k^{1+\frac{1}{p_c}} c(\lambda_k)^{\frac{1}{d+1}} \right) = o \left(\lambda_k^{\frac{1}{p_c}} (\log \lambda_k)^{\frac{2\varepsilon_1}{\delta_d} \left(\frac{p_c+1}{p_c} \right)} (\log \lambda_k)^{-\frac{1}{d+1}} \right).$$

It is now an easy algebraic computation to see that the choice of ε_0 in Remark 1.2 means that ε_1 satisfies (3.26). The improvements on the exponent for negatively curved manifolds claimed in Remarks 1.2 and 2.4 follow since the equation for $d \geq 4$ in (3.26) is now satisfied for $d = 2, 3$.

Given (3.22) and (3.23), for k large enough and independently of L ,

$$\frac{3}{4} \frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < \alpha_k \left| \left\{ |\tilde{\sigma}_\lambda h_1| > \frac{\alpha_k}{2} \right\} \right|^{\frac{1}{p_c}}. \quad (3.27)$$

We are now left to show that there exists $Q_{v_{L+1}}$, distinct from those previously chosen, such that $Q_{v_{L+1}} \rho_k f_k$ also satisfies the bounds in (3.20), i.e.,

$$\frac{1}{2} \frac{B_k}{(\log \lambda_k)^{\varepsilon_1}} \leq C \|Q_{v_{L+1}} \rho_k f_k\|_{L^2(\Omega)}^{\delta_d}. \quad (3.28)$$

Indeed, once this is shown (3.21) can be concluded by taking $h_1 = \rho_k f_k - \sum_{l=1}^{L+1} Q_{v_l} \rho_k f_k$ and $h_2 = \sum_{l=1}^{L+1} Q_{v_l} \rho_k f_k$ in (3.22) and using (3.23) once again.

Given (3.13), we have $\|h_1\|_{L^2} \leq 4$ and hence by (3.27) and Theorem 3.2, there exists v_{\max} such that

$$\frac{3}{4} \frac{B_k \lambda_k^{1/p_c}}{(\log \lambda_k)^{\varepsilon_1}} < C(2C_d)^{-\delta_d} \lambda^{1/p_c} \left\| \sum_{v \neq v_l} Q_{v_{\max}} Q_v \rho_k f_k \right\|_{L^2(\Omega)}^{\delta_d}, \quad (3.29)$$

where the sum in the expression on the right is over all v distinct from each of the v_l , $l = 1, \dots, L$. Here we have used our assumption that $C(2C_d)^{-\delta_d}$ exceeds the implicit constant in (3.18). Now take v_{L+1} so that

$$\|Q_{v_{\max}} Q_{v_{L+1}} \rho_k f_k\|_{L^2} = \max \{ \|Q_{v_{\max}} Q_v \rho_k f_k\|_{L^2} : v \neq v_1, \dots, v_L \}, \quad (3.30)$$

so that $Q_{v_{L+1}}$ is distinct from the previously chosen operators. Note that by the symbolic calculus

$$\|Q_v \circ Q_{v'}\|_{L^2 \rightarrow L^2} \lesssim_N \lambda^{-N}, \quad \text{if } \text{supp}(q_v) \cap \text{supp}(q_{v'}) = \emptyset. \quad (3.31)$$

We therefore must have $\text{supp}(q_{v_{\max}}) \cap \text{supp}(q_{v_{L+1}}) \neq \emptyset$ in (3.30), since otherwise (3.31) would imply a contradiction of (3.29). Hence (3.11), (3.30), and taking $N > 1 + \frac{d-1}{8}$ in (3.31) then yields for k large enough

$$\begin{aligned} \left\| \sum_{\nu \neq \nu_l} Q_{\nu_{\max}} Q_{\nu} \rho_k f_k \right\|_{L^2(\Omega)} &\leq C_d \| Q_{\nu_{\max}} Q_{\nu_{L+1}} \rho_k f_k \|_{L^2} + \lambda_k^{-1} \\ &\leq 2C_d \| Q_{\nu_{L+1}} \rho_k f_k \|_{L^2} + \lambda_k^{-1}. \end{aligned}$$

Combining this with (3.29) then gives (3.28) for large enough k .

Remark 3.3 The condition (3.26) is the strongest limitation on ε_1 , which in turn gives ε_0 as in (1.10). Indeed, the only other assumption was $\varepsilon_1 \leq \frac{1}{d+1} < 1$ in (3.2) and (3.3). Also, with some small changes in the exposition, the arguments here show that in the $d = 3$ case of Theorem 1.1, we have

$$\| \rho(T(\lambda - P)) \|_{L^2(M) \rightarrow L^4(M)} \lesssim \frac{\lambda^{1/p_c} (\log \log \lambda)^{1/4}}{(\log \lambda)^{1/48}}.$$

4 Almost orthogonality and bilinear estimates

Here we prove Theorem 3.2, which involves bilinear estimates and almost orthogonality in the spirit of the prior works of the authors [7–9, 27].

4.1 Whitney-type decompositions and the key lemmas

Recall that ν indexes a $\approx \lambda^{-1/8}$ separated set in a neighborhood of $(1, 0, \dots, 0)$ on \mathbb{S}^{d-1} . Given (3.10), we may write

$$(\tilde{\sigma}_\lambda h)^2 = \sum_{\nu, \tilde{\nu}} (\tilde{\sigma}_\lambda Q_\nu h) (\tilde{\sigma}_\lambda Q_{\tilde{\nu}} h). \quad (4.1)$$

We may thus view this neighborhood of $(1, 0, \dots, 0)$ as a graph in the last $d - 1$ variables, and given $\nu, \tilde{\nu} \in \mathbb{S}^{d-1}$, we let $\nu', \tilde{\nu}'$ denote the projection of these vectors onto the last $d - 1$ coordinates. This allows us to organize the sum here in a fashion similar to that in [9, p. 513], which in turn is analogous to the Whitney decomposition taken in [36, §2.5].

Consider the standard family of dyadic cubes in \mathbb{R}^{d-1} with $\tau_{\mu'}^j$ denoting the translation of $[0, 2^j]^{d-1}$ by $\mu' \in 2^j \mathbb{Z}^{d-1}$. Two dyadic cubes of sidelength 2^j are declared to be *close* if they are not adjacent, but have adjacent parents of sidelength 2^{j+1} , and in this case we write $\tau_{\mu'}^j \sim \tau_{\tilde{\mu}'}^j$. Note that close cubes satisfy $d(\tau_{\mu'}^j, \tau_{\tilde{\mu}'}^j) \approx 2^j$. As noted in [36, p. 971], any two distinct points $\nu', \tilde{\nu}' \in \mathbb{R}^{d-1}$ lie in a unique pair of close cubes, that is, there exists a unique triple j, μ, μ' such that $(\nu', \tilde{\nu}') \in \tau_{\mu'}^j \times \tau_{\tilde{\mu}'}^j$ and $\tau_{\mu'}^j \sim \tau_{\tilde{\mu}'}^j$. We remark that in what follows we only need to consider $j \leq 0$.

Let J be the integer satisfying $2^{J-1} < 8\lambda^{-1/8} \leq 2^J$. The observations above now imply that the sum in (4.1) can be organized as

$$\left(\sum_{j=J+1}^0 \sum_{(\nu', \tilde{\nu}') \in \tau_{\mu'}^j \times \tau_{\tilde{\mu}'}^j : \tau_{\mu'}^j \sim \tau_{\tilde{\mu}'}^j} + \sum_{(\nu', \tilde{\nu}') \in \Xi_J} \right) (\tilde{\sigma}_\lambda Q_\nu h)(\tilde{\sigma}_\lambda Q_{\tilde{\nu}} h), \quad (4.2)$$

where Ξ_J indexes the remaining pairs such that $|\nu' - \tilde{\nu}'| \lesssim \lambda^{-1/8}$, including the on-diagonal pairs $\nu' = \tilde{\nu}'$. To see that this does indeed rewrite the sum in (4.1), note that if $\nu' \neq \tilde{\nu}'$, then as observed above $(\nu', \tilde{\nu}') \in \tau_{\mu'}^j \times \tau_{\tilde{\mu}'}^j$ for some unique pair of close cubes. If $j \leq J$, then we say $(\nu', \tilde{\nu}') \in \Xi_J$. Otherwise, it is included in the first sum in (4.2). Note that

$$\text{for each } \nu', \# \{ \tilde{\nu}' : (\nu', \tilde{\nu}') \in \Xi_J \} = \mathcal{O}(1). \quad (4.3)$$

For $j > J$, we define Ξ_j differently, indexing

$$\Xi_j := \{(\mu', \tilde{\mu}') \in 2^j \mathbb{Z}^{2(d-1)} : \tau_{\mu'}^j \sim \tau_{\tilde{\mu}'}^j\}.$$

Let $\mu \in \mathbb{S}^{d-1}$ be the vector with positive first coordinate and last $d-1$ coordinates given by μ' . Define

$$Q_{j,\mu} h := \sum_{\nu' \in \tau_{\mu'}^j} Q_\nu h,$$

so that

$$\sum_{(\nu', \tilde{\nu}') \in \tau_{\mu'}^j \times \tau_{\tilde{\mu}'}^j : \tau_{\mu'}^j \sim \tau_{\tilde{\mu}'}^j} (\tilde{\sigma}_\lambda Q_\nu h)(\tilde{\sigma}_\lambda Q_{\tilde{\nu}} h) = \sum_{(\mu', \tilde{\mu}') \in \Xi_j} (\tilde{\sigma}_\lambda Q_{j,\mu} h)(\tilde{\sigma}_\lambda Q_{j,\tilde{\mu}} h).$$

Now define a semiclassical symbol $\tilde{q}_{j,\mu}$ satisfying

$$\begin{aligned} \tilde{q}_{j,\mu}(x, \xi) \left(\sum_{\nu' \in \tau_{\mu'}^j} \tilde{q}_\nu(x, \xi) \right) &= \sum_{\nu' \in \tau_{\mu'}^j} \tilde{q}_\nu(x, \xi), \\ \left| \langle \omega(x, \mu), d_\xi \rangle^j \partial_{x,\xi}^\beta \tilde{q}_{j,\nu} \right| &\lesssim_{\beta,j} 2^{-j|\beta|}, \\ \text{supp}(\tilde{q}_{j,\mu}) &\subset \left\{ (x, \xi) \in T^*\Omega : \left| \xi / |\xi|_{g(x)} - \omega(x, \mu) \right|_{g(x)} \lesssim 2^j, |\xi| \approx 1 \right\}, \end{aligned} \quad (4.4)$$

where $\omega(x, \mu) \in S_x^* \Omega$ is the covector of the unit speed geodesic passing through x and whose covector takes the form $\mu/|\mu|_{g(x)}$ as it passes through the $x_1 = 0$ plane (see the discussion following (3.9)). As usual, denote $\tilde{Q}_{j,\mu} := \text{Op}(\tilde{q}_{j,\mu})$ with the usual quantization (1.14). Taking the support of $\tilde{q}_{j,\mu}$ suitably we may assume

$$\begin{aligned} d(\text{supp}(\tilde{q}_{j,\mu}), \text{supp}(\tilde{q}_{j,\tilde{\mu}})) &\approx 2^j, \quad \text{for } (\mu', \tilde{\mu}') \in \Xi_j, \quad j = J+1, \dots, 0, \\ \| (I - \tilde{Q}_{j,\mu}) \circ Q_{j,\mu} h \|_{L^2} &\lesssim_N \lambda^{-N}. \end{aligned} \tag{4.5}$$

Next define the bilinear operators

$$\begin{aligned} \Upsilon(h_1, h_2)(x) &:= (\tilde{\sigma}_\lambda h_1)(x) (\tilde{\sigma}_\lambda h_2)(x), \\ \Upsilon_{j,\mu,\tilde{\mu}}(h_1, h_2) &:= \Upsilon(\tilde{Q}_{j,\mu} h_1, \tilde{Q}_{j,\tilde{\mu}} h_2), \quad j = J+1, \dots, 0. \end{aligned}$$

This allows (4.1) to be rewritten as

$$\begin{aligned} (\tilde{\sigma}_\lambda h)^2 &= \Upsilon^{\text{diag}}(h) + \Upsilon^{\text{off}}(h) + \Upsilon^{\text{smooth}}(h), \\ \Upsilon^{\text{diag}}(h) &:= \sum_{(\nu', \tilde{\nu}') \in \Xi_J} (\tilde{\sigma}_\lambda Q_\nu h) (\tilde{\sigma}_\lambda Q_{\tilde{\nu}} h), \\ \Upsilon^{\text{off}}(h) &:= \sum_{j=J+1}^0 \sum_{(\mu, \tilde{\mu}) \in \Xi_j} \Upsilon_{j,\mu,\tilde{\mu}}(Q_{j,\mu} h, Q_{j,\tilde{\mu}} h), \\ \Upsilon^{\text{smooth}}(h) &:= \sum_{j=J+1}^0 \sum_{(\mu, \tilde{\mu}) \in \Xi_j} (\Upsilon(Q_{j,\mu} h, Q_{j,\tilde{\mu}} h) - \Upsilon_{j,\mu,\tilde{\mu}}(Q_{j,\mu} h, Q_{j,\tilde{\mu}} h)). \end{aligned} \tag{4.6}$$

Each term in the sum defining Υ^{smooth} can be rewritten as a sum of 3 terms, each of which contains a factor of the form $\sigma_\lambda((I - \tilde{Q}_{j,\mu}) \circ Q_{j,\mu} h)$ (or one with μ replacing $\tilde{\mu}$). Hence linear estimates on σ_λ in (3.7), almost orthogonality in (3.17), and taking N large in (4.5) implies

$$\|\Upsilon^{\text{smooth}}(h)\|_{L^{pc/2}} \lesssim \lambda^{-N} \|h\|_{L^2}^2 \lesssim \lambda^{\frac{2}{pc}} \left(\max_\nu \|Q_\nu h\|_{L^2}^{\delta_d} \right)^2.$$

Hence by (4.6)

$$\begin{aligned} |\{|\tilde{\sigma}_\lambda h| > \alpha\}| &\leq \left| \left\{ |\Upsilon^{\text{diag}}(h)| > \frac{\alpha^2}{3} \right\} \right| \\ &+ \left| \left\{ |\Upsilon^{\text{off}}(h)| > \frac{\alpha^2}{3} \right\} \right| + \left| \left\{ |\Upsilon^{\text{smooth}}(h)| > \frac{\alpha^2}{3} \right\} \right|. \end{aligned}$$

As observed, the last term here is easily bounded by Chebyshev's inequality. The following lemma shows that Υ^{off} satisfies stronger estimates as well and is closely related to [8, Theorem 2.1] and [9, Theorems 3.3, 3.4]:

Lemma 4.1 *Suppose $\frac{2(d+2)}{d} < q < \frac{2(d+1)}{d-1} = p_c$. Then for $j = J+1, \dots, 0$,*

$$\left\| \sum_{(\mu, \tilde{\mu}) \in \Xi_j} \Upsilon_{j, \mu, \tilde{\mu}} (Q_{j, \mu} h, Q_{j, \tilde{\mu}} h) \right\|_{L^{q/2}} \lesssim_q \lambda^{d-1-\frac{2d}{q}} 2^{j(d-1-\frac{2(d+1)}{q})} \|h\|_{L^2}^2. \quad (4.7)$$

To appreciate the gain furnished by this lemma in weak- L^{p_c} , the first two factors on the right in (4.7) should be raised to the power of $\frac{q}{2p_c}$, which is $\lambda^{\frac{1}{p_c}} (\lambda 2^j)^{\frac{d-1}{2}(\frac{q}{p_c}-1)}$. Summation in j thus gives for any fixed $q \in (\frac{2(d+2)}{d}, p_c)$

$$\begin{aligned} \alpha \left| \left\{ |\Upsilon^{\text{off}}(h)| > \frac{\alpha^2}{3} \right\} \right|^{\frac{1}{p_c}} &\lesssim \alpha^{1-\frac{q}{p_c}} \lambda^{\frac{1}{p_c}} (\lambda^{\frac{7}{8}})^{\frac{d-1}{2}(\frac{q}{p_c}-1)} \|h\|_{L^2}^{\frac{q}{p_c}} \\ &\lesssim \lambda^{\frac{1}{p_c}} (\alpha \lambda^{\frac{7-7d}{16}})^{1-\frac{q}{p_c}}. \end{aligned} \quad (4.8)$$

Since $\alpha \leq \lambda^{\frac{d-1}{4}+\frac{1}{8}}$, the quantity in parentheses in the last line is $\mathcal{O}(\lambda^{\frac{5-3d}{16}})$. Thus the right hand side can be bounded by the second term on right hand side of (3.18).

A step in the proof of (4.7) and in the treatment of $\Upsilon^{\text{diag}}(h)$ is to show the following almost orthgonality lemma, akin to [9, Theorem 3.3]. This establishes an almost orthgonality principle in L^r spaces, but with respect to the operators $\Upsilon_{j, \mu, \mu'}$ and their counterparts in the definition of Υ^{diag} . Hence this is a substantial variation on the L^2 almost orthgonality principle in Proposition 3.1.

Lemma 4.2 *For $1 \leq r \leq \infty$, set $r^* = \min(r, r')$ where r' is the Hölder conjugate of r . Then for any $j = J+1, \dots, 0$, and any $N \in \mathbb{N}$ large*

$$\begin{aligned} &\left\| \sum_{(\mu, \tilde{\mu}) \in \Xi_j} \Upsilon_{j, \mu, \tilde{\mu}} (Q_{j, \mu} h, Q_{j, \tilde{\mu}} h) \right\|_{L^r} \\ &\lesssim_N \left(\sum_{(\mu, \tilde{\mu}) \in \Xi_j} \|\Upsilon_{j, \mu, \tilde{\mu}} (Q_{j, \mu} h, Q_{j, \tilde{\mu}} h)\|_{L^r}^{r^*} \right)^{1/r^*} + \lambda^{-N} \|h\|_{L^r}, \quad (4.9) \end{aligned}$$

$$\|\Upsilon^{\text{diag}}(h)\|_{L^r} \lesssim_N \left(\sum_{(\nu, \tilde{\nu}) \in \Xi_J} \|(\tilde{\sigma}_\lambda Q_\nu h)(\tilde{\sigma}_\lambda Q_{\tilde{\nu}} h)\|_{L^r}^{r^*} \right)^{1/r^*} + \lambda^{-N} \|h\|_{L^r}. \quad (4.10)$$

Given Lemma 4.2, we have that with $r = p_c/2$,

$$\alpha \left| \left\{ |\Upsilon^{\text{diag}} h| > \frac{\alpha^2}{3} \right\} \right|^{\frac{1}{p_c}} \lesssim \left(\sum_{(v, \tilde{v}) \in \Xi_J} \|(\tilde{\sigma}_\lambda Q_v h)(\tilde{\sigma}_\lambda Q_{\tilde{v}} h)\|_{L^r}^{r^*} \right)^{1/r^*} + \lambda^{-N} \|h\|_{L^r}.$$

The factor λ^{-N} means that the second term here can be harmlessly absorbed into the second term in (3.18). Indeed, when $d \geq 3$, $r = p_c/2 \leq 2$ so Hölder's inequality can be applied as h is supported in Ω . When $d = 2$, Sobolev embedding can be applied instead. For the first sum on the right, the linear estimates (3.7) for $\tilde{\sigma}_\lambda$ shows it is bounded by

$$\lambda^{\frac{1}{p_c}} \left(\sum_{(v, \tilde{v}) \in \Xi_J} \|Q_v h\|_{L^2}^{r^*} \|Q_{\tilde{v}} h\|_{L^2}^{r^*} \right)^{\frac{1}{2r^*}}. \quad (4.11)$$

This sum can be treated very similarly to [9, §3] and previous works. We show the details for $d \geq 4$, as the other cases are handled similarly. When $d \geq 4$, we have $p_c/2 < 2$ and hence $(p_c/2)^* = p_c/2$. Using Hölder's inequality with $\frac{1}{r^*} = \frac{2}{p_c} = \frac{1}{2} + \frac{4-p_c}{2p_c}$ and (4.3), the sum is bounded by

$$\begin{aligned} & \left(\sum_v \|Q_v h\|_{L^2}^2 \right)^{\frac{1}{4}} \left(\sum_{\tilde{v}} \|Q_{\tilde{v}} h\|_{L^2}^{\frac{2p_c}{4-p_c}} \right)^{\frac{4-p_c}{4p_c}} \\ & \lesssim \left(\sum_v \|Q_v h\|_{L^2}^2 \right)^{\frac{1}{p_c}} \left(\sup_v \|Q_v h\|_{L^2}^{1-\frac{2}{p_c}} \right) \lesssim \|h\|_{L^2}^{\frac{2}{p_c}} \left(\sup_v \|Q_v h\|_{L^2}^{1-\frac{2}{p_c}} \right), \end{aligned}$$

and in the first inequality we use $\frac{2p_c}{4-p_c} = 2 + \frac{4p_c-8}{4-p_c}$ and (3.14). When $d = 3$, $\frac{p_c}{2} = 2$ so we use Hölder's inequality again but with $\frac{2}{p_c} = \frac{1}{2} + \frac{1}{\infty}$, making the first inequality extraneous. When $d = 2$, the argument is similar to the $d \geq 4$ case, but $(p_c/2)^* = 3/2 < p_c/2$, and the resulting exponent is $\delta_0 = \frac{1}{3}$. This completes the proof of Theorem 3.2 once we show Lemmas 4.1 and 4.2.

4.2 Composition of $\tilde{\sigma}_\lambda$ with PDO and bilinear estimates

Here we prove Lemma 4.1, assuming Lemma 4.2 for now. As stated above, it is nearly the same as [8, Theorem 2.1] or [9, Theorems 3.3, 3.4], but we are somewhat thorough here as there are differences in the constructions.

Given Lemma 4.2, we are reduced to showing that for $\frac{2(d+2)}{d} < q < \frac{2(d+1)}{d-1}$,

$$\begin{aligned} & \left\| \Upsilon_{j,\mu,\tilde{\mu}} (Q_{j,\mu} h, Q_{j,\tilde{\mu}} h) \right\|_{L^{q/2}} \\ & \lesssim_q \lambda^{d-1-\frac{2d}{q}} 2^{j(d-1-\frac{2(d+1)}{q})} \|Q_{j,\mu} h\|_{L^2} \|Q_{j,\tilde{\mu}} h\|_{L^2}. \end{aligned} \quad (4.12)$$

Indeed, if this holds, then given that for each $\mu \in 2^j \mathbb{Z}^{d-1}$ there are $\mathcal{O}(1)$ elements $\tilde{\mu}$ satisfying $(\mu, \tilde{\mu}) \in \Xi_j$ (similar to (4.3)), hence Lemma 4.2 with $r = q/2$ and Cauchy–Schwarz means it suffices to bound

$$\left(\sum_{(\mu, \tilde{\mu}) \in \Xi_j} \|Q_{j,\mu} h\|_{L^2}^{(\frac{q}{2})^*} \|Q_{j,\tilde{\mu}} h\|_{L^2}^{(\frac{q}{2})^*} \right)^{\frac{1}{(\frac{q}{2})^*}} \lesssim \left(\sum_{\mu \in 2^j \mathbb{Z}^{d-1}} \|Q_{j,\mu} h\|_{L^2}^{2(\frac{q}{2})^*} \right)^{\frac{1}{(\frac{q}{2})^*}}.$$

But given the almost orthogonality of the $\{Q_{j,\mu}\}_\mu$ (proved similarly to (3.13) and (3.14)) and the embedding $\ell^2 \hookrightarrow \ell^{2(\frac{q}{2})^*}$, the right hand side is $\mathcal{O}(\|h\|_{L^2}^2)$.

4.2.1 Composition of $\tilde{\sigma}_\lambda$ with PDO

To set the stage for bilinear estimates, we need to examine the effect of composing $\tilde{\sigma}_\lambda$ with the $\tilde{Q}_{j,\mu}$. Recall that $\omega(z, \mu)$ is the covector at z of the geodesic through z whose covector at the intersection of the $y^1 = 0$ plane is $\mu/|\mu|_g$. We use “#” to denote the isomorphism from $T_x^* M$ to $T_x M$ determined by the metric g (the “musical isomorphism”). Its inverse is denoted by “ \flat ”.

Lemma 4.3 *For any $N \in \mathbb{N}$, the kernel of $\tilde{\sigma}_\lambda \circ \tilde{Q}_{j,\mu}$ can be written*

$$(\tilde{\sigma}_\lambda \circ \tilde{Q}_{j,\mu})(x, z) = \lambda^{\frac{d-1}{2}} e^{i\lambda\Theta_g(x,z)} V_{j,\mu}(x, z) + \mathcal{O}(\lambda^{-N}), \quad (4.13)$$

where $V_{j,\mu} = 0$ unless $\Theta_g(x, z) \in (\frac{\tilde{c}_0}{4}, \frac{\tilde{c}_0}{2})$ and the unit covector at z of the geodesic from z to x lies in $\text{supp}(\tilde{q}_{j,\mu}(z, \cdot))$. Moreover, denoting $(\omega(x, \mu)^\#)^k$ as k applications of the vector field obtained by raising the indices of $\omega(x, \mu)$,

$$\left| (\omega(x, \mu)^\#)^k \partial_x^\beta V_{j,\mu}(x, z) \right| \lesssim_{k,\beta} 2^{-j|\beta|} \quad \text{for any } k, \beta. \quad (4.14)$$

Proof This is a small variation on the stationary phase arguments in [26, Lemma 5.1.3]. The kernel $(\tilde{\sigma}_\lambda \circ \tilde{Q}_{j,\mu})(x, z)$ is given by

$$\begin{aligned} & \frac{\lambda^{2d}}{(2\pi)^d} \int e^{i\lambda(t+\varphi(t,x,\xi)-y\cdot\xi+(y-z)\cdot\eta)} \hat{\rho}(t/\tilde{c}_0) \\ & \quad \times v(t, x, \xi) \tilde{\psi}(y) \tilde{q}_{j,\mu}(y, \eta) dt dy d\eta d\xi. \end{aligned} \quad (4.15)$$

The critical points of this oscillatory integral satisfy

$$\eta = \xi, \quad y = z, \quad p(x, d_x \varphi(t, x, \xi)) = 1, \quad y = d_\xi \varphi(t, x, \xi), \quad (4.16)$$

which arises from differentiation in y, η, t, ξ respectively. The third identity here uses the eikonal equation for φ . The last 2 identities fix t, y, ξ so that ξ lies in the cosphere bundle and $t = \Theta_g(x, z)$ is the time at which the minimizing geodesic through $(y, \xi) = (z, \xi)$ passes through $(x, d_x \varphi(t, x, \xi))$. Hence the kernel is $\mathcal{O}(\lambda^{-N})$ when $\Theta_g(x, z) \notin (\tilde{c}_0/4, \tilde{c}_0/2)$. Moreover, at the critical points, $\varphi(t, x, \xi) - y \cdot \xi = 0$ since φ is homogeneous of degree 1 in ξ . Stationary phase can be applied to (4.15) since the mixed Hessian in (t, ξ) is nonsingular, which follows from the same idea as in [26, p.140]. This yields the expression (4.13) and the claim concerning the support of $V_{j,\mu}$.

We are left to verify that applying $\omega(x, \mu)^\#$ to $V_{j,\mu}$ yields no loss in 2^{-j} . Note that $\xi = \eta$ as a function of (x, z) is the unit covector over z of the geodesic joining x and z . If $\gamma(t)$ parameterizes this geodesic with $\gamma(0) = z, \xi(\gamma(t), z)$ is constant in t and using the summation convention we have

$$\begin{aligned} 0 &= \partial_t(\xi_i(\gamma(t), z)) = \partial_{x^k} \xi_i(\gamma(t), z) \dot{\gamma}^k(t) \\ &= \partial_{x^k} \xi_i(\gamma(t), z) g^{jk}(\gamma(t)) (\dot{\gamma}^\flat)_j(t), \end{aligned}$$

where $\dot{\gamma}^\flat$ gives the unit covector of the geodesic at t . Since $V_{j,\mu} = 0$ unless $|\dot{\gamma}^\flat(0) - \omega(z, \mu)| \lesssim 2^j$, this shows the rest of (4.14). \square

Proof of (3.24) The main idea is that the leading order term in the stationary phase expansion of the integral kernels of $\tilde{\sigma}_\lambda \circ Q_v$ and $Q_v \circ \tilde{\sigma}_\lambda$. The remainder terms are then $\mathcal{O}(\lambda^{-3/4})$, at which point Sobolev embedding yields the desired gain for the remainder as $\frac{d}{2} - \frac{d}{p_c} = \frac{1}{p_c} + \frac{1}{2}$.

The kernel of $\tilde{\sigma}_\lambda \circ Q_v$ is just (4.15) but with $\tilde{q}_{j,\mu}(y, \eta)$ replaced by $q_v(y, \eta)$. The critical points are thus determined by (4.16) and since we have $\eta = \xi$ and $y = z$, we only need to consider ξ and t as functions of x, z . Recall that $t(x, z) = \Theta_g(x, z)$. Furthermore, the equation $z = d_\xi \varphi(t, x, \xi)$ means that $\xi(x, z)$ is the covector at z of the unit speed geodesic from z to x .

On the other hand, the kernel of $Q_v \circ \tilde{\sigma}_\lambda$ is

$$\begin{aligned} &\frac{\lambda^{2d}}{(2\pi)^d} \int e^{i\lambda(t+\varphi(t,y,\xi)-z\cdot\xi+(x-y)\cdot\eta)} \hat{\rho}(t/\tilde{c}_0) \\ &\quad \times v(t, y, \xi) \tilde{\psi}(z) q_v(x, \eta) dt dy d\eta d\xi. \end{aligned}$$

This time the critical points of the phase are given by

$$\eta = d_y \varphi(t, y, \xi), \quad x = y, \quad p(y, d_y \varphi(t, y, \xi)) = 1, \quad z = d_\xi \varphi(t, y, \xi).$$

This time we must treat t, η, ξ as functions of (x, z) . Once again, $t(x, z) = \Theta_g(x, z)$ and since $x = y$ on the critical set the equation $z = d_\xi \varphi(t, x, \xi)$ means that $\xi(x, z)$ has the same role as above. The equation $\eta = d_y \varphi(t, x, \xi)$ then implies that $\eta(x, z)$ is the covector at x of the unit speed geodesic from z to x .

Given the observations, the only possible difference between the leading order terms in the stationary phase expansions result from evaluating q_v at the two different sets of critical points. The first is $q_v(d_\xi \varphi(t, x, \xi), \xi(x, z))$, the second is $q_v(x, \eta(x, z))$. We now appeal to the observation following the definition of Q_v that q_v is invariant under the flow χ_t at points within the support of v . Thus since $\chi_t(d_\xi \varphi(t, x, \xi), \xi(x, z)) = \chi_t(z, \xi(x, z)) = (x, \eta(x, z))$ by the observations above, the proof is now complete. \square

4.2.2 Preliminaries for bilinear estimates and spatial localization

Taking N large in Lemma 4.3, we have that up to an error term which is a bilinear operator in (h_1, h_2) satisfying stronger $L^{q/2}$ bounds,

$$\begin{aligned} & \Upsilon_{j,\mu,\tilde{\mu}}(h_1, h_2)(x) \\ &= \lambda^{d-1} \iint e^{i\lambda(\Theta_g(x,z) + \Theta_g(x,\tilde{z}))} V_{j,\mu}(x, z) V_{j,\tilde{\mu}}(x, \tilde{z}) h_1(z) h_2(\tilde{z}) dz d\tilde{z} \\ & \quad + \text{error.} \end{aligned} \tag{4.17}$$

Next, we claim that when $V_{j,\mu}(x, z) V_{j,\tilde{\mu}}(x, \tilde{z}) \neq 0$,

$$|d_x \Theta_g(x, z) - d_x \Theta_g(x, \tilde{z})|_{g(x)} \approx 2^j. \tag{4.18}$$

To see this, recall $d_x \Theta_g(x, z), d_x \Theta_g(x, \tilde{z})$ give the unit covector at x of the unit speed geodesic from z to x , \tilde{z} to x respectively. Now consider the coordinates w', \tilde{w}' on the $w_1 = 0$ plane where the geodesics through (x, z) and (x, \tilde{z}) intersect this plane and let $\omega, \tilde{\omega}$ denote the unit covectors of the geodesic at the respective intersection points, see Fig. 1. The assumption on x, z, \tilde{z} ensures that with respect to Euclidean distance determined by the coordinate system, we have $|\omega - \tilde{\omega}| \approx 2^j$. Also note that it suffices to show (4.18) with the intrinsic distance replaced by Euclidean. Since the geodesics can be reparameterized in terms of the first coordinate x_1 , we have that

$$|d_x \Theta_g(x, z) - d_x \Theta_g(x, \tilde{z})| \approx |w' - \tilde{w}'| + |\omega - \tilde{\omega}| \approx |w' - \tilde{w}'| + 2^j,$$

as the covectors on the very left here are over the same point. The lower bound in (4.18) now follows. For the upper bound, integrating the equation for the

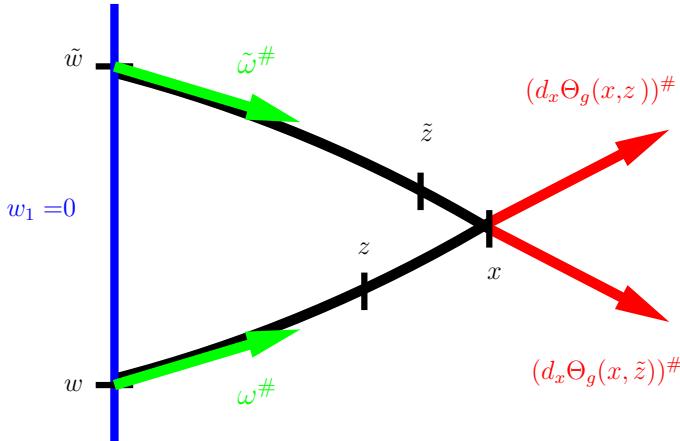


Fig. 1 Intersecting geodesics at x passing through z, \tilde{z} and their tangent vectors as raised covectors

w' components of the geodesics parameterized by x_1 gives

$$|w' - \tilde{w}'| \lesssim |x_1| |d_x \Theta_g(x, z) - d_x \Theta_g(x, \tilde{z})| \lesssim \epsilon |d_x \Theta_g(x, z) - d_x \Theta_g(x, \tilde{z})|.$$

Taking $\epsilon > 0$ small in (3.5), the rest of (4.18) follows.

Redefine β to be a smooth bump function satisfying for $w' \in \mathbb{R}^{d-1}$,

$$\sum_{l \in 2^j \mathbb{Z}^{d-1}} \beta^2 \left(2^{-j} (w' - l) \right) = 1, \quad \text{supp}(\beta) \subset [-1, 1]^{d-1}. \quad (4.19)$$

We now return to Φ as in (3.9) and let $\Upsilon_{j, \mu, \tilde{\mu}, l}$ denote the operator defined by replacing the amplitude $V_{j, \mu}(x, z) V_{j, \tilde{\mu}}(x, \tilde{z})$ in (4.17) by

$$V_{j, \mu, \tilde{\mu}, l}(x, z, \tilde{z}) := V_{j, \mu}(x, z) V_{j, \tilde{\mu}}(x, \tilde{z}) \beta(2^{-j}(\Phi(x, \omega(x, \mu)) - l)), \quad (4.20)$$

where as before, $\omega(x, \mu)$ is the covector of the unit speed geodesic passing through x such that the covector at the intersection point with $y_1 = 0$ is $\mu/|\mu|_g$ (cf. (4.4) and the discussion after (3.9)). Hence this geodesic has coordinates $(\Phi(x, \omega(x, \mu)), \mu)$ at its intersection point with $y_1 = 0$. The introduction of the bump function thus has the effect of localizing the amplitude of the oscillatory integral operator to a 2^j neighborhood of the geodesic which passes through the $y_1 = 0$ plane at $(0, l)$ with unit covector $\mu/|\mu|_g$.

We now claim that while (4.20) localizes the kernel in the x coordinates, the support properties of the $V_{j, \mu}$ in Lemma 4.3 imply that

$$\text{if } V_{j, \mu}(x, z) \beta(2^{-j}(\Phi(x, \omega(x, \mu)) - l)) \neq 0,$$

$$\text{then } |\Phi(z, \omega(z, \mu)) - l| \lesssim 2^j. \quad (4.21)$$

To see this, consider the unit speed geodesic joining z to x and let the unit covectors of the geodesic at these point be denoted by $\tilde{\omega}_z, \tilde{\omega}_x$ respectively. Hence $\Phi(z, \tilde{\omega}_z) = \Phi(x, \tilde{\omega}_x)$ as this geodesic segment will intersect the $y_1 = 0$ at a unique point. We now have if the hypothesis in (4.21) is satisfied, then

$$\begin{aligned} & |\Phi(z, \omega(z, \mu)) - l| \\ & \leq |\Phi(z, \omega(z, \mu)) - \Phi(z, \tilde{\omega}_z)| + |\Phi(x, \tilde{\omega}_x) - \Phi(x, \omega(x, \mu))| \\ & \quad + |\Phi(x, \omega(x, \mu)) - l| \\ & \lesssim |\omega(z, \mu) - \tilde{\omega}_z| + |\omega(x, \mu) - \tilde{\omega}_x| + 2^j, \end{aligned}$$

where we have used the Lipschitz bounds on Φ . Since $V_{j,\mu}(x, z) \neq 0$, Lemma 4.3 and (4.4) imply that $(z, \omega_z) \in \text{supp}(\tilde{q}_{j,\mu}(z, \cdot))$ and $|\omega(z, \mu) - \tilde{\omega}_z| \lesssim 2^j$. Recalling the discussion after (3.9), $\Psi(x, \cdot), \Psi(z, \cdot)$ are invertible and hence

$$\begin{aligned} |\omega(x, \mu) - \omega_x| & \lesssim |\Psi(x, \omega(x, \mu)) - \Psi(x, \omega_x)| \\ & = |\mu - \Psi(z, \omega_z)| \lesssim |\omega(z, \mu) - \tilde{\omega}_z|, \end{aligned}$$

using that $\omega(x, \cdot)$ inverts $\Psi(x, \cdot)$. The claim in (4.21) now follows.

We now claim it suffices to prove (4.12), with $\Upsilon_{j,\mu,\tilde{\mu}}$ replaced by $\Upsilon_{j,\mu,\tilde{\mu},l}$. First note that while the identity in (4.19) applies to the square sum over the β , we have more generally

$$\sum_{l \in 2^j \mathbb{Z}^{d-1}} \beta^r \left(2^{-j}(w' - l) \right) \approx_r 1, \quad 0 < r < \infty, \quad (4.22)$$

since for any l ,

$$\#\{\tilde{l} \in 2^j \mathbb{Z}^{d-1} : \text{supp}(\beta(2^{-j}(\cdot - l))) \cap \text{supp}(\beta(2^{-j}(\cdot - \tilde{l})))\} = \mathcal{O}(1).$$

Hence Hölder's inequality gives

$$\|\Upsilon_{j,\mu,\tilde{\mu}}(Q_{j,\mu}h, Q_{j,\tilde{\mu}}h)\|_{L^{q/2}} \lesssim \left(\sum_l \|\Upsilon_{j,\mu,\tilde{\mu},l}(Q_{j,\mu}h, Q_{j,\tilde{\mu}}h)\|_{L^{q/2}}^{\frac{q}{2}} \right)^{\frac{2}{q}} \quad (4.23)$$

We now apply (4.21) to see that there exists a bump function $\tilde{\beta}$ such that $\tilde{\beta}_{\mu,l}(z) := \tilde{\beta}(2^{-j}(\Phi(z, \omega(z, \mu)) - l))$ satisfies

$$\Upsilon_{j,\mu,\tilde{\mu},l}(Q_{j,\mu}h, Q_{j,\tilde{\mu}}h) = \Upsilon_{j,\mu,\tilde{\mu},l}(\tilde{\beta}_{\mu,l}Q_{j,\mu}h, \tilde{\beta}_{\tilde{\mu},l}Q_{j,\tilde{\mu}}h)$$

If (4.12) holds with the additional localization, i.e. this holds with $\Upsilon_{j,\mu,\tilde{\mu},l}$ replacing $\Upsilon_{j,\mu,\tilde{\mu}}$, then the more general bound holds since

$$\begin{aligned} & \left(\sum_l \left\| \tilde{\beta}_{\mu,l} Q_{j,\mu} h \right\|_{L^2}^{\frac{q}{2}} \left\| \tilde{\beta}_{\tilde{\mu},l} Q_{j,\tilde{\mu}} h \right\|_{L^2}^{\frac{q}{2}} \right)^{\frac{2}{q}} \\ & \lesssim \left(\sum_l \left\| \tilde{\beta}_{\mu,l} Q_{j,\mu} h \right\|_{L^2}^q \right)^{\frac{1}{q}} \left(\sum_l \left\| \tilde{\beta}_{\tilde{\mu},l} Q_{j,\tilde{\mu}} h \right\|_{L^2}^q \right)^{\frac{1}{q}} \\ & \lesssim \| Q_{j,\mu} h \|_{L^2} \| Q_{j,\tilde{\mu}} h \|_{L^2}, \end{aligned}$$

where the first inequality here uses Cauchy–Schwarz and the second uses the embedding $\ell^2 \hookrightarrow \ell^q$ along with the fact that we may assume that (4.22) holds with β replaced by $\tilde{\beta}$.

4.2.3 Fermi coordinates, parabolic scaling, and the proof of Lemma 4.1

As noted above, the additional localization means that $V_{j,\mu,\tilde{\mu},l}$ vanishes unless x is in a 2^j -neighborhood of the image of the geodesic $\gamma_{l,\mu}$ which passes through the $x_1 = 0$ plane at $(0, l)$ with unit covector μ . Given x , let $\omega_{l,\mu}(x)$ denote the unit covector on $\gamma_{l,\mu}$ at the closest point to x . The same idea in Lemma 4.3 shows that since $\omega_{l,\mu}(x)$ is within a distance of $\mathcal{O}(2^j)$ to both $\omega(x, \mu)$ and $\omega(x, \tilde{\mu})$,

$$\left| (\omega_{l,\mu}(x)^\#)^k \partial_x^\beta V_{j,\mu,\tilde{\mu},l}(x, z, \tilde{z}) \right| \lesssim_{k,\beta} 2^{-j|\beta|}.$$

The additional localization of $\Upsilon_{j,\mu,\tilde{\mu},l}$ now allows us to change to Fermi coordinates which straighten $\gamma_{l,\mu}$ so that the x_1 -coordinate parameterizes the geodesic and $|x'| = \Theta_g(x, \gamma)$. The regularity bounds for $V_{j,\mu,\tilde{\mu},l}$ transform as

$$\left| \partial_{x_1}^k \partial_x^\beta V_{j,\mu,\tilde{\mu},l}(x, z, \tilde{z}) \right| \lesssim_{k,\beta} 2^{-j|\beta|}.$$

Note that (4.18) holds in these coordinates as it is intrinsic to (M, g) . Moreover, in these coordinates we may still view the cosphere bundle as a graph in the last $d - 1$ variables, and in particular for each x , there exists a strictly concave function $\xi' \mapsto r(x, \xi')$ which defines $|\xi|_{g(x)} = 1$ in that

$$\partial_{x_1} \Theta_g(x, z) = r(x, d_{x'} \Theta_g(x, z)).$$

For $i = 1, 2$, let $h_i^{z_1}(z') = h_i(z_1, z')$. After an application of Minkowski's inequality, (4.12) is reduced to showing that uniformly in z_1, \tilde{z}_1 we have

$$\begin{aligned} & \left\| \iint e^{i\lambda(\Theta_g(x,z)+\Theta_g(x,\tilde{z}))} V_{j,\mu,\tilde{\mu},l}(x,z,\tilde{z}) h_1^{z_1}(z') h_2^{\tilde{z}_1}(\tilde{z}') dz' d\tilde{z}' \right\|_{L^{\frac{q}{2}}(\mathbb{R}^d)} \\ & \lesssim \lambda^{-\frac{2d}{q}} 2^{j(d-1-\frac{2(d+1)}{q})} \|h_1^{z_1}\|_{L^2(\mathbb{R}^{d-1})} \|h_2^{\tilde{z}_1}\|_{L^2(\mathbb{R}^{d-1})}, \end{aligned} \quad (4.24)$$

where we have cancelled the factor λ^{d-1} from (4.17) and the right hand side of (4.12). This will follow from the parabolic rescaling $(x', z') \mapsto (2^j x', 2^j z')$, rewriting the oscillatory factor as

$$e^{i\lambda 2^{2j}(\tilde{\Theta}(x,z)+\tilde{\Theta}(x,\tilde{z}))} \quad \text{with} \quad \tilde{\Theta}(x,z) = 2^{-2j} \Theta_g(x_1, 2^j x', z_1, 2^j z').$$

The dilation ensures that the derivatives of the amplitude are uniformly bounded in λ and j . Moreover, in the new coordinates $\partial_{x_1} \tilde{\Theta} = \tilde{r}(x, d_{x'} \tilde{\Theta})$ with $\tilde{r}(x, \xi') = 2^{-2j} r(x_1, 2^j x', 2^j \xi')$, whose Hessian in ξ' satisfies the same bounds as r . Hence (4.24) follows from [8, Theorem 3.3], which removes the ε -loss in the bilinear estimates of Lee [21, Theorem 1.1]. The latter in turn generalizes bilinear Fourier restriction estimates of [35, 37]. Indeed, (4.18) ensures that over the support of the dilated amplitude,

$$|d_{x'} \tilde{\Theta}(x, z) - d_{x'} \tilde{\Theta}(x, \tilde{z})| \approx 1,$$

which with the concavity of $\tilde{r}(x, \cdot)$, can be seen to be sufficient for the condition⁵ in these theorems, as the differentials are uniformly transverse in the graph of \tilde{r} . When $d = 2$, one could also show (4.24) by the method in [7, Lemma 3.3] which follows the approach of Hörmander [17] at the valid endpoint $q = 4$ and does not require a dilation of coordinates.

4.3 Almost orthogonality

In this section, we prove Lemma 4.2, primarily focusing on (4.9). The principle is essentially the same as in the proofs of [7, (3-4), (3-10)] or [23, Lemma 6.7] and is a “variable coefficient” version of the almost orthogonality principle in [36, Lemma 6.1]. The cases $r = 1, \infty$ follow from the triangle inequality, so it suffices to consider $r = 2$ and interpolate. Note that when $r = 2$, the left hand side of (4.9) is

$$\sum_{(\mu_1, \tilde{\mu}_1), (\mu_2, \tilde{\mu}_2) \in \Xi_j} \langle \Upsilon_{j,\mu_1, \tilde{\mu}_1} (Q_{j,\mu_1} h, Q_{j,\tilde{\mu}_1} h), \Upsilon_{j,\mu_2, \tilde{\mu}_2} (Q_{j,\mu_2} h, Q_{j,\tilde{\mu}_2} h) \rangle_{L^2}.$$

⁵ In comparison to [8], x in the present work plays the role of $z = (x, s)$ there.

For any fixed C , and any μ_1 , we have $\#\{\mu_2 : |\mu_1 - \mu_2| \leq C2^j\} = \mathcal{O}(1)$. Therefore summing over the pairs $(\mu_1, \tilde{\mu}_1), (\mu_2, \tilde{\mu}_2) \in \Xi_j$ such that $|\mu_1 - \mu_2| \leq C2^j$ satisfies the bound

$$\begin{aligned} & \sum_{|\mu_1 - \mu_2| \leq C2^j} \langle \Upsilon_{j, \mu_1, \tilde{\mu}_1} (Q_{j, \mu_1} h, Q_{j, \tilde{\mu}_1} h), \Upsilon_{j, \mu_2, \tilde{\mu}_2} (Q_{j, \mu_2} h, Q_{j, \tilde{\mu}_2} h) \rangle_{L^2} \\ & \lesssim \sum_{(\mu, \tilde{\mu}) \in \Xi_j} \|\Upsilon_{j, \mu, \tilde{\mu}} (Q_{j, \mu_1} h, Q_{j, \tilde{\mu}_1} h)\|_{L^2}^2. \end{aligned}$$

It thus suffices to show that there exists C such that the sum over the off-diagonal pairs satisfies

$$\begin{aligned} & \sum_{|\mu_1 - \mu_2| > C2^j} \langle \Upsilon_{j, \mu_1, \tilde{\mu}_1} (Q_{j, \mu_1} h, Q_{j, \tilde{\mu}_1} h), \Upsilon_{j, \mu_2, \tilde{\mu}_2} (Q_{j, \mu_2} h, Q_{j, \tilde{\mu}_2} h) \rangle_{L^2} \\ & \lesssim \lambda^{-2N} \|h\|_{L^2}^2. \end{aligned}$$

Recalling the form of the kernel of $\Upsilon_{j, \mu, \tilde{\mu}}$ in (4.17), the main idea is that if $(\mu_1, \tilde{\mu}_1), (\mu_2, \tilde{\mu}_2) \in \Xi_j$ and $2^{-j}|\mu_1 - \mu_2|$ is sufficiently large then for any $N \in \mathbb{N}$ (possibly larger than that in the previous display),

$$\begin{aligned} & \left| \int e^{i\lambda(\Theta_g(x, z) + \Theta_g(x, \tilde{z}) - \Theta_g(x, w) - \Theta_g(x, \tilde{w}))} V_{j, \mu_1, \tilde{\mu}_1, \mu_2, \tilde{\mu}_2}(x, z, \tilde{z}, w, \tilde{w}) dx \right| \\ & \lesssim_N \lambda^{-N}, \\ & V_{j, \mu_1, \tilde{\mu}_1, \mu_2, \tilde{\mu}_2}(x, z, \tilde{z}, w, \tilde{w}) \\ & := V_{j, \mu_1}(x, z) V_{j, \tilde{\mu}_1}(x, \tilde{z}) V_{j, \mu_2}(x, w) V_{j, \tilde{\mu}_2}(x, \tilde{w}). \end{aligned}$$

Given the regularity estimate (4.14) and $2^j \gtrsim \lambda^{-1/8}$, this in turn follows from integration by parts and the bound

$$\left| d_x(\Theta_g(x, z) + \Theta_g(x, \tilde{z}) - \Theta_g(x, w) - \Theta_g(x, \tilde{w})) \right| \gtrsim 2^j.$$

Indeed, each integration by parts will yield a gain of $\lambda^{-1}2^j$ which is at least $\lambda^{-3/4}$. Given (4.18), this in turn will follow from

$$\left| d_x(\Theta_g(x, z) - \Theta_g(x, w)) \right| \gtrsim 2^j,$$

again assuming $2^{-j}|\mu_1 - \mu_2|$ is sufficiently large. Since $d_x\Theta_g(x, z)$, $d_x\Theta_g(x, w)$ give the covectors along the geodesics joining z, w to x , this follows from the same principle as in (4.18): if $V_{j, \mu_1}(x, z) V_{j, \mu_2}(x, w) \neq 0$, the two geodesics through x passing through z, w respectively intersect the

$y_1 = 0$ hyperplane with covectors pointing in the direction μ_1, μ_2 respectively. Since these two vectors are separated by a distance $\gtrsim C2^j$, this is enough.

The bound (4.10) is shown similarly, the only difference is that we did not multiply the kernel $\tilde{\sigma}_\lambda$ by a localizing factor⁶ akin to the $\tilde{q}_{\mu,j}$ (cf. (4.5)). However, if we consider the composition $\tilde{\sigma}_\lambda \circ Q_\nu$ as in Lemma 4.3 the proof is nearly identical to the $2^j \approx \lambda^{-1/8}$ case here.

5 Weaker geometric conditions

We conclude this work with a discussion of the prospects for proving Theorem 1.1 under weaker hypotheses on the sectional curvatures of (M, g) . Assuming that (M, g) has no conjugate points, the second author showed in [28] that if one had a $o(\lambda^{\delta(p,d)})$ gain in the L^p bounds when $2 < p < p_c$, then this would imply a $o(\lambda^{\frac{1}{p_c}})$ gain in the L^{p_c} bounds. Here we show that there are intermediate hypotheses, stronger than assuming no conjugate points, but weaker than nonpositive curvature, that yield a bound with a logarithmic gain of the form (1.8) with a possibly smaller value of ε_0 .

There are only two places in the argument above where the nonpositive curvature hypothesis was used in the arguments above over the implicit no conjugate point hypothesis: in (2.10) and in the proof of Theorem 2.1 when bounding the expression of $V(\tilde{x}, \tilde{y})$ in (2.21), (2.22). In the latter case, the observations in [32] were recalled, showing that the leading coefficient in the Hadamard parametrix $\vartheta(\tilde{x}, \tilde{y})$ is uniformly bounded when the curvatures are nonpositive, yielding (2.25). The other bounds (2.23) follow from lower bounds on the curvature and Jacobi field estimates, and here one can allow for the following algebraic growth in ϑ , which is also enough to show (2.10),

$$\vartheta(\tilde{x}, \tilde{y}) \lesssim \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})^{\frac{d-1}{2}}. \quad (5.1)$$

Recall that ϑ is characterized by $dV_g = \vartheta^{-2}(\tilde{x}, \tilde{y})d\mathcal{L}(\tilde{y})$ in normal coordinates at \tilde{x} , with \mathcal{L} denoting Lebesgue measure. To motivate the proof of (5.1), let $\tilde{\gamma}(t) : \mathbb{R} \rightarrow \tilde{M}$ be the unit speed geodesic joining \tilde{x} to \tilde{y} with $\tilde{\gamma}(0) = \tilde{x}$. Let $E_1, \dots, E_{d-1} \in T_{\tilde{x}}\tilde{M}$ be an orthonormal basis for the orthogonal complement of $\dot{\tilde{\gamma}}(0)$. Then let $Y_j(t)$ be the normal Jacobi field along $\tilde{\gamma}(t)$ with initial condition $Y_j(0) = 0$ and covariant derivative $D_t Y_j(0) = E_j$. Also let $Z_j(t)$ denote parallel translation of E_j along $\tilde{\gamma}(t)$. The fields Y_1, \dots, Y_{d-1} thus determine a fundamental matrix for solutions $Y(t)$ to the Jacobi equation along $\tilde{\gamma}(t)$ with $Y(0) = 0$ and $\dot{Y}(0) \perp Y(t)$; we denote this linear transformation as $\mathcal{Y}(t)$. Taking polar coordinates it can be seen that

⁶ This meant that the application of the linear theory in (4.11) was straightforward as the bounds on the amplitude defining σ_λ and its derivatives are uniform in λ .

$$\vartheta^{-2}(\tilde{x}, \tilde{y}) = t^{1-d} |\det(Y_1(t), \dots, Y_{d-1}(t))|, \quad t = \Theta_{\tilde{g}}(\tilde{x}, \tilde{y}), \quad (5.2)$$

though the right hand side of this identity is more accurately taken to be the determinant of the change of basis matrix from Y_1, \dots, Y_{d-1} to Z_1, \dots, Z_{d-1} .

The aforementioned observations in [32] (Günther comparison theorem), use that if M has nonpositive sectional curvatures, comparison theorems for Jacobi fields imply that $|Y_j(t)|_g \geq t$, and hence after other considerations, ϑ is in fact uniformly bounded. Otherwise, when (M, g) merely lacks conjugate points, the proof of (5.1) uses that the determinant in (5.2) is uniformly bounded from below. When $d = 2$, it was observed in [1] that such lower bounds follow from results of Green [15]. When $d \geq 3$, these were formalized in [3], though it seems that the crucial bounds were known to Freire and Mañé [14, Lemma I.3].

However, (5.1) by itself is not enough to imply the kernel estimates (2.17) in Theorem 2.1. Indeed, any reasonable substitute for the summation argument (2.27) would require that

$$\vartheta(\tilde{x}, \tilde{y}) \lesssim \Theta_{\tilde{g}}(\tilde{x}, \tilde{y})^{\frac{d-1}{2}(1-\delta_0)} \text{ for some } 0 < \delta_0 < 1, \quad (5.3)$$

at which point one would have (with $T = c_0 \log \lambda$ as throughout)

$$\sum_{\alpha \in \Gamma_{T_R}} |U_\alpha(\tilde{w}, \tilde{z})| \lesssim \begin{cases} \lambda^{\frac{d-1}{2} T - \min(1, \frac{(d-1)\delta_0}{2})}, & \delta_0 \neq \frac{2}{d-1}, \\ \lambda^{\frac{d-1}{2}} (\log T) T^{-1}, & \delta_0 = \frac{2}{d-1}. \end{cases} \quad (5.4)$$

The same considerations in Sect. 2 would then hold, though the definition of $c(\lambda)$ in (2.17) would have to be adjusted to be consistent with right hand side here. This in turn leads to adjustments in the exponents of $\log \lambda$ throughout Sects. 2 and 3 which are not difficult to compute. The remaining considerations in Remark 3.3 would then determine the exponent ε_0 .

We have outlined the proof of:

Theorem 5.1 *Suppose M has no conjugate points and that (5.3) is satisfied. In particular, this holds if along any unit speed geodesic $\tilde{\gamma} : \mathbb{R} \rightarrow \tilde{M}$, the linear transformation \mathcal{Y} determined by the Y_1, \dots, Y_{d-1} above satisfies*

$$|\mathcal{Y}(t)X|_{g(\tilde{\gamma}(t))} \gtrsim t^{\delta_0} |X|_{g(\tilde{\gamma}(0))}, \quad t \gtrsim 1, \quad X \in T_{\tilde{\gamma}(0)}\tilde{M}, \quad X \perp \dot{\tilde{\gamma}}(0). \quad (5.5)$$

for some implicit constant depending only on (\tilde{M}, \tilde{g}) . Then there exists $\varepsilon_0 > 0$, possibly different from that in Remark 1.2, such that (1.8) holds as in the conclusion in Theorem 1.1.

Remark 5.2 In general, the validity of (5.3) and (5.5) on an arbitrary manifold without conjugate points appears to be a long standing open problem. As

observed above, they are satisfied with $\delta_0 = 1$ when the sectional curvatures are nonpositive. An argument of Berger [2, §3], shows that (5.3) holds with $\delta_0 = 1$ if one has the intermediate hypothesis that M has no focal points.⁷ In particular, we could have stated our original theorem for Riemannian manifolds without focal points instead of those with nonpositive curvatures, with exactly the same exponent ε_0 . Results of Klingenberg [20] and Mañé [22] show that when the geodesic flow is Anosov, then in fact exponentially growing lower bounds in (5.5) are satisfied, yield the same exponent when $d \geq 4$ and even better ones for $d = 2, 3$ as in Remark 2.4. Finally, Eschenburg [13, Proposition 6] showed that (5.5) holds with $\delta_0 = \frac{1}{2}$ for manifolds with a so-called “ ρ -bounded asymptote” condition, which is stronger than assuming there are no conjugate points, but weaker than assuming there are no focal points.

Remark 5.3 The observations here carry over similarly to the authors’ work [10] where the key issue was to obtain a bound of the form (5.4) for a slightly different choice of Q_λ and use Lemma 2.3 similarly to achieve bounds analogous to (2.28). Consequently, if (5.3) is satisfied, we have for $\|f\|_{L^2(M)} = 1$ and $\mathcal{T}_{\lambda^{-1/2}}(\gamma)$ as the tubular neighborhood of diameter $\lambda^{-1/2}$ about a geodesic segment γ in M :

$$\int_{\mathcal{T}_{\lambda^{-1/2}}(\gamma)} |\mathbf{1}_{[\lambda, \lambda(\log \lambda)^{-1}]}(P)f|^2 dV_g \lesssim \begin{cases} (\log \lambda)^{-\min(\frac{d-1}{2}\delta_0, 1)}, & \delta_0 \neq \frac{2}{d-1}, \\ (\log \lambda)^{-1} \log \log \lambda, & \delta_0 = \frac{2}{d-1}. \end{cases} \quad (5.6)$$

In particular, the results in [10] hold for (M, g) without focal points. Given the results in [9], these considerations in turn yields a logarithmic gain in the known $L^p(M)$ bounds on spectral clusters when $2 < p < p_c$ (possibly with a larger exponent of $(\log \lambda)^{-1}$ than what would result from interpolating the main bound (1.8) with the trivial $p = 2$ bounds).

Acknowledgements The authors are grateful to the anonymous referee for numerous comments which improved the exposition in this work.

References

1. Bérard, P.H.: On the wave equation on a compact Riemannian manifold without conjugate points. *Math. Z.* **155**(3), 249–276 (1977)
2. Berger, M.: Some relations between volume, injectivity radius, and convexity radius in Riemannian manifolds. In: Differential Geometry and Relativity. Mathematical Physics and Applied Mathematics, vol. 3, pp. 33–42. Reidel, Dordrecht (1976)

⁷ While the results in Berger’s work are phrased in terms of the convexity radius, the argument fundamentally uses the no focal points hypothesis that for any nontrivial normal Jacobi field with $Y(0) = 0$ as above, $|Y(t)|_g$ is an increasing function of t .

3. Bonthonneau, Y.: A lower bound for the Θ function on manifolds without conjugate points, (2016) arXiv preprint [arXiv:1603.05697](https://arxiv.org/abs/1603.05697)
4. Bourgain, J.: Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.* **1**(2), 147–187 (1991)
5. Bourgain, J.: Refinements of Strichartz' inequality and applications to 2D-NLS with critical nonlinearity. *Int. Math. Res. Not.* **5**, 253–283 (1998)
6. Bak, J.-G., Seeger, A.: Extensions of the Stein–Tomas theorem. *Math. Res. Lett.* **18**(4), 767–781 (2011)
7. Blair, M.D., Sogge, C.D.: Refined and microlocal Kakeya–Nikodym bounds for eigenfunctions in two dimensions. *Anal. PDE* **8**(3), 747–764 (2014)
8. Blair, M.D., Sogge, C.D.: On Kakeya–Nikodym averages, L^p -norms and lower bounds for nodal sets of eigenfunctions in higher dimensions. *J. Eur. Math. Soc. (JEMS)* **17**(10), 2513–2543 (2015)
9. Blair, M.D., Sogge, C.D.: Refined and microlocal Kakeya–Nikodym bounds of eigenfunctions in higher dimensions. *Commun. Math. Phys.* **356**(2), 501–533 (2017)
10. Blair, M.D., Sogge, C.D.: Concerning Toponogov's theorem and logarithmic improvement of estimates of eigenfunctions. *J. Differ. Geom.* **109**(2), 189–221 (2018)
11. Béguin, P., Vargas, A.: Mass concentration phenomena for the L^2 -critical nonlinear Schrödinger equation. *Trans. Am. Math. Soc.* **359**(11), 5257–5282 (2007)
12. Chen, X., Sogge, C.D.: A few endpoint geodesic restriction estimates for eigenfunctions. *Commun. Math. Phys.* **329**(2), 435–459 (2014)
13. Eschenburg, J.-H.: Horospheres and the stable part of the geodesic flow. *Math. Z.* **153**(3), 237–251 (1977)
14. Freire, A., Mañé, R.: On the entropy of the geodesic flow in manifolds without conjugate points. *Invent. Math.* **69**(3), 375–392 (1982)
15. Green, L.W.: Surfaces without conjugate points. *Trans. Am. Math. Soc.* **76**, 529–546 (1954)
16. Hörmander, L.: The spectral function of an elliptic operator. *Acta Math.* **121**, 193–218 (1968)
17. Hörmander, L.: Oscillatory integrals and multipliers on FL^p . *Ark. Mat.* **11**, 1–11 (1973)
18. Hezari, H., Rivière, G.: L^p norms, nodal sets, and quantum ergodicity. *Adv. Math.* **290**, 938–966 (2016)
19. Hassell, A., Tacy, M.: Improvement of eigenfunction estimates on manifolds of nonpositive curvature. *Forum Mathematicum* **27**(3), 1435–1451 (2015)
20. Klingenberg, W.: Riemannian manifolds with geodesic flow of Anosov type. *Ann. Math.* (2) **99**, 1–13 (1974)
21. Lee, S.: Linear and bilinear estimates for oscillatory integral operators related to restriction to hypersurfaces. *J. Funct. Anal.* **241**(1), 56–98 (2006)
22. Mañé, R.: On a theorem of Klingenberg. In: *Dynamical Systems and Bifurcation Theory* (Rio de Janeiro, 1985), Pitman Research Notes in Mathematics Series, vol. 160, pp. 319–345. Longman Science and Technology, Harlow (1987)
23. Mockenhaupt, G., Seeger, A., Sogge, C.D.: Local smoothing of Fourier integral operators and Carleson–Sjölin estimates. *J. Am. Math. Soc.* **6**(1), 65–130 (1993)
24. Sogge, C.D.: Oscillatory integrals and spherical harmonics. *Duke Math. J.* **53**(1), 43–65 (1986)
25. Sogge, C.D.: Concerning the L^p norm of spectral clusters for second-order elliptic operators on compact manifolds. *J. Funct. Anal.* **77**(1), 123–138 (1988)
26. Sogge, C.D.: *Fourier Integrals in Classical Analysis*, Cambridge Tracts in Mathematics, vol. 105. Cambridge University Press, Cambridge (1993)
27. Sogge, C.D.: Kakeya–Nikodym averages and L^p -norms of eigenfunctions. *Tohoku Math. J. (2)* **63**(4), 519–538 (2011)
28. Sogge, C.D.: Localized L^p -estimates for eigenfunctions: II. arXiv preprint [arXiv:1610.06639](https://arxiv.org/abs/1610.06639) (2016)

29. Sogge, C.D.: Improved critical eigenfunction estimates on manifolds of nonpositive curvature. *Math. Res. Lett.* **24**(2), 549–570 (2017)
30. Sogge, C.D., Toth, J.A., Zelditch, S.: About the blowup of quasimodes on Riemannian manifolds. *J. Geom. Anal.* **21**(1), 150–173 (2011)
31. Sogge, C.D., Zelditch, S.: Riemannian manifolds with maximal eigenfunction growth. *Duke Math. J.* **114**(3), 387–437 (2002)
32. Sogge, C.D., Zelditch, S.: On eigenfunction restriction estimates and L^4 -bounds for compact surfaces with nonpositive curvature. In: *Advances in Analysis: The Legacy of Elias M. Stein*, Princeton Mathematical Series, vol. 50, pp. 447–461. Princeton University Press, Princeton (2014)
33. Sogge, C.D., Zelditch, S.: Focal points and sup-norms of eigenfunctions. *Rev. Mat. Iberoam.* **32**(3), 971–994 (2016)
34. Tacy, M.: A note on constructing families of sharp examples for L^p growth of eigenfunctions and quasimodes. *Proc. Am. Math. Soc.* **146**(7), 2909–2924 (2018)
35. Tao, T.: A sharp bilinear restriction estimate for paraboloids. *Geom. Funct. Anal.* **13**(6), 1359–1384 (2003)
36. Tao, T., Vargas, A., Vega, L.: A bilinear approach to the restriction and Kakeya conjectures. *J. Am. Math. Soc.* **11**(4), 967–1000 (1998)
37. Wolff, T.: A sharp bilinear cone restriction estimate. *Ann. Math.* (2) **153**(3), 661–698 (2001)
38. Zworski, M.: *Semiclassical Analysis*, Graduate Studies in Mathematics. American Mathematical Society, Providence (2012)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.