



# Refined and Microlocal Kakeya–Nikodym Bounds of Eigenfunctions in Higher Dimensions

Matthew D. Blair<sup>1</sup>, Christopher D. Sogge<sup>2</sup>

<sup>1</sup> Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131, USA.  
E-mail: blair@math.unm.edu

<sup>2</sup> Department of Mathematics, Johns Hopkins University, Baltimore, MD 21093, USA.  
E-mail: sogge@jhu.edu

Received: 15 December 2016 / Accepted: 8 June 2017

Published online: 4 August 2017 – © Springer-Verlag GmbH Germany 2017

**Abstract:** We prove a Kakeya–Nikodym bound on eigenfunctions and quasimodes, which sharpens a result of the authors (Blair and Sogge in Anal PDE 8:747–764, 2015) and extends it to higher dimensions. As in the prior work, the key intermediate step is to prove a microlocal version of these estimates, which involves a phase space decomposition of these modes that is essentially invariant under the bicharacteristic/geodesic flow. In a companion paper (Blair and Sogge in J Differ Geom, 2015), it will be seen that these sharpened estimates yield improved  $L^q(M)$  bounds on eigenfunctions in the presence of nonpositive curvature when  $2 < q < \frac{2(d+1)}{d-1}$ .

## 1. Introduction and Main Results

Let  $(M, g)$  be a smooth, compact, connected, boundaryless Riemannian manifold of dimension  $d \geq 2$ . Let  $\Delta_g$  be the nonpositive Laplace–Beltrami operator which is self adjoint with respect to the Riemannian measure  $dV_g$ . The compactness assumption ensures that the spectrum of  $\Delta_g$  is discrete, so that there exists an orthonormal basis for  $L^2(M)$  consisting of eigenfunctions  $e_\lambda$  satisfying

$$-\Delta_g e_\lambda = \lambda^2 e_\lambda, \quad \text{or equivalently,} \quad \sqrt{-\Delta_g} e_\lambda = \lambda e_\lambda. \quad (1.1)$$

Given such an eigenfunction with  $\lambda > 0$ , we set  $h = 1/\lambda$ , obtaining a solution to the semiclassical problem  $(h^2 \Delta_g + 1)e_{h^{-1}} = 0$ .

In this work, we are interested in  $L^q$  bounds on eigenfunctions and quasimodes associated to the semiclassical operator  $h^2 \Delta_g + 1$ . A family of functions  $\{\psi_h\}$  either defined for  $h$  in some subinterval  $h \in (0, h_0] \subset (0, 1]$  or a decreasing sequence in this interval tending to 0 will be considered to be an admissible family of quasimodes if

---

The first author was supported in part by the National Science Foundation Grant DMS-1301717, and the second by the National Science Foundation Grant DMS-1361476.

$$\|\psi_h\|_{L^2(M)} = 1, \\ \|(h^2 \Delta_g + 1)^k \psi_h\|_{L^2(M)} = O(h^k), \quad k = 1, \dots, \lceil \frac{d}{d+1} + \frac{d-1}{2} \rceil. \quad (1.2)$$

Eigenfunctions normalized in  $L^2$  define admissible quasimodes, as are normalized functions in the range of  $\mathbf{1}_{[1,1+h]}(h\sqrt{-\Delta_g})$ . In the classical notation (1.1), it is easily verified that  $L^2$ -normalized functions in the ranges of

$$\mathbf{1}_{[\lambda,\lambda+1/T]}(\sqrt{-\Delta_g}) \quad \text{and} \quad \rho(T(\lambda - \sqrt{-\Delta_g})) \quad (1.3)$$

define admissible quasimodes for  $T = T(\lambda) \geq 1$  a nondecreasing function, and in the latter case  $\rho \in \mathcal{S}(\mathbb{R})$ .

In this work, the interest is in upper bounds on the norm of these eigenfunctions and quasimodes in  $L^q(M)$  for  $2 < q < \frac{2(d+1)}{d-1}$ . In [13], the second author proved that  $\|\psi_h\|_{L^q(M)} = O(h^{-\delta(q)})$  where  $\delta(q) = \max(\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q}), \frac{d-1}{2} - \frac{d}{q})$  as  $h \rightarrow 0$  (or equivalently  $\lambda \rightarrow \infty$ ). These bounds are universal in that they apply to any  $(M, g)$  as above and any admissible quasimode as in (1.2). Moreover, if one considers the large classes of quasimodes given by functions in the range of  $\mathbf{1}_{[1,1+h]}(h\sqrt{-\Delta_g})$  (or in the classical notation,  $\mathbf{1}_{[\lambda,\lambda+1]}(\sqrt{-\Delta_g})$ ), it can be seen that this exponent cannot be improved. Indeed, when  $2 < q \leq \frac{2(d+1)}{d-1}$  one can find functions  $\psi_h$  in the range of these spectral projectors for which  $|\psi_h(x)|$  is roughly constant in a tubular neighborhood of radius  $\approx h^{1/2}$  about a geodesic segment when  $2 < q \leq \frac{2(d+1)}{d-1}$  and rapidly decreasing outside of this set so that  $\|\psi_h\|_{L^q(M)} \approx h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})}$ . When  $\frac{2(d+1)}{d-1} \leq q \leq \infty$ , there exist modes weakly decaying outside of a ball of radius  $h$  which similarly saturate these bounds.

On the round sphere  $\mathbb{S}^d$ , it is well known that the spectrum of  $\sqrt{-\Delta_{\mathbb{S}^d}}$  is of the form  $\sqrt{k(k+d-1)} = k + O(\frac{1}{k})$ ,  $k \in \mathbb{N} \cup \{0\}$ . Consequently, these spectral projectors in the previous paragraph are nearly the same as projections onto an eigenspace, and the  $L^q(M)$  bounds are saturated by families of exact eigenfunctions, in particular the highest weight spherical harmonics when  $2 < q \leq \frac{2(d+1)}{d-1}$  and the zonal harmonics when  $\frac{2(d+1)}{d-1} \leq q \leq \infty$ . However, the spectrum of  $\sqrt{-\Delta_{\mathbb{S}^d}}$  and the dynamics of the geodesic flow on  $\mathbb{S}^d$  are very unique and these universal bounds are not expected to be optimal for eigenfunctions in most other geometries (save those with elliptic closed geodesics or a full measure of periodic geodesics in the sense of [18]). A problem of great interest is to determine geometric conditions that show that these universal bounds can be improved for eigenfunctions or even quasimodes of shrinking width with  $T(\lambda) \nearrow \infty$  in (1.3).

This work develops a condition for obtaining improvements on the universal bounds when  $2 < q < \frac{2(d+1)}{d-1}$ . Let  $\Pi$  denote the space of unit length geodesic segments. Given  $\gamma \in \Pi$ , define  $\mathcal{T}_\varepsilon(\gamma)$  to be its  $\varepsilon$ -neighborhood

$$\mathcal{T}_\varepsilon(\gamma) = \{x \in M : d_g(x, \gamma) < \varepsilon\}. \quad (1.4)$$

For  $h > 0$  we then define the Kakeya–Nikodym norm of a quasimode  $\psi_h$  as

$$\|\psi_h\|_{KN}^2 = \sup_{\gamma \in \Pi} h^{-\frac{d-1}{2}} \int_{\mathcal{T}_{h^{1/2}}(\gamma)} |\psi_h|^2 dV_g. \quad (1.5)$$

It is also convenient to use the analogous norm without (approximate) averages

$$|||\psi_h|||_{KN}^2 = \sup_{\gamma \in \Pi} \int_{T_{h^{1/2}}(\gamma)} |\psi_h|^2 dV_g. \quad (1.6)$$

The norm thus depends on the frequency scale  $h$  of the quasimode under consideration. Note that if  $\psi_h$  is rapidly decreasing outside a tube  $T_{h^{1/2}}(\gamma)$ , as in the case of the highest weight spherical harmonics, then  $|||\psi_h|||_{KN} \approx 1$ . Otherwise, one expects better bounds to be satisfied.

The following is the primary result of this work.

**Theorem 1.1.** *Let  $\psi_h$  be a member of an admissible family of quasimodes as in (1.2). If  $d \geq 2$  and  $\frac{2(d+2)}{d} < q < \frac{2(d+1)}{d-1}$ , then*

$$\|\psi_h\|_{L^q(M)} \lesssim h^{-\frac{d-1}{4} + \frac{d-1}{2q}} |||\psi_h|||_{KN}^{\frac{2(d+1)}{q(d-1)} - 1} = h^{-\delta(q)} |||\psi_h|||_{KN}^{\frac{2(d+1)}{q(d-1)} - 1}. \quad (1.7)$$

Note that when  $d = 2$ , this means that for  $4 < q < 6$ ,

$$\|\psi_h\|_{L^q(M)} \lesssim h^{-\frac{1}{4} + \frac{1}{2q}} |||\psi_h|||_{KN}^{\frac{6}{q} - 1} = h^{-\delta(q)} |||\psi_h|||_{KN}^{\frac{6}{q} - 1}. \quad (1.8)$$

Moreover, when  $d = 2$  and  $q = 4$ , we also have that

$$\|\psi_h\|_{L^4(M)} \lesssim h^{-\frac{1}{8}} \sqrt{|\log h|} |||\psi_h|||_{KN}^{\frac{1}{2}} = h^{-\delta(4)} \sqrt{|\log h|} |||\psi_h|||_{KN}^{\frac{1}{2}}. \quad (1.9)$$

Here  $\delta(q) = \frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})$  as in the classical bounds of the second author [13]. The theorem thus furnishes a sufficient condition for improving on these classical bounds, as it means that improvements on the trivial  $|||\psi_h|||_{KN} = O(1)$  bounds translates to improvements on the  $L^q(M)$  bounds (see our discussion of [5] below).

Expressing our main theorem in this manner allows for a good discussion of its consequences, which we do below and in our companion paper [5]. However, in proving the theorem, it is convenient to instead work with the averaged norms (1.5) and show the following bounds, which are equivalent to (1.7), (1.9):

$$\|\psi_h\|_{L^q(M)} \lesssim h^{-\frac{d-1}{2} + \frac{d}{q}} \|\psi_h\|_{KN}^{\frac{2(d+1)}{q(d-1)} - 1}, \quad d \geq 2, \quad (1.10)$$

$$\|\psi_h\|_{L^4(M)} \lesssim \sqrt{|\log h|} \|\psi_h\|_{KN}^{\frac{1}{2}}, \quad d = 2. \quad (1.11)$$

In [4], the authors proved the bound (1.11), but with the factor  $\sqrt{|\log h|}$  on the right replaced by  $h^{-\varepsilon}$ , where  $\varepsilon > 0$  can be taken to be arbitrarily small. This is because that work actually proves lossless estimates involving averages of the mass over slightly thicker tubular neighborhoods  $h^{1/2-\varepsilon}$ . The present work thus sharpens this result and extends it to higher dimensions. Estimates of this type extend results of Bourgain [6] and the second author [16] in two dimensions, and the authors [3] in higher dimensions. This latter set of results show that the universal  $L^q$  bounds on eigenfunctions due to the second author [13] can be improved to  $o(h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})})$  when  $|||\psi_h|||_{KN} = o(1)$  and  $2 < q < \frac{2(d+1)}{d-1}$ .

Obtaining lossless estimates involving tubes of diameter  $h^{1/2}$  is significant for applications. In a companion paper [5], it will be shown that

$$|||\psi_h|||_{KN} = O(|\log h|^{-\sigma}), \quad \text{for some } \sigma = \sigma(d) > 0,$$

for quasimodes in the range of

$$\rho(T(\lambda - \sqrt{-\Delta_g})) = \rho(Th^{-1}(1 - h\sqrt{-\Delta_g})), \quad h = 1/\lambda, \quad (1.12)$$

as in (1.3) with  $T \approx \log \lambda = |\log h|$ . The small  $h^{1/2} = \lambda^{-1/2}$  thickness of the tubes turns out to be significant in the proof of such bounds. Combining these with Theorem 1.1 thus yields a logarithmic gain in the  $L^q$  bounds when  $2 < q < \frac{2(d+1)}{d-1}$  for quasimodes in the range of

$$\mathbf{1}_{[\lambda, \lambda + (\log \lambda)^{-1}]}(\sqrt{-\Delta_g}) = \mathbf{1}_{[1, 1+h|\log h|^{-1}]}(h\sqrt{-\Delta_g}). \quad (1.13)$$

Namely, for such quasimodes we have

$$\|\psi_h\|_{L^q(M)} \lesssim h^{-\frac{d-1}{4} + \frac{d-1}{2q}} |\log h|^{-\sigma(\frac{2(d+1)}{q(d-1)} - 1)}.$$

Moreover, the second author has shown [15] that averaging over these smaller tubes allows one to see that upper bounds on the  $L^1(M)$  norm allows one to detect scarring phenomena for eigenfunctions and quasimodes. More precisely, it is observed that using Hölder's inequality,

$$\|\psi_h\|_{K_N}^{-\frac{2(d+1)-(d-1)q}{(d-1)(q-2)}} \lesssim h^{-\frac{d-1}{4}} \|\psi_h\|_{L^1(M)} \quad \text{for } \frac{2(d+2)}{d} < q < \frac{2(d+1)}{d-1}.$$

Thus if the right hand side is  $\approx 1$ , it can be seen that the mass of  $\psi_h$  must concentrate in an  $h^{1/2}$  tubular neighborhood about a geodesic segment. Indeed, if the mass of  $\psi_h$  were mostly concentrated in a  $\mathcal{T}_{h^{1/2}}(\gamma)$ , then it is not hard to see that  $\|\psi_h\|_{L^1(M)} \lesssim h^{\frac{d-1}{4}}$ . The main idea in [15] is to show these upper bounds on the  $L^1(M)$  norm imply such concentration must occur.

The present work does not address how to improve upon the universal  $L^q(M)$  bounds of the second author in the range  $\frac{2(d+1)}{d-1} < q \leq \infty$  or when  $q = \frac{2(d+1)}{d-1}$ . Improvements in the former range, based on the measure of closed loops in the cosphere spaces  $S_x^* M$ , appear in [17–19] and stronger improvements for nonpositive curvature were shown in [2, 8]. As alluded to above, any condition for improvements on the universal  $L^q$  bounds in these cases must address the existence of modes which decay weakly outside of balls of radius  $h$ , and the norms in (1.5) above are insufficient for this purpose. The results in [9] obtain a logarithmic gain in the universal estimates in the presence of negative curvature, for all  $2 < q \leq \infty$  via an estimate at  $q = \frac{2(d+1)}{d-1}$  and interpolation, but the methods rely on generalized Weyl laws, and hold only for a full density subsequence of eigenfunctions.

The discussion above shows that the exponent in the two dimensional bound (1.8) is perhaps not so surprising as it would be the outcome of interpolating between the  $L^4(M)$  bound (without the logarithmic factor) and the universal  $L^6(M)$  bound of the second author. The same idea holds for the higher dimensional bound (1.10): if this were to hold when  $q = \frac{2d}{d-1}$  (so that the exponent of  $h$  vanishes), then the estimates appearing here would be the result of interpolating between this bound and the classical bounds of the second author at  $q = \frac{2(d+1)}{d-1}$ . However, the validity of (1.10) when  $\frac{2d}{d-1} \leq q \leq \frac{2(d+2)}{d}$  and  $d \geq 3$  could be subtle. Indeed, when  $d \geq 3$ , the present work relies on bilinear estimates of Lee [12] arising from the bilinear approach to the Fourier restriction problem that originated in works of Tao et al. Vega [21], Wolff [23], and Tao [20]. Explicit examples show that such bilinear estimates fail to hold in the regions  $\frac{2d}{d-1} \leq q \leq \frac{2(d+2)}{d}$  (see [21]).

*Remark 1.2.* The present work uses the semiclassical formalism of Koch et al. [11]. It is thus likely that Theorem 1.1 can be generalized to semiclassical pseudodifferential operators  $P(x, hD)$  with symbols  $p(x, \xi)$  whose characteristic sets  $\{\xi \in T_x^*M : p(x, \xi) = 0\}$  have nonvanishing second fundamental form. This is provided sets analogous to the  $\mathcal{T}_{h^{1/2}}(\gamma)$  above can be defined in a manner that allows for a substitute for Lemma 3.2 below. While we do not examine this issue in detail, our methods do easily treat semiclassical operators of the form  $h^2 \Delta_g + V$ ,  $V \in C^\infty$  with  $0 < \inf_{x \in M} V(x)$ , by passing to the “Jacobi metric”  $\tilde{g}(x) := V(x)g(x)$  (see [1, p. 228]). In this case, if one has an admissible quasimode defined by replacing  $(h^2 \Delta_g + 1)\psi_h$  by  $(h^2 \Delta_g + V)\psi_h$  in (1.2), then  $\psi_h$  is a quasimode with respect to  $h^2 \Delta_{\tilde{g}} + 1$  in that

$$\|(h^2 \Delta_{\tilde{g}} + 1)^k \psi_h\|_{L^2(M)} = O(h^k), \text{ for } k = 1, \dots, \left\lceil \frac{d}{d+1} + \frac{d-1}{2} \right\rceil.$$

Indeed, it is verified that the difference  $V \cdot \Delta_{\tilde{g}} - \Delta_g$  is a first order differential operator on  $M$ , which implies that Theorem 1.1 applies in this setting, defining the sets  $\mathcal{T}_{h^{1/2}}(\gamma)$  with respect to  $\tilde{g}$ .

*Outline of the paper.* In Sect. 2, we review aspects of the semiclassical approach to  $L^q$  bounds for eigenfunctions introduced in [11]. We also prove an Egorov theorem for symbols in the critical “ $S_{1/2}(1)$ ” class that is needed in our work. Section 3 provides the main outline of the proof of Theorem 1.1, considering “microlocal” Kakeya–Nikodym norms and reducing matters to showing certain almost orthogonality bounds and bilinear estimates in  $L^{\frac{q}{2}}$ . The last two sections, Sects. 4 and 5, treat these almost orthogonality bounds and bilinear estimates respectively.

## 2. Semiclassical Analysis

**2.1. Preliminary reductions.** Since  $\|(h^2 \Delta_g)^k \psi_h\|_{L^2(M)} = O(1)$  for the  $k$  in (1.2), elliptic regularity shows that  $\|\psi_h\|_{H_h^{2k}(M)} = O(1)$  in the semiclassical Sobolev spaces. Hence commutator estimates show that for a smooth bump function  $\varphi$ ,

$$\|(h^2 \Delta_g + 1)^k (\varphi \psi_h)\|_{L^2(M)} = O(h^k) \text{ for } k = 1, \dots, \left\lceil \frac{d}{d+1} + \frac{d-1}{2} \right\rceil.$$

Hence by multiplying  $\psi_h$  by a bump function within a finite partition of unity, we may assume it is supported in a suitable coordinate chart given by a cube of sidelength  $2\varepsilon$  centered at the origin in  $\mathbb{R}^d$ , and in particular

$$\text{supp}(\psi_h) \subset \left[-\frac{\varepsilon}{2}, \frac{\varepsilon}{2}\right]^d, \quad (2.1)$$

for some  $\varepsilon > 0$  sufficiently small. Moreover, we may assume that  $g^{ij}(0) = \delta^{ij}$ .

Next, we let  $\chi$  be a suitable smooth cutoff to a neighborhood of the unit sphere such that for  $k$  as in (1.2)

$$\|(1 - \chi(hD))\psi_h\|_{L^2(M)} \lesssim \|(h^2 \Delta_g + 1)^k \psi_h\|_{L^2(M)} + h^k \|\psi_h\|_{L^2(M)} \lesssim h^k. \quad (2.2)$$

Indeed, for a suitable choice of  $\chi$ ,  $h^2\Delta_g + 1$  is elliptic on  $\text{supp}(1 - \chi)$ , so elliptic regularity bounds yield the first inequality and quasimode properties the second. Taking  $k = \lceil \frac{d}{d+1} + \frac{d-1}{2} \rceil$ , semiclassical Sobolev embedding gives for  $2 < q < \frac{2(d+1)}{d-1}$

$$\begin{aligned} \|(1 - \chi(hD))\psi_h\|_{L^q(M)} &\lesssim h^{-\left(\frac{d}{2} - \frac{d}{q}\right)} \sum_{|\alpha| \leq 1} \|(hD)^\alpha (1 - \chi(hD))\psi_h\|_{L^2(M)} \\ &\lesssim h^{-\frac{d}{d+1}} \left( \|(h^2\Delta_g + 1)^k (1 - \chi(hD))\psi_h\|_{L^2(M)} + \|(1 - \chi(hD))\psi_h\|_{L^2(M)} \right) \\ &\lesssim h^{k - \frac{d}{d+1}} \leq h^{\frac{d-1}{2}}, \end{aligned}$$

where the second inequality follows from elliptic regularity (as in [11, Lemma 2.6]) and the last follows from the observation (2.2) above (for a slightly different choice of  $\chi$ ) as the symbol of  $[(h^2\Delta_g + 1)^k, (1 - \chi(hD))]$  is supported away from the unit sphere, at least up to  $O(h^k)$  error. Since we may cover  $M$  with  $O(h^{-\frac{d-1}{2}})$  tubular neighborhoods  $T_{h^{1/2}}(\gamma)$ ,  $\|(1 - \chi(hD))\psi_h\|_{L^q(M)}$  is in turn bounded by the right hand side of the inequalities in Theorem 1.1.

Taking a further partition of unity in the frequency variable, we may alter  $\chi(\xi)$  so that it is supported in a conic set of the form

$$\{\xi : -\xi_1 \geq C|\xi'|, |\xi| \approx 1\}, \quad \xi' = (\xi_2, \dots, \xi_n), \quad (2.3)$$

for some  $C > 0$  sufficiently large. The main idea in [11] is that after this further microlocalization, we may write the characteristic surface  $\sum_{ij} g^{ij}(z)\xi_i \xi_j = 1$  as a graph in the  $\xi'$  variables, which leads us to a evolution equation in the first variable. Indeed, over (2.3), the principal symbol of  $h^2\Delta_g + 1$  can be written in the form

$$-\sum_{i,j=1}^d g^{ij}(z)\xi_i \xi_j + 1 = -a(z, \xi)(\xi_1 + p(z, \xi')), \quad (z, \xi) \in \text{supp}(\psi_h) \times \text{supp}(\chi), \quad (2.4)$$

with  $a(z, \xi), p(z, \xi') > 0$ , by taking  $\varepsilon > 0$  sufficiently small in (2.1) and the aperture of the cone sufficiently small in (2.3). Moreover taking  $\varepsilon$  small above, we may assume that the restricted matrix  $\{g^{ij}(z)\}_{i,j=2}^d$  is positive definite when  $z \in [-\varepsilon, \varepsilon]^d$ .

We now have that

$$\begin{aligned} (hD_{z_1} + P(z, hD_{z'}))(\chi(hD)\psi_h) &= hf_h, \\ f_h := \frac{1}{h}b(z, hD)[h^2\Delta_g, \chi(hD)]\psi_h, \end{aligned}$$

for some symbol  $b$  compactly supported in all variables. In particular, we may assume that  $\text{supp}(b(\cdot, \xi)) \subset [-\varepsilon, \varepsilon]^d$  for every  $\xi$ . Moreover, we have that

$$\|f_h\|_{L^2(\mathbb{R}^d)} \lesssim \|\psi_h\|_{L^2(\mathbb{R}^d)} \quad (2.5)$$

and up to an error term which is  $O(h^\infty)$  as an operator on  $L^2$ , the composition

$$b(z, hD)[h^2\Delta_g, \chi(hD)] \quad (2.6)$$

is an operator with a compactly supported symbol.

**Notation.** Since the operator with symbol  $(\zeta_1 + p(z, \zeta'))$  is naturally an evolution equation, it is now convenient to take a different convention on the notation. We will use  $r, s, t$  to denote variables in  $\mathbb{R}$  and  $w, x, y, \xi, \eta$  to denote those in  $\mathbb{R}^{d-1}$  so that expressions such as  $(t, x)$  give variables in  $\mathbb{R}^d$ . The variable  $z$  will still denote those in  $\mathbb{R}^d$  on the few occasions this is needed. The relation for  $\psi_h$  thus reads

$$(hD_t + P(t, x, hD))\psi_h = hf_h$$

where it is understood that  $P(t, x, hD)$  (or  $P(t)$  when the dependence on  $x, h$  is not crucial) is a family of semiclassical pseudodifferential operators acting on the  $x$  variables. Keeping in line with these conventions, the symbols we consider in the remainder of this work will be quantized in the “ $x$ ” variables over  $\mathbb{R}^{d-1}$  and never in the first variable. Note that given the restriction of  $\chi$  above to a cone of sufficiently small aperture, we may also assume that for some smooth bump function  $\tilde{\chi}$  with  $\text{supp}(\tilde{\chi}) \subset \{|\xi| \ll 1\}$  (so that the quantization is only in the  $x$  variables)

$$\|(1 - \tilde{\chi})(hD)\psi_h(t, \cdot)\|_{L^2}, \|(1 - \tilde{\chi})(hD)f_h(t, \cdot)\|_{L^2} = O(h^\infty), \quad |t| \leq \varepsilon. \quad (2.7)$$

Let  $\tilde{\psi}_h := \chi(hD)\psi_h$  and let  $S(t, s)$  be the family of unitary operators satisfying

$$(hD_t + P(t, x, hD))S(t, s) = 0, \quad S(t, s)|_{t=s} = I.$$

Hence in vector valued notation, for  $f_h \in L^2([-\varepsilon, \varepsilon]_s; L^2(\mathbb{R}_x^{d-1}))$  and  $t \in [-\varepsilon, \varepsilon]$ ,

$$\tilde{\psi}_h(t) = \int_{-\varepsilon}^t S(t, s) f_h(s) ds. \quad (2.8)$$

Let  $\kappa_{t,s}(x, \xi)$  denote the time  $t$  value of the integral curve determined by  $H_{p_r}$ , the Hamiltonian vector field of  $p_r(x, \xi) = p(r, x, \xi)$  whose value at time  $s$  is  $(x, \xi)$ . Taking  $\varepsilon > 0$  sufficiently small in (2.1), we may assume that  $\kappa_{t,s}(x, \xi)$  defines a canonical transformation for  $t, s \in [-\varepsilon, \varepsilon]$ . We denote the components of this map in  $\mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$  as  $\kappa_{t,s}(x, \xi) = (x_{t,s}(x, \xi), \xi_{t,s}(x, \xi))$ . Moreover, standard construction (see e.g. [24, Chap. 10]) shows that there exists a phase function  $\phi(t, s, x, \eta)$  and a smooth, compactly supported amplitude  $a(t, s, x, \eta)$  such that

$$(S(t, s)f)(x) = \frac{1}{(2\pi h)^{d-1}} \iint e^{\frac{i}{h}(\phi(t, s, x, \eta) - \langle y, \eta \rangle)} a(t, s, x, \eta) f(y) dy d\eta + Ef, \quad (2.9)$$

for  $\text{supp}(f) \subset [-\varepsilon, \varepsilon]^{d-1}$ . Here  $E = E(t, s)$  satisfies  $\|E(t, s)\|_{L^2 \rightarrow L^2} \lesssim_N h^N$  for any  $N > 0$  (i.e. it is “ $O(h^\infty)$ ”) and given Sobolev embedding, it has a negligible contribution to the estimates in the present work. Hence we will often make a slight abuse of notation, treating  $S(t, s)$  as the same as this oscillatory integral operator.

We further recall that the phase function has the properties

$$\begin{aligned} \partial_t \phi(t, s, x, \eta) + p(t, x, d_x \phi(t, s, x, \eta)) &= 0, \\ \kappa_{t,s}(y, \eta) = (x, \xi) \quad \text{if and only if} \quad \xi &= d_x \phi(t, s, x, \eta), \quad y = d_\eta \phi(t, s, x, \eta). \end{aligned} \quad (2.10)$$

The operators  $S(t, s)$  are thus semiclassical Fourier integral operators associated to the canonical transformations  $\kappa_{t,s}(x, \xi)$ .

It is implicit in the work of Koch et al. [11] that for  $s$  fixed, the phase functions  $\phi(t, s, x, \xi)$  are of Carleson–Sjölin type as defined in [14, Chap. 2]. Indeed, for  $t, s \in [\varepsilon, \varepsilon]$ , we have that

$$d_x d_\xi \phi(t, s, x, \xi) = I + O(\varepsilon). \quad (2.11)$$

Moreover, it can be checked that

$$\{(-p(t, x, d_x \phi(t, s, x, \xi)), d_x \phi(t, s, x, \xi)) : \xi \in \text{supp}(a(t, s, x, \cdot))\} \quad (2.12)$$

is an embedded hypersurface in  $T_{(t,x)}^* \mathbb{R}^d$  with nonvanishing principal curvatures, all of the same sign. This follows as (2.11) implies that this is a local reparameterization of the graph  $\{(-p(t, x, \eta), \eta) : |\eta| \leq \frac{1}{2}\}$  and this is a subset of the strictly convex hypersurface determined by the zero set of the left hand side of (2.4).

Stein's theorem on Carleson–Sjölin phases (see e.g. [14, Theorem 2.2.1]) thus shows that

$$\left( \int_{-\varepsilon}^{\varepsilon} \|S(t, s)g\|_{L^q(\mathbb{R}^{d-1})}^q dt \right)^{\frac{1}{q}} \lesssim h^{\frac{d-1}{2}(\frac{1}{2} + \frac{1}{q}) - (d-1)} \|\mathcal{F}_h(g)\|_{L^2} \lesssim h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})} \|g\|_{L^2} \quad (2.13)$$

Indeed, Stein's theorem treats the case  $q = \frac{2(d+1)}{d-1}$ , so that the first inequality follows from interpolation with Hörmander's theorem [10] for nondegenerate phase functions in  $L^2(\mathbb{R}^{d-1})$  (cf. (2.11)). The second inequality follows from the semiclassical Plancherel identity. An application of Minkowski's inequality for integrals then shows that this yields the same linear bounds of the second author in [13]:

$$\|\tilde{\psi}_h\|_{L^q(\mathbb{R}^d)} \lesssim \left( \int_{-\varepsilon}^{\varepsilon} \left\| \int_{-\varepsilon}^t S(t, s) f_h(s) ds \right\|_{L^q(\mathbb{R}^{d-1})}^q dt \right)^{\frac{1}{q}} \lesssim h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})} \|\psi_h\|_{L^2(M)}.$$

In our case, linear bounds of this type will play a role, but ultimately the key is bilinear estimates, specifically those due to Hörmander [10] when  $d = 2$  and Lee [12] and the epsilon removal lemma of the authors [3] when  $d \geq 3$ . For now, we remark that the reductions above mean that it now suffices to prove estimates on

$$\|(\tilde{\psi}_h(t))^2\|_{L^{\frac{q}{2}}([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})} = \left\| \left( \int_{-\varepsilon}^t S(t, s) f_h(s) ds \right)^2 \right\|_{L^{\frac{q}{2}}([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})}, \quad (2.14)$$

showing they are bounded by the square of the right hand sides in (1.10), (1.11).

**2.2. An Egorov theorem for critical and near-critical symbols.** We conclude this section with an Egorov type theorem in the symbol class  $S_{1/2-\delta}(1)$ ,  $\delta \in [0, \frac{1}{2}]$  (see [24, p. 73] for the notation) which are symbols of the form (2.15) with  $\theta = h^{1/2-\delta}$  below. The  $\delta = 0$  case is considered critical as the usual symbolic calculus does not furnish terms of higher order in  $h$ . For example, the usual stationary phase expansions one often uses to prove the calculus do not yield these higher order terms. Consequently, one cannot apply the usual Egorov theorem to such symbols.

Here we show a version of the Egorov theorem for  $h$ -pseudodifferential operators (PDO) with symbols in these classes. As observed in [22, Chap. VIII, Sect. 8] and [7], there are Egorov theorems for classical PDO in these critical symbol classes, but here we are interested in a more qualitative version of the theorem. While we cannot show that conjugation by an  $h$ -Fourier integral operator (FIO) results in a strict propagation of support for the symbols via the associated canonical transformation, the symbol is rapidly decreasing on a scale of at least  $h^{-1/2}$  in terms of the distance to this propagated region, which is enough for our applications. The theorem not only applies to the operator  $S(t, s)$  defined above, but to more general local  $h$ -FIO.

**Theorem 2.1.** *Let  $S$  be a local  $h$ -Fourier integral operator*

$$(Sf)(x) = \frac{1}{(2\pi h)^n} \iint_{\mathbb{R}^{2n}} e^{\frac{i}{h}(\phi(x, \eta) - \langle y, \eta \rangle)} a(x, \eta) f(y) dy d\eta,$$

for some compactly supported symbol  $a \in S(1)$  over  $\mathbb{R}^{2n}$  and associated to a canonical transformation  $\kappa : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$  in that  $\kappa(y, \eta) = (x, \xi)$  if and only if  $\xi = d_x \phi(x, \eta)$  and  $y = d_\eta \phi(x, \eta)$ . Furthermore, let  $B$  be an  $h$ -pseudodifferential operator with symbol satisfying

$$|\partial^\alpha b(y, \xi)| \lesssim_\alpha \theta^{-|\alpha|} \quad \text{where } \theta = h^{\frac{1}{2} - \delta}, \delta \in [0, \frac{1}{2}]. \quad (2.15)$$

Then  $C = SBS^*$  defines an  $h$ -pseudodifferential operator with symbol  $c(x, \rho)$  also satisfying (2.15). If in addition,  $\text{supp}(b) \subset \mathcal{D}$  for some compact set  $\mathcal{D} \subset \mathbb{R}^n$ , then for every  $N \geq 0$ ,

$$|\partial^\alpha c(x, \rho)| \lesssim_{\alpha, N} \theta^{-|\alpha|} (1 + h^{-1} \theta d(x, \rho; \kappa(\mathcal{D})))^{-N}, \quad (2.16)$$

with  $d(x, \rho; \kappa(\mathcal{D}))$  denoting the distance to  $\kappa(\mathcal{D})$ , the image of  $\mathcal{D}$  under  $\kappa$ . Furthermore, for any  $M > 0$  there exists a symbol  $c_M(y, \eta)$  satisfying (2.15) such that  $\text{supp}(c_M) \subset \kappa(\mathcal{D})$  and for every  $N \geq 0$ ,

$$|\partial^\alpha (c - c_M)(y, \eta)| \lesssim_{\alpha, N} \theta^{-|\alpha|} (h \theta^{-2})^M (1 + h^{-1} \theta d(y, \eta; \kappa(\mathcal{D})))^{-N}. \quad (2.17)$$

Tracing through the proof, it is seen that the implicit constants appearing in (2.16) and (2.17) do not depend on  $\delta$ , but only on those in (2.15) and the derivative bounds for  $a$ ,  $\kappa$ , and  $\phi$ . In this work, we are mainly interested applying this theorem in the  $\delta = 0$  case, but record the more general case as a point of interest. We also stress that  $B$  is defined by taking the standard quantization of the symbol  $b(y, \xi)$ .

*Proof.* The composition  $C = SBS^*$  has a Schwartz kernel  $K(x, \tilde{x})$  given by an oscillatory integral. For sufficiently regular  $f$ , the compact support of the symbols ensures that we may write

$$(Cf)(x) = \frac{1}{(2\pi h)^n} \int e^{\frac{i}{h}\langle x, \rho \rangle} \left( \int e^{\frac{i}{h}\langle \tilde{x} - x, \rho \rangle} K(x, \tilde{x}) d\tilde{x} \right) \mathcal{F}_h(f)(\rho) d\rho.$$

The expression in parentheses determines the symbol  $c(x, \rho)$  and can be written as

$$\begin{aligned} c(x, \rho) &= \frac{1}{(2\pi h)^{3n}} \int_{\mathbb{R}^{6n}} e^{\frac{i}{h}(\langle y - \tilde{y}, \xi \rangle - \langle y, \eta \rangle + \langle \tilde{y}, \tilde{\eta} \rangle + \phi(x, \eta) - \phi(\tilde{x}, \tilde{\eta}) + \langle x - \tilde{x}, \rho \rangle)} \\ &\quad \times a(x, \eta) \overline{a(\tilde{x}, \tilde{\eta})} b(y, \xi) d(y, \tilde{y}, \eta, \tilde{\eta}, \xi, \tilde{x}) \end{aligned}$$

where the integrations here must be interpreted in the sense of distributions since the amplitude is independent of  $\tilde{y}$ . A standard limiting procedure shows that this presents no difficulty, allowing us to integrate in the  $\tilde{y}$  variable first, which yields the distribution  $(2\pi h)^n \delta(\tilde{\eta} - \xi)$ . Hence  $c(x, \rho)$  can be written as

$$\frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{4n}} e^{\frac{i}{h}(\phi(x, \eta) - \phi(\tilde{x}, \tilde{\eta}) + \langle \tilde{x} - x, \rho \rangle + \langle y, \tilde{\eta} - \eta \rangle)} a(x, \eta) \overline{a(\tilde{x}, \tilde{\eta})} b(y, \tilde{\eta}) d\eta d\tilde{\eta} dy d\tilde{x},$$

and the integration here is no longer formal. We rewrite this as

$$\begin{aligned} &\frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{4n}} e^{\frac{i}{h}(\phi(x, \eta) - \phi(\tilde{x} + x, \tilde{\eta} + \eta) + \langle \tilde{x}, \rho \rangle + \langle y, \tilde{\eta} \rangle)} \\ &\quad \times a(x, \eta) \overline{a(\tilde{x} + x, \tilde{\eta} + \eta)} b(y, \tilde{\eta} + \eta) d\eta d\tilde{\eta} dy d\tilde{x}. \end{aligned}$$

For each  $(x, \tilde{x}, \tilde{\eta}, \rho)$ , we make a further change of variables  $(y, \eta) \mapsto (w, \sigma)$  with

$$w = y - \int_0^1 d_\eta \phi(x + s\tilde{x}, \eta + s\tilde{\eta}) ds = y - d_\eta \phi(x + \tilde{x}, \eta + \tilde{\eta}) + O(|(\tilde{x}, \tilde{\eta})|) \quad (2.18)$$

$$\sigma = \rho - \int_0^1 d_x \phi(x + s\tilde{x}, \eta + s\tilde{\eta}) ds = \rho - d_x \phi(x, \eta) + O(|(\tilde{x}, \tilde{\eta})|) \quad (2.19)$$

so that the phase reads  $\langle(\tilde{x}, w), (\sigma, \tilde{\eta})\rangle$ . The invertibility of the mixed Hessian  $d_x d_\eta \phi$  ensures that locally,  $(y, \eta) \mapsto (w, \sigma)$  defines an invertible transformation. The last expression in (2.18) means that we may write  $y = d_\eta \phi(x + E_1, \eta + \tilde{\eta})$  where  $E_1(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) = O(|(w, \tilde{x}, \tilde{\eta})|)$  as  $(w, \tilde{x}, \tilde{\eta}) \rightarrow 0$ . Hence

$$b(y, \eta + \tilde{\eta}) = b(d_\eta \phi(x + E_1, \eta + \tilde{\eta}), \eta + \tilde{\eta}) = (b \circ \kappa^{-1})(x + E_1, d_x \phi(x + E_1, \eta + \tilde{\eta}))$$

where  $y, \eta$  are understood as functions of  $(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho)$ . Moreover, the last expression in (2.19) means that  $d_x \phi(x + E_1, \eta + \tilde{\eta}) = \rho + E_2$ , with  $E_2(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) = O(|(w, \tilde{x}, \tilde{\eta}, \sigma)|)$  near 0. Up to  $O(h^\infty)$  error,  $c(x, \rho)$  can thus be rewritten as

$$\frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{4n}} e^{\frac{i}{h} \langle(\tilde{x}, w), (\sigma, \tilde{\eta})\rangle} A(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) (b \circ \kappa^{-1})(x + E_1, \rho + E_2) d\eta d\tilde{\eta} dy d\tilde{x} \quad (2.20)$$

where  $A$  and its partial derivatives are uniformly bounded in  $h$ .

As observed in [24, p. 51], the phase function in (2.20) can be rewritten as

$$\langle(\tilde{x}, w), (\sigma, \tilde{\eta})\rangle = \frac{1}{2} \langle Q(\tilde{x}, w, \sigma, \tilde{\eta}), (\tilde{x}, w, \sigma, \tilde{\eta})\rangle, \text{ where } Q = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = Q^{-1}.$$

At this point, it can be treated using the method in [24, Theorem 3.13]. We momentarily treat  $x, \rho$  as fixed, and let  $u(\tilde{x}, w, \sigma, \tilde{\eta})$  be the amplitude in (2.20) formed by fixing these variables. Let  $\zeta \in \mathbb{R}^{4n}$  denote variables which are dual to  $(\tilde{x}, w, \sigma, \tilde{\eta})$ , then using the classical Fourier transform, (2.20) can be rewritten as

$$J(h, u) = \frac{1}{(2\pi)^{4n}} \int_{\mathbb{R}^{4n}} e^{-\frac{ih}{2} \langle Q\zeta, \zeta \rangle} \widehat{u}(\zeta) d\zeta.$$

Applying Taylor's formula in  $h$  we have for every  $M \in \mathbb{N}$ , the differential operator  $\tilde{P} = -i \langle D_{(\tilde{x}, w)}, D_{(\sigma, \tilde{\eta})} \rangle$  can be used to write

$$J(h, u) = \sum_{k=0}^{M-1} \frac{h^k}{k!} J(0, \tilde{P}^k u) + \frac{h^M}{(M-1)!} \int_0^t (1-t)^{M-1} J(th, \tilde{P}^M u) dt. \quad (2.21)$$

The sum in  $k$  can be rewritten as  $\sum_{k=0}^{M-1} (h\theta^{-2})^k \tilde{c}_k(x, \rho)$  with  $|\partial^\alpha \tilde{c}_k| \lesssim \theta^{-|\alpha|}$ , and  $\text{supp}(c_k) \subset \text{supp}(b \circ \kappa^{-1}) = \kappa(\text{supp}(b))$ .

A change of variables  $\zeta \mapsto t^{-\frac{1}{2}}\zeta$  in the remainder  $J(th, \tilde{P}^M u)$  yields

$$J(th, \tilde{P}^M u) = \frac{1}{(2\pi \sqrt{t})^{4n}} \int_{\mathbb{R}^{4n}} e^{-\frac{ih}{2} \langle Q\zeta, \zeta \rangle} \widehat{\tilde{P}^M u}(t^{-\frac{1}{2}}\zeta) d\zeta \quad (2.22)$$

We are thus led to study integrals of the following form, for  $t \in (0, 1]$

$$\tilde{c}_t(x, \rho) = \frac{1}{(2\pi h)^{2n}} \int_{\mathbb{R}^{4n}} e^{\frac{i}{h} \langle(\tilde{x}, w), (\sigma, \tilde{\eta})\rangle} \widetilde{A}(x, t^{\frac{1}{2}}\tilde{x}, t^{\frac{1}{2}}w, t^{\frac{1}{2}}\sigma, t^{\frac{1}{2}}\tilde{\eta}, \rho) d\tilde{x} dw d\sigma d\tilde{\eta}. \quad (2.23)$$

for some amplitude  $\tilde{A}(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho)$  satisfying  $|\partial^\alpha \tilde{A}| \lesssim_\alpha \theta^{-|\alpha|}$  and

$$\begin{aligned} \{(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) : \tilde{A}(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) \neq 0\} \\ \subset \{(x, \tilde{x}, w, \sigma, \tilde{\eta}, \rho) : (b \circ \kappa^{-1})(x + E_1, \rho + E_2) \neq 0\}. \end{aligned}$$

Indeed, the identity in (2.22), combined with the Fourier transform, shows that the remainder term in (2.21) can be expressed as

$$\frac{(h\theta^{-2})^M}{(M-1)!} \int_0^1 \tilde{c}_t(x, \rho) (1-t)^{M-1} dt$$

for some integral  $\tilde{c}_t(x, \rho)$ . Moreover, (2.20) is of the form (2.23) with  $t = 1$ , so studying these integrals provides a unified approach to (2.16) and (2.17).

We now show that  $(\theta D_{x_i})^\alpha \tilde{c}_t(x, \rho)$  has the effect of replacing the amplitude in the definition by one of the same regularity. By induction, it suffices to prove this for  $|\alpha| = 1$ . This is clear for any such weighted derivative applied to the amplitude, hence it suffices to consider the effect of such derivatives on the phase. Recall that  $\tilde{x}, \tilde{\eta}$  are independent of  $(x, \rho)$  even though  $(w, \sigma)$  are not. Hence

$$\theta D_{x_i} e^{\frac{i}{h} \langle (\tilde{x}, w), (\sigma, \tilde{\eta}) \rangle} = (\partial_{x_i} \sigma \cdot (\theta D_\sigma) + \partial_{x_i} w \cdot (\theta D_w)) e^{\frac{i}{h} \langle (\tilde{x}, w), (\sigma, \tilde{\eta}) \rangle},$$

and integration by parts completes the proof of the claim. Derivatives in  $\rho$  are handled analogously.

To see (2.16), (2.17) with  $N = 0$ , it remains to show that  $|\tilde{c}_t(x, \rho)|$  is uniformly bounded in  $h$ . Note that the first order differential operator

$$\frac{1 + (\sigma, \tilde{\eta}, \tilde{x}, w) \cdot D_{(\tilde{x}, w, \sigma, \tilde{\eta})}}{1 + h^{-1} |(\sigma, \tilde{\eta}, \tilde{x}, w)|^2} = \frac{1 + (h^{-\frac{1}{2}} (\sigma, \tilde{\eta}, \tilde{x}, w)) \cdot (h^{\frac{1}{2}} D_{(\tilde{x}, w, \sigma, \tilde{\eta})})}{1 + h^{-1} |(\sigma, \tilde{\eta}, \tilde{x}, w)|^2} \quad (2.24)$$

preserves the exponential factor in (2.23). Writing the differential operator in the second fashion illustrates that the transpose of this operator applied to the class of amplitudes here yields an amplitude of the same regularity (in that there is no loss in  $h$ ). Integration by parts sufficiently many times yields an amplitude which is  $O((1 + h^{-1} |(\sigma, \tilde{\eta}, \tilde{x}, w)|^2)^{-4n})$ , so that integration in these variables counterbalances the loss of  $h^{-2n}$  in front of (2.23). Alternatively, one could verify (2.16), (2.17) when  $N = 0$  by changing variables  $(\tilde{x}, w, \sigma, \tilde{\eta}) \mapsto h^{1/2}(\tilde{x}, w, \sigma, \tilde{\eta})$  from the outset in (2.23).

To show (2.16), (2.17) for  $N \geq 1$ , we observe that if  $d(x, \rho; \kappa(\mathcal{D})) \geq h\theta^{-1}$ , then

$$|c_t(x, \rho)| \lesssim_N \left( h^{-1} \theta d(x, \rho; \kappa(\mathcal{D})) \right)^{-N}. \quad (2.25)$$

Indeed, since  $(b \circ \kappa^{-1})(x + E_1, \rho + E_2) \neq 0$  over the domain of integration in (2.23), our previous observation that  $E_1, E_2 = O(|(\tilde{x}, w, \sigma, \tilde{\eta})|)$  means that for such  $(x, \rho)$ ,

$$h\theta^{-1} \leq d(x, \rho; \kappa(\mathcal{D})) \leq |(E_1, E_2)| \lesssim |t^{\frac{1}{2}}(\tilde{x}, w, \sigma, \tilde{\eta})| \lesssim |(\tilde{x}, w, \sigma, \tilde{\eta})|,$$

Hence the phase function in (2.23) has no critical points, and each integration by parts with respect to  $D_{(\tilde{x}, w, \sigma, \tilde{\eta})}$  yields a gain of at least  $O(h\theta^{-1} |(\tilde{x}, w, \sigma, \tilde{\eta})|^{-1})$ . Integrating by parts sufficiently many times with this operator followed by a successive integration by parts using the operator in (2.24) thus implies (2.25).  $\square$

*Remark 2.2.* It is interesting to note that if one does not assume  $\text{supp}(b) \subset \mathcal{D}$ , but instead the weaker decay condition

$$|\partial^\alpha b(x, \xi)| \lesssim_{\alpha, N} \theta^{-|\alpha|} (1 + \theta^{-1} d(x, \xi; \mathcal{D}))^{-N},$$

then the symbol of  $c$  satisfies the same bound, but with  $\mathcal{D}$  replaced by  $\kappa(\mathcal{D})$ . This is a matter of replacing the differential operator in (2.24) by

$$\frac{1 + (\theta^{-1}(\sigma, \tilde{\eta}, \tilde{x}, w)) \cdot (h\theta^{-1} D_{(\tilde{x}, w, \sigma, \tilde{\eta})})}{1 + \theta^{-2} |(\tilde{x}, w, \sigma, \tilde{\eta})|^2}$$

and proceeding in a similar fashion to the  $t = 1$  case in the analysis of (2.23) above.

### 3. Microlocal Kakeya–Nikodym Norms and the Outline of the Proof

Let  $J$  be the negative integer satisfying  $2^J \geq \sqrt{h} > 2^{J-1}$ . Given  $v \in 2^J \mathbb{Z}^{2(d-1)}$ , let  $b_v^J$  be a smooth bump function such that for some  $\tilde{\varepsilon} > 0$  sufficiently small

$$\begin{aligned} \text{supp}(b_v^J) &\subset v + \left[ -\left(\frac{1}{2} + \tilde{\varepsilon}\right) 2^J, \left(\frac{1}{2} + \tilde{\varepsilon}\right) 2^J \right]^{2(d-1)}, \\ \sum_{v \in 2^J \mathbb{Z}^{2(d-1)}} b_v^J(x, \xi) &= 1 \quad \text{and} \quad \left| \partial_{y, \eta}^\alpha b_v^J(y, \eta) \right| \lesssim_\alpha 2^{-J|\alpha|} \approx h^{-\frac{|\alpha|}{2}}. \end{aligned} \quad (3.1)$$

We then define  $B_v^J(0)$  (to be extended to a family of operators  $\{B_v^J(t)\}_{t \in [-\varepsilon, \varepsilon]}$ ), to be the semiclassical  $h$ -pseudodifferential operator with symbol  $b_v^J \in S_{1/2}(1)$

$$(B_v^J(0)g)(x) = \frac{1}{(2\pi h)^{d-1}} \iint_{\mathbb{R}^{2(d-1)}} e^{\frac{i}{h} \langle x - y, \eta \rangle} b_v^J(x, \xi) g(y) dy d\eta. \quad (3.2)$$

We remark that by a standard rescaling argument and the Calderón–Vaillancourt theorem,  $B_v^J(0)$  is bounded on  $L^2(\mathbb{R}_x^{d-1})$  (as is any operator with symbol in  $S_{1/2}(1)$ ). Extend  $B_v^J(0)$  to a family of operators by defining

$$B_v^J(t) = S(t, 0) B_v^J(0) S(0, t).$$

Given  $g \in L^2([-\varepsilon, \varepsilon]_s; L^2(\mathbb{R}_x^{d-1}))$ , let  $T$  denote the operator defined in vector valued notation

$$(Tg)(t) = \int_{-\varepsilon}^t S(t, s) g(s) ds, \quad \text{that is, } (Tg)(t, x) = \int_{-\varepsilon}^t (S(t, s) g(s, \cdot))(x) ds,$$

so that  $g(s), (Tg)(t) \in L^2(\mathbb{R}_x^{d-1})$  for  $t, s \in [-\varepsilon, \varepsilon]$ . As suggested in (2.14), the main strategy for proving the estimates in (1.10), (1.11) will be to prove  $L^{q/2}$  bounds on  $(\tilde{\psi}_h)^2$ , leading us to consider  $(Tf_h)^2$ . Since  $f_h(s) = \sum_v B_v^J(s) f_h(s)$ , we may write

$$\begin{aligned} (Tf_h(t))^2 &= \sum_{v, v'} \left( \int_{-\varepsilon}^t S(t, s) B_v^J(s) f_h(s) ds \right) \left( \int_{-\varepsilon}^t S(t, r) B_{v'}^J(r) f_h(r) dr \right) \\ &= \sum_{v, v'} T(B_v^J f_h)(t) T(B_{v'}^J f_h)(t) \end{aligned} \quad (3.3)$$

where  $B_v^J f_h$  abbreviates the vector valued function  $s \mapsto B_v^J(s) f_h(s)$ .

We next consider the family of dyadic cubes in  $\mathbb{R}^{2(d-1)}$ ,  $\tau_\mu^j$  formed by taking the cube  $[-2^{j-1}, 2^{j-1}]^{2(d-1)}$ , and translating by an element  $\mu \in 2^j \mathbb{Z}^{2(d-1)}$ . As in [21], two dyadic cubes are said to be *close*, writing  $\tau_\mu^j \sim \tau_{\mu'}^j$ , if they are not adjacent, but have adjacent parents of sidelength  $2^{j+1}$ . We use this to organize the sum in (3.3): given a pair  $(v, v')$  such that  $2^{-J} |v - v'|$  is sufficiently large, there exist close dyadic cubes  $\tau_\mu^j \sim \tau_{\mu'}^j$  such that  $v \in \tau_\mu^j$  and  $v' \in \tau_{\mu'}^j$  (which means  $|v - v'| \approx 2^j$ ), allowing us to write

$$\begin{aligned} \sum_{v, v'} T(B_v^J f_h) T(B_{v'}^J f_h) &= \sum_{(v, v') \in \Xi_J} T(B_v^J f_h) T(B_{v'}^J f_h) \\ &\quad + \sum_{j=J+1}^0 \sum_{(v, v') \in \tau_\mu^j \times \tau_{\mu'}^j : \tau_\mu^j \sim \tau_{\mu'}^j} T(B_v^J f_h) T(B_{v'}^J f_h) + O(h^\infty), \end{aligned} \quad (3.4)$$

where we let  $\Xi_J$  denote pairs  $(v, v')$  lying in adjacent cubes. Note that by (2.7), we may limit ourselves to cubes of distance less than 1 at the cost of the negligible  $O(h^\infty)$  error term on the right. When  $0 \leq j \leq J+1$ , we define

$$\Xi_j := \{(\mu, \mu') : \tau_\mu^j \sim \tau_{\mu'}^j\},$$

noting the slight variation in these definitions. For  $\mu \in 2^j \mathbb{Z}^{2(d-1)}$ , we define  $B_\mu^j(0)$  as the  $h$ -PDO with symbol

$$b_\mu^j(y, \eta) := \sum_{v \in \tau_\mu^j} b_v^J(y, \eta).$$

For convenience, set

$$\mathcal{D}_\mu^j := \text{supp}(b_\mu^j). \quad (3.5)$$

Taking  $\tilde{\varepsilon}$  small in (3.1), we have for  $(\mu, \mu') \in \Xi_j$ , we have

$$d(\mathcal{D}_\mu^j, \mathcal{D}_{\mu'}^j) \approx 2^j, \quad 0 \geq j \geq J+1. \quad (3.6)$$

Also define

$$B_\mu^j(s) := \sum_{v \in \tau_\mu^j} B_v^J(s) = S(s, 0) B_\mu^j(0) S(0, s),$$

and denote the symbol of  $B_\mu^j(s)$  as  $b_\mu^j(s, y, \eta)$ . Given Theorem 2.1,

$$\left| \partial_{y, \eta}^\alpha b_\mu^j(s, y, \eta) \right| \lesssim_{\alpha, N} h^{-\frac{|\alpha|}{2}} \left( 1 + h^{-\frac{1}{2}} d \left( y, \eta; \kappa_{s, 0} \left( \mathcal{D}_\mu^j \right) \right) \right)^{-N}, \quad (3.7)$$

and the same holds when  $j = J$ .

With this notation, the first sum on the right in (3.4) can be rewritten as

$$\sum_{j=J+1}^0 \sum_{(\mu, \mu') \in \Xi_j} T(B_\mu^j f_h) T(B_{\mu'}^j f_h). \quad (3.8)$$

Hence  $\|(\tilde{\psi}_h)^2\|_{L^{\frac{q}{2}}([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})}$  is bounded above by

$$\sum_{j=J+1}^0 \left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_\mu^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} + \left\| \sum_{(v, v') \in \Xi_J} T(B_v^J f_h) T(B_{v'}^J f_h) \right\|_{L^{\frac{q}{2}}}. \quad (3.9)$$

To estimate (3.9), we follow the strategy in [4], first proving almost orthogonality of the  $T(B_\mu^j f_h) T(B_{\mu'}^j f_h)$  for  $(\mu, \mu') \in \Xi_j$ , then using bilinear estimates to bound each product. However, to make this rigorous, we will need to instead consider a family of  $h$ -PDO for  $|s| \leq \varepsilon$  where  $\varepsilon$  is as in (2.1)

$$\left\{ B_{\mu, \omega}^j(s) : \omega \in \Omega, s \in [-\varepsilon, \varepsilon] \right\} \quad (3.10)$$

where  $\Omega$  is a finite index set and each corresponding symbol  $b_{\mu, \omega}^j(s, y, \eta)$  in the family satisfies the same bound as in (3.7)

$$\left| \partial_{y, \eta}^\alpha b_{\mu, \omega}^j(s, y, \eta) \right| \lesssim_{\alpha, N} h^{-\frac{|\alpha|}{2}} \left( 1 + h^{-\frac{1}{2}} d(y, \eta; \kappa_{s, 0}(\mathcal{D}_\mu^j)) \right)^{-N}, \quad (3.11)$$

and for each  $\alpha$  and  $N$ , the implicit constants here are uniformly bounded in  $\mu, \omega, j$ . This collection will be defined below. A  $B_{\mu, \omega}^j(s)$  can thus be viewed as akin to a  $B_\mu^j(s)$ , but whose symbol as been distorted (though compact support of the symbol may be lost when  $s = 0$ ).

Before getting to the crux of our argument, we prove two lemmas.

**Lemma 3.1.** *Suppose  $\omega_\mu \in \Omega$  is any sequence parameterized by  $\mu \in 2^j \mathbb{Z}^{2(d-1)}$ . Then for any  $J \leq j \leq 0$ ,*

$$\sum_\mu \left\| B_{\mu, \omega_\mu}^j f_h \right\|_{L^2([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})}^2 \lesssim \|f_h\|_{L^2}^2. \quad (3.12)$$

*Proof.* Observe that it suffices to show that for each  $s \in [-\varepsilon, \varepsilon]$

$$\sum_\mu \left\| B_{\mu, \omega_\mu}^j(s) f_h(s) \right\|_{L^2(\mathbb{R}^{d-1})}^2 \lesssim \|f_h(s)\|_{L^2(\mathbb{R}^{d-1})}^2,$$

and integrate both sides over  $s \in [-\varepsilon, \varepsilon]$ . By Khintchine's inequality, this reduces to showing that for an arbitrary sequence  $\{\nu_\mu\}_\mu$  taking values  $\nu_\mu = \pm 1$

$$\left\| \sum_\mu \nu_\mu B_{\mu, \omega_\mu}^j(s) f_h(s) \right\|_{L^2(\mathbb{R}^{d-1})} \lesssim \|f_h(s)\|_{L^2(\mathbb{R}^{d-1})},$$

where the implicit constant on the right is independent of the sequence. By the aforementioned application of the Calderón-Vaillancourt theorem, this further reduces to showing that the symbol

$$\sum_\mu \nu_\mu b_{\mu, \omega_\mu}^j(s, y, \eta) \quad (3.13)$$

is in  $S_{1/2}(1)$  with bounds uniform in the sequence  $\{v_\mu\}_\mu$ . This in turn follows from verifying that for  $N \geq 2d$ ,

$$\sum_\mu \left(1 + h^{-\frac{1}{2}} d\left(y, \eta; \kappa_{s,0}(\mathcal{D}_\mu^j)\right)\right)^{-N} \lesssim \sum_\mu \left(1 + 2^{-j} d\left(\kappa_{0,s}(y, \eta); \mathcal{D}_\mu^j\right)\right)^{-N} \lesssim 1. \quad (3.14)$$

Indeed, if this holds, then any weighted derivative  $(h^{1/2} \partial_{y,\eta})^\alpha$  of the symbol in (3.13) is  $O(1)$  by (3.11) and the triangle inequality.

To see (3.14), note that the first inequality follows since  $\kappa_{s,0}$  and its inverse are Lipschitz maps and  $2^{-j} \leq h^{-1/2}$ . For the second inequality, observe that

$$d\left(\kappa_{0,s}(y, \eta); \mathcal{D}_\mu^j\right) \leq 4|\kappa_{0,s}(y, \eta) - \mu| \text{ if } d\left(\kappa_{0,s}(y, \eta); \mathcal{D}_\mu^j\right) \geq 2^{j+3}.$$

The rest of (3.14) then follows as a consequence of

$$\sum_{\mu \in 2^j \mathbb{Z}^{2(d-1)}} \left(1 + 2^{-j} |\kappa_{0,s}(y, \eta) - \mu|\right)^{-N} \lesssim 1.$$

□

We now define the microlocal Kakeya–Nikodym norm of  $f_h$  as

$$\|f_h\|_{MKN}^2 := \sup_{J \leq j \leq 0} \sup_{\omega \in \Omega} \sup_{\mu \in 2^j \mathbb{Z}^{2(d-1)}} 2^{-j(d-1)} \int_{-\varepsilon}^{\varepsilon} \|B_{\mu,\omega}^j(s) f_h(s)\|_{L^2(\mathbb{R}^{d-1})}^2 ds.$$

**Lemma 3.2.** *Using the norms in (1.5) involving averages over tubes, we have that*

$$\|f_h\|_{MKN} \lesssim \|\psi_h\|_{KN}. \quad (3.15)$$

*Proof.* The bound (3.15) will follow by showing that for any  $j, \mu, \omega$

$$2^{-j(d-1)} \int_{-\varepsilon}^{\varepsilon} \|B_{\mu,\omega}^j(s) f_h(s)\|_{L^2(\mathbb{R}^{d-1})}^2 ds \lesssim \sup_{\gamma \in \Pi} h^{-\frac{d-1}{2}} \int_{\mathcal{T}_{h^{1/2}}(\gamma)} |\psi_h(t, x)|^2 dx dt.$$

In what follows, let  $\pi : T^*\mathbb{R}^{d-1} \rightarrow \mathbb{R}^{d-1}$  denote the projection  $\pi(x, \xi) = x$ . Observe that the integral curves of the Hamiltonian vector field  $H_{p_t}$  are nonlinear reparameterizations of unit speed geodesics in  $T^*\mathbb{R}^d$ , which are integral curves of the Hamiltonian vector field determined by  $\frac{1}{2}(\sum_{i,j} g^{ij}(z) \zeta_i \zeta_j - 1)$ . So while

$$\mathcal{E}_\mu^j := \cup_{s \in [-\varepsilon, \varepsilon]} \pi(\kappa_{s,0}(\mathcal{D}_\mu^j))$$

does not define a tubular neighborhood, it is contained in a tubular neighborhood of width  $O(2^j)$  about the unit speed geodesic segment which intersects the  $s = 0$  plane with coordinates  $\mu \in T^*(\{s = 0\})$ .

We now claim that there exists a cover of our coordinate chart by family of (not necessarily tubular) neighborhoods about geodesic segments  $\{\tilde{\mathcal{T}}_\mathbf{k}\}_\mathbf{k}$  indexed by  $\mathbf{k} \in 2^J \mathbb{Z}^{d-1}$  ( $2^J \approx h^{1/2}$  as above) and an associated partition of unity  $\sum \chi_\mathbf{k}^2 = 1$  in the chart with  $\text{supp}(\chi_\mathbf{k}) \subset \tilde{\mathcal{T}}_\mathbf{k}$  with the following properties for some  $C$  uniform and  $\mathbf{k}_0$  fixed:

$$d(\tilde{\mathcal{T}}_\mathbf{k}, \mathcal{E}_\mu^j) \approx |\mathbf{k} - \mathbf{k}_0| \text{ whenever } d(\tilde{\mathcal{T}}_\mathbf{k}, \mathcal{E}_\mu^j) \geq C2^j, \quad (3.16)$$

$$\#\left\{\mathbf{k} : d\left(\tilde{\mathcal{T}}_{\mathbf{k}}, \mathcal{E}_{\mu}^j\right) < C2^j\right\} = O\left((h^{-1/2}2^j)^{d-1}\right), \quad (3.17)$$

each  $\tilde{\mathcal{T}}_{\mathbf{k}}$  can be covered by  $O(1)$  tubular neighborhoods  $\mathcal{T}_{h^{1/2}}(\gamma)$ ,  $(3.18)$

$$\tilde{\mathcal{T}}_{\mathbf{k}} \cap \tilde{\mathcal{T}}_{\tilde{\mathbf{k}}} = \emptyset \text{ and } d(\tilde{\mathcal{T}}_{\mathbf{k}}, \tilde{\mathcal{T}}_{\tilde{\mathbf{k}}}) \approx |\mathbf{k} - \tilde{\mathbf{k}}| \quad \text{if} \quad |\mathbf{k} - \tilde{\mathbf{k}}| \geq 2^{J+3}. \quad (3.19)$$

To see that this cover exists, we momentarily take Fermi coordinates  $(z_1, z') \in \mathbb{R}^d$  adapted to a hypersurface orthogonal to the geodesic segment  $\{\kappa_{s,0}(\mu) : |s| \leq \varepsilon\}$ , so that  $\{z' = 0\}$  identifies the hypersurface and  $r \mapsto (r, z')$  always parameterizes a unit speed geodesic. In particular, we assume that  $r \mapsto (r, 0)$  parameterizes the geodesic segment given by  $s \mapsto \kappa_{s,0}(\mu)$  so that  $\mathcal{E}_{\mu}^j \subset \{z : |z'| \lesssim 2^j\}$ . In these coordinates, we take a partition of unity  $\sum \chi_{\mathbf{k}}^2 = 1$  so that with  $\mathbf{k} = (k_2, \dots, k_d)$

$$\text{supp}(\chi_{\mathbf{k}}) \subset \tilde{\mathcal{T}}_{\mathbf{k}} := \left\{(z_1, z') : |z_1| \lesssim \varepsilon, \max_{2 \leq i \leq d} |z_i - k_i| \leq 2^{J+1}\right\}.$$

The property  $(3.19)$  is immediate in these coordinates. The set  $\tilde{\mathcal{T}}_{\mathbf{k}}$  is a neighborhood about the geodesic segment  $\gamma_{\mathbf{k}}$  given as the image of  $r \mapsto (r, \mathbf{k})$ . Since  $2^J \approx h^{1/2}$ , we have that  $\tilde{\mathcal{T}}_{\mathbf{k}} \subset \mathcal{T}_{C_1 h^{1/2}}(\gamma_{\mathbf{k}})$  for some  $C_1$  large, and  $(3.18)$  follows easily in any coordinate system. We now take  $\mathbf{k}_0 = 0$  and it is then verified that we may take  $C$  so that  $(3.17)$  is satisfied and  $d(\tilde{\mathcal{T}}_{\mathbf{k}}, \tilde{\mathcal{T}}_{\mathbf{k}_0}) \gtrsim |\mathbf{k} - \mathbf{k}_0|$  whenever  $d\left(\tilde{\mathcal{T}}_{\mathbf{k}}, \mathcal{E}_{\mu}^j\right) \geq C2^j$ . We now revert back to the original coordinates and at the cost of enlarging  $C$  and the other implicit constants,  $(3.16)$ ,  $(3.17)$ , and  $(3.19)$  are all satisfied here since the diffeomorphism can be taken to satisfy a Lipschitz bound with uniform constant.

Given the above,  $(3.18)$  means that it suffices to show the two inequalities

$$\|B_{\mu,\omega}^j(s) f_h(s)\|_{L^2(\mathbb{R}^{d-1})}^2 \lesssim (h^{-1/2}2^j)^{d-1} \sup_{\mathbf{k}} \|(\chi_{\mathbf{k}} f_h)(s, \cdot)\|_{L^2(\mathbb{R}^{d-1})}^2, \quad (3.20)$$

$$\int_{-\varepsilon}^{\varepsilon} \|(\chi_{\mathbf{k}} f_h)(s, \cdot)\|_{L^2(\mathbb{R}^{d-1})}^2 ds \lesssim \sup_{\tilde{\mathbf{k}}} \int \|(\chi_{\tilde{\mathbf{k}}} \psi_h)(t, \cdot)\|_{L^2(\mathbb{R}^{d-1})}^2 dt, \quad (3.21)$$

We primarily focus on  $(3.20)$ , as the second bound will be seen to follow from very similar ideas. Let  $B_{\mu,\omega,\mathbf{k}}(s)$  be the operator with Schwartz kernel given by

$$\frac{1}{(2\pi h)^{d-1}} \int e^{\frac{i}{h} \langle y-x, \eta \rangle} b_{\mu,\omega}(s, y, \eta) \chi_{\mathbf{k}}(s, x) d\eta, \quad (3.22)$$

and since  $\sum_{\mathbf{k}} \chi_{\mathbf{k}}^2 = 1$ , this means that

$$B_{\mu,\omega}^j(s)(f_h(s, \cdot)) = B_{\mu,\omega}^j(s) \left( \sum_{\mathbf{k}} (\chi_{\mathbf{k}}^2 f_h)(s, \cdot) \right) = \sum_{\mathbf{k}} B_{\mu,\omega,\mathbf{k}}(s) ((\chi_{\mathbf{k}} f_h)(s, \cdot)).$$

Partition the indices  $\mathbf{k}$  into sets  $K_1, K_2$  where  $\mathbf{k} \in K_1$  if  $d\left(\tilde{\mathcal{T}}_{\mathbf{k}}, \mathcal{E}_{\mu}^j\right) < C2^j$  and  $\mathbf{k} \in K_2$  otherwise. The compound symbol of  $B_{\mu,\omega,\mathbf{k}}(s)$  is in  $S_{1/2}(1)$  so the operator is uniformly bounded in  $L^2(\mathbb{R}^{d-1})$ . Hence

$$\begin{aligned} \left\| B_{\mu, \omega}(s) \left( \sum_{\mathbf{k} \in K_1} \chi_{\mathbf{k}} f_h(s, \cdot) \right) \right\|_{L^2(\mathbb{R}^{d-1})}^2 &\lesssim \int \left| \sum_{k \in K_1} \chi_{\mathbf{k}}^2(x) f_h(s, x) \right|^2 dx \\ &\lesssim \sum_{k \in K_1} \int |\chi_{\mathbf{k}}(x) f_h(s, x)|^2 dx. \end{aligned}$$

where we have used (3.19) and  $\chi_{\mathbf{k}}^2(x) \leq \chi_{\mathbf{k}}(x)$  as a bump function. Given (3.17), the expression on the right is dominated by the right hand side of (3.20).

We conclude (3.20) by showing that for any  $N \geq 0$ ,

$$\|B_{\mu, \omega, \mathbf{k}}(s)\|_{L^2 \rightarrow L^2} \lesssim_N (h^{-\frac{1}{2}} |\mathbf{k} - \mathbf{k}_0|)^{-N} \quad (3.23)$$

so that by taking  $N$  large

$$\left\| \sum_{\mathbf{k} \in K_2} B_{\mu, \omega, \mathbf{k}}(s) ((\chi_{\mathbf{k}} f_h)(s, \cdot)) \right\|_{L^2} \lesssim \sum_{\mathbf{k} \in K_2} (2^{-J} |\mathbf{k} - \mathbf{k}_0|)^{-N} \|(\chi_{\mathbf{k}} f_h)(s, \cdot)\|_{L^2} \quad (3.24)$$

and the last expression is seen to be bounded by the right hand side of (3.20). To see (3.23), observe that an integration by parts in (3.22) similar to (2.24) shows that we may take the compound symbol  $b_{\mu, \omega, \mathbf{k}}^j(s, y, x, \eta)$  of  $B_{\mu, \omega, \mathbf{k}}^j(s)$  to satisfy

$$|(h^{\frac{1}{2}} \partial_{y, x, \eta})^\alpha b_{\mu, \omega, \mathbf{k}}^j(s, y, x, \eta)| \lesssim_{\alpha, N} (1 + h^{-1} |y - x|^2)^{-N} \left( 1 + h^{-\frac{1}{2}} d(y, \eta; \kappa_{s, 0}(\mathcal{D}_\mu^j)) \right)^{-2N}.$$

The right hand side here is thus  $O((1 + h^{-1/2} d(x, \mathcal{E}_\mu^j))^{-N})$  and since we may restrict to  $x$  such that  $\chi_{\mathbf{k}}(s, x) \neq 0$ , this in turn is  $O((h^{-\frac{1}{2}} |\mathbf{k} - \mathbf{k}_0|)^{-N})$  by (3.16).

The proof of (3.21) is similar, the primary difference is that  $f_h$  is the image of  $\psi_h$  under the operator in (2.6), so the estimates are over  $\mathbb{R}^d$  instead of  $\mathbb{R}^{d-1}$ . But given (3.19), the same principles as before apply here, as the kernel of the operator in (2.6) rapidly decreases on the scale of  $h^{1/2}$  away from the diagonal. The only other significant difference is that the analog of (3.24) should be adjusted to read

$$\left\| \sum_{\tilde{\mathbf{k}} \in K_2} \chi_{\tilde{\mathbf{k}}} f_h \right\|_{L^2(\mathbb{R}^d)} \lesssim \sum_{\tilde{\mathbf{k}} \in K_2} (2^{-J} |\mathbf{k} - \tilde{\mathbf{k}}|)^{-N} \|\chi_{\tilde{\mathbf{k}}} \psi_h\|_{L^2(\mathbb{R}^d)}.$$

□

We now return to estimating (3.9) and state the two main theorems which form the crux of the proof of Theorem 1.1. The first step in is to use an almost orthogonality theorem, valid when  $j = J$  or when  $j \geq J + 1$ , which will be proved in Sect. 4. Note that  $\kappa_{t, 0}(\mu) = (x_{t, 0}(\mu), \xi_{t, 0}(\mu))$  is the image of  $\mu \in 2^j \mathbb{Z}^{2(d-1)}$  under the map  $\kappa_{t, 0}$  defined after (2.8).

**Theorem 3.3.** *Suppose  $1 \leq p \leq \infty$  and  $p^* = \min(p, p')$ . There exists a family of  $h$ -pseudodifferential operators  $B_{\mu, \omega}^j(s)$  which satisfy (3.11) and with  $\omega$  ranging over a finite collection of indices depending only on the dimension, such that*

$$\begin{aligned} & \left\| \sum_{\mu, \mu' \in \Xi_j} T(B_\mu^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^p} \\ & \lesssim \sum_{\omega, \omega'} \left( \sum_{\mu, \mu' \in \Xi_j} \left\| T_{\mu, \omega}^j (B_{\mu, \omega}^j f_h) T_{\mu', \omega'} (B_{\mu', \omega'}^j f_h) \right\|_{L^p}^{p^*} \right)^{\frac{1}{p^*}}, \end{aligned} \quad (3.25)$$

when  $0 \geq j \geq J+1$  or when  $j = J$ . Here  $T_{\mu, \omega}^j$  (and similarly  $T_{\mu', \omega'}^j$ ) is defined by

$$(T_{\mu, \omega}^j F)(t, x) = \int_{-\varepsilon}^t \left( T_{\mu, \omega}^{j, s} F(s, \cdot) \right) (x) ds \quad (3.26)$$

where  $T_{\mu, \omega}^{j, s}$  is the fixed  $s$  operator defined by the formula

$$(T_{\mu, \omega}^{j, s} f)(t, x) = \frac{1}{(2\pi h)^{d-1}} \int e^{\frac{i}{h}(\phi(t, s, x, \eta) - \langle y, \eta \rangle)} a_{\mu, \omega}^j(t, s, x, \eta) f(y) dy d\eta \quad (3.27)$$

with  $\phi$  as in (2.9) but with symbol  $a_{\mu, \omega}^j(t, s, x, \eta)$  satisfying

$$|\partial_{x, \eta}^\alpha a_{\mu, \omega}^j(t, s, x, \eta)| \lesssim 2^{-j|\alpha|} (1 + 2^{-j}|x - x_{t, 0}(\mu)|)^{-2d} (1 + 2^{-j}|\eta - \xi_{s, 0}(\mu)|)^{-2d}.$$

The second step in estimating this is to employ the bilinear estimates in the following theorem, which we prove in Sect. 5:

**Theorem 3.4.** Suppose  $j = J, \dots, 0$ ,  $\mu, \mu' \in \Xi_j$  and  $T_{\mu, \omega}^j, T_{\mu', \omega'}^j, B_{\mu, \omega}^j, B_{\mu', \omega'}^j$  are as in Theorem 3.3. There exist corresponding ( $s$  dependent)  $h$ -PDO  $\tilde{B}_{\mu, \omega}^j$  and  $\tilde{B}_{\mu', \omega'}^j$  satisfying (3.11) such that the bilinear operator  $R$  defined for fixed  $s, s' \in [-\varepsilon, \varepsilon]$  by

$$\begin{aligned} & R(B_{\mu, \omega}^j(s) f, B_{\mu', \omega'}^j(s') g)(t, x) \\ & := \left( T_{\mu, \omega}^{j, s} (B_{\mu, \omega}^j(s) f) \right) (t, x) \cdot \left( T_{\mu', \omega'}^{j, s'} (B_{\mu', \omega'}^j(s') g) \right) (t, x). \end{aligned}$$

satisfies bounds

$$\begin{aligned} & h^{-\frac{2d}{q} + (d-1)} 2^{-j(d-1 - \frac{2(d+1)}{q})} \left\| R(B_{\mu, \omega}^j(s) f, B_{\mu', \omega'}^j(s') g) \right\|_{L^{\frac{q}{2}}([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})} \\ & \lesssim \|B_{\mu, \omega}^j(s) f\|_{L^2} \|B_{\mu', \omega'}^j(s') g\|_{L^2} + \|\tilde{B}_{\mu, \omega}^j(s) f\|_{L^2} \|B_{\mu', \omega'}^j(s') g\|_{L^2} \\ & \quad + \|B_{\mu, \omega}^j(s) f\|_{L^2} \|\tilde{B}_{\mu', \omega'}^j(s') g\|_{L^2} + \|\tilde{B}_{\mu, \omega}^j(s) f\|_{L^2} \|\tilde{B}_{\mu', \omega'}^j(s') g\|_{L^2}, \end{aligned} \quad (3.28)$$

where the norms on the right hand side are taken to be  $L^2(\mathbb{R}^{d-1})$ .

Note that the first factor on the left rewrites as  $h^{(d-1)(\frac{1}{2} - \frac{1}{q})} (h^{\frac{1}{2}} 2^{-j})^{\frac{d-1}{q}(q - \frac{2(d+1)}{d-1})}$ , showing that the phase space separation of the  $B_{\mu, \omega}^j$  and  $B_{\mu', \omega'}^j$  (up to rapidly decreasing tails) results in a gain of  $(h^{\frac{1}{2}} 2^{-j})^{\frac{d-1}{q}(\frac{2(d+1)}{d-1} - q)}$  over what would be obtained by applying the linear estimates (2.13). The  $\tilde{B}_{\mu, \omega}^j$  are slight distortions of the  $B_{\mu, \omega}^j$  that arise when replacing the rapidly decaying symbol of  $T_{\mu, \omega}^j$  by a compactly supported one. We take the

family (3.10) to consist of the union of the operators constructed in these two theorems, that is, operators of the form  $B_{\mu,\omega}^j(s)$  or  $\tilde{B}_{\mu,\omega}^j(s)$ .

We can now begin discussing how to estimate the terms in (3.9), showing that the two theorems here imply (1.10), (1.11), and hence Theorem 1.1. Given Theorem 3.3, we have for any  $J \leq j \leq 0$ ,

$$\begin{aligned} & \left\| \sum_{(\mu,\mu') \in \Xi_j} T(B_{\mu}^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \\ & \lesssim \sum_{\omega,\omega'} \left( \sum_{\mu,\mu' \in \Xi_j} \left\| T_{\mu,\omega}^j (B_{\mu,\omega}^j f_h) T_{\mu',\omega'}^j (B_{\mu',\omega'}^j f_h) \right\|_{L^{\frac{q}{2}}} \right)^{\frac{1}{(\frac{q}{2})^*}}. \end{aligned}$$

where the  $L^{\frac{q}{2}}$  norms are taken over  $[-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1}$ . We now fix  $\omega, \omega'$  and show that Theorem 3.4 supplies bounds on each term on the right hand side here. In this case Minkowski's inequality for integrals shows that

$$\begin{aligned} & \left\| T_{\mu,\omega}^j (B_{\mu,\omega}^j f_h) T_{\mu',\omega'}^j (B_{\mu',\omega'}^j f_h) \right\|_{L^{\frac{q}{2}}} \\ & \lesssim \int_{[-\varepsilon, \varepsilon]^2} \left( \int_{-\varepsilon}^{\varepsilon} \left\| T_{\mu,\omega}^{j,s} (B_{\mu,\omega}^j(s) f_h(s)) \cdot T_{\mu',\omega'}^{j,s'} (B_{\mu',\omega'}^j(s') f_h(s')) \right\|_{L^{\frac{q}{2}}(\mathbb{R}^{d-1})}^{\frac{q}{2}} dt \right)^{\frac{2}{q}} ds ds' \end{aligned}$$

Hence (3.28) and Hölder's inequality show that this in turn is bounded by

$$\begin{aligned} & h^{\frac{2d}{q} - (d-1)} 2^{j(d-1 - \frac{2(d+1)}{q})} \left( \|B_{\mu,\omega}^j f_h\|_{L^2(\mathbb{R}^d)} \|B_{\mu',\omega'}^j f_h\|_{L^2(\mathbb{R}^d)} \right. \\ & \quad + \|\tilde{B}_{\mu,\omega}^j f_h\|_{L^2(\mathbb{R}^d)} \|B_{\mu',\omega'}^j f_h\|_{L^2(\mathbb{R}^d)} + \|B_{\mu,\omega}^j f_h\|_{L^2(\mathbb{R}^d)} \|\tilde{B}_{\mu',\omega'}^j f_h\|_{L^2(\mathbb{R}^d)} \\ & \quad \left. + \|\tilde{B}_{\mu,\omega}^j f_h\|_{L^2(\mathbb{R}^d)} \|\tilde{B}_{\mu',\omega'}^j f_h\|_{L^2(\mathbb{R}^d)} \right). \end{aligned} \quad (3.29)$$

We will now bound the right hand side of (3.29) in two different ways, and then show how to optimize the choice. If  $(\mu, \mu') \in \Xi_j$ , then  $2^{-j}(\mu - \mu')$  lies in a fixed, finite collection of vectors (of size  $\approx 1$  when  $j \geq J+1$  or  $O(1)$  when  $j = J$ ). Hence

$$\#\{\mu' : (\mu, \mu') \in \Xi_j\} = O(1) \text{ for any fixed } \mu \in 2^j \mathbb{Z}^{2(d-1)}. \quad (3.30)$$

We now have for any two  $\omega, \omega' \in \Omega$ ,

$$\begin{aligned} & \left\| T_{\mu,\omega}^j f_h T_{\mu',\omega'}^j f_h \right\|_{\ell_{\mu}^{(\frac{q}{2})^*} L^{\frac{q}{2}}} \\ & \lesssim h^{\frac{2d}{q} - (d-1)} 2^{j(d-1 - \frac{2(d+1)}{q})} \left( \sum_{\mu} \sum_{\mu':(\mu,\mu') \in \Xi_j} \|B_{\mu,\omega}^j f_h\|_{L^2}^{(\frac{q}{2})^*} \|B_{\mu',\omega'}^j f_h\|_{L^2}^{(\frac{q}{2})^*} \right)^{\frac{1}{(\frac{q}{2})^*}} + \dots \end{aligned} \quad (3.31)$$

where the dots denote the corresponding contributions of the last three terms in (3.29). Observe that when  $d \geq 3$ ,  $(\frac{q}{2})^* = \frac{q}{2}$  (since  $q < \frac{2(d+1)}{d-1} \leq 4$  in this case) and hence by

Cauchy–Schwarz and (3.30), the right hand side in (3.31) is dominated by

$$h^{\frac{2d}{q}-(d-1)} 2^{j(d-1-\frac{2(d+1)}{q})} \left( \sum_{\mu} \|B_{\mu, \omega}^j f_h\|_{L^2}^q \right)^{\frac{1}{q}} \left( \sum_{\mu'} \|B_{\mu', \omega'}^j f_h\|_{L^2}^q \right)^{\frac{1}{q}} + \dots$$

Indeed given (3.30), we may view  $\mu'$  as a function of  $\mu$  at the cost of increasing the implicit constant by a factor determined the  $O(1)$  constant there. By the embedding  $\ell^2 \hookrightarrow \ell^q$  and Lemma 3.1, we have that this in turn is bounded by  $h^{\frac{2d}{q}-(d-1)} 2^{j(d-1-\frac{2(d+1)}{q})} \|f_h\|_{L^2}^2$ .

When  $d = 2$ , the same conclusion holds since  $(\frac{q}{2})^* = (\frac{q}{2})' = \frac{q}{q-2}$  and  $\ell^2 \hookrightarrow \ell^{\frac{2q}{q-2}}$ .

On the other hand, when  $d \geq 3$  we may estimate the sum appearing on the right in (3.31) (and similarly the other terms) by

$$\begin{aligned} & \left( \sum_{\mu} \sum_{\mu': (\mu, \mu') \in \Xi_j} \|B_{\mu, \omega}^j f_h\|_{L^2}^{\frac{q}{2}} \|B_{\mu', \omega'}^j f_h\|_{L^2}^{\frac{q}{2}} \right)^{\frac{2}{q}} \lesssim \|f_h\|_{L^2} \left( \sum_{\mu'} \|B_{\mu', \omega'}^j f_h\|_{L^2}^{\frac{2q}{4-q}} \right)^{\frac{4-q}{2q}} \\ & \lesssim \|f_h\|_{L^2}^{\frac{4}{q}} \left( \sup_{\mu'} \|B_{\mu', \omega'}^j f_h\|_{L^2} \right)^{2-\frac{4}{q}} \lesssim 2^{j(d-1)(1-\frac{2}{q})} \|f_h\|_{L^2}^{\frac{4}{q}} \|f_h\|_{MKN}^{2-\frac{4}{q}}, \end{aligned}$$

where the first inequality follows by Hölder's inequality with exponents satisfying  $\frac{2}{q} = \frac{1}{2} + \frac{4-q}{2q}$  and the second inequality follows by using that  $\frac{2q}{4-q} = 2 + \frac{4q-8}{4-q}$ . Note that we have used Lemma 3.1 twice in the process and also (3.30) similar to before. When  $d = 2$ , a similar argument checks that the sum appearing on the right in (3.31) is majorized by

$$\|f_h\|_{L^2}^{2-\frac{4}{q}} \left( \sup_{\mu'} \|B_{\mu', \omega'}^j f_h\|_{L^2} \right)^{\frac{4}{q}} \lesssim 2^{\frac{2j}{q}} \|f_h\|_{L^2}^{2-\frac{4}{q}} \|f_h\|_{MKN}^{\frac{4}{q}}.$$

Repeating this for the finite collection of pairs  $\omega, \omega'$ , we have that when  $d \geq 3$ ,

$$\left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_{\mu}^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \lesssim h^{\frac{2d}{q}-(d-1)} 2^{j((d-1)(2-\frac{2}{q})-\frac{2(d+1)}{q})} \|f_h\|_{L^2}^{\frac{4}{q}} \|f_h\|_{MKN}^{2-\frac{4}{q}}, \quad (3.32)$$

though the second factor on the right rewrites as  $2^{2j(\frac{d-1}{q})(q-\frac{2d}{d-1})}$ . When  $d = 2$ ,

$$\left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_{\mu}^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \lesssim h^{\frac{2d}{q}-(d-1)} 2^{j(1-\frac{4}{q})} \|f_h\|_{L^2}^{2-\frac{4}{q}} \|f_h\|_{MKN}^{\frac{4}{q}}. \quad (3.33)$$

At the same time, in all cases we have

$$\left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_{\mu}^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \lesssim h^{\frac{2d}{q}-(d-1)} 2^{j(d-1-\frac{2(d+1)}{q})} \|f_h\|_{L^2}^2. \quad (3.34)$$

Consequently, if we fix an integer  $J \leq L \leq 0$ , then when  $d \geq 3$ , (3.32) yields

$$\sum_{j=J}^L \left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_\mu^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \lesssim h^{\frac{2d}{q} - (d-1)} 2^{L((d-1)(2 - \frac{2}{q}) - \frac{2(d+1)}{q})} \|f_h\|_{L^2}^{\frac{4}{q}} \|f_h\|_{MKN}^{2 - \frac{4}{q}},$$

since  $q > \frac{2(d+2)}{d} > \frac{2d}{d-1}$  implies that  $(d-1)(2 - \frac{2}{q}) - \frac{2(d+1)}{q} > 0$ . Moreover,  $d-1 - \frac{2(d+1)}{q} < 0$ , so (3.34) yields

$$\sum_{j=L}^0 \left\| \sum_{(\mu, \mu') \in \Xi_j} T(B_\mu^j f_h) T(B_{\mu'}^j f_h) \right\|_{L^{\frac{q}{2}}} \lesssim h^{\frac{2d}{q} - (d-1)} 2^{L(d-1 - \frac{2(d+1)}{q})} \|f_h\|_{L^2}^2.$$

It is then verified that we optimize the bounds on (3.9) by taking  $L$  such that  $2^{L(d-1)/2} \approx \|f_h\|_{L^2}/\|f_h\|_{MKN}$ , which is possible since the latter quantity is  $O(1)$  and bounded below by  $h^{\frac{d-1}{4}}$ . This shows that (3.9) is bounded by

$$h^{\frac{2d}{q} - (d-1)} \|f_h\|_{L^2}^{4 - \frac{4(d+1)}{q(d-1)}} \|f_h\|_{MKN}^{\frac{4(d+1)}{q(d-1)} - 2}$$

When  $d = 2$  and  $4 < q < 6$ , a similar argument using (3.33) and (3.34) shows the same bound. Finally, when  $d = 2$  and  $q = 4$  we obtain that (3.9) is bounded by

$$\log(h^{-1}) \|f_h\|_{L^2} \|f_h\|_{MKN}.$$

Given (2.5) and Lemma 3.2, this completes the proofs of (1.10), (1.11).

Another perspective on (3.32) and (3.34) (and similarly when  $d = 2$ ) results from noting that right hand sides of these two inequalities are respectively bounded by

$$h^{-2\delta(q)} (h 2^{-2L})^{\frac{d-1}{q}(q - \frac{2d}{d-1})} \|\psi_h\|_{L^2}^{\frac{4}{q}} \|\psi_h\|_{KN}^{2 - \frac{4}{q}},$$

$$h^{-2\delta(q)} (h 2^{-2L})^{\frac{d-1}{2q}(\frac{2(d+1)}{d-1} - q)} \|\psi_h\|_{L^2}^2,$$

after an application of (3.15) and  $\|f_h\|_{L^2} \lesssim \|\psi_h\|_{L^2}$  (as well as recalling the differences between (1.5), (1.6)). The latter of these underscores the gain resulting from the bilinear estimates while the former trades a loss in  $h 2^{-2L}$  for a gain in the Kakeya–Nikodym norm of  $\psi_h$ . This also shows that  $L$  is taken so that

$$\|\psi_h\|_{KN} \approx \|\psi_h\|_{L^2} (h 2^{-2L})^{\frac{d-1}{4}}.$$

#### 4. Almost Orthogonality

In this section, we prove Theorem 3.3. We begin with a lemma, recalling the notation for  $(x_{t,0}(\mu), \xi_{t,0}(\mu))$  following (2.8) and prior to Theorem 3.3.

**Lemma 4.1.** *Let  $2^j \in [h^{\frac{1}{2}}, 1]$ ,  $\mu \in 2^j \mathbb{Z}^{2(d-1)}$  and suppose that  $\alpha, \beta$  are multiindices such that  $|\alpha|, |\beta| \leq 2d$ . Given  $f \in \mathcal{S}(\mathbb{R}^{d-1})$ , let*

$$F_\mu^s(y) := \langle 2^{-j}(y - x_{s,0}(\mu)) \rangle^{4d} \langle (2^{-j}(hD_y - \xi_{s,0}(\mu)))^{4d} f \rangle(y),$$

with  $\langle w \rangle = (1 + |w|^2)^{\frac{1}{2}}$ . There exists a symbol  $a_{\mu,\alpha,\beta}^j(t, s, x, \xi)$  satisfying

$$\left| D_{x,\xi}^\gamma a_{\mu,\alpha,\beta}^j(t, s, x, \xi) \right| \lesssim_\gamma 2^{-j|\gamma|} \langle 2^{-j}(y - x_{t,0}(\mu)) \rangle^{-2d} \langle 2^{-j}(\xi - \xi_{s,0}(\mu)) \rangle^{-2d} \quad (4.1)$$

such that up to negligible  $O(h^\infty)$  error,

$$\begin{aligned} & 2^{-j(|\alpha|+|\beta|)}(x - x_{t,0}(\mu))^\alpha (hD_x - \xi_{t,0}(\mu))^\beta (S(t, s)f)(x) \\ &= \frac{1}{(2\pi h)^{d-1}} \iint e^{\frac{i}{h}(\phi(t,s,x,\xi) - \langle y, \xi \rangle)} a_{\mu,\alpha,\beta}^j(t, s, x, \xi) F_\mu^s(y) dy d\xi. \end{aligned} \quad (4.2)$$

*Proof.* Let  $A(t, s, x, y, \eta)$  be defined by

$$A(t, s, x, y, \eta) = a(t, s, x, \eta) \langle 2^{-j}(\eta - \xi_{s,0}(\mu)) \rangle^{-4d} \langle 2^{-j}(y - x_{s,0}(\mu)) \rangle^{-4d}$$

so that up to  $O(h^\infty)$  error (resulting only from the error term in (2.9)),

$$\begin{aligned} & (2\pi h)^{d-1} S(t, s)f(x) \\ &= \int e^{\frac{i}{h}\phi(t,s,x,\eta)} a(t, s, x, \eta) \langle 2^{-j}(\eta - \xi_{s,0}(\mu)) \rangle^{-4d} \mathcal{F}_h(\langle 2^{-j}(hD - \xi_{s,0}(\mu))^{-4d} f \rangle)(\eta) d\eta \\ &= \iint e^{\frac{i}{h}(\phi(t,s,x,\eta) - \langle y, \eta \rangle)} A(t, s, x, y, \eta) F_\mu^s(y) dy d\eta, \end{aligned}$$

where we have inverted  $\mathcal{F}_h$  in the last expression and used that the argument of  $\mathcal{F}_h$  in the second expression is  $\langle 2^{-j}(y - x_{s,0}(\mu)) \rangle^{-4d} F_\mu^s(y)$ . Next we write

$$\begin{aligned} & 2^{-j(|\alpha|+|\beta|)}(x - x_{t,0}(\mu))^\alpha (hD_x - \xi_{t,0}(\mu))^\beta \int e^{\frac{i}{h}(\phi(t,s,x,\eta) - \langle y, \eta \rangle)} A(t, s, x, y, \eta) d\eta \\ &= \int e^{\frac{i}{h}(\phi(t,s,x,\eta) - \langle y, \eta \rangle)} A_1(t, s, x, y, \eta) d\eta \end{aligned} \quad (4.3)$$

where

$$\begin{aligned} & 2^{j|\gamma|} |D_{x,y,\eta}^\gamma A_1(t, s, \cdot)| \lesssim_\gamma \langle 2^{-j}(\eta - \xi_{s,0}(\mu)) \rangle^{-4d} \langle 2^{-j}(y - x_{s,0}(\mu)) \rangle^{-4d} \\ & \quad \times \langle 2^{-j}(x - x_{t,0}(\mu)) \rangle^{|\alpha|} \langle 2^{-j}(d_x \phi(t, s, x, \eta) - \xi_{t,0}(\mu)) \rangle^{|\beta|}. \end{aligned}$$

The differential operator

$$\frac{1 + 2^{-j}(d_\eta \phi(t, s, x, \eta) - y) \cdot (h2^{-j} D_\eta)}{1 + 2^{-2j}|d_\eta \phi(t, s, x, \eta) - y|^2} \quad (4.4)$$

preserves the exponential function in the integral on the right in (4.3). Moreover, since  $h2^{-j} \leq 2^j$ , the weighted derivative  $h2^{-j} D_\eta$  appearing here means that the adjoint of the differential operator applied to  $A_1$  has the effect of replacing that amplitude by one of the same regularity. Using that

$$\begin{aligned} & \langle 2^{-j} |y - x_{s,0}(\mu)| \rangle^{-4d} \langle 2^{-j} |d_\eta \phi(t, s, x, \eta) - y| \rangle^{-4d} \\ & \lesssim \langle 2^{-j} |d_\eta \phi(t, s, x, \eta) - x_{s,0}(\mu)| \rangle^{-4d}, \end{aligned}$$

integration by parts shows that (4.3) can further be rewritten as

$$\int e^{\frac{i}{h}(\phi(t,s,x,\eta)-\langle y,\eta \rangle)} A_2(t, s, x, y, \eta) d\eta$$

where

$$\begin{aligned} & 2^{j|\gamma|} |D_{x,y,\eta}^\gamma A_2(t, s, \cdot)| \\ & \lesssim_\gamma \langle 2^{-j} (\eta - \xi_{s,0}(\mu)) \rangle^{-4d} \langle 2^{-j} (d_\eta \phi(t, s, x, \eta) - x_{s,0}(\mu)) \rangle^{-4d} \\ & \quad \times \langle 2^{-j} (x - x_{t,0}(\mu)) \rangle^{|\alpha|} \langle 2^{-j} (d_x \phi(t, s, x, \eta) - \xi_{t,0}(\mu)) \rangle^{|\beta|} \\ & \lesssim \langle 2^{-j} (\eta - \xi_{s,0}(\mu)) \rangle^{-2d} \langle 2^{-j} (x - x_{t,0}(\mu)) \rangle^{-2d}, \end{aligned}$$

where the last inequality follows since  $|\alpha|, |\beta| \leq 2d$  and  $\kappa_{t,s}$  is Lipschitz in  $(x, \xi)$ .

We now define

$$a_{\mu,\alpha,\beta}^j(t, s, x, \xi) := \frac{e^{-\frac{i}{h}\phi(t,s,x,\xi)}}{(2\pi h)^{d-1}} \iint e^{\frac{i}{h}(\phi(t,s,x,\eta)+\langle y,\xi-\eta \rangle)} A_2(t, s, x, y, \eta) dy d\eta \quad (4.5)$$

and are left to show (4.1). Abbreviate the phase function here as

$$\Phi(t, s, x, y, \eta, \xi) = \phi(t, s, x, \eta) - \phi(t, s, x, \xi) + \langle y, \xi - \eta \rangle.$$

First observe that for a single derivative

$$D_{x_m} e^{\frac{i}{h}\Phi} = - \sum_{l=1}^{d-1} \left( \int_0^1 \partial_{x_m \xi_l}^2 \phi(t, s, x, (1-r)\xi + r\eta) dr \right) D_{y_l} e^{\frac{i}{h}\Phi}.$$

Integration by parts thus shows that applying a weighted derivative  $2^j D_{x_m}$  to  $a_{\mu,\alpha,\beta}^j$  has the effect of replacing  $A_2$  by an amplitude of the same regularity. For derivatives in  $\xi$ , note that

$$D_{\xi_m} e^{\frac{i}{h}\Phi} = h^{-1} (\partial_{\eta_m} \phi(t, s, x, \eta) - \partial_{\xi_m} \phi(t, s, x, \xi)) e^{\frac{i}{h}\Phi} - D_{\eta_m} e^{\frac{i}{h}\Phi}.$$

As before, the first term on the right here is thus equivalent to the action of a vector field acting on the exponential in the  $y$  variable. Hence integration by parts in  $y, \eta$  similar to before shows that applying a weighted derivative  $2^j D_{\xi_m}$  to  $a_{\mu,\alpha,\beta}^j$  again has the effect of replacing  $A_2$  by an amplitude of the same regularity.

We now conclude the proof by observing that the expression in (4.5) is uniformly bounded in  $h$ . But this follows from integrating by parts with respect to the operator in (4.4) with  $2^{-j} = h^{-1/2}$  sufficiently many times so that integration in  $y, \eta$  yields a gain of  $O(h^{d-1})$ .  $\square$

*Proof of Theorem 3.3.* Let  $Q_{1,\mu}^j(t)$ ,  $Q_{2,\mu}^j(t)$  be the multiplier and Fourier multiplier  $Q_{1,\mu}^j(t) := (1 + 2^{-2j}|x - x_{t,0}(\mu)|^2)^{-2d}$ ,  $Q_{2,\mu}^j(t) := (1 + 2^{-2j}|hD_x - \xi_{t,0}(\mu)|^2)^{-2d}$ .

Thus if  $c(w)$  denotes the polynomial  $c(w) = (1 + 2^{-2j}|w|^2)^{2d}$  for  $w \in \mathbb{R}^{d-1}$ , we have

$$\begin{aligned} & T(B_\mu^j f_h) \cdot T(B_{\mu'}^j f_h) \\ &= Q_{1,\mu}^j(t) Q_{2,\mu}^j(t) c(x - x_{t,0}(\mu)) c(hD_x - \xi_{t,0}(\mu)) \left( T(B_\mu^j f_h)(t) \cdot T(B_{\mu'}^j f_h)(t) \right) \\ &+ Q_{1,\mu}^j(t) [c(x - x_{t,0}(\mu)), Q_{2,\mu}^j(t)] c(hD_x - \xi_{t,0}(\mu)) \left( T(B_\mu^j f_h)(t) \cdot T(B_{\mu'}^j f_h)(t) \right). \end{aligned} \quad (4.6)$$

In turn the commutator  $[c(x - x_{t,0}(\mu)), Q_{2,\mu}^j(t)]$  can be realized as

$$[c(x - x_{t,0}(\mu)), Q_{2,\mu}^j(t)] = \sum_{|\alpha| \leq 2d} Q_{2,\mu}^{j,\alpha}(t) c_\alpha(x - x_{t,0}(\mu)) \quad (4.7)$$

where the symbol of  $Q_{2,\mu}^{j,\alpha}(t)$  is  $(h2^{-j}\partial_\xi)^\alpha (1 + 2^{-2j}|\xi - \xi_{t,0}(\mu)|^2)^{-2d}$  and  $c_\alpha$  is a polynomial of degree  $|\alpha|$ . We thus concern ourselves with the contribution of the middle line in (4.6), as similar arguments treat the last line.

We now observe that

$$c(x - x_{t,0}(\mu)) c(hD_x - \xi_{t,0}(\mu)) \left( T(B_\mu^j f_h) T(B_{\mu'}^j f_h) \right) \quad (4.8)$$

is a sum consisting of terms of the form

$$2^{-j(|\alpha|+|\beta|+|\gamma|)} (x - x_{t,0}(\mu))^\alpha (hD_x - \xi_{t,0}(\mu))^\beta T(B_\mu^j f_h) \cdot (hD_x - \xi_{t,0}(\mu))^\gamma T(B_{\mu'}^j f_h).$$

The operator  $2^{-j|\gamma|}(hD_x - \xi_{t,0}(\mu))^\gamma$  can be written as a sum consisting of terms

$$2^{-j|\gamma_1|} (\xi_{t,0}(\mu) - \xi_{t,0}(\mu'))^{\gamma_1} \cdot 2^{-j|\gamma_2|} (hD_x - \xi_{t,0}(\mu))^{\gamma_2}, \quad \gamma_1 + \gamma_2 = \gamma,$$

but since  $|\xi_{t,0}(\mu) - \xi_{t,0}(\mu')| \lesssim 2^j$ , the first factor here is a bounded function.

We next claim that

$$2^{-j(|\alpha|+|\beta|)} (x - x_{t,0}(\mu))^\alpha (hD_x - \xi_{t,0}(\mu))^\beta B_\mu^j f_h = B_{\mu,\alpha,\beta}^j f_h$$

for some  $h$ -PDO  $B_{\mu,\alpha,\beta}^j$  whose symbol satisfies (3.11). Its symbol is

$$2^{-j(|\alpha|+|\beta|)} (x - x_{t,0}(\mu))^\alpha (\xi - \xi_{t,0}(\mu))^\beta b_\mu^j(t, x, \xi).$$

For each  $(x, \xi)$ , we claim that the absolute value of this symbol is bounded above by  $(1 + h^{-\frac{1}{2}}d(x, \xi; \kappa_{t,0}(\mathcal{D}_\mu^j)))^{-N}$ . Take  $(\tilde{x}, \tilde{\xi}) \in \kappa_{t,0}(\mathcal{D}_\mu^j)$  so that

$$|(x, \xi) - (\tilde{x}, \tilde{\xi})| \leq 2d(x, \xi; \kappa_{t,0}(\mathcal{D}_\mu^j))$$

and by a multinomial expansion the symbol rewrites as

$$2^{-j(|\alpha|+|\beta|)} \sum_{\gamma_1 \leq \alpha, \gamma_2 \leq \beta} c_{\gamma_1, \gamma_2} (x - \tilde{x})^{\gamma_1} (\tilde{x} - x_{t,0}(\mu))^{\alpha - \gamma_1} (\xi - \tilde{\xi})^{\gamma_2} (\tilde{\xi} - \xi_{t,0}(\mu))^{\beta - \gamma_2} b_\mu^j,$$

for some coefficients  $c_{\gamma_1, \gamma_2}$ . Since  $|\tilde{x} - x_{t,0}(\mu)|, |\tilde{\xi} - \xi_{t,0}(\mu)| = O(2^j)$ , and  $2^{-j} \leq h^{-1/2}$  the absolute value of the symbol is dominated by

$$\begin{aligned} & \sum_{\gamma_1 \leq \alpha, \gamma_2 \leq \beta} 2^{-j(|\gamma_1| + |\gamma_2|)} |x - \tilde{x}|^{|\gamma_1|} |\xi - \tilde{\xi}|^{|\gamma_2|} \left(1 + h^{-\frac{1}{2}} d(x, \xi; \kappa_{t,0}(\mathcal{D}_\mu^j))\right)^{-N-|\alpha|-|\beta|} \\ & \lesssim \left(1 + h^{-\frac{1}{2}} d(x, \xi; \kappa_{t,0}(\mathcal{D}_\mu^j))\right)^{-N}. \end{aligned}$$

Derivatives of the symbol of  $B_{\mu, \alpha, \beta}^j$  are handled similarly as they will be a sum of terms involving derivatives of  $b_\mu^j$  and monomials as above with smaller powers.

We now apply Lemma 4.1, and reindex the sum formed by (4.8) terms by  $\omega, \omega'$ , so that it writes as a sum

$$\sum_{\omega, \omega'} T_{\mu, \omega}^j (B_{\mu, \omega}^j f_h) T_{\mu', \omega'}^j (B_{\mu', \omega'}^j f_h),$$

where  $T_{\mu, \omega}^j$  is of the form (3.26), (3.27).

Fix  $\omega, \omega'$ , and set  $G_{\mu, \mu'} = T_{\mu, \omega}^j (B_{\mu, \omega}^j f_h) T_{\mu', \omega'}^j (B_{\mu', \omega'}^j f_h)$ . We claim that

$$\left\| \sum_{\mu, \mu' \in \Xi_j} Q_{1, \mu}^j Q_{2, \mu}^j (G_{\mu, \mu'}) \right\|_{L^p} \lesssim \left( \sum_{\mu, \mu' \in \Xi_j} \|G_{\mu, \mu'}\|_{L^p}^{p^*} \right)^{\frac{1}{p^*}}, \quad (4.9)$$

Observe that when  $p = 1$  or  $p = \infty$ , (4.9) follows from the fact that  $Q_{1, \mu}^j(t)$  is multiplication by a bounded function and the convolution kernel of  $Q_{2, \mu}^j(t)$  is a function with uniform  $L^1$  norm. By interpolation, it suffices to prove (4.9) when  $p = 2$ . Moreover, it is sufficient to prove it for  $t$  fixed.

We thus write the left hand side of (4.9) as

$$\sum_{\substack{\mu, \mu' \in \Xi_j \\ \tilde{\mu}, \tilde{\mu}' \in \Xi_j}} \int_{\mathbb{R}^{d-1}} Q_{1, \mu}^j(t) Q_{2, \mu}^j(t) (G_{\mu, \mu'}(t, \cdot)) \overline{Q_{1, \tilde{\mu}}(t) Q_{2, \tilde{\mu}}(t) (G_{\tilde{\mu}, \tilde{\mu}'}(t, \cdot))}.$$

Given (3.30), the remainder of the proof for  $p = 2$  now follows by showing that the absolute value of each integral here is bounded by

$$(1 + 2^{-2j} |\mu - \tilde{\mu}|^2)^{-2d} \|G_{\mu, \mu'}(t, \cdot)\|_{L^2} \|G_{\tilde{\mu}, \tilde{\mu}'}(t, \cdot)\|_{L^2}. \quad (4.10)$$

Begin by writing  $\mu = (\mu_x, \mu_\xi) \in \mathbb{R}^{d-1} \times \mathbb{R}^{d-1}$  and consider cases

$$|\mu_x - \tilde{\mu}_x| \geq 8|\mu_\xi - \tilde{\mu}_\xi| \text{ and } |\mu_x - \tilde{\mu}_x| < 8|\mu_\xi - \tilde{\mu}_\xi|.$$

In the first case, we have  $|x_{t,0}(\mu) - x_{t,0}(\mu')| \approx |\mu - \mu'|$  which shows that

$$(1 + 2^{-2j} |x - x_{t,0}(\mu)|^2)^{-2d} (1 + 2^{-2j} |x - x_{t,0}(\tilde{\mu})|^2)^{-2d} \lesssim (1 + 2^{-2j} |\mu - \tilde{\mu}|^2)^{-2d},$$

and hence (4.10) follows. In the second case,  $|\xi_{t,0}(\mu) - \xi_{t,0}(\mu')| \approx |\mu - \mu'|$  instead. We thus use the Plancherel identity for the semiclassical Fourier transform and simple convolution estimates to see that  $Q_{1, \mu}^j(t) Q_{2, \mu}^j(t) (G_{\mu, \mu'}(t, \cdot))$  is sufficiently concentrated in a  $2^{-j}$  neighborhood of  $\xi_{t,0}(\mu)$ .  $\square$

## 5. Proof of Theorem 3.4

The bound (3.28) asserts that the phase space separation of  $B_{\mu, \omega}^j(s)f, B_{\mu', \omega'}^j(s')g$  yields a gain of  $(h2^{-2j})^{\frac{d+1}{q} - \frac{d-1}{2}}$  over what would be obtained by a formal application of the linear bounds (2.13). The only difficulty with linear bounds is that the symbols  $a_{\mu, \omega}^j$  from (3.27) are not in  $S_0(1)$  since there are losses when they are differentiated in  $(x, \eta)$ . However, we will see that the decay of the symbol in  $x, \xi$  ensures that linear bounds for these symbols are valid nonetheless. In particular, linear bounds will apply to the case  $h \approx 2^{2j}$ , where there is insufficient phase space separation to apply bilinear bounds.

Our first task is to reduce matters to working with compactly supported symbols instead of the rapidly decaying symbols  $a_{\mu, \omega}^j$ . This step is not necessary when  $h \approx 2^{2j}$ , so the discussion from here through Lemma 5.2 will assume that  $J+1 \leq j \leq 0$ . Recalling from (3.6) that

$$d(\mathcal{D}_\mu^j, \mathcal{D}_{\mu'}^j) \approx 2^j, \quad J+1 \leq j \leq 0$$

we may take smooth bump functions  $c_\mu^j(x, \xi)$  and  $c_{\mu'}^j(x, \xi)$  such that

$$d(\text{supp}(1 - c_\mu^j), \mathcal{D}_\mu^j) \geq \delta 2^j$$

for some sufficiently small, but uniform constant  $\delta > 0$  while

$$d(\text{supp}(c_\mu^j), \text{supp}(c_{\mu'}^j)) \approx 2^j. \quad (5.1)$$

Now extend the  $c_\mu^j$  to a family  $\{c_\mu^j(s, \cdot)\}_{s \in [-\varepsilon, \varepsilon]}$  by defining

$$c_\mu^j(s, x, \xi) := c_\mu^j(\kappa_{0,s}(x, \xi)) = c_\mu^j((\kappa_{s,0})^{-1}(x, \xi))$$

Taking  $\varepsilon > 0$  sufficiently small in (2.1), we may assume that for every  $s \in [-\varepsilon, \varepsilon]$

$$d(\text{supp}(1 - c_\mu^j)(s, \cdot), \kappa_{s,0}(\mathcal{D}_\mu^j)) \geq \delta 2^j, \quad (5.2)$$

$$d(\text{supp}(c_\mu^j(s, \cdot)), \text{supp}(c_{\mu'}^j(s, \cdot))) \approx 2^j. \quad (5.3)$$

Now let  $C_\mu^j(s)$  denote the  $h$ -pseudodifferential operator with symbol  $c_\mu^j(s, \cdot)$ .

We will first show that linear estimates can be applied to

$$\begin{aligned} & h^{-\frac{2d}{q} + (d-1)} 2^{-j(d-1 - \frac{2(d+1)}{q})} \\ & \times \left\| R \left( B_{\mu, \omega}^j(s)f, B_{\mu', \omega'}^j(s')g \right) - R \left( (C_\mu^j(s) \circ B_{\mu, \omega}^j(s))f, (C_{\mu'}^j(s') \circ B_{\mu', \omega'}^j(s'))g \right) \right\|_{L^{\frac{q}{2}}} \end{aligned} \quad (5.4)$$

so that it is bounded by the last three terms on the right hand side of (3.28). In particular, this process will lead us to define

$$\tilde{B}_{\mu, \omega}^j(s) := (h2^{-2j})^{-\frac{d+1}{q} + \frac{d-1}{2}} (I - C_\mu^j(s)) \circ B_{\mu, \omega}^j(s). \quad (5.5)$$

After applying linear bounds, we will be reduced to

$$\begin{aligned} & \left\| R \left( (C_\mu^j(s) \circ B_{\mu,\omega}^j(s)) f, (C_{\mu'}^j(s') \circ B_{\mu,\omega}^j(s')) g \right) \right\|_{L^{\frac{q}{2}}([-\varepsilon, \varepsilon]_t \times \mathbb{R}_x^{d-1})} \\ & \lesssim h^{\frac{2d}{q} - (d-1)} 2^{j(d-1 - \frac{2(d+1)}{q})} \|B_{\mu,\omega}^j(s) f\|_{L^2(\mathbb{R}^{d-1})} \|B_{\mu',\omega'}^j(s') g\|_{L^2(\mathbb{R}^{d-1})}, \end{aligned} \quad (5.6)$$

which will follow from bilinear estimates of S. Lee [12], using the separation in (5.3).

**Lemma 5.1.** *The composition  $(I - C_\mu^j(s)) \circ B_{\mu,\omega}^j(s)$  is an  $h$ -PDO with whose symbol  $q(s, x, \xi)$  satisfies*

$$|\partial_{x,\xi}^\alpha q(s, x, \xi)| \lesssim_{\alpha, N} (h 2^{-2j}) h^{-\frac{|\alpha|}{2}} (1 + h^{-\frac{1}{2}} d(x, \xi; \kappa_{s,0}(\mathcal{D}_\mu^j)))^{-N}.$$

The lemma thus ensures that  $\tilde{B}_{\mu,\omega}^j(s)$  as defined in (5.5) satisfies (3.11) since  $h 2^{-2j} \leq (h 2^{-2j})^{\frac{d+1}{q} - \frac{d-1}{2}}$  for  $q \geq 2$ .

*Proof.* Observe that

$$q(s, x, \xi) = \frac{1}{(2\pi h)^{d-1}} \int_{\mathbb{R}^{2(d-1)}} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle} (1 - c_\mu^j(s, x, \eta)) b_{\mu,\omega}^j(s, y, \xi) dy d\eta,$$

and consider more general oscillatory integrals of the form

$$\frac{1}{(2\pi h)^{d-1}} \int_{\mathbb{R}^{2(d-1)}} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle} A(x, \eta, y, \xi) dy d\eta. \quad (5.7)$$

where  $\text{supp}(A(\cdot, y, \xi)) \subset \text{supp}((1 - c_\mu^j)(s, \cdot))$  for every  $(y, \xi)$  and

$$|\partial^\alpha A(x, \eta, y, \xi)| \lesssim_{\alpha, N} h^{-\frac{|\alpha|}{2}} (1 + h^{-\frac{1}{2}} d(y, \xi; \kappa_{s,0}(\mathcal{D}_\mu^j)))^{-N}$$

with implicit constants depending only on those defining  $c_\mu^j$ ,  $b_\mu^j$  and their derivatives. The main idea is that through induction,  $(h^{\frac{1}{2}} \partial_{x,\xi})^\alpha q(s, x, \xi)$  is an oscillatory integral of this type, which follows easily by integration by parts since

$$\partial_{x_i} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle} = -\partial_{y_i} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle}, \quad \partial_{\xi_i} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle} = -\partial_{\eta_i} e^{\frac{i}{h} \langle x-y, \eta-\xi \rangle}.$$

Hence we are left to show that

$$|(5.7)| \lesssim h 2^{-2j} (1 + h^{-\frac{1}{2}} d(x, \xi; \kappa_{s,0}(\mathcal{D}_\mu^j)))^{-N}. \quad (5.8)$$

The differential operator

$$\frac{1 + (h^{-\frac{1}{2}}(\xi - \eta, x - y)) \cdot (h^{\frac{1}{2}} D_{(y,\eta)})}{1 + h^{-1} |(\xi - \eta, x - y)|^2}$$

preserves the phase function in (5.7), so integration by parts shows that

$$\begin{aligned} |(5.7)| & \lesssim_N \\ & \frac{1}{(2\pi h)^{d-1}} \int (1 + h^{-\frac{1}{2}} |(x-y, \xi - \eta)|)^{-3N} \left( 1 + h^{-\frac{1}{2}} d(y, \xi; \kappa_{s,0}(\mathcal{D}_\mu^j)) \right)^{-2N} dy d\eta \end{aligned}$$

and the domain of integration can be restricted to  $\eta \in \text{supp}((1 - c_\mu^j)(s, x, \cdot))$ . The first factor in the integrand allows us to replace the pair  $(y, \xi)$  in the second by either  $(x, \eta)$  or  $(x, \xi)$ . In particular, when  $N \geq 2$ , the integrand is dominated by

$$(1 + h^{-\frac{1}{2}} |(x - y, \xi - \eta)|)^{-N} (h^{-\frac{1}{2}} d(x, \eta; \kappa_{s,0}(\mathcal{D}_\mu^j)))^{-2} \left(1 + h^{-\frac{1}{2}} d(x, \xi; \kappa_{s,0}(\mathcal{D}_\mu^j))\right)^{-N}.$$

Since  $(x, \eta) \in \text{supp}((1 - c_\mu^j)(s, \cdot))$ , (5.2) shows that the second factor here is  $O(h^{2-2j})$ . Since the first factor means that integration in  $y, \eta$  yields a gain of  $O(h^{d-1})$ , the bound (5.8) follows.  $\square$

**Lemma 5.2.** *Let  $\tilde{c}_\mu^j(t, s, x, \xi) = c_\mu^j(s, d_\eta \phi(t, s, x, \xi), \xi)$ . There exists symbols  $q_{\mu,\omega}^j$  and  $r_{\mu,\omega}^j$  such that*

$$T_{\mu,\omega}^{j,s}(C_\mu^j(s)f) = \frac{1}{(2\pi h)^{d-1}} \iint e^{\frac{i}{h}(\phi(t,s,x,\xi) - \langle y, \xi \rangle)} (q_{\mu,\omega}^j + r_{\mu,\omega}^j)(t, s, x, \xi) f(y) dy d\xi,$$

with  $|\partial_{x,\xi}^\alpha q_{\mu,\omega}^j(t, s, \cdot)| \lesssim_\alpha 2^{-j|\alpha|}$  and

$$\text{supp}(q_{\mu,\omega}^j(t, s, \cdot)) \subset \text{supp}(\tilde{c}_\mu^j(t, s, \cdot)) = \overline{\{(x, \xi) : c_\mu^j(s, d_\eta \phi(t, s, x, \xi), \xi) \neq 0\}} \quad (5.9)$$

while

$$|\partial_{x,\xi}^\alpha r_{\mu,\omega}^j(t, s, \cdot)| \lesssim_{\alpha, N} (h 2^{-2j}) 2^{-j|\alpha|} \left(1 + h^{-1} 2^j d(x, \xi; \text{supp}(\tilde{c}_\mu^j(t, s, \cdot)))\right)^{-N}. \quad (5.10)$$

Again, since  $h 2^{-2j} \leq (h 2^{-2j})^{\frac{d+1}{q} - \frac{d-1}{2}}$  the gain satisfied by  $r_{\mu,\omega}^j$  will be sufficient.

*Proof.* The Schwartz kernel of the composition  $T_{\mu,\omega}^{j,s} \circ C_\mu^j(s)$  is the oscillatory integral

$$\frac{1}{(2\pi h)^{d-1}} \iint e^{\frac{i}{h}(\phi(t,s,x,\eta) + \langle y, \xi - \eta \rangle)} a_{\mu,\omega}^j(t, s, x, \eta) c_\mu^j(t, s, y, \xi) dy d\eta. \quad (5.11)$$

Its critical points satisfy  $d_\eta \phi(t, s, x, \eta) = y$  and  $\xi = \eta$ . The lemma thus follows from a variation on the stationary phase arguments in Theorem 2.1 (taking  $M = 1$  in the analog of (2.21)).  $\square$

**Lemma 5.3.** *Let  $T_\mu^j f$  be an oscillatory integral operator defined by*

$$(T_\mu^j f)(t, x) = \frac{1}{(2\pi h)^{d-1}} \iint_{\mathbb{R}^{2(d-1)}} e^{\frac{i}{h}(\phi(t,s,x,\xi) - \langle y, \xi \rangle)} a_\mu^j(t, s, x, \xi) f(y) dy d\xi,$$

where  $J \leq j \leq 0$  and

$$|\partial_{x,\xi}^\alpha a_\mu^j(t, s, \cdot)| \lesssim_\alpha 2^{-j|\alpha|} (1 + 2^{-j} |x - x_{t,0}(\mu)|)^{-2d} (1 + 2^{-j} |\xi - \xi_{s,0}(\mu)|)^{-2d}. \quad (5.12)$$

Then for some implicit constant depending only on the bounds for  $a_\mu^j$ ,  $\phi$  and their derivatives, we have

$$\|T_\mu^j f\|_{L^q([-\varepsilon, \varepsilon], \mathbb{R}^{d-1})} \lesssim h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})} \|f\|_{L^2(\mathbb{R}^{d-1})}.$$

We postpone a thorough proof of these bounds to the end of the section, though as noted, they essentially are a consequence of the standard theory of Carleson–Sjölin integral operators. For now we note that this will settle Theorem 3.4 when  $j = J$  (strictly speaking this means that we take  $\tilde{B}_{\mu,\omega}^J(s) = 0$  as the enlarged family is not needed in this case). Moreover, the linear estimates imply that (5.4) is bounded by the last three terms in (3.28). Indeed, the lemma and the definition of  $\tilde{B}_{\mu,\omega}^j(s)$  establishes that

$$\begin{aligned} & \left\| T_{\mu,\omega}^{j,s} \left( (I - C_{\mu}^j(s)) \circ B_{\mu,\omega}^j(s) f \right) \right\|_{L^q([-\varepsilon,\varepsilon],_t \times \mathbb{R}_x^{d-1})} \\ & \lesssim h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})} (h 2^{-2j})^{\frac{d+1}{q} - \frac{d-1}{2}} \|\tilde{B}_{\mu,\omega}^j(s) f\|_{L^2(\mathbb{R}_y^{d-1})}, \end{aligned}$$

which yields the desired gain when paired with the linear bound on the second factor. Moreover, the gain of  $h 2^{-2j} \leq (h 2^{-2j})^{\frac{d+1}{q} - \frac{d-1}{2}}$  in the symbol of  $r_{\mu,\omega}^j$  shown in (5.10) means that it suffices to prove (5.6) with the compactly supported  $q_{\mu,\omega}^j$  replacing  $a_{\mu,\omega}^j$  in the definition of  $T_{\mu,\omega}^{j,s}$ .

**5.1. Bilinear estimates.** We now complete the proof of (5.6). Set

$$u_1(\eta) = \frac{2^{j(d-1)}}{(2\pi h)^{(d-1)}} \mathcal{F}_h(B_{\mu,\omega}^j(s) f)(2^j \eta), \quad u_2(\eta) = \frac{2^{j(d-1)}}{(2\pi h)^{(d-1)}} \mathcal{F}_h(B_{\mu',\omega'}^j(s') g)(2^j \eta) \quad (5.13)$$

where  $\mathcal{F}_h$  denotes the semiclassical Fourier transform. Observe that by the semiclassical Plancherel identity,

$$\|u_1\|_{L^2(\mathbb{R}^{d-1})} = \left( \frac{2^j}{2\pi h} \right)^{\frac{d-1}{2}} \|B_{\mu,\omega}^j(s) f\|_{L^2(\mathbb{R}^{d-1})}, \quad (5.14)$$

and an analogous computation establishes the  $L^2$  norm of  $u_2$ .

By the discussion following Lemma 5.3, we may replace the symbol  $a_{\mu,\omega}^j$  in the definition of  $T_{\mu,\omega}^{j,s}$  by the compactly supported  $q_{\mu,\omega}^j$  defined in Lemma 5.2. Next set

$$\tilde{\phi}(t, s, x, \eta) = 2^{-2j} \phi(t, s, 2^j x, 2^j \eta), \quad \tilde{p}(t, x, \xi) = 2^{-2j} p(t, 2^j x, 2^j \xi),$$

so that  $\tilde{\phi}$  solves the eikonal equation

$$\partial_t \tilde{\phi}(t, s, x, \eta) + \tilde{p}(t, x, d_x \tilde{\phi}(t, s, x, \eta)) = 0. \quad (5.15)$$

Also, let  $\tilde{\kappa}_{t,s}(x, \xi)$  denote the time  $t$  value of the integral curve determined by the Hamiltonian vector field of  $\tilde{p}(r, x, \xi)$  such that  $\tilde{\kappa}_{t,s}(x, \xi)|_{r=s} = (x, \xi)$ . It is verified that the  $\tilde{\kappa}_{t,s}$  are related to the original mappings  $\kappa_{t,s}$  by

$$\tilde{\kappa}_{t,s}(x, \xi) = 2^{-j} \kappa_{t,s}(2^j x, 2^j \xi).$$

Hence  $\tilde{\phi}$  is a generating function for  $\tilde{\kappa}_{t,s}$ , satisfying relations analogous to (2.10). The mixed Hessian  $d_x d_{\xi} \tilde{\phi} = I + O(\varepsilon)$  is nonsingular by the chain rule and (2.11).

Now make the following definitions

$$\tilde{q}_{\mu,\omega}^j(t, s, x, \xi) := q_{\mu,\omega}^j(t, s, 2^j x, 2^j \xi), \quad \tilde{q}_{\mu',\omega'}^j(t, s, x, \xi) := q_{\mu',\omega'}^j(t, s', 2^j x, 2^j \xi),$$

$$T_1 u_1(t, x) := \int e^{\frac{i2^j}{h} \tilde{\phi}(t, s, x, \xi)} \tilde{q}_{\mu, \omega}^j(t, s, x, \xi) u_1(\xi) d\xi, \quad (5.16)$$

$$T_2 u_2(t, x) := \int e^{\frac{i2^j}{h} \tilde{\phi}(t, s', x, \eta)} \tilde{q}_{\mu', \omega'}^j(t, s', x, \eta) u_2(\eta) d\eta. \quad (5.17)$$

so that after the symbol replacement above, we have that up to acceptable error,

$$R \left( (C_\mu^j(s) \circ B_{\mu, \omega}^j(s)) f, (C_{\mu'}^j(s') \circ B_{\mu', \omega'}^j(s')) g \right) (t, 2^j x) = T_1 u_1(t, x) \cdot T_2 u_2(t, x)$$

The bound (5.6) will thus follow by dilating variables  $x \mapsto 2^j x$  and the bound

$$\|T_1 u_1 T_2 u_2\|_{L^{\frac{q}{2}}([- \varepsilon, \varepsilon], \mathbb{R}_x^{d-1})} \lesssim (h 2^{-2j})^{\frac{2d}{q}} \|u_1\|_{L^2(\mathbb{R}^{d-1})} \|u_2\|_{L^2(\mathbb{R}^{d-1})}. \quad (5.18)$$

To show (5.18), we will use bilinear estimates. When  $d = 2$ , these are implicit in the work of Hörmander [10] and when  $d \geq 3$ , they are a consequence of results of S. Lee [12, Theorem 1.1] and the epsilon removal lemma in [3, Theorem 3.3]. To align the notation here with the latter works, define

$$\delta(t, x, \xi, \eta) = (d_\xi \tilde{p})(t, x, d_x \tilde{\phi}(t, s, x, \xi)) - (d_\eta \tilde{p})(t, x, d_x \tilde{\phi}(t, s', x, \eta)).$$

Since the Hessian  $\partial_{\eta_i \eta_j}^2 \tilde{p}(t, x, \eta) = (\partial_{\eta_i \eta_j}^2 p)(t, 2^j x, 2^j \eta)$  is positive definite,

$$|\delta(t, x, \xi, \eta)| \approx |d_x \tilde{\phi}(t, s, x, \xi) - d_x \tilde{\phi}(t, s', x, \eta)|.$$

The crucial hypothesis to be verified to apply these bounds is that the following expression is uniformly bounded from below when  $(2^j x, 2^j \xi) \in \text{supp}(q_\mu^j(s, \cdot))$  and  $(2^j x, 2^j \eta) \in \text{supp}(q_{\mu'}^j(s', \cdot))$ ,

$$\left| \left\langle \left[ d_x d_\xi \tilde{\phi} \right] \delta(t, x, \xi, \eta), \left[ d_x d_\xi \tilde{\phi} \right]^{-1} \left[ d_\xi d_\xi \tilde{p}(t, x, d_x \tilde{\phi}) \right]^{-1} \delta(t, x, \xi, \eta) \right\rangle \right|,$$

where the derivatives of  $d_x \tilde{\phi}$ ,  $d_x d_\xi \tilde{\phi}$  are either both evaluated at  $(t, s, x, \xi)$  or both evaluated at  $(t, s', x, \eta)$ . Given that  $d_x d_\xi \tilde{\phi} = I + O(\varepsilon)$ , it suffices to see that

$$\left| d_x \tilde{\phi}(t, s, x, \xi) - d_x \tilde{\phi}(t, s', x, \eta) \right| \approx 1. \quad (5.19)$$

To see (5.19), note that by (5.9),  $(x, \xi) \in \text{supp}(\tilde{q}_\mu^j(s, \cdot))$  implies that

$$\left( d_\xi \phi(t, s, 2^j x, 2^j \xi), 2^j \xi \right) \in \text{supp}(c_\mu^j(s, \cdot)) = \kappa_{s,0} \left( \text{supp}(c_\mu^j(0, \cdot)) \right).$$

Applying  $\kappa_{t,s}$  to both sides of this, we see that by (2.10) this is equivalent to

$$\left( 2^j x, d_x \phi(t, s, 2^j x, 2^j \xi) \right) \in \kappa_{t,0} \left( \text{supp}(c_\mu^j(0, \cdot)) \right).$$

The same reasoning shows that if  $(2^j x, 2^j \eta) \in \text{supp}(q_{\mu'}^j(s', \cdot))$ , then

$$\left( 2^j x, d_x \phi(t, s', 2^j x, 2^j \eta) \right) \in \kappa_{t,0} \left( \text{supp}(c_{\mu'}^j(0, \cdot)) \right).$$

By (5.1), this gives  $|d_x \phi(t, s, 2^j x, 2^j \xi) - d_x \phi(t, s', 2^j x, 2^j \eta)| \approx 2^j$ , and (5.19) follows.

**5.2. Linear estimates.** Here we prove Lemma 5.3, using the rescaling from the previous Sect. 5.1. Define  $\tilde{a}_{\tilde{\mu}}^j(t, s, x, \xi) := a_{\mu}^j(t, s, 2^j x, 2^j \xi)$ , let  $u_1$  be as in (5.13) but replacing  $B_{\mu, \omega}^j(s) f$  by  $f$ , and let  $T_1$  be the operator with Schwartz kernel given by the oscillatory integral

$$\int e^{\frac{i2^j}{h}(\tilde{\phi}(t, s, x, \xi) - \langle y, \xi \rangle)} \tilde{a}_{\tilde{\mu}}^j(t, s, x, \xi) d\xi.$$

We thus have that  $\tilde{a}_{\tilde{\mu}}^j(t, s, \cdot) \in S_0(1)$ , with

$$|\partial_{x, \xi}^{\alpha} \tilde{a}_{\tilde{\mu}}^j(t, s, \cdot)| \lesssim_{\alpha} \langle x - \tilde{x}_{t, 0}(\tilde{\mu}) \rangle^{-2d} \langle \xi - \tilde{\xi}_{s, 0}(\tilde{\mu}) \rangle^{-2d}, \quad \tilde{\mu} = 2^{-j} \mu \quad (5.20)$$

writing  $\tilde{\kappa}_{r, 0}(\tilde{\mu})$  in terms of components  $(\tilde{x}_{r, 0}(\tilde{\mu}), \tilde{\xi}_{r, 0}(\tilde{\mu}))$ . It now suffices to show

$$\|T_1 u_1\|_{L^q([-\varepsilon, \varepsilon], \times \mathbb{R}_x^{d-1})} \lesssim (h 2^{-2j})^{\frac{d-1}{2}(\frac{1}{2} + \frac{1}{q})} \|u_1\|_{L^2(\mathbb{R}_y^{d-1})}. \quad (5.21)$$

Indeed, given (5.14), the right hand side is  $\approx h^{-\frac{d-1}{2}(\frac{1}{2} - \frac{1}{q})} 2^{-\frac{j(d-1)}{q}} \|f\|_{L^2}$ , and the loss in  $2^j$  cancelled by reverting back to the original coordinates on the left hand side. This is the same gain observed in the first inequality in (2.13), but with frequency  $h^{-1} 2^{2j}$ . The main observation is that  $T_1$  is an oscillatory integral operator with a Carleson–Sjölin phase as defined in [14, Chap. 2]. Indeed,

$$d_x d_{\xi} \tilde{\phi}(t, s, x, \xi) = (d_x d_{\xi} \phi)(t, s, 2^j x, 2^j \xi), \quad (5.22)$$

and the latter defines a nonsingular  $(d-1) \times (d-1)$  matrix. Recalling (5.15), it suffices to check that for  $(t, x) \in \text{supp}(\tilde{q}_{\mu}^j(\cdot, s, \cdot, \xi))$  fixed

$$\left\{ (-\tilde{p}(t, x, d_x \tilde{\phi}(t, s, x, \xi)), d_x \tilde{\phi}(t, s, x, \xi)) : \xi \in \text{supp}(\tilde{q}_{\mu}^j(t, s, x, \cdot)) \right\} \quad (5.23)$$

is an embedded hypersurface in  $T_{(t, x)}^* \mathbb{R}^d$  with nonvanishing principal curvatures. But this follows from the fact that this is a local reparameterization of the graph of  $\eta \mapsto -\tilde{p}(t, x, \eta)$  and  $\partial_{\eta_i \eta_j}^2 \tilde{p}(t, x, \eta) = (\partial_{\eta_i \eta_j}^2 p)(t, 2^j x, 2^j \eta)$ .

At this point, (5.21) formally follows from the standard theory of Carleson–Sjölin integrals due to Hörmander and Stein (see e.g. [14, Theorem 2.2.1]). However, care must be taken here since the amplitude  $\tilde{a}_{\tilde{\mu}}^j$  is not compactly supported in a set of uniform diameter. However, it does decay rapidly outside such a set, which is enough. Seeing this is a matter of tracing through the proof in that theorem, which we do here. We first observe that given (5.22), the bound (5.21) with  $q = 2$  is a consequence of Hörmander's  $L^2$  theorem [10] and the decay in (5.20). Indeed, this decay allows us to write the amplitude as a sum

$$\tilde{a}_{\tilde{\mu}}^j(t, s, x, \xi) = \sum_{\tilde{v} = (\tilde{v}_x, \tilde{v}_{\xi}) \in \mathbb{Z}^{2(d-1)}} \langle \tilde{v}_x - \tilde{x}_{t, 0}(\tilde{\mu}) \rangle^{-2d} \langle \tilde{v}_{\xi} - \tilde{\xi}_{s, 0}(\tilde{\mu}) \rangle^{-2d} \tilde{a}_{\tilde{\mu}, \tilde{v}}^j(t, s, x, \xi),$$

and we may take  $\tilde{a}_{\tilde{\mu}, \tilde{v}}^j(t, s, \cdot) \in S_0(1)$  with  $\text{supp}(\tilde{a}_{\tilde{\mu}, \tilde{v}}^j(t, s, \cdot)) \subset \{|(x, \xi) - \tilde{v}| \lesssim 1\}$ . For fixed  $t$ , Hörmander's theorem can now be applied in  $L^2(\mathbb{R}^{d-1})$  to each operator

determined by the decomposition here, and the decay in  $\tilde{v}$  ensures that we may sum over the contribution of each of these operators without loss. This results in

$$\|T_1\|_{L^2(\mathbb{R}^{d-1}) \rightarrow L^2([- \varepsilon, \varepsilon] \times \mathbb{R}^{d-1})} = O((h2^{-2j})^{\frac{d-1}{2}}).$$

It now suffices to prove (5.21) with  $q = \frac{2(d+1)}{d-1}$ , that is,

$$\|T_1\|_{L^2(\mathbb{R}^{d-1}) \rightarrow L^{\frac{2(d+1)}{d-1}}([- \varepsilon, \varepsilon] \times \mathbb{R}^{d-1})} = O((h2^{-2j})^{\frac{d(d-1)}{2(d+1)}}).$$

Defining the fixed  $t$  operator  $(T_1^t u_1)(x) = (T_1 u_1)(t, x)$ , a duality and fractional integration argument along with the aforementioned  $L^2$  bounds means it suffices to show that

$$\|T_1^t (\tilde{T}_1^{\tilde{t}})^*\|_{L^1(\mathbb{R}^{d-1}) \rightarrow L^\infty(\mathbb{R}^{d-1})} \lesssim (h2^{-2j}|t - \tilde{t}|)^{-\frac{d-1}{2}}$$

which is a consequence of

$$\left| \int e^{\frac{i2^{2j}}{h}(\tilde{\phi}(t, s, x, \xi) - \tilde{\phi}(\tilde{t}, s, \tilde{x}, \xi))} \tilde{a}_{\tilde{\mu}}^j(t, s, x, \xi) \overline{\tilde{a}_{\tilde{\mu}}^j(\tilde{t}, s, \tilde{x}, \xi)} d\xi \right| \lesssim (h2^{-2j}|t - \tilde{t}|)^{-\frac{d-1}{2}}.$$

As in [14, Chap. 2], this in turn follows from the fact that (5.23) ensures that any critical points in the phase are nondegenerate. As before, this requires some care since the integration is not over a set of uniform size, but this can be overcome by using a partition of unity adapted to cubes of sidelength 1 in  $\xi$ , then using the decay in  $\xi$  away from  $\tilde{\xi}_{s,0}(\tilde{\mu})$ .

*Acknowledgements.* The authors are grateful to Steve Zelditch for several helpful conversations and to the anonymous referee for numerous thoughtful suggestions which improved the exposition in this work.

## References

1. Abraham, R., Marsden, J.E.: Foundations of Mechanics, 2nd edn. Benjamin/Cummings Publishing Co., Reading, Mass. (1978)
2. Bérard, P.: On the wave equation on a compact Riemannian manifold without conjugate points. Math. Z. **155**(3), 249–276 (1977)
3. 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. **17**, 2513–2543 (2015)
4. Blair, M.D., Sogge, C.D.: Refined and microlocal Kakeya–Nikodym bounds of eigenfunctions in two dimensions. Anal. PDE **8**, 747–764 (2015)
5. Blair, M.D., Sogge, C.D.: Toponogov’s theorem and improved estimates of eigenfunctions. J. Differ. Geom. (2015) (to appear in print 2017)
6. Bourgain, J.: Geodesic restrictions and  $L^p$ -estimates for eigenfunctions of Riemannian surfaces. In: Linear and Complex Analysis, Amer. Math. Soc. Transl. 226, pp. 27–35. Amer. Math. Soc., Providence (2009)
7. Christianson, H.: Quantum monodromy and nonconcentration near a closed semi-hyperbolic orbit. Trans. Am. Math. Soc. **363**(7), 3373–3438 (2011)
8. Hassell, A., Tacy, M.: Improvement of eigenfunction estimates on manifolds of nonpositive curvature. Forum Math. **27**(3), 1435–1451 (2015)
9. Hezari, H., Rivière, G.:  $L^p$  norms, nodal sets, and quantum ergodicity. Adv. Math. **290**, 938–966 (2016)
10. Hörmander, L.: Oscillatory integrals and multipliers on  $FL^p$ . Ark. Math. **11**, 1–11 (1973)
11. Koch, H., Tataru, D., Zworski, M.: Semiclassical  $L^p$  estimates. Ann. Henri Poincaré **8**, 885–916 (2007)
12. Lee, S.: Linear and bilinear estimates for oscillatory integral operators related to restriction to hypersurfaces. J. Funct. Anal. **241**, 56–98 (2006)

13. Sogge, C.D.: Concerning the  $L^p$  norm of spectral clusters for second-order elliptic operators on compact manifolds. *J. Funct. Anal.* **77**, 123–138 (1988)
14. Sogge, C.D.: Fourier Integrals in Classical Analysis. Cambridge Tracts in Math. Cambridge University Press, Cambridge (1993)
15. Sogge, C.D.: Problems related to the concentration of eigenfunctions. Preprint
16. Sogge, C.D.: Kakeya–Nikodym averages and  $L^p$ -norms of eigenfunctions. *Tohoku Math. J.* **63**, 519–538 (2011)
17. Sogge, C.D., Toth, J., Zelditch, S.: About the blowup of quasimodes on Riemannian manifolds. *J. Geom. Anal.* **21**(1), 150–173 (2011)
18. Sogge, C.D., Zelditch, S.: Riemannian manifolds with maximal eigenfunction growth. *Duke Math. J.* **114**, 387–437 (2002)
19. Sogge, C.D., Zelditch, S.: Focal points and sup-norms of eigenfunctions. *Rev. Mat. Iberoam.* **32**(3), 971–994 (2016)
20. Tao, T.: A sharp bilinear restriction estimates for paraboloids. *Geom. Funct. Anal.* **13**, 1359–1384 (2003)
21. Tao, T., Vargas, A., Vega, L.: A bilinear approach to the restriction and Kakeya conjectures. *J. Am. Math. Soc.* **11**, 967–1000 (1998)
22. Taylor, M.E.: Pseudodifferential Operators. Princeton Mathematical Series. Princeton University Press, Princeton (1981)
23. Wolff, T.: A sharp cone restriction estimate. *Ann. Math.* **153**, 661–698 (2001)
24. Zworski, M.: Semiclassical Analysis. Graduate Studies in Mathematics, vol. 138. Amer. Math. Soc., Providence (2012)

Communicated by J. Marklof