# GROWTH OF HIGH $L^p$ NORMS FOR EIGENFUNCTIONS: AN APPLICATION OF GEODESIC BEAMS

YAIZA CANZANI AND JEFFREY GALKOWSKI

ABSTRACT. This work concerns  $L^p$  norms of high energy Laplace eigenfunctions,  $(-\Delta_g - \lambda^2)\phi_{\lambda} = 0$ ,  $\|\phi_{\lambda}\|_{L^2} = 1$ . In [Sog88], Sogge gave optimal estimates on the growth of  $\|\phi_{\lambda}\|_{L^p}$  for a general compact Riemannian manifold. The goal of this article is to give general dynamical conditions guaranteeing quantitative improvements in  $L^p$  estimates for  $p > p_c$ , where  $p_c$  is the critical exponent. We also apply the results of [CG19a] to obtain quantitative improvements in concrete geometric settings including all product manifolds. These are the first results improving estimates for the  $L^p$  growth of eigenfunctions that only require dynamical assumptions. In contrast with previous improvements, our assumptions are local in the sense that they depend only on the geodesics passing through a shrinking neighborhood of a given set in M. Moreover, the article gives a structure theorem for eigenfunctions which saturate the quantitatively improved  $L^p$  bound. Modulo an error, the theorem describes these eigenfunctions as finite sums of quasimodes which, roughly, approximate zonal harmonics on the sphere scaled by  $1/\sqrt{\log \lambda}$ .

#### 1. INTRODUCTION

Let (M, g) be a smooth, compact, Riemannian manifold of dimension n and consider normalized Laplace eigenfunctions: solutions to

$$(-\Delta_g - \lambda^2)\phi_\lambda = 0, \qquad \|\phi_\lambda\|_{L^2(M)} = 1.$$

This article studies the growth of  $L^p$  norms of the eigenfunctions,  $\phi_{\lambda}$ , as  $\lambda \to \infty$ . Since the work of Sogge [Sog88], it has been known that there is a change of behavior in the growth of  $L^p$  norms for eigenfunctions at the *critical exponent*  $p_c := \frac{2(n+1)}{n-1}$ . In particular,

$$\|\phi_{\lambda}\|_{L^{p}(M)} \leq C\lambda^{\delta(p)}, \qquad \delta(p) := \begin{cases} \frac{n-1}{2} - \frac{n}{p} & p_{c} \leq p\\ \frac{n-1}{4} - \frac{n-1}{2p} & 2 \leq p \leq p_{c}. \end{cases}$$
(1.1)

For  $p \ge p_c$ , (1.1) is saturated by the zonal harmonics on the round sphere  $S^n$ . On the other hand, for  $p \le p_c$ , these bounds are saturated by the highest weight spherical harmonics on  $S^n$ , also known as Gaussian beams. In a very strong sense, the authors showed in [CG20, page 4] that any eigenfunction saturating (1.1) for  $p > p_c$  behaves like a zonal harmonic, while Blair– Sogge [BS15a, BS17] showed that for  $p < p_c$  such eigenfunctions behave like Gaussian beams. In the  $p \le p_c$ , Blair–Sogge have recently made substantial progress on improved  $L^p$  estimates on manifolds with non-positive curvature [BS19, BS18, BS15b]

This article concerns the behavior of  $L^p$  norms for high p; that is, for  $p > p_c$ . While there has been a great deal of work on  $L^p$  norms of eigenfunctions [KTZ07, HR16, Tac19, Tac18, STZ11, SZ02, SZ16, TZ02, TZ03] this article departs from the now standard approaches. We both adapt the geodesic beam methods developed by the authors in [GT17, Gal19, Gal18, CGT18, CG19b, GT20, CG19a, CG20] and develop a new second microlocal calculus used to understand the number of points at which  $|u_{\lambda}|$  can be large. By doing this we give general dynamical conditions guaranteeing quantitative improvements over (1.1) for  $p > p_c$ . In order to work in compact subsets of phase space, we semiclassically rescale our problem. Let  $h = \lambda^{-1}$  and, abusing notation slightly, write  $\phi_{\lambda} = \phi_h$  so that

$$(-h^2\Delta_g - 1)\phi_h = 0, \qquad \|\phi_h\|_{L^2(M)} = 1.$$

We also work with the semiclassical Sobolev spaces  $H^s_{scl}(M)$ ,  $s \in \mathbb{R}$ , defined by the norm

$$||u||_{H^s_{\rm scl}(M)}^2 := \langle (-h^2 \Delta_g + 1)^s u, u \rangle_{L^2(M)}.$$

We start by stating a consequence of our main theorem. Let  $\Xi$  denote the collection of maximal unit speed geodesics for (M, g). For m a positive integer,  $r > 0, t \in \mathbb{R}$ , and  $x \in M$  define

 $\Xi_x^{m,r,t} := \big\{ \gamma \in \Xi : \gamma(0) = x, \, \exists \text{ at least } m \text{ conjugate points to } x \text{ in } \gamma(t-r,t+r) \big\},$ 

where we count conjugate points with multiplicity. Next, for a set  $V \subset M$  write

$$\mathcal{C}_{\!_{V}}^{m,r,t}:=\bigcup_{x\in V}\{\gamma(t):\gamma\in\Xi_{\!x}^{m,r,t}\}$$

Note that if  $r_t \to 0^+$  as  $|t| \to \infty$ , then saying  $y \in C_x^{n-1,r_t,t}$  for t large indicates that y behaves like a point that is maximally conjugate to x. This is the case for every point x on the sphere when y is either equal to x or its antipodal point. The following result applies under the assumption that points are not maximally conjugate and obtains quantitative improvements.

**Theorem 1.** Let  $p > p_c$ ,  $U \subset M$ , and assume there exist  $t_0 > 0$  and a > 0 so

$$\inf_{c_1, x_2 \in U} d(x_1, \mathcal{C}_{x_2}^{n-1, r_t, t}) \ge r_t, \qquad \text{for } t \ge t_0,$$

with  $r_t = \frac{1}{a}e^{-at}$ . Then, there exist C > 0 and  $h_0 > 0$  so that for  $0 < h < h_0$  and  $u \in \mathcal{D}'(M)$ 

$$\|u\|_{L^{p}(U)} \leq Ch^{-\delta(p)} \left( \frac{\|u\|_{L^{2}(M)}}{\sqrt{\log h^{-1}}} + \frac{\sqrt{\log h^{-1}}}{h} \|(-h^{2}\Delta_{g}-1)u\|_{H^{\frac{n-3}{2}-\frac{n}{p}}(M)} \right).$$

The assumption in Theorem 1 rules out maximal conjugacy of any two points  $x, y \in U$  uniformly up to time  $\infty$ , and we expect it to hold on a generic manifold M with U = M. Since Theorem 1 includes the case of manifolds without conjugate points, it generalizes the work of [HT15], where it was shown that logarithmic improvements in  $L^p$  norms for  $p > p_c$  are possible on manifolds with non-positive curvature. One family of examples where the assumptions of Theorem 1 hold is that of product manifolds [CG20, Lemma 1.1] i.e.  $(M_1 \times M_2, g_1 \oplus g_2)$  where  $(M_i, g_i)$  are non-trivial compact Riemannian manifolds. Note that this family of examples includes manifolds with large numbers of conjugate points e.g.  $S^2 \times M$ .

The proof of Theorem 1 gives a great deal of information about eigenfunctions which may saturate  $L^p$  bounds  $(p > p_c)$ . Our next theorem describes the structure of such eigenfunctions. This theorem shows that an eigenfunction can saturate the *logarithmically improved*  $L^p$  norm near at most *boundedly many* points. Moreover, modulo an error small in  $L^p$ , near each of these points the eigenfunction can be decomposed as a sum of quasimodes which are similar to the highest weight spherical harmonics scaled by  $h^{\frac{n-1}{4}}/\sqrt{\log h^{-1}}$  whose number is nearly proportional to  $h^{\frac{1-n}{2}}$ . In the theorem below the quasimodes are denoted by  $v_j$  and, while similar to highest weight spherical harmonics (a.k.a Gaussian beams), they are not as tightly localized to a geodesic segment and do not have Gaussian profiles. We refer to these quasimodes as geodesic beams (see Remark 3). **Theorem 2.** Let  $p > p_c$ . There exist c, C > 0 such that the following holds. Suppose the same assumptions as Theorem 1. Let  $0 < \delta_1 < \delta_2 < \frac{1}{2}$ ,  $h^{\delta_2} \leq R(h) \leq h^{\delta_1}$ , and  $\{x_{\alpha}\}_{\alpha \in \mathcal{I}(h)} \subset M$  be a maximal R(h)-separated set. Let  $u \in \mathcal{D}'(M)$  with  $\|(-h^2\Delta_g - 1)u\|_{H_h^{\frac{n-3}{2}}} = o(\frac{h}{\log h^{-1}}\|u\|_{L^2})$ , and for  $\varepsilon > 0$ 

$$\mathcal{S}_{U}(h,\varepsilon,u) := \Big\{ \alpha \in \mathcal{I}(h) : \|u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \ge \frac{\varepsilon h^{\frac{1-n}{2}}}{\sqrt{\log h^{-1}}} \|u\|_{L^{2}(M)}, \ B(x_{\alpha},R(h)) \cap U \neq \emptyset \Big\}.$$

Then, for all  $\varepsilon > 0$  there are  $N_{\varepsilon} > 0$  and  $h_0 > 0$  such that  $|\mathcal{S}_U(h, \varepsilon, u)| \le N_{\varepsilon}$  for all  $0 < h \le h_0$ .

Moreover, there is collection of geodesic tubes  $\{\mathcal{T}_j\}_{j \in \mathcal{L}(\varepsilon, u)}$  of radius R(h) (see Definition 1), with indices satisfying  $\mathcal{L}(\varepsilon, u) = \bigcup_{i=1}^C \mathcal{J}_i$  and  $\mathcal{T}_k \cap \mathcal{T}_\ell = \emptyset$  for  $k, \ell \in \mathcal{J}_i$  with  $k \neq \ell$ , such that

$$u = u_e + \frac{1}{\sqrt{\log h^{-1}}} \sum_{j \in \mathcal{L}(\varepsilon, u)} v_j$$

where  $v_j$  is microsupported in  $\mathcal{T}_j$ ,  $|\mathcal{L}(\varepsilon, u)| \leq C \varepsilon^{-2} R(h)^{1-n}$ , and for all  $p \leq q \leq \infty$ ,

$$\|u_e\|_{L^q} \le \varepsilon h^{-\delta(q)} (\log h^{-1})^{-\frac{1}{2}} \|u\|_{L^2},$$
  
$$\|v_j\|_{L^2} \le C\varepsilon^{-1} R(h)^{\frac{n-1}{2}} \|u\|_{L^2}, \qquad \|Pv_j\|_{L^2} \le C\varepsilon^{-1} R(h)^{\frac{n-1}{2}} h \|u\|_{L^2}.$$

Finally, with  $\mathcal{L}(\varepsilon, u, \alpha) := \{ j \in \mathcal{L}(\varepsilon, u) : \pi(\mathcal{T}_j) \cap B(x_\alpha, 3R(h)) \neq \emptyset \}$ , for every  $\alpha \in \mathcal{S}_u(h, \varepsilon, u)$ ,

$$c\varepsilon^2 R(h)^{1-n} \le |\mathcal{L}(\varepsilon, u, \alpha)| \le CR(h)^{1-n}, \qquad \sum_{j \in \mathcal{L}(\varepsilon, u, \alpha)} \|v_j\|_{L^2}^2 \ge c^2 \varepsilon^2.$$

The decomposition of u into geodesic beams  $v_j$  is illustrated in Figure 1. One covers  $S^*M$  with a collection of tubes  $\{\mathcal{T}_j\}$  of radius R(h) that run along a geodesic. Each geodesic beam  $v_j$  corresponds to microlocalizing u to the tube  $\mathcal{T}_j$ .

Let u be a quasimode with  $||Pu|| = o(h/\log h^{-1})||u||$ . Note that, by interpolation Theorem 1 implies that for each  $p > p_c$  there is N > 0 such that if  $||u||_{L^p} \ge \varepsilon h^{-\delta(p)}/\sqrt{\log h^{-1}}||u||_{L^2}$ , then  $||u||_{L^{\infty}} \ge \varepsilon^N h^{\frac{1-n}{2}} ||u||_{L^2}$ . In particular, for u to saturate the logarithmically improved  $L^p$  bound, it follows that  $\mathcal{S}_M(h, \varepsilon^N, u)$  is non-empty. Theorem 2 then gives that  $\mathcal{S}_M(h, \varepsilon^N, u)$  has a uniformly bounded number of points and at these points the quasimode u needs to consist of at least  $c\varepsilon^{2N}R(h)^{1-n}$  geodesic beams whose combined  $L^2$  mass is at least  $c\varepsilon^N$ . Since  $\dim(S^*_{x_{\alpha}}M) = n-1$ , this implies that there is a positive measure set of directions through  $x_{\alpha}$  among which u is spreading its mass nearly uniformly.

The proofs of Theorems 1 and 2 hinge on a much more general theorem which does not require global geometric assumptions on (M, g) and, in particular, Theorem 2 holds without modification under the assumptions of Theorem 3 below. (We actually prove Theorem 2 under the more general assumptions, see Section 3.7). As far as the authors are aware, Theorem 1 is the first result giving quantitative estimates for the  $L^p$  growth of eigenfunctions that only requires dynamical assumptions. We emphasize that, in contrast with previous improvements on Sogge's  $L^p$  estimates, the assumptions in Theorem 3 below are purely dynamical and, moreover, are local in the sense that they depend only on the geodesics passing through a shrinking neighborhood of a given set in M. Moreover, the techniques do not require long-time wave parametrices.

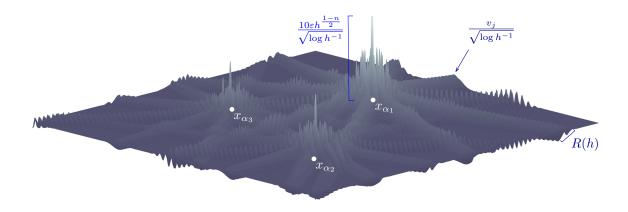


FIGURE 1. The figure illustrates a function u that saturates the  $L^{\infty}$  bound at three points  $x_{\alpha_1}, x_{\alpha_2}, x_{\alpha_3}$  viewed as a superposition of geodesic beams  $v_j$ . Each ridge corresponds to a beam  $v_j$  and is microsupported on a tube  $\mathcal{T}_j$  of radius R(h).

Theorem 3 below controls  $||u||_{L^p(U)}$  using an assumption on the maximal volume of long geodesics joining any two given points in U. For our proof, it is necessary to control the number points in U where the  $L^{\infty}$  norm of u can be large. This is a very delicate and technical part of the argument, as the points in question may be approaching one another at rates  $\sim h^{\delta}$  as  $h \to 0^+$ , with  $0 < \delta < \frac{1}{2}$ . We overcome this problem by developing a second microlocal calculus in Section 6.1 which, after a delicate microlocal argument, yields an uncertainty type principle controlling the amount of  $L^2$  mass shared along short geodesics connecting two nearby points. We expect that additional development of these counting techniques will have many other applications, e.g. to estimates on  $L^p$  norms  $p \leq p_c$ .

To state our theorem, we need to introduce a few geometric objects. First, consider the Hamiltonian function  $p \in C^{\infty}(T^*M \setminus \{0\})$ ,

$$p(x,\xi) = |\xi|_g - 1,$$

and let  $\varphi_t : T^*M \setminus 0 \to T^*M \setminus 0$  denote the Hamiltonian flow for p at time t. We also define the maximal expansion rate and the Ehrenfest time at frequency  $h^{-1}$  respectively:

$$\Lambda_{\max} := \limsup_{|t| \to \infty} \frac{1}{|t|} \log \sup_{S^*M} \|d\varphi_t(x,\xi)\|, \qquad T_e(h) := \frac{\log h^{-1}}{2\Lambda_{\max}}, \tag{1.2}$$

where  $\|\cdot\|$  denotes the norm in any metric on  $T(T^*M)$ . Note that  $\Lambda_{\max} \in [0, \infty)$ , and if  $\Lambda_{\max} = 0$  we may replace it by an arbitrarily small positive constant. We next describe a cover of  $S^*M$  by geodesic tubes.

For each  $\rho_0 \in S^*M$ , the co-sphere bundle to M, let  $H_{\rho_0} \subset M$  be a hypersurface so that  $\rho_0 \in SN^*H_{\rho_0}$ , the unit conormal bundle to  $H_{\rho_0}$ . Then, let

$$\mathcal{H}_{\rho_0} \ \subset \ T^*_{H_{\rho_0}} M = \{ (x,\xi) \in T^*M : \ x \in H_{\rho_0} \}$$

be a hypersurface containing  $SN^*H_{\rho_0}$ . Next, for  $q \in \mathcal{H}_{\rho_0}$ ,  $\tau > 0$ , we define the tube through q of radius R(h) > 0 and 'length'  $\tau + R(h)$  as

$$\Lambda_{q}^{\tau}(R(h)) := \bigcup_{|t| \le \tau + R(h)} \varphi_{t}(B_{\mathcal{H}_{p_{0}}}(q, R(h))), B_{\mathcal{H}_{p_{0}}}(q, R(h)) := \{\rho \in \mathcal{H}_{\rho_{0}} : d(\rho, q) \le R(h)\}, \quad (1.3)$$

and d is distance induced by the Sasaki metric on  $T^*M$  (See e.g. [Bla10, Chapter 9] for a description of the Sasaki metric). Note that the tube runs along the geodesic through  $q \in H_{\rho_0}$ . Similarly, for  $A \subset S^*M$ , we define  $\Lambda_A^{\tau}(R(h))$  in the same way, replacing q with A in (1.3).

**Definition 1.** Let  $A \subset S^*M$ , r > 0, and  $\{\rho_j(r)\}_{j=1}^{N_r} \subset A$  for some  $N_r > 0$ . We say the collection of tubes  $\{\Lambda_{\rho_j}^\tau(r)\}_{j=1}^{N_r}$  is a  $(\tau, r)$ -cover of a set  $A \subset S^*M$  provided

$$\Lambda_A^{\tau}(\frac{1}{2}r) \subset \bigcup_{j=1}^{N_r} \mathcal{T}_j, \qquad \mathcal{T}_j := \Lambda_{\rho_j}^{\tau}(r).$$

Given a  $(\tau, r)$  cover  $\{\mathcal{T}_j\}_{j \in \mathcal{J}}$  for  $S^*M$ , for each  $x \in M$  we define

$$\mathcal{J}_x := \{ j \in \mathcal{J} : \ \pi(\mathcal{T}_j) \cap B(x, r) \neq \emptyset \}.$$

We are now ready to state Theorem 3, where we give *explicit dynamical conditions* guaranteeing quantitative improvements in  $L^p$  norms.

**Theorem 3.** There exists  $\tau_M > 0$  such that for all  $p > p_c$  and  $\varepsilon_0 > 0$  the following holds. Let  $U \subset M$ ,  $0 < \delta_1 < \delta_2 < \frac{1}{2}$  and let  $h^{\delta_2} \leq R(h) \leq h^{\delta_1}$  for all h > 0. Let  $1 \leq T(h) \leq (1 - 2\delta_2)T_e(h)$  and let  $t_0 > 0$  be h-independent. Let  $\{\mathcal{T}_j\}_{j \in \mathcal{J}}$  be a  $(\tau, R(h))$  cover for S\*M for some  $0 < \tau < \tau_M$ .

Suppose that for any pair of points  $x_1, x_2 \in U$ , the tubes over  $x_1$  can be partitioned into a disjoint union  $\mathcal{J}_{x_1} = \mathcal{B}_{x_1, x_2} \sqcup \mathcal{G}_{x_1, x_2}$  where

$$\bigcup_{j \in \mathcal{G}_{x_1, x_2}} \varphi_t(\mathcal{T}_j) \cap S^*_{B(x_2, R(h))} M = \emptyset, \qquad t \in [t_0, T(h)].$$

Then, there are  $h_0 > 0$  and C > 0 so that for all  $u \in \mathcal{D}'(M)$ , and  $0 < h < h_0$ ,

$$\begin{aligned} \|u\|_{L^{p}(U)} &\leq Ch^{-\delta(p)} \left( \frac{\sqrt{t_{0}}}{\sqrt{T(h)}} + \left[ \sup_{x_{1},x_{2} \in U} |\mathcal{B}_{x_{1},x_{2}}|R(h)^{n-1} \right]^{\frac{1}{6+\varepsilon_{0}}(1-\frac{p_{c}}{p})} \right) \\ & \times \left( \|u\|_{L^{2}} + \frac{T(h)}{h} \|(-h^{2}\Delta_{g}-1)u\|_{H^{\frac{n-3}{2}-\frac{n}{p}}_{h}} \right). \end{aligned}$$
(1.4)

In order to interpret (1.4), note that we think of the tubes  $\mathcal{G}_{x_1,x_2}$  and  $\mathcal{B}_{x_1,x_2}$  as respectively good (or non- looping) and bad (or looping) tubes. Then, observe that  $|\mathcal{B}_{x_1,x_2}|R(h)^{n-1} \sim$  $\operatorname{vol}(\bigcup_{j\in\mathcal{B}_{x_1,x_2}}\mathcal{T}_j\cap S^*_{x_1}M)$ , and  $\bigcup_{j\in\mathcal{B}_{x_1,x_2}}\mathcal{T}_j$  is the set of points over  $x_1$  which may loop through  $x_2$  in time T(h). Therefore, if the volume of points in  $S^*_{x_1}M$  looping through  $x_2$  is bounded by  $T(h)^{-(3+\varepsilon_0)(1-\frac{p_c}{p})^{-1}}$ , (1.4) provides  $T(h)^{-\frac{1}{2}}$  improvements over the standard  $L^p$  bounds. We expect these non-looping type assumptions to be valid on generic manifolds. Theorem 3 can be used to obtain improved  $L^p$  resolvent bounds [Cue20, Theorem 2.21] and, as shown their, are stable by certain rough perturbations. These estimates in turn can be used to construct complex geometric optics solutions and solve certain inverse problems [DSFKS13].

As in [CG20, Theorem 5 and Section 5], the assumptions of Theorem 3 can be verified in certain integrable situations with  $T(h) \gg \log h^{-1}$ , thus producing  $o((\log h^{-1})^{-\frac{1}{2}})$  improvements. Moreover, in [CG19a], we used these types of good and bad tubes to understand averages and  $L^{\infty}$ -norms under various assumptions on M, including that it has Anosov geodesic flow or non-positive curvature. Since our results do not require parametrices for the wave-group, we expect that the arguments leading to Theorem 3 will provide *polynomial* improvements over Sogge's estimates on manifolds where Egorov type theorems hold for longer than logarithmic times.

**Remark 1.** The proofs below adapt directly to the case of quasimodes for real principal type semiclassical pseudodifferential operators of Laplace type. That is, to operators with principal symbol p satisfying both  $\partial_{\xi}p \neq 0$  on  $\{p = 0\}$  and that  $\{p = 0\} \cap T_x^*M$  has positive definite second fundamental form. This is the case, for example, for Schrödinger operators away from the forbidden region. However, for concreteness and simplicity of exposition, we have chosen to consider only the Laplace operator.

1.1. Discussion of the proof of Theorem 3. Our method for proving Theorem 3 differs from the standard approaches for treating  $L^p$  norms in two major ways. It hinges on adapting the geodesic beam techniques constructed by the authors [CG20], and on the development of a new second-microlocal calculus.

We start in Section 2 by covering  $S^*M$  with tubes of radius R(h). Then, in Section 3.2, we decompose the function u, whose  $L^p$  norm we wish to study, into geodesic beams i.e. into pieces microlocalized along each of these tubes. We then sort these beams into collections which carry  $\sim 2^{-k} \|u\|_{L^2}$  mass and study the collections for each k separately.

In order to understand the  $L^p$  norm of u, we next decompose the manifold into balls of radius R(h). By constructing a good cover of M, we are able to think the  $L^p$  norm of a function on M as the  $L^p$  norm of a function on a disjoint union of balls of radius R(h). In each ball, B, we are able to apply the methods from [CG20] to understand the  $L^{\infty}$  norm of u on B in terms of the number of tubes with mass  $\sim 2^{-k} ||u||_{L^2}$  passing over that ball.

To bound the  $L^p$  norm with  $p < \infty$ , it then remains to understand the number of balls on which the function u can have a certain  $L^{\infty}$  norm. In Section 3.4 we first observe that when uhas relatively low  $L^{\infty}$  norm on a ball, this ball can be neglected by interpolation with Sogge's  $L^{p_c}$  estimate. It thus remains to understand the number of balls B on which the  $L^{\infty}$  norm of ucan be large (i.e. close to extremal). In fact, we will show that the number of balls such that  $\|u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \geq Ch^{\frac{1-n}{2}}/\sqrt{\log h^{-1}}$  is bounded uniformly in h. That is, there is some number N such that there are at most N such balls for any value of h > 0. This is the content of Theorem 2 and is proved in Section 3.7. It is in this step where a crucial new ingredient is input.

The new method allows us to control the size of the set on which an eigenfunction (or quasimode) can have high  $L^{\infty}$  norm. The method relies on understanding how much  $L^2$  mass can be effectively shared along short geodesics joining two nearby points in such a way as to produce large  $L^{\infty}$  norm at both points. That is, if  $x_{\alpha}$  and  $x_{\beta}$  are nearby points on M, and if  $|u(x_{\alpha})|$  and  $|u(x_{\beta})|$  are near extremal, how much total  $L^2$  mass must the tubes over  $x_{\alpha}$  and  $x_{\beta}$  carry? In order to understand this sharing phenomenon, we develop a new second microlocal calculus associated to a Lagrangian foliation L over a co-isotropic submanifold  $\Gamma \subset T^*M$ . This calculus allows for simultaneous localization along a leaf of L and along  $\Gamma$ . The calculus, which is developed in Section 5, can be seen as an interpolation between those in [DZ16] and [SZ99]. It is then the incompatibility between the calculi coming from two nearby points which allows us to control this sharing of mass. This incompatibility is demonstrated in Section 6 in the form of an uncertainty principle type of estimate.

Once the number of balls with high  $L^{\infty}$  norm is understood, it remains to employ the nonlooping techniques from [CG20] where the  $L^2$  mass on a collection of tubes is estimated using its non-looping time (see Section 3.5.2).

1.2. Outline of the paper. In section 2, we construct the covers of  $S^*M$  by tubes and  $T^*M$  by balls which are necessary in the rest of the article. Section 3 contains the proof of Theorems 2 and 3. This proof uses the anisotropic calculus developed in Section 5 and the almost orthogonality results from Section 6. Section 4 contains the necessary dynamical arguments to prove Theorem 1 using Theorem 3.

ACKNOWLEDGEMENTS. The authors are grateful to the National Science Foundation for support under grants DMS-1900519 (Y.C) and DMS-1502661, DMS-1900434 (J.G.). Y.C. is grateful to the Alfred P. Sloan Foundation.

## 2. Tubes Lemmata

The next few lemmas are aimed at constructing  $(\tau, r)$ -good covers and partitions of various subsets of  $T^*M$  (see also [CG20, Section 3.2]).

**Definition 2** (good covers and partitions). Let  $A \subset T^*M$ , r > 0, and  $\{\rho_j(r)\}_{j=1}^{N_r} \subset A$  be a collection of points, for some  $N_r > 0$ . Let  $\mathfrak{D}$  be a positive integer. We say that the collection of tubes  $\{\Lambda_{\rho_j}^{\tau}(r)\}_{j=1}^{N_r}$  is a  $(\mathfrak{D}, \tau, r)$ -good cover of  $A \subset T^*M$  provided it is a  $(\tau, r)$ -cover of A and there exists a partition  $\{\mathcal{J}_{\ell}\}_{\ell=1}^{\mathfrak{D}}$  of  $\{1, \ldots, N_r\}$  so that for every  $\ell \in \{1, \ldots, \mathfrak{D}\}$ 

$$\Lambda_{\rho_i}^{\tau}(3r) \cap \Lambda_{\rho_i}^{\tau}(3r) = \emptyset, \qquad i, j \in \mathcal{J}_{\ell}, \qquad i \neq j.$$

In addition, for  $0 \le \delta \le \frac{1}{2}$  and  $R(h) \ge 8h^{\delta}$ , we say that a collection  $\{\chi_j\}_{j=1}^{N_h} \subset S_{\delta}(T^*M; [0, 1])$  is a  $\delta$ -good partition for A associated to a  $(\mathfrak{D}, \tau, R(h))$ -good cover if  $\{\chi_j\}_{j=1}^{N_h}$  is bounded in  $S_{\delta}$  and

(1) supp 
$$\chi_j \subset \Lambda_{\rho_j}^{\tau}(R(h)),$$
 (2)  $\sum_{j=1}^{N_h} \chi_j \ge 1 \text{ on } \Lambda_A^{\tau/2}(\frac{1}{2}R(h)).$ 

**Remark 2.** We show below that for any compact Riemannian manifold M, there are  $\mathfrak{D}_M, R_0, \tau_0 > 0$ , depending only on (M, g), such that for  $0 < \tau < \tau_0$ ,  $0 < r < R_0$ , there exists a  $(\mathfrak{D}_M, \tau, r)$  good cover for  $S^*M$ .

We start by constructing a useful cover of any Riemannian manifold with bounded curvature.

**Lemma 2.1.** Let  $\tilde{M}$  be a compact Riemannian manifold. There exist  $\mathfrak{D}_n > 0$ , depending only on n, and  $R_0 > 0$  depending only on n and a lower bound for the sectional curvature of  $\tilde{M}$ , so

that the following holds. For  $0 < r < R_0$ , there exist a finite collection of points  $\{x_{\alpha}\}_{\alpha \in \mathcal{I}} \subset M$ ,  $\mathcal{I} = \{1, \ldots, N_r\}$ , and a partition  $\{\mathcal{I}_i\}_{i=1}^{\mathfrak{D}_n}$  of  $\mathcal{I}$  so that

$$\begin{split} \tilde{M} \subset \bigcup_{\alpha \in \mathcal{I}} B(x_{\alpha}, r), & B(x_{\alpha_1}, 3r) \cap B(x_{\alpha_2}, 3r) = \emptyset, \qquad \alpha_1, \alpha_2 \in \mathcal{I}_i, \quad \alpha_1 \neq \alpha_2, \\ \{x_{\alpha}\}_{\alpha \in \mathcal{I}} \text{ is an } \frac{r}{2} \text{ maximal separated set in } \tilde{M}. \end{split}$$

Proof. Let  $\{x_{\alpha}\}_{\alpha \in \mathcal{I}}$  be a maximal  $\frac{r}{2}$  separated set in  $\tilde{M}$ . Fix  $\alpha_0 \in \mathcal{I}$  and suppose that  $B(x_{\alpha_0}, 3r) \cap B(x_{\alpha}, 3r) \neq \emptyset$  for all  $\alpha \in \mathcal{K}_{\alpha_0} \subset \mathcal{I}$ . Then for all  $\alpha \in \mathcal{K}_{\alpha_0}, B(x_{\alpha}, \frac{r}{2}) \subset B(x_{\alpha_0}, 8r)$ . In particular,

$$\sum_{\alpha \in \mathcal{K}_{\alpha_0}} \operatorname{vol}(B(x_{\alpha}, \frac{r}{2})) \le \operatorname{vol}(B(x_{\alpha_0}, 8r)).$$

Now, there exist  $R_0 > 0$  depending on n and a lower bound on the sectional curvature of M, and  $\mathfrak{D}_n > 0$  depending only on n, so that for all  $0 < r < R_0$ ,

$$\operatorname{vol}(B(x_{\alpha_0}, 8r)) \le \operatorname{vol}(B(x_{\alpha}, 14r)) \le \mathfrak{D}_n \operatorname{vol}(B(x_{\alpha}, \frac{r}{2})).$$
(2.1)

Hence, it follows from (2.1) that

 $\alpha$ 

$$\sum_{\in \mathcal{K}_{\alpha_0}} \operatorname{vol}(B(x_{\alpha}, \frac{r}{2})) \le \operatorname{vol}(B(\rho_{\alpha_0}, 8r)) \le \frac{\mathfrak{D}_n}{|\mathcal{K}_{\alpha_0}|} \sum_{\alpha \in \mathcal{K}_{\alpha_0}} \operatorname{vol}(B(x_{\alpha}, \frac{r}{2})).$$

In particular,  $|\mathcal{K}_{\alpha_0}| \leq \mathfrak{D}_n$ .

At this point we have proved that each of the balls  $B(x_{\alpha}, 3r)$  intersects at most  $\mathfrak{D}_n - 1$  other balls. We now construct the sets  $\mathcal{I}_1, \ldots, \mathcal{I}_{\mathfrak{D}_n}$  using a greedy algorithm. We will say that the index  $\alpha_1$  intersects the index  $\alpha_2$  if

$$B(x_{\alpha_1}, 3r) \cap B(x_{\alpha_2}, 3r) \neq \emptyset.$$

We place the index  $1 \in \mathcal{I}_1$ . Then suppose we have placed the indices  $\{1, \ldots, \alpha\}$  in  $\mathcal{I}_1, \ldots, \mathcal{I}_{\mathfrak{D}_n}$  so each of the  $\mathcal{I}_i$ 's consists of disjoint indices. Then, since  $\alpha + 1$  intersects at most  $\mathfrak{D}_n - 1$  indices, it is disjoint from  $\mathcal{I}_i$  for some *i*. We add the index  $\alpha$  to  $\mathcal{I}_i$ . By induction we obtain the partition  $\mathcal{I}_1, \ldots, \mathcal{I}_{\mathfrak{D}_n}$ .

Now, suppose that there exists  $x \in \tilde{M}$  so that  $x \notin \bigcup_{\alpha \in \mathcal{I}} B(x_{\alpha}, r)$ . Then,  $\min_{\alpha \in \mathcal{I}} d(x, x_{\alpha}) \geq r$ , a contradiction of the r/2 maximality of  $x_{\alpha}$ .

In order to construct our microlocal partition, we first fix a smooth hypersurface  $H \subset M$ , and choose Fermi normal coordinates  $x = (x_1, x')$  in a neighborhood of  $H = \{x_1 = 0\}$ . We write  $(\xi_1, \xi') \in T_x^*M$  for the dual coordinates. Let

$$\Sigma_{H} := \left\{ (x,\xi) \in S_{H}^{*}M | |\xi_{1}| \ge \frac{1}{2} \right\}$$
(2.2)

We then consider

$$\mathcal{H}_{\Sigma_{H}} := \{ (x,\xi) \in T_{H}^{*}M \mid |\xi_{1}| \ge \frac{1}{2}, \ \frac{1}{2} < |\xi|_{g(x)} < \frac{3}{2} \}.$$

$$(2.3)$$

Then,  $\mathcal{H}_{\Sigma_{tr}}$  is transverse to the geodesic flow and there is  $0 < \tau_{\text{inj}H} < 1$  so that the map

$$\Psi: [-\tau_{\mathrm{inj}H}, \tau_{\mathrm{inj}H}] \times \mathcal{H}_{\Sigma_H} \to T^*M, \qquad \Psi(t, \rho) := \varphi_t(\rho), \tag{2.4}$$

is injective. Our next lemma shows that there is  $\mathfrak{D}_n > 0$  depending only on n such that one can construct a  $(\mathfrak{D}_n, \tau, r)$ -good cover of  $\Sigma_H$ .

**Lemma 2.2.** There exist  $\mathfrak{D}_n > 0$  depending only on n,  $R_0 = R_0(n, H) > 0$ , such that for  $0 < r_1 < R_0$ ,  $0 < r_0 \leq \frac{r_1}{2}$ , there exist points  $\{\rho_j\}_{j=1}^{N_{r_1}} \subset \Sigma_H$  and a partition  $\{\mathcal{J}_i\}_{i=1}^{\mathfrak{D}_n}$  of  $\{1, \ldots, N_{r_1}\}$  so that for all  $0 < \tau < \frac{\tau_{injH}}{2}$ 

• 
$$\Lambda_{\Sigma_{H}}^{\tau}(r_{0}) \subset \bigcup_{j=1}^{N_{r_{1}}} \Lambda_{\rho_{j}}^{\tau}(r_{1}),$$
 •  $\Lambda_{\rho_{j}}^{\tau}(3r_{1}) \cap \Lambda_{\rho_{\ell}}^{\tau}(3r_{1}) = \emptyset, \quad j, \ell \in \mathcal{J}_{i}, \quad j \neq \ell.$ 

*Proof.* We first apply Lemma 2.1 to  $\tilde{M} = \Sigma_H$  to obtain  $R_0 > 0$  depending only on n and the sectional curvature of H and that of M near H, so that for  $0 < r_1 < R_0$ , there exist  $\{\rho_j\}_{j=1}^{N_{r_1}} \subset \Sigma_H$  and a partition  $\{\mathcal{J}_i\}_{i=1}^{\mathfrak{D}_n}$  of  $\{1, \ldots, N_{r_1}\}$  such that

$$\begin{split} \Sigma_{\!_{H}} &\subset \bigcup_{j=1}^{N_{r_1}} B(\rho_j, r_1), \qquad B(\rho_j, 3r_1) \cap B(\rho_\ell, 3r_1) = \emptyset, \qquad j, \ell \in \mathcal{J}_i, \quad j \neq \ell, \\ & \{\rho_j\}_{j=1}^{N_{r_1}} \text{ is an } \frac{r_1}{2} \text{ maximal separated set in } \Sigma_{\!_{H}}. \end{split}$$

Now, suppose that  $j, \ell \in \mathcal{J}_i$  and

$$\Lambda^{\tau}_{\rho_{\ell}}(3r_1) \cap \Lambda^{\tau}_{\rho_i}(3r_1) \neq \emptyset.$$

Then, there exist  $q_{\ell} \in B(\rho_{\ell}, 3r_1) \cap \mathcal{H}_{\Sigma_H}$ ,  $q_j \in B(\rho_j, 3r_1) \cap \mathcal{H}_{\Sigma_H}$  and  $t_{\ell}, t_j \in [-\tau, \tau]$  so that  $\varphi_{t_{\ell}-t_j}(q_{\ell}) = q_j$ . Here,  $\mathcal{H}_{\Sigma}$  is the hypersurface defined in (2.3). In particular, for  $\tau < \tau_{injH}/2$ , this implies that  $q_{\ell} = q_j$ ,  $t_{\ell} = t_j$  and hence  $B(\rho_{\ell}, 3r_1) \cap B(\rho_j, 3r_1) \neq \emptyset$  a contradiction.

Now, suppose  $r_0 \leq r_1$  and that there exists  $\rho \in \Lambda_{\Sigma_H}^{\tau}(r_0)$  so that  $\rho \notin \bigcup_{j=1,\dots,N_{r_1}} \Lambda_{\rho_j}^{\tau}(r_1)$ . Then, there are  $|t| < \tau + r_0$  and  $q \in \mathcal{H}_{\Sigma_H}$  so that

$$\rho = \varphi_t(q), \qquad d(q, \Sigma_H) < r_0, \qquad \min_{j=1,\dots,N_{r_1}} d(q, \rho_j) \ge r_1.$$

In particular, there exists  $\tilde{\rho} \in \Sigma_{H}$  with  $d(q, \tilde{\rho}) < r_0$  such that for all  $j = 1, \ldots, N_{r_1}$ ,

$$d(\tilde{\rho}, \rho_j) \ge d(q, \rho_j) - d(q, \tilde{\rho}) > r_1 - r_0.$$

This contradicts the maximality of  $\{\rho_j\}_{j=1}^{N_{r_1}}$  if  $r_0 \leq r_1/2$ .

We proceed to build a  $\delta$ -good partition of unity associated to the cover we constructed in Lemma 2.2. The key feature in this partition is that it is invariant under the geodesic flow. Indeed, the partition is built so that its quantization commutes with the operator  $P = -h^2 \Delta - I$  in a neighborhood of  $\Sigma_{\mu}$ .

**Proposition 2.3.** There exist  $\tau_1 = \tau_1(\tau_{injH}) > 0$  and  $\varepsilon_1 = \varepsilon_1(\tau_1) > 0$ , and given  $0 < \delta < \frac{1}{2}$ ,  $0 < \varepsilon \leq \varepsilon_1$ , there exists  $h_1 > 0$ , so that for any  $0 < \tau \leq \tau_1$ , and  $R(h) \geq 2h^{\delta}$ , the following holds.

There exist  $C_1 > 0$  so that for all  $0 < h \leq h_1$  and every  $(\tau, R(h))$ -cover of  $\Sigma_{\!_H}$  there exists a partition of unity  $\chi_j \in S_\delta \cap C_c^\infty(T^*M; [-C_1h^{1-2\delta}, 1+C_1h^{1-2\delta}])$  on  $\Lambda_{\Sigma_{\!_H}}^\tau(\frac{1}{2}R(h))$  for which

$$\operatorname{supp} \chi_j \subset \Lambda_{\rho_j}^{\tau+\varepsilon}(R(h)), \qquad \operatorname{MS_h}([P, Op_h(\chi_j)]) \cap \Lambda_{\Sigma_H}^{\tau}(\varepsilon) = \emptyset,$$
$$\sum_j \chi_j \equiv 1 \ on \ \Lambda_{\Sigma_H}^{\tau}(\frac{1}{2}R(h)),$$

and  $\{\chi_j\}_j$  is bounded in  $S_{\delta}$ , and  $[-h^2\Delta_g, Op_h(\chi_j)]$  is bounded in  $\Psi_{\delta}$ .

*Proof.* The proof is identical to that of [CG20, Proposition 3.4]. Although the claim that  $\sum_{j} \chi_{j} \equiv 1$  on  $\Lambda^{\tau}_{\Sigma_{H}}(\frac{1}{2}R(h))$  does not appear its statement, it is included in its proof.  $\Box$ 

### 3. Proof of Theorem 3

For each  $q \in S^*M$ , choose a hypersurface  $H_q \subset M$  with  $q \in SN^*H_q$  and  $\tau_{injH_q} > \frac{inj(M)}{2}$ , where  $\tau_{injH_q}$  is defined in (2.4) and inj(M) is the injectivity radius of M. We next use Lemma 2.2 to generate a cover of  $\Sigma_{H_q}$ . Lemma 2.2 yields the existence of  $\mathfrak{D}_n > 0$  depending only on n and  $R_0 = R_0(n, H_q) > 0$ , such that the following holds. Since by assumption  $R(h) \leq h^{\delta_1}$ , there is  $h_0 > 0$  such that  $h^{\delta_2} \leq R(h) \leq R_0$  for all  $0 < h < h_0$ . Also, set  $r_1 := R(h)$  and  $r_0 := \frac{1}{2}R(h)$ . Then, by Lemma 2.2 there exist  $N_{R(h)} = N_{R(h)}(q, R(h)) > 0$ ,  $\{\rho_j\}_{j\in\mathcal{J}_q} \subset \Sigma_{H_q}$  with  $\mathcal{J}_q = \{1, \ldots, N_{R(h)}\}$ , and a partition  $\{\mathcal{J}_{q,i}\}_{i=1}^{\mathfrak{D}_n}$  of  $\mathcal{J}_q$ , so that for all  $0 < \tau < \frac{\tau_{injH_q}}{2}$ 

• 
$$\Lambda_{\Sigma_{H_q}}^{\tau}(\frac{1}{2}R(h)) \subset \bigcup_{j \in \mathcal{J}_q} \Lambda_{\rho_j}^{\tau}(R(h)),$$
 (3.1)

• 
$$\bigcup_{i=1}^{\mathfrak{D}_n} \mathcal{J}_{q,i} = \mathcal{J}_q, \tag{3.2}$$

• 
$$\Lambda^{\tau}_{\rho_{j_1}}(3R(h)) \cap \Lambda^{\tau}_{\rho_{j_2}}(3R(h)) = \emptyset, \qquad j_1, j_2 \in \mathcal{J}_{q,i}, \quad j_1 \neq j_2.$$
 (3.3)

By (3.1) there is an *h*-independent open neighborhood of q,  $V_q \subset S^*M$ , covered by tubes as in Lemma 2.2. Since  $S^*M$  is compact, we may choose  $\{q_\ell\}_{\ell=1}^L$  with L independent of h, so that  $S^*M \subset \bigcup_{\ell=1}^L V_{q_\ell}$ . In particular, if  $0 < \tau \leq \min_{1 \leq \ell \leq L} \tau_{H_{q_\ell}}$ , and for each  $\ell \in \{1, \ldots, L\}$  we let

$$\mathcal{T}_{q_{\ell},j} = \Lambda_{\rho_j}^{\tau}(R(h)),$$

then there is  $\mathfrak{D}_M > 0$  such that  $\bigcup_{\ell=1}^L \{\mathcal{T}_{q_\ell,j}\}_{j \in \mathcal{J}_{q_\ell}}$  is a  $(\mathfrak{D}_M, \tau, R(h))$ -good cover for  $S^*M$ . Let  $\{\psi_{q_\ell}\}_{\ell=1}^L \subset C_c^{\infty}(T^*M)$  satisfy

$$\operatorname{supp} \psi_{q_{\ell}} \subset \{(x,\xi) \in T^*M \setminus \{0\} \mid \left(x, \frac{\xi}{|\xi|_g}\right) \in V_{q_{\ell}}\} \qquad \forall \ell = 1, \dots, L$$
$$\sum_{\ell=1}^{L} \psi_{q_{\ell}} \equiv 1 \text{ in an } h \text{-independent neighborhood of } S^*M.$$

We split the analysis of u in two parts: near and away from the characteristic variety  $\{p = 0\} = S^*M$ . In what follows we use C to denote a positive constant that may change from line to line.

3.1. It suffices to study u near the characteristic variety. In this section we reduce the study of  $||u||_{L^p(U)}$  to an *h*-dependent neighborhood of the characteristic variety  $\{p = 0\} = S^*M$ . We will use repeatedly the following result.

**Lemma 3.1.** For all  $\varepsilon > 0$  and all  $p \ge 2$ , there exists C > 0 such that

$$\|u\|_{L^p} \le Ch^{n(\frac{1}{p} - \frac{1}{2})} \|u\|_{H_h^{n(\frac{1}{2} - \frac{1}{p}) + \varepsilon}}.$$
(3.4)

 $\begin{array}{l} \textit{Proof. By [Gal19, Lemma 6.1] (or more precisely its proof), for any $\varepsilon > 0$, there exists $C_{\varepsilon} \geq 1$ so that } \| \operatorname{Id} \|_{H_h^{\frac{n}{2} + \varepsilon} \to L^{\infty}} \leq C_{\varepsilon} h^{-\frac{n}{2}}. \text{ By complex interpolation of Id} : L^2 \to L^2 \text{ and Id} : H_h^{\frac{n}{2} + \varepsilon} \to L^{\infty} \\ \text{with } \theta = \frac{2}{p} \text{ we obtain } \| \operatorname{Id} \|_{H_h^{(\frac{n}{2} + \varepsilon)(1-\theta)} \to L^p} \leq C_{\varepsilon}^{1-\theta} h^{-\frac{n}{2}(1-\theta)}, \text{ and this yields (3.4).} \end{array}$ 

Observe that

$$u = \sum_{\ell=1}^{L} Op_h(\psi_{q_\ell}) u + \left(1 - \sum_{\ell=1}^{L} Op_h(\psi_{q_\ell})\right) u.$$

Note that since  $1 - \sum_{\ell=1}^{L} \psi_{q_{\ell}} = 0$  in an *h*-independent neighborhood of  $S^*M = \{p = 0\}$ , by the standard elliptic parametrix construction (e.g. [DZ19, Appendix E]) there is  $E \in \Psi^{-2}(M)$  with

$$1 - \sum_{\ell=1}^{L} Op_h(\psi_{q_\ell}) = EP + O(h^{\infty})_{\Psi^{-\infty}}.$$
 (3.5)

Next, combining (3.5) with Lemma 3.1, and using that  $h^{n(\frac{1}{p}-\frac{1}{2})} = h^{-\delta(p)+\frac{1}{2}}h^{-1}$ , we have

$$\begin{aligned} \left\| \left( 1 - \sum_{\ell=1}^{L} Op_{h}(\psi_{q_{\ell}}) \right) u \right\|_{L^{p}} &\leq Ch^{n(\frac{1}{p} - \frac{1}{2})} \|EPu\|_{H_{h}^{n(\frac{1}{2} - \frac{1}{p}) + \varepsilon}} + O(h^{\infty}) \|u\|_{L^{2}} \\ &\leq Ch^{-\delta(p) + \frac{1}{2}} h^{-1} \|Pu\|_{H_{h}^{n(\frac{1}{2} - \frac{1}{p}) + \varepsilon - 2}} + O(h^{\infty}) \|u\|_{L^{2}}. \end{aligned}$$
(3.6)

It remains to understand the terms  $Op_h(\psi_{q_\ell})u$ . Since there are finitely many such terms,

$$\left\|\sum_{\ell=1}^{L} Op_{h}(\psi_{q_{\ell}})u\right\|_{L^{p}} \leq \sum_{\ell=1}^{L} \|Op_{h}(\psi_{q_{\ell}})u\|_{L^{p}},\tag{3.7}$$

and consider each term  $\|Op_h(\psi_{q_\ell})u\|_{L^p}$  individually.

By Proposition 2.3 for each  $\ell = 1, \ldots, L$  there exist  $\tau_1(q_\ell) > 0$ ,  $\varepsilon_1(q_\ell) > 0$ , and a family of cut-offs  $\{\tilde{\chi}_{\tau_{q_\ell,j}}\}_{j \in \mathcal{J}_{q_\ell}}$ , with  $\tilde{\chi}_{\tau_{q_\ell,j}}$  supported in  $\Lambda_{\rho_j}^{\tau+\varepsilon_1(q_\ell)}(R(h))$  and such that for all  $0 < \tau < \tau_1(q_\ell)$ 

$$\sum_{j \in \mathcal{J}_{q_{\ell}}} \tilde{\chi}_{\tau_{q_{\ell},j}} \equiv 1 \qquad \text{on} \quad \Lambda^{\tau}_{\Sigma_{H_{q_{\ell}}}}(\frac{1}{2}R(h)).$$
(3.8)

Let  $\tau_0(q_\ell)$  from [CG20, Theorem 8]. Then, set

$$\tau_{M} := \min_{1 \le \ell \le L} \left\{ \frac{\operatorname{inj}(M)}{4}, \, \tau_{0}(q_{\ell}), \, \tau_{1}(q_{\ell}), \, \frac{1}{2}\tau_{\operatorname{inj}Hq_{\ell}} \right\}.$$

From now on we work with tubes  $\mathcal{T}_{q_{\ell},j} = \Lambda^{\tau}_{\rho_j}(R(h))$  for some  $0 < \tau < \tau_M$ . Next, we localize u near and away from  $\Lambda^{\tau}_{\Sigma_{H_{as}}}(h^{\delta})$ :

$$Op_h(\psi_{q_\ell})u = \sum_{j \in \mathcal{J}_{q_\ell}} Op_h(\tilde{\chi}_{\tau_{q_\ell,j}}) Op_h(\psi_{q_\ell})u + \Big(1 - \sum_{j \in \mathcal{J}_{q_\ell}} Op_h(\tilde{\chi}_{\tau_{q_\ell,j}})\Big) Op_h(\psi_{q_\ell})u.$$

**Remark 3.** We refer to functions of the form  $Op_h(\tilde{\chi}_{\tau_{q_{\ell},j}})u$  as geodesic beams. One can check using Proposition 2.3, that if u solves Pu = O(h), then the geodesic beams also solve Pu = O(h) and are localized to an R(h) neighborhood of a length~1 segment of a geodesic.

In particular, by (3.8),  $\frac{1}{2}R(h) \ge \frac{1}{2}h^{\delta_2}$ , and [CG20, Lemma 3.6], there is  $E \in h^{-\delta_2}\Psi_{\delta_2}^{\text{comp}}$  so that

$$\left(1 - \sum_{j \in \mathcal{J}_{q_{\ell}}} Op_h(\tilde{\chi}_{\tau_{q_{\ell},j}})\right) Op_h(\psi_{q_{\ell}}) = EP + O_{\Psi^{-\infty}}(h^{\infty}).$$
(3.9)

Since  $h^{n(\frac{1}{p}-\frac{1}{2})-\delta_2} = h^{-\delta(p)+\frac{1}{2}-\delta_2}h^{-1}$ , combining (3.9) with Lemma 3.1 yields

$$\left\| \left( 1 - \sum_{j \in \mathcal{J}_{q_{\ell}}} Op_h(\tilde{\chi}_{\tau_{q_{\ell},j}}) \right) Op_h(\psi_{q_{\ell}}) u \right\|_{L^p} \le Ch^{-\delta(p) - \frac{1}{2} - \delta_2} \|Pu\|_{H_h^{n(\frac{1}{2} - \frac{1}{p}) + \varepsilon - 2}} + O(h^{\infty}) \|u\|_{L^2}.$$
(3.10)

Combining (3.6), (3.7) and (3.10) we have proved that for  $U \subset M$ 

$$\|u\|_{L^{p}(U)} \leq \sum_{\ell=1}^{L} \left\| \sum_{j \in \mathcal{J}_{q_{\ell}}} Op_{h}(\tilde{\chi}_{\tau_{q_{\ell},j}}) Op_{h}(\psi_{q_{\ell}}) u \right\|_{L^{p}(U)} + Ch^{-\delta(p) + \frac{1}{2} - \delta_{2}} h^{-1} \|Pu\|_{H^{n(\frac{1}{2} - \frac{1}{p}) + \varepsilon - 2}_{h}} + O(h^{\infty}) \|u\|_{L^{2}}.$$
(3.11)

3.2. Filtering tubes by  $L^2$ -mass. By (3.11) it only remains to control terms of the form  $\|\sum_{j\in \mathcal{J}_{q_\ell}} Op_h(\tilde{\chi}_{\tau_{q_\ell,j}}) Op_h(\psi_{q_\ell}) u\|_{L^p}$ , where u is localized to  $V_{q_\ell}$  within the characteristic variety  $S^*M$  and, more importantly, to the tubes  $\mathcal{T}_{q_\ell,j}$ . We fix  $\ell$  and, abusing notation slightly, write

$$\psi := \psi_{q_{\ell}}, \qquad \mathcal{J} = \mathcal{J}_{q_{\ell}}, \qquad \mathcal{T}_{j} = \mathcal{T}_{q_{\ell},j}, \qquad \tilde{\chi}_{\tau_{j}} := \tilde{\chi}_{\tau_{q_{\ell},j}}, v := \sum_{j \in \mathcal{J}} Op_{h}(\tilde{\chi}_{\tau_{j}}) Op_{h}(\psi) u.$$
(3.12)

Let  $T = T(h) \ge 1$ . For each  $j \in \mathcal{J}$  let

$$\chi_{\tau_i} \in C_c^{\infty}(T^*M; [0, 1]) \cap S_{\delta}$$

$$(3.13)$$

be a smooth cut-off function with  $\operatorname{supp} \chi_{\tau_j} \subset \mathcal{T}_j, \ \chi_{\tau_j} \equiv 1$  on  $\operatorname{supp} \tilde{\chi}_{\tau_j}$ , and such that  $\{\chi_j\}_j$  is bounded in  $S_{\delta}$ . We shall work with the modified norm

$$||u||_{P,T} := ||u||_{L^2} + \frac{T}{h} ||Pu||_{L^2}$$

Note that this norm is the natural norm for obtaining  $T^{-\frac{1}{2}}$  improved estimates in  $L^p$  bounds since the fact that u is an  $o(T^{-1}h)$  quasimode implies, roughly, that u is an accurate solution to  $(hD_t + P)u = 0$  for times  $t \leq T$ . For each integer  $k \geq -1$  we consider the set

$$\mathcal{A}_{k} = \left\{ j \in \mathcal{J} : \frac{1}{2^{k+1}} \|u\|_{P,T} \le \|Op_{h}(\chi_{\tau_{j}})u\|_{L^{2}} + h^{-1} \|Op_{h}(\chi_{\tau_{j}})Pu\|_{L^{2}} \le \frac{1}{2^{k}} \|u\|_{P,T} \right\}.$$
(3.14)

It follows that  $\mathcal{A}_k$  consists of those tubes  $\mathcal{T}_j$  with  $L^2$  mass comparable to  $2^{-k}$ .

Observe that since  $|\chi_{\tau_j}| \leq 1$ , for *h* small enough depending on finitely many seminorms of  $\chi_j$ ,  $||Op_h(\chi_{\tau_j})||_{L^2 \to L^2} \leq 2$ . In particular, this together with  $T \geq 1$ , implies that

$$\mathcal{J} = \bigcup_{k \ge -1} \mathcal{A}_k \,. \tag{3.15}$$

**Lemma 3.2.** There exists  $C_n > 0$  so that for all  $k \ge -1$ 

$$|\mathcal{A}_k| \le C_n 2^{2k}.\tag{3.16}$$

*Proof.* According to (3.2), the collection  $\{\mathcal{T}_j\}_{j\in\mathcal{J}}$  can be partitioned into  $\mathfrak{D}_n$  sets of disjoint tubes. Thus, we have  $\sum_{j\in\mathcal{J}} |\chi_{\mathcal{T}_j}|^2 \leq \mathfrak{D}_n$  and there is  $C_n > 0$  depending only on n such that

$$\left\|\sum_{j\in\mathcal{J}}Op_h(\chi_{\tau_j})^*Op_h(\chi_{\tau_j})\right\|_{L^2\to L^2}\leq C_n.$$

In particular,

$$\sum_{j \in \mathcal{J}} \|Op_h(\chi_{\tau_j})u\|_{L^2}^2 \le C_n \|u\|_{L^2}^2 \quad \text{and} \quad \sum_{j \in \mathcal{J}} \|Op_h(\chi_{\tau_j})Pu\|_{L^2}^2 \le C_n \|Pu\|_{L^2}^2.$$

Therefore,

$$\|\mathcal{A}_{k}\|^{2^{-2k-2}}\|u\|^{2}_{P,T} \leq 2\Big(\sum_{j\in\mathcal{A}_{k}}\|Op_{h}(\chi_{\tau_{j}})u\|^{2}_{L^{2}} + h^{-2}\|Op_{h}(\chi_{\tau_{j}})Pu\|^{2}_{L^{2}}\Big) \leq C_{n}\|u\|^{2}_{P,T}.$$

Next, let

$$w_k := \sum_{j \in \mathcal{A}_k} Op_h(\tilde{\chi}_{\tau_j}) Op_h(\psi) u.$$
(3.17)

Then, by (3.12) and (3.15) we have

$$v = \sum_{k=-1}^{\infty} w_k. \tag{3.18}$$

The goal is therefore to control  $||w_k||_{L^p(U)}$  for each k since the triangle inequality yields

$$||v||_{L^p(U)} \le \sum_{k=-1}^{\infty} ||w_k||_{L^p(U)}.$$

3.3. Filtering tubes by  $L^{\infty}$  weight on shrinking balls. By Lemma 2.1, there are points  $\{x_{\alpha}\}_{\alpha \in \mathcal{I}} \subset M$  such that there exists a partition  $\{\mathcal{I}_i\}_{i=1}^{\mathfrak{D}_n}$  of  $\mathcal{I}$  so that

- $M \subset \bigcup_{\alpha \in \mathcal{I}} B(x_{\alpha}, R(h)),$
- $B(x_{\alpha_1}, 3R(h)) \cap B(x_{\alpha_2}, 3R(h)) = \emptyset, \qquad \alpha_1, \alpha_2 \in \mathcal{I}_i, \quad \alpha_1 \neq \alpha_2.$

Then, for  $m \in \mathbb{Z}$  define

$$\mathcal{I}_{k,m} := \left\{ \alpha \in \mathcal{I}_U : \quad 2^{m-1} \le h^{\frac{n-1}{2}} R(h)^{\frac{1-n}{2}} 2^k \frac{\|w_k\|_{L^{\infty}(B(x_{\alpha}, R(h)))}}{\|u\|_{P,T}} \le 2^m \right\},$$
(3.19)

where  $\mathcal{I}_U := \{ \alpha \in \mathcal{I} : B(x_\alpha, R(h)) \cap U \neq \emptyset \}$ . For each  $k \in \mathbb{Z}_+$  and  $\alpha \in \mathcal{I}$  consider the sets

$$\mathcal{A}_k(\alpha) := \{ j \in \mathcal{A}_k: \ \pi_{_M}(\mathcal{T}_j) \cap B(x_\alpha, 2R(h)) \neq \emptyset \},$$

where  $\pi_M : T^*M \to M$  is the standard projection. The indices in  $\mathcal{A}_k$  are those that correspond to tubes with mass comparable to  $\frac{1}{2^k} \|u\|_{P,T}$ , while indices in  $\mathcal{A}_k(\alpha)$  correspond to tubes of mass  $\frac{1}{2^k} \|u\|_{P,T}$  that run over the ball  $B(x_{\alpha}, 2R(h))$ . In particular, Lemma 3.2 and [CG20, Lemma 3.7] yield the existence of  $C_n, c_M > 0$  such that

$$c_M 2^m \le |\mathcal{A}_k(\alpha)| \le C_n 2^{2k}, \qquad \alpha \in \mathcal{I}_{k,m}.$$
(3.20)

Indeed, for  $\alpha \in \mathcal{I}_{k,m}$ ,

$$2^{m-1}h^{\frac{1-n}{2}}R(h)^{\frac{1-n}{2}}2^{-k}\|u\|_{P,T} \le \|w_k\|_{L^{\infty}(B(x_{\alpha},R(h)))}.$$
(3.21)

In addition, (3.14) and Lemma [CG20, Lemma 3.7] imply that there exist  $c_M > 0$ ,  $\tau_M > 0$ , and  $C_n > 0$ , depending on M and n respectively, such that for all N > 0 there exists  $C_N > 0$  with

$$\begin{split} \|w_k\|_{L^{\infty}(B(x_{\alpha},R(h)))} &\leq \frac{C_n R(h)^{\frac{n-1}{2}}}{\tau_M^{1/2} h^{\frac{n-1}{2}} \sum_{j \in \mathcal{A}_k(\alpha)}} \|Op_h(\tilde{\chi}_{\tau_j})Op_h(\psi)u\|_{L^2} + h^{-1} \|Op_h(\tilde{\chi}_{\tau_j})POp_h(\psi)u\|_{L^2} + C_N h^N \|u\|_{P,T} \\ &\leq c_M^{-1} h^{-\frac{n-1}{2}} R(h)^{\frac{n-1}{2}} 2^{-k} \|u\|_{P,T} |\mathcal{A}_k(\alpha)| + C_N h^N \|u\|_{P,T}, \end{split}$$

which, combined with (3.21), proves the lower bound in (3.20). To simplify notation, let

$$\mathcal{A}_{k,m} := \bigcup_{\alpha \in \mathcal{I}_{k,m}} \mathcal{A}_k(\alpha).$$
(3.22)

Note that for each  $\alpha \in \mathcal{I}_{k,m}$  there is  $\tilde{x}_{\alpha} \in B(x_{\alpha}, R(h))$  such that

$$|w_k(\tilde{x}_{\alpha})| \ge 2^{m-1} h^{\frac{1-n}{2}} R(h)^{\frac{n-1}{2}} 2^{-k} ||u||_{P,T}.$$
(3.23)

We finish this section with a result that controls the size of  $\mathcal{I}_{k,m}$  in terms of that of  $\mathcal{A}_{k,m}$ . Let

$$\frac{1}{2}(\delta_2 + 1) < \rho < 1, \tag{3.24}$$

 $0 < \varepsilon < \delta, \ \tilde{\chi} \in C_c^{\infty}((-1,1)),$  and define the operator  $\chi_{h,\tilde{x}_{\alpha}}$  by

$$\chi_{h,\tilde{x}_{\alpha}}u(x) := \tilde{\chi}(\frac{1}{\varepsilon}h^{-\rho}d(x,\tilde{x}_{\alpha})) \ [Op_h(\tilde{\chi}(\frac{1}{\varepsilon}(|\xi|_g-1)))u](x).$$

In Lemma 6.1 we prove that  $\chi_{h,\tilde{x}_{\alpha}} \in \Psi_{\Gamma_{\tilde{x}_{\alpha}},L_{\tilde{x}_{\alpha}},\rho}^{-\infty}$ , where

$$\Omega_{\tilde{x}_{\alpha}} = \{ \xi \in T^*_{\tilde{x}_{\alpha}} M : |1 - |\xi|_{g(\tilde{x}_{\alpha})}| < \delta \}, \qquad \Gamma_{\tilde{x}_{\alpha}} = \bigcup_{|t| < \frac{1}{2} \operatorname{inj}(M)} \varphi_t(\Omega_{\tilde{x}_{\alpha}})$$

and  $\Psi_{\Gamma_{\tilde{x}_{\alpha}},L_{\tilde{x}_{\alpha}},\rho}^{-\infty}$  is a class of smoothing pseudodifferential operators that allows for localization to  $h^{\rho}$  neighborhoods of  $\Gamma_{\tilde{x}_{\alpha}}$  and is compatible with localization to  $h^{\rho}$  neighborhoods of the foliation  $L_{\tilde{x}_{\alpha}}$  of  $\Gamma_{\tilde{x}_{\alpha}}$  generated by  $\Omega_{\tilde{x}_{\alpha}}$ .

In Theorem 4 for  $\varepsilon > 0$  we explain how to build a cut-off operator  $X_{\tilde{x}_{\alpha}} \in \Psi_{\Gamma_{\tilde{x}_{\alpha}}, L_{\tilde{x}_{\alpha}}, \rho}^{-\infty}$  such that

$$\begin{cases} \chi_{h,\tilde{x}_{\alpha}} X_{\tilde{x}_{\alpha}} = \chi_{h,\tilde{x}_{\alpha}} + O(h^{\infty})_{\Psi^{-\infty}}, \\ WF_{h}'([P, X_{\tilde{x}_{\alpha}}]) \cap \{(x, \xi) : x \in B(\tilde{x}_{\alpha}, \frac{1}{2} \text{inj} M), \ \xi \in \Omega_{x} \} = \emptyset, \end{cases}$$
(3.25)

where  $\operatorname{inj} M$  denotes the injectivity radius of M.

**Lemma 3.3.** Let  $\frac{1}{2}(\delta_2+1) < \rho \leq 1$ . There exists C > 0 so that for every  $k \geq -1$  and  $m \in \mathbb{Z}$  the following holds. If

$$|\mathcal{A}_{k,m}| \leq C \, 2^{2m} R(h)^{n-1} \left( h^{\rho - \frac{1}{2}} R(h)^{-\frac{1}{2}} \right)^{-\frac{2n(n-1)}{3n+1}},$$

$$|\mathcal{I}_{k,m}| \leq C |\mathcal{A}_{k,m}| 2^{-2m} R(h)^{1-n}.$$
(3.26)

then

$$|\mathcal{I}_{k,m}| \le C |\mathcal{A}_{k,m}| 2^{-2m} R(h)^{1-n}.$$
 (3.26)

*Proof.* We claim that by (3.17), for  $\alpha \in \mathcal{I}_{k,m}$ ,

$$\chi_{h,\tilde{x}_{\alpha}}w_{k} = \chi_{h,\tilde{x}_{\alpha}}w_{k,m} + O(h^{\infty} \|u\|_{L^{2}}), \qquad w_{k,m} := \sum_{j \in \mathcal{A}_{k,m}} Op_{h}(\tilde{\chi}_{\tau_{j}})Op_{h}(\psi)u.$$
(3.27)

Indeed, it suffices to show that  $\chi_{h,\tilde{x}_{\alpha}}Op_h(\tilde{\chi}_{\tau_j})Op_h(\psi)u = O(h^{\infty}||u||_{L^2})$  for  $\alpha \in \mathcal{I}_{k,m}$  and  $j \notin \mathcal{A}_{k,m}$ . Note that for such indices  $\pi_M(\mathcal{T}_j) \cap B(\tilde{x}_{\alpha}, 2R(h)) = \emptyset$  while

$$\operatorname{supp} \tilde{\chi}(\tfrac{1}{\varepsilon}h^{-\rho}d(x,\tilde{x}_{\alpha})) \subset B(\tilde{x}_{\alpha},C\varepsilon h^{\rho}) \subset B(x_{\alpha},\tfrac{3}{2}R(h))$$

for some C > 0 and all h small enough.

Our next goal is to produce a lower bound for  $|\mathcal{A}_{k,m}|$  in terms of  $|\mathcal{I}_{k,m}|$  by using the lower bound (3.23) on  $\|\chi_{h,\tilde{z}_{\alpha}} w_{k,m}\|_{L^{\infty}}$  for indices  $\alpha \in \mathcal{I}_{k,m}$ . By (3.25), we have

$$\chi_{h,\tilde{x}_{\alpha}}w_{k,m} = \chi_{h,\tilde{x}_{\alpha}}X_{\tilde{x}_{\alpha}}w_{k,m} + O(h^{\infty})_{L^{\infty}},$$

for  $\alpha \in \mathcal{I}_{k,m}$ . In particular, by (3.23) and (3.27),

$$2^{m-1}h^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}2^{-k}\|u\|_{P,T} \le \|\chi_{h,\tilde{x}_{\alpha}}w_{k}\|_{L^{\infty}} \le \|X_{\tilde{x}_{\alpha}}w_{k,m}\|_{L^{\infty}} + O(h^{\infty})\|u\|_{P,T}).$$
(3.28)

Therefore, applying the standard  $L^{\infty}$  bound for quasimodes of the Laplacian (see e.g. [Zwo12, Theorem 7.12]) and using that by (3.25) we have that  $X_{\tilde{x}_{\alpha}}$  nearly commutes with P on  $B(\tilde{x}_{\alpha}, \frac{1}{2} \text{inj } M)$ ,

$$2^{m-1}R(h)^{\frac{n-1}{2}}2^{-k}\|u\|_{P,T} \leq C(\|X_{\tilde{x}_{\alpha}}w_{k,m}\|_{L^{2}} + h^{-1}\|PX_{\tilde{x}_{\alpha}}w_{k,m}\|_{L^{2}(B)}) + O(h^{\infty}\|u\|_{P,T}).$$

$$\leq C(\|X_{\tilde{x}_{\alpha}}w_{k,m}\|_{L^{2}} + h^{-1}\|X_{\tilde{x}_{\alpha}}Pw_{k,m}\|_{L^{2}}) + O(h^{\infty}\|u\|_{P,T}).$$
(3.29)

Note that we have canceled the factor  $h^{\frac{1-n}{2}}$  which appears both in (3.28) and the standard  $L^{\infty}$  bounds for quasimodes. Using that  $h^{2\rho-1}R(h)^{-1} = o(1)$ , Proposition 6.4 proves that for all  $\tilde{\mathcal{I}} \subset \mathcal{I}_{k,m}$  and  $v \in L^2(M)$ 

$$\sum_{\alpha \in \tilde{I}} \|X_{\tilde{x}_{\alpha}}v\|_{L^{2}}^{2} \leq C\left(1 + a_{h}|\tilde{I}|^{\frac{3n+1}{2n}}\right) \|v\|_{L^{2}}^{2},$$

where  $a_h = (h^{\rho - \frac{1}{2}} R(h)^{-\frac{1}{2}})^{n-1}$ . As a consequence, (3.29) gives

$$\begin{split} |\tilde{\mathcal{I}}|R(h)^{n-1}2^{-2k}2^{2(m-1)}||u||_{P,T}^2 &\leq C\Big(\sum_{\alpha\in\tilde{I}}||X_{\tilde{x}_{\alpha}}w_{k,m}||_{L^2}^2 + h^{-2}\sum_{\alpha\in\tilde{I}}||X_{\tilde{x}_{\alpha}}Pw_{k,m}||_{L^2}^2\Big)\\ &\leq C\Big(1+a_h|\tilde{\mathcal{I}}|^{\frac{3n+1}{2n}}\Big)(||w_{k,m}||_{L^2}^2 + h^{-2}||Pw_{k,m}||_{L^2}^2)\\ &\leq C\Big(1+a_h|\tilde{\mathcal{I}}|^{\frac{3n+1}{2n}}\Big)2^{-2k}|\mathcal{A}_{k,m}|||u||_{P,T}^2. \end{split}$$

The last inequality follows from the definition of  $w_{k,m}$  together with the definition (3.14) of  $\mathcal{A}_k$ .

In particular, we have proved that there is C > 0 such that for all  $\tilde{\mathcal{I}} \subset \mathcal{I}_{k,m}$ 

$$\left|\tilde{\mathcal{I}}\right|R(h)^{n-1}2^{2m} \le C \max\left(1, a_h |\tilde{\mathcal{I}}|^{\frac{3n+1}{2n}}\right) |\mathcal{A}_{k,m}|.$$
(3.30)

Suppose that  $a_h |\mathcal{I}_{k,m}|^{\frac{3n+1}{2n}} \geq 1$ . Then, there exists  $\tilde{\mathcal{I}} \subset \mathcal{I}_{k,m}$  such that  $a_h |\tilde{\mathcal{I}}|^{\frac{3n+1}{2n}} = 1$ . In particular,  $|\tilde{\mathcal{I}}|R(h)^{n-1}2^{2m} \leq C |\mathcal{A}_{k,m}|$ . This implies that if  $|\mathcal{A}_{k,m}| \leq \frac{1}{C}a_h^{-\frac{2n}{3n+1}}R(h)^{n-1}2^{2m}$ , then  $a_h |\mathcal{I}_{k,m}|^{\frac{3n+1}{2n}} \leq 1$  and so by (3.30)

$$|\mathcal{I}_{k,m}| R(h)^{n-1} 2^{2m} \le C |\mathcal{A}_{k,m}|.$$

Note that for  $w_{k,m}$  defined as in (3.27),

$$\|w_k\|_{L^p(U)}^p \le \mathfrak{D}_n \sum_{m=-\infty}^{\infty} \|w_k\|_{L^p(U_{k,m})}^p = \mathfrak{D}_n \sum_{m=-\infty}^{\infty} \|w_{k,m}\|_{L^p(U_{k,m})}^p + O(h^{\infty} \|u\|_{P,T}),$$
(3.31)

where

$$U_{k,m} := \bigcup_{\alpha \in \mathcal{I}_{k,m}} B(x_{\alpha}, R(h)).$$
(3.32)

Finally, we split the study of  $||w_k||_{L^p(U)}$  into two regimes: tubes with low or high  $L^{\infty}$  mass. Fix N > 0 large, to be determined later. (Indeed, we will see that it suffices to take  $N > \frac{1}{2}(1 - \frac{p_c}{p})^{-1}$ .) Then, we claim that for each  $k \ge -1$ ,

$$\|w_k\|_{L^p(U)}^p \le \mathfrak{D}_n \sum_{m=-\infty}^{m_{1,k}} \|w_{k,m}\|_{L^p(U_{k,m})}^p + \mathfrak{D}_n \sum_{m=m_{1,k}+1}^{m_{2,k}} \|w_{k,m}\|_{L^p(U_{k,m})}^p + O(h^\infty \|u\|_{P,T}), \quad (3.33)$$

where  $m_{1,k}$  and  $m_{2,k}$  are defined by

$$2^{m_{1,k}} = \min\left(\frac{2^k R(h)^{\frac{1-n}{2}}}{T^N}, c_n 2^{2k}, c_0 R(h)^{1-n}\right), \quad 2^{m_{2,k}} = \min\left(c_n 2^{2k}, c_0 R(h)^{1-n}\right),$$

where  $c_0, c_n$  are described in what follows. Indeed, note that the bound (3.20) yields that  $2^m$  is bounded by  $|\mathcal{A}_k(\alpha)|$  for all  $\alpha \in \mathcal{I}_{k,m}$  and the latter is controlled by  $c_0 R(h)^{n-1}$  for some  $c_0 > 0$ , depending only on (M, g). Also, note that by (3.20) the  $w_{k,m}$  are only defined for m satisfying  $2^m \leq c_n 2^{2k}$ . These observations justify that the second sum in (3.33) runs only up to  $m_{2,k}$ .

3.4. Control of the low  $L^{\infty}$  mass term,  $m \leq m_{1,k}$ . We first estimate the small m term in (3.33). The estimates here essentially amount to interpolation between  $L^{p_c}$  and  $L^{\infty}$ . From the definition (3.19) of  $\mathcal{I}_{km}$ , together with  $\frac{1-n}{2}(p-p_c)-1 = -p\delta(p)$  and  $\|w_{k,m}\|_{L^{p_c}(U_{k,m})} \leq h^{-\frac{1}{p_c}}\|u\|_{P,T}$ ,

$$\sum_{m=-\infty}^{m_{1,k}} \|w_{k,m}\|_{L^{p}(U_{k,m})}^{p} \leq C \sum_{m=-\infty}^{m_{1,k}} \|w_{k,m}\|_{L^{\infty}(U_{k,m})}^{p-p_{c}} \|w_{k,m}\|_{L^{p_{c}}(U_{k,m})}^{p_{c}}$$
$$\leq Ch^{-p\delta(p)}R(h)^{\frac{n-1}{2}(p-p_{c})}2^{-k(p-p_{c})} \sum_{m=-\infty}^{m_{1,k}} 2^{m(p-p_{c})} \|u\|_{P,T}^{p}$$
$$\leq Ch^{-p\delta(p)}R(h)^{\frac{n-1}{2}(p-p_{c})}2^{(m_{1,k}-k)(p-p_{c})} \|u\|_{P,T}^{p}.$$

It follows that

$$\sum_{k\geq -1} \left( \sum_{m=-\infty}^{m_{1,k}} \|w_{k,m}\|_{L^p(U_{k,m})}^p \right)^{\frac{1}{p}} \leq Ch^{-\delta(p)} R(h)^{\frac{n-1}{2}(1-\frac{p_c}{p})} \|u\|_{P,T} \sum_{k\geq -1} 2^{(m_{1,k}-k)(1-\frac{p_c}{p})}.$$
 (3.34)

Finally, define  $k_1, k_2$  such that

$$2^{k_1} = \frac{R(h)^{\frac{1-n}{2}}}{c_n T^N}, \qquad 2^{k_2} = c_0 R(h)^{\frac{1-n}{2}} T^N.$$
(3.35)

If  $k \leq k_1$ , then  $2^{m_{1,k}} = c_n 2^{2k}$ , so there exists  $C_{n,p} > 0$  such that

$$\sum_{k=-1}^{k_1} 2^{(m_{1,k}-k)(1-\frac{p_c}{p})} \le C_{n,p} \frac{R(h)^{\frac{1-n}{2}(1-\frac{p_c}{p})}}{T^{N(1-\frac{p_c}{p})}}.$$

If  $k_1 \leq k \leq k_2$ , then  $2^{m_{1,k}} = \frac{2^k R(h)^{\frac{1-n}{2}}}{T^N}$ . Therefore, since  $|k_2 - k_1| \leq cN \log T$  for some c > 0, there exists C > 0 such that

$$\sum_{k=k_1}^{k_2} 2^{(m_{1,k}-k)(1-\frac{p_c}{p})} \le CN \log T \frac{R(h)^{\frac{1-n}{2}(1-\frac{p_c}{p})}}{T^{N(1-\frac{p_c}{p})}}.$$

Last, if  $k \ge k_2$ , then  $2^{m_{1,k}} = c_0 R(h)^{1-n}$ , so there exists  $C_p > 0$  such that

$$\sum_{k=k_2}^{\infty} 2^{(m_{1,k}-k)(1-\frac{p_c}{p})} \le C_p \frac{R(h)^{\frac{1-n}{2}(1-\frac{p_c}{p})}}{T^{N(1-\frac{p_c}{p})}}.$$

Putting these three bounds together with (3.34), we obtain

$$\sum_{k \ge -1} \left( \sum_{m = -\infty}^{m_{1,k}} \|w_{k,m}\|_{L^p(U_{k,m})}^p \right)^{\frac{1}{p}} \le Ch^{-\delta(p)} \frac{N\log T}{T^{N(1-\frac{p_c}{p})}} \|u\|_{P,T}.$$
(3.36)

3.5. Control of the high  $L^{\infty}$  mass term,  $m \ge m_{1,k}$ . In this section we estimate the large m term in (3.33). To do this we split

$$\mathcal{A}_{k,m} = \mathcal{G}_{k,m} \sqcup \mathcal{B}_{k,m},$$

where the set of 'good' tubes  $\bigcup_{j \in \mathcal{G}_{k,m}} \mathcal{T}_j$  is  $[t_0, T]$  non-self looping and the number of 'bad' tubes  $|\mathcal{B}_{k,m}|$  is small. To do this, let

$$\mathcal{B}_{U}(\alpha,\beta) := \left\{ j \in \bigcup_{k} \mathcal{A}_{k}(\alpha) : \bigcup_{t=t_{0}}^{T} \varphi_{t}(\mathcal{T}_{j}) \cap S^{*}_{B(x_{\beta},2R(h))}M \neq \emptyset \right\}.$$
(3.37)

Then, we define

$$\mathcal{B}_{k,m} := \bigcup_{\alpha,\beta \in \mathcal{I}_{k,m}} \mathcal{B}_U(\alpha,\beta) \cap \mathcal{A}_k(\alpha).$$

Let  $\mathcal{G}_{k,m} := \mathcal{A}_{k,m} \setminus \mathcal{B}_{k,m}$ . Then, by construction,  $\bigcup_{j \in \mathcal{G}_{k,m}} \mathcal{T}_j$  is  $[t_0, T]$  non-self looping and we have  $|\mathcal{B}_{k,m}| \le c |\mathcal{I}_{k,m}|^2 |\mathcal{B}_{U}|$ (3.38)

for some c > 0, where

$$|\mathcal{B}_{U}| := \sup\{|\mathcal{B}_{U}(\alpha,\beta)|: \ \alpha,\beta \in \mathcal{I}\},\tag{3.39}$$

That is,  $|\mathcal{B}_U|$  is the maximum number of loops of length in  $[t_0, T]$  joining any two points in U. Then, define

$$w_{k,m}^{\mathcal{G}} := \sum_{j \in \mathcal{G}_{k,m}} Op_h(\tilde{\chi}_{\tau_j}) Op_h(\psi) u, \qquad w_{k,m}^{\mathcal{B}} := \sum_{j \in \mathcal{B}_{k,m}} Op_h(\tilde{\chi}_{\tau_j}) Op_h(\psi) u.$$
(3.40)

Next, consider

$$\left(\sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}\|_{L^{p}(U_{k,m})}^{p}\right)^{\frac{1}{p}} \leq \left(\sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{p}(U_{k,m})}^{p}\right)^{\frac{1}{p}} + \left(\sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^{p}(U_{k,m})}^{p}\right)^{\frac{1}{p}}.$$
 (3.41)

3.5.1. Bound on the looping piece. We start by estimating the 'bad' piece

$$\sum_{k \ge -1} \left( \sum_{m=m_{1,k}}^{m_{2,k}} \| w_{k,m}^{\mathcal{B}} \|_{L^{p}(U_{k,m})}^{p} \right)^{\frac{1}{p}}.$$

Observe that if  $2^{m_{1,k}} = \min(c_0 R(h)^{1-n}, c_n 2^{2k})$ , then  $m_{1,k} = m_{2,k}$  and we need not consider this part of the sum. Therefore, the high  $L^{\infty}$  mass term has

$$2^{m_{1,k}} = \frac{2^k R(h)^{\frac{1-n}{2}}}{T^N}$$
(3.42)

and  $k_1 \leq k \leq k_2$ . Hence, for  $m_{1,k} < m \leq m_{2,k}$ , Lemma 3.2 gives that there is  $C_n > 0$  with

$$|\mathcal{A}_{k,m}| \le C_n 2^{2k} \le C_n R(h)^{n-1} 2^{2m} T^{2N}.$$

Furthermore, since  $R(h) \ge h^{\delta_2}$  with  $\delta_2 < \frac{1}{2}$ , (3.24) yields that there is  $\varepsilon = \varepsilon(n, N) > 0$  such that  $h^{\rho-\frac{1}{2}}R(h)^{-\frac{1}{2}} < h^{\varepsilon}$ , and hence, since  $T = O(\log h^{-1})$ ,

$$|\mathcal{A}_{k,m}| = o\left(R(h)^{n-1}2^{2m}\left(h^{\rho-\frac{1}{2}}R(h)^{-\frac{1}{2}}\right)^{-\frac{2n(n-1)}{3n+1}}\right).$$

In particular, a consequence of Lemma 3.3 is the existence of  $h_0 > 0$  and C > 0 such that

$$\left|\mathcal{I}_{k,m}\right| \le CR(h)^{1-n} 2^{-2m} \left|\mathcal{A}_{k,m}\right| \tag{3.43}$$

$$\leq CR(h)^{1-n}2^{2k-2m},\tag{3.44}$$

for all  $0 < h \le h_0$ , where we have used again Lemma 3.2 to bound  $|\mathcal{A}_{k,m}|$ .

Next, note that for each point in  $\mathcal{I}_{k,m}$  there are at most  $c|\mathcal{I}_{k,m}||\mathcal{B}_U|$  tubes in  $\mathcal{B}_{k,m}$  touching it. Therefore, we may apply [CG20, Lemma 3.7] to obtain C > 0 such that

$$\|w_{k,m}^{\mathcal{B}}\|_{L^{\infty}(U_{k,m})} \le Ch^{\frac{1-n}{2}} R(h)^{\frac{n-1}{2}} |\mathcal{I}_{k,m}| |\mathcal{B}_{U}| 2^{-k} \|u\|_{P,T}.$$
(3.45)

Using (3.45) and interpolating between  $L^{\infty}$  and  $L^{p_c}$  we obtain

$$\|w_{k,m}^{\mathcal{B}}\|_{L^{p}(U_{k,m})}^{p} \leq Ch^{-p\delta(p)} \left(R(h)^{\frac{n-1}{2}} |\mathcal{I}_{k,m}| |\mathcal{B}_{U}| 2^{-k} \|u\|_{P,T}\right)^{p-p_{c}} \|w_{k,m}^{\mathcal{B}}\|_{L^{2}(U_{k,m})}^{p_{c}}.$$
 (3.46)

In addition, since combining (3.14) with (3.38) yields

$$\|w_{k,m}^{\mathcal{B}}\|_{L^{2}(U_{k,m})} \leq C|\mathcal{B}_{k,m}|^{\frac{1}{2}}2^{-k}\|u\|_{P,T} \leq C2^{-k}|\mathcal{I}_{k,m}||\mathcal{B}_{U}|^{\frac{1}{2}}\|u\|_{P,T}$$

the bounds in (3.46) and (3.44), together with the definition of  $m_{1,k}$  (3.42) yield

$$\begin{split} \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^{p}(U_{k,m})}^{p} &\leq Ch^{-p\delta(p)}R(h)^{\frac{n-1}{2}(p-p_{c})}\sum_{m=m_{1,k}}^{m_{2,k}} |\mathcal{I}_{k,m}|^{p}|\mathcal{B}_{U}|^{p-\frac{p_{c}}{2}}2^{-kp}\|u\|_{P,T}^{p} \\ &\leq Ch^{-p\delta(p)}R(h)^{\frac{n-1}{2}(-p-p_{c})}2^{kp}|\mathcal{B}_{U}|^{p-\frac{p_{c}}{2}}\|u\|_{P,T}^{p}\sum_{m=m_{1,k}}^{m_{2,k}}2^{-2mp} \\ &\leq Ch^{-p\delta(p)}R(h)^{\frac{n-1}{2}(p-p_{c})}|\mathcal{B}_{U}|^{p-\frac{p_{c}}{2}}T^{2Np}2^{-kp}\|u\|_{P,T}^{p}. \end{split}$$

Then, with  $k_1, k_2$  defined as in (3.35), we have that

$$\begin{split} \sum_{k=k_1}^{k_2} \Big( \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^p(U_{k,m})}^p \Big)^{\frac{1}{p}} &\leq Ch^{-\delta(p)} R(h)^{\frac{n-1}{2}(1-\frac{p_c}{p})} |\mathcal{B}_U|^{1-\frac{p_c}{2p}} T^{2N} \|u\|_{P,T} \sum_{k=k_1}^{k_2} 2^{-k} \\ &\leq Ch^{-\delta(p)} (R(h)^{n-1} |\mathcal{B}_U|)^{1-\frac{p_c}{2p}} T^{3N} \|u\|_{P,T}. \end{split}$$

Finally, since we only need to consider  $k_1 \leq k \leq k_2$ ,

$$\sum_{k\geq -1} \left( \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^{p}(U_{k,m})}^{p} \right)^{\frac{1}{p}} \leq Ch^{-\delta(p)} (R(h)^{n-1} |\mathcal{B}_{U}|)^{1-\frac{p_{c}}{2p}} T^{3N} \|u\|_{P,T}.$$
(3.47)

3.5.2. Bound on the non self-looping piece. In this section we aim to control the 'good' piece

$$\sum_{k \ge -1} \left( \sum_{m=m_{1,k}}^{m_{2,k}} \| w_{k,m}^{\mathcal{G}} \|_{L^{p}(U_{k,m})}^{p} \right)^{\frac{1}{p}}.$$
(3.48)

So far all  $L^p$  bounds appearing have been  $\ll h^{\frac{1-n}{2}}/\sqrt{T}$ . The reason for this is that the bounds were obtained by interpolation with an  $L^{\infty}$  estimate which is substantially stronger than  $h^{\frac{1-n}{2}}/\sqrt{T}$ .

We now estimate the number of non-self looping tubes  $\mathcal{T}_j$  with  $j \in \mathcal{A}_k$ . That is, tubes on which the  $L^2$  mass of u is comparable  $2^{-k} ||u||_{P,T}$ .

**Lemma 3.4.** Let  $k \in \mathbb{Z}$ ,  $k \geq -1$ , and  $t_0 > 1$ . Suppose that  $\mathcal{G} \subset \mathcal{A}_k$  is such that

$$\bigcup_{j \in \mathcal{G}} \mathcal{T}_j \quad is \ [t_0, T] \ non-self \ looping.$$

Then, there exists a constant  $C_n > 0$ , depending only on n, such that  $|\mathcal{G}| \leq \frac{C_n t_0}{T} 2^{2k}$ .

*Proof.* Using that  $\mathcal{G} \subset \mathcal{A}_k$ , we have

$$|\mathcal{G}|\frac{\|u\|_{P,T}^2}{2^{2(k+1)}} \le 2\sum_{j\in\mathcal{G}} \Big(\|Op_h(\chi_{\tau_j})u\|^2 + h^{-2}\|Op_h(\chi_{\tau_j})Pu\|_{L^2}^2\Big).$$
(3.49)

Since  $\{\mathcal{T}_j\}_{j\in\mathcal{G}}$  is  $(\mathfrak{D}_n, \tau, R(h))$ -good, there are  $\{\mathcal{G}_i\}_{i=1}^{\mathfrak{D}_n} \subset \mathcal{G}$ , such that for each  $i = 1, \ldots, \mathfrak{D}_n$ ,

$$\mathcal{T}_j \cap \mathcal{T}_k = \emptyset, \qquad j, k \in \mathcal{G}_i, \quad j \neq k.$$

By [CG20, Lemma 4.1] with  $t_{\ell} = t_0$  and  $T_{\ell} = T$  for all  $\ell$ ,

$$\sum_{j \in \mathcal{G}} \|Op_h(\chi_{\tau_j})u\|_{L^2}^2 \le \sum_{i=1}^{\mathfrak{D}_n} \sum_{j \in \mathcal{G}_i} \|Op_h(\chi_{\tau_j})u\|_{L^2}^2 \le \frac{\mathfrak{D}_n 4t_0}{T} \|u\|_{P,T}^2.$$
(3.50)

On the other hand, since  $\sum_{j \in \mathcal{G}_i} \|Op_h(\chi_{\tau_j})\|^2 \leq 2$  for each i,

$$\sum_{j \in \mathcal{G}} \|Op_h(\chi_{\tau_j})Pu\|_{L^2}^2 \le 2\mathfrak{D}_n \|Pu\|_{L^2}^2.$$
(3.51)

Combining (3.49), (3.50), and (3.51) yields

$$|\mathcal{G}|\frac{\|u\|_{P,T}^2}{2^{2(k+1)}} \le \frac{8\mathfrak{D}_n t_0}{T} \|u\|_{P,T}^2 + \frac{4\mathfrak{D}_n}{h^2} \|Pu\|_{L^2}^2 \le \frac{8\mathfrak{D}_n t_0 + \frac{4\mathfrak{D}_n}{T}}{T} \|u\|_{P,T}^2.$$

We may now proceed to estimate the  $L^p$ -norm of the non-looping piece (3.48). The first step is to notice that we only need to sum up to  $m \leq m_{3,k}$ , where  $m_{3,k}$  is defined by

$$2^{m_{3,k}} := \min\left(\frac{C_n t_0 2^{2k}}{c_M T}, \ c_0 R(h)^{1-n}\right),$$

and  $c_M > 0$  is as defined in (3.20) and  $C_n > 0$  is the constant in Lemma 3.4. To see this, first observe that, using (3.19), (3.43) and (3.45), for each  $\alpha \in \mathcal{I}_{k,m}$ 

$$\begin{aligned} \|w_{k,m}^{\mathcal{G}}\|_{L^{\infty}(B((x_{\alpha},R(h))))} &\leq \|w_{k,m}\|_{L^{\infty}(B((x_{\alpha},R(h))))} + \|w_{k,m}^{\mathcal{B}}\|_{L^{\infty}(B((x_{\alpha},R(h))))} \\ &\leq C(2^{m} + |\mathcal{I}_{k,m}||\mathcal{B}_{U}|)2^{-k}h^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}\|u\|_{P,T} \\ &\leq C(1 + R(h)^{1-n}2^{-3m}|\mathcal{A}_{k,m}||\mathcal{B}_{U}|)2^{m-k}h^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}\|u\|_{P,T}. \end{aligned}$$
(3.52)

Furthermore, since  $|\mathcal{G}_{k,m}| \ge |\mathcal{A}_{k,m}| - |\mathcal{I}_{k,m}|^2 |\mathcal{B}_U|$  and  $\mathcal{G}_{k,m}$  is  $[t_0, T]$  non-self looping, Lemma 3.4 yields the existence of  $C_n > 0$  such that

$$|\mathcal{A}_{k,m}| - |\mathcal{I}_{k,m}|^2 |\mathcal{B}_U| \le C_n \frac{t_0}{T} 2^{2k}$$

Next, since  $m_{1,k} \leq m \leq m_{2,k}$ , we may apply Lemma 3.3 to bound  $|\mathcal{I}_{k,m}|$  as in (3.43) to obtain that for some C > 0

$$|\mathcal{A}_{k,m}|(1 - CR(h)^{2(1-n)}2^{-4m}|\mathcal{A}_{k,m}||\mathcal{B}_{U}|) \le C_n \frac{t_0}{T} 2^{2k}.$$
(3.53)

In addition, provided

$$|\mathcal{B}_U| R(h)^{n-1} \ll T^{-6N},$$
 (3.54)

we have, that for  $m \ge m_{1,k}$  and  $k_1 \le k \le k_2$ 

$$R(h)^{2(1-n)}2^{-4m} |\mathcal{A}_{k,m}| |\mathcal{B}_{U}| \le R(h)^{2(1-n)}2^{-4m+2k} |\mathcal{B}_{U}| \le 2^{-2k}T^{4N} |\mathcal{B}_{U}| \le R(h)^{n-1}T^{6N} |\mathcal{B}_{U}| \ll 1,$$
(3.55)

where we used that by (3.20),  $|A_{k,m}|$  is controlled by  $2^{2k}$  to get the first inequality, that  $m \ge m_{1,k}$  to get the second, and that  $k \ge k_1$  to get the third. Combining (3.53) and the bound in (3.55) we obtain  $|\mathcal{A}_{k,m}| \le C_n \frac{t_0 2^{2k}}{T}$ , and so, by (3.20),  $2^m \le C_n \frac{t_0 2^{2k}}{c_M T}$ . As claimed, this shows that to deal with (3.48) we only need to sum up to  $m \le m_{3,k}$ .

The next step is to use interpolation to control the first sum in (3.48) by

$$\sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{p}(U_{k,m})}^{p} = \sum_{m=m_{1,k}}^{m_{3,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{p}(U_{k,m})}^{p} \le \sum_{m=m_{1,k}}^{m_{3,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{\infty}(U_{k,m})}^{p-p_{c}} \|w_{k,m}^{\mathcal{G}}\|_{L^{p_{c}}(U_{k,m})}^{p_{c}}.$$
 (3.56)

We claim that (3.52) yields

$$\|w_{k,m}^{\mathcal{G}}\|_{L^{\infty}(B(x_{\alpha},R(h)))} \le C2^{m-k}h^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}\|u\|_{P,T}.$$
(3.57)

Indeed, using the bound (3.54) on  $|\mathcal{B}_U|$ , that  $|\mathcal{A}_{k,m}|$  is controlled by  $2^{2k}$ , that  $m \geq m_{1,k}$  as in (3.42), and that  $k_1 \leq k \leq k_2$ , we have

$$R(h)^{1-n}2^{-3m} |\mathcal{A}_{k,m}| |\mathcal{B}_{U}| \ll R(h)^{2(1-n)}2^{-3m+2k}T^{-6N} \le T^{-2N}$$

Using (3.57), the standard bound on  $\|w_{k,m}^{\mathcal{G}}\|_{L^{p_c}(U_{k,m})}^{p_c}$ , and  $\|w_{k,m}^{\mathcal{G}}\|_{L^2}^2 \leq C\frac{t_0}{T}$ , we obtain

$$\|w_{k,m}^{\mathcal{G}}\|_{L^{p}(U_{k,m})}^{2} \leq Ch^{-p\delta(p)}(R(h)^{\frac{n-1}{2}}2^{m-k})^{p-p_{c}}\frac{t_{0}^{\frac{p_{c}}{2p}}}{T^{\frac{p_{c}}{2p}}} + O(h^{\infty}\|u\|_{P,T}^{p}).$$
(3.58)

Using this, we estimate (3.56)

$$\sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{p}(U_{k,m})}^{p} \leq Ch^{-p\delta(p)} (R(h)^{\frac{n-1}{2}} 2^{(m_{3,k}-k)})^{p-p_{c}} \|u\|_{P,T}^{p} \frac{t_{0}^{\frac{p_{c}}{2}}}{T^{\frac{p_{c}}{2}}} + O(h^{\infty} \|u\|_{P,T}^{p}).$$
(3.59)

Then, summing in k, and again using that only  $k_1 \leq k \leq k_2$  contribute,

$$\sum_{k=-1}^{\infty} \left( \sum_{m=m_{1,k}}^{m_{2,k}} \| w_{k,m}^{\mathcal{G}} \|_{L^{p}(U_{m})}^{p} \right)^{\frac{1}{p}} \leq Ch^{-\delta(p)} \| u \|_{P,T} \frac{t_{0}^{\frac{pc}{2p}}}{T^{\frac{pc}{2p}}} \sum_{k=k_{1}}^{k_{2}} \left( R(h)^{\frac{n-1}{2}} 2^{(m_{3,k}-k)} \right)^{1-\frac{pc}{p}} + O(h^{\infty} \| u \|_{P,T})$$

$$\leq Ch^{-\delta(p)} \frac{t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} \| u \|_{P,T} + O(h^{\infty} \| u \|_{P,T}). \tag{3.60}$$

Note that the sum over k in (3.60) is controlled by the value of k for which  $\frac{C_n t_0 2^{2k}}{c_M T} = c_0 R(h)^{1-n}$ , since the sum is geometrically increasing before such k and geometrically decreasing afterward.

3.6. Wrapping up the proof of Theorem 3. Combining (3.36), (3.47), (3.60), with (3.41) and (3.33), and taking  $N > \frac{1}{2}(1 - \frac{p_c}{p})^{-1}$ , provided  $R(h)^{n-1}|\mathcal{B}_U| \leq CT^{-6N}$ , for some C > 0, we obtain

$$\|v\|_{L^{p}(U)} \leq \sum_{k=-1}^{\infty} \|w_{k}\|_{L^{p}(U)} \leq Ch^{-\delta(p)} \left(\frac{t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} + (R(h)^{n-1}|\mathcal{B}_{U}|)^{1-\frac{p_{c}}{2p}}T^{3N}\right) \|u\|_{P,T}$$

as requested in (3.54). Since this estimate holds only when  $|\mathcal{B}_U|R(h)^{n-1} \leq CT^{-6N}$ , we replace T by  $T_0 := \min\{\frac{1}{C}(R(h)^{n-1}|\mathcal{B}_U|)^{-\frac{1}{6N}}, T\}$  so that

$$\begin{aligned} \|v\|_{L^{p}(U)} &\leq Ch^{-\delta(p)} \left( \frac{t_{0}^{\frac{1}{2}}}{T_{0}^{\frac{1}{2}}} + (R(h)^{n-1}|\mathcal{B}_{U}|)^{1-\frac{p_{c}}{2p}} T_{0}^{3N} \right) \|u\|_{P,T} \\ &\leq Ch^{-\delta(p)} \left( \frac{t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} + t_{0}^{\frac{1}{2}} (R(h)^{n-1}|\mathcal{B}_{U}|)^{\frac{1}{12N}} + (R(h)^{n-1}|\mathcal{B}_{U}|)^{\frac{1}{2}(1-\frac{p_{c}}{p})} \right) \|u\|_{P,T} \\ &\leq Ch^{-\delta(p)} \left( \frac{t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} + (R(h)^{n-1}|\mathcal{B}_{U}|)^{\frac{1}{12N}} \right) \|u\|_{P,T}, \end{aligned}$$
(3.61)

where the constant C is adjusted from line to line.

Next, combining (3.61) with (3.11) and the definition of v in (3.12), we obtain

$$\|u\|_{L^{p}(U)} \leq Ch^{-\delta(p)} \left( \frac{t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} + (R(h)^{n-1}|\mathcal{B}_{U}|)^{\frac{1}{12N}} \right) \|u\|_{P,T} + Ch^{-\delta(p)+\frac{1}{2}-\delta_{2}}h^{-1}\|Pu\|_{H^{n(\frac{1}{2}-\frac{1}{p})+\varepsilon-2}_{h}}.$$

Putting  $\varepsilon = \frac{1}{2}$  and setting  $N = \frac{1}{2}(1 + \frac{\varepsilon_0}{6})(1 - \frac{p_c}{p})^{-1}$ , the estimate (1.4) will follow once we relate  $|\mathcal{B}_U|$  for a given  $(\tau, R(h))$  cover to  $|\mathcal{B}_U|$  for the  $(\mathfrak{D}, \tau, R(h))$  cover used in our proof.

Finally, to finish the proof of Theorem 3, we need to show that for any  $(\tau, R(h))$  cover  $\{\mathcal{T}_j\}_j$  of  $S^*M$ , up to a constant depending only on M,  $|\mathcal{B}_U|$  can be bounded by  $|\tilde{\mathcal{B}}_U|$  where  $\tilde{\mathcal{B}}_U$  is defined as in (3.39) using a  $(\mathfrak{D}, \tau, R(h))$  good cover  $\{\mathcal{T}_k\}_k$  of  $S^*M$ .

**Lemma 3.5.** There exists  $C_M > 0$  depending only on M so that if  $\{\mathcal{T}_j\}_{j \in \mathcal{J}}$  and  $\{\tilde{\mathcal{T}}_k\}_{k \in \mathcal{K}}$  are respectively a  $(\tau, R(h))$  cover  $S^*M$  and a  $(\tilde{\mathfrak{D}}, \tau, R(h))$  good cover of  $S^*M$ , and  $|\mathcal{B}_U|$ ,  $|\tilde{\mathcal{B}}_U|$  are defined as in (3.39) for respectively the covers  $\{\mathcal{T}_j\}_{j \in \mathcal{J}}, \{\tilde{\mathcal{T}}_k\}_k$ , then

$$|\mathcal{B}_U| \le C_M \mathfrak{D}|\mathcal{B}_U|.$$

*Proof.* Fix  $\alpha, \beta$  such that  $x_{\alpha}, x_{\beta} \in U$ . Suppose that  $j \in \mathcal{B}_{U}(\alpha, \beta)$  where  $\mathcal{B}_{U}(\alpha, \beta)$  is as in (3.37). Then, there is  $k \in \tilde{\mathcal{B}}_{U}(\alpha, \beta)$  such that  $\tilde{\mathcal{T}}_{k} \cap \mathcal{T}_{j} \neq \emptyset$ . Now, fix  $j \in \mathcal{J}$  and let

$$\mathcal{C}_j := \{k \in \mathcal{K} : \ \mathcal{T}_j \cap \tilde{\mathcal{T}}_k \neq \emptyset\}.$$

We claim that there is  $c_M > 0$  such that for each  $k \in C_j$ 

$$\tilde{\mathcal{T}}_k \subset \Lambda_{\rho_j}^{c_M \tau}(c_M R(h)).$$
(3.62)

Assuming (3.62) for now, there exists  $C_M > 0$  such that

$$|\mathcal{C}_j| \leq \tilde{\mathfrak{D}} \frac{\operatorname{vol}(\Lambda_{\rho_j}^{c_M \gamma}(c_M R(h)))}{\inf_{k \in \mathcal{K}} \operatorname{vol}(\tilde{\mathcal{T}}_k)} \leq \tilde{\mathfrak{D}} C_{\!_M}.$$

Thus, for each  $j \in \mathcal{B}_U(\alpha, \beta)$ , there are at most  $C_M \tilde{\mathfrak{D}}$  elements in  $\tilde{\mathcal{B}}_U(\alpha, \beta)$  and hence  $|\mathcal{B}_U(\alpha, \beta)| \geq |\tilde{\mathcal{B}}_U(\alpha, \beta)|/(C_M \tilde{\mathfrak{D}})$  as claimed.

We now prove (3.62). Let  $q \in \tilde{\mathcal{T}}_k$ . Then, there are  $\rho'_k, \rho'_j, q' \in S^*M$  and  $t_k, t_j, s \in [\tau - R(h), \tau + R(h)]$  such that

$$d(\rho_k, \rho'_k) < R(h), \qquad d(\rho_j, \rho'_j) < R(h), \qquad d(\rho_k, q') < R(h)$$
$$\varphi_{t_k}(\rho'_k) = \varphi_{t_j}(\rho'_j), \qquad \varphi_s(q') = q.$$

In particular,  $d(q', \rho'_k) < 2R(h)$ , so there is  $c_M > 0$  such that  $d(\varphi_{t_k}(\rho'_k), \varphi_{t_k-s}(q)) < c_M R(h)$ . Applying  $\varphi_{-t_j}$ , and adjusting  $c_M$  in a way depending only on M,  $d(\rho'_j, \varphi_{t_k-t_j-s}(q)) < c_M R(h)$ . In particular, adjusting  $c_M$  again,  $d(\rho_j, \varphi_{t_k-t_j-s}(q)) < c_M R(h)$  and the claim follows.

3.7. Proof of Theorem 2. As explained in the introduction, Theorem 2 actually holds under the more general assumptions of Theorem 3. Let  $p > p_c$  and assume that there is  $\delta > 0$  such that

$$T = T(h) \to \infty, \qquad \qquad |\mathcal{B}_U| R(h)^{n-1} T^{\frac{3p}{p-p_c}+\delta} = o(1).$$

In the general setup we work with

$$\mathcal{S}_{U}(h,\varepsilon,u) := \Big\{ \alpha \in \mathcal{I}(h) : \|u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \ge \frac{\varepsilon h^{\frac{1-n}{2}}\sqrt{t_{0}}}{\sqrt{T(h)}} \|u\|_{L^{2}(M)}, \ B(x_{\alpha},R(h)) \cap U \neq \emptyset \Big\}.$$

We proceed to prove Theorem 2 in this setup, using the decompositions introduced in the previous sections. Throughout this proof we assume that

$$\|Pu\|_{H_h^{\frac{n-3}{2}}} = o\left(\frac{h}{T}\|u\|_{L^2}\right).$$
(3.63)

3.7.1. Proof of the bound on  $|\mathcal{S}_{U}(h,\varepsilon,u)|$ . We claim that there is c > 0 such that for  $\alpha \in S_{U}(h,\varepsilon,u)$ 

$$\frac{c\varepsilon\sqrt{t_0}}{\sqrt{T}}h^{-\frac{1}{p}}\|u\|_{P,T} \le \|u\|_{L^p(B(x_\alpha,2R(h)))}.$$
(3.64)

To see (3.64), first let  $\chi_0, \chi_1 \in C_c^{\infty}(-2,2), \chi \equiv 1$  on  $[-3/2,3/2], \chi_1 \equiv 1$  on supp  $\chi_0$  and note that by Lemma 3.1, the elliptic parametrix construction for P, and (3.63)

$$\|(1-\chi_0(-h^2\Delta_g))u\|_{L^p} \le Ch^{-\delta(p)-\frac{1}{2}} \|Pu\|_{H_h^{\frac{n-3}{2}}} = o\Big(\frac{h^{-\delta(p)+\frac{1}{2}}}{T}\Big) \|u\|_{L^2}.$$
(3.65)

Therefore, for  $\alpha \in S_{U}(h, \varepsilon, u)$  we have

$$\|\chi_0(-h^2\Delta_g)u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \ge \frac{\varepsilon h^{\frac{1-n}{2}}}{2\sqrt{T}} \|u\|_{L^{2}(M)}$$
(3.66)

for h small enough. Next, set  $\chi_{\alpha,h}(x) := \chi(R(h)^{-1}d(x,x_{\alpha}))$  and note

$$\chi_1(-h^2\Delta_g)\chi_{\alpha,h}\chi_0(-h^2\Delta_g)u = \chi_{\alpha,h}\chi_0(-h^2\Delta_g)u + O(h^{\infty}||u||_{L^2})_{C^{\infty}}$$

Then, by (3.66) and [Zwo12, Theorem 7.15]

$$\frac{\varepsilon h^{\frac{1-n}{2}}}{2\sqrt{T}} \|u\|_{L^{2}(M)} \leq \|\chi_{0}(-h^{2}\Delta_{g})u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \leq \|\chi_{\alpha,h}\chi_{0}(-h^{2}\Delta_{g})u\|_{L^{\infty}(B(x_{\alpha},R(h)))} \leq Ch^{-\frac{n}{p}} \Big(\|\chi_{0}(-h^{2}\Delta_{g})u\|_{L^{p}(B(x_{\alpha},2R(h))} + O(h^{\infty})\|u\|_{L^{2}}\Big), \quad (3.67)$$

Combining (3.67) and (3.65) yields the claim in (3.64). It then follows that, if  $\{\alpha_i\}_{i=1}^N \subset S_U(h,\varepsilon,u)$  with  $B(x_{\alpha_i},2R(h)) \cap B(x_{\alpha_j},2R(h)) = \emptyset$  for  $i \neq j$ , then using Theorem 3,

$$N^{\frac{1}{p}} \frac{c\varepsilon\sqrt{t_0}}{\sqrt{T}} h^{-\frac{1}{p}} \|u\|_{P,T} \le \|u\|_{L^p} \le Ch^{-\frac{1}{p}} \|u\|_{L^2} \le Ch^{-\frac{1}{p}} \frac{\sqrt{t_0}}{\sqrt{T}} \|u\|_{P,T}$$

Then,  $N^{\frac{1}{p}} \leq C\varepsilon^{-1}$ . Since at most  $\mathfrak{D}_n$  balls  $B(x_{\alpha}, 2R(h))$  intersect,  $|S_U(h, \varepsilon, u)| \leq C\mathfrak{D}_n\varepsilon^{-p}$ .

3.7.2. Preliminaries for the decomposition of u. Let  $q \in \mathbb{R}$  such that  $p \leq q \leq \infty$ . Below, all implicit constants are uniform for  $p \leq q \leq \infty$ . As above, it suffices to prove the statement for v as in (3.12) instead of u. Then, we decompose  $v = \sum_{k=-1}^{\infty} w_k$  as in (3.18). For  $V \subset U$ , by the same analysis that led to (3.33),

$$\|w_k\|_{L^q(V)}^q \le \mathfrak{D}_n \sum_{m=-\infty}^{m_{2,k}} \|w_{k,m}\|_{L^q(V \cap U_{k,m})}^q + O(h^\infty) \|u\|_{P,T},$$

where  $w_{k,m}$  is as in (3.27). Then, by (3.36), with  $N = \frac{q}{2(q-p_c)} + \frac{\delta}{6}$ 

$$\sum_{k\geq -1} \left( \sum_{m=-\infty}^{m_{1},k} \|w_{k,m}\|_{L^{q}(U_{k,m})}^{q} \right)^{\frac{1}{q}} \leq Ch^{-\delta(q)} \frac{\log T}{T^{\frac{1}{2} + \frac{\delta(q-p_{c})}{6q}}} \|u\|_{P,T},$$
(3.68)

for h small enough. Then, splitting  $w_{k,m} = w_{k,m}^{\mathcal{B}} + w_{k,m}^{\mathcal{G}}$ , as in (3.40), we have by (3.47)

$$\sum_{k\geq -1} \Big( \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^q(U_{k,m})}^q \Big)^{\frac{1}{q}} \leq Ch^{-\delta(q)} (R(h)^{n-1} |\mathcal{B}_U|)^{1-\frac{p_c}{2q}} T^{\frac{3q}{2(q-p_c)} + \frac{\delta}{2}} \|u\|_{P,T}.$$
(3.69)

Define  $k_1^{\varepsilon}$  and  $k_2^{\varepsilon}$ , by

$$2^{2k_1^{\varepsilon}} = \frac{C^{-2}\mathfrak{D}_n^{-2}\varepsilon^2 R(h)^{1-n} c_M T}{4C_n t_0}, \qquad 2^{2k_2^{\varepsilon}} = \frac{C^2\mathfrak{D}_n^2\varepsilon^{-2} R(h)^{1-n} c_M T}{4C_n t_0}, \qquad (3.70)$$

where C is as in (3.60). Then, define  $\mathcal{K}(\varepsilon) := \{k : k_1^{\varepsilon} \leq k \leq k_2^{\varepsilon}\}$  and note that, since  $2^{(k_2^{\varepsilon} - k_1^{\varepsilon})} = C^2 \mathfrak{D}_n^2 \varepsilon^{-2}, |\mathcal{K}(\varepsilon)| \leq \log_2(4C^2 \mathfrak{D}_n^2 \varepsilon^{-2}) =: K_{\varepsilon}$ . Using (3.58) and summing over  $k \notin \mathcal{K}(\varepsilon)$ , it follows that we have

$$\sum_{k \notin \mathcal{K}(\varepsilon)} \left( \sum_{m=m_{1,k}}^{m_{3,k}} \| w_{k,m}^{\mathcal{G}} \|_{L^q(U_{k,m})}^q \right)^{\frac{1}{q}} \le \frac{\varepsilon}{4\mathfrak{D}_n} \frac{h^{-\delta(q)}\sqrt{t_0}}{\sqrt{T}} \| u \|_{P,T}.$$
(3.71)

Next, for  $k \in \mathcal{K}(\varepsilon)$  let

$$\mathcal{M}(k,\varepsilon) := \{m : m_{3,k}^{\varepsilon} \le m \le m_{3,k}\}, \qquad m_{3,k}^{\varepsilon} := m_{3,k} - \frac{p}{p-p_c} \log_2(\varepsilon^{-1} 2C\mathfrak{D}_n),$$

and note  $|\mathcal{M}(k,\varepsilon)| \leq \frac{p}{p-p_c} \log_2(\varepsilon^{-1} 2C\mathfrak{D}_n) := M_{\varepsilon}$ . Using (3.58) and summing over  $k \in \mathcal{K}(\varepsilon)$ ,  $m \notin \mathcal{M}(k,\varepsilon)$ , it follows that

$$\sum_{k\in\mathcal{K}(\varepsilon)} \left( \sum_{m\notin\mathcal{M}(k,\varepsilon)} \|w_{k,m}^{\mathcal{G}}\|_{L^{q}(U_{k,m})}^{q} \right)^{\frac{1}{q}} \leq Ch^{-\delta(q)} \frac{t_{0}^{\frac{p_{c}}{2q}}}{T^{\frac{p_{c}}{2q}}} \sum_{k\in\mathcal{K}(\varepsilon)} (R(h)^{\frac{n-1}{2}} 2^{m_{3,k}^{\varepsilon}-k})^{1-\frac{p_{c}}{q}} \|u\|_{P,T} + O(h^{\infty} \|u\|_{P,T}) \\
\leq \frac{\varepsilon}{4\mathfrak{D}_{n}} \frac{h^{-\delta(q)} t_{0}^{\frac{1}{2}}}{T^{\frac{1}{2}}} \|u\|_{P,T}.$$
(3.72)

Let

$$\mathcal{N}_{k,m}(\varepsilon) := \left\{ \alpha \in \mathcal{I}_{k,m} : \| w_{k,m}^{\mathcal{G}} \|_{L^{\infty}(B(x_{\alpha}, R(h)))} \ge \frac{\varepsilon}{4\mathfrak{D}_{n}M_{\varepsilon}K_{\varepsilon}} \frac{h^{\frac{1-n}{2}}\sqrt{t_{0}}}{\sqrt{T}} \| u \|_{P,T} \right\}.$$
(3.73)

We claim

$$\mathcal{S}_{U}(h,\varepsilon,u) \subset \bigcup_{k \in \mathcal{K}(\varepsilon)} \bigcup_{m \in \mathcal{M}(k,\varepsilon)} \mathcal{N}_{k,m}(\varepsilon).$$
(3.74)

To prove the claim (3.74), suppose  $\alpha \notin \bigcup_{k \in \mathcal{K}(\varepsilon)} \bigcup_{m \in \mathcal{M}(k,\varepsilon)} \mathcal{N}_{k,m}(\varepsilon)$ . Then, using (3.68) with  $q = \infty$  and  $N = \frac{1}{2} + \frac{\delta}{6}$ ,

$$\frac{1}{\mathfrak{D}_n} \|v\|_{L^{\infty}(B(x_{\alpha}, R(h)))} \le \frac{Ch^{\frac{1-n}{2}} \log T}{T^{\frac{1}{2} + \frac{\delta}{6}}} \|u\|_{P,T} + \sum_{k \ge -1} \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}\|_{L^{\infty}(U_{k,m})}.$$
(3.75)

Next, we decompose the second term in the RHS of (3.75) as

$$\sum_{k \ge -1} \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{B}}\|_{L^{\infty}(U_{k,m})} + \sum_{k \notin \mathcal{K}(\varepsilon)} \sum_{m=m_{1,k}}^{m_{3,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{\infty}(U_{k,m})} + \sum_{k \in \mathcal{K}(\varepsilon)} \sum_{m=m_{1,k}}^{m_{2,k}} \|w_{k,m}^{\mathcal{G}}\|_{L^{\infty}(U_{k,m})}$$
(3.76)

Note that in the term with the sum over  $k \notin \mathcal{K}(\varepsilon)$  we only sum in  $m \leq m_{3,k}$  for the same reason as in (3.56). We bound the three terms in (3.76) using (3.69), (3.71), (3.72), and (3.73), with  $q = \infty$  and  $N = \frac{1}{2} + \frac{\delta}{6}$ . Combining it with (3.75) this yields

$$\frac{1}{\mathfrak{D}_n} \|v\|_{L^{\infty}(B(x_{\alpha}, R(h)))} \le Ch^{\frac{1-n}{2}} \|u\|_{P, T} \Big( \frac{\log T}{T^{\frac{1}{2} + \frac{\delta}{6}}} + R(h)^{n-1} |\mathcal{B}_U| T^{\frac{3}{2} + \frac{\delta}{2}} + \frac{3\varepsilon}{4\mathfrak{D}_n} \frac{\sqrt{t_0}}{\sqrt{T}} + O(h^{\infty}) \Big).$$

Thus, if  $\alpha \notin \bigcup_{k \in \mathcal{K}(\varepsilon)} \bigcup_{m \in \mathcal{M}(k,\varepsilon)} \mathcal{N}_{k,m}(\varepsilon)$ , then  $\|v\|_{L^{\infty}(B(x_{\alpha},R(h)))} \leq \varepsilon h^{\frac{1-n}{2}} \frac{\sqrt{t_0}}{\sqrt{T}} \|u\|_{P,T}$  for h small enough. In particular,  $\alpha \notin \mathcal{S}_U(h,\varepsilon,u)$ . This proves the claim (3.74)

3.7.3. Decomposition of u. We next decompose u as described in the theorem. First, put

$$u_{e,1} := \sum_{k \ge -1} \sum_{m = -\infty}^{m_1, k} w_{k,m} + \sum_{k \ge -1} \sum_{m = m_{1,k}}^{m_{2,k}} w_{k,m}^{\mathcal{B}} + \sum_{k \notin \mathcal{K}(\varepsilon)} \sum_{m = m_{1,k}}^{m_{3,k}} w_{k,m}^{\mathcal{G}} + \sum_{k \in \mathcal{K}(\varepsilon)} \sum_{m \notin \mathcal{M}(k,\varepsilon)} w_{k,m}^{\mathcal{G}}$$
$$u_{big} := \sum_{k \in \mathcal{K}(\varepsilon)} \sum_{m \in \mathcal{M}(k,\varepsilon)} w_{k,m}^{\mathcal{G}}.$$

and  $u_{e,2} := u - u_{big} - u_{e,1}$ . Note that

$$\|u_{e,1}\|_{L^q} \le \frac{3\varepsilon}{4} h^{-\delta(q)} \frac{\sqrt{t_0}}{\sqrt{T}} \|u\|_{P,T}, \qquad \|u_{e,2}\|_{L^q} \le C h^{-\delta(q) + \frac{1}{2} - \delta_2} h^{-1} \|Pu\|_{H_h^{\frac{n-3}{2}}},$$

where we use (3.69), (3.71), (3.72), (3.75), and (3.76) to obtain the frist estimate, and (3.11) to obtain the second. These two estimates prove the claim on  $||u_{\varepsilon}||_{L^q}$  after combining them with (3.63). Next, observe that

$$u_{big} = \sum_{j \in \mathcal{L}(\varepsilon)} u_j, \qquad u_j := Op_h(\tilde{\chi}_{\tau_j}) Op_h(\psi) u, \qquad \mathcal{L}(\varepsilon) := \bigcup_{k \in \mathcal{K}(\varepsilon)} \bigcup_{m \in \mathcal{M}(k,\varepsilon)} \mathcal{G}_{k,m}.$$

We claim that the statement of the theorem holds with  $v_j = \sqrt{T}u_j$ . Note that  $v_j$  are manifestly microsupported inside  $\mathcal{T}_j$ .

Let  $\alpha \in \mathcal{S}_{U}(h,\varepsilon,u)$ , then by definition,

$$\|u_{big}\|_{L^{\infty}(B(x_{\alpha},R(h)))} \ge \frac{\varepsilon}{4} h^{\frac{1-n}{2}} \frac{\sqrt{t_0}}{\sqrt{T}} \|u\|_{P,T}.$$
(3.77)

Note that for all  $j \in \mathcal{L}(\varepsilon)$ , the estimate

$$\|Op_h(\tilde{\chi}_{\tau_j})Op_h(\psi)u\| + h^{-1}\|Op_h(\tilde{\chi}_{\tau_j})Op_h(\psi)Pu\|_{L^2} \le 2^{-k_1^{\varepsilon}+1}\|u\|_{P,T}$$
(3.78)

follows from the definition, (3.14), of  $\mathcal{A}_k$  and the fact that  $\chi_{\tau_j} \equiv 1$  on  $\operatorname{supp} \tilde{\chi}_{\tau_j}$ . To see that  $u_j$  is a quasimode, we use the definition of  $\mathcal{A}_k$  again, together with Proposition 2.3, and obtain

$$\|Pu_j\|_{L^2} \le \|[-h^2\Delta_g, Op_h(\tilde{\chi}_{\tau_j})]u_j\|_{L^2} + \|Op_h(\tilde{\chi}_{\tau_j})Pu\|_{L^2} \le C2^{-k_1^{\varepsilon}}h\|u\|_{P,T}.$$
(3.79)

The definition of  $k_1^{\varepsilon}$ , together with (3.78) and (3.79) give the required bounds on  $v_j$  and  $Pv_j$ .

Next, define

$$\mathcal{L}(\varepsilon, u, \alpha) := \{ j \in \mathcal{L} : \pi_M(\mathcal{T}_j) \cap B(x_\alpha, 3R(h)) \neq \emptyset \},\$$

and note that by [CG20, Lemma 3.7]

$$\begin{aligned} \|u_{big}\|_{L^{\infty}(B(x_{\alpha},R(h)))} &\leq Ch^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}\sum_{j\in\mathcal{L}(\varepsilon,u,\alpha)}\|Op_{h}(\tilde{\chi}_{\tau_{j}})Op_{h}(\psi)u\| + h^{-1}\|Op_{h}(\tilde{\chi}_{\tau_{j}})Op_{h}(\psi)Pu\|_{L^{2}} + O(h^{\infty})\|u\|_{L^{2}} \\ &\leq Ch^{\frac{1-n}{2}}R(h)^{\frac{n-1}{2}}2^{-k_{1}^{\varepsilon}}|\mathcal{L}(\varepsilon,u,\alpha)|\|u\|_{P,T} + O(h^{\infty})\|u\|_{P,T}. \end{aligned}$$
(3.80)

Therefore, combining (3.77) with (3.80) yields

$$\varepsilon \frac{\sqrt{t_0}}{\sqrt{T}} \le CR(h)^{\frac{n-1}{2}} 2^{-k_1^{\varepsilon}} |\mathcal{L}(\varepsilon, \alpha, u)| + O(h^{\infty}).$$

Moreover,  $\bigcup_{j \in \mathcal{L}(\varepsilon, u)} \mathcal{T}_j$  is  $[t_0, T]$  non-self looping and so by Lemma 3.4  $|\mathcal{L}(\varepsilon, u)| \leq \frac{C_n t_0}{T} 2^{2k_2^{\varepsilon}}$ . Using the definition (3.70) of  $k_1^{\varepsilon}, k_2^{\varepsilon}$  we have for h small enough,

$$c\varepsilon^2 R(h)^{1-n} = \varepsilon \frac{\sqrt{t_0}}{\sqrt{T}} R(h)^{\frac{1-n}{2}} 2^{k_1^{\varepsilon}} \le |\mathcal{L}(\varepsilon, u, \alpha)| \le |\mathcal{L}(\varepsilon, u)| \le \frac{C_n t_0}{T} 2^{2k_2^{\varepsilon}} \le C\varepsilon^{-2} R(h)^{1-n},$$

which yields the upper bound on  $|\mathcal{L}(\varepsilon, u)|$  and the lower bound on  $|\mathcal{L}(\varepsilon, u, \alpha)|$ . Note that, the upper bound on  $|\mathcal{L}(\varepsilon, u, \alpha)|$  follows from the fact that the total number of tubes over  $B(x_{\alpha}, 3R(h))$  is bounded by  $CR(h)^{1-n}$ . Next, we note that the fact that at most  $\mathfrak{D}_n$  tubes  $\mathcal{T}_j$  overlap implies

$$\sum_{j \in \mathcal{L}(\varepsilon, u, \alpha)} \| Op_h(\tilde{\chi}_{\tau_j}) Op_h(\psi) Pu \|_{L^2}^2 \le C \| Pu \|_{L^2}^2 + O(h^{\infty} \| u \|_{L^2})$$

Therefore, using the first inequality in (3.80) again, applying Cauchy-Schwarz, and using that there is C > 0 such that  $|\mathcal{L}(\varepsilon, u, \alpha)| \leq CR(h)^{1-n}$  we have

$$\begin{split} \frac{\varepsilon}{4} \frac{\sqrt{t_0}}{\sqrt{T}} \|u\|_{P,T} &\leq CR(h)^{\frac{n-1}{2}} |\mathcal{L}(\varepsilon, u, \alpha)|^{\frac{1}{2}} \Big(\sum_{j \in \mathcal{L}(\varepsilon, u, \alpha)} \|u_j\|_{L^2}^2 \Big)^{\frac{1}{2}} + Ch^{-1} \|Pu\|_{L^2} + O(h^{\infty}) \|u\|_{L^2} \\ &\leq C \Big(\sum_{j \in \mathcal{L}(\varepsilon, u, \alpha)} \|u_j\|_{L^2}^2 \Big)^{\frac{1}{2}} + o(T^{-1} \|u\|_{L^2}). \end{split}$$

In particular, for h small enough,  $c_{\sqrt{T}}^{\sqrt{t_0}} \|u\|_{P,T} \leq \left(\sum_{j \in \mathcal{L}(\varepsilon, u, \alpha)} \|u_j\|^2\right)^{\frac{1}{2}}$ . This completes the proof.

## 4. Proof of theorem 1

In order to finish the proof of Theorem 1, we need to verify that the hypotheses of Theorem 3 hold with  $T(h) = b \log h^{-1}$  for some b > 0, and such that for all  $x_1, x_2 \in U$  there is some splitting  $\mathcal{J}_{x_1} = \mathcal{G}_{x_1,x_2} \cup \mathcal{B}_{x_1,x_2}$  of the set of tubes over  $x_1 \in M$  with a set of 'bad' tubes  $\mathcal{B}_{x_1,x_2}$  satisfying

$$(|\mathcal{B}_{x_1,x_2}|R(h)^{n-1})^{\frac{1}{6+\varepsilon_0}(1-\frac{p_c}{p})} \le T(h)^{-\frac{1}{2}}$$

and  $\varepsilon_0 > 0$ . Fix  $x_1, x_2 \in U$  and let  $F_1, F_2: T^*M \to \mathbb{R}^{n+1}$  be smooth functions so that for i = 1, 2,

$$S_{x_{i}}^{*}M = F_{i}^{-1}(0), \qquad \frac{1}{2}d(q, S_{x_{i}}^{*}M) \le |F_{i}(q)| \le 2d(q, S_{x_{i}}^{*}M), \qquad \max_{|\alpha| \le 2}(|\partial^{\alpha}F_{i}(q)|) \le 2$$

$$dF_{i}(q) \text{ has a right inverse } R_{F_{i}}(q) \text{ with } ||R_{F_{i}}(q)|| \le 2.$$
(4.1)

Define also  $\psi_i : \mathbb{R} \times T^*M \to \mathbb{R}^{n+1}$  by  $\psi_i(t, \rho) = F_i \circ \varphi_t(\rho)$ .

To find  $\mathcal{B}_{x_1,x_2}$ , we apply the arguments from [CG19a, Sections 2, 4]. In particular, fix a > 0 and let  $r_t := a^{-1}e^{-a|t|}$ . Suppose that  $d(x_2, \mathcal{C}_{x_1}^{n-1,r_t,t_0,t_0}) > r_{t_0}$ , then for  $\rho_0 \in S_{x_1}^*M$  with

 $d(S^*_{x_2}M,\varphi_{t_0}(\rho_0)) < r_{t_0}$  we have by [CG19a, Lemma 4.1], that there exists  $\mathbf{w} \in T_{\rho_0}S^*_{x_1}M$  so that

$$d(\psi_2)_{(t_0,\rho_0)} : \mathbb{R}\partial_t \times \mathbb{R}\mathbf{w} \to T_{\psi_2(t_0,\rho_0)}\mathbb{R}^{n+1}$$

has a left inverse  $L_{(t_0,\rho_0)}$  with

$$||L_{(t_0,\rho_0)}|| \le C_M \max(ae^{C_M(a+\Lambda)|t_0|}, 1).$$

Next, let { $\Lambda_{\rho_j}^{\tau}(r_1)$  be a  $(\mathfrak{D}_M, \tau, r_1)$ -good cover for  $S^*M$ . We apply [CG19a, Proposition 2.2] to construct  $\mathcal{B}_{x_1,x_2}$  and  $\mathcal{G}_{x_1,x_2}$ .

**Remark 4.** We must point out that we are applying the proof of that proposition rather than the proposition as stated. The only difference here is that the loops we are interested in go from a point  $x_1$  to a point  $x_2$  where  $x_1$  and  $x_2$  are not necessarily equal. This does not affect the proof.

We use [CG19a, Proposition 2.2] to see that there exist  $\alpha_1 = \alpha_1(M) > 0$ ,  $\alpha_2 = \alpha_2(M, a)$ , and  $\mathbf{C_0} = \mathbf{C_0}(M, a)$  so that the following holds. Let  $r_0, r_1, r_2 > 0$  satisfy

$$r_0 < r_1, \qquad r_1 < \alpha_1 r_2, \qquad r_2 \le \min\{R_0, 1, \alpha_2 e^{-\gamma T}\}, \qquad r_0 < \frac{1}{3} e^{-\Lambda T} r_2,$$
(4.2)

where  $\gamma = 5\Lambda + 2a$  and  $\Lambda > \Lambda_{\max}$  where  $\Lambda_{\max}$  is as in (1.2). Then, for all balls  $B \subset S_{x_1}^*M$  of radius  $R_0 > 0$ , there is a family of points  $\{\rho_j\}_{j \in \mathcal{B}_B} \subset S_{x_1}^*M$  such that

$$|\mathcal{B}_B| \leq \mathbf{C}_0 \mathfrak{D}_n \ r_2 \ rac{R_0^{n-1}}{r_1^{n-1}} \ T \ e^{4(2\Lambda+a)T},$$

and for  $j \in \mathcal{G}_{\scriptscriptstyle B} := \{j \in \mathcal{J}_{x_1} : B(\rho_j, 2r_1) \cap B \neq \emptyset\} \setminus \mathcal{B}_{\scriptscriptstyle B}\}$ 

$$\bigcup_{t \in [t_0,T]} \varphi_t \Big( \Lambda_{\rho_j}^{\tau}(r_1) \Big) \cap \Lambda_{S_{x_2}}^{\tau} M(r_1) = \emptyset.$$

We proceed to apply [CG19a, Proposition 2.2]. There is  $c_M r^{1-n} \ge N_r > 0$  so that for all  $x_1 \in M$ ,  $S_{x_1}^*M$  can be covered by  $N_r$  balls. Let  $0 < R_0 < 1$  and  $\{B_i\}_{i=1}^{N_{R_0}}$  be such a cover. Fix  $0 < \varepsilon < \varepsilon_1 < \frac{1}{4}$  and set

$$r_0 := h^{\varepsilon_1}, \qquad r_1 := h^{\varepsilon}, \qquad r_2 := \frac{2}{\alpha_1} h^{\varepsilon}.$$

Let  $T(h) = b \log h^{-1}$  with  $0 < b < \frac{1}{4\Lambda_{\max}} < \frac{1-2\varepsilon_1}{2\Lambda_{\max}}$  to be chosen later. Then, the assumptions in (4.2) hold provided

$$h^{\varepsilon} < \min\left\{\frac{\alpha_1\alpha_2}{2}e^{-\gamma T}, \frac{\alpha_1R_0}{2}\right\}, \qquad h^{\varepsilon_1-\varepsilon} < \frac{2}{3\alpha_1}e^{-\Lambda T}$$

In particular, if we set  $\alpha_3 := \frac{\alpha_1 \alpha_2}{2}$ ,  $\alpha_4 = \frac{2}{3\alpha_1}$ , the assumptions in (4.2) hold provided  $h < \left(\frac{\alpha_1 R(h)}{2}\right)^{\frac{1}{\varepsilon}}$  and

$$T(h) < \min\left\{\frac{\varepsilon}{\gamma}\log h^{-1} + \frac{\log\alpha_3}{\gamma}, \frac{\varepsilon_1 - \varepsilon}{\Lambda}\log h^{-1} + \frac{\log(\alpha_4)}{\Lambda}\right\}.$$
(4.3)

Fix b > 0 and  $h_0 > 0$  so that  $b < \frac{\min(\varepsilon, \varepsilon_1 - \varepsilon)}{12(2\Lambda + a)}$  and (4.3) is satisfied for all  $h < h_0$ . Note that this implies that  $b = b(M, a), h_0 = h_0(M, a)$ . Let  $\mathcal{B}_{x_1, x_2} := \bigcup_{i=1}^{N_{R_0}} \mathcal{B}_{B_i}$ . Then, for  $j \in \mathcal{G}_{x_1, x_2} :=$ 

 $\mathcal{J}_{x_1} \setminus \mathcal{B}_{x_1,x_2},$ 

$$\bigcup_{\in [t_0,T]} \varphi_t \left( \Lambda_{\rho_j}^{\tau}(r_1) \right) \cap \Lambda_{S_{x_2}^*M}^{\tau}(r_1) = \emptyset.$$

Moreover, shrinking  $h_0$  in a way depending only on  $(M, a, \varepsilon)$ , we have for  $0 < h < h_0$ ,

$$r_1^{n-1}|\mathcal{B}_{x_1,x_2}| \le C_M \mathbf{C}_0 \mathfrak{D}_n r_2 T e^{4(2\Lambda+a)T} \le h^{\frac{\varepsilon}{3}}.$$

Therefore, putting  $R(h) = r_1 = h^{\varepsilon}$  and  $T = T(h) = b \log h^{-1}$  in Theorem 3 proves Theorem 1.

## 5. Anisotropic Pseudodifferential calculus

In this section, we develop the second microlocal calculi necessary to understand 'effective sharing' of  $L^2$  mass between two nearby points. That is, to answer the question: how much  $L^2$ mass is necessary to produce high  $L^{\infty}$  growth at two nearby points? To that end, we develop a calculus associated to the co-isotropic

$$\Gamma_x := \bigcup_{|t| < \frac{1}{2} \operatorname{inj}(M)} \varphi_t(\Omega_x), \qquad \Omega_x := \{\xi \in T_x^*M : |1 - |\xi|_g| < \delta\},$$

which allows for localization to a Lagrangian leaves  $\varphi_t(\Omega_x)$ . In Section 6.2 we will see, using a type of uncertainty principle, that the calculi associated to two distinct points,  $x_{\alpha}, x_{\beta} \in M$ , are incompatible in the sense that, despite the fact that  $\Gamma_{x_{\alpha}}$  and  $\Gamma_{x_{\beta}}$  intersect in a dimension 2 submanifold, for operators  $X_{x_{\alpha}}$  and  $X_{x_{\beta}}$  localizing to  $\Gamma_{x_{\alpha}}$  and  $\Gamma_{x_{\beta}}$  respectively,

$$\|X_{x_{\alpha}}X_{x_{\beta}}\|_{L^{2}\to L^{2}} \ll \|X_{x_{\alpha}}\|_{L^{2}\to L^{2}}\|X_{x_{\beta}}\|_{L^{2}\to L^{2}}.$$

Let  $\Gamma \subset T^*M$  be a co-isotropic submanifold and  $L = \{L_q\}_{q \in \Gamma}$  be a family of Lagrangian subspaces  $L_q \subset T_q\Gamma$  that is integrable in the sense that if U is a neighborhood of  $\Gamma$ , and V, Ware smooth vector fields on  $T^*M$  such that  $V_q, W_q \in L_q$  for all  $q \in \Gamma$ , then  $[V, W]_q \in L_q$  for all  $q \in \Gamma$ . The aim of this section is to introduce a calculus of pseudodifferential operators associated to  $(L, \Gamma)$  that allows for localization to  $h^{\rho}$  neighborhoods of  $\Gamma$  with  $0 \leq \rho < 1$  and is compatible with localization to  $h^{\rho}$  neighborhoods of the foliation of  $\Gamma$  generated by L. This calculus is close in spirit to those developed in [SZ99] and [DZ16]. To see the relationships between these calculi, note that the calculus in [DZ16] allows for localization to any leaf of a Lagrangian foliation defined over an open subset of  $T^*M$  and that in [SZ99] allows for localization along leaves of a Lagrangian foliation defined only over a co-isotropic submanifold of  $T^*M$ . In the case that the co-istropic is a whole open set, this calculus is the same as the one developed in [DZ16]. Similarly, in the case that the co-isotropic is a hypersurface and no Lagrangian foliation is prescribed, the calculus becomes that developed in [SZ99].

**Definition 3.** Let  $\Gamma$  be a co-isotropic submanifold and L a Lagrangian foliation on  $\Gamma$ . Fix  $0 \leq \rho < 1$  and let k be a positive integer. We say that  $a \in S^k_{\Gamma,L,\rho}$  if  $a \in C^{\infty}(T^*M)$ , a is supported in an h-independent compact set, and

$$V_1 \dots V_{\ell_1} W_1 \dots W_{\ell_2} a = O(h^{-\rho \ell_2} \langle h^{-\rho} d(\Gamma, \cdot) \rangle^{k-\ell_2})$$
(5.1)

where  $W_1, \ldots, W_{\ell_2}$  are any vector fields on  $T^*M, V_1, \ldots, V_{\ell_1}$  are vector fields on  $T^*M$  with  $(V_1)_q, \ldots, (V_{\ell_1})_q \in L_q$  for  $q \in \Gamma$ , and  $q \mapsto d(\Gamma, q)$  is the distance from q to  $\Gamma$  induced by the Sakai metric on  $T^*M$ .

We also define symbol classes associated to only the co-isotropic

**Definition 4.** Let  $\Gamma$  be a co-isotropic submanifold. We say that  $a \in S_{\Gamma,\rho}^k$  if  $a \in C^{\infty}(T^*M)$ , a is supported in an *h*-independent compact set, and

$$V_1 \dots V_{\ell_1} W_1 \dots W_{\ell_2} a = O(h^{-\rho \ell_2} \langle h^{-\rho} d(\Gamma, \cdot) \rangle^{k-\ell_2})$$

where  $V_1, \ldots V_{\ell_1}$  are tangent vector fields to  $\Gamma$ , and  $W_1, \ldots W_{\ell_2}$  are any vector fields.

5.1. Model case. The goal of this section is to define the quantization of symbols in  $S^k_{\Gamma_0,L_0,\rho}$ , where  $\Gamma_0, L_0$  are a model pair of co-isotropic and Lagrangian foliation defined below. The model co-isotropic submanifolds of dimension 2n - r is

$$\Gamma_0 := \{ (x', x'', \xi', \xi'') \in \mathbb{R}^r \times \mathbb{R}^{n-r} \times \mathbb{R}^r \times \mathbb{R}^{n-r} : x' = 0 \}$$

with Lagrangian foliation

$$L_0 := \{L_{0,q}\}_{q \in \Gamma_0}, \qquad L_{0,q} = \operatorname{span}\{\partial_{\xi_i}, i = 1, \dots n\} \subset T_q \Gamma_0.$$

Note that in this model case the distance from a point  $(x, \xi)$  to  $\Gamma_0$  is controlled by |x'|. Therefore,  $a \in S^k_{\Gamma_0, L_0, \rho}$  if and only if a is supported in an h-independent compact set and for all  $(\alpha, \beta) \in \mathbb{N}^n \times \mathbb{N}^n$  there exists  $C_{\alpha,\beta} > 0$  such that

$$|\partial_x^{\alpha}\partial_{\xi}^{\beta}a| \le C_{\alpha,\beta}h^{-\rho|\alpha|}\langle h^{-\rho}|x'|\rangle^{k-|\alpha|}$$

In the model case, it will be convenient to define  $\tilde{a} \in C^{\infty}(\mathbb{R}^n_x \times \mathbb{R}^n_{\xi} \times \mathbb{R}^r_{\lambda})$  such that

$$a(x,\xi) = \tilde{a}(x,\xi,h^{-\rho}x')$$

and for all  $(\alpha', \alpha'', \beta, \gamma) \in \mathbb{N}^r \times \mathbb{N}^{n-r} \times \mathbb{N}^n \times \mathbb{N}^r$  there exists  $C_{\alpha, \beta, \gamma} > 0$  such that

$$\left|\partial_{x'}^{\alpha'}\partial_{x''}^{\alpha''}\partial_{\xi}^{\beta}\partial_{\lambda}^{\gamma}\tilde{a}(x,\xi,\lambda)\right| \le C_{\alpha,\beta,\gamma}h^{-\rho|\alpha''|}\langle\lambda\rangle^{k-|\gamma|-|\alpha''|}.$$
(5.2)

Similarly, if  $a \in S^k_{\Gamma_0,\rho}$ , then for  $(\alpha', \alpha'', \beta, \gamma) \in \mathbb{N}^r \times \mathbb{N}^{n-r} \times \mathbb{N}^n \times \mathbb{N}^r$  there exists  $C_{\alpha,\beta,\gamma} > 0$ 

$$|\partial_{x'}^{\alpha'}\partial_{x''}^{\alpha''}\partial_{\xi}^{\beta}\partial_{\lambda}^{\gamma}\tilde{a}(x,\xi,\lambda)| \le C_{\alpha,\beta,\gamma}\langle\lambda\rangle^{k-|\gamma|}.$$
(5.3)

**Definition 5.** The symbols associated with this submanifold are as follows. We say  $a \in S^k_{\Gamma_0,L_0,\rho}$  if  $a \in C^{\infty}(\mathbb{R}^n_x \times \mathbb{R}^n_{\xi} \times \mathbb{R}^n_{\lambda})$  satisfies (5.2) and a is supported in an h-independent set in  $(x,\xi)$ . If we have the improved estimates (5.3) then we say that  $a \in \widetilde{S^k_{\Gamma_0,\rho}}$ .

**Remark 5.** While there is no  $\rho$  in the definition of  $\widetilde{S_{\Gamma_0,\rho}^k}$ , we keep it in the notation for consistency.

Let  $a \in \widetilde{S^k_{\Gamma_0,L_0,\rho}}$ . We then define

$$[\widetilde{Op}_{h}(a)]u(x) := \frac{1}{(2\pi h)^{n}} \int e^{\frac{i}{h}\langle x-y,\xi\rangle} a(x,\xi,h^{-\rho}x')u(y)dyd\xi$$

Since  $a \in \widetilde{S_{\Gamma_0,L_0,\rho}^k}$  is compactly supported in x, there exists C > 0 such that on the support of the integrand  $\lambda \leq Ch^{-\rho}$  and hence  $h \leq Ch^{1-\rho} \langle \lambda \rangle^{-1}$ . This will be important when computing certain asymptotic expansions.

**Lemma 5.1.** Let  $k \in \mathbb{R}$  and  $a \in \widetilde{S_{\Gamma_0, L_0, \rho}^k}$ . Then,

$$\|\widetilde{Op}_{h}(a)\|_{L^{2}\to L^{2}} \leq C \sup_{\mathbb{R}^{2n}} |a(x,\xi,h^{-\rho}x')| + O(h^{-\rho\max(k,0) + \frac{1-\rho}{2}})$$

Proof. Define  $T_{\delta}: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$  by

$$T_{\delta}u(x) := h^{\frac{n}{2}\delta}u(h^{\delta}x).$$
(5.4)

Then  $T_{\delta}$  is unitary and, for  $a \in \widetilde{S^k_{\Gamma_0, L_0, \rho}}$ ,

$$\widetilde{Op}_h(a)u = T_{\frac{1+\rho}{2}}^{-1}Op_1(a_h)T_{\frac{1+\rho}{2}}u, \qquad a_h(x,\xi) := a(h^{\frac{1+\rho}{2}}x, h^{\frac{1-\rho}{2}}\xi, h^{\frac{1-\rho}{2}}x').$$

Then, for all  $\alpha, \beta \in \mathbb{N}^n$  there exists  $C_{\alpha,\beta}$  such that  $|\partial_x^{\alpha}\partial_{\xi}^{\beta}a_h| \leq C_{\alpha,\beta}h^{\frac{1-\rho}{2}(|\alpha|+|\beta|)}\langle h^{\frac{1-\rho}{2}}x'\rangle^{k-|\alpha|}$ . Now, since  $a_h, b_h \in S_{\frac{1-\rho}{2}}$ , by [Zwo12, Theorem 4.23] there is a universal constant M > 0 with

$$\|\widetilde{Op}_1(a_h)\|_{L^2 \to L^2} \le C \sum_{|\alpha| \le Mn} \sup_{\mathbb{R}^{2n}} |\partial^{\alpha} a_h| \le C \sup |a| + C_a h^{-\max(\rho k, 0) + \frac{1-\rho}{2}}.$$

**Lemma 5.2.** Suppose that  $a \in \widetilde{S_{\Gamma_0,L_0,\rho}^{k_1}}$ ,  $b \in \widetilde{S_{\Gamma_0,L_0,\rho}^{k_2}}$ . Then,  $\widetilde{Op}_h(a)\widetilde{Op}_h(b) = \widetilde{Op}_h(c) + O(h^{\infty})_{L^2 \to L^2}$ where  $c \in \widetilde{S_{\Gamma_0,L_0,\rho}^{k_1+k_2}}$  satisfies

$$c = ab + O(h^{1-\rho})_{S_{\Gamma_0,L_0,\rho}^{\widetilde{k_1+k_2-1}}}.$$
(5.5)

In particular,

$$c \sim \sum_{j} \sum_{|\alpha|=j} \frac{i^{j}}{j!} ((hD_{x''})^{\alpha''} (hD_{x'} + h^{1-\rho}D_{\lambda})^{\alpha'} b) D_{\xi}^{\alpha} a.$$
(5.6)

If instead,  $a \in \widetilde{S_{\Gamma_0,\rho}^{k_1}}$  and  $b \in \widetilde{S_{\Gamma_0,\rho}^{k_2}}$ , then the remainder in (5.5) lies in  $h^{1-\rho} \widetilde{S_{\Gamma_0,\rho}^{k_1+k_2-1}}$ .

*Proof.* With  $T_{\delta}$  as in (5.4), we have  $\widetilde{Op}_{h}(a)\widetilde{Op}_{h}(b) = T_{\rho/2}^{-1}Op_{h}(a_{h})Op_{h}(b_{h})T_{\rho/2}$  where  $a_{h} = a(h^{\frac{\rho}{2}}x, h^{-\frac{\rho}{2}}\xi, h^{-\frac{\rho}{2}}x'), \qquad b_{h} = b(h^{\frac{\rho}{2}}x, h^{-\frac{\rho}{2}}\xi, h^{-\frac{\rho}{2}}x').$ 

Now, for all  $\alpha, \beta \in \mathbb{N}^n$  there exists  $C_{\alpha,\beta}$  such that

$$|\partial_x^{\alpha}\partial_\xi^{\beta}a_h| \le C_{\alpha,\beta}h^{-\frac{\rho}{2}(|\alpha|+|\beta|)} \langle h^{-\frac{\rho}{2}}x'\rangle^{k_1-|\alpha|}, \quad |\partial_x^{\alpha}\partial_\xi^{\beta}b_h| \le C_{\alpha,\beta}h^{-\frac{\rho}{2}(|\alpha|+|\beta|)} \langle h^{-\frac{\rho}{2}}x'\rangle^{k_2-|\alpha|}.$$

In particular, using that a and b are compactly supported,  $a_h \in h^{-\max(\rho k_1,0)} S_{\rho/2}$  and  $b_h \in h^{-\max(\rho k_2,0)} S_{\rho/2}$  and hence [Zwo12, Theorems 4.14,4.17] apply. In particular, if we let M > 0 and  $\tilde{k} := \max(k_1,0) + \max(k_2,0)$ , we obtain  $Op_h(a_h)Op_h(b_h) = Op_h(c_h)$  where, for any N > 0,

$$\begin{aligned} c_h(x,\xi) &= \sum_{j=0}^{N-1} \sum_{|\alpha|=j} \frac{h^j i^j}{j!} (D_{\xi}^{\alpha} a_h(x,\xi)) (D_x^{\alpha} b_h(x,\xi)) + O(h^{-\rho \tilde{k} + N(1-\rho)})_{S_{\rho/2}} \\ &= \sum_{j=0}^{N-1} \sum_{|\alpha|=j} \sum_{\alpha'+\alpha''=\alpha} \frac{h^{(1-\rho)j} i^j}{j!} (D_{\xi}^{\alpha} a)_h [(h^{\rho} D_{x''})^{\alpha''} (h^{\rho} D_{x'} + D_{\lambda})^{\alpha'} b]_h + O(h^{-\rho \tilde{k} + N(1-\rho)})_{S_{\rho/2}}. \end{aligned}$$

Choosing  $N = \max\left(k_1 + k_2, \frac{\rho \tilde{k} + M}{1 - \rho}\right)$ , the remainder is  $O(h^M)_{S_{\rho/2}}$ . Moreover, since a and b were compactly supported, we may assume introducing an  $h^{\infty}$  error, that the remainder is supported in  $\{(x,\xi) : |(x,\xi)| \leq Ch^{-\frac{\rho}{2}}\}$ . Putting

$$c = \sum_{j=0}^{N-1} \sum_{|\alpha|=j} \sum_{\alpha'+\alpha''=\alpha} \frac{i^j}{j!} (D_{\xi}^{\alpha} a) [(hD_{x''})^{\alpha''} (hD_{x'} + h^{1-\rho}D_{\lambda})^{\alpha'} b]_{\xi}$$

we thus have  $T_{\rho/2}^{-1}Op_h(c_h)T_{\rho/2} = \widetilde{Op}_h(c) + O(h^M)_{\mathcal{D}' \to C^{\infty}}$  as claimed.

**Lemma 5.3.** Suppose that  $a \in \widetilde{S_{\Gamma_0,L_0,\rho}^{m_1}}$ ,  $b \in \widetilde{S_{\Gamma_0,L_0,\rho}^{m_2}}$ . Then,  $[\widetilde{Op}_h(a), \widetilde{Op}_h(b)] = -ih^{1-\rho}\widetilde{Op}_h(c) + O(h^{\infty})_{L^2 \to L^2}$ 

where  $c \in \widetilde{S_{\Gamma_0,L_0,\rho}^{m_1+m_2-2}}$  satisfies

$$c = h^{\rho} \sum_{i=1}^{n} (\partial_{\xi_i} a \partial_{x_i} b - \partial_{\xi_i} b \partial_{x_i} a) + \sum_{i=1}^{r} (\partial_{\xi_i} a \partial_{\lambda_i} b - \partial_{\lambda_i} a \partial_{\xi_i} b) + O(h^{1-\rho})_{S_{\Gamma_0, L_0, \rho}^{\widetilde{m_1 + m_2 - 2}}}.$$

If instead,  $a \in \widetilde{S_{\Gamma_0,\rho}^{m_1}}$  and  $b \in \widetilde{S_{\Gamma_0,\rho}^{m_2}}$ , then the remainder lies in  $h^{1-\rho}S_{\Gamma_0,\rho}^{\widetilde{m_1+m_2}-2}$ . Moreover, if  $a \in S^{\text{comp}}(\mathbb{R}^{2n})$  is independent of  $\lambda$  and  $\partial_{\xi'}a = e(x,\xi)x'$  with  $e(x,\xi) : \mathbb{R}^r \to \mathbb{R}^r$  for all  $(x,\xi)$ , then

$$[\widetilde{Op}_h(a),\widetilde{Op}_h(b)] = -ih\widetilde{Op}_h(c) + O(h^{\infty})_{\Psi^{-\infty}}$$

with  $c = H_a b + \sum_{i=1}^r (e\lambda)_i \partial_{\lambda_i} b + O(h^{1-\rho})_{S_{\Gamma_0, L_0, \rho}^{m_2-1}}$ . Similarly, the same conclusion holds if  $b \in \widetilde{S_{\Gamma_0, \rho}^{m_2}}$  with the error term in c being  $O(h^{1-\rho})_{\widetilde{S_{\Gamma_0, \rho}^{m_2-1}}}$ .

*Proof.* In each case, we need only apply the formula (5.6).

5.2. Reduction to normal form. In order to define the quantization of symbols in  $S_{\Gamma,L,\rho}$  for general  $(\Gamma, L)$ , we first explain how to reduce the problem to the model case  $(\Gamma_0, L_0)$ .

**Lemma 5.4.** Let L be a Lagrangian foliation over a co-isotropic submanifold  $\Gamma \subset \mathbb{R}^{2n}$  of dimension 2n-r. Then for each  $(x_0,\xi_0) \in \Gamma$  there is a neighborhood  $U_0$  of  $(x_0,\xi_0)$  and a symplectomorphism  $\kappa : U_0 \to V_0 \subset T^* \mathbb{R}^n$  such that

$$\kappa(\Gamma \cap U_0) = \Gamma_0 \cap V_0$$
 and  $(\kappa_*)_q L_q = L_{0,q}$  for  $q \in \Gamma \cap U_0$ .

*Proof.* We first put  $\Gamma$  in normal form. That is, we build symplectic coordinates  $(y, \eta)$  such that

$$\Gamma = \{(y,\eta) : y_1 = \dots = y_r = 0\}.$$
(5.7)

First, assume r = 1 and let  $f_1 \in C^{\infty}(T^*M)$  define  $\Gamma$ . By Darboux's theorem there are symplectic coordinates such that  $y_1 = f_1$  and the proof of (5.7) is complete for r = 1.

Next, assume that we can put any coisotropic of co-dimension r-1 in normal form. Let  $f_1, \ldots, f_r \in C^{\infty}(T^*M)$  define  $\Gamma$ . Then, since  $\Gamma$  is co-isotropic, for  $X \in T\Gamma$  and  $i = 1, \ldots, r$ 

$$\sigma(X, H_{f_i}) = df_i(X) = 0.$$

32

In addition, since  $\Gamma$  is co-isotropic,  $(T\Gamma)^{\perp} \subset T\Gamma$  and so  $H_{f_i} \in T\Gamma$  for all  $i = 1, \ldots, r$ . In particular,

$$\{f_i, f_j\} = H_{f_i}f_j = df_j(H_{f_i}) = 0$$

on  $\Gamma$ . Now, using Darboux' theorem, choose symplectic coordinates  $(y, \eta) = (y_1, y', \eta_1, \eta')$  so that  $y_1 = f_1$  and  $(x_0, \xi_0) \mapsto (0, 0)$ . Then,  $\partial_{\eta_1} f_j = \{f_j, y_1\} = 0$  on  $\Gamma$ , for  $j = 2, \ldots, r$ . Next, observe that  $\Gamma = \{(y, \eta) : y_1 = f_2 = \cdots = f_r = 0\}$ , and  $dy_1, \{df_j\}_{j=2}^r$  are independent. Thus, since  $\partial_{\eta_1} f_j = 0$  on  $\Gamma$ ,

$$\Gamma = \{(y,\eta) : y_1 = 0, \ f_j(0,y',0,\eta') = 0, \ j = 2,\dots,r\}.$$

Now,  $\{y_1 = \eta_1 = 0\} \cap \Gamma$  is a co-isotropic submanifold of co-dimension r-1 in  $T^*\{y_1 = 0\}$ . Hence, by induction, there are symplectic coordinates  $(y_2, \ldots, y_n, \eta_2, \ldots, \eta_n)$  on  $T^*\{y_1 = 0\}$  such that

$$\Gamma \cap \{y_1 = \eta_1 = 0\} = \{y_1 = \eta_1 = 0, \ y_2 = \dots = y_r = 0\}$$

In particular,

$$\{(y',\eta'): f_j(0,y',0,\eta')=0, \ j=2,\ldots,r\}=\{y_2=\cdots=y_r=0\}$$

Thus, extending  $(y_2, \ldots, y_n, \eta_2, \ldots, \eta_n)$  to be independent of  $(y_1, \eta_1)$  puts  $\Gamma$  in the form (5.7).

Next, we adjust the coordinates to be adapted to L along  $\Gamma$ . First, define  $\tilde{y}_i := y_i$  for  $i = 1, \ldots, r$ . Then, since  $L \subset T\Gamma$ , for every  $i = 1, \ldots, r$  we have that  $d\tilde{y}_i(X)|_{\Gamma}$  is well defined for  $X \in L$  and  $d\tilde{y}_i(X)|_{\Gamma} = 0$ . Next, since L is integrable, the Frobenius theorem [Lee13, Theorem 19.21] shows that there are coordinates  $(\tilde{y}_{r+1}, \ldots, \tilde{y}_n, \tilde{\xi}_1, \ldots, \tilde{\xi}_n)$  on  $\Gamma$ , defined in a neighborhood of (0, 0), such that L is the annihilator of  $d\tilde{y}$ . Since we know that for every  $X \in L$ 

$$\sigma(X, H_{\tilde{y}_i}) = d\tilde{y}_i(X) = 0,$$

and L is Lagrangian, we conclude that  $H_{\tilde{y}_i} \in L$ . In particular, since L is the annihilator of  $d\tilde{y}$ ,

$$\{\tilde{y}_i, \tilde{y}_j\} = H_{\tilde{y}_i}\tilde{y}_j = d\tilde{y}_j(H_{\tilde{y}_i}) = 0$$

Now, extend  $(\tilde{y}_{r+1}, \ldots, \tilde{y}_n, \tilde{\xi}_1, \ldots, \tilde{\xi}_n)$  outside  $\Gamma$  to be independent of  $(\tilde{y}_1, \ldots, \tilde{y}_r)$ . Then,  $\{\tilde{y}_i, \tilde{y}_j\} = 0$  in a neighborhood of  $(x_0, \xi_0)$  and hence, by Darboux's theorem, there are functions  $\{\tilde{\eta}_j\}_{j=1}^n$ , such that  $\{\tilde{y}_i, \tilde{\eta}_j\} = \delta_{ij}$  and  $\{\tilde{\eta}_i, \tilde{\eta}_j\} = 0$ . In particular, in the  $(\tilde{y}, \tilde{\eta})$  coordinates,  $\Gamma = \{(\tilde{y}, \tilde{\eta}) : \tilde{y}_1 = \cdots = \tilde{y}_r = 0\}$ , and  $d\tilde{y}(L)|_{\Gamma} = 0$ . In particular,  $L = \operatorname{span}\{\partial\tilde{\eta}_i\}$  as claimed.  $\Box$ 

In order to create a well-defined global calculus of psuedodifferential operators associated to  $(\Gamma, L)$ , we will need to show invariance under conjugation by FIOs preserving the pair  $(L_0, \Gamma_0)$ .

**Proposition 5.5.** Suppose that  $U_0, V_0$  are neighborhoods of (0, 0) in  $T^*\mathbb{R}^n$  and  $\kappa : U_0 \to V_0$  is a symplectomorphism such that

$$\kappa(0,0) = (0,0), \qquad \kappa(\Gamma_0 \cap U_0) = \Gamma_0 \cap V_0, \qquad \kappa_*|_{\Gamma_0} L_0 = L_0|_{\Gamma_0}. \tag{5.8}$$

Next, let T be a semiclassically elliptic FIO microlocally defined in a neighborhood of

$$((0,0),(0,0)) \in T^* \mathbb{R}^n \times T^* \mathbb{R}^n$$

quantizing  $\kappa$ . Then, for  $a \in \widetilde{S_{\Gamma_0,L_0,\rho}^k}$ , there are  $b \in \widetilde{S_{\Gamma_0,L_0,\rho}^k}$  and  $c \in \widetilde{S_{\Gamma_0,L_0,\rho}^{k-1}}$  such that

$$T^{-1}Op_h(a)T = Op_h(b), \qquad b = a \circ K_\kappa + h^{1-\rho_h}$$

where  $K_{\kappa}: T^*\mathbb{R}^n \times \mathbb{R}^r \to T^*\mathbb{R}^n \times \mathbb{R}^r$  is defined by

$$K_{\kappa}(y,\eta,\mu) = \left(\kappa(y,\eta), \pi_{x'}(\kappa(y,\eta))\frac{|\mu|}{|y'|}\right),$$

and  $\pi_{x'}: T^*\mathbb{R}^n \to \mathbb{R}^r$  is the projection onto the first r-spatial coordinates. In addition, if  $a \in \widetilde{S^k_{\Gamma_0,\rho}}$ , then  $c \in \widetilde{S^{k-1}_{\Gamma_0,\rho}}$  and  $b \in \widetilde{S^k_{\Gamma_0,\rho}}$ .

To prove Proposition 5.5, we follow [SZ99]. First, observe that the proposition holds with  $\kappa = \text{Id}$  since then T is a standard pseudodifferential operator. In addition, the proposition also holds whenever for a given  $j \in \{1, ..., n\}$  we work with

$$\kappa(y,\eta) := (y_1, \dots, y_{j-1}, -y_j, y_{j+1}, \dots, y_n, \eta_1, \dots, \eta_{j-1}, -\eta_j, \eta_{j+1}, \dots, \eta_n).$$

Indeed, this follows from the fact that in this case an FIO quantizing  $\kappa$  is

 $Tu(x) = u(x_1, \dots, x_{j-1}, -x_j, x_{j+1}, \dots, x_n)$ 

and so the conclusion of the proposition follows from a direct computation together with the identity case. Thus, we may assume that

$$\kappa(y,\eta) = (x,\xi) \qquad \text{implies} \qquad x_i y_i \ge 0, \qquad i = 1, \dots n. \tag{5.9}$$

**Lemma 5.6.** Let  $\kappa$  be a symplectomorphism satisfying (5.8) and (5.9). Then, there is a piecewise smooth family of symplectomorphisms  $[0,1] \ni t \mapsto \kappa_t$  such that  $\kappa_t$  satisfies (5.8), (5.9),  $\kappa_0 = \text{Id}$ , and  $\kappa_1 = \kappa$ .

*Proof.* In what follows we assume that  $\kappa(y,\eta) = (x,\xi)$  but reorder the coordinates  $(y',y'',\eta',\eta'') \in T^*\mathbb{R}^n$  is represented as  $(y',\eta',y'',\eta'') \in \mathbb{R}^{2r} \times \mathbb{R}^{2(n-r)}$ . Let  $\xi'$  and  $\kappa'' = (x''(y',\eta),\xi''(y',\eta))$  with

$$\kappa|_{\Gamma_0}: (0,\eta',y'',\eta'') \mapsto (0,\xi'(y'',\eta),\kappa''(y'',\eta)).$$

Now, since  $(\kappa_*)|_{\Gamma_0}L_0 = L_0$ , we have for  $i = 1, \ldots n$ ,

$$\kappa_* \partial_{\eta_i} = \frac{\partial x_j}{\partial \eta_i} \partial_{x_j} + \frac{\partial \xi_j}{\partial \eta_i} \partial_{\xi_j} \in L_0$$
(5.10)

and hence,

$$\partial_{\eta} x|_{\Gamma_0} \equiv 0. \tag{5.11}$$

Next, since  $\kappa$  preserves  $\Gamma_0$ ,  $\{\kappa^* x_i\}_{i=1}^r$  defines  $\Gamma_0$  and  $\operatorname{span}\{d\kappa^* x_i|_{\Gamma_0}\}_{i=1}^r = \operatorname{span}\{dy_i|_{\Gamma_0}\}_{i=1}^r$ ,

$$\operatorname{span}\{H_{\kappa^* x_i}|_{\Gamma_0}\}_{i=1}^r = \operatorname{span}\{H_{y_i}|_{\Gamma_0}\}_{i=1}^r.$$

By Jacobi's theorem,  $\kappa_* H_{\kappa^* x^i} = H_{x_i}$ . Therefore,

$$(\kappa|_{\Gamma_0})_* \left( \operatorname{span}\{H_{y_i}\}_{i=1}^r \Big|_{\Gamma_0} \right) = \operatorname{span}\{H_{x_i}\}_{i=1}^r \Big|_{\Gamma_0},$$

and we conclude from (5.10)  $\xi''|_{\Gamma_0}$  is independent of  $\eta'$  and hence that  $\kappa''$  is independent of  $\eta'$ . In particular,  $\kappa''$  is a symplectomorphism on  $T^*\mathbb{R}^{n-r}$ . This also implies that for each fixed  $(y'', \eta'')$ , the map  $\eta' \mapsto \xi'(y'', \eta', \eta'')$  is a diffeomorphism. Writing

$$\kappa''(y'',\eta'') = (x''(y'',\eta''),\xi''(y'',\eta'')),$$

we have by (5.11) that  $\partial_{\eta''}x'' = 0$ , and hence x'' = x''(y''). Now, since  $\kappa''$  is symplectic,

$$(\partial_{\eta''}\xi''d\eta'' + \partial_{y''}\xi''dy'') \wedge \partial_{y''}x''dy'' = d\eta'' \wedge dy'',$$

and so we conclude that

$$(\partial_{y''}x'')^t \partial_{\eta''}\xi'' = \mathrm{Id}, \qquad (\partial_{y''}x'')^t \partial_{y''}\xi'' \quad \text{is diagonal.}$$
(5.12)

The first equality in (5.12) gives that  $\partial_{\eta''}\xi''$  is a function of y'' only, and hence there exists a function F = F(y'') such that  $\xi''(y'', \eta'') = [(\partial x''(y''))^t]^{-1}(\eta'' - F(y''))$ . Therefore, calculating on  $\eta'' = F(y'')$ , the second statement in (5.12) implies that  $-\partial_{y''}F(y'')dy'' \wedge dy'' = 0$ . In particular,  $d(F(y'') \cdot dy'') = 0$ . It follows from the Poincaré lemma that, shrinking the neighborhood of (0,0) to be simply connected if necessary,  $F(y'') \cdot dy'' = d\psi(y'')$  for some function  $\psi = \psi(y'')$ . Hence,

$$\kappa''(y'',\eta'') = \left(x''(y''), \ \left[(dx''(y''))^t\right]^{-1}(\eta'' - \partial\psi(y''))\right).$$
(5.13)

Now, every symplectomorphism of the form (5.13) preserves  $L_0$ . Hence, we can deform  $\kappa''$  to the identity by putting  $\psi_t = t\psi$  and deforming x'' to the identity. Since the assumption in (5.9) implies  $\partial_{y''}x'' > 0$ , this can be done simply by taking  $x''_t = (1-t) \operatorname{Id} + tx''$ . Putting  $\kappa''_t = (x''_t, \xi''_t)$ , there is  $\kappa''_t$  such that  $\kappa''_0 = \operatorname{Id}$  and  $\kappa''_1 = \kappa''$ . Now, composing  $\kappa$  with

$$(y',\eta';y'',\eta'')\mapsto (y',\eta';(\kappa_t'')^{-1}(y'',\eta''))$$

we reduce to the case that  $\kappa'' = \text{Id.}$  In particular, we need only consider the case in which

$$\kappa(y',\eta',y'',\eta'') = \left(f(y,\eta)y', \ \xi'(y'',\eta) + h_0(y,\eta)y', \ (y'',\eta'') + h_1(y,\eta)y'\right).$$
(5.14)

where  $f(y,\eta) \in \mathbb{GL}_r$ ,  $h_0(y,\eta)$  is an  $r \times r$  matrix, and  $h_1(y,\eta)$  is an  $2(n-r) \times r$  matrix. Next, we claim that the projection map from graph( $\kappa$ ) to  $\mathbb{R}^{2n}$  defined as  $(x,\xi;y,\eta) \mapsto (x,\eta)$  is a local diffeomorphism. To see this, note that for |y'| small the map  $(x'',\eta'') \mapsto (y'',\xi'')$  is a diffeomorphism, that for each fixed  $(y'',\eta'')$  the map  $\eta' \mapsto \xi'$  is a diffeomorphism, and that  $\det \partial_{y'}x'|_{\Gamma_0} \neq 0$ . Thus,  $\kappa$  has a generating function  $\phi$ :  $\kappa : (\partial_{\eta}\phi(x,\eta),\eta) \mapsto (x,\partial_x\phi(x,\eta))$ , such that

$$\det \partial_{xn}^2 \phi(0,0) \neq 0, \qquad \partial_{n'} \phi(0,x'',\eta) = 0.$$

Now, writing  $\kappa = (\kappa', \kappa'')$ , we have  $\kappa'' = \text{Id at } x' = 0$ . Therefore,

$$\partial_{\eta''}\phi(0,x'',\eta) = x'', \qquad \partial_{x''}\phi(0,x'',\eta) = \eta''$$

and we have  $\phi(0, x'', \eta) = \langle x'', \eta'' \rangle + C$  for some  $C \in \mathbb{R}$ . We may choose C = 0 to obtain

$$\phi(x,\eta) = \langle x'',\eta'' \rangle + g(x,\eta)x', \qquad (5.15)$$

for some  $g : \mathbb{R}^{2n} \to \mathbb{M}_{1 \times r}$ . Finally, since  $\kappa(0,0) = (0,0)$  and  $\partial_{x\eta}^2 \phi$  is non-degenerate, we have  $\partial_{x'} \phi(0,0) = g(0,0) = 0$  and  $\partial_{\eta'} g$  is non-degenerate. In fact (5.9) implies that as a quadratic form

$$\partial_{\eta'}g > 0. \tag{5.16}$$

Observe next that for every  $\phi$ , such that (5.15) holds for some g satisfying (5.16) and g(0,0) = 0 generates a canonical transformation satisfying (5.14) and (5.9). In particular, the symplectomorphism satisfies (5.8). Thus, we can deform from the identity by putting  $g_t = (1-t)\eta' + tg$ .  $\Box$ 

Finally, we proceed with the proof of Proposition 5.5.

Proof of Proposition 5.5. Let  $\kappa_t$  be as in Lemma 5.6. That is, a piecewise smooth deformation from  $\kappa_0 = \text{Id to } \kappa_1 = \kappa$  such that  $\kappa_t$  preserves  $\Gamma_0$  and  $(\kappa_t)_*|_{\Gamma_0}$  perserves  $L_0$ . Let  $T_t$  be piecewise smooth family of elliptic FIOs defined microlocally near (0,0), quantizing  $\kappa_t$ , and satisfying

$$hD_tT_t + T_tQ_t = 0, \qquad T_0 = \text{Id}.$$
 (5.17)

Here,  $Q_t$  is a smooth family of pseudodifferential operator with symbol  $q_t$  satisfying  $\partial_t \kappa_t = (\kappa_t)_* H_{q_t}$ . (Such an FIO exists, for example, by [Zwo12, Chapter 10] and  $q_t$  exists by [Zwo12, Theorems 11.3, 11.4]) Next, define

$$A_t := T_t^{-1} \widetilde{Op}_h(a) T_t.$$

Note that  $T^{-1}\widetilde{Op}(a)T = T^{-1}T_1T_1^{-1}\widetilde{Op}(a)T_1T_1^{-1}T + O(h^{\infty})_{\Psi^{-\infty}}$ . Hence, since the Proposition follows by direct calculation when  $\kappa = \text{Id}$ , we may assume that  $T = T_1$ .

In that case, our goal is to find a symbol b such that  $A_1 = Op_h(b)$ . First, observe that (5.17) implies that  $hD_tT_t^{-1} - Q_tT_t^{-1} = 0$  and so

$$hD_tA_t = [Q_t, A_t], \qquad A_0 = \widetilde{Op}_h(a).$$
(5.18)

We will construct  $b_t \in \widetilde{S_{\Gamma_0,L_0,\rho}^k}$  such that  $B_t := \widetilde{Op}_h(b_t)$  satisfies

$$hD_tB_t = [Q_t, B_t] + O(h^{\infty})_{\Psi^{-\infty}}, \qquad B_0 = \widetilde{Op}_h(a).$$
(5.19)

This would yield that  $B_t - A_t = O(h^{\infty})_{L^2 \to L^2}$  and the argument would then be finished by setting  $b = b_1$ . Indeed, that  $B_t - A_t = O(h^{\infty})_{L^2 \to L^2}$  would follow from the fact that by (5.19)

$$hD_t(T_tB_tT_t^{-1}) = O(h^\infty)_{\Psi^{-\infty}},$$

and hence, since  $T_0 = \text{Id}$  and  $B_0 = \widetilde{Op}_h(a)$ , we have  $T_t B_t T_t^{-1} - \widetilde{Op}_h(a) = O(h^{\infty})_{\Psi^{-\infty}}$ . Combining this with the fact that both  $T_t$  and  $T_t^{-1}$  are bounded on  $H_h^k$  completes the proof.

To find  $b_t$  as in (5.19), note that since  $\kappa_t$  preserves  $\Gamma_0$  and  $L_0$ ,  $\partial_t \kappa_t = H_{q_t}$ , and  $H_{q_t}$  is tangent to  $L_0$  on  $\Gamma_0$ . Therefore,  $\partial_{\eta'} q_t = 0$  on y' = 0 and so there exists  $r_t(y, \eta)$  such that  $\partial_{\eta'} q_t(y, \eta) = r_t(y, \eta)y'$ . Hence, by Lemma 5.3 for any  $b \in S_{\Gamma_0, L_0, \rho}^k$ 

$$[Q_t, \widetilde{Op}_h(b)] = -ih\widetilde{Op}_h(f) + O(h^{\infty})_{\Psi^{-\infty}}, \quad f = H_{qt}b + \sum_{j=1}^r (r_t\lambda)_j (\partial_\lambda b)_j + O(h^{1-\rho})_{S_{\Gamma_0, L_0, \rho}^{k-2}}.$$

Then, letting  $b_t^0 := a \circ K_{\kappa_t} \in \widetilde{S^k_{\Gamma_0,L_0,\rho}}$  and  $B_t^0 = \widetilde{Op}_h(b_t^0)$  yields

$$hD_tB_t^0 = -ih\widetilde{Op}_h\left(H_{q_t}b_t^0 + (r_t\mu)\cdot\partial_\mu b_t^0\right) = [Q_t, B_t^0] + h^{2-\rho}\widetilde{Op}_h(e_t^0)$$

where  $e_t^0 \in \widetilde{S_{\Gamma_0,L_0,\rho}^{k-2}}$ . This follows from the fact that if we set  $\mu(y) = y'h^{-\rho}$ , then  $\partial_t(b_t^0(y,\eta,\mu(y))) = H_{q_t}b_t^0(y,\eta,\mu(y)) + \partial_\mu b_t^0(y,\eta,\mu(y))H_{q_t}(\mu(y))$ and  $H_{q_t}\mu(y) = r_t(y,\eta)\mu(y)$ .

Iterating this procedure and solving away successive errors finishes the proof of Proposition 5.5. If  $a \in \widetilde{S_{\Gamma_0,\rho}^k}$ , then we need only use that  $\partial_{\xi'}q_t = r_t x'$  and we obtain the remaining results. Our next lemma follows [SZ99, Lemma 4.1] and gives a characterization of our second microlocal calculus in terms of the action of an operator. In what follows, given operators A and B, we define the operator  $\operatorname{ad}_A$  by  $\operatorname{ad}_A B = [A, B]$ .

**Lemma 5.7** (Beal's criteria). Let  $A_h : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$  and  $k \in \mathbb{Z}$ . Then,  $A_h = \widetilde{Op}_h(a)$  for some  $a \in \widetilde{S^k_{\Gamma_0, L_0, \rho}}$  if and only if for any  $\alpha, \beta \in \mathbb{R}^n$  there exists C > 0 with

$$\|\operatorname{ad}_{h^{-\rho_{x}}}^{\alpha}\operatorname{ad}_{hD_{x}}^{\beta}A_{h}u\|_{|\beta|-\min(k,0)} \leq Ch^{(1-\rho)(|\alpha|+|\beta|)}\|u\|_{\max(k,0)}$$

where  $||u||_r := ||u||_{L^2} + ||h^{-\rho r}|x'|^r u||_{L^2}$ , for  $r \ge 0$ . Similarly,  $A_h = Op_h(a)$  for some  $a \in \widetilde{S_{\Gamma_0,\rho}^k}$  if and only if

$$\|\mathrm{ad}_{h-\rho_{x'}}^{\alpha'} \mathrm{ad}_{x''}^{\alpha''} \mathrm{ad}_{hD_{x'}}^{\beta'} \mathrm{ad}_{hD_{x''}}^{\beta''} A_h u\|_{|\beta'|-\min(k,0)} \le Ch^{(1-\rho)(|\alpha'|+|\beta'|)+|\alpha''|+|\beta''|} \|u\|_{\max(k,0)} \le Ch^{(1-\rho)(|\alpha'|+|\beta'|)+|\alpha''|+|\beta''|} \le Ch^{(1-\rho)(|\alpha'|+|\beta'|)+|\alpha''|+|\beta''|+|\beta''|} \le Ch^{(1-\rho)(|\alpha'|+|\beta'|)} \le Ch^{(1-\rho)(|\alpha'|+|\beta'|)}$$

then  $A_h = Op_h(a)$  for some  $a \in S^k_{\Gamma_0,\rho}$ .

*Proof.* The fact that  $A_h = \widetilde{Op}_h(a)$  for some  $a \in \widetilde{S_{\Gamma_0,L_0,\rho}^k}$  implies the estimates above follows directly from the model calculus. Let  $U_h$  be the unitary (on  $L^2$ ) operator,  $U_h u(x) = h^{\frac{n}{2}} u(hx)$ , and note that

$$||U_h^{-1}u||_r = ||u||_{L^2} + ||h^{(1-\rho)r}|x'|^r u||_{L^2}.$$

Then, consider  $\tilde{A}_h := U_h A_h U_h^{-1}$ . For fixed h, we can use Beal's criteria (see e.g. [Zwo12, Theorem 8.3]) to see that there is  $a_h$  such that  $\tilde{A}_h = a_h(x, D)$ . Define a such that  $a(hx, \xi; h) = a_h(x, \xi)$  and hence,  $A_h = Op_h(a)$ . Note that for  $\phi, \psi \in \mathcal{S}(\mathbb{R}^n)$ ,

$$\langle \tilde{A}_h \psi, \phi \rangle = \frac{1}{(2\pi)^n} \iint e^{i\langle x, \xi \rangle} a_h(x, \xi) \hat{\psi}(\xi) \overline{\phi(x)} dx d\xi,$$
(5.20)

where  $\hat{\psi}(\xi) = (\mathcal{F}\psi)(\xi) = \int e^{-i\langle y,\xi\rangle} \psi(y) dy$ . Next, define

$$B_h := U_h \operatorname{ad}_{h^{-\rho}x}^{\alpha} (\operatorname{ad}_{hD_x}^{\beta}(A_h)) U_h^{-1}.$$

Since  $D_x U_h = U_h h D_x$  and  $U_h^{-1} D_x = h D_x U_h^{-1}$ , we have

$$B_h = \operatorname{ad}_{h^{1-\rho_x}}^{\alpha} \operatorname{ad}_{D_x}^{\beta} \tilde{A}_h = (-i)^{|\alpha|+|\beta|} h^{(1-\rho)|\alpha|} b_h(x, D),$$

where  $b_h(x,\xi) = (-\partial_\xi)^{\alpha} \partial_x^{\beta} a_h(x,\xi)$ . Our goal is then to understand the behavior of  $b_h(x,\xi)$  in terms of h and  $\langle h^{1-\rho}x' \rangle$ . Let  $\tau_{x_0}$  and  $\hat{\tau}_{\xi_0}$  be the physical and frequency shift operators

$$\tau_{x_0}u(x) = u(x - x_0), \qquad \hat{\tau}_{\xi_0}u(x) = e^{i\langle x,\xi_0 \rangle}u(x),$$

with  $\mathcal{F}\hat{\tau}_{\xi_0} = \tau_{\xi_0}\mathcal{F}$  and  $\mathcal{F}\tau_{x_0} = \hat{\tau}_{-x_0}$ . In addition, write  $\|u\|_{(-r)} := \|\langle h^{1-\rho}x'\rangle^{-r}u\|_{L^2}$  for the dual norm to  $\|u\|_{(r)} := \|U_h^{-1}u\|_r$ .

Assume that  $k \geq 0$ . Then, the definition of  $B_h$  combined with the assumptions yield

$$|\langle B\tau_{x_0}\hat{\tau}_{\xi_0}\psi, \tau_{y_0}\hat{\tau}_{\eta_0}\phi\rangle| \le h^{(1-\rho)(|\alpha|+|\beta|)} \|\tau_{x_0}\hat{\tau}_{\xi_0}\psi\|_{(k)} \|\tau_{y_0}\hat{\tau}_{\eta_0}\phi\|_{(-|\beta|)}.$$

In addition, note that for fixed  $\psi, \phi \in \mathcal{S}$ ,

$$\|\tau_{x_0}\hat{\tau}_{\xi_0}\psi\|_{(k)} \sim \langle h^{1-\rho}(x_0)'\rangle^k, \qquad \|\tau_{y_0}\hat{\tau}_{\eta_0}\psi\|_{(-|\beta|)} \sim \langle h^{1-\rho}(y_0)'\rangle^{-|\beta|}.$$

Therefore, (5.2) leads to

$$|\langle B\tau_{x_0}\hat{\tau}_{\xi_0}\psi, \tau_{y_0}\hat{\tau}_{\eta_0}\phi\rangle| \le Ch^{(1-\rho)(|\alpha|+|\beta|)}\langle h^{1-\rho}(x_0)'\rangle^k \langle h^{1-\rho}(y_0)'\rangle^{-|\beta|}.$$
(5.21)

On the other hand, we have by (5.20) that

$$\begin{aligned} |\langle B\tau_{x_0}\hat{\tau}_{\xi_0}\psi,\tau_{y_0}\hat{\tau}_{\eta_0}\phi\rangle| &= \frac{h^{(1-\rho)|\alpha|}}{(2\pi)^n} \Big| \iint e^{i\langle x,\xi\rangle} b_h(x,\xi)\hat{\psi}(\xi-\xi_0)e^{-i\langle x_0,\xi-\xi_0\rangle-i\langle \eta_0,x-y_0\rangle}\bar{\phi}(x-y_0)dxd\xi \Big| \\ &= h^{(1-\rho)|\alpha|} |\mathcal{F}((\tau_{y_0,\xi_0}\chi)b_h)(\eta_0-\xi_0,x_0-y_0)|, \end{aligned}$$
(5.22)

where  $\chi(x,\xi) = e^{i\langle x,\xi\rangle}\hat{\psi}(\xi)\bar{\phi}(x)$ . Combining (5.22) with (5.21) we then have

$$|\mathcal{F}((\tau_{y_0,\xi_0}\chi)\partial_{\xi}^{\alpha}\partial_x^{\beta}a_h)(\eta_0-\xi_0,x_0-y_0)| \le Ch^{(1-\rho)|\beta|} \langle h^{1-\rho}(x_0)' \rangle^k \langle h^{1-\rho}(y_0)' \rangle^{-|\beta|}.$$

Next, note that  $\chi$  can be replaced by any fixed function in  $C_c^{\infty}$  by taking  $\psi, \phi$  with  $\hat{\psi}(\xi)\phi(x) \neq 0$ on supp  $\chi$ . Putting  $\zeta = \eta_0 - \xi_0$  and  $z = x_0 - y_0$ , we obtain that for every  $\tilde{\alpha}, \tilde{\beta} \in \mathbb{N}^n$ 

$$|\mathcal{F}(\partial_{\xi}^{\tilde{\alpha}}\partial_{x}^{\beta}(\tau_{y_{0},\xi_{0}}\chi)\partial_{\xi}^{\alpha}\partial_{x}^{\beta}a_{h})(\zeta,z)| \leq Ch^{(1-\rho)|\beta|} \langle h^{1-\rho}(x_{0})' \rangle^{k} \langle h^{1-\rho}(x_{0}-z)' \rangle^{-|\beta|}$$

Hence,

$$|z^{\tilde{\alpha}}\zeta^{\tilde{\beta}}\mathcal{F}((\tau_{y_0,\xi_0}\chi)\partial_{\xi}^{\alpha}\partial_{x}^{\beta}a_h)(\zeta,z)| \leq Ch^{(1-\rho)|\beta|} \langle h^{1-\rho}(x_0)' \rangle^k \langle h^{1-\rho}(x_0-z)' \rangle^{-|\beta|}.$$

In particular, for every N > 0

$$|\mathcal{F}((\tau_{y_0,\xi_0}\chi)\partial_{\xi}^{\alpha}\partial_x^{\beta}a_h)(\zeta,z)| \le Ch^{(1-\rho)|\beta|} \langle h^{1-\rho}(x_0)' \rangle^{k-|\beta|} \langle \zeta \rangle^{-N} \langle z \rangle^{-N},$$

and, as a consequence, we obtain

$$\partial_{\xi}^{\alpha}\partial_{x}^{\beta}a_{h}(x,\xi) = \partial_{\xi}^{\alpha}\partial_{x}^{\beta}(a(hx,\xi)) = O(h^{(1-\rho)|\beta|}\langle h^{1-\rho}x'\rangle^{k-|\beta|}).$$

This gives the first claim of the lemma for  $k \ge 0$ . For  $k \le 0$ , we consider  $\langle h^{-\rho} x' \rangle^{-k} A$  and use the composition formulae. A nearly identical argument yields the second claim.

5.3. Definition of the second microlocal class. With Proposition 5.5 in place, we are now in a position to define the class of operators with symbols in  $S_{\Gamma,L,\rho}^k$ .

**Definition 6.** Let  $\Gamma \subset U \subset T^*M$  be a co-isotropic submanifold, U an open set, and L a Lagrangian folation on  $\Gamma$ . A *chart for*  $(\Gamma, L)$  is a symplectomorphism

$$\kappa: U_0 \to V, \qquad U_0 \subset U, \qquad V \subset T^* \mathbb{R}^n,$$

such that  $\kappa(U_0 \cap \Gamma) \subset V \cap \Gamma_0$  and  $\kappa_{*,q}L_q = (L_0)_{\kappa(q)}$  for  $q \in \Gamma \cap U$ .

We now define the pseudodifferential operators associated to  $(\Gamma, L)$ .

**Definition 7.** Let M be a smooth, compact manifold and  $U \subset T^*M$  open,  $\Gamma \subset U$  a co-isotropic submanifold, L a Lagrangian foliation on  $\Gamma$  and  $\rho \in [0, 1)$ . We say that  $A : \mathcal{D}'(M) \to C_c^{\infty}(M)$  is a *semiclassical pseudodifferential operator with symbol class*  $S^k_{\Gamma,L,\rho}(U)$  (and write  $A \in \Psi^k_{\Gamma,L,\rho}(U)$ )

if there are charts  $\{\kappa_\ell\}_{\ell=1}^N$  for  $(\Gamma, L)$  and symbols  $\{a_\ell\}_{\ell=1}^N \subset S_{\Gamma,L,\rho}^k(U)$  such that A can be written in the form

$$A = \sum_{\ell=1}^{N} T'_{\ell} \widetilde{Op}_h(a_{\ell}) T_{\ell} + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}$$
(5.23)

where  $T_{\ell}$  and  $T'_{\ell}$  are FIOs quantizing  $\kappa_{\ell}$  and  $\kappa_{\ell}^{-1}$  for  $\ell = 1, \ldots, N$ .

We say that A is a semiclassical pseudodifferential operator with symbol class  $S^k_{\Gamma,\rho}(U)$ , and write  $A \in \Psi^k_{\Gamma,\rho}(U)$ , if there are symbols  $\{a_\ell\}_{\ell=1}^N \subset \widetilde{S^k_{\Gamma,\rho}}(U)$  such that A can be written in the form (5.23).

**Lemma 5.8.** Suppose that  $\kappa : U \to T^* \mathbb{R}^n$  is a chart for  $(\Gamma, L)$ , T quantizes  $\kappa$ , and T' quantizes  $\kappa^{-1}$ . If  $A \in \Psi^k_{\Gamma,L,\rho}(U)$ , then there is  $a \in \widetilde{S^k_{\Gamma,L,\rho}}(U)$ , with  $\operatorname{supp} a(\cdot, \cdot, \lambda) \subset \kappa(U)$ , such that  $TAT' = \widetilde{Op}_h(a) + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}$ . Moreover, if A is given by (5.23), then

$$a \circ K_{\kappa} = \sigma(T'T) \sum_{\ell=1}^{N} \sigma(T'_{\ell}T_{\ell}) \left(a_{\ell} \circ K_{\kappa_{\ell}}\right) + O(h^{1-\rho})_{\widetilde{S^{k-1}_{\Gamma,L,\rho}}}.$$

*Proof.* Note that we can write  $TAT' = \sum_{\ell=1}^{N} TT'_{\ell} \widetilde{Op}_h(a_{\ell}) T_{\ell}T' + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}$ . Next, note that  $TT'_{\ell}$  quantizes  $\kappa \circ \kappa_{\ell}^{-1}$  and that  $T_{\ell}T'$  quantizes  $\kappa_{\ell} \circ \kappa^{-1}$ . Letting  $F_{\ell}$  be a microlocally unitary FIO quantizing  $\kappa_{\ell} \circ \kappa^{-1}$ ,  $F_{\ell}$  satisfies the hypotheses of Proposition 5.5 and we can write

$$T_{\ell}T' = C_{L}F_{\ell}, \qquad TT'_{\ell} = F_{\ell}^{-1}C_{R}$$
  
with  $C_{L}, C_{R} \in \Psi(M)$  satisfying  $\sigma(C_{R}C_{L}) = \left(\sigma(T'_{\ell}T_{\ell}) \circ \kappa_{\ell}^{-1}\right)\left(\sigma(T'T) \circ \kappa_{\ell}^{-1}\right).$  Therefore,  
 $TT'_{\ell}\widetilde{Op}_{h}(a_{\ell})T_{\ell}T' = F_{\ell}^{-1}C_{R}\widetilde{Op}_{h}(a_{\ell})C_{L}F_{\ell} = Op_{h}(b_{\ell}) + (h^{\infty})_{\mathcal{D}' \to C^{\infty}},$   
 $b_{\ell} = \left(\sigma(C_{R}C_{L}) \circ \kappa_{\ell} \circ \kappa^{-1}\right)\left(a_{\ell} \circ K_{\kappa_{\ell} \circ \kappa^{-1}}\right) + O(h^{1-\rho})_{\widetilde{S_{\Gamma,L,\rho}^{k-1}}}.$ 

The lemma follows.

**Lemma 5.9.** Let  $\Gamma \subset U \subset T^*M$  be a co-isotropic submanifold, U an open set, and L a Lagrangian foliation on  $\Gamma$ . There is a principal symbol map

$$\sigma_{\Gamma,L}: \Psi^k_{\Gamma,L,\rho}(U) \to S^k_{\Gamma,L,\rho}(U)/h^{1-\rho}S^{k-1}_{\Gamma,L,\rho}(U)$$

such that for  $A \in \Psi_{\Gamma,L,\rho}^{k_1}(U)$ ,  $B \in \Psi_{\Gamma,L,\rho}^{k_2}(U)$ ,

$$\sigma_{\Gamma,L}(AB) = \sigma_{\Gamma,L}(A)\sigma_{\Gamma,L}(B), \quad \sigma_{\Gamma,L}([A,B]) = -ih\{\sigma_{\Gamma,L}(A), \sigma_{\Gamma,L}(B)\}.$$
(5.24)

Furthermore, the sequence

$$0 \mapsto h^{1-\rho} \Psi^{k-1}_{\Gamma,L,\rho}(U) \xrightarrow{\sigma_{\Gamma,L}} S^k_{\Gamma,L,\rho}(U) / h^{1-\rho} S^{k-1}_{\Gamma,L,\rho}(U) \to 0$$

is exact. The same holds with  $\sigma_{\Gamma}$ ,  $\Psi_{\Gamma,\rho}$  and  $S_{\Gamma,\rho}^k$ .

*Proof.* For A as in (5.23), we define

$$\sigma_{\Gamma,L}(A) = \sum_{\ell=1}^{N} \sigma(T_{\ell}T_{\ell}')(\tilde{a}_{\ell} \circ \kappa)$$

where  $\tilde{a}_{\ell}(x,\xi) := a_{\ell}(x,\xi,h^{-\rho}x')$ . The fact that  $\sigma$  is well defined then follows from Lemma 5.8, and the formulae (5.24) follow from Lemma 5.2.

To see that the sequence is exact, we only need to check that if  $A \in \Psi_{\Gamma,L,\rho}^k$  and  $\sigma_{\Gamma,L}(A) = 0$ , then  $A \in h^{1-\rho}\Psi_{\Gamma,L,\rho}^{k-1}$ . To do this, we may assume that  $WF_h'(A) \subset U$  such that there is a chart  $(\kappa, U)$  for  $(\Gamma, L)$ . Let T be a microlocally unitary FIO quantizing  $\kappa$  and suppose that  $\sigma_{\Gamma,L}(A) \in h^{1-\rho}S_{\Gamma,L,\rho}^{k-1}$ . Then, by the first part of Lemma 5.8 we know  $TAT^{-1} = \widetilde{Op}_h(a) + O(h^{\infty})$  for some  $a \in \widetilde{S_{\Gamma,L,\rho}^k}$ . Then, by the second part of Lemma 5.8, since  $\sigma_{\Gamma,L}(A) \in h^{1-\rho}S_{\Gamma,L,\rho}^{k-1}$ ,  $a \in h^{1-\rho}\widetilde{S_{\Gamma,L,\rho}^{k-1}}$  and in particular,  $A \in h^{1-\rho}\Psi_{\Gamma,L,\rho}^{k-1}$ .

Note that if  $A \in \Psi^{\text{comp}}(M)$ , then  $A \in \Psi^{0}_{\Gamma,L,\rho}$  and  $\sigma(A) = \sigma_{\Gamma}(A)$ . Furthermore, if  $A \in \Psi^{k}_{\Gamma,\rho}$ , then  $A \in \Psi^{k}_{\Gamma,L,\rho}$  and  $\sigma_{\Gamma}(A) = \sigma_{\Gamma,L}(A)$ .

**Lemma 5.10.** Let  $\Gamma \subset U \subset T^*M$  be a co-isotropic submanifold, U an open set, and L a Lagrangian foliation on  $\Gamma$ . There is a non-canonical quantization procedure

$$Op_h^{\Gamma,L}: S^k_{\Gamma,L,\rho}(U) \to \Psi^k_{\Gamma,L,\rho}(U)$$

such that for all  $A \in \Psi_{\Gamma,L,\rho}^k(U)$  there is  $a \in S_{\Gamma,L,\rho}^k(U)$  such that  $Op_h^{\Gamma,L}(a) = A + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}$ . and  $\sigma_{\Gamma,L} \circ Op_h^{\Gamma,L} : S_{\Gamma,L,\rho}^k(U) \to S_{\Gamma,L,\rho}^k(U)/h^{1-\rho}S_{\Gamma,L,\rho}^{k-1}(U)$  is the natural projection map.

Proof. Let  $\{(\kappa_{\ell}, U_{\ell})\}_{\ell=1}^{N}$  be charts for  $(\Gamma, L)$  such that  $\{U_{\ell}\}_{\ell=1}^{N}$  is a locally finite cover for  $U, T_{\ell}$ and  $T'_{\ell}$  quantize respectively  $\kappa_{\ell}$  and  $\kappa_{\ell}^{-1}$ , and  $\sigma(T'_{\ell}T_{\ell}) \in C^{\infty}_{c}(U_{\ell})$  is a partition of unity on U. Let  $a \in S^{k}_{\Gamma,L,\rho}(U)$ . Then, define  $a_{\ell} \in S^{k}_{\Gamma_{0},L_{0},\rho}$  such that  $a_{\ell}(x,\xi,h^{-\rho}x') := (\chi_{\ell}a) \circ \kappa^{-1}(x,\xi)$  where  $\chi_{\ell} \equiv 1$  on  $\operatorname{supp} \sigma(T'_{\ell}T_{\ell})$ . We then define the quantization map

$$Op_h^{\Gamma,L}(a) := \sum_{\ell=1}^N T'_\ell \widetilde{Op}_h(a_\ell) T_\ell.$$

The fact that  $\sigma_{\Gamma,L} \circ Op_h^{\Gamma,L}$  is the natural projection follows immediately. Now, fix  $A \in \Psi_{\Gamma,L,\rho}^k(U)$ . Put  $a_0 = \sigma_{\Gamma,L}(A)$ . Then,  $A = Op_h^{\Gamma,L}(a_0) + h^{1-\rho}A_1$  where  $A_1 \in \Psi_{\Gamma,L,\rho}^{k-1}$ . We define  $a_k = \sigma_{\Gamma,L}(A_k)$  inductively for  $k \ge 1$  by

$$h^{(k+1)(1-\rho)}A_{k+1} = A - \sum_{k=0}^{k} h^{k(1-\rho)}Op_{h}^{\Gamma,L}(a_{k}).$$

Then, letting  $a \sim \sum_k h^{k(1-\rho)} a_k$ , we have  $A = Op_h^{\Gamma,L}(a) + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}$  as claimed.

**Remark 6.** Note that  $E := \sum_{\ell=1}^{N} T_{\ell} T'_{\ell}$  is an elliptic pseudodifferential operator with symbol 1. Therefore, there is  $E' \in \Psi^0$  with  $\sigma(E') = 1$  such that E'EE' = Id. Replacing  $T_{\ell}$  by  $E'T_{\ell}$  and  $T'_{\ell}$  by  $T'_{\ell}E'$ , we may (and will) ask for  $\sum_{\ell=1}^{N} T_{\ell}T'_{\ell} = \text{Id}$ , and so  $Op_h^{\Gamma,L}(1) = \text{Id}$ .

**Lemma 5.11.** Let  $\Gamma \subset U \subset T^*M$  be a co-isotropic submanifold. If  $A \in \Psi^k_{\Gamma,\rho}(U)$  and  $P \in \Psi^m(U)$ with symbol p such that for every  $q \in \Gamma$  we have  $H_p(q) \in T_q\Gamma$ . Then,

$$\frac{i}{h}[P,A] = Op_h^{\Gamma}(H_p a) + O(h^{1-\rho})_{\Psi_{\Gamma,\rho}^{k-1}},$$

where  $a(x,\xi;h) = \sigma_{\Gamma}(A)(x,\xi,h^{-\rho}x').$ 

Proof. Suppose that  $WF'_h(A) \subset U_\ell$  for  $U_\ell \subset U$  open, and suppose that  $\kappa : U_\ell \to T^*\mathbb{R}^n$  is a chart for  $(\Gamma, L)$ . Note that we may assume that  $WF_h(A)' \subset U_\ell$  and then use a partition of unity to cover U with a family  $\{U_\ell\}_\ell$ . Therefore, there exist  $a \in \widetilde{S^k_{\Gamma,\rho}}$  and a Fourier integral operator Tthat is microlocally elliptic on  $U_\ell$  and quantizes  $\kappa$ , such that  $A = T^{-1}\widetilde{Op}_h(a)T + O(h^\infty)_{\mathcal{D}'\to C^\infty}$ . Then, on  $WF_h'(A)$ ,

$$T[P,A]T^{-1} = [TPT^{-1}, \widetilde{Op}_h(a)] + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}.$$

Now,  $TPT^{-1} = Op_h(p \circ \kappa^{-1}) + O(h)_{\Psi^{m-1}}$ . Hence, a direct computation using Lemma 5.3 gives

$$[TPT^{-1}, \widetilde{Op}_h(a)] = -ih\widetilde{Op}_h(c) + O(h^{2-\rho})_{\widetilde{\Psi^{k-2}_{\Gamma_0,\rho}}},$$

with  $c(x,\xi,h^{-\rho}x')=H_{p\circ\kappa^{-1}}(a(x,\xi,h^{-\rho}x'))\in S^{k-1}_{\Gamma,\rho}(U_\ell).$  In particular,

$$[P,A] = -ihT^{-1}\widetilde{Op}_h(c)T + O(h^{2-\rho})_{\Psi_{\Gamma,\rho}^{k-2}}$$

Therefore,  $[P, A] \in h\Psi_{\Gamma, \rho}^{k-1}$ . with symbol  $\sigma_{\Gamma}(ih^{-1}[P, A]) = H_p(a(x, \xi, h^{-\rho}x'))$ .

## 6. An Uncertainty principle for co-isotropic localizers

The first goal of this section is to build a family of cut-off operators  $X_y$  with  $y \in M$  that act as the identity on the shrinking ball  $B(y, h^{\rho})$  and such that they commute with P in a fixed size neighborhood of y. This is the content of section 6.1. The second goal is to control  $||X_{y_1}X_{y_2}||_{L^2 \to L^2}$ in terms of the distance  $d(y_1, y_2)$ , as this distance shrinks to 0. We do this in Section 6.2. Finally, in Section 6.3, we study the consequences of these estimates for the almost orthogonality of  $X_{y_i}$ .

In order to localize to the ball  $B(y, h^{\rho})$  in a way compatible with microlocalization we need to make sense of

$$\chi_y(x) = \tilde{\chi}\big(\tfrac{1}{\varepsilon}h^{-\rho}d(x,y)\big) \qquad \tilde{\chi} \in C_c^{\infty}((-1,1)),$$

as an operator in some anisotropic pseudodifferential calculus. As a function,  $\chi_y$  is in the symbol class  $S_{\Gamma_y,L_y}^{-\infty}$ , where  $\Gamma_y, L_y$  are the co-isotropic submanifold and Lagrangian foliation defined as follows: Fix  $\delta > 0$ , to be chosen small later, and for each  $x \in M$  let

$$\Gamma_y := \bigcup_{|t| < \frac{1}{2} \operatorname{inj}(M)} \varphi_t(\Omega_y), \qquad \Omega_y := \left\{ \xi \in T_y^* M : \ \left| 1 - |\xi|_g \right| < \delta \right\}.$$
(6.1)

In this section, we construct localizers to  $\Gamma_y$  adapted to the Laplacian and study the incompatibility between localization to  $\Gamma_{y_1}$  and  $\Gamma_{y_2}$  as a function of the distance between  $y_1, y_2 \in M$ . Let  $y \in M$ . In what follows we work with the Lagrangian foliation  $L_y$  of  $\Gamma_y$  given by

 $L_y = \{L_{y,\tilde{q}}\}_{\tilde{q}\in\Gamma_y}, \qquad L_{y,\tilde{q}} = (\varphi_t)_*(T_q T_y^* M),$ 

where  $\tilde{q} = \varphi_t(q)$  for some  $|t| < \frac{1}{2} \operatorname{inj}(M)$  and  $q \in \Omega_y$ .

**Remark 7.** In fact, it will be enough for us to show that  $\chi_y(x)\tilde{\chi}(\delta^{-1}(|hD|_g-1)) \in \Psi_{\Gamma_y,L_y,\rho}$  since we will be working near the characteristic variety for the Laplacian.

## 6.1. Co-isotropic cutoffs adapted to the Laplacian.

**Lemma 6.1.** Let  $y \in M$ ,  $0 < \varepsilon < \delta$ ,  $0 \le \rho < 1$ ,  $\tilde{\chi} \in C_c^{\infty}((-1,1))$ , and define the operator  $\chi_{h,y}$  by

$$\chi_{h,y}u(x) := \tilde{\chi}(\frac{1}{\varepsilon}h^{-\rho}d(x,y)) \left[Op_h(\tilde{\chi}(\frac{1}{\varepsilon}(|\xi|_g - 1)))u\right](x).$$
(6.2)

Then,  $\chi_{h,y} \in \Psi_{\Gamma_y,L_y,\rho}^{-\infty}$ .

*Proof.* We will use Lemma 5.7 to prove the claim. First, observe that we may work in a single chart for  $(\Gamma_y, L_y)$  by using a partition of unity. Therefore, suppose that  $B \in \Psi^0$  and  $\kappa : U_0 \to T^* \mathbb{R}^n$  is a chart for  $(\Gamma_y, L_y), V_0 \Subset U_0$ , and T is an FIO quantizing  $\kappa$  that is microlocally unitary on  $V_0$ . Furthermore, since  $\kappa_* L_y = L_0$ , we may assume that  $\kappa(U_0 \cap T_y^* M) \subset T_0^* \mathbb{R}^n$ . Denote the microlocal inverse of T by T'. Then, observe that for A and B with wavefront set in  $V_0$ 

$$\operatorname{ad}_A(TBT') = T \operatorname{ad}_{T'AT}(B)T' + O(h^\infty)_{\mathcal{D}' \to C^\infty}.$$

By a partition of unity, we will work as though  $\chi_{h,y}$  were microsupported in  $U_0$ . We then consider for all N > 0, and  $\alpha, \beta \in \mathbb{N}^n$ ,

$$\begin{split} h^{-2N\rho} |x'|^{2N} \operatorname{ad}_{h^{-\rho}x}^{\alpha} \operatorname{ad}_{hD_x}^{\beta} \left( T\chi_{h,y} T' \right) \\ &= h^{-2\rho N} T(T'|x'|^2 T)^N \operatorname{ad}_{h^{-\rho}T'xT}^{\alpha} \left( \operatorname{ad}_{T'hD_xT}^{\beta} (\chi_{h,y}) \right) T' + O(h^{\infty})_{\mathcal{D}' \to C^{\infty}}. \end{split}$$

In order to prove the requisite estimates, we will actually view  $\chi_{h,y}$  first as an element of the model microlocal class. In particular, we work with  $x \in M$  written in geodesic normal coordinates centered at y, so that

$$\chi_{h,y}u(x) = \tilde{\chi}(\frac{1}{\varepsilon}h^{-\rho}|x|) \left[Op_h(\tilde{\chi}(\frac{1}{\varepsilon}(|\xi|_g - 1)))u\right](x).$$

Then,  $\chi_{h,y} = \widetilde{Op}_h(\frac{1}{\varepsilon}\widetilde{\chi}(\lambda)) Op_h(\widetilde{\chi}(\frac{1}{\varepsilon}(|\xi|-1)))$  is an element of  $\Psi_{\Gamma_0,L_0,\rho}^{-\infty}$  with r = n, and so we can apply Lemma 5.3 to compute  $\operatorname{ad}_A(\chi_{h,y})$  for  $A \in \Psi^{-\infty}(M)$ . In particular,

$$\mathrm{ad}_{T'hD_xT}(\chi_{h,y}) = \widetilde{Op}_h(c) + O(h^\infty)$$
(6.3)

where  $c \in h^{1-\rho} \widetilde{S_{\Gamma_0,L_0,\rho}}$  is supported on  $\{(x,\xi,\lambda) : |x| \leq \varepsilon h^{\rho}, |\lambda| \leq \varepsilon\}$ . Now, suppose  $c \in \widetilde{S_{\Gamma_0,L_0,\rho}}$  is supported on  $\{(x,\xi,\lambda) : |x| \leq \varepsilon h^{\rho}, |\lambda| \leq \varepsilon\}$  and  $B \in \Psi^{-\infty}$  with  $\sigma(B)(0,\xi) = 0$ . Then, again using Lemma 5.3,

$$\mathrm{ad}_B(\widetilde{Op}_h(c)) = \widetilde{Op}_h(c') + O(h^\infty)$$
(6.4)

where  $c' \in h\widetilde{S_{\Gamma_0,L_0,\rho}^{-\infty}}$  is supported on  $\{(x,\xi,\lambda): |x| \leq \varepsilon h^{\rho}, |\lambda| \leq \varepsilon\}.$ 

Now, note that since  $\kappa(T_y^*M) \subset T_0^*\mathbb{R}^n$ , then for all  $i = 1, \ldots n$ ,  $B = T'x_iT$  has symbol  $\sigma(B) = [b(x,\xi)x]_i$  for some  $b \in C^{\infty}(T^*M; \mathbb{M}_{n \times n})$ . Therefore, (6.3) and (6.4) yield

$$\mathrm{ad}_{h^{-\rho}T'xT}^{\alpha}(\mathrm{ad}_{T'hD_xT}^{\beta}(\chi_{h,y})) = h^{(1-\rho)(|\alpha|+|\beta|)}\widetilde{Op}_h(c') + O(h^{\infty}),$$

where  $c' \in \widetilde{S_{\Gamma_0,L_0,\rho}^{-\infty}}$  is supported on  $\{(x,\xi,\lambda): |x| \leq \varepsilon h^{\rho}, |\lambda| \leq \varepsilon\}$ . Finally, using again that  $T'x_iT$  has symbol  $[b(x,\xi)x]_i$ , we have that (6.4) gives

$$\|h^{-2N\rho}|x'|^{2N} \operatorname{ad}_{h^{-\rho}T'xT}^{\alpha}(\operatorname{ad}_{T'hD_{x}T}^{\beta}(\chi_{h,y}))\|_{L^2 \to L^2} \le Ch^{(1-\rho)(|\alpha|+|\beta|)}.$$

42

We next construct a pseudodifferential cutoff,  $X_y \in \Psi_{\Gamma_y,\rho}^{-\infty}$  which is microlocally the identity near  $S_y^*M$  and which essentially commutes with P near y. In particular, we will have

$$\chi_{h,y}X_y = \chi_{h,y} + O(h^\infty)_{\Psi^{-\infty}}.$$

When considering the value of a quasimode u, that is  $h^{\rho}$  close to the point y, this will allow us to effectively work with  $X_y u$  instead.

**Theorem 4.** Let  $y \in M$ ,  $0 < \varepsilon < \delta$ ,  $0 \le \rho < 1$ . Then, there exists  $X_y \in \Psi_{\Gamma_y,\rho}^{-\infty} \subset \Psi_{\Gamma_y,L_y,\rho}^{-\infty}$  with

(1) if  $\chi_{h,y}$  is defined as in (6.2), then

$$\chi_{h,y} X_y = \chi_{h,y} + O(h^{\infty})_{\Psi^{-\infty}}.$$
(6.5)  
(2) WF<sub>h</sub>'([P, X<sub>y</sub>])  $\cap \{(x, \xi) : x \in B(y, \frac{1}{2} \operatorname{inj}(M)), \xi \in \Omega_x\} = \emptyset.$ 

*Proof.* First, we note that we will actually prove that  $X_y \in \Psi_{\Gamma_y,\rho}^{-\infty}$ , and so the result will follow since  $\Psi_{\Gamma_y,\rho}^{-\infty} \subset \Psi_{\Gamma_y,L_y,\rho}^{-\infty}$ . Let  $\mathcal{H} \subset T^*M$  be transverse to the Hamiltonian flow  $H_p$  such that  $\Omega_y \subset \mathcal{H}$ . Next, let  $\varkappa \in C_c^{\infty}((-2,2))$  with  $\varkappa \equiv 1$  on [-1,1] and define  $\varkappa_0 \in C_c^{\infty}(\mathcal{H})$ , by

$$\varkappa_0 = \varkappa (h^{-\rho} d(x, y)) \varkappa (\frac{2}{\delta} (1 - |\xi|_g))$$

where  $\delta$  is as in the definition of  $\Omega_y$ . Let  $\psi \in C_c^{\infty}(T^*M)$  with

 $\psi \equiv 1 \text{ on } B(y, \tfrac{1}{2}\operatorname{inj}(M)) \cap \{ |\xi|_g < 2 \}, \qquad \operatorname{supp} \psi \subset B(y, \tfrac{3}{4}\operatorname{inj}(M)).$ 

Then, let  $\chi_0$  be defined locally by  $H_p\chi_0 \equiv 0$  and  $\chi_0|_{\mathcal{H}} = \varkappa_0$ . so that  $\chi_0 \in S^{-\infty}_{\Gamma_y,\rho}$ . That is,  $\chi_0(\varphi_t(q)) = \psi(\varphi_t(q))\chi_0(q)$  for  $|t| < \operatorname{inj}(M)$  and  $q \in \mathcal{H}$ . Next, observe that there is  $e_0 \in S^{-\infty}_{\Gamma_y,\rho}$  such that

$$-\frac{i}{\hbar}[P, Op_h^{\Gamma_y}(\chi_0)] = h^{1-\rho} Op_h^{\Gamma_y}(e_0), \qquad \operatorname{supp} e_0 \cap B(y, \frac{1}{2}\operatorname{inj}(M)) \subset \bigcup_{|t| < \frac{3}{4}\operatorname{inj}(M)} \varphi_t(\mathcal{H} \cap \operatorname{supp} \partial \varkappa_0).$$

Suppose that there exist  $\chi_{k-1}, e_{k-1} \in S^{-\infty}_{\Gamma_{y},\rho}$  such that

$$-\frac{i}{h}[P, Op_h^{\Gamma_y}(\chi_{k-1})] = h^{k(1-\rho)}Op_h(e_{k-1}), \qquad \operatorname{supp} e_{k-1} \cap B(y, \frac{1}{2}\operatorname{inj}(M)) \subset \bigcup_{|t| < \frac{3}{4}\operatorname{inj}(M)} \varphi_t(\mathcal{H} \cap \operatorname{supp} \partial \varkappa_0).$$

Then, define  $\tilde{\chi}_k \in S^{-\infty}_{\Gamma_y,\rho}$  by solving locally  $H_p \tilde{\chi}_k = e_{k-1}$  and  $\tilde{\chi}_k|_{\mathcal{H}} = 0$ . Note that then

$$\operatorname{supp} \tilde{\chi}_k \cap B(y, \frac{1}{2}\operatorname{inj}(M)) \subset \bigcup_{|t| < \frac{3}{4}\operatorname{inj}(M)} \varphi_t(\mathcal{H} \cap \operatorname{supp} \partial \varkappa_0)$$

and

$$h^{-k(1-\rho)}\sigma\Big(\frac{i}{h}\Big[P,Op_{h}^{\Gamma_{y}}(\chi_{k-1}+h^{k(1-\rho)}\tilde{\chi}_{k})\Big]\Big) = H_{p}\tilde{\chi}_{k} - e_{k-1} = 0.$$

In particular, with  $\chi_k := \chi_{k-1} + h^{k(1-\rho)} \tilde{\chi}_k$ , we obtain  $-\frac{i}{h} [P, Op_h^{\Gamma_y}(\chi_k)] = h^{(k+1)(1-\rho)} Op_h(e_k)$  with  $e_k \in S_{\Gamma_y,\rho}^{-\infty}$  and

$$\operatorname{supp} e_k \cap B(y, \frac{1}{2} \operatorname{inj}(M)) \subset \bigcup_{|t| < \frac{3}{4} \operatorname{inj}(M)} \varphi_t(\mathcal{H} \cap \operatorname{supp} \partial \varkappa_0).$$

Setting

$$X_y = Op_h^{\Gamma_y}(\chi_\infty), \qquad \chi_\infty \sim \left(\chi_0 + \sum_k \chi_{k+1} - \chi_k\right) \Big),$$

we have that  $X_y$  satisfies the second claim and, moreover,  $\chi_{\infty} \equiv 1$  on

$$\bigcup_{|t| \le \frac{1}{4} \operatorname{inj}(M)} \varphi_t \Big( \mathcal{H} \cap \{ d(x, y) < h^{\rho} \} \cap \{ \big| |\xi|_g - 1 \big| < \frac{\delta}{2} \} \Big).$$

To see the first claim, observe that for  $\varepsilon > 0$  small enough,

$$B(y,\varepsilon h^{\rho}) \cap \left\{ \left| |\xi|_g - 1 \right| < \delta \right\} \subset \bigcup_{|t| \le \frac{1}{4} \operatorname{inj}(M)} \varphi_t \Big( \mathcal{H} \cap \left\{ d(x,y) < h^{\rho} \right\} \cap \left\{ \left| |\xi|_g - 1 \right| < \frac{\delta}{2} \right\} \Big).$$

and hence  $\chi_{h,y}X_y = \chi_{h,y}Op_h^{\Gamma,L}(1) + O(h^\infty)_{\Psi^{-\infty}} = \chi_{h,y} + O(h^\infty)_{\Psi^{-\infty}}.$ 

6.2. An uncertainty principle for co-isotropic localizers. Let  $\Gamma(t) \subset T^*\mathbb{R}^n$ ,  $t \in (-\varepsilon_0, \varepsilon_0)$  be a smooth family of co-isotropic submanifolds of dimension n + 1 with

$$\Gamma(0) = \{ (0, x_n, \xi', \xi_n) : x_n \in \mathbb{R}, \, \xi' \in \mathbb{R}^{n-1}, \, \xi_n \in \mathbb{R} \}$$

Assume that  $\Gamma(t)$  is defined by  $\{q_i(t)\}_{i=1}^{n-1}$  with  $q_i(0) = x_i$ . Moreover, assume that there are c, C > 0 such that for  $i = 1, \ldots, n-1$ ,

$$|\{q_i(t), x_i\}| \ge c|t|$$
 on  $\Gamma(0) \cap \Gamma(t), \quad |t| > 0,$  (6.6)

and for all i, j = 1, ..., n - 1, and all  $t \in (-\varepsilon_0, \varepsilon_0)$ ,

$$\{q_i(t), q_j(t)\} = 0, \qquad \{q_i(t), \xi_n\} = 0, \qquad |\{q_i(t), x_j\}| \le Ct^2, \ i \ne j.$$
(6.7)

The main goal of this section is to prove the following proposition

**Proposition 6.2.** Let  $0 < \rho < 1$  and  $\{\Gamma(t) : t \in (-\varepsilon_0, \varepsilon_0)\}$  be as above. Suppose that  $X(t) \in \Psi_{\Gamma(t),\rho}^{-\infty}$  for all  $t \in (-\varepsilon_0, \varepsilon_0)$ , and that there is  $\varepsilon > 0$  such that  $h^{\rho-\varepsilon} \leq |t| < \varepsilon_0$ . Then,

$$||X(0)X(t)||_{L^2 \to L^2} \le Ch^{\frac{n-1}{2}(2\rho-1)} t^{\frac{1-n}{2}}.$$

*Proof.* We begin by finding a convenient chart for  $\Gamma(t)$ . By Darboux's theorem, there is a smooth family of sympletomorphisms  $\kappa_t : V_1 \to V_2$  such that for  $j = 1, \ldots, n-1$ ,

$$\kappa_t^*(q_j(t)) = y_j, \qquad \kappa_t^* \xi_n = \eta_n, \tag{6.8}$$

where  $V_1, V_2$  are simply connected neighborhoods of 0. Note that  $\kappa_t(\Gamma(0)) = \Gamma(t)$  with this setup, so  $\kappa_t^{-1}$  is a chart for  $\Gamma(t)$ . By [Zwo12, Theorem 11.4], the symplectomorphism  $\kappa_t$  can be extended to a family of symplectomorphisms on  $T^*\mathbb{R}^n$  that is the identity outside a compact set, and such that there is a smooth family of symbols  $p_t \in C^{\infty}(T^*\mathbb{R}^n)$  satisfying  $\partial_t \kappa_t = (\kappa_t)_* H_{p_t}$ .

Now, let  $U(t): L^2 \to L^2$  solve

$$(hD_t + Op_h(p_t))U(t) = 0, \qquad U(0) = \text{Id}$$

Then, U(t) is microlocally unitary from  $V_1$  to  $V_2$  and quantizes  $\kappa_t$ . Moreover,

$$U(t) = \frac{1}{(2\pi h)^n} \int_{\mathbb{R}^n} e^{\frac{i}{h}(\phi(t,x,\eta) - \langle y,\eta \rangle)} b(t,x,\eta;h) d\eta$$

where  $b(t, \cdot) \in S^{\text{comp}}(T^*\mathbb{R}^n)$  and the phase function  $\phi(t, \cdot) \in C^{\infty}(T^*\mathbb{R}^n; \mathbb{R})$  satisfies

$$\partial_t \phi + p_t(x, \partial_x \phi) = 0, \qquad \phi(0, x, \eta) = \langle x, \eta \rangle,$$

for all  $t \in (-\varepsilon_0, \varepsilon_0)$ . Since U(t) is microlocally unitary, it is enough to estimate the operator

A(t) := X(0)X(t)U(t).

First, note that since  $X(t) \in \Psi_{\Gamma(t),\rho}^{-\infty}$ , and U(t) quantizes  $\kappa_t$ , there exist  $a(t) \in \widetilde{S_{\Gamma_0,\rho}^{-\infty}}$  with  $t \in (-\varepsilon_0, \varepsilon_0)$  such that  $X(t) = U(t)\widetilde{Op}_h(a(t))[U(t)]^* + O(h^{\infty})_{L^2 \to L^2}$  and so

$$A(t) = \widetilde{Op}_h(a(0))U(t)\widetilde{Op}_h(a(t)) + O(h^{\infty})_{L^2 \to L^2}.$$

Fix N > n-1 and let  $\chi = \chi(\lambda) \in \widetilde{S_{\Gamma_0,\rho}^{-N}}$  be such that  $|\chi(\lambda)| \ge c \langle \lambda \rangle^{-N}$ . Now, since  $a(t) \in \widetilde{S_{\Gamma_0,\rho}^{-\infty}}$ , by the elliptic parametrix construction there are  $e_L(t), e_R(t) \in \widetilde{S_{\Gamma_0,\rho}^{-\infty}}$  such that

 $\widetilde{Op}_h(e_L(t))\widetilde{Op}_h(\chi) = \widetilde{Op}_h(a(t)) + O(h^{\infty})_{L^2 \to L^2}, \quad \widetilde{Op}_h(\chi)\widetilde{Op}_h(e_R(t)) = \widetilde{Op}_h(a(t)) + O(h^{\infty})_{L^2 \to L^2},$ for all  $t \in (-\varepsilon_0, \varepsilon_0)$ . Note that we are implicitly using the fact that a(t) is compactly supported in  $(x, \xi)$  to handle the fact that  $\chi$  is not compactly supported in  $(x, \xi)$ . Thus,

$$A(t) = \widetilde{Op}_h(e_L(0))\widetilde{Op}_h(\chi)U(t)\widetilde{Op}_h(\chi)\widetilde{Op}_h(e_R(t)) + O(h^{\infty})_{L^2 \to L^2}.$$

Since  $\widetilde{Op}_h(e_L(t))$  and  $\widetilde{Op}_h(e_R(t))$  are  $L^2$  bounded uniformly in  $t \in (-\varepsilon_0, \varepsilon_0)$ , we estimate

$$\widetilde{A}(t) := \widetilde{Op}_h(\chi) U(t) \widetilde{Op}_h(\chi)$$

In fact, we estimate  $B(t) := \tilde{A}(t)(\tilde{A}(t))^*$  by considering its kernel.

$$B(t;x,y) = \int U(t)(x,w)U(t)^{*}(w,y)\chi(h^{-\rho}x')\chi(h^{\rho}y')\chi(h^{-\rho}w')^{2}dw$$
$$= \frac{1}{(2\pi h)^{2n}}\int e^{\frac{i}{h}\Phi(t,x,w,y,\eta,\xi)}b(t,x,\eta)\bar{b}(t,y,\xi)\chi(h^{-\rho}x')\chi(h^{-\rho}y')\chi(h^{-\rho}w')^{2}dwd\eta d\xi,$$

with  $\Phi(t, x, w, y, \eta, \xi) = \phi(t, x, \eta) - \phi(t, y, \xi) + \langle w, \xi - \eta \rangle$ . First, performing stationary phase in  $(w_n, \eta_n)$  yields

$$B(t;x,y) = \frac{1}{(2\pi h)^{2n-1}} \int F(t,x,w',\xi_n) \overline{F(t,y,w',\xi_n)} dw' d\xi_n,$$
  

$$F(t,x,w',\xi_n) := \int e^{\frac{i}{h}(\phi(t,x,\eta',\xi_n) - \langle w',\eta' \rangle)} b_1(t,x,\eta',\xi_n) \chi(h^{-\rho}x') \chi(h^{-\rho}w')^2 d\eta$$

for some  $b_1 \in S^{\text{comp}}(T^*\mathbb{R}^n)$ . Next, note that since  $\phi(0, x, \eta) = \langle x, \eta \rangle$ ,

$$\phi(t, x, \eta) - \langle x, \eta \rangle = t\phi(t, x, \eta)$$

with  $\tilde{\phi}$  such that for every multi-index  $\alpha$  there exists  $C_{\alpha} > 0$  with  $|\partial_{t,x,\eta}^{\alpha} \tilde{\phi}| \leq C_{\alpha}$ .

Next, we claim that there exists C > 0 such that

$$\|(\partial_{\eta'}^2 \tilde{\phi}(t, x, \eta))^{-1}\| \le C \qquad \text{if} \quad (x, \eta) \in \Gamma(0), \ \partial_{\eta'} \phi(t, x, \eta) = 0.$$

$$(6.9)$$

We postpone the proof of (6.9) and proceed to finish the proof of the lemma.

To continue the proof, note that modulo an  $O(h^{N\varepsilon})$  error, we may assume that the integrand of B(t;x,y) is supported, in  $\{(x,y,w'): |x'| \leq h^{\rho-\varepsilon}, |y'| \leq h^{\rho-\varepsilon}, |w'| \leq h^{\rho-\varepsilon}\}$ , and  $h^{\rho-\varepsilon} \leq |t|$ . Therefore, the bound in (6.9) continues to hold on the support of the integrand. By (6.9) and

$$\partial_{\eta'}^2 \left( \phi(t, x, \eta) - t \dot{\phi}(t, x, \eta) \right) = 0, \tag{6.10}$$

there is a unique critical point  $\eta'_c(t, x, w', \xi_n)$  for the map  $\eta' \mapsto \phi(t, x, \eta', \xi_n) - \langle w', \eta' \rangle$ , in an O(1) neighborhood of  $\eta'_c$ . In particular,  $\eta'_c$  is the unique solution to  $\partial_{\eta'}\phi(t, x, \eta'_c, \xi_n) - w' = 0$ .

Next, again using (6.10), by applying the method of stationary phase in  $\eta'$  to F, with small parameter h/t, we obtain

$$B(t;x,y) = \frac{1}{(2\pi h)^n t^{n-1}} \int e^{\frac{i}{h} \Phi_1(t,x,w',y,\xi_n)} B_1(t;x,y,w',\eta'_c,\xi) dw' d\xi_n,$$
  

$$\Phi_1(t,x,w',y,\xi_n) := \Psi(t,x,w',\xi_n) - \Psi(t,y,w',\xi_n),$$
  

$$\Psi(t,x,w',\xi_n) := \phi(t,x,\eta'_c(t,x,w',\xi_n),\xi_n) - \langle w',\eta'_c(t,x,w',\xi_n) \rangle,$$
  

$$B_1(t;x,y,w',\eta',\xi) := b_2(t,x,\eta',\xi_n) \bar{b}(t,y,\xi',\xi_n) \chi(h^{-\rho}x') \chi(h^{-\rho}y') \chi(h^{-\rho}w')^2.$$

for some  $b_2 \in S^{\text{comp}}(T^*\mathbb{R}^n)$ . Next, observe that  $\partial_{x_n}\partial_{\xi_n}\Psi(t, x, w', \xi_n) = 1 + O(t)$ , and therefore, there exist c > 0 and a function  $g = g(x', y, w', \xi_n)$  so that  $|\partial_{\xi_n}\Phi_1| \ge c|x_n - g_n|$ . In particular, integration by parts in  $\xi_n$  shows that for any N > 0 there is  $C_N > 0$  such that

$$|B(t;x,y)| \le C_N h^{-n} t^{1-n} h^{\rho(n-1)} \chi(h^{-\rho} y') \chi(h^{-\rho} x') \frac{h^{2N} + h^N |x_n - g_n|}{(h^2 + |x_n - g_n|^2)^N}$$

Applying Schur's lemma together with the fact that there exists C > 0 such that for all t

$$\sup_{x} \int |B(t;x,y)| dy + \sup_{y} \int |B(t;x,y)| dy \le Ch^{(2\rho-1)(n-1)} t^{1-n}$$

yields that  $||B(t)||_{L^2 \to L^2} \leq Ch^{(2\rho-1)(n-1)}t^{1-n}$ , for all  $t \in (-\varepsilon_0, \varepsilon_0)$ , and hence  $||X(0)X(t)||_{L^2 \to L^2} \leq Ch^{\frac{n-1}{2}(2\rho-1)}t^{\frac{1-n}{2}}$ , as claimed.

Proof of the bound in (6.9). Let  $\phi_t(x,\eta) := \phi(t,x,\eta)$  and  $\varphi_t(x,y,\eta) := \phi_t(x,\eta) + \langle y,\eta \rangle$ . Then,  $C_{\varphi_t} = \{(x,y,\eta): \partial_\eta \phi_t(x,\eta) = y\}$  and so

$$\Lambda_{\varphi_t} = \{ (x, \partial_x \phi_t(x, \eta), \partial_\eta \phi_t(x, \eta), -\eta) \} \subset T^* \mathbb{R}^n \times T^* \mathbb{R}^n$$

In particular, since  $\Lambda_{\varphi_t}$  is the twisted graph of  $\kappa_t$ , we have that  $\kappa_t$  is characterized by

$$\kappa_t(\partial_\eta \phi_t(x,\eta),\eta) = (x,\partial_x \phi_t(x,\eta)).$$

Furthermore, since  $\kappa_t(\Gamma(0)) = \Gamma(t)$ , we have

 $\Gamma(t) = \{(x,\xi) : \kappa_t(y,\eta) = (x,\xi), \ y = \partial_\eta \phi_t(x,\eta), \ \xi = \partial_x \phi_t(x,\eta), \ (y,\eta) \in \Gamma(0) \}.$ Then, using  $\kappa_t^* \xi_n = \eta_n$ ,

$$\Gamma(t) = \{(x,\xi): \xi' = \partial_{x'}\phi_t(x,\eta), \ \partial_{\eta'}\phi_t(x,\eta) = 0, \ \xi_n = \eta_n, \ \eta \in \mathbb{R}^n\}.$$

Next, let  $\tilde{p} := (\tilde{x}, \tilde{\eta}) \in \Gamma(0)$  be such that  $\partial_{\eta'} \phi_t(\tilde{x}, \tilde{\eta}) = 0$ . Without loss of generality, in what follows we assume that  $\tilde{x}_n = 0$ . Letting  $\Gamma_0(t) := \Gamma(t)|_{\{x_n=0\}}$  we have that

$$\Gamma_0(t) = \{ (x,\xi) : \xi' = \partial_{x'} \phi_t(x,\eta), \ \partial_{\eta'} \phi_t(x,\eta) = 0, \ x_n = 0, \ \xi_n = \eta_n, \ \eta \in \mathbb{R}^n \}.$$

In particular, letting  $\tilde{\xi} := (\partial_{x'} \phi_t(\tilde{p}), \tilde{\eta}_n)$ , and  $\tilde{p}_0 := (\tilde{x}, \tilde{\xi})$ , we have  $\tilde{p}_0 \in \Gamma_0(t) \cap \Gamma_0(0)$ , and

$$T_{\tilde{p}_0}\Gamma_0(t) = \{ (\delta_x, \delta_{\xi}) : \delta_{\xi'} = \partial_x \partial_{x'} \phi_t(\tilde{p}) \delta_x + \partial_\eta \partial_{x'} \phi_t(\tilde{p}) \delta_\eta, \\ \partial_x \partial_{\eta'} \phi_t(\tilde{p}) \delta_x + \partial_\eta \partial_{\eta'} \phi_t(\tilde{p}) \delta_\eta = 0, \ \delta_{x_n} = 0, \ \delta_{\xi_n} = \delta_{\eta_n}, \ \delta_\eta \in \mathbb{R}^n \}.$$

Next, we note that  $\partial_{x_n} \in T_{\tilde{p}_0}\Gamma(t)$  and  $H_{q_i(t)} \in T_{\tilde{p}_0}\Gamma(t)$  for all  $i = 1, \ldots, n-1$ . Therefore, since  $\partial_{x_n}q_i(t) = 0$ , we also know that  $H'_{q_i(t)} := (\partial_{\xi'}q_i(t), 0, -\partial_{x'}q_i(t), 0) \in T_{\tilde{p}_0}\Gamma_0(t)$  for all  $i = 1, \ldots, n-1$ . We claim that there exists C > 0 such that for all  $v = (\delta_{x'}, 0, \delta_{\xi'}, 0) \in \operatorname{span}\{H'_{q_i(t)}\}_{i=1}^{n-1} \subset T_{\tilde{p}_0}\Gamma(t)$  we have

$$\|\delta_{x'}\| \ge Ct \|\delta_{\xi'}\|. \tag{6.11}$$

Suppose that the claim in (6.11) holds. Then, note that for each such v, since  $\delta_{x_n} = 0$  and  $\delta_{\xi_n} = 0$ , we have that there is  $\delta_{\eta'} \in \mathbb{R}^{n-1}$  such that

$$\delta_{\xi'} = \partial_{x'}^2 \phi_t(\tilde{p}) \delta_{x'} + \partial_{\eta'x'}^2 \phi_t(\tilde{p}) \delta_{\eta'}, \qquad \partial_{x'\eta'}^2 \phi_t(\tilde{p}) \delta_{x'} + \partial_{\eta'}^2 \phi_t(\tilde{p}) \delta_{\eta'} = 0.$$

Using that  $\partial_{x'\eta'}^2 \phi_t(\tilde{p}) = \mathrm{Id} + O(t)$  and  $\partial_{x'}^2 \phi_t(\tilde{p}) = O(t)$ , we conclude that

$$\partial_{\eta'}^2 \phi_t(\tilde{p}) [\partial_{\eta'x'}^2 \phi_t(\tilde{p})]^{-1} \delta_{\xi'} = \left( \partial_{\eta'}^2 \phi_t(\tilde{p}) [\partial_{\eta'x'}^2 \phi_t(\tilde{p})]^{-1} \partial_{x'}^2 \phi_t(\tilde{p}) - \partial_{x'\eta'}^2 \phi_t(\tilde{p}) \right) \delta_{x'},$$

and so

$$\partial_{\eta'}^2 \phi_t(\tilde{p}) (\mathrm{Id} + O(t)) \delta_{\xi'} = (-\mathrm{Id} + O(t)) \delta_{x'}.$$
(6.12)

Let  $H'_{q_i(t)} = (\delta^{(i)}_{x'}, 0, \delta^{(i)}_{\xi'}, 0)$ . Since  $\tilde{p}_0 \in \Gamma(t) \cap \Gamma(0)$ , assumptions (6.6) and (6.7) yield that the vectors  $\{\delta^{(i)}_{x'}\}_{i=1}^{n-1}$  are linearly independent. Indeed, setting  $e_i := (\delta_{ij})_{j=1}^{n-1} \in \mathbb{R}^{n-1}$ ,

$$\delta_{x'}^{(i)} = \partial_{\xi_i} q_i(t) e_i + O(t^2), \qquad |\partial_{\xi_i} q_i(t)| \ge Ct, \tag{6.13}$$

for t small enough. Furthermore, (6.12) then yields that  $\{\delta_{\xi'}^{(i)}\}_{i=1}^{n-1}$  are linearly independent. Then, combining (6.12) with (6.11) yields (6.9) as claimed.

To finish it only remains to prove (6.11). Let  $v = (\delta_{x'}, 0, \delta_{\xi'}, 0) \in \operatorname{span}\{H'_{q_i(t)}\}_{i=1}^{n-1}$ . Then, there is  $a \in \mathbb{R}^{n-1}$  such that  $\delta_{x'} = \sum_{i=1}^{n-1} a_i \delta_{x'}^{(i)}$  and  $\delta_{\xi'} = \sum_{i=1}^{n-1} a_i \delta_{\xi'}^{(i)}$ . Next, note that by (6.13) we have  $\|\delta_{x'}\| \ge \|a\|(Ct + O(t^2))$ . Since  $\|\delta_{\xi'}\| \le C_0 \|a\|$  for some  $C_0 > 0$ , the claim in (6.11) follows.  $\Box$ 

For each  $x \in M$  let  $\Gamma_x$  be as in (6.1). (See Figure 6.2 for a schematic representation of these two co-isotropic submanifolds.) Then we have the following result.

**Corollary 6.3.** Let  $0 < \rho < 1$ ,  $0 < \varepsilon < \rho$ , and  $\gamma(t) : (-\varepsilon_0, \varepsilon_0) \to M$  be a unit speed geodesic. Then, for  $X(t) \in \Psi_{\Gamma_{\gamma(t)},\rho}^{-\infty}$  and h such that  $h^{\rho-\varepsilon} \leq |t| < \varepsilon_0$ ,

$$||X(0)X(t)||_{L^2 \to L^2} \le Ch^{\frac{n-1}{2}(2\rho-1)} t^{\frac{1-n}{2}}.$$

*Proof.* To do this, we study the geometry of the flow-out co-isotropics  $\Gamma_{\gamma(t)}$ . Namely, we prove that  $\Gamma_{\gamma(t)}$  is defined by some functions  $\{q_i(t)\}_{i=1}^n$  with  $q_i(0) = x_i$  and satisfying (6.6) and (6.7). We then apply Proposition 6.2 to  $\Gamma(t) = \kappa^{-1}(\Gamma_{\gamma(t)})$ , for a suitable symplectomorphism  $\kappa$ .

Fix coordinates  $(x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$  on M so that  $\gamma(t) = (0, t)$ , and for each  $t \in (-\varepsilon_0, \varepsilon_0)$  let  $\mathcal{H}_t$  be the submanifold transverse to the Hamiltonian vector field  $H_p$  defined by

$$\mathcal{H}_t := \{ (x', t, \xi', \xi_n) : 2\xi_n > |\xi|_g, |x'| \le \delta_0 \},\$$

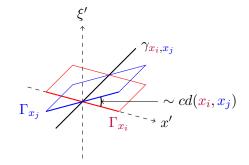


FIGURE 2. A pictorial representation of the co-isotropics involved in Corollary 6.3 where  $\gamma_{x_i,x_j}$  is the geodesic from  $x_i$  to  $x_j$ . Localization to both  $\Gamma_{x_i}$  and  $\Gamma_{x_j}$ implies localization in the non-symplectically orthogonal directions, x' and  $\xi'$ . The uncertainty principle then rules this behavior out.

where  $\delta_0 > 0$  is chosen such that  $\Gamma_{\gamma(t)} \cap \mathcal{H}_t = \{(0, t, \xi', \xi_n) : 2\xi_n > |\xi|_g, ||\xi|_g - 1| < \delta\}.$ 

In particular, as a subset of  $\{ ||\xi|_g - 1| < \delta \}$ ,  $\Gamma_{\gamma(t)} \cap \mathcal{H}_t$  is defined by the coordinate functions  $\{x_i\}_{i=1}^{n-1}$ . For each  $t \in (-\varepsilon_0, \varepsilon_0)$  let  $\tilde{q}_i(t) : \mathcal{H}_t \to \mathbb{R}$  be given by  $\tilde{q}_i(t) = x_i$  for  $i = 1, \ldots, n-1$ . Then, define  $\{q_i(t)\}_{i=1}^{n-1}$  on  $T^*M$  by

$$H_p q_i(t) = 0, \qquad q_i(t)|_{\mathcal{H}_t} = \tilde{q}_i(t)$$

Note that for all  $t H_p(H_{q_i(t)}q_j(t)) = 0$  and

$$\{q_i(t), q_j(t)\} \mid_{\mathcal{H}_t} = \partial_{\xi_n} q_i(t) \partial_{x_n} q_j(t) - \partial_{\xi_n} q_j(t) \partial_{x_n} q_i(t) + H_{q_i(t)} q_j(t),$$

where  $\tilde{H}$  is the Hamiltonian vector field in  $T^*\{x_n = t\}$ . In particular, since  $\partial_{\xi_n} \tilde{q}_i(t) = 0$  and  $\tilde{H}_{q_i(t)}$  is tangent to  $\mathcal{H}_t$ , we have  $\{q_i(t), q_j(t)\}|_{\mathcal{H}_t} = 0$ . Hence,  $\{q_i(t), q_j(t)\} \equiv 0, \{q_i(t), p\} = 0, q_i(0) = x_i$ , and  $\{q_i(t)\}_{i=1}^{n-1}$  define  $\Gamma_{\gamma(t)}$ . Next, observe that there exists  $s \in \mathbb{R}$  such that for each  $i = 1, \ldots, n-1, q_i(0)(x,\xi) = x_i(\varphi_s(x,\xi))$  with  $\varphi_s(x,\xi) \in \mathcal{H}_0$ . Since  $\partial_{\xi_n} p \neq 0$  on  $\mathcal{H}_0$ , for E near 0 there exist  $a_E$  and  $e_E$  such that

$$p(x,\xi) - E = e_E(x,\xi)(\xi_n - a_E(x,\xi'))$$

with  $e_{E} > c$  for some constant c > 0. In particular,  $\varphi_{s} = \exp(sH_{p})$  is a reparametrization of  $\tilde{\varphi}_{s} := \exp(s(H_{\xi_{n}-a_{E}}))$  on  $\{p = E\}$  and we have that for  $(x,\xi) \in \{p = E\}$ , and all  $i = 1, \ldots, n-1$ ,

$$q_i(0)(x,\xi) = x_i(\tilde{\varphi}_{-x_n}(x,\xi)) = x_i + x_n \partial_{\xi_i} a_{_E}(x,\xi') + O(x_n^2)_{C^{\infty}}$$

In particular, on  $\mathcal{H}_t \cap \{p = E\}$ , since  $H_{q_i(t)}$  is tangent to  $\{p = E\}$ , we have

$$\{q_j(t), q_i(0)\}|_{\mathcal{H}_t \cap \{p=E\}} = \partial_{\xi_n} q_j(t) \partial_{x_n} q_i(0) - \partial_{x_n} q_j(t) \partial_{\xi_n} q_i(0) + \tilde{H}_{q_j(t)} q_i(0) = O(t^2) + \partial_{\xi_j} (t \partial_{\xi_i} a_E).$$

Now, since  $\partial_{\xi}^2 p|_{T\{p=E\}} > 0$ , and for all  $i, j = 1, \dots, n$ 

$$\partial_{\xi_i\xi_j}p = \partial_{\xi_j}\partial_{\xi_i}e_{_E}(\xi_n - a_{_E}) + \partial_{\xi_i}e_{_E}(\delta_{nj} - \partial_{\xi_j}a_{_E}) + \partial_{\xi_j}e_{_E}(\delta_{ni} - \partial_{\xi_i}a_{_E}) - e_{_E}\partial_{\xi_j}\partial_{\xi_i}a_{_E},$$

then, as quadratic forms,  $\partial_{\xi}^2 p|_{T_{\{p=E\}}} = -e_E \partial_{\xi}^2 a_E|_{T_{\{p=E\}}}$ . Hence,  $\partial_{\xi'}^2 a_E < 0$ , and there is c > 0 with

$$c\delta_{ij}t + O(t^2) \le |\{q_i(t), q_j(0)\}|_{\mathcal{H}_t \cap \{p=E\}}| \le C\delta_{ij}t + O(t^2).$$

Then,  $c\delta_{ij}t + O(t^2) \leq |\{q_i(t), q_j(0)\}|_{\{p=E\}}| \leq C\delta_{ij}t + O(t^2)$  by invariance under  $H_p$ . Since E small is arbitrary, this holds on  $\Gamma_{\gamma(0)} \cap \Gamma_{\gamma(t)}$ .

Now, by Darboux's theorem, there is a symplectomorphism  $\kappa$  such that for all  $i = 1, \ldots, n-1$  $\kappa^* q_i(0) = x_i$  and  $\kappa^* p = \xi_n$ . In particular,  $\kappa^{-1}(\Gamma_{\gamma(0)}) \subset \Gamma(0) = \{(0, x_n, \xi', \xi_n) : x_n \in \mathbb{R}, \xi \in \mathbb{R}^{n-1} \times \mathbb{R}\}$  and, abusing notation slightly by relabeling  $q_i(t) = \kappa^* q_i(t)$ , we have that (6.6) and (6.7) hold. In particular, Proposition 6.2 applies to  $\Gamma(t) = \kappa^{-1}(\Gamma_{\gamma(t)})$ .

Now, let U be a microlocally unitary quantization of  $\kappa$ , and  $X(t) \in \Psi^{-\infty}_{\Gamma_{\gamma(t)},\rho}$ . Then,  $U^{-1}X(t)U \in \Psi^{-\infty}_{\Gamma(t),\rho}$  and hence the corollary is proved.

6.3. Almost orthogonality for coisotropic cutoffs. In this section, we finally prove an estimate which shows that co-isotropic cutoffs associated with  $\Gamma_{x_i}$  for many  $x_i$  are almost orthogonal. This, together with the fact that these cutoffs respect pointwise values near  $x_i$ , is what allows us to control the number of points at which a quasimode may be large.

**Proposition 6.4.** Let  $\{B(x_i, R)\}_{i=1}^{N(h)}$  be a  $(\mathfrak{D}, R)$ -good cover for M, and  $X_i \in \Psi_{\Gamma_{x_i}, \rho}^{-\infty}$   $i = 1, \ldots, N(h)$ , with uniform symbol estimates. Then, there are C > 0 and  $h_0 > 0$  such that for all  $0 < h < h_0$ ,  $\mathcal{J} \subset \{1, \ldots, N(h)\}$ , and  $u \in L^2(M)$ , we have

$$\sum_{j \in \mathcal{J}} \|X_j u\|_{L^2}^2 \le C \Big( 1 + (h^{2\rho-1} R^{-1})^{\frac{n-1}{2}} |\mathcal{J}|^{\frac{3n+1}{2n}} \Big( 1 + (h^{2\rho-1} R^{-1})^{\frac{n-1}{4}} \Big) \Big) \|u\|_{L^2}^2 \tag{6.14}$$

*Proof.* To prove this bound we will decompose the sum in (6.14) as

$$\sum_{i \in \mathcal{J}} \|X_i u\|_{L^2}^2 = \left\| \sum_{i \in \mathcal{J}} X_i u \right\|_{L^2}^2 - \left\langle \sum_{\substack{i,j \in \mathcal{J} \\ i \neq j}} X_j^* X_i u, u \right\rangle.$$
(6.15)

First, we note that by Corollary 6.3, there exists C > 0 such that for  $i \neq j$ 

$$||X_j^*X_i|| \le Ch^{(n-1)(\rho-\frac{1}{2})}d(x_i, x_j)^{\frac{1-n}{2}}.$$

Therefore, by the the Cotlar-Stein lemma,

$$\Big\| \sum_{j \in \mathcal{J}} X_j \Big\| \le \sup_{j \in \mathcal{J}} \Big( \|X_j\| + \sum_{i \in \mathcal{J} \setminus \{j\}} \|X_j^* X_i\|^{\frac{1}{2}} \Big) \le 2 + Ch^{\frac{n-1}{2}(\rho - \frac{1}{2})} \sup_{j \in \mathcal{J}} \sum_{i \in \mathcal{J} \setminus \{j\}} d(x_i, x_j)^{\frac{1-n}{4}}.$$

To estimate the sum, observe that there exists C > 0 such that for any  $j \in \mathcal{J}$  and any positive integer  $k \frac{1}{C} 2^{kn} \leq \#\{i: 2^k R \leq d(x_i, x_j) \leq 2^{k+1} R\} \leq C 2^{(k+1)n}$ . In particular, there is C > 0 such that for any  $j \in \mathcal{J}$ 

$$\sum_{i \in \mathcal{J} \setminus \{j\}} d(x_i, x_j)^{\frac{1-n}{4}} \le C \sum_{k=0}^{\frac{1}{n} \log_2 |\mathcal{J}|} 2^{kn} (2^k R)^{\frac{1-n}{4}} \le C |\mathcal{J}|^{\frac{3n+1}{4n}} R^{\frac{1-n}{4}}.$$
 (6.16)

Therefore, we shall bound the first term in (6.15) using

$$\left\|\sum_{j\in\mathcal{J}}X_{j}\right\| \leq C + Ch^{\frac{n-1}{2}(\rho-\frac{1}{2})}R^{\frac{1-n}{4}}|\mathcal{J}|^{\frac{3n+1}{4n}}.$$
(6.17)

We next proceed to control the second term in (6.15). Let  $\tilde{X}_j \in \Psi_{\Gamma_{x_j},\rho}^{-\infty}$  such that  $\tilde{X}_j X_j = X_j + O(h^{\infty})_{L^2 \to L^2}$ . By the Cotlar-Stein Lemma,

$$\left\|\sum_{\substack{i,j\in\mathcal{J}\\i\neq j}} X_j^* X_i\right\| \le \sup_{\substack{k,\ell\in\mathcal{J}\\k\neq\ell}} \sum_{\substack{i,j\in\mathcal{J}\\i\neq j}} \|X_k^* \tilde{X}_\ell X_\ell X_j^* \tilde{X}_j^* X_i\|^{\frac{1}{2}} + O(h^\infty |\mathcal{J}|^2).$$
(6.18)

By Corollary 6.3 there exists C > 0 such that for  $k \neq \ell, i \neq j$ ,

$$\|X_k^* \tilde{X}_\ell X_\ell X_j^* \tilde{X}_j^* X_i\| \le Ch^{(n-1)(2\rho-1)} \min\{1, h^{\frac{(n-1)}{2}(2\rho-1)} d(x_j, x_\ell)^{-\frac{n-1}{2}}\} (d(x_k, x_\ell) d(x_j, x_i))^{\frac{1-n}{2}}.$$

Using that  $\sup_{\substack{k,\ell\in\mathcal{J}\\k\neq\ell}} d(x_k,x_\ell)^{\frac{1-n}{4}} \leq R^{\frac{1-n}{4}}$ , adding first in  $i \in \mathcal{J}\setminus\{j\}$  in (6.18), and combining with the bound in (6.16), yields

$$\left\|\sum_{\substack{i,j\in\mathcal{J}\\i\neq j}} X_j^* X_i\right\| \le Ch^{\frac{n-1}{2}(2\rho-1)} (1+h^{\frac{n-1}{4}(2\rho-1)} |\mathcal{J}|^{\frac{3n+1}{4n}} R^{\frac{1-n}{4}}) |\mathcal{J}|^{\frac{3n+1}{4n}} R^{\frac{1-n}{2}}.$$
 (6.19)

In particular, combining (6.17) and (6.19) into (6.15) we obtain

$$\begin{split} \sum_{i\in\mathcal{J}} \|X_i u\|^2 &\leq C \Big( 1 + h^{(n-1)(\rho - \frac{1}{2})} R^{\frac{1-n}{2}} |\mathcal{J}|^{\frac{3n+1}{2n}} + h^{\frac{3(n-1)}{4}(2\rho - 1)} R^{\frac{3(1-n)}{4}} |\mathcal{J}|^{\frac{3n+1}{2n}} \Big) \|u\|_{L^2}^2 \\ &\leq C \big( 1 + h^{(n-1)(\rho - \frac{1}{2})} R^{\frac{1-n}{2}} (1 + (h^{2\rho - 1} R^{-1})^{\frac{n-1}{4}}) |\mathcal{J}|^{\frac{3n+1}{2n}} ) \|u\|_{L^2}^2. \end{split}$$

## References

- [Bla10] David E. Blair. Riemannian geometry of contact and symplectic manifolds, volume 203 of Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, second edition, 2010.
- [BS15a] Matthew Blair and Christopher Sogge. Refined and microlocal Kakeya–Nikodym bounds for eigenfunctions in two dimensions. Analysis & PDE, 8(3):747–764, 2015.
- [BS15b] Matthew D. Blair and Christopher D. Sogge. On Kakeya-Nikodym averages, L<sup>p</sup>-norms and lower bounds for nodal sets of eigenfunctions in higher dimensions. J. Eur. Math. Soc. (JEMS), 17(10):2513–2543, 2015.
- [BS17] Matthew D Blair and Christopher D Sogge. Refined and microlocal Kakeya–Nikodym bounds of eigenfunctions in higher dimensions. *Communications in Mathematical Physics*, 356(2):501–533, 2017.
- [BS18] Matthew D. Blair and Christopher D. Sogge. Concerning Toponogov's theorem and logarithmic improvement of estimates of eigenfunctions. J. Differential Geom., 109(2):189–221, 2018.
- [BS19] Matthew D. Blair and Christopher D. Sogge. Logarithmic improvements in  $L^p$  bounds for eigenfunctions at the critical exponent in the presence of nonpositive curvature. *Invent. Math.*, 217(2):703–748, 2019.
- [CG19a] Yaiza Canzani and Jeffrey Galkowski. Improvements for eigenfunction averages: an application of geodesic beams. arXiv:1809.06296, 2019.
- [CG19b] Yaiza Canzani and Jeffrey Galkowski. On the growth of eigenfunction averages: microlocalization and geometry. Duke Math. J., 168(16):2991–3055, 2019.
- [CG20] Yaiza Canzani and Jeffrey Galkowski. Eigenfunction concentration via geodesic beams. Journal für die reine und angewandte Mathematik, 1(ahead-of-print), 2020.

- [CGT18] Yaiza Canzani, Jeffrey Galkowski, and John A. Toth. Averages of eigenfunctions over hypersurfaces. Comm. Math. Phys., 360(2):619–637, 2018.
- [Cue20] Jean-Claude Cuenin. From spectral cluster to uniform resolvent estimates on compact manifolds. arXiv:2011.07254, 2020.
- [DSFKS13] David Dos Santos Ferreira, Carlos E. Kenig, and Mikko Salo. Determining an unbounded potential from Cauchy data in admissible geometries. *Comm. Partial Differential Equations*, 38(1):50–68, 2013.
- [DZ16] Semyon Dyatlov and Joshua Zahl. Spectral gaps, additive energy, and a fractal uncertainty principle. Geometric and Functional Analysis, 26(4):1011–1094, 2016.
- [DZ19] Semyon Dyatlov and Maciej Zworski. Mathematical theory of scattering resonances. 200:xi+634, 2019.
- [Gal18] Jeffrey Galkowski. A microlocal approach to eigenfunction concentration. Journées équations aux dérivées partielles, 2018. talk:3.
- [Gal19] Jeffrey Galkowski. Defect measures of eigenfunctions with maximal  $L^{\infty}$  growth. Ann. Inst. Fourier (Grenoble), 69(4):1757–1798, 2019.
- [GT17] Jeffrey Galkowski and John A Toth. Eigenfunction scarring and improvements in  $L^{\infty}$  bounds. Analysis & PDE, 11(3):801–812, 2017.
- [GT20] Jeffrey Galkowski and John A. Toth. Pointwise bounds for joint eigenfunctions of quantum completely integrable systems. *Comm. Math. Phys.*, 375(2):915–947, 2020.
- [HR16] Hamid Hezari and Gabriel Rivière.  $L^p$  norms, nodal sets, and quantum ergodicity. Adv. Math., 290:938–966, 2016.
- [HT15] Andrew Hassell and Melissa Tacy. Improvement of eigenfunction estimates on manifolds of nonpositive curvature. *Forum Math.*, 27(3):1435–1451, 2015.
- [KTZ07] Herbert Koch, Daniel Tataru, and Maciej Zworski. Semiclassical L<sup>p</sup> estimates. Ann. Henri Poincaré, 8(5):885–916, 2007.
- [Lee13] John M. Lee. Introduction to smooth manifolds, volume 218 of Graduate Texts in Mathematics. Springer, New York, second edition, 2013.
- [Sog88] Christopher D. Sogge. Concerning the  $L^p$  norm of spectral clusters for second-order elliptic operators on compact manifolds. J. Funct. Anal., 77(1):123–138, 1988.
- [STZ11] Christopher D. Sogge, John A. Toth, and Steve Zelditch. About the blowup of quasimodes on Riemannian manifolds. J. Geom. Anal., 21(1):150–173, 2011.
- [SZ99] Johannes Sjöstrand and Maciej Zworski. Asymptotic distribution of resonances for convex obstacles. Acta Math., 183(2):191–253, 1999.
- [SZ02] Christopher D. Sogge and Steve Zelditch. Riemannian manifolds with maximal eigenfunction growth. Duke Math. J., 114(3):387–437, 2002.
- [SZ16] Christopher D. Sogge and Steve Zelditch. A note on  $L^p$ -norms of quasi-modes. 34:385–397, 2016.
- [Tac18] Melissa Tacy. A note on constructing families of sharp examples for  $L^p$  growth of eigenfunctions and quasimodes. *Proc. Amer. Math. Soc.*, 146(7):2909–2924, 2018.
- $\begin{array}{ll} [{\rm Tac19}] & {\rm Melissa\ Tacy.}\ L^p\ {\rm estimates\ for\ joint\ quasimodes\ of\ semiclassical\ pseudodifferential\ operators.}\ Israel\ J.\\ Math.,\ 232(1):401-425,\ 2019. \end{array}$
- [TZ02] John A. Toth and Steve Zelditch. Riemannian manifolds with uniformly bounded eigenfunctions. *Duke Math. J.*, 111(1):97–132, 2002.
- [TZ03] John A. Toth and Steve Zelditch.  $L^p$  norms of eigenfunctions in the completely integrable case. Ann. Henri Poincaré, 4(2):343–368, 2003.
- [Zwo12] Maciej Zworski. Semiclassical analysis, volume 138 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL, CHAPEL HILL, NC, USA *Email address*: canzani@email.unc.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY COLLEGE LONDON, LONDON, UNITED KINGDOM *Email address*: j.galkowski@ucl.ac.uk