

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Guth, Larry and Hickman, Jonathan and Iliopoulou, Marina (2019) Sharp estimates for oscillatory integral operators via polynomial partitioning. *Acta Mathematica* . (In press)

### DOI

### Link to record in KAR

<https://kar.kent.ac.uk/75575/>

### Document Version

Pre-print

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

# SHARP ESTIMATES FOR OSCILLATORY INTEGRAL OPERATORS VIA POLYNOMIAL PARTITIONING

LARRY GUTH, JONATHAN HICKMAN, AND MARINA ILIOPOULOU

ABSTRACT. The sharp range of  $L^p$ -estimates for the class of Hörmander-type oscillatory integral operators is established in all dimensions under a positive-definite assumption on the phase. This is achieved by generalising a recent approach of the first author for studying the Fourier extension operator, which utilises polynomial partitioning arguments.

## 1. INTRODUCTION

1.1. **Statement of results.** Let  $B^d$  denote the unit ball in  $\mathbb{R}^d$  and  $\Sigma: B^{n-1} \rightarrow \mathbb{R}^n$  be a smooth<sup>1</sup> parametrisation of a hypersurface. Further let  $a \in C_c^\infty(\mathbb{R}^{n-1})$  be non-negative and supported in  $B^{n-1}$  and suppose  $\Sigma$  has of non-vanishing Gaussian curvature on the support  $\text{supp } a$  of  $a$ . Analytically, this means that  $\Sigma$  satisfies the following conditions:

- E1)  $\text{rank } \partial_\omega \Sigma(\omega) = n - 1$  for all  $\omega \in B^{n-1}$ .
- E2) Defining the Gauss map  $G: B^{n-1} \rightarrow S^{n-1}$  by  $G(\omega) := \frac{G_0(\omega)}{|G_0(\omega)|}$  where

$$G_0(\omega) := \bigwedge_{j=1}^{n-1} \partial_{\omega_j} \Sigma(\omega),$$

the curvature condition

$$\det \partial_{\omega\omega}^2 \langle \Sigma(\omega), G(\omega_0) \rangle |_{\omega=\omega_0} \neq 0$$

holds for all  $\omega_0 \in \text{supp } a$ .

Here the wedge product of  $n - 1$  vectors in  $\mathbb{R}^n$  is identified with a vector in the usual manner.

A central problem in harmonic analysis is to understand the Lebesgue space mapping properties of the extension operator  $E$  associated to such a parametrised hypersurface. This operator is defined by the formula

$$(1.1) \quad Ef(x) := \int_{B^{n-1}} e^{2\pi i \langle x, \Sigma(\omega) \rangle} a(\omega) f(\omega) d\omega$$

for all integrable  $f: B^{n-1} \rightarrow \mathbb{C}$ . Thus,  $E$  is an oscillatory integral operator with associated phase function

$$(1.2) \quad \phi(x; \omega) := \langle x, \Sigma(\omega) \rangle.$$

---

<sup>1</sup>In view of the methods of the present article it is convenient to work in the  $C^\infty$  category, but the forthcoming definitions and questions certainly make sense at lower levels of regularity (in particular, in the  $C^2$  class).

Observe that the parametrisation  $\Sigma$  can be recovered from the phase by differentiation; that is,

$$\partial_x \phi(x; \omega) = \Sigma(\omega).$$

Typically, one is interested in proving local estimates for (1.1) of the form<sup>2</sup>

$$(1.3) \quad \|Ef\|_{L^p(B(0,\lambda))} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})};$$

here the left-hand norm has been localised to a ball of radius  $\lambda \geq 1$  and the right-hand constant is allowed some weak dependence on  $\lambda$ . In particular, the Fourier restriction conjecture asserts that (1.3) should hold for any  $\varepsilon > 0$  in the range  $p \geq 2n/(n-1)$ .

In this article the natural variable coefficient generalisations of such extension operators (1.1) and estimates (1.3) are studied. In particular, here more general oscillatory integral operators are considered whose associated phase function  $\phi(x; \omega)$  shares the property of the extension operator that for each  $x$  the map  $\omega \mapsto \partial_x \phi(x; \omega)$  parametrises a hypersurface of non-vanishing Gaussian curvature. Crucially, however, the choice of hypersurface is now allowed to smoothly vary with  $x$ .

To formalise this discussion, let  $n \geq 2$ ,  $a \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$  be non-negative and supported in  $B^n \times B^{n-1}$  and  $\phi: B^n \times B^{n-1} \rightarrow \mathbb{R}$  be a smooth function which satisfies the following conditions:

H1)  $\text{rank } \partial_{\omega x}^2 \phi(x; \omega) = n-1$  for all  $(x; \omega) \in B^n \times B^{n-1}$ .

H2) Defining the map  $G: B^n \times B^{n-1} \rightarrow S^{n-1}$  by  $G(x; \omega) := \frac{G_0(x; \omega)}{|G_0(x; \omega)|}$  where

$$G_0(x; \omega) := \bigwedge_{j=1}^{n-1} \partial_{\omega_j} \partial_x \phi(x; \omega),$$

the curvature condition

$$\det \partial_{\omega \omega}^2 \langle \partial_x \phi(x; \omega), G(x; \omega_0) \rangle|_{\omega=\omega_0} \neq 0$$

holds for all  $(x; \omega_0) \in \text{supp } a$ .

Clearly H1) and H2) agree with E1) and E2) when one restricts to phases of the form (1.2), and this definition therefore leads to a generalisation of the operator  $E$  introduced above.

Suppose  $\phi$  satisfies H1) and H2), for any  $\lambda \geq 1$  let  $a^\lambda(x; \omega) := a(x/\lambda; \omega)$ ,  $\phi^\lambda(x; \omega) := \lambda \phi(x/\lambda; \omega)$  and define the operator  $T^\lambda$  by

$$(1.4) \quad T^\lambda f(x) := \int_{B^{n-1}} e^{2\pi i \phi^\lambda(x; \omega)} a^\lambda(x; \omega) f(\omega) d\omega$$

for all integrable  $f: B^{n-1} \rightarrow \mathbb{C}$ . In this case  $T^\lambda$  is said to be a *Hörmander-type operator*. Note that the spatial localisation featured in (1.3) is now built into the operator.

**Theorem 1.1** (Stein [18], Bourgain–Guth [8]). *Suppose  $T^\lambda$  is a Hörmander-type operator. For all  $\varepsilon > 0$  the estimate*

$$(1.5) \quad \|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}$$

<sup>2</sup>Given a (possibly empty) list of objects  $L$ , for real numbers  $A_p, B_p \geq 0$  depending on some Lebesgue exponent  $p$  the notation  $A_p \lesssim_L B_p$  or  $B_p \gtrsim_L A_p$  signifies that  $A_p \leq C B_p$  for some constant  $C = C_{L,n,p} \geq 0$  depending on the objects in the list,  $n$  and  $p$ . In addition,  $A_p \sim_L B_p$  is used to signify that  $A_p \lesssim_L B_p$  and  $A_p \gtrsim_L B_p$ .

holds uniformly for  $\lambda \geq 1$  whenever  $p$  satisfies

$$(1.6) \quad \begin{aligned} p &\geq 2 \cdot \frac{n+1}{n-1} && \text{if } n \text{ is odd;} \\ p &\geq 2 \cdot \frac{n+2}{n} && \text{if } n \text{ is even.} \end{aligned}$$

The odd dimensional case is due to Stein [18], who in fact showed that the above estimates are valid for  $p \geq 2(n+1)/(n-1)$  in all dimensions without the  $\lambda^\varepsilon$ -loss. The strengthened results in even dimensions were established much later by Bourgain and the first author [8].<sup>3</sup> A detailed history of this problem is provided later in the introduction. It is remarked that Theorem 1.1 is sharp in the sense that there are examples of Hörmander-type operators for which (1.5) fails whenever  $p$  does not satisfy (1.6).

The majority of this work concerns the case where the phase satisfies a strengthened version of H2), namely:

H2<sup>+</sup>) For all  $(x; \omega_0) \in \text{supp } a$  the matrix

$$\partial_{\omega\omega}^2 \langle \partial_x \phi(x; \omega), G(x; \omega_0) \rangle|_{\omega=\omega_0}$$

is positive-definite.

If  $\phi$  satisfies H1) and H2<sup>+</sup>), then  $T^\lambda$  is said to be a *Hörmander-type operator with positive-definite phase*. Geometrically, this condition implies that the principal curvatures of the hypersurface parametrised by  $\omega \mapsto \partial_x \phi(x; \omega)$  are everywhere positive. A hypersurface satisfying this condition is said to be *positively-curved*.

Lee [17] observed that for positive-definite phases one may prove estimates beyond the range of Theorem 1.1. The main result of this article provides sharp estimates in this setting.

**Theorem 1.2.** *Suppose  $T^\lambda$  is a Hörmander-type operator with positive-definite phase. For all  $\varepsilon > 0$  the estimate*

$$(1.7) \quad \|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}$$

holds for all  $\lambda \geq 1$  whenever  $p$  satisfies

$$(1.8) \quad \begin{aligned} p &\geq 2 \cdot \frac{3n+1}{3n-3} && \text{if } n \text{ is odd;} \\ p &\geq 2 \cdot \frac{3n+2}{3n-2} && \text{if } n \text{ is even.} \end{aligned}$$

This result settles a recent conjecture of the first author [13] and improves upon the previous best results of Lee [17] and Bourgain and the first author [8]. It is sharp in the sense that there are examples of Hörmander-type operators with positive-definite phase for which (1.7) fails whenever  $p$  does not satisfy (1.8). Furthermore, away from the endpoint values one may apply  $\varepsilon$ -removal techniques to establish (1.7) without the  $\lambda^\varepsilon$ -loss in the constant (see §12, below).

---

<sup>3</sup>Strictly speaking, in [8] weaker  $L^\infty - L^p$  bounds are proven, but the methods can be used to establish the  $L^p - L^p$  strengthening.

**1.2. Historical remarks.** The problem of determining  $L^p$  estimates for Hörmander-type operators has an interesting history. Hörmander [16] conjectured that (1.7) should hold for  $p > 2n/(n-1)$  (without  $\varepsilon$ -loss) under the hypotheses H1) and H2) only and proved that this is indeed the case when  $n = 2$ . This numerology agrees with the Fourier restriction conjecture and also the Bochner–Riesz multiplier problem, both of which would follow as simple consequences of Hörmander’s conjecture.<sup>4</sup> Stein [18] provided further evidence for this conjecture by proving the estimate

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^2(B^{n-1})} \quad \text{for all } p \geq 2(n+1)/(n-1).$$

It was therefore somewhat surprising when Bourgain [5] showed that, in general, Stein’s theorem is sharp. In particular, he demonstrated that for every odd dimension  $n \geq 3$  there exists a Hörmander-type operator for which

$$(1.9) \quad \|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{L^\infty(B^{n-1})}$$

fails to hold uniformly in  $\lambda \geq 1$  whenever  $p < 2(n+1)/(n-1)$ . Aside from disproving Hörmander’s original conjecture, Bourgain’s work hinted at an interesting divergence between the odd and even-dimensional theory. Moreover, it was demonstrated in [7] that in even dimensions the  $L^\infty - L^p$  estimates always hold in a wider range than that of Stein’s theorem. Thus, in general, the even-dimensional case is better behaved than the odd-dimensional case. This was further highlighted by Bourgain and the first author [8] who showed that in even dimensions (1.9) holds for  $p > 2(n+2)/n$ . Furthermore, in [8] and also the work of Wisewell [25], examples were found in even dimensions which show that (1.9) can fail for  $p < 2(n+2)/n$ . Thus, Hörmander’s original conjecture is valid only when  $n = 2$ .

At this point it is useful to describe the nature of the counter-examples of [5], [25] and [8] and provide some explanation for the difference between the odd and even-dimensional cases. Roughly speaking, in the odd-dimensional case  $T^\lambda$  and  $f$  can be chosen so that  $|T^\lambda f|$  is concentrated in the 1-neighbourhood of a low degree algebraic variety  $Z$  of dimension  $(n+1)/2$ . This is the smallest possible dimension for which such concentration is possible. In the even-dimensional case  $(n+1)/2$  is no longer an integer, and it transpires that  $|T^\lambda f|$  can only be concentrated into the 1-neighbourhood of a variety of relatively large dimension  $(n+2)/2$ . These observations are related to Kakeya compression phenomena for sets of space curves (see [25] for a thorough introduction to this topic). They also hint at some underlying algebraic structure in the problem.

So far the discussion has focused on operators satisfying the original H1) and H2) hypotheses of Hörmander. Lee [17] observed that under the positive-definite hypothesis H2<sup>+</sup>) one can establish improvements over the range given by Stein’s theorem in all dimensions. In particular, he showed that for any Hörmander-type operator with positive-definite phase (1.7) holds for  $p \geq 2(n+2)/2$ . This coincides with Theorem 1.2 when  $n = 3$ , but is weaker in higher dimensions. Wisewell [25] produced examples (again relying on Kakeya compression phenomena) to show that this result is sharp when  $n = 3$  (see also [8]).

Comparing the sharp examples under the H2) and H2<sup>+</sup>) hypotheses highlights another important consideration in addition to Kakeya compression phenomena. This feature relates to how the mass of  $|T^\lambda f|$  can be distributed in neighbourhoods

---

<sup>4</sup>The connection with Bochner–Riesz multipliers is made via the classical reduction of Carleson–Sjölin [9, 16].

	$n$ odd	$n$ even
H2)	$2 \cdot \frac{n+1}{n-1}$	$2 \cdot \frac{n+2}{n}$
H2 <sup>+</sup> )	$2 \cdot \frac{3n+1}{3n-3}$	$2 \cdot \frac{3n+2}{3n-2}$

FIGURE 1. Endpoint values for  $p$  for Hörmander-type operators under various hypotheses.

of low degree varieties. It accounts for the improved behaviour demonstrated by operators with positive-definite phase and is described in detail in §2.

Given the results of this article, the sharp range of estimates for this problem are now understood under either the H2) or H2<sup>+</sup>) hypothesis. The corresponding endpoint values for  $p$  are concisely tabulated in Figure 1.

It is remarked that it is possible to prove estimates beyond the range of Theorem 1.2 under additional assumptions on the phase function. For example, the first author [14] has shown that for  $n = 3$  and all  $\varepsilon > 0$  the extension operator  $E$  associated to the paraboloid satisfies

$$\|Ef\|_{L^p(B(0,\lambda))} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^2)}$$

for all  $p \geq 13/4$ . Furthermore, the aforementioned restriction conjecture asserts that the above inequality should be valid in the wider range  $p \geq 3$ .

**1.3. Multilinear estimates.** The proof of Theorem 1.2 follows the strategy introduced by the first author in [13]. The argument relies on establishing (weakened versions of) multilinear estimates for Hörmander-type operators. The multilinear approach was introduced in the late 1990s to study oscillatory integral operators (although it was arguably already implicit in many earlier foundational works in the subject [12, 9]) and has proven an invaluable tool. To describe the  $k$ -linear setup one first requires the notion of transversality.

**Definition 1.3.** Let  $1 \leq k \leq n$  and  $\mathbf{T}^\lambda = (T_1^\lambda, \dots, T_k^\lambda)$  be a  $k$ -tuple of Hörmander-type operators, where  $T_j^\lambda$  has associated phase  $\phi_j^\lambda$ , amplitude  $a_j^\lambda$  and generalised Gauss map  $G_j$  for  $1 \leq j \leq k$ . Then  $\mathbf{T}^\lambda$  is said to be  $\nu$ -transverse for some  $0 < \nu \leq 1$  if

$$\left| \bigwedge_{j=1}^k G_j(x; \omega_j) \right| \geq \nu \quad \text{for all } (x; \omega_j) \in \text{supp } a_j \text{ for } 1 \leq j \leq k.$$

The following conjecture is a natural generalisation of an existent conjecture of Bennett [3] for Fourier extension operators.

**Conjecture 1.4** ( $k$ -linear Hörmander conjecture). *Let  $1 \leq k \leq n$  and suppose  $(T_1^\lambda, \dots, T_k^\lambda)$  is a  $\nu$ -transverse  $k$ -tuple of Hörmander-type operators which all share the same positive-definite phase function  $\phi^\lambda$ . For all  $p \geq \bar{p}(k, n) := 2(n+k)/(n+k-2)$  and  $\varepsilon > 0$  the estimate*

$$(1.10) \quad \left\| \prod_{j=1}^k |T_j^\lambda f_j|^{1/k} \right\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon, \nu, \phi} \lambda^\varepsilon \prod_{j=1}^k \|f_j\|_{L^2(B^{n-1})}^{1/k}$$

holds for all  $\lambda \geq 1$ .

Techniques have been developed by Tao–Vargas–Vega [24] and Bourgain and the first author [8] to convert  $k$ -linear into linear inequalities. There are a number of features of the multilinear theory which suggest that it is more approachable than directly tackling the linear estimates. For instance, here the desired inequalities are  $L^2$ -based, giving greater scope for orthogonality methods.

Some instances of the conjecture are known.

- The  $k = 1$  case corresponds to Stein’s theorem [18] (which holds without the positive-definite hypothesis).
- The  $k = 2$  case was established by Lee [17], who then used the method of Tao–Vargas–Vega [24] to derive estimates for the linear problem. This approach yields Theorem 1.2 in the  $n = 3$  case, but produces strictly weaker results in higher dimensions (see the discussion in §1.2).
- The  $k = n$  case was established by Bennett–Carbery–Tao [4] who also gave partial results at all levels of multilinearity (see also [3] for further discussion of this work). Bourgain and the first author [8] later developed a method to deduce improved linear estimates from these multilinear inequalities.

The precise statement of the Bennett–Carbery–Tao theorem [4] is as follows.

**Theorem 1.5** (Bennett–Carbery–Tao [4]). *Let  $2 \leq k \leq n$  and suppose that  $(T_1^\lambda, \dots, T_k^\lambda)$  is a  $\nu$ -transverse  $k$ -tuple of Hörmander-type operators. For all  $p \geq 2k/(k-1)$  and  $\varepsilon > 0$  the estimate*

$$\left\| \prod_{j=1}^k |T_j^\lambda f_j|^{1/k} \right\|_{L^p(\mathbb{R}^n)} \lesssim_{\varepsilon, (\phi_j)_{j=1}^k} \nu^{-C\varepsilon} \lambda^\varepsilon \prod_{j=1}^k \|f_j\|_{L^2(B^{n-1})}^{1/k}$$

holds for all  $\lambda \geq 1$ .

The positive-definite assumption does not appear in the hypotheses of Theorem 1.5.<sup>5</sup> Combining Theorem 1.5 with the method of [8] leads to the sharp estimates for Hörmander-type operators stated in Theorem 1.1. For completeness, the details of this argument are given in §11.

**1.4.  $k$ -broad estimates.** In [13] it was observed that, in the context of Fourier extension operators, the method of [8] does not require the full power of the  $k$ -linear theory, but rather can take as its input inequalities of a weaker form than (1.10) known as  *$k$ -broad estimates*. By applying polynomial partitioning techniques, the first author [13] was further able to prove the sharp range of  $k$ -broad estimates for the Fourier extension operator associated to the paraboloid. This led to an improvement on the known range of estimates for parabolic restriction in dimensions  $n \geq 4$ . The main goal of this paper is to extend the theory of  $k$ -broad estimates to the more general context of Hörmander-type operators with positive-definite phase.

The  $k$ -broad setup involves the notion of a  $k$ -broad norm, which was introduced in [13]. Decompose  $B^{n-1}$  into finitely-overlapping balls  $\tau$  of radius  $K^{-1}$ , where  $K$  is a large constant. These balls will be frequently referred to as  $K^{-1}$ -caps. Given a function  $f: B^{n-1} \rightarrow \mathbb{C}$  write  $f = \sum_\tau f_\tau$  where  $f_\tau$  is supported in  $\tau$ . In view of the rescaling  $\phi^\lambda$  of the phase function, define the rescaled generalised Gauss map

$$G^\lambda(x; \omega) := G(x/\lambda; \omega) \quad \text{for } (x; \omega) \in \text{supp } a^\lambda.$$

<sup>5</sup>In particular, of the results mentioned above only Lee’s bilinear estimate [17] exploits the positive-definite hypothesis.

For each  $x \in B(0, \lambda)$  there is a range of normal directions associated to the cap  $\tau$  given by

$$G^\lambda(x; \tau) := \{G^\lambda(x; \omega) : \omega \in \tau, (x; \omega) \in \text{supp } a^\lambda\}.$$

If  $V \subseteq \mathbb{R}^n$  is a linear subspace, then let  $\not\prec(G^\lambda(x; \tau), V)$  denote the smallest angle between any non-zero vector  $v \in V$  and  $v' \in G^\lambda(x; \tau)$ .

The spatial ball  $B(0, \lambda)$  is also decomposed into relatively small balls  $B_{K^2}$  of radius  $K^2$ . In particular, fix  $\mathcal{B}_{K^2}$  a collection of finitely-overlapping  $K^2$ -balls which are centred in and cover  $B(0, \lambda)$ . For  $B_{K^2} \in \mathcal{B}_{K^2}$  centred at some point  $\bar{x} \in B(0, \lambda)$  define

$$(1.11) \quad \mu_{T^\lambda f}(B_{K^2}) := \min_{V_1, \dots, V_A \in \text{Gr}(k-1, n)} \left( \max_{\tau: \not\prec(G^\lambda(\bar{x}; \tau), V_a) > K^{-1} \text{ for } 1 \leq a \leq A} \|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p \right);$$

here  $\text{Gr}(k-1, n)$  is the Grassmannian manifold of all  $(k-1)$ -dimensional subspaces in  $\mathbb{R}^n$ . It will often be notationally convenient to write  $\tau \notin V_a$  to mean that  $\not\prec(G^\lambda(\bar{x}; \tau), V_a) > K^{-1}$  (the choice of centre  $\bar{x}$  should always be clear from the context); with this notation the above expression becomes

$$\mu_{T^\lambda f}(B_{K^2}) := \min_{V_1, \dots, V_A \in \text{Gr}(k-1, n)} \left( \max_{\tau: \tau \notin V_a \text{ for } 1 \leq a \leq A} \|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p \right).$$

For  $U \subseteq \mathbb{R}^n$  the  $k$ -broad norm over  $U$  is then defined to be

$$(1.12) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(U)} := \left( \sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap U \neq \emptyset}} \mu_{T^\lambda f}(B_{K^2}) \right)^{1/p}.$$

It is remarked that  $\|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}$  is not a norm in the traditional sense, but it does satisfy weak variants of certain key properties of  $L^p$ -norms, as discussed below in §6.

Theorem 1.2 will be a consequence of certain estimates for  $k$ -broad norms. These estimates are proved under a further technical assumption that the phase is of *reduced form*. The details of this condition are postponed until §4.

**Theorem 1.6.** *For  $2 \leq k \leq n$  and all  $\varepsilon > 0$  there exists a constant  $C_\varepsilon > 1$  and an integer  $A$  such that whenever  $T^\lambda$  is a Hörmander-type operator with positive-definite reduced phase the estimate*

$$(1.13) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(\mathbb{R}^n)} \lesssim_\varepsilon K^{C_\varepsilon} \lambda^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds for all  $\lambda \geq 1$  and  $K \geq 1$  whenever  $p \geq \bar{p}(k, n) := 2(n+k)/(n+k-2)$ .

The range of  $p$  is sharp for this theorem, as can be seen by considering the extension operator associated to the (elliptic) paraboloid (see [13]). As explained in §6 below, the  $k$ -broad estimate (1.13) is weaker than the corresponding  $k$ -linear estimate (1.10), and so Theorem 1.6 can be viewed as a weak substitute for Conjecture 1.4.

Theorem 1.6 implies that for every ball  $B_{K^2}$  one can find a collection of  $(k-1)$ -dimensional subspaces  $V_1, \dots, V_A$  for which

$$\max_{\tau \notin V_a \text{ for } 1 \leq a \leq A} \|T^\lambda f_\tau\|_{L^p(B_{K^2})} \lesssim_{K, \varepsilon} \lambda^\varepsilon \|f\|_{L^2(B^{n-1})}.$$

Thus, the contribution to  $\|T^\lambda f\|_{L^p(B_{K^2})}$  arising from caps  $\tau$  whose directions lie away from all of the  $V_a$  can be effectively controlled. The problem of estimating

$\|T^\lambda f\|_{L^p(B_{K^2})}$  is therefore reduced to studying expressions of the form

$$\left\| \sum_{\tau \in V} T^\lambda f_\tau \right\|_{L^p(B_{K^2})}$$

where the sum is over caps  $\tau$  which make a small angle with some  $(k-1)$ -dimensional subspace  $V$ . This term can then be controlled using a combination of  $\ell^p$ -decoupling and an induction-on-scale argument, leading to the proof of Theorem 1.2. The full details of this argument are given in §11.

**Structure of the article.** The structure of this article is as follows:

- In §2 sharp examples for Theorem 1.1 and Theorem 1.2 are discussed in detail.
- In §3 the key features of the problem are identified in order to motivate the forthcoming analysis.
- In §4 some basic reductions are described which allow one to assume the phase is of a certain *reduced form* in the proof of Theorem 1.2.
- In §5 and §6 some basic analytic tools are introduced. In particular, the wave packet decomposition for Hörmander-type operators is defined and studied, some elementary aspects of the  $L^2$ -theory for Hörmander-type operators are reviewed, and there is also a discussion of the basic properties of the  $k$ -broad norms and their relation to  $k$ -linear estimates.
- In §7 certain algebraic tools from combinatorial geometry are introduced. In particular, polynomial partitioning techniques are reviewed and some important geometric lemmas are proved; these techniques will play a fundamental rôle in the proof of Theorem 1.6.
- In §8 and §9 *transverse equidistribution estimates* for  $T^\lambda$  are introduced and studied. These estimates rely heavily on the positive-definite hypothesis and partially account for the improved behaviour exhibited by operators satisfying the  $H2^+$  hypothesis.
- In §10 the proof of  $k$ -broad estimates of Theorem 1.6 is given.
- In §11 the linear estimates of Theorem 1.2 are deduced as a consequence of the  $k$ -broad estimates of Theorem 1.6. For completeness, the same methods are also applied to deduce Theorem 1.1 as a consequence of Corollary 6.5.
- In §12 standard  $\varepsilon$ -removal lemmas are generalised to the variable coefficient setting. This allows one to strengthen Theorem 1.2 away from the endpoint by removing the  $\lambda^\varepsilon$ -dependence in the constant.

**Acknowledgment.** The second author would like to thank David Beltran for engaging discussions on topics related to this article. The first author is supported by a Simons Investigator Award. This material is based upon work supported by the National Science Foundation under Grand No. DMS-1440140 while the second and third authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Spring 2017 semester.

## 2. NECESSARY CONDITIONS

**2.1. An overview of the sharp examples.** In this section examples of Hörmander-type operators are studied in view of establishing the necessity of the conditions on the  $p$  exponent in the linear estimates of Theorem 1.1 and Theorem 1.2. This

	$n$ odd	$n$ even
H2)	$\left(\frac{n+1}{2}, 0\right)$	$\left(\frac{n}{2} + 1, 0\right)$
H2 <sup>+</sup> )	$\left(\frac{n+1}{2}, \frac{1}{2}\right)$	$\left(\frac{n}{2} + 1, \frac{1}{2}\right)$

FIGURE 2. Optimal values of  $(m, \sigma)$  for the sharp examples.

analysis will also identify some key features of operators with positive-definite phase which will later be exploited in the proof of Theorem 1.2.

All the examples considered here are of the following general form: for a fixed operator  $T^\lambda$ , a function  $f$  is chosen so that  $|f|$  is constant whilst  $|T^\lambda f|$  is concentrated in  $N_{\lambda^\sigma}(Z) \cap B(0, \lambda)$  for some low-degree algebraic variety  $Z$  with  $\dim Z = m$ ; here  $N_{\lambda^\sigma}(Z)$  is the  $\lambda^\sigma$ -neighbourhood of  $Z$ . In particular, one has

$$(2.1) \quad \|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \|T^\lambda f\|_{L^2(N_{\lambda^\sigma}(Z) \cap B(0, \lambda))}.$$

The examples will further be chosen so that

$$(2.2) \quad \|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \lambda^{1/2} \|f\|_{L^2(B^{n-1})};$$

note that, by Hörmander's generalisation of the Hausdorff–Young inequality [16], the inequality  $\|T^\lambda f\|_{L^2(\mathbb{R}^n)} \lesssim \lambda^{1/2} \|f\|_{L^2(B^{n-1})}$  always holds (see also §5, below).

Playing (2.1) and (2.2) off against one another yields the necessary conditions on  $p$ . Indeed, for  $f$  as above

$$\|f\|_{L^p(B^{n-1})} \sim \|f\|_{L^2(B^{n-1})} \sim \lambda^{-1/2} \|T^\lambda f\|_{L^2(\mathbb{R}^n)} \sim \lambda^{-1/2} \|T^\lambda f\|_{L^2(N_{\lambda^\sigma}(Z) \cap B(0, \lambda))}.$$

Now, assuming the estimate  $\|T^\lambda g\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|g\|_{L^p(B^{n-1})}$  holds for all  $\varepsilon > 0$  and applying Hölder's inequality, it follows that

$$\|f\|_{L^p(B^{n-1})} \lesssim_\varepsilon |N_{\lambda^\sigma}(Z) \cap B(0, \lambda)|^{1/2-1/p} \lambda^{-1/2} \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}.$$

By a simple computation one concludes that

$$(2.3) \quad p \geq 2 \cdot \frac{\sigma(n-m) + m}{\sigma(n-m) + m - 1}.$$

This condition depends on the two parameters  $m$  and  $\sigma$ , which one wishes to minimise to obtain the strongest possible restriction on  $p$ .

- The optimal choice of  $m$  is  $n - \lfloor \frac{n-1}{2} \rfloor$ . This value arises directly from the theory of Kakeya sets of curves, and will be discussed in more detail in the next subsection.
- The optimal choice of  $\sigma$  depends on the signature of the phase. For general Hörmander-type operators, one may find examples for which  $\sigma = 0$ . If one assumes the positive-definite condition H2<sup>+</sup>), then  $\sigma = 1/2$  is the lowest possible value. This difference in behaviour is governed by *transverse equidistribution estimates* for  $T^\lambda$ , which were introduced in the context of Fourier extension operators in [13]. This will be discussed in detail in §2.4.

The optimal pairs  $(m, \sigma)$  under the various hypotheses are tabulated in Figure 2. Plugging these values into (2.3) gives the corresponding sharp range of estimates for  $T^\lambda$ .

**2.2. Model operators.** The examples described above will arise from operators with phase of the relatively simple form

$$(2.4) \quad \phi(x; \omega) := \langle x', \omega \rangle + \frac{1}{2} \langle \mathbf{A}(x_n) \omega, \omega \rangle$$

where  $\mathbf{A}: \mathbb{R} \rightarrow \text{Sym}(n-1, \mathbb{R})$  is a polynomial function taking values in the class of real symmetric matrices which satisfies  $\mathbf{A}(0) = 0$ . For such a phase the condition H1) always holds whilst H2) (respectively, H2<sup>+</sup>) holds if and only if the component-wise derivative  $\mathbf{A}'(x_n) \in \text{GL}(n-1, \mathbb{R})$  (respectively,  $\mathbf{A}'(x_n)$  is positive-definite) for all relevant  $x_n \in [-1, 1]$ . Observe that if  $\mathbf{A}(x_n) := x_n A$  for some fixed  $A \in \text{Sym}(n-1, \mathbb{R}) \cap \text{GL}(n-1, \mathbb{R})$ , then the resulting operator is the extension operator associated to the graph of the non-degenerate quadratic form  $\omega \mapsto \frac{1}{2} \langle A \omega, \omega \rangle$ . For the present purpose, one is interested in examples with higher-order dependence on  $x_n$ .

Let  $T^\lambda$  be an operator associated to the phase function (2.4) for some  $\mathbf{A}$ . Cover  $B^{n-1}$  by finitely-overlapping balls  $\theta$  of radius  $\lambda^{-1/2}$ ; these balls will frequently be referred to as  $\lambda^{-1/2}$ -caps. Let  $\psi_\theta$  be a smooth partition of unity adapted to this cover. Consider a *wave packet* of the form

$$f_{\theta, v_\theta}(\omega) := e^{-2\pi i \lambda \langle v_\theta, \omega - \omega_\theta \rangle} \psi_\theta(\omega)$$

for some choice of  $v_\theta \in \mathbb{R}^{n-1}$  and  $\omega_\theta$  the centre of the cap  $\theta$ . To obtain the necessary conditions for  $L^p$ -boundedness of  $T^\lambda$ , the operator will be tested against functions given by superpositions of these basic wave packets.

A simple stationary phase analysis shows that  $T^\lambda f_{\theta, v_\theta}$  is concentrated on a tubular region in  $\mathbb{R}^n$ . In particular, define the curve

$$(2.5) \quad \gamma_{\theta, v_\theta}(t) := v_\theta - \mathbf{A}(t) \omega_\theta \quad \text{for } t \in (-1, 1)$$

and let  $T_{\theta, v_\theta}$  be the curved tube

$$T_{\theta, v_\theta} := \{x \in B(0, \lambda) : |x' - \lambda \gamma_{\theta, v_\theta}(x_n/\lambda)| < c \lambda^{1/2}\}$$

for some choice of constant  $c > 0$ . It is not difficult to show that

$$|T^\lambda f_{\theta, v_\theta}(x)| \gtrsim \lambda^{-(n-1)/2} \chi_{T_{\theta, v_\theta}}(x) \quad \text{for all } x \in B(0, \lambda),$$

provided that  $c$  is chosen suitably small.

**2.3. Kakeya sets of curves.** By studying the geometry of the family of tubes  $T_{\theta, v_\theta}$  one may construct sharp examples for Theorem 1.2. These examples arise owing to Kakeya compression phenomena, whereby the tubes are arranged to lie in a neighbourhood of a low-dimensional set. For  $n = 3$ , the following example appears in Bourgain–Guth [8]. Let  $\phi$  be of the form (2.4) where  $\mathbf{A}(t)$  is taken to be the  $(n-1) \times (n-1)$  block-diagonal matrix

$$\mathbf{A}(t) := \underbrace{\begin{pmatrix} t & t^2 \\ t^2 & t+t^3 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} t & t^2 \\ t^2 & t+t^3 \end{pmatrix}}_{\lfloor \frac{n-1}{2} \rfloor\text{-fold}} \oplus (t).$$

Here it is understood that the final  $1 \times 1$  block appears only when  $n$  is even. Observe that the resulting phase (2.4) satisfies H1) and H2<sup>+</sup>) on  $B^n \times B^{n-1}$ .

Suppose that  $T^\lambda$  is the associated oscillatory integral operator. The estimate

$$(2.6) \quad \|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_\varepsilon \lambda^\varepsilon \|f\|_{L^p(B^{n-1})}$$

is tested against a superposition of wave packets

$$f := \sum_{\theta: \lambda^{-1/2}\text{-cap}} \epsilon_\theta \cdot f_{\theta, v_\theta}$$

where the  $\epsilon_\theta \in \{1, -1\}$  are uniformly distributed independent random signs. By Khinchine's theorem (see, for instance, [19, Appendix D]), the expected value of  $|T^\lambda f(x)|$  is given by

$$\mathbf{E}[|T^\lambda f(x)|] \sim \left( \sum_{\theta: \lambda^{-1/2}\text{-cap}} |T^\lambda f_{\theta, v_\theta}(x)|^2 \right)^{1/2} \gtrsim \lambda^{-(n-1)/2} \left( \sum_{\theta: \lambda^{-1/2}\text{-cap}} \chi_{T_{\theta, v_\theta}}(x) \right)^{1/2}$$

for all  $x \in B(0, \lambda)$ . Thus, by Hölder's and Minkowski's inequalities,

$$\lambda^{-(n-1)/2} \left( \int \sum_{\theta: \lambda^{-1/2}\text{-cap}} \chi_{T_{\theta, v_\theta}} \right)^{1/2} \lesssim \left| \bigcup_{\theta: \lambda^{-1/2}\text{-cap}} T_{\theta, v_\theta} \right|^{1/2-1/p} \mathbf{E}[\|T^\lambda f\|_{L^p(\mathbb{R}^n)}].$$

The hypothesis (2.6) together with a direct computation now gives

$$(2.7) \quad \|f\|_{L^p(B^{n-1})} \lesssim \epsilon \left| \bigcup_{\theta: \lambda^{-1/2}\text{-cap}} T_{\theta, v_\theta} \right|^{1/2-1/p} \lambda^{-1/2+\epsilon} \|f\|_{L^p(B^{n-1})},$$

since  $\|f\|_{L^p(B^{n-1})} \sim 1$  is independent of the outcomes of the  $\epsilon_\theta$ .

Varying  $v_\theta$  corresponds to translating the tube  $T_{\theta, v_\theta}$  in space. In view of (2.7), one wishes to choose the  $v_\theta$  in order to arrange the tubes so that their union has small measure. For the above choice of phase it is in fact possible to select the  $v_\theta$  so that the tubes all lie in the  $\lambda^{1/2}$ -neighbourhood of a low-dimensional, low degree algebraic variety. In particular, let  $m := n - \lfloor \frac{n-1}{2} \rfloor$  and  $Z := Z(P_1, \dots, P_{n-m})$  be the common zero set of the polynomials

$$P_j(x) := \lambda x_{2j} - x_{2j-1} x_n \quad \text{for } 1 \leq j \leq \lfloor \frac{n-1}{2} \rfloor.$$

Thus,  $Z$  is an algebraic variety of dimension  $m$  and degree 2. If one defines

$$v_{\theta, 2j-1} := -\omega_{\theta, 2j} \quad \text{and} \quad v_{\theta, 2j} = v_{\theta, n-1} = 0 \quad \text{for } 1 \leq j \leq \lfloor \frac{n-1}{2} \rfloor$$

for each cap  $\theta$ , then a simple computation shows that the curve  $t \mapsto (\lambda \gamma_{\theta, v_\theta}(t/\lambda), t)$  lies in  $Z$ . Thus,

$$\bigcup_{\theta: \lambda^{-1/2}\text{-cap}} T_{\theta, v_\theta} \subseteq N_{\lambda^{1/2}}(Z) \cap B(0, \lambda)$$

and the desired necessary conditions on  $p$  follow from (2.7).

In conclusion, here the necessary conditions arise due the fact that it is possible to compress an  $(n-1)$ -dimensional family of curves into a set of small dimension  $m$ . The value  $m = n - \lfloor \frac{n-1}{2} \rfloor$  is optimal for this kind of behaviour, in view of known estimates for associated Kakeya maximal functions: see [25] and [8].

**2.4. Mass concentration.** It will be useful to contrast the behaviour in the positive-definite and indefinite cases by considering sharp examples for Theorem 1.1 (that is, for the class of operators satisfying H1) and the weaker hypothesis H2)). As before, Kakeya compression plays a significant rôle in the argument, but one can introduce additional interference between the wave packets which leads to stronger necessary conditions. In particular, this interference causes the mass of  $|T^\lambda f|$  to concentrate in a tiny  $O(1)$ -neighbourhood of a variety  $Z$ ; such tight concentration is not possible under the H2<sup>+</sup>) hypothesis (as demonstrated by Theorem 1.2).

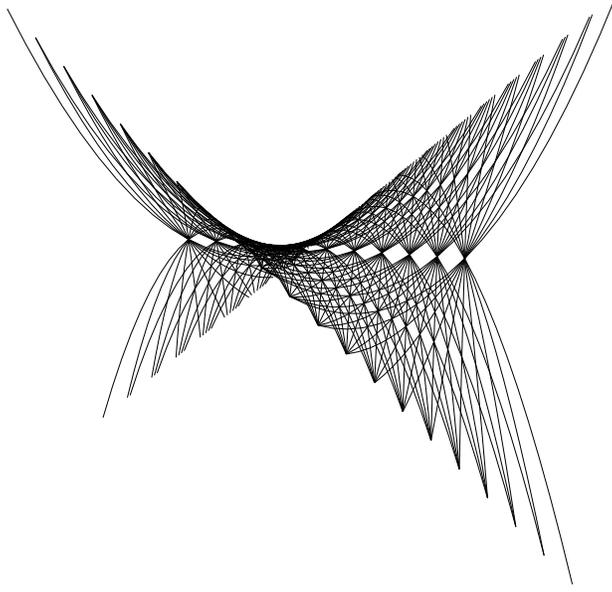


FIGURE 3. The Kakeya compression phenomenon for the curves arising from the matrix (2.8).

The following example was introduced by Bourgain [5] (see also [7]). Once again, the phase is taken to be of the form (2.4). This time  $\mathbf{A}(t)$  is defined to be the  $(n-1) \times (n-1)$  block-diagonal matrix

$$(2.8) \quad \mathbf{A}(t) := \underbrace{\begin{pmatrix} 0 & t \\ t & t^2 \end{pmatrix} \oplus \cdots \oplus \begin{pmatrix} 0 & t \\ t & t^2 \end{pmatrix}}_{\lfloor \frac{n-1}{2} \rfloor\text{-fold}} \oplus (t).$$

Clearly, the corresponding phase satisfies H1) and H2), but H2<sup>+</sup>) fails. Define the curves  $\gamma_{\theta, v_\theta}$  as in (2.5) so that

$$(2.9) \quad |T^\lambda f_{\theta, v_\theta}(x)| \gtrsim \lambda^{-(n-1)/2} \chi_{T_{\theta, v_\theta}}(x) \quad \text{for all } x \in B(0, \lambda).$$

If one takes

$$v_{\theta, 2j-1} := -\omega_{\theta, 2j-1} \quad \text{and} \quad v_{\theta, 2j} = v_{\theta, n-1} = 0 \quad \text{for } 1 \leq j \leq \lfloor \frac{n-1}{2} \rfloor,$$

then it follows that the curve  $t \mapsto (\lambda \gamma_{\theta, v_\theta}(t/\lambda), t)$  lies in  $Z$  for all  $\lambda^{-1/2}$ -caps  $\theta$ , where  $Z$  is the same variety as that appearing in the previous subsection (see Figure 3).

One may repeat the analysis of §2.3 by taking  $f$  to be a linear combination of wave packets  $f_{\theta, v_\theta}$  with random signs. This leads to the same necessary conditions as in the positive-definite case. However, certain deterministic choices of  $f$  lead to stronger conditions on  $p$ . In particular, consider the function

$$(2.10) \quad \tilde{f} := \sum_{\theta: \lambda^{-1/2}\text{-cap}} e^{2\pi i \lambda Q(\omega_\theta)} f_{\theta, v_\theta},$$

where the  $v_\theta$  are as defined above and  $Q$  is the quadratic polynomial

$$Q(\omega) := \frac{1}{2} \sum_{j=1}^{\lfloor \frac{n-1}{2} \rfloor} \omega_{2j-1}^2.$$

Each modulated wave packet appearing in (2.10) has a phase given by

$$\lambda(Q(\omega_\theta) - \langle v_\theta, \omega - \omega_\theta \rangle) = \lambda Q(\omega) - \frac{\lambda}{2} \sum_{j=1}^{\lfloor \frac{n-1}{2} \rfloor} (\omega_{2j-1} - \omega_{\theta, 2j-1})^2.$$

Since the  $\lambda(\omega_{2j-1} - \omega_{\theta, 2j-1})^2$  terms are bounded functions on the support of  $\psi_\theta$ , they do not contribute any significant oscillation. One may therefore heuristically identify  $\tilde{f}$  with the function

$$f(\omega) := e^{2\pi i \lambda Q(\omega)} \psi(\omega)$$

where  $\psi$  is a bump function supported in  $B^{n-1}$ . Using a simple stationary phase argument, it was shown in [5] (see also [7, 8]) that

$$(2.11) \quad |T^\lambda f(x)| \gtrsim \lambda^{-\lfloor n/2 \rfloor / 2} \chi_{N_c(Z)}(x) \quad \text{for all } x \in B(0, \lambda).$$

Here  $0 < c < 1$  is some small, fixed constant (which is independent of  $\lambda$ ) and  $Z$  is as defined in §2.3. With this estimate, one readily deduces the desired necessary conditions on  $p$ .

Comparing (2.9) and (2.11) (in view of the heuristic identification between  $\tilde{f}$  and  $f$ ) suggests that there is significant constructive interference between the various  $T^\lambda f_{\theta, v_\theta}$  in a region very close to  $Z$ . This forces the mass of  $|T^\lambda f(x)|$  to concentrate around a tiny neighbourhood of the variety and results in stronger necessary conditions for the  $L^p$ -boundedness of  $T^\lambda$ .

For completeness we briefly detail the argument used in [5, 7] to prove (2.11). Observe that  $T^\lambda f(x)$  is an oscillatory integral with smooth amplitude  $\psi$  and phase

$$(2.12) \quad \sum_{j=1}^{\lfloor \frac{n-1}{2} \rfloor} x_{2j-1} \omega_{2j-1} + x_{2j} \omega_{2j} + \frac{\lambda}{2} (\omega_{2j-1} + \lambda^{-1} x_n \omega_{2j})^2 + \delta_e (x_{n-1} \omega_{n-1} + \frac{x_n}{2} \omega_{n-1}^2)$$

where  $\delta_e = 0$  if  $n$  is odd and  $\delta_e = 1$  if  $n$  is even. Introduce new variables  $z_j := \omega_{2j-1} + \lambda^{-1} x_n \omega_{2j}$  for  $1 \leq j \leq \lfloor \frac{n-1}{2} \rfloor$ . If  $x \in Z$ , then the phase function (2.12) can be expressed as

$$\sum_{j=1}^{\lfloor \frac{n-1}{2} \rfloor} x_{2j-1} z_j + \frac{\lambda}{2} z_j^2 + \delta_e (x_{n-1} \omega_{n-1} + \frac{x_n}{2} \omega_{n-1}^2).$$

The integral  $T^\lambda f$  can now be reduced to a product of  $\lfloor \frac{n-1}{2} \rfloor + \delta_e = \lfloor \frac{n}{2} \rfloor$  integrals, each in a single variable. For  $x \in Z$  the inequality (2.11) follows as a consequence of standard stationary phase estimates applied to each of these integrals (see, for instance, [20, Chapter VIII, Proposition 3]). This lower bound can then be extended to some  $c$ -neighbourhood of  $Z$  via a simple estimate on the gradient of  $T^\lambda f(x)$ .

### 3. KEY FEATURES OF THE ANALYSIS

The examples of the previous section highlight several key features of Hörmander-type operators. All these features are exploited in the proofs of the linear and  $k$ -broad estimates.

- 1) **Algebraic structure.** The sharp examples were given by arranging collections of wave packets to lie in a relatively small neighbourhood of a low degree, low dimensional algebraic variety  $Z$ . It turns out that this is an essential feature of both the linear and  $k$ -broad problems. To exploit this underlying algebraic structure, the proof of Theorem 1.6 will rely on a variant of the polynomial partitioning method introduced by Katz and the first author [15]. Roughly speaking, this method allows one to reduce to the case where  $|T^\lambda f|$  concentrates around some low degree, low dimensional variety, as in the sharp examples. This can be thought of as a dimensional reduction and, indeed, the proof of Theorem 1.6 will proceed by an induction on dimension. Polynomial partitioning has played an increasingly important rôle in the theory of oscillatory integral operators, beginning with the work on the restriction problem in [14, 13] and, more recently, in [11].
- 2) **Non-concentration/transverse equidistribution.** Suppose one does not assume the phase is positive-definite. The example of §2.4 then shows that interference between the wave packets can cause  $|T^\lambda f|$  to be concentrated in a tiny neighbourhood of  $Z$ . In order to prove the sharp range of estimates in the positive-definite case one must rule out the possibility of such concentration. This is achieved by extending the theory of so-called *transverse equidistribution estimates* introduced in [13] to the variable coefficient setting. These estimates can be interpreted as showing that  $|T^\lambda f(x)|$  is morally constant along transverse directions to  $Z$  in a  $\lambda^{1/2}$ -neighbourhood of the variety. Consequently,  $|T^\lambda f(x)|$  cannot concentrate in a smaller neighbourhood.
- 3) **Parity of the dimension.** Another key feature of the examples discussed in the previous section is their dependence on the parity of the ambient dimension  $n$ . Recall that this is directly related to the minimal dimension

$$d(n) := \begin{cases} \frac{n+1}{2} & \text{if } n \text{ is odd} \\ \frac{n+2}{2} & \text{if } n \text{ is even} \end{cases}$$

of Kakeya sets of curves in  $\mathbb{R}^n$ . The parity of the dimension does not play an overt rôle in the proof of the  $k$ -broad estimates, but it becomes a noticeable feature when one wishes to pass from  $k$ -broad to linear estimates in the proof of Theorem 1.2. In particular, for each fixed value of  $2 \leq k \leq n$ , the method of §11 shows that the  $k$ -broad estimates imply a (possibly empty/trivial) range of linear estimates. It transpires that to optimise the range of linear estimates obtained in this manner one should choose  $k$  to correspond to the dimension  $d(n)$  from the Kakeya problem.

The proof of the  $k$ -broad estimates follows the same general scheme as that used to study Fourier extension operators in [13], and heavily exploits the features 1) and 2) of the problem highlighted above. A detailed sketch of the argument in the extension context is provided in [13]; this sketch is likely to be beneficial to readers new to these ideas.

## 4. REDUCTIONS

**4.1. Basic reductions.** The prototypical example of a positive-definite phase function is given by

$$\phi_{\text{par}}(x; \omega) := \langle x', \omega \rangle + x_n \frac{|\omega|^2}{2}.$$

This is the phase associated to the extension operator for the (elliptic) paraboloid. In general, to prove Theorem 1.2 it suffices to only consider phases which are given by small perturbations of  $\phi_{\text{par}}$ . To see why this is so, first recall that the class of operators under consideration are those of the form

$$(4.1) \quad T^\lambda f(x) := \int_{\mathbb{R}^{n-1}} e^{2\pi i \phi^\lambda(x; \omega)} a^\lambda(x; \omega) f(\omega) d\omega$$

where  $\phi$  satisfies H1) and H2<sup>+</sup>). In addition, one may assume a number of fairly stringent conditions on the form of  $\phi$  on the support of  $a$ .

**Lemma 4.1.** *To prove Theorem 1.2 it suffices to consider the case where  $a$  is supported on  $X \times \Omega$  where  $X := X' \times X_n$  and  $X' \subset B^{n-1}$ ,  $X_n \subset B^1$  and  $\Omega \subset B^{n-1}$  are small balls centred at 0 upon which the phase  $\phi$  has the form*

$$(4.2) \quad \phi(x; \omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x; \omega).$$

Here  $h$  and  $\mathcal{E}$  are smooth functions,  $h$  is quadratic in  $\omega$  and  $\mathcal{E}$  is quadratic in  $x$  and  $\omega$ .<sup>6</sup> Furthermore, letting  $c_{\text{par}} > 0$  be a small dimensional constant, one may assume

$$(4.3) \quad \|\partial_{x'\omega}^2 \phi(x; \omega) - \mathbf{I}_{n-1}\|_{\text{op}} < c_{\text{par}} \quad \text{and} \quad \|\partial_{\omega\omega}^2 \partial_{x_k} \phi(x; \omega) - \delta_{kn} \mathbf{I}_{n-1}\|_{\text{op}} < c_{\text{par}}$$

for all  $(x; \omega) \in X \times \Omega$  and  $1 \leq k \leq n$ .

Here  $\mathbf{I}_{n-1}$  denotes the  $(n-1) \times (n-1)$  identity matrix,  $\delta_{ij}$  the Kronecker  $\delta$ -function and  $\|\cdot\|_{\text{op}}$  the operator norm.

The proof of the lemma is based upon three elementary principles.

*Localisation.* If a property  $P$  of a phase holds locally on  $\text{supp } a$ , then typically one may assume  $P$  holds on the whole of  $\text{supp } a$  by applying a partition of unity, the triangle inequality and shifting co-ordinates.

*Parametrisation invariance.* By the change of variables formula, one may compose  $\phi$  with a smooth change of either the  $x$  or  $\omega$  variables.

*Modulation invariance.* One is free to add smooth functions to the phase which depend only on either the  $x$  or on the  $\omega$  variables. In particular,  $\phi$  can be replaced by

$$\phi(x; \omega) + \phi(0; 0) - \phi(0; \omega) - \phi(x; 0)$$

and therefore one may assume that

$$(4.4) \quad \partial_x^\alpha \phi(x; 0) = 0 \quad \text{and} \quad \partial_\omega^\beta \phi(0; \omega) = 0$$

for all multi-indices  $(\alpha, \beta) \in \mathbb{N}_0^n \times \mathbb{N}_0^{n-1}$ .

The following argument provides an example of these three principles working together. Rotating the  $x$ -coordinates, one may assume that  $\partial_\omega \partial_{x_n} \phi(0; 0) = 0$ . By H2) it follows that

$$\det \partial_{x'\omega}^2 \phi(0; 0) \neq 0.$$

The inverse function theorem now implies the existence of local inverses to the functions  $\omega \mapsto \partial_{x'} \phi(x; \omega)$  and  $x' \mapsto \partial_\omega \phi(x; \omega)$  in a neighbourhood of 0. Thus, by localisation, one may assume  $\text{supp } a$  is contained in  $X \times \Omega$  where  $X = X' \times X_n$  for

<sup>6</sup>Explicitly, if  $(\alpha, \beta) \in \mathbb{N}_0 \times \mathbb{N}_0^{n-1}$  is a pair of multi-indices, then:

- i)  $\partial_\omega^\beta h(0) = \partial_\omega^\beta \partial_x^\alpha \mathcal{E}(x; 0) = 0$  whenever  $x \in X$  and  $|\beta| \leq 1$ ;
- ii)  $\partial_\omega^\beta \partial_x^\alpha \mathcal{E}(0; \omega) = 0$  whenever  $\omega \in \Omega$  and  $|\alpha| \leq 1$ .

$X' \subset B^{n-1}$ ,  $X_n \subset B^1$  and  $\Omega \subset B^{n-1}$  small balls centred at 0 and that there exist smooth functions  $\Phi$  and  $\Psi$  taking values in  $X$  and  $\Omega$ , respectively, such that

$$(4.5) \quad \partial_{x'}\phi(x; \Psi(x; u)) = u \quad \text{and} \quad \partial_\omega\phi(\Phi(z', x_n; \omega), x_n; \omega) = z'.$$

The former identity can be thought of as a generalisation of the fact that any hypersurface can be locally parametrised as a graph. The latter identity features in the proof of Lemma 4.1 and it is useful to highlight some further properties of  $\Phi$ . For each  $(x_n, \omega) \in X_n \times \Omega$  the map  $z' \mapsto \Phi(z', x_n; \omega)$  is a diffeomorphism from its domain onto  $X'$ ; this provides a useful change of variables on  $X'$ . Furthermore, it is easy to see that 0 lies in the domain of this map when  $x_n = 0$ ,  $\omega = 0$  and that

$$(4.6) \quad \Phi(0; 0) = 0, \quad \partial_{x_n}\Phi(0; 0) = 0 \quad \text{and} \quad \partial_{x'}\Phi(0; 0) = \partial_{x'\omega}\phi(0; 0)^{-1}.$$

Indeed, the first identity follows directly from (4.4) whilst the remaining identities are obtained by differentiating the defining expression for  $\Phi$  from (4.5).

*Proof (of Lemma 4.1).* By (4.4) one may assume that

$$(4.7) \quad \phi(x; \omega) = \langle \partial_\omega\phi(x; 0), \omega \rangle + \rho(x; \omega)$$

where  $\rho(0; \omega) = 0$  for all  $\omega \in \Omega$ . Let  $\Phi$  be the function defined in (4.5) and, using localisation and parametrisation invariance, perform the change of variables  $x' \mapsto \Phi(x', x_n; 0)$  on  $X'$  so that the phase becomes

$$\phi(x; \omega) = \langle x', \omega \rangle + \rho(\Phi(x; 0), x_n; \omega).$$

By (4.6) one has  $\partial_{x_n}\Phi(x', 0; 0) = O(|x|)$  and, taking a Taylor expansion of  $\rho$  in  $x_n$ ,

$$\rho(\Phi(x; 0), x_n; \omega) = \rho(\Phi(x', 0; 0), 0; \omega) + (\partial_{x_n}\rho)(\Phi(x', 0; 0), 0; \omega)x_n + O(|x|^2).$$

Note that the first expression on the right-hand side satisfies

$$\rho(\Phi(x', 0; 0), 0; \omega) = \langle \partial_{x'\omega}\phi(0; 0)^{-1}x', (\partial_{x'}\rho)(0; \omega) \rangle + O(|x|^2)$$

whilst, Taylor-expanding now in  $x'$ , it follows that

$$(\partial_{x_n}\rho)(\Phi(x', 0; 0), 0; \omega) = (\partial_{x_n}\rho)(0; \omega) + O(|x|).$$

Combining these observations, and noting, for instance, that (4.4) implies that  $\partial_x^\alpha \rho(x; 0) = 0$  for all  $\alpha \in \mathbb{N}_0^n$  and  $x \in B^n$ , one deduces that

$$\phi(x; \omega) = \langle x', \omega + \partial_{x'\omega}\phi(0; 0)^{-\top}(\partial_{x'}\rho)(0; \omega) \rangle + x_n(\partial_{x_n}\rho)(0; \omega) + O(|x|^2|\omega|^2);$$

Here the symbol  $\top$  is used to denote the matrix transpose and  $-\top$  the inverse matrix transpose.

Since  $\rho$  is quadratic in  $\omega$ , it follows that  $\omega \mapsto \omega + \partial_{x'\omega}\phi(0; 0)^{-\top}(\partial_{x'}\rho)(0; \omega)$  is a well-defined change of variables in a neighbourhood of the origin and so, once again by localisation and parametrisation invariance, the problem is reduced to considering phase functions of the form (4.2). By the construction  $h$  and  $\mathcal{E}$  are quadratic. Finally, the condition H2<sup>+</sup>) implies that the matrix  $\partial_{\omega\omega}^2\partial_{x_n}\phi(0; 0)$  is positive definite. Applying a linear coordinate change, one may therefore suppose that  $\partial_{\omega\omega}^2\partial_{x_n}\phi(0; 0) = \mathbf{I}_{n-1}$ . On the other hand, clearly  $\partial_{\omega\omega}^2\partial_{x_k}\phi(0; 0) = 0_{n-1}$  for  $1 \leq k \leq n-1$  and  $\partial_{\omega x'}^2\phi(0; 0) = \mathbf{I}_{n-1}$ . By continuity, if the support of  $a$  is sufficiently small, then the conditions of (4.3) are valid on the support of  $a$ .  $\square$

**4.2. Parabolic rescaling.** In addition to the reductions of Lemma 4.1, it will be useful to have control over higher order derivatives of the phase, and also the amplitude function.

**Lemma 4.2.** *To prove Theorem 1.2 for some fixed  $\varepsilon > 0$  it suffices to consider the case where, in addition to the properties described in Lemma 4.1, the phase satisfies*

$$\|\partial_\omega^\beta \partial_x^\alpha \phi\|_{L^\infty(X \times \Omega)} < c_{\text{par}} \quad \text{for } |\alpha| \leq N_{\text{par}}, 3 \leq |\beta| \leq N_{\text{par}}$$

for some small dimensional constant  $c_{\text{par}}$  and large integer  $N_{\text{par}} \in \mathbb{N}$ , which can be chosen to depend on  $n$  and  $\varepsilon$ . Furthermore, one may assume that the amplitude satisfies

$$\|\partial_\omega^\beta \partial_x^\alpha a\|_{L^\infty(X \times \Omega)} \lesssim_{\alpha, \beta} 1 \quad \text{for all } (\alpha, \beta) \in \mathbb{N}_0^{n-1} \times \mathbb{N}_0^n.$$

Henceforth  $c_{\text{par}} > 0$  and  $N_{\text{par}} \in \mathbb{N}$  are assumed to be fixed constants, chosen to satisfy the requirements of the forthcoming arguments. A phase satisfying the properties of Lemma 4.2 for this choice of  $c_{\text{par}}$  and  $N_{\text{par}}$  is said to be *reduced*. This notion of reduced phase is precisely that which appears, then undefined, in the statement of the  $k$ -broad estimates of Theorem 1.6.

*Proof (of Lemma 4.2).* The proof of Lemma 4.2 relies upon a standard parabolic rescaling. One may assume that the phase of  $T^\lambda$  is given by  $\phi^\lambda(x; \omega) := \lambda\phi(x/\lambda; \omega)$  where

$$\phi(x; \omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x; \omega) \quad \text{for } (x; \omega) \in X \times \Omega.$$

Let  $\rho \geq 1$ ,  $f \in L^1(B^{n-1})$  and cover  $B^{n-1}$  by finitely-overlapping  $\rho^{-1}$ -balls. Provided  $\rho$  is chosen to depend only on  $\phi$  and  $\varepsilon > 0$ , by the triangle inequality one may assume that  $f$  is supported on one such ball, say  $B(\bar{\omega}, \rho^{-1})$  where  $\bar{\omega} \in B^{n-1}$ . Thus, by a linear change of variables,

$$|T^\lambda f(x)| = \left| \int_{B^{n-1}} e^{2\pi i(\phi^\lambda(x; \bar{\omega} + \rho^{-1}\omega) - \phi^\lambda(x; \bar{\omega}))} a^\lambda(x; \bar{\omega} + \rho^{-1}\omega) \tilde{f}(\omega) d\omega \right|$$

where  $\tilde{f}(\omega) := \rho^{-(n-1)} f(\bar{\omega} + \rho^{-1}\omega)$ . The phase function appearing in the above oscillatory integral may be expressed as

$$\rho^{-1} \langle \partial_\omega \phi^\lambda(x; \bar{\omega}), \omega \rangle + \rho^{-2} (x_n \tilde{h}(\omega) + \lambda \tilde{\mathcal{E}}_1(x/\lambda; \omega))$$

where

$$\tilde{h}(\omega) := \rho^2 (h(\bar{\omega} + \rho^{-1}\omega) - h(\bar{\omega}) - \rho^{-1} \langle \partial_\omega h(\bar{\omega}), \omega \rangle)$$

and

$$\tilde{\mathcal{E}}_1(x; \omega) := \rho^2 (\mathcal{E}(x; \bar{\omega} + \rho^{-1}\omega) - \mathcal{E}(x; \bar{\omega}) - \rho^{-1} \langle \partial_\omega \mathcal{E}(x; \bar{\omega}), \omega \rangle).$$

Defining  $\tilde{\mathcal{E}}(x; \omega) := \tilde{\mathcal{E}}_1(\Phi(\rho^{-1}x', x_n; \bar{\omega}), x_n; \omega)$  where  $\Phi$  is the function introduced in (4.5), it follows that under a change of the  $x$  variables the phase and amplitude are transformed into  $\tilde{\phi}^{\lambda/\rho^2}(x; \omega)$  and  $\tilde{a}^{\lambda/\rho^2}(x; \omega)$ , respectively, where

$$\tilde{\phi}(x; \omega) := \langle x', \omega \rangle + x_n \tilde{h}(\omega) + \tilde{\mathcal{E}}(x; \omega)$$

and

$$\tilde{a}(x; \omega) := a(\Phi(\rho^{-1}x', x_n; \bar{\omega}), x_n; \bar{\omega} + \rho^{-1}\omega).$$

In particular, defining

$$(4.8) \quad \tilde{T}^{\lambda/\rho^2} g(x) := \int_{\mathbb{R}^{n-1}} e^{2\pi i \tilde{\phi}^{\lambda/\rho^2}(x; \omega)} \tilde{a}^{\lambda/\rho^2}(x; \omega) g(\omega) d\omega$$

it follows that

$$\|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim \rho^{(n+1)/p} \|\tilde{T}^{\lambda/\rho^2} \tilde{f}\|_{L^p(\mathbb{R}^n)}.$$

It is easy to verify that the phase  $\tilde{\phi}$  satisfies the conditions of Lemma 4.1 and, provided  $\rho$  is chosen appropriately, it also satisfies the additional conditions described in Lemma 4.2. This facilitates the desired reduction.  $\square$

**4.3. Geometric consequences.** Henceforth all phase functions  $\phi$  are assumed to be reduced in the sense described above. The strategy is to obtain uniform estimates over this class of phases.

By the definition of Hörmander-type operators, for each  $x \in X$  the map  $\omega \mapsto \partial_x \phi(x; \omega)$  parametrises a smooth hypersurface  $\Sigma_x$ . In many respects, these hypersurfaces are geometrically very similar to the paraboloid  $\omega \mapsto (\omega, \frac{|\omega|^2}{2})$ . To see this, recall that  $\Psi: U \rightarrow \Omega$  is a smooth function which satisfies

$$(4.9) \quad \partial_{x'} \phi(x; \Psi(x; u)) = u$$

for all  $(x; u) \in U \subset X \times \mathbb{R}^{n-1}$ . On each of the fibres  $U_x := \{u \in \mathbb{R}^{n-1} : (x; u) \in U\}$  of the domain  $U$ , the map  $u \mapsto \Psi(x; u)$  is a diffeomorphism. Thus, (4.9) implies that  $\Sigma_x$  is the graph of the function

$$h_x(u) := \partial_{x_n} \phi(x; \Psi(x; u))$$

over the fibre  $U_x$ . Each  $h_x$  is a perturbation of  $\frac{|u|^2}{2}$  in the following sense.

**Lemma 4.3.** *The function  $h_x$  satisfies  $h_x(0) = 0$ ,  $\partial_u h_x(0) = 0$  and*

$$\|\partial_{uu}^2 h_x(u) - \mathbf{I}_{n-1}\|_{\text{op}} = O(c_{\text{par}})$$

for all  $u \in U_x$ .

Before proving the lemma, some simple properties of  $\Psi$  are recorded. By (4.4) it follows that  $\Psi(x; 0) = 0$ . The implicit function theorem implies that  $\partial_u \Psi(x; u) = \partial_{x'\omega}^2 \phi(x; \Psi(x; u))^{-1}$  so that, by (4.3) and the local Lipschitz continuity of taking matrix inverses,

$$(4.10) \quad \|\partial_u \Psi(x; u) - \mathbf{I}_{n-1}\|_{\text{op}} = O(c_{\text{par}}).$$

As a consequence of this identity (and choosing  $c_{\text{par}}$  to be sufficiently small),

$$(4.11) \quad |\Psi(x; u) - \Psi(x; u')| \sim |u - u'| \quad \text{for all } u, u' \in U_x,$$

where the implied constant depends only on  $n$ . In addition, if  $1 \leq k \leq n-1$ , then by twice differentiating  $\partial_{x_k} \phi(x; \Psi(x; u)) = u_k$  in the  $u$  variables one may deduce that

$$(4.12) \quad \|\partial_{uu}^2 \Psi_k(x; u)\|_{\text{op}} = O(c_{\text{par}}).$$

The stated properties of  $h_x$  now easily follow.

*Proof (of Lemma 4.3).* By (4.4) one has

$$h_x(0) = \partial_{x_n} \phi(x; \Psi(x; 0)) = \partial_{x_n} \phi(x; 0) = 0.$$

Similarly,  $\partial_u h_x(0) = 0$  since  $\partial_\omega \partial_{x_n} \phi(x; 0) = 0$ . Finally,

$$\partial_{uu}^2 h_x(u) = (\partial_u \Psi)^\top(x; u) (\partial_{\omega\omega}^2 \partial_{x_n} \phi)(x; \Psi(x; u)) \partial_u \Psi(x; u) + E(x; u)$$

where  $E(x; \omega)$  is the  $(n-1) \times (n-1)$  matrix whose  $(i, j)$ th entry is given by

$$E_{ij}(x; u) = \langle (\partial_\omega \partial_{x_n} \phi)(x; \Psi(x; u)), \partial_{u_i u_j} \Psi(x; u) \rangle.$$

By (4.3) and (4.12) it follows that  $\|E(x; u)\|_{\text{op}} = O(c_{\text{par}})$ , whilst (4.3) and multiple applications of (4.10) imply that

$$\|(\partial_u \Psi)^\top(x; u)(\partial_{\omega\omega}^2 \partial_{x_n} \phi)(x; \Psi(x; u))\partial_u \Psi(x; u) - \mathbf{I}_{n-1}\|_{\text{op}} = O(c_{\text{par}}).$$

This concludes the proof.  $\square$

Similar reasoning can be used to provide useful uniform estimates for the generalised Gauss map associated to  $T^\lambda$ . To state the result, let  $X^\lambda := \{x \in \mathbb{R}^n : x/\lambda \in X\}$  denote the  $\lambda$ -dilate of  $X$ , so that  $a^\lambda$  is supported in  $X^\lambda \times \Omega$ .

**Lemma 4.4.** *For all  $x, x' \in X^\lambda$  and  $\omega, \omega' \in \Omega$  the estimates*

$$\chi(G^\lambda(x; \omega), G^\lambda(x; \omega')) \sim |\omega - \omega'| \quad \text{and} \quad \chi(G^\lambda(x; \omega), G^\lambda(x'; \omega)) \lesssim \lambda^{-1}|x_1 - x_2|$$

*hold with implied constants which depend only on the dimension.*

The proof, which is an elementary calculus exercise in the style of the proof of Lemma 4.3, is omitted.

## 5. BASIC ANALYTIC PRELIMINARIES

**5.1. Wave packet decomposition.** Throughout the following sections  $\varepsilon > 0$  is a fixed small parameter and  $\delta > 0$  is a tiny number satisfying<sup>7</sup>  $\delta \ll \varepsilon$  and  $\delta \sim_\varepsilon 1$ .

A wave packet decomposition is carried out with respect to some spatial parameter  $1 \ll R \ll \lambda$ . Cover  $B^{n-1}$  by finitely-overlapping balls  $\theta$  of radius  $R^{-1/2}$  and let  $\psi_\theta$  be a smooth partition of unity adapted to this cover. These  $\theta$  will frequently be referred to as  $R^{-1/2}$ -caps. Cover  $\mathbb{R}^{n-1}$  by finitely-overlapping balls of radius  $cR^{(1+\delta)/2}$  centred on points belonging to the lattice  $R^{(1+\delta)/2}\mathbb{Z}^{n-1}$ . By Poisson summation one may find a bump function adapted to  $B(0, R^{(1+\delta)/2})$  so that the functions  $\eta_v(z) := \eta(z - v)$  for  $v \in R^{(1+\delta)/2}\mathbb{Z}^{n-1}$  form a partition of unity for this cover. Let  $\mathbb{T}$  denote the collection of all pairs  $(\theta, v)$ . Thus, for  $f: \mathbb{R}^{n-1} \rightarrow \mathbb{C}$  with support in  $B^{n-1}$  and belonging to some suitable *a priori* class one has

$$f = \sum_{(\theta, v) \in \mathbb{T}} (\eta_v(\psi_\theta f))^\wedge = \sum_{(\theta, v) \in \mathbb{T}} \hat{\eta}_v * (\psi_\theta f).$$

For each  $R^{-1/2}$ -cap  $\theta$  let  $\omega_\theta \in B^{n-1}$  denote its centre. Choose a real-valued smooth function  $\tilde{\psi}$  so that the function  $\tilde{\psi}_\theta(\omega) := \tilde{\psi}(R^{1/2}(\omega - \omega_\theta))$  is supported in  $\theta$  and  $\tilde{\psi}_\theta(\omega) = 1$  whenever  $\omega$  belongs to a  $cR^{-1/2}$  neighbourhood of the support of  $\psi_\theta$  for some small constant  $c > 0$ . Finally, define

$$f_{\theta, v} := \tilde{\psi}_\theta \cdot [\hat{\eta}_v * (\psi_\theta f)].$$

The function  $\hat{\eta}_v(\omega)$  is rapidly decaying for  $|\omega| \gtrsim R^{-(1+\delta)/2}$  and, consequently,

$$\|f_{\theta, v} - \hat{\eta}_v * (\psi_\theta f)\|_{L^\infty(\mathbb{R}^{n-1})} \leq \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

Here the notation  $\text{RapDec}(R)$  is used to denote any quantity  $C_R$  which is rapidly decaying in  $R$ ; that is, any  $C_R \in \mathbb{R}$  such that  $|C_R| \lesssim_{\varepsilon, N} R^{-N}$  for all  $N \in \mathbb{N}$ . It follows that

$$\|f - \sum_{\theta, v} f_{\theta, v}\|_{L^\infty(\mathbb{R}^{n-1})} \leq \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

<sup>7</sup>For  $A, B \geq 0$  the notation  $A \ll B$  or  $B \gg A$  is used to denote that  $A$  is ‘much smaller’ than  $B$ ; a more precise interpretation of this is that  $A \leq C_\varepsilon^{-1} B$  for some constant  $C_\varepsilon \geq 1$  which can be chosen to be large depending on  $n$  and  $\varepsilon$ .

The functions  $f_{\theta,v}$  are almost orthogonal: if  $\mathbb{S} \subseteq \mathbb{T}$ , then

$$\left\| \sum_{(\theta,v) \in \mathbb{S}} f_{\theta,v} \right\|_{L^2(\mathbb{R}^{n-1})}^2 \sim \sum_{(\theta,v) \in \mathbb{S}} \|f_{\theta,v}\|_{L^2(\mathbb{R}^{n-1})}^2.$$

Let  $T^\lambda$  be an operator with reduced phase  $\phi$  and amplitude  $a$  supported in  $X \times \Omega$  as in Lemma 4.1. For  $(\theta, v) \in \mathbb{T}$  define the curve  $\gamma_{\theta,v}^1: I_{\theta,v}^1 \rightarrow \mathbb{R}^{n-1}$  by setting  $\gamma_{\theta,v}^1(t) := \Phi(v, t; \omega_\theta)$ , where  $\Phi$  is the function introduced in §4 and

$$I_{\theta,v}^1 := \{t \in X_n : \partial_\omega \phi(x', t; \omega_\theta) = v \text{ for some } x' \in X'\}.$$

Thus,  $\partial_\omega \phi(\gamma_{\theta,v}^1(t), t; \omega_\theta) = v$  for all  $t \in I_{\theta,v}^1$ . Moreover, the rescaled curve  $\gamma_{\theta,v}^\lambda(t) := \lambda \gamma_{\theta,v}^1(t/\lambda)$  satisfies

$$\partial_\omega \phi^\lambda(\gamma_{\theta,v}^\lambda(t), t; \omega_\theta) = v \quad \text{for all } t \in I_{\theta,v}^\lambda := \{t \in \mathbb{R} : t/\lambda \in I_{\theta,v}^1\}.$$

Let  $\Gamma_{\theta,v}^\lambda: I_{\theta,v}^\lambda \rightarrow \mathbb{R}^n$  denote the graphing map  $\Gamma_{\theta,v}^\lambda(t) := (\gamma_{\theta,v}^\lambda(t), t)$ ; by an abuse of notation  $\Gamma_{\theta,v}^\lambda$  will also be used to denote the image of this mapping. It is remarked that one may easily verify that the tangent vector to the curve  $\Gamma_{\theta,v}^\lambda$  at  $\Gamma_{\theta,v}^\lambda(t)$  is parallel to  $G^\lambda(\Gamma_{\theta,v}^\lambda(t); \omega_\theta)$  for all  $t \in I_{\theta,v}^\lambda$ .

Define the curved  $R^{1/2+\delta}$ -tube

$$T_{\theta,v} := \{(x', x_n) \in B(0, R) : x_n \in I_{\theta,v}^\lambda \text{ and } |x' - \gamma_{\theta,v}^\lambda(x_n)| \leq R^{1/2+\delta}\}.$$

The curve  $\Gamma_{\theta,v}^\lambda$  is referred to as the *core* of  $T_{\theta,v}$ . Observe that, since  $\phi$  is of the reduced form defined in §4, one has

$$(5.1) \quad |x' - \gamma_{\theta,v}^\lambda(x_n)| \sim |\partial_\omega \phi^\lambda(x; \omega_\theta) - v|,$$

for all  $x = (x', x_n) \in X^\lambda$  with  $x_n \in I_{\theta,v}^\lambda$ .

**Example 5.1.** Let  $\phi^\lambda(x; \omega) := \langle x', \omega \rangle + x_n h(\omega)$  and observe that  $\gamma_{\theta,v}^\lambda(t) = v - t \partial_\omega h(\omega_\theta)$  parametrises a straight line through  $v$  in the direction of  $\partial_\omega h(\omega_\theta)$ . The tube is given by

$$T_{\theta,v} = \{x \in B(0, R) : |x' + x_n \partial_\omega h(\omega_\theta) - v| \leq R^{1/2+\delta}\},$$

which agrees with those studied in the case of the extension operator.

**Lemma 5.2.** *If  $1 \ll R \ll \lambda$  and  $x \in B(0, R) \setminus T_{\theta,v}$ , then*

$$|T^\lambda f_{\theta,v}(x)| \leq (1 + R^{-1/2} |\partial_\omega \phi^\lambda(x; \omega_\theta) - v|)^{-(n+1)} \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

*Proof.* Observe that

$$T^\lambda f_{\theta,v}(x) = \int_{\mathbb{R}^{n-1}} \hat{\eta}_v * (\psi_\theta f) \cdot \overline{G_x}$$

where

$$G_x(\omega) := e^{-2\pi i \phi^\lambda(x; \omega)} a^\lambda(x; \omega) \tilde{\psi}_\theta(\omega)$$

and so, by Plancherel,

$$T^\lambda f_{\theta,v}(x) = \int_{\mathbb{R}^{n-1}} \eta_v \cdot (\psi_\theta f)^\vee \cdot \overline{\tilde{G}_x}.$$

By a simple change of variables, if one defines the phase function  $\phi_{x,z}^\lambda(\omega) := R^{1/2} \phi^\lambda(x; \omega_\theta + R^{-1/2} \omega) - \langle z, \omega \rangle$ , then one has

$$|\tilde{G}_x(z)| = R^{-(n-1)/2} \left| \int_{\mathbb{R}^{n-1}} e^{-2\pi i R^{-1/2} \phi_{x,z}^\lambda(\omega)} a^\lambda(x; \omega_\theta + R^{1/2} \omega) \tilde{\psi}(\omega) d\omega \right|.$$

This integral is analysed using (non-)stationary phase. Computing the derivative of the phase function,

$$\begin{aligned}\partial_\omega \phi_{x,z}^\lambda(\omega) &= \partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - z \\ &= [\partial_\omega \phi^\lambda(x; \omega_\theta) - v] + [v - z] + [\partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(x; \omega_\theta)].\end{aligned}$$

Observe that  $|v - z| \lesssim R^{(1+\delta)/2}$  whenever  $z \in \text{supp } \eta_v$ . Moreover,

$$|\partial_{\omega_j} \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_{\omega_j} \phi^\lambda(x; \omega_\theta)| \leq R^{-1/2} \int_0^1 |\langle \partial_\omega \partial_{\omega_j} \phi^\lambda(x; \omega_\theta + tR^{-1/2}\omega), \omega \rangle| dt.$$

Since  $\partial_{\omega_i \omega_j}^2 \phi^\lambda(0; \omega) = 0$  for all  $\omega \in \Omega$  and  $1 \leq i, j \leq n-1$ , it follows that

$$\begin{aligned}|\partial_{\omega_i \omega_j}^2 \phi^\lambda(x; \omega_\theta + tR^{-1/2}\omega)| &= \int_0^1 |\langle \partial_x \partial_{\omega_i \omega_j}^2 \phi^\lambda(sx; \omega_\theta + tR^{-1/2}\omega), x \rangle| ds \\ &\leq \| \partial_x \partial_{\omega_i \omega_j}^2 \phi^\lambda \|_{L^\infty(X^\lambda \times \Omega)} |x| \lesssim R\end{aligned}$$

for  $x \in B(0, R) \cap X^\lambda$ . Combining these estimates,

$$|\partial_\omega \phi^\lambda(x; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(x; \omega_\theta)| \lesssim R^{1/2}.$$

On the other hand, for  $x \in B(0, R) \setminus T_{\theta,v}$  it is claimed that

$$(5.2) \quad |\partial_\omega \phi^\lambda(x; \omega_\theta) - v| \gtrsim R^{1/2+\delta}.$$

If  $x_n \in I_{\theta,v}^\lambda$ , then (5.2) follows directly from the definition of  $T_{\theta,v}$  and (5.1). Temporarily assuming (5.2) holds in general, it follows that the  $\partial_\omega \phi^\lambda(x; \omega_\theta) - v$  term dominates in the above expansion of  $\partial_\omega \phi_{x,z}^\lambda(\omega)$ . In particular, for all  $z \in \text{supp } \eta_v$ , one concludes that

$$R^{-1/2} |\partial_\omega \phi_{x,z}^\lambda(\omega)| \gtrsim R^{-1/2} |\partial_\omega \phi^\lambda(x; \omega_\theta) - v| \gtrsim R^\delta$$

whenever  $R \gg 1$ . Thus, by repeated integration by parts,

$$|\check{G}_x(z)| = (1 + R^{-1/2} |\partial_\omega \phi^\lambda(x; \omega_\theta) - v|)^{-(n+1)} \text{RapDec}(R).$$

The desired inequality is now an immediate consequence of Young's inequality and Plancherel's theorem.

It remains to establish (5.2) for  $x \in B(0, R) \setminus T_{\theta,v}$  with  $x_n \notin I_{\theta,v}^\lambda$ . The condition on  $x_n$  implies that  $v \notin \tilde{X}'$  where  $\tilde{X}'$  is defined to be the image of  $X'$  under the diffeomorphism  $z' \mapsto \lambda \partial_\omega \phi(z', x_n/\lambda; \omega_\theta)$ . The set  $\tilde{X}'$  will contain a ball centred at  $\lambda \partial_\omega \phi(0, x_n/\lambda; \omega_\theta)$  of radius  $2c\lambda$  for some dimensional constant  $c > 0$ . Since  $|x_n| < R$  and  $\partial_\omega \phi(0; \omega_\theta) = 0$ , one observes that  $\lambda |\partial_\omega \phi(0, x_n/\lambda; \omega_\theta)| \lesssim R$  and so  $B(0, c\lambda) \subseteq \tilde{X}'$ , provided  $R \ll \lambda$ . Consequently,  $|v| \gtrsim \lambda$  whilst, on the other hand,  $|\partial_\omega \phi^\lambda(x; \omega_\theta)| \lesssim R$  and so, again provided  $R \ll \lambda$ , the estimate (5.2) immediately follows.<sup>8</sup>  $\square$

<sup>8</sup>This argument can also be used to show that for  $|v| \lesssim \lambda$ , the domain  $I_{\theta,v}^\lambda$  contains an interval about 0 of length  $\sim \lambda$ .

**5.2. An  $L^2$  estimate.** The following standard  $L^2$ -bound, which has been mentioned previously in §2, will play a significant rôle in the forthcoming analysis.

**Lemma 5.3** (Hörmander [16]). *If  $1 \leq R \leq \lambda$  and  $B_R$  is any ball of radius  $R$ , then*

$$\|T^\lambda f\|_{L^2(B_R)} \lesssim R^{1/2} \|f\|_{L^2(B^{n-1})}.$$

This lemma is a direct corollary of the following lemma which, in turn, is a consequence of Hörmander's generalisation of the Hausdorff–Young inequality [16].

**Lemma 5.4.** *For any fixed  $x_n \in \mathbb{R}$ , the estimate*

$$\|T^\lambda f\|_{L^2(\mathbb{R}^{n-1} \times \{x_n\})} \lesssim \|f\|_{L^2(B^{n-1})}$$

*holds.*

*Proof.* Defining  $Sf(x') := T^\lambda f(\lambda x', x_n)$ , the problem is to show that

$$(5.3) \quad \|Sf\|_{L^2(\mathbb{R}^{n-1})} \lesssim \lambda^{-(n-1)/2} \|f\|_{L^2(\mathbb{R}^{n-1})}.$$

Observe that

$$Sf(x') = \int_{\mathbb{R}^{n-1}} e^{2\pi i \lambda \phi(x', x_n/\lambda; \omega)} a(x', x_n/\lambda; \omega) f(\omega) \, d\omega.$$

The original hypotheses on the phase  $\phi$  imply that

$$|\det \partial_{x'\omega}^2 \phi(x', x_n/\lambda; \omega)| \gtrsim 1$$

whilst  $(x'; \omega) \mapsto a(x', x_n/\lambda; \omega)$  has support in some bounded subset of  $\mathbb{R}^{n-1} \times \mathbb{R}^{n-1}$ . Both these conditions hold uniformly in  $x_n$  and  $\lambda$ . Thus, the operator  $S$  satisfies the conditions of Hörmander's generalisation of the Hausdorff–Young inequality [16] (see also, for instance, [20, p. 377]) uniformly in  $x_n$  and  $\lambda$ . Applying Hörmander's theorem immediately yields (5.3).  $\square$

**5.3. The locally constant property.** As a final analytic preliminary, some simple consequences of the uncertainty principle are discussed. It is remarked that the results of this subsection (that is, Lemma 5.6 and Corollary 5.7) only play a rôle in the proof of Theorem 1.2 much later in the argument (namely, in the parabolic rescaling argument in §11). They do, however, feature in an independent discussion in the following section.

**Definition 5.5.** A function  $\zeta: \mathbb{R}^n \rightarrow [0, \infty)$  is said to be *locally constant at scale  $\rho$*  for some  $\rho > 0$  if  $\zeta(x) \sim \zeta(y)$  for all  $x, y \in \mathbb{R}^n$  with  $|x - y| \lesssim \rho$ .

Owing to the uncertainty principle, heuristically one expects the following: if  $f$  is supported on a  $\rho^{-1}$ -cap, then  $|T^\lambda f|$  is essentially constant at scale  $\rho$ . For extension operators this is due to the fact that, under the support hypothesis on the input function,  $Ef$  has (distributional) Fourier support inside a  $\rho^{-1}$ -ball. For general Hörmander-type operators  $T^\lambda$  the Fourier transform of  $T^\lambda f$  does not necessarily have compact support. It will, however, be concentrated in some  $\rho^{-1}$ -ball and this is sufficient to ensure the locally constant property holds. This discussion is formalised by the following lemma and its corollary.

**Lemma 5.6.** *Let  $T^\lambda$  be a Hörmander-type operator. There exists a smooth, rapidly decreasing function  $\eta: \mathbb{R}^n \rightarrow \mathbb{R}$  with the following property: if  $\delta > 0$  and  $1 \leq \rho \leq \lambda^{1-\delta}$ , then*

$$(5.4) \quad e^{-2\pi i \phi^\lambda(\cdot; \bar{\omega})} T^\lambda f = [e^{-2\pi i \phi^\lambda(\cdot; \bar{\omega})} T^\lambda f] * \eta_\rho + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}$$

holds whenever  $f$  is supported in a  $\rho^{-1}$ -ball  $B(\bar{\omega}, \rho^{-1})$  for some  $\bar{\omega} \in \Omega$ . Here  $\eta_\rho(x) := \rho^{-n}\eta(x/\rho)$ .

The identity (5.4) can be roughly interpreted as saying that  $e^{-2\pi i\phi^\lambda(x;\bar{\omega})}T^\lambda f(x)$  is equal to an average of itself over a  $\rho$ -ball centred at  $x$ . This is a rigorous formulation of the locally constant heuristic discussed above.

*Proof (of Lemma 5.6).* Suppose that  $\text{supp } f \subset B(\bar{\omega}, \rho^{-1})$  where  $\bar{\omega} \in \Omega$  and observe that

$$(5.5) \quad [e^{-2\pi i\phi^\lambda(\cdot;\bar{\omega})}T^\lambda f]^\wedge(\xi) = \int_{\mathbb{R}^{n-1}} K^\lambda(\xi;\omega)f(\omega) d\omega$$

where the function  $K^\lambda$  is given by

$$K^\lambda(\xi;\omega) = \lambda^n \int_{\mathbb{R}^n} e^{-2\pi i\lambda(\langle x,\xi \rangle - \phi(x;\omega) + \phi(x;\bar{\omega}))} a(x;\omega) dx.$$

This oscillatory integral is estimated via (non-)stationary phase, using the simple estimate

$$|\partial_x \phi(x;\omega) - \partial_x \phi(x;\bar{\omega})| \lesssim \rho^{-1} \quad \text{for } (x;\omega) \in X \times \Omega \text{ with } \omega \in \text{supp } f.$$

In particular, if  $|\xi| \geq C\rho^{-1}$  for a suitably large constant  $C \geq 1$ , then repeated integration-by-parts shows that

$$|K^\lambda(\xi;\omega)| \leq \text{RapDec}(\lambda)(1 + |\xi|)^{-(n+1)}.$$

Let  $\eta$  be a Schwartz function on  $\mathbb{R}^n$  with  $\hat{\eta}(\xi) = 1$  for all  $|\xi| < C$  for a suitable constant  $C \geq 1$  and support in  $B(0, 2C)$ . Such a function can further be chosen to be locally constant at scale 1. From the above observations it follows that the left-hand side of (5.5) can be expressed as

$$[e^{-2\pi i\phi^\lambda(\cdot;\bar{\omega})}T^\lambda f]^\wedge(\xi)\hat{\eta}_\rho(\xi) + \text{RapDec}(\lambda)(1 + |\xi|)^{-(n+1)}\|f\|_{L^2(B^{n-1})}$$

and the desired identity is now immediate.  $\square$

In order to apply the locally constant property effectively in  $k$ -linear settings, it is useful to establish a variant of the above lemma involving expressions of the form  $|T^\lambda f|^{1/k}$ . This can be achieved by combining Lemma 5.6 together with Bernstein's inequality.

**Corollary 5.7.** *Let  $T^\lambda$  be a Hörmander-type operator and  $1 \leq k \leq n$ . There exists a smooth, rapidly decreasing function  $\zeta: \mathbb{R}^n \rightarrow [0, \infty)$  with the following properties:*

- 1)  $\zeta$  is locally constant at scale 1.
- 2) If  $\delta > 0$  and  $1 \leq \rho \leq \lambda^{1-\delta}$ , then the pointwise inequality

$$(5.6) \quad |T^\lambda f|^{1/k} \lesssim |T^\lambda f|^{1/k} * \zeta_\rho + \text{RapDec}(\lambda)\|f\|_{L^2(B^{n-1})}^{1/k}$$

holds whenever  $f$  is supported in some  $\rho^{-1}$ -ball. Here  $\zeta_\rho(x) := \rho^{-n}\zeta(x/\rho)$ .

The locally constant property of  $\zeta$  implies that

$$|T^\lambda f|^{1/k} * \zeta_\rho(x) \sim |T^\lambda f|^{1/k} * \zeta_\rho(y) \quad \text{for all } x, y \in \mathbb{R}^n \text{ with } |x - y| \lesssim \rho;$$

namely,  $|T^\lambda f|^{1/k} * \zeta_\rho$  is locally constant at scale  $\rho$ . This is again consistent with uncertainty principle heuristics.

*Proof (of Corollary 5.7).* From the proof of Lemma 5.6 it is clear that the function  $\eta$  may be chosen so that  $|\eta|^{1/k}$  admits a smooth, non-negative, rapidly decreasing majorant  $\zeta$  which is locally constant at scale 1. Iterating the formula (5.4), one deduces that

$$|T^\lambda f(x)| \leq \left| \int_{\mathbb{R}^n} [e^{-2\pi i \phi(\cdot; \bar{\omega})} T^\lambda f] * \eta_\rho(x-y) \eta_\rho(y) dy \right| + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}.$$

Observe that the above integrand has Fourier support in a ball of radius  $O(\rho^{-1})$ . Bernstein's inequality<sup>9</sup> may therefore be applied to dominate the right-hand side by

$$\left( \int_{\mathbb{R}^n} |[e^{-2\pi i \phi(\cdot; \bar{\omega})} T^\lambda f] * \eta_\rho(x-y)|^{1/k} \zeta_\rho(y) dy \right)^k + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}.$$

A final application of (5.4) allows one to replace  $|[e^{-2\pi i \phi(\cdot; \bar{\omega})} T^\lambda f] * \eta_\rho|$  with  $|T^\lambda f|$  in this expression, concluding the proof.  $\square$

## 6. PROPERTIES OF THE $k$ -BROAD NORMS

**6.1.  $k$ -broad triangle inequality and logarithmic convexity.** The functional  $f \mapsto \|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}$  is not a norm in a literal sense, but it does exhibit some properties similar to those of  $L^p$ -norms. For instance, the map  $U \mapsto \|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}$  behaves similarly to a measure.

**Lemma 6.1** (Finite (sub)-additivity). *Let  $U_1, U_2 \subseteq \mathbb{R}^n$  and  $U := U_1 \cup U_2$ . If  $1 \leq p < \infty$  and  $A$  is a non-negative integer, then*

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}^p \leq \|T^\lambda f\|_{\text{BL}_{k,A}^p(U_1)}^p + \|T^\lambda f\|_{\text{BL}_{k,A}^p(U_2)}^p$$

holds for all integrable  $f: B^{n-1} \rightarrow \mathbb{C}$ .

This is an immediate consequence of the definition of the  $k$ -broad norms. A slightly less trivial observation is that  $\|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}$  also satisfies weak versions of the triangle and logarithmic convexity inequalities.

**Lemma 6.2** (Triangle inequality [13]). *If  $U \subseteq \mathbb{R}^n$ ,  $1 \leq p < \infty$  and  $A := A_1 + A_2$  for  $A_1, A_2$  non-negative integers, then*

$$\|T^\lambda(f_1 + f_2)\|_{\text{BL}_{k,A}^p(U)} \lesssim \|T^\lambda f_1\|_{\text{BL}_{k,A_1}^p(U)} + \|T^\lambda f_2\|_{\text{BL}_{k,A_2}^p(U)}$$

holds for all integrable  $f_1, f_2: B^{n-1} \rightarrow \mathbb{C}$ .

**Lemma 6.3** (Logarithmic convexity [13]). *Suppose that  $U \subseteq \mathbb{R}^n$ ,  $1 \leq p, p_1, p_2 < \infty$  and  $0 \leq \alpha_1, \alpha_2 \leq 1$  satisfy  $\alpha_1 + \alpha_2 = 1$  and*

$$\frac{1}{p} = \frac{\alpha_1}{p_1} + \frac{\alpha_2}{p_2}.$$

*If  $A := A_1 + A_2$  for  $A_1, A_2$  non-negative integers, then*

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(U)} \lesssim \|T^\lambda f_1\|_{\text{BL}_{k,A_1}^{p_1}(U)}^{\alpha_1} \|T^\lambda f_2\|_{\text{BL}_{k,A_2}^{p_2}(U)}^{\alpha_2}$$

<sup>9</sup>More precisely, here the proof uses a general form of Bernstein's inequality, valid for exponents less than 1. In particular, if  $0 < p \leq q \leq \infty$  and  $g$  is an integrable function on  $\mathbb{R}^n$  satisfying  $\text{supp } \hat{g} \subseteq B_r$ , then

$$\|g\|_{L^q(\mathbb{R}^n)} \lesssim r^{n(1/p-1/q)} \|g\|_{L^p(\mathbb{R}^n)}.$$

This extension follows from the classical Bernstein inequality (that is, the above estimate in the restricted range  $1 \leq p \leq q \leq \infty$ ) in a rather straight-forward manner.

holds for all integrable  $f: B^{n-1} \rightarrow \mathbb{C}$ .

These estimates are proven in the context of Fourier extension operators in [13]. The arguments are entirely elementary and readily generalise to the variable coefficient case. It is remarked that the parameter  $A$  appears in the definition of the  $k$ -broad norm to allow for these weak triangle and logarithmic convexity inequalities.

**6.2.  $k$ -broad versus  $k$ -linear.** A relationship between  $k$ -broad and  $k$ -linear estimates is given by the following proposition.

**Proposition 6.4.** *Let  $\mathcal{T}$  be a class of Hörmander-type operators which is closed under translation,<sup>10</sup>  $2 \leq p < \infty$ ,  $2 \leq k \leq n$  and  $\varepsilon > 0$ . Suppose that for all  $1 \ll R \leq \lambda$  and  $R$ -balls  $B_R$  the  $k$ -linear inequality*

$$\left\| \prod_{j=1}^k |T_j^\lambda f_j|^{1/k} \right\|_{L^p(B_R)} \lesssim_{\varepsilon, (\phi_j)_{j=1}^k} \nu^{-C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds whenever  $(T_1^\lambda, \dots, T_k^\lambda) \in \mathcal{T}^k$  is a  $\nu$ -transverse  $k$ -tuple of Hörmander-type operators. Then for all  $1 \ll R \leq \lambda$  and  $R$ -balls  $B_R$  the  $k$ -broad inequality

$$\|T^\lambda f\|_{\text{BL}_{k,1}^p(B_R)} \lesssim_{\varepsilon, \phi} K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds for any  $T^\lambda \in \mathcal{T}$ .

The parameter  $K$  in the above theorem is the same as that which appears in the definition of the  $k$ -broad norms; the  $C_\varepsilon$  denote constants, which may vary from line to line, which depend only on  $n$  and  $\varepsilon$ .

The (local version of the) Bennett–Carbery–Tao theorem [4] therefore implies a version of Theorem 1.6 which holds for all Hörmander-type operators (that is, without the positive-definite hypothesis) with a restricted range of  $p$ .

**Corollary 6.5.** *Let  $T^\lambda$  be a Hörmander-type operator. For all  $2 \leq k \leq n$ ,  $p \geq 2k/(k-1)$  and  $\varepsilon > 0$  the estimate*

$$\|T^\lambda f\|_{\text{BL}_{k,1}^p(B_R)} \lesssim_{\varepsilon, \phi} K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})}$$

holds for all  $\lambda \geq 1$ .

For completeness the proof of Proposition 6.4 is given; the result itself will not be used in the proof of Theorem 1.2 and is included mainly for expository purposes. Thus, readers interested only in the proof of Theorem 1.2 may safely skip to the next section.

*Proof (of Proposition 6.4).* Let  $\mathcal{Z} \subset B_R$  be a maximal set of points with the property that the balls  $B(z, R/2\bar{C}K)$  for  $z \in \mathcal{Z}$  are pairwise disjoint. Here  $\bar{C} \geq 1$  is a suitable constant, chosen to meet the forthcoming requirements of the proof. Letting  $B_z := B(z, R/\bar{C}K)$  for  $z \in \mathcal{Z}$ , it follows that  $\#\mathcal{Z} \lesssim K^n$  and

$$\|T^\lambda f\|_{\text{BL}_{k,1}^p(B_R)}^p \leq \sum_{z \in \mathcal{Z}} \|T^\lambda f\|_{\text{BL}_{k,1}^p(B_z)}^p.$$

Fixing  $z \in \mathcal{Z}$  it therefore suffices to show that

$$\|T^\lambda f\|_{\text{BL}_{k,1}^p(B_z)} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^2(B^{n-1})},$$

<sup>10</sup>That is, if  $T^\lambda \in \mathcal{T}$  and  $a \in \mathbb{R}^n$ , then the translated operator  $T_a^\lambda$  defined by  $T_a^\lambda f(x) := T^\lambda f(x+a)$  also belongs to  $\mathcal{T}$ .

since summing the contributions from each choice of  $z \in \mathcal{Z}$  only introduces an acceptable  $K^n$  factor into the estimate. By introducing a bump function into the definition of the operator, one may further assume that the amplitude  $a^\lambda$  has  $x$ -support in the ball concentric to  $B_z$  with twice the radius.

Fix a ball  $B_{K^2} = B(\bar{x}, K^2) \in \mathcal{B}_{K^2}$  with  $B_{K^2} \cap B_z \neq \emptyset$  and suppose that  $V$  is a  $(k-1)$ -dimensional subspace which realises the minimum in  $\mu_{T^\lambda f}(B_{K^2})$ . Thus, by definition, if  $\tau$  is a  $K^{-1}$ -cap for which

$$\|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p > \mu_{T^\lambda f}(B_{K^2}),$$

then  $\tau \in V$ , where the inclusion symbol is used in the non-standard sense described in the introduction. Amongst all such subspaces  $V$  choose one which maximises the cardinality of the set

$$\mathcal{T}(V) := \{\tau \in V : \|T^\lambda f_\tau\|_{L^p(B_{K^2})}^p \geq \mu_{T^\lambda f}(B_{K^2})\}.$$

By definition there exists some cap  $\tau^* \notin V$  such that  $\|T^\lambda f_{\tau^*}\|_{L^p(B_{K^2})}^p = \mu_{T^\lambda f}(B_{K^2})$ .

Suppose there exists a  $(k-2)$ -dimensional subspace  $W \subset \mathbb{R}^n$  such that  $\tau \in W$  for all  $\tau \in \mathcal{T}(V)$ . Then defining  $V' := \text{span}(W \cup \{G^\lambda(\bar{x}, \omega_{\tau^*})\})$  where  $\omega_{\tau^*}$  is the centre of  $\tau^*$ , it follows that  $\tau^* \in V'$  and  $\tau \in V'$  for all  $\tau \in \mathcal{T}(V)$ . On the other hand,  $V'$  also realises the minimum in the definition in  $\mu_{T^\lambda f}(B_{K^2})$ : to see this it suffices to show that

$$(6.1) \quad \|T^\lambda f\|_{L^p(B_{K^2})}^p \leq \mu_{T^\lambda f}(B_{K^2}) \quad \text{for all } \tau \notin V' \text{ with } \tau \in V;$$

if  $\tau \notin V'$ , then  $\tau \notin W$  so that  $\tau$  does not belong to  $\mathcal{T}(V)$  and so (6.1) is immediate. These observations contradict the maximality of  $V$  and, consequently, no such subspace  $W$  can exist.

By the preceding discussion, one may find a family of caps  $\tau_1^*, \dots, \tau_{k-1}^* \in \mathcal{T}(V)$  satisfying

$$(6.2) \quad \left| \bigwedge_{j=1}^k G^\lambda(\bar{x}, \omega_j) \right| \gtrsim K^{-(k-1)} \quad \text{for all } \omega_j \in \tau_j^*, 1 \leq j \leq k.$$

Thus,

$$(6.3) \quad \mu_{T^\lambda f}(B_{K^2}) \leq \prod_{j=1}^k \|T^\lambda f_{\tau_j^*}\|_{L^p(B_{K^2})}^{p/k}$$

for  $\tau_k^* := \tau^*$ . To apply the hypothesised multilinear estimate one wishes to exchange the order of taking the norm and product on the right-hand side of the above expression; that is, one wishes to prove an estimate of the form

$$\prod_{j=1}^k \|T^\lambda f_{\tau_j^*}\|_{L^p(B_{K^2})}^{p/k} \lesssim K^{O(1)} \left\| \prod_{j=1}^k |T^\lambda f_{\tau_j^*}|^{1/k} \right\|_{L^p(B_{K^2})}^p.$$

This is achieved by exploiting the locally constant property of the  $T^\lambda f_\tau$ , as discussed in §5.3. In particular, by Corollary 5.7 and Hölder's inequality there exists a non-negative, rapidly decreasing, locally constant function  $\zeta$  such that

$$(6.4) \quad |T^\lambda f_\tau|^{p/k} \lesssim |T^\lambda f_\tau|^{p/k} * \zeta_K + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}^{1/k}$$

holds for all  $K^{-1}$ -caps  $\tau$ . Since rapidly decaying error terms are entirely harmless, henceforth they will be suppressed in the notation. Observe that for all

$x \in B(\bar{x}, K^2)$  and  $y \in \mathbb{R}^n$  one has  $\zeta_K(x - y) \lesssim K^{O(1)} w_K(\bar{x} - y)$  where  $w_K(y) := (1 + |y|/K)^{-N}$  for some choice suitable of large exponent  $N$  satisfying  $N = O(1)$ . Combining these observations with a second application of (6.4) yields

$$(6.5) \quad \|T^\lambda f_\tau\|_{L^p(B_{K^2})}^{p/k} \lesssim K^{O(1)} \int_{\mathbb{R}^n} |T^\lambda f_\tau|^{p/k} * \zeta_K(y) w_K(\bar{x} - y) dy.$$

Fix  $x \in B(0, K)$  and note that the locally constant property of  $\zeta$  implies that

$$(6.6) \quad |T^\lambda f_\tau|^{p/k} * \zeta_K(y) \lesssim |T^\lambda f_\tau|^{p/k} * \zeta_K(x + y) \quad \text{for all } y \in \mathbb{R}^n.$$

Thus, by (6.3), (6.5) and (6.6), one deduces that

$$\mu_{T^\lambda f}(B_{K^2}) \lesssim K^{O(1)} \int_{(\mathbb{R}^n)^k} \prod_{j=1}^k |T^\lambda f_{\tau_j^*}|^{p/k} * \zeta_K(x + y_j) w_K(\bar{x} - y_j) dy.$$

Taking the average of both sides of this estimate over all  $x \in B(0, K)$  and shifting the  $y_j$  variables,

$$\mu_{T^\lambda f}(B_{K^2}) \lesssim K^{O(1)} \int_{(\mathbb{R}^n)^{2k}} \int_{B(\bar{x}, K)} \prod_{j=1}^k |T^\lambda f_{\tau_j^*} \circ \sigma_{y_j, z_j}(x)|^{p/k} dx Z_K(\mathbf{y}, \mathbf{z}) dy dz$$

where  $Z_K(\mathbf{y}, \mathbf{z}) := \prod_{j=1}^k w_K(y_j) \zeta_K(z_j)$  and  $\sigma_{y_j, z_j}(x) := x + y_j - z_j$ . Since both  $w_K$  and  $\zeta_K$  decay (at least) as rapidly as  $|y|^{-N}$  away from  $B(0, K)$ , one may restrict the integral in  $\mathbf{y}$  and  $\mathbf{z}$  from the whole space  $(\mathbb{R}^n)^{2k}$  to the bounded region  $B(0, \lambda/\bar{C}K)^{2k}$  at the expense of an additional harmless error term.

Given a  $K^{-1}$ -cap  $\tau$ , let  $T_\tau^\lambda$  be a Hörmander-type operator given by replacing the amplitude  $a^\lambda$  in the definition of  $T^\lambda$  with some amplitude  $a_\tau^\lambda$  which has  $\omega$ -support in a  $2K^{-1}$ -cap concentric to  $\tau$  and which satisfies  $T_\tau^\lambda f_\tau = T^\lambda f_\tau$ . One now wishes to bound

$$(6.7) \quad \int_{B(0, \lambda/\bar{C}K)^{2k}} \int_{B(\bar{x}, K)} \prod_{j=1}^k |T_{\tau_j^*}^\lambda f_{\tau_j^*} \circ \sigma_{y_j, z_j}(x)|^{p/k} dx Z_K(\mathbf{y}, \mathbf{z}) dy dz$$

For the purposes of this proof a  $k$ -tuple  $(\tau_1, \dots, \tau_k)$  of  $K^{-1}$ -caps is said to be *transverse* if  $(T_{\tau_1}^\lambda, \dots, T_{\tau_k}^\lambda)$  is a  $cK^{-(k-1)}$ -transverse  $k$ -tuple of Hörmander-type operators, for a suitable choice of small constant  $c > 0$ . If the constant  $\bar{C}$  from the original decomposition is sufficiently large, then using the estimate

$$|G^\lambda(\bar{x}, \omega) - G^\lambda(x, \omega)| \lesssim \frac{|\bar{x} - x|}{\lambda} \lesssim \bar{C}^{-1} K^{-1} \quad \text{for } (x, \omega) \in \text{supp } a,$$

it is clear that, in addition to (6.2), one may assume that  $(\tau_1^*, \dots, \tau_k^*)$  is transverse. The expression (6.7) is therefore dominated by

$$\sum_{(\tau_j)_{j=1}^k \text{ trans.}} \int_{B(0, \lambda/\bar{C}K)^{2k}} \int_{B(\bar{x}, K)} \prod_{j=1}^k |T_{\tau_j}^\lambda f_{\tau_j} \circ \sigma_{y_j, z_j}(x)|^{p/k} dx Z_K(\mathbf{y}, \mathbf{z}) dy dz,$$

where the sum is over all choices of transverse  $k$ -tuples of caps. Summing both sides of this inequality over all  $B(\bar{x}, K^2) \in \mathcal{B}_{K^2}$  with  $B(\bar{x}, K^2) \cap B_z \neq \emptyset$ , it suffices to show that

$$\int_{\mathbb{R}^n} \prod_{j=1}^k |T_{\tau_j}^\lambda f_{\tau_j} \circ \sigma_{y_j, z_j}(x)|^{p/k} dx \lesssim R^\varepsilon \|f\|_{L^2(B^{n-1})}^p$$

for any choice of  $K^{-(k-1)}$ -transverse tuple  $(\tau_1, \dots, \tau_k)$  and any  $\mathbf{y}, \mathbf{z} \in B(0, \lambda/\bar{C}K)^k$ . However, defining  $T_j^\lambda f(x) := T_{\tau_j}^\lambda f \circ \sigma_{y_j, z_j}(x)$  and again choosing  $\bar{C}$  to be sufficiently large, this estimate is an immediate consequence of the hypothesised multilinear estimate.  $\square$

## 7. ALGEBRAIC PRELIMINARIES

**7.1. Basic definitions and results.** Let  $0 \leq m \leq n$  and consider a collection of real polynomials  $P_j \in \mathbb{R}[X_1, \dots, X_n]$ ,  $1 \leq j \leq n - m$ . Let  $Z(P_1, \dots, P_{n-m})$  denote their zero locus; that is,

$$Z(P_1, \dots, P_{n-m}) := \{x \in \mathbb{R}^n : P_j(x) = 0 \text{ for } 1 \leq j \leq n - m\}.$$

A set of the above form is referred to as a *variety* and the *maximum degree* of  $Z(P_1, \dots, P_{n-m})$  is defined to be the number

$$\overline{\deg} Z(P_1, \dots, P_{n-m}) := \max_{1 \leq j \leq n-m} \deg P_j.$$

Throughout this article it will be convenient to work with varieties which satisfy the additional property that the  $n \times (n - m)$  matrix  $(\nabla P_1(z) \dots \nabla P_{n-m}(z))$  has full rank whenever  $z \in Z(P_1, \dots, P_{n-m})$ . In this case  $Z(P_1, \dots, P_{n-m})$  is said to be a *transverse complete intersection*. Clearly any transverse complete intersection is a smooth  $m$ -dimensional submanifold of  $\mathbb{R}^n$ .

For 0-dimensional transverse complete intersections the following well-known variant of the classical Bézout theorem holds (see, for instance, [10]).

**Theorem 7.1** (Bézout's theorem). *Suppose  $Z = Z(P_1, \dots, P_n)$  is a transverse complete intersection. Then  $Z$  is finite and  $\#Z \leq \prod_{j=1}^n \deg P_j$ .*

A key tool in the present analysis of Hörmander-type operators is the following polynomial partitioning result, which is a variant of the polynomial partitioning theorem introduced in [15] and is based on the classical polynomial ham sandwich theorem of Stone and Tukey [21].

**Theorem 7.2** (Polynomial partitioning [13]). *Suppose  $W \in L^1(\mathbb{R}^n)$  is non-negative. For any degree  $D \in \mathbb{N}$  there is a polynomial  $P$  of degree  $\deg P \lesssim D$  such that the following hold.*

- i) The set  $Z(P)$  is a finite union of  $\sim \log D$  transverse complete intersections.*
- ii) If  $\{O_i\}_{i \in \mathcal{I}}$  denotes the set of connected components of  $\mathbb{R}^n \setminus Z(P)$ , then  $\#\mathcal{I} \lesssim D^n$  and*

$$\int_{O_i} W \sim D^{-n} \int_{\mathbb{R}^n} W \quad \text{for all } i \in \mathcal{I}.$$

The connected components  $O_i$  of the set  $\mathbb{R}^n \setminus Z(P)$  are referred to as *cells*. It is remarked that in [13] a stronger version of the above theorem is stated and proved, which provides further structural information about the polynomial  $P$  (in particular, the full result is stable under certain small perturbations of  $P$ ). Whilst the methods of this article will require this strengthened version of Theorem 7.2, the full statement of the result is not reproduced here (it is only needed to address certain technical aspects of the analysis).

It was observed in recent work of the first author [14, 13] that polynomial partitioning is a useful tool for studying oscillatory integral operators. Roughly speaking, Theorem 7.2 can be used to effectively reduce the problem to situations where the

mass of  $T^\lambda f$  is concentrated in the neighbourhood of some low-degree algebraic variety; note that this is precisely the setup in the sharp examples discussed in §2.

**7.2. Polynomial approximation.** Recall that the operators  $T^\lambda$  are defined with respect to data belonging to the  $C^\infty$  category. In order to apply algebraic methods to the problem, one must approximate certain  $C^\infty$  objects by polynomials. This applies, in particular, to the core curves  $\Gamma_{\theta,v}^\lambda$  which appear in the definition of the wave packets in §5. Similar issues were addressed in [8, 27] via a Jackson-type approximation theorem (see, for instance, [1]); for the present purpose an entirely elementary Taylor approximation argument is all that is required.

Let  $\varepsilon > 0$  be a small parameter and define  $N = N_\varepsilon := \lceil 1/2\varepsilon \rceil \in \mathbb{N}$ . Suppose that  $\Gamma: (-1, 1) \rightarrow \mathbb{R}^n$  is a smooth curve satisfying

$$\|\Gamma\|_{C^{N+1}(-1,1)} := \max_{1 \leq k \leq N+1} \sup_{|t| < 1} |\Gamma^{(k)}(t)| \lesssim 1;$$

by the reductions made in §4, the curves  $\Gamma_{\theta,v}$  from §5 satisfy this condition. Let  $[\Gamma]_\varepsilon: \mathbb{R} \rightarrow \mathbb{R}^n$  denote the polynomial curve given by the degree  $N$  Taylor approximation of  $\Gamma$  around 0. Observe that

$$\|[\Gamma]_\varepsilon\|_{C^\infty(-2,2)} \leq e^2 \|\Gamma\|_{C^N(-1,1)} \lesssim 1.$$

Given  $\lambda \gg 1$ , noting that  $\lambda^{-\varepsilon N} \leq \lambda^{-1/2}$ , Taylor's theorem yields

$$|\Gamma^{(i)}(t) - [\Gamma]_\varepsilon^{(i)}(t)| \lesssim_\varepsilon \lambda^{-1/2} |t|^{1-i} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{-\varepsilon} \text{ and } i = 0, 1.$$

Letting  $\Gamma^\lambda: (-\lambda, \lambda) \rightarrow \mathbb{R}^n$  denote the rescaled curve  $\Gamma^\lambda(t) := \lambda\Gamma(t/\lambda)$ , the above inequalities trivially imply that

$$(7.1) \quad \|[\Gamma^\lambda]_\varepsilon'\|_{C^\infty(-2\lambda, 2\lambda)} \lesssim 1 \quad \text{and} \quad \|[\Gamma^\lambda]_\varepsilon''\|_{C^\infty(-2\lambda, 2\lambda)} \lesssim \lambda^{-1}$$

and

$$(7.2) \quad |(\Gamma^\lambda)^{(i)}(t) - ([\Gamma^\lambda]_\varepsilon)^{(i)}(t)| \lesssim_\varepsilon \lambda^{-1/2} |t|^{1-i} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{1-\varepsilon} \text{ and } i = 0, 1.$$

Combining the  $i = 1$  case of the above estimate with the elementary inequality

$$|x \wedge y| \leq |x||y| \left| \frac{x}{|x|} - \frac{y}{|y|} \right| \leq 2 \min\{|x|, |y|\} |x - y| \quad \text{for all } x, y \in \mathbb{R}^n \setminus \{0\},$$

one observes that the tangent spaces to the curves  $\Gamma^\lambda$  and  $[\Gamma^\lambda]_\varepsilon$  have a small angular separation; more precisely,

$$(7.3) \quad \angle(T_{\Gamma^\lambda(t)}\Gamma^\lambda, T_{[\Gamma^\lambda]_\varepsilon(t)}[\Gamma^\lambda]_\varepsilon) \lesssim_\varepsilon \lambda^{-1/2} \quad \text{for all } |t| \lesssim_\varepsilon \lambda^{1-\varepsilon}.$$

**7.3. Transverse interactions between curved tubes and varieties.** Let  $Z = Z(P_1, \dots, P_{n-m})$  be a transverse complete intersection and fix a polynomial curve  $\Gamma: \mathbb{R} \rightarrow \mathbb{R}^n$ . The purpose of this subsection is to study transverse interactions between  $\Gamma$  and an  $r$ -neighbourhood of  $Z$ ; that is, roughly speaking, one wishes to understand the set of points at which the curve  $\Gamma$  enters  $\mathcal{N}_r Z$  at a large angle. More precisely, given  $\alpha, r > 0$  the problem is to estimate the size of the set

$$Z_{>\alpha, r, \Gamma} := \{z \in Z : \exists x \in \Gamma \text{ with } |x - z| < r \text{ and } \angle(T_z Z, T_x \Gamma) > \alpha\}.$$

It will be convenient to assume that  $\Gamma$  is a *polynomial graph*, by which it is meant that the curve can be rotated so that it is given by  $\Gamma(t) = (\gamma(t), t)$  for some polynomial mapping  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^{n-1}$ .

**Lemma 7.3.** *Let  $n \geq 2$ ,  $1 \leq m \leq n$  and  $Z = Z(P_1, \dots, P_{n-m}) \subseteq \mathbb{R}^n$  be a transverse complete intersection. Suppose  $\Gamma: \mathbb{R} \rightarrow \mathbb{R}^n$  is a polynomial graph satisfying*

$$(7.4) \quad \|\Gamma'\|_{L^\infty(-2\lambda, 2\lambda)} \lesssim 1 \quad \text{and} \quad \|\Gamma''\|_{L^\infty(-2\lambda, 2\lambda)} \leq \delta$$

for some  $\lambda, \delta > 0$ . There exists a dimensional constant  $\bar{C} > 0$  such that for all  $\alpha > 0$  and  $0 < r < \lambda$  satisfying  $\alpha \geq \bar{C}\delta r$  the set  $Z_{>\alpha, r, \Gamma} \cap B(0, \lambda)$  is contained in a union of  $O((\overline{\deg} Z \cdot \deg \Gamma)^n)$  balls of radius  $r/\alpha$ .

The case of interest is given by taking  $\Gamma := [\Gamma_{\theta, v}^\lambda]_\varepsilon$  to be the polynomial approximant of the curve  $\Gamma_{\theta, v}^\lambda$  introduced in the previous subsection. Here  $\deg \Gamma \lesssim_\varepsilon 1$  and, by (7.1), the condition (7.4) holds with  $\delta \sim_\varepsilon 1/\lambda$ ; thus, Lemma 7.3 implies that for  $\alpha > 0$  and  $0 < r < \lambda$  satisfying  $\alpha \gtrsim r/\lambda$ , the set  $Z_{>\alpha, r, \Gamma} \cap B(0, \lambda)$  is contained in a union of  $O_\varepsilon((\overline{\deg} Z)^n)$  balls of radius  $r/\alpha$ .

Using Bézout's theorem (that is, Theorem 7.1), Lemma 7.3 was established in the case where  $\Gamma$  is a line by the first author in [Lemma 5.7] [13]. If  $\Gamma = \ell$  is a line, then the condition (7.4) holds for any  $\lambda > 0$  and any  $\delta > 0$  and therefore the lemma implies that for any  $\alpha, r > 0$  the set  $Z_{>\alpha, r, \ell}$  is contained in a union of  $O((\overline{\deg} Z)^n)$  balls of radius  $r/\alpha$ . The result for general curves  $\Gamma$  is, in fact, a rather straight-forward consequence of the special case of lines.

*Proof (of Lemma 7.3).* Since the problem is rotationally invariant, one may assume that  $\Gamma(t) = (\gamma(t), t)$  where  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^{n-1}$  is a polynomial mapping.

The function  $\Upsilon: \mathbb{R}^n \rightarrow \mathbb{R}^n$  given by  $\Upsilon(x', x_n) := (x' - \gamma(x_n), x_n)$  is clearly a diffeomorphism which maps bijectively between  $\Gamma$  and the vertical line  $\ell = \text{span}\{e_n\}$ . Furthermore, it easily follows that the image set

$$\tilde{Z} := \Upsilon(Z(P_1, \dots, P_{n-m})) = Z(P_1 \circ \Upsilon^{-1}, \dots, P_{n-m} \circ \Upsilon^{-1})$$

is a transverse complete intersection of maximum degree  $\overline{\deg} \tilde{Z} \leq \overline{\deg} Z \cdot \deg \Gamma$ .

Let  $\lambda, \alpha, r$  satisfy the hypotheses of the lemma for some suitably large dimensional constant  $\bar{C} \geq 1$ . The key observation is as follows.

**Claim.** *There exist dimensional constants  $0 < c \leq 1$  and  $C \geq 1$  such that*

$$Z_{>\alpha, r, \Gamma} \cap (\mathbb{R}^{n-1} \times (-\lambda, \lambda)) \subseteq \Upsilon^{-1}(\tilde{Z}_{>c\alpha, Cr, \ell}).$$

Once this claim is verified, Lemma 7.3 easily follows. Indeed, one may apply the special case of Lemma 7.3 for lines (which, as previously remarked, is proved in [Lemma 5.7] [13]) to conclude that  $\tilde{Z}_{>c\alpha, Cr, \ell}$  is contained in a union of  $O((\overline{\deg} Z \cdot \deg \Gamma)^n)$  balls of radius  $r/\alpha$ . On the other hand, as a consequence of the first hypothesis in (7.4),

$$(7.5) \quad |\Upsilon(x) - \Upsilon(x')| \sim |x - x'| \quad \text{for all } x, x' \in \mathbb{R}^{n-1} \times (-\lambda, \lambda).$$

Combining these observations, it follows that the set  $Z_{>\alpha, r, \Gamma} \cap B(0, \lambda)$  can be covered by  $O((\overline{\deg} Z \cdot \deg \Gamma)^n)$  balls of radius  $r/\alpha$ , as required.

Turning to the proof of the claim, let  $z \in Z_{>\alpha, r, \Gamma} \cap B(0, \lambda)$  and note that there exists some  $x = \Gamma(x_n) \in \Gamma$  with  $|x - z| < r$  and  $\angle(T_z Z, T_x \Gamma) > \alpha$ . Defining  $\tilde{z} := \Upsilon(z) \in \tilde{Z}$  and  $\tilde{x} := \Upsilon(x) \in \ell$ , it follows from (7.5) that  $|\tilde{x} - \tilde{z}| \lesssim r$ . Thus, the problem is reduced to showing that  $\angle(T_{\tilde{z}} \tilde{Z}, e_n) = \angle(T_{\tilde{z}} \tilde{Z}, T_{\tilde{x}} \ell) \gtrsim \alpha$ .

Observe that, provided  $\bar{C}$  is sufficiently large depending only on  $n$ ,

$$(7.6) \quad \angle(T_z Z, T_{\Gamma(z_n)} \Gamma) > \alpha/2.$$

Indeed,  $|z_n| < \lambda$  and  $|x_n| < \lambda + r < 2\lambda$  and so, by the second condition in (7.4),

$$|\Gamma'(x_n) - \Gamma'(z_n)| \leq \delta|x_n - z_n| < \delta r < \bar{C}^{-1}\alpha.$$

Thus, if  $\bar{C}$  is appropriately chosen, then  $\sphericalangle(T_{\Gamma(x_n)}\Gamma, T_{\Gamma(z_n)}\Gamma) < \alpha/2$ , which immediately yields (7.6).

Combining the observations of the previous paragraphs, the claim follows provided one can show that  $\sphericalangle(T_{\tilde{z}}\tilde{Z}, e_n) \sim \sphericalangle(T_z Z, T_{\Gamma(z_n)}\Gamma)$ . Let  $\eta$  be a smooth curve in  $Z$  containing  $z$  and define  $\tilde{\eta} := \Upsilon(\eta)$ ; thus,  $\tilde{\eta}$  is a smooth curve in  $\tilde{Z}$  containing  $\tilde{z}$ . The problem is now reduced to proving that

$$(7.7) \quad \sphericalangle(T_{\tilde{z}}\tilde{\eta}, e_n) \sim \sphericalangle(T_z\eta, T_{\Gamma(z_n)}\Gamma).$$

If  $T_z\eta$  lies in the hyperplane  $e_n^\perp$  orthogonal to  $e_n$ , then the above estimate easily follows. Indeed, the tangent space  $T_{\Gamma(z_n)}\Gamma$  is spanned by  $\Gamma'(z_n) = (\gamma'(z_n), 1)$  and therefore, by (7.4), one has  $\sphericalangle(T_z\eta, T_{\Gamma(z_n)}\Gamma) \sim 1$ . On the other hand, it is clear from the definition of  $\Upsilon$  that  $T_{\tilde{z}}\tilde{\eta}$  also lies in  $e_n^\perp$  and so  $\sphericalangle(T_{\tilde{z}}\tilde{\eta}, e_n) = \pi/2$ . Thus, (7.7) holds in this case.

If  $T_z\eta$  does not lie in the hyperplane  $e_n^\perp$ , then  $\eta$  can be locally parametrised as a graph over the  $x_n$ -variable. By an abuse of notation, let  $\eta$  denote this graph parametrisation and  $\tilde{\eta} := \Upsilon \circ \eta$  so that  $\eta(z_n) = z$  and  $\tilde{\eta}(z_n) = \tilde{z}$ . One may easily verify that  $|\tilde{\eta}'(z_n) \wedge e_n| = |\eta'(z_n) \wedge \Gamma'(z_n)|$  and so

$$\sin \sphericalangle(T_{\tilde{z}}\tilde{\eta}, e_n)|\tilde{\eta}'(z_n)| = \sin \sphericalangle(T_z\eta, T_{\Gamma(z_n)}\Gamma)|\eta'(z_n)||\Gamma'(z_n)|.$$

By the first hypothesis in (7.4), one has  $|\tilde{\eta}'(z_n)| \sim |\eta'(z_n)|$  and  $|\Gamma'(z_n)| \sim 1$ , and (7.7) follows.  $\square$

## 8. TRANSVERSE EQUIDISTRIBUTION ESTIMATES

**8.1. Tangential wave packets and transverse equidistribution.** In this section the theory of transverse equidistribution estimates, as introduced in [13], is extended to the variable coefficient setting. This is a key step in the proof of Theorem 1.6 and here the positive-definite hypothesis H2<sup>+</sup>) plays a crucial rôle in the argument.

The first step is to give a precise definition of what it means for a wave packet to be ‘tangential’ to a transverse complete intersection  $Z$ . Throughout this section let  $T^\lambda$  be a Hormander-type operator with reduced phase  $\phi$  and for some  $R \ll \lambda$  define the tubes  $T_{\theta,v}$  as in §5. Furthermore, let  $\delta_m$  denote a small parameter satisfying  $0 < \delta \ll \delta_m \ll 1$  (here  $\delta$  is the same parameter as that which appears in the definition of the wave packets).

**Definition 8.1.** Suppose  $Z = Z(P_1, \dots, P_{n-m})$  is a transverse complete intersection. A tube  $T_{\theta,v}$  is  $R^{-1/2+\delta_m}$ -tangent to  $Z$  in  $B(0, R)$  if

$$(8.1) \quad T_{\theta,v} \subseteq N_{R^{1/2+\delta_m}}(Z)$$

and

$$(8.2) \quad \sphericalangle(G^\lambda(x; \omega_\theta), T_z Z) \leq \bar{c}_{\text{tang}} R^{-1/2+\delta_m}$$

for any  $x \in T_{\theta,v}$  and  $z \in Z \cap B(0, 2R)$  with  $|x - z| \leq \bar{C}_{\text{tang}} R^{1/2+\delta_m}$ .

Here  $\bar{c}_{\text{tang}} > 0$  (respectively,  $\bar{C}_{\text{tang}} \geq 1$ ) is a dimensional constant, chosen to be sufficiently small (respectively, large) for the purposes of the following arguments.

**Definition 8.2.** If  $\mathbb{S} \subseteq \mathbb{T}$ , then  $f$  is said to be concentrated on wave packets from  $\mathbb{S}$  if

$$f = \sum_{(\theta, v) \in \mathbb{S}} f_{\theta, v} + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

One wishes to study functions concentrated on wave packets from the collection

$$\mathbb{T}_Z := \{(\theta, v) \in \mathbb{T} : T_{\theta, v} \text{ is } R^{-1/2+\delta_m}\text{-tangent to } Z \text{ in } B(0, R)\}.$$

Let  $B \subseteq \mathbb{R}^n$  be a fixed ball of radius  $R^{1/2+\delta_m}$  with centre  $\bar{x} \in B(0, R)$ . Throughout this section the analysis will be essentially confined to a spatially localised operator  $\eta_B \cdot T^\lambda g$  where  $\eta_B$  is a suitable choice of Schwartz function concentrated on  $B$ . For any  $(\theta, v)$  the frequencies of  $\eta_B \cdot T^\lambda g_{\theta, v}$  are concentrated near the surface

$$(8.3) \quad \Sigma := \{\Sigma(\omega) : \omega \in \Omega\} \quad \text{where} \quad \Sigma(\omega) := \partial_x \phi^\lambda(\bar{x}; \omega).$$

Now consider the refined set of wave packets

$$\mathbb{T}_{Z, B} := \{(\theta, v) \in \mathbb{T}_Z : T_{\theta, v} \cap B \neq \emptyset\}.$$

If  $(\theta, v) \in \mathbb{T}_{Z, B}$ , then the direction  $G^\lambda(\bar{x}; \omega_\theta)$  of the curved tube  $T_{\theta, v}$  on the ball  $B$  must make a small angle with each of the tangent spaces  $T_z Z$  for all  $z \in Z \cap B$ . It transpires that this essentially constrains the frequency  $\Sigma(\omega_\theta)$  to lie in a small neighbourhood of some fixed (depending on the choice of ball  $B$ )  $m$ -dimensional manifold  $S_\xi$  (here  $m = \dim Z$ ). In the case of the parabolic extension operator  $E_{\text{par}}$ , which is studied in [13], the relationship between the normal direction  $G^\lambda(\bar{x}; \omega_\theta)$  and the frequency  $\Sigma(\omega_\theta)$  is particularly simple. Here  $\Sigma(\omega_\theta) = (\omega_\theta, \frac{|\omega_\theta|^2}{2})$  is constrained to lie in roughly the  $R^{-1/2}$ -neighbourhood some *affine* subspace  $A_\xi$ . The uncertainty principle then suggests that if  $g$  concentrated on wave packets from  $\mathbb{T}_{Z, B}$ , then the function  $|E_{\text{par}}g(x)|$  is morally constant as one varies  $x$  by  $R^{1/2}$  in directions perpendicular to  $A_\xi$ . Furthermore, it can be shown that the affine subspace  $A_\xi$  makes a small angle with the tangent planes  $T_z Z$  for  $z \in Z \cap B$  and so  $|E_{\text{par}}g(x)|$  is morally constant as one varies  $x$  by  $R^{1/2}$  in directions transverse to  $Z \cap B$ .

One wishes to extend the above observations for  $E_{\text{par}}$  to the variable coefficient setting; that is, for  $g$  concentrated on wave packets from  $\mathbb{T}_{Z, B}$ ,<sup>11</sup> the problem is to show that  $|T^\lambda g|$  is morally constant in directions transverse to  $Z \cap B$ . More precisely, one wishes to establish an inequality roughly of the form

$$(8.4) \quad \int_{N_{\rho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim \int_B |T^\lambda g|^2$$

for  $0 < \rho < R$ ; this would show that the  $L^2$  mass of  $T^\lambda g$  is unable to concentrate in a small neighbourhood of  $Z \cap B$ . For the parabolic extension operator the observations of the previous paragraph can be used to prove (8.4) (up to a rapidly decaying error term). The general case is more complicated, however. First of all, the surface  $S_\xi$  described above is no longer necessarily an affine subspace and may possess curvature. One way to circumvent this issue is to introduce a further constraint on the family of wave packets. Let  $R^{1/2} < \rho \ll R$  and throughout this

<sup>11</sup>In fact, in the general case a more stringent hypothesis on  $g$  is required, as discussed below.

section let  $\tau \subset \mathbb{R}^{n-1}$  be a cap of radius  $O(\rho^{-1/2+\delta_m})$  centred at a point in  $B^{n-1}$ . Now define

$$\mathbb{T}_{Z,B,\tau} := \{(\theta, v) \in \mathbb{T}_Z : \theta \cap \tau \neq \emptyset \text{ and } T_{\theta,v} \cap B \neq \emptyset\}.$$

The frequencies  $\Sigma(\omega_\theta)$  for  $(\theta, v) \in \mathbb{T}_{Z,B,\tau}$  are further constrained to lie in a small region of  $\Sigma$  upon which the curved space  $S_\xi$  can be effectively approximated by an affine space  $A_\xi$ . Consequently, one can carry out a similar analysis as in the parabolic extension case.

The second issue is to ensure that the resulting affine space  $A_\xi$  makes a small angle with the tangent spaces  $T_z Z$  for  $z \in Z \cap B$ . This is crucial to ensure that the morally constant property holds in directions transverse to  $Z$ . For general Hörmander-type operators this property can fail (a simple example is given by the extension operator associated to the hyperbolic paraboloid, as discussed below). In order to ensure the angle condition one needs to exploit the additional positive-definite hypothesis  $H2^+$ .

In practice, the rigorous formulation of these heuristics is somewhat messier than (8.4), and it is convenient to state the key estimate in the following manner.

**Lemma 8.3.** *With the above setup, if  $\overline{\dim} Z \lesssim_\varepsilon 1$  and  $g$  is concentrated on wave packets from  $\mathbb{T}_{Z,B,\tau}$ , then*

$$(8.5) \quad \int_{N_{\rho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim R^{1/2+O(\delta_m)} (\rho/R)^{(n-m)/2} \|g\|_{L^2(B^{n-1})}^2.$$

The inequality (8.5) is related to the heuristic inequality (8.4) via Hörmander's  $L^2$  bound

$$\|T^\lambda g\|_{L^2(B)}^2 \lesssim R^{1/2+\delta_m} \|g\|_{L^2(B^{n-1})}^2.$$

The estimate is presented in this way (rather than in a form more closely resembling (8.4)) as it provides a relatively clean statement and, moreover, (8.5) happens to be the precise bound required later in the proof.

**8.2. Uncertainty principle preliminaries.** If  $G: \mathbb{R}^n \rightarrow \mathbb{C}$  is frequency supported on a ball of radius  $r > 0$ , then the uncertainty principle dictates that  $G$  should be essentially constant at spatial scale  $r^{-1}$ . In particular, the  $L^2$ -mass of  $G$  cannot be highly concentrated in any ball of radius  $\rho < r^{-1}$  and so one has

$$\int_{B(x_0, \rho)} |G|^2 \lesssim \int_{B(x_0, r^{-1})} |G|^2.$$

Strictly speaking, for this inequality to hold the right-hand integral should be taken with respect to a rapidly decaying weight function rather than over the compact region  $B(x_0, r^{-1})$  (see, for instance, [13, Section 6]). There is a variant of this estimate which is effective in cases where  $G$  has the property that  $\hat{G}$  is merely concentrated in (rather than supported in) an  $r$ -ball.

**Lemma 8.4.** *If  $r^{-1/2} \leq \rho \leq r^{-1}$ , then for any ball  $B(x_0, \rho)$ ,  $\xi_0 \in \hat{\mathbb{R}}^n$  and  $\delta > 0$  one has*

$$\int_{B(x_0, \rho)} |G|^2 \lesssim_\delta \|\hat{G} w_{B(\xi_0, r)}^{-1}\|_\infty^{2\delta/(1+\delta)} \frac{1}{|B(0, r^{-1})|} \left( \int_{\mathbb{R}^n} |G|^2 \right)^{1/(1+\delta)}.$$

Here  $w_{B(\xi_0, r)}$  is a weight concentrated on  $B(\xi_0, r)$  given by

$$(8.6) \quad w_{B(\xi_0, r)}(\xi) := (1 + r^{-1}|\xi - \xi_0|)^{-N}$$

for some large  $N = N_\delta \in \mathbb{N}$ .

Hence, if  $\hat{G}$  is concentrated in  $B(\xi_0, r)$  in the sense that  $|\hat{G}(\xi)| \lesssim M w_{B(\xi_0, r)}(\xi)$  for some controllable constant  $M \geq 0$ , then the lemma produces a favourable estimate.

*Proof (of Lemma 8.4).* Define  $\psi(x) := \psi_0(\rho^{-1}(x - x_0))$  where  $\psi_0$  is a Schwartz function which satisfies  $|\psi_0(x)| \gtrsim 1$  on  $B(0, 1)$ . Thus, by Plancherel,

$$\int_{B(x_0, \rho)} |G|^2 \lesssim \int |\psi G|^2 = \int |\hat{\psi} * \hat{G}|^2.$$

Using the rapid decay of  $\hat{\psi}$  and Hölder's inequality, one deduces that

$$|\hat{\psi} * \hat{G}(\xi)| \lesssim \rho^n \int_{\mathbb{R}^n} w_{B(\xi, \rho^{-1})}(\eta)^{\delta/(1+\delta)} |\hat{G}(\eta)| d\eta.$$

for all  $\xi \in \mathbb{R}^n$ . By expressing the right-hand integral as

$$\int_{\mathbb{R}^n} \left( w_{B(\xi, \rho^{-1})}(\eta) w_{B(\xi_0, r)}(\eta)^{1/2} \right)^{\delta/(1+\delta)} \left( |\hat{G}(\eta)|^{1+\delta} w_{B(\xi_0, r)}(\eta)^{-\delta/2} \right)^{1/(1+\delta)} d\eta$$

and applying Hölder's inequality, it follows that

$$|\hat{\psi} * \hat{G}(\xi)| \lesssim |B(x_0, \rho)| \cdot \text{I}(\xi)^{\delta/(1+\delta)} \cdot \text{II}(\xi)^{1/(1+\delta)}$$

where

$$\text{I}(\xi) := \int_{\mathbb{R}^n} w_{B(\xi, \rho^{-1})}(\eta) w_{B(\xi_0, r)}(\eta)^{1/2} d\eta,$$

$$\text{II}(\xi) := \int_{\mathbb{R}^n} |\hat{G}(\eta)|^{1+\delta} w_{B(\xi_0, r)}(\eta)^{-\delta/2} d\eta.$$

To estimate  $\text{I}(\xi)$  first perform the variable shift  $\eta \mapsto \eta + \xi_0$  and then decompose the range of integration into the regions  $|\eta| < |\xi - \xi_0|/2$  and  $|\eta| \geq |\xi - \xi_0|/2$ . Since  $w_{B(\xi - \xi_0, \rho^{-1})}(\eta) \lesssim_\delta w_{B(\xi_0, \rho^{-1})}(\xi)$  for  $|\eta| < |\xi - \xi_0|/2$  and  $w_{B(0, \rho^{-1})}(\eta) \lesssim_\delta w_{B(\xi_0, \rho^{-1})}(\xi)$  for  $|\eta| \geq |\xi - \xi_0|/2$ , it follows that

$$\text{I}(\xi) \lesssim_\delta |B(x_0, \rho)|^{-1} w_{B(\xi_0, \rho^{-1})}(\xi)^{1/2}.$$

To estimate  $\text{II}(\xi)$  note that, provided  $N_\delta$  is chosen sufficiently large, by Cauchy-Schwarz and Plancherel's theorem one has

$$\begin{aligned} \text{II}(\xi) &\leq \|\hat{G} w_{B(\xi_0, r)}^{-1}\|_\infty^\delta \int |\hat{G}(\eta)| w_{B(\xi_0, r)}(\eta)^{\delta/2} d\eta \\ &\lesssim |B(0, r^{-1})|^{-1/2} \|\hat{G} w_{B(\xi_0, r)}^{-1}\|_\infty^\delta \|G\|_{L^2(\mathbb{R}^n)}. \end{aligned}$$

Combining these observations, one obtains the desired estimate but with an additional factor of  $(\rho r^{1/2})^{-2n\delta/(1+\delta)}$  on the right-hand side. Since  $1 \leq \rho r^{1/2}$ , the result immediately follows.  $\square$

**8.3. Wave packets tangential to linear subspaces.** Here, as a step towards Lemma 8.3, transverse equidistribution estimates are proven for functions concentrated on wave packets tangential to some fixed linear subspace  $V \subseteq \mathbb{R}^n$ . As before, let  $B$  be a ball of radius  $R^{1/2+\delta_m}$  with centre  $\bar{x} \in \mathbb{R}^n$  and define

$$\mathbb{T}_{V, B} := \{(\theta, v) : \not\perp(G^\lambda(\bar{x}, \omega_\theta), V) \lesssim R^{-1/2+\delta_m} \text{ and } T_{\theta, v} \cap B \neq \emptyset\}.$$

Let  $R^{1/2} < \rho < R$  and for  $\tau \subset \mathbb{R}^{n-1}$  a ball of radius  $O(\rho^{-1/2+\delta_m})$  centred at a point in  $B^{n-1}$  define

$$\mathbb{T}_{V, B, \tau} := \{(\theta, v) \in \mathbb{T}_{V, B} : \theta \cap (\frac{1}{10} \cdot \tau) \neq \emptyset\}$$

where  $(\frac{1}{10} \cdot \tau)$  is the cap concentric to  $\tau$  but with 1/10th of the radius.

The key estimate is the following.

**Lemma 8.5.** *If  $V \subseteq \mathbb{R}^n$  is a linear subspace, then there exists a linear subspace  $V'$  with the following properties:*

- 1)  $\dim V + \dim V' = n$ .
- 2)  $V, V'$  are quantitatively transverse in the sense that there exists a uniform constant  $c_{\text{trans}} > 0$  such that

$$\sphericalangle(v, v') \geq 2c_{\text{trans}} \quad \text{for all non-zero vectors } v \in V \text{ and } v' \in V'.$$

- 3) If  $g$  is concentrated on wave packets from  $\mathbb{T}_{V, B, \tau}$ ,  $\Pi$  is any plane parallel to  $V'$  and  $x_0 \in \Pi \cap B$ , then the inequality

$$\int_{\Pi \cap B(x_0, \rho^{1/2+\delta_m})} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)} (\rho/R)^{\dim V'/2} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left( \int_{\Pi \cap 2B} |T^\lambda g|^2 \right)^{1/(1+\delta)}$$

holds up to the inclusion of a  $\text{RapDec}(R)\|g\|_{L^2(B^{n-1})}$  term on the right-hand side.

*Proof (of Lemma 8.5).* The argument is presented in a number of stages.

**Constructing the subspace  $V'$ .** Recall that  $\omega \mapsto G^\lambda(\bar{x}; \omega) := \frac{G_0^\lambda(\bar{x}; \omega)}{|G_0^\lambda(\bar{x}; \omega)|}$  is the Gauss map associated to the hypersurface  $\Sigma$ , defined in (8.3). Since  $G(x; 0) = e_n$  for all  $x \in X$ , Lemma 4.4 implies that

$$\sphericalangle(G^\lambda(\bar{x}; \omega), e_n) \sim |\omega| \quad \text{for all } \omega \in \Omega.$$

Consequently, by choosing  $\text{diam } \Omega$  to be sufficiently small in the initial reductions, one may assume that the Gauss map  $\omega \mapsto G^\lambda(\bar{x}; \omega)$  always makes a wide angle with the hyperplane  $e_n^\perp = \mathbb{R}^{n-1} \times \{0\}$ . In particular,

$$\sphericalangle(G^\lambda(\bar{x}; \omega), e_n^\perp) \gtrsim 1 \quad \text{for all } \omega \in \Omega.$$

Since the situation is trivial if  $\mathbb{T}_{V, B} = \emptyset$ , one may assume that

$$(8.7) \quad \overline{\sphericalangle}(V, e_n^\perp) := \max_{v \in V \cap S^{n-1}} \sphericalangle(v, e_n^\perp) \gtrsim 1.$$

Define  $S_\omega \subset \mathbb{R}^{n-1}$  by

$$S_\omega := \{\omega \in \Omega : G^\lambda(\bar{x}; \omega) \in V\}.$$

Fixing an orthonormal basis  $\{N_1, \dots, N_{n-\dim V}\}$  for  $V^\perp$ , one has

$$S_\omega = \{\omega \in \Omega : \langle G_0^\lambda(\bar{x}; \omega), N_k \rangle = 0 \text{ for } 1 \leq k \leq n - \dim V\}.$$

**Claim.** *If  $S_\omega \neq \emptyset$ , then  $S_\omega$  a smooth surface in  $\mathbb{R}^{n-1}$  of dimension  $\dim V - 1$ .*

*Proof.* Let  $\omega \in S_\omega$  and note that each  $N_k$  is tangential to  $\Sigma$  at  $\Sigma(\omega)$ . Hence, one may write

$$N_k = \sum_{j=1}^{n-1} N_k^{(j)}(\omega) \partial_{\omega_j} \partial_x \phi^\lambda(\bar{x}; \omega)$$

for some choice of coefficients  $N_k^{(j)}(\omega) \in \mathbb{R}$ . A computation now shows that

$$\partial_{\omega_i} \langle G_0^\lambda(\bar{x}; \omega), N_k \rangle = - \sum_{j=1}^{n-1} \langle \partial_{\omega_i \omega_j}^2 \partial_x \phi^\lambda(\bar{x}; \omega), G_0^\lambda(\bar{x}; \omega) \rangle N_k^{(j)}(\omega).$$

The condition H2) implies the invertibility of the  $(n-1) \times (n-1)$  matrix whose  $(i, j)$ th entry is given by

$$\langle \partial_{\omega_i \omega_j}^2 \partial_x \phi^\lambda(\bar{x}; \omega), G_0^\lambda(\bar{x}; \omega) \rangle.$$

Thus, the Jacobian of  $\omega \mapsto (\langle G_0^\lambda(\bar{x}; \omega), N_k \rangle)_{k=1}^{n-\dim V}$  has maximal rank, and the claim follows by the implicit function theorem.  $\square$

At this point it is convenient to switch to a graph parametrisation of  $\Sigma$  via the change of variables  $u \mapsto \Psi^\lambda(\bar{x}; u)$  where  $\Psi^\lambda$  is the (appropriate  $\lambda$ -rescaling of the) function introduced in §4. For convenience, let  $\Psi: U \rightarrow \Omega$  denote this mapping; that is,  $\Psi(u) := \Psi^\lambda(\bar{x}; u)$ . Recall that the hypersurface  $\Sigma$  coincides with the graph of the function

$$(8.8) \quad \bar{h}: U \rightarrow \mathbb{R}, \quad \bar{h}(u) := \partial_{x_n} \phi^\lambda(\bar{x}; \Psi(u)).$$

If  $S_\omega \cap \tau = \emptyset$ , then it follows that

$$\not\prec(G^\lambda(\bar{x}; \theta), V) \gtrsim \rho^{-1/2+\delta_m} > R^{-1/2+\delta_m}$$

whenever  $\theta \cap (\frac{1}{10} \cdot \tau) \neq \emptyset$ . Consequently,  $\mathbb{T}_{B, \tau, V} = \emptyset$  and the situation is trivial. Thus, one may assume without loss of generality that  $S_\omega \cap \tau \neq \emptyset$  and so, letting

$$S_u := \Psi^{-1}(S_\omega) = \{u \in U : G_0^\lambda(\bar{x}; \Psi(u)) \in V\},$$

it follows that  $S_u \cap \Psi^{-1}(\tau) \neq \emptyset$ . The properties of the mapping  $\Psi$  discussed in §4 imply that  $\Psi^{-1}(\tau)$  is roughly a ball of radius  $O(\rho^{-1/2+\delta_m})$ .

Fix some  $u_0 \in S_u \cap \Psi^{-1}(\tau)$  and let  $A_u$  be the tangent plane to  $S_u$  at  $u_0$ . This is interpreted as a  $(\dim V - 1)$ -dimensional affine subspace of  $\mathbb{R}^{n-1}$  through  $u_0$ . Now define  $A_\xi := A_u \times \mathbb{R} \subseteq \mathbb{R}^n$  so that  $\dim V_\xi = \dim V$  and let  $V_u$  and  $V_\xi$  be the linear subspaces parallel to  $A_u$  and  $A_\xi$ , respectively. Finally, let  $V' := V_\xi^\perp$  so that  $\dim V + \dim V' = n$ .

**Verifying the transverse equidistribution estimate in 3).** Suppose  $\Pi \subseteq \mathbb{R}^n$  is an affine subspace parallel to  $V'$  which intersects  $B$  and  $x_0 \in \Pi \cap B$ . Let  $\eta_B(x) := \eta((x - \bar{x})/R^{1/2+\delta_m})$  where  $\eta$  is a Schwartz function which satisfies  $\eta(x) = 1$  for  $x \in B(0, 2)$  and, for any  $(\theta, v) \in \mathbb{T}$ , consider

$$(\eta_B \cdot T^\lambda g_{\theta, v}|_\Pi)^\wedge(\xi) = e^{-2\pi i \langle x_0, \xi \rangle} R^{\dim V'(1/2+\delta_m)} \int_{B^{n-1}} K^\lambda(\xi; \omega) g_{\theta, v}(\omega) d\omega$$

where  $K^\lambda(\xi; \omega)$  is given by

$$\int_{V'} e^{2\pi i (\phi^\lambda(x_0 + R^{1/2+\delta_m} z; \omega) - R^{1/2+\delta_m} \langle z, \xi \rangle)} a^\lambda(x_0 + R^{1/2+\delta_m} z; \omega) \tilde{\eta}(z) dz$$

for  $\tilde{\eta}(z) := \eta(z + (x_0 - \bar{x})/R^{1/2+\delta_m})$ . The  $z$ -gradient of the phase in the above oscillatory integral is equal to

$$R^{1/2+\delta_m} [(\xi - \text{proj}_{V'} \Sigma(\omega)) + \text{proj}_{V'} [(\partial_x \phi^\lambda)(\bar{x}; \omega) - (\partial_x \phi^\lambda)(x_0 + R^{1/2+\delta_m} z; \omega)]]$$

where the second term satisfies

$$|\text{proj}_{V'} [(\partial_x \phi^\lambda)(\bar{x}; \omega) - (\partial_x \phi^\lambda)(x_0 + R^{1/2+\delta_m} z; \omega)]| \lesssim R^{1/2+\delta_m} / \lambda \ll R^{-1/2+\delta_m}.$$

Repeated integration-by-parts now shows that  $K^\lambda$  is rapidly decaying whenever  $|\xi - \text{proj}_{V'} \Sigma(\omega)| \gtrsim R^{-1/2+\delta_m}$  and, moreover,

$$|K^\lambda(\xi; \omega)| \lesssim_N (1 + R^{1/2} |\xi - \text{proj}_{V'} \Sigma(\omega)|)^{-N}.$$

If  $\omega \in \text{supp } g_{\theta, v}$ , then  $|\omega - \omega_{\theta}| < R^{-1/2}$  and so  $|\Sigma(\omega) - \xi_{\theta}| \lesssim R^{-1/2}$  where  $\xi_{\theta} := \Sigma(\omega_{\theta})$ . Consequently,

$$(8.9) \quad |(\eta_B \cdot T^{\lambda} g_{\theta, v}|_{\Pi})^{\wedge}(\xi)| \lesssim_N R^{O(1)} w_{B(\text{proj}_{V'} \xi_{\theta}, R^{-1/2})}(\xi) \|g_{\theta, v}\|_{L^2(B^{n-1})}$$

where the definition of the weight function

$$w_{B(\text{proj}_{V'} \xi_{\theta}, R^{-1/2})}(\xi) := (1 + R^{1/2} |\xi - \xi_{\theta}|)^{-N}$$

agrees with that of Lemma 8.4 (although here the weights are thought of as functions on  $V'$ ), and so  $N = N_{\delta}$  is a large integer, depending on  $\delta$ .

The following geometric observation is key to the proof of property 3).

**Claim.** *If  $(\theta, v) \in \mathbb{T}_{B, \tau, V}$ , then  $\text{dist}(\xi_{\theta}, A_{\xi}) \lesssim R^{-1/2 + \delta_m}$ .*

Temporarily assume this claim and recall that  $V' := V_{\xi}^{\perp}$  where  $V_{\xi}$  is the linear subspace parallel to the affine subspace  $A_{\xi}$ . Thus, if  $(\theta, v) \in \mathbb{T}_{B, \tau, V}$ , then  $\text{proj}_{V'} \xi_{\theta}$  lies in some fixed ball of radius  $O(R^{-1/2 + \delta_m})$ . Letting  $\xi_{*} \in V'$  denote the centre of this ball, it follows that  $w_{B(\text{proj}_{V'} \xi_{\theta}, R^{-1/2})} \lesssim_{\delta} w_{B(\xi_{*}, R^{-1/2 + \delta_m})}$  and so

$$\sum_{(\theta, v) \in \mathbb{T}_{B, \tau, V}} w_{B(\text{proj}_{V'} \xi_{\theta}, R^{-1/2})} \lesssim_{\delta} R^{O(1)} w_{B(\xi_{*}, R^{-1/2 + \delta_m})}.$$

Recalling (8.9),

$$\|(\eta_B \cdot T^{\lambda} g|_{\Pi})^{\wedge} w_{B(\xi_{*}, R^{-1/2 + \delta_m})}^{-1}\|_{\infty} \lesssim R^{O(1)} \|g\|_{L^2(B^{n-1})}$$

and, applying Lemma 8.4, one concludes that

$$\int_{B(x_0, \rho^{1/2 + \delta_m}) \cap \Pi} |T^{\lambda} g|^2 \lesssim_{\delta} R^{O(\delta_m)} \left(\frac{\rho^{1/2}}{R^{1/2}}\right)^{-\dim V'} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left(\int_{\Pi} |T^{\lambda} g|^2 |\eta_B|^2\right)^{1/(1+\delta)}.$$

If  $\xi \notin 2B$ , then  $\xi \notin \bigcup_{(\theta, v) \in \mathbb{T}_B} T_{\theta, v}$  and so  $|T^{\lambda} g_{\theta, v}(\xi)| = \text{RapDec}(R) \|g\|_{L^2(B^{n-1})}$  for all  $(\theta, v) \in \mathbb{T}_{B, \tau, V}$ . Hence

$$\int_{\Pi} |T^{\lambda} g|^2 |\eta_B|^2 \leq \int_{2B \cap \Pi} |T^{\lambda} g|^2 + \text{RapDec}(R) \|g\|_{L^2(B^{n-1})},$$

completing the proof of property 3) under the assumption that the above claim holds.

*Proof (of Claim).* Fix  $(\theta, v) \in \mathbb{T}_{B, \tau, V}$  and let  $u_{\theta} := \text{proj}_{x_{\Pi}^{\perp}} \Sigma(\omega_{\theta})$ . Recalling that  $A_{\xi} = A_u \times \mathbb{R}$  and applying triangle inequality, one deduces that

$$\text{dist}(\xi_{\theta}, A_{\xi}) = \text{dist}(u_{\theta}, A_u) \leq \text{dist}(u_{\theta}, S_u \cap \Psi^{-1}(\tau)) + \sup_{u_* \in S_u \cap \Psi^{-1}(\tau)} \text{dist}(u_*, A_u).$$

Furthermore,

$$\text{dist}(u_{\theta}, S_u \cap \Psi^{-1}(\tau)) \sim \text{dist}(\omega_{\theta}, S_u \cap \tau) \lesssim \not\prec (G^{\lambda}(\bar{x}; \omega_{\theta}), V) \lesssim R^{-1/2 + \delta_m},$$

where the last inequality is by the definition of  $\mathbb{T}_{B, \tau, V}$ . On the other hand, fixing  $u_* \in S_u \cap \Psi^{-1}(\tau)$ , one wishes to estimate  $\text{dist}(u_*, A_u)$ . Provided  $\rho$  is sufficiently large (so that  $\text{diam } \tau$  is sufficiently small), the surface  $S_u \cap \Psi^{-1}(\tau)$  can be parametrised as the graph of some function  $\psi: \mathcal{W} \rightarrow V_u^{\perp} \subseteq \mathbb{R}^{n-1}$  where  $\mathcal{W} \subset V_u$  is an open set about the origin of diameter  $O(\rho^{-1/2 + \delta_m})$ . In particular, one may write

$$S_u \cap \Psi^{-1}(\tau) = \{w + \psi(w) : w \in \mathcal{W}\} + u_0$$

where  $\psi(0) = 0$  and  $\partial_w \psi_j(0) = 0$  for  $1 \leq j \leq n - \dim V_u$ . Thus,  $u_* = w_* + \psi(w_*) + u_0$  for some  $w_* \in \mathcal{W}$  and, since  $w_* \in V_u + u_0 = A_u$ , it follows that  $\text{dist}(u_*, A_u) \leq |\psi(w_*)|$ . By Taylor's theorem,

$$|\psi_j(w_*)| \leq \int_0^1 (1-t) |\langle \partial_{ww}^2 \psi_j(tw_*) w_*, w_* \rangle| dt \lesssim |w_*|^2 \lesssim \rho^{-1+2\delta_m} \leq R^{-1/2+\delta_m}$$

for  $1 \leq j \leq n - \dim V_u$  and combining these observations yields the desired estimate.  $\square$

**Verifying the transversality condition in 1).** The transversality of the planes  $V$  and  $V'$  heavily relies upon the positive-definite hypothesis H2<sup>+</sup>); the property does not hold in general if one only assumes the weaker condition H2).

**Example 8.6.** For  $\phi_{\text{hyp}}(x; \omega) := \langle x', \omega \rangle + x_3 \omega_1 \omega_2$  for  $(x; \omega) \in \mathbb{R}^3 \times \mathbb{R}^2$  the oscillatory integral operator

$$E_{\text{hyp}} f(x) := \int_{B^2} e^{2\pi i \phi_{\text{hyp}}(x; \omega)} f(\omega) d\omega$$

is the extension operator associated to the hyperbolic paraboloid. This is the prototypical example of a Hörmander-type operator for which H2<sup>+</sup>) fails. Here  $G_0(\omega) = (-\omega_2, -\omega_1, 1)^\top$  and therefore if  $V := \{x \in \mathbb{R}^3 : x_1 = 0\}$ , then

$$S_\omega := \{\omega \in B^2 : G_0(\omega) \in V\} = \{\omega \in B^2 : \omega_2 = 0\}.$$

It follows that  $V_\xi = \{\xi \in \hat{\mathbb{R}}^3 : \xi_2 = 0\}$  and so  $V' := V_\xi^\perp$  is a vector subspace of  $V$ . Clearly, in this situation the desired transversality condition completely fails.

The present analysis concerns Hörmander-type operators with positive-definite reduced phase  $\phi^\lambda$ , so that  $\phi^\lambda$  is a small perturbation of  $\phi_{\text{par}}$ . Such phases do not exhibit the phenomenon observed in the above example: the following claim is key to understanding this.

**Claim.** *Let  $c_{\text{par}}$  be the constant defined in §4. Then*

$$\max_{v^* \in V \cap (S^{n-2} \times \{0\})} \not\prec(v^*, V_u) = O(c_{\text{par}}).$$

Here  $V_u$  is identified with a subspace of  $e_n^\perp = \mathbb{R}^{n-1} \times \{0\}$  in the natural manner.

**Example 8.7.** Returning to the example of the hyperbolic paraboloid with  $V := \{x \in \mathbb{R}^3 : x_1 = 0\}$ , the spaces  $V \cap (\mathbb{R}^2 \times \{0\})$  and  $V_u := \{x \in \mathbb{R}^3 : x_2 = x_3 = 0\}$  are mutually orthogonal, and so the claim does not hold in this case.

Provided  $c_{\text{par}} > 0$  is chosen sufficiently small, the claim implies the transversality condition. Indeed, let  $\{v_1^*, \dots, v_{\dim V - 1}^*\}$  be an orthonormal basis for  $V \cap e_n^\perp$ . Fix a unit vector  $v_{\dim V}^* \in V$  which is perpendicular to  $V \cap e_n^\perp$  so that  $\{v_1^*, \dots, v_{\dim V}^*\}$  forms an orthonormal basis for  $V$ . By the above claim, there exist  $v_k \in V_u \cap S^{n-2} \subset e_n^\perp$  such that

$$\not\prec(v_k^*, v_k) = O(c_{\text{par}}) \quad \text{for } 1 \leq k \leq \dim V - 1.$$

Applying the Gram–Schmidt process, one may further assume that  $\{v_1, \dots, v_{\dim V - 1}\}$  forms an orthonormal basis of  $V_u$ ; adjoining  $e_n$  to this set then gives an orthonormal basis of  $V_\xi$ . Given  $v \in V \cap S^{n-1}$  and writing

$$v = \sum_{k=1}^{\dim V - 1} \langle v, v_k^* \rangle v_k + \langle v, v_{\dim V}^* \rangle v_{\dim V}^* + \sum_{k=1}^{\dim V - 1} \langle v, v_k^* \rangle (v_k^* - v_k),$$

since  $\sin \not\prec(v, V') = |\text{proj}_{V_\xi} v|$ , it follows that

$$\begin{aligned} \sin \not\prec(v, V') &= \left( \sum_{k=1}^{\dim V-1} |\langle v, v_k^* \rangle|^2 + |\langle v, v_{\dim V}^* \rangle|^2 |\langle v_{\dim V}^*, e_n \rangle|^2 \right)^{1/2} - O(c_{\text{par}}) \\ &\geq \left( \sum_{k=1}^{\dim V} |\langle v, v_k^* \rangle|^2 \right)^{1/2} \cdot |\langle v_{\dim V}^*, e_n \rangle| - O(c_{\text{par}}). \end{aligned}$$

Consequently, provided that  $c_{\text{par}}$  is chosen to be sufficiently small,

$$\sin \not\prec(v, V') \geq |\langle v_{\dim V}^*, e_n \rangle| - O(c_{\text{par}}) \gtrsim 1;$$

indeed, the last inequality holds since (8.7) implies that

$$|\langle v_{\dim V}^*, e_n \rangle| = \not\prec(v_{\dim V}^*, e_n^\perp) = \overline{\not\prec}(V, e_n^\perp) \gtrsim 1.$$

This concludes the proof of the transversality condition, conditional on the above claim.

*Proof of Claim.* Fix  $v^* \in V \cap (S^{n-2} \times \{0\})$  and let  $v \in V_u \cap S^{n-2}$  denote the unit normalisation of the vector  $\text{proj}_{V_u} v^*$ . It suffices to show that  $\not\prec(v^*, v) = O(c_{\text{par}})$ . Since

$$v^* - v = (|\text{proj}_{V_u} v^*| - 1)v + \text{proj}_{V_u^\perp} v^*$$

and  $||\text{proj}_{V_u} v^*| - 1| \leq |\text{proj}_{V_u^\perp} v^*|$ , the problem is further reduced to proving that

$$(8.10) \quad |\text{proj}_{V_u^\perp} v^*| = O(c_{\text{par}}).$$

The surface  $S_u$  is defined by the equations  $-\langle \partial_u \bar{h}(u), N'_k \rangle + N_{k,n} = 0$  for  $1 \leq k \leq n - \dim V$  where, as above,  $\{N_1, \dots, N_{n-\dim V}\}$  is a choice of orthonormal basis for  $V^\perp$  and  $\bar{h}$  is the function introduced in (8.8). By differentiating these expressions, one deduces that a basis for  $V_u^\perp$  is given by  $\{M_1, \dots, M_{n-\dim V}\}$  where

$$M_k := \partial_{uu}^2 \bar{h}(u_0) N'_k \quad \text{for } 1 \leq k \leq n - \dim V.$$

Let  $1 \leq k \leq n - \dim V$  and recall from Lemma 4.3 that

$$(8.11) \quad \|\partial_{uu}^2 \bar{h}(u_0) - \mathbf{I}_{n-1}\|_{\text{op}} = O(c_{\text{par}}).$$

Consequently,

$$(8.12) \quad |M_k - N'_k| = O(c_{\text{par}})$$

and, combining this with the fact that  $\langle v^*, N'_k \rangle = 0$  for  $1 \leq k \leq n - \dim V$  (where  $v^*$  is identified with a vector in  $\mathbb{R}^{n-1}$  in the natural manner), it follows that

$$(8.13) \quad \langle v^*, M_k \rangle = \langle v^*, M_k - N'_k \rangle = O(c_{\text{par}}).$$

Let  $\mathbf{M}$  be the  $(n-1) \times (n - \dim V)$  matrix whose  $k$ th column is given by the vector  $M_k$ . The orthogonal projection of  $v^*$  onto the subspace  $V_u^\perp$  can be expressed in terms of  $\mathbf{M}$  via the formula

$$\text{proj}_{V_u^\perp} v^* := \mathbf{M}(\mathbf{M}^\top \mathbf{M})^{-1} \mathbf{M}^\top v^*.$$

By (8.13), the components of the vector  $\mathbf{M}^\top v^*$  are all  $O(c_{\text{par}})$ . Thus, to prove (8.10) (and thereby establish the claim) it suffices to show that  $\|\mathbf{M}(\mathbf{M}^\top \mathbf{M})^{-1}\|_{\text{op}} \lesssim 1$ , which would in turn follow from

$$\|\mathbf{M}\|_{\text{op}} \lesssim 1 \quad \text{and} \quad \|(\mathbf{M}^\top \mathbf{M})^{-1}\|_{\text{op}} \lesssim 1.$$

The bound for  $\mathbf{M}$  is an immediate consequence of the definition of the  $M_k$  and (8.11). The remaining estimate would follow if one could show that, provided  $c_{\text{par}}$

is sufficiently small,  $|\lambda| \gtrsim 1$  for every eigenvalue  $\lambda$  of the symmetric matrix  $\mathbf{M}^\top \mathbf{M}$ . By (8.12) and continuity of eigenvalues, it suffices to show that the matrix  $\mathbf{N}^\top \mathbf{N}$  satisfies the same property, where  $\mathbf{N}$  is the  $(n-1) \times (n - \dim V)$  matrix whose  $k$ th column is given by the vector  $N'_k$ . By (8.7) the vectors  $N'_1, \dots, N'_{n-\dim V} \in \mathbb{R}^{n-1}$  are linearly independent and, moreover, satisfy

$$|\det \mathbf{N}^\top \mathbf{N}| = |N'_1 \wedge \dots \wedge N'_{n-\dim V}|^2 \gtrsim 1.$$

Therefore, the desired condition on the eigenvalues holds if the spectral radius of  $\mathbf{N}^\top \mathbf{N}$  is  $O(1)$ . But the latter property is an obvious consequence of the Newton-Girard identity

$$\sum_{i=1}^m a_i^2 = \left( \sum_{i=1}^m a_i \right)^2 - 2 \sum_{1 \leq i_1 < i_2 \leq m} a_{i_1} a_{i_2} \quad a_i \in \mathbb{R} \text{ for } 1 \leq i \leq m$$

and the fact that the entries of  $\mathbf{N}^\top \mathbf{N}$  are all  $O(1)$ . □

This concludes the proof of Lemma 8.5. □

**8.4. The proof of the transverse equidistribution estimate.** It remains to demonstrate how to pass from Lemma 8.5 to Lemma 8.3. At this stage, the proof is very similar to the argument found in [13], but the details are nevertheless included for completeness.

There are two additional ingredients needed for the proof of Lemma 8.3. The first is the following theorem of Wongkew [26] (see also [14, 28]), which controls the size of a neighbourhood of a variety.

**Theorem 8.8** (Wongkew [26]). *Suppose  $Z = Z(P_1, \dots, P_{n-m})$  is an  $m$ -dimensional transverse complete intersection in  $\mathbb{R}^n$  with  $\deg Z \leq D$ . For any  $0 < \rho \leq R$  and  $R$ -ball  $B_R$  the neighbourhood  $N_\rho(Z \cap B_R)$  can be covered by  $O(D^n (R/\rho)^{n-m})$  balls of radius  $\rho$ .*

The second ingredient is a geometric lemma concerning planar slices of neighbourhoods of varieties. The statement of this result requires a general quantitative notion of transversality for pairs of linear subspaces in  $\mathbb{R}^n$ . Any  $m$ -dimensional linear subspace  $V$  can be expressed as a transverse complete intersection  $V = Z(P_{N_1}, \dots, P_{N_{n-m}})$  where  $\{N_1, \dots, N_{n-m}\}$  forms an orthonormal basis of  $V^\perp$  and  $P_{N_j}(x) := \langle x, N_j \rangle$ . Suppose  $V_1, V_2$  are linear subspaces in  $\mathbb{R}^n$  satisfying

$$(8.14) \quad \dim V_1 + \dim V_2 \geq n.$$

It is easy to verify that the subspace  $V_1 \cap V_2$  is a transverse complete intersection if and only if  $\dim V_1 \cap V_2 = \dim V_1 + \dim V_2 - n$  (of course, the inequality  $\dim V_1 \cap V_2 \geq \dim V_1 + \dim V_2 - n$  always holds so the latter condition says that  $V_1 \cap V_2$  is as small as possible).

**Definition 8.9.** A pair  $V_1, V_2$  of linear subspaces in  $\mathbb{R}^n$  satisfying (8.14) is said to be quantitatively transverse if the following hold:

- i)  $\dim(V_1 \cap V_2) = \dim V_1 + \dim V_2 - n$ ;
- ii)  $\angle(v_1, v_2) \geq c_{\text{trans}}$  for all non-zero vectors  $v_j \in (V_1 \cap V_2)^\perp \cap V_j$ ,  $j = 1, 2$ .

**Remark 8.10.** In the special case where  $\dim V_1 + \dim V_2 = n$ , it follows that the pair  $V_1, V_2$  is quantitatively transverse if and only if  $\angle(v_1, v_2) \geq c_{\text{trans}}$  for all non-zero vectors  $v_1 \in V_1$ ,  $v_2 \in V_2$ . Thus, up to the minor disparity between the choice

of constant  $c_{\text{trans}}$ , this agrees with the transversality condition appearing in the statement of Lemma 8.5.

**Lemma 8.11.** *There exists some dimensional constant  $C > 0$  such that the following holds. Let  $B_r \subseteq \mathbb{R}^n$  be an  $r$ -ball,  $V \subseteq \mathbb{R}^n$  a linear subspace,  $Z$  a transverse complete intersection and suppose that  $\dim Z + \dim V \geq n$  and  $T_z Z, V$  is a quantitatively transverse pair for all  $z \in Z \cap 2B_r$ . Then*

$$V \cap B_r \cap N_\rho(Z) \subseteq N_{C\rho}(V \cap Z)$$

for all  $0 < \rho \ll r$ .

The proof of this simple lemma is postponed until the end of this section.

*Proof (of Lemma 8.3).* If  $T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B = \emptyset$ , then it follows that

$$|T^\lambda g_{\theta,v}(x)| = \text{RapDec}(R) \|g\|_{L^2(B^{n-1})} \quad \text{for all } x \in N_{\rho^{1/2+\delta_m}}(Z) \cap B.$$

Consequently, one may assume that  $g$  is concentrated on only those wave packets from  $\mathbb{T}_{B,\tau,Z}$  for which  $T_{\theta,v}$  intersects  $N_{R^{1/2+\delta_m}}(Z) \cap B$  non-trivially. Suppose  $(\theta, v) \in \mathbb{T}_{B,\tau,Z}$  has this property and let  $x \in T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B$ . If  $z \in Z \cap 2B$ , then  $|x - z| \lesssim R^{1/2+\delta_m}$  and, by the  $R^{1/2+\delta_m}$ -tangent condition,

$$\angle(G^\lambda(x; \theta), T_z Z) \lesssim R^{-1/2+\delta_m}.$$

Since  $|G^\lambda(\bar{x}; \theta) - G^\lambda(x; \theta)| \lesssim |\bar{x} - x|/\lambda \lesssim R^{-1/2+\delta_m}$ , one concludes that

$$\angle(G^\lambda(\bar{x}; \theta), T_z Z) \lesssim R^{-1/2+\delta_m} \quad \text{for all } z \in Z \cap 2B.$$

Thus, there exists a subspace  $V \subseteq \mathbb{R}^n$  of minimal dimension  $\dim V \leq \dim Z$  such that

$$\angle(G^\lambda(\bar{x}; \theta), V) \lesssim R^{-1/2+\delta_m}$$

for all  $(\theta, v) \in \mathbb{T}_{B,\tau,Z}$  for which  $T_{\theta,v} \cap N_{R^{1/2+\delta_m}}(Z) \cap B \neq \emptyset$ . In particular,  $g$  is concentrated on wave packets from  $\mathbb{T}_{B,\tau,V}$ . One may apply Lemma 8.5 to find a subspace  $V'$  of dimension  $n - \dim V$  such that

$$(8.15) \quad \angle(v, v') \geq 2c_{\text{trans}} \quad \text{for all non-zero vectors } v \in V \text{ and } v' \in V'$$

and

$$(8.16) \quad \int_{\Pi \cap B(x_0, \rho^{1/2+\delta_m})} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)} (\rho/R)^{\dim V'/2} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left( \int_{\Pi \cap 2B} |T^\lambda g|^2 \right)^{1/(1+\delta)}$$

for every affine subspace  $\Pi$  parallel to  $V'$ . More precisely, the above estimate holds up to the inclusion of some additional rapidly decreasing term. This small error will propagate through the remainder of the argument but in the end will be harmless and is therefore suppressed in the notation.

It is claimed that for each  $z \in Z \cap 2B$  the tangent space  $T_z Z$  forms a quantitatively transverse pair  $T_z Z, V'$  with  $V'$ . Indeed, if this fails, then it is easy to see that for some  $z \in Z \cap 2B$  there exists a subspace  $W \subseteq T_z Z$  of dimension

$$\dim W > \dim Z - \dim V$$

with the property that  $\angle(w, V') < c_{\text{trans}}$  for all  $w \in W \setminus \{0\}$ . Consequently, the crucial angle condition (8.15) guarantees that

$$\angle(w, V) \geq c_{\text{trans}} \quad \text{for all } w \in W \setminus \{0\}.$$

This implies that there exists a linear map  $L: \mathbb{R}^n \rightarrow V$  such that  $L$  restricted to  $V$  is the identity,  $L$  restricted to  $W$  is zero and  $\|L\|_{\text{op}} \lesssim 1$ . Recall that for each  $(\theta, v) \in \mathbb{T}_{B,Z}$  one has  $\not\prec(G^\lambda(\bar{x}, \theta), V) \lesssim R^{-1/2+\delta_m}$  and so

$$\sup_{\omega, \omega' \in \theta} |L(G^\lambda(\bar{x}, \omega)) - G^\lambda(\bar{x}, \omega')| \lesssim R^{-1/2+\delta_m}.$$

On the other hand,  $G^\lambda(\bar{x}, \theta) \subset N_{C_1 R^{-1/2+\delta_m}}(T_z Z) \cap S^{n-1}$  and so  $L(G^\lambda(\bar{x}, \theta))$  lies in

$$L(N_{C_1 R^{-1/2+\delta_m}}(T_z Z) \cap S^{n-1}) \subseteq N_{C_2 R^{-1/2+\delta_m}}(L(T_z Z)).$$

This shows that for all  $(\theta, v) \in \mathbb{T}_{B,Z}$  one has

$$\not\prec(G^\lambda(\bar{x}, \theta), L(T_z Z)) \lesssim R^{-1/2+\delta_m}.$$

Since  $L$  vanishes on  $W$ , by rank-nullity  $L(T_z Z)$  is a subspace of dimension at most  $\dim Z - \dim W < \dim V$ . This contradicts the minimality of  $V$  and so  $(T_z Z, V')$  is a quantitatively transverse pair for all  $z \in Z \cap 2B$ .

By Lemma 8.11 one deduces that

$$\Pi \cap N_{\rho^{1/2+\delta_m}}(Z) \cap B \subseteq N_{C\rho^{1/2+\delta_m}}(\Pi \cap Z) \cap 2B.$$

Since  $\Pi \cap Z$  is a transverse complete intersection of dimension  $\dim V' + \dim Z - n$ , Wongkeew's theorem now implies that  $\Pi \cap N_{\rho^{1/2+\delta_m}}(Z) \cap B$  can be covered by

$$O(R^{O(\delta_m)}(R/\rho)^{(\dim V' + \dim Z - n)/2})$$

balls of radius  $\rho^{1/2+\delta_m}$ . Applying the estimate (8.16) to each of these balls and summing, one deduces that

$$\int_{\Pi \cap N_{\rho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)}(\rho/R)^{(n-m)/2} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left( \int_{\Pi \cap 2B} |T^\lambda g|^2 \right)^{1/(1+\delta)}.$$

Integrating over planes  $\Pi$  parallel to  $V'$  and applying Hölder's inequality, it follows that

$$\int_{N_{\rho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim_\delta R^{O(\delta_m)}(\rho/R)^{(n-m)/2} \|g\|_{L^2(B^{n-1})}^{2\delta/(1+\delta)} \left( \int_{2B} |T^\lambda g|^2 \right)^{1/(1+\delta)}.$$

Finally, recalling Hörmander's bound,

$$\left( \int_{2B} |T^\lambda g|^2 \right)^{1/(1+\delta)} \lesssim R^{1/2+O(\delta_m)} \left( \int_{B^{n-1}} |g|^2 \right)^{1/(1+\delta)}$$

and absorbing the implied rapidly decaying error into the main term, one concludes that

$$\int_{N_{\rho^{1/2+\delta_m}}(Z) \cap B} |T^\lambda g|^2 \lesssim_\delta R^{1/2+O(\delta_m)}(\rho/R)^{(n-m)/2} \|g\|_{L^2(B^{n-1})}^2,$$

which is the desired estimate.  $\square$

It remains to prove Lemma 8.11.

*Proof (of Lemma 8.11).* Applying a rotation, one may assume that  $V$  is the span of the co-ordinate vectors  $e_1, \dots, e_{\dim V}$ . For the purposes of this proof,  $\gamma_V := (\gamma_1, \dots, \gamma_{\dim V})$  and  $\gamma_{V^\perp} = (\gamma_{\dim V+1}, \dots, \gamma_n)$  will denote the orthogonal projections of a space curve  $\gamma$  onto  $V$  and  $V^\perp$ , respectively.

Suppose that  $x \in V \cap B_r \cap N_\rho(Z)$  and fix some  $z_0 \in Z \cap N_\rho(B_r)$  with  $0 < |x - z_0| < \rho$ . Let  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^n$  be the constant speed parametrisation of the line through  $\gamma(0) := z_0$  and  $\gamma(1) := x$ . To prove the lemma it suffices to show that there exists a curve  $\tilde{\gamma}: [0, 1] \rightarrow \mathbb{R}^n$  such that for all  $t \in [0, 1]$  the following hold:

- 1)  $\tilde{\gamma}(0) = \gamma(0) = z_0$ ,
- 2)  $\tilde{\gamma}(t) \in Z$ ,
- 3)  $\tilde{\gamma}_{V^\perp}(t) = \gamma_{V^\perp}(t)$ ,
- 4)  $|\tilde{\gamma}'(t)| \leq \bar{C} |\tilde{\gamma}'_{V^\perp}(t)|$  where  $\bar{C} := (\sin c_{\text{trans}})^{-1}$ .

Indeed, once this is established, observe that  $z_1 := \tilde{\gamma}(1) \in Z \cap V$  by properties 2) and 3). Furthermore, 3) and 4) ensure that

$$|\tilde{\gamma}'(t)| \leq \bar{C} |\tilde{\gamma}'_{V^\perp}(t)| \leq \bar{C} |\gamma'(t)| < \bar{C} \rho$$

and so, combining this observation with 1),

$$(8.17) \quad |x - \tilde{\gamma}(t)| \leq |x - z_0| + |z_0 - \tilde{\gamma}(t)| < (1 + \bar{C}t)\rho$$

for all  $t \in [0, 1]$ . In particular,  $|x - z_1| \lesssim \rho$ , giving the desired conclusion.

The transversality condition implies that the distribution  $W_z := (T_z Z \cap V)^\perp \cap T_z Z$  has rank  $n - \dim V$  on  $Z \cap 2B_r$  and, moreover,  $\text{proj}_{W_z^\perp}|_V: V \rightarrow W_z^\perp$  is an isomorphism for all  $z \in Z \cap 2B_r$ . Smoothly extend  $W_z$  to a small neighbourhood  $U$  of  $Z \cap 2B_r$  so that

$$(8.18) \quad \text{proj}_{W_x^\perp}|_V: V \rightarrow W_x^\perp \quad \text{is an isomorphism for all } x \in U.$$

The curve  $\tilde{\gamma}$  will be chosen so that its tangent always lies in this distribution. Given that  $\tilde{\gamma}_{V^\perp}$  is already defined by property 3), to satisfy this condition  $\tilde{\gamma}_V$  must be a solution to the ODE

$$\begin{cases} \text{proj}_{W_{x(t)}^\perp}(y'(t), \gamma'_{V^\perp}(t)) &= 0 \\ y(0) &= \text{proj}_V z_0 \end{cases}$$

where  $x(t) := (y(t), \gamma_{V^\perp}(t))$ . By (8.18), solving the above ODE is equivalent to solving a system of the form

$$(8.19) \quad \begin{cases} y' &= g(t, y) \\ y(0) &= \text{proj}_V z_0 \end{cases}$$

for  $g$  a smooth function defined on  $\{(t, y) \in \mathbb{R} \times V : (y, \gamma_{V^\perp}(t)) \in U\}$ . Note that  $g$  can be described explicitly in terms of the inverse of  $\text{proj}_{W_x^\perp}|_V$  and, provided  $U$  is appropriately chosen, the derivatives of  $g$  are bounded.

The Picard–Lindelöf existence theorem implies that the system (8.19) has a solution  $\tilde{\gamma}_V$  defined on an interval  $[0, T]$  for some  $T > 0$  such that  $\tilde{\gamma} := (\tilde{\gamma}_V, \gamma_{V^\perp})$  satisfies  $\tilde{\gamma}(t) \in 2B_r$  for all  $t \in [0, T]$ . On this interval the curve  $\tilde{\gamma}$  clearly further satisfies 1) and 3) and, by the tangency condition which motivated the definition of the ODE, 2) also holds. If  $t \in [0, T]$ , then it follows that  $\tilde{\gamma}(t) \in Z \cap 2B_r$  and  $\tilde{\gamma}'(t) \in W_{\tilde{\gamma}(t)}$  and so the transversality hypothesis implies that

$$\bar{C}^{-1} = \sin c_{\text{trans}} \leq \sin \angle(\tilde{\gamma}'(t), V) = \frac{|\tilde{\gamma}'_{V^\perp}(t)|}{|\tilde{\gamma}'(t)|}.$$

Rearranging, one concludes that properties 1) to 4) all hold on  $[0, T]$ .

It remains to show that  $T$  can be chosen to satisfy  $T \geq 1$ . If  $\text{dist}(\tilde{\gamma}(T), U^c) \geq 1$ , then the regularity of  $g$  implies that the interval of existence can be extended by a fixed increment. Thus, one may assume that at least one of the following holds:

$T \geq 1$  or  $|\tilde{\gamma}(T) - x| \geq r/2$ .<sup>12</sup> Supposing the latter holds, by the choice of  $T$  and (8.17), one deduces that

$$r/2 \leq |\tilde{\gamma}(T) - x| \leq (1 + \bar{C}T)\rho.$$

Provided  $r$  is chosen to be sufficiently large compared to  $\rho$ , the desired bound immediately follows.  $\square$

## 9. COMPARING WAVE PACKETS AT DIFFERENT SPATIAL SCALES

**9.1. Wave packet decomposition at scale  $\rho$ .** The proof of Theorem 1.6 relies on a multi-scale analysis and for this it is necessary to compare wave packets at different scales.

Let  $1 \ll R \ll \lambda$  and recall the decomposition

$$T^\lambda f(x) = \sum_{(\theta, v) \in \mathbb{T}} T^\lambda f_{\theta, v}(x) + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}$$

described in §5. Consider a smaller spatial scale<sup>13</sup>  $R^{1/2} \leq \rho \leq R$  and fix  $B(y, \rho) \subset B(0, R)$  with centre  $y \in X^\lambda$ . Each of the  $T^\lambda f_{\theta, v}$  can be further decomposed into wave packets at scale  $\rho$  over  $B(y, \rho)$ . To do this, first apply a transformation to recentre  $B(y, \rho)$  at the origin. For  $g: B^{n-1} \rightarrow \mathbb{C}$  integrable define  $\tilde{g} := e^{2\pi i \phi^\lambda(y; \cdot)} g$  so that

$$(9.1) \quad T^\lambda g(x) = \tilde{T}^\lambda \tilde{g}(\tilde{x}) \quad \text{for } \tilde{x} = x - y$$

where  $\tilde{T}^\lambda$  is the Hörmander-type operator with phase  $\tilde{\phi}^\lambda$  and amplitude  $\tilde{a}^\lambda$  given by

$$(9.2) \quad \tilde{\phi}(x; \omega) := \phi\left(x + \frac{y}{\lambda}; \omega\right) - \phi\left(\frac{y}{\lambda}; \omega\right) \quad \text{and} \quad \tilde{a}(x; \omega) := a\left(x + \frac{y}{\lambda}; \omega\right).$$

Applying this identity to the wave packet decomposition above,

$$T^\lambda f(x) = \sum_{(\theta, v) \in \mathbb{T}} \tilde{T}^\lambda((f_{\theta, v})^\sim)(\tilde{x}) + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

Each  $T^\lambda f_{\theta, v}$  is (spatially) concentrated on the curved  $R^{1/2+\delta}$ -tube  $T_{\theta, v}$  and, consequently, each  $\tilde{T}^\lambda(f_{\theta, v})^\sim$  is concentrated on the translate  $T_{\theta, v} - y$ . Since

$$(9.3) \quad \begin{aligned} \partial_\omega \tilde{\phi}^\lambda((\gamma_{\omega, v}^\lambda(t), t) - y; \omega) &= \partial_\omega \phi^\lambda((\gamma_{\omega, v}^\lambda(t), t); \omega) - \partial_\omega \phi^\lambda(y; \omega) \\ &= v - \partial_\omega \phi^\lambda(y; \omega), \end{aligned}$$

the core curve  $\Gamma_{\theta, v}^\lambda - y$  of  $T_{\theta, v} - y$  is equal to  $\Gamma_{\theta, v - \bar{v}(y; \omega_\theta)}$  where

$$\bar{v}(y; \omega) := \partial_\omega \phi^\lambda(y; \omega).$$

Now repeat the construction of the wave packets for each  $\tilde{T}^\lambda(f_{\theta, v})^\sim$ , but at scale  $\rho$ . In particular, cover  $\Omega$  by finitely overlapping caps  $\tilde{\theta}$  of radius  $\rho^{-1/2}$  and  $\mathbb{R}^{n-1}$

<sup>12</sup>Indeed, suppose both conditions fail for  $T$ . The failure of the latter condition implies that  $\text{dist}(\tilde{\gamma}(T), (2B_r)^c) \geq r/2$ . Since  $\tilde{\gamma}(T) \in Z$  by property 2), one concludes that  $\tilde{\gamma}(T)$  is far from  $U^c$  and thus the interval of existence for  $\tilde{\gamma}$  can be extended by a fixed increment. One may redefine  $T$  to be some value in the interval of existence incrementally larger than the original value of  $T$  and repeat this procedure until at least one of the stated conditions hold.

<sup>13</sup>Later it will be useful to assume the more stringent condition  $R^{1/2} \leq \rho \leq R^{1-2\delta}$ .

by finitely-overlapping balls of radius  $\rho^{(1+\delta)/2}$  centered at vectors  $\tilde{v} \in \rho^{(1+\delta)/2}\mathbb{Z}^{n-1}$ . Let  $\tilde{\mathbb{T}}$  denote the set of all pairs  $(\tilde{\theta}, \tilde{v})$ . For each  $(\theta, v) \in \mathbb{T}$  one may decompose

$$(f_{\theta,v})^\sim = \sum_{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}} (f_{\theta,v})_{\tilde{\theta}, \tilde{v}}^\sim + \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}$$

as in §5. The significant contributions to this sum arise from pairs  $(\tilde{\theta}, \tilde{v})$  belonging to

$$\tilde{\mathbb{T}}_{\theta,v} := \{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}} : \text{dist}(\theta, \tilde{\theta}) \lesssim \rho^{-1/2} \text{ and } |v - \bar{v}(y; \omega_\theta) - \tilde{v}| \lesssim R^{(1+\delta)/2}\},$$

as demonstrated by the following lemma.

**Lemma 9.1.** *If  $R^{1/2} \leq \rho \leq R$ , then, with the above definitions, the function  $(f_{\theta,v})^\sim$  is concentrated on wave packets from  $\tilde{\mathbb{T}}_{\theta,v}$ ; that is,*

$$(f_{\theta,v})^\sim = \sum_{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta,v}} (f_{\theta,v})_{\tilde{\theta}, \tilde{v}}^\sim + \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

*Proof.* Since  $(f_{\theta,v})^\sim$  is supported in  $\theta$ , clearly the wave packets of  $(f_{\theta,v})^\sim$  at scale  $\rho$  are all supported in

$$\bigcup_{\tilde{\theta}: \text{dist}(\tilde{\theta}, \theta) \lesssim \rho^{-1/2}} \tilde{\theta}.$$

Thus, it suffices to show that for each  $(\theta, v) \in \mathbb{T}$  and  $\rho^{-1/2}$ -cap  $\tilde{\theta}$  one has

$$\sum_{\tilde{v}: |v - \bar{v}(y; \omega_\theta) - \tilde{v}| \gtrsim R^{(1+\delta)/2}} (f_{\theta,v})_{\tilde{\theta}, \tilde{v}}^\sim = \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

Fixing  $(\theta, v) \in \mathbb{T}$  and  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}$  with  $|v - \bar{v}(y; \omega_\theta) - \tilde{v}| \gtrsim R^{(1+\delta)/2}$ , the above estimate would follow if one could show that

$$(f_{\theta,v})_{\tilde{\theta}, \tilde{v}}^\sim = (1 + R^{-1/2}|v - \bar{v}(y; \omega_\theta) - \tilde{v}|)^{-(n+1)} \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}.$$

Fourier inversion yields the pointwise bound

$$\begin{aligned} |(f_{\theta,v})_{\tilde{\theta}, \tilde{v}}^\sim(\omega)| &\leq \|(\tilde{\psi}_{\tilde{\theta}})^\sim\|_{L^1(\mathbb{R}^n)} \|\eta_{\tilde{v}}(\psi_{\tilde{\theta}}(f_{\theta,v})^\sim)^\sim\|_{L^\infty(\mathbb{R}^n)} \\ &\lesssim \|(\psi_{\tilde{\theta}})^\sim * ((f_{\theta,v})^\sim)^\sim\|_{L^\infty(B(\tilde{v}, CR^{(1+\delta)/2}))} \end{aligned}$$

for all  $\omega \in \mathbb{R}^{n-1}$ . Since  $(\psi_{\tilde{\theta}})^\sim$  is concentrated in  $B(0, \rho^{1/2})$ , the problem is further reduced to showing that

$$((f_{\theta,v})^\sim)^\sim(z) = (1 + R^{-1/2}|z - v + \bar{v}(y; \omega_\theta)|)^{-(n+1)} \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}$$

whenever  $|z - v + \bar{v}(y; \omega_\theta)| \gtrsim R^{(1+\delta)/2}$ .

Let  $\tilde{\psi}$  be a Schwartz function on  $\mathbb{R}^{n-1}$  satisfying  $\tilde{\psi}(\omega) = 1$  for  $\omega \in B^{n-1}$  and define  $\tilde{\psi}_\theta(\omega) := \tilde{\psi}(R^{-1/2}(\omega - \omega_\theta))$  so that

$$((f_{\theta,v})^\sim)^\sim = (\tilde{\psi}_\theta e^{2\pi i \phi^\lambda(y; \cdot)})^\sim * (f_{\theta,v})^\sim.$$

On the one hand, since  $\text{supp } \eta_v \subset B(v, CR^{(1+\delta)/2})$ , it is not difficult to show that

$$(9.4) \quad |(f_{\theta,v})^\sim(z)| = (1 + R^{-1/2}|z - v|)^{-(n+1)} \text{RapDec}(R)\|f\|_{L^2(B^{n-1})}$$

whenever  $|z - v| \gtrsim R^{(1+\delta)/2}$ . On the other hand, it is claimed that

$$(9.5) \quad |(\tilde{\psi}_\theta e^{2\pi i \phi^\lambda(y; \cdot)})^\sim(z)| = (1 + R^{-1/2}|z + \bar{v}(y; \omega_\theta)|)^{-(n+1)} \text{RapDec}(R)$$

whenever  $|z + \bar{v}(y; \omega_\theta)| \gtrsim R^{(1+\delta)/2}$ . A routine argument then shows that (9.4) and (9.5) imply the desired estimate for  $((f_{\theta,v})^\sim)^\sim$  and so it only remains to prove the claim.

The inverse Fourier transform in (9.5) can be expressed as

$$R^{-(n-1)/2} \int_{\mathbb{R}^{n-1}} e^{2\pi i(\langle z, \omega_\theta + R^{-1/2}\omega \rangle + \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega))} \tilde{\psi}(\omega) d\omega,$$

where the  $\omega$ -gradient of the phase is given by

$$(9.6) \quad R^{-1/2}([\partial_\omega \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(y; \omega_\theta)] + [\bar{v}(y; \omega_\theta) + z]).$$

Using the fact that  $\partial_{\omega\omega}^2 \phi^\lambda(0; \omega) = 0$ , the first term in (9.6) can be estimated thus:

$$|\partial_\omega \phi^\lambda(y; \omega_\theta + R^{-1/2}\omega) - \partial_\omega \phi^\lambda(y; \omega_\theta)| \lesssim R^{-1/2}|y| \leq R^{1/2}.$$

Consequently, if  $z \notin B(-\bar{v}(y; \omega_\theta), R^{(1+\delta)/2})$ , then the second term in (9.6) dominates the  $\omega$ -gradient of the phase and (9.6) is  $\gtrsim R^{\delta/2}$  in norm. Repeated integration by parts now implies (9.4), concluding the proof.  $\square$

**9.2. Tangency properties.** Let  $Z$  be a transverse complete intersection of dimension  $m$  and suppose  $h$  is a function which is concentrated on wave packets from

$$\mathbb{T}_{Z, B(y, \rho)} := \{(\theta, v) \in \mathbb{T}_Z : T_{\theta, v} \cap B(y, \rho) \neq \emptyset\}.$$

What can be said about the scale  $\rho$  wave packets of  $\tilde{h}$ ? In particular, do the lower scale wave packets inherit the tangency property; namely, is  $\tilde{h}$  concentrated on scale  $\rho$  wave packets which are  $\rho^{-1/2+\delta_m}$ -tangent to some variety? It transpires that this is **not** true in general. It is true, however, that  $\tilde{h}$  can be broken up into pieces which are each made up of scale  $\rho$  wave packets tangential to some translate of  $Z$ .

To make the above discussion precise, let  $\tilde{\gamma}_{\omega, v}^\lambda$  be the curve defined by

$$(9.7) \quad \partial_\omega \tilde{\phi}^\lambda(\tilde{\gamma}_{\omega, v}^\lambda(t), t; \omega) = v \quad \text{for } t \in (-\rho, \rho).$$

It is remarked that (9.3) implies the relation

$$(9.8) \quad \gamma_{\omega, v}^\lambda(t) = \tilde{\gamma}_{\omega, v-\bar{v}(y; \omega)}^\lambda(t - y_n) + y'.$$

Let  $\tilde{T}_{\omega, v}$  be the  $\rho^{1/2+\delta}$ -tube with core curve  $\tilde{\Gamma}_{\omega, v}^\lambda = (\tilde{\gamma}_{\omega, v}^\lambda(t), t)$  (defined analogously to the  $R^{1/2+\delta}$ -tube  $T_{\theta, v}$ ) and for  $b \in \mathbb{R}^n$  define

$$\tilde{\mathbb{T}}_{Z+b} := \{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}} : \tilde{T}_{\tilde{\theta}, \tilde{v}} \text{ is } \rho^{-1/2+\delta_m}\text{-tangent to } Z + b \text{ in } B(0, \rho)\}.$$

The key observation is as follows.

**Proposition 9.2.** *Let  $R^{1/2} \leq \rho \leq R^{1-\delta}$  and  $Z \subseteq \mathbb{R}^n$  be a transverse complete intersection.*

1) *Let  $(\theta, v) \in \mathbb{T}_Z$  and  $b \in B(0, 2R^{1/2+\delta_m})$ . If  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$  satisfies*

$$\tilde{T}_{\tilde{\theta}, \tilde{v}} \cap N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z - y + b) \neq \emptyset,$$

*then  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{Z-y+b}$ .*

2) *If  $h$  is concentrated on wave packets in  $\mathbb{T}_{Z, B(y, \rho)}$ , then  $\tilde{h}$  is concentrated on wave packets in  $\bigcup_{|b| \leq R^{1/2+\delta_m}} \tilde{\mathbb{T}}_{Z-y+b}$ .*

Under the hypotheses of the proposition, if one defines

$$\tilde{\mathbb{T}}_b := \left\{ (\tilde{\theta}, \tilde{v}) \in \bigcup_{(\theta, v) \in \mathbb{T}_{Z, B(y, \rho)}} \tilde{\mathbb{T}}_{\theta, v} : \tilde{T}_{\tilde{\theta}, \tilde{v}} \cap N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z - y + b) \neq \emptyset \right\},$$

then it follows that  $\tilde{\mathbb{T}}_b \subseteq \tilde{\mathbb{T}}_{Z-y+b}$ . The proof of Proposition 9.2 will in fact yield the following stronger result.

**Corollary 9.3.** *Under the hypotheses of 1) of Proposition 9.2, if  $(\theta, v) \in \mathbb{T}$  and  $\tilde{\mathbb{T}}_{\theta, v} \cap \tilde{\mathbb{T}}_b \neq \emptyset$ , then  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \subseteq N_{\rho^{1/2+\delta_m}}(Z - y + b)$  for all  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$ .*

In view of the forthcoming analysis, before proving these statements a simple application is discussed. Given a function  $h$  concentrated on wave packets in  $\mathbb{T}_{Z, B(y, \rho)}$ , consider a function of the form

$$(9.9) \quad \tilde{h}_b := \sum_{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_b} \tilde{h}_{\tilde{\theta}, \tilde{v}}.$$

Expressions of the form (9.9) will play an important rôle in later arguments. One may easily verify that  $\|\tilde{h}_b\|_{L^2(B^{n-1})} \lesssim \| [h_b^*]^\sim \|_{L^2(B^{n-1})}$  where

$$(9.10) \quad h_b^* := \sum_{\substack{(\theta, v) \in \mathbb{T}_{Z, B(y, \rho)} \\ \tilde{\mathbb{T}}_{\theta, v} \cap \tilde{\mathbb{T}}_b \neq \emptyset}} h_{\theta, v}.$$

The latter function is slightly easier to analyse; in particular, Corollary 9.3 implies that

$$(9.11) \quad \tilde{T}^\lambda [h_b^*]^\sim(\tilde{x}) = T^\lambda h_b^*(x) \chi_{N_{\rho^{1/2+\delta_m}}(Z+b)}(x) + \text{RapDec}(R) \|h\|_{L^2(B^{n-1})}$$

for all  $x = \tilde{x} + y \in B(y, \rho)$ .

The proof of Proposition 9.2 relies on the following lemma.

**Lemma 9.4.** *If  $(\theta, v) \in \mathbb{T}$  and  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$ , then*

$$|\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(t) - (\Gamma_{\theta, v}^\lambda(t + y_n) - y)| \lesssim R^{(1+\delta)/2} \quad \text{for all } t \in (-\rho, \rho).$$

*Proof.* By the identity (9.8) and the definition of  $\tilde{\mathbb{T}}_{\theta, v}$ , it suffices to show that if  $(\omega_1, v_1), (\omega_2, v_2) \in \Omega \times \mathbb{R}^{n-1}$  satisfy  $|\omega_1 - \omega_2| \lesssim \rho^{-1/2}$  and  $|v_1 - v_2| \lesssim R^{(1+\delta)/2}$ , then

$$|\tilde{\gamma}_{\omega_1, v_1}^\lambda(t) - \tilde{\gamma}_{\omega_2, v_2}^\lambda(t)| \lesssim R^{(1+\delta)/2} \quad \text{for all } t \in (-\rho, \rho).$$

Fixing  $t \in (-\rho, \rho)$ , let  $x_t := (\tilde{\gamma}_{\omega_1, v_1}^\lambda(t), t)$  and  $v_t := \partial_\omega \tilde{\phi}^\lambda(x_t; \omega_2)$  and note that, since the value of  $\tilde{\gamma}_{\omega, v}^\lambda(t)$  is uniquely determined by (9.7),  $x_t = (\tilde{\gamma}_{\omega_2, v_t}^\lambda(t), t)$ . Observe that

$$|v_1 - v_t| = |\partial_\omega \tilde{\phi}^\lambda(x_t; \omega_1) - \partial_\omega \tilde{\phi}^\lambda(x_t; \omega_2)| \lesssim |\omega_1 - \omega_2| |x_t| \lesssim \rho^{1/2}.$$

Since  $|v_1 - v_2| \lesssim R^{(1+\delta)/2}$ , it follows that  $|v_t - v_2| \lesssim R^{(1+\delta)/2}$ . Therefore,

$$|\tilde{\gamma}_{\omega_1, v_1}^\lambda(t) - \tilde{\gamma}_{\omega_2, v_2}^\lambda(t)| = |\tilde{\gamma}_{\omega_2, v_t}^\lambda(t) - \tilde{\gamma}_{\omega_2, v_2}^\lambda(t)| \sim |v_t - v_2| \lesssim R^{(1+\delta)/2},$$

which establishes the lemma.  $\square$

One may now turn to the proof of Proposition 9.2.

*Proof (of Proposition 9.2).* The proof is broken into stages.

*The angle condition.* Fix  $(\theta, v) \in \mathbb{T}_Z$  and  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$ . Motivated by the definition of tangency, let  $x \in \tilde{T}_{\tilde{\theta}, \tilde{v}}$  and suppose  $z \in Z$  and  $b \in B(0, 2R^{1/2+\delta_m})$  are such that  $z - y + b \in B(0, 4\rho)$  and  $|x - (z - y + b)| \leq \bar{C}_{\text{tang}}\rho^{1/2+\delta_m}$ . It is claimed that

$$(9.12) \quad \not\prec(\tilde{G}^\lambda(x; \omega_{\tilde{\theta}}), T_{z-y+b}(Z - y + b)) \leq \bar{c}_{\text{tang}}\rho^{-1/2+\delta_m},$$

where  $\tilde{G}^\lambda$  is the generalised Gauss map associated to the phase  $\tilde{\phi}^\lambda$ . It is easy to see that  $\tilde{G}^\lambda(x; \omega) = G^\lambda(x + y; \omega)$  and so the above estimate can be written as

$$(9.13) \quad \not\prec(G^\lambda(x + y; \omega_{\tilde{\theta}}), T_z Z) \leq \bar{c}_{\text{tang}}\rho^{-1/2+\delta_m}.$$

By Lemma 9.4, the definition of  $\tilde{T}_{\tilde{\theta}, \tilde{v}}$  and the hypothesis  $\rho \leq R^{1-\delta}$ , it follows that

$$(9.14) \quad |x + y - \Gamma_{\theta, v}^\lambda(x_n + y_n)| \lesssim R^{(1+\delta)/2}$$

and this implies that

$$\not\prec(G^\lambda(x + y; \omega_{\tilde{\theta}}), T_z Z) \lesssim \not\prec(G^\lambda(\Gamma_{\theta, v}^\lambda(x_n + y_n); \omega_\theta), T_z Z) + \rho^{-1/2}.$$

Finally,  $\Gamma_{\theta, v}^\lambda(x_n + y_n) \in T_{\theta, v}$  and this tube is  $R^{-1/2+\delta_m}$ -tangent to  $Z$ . Note that  $z \in Z \cap B(0, 2R)$  whilst, recalling (9.14), one has

$$|\Gamma_{\theta, v}^\lambda(x_n + y_n) - z| \leq |x + y - \Gamma_{\theta, v}^\lambda(x_n + y_n)| + |x - (z - y + b)| + |b| \lesssim R^{1/2+\delta_m}.$$

Thus, if the constant  $\bar{C}_{\text{tang}}$  in Definition 8.1 is appropriately chosen, then the tangency of  $T_{\theta, v}$  implies that

$$\not\prec(G^\lambda(\Gamma_{\theta, v}^\lambda(x_n + y_n); \omega_\theta), T_z Z) \leq \bar{c}_{\text{tang}}R^{-1/2+\delta_m}$$

and, provided that  $R$  is sufficiently large, (9.13) (and therefore (9.12)) immediately follows.

*Containment properties.* The angle condition (9.12) implies Corollary 9.3: any tube  $\tilde{T}_{\tilde{\theta}, \tilde{v}}$  which intersects  $N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z - y + b) \cap B(0, \rho)$  is actually contained in  $N_{\rho^{1/2+\delta_m}}(Z - y + b)$ . To demonstrate this containment property, continue with the setup of the previous stage, but now assume the slightly stronger conditions that  $z - y + b \in B(0, \rho)$  and  $|x - (z - y + b)| \leq (1/2)\rho^{1/2+\delta_m}$ . Define a time-dependent vector field  $X_{\tilde{\theta}, \tilde{v}}: (-\rho, \rho) \times Z \cap B(0, 2R) \rightarrow \mathbb{R}^m$  on  $Z \cap B(0, 2R)$  by

$$X_{\tilde{\theta}, \tilde{v}}(t, z) := \text{proj}_{T_z Z}(\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda)'(t) \quad \text{for all } (t, z) \in [-\rho, \rho] \times Z \cap B(0, 2R).$$

This can be smoothly extended to a map on  $[-\rho, \rho] \times U$  where  $U \subseteq \mathbb{R}^n$  is a small open neighbourhood of  $Z \cap B(0, 2R)$ . By the Picard–Lindelöf existence theorem for ODE there exists some smooth mapping  $Z_{\tilde{\theta}, \tilde{v}}: (-\rho, \rho) \rightarrow Z$  such that  $Z_{\tilde{\theta}, \tilde{v}}(x_n) = z$  and  $Z_{\tilde{\theta}, \tilde{v}}'(t) = X_{\tilde{\theta}, \tilde{v}}(t, Z_{\tilde{\theta}, \tilde{v}}(t))$  for all  $t \in (-\rho, \rho)$ . Here  $x = (x', x_n) \in \tilde{T}_{\tilde{\theta}, \tilde{v}}$  are the points fixed above.

Observe that

$$|\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(x_n) - (z - y + b)| \leq |\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(x_n) - x| + |x - (z - y + b)| < (2/3)\rho^{1/2+\delta_m}.$$

Let  $I$  denote the set of all  $t \in (-\rho, \rho)$  such that  $t \geq x_n$  and

$$|\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(s) - (Z_{\tilde{\theta}, \tilde{v}}(s) - y + b)| \leq (9/10)\rho^{1/2+\delta_m} \text{ for all } x_n \leq s \leq t.$$

It is claimed that  $t_* := \sup I = \rho$ . To see this, first note that if  $t_* < \rho$ , then

$$(9/10)\rho^{1/2+\delta_m} = |\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(t_*) - (Z_{\tilde{\theta}, \tilde{v}}(t_*) - y + b)|.$$

The angle condition (9.12) implies that

$$\not\prec((\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda)'(t), T_{Z_{\tilde{\theta}, \tilde{v}}(t)}Z) \leq \bar{c}_{\text{tang}}\rho^{-1/2+\delta_m} \quad \text{for all } x_n \leq t \leq t_*.$$

Combining the previous two displays with the identity

$$\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(t_*) - Z_{\tilde{\theta}, \tilde{v}}(t_*) = \int_{x_n}^{t_*} \text{proj}_{(T_{Z_{\tilde{\theta}, \tilde{v}}(t)}Z)^\perp}(\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda)'(t) dt + (\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(x_n) - z),$$

one concludes that

$$\begin{aligned} (9/10)\rho^{1/2+\delta_m} &< \int_{x_n}^{t_*} \sin \not\prec((\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda)'(t), T_{Z_{\tilde{\theta}, \tilde{v}}(t)}Z) |(\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda)'(t)| dt + (2/3)\rho^{1/2+\delta_m} \\ &\leq 2\bar{c}_{\text{tang}}\rho^{-1/2+\delta_m}|t_* - x_n| + (2/3)\rho^{1/2+\delta_m}. \end{aligned}$$

Since  $|t_* - x_n| \leq 2\rho$ , if  $\bar{c}_{\text{tang}}$  is appropriately chosen, then this yields a contradiction and, consequently,  $[x_n, \rho] \subseteq I$ . A similar argument shows that  $(-\rho, x_n] \subseteq I$  and so  $\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda((-\rho, \rho)) \subseteq N_{\frac{\rho}{10}\rho^{1/2+\delta_m}}(Z - y + b)$ . One therefore concludes that  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \subseteq N_{\rho^{1/2+\delta_m}}(Z - y + b)$ .

*Proof of Proposition 9.2, 1).* Let  $b \in B(0, 2R^{1/2+\delta_m})$  and suppose that  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \cap N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z - y + b) \cap B(0, \rho) \neq \emptyset$ ; the problem is to show that  $\tilde{T}_{\tilde{\theta}, \tilde{v}}$  is  $\rho^{-1/2+\delta_m}$ -tangential to  $Z - y + b$ . By hypothesis, there exists some  $x \in \tilde{T}_{\tilde{\theta}, \tilde{v}}$  and  $z \in Z$  such that  $z - y + b \in B(0, \rho)$  and  $|x - (z - y + b)| \leq (1/2)\rho^{1/2+\delta_m}$ . The preceding analysis therefore implies that  $\tilde{\mathbb{T}}_{\tilde{\theta}, \tilde{v}} \subseteq N_{\rho^{1/2+\delta_m}}(Z - y + b)$ , which is the desired containment condition for tangency. On the other hand, the angle condition for tangency always holds by (9.12), and so the proof of part 1) is complete.

*Proof of Proposition 9.2, 2).* By Lemma 9.1 it suffices to prove that

$$\bigcup_{(\theta, v) \in \mathbb{T}_{Z, B(y, \rho)}} \tilde{\mathbb{T}}_{\theta, v} \subseteq \bigcup_{|b| \lesssim R^{1/2+\delta_m}} \tilde{\mathbb{T}}_{Z-y+b}.$$

Fixing  $(\theta, v) \in \mathbb{T}_{Z, B(y, \rho)}$  and  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$ , by (9.12) the problem is further reduced to showing that there exists some  $|b| \lesssim R^{1/2+\delta_m}$  such that  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \subseteq N_{\rho^{1/2+\delta_m}}(Z - y + b)$ . Lemma 9.4 implies that  $\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(t) \in N_{CR^{1/2+\delta_m}}(Z - y)$  for  $t \in [-\rho, \rho]$ . Consequently, fixing  $t_0 \in [-\rho, \rho]$  there exists some  $|b| \lesssim R^{1/2+\delta_m}$  such that  $\tilde{\Gamma}_{\tilde{\theta}, \tilde{v}}^\lambda(t_0) \in Z - y + b$ . The desired inclusion now follows from the containment property discussed earlier in the proof.  $\square$

**9.3. Sorting the wave packets.** If  $(\theta, v) \in \mathbb{T}$  and  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{\theta, v}$ , then Lemma 9.4 implies that<sup>14</sup>

$$(9.15) \quad \text{dist}_H(T_{\theta, v} \cap B(y, \rho), \tilde{T}_{\tilde{\theta}, \tilde{v}} + y) \lesssim R^{1/2+\delta}.$$

In particular, if a pair of wave packets  $(\theta_1, v_1), (\theta_2, v_2) \in \mathbb{T}$  are such that  $\tilde{\mathbb{T}}_{\theta_1, v_1} \cap \tilde{\mathbb{T}}_{\theta_2, v_2} \neq \emptyset$ , then the tubes  $T_{\theta_1, v_1}, T_{\theta_2, v_2}$  are approximately equal on  $B(y, \rho)$ .<sup>15</sup> This suggests sorting the scale  $R$  wave packets  $(\theta, v) \in \mathbb{T}$  into disjoint sets for which the associated tubes essentially agree on  $B(y, \rho)$ .

<sup>14</sup>Here  $\text{dist}_H$  denotes the Hausdorff distance.

<sup>15</sup>More precisely, enlarging the radius of either one of the  $T_{\theta_j, v_j}$  by a constant factor produces a tube which contains  $(T_{\theta_1, v_1} \cup T_{\theta_2, v_2}) \cap B(y, \rho)$ .

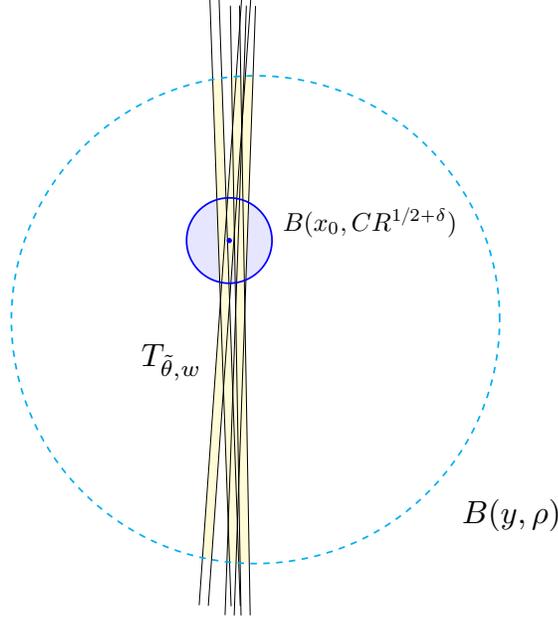


FIGURE 4. The set  $T_{\tilde{\theta}, w} := \bigcup_{(\theta, v) \in \mathcal{T}_{\tilde{\theta}, w}} T_{\theta, v} \cap B(y, \rho)$  is highlighted in yellow. Fixing  $x_0 \in T_{\tilde{\theta}, w}$ , for every  $(\theta, v) \in \mathcal{T}_{\tilde{\theta}, w}$  the tube  $T_{\theta, v}$  intersects the ball  $B(x_0, CR^{1/2+\delta})$ .

Let  $\mathcal{T}$  denote the collection of all pairs  $(\tilde{\theta}, w)$  formed by a  $\rho^{-1/2}$ -cap  $\tilde{\theta}$  and  $w \in R^{(1+\delta)/2}\mathbb{Z}^{n-1}$ . For each  $(\tilde{\theta}, w) \in \mathcal{T}$  choose some

$$\mathcal{T}_{\tilde{\theta}, w} \subseteq \{(\theta, v) \in \mathbb{T} : \text{dist}(\theta, \tilde{\theta}) \lesssim \rho^{-1/2} \text{ and } |v - \bar{v}(y; \omega_\theta) - w| \lesssim R^{(1+\delta)/2}\}$$

so that the family  $\{\mathcal{T}_{\tilde{\theta}, w} : (\tilde{\theta}, w) \in \mathcal{T}\}$  forms a covering of  $\mathbb{T}$  by disjoint sets. Defining

$$T_{\tilde{\theta}, w} := \bigcup_{(\theta, v) \in \mathcal{T}_{\tilde{\theta}, w}} T_{\theta, v} \cap B(y, \rho),$$

one obtains the following corollary to Lemma 9.4.

**Corollary 9.5.** *If  $(\tilde{\theta}, w) \in \mathcal{T}$  and  $(\theta, v) \in \mathcal{T}_{\tilde{\theta}, w}$ , then*

$$\text{dist}_H(T_{\theta, v} \cap B(y, \rho), T_{\tilde{\theta}, w}) \lesssim R^{1/2+\delta}.$$

Let  $g: B^{n-1} \rightarrow \mathbb{C}$  be integrable and define

$$g_{\tilde{\theta}, w} := \sum_{(\theta, v) \in \mathcal{T}_{\tilde{\theta}, w}} g_{\theta, v} \quad \text{for all } (\tilde{\theta}, w) \in \mathcal{T}.$$

Since the  $\mathcal{T}_{\tilde{\theta}, w}$  cover  $\mathbb{T}$  and are disjoint, it follows that

$$(9.16) \quad g = \sum_{(\tilde{\theta}, w) \in \mathcal{T}} g_{\tilde{\theta}, w} + \text{RapDec}(R) \|g\|_{L^2(B^{n-1})};$$

furthermore, the functions  $g_{\tilde{\theta},w}$  are almost orthogonal and, consequently,

$$(9.17) \quad \|g\|_{L^2(B^{n-1})}^2 \sim \sum_{(\tilde{\theta},w) \in \mathcal{T}} \|g_{\tilde{\theta},w}\|_{L^2(B^{n-1})}^2.$$

By Lemma 9.1, the function  $(g_{\tilde{\theta},w})^\sim$  is concentrated on scale  $\rho$  wave packets belonging to  $\bigcup_{(\theta,v) \in \mathcal{T}_{\tilde{\theta},w}} \tilde{\mathbb{T}}_{\theta,v}$ . This union is contained in

$$\tilde{\mathcal{T}}_{\tilde{\theta},w} := \{(\tilde{\theta}', \tilde{v}) \in \tilde{\mathbb{T}} : \text{dist}(\tilde{\theta}', \tilde{\theta}) \lesssim \rho^{-1/2} \text{ and } |\tilde{v} - w| \lesssim R^{(1+\delta)/2}\}$$

and therefore each  $(g_{\tilde{\theta},w})^\sim$  is concentrated on wave packets from  $\tilde{\mathcal{T}}_{\tilde{\theta},w}$ . The family  $\{\tilde{\mathcal{T}}_{\tilde{\theta},w} : (\tilde{\theta}, w) \in \mathcal{T}\}$  forms a covering of  $\tilde{\mathbb{T}}$  by almost disjoint sets. This implies almost orthogonality between the scale  $\rho$  wave packets of the different functions  $(g_{\tilde{\theta},w})^\sim$ . A particular consequence of this observation is that

$$(9.18) \quad \left\| \sum_{(\tilde{\theta},w) \in \mathcal{T}} (g_{\tilde{\theta},w})^\sim \right\|_{L^2(B^{n-1})}^2 \sim \sum_{(\tilde{\theta},w) \in \mathcal{T}} \|(g_{\tilde{\theta},w})^\sim\|_{L^2(B^{n-1})}^2,$$

where  $\tilde{h}_b$  is defined for a given function  $h$  as in (9.9).

**9.4. Transverse equidistribution revisited.** Let  $Z$  be an  $m$ -dimensional transverse complete intersection,  $(\tilde{\theta}, w) \in \mathcal{T}$  and  $h$  be a function concentrated on wave packets in  $\mathbb{T}_{Z \cap B(y,\rho)} \cap \mathcal{T}_{\tilde{\theta},w}$ . Here the key example to have in mind is  $h = g_{\tilde{\theta},w}$  for some function  $g$  concentrated on wave packets in  $\mathbb{T}_{Z,B(y,\rho)}$ .

Every scale  $R$  wave packet of  $h$  intersects  $B(y,\rho)$  on the set  $T_{\tilde{\theta},w}$  which, by Corollary 9.5, is comparable to  $T_{\theta,v} \cap B(y,\rho)$  for any  $(\theta,v) \in \mathcal{T}_{\tilde{\theta},w}$ .<sup>16</sup> Consequently, if  $x_0 \in T_{\tilde{\theta},w}$ , then all the scale  $R$  wave packets of  $h$  intersect  $B(x_0, CR^{1/2+\delta_m})$  (see Figure 4). Moreover, (9.15) implies that the scale  $\rho$  wave packets of  $\tilde{h}$  will intersect  $B(x_0 - y, CR^{1/2+\delta_m})$ . Under these conditions a useful reverse-type version of Hörmander's  $L^2$  bound holds.

**Lemma 9.6.** *Let  $T^\lambda$  be a Hörmander-type operator with phase  $\phi^\lambda$  given by a translate of a reduced phase in the sense of (9.2) and  $1 \leq R^{1/2+\delta} \leq r \lesssim \lambda^{1/2}$ . If  $f$  is concentrated on wave packets which intersect some  $B(\bar{x}, r) \subseteq B(0, R)$ , then<sup>17</sup>*

$$(9.19) \quad \|f\|_{L^2(B^{n-1})}^2 \lesssim r^{-1} \|T^\lambda f\|_{L^2(B(\bar{x}, Cr))}^2.$$

Lemma 9.6 can be proven for extension operators fairly directly via Plancherel's theorem (see [13, §3]). Establishing the general (variable coefficient) version of Lemma 9.6 involves a number of additional technicalities and the proof is therefore postponed until the end of the section.

For  $h$  as above,  $x_0 \in T_{\tilde{\theta},w}$  and  $|b| \lesssim R^{1/2+\delta_m}$  the preceding discussion implies that the function  $[h_b^*]^\sim$ , as defined in (9.10), is a sum of wave packets which intersect  $B(x_0 - y, CR^{1/2+\delta_m})$ . Lemma 9.6 can therefore be applied at scale  $\rho$  with  $r \sim R^{1/2+\delta_m}$  to yield

$$\|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim \|[h_b^*]^\sim\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2-\delta_m} \|\tilde{T}^\lambda [h_b^*]^\sim\|_{L^2(B(x_0-y, CR^{1/2+\delta_m}))}^2.$$

<sup>16</sup>Here 'the scale  $R$  wave packets of  $h$ ' refers to the scale  $R$  wave packets upon which  $h$  is concentrated.

<sup>17</sup>The operator  $T^\lambda$  appearing on the right-hand side of (9.19) will be defined with respect to an amplitude with slightly larger support than the  $a^\lambda$  appearing in the original definition of the operator. The choice amplitude function in  $T^\lambda$  is essentially irrelevant in future applications of (9.19), so this technicality does not present any difficulties.

By (9.11), one concludes that

$$(9.20) \quad \|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2-\delta_m} \|T^\lambda h_b^*\|_{L^2(N_{\rho^{1/2+\delta_m}}(Z+b) \cap B(x_0, CR^{1/2+\delta_m}))}^2.$$

This observation has several useful consequences. First of all, by applying Hörmander's  $L^2$  bound one obtains the following.

**Lemma 9.7.** *Let  $h$  be concentrated on wave packets from  $\mathbb{T}_{Z \cap B(y, \rho)} \cap \mathcal{T}_{\tilde{\theta}, w}$ , for some  $(\tilde{\theta}, w) \in \mathcal{T}$ . Let  $\mathfrak{B} \subset B(0, CR^{1/2+\delta_m})$  be such that the sets  $N_{\rho^{1/2+\delta_m}}(Z+b) \cap B(x_0, CR^{1/2+\delta_m})$  are pairwise disjoint over all  $b \in \mathfrak{B}$ . Then,*

$$\sum_{b \in \mathfrak{B}} \|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim \|h\|_{L^2(B^{n-1})}^2.$$

A further consequence of (9.20) is the following transverse equidistribution result.

**Lemma 9.8.** *Let  $(\tilde{\theta}, w) \in \mathcal{T}$ ,  $|b| \lesssim R^{1/2+\delta_m}$  and  $Z$  be an  $m$ -dimensional transverse complete intersection with  $\overline{\deg} Z \lesssim_\varepsilon 1$ . If  $h$  is concentrated on wave packets from  $\mathbb{T}_{Z \cap B(y, \rho)} \cap \mathcal{T}_{\tilde{\theta}, w}$ , then*

$$\|\tilde{h}_b\|_{L^2(B^{n-1})}^2 \lesssim R^{O(\delta_m)} (\rho/R)^{(n-m)/2} \|h\|_{L^2(B^{n-1})}^2.$$

*Proof.* The transverse equidistribution estimate in Lemma 8.3 implies that

$$\|T^\lambda \tilde{h}_b^*\|_{L^2(N_{\rho^{1/2+\delta_m}}(Z+b) \cap B(x_0, CR^{1/2+\delta_m}))}^2 \lesssim R^{1/2+O(\delta_m)} (\rho/R)^{(n-m)/2} \|h\|_{L^2(B^{n-1})}^2.$$

Combining this with (9.20) completes the proof.  $\square$

Let  $g$  be concentrated on wave packets of  $\mathbb{T}_{Z, B(y, \rho)}$ . For each  $(\tilde{\theta}, w) \in \mathcal{T}$  the function  $g_{\tilde{\theta}, w}$  is concentrated on wave packets in  $\mathbb{T}_{Z \cap B(y, \rho)} \cap \mathcal{T}_{\tilde{\theta}, w}$ . It follows that Lemma 9.7 and Lemma 9.8 hold for  $h = g_{\tilde{\theta}, w}$ . Combining the contributions from distinct  $\mathcal{T}_{\tilde{\theta}, w}$ , one obtains the following.

**Lemma 9.9.** *Let  $|b| \lesssim R^{1/2+\delta_m}$  and  $Z$  be an  $m$ -dimensional transverse complete intersection with  $\overline{\deg} Z \lesssim_\varepsilon 1$ . If  $g$  is concentrated on wave packets from  $\mathbb{T}_{Z, B(y, \rho)}$ , then*

$$\|\tilde{g}_b\|_{L^2(B^{n-1})} \lesssim R^{O(\delta_m)} (\rho/R)^{(n-m)/2} \|g\|_{L^2(B^{n-1})}.$$

*Proof.* By (9.16) and the linearity of the map  $h \mapsto \tilde{h}_b$  it follows that

$$\tilde{g}_b = \sum_{(\tilde{\theta}, w) \in \mathcal{T}} (g_{\tilde{\theta}, w})_{\tilde{b}} + \text{RapDec}(R) \|g\|_{L^2(B^{n-1})}.$$

The almost orthogonality relation (9.18) between the  $(g_{\tilde{\theta}, w})_{\tilde{b}}$  implies that

$$\|\tilde{g}_b\|_{L^2(B^{n-1})}^2 \lesssim \sum_{(\tilde{\theta}, w) \in \mathcal{T}} \|(g_{\tilde{\theta}, w})_{\tilde{b}}\|_{L^2(B^{n-1})}^2 + \text{RapDec}(R) \|g\|_{L^2(B^{n-1})}^2.$$

By Lemma 9.8, the right-hand side of the above display is in turn dominated by

$$R^{O(\delta_m)} (\rho/R)^{(n-m)/2} \sum_{(\tilde{\theta}, w) \in \mathcal{T}} \|g_{\tilde{\theta}, w}\|_{L^2(B^{n-1})}^2 + \text{RapDec}(R) \|g\|_{L^2(B^{n-1})}^2.$$

An application of (9.17) yields the desired estimate.  $\square$

### 9.5. The proof of the reverse $L^2$ bound.

*Proof (of Lemma 9.6).* One may assume without loss of generality that  $\bar{x} = 0$  and  $r = R^{1/2+\delta}$ . Indeed, the first reduction follows from the formula (9.1), which can be used to replace  $T^\lambda$  and  $f$  with  $\tilde{T}^\lambda$  and  $\tilde{f}$ , here defined with  $y := \bar{x}$ . Lemma 9.1 and the identity (9.8) imply that  $\tilde{f}$  is concentrated on scale  $R$  wave packets associated to  $\tilde{T}^\lambda$  which intersect  $B(0, r)$ . For the second reduction, suppose the result is known for  $r = R^{1/2+\delta}$  and let  $R^{1/2+\delta} \leq r \leq R$  and  $f$  be as in the statement of the theorem. For  $a \in \mathbb{R}$  consider the slab

$$S_a := \mathbb{R}^{n-1} \times [a - R^{1/2+\delta}, a + R^{1/2+\delta}] \cap B(0, Cr)$$

where  $C \geq 2$  is a constant, chosen large enough for the purposes of the argument. Cover  $S_a$  with a collection  $\{B_j\}_{j \in \mathcal{J}}$  of finitely-overlapping  $R^{1/2+\delta}$ -balls satisfying  $B_j \cap S_a \neq \emptyset$  for all  $j \in \mathcal{J}$ . By the initial reductions, any tube  $T_{\theta, v}$  makes a small angle with the  $e_n$  direction and thus intersects at most  $O(1)$  of these balls. Orthogonality of the wave packets together with the hypothesised estimate therefore imply that

$$(9.21) \quad \|g\|_{L^2(B^{n-1})}^2 \lesssim R^{-1/2-\delta} \|T^\lambda g\|_{L^2(N_{CR^{1/2+\delta}}(S_a))}^2$$

for any  $g$  concentrated on wave packets at scale  $R$  which intersect  $S_a$ . If  $(\theta, v) \in \mathbb{T}_{B(0, r)}$ , then the aforementioned angle condition implies that  $T_{\theta, v}$  intersects every slab  $S_a$  for which  $S_a \cap B(0, r) \neq \emptyset$ . Hence, if  $f$  is concentrated on wave packets from  $\mathbb{T}_{B(0, r)}$ , then one may sum (9.21) over a collection of  $\sim r/R^{1/2+\delta}$  slabs which cover  $B(0, r)$  to obtain the desired result.

Fix a function  $f$  satisfying the hypotheses of the lemma with  $\bar{x} = 0$  and  $r = R^{1/2+\delta}$  and note that

$$(9.22) \quad \|f\|_{L^2(B^{n-1})} \sim \left\| \sum_{(\theta, v): T_{\theta, v} \cap B(0, R^{1/2+\delta}) \neq \emptyset} f_{\theta, v} \right\|_{L^2(B^{n-1})}.$$

Let  $\Psi: U \rightarrow \Omega$  be the diffeomorphism  $\Psi(u) := \Psi^\lambda(0; u)$ , where  $\Psi^\lambda$  is the rescaled version of map  $\Psi$  introduced in §4. Recall that  $\Psi$  reparametrises the surface  $\{\partial_x \phi^\lambda(0; \omega) : \omega \in \Omega\}$  as the graph of the function  $\bar{h}(u) := \partial_{x_n} \phi^\lambda(0; \Psi(u))$ ; in particular,  $\partial_{x'} \phi^\lambda(0; \Psi(u)) = u$  for all  $u \in U$ . Applying the change of variables  $\omega \mapsto \Psi(u)$ , and denoting the corresponding Jacobian by  $J_\Psi$ , it follows that  $\|f\|_{L^2(B^{n-1})} \sim \|f_\Psi\|_{L^2(B^{n-1})}$  where

$$f_\Psi := \left( \sum_{(\theta, v): T_{\theta, v} \cap B(0, R^{1/2+\delta}) \neq \emptyset} f_{\theta, v} \circ \Psi \right) \cdot J_\Psi.$$

Let  $E$  denote the extension operator

$$Eg(x) := \int_U e^{2\pi i(\langle x', u \rangle + x_n \bar{h}(u))} g(u) \, du$$

associated to the graph  $u \mapsto (u, \bar{h}(u))$ . For any  $x_n \in \mathbb{R}$  and  $g$  a square integrable function supported on  $U$ , Plancherel's theorem implies that

$$\|g\|_{L^2(B^{n-1})}^2 = \|e^{2\pi i x_n \bar{h}} g\|_{L^2(B^{n-1})}^2 = \|Eg\|_{L^2(\mathbb{R}^{n-1} \times \{x_n\})}^2.$$

Hence, averaging this estimate over  $|x_n| < R^{1/2+\delta}$  one obtains

$$(9.23) \quad \|g\|_{L^2(B^{n-1})}^2 \sim R^{-1/2-\delta} \|Eg\|_{L^2(\mathbb{R}^{n-1} \times (-R^{1/2+\delta}, R^{1/2+\delta}))}^2.$$

The key observation is that the hypothesis on  $f$  implies that the right-hand  $L^2$ -norm can be localised.

**Claim.** *If  $|x_n| < R^{1/2+\delta}$  and  $x' \notin B(0, CR^{1/2+\delta})$ , then*

$$Ef_{\Psi}(x) = (1 + R^{-1/2}|x'|)^{-(n+1)} \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

This concentration property is very similar to that detailed in Lemma 5.2, the main difference being that the condition  $(\theta, v) \in \mathbb{T}_{B(0, R^{1/2+\delta})}$  is defined with respect to the operator  $T^\lambda$  whilst the above identity concerns the linearised version  $E$ . The proof is a minor adaptation of the stationary phase analysis used to establish Lemma 5.2 and is therefore omitted.

For the specific choice of function  $g = f_{\Psi}$ , the claim implies that (9.23) may be strengthened to

$$(9.24) \quad \|f\|_{L^2(B^{n-1})}^2 \sim \|f_{\Psi}\|_{L^2(B^{n-1})}^2 \sim R^{-1/2-\delta} \|Ef_{\Psi}\|_{L^2(B(0, CR^{1/2+\delta}))}^2.$$

This is easily seen to imply the lemma. Indeed, reversing the earlier change of variables,

$$Ef_{\Psi}(z) = \int_{\Omega} e^{2\pi i \phi^\lambda(z; \omega)} e^{-2\pi i \lambda \Omega(z/\lambda; \omega)} f(\omega) d\omega + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}$$

for all  $z \in B(0, CR^{1/2+\delta})$ , where  $\Omega$  is the error term in the Taylor expansion

$$\phi(z; \omega) = \phi(0; \omega) + \langle \partial_x \phi(0; \omega), z \rangle + \Omega(z; \omega).$$

Were it not for the factor  $e^{-2\pi i \lambda \Omega(z/\lambda; \omega)}$  the functions  $Ef_{\Psi}(z)$  and  $T^\lambda f(z)$  would be equal up to a negligible error term and (9.24) would directly imply the desired estimate. This unwanted additional factor can be removed via a Fourier series decomposition. Recalling the hypothesis  $R^{1/2+\delta} \lesssim \lambda^{1/2}$ , it follows that

$$(9.25) \quad |\partial_\omega^\beta \partial_x^\alpha \lambda \Omega(R^{1/2+\delta} z/\lambda; \omega)| \lesssim_{\alpha, \beta} 1$$

for all multi-indices  $(\alpha, \beta) \in \mathbb{N}_0^n \times \mathbb{N}_0^{n-1}$ . Let  $\psi \in C_c^\infty(\mathbb{R}^n \times \mathbb{R}^{n-1})$  be supported on  $X \times \Omega$  and equal to 1 on  $\text{supp } a$ . By forming the Fourier series expansion in both the  $x$  and  $\omega$  variables one obtains

$$e^{-2\pi i \lambda \Omega(z/\lambda; \omega)} \psi(z/R^{1/2+\delta}; \omega) = \sum_{\substack{k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1} \\ k = (k_1, k_2)}} (1 + |k|)^{-N} c_k e^{2\pi i (\langle R^{1/2+\delta} z, k_1 \rangle + \langle \omega, k_2 \rangle)},$$

where the  $c_k$  are weighted Fourier coefficients and  $N \geq n + 1$  is an integer which depends only on  $n$  and is chosen large enough to satisfy the forthcoming requirements of the proof. Observe that (9.25) implies that  $|c_k| \lesssim 1$  for all  $k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1}$ . Thus, (9.24) now yields

$$\|f\|_{L^2(B^{n-1})} \lesssim R^{1/4+\delta/2} \sum_{\substack{k \in \mathbb{Z}^n \times \mathbb{Z}^{n-1} \\ k = (k_1, k_2)}} (1 + |k|)^{-N} \|T^\lambda(e^{2\pi i \langle \cdot, k_2 \rangle} f)\|_{L^2(B(0, CR^{1/2+\delta}))}.$$

Recalling Hörmander's  $L^2$ -estimate from Lemma 5.3, there exists a constant  $A$ , depending only on the dimension, such that

$$\|f\|_{L^2(B^{n-1})} \leq A(R^{1/4+\delta/2} \|T^\lambda f\|_{L^2(B(0, CR^{1/2+\delta}))}) + \sum_{|k| \neq 0} (1 + |k|)^{-N} \|f\|_{L^2(B^{n-1})}.$$

If  $N$  is chosen appropriately, then  $A \sum_{|k| \neq 0} (1 + |k|)^{-N} \leq 1/2$  and the desired estimate follows.  $\square$

10. PROOF OF THE  $k$ -BROAD ESTIMATE

**10.1. A more general result.** In this section the proof of the  $k$ -broad estimate in Theorem 1.6 is given. In order to facilitate an inductive argument, a more general result will be established, which is described presently.

Throughout this section  $T^\lambda$  denotes an arbitrary choice of a translate of a Hörmander-type operator with reduced positive-definite phase. That is,  $T^\lambda$  is of the form of the operator  $\tilde{T}^\lambda$  discussed in the previous section, with phase and amplitude of the type described in (9.2). Many of the estimates involving  $T^\lambda$ , such as (10.1) below, are understood to hold uniformly for the entire class of such operators; it is important to work with the whole class rather than a single choice of  $T^\lambda$  in order to run certain induction arguments.

In order to state the result, first define the exponent

$$e_{k,n}(p) := \frac{1}{2} \left( \frac{1}{2} - \frac{1}{p} \right) (n+k)$$

and note that  $-e_{k,n}(p) + 1/2 \leq 0$  if and only if  $p \geq \bar{p}(k,n)$  where  $\bar{p}(k,n) = 2(n+k)/(n+k-2)$  is the exponent appearing in Theorem 1.6.

**Proposition 10.1.** *Given  $\varepsilon > 0$  sufficiently small and  $1 \leq m \leq n$  there exist*

$$0 < \delta \ll \delta_{n-1} \ll \delta_{n-2} \ll \dots \ll \delta_1 \ll \varepsilon$$

and constants  $\bar{C}_\varepsilon, \bar{A}_\varepsilon$  dyadic,  $D_{m,\varepsilon} \lesssim_\varepsilon 1$  and  $\vartheta_m < \varepsilon$  such that the following holds.

Suppose  $Z = Z(P_1, \dots, P_{n-m})$  is a transverse complete intersection with  $\deg Z \leq D_{m,\varepsilon}$ . For all  $2 \leq k \leq n$ ,  $1 \leq A \leq \bar{A}_\varepsilon$  dyadic and  $1 \leq K \leq R \leq \lambda$  the inequality

$$(10.1) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim_\varepsilon K^{\bar{C}_\varepsilon} R^{\vartheta_m + \delta (\log \bar{A}_\varepsilon - \log A) - e_{k,n}(p) + 1/2} \|f\|_{L^2(B^{n-1})}$$

holds whenever  $f$  is concentrated on wave packets from  $\mathbb{T}_Z$  and

$$(10.2) \quad 2 \leq p \leq \bar{p}_0(k,m) := \begin{cases} \bar{p}(k,m) & \text{if } k < m \\ \bar{p}(m,m) + \delta & \text{if } k = m \end{cases} .$$

Here  $\mathbb{T}_Z$  is defined as in §8; that is,

$$\mathbb{T}_Z := \{(\theta, v) \in \mathbb{T} : T_{\theta,v} \text{ is } R^{-1/2 + \delta_m}\text{-tangent to } Z \text{ in } B(0, R)\}.$$

Proposition 10.1 immediately yields the desired  $k$ -broad estimate.

*Proof (of Theorem 1.6).* Theorem 1.6 is a special case of Proposition 10.1. Indeed,  $Z = \mathbb{R}^n$  is a transverse complete intersection of dimension  $n$  and  $\mathbb{T}_Z$  contains all wave packets in  $B(0, R)$ . Thus, taking  $A = \bar{A}_\varepsilon$  and  $p = \bar{p}(k,n)$  yields the endpoint case of Theorem 1.6. The general result follows by interpolating with the trivial  $p = \infty$  estimate via the logarithmic convexity of the  $k$ -broad norms (that is, Lemma 6.3).  $\square$

**10.2. Reducing to  $R \lesssim \lambda^{1-\varepsilon}$ .** Turning to the proof of Proposition 10.1, the first step is a technical reduction on the radii  $R$  which is needed to facilitate certain polynomial partitioning arguments. In particular, it will be necessary to approximate the curves  $\Gamma_{\theta,v}^\lambda$  by polynomial curves of degree independent of  $R$ ; by the observations of §7.2, this is possible when  $1 \leq R \lesssim_\varepsilon \lambda^{1-\varepsilon}$  and thus (10.1) will first be proved for this restricted range of  $R$ . The result can then be extended to the full  $R \leq \lambda$  range by a triangle inequality argument (incurring a permissible  $R^{O(\varepsilon)}$  loss). The concentration hypothesis on  $f$  in Proposition 10.1 creates some difficulties here, which are addressed by the following lemma.

**Lemma 10.2.** *Let  $1 \leq \rho \leq R \leq \lambda$  and assume that for any transverse complete intersection  $Z = Z(P_1, \dots, P_{n-m})$  with  $\overline{\deg} Z \leq D$  the inequality*

$$(10.3) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,\rho))} \leq E \|f\|_{L^2(B^{n-1})}$$

*holds with some constant  $E > 0$  whenever  $f$  is concentrated on wave packets from  $\tilde{\mathbb{T}}_Z$ . Then for all  $Z$  as above, the inequality*

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim (R/\rho)^{O(1)} E \|f\|_{L^2(B^{n-1})}$$

*holds for all functions  $f$  concentrated on wave packets in  $\mathbb{T}_Z$ .*

Here  $\tilde{\mathbb{T}}_Z$  denotes the collection of wave packets at scale  $\rho$  that are  $\rho^{1/2+\delta_m}$ -transverse to  $Z$  in  $B(0, \rho)$ ; this notation is consistent with that used in §9.

*Proof (of Lemma 10.2).* Let  $f$  be a function concentrated on wave packets in  $\mathbb{T}_Z$  for some transverse complete intersection  $Z$  as in the statement of the lemma. Fix a cover  $\mathcal{B}_\rho$  of  $B(0, R)$  by finitely overlapping  $\rho$ -balls. By the sub-additivity of the  $k$ -broad norms and Lemma 5.2, there exists some  $B = B(y, \rho) \in \mathcal{B}_\rho$  such that

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim (R/\rho)^{O(1)} \|T^\lambda h\|_{\text{BL}_{k,A}^p(B)}^p + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}^p,$$

where  $h := \sum_{(\theta,v) \in \mathbb{T}_{Z,B}} f_{\theta,v}$ ; here and below the notation  $\mathbb{T}_{Z,B}$  is consistent with that used in §8. As in §9, write  $T^\lambda h(x+y) = \tilde{T}^\lambda \tilde{h}(x)$  so that, suppressing the harmless rapidly decaying term in the notation, one has<sup>18</sup>

$$(10.4) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim (R/\rho)^{O(1)} \|\tilde{T}^\lambda \tilde{h}\|_{\text{BL}_{k,A}^p(B(0,\rho))}^p.$$

In general,  $\tilde{h}$  is not concentrated on wave packets which are  $\rho^{-1/2+\delta_m}$ -tangential to a suitable variety inside  $B(0, \rho)$ ; thus, hypothesis (10.3) cannot be applied directly to estimate the right-hand side of (10.4). Rather, one approaches the problem via the methods of §9. By Proposition 9.2,  $\tilde{h}$  is concentrated on wave packets in  $\bigcup_{b \in \mathfrak{B}} \tilde{\mathbb{T}}_{Z-y+b}$  where  $\mathfrak{B}$  is a discrete set of cardinality  $\lesssim (R/\rho)^{O(1)}$  such that  $|b| \lesssim R^{1/2+\delta_m}$  for all  $b \in \mathfrak{B}$ . Consequently, by the sub-additivity of the  $k$ -broad norms and Lemma 5.2 and Corollary 9.3, there exists some  $b \in \mathfrak{B}$  such that

$$(10.5) \quad \|\tilde{T}^\lambda \tilde{h}\|_{\text{BL}_{k,A}^p(B(0,\rho))}^p \lesssim (R/\rho)^{O(1)} \|\tilde{T}^\lambda \tilde{h}_b\|_{\text{BL}_{k,A}^p(B(0,\rho))}^p,$$

for  $\tilde{h}_b$  as defined in §9. Recall from Proposition 9.2 that  $\tilde{h}_b$  is concentrated on wave packets in  $\tilde{\mathbb{T}}_{Z-y+b}$  and satisfies

$$(10.6) \quad \|\tilde{h}_b\|_{L^2(B^{n-1})} \lesssim \|f\|_{L^2(B^{n-1})}.$$

Combining (10.4) and (10.5), the desired estimate now follows by applying hypothesis (10.3) to the function  $\tilde{h}_b$  and then using (10.6) to bound the resulting right-hand expression.  $\square$

<sup>18</sup> Strictly speaking, in order for (10.4) to hold the  $k$ -broad norm on the right-hand side should be defined with respect to a translate of the family of balls  $\mathcal{B}_{K^2}$ . Since the estimates will be uniform over all choices of families  $\mathcal{B}_{K^2}$  of bounded multiplicity, this slight technicality does not affect the argument.

**10.3. Setting up the induction argument.** Henceforth it is assumed that  $1 \leq R \lesssim_\varepsilon \lambda^{1-\varepsilon}$ . Under this hypothesis, given  $\varepsilon > 0$  sufficiently small, Proposition 10.1 will be established for the following choice of parameters:

$$(10.7) \quad D_{m,\varepsilon} := \varepsilon^{-\delta^{-(2n-m)}}, \quad \vartheta(\varepsilon) := \varepsilon - c_n \delta_m, \quad \bar{A}_\varepsilon \sim \delta^{-\varepsilon},$$

$$\delta_i = \delta_i(\varepsilon) := \varepsilon^{2i+1} \quad \text{for all } i = 1, \dots, n-1.$$

Here  $0 < \delta = \delta(\varepsilon) \ll \delta_{n-1}(\varepsilon)$  and  $c_n > 0$  is a fixed dimensional constant.

The proof is by induction on the radius  $R$  and the dimension  $m$ ; presently the base cases for this induction are established.

*Base case for the radius:*  $R \lesssim_\varepsilon K^n$ . Provided that the implied constant in (10.1) and  $C_\varepsilon$  are chosen to be sufficiently large, in this case Proposition 10.1 follows immediately from the trivial inequality

$$(10.8) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim R^{n/p} \|f\|_{L^2(B^{n-1})},$$

valid for all  $A \in \mathbb{N}$  and  $1 \leq p \leq \infty$ .

*Base case for the dimension:*  $m \leq k-1$ . Assuming (without loss of generality) that  $K \lesssim_\varepsilon R^{1/n}$ , one can show in this case that

$$(10.9) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} = \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}.$$

Indeed, fix a ball  $B_{K^2} \in \mathcal{B}_{K^2}$  with  $B_{K^2} \cap B(0,R) \neq \emptyset$ ; here  $\mathcal{B}_{K^2}$  denotes the collection of  $K^2$ -balls featured in the definition of the  $k$ -broad norm (1.12). Let  $\mathfrak{T}_{B_{K^2}}$  denote the collection of all  $K^{-1}$ -caps  $\tau$  for which there exists some  $(\theta, v) \in \mathbb{T}_{Z, B_{K^2}}$  with  $\theta \cap \tau \neq \emptyset$ . Observe that if  $\tau \notin \mathfrak{T}_{B_{K^2}}$ , then

$$(10.10) \quad \int_{B_{K^2}} |T^\lambda f_\tau|^p = \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}^p,$$

since  $f$  is concentrated on wave packets in  $\mathbb{T}_Z$ .

It is claimed that there exists some  $V \in \text{Gr}(k-1, n)$ , such that

$$(10.11) \quad \not\prec(G^\lambda(\bar{x}; \tau), V) \leq K^{-1} \quad \text{for all } \tau \in \mathfrak{T}_{B_{K^2}},$$

where  $\bar{x} \in \mathbb{R}^n$  denotes the centre of  $B_{K^2}$ . Indeed, by (10.10) one may assume without loss of generality that  $\mathfrak{T}_{B_{K^2}} \neq \emptyset$  and hence  $\mathbb{T}_{Z, B_{K^2}} \neq \emptyset$ . Thus, there exists  $z \in Z \cap B(0,R)$  with  $|z - \bar{x}| \lesssim R^{1/2+\delta_m}$  and, taking  $V \in \text{Gr}(k-1, n)$  to be any subspace that contains  $T_z Z$ , the claim is easily deduced from the definition of  $R^{1/2+\delta_m}$ -tangency (together with the hypothesis  $K \lesssim R^{1/n}$ ).

Recalling the definition of  $\mu_{T^\lambda f}(B_{K^2})$  from (1.11), it follows from (10.11) that

$$\mu_{T^\lambda f}(B_{K^2}) \leq \max_{\tau \notin V} \int_{B_{K^2}} |T^\lambda f_\tau|^p \leq \max_{\tau \notin \mathfrak{T}_{B_{K^2}}} \int_{B_{K^2}} |T^\lambda f_\tau|^p$$

and the desired estimate (10.9) is now a consequence of (10.10).

*Reduction to  $A \geq 2$ :* Recall that  $\bar{A}_\varepsilon \sim \delta^{-\varepsilon}$ . Although the argument does not require one to induct on  $A$ , it is useful to note that (10.8) implies that Proposition 10.1 holds for  $A = 1$ . This allows one to assume  $A \geq 2$  throughout the following argument, and therefore facilitates the use of the  $k$ -broad triangle and logarithmic-convexity inequalities.

**10.4. An overview of the inductive step.** Let  $2 \leq k \leq n-1$ ,  $k \leq m \leq n$  and  $R \gtrsim_\varepsilon K^n$ . Assume, by way of induction hypothesis, that (10.1) holds whenever the dimension of the transverse complete intersection  $Z$  is at most  $m-1$  or the radial parameter is at most  $R/2$ .

Fix  $\varepsilon > 0$ ,  $1 < A \leq \bar{A}_\varepsilon$  and a transverse complete intersection  $Z = Z(P_1, \dots, P_{n-m})$  with  $\overline{\deg} Z \leq D_{m,\varepsilon}$ , where the parameters  $\bar{A}_\varepsilon$  and  $D_{m,\varepsilon}$  are as defined in (10.7). Let  $f$  be concentrated on wave packets from  $\mathbb{T}_Z$ .

It suffices to show that (10.1) holds at the endpoint  $p = \bar{p}_0(k, m)$ . Indeed, observe that Lemma 5.3 implies the  $L^2$ -bound

$$\|T^\lambda f\|_{\text{BL}_{k,A}^2(B(0,R))}^2 \lesssim \sum_{\tau: K^{-1}\text{-cap}} \|T^\lambda f_\tau\|_{L^2(B(0,R))}^2 \lesssim R \|f\|_{L^2(B^{n-1})}^2.$$

Once (10.1) is established for  $p = \bar{p}_0(k, m)$ , one may use the logarithmic convexity of the  $k$ -broad norms to interpolate the  $p = \bar{p}_0(k, m)$  estimate against the above inequality and thereby obtain (10.1) in the desired range.

The analysis proceeds by considering two different cases.

*The algebraic case.* There exists a transverse complete intersection  $Y^l \subseteq Z$  of dimension  $1 \leq l \leq m-1$  of maximum degree at most  $(D_{m,\varepsilon})^n$  such that

$$(10.12) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(N_{\frac{1}{4}R^{1/2+\delta_m}}(Y^l) \cap B(0,R))}^p \geq c_{\text{alg}} \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p.$$

Here  $c_{\text{alg}} > 0$  is a constant depending only on  $n$  and  $\varepsilon$  which is chosen to be sufficiently small to suit the needs of the forthcoming argument.

*The cellular case.* The negation of the algebraic case: for every transverse complete intersection  $Y^l \subseteq Z$  of dimension  $1 \leq l \leq m-1$  and maximum degree at most  $(D_{m,\varepsilon})^n$  the inequality

$$(10.13) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(N_{\frac{1}{4}R^{1/2+\delta_m}}(Y^l) \cap B(0,R))}^p < c_{\text{alg}} \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p$$

holds.

The cellular case is the simplest situation and will be treated first. Here a polynomial partitioning argument is employed which splits the mass of the  $k$ -broad norm into small pieces; these pieces can then be treated individually using the radial induction hypothesis. The algebraic case is the most involved situation; it encapsulates the kind of behaviour experienced in the sharp examples in §2. In this case  $T^\lambda f$  can be thought of as concentrated near a low-dimensional and low-degree variety  $Y^l$  (in a  $k$ -broad sense). If the wave packets of  $f$  are also tangential to this variety, then one may use induction on the dimension to conclude the argument. This might not be the case, however, and if many of the wave packets of  $f$  are transverse to  $Y^l$ , then an alternative argument is required. Thus, the algebraic case naturally splits into two sub-cases, a tangential and a transverse sub-case, which are discussed in detail below. Lemma 7.3 can be applied to show that a given tube  $T_{\theta,v}$  can only interact transversely with the variety  $Y^l$  on a small number of balls, which eventually allows one to also close the induction in the transverse situation.

**10.5. The cellular case.** The cellular case can be treated using polynomial partitioning. Roughly speaking, since by hypothesis  $T^\lambda f$  is concentrated in a neighbourhood of an  $m$ -dimensional transverse complete intersection, for any  $D \geq 1$  Theorem 7.2 can be applied in  $m$  dimensions to show that there exists a non-zero polynomial  $P$  of degree at most  $D$  such that, letting  $\{O_i\}_{i \in \mathcal{I}}$  denote the connected components of  $\mathbb{R}^n \setminus Z(P)$  (which, recall, are referred to as *cells*), one has  $\#\{O_i : i \in \mathcal{I}\} \sim D^m$  and

$$(10.14) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(O_i)}^p \sim D^{-m} \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \quad \text{for all } i \in \mathcal{I}.$$

Thus, the mass of the  $k$ -broad norm is essentially equally distributed amongst the cells. Moreover, using the hypothesis of the cellular case, one can construct  $P$  so that very little of the mass lies near the cell wall

$$W := N_{2R^{1/2+\delta}}(Z(P)) \cap B(0, R).$$

In particular, the estimate (10.14) essentially still holds if the  $O_i$  are replaced with the shrunken cells  $O'_i := O_i \setminus W$ . The  $O'_i$  can be thought of as well-separated<sup>19</sup> and this facilitates a divide-and-conquer-style argument. More precisely, the fact that a non-zero univariate polynomial of degree at most  $D$  has at most  $D$  roots quickly leads to the following observation.

**Lemma 10.3.** *If  $P$  is a polynomial of degree  $\deg P \leq D$  and  $\{O'_i\}_{i \in \mathcal{I}}$  are defined as above, then each tube  $T_{\theta,v}$  enters at most  $D/\varepsilon$  of the cells  $O'_i$ .*

It is important to note that, in general, Lemma 10.3 does not hold if the  $O'_i$  are replaced with the cells  $O_i$ .

*Proof (of Lemma 10.3).* Let  $[\Gamma_{\theta,v}^\lambda]_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}^{n-1}$  denote the polynomial approximant of  $\Gamma_{\theta,v}^\lambda$ , as defined in §7.2. Thus,  $\deg[\Gamma_{\theta,v}^\lambda]_\varepsilon \leq [1/2\varepsilon]$  and (7.2) implies that

$$|[\Gamma_{\theta,v}^\lambda]_\varepsilon(t) - \Gamma_{\theta,v}^\lambda(t)| \leq R^{1/2} \quad \text{for all } t \in (-R, R).$$

Suppose that  $x \in O'_i \cap T_{\theta,v}$ . By the definition of  $O'_i$ , the ball  $B(x, 2R^{1/2+\delta})$  contains no points of  $Z(P)$ , and is therefore contained in  $O_i$ . On the other hand,  $\text{dist}(x, \Gamma_{\theta,v}^\lambda) < R^{1/2+\delta}$  and therefore  $\text{dist}(x, [\Gamma_{\theta,v}^\lambda]_\varepsilon) < 2R^{1/2+\delta}$ . Consequently,  $[\Gamma_{\theta,v}^\lambda]_\varepsilon$  enters  $B(x, 2R^{1/2+\delta}) \subseteq O_i$ . Thus, if  $T_{\theta,v}$  enters a cell  $O'_i$ , then the polynomial curve  $[\Gamma_{\theta,v}^\lambda]_\varepsilon$  enters  $O_i$  whilst, by the simple property of univariate polynomials quoted above,  $[\Gamma_{\theta,v}^\lambda]_\varepsilon$  can enter at most  $\deg P \cdot \deg[\Gamma_{\theta,v}^\lambda]_\varepsilon + 1 \leq D/\varepsilon$  cells  $O_i$ .  $\square$

Some aspects of the discussion prior to Lemma 10.3 are not entirely precise; for instance, to apply the polynomial partitioning theorem one must work with an  $L^1$  function rather than a  $k$ -broad norm. In view of this, let  $\mu$  denote the measure on  $\mathbb{R}^n$  with Radon–Nikodym derivative

$$\sum_{B_{K^2} \in \mathcal{B}_{K^2}} \mu_{T^\lambda f}(B_{K^2}) \frac{1}{|B_{K^2}|} \chi_{B_{K^2}}$$

with respect to the Lebesgue measure. One may easily verify that

$$(10.15) \quad \mu(U) \leq \|T^\lambda f\|_{\text{BL}_{k,A}^p(U)}^p \quad \text{and} \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \leq \mu(B(0, 2R))$$

<sup>19</sup>In particular, the distance between a pair of distinct  $O'_i$  is wider than the width  $R^{1/2+\delta}$  of any tube  $T_{\theta,v}$ .

for all Lebesgue measurable sets  $U \subseteq \mathbb{R}^n$ . These inequalities ensure that the measure  $\mu$  acts as an effective surrogate for the  $k$ -broad norm in the polynomial partitioning argument.

By combining the cellular hypothesis with Theorem 7.2, one obtains the following partitioning result.

**Lemma 10.4** (Polynomial partitioning [13]). *There exists a polynomial  $P$  of degree  $\deg P \leq D_{m,\varepsilon}$  such that, if  $\{O_i\}_{i \in \mathcal{I}}$  and  $W$  are defined as above and  $O'_i := O_i \setminus W$  for all  $i \in \mathcal{I}$ , then  $\#\mathcal{I} \lesssim (D_{m,\varepsilon})^m$  and*

$$(10.16) \quad \mu(O'_i) \sim (D_{m,\varepsilon})^{-m} \mu(B(0, 2R))$$

for at least 99% of the cells  $O'_i$ .

This lemma is contained in the work of the second author [13, §8.1] and the details of the proof are not reproduced here. The basic idea is as follows: by hypothesis, the mass of  $\mu$  is concentrated in  $N_{R^{1/2+\delta_m}}(Z)$  where  $Z$  is an  $m$ -dimensional algebraic variety; this allows one to apply Theorem 7.2 in  $m$ -dimensions to construct a polynomial  $P$  which satisfies the desired properties with  $O_i$  in place of  $O'_i$ . The hypothesis of the cellular case implies that the mass of  $\mu$  cannot concentrate in a neighbourhood of a certain type of algebraic variety and this can be used to show, in particular, that the mass cannot concentrate around the cell wall  $W$ . Provided the constant  $c_{\text{alg}}$  is chosen to be sufficiently small, this allows one to pass to the shrunken cells  $O'_i$  in (10.16) (at least for 99% of the cells).

There are a number of technicalities involved in rigorously carrying out this argument. In particular, one must justify the application of Theorem 7.2 in dimension  $m$ ; this requires locally approximating  $Z$  by some tangent plane  $T_z Z$  and applying the theorem to the push-forward of  $\mu$  onto  $T_z Z$  under orthogonal projection. The partitioning variety in  $T_z Z$  is lifted to a variety  $\tilde{Z}$  in  $\mathbb{R}^n$  by taking the pre-image under the orthogonal projection; it is possible to define  $\tilde{Z}$  in this way so that it is transverse to  $Z$ . The cells  $O_i$  are then defined with respect to  $\tilde{Z}$ .<sup>20</sup>

Presently, it is shown how together Lemma 10.3 and Lemma 10.4 easily yield the proof of Proposition 10.1 in the cellular case. Applying Lemma 10.4 one obtains a partition of  $\mathbb{R}^n \setminus W$  into disjoint cells  $\{O_i\}_{i \in \mathcal{I}}$ . For each  $i \in \mathcal{I}$  let

$$\mathbb{T}_i := \{(\theta, v) \in \mathbb{T}_Z : T_{\theta, v} \cap O'_i \neq \emptyset\} \quad \text{and} \quad f_i := \sum_{(\theta, v) \in \mathbb{T}_i} f_{\theta, v}.$$

By Lemma 5.2 one has

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(O'_i)}^p \leq \|T^\lambda f_i\|_{\text{BL}_{k,A}^p(O'_i)}^p + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}^p.$$

Combining this inequality with (10.15) and Lemma 10.4, one deduces that at least 99% of the cells  $O'_i$  have the property that

$$(10.17) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim (D_{m,\varepsilon})^m \|T^\lambda f_i\|_{\text{BL}_{k,A}^p(O'_i)}^p + \text{RapDec}(R) \|f\|_{L^2(B^{n-1})}^p.$$

---

<sup>20</sup>To carry out this argument rigorously, one must further ensure that all the relevant varieties are transverse complete intersections of dimension at most  $m-1$  and controlled degree in order to invoke (10.13). Such technicalities account for the choice of maximum degree  $(D_{m,\varepsilon})^n$  in the definition of the algebraic and cellular cases.

On the other hand, by Lemma 10.3 and orthogonality between the  $f_{\theta,v}$ , one has

$$\begin{aligned} \sum_{i \in \mathcal{I}} \|f_i\|_{L^2(B^{n-1})}^2 &\sim \sum_{(\theta,v) \in \mathbb{T}_Z} \#\{i \in \mathcal{I} : (\theta,v) \in \mathbb{T}_i\} \|f_{\theta,v}\|_{L^2(B^{n-1})}^2 \\ &\lesssim_\varepsilon D_{m,\varepsilon} \|f\|_{L^2(B^{n-1})}^2. \end{aligned}$$

Since there are roughly  $(D_{m,\varepsilon})^m$  cells in total, Markov's inequality shows that at least 99% of the cells  $O'_i$  have the property that

$$(10.18) \quad \|f_i\|_{L^2(B^{n-1})}^2 \lesssim_\varepsilon D_{m,\varepsilon}^{-(m-1)} \|f\|_{L^2(B^{n-1})}^2.$$

Therefore, there exists some cell  $O'_i$  for which (10.17) and (10.18) simultaneously hold; henceforth, attention is focused on a single such cell  $O'_i$ .

Let  $E_{m,A}(R)$  denote the constant appearing on the right-hand side of (10.1); namely,

$$E_{m,A}(R) := C_{m,\varepsilon} K^{\bar{C}_\varepsilon} R^{\varepsilon - c_n \delta_m + \delta(\log \bar{A}_\varepsilon - \log A) - e_{k,n}(p) + 1/2}.$$

By the radial induction hypothesis, Proposition 10.1 holds for the radius  $R/2$ . Applying Lemma 10.2 to  $f_i$  for the fixed choice of  $i$  as above and  $\rho = R/2$ , one obtains

$$\|T^\lambda f_i\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim E_{m,A}(R/2) \|f_i\|_{L^2(B^{n-1})} \lesssim E_{m,A}(R) \|f_i\|_{L^2(B^{n-1})}.$$

Combining the above estimate with (10.17) and (10.18), one deduces that

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \leq C_\varepsilon (D_{m,\varepsilon})^{-(m-1)/2 + m/p} E_{m,A}(R) \|f\|_{L^2(B^{n-1})}$$

for some constant  $C_\varepsilon > 0$ . The  $D_{m,\varepsilon}$  exponent is negative if and only if  $p > 2m/(m-1)$ ; note this is the case for the choice of exponent  $p = \bar{p}_0(k, m)$ .<sup>21</sup> Thus, recalling the definition of  $D = D_{m,\varepsilon}$  and assuming  $\varepsilon$  is sufficiently small depending on  $n$ , it follows that  $C_\varepsilon (D_{m,\varepsilon})^{-(m-1)/2 + m/p} \leq 1$ . This establishes the desired estimate (10.1) and closes the induction in the cellular case.

**10.6. The algebraic case.** Fix a transverse complete intersection  $Y^l$  of dimension  $1 \leq l \leq m-1$  and maximum degree  $\overline{\deg} Y^l \leq (D_{m,\varepsilon})^n$  which satisfies (10.12). Let  $1 \leq \rho \ll R$  be such that  $\rho^{1/2 + \delta_l} = R^{1/2 + \delta_m}$  and note that

$$(10.19) \quad R \leq R^{2\delta_l} \cdot \rho \quad \text{and} \quad \rho \leq R^{-\delta_l/2} \cdot R.$$

Let  $\mathcal{B}_\rho$  be a finitely-overlapping cover of  $B(0, R)$  by  $\rho$ -balls and for each  $B \in \mathcal{B}_\rho$  define

$$\mathbb{T}_B := \{(\theta, v) \in \mathbb{T} : T_{\theta,v} \cap N_{\frac{1}{4}R^{1/2 + \delta_m}}(Y^l) \cap B \neq \emptyset\}$$

and  $f_B := \sum_{(\theta,v) \in \mathbb{T}_B} f_{\theta,v}$ . Thus, by (10.12) and Lemma 5.2,

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\rho} \|T^\lambda f_B\|_{\text{BL}_{k,A}^p(N_{\frac{1}{4}R^{1/2 + \delta_m}}(Y^l) \cap B)}^p$$

holds up to the inclusion of a rapidly decreasing error term. In what follows, such error terms, which are harmless, are suppressed in the notation.

For  $B = B(y, \rho) \in \mathcal{B}_\rho$  let  $\mathbb{T}_{B, \text{tang}}$  denote the set of all  $(\theta, v) \in \mathbb{T}_B$  with the property that whenever  $x \in T_{\theta,v}$  and  $z \in Y^l \cap B(y, 2\rho)$  satisfy  $|x-z| \leq 2\bar{C}_{\text{tang}} \rho^{1/2 + \delta_l}$ , it follows that

$$(10.20) \quad \not\prec (G^\lambda(x; \omega_\theta), T_z Y^l) \leq \frac{\bar{C}_{\text{tang}}}{2} \rho^{-1/2 + \delta_l},$$

<sup>21</sup>It is for this reason that one works with the modified exponent  $\bar{p}_0(k, m)$  rather than  $\bar{p}(k, m)$ .

where  $\bar{C}_{\text{tang}}$  and  $\bar{c}_{\text{tang}}$  are the constants appearing in the definition of tangency. Furthermore, let  $\mathbb{T}_{B,\text{trans}} := \mathbb{T}_B \setminus \mathbb{T}_{B,\text{tang}}$  and define

$$f_{B,\text{tang}} := \sum_{(\theta,v) \in \mathbb{T}_{B,\text{tang}}} f_{\theta,v} \quad \text{and} \quad f_{B,\text{trans}} := \sum_{(\theta,v) \in \mathbb{T}_{B,\text{trans}}} f_{\theta,v}.$$

It follows that  $f_B = f_{B,\text{tang}} + f_{B,\text{trans}}$  and, by the triangle inequality for broad norms (that is, Lemma 6.2), one concludes that

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\rho} \|T^\lambda f_{B,\text{tang}}\|_{\text{BL}_{k,A/2}^p(B)}^p + \sum_{B \in \mathcal{B}_\rho} \|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B)}^p.$$

Either the tangential or transverse contribution to the above sum dominates, and each case is treated separately.

**The tangential sub-case.** Suppose that the tangential term dominates; that is,

$$(10.21) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\rho} \|T^\lambda f_{B,\text{tang}}\|_{\text{BL}_{k,A/2}^p(B)}^p.$$

Each term in the right-hand sum is bounded using the dimensional induction hypothesis. Fix  $B = B(y, \rho) \in \mathcal{B}_\rho$  and, as in §9, let  $\tilde{T}^\lambda(f_{B,\text{tang}})^\sim(x) = T^\lambda f_{B,\text{tang}}(x+y)$  so that<sup>22</sup>

$$(10.22) \quad \|T^\lambda f_{B,\text{tang}}\|_{\text{BL}_{k,A/2}^p(B(y,\rho))} = \|\tilde{T}^\lambda(f_{B,\text{tang}})^\sim\|_{\text{BL}_{k,A/2}^p(B(0,\rho))}.$$

Since  $\overline{\text{deg}} Y^l \leq D_{l,\epsilon}$ , in order to apply the induction hypothesis, one must verify that  $(f_{B,\text{tang}})^\sim$  is concentrated on scale  $\rho$  wave packets that are  $\rho^{-1/2+\delta_l}$ -tangent to  $Y^l - y$  in  $B(0, \rho)$ . By Lemma 9.1,  $(f_{B,\text{tang}})^\sim$  is concentrated on scale  $\rho$  wave packets from

$$\tilde{\mathbb{T}}_{B,\text{tang}} := \bigcup_{(\theta,v) \in \mathbb{T}_{B,\text{tang}}} \tilde{\mathbb{T}}_{\theta,v},$$

where the  $\tilde{\mathbb{T}}_{\theta,v}$  are as defined in §9. Fix  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_{B,\text{tang}}$  and recall from (9.15) that

$$(10.23) \quad \text{dist}_H(\tilde{T}_{\tilde{\theta},\tilde{v}}, (T_{\theta,v} - y) \cap B(0, \rho)) \lesssim R^{1/2+\delta} \ll \rho^{1/2+\delta_l}.$$

Combining this with the definition of  $\mathbb{T}_{B,\text{tang}}$ , it is easy to see that  $\tilde{T}_{\tilde{\theta},\tilde{v}}$  satisfies the angle condition for tangency and it remains to verify the containment property  $\tilde{T}_{\tilde{\theta},\tilde{v}} \subseteq N_{\rho^{1/2+\delta_l}}(Y^l - y)$ . The definition of  $\mathbb{T}_B$  and (10.23) imply that  $\tilde{T}_{\tilde{\theta},\tilde{v}} \cap N_{\frac{1}{2}\rho^{1/2+\delta_l}}(Y^l - y) \cap B(0, \rho) \neq \emptyset$  and so the containment property follows from the angle condition, as in the proof of Proposition 9.2.

Thus, the dimensional induction hypothesis may be applied to  $(f_{B,\text{tang}})^\sim$ , and one therefore deduces that

$$\|\tilde{T}^\lambda(f_{B,\text{tang}})^\sim\|_{\text{BL}_{k,A/2}^p(B(0,\rho))} \leq E_{l,A/2}(\rho) \|f_{B,\text{tang}}\|_{L^2(B^{n-1})}.$$

Combining this estimate with (10.21) and (10.22), one concludes that

$$\|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \leq R^{O(\delta_l)} E_{l,A/2}(\rho) \|f\|_{L^2(B^{n-1})}^2.$$

To close the induction in this case, it remains to show that

$$R^{O(\delta_l)} E_{l,A/2}(\rho) \leq E_{m,A}(R).$$

<sup>22</sup>See footnote 18 on page 56.

Clearly,  $\rho^{\delta(\log \bar{A}_\varepsilon - \log(A/2))} \leq R^{\delta(\log \bar{A}_\varepsilon - \log A)}$  and  $\rho^{-e_{k,n}(p)+1/2} \leq R^{O(\delta_l)} R^{-e_{k,n}(p)+1/2}$ , where both inequalities are by (10.19). Combining these observations, the problem is further reduced to showing that  $\rho^{\varepsilon - c_n \delta_l} \leq R^{-c \delta_l} R^{\varepsilon - c_n \delta_m}$ , where  $c > 0$  is a dimensional constant. By (10.19), one may ensure that this inequality holds by choosing the constant  $c_n$  in (10.7) at the outset to be large relative to  $c$ .

**The transverse sub-case.** Now suppose the transverse term dominates; that is,

$$(10.24) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))}^p \lesssim \sum_{B \in \mathcal{B}_\rho} \|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B)}^p.$$

In this case one may exploit the following key inequality.

**Claim.**

$$(10.25) \quad \sum_{B \in \mathcal{B}_\rho} \|f_{B,\text{trans}}\|_{L^2(B^{n-1})}^2 \lesssim_\varepsilon \|f\|_{L^2(B^{n-1})}^2.$$

*Proof.* This is a fairly direct consequence of the hypothesis of the transverse case together with Lemma 7.3. Indeed, note that

$$\sum_{B \in \mathcal{B}_\rho} \|f_{B,\text{trans}}\|_{L^2(B^{n-1})}^2 \sim \sum_{(\theta,v) \in \mathbb{T}} \#\{B \in \mathcal{B}_\rho : (\theta,v) \in \mathbb{T}_{B,\text{trans}}\} \|f_{\theta,v}\|_{L^2(B^{n-1})}^2.$$

and so, to prove (10.25), it suffices to fix an arbitrary  $(\theta,v) \in \mathbb{T}_{B,\text{trans}}$  and show that

$$(10.26) \quad \#\{B \in \mathcal{B}_\rho : (\theta,v) \in \mathbb{T}_{B,\text{trans}}\} \lesssim_\varepsilon 1.$$

Let  $\Gamma := [\Gamma_{\theta,v}^\lambda]_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}^n$  be the polynomial approximant of the core curve  $\Gamma_{\theta,v}^\lambda$  defined in §7.2. Thus,  $\deg \Gamma \lesssim_\varepsilon 1$  and, recalling that  $R \lesssim_\varepsilon \lambda^{1-\varepsilon}$ , property (7.2) of the approximant implies that

$$(10.27) \quad |\Gamma(t) - \Gamma_{\theta,v}^\lambda(t)| \leq R^{1/2} \quad \text{for all } t \in (-R, R).$$

Let  $u \in T_{\theta,v}$  and  $x \in \Gamma \cap B(0, R)$  with  $|u - x| \lesssim R^{1/2+\delta}$ . It follows from the definition of  $T_{\theta,v}$  and (10.27) that there exists some  $t \in (-R, R)$  such that

$$|u - \Gamma_{\theta,v}^\lambda(t)| \lesssim R^{1/2+\delta} \quad \text{and} \quad |x - \Gamma(t)| \lesssim R^{1/2+\delta}.$$

Consequently, recalling Lemma 4.4,

$$\not\prec(G^\lambda(u; \omega_\theta), T_x \Gamma) \lesssim \not\prec(T_{\Gamma_{\theta,v}^\lambda(t)} \Gamma_{\theta,v}^\lambda, T_{\Gamma(t)} \Gamma) + R^{1/2+\delta}/\lambda$$

and therefore, by property (7.3) of the approximant,

$$\not\prec(G^\lambda(u; \omega_\theta), T_x \Gamma) \lesssim_\varepsilon \lambda^{-1/2} + R^{1/2+\delta}/\lambda < \frac{\bar{c}_{\text{tang}}}{4} \rho^{-1/2+\delta_l}.$$

Using the above inequality, one may easily verify that if  $B = B(y, \rho) \in \mathcal{B}_\rho$  and  $(\theta, v) \in \mathbb{T}_{B,\text{trans}}$ , then  $Y_{>\alpha,r,\Gamma}^l \cap B(y, 2\rho) \neq \emptyset$  for  $\alpha \sim \rho^{-1/2+\delta_l}$  and  $r \sim \rho^{1/2+\delta_l}$ . Here  $Y_{>\alpha,r,\Gamma}^l$  is as defined in §7; that is

$$Y_{>\alpha,r,\Gamma}^l := \{z \in Y^l : \exists x \in \Gamma \text{ with } |x - z| < r \text{ and } \not\prec(T_z Y^l, T_x \Gamma) > \alpha\}.$$

By Lemma 7.3, the number of balls  $B = B(y, \rho) \in \mathcal{B}_\rho$  for which  $B(y, 2\rho)$  intersects  $Y_{>\alpha,r,\Gamma}^l$  non-trivially is at most  $O((\deg \Gamma)^n \cdot (\deg Y^l)^n) = O_\varepsilon(1)$ . Combining these observations, one immediately deduces (10.26), as required.  $\square$

In view of (10.25), the strategy in the transverse case is to use the radial induction hypothesis to show that for some constant  $\bar{c}_\varepsilon > 0$  one has

$$(10.28) \quad \|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B)} \leq \bar{c}_\varepsilon E_{m,A}(R) \|f_{B,\text{trans}}\|_{L^2(B^{n-1})} \quad \text{for all } B \in \mathcal{B}_\rho.$$

Indeed, provided  $\bar{c}_\varepsilon > 0$  is sufficiently small, depending only on  $n$  and  $\varepsilon$ , the preceding inequality may be combined with (10.24), (10.25) and the simple estimate

$$\|f_{B,\text{trans}}\|_{L^2(B^{n-1})} \lesssim \|f\|_{L^2(B^{n-1})},$$

to yield

$$\begin{aligned} \|T^\lambda f\|_{\text{BL}_{k,A/2}^p(B(0,R))} &\lesssim \bar{c}_\varepsilon E_{m,A}(R) \|f\|_{L^2(B^{n-1})}^{1-2/p} \left( \sum_{B \in \mathcal{B}_\rho} \|f_{B,\text{trans}}\|_{L^2(B^{n-1})}^2 \right)^{1/p} \\ &\leq E_{m,A}(R) \|f\|_{L^2(B^{n-1})}, \end{aligned}$$

closing the induction in this case.

The main obstacle in carrying out this programme is that the  $f_{B,\text{trans}}$  do not, in general, satisfy the hypothesis of Proposition 10.1 at scale  $\rho$ , and therefore one cannot directly apply the radial induction hypothesis to these functions. However, one can appeal to the theory developed in §9, which essentially allows  $f_{B,\text{trans}}$  to be broken into pieces  $f_{B,\text{trans},b}$  which do satisfy the hypothesis of the proposition at scale  $\rho$ . Here is a sketch of how the argument works. By choosing a suitable set of translates  $\mathfrak{B}$ , one may essentially write

$$(10.29) \quad \|T^\lambda f_{B,\text{trans}}\|_{L^p(B(0,R))} \lesssim \left( \sum_{b \in \mathfrak{B}} \|T^\lambda f_{B,\text{trans},b}\|_{L^p(B(0,R))}^p \right)^{1/p}$$

where each piece  $f_{B,\text{trans},b}$  is defined so that it is concentrated on scale  $\rho$  wave packets which are tangential to some translate  $Z - y + b$  of  $Z$ . By the theory of transverse equidistributions developed in §8 and §9, the  $f_{B,\text{trans},b}$  satisfy favourable  $L^2$  estimates and, in particular, the inequality (10.32) below holds. The radial induction hypothesis is applied to each of the  $T^\lambda f_{B,\text{trans},b}$ . To close the induction, one must estimate the resulting sum

$$\left( \sum_{b \in \mathfrak{B}} \|f_{B,\text{trans},b}\|_{L^p(B^{n-1})}^p \right)^{1/p}$$

in terms of  $\|f_{B,\text{trans}}\|_{L^2(B^{n-1})}$ . Here the gain in  $\rho/R$  in (10.32), afforded by transverse equidistribution, is crucial to the argument: it allows one to sum up the contributions from the individual pieces  $f_{B,\text{trans},b}$  without any (significant) loss in  $R$ . It is this gain which accounts for the improved range of estimates for the  $k$ -broad inequalities under the positive-definite hypothesis (recall, the proof of the transverse equidistribution lemma relied heavily on the positive-definite condition).

As part of this argument, to ensure that the  $f_{B,\text{trans},b}$  form a reasonable decomposition of  $f_{B,\text{trans}}$  so that (10.29) essentially holds, the set of translates  $\mathfrak{B}$  must be chosen so that  $\bigcup_{b \in \mathfrak{B}} N_{\rho^{1/2+\delta_m}}(Z - y + b)$  covers  $N_{R^{1/2+\delta_m}}(Z)$  (recall, by hypothesis  $f_{B,\text{trans}}$  is concentrated on wave packets in  $\mathbb{T}_Z$  and so the mass of  $T^\lambda f_{B,\text{trans}}$  is concentrated in  $N_{R^{1/2+\delta_m}}(Z)$ ) and so that the  $N_{\rho^{1/2+\delta_m}}(Z - y + b)$  are essentially disjoint. This can be achieved using a probabilistic construction. More precisely, fixing  $B = B(y, \rho) \in \mathcal{B}_\rho$ , one may show the following.

**Lemma 10.5.** *There exist a finite set  $\mathfrak{B} \subset B(0, 2R^{1/2+\delta_m})$  and a collection  $\mathcal{B}' \subseteq \{B_{K^2} \in \mathcal{B}_{K^2} : B_{K^2} \cap B(y, \rho) \neq \emptyset\}$  such that, up to inclusion of a rapidly decreasing*

error term,

$$(10.30) \quad \|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B(0,R))} \lesssim (\log R)^2 \left( \sum_{B_{K^2} \in \mathcal{B}'} \mu_{T^\lambda f_{B,\text{trans}}}(B_{K^2}) \right)^{1/p}$$

and for each  $B_{K^2} \in \mathcal{B}'$  the following hold:

i) There exists some  $b \in \mathfrak{B}$  such that

$$(10.31) \quad B_{K^2} \subset N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z+b);$$

ii) There exist at most  $O(1)$  vectors  $b \in \mathfrak{B}$  for which

$$B_{K^2} \cap N_{\rho^{1/2+\delta_m}}(Z+b) \neq \emptyset.$$

The proof of the lemma, which is slightly technical, is postponed until the end of the section. Temporarily assuming this result, one may argue as follows to complete the proof of Proposition 10.1.

For each  $b \in \mathfrak{B}$  let  $\mathcal{B}'_b$  denote the collection of all  $B_{K^2} \in \mathcal{B}'$  for which (10.31) holds. Thus, by (10.30) and property i) in the lemma,

$$\|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B(0,R))} \lesssim (\log R)^2 \left( \sum_{b \in \mathfrak{B}} \sum_{B_{K^2} \in \mathcal{B}'_b} \mu_{\tilde{T}^\lambda(f_{B,\text{trans}})^\sim}(B_{K^2}-y) \right)^{1/p},$$

up to a rapidly decreasing error term.

Define the collection of wave packets

$$\tilde{\mathbb{T}}'_b := \left\{ (\tilde{\theta}, \tilde{v}) \in \bigcup_{(\theta,v) \in \mathbb{T}_{B(y,\rho),\text{trans}}} \tilde{\mathbb{T}}_{\theta,v} : \tilde{\mathbb{T}}_{\tilde{\theta},\tilde{v}} \cap \bigcup_{B_{K^2} \in \mathcal{B}'_b} (B_{K^2}-y) \neq \emptyset \right\};$$

note this set is a subset of the collection  $\tilde{\mathbb{T}}_b$  defined in §9 and so, by Proposition 9.2, one has  $\tilde{\mathbb{T}}'_b \subseteq \tilde{\mathbb{T}}_{Z-y+b}$ . Therefore, if  $f_{B,\text{trans},b}$  is defined by

$$(f_{B,\text{trans},b})^\sim = \sum_{(\tilde{\theta},\tilde{v}) \in \tilde{\mathbb{T}}'_b} (f_{B,\text{trans}})_{\tilde{\theta},\tilde{v}}^\sim,$$

then  $(f_{B,\text{trans},b})^\sim$  is concentrated on wave packets that are  $\rho^{-1/2+\delta_m}$ -tangent to  $Z-y+b$ . Furthermore, again up to a rapidly decreasing error term, one has

$$\|T^\lambda f_{B,\text{trans}}\|_{\text{BL}_{k,A/2}^p(B(0,R))} \lesssim (\log R)^2 \left( \sum_{b \in \mathfrak{B}} \|\tilde{T}^\lambda (f_{B,\text{trans},b})^\sim\|_{\text{BL}_{k,A/2}^p(B(0,\rho))}^p \right)^{1/p}.$$

The function  $(f_{B,\text{trans},b})^\sim$  satisfies the hypotheses of Proposition 10.1 at scale  $\rho$  and therefore the radial induction hypothesis yields

$$\left( \sum_{b \in \mathfrak{B}} \|\tilde{T}^\lambda (f_{B,\text{trans},b})^\sim\|_{\text{BL}_{k,A/2}^p(B(0,\rho))}^p \right)^{1/p} \leq E_{m,A/2}(\rho) \left( \sum_{b \in \mathfrak{B}} \|f_{B,\text{trans},b}\|_{L^2(B^{n-1})}^p \right)^{1/p}.$$

On the other hand, it is claimed that

$$\left( \sum_{b \in \mathfrak{B}} \|f_{B,\text{trans},b}\|_{L^2(B^{n-1})}^p \right)^{1/p} \lesssim R^{O(\delta_m)} (\rho/R)^{(n-m)(1/4-1/2p)} \|f_{B,\text{trans}}\|_{L^2(B^{n-1})}.$$

Clearly it suffices to prove the above inequality for  $p = 2$  and  $p = \infty$ ; the desired estimate for  $p = \bar{p}_0(k, m)$  then follows by interpolation (via Hölder's inequality).

$p = 2$ . Observe that, by the orthogonality between the wave packets,

$$\sum_{b \in \mathfrak{B}} \|f_{B, \text{trans}, b}\|_{L^2(B^{n-1})}^2 \sim \sum_{(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}} \#\mathfrak{B}_{\tilde{\theta}, \tilde{v}} \cdot \|(f_{B, \text{trans}})_{\tilde{\theta}, \tilde{v}}\|_{L^2(B^{n-1})}^2$$

where  $\mathfrak{B}_{\tilde{\theta}, \tilde{v}} := \{b \in \mathfrak{B} : (\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}'_b\}$ . Fixing  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}$ , it suffices to show that  $\#\mathfrak{B}_{\tilde{\theta}, \tilde{v}} \lesssim 1$ . Supposing  $\mathfrak{B}_{\tilde{\theta}, \tilde{v}} \neq \emptyset$ , there exists some  $B_{K^2} \in \mathcal{B}'$  with  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \cap (B_{K^2} - y) \neq \emptyset$ . For any  $b \in \mathfrak{B}_{\tilde{\theta}, \tilde{v}}$  it follows that  $(\tilde{\theta}, \tilde{v}) \in \tilde{\mathbb{T}}_b$  and so  $\tilde{T}_{\tilde{\theta}, \tilde{v}} \subseteq N_{\rho^{1/2+\delta_m}}(Z - y + b)$  by Corollary 9.3. Consequently,  $B_{K^2} \cap N_{\rho^{1/2+\delta_m}}(Z + b) \neq \emptyset$  for all  $b \in \mathfrak{B}_{\tilde{\theta}, \tilde{v}}$  and so the desired bound follows from property ii) of Lemma 10.5.

$p = \infty$ . In this case, the estimate is a direct consequence of the transverse equidistribution estimates established in §8 and §9. In particular, the function  $f_{B, \text{trans}}$  is concentrated on wave packets belonging to  $\mathbb{T}_{Z, B}$  and so, by Lemma 9.9, one deduces that

$$(10.32) \quad \|f_{B, \text{trans}, b}\|_{L^2(B^{n-1})} \lesssim R^{O(\delta_m)} (\rho/R)^{(n-m)/4} \|f_{B, \text{trans}}\|_{L^2(B^{n-1})},$$

as required.

. The preceding analysis shows that  $\|T^\lambda f_{B, \text{trans}}\|_{\text{BL}_{k, A/2}^p(B(0, R))}$  is bounded above by

$$R^{O(\delta_m)} E_{m, A}(\rho) (\rho/R)^{(n-m)(1/4-1/2p)} \|f_{B, \text{trans}}\|_{L^2(B^{n-1})}$$

and therefore, to prove (10.28) and thereby close the induction argument in this case, it suffices to show that

$$(10.33) \quad R^{O(\delta_m)} E_{m, A}(\rho) (\rho/R)^{(n-m)(1/4-1/2p)} \leq \bar{c}_\varepsilon E_{m, A}(R).$$

For the exponent  $p = \bar{p}(k, m)$  one has

$$\rho^{-e_{k, n}(p)+1/2} (\rho/R)^{(n-m)(1/4-1/2p)} \leq R^{-e_{k, n}(p)+1/2}$$

whilst for the perturbed exponent  $p = \bar{p}_0(k, m)$  the same inequality holds up to a  $R^{O(\delta)}$  factor. Thus, the left-hand side of (10.33) is dominated by

$$R^{O(\delta_m)} (\rho/R)^\varepsilon E_{m, A}(R).$$

Recalling (10.19) and the choice of parameters  $\delta_l$  and  $\delta_m$ , one obtains the desired inequality.

**The probabilistic argument.** The above argument establishes Proposition 10.1 except for the details of the probabilistic argument used to prove Lemma 10.5.

*Proof (of Lemma 10.5).* Before commencing the argument proper, a few technical reductions are necessary. By a standard dyadic pigeonholing argument, one may assume that

$$(10.34) \quad \|T^\lambda f_{B, \text{trans}}\|_{\text{BL}_{k, A/2}^p(B(0, R))} \lesssim \log R \left( \sum_{B_{K^2} \in \mathcal{B}''} \mu_{T^\lambda f_{B, \text{trans}}}(B_{K^2}) \right)^{1/p}$$

for some sub-collection  $\mathcal{B}'' \subseteq \mathcal{B}_{K^2}$  with the property that

$$(10.35) \quad \mu_{T^\lambda f_{B, \text{trans}}}(B_{K^2}) \sim \mu_{T^\lambda f_{B, \text{trans}}}(\bar{B}_{K^2}) \quad \text{for all } B_{K^2}, \bar{B}_{K^2} \in \mathcal{B}''.$$

Since  $f_{B, \text{trans}}$  is concentrated on wave packets from  $\mathbb{T}_{Z, B}$ , one may further assume that  $B_{K^2} \cap B(y, \rho) \cap N_{R^{1/2+\delta_m}}(Z) \neq \emptyset$  for all  $B_{K^2} \in \mathcal{B}''$ , at the cost of a rapidly decaying term on the right-hand side of (10.34).

A set of translates  $\mathfrak{B}$  will be selected at random from  $\mathbb{R}^n$  according to a probability measure  $\mathbf{P}$ . The distribution  $\mathbf{P}$  is taken to be a mollified version of the uniform probability distribution  $\mathbf{P}_{\text{unif}}$  on  $B(0, R^{1/2+\delta_m})$ . In particular, let  $\omega: \mathbb{R}^n \rightarrow [0, \infty)$  be given by<sup>23</sup>

$$\omega(x) := \exp\left(\frac{-(|x| - R^{1/2+\delta_m})_+}{\rho^{1/2+\delta_m}}\right) \quad \text{for all } x \in \mathbb{R}^n$$

and  $\mathbf{P}$  be the continuous probability measure on  $\mathbb{R}^n$  with Radon–Nikodym derivative  $(\int_{\mathbb{R}^n} \omega)^{-1} \omega$  (with respect to Lebesgue measure). This measure approximates  $\mathbf{P}_{\text{unif}}$  in the sense that

$$(10.36) \quad \mathbf{P}(\mathbb{R}^n \setminus B(0, 2R^{1/2+\delta_m})) = \text{RapDec}(R).$$

The motivation behind the definition of  $\mathbf{P}$  is that, in contrast with  $\mathbf{P}_{\text{unif}}$ , it enjoys the doubling property

$$\mathbf{P}(B(x, 2r)) \lesssim \mathbf{P}(B(x, r)) \quad \text{for all } x \in \mathbb{R}^n \text{ and } 0 < r \lesssim \rho^{1/2+\delta_m}.$$

Consequently, by the Vitali covering lemma, for any  $E \subseteq \mathbb{R}^n$  one has

$$(10.37) \quad \mathbf{P}(N_{2r}(E)) \lesssim \mathbf{P}(N_r(E)) \quad \text{for all } 0 < r \lesssim \rho^{1/2+\delta_m}.$$

Recall, if  $B(x, K^2) \in \mathcal{B}''$ , then  $B(x, K^2) \cap N_{R^{1/2+\delta_m}}(Z) \neq \emptyset$  and so

$$|B(0, R^{1/2+\delta_m}) \cap N_{\rho^{1/2+\delta_m}}(Z - x)| \gtrsim |B(0, \rho^{1/2+\delta_m})|$$

which implies that

$$\mathbf{P}(N_{\rho^{1/2+\delta_m}}(Z - x)) \gtrsim \frac{|B(0, \rho^{1/2+\delta_m})|}{|B(0, R^{1/2+\delta_m})|}.$$

For any  $s \in \mathbb{N}$  with  $2^s \gtrsim |B(0, \rho^{1/2+\delta_m})|$ , define

$$\mathcal{B}^s := \left\{ B(x, K^2) \in \mathcal{B}'' : \mathbf{P}(N_{\rho^{1/2+\delta_m}}(Z - x)) \sim \frac{2^s}{|B(0, R^{1/2+\delta_m})|} \right\}.$$

By a further pigeonholing argument, there exists some value of  $s$  as above such that (10.30) holds with  $\mathcal{B}^s$  in place of  $\mathcal{B}'$ .

Let  $\bar{C} \geq 1$  be a dimensional constant, chosen to be sufficiently large for the purposes of the following argument, such that  $N := \bar{C}2^{-s}|B(0, R^{1/2+\delta_m})| \in \mathbb{N}$ . Recalling (10.19), it follows that

$$(10.38) \quad N \lesssim \frac{|B(0, R^{1/2+\delta_m})|}{|B(0, \rho^{1/2+\delta_m})|} \lesssim R^{2n\delta_l}.$$

Suppose  $\mathfrak{B} = \{b_1, \dots, b_N\}$  is a sequence of vectors in  $\mathbb{R}^n$  formed by choosing each term independently at random according to the probability distribution  $\mathbf{P}$ . The problem is to show that  $\mathfrak{B}$  satisfies each of the desired properties with high probability.

---

<sup>23</sup>Here  $(u)_+ := \begin{cases} u & \text{if } u \geq 0 \\ 0 & \text{if } u < 0 \end{cases}$  for all  $u \in \mathbb{R}$ .

The containment property  $\mathfrak{B} \subset B(0, 2R^{1/2+\delta_m})$ . Recalling (10.36) and (10.38), it follows that

$$\begin{aligned} \mathbf{P}(\mathfrak{B} \subseteq B(0, 2R^{1/2+\delta_m})) &= 1 + \sum_{k=1}^N \binom{N}{k} (-1)^k \mathbf{P}(\mathbb{R}^n \setminus B(0, 2R^{1/2+\delta_m}))^k \\ &= 1 + \text{RapDec}(R). \end{aligned}$$

Thus, if  $R \geq 1$  is sufficiently large depending only on  $n$  and  $\varepsilon$ , then

$$(10.39) \quad \mathbf{P}(\mathfrak{B} \subseteq B(0, 2R^{1/2+\delta_m})) \geq \frac{99}{100},$$

which verifies that the desired containment property holds with high probability.

*Property i).* Let  $B(x, K^2) \in \mathcal{B}^s$  and observe that

$$\begin{aligned} \mathbf{P}\left(B(x, K^2) \subseteq \bigcup_{j=1}^N N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z + b_j)\right) &\geq \mathbf{P}\left(x \in \bigcup_{j=1}^N N_{\frac{1}{4}\rho^{1/2+\delta_m}}(Z + b_j)\right) \\ &= 1 - \left(1 - \mathbf{P}\left(N_{\frac{1}{4}\rho^{1/2+\delta_m}}(Z - x)\right)\right)^N. \end{aligned}$$

By the definition of  $\mathcal{B}^s$  and the doubling property (10.37) of  $\mathbf{P}$ , it follows that  $\mathbf{P}\left(N_{\frac{1}{4}\rho^{1/2+\delta_m}}(Z - x)\right) = c\bar{C}/N$  for some dimensional constant  $c > 0$  and, consequently,

$$\mathbf{P}\left(B(x, K^2) \subseteq \bigcup_{j=1}^N N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z + b_j)\right) \geq 1 - (1 - c\bar{C}/N)^N > 1 - e^{-c\bar{C}}.$$

Let  $X$  denote the random variable that counts the number of  $B_{K^2} \in \mathcal{B}^s$  for which  $B_{K^2} \subseteq N_{\frac{1}{2}\rho^{1/2+\delta_m}}(Z + b)$  for some  $b \in \mathfrak{B}$ . If  $\bar{C}$  is suitably chosen, then the above inequality implies that the expected value of  $X$  satisfies  $\mathbf{E}[X] \geq (1 - 10^{-4})\#\mathcal{B}^s$ . By Markov's inequality,

$$(10.40) \quad \mathbf{P}\left(X > \frac{99}{100}\#\mathcal{B}^s\right) \geq 1 - \frac{100}{\#\mathcal{B}^s}\mathbf{E}[\#\mathcal{B}^s - X] \geq \frac{99}{100},$$

which verifies that property i) of the lemma holds with high probability.

*Property ii).* For each  $x \in \mathbb{R}^n$  let  $M_x$  denote the random variable that counts the number of sets  $N_{\rho^{1/2+\delta_m}}(Z + b_j)$  containing  $x$ ; that is,

$$M_x(b_1, \dots, b_N) := \sum_{j=1}^N \chi_{N_{\rho^{1/2+\delta_m}}(Z + b_j)}(x).$$

If  $B(x, K^2) \in \mathcal{B}^s$ , then

$$\mathbf{E}[M_x] = \sum_{j=1}^N \mathbf{P}(N_{\rho^{1/2+\delta_m}}(Z - x)) \sim N \frac{2^s}{|B(0, R^{1/2+\delta_m})|} \sim 1.$$

Now let  $C \geq 1$  be a dimensional constant and  $Y$  denote the random variable that counts the number of  $B(x, K^2) \in \mathcal{B}^s$  for which  $M_x \leq C$ . By a two-fold application of Markov's inequality, if  $C$  is chosen to be sufficiently large, then

$$(10.41) \quad \mathbf{P}\left(Y > \frac{99}{100}\#\mathcal{B}^s\right) \geq \mathbf{P}\left(\#\mathcal{B}^s - \frac{1}{C} \sum_{B(x, K^2) \in \mathcal{B}^s} M_x > \frac{99}{100}\#\mathcal{B}^s\right) \geq \frac{99}{100},$$

which verifies that property ii) of the lemma holds with high probability.

In view of (10.39), (10.40) and (10.41), there exists a choice of  $\mathfrak{B} \subseteq B(0, 2R^{1/2+\delta_m})$  and a subset  $\mathcal{B}' \subseteq \mathcal{B}^s$  of cardinality comparable to that of  $\mathcal{B}^s$  for which the desired properties i) and ii) hold. Finally, by (10.35), the inequality (10.30) also holds for the sub-collection  $\mathcal{B}'$ .  $\square$

## 11. GOING FROM $k$ -BROAD TO LINEAR ESTIMATES

**11.1. Applying the Bourgain–Guth method.** Theorem 1.2 can be deduced as a consequence of the  $k$ -broad estimates via the method of [8]. The key proposition is as follows.

**Proposition 11.1.** *Suppose that for all  $K \geq 1$  and all  $\varepsilon > 0$  any Hörmander-type operator  $T^\lambda$  with reduced positive-definite phase obeys the  $k$ -broad inequality*

$$(11.1) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

for some fixed  $k, A, p, q, C_\varepsilon$  and all  $R \geq 1$ . If

$$2 \cdot \frac{2n-k+2}{2n-k} \leq p \leq 2 \cdot \frac{k-1}{k-2},$$

then any Hörmander-type operator  $T^\lambda$  with positive-definite phase satisfies

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_{\phi,\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}.$$

Theorem 1.2 is now a direct consequence of Proposition 11.1 and Theorem 1.6.

*Proof (of Theorem 1.2).* Theorem 1.6 implies that for each  $2 \leq k \leq n$  the estimate (11.1) is valid for all  $p \geq \bar{p}(n, k)$ . Thus, for each  $k$  one may apply Proposition 11.1 with  $\bar{p}(n, k) \leq p \leq 2(k-1)/(k-2)$  to obtain a (potentially empty) range of estimates for the linear problem. Since  $\bar{p}(n, k)$  is a decreasing function of  $k$ , the optimal estimate is given by applying Proposition 11.1 as above with  $k$  chosen to be as large as possible subject to the constraint

$$2 \cdot \frac{2n-k+2}{2n-k} \leq \bar{p}(n, k) = 2 \cdot \frac{n+k}{n+k-2}.$$

Rearranging this inequality yields  $k \leq n/2 + 1$ . Hence, defining  $k_* := n/2 + 1$  for  $n$  even and  $k_* := (n+1)/2$  for  $n$  odd, the linear estimate holds for all  $p \geq \bar{p}(n, k_*)$ . A simple computation shows that this corresponds to the range of estimates stated in Theorem 1.2.  $\square$

For contrast, it is noted that there is also a version of Proposition 11.1 which holds without the positive-definite assumption. This can be combined with the multilinear estimates of Bennett–Carbery–Tao [4] to prove Theorem 1.1 (this is essentially the argument used in [8]).

**Proposition 11.2.** *Suppose that for all  $K \geq 1$  and all  $\varepsilon > 0$  any Hörmander-type operator  $T^\lambda$  with reduced phase obeys the  $k$ -broad inequality*

$$(11.2) \quad \|T^\lambda f\|_{\text{BL}_{k,A}^p(B(0,R))} \lesssim_\varepsilon K^{C_\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

for some fixed  $k, A, p, q, C_\varepsilon$  and all  $R \geq 1$ . If

$$2 \cdot \frac{n-k+2}{n-k+1} \leq p,$$

then any Hörmander-type operator  $T^\lambda$  satisfies

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_{\phi,\varepsilon} R^\varepsilon \|f\|_{L^p(B^{n-1})}.$$

Theorem 1.1 is now a direct consequence of Proposition 11.2 and the Bennett–Carbery–Tao theorem.

*Proof (of Theorem 1.1).* The proof is precisely the same as that of Theorem 1.2 above, but with the exponent  $2k/(k-1)$  from the Bennett–Carbery–Tao theorem (that is, Theorem 1.5 or, more precisely, the  $k$ -broad version given by Corollary 6.5) playing the rôle of  $\bar{p}(n, k)$ .  $\square$

To establish the main result, Theorem 1.2, it remains to prove Proposition 11.1. Both Proposition 11.1 and Proposition 11.2 can be established using very similar arguments: in fact, the proofs differ only at one (crucial) point. To highlight the essential differences between the positive-definite and mixed-signature cases, at the end of this subsection it is indicated how one may adapt the proof of Proposition 11.1 to yield Proposition 11.2.

The proof of Proposition 11.1 is an induction on scale argument. The induction quantity is defined as follows.

**Definition 11.3.** For  $1 \leq p \leq \infty$  and  $R \geq 1$  let  $Q_p(R)$  denote the infimum over all constants  $C$  for which the estimate

$$\|T^\lambda f\|_{L^p(B(0,r))} \leq C \|f\|_{L^p(B^{n-1})}$$

holds for  $1 \leq r \leq R$  and all Hörmander-type operators  $T^\lambda$  with reduced positive-definite phase and all  $\lambda \geq R$ .

With this definition, the problem is now to show that, under the hypotheses of Proposition 11.1, one has

$$(11.3) \quad Q_p(R) \lesssim_\varepsilon R^\varepsilon$$

for all  $\varepsilon > 0$  and  $1 \leq R \leq \lambda$ . Indeed, this establishes the linear estimates in the case of reduced phases, and then the arguments of §4 extend the result to general Hörmander-type operators with positive-definite phase.

It is useful to introduce some of the ingredients of the proof of (11.3). Decompose  $B(0, R)$  into balls  $B_{K^2}$  of radius  $K^2$  and consider  $\|T^\lambda f\|_{L^p(B_{K^2})}$  for some fixed  $B_{K^2}$  with centre  $\bar{x}$ . To bound this quantity one expresses  $f$  as a sum of two terms: a “narrow” and a “broad” term. The narrow term is of the form

$$(11.4) \quad \sum_{\tau \in V_a \text{ for some } a} f_\tau,$$

consisting of contributions to  $f$  from caps for which  $G^\lambda(\bar{x}; \tau)$  makes a small angle with some member of a family of  $(k-1)$ -planes. The broad term consists of the contributions to  $f$  from all the remaining caps. One may choose the planes  $V_1, \dots, V_A$  so that the broad term can be bounded by the  $k$ -broad inequality from the hypothesis. Thus, the problem is roughly reduced to studying the case where  $f$  is of the form (11.4). To treat this case, the first step is to apply an  $\ell^p$ -decoupling inequality to isolate the contributions of the different  $f_\tau$ .

**Theorem 11.4.** *Suppose that  $T^\lambda$  is a Hörmander-type operator with reduced positive-definite phase. If  $V \subseteq \mathbb{R}^n$  is an  $m$ -dimensional linear subspace, then for  $2 \leq p \leq 2m/(m-1)$  and  $\delta > 0$  one has*

$$\left\| \sum_{\tau \in V} T^\lambda g_\tau \right\|_{L^p(B_{K^2})} \lesssim_\delta K^{(m-1)(1/2-1/p)+\delta} \left( \sum_{\tau \in V} \|T^\lambda g_\tau\|_{L^p(w_{B_{K^2}})}^p \right)^{1/p}.$$

Here the sums are over all caps  $\tau$  for which  $\not\prec(G^\lambda(\bar{x}, \tau), V) \leq K^{-1}$  where  $\bar{x}$  is the centre of  $B_{K^2}$  and  $w_{B_{K^2}}$  is a rapidly decaying weight of the form of that defined in (8.6).

This theorem is a variable coefficient generalisation of a decoupling inequality due to Bourgain [6]. It can be established by adapting the argument of [6] using many of the techniques employed in the current article: see also [2].<sup>24</sup>

Summing together the contributions from the various spatial balls  $B_{K^2}$ , it remains to estimate the decoupled contributions  $\|T^\lambda f_\tau\|_{L^p(B_R)}$ . Since each  $f_\tau$  has small support, after rescaling one obtains favourable estimates for  $\|T^\lambda f_\tau\|_{L^p(B(0,R))}$  by invoking the induction hypothesis. This is made precise by the following lemma.

**Lemma 11.5** (Parabolic rescaling). *Let  $1 \leq R \leq \lambda$  and suppose  $f$  is supported on a ball of radius  $\rho^{-1}$  where  $1 \leq \rho \leq R$ . For all  $p \geq 2$  and  $\delta > 0$  one has*

$$\|T^\lambda f\|_{L^p(B(0,R))} \lesssim_\delta Q_p(R) R^\delta \rho^{2n/p-(n-1)} \|f\|_{L^p(B^{n-1})}.$$

The proof of the parabolic rescaling lemma for extension operators is simple, consisting of an affine change of variables. In the variable coefficient case some significant additional complications arise; the proof is therefore postponed until the following subsection.

Having introduced the main tools, the proof of Proposition 11.1 follows easily.

*Proof (of Proposition 11.1).* It suffices to demonstrate the linear estimate for  $p$  satisfying the additional condition

$$(11.5) \quad 2 \cdot \frac{2n-k+2}{2n-k} < p;$$

the result for the remaining value of  $p$  then follows immediately by Hölder's inequality.

Let  $\varepsilon > 0$  be given. By hypothesis,

$$\sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \emptyset}} \min_{V_1, \dots, V_A} \max_{\tau \notin V_a} \int_{B_{K^2}} |T^\lambda f_\tau|^p \leq C(K, \varepsilon) R^{p\varepsilon/2} \|f\|_{L^p(B^{n-1})}^p$$

where  $V_1, \dots, V_A$  are  $(k-1)$ -planes and the notation  $\tau \notin V_a$  signifies that  $\not\prec(G^\lambda(\bar{x}, \tau), V_a) > K^{-1}$  for  $\bar{x}$  the centre of the corresponding  $K^2$ -ball  $B_{K^2}$ .

For each  $B_{K^2}$  fix a choice of  $V_1, \dots, V_A$  which achieves the minimum above. Then one may write

$$\int_{B_{K^2}} |T^\lambda f|^p \lesssim K^{O(1)} \max_{\tau \notin V_a} \int_{B_{K^2}} |T^\lambda f_\tau|^p + \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p.$$

<sup>24</sup>It is remarked that since the decoupling estimate is applied at a small spatial scale  $K^2 \ll \lambda^{1/2}$  one can avoid the use of the full statement of Theorem 11.4 by appealing to an approximation argument. If one argues in this way, then only Theorem 11.4 for extension operators associated to elliptic-type hypersurfaces is required.

The first term can be estimated using the hypothesised  $k$ -broad estimate; in particular,

$$\int_{B(0,R)} |T^\lambda f|^p \lesssim K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} \|f\|_{L^p(B^{n-1})}^p + \sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \emptyset}} \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p.$$

It remains to bound the narrow term, where the contributions come from caps whose directions make a small angle with one of planes  $V_a$ . By Theorem 11.4, for any  $\delta' > 0$  one has

$$\int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p \lesssim_{\delta'} K^{(k-2)(p/2-1)+\delta'} \sum_{\tau \in V_a} \int_{\mathbb{R}^n} |T^\lambda f_\tau|^p w_{B_{K^2}}$$

for each  $1 \leq a \leq A$ . Thus, summing over the  $a$  and all the relevant balls  $B_{K^2}$ , one concludes that

$$\sum_{\substack{B_{K^2} \in \mathcal{B}_{K^2} \\ B_{K^2} \cap B(0,R) \neq \emptyset}} \sum_{a=1}^A \int_{B_{K^2}} \left| \sum_{\tau \in V_a} T^\lambda f_\tau \right|^p \lesssim_{\delta'} K^{(k-2)(p/2-1)+\delta'} \sum_{\tau: K^{-1}\text{-cap}} \int_{B(0,2R)} |T^\lambda f_\tau|^p.$$

Since each  $f_\tau$  is supported on a  $K^{-1}$ -cap, the summands appearing in the right-hand expression are amenable to parabolic rescaling. In particular, letting  $\delta > 0$  be a small number chosen to satisfy the requirements of the forthcoming argument, Lemma 11.5 implies that

$$\int_{B(0,2R)} |T^\lambda f_\tau|^p \lesssim_\delta Q_p(R)^p R^\delta K^{2n-(n-1)p} \|f_\tau\|_{L^p(B^{n-1})}^p.$$

Defining

$$e(k, p) := (k-2)(1-p/2) - 2n + (n-1)p$$

and combining these estimates,

$$\int_{B(0,R)} |T^\lambda f|^p \leq (K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} + C_{\delta, \delta'} Q_p(R)^p R^\delta K^{-e(k,p)+\delta'}) \|f\|_{L^p(B^{n-1})}^p$$

and so, by definition,

$$Q_p(R)^p \leq K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} + C_{\delta, \delta'} Q_p(R)^p R^\delta K^{-e(k,p)+\delta'}.$$

Since  $p$  satisfies (11.5), it follows that  $e(p, k) > 0$  and one may choose  $\delta' = e(k, p)/2$  so that the  $K$  exponent in the right-hand term is negative. Thus, if  $K = K_0 R^{2\delta/e(k,p)}$  for a sufficiently large constant  $K_0$ , depending only on  $\varepsilon, \delta, p$  and  $n$ , it follows that

$$Q_p(R)^p \leq K^{O(1)} C(K_0 R^{2\delta/e(k,p)}, \varepsilon) R^{p\varepsilon/2} + Q_p(R)^p/2.$$

Recall that, by hypothesis, the constant  $C(K, \varepsilon)$  arising from the  $k$ -broad estimate grows at most polynomially in  $K$ . Consequently, one may choose  $\delta$  to be small enough (depending only on admissible parameters) so that  $Q_p(R) \lesssim_\varepsilon R^\varepsilon$ , as required.  $\square$

As mentioned above, this argument can be adapted to study the case of general Hörmander-type operators (with potentially mixed signature) to prove Proposition 11.2. The induction quantity  $Q_p(R)$  is defined as before, but now the supremum is taken over the larger class of all Hörmander-type operators  $T^\lambda$  which are in a suitably reduced form. The proof of the parabolic rescaling lemma then extends to

this setting *mutatis mutandis*. The key differences arise in the decoupling inequality. In particular, Theorem 11.4 does not hold at the required level of generality. To see why this is so, consider the example of the extension operator  $E$  associated to (a compact piece of) the hyperbolic paraboloid given by the graph of  $h(\omega) := \omega_1\omega_2$ . If  $V := \{x \in \mathbb{R}^3 : x_1 = 0\}$  and  $G$  is the relevant Gauss map, then

$$S_\omega := \{\omega \in B^2 : G(\omega) \in V\} = \{\omega \in B^2 : \omega_2 = 0\}.$$

Thus, the  $Eg_\tau$  for  $\tau \in V$  are (distributionally) Fourier supported in a neighbourhood of the  $\xi_1$ -axis (which is, in particular, a curve of everywhere zero curvature). As is well-known, in the absence of curvature, no non-trivial decoupling estimates are possible.

The following simple result acts as a substitute for Theorem 11.4.

**Lemma 11.6** (Bourgain–Guth [8]). *Suppose that  $T^\lambda$  is a Hörmander-type operator with reduced phase. If  $V \subseteq \mathbb{R}^n$  is an  $m$ -dimensional linear subspace, then for all  $p \geq 2$  and  $\delta > 0$  one has*

$$\left\| \sum_{\tau \in V} T^\lambda g_\tau \right\|_{L^p(B_{K^2})} \lesssim_\delta K^{(m-1)(1-2/p)+\delta} \left( \sum_{\tau \in V} \|T^\lambda g_\tau\|_{L^p(w_{B_{K^2}})}^p \right)^{1/p}.$$

This lemma provides much weaker estimates than those guaranteed by the  $\ell^p$ -decoupling theorem in the positive-definite case: here the  $K$  exponent is larger by a factor of 2 than that appearing in Theorem 11.4. To prove Proposition 11.2 one proceeds as in the proof of Proposition 11.1, first decomposing  $B(0, R)$  into balls of radius  $K^2$ . For each such ball the broad term is bounded using the hypothesised  $k$ -broad estimate whilst the narrow term is bounded by Lemma 11.6 together with the induction hypothesis (via parabolic rescaling). The larger exponent incurred by Lemma 11.6 propagates through the argument until one arrives at the estimate

$$Q_p(R)^p \leq K^{O(1)} C(K, \varepsilon) R^{p\varepsilon/2} + C_{\delta, \delta'} Q_p(R)^p R^\delta K^{(k-2)(p-2)+2n-(n-1)p+\delta'}.$$

In order to close the induction, once again one must ensure that the  $K$  exponent is negative. By choosing  $\delta'$  appropriately, this is possible if  $p$  satisfies the stronger hypothesis  $p > 2(n-k+2)/(n-k+1)$ , which is precisely the condition featured in the statement of Proposition 11.2.

**11.2. Proof of Lemma 11.5.** It remains to establish the parabolic rescaling lemma, which is achieved by adapting arguments implicit in [8]. As mentioned in the previous section, some additional complications arise in the case of Hörmander operators (as opposed to the extension case) and the proof of the parabolic rescaling is slightly involved.

It will be useful to work with the following discrete reformulation of the main estimate for the operator  $T^\lambda$ .

**Lemma 11.7.** *If  $\mathcal{D}$  is a maximal  $R^{-1}$ -separated discrete subset of  $\Omega$ , then*

$$(11.6) \quad \left\| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(\cdot; \omega_\theta)} G(\omega_\theta) \right\|_{L^p(B(0, R))} \lesssim Q_p(R) R^{(n-1)/p'} \|G\|_{\ell^p(\mathcal{D})}$$

for all  $G: \mathcal{D} \rightarrow \mathbb{C}$ .

*Proof.* Fix  $\psi \in C_c^\infty(\mathbb{R}^{n-1})$  supported on  $B(0, 2)$  which satisfies  $0 \leq \psi \leq 1$  and  $\psi(\omega) = 1$  for all  $\omega \in B^{n-1}$  and for each  $\omega_\theta \in \mathcal{D}$  define  $\psi_\theta(\omega) := \psi(10R(\omega - \omega_\theta))$ .

Thus, for all  $x \in B(0, R)$  the exponential sum appearing in the left-hand side of (11.6) can be expressed as a constant multiple of

$$R^{n-1} \int_{\mathbb{R}^{n-1}} e^{2\pi i \phi^\lambda(x; \omega)} a^\lambda(x; \omega) [\tilde{\psi}(x/R) \sum_{\omega_\theta \in \mathcal{D}} e^{-2\pi i \lambda \Omega_\theta(x/\lambda; \omega)} G(\omega_\theta) \psi_\theta(\omega)] d\omega$$

where  $\Omega_\theta(x; \omega) := \phi(x; \omega) - \phi(x; \omega_\theta)$ ,  $\tilde{\psi} \in C_c^\infty(\mathbb{R}^n)$  is a function of  $n$  variables which enjoys properties similar to those of  $\psi$  and  $a^\lambda$  is a suitable choice of amplitude. Since

$$\sup_{\omega \in \text{supp } \psi_\theta} |\partial_x^\beta \Omega_\theta(x; \omega)| \lesssim_\beta R^{-1} |x| \quad \text{for all } \beta \in \mathbb{N}_0^n \text{ and } x \in X,$$

one may safely remove the  $\lambda \Omega_\theta(x/\lambda; \omega)$  term from the phase. More precisely, by expanding  $\tilde{\psi}(x) e^{-2\pi i \lambda \Omega_\theta(x/\lambda; \omega)}$  as a Fourier series in the variable  $x$ , one can show that

$$\left| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(x; \omega_\theta)} G(\omega_\theta) \right| \lesssim R^{n-1} \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-(n+1)} |T^\lambda g_k(x)|$$

where  $T^\lambda$  is a Hörmander-type operator with phase  $\phi^\lambda$  and

$$g_k(\omega) := \sum_{\omega_\theta \in \mathcal{D}} G(\omega_\theta) c_{k, \theta}(\omega) \psi_\theta(\omega)$$

for some choice of smooth functions  $c_{k, \theta}$  satisfying the uniform bound  $\|c_{k, \theta}\|_{L^\infty(B^{n-1})} \lesssim 1$ . Thus, by the definition of  $Q_p(R)$  it follows that

$$\left\| \sum_{\omega_\theta \in \mathcal{D}} e^{2\pi i \phi^\lambda(\cdot; \omega_\theta)} G(\omega_\theta) \right\|_{L^p(B(0, R))} \lesssim Q_p(R) R^{n-1} \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-(n+1)} \|g_k\|_{L^p(B^{n-1})}$$

and, since the  $\psi_\theta$  are supported on pairwise disjoint sets,

$$\|g_k\|_{L^p(B^{n-1})} \lesssim R^{-(n-1)/p'} \left( \sum_{\omega_\theta \in \mathcal{D}} |G(\omega_\theta)|^p \right)^{1/p},$$

concluding the proof.  $\square$

*Proof (of Lemma 11.5).* Recall, the phase of  $T^\lambda$  is given by  $\phi^\lambda(x; \omega) := \lambda \phi(x/\lambda; \omega)$  where

$$(11.7) \quad \phi(x; \omega) = \langle x', \omega \rangle + x_n h(\omega) + \mathcal{E}(x; \omega).$$

Let  $B(\bar{\omega}, \rho^{-1})$  be a ball supporting  $f$ , where  $\bar{\omega} \in B^{n-1}$ . If  $\tilde{T}^{\lambda/\rho^2}$  denotes the parabolically rescaled operator defined in (4.8), with rescaled phase function

$$(11.8) \quad \tilde{\phi}(x; \omega) := \langle x', \omega \rangle + x_n \tilde{h}(\omega) + \tilde{\mathcal{E}}(x; \omega),$$

then it follows that

$$\|T^\lambda f\|_{L^p(B(0, R))} \lesssim \rho^{(n+1)/p} \|\tilde{T}^{\lambda/\rho^2} \tilde{f}\|_{L^p(\tilde{D}_R)}$$

where now  $\tilde{D}_R$  is an ellipse with principal axes parallel to the co-ordinate axes and dimensions  $O(R/\rho) \times \cdots \times O(R/\rho) \times O(R/\rho^2)$  and  $\tilde{f}(\omega) := \rho^{-(n-1)} f(\bar{\omega} + \rho^{-1}\omega)$ . Since

$$\|\tilde{f}\|_{L^p(B^{n-1})} = \rho^{-(n-1)+(n-1)/p} \|f\|_{L^p(B^{n-1})},$$

given  $\delta > 0$ , the problem is to show that

$$\|\tilde{T}^{\lambda/\rho^2} \tilde{f}\|_{L^p(\tilde{D}_R)} \lesssim_\delta Q_p(R) R^\delta \|\tilde{f}\|_{L^p(B^{n-1})}.$$

Observe that the phase  $\tilde{\phi}$  defined in (11.8) is also positive-definite and of reduced form. To lighten the notation, consider once again a general positive-definite reduced phase  $\phi$  as in (11.7) and let  $T^\lambda$  is a Hörmander-type operator associated to  $\phi^\lambda$ . It suffices to show

$$\|T^\lambda f\|_{L^p(D_{\mathbf{R}})} \lesssim_\delta Q_p(R) R^\delta \|f\|_{L^p(B^{n-1})}$$

for all  $1 \ll R \leq R' \leq \lambda$  and  $\delta > 0$  where

$$D_{\mathbf{R}} := \left\{ x \in \mathbb{R}^n : \left( \frac{|x'|}{R'} \right)^2 + \left( \frac{|x_n|}{R} \right)^2 \leq 1 \right\}$$

is an ellipse. Of course, if  $R = R'$ , then this inequality is immediate from the definition of  $Q_p(R)$ .

Cover  $B^{n-1}$  by a collection of essentially disjoint  $R^{-1}$ -caps  $\theta$  and decompose  $f$  as  $f = \sum_\theta f_\theta$ . Define

$$T_\theta^\lambda f(x) := e^{-2\pi i \phi^\lambda(x; \omega_\theta)} T_\theta^\lambda f_\theta(x)$$

so that

$$T^\lambda f(x) = \sum_{\theta: R^{-1}\text{-cap}} e^{2\pi i \phi^\lambda(x; \omega_\theta)} T_\theta^\lambda f_\theta(x).$$

Fix  $\delta > 0$  to be sufficiently small for the purposes of the forthcoming argument. Each  $f_\theta$  is supported on an  $R^{-1}$ -ball and is therefore, of course, supported on an  $R^{-1+\delta}$ -ball. Since  $(R^{-1+\delta})^{-1} \leq \lambda^{1-\delta}$ , one may apply Lemma 5.6 to deduce that

$$T_\theta^\lambda f_\theta(x) = T_\theta^\lambda f_\theta * \eta_{R^{1-\delta}}(x) + \text{RapDec}(\lambda) \|f\|_{L^2(B^{n-1})}$$

for some choice of rapidly decreasing function  $\eta$ . As in the proof of Corollary 5.7, the function  $\eta$  may be chosen so that  $|\eta|$  admits a smooth, rapidly decreasing majorant  $\zeta: \mathbb{R}^n \rightarrow [0, \infty)$  which is locally constant at scale 1. In particular, it follows that

$$(11.9) \quad \zeta_{R^{1-\delta}}(x) \lesssim R^\delta \zeta_{R^{1-\delta}}(y) \quad \text{if } |x - y| \lesssim R.$$

Cover  $D_{\mathbf{R}}$  by finitely-overlapping  $R$ -balls and let  $B_R$  be some member of this cover. Combining the above observations, if  $\bar{x}$  denotes the centre of  $B_R$  and  $z \in B(0, R)$ , then

$$|T^\lambda f(\bar{x} + z)| \lesssim R^\delta \int_{\mathbb{R}^n} \left| \sum_{\theta: R^{-1}\text{-cap}} e^{2\pi i \tilde{\phi}^\lambda(z; \omega_\theta)} e^{2\pi i \phi^\lambda(\bar{x}; \omega_\theta)} T_\theta^\lambda f_\theta(y) \right| \zeta_{R^{1-\delta}}(\bar{x} - y) \, dy,$$

where  $\tilde{\phi}^\lambda(z; \omega_\theta) := \phi^\lambda(\bar{x} + z; \omega_\theta) - \phi^\lambda(\bar{x}; \omega_\theta)$ . Taking the  $L^p$ -norm in  $z$  it follows from Minkowski's inequality that  $\|T^\lambda f\|_{L^p(B_R)}$  is dominated by

$$R^\delta \int_{\mathbb{R}^n} \left\| \sum_{\theta: R^{-1}\text{-cap}} e^{2\pi i \tilde{\phi}^\lambda(\cdot; \omega_\theta)} e^{2\pi i \phi^\lambda(\bar{x}; \omega_\theta)} T_\theta^\lambda f_\theta(y) \right\|_{L^p(B(0, R))} \zeta_{R^{1-\delta}}(\bar{x} - y) \, dy.$$

By Lemma 11.7 the  $L^p$ -norm appearing in the above integrand is bounded by a constant multiple of

$$Q_p(R) R^{(n-1)/p'} \left( \sum_{\theta: R^{-1}\text{-cap}} |T_\theta^\lambda f_\theta(y)|^p \right)^{1/p}.$$

Applying Hölder's inequality and the locally-constant property (11.9), one deduces that

$$\|T^\lambda f\|_{L^p(B_R)} \lesssim Q_p(R) R^{(n-1)/p' + O(\delta)} \left( \int_{\mathbb{R}^n} \sum_{\theta: R^{-1}\text{-cap}} |T_\theta^\lambda f_\theta(\bar{x} + z - y)|^p \zeta_{R^{1-\delta}}(y) \, dy \right)^{1/p}$$

for all  $z \in B(0, R)$ . By raising both sides of this estimate to the  $p$ th power, averaging in  $z$  and summing over all balls  $B_R$  in the covering, it follows that  $\|T^\lambda f\|_{L^p(D_{\mathbf{R}})}$  is dominated by

$$Q_p(R)R^{(n-1)/p' - n/p + O(\delta)} \left( \int_{\mathbb{R}^n} \sum_{\theta: R^{-1}\text{-cap}} \|T^\lambda f_\theta\|_{L^p(D_{\mathbf{R}-y})}^p \zeta_{R^{1-\delta}}(y) dy \right)^{1/p}.$$

Observe that, by Hörmander's theorem (Lemma 5.4) and Hölder's inequality, one has

$$\|T^\lambda f_\theta\|_{L^2(D_{\mathbf{R}-y})} \lesssim R^{-(n-1)(1/2-1/p)+1/2} \|f_\theta\|_{L^p(B^{n-1})}.$$

On the other hand, the trivial estimate

$$\|T^\lambda f_\theta\|_{L^\infty(D_{\mathbf{R}-y})} \lesssim R^{-(n-1)/p'} \|f_\theta\|_{L^p(B^{n-1})}$$

holds, simply due to Hölder's inequality. Combining the above,

$$\|T^\lambda f_\theta\|_{L^p(D_{\mathbf{R}-y})} \lesssim R^{-(n-1)/p' + n/p} \|f_\theta\|_{L^p(B^{n-1})}.$$

The desired inequality is now immediate.  $\square$

## 12. AN $\varepsilon$ -REMOVAL LEMMA

The  $\lambda^\varepsilon$ -loss in the linear estimates of Theorems 1.1 and 1.2 can be removed away from the endpoint by an appeal to an  $\varepsilon$ -removal lemma of the type introduced in [23] (see also [8, 22]). The precise form of the required lemma does not appear in the literature, but it can be deduced by a minor modification of an argument from [23]. For completeness, the details are given presently.

Suppose  $T^\lambda$  is a Hörmander-type operator with associated phase function  $\phi^\lambda$ . Let  $\bar{p} \geq 2$  and suppose for all  $\varepsilon > 0$  the estimate

$$(12.1) \quad \|T^\lambda f\|_{L^p(B_R)} \lesssim_{\varepsilon, \phi, a} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

holds for all  $p \geq \bar{p}$ , all  $R$ -balls  $B_R$  for  $1 \leq R \leq \lambda$  and any choice of amplitude function. Under this hypothesis, one wishes to show that the global estimate

$$(12.2) \quad \|T^\lambda f\|_{L^p(\mathbb{R}^n)} \lesssim_{\phi, a} \|f\|_{L^p(B^{n-1})}$$

is valid for all  $p > \bar{p}$ .

**Definition 12.1** (Tao [23]). Let  $R \geq 1$ . A collection  $\{B(x_j, R)\}_{j=1}^N$  of  $R$ -balls in  $\mathbb{R}^n$  is *sparse* if the centres  $\{x_1, \dots, x_N\}$  are  $(RN)^{\bar{C}}$ -separated. Here  $\bar{C} \geq 1$  is a fixed constant, chosen large enough for the purposes of the proof.

Following [23], the first step towards establishing (12.2) is to reduce the problem to proving estimates for  $T^\lambda$  over sparse families of balls.

**Lemma 12.2.** *To prove (12.2) for all  $p > \bar{p}$  it suffices to show that for all  $\varepsilon > 0$  the estimate*

$$(12.3) \quad \|T^\lambda f\|_{L^{\bar{p}}(S)} \lesssim_{\varepsilon, \phi, a} R^\varepsilon \|f\|_{L^{\bar{p}}(B^{n-1})}$$

*holds whenever  $R \geq 1$  and  $S \subseteq \mathbb{R}^n$  is a union of  $R$ -balls belonging to a sparse collection, for any choice of amplitude function.*

The key step in the proof of Lemma 12.2 is the following covering lemma.

**Lemma 12.3** (Tao [22, 23]). *Suppose  $E \subseteq \mathbb{R}^n$  is the union of 1-cubes and  $N \geq 1$ . Define the radii  $R_j$  inductively by*

$$R_0 := 1, \quad R_j := R_{j-1}^{\bar{C}} |E|^{\bar{C}} \text{ for } 1 \leq j \leq N-1.$$

*Then for each  $0 \leq j \leq N-1$  there exists a family of sparse collections  $(\mathcal{B}_{j,\alpha})_{\alpha \in A_j}$  of balls of radius  $R_j$  such that the index sets  $A_k$  have cardinality  $O(|E|^{1/N})$  and*

$$E \subseteq \bigcup_{j=0}^{N-1} \bigcup_{\alpha \in A_j} S_{j,\alpha}$$

where  $S_{j,\alpha}$  is the union of all the balls belonging to the family  $\mathcal{B}_{j,\alpha}$ .

*Proof (of Lemma 12.2).* Let  $E \subseteq \mathbb{R}^n$  be a finite union of 1-cubes. For  $N \geq 1$ , the covering lemma together with the hypothesis (12.3) imply that

$$\|T^\lambda f\|_{L^{\bar{p}}(E)} \lesssim_{\varepsilon, \phi, a} N |E|^{1/N + \varepsilon \bar{C}^N} \|f\|_{L^{\bar{p}}(B^{n-1})}.$$

Choosing  $N \sim \log(1/\varepsilon)$ , it follows that

$$\|T^\lambda f\|_{L^{\bar{p}}(E)} \lesssim_{\varepsilon, \phi, a} |E|^{\bar{C}/\log(1/\varepsilon)} \|f\|_{L^{\bar{p}}(B^{n-1})}.$$

It will be convenient to work with the dual operator

$$T^*g(\omega) := \int_{\mathbb{R}^n} e^{-2\pi i \phi^\lambda(x; \omega)} a^\lambda(x; \omega) g(x) dx$$

so that the above estimate can be reformulated as

$$(12.4) \quad \|T^*g\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\varepsilon, \phi, a} |E|^{\bar{C}/\log(1/\varepsilon)} \|g\|_{L^{\bar{p}'}(E)}$$

for  $g$  supported on the set  $E$ .

Fix  $p > \bar{p}$  and  $\tau \in [-1/2, 1/2]^n$ . Suppose that  $g \in L^{p'}(\mathbb{R}^n)$  satisfies  $\|g\|_{L^{p'}(\mathbb{R}^n)} = 1$  and is constant on the mesh of 1-cubes centred on points of the lattice  $\tau + \mathbb{Z}^n$ . Form a level set decomposition of  $g$  by writing  $g = \sum_{k \in \mathbb{Z}} g_k$  where  $g_k := g \chi_{E_k}$  for

$$E_k := \{x \in \mathbb{R}^n : 2^{-k} \leq |g(x)| < 2^{-k+1}\}.$$

Chebyshev's inequality implies that  $|E_k| \leq 2^{kp'}$  for all  $k \in \mathbb{Z}$ . Furthermore, each set  $E_k$  is a union of 1-cubes and therefore if  $E_k \neq \emptyset$ , then  $|E_k| \geq 1$ . Combining these observations, one deduces that  $E_k = \emptyset$  for all  $k < 0$ . Since  $g_k$  is supported on  $E_k$ , one may apply (12.4) to conclude that

$$(12.5) \quad \|T^*g_k\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\varepsilon, \phi, a} |E_k|^{\bar{C}/\log(1/\varepsilon)} \|g_k\|_{L^{p'}(\mathbb{R}^n)}.$$

Using a simple base-times-height estimate, the right-hand side of (12.5) can be bounded by (a constant multiple of)

$$2^{-k} |E_k|^{\bar{C}/\log(1/\varepsilon) + 1/\bar{p}'} \lesssim 2^{-k(1 - \bar{C}p'/\log(1/\varepsilon) - p'/\bar{p}')}.$$

Since  $p' < \bar{p}'$ , by choosing  $\varepsilon$  sufficiently small one can ensure that the right-hand exponent is negative and therefore

$$\|T^*g\|_{L^{p'}(B^{n-1})} \lesssim \|T^*g\|_{L^{\bar{p}'}(B^{n-1})} \leq \sum_{k \geq 0} \|T^*g_k\|_{L^{\bar{p}'}(B^{n-1})} \lesssim_{\phi, a} 1 = \|g\|_{L^{p'}(\mathbb{R}^n)}.$$

This establishes the dual of the desired estimate (12.2) under the additional hypothesis that the function  $g$  is constant on 1-cubes.

It remains to remove the condition that  $g$  is constant on 1-cubes. The key observation is that this special case of (12.2) implies the discrete inequality

$$(12.6) \quad \left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma + \tau; \cdot)} a^\lambda(\sigma + \tau; \cdot) G(\sigma) \right\|_{L^{p'}(B^{n-1})} \lesssim_{\phi, a} \|G\|_{\ell^{p'}(\mathbb{Z}^n)}$$

for all  $G \in \ell^{p'}(\mathbb{Z}^n)$  and  $\tau \in [-1/2, 1/2]^n$ . Indeed, once (12.6) is established, taking  $g \in L^{p'}(\mathbb{R}^n)$  belonging to a suitable *a priori* class and applying Minkowski's inequality one deduces that

$$\|T^*g\|_{L^{p'}(B^{n-1})} \leq \int_{[-1/2, 1/2]^n} \left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma + \tau; \cdot)} a^\lambda(\sigma + \tau; \cdot) g(\sigma + \tau) \right\|_{L^{p'}(B^{n-1})} d\tau.$$

Combining this with (12.6) and Hölder's inequality yields (12.4).

Thus, the problem is now reduced to proving (12.6). Fix  $G \in \ell^{p'}(\mathbb{Z}^n)$  and define

$$\tilde{g}(x) := \sum_{\sigma \in \mathbb{Z}^n} G(\sigma) \chi(x - \sigma - \tau)$$

where  $\chi$  is the characteristic function of  $[-1/2, 1/2]^n$ . Since  $\tilde{g}$  is constant on 1-cubes, one is free to apply (12.4) to this function. In particular, let  $\tilde{T}^*$  be the dual of a Hörmander-type operator with phase  $\phi^\lambda$  and amplitude  $\tilde{a}^\lambda$  where

$$\tilde{a}(x; \omega) := \left( \prod_{j=1}^n \frac{\sin \pi(\partial_{x_j} \phi)(x; \omega)}{\pi(\partial_{x_j} \phi)(x; \omega)} \right)^{-1} a_0(x; \omega)$$

for  $a_0$  a smooth amplitude which is supported on  $X \times \Omega$  and satisfies  $a_0(x; \omega) = 1$  for  $(x; \omega) \in \text{supp } a$ . By the usual reductions (see §4) one may assume from the outset that  $|\partial_{x_j} \phi(x; \omega)| \leq 1/2$  for  $(x; \omega) \in X \times \Omega$  and  $1 \leq j \leq n$  and hence  $\tilde{a}$  is a well-defined, smooth function. Thus, the estimate

$$\|\tilde{T}^* \tilde{g}\|_{L^{p'}(B^{n-1})} \lesssim_{\phi, a} \|\tilde{g}\|_{L^{p'}(\mathbb{R}^n)}$$

holds, which can be rewritten as

$$(12.7) \quad \left\| \sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma + \tau; \cdot)} (A_\lambda)^\lambda(\sigma + \tau; \cdot) G(\sigma) \right\|_{L^{p'}(B^{n-1})} \lesssim_{\phi, a} \|G\|_{\ell^{p'}(\mathbb{Z}^n)}$$

where

$$(12.8) \quad A_\lambda(x; \omega) := \int_{[-1/2, 1/2]^n} e^{-2\pi i \lambda(\phi(x+y/\lambda; \omega) - \phi(x; \omega))} \tilde{a}(x + y/\lambda; \omega) dy.$$

Note that (12.7) is almost the desired expression (12.6) except for the disparity between the amplitude functions. To deal with this slight technicality, observe that, since

$$\lim_{\lambda \rightarrow \infty} A_\lambda(x; \omega) = a_0(x; \omega) \quad \text{uniformly,}$$

one may assume that  $\lambda$  is sufficiently large so that  $|A_\lambda(x; \omega)| \gtrsim 1$  for all  $(x; \omega) \in \text{supp } a$ . Thus, the expression appearing in the norm on the left-hand side of (12.6) is given by

$$\sum_{\sigma \in \mathbb{Z}^n} e^{-2\pi i \phi^\lambda(\sigma + \tau; \omega)} (A_\lambda)^\lambda(\sigma + \tau; \omega) G(\sigma) (\rho_\lambda)^\lambda(\sigma + \tau; \omega)$$

where, by taking a Fourier series expansion, the ratio  $\rho_\lambda$  satisfies

$$\rho_\lambda(x; \omega) := \frac{a(x; \omega)}{A_\lambda(x; \omega)} = \sum_{k \in \mathbb{Z}^n} (1 + |k|)^{-(n+1)} c_{\lambda, k}(\omega) e^{2\pi i \langle x, k \rangle}$$

where the  $c_{\lambda,k}$  are bounded functions, uniformly in  $\lambda$  and  $k$ . One may therefore bound the left-hand side of (12.6) by a  $(1+|k|)^{-(n+1)}$ -weighted sum of the left-hand side of (12.7) applied to modulated versions of  $G$ . Estimating each summand using (12.7) and summing in  $k$  concludes the proof.  $\square$

Given the above reduction, it remains to establish the estimates for  $T^\lambda$  over sparse collections of  $R$ -balls.

**Lemma 12.4.** *Under the hypothesis (12.1), if  $p \geq \bar{p}$ , then the estimate*

$$\|T^\lambda f\|_{L^p(S)} \lesssim_{\varepsilon,\phi,a} R^\varepsilon \|f\|_{L^p(B^{n-1})}$$

holds for all  $\varepsilon > 0$  whenever  $S \subseteq \mathbb{R}^n$  is a union of  $R$ -balls belonging to a sparse collection.

*Proof.* The proof uses a crude form of wave packet analysis and has much in common with the arguments described in §5. Let  $\{B(x_j, R)\}_{j=1}^N$  be the sparse collection of balls whose union is the set  $S$ . Clearly it suffices to assume that  $R \ll \lambda$  and that all the  $B(x_k, R)$  intersect the  $x$ -support of  $a^\lambda$ . Furthermore, letting  $c_{\text{diam}} > 0$  be a small constant chosen to satisfy the requirements of the forthcoming argument, by applying a partition of unity one may assume that  $\text{diam } X < c_{\text{diam}}$  and so

$$(12.9) \quad \frac{|x_{j_1} - x_{j_2}|}{\lambda} \lesssim c_{\text{diam}} \quad \text{for all } 1 \leq j_1, j_2 \leq N.$$

Fix  $\eta \in C^\infty(\mathbb{R}^{n-1})$  satisfying  $0 \leq \eta \leq 1$ ,  $\text{supp } \eta \in B^{n-1}$  and  $\eta(z) = 1$  for all  $z \in B(0, 1/2)$ . For  $R_1 := CNR$ , where  $C \geq 1$  is a large constant, define  $\eta_{R_1}(z) := \eta(z/R_1)$ . Further, let  $\psi \in C_c^\infty(\mathbb{R}^{n-1})$  satisfy  $0 \leq \psi \leq 1$ ,  $\text{supp } \psi \subset \Omega$  and  $\psi(\omega) = 1$  for  $\omega$  belonging to the  $\omega$ -support of  $a^\lambda$ . Fix  $1 \leq j \leq N$  and write

$$e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f = P_j f + (e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f - P_j f) =: P_j f + f_{j,\infty}$$

where  $P_j f := \hat{\eta}_{R_1} * [e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f]$ . If one defines

$$\text{Err}(x) := \int_{\mathbb{R}^{n-1}} e^{2\pi i(\phi^\lambda(x;\omega) - \phi^\lambda(x_j;\omega))} a^\lambda(x;\omega) f_{j,\infty}(\omega) d\omega,$$

then it follows that

$$T^\lambda f(x) = T^\lambda [e^{-2\pi i \phi^\lambda(x_j; \cdot)} P_j f](x) + \text{Err}(x).$$

For  $x \in B(x_j, R)$  the term  $\text{Err}(x)$  is negligible. Indeed, by Plancherel's theorem

$$\text{Err}(x) = \int_{\mathbb{R}^{n-1}} \overline{\check{G}_x(z)} \cdot (1 - \eta_{R_1}(z)) [e^{2\pi i \phi^\lambda(x_j; \cdot)} \psi f]^\vee(z) dz$$

where

$$\check{G}_x(z) = \int_{\mathbb{R}^{n-1}} e^{2\pi i(\langle z, \omega \rangle - \phi^\lambda(x;\omega) + \phi^\lambda(x_j;\omega))} a^\lambda(x;\omega) d\omega.$$

Taking the  $\omega$ -gradient of the phase of  $\check{G}_x(z)$ , one obtains

$$z - \lambda(\partial_\omega \phi(x/\lambda; \omega) - \partial_\omega \phi(x_j/\lambda; \omega)) = z + O(R).$$

Thus, if  $z$  belongs to the support of  $1 - \eta_{R_1}$ , then  $G_x(z)$  is rapidly decaying in  $R_1$  and therefore

$$|\text{Err}(x)| \leq \text{RapDec}(R_1) \|f\|_{L^p(B^{n-1})}.$$

It remains to bound the contributions arising from the frequency localised pieces. By applying the estimate for  $T^\lambda$  with  $R^\varepsilon$ -loss over each ball  $B(x_j, R)$  one obtains

$$\begin{aligned} \|T^\lambda f\|_{L^p(S)} &\leq \left( \sum_{j=1}^N \|T^\lambda [e^{-2\pi i \phi^\lambda(x_j; \cdot)} P_j f]\|_{L^p(B(x_j, R))}^p \right)^{1/p} + \text{RapDec}(R_1) \|f\|_{L^p(B^{n-1})} \\ &\lesssim_{\varepsilon, \phi, a} R^\varepsilon \left( \sum_{j=1}^N \|P_j f\|_{L^p(B^{n-1})}^p \right)^{1/p} + \|f\|_{L^p(B^{n-1})}. \end{aligned}$$

Thus, it now suffices to show that

$$\left( \sum_{j=1}^N \|P_j f\|_{L^p(\mathbb{R}^{n-1})}^p \right)^{1/p} \lesssim \|f\|_{L^p(B^{n-1})}.$$

This estimate follows via interpolation between the endpoint cases  $p = 2$  and  $p = \infty$ , which are established presently. The  $p = \infty$  case is a trivial consequence of Young's inequality and so it suffices to consider  $p = 2$ . By duality, the desired inequality is equivalent to

$$(12.10) \quad \left\| \sum_{j=1}^N e^{-2\pi i \phi^\lambda(x_j; \cdot)} \psi \cdot [\hat{\eta}_{R_1} * g_j] \right\|_{L^2(\mathbb{R}^{n-1})} \lesssim \left( \sum_{j=1}^N \|g_j\|_{L^2(B^{n-1})}^2 \right)^{1/2}.$$

By squaring the left-hand side of (12.10) one obtains

$$\sum_{j_1, j_2=1}^N \int_{\mathbb{R}^{n-1}} \overline{G_{j_1, j_2}(\omega)} \hat{\eta}_{R_1} * g_{j_1}(\omega) \overline{\hat{\eta}_{R_1} * g_{j_2}(\omega)} d\omega$$

where

$$G_{j_1, j_2}(\omega) := e^{2\pi i(\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega))} \psi(\omega)^2.$$

By Plancherel's theorem, each summand of the above expression can be written as

$$(12.11) \quad \int_{\mathbb{R}^{n-1}} \overline{\check{G}_{j_1, j_2}(z)} (\eta_{R_1} \check{g}_{j_1}) * (\eta_{R_1} \check{g}_{j_2})^\sim(z) dz;$$

here  $(\eta_{R_1} \check{g}_{j_2})^\sim(z) := \overline{(\eta_{R_1} \check{g}_{j_2})(-z)}$ . Note that the integrand in (12.11) is supported on a ball of radius  $O(R_1)$  about the origin.

Fix  $1 \leq j_1, j_2 \leq N$  with  $j_1 \neq j_2$ , let  $z \in \mathbb{R}^{n-1}$  with  $|z| \lesssim R_1 < |x_{j_2} - x_{j_1}|$  and consider

$$\check{G}_{j_1, j_2}(z) = \int_{\mathbb{R}^{n-1}} e^{2\pi i(\langle z, \omega \rangle + \phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega))} \psi(\omega)^2 d\omega.$$

This oscillatory integral can be bounded by a simple stationary phase analysis. For  $\alpha \in \mathbb{N}^{n-1}$  with  $|\alpha| \leq 2$  consider the function

$$\partial_\omega^\alpha [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)] = \partial_\omega^\alpha \langle \partial_x^\lambda \phi(x_{j_1}; \omega), x_{j_2} - x_{j_1} \rangle + O(c_{\text{diam}} |x_{j_2} - x_{j_1}|),$$

where the remainder term has been estimated using (12.9).

Let  $c_{\text{crit}} > 0$  be another small constant, chosen to satisfy the requirements of the forthcoming argument, and  $\omega_0 \in \Omega$ . Suppose that

$$(12.12) \quad \left| \pm \frac{x_{j_2} - x_{j_1}}{|x_{j_2} - x_{j_1}|} - G^\lambda(x_{j_1}; \omega_0) \right| \geq c_{\text{crit}},$$

where the estimate is interpreted as holding for both choices of sign. Condition H1) on the phase implies that for each  $\omega_0 \in \Omega$  the vector  $G^\lambda(x; \omega_0)$  spans the kernel of  $\partial_{\omega x}^2 \phi^\lambda(x; \omega_0)$ . Consequently, in view of (12.12) one has

$$|\partial_\omega [\langle \partial_x \phi^\lambda(x_{j_1}; \omega), x_{j_2} - x_{j_1} \rangle] |_{\omega=\omega_0}| \gtrsim |x_{j_2} - x_{j_1}|$$

and therefore

$$|\partial_\omega [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)] |_{\omega=\omega_0}| \gtrsim |x_{j_2} - x_{j_1}|,$$

provided  $c_{\text{diam}}$  is sufficiently small. On the other hand, if (12.12) fails, then

$$\partial_\omega^\alpha \left\langle \partial_x \phi^\lambda(x_{j_1}; \omega), \frac{x_{j_2} - x_{j_1}}{|x_{j_2} - x_{j_1}|} \right\rangle |_{\omega=\omega_0} = \partial_\omega^\alpha \langle \partial_x \phi^\lambda(x_{j_1}; \omega), G^\lambda(x; \omega_0) \rangle |_{\omega=\omega_0} + O(c_{\text{crit}}).$$

If  $c_{\text{crit}}$  and  $c_{\text{diam}}$  are both chosen to be sufficiently small, then condition H2) implies that

$$|\det \partial_{\omega\omega}^2 [\phi^\lambda(x_{j_1}; \omega) - \phi^\lambda(x_{j_2}; \omega)] |_{\omega=\omega_0}| \gtrsim |x_{j_2} - x_{j_1}|^{n-1}.$$

Thus, any critical point of the phase must be (quantitatively) non-degenerate and one may apply higher dimensional versions of van der Corput's lemma (see, for instance, Chapter VIII, Proposition 6 of [20]) to estimate the oscillatory integral. In particular,

$$|\check{G}_{j_1, j_2}(z)| \lesssim |x_{j_2} - x_{j_1}|^{-(n-1)/2} \lesssim R_1^{-\bar{C}}$$

so that the absolute value of (12.11) is bounded by

$$\begin{aligned} R_1^{-\bar{C}} \|(\eta_{R_1} \check{g}_{j_1}) * (\eta_{R_1} \check{g}_{j_2})^\sim\|_{L^1(\mathbb{R}^{n-1})} &\lesssim R_1^{-\bar{C}} \prod_{i=1}^2 \|\eta_{R_1} \check{g}_{j_i}\|_{L^1(\mathbb{R}^{n-1})} \\ &\lesssim R_1^{-\bar{C}+n-1} \prod_{i=1}^2 \|g_{j_i}\|_{L^2(\mathbb{R}^{n-1})}. \end{aligned}$$

Since there are only  $O(N^2)$  choice of indices  $j_1, j_2$ , one may invoke the trivial estimate

$$\prod_{i=1}^2 \|g_{j_i}\|_{L^2(\mathbb{R}^{n-1})} \lesssim \sum_{j=1}^N \|g_j\|_{L^2(B^{n-1})}^2$$

and then sum all the contributions from all pairs  $j_1, j_2$  to bound the off-diagonal terms arising from the left-hand side of (12.10). On the other hand, the diagonal terms provide a favourable contribution of

$$\left( \sum_{j=1}^N \|\hat{\eta}_{R_1} * g_j\|_{L^2(B^{n-1})}^2 \right)^{1/2} \lesssim \left( \sum_{j=1}^N \|g_j\|_{L^2(B^{n-1})}^2 \right)^{1/2}.$$

Combining these observations concludes the proof of (12.10) and thereby establishes the lemma.  $\square$

## REFERENCES

- [1] T. Bagby, L. Bos, and N. Levenberg, *Multivariate simultaneous approximation*, Constr. Approx. **18** (2002), no. 4, 569–577. MR 1920286
- [2] David Beltran and Jonathan Hickman, *Variable coefficient Wolff-type inequalities and sharp local smoothing estimates for wave equations on manifolds*, In preparation.
- [3] Jonathan Bennett, *Aspects of multilinear harmonic analysis related to transversality*, Harmonic analysis and partial differential equations, Contemp. Math., vol. 612, Amer. Math. Soc., Providence, RI, 2014, pp. 1–28. MR 3204854

- [4] Jonathan Bennett, Anthony Carbery, and Terence Tao, *On the multilinear restriction and Keakeya conjectures*, Acta Math. **196** (2006), no. 2, 261–302. MR 2275834
- [5] J. Bourgain,  *$L^p$ -estimates for oscillatory integrals in several variables*, Geom. Funct. Anal. **1** (1991), no. 4, 321–374. MR 1132294
- [6] ———, *Moment inequalities for trigonometric polynomials with spectrum in curved hypersurfaces*, Israel J. Math. **193** (2013), no. 1, 441–458. MR 3038558
- [7] Jean Bourgain, *Some new estimates on oscillatory integrals*, Essays on Fourier analysis in honor of Elias M. Stein (Princeton, NJ, 1991), Princeton Math. Ser., vol. 42, Princeton Univ. Press, Princeton, NJ, 1995, pp. 83–112. MR 1315543
- [8] Jean Bourgain and Larry Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295. MR 2860188
- [9] Lennart Carleson and Per Sjölin, *Oscillatory integrals and a multiplier problem for the disc*, Studia Math. **44** (1972), 287–299. (errata insert), Collection of articles honoring the completion by Antoni Zygmund of 50 years of scientific activity, III. MR 0361607
- [10] Xi Chen, Neeraj Kayal, and Avi Wigderson, *Partial derivatives in arithmetic complexity and beyond*, Found. Trends Theor. Comput. Sci. **6** (2010), no. 1–2, front matter, 1–138 (2011). MR 2901512
- [11] Xiumin Du, Larry Guth, and Xiaochun Li, *A sharp schrödinger maximal estimate in  $\mathbb{R}^2$* , Ann. of Math. (2) **186** (2017), 607–640.
- [12] Charles Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36. MR 0257819
- [13] Larry Guth, *Restriction estimates using polynomial partitioning II*, Preprint: arXiv:1603.04250.
- [14] ———, *A restriction estimate using polynomial partitioning*, J. Amer. Math. Soc. **29** (2016), no. 2, 371–413. MR 3454378
- [15] Larry Guth and Nets Hawk Katz, *On the Erdős distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), no. 1, 155–190. MR 3272924
- [16] Lars Hörmander, *Oscillatory integrals and multipliers on  $FL^p$* , Ark. Mat. **11** (1973), 1–11. MR 0340924 (49 #5674)
- [17] Sanghyuk Lee, *Linear and bilinear estimates for oscillatory integral operators related to restriction to hypersurfaces*, J. Funct. Anal. **241** (2006), no. 1, 56–98. MR 2264247
- [18] E. M. Stein, *Oscillatory integrals in Fourier analysis*, Beijing lectures in harmonic analysis (Beijing, 1984), Ann. of Math. Stud., vol. 112, Princeton Univ. Press, Princeton, NJ, 1986, pp. 307–355. MR 864375
- [19] Elias M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, No. 30, Princeton University Press, Princeton, N.J., 1970. MR 0290095
- [20] ———, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III. MR 1232192 (95c:42002)
- [21] A. H. Stone and J. W. Tukey, *Generalized “sandwich” theorems*, Duke Math. J. **9** (1942), 356–359. MR 0007036
- [22] Terence Tao, *The weak-type endpoint Bochner–Riesz conjecture and related topics*, Indiana Univ. Math. J. **47** (1998), no. 3, 1097–1124. MR 1665753
- [23] ———, *The Bochner–Riesz conjecture implies the restriction conjecture*, Duke Math. J. **96** (1999), no. 2, 363–375. MR 1666558
- [24] Terence Tao, Ana Vargas, and Luis Vega, *A bilinear approach to the restriction and Keakeya conjectures*, J. Amer. Math. Soc. **11** (1998), no. 4, 967–1000. MR 1625056
- [25] L. Wisewell, *Keakeya sets of curves*, Geom. Funct. Anal. **15** (2005), no. 6, 1319–1362. MR 2221250
- [26] Richard Wongkew, *Volumes of tubular neighbourhoods of real algebraic varieties*, Pacific J. Math. **159** (1993), no. 1, 177–184. MR 1211391
- [27] Joshua Zahl, *On the Wolff circular maximal function*, Illinois J. Math. **56** (2012), no. 4, 1281–1295. MR 3231483
- [28] Ruixiang Zhang, *Polynomials with dense zero sets and discrete models of the Keakeya conjecture and the Furstenberg set problem*, Selecta Math. (N.S.) **23** (2017), no. 1, 275–292. MR 3595894

LARRY GUTH: DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY,  
182 MEMORIAL DRIVE, CAMBRIDGE, MASSACHUSETTS, 02139, USA.

*E-mail address:* [lguth@math.mit.edu](mailto:lguth@math.mit.edu)

JONATHAN HICKMAN: ECKHART HALL ROOM 414, DEPARTMENT OF MATHEMATICS, UNIVERSITY  
OF CHICAGO, 5734 S. UNIVERSITY AVENUE, CHICAGO, ILLINOIS, 60637, USA.

*E-mail address:* [jehickman@uchicago.edu](mailto:jehickman@uchicago.edu)

MARINA ILIOPOULOU: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKE-  
LEY, CA, 94720-3840, USA.

*E-mail address:* [m.iliopoulou@berkeley.edu](mailto:m.iliopoulou@berkeley.edu)