A note on Fourier restriction and nested Polynomial Wolff axioms

A NOTE ON FOURIER RESTRICTION AND NESTED POLYNOMIAL WOLFF AXIOMS

JONATHAN HICKMAN AND JOSHUA ZAHL

Abstract. This note records an asymptotic improvement on the known L^p range for the Fourier restriction conjecture in high dimensions. This is obtained by combining Guth's polynomial partitioning method with recent geometric results regarding intersections of tubes with nested families of varieties.

1. Introduction

1.1. **Main result.** For $n \geq 2$ let B^{n-1} denote the unit ball in $\mathbb{R}^{n-1}$ and consider the *Fourier extension* operator E defined by

$$Ef(x) := \int_{B^{n-1}} f(\omega)\, e^{i(x_1\omega_1+\cdots+x_{n-1}\omega_{n-1}+x_n|\omega|^2)}\, \mathrm{d}\omega$$

for $f \in L^1(B^{n-1})$. The adjoint form of Stein's restriction conjecture [16] asserts that the estimate

$$\|Ef\|_{L^p(\mathbb{R}^n)} \leq C_{n,p}\|f\|_{L^p(\mathbb{R}^n)} \qquad (\mathrm{R}^*_p)$$

holds for all $p > \frac{2n}{n-1}$. This conjecture is known to hold for $n = 2$ due to Fefferman–Stein [6] but remains open in all other dimensions. The problem has a rich history and is related to a wide range of important questions in harmonic analysis, geometric measure theory, PDE and number theory: see, for instance, the surveys [22, 19, 1] for further details.

The current best partial results on the restriction conjecture are based on the polynomial method [21, 10, 12], which was introduced in this context in a seminal work of Guth [10]. The purpose of this note is to combine Guth's polynomial method approach with a recent Kakeya-type geometric result from [13, 25] in order improve the known bounds on the problem in the high-dimensional regime.

The attendant numerology is somewhat complicated and consequently it is convenient to express the results in an asymptotic form. In particular, if the restriction conjecture were true, then (R^*_p) would hold for

$$p > 2 + 2n^{-1} + O(n^{-2}).$$

Thus, one may consider the $\lambda \geq 2$ for which (R^*_p) can be confirmed in the range

$$p > 2 + \lambda n^{-1} + O(n^{-2}). \qquad (1.1)$$

Theorem 1.1. (R^*_p) *holds in the range* (1.1) *with* $\lambda = 2.596\ldots$.

The precise form of λ, involving the irrational real root of a cubic equation, is described in the appendix.

A comparison of the numerology of Theorem 1.1 with previous partial results on the restriction conjecture is presented in Figure 1. It is remarked that in certain low dimensions the methods of this paper do not improve upon known results and, in particular, stronger estimates are known in the $n = 3$ due to Wang [21]. For the current best bounds in low dimensions see Figure 2.

$\lambda =$	
4	Tomas [20]
3	Bourgain–Guth [3]
8/3	Guth [10]
2.604...	Hickman–Rogers [12]
2.596...	Theorem 1.1

FIGURE 1. Comparison with previous asymptotic results on the restriction conjecture.

1.2. **Polynomial Wolff axioms.** Underlying the restriction conjecture are deep geometric problems concerning the continuum incidence theory of long, thin tubes in $\mathbb{R}^n$. To be more precise, given a large parameter $R \geq 1$ define an $R \times R^{1/2}$*-tube* to be a cylinder $T \subset \mathbb{R}^n$ of height R and radius $R^{1/2}$ with arbitrary position and arbitrary orientation. The R versus $R^{1/2}$ scaling arises naturally in the analysis of Ef owing to the quadratic nature of the phase function. The *direction* of an $R \times R^{1/2}$-tube T is defined to be the direction of its coaxial line, which is denoted here by $\mathrm{dir}(T) \in S^{n-1}$. A family $\mathbf{T}$ of $R \times R^{1/2}$-tubes is *direction-separated* if $\{\mathrm{dir}(T) : T \in \mathbf{T}\}$ forms an $R^{-1/2}$-separated subset of the unit sphere.

Effective analysis of Ef relies on understanding of the incidence geometry of direction-separated tube families $\mathbf{T}$. This is formalised by the well-known and celebrated fact that the restriction conjecture implies the *Kakeya conjecture*, where the latter may be loosely interpreted as a bound on the number of possible incidences between tubes in $\mathbf{T}$.

In relation to the incidence theory, a critical case occurs when the tubes from $\mathbf{T}$ tend to align around neighbourhoods of low dimensional, low degree algebraic varieties. The significance of this situation was first highlighted in the context of the restriction conjecture by Guth [7], and in the context of the Kakeya conjecture by Guth and the second author [8]. Thus, it is important to understand possible interactions between tubes and varieties, which are typically constrained by the dimension and degree of the variety. A fundamental tool in this direction is the following theorem which, in the language of [8], states that direction-separated families of tubes satisfy the *polynomial Wolff axioms* (see [8] for a discussion of this terminology).

Theorem 1.2 (Polynomial Wolff axioms [14]). *For all $n > j \geq 1$, $d \geq 1$ and $\varepsilon > 0$, there is a constant $C_{n,d,\varepsilon} > 0$ such that*

$$\#\big\{T \in \mathbf{T} : |T \cap B_r \cap N_{R^{1/2}}\mathbf{Z}| \geq r|T|\big\} \leq C_{n,d,\varepsilon} r^{-j} R^{(n+j-1)/2+\varepsilon}$$

whenever $1 \leq R^{1/2} \leq r \leq R$, $\mathbf{T}$ is a direction-separated family of $R \times R^{1/2}$-tubes and $\mathbf{Z} \subset \mathbb{R}^n$ is an algebraic variety of codimension j and degree at most d.

See also [7, 24] for earlier partial results. Here $N_r E$ denotes the r-neighbourhood of E for any $r > 0$ and $E \subseteq \mathbb{R}^n$ and B_r is a choice of ball in $\mathbb{R}^n$ of radius r. The relevant algebraic definitions are recalled in §3.2 below.

Using the $n = 3$ case of Theorem 1.2, Guth [7] was able to improve the then best bound on the restriction conjecture in $\mathbb{R}^3$. This argument was extended to higher dimensions by Rogers and the first author [12], combining the ideas from [7] with those of Guth's study of the higher dimensional problem in [10]. This led to the previous best known asymptotic for the restriction conjecture (see Figure 1).

$n =$	$p >$		$n =$	$p >$	
2	4	Fefferman–Stein [6]	11	$2 + \frac{12597}{49670}$	Theorem 1.4
3	$3 + \frac{3}{13}$	Wang [21]	12	$2 + \frac{4}{17}$	Guth [10]
4	$2 + \frac{1407}{1759}$	Hickman–Rogers [12][1]	13	$2 + \frac{185725}{878068}$	Theorem 1.4
5	$2 + \frac{63}{100}$	Theorem 1.4	14	$2 + \frac{1671525}{8414731}$	Theorem 1.4
6	$2 + \frac{1}{2}$	Guth [10]	15	$2 + \frac{2}{11}$	Theorem 1.4
7	$2 + \frac{429}{1018}$	Theorem 1.4	16	$2 + \frac{20036013}{116580449}$	Theorem 1.4
8	$2 + \frac{4}{11}$	Guth [10]	17	$2 + \frac{4}{25}$	Theorem 1.4
9	$2 + \frac{7293}{23032}$	Theorem 1.4	18	$2 + \frac{123751845}{817128103}$	Theorem 1.4
10	$2 + \frac{2}{7}$	Guth [10]	19	$2 + \frac{1}{7}$	Theorem 1.4

FIGURE 2. The current state-of-the-art for the restriction problem in low dimensions. New results are highlighted and are deduced by combining Theorem 1.4 with the linear to k-broad reduction from [3, 10].[2]

In [13] and [25] a non-trivial extension of Theorem 1.2 was concurrently and independently established which, rather than controlling interactions between direction-separated tubes lying close to a single variety, controls interactions between direction-separated tubes and *nested families of varieties.*

Theorem 1.3 (Nested polynomial Wolff axioms [13, 25]). *For all $n > m \geq 1$, $d \geq 1$ and $\varepsilon > 0$, there is a constant $C_{n,d,\varepsilon} > 0$ such that*

$$\# \bigcap_{j=1}^{m} \Big\{ T \in \mathbf{T} : |T \cap B_{r_j} \cap N_{R^{1/2}}\mathbf{Z}_j| \geq r_j |T| \Big\} \leq C_{n,d,\varepsilon} \Big(\prod_{j=1}^{m} r_j^{-1} \Big) R^{(n+m-1)/2+\varepsilon}$$

holds whenever:

- $1 \leq R^{1/2} \leq r_j \leq R$ *for* $1 \leq j \leq m$ *and* $B_{r_m} \subseteq \dots \subseteq B_{r_1} \subset \mathbb{R}^n$*;*
- $\mathbf{T}$ *is a direction-separated family of* $R \times R^{1/2}$*-tubes;*
- *Each* $\mathbf{Z}_j \subset \mathbb{R}^n$ *is an algebraic variety of codimension* j *and degree at most* d.

In [13, 25] Theorem 1.3 was applied to give new bounds on the Kakeya conjecture in the high dimensional regime. In this note the same ideas are transferred into the context of the restriction problem. In particular, by adapting the arguments used to study the restriction problem from [7, 10, 12], one is led to consider tube interactions with nested families of varieties. There are some differences between the geometric setup in the Kakeya and restriction problems, however, and consequently Theorem 1.3 is not used in the forthcoming analysis *per se*, but rather a variant which is better adapted to the restriction problem. This variant is stated in terms of lines rather than tubes and follows directly from [25, Lemma 2.11]: see Theorem 3.8 below.

[1]See also [5].

[2]These computations were carried out using the following Maple [15] code:

```
n := [insert dimension];
p_broad := 2+6/(2*(n-1)+(k-1)*4^(n-k)*(factorial(n-1)/factorial(k-1))^2
*factorial(2*k-1)/factorial(2*n-1)):   p_limit :=2+ 4/(2*n-k):
p_seq := [seq(max(eval(p_broad, k = i), eval(p_limit, k = i)), i = 2 .. n)]:
new_exponent := min(p_seq);
```

1.3. ***k*-broad estimates.** Rather than attempt to prove (R^*_p) directly, a number of standard reductions are applied to reduce matters to a simpler class of estimates.

By a now standard *ε-removal argument* (see [17]) and factorisation theory (see [2] or [4, Lemma 1]), the inequality (R^*_p) holds for all p in an open range if and only if for all $\varepsilon > 0$ and all $R \geq 1$ the local estimates

$$\|Eg\|_{L^p(B_R)} \leq C_{n,p,\varepsilon} R^{\varepsilon} \|g\|_{L^\infty(\mathbb{R}^{n-1})} \qquad (\mathrm{R}^*_{p,\mathrm{loc}})$$

hold in the same range. Here B_R denotes an arbitrary ball of radius R in $\mathbb{R}^n$.

Using the Bourgain–Guth method [3, 10], one may further reduce the problem to working with weaker *k-broad estimates*, which take the form

$$\|Eg\|_{\mathrm{BL}^p_k(B_R)} \leq C_{n,p,\varepsilon} R^{\varepsilon} \|g\|_{L^\infty(\mathbb{R}^{n-1})}. \qquad (\mathrm{BL}^p_k)$$

The reader is referred to [10] for the definition and basic properties of the *k-broad norm* appearing on the left-hand side of this inequality.

The main result of this article is the following theorem.

Theorem 1.4. *Let* $2 \leq k \leq n-1$ *and*

$$p \geq p_n(k) := 2 + \frac{6}{2(n-1) + (k-1)\prod_{i=k}^{n-1} \frac{2i}{2i+1}}. \qquad (1.2)$$

Then (BL^p_k) *holds for all* $\varepsilon > 0$ *and* $R \gg 1$.

Theorem 1.1 follows from Theorem 1.4 in view of the aforementioned Bourgain–Guth method [3, 10]. In particular, it follows from [3, 10] that (BL^p_k) implies ($\mathrm{R}^*_{p,\mathrm{loc}}$) whenever $n \geq 3$ and

$$2 + \frac{4}{2n-k} \leq p \leq 2 + \frac{2}{k-2}. \qquad (1.3)$$

Thus, to obtain the best possible estimate (R^*_p) from Theorem 1.4, one wishes to choose a value of k which optimises the range of p in (1.2) subject to the constraints in (1.3). For a given dimension it is a straightforward exercise to compute the optimal choice of k and the resulting range of linear restriction estimates in low dimensions are tabulated in Figure 2. However, in general the optimisation procedure does not produce a compact formula for the explicit p range, hence the convenience of the asymptotic formulation in Theorem 1.1. The derivation of the value $\lambda = 2.596...$ for the asymptotic is described in the appendix.

1.4. **Structure of the article.** This article is **not** self-contained and refers heavily back to the work of Guth [10], and the reformulation of Guth's induction-on-scale argument as a recursive algorithm in [12]. In §2 certain notational conventions are set up; §3 describes various preliminaries including the wave packet decomposition and basic algebraic definitions; §4 provides an overview of a modified form of the polynomial structural decomposition described in [10, 12] and explains how this modification can be combined with Theorem 1.3 to give Theorem 1.4; §5 deals with basic orthogonality results used to establish the decomposition in §4 whilst the decomposition itself is proven in §6, using arguments from [10, 12]; appended is a discussion of the numerology of Theorem 1.1.

Acknowledgement. The authors would like to thank Keith M. Rogers for numerous helpful discussions.

2. Notational conventions

In the arguments that follow, the parameters n, k, ε are fixed and satisfy the hypotheses of Theorem 1.4. In particular, all implicit constants are allowed to depend on n, k, and ε.

Our arguments will involve a number of additional admissible parameters

$$\varepsilon^C \leq \delta \ll_\varepsilon \delta_0 \ll_\varepsilon \delta_1 \ll_\varepsilon \cdots \ll_\varepsilon \delta_{n-k} \ll_\varepsilon \varepsilon_\circ \ll_\varepsilon \varepsilon. \tag{2.1}$$

Here C is some dimensional constant and the notation $A \ll_\varepsilon B$ for $A, B \geq 0$ indicates that $A \leq \mathbf{C}_{n,\varepsilon}^{-1} B$ for some fixed large admissible constant $\mathbf{C}_{n,\varepsilon} \geq 1$ chosen to satisfy the requirements of the following arguments.

Given $A, B \geq 0$ and a (possibly empty) list of objects L, the notation $A \lesssim_L B$, $A = O_L(B)$ or $B \gtrsim_L A$ indicates that $A \leq C_{n,L} B$ for some constant $C_{n,L} > 0$ depending only on n and the objects in L, whilst $A \sim_L B$ denotes that $A \lesssim_L B$ and $B \lesssim_L A$. Given a large parameter $r \geq 1$, the notation $\mathrm{RapDec}(r)$ is used to denote a non-negative term which is rapidly decreasing in r in the sense that

$$\mathrm{RapDec}(r) \lesssim_{\varepsilon,N} r^{-N} \qquad \text{for all } N \in \mathbb{N}.$$

Such terms frequently appear as 'errors' in the arguments.

3. Preliminaries

3.1. **Wave packet decomposition.** For $r \geq 1$ let $\Theta[r]$ denote the set of all balls θ of radius $r^{-1/2}$ in $\mathbb{R}^{n-1}$ with centres ω_θ lying in $c_n r^{-1/2}\mathbb{Z}^{n-1} \cap B^{n-1}$ for $c_n := 2^{-1}(n-1)^{-1/2}$. Fix $\psi \in C_c^\infty(\mathbb{R}^{n-1})$ with $\operatorname{supp}\psi \subseteq [-c_n, c_n]^{n-1}$ satisfying

$$\sum_{k \in \mathbb{Z}^{n-1}} \psi(\,\cdot\, - c_n k) \equiv 1;$$

such a function may be constructed using the Poisson summation formula. In addition, let $\tilde{\psi} \in C_c^\infty(\mathbb{R}^{n-1})$ satisfy $\operatorname{supp}\tilde{\psi} \subseteq B^{n-1}$ and $\tilde{\psi}(\omega) = 1$ whenever $\omega \in \operatorname{supp}\psi$. For $\theta \in \Theta[r]$ with $\omega_\theta = c_n r^{-1/2} k_\theta$ define $\psi_\theta(\omega) := \psi(r^{1/2}\omega - c_n k_\theta)$ and $\tilde{\psi}_\theta(\omega) := \tilde{\psi}(r^{1/2}\omega - c_n k_\theta)$ so that both ψ_θ and $\tilde{\psi}_\theta$ are supported in θ and $\tilde{\psi}_\theta(\omega) = 1$ whenever $\omega \in \operatorname{supp}\psi_\theta$.

Writing $\mathbb{T}[r] := \Theta[r] \times r^{1/2}\mathbb{Z}^{n-1}$, the inversion formula for Fourier series allows one to decompose

$$f = \sum_{(\theta,v) \in \mathbb{T}[r]} f_{\theta,v} \tag{3.1}$$

where[3]

$$f_{\theta,v}(\omega) := \Big(\frac{r^{1/2}}{2\pi}\Big)^{n-1} e^{i\langle v,\omega\rangle} (f \cdot \psi_\theta)\hat{\ }(v) \tilde{\psi}_\theta(\omega).$$

The sum (3.1) is referred to as the *wave packet decomposition of f at scale r*. The pairs $(\theta, v) \in \mathbb{T}[r]$ and functions $f_{\theta,v}$ will both be referred to as *(scale r) wave packets*.

The key properties of this decomposition are as follows:

- **Orthogonality between the wave packets.** Combining spatial orthogonality with the Plancherel identity for Fourier series,

$$\max_{\theta_* \in \Theta[\rho]} \Big\| \sum_{(\theta,v) \in \mathbb{W}} f_{\theta,v} \Big\|_{L^2(\theta_*)}^2 \sim \max_{\theta_* \in \Theta[\rho]} \sum_{(\theta,v) \in \mathbb{W}} \|f_{\theta,v}\|_{L^2(\theta_*)}^2$$

  for any collection of wave packets $\mathbb{W} \subseteq \mathbb{T}[r]$ and $1 \leq \rho \leq r$.
- **Spatial concentration.** Given $0 < \delta \ll 1$ as in (2.1) and any wave packet $(\theta, v) \in \mathbb{T}[r]$, define the tube

$$T_{\theta,v} := \big\{ x \in B(0,r) : |x' + 2x_n\omega_\theta + v| \leq r^{1/2+\delta} \big\}.$$

  By a simple stationary phase argument (see, for instance, [18]) the function $Ef_{\theta,v}$ is concentrated on $T_{\theta,v}$ in the sense that

$$|Ef_{\theta,v}(x)\chi_{B(0,r)\setminus T_{\theta,v}}(x)| = \mathrm{RapDec}(r)\|f\|_2 \qquad \text{for all } x \in \mathbb{R}^n.$$

[3]Here $\hat{g}$ denotes the Fourier transform of $g \in L^1(\mathbb{R}^d)$; that is: $\hat{g}(\xi) := \int_{\mathbb{R}^d} e^{-i\langle x,\xi\rangle} g(x)\,\mathrm{d}x$.

Each $T_{\theta,v}$ is an $R\times R^{1/2}$-tube with coaxial line passing through the point $(-v,0)^\top \in \mathbb{R}^n$ in the direction $G(\omega_\theta) := (-2\omega_\theta, 1)^\top$.

Definition 3.1. Given $\mathbb{W} \subseteq \mathbb{T}[r]$, a function $f \in L^1(B^{n-1})$ is said to be concentrated on wave packets from $\mathbb{W}$ if

$$\Big\| \sum_{(\theta,v)\notin\mathbb{W}} f_{\theta,v}\Big\|_\infty = \mathrm{RapDec}(r)\|f\|_2.$$

Here the $\mathrm{RapDec}(r)$ notation is as defined in §2.

3.2. **Algebraic/geometric definitions.** Given real polynomials $P_1, \dots, P_m \in \mathbb{R}[X_1, \dots, X_n]$ let

$$Z(P_1, \dots, P_m) := \big\{z \in \mathbb{R}^n : P_j(z) = 0 \text{ for } 1 \leq j \leq m\big\}$$

denote their common zero set. For the purposes of this article, such a set $\mathbf{Z}$ is referred to as a *variety*.

Suppose $\mathbf{Z} := Z(P_1, \dots, P_m)$ is a variety satisfying the additional condition

$$\bigwedge_{j=1}^m \nabla P_j(z) \neq 0 \qquad \text{for all } z \in \mathbf{Z}.$$

In this case, $\mathbf{Z}$ is said to be a *transverse complete intersection*. Such variety $\mathbf{Z}$ is a smooth submanifold of $\mathbb{R}^n$ of codimension m and, in particular, has a tangent plane $T_z\mathbf{Z}$ at every point $z \in \mathbf{Z}$. It is remarked that the entire Euclidean space $\mathbb{R}^n$ is trivially considered a transverse complete intersection.

The proof of Theorem 1.4 relies on analysing geometric interactions between the tubes $T_{\theta,v}$ arising from the wave packet decomposition and varieties $\mathbf{Z}$. Given a wave packet $(\theta, v) \in \mathbb{T}[r]$ and $y \in \mathbb{R}^n$ let $T_{\theta,v}(y) := y + T_{\theta,v}$.

Definition 3.2. Let $\mathbf{Z}$ be a codimension m transverse complete intersection and fix a ball $B(y, r) \subseteq \mathbb{R}^n$. The tube $T_{\theta,v}(y)$ associated to a wave packet $(\theta, v) \in \mathbb{T}[r]$ is said to be $r^{-1/2+\delta_m}$-*tangent* to $\mathbf{Z}$ in $B(y, r)$ if the following conditions hold:

i) $T_{\theta,v}(y) \subseteq N_{r^{1/2+\delta_m}}\mathbf{Z} \cap B(y, r)$;
ii) For any $x \in T_{\theta,v}(y)$ and $z \in \mathbf{Z} \cap B(y, r)$ with $|x - z| \lesssim r^{1/2+\delta_m}$ one has

$$\angle\big(G(\omega_\theta), T_z\mathbf{Z}\big) \lesssim r^{-1/2+\delta_m}. \tag{3.2}$$

Here $G(\omega_\theta)$ is the direction of the tube $T_{\theta,v}$, as defined at the end of §3.1, and the left-hand side of (3.2) denotes the (unsigned) angle between this vector and $T_z\mathbf{Z}$. The δ_m exponent is as described in §2.

Such notions of tangency may be expressed more succinctly by introducing the concept of a 'grain', similar to that used in [9, 25]

Definition 3.3. For the purposes of this article, a *grain* is defined to be a pair (S, B_r) where $S \subseteq \mathbb{R}^n$ is a transverse complete intersection and $B_r \subset \mathbb{R}^n$ is a ball of some radius $r > 0$. The *(co)dimension* of a grain (S, B_r) is the (co)dimension of the variety S, whilst its *degree* is the degree of S and its *scale* is the value of the radial parameter r.

Definition 3.4. Let $\big(S, B(y, r)\big)$ be a codimension m grain. A function $f \in L^1(B^{n-1})$ is said to be *tangent* to $\big(S, B(y, r)\big)$ if it is concentrated on scale r wave packets belong to the collection

$$\big\{(\theta, v) \in \mathbb{T}[r] : \ T_{\theta,v}(y) \text{ is } r^{-1/2+\delta_m}\text{-tangent to } S \text{ in } B(y, r)\big\}. \tag{3.3}$$

The polynomial Wolff axioms may be used to study the geometry of tubes $T_{\theta,v}(y)$ satisfying the condition in (3.3); in particular, Theorem 1.2 was used in this way to prove restriction estimates in [7, 12]. Here more complex nested geometric structures are considered, which are described by the following definition (see also [13, 25]).

Definition 3.5 (Multigrain)**.** A *multigrain* is an $(m+1)$-tuple of grains

$$\vec{S}_m = (\mathcal{G}_0, \dots, \mathcal{G}_m), \qquad \mathcal{G}_i = (S_i, B_{r_i}) \text{ for } 0 \le i \le m \le n$$

satisfying

- $\operatorname{codim} S_i = i$ for $0 \le i \le m$,
- $S_m \subset S_{m-1} \subset \cdots \subset S_0$,
- $B_{r_m} \subseteq B_{r_{m-1}} \subseteq \cdots \subseteq B_{r_0}$.

The parameter m is referred to as the *level* of the multigrain. The *complexity* of the multigrain is defined to be the maximum of the degrees $\deg S_i$ over all $0 \le i \le m$. Finally, the *multiscale* of $\vec{S}_m$ is the tuple $\vec{r} = (r_0, r_1, \dots, r_m)$.

Given multigrains $\vec{S}_\ell$ and $\vec{S}_m$ of levels ℓ and m, respectively, write $\vec{S}_m \preceq \vec{S}_\ell$ if $\ell \le m$ and the grains forming the first $\ell + 1$ components of $\vec{S}_m$ agree those of $\vec{S}_\ell$.

3.3. **Nested tubes and the nested polynomial Wolff axioms.** The proof of Theorem 1.4 relies on an incidence estimate for families of tubes which have a certain multi-scale structure.

Definition 3.6. Let $\vec{S}_m = (\mathcal{G}_0, \dots, \mathcal{G}_m)$ be a multigrain with

$$\mathcal{G}_i = \big(S_i, B(y_i, r_i)\big) \quad \text{for } 0 \le i \le m.$$

Define $\mathbb{T}[\vec{S}_m]$ to be the set of scale $R := r_0$ wave packets $(\theta_0, v_0) \in \mathbb{T}[R]$ that satisfy:

Nested tube hypothesis. There exists $(\theta_i, v_i) \in \mathbb{T}[r_i]$ for $1 \le i \le m$ such that

i) $\operatorname{dist}(\theta_i, \theta_j) \lesssim r_j^{-1/2}$,

ii) $\operatorname{dist}\big(T_{\theta_j,v_j}(y_j),\, T_{\theta_i,v_i}(y_i) \cap B(y_j, r_j)\big) \lesssim r_i^{1/2+\delta}$,

iii) $T_{\theta_j,v_j}(y_j) \subset N_{r_j^{1/2+\delta_j}} S_j$

hold for all $0 \le i \le j \le m$. In each case dist can be taken to be the Hausdorff distance.

The direction set associated to $\mathbb{T}[\vec{S}_m]$ is given by

$$\Theta[\vec{S}_m] := \big\{\theta \in \Theta[R] : (\theta, v) \in \mathbb{T}[\vec{S}_m] \text{ for some } v \in R^{1/2}\mathbb{Z}^{n-1}\big\}.$$

Trivially, $\#\Theta[\vec{S}_m] \lesssim R^{(n-1)/2}$. However, the nested tube hypothesis further constrains the number of possible directions of the tubes in $\mathbb{T}[\vec{S}_m]$.

Lemma 3.7. *Let $\vec{S}_m$ be a level m multigrain with multiscale $\vec{r}_m = (r_0, \dots, r_m)$ and complexity at most d. If $R := r_0$ and the constants in* (2.1) *are chosen appropriately, then*

$$\#\Theta[\vec{S}_m] \lesssim_{n,d} \Big(\prod_{j=1}^{m} r_j^{-1/2}\Big) R^{(n-1)/2+\varepsilon_\circ}.$$

Lemma 3.7 is a direct consequence of the following variant of Theorem 1.3, which is deduced by combining [25, Lemma 2.11] with Wongkew's theorem [23].

Theorem 3.8 (Nested polynomial Wolff axioms [13, 25])**.** *For all $n > m \ge 1$, $d \ge 1$ and $\varepsilon > 0$, the bound*

$$\# \bigcap_{j=1}^{m} \Big\{ L \in \mathbf{L} : \mathcal{H}^1\big(L \cap B_{r_j} \cap N_{\rho_j}\mathbf{Z}_j\big) \ge r_j \Big\} \lesssim_{n,d,\varepsilon} \Big(\prod_{j=1}^{m} \frac{\rho_j}{r_j}\Big) R^{(n-1)/2+\varepsilon}$$

holds whenever:

- $(\rho_j)_{j=1}^m$ *and* $(r_j)_{j=1}^m$ *are non-increasing sequences lying in the interval* $[1, R]$ *and* $R^{-1/2} \leq \rho_1/r_1$*;*
- *The balls* B_{r_j} *are nested:* $B_{r_m} \subseteq \ldots \subseteq B_{r_1} \subset \mathbb{R}^n$*;*
- $\mathbf{L}$ *is a set of lines pointing in* $R^{-1/2}$*-separated directions;*
- *Each* $\mathbf{Z}_j \subset \mathbb{R}^n$ *is an algebraic variety of codimension* j *and degree at most* d*, and the varieties are nested:* $\mathbf{Z}_m \subset \mathbf{Z}_{m-1} \subset \cdots \subset \mathbf{Z}_1$.

The advantage of Theorem 3.8 compared with Theorem 1.3 is that the former allows for additional flexibility in the choice of the widths ρ_j of the neighbourhoods of the $\mathbf{Z}_j$.

Proof (of Lemma 3.7). Let $\mathbb{T} \subset \mathbb{T}[\vec{S}_m]$ be a set of wave packets pointing in different directions with $\#\mathbb{T} = \#\Theta[\vec{S}_m]$. For each wave packet $(\theta_0, v_0) \in \mathbb{T}$, let $(\theta_i, v_i) \in \mathbb{T}[r_i]$ for $i = 1, \ldots, m$ be the wave packets described in Definition 3.6. Let L_{θ_0,v_0} be the line parallel to T_{θ_0,v_0} that passes through the midpoint of the wave packet $T_{\theta_m,v_m}(y_m)$. It suffices to bound the cardinality of the family of lines $\mathbf{L} := \{L_{\theta_0,v_0} \colon (\theta_0, v_0) \in \mathbb{T}\}$. By construction, (after pigeonholing) one may assume that the lines in $\mathbf{L}$ point in $R^{-1/2}$-separated directions.

Let $L \in \mathbf{L}$ and let (θ_i, v_i), $i = 1, \ldots, m$, be the corresponding wave packets. By item i) and ii) from Definition 3.6, for each index $j = 0, \ldots, m$ it follows that

$$L \cap B(y_j, r_j) \subset N_{Cr_j^{1/2+\delta}} T_{\theta_j,v_j}(y_j).$$

Since $\delta < \delta_j$, by item iii),

$$\mathcal{H}^1\big(L \cap N_{Cr_j^{1/2+\delta_j}} S_j \cap B(y_j, r_j)\big) \geq r_j.$$

Applying Theorem 3.8, one concludes that

$$\begin{aligned} \#\mathbf{L} &\lesssim_{n,d} \Big(\prod_{j=1}^m \frac{r_j^{1/2+\delta_j}}{r_j}\Big) R^{(n-1)/2+\varepsilon_\circ/2} \\ &\leq \Big(\prod_{j=1}^m r_j^{-1/2}\Big) R^{(n-1)/2+\varepsilon_\circ/2+\delta_0+\ldots+\delta_m}. \end{aligned}$$

The result now follows, provided $\varepsilon_\circ > 2(\delta_0 + \ldots + \delta_m)$. □

4. An overview of the argument

4.1. **Multiscale grains decomposition.** The induction-on-scale argument from [10] may be interpreted as a procedure for decomposing the broad norm $\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)}$ into pieces with certain structural properties. Moreover, the relevant structure may be described in terms of multigrains $\vec{S}_\ell$ and the tube families $\mathbb{T}[\vec{S}_\ell]$, as introduced in §3. Here a succinct description of this decomposition is provided, based on the algorithms `[alg 1]` and `[alg 2]` from [12].

Consider a family of Lebesgue exponents p_i for $0 \leq i \leq n-k$ satisfying

$$p_{n-k} \geq p_{n-k+1} \geq \cdots \geq p_0 =: p \geq 2.$$

and define $0 \leq \alpha_i, \beta_i \leq 1$ in terms of the p_i by

$$\alpha_i := \Big(\frac{1}{2} - \frac{1}{p_i}\Big)\Big(\frac{1}{2} - \frac{1}{p_{i-1}}\Big)^{-1} \quad \text{and} \quad \beta_i := \Big(\frac{1}{2} - \frac{1}{p_i}\Big)\Big(\frac{1}{2} - \frac{1}{p_0}\Big)^{-1}$$

for $1 \leq i \leq n-k$ and $\alpha_0 :=: \beta_0 := 1$.

Input. Fix $R \gg 1$ and let $f\colon B^{n-1} \to \mathbb{C}$ be smooth and bounded and, without loss of generality, assume that f satisfies the *non-degeneracy hypothesis*

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \geq C_{\mathrm{hyp}} R^{\varepsilon} \|f\|_{L^2(B^{n-1})}$$

where C_{hyp} and $A \in \mathbb{N}$ are admissible constants which are chosen sufficiently large so as to satisfy the forthcoming requirements.

Output. The algorithm outputs the following objects:

- $\mathcal{O}$ a finite collection of open subsets of $\mathbb{R}^n$ of diameter at most $R^{\varepsilon_\circ}$.
- A codimension $0 \leq m \leq n-k$ integer parameter $1 \leq A_{m+1} \leq A$.
- An $(m+1)$-tuple of:
  - Scales $\vec{r} = (r_0, \dots, r_m)$ satisfying $R = r_0 > r_1 > \cdots > r_m$;
  - Large and (in general) non-admissible parameters $\vec{D} = (D_1, \dots, D_{m+1})$.
- For $0 \leq \ell \leq m$ a family $\vec{\mathcal{S}}_\ell$ of level ℓ multigrains. Each $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ has multiscale $\vec{r}_\ell = (r_0, \cdots, r_\ell)$ and complexity $O_\varepsilon(1)$. The families have a nested structure in the sense that for each $1 \leq \ell \leq m$ and each $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$, there exists some $\vec{S}_{\ell-1} \in \vec{\mathcal{S}}_{\ell-1}$ such that $\vec{S}_\ell \preceq \vec{S}_{\ell-1}$.
- For $0 \leq \ell \leq m$ an assignment of a function $f_{\vec{S}_\ell}$ to each $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$. Each $f_{\vec{S}_\ell}$ is tangent to (S_ℓ, B_{r_ℓ}), the final component of $\vec{S}_\ell$, in the sense of Definition 3.4.

The above data is chosen so that the following properties hold:

Property i). The inequality

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \leq M(\vec{r}, \vec{D}) \|f\|_{L^2(B^{n-1})}^{1-\beta_m} \Big(\sum_{O \in \mathcal{O}} \|Ef_O\|^{p_m}_{\mathrm{BL}^{p_m}_{k,A_{m+1}}(O)} \Big)^{\frac{\beta_m}{p_m}} \tag{P-i}$$

holds for

$$M(\vec{r}, \vec{D}) := \Big(\prod_{i=1}^{m} D_i \Big)^{m\delta} \Big(\prod_{i=1}^{m} r_i^{(\beta_{i-1}-\beta_i)/2} D_i^{(\beta_{i-1}-\beta_m)/2} \Big).$$

Property ii).

$$\sum_{O \in \mathcal{O}} \|f_O\|_2^2 \lesssim_\varepsilon \Big(\prod_{i=1}^{m+1} D_i^{1+\delta} \Big) R^{O(\varepsilon_\circ)} \|f\|^2_{L^2(B^{n-1})}. \tag{P-ii}$$

Property iii). For $1 \leq \ell \leq m$,

$$\max_{O \in \mathcal{O}} \|f_O\|_2^2 \lesssim_\varepsilon r_\ell^{-\ell/2} \prod_{i=\ell+1}^{m+1} r_i^{-1/2} D_i^{-(n-i)+\delta} R^{O(\varepsilon_\circ)} \max_{\vec{S}_\ell \in \vec{\mathcal{S}}_\ell} \|f_{\vec{S}_\ell}\|_2^2, \tag{P-iii}$$

where $r_{m+1} := 1$.

Properties i), ii) and iii) are stated explicitly in [12]: see Remark 4.1 below. The present argument requires one further property which does not appear in [12] but nevertheless follows as a consequence of the decomposition procedure described there. In order to state this property, for each multigrain $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ let

$$f^{\#}_{\vec{S}_\ell} := \sum_{(\theta,v) \in \mathbb{T}[\vec{S}_\ell]} f_{\theta,v}$$

where $\mathbb{T}[\vec{S}_\ell]$ is the collection of scale R wave packets introduced in Definition 3.6.

Property iv). For $1 \leq \ell \leq m$,

$$\|f_{\vec{S}_\ell}\|^2_{L^2(B^{n-1})} \lesssim_\varepsilon r_\ell^{\ell/2}\Big(\prod_{i=1}^{\ell} r_i^{-1/2} D_i^\delta\Big) R^{O(\varepsilon_\circ)} \|f^\#_{\vec{S}_\ell}\|_2^2 \tag{P-iv}$$

holds for all $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$.

To verify (P-iv) it is necessary to unpack some of the details of the polynomial partitioning algorithms described in [10, 12]. This is postponed until §6 below. Presently the above properties are combined with Lemma 3.7 to conclude the proof of the Theorem 1.4.

Remark 4.1. The decomposition corresponds to the final output of the algorithm `[alg 2]` from [12]. In particular:

- (P-i) corresponds to [12, (55)] and follows from Property I of `[alg 1]` and Property 1 of `[alg 2]`.
- (P-ii) corresponds to the second displayed equation on p.269 of [12] and follows from Property II of `[alg 1]` and Property 2 of `[alg 2]`.
- (P-iii) corresponds to [12, (57)] and follows from Property III of `[alg 1]` and Property 3 of `[alg 2]`.
- (P-iv) is related to [12, (63)], which follows from the local versions of the estimates in Property III of `[alg 1]` and Property 3 of `[alg 2]`. In §6 below (P-iv) is established by adapting the argument used to prove [12, (63)]. For this various auxiliary results are required, which are discussed in §5.

Note that the indexing used above is slightly different to that appearing in [12] since here the $\vec{S}_\ell$ are indexed according to codimension rather than dimension.

Remark 4.2. Note that the multiscale grains decomposition detailed above outputs a set of functions $\{f_{\vec{S}_\ell} \colon \vec{S}_\ell \in \vec{\mathcal{S}}_\ell\}$ and states certain inequalities that these functions satisfy. This remark provides an informal description of these functions and how they are constructed.

Recall that a multigrain $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ is a tuple $(\mathcal{G}_0, \mathcal{G}_1, \dots, \mathcal{G}_m)$. Here $\mathcal{G}_0$ corresponds to a choice of ball of radius R, and for each index $i = 1, \dots, m$, $\mathcal{G}_i$ is a pair (S_i, B_{r_i}), where S_i is a variety of codimension i and $B_{r_i} = B(y_i, r_i)$ is a ball of radius r_i.

When $\ell = 0$ and $\vec{S}_0 = (\mathcal{G}_0)$, then $f_{\vec{S}_0} = \sum_{(\theta,v)\in\mathbb{W}_0} f_{\theta,v}$, where $\mathbb{W}_0 \subset \mathbb{T}[r_0]$ consists of those scale $r_0 = R$ wave packets (θ, v) for which the associated tube $T_{\theta,v}$ intersects B_{r_0}.

Now suppose $1 \leq i \leq m$ and $f_{\vec{S}_{i-1}}$ has been defined for all multigrains $\vec{S}_{i-1}$ of level $i-1$. Fixing $\vec{S}_i$ a multigrain of level i, the function $f_{\vec{S}_i}$ may be described as follows. Let $\vec{S}_{i-1}$ be the (unique) multigrain from $\vec{\mathcal{S}}_{i-1}$ with $\vec{S}_i \preceq \vec{S}_{i-1}$. Then heuristically, $f_{\vec{S}_i}$ should be thought of as

$$f_{\vec{S}_i} \text{“}=\text{”} \sum_{(\theta,v)\in\mathbb{W}_i} (f_{\vec{S}_{i-1}})_{\theta,v}, \tag{4.1}$$

where $\mathbb{W}_i$ consists of those wave packets $(\theta, v) \in \mathbb{T}[r_i]$ for which the associated scale r_i tube $T_{\theta,v}(y_i)$ is $r_i^{-1/2+\delta_i}$ tangent to S_i in $B(y_i, r_i)$.

Note the quotation marks around the equality in (4.1), which are intended warn the reader that (4.1) is not true in a literal sense. The reason for this is that $f_{\vec{S}_i}$ is constructed from $f_{\vec{S}_{i-1}}$ over many steps using an iterative process. In each one of these steps one performs a new wave packet decomposition at some intermediate scale between r_{i-1} and r_i, and some of the wave packets might be discarded at each stage. The arguments have been carefully crafted so that these discarded wave packets can be ignored without adversely affecting our estimates, but caution is needed because some of the more straightforward, naïve statements which would follow from (4.1) are not true.

Since each $O \in \mathcal{O}$ has diameter at most $R^{\varepsilon_\circ}$, trivially one may bound

$$\|Ef_O\|_{\mathrm{BL}^{p_m}_{k,A_{m-1}}(O)} \lessapprox_\varepsilon R^{O(\varepsilon_\circ)}\|f_O\|_{L^2(B^{n-1})}$$

This trivial bound can be applied to the right-hand side of (P-i) and combined with (P-ii) and the definition of $M(\vec{r}, \vec{D})$ to deduce that

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lessapprox_\varepsilon \prod_{i=1}^{m+1} r_i^{\frac{\beta_{i-1}-\beta_i}{2}} D_i^{\frac{\beta_{i-1}}{2}-(\frac{1}{2}-\frac{1}{p})+O(\delta)} R^{O(\varepsilon_\circ)}\|f\|_2^{2/p} \max_{O\in\mathcal{O}}\|f_O\|_2^{1-2/p}.$$

The problem is now to bound the maximum appearing on the right-hand side of this expression.

4.2. **Improvement using the multiscale polynomial Wolff axioms.** To obtain an improved result, the multigrain structure is exploited using Lemma 3.7.

Lemma 4.3. *For $m \le \ell \le n$,*

$$\max_{\vec{S}_\ell\in\vec{\mathcal{S}}_\ell}\|f^{\#}_{\vec{S}_\ell}\|_2^2 \lessapprox_\varepsilon \Big(\prod_{i=1}^{\ell} r_i^{-1/2}\Big)R^{\varepsilon_\circ}\|f\|^2_{L^\infty(B^{n-1})}.$$

Proof. Letting $\|\cdot\|_{L^2_{\mathrm{avg}}(\theta)}$ denote the L^2-norm taken with respect to the normalised (to have mass 1) Lebesgue measure on θ, one may write

$$\|f^{\#}_{\vec{S}_\ell}\|_2^2 \sim R^{-(n-1)/2}\sum_{\theta\in\Theta(R)}\|f^{\#}_{\vec{S}_\ell}\|^2_{L^2_{\mathrm{avg}}(\theta)},$$

where the right-hand sum can of course be restricted to those θ that intersect $\operatorname{supp} f^{\#}_{\vec{S}_\ell}$. Since $f^{\#}_{\vec{S}_\ell}$ is the sum of $f_{\theta,v}$ over all $(\theta, v) \in \mathbb{T}[\vec{S}_\ell]$, it follows that

$$\#\big\{\theta\in\Theta(R) : \theta\cap\operatorname{supp} f^{\#}_{\vec{S}_\ell} \neq \emptyset\big\} \lessapprox \#\Theta[\vec{S}_\ell]$$

where $\Theta[\vec{S}_\ell]$ is as defined in Definition 3.6. Consequently,

$$\|f^{\#}_{\vec{S}_\ell}\|^2_{L^2(B^{n-1})} \lessapprox R^{-(n-1)/2}\cdot\#\Theta[\vec{S}_\ell]\cdot\max_{\theta\in\Theta(R)}\|f^{\#}_{\vec{S}_\ell}\|^2_{L^2_{\mathrm{avg}}(\theta)}.$$

By orthogonality between the wave packets,

$$\max_{\theta\in\Theta(R)}\|f^{\#}_{\vec{S}_\ell}\|^2_{L^2_{\mathrm{avg}}(\theta)} \lessapprox \max_{\theta\in\Theta(R)}\|f\|^2_{L^2_{\mathrm{avg}}(\theta)} \le \|f\|^2_{L^\infty(B^{n-1})}.$$

On the other hand, Lemma 3.7 implies that

$$\#\Theta[\vec{S}_\ell] \lessapprox_\varepsilon \big(\prod_{i=1}^{\ell} r_i^{-1/2}\big)R^{(n-1)/2+\varepsilon_\circ}.$$

Combining the three previous displays yields the desired result. ☐

Combining Lemma 4.3 with property (P-iv) of the decomposition,

$$\max_{\vec{S}_\ell\in\vec{\mathcal{S}}_\ell}\|f_{\vec{S}_\ell}\|_2^2 \lessapprox_\varepsilon r^{\ell/2}\Big(\prod_{i=1}^{\ell} r_i^{-1/2}\Big)^2\Big(\prod_{i=1}^{\ell} D_i^\delta\Big)R^{O(\varepsilon_\circ)}\|f\|^2_{L^\infty(B^{n-1})}.$$

Substituting this estimate into the right-hand side of (P-iii), one concludes that

$$\max_{O\in\mathcal{O}}\|f_O\|_2^2 \lessapprox_\varepsilon \Big(\prod_{i=1}^{m} r_i^{-1/2}D_i^\delta\Big)\Big(\prod_{i=1}^{\ell} r_i^{-1/2}\Big)\Big(\prod_{i=\ell+1}^{m+1} D_i^{-(n-i)}\Big)R^{O(\varepsilon_\circ)}\|f\|^2_{L^\infty(B^{n-1})} \tag{4.2}$$

for all $0 \le \ell \le m$.

4.3. **Fixing the exponents and concluding the argument.** Varying the ℓ parameter in (4.2) produces $m+1$ different bounds. Combining these inequalities by taking a weighted geometric mean, one arrives at the following key estimate.

Key estimate. *Let $0 \leq \gamma_0, \dots, \gamma_m \leq 1$ satisfy $\sum_{j=0}^m \gamma_j = 1$. Then*

$$\max_{O \in \mathcal{O}} \|f_O\|_2^2 \lesssim_\varepsilon \prod_{i=1}^{m+1} r_i^{-\frac{1+\sigma_i}{2}} D_i^{-(n-i)(1-\sigma_i)+O(\delta)} R^{O(\varepsilon_\circ)} \|f\|_{L^\infty(B^{n-1})}^2,$$

where $\sigma_i := \sum_{j=i}^m \gamma_j$ for $0 \leq i \leq m$ and $\sigma_{m+1} := 0$.

The key estimate may be plugged into the earlier inequality

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim_\varepsilon \prod_{i=1}^{m+1} r_i^{\frac{\beta_{i-1}-\beta_i}{2}} D_i^{\frac{\beta_{i-1}}{2}-(\frac{1}{2}-\frac{1}{p})+O(\delta)} R^{O(\varepsilon_\circ)} \|f\|_2^{2/p} \max_{O \in \mathcal{O}} \|f_O\|_2^{1-2/p}$$

to yield the bound

$$\|Ef\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim_\varepsilon \prod_{i=1}^{m+1} r_i^{X_i} D_i^{Y_i+O(\delta)} R^{O(\varepsilon_\circ)} \|f\|_{L^\infty(B^{n-1})} \tag{4.3}$$

where, recalling $p = p_0$, the X_i, Y_i exponents are given by

$$X_i := \frac{\beta_{i-1}-\beta_i}{2} - \frac{(1+\sigma_i)}{2}\Big(\frac{1}{2} - \frac{1}{p_0}\Big);$$
$$Y_i := \frac{\beta_{i-1}}{2} - \big(1 + (n-i)(1-\sigma_i)\big)\Big(\frac{1}{2} - \frac{1}{p_0}\Big).$$

At this point the values of the free parameters p_i and γ_i are fixed. Define

$$\gamma_j := \frac{n-m-1}{2} \cdot \frac{1}{(n-j)(n-j-1)} \cdot \prod_{i=n-m}^{n-j} \frac{2i}{2i+1} \quad \text{for } 1 \leq j \leq m,$$

$$\gamma_0 := 1 - \sum_{j=1}^m \gamma_j,$$

so that the γ_j sum to 1. One may also show, using some algebra, that $0 \leq \gamma_j \leq 1$. Let $p_m := 2 \cdot \frac{n-m}{n-m-1}$ and define the remaining p_i in terms of the γ_j via the equation

$$\Big(\frac{1}{2} - \frac{1}{p_i}\Big)^{-1} = 2n - m - i + \sum_{j=i+1}^m (j-i)\gamma_j.$$

With these parameter choices one may verify using simple (yet rather lengthy) algebraic manipulations that $X_i, Y_i = 0$ for all $1 \leq i \leq m$ and $Y_{m+1} = 0$ and, furthermore,

$$p_0 = 2 + \frac{6}{2(n-1) + (n-m-1)\prod_{i=n-m}^{n-1} \frac{2i}{2i+1}}.$$

These computations are similar to those appearing in [12] and are left to the interested reader. The worst case scenario (in the sense that p_0 is maximised) occurs when m is as large as possible. Recall that $0 \leq m \leq n-k$; by taking $m = n-k$ one obtains the exponent $p_n(k)$ featured in Theorem 1.4.

With the above choice of exponents, the $r_i^{X_i} D_i^{Y_i}$ factors in (4.3) are admissible (indeed, they are equal to 1). However, the $D_i^{O(\delta)}$ factors may still be large. To deal with this, one may slightly perturb the exponents to decrease the Y_i value so as to ensure that the $Y_i + O(\delta)$ is non-positive. This results in a slightly larger p value and, since $\delta > 0$ is arbitrary, establishes Theorem 1.4 in the open range $p > p_n(k)$. The closed range of estimates then follows trivially via Hölder's inequality, using the fact that R^ε losses are permitted in the constants. $\square$

5. L^2 ORTHOGONALITY AND TRANSVERSE EQUIDISTRIBUTION REVISITED

5.1. **Comparing wave packets at different scales.** Fix a large scale $r \geq 1$ and a smaller scale $r^{1/2} \leq \rho \leq r$. Given $g \in L^1(B^{n-1})$ one may form the scale r wave packet decomposition

$$g = \sum_{(\theta,v)\in\mathbb{T}[r]} g_{\theta,v} + \mathrm{RapDec}(r)\|g\|_2. \tag{5.1}$$

Alternatively, given a ball $B(y,\rho)$ with $y \in B(0,r)$ one may form the scale ρ wave packet decomposition of g over $B(y,\rho)$. In particular, let $\phi\colon \mathbb{R}^n \times \mathbb{R}^{n-1} \to \mathbb{R}$ denote the phase function associated to the extension operator E; that is,

$$\phi(x;\omega) := \langle x', \omega\rangle + x_n|\omega|^2, \qquad x = (x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R},\ \omega \in \mathbb{R}^{n-1}.$$

Write

$$\tilde{g}(\omega) := e^{i\phi(y;\omega)}g(\omega) \tag{5.2}$$

so that $Eg(x) = E\tilde{g}(\tilde{x})$ for $x = y + \tilde{x}$. One may then decompose

$$\tilde{g} = \sum_{(\tilde{\theta},\tilde{v})\in\mathbb{T}[\rho]} \tilde{g}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2. \tag{5.3}$$

Following the discussion in [10, §7], the purpose of this section is to compare properties of the decompositions (5.1) and (5.3) under various hypotheses on g. In particular, it is useful to study properties of functions formed by restricted sums of wave packets.

Definition 5.1. For $\widetilde{\mathbb{W}} \subseteq \mathbb{T}[\rho]$ define $\uparrow\widetilde{\mathbb{W}}$ to be the set of all $(\theta, v) \in \mathbb{T}[r]$ for which there exists some $(\tilde{\theta},\tilde{v}) \in \widetilde{\mathbb{W}}$ satisfying

i) $\mathrm{dist}(\tilde{\theta},\theta) \lesssim \rho^{-1/2}$;
ii) $\mathrm{dist}\big(T_{\tilde{\theta},\tilde{v}}(y),\, T_{\theta,v} \cap B(y,\rho)\big) \lesssim r^{1/2+\delta}$.

Furthermore, for any $g \in L^1(B^{n-1})$ let

$$\tilde{g}|_{\widetilde{\mathbb{W}}} := \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{W}}} \tilde{g}_{\tilde{\theta},\tilde{v}}, \qquad g|_{\uparrow\widetilde{\mathbb{W}}} := \sum_{(\theta,v)\in\uparrow\widetilde{\mathbb{W}}} g_{\theta,v}.$$

The definition of $\uparrow\widetilde{\mathbb{W}}$ is motivated by the following.

Lemma 5.2. *Given $g \in L^1(B^{n-1})$ and $\widetilde{\mathbb{W}} \subseteq \mathbb{T}[\rho]$, one has*

$$\tilde{g}|_{\widetilde{\mathbb{W}}} = \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{W}}} \big(g|_{\uparrow\widetilde{\mathbb{W}}}\big)^{\sim}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2.$$

In particular, combining Lemma 5.2 with the orthogonality property of the wave packets,

$$\|\tilde{g}|_{\widetilde{\mathbb{W}}}\|_2^2 \lesssim \|g|_{\uparrow\widetilde{\mathbb{W}}}\|_2^2 + \mathrm{RapDec}(r)\|g\|_2^2. \tag{5.4}$$

The proof of Lemma 5.2 follows directly from the observations in [10, §7].

Proof (of Lemma 5.2). Forming the scale r-wave decomposition and using the linearity of the mapping $f \mapsto \tilde{f}_{\tilde{\theta},\tilde{v}}$, one may write

$$\tilde{g}|_{\widetilde{W}} = \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{W}}}\ \sum_{(\theta,v)\in\mathbb{T}[r]} \big(g|_{\theta,v}\big)^{\sim}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2.$$

An integration-by-parts argument (see, for instance, [10, Lemma 7.1]) shows that the function $\big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}$ is rapidly decaying whenever $(\tilde{\theta},\tilde{v})$ fails to satisfy the conditions

$$\mathrm{dist}(\theta,\tilde{\theta}) \lesssim \rho^{-1/2} \quad \text{and} \quad |v-\tilde{v}| \lesssim r^{1/2+\delta}. \tag{5.5}$$

On the other hand, if $(\tilde{\theta},\tilde{v})$ does satisfy (5.5), then it is not difficult to show that

$$T_{\tilde{\theta},\tilde{v}}(y) \subseteq C\cdot T_{\theta,v} \cap B(y,\rho);$$

indeed, this is essentially part of the content of [10, Lemma 7.2]. Combining these observations, one deduces that $\big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}$ is rapidly decaying whenever $(\theta,v)\notin\uparrow\widetilde{\mathbb{W}}$, and the desired identity follows. $\square$

5.2. **Transverse equidistribution estimates revisited.** Continuing to work with the scales $r^{1/2}\le\rho\le r$ from the previous section, fix a transverse complete intersection $\mathbf{Z}$ in $\mathbb{R}^n$ of codimension j and degree at most d. Suppose $g\in L^1(B^{n-1})$ concentrated on scale r wave packets belonging to the family

$$\mathbb{T}_{\mathbf{Z}} := \big\{(\theta,v)\in\mathbb{T}[r] : T_{\theta,v} \text{ is } r^{-1/2+\delta_j}\text{-tangent to } \mathbf{Z} \text{ in } B(0,r)\big\}.$$

It is shown in [10, §7] that $\tilde{g}$ is concentrated on scale ρ wave packets which belong to the union of the sets

$$\widetilde{\mathbb{T}}_{\mathbf{Z}+b} := \big\{(\tilde{\theta},\tilde{v})\in\mathbb{T}[\rho] : T_{\tilde{\theta},\tilde{v}}(y) \text{ is } \rho^{-1/2+\delta_j}\text{-tangent to } \mathbf{Z}+b \text{ in } B(y,\rho)\big\} \tag{5.6}$$

as b varies over vectors in $\mathbb{R}^n$ satisfying $|b|\lesssim r^{1/2+\delta_j}$. Thus, one is led to consider the functions

$$\tilde{g}_b := \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\mathbf{Z}+b}} \tilde{g}_{\tilde{\theta},\tilde{v}}.$$

In this subsection certain L^2 bounds for the functions $\tilde{g}_b$ obtained in [10] are generalised, in view of establishing property (P-iv) of the decomposition from the previous section.

Lemma 5.3. *Suppose $\mathbf{Z}$ is a transverse complete intersection of codimension j and degree at most d and $b\in\mathbb{R}^n$ with $|b|\lesssim r^{1/2+\delta_j}$. If $g\in L^1(B^{n-1})$ is concentrated on wave packets from $\mathbb{T}_{\mathbf{Z}}$ and $\widetilde{\mathbb{W}}\subseteq\widetilde{\mathbb{T}}_{\mathbf{Z}+b}$, then*

$$\|\tilde{g}|_{\widetilde{\mathbb{W}}}\|_2^2 \lesssim r^{O(\delta_j)}(r/\rho)^{-j/2}\|g|_{\uparrow\widetilde{\mathbb{W}}}\|_2^2 + \mathrm{RapDec}(r)\|g\|_2^2.$$

Remark 5.4. By the orthogonality properties of the wave packets,

$$\|g|_{\uparrow\widetilde{\mathbb{W}}}\|_2^2 \lesssim \|g\|_2^2.$$

On the other hand, if $\widetilde{\mathbb{W}}=\widetilde{\mathbb{T}}_{\mathbf{Z}+b}$, then $\tilde{g}|_{\widetilde{\mathbb{W}}}=\tilde{g}_b$ and so Lemma 5.3 implies that

$$\|\tilde{g}_b\|_2^2 \lesssim r^{O(\delta_j)}(r/\rho)^{-j/2}\|g\|_2^2,$$

which is precisely the estimate from [10, Lemma 7.6].

The proof of Lemma 5.3 follows from a minor modification of the argument used to establish [10, Lemma 7.6]. In particular, the key ingredient is an auxiliary inequality from [10, Lemma 7.5]; in order to recall this lemma, a few preliminary definitions are in order.

Partition $\mathbb{T}[r]$ into disjoint sets $\mathbb{T}_{\kappa,w}$ indexed by $(\kappa,w)\in\mathcal{T}:=\Theta[\rho]\times r^{1/2}\mathbb{Z}^{n-1}$ satisfying

$$\mathrm{dist}(\theta,\kappa)\lesssim\rho^{-1/2} \quad \text{and} \quad |v+(\partial_\omega\phi)(y;\omega_\theta)-w|\lesssim r^{1/2} \quad \text{for all } (\theta,v)\in\mathbb{T}_{\kappa,w}.$$

Accordingly, write

$$g_{\kappa,w} := \sum_{(\theta,v)\in\mathbb{T}_{\kappa,w}} g_{\theta,v} \qquad \text{for all } (\kappa,w)\in\mathcal{T}$$

so that

$$g = \sum_{(\kappa,w)\in\mathcal{T}} g_{\kappa,w} + \mathrm{RapDec}(r)\|g\|_2. \tag{5.7}$$

This decomposition satisfies the following properties:

- By the L^2-orthogonality between the wave packets,

$$\|g\|_2^2 \sim \sum_{(\kappa,w)\in\mathcal{T}} \|g_{\kappa,w}\|_2^2 \qquad \text{and} \qquad \|g_{\kappa,w}\|_2^2 \sim \sum_{(\theta,v)\in\mathbb{T}_{\kappa,w}} \|g_{\theta,v}\|_2^2. \tag{5.8}$$

- Each $(g_{\kappa,w})\tilde{}$ is concentrated on scale ρ wave packets belonging to

$$\widetilde{\mathbb{T}}_{\kappa,w} := \big\{(\tilde{\theta},\tilde{v}) \in \mathbb{T}[\rho] : \mathrm{dist}(\tilde{\theta},\kappa) \lesssim \rho^{-1/2} \text{ and } |\tilde{v}-w| \lesssim r^{1/2}\big\};$$

  see [10, Lemma 7.3]. The sets $\widetilde{\mathbb{T}}_{\kappa,w}$ form a finitely-overlapping cover of $\mathbb{T}[\rho]$ as (κ,w) varies over $\mathcal{T}$.
- If $(\theta,v)\in\mathbb{T}_{\kappa,w}$ and $(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\kappa,w}$ for some $(\kappa,w)\in\mathcal{T}$, then

$$\mathrm{dist}(\theta,\tilde{\theta}) \lesssim \rho^{-1/2} \quad \text{and} \quad \mathrm{dist}\big(T_{\tilde{\theta},\tilde{v}}(y),\, T_{\theta,v}\cap B(y,\rho)\big) \lesssim r^{1/2+\delta}. \tag{5.9}$$

  Indeed, by the definition of $\mathbb{T}_{\kappa,w}$ and $\widetilde{\mathbb{T}}_{\kappa,w}$, it follows directly that

$$\mathrm{dist}(\theta,\tilde{\theta}) \lesssim \rho^{-1/2} \quad \text{and} \quad |v + (\partial_\omega\phi)(y;\omega_\theta) - \tilde{v}| \lesssim r^{1/2}. \tag{5.10}$$

  This immediately establishes the first inequality in (5.9) and, in fact, the second inequality in (5.9) also follows from (5.10); see [10, Lemma 7.2].

Furthermore, the following *transverse equidistribution estimate* holds for functions simultaneously concentrated on wave packets from $\mathbb{T}_{\mathbf{Z}}$ and some $\mathbb{T}_{\kappa,w}$.

Lemma 5.5 ([10, Lemma 7.5])**.** *Suppose $h \in L^1(B^{n-1})$ is concentrated on wave packets from $\mathbb{T}_{\mathbf{Z}}\cap\mathbb{T}_{\kappa,w}$ for some $(\kappa,w)\in\mathcal{T}$ and $|b|\lesssim r^{1/2+\delta_j}$. Then,*

$$\|\tilde{h}_b\|_2^2 \lesssim r^{O(\delta_j)}(r/\rho)^{-j/2}\|h\|_2^2.$$

The auxillary estimate from Lemma 5.5 may be combined with the various properties of the functions $\tilde{g}_{\kappa,w}$ described above in order to establish Lemma 5.3.

Proof (of Lemma 5.3). Since $g \mapsto \tilde{g}|_{\widetilde{\mathbb{W}}}$ is a linear operation on $L^1(\mathbb{R}^{n-1})$, it follows from (5.7) that

$$\tilde{g}|_{\widetilde{\mathbb{W}}} = \sum_{(\kappa,w)\in\mathcal{T}} (g_{\kappa,w})\tilde{}|_{\widetilde{\mathbb{W}}} + \mathrm{RapDec}(r)\|g\|_2. \tag{5.11}$$

Each $(g_{\kappa,w})\tilde{}$ is concentrated on scale ρ wave packets belonging to $\widetilde{\mathbb{T}}_{\kappa,w}$ and, consequently,

$$(g_{\kappa,w})\tilde{}|_{\widetilde{\mathbb{W}}} = \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}} (g_{\kappa,w})\tilde{}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2. \tag{5.12}$$

Combining (5.11) and (5.12), one deduced that

$$\tilde{g}|_{\widetilde{\mathbb{W}}} = \sum_{\substack{(\kappa,w)\in\mathcal{T}\\ \widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}\neq\emptyset}} \sum_{(\tilde{\theta},\tilde{v})\in\widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}} (g_{\kappa,w})\tilde{}_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2$$

and thus, since the $\widetilde{\mathbb{T}}_{\kappa,w}$ are finite-overlapping, by the orthogonality between the wave packets,

$$\|\tilde{g}|_{\widetilde{\mathbb{W}}}\|_2^2 \lesssim \sum_{\substack{(\kappa,w)\in\mathcal{T}\\ \widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}\neq\emptyset}} \|(g_{\kappa,w})\tilde{}|_{\widetilde{\mathbb{W}}}\|_2^2 + \mathrm{RapDec}(r)\|g\|_2^2. \tag{5.13}$$

As $\widetilde{\mathbb{W}} \subseteq \widetilde{\mathbb{T}}_{\mathbf{Z}+b}$, again using the orthogonality between the wave packets

$$\|(g_{\kappa,w})\tilde{}|_{\widetilde{\mathbb{W}}}\|_2^2 \lesssim \|(g_{\kappa,w})\tilde{}_b\|_2^2. \tag{5.14}$$

Since $g_{\kappa,w}$ is concentrated on wave packets belonging to $\mathbb{T}_{\mathbf{Z}} \cap \mathbb{T}_{\kappa,w}$, one may apply Lemma 5.5 to conclude that

$$\|(g_{\kappa,w})_{\widetilde{b}}\|_2^2 \lesssim r^{O(\delta_j)}(r/\rho)^{-j/2}\|g_{\kappa,w}\|_2^2. \tag{5.15}$$

Combining (5.13), (5.14) and (5.15) together with the second orthogonality relation in (5.8), one obtains

$$\|\tilde{g}|_{\widetilde{\mathbb{W}}}\|_2^2 \lesssim \sum_{\substack{(\kappa,w)\in\mathcal{T} \\ \widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}\neq\emptyset}} \sum_{(\theta,v)\in\mathbb{T}_{\kappa,w}} \|g_{\theta,v}\|_2^2.$$

Thus, by yet another application of the orthogonality property, the problem is reduced to showing that

$$\bigcup_{\substack{(\kappa,w)\in\mathcal{T} \\ \widetilde{\mathbb{T}}_{\kappa,w}\cap\widetilde{\mathbb{W}}\neq\emptyset}} \mathbb{T}_{\kappa,w} \subseteq \uparrow\widetilde{\mathbb{W}}. \tag{5.16}$$

This last step follows from (5.9). Indeed, suppose $(\theta, v) \in \mathbb{T}[r]$ belongs to the left-hand set in (5.16) so that there exists some $(\kappa, w) \in \mathcal{T}$ such that $(\theta, v) \in \mathbb{T}_{\kappa,w}$ and $\widetilde{\mathbb{T}}_{\kappa,w} \cap \widetilde{\mathbb{W}} \neq \emptyset$. If $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{T}}_{\kappa,w} \cap \widetilde{\mathbb{W}}$, then $(\tilde{\theta}, \tilde{v}) \in \widetilde{\mathbb{W}}$ satisfies (5.9), and therefore $(\theta, v) \in \uparrow\widetilde{\mathbb{W}}$ by the definition of the latter set. □

5.3. **Repeatedly refining the wave packets.** This subsection deals with a technical lemma which is useful when one wishes to repeatedly form refinements of the wave packet decomposition at a given scale.

Given $\mathbb{W} \subseteq \mathbb{T}[r]$ let $\mathbb{W}^* \subseteq \mathbb{T}[r]$ denote the slightly enlarged set of wave packets

$$\mathbb{W}^* := \big\{(\theta, v) \in \mathbb{T}[r] : \mathrm{dist}(\theta, \tilde{\theta}) \lesssim r^{-1/2} \text{ and } |v - \tilde{v}| \lesssim r^{1/2+\delta} \text{ for some } (\tilde{\theta}, \tilde{v}) \in \mathbb{W}\big\}.$$

Lemma 5.6. *If* $\mathbb{W}_1, \mathbb{W}_2 \subseteq \mathbb{T}[r]$ *and* $g \in L^1(B^{n-1})$*, then*

$$\big\|(g|_{\mathbb{W}_1})|_{\mathbb{W}_2}\big\|_2 \lesssim \big\|g|_{\mathbb{W}_1\cap\mathbb{W}_2^*}\big\|_2 + \mathrm{RapDec}(r)\|g\|_2.$$

Proof. Fix $(\tilde{\theta}, \tilde{v}) \in \mathbb{W}_2$ and note that

$$\big(g|_{\mathbb{W}_1}\big)_{\tilde{\theta},\tilde{v}} = \sum_{(\theta,v)\in\mathbb{W}_1} \big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}. \tag{5.17}$$

As in the proof of Lemma 5.2, the function $\big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}$ is rapidly decaying whenever $(\tilde{\theta}, \tilde{v}) \notin \mathbb{T}_{\theta,v}$ where

$$\mathbb{T}_{\theta,v} := \big\{(\tilde{\theta}, \tilde{v}) \in \mathbb{T}[r] : \mathrm{dist}(\theta, \tilde{\theta}) \lesssim r^{-1/2} \text{ and } |v - \tilde{v}| \lesssim r^{1/2+\delta}\big\};$$

see, for instance, [10, Lemma 7.1]. The condition $(\tilde{\theta}, \tilde{v}) \notin \mathbb{T}_{\theta,v}$ is equivalent to $(\theta, v) \notin \mathbb{T}_{\tilde{\theta},\tilde{v}}$ and so (5.17) implies that

$$\big(g|_{\mathbb{W}_1}\big)_{\mathbb{W}_2} = \sum_{(\tilde{\theta},\tilde{v})\in\mathbb{W}_2} \sum_{(\theta,v)\in\mathbb{W}_1\cap\mathbb{T}_{\tilde{\theta},\tilde{v}}} \big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}} + \mathrm{RapDec}(r)\|g\|_2.$$

Since $\#\mathbb{T}_{\tilde{\theta},\tilde{v}} = O(1)$, one may apply L^2-othogonality together with the Cauchy–Schwarz inequality to deduce that the inequalities

$$\begin{aligned}\big\|\big(g|_{\mathbb{W}_1}\big)_{\mathbb{W}_2}\big\|_2^2 &\lesssim \sum_{(\tilde{\theta},\tilde{v})\in\mathbb{W}_2} \sum_{(\theta,v)\in\mathbb{W}_1\cap\mathbb{T}_{\tilde{\theta},\tilde{v}}} \big\|\big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}\big\|_2^2 \\ &\lesssim \sum_{(\theta,v)\in\mathbb{W}_1\cap\mathbb{W}_2^*} \sum_{(\tilde{\theta},\tilde{v})\in\mathbb{T}[r]} \big\|\big(g_{\theta,v}\big)_{\tilde{\theta},\tilde{v}}\big\|_2^2\end{aligned}$$

hold up to the inclusion of a rapidly decaying error. Further application of L^2-orthogonality then yields the desired estimate. □

6. Relating the scales: verifying Property iv)

6.1. **The first algorithm.** Throughout this section let $p \geq 2$ be fixed and

$$\varepsilon^C \leq \delta \ll_\varepsilon \delta_0 \ll_\varepsilon \delta_1 \ll_\varepsilon \cdots \ll_\varepsilon \delta_{n-k} \ll_\varepsilon \varepsilon_\circ \ll_\varepsilon \varepsilon$$

be the family of small parameters described in §2. It will be useful to also work with auxiliary numbers $\tilde{\delta}_j$ defined by

$$\big(1 - \tilde{\delta}_{j+1}\big)\big(\tfrac{1}{2} + \delta_{j+1}\big) = \tfrac{1}{2} + \delta_j,$$

so that $\delta_j/2 \leq \tilde{\delta}_j \leq 2\delta_j$ for all $0 \leq j \leq n-k$.

<u>`Input`</u> The algorithm `[alg 1*]` takes as its input:

- A grain $\big(\mathbf{Z}, B(y,r)\big)$ of codimension m.
- A function $f \in L^1(B^{n-1})$ which is tangent to $\big(\mathbf{Z}, B(y,r)\big)$.
- An admissible large integer $A \in \mathbb{N}$.

<u>`Output`</u> The jth stage of `[alg 1*]` outputs:

- A choice of spatial scale $\rho_j \geq 1$ satisfying $\rho_j \leq \rho_{j-1}/2$ and
$$\rho_j \leq r^{(1-\tilde{\delta}_{m+1})^{\#\mathtt{a}(j)}} \quad \text{and} \quad \rho_j \leq \frac{r}{2^{\#\mathtt{c}(j)}}$$
for certain integers $\#\mathtt{a}(j), \#\mathtt{c}(j) \in \mathbb{N}_0$ satisfying $\#\mathtt{a}(j) + \#\mathtt{c}(j) = j$.
- A family of subsets $\mathcal{O}_j$ of $\mathbb{R}^n$ referred to as *cells*. Each cell $O_j \in \mathcal{O}_j$ is contained in some ρ_j-ball $B_{O_j} = B(y_{O_j}, \rho_j)$.
- A collection of functions $(f_{O_j})_{O_j \in \mathcal{O}_j}$. For each cell O_j there is a translate $\mathbf{Z}_{O_j} := \mathbf{Z} + x_{O_j}$ such that f_{O_j} is tangent to the grain $(\mathbf{Z}_{O_j}, B_{O_j})$.
- A large integer $d \in \mathbb{N}$ which depends only on the admissible parameters and $\deg \mathbf{Z}$.

Moreover, the components of the ensemble are defined so as to ensure that, for certain coefficients

$$C^{\mathrm{I}}_{j,\delta}(d,r),\ C^{\mathrm{II}}_{j,\delta}(d),\ C^{\mathrm{III}}_{j,\delta}(d,r),\ C^{\mathrm{IV}}_{j,\delta}(d,r) \lesssim_{d,\delta} r^{\varepsilon_\circ} d^{\#\mathtt{c}(j)\delta} \tag{6.1}$$

and $A_j := 2^{-\#\mathtt{a}(j)} A \in \mathbb{N}$, the following properties hold:

<u>Property I</u>. Most of the mass of $\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_r)}$ is concentrated over the $O_j \in \mathcal{O}_j$:

$$\|Ef\|^p_{\mathrm{BL}^p_{k,A}(B_r)} \leq C^{\mathrm{I}}_{j,\delta}(d,r) \sum_{O_j \in \mathcal{O}_j} \|Ef_{O_j}\|^p_{\mathrm{BL}^p_{k,A_j}(O_j)} + jr^{-N}\|f\|^p_{L^2(B^{n-1})} \qquad (\mathrm{I})_j$$

for some large fixed $N \in \mathbb{N}$.

<u>Property II</u>. The functions f_{O_j} satisfy

$$\sum_{O_j \in \mathcal{O}_j} \|f_{O_j}\|_2^2 \leq C^{\mathrm{II}}_{j,\delta}(d) d^{\#\mathtt{c}(j)} \|f\|_2^2. \qquad (\mathrm{II})_j$$

<u>Property III</u>. Each f_{O_j} satisfies

$$\|f_{O_j}\|^2_{L^2(B^{n-1})} \leq C^{\mathrm{III}}_{j,\delta}(d,r)\Big(\frac{r}{\rho_j}\Big)^{-m/2} d^{-\#\mathtt{c}(j)(n-m-1)} \|f\|^2_{L^2(B^{n-1})}. \qquad (\mathrm{III})_j$$

Properties I, II and III are stated explicitly in the description of `[alg 1]` from [12, §9]. The modified algorithm `[alg 1*]` includes an additional property, described presently.

For $\mathbb{W} \subseteq \mathbb{T}[\rho_j]$ let $\uparrow^j \mathbb{W}$ denote the set of wave packets $(\theta, v) \in \mathbb{T}[r]$ satisfying

$$\mathrm{dist}(\theta, \theta_j) \leq c_j \rho_j^{-1/2} \quad \text{and} \quad \mathrm{dist}\big(T_{\theta_j, v_j}(y_{O_j}),\, T_{\theta,v}(y) \cap B_{O_j}\big) \leq c_j r^{1/2+\delta}$$

for some $(\theta_j, v_j) \in \mathbb{W}$. Here $(c_j)_{j=0}^{\infty}$ is positive sequence which is bounded above by an absolute constant $C_\circ$ and chosen so as to satisfy the forthcoming requirements of the argument. Furthermore, let

$$f_{O_j}|_{\mathbb{W}} := \sum_{(\theta_j, v_j) \in \mathbb{W}} (f_{O_j})|_{\theta_j, v_j} \qquad \text{and} \qquad f|_{\uparrow^j \mathbb{W}} := \sum_{\theta, v \in \uparrow^j \mathbb{W}} f_{\theta, v}.$$

Property IV. For any $\mathbb{W} \subseteq \mathbb{T}[\rho_j]$, each f_{O_j} satisfies

$$\|f_{O_j}|_{\mathbb{W}}\|_2^2 \leq C^{\mathrm{IV}}_{j,\delta}(d, r)\Big(\frac{r}{\rho_j}\Big)^{-m/2}\|f|_{\uparrow^j \mathbb{W}}\|_2^2. \tag{IV$_j$}$$

This concludes the description of the output of `[alg 1*]`.

`Stopping conditions`. The algorithm has two stopping conditions which are labelled `[tiny]` and `[tang*]`.

`Stop: [tiny]` The algorithm terminates if $\rho_j \leq r^{\tilde{\delta}_{m+1}}$.

In view of the additional Property IV above, the second stopping condition is slightly modified compared with that of `[alg 1]` of [12, §9].

`Stop: [tang*]` Let $C_{\mathtt{tang}}$ and C_{alg} be large, fixed dimensional constants and $\tilde{\rho} := \rho_j^{1-\tilde{\delta}_m}$. The algorithm terminates if there exist

- $\mathcal{S}$ a collection of grains $(S, B_{\tilde{\rho}})$ of codimension $m+1$, scale $\tilde{\rho}$ and degree at most $C_{\mathrm{alg}}d$;
- An assignment of a function f_S to each[4] $S \in \mathcal{S}$ which is tangent to $(S, B_{\tilde{\rho}})$

such that the following inequalities hold:

Condition I.

$$\sum_{O_j \in \mathcal{O}_j} \|Ef_{O_j}\|^p_{\mathrm{BL}^p_{k,A_j}(O_j)} \leq C_{\mathtt{tang}} \sum_{S \in \mathcal{S}} \|Ef_S\|^p_{\mathrm{BL}^p_{k,A_j/2}(B_{\tilde{\rho}})}.$$

Condition II.

$$\sum_{S \in \mathcal{S}} \|f_S\|^2_{L^2(B^{n-1})} \leq C_{\mathtt{tang}} r^{n\tilde{\delta}_m} \sum_{O_j \in \mathcal{O}_j} \|f_{O_j}\|^2_{L^2(B^{n-1})}.$$

Condition III.

$$\max_{S \in \mathcal{S}} \|f_S\|_2^2 \leq C_{\mathtt{tang}} \max_{O_j \in \mathcal{O}_j} \|f_{O_j}\|_2^2.$$

Conditions I, II and III are stated explicitly in the description of the stopping condition for `[alg 1]` from [12, §9]. The modified algorithm `[alg 1*]` includes an additional condition.

Condition IV. Given $(S, B(\tilde{y}, \tilde{\rho})) \in \mathcal{S}$ there exists some $O_j \in \mathcal{O}_j$ such that

$$\|\tilde{f}_S|_{\mathbb{W}}\|_2^2 \leq C_{\mathtt{tang}}\|f_{O_j}|_{\uparrow \mathbb{W}}\|_2^2$$

holds for all $\mathbb{W} \subseteq \mathbb{T}[\tilde{\rho}]$. Here $\uparrow \mathbb{W}$ is the set of all $(\theta, v) \in \mathbb{T}[\rho_j]$ for which there exists some $(\tilde{\theta}, \tilde{v}) \in \mathbb{W}$ satisfying

$$\mathrm{dist}(\tilde{\theta}, \theta) \lesssim \tilde{\rho}^{-1/2}, \quad \mathrm{dist}\big(T_{\tilde{\theta}, \tilde{v}}(\tilde{y}),\, T_{\theta, v}(y_{O_j}) \cap B(\tilde{y}, \tilde{\rho})\big) \lesssim \rho_j^{1/2+\delta}$$

for y_{O_j} the centre of B_{O_j}, whilst

$$\tilde{f}_S|_{\mathbb{W}} := \sum_{(\tilde{\theta}, \tilde{v}) \in \mathbb{W}} (f_S)^{\sim}_{\tilde{\theta}, \tilde{v}} \qquad \text{and} \qquad f_{O_j}|_{\uparrow \mathbb{W}} := \sum_{(\theta, v) \in \uparrow \mathbb{W}} (f_{O_j})_{\theta, v}.$$

The function $\tilde{f}_S$ is as defined in (5.2), taking y to be the centre of $B_{\tilde{\rho}}$.

[4]Here, by an abuse of notation, S is used to denote the grain $(S, B_{\tilde{\rho}})$.

6.2. **Ensuring Property IV.** The modified algorithm `[alg 1*]` is obtained by combining the recursive step from `[alg 1]` from [12, §9] with the L^2-orthogonality results from §5. The main task is to verify Property IV holds. For this, it is useful to work with an explicit formula for the constant

$$C_{j,\delta}^{\mathrm{IV}}(d,r) := d^{j\delta} r^{\bar{C}\#\mathtt{a}(j)\delta_m}, \tag{6.2}$$

where $\bar{C} \geq 1$ is a suitably large absolute constant. Note that this agrees with the definition of $C_{j,\delta}^{\mathrm{III}}(d,r)$ used in [12, §9] and the stopping condition `[tiny]` together with (6.1) ensure (6.1) holds in this case.

The f_{O_j} are defined recursively exactly as in `[alg 1]`. Recall that there are two cases to consider: the *cellular-dominant* case and *algebraic-dominant* case. For either situation the functions $f_{O_{j+1}}$ are obtained from the f_{O_j} via the same procedure. In particular, supposing ρ_{j+1}, O_{j+1}, $B_{O_{j+1}}$ and $x_{O_{j+1}}$ have already been defined, the function $f_{O_{j+1}}$ has the following form: for two sets of wave packets $\mathbb{T}[O_{j+1}] \subseteq \mathbb{T}[\rho_j]$ and $\widetilde{\mathbb{T}}[O_{j+1}] \subseteq \mathbb{T}[\rho_{j+1}]$ define

$$\tilde{f}_{O_{j+1}} := \tilde{g}_{O_{j+1}}|_{\widetilde{\mathbb{T}}[O_{j+1}]} \quad \text{for} \quad g_{O_{j+1}} := f_{O_j}|_{\mathbb{T}[O_{j+1}]}.$$

Here, given $g \in L^1(B^{n-1})$, the function $\tilde{g}$ is defined with respect to the ball $B_{O_{j+1}}$ as in §5.1 (that is, taking $y = y_{O_{j+1}}$ in (5.2)). The following table shows the choices of $\mathbb{T}[O_{j+1}]$ and $\widetilde{\mathbb{T}}[O_{j+1}]$ used in the cellular-dominant and algebraic-dominant cases in [12, §9] or [10].

Case	$\mathbb{T}[O_{j+1}]$	$\widetilde{\mathbb{T}}[O_{j+1}]$
Cellular-dominant	$\{(\theta,v) \in \mathbb{T}[\rho_j] : T_{\theta,v} \cap O_{j+1} \neq \emptyset\}$	$\widetilde{\mathbb{T}}_{Z_{O_{j+1}}}$
Algebraic-dominant	$\mathbb{T}_{B,\mathrm{trans}}$	$\widetilde{\mathbb{T}}_{Z_{O_{j+1}}}$

The set $\mathbb{T}_{B,\mathrm{trans}}$ is as defined in [12, p.257] or [10, p.129], but the precise choice of $\mathbb{T}[O_{j+1}]$ set is in fact unimportant for the purpose of establishing Property IV (the information is only included here as a reference to the arguments in [10, 12]). The sets $\widetilde{\mathbb{T}}_{Z_{O_{j+1}}}$ are as defined in (5.6) for $\rho := \rho_{j+1}$ and $y = y_{O_{j+1}}$.

Remark 6.1. The additional decomposition according to the $\widetilde{\mathbb{T}}_{Z_{O_{j+1}}}$ wave packets is not carried out in the cellular-dominant case in either [10] or [12] but is nevertheless useful in the argument: see [12, p.254, fn 11] or [11, Lemma 10.2] for further details.

With the general setup above, $(\mathrm{IV})_{j+1}$ may be established as follows. By a combination of Lemma 5.6 and the basic orthogonality between the wave packets,

$$\|\tilde{f}_{O_{j+1}}|_{\mathbb{W}}\|_2^2 \lesssim \|\tilde{g}_{O_{j+1}}|_{\widetilde{\mathbb{T}}[O_{j+1}]\cap\mathbb{W}^*}\|_2^2 \lesssim \|\tilde{g}_{O_{j+1}}|_{\mathbb{W}^*}\|_2^2$$

where $\mathbb{W}^*$ is the enlarged version of $\mathbb{W}$ defined in §5.3. Strictly speaking, these bounds should include additional rapidly decreasing error terms but, for simplicity, here and below these minor contributions are omitted. The transverse equidistribution estimate from Lemma 5.3 implies that

$$\|\tilde{g}_{O_{j+1}}|_{\mathbb{W}^*}\|_2^2 \lesssim \rho_j^{O(\delta_m)} \Big(\frac{\rho_j}{\rho_{j+1}}\Big)^{-m/2} \|g_{O_{j+1}}|_{\uparrow\mathbb{W}^*}\|_2^2 \tag{6.3}$$

whilst a second application of Lemma 5.6 and orthogonality yields

$$\|g_{O_{j+1}}|_{\uparrow\mathbb{W}^*}\|_2^2 \lesssim \|f_{O_j}|_{\mathbb{T}[O_{j+1}]\cap(\uparrow\mathbb{W}^*)^*}\|_2^2 \lesssim \|f_{O_j}|_{(\uparrow\mathbb{W}^*)^*}\|_2^2.$$

The set $(\uparrow \mathbb{W}^*)^*$ agrees with a set that is almost identical to $\uparrow \mathbb{W}$ except that certain constants in the definition of $\uparrow \mathbb{W}$ are slightly enlarged. By redefining $\uparrow \mathbb{W}$ using these larger constants, one concludes that

$$\|\tilde{f}_{O_{j+1}}|_{\mathbb{W}}\|_2^2 \lesssim \rho_j^{O(\delta_m)}\Big(\frac{\rho_j}{\rho_{j+1}}\Big)^{-m/2}\|f_{O_j}|_{\uparrow\mathbb{W}}\|_2^2. \tag{6.4}$$

The inequality (6.4) is in fact only useful in the algebraic-dominant case. The estimate is true in the cellular case but is not efficient since here the ratio of the scales ρ_j/ρ_{j+1} is small compared to the $\rho^{O(\delta_m)}$ factor. Instead, one may replace (6.3) in the above argument with the more elementary bound

$$\|\tilde{g}_{O_{j+1}}|_{\mathbb{W}^*}\|_2^2 \lesssim \|g_{O_{j+1}}|_{\uparrow\mathbb{W}^*}\|_2^2,$$

which follows directly from (5.4). Arguing in this way, one may strengthen (6.4) in the cellular case to an estimate without any additional $\rho_j^{O(\delta_m)}$.

Henceforth assume the algebraic-dominant case holds; the cellular-dominant case follows almost identically using the refined version of (6.4) described in the previous paragraph. Apply $(\mathrm{IV})_j$ to the right-hand side of (6.4) to deduce that

$$\|\tilde{f}_{O_{j+1}}|_{\mathbb{W}}\|_2^2 \lesssim \rho_j^{O(\delta_m)} C^{\mathrm{IV}}_{j,\delta}(d,r)\Big(\frac{r}{\rho_{j+1}}\Big)^{-m/2}\|f_{O_j}|_{\uparrow^j(\uparrow\mathbb{W})}\|_2^2. \tag{6.5}$$

It is claimed that $\uparrow^j(\uparrow \mathbb{W}) \subseteq \uparrow^{j+1}\mathbb{W}$. Once this is established, combing (6.5) with basic L^2-orthogonality gives

$$\|\tilde{f}_{O_{j+1}}|_{\mathbb{W}}\|_2^2 \leq C(\deg \mathbf{Z},\delta)\rho_j^{\bar{C}\delta_m} C^{\mathrm{IV}}_{j,\delta}(d,r)\Big(\frac{r}{\rho_{j+1}}\Big)^{-m/2}\|f_{O_j}|_{\uparrow^j(\uparrow\mathbb{W})}\|_2^2.$$

for suitable constants $C(\deg \mathbf{Z},\delta)$, $\bar{C} \geq 1$. Finally, from the formula (6.2) and the assumption that the algebraic-dominant case holds,

$$C(\deg \mathbf{Z},\delta)\rho_j^{\bar{C}\delta_m} C^{\mathrm{IV}}_{j,\delta}(d,r) \leq C^{\mathrm{IV}}_{j+1,\delta}(d,r),$$

provided the parameter d is chosen to be sufficiently large, depending only on the admissible parameters and $\deg \mathbf{Z}$. To see this, recall from the description of `[alg 1]` from [12, §9] that $\#_{\mathtt{a}}(j+1) = \#_{\mathtt{a}}(j)+1$ if the algebraic-dominant case holds. This concludes the proof of $(\mathrm{IV})_j$, except for establishing the inclusion $\uparrow^j(\uparrow \mathbb{W}) \subseteq \uparrow^{j+1}\mathbb{W}$.

Let $(\theta, v) \in \uparrow^j(\uparrow \mathbb{W})$ so that there exists some $(\theta_j, v_j) \in \uparrow \mathbb{W}$ such that

$$\mathrm{dist}(\theta,\theta_j) \leq c_j\rho_j^{-1/2} \quad \text{and} \quad \mathrm{dist}\big(T_{\theta_j,v_j}(y_{O_j}),\, T_{\theta,v}(y)\cap B_{O_j}\big) \leq c_j r^{1/2+\delta}.$$

On the other hand, since $(\theta_j, v_j) \in \uparrow \mathbb{W}$ there exists some $(\theta_{j+1}, v_{j+1}) \in \mathbb{W}$ such that

$$\mathrm{dist}(\theta_j,\theta_{j+1}) \leq C\rho_{j+1}^{-1/2};$$
$$\mathrm{dist}\big(T_{\theta_{j+1},v_{j+1}}(y_{O_{j+1}}),\, T_{\theta_j,v_j}(y_{O_j})\cap B_{O_{j+1}}\big) \leq C\rho_j^{1/2+\delta},$$

for an appropriate choice of C. At this point, fix the values of c_j to be

$$c_j := C\sum_{i=0}^{j-1}2^{-i/2} \leq C_\circ := C\sum_{i=0}^{\infty}2^{-i/2}.$$

Since $\rho_{i+1} \leq \rho_i/2$ for all $0 \leq i \leq j$, it follows from the preceding displays that

$$\mathrm{dist}(\theta,\theta_{j+1}) \leq \big(c_j(\rho_{j+1}/\rho_j)^{1/2} + C\big)\rho_{j+1}^{-1/2} \leq c_{j+1}\rho_{j+1}^{-1/2};$$
$$\mathrm{dist}\big(T_{\theta_{j+1},v_{j+1}}(y_{O_{j+1}}),\, T_{\theta,v}(y)\cap B_{O_{j+1}}\big) \leq \big(c_j + C2^{-j/2}\big)r^{1/2+\delta} = c_{j+1}r^{1/2+\delta}.$$

Thus, $(\theta, v) \in \uparrow^{j+1}\mathbb{W}$, as required. □

6.3. **The modified stopping condition.** The condition `[tang*]` in `[alg 1*]` is slightly different from the corresponding condition `[tang]` appearing in `[alg 1]` and, in particular, Condition IV must hold in order to trigger `[tang*]`. To incorporate this extra condition, the algorithm described in [12, §9] requires Conditions II, III and IV to hold for certain functions $f_{B,\mathrm{tang}}$. These functions are of the form $f_S := f_{O_j}|_{\mathbb{T}[S]}$ for some $\mathbb{T}[S] \subseteq \mathbb{T}[\rho_j]$, similar to those encountered in the previous subsection. This time $\mathbb{T}[S] := \mathbb{T}_{B,\mathrm{tang}}$, where the latter set is as defined in [12, p.237] or [10, p.129]. As before, the exact form of the set $\mathbb{T}[S]$ is not important for the purposes of verifying the Condition IV. Thus, one wishes to show that

$$\|\tilde{f}_S|_{\mathbb{W}}\|_2^2 \lesssim \|f_{O_j}|_{\uparrow\mathbb{W}}\|_2^2. \tag{6.6}$$

This inequality is easily deduced using the arguments of the previous subsection. In particular, (5.4) and Lemma 5.6 together imply that

$$\|(f_S)^{\sim}|_{\mathbb{W}}\|_2^2 \lesssim \|f_S|_{\uparrow\mathbb{W}}\|_2^2 \lesssim \|f_{O_j}|_{\mathbb{T}[S]\cap(\uparrow\mathbb{W})^*}\|_2^2.$$

The desired estimate (6.6) now follows from the L^2-orthogonality between the wave packets, provided, as before, that the constants in the definition of $\uparrow \mathbb{W}$ are slightly enlarged.

6.4. **The second algorithm: a sketch.** The multigrain decomposition from §4.1 is obtained by repeatedly applying `[alg 1*]` as part of a recursive procedure described in `[alg 2]` in [12, §10]. Here a brief sketch of this process is given.

At stage 0, one begins with the input of the multigrain decomposition from §4.1. After the ℓth stage a family of functions $f_{\vec{S}_\ell}$ has been constructed, indexed by $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ where $\vec{\mathcal{S}}_\ell$ is a collection of level ℓ multigrains at some scale r_ℓ and of complexity $O_\varepsilon(1)$. Each function $f_{\vec{S}_\ell}$ is tangent to S_ℓ, the codimension ℓ grain forming the final component of $\vec{S}_\ell$. Furthermore, the functions satisfy suitable "level ℓ variants" of the properties (P-i) to (P-iv): see [12, §10] for details.

To pass to the next stage of the construction, apply `[alg 1*]` to each function $f_{\vec{S}_\ell}$. Notice that these functions satisfy the tangency conditions required in the input of the algorithm. In each case, either `[alg 1*]` terminates due to `[tiny]` or due to `[tang*]`. Suppose that the inequality

$$\sum_{\vec{S}_\ell \in \vec{\mathcal{S}}_\ell} \|Ef_{\vec{S}_\ell}\|^{p_\ell}_{\mathrm{BL}^{p_\ell}_{k,A_\ell}(B_{r_\ell})} \leq 2 \sum_{\vec{S}_\ell \in \vec{\mathcal{S}}_{\ell,\mathtt{tiny}}} \|Ef_{\vec{S}_\ell}\|^{p_\ell}_{\mathrm{BL}^{p_\ell}_{k,A_\ell}(B_{r_\ell})} \tag{6.7}$$

holds, where the right-hand summation is restricted to those $S_\ell \in \vec{\mathcal{S}}_\ell$ for which `[alg 1*]` terminates owing to the stopping condition `[tiny]`. In this case, the process terminates and $m := \ell$. Defining the functions f_O appropriately via `[alg 1*]`, one may verify the properties of the multigrain decomposition from §4.1: see [12, §§9-10] and the following discussion for further details.

Alternatively, if (6.7) fails, then necessarily

$$\sum_{\vec{S}_\ell \in \vec{\mathcal{S}}_\ell} \|Ef_{\vec{S}_\ell}\|^{p_\ell}_{\mathrm{BL}^{p_\ell}_{k,A_\ell}(B_{r_\ell})} \leq 2 \sum_{\vec{S}_\ell \in \vec{\mathcal{S}}_{\ell,\mathtt{tang}}} \|Ef_{\vec{S}_\ell}\|^{p_\ell}_{\mathrm{BL}^{p_\ell}_{k,A_\ell}(B_{r_\ell})},$$

where the right-hand summation is restricted to those $S_\ell \in \vec{\mathcal{S}}_\ell$ for which `[alg 1*]` does not terminate owing to `[tiny]` and therefore terminates owing to `[tang*]`. For each $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ there exists a collection $\mathcal{S}_{\ell+1}[\vec{S}_\ell]$ of codimension $\ell + 1$ grains of degree $O_\varepsilon(1)$ which arise from `[tang*]` and, in particular, satisfy the Conditions I to IV with f replaced with $f_{\vec{S}_\ell}$. By appropriately pigeonholing, one may further assume that all the grains in $\mathcal{S}_{\ell+1}[\vec{S}_\ell]$ have a common scale $r_{\ell+1} < r_\ell$. A family of level $\ell+1$ multigrains is then defined by

$$\vec{\mathcal{S}}_{\ell+1} := \big\{(\vec{S}_\ell, S_{\ell+1}) : S_{\ell+1} \in \mathcal{S}_{\ell+1}[\vec{S}_\ell]\big\},$$

whilst $f_{\vec{S}_{\ell+1}} := (f_{\vec{S}_\ell})_{S_{\ell+1}}$, where the right-hand function satisfies the properties stated in `[tang*]`.

6.5. **Ensuring Property iv).** The procedure sketched in the previous subsection is precisely `[alg 2]` from [12] (which in turn corresponds to the induction-on-dimension used in [10]). The only modification required for the purposes of this article is to construct the functions $f^{\#}_{\vec{S}_\ell}$ described in §4.1 and ensure that Property iv) from §4.1 holds. This is achieved using the additional Property IV of `[alg 1*]`.

For $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ and $0 \le j \le \ell$, if $\big(S_j, B(y_j, r_j)\big)$ denotes the codimension j component of $\vec{S}_\ell$, then let $\mathbb{T}_{\mathrm{tang}}[S_j]$ denote the set of all scale r_j wave packets which are $r_j^{-1/2+\delta_j}$-tangent to S_j in $B(y_j, r_j)$. Thus, if $\vec{S}_\ell \preceq \vec{S}_j$ for some $0 \le j \le \ell$, then the function $f_{\vec{S}_j}$ is concentrated on wave packets belonging to $\mathbb{T}_{\mathrm{tang}}[S_j]$.

For $1 \le \ell \le m$, given $\mathbb{W} \subseteq \mathbb{T}[r_\ell]$, let $\uparrow\uparrow_\ell \mathbb{W}$ denote the set of wave packets $(\theta, v) \in \mathbb{T}[r_{\ell-1}]$ for which there exists some $(\tilde{\theta}, \tilde{v}) \in \mathbb{W}$ satisfying

$$\mathrm{dist}(\tilde{\theta}, \theta) \le C_\circ r_\ell^{-1/2} \quad \text{and} \quad \mathrm{dist}\big(T_{\tilde{\theta},\tilde{v}}(y_\ell),\, T_{\theta,v}(y_{\ell-1}) \cap B(y_\ell, r_\ell)\big) \le C_\circ r_{\ell-1}^{-1/2+\delta}.$$

Property IV of `[alg 1*]` and the definition of the stopping condition `[tang*]` together imply that for $1 \le \ell \le m$ and $\mathbb{W} \subseteq \mathbb{T}[r_\ell]$, the inequality

$$\big\|f_{\vec{S}_\ell}|_{\mathbb{W}}\big\|_2^2 \lesssim_\varepsilon \Big(\frac{r_{\ell-1}}{r_\ell}\Big)^{-(\ell-1)/2} D_\ell^\delta R^{O(\varepsilon_\circ)} \big\|f_{\vec{S}_{\ell-1}}|_{\uparrow\uparrow_\ell \mathbb{W}}\big\|_2^2 \tag{6.8}$$

holds whenever $\vec{S}_{\ell-1} \in \vec{\mathcal{S}}_{\ell-1}$, $\vec{S}_\ell \in \vec{\mathcal{S}}_\ell$ and $\vec{S}_\ell \preceq \vec{S}_{\ell-1}$.

Construct a sequence of sets $\mathbb{W}_j \subseteq \mathbb{T}_{\mathrm{tang}}[S_j]$ for $0 \le j \le \ell$ recursively as follows:

- Set $\mathbb{W}_\ell := \mathbb{T}_{\mathrm{tang}}[S_\ell]$.
- Assuming $\mathbb{W}_t, \dots, \mathbb{W}_\ell$ have already been constructed for some $1 \le t \le \ell$, define
$$\mathbb{W}_{t-1} := \mathbb{T}_{\mathrm{tang}}[S_{t-1}] \cap \big(\uparrow\uparrow_t \mathbb{W}_t\big)^*.$$

For each $0 \le t \le \ell$, the tubes belonging to $\mathbb{W}_t$ satisfy a version of the hypothesis from Definition 3.6. In particular, each $(\theta_t, v_t) \in \mathbb{W}_t$ satisfies the following:

Nested tube hypothesis. There exist $(\theta_i, v_i) \in \mathbb{T}[r_i]$ for $t+1 \le i \le \ell$ such that

i) $\mathrm{dist}(\theta_i, \theta_j) \lesssim r_j^{-1/2}$,
ii) $\mathrm{dist}\big(T_{\theta_j,v_j}(y_j),\, T_{\theta_i,v_i}(y_i) \cap B(y_j, r_j)\big) \lesssim r_i^{1/2+\delta}$,
iii) $T_{\theta_j,v_j}(y_j) \subset N_{r_j^{1/2+\delta_j}} S_j$

hold for all $t \le i \le j \le \ell$.

Furthermore, for each $0 \le t \le \ell$ the inequality

$$\|f_{\vec{S}_\ell}\|_2^2 \lesssim \prod_{j=t+1}^{\ell} \Big(\frac{r_{j-1}}{r_j}\Big)^{-\frac{j-1}{2}} D_j^\delta \|f_{\vec{S}_t}|_{\mathbb{W}_t}\|_2^2 \tag{6.9}$$

holds up to the inclusion of a rapidly decaying error term. Indeed, by (6.8) above,

$$\|f_{\vec{S}_j}|_{\mathbb{W}_j}\|_2^2 \lesssim_\varepsilon \Big(\frac{r_{j-1}}{r_j}\Big)^{-\frac{j-1}{2}} D_j^\delta R^{O(\varepsilon_\circ)} \big\|f_{\vec{S}_{j-1}}|_{\uparrow_j \mathbb{W}_j}\big\|_2^2.$$

Since $f_{\vec{S}_{j-1}}$ is concentrated on wave packets belonging to $\mathbb{T}_{\mathrm{tang}}[S_{j-1}]$, one has

$$f_{\vec{S}_{j-1}} = f_{\vec{S}_{j-1}}|_{\mathbb{T}_{\mathrm{tang}}[S_{j-1}]} + \mathrm{RapDec}(r)\|f\|_2.$$

Combining the two preceding displays together with Lemma 5.6, one deduces that

$$\|f_{\vec{S}_j}|_{\mathbb{W}_j}\|_2^2 \lesssim_\varepsilon \Big(\frac{r_{j-1}}{r_j}\Big)^{-\frac{j-1}{2}} D_j^\delta R^{O(\varepsilon_\circ)} \big\|f_{\vec{S}_{j-1}}|_{\mathbb{W}_{j-1}}\big\|_2^2$$

and this inequality may be applied recursively to deduce (6.9).

To conclude the proof, simply define $f^{\#}_{\vec{S}_\ell} := f|_{\mathbb{W}_0}$, noting that the desired properties then immediately follow from the preceding discussion. □

Appendix A. Deriving the asymptotic for the linear exponents

Here the derivation of the λ coefficient featured in Theorem 1.1 is described. The first author thanks Keith M. Rogers for the following argument. Begin by noting that

$$\frac{2i+1}{2(i+1)+1} \geq \frac{2i}{2i+1}\frac{2(i+1)}{2(i+1)+1} \geq \frac{i}{i+1}, \tag{A.1}$$

so that, by using the lower bound and telescoping,

$$\Big(\prod_{i=k}^{n-1}\frac{2i}{2i+1}\Big)^2 = \frac{2k}{2k+1}\frac{2n+1}{2n}\prod_{i=k}^{n-1}\frac{2i}{2i+1}\frac{2(i+1)}{2(i+1)+1} \geq \frac{2n+1}{2k+1}\frac{k^2}{n^2}.$$

Taking the square root and plugging this into the definition of $p_n(k)$, one obtains

$$p_n(k) \leq 2 + \frac{6}{2(n-1)+(k-1)(\frac{2n+1}{2k+1})^{1/2}\frac{k}{n}}.$$

The analogous argument, using the upper bound from (A.1), yields

$$p_n(k) \geq 2 + \frac{6}{2(n-1)+(k-1)(\frac{k}{n})^{1/2}}.$$

Taking $k = \nu n + O(1)$ for some $0 < \nu < 1$, it follows that, asymptotically,

$$p_n(k) = 2 + \frac{6}{2+\nu^{3/2}}n^{-1} + O(n^{-2}). \tag{A.2}$$

On the other hand, for $k = \nu n + O(1)$, the constraint

$$p \geq 2 + \frac{4}{2n-k}$$

coming from the Bourgain–Guth argument (c.f. (1.3)) can be rewritten as

$$p \geq 2 + \frac{4}{2-\nu}n^{-1} + O(n^{-2}). \tag{A.3}$$

Optimal choice of ν corresponds to the value at which linear coefficients in (A.2) and (A.3) are equal. This occurs when $\nu^{1/2}$ solves the cubic equation

$$2x^3 + 3x^2 - 2 = 0;$$

the derivation of this condition is presented in the appendix. Cardano's formula shows that the unique real root of this equation is given by the irrational number[5]

$$\nu^{1/2} = \Big(\frac{3}{8}+\frac{1}{8^{1/2}}\Big)^{1/3} + \Big(\frac{3}{8}-\frac{1}{8^{1/2}}\Big)^{1/3} - \frac{1}{2} = 0.67765....$$

Plugging this back into (1.3) yields (R^*_p) in the range

$$p > 2 + \lambda n^{-1} + O(n^{-2})$$

with $\lambda = \frac{4}{2-\nu} = 2.59607....$

[5]One can immediately see that the root must be irrational by applying Eisenstein's criterion to the shifted polynomial $2(x+1)^3 + 3(x+1)^2 - 2$.

References

[1] 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

[2] J. Bourgain, *Besicovitch type maximal operators and applications to Fourier analysis*, Geom. Funct. Anal. **1** (1991), no. 2, 147–187. MR 1097257

[3] J. Bourgain and L. Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, Geom. Funct. Anal. **21** (2011), no. 6, 1239–1295. MR 2860188

[4] A. Carbery, *Restriction implies Bochner-Riesz for paraboloids*, Math. Proc. Cambridge Philos. Soc. **111** (1992), no. 3, 525–529. MR 1151328

[5] C. Demeter, *On the restriction theorem for paraboloid in* $\mathbb{R}^4$, Preprint: arXiv:1701.03523.

[6] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta Math. **124** (1970), 9–36. MR 0257819

[7] L. Guth, *A restriction estimate using polynomial partitioning*, J. Amer. Math. Soc. **29** (2016), no. 2, 371–413. MR 3454378

[8] L. Guth and N.H. Katz, *On the Erdős distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), no. 1, 155–190. MR 3272924

[9] L. Guth and J. Zahl, *Polynomial Wolff axioms and Kakeya-type estimates in* $\mathbb{R}^4$, Preprint: arXiv:1701.07045.

[10] Larry Guth, *Restriction estimates using polynomial partitioning II*, Acta Math. **221** (2018), no. 1, 81–142. MR 3877019

[11] Larry Guth, Jonathan Hickman, and Marina Iliopoulou, *Sharp estimates for oscillatory integral operators via polynomial partitioning*, Acta Math. **223** (2019), no. 2, 251–376. MR 4047925

[12] Jonathan Hickman and Keith M. Rogers, *Improved Fourier restriction estimates in higher dimensions*, Camb. J. Math. **7** (2019), no. 3, 219–282. MR 4010062

[13] Jonathan Hickman, Keith M. Rogers, and Ruixiang Zhang, *Improved bounds for the Kakeya maximal conjecture in higher dimensions*, Preprint: arXiv:1908.05589.

[14] Nets Hawk Katz and Keith M. Rogers, *On the polynomial Wolff axioms*, Geom. Funct. Anal. **28** (2018), no. 6, 1706–1716. MR 3881832

[15] Maplesoft, a division of Waterloo Maple Inc., *Maple 17*.

[16] E.M. Stein, *Some problems in harmonic analysis*, Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979, pp. 3–20. MR 545235

[17] T. Tao, *The Bochner-Riesz conjecture implies the restriction conjecture*, Duke Math. J. **96** (1999), no. 2, 363–375. MR 1666558

[18] ______, *A sharp bilinear restrictions estimate for paraboloids*, Geom. Funct. Anal. **13** (2003), no. 6, 1359–1384. MR 2033842

[19] Terence Tao, *Some recent progress on the restriction conjecture*, Fourier analysis and convexity, Appl. Numer. Harmon. Anal., Birkhäuser Boston, Boston, MA, 2004, pp. 217–243. MR 2087245

[20] P.A. Tomas, *A restriction theorem for the Fourier transform*, Bull. Amer. Math. Soc. **81** (1975), 477–478. MR 0358216

[21] H. Wang, *A restriction estimate in* $\mathbb{R}^3$ *using brooms*, Preprint: arXiv:1802.04312.

[22] Thomas Wolff, *Recent work connected with the Kakeya problem*, Prospects in mathematics (Princeton, NJ, 1996), Amer. Math. Soc., Providence, RI, 1999, pp. 129–162. MR 1660476

[23] R. Wongkew, *Volumes of tubular neighbourhoods of real algebraic varieties*, Pacific J. Math. **159** (1993), no. 1, 177–184. MR 1211391

[24] J. Zahl, *A discretized Severi-type theorem with applications to harmonic analysis*, Geom. Funct. Anal. **28** (2018), no. 4, 1131–1181. MR 3820441

[25] Joshua Zahl, *New Kakeya estimates using Gromov's algebraic lemma*, Preprint: arXiv:1908.05314.

School of Mathematics, The University of Edinburgh, Edinburgh, EH9 3JZ, UK
Email address: `jonathan.hickman@ed.ac.uk`

Department of Mathematics, University of British Columbia, Vancouver BC, V6T 1Z2, Canada
Email address: `jzahl@math.ubc.ca`