A dichotomy for Hörmander-type oscillatory integral operators

Shaoming Guo[1,2] · Hong Wang[3,4] · Ruixiang Zhang[5]

Abstract

In this paper, we first generalize the work of Bourgain (Geom. Funct. Anal. 1(4):321–374, 1991) and state a curvature condition for Hörmander-type oscillatory integral operators, which we call Bourgain's condition. This condition is notably satisfied by the phase functions for the Fourier restriction problem and the Bochner-Riesz problem. We conjecture that for Hörmander-type oscillatory integral operators satisfying Bourgain's condition, they satisfy the same L^p bounds as in the Fourier Restriction Conjecture. To support our conjecture, we show that whenever Bourgain's condition fails, then the $L^\infty \to L^q$ boundedness always fails for some $q = q(n) > \frac{2n}{n-1}$, extending Bourgain's three-dimensional result (Geom. Funct. Anal. 1(4):321–374, 1991). On the other hand, if Bourgain's condition holds, then we prove L^p bounds for Hörmander-type oscillatory integral operators for a range of p that extends the currently best-known range for the Fourier restriction conjecture in high dimensions, given by Hickman and Zahl (A note on Fourier restriction and nested polynomial wolff axioms, 2020, arXiv:2010.02251). This gives new progress on the Fourier restriction problem, the Bochner-Riesz problem on $\mathbb{R}^n$, the Bochner-Riesz problem on spheres S^n, etc.

✉ S. Guo
shaomingguo2018@gmail.com

H. Wang
hongwang@math.ucla.edu

R. Zhang
ruixiang@berkeley.edu

1 Chern Institute of Mathematics, Nankai University, Tianjin, China

2 University of Wisconsin-Madison, Madison, WI, USA

3 NYU, New York, USA

4 UCLA, Los Angeles, USA

5 UC Berkeley, Berkeley, CA, USA

1 Introduction

Let us recall Hörmander's problem in [15]. Let $n \geq 2$. Let B^n be the unit ball in $\mathbb{R}^n$ and $a : \mathbb{R}^n \times \mathbb{R}^{n-1} \mapsto \mathbb{R}$ be a smooth function supported on $B^n \times B^{n-1}$. Let $\phi : B^n \times B^{n-1} \to \mathbb{R}$ be a smooth function satisfying the following conditions:

(H1) rank $\nabla_{\mathbf{x}} \nabla_\xi \phi(\mathbf{x}; \xi) = n - 1$ for all $(\mathbf{x}; \xi) \in B^n \times B^{n-1}$;
(H2) with the map $G_0 : B^n \times B^{n-1} \to \mathbb{R}^n$ defined by

$$G_0(\mathbf{x}; \xi) := \bigwedge_{j=1}^{n-1} \partial_{\xi_j} \nabla_{\mathbf{x}} \phi(\mathbf{x}; \xi), \tag{1.1}$$

the curvature condition

$$\det \nabla_\xi^2 \langle \nabla_{\mathbf{x}} \phi(\mathbf{x}; \xi), G_0(\mathbf{x}; \xi_0) \rangle \big|_{\xi = \xi_0} \neq 0 \tag{1.2}$$

holds for all $(\mathbf{x}; \xi_0) \in \mathrm{supp}(a)$.

Define the oscillatory integral operator

$$T_N f(\mathbf{x}) := \int e^{iN\phi(\mathbf{x};\xi)} f(\xi) a(\mathbf{x}; \xi) d\xi. \tag{1.3}$$

If one takes $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$ where $\mathbf{x} = (x, t)$, then one can check easily that Hypothesis (H1) and (H2) are satisfied, and T_N becomes the standard Fourier extension operator for the paraboloid. Hörmander [15] asked the question whether T_N satisfies similar L^p-boundedness properties to those of the Fourier extension operator. To be more precise, he asked whether it holds that

$$\|T_N f\|_q \lesssim_{n,q} N^{-n/q} \|f\|_\infty, \tag{1.4}$$

for all $q > \frac{2n}{n-1}$.

In dimension $n = 2$, the answer to Hörmander's question is affirmative, see for instance Carleson and Sjölin [5], Hörmander [15] and Fefferman [6]. However, in dimension $n = 3$, Bourgain [3] showed that if one takes

$$\phi(\mathbf{x}; \xi) = x_1 \xi_1 + x_2 \xi_2 + t \xi_1 \xi_2 + \frac{1}{2} t^2 \xi_1^2, \quad \mathbf{x} = (x_1, x_2, t), \tag{1.5}$$

then (1.4) may fail for every $q < 4$. Indeed, he showed that even if one replaces (H2) by the following stronger assumption

$$\text{(H2}^+\text{)} \quad \nabla_\xi^2 \langle \nabla_{\mathbf{x}} \phi(\mathbf{x}; \xi), G_0(\mathbf{x}; \xi_0) \rangle \big|_{\xi = \xi_0} \text{ is positive definite,} \tag{1.6}$$

the estimate (1.4) may still fail for some $q > 3$, and it may even fail generically. Let us be more precise about this generic failure. By some elementary change of variables, phase functions $\phi(\mathbf{x}; \xi)$ satisfying (H1) and (H2) can be taken to be

$$\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t \langle A\xi, \xi \rangle + O(|t||\xi|^3 + |\mathbf{x}|^2 |\xi|^2), \tag{1.7}$$

where A is a symmetric non-degenerate matrix. Condition (H2$^+$) amounts to that A is positive definite. The form (1.7) is called a normal form of the phase function $\phi(\mathbf{x}; \xi)$ at the origin (see [3, page 323]). Bourgain [3] proved in dimension $n = 3$ that, if

$$\nabla_\xi^2 \partial_t^2 \phi \big|_{\mathbf{x}=0, \xi=0} \text{ is not a multiple of } \nabla_\xi^2 \partial_t \phi \big|_{\mathbf{x}=0, \xi=0} \tag{1.8}$$

where $\phi(\mathbf{x}; \xi)$ is as in (1.7), then (1.4) fails for every $q < \frac{118}{39}$, which is > 3.

On the other hand, for phase functions $\phi(\mathbf{x}; \xi)$ satisfying (H1) and (H2$^+$), Guth, Hickman and Iliopoulou [10] (basing on earlier work Guth [9]) obtained the *optimal* range of q for which (1.4) holds. Denote

$$q_{n,\text{GHI}} := \begin{cases} \frac{2(3n+1)}{3n-3}, & \text{if } n \text{ is odd,} \\ \frac{2(3n+2)}{3n-2}, & \text{if } n \text{ is even.} \end{cases} \tag{1.9}$$

Then (1.4) holds for all $q > q_{n,\text{GHI}}$.

In the first result of the paper, we show that generic failure in the spirit of Bourgain [3] occurs in every dimension $n \geq 3$. Let us first introduce some terminology. At a given point $(\mathbf{x}_0; \xi_0)$, consider a new phase function $\phi'(\mathbf{x}; \xi) := \phi(\mathbf{x}_0 + \mathbf{x}; \xi_0 + \xi)$. Let us use $\phi''(\mathbf{x}; \xi)$ to denote a normal form of $\phi'(\mathbf{x}; \xi)$ at the origin $\mathbf{x} = 0, \xi = 0$. We say that Bourgain's condition holds at $(\mathbf{x}_0; \xi_0)$ if

$$\nabla_\xi^2 \partial_t^2 \phi'' \big|_{\mathbf{x}=0, \xi=0} \text{ is a multiple of } \nabla_\xi^2 \partial_t \phi'' \big|_{\mathbf{x}=0, \xi=0}, \tag{1.10}$$

where the implicit constant is allowed to depend on $\mathbf{x}_0$ and ξ_0. Otherwise, we say that Bourgain's condition fails at this point. As normal forms are not unique, we need to show that Bourgain's condition is well-defined, and this is done in Corollary 2.2.

Theorem 1.1 (Generic failure) *Let $n \geq 3$ and*

$$q_{n,1} := \frac{2(2n^2 + n - 1)}{2n^2 - n - 2}. \tag{1.11}$$

Let $\phi : B^n \times B^{n-1} \to \mathbb{R}$ be a smooth function satisfying (H1) and (H2). If Bourgain's condition fails at some $(\mathbf{x}_0; \xi_0) \in \text{supp}(a)$, then (1.4) may fail for every $q < q_{n,1}$.

Note that when $n = 3$, $q_{n,1} = 40/13$, which is slightly better than Bourgain's exponent 118/39.

Based on the above theorem, we think it is very natural to conjecture that (1.4) holds for every $q > \frac{2n}{n-1}$ if Bourgain's condition holds at every point. The following positive results provide some further evidence for such a conjecture.

For $\delta > 0$, we define δ-tubes. Fix a dyadic cube $\theta \subset B^{n-1}$ of side length δ and let ξ_θ be the center of θ. For $v \in B^{n-1}$ with $v \in \delta\mathbb{Z}^{n-1}$, let $X_t(\xi_\theta, v) \in B^{n-1}$ denote the unique solution in the x variable to

$$\nabla_\xi \phi(x, t; \xi_\theta) = v. \tag{1.12}$$

By a δ-tube, we mean

$$T_{\theta,v} := \{(x,t) : |x - X_t(\xi_\theta, v)| \leq \delta, |t| \leq 1\}. \tag{1.13}$$

For a collection $\mathbb{T}$ of tubes $\{T_{\theta,v}\}$, we say that the tubes in $\mathbb{T}$ point in different directions if for two different tubes T_{θ_1,v_1} and T_{θ_2,v_2} from $\mathbb{T}$, we always have $\theta_1 \neq \theta_2$.

Theorem 1.2 (Polynomial Wolff Axiom for ϕ) *Let $n \geq 3$. If Bourgain's condition holds for the phase function ϕ at every $(\mathbf{x}_0; \xi_0) \in \operatorname{supp}(a)$, then the following polynomial Wolff axiom for ϕ holds: Let $E \geq 2$ be an integer. For every $\epsilon > 0$, there exists $C(n, E, \epsilon) > 0$ such that for every collection $\mathbb{T}$ of δ-tubes pointing in different directions,*

$$\#\{T \in \mathbb{T} : T \subset S\} \leq C(n, E, \epsilon)|S|\delta^{1-n-\epsilon} \tag{1.14}$$

whenever $S \subset B^n$ is a semialgebraic set of complexity $\leq E$.

The above polynomial Wolff axiom for ϕ satisfying Bourgain's condition is a generalization of that for $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$, proven by Katz and Rogers [11]. Hickman and Rogers [12] used the polynomial Wolff axiom of Katz and Rogers and proved that (1.4) holds with $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$ for

$$q > q_{n,\mathrm{HR}} := 2 + \frac{\lambda_{\mathrm{HR}}}{n} + O(n^{-2}), \tag{1.15}$$

where

$$\lambda_{\mathrm{HR}} = 4/(5 - 2\sqrt{3}) = 2.60434\ldots \tag{1.16}$$

After verifying the polynomial Wolff axiom for general ϕ satisfying Bourgain's condition, one can expect to combine the argument of [10] and [12], and prove (1.4) for all q satisfying (1.15). We will indeed prove something stronger. Before stating the next theorem, we first recall the result of Hickman and Zahl [13]. Let $\nu^{1/2}$ be the real number that solves the equation

$$2x^3 + 3x^2 - 2 = 0. \tag{1.17}$$

Denote

$$\lambda_{\mathrm{HZ}} := \frac{4}{2 - \nu} = 2.59607\ldots \tag{1.18}$$

Hickman and Zahl [13] used the strong polynomial Wolff axiom by Hickman-Rogers-Zhang [14] and independently Zahl [24] to further improve the result in [12] and obtained that (1.4) holds for

$$q > q_{n,\mathrm{HZ}} := 2 + \frac{\lambda_{\mathrm{HZ}}}{n} + O(n^{-2}), \tag{1.19}$$

with $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$. This result gives the best asymptotic formula (as $n \to \infty$) in the literature for the Fourier restriction conjecture.

Theorem 1.3 *Let* $\phi : B^n \times B^{n-1} \to \mathbb{R}$ *be a smooth function satisfying* (H1) *and* (H2$^+$). *If Bourgain's condition holds for the phase function* $\phi(\mathbf{x}; \xi)$ *at every point* $(\mathbf{x}; \xi) \in \mathrm{supp}(a)$, *then* (1.4) *holds for*

$$q > q_{n,2} := 2 + \frac{2.5921}{n} + O(n^{-2}). \tag{1.20}$$

Note that

$$q_{n,1} = 2 + \frac{4n + 2}{2n^2 - n - 2} = 2 + \frac{2}{n} + O(n^{-2}), \tag{1.21}$$

and therefore for large n, we have the order

$$q_{n,1} < q_{n,2} < q_{n,\mathrm{HZ}} < q_{n,\mathrm{HR}} < q_{n,\mathrm{GHI}}. \tag{1.22}$$

When $n = 3$, the exponents in (1.22) are given by

$$3 + \frac{1}{13} < 3 + \frac{3}{13} < 3 + \frac{1}{4} \leq 3 + \frac{1}{4} < 3 + \frac{1}{3}, \tag{1.23}$$

respectively. Here $q_{n,2}$ with $n = 3$ matches Wang's exponent [21].

As $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$ also satisfies Bourgain's condition, we obtain the following immediate corollary of Theorem 1.3.

Corollary 1.4 (Improved Fourier restriction estimate) *For* $q > q_{n,2}$, *it holds that*

$$\left\| \int_{[0,1]^{n-1}} e^{iN(\langle x, \xi \rangle + t|\xi|^2)} f(\xi) d\xi \right\|_q \lesssim_{n,q} N^{-n/q} \|f\|_\infty. \tag{1.24}$$

Recall Bourgain's observation [3, Remark 3.43] that the phase function for the Bochner-Riesz problem also satisfies Bourgain's condition, we see Theorem 1.3 also gives the currently best known bounds for the Bochner-Riesz problem. More precisely, for $\alpha \geqslant 0$, the Bochner-Riesz multiplier of order α is defined by

$$m^\alpha(\xi) := \left(1 - |\xi|^2 \right)_+^\alpha, \quad \xi \in \mathbb{R}^n.$$

As a corollary of Theorem 1.3 (also Theorem 4.1 below), we obtain

Corollary 1.5 (Improved Bochner-Riesz estimate on $\mathbb{R}^n$) *For* $q > q_{n,2}$, *it holds that*

$$\left\| m^\alpha(D) f \right\|_{L^q(\mathbb{R}^n)} \lesssim_{n,\alpha,q} \|f\|_{L^q(\mathbb{R}^n)}, \tag{1.25}$$

whenever $\alpha > n(\frac{1}{2} - \frac{1}{p}) - \frac{1}{2}$.

Before stating the next corollary, let us discuss a prior attempt in breaking the critical range of exponents in (1.9) for general Hörmander operators. In [7], Gao, Li and Wang, under the assumptions (H1), (H2)$^+$, and the additional assumption

$$G_0(\mathbf{x}; \xi)/|G_0(\mathbf{x}; \xi)| \text{ is constant in } \mathbf{x}, \tag{1.26}$$

proved that (1.4) holds for q satisfying the range of Hickman and Zahl (1.19).

By Theorem 2.1 and Lemma 2.3, one can check directly that phase functions satisfying (1.26) also satisfy Bourgain's condition. Therefore the bounds Gao, Li and Wang obtained in [7] can also be improved to the one stated in Theorem 1.3.

The assumption (1.26) appears naturally in several interesting applications, including the generalized Bochner-Riesz problem for non-degenerate hyper-surfaces, the local smoothing estimates for fractional Schrödinger equations and sharp resolvent estimates outside of the uniform boundedness range. These applications were worked out carefully in [7]. We include the latter two here.

Let $u : \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{C}$ be the solution to the equation

$$\begin{cases} i\partial_t u + (-\Delta)^{\frac{\alpha}{2}} u = 0, & (x,t) \in \mathbb{R}^{n-1} \times R, \\ u(x,0) = f(x), & x \in \mathbb{R}^{n-1}, \end{cases} \tag{1.27}$$

where $\alpha > 1$ and f is a Schwartz function.

Corollary 1.6 (Local smoothing estimates for fractional Schrödinger equations) *Let $\alpha > 1$ and let u be a solution to* (1.27). *Then*

$$\|u\|_{L^q(\mathbb{R}^{n-1} \times [1,2])} \lesssim_{\alpha,\beta,n,\epsilon,p} \|f\|_{L^q_\beta(\mathbb{R}^{n-1})}, \tag{1.28}$$

whenever

$$\beta > (n-1)\alpha\left(\frac{1}{2} - \frac{1}{q}\right) - \frac{\alpha}{q}, \tag{1.29}$$

and q satisfies (1.20). *Moreover, for each fixed q, the range of β is sharp.*

The conjectured range for (1.28) to hold is $q > \frac{2n}{n-1}$, the same as the range in the Fourier restriction conjecture.

The resolvent estimate for the Laplacian on $\mathbb{R}^n$ is of the form

$$\left\|(-\Delta - z)^{-1} f\right\|_{L^q(\mathbb{R}^n)} \leq C_{p,q,n}(z)\|f\|_{L^p(\mathbb{R}^n)}, \quad z \in \mathbb{C} \setminus [0, \infty). \tag{1.30}$$

Here $C_{p,q,n}(z)$ is a constant that is allowed to depend on p, q, n and z. We are particularly interested in tracking the dependence on z, for fixed n, p and q.

Corollary 1.7 (Resolvent estimates) *For q satisfying* (1.20), *we have*

$$\left\|(-\Delta - z)^{-1} f\right\|_{L^q(\mathbb{R}^n)} \lesssim_{q,n} |z|^{-1+\gamma_q} \operatorname{dist}(z, [0, \infty))^{-\gamma_q} \|f\|_{L^q(\mathbb{R}^n)}, \tag{1.31}$$

where

$$\gamma_q := \frac{n+1}{2} - \frac{n}{q}. \tag{1.32}$$

Moreover, for fixed q, the bound is optimal in z.

The study of resolvent estimates in (1.30) has a long history and can be dated back to the work of Kenig, Ruiz and Sogge [17]. The authors there proved (1.30) for optimal ranges of (p, q) for which the constant $C_{p,q,n}(z)$ is independent of z. These are called uniform Sobolev inequalities, and have found numerous applications including unique continuation properties, limiting absorption principles, etc.

The conjectured range for (1.31) to hold is $q > \frac{2n}{n-1}$, the same as the Fourier restriction exponent. Moreover, if this conjecture turns out to be true, then by interpolation with known results, it would imply (1.30) for all combinations of (p, q), with an optimal dependence of $C_{p,q,n}(z)$ on z.

In the last corollary, we discuss the connection of Theorem 1.3 to Sogge's work [18] and [20]. In [20], Sogge studied the Nikodym problem on general Riemannian manifolds and proved that the Nikodym maximal operator satisfies better bounds on manifolds of constant scalar curvature than on general Riemannian manifolds. This is a strong indication that distance functions on manifolds of constant curvature satisfy Bourgain's condition. In the next corollary, we show that this is indeed the case for S^n, the n-dimensional Euclidean sphere.

Corollary 1.8 *Let* $a : S^n \times S^n \to \mathbb{R}$ *be a smooth function supported away from the diagonal. Let* dist *be the distance function on* S^n. *Then*

$$\left\| \int_{S^n} e^{iN \operatorname{dist}(x, y)} a(x, y) f(y) dy \right\|_{L^q(S^n)} \lesssim N^{-n/q} \| f \|_{L^q(S^n)}, \qquad (1.33)$$

for every q satisfying (1.20).

Proof of Corollary 1.8 By cutting the support of a into finitely many pieces, we without loss of generality assume that the support of $a(x, y)$ is such that x is around the north pole and y is slightly away from the north pole. Write

$$x = (x_1, \ldots, x_n, \sqrt{1 - |x'|^2}), \quad y = (y_1, \ldots, y_n, \sqrt{1 - |y'|^2}), \qquad (1.34)$$

where $x' = (x_1, \ldots, x_n)$, $y' = (y_1, \ldots, y_n)$. Note that $\operatorname{dist}(x, y) = \arccos(x \cdot y)$. When integrating in y on the left hand side of (1.33), we apply Fubini's theorem and integrate in y with $\sqrt{1 - |y'|^2} = r$ for some r and then integrate in r. For a fixed r, our distance function can be written as

$$\arccos(x_1 y_1 + \cdots + x_{n-1} y_{n-1} + x_n \sqrt{1 - r^2 - |y''|^2} + r\sqrt{1 - |x'|^2}), \qquad (1.35)$$

where $y'' := (y_1, \ldots, y_{n-1})$. Therefore, to prove Corollary 1.8, it suffices to show that (1.35) satisfies Bourgain's condition in x' and y'' variables. We will prove this by checking the definition of Bourgain's condition as in (1.10). By rotation symmetry, it suffices to consider the phase function (1.35) near the north pole $x' = 0$ and $y'' = 0$. Next, we apply a Taylor expansion for (1.35) about $x' = 0$ and $y'' = 0$. After the Taylor expansion, note that all linear terms in y'' can be written as

$$(x_1 + O(|x'|^2)) y_1 + \cdots + (x_{n-1} + O(|x'|^2)) y_{n-1}. \qquad (1.36)$$

We therefore apply the change of variables

$$x_1 + O(|x'|^2) \mapsto x_1, \ldots, x_{n-1} + O(|x'|^2) \mapsto x_{n-1}, \tag{1.37}$$

to turn our phase function into a normal form. Let ϕ'' denote this normal form. In the end, we just need to note that both matrices

$$\nabla^2_{y''} \partial^2_t \phi''|_{x'=0, y''=0}, \quad \nabla^2_{y''} \partial_t \phi''|_{x'=0, y''=0} \tag{1.38}$$

are multiples of the identity matrix. This finishes the proof. $\qquad\qquad\qquad\square$

It is conjectured (see Sogge [18]) that (1.33) holds for all $q > 2n/(n-1)$. The operator studied in (1.33) appears in the study of the Bochner-Riesz problem on Euclidean spheres S^n. Let Δ_g denote the Laplace-Beltrami operator, and

$$0 < \lambda_1 \le \lambda_2 \le \ldots \tag{1.39}$$

the eigenvalues of $-\Delta_g$. Let E_j be the one-dimensional eigenspace for $-\Delta_g$ with eigenvalue λ_j, and $e_j : L^2(S^n) \to L^2(S^n)$ be the projection operator onto the eigenspace E_j. We define the Riesz means of index $\alpha \ge 0$ as

$$S^\alpha_L(f) := \sum_{j=1}^{\infty} \left(1 - \frac{\lambda_j}{L}\right)^\alpha_+ e_j(f). \tag{1.40}$$

One can follow the work Sogge [18] and Huang and Sogge [16], and deduce the following corollary from Corollary 1.8.

Corollary 1.9 (Bochner-Riesz for spheres) *Assume that q satisfies* (1.20). *We have that*

$$\left\| S^\alpha_L(f) \right\|_{L^q(S^n)} \lesssim \left\| f \right\|_{L^q(S^n)} \text{ uniformly in } L, \tag{1.41}$$

whenever $\alpha > n(\frac{1}{2} - \frac{1}{p}) - \frac{1}{2}$. Moreover, the range of α is sharp for fixed q.

At the end of the introducion, we discuss the proof of Theorem 1.3. To prove Theorem 1.3, we first prove a strong Polynomial Wolff Axiom (SPWA) for the phase function ϕ satisfying Bourgain's condition, see Sect. 6. It is a nested version of the Polynomial Wolff Axiom and generalizes the SPWA by Hickman and Zahl [13], which was built on the work of Hickman-Rogers-Zhang [14] and Zahl [24]. One can combine this strong polynomial Wolff axiom with the argument in [13], and prove (1.4) for q satisfying (1.19).

To improve the range in [13], we further develop the idea of "brooms" in dimension $n = 3$ in Wang [21] for the Fourier restriction problem. In [21], the second author introduced a notion of *brooms*. They enable one to exploit the feature that if the sum of a collection of wave packets is highly concentrated locally in space, then this collection must spread out on the far end, leading to new improvements on the range of exponent for the Fourier restriction conjecture in $\mathbb{R}^3$.

However, a key geometric argument in [21] regarding brooms relies heavily on the space being three-dimensional. Even in the Fourier restriction setting, it was not clear how one can most efficiently generalize the notion and construction of brooms in [21] to higher dimensions. In the current paper, we come up with a slightly different notion of brooms, which works in all dimensions and also in the setting of oscillatory integral operators satisfying Hörmander's conditions. This is done in Sect. 7.1. When proving the relevant broom estimates (see Theorem 7.8 in Sect. 7.3), we use an argument that can be viewed as a generalized pseudo-conformal transformation (Lemma 7.12). By this transformation and a counting lemma (Lemma 7.10 below) the key broom estimate is then reduced to what we call a *Variety Uncertainty Principle*. We state it in the Fourier transform case below as we will use it to prove the general case.

Lemma 1.10 (Variety Uncertainty Principle) *Given two $(m-1)$-dimensional algebraic varieties Y_1, Y_2 in $\mathbb{R}^{n-1}$ that are transverse complete intersections (see Definition 5.1 below). Let $Z_i \subset Y_i$ be the part of Y_i where every point is non-singular and the angle formed by $T_{\mathbf{z}_i}(Z_i)$ and the space spanned by $\{\vec{e}_1, \ldots, \vec{e}_{m-1}\}$ is $\leq 1/(100n)$, for every $i = 1, 2$ and every $\mathbf{z}_i \in Z_i$. Here $T_{\mathbf{z}_i}(Z_i)$ refers to the tangent space and $\vec{e}_j$ refers to a coordinate vector. Let $1 \leq R_1 \leq R_2$. Denote[1]*

$$\Omega_1 = \mathcal{N}_{\sqrt{R_1}}(Z_1), \quad \Omega_2 = \mathcal{N}_{1/\sqrt{R_2}}(Z_2). \tag{1.42}$$

Assume that $F : \mathbb{R}^{n-1} \to \mathbb{C}$ satisfies $\operatorname{supp}(F) \subset \Omega_2$. Then

$$\left\| \widehat{F} \right\|_{L^2(\Omega_1)}^2 \lesssim \left(\frac{R_1}{R_2} \right)^{\frac{n-m}{2} - \delta} \| F \|_{L^2}^2, \tag{1.43}$$

for every $\delta > 0$, where the implicit constant depends on $n, m, \deg(Z_1), \deg(Z_2)$ and δ.

Lemma 1.10 may also be of independent interest as it can be viewed as a variant of the generalized Mizohata-Takeuchi Conjecture by Jonathan Bennett and Tony Carbery, as stated in (9) in [2]. For the original Mizohata-Takeuchi Conjecture and its influences, see e.g. the references in [2]. As one sees from the proof, Lemma 1.10 is proved by using the geometric information of neighborhoods of both the spatial set and the hypersurface in the Fourier space at many scales, and can be viewed as a result in the vein of Mizohata-Takeuchi but with much stronger assumptions about the neighborhoods of the underlying sets at many scales.

Structure of the paper. In Sect. 2 we first give an equivalent characterization of Bourgain's condition, which is more straightforward to check, and then prove Theorem 1.1. In dimension $n = 3$, the improvement of our result over Bourgain's [3] comes from a slightly more efficient way of constructing (curved) tubes that have high overlapping.

In Sect. 3, we show that for phase functions $\phi(\mathbf{x}; \xi)$ satisfying Bourgain's condition, the corresponding tubes satisfy the polynomial Wolff axiom. We follow largely the argument of Katz and Rogers [11].

---

[1]Here $\mathcal{N}$ means neighborhood in $\mathbb{R}^{n-1}$; in this lemma there is no $\mathbb{R}^n$.

In Sect. 4, we introduce the standard wave packet decomposition and standard reduction of Theorem 1.3 to a broad norm estimate (Theorem 4.2).

In Sect. 5 we apply a polynomial partitioning algorithm to decompose the broad norm in Theorem 4.2. The algorithm is a slight variant of that in Hickman and Rogers [12], with one difference that we need to have a better control of how fast cells shrink.

In Sect. 6, we prove the strong polynomial Wolff axiom (mentioned below Theorem 1.3) for phase functions satisfying Bourgain's condition.

In Sect. 7, we define brooms and prove the broom estimate (Theorem 7.8), which is key to the proof of Theorem 4.2. It is worth noting that the broom estimate holds for all phase functions $\phi(\mathbf{x}; \xi)$ satisfying (H1) and (H2$^+$), and does not rely on Bourgain's condition.

In Sect. 8, we define bushes and prove bush estimates. They are used to handle "small" grain resulting from the polynomial partitioning algorithm in Sect. 5.

In Sect. 9, we put all the ingredients together and finish the proof of Theorem 4.2, the broad norm estimate.

Notation. We use $\mathbf{x} = (x, t)$ to refer to a spatial points in $\mathbb{R}^n$, and ξ or ω for a frequency point in $\mathbb{R}^{n-1}$. Denote $\partial_i = \partial_{x_i}$ if $1 \leq i \leq n - 1$ and $\partial_n = \partial_t$.

We use d for the degree of polynomials that we will use in the polynomial partitioning lemmas. It is a large constant and is not allowed to depend on parameters like R or λ.

We will use a few admissible parameters

$$\epsilon^C \leqslant \delta \ll_\epsilon \delta_n \ll_\epsilon \delta_{n-1} \ll_\epsilon \cdots \ll_\epsilon \delta_1 \ll_\epsilon \delta_0 \ll_\epsilon \epsilon_\circ \ll_\epsilon \epsilon. \tag{1.44}$$

Here C is some dimensional constant and the notation $A \ll_\epsilon B$ indicates that $A \leqslant C_{n,\epsilon}^{-1} B$ for some large admissible constant $C_{n,\epsilon} \geqslant 1$. These parameters have exactly the same meaning as their counterparts $\delta, \delta_n, \ldots, \delta_0, \epsilon$ in Hickman-Rogers' work [12]. In the current paper, we need more of these parameters. For each $1 \leq n' \leq n$, let $\delta_{n'-1/2}$ be such that

$$\delta_{n'} \ll_\epsilon \delta_{n'-1/2} \ll_\epsilon \delta_{n'-1}. \tag{1.45}$$

These new parameters will be used when we modify the polynomial partitioning algorithm of [12].

For two positive constant A, B, by $A \lesssim B$ we mean $A \leq R^{O(\delta)} B$.

For a function $F : \mathbb{R}^n \to \mathbb{C}$ and a region $\Omega \subset \mathbb{R}^n$, we say that F is essentially supported on Ω if it decay rapidly outside Ω. In other words, for $\mathbf{x} \in \mathbb{R}^n \setminus \Omega$, and every $N \in \mathbb{N}$, it holds

$$|F(\mathbf{x})| \leq C_{n,N} R^{-N}, \tag{1.46}$$

for some constant $C_{n,N}$. Here $R > 1$ is as in Theorem 4.2.

2 Bourgain's condition and proof of Theorem 1.1

This section consists of two subsections. In Sect. 2.1, we will give an equivalent form of Bourgain's condition; in Sect. 2.2, we will prove Theorem 1.1.

Bourgain's formulation of Bourgain's condition in [3] requires that one first writes a phase function in its normal form, which differs from point to point. The equivalent form below (see Theorem 2.1) can be checked directly for a phase function, and can be more conveniently applied later when checking polynomial Wolff axioms (see Sect. 3).

The proof of Theorem 1.1 in Sect. 2.2 generalizes Bourgain's argument in [3] for $n = 3$. Bourgain discovered that if a phase function fails Bourgain's condition, then we always have "Kakeya compression" phenomenon. When reflected in the proof, this phenomenon means that the Jacobian determinant of the map $X_t(\xi)$ in Lemma 2.4 can be "small", which further means that the image of the map (the relevant Kakeya set) can also be "small". This ends up being how we generalize Bourgain's result to higher dimensions.

2.1 An equivalent formulation of Bourgain's condition

In this subsection, we will provide an equivalent formulation of Bourgain's condition. This equivalent formulation will be used in a few places below; for instance, it will play a crucial role in Sect. 3 and Sect. 6 when proving polynomial Wolff axioms for Hörmander's operators satisfying Bourgain's condition.

Define

$$\vec{T}_j(\mathbf{x}; \xi) = \partial_{\xi_j} \nabla_{\mathbf{x}} \phi(\mathbf{x}; \xi), \tag{2.1}$$

and

$$\vec{V}(\mathbf{x}; \xi) = \vec{T}_1(\mathbf{x}; \xi) \wedge \cdots \wedge \vec{T}_{n-1}(\mathbf{x}; \xi). \tag{2.2}$$

The main result in this subsection is

Theorem 2.1 *Let $\phi(\mathbf{x}; \xi)$ be a phase function that satisfies conditions* (H1) *and* (H2). *It satisfies Bourgain's condition at* $(\mathbf{x}_0; \xi_0)$ *if and only if*

$$((\vec{V} \cdot \nabla_{\mathbf{x}})^2 \nabla_\xi^2 \phi)(\mathbf{x}_0; \xi_0) \text{ is a multiple of } ((\vec{V} \cdot \nabla_{\mathbf{x}}) \nabla_\xi^2 \phi)(\mathbf{x}_0; \xi_0). \tag{2.3}$$

The constant is allowed to depend on $\mathbf{x}_0$ *and* ξ_0.

Corollary 2.2 *Let $\phi(\mathbf{x}; \xi)$ be a phase function satisfying conditions* (H1) *and* (H2). *That Bourgain's condition holds at* $(\mathbf{x}_0; \xi_0)$ *is independent of the choice of normal forms.*

Here by $\vec{V} \cdot \nabla_{\mathbf{x}}$, we mean

$$V_1(\mathbf{x}; \xi)\partial_1 + \cdots + V_{n-1}(\mathbf{x}; \xi)\partial_{n-1} + V_n(\mathbf{x}; \xi)\partial_n, \tag{2.4}$$

where $\vec{V} = (V_1, \ldots, V_n)^T$ and for the sake of simplicity we introduced the notation $\partial_i = \partial_{x_i}$ if $i \leq n - 1$ and $\partial_n = \partial_t$. It is perhaps worthy emphasizing that because of the dependence of $\vec{V}$ on $\mathbf{x}$, the differential operator $(\vec{V} \cdot \nabla_{\mathbf{x}})^2$ is equal to

$$\sum_{1 \leq i, j \leq n} V_i V_j \cdot \partial_i \partial_j + \sum_{1 \leq i \leq n} V_i \sum_{1 \leq j \leq n} \partial_i V_j \cdot \partial_j. \tag{2.5}$$

Proof of Theorem 2.1. We first observe that if (2.3) is satisfied everywhere, then it is also satisfied everywhere if one replaces V by $\lambda(\mathbf{x}; \xi)V$ where $\lambda(\mathbf{x}; \xi)$ is any smooth scalar function. This can be seen by the straightforward computation

$$(\lambda \vec{V} \cdot \nabla_{\mathbf{x}})^2 f = \lambda^2 (\vec{V} \cdot \nabla_{\mathbf{x}})^2 f + \lambda \cdot (\vec{V} \cdot \nabla_{\mathbf{x}} \lambda) \vec{V} \cdot \nabla_{\mathbf{x}} f \tag{2.6}$$

for every smooth function $f = f(\mathbf{x}; \xi)$.

Back to the proof of the Theorem. We first prove that if ϕ satisfies (2.3) everywhere, then it also satisfies (2.3) everywhere after any diffeomorphism in $\mathbf{x}$ only or in ξ only. Without loss of generality, we only need to verify this at $(\mathbf{0}, 0)$ and can assume the diffeomorphism always preserves the origin.

In order to show that (2.3) continues to hold after any diffeomorphism in $\mathbf{x}$, in light of the above property, we just need to show the "direction" of $\vec{V} \cdot \nabla_{\mathbf{x}}$ is invariant. More specifically, if h is a diffeomorphism that preserves the origin and write $h(\mathbf{x}) = \mathbf{y}$, we use h_* to denote the tangent map of h at the origin and only need to prove

$$h_*(\vec{T}_1 \wedge \cdots \wedge \vec{T}_{n-1}) \parallel \vec{S}_1 \wedge \cdots \wedge \vec{S}_{n-1} \tag{2.7}$$

where $\vec{T}_j$ is the original $\vec{T}_j(\mathbf{0}; 0) = \nabla_{\mathbf{x}} \partial_{\xi_j} \phi(\mathbf{0}; 0)$ as before and $\vec{S}_j = \nabla_{\mathbf{y}} \partial_{\xi_j} \phi(\mathbf{0}; 0)$ is defined similarly in the tangent space of $(\mathbf{0}; 0)$ in $\mathbf{y}$ coordinates.

Now everything can be computed in the tangent space of $(\mathbf{0}; 0)$ in terms of n-variate functions $\partial_{\xi_j} \phi(\cdot; 0)$ and we now check (2.7) using linear algebra. View all vectors $\vec{T}, \vec{S}$ as column vectors as before. Let J denote the Jacobian $\frac{\partial \mathbf{x}}{\partial \mathbf{y}}|_{\mathbf{0}}$ with $J_{ij} = \frac{\partial x_i}{\partial y_j}|_{\mathbf{0}}$. Then

$$\vec{S}_j = J^T \cdot \vec{T}_j, 1 \le j \le n - 1. \tag{2.8}$$

It is then easy to compute by considering the $(n-1)$-th tensor product of J acting on the $(n-1)$-fold wedge algebra that

$$\wedge_{j=1}^{n-1} \vec{S}_j = \det(J) \cdot (J^{-1}) \cdot (\wedge_{j=1}^{n-1} \vec{T}_j). \tag{2.9}$$

Finally, note that the matrix of h_* is $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}|_{\mathbf{0}} = J^{-1}$. By (2.9) we see both sides of (2.7) are parallel to $(J^{-1}) \cdot (\wedge_{j=1}^{n-1} \vec{T}_j)$ and thus (2.7) holds.

Next we show that (2.3) continues to hold after any diffeomorphism in ξ. First note that this property is preserved if we do any diffeomorphism in ξ. Indeed, this will multiply a nonzero scalar (equal to the determinant of the linear change of variable in ξ) to the whole vector field $\vec{V}$, and will result in a constant congruent transformation in $\nabla_\xi^2 \phi$ everywhere.

Hence, it suffices to show that (2.3) gets preserved if one does a change of variables of the following shape:

$$\xi_j \mapsto \xi_j + \xi^T A_j \xi + \text{ higher order terms} \tag{2.10}$$

where $A_j (1 \le j \le n-1)$ is a symmetric $(n-1) \times (n-1)$ matrix. Assume the Taylor expansion of ϕ near the origin is

$$\phi(\mathbf{x}; \xi) = a(\mathbf{x}) + b(\xi) + \sum_{i,j} c_{i,j} x_i \xi_j + \sum_i x_i \cdot \xi^T D_i \xi + \sum_j \xi_j \cdot \mathbf{x}^T E_j \mathbf{x}$$

$$+ \sum_{i_1,i_2,j_1,j_2} f_{i_1,i_2,j_1,j_2} x_{i_1} x_{i_2} \xi_{j_1} \xi_{j_2} + g(\mathbf{x}; \xi) \tag{2.11}$$

where to simplify notation we wrote $t = x_n$, the functions a, b are polynomials of degree ≤ 3 and g is a sum of terms of order ≥ 3 in $\mathbf{x}$ and terms of order ≥ 3 in ξ. These terms will not play roles in verifying (2.3) at the origin. Here D_i and E_j refer to square matrices.

We use $\tilde{\phi}$ to denote the new expression of ϕ under the change of variable (2.10), and use $\tilde{\vec{T}}_j$ and $\tilde{\vec{V}}$ to denote the counterpart of $\vec{T}_j$ and $\vec{V}$ after the change of variable (2.10).

Note that $\tilde{\phi}$ has the Taylor expansion

$$\tilde{\phi}(\mathbf{x}; \xi) = \sum_{i,j} c_{i,j} x_i \xi_j + \sum_i x_i \cdot \xi^T (D_i + \sum_j c_{i,j} A_j) \xi + \sum_j \xi_j \cdot \mathbf{x}^T E_j \mathbf{x}$$

$$+ \sum_{i_1,i_2,j_1,j_2} f_{i_1,i_2,j_1,j_2} x_{i_1} x_{i_2} \xi_{j_1} \xi_{j_2} + \sum_j \xi^T A_j \xi \cdot \mathbf{x}^T E_j \mathbf{x} \tag{2.12}$$

$$+ \text{ terms playing no role}$$

Comparing (2.11) and (2.12), we see $\vec{V}$ and $\tilde{\vec{V}}$ differ at the origin by terms of order at least 1 in $\mathbf{x}$ or ξ. Denote $\vec{V}_0 = (V_{0,1}, \ldots, V_{0,n})$ to be the common value of $\vec{V}$ and $\tilde{\vec{V}}$ at $(\mathbf{0}, 0)$. Now

$$(\tilde{\vec{V}} \cdot \nabla_{\mathbf{x}}) \nabla_\xi^2 \tilde{\phi}(\mathbf{0}; 0) - (\vec{V} \cdot \nabla_{\mathbf{x}}) \nabla_\xi^2 \phi(\mathbf{0}; 0)$$

$$= \sum_i \sum_j V_{0,i} c_{ij} (2A_j) = 0$$

where the last equality is because by definition, $\vec{V}_0$ is orthogonal to each $(c_{1j}, \ldots, c_{nj})^T$.

Next we compare $(\tilde{\vec{V}} \cdot \nabla_{\mathbf{x}})^2 \nabla_\xi^2 \tilde{\phi}$ and $(\vec{V} \cdot \nabla_{\mathbf{x}})^2 \nabla_\xi^2 \phi$. We will show that they are also equal by using (2.11) and (2.12) to compute their difference. We begin by showing that near the origin,

$$(\tilde{\vec{V}} - \vec{V})(\mathbf{x}; \xi) = O(|\xi| + |\mathbf{x}|^2), \tag{2.13}$$

which will greatly simplify our computation. Indeed, use the definition (2.2) of $\tilde{\vec{V}}$ and $\vec{V}$, we see that both vectors have the same constant terms. Moreover, observe that in the wedge definition (2.2) of $\tilde{\vec{V}}$ and $\vec{V}$, all $\vec{T}$ have the same linear term in $\mathbf{x}$, since the coefficients of all $x_{i_1} x_{i_2} \xi_j$ terms are the same for ϕ and $\tilde{\phi}$. Hence $\tilde{\vec{V}}$ and $\vec{V}$ also have the same linear term in $\mathbf{x}$ and (2.13) is seen to hold.

By (2.13), if we write

$$\vec{V}(\mathbf{x}; \xi) = \vec{V}_0 + \sum_{i=1}^n x_i \vec{U}_i + O(|\xi| + |\mathbf{x}|^2), \tag{2.14}$$

we can reduce the effect of both $(\tilde{\vec{V}} \cdot \nabla_\mathbf{x})^2$ and $(\vec{V} \cdot \nabla_\mathbf{x})^2$ at the origin to the action of a constant coefficient differential operator $\mathcal{D}_0 = (\vec{V}_0 \cdot \nabla_\mathbf{x})^2 + \sum_j V_{0,j}(\vec{U}_j \cdot \nabla_\mathbf{x})$. Hence

$$(\tilde{\vec{V}} \cdot \nabla_\mathbf{x})^2 \nabla_\xi^2 \tilde{\phi}(\mathbf{0}; 0) - (\vec{V} \cdot \nabla_\mathbf{x})^2 \nabla_\xi^2 \phi(\mathbf{0}; 0)$$
$$= \mathcal{D}_0(\tilde{\phi} - \phi)(\mathbf{0}; 0) = \sum_j 4(\vec{V}_0^T E_j \vec{V}_0) A_j + 2 \sum_j \vec{C}_j^T U \vec{V}_0 A_j \qquad (2.15)$$

where the column vector $\vec{C}_j = (c_{1,j}, \ldots, c_{n,j})^T$ and the $n \times n$ matrix U has the k-th column equal to $\vec{U}_k$ and $(\cdot)_j$ means the j-th component.

In order to show (2.15) gives 0, it suffices to show the stronger statement

$$2E_j \vec{V}_0 + U^T \vec{C}_j = 0, \forall 1 \le j \le n - 1. \qquad (2.16)$$

We prove (2.16) entrywisely. Take an arbitrary $1 \le k \le n$, we need to prove

$$2\vec{E}_{j;k} \cdot \vec{V}_0 + \vec{U}_k \cdot \vec{C}_j = 0 \qquad (2.17)$$

where $\vec{E}_{j;k}$ is the k-th row (or column) of E_j. To show this we recall how one obtain $\vec{V}_0$ and $\vec{U}_k$. Recall from (2.2) and (2.11),

$$\vec{V}_0 = \vec{C}_1 \wedge \vec{C}_2 \wedge \cdots \wedge \vec{C}_{n-1}. \qquad (2.18)$$

To compute $\vec{U}_k$, note that it is the coefficient of x_k in (2.14). Its computation boils down to expanding the $\vec{T}_j$ in (2.1) and (2.2) into the constant term, the x_k term and higher terms. We see

$$\vec{U}_k = 2\vec{E}_{1;k} \wedge \vec{C}_2 \wedge \cdots \wedge \vec{C}_{n-1} + 2\vec{C}_1 \wedge \vec{E}_{2;k} \wedge \cdots \wedge \vec{C}_{n-1} + \cdots$$
$$+ 2\vec{C}_1 \wedge \cdots \wedge \vec{C}_{n-2} \wedge \vec{E}_{n-1;k}. \qquad (2.19)$$

Now there is only one nonzero term in the above expression, namely $2\vec{C}_1 \wedge \cdots \wedge \vec{E}_{j;k} \wedge \cdots \wedge \vec{C}_{n-1}$, that contributes to $\vec{U}_k \cdot \vec{C}_j$. Hence

$$\vec{U}_k \cdot \vec{C}_j = 2 \det(\vec{C}_1, \ldots, \vec{C}_{j-1}, \vec{E}_{j;k}, \vec{C}_{j+1}, \ldots, \vec{C}_{n-1}, \vec{C}_j). \qquad (2.20)$$

But we also have

$$2\vec{E}_{j;k} \cdot \vec{V}_0 = 2 \det(\vec{C}_1, \ldots, \vec{C}_{n-1}, \vec{E}_{j;k}). \qquad (2.21)$$

Since the two add up to 0, we see (2.17), and thus (2.16) holds. This concludes the proof that both terms in (2.3) at $(\mathbf{0}; 0)$ remain the same under the change of variables (2.10), finishing the proof that the property (2.3) is preserved under every diffeomorphism in $\mathbf{x}$ only or in ξ only.

Now we just need to prove that if the phase function is in the normal form (1.7), Bourgain's condition (1.10) at the origin coincides with (2.3). Indeed, in the normal form, the expression of every $\vec{T}_j$ has no linear term in $\mathbf{x}$ and thus the action of $\vec{V}$ at the origin is the same as ∂_t and the action of $\vec{V}^2$ at the origin is the same as ∂_t^2, verifying the above claim. $\qquad \square$

The first part of the above proof immediately implies the following useful corollary.

Lemma 2.3 *Let $\lambda(\mathbf{x}; \xi)$ be a smooth scalar function that does not take value zero. Then for a phase function $\phi(\mathbf{x}; \xi)$ satisfying conditions (H1) and (H2), it satisfies Bourgain's condition at $(\mathbf{x}_0; \xi_0)$ if and only if (2.3) holds with $\vec{V}$ replaced by $\lambda \cdot \vec{V}$.*

2.2 Proof of Theorem 1.1

Given a phase function $\phi(\mathbf{x}; \xi)$, we would like to show that (1.4) may fail for some $q > \frac{2n}{n-1}$ and some $f \in L^\infty$. Let us turn to the dual form of it:

$$\int \left| \int g(\mathbf{x}) e^{iN\phi(\mathbf{x};\xi)} a(\mathbf{x}; \xi) d\mathbf{x} \right| d\xi \lesssim N^{-n/q} \|g\|_{q'}, \tag{2.22}$$

for $q > \frac{2n}{n-1}$. Let $\delta \simeq N^{-1/2}$. Consider a δ-net $\{\xi_\alpha\}$ in the ξ variable. For each α, we will introduce a curved tube T_α, whose bottom is a disc of radius δ and length is about δ^λ with λ to be determined. Let $\mathbf{T}$ denote the collection of tubes $\{T_\alpha\}$ and let $\#\mathbf{T}$ denote the number of tubes. Moreover, we will find a function $\mathbf{\Omega} = (\Omega_1(\xi), \ldots, \Omega_{n-1}(\xi))$ such that for every α and every $\mathbf{x} \in T_\alpha$, we have

$$\left| \nabla_\xi \phi(\mathbf{x}; \xi_\alpha) - \mathbf{\Omega}(\xi_\alpha) \right| \leq \delta. \tag{2.23}$$

Afterwards, let us set

$$g(\mathbf{x}) = g_\epsilon(\mathbf{x}) = \sum_\alpha \epsilon_\alpha e^{-iN\phi(\mathbf{x};\xi_\alpha)} \chi_{T_\alpha}(\mathbf{x}) \tag{2.24}$$

where χ_{T_α} is the indicator function of T_α and ϵ_α takes ± 1 randomly. If (2.22) holds, then

$$\int \max_\alpha \left| \int_{T_\alpha} e^{iN[\phi(\mathbf{x};\xi) - \phi(\mathbf{x};\xi_\alpha)]} a(\mathbf{x}; \xi) d\mathbf{x} \right| d\xi$$
$$\leq \int \left(\sum_\alpha \left| \int_{T_\alpha} e^{iN[\phi(\mathbf{x};\xi) - \phi(\mathbf{x};\xi_\alpha)]} a(\mathbf{x}; \xi) d\mathbf{x} \right|^2 \right)^{1/2} d\xi \tag{2.25}$$

which, by Khintchine's inequality, is bounded by

$$N^{-n/q} \left(\int \left(\sum_\alpha \chi_{T_\alpha} \right)^{q'/2} d\mathbf{x} \right)^{1/q'}. \tag{2.26}$$

We apply Hölder's inequality, and obtain

$$(2.26) \lesssim N^{-n/q} \left| \bigcup_\alpha T_\alpha \right|^{\frac{1}{2} - \frac{1}{q}} \left(\sum |T_\alpha| \right)^{\frac{1}{2}}. \tag{2.27}$$

Next, we give a lower bound for the left hand side of (2.25). Recall the function $\mathbf{\Omega}$. Write

$$
\begin{aligned}
\phi(\mathbf{x}; \xi) &= \phi(\mathbf{x}; \xi_\alpha) + \langle \nabla_\xi \phi(\mathbf{x}; \xi_\alpha), \xi - \xi_\alpha \rangle + O(\delta^2) \\
&= \phi(\mathbf{x}; \xi_\alpha) + \langle \mathbf{\Omega}(\xi_\alpha), \xi - \xi_\alpha \rangle + O(\delta^2), \ \mathbf{x} \in T_\alpha.
\end{aligned}
\tag{2.28}
$$

Therefore the left hand side of (2.25) is at least

$$
\delta^{n-1} \sum |T_\alpha|.
\tag{2.29}
$$

We combine (2.27) and (2.29), and obtain

$$
\delta^{\lambda/2} \lesssim N^{-n/q} \delta^{-(n-1)} \Big| \bigcup_\alpha T_\alpha \Big|^{1/2 - 1/q}.
\tag{2.30}
$$

In the remaining part, we will construct $\{T_\alpha\}$ so that the union of these tubes is small. So far we have been following Bourgain's framework in [3]. The improvement over Bourgain's result in dimension $n = 3$ comes from the construction of the tubes $\{T_\alpha\}$.

Let us write our phase function in its normal form at the origin, that is,

$$
\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t\langle A\xi, \xi \rangle + O(|t||\xi|^3 + |\mathbf{x}|^2 |\xi|^2).
\tag{2.31}
$$

The following lemma is the key for the construction of $\{T_\alpha\}$.

Lemma 2.4 *If Bourgain's condition fails at the origin, then we can find $\mathbf{\Omega}$ with $\mathbf{\Omega}(0) = 0$ such that the following holds*: *Let $X_t(\xi) : \mathbb{R}^{n-1} \mapsto \mathbb{R}^{n-1}$ denote the unique solution to $\nabla_\xi \phi(x, t; \xi) = \mathbf{\Omega}(\xi)$ in the x variable, then*

$$
\big| \det \nabla_\xi X_t \big| = O(|(t, \xi)|^n),
\tag{2.32}
$$

for t, ξ small.

Let us assume Lemma 2.4 and finish the proof of Theorem 1.1. As a consequence of this lemma, if we define

$$
T_\alpha = \{(x, t) : |x - X_t(\xi_\alpha)| \le \delta, 0 \le t \le \delta^\lambda\},
\tag{2.33}
$$

then (2.23) holds by mean value theorems. Here the value that $\lambda > 0$ takes is not relevant, that is, λ can even be very close to zero. Next, pick $\lambda = 1/(n + 1)$. Lemma 2.4 then says that the union of the tubes is small:

Claim 2.5

$$
\Big| \bigcup_{\alpha : |\xi_\alpha| \le \delta^\lambda} T_\alpha \Big| \lesssim_\lambda \delta^{2n\lambda}.
\tag{2.34}
$$

Let us first accept the above claim. For $|\xi_\alpha| \geq \delta^\lambda$, we will construct tubes in the same way. More precisely, we will cut frequency ξ space into balls of radius δ^λ, and on each ball, we repeat the above construction. Therefore

$$\left| \bigcup_\alpha T_\alpha \right| \lesssim \delta^{2n\lambda} \delta^{-(n-1)\lambda}. \tag{2.35}$$

Substituting this into (2.30) will give us

$$q \geq \frac{2(2n^2 + n - 1)}{2n^2 - n - 2} > \frac{2n}{n-1}. \tag{2.36}$$

This finishes the proof of the theorem, modulo the proof of Lemma 2.4 and the proof of Claim 2.5.

Proof of Lemma 2.4. Let us start with (2.31). Write

$$\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t \langle A\xi, \xi \rangle + t^2 Q_2(\xi) + \phi_4(\mathbf{x}; \xi), \tag{2.37}$$

where $Q_2(\xi)$ is a quadratic form in ξ. The assumption that Bourgain's condition fails at the origin is then equivalent to saying that

$$\text{Hessian}(Q_2) \text{ is not a multiple of } A. \tag{2.38}$$

Let $\mathbf{\Omega} = (\Omega_1(\xi), \ldots, \Omega_{n-1}(\xi))$ be smooth with $\mathbf{\Omega}(0) = 0$. We need to solve

$$x + tA\xi + t^2 \nabla_\xi Q_2(\xi) + \nabla_\xi \phi_4(\mathbf{x}; \xi) = \mathbf{\Omega}(\xi). \tag{2.39}$$

It is not difficult to see that when $\mathbf{x}$ and ξ are small, the solution is unique. We solve (2.39) and write the solution as

$$X_t(\xi) = -tA\xi - t^2 B\xi + t P_1(\xi) + t^2 P_2(\xi)$$
$$+ \sum_{j=3}^{n} t^j P_j(\xi) + \widetilde{\mathbf{\Omega}}(\xi) + O(|(t, \xi)|^{n+1}), \tag{2.40}$$

where B is a $(n-1) \times (n-1)$ matrix that is not a constant multiple of A, $B\xi = \nabla_\xi Q_2(\xi)$, $\widetilde{\mathbf{\Omega}}$ depends on $\mathbf{\Omega}$, $P_i(\xi)$ is a polynomial of degree $n-i$ with lowest order term of degree 2 for $i = 1, 2$, and $P_j(\xi)$ is a polynomial of degree $n-i$ with lowest order term of degree 1 for $j \geq 3$. Here $P_i (1 \leq i \leq n)$ depends on $\mathbf{\Omega}$, but it is important that the matrix B does not depend on $\mathbf{\Omega}$. Before computing $\nabla_\xi X_t$, let us do the change of variables

$$-A\xi + P_1(\xi) \mapsto A\eta. \tag{2.41}$$

Write the right hand side of (2.40) in the η variable:

$$tA\eta + t^2 B\eta + t^2 P_2'(\eta) + \sum_{j=3}^{n} t^j P_j'(\eta) + \widetilde{\mathbf{\Omega}}'(\eta) + O(|(t, \eta)|^{n+1}) =: X_t'(\eta), \tag{2.42}$$

where $P_2'(\eta)$ is a polynomial of degree $n-2$ with lowest order term of degree 2, and $P_j'(\eta)$ is a polynomial of degree $n-i$ with lowest order term of degree 1 for $j \geq 3$. As the change of variables in (2.41) is non-degenerate, in order to guarantee (2.32), we just need to show that

$$\left| \det \nabla_\eta X_t' \right| = O(|(t, \eta)|^n). \tag{2.43}$$

When computing (2.43), we will see more clearly why it is convenient to do the change of variables in (2.41). Write

$$X_t'(\eta) = t A \eta + t^2 B \eta + \widetilde{\mathbf{\Omega}}'(\eta) + \phi_4'(t, \eta). \tag{2.44}$$

Note that the lowest order term in ϕ_4', jointly in t and η variables, is four. Compute

$$\nabla_\eta X_t' = t A + t^2 B + \nabla_\eta \widetilde{\mathbf{\Omega}}'(\eta) + \nabla_\eta \phi_4'(t, \eta). \tag{2.45}$$

Next we will compute the determinant. As A is non-degenerate, when computing the determinant, we can without loss of generality assume that A is the identity matrix. Moreover, by using Jordan normal forms for B, and using the fact that B is not a multiple of A, we can therefore without loss of generality assume that B is of the form $[b_{ij}]_{1 \leq i, j \leq n-1}$ with $b_{i1} = b_{i2} = 0$ for every $i \geq 3$, and the leading principal minor of order 2 is one of the following forms

$$\begin{bmatrix} \gamma, & 1 \\ 0, & \gamma \end{bmatrix} \text{ or } \begin{bmatrix} \gamma', & 0 \\ 0, & 0 \end{bmatrix} \text{ or } \begin{bmatrix} \gamma_1, & -\gamma_2 \\ \gamma_2, & \gamma_1 \end{bmatrix} \tag{2.46}$$

where $\gamma, \gamma', \gamma_1, \gamma_2 \in \mathbb{R}$ and $\gamma' \neq 0, \gamma_2 \neq 0$. Write

$$\nabla_\eta X_t'(\eta) = \nabla_\eta \widetilde{\mathbf{\Omega}}'(\eta) + \begin{bmatrix} *, & *, & 1_{t,y}, & 1_{t,y}, & \cdots \\ *, & *, & 1_{t,y}, & 1_{t,y}, & \cdots \\ 3_{t,y}, & 3_{t,y}, & 1_{t,y}, & 1_{t,y}, & \cdots \\ 3_{t,y}, & 3_{t,y}, & 1_{t,y}, & 1_{t,y}, & \cdots \\ \cdots, & \cdots, & \cdots, & \cdots, & \cdots \end{bmatrix} \tag{2.47}$$

where $i_{t,y}$ means that the lowest order in t, y is i, for $i = 1, 3$. If we are in the first case in (2.46), then we pick $\widetilde{\mathbf{\Omega}}'$ such that

$$\nabla_\eta \widetilde{\mathbf{\Omega}}' = \begin{bmatrix} 0, & 0, & 0, & \cdots \\ 1, & 0, & 0, & \cdots \\ 0, & 0, & 0, & \cdots \\ \cdots, & \cdots, & \cdots, & \cdots \end{bmatrix} \tag{2.48}$$

If we are in the second case, then we pick $\widetilde{\mathbf{\Omega}}'$ such that

$$\nabla_\eta \widetilde{\mathbf{\Omega}}' = \frac{1}{\gamma'} \begin{bmatrix} 1, & 1, & 0, & \cdots \\ -1, & -1, & 0, & \cdots \\ 0, & 0, & 0, & \cdots \\ \cdots, & \cdots, & \cdots, & \cdots \end{bmatrix} \tag{2.49}$$

The last case in (2.46) can be handled in the same way. These choices of $\widetilde{\boldsymbol{\Omega}}'$ will guarantee that (2.43) holds. In the end, to find $\boldsymbol{\Omega}$ from $\widetilde{\boldsymbol{\Omega}}'$, we just need to revert the change of variables in (2.41). $\qquad\square$

Proof of Claim 2.5 Let us start by sketching the ideas in the proof. To begin with, we replace the set $\bigcup_{\alpha:|\xi_\alpha|\leq\delta^\lambda} T_\alpha$ by a larger set $\mathcal{N}_{2\delta}(X(B))$, where X is the map $(\xi,t)\mapsto(X_t(\xi),t)$, B is a ball of radius $O(\delta^\lambda)$ around the origin and $\mathcal{N}_{2\delta}$ refers to the 2δ neighbourhood. Now intuitively the volume of $\mathcal{N}_{2\delta}(X(B))$ depends on the volume of $X(B)$ and the "surface area" of $X(B)$. We will control the volume of $X(B)$ by Lemma 2.4. For its "surface volume", we first observe that if ϕ and $\boldsymbol{\Omega}$ are polynomials, $\partial X(B)$ is contained in a nice semialgebraic set of dimension $< n$ and hence has a controlled "surface volume" by tools in real algebraic geometry. Finally, the general situation can be reduced to the above polynomial situation by a Taylor series approximation. To establish the semialgebracity above, we will use quantifier elimination based on the Tarski-Seidenberg theorem. For a recent application of quantifier elimination in Kakeya and restriction that also helped us to motivate the present proof, see [11]. The tools we need from real algebraic geometry can be found in references [1, 23].

We now present the proof details. First we claim that without loss of generality, one may assume ϕ and all components of $\boldsymbol{\Omega}$ are polynomials of degree $O_\lambda(1)$. To see this, let $M \geq 5$ be a large positive integer to be determined later and replace ϕ and $\boldsymbol{\Omega}$ by their degree M Taylor approximations. Since $M \geq 5$, ϕ will stay as a legitimate phase function and Bourgain's condition continues to fail at the origin. Moreover, whenever $|t|, |\xi_\alpha| \leq \delta^\lambda$, the change of $\boldsymbol{\Omega}(\xi_\alpha)$ is $O(\delta^{(M+1)\lambda})$. Thus for these t and ξ_α, the distance between the new $X_t(\xi_\alpha)$ from the old one is $O(\delta^{(M+1)\lambda})$ by the nondegeneracy of ϕ. Since we only care about the volume of the union of δ-neighborhoods, it suffices to choose $M > \frac{1}{\lambda}$ so that each old T_α is contained in the twice-thickening of the corresponding new T_α. Now the old situation is reduced to the new situation where ϕ and all components of $\boldsymbol{\Omega}$ are polynomials of degree $O_\lambda(1)$.

Take B to be a ball of radius $O(\delta^\lambda)$ centered at the origin in the (ξ, t) space containing all (ξ, t) with $|\xi|, |t| \leq \delta^\lambda$. Let X denote the map $(\xi, t) \mapsto (X_t(\xi), t)$. X is smooth near the origin by the implicit function theorem. By definition, $\bigcup_{\alpha:|\xi_\alpha|\leq\delta^\lambda} T_\alpha$ is contained in $\mathcal{N}_{2\delta}(X(B))$. It suffices to prove

$$\left|\mathcal{N}_{2\delta}(X(B))\right| \lesssim_\lambda \delta^{2n\lambda}. \tag{2.50}$$

Let us understand the geometry of $X(B)$. For a point in $\partial X(B)$, either it is in $X(B)$ and hence in $X(\mathrm{Sing}(X; B))$ where $\mathrm{Sing}(X; B)$ is the singular set of X inside B, or it is outside of $X(B)$ and hence by a compactness argument it is in $X(\partial B)$. Since

$$\mathcal{N}_{2\delta}(X(B)) \in X(B) \bigcup \mathcal{N}_{2\delta}(\partial X(B)), \tag{2.51}$$

we have

$$\mathcal{N}_{2\delta}(X(B)) \subseteq X(B) \bigcup \mathcal{N}_{2\delta}(X(\mathrm{Sing}(X; B))) \bigcup \mathcal{N}_{2\delta}(X(\partial B)) \tag{2.52}$$

and will next bound the measures of all three sets on the right-hand side from above.

First we bound $|X(B)|$. By the definition of X and Lemma 2.4, $|\det \nabla X| = |\det \nabla_\xi X_t| = O(|(t, \xi)|^n)$. Integrating on B we get

$$\left|X(B)\right| \lesssim |B| \sup_{(\xi,t)\in B} |(\xi, t)|^n \lesssim \delta^{2n\lambda}. \tag{2.53}$$

Next we bound $|\mathcal{N}_{2\delta}(X(\mathrm{Sing}(X; B)))|$. By the chain rule and the non-degeneracy of $\nabla_x \nabla_\xi \phi$ near 0, we can rewrite

$$X(\mathrm{Sing}(X; B)) = \{(x, t) : \exists (\xi, t) \in B \text{ s.t. } \nabla_\xi \phi(x, t; \xi) = \mathbf{\Omega}(\xi)$$
$$\text{and } \nabla_\xi^2 \phi(x, t; \xi) = \nabla_\xi \mathbf{\Omega}(\xi)\}. \tag{2.54}$$

We will analyze this set using tools in real algebraic geometry and first do some setup. We recall a subset of some $\mathbb{R}^N$ is *semialgebraic* if it can be obtained by finitely many steps of taking unions, intersections, or complements from algebraic sets. The *complexity* of a semialgebraic set is the smallest possible sum of the degrees of all polynomials appearing in a complete description of it. Section 2 of [11] has a good introduction to basic properties of the above notions, as well as a quantitative quantifier elimination (or a quantitative Tarski-Seidenberg theorem) that we will use below, from analysts' viewpoint. A semialgebraic set in $\mathbb{R}^N$ has a *dimension* that is a non-negative integer $\leq N$. See Chap. 5 of [1] for its basic properties.

Note that the ball B is a semialgebraic set of complexity $O(1)$. Moreover, ϕ and components of $\mathbf{\Omega}$ are already polynomials of degree $O_\lambda(1)$. Hence by quantitative quantifier elimination (or the quantitative Tarski-Seidenberg theorem, see Theorem 14.16 of [1]), $X(\mathrm{Sing}(X; B))$ is a semialgebraic set of complexity $O_\lambda(1)$. By Sard's theorem, $X(\mathrm{Sing}(X; B))$ has measure zero and thus has dimension $< n$ (by Proposition 5.53 of [1]). Now by Corollary 5.7 of [23], $X(\mathrm{Sing}(X; B))$ can be covered by $O_\lambda(1) \times (\delta^{\lambda-1})^{(n-1)}$ many δ-balls. Hence

$$\left|\mathcal{N}_{2\delta}(X(\mathrm{Sing}(X; B)))\right| \lesssim_\lambda \delta^{1+(n-1)\lambda}. \tag{2.55}$$

We remark that Corollary 5.7 of [23] can be viewed qualitatively as a generalization of Wongkew's theorem [22] to the semialgebraic setting.

Finally we bound $|\mathcal{N}_{2\delta}(X(\partial B))|$ via a similar application of real algebraic geometrical tools. First write

$$X(\partial B) = \{(x, t) : \exists (\xi, t) \in \partial B \text{ s.t. } \nabla_\xi \phi(x, t; \xi) = \mathbf{\Omega}(\xi)\}$$

and we see $X(\partial B)$ is a semialgebraic set of complexity $O_\lambda(1)$. Since X is smooth on the dilation $2B$, $X(\partial B)$ has zero measure and thus has dimension $< n$. Applying Corollary 5.7 of [23] as before we get

$$\left|\mathcal{N}_{2\delta}(X(\partial B))\right| \lesssim_\lambda \delta^{1+(n-1)\lambda}. \tag{2.56}$$

Combining (2.52), (2.53), (2.55) and (2.56) and noticing that our $\lambda = \frac{1}{n+1}$, we finish the proof of the claim. $\qquad\square$

3 Polynomial Wolff axiom: proof of Theorem 1.2

In this section, we will prove Theorem 1.2. In particular, we will see that Bourgain's condition also appears very naturally in the proof, see (3.32)-(3.34). This may not be surprising, as polynomial Wolff axioms are "opposite" to Bourgain's Kakeya compression phenomenon (see the beginning of Sect. 2 for a brief discussion): Polynomial Wolff axioms say that if a family of tubes are supported near algebraic varieties in $\mathbb{R}^n$, then this family of tubes must point in a "small" number of directions; on the other hand, Bourgain's Kakeya compression phenomenon says that (curved) Kakeya sets may concentrate near hyper-surfaces in $\mathbb{R}^n$.

In order to prove this theorem, we first state and prove a generalized version of Polynomial Wolff Axiom by Katz-Rogers [11] as follows. We first introduce more notation. Let $n \geq 2$. Suppose the map

$$\Phi : \mathbb{R}^{n-1} \times \mathbb{R} \times \mathbb{R}^{n-1} \to \mathbb{R}^{n-1} \tag{3.1}$$

is smooth on a neighborhood of $[-1, 1]^{2n-1}$ with $\|\Phi\|_{C^k} \lesssim_k 1, \forall k \geq 1..$[2]

By a *δ-tube for cap θ with respect to* Φ, we mean some

$$T_{\xi_\theta, v, \Phi}(\delta, 1) := \{(x, t) \in \mathbb{R}^n : |x - \Phi(v, t, \xi_\theta)| \leq \delta, |t| \leq 1\} \tag{3.2}$$

where the δ in the name indicates the "thickness" and the 1 in the name indicates the time span of the tube. For a collection $\mathbb{T}$ of tubes $\{T_{\xi_\theta, v, \Phi}(\delta, 1)\}$, we say that the tubes in $\mathbb{T}$ point in different directions if all the underlying θ for them are distinct.

Theorem 3.1 (Generalized Polynomial Wolff Axiom) *Suppose that for every choice of $v \in [-1, 1]^{n-1}$ and $\xi \in [-1, 1]^{n-1}$,*

$$\int_{-\delta^\epsilon}^{\delta^\epsilon} |\det(\nabla_v \Phi(v, t, \xi) \cdot M + \nabla_\xi \Phi(v, t, \xi))| \mathrm{d}t \gtrsim_\epsilon \delta^{C\epsilon}, \forall M \in \mathrm{Mat}_{(n-1)\times(n-1)}(\mathbb{R}) \tag{3.3}$$

for some constant C that depends only on the dimension n, where $\mathrm{Mat}_{(n-1)\times(n-1)}(\mathbb{R})$ stands for the set of $(n-1) \times (n-1)$ real matrices and the bound is uniform and independent of the choices of v, ξ and M. Then for every collection $\mathbb{T}$ of δ-tubes pointing in different directions,

$$\#\{T \in \mathbb{T} : T \subset S\} \leq C(n, E, \epsilon)|S|\delta^{1-n-\epsilon} \tag{3.4}$$

whenever $S \subset B^n$ is a semialgebraic set of complexity $\leq E$.

Moreover the implied constant only depends on bounds of finitely many (depending on n, E, ϵ) derivatives of Φ.

Proof of Theorem 3.1 When first reading the proof we recommend fixing Φ, and it will be easy to see from the proof that the constant only depends on bounds of finitely many derivatives of Φ when we are allowed to change it.

---

[2]Depending on the choice of ϵ, we will only use the boundedness of finitely many derivatives of Φ in the proof.

In the proof we always think of ϵ as fixed and always assume δ is sufficiently small (that can depend on ϵ) since otherwise the conclusion is easily seen to hold. Without loss of generality, we always assume $|S| \gtrsim \delta^{n-1}$ with a suitable absolute constant, since otherwise no T can lie in S. Our proof will largely follow that of Theorem 3.1 in [11].

Our Φ is not necessarily a semialgebraic map and we would like to first make it semialgebraic by a Taylor approximation. Fix a $0 < \epsilon_1 < \min\{0.1, \epsilon\}$ and fix a large $K > \frac{2n^3}{\epsilon_1}$ (with more constraints to be determined on both parameters). Let N be the quantity on the left hand side of (3.4). We may assume $N \geq 1$. Without loss of generality we can assume

$$\#\{T_{\xi_\theta, v, \Phi}(\delta, 1) \in \mathbb{T} : T_{\xi_\theta, v, \Phi}(\delta, 1) \subset S, |\xi_\theta| \leq \delta^{\epsilon_1}, |v| \leq \delta^{\epsilon_1}\} \gtrsim \delta^{2(n-1)\epsilon_1} N \quad (3.5)$$

and will focus on giving an upper bound of the number of these $T_{\xi_\theta, v, \Phi}(\delta, 1)$. Without loss of generality we assume $\Phi(0) = 0$.

Now replace Φ by its K-th Taylor approximation at the origin, called Φ_1. The new map Φ_1 is still in C^∞. For each $T_{\xi_\theta, v, \Phi}(\delta, 1)$ in $\mathbb{T}$ such that $|\xi_\theta| \leq \delta^{\epsilon_1}$ and that $|v| \leq \delta^{\epsilon_1}$, we form a corresponding shrunken tube $T_{\xi_\theta, v, \Phi_1}(\delta/2, 2\delta^{\epsilon_1})$ that is defined similar to (3.2) but have thickness $\delta/2$ and time span $|t| \leq 2\delta^{\epsilon_1}$. By Taylor approximation it is easy to see the entire

$$T_{\xi_\theta, v, \Phi_1}(\delta/2, 2\delta^{\epsilon_1}) \subset T_{\xi_\theta, v, \Phi}(\delta, 1) \subset S. \quad (3.6)$$

Moreover, by Taylor approximation, we see that the following analogue of (3.3) continues to hold for each $|\xi| \leq \delta^{\epsilon_1}$ and $|v| \leq \delta^{\epsilon_1}$ as long as $K \gtrsim 1$ (when δ is sufficiently small) and $M \in \text{Mat}_{(n-1)\times(n-1)}(\mathbb{R})$ are such that all entries of M are $\leq \delta^{-n}$:

$$\int_{-\delta^{\epsilon_1}}^{\delta^{\epsilon_1}} |\det(\nabla_v \Phi_1(v, t, \xi) \cdot M + \nabla_\xi \Phi_1(v, t, \xi))| dt \gtrsim_{\epsilon_1} \delta^{C\epsilon_1} \quad (3.7)$$

Like Katz-Rogers did in [11], we define a set

$$L = \{(\xi, v) : T_{\xi, v, \Phi_1}(\delta/3, 2\delta^{\epsilon_1}) \subset S\}. \quad (3.8)$$

Now the map Φ_1 is algebraic, by definition L is semialgebraic with complexity $O_{n, E, \epsilon_1, K}(1)$ (for the definition of semialgebraic sets and their complexity and how to arrive at the present claim, see Sect. 2 of [11]). Note that if a tube $T_{\xi_\theta, v, \Phi_1}(\delta/2, 2\delta^{\epsilon_1}) \subset S$, then keeping v and perturbing ξ_θ by an arbitrary small distance proportional to δ, the resulting (ξ, v) will end up in L by definition. Hence we know the measure

$$|\{\xi : \exists v \text{ s.t. } (\xi, v) \in L\}| \gtrsim \delta^{2(n-1)\epsilon_1} N \cdot \delta^{n-1} \simeq \delta^{(n-1)+2(n-1)\epsilon_1} N. \quad (3.9)$$

Like in the proof of Theorem 3.1 in [11], next we apply the Tarski-Seidenberg theorem to obtain a semialgebraic section $L' \subset L$ of complexity $O_{n, E, \epsilon_1, K}(1)$ consisting of a single (ξ, v) for each ξ appearing in the set in (3.9). Arguing like [11], we see L' is an $(n-1)$-dimensional subset of $\mathbb{R}^{2n-2}$. Using Gromov's lemma (cited as Lemma

2.3 in [11]) in the identical way as in pages 1711-1712 of [11], we find two polynomial maps F and $G : [0, \delta^{\epsilon_1}]^{n-1} \to \mathbb{R}^{n-1}$ with $\deg F, \deg G = O_{n,E,\epsilon_1,K}(1)$ and $\|F\|_{C^1}, \|G\|_{C^1} \leq 1$ such that

$$|G([0, \delta^{\epsilon_1}]^{n-1})| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N \tag{3.10}$$

and that

$$(\Phi_1(F(x), t, G(x)), t) \in S, \forall x \in [0, \delta^{\epsilon_1}]^{n-1} \text{ and } \forall |t| \leq \delta^{\epsilon_1}. \tag{3.11}$$

For technical reasons that we replaced Φ by Φ_1 which does not need to satisfy (3.3) but instead only satisfies the weaker (3.7), we now pass to a subset B of $[0, \delta^{\epsilon_1}]^{n-1}$ where G has reasonably large Jacobian. Define

$$B = \{x \in [0, \delta^{\epsilon_1}]^{n-1} : |\det \nabla_x G| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N\}. \tag{3.12}$$

We see that if the implied constant above is carefully chosen, by Chebyshev we continue to have the following inequality similar to (3.10):

$$|G(B)| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N. \tag{3.13}$$

Now like in [11], we look at the volume of

$$\Delta = \{(\Phi_1(F(x), t, G(x)), t) : x \in B, |t| \leq \delta^{\epsilon_1}\}. \tag{3.14}$$

On one hand, Δ is contained in S and thus

$$|\Delta| \leq |S|. \tag{3.15}$$

On the other hand, we can bound $|\Delta|$ below by calculus. Note that Φ_1 is a polynomial of degree K and that F and G are polynomials of degree $O_{n,E,\epsilon_1,K}(1)$. Hence by Bézout's theorem for every fixed t the map

$$x \mapsto \Phi_1(F(x), t, G(x)) \tag{3.16}$$

is $O_{n,E,\epsilon_1,K}(1)$ to 1. Thus

$$|\Delta| \simeq_{n,E,\epsilon_1,K} \int_{-\delta^{\epsilon_1}}^{\delta^{\epsilon_1}} \int_B |\nabla_x(\Phi_1(F(x), t, G(x)))| \mathrm{d}x \mathrm{d}t$$

$$= \int_{-\delta^{\epsilon_1}}^{\delta^{\epsilon_1}} \int_B |\det(\nabla_v \Phi_1(F(x), t, G(x)) \cdot \nabla_x F + \nabla_\xi \Phi_1(F(x), t, G(x)) \cdot \nabla_x G| \mathrm{d}x \mathrm{d}t$$

$$= \int_B |\det(\nabla_x G)|$$

$$\cdot \int_{-\delta^{\epsilon_1}}^{\delta^{\epsilon_1}} |\det(\nabla_v \Phi_1(F(x), t, G(x)) \cdot (\nabla_x F \cdot (\nabla_x G)^{-1})$$

$$+ \nabla_\xi \Phi_1(F(x), t, G(x)))| \mathrm{d}t \mathrm{d}x. \tag{3.17}$$

Note that $\|F\|_{C^1}, \|G\|_{C^1} \le 1$ and that for all $x \in B$,

$$|\det \nabla_x G| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N \ge \delta^{(n-1)+C\epsilon_1}, \tag{3.18}$$

we see that each entry of $(\nabla_x F \cdot (\nabla_x G)^{-1})$ is

$$\lesssim_{n,E,\epsilon_1,K} \delta^{-(n-1)-C\epsilon_1} \le \delta^{-n} \tag{3.19}$$

if δ is sufficiently small. This allows us to invoke (3.7) to obtain

$$|\Delta| \gtrsim_{n,E,\epsilon_1,K} \delta^{C\epsilon_1} \int_B |\det \nabla_x G(x)| \mathrm{d}x. \tag{3.20}$$

Use Bézout again and notice (3.13), the right hand side is

$$\simeq_{n,E,\epsilon_1,K} |G(B)| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N. \tag{3.21}$$

Hence

$$|\Delta| \gtrsim_{n,E,\epsilon_1,K} \delta^{(n-1)+C\epsilon_1} N. \tag{3.22}$$

Combine (3.15) and (3.22), we obtain

$$N \lesssim_{n,E,\epsilon_1,K} |S|\delta^{1-n-C\epsilon_1}. \tag{3.23}$$

It suffices to take ϵ_1 to be a suitable multiple of ϵ depending on the above constant C (and fix K accordingly as in the beginning of the proof). $\qquad\square$

Before proving Theorem 1.2, we also state and prove an elementary lemma on the averaged size of determinants.

Lemma 3.2 *For $A, B \in Mat_{k \times k}(\mathbb{R})$ and a measurable $E \subset \mathbb{R}$, we have*

$$\int_E |\det(tA + B)| \mathrm{d}t \gtrsim_k |E|^{k+1} |\det(A)|. \tag{3.24}$$

Proof of Theorem 3.2 Without loss of generality we can assume $\det A \ne 0$. We may further assume $A = I$ since otherwise we can replace A by I and replace B by BA^{-1} and notice $tA + B = A \cdot (tI + BA^{-1})$.

Now $\det(tI + B)$ is a monic polynomial $g_B(t)$ in variable t of degree k. Factorize g_B over $\mathbb{C}$ and notice that the set of $t \in E$ with distance $\ge \frac{|E|}{2k}$ against each root of g_B has measure at least $\frac{|E|}{2}$. For each such t,

$$|\det(tI + B)| \ge \frac{|E|^k}{(2k)^k}. \tag{3.25}$$

This finishes the proof. $\qquad\square$

Note that the key point of Lemma 3.2 is that the estimate (3.24) is independent of B. With the above preparation we now prove Theorem 1.2.

Proof of Theorem 1.2 By the non-degenerate assumption of ϕ, for sufficiently small (v, t, ξ) one can find a unique $\Phi = \Phi(v, t, x)$ near 0 such that

$$(\nabla_\xi \phi)(\Phi, t; \xi) = v. \tag{3.26}$$

Let us assume the above can be done for all $(v, t, \xi) \in [-1.5, 1.5]^{2n-1}$ without loss of generality since otherwise we can perform a constant rescaling. It suffices to show that our Φ satisfies the condition (3.3) since we can then use Theorem 3.1 to conclude the proof.

By (3.26), we get

$$(\nabla_\xi \phi)(\Phi(v, t, \xi), t; \xi) = v. \tag{3.27}$$

Differentiating with respect to ξ and v respectively, we deduce

$$\nabla_x \nabla_\xi \phi \cdot \nabla_\xi \Phi + \nabla_\xi^2 \phi = 0 \tag{3.28}$$

and

$$\nabla_x \nabla_\xi \phi \cdot \nabla_v \Phi = I. \tag{3.29}$$

Note that we have adopted the abbreviation that ϕ is evaluated at $(\Phi(v, t, \xi), t; \xi)$. By the non-degeneracy of ϕ, we know $|\nabla_x \nabla_\xi \phi| \simeq 1$. Hence

$$\int_{-\delta^\epsilon}^{\delta^\epsilon} |\det(\nabla_v \Phi(v, t, \xi) \cdot M + \nabla_\xi \Phi(v, t, \xi))| dt$$

$$\simeq \int_{-\delta^\epsilon}^{\delta^\epsilon} |\det(\nabla_x \nabla_\xi \phi \cdot \nabla_v \Phi(v, t, \xi) \cdot M + \nabla_x \nabla_\xi \phi \cdot \nabla_\xi \Phi(v, t, \xi))| dt$$

$$\simeq \int_{-\delta^\epsilon}^{\delta^\epsilon} |\det(M - \nabla_\xi^2 \phi(\Phi(v, t, \xi), t; \xi))| dt. \tag{3.30}$$

Recall we only need to verify (3.3) for Φ. In light of (3.30), it now suffices to show that the right hand side of (3.30) is $\gtrsim_\epsilon \delta^{n\epsilon}$ (independent of the choice of M, $v \in [-1, 1]^{n-1}$ and $\xi \in [-1, 1]^{n-1}$).

From now on we fix v and omit it from place to place, and all the estimates will be uniform in v. For simplicity we use $X_t(\xi)$ to denote $\Phi(v, t, \xi)$ below. Denote

$$A(t; \xi) = \nabla_\xi^2 \phi(X_t(\xi), t; \xi). \tag{3.31}$$

We claim that for all $t \in [0, 1]$, $\xi \in [0, 1]^{n-1}$, $A(t; \xi) - A(0; \xi)$ as a matrix is proportional to a matrix $B(\xi)$ independent of t. This will be refered to as Claim $(*)$. We will prove this claim by finding $B(\xi)$ explicitly. Compute

$$\partial_t A(t; \xi) = \left(\nabla_x \cdot \partial_t X_t\right) \nabla_\xi^2 \phi(X_t(\xi), t; \xi) + \partial_t \nabla_\xi^2 \phi(X_t(\xi), t; \xi)$$

$$= \left[\left(\vec{V} \cdot \nabla_\mathbf{x}\right) \nabla_\xi^2 \phi\right](X_t(\xi), t; \xi) \tag{3.32}$$

where $\vec{V} := (\partial_t X_t, 1)$. Note that if we differentiate both side of (3.26) in t, then we obtain

$$\partial_t X_t \cdot \nabla_x \nabla_\xi \phi(X_t(\xi), t; \xi) + \partial_t \nabla_\xi \phi(X_t(\xi), t; \xi) = 0. \qquad (3.33)$$

This means at the point $(X_t(\xi), t; \xi)$, the vector field $\vec{V}$ is parallel to the one defined in (2.2). Moreover by a similar computation, we obtain that

$$\partial_t^j A(t; \xi) = \left[\left(\vec{V} \cdot \nabla_\mathbf{x} \right)^j \nabla_\xi^2 \phi \right] (X_t(\xi), t; \xi) \qquad (3.34)$$

for every $j \geq 1$. By Lemma 2.3, $\partial_t^2 A(t; \xi)$ is always parallel to $\partial_t A(t; \xi)$. Hence by considering the time derivative of the quotient of two entries we see $\partial_t A(t; \xi)$ is always parallel to $\partial_t A(0; \xi)$. By our non-degeneracy assumption (1.2) for ϕ, we see $\partial_t A(0; \xi)$ has norm and determinant $\simeq 1$ and call it $B(\xi)$. Since $\partial_t A(t; \xi)$ is always parallel to $B(\xi)$, we see that Claim ($*$) holds.

Now by Claim ($*$), we assume

$$A(t; \xi) = f(t; \xi) B(\xi) + A(0; \xi) \qquad (3.35)$$

for some scalar function f. Since

$$\partial_t A(t; \xi)|_{t=0} = B(\xi), \qquad (3.36)$$

we see that the time derivative of f is 1 at $t = 0$. By compactness, the range of $f(t; \xi)$ for $|t| \leq \delta^\epsilon$ has measure $\geq C\delta^\epsilon$ for some universal $C > 0$ independent of $\xi \in [-1, 1]$ and ϵ.

We are in a position to apply Lemma 3.2 to the right hand side of (3.30). Note that

$$M - \nabla_\xi^2 \phi(\Phi(v, t, \xi), t; \xi) = M - A(t; \xi) = (M - A(0; \xi)) - f(t; \xi) B(\xi). \quad (3.37)$$

We see the right hand side of (3.30) is bounded below by $C\delta^{n\epsilon} \det B(\xi) \geq C\delta^{n\epsilon}$, thus concluding the proof. $\qquad \square$

4 Preliminaries for the proof of Theorem 1.3

This section is a preparation for the proof of Theorem 1.3. Materials in this section are standard: We follow the frameworks of Bourgain and Guth [4], Guth [9], Guth, Hickman and Iliopoulou [10], Hickman and Rogers [12] and reduce Theorem 1.3 to a broad-norm estimate (see Theorem 4.2 below).

For $\lambda \geq 1$, denote

$$\phi^\lambda(\mathbf{x}; \xi) = \lambda \phi(\mathbf{x}/\lambda; \xi), \quad a^\lambda(\mathbf{x}; \xi) = a(\mathbf{x}/\lambda; \xi). \qquad (4.1)$$

Define an operator

$$T^\lambda f(\mathbf{x}) := \int e^{i\phi^\lambda(\mathbf{x}; \xi)} a^\lambda(\mathbf{x}; \xi) f(\xi) d\xi. \qquad (4.2)$$

Note that $T^\lambda f$ is just a rescaled version of the operator in Theorem 1.3, and we use this rescaled version as we will use the wave packet decomposition and uncertainty principles to bound T^λ. In the rest of the paper, we will prove the following Theorem.

Theorem 4.1 *If ϕ is assumed to satisfy (H1), (H2$^+$) and Bourgain's condition at every point, then*

$$\left\| T^\lambda f \right\|_{L^p(B_\lambda)} \lesssim_{\epsilon, p, \phi, a} \lambda^\epsilon \| f \|_{L^p} \tag{4.3}$$

for every $p > q_{n,2}$, defined in (1.20), ball $B_\lambda \subset \mathbb{R}^n$ of radius $\lambda \geq 1$ and every $\epsilon > 0$.

For the sake of simplicity, we will assume that our phase function ϕ is of the normal form. To prove (4.3), it suffices to prove

$$\left\| T^\lambda f \right\|_{L^p(B_R)} \lesssim_{\epsilon, p, \phi, a} R^\epsilon \| f \|_{L^p}, \tag{4.4}$$

for every $1 \leq R \leq \lambda^{1-\epsilon}$ and every cube $B_R \subset B_\lambda$. We will run an induction on both parameters λ and R. The base case of the induction $\lambda = R = 1$ is trivial. Let us assume that we have proven

$$\left\| T^{\lambda'} f \right\|_{L^p(B_{R'})} \lesssim_{\epsilon, p, \phi, a} (R')^\epsilon \| f \|_{L^p}, \tag{4.5}$$

for every $\lambda' \leq \lambda/2$, $R' \leq (\lambda')^{1-\epsilon}$ and every cube $B_{R'} \subset B_{\lambda'}$. Our goal is to prove that the same holds with λ and R.

4.1 Wave packet decomposition

Let $1 \leq r \leq R$ and take a collection Θ_r of dyadic cubes of side length $\frac{9}{11} r^{-1/2}$ covering the ball $B^{n-1}(0, 2)$. We take a smooth partition of unity $(\psi_\theta)_{\theta \in \Theta_r}$ with supp $\psi_\theta \subset \frac{11}{10}\theta$ for the ball $B^{n-1}(0, 2)$ such that

$$\left\| \partial_w^\alpha \psi_\theta \right\|_{L^\infty} \lesssim_\alpha r^{|\alpha|/2}$$

for any $\alpha \in \mathbb{N}_0^{n-1}$. We denote by ω_θ the center of θ. Given a function g, we perform a Fourier series decomposition to the function $g\psi_\theta$ on the region $\frac{11}{9}\theta$ and obtain

$$g(w)\psi_\theta(w) \cdot \mathbb{1}_{\frac{11}{10}\theta}(\omega) = \left(\frac{r^{1/2}}{2\pi} \right)^{n-1} \sum_{v \in r^{1/2}\mathbb{Z}^{n-1}} (g\psi_\theta)^\wedge(v) e^{2\pi i v \cdot w} \mathbb{1}_{\frac{11}{10}\theta}(\omega).$$

Let $\widetilde{\psi}_\theta$ be a non-negative smooth cutoff function supported on $\frac{11}{9}\theta$ and equal to 1 on $\frac{11}{10}\theta$. We can therefore write

$$g(w)\psi_\theta(w) \cdot \widetilde{\psi}_\theta(\omega) = \left(\frac{r^{1/2}}{2\pi} \right)^{n-1} \sum_{v \in r^{1/2}\mathbb{Z}^{n-1}} (g\psi_\theta)^\wedge(v) e^{2\pi i v \cdot w} \widetilde{\psi}_\theta(\omega)$$

If we also define

$$g_{\theta,v}(w) := \left(\frac{r^{1/2}}{2\pi}\right)^{n-1} (g\psi_\theta)^\wedge(v) e^{2\pi i v \cdot \omega} \widetilde{\psi}_\theta(\omega)$$

then we have

$$g = \sum_{(\theta,v) \in \Theta_r \times r^{1/2}\mathbb{Z}^{n-1}} g_{\theta,v}$$

For $\omega \in B^{n-1}$, $z' \in B^{n-1}$ and $t \in [0,1]$, let us define a function $\Phi = \Phi(z',t;\omega)$ by

$$\partial_\omega \phi(\Phi(z',t;\omega),t;\omega) = z'. \tag{4.6}$$

We refer to equation (4.6) in [10, page 275] for a discussion on the definition of Φ. For $\theta \in B^{n-1}$ and $v \in B^{n-1}$, define the curve $\gamma^1_{\theta,v} : [0,1] \to \mathbb{R}^{n-1}$ by

$$\gamma^1_{\theta,v}(t) := \Phi(v,t;\omega_\theta), \tag{4.7}$$

where ω_θ is the center of θ. Moreover, for given (θ,v) let us define the rescaled curve

$$\gamma^\lambda_{\theta,v}(t) := \lambda \gamma^1_{\theta,v/\lambda}\left(\frac{t}{\lambda}\right). \tag{4.8}$$

Let $\Gamma^\lambda_{\theta,v}$ be the map

$$\Gamma^\lambda_{\theta,v} := (\gamma^\lambda_{\theta,v}(t), t), \tag{4.9}$$

where $t \in [0,\lambda]$. Define the curved $r^{\frac{1}{2}+\delta}$-tube as

$$T_{\theta,v} := \left\{(x,t) : |x - \gamma^\lambda_{\theta,v}(t)| \le r^{1/2+\delta}, \ t \in [0,r]\right\}. \tag{4.10}$$

The curve $\Gamma^\lambda_{\theta,v}$ is referred to as the core of $T_{\theta,v}$. This finishes our wave packet decomposition for a ball of radius r centered at the origin.

Next, let us define the wave packet decomposition for a ball not centered at the origin. Fix $\mathbf{x}_0 \in B(0,\lambda)$ and consider the ball $B(\mathbf{x}_0, r)$. For $g : B^{n-1} \to \mathbb{C}$ integrable, define

$$\tilde{g}(\omega) := e^{2\pi i \phi^\lambda(\mathbf{x}_0;\omega)} g(\omega)$$

so that

$$T^\lambda g(\mathbf{x}) = \tilde{T}^\lambda \tilde{g}(\tilde{\mathbf{x}}) \quad \text{for } \tilde{\mathbf{x}} = \mathbf{x} - \mathbf{x}_0,$$

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

$$\tilde{\phi}(\mathbf{x};\omega) := \phi\left(\mathbf{x} + \frac{\mathbf{x}_0}{\lambda};\omega\right) - \phi\left(\frac{\mathbf{x}_0}{\lambda};\omega\right) \text{ and } \tilde{a}(\mathbf{x};\omega) := a\left(\mathbf{x} + \frac{\mathbf{x}_0}{\lambda};\omega\right).$$

If $\mathbf{x} \in B(\mathbf{x}_0, r)$, then $\tilde{\mathbf{x}} \in B(0, r)$, and we can therefore apply the wave packet decomposition above to $\tilde{T}^\lambda \tilde{g}$. Moreover, notice that the core curve of $T_{\theta,v}$ is given by the collection of $\tilde{\mathbf{x}} \in B(0, r)$ satisfying

$$\partial_\omega \phi \left(\frac{\tilde{\mathbf{x}}}{\lambda} + \frac{\mathbf{x}_0}{\lambda}; \omega_\theta \right) = \frac{v}{\lambda} + \partial_\omega \phi \left(\frac{\mathbf{x}_0}{\lambda}; \omega_\theta \right). \tag{4.11}$$

Set

$$v^\lambda(\mathbf{x}_0; \omega) := \partial_\omega \phi^\lambda(\mathbf{x}_0; \omega). \tag{4.12}$$

Under this notation, the core curve of $T_{\theta,v}$ can be written as the image of the map

$$\Gamma^\lambda_{\theta, v + v^\lambda(\mathbf{x}_0; \omega_\theta)} = \left(\gamma^\lambda_{\theta, v + v^\lambda(\mathbf{x}_0; \omega_\theta)}(t), t \right) \tag{4.13}$$

with $t \in [0, \lambda]$. Define curved tubes

$$T_{\theta, v}(\mathbf{x}_0) := \mathbf{x}_0 + \left\{ (x, t) : |x - \gamma^\lambda_{\theta, v + v^\lambda(\mathbf{x}_0; \omega_\theta)}(t)| \leq r^{1/2+\delta}, \ t \in [0, r] \right\}. \tag{4.14}$$

Thus the function

$$\tilde{T}^\lambda(\tilde{g})_{\theta, v}(\mathbf{x} - \mathbf{x}_0) \tag{4.15}$$

is essentially supported on $T_{\theta, v}(\mathbf{x}_0)$ if we restrict $t \in [t_0, t_0 + r]$ where $\mathbf{x}_0 = (x_0, t_0)$. We will use $\mathbb{T}[B(\mathbf{x}_0, r)]$ to denote the collection $\{T_{\theta, v}(\mathbf{x}_0)\}_{(\theta, v)}$. Moreover, we write $\theta(T) = \theta$ for a tube $T = T_{\theta, v}(\mathbf{x}_0)$. To simplify notation, we also define

$$g_{T_{\theta, v}(\mathbf{x}_0)}(\omega) := e^{-2\pi i \phi^\lambda(\mathbf{x}_0; \omega)}(\tilde{g})_{\theta, v}(\omega). \tag{4.16}$$

Under this notation, we can write

$$T^\lambda g(\mathbf{x}) = \sum_{T \in \mathbb{T}[B(\mathbf{x}_0, r)]} T^\lambda g_T(\mathbf{x}). \tag{4.17}$$

This finishes our wave packet decomposition associated to the ball $B(\mathbf{x}_0, r)$.

4.2 Reducing to broad term estimates

We define Gauss maps and rescaled Gauss maps. Define

$$G_0(\mathbf{x}; \xi) := \partial_{\xi_1} \nabla_\mathbf{x} \phi \wedge \cdots \wedge \partial_{\xi_{n-1}} \nabla_\mathbf{x} \phi. \tag{4.18}$$

Moreover, define

$$G(\mathbf{x}; \xi) := \frac{G_0(\mathbf{x}; \xi)}{|G_0(\mathbf{x}; \xi)|}. \tag{4.19}$$

Define the rescaled Gauss map

$$G^\lambda(\mathbf{x}; \xi) := G(\mathbf{x}/\lambda; \xi). \tag{4.20}$$

Let $K \geq 1$. We divide B^{n-1} into caps τ of side length K^{-1}. Let g_τ denote $g \cdot \mathbb{1}_\tau$. For $\mathbf{x} \in B_R$, denote

$$G^\lambda(\mathbf{x}; \tau) := \{G^\lambda(\mathbf{x}; \xi) : \xi \in \tau\}. \tag{4.21}$$

Let $V \subset \mathbb{R}^n$ be a linear subspace. Let $\angle(G^\lambda(\mathbf{x}; \tau), V)$ denote the smallest angle between any non-zero vector $v \in V$ and $v' \in G^\lambda(\mathbf{x}; \tau)$. Moreover, we say that $\tau \notin_{\mathbf{x},K} V$ if $\angle(G^\lambda(\mathbf{x}; \tau), V) \geq K^{-1}$; otherwise, we say $\tau \in_{\mathbf{x},K} V$. If $\mathbf{x}$ and K are clear from the context, we often abbreviate $\tau \notin_{\mathbf{x},K} V$ to $\tau \notin V$. Next, let us introduce the notion of broad norms. Fix $B_{K^2} \subset B_R$ centered at $\mathbf{x}_0$. Define

$$\mu_{T^\lambda g}(B_{K^2}) := \min_{V_1,\ldots,V_A \in \mathrm{Gr}(k-1,n)} \left(\max_{\substack{\tau \notin V_a \\ \text{for any } 1 \leq a \leq A}} \|T^\lambda g_\tau\|^p_{L^p(B_{K^2})} \right). \tag{4.22}$$

Here $\mathrm{Gr}(k-1, n)$ is the Grassmannian of all $(k-1)$-dimensional subspaces in $\mathbb{R}^n$, and k is to be determined, and A is a parameter that is less important and its choice will become clear later. For $U \subset \mathbb{R}^n$, define

$$\|T^\lambda g\|_{\mathrm{BL}^p_{k,A}(U)} := \left(\sum_{B_{K^2}} \frac{|B_{K^2} \cap U|}{|B_{K^2}|} \mu_{T^\lambda g}(B_{K^2}) \right)^{1/p}. \tag{4.23}$$

This is called the broad part of $T^\lambda g$.

For $2 \leq k \leq n - 1$, denote

$$\Gamma_{\mathrm{HZ}}(n, k) := \frac{n-1}{3} + \frac{k-1}{6} \prod_{i=k}^{n-1} \frac{2i}{2i+1}. \tag{4.24}$$

We will prove

Theorem 4.2 (Broad norm estimate) *Let* $2n/5 \leq k \leq n/2$, *and*

$$p > p_n(k) := 2 + \frac{1}{\Gamma_{\mathrm{HZ}}(n, k) + \frac{n}{10^3}}. \tag{4.25}$$

Then for every $\epsilon > 0$, there exists A such that

$$\|T^\lambda g\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim_{K,\epsilon} R^\epsilon \|g\|^{2/p}_{L^2} \|g\|^{1-2/p}_{L^\infty}, \tag{4.26}$$

for every $K \geq 1$, $1 \leq R \leq \lambda$, where $B_R \subset B_\lambda$ is a ball of radius R. Moreover, the implicit constant depends polynomially on K.

Recall that when $\phi(\mathbf{x}; \xi) = \langle x, \xi \rangle + t|\xi|^2$, Hickman and Zahl [13] proved (4.26) for all

$$p > 2 + \frac{1}{\Gamma_{\mathrm{HZ}}(n, k)}. \tag{4.27}$$

Theorem 4.2 provides a slight improvement in this case. When k is outside the range $[2n/5, n/2]$, we also obtain improved broad norm estimates. As they are irrelevant for the asymptotic formula in (1.20), we do not state them here.

It is standard in the literature to reduce Theorem 4.1 to Theorem 4.2. For instance, by Proposition 11.1 in Guth-Hickman-Iliopoulou's work [10], if

$$2 + \frac{4}{2n - k} \leq p \leq 2 + \frac{2}{k - 2},$$ (4.28)

then Theorem 4.2 (for some fixed k) implies Theorem 4.1 for the same p. To see the asymptotic formula in (1.20), we set $k = \nu n$ and

$$p_n(k) = 2 + \frac{4}{2n - k},$$ (4.29)

solve for ν, and then obtain

$$q_{n,2} = 2 + \frac{4}{2 - \nu} \frac{1}{n} + O(n^{-2}).$$ (4.30)

We refer to the appendix of [13] on how to control $\Gamma_{\mathrm{HZ}}(n, k)$.

When proving Theorem 4.2, we will apply a wave packet decomposition

$$g = \sum_{T \in \mathbb{T}[B_R]} g_T.$$ (4.31)

By pigeonholing, we can assume that $\left\| g_T \right\|_2 \simeq \left\| g_{T'} \right\|_2$ for every T and T'. [3]

5 Polynomial partitioning

5.1 Preparatory work

In this subsection, we state a few definitions that will be useful in the forthcoming polynomial partitioning algorithms.

Definition 5.1 (Transverse complete intersection) Let $P_1, \ldots, P_{n-m} : \mathbb{R}^n \to \mathbb{R}$ be polynomials. We consider the common zero set

$$Z(P_1, \ldots, P_{n-m}) := \left\{ x \in \mathbb{R}^n : P_1(x) = \cdots = P_{n-m}(x) = 0 \right\}.$$ (5.1)

Suppose that for all $z \in Z(P_1, \ldots, P_{n-m})$, one has

$$\bigwedge_{j=1}^{n-m} \nabla P_j(z) \neq 0.$$

---

[3]This will be used in a counting lemma below (Lemma 7.10).

Then a connected branch of this set, or a union of connected branches of this set, is called an m-dimensional transverse complete intersection. Given a set Z of the form (5.1), the degree of Z is defined by

$$\min \left(\prod_{i=1}^{n-m} \deg (P_i) \right)$$

where the minimum is taken over all possible representations of $Z = Z(P_1, \ldots, P_{n-m})$.

Definition 5.2 (Tangent tubes) Recall the parameters in (1.44). Let $r \geqslant 1$ and Z be an m-dimensional transverse complete intersection. A tube $T_{\theta,v}(\mathbf{x}_0) \in \mathbb{T}[B(\mathbf{x}_0, r)]$ is said to be $r^{-1/2+\delta_m}$-tangent to Z in $B(\mathbf{x}_0, r)$ if it satisfies

1. $T_{\theta,v}(\mathbf{x}_0) \subset \mathcal{N}_{r^{1/2+\delta_m}}(Z) \cap B(\mathbf{x}_0, r)$;
2. For every $\mathbf{z} \in Z \cap B(\mathbf{x}_0, r)$, if $\mathbf{y} \in T_{\theta,v}(\mathbf{x}_0)$ with $|\mathbf{z} - \mathbf{y}| \lesssim r^{1/2+\delta_m}$, then one has

$$\angle \left(G^\lambda(\mathbf{y}; w_\theta), T_{\mathbf{z}} Z \right) \lesssim r^{-1/2+\delta_m}.$$

Here, $T_{\mathbf{z}} Z$ is the tangent space of Z at $\mathbf{z}$.

Definition 5.3 Given a function $f : B^{n-1} \to \mathbb{C}$, we say that it is concentrated on wave packets from $\mathbb{W} \subset \mathbb{T}[B_R]$ if

$$\Big\| \sum_{T \notin \mathbb{W}} f_T \Big\|_\infty \lesssim R^{-100n} \| f \|_2. \tag{5.2}$$

Here R is from Theorem 4.2.

5.2 Partitioning algorithms: part I

In this subsection, we run the first part of the polynomial partitioning algorithm. It is a variant of the algorithm in Hickman-Rogers' work [12] with two main differences.

The first difference is that, after reaching an algebraic dominant case, we will not compare contributions from the tangential case and the transverse case, but instead keep both terms and continue to do polynomial partitioning for both terms. This will be needed in Sect. 7 when we construct brooms.

Let us explain the second difference. In the first algorithm in [12, page 247], the authors there did not need to control how fast cells shrink. In other words, each time when they see a cellular case, they simply decrease the radius parameter ρ_j by a factor of 2 (see for instance equation (31) in [12, page 253]). If in the current paper we simply repeat their algorithm, then we will not have good control of the non-admissible parameters $D_n, D_{n-1}, \ldots$ by R (see Lemma 5.10 below), which was not needed in [12] and is crucial in our inductive argument in Sect. 9.

In order to control how fast cells shrink, in Lemma 5.4 we require cells to have diameter at most r/d, instead of $r/2$. This change also brings in changes in how the algorithm runs. For instance, in the last equation in [12, page 255], the authors there

simply let $d^{-\delta}$ absorb the constant 2 from the above $r/2$, and this steps needs to be done more carefully in the current paper as $d^{-\delta}$ certainly can not absorb the factor d.[4]

By pigeonholing, we can find a collection $\mathbb{B}_{K^2}$ of balls of radius K^2 such that

$$\frac{1}{2}\left\|T^\lambda g\right\|_{\mathrm{BL}^p_{k,A}(B'_{K^2})} \le \left\|T^\lambda g\right\|_{\mathrm{BL}^p_{k,A}(B_{K^2})} \le 2\left\|T^\lambda g\right\|_{\mathrm{BL}^p_{k,A}(B'_{K^2})} \tag{5.3}$$

for two arbitrary $B_{K^2}, B'_{K^2} \in \mathbb{B}_{K^2}$ and

$$\left\|T^\lambda g\right\|^p_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim (\log R)^{10} \sum_{B_{K^2}\in\mathbb{B}_{K^2}} \left\|T^\lambda g\right\|^p_{\mathrm{BL}^p_{k,A}(B_{K^2})}. \tag{5.4}$$

Denote

$$\mathfrak{Y} = \bigcup_{B_{K^2}\in\mathbb{B}_{K^2}} B_{K^2}. \tag{5.5}$$

Next, we apply polynomial partitioning to $T^\lambda g$ restricted to $\mathfrak{Y}$. For a polynomial $P:\mathbb{R}^n \to \mathbb{R}$, we let $Z(P) := \{z \in \mathbb{R}^n : P(z) = 0\}$ and let $\mathrm{cell}(P)$ denote the set of connected components of $\mathbb{R}^n \setminus Z(P)$.

Lemma 5.4 (Polynomial partitioning, Guth [9], Hickman and Rogers [12]) *Fix $r \gg 1, d \in \mathbb{N}$ and suppose $F \in L^1(\mathbb{R}^n)$ is non-negative and supported on $B_r \cap \mathcal{N}_{r^{1/2+\delta_\circ}}(Z)$ for some $0 < \delta_\circ \ll 1$, where Z is an m-dimensional transverse complete intersection of degree at most d. At least one of the following cases holds*:

Cellular case. There exists a polynomial $P:\mathbb{R}^n \to \mathbb{R}$ of degree $O(d)$ with the following properties:

(1) *$\#\mathrm{cell}(P) \simeq d^m$ and each $O' \in \mathrm{cell}(P)$ has diameter at most r/d.*
(2) *If we define*

$$\mathcal{O} := \{O' \setminus \mathcal{N}_{r^{1/2+\delta_\circ}}(Z(P)) : O' \in \mathrm{cell}(P)\}, \tag{5.6}$$

  then

$$\int_O F \simeq d^{-m} \int_{\mathbb{R}^n} F \quad \text{for all } O \in \mathcal{O}. \tag{5.7}$$

Algebraic case. There exists an $(m-1)$-dimensional transverse complete intersection Y of degree at most $O(d)$ such that

$$\int_{B_r\cap\mathcal{N}_{r^{1/2+\delta_\circ}}(Z)} F \lesssim \int_{B_r\cap\mathcal{N}_{r^{1/2+\delta_\circ}}(Y)} F$$

Here the diameter of a cell in Lemma 5.4 is $O(r/d)$ instead of $O(r/2)$ as in [9] and [12]. See the proof sketch of Theorem 2.12 in [21] for a discussion.

---

[4] This will be addressed at the end of Sect. 5.4.

We now start our polynomial partitioning algorithm. This algorithm will produce a tree consisting of many nodes. Each node will have no child (algorithm for that node stops), one child (cellular case) or two children (algebraic case).

Let $\mathfrak{n}$ be a node. We apply Lemma 5.4 to $\mathfrak{n}$, and see whether we are in the cellular case or the algebraic case. If we are in the cellular case, then $\mathfrak{n}_j$ has only one child, which will be denoted as $\mathfrak{n}_{j+1,L} = \mathfrak{n}_{j+1,L}(\mathfrak{n}_j)$; here "$L$" refers to "left" and $\mathfrak{n}_{j+1,L}$ is called the L-child of $\mathfrak{n}_j$. If we are the algebraic case, then $\mathfrak{n}_j$ will have two children, which will be denoted as $\mathfrak{n}_{j+1,M} = \mathfrak{n}_{j+1,M}(\mathfrak{n}_j)$ and $\mathfrak{n}_{j+1,R} = \mathfrak{n}_{j+1,R}(\mathfrak{n}_j)$; here "$M$" refers to "middle" and $\mathfrak{n}_{j+1,M}$ is called the M-child of $\mathfrak{n}_j$, and "R" refers to "right" and $\mathfrak{n}_{j+1,R}$ is called the R-child of $\mathfrak{n}_j$. [5]

For two nodes $\mathfrak{n}$ and $\mathfrak{n}'$, if $\mathfrak{n}$ is a descendant of $\mathfrak{n}'$, then we write $\mathfrak{n} \preccurlyeq \mathfrak{n}'$; similarly we define $\succcurlyeq$. Here we make the convention that $\mathfrak{n} \preccurlyeq \mathfrak{n}$ and $\mathfrak{n} \succcurlyeq \mathfrak{n}$. Recall the parameters in (1.44). Moreover, define $\tilde{\delta}_{m-1}$ by

$$(1 - \tilde{\delta}_{m-1})(1/2 + \delta_{m-1}) = 1/2 + \delta_m. \tag{5.8}$$

Note that $\delta_{m-1}/2 \le \tilde{\delta}_{m-1} \le 2\delta_{m-1}$.

Step 0. In this step, we create the root of the tree. Denote

$$\mathfrak{n}_0 = \{O_{i_0}\}, \tag{5.9}$$

with $O_{i_0} = B_R \cap \mathfrak{Y}$. Moreover, define $\dim(\mathfrak{n}_0) = n$, $\rho(\mathfrak{n}_0) = R$ and $\mathrm{j}(\mathfrak{n}_0) = 0$. Later we will define j for every node. It will play the role of the parameter j in the recursive step of the first algorithm in [12, page 247]. In the end of this step, define

$$\#_a(\mathrm{j}(\mathfrak{n}_0)) = 0, \quad \#_c(\mathrm{j}(\mathfrak{n}_0)) = 0. \tag{5.10}$$

Here "a" is short for "algebraic" and "c" is short of "cellular", and we use $\#_a$ to record the number of algebraic cases we have so far when applying Lemma 5.4, and similarly, we use $\#_c$ to record the number of cellular cases. Initialize

$$\mathfrak{M}'_{\ell'} = \emptyset, \quad \forall \ell' \in \mathbb{N}. \tag{5.11}$$

This collection will appear at almost the end of the algorithm; we will keep adding elements to it as we run the algorithm.

Step 1. Creating nodes at the first level. The root node $\mathfrak{n}_0$ has either one or two children, given as follows. Apply Lemma 5.4 to the function $T^\lambda g \cdot \mathbb{1}_{O_{i_0}}$, we obtain a collection of cells $\{O_{i_1}\}_{i_1}$ (as in (5.6)) and a wall $W = \mathcal{N}_{(\rho(\mathfrak{n}_0))^{1/2+\delta_m}}(Z)$ with $m = \dim(\mathfrak{n}_0)$ for some variety Z of dimension $m-1$. We without loss of generality assume that all these regions are unions of balls B_{K^2}. Compare

$$\#\{B_{K^2} \in \mathbb{B}_{K^2} : B_{K^2} \subset \bigcup_{i_1} O_{i_1}\} \text{ and } \#\{B_{K^2} \in \mathbb{B}_{K^2} : B_{K^2} \subset W\}. \tag{5.12}$$

If the former term in (5.12) is larger, then we say that we are in the cellular case of this step, and otherwise we say that we are in the algebraic case. In the cellular

---

[5]For readers familiar with the polynomial partitioning algorithms in Guth [9] or Hickman and Rogers [12]: Here $\mathfrak{n}_M$ collects transverse contributions, and $\mathfrak{n}_R$ collects tangential contributions.

case, the node $\mathfrak{n}_0$ has only one child (defined in (5.13)-(5.14)), and will be denoted by $\mathfrak{n}_{1,L} = \mathfrak{n}_{1,L}(\mathfrak{n}_0)$ and called the L-child; here L refers to "left". Define

$$\rho(\mathfrak{n}_{1,L}) = \rho(\mathfrak{n}_0)/d, \quad \mathfrak{n}_{1,L} = \{O_{i_1}\}_{i_1}, \quad \dim(\mathfrak{n}_{1,L}) = m, \tag{5.13}$$

and

$$j(\mathfrak{n}_{1,L}) = j(\mathfrak{n}_0) + 1, \tag{5.14}$$

$$\#_c(j(\mathfrak{n}_{1,L})) = \#_c(j(\mathfrak{n}_0)) + 1, \quad \#_a(j(\mathfrak{n}_{1,L})) = \#_a(j(\mathfrak{n}_0)). \tag{5.15}$$

In other words, we have had one cellular cases so far, and zero algebraic cases. Define $\mathfrak{L}_1 = \{\mathfrak{n}_{1,L}\}$, that is, we use $\mathfrak{L}_1$ to collect all the L-children in this step. Moreover, set $\mathfrak{M}_1 = \mathfrak{R}_1 = \emptyset$. This finishes defining the node $\mathfrak{n}_{1,L}$ and its information.

If we are in the algebraic case, then the node $\mathfrak{n}_0$ has two children, which will be denoted by $\mathfrak{n}_{1,M} = \mathfrak{n}_{1,M}(\mathfrak{n}_0)$ and $\mathfrak{n}_{1,R} = \mathfrak{n}_{1,R}(\mathfrak{n}_0)$ and be called the M-child and the R-child; here M refers to "middle", and R to "right". Define

$$\rho(\mathfrak{n}_{1,M}) = \rho(\mathfrak{n}_{1,R}) = \rho(\mathfrak{n}_0)^{1-\tilde{\delta}_{m-1}}, \tag{5.16}$$

and

$$\dim(\mathfrak{n}_{1,M}) = n, \quad \dim(\mathfrak{n}_{1,R}) = n - 1, \quad \mathfrak{n}_{1,M} = \mathfrak{n}_{1,R} = \{O_{i_1'}\}_{i_1'}, \tag{5.17}$$

where each $O_{i_1'}$ is given by $W \cap B_{\rho(\mathfrak{n}_{1,M})}$ and we let $B_{\rho(\mathfrak{n}_{1,M})}$ run through a collection of finitely overlapping balls of radius $\rho(\mathfrak{n}_{1,M})$ inside $B_{\rho(\mathfrak{n}_0)}$. Moreover, define

$$j(\mathfrak{n}_{1,M}) = j(\mathfrak{n}_0) + 1, \tag{5.18}$$

$$\#_c(j(\mathfrak{n}_{1,L})) = \#_c(j(\mathfrak{n}_0)), \quad \#_a(j(\mathfrak{n}_{1,L})) = \#_a(j(\mathfrak{n}_0)) + 1, \tag{5.19}$$

and

$$j(\mathfrak{n}_{1,R}) = 0, \quad \#_c(j(\mathfrak{n}_{1,R})) = \#_a(j(\mathfrak{n}_{1,R})) = 0. \tag{5.20}$$

Here let us explain the rule of defining j: It is always reset to be 0 whenever we see an R-child, and otherwise its values is increased by 1.

Let $\mathfrak{M}_1, \mathfrak{R}_1$ collect all the M-children and R-children at Step 1, respectively. As $\mathfrak{n}_0$ has no L-child and we always use $\mathfrak{L}_1$ to collection L-children, we therefore set $\mathfrak{L}_1 = \emptyset$. This finishes the first step.

Step 2. Creating nodes at the second level. Take a node $\mathfrak{n}_1$ from the previous step. It has one or two children.

If $\mathfrak{n}_1 \in \mathfrak{L}_1$ or $\mathfrak{M}_1$, then its children, which will be named either as $\mathfrak{n}_{2,L}$ or as $\mathfrak{n}_{2,M}, \mathfrak{n}_{2,R}$, will be given as follows. For each $O_{i_1} \in \mathfrak{n}_1$, we apply Lemma 5.4 with dimension parameter $\dim(\mathfrak{n}_1)$, and obtain a collection of cells $\{O_{i_2}\}_{i_2}$ and a wall $W_{m-1} = \mathcal{N}_{(\rho(\mathfrak{n}_1))^{1/2+\delta_m}}(Z_{m-1})$ with $m = \dim(\mathfrak{n}_1)$ for some variety Z_{m-1} of dimension $m - 1$. We make a comparison similar to (5.12). If we are in the cellular case, then define

$$\rho(\mathfrak{n}_{2,L}) = \rho(\mathfrak{n}_1)/d, \quad \mathfrak{n}_{2,L} = \bigcup_{O_{i_1} \in \mathfrak{n}_1} \{O_{i_2}\}_{i_2}, \quad \dim(\mathfrak{n}_{2,L}) = \dim(\mathfrak{n}_1), \tag{5.21}$$

and

$$j(\mathfrak{n}_{2,L}) = j(\mathfrak{n}_1) + 1, \tag{5.22}$$

$$\#_c(j(\mathfrak{n}_{2,L})) = \#_c(j(\mathfrak{n}_1)) + 1, \quad \#_a(j(\mathfrak{n}_{2,L})) = \#_a(j(\mathfrak{n}_1)). \tag{5.23}$$

If we are in the algebraic case, then define

$$\rho(\mathfrak{n}_{2,M}) = \rho(\mathfrak{n}_{2,R}) = \rho(\mathfrak{n}_1)^{1-\tilde{\delta}_{m-1}}, \tag{5.24}$$

and

$$\dim(\mathfrak{n}_{2,M}) = \dim(\mathfrak{n}_1), \quad \dim(\mathfrak{n}_{2,R}) = \dim(\mathfrak{n}_1) - 1, \tag{5.25}$$

$$\mathfrak{n}_{2,M} = \mathfrak{n}_{2,R} = \bigcup_{O_{i_1} \in \mathfrak{n}_1} \{O_{i_2'}\}_{i_2'}, \tag{5.26}$$

where each $O_{i_2'}$ is given by $W_{m-1} \cap B_{\rho(\mathfrak{n}_{2,M})}$ and we let $B_{\rho(\mathfrak{n}_{2,M})}$ run through all balls of radius $\rho(\mathfrak{n}_{2,M})$ inside $B_{\rho(\mathfrak{n}_1)}$. Moreover, define

$$j(\mathfrak{n}_{2,M}) = j(\mathfrak{n}_1) + 1, \tag{5.27}$$

$$\#_c(j(\mathfrak{n}_{2,L})) = \#_c(j(\mathfrak{n}_1)), \quad \#_a(j(\mathfrak{n}_{2,L})) = \#_a(j(\mathfrak{n}_1)) + 1, \tag{5.28}$$

and

$$j(\mathfrak{n}_{2,R}) = 0, \quad \#_c(j(\mathfrak{n}_{2,R})) = \#_a(j(\mathfrak{n}_{2,R})) = 0. \tag{5.29}$$

Let $\mathfrak{L}_2$ collect all the L-children at Step 2, and similarly, we define $\mathfrak{M}_2$ and $\mathfrak{R}_2$.

Next, consider the remaining case $\mathfrak{n}_1 \in \mathfrak{R}_1$. This step is quite similar to the above case where $\mathfrak{n}_1 \in \mathfrak{L}_1$ or $\mathfrak{M}_1$, with the main difference in how $\mathfrak{n}_{2,M}$ will be defined.[6] The children of $\mathfrak{n}_1$ are given as follows. For each $O_{i_1} \in \mathfrak{n}_1$, we apply Lemma 5.4 with dimension parameter $m = \dim(\mathfrak{n}_1)$ and $\delta_\circ = \delta_m$, and obtain a collection of cells $\{O_{i_2}\}_{i_2}$ and a wall $W_{m-1} = \mathcal{N}_{(\rho(\mathfrak{n}_1))^{1/2+\delta_m}}(Z_{m-1})$ for some variety Z_{m-1} of dimension $m - 1$. If we are in the cellular case, then define $\rho(\mathfrak{n}_{2,L}), \mathfrak{n}_{2,L}, \dim(\mathfrak{n}_{2,L})$ in the same way as in (5.21) and $j(\mathfrak{n}_{2,L}), \#_c(j(\mathfrak{n}_{2,L})), \#_a(j(\mathfrak{n}_{2,L}))$ in the same way as in (5.22). If we are in the algebraic case, then define

$$\rho(\mathfrak{n}_{2,M}) = \rho(\mathfrak{n}_{2,R}) = \rho(\mathfrak{n}_1)^{1-\tilde{\delta}_{m-1}}, \tag{5.30}$$

and

$$\dim(\mathfrak{n}_{2,M}) = \dim(\mathfrak{n}_1), \quad \dim(\mathfrak{n}_{2,R}) = \dim(\mathfrak{n}_1) - 1, \tag{5.31}$$

$$\mathfrak{n}_{2,R} = \bigcup_{O_{i_1} \in \mathfrak{n}_1} \{O_{i_2'}\}_{i_2'}, \tag{5.32}$$

---

[6]The difference comes from the fact that the dimension parameter $\dim(\mathfrak{n}_1)$ drops by one whenever we meet an R-child, which of course means that we are in the algebraic case when applying Lemma 5.4. In this case, we need to cut walls into thinner layers to define new nodes, see (5.35).

where each $O_{i'_2}$ is given by $W_{m-1} \cap B_{\rho(\mathfrak{n}_{2,M})}$ and we let $B_{\rho(\mathfrak{n}_{2,M})}$ run through a collection of finitely overlapping balls of radius $\rho(\mathfrak{n}_{2,M})$ that intersect O_{i_1}. Our choice of parameters guarantees that

$$\rho(\mathfrak{n}_{2,R})^{1/2+\delta_{m-1}} = \rho(\mathfrak{n}_1)^{1/2+\delta_m}. \tag{5.33}$$

We still need to define $\mathfrak{n}_{2,M}$. Roughly speaking, for each $O_{i'_2}$ given by $W_{m-1} \cap B_{\rho(\mathfrak{n}_{2,M})}$, which is of thickness $\rho(\mathfrak{n}_1)^{1/2+\delta_m}$, we will cut it into thinner layers $W_{m-1,b}$ of thickness $\rho(\mathfrak{n}_{2,M})^{1/2+\delta_m}$. Then we set

$$\mathfrak{n}_{2,M} = \bigcup_{O_{i_1} \in \mathfrak{n}_1} \bigcup_{O_{i'_2}} \{O_{i'_2} \cap W_{m-1,b}\}_b. \tag{5.34}$$

To make this precise, we follow Hickman-Rogers' treatment [12]. For each $B_{\rho(\mathfrak{n}_{2,M})}$, we follow page 258 in [12], find a finite set of translates $\mathfrak{B} \subset B(0, \rho(\mathfrak{n}_1)^{1/2+\delta_m})$ and then set (following the last equation in [12, page 258])

$$\mathfrak{n}_{2,M} = \bigcup_{O_{i_1} \in \mathfrak{n}_1} \bigcup_{O_{i'_2}} \{O_{i'_2} \cap \mathcal{N}_{\rho(\mathfrak{n}_{2,M})^{1/2+\delta_m}}(Z_{m-1}+b) : b \in \mathfrak{B}\}. \tag{5.35}$$

In the end, define $\mathfrak{j}(\mathfrak{n}_{2,M})$, $\mathfrak{j}(\mathfrak{n}_{2,R})$, $\#_a$, $\#_c$ in the same way as in (5.27)–(5.29).

Let $\mathfrak{L}_2$ be the collection of all the L-children at Step 2, and similarly, we define $\mathfrak{M}_2$ and $\mathfrak{R}_2$. This finishes Step 2.

Step ℓ. Creating nodes at the ℓ-th level. How we proceed in a general step is similar to what we did in Step 2, with one difference mentioned at the beginning of this section that we need to control how fast cells shrink. We will sketch the part in this step that is similar to Step 2, and explain in more details the difference. [7]

Take a node $\mathfrak{n}_{\ell-1}$ from the previous step. There are a few parameters associated to it: A dimension parameter $\dim(\mathfrak{n}_{\ell-1}) =: m$, a radius parameter $\rho(\mathfrak{n}_{\ell-1})$, the parameters $\mathfrak{j}(\mathfrak{n}_{\ell-1})$, $\#_c(\mathfrak{j}(\mathfrak{n}_{\ell-1}))$ and $\#_a(\mathfrak{j}(\mathfrak{n}_{\ell-1}))$ satisfying

$$\mathfrak{j}(\mathfrak{n}_{\ell-1}) = \#_c(\mathfrak{j}(\mathfrak{n}_{\ell-1})) + \#_a(\mathfrak{j}(\mathfrak{n}_{\ell-1})). \tag{5.36}$$

Before we proceed, we need to introduce new notation. Let $\mathfrak{n}^{\uparrow}_{\ell-1}$ denote the closest ancestor (itself included) to $\mathfrak{n}_{\ell-1}$ that belongs to $\mathfrak{M}_{\ell'}$, $\mathfrak{M}'_{\ell'}$ or $\mathfrak{R}_{\ell'}$ for some ℓ'. (Recall the initialization of $\mathfrak{M}'_{\ell'}$ in (5.11).) Note that

$$\dim(\mathfrak{n}_{\ell-1}) = \dim(\mathfrak{n}^{\uparrow}_{\ell-1}). \tag{5.37}$$

For each $O_{i_{\ell-1}} \in \mathfrak{n}_{\ell-1}$, we apply Lemma 5.4 with dimension parameter m and $\delta_\circ = \delta_\circ(\mathfrak{n}_{\ell-1})$ satisfying

$$\rho(\mathfrak{n}_{\ell-1})^{\frac{1}{2}+\delta_\circ} = \rho(\mathfrak{n}^{\uparrow}_{\ell-1})^{\frac{1}{2}+\delta_m} \tag{5.38}$$

---

[7]This is the second difference explained at the beginning of Sect. 5.2: We need to apply (for a second time) the transverse equidistribution properties in [10, Lemma 8.4], see (5.44)-(5.50). In (5.115)-(5.116), it will become clear why this step is necessary.

and obtain a collection of cells $\{O_{i_\ell}\}_{i_\ell}$ and a wall

$$W_{m-1} = \mathcal{N}_{(\rho(\mathfrak{n}_{\ell-1}^{\uparrow}))^{1/2+\delta_m}} (Z_{m-1}) \tag{5.39}$$

for some variety Z_{m-1}. If we are in the algebraic case, then $\mathfrak{n}_{\ell-1}$ has two children, called $\mathfrak{n}_{\ell,M}$ and $\mathfrak{n}_{\ell,R}$, and similarly to (5.30)–(5.35), we define

$$\rho(\mathfrak{n}_{\ell,M}) = \rho(\mathfrak{n}_{\ell,R}) = \rho(\mathfrak{n}_{\ell-1})^{1-\tilde{\delta}_{m-1}}, \quad \mathfrak{n}_{\ell,R} = \bigcup_{O_{i_{\ell-1}} \in \mathfrak{n}_{\ell-1}} \{O_{i_\ell'}\}_{i_\ell'}; \tag{5.40}$$

moreover, define

$$\mathfrak{n}_{\ell,M} = \bigcup_{O_{i_{\ell-1}} \in \mathfrak{n}_{\ell-1}} \bigcup_{O_{i_\ell'}} \{O_{i_\ell'} \cap \mathcal{N}_{\rho(\mathfrak{n}_{\ell,M})^{1/2+\delta_m}} (Z_{m-1} + b) : b \in \mathfrak{B}\}, \tag{5.41}$$

where $\mathfrak{B}$ is a finite set of points in $B(0, \rho(\mathfrak{n}_{\ell-1})^{1/2+\delta_\circ})$, and

$$\dim(\mathfrak{n}_{\ell,M}) = \dim(\mathfrak{n}_{\ell-1}), \quad \dim(\mathfrak{n}_{\ell,R}) = \dim(\mathfrak{n}_{\ell-1}) - 1. \tag{5.42}$$

If we are in the cellular case, then we proceed differently. [8] There are two further cases. If we are in the case

$$\frac{\rho(\mathfrak{n}_{\ell-1})}{d} \geq \rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1-\delta_{m-1/2}}, \tag{5.43}$$

where $\delta_{m-1/2}$ is as in (1.45), then we define $\mathfrak{n}_{\ell,L}$, $\rho(\mathfrak{n}_{\ell,L})$ and $\dim(\mathfrak{n}_{\ell,L})$ in the same way as in (5.21), and $\mathfrak{j}(\mathfrak{n}_{\ell,\mathfrak{L}})$, $\#_c(\mathfrak{j}(\mathfrak{n}_{\ell,\mathfrak{L}}))$ and $\#_a(\mathfrak{j}(\mathfrak{n}_{\ell,\mathfrak{L}}))$ in the same way as in (5.22). If (5.43) is violated, then we first update

$$\mathfrak{M}_\ell' = \mathfrak{M}_\ell' \bigcup \{\mathfrak{n}_{\ell,L}(\mathfrak{n}_{\ell-1})\}. \tag{5.44}$$

The next step is to cut each O_{i_ℓ} into thinner layers, in a way that is essentially the same as in (5.35). Let us be more precise. By the way we run the partitioning algorithm, in particular, due to the choice of the parameter $\delta_\circ$ in (5.38), we know that

$$\rho(\mathfrak{n})^{\frac{1}{2}+\delta_\circ(\mathfrak{n})} = \rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{\frac{1}{2}+\delta_m} \tag{5.45}$$

for every node $\mathfrak{n}$ with $\mathfrak{n}_{\ell-1} \preccurlyeq \mathfrak{n} \preccurlyeq \mathfrak{n}_{\ell-1}^{\uparrow}$. Therefore we have

$$O_{i_\ell} \subset B_{\rho(\mathfrak{n}_{\ell,L})} \cap \mathcal{N}_{\rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1/2+\delta_m}} (Z_m), \tag{5.46}$$

where

$$\rho(\mathfrak{n}_{\ell,L}) := \rho(\mathfrak{n}_{\ell-1})/d, \tag{5.47}$$

---

[8]Indeed Step 2 also falls into the same framework; the forthcoming difference only occurs when ℓ is large.

and similarly as before we define $\rho(\mathfrak{n}_{\ell,L})$ before defining $\mathfrak{n}_{\ell,L}$, and Z_m is an m-dimensional variety. Note that as (5.43) is violated, we have

$$\rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1-\delta_{m-1/2}}/d \leq \rho(\mathfrak{n}_{\ell,L}) \leq \rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1-\delta_{m-1/2}}. \tag{5.48}$$

We will cut the right hand side of (5.46) into thinner layers of thickness $\rho(\mathfrak{n}_{\ell,L})^{1/2+\delta_m}$ and apply transverse equidistribution properties (for instance Lemma 8.4 in [10]). This can be done in exactly the same way as in (5.35) and (5.41), with the only difference that the parameter $\tilde{\delta}_{m-1}$ that appears in the radius $\rho(\mathfrak{n}_{2,M})$ is replaced by $\delta_{m-1/2}$ (which appears in the radius $\rho(\mathfrak{n}_{\ell,L})$ because of the relation (5.48)). Therefore, we follow [12, page 258] and find a finite set of translates $\mathfrak{B} \subset B(0, \rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1/2+\delta_m})$, and then set

$$\mathfrak{n}_{\ell,L} = \bigcup_{O_{i_{\ell-1}} \in \mathfrak{n}_{\ell-1}} \bigcup_{O_{i_\ell'}} \{O_{i_\ell'} \cap \mathcal{N}_{\rho(\mathfrak{n}_{\ell,L})^{1/2+\delta_m}}(Z_{m-1} + b) : b \in \mathfrak{B}\}. \tag{5.49}$$

Moreover, define

$$\dim(\mathfrak{n}_{\ell,L}) = \dim(\mathfrak{n}_{\ell-1}), \quad \mathrm{j}(\mathfrak{n}_{\ell,L}) = \mathrm{j}(\mathfrak{n}_{\ell-1}) + 1, \tag{5.50}$$

$$\#_c(\mathrm{j}(\mathfrak{n}_{\ell,L})) = \#_c(\mathfrak{n}_{\ell-1}) + 1, \quad \#_a(\mathrm{j}(\mathfrak{n}_{\ell,L})) = \#_a(\mathfrak{n}_{\ell-1}). \tag{5.51}$$

In the end, we let $\mathfrak{L}_\ell$ collect all the L-children at the ℓ-th level, and similarly we let $\mathfrak{M}_\ell$ collect all the M-children and $\mathfrak{R}_\ell$ collect all the R-children. This finishes the ℓ-th step of the algorithm.

Before we proceed to the next step, we make a remark on the size of the parameter $\delta_\circ$ in (5.38). By (5.38) and (5.43), we have

$$\rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{\frac{1+2\delta_m}{1+2\delta_\circ}} = \rho(\mathfrak{n}_{\ell-1}) \geq d\rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{1-\delta_{m-1/2}}, \tag{5.52}$$

which further implies

$$\delta_m \leq \delta_\circ \leq \delta_{m-1/2}. \tag{5.53}$$

In other words, $\delta_\circ$ is still quite close to δ_m, and is very far from δ_{m-1}.

Remark 5.5 In (5.40), each element $O_{i_\ell'}$ in $\mathfrak{n}_{\ell,R}$ is given by $B_{\rho(\mathfrak{n}_{\ell,R})} \cap W_{m-1}$, where $B_{\rho(\mathfrak{n}_{\ell,R})}$ is a ball of radius $\rho(\mathfrak{n}_{\ell,R})$ and W_{m-1} is given in (5.39). Their counterpart in [12] is given by $B \cap \mathcal{N}_{\rho_j^{1/2+\delta_m}}(\mathbf{Y})$ at the bottom of [12, page 256], where B is a ball of radius ρ_{j+1} and $\rho_{j+1} = \rho(\mathfrak{n}_{\ell,R})$ in our notation, $\rho_j = \rho(\mathfrak{n}_{\ell-1})$, $\mathbf{Y} = W_{m-1}$. The slight difference is that the neighborhood scale $(\rho(\mathfrak{n}_{\ell-1}^{\uparrow}))^{1/2+\delta_m}$ in (5.39) is bigger than $\rho_j^{1/2+\delta_m}$. However, by (5.43),

$$\frac{\rho(\mathfrak{n}_{\ell-1}^{\uparrow})}{\rho(\mathfrak{n}_{\ell-1})} \leq \rho(\mathfrak{n}_{\ell-1}^{\uparrow})^{\delta_{m-1/2}}. \tag{5.54}$$

We will lose about δ_{m-1}^{-1} many of these multiplicative factors. As $\delta_{m-1/2} \ll_\epsilon \delta_{m-1}$, we see that they are harmless.

Stopping condition. Suppose we have arrived at the ℓ_0-th level. Take a node $\mathfrak{n}_{\ell_0}$. The algorithm will not continue at this node (but may still continue at other nodes at the same level) if either $\rho(\mathfrak{n}_{\ell_0}) \leq R^{\delta_0}$ or $\dim(\mathfrak{n}_{\ell_0}) \leq k - 1$. Here δ_0 is given in (1.44) and k is given in Theorem 4.2. In other words, we will not continue our algorithm if the radius of the node is too small, or the dimension is too small.

We state one lemma that will be used later.

Lemma 5.6 *We have*

$$\#\Big(\bigcup_\iota (\mathfrak{M}_\iota \cup \mathfrak{R}_\iota)\Big) \lesssim_{n,\delta} 1. \tag{5.55}$$

Proof of Lemma 5.6 Note that the left hand side of (5.55) would not change if we assume that there is no cellular case in the algorithm. In this case, each node has either zero or two children. The total number of levels $\ell_0 \leq \delta^{-1}$. Moreover, note that the algorithm stops at a node $\mathfrak{n}$ if $\dim(\mathfrak{n}) \leq k - 1$. As a consequence, we see that the left hand side of (5.55) is $\leq n\delta^{-1}$. $\square$

5.3 The related case

The proof of Theorem 4.2 relies on a two-ends argument. This requires a relation, denoted by $\sim$, which is defined between tubes $T \in \mathbb{T}[B_R]$ and balls $B_\iota \subset B_R$ of radius $R^{1-\delta}$. The definition of $\sim$ is a bit complicated and relies on the definition of brooms; it will be given in Sect. 7. At this point, we only need the fact that

$$\#\{B_\iota \subset B_R : B_\iota \sim T\} \lesssim 1, \tag{5.56}$$

for every tube $T \in \mathbb{T}[B_R]$.

Definition 5.7 For each ball $B_\iota \subset B_R$ of radius $R^{1-\delta}$ and $\mathbf{x} \in B_\iota$, define

$$T^\lambda g^{\sim}(\mathbf{x}) := \sum_{T \in \mathbb{T}[B_R], T \sim B_\iota} T^\lambda g_T(\mathbf{x}), \tag{5.57}$$

and define $T^\lambda g^{\nsim}(\mathbf{x})$ to be the difference of $T^\lambda g(\mathbf{x})$ and $T^\lambda g^{\sim}(\mathbf{x})$. Moreover, for a given ι, define

$$g_\iota^{\nsim} := \sum_{T \in \mathbb{T}[B_R], T \nsim B_\iota} g_T. \tag{5.58}$$

Under the above notation, it holds that

$$T^\lambda g^{\nsim}(\mathbf{x}) = T^\lambda g_\iota^{\nsim}(\mathbf{x}), \tag{5.59}$$

whenever $\mathbf{x} \in B_\iota$. Recall the definition of $\mathbb{B}_{K^2}$ at the beginning of Sect. 5.2. Denote

$$\mathbb{B}'_{K^2} := \Big\{ B_{K^2} \in \mathbb{B}_{K^2} : \big\| T^\lambda g^{\sim} \big\|_{\mathrm{BL}^p_{k,A}(B_{K^2})} \leq \big\| T^\lambda g^{\nsim} \big\|_{\mathrm{BL}^p_{k,A}(B_{K^2})} \Big\}. \tag{5.60}$$

If $|\mathbb{B}'_{K^2}| \leq |\mathbb{B}_{K^2}|/2$, then we say that we are in the related case (of Theorem 4.2), and otherwise we say that we are in the non-related case. Because of the pigeonholing step in (5.3), if we are in the related case, then the contribution to the broad-norm $\mathrm{BL}^p_{k,A}(B_R)$ from $\mathbb{B}_{K^2} \setminus \mathbb{B}'_{K^2}$ is bigger. In this case, we can use the induction hypothesis (4.5) and the fact that there is only a small number of balls related to each tube, as in (5.56), to finish the proofs of Theorem 4.2 and Theorem 4.1.

Lemma 5.8 *If we are in the related case, then* (4.26) *holds.*

The proof of this lemma is a standard induction-on-scales argument, and is the same as that of Lemma 2.20 in [21]. We leave out the proof.

5.4 Partitioning algorithm: part II

The rest of the paper is to handle the case that

$$|\mathbb{B}'_{K^2}| \geq |\mathbb{B}_{K^2}|/2 \tag{5.61}$$

that is, the unrelated component $T^\lambda g^\sim$ dominates. Recall that we need to bound

$$\sum_{B_{K^2} \in \mathbb{B}_{K^2}} \|T^\lambda g^\sim\|^p_{\mathrm{BL}^p_{k,A}(B_{K^2})} \simeq \sum_{B_\iota} \sum_{B_{K^2} \subset B_\iota} \|T^\lambda g^\sim\|^p_{\mathrm{BL}^p_{k,A}(B_{K^2})}. \tag{5.62}$$

We run the previous algorithm again with $T^\lambda g$ replaced by $T^\lambda g^\sim$. Note that in the algorithm in Sect. 5.2, we did not compare contributions between the transverse case and the tangential case, which is exactly because we will further run the algorithm below. In what follows, we often abbreviate $g^\sim_\iota$ to g_ι.

Step 0. Define $\mathfrak{n}^*_0 = \mathfrak{n}_0$.

Step 1. We will define three quantities; they correspond to contributions from the cellular case $C(\mathfrak{L}_1)$, the transverse case $C(\mathfrak{M}_1)$ and the tangential case $C(\mathfrak{R}_1)$. Take $B_\iota \subset B_R$, a ball of radius $R^{1-\delta}$, and $\mathfrak{n}_1$, a child of $\mathfrak{n}^*_0$ with $\mathfrak{n}_1 \in \mathcal{L}_1$, take $O_{i_1} \in \mathfrak{n}_1$ with $O_{i_1} \subset B_{\rho_1} \subset B_\iota$ with $\rho_1 = \rho(\mathfrak{n}_1)$ for some $B_{\rho(\mathfrak{n}_1)}$, denote

$$g_{\iota,O_{i_1}} = \sum_{T \in \mathbb{T}[B_R], T \cap O_{i_1} \neq \emptyset} (g_\iota)_T. \tag{5.63}$$

Define

$$C(\mathfrak{L}_1) := \sum_{O_{i_1} \in \mathfrak{n}_1} \sum_\iota \|T^\lambda g_{\iota,O_{i_1}}\|^p_{\mathrm{BL}^p_{k,A}(O_{i_1} \cap B_\iota)}. \tag{5.64}$$

If $\mathfrak{n}^*_0$ does not have any children in $\mathfrak{L}_1$, then we simply set $C(\mathfrak{L})_1 = 0$.

To define the other two quantities, we need more notation. Let $\mathfrak{n}_1$ be a child of $\mathfrak{n}^*_0$ with $\mathfrak{n}_1 \in \mathfrak{M}_1$ or $\mathfrak{R}_1$, and take $O_{i'_1} \in \mathfrak{n}_1$ given by $O_{i'_1} = B_{\rho_1} \cap W$ for some $B_{\rho_1} \subset B_{\rho_0}$, with $\rho_1 = \rho(\mathfrak{n}_1)$ and $\rho_0 = \rho(\mathfrak{n}^*_0)$. Similarly to (5.63), we define $g_{\iota,O_{i'_1}}$. Let $\mathbb{T}_{O_{i'_1}}$ denote the collection of all $T \in \mathbb{T}[B_{\rho_0}]$ for which

$$T \cap B_{\rho_1} \cap W \neq \emptyset. \tag{5.65}$$

Moreover, we will partition $\mathbb{T}_{O_{i_1'}}$ into two parts

$$\mathbb{T}_{O_{i_1'}} = \mathbb{T}_{O_{i_1'},\text{tang}} \bigcup \mathbb{T}_{O_{i_1'},\text{trans}}, \tag{5.66}$$

where

$$\mathbb{T}_{O_{i_1'},\text{tang}} := \left\{ T \in \mathbb{T}_{O_{i_1'}} : T \text{ is } \rho_1^{-\frac{1}{2}+\delta_{m-1}}\text{-tangent to } Z \text{ on } B_{\rho_1} \right\}, \tag{5.67}$$

where $m = \dim(\mathfrak{n}_0^*)$. We refer the definition of $\mathbb{T}_{O_{i_1'},\text{tang}}$ to Definition 9.3 in [12, page 257]; it needs some clarification as T is a wave packet at the scale ρ_0 and we are talking about tangency at the smaller scale ρ_1.

After defining $\mathbb{T}_{O_{i_1'},\text{tang}}$, we will just set

$$\mathbb{T}_{O_{i_1'},\text{trans}} := \mathbb{T}_{O_{i_1'}} \setminus \mathbb{T}_{O_{i_1'},\text{tang}}. \tag{5.68}$$

Moreover, define

$$g_{\iota,O_{i_1'},\text{tang}} := \sum_{T \in \mathbb{T}_{O_{i_1'},\text{tang}}} (g_{\iota,O_{i_1'}})_T \tag{5.69}$$

and

$$g_{\iota,O_{i_1'},\text{trans}} := \sum_{T \in \mathbb{T}_{O_{i_1'},\text{trans}}} (g_{\iota,O_{i_1'}})_T. \tag{5.70}$$

We continue to define the other two quantities. Define

$$C(\mathfrak{M}_1) := \sum_{O_{i_1'} \in \mathfrak{n}_{1,M}(\mathfrak{n}_0^*)} \sum_{\iota} \left\| T^\lambda g_{\iota,O_{i_1'},\text{trans}} \right\|_{\mathrm{BL}_{k,A}^p(O_{i_1'} \cap B_\iota)}^p \tag{5.71}$$

and

$$C(\mathfrak{R}_1) := \sum_{O_{i_1'} \in \mathfrak{n}_{1,R}(\mathfrak{n}_0^*)} \sum_{\iota} \left\| T^\lambda g_{\iota,O_{i_1'},\text{tang}} \right\|_{\mathrm{BL}_{k,A}^p(O_{i_1'} \cap B_\iota)}^p. \tag{5.72}$$

In the end, we compare $C(\mathfrak{L}_1)$, $C(\mathfrak{M}_1)$, $C(\mathfrak{R}_1)$ and see which one is the largest. For the one that is the largest, its node $\mathfrak{n}_1$ will be called $\mathfrak{n}_1^*$. This finishes the first step.

Before we proceed to the next step, we introduce more notations which will be used later and also in the forthcoming broom estimates. If $\mathfrak{n}_1^* \in \mathcal{L}_1$, then

$$g_{\iota,O_{i_1}}^* := g_{\iota,O_{i_1}}, \tag{5.73}$$

for which we refer to (5.63). If $\mathfrak{n}_1^* \in \mathcal{M}_1$, then

$$g_{\iota,O_{i_1}}^* := g_{\iota,O_{i_1},\text{trans}} \tag{5.74}$$

for which we refer to (5.69). If $\mathfrak{n}_1^* \in \mathcal{R}_1$, then

$$g_{\iota,O_{i_1}}^* := g_{\iota,O_{i_1},\text{tang}} \tag{5.75}$$

for which we refer to (5.70).

Step $2 \le \ell \le \ell_0$. Here $\ell_0 \in \mathbb{N}$ is the last step in the algorithm in Sect. 5.2. Step ℓ will be similar to Step 1. Our goal is to define $C(\mathfrak{L}_\ell), C(\mathfrak{M}_\ell), C(\mathfrak{R}_\ell)$. We consider the case $\mathfrak{n}_{\ell,L}(\mathfrak{n}_{\ell-1}^*)$ and the case $\mathfrak{n}_{\ell,M}(\mathfrak{n}_{\ell-1}^*), \mathfrak{n}_{\ell,R}(\mathfrak{n}_{\ell-1}^*)$ separately.

Take $\mathfrak{n}_\ell = \mathfrak{n}_{\ell,L}(\mathfrak{n}_{\ell-1}^*)$ and $O_{i_\ell} \in \mathfrak{n}_\ell$. Suppose that $O_{i_\ell} \subset B_{\rho_\ell} \cap O_{i_{\ell-1}}$ with $\rho_\ell = \rho(\mathfrak{n}_\ell)$, $O_{i_{\ell-1}} \in \mathfrak{n}_{\ell-1}^*$ and $O_{i_{\ell-1}} \subset B_{\rho_{\ell-1}}$, $\rho_{\ell-1} = \rho(\mathfrak{n}_{\ell-1}^*)$. For a given ι, denote

$$g_{\iota,O_{i_\ell}} = \sum_{\substack{T \in \mathbb{T}[B_{\rho_{\ell-1}}] \\ T \cap O_{i_\ell} \ne \emptyset}} (g_{\iota,O_{i_{\ell-1}}}^*)_T. \tag{5.76}$$

Define

$$C(\mathfrak{L}_\ell) := \sum_{O_{i_\ell} \in \mathfrak{n}_\ell} \sum_\iota \left\| T^\lambda g_{\iota,O_{i_\ell}} \right\|_{\text{BL}_{k,A}^p(O_{i_\ell} \cap B_\iota)}^p. \tag{5.77}$$

Next, take $\mathfrak{n}_\ell = \mathfrak{n}_{\ell,M}(\mathfrak{n}_{\ell-1}^*)$ or $\mathfrak{n}_{\ell,R}(\mathfrak{n}_{\ell-1}^*)$ and $O_{i_\ell'} \in \mathfrak{n}_\ell$. Suppose that $O_{i_\ell'} \subset B_{\rho_\ell} \cap O_{i_{\ell-1}'}$ with $\rho_\ell = \rho(\mathfrak{n}_\ell)$, $O_{i_{\ell-1}'} \in \mathfrak{n}_{\ell-1}^*$ and $O_{i_{\ell-1}'} \subset B_{\rho_{\ell-1}}$, $\rho_{\ell-1} = \rho(\mathfrak{n}_{\ell-1}^*)$. Let $\mathbb{T}_{O_{i_\ell'}}$ denote the collection of all $T \in \mathbb{T}[B_{\rho_\ell}]$ for which

$$T \cap B_{\rho_\ell} \cap W_{m-1} \ne \emptyset, \tag{5.78}$$

with $m = \dim(\mathfrak{n}_{\ell-1}^*)$. Moreover, we will partition $\mathbb{T}_{O_{i_\ell'}}$ into two parts

$$\mathbb{T}_{O_{i_\ell'}} = \mathbb{T}_{O_{i_\ell'},\text{tang}} \bigcup \mathbb{T}_{O_{i_\ell'},\text{trans}}, \tag{5.79}$$

where

$$\mathbb{T}_{O_{i_\ell'},\text{tang}} := \left\{ T \in \mathbb{T}_{O_{i_\ell'}} : T \text{ is } \rho_\ell^{-\frac{1}{2}+\delta_{m-1}}\text{-tangent to } W_{m-1} \text{ on } B_{\rho_\ell} \right\}, \tag{5.80}$$

and

$$\mathbb{T}_{O_{i_\ell'},\text{trans}} := \mathbb{T}_{O_{i_\ell'}} \setminus \mathbb{T}_{O_{i_\ell'},\text{tang}}. \tag{5.81}$$

We continue to define the other two quantities. Define

$$C(\mathfrak{M}_\ell) := \sum_{O_{i_\ell'} \in \mathfrak{n}_{\ell,M}(\mathfrak{n}_{\ell-1}^*)} \sum_\iota \left\| T^\lambda g_{\iota,O_{i_\ell'},\text{trans}} \right\|_{\text{BL}_{k,A}^p(O_{i_\ell'} \cap B_\iota)}^p \tag{5.82}$$

and

$$C(\mathfrak{R}_\ell) := \sum_{O_{i_\ell'} \in \mathfrak{n}_{\ell,R}(\mathfrak{n}_{\ell-1}^*)} \sum_\iota \left\| T^\lambda g_{\iota,O_{i_\ell'},\text{tang}} \right\|_{\text{BL}_{k,A}^p(O_{i_\ell'} \cap B_\iota)}^p. \tag{5.83}$$

In the end, we compare $C(\mathfrak{L}_\ell)$, $C(\mathfrak{M}_\ell)$, $C(\mathfrak{R}_\ell)$ and see which one is the largest. For the one that is the largest, its node $\mathfrak{n}_\ell$ will be called $\mathfrak{n}_\ell^*$.

In the end, we define $g_{i,O_{i_\ell}}^*$ in the same way as in (5.73)–(5.75).

The above algorithm outputs a sequence of nodes

$$\mathfrak{n}_0^*, \mathfrak{n}_1^*, \ldots, \mathfrak{n}_{\ell_0}^*. \tag{5.84}$$

The parameter $\dim(\mathfrak{n}_\ell^*)$ is non-increasing in ℓ. If $\mathfrak{n}_\ell^*$ is an R-child, then

$$\dim(\mathfrak{n}_\ell^*) = \dim(\mathfrak{n}_{\ell-1}^*) - 1; \tag{5.85}$$

otherwise the dimension does not decrease. Denote $m := \dim(\mathfrak{n}_{\ell_0}^*)$. We know that $m \geq k$, where k is as in Theorem 4.2, as otherwise the desired estimate (4.26) there would be trivial. Let

$$\mathfrak{S}_n, \mathfrak{S}_{n-1}, \ldots, \mathfrak{S}_m \tag{5.86}$$

denote the nodes from (5.84) that are R-children, where $\mathfrak{S}_n := \mathfrak{n}_0^*$ is also included. Here $\mathfrak{S}$ is short for "surface", as elements in $\mathfrak{S}_{n'}$ are neighborhoods of algebraic varieties for each n'. We therefore have

$$\dim(\mathfrak{S}_{n'}) = n', \quad \forall m \leq n' \leq n. \tag{5.87}$$

Moreover, denote

$$r_{n'} := \rho(\mathfrak{S}_{n'}), \quad \forall m \leq n' \leq n, \quad r_{m-1} := 1. \tag{5.88}$$

Elements in $\mathfrak{S}_{n'}$ are of the form $B_{r_{n'}} \cap \mathcal{N}_{r_{n'}^{1/2+\delta_{n'}}}(S_{n'})$ where $S_{n'}$ is some algebraic variety of dimension n'. To simplify notation, we will often identify $B_{r_{n'}} \cap \mathcal{N}_{r_{n'}^{1/2+\delta_{n'}}}(S_{n'})$ with $S_{n'}$ if it is clear from the context that we are talking about the node $\mathfrak{S}_{n'}$. We follow [12] and introduce a few new notions.

The pair $(S_{n'}, B_{r_{n'}})$ is called a grain, with its dimension given by n' and degree given by the degree of S.

A multigrain $\vec{S}_{n'}$ is a tuple of grains

$$\vec{S}_{n'} = (\mathcal{G}_n, \ldots, \mathcal{G}_{n'}), \quad \mathcal{G}_i = (S_i, B_{r_i}) \quad \text{for } n' \leqslant i \leqslant n$$

satisfying

1. $\dim(S_i) = i$ for $n' \leqslant i \leqslant n$;
2. $S_n \supset S_{n-1} \supset \cdots \supset S_{n'}$;
3. $B_{r_n} \supset B_{r_{n-1}} \supset \cdots \supset B_{r_{n'}}$.

Sometimes we also write $\vec{S}_{n'} = (S_n, \ldots, S_{n'})$. The parameter $n - n'$ is referred to as the level of the multigrain $\vec{S}_{n'}$. The complexity of the multigrain is defined to be the maximum of the degrees $\deg S_i$ over all $n' \leqslant i \leqslant n$.

Definition 5.9 (Nested tubes, [13]) Let $\vec{S}_{n'} = (\mathcal{G}_n, \ldots, \mathcal{G}_{n'})$ be a multigrain and

$$\mathcal{G}_i = \left(S_i, B_{r_i} \right) \quad \text{for } n' \leqslant i \leqslant n.$$

Define $\mathbb{T}_{r_i}[\vec{S}_{n'}]$ to be the set of length r_i tubes $T \in \mathbb{T}[B_{r_i}]$ that are $r_i^{-\frac{1}{2}+\delta_i}$-tangent to S_i on B_{r_i} (see Definition 5.2 above or Definition 9.3 in [12], page 257) and satisfy that there exists $T_j \in \mathbb{T}[B_{r_j}]$ for $n' \leq j < i$ such that

$$T_j \subset \mathcal{N}_{r_j^{1/2+\delta_j}} S_j, \quad \text{dist}\left(\theta(T_i), \theta(T_j)\right) \lesssim r_j^{-1/2}, \tag{5.89}$$

and

$$\text{dist}\left(T_j, T_i \cap B_{r_j}\right) \lesssim r_i^{(1+\delta)/2} \tag{5.90}$$

hold true for all i, j with $n' \leqslant j < i$. The direction set of $\mathbb{T}_{r_i}[\vec{S}_{n'}]$ is defined to be

$$\Theta_{r_i}[\vec{S}_{n'}] := \{\theta(T) : T \in \mathbb{T}_{r_i}[\vec{S}_{n'}]\}. \tag{5.91}$$

For each $m \leq n' < n$, define

$$D_{n'} = d^{\#_c(\mathfrak{j}(\mathfrak{n}))}, \tag{5.92}$$

where $\mathfrak{n}$ is the parent node of $\mathfrak{S}_{n'}$. Moreover, define

$$D_{m-1} = d^{\#_c(\mathfrak{j}(n_{\ell_0}^*))}, \quad D_n = 1. \tag{5.93}$$

This defines the same quantity as $D_{\ell-1}$ in [12, page 265].

Lemma 5.10 *For each $n' \geq m-1$, it holds that*

$$r_{n'} \prod_{i=n'}^{n} D_i \leq R. \tag{5.94}$$

Proof of Lemma 5.10 It suffices to show that

$$D_i \leq r_{i+1}/r_i, \ \forall i < n. \tag{5.95}$$

Recall that $r_{i+1} = \rho(\mathfrak{S}_{i+1})$ and $r_i = \rho(\mathfrak{S}_i)$. As $\mathfrak{S}_{i+1}$ is an R-child, by definition (see equation (5.20) and the line below), we have

$$\mathfrak{j}(\mathfrak{S}_{i+1}) = \#_a(\mathfrak{j}(\mathfrak{S}_{i+1})) = \#_c(\mathfrak{j}(\mathfrak{S}_{i+1})) = 0. \tag{5.96}$$

Let $\mathfrak{n}$ be the parent node of $\mathfrak{S}_i$, and therefore $D_i = d^{\#_c(\mathfrak{j}(\mathfrak{n}))}$. When the algorithm runs from $\mathfrak{S}_{i+1}$ to $\mathfrak{n}$, the radius parameter ρ decreases to ρ/d each time $\#_c$ increases by 1, and (5.95) follows immediately. $\qquad\square$

In the end of this subsection, we describe a few output functions of the above algorithm. Take n' with $m \leq n' \leq n$ and consider the node $\mathfrak{S}_{n'}$. As $\mathfrak{S}_{n'}$ is an R-child, it means the tangential case $C(\mathfrak{R}_\ell)$, which was defined in (5.83), dominates. Here ℓ is the level that $\mathfrak{S}_{n'}$ belongs to. As elements in $\mathfrak{S}_{n'}$ are neighborhoods of algebraic varieties, from now on we will always use $S_{n'}$ to refer to an element in $\mathfrak{S}_{n'}$. Consequently, $g^*_{\iota, O_{i'_\ell}}$ will be called $g^*_{\iota, S_{n'}}$, and its wave packets in the ball $B_{r_{n'}}$ with $S_{n'} \subset B_{r_{n'}}$ are all tangent to $S_{n'}$.

Regarding these functions, we have the following properties. Let $p_{n'}$ with $n' \geq m \geq k$ be a Lebesgue exponent that is fixed later. These exponents satisfy

$$p_m \geq p_{m+1} \geq \cdots \geq p_n = p \geq 2, \tag{5.97}$$

where p is the exponent in Theorem 4.2. Define $\alpha_{n'}, \beta_{n'} \in [0, 1]$ by

$$\frac{1}{p_{n'}} = \frac{1 - \alpha_{n'-1}}{2} + \frac{\alpha_{n'-1}}{p_{n'-1}}, \quad \beta_{n'} = \prod_{i=n'}^{n-1} \alpha_i, \tag{5.98}$$

for $m + 1 \leq n' \leq n - 1$, and $\alpha_n = \beta_n = 1$. We have

Property 1. The inequality

$$\|T^\lambda g\|_{\mathrm{BL}^p_{k,A}(B_R)} \lesssim M(\vec{r}_{n'}, \vec{D}_{n'}) \|g\|_{L^2}^{1-\beta_{n'}} \tag{5.99}$$

$$\left(\sum_{S_{n'} \in \mathfrak{S}_{n'}} \sum_\iota \left\| T^\lambda g^*_{\iota, S_{n'}} \right\|_{\mathrm{BL}^{p_{n'}}_{k, A_{n'}}(B_{r_{n'}})}^{p_{n'}} \right)^{\frac{\beta_{n'}}{p_{n'}}}, \tag{5.100}$$

where $B_{r_{n'}}$ is the ball of radius $r_{n'}$ that contains $S_{n'}$ and

$$\vec{r}_{n'} := (r_n, r_{n-1}, \ldots, r_{n'}), \quad \vec{D}_{n'} := (D_n, D_{n-1}, \ldots, D_{n'}), \tag{5.101}$$

holds for

$$M(\vec{r}_{n'}, \vec{D}_{n'}) := \left(\prod_{i=n'}^{n-1} D_i \right)^{(n-n')\delta} \left(\prod_{i=n'}^{n-1} r_i^{(\beta_{i+1} - \beta_i)/2} D_i^{(\beta_{i+1} - \beta_{n'})/2} \right). \tag{5.102}$$

Property 2. For $n' \leq n - 1$, we have

$$\sum_{S_{n'} \in \mathfrak{S}_{n'}} \left\| g^*_{\iota, S_{n'}} \right\|_2^2 \lesssim D_{n'}^{1+\delta} \sum_{S_{n'+1} \in \mathfrak{S}_{n'+1}} \left\| g^*_{\iota, S_{n'+1}} \right\|_2^2, \tag{5.103}$$

for every $B_\iota \subset B_R$ of radius $R^{1-\delta}$. Here when $n' = n - 1$, $g^*_{\iota, S_{n'+1}}$ was not defined before and we simply set $g^*_{\iota, S_{n'+1}} = g$.

Property 3. For $n' \leq n - 1$, we have

$$\max_{S_{n'} \in \mathfrak{S}_{n'}} \left\| g^*_{\iota, S_{n'}} \right\|_2^2 \lesssim \left(\frac{r_{n'+1}}{r_{n'}} \right)^{-\frac{n-n'-1}{2}} D_{n'}^{-n'+\delta} \max_{S_{n'+1} \in \mathfrak{S}_{n'+1}} \left\| g^*_{\iota, S_{n'+1}} \right\|_2^2 \tag{5.104}$$

and

$$\max_{S_{n'}\in\mathfrak{S}_{n'}}\max_\theta \left\| g_{\iota,S_{n'}}^* \right\|_{L^2_{\mathrm{avg}}(\theta)}^2 \tag{5.105}$$

$$\lesssim \left(\frac{r_{n'+1}}{r_{n'}}\right)^{-\frac{n-n'-1}{2}} D_{n'}^\delta \max_{S_{n'+1}\in\mathfrak{S}_{n'+1}}\max_\theta \left\| g_{\iota,S_{n'+1}}^* \right\|_{L^2_{\mathrm{avg}}(\theta)}^2, \tag{5.106}$$

where θ is a frequency cap defined in Sect. 4.1 of side length $\rho^{-1/2}$, hold for all $1 \le \rho \le r_{n'}$.

Property 4. For $n' \le n'' \le n$, it holds that

$$\left\| g_{\iota,S_{n'}}^* \right\|_2^2 \lesssim_\epsilon r_{n'}^{\frac{n-n'}{2}} \left(\prod_{i=n'}^{n-1} r_i^{-\frac{1}{2}}\right) r_{n''}^{-\frac{n-n''}{2}} \left(\prod_{i=n''}^{n-1} r_i^{\frac{1}{2}}\right) R^{O(\epsilon_\circ)} \left\| g_{\iota,S_{n'}}^{*(n'')} \right\|_2^2, \tag{5.107}$$

where

$$g_{\iota,S_{n'}}^{*(n'')} := \sum_{T\in\mathbb{T}_{r_{n''}}[\vec{S}_{n'}]} (g_\iota)_T, \tag{5.108}$$

and $\mathbb{T}_{r_{n''}}[\vec{S}_{n'}]$ is from Definition 5.9.

If one takes $n'' = n$, then $g_{\iota,S_{n'}}^{*(n)}$ becomes $g_{S_{n'}}^\#$ in the last equation in [13, page 9]. If $n'' = n'$, then $g_{\iota,S_{n'}}^* = g_{\iota,S_{n'}}^{*(n'')}$.

These four properties are taken essentially from [13, page 9]. The only main difference is that Hickman and Zahl [13] only introduced and used $n'' = n$ in (5.108). The proofs of the first three properties are given in [12], the proof of the fourth property is the same as that of Property iv) in [13, page 10] and relies on transverse equidistribution properties (for our setting, the needed property is in [10]), and therefore we will not repeat it.

The only explanation that is needed is as follows. In the last equation in [12, page 255], the authors there used the fact that $\rho_{j+1} \simeq \rho_j$, where the implicit constant is universal. In our case, these two radius parameters differ by d, which is a large constant. Consequently, Property (III)$_j$ in [12, page 250] may not hold as is written there. However, we can still obtain some good substitute for it. For a node $\mathfrak{n}$, let $\mathfrak{n}^{\Uparrow}$ denote the closest ancestor (itself included) that is an R-child. Let $\mathfrak{n}_\ell$ be a node in $\mathfrak{M}_\ell' \cup \mathfrak{M}_\ell$ with $\mathfrak{n}_\ell^{\Uparrow} = \mathfrak{S}_{n'}$. We will prove

$$\left\| g_{\iota,O_\ell} \right\|_2^2 \le C_{\ell,\delta}^{\mathrm{III}} \cdot \left(\frac{r_{n'}}{\rho(\mathfrak{n}_\ell)}\right)^{-\frac{n-n'}{2}} d^{-\#_c(\mathrm{j}(\mathfrak{n}_\ell))(n'-1)} \left\| g \right\|_2^2, \tag{5.109}$$

for every $O_\ell \in \mathfrak{n}_\ell$ and ball B_ι of radius $R^{1-\delta}$, where

$$C_{\ell,\delta}^{\mathrm{III}} := d^{\#_c(\mathrm{j}(\mathfrak{n}_\ell))\delta + \#_a(\mathrm{j}(\mathfrak{n}_\ell))\delta} (r_{n'})^{\#_a(\mathrm{j}(\mathfrak{n}_\ell))O(\delta_{n'-1/2}) + \#_{a'}(\mathrm{j}(\mathfrak{n}_\ell))O(\delta_{n'})} \tag{5.110}$$

and

$$\#_{a'}(\mathrm{j}(\mathfrak{n}_\ell)) := \#\{\mathfrak{n} : \mathfrak{n}_\ell \preccurlyeq \mathfrak{n} \preccurlyeq \mathfrak{S}_{n'}, \mathfrak{n} \in \mathfrak{M}_{\ell'}' \text{ for some } \ell'\}. \tag{5.111}$$

Note that

$$\#_a(\mathfrak{j}(\mathfrak{n}_\ell)) \lesssim \frac{|\log \delta_{n'-1}|}{\delta_{n'-1}}, \quad \#_{a'}(\mathfrak{j}(\mathfrak{n}_\ell)) \lesssim \frac{|\log \delta_{n'-1/2}|}{\delta_{n'-1/2}}, \tag{5.112}$$

and therefore by applying Lemma 5.10, we always have (5.110) $\lessapprox 1$.

Let us prove (5.109). Let ℓ_1 be the largest integer smaller than ℓ such that there exists $\mathfrak{n}_{\ell_1} \in \mathfrak{M}'_{\ell_1} \cup \mathfrak{M}_{\ell_1}$ with $\mathfrak{n}_\ell \preccurlyeq \mathfrak{n}_{\ell_1}$ and $\mathfrak{n}_{\ell_1}^{\Uparrow} = \mathfrak{S}_{n'}$; if no such nodes exist, then we simply take $\mathfrak{n}_{\ell_1} = \mathfrak{S}_{n'}$. Assume that (5.109) has been proved for $\mathfrak{n}_{\ell_1}$, and we will prove it for $\mathfrak{n}_\ell$. There are two cases: $\mathfrak{n}_\ell \in \mathfrak{M}_\ell$ or $\mathfrak{n}_\ell \in \mathfrak{M}'_\ell$. We only prove the latter case, and the former case can be done in a similar way. List all the nodes between $\mathfrak{n}_{\ell_1}$ and $\mathfrak{n}_\ell$ in a descending order:

$$\mathfrak{n}_{\ell_1}, \mathfrak{n}_{\ell_1+1}, \ldots, \mathfrak{n}_{\ell-1}, \mathfrak{n}_\ell. \tag{5.113}$$

Note that $\mathfrak{n}_{\ell'}$ is an L-child for every $\ell_1 < \ell' < \ell$. By orthogonality between wave packets and the fundamental theorem of algebra, we have

$$\left\| g_{\iota, O_{\ell'}} \right\|_2^2 \lesssim d^{-(n'-1)} \left\| g_{\iota, O_{\ell'-1}} \right\|_2^2, \tag{5.114}$$

for every $\ell' < \ell$, $O_{\ell'} \in \mathfrak{n}_{\ell'}$, $O_{\ell'-1} \in \mathfrak{n}_{\ell'-1}$ and $O_{\ell'} \subset O_{\ell'-1}$. This further implies

$$\left\| g_{\iota, O_{\ell-1}} \right\|_2^2 \leq C_{\ell-1,\delta}^{\mathrm{III}} \cdot \left(\frac{r_{n'}}{\rho(\mathfrak{n}_{\ell_1})} \right)^{-\frac{n-n'}{2}} d^{-\#_c(\mathfrak{j}(\mathfrak{n}_{\ell-1}))(n'-1)} \|g\|_2^2. \tag{5.115}$$

Here note that the denominator is $\rho(\mathfrak{n}_{\ell_1})$ instead of $\rho(\mathfrak{n}_{\ell-1})$, as remarked at the beginning of Sect. 5.2. When passing from $\mathfrak{n}_{\ell-1}$ to $\mathfrak{n}_\ell$, recall that in (5.46) and (5.49), we cut the neighborhood $\mathcal{N}_{\rho(\mathfrak{n}_{\ell-1}^\uparrow)^{1/2+\delta_m}} = \mathcal{N}_{\rho(\mathfrak{n}_{\ell_1})^{1/2+\delta_m}}$ into thinner layers $\mathcal{N}_{\rho(\mathfrak{n}_\ell)^{1/2+\delta_m}}$. Therefore by transverse equidistribution properties (for instance Lemma 8.4 in [10]), we have

$$\left\| g_{\iota, O_\ell} \right\|_2^2 \lesssim (r_{n'})^{O(\delta_{n'})} d^{-(n'-1)} \left(\frac{\rho(\mathfrak{n}_{\ell_1})}{\rho(\mathfrak{n}_\ell)} \right)^{-\frac{n-n'}{2}} \left\| g_{\iota, O_{\ell-1}} \right\|_2^2, \tag{5.116}$$

for $O_\ell \in \mathfrak{n}_\ell$ with $O_\ell \subset O_{\ell-1}$. This, combined with (5.115) and the choice of the constant in (5.110), gives us the desired bound.

6 Strong polynomial Wolff axioms

The goal of this section is to prove a strong polynomial Wolff axiom (see Theorem 6.2 below), a stronger version of the polynomial Wolff axiom in Sect. 3. As a direct consequence, we will obtain the key Lemma 6.1, which controls the number of tubes concentrated near a multi-grain. Strong polynomial axioms already appeared in earlier works in [14, 24] and [13]. The proof in this section combines the argument in the papers mentioned above and Bourgain's conditions as formulated in Theorem 2.1.

Lemma 6.1 *Let $\vec{S}_{n'}$ be the multi-grain given in Definition 5.9 with complexity at most d. We have*

$$\#\Theta_{r_i}[\vec{S}_{n'}] \lesssim_{\epsilon_\circ, d} \left(\prod_{j=n'}^{i-1} r_j^{-1/2} \right) r_i^{\frac{i-1}{2} + \epsilon_\circ}, \tag{6.1}$$

for all $n' \leq i \leq n$, where $\epsilon_\circ$ is given in (1.44).

As with Theorem 1.2, we will deduce Lemma 6.1 from a geometric theorem.

Theorem 6.2 [*Strong Polynomial Wolff Axiom for our ϕ*] *Let $n \geq 3$. If Bourgain's condition holds for the phase function ϕ at every $(\mathbf{x}_0; \xi_0) \in \mathrm{supp}(a)$, then the following strong polynomial Wolff axiom for ϕ holds: Let $E \geq 2$ be an integer and we fix an integer k. For every $\epsilon > 0$, there exists $C(n, E, k, \epsilon) > 0$ such that if we have a sequence of balls*

$$B(\mathbf{x}_1, s_1) \subset \frac{1}{2} B(\mathbf{x}_2, s_2) \subset \cdots \subset \frac{1}{2^{k-1}} B(\mathbf{x}_k, s_k) \subset \frac{1}{2^k} B^n, \tag{6.2}$$

numbers $\kappa^{C_0} \leq \kappa_1 \leq \cdots \leq \kappa_k \leq \kappa$ and subsets $S_j \subset B(\mathbf{x}_j, s_j)$, $j = 1, 2, \ldots, k$ satisfying:

- $\kappa_j \leq s_j, \forall 1 \leq j \leq k$ *and* $\varphi_j := \frac{s_j}{\kappa_j}$ *satisfy* $\varphi_1 \geq \varphi_2 \geq \cdots \geq \varphi_k$,
- S_j *is a semialgebraic set of complexity $\leq E$ whose κ_j-neighborhood has volume $\simeq |S_j|$,*
- *The intersection between S_j and any ball of radius $r \in [\kappa_j, s_j]$ has volume $\leq C_j r^{d_j} \kappa_j^{n-d_j}$,*

then the following holds uniformly:

For every collection $\mathbb{T}$ of κ-tubes pointing in different directions (defined before Theorem 1.2), if we use $c(T)$ to denote the core curve of T, then

$$\#\{T \in \mathbb{T} : c(T) \bigcap \mathcal{B}(\mathbf{x}_j, \frac{1}{2} s_j) \subset S_j, \forall 1 \leq j \leq k\}$$

$$\leq C(C_0, n, E, k, \epsilon) \prod_{j=1}^{k} \left(\frac{\varphi_{j-1}}{\varphi_j} \right)^{d_j - 1} \left(\frac{\varphi_k}{\kappa} \right)^{n-1} \kappa^{-\epsilon} \tag{6.3}$$

where $\varphi_0 = 1$ and the horizontal slab[9] $\mathcal{B}(\mathbf{x}_j, \frac{1}{2} s_j)$ is defined to be $\pi_t^{-1}(\pi_t(B(\mathbf{x}_j, \frac{1}{2} s_j)))$. Here $\pi_t : \mathbb{R}^n \to \mathbb{R}$ is the orthogonal projection to the t-variable.

Moreover the implied constant only depends on bounds of finitely many (depending on C_0, n, E, k, ϵ) derivatives of ϕ.

Like in Sect. 3, we are going to deduce Theorem 6.2 when the phase function satisfies a concrete derivative condition. Then we simply check that the condition is satisfied by our phase function.

---

[9]In general, we call a set to be a *horizontal slab* if it is $\pi_t^{-1}(I)$ for some interval I.

Theorem 6.3 [*Generalized Strong Polynomial Wolff Axiom*] *Let $n \geq 3$. Suppose for a Φ as in the beginning of Sect.* 3:

(a) *For every choice of $v \in [-1, 1]^{n-1}$, $\xi \in [-1, 1]^{n-1}$, subinterval $I \subset [-1, 1]$ and $t \in [-1, 1]$, we have both*

$$|\det(\nabla_v \Phi(v, t, \xi) \cdot M + \nabla_\xi \Phi(v, t, \xi))|$$

$$\lesssim \left(1 + \frac{\mathrm{dist}(t, I)}{|I|}\right)^{n-1} \frac{1}{|I|} \int_I |\det(\nabla_v \Phi(v, s, \xi) \cdot M + \nabla_\xi \Phi(v, s, \xi))| \mathrm{d}s \quad (6.4)$$

and (3.3) *for some implied constants independent of the choices of v, ξ, I and M.*

(b) *If $t_1 \neq t_2$, $\Phi(v, t_1, \xi) = x_1$, and $\Phi(v', t_1, \xi') = x_1$, then*

$$|\Phi(v', t_2, \xi')) - \Phi(v, t_2, \xi))| \lesssim |t_1 - t_2| \cdot |\xi - \xi'|. \quad (6.5)$$

(c) *If $t_1 \neq t_2$ and $\Phi(v, t_1, \xi)) = x_1$, $\Phi(v, t_2, \xi)) = x_2$, then for x_2' with distance $\mu |t_1 - t_2|$ from x_2 ($\mu \leq 10$), there are v' and ξ' with $\Phi(v', t_1, \xi')) = x_1$, $\Phi(v', t_2, \xi')) = x_2'$ and $|\xi' - \xi| \lesssim \mu$.*

Then the conclusion of Theorem 6.2 (*with the notion "pointing in different directions" now defined as in the beginning of Sect.* 3) *holds with the implied constant only depending on bounds of finitely many* (*depending on C_0, n, E, k, ϵ*) *derivatives of Φ and the implied constants in* (a) - (c).

Remark 6.4 We explain the intuition behind (b) and (c) a bit. For convenience we introduce the following notation. For fixed $v \in \mathbb{R}^{n-1}$ and $\xi \in \mathbb{R}^{n-1}$, we call the curve

$$c_{v,\xi} = \{(x, t) \in \mathbb{R}^{n-1} \times [-1, 1] : x = \Phi(v, t, \xi)\} \quad (6.6)$$

to be a Φ-*curve*. Intuitively, if we know a Φ-curve passes through (x_1, t_1) and want to perturb the "direction variable" ξ and the "initial position variable" v so that (x_1, t_1) is still on the curve, then (b) says whenever the perturbation on the direction ξ is $O(\mu)$ we always have the perturbation of the curve at time $t = t_2$ is $O(|t_1 - t_2|\mu)$, and (c) says if we want the x coordinate at time $t = t_2$ to be shifted by a distance $\simeq \mu |t_1 - t_2|$, we can always succeed with the amount of the perturbation needed on ξ being $O(\mu)$.

Proof of Theorem 6.3 Similarly to the proof of Theorem 3.1, we only do the proof when Φ is fixed and after seeing the proof it will be clear that the estimate only depends on finitely many derivatives of Φ (in particular since the smallest scale (κ^{C_0}) we consider is polynomial in κ, in the approximation argument described below one only needs a Taylor approximation of order $O_{C_0, n, E, k, \epsilon}(1)$).

Like the proof of Theorem 3.1, we can reduce the situation to the case where Φ is a polynomial of degree $O_{C_0, n, E, k, \epsilon}(1)$ by a Taylor approximation argument. For general Φ by this argument one reduces to a problem with two modified conditions: (a') a slightly weaker condition than (6.4) in (a) (similar to (3.7) versus (3.3)) that has very small error term (of the form $\kappa^{-1000n(1+C_0)}$) which makes no difference (see

the proof below and how to deal with this issue in the similar situation in the proof of Theorem 3.1), and (b') (c') two slightly weaker conditions than (b) and (c) with an error term κ^{1+C_0} for $|t| \leq \kappa^\epsilon$, both not affecting our framework (for (b) and (c), note that the only place they are needed is the verification of a claim in the beginning of the induction step). From now on we always assume Φ is a polynomial of degree $O_{C_0,n,E,k,\epsilon}(1)$.

Let us assume κ and κ_j are all sufficiently small (allowed to depend on derivatives of Φ). Otherwise we simply ignore some constraints. This assumption will enable us to use the implicit function theorem at scales κ or κ_j freely.

We make one more comment before starting: the present theorem is a generalization of Lemma 3.7 in [13], which was in turn developed based on Theorem 1.4 in [14] or Theorem 1.9 in [24]). Our proof will have a lot in common with these, and will be a natural generalization of Theorem 3.1.

We also refine the slabs $\mathcal{B}(\mathbf{x}_j, \frac{1}{2}s_j)$ a bit before starting. Since each $B(\mathbf{x}_j, s_j)$ is contained in $\frac{1}{2}B(\mathbf{x}_{j+1}, s_{j+1})$ $(\forall j < k)$, we can take horizontal slabs $\mathcal{B}_1, \ldots, \mathcal{B}_k$ such that:

(i) $\mathcal{B}_j \subset \mathcal{B}(\mathbf{x}_j, \frac{1}{2}s_j)$.
(ii) The thickness of $\mathcal{B}_j$ is $\simeq s_j$.
(iii) The distance between $\mathcal{B}_{j_1}$ and $\mathcal{B}_{j_2}$ is $\gtrsim s_{j_2}$ for all pairs $j_1 < j_2$.

For $1 \leq l \leq k$ and $t \in [-1, 1]$, we define

$$S_{l,t} = \{y \in \mathbb{R}^{n-1} : \exists \text{ a } \Phi - \text{curve } c_{v,\xi} \text{ s.t.}$$

$$c_{v,\xi} \bigcap \mathcal{B}_j \subset S_j, \forall 1 \leq j \leq l \text{ and } (y, t) \in c_{v,\xi}\} \tag{6.7}$$

and we are going to prove inductively that

$$m_{n-1}^*(S_{l,t}) \leq C(C_0, n, E, k, \epsilon)(d_l(t))^{n-1}$$

$$\times \prod_{j=1}^{l} (\frac{\varphi_{j-1}}{\varphi_j})^{d_j-1} \varphi_l^{n-1} \kappa^{-2^{l-k-1}\epsilon}, \forall t \in [-1, 1] \tag{6.8}$$

where $d_l(t)$ is defined to be s_l plus the distance between $B(\mathbf{x}_l, s_l)$ and the hyperplane $\{x_n = t\}$, and $m_{n-1}^*(\cdot)$ is the $(n-1)$-dimensional Lebesgue outer measure.

For convenience, we choose a large constant K depending on the implied constant in (b) and (c) and also consider a companion set

$$\tilde{S}_{l,t} = \{y \in \mathbb{R}^{n-1} : \exists \text{ a } \Phi - \text{curve } c_{v,\xi} \text{ s.t.}$$

$$c_{v,\xi} \bigcap \mathcal{B}_j \subset \mathcal{N}_{K\kappa_j}(S_j), \forall 1 \leq j \leq l \text{ and } (y, t) \in c_{v,\xi}\} \tag{6.9}$$

Note that once we prove (6.8), by the same reasoning and the assumption about $\mathcal{N}_{\kappa_j}(S_j)$, we also prove the same upper bound for $m_{n-1}^*(\tilde{S}_{l,t})$.

Base case. Our base case is when $l = 0$. Since Φ has a C^1-derivative bound, we observe that all $S_{l,t}$ lie in a uniformly bounded set Ω_0. We make the convention that $S_{0,t}$ is the part of the set defined in (6.7) with $l = 0$ (hence with a vacuous condition)

lying in Ω_0. (6.8) then trivially holds for $l = 0$ with the convention $d_0(t) = 1$. We will see the first induction step is similar to the steps afterward with this setup.

The inductive step. Suppose we have (6.8) for some l in $[1, k)$. Next we prove it for $l + 1$.

Integrate the induction hypothesis over t and temporally ignore measurability issues, by Fubini we formally deduce

$$m_n^*((\bigcup_t S_{l,t}) \bigcap \mathcal{B}_{l+1})$$

$$\leq C(C_0, n, E, k, \epsilon) \prod_{j=1}^{l} (\frac{\varphi_{j-1}}{\varphi_j})^{d_j - 1} \varphi_l^{n-1} s_{l+1}^n \kappa^{-2^{l-k-1}\epsilon}. \quad (6.10)$$

We assert a stronger conclusion: In fact, $(\bigcup_t S_{l,t}) \bigcap \mathcal{B}_{l+1}$ is contained in a set U_l such that U_l is a union of (n-dim) balls of radii $\varphi_l s_{l+1}$ and that

$$m_n^*(U_l) \leq C(C_0, n, E, k, \epsilon) \prod_{j=1}^{l} (\frac{\varphi_{j-1}}{\varphi_j})^{d_j - 1} \varphi_l^{n-1} s_{l+1}^n \kappa^{-2^{l-k-1}\epsilon}. \quad (6.11)$$

To construct such a U_l, we take the $\varphi_l s_{l+1}$-neighborhood of $(\bigcup_t S_{l,t}) \bigcap \mathcal{B}_{l+1}$ and cover it by a finitely overlapping collection of balls of radii $\varphi_l s_{l+1}$. Define the union of these balls to be U_l.

It remains to derive the volume bound (6.11). We claim that we can take K in (6.9) large (and the constraint here will be the only one affecting the choice of K) such that this U_l is contained in $\bigcup_t \tilde{S}_{l,t}$.

To verify this claim, by definition we see U_l is contained in the $3\varphi_l s_{l+1}$-neighborhood of $(\bigcup_t S_{l,t}) \bigcap \mathcal{B}_{l+1}$. This means for every point $(\tilde{y}, t)$ in U_l, we can find a Φ-curve that is $3\varphi_l s_{l+1}$-close to this point and the part of that curve in $\mathcal{B}_j$ completely lies in S_j, $\forall 1 \leq j \leq l$. Now keep a point of that Φ-curve in $\mathcal{B}_1$ fixed and by assumption (c) (see Remark 6.4 for more intuition), one can change the "direction" ξ by up to $O(\varphi_l)$ so that the new Φ-curve now passes through $(\tilde{y}, t)$. By assumption (b) for each $t \in \pi_t \mathcal{B}_j (1 \leq j \leq l)$, the perturbation amount of the x variable is $\lesssim \varphi_l s_j \lesssim \varphi_j s_j = \kappa_j$. Hence the intersection between the new Φ-curve and $\mathcal{B}_j$ lies in $\mathcal{N}_{K\kappa_j}(S_j)$ if K is sufficiently large depending on the implied constants in (b) and (c). For this choice of K we just proved that the claim holds. Applying the induction hypothesis to each t-slice of $\tilde{S}_{l,t}$ and integrate, we deduce (6.11).

Our U_l is a union of $\varphi_l s_{l+1}$-balls that contains $(\bigcup_t S_{l,t}) \bigcap \mathcal{B}_{l+1}$ and obeys the volume bound (6.11). By a covering lemma we may assume without loss of generality that the $\varphi_l s_{l+1}$-balls are finite-overlapping. Now we use the volume upper bound of the intersection between S_{l+1} and r-balls in the assumption of Theorem 6.2. We deduce

$$m_n^*((\bigcup_t S_{l+1,t}) \bigcap \mathcal{B}_{l+1})$$

$$\leq C(C_0, n, E, k, \epsilon) \prod_{j=1}^{l} (\frac{\varphi_{j-1}}{\varphi_j})^{d_j - 1} \varphi_l^{n-1} s_{l+1}^n (\frac{\kappa_{l+1}}{\varphi_l s_{l+1}})^{n - d_{l+1}} \kappa^{-2^{l-k-1}\epsilon}$$

$$=C(C_0,n,E,k,\epsilon)\prod_{j=1}^{l}(\frac{\varphi_{j-1}}{\varphi_j})^{d_j-1}\varphi_l^{n-1}s_{l+1}^n(\frac{\varphi_{l+1}}{\varphi_l})^{n-d_{l+1}}\kappa^{-2^{l-k-1}\epsilon}$$

$$=C(C_0,n,E,k,\epsilon)\prod_{j=1}^{l+1}(\frac{\varphi_{j-1}}{\varphi_j})^{d_j-1}\varphi_{l+1}^{n-1}s_{l+1}^n\kappa^{-2^{l-k-1}\epsilon}. \tag{6.12}$$

We will use (6.12) to close the induction step by an argument developed in [14] and [24]. Below we fix an arbitrary $t = t_0$ to do the proof. This step is a lot similar to the proof of Theorem 3.1. So we will present some steps in sketch only.

Define the set

$$L_{l+1,t} = \{(v,\xi) : c_{v,\xi} \bigcap \mathcal{B}_j \subset S_j, \forall 1 \le j \le l+1\}. \tag{6.13}$$

We already assumed Φ is a polynomial of degree $O_{C_0,n,E,k,\epsilon}(1)$. Thus in the expression of a Φ-curve, x is polynomial in v,t,ξ with the same degree bound. For simplicity we use b_{t_0} to denote the hyperplane $\{x_n = t_0\}$.

By effective quantifier elimination (i.e. the Tarski-Seidenberg theorem) we find a semialgebraic subset $L'_{l+1,t} \subset L_{l+1,t}$ of complexity $O_{n,E,\epsilon_1,K}(1)$ such that all $c_{v,\xi} \bigcap b_{t_0}$ are distinct for $(v,\xi) \in L'_{l+1,t}$, and that $\{c_{v,\xi} \bigcap b_{t_0} : (v,\xi) \in L'_{l+1,t}\} = \{c_{v,\xi} \bigcap b_{t_0} : (v,\xi) \in L_{l+1,t}\}$. (To see this one can first add $(n-1)$ more coordinates to each $(\xi,v) \in L_{l+1,t}$ denoting the "position", i.e. the first $(n-1)$ coordinates, of the intersection $c_{v,\xi} \bigcap b_{t_0}$. This is still a semialgebraic set of bounded complexity. Then one applies the quantifier elimination to find a subset such that the last $(n-1)$ coordinates are distinct among different points in the subset and the set of the last $(n-1)$ coordinates does not change. From the construction we see easily that $L'_{l+1,t}$ has dimension $\le n-1$.

Using Gromov's lemma to approximate $L'_{l+1,t}$ by images of smooth maps in the same way as we did to prove Theorem 3.1, we see (for arbitrary $\epsilon_1 > 0$) there exist two polynomial maps F and $G : [0,\kappa^{\epsilon_1}]^{n-1} \to \mathbb{R}^{n-1}$ (whose images are the v and the ξ variables, respectively) with $\deg F, \deg G = O_{n,E,\epsilon_1}(1)$ and $\|F\|_{C^1}, \|G\|_{C^1} \le 1$ such that

$$c_{F(x),G(x)} \bigcap \mathcal{B}_j \subset \mathcal{N}_{\kappa_j}(S_j), \forall x, \forall 1 \le j \le l+1 \tag{6.14}$$

and that

$$\mathcal{H}^{n-1}(\{c_{F(x),G(x)} \bigcap b_{t_0} : x \in [0,\kappa^{\epsilon_1}]^{n-1}\}) \gtrsim_{n,E,\epsilon_1,K} \kappa^{C\epsilon_1} m_{n-1}^*(S_{l+1,t_0}) \tag{6.15}$$

where $\mathcal{H}^{n-1}(\cdot)$ stands for the $(n-1)$-dimensional Hausdorff measure on the hyperplane b_{t_0}.

Now look at the n-dimensional volume of

$$M = \{(\Phi(F(x),t,G(x)),t) : x \in [0,\kappa^{\epsilon_1}]^{n-1}\} \bigcap \mathcal{B}_{l+1}$$

$$= (\bigcup c_{F(x),G(x)}) \bigcap \mathcal{B}_{l+1} \tag{6.16}$$

and the $(n-1)$-dimensional volume of

$$H = \{(\Phi(F(x), t_0, G(x)), t_0) : x \in [0, \kappa^{\epsilon_1}]^{n-1}\} = \left(\bigcup c_{F(x), G(x)}\right) \bigcap b_{t_0} \quad (6.17)$$

and compare them.

Suppose the time interval (i.e. the range of the last coordinate) of $\mathcal{B}_{l+1}$ is I_{l+1}. Then using Bézout like in the proof of Theorem 3.1, we have

$$\begin{aligned}
|M| &\sim_{n, E, \epsilon_1} \int_{I_{l+1}} \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\nabla_x (\Phi(F(x), t, G(x)))| \mathrm{d}x \mathrm{d}t \\
&= \int_{I_{l+1}} \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\det(\nabla_v \Phi(F(x), t, G(x)) \cdot \nabla_x F \\
&\quad + \nabla_\xi \Phi(F(x), t, G(x)) \cdot \nabla_x G)| \mathrm{d}x \mathrm{d}t \\
&= \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\det(\nabla_x G)| \\
&\quad \cdot \int_{I_{l+1}} |\det(\nabla_v \Phi(F(x), t, G(x)) \cdot (\nabla_x F \cdot (\nabla_x G)^{-1}) \\
&\quad + \nabla_\xi \Phi(F(x), t, G(x)))| \mathrm{d}t \mathrm{d}x. \quad (6.18)
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\mathcal{H}^{n-1}(H) \\
&\sim_{n, E, \epsilon_1} \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\nabla_x (\Phi(F(x), t_0, G(x)))| \mathrm{d}x \\
&= \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\det(\nabla_v \Phi(F(x), t_0, G(x)) \cdot \nabla_x F + \nabla_\xi \Phi(F(x), t_0, G(x)) \cdot \nabla_x G)| \mathrm{d}x \\
&= \int_{[0, \kappa^{\epsilon_1}]^{n-1}} |\det(\nabla_x G)| \cdot |\det(\nabla_v \Phi(F(x), t_0, G(x)) \cdot (\nabla_x F \cdot (\nabla_x G)^{-1}) \\
&\quad + \nabla_\xi \Phi(F(x), t_0, G(x)))| \mathrm{d}x. \quad (6.19)
\end{aligned}$$

Now by assumption (a), the left hand side of (6.18) is $\gtrsim |I_{l+1}|^n \cdot d_{l+1}(t_0)^{-(n-1)} \simeq s_{l+1}^n \cdot d_{l+1}(t_0)^{-(n-1)}$ times the left hand side of (6.19). Note that (the $\tilde{S}_{l+1,t}$ version of) (6.12) gives an upper bound of the left hand side of (6.18). Moreover (6.15) gives a lower bound of the left hand side of (6.19) in terms of $m_{n-1}^*(S_{l+1,t_0})$. Combining everything, we can take ϵ_1 to be a sufficiently small multiple of ϵ to finish the induction step and (6.8) is proved.

From (6.8) the conclusion will follow easily. Take the union of all $S_{k,t}$ for $t \in [-1, 1]$. We notice by the definition and effective quantifier elimination that this is a semialgebraic set of complexity $O_{n, E, \epsilon}(1)$. Use (6.8), we see its measure is

$$\leq C(C_0, n, E, k, \epsilon) \prod_{j=1}^k \left(\frac{\varphi_{j-1}}{\varphi_j}\right)^{d_j - 1} \varphi_k^{n-1} \kappa^{-2^{-1}\epsilon}. \quad (6.20)$$

Note that we already have (3.3) holds. In exactly the same way as we proved Theorem 3.1 (the only difference is that Theorem 3.1 was stated for tubes and we need a version for Φ-curves but notice that in Theorem 3.1 in fact a Φ-curves version was proven), we can bound the left hand side of (6.3) by

$$C(C_0, n, E, k, \epsilon) \prod_{j=1}^{k} \left(\frac{\varphi_{j-1}}{\varphi_j}\right)^{d_j-1} \varphi_k^{n-1} \kappa^{(1-n)-\epsilon}. \tag{6.21}$$

This concludes the proof. □

Proof of Theorem 6.2 The proof will be similar to the proof of Theorem 1.2 in §3. As with that Theorem, take unique smooth $\Phi = \Phi(v, t, x)$ near 0 such that (3.26) holds. We can assume the above can be done for all $(v, t, \xi) \in [-1.5, 1.5]^{2n-1}$ without loss of generality like in the other proof. It suffices to show that our Φ satisfies conditions (a)-(c) in Theorem 6.3. That Theorem then immediately leads to the desired conclusion. When reading the proof one naturally sees the implicit constants in (b) and (c) only depend on finitely many derivatives of ϕ. We also note that (b) and (c) comes from the non-degeneracy property of ϕ, and (a) comes from Bourgain's condition.

For (a), (3.3) is already verified in the proof of Theorem 1.2. Recall we defined $A(t; \xi) = \nabla_\xi^2 \phi(X_t(\xi), t; \xi)$ in (3.31) and deduced $A(t; \xi) = f(t; \xi)B(\xi) + A(0; \xi)$ and the time derivative of f is 1 at $t = 0$ around (3.35). We make one more harmless assumption that the time derivative of f is always in $(\frac{1}{2}, 2)$ since otherwise we can do a constant rescaling that only causes loss of a constant. Now we can do the reduction to both sides of (6.4) like in the proof of Theorem 1.2, which reduces (6.4) to proving that for every polynomial $P(t)$ of degree $n - 1$,

$$|P(t)| \lesssim \left(1 + \frac{\text{dist}(t, I)}{|I|}\right)^{n-1} \frac{1}{|I|} \int_I |P(s)| \mathrm{d}s, \tag{6.22}$$

which is an elementary Theorem proved in e.g. Lemma 3.8 in [14]. We have completed the verification of (a).

As before, (b) and (c) are more general properties that do not depend on Bourgain's condition. Next we verify them.

Differentiate (3.26) with respect to t, we see

$$\nabla_x \nabla_\xi \phi \cdot \partial_t \Phi + \partial_t \nabla_\xi \phi = 0. \tag{6.23}$$

Hence Φ-curves can be viewed as integral curves of the vector field $V_\xi(\mathbf{x}, t) = (\nabla_x \nabla_\xi \phi(\mathbf{x}, t, \xi)^{-1} \cdot \partial_t \nabla_\xi \phi(\mathbf{x}, t, \xi), 1)$ parameterized by ξ. Note that the non-degeneracy of ϕ implies $|\nabla_x \nabla_\xi \phi| \simeq 1$ and a μ-perturbation on ξ only causes a $O(\mu)$-perturbation of the above vector field. We see (b) follows from stability of ODE solutions.

Next we check (c), which is the "opposite direction" to (b). When checking it we can assume t_1 and t_2 are sufficiently close and that x_1 is sufficiently close to 0 (and in application the honest (c) will always be satisfied after a harmless constant-rescaling of (x, t)). (c) basically asks: if we start from

$$x_0 = \Phi(v_0, t_1, \xi_0) \tag{6.24}$$

and start to change ξ and solve v from the equation

$$x_0 = \Phi(v, t_1, \xi), \tag{6.25}$$

how would $y = \Phi(v, t_2, \xi)$ change? For convenience denote $y_0 = \Phi(v_0, t_2, \xi_0)$. We use differentiation to compute this change. Differentiating (6.25), we see

$$\nabla_v \Phi|_{(v_0, t_1, \xi_0)} \cdot \nabla_\xi v|_{v_0} + \nabla_\xi \Phi|_{(v_0, t_1, \xi_0)} = 0. \tag{6.26}$$

Hence by the chain rule,

$$\nabla_\xi y = \nabla_v \Phi|_{(v_0, t_2, \xi_0)} \cdot (-\nabla_v \Phi|_{(v_0, t_1, \xi_0)})^{-1} \cdot \nabla_\xi \Phi|_{(v_0, t_1, \xi_0)} + \nabla_\xi \Phi|_{(v_0, t_2, \xi_0)}. \tag{6.27}$$

By (3.28) and (3.29), this simplifies to

$$\nabla_\xi y = (\nabla x \nabla_\xi \phi|_{y_0, t_2, \xi_0})^{-1} \cdot (\nabla_\xi^2 \phi|_{x_0, t_1, \xi_0} - \nabla_\xi^2 \phi|_{y_0, t_1, \xi_0}). \tag{6.28}$$

The first factor $(\nabla x \nabla_\xi \phi|_{y_0, t_2, \xi_0})^{-1}$ has entries $\lesssim 1$ and determinant $\simeq 1$ and is harmless. We focus on the second factor. It is equal to $(t_2 - t_1)$ times some $(\sum_{j=1}^{n-1} c_j \partial_{x_j} \nabla_\xi^2 \phi + \partial_t \nabla_\xi^2 \phi)|_{x_0, t_1, \xi_0}$ plus a higher order term in $(t_2 - t_1)$, where each $|c_j| \lesssim 1$. By a familiar technique of parabolic rescaling (see for example the reduction Lemmas 4.1-4.3 in [10]), one can assume all $\|\partial_{x_j} \nabla_\xi^2 \phi\|$ are uniformly very small and since we have the nondegeneracy condition on $\partial_t \nabla_\xi^2 \phi$ from (1.7), we see $\nabla_\xi y$ is equal to $(t_2 - t_1)$ times a nondegenerate matrix of bounded entries. From here we see (c) holds by an application of the implicit function theorem.

Now that (a)-(c) are all verified, we apply Theorem 6.3 and conclude the proof. □

Proof of Lemma 6.1 At this point, the Lemma is just a straightforward consequence of Theorem 6.2. We rescale the whole $B_{r_{n'}}$ to the unit ball and rescale all $\mathcal{N}_{r_j^{1/2+\delta_j}} S_j$ in Definition 5.9 accordingly (to be our S_j in Theorem 6.2). For each possible $\theta(T)$ in $\Theta_{r_i}[\vec{S}_{n'}]$, we pick the core curve of the corresponding $T_{n'}$, rescale it into the unit ball and extend it into a Φ-curve (Φ defined from ϕ as in the beginning of §3). Then we choose $\kappa = R^{-\frac{1}{2}}$ and see by Definition 5.9 that the set of κ-tubes around all above Φ-curves satisfies the assumption of Theorem 6.2 with the balls having radii $s_j = \frac{r_{i-1+j}}{r_{n'}}$ and corresponding $\kappa_j = \frac{r_{i-1+j}^{\frac{1}{2}+\delta}}{r_{n'}}$. Hence $\varphi_j = r_{n'+1-j}^{-\frac{1}{2}+\delta_{i-1+j}}$ By Wongkew's theorem [22] of intersections between neighborhoods of algebraic varieties we can take $d_j = i - 1 + j$. Lastly we can surely take $E = O_{d,n}(1)$ to be a constant by the definition of S_j in Definition 5.9.

By Theorem 6.2, we see the left hand side of (6.1) is

$$\lesssim r_i^{\delta_i} r_{n'}^{\frac{n'-1}{2}} \prod_{j=n'+1}^{i} \left(\frac{r_j}{r_{j-1}} \right)^{\frac{j-1}{2}} \tag{6.29}$$

and is thus bounded by the right hand side. □

As a corollary of Lemma 6.1, we obtain

Corollary 6.5 *For* $m \leq n' \leq n''$, *we have*

$$\left\| g_{\iota, S_{n'}}^{*(n'')} \right\|_2^2 \lesssim \Big(\prod_{j=n'}^{n''} r_j^{-\frac{1}{2}} \Big) r_{n''}^{-\frac{n-n''-1}{2}} \max_{\tau : \ell(\tau) = r_{n''}^{-1/2}} \left\| g_{\iota, S_{n'}}^{*(n'')} \right\|_{L^2_{\mathrm{avg}}(\tau)}^2 \tag{6.30}$$

7 Brooms

7.1 Definition of brooms

Let $(S, B(\mathbf{x}_0, r))$ be a grain of dimension n' with $S \in \mathfrak{S}_{n'}$ and assume that it is the last entry of a multi-grain $\vec{S}$. Throughout this section, we always assume that

$$r \geq \sqrt{R}. \tag{7.1}$$

Recall Definition 5.2 and Definition 5.9. Define $\mathbb{T}[S] \subset \mathbb{T}[B(\mathbf{x}_0, r)]$ to be the collection of tubes that are tangent to S in the ball $B(\mathbf{x}_0, r)$. Define

$$\Theta[S] := \{\theta(T) : T \in \mathbb{T}[S]\}. \tag{7.2}$$

Moreover, define $\mathbb{T}_R[S] := \mathbb{T}_R[\vec{S}]$.

Before we define brooms, we cut each S into $O_{d,n}(1)$ many pieces so that the tangent spaces of each piece form a small angle with each other. Let us be more precise. For each $\mathbf{z} \in S$, let $T_{\mathbf{z}}S$ denote the tangent space of S at $\mathbf{z}$. We cut S into $O_{d,n}(1)$ many pieces $\{S', S'', \dots\}$ so that for each such piece, say S', it holds that

$$\angle(T_{\mathbf{z}_1} S', T_{\mathbf{z}_2} S') \leq \frac{1}{100n}, \tag{7.3}$$

for $\mathbf{z}_1, \mathbf{z}_2 \in S'$. Similarly, we define $\mathbb{T}[S']$, $\Theta[S']$ and $\mathbb{T}_R[S']$. Such a decomposition only appears in this section. To simplify notation, in the rest of this section we will still use S to refer to each such piece, and still call it a grain.

Fix $\tau \in \Theta[S]$ and a grain S satisfying (7.3). Define

$$\mathbb{T}_{\tau, R}[S] := \{T \in \mathbb{T}_R[S] : \theta(T) \subset \tau\}. \tag{7.4}$$

We let R-tubes $T \in \mathbb{T}_{\tau, R}[S]$ intersect S. Morally speaking, $T \cap S$ can be thought of as a "curved" rectangular box of dimensions

$$r \times \underbrace{R^{1/2+\delta} \times \cdots \times R^{1/2+\delta}}_{(n'-1)\text{copies}} \times \underbrace{r^{1/2+\delta_m} \times \cdots \times r^{1/2+\delta_m}}_{(n-n')\text{copies}}. \tag{7.5}$$

For two tubes $T_1, T_2 \in \mathbb{T}_{\tau, R}[S]$, we say that

$$T_1 \cap S \cap B(\mathbf{x}_0, r) \approx T_2 \cap S \cap B(\mathbf{x}_0, r) \tag{7.6}$$

if

$$T_1 \cap S \cap B(\mathbf{x}_0, r) \subset (10nT_2) \cap S \cap B(\mathbf{x}_0, r), \tag{7.7}$$

or the other way around. Before we study the geometry of $S_\square := T \cap S \cap B(\mathbf{x}_0, r)$, let us assume without loss of generality that the tangent space $T_{\mathbf{z}}(S)$ forms an angle $\leq 1/(100n)$ with the subspace spanned by $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}, \vec{e}_n\}$, the first $(n'-1)$ vectors from the orthonormal basis and the vertical t coordinate direction $\vec{e}_n$, for every $\mathbf{z} \in S$.

Lemma 7.1 (1) *We can write*

$$S \supseteq \bigcup_\square S_\square \tag{7.8}$$

where $S_\square = T \cap S \cap B(\mathbf{x}_0, r)$ for some $T \in \mathbb{T}_{\tau, R}[S]$ and $\{S_\square\}_\square$ is a disjoint collection. Moreover, for every $T \in \mathbb{T}_{\tau, R}[S]$, we can find $S_\square$ such that $T \cap S \cap B(\mathbf{x}_0, r) \approx S_\square$.

(2) *Take $(x_1, t_1) \in S$. For each $S_\square$, we can find an algebraic variety $Z \subset \{t = t_1\}$ of dimension $n' - 1$ and complexity $O(\deg(S))$ satisfying that the angle between $T_{\mathbf{z}}(Z)$ and the subspace $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}\}$ is $\leq 1/(100n)$ for every $\mathbf{z} \in Z \cap S_\square$, such that*

$$(B(x_1, r) \times \{t_1\}) \cap S_\square \subset \mathcal{N}_{r^{1/2}}(Z). \tag{7.9}$$

Here $B(x_1, r)$ is the ball in $\mathbb{R}^n$ of radius r centered at x_1.

Proof of Lemma 7.1 If $\mathbb{T}_{\tau, R}[S]$ is empty, then define the right-hand side of (7.8) as an empty set. Now assume that $\mathbb{T}_{\tau, R}[S]$ is nonempty. Pick $T \in \mathbb{T}_{\tau, R}[S]$ and define $S_\square = T \cap S \cap B(\mathbf{x}_0, r)$. If there exists $T' \in \mathbb{T}_{\tau, R}[S]$ and $T \cap T' \cap S \cap B(\mathbf{x}_0, r) \neq \emptyset$, then $T \cap S \cap B(\mathbf{x}_0, r) \approx T' \cap S \cap B(\mathbf{x}_0, r)$. To see this, if $\mathbf{x} \in T \cap T'$ and $\theta(T), \theta(T') \subset \tau$, then

$$T \cap B(\mathbf{x}, 10r) \approx T' \cap B(\mathbf{x}, 10r)$$

and $B(\mathbf{x}_0, r) \subset B(\mathbf{x}, r)$.

If there exists $T' \in \mathbb{T}_{\tau, R}[S]$, then we add $S'_\square := T' \cap S \cap B(\mathbf{x}_0, r)$ to the right-hand side of (7.8). Continue until for every $T \in \mathbb{T}_{\tau, R}[S]$, there exists $S_\square$ from the right-hand side of (7.8) such that $S_\square = T \cap S \cap B(\mathbf{x}_0, r)$.

For each $S_\square$, define $Z = S \cap \{t = t_1\}$. $\square$

To define brooms, we fix $(S, B(\mathbf{x}_0, r))$, $\vec{S}$, τ and $S_\square$. Write $\mathbf{x}_0 = (x_0, t_0)$. Define

$$\mathbb{T}_{\tau, R}[S_\square] := \{T \in \mathbb{T}_{\tau, R}[S] : T \cap S_\square \neq \emptyset\}. \tag{7.10}$$

Let us record the following lemma that will be useful later.

Lemma 7.2 *Under the above notation, we have that*

$$\bigcup_{T \in \mathbb{T}_{\tau, R}[S_\square]} \left(T \cap \{(x, t_1) \in \mathbb{R}^n : x \in \mathbb{R}^{n-1}\} \right) \tag{7.11}$$

is contained in an $(n-1)$ dimensional ball of radius $R^{1+\delta}r^{-1/2}$, for every $|t_1 - t_0| \leq R$.

Proof of Lemma 7.2 By translation, we assume that $S_\square$ contains the origin. Let ω_0 be the center of τ. For $\omega \in \tau$, let $x = X_\omega(t)$ denote the solution to

$$\nabla_\omega \phi(x, t; \omega) = 0, \tag{7.12}$$

for $|t| \leq 1$. Then we need to show that

$$|X_{\omega_0}(t) - X_\omega(t)| \lesssim r^{-1/2}, \tag{7.13}$$

for every t. Note that

$$\nabla_\omega \phi(X_\omega(t), t; \omega) - \nabla_\omega \phi(X_{\omega_0}(t), t; \omega_0) = 0. \tag{7.14}$$

By Taylor's expansion, this further implies

$$\nabla_x \nabla_\omega \phi(x', t; \omega)(X_\omega(t) - X_{\omega_0}(t)) + \nabla_\omega^2 \phi(X_{\omega_0}(t), t; \omega')(\omega - \omega_0) = 0 \tag{7.15}$$

for some x', ω'. The desired bound follows from the fact that $\nabla_x \nabla_\omega \phi$ is non-degenerate. $\qquad\square$

We apply the following algorithm. Initialize

$$\mathbb{T}_0 := \mathbb{T}_{\tau, R}[S_\square], \quad \mathbb{D}_0 = \{T \cap \{t = t_0 + R\} : T \in \mathbb{T}_{\tau, R}[S_\square]\}. \tag{7.16}$$

Suppose we are at the ℓ'-th step of the algorithm and from now on $\ell' \geq 1$. Let $Z \subset \{t = t_0 + R\}$ be an algebraic variety of dimension $n' - 1$ and complexity $O(\deg(S))$ which satisfies that the angle between $T_{\mathbf{z}}(Z)$ and the subspace spanned by $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}\}$ is $\leq 1/(100n)$, for every $\mathbf{z} \in Z$. Find such a Z that maximizes

$$\#\{D_0 \in \mathbb{D}_0 : D_0 \subset \mathcal{N}_{10nR^{1/2+\delta}}(Z)\}; \tag{7.17}$$

use $b_{\ell'}$ to refer to the number in (7.17). Remove the discs (7.17) from $\mathbb{D}_0$, use $\mathbb{T}_{\ell'}$ to collect the tubes T from $\mathbb{T}_0$ for which $T \cap \{t = t_0 + R\}$ is removed from $\mathbb{D}_0$ in this step, and repeat this process until there is no any discs left.

Suppose this algorithm terminates after L steps. We obtain a collection of positive integers

$$b_1 \geq b_2 \geq \cdots \geq b_L \tag{7.18}$$

and a collection of tubes

$$\mathbb{T}_1, \mathbb{T}_2, \ldots, \mathbb{T}_L. \tag{7.19}$$

We group $\{b_{\ell'}\}_{\ell'}$ by checking which interval from

$$[1, R^\delta), (R^\delta, R^{2\delta}], (R^{2\delta}, R^{3\delta}], \ldots \tag{7.20}$$

they belong to:

$$\{b_{\ell_0+1}, \ldots, b_{\ell_1}\}, \{b_{\ell_1+1}, \ldots, b_{\ell_2}\}, \ldots \tag{7.21}$$

with $\ell_0 = 0$, and therefore two $b_{\ell'}, b_{\ell''}$ in the same group are comparable up to factor R^δ. Now we are ready to define brooms.

Definition 7.3 (Brooms) Fix $(S, B(\mathbf{x}_0, r))$, $\vec{S}$, τ and $S_\square$. Each

$$\mathcal{B}_{\ell,b} := \bigcup_{\ell_m+1 \leq \ell' \leq \ell_{m+1}} \mathbb{T}_{\ell'}, \tag{7.22}$$

(see (7.19) for definition of $\mathbb{T}_{\ell'}$), with level

$$\ell := \lfloor R^{w\delta} \rfloor, \text{ with } w \in \mathbb{N}, R^{w\delta} \leq \ell_{m+1} - \ell_m < R^{(w+1)\delta}, \tag{7.23}$$

is called a broom. Here

$$b := R^{w'\delta} \text{ with } w' \in \mathbb{N}, R^{w'\delta} \leq b_{\ell_m+1} < R^{(w'+1)\delta}, \tag{7.24}$$

will be called the length of the broom. For the broom $\mathcal{B}_{\ell,b}$ in (7.22), we say that it is rooted at $S_\square$.

In the previous definition, we used tubes from $\mathbb{T}_R[S]$. For some perhaps technical reasons, we need to introduce the notion of brooms by using a sub-collection of tubes from $\mathbb{T}_R[S]$. Fix $(S, B(\mathbf{x}_0, r))$, $\vec{S}$, τ and $S_\square$, and a sub-collection $\mathbb{T}'_R[S] \subset \mathbb{T}_R[S]$. We repeat the above definition of brooms with $\mathbb{T}_R[S]$ replaced by $\mathbb{T}'_R[S]$, and obtain a unique decomposition

$$\mathbb{T}'_R[S] = \bigcup_{\ell,b} \mathcal{B}_{\ell,b}(\mathbb{T}'_R[S]), \tag{7.25}$$

where each $\mathcal{B}_{\ell,b}(\mathbb{T}'_R[S])$ is called a broom of level ℓ and length b, and generated by tubes from $\mathbb{T}'_R[S]$.

7.2 Definition of the two-ends relation

Recall the algorithm in Sect. 5.2. For each node $\mathfrak{n} \in \cup_\iota \mathfrak{R}_\iota$, we will define a relation $\sim_\mathfrak{n}$; Lemma 5.6 guarantees that we have a small number of these relations.

We define a few auxiliary functions $\chi_{\mathfrak{n},\kappa} = \chi_\kappa$, taking values 0 or 1, where $\kappa = ((\ell_1, b_1), \mu_1, \ldots, (\ell_\iota, b_\iota), \mu_\iota)$, $\iota \in \mathbb{N}$ and $\ell_{\iota'}, b_{\iota'}, \mu_{\iota'} \in \{R^{w\delta} : w \in \mathbb{N}\}$ for every $1 \leq \iota' \leq \iota$. Denote

$$r = \rho(\mathfrak{n}). \tag{7.26}$$

Step 1. For $S \in \mathfrak{n}$ and a tube $T \in \mathbb{T}[B_R]$, we say that

$$\chi_{(\ell_1, b_1)}(S, T) = 1 \tag{7.27}$$

if T belongs to a broom rooted at some $S_\square \subset S$ with level ℓ_1 and length b_1. Moreover, we say that

$$\chi_{(\ell_1,b_1),\mu_1}(S,T) = 1 \tag{7.28}$$

if

$$\chi_{(\ell_1,b_1)}(S,T) = 1 \tag{7.29}$$

and

$$\mu_1 \le \sum_{S' \in \mathfrak{n}} \chi_{(\ell_1,b_1)}(S',T) < \mu_1 R^\delta. \tag{7.30}$$

A general step. Suppose we have defined χ_κ for $\kappa = ((\ell_1,b_1),\mu_1,\ldots,(\ell_\iota,b_\iota),\mu_\iota)$, and $\iota \ge 1$. Let us define

$$\chi_{\kappa,(\ell_{\iota+1},b_{\iota+1})}, \quad \chi_{\kappa,(\ell_{\iota+1},b_{\iota+1}),\mu_{\iota+1}}. \tag{7.31}$$

For fixed $S \in \mathfrak{n}$, define

$$\mathbb{T}_{S,\kappa} := \{T' \in \mathbb{T}[B_R] : \chi_\kappa(S,T') = 1\}. \tag{7.32}$$

Recall (7.25). Write

$$\mathbb{T}_{S,\kappa} = \bigcup_{\ell_{\iota+1},b_{\iota+1}} \bigcup_{\tau,S_\square} \mathcal{B}_{\ell_{\iota+1},b_{\iota+1},\tau,S_\square}(\mathbb{T}_{S,\kappa}), \tag{7.33}$$

where τ runs through all frequency caps of side length $\rho(\mathfrak{n})^{-1/2}$, $S_\square$ is as given in Lemma 7.1, $\mathcal{B}_{\ell_{\iota+1},b_{\iota+1},\tau,S_\square}(\mathbb{T}_{S,\kappa})$ is a broom of level $\ell_{\iota+1}$, length $b_{\iota+1}$, rooted at $S_\square$ and generated by tubes from $\mathbb{T}_{S,\kappa}$. We then say that

$$\chi_{\kappa,(\ell_{\iota+1},b_{\iota+1})}(S,T) = 1 \text{ if } T \in \mathcal{B}_{\ell_{\iota+1},b_{\iota+1},\tau,S_\square}(\mathbb{T}_{S,\kappa,\tau}), \tag{7.34}$$

for some τ and $S_\square$. Next, set

$$\chi_{\kappa,(\ell_{\iota+1},b_{\iota+1}),\mu_{\iota+1}} = 1 \tag{7.35}$$

if

$$\chi_{\kappa,(\ell_{\iota+1},b_{\iota+1})}(S,T) = 1 \tag{7.36}$$

and

$$\mu_{\iota+1} \le \sum_{S' \in \mathfrak{n}} \chi_{\kappa,(\ell_{\iota+1},b_{\iota+1}),\mu_{\iota+1}}(S',T) < \mu_{\iota+1} R^\delta. \tag{7.37}$$

This finishes the definition of the auxiliary functions we need.

For $\kappa = ((\ell_1,b_1),\mu_1,\ldots,(\ell_\iota,b_\iota),\mu_\iota)$, we say that κ is admissible if there exists exactly one pair (ι_1,ι_2) with $\iota_1 \neq \iota_2$ such that

$$((\ell_{\iota_1},b_{\iota_1}),\mu_{\iota_1}) = ((\ell_{\iota_2},b_{\iota_2}),\mu_{\iota_2}). \tag{7.38}$$

Lemma 7.4 *The number of admissible κ is $O_\delta(1)$.*

Proof of Lemma 7.4 Note that by (7.23), the number of values that $\ell_{\iota'}$ can take is δ^{-1}; the same is true is for $b_{\iota'}$ and $\mu_{\iota'}$, for each ι'. The lemma follows. $\qquad\square$

Definition 7.5 For a ball $B \subset B_R$ of radius $R^{1-\delta}$, a tube $T \in \mathbb{T}[B_R]$, a node $\mathfrak{n} \in \cup_\iota \mathfrak{R}_\iota$ and an admissible multi-index κ, we say that $B \sim_{\mathfrak{n},\kappa} T$ if B maximizes

$$\#\{S' \in \mathfrak{n} : S' \subset B', \chi_{\mathfrak{n},\kappa}(S', T) = 1\}, \tag{7.39}$$

among all B' of radius $R^{1-\delta}$.

Definition 7.6 (Relation) For a ball B of radius $R^{1-\delta}$ and a tube $T \in \mathbb{T}[B_R]$, we say that $B \sim T$ if

$$B \sim_{\mathfrak{n},\kappa} T, \tag{7.40}$$

for some node $\mathfrak{n} \in \cup_\iota \mathfrak{R}_\iota$ and admissible κ.

By a simple inductive argument on κ, we have:

Lemma 7.7 *Let $\vec{S}$ be a multi-grain with the last component given by S. For every $T \in \mathbb{T}_R[S]$, there exists exactly one admissible κ such that $T \in \mathbb{T}_{S,\kappa}$.*

7.3 Broom estimates

Let $\vec{S}_{n'}$ be a multigrain from Definition 5.9 with the last component given by $S_{n'}$. Let $B_\iota \subset B_R$ be the ball of radius $R^{1-\delta}$ that contains $S_{n'}$. Recall the definition of $f^*_{\iota,S_{n'}}$ from Sect. 5.4 and the definition of $f^{*(n'')}_{\iota,S_{n'}}$ with $n' \le n'' \le n$ from (5.108). The notation $*(n'')$ means that we start with the function $f^*_{\iota,S_{n'}}$, which is defined via wave packets from $\mathbb{T}[B_{r_{n''}}]$, and "trace" back by Definition 5.9 along the nodes in (5.86) to wave packets in $\mathbb{T}[B_{r_{n''}}]$. Note that

$$f^{*(n'')}_{\iota,S_{n'}} = f^*_{\iota,S_{n'}}, \quad \text{when } n'' = n'. \tag{7.41}$$

Define

$$f^{\nsim}_{S_{n'},\tau} := (f_\tau)^*_{\iota,S_{n'}}. \tag{7.42}$$

Here we have suppressed the dependence on ι and replaced it by $\nsim$, as B_ι is uniquely determined by S, and we would also like to emphasize that we are in the non-related case. Moreover, define

$$f^{\nsim(n'')}_{S_{n'},\tau} := (f_\tau)^{*(n'')}_{\iota,S_{n'}}. \tag{7.43}$$

The main goal of this subsection is to prove the following broom estimate.

Theorem 7.8 *Let* $\vec{S}_{n'} = (S_n, \dots, S_{n'})$ *be a multi-grain and* $S_{n'} = \mathcal{N}_{r_{n'}^{1/2+\delta_{n'}}}(Z_{n'}) \cap B_{r_{n'}}$ *with* $r_{n'} \geq \sqrt{R}$, *where* $Z_{n'}$ *is an* n'-*dimensional algebraic variety of degree* $\lesssim_{n'} 1$ *in* $\mathbb{R}^n$. *Then, for* τ *of scale* $r_{n'}^{-1/2}$ *and* $n' \leq n'' < n$, *it holds that*

$$\| f_{S_{n'},\tau}^{\infty(n'')} \|_{L^2}^2 \lessapprox \left(\frac{r_{n''}}{R} \right)^{\frac{n-n'}{2}} \| f_\tau \|_{L^2}^2. \tag{7.44}$$

Here $r_i = \rho(S_i)$ *for each* $n' \leq i \leq n$.

Proof of Theorem 7.8 We will first write down the details for the case $n'' = n'$, as it requires less notation and contains all the ideas of the proof; the general case $n'' \geq n'$ will be remarked in the end. Our goal is to prove

$$\| f_{S_{n'},\tau}^{\infty} \|_{L^2}^2 \lessapprox \left(\frac{r_{n'}}{R} \right)^{\frac{n-n'}{2}} \| f_\tau \|_{L^2}^2. \tag{7.45}$$

To simplify notation, we will abbreviate $S_{n'}$ to S, $\vec{S}_{n'}$ to $\vec{S}$ and $r_{n'}$ to r. For the ball B_ι of radius $R^{1-\delta}$ containing S, the node $\mathfrak{n}$ containing S and each admissible multi-index κ, recall the notation in (7.32) and define

$$\mathbb{T}_{S,\kappa,\tau}^{\infty} := \{ T \in \mathbb{T}_{S,\kappa} : \theta(T) \subset \tau, B_\iota \sim_{\mathfrak{n},\kappa} T \}, \tag{7.46}$$

and

$$f_{\kappa,S,\tau}^{\infty} := \sum_{T \in \mathbb{T}_{S,\kappa,\tau}^{\infty}} f_T \quad \text{and} \quad f_{\kappa,S,\tau}^{\infty,*} := (f_{\kappa,S,\tau}^{\infty})_{\iota,S}^*.$$

Then we claim that

$$f_{S,\tau}^{\infty} = \sum_{\kappa} f_{\kappa,S,\tau}^{\infty,*}. \tag{7.47}$$

To see this, note that for every $T \in \mathbb{T}[B_R]$ that is tangent to S, by Lemma 7.7, there exists exactly one admissible κ such that $T \in \mathbb{T}_{S,\kappa}$. Moreover, if $T \sim B_\iota$, then for the above given $\mathfrak{n}$ and κ, we also have $T \sim_{\mathfrak{n},\kappa} B_\iota$.

By the Cauchy-Schwarz inequality,

$$\| f_{S,\tau}^{\infty} \|_2^2 \lesssim_\delta \sum_{\kappa} \| f_{\kappa,S,\tau}^{\infty,*} \|_2^2. \tag{7.48}$$

Here we used Lemma 7.4. Next, we localize $f_{\kappa,S,\tau}^{\infty}$ further in space. Recall from Lemma 7.1 that $S = \cup_\square S_\square$. Denote

$$\mathbb{T}_{S_\square,\kappa,\tau}^{\infty} := \{ T \in \mathbb{T}_{S,\kappa} : T \cap S_\square \neq \emptyset, \theta(T) \subset \tau, B_\iota \sim_{\mathfrak{n},\kappa} T \}, \tag{7.49}$$

and then

$$f_{\kappa,S_\square,\tau}^{\infty} := \sum_{T \in \mathbb{T}_{S_\square,\kappa,\tau}^{\infty}} f_T \quad \text{and} \quad f_{\kappa,S_\square,\tau}^{\infty,*} := (f_{\kappa,S_\square,\tau}^{\infty})_{\iota,S}^*. \tag{7.50}$$

By spatial disjointness of $\{S_\square\}_\square$, we have

$$(7.48) \lesssim \sum_\kappa \sum_\square \| f^{\varkappa,*}_{\kappa,S_\square,\tau} \|_2^2. \tag{7.51}$$

Let (x_1, t_1) be a point in S, then

$$\| f^{\varkappa,*}_{\kappa,S_\square,\tau} \|_{L^2}^2 \lesssim \| T^\lambda f^{\varkappa,*}_{\kappa,S_\square,\tau} \|_{L^2((B(x_1,2r_{n'})\times\{t_1\})\cap S_\square)}^2 \tag{7.52}$$

Here $B(x_1, 2r_{n'})$ is an $n - 1$ dimensional ball. Indeed, inequalities of the form (7.52) hold for more general data. Let h_τ be supported on τ. Assume that $T^\lambda h_\tau(\cdot, t_1)$ is essentially supported on $B_{\sqrt{R}} \subset \mathbb{R}^{n-1}$, a ball of radius $\sqrt{R}$. Then

$$\| h_\tau \|_2^2 \lesssim \| T^\lambda h_\tau \|_{L^2(B_{\sqrt{R}}\times\{t_1\})}. \tag{7.53}$$

To see this, we first apply the change of variables

$$x \mapsto x + x_0, t \mapsto t + t_1, \xi \mapsto \xi + \xi_0, \tag{7.54}$$

where x_0 is the center of $B_{\sqrt{R}}$ and ξ_0 is the center of τ. We obtain

$$T^\lambda h_\tau(x + x_0, t + t_1) = \int h_\tau(\xi + \xi_0) e^{i\phi^\lambda(x+x_0,t+t_1;\xi+\xi_0)} a^\lambda(x + x_0, t + t_1; \xi + \xi_0) d\xi. \tag{7.55}$$

Its absolute value can be written as the absolute value of

$$\int \widetilde{h}_\tau(\xi) e^{i\phi_0^\lambda(x,t;\xi)} a^\lambda(x + x_0, t + t_1; \xi + \xi_0) d\xi \tag{7.56}$$

where

$$\begin{aligned} \phi_0^\lambda(x,t;\xi) :=&\phi^\lambda(x + x_0, t + t_1; \xi + \xi_0) - \phi^\lambda(x_0, t_1; \xi + \xi_0) \\ &- \phi^\lambda(x + x_0, t + t_1; \xi_0) + \phi^\lambda(x_0, t_1; \xi_0) \end{aligned} \tag{7.57}$$

and

$$\widetilde{h}_\tau(\xi) := h_\tau(\xi + \xi_0) e^{i\phi^\lambda(x_0,t_1;\xi+\xi_0)} \tag{7.58}$$

This change of variables tells us that, if we denote

$$\widetilde{T}^\lambda h_\tau(x, t) := \int h_\tau(\xi) e^{i\phi_0^\lambda(x,t;\xi)} a^\lambda(\mathbf{x}; \xi) d\xi, \tag{7.59}$$

then to prove (7.53), it suffices to prove it for the operator as defined in (7.59) with $t_1 = 0, x_0 = 0, \xi_0 = 0$.

We apply Taylor expansion to ϕ_0^λ in the x variable about the origin, and write

$$\phi_0^\lambda(x, 0; \xi) = \nabla_x \phi_0^\lambda(0; \xi) \cdot x + w_1(x; \xi). \tag{7.60}$$

Note that

$$|\nabla_x^2 \phi_0^\lambda| \lesssim 1/\lambda, \quad |x| \lesssim \sqrt{R}, \tag{7.61}$$

and therefore $|w_1(x;\xi)| \lesssim 1$. We apply Taylor expansion in the ξ variable about the origin, and obtain

$$\phi_0^\lambda(x,0;\xi) = \langle x,\xi \rangle + \underbrace{w_2(x;\xi) + w_1(x;\xi)}_{=:w(x;\xi)}. \tag{7.62}$$

Note that

$$|\nabla_x \nabla_\xi^\beta \phi_0^\lambda| \lesssim_\beta 1, \tag{7.63}$$

for all multi-indices β and therefore $|w_2(x;\xi)| \lesssim r^{-1}\sqrt{R} \lesssim 1$. Now the claimed estimate (7.53) follows from Plancherel's theorem and Taylor's expansion.

Suppose that

$$\kappa = \{(\ell_1, b_1), \mu_1, \ldots, (\ell_j, b_j), \mu_j\}. \tag{7.64}$$

For each plank $S_\square$,

$$\mathbb{T}_{\tau,R}[S_\square] \cap \mathbb{T}_{S,\kappa},$$

where $\mathbb{T}_{\tau,R}[S_\square]$ was defined in (7.10), is contained in a broom $\mathcal{B}_{\ell_j,b_j}$. Recall the definition of brooms and the notation in (7.19). Write

$$\mathcal{B}_{\ell_j,b_j} = \bigcup_{1 \le \ell' \le \ell_j} \mathbb{T}_{\ell'}, \tag{7.65}$$

and

$$\bigcup_{T \in \mathbb{T}_{\ell'}} T \cap \{t = t_1 + R\} \subset \mathcal{N}_{10nR^{1/2+\delta}}(Z_{\ell'}) \tag{7.66}$$

for an algebraic variety $Z_{\ell'}$ of dimension $n' - 1$ and complexity $O(\deg(S))$ satisfying that the angle between $T_{\mathbf{z}}(Z_{\ell'})$ and the space spanned by $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}\}$ is $\le 1/(100n)$, for every $\mathbf{z} \in Z_{\ell'}$. Write

$$f_{\kappa,S_\square,\tau,\ell'}^\infty = \sum_{T \in \mathbb{T}_{S_\square,\kappa,\tau}^\infty \cap \mathbb{T}_{\ell'}} f_T$$

and

$$f_{\kappa,S_\square,\tau,\ell'}^{\infty,*} := (f_{\kappa,S_\square,\tau,\ell'}^\infty)_{i,S}^*.$$

Then by the triangle inequality and Cauchy-Schwarz inequality,

$$\begin{aligned}
&\|T^\lambda f_{\kappa,S_\square,\tau}^{\infty,*}\|_{L^2((B(x_1,2r)\times\{t_1\})\cap S_\square)}^2 \\
&\lesssim \ell_j \sum_{1 \le \ell' \le \ell_j} \|T^\lambda f_{\kappa,S_\square,\tau,\ell'}^{\infty,*}\|_{L^2((B(x_1,2r)\times\{t_1\})\cap S_\square)}^2.
\end{aligned} \tag{7.67}$$

Claim 7.9 *For each ℓ', it holds that*

$$\|T^\lambda f_{\kappa,S_\square,\tau,\ell'}^{\approx,*}\|_{L^2((B(x_1,2r)\times\{t_1\})\cap S_\square)}^2$$

$$\lesssim (\frac{r}{R})^{\frac{n-n'}{2}} \|T^\lambda f_{\kappa,S_\square,\tau,\ell'}^{\approx}\|_{L^2(\{t=t_2\})}^2 \tag{7.68}$$

where $t_2 := t_1 + R$.

We first accept Claim 7.9, and continue with the L^2 estimate:

$$(7.68) \lesssim (\frac{r}{R})^{\frac{n-n'}{2}} \|f_{\kappa,S_\square,\tau,\ell'}^{\approx}\|_{L^2}^2. \tag{7.69}$$

Summing over all ℓ' and $S_\square$, we obtain

$$\|T^\lambda f_{S,\tau}^{\approx}\|_{L^2(B(x_1,2r)\times\{t_1\})}^2 \lesssim \ell_j (\frac{r}{R})^{\frac{n-n'}{2}} \|f_{\kappa,S,\tau}^{\approx}\|_{L^2}^2.$$

By Lemma 7.10 below and the assumption that all wave packets have comparable coefficients, we conclude that

$$\|f_{\kappa,S,\tau}^{\approx}\|_{L^2}^2 \lesssim \ell_j^{-1} R^{O(n\delta)} \|f_\tau\|_{L^2}^2.$$

This finishes the proof of the theorem, modulo the proofs of Lemma 7.10 and Claim 7.9. □

Lemma 7.10 *Let $\kappa = ((\ell_1, b_1), \mu_1, \ldots, (\ell_j, b_j), \mu_j)$ be an admissible multi-index and $\mathbb{T}_{S,\kappa,\tau}^{\approx}$ be defined as in (7.46). We have*

$$|\mathbb{T}_{S,\kappa,\tau}^{\approx}| \lesssim \ell_j^{-1} R^{O(n\delta)} |\mathbb{T}_\tau|,$$

where $\mathbb{T}_\tau$ collects tubes $T \in \mathbb{T}[B_R]$ with $\theta(T) \subset \tau$.

Proof of Lemma 7.10 Since κ is admissible, there exists

$$\kappa' = ((\ell_1, b_1), \mu_1, \ldots, (\ell_j, b_{j'}), \mu_{j'}) \tag{7.70}$$

such that

$$\kappa = (\kappa', (\ell_{j'+1}, b_{j'+1}), \mu_{j'+1}, \ldots (\ell_j, b_j), \mu_j) \tag{7.71}$$

and

$$((\ell_{j'}, b_{j'}), \mu_{j'}) = ((\ell_j, b_j), \mu_j). \tag{7.72}$$

Let B be a ball of radius $R^{1-\delta}$ containing S and $\mathfrak{n}$ be the node containing S, then for each $S \in \mathfrak{n}$ and $S' \not\subset 2B$, we have

$$\sum_{T \nsim B, T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa'}(S', T) \chi_{\mathfrak{n},\kappa}(S, T) \lesssim R^{O(n\delta)} \ell_{j'}^{-1} \sum_{T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa'}(S', T) \tag{7.73}$$

Summing over all $S' \in \mathfrak{n}$, $S' \not\subset 2B$,

$$
\begin{aligned}
\sum_{S \in \mathfrak{n}, S \not\subset 2B} &\sum_{T \nsim B, T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa}(S, T) \chi_{\mathfrak{n},\kappa'}(S', T) \\
&\leq R^{O(n\delta)} \ell_{j'}^{-1} \sum_{S \in \mathfrak{n}, S \not\subset 2B} \sum_{T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa'}(S', T).
\end{aligned}
\tag{7.74}
$$

On the other hand, for each $T \nsim B$ and $\chi_{\mathfrak{n},\kappa}(S, T) = 1$,

$$
\sum_{S' \in \mathfrak{n}, S' \not\subset 2B} \chi_{\mathfrak{n},\kappa'}(S', T) \geq \mu_j R^{-\delta}.
$$

As a consequence,

$$
|\mathbb{T}^{\sim}_{S,\kappa,\tau}| \leq R^\delta \mu_j^{-1} \sum_{S' \in \mathfrak{n}, S' \not\subset 2B} \sum_{T \nsim B, T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa}(S, T) \chi_{\mathfrak{n},\kappa'}(S', T).
$$

Moreover, note that

$$
\sum_{S' \in \mathfrak{n}, S' \not\subset 2B} \sum_{T \in \mathbb{T}_\tau} \chi_{\mathfrak{n},\kappa'}(S', T) \lesssim R^\delta \mu_{j'} |\mathbb{T}_\tau|
$$

The conclusion follows from (7.72). $\qquad\qquad\square$

In the rest of this section, we will prove Claim 7.9. We will start with the proof of Lemma 1.10, and this lemma will be an important ingredient in the forthcoming proof of Claim 7.9.

Proof of Lemma 1.10 Denote $\sigma := \sqrt{R_1}/\sqrt{R_2}$, $\Omega_1' = \mathcal{N}_1(Z_1)$ and $\Omega_2' := \mathcal{N}_\sigma(Z_2)$. By scaling, let us assume that $\mathrm{supp}(F) \subset \Omega_2'$, and we need to prove

$$
\left\| \widehat{F} \right\|_{L^2(\Omega_1')}^2 \lesssim \sigma^{n-m-\delta} \left\| F \right\|_{L^2}^2.
\tag{7.75}
$$

Let K be a large number depending on n, $\deg(Z_2)$, δ, which is to be determined. We cut Z_2 into $O_{n,\deg(Z_2),\delta}(1)$ many pieces so that for each piece $Z_2' \subset Z_2$, there exists a linear subspace of dimension $m - 1$ satisfying that the angle between $T_{\mathbf{z}}(Z_2')$ and the linear subspace is $\leq 1/K$ for every $\mathbf{z} \in Z_2'$. As our constant is allowed to depend on K, we only need to prove Lemma 1.10 for each Z_2'.

Claim 7.11 *Fix Z_2' as above and $K' \geq K$. Let $V \subset \mathbb{R}^{n-1}$ be an $(n-m)$-dimensional affine subspace such that the angle between $T_{\mathbf{z}}(Z_2')$ and $V^\perp$ is $\leq 1/K$ for every $\mathbf{z} \in Z_2'$. Let S_V denote the $1/K'$-neighbourhood of V. Then $Z_2' \cap S_V$ is contained in a union of $O(\deg(Z_2)^{n-1})$ many rectangular boxes of dimensions*

$$
\underbrace{\frac{1}{K'} \times \ldots \times \frac{1}{K'}}_{(m-1) \ copies} \times \underbrace{\frac{1}{KK'} \times \ldots \frac{1}{KK'}}_{(n-m) \ copies}
\tag{7.76}
$$

whose long sides are parallel to $V^\perp$.

Proof of Claim 7.11 Without loss of generality, we may assume $K = e^{-n}$ (all we need here is a small constant) and $K' = 1$ since we can do an (anisotropic) rescaling otherwise. Since the angle between every $T_{\mathbf{z}}(Z_2')$ and $V^{\perp}$ is $\leq e^{-n}$, we see the angle between every $T_{\mathbf{z}}(Z_2')$ and V is $\gtrsim 1$. Hence if we take the union of all points $\mathbf{z} \in Z_2 \bigcap S_V$ such that the angle between $T_{\mathbf{z}}(Z_2)$ and V is $\gtrsim 1$, it suffices to prove this whole set can be contained in a union of $O(\deg(Z_2)^{n-1})$ many unit balls. Note that when $n - m = 1$, this is proved by Guth as a special case of Lemma 5.7 in [9] (when one takes $r \simeq \alpha \simeq 1$ there). For general n and m this can also be proved by induction on dimension exactly in the same way as in the proof of Lemma 5.7 in [9] and we thus omit the details. $\qquad\square$

We continue to prove Lemma 1.10. We start with a trivial estimate:

$$\|\widehat{F}\|_{L^2(\mathcal{N}_1(Z_1))}^2 \leq \|\widehat{F}\|_{L^2(\mathcal{N}_K(Z_1))}^2. \tag{7.77}$$

For a given $K' \geq 1$, let $\mathcal{P}_{K'}$ be a partition of Ω_2' into disjoint pieces $\{\Omega_{2,K'}'\}$ so that the orthogonal projection of each $\Omega_{2,K'}'$ into $\mathrm{span}\{\vec{e}_1, \ldots, \vec{e}_{m-1}\}$ is roughly a dyadic cube of side length $1/K'$. Denote

$$F_{\Omega_{2,K'}'} = F \cdot \mathbb{1}_{\Omega_{2,K'}'}. \tag{7.78}$$

By local L^2 orthogonality (see for instance [8, Appendix B]) and Claim 7.11,

$$(7.77) \lesssim \sum_{\Omega_{2,K}'} \|\widehat{F_{\Omega_{2,K}'}}\|_{L^2(\mathcal{N}_K(Z_1))}^2. \tag{7.79}$$

When applying the local orthogonality lemma in [8, Appendix B], the L^2 norm on the right hand side of (7.79) carries a weight that is associated to $\mathcal{N}_K(Z_1)$ and decays rapidly outside $\mathcal{N}_K(Z_1)$; this is standard and we leave out the technical description. By the assumption of Lemma 1.10, we see that $\Omega_{2,K}'$ is contained in a rectangular box of dimensions

$$\underbrace{\frac{1}{K} \times \ldots \frac{1}{K}}_{(m-1) \text{ copies}} \times \underbrace{\frac{1}{K^2} \times \ldots \frac{1}{K^2}}_{(n-m) \text{ copies}}, \tag{7.80}$$

such that its long axes span a subspace that has angle $\lesssim 1/K$ against $\mathrm{span}\{\vec{e}_1, \ldots, \vec{e}_{m-1}\}$ the for every $\mathbf{z}$.

We now use the uncertainty principle to analyze $\widehat{F_{\Omega_{2,K}'}}$ again with the help of (the $K \simeq K' \simeq 1$ version of) Claim 7.11. Each $|\widehat{F_{\Omega_{2,K}'}}|$ is essentially a constant on dual boxes of dimensions $\underbrace{K \times \ldots K}_{(m-1) \text{ copies}} \times \underbrace{K^2 \times \ldots K^2}_{(n-m) \text{ copies}}$. If we use Claim 7.11, set all parameters in that claim to be $\simeq 1$ and rescale the conclusion by K times, we see that if we tile any given $\underbrace{K \times \ldots K}_{(m-1) \text{ copies}} \times \underbrace{K^2 \times \ldots K^2}_{(n-m) \text{ copies}}$ box as above by K-balls, then Z_1 only intersects $\simeq 1$ of them.

Therefore, the uncertainty principle (for instance the version in [8, Appendix B]) and the above geometric observation imply that

$$
\begin{aligned}
(7.79) &\lesssim \frac{1}{K^{n-m}} \sum_{\Omega'_{2,K}} \big\| \widehat{F_{\Omega'_{2,K}}} \big\|^2_{L^2(\mathcal{N}_{K^2}(Z_1))} \\
&\lesssim \frac{1}{K^{n-m}} \sum_{\Omega'_{2,K^2}} \big\| \widehat{F_{\Omega'_{2,K^2}}} \big\|^2_{L^2(\mathcal{N}_{K^2}(Z_1))},
\end{aligned}
\tag{7.81}
$$

where in the second inequality we use L^2 orthogonality. Similarly to (7.80), Ω'_{2,K^2} is contained in a rectangular box of dimensions

$$
\underbrace{\frac{1}{K^2} \times \cdots \frac{1}{K^2}}_{(m-1) \text{ copies}} \times \underbrace{\frac{1}{K^3} \times \cdots \frac{1}{K^3}}_{(n-m) \text{ copies}}
\tag{7.82}
$$

with nearby directions.

We continue this process repeatedly, and in the end arrive at

$$
\big\| \widehat{F} \big\|_{L^2(\mathcal{N}_1(Z))} \le C^W \Big(\frac{1}{K^W} \Big)^{n-m} \| F \|^2_{L^2},
\tag{7.83}
$$

where $K^W = 1/\sigma$. In the end, we pick K to be large enough. $\qquad\square$

Next, we will prove a simpler version of Claim 7.9, see Lemma 7.12 below. The proof of this lemma contains the main idea of that of Claim 7.9, and indeed we will apply Lemma 7.12 iteratively to prove Claim 7.9. To simplify notation, let us denote

$$
F(x,t) := \int f_\tau(\xi) e^{i\phi^\lambda(x,t;\xi)} a^\lambda(\mathbf{x};\xi) d\xi,
\tag{7.84}
$$

where τ is a frequency cap.

Lemma 7.12 *Let $\sqrt{R} \le r \le R'_1 \le R'_2 \le R$. Let τ be a frequency cap of side length $r^{-1/2}$. Given $t_1, t_2 \in [0, \lambda]$ with $R = |t_1 - t_2|$. Assume that we are given two $(m-1)$-dimensional algebraic varieties $Z_1 \subset \{t = t_1\}$ and $Z_2 \subset \{t = t_2\}$ satisfying that the angle formed by $T_{\mathbf{z}_i}(Z_i)$ and the space spanned by $\{\vec{e}_1, \ldots, \vec{e}_{m-1}\}$ is $\le 1/(100n)$, for every $i = 1, 2$ and every $\mathbf{z}_i \in Z_i$. Here $T_{\mathbf{z}_i}(Z)$ refers to the tangent space. Denote*

$$
\Omega_1 = \mathcal{N}_{\sqrt{R'_1}}(Z_1) \cap B_{\sqrt{R}},
\tag{7.85}
$$

for a given ball $B_{\sqrt{R}}$ of radius $\sqrt{R}$; denote

$$
\Omega_2 = \mathcal{N}_{R/\sqrt{R'_2}}(Z_2) \cap B_{R/\sqrt{r}}.
\tag{7.86}
$$

Assume that $F(\cdot, t_1)$ is essentially supported on $B_{\sqrt{R}}$ and $F(\cdot, t_2)$ is essentially supported on Ω_2, then

$$
\| F(x, t_1) \|^2_{L^2(\Omega_1)} \lesssim \Big(\frac{R'_1}{R'_2} \Big)^{\frac{n-m}{2} - O(\delta)} \| F(x, t_2) \|^2_{L^2(\Omega_2)}.
\tag{7.87}
$$

Proof of Lemma 7.12 Let x_0 be the center of $B_{\sqrt{R}}$ and ξ_0 the center of τ. We apply the same change of variables as in (7.54). Recall the new phase function in (7.57). If we denote [10]

$$F(x,t) := \int f_\tau(\xi) e^{i\phi_0^\lambda(x,t;\xi)} a^\lambda(\mathbf{x};\xi) d\xi, \tag{7.88}$$

then to prove the lemma, we need to prove it for the function $F(x,t)$ as in (7.88) with $t_1 = 0$, $x_0 = 0$ and $\xi_0 = 0$. Before we finish this reduction step, let us do another linear change of variables so that $\nabla_x \nabla_\xi \phi_0^\lambda(0;0)$ is the identity matrix.

We claim that $\check{f}_\tau$ is essentially supported on $2B_{\sqrt{R}}$. By the same Taylor expansion as in (7.60)–(7.62), we can write

$$\check{f}_\tau(x) = \int f_\tau(\xi) e^{ix\cdot\xi} d\xi = \int f_\tau(\xi) e^{i\phi_0^\lambda(x,0;\xi)} e^{-iw(x;\xi)} d\xi$$

$$= \sum_{k\in\mathbb{N}} \frac{(-i)^k}{k!} \int f_\tau(\xi) e^{i\phi_0^\lambda(x,0;\xi)} w^k(x;\xi) d\xi. \tag{7.89}$$

Next, we do Fourier expansion for $w^k(x;\xi) a^\lambda(x,0;\xi)$ in the ξ variable at the unit scale, and write it as

$$\sum_{\beta\in\mathbb{N}^{n-1}} c_{k,\beta}(x) e^{i\beta\cdot\xi}, \tag{7.90}$$

where the coefficients satisfy

$$|c_{k,\beta}(x)| \lesssim_N 2^k |\beta|^{-N}, \tag{7.91}$$

for every large N, uniformly in x. Now the claim that $\check{f}_\tau$ is essentially supported on $2B_{\sqrt{R}}$ follows from the assumption on F and the rapid decay in (7.91). Moreover, by a similar Taylor expansion, we obtain

$$\left\| F(x,t_1) \right\|_{L^2(\Omega_1)}^2 \lesssim \left\| \check{f}_\tau \right\|_{L^2(\Omega_1)}^2. \tag{7.92}$$

It therefore remains to control $\left\| \check{f}_\tau \right\|_{L^2(\Omega_1)}^2$ by $\left\| F(x,t_2) \right\|_{L^2(\Omega_2)}^2$.

Consider $t = t_2$, and write

$$F(x,t_2) = \int f_\tau(\xi) e^{i\phi^\lambda(x,t_2;\xi)} a^\lambda(x,t_2;\xi) d\xi$$

$$= \int \check{f}_\tau(y) \left(\int e^{-iy\cdot\xi} e^{i\phi_0^\lambda(x,t_2;\xi)} a^\lambda(x,t_2;\xi) d\xi \right) dy \tag{7.93}$$

Consider the critical point of the phase function:

$$\nabla_\xi \phi_0^\lambda(x,t_2;\xi) = y. \tag{7.94}$$

---

[10]Here we still use F just to avoid new notation.

Let $\xi_c = \xi_c(x, t_2; y)$ denote the critical point. Note that by (7.57) and the assumption that ϕ^λ is in its normal form, we see that

$$\nabla_\xi^2 \phi_0^\lambda(x, t_2; \xi) = t_2 \cdot I_{(n-1)\times(n-1)} + \text{small perturbation} \tag{7.95}$$

where $I_{(n-1)\times(n-1)}$ is the identity matrix of order $(n-1)$. Denote

$$\psi_{t_2}^\lambda(x, y) = \phi_0^\lambda(x, t_2; \xi_c) - y \cdot \xi_c. \tag{7.96}$$

Then by the stationary phase principle (see for instance Sogge [19, Theorem 1.2.1]), we obtain

$$F(x, t_2) = t_2^{-\frac{n-1}{2}} \int \check{f}_\tau(y) e^{i\psi_{t_2}^\lambda(x,y)} a_{t_2}^\lambda(x, y) dy \tag{7.97}$$

where

$$a_{t_2}^\lambda(x, y) := a_{t_2}(\frac{x}{\lambda}, \frac{y}{\lambda}), \tag{7.98}$$

and a_{t_2} is a compactly supported smooth function in both variables. To continue, we apply Taylor expansion of $\psi_{t_2}^\lambda(x, y)$ in y. Take ∇_y on both sides of (7.96):

$$\nabla_y \psi_{t_2}^\lambda(x, y) = \nabla_\xi \phi_0^\lambda(x, t_2; \xi_c) \cdot \nabla_y \xi_c - \xi_c - y \cdot \nabla_y \xi_c = -\xi_c. \tag{7.99}$$

Hence

$$\nabla_y^2 \psi_{t_2}^\lambda = -\nabla_y \xi_c.$$

We claim that

$$|\nabla_y \xi_c| \leq \frac{1}{t_2}. \tag{7.100}$$

Here by $|\cdot|$ of the matrix $\nabla_y \xi_c$, we mean the maximum of all the entries. To see this, we go back to the definition of ξ_c in (7.94), and differentiate both side:

$$\nabla_\xi^2 \phi_0^\lambda(x, t_2; \xi_c) \nabla_y \xi_c = I_{(n-1)\times(n-1)}. \tag{7.101}$$

The claim now follows from (7.95). Moreover, taking ∇_x on both sides of $\nabla_\xi \phi_0^\lambda(x, t_2; \xi_c) = y$, we obtain

$$\nabla_x \nabla_\xi \phi_0^\lambda + \nabla_\xi^2 \phi_0^\lambda \cdot \nabla_x \xi_c = 0. \tag{7.102}$$

If we keep differentiating both sides of (7.101) and (7.102) in x, y, then we will be able to obtain

$$|\nabla_x^\alpha \nabla_y^{\alpha'} \xi_c| \lesssim_\alpha t_2^{-\alpha} t_2^{-\alpha'}, \tag{7.103}$$

for every $\alpha, \alpha' \in \mathbb{N}$. The proof is left out.

From (7.100) and the fact that $\check{f}_\tau$ is essentially supported on $B_{\sqrt{R}}$, we see that we can "ignore" the quadratic term in the Taylor expansion of $\psi_{t_2}^\lambda(x, y)$ in the y variable. Write

$$F(x, t_2) = e^{i\psi_{t_2}^\lambda(x,0)} \int \check{f}_\tau(y) e^{iy \cdot \nabla_y \psi_{t_2}^\lambda(x,0) + iw_{3,t_2}(x,y)} a_{t_2}^\lambda(x, y) dy, \qquad (7.104)$$

for some error function $w_{3,t_2}(x, y)$. Next, we do Taylor's expansion for $\nabla_y \psi_{t_2}^\lambda(x, 0)$ in the x variable. Recall our assumption that $F(x, t_2)$ is essentially supported on a ball of radius $R/\sqrt{r}$; let x_0 denote its center. Write

$$y \cdot \nabla_y \psi_{t_2}^\lambda(x, 0) = y \cdot \nabla_y \psi_{t_2}^\lambda(x_0, 0) + y \cdot \nabla_x \nabla_y \psi_{t_2}^\lambda(x_0, 0)x + w_{4,t_2}(x, y), \quad (7.105)$$

for some error function $w_{4,t_2}(x, y)$. In particular,

$$|w_{4,t_2}(x, y)| \lesssim \frac{1}{R^2} \left(\frac{R}{\sqrt{r}}\right)^2 |y| \lesssim 1. \qquad (7.106)$$

Next, by (7.99), (7.102) and (7.95), we see that $\nabla_x \nabla_y \psi_{t_2}^\lambda(x_0, 0)$ is a small perturbation of $\frac{1}{t_2} I_{(n-1)\times(n-1)}$. The desired bound

$$\left\| \check{f}_\tau \right\|_{L^2(\Omega_1)}^2 \lesssim \left(\frac{R_1'}{R_2'}\right)^{\frac{n-m}{2} - O(\delta)} \left\| F(x, t_2) \right\|_{L^2(\Omega_2)}^2 \qquad (7.107)$$

now follows from Taylor's expansion and Lemma 1.10. $\square$

In the end of this section, we prove Claim 7.9. As mentioned above, the main idea is already explained in Lemma 1.10 and the proof of Lemma 7.12. The extra work we need to do is to take care of the refinement process of wave packets in the polynomial partitioning algorithm. In other words, in the partitioning algorithm, each time we have an algebraic dominating case, we will need to remove certain wave packets, and therefore the input function $f_{\kappa, S_\square, \tau, \ell'}^\infty$ also changes as the algorithm proceeds.

Proof of Claim 7.9 To simplify notation, let us write

$$f_R := f_{\kappa, S_\square, \tau, \ell'}^\infty. \qquad (7.108)$$

Here we use R to emphasize that the function f_R is built on wave packets from $\mathbb{T}[B_R]$. Our goal is to prove

$$\left\| T^\lambda (f_R)_{\iota,S}^* \right\|_{L^2((B(x_1, 2r) \times \{t_1\}) \cap S_\square)}^2 \lesssim \left(\frac{r}{R}\right)^{\frac{n-n'}{2}} \left\| T^\lambda f_R \right\|_{L^2(\{t=t_2\})}^2 \qquad (7.109)$$

We apply Lemma 7.1 and find an algebraic variety $Z_1 \subset \{t = t_1\}$ satisfying that the angle between $T_{\mathbf{z}_1}(Z_1)$ and $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}\}$ is $\leq 1/(100n)$, such that

$$(B(x_1, 2r) \times \{t_1\}) \cap S_\square \subset \mathcal{N}_{\sqrt{r}}(Z_1) \cap B_{\sqrt{R}} \cap \{t = t_1\} =: \Omega_1, \qquad (7.110)$$

for some ball $B_{\sqrt{R}}$ of radius $\sqrt{R}$. Moreover, by Lemma 7.2 and the definition of brooms, we can find an algebraic variety $Z_2 \subset \{t = t_1 + R\}$ satisfying that the angle between $T_{\mathbf{z}_2}(Z_2)$ and $\{\vec{e}_1, \ldots, \vec{e}_{n'-1}\}$ is $\leq 1/(100n)$ for every $\mathbf{z}_2 \in Z_2$, such that $T^\lambda f_R(\cdot, t_2)$ is essentially supported on

$$\mathcal{N}_{\sqrt{R}}(Z_2) \cap B_{R^{1+\delta}/\sqrt{r}} \cap \{t = t_2\} =: \Omega_2. \tag{7.111}$$

Under the above notation, (7.109) can be written as

$$\|T^\lambda (f_R)^*_{t,S}\|^2_{L^2(\Omega_1)} \lesssim (\frac{r}{R})^{\frac{n-n'}{2}} \|T^\lambda f_R\|^2_{L^2(\Omega_2)}. \tag{7.112}$$

To proceed, we need to recall notation and definitions from Sect. 5.4. Let $\mathfrak{n}^*_r \in \{\mathfrak{n}^*_0, \mathfrak{n}^*_1, \ldots\}$ be such that $S \in \mathfrak{n}^*_r$. Collect all the ancestors of $\mathfrak{n}^*_r$ that are in $\cup_j (\mathfrak{M}_j \cup \mathfrak{R}_j)$, and list them in descending order

$$\mathfrak{n}^*_{R_1}, \mathfrak{n}^*_{R_2}, \ldots, \mathfrak{n}^*_{R_{W-1}}, \tag{7.113}$$

where $R > R_1 > R_2 > \cdots > R_{W-1} > r$. Moreover, denote $R_0 := R$, $\mathfrak{n}^*_{R_0} := \mathfrak{n}^*_0$, $R_W := r$ and $\mathfrak{n}^*_{R_W} := \mathfrak{n}^*_r$. Note that we have the trivial bound $W \leq \delta_n^{-10}$. Next, find $S_{R_w} \in \mathfrak{n}^*_{R_w}$ for each $1 \leq w \leq W$ such that

$$S = S_{R_W} \subset S_{R_{W-1}} \subset \cdots \subset S_{R_1} \subset S_{R_0} = B_R, \tag{7.114}$$

and each S_{R_w} is contained in a ball B_{R_w} of radius R_w.

We will prove Claim 7.9 by applying Lemma 7.12 iteratively. Denote

$$\mathcal{N}_{\sqrt{R_w}}(\Omega_1) \cap \{t = t_1\} =: \Omega^{(1)}_{R_w}, \quad \mathcal{N}_{R/\sqrt{R_w}}(\Omega_2) \cap \{t = t_1 + R\} =: \Omega^{(2)}_{R_w}, \tag{7.115}$$

for every w. Recall that $T^\lambda f_{R_0}(\cdot, t_2)$ is supported on $\Omega^{(2)}_{R_0}$. By Lemma 7.12 with $m = n'$, $R'_1 = R_1$ and $R'_2 = R_0$, we obtain

$$\left\| T^\lambda f_{R_0} \right\|_{L^2(\Omega^{(1)}_{R_1})} \lesssim \left(\frac{R_1}{R_0} \right)^{\frac{n-n'}{2} - O(\delta)} \left\| T^\lambda f_{R_0} \right\|_{L^2(\Omega^{(2)}_{R_0})} \tag{7.116}$$

On the ball B_{R_1}, we have the new wave packet decomposition

$$f_{R_0} = \sum_{T \in \mathbb{T}[B_{R_1}]} (f_{R_0})_T. \tag{7.117}$$

Let $\mathbb{T}'[B_{R_1}]$ be a subset of $\mathbb{T}[B_{R_1}]$ such that

$$(f_{R_0})^*_{t, S_{R_1}} = \sum_{T \in \mathbb{T}'[B_{R_1}]} (f_{R_0})_T, \tag{7.118}$$

that is, we throw away certain wave packets, depending on whether we are in the tangential case or the transverse case. Denote

$$f_{R_1} := \sum_{\substack{T \in \mathbb{T}'[B_{R_1}] \\ T \cap \Omega^{(1)}_{R_1} \neq \emptyset}} (f_{R_0})_T. \tag{7.119}$$

Therefore by L^2 orthogonality, we have

$$\left\| T^\lambda f_{R_1} \right\|_{L^2(\Omega_{R_1}^{(1)})} \lesssim \left\| T^\lambda f_{R_0} \right\|_{L^2(\Omega_{R_1}^{(1)})}. \tag{7.120}$$

Note that by the construction of f_{R_1}, the function $T^\lambda f_{R_1}(\cdot, t_1)$ is essentially supported on $\Omega_{R_1}^{(1)}$. Therefore, by (7.53) and L^2 bounds for Hörmander's operator at fixed time (see for instance Hörmander [15]), we have

$$\left\| T^\lambda f_{R_1}(x, t_2) \right\|_{L_x^2(\mathbb{R}^{n-1})} \lesssim \left\| f_{R_1} \right\|_2 \lesssim \left\| T^\lambda f_{R_1} \right\|_{L^2(\Omega_{R_1}^{(1)})}. \tag{7.121}$$

The main observation is that $T^\lambda f_{R_1}(\cdot, t_2)$ is essentially supported on $\Omega_{R_1}^{(2)}$. Once this is proven, we see that we can repeat the above process: By Lemma 7.12 with $R_1' = R_2$ and $R_2' = R_1$, we obtain

$$\begin{aligned} \left\| T^\lambda f_{R_1} \right\|_{L^2(\Omega_{R_2}^{(1)})} &\lesssim \left(\frac{R_2}{R_1} \right)^{\frac{n-n'}{2} - O(\delta)} \left\| T^\lambda f_1 \right\|_{L^2(\Omega_{R_1}^{(2)})} \\ &\lesssim \left(\frac{R_2}{R_0} \right)^{\frac{n-n'}{2} - O(\delta)} \left\| T^\lambda f_{R_0} \right\|_{L^2(\Omega_{R_0}^{(2)})}. \end{aligned} \tag{7.122}$$

We define f_{R_2} similarly as above, and then observe that $T^\lambda f_{R_2}(\cdot, t_2)$ is essentially supported on $\Omega_{R_2}^{(2)}$. This allows us to repeat the above iteration one more time. In the end, we will obtain the desired bound (7.112).

It remains to prove that $T^\lambda f_{R_w}(\cdot, t_2)$ is essentially supported on $\Omega_{R_w}^{(2)}$ for every w. We will only prove the case $w = 1$, and the other cases are the same. Recall the definition of f_{R_1} from (7.119). Write

$$f_{R_0} = \sum_{T_0 \in \mathbb{T}'[B_{R_0}]} f_{T_0}, \tag{7.123}$$

for some $\mathbb{T}'[B_{R_0}] \subset \mathbb{T}[B_{R_0}]$. Denote

$$f_{T_0, R_1} := \sum_{\substack{T \in \mathbb{T}'[B_{R_1}] \\ T \cap \Omega_{R_1}^{(1)} \neq \emptyset}} (f_{T_0})_T. \tag{7.124}$$

It suffices to prove that for each T_0, we have

$$\operatorname{supp} T^\lambda f_{T_0, R_1}(\cdot, t_2) \subset \mathcal{N}_{R/\sqrt{R_1}}(\operatorname{supp} T^\lambda f_{T_0}(\cdot, t_2)), \tag{7.125}$$

where by $\mathcal{N}$ we mean neighborhood in $(n-1)$ dimensions. Note that f_{T_0, R_1} consists of wave packets of frequency scale $R_1^{-1/2}$. In order to see where $T^\lambda f_{T_0, R_1}(\cdot, t_2)$ is supported, we do a wave packet decomposition for f_{T_0, R_1} by using wave packets of frequency scale $R_0^{-1/2}$:

$$f_{T_0, R_1} = \sum_{T_0' \in \mathbb{T}[B_{R_0}]} (f_{T_0, R_1})_{T_0'}. \tag{7.126}$$

In order for $T_0' \in \mathbb{T}[B_{R_0}]$ to have non-trivial contribution, we need that $T_0 \cap T_0' \neq \emptyset$ and $\text{dist}(\theta(T_0), \theta(T_0')) \lesssim R_1^{-1/2}$. Under these two conditions, by the same Taylor expansion argument as in Lemma 7.2, (7.125) follows immediately. $\qquad\square$

8 Bushes: small grains

Let $\vec{S}$ be a grain with its last component given by $(S, B(\mathbf{x}_0, r))$. In the previous section, we considered the case $r \geq \sqrt{R}$. In this section, we will consider the case $r \leq \sqrt{R}$. As the grain S is small, in particular, it is smaller than the scale $\sqrt{R}$ of a wave packet $T \in \mathbb{T}[B_R]$, we will see that this case is much easier to handle.

The goal of this section is to prove the following result.

Theorem 8.1 *Let* $\vec{S}_{n'} = (S_n, \ldots, S_{n'})$ *be a multi-grain with* $S_{n'} = \mathcal{N}_{r_{n'}^{1/2+\delta_{n'}}}(Z_{n'}) \cap B_{r_{n'}}$, $r_{n'} \leq \sqrt{R}$ *and* $Z_{n'}$ *an* n'-*dimensional algebraic variety of degree* $\lesssim_{n'} 1$ *in* $\mathbb{R}^n$. *Then for frequency caps* τ *of side length* $r_{n'}^{-1/2}$ *and* $n' \leq n'' \leq n$, *we have*

$$\|f_{\vec{S}_{n'}, \tau}^{\varpropto(n'')}\|_{L^2}^2 \lessapprox \left(\frac{r_{n''}}{R}\right)^{\frac{n-n'}{2}} \|f_\tau\|_{L^2}^2, \tag{8.1}$$

where $f_{\vec{S}_{n'}, \tau}^{\varpropto(n'')}$ *is defined in* (7.42) *and* (7.43), *with the definition of* $\varpropto$ *given in Definition* 8.3 *below.*

In Theorem 8.1, the scale $r_{n'}$ is so small that we do not see the broom structure; we will replace brooms by bushes.

Definition 8.2 (Bushes) Given a grain $(S, B(\mathbf{x}_0, r))$ with $r \leq \sqrt{R}$, a collection of tubes $\mathbb{T}'[B_R] \subset \mathbb{T}[B_R]$ and a frequency cap τ of side length $r^{-1/2}$, the collection of tubes

$$\mathcal{B}(\mathbb{T}'[B_R]) := \{T \in \mathbb{T}'[B_R] : \theta(T) \subset \tau, T \cap S \neq \emptyset\} \tag{8.2}$$

is called a bush generated by $\mathbb{T}'[B_R]$. The size b of the bush $\mathcal{B}(\mathbb{T}'[B_R])$ is defined to be $R^{w\delta}$ with $w \in \mathbb{N}$, $R^{w\delta} \leq \#\mathcal{B}(\mathbb{T}'[B_R]) < R^{(w+1)\delta}$. Often $\mathcal{B}(\mathbb{T}'[B_R])$ will be written as $\mathcal{B}_b(\mathbb{T}'[B_R])$. In particular, if $T'[B_R] = \mathbb{T}[B_R]$, then we will simply write $\mathcal{B}_b$ for a bush.

Next, we define a two-ends relation $\sim_{\mathfrak{n}}$ for nodes $\mathfrak{n} \in \cup_l \mathfrak{R}_l$ with $\rho(\mathfrak{n}) \leq \sqrt{R}$. Similarly as above, we start by introducing a few auxiliary functions $\chi_{\mathfrak{n},\kappa} = \chi_\kappa$, taking values 0 or 1, where $\kappa = (b_1, \mu_1, \ldots, b_\ell, \mu_\ell)$ and $b_{\ell'}, \mu_{\ell'} \in \{R^{w\delta} : w \in \mathbb{N}\}$ for every $1 \leq \ell' \leq \ell$. Denote $r = \rho(\mathfrak{n})$. For $S \in \mathfrak{n}$ and a tube $T \in \mathbb{T}[B_R]$, we say that

$$\chi_{b_1}(S, T) = 1, \tag{8.3}$$

if T belongs to a bush of size b_1 rooted at S. Moreover, we say that

$$\chi_{b_1, \mu_1}(S, T) = 1 \tag{8.4}$$

if

$$\chi_{b_1}(S, T) = 1, \quad \mu_1 \leq \sum_{S' \in \mathfrak{n}} \chi_{b_1}(S', T) < \mu_1 R^\delta. \tag{8.5}$$

Now let us assume that we have defined χ_κ with $\kappa = (b_1, \mu_1, \ldots, b_\ell, \mu_\ell)$ already, and we would like to define $\chi_{\kappa, b_{\ell+1}}$ and $\chi_{\kappa, b_{\ell+1}, \mu_{\ell+1}}$. For a fixed $S \in \mathfrak{n}$, define

$$\mathbb{T}_{S,\kappa} := \{T' \in \mathbb{T}[B_R] : \chi_\kappa(S, T') = 1\}. \tag{8.6}$$

We write $\mathbb{T}_{S,\kappa}$ as a disjoint union of bushes

$$\bigcup_{b_{\ell+1}} \bigcup_\tau \mathcal{B}_{b_{\ell+1}, \tau}(\mathbb{T}_{S,\kappa}) \tag{8.7}$$

where τ runs through all caps of side length $\rho(\mathfrak{n})^{-1/2}$, and $\mathcal{B}_{b_{\ell+1}, \tau}(\mathbb{T}_{S,\kappa})$ is a bush of size $b_{\ell+1}$ with tubes coming from τ. We say that

$$\chi_{\kappa, b_{\ell+1}}(S, T) = 1, \text{ if } T \in \mathcal{B}_{b_{\ell+1}, \tau}(\mathbb{T}_{S,\kappa}), \tag{8.8}$$

for some τ. Moreover,

$$\chi_{\kappa, b_{\ell+1}, \mu_{\ell+1}}(S, T) = 1, \tag{8.9}$$

if

$$\chi_{\kappa, b_{\ell+1}}(S, T) = 1, \quad \mu_{\ell+1} \leq \sum_{S' \in \mathfrak{n}} \chi_{\kappa, b_{\ell+1}}(S', T) < \mu_{\ell+1} R^\delta. \tag{8.10}$$

This finishes the definition of the auxiliary functions.

For $\kappa = (b_1, \mu_1, \ldots, b_\ell, \mu_\ell)$, we say that κ is admissible if there exists exactly one pair (ℓ_1, ℓ_2) with $\ell_1 \neq \ell_2$ such that

$$(b_{\ell_1}, \mu_{\ell_1}) = (b_{\ell_2}, \mu_{\ell_2}). \tag{8.11}$$

Similarly to Lemma 7.4, it is elementary to see that the number of admissible κ is $O_\delta(1)$.

Definition 8.3 For a ball $B \subset B_R$ of radius $R^{1-\delta}$, a tube $T \in \mathbb{T}[B_R]$, a node $\mathfrak{n} \in \cup_l \mathfrak{R}_l$ with $\rho(\mathfrak{n}) \leq \sqrt{R}$ and an admissible κ, we say that $T \sim_{\mathfrak{n}, \kappa} B$ if B maximizes

$$\#\{S' \in \mathfrak{n} : S' \subset B', \chi_{\mathfrak{n}, \kappa}(S', T) = 1\}, \tag{8.12}$$

among all B' of radius $R^{1-\delta}$. Moreover, we say that $T \sim_{\mathfrak{n}} B$ if $T \sim_{\mathfrak{n}, \kappa}$ for some admissible κ; we say that $T \sim B$ if $T \sim_{\mathfrak{n}} B$ for some node $\mathfrak{n} \in \cup_l \mathfrak{R}_l$.

Now we are ready to prove Theorem 8.1.

Proof of Theorem 8.1. Similarly to what we did in the proof of Theorem 7.8, we will write down more details for the case $n'' = n'$, and the case $n'' > n'$ is essentially the same.

We abbreviate $S_{n'}$ to S, $\vec{S}_{n'}$ to $\vec{S}$ and $r_{n'}$ to r. Let B_ι be the ball of radius $R^{1-\delta}$ containing S; let $\mathfrak{n}$ be the node such that $S \in \mathfrak{n}$. For an admissible multi-index κ of the form $(b_1, \mu_1, \ldots, b_\ell, \mu_\ell)$, denote

$$\mathbb{T}_{S,\kappa,\tau} := \{T \in \mathbb{T}_{S,\kappa} : \theta(T) \subset \tau\}, \tag{8.13}$$

$$\mathbb{T}_{S,\kappa,\tau}^\infty := \{T \in \mathbb{T}_{S,\kappa} : \theta(T) \subset \tau, B_\iota \sim_{\mathfrak{n},\kappa} T\}, \tag{8.14}$$

and

$$f_{\kappa,S,\tau}^\infty := \sum_{T \in \mathbb{T}_{S,\kappa,\tau}^\infty} f_T \quad \text{and} \quad f_{\kappa,S,\tau}^{\infty,*} := (f_{\kappa,S,\tau}^\infty)_{\iota,S}^*. \tag{8.15}$$

Then similarly to (7.47), we have

$$f_{S,\tau}^\infty = \sum_\kappa f_{\kappa,S,\tau}^{\infty,*}. \tag{8.16}$$

Next, similarly to (7.48) and (7.52), we have

$$\left\| f_{S,\tau}^\infty \right\|_2^2 \lesssim_\delta \sum_\kappa \left\| T^\lambda f_{\kappa,S,\tau}^{\infty,*} \right\|_{L^2(\{t=t_1\} \cap S)}^2, \tag{8.17}$$

where (x_1, t_1) is a point in S. By the definition of κ, $\mathbb{T}_{S,\kappa,\tau}$ is contained in a bush $\mathcal{B}_{b_\ell}$. Write

$$f_{\kappa,S,\tau}^{\infty,*} = \sum_{T \in \mathbb{T}_{S,\kappa,\tau}^\infty \cap \mathcal{B}_{b_\ell}} (f_T)_{\iota,S}^*. \tag{8.18}$$

By the Cauchy-Schwarz inequality, we have

$$\left\| T^\lambda f_{\kappa,S,\tau}^{\infty,*} \right\|_{L^2(\{t=t_1\} \cap S)}^2 \lesssim b_\ell \sum_{T \in \mathbb{T}_{S,\kappa,\tau}^\infty \cap \mathcal{B}_{b_\ell}} \left\| T^\lambda (f_T)_{\iota,S}^* \right\|_{L^2(\{t=t_1\} \cap S)}^2. \tag{8.19}$$

It is elementary to see that[11]

$$\left\| T^\lambda (f_T)_{\iota,S}^* \right\|_{L^2(\{t=t_1\} \cap S)}^2 \lesssim \left(\frac{r}{R} \right)^{\frac{n-n'}{2}} \left\| f_T \right\|_{L^2}^2. \tag{8.20}$$

By summing over T, we obtain

$$\left\| T^\lambda f_{\kappa,S,\tau}^{\infty,*} \right\|_{L^2(\{t=t_1\} \cap S)}^2 \lesssim b_\ell \left(\frac{r}{R} \right)^{\frac{n-n'}{2}} \left\| f_{\kappa,S,\tau}^\infty \right\|_{L^2}^2. \tag{8.21}$$

---

[11] This can be viewed as a trivial version of Claim 7.9.

In the end, we just need to show that

$$|\mathbb{T}^{\infty}_{S,\kappa,\tau}| \lesssim b_{\ell}^{-1} R^{O(n\delta)} |\mathbb{T}_{\tau}|, \tag{8.22}$$

which is an analogue of Lemma 7.10. As the proof of (8.22) is also more or less the same as that of Lemma 7.10 (indeed simpler), we leave it out. $\qquad\square$

9 Finishing the proof of Theorem 4.2

In the last section, we combine the polynomial Wolff axiom in Lemma 6.1 and Corollary 6.5, Properties 1-4 in Sect. 5.4, the broom estimate in Theorem 7.8 and the bush estimate in Theorem 8.1 to finish the proof of Theorem 4.2.

First of all, by Property 1 and repeated applications of Property 2 in Sect. 5.4 in the same way as [12, page 269] obtain equation (56) and (57), we obtain

$$\|T^{\lambda} g\|_{\mathrm{BL}^{p}_{k,A}(B_R)} \lessapprox \prod_{i=m-1}^{n-1} r_i^{\frac{\beta_{i+1}-\beta_i}{2}} D_i^{\frac{\beta_{i+1}}{2} - \left(\frac{1}{2} - \frac{1}{p_n}\right)} \tag{9.1}$$

$$\|g\|_2^{\frac{2}{p_n}} \max_{O \in \mathfrak{n}^*_{\ell_0}} \|g_{\iota,O}\|_2^{1 - \frac{2}{p_n}}, \tag{9.2}$$

where ι refers to $B_{\iota} \subset B_R$, the ball of radius $R^{1-\delta}$ containing O, the parameter m comes from (5.86) and $\mathfrak{n}^*_{\ell_0}$ is as in (5.84). Next, by repeated applications of Property 3, we obtain

$$\max_{O \in \mathfrak{n}^*_{\ell_0}} \|g_{\iota,O}\|_2^2 \lessapprox r_{n'}^{-\frac{n-n'}{2}} \prod_{i=m-1}^{n'-1} r_i^{-1/2} D_i^{-i+\delta} \max_{S_{n'} \in \mathfrak{S}_{n'}} \left\|g^*_{\iota,S_{n'}}\right\|_2^2. \tag{9.3}$$

By Property 4,

$$\left\|g^*_{\iota,S_{n'}}\right\|_2^2 \lessapprox r_{n'}^{\frac{n-n'}{2}} \left(\prod_{i=n'}^{n-1} r_i^{-\frac{1}{2}}\right) r_{n''}^{-\frac{n-n''}{2}} \left(\prod_{i=n''}^{n-1} r_i^{\frac{1}{2}}\right) R^{O(\epsilon_{\circ})} \left\|g^{*(n'')}_{\iota,S_{n'}}\right\|_2^2. \tag{9.4}$$

We then apply Corollary 6.5 and obtain

$$\left\|g^{*(n'')}_{\iota,S_{n'}}\right\|_2^2 \lessapprox \left(\prod_{j=n'}^{n''} r_j^{-\frac{1}{2}}\right) r_{n''}^{-\frac{n-n''-1}{2}} \max_{\tau:\ell(\tau)=r_{n''}^{-1/2}} \left\|g^{*(n'')}_{\iota,S_{n'}}\right\|_{L^2_{\mathrm{avg}}(\tau)}^2. \tag{9.5}$$

Recall the notation (7.42) and (7.43). By the broom estimate in Theorem 7.8 and the bush estimate in Theorem 8.1, we have

$$\left\|g^{*(n'')}_{\iota,S_{n'}}\right\|_{L^2_{\mathrm{avg}}(\tau)}^2 \lessapprox \left(\frac{R}{r_{n''}}\right)^{-\frac{n-n'}{2}} \|g\|_{\infty}^2. \tag{9.6}$$

Putting everything together, we obtain

$$\max_O \|g_{\iota,O}\|_2^2 \lesssim_{\approx} R^{O(\epsilon_\circ)} \Big(\prod_{i=m-1}^{n'-1} D_i^{-i} \Big) \Big(\prod_{j=m-1}^{n'-1} r_j^{\frac{1}{2}} \Big) \Big(\prod_{i=m-1}^{n-1} r_i^{-1} \Big) \tag{9.7}$$

$$r_{n''}^{-(n-n''-1)} \Big(\prod_{i=n''+1}^{n-1} r_i \Big) \Big(\frac{R}{r_{n''}} \Big)^{-\frac{n-n'}{2}} \|g\|_\infty^2. \tag{9.8}$$

We pick n'' such that

$$n - n'' = \Big[\frac{n-n'}{3} \Big] + 1, \tag{9.9}$$

bound r_i by R and obtain

$$\max_O \|g_{\iota,O}\|_2^2 \lesssim_{\approx} \Big(\prod_{i=m-1}^{n-1} r_i^{-1} \Big) \Big(\prod_{i=m-1}^{n'-1} r_i^{1/2} \Big) \Big(\prod_{i=m-1}^{n'-1} D_i^{-i} \Big) \Big(\frac{R}{r_{n''}} \Big)^{-\frac{n-n'}{6}} \|g\|_\infty^2. \tag{9.10}$$

Note that when

$$n' \le n/100 + 99m/100 =: W_m^n, \tag{9.11}$$

it holds that

$$n'' \le \frac{2n}{3} + \frac{n'}{3} \le \frac{67n}{100} + \frac{33m}{100} =: M, \tag{9.12}$$

and therefore we have

$$\max_O \|g_{\iota,O}\|_2^2 \lesssim_{\approx} \Big(\prod_{i=m-1}^{n-1} r_i^{-1} \Big) \Big(\prod_{i=m-1}^{n'-1} r_i^{1/2} \Big) \Big(\prod_{i=m-1}^{n'-1} D_i^{-i} \Big) \tag{9.13}$$

$$\times \begin{cases} \Big(\dfrac{R}{r_M} \Big)^{-\frac{33(n-m)}{200}} \|g\|_\infty^2 & \text{if } n' \le W_m^n, \\ \|g\|_\infty^2, & \text{otherwise} \end{cases} \tag{9.14}$$

By taking a weighted geometric average in $n' \in \{m, m+1, \dots, n-1\}$, and substituting into (9.1), we obtain

$$\big\| T^\lambda g \big\|_{\mathrm{BL}_{k,A}^p(B_R)} \lesssim_{\approx} R^{-\Lambda} \Big(\prod_{i=m-1}^{n-1} r_i^{X_i} D_i^{Y_i} \Big) \|g\|_2^{\frac{2}{p}} \|g\|_\infty^{1-\frac{2}{p}}, \tag{9.15}$$

where

$$\Lambda = \Big(\sum_{j=m}^{W_m^n} \gamma_j \Big) \Big(\frac{1}{2} - \frac{1}{p} \Big) \frac{33(n-m)}{200}, \tag{9.16}$$

$$X_i' = \frac{\beta_{i+1} - \beta_i}{2} - \left(\frac{1}{2} - \frac{1}{p}\right) + \frac{1}{2}\left(1 - \sum_{j=m}^{i} \gamma_j\right)\left(\frac{1}{2} - \frac{1}{p}\right), \qquad (9.17)$$

$$X_i = X_i' + \begin{cases} 0 \text{ if } i \neq M, \\ \Lambda \text{ if } i = M, \end{cases} \qquad (9.18)$$

$$Y_i = \frac{\beta_{i+1}}{2} - \left(1 + i(1 - \sum_{j=m}^{i} \gamma_j)\left(\frac{1}{2} - \frac{1}{p}\right)\right) \qquad (9.19)$$

and

$$\gamma_{m-1} := 0, \; 0 \leq \gamma_m, \ldots, \gamma_n \leq 1, \; \gamma_m + \cdots + \gamma_n = 1, \qquad (9.20)$$

$$r_{m-1} = 1, \; D_n = 1, \; \beta_n = 1, \; \beta_n \geq \beta_{n-1} \geq \cdots \geq \beta_m. \qquad (9.21)$$

Lemma 5.10 suggests that we write the coefficients on the right hand side of (9.15) as

$$R^{-\Lambda}(D_{m-1})^{\frac{\beta_m}{2} - m(\frac{1}{2} - \frac{1}{p})}\left(\prod_{i=m}^{n-1} D_i^{Y_i - (\sum_{j=m}^{i} X_j)}\right)\prod_{i=m}^{n-1}\left(r_i \prod_{j=i}^{n-1} D_j\right)^{X_i}. \qquad (9.22)$$

We pick γ_i and β_i so that

$$Y_{n-1} - X_{n-1} - \cdots - X_m = \cdots = Y_m - X_m = \frac{\beta_m}{2} - m(\frac{1}{2} - \frac{1}{p}) = 0, \qquad (9.23)$$

$$X_M = \Lambda, \; X_i = 0, \text{ when } i \neq M. \qquad (9.24)$$

One can check directly that if we set $p_m = 2m/(m-1)$,

$$\gamma_i' := \frac{1}{2(i-1)} \prod_{j=m}^{i} \frac{2(j-1)}{2j+1}, \forall m \leq i \leq n-1, \qquad (9.25)$$

and let γ_i be given by the solution to the system

$$\gamma_i = \gamma_i', \; \forall m \leq i \leq M-1,$$

$$(M + \frac{1}{2})(\gamma_M - \gamma_M') = \Gamma_m^n \frac{33(n-m)}{200}, \; \text{with } \Gamma_m^n := \sum_{j=m}^{W_m^n} \gamma_j,$$

$$(M + \frac{2i+1}{2})(\gamma_{M+i} - \gamma_{M+i}') + \frac{3}{2} \sum_{j=M}^{M+i-1} (\gamma_j - \gamma_j') = 0, \; \forall 1 \leq i \leq n - M - 1,$$

$$(9.26)$$

then (9.23) is satisfied.

Claim 9.1 *Let γ_i be given as above. Then $\gamma_i \geq 0$ for every $m \leq i \leq n-1$ and*

$$\sum_{i=m}^{n-1} \gamma_i \leq 1. \tag{9.27}$$

Proof of Claim 9.1 By taking the logarithm and Taylor's expansion, we see that

$$\sum_{i=m}^{n-1} \gamma_i' \leq 1/3. \tag{9.28}$$

Moreover, under the assumption that $k \geq 2n/5$ in Theorem 4.2, it holds that

$$\begin{aligned}
\gamma_i' &= \frac{1}{2(i-1)} \frac{2(m-1)2m \ldots 2(i-1)}{(2m+1)(2m+3) \ldots (2i+1)} \\
&\geq \frac{1}{2(i-1)} \frac{2(m-1)2m}{(2i-1)(2i+1)} \geq \frac{2n}{25},
\end{aligned} \tag{9.29}$$

and

$$\gamma_M - \gamma_M' = (M+1/2)^{-1} \Gamma_m^n \frac{33(n-m)}{200} \leq \frac{1}{24}. \tag{9.30}$$

The claim now follows from checking the system (9.26). $\qquad\square$

So far we have picked the values for γ_i with $m \leq i < n-1$ and β_m. By Claim 9.1, we can choose $\gamma_n = 1 - \gamma_m - \cdots - \gamma_{n-1}$. Now we pick β_i satisfying (9.21) so that (9.24) is satisfied. Elementary computation shows that

$$\frac{1}{2} = \left(\frac{1}{2} - \frac{1}{p}\right)\left(\frac{n+m}{2} + \frac{1}{2}\sum_{j=m}^{n-1}(n-j)\gamma_j\right) \tag{9.31}$$

Therefore, to prove Theorem 4.2, it remains to prove that $p \leq p_n(k)$, where p is given in (9.31). This can done by using the trick in the appendix of [13], together with elementary but tedious computation, which will be skipped.

References

1. Basu, S., Pollack, R., Roy, M.-F.: Algorithms in Real Algebraic Geometry, 2nd edn. Algorithms and Computation in Mathematics, vol. 10. Springer, Berlin (2006)
2. Bennett, J., Nakamura, S., Shiraki, S.: Tomographic Fourier extension identities for submanifolds of $\mathbb{R}^n$ (2022). ArXiv preprint. arXiv:2212.12348
3. Bourgain, J.: L^p-estimates for oscillatory integrals in several variables. Geom. Funct. Anal. **1**(4), 321–374 (1991)
4. Bourgain, J., Guth, L.: Bounds on oscillatory integral operators based on multilinear estimates. Geom. Funct. Anal. **21**(6), 1239–1295 (2011)
5. Carleson, L., Sjölin, P.: Oscillatory integrals and a multiplier problem for the disc. Studia Math. **44**, 287–299 (1972)
6. Fefferman, C.: A note on spherical summation multipliers. Israel J. Math. **15**, 44–52 (1973)
7. Gao, C., Li, J., Wang, L.: A type of oscillatory integral operator and its applications. Math. Z. **302**(3), 1551–1584 (2022)
8. Guo, S., Zorin-Kranich, P.: Decoupling for moment manifolds associated to Arkhipov–Chubarikov–Karatsuba systems. Adv. Math. **360**, Article ID 106889 (2020)
9. Guth, L.: Restriction estimates using polynomial partitioning ii. Acta Math. **221**(1), 81–142 (2018)
10. Guth, L., Hickman, J., Iliopoulou, M.: Sharp estimates for oscillatory integral operators via polynomial partitioning. Acta Math. **223**(2), 251–376 (2019)
11. Hawk Katz, N., Rogers, K.M.: On the polynomial Wolff axioms. Geom. Funct. Anal. **28**(6), 1706–1716 (2018)
12. Hickman, J., Rogers, K.M.: Improved Fourier restriction estimates in higher dimensions. Camb. J. Math. **7**(3), 219–282 (2019)
13. Hickman, J., Zahl, J.: A note on Fourier restriction and nested polynomial wolff axioms (2020). arXiv:2010.02251
14. Hickman, J., Rogers, K., Zhang, R.: Improved bounds for the Kakeya maximal conjecture in higher dimensions (2019). arXiv:1908.05589
15. Hörmander, L.: Oscillatory integrals and multipliers on FL^p. Ark. Mat. **11**, 1–11 (1973)
16. Huang, S., Sogge, C.D.: Concerning L^p resolvent estimates for simply connected manifolds of constant curvature. J. Funct. Anal. **267**(12), 4635–4666 (2014)
17. Kenig, C.E., Ruiz, A., Sogge, C.D.: Uniform Sobolev inequalities and unique continuation for second order constant coefficient differential operators. Duke Math. J. **55**(2), 329–347 (1987)
18. Sogge, C.D.: Oscillatory integrals and spherical harmonics. Duke Math. J. **53**(1), 43–65 (1986)
19. Sogge, C.D.: Fourier Integrals in Classical Analysis. Cambridge Tracts in Mathematics, vol. 105. Cambridge University Press, Cambridge (1993)
20. Sogge, C.D.: Concerning Nikodým-type sets in 3-dimensional curved spaces. J. Amer. Math. Soc. **12**(1), 1–31 (1999)
21. Wang, H.: A restriction estimate in $\mathbb{R}^3$ using brooms (2018). ArXiv preprint. arXiv:1802.04312
22. Wongkew, R.: Volumes of tubular neighbourhoods of real algebraic varieties. Pacific J. Math. **159**(1), 177–184 (1993)
23. Yomdin, Y., Comte, G.: Tame Geometry with Application in Smooth Analysis. Lecture Notes in Mathematics, vol. 1834. Springer, Berlin (2004)
24. Zahl, J.: New Kakeya estimates using Gromov's algebraic lemma. Adv. Math. **380**, 107596 (2021)