

On the multilinear restriction and Keakeya conjectures

by

JONATHAN BENNETT

*University of Birmingham
Birmingham, U.K.*

ANTHONY CARBERY

*University of Edinburgh
Edinburgh, U.K.*

TERENCE TAO

*University of California
Los Angeles, CA, U.S.A.*

1. Introduction

For $d \geq 2$, let U be a compact neighbourhood of the origin in \mathbf{R}^{d-1} and $\Sigma: U \rightarrow \mathbf{R}^d$ be a smooth parametrisation of a $(d-1)$ -dimensional submanifold S of \mathbf{R}^d (for instance, S could be a small portion of the unit sphere \mathbf{S}^{d-1}). With Σ we associate the *extension operator* \mathcal{E} , given by

$$\mathcal{E}g(\xi) := \int_U g(x) e^{i\xi \cdot \Sigma(x)} dx,$$

where $g \in L^1(U)$ and $\xi \in \mathbf{R}^d$. This operator is sometimes referred to as the *adjoint restriction operator*, since its adjoint \mathcal{E}^* is given by $\mathcal{E}^*f = \widehat{f \circ \Sigma}$, where $\widehat{\cdot}$ denotes the d -dimensional Fourier transform. It was observed by E.M. Stein in the late 1960s that if the submanifold parametrised by Σ has everywhere nonvanishing gaussian curvature, then nontrivial $L^p(U) \rightarrow L^q(\mathbf{R}^d)$ estimates for \mathcal{E} may be obtained. The classical *restriction conjecture* concerns the full range of exponents p and q for which such bounds hold.

CONJECTURE 1.1. (Linear restriction) *If S has everywhere nonvanishing gaussian curvature, $q > 2d/(d-1)$ and $p' \leq q(d-1)/(d+1)$, then there exists a constant $0 < C < \infty$ depending only on q , d and Σ such that*

$$\|\mathcal{E}g\|_{L^q(\mathbf{R}^d)} \leq C \|g\|_{L^p(U)}$$

The first author was supported by EPSRC Postdoctoral Fellowship GR/S27009/02, the second by EC project ‘‘HARP’’ and a Leverhulme Study Abroad Fellowship, and the third by a grant from the Packard Foundation.

for all $g \in L^p(U)$. (Here, as usual, p' denotes the exponent dual to p .)

See for example [29] for a discussion of the progress made on this problem, the rich variety of techniques that have developed in its wake, and the connection to other problems in harmonic analysis, partial differential equations, and geometric analysis. In particular the restriction problem is intimately connected to the *Keakeya problem*, which we shall discuss later in this introduction.

In recent years, certain *bilinear* analogues of the restriction problem have come to light in natural ways from a number of sources (see for example [10], [19], [23], [30], [34], [27], [28], [31], [20], [14] concerning the well-posedness theory of nonlinear dispersive equations, and applications to a variety of central problems in harmonic and geometric analysis). More specifically, given two such smooth mappings Σ_1 and Σ_2 , with associated extension operators \mathcal{E}_1 and \mathcal{E}_2 , one may ask for which values of the exponents p and q the bilinear operator $(g_1, g_2) \mapsto \mathcal{E}_1 g_1 \mathcal{E}_2 g_2$ may be bounded from $L^p \times L^p$ to $L^{q/2}$. The essential point here is that if the submanifolds parametrised by Σ_1 and Σ_2 are assumed to be *transversal* (up to translations), then one can expect the range of such exponents to broaden; again see [30]. However, one of the more puzzling features of such bilinear problems is that, in three dimensions and above, they seem to somewhat confuse the role played by the curvature of the associated submanifolds. For example, it is known that the bilinear restriction theories for the cone and paraboloid are almost identical, whereas the linear theories for these surfaces are not (see [29] for further discussion of this). Moreover, simple heuristics suggest that the optimal “ k -linear” restriction theory requires at least $d-k$ nonvanishing principal curvatures, but that further curvature assumptions have no further effect. In d dimensions it thus seems particularly natural to consider a d -linear set-up, as one then does not expect to require any curvature conditions.

For each $1 \leq j \leq d$, let $\Sigma_j: U_j \rightarrow \mathbf{R}^d$ be such a smooth mapping and let \mathcal{E}_j be the associated extension operator. Our analogue of the bilinear transversality condition will essentially amount to requiring that the normals to the submanifolds parametrised by the Σ_j 's span at all points of the parameter space. In order to express this in an appropriately uniform manner, let $A, \nu > 0$ be given, and for each $1 \leq j \leq d$ let Y_j be the $(d-1)$ -form

$$Y_j(x) := \bigwedge_{k=1}^{d-1} \frac{\partial}{\partial x_k} \Sigma_j(x)$$

for all $x \in U_j$; by duality, we can view Y_j as a vector field on U_j . We will not impose any curvature conditions (in particular, we permit the vector fields Y_j to be constant), but we will impose the “transversality” (or “spanning”) condition

$$\det(Y_1(x^{(1)}), \dots, Y_d(x^{(d)})) \geq \nu \tag{1}$$

for all $x^{(1)} \in U_1, \dots, x^{(d)} \in U_d$, along with the smoothness condition

$$\|\Sigma_j\|_{C^2(U_j)} \leq A \quad \text{for all } 1 \leq j \leq d. \tag{2}$$

Remark 1.2. If U_j is sufficiently small, then $\mathcal{E}_j g_j = \widehat{G_j d\sigma_j}$, where $G_j: \Sigma_j(U_j) \rightarrow \mathbf{C}$ is the “normalised lift” of g_j , given by $G_j(\Sigma_j(x)) = |Y_j(x)|^{-1} g_j(x)$, and $d\sigma_j$ is the induced Lebesgue measure on $\Sigma_j(U_j)$.

By testing on the standard examples that generate the original linear restriction conjecture (characteristic functions of small balls in \mathbf{R}^{d-1} ; see [26]), we are led to the following conjecture.⁽¹⁾

CONJECTURE 1.3. (Multilinear restriction) *Suppose that the conditions (1) and (2) hold, and let $q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$. Then, there exists a constant C , depending only on A, ν, d , and U_1, \dots, U_d , for which*

$$\left\| \prod_{j=1}^d \mathcal{E}_j g_j \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \prod_{j=1}^d \|g_j\|_{L^{p'}(U_j)} \tag{3}$$

for all $g_1 \in L^{p'}(U_1), \dots, g_d \in L^{p'}(U_d)$.

Remark 1.4. Using a partition of unity and an appropriate affine transformation, we may assume that $\nu \sim 1$ and that for each $1 \leq j \leq d$, $\Sigma_j(U_j)$ is contained in a sufficiently small neighbourhood of the j th standard basis vector $e_j \in \mathbf{R}^d$.

Remark 1.5. By multilinear interpolation (see for example [6]) and Hölder’s inequality, Conjecture 1.3 may be reduced to the endpoint case $p=2$ and $q=2d/(d-1)$; i.e. the L^2 estimate

$$\left\| \prod_{j=1}^d \mathcal{E}_j g_j \right\|_{L^{2/(d-1)}(\mathbf{R}^d)} \leq C \prod_{j=1}^d \|g_j\|_{L^2(U_j)}. \tag{4}$$

We emphasise that at this d -linear level, the optimal estimate is on L^2 , rather than $L^{2d/(d-1)}$. It should also be pointed out that the conjectured range of exponents p and q is independent of any additional curvature assumptions that one might make on the submanifolds parametrised by the Σ_j ’s. This is very much in contrast with similar claims at lower levels of multilinearity. It is instructive to observe that if the mappings Σ_j are *linear*, then, by an application of Plancherel’s theorem, the conjectured

⁽¹⁾ Strictly speaking, this is a “multilinear extension” or “multilinear adjoint restriction” conjecture rather than a multilinear restriction conjecture, but the use of the term “restriction” is well established in the literature.

inequality (4) (for an appropriate constant C) is equivalent to the classical Loomis–Whitney inequality [22]. This elementary inequality states that if $\pi_j: \mathbf{R}^d \rightarrow \mathbf{R}^{d-1}$ is given by $\pi_j(x) := (x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_d)$, then

$$\int_{\mathbf{R}^d} f_1(\pi_1(x)) \dots f_d(\pi_d(x)) \, dx \leq \|f_1\|_{d-1} \dots \|f_d\|_{d-1} \tag{5}$$

for all $f_j \in L^{d-1}(\mathbf{R}^{d-1})$. One may therefore view the multilinear restriction conjecture as a certain (rather oscillatory) generalisation of the Loomis–Whitney inequality. The nature of this generalisation is clarified in §2.

Remark 1.6. Conjecture 1.3 in two dimensions is elementary and classical, and is implicit in arguments of C. Fefferman and Sjölin. In three dimensions this (trilinear) problem was considered in [5] (and previously in [3]), where some partial results on the sharp line $p' = q(d-1)/d$ were obtained.

It is a well-known fact that the linear restriction conjecture implies the so-called (linear) *Kakeya conjecture*. This conjecture takes several forms. A particularly simple one is the assertion that any compact set in \mathbf{R}^n which contains a unit line segment in every direction must have full Hausdorff (and thus Minkowski) dimension. Here we shall consider a more quantitative version of the conjecture, which is stronger than the one just described. For $0 < \delta \ll 1$ we define a δ -tube to be any rectangular box T in \mathbf{R}^d with $d-1$ sides of length δ and one side of length 1; observe that such tubes have volume $|T| \sim \delta^{d-1}$. Let \mathbf{T} be an arbitrary collection of such δ -tubes whose orientations form a δ -separated set of points on \mathbf{S}^{d-1} . We use $\#\mathbf{T}$ to denote the cardinality of \mathbf{T} , and χ_T to denote the characteristic function of T (thus $\chi_T(x) = 1$ when $x \in T$ and $\chi_T(x) = 0$ otherwise).

CONJECTURE 1.7. (Linear Kakeya) *Let \mathbf{T} and δ be as above. For each $d/(d-1) < q \leq \infty$ there exists a constant C , independent of δ and the collection \mathbf{T} , such that*

$$\left\| \sum_{T \in \mathbf{T}} \chi_T \right\|_{L^q(\mathbf{R}^d)} \leq C \delta^{(d-1)/q} (\#\mathbf{T})^{1-1/q(d-1)}. \tag{6}$$

The proof that Conjecture 1.1 implies Conjecture 1.7 follows a standard Rademacher-function argument going back implicitly to [15] and [2]. The endpoint $q = d/(d-1)$ of (6) can be seen to be false (unless one places an additional logarithmic factor in δ on the right-hand side), either by considering a collection of tubes passing through the origin, or by Besicovitch set examples. See [35] for a detailed account of these facts, and discussion of some of the progress made on this conjecture.

By a straightforward adaptation⁽²⁾ of the techniques in the linear situation, the multilinear restriction conjecture can be seen to imply a corresponding multilinear Kakeya-

⁽²⁾ In the multilinear setting one considers wave-packet decompositions of $\mathcal{E}_j g_j$ for general $g_j \in L^2(U_j)$, and places random \pm signs on each summand.

type conjecture. Suppose $\mathbf{T}_1, \dots, \mathbf{T}_d$ are families of δ -tubes in \mathbf{R}^d . We now allow the tubes within a single family \mathbf{T}_j to be parallel. However, we assume that for each $1 \leq j \leq d$, the tubes in \mathbf{T}_j have long sides pointing in directions belonging to some sufficiently small *fixed* neighbourhood of the j th standard basis vector e_j in \mathbf{S}^{d-1} . It will be convenient to refer to such a family of tubes as being *transversal*. (The vectors e_1, \dots, e_d may be replaced by any fixed linearly independent set of vectors in \mathbf{R}^d here, as affine invariance considerations reveal.)

CONJECTURE 1.8. (Multilinear Kakeya) *Let $\mathbf{T}_1, \dots, \mathbf{T}_d$ and δ be as above. For each $d/(d-1) \leq q \leq \infty$ there exists a constant C , independent of δ and the families of tubes $\mathbf{T}_1, \dots, \mathbf{T}_d$, such that*

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \prod_{j=1}^d (\delta^{d/q} \#\mathbf{T}_j). \tag{7}$$

Remark 1.9. Since the case $q = \infty$ is trivially true, the above conjecture is equivalent, via Hölder’s inequality, to the endpoint case $q = d/(d-1)$. In contrast to the linear setting, there is no obvious counterexample prohibiting this claim holding at the endpoint $q = d/(d-1)$, and indeed in the $d = 2$ case it is easy to verify this endpoint estimate.

Remark 1.10. By contrast with similar statements at lower levels of multilinearity, each family \mathbf{T}_j is permitted to contain parallel tubes, and even arbitrary repetitions of tubes. By a limiting argument, we thus see that the conjectured inequality reduces to the superficially stronger inequality

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} * \mu_{T_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \prod_{j=1}^d \left(\delta^{d/q} \sum_{T_j \in \mathbf{T}_j} \|\mu_{T_j}\| \right) \tag{8}$$

for all finite measures μ_{T_j} ($T_j \in \mathbf{T}_j$, $1 \leq j \leq d$) on \mathbf{R}^d .

Remark 1.11. The decision to formulate Conjecture 1.8 in terms of $\delta \times \dots \times \delta \times 1$ tubes is largely for historical reasons. However, just by scaling, it is easily seen that (7) is equivalent to the inequality

$$\left\| \prod_{j=1}^d \left(\sum_{\tilde{T}_j \in \tilde{\mathbf{T}}_j} \chi_{\tilde{T}_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \prod_{j=1}^d (\#\tilde{\mathbf{T}}_j), \tag{9}$$

where the collections $\tilde{\mathbf{T}}_j$ consist of tubes of width 1 and arbitrary (possibly infinite) length. (Of course we continue to impose the appropriate transversality condition on the families $\tilde{\mathbf{T}}_1, \dots, \tilde{\mathbf{T}}_d$ here.)

Remark 1.12. As may be expected given Remark 1.5, the special case of the conjectured inequality (or rather the equivalent form (8) with $q=d/(d-1)$ and an appropriate constant C) where all of the tubes in each family \mathbf{T}_j are *parallel*, is easily seen to be equivalent to the Loomis–Whitney inequality. We may therefore also view the multilinear Kakeya conjecture as a generalisation of the Loomis–Whitney inequality. The geometric nature of this generalisation is of course much more transparent than that of Conjecture 1.3. In particular, one may find it enlightening to reformulate (8) (with $q=d/(d-1)$) as an l^1 vector-valued version of (5).

Remark 1.13. As mentioned earlier, the linear Kakeya conjecture implies something about the dimension of sets which contain a unit line segment in every direction. The multilinear Kakeya conjecture does not have a similarly simple geometric implication, however there is a connection in a similar spirit between this conjecture and the joints problem; see §7.

Remarkably, at this d -linear level it turns out that the restriction and Kakeya conjectures are essentially *equivalent*. This “equivalence”, which is the subject of §2, follows from multilinearising a well-known induction-on-scales argument of Bourgain [8] (see also [30] for this argument in the bilinear setting). Once we have this equivalence, we may of course focus our attention on Conjecture 1.8, the analysis of which is the main innovation of this paper. The general idea behind our approach to this conjecture is sufficiently simple to warrant discussion here in the introduction. First let us observe that if each $T_j \in \mathbf{T}_j$ is centred at the origin (for all $1 \leq j \leq d$), then the two sides of the conjectured inequality (7) are trivially comparable. This observation leads to the suggestion that such configurations of tubes might actually be *extremal* for the left-hand side of (7).

Question 1.14. Is it reasonable to expect a quantity such as

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)}$$

to be monotone increasing for $q \geq d/(d-1)$ as the constituent tubes “simultaneously slide” to the origin?

For both analytic and algebraic reasons, in pursuing this idea it seems natural to replace the rough characteristic functions of tubes by gaussians (of the form $e^{-\pi \langle A(x-v), x-v \rangle}$, for appropriate positive definite $d \times d$ matrices A and vectors $v \in \mathbf{R}^d$) adapted to them. As we shall see in §3 and §4, with this gaussian reformulation the answer to the above question is, to all intents and purposes, yes for $q > d/(d-1)$. In §3 we illustrate this by giving a new proof of the Loomis–Whitney inequality, which we then

are able to perturb in §4. As a corollary of our perturbed result in §4, we obtain the multilinear Kakeya conjecture up to the endpoint, and a “weak” form of the multilinear restriction conjecture.

More precisely, our main results are as follows.

THEOREM 1.15. (Near-optimal multilinear Kakeya) *If $d/(d-1) < q \leq \infty$, then there exists a constant C , independent of δ and the transversal families of tubes $\mathbf{T}_1, \dots, \mathbf{T}_d$, such that*

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \prod_{j=1}^d (\delta^{d/q} \#\mathbf{T}_j).$$

Furthermore, for each $\varepsilon > 0$ there is a similarly uniform constant C for which

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{L^{1/(d-1)}(B(0,1))} \leq C \delta^{-\varepsilon} \prod_{j=1}^d (\delta^{d-1} \#\mathbf{T}_j).$$

THEOREM 1.16. (Near-optimal multilinear restriction) *For each $\varepsilon > 0$, $q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$, there exists a constant C , depending only on $A, \nu, \varepsilon, d, p$ and q , for which*

$$\left\| \prod_{j=1}^d \mathcal{E}_j g_j \right\|_{L^{q/d}(B(0,R))} \leq CR^\varepsilon \prod_{j=1}^d \|g_j\|_{L^p(U_j)}$$

for all $g_j \in L^p(U_j)$, $1 \leq j \leq d$, and all $R \geq 1$.

In §2 we show that Theorem 1.16 follows from Theorem 1.15. In §4 we prove Theorem 1.15, and in §5 we discuss the applications of our techniques to lower orders of multilinearity, and to more general multilinear k -plane transforms. In §6 we derive the natural variable coefficient extensions of our results using further bootstrapping arguments closely related to those of Bourgain. Finally, in §7 we give an application of our results to a variant of the classical “joints” problem considered in [12], [25] and [16].

Remark 1.17. The monotonicity approach that we take here arose from an attempt to devise a continuous and more efficient version of an existing induction-on-scales argument⁽³⁾ introduced by Wolff (and independently by the third author). This inductive argument allows one to deduce linear (and multilinear) Kakeya estimates for families of δ -tubes from corresponding ones for families of $\sqrt{\delta}$ -tubes. However, unfortunately, there are inefficiencies present which prevent one from keeping the constants in the inequalities under control from one iteration to the next. Our desire to minimise these inefficiencies led to the introduction of the formulations in terms of gaussians adapted

⁽³⁾ This induction-on-scales argument is closely related to that of Bourgain, and plays an important role in our applications in §6.

to tubes (rather than rough characteristic functions of tubes). The suggestion that one might then proceed by an induction-on-scales argument, incurring constant factors of at most 1 at each scale, is then tantamount to a certain monotonicity property. We should emphasise that this reasoning served mainly as philosophical motivation, and that there are important differences between the arguments presented here and the aforementioned induction arguments. (Curiously however, one of the most natural seeming formulations of monotonicity fails at the endpoint $q=d/(d-1)$ when $d \geq 3$; see Proposition 4.6.) We also remark that closely related monotonicity arguments for gaussians are effective in analysing the Brascamp–Lieb inequalities [11], [4]; see Remark 3.4 below.

Remark 1.18. There is perhaps some hope that variants of the techniques that we introduce here may lead to progress on the original linear form of the Kakeya conjecture. For $d \geq 3$ it seems unlikely that our multilinear estimates (in their current forms) may simply be “reassembled” in order to achieve this. However, our multilinear results do suggest (in some nonrigorous sense) that if there were some counterexamples to either the linear restriction or Kakeya conjectures, then they would have to be somewhat “non-transverse” (or “plany”, in the terminology of [18]). Issues of this nature arise in our application to “joints” problems in §7.

Remark 1.19. The d -linear transversality condition (1) has also turned out to be decisive when estimating spherical averages of certain multilinear extension operators $(g_1, \dots, g_d) \mapsto \mathcal{E}_1 g_1 \dots \mathcal{E}_d g_d$ of the type considered here. See [1] for further details.

Notation. For non-negative quantities X and Y , we will use the statement $X \lesssim Y$ to denote the existence of a constant C for which $X \leq CY$. The dependence of this constant on various parameters will depend on the context, and will be clarified where appropriate.

Acknowledgement. We would like to thank Jim Wright for many helpful discussions on a variety of techniques touched on in this paper.

2. Multilinear Restriction \iff Multilinear Kakeya

It is convenient to introduce some notation. For $\alpha \geq 0$, $q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$, we use

$$\mathcal{R}^*(p \times \dots \times p \rightarrow q; \alpha)$$

to denote the estimate

$$\left\| \prod_{j=1}^d \mathcal{E}_j g_j \right\|_{L^{q/d}(B(0,R))} \leq CR^\alpha \prod_{j=1}^d \|g_j\|_{L^p(U_j)}$$

for some constant C , depending only on A, ν, α, d, p and q , for all $g_j \in L^p(U_j), 1 \leq j \leq d$, and all $R \geq 1$. Similarly, for $d/(d-1) \leq q \leq \infty$, we use

$$\mathcal{K}^*(1 \times \dots \times 1 \rightarrow q; \alpha)$$

to denote the estimate

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{q/d} \leq C \delta^{-\alpha} \prod_{j=1}^d (\delta^{d/q} \#\mathbf{T}_j) \tag{10}$$

for some constant C , depending only on α, d and q , for all transversal collections of families of δ -tubes in \mathbf{R}^d , and all $0 < \delta \leq 1$. We again note that (10) is equivalent, by standard density arguments (in suitable weak topologies), to the superficially stronger inequality

$$\left\| \prod_{j=1}^d \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} * \mu_{T_j} \right) \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \delta^{-\alpha} \prod_{j=1}^d \left(\delta^{d/q} \sum_{T_j \in \mathbf{T}_j} \|\mu_{T_j}\| \right) \tag{11}$$

for all finite measures $\mu_{T_j} (T_j \in \mathbf{T}_j, 1 \leq j \leq d)$ on \mathbf{R}^d .

With this notation, Theorem 1.15 is equivalent to the statements $\mathcal{K}^*(1 \times \dots \times 1 \rightarrow q; 0)$ for all $d/(d-1) < q \leq \infty$, and $\mathcal{K}^*(1 \times \dots \times 1 \rightarrow d/(d-1); \varepsilon)$ for all $\varepsilon > 0$; Theorem 1.16 is equivalent to $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow 2d/(d-1); \varepsilon)$ for all $\varepsilon > 0$.

As we have already discussed, a standard Rademacher-function argument allows one to deduce the multilinear Kakeya conjecture from the multilinear restriction conjecture. In the localised setting, this of course continues to be true; i.e. for any $\alpha \geq 0$,

$$\mathcal{R}^*\left(2 \times \dots \times 2 \rightarrow \frac{2d}{d-1}; \alpha\right) \implies \mathcal{K}^*\left(1 \times \dots \times 1 \rightarrow \frac{d}{d-1}; 2\alpha\right). \tag{12}$$

Multilinearising a well-known bootstrapping argument of Bourgain [8] (again see [30] for this argument in the bilinear setting) we shall obtain the following reverse mechanism.

PROPOSITION 2.1. *For all $\alpha, \varepsilon \geq 0$ and $2d/(d-1) \leq q \leq \infty$,*

$$\mathcal{R}^*(2 \times \dots \times 2 \rightarrow q; \alpha) \text{ and } \mathcal{K}^*(1 \times \dots \times 1 \rightarrow \frac{1}{2}q; \varepsilon) \implies \mathcal{R}^*(2 \times \dots \times 2 \rightarrow q; \frac{1}{2}\alpha + \frac{1}{4}\varepsilon).$$

Using elementary estimates, we may easily verify $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow 2d/(d-1); \alpha)$ for some large positive value of α . For example, noting that $|B(0, R)| = c_d R^d$ for some constant c_d , we have that

$$\left\| \prod_{j=1}^d \mathcal{E}_j g_j \right\|_{L^{2/(d-1)}(B(0, R))} \leq c_d R^{d(d-1)/2} \prod_{j=1}^d \|\mathcal{E}_j g_j\|_\infty \leq c_d R^{d(d-1)/2} \prod_{j=1}^d \|g_j\|_{L^1(U_j)},$$

which, by the Cauchy–Schwarz inequality, yields $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow 2d/(d-1); \frac{1}{2}d(d-1))$. In the presence of appropriately favourable Keakeya estimates, this value of α may then be reduced by a repeated application of the above proposition. In particular, Proposition 2.1, along with implication (12), easily allows one to deduce the equivalence

$$\mathcal{R}^*\left(2 \times \dots \times 2 \rightarrow \frac{2d}{d-1}; \varepsilon\right) \text{ for all } \varepsilon > 0 \iff \mathcal{K}^*\left(1 \times \dots \times 1 \rightarrow \frac{d}{d-1}; \varepsilon\right) \text{ for all } \varepsilon > 0.$$

Arguing in much the same way allows us to reduce the proof of Theorem 1.16 to Theorem 1.15, as claimed in the introduction.

Our proof of Proposition 2.1 is very similar to that of [30, Lemma 4.4], and on a technical level is slightly more straightforward. We begin by stating a lemma which, given Remark 1.2 and the control of $|Y_j|$ implicit in (1) and (2), is a standard manifestation of the uncertainty principle (see [13] for the origins of this, and [30, Proposition 4.3] for a proof in the bilinear case which immediately generalises to the multilinear case).

LEMMA 2.2. $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow q; \alpha)$ is true if and only if

$$\left\| \prod_{j=1}^d \widehat{f}_j \right\|_{L^{q/d}(B(0,R))} \leq CR^{\alpha-d/2} \prod_{j=1}^d \|f_j\|_2 \quad (13)$$

for all $R \geq 1$ and functions f_j supported on $A_j^R := \Sigma_j(U_j) + O(R^{-1})$, $1 \leq j \leq d$.

We now turn to the proof of the proposition, where the implicit constants in the \lesssim notation will depend on at most A, ν, d, p, α and ε .

Proof of Proposition 2.1. From the above lemma, it suffices to show that

$$\left\| \prod_{j=1}^d \widehat{f}_j \right\|_{L^{q/d}(B(0,R))} \lesssim R^{\alpha/2+\varepsilon/4-d/2} \prod_{j=1}^d \|f_j\|_{L^2(A_j^R)}$$

for all f_j supported in A_j^R , $1 \leq j \leq d$. To this end, we let ϕ be a real-valued bump function adapted to $B(0, C)$, such that its Fourier transform is non-negative on the unit ball. For each $R \geq 1$ and $x \in \mathbf{R}^d$ let $\phi_{R^{1/2}}^x(\xi) := e^{-ix \cdot \xi} R^{d/2} \phi(R^{1/2} \xi)$. Observe that $\phi_{R^{1/2}}^x$ is an L^1 -normalised modulated bump function adapted to $B(0, C/R^{1/2})$, whose Fourier transform is non-negative and bounded below on $B(x, R^{1/2})$, uniformly in x . From the hypothesis $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow q; \alpha)$ and Lemma 2.2 we have

$$\left\| \prod_{j=1}^d \widehat{\phi_{R^{1/2}}^x f_j} \right\|_{L^{q/d}(B(x, R^{1/2}))} \lesssim R^{\alpha/2-d/4} \prod_{j=1}^d \|f_j * \phi_{R^{1/2}}^x\|_{L^2(A_j^R)} \quad (14)$$

for all x . Averaging this over $x \in B(0, R)$, we obtain

$$\left\| \prod_{j=1}^d \widehat{f_j} \right\|_{L^{q/d}(B(0,R))} \lesssim R^{\alpha/2-d/4} \left(R^{-d/2} \int_{B(0,R)} \left(\prod_{j=1}^d \|f_j * \phi_{R^{1/2}}^x\|_{L^2(\mathbf{R}^d)}^2 \right)^{q/2d} dx \right)^{d/q}. \quad (15)$$

Now, for each $1 \leq j \leq d$, we cover A_j^R by a boundedly overlapping collection of discs $\{\varrho_j\}$ of diameter $R^{-1/2}$, and set $f_{j,\varrho_j} := \chi_{\varrho_j} f_j$. Since (for each j) the supports of the functions $f_{j,\varrho_j} * \phi_{R^{1/2}}^x$ have bounded overlap, it suffices to show that

$$\left\| \prod_{j=1}^d \widehat{f_j} \right\|_{L^{q/d}(B(0,R))} \lesssim R^{\alpha/2-d/4} \left(R^{-d/2} \int_{B(0,R)} \left(\prod_{j=1}^d \sum_{\varrho_j} \|f_{j,\varrho_j} * \phi_{R^{1/2}}^x\|_{L^2(\mathbf{R}^d)}^2 \right)^{q/2d} dx \right)^{d/q}. \quad (16)$$

The function $\widehat{\phi_{R^{1/2}}^x}$ is rapidly decreasing away from $B(x, R^{1/2})$, and so, by Plancherel's theorem, the left-hand side of (16) is bounded by

$$CR^{\alpha/2-d/4} \left(R^{-d/2} \int_{B(0,R)} \left(\prod_{j=1}^d \sum_{\varrho_j} \|\widehat{f_{j,\varrho_j}}\|_{L^2(B(x,R^{1/2}))}^2 \right)^{q/2d} dx \right)^{d/q},$$

since the portions of $\widehat{\phi_{R^{1/2}}^x}$ on translates of $B(x, R^{1/2})$ can be handled by translation symmetry. For each ϱ_j let ψ_{ϱ_j} be a Schwartz function which is comparable to 1 on ϱ_j , and whose Fourier transform satisfies

$$|\widehat{\psi_{\varrho_j}}(x+y)| \lesssim R^{-(d+1)/2} \chi_{\varrho_j^*}(x)$$

for all $x, y \in \mathbf{R}^d$ with $|y| \leq R^{1/2}$, where ϱ_j^* denotes an $O(R) \times O(R^{1/2}) \times \dots \times O(R^{1/2})$ -tube, centred at the origin, and with long side pointing in the direction normal to the disc ϱ_j ; the implicit constants here depending only on A, ν and d . We point out that this is where we use the full $C^2(U_j)$ control given by condition (2). If we define $\tilde{f}_{j,\varrho_j} := f_{j,\varrho_j} / \psi_{\varrho_j}$, then f_{j,ϱ_j} and \tilde{f}_{j,ϱ_j} are pointwise comparable, and furthermore, by Jensen's inequality,

$$|\widehat{f_{j,\varrho_j}}(x+y)|^2 = |\widehat{\tilde{f}_{j,\varrho_j} * \psi_{\varrho_j}}(x+y)|^2 \lesssim R^{-(d+1)/2} |\widehat{\tilde{f}_{j,\varrho_j}}|^2 * \chi_{\varrho_j^*}(x)$$

whenever $x \in \mathbf{R}^d$ and $|y| \leq R^{1/2}$. Integrating this in y , we conclude that

$$\|\widehat{f_{j,\varrho_j}}\|_{L^2(B(x,R^{1/2}))}^2 \lesssim R^{-1/2} |\widehat{\tilde{f}_{j,\varrho_j}}|^2 * \chi_{\varrho_j^*}(x),$$

and hence, by rescaling the hypothesis $\mathcal{K}(1 \times \dots \times 1 \rightarrow \frac{1}{2}q; \varepsilon)$ (in its equivalent form (11)), we obtain

$$\begin{aligned} \left\| \prod_{j=1}^d \widehat{f_j} \right\|_{L^{q/d}(B(0,R))} &\lesssim R^{\alpha/2-d/4} \left(R^{-d/2} \int_{B(0,R)} \left(\prod_{j=1}^d \sum_{\varrho_j} R^{-1/2} |\widehat{\tilde{f}_{j,\varrho_j}}|^2 * \chi_{\varrho_j^*}(x) \right)^{q/2d} dx \right)^{d/q} \\ &\lesssim R^{\alpha/2+\varepsilon/4-d/2} \prod_{j=1}^d \left(\sum_{\varrho_j} \|\tilde{f}_{j,\varrho_j}\|_{L^2(A_j^R)}^2 \right)^{1/2} \\ &\lesssim R^{\alpha/2+\varepsilon/4-d/2} \prod_{j=1}^d \|f_j\|_{L^2(A_j^R)}. \end{aligned}$$

In the last two lines, we have used Plancherel’s theorem, disjointness and the pointwise comparability of \tilde{f}_{j,ϱ_j} and f_{j,ϱ_j} . This completes the proof of Proposition 2.1. \square

3. The Loomis–Whitney case

As we have already discussed, the “Loomis–Whitney case” of Conjecture 1.8 corresponds to the situation where each family \mathbf{T}_j consists of translates of a fixed tube with direction e_j ; the transversality hypothesis allows us to assume (after a linear change of variables) that e_1, \dots, e_d is the standard orthonormal basis of \mathbf{R}^d . The gaussian reformulation of this inequality (or rather the equivalent inequality (9)) alluded to in the introduction is now

$$\left\| \prod_{j=1}^d \left(\sum_{v_j \in \mathbf{V}_j} e^{-\pi \langle A_j^0(\cdot - v_j), (\cdot - v_j) \rangle} \right) \right\|_{L^{1/(d-1)}(\mathbf{R}^d)} \leq C \prod_{j=1}^d \#\mathbf{V}_j, \tag{17}$$

where for each $1 \leq j \leq d$, A_j^0 is the orthogonal projection to the j th coordinate hyperplane $\{x \in \mathbf{R}^d : x_j = 0\}$, and \mathbf{V}_j is an arbitrary finite subset of \mathbf{R}^d . The matrix A_j^0 , which is just the diagonal matrix whose diagonal entries are all 1 except for the j th entry, which is zero, we refer to as the j th *Loomis–Whitney matrix*.

We will actually consider a rather more general n -linear setup, where there are n distinct matrices A_j , which are not necessarily commuting, and no relation between n and d is assumed. We adopt the notation that $A \leq_{\text{pd}} B$ if $B - A$ is positive semi-definite, and $A <_{\text{pd}} B$ if $B - A$ is positive definite. The observations that if $A \leq_{\text{pd}} B$, then $D^*AD \leq_{\text{pd}} D^*BD$ for any matrix D , and also if $A \geq_{\text{pd}} B >_{\text{pd}} 0$ then $B^{-1} \geq_{\text{pd}} A^{-1} >_{\text{pd}} 0$ (as can be seen by comparing the norms $\langle Ax, x \rangle^{1/2}$ and $\langle Bx, x \rangle^{1/2}$ on \mathbf{R}^d , and then using duality) will be useful at the end of the proof of the next proposition.

PROPOSITION 3.1. *Let $d, n \geq 1$ and A_1, \dots, A_n be positive semi-definite real symmetric $d \times d$ matrices. Let μ_1, \dots, μ_n be finite compactly supported positive Borel measures on \mathbf{R}^d . For $t \geq 0$, $x \in \mathbf{R}^d$ and $1 \leq j \leq n$, let $f_j(t, x)$ denote the non-negative quantity*

$$f_j(t, x) := \int_{\mathbf{R}^d} e^{-\pi \langle A_j(x - v_j t), (x - v_j t) \rangle} d\mu_j(v_j).$$

Then if $\bigcap_{j=1}^n \ker A_j = \{0\}$ ⁽⁴⁾ and $p = (p_1, \dots, p_n) \in (0, \infty)^n$ is such that

$$A_1, \dots, A_n \leq_{\text{pd}} p_1 A_1 + \dots + p_n A_n, \tag{18}$$

⁽⁴⁾ Note that this is equivalent to the nonsingularity of $p_1 A_1 + \dots + p_n A_n$. See [4] for further analysis of the condition (18).

the quantity

$$Q_p(t) := \int_{\mathbf{R}^d} \prod_{j=1}^n f_j(t, x)^{p_j} dx$$

is nonincreasing in time.

COROLLARY 3.2. Under the conditions of Proposition 3.1,

$$\int_{\mathbf{R}^d} \prod_{j=1}^n f_j(1, x)^{p_j} dx \leq \int_{\mathbf{R}^d} \prod_{j=1}^n f_j(0, x)^{p_j} dx = \left(\frac{1}{\det A_*} \right)^{1/2} \prod_{j=1}^n \|\mu_j\|^{p_j},$$

where $A_* := p_1 A_1 + \dots + p_n A_n$.

Remark 3.3. The Loomis–Whitney case (17) (with $C=1$) now follows from Corollary 3.2, on setting $n=d$, $A_j=A_j^0$, $p_j=1/(d-1)$ for all j , and the measures μ_j to be arbitrary sums of Dirac masses.

Remark 3.4. Variants of Proposition 3.1 and Corollary 3.2, by the current authors and M. Christ [4], have recently lead to new proofs of the fundamental theorem of Lieb [21] concerning the exhaustion by gaussians of the Brascamp–Lieb inequalities (of which the Loomis–Whitney inequality is an important special case). Although our proof of Proposition 3.1 is rather less direct than the one given in [4] (which is closely related to the heat-flow approach of [11]), it does seem to lend itself much better to the perturbed situation, as we will discover in the next section.

Remark 3.5. In the statement of Proposition 3.1, we interpret t as the “time” variable so that the v_j ’s are the velocities with which the gaussians slide to the origin.

Our proof of Proposition 3.1 is rather unusual. We begin by considering the integer exponent case when $p=(p_1, \dots, p_n) \in \mathbf{N}^n$. By multiplying the p_j -th powers in the expression for $Q_p(t)$, and using Fubini’s theorem, we may obtain an explicit formula for the time derivative $Q'_p(t)$. With some careful algebraic and combinatorial manipulation, we are then able to rewrite this expression in a way that makes sense for $p \notin \mathbf{N}^n$, and is manifestly nonpositive whenever $A_1, \dots, A_n \leq_{\text{pd}} p_1 A_1 + \dots + p_n A_n$. Finally, we appeal to an extrapolation lemma (see the appendix) to conclude that the formula must in fact also hold for $p \in (0, \infty)^n$. It may also be interesting to consider this approach in the light of the notion of a fractional cartesian product of a set (see [7]).

Proof of Proposition 3.1. We begin by considering the case when $p \in \mathbf{N}^n$. Then, the quantity $Q_p(t)$ defined in this proposition can be expanded as

$$\int_{\mathbf{R}^d} \int_{(\mathbf{R}^d)^{p_1}} \dots \int_{(\mathbf{R}^d)^{p_n}} e^{-\pi \sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_j(x-v_{j,k}t), (x-v_{j,k}t) \rangle} \prod_{j=1}^n \prod_{k=1}^{p_j} d\mu_j(v_{j,k}) dx.$$

On completing the square, we find that

$$\sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_j(x - v_{j,k}t), (x - v_{j,k}t) \rangle = \langle A_* (x - \bar{v}t), (x - \bar{v}t) \rangle + \delta t^2,$$

where $A_* := \sum_{j=1}^n p_j A_j$ is a positive definite matrix, $\bar{v} := A_*^{-1} \sum_{j=1}^n A_j \sum_{k=1}^{p_j} v_{j,k}$ is the weighted average velocity, and δ is the weighted variance of the velocity:

$$\delta := \sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_j v_{j,k}, v_{j,k} \rangle - \langle A_* \bar{v}, \bar{v} \rangle. \quad (19)$$

Using translation invariance in x , we thus have

$$Q'_p(t) = -2\pi t \int_{\mathbf{R}^d} \int_{(\mathbf{R}^d)^{p_1}} \dots \int_{(\mathbf{R}^d)^{p_n}} \delta \prod_{j=1}^n \prod_{k=1}^{p_j} e^{-\pi \langle A_j(x - v_{j,k}t), (x - v_{j,k}t) \rangle} d\mu_j(v_{j,k}) dx.$$

If for each j we let \mathbf{v}_j be v_j regarded as a random variable associated to the probability measure

$$\frac{e^{-\pi \langle A_j(x - v_j t), (x - v_j t) \rangle} d\mu_j(v_j)}{f_j(t, x)},$$

and let $\mathbf{v}_{j,1}, \dots, \mathbf{v}_{j,p_j}$ be p_j independent samples of these random variables (with the $\mathbf{v}_{j,k}$'s being independent in both j and k), then we can write the equation above as

$$Q'_p(t) = -2\pi t \int_{\mathbf{R}^d} \mathbf{E}(\delta) \prod_{j=1}^n f_j(t, x)^{p_j} dx,$$

where δ is now considered a function of the $\mathbf{v}_{j,k}$'s, and $\mathbf{E}(\cdot)$ denotes probabilistic expectation. By linearity of expectation, we have

$$\mathbf{E}(\delta) = \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{E}(\langle A_j \mathbf{v}_{j,k}, \mathbf{v}_{j,k} \rangle) - \mathbf{E}(\langle A_* \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle) \quad (20)$$

(where of course $\bar{\mathbf{v}}$ is \bar{v} regarded as a random variable). By symmetry, the first term on the right-hand side is $\sum_{j=1}^n p_j \mathbf{E}(\langle A_j \mathbf{v}_j, \mathbf{v}_j \rangle)$. As for the second term, by definition of $\bar{\mathbf{v}}$, we have

$$\begin{aligned} \mathbf{E}(\langle A_* \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle) &= \mathbf{E} \left(\left\langle A_*^{-1} \sum_{j=1}^n \sum_{k=1}^{p_j} A_j \mathbf{v}_{j,k}, \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} A_{j'} \mathbf{v}_{j',k'} \right\rangle \right) \\ &= \sum_{j=1}^n \sum_{j'=1}^n \sum_{k=1}^{p_j} \sum_{k'=1}^{p_{j'}} \mathbf{E}(\langle A_*^{-1} A_j \mathbf{v}_{j,k}, A_{j'} \mathbf{v}_{j',k'} \rangle). \end{aligned}$$

When $(j, k) \neq (j', k')$, we can factorise the expectation using independence and symmetry to obtain

$$\begin{aligned} \mathbf{E}(\langle A_* \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle) &= \sum_{j=1}^n p_j \mathbf{E}(\langle A_*^{-1} A_j \mathbf{v}_j, A_j \mathbf{v}_j \rangle) - \sum_{j=1}^n p_j \langle A_*^{-1} A_j \mathbf{E}(\mathbf{v}_j), A_j \mathbf{E}(\mathbf{v}_j) \rangle \\ &\quad + \sum_{1 \leq j, j' \leq n} p_j p_{j'} \langle A_*^{-1} A_j \mathbf{E}(\mathbf{v}_j), A_{j'} \mathbf{E}(\mathbf{v}_{j'}) \rangle. \end{aligned}$$

Combining these observations, we obtain

$$Q'_p(t) = -2\pi t \int_{\mathbf{R}^d} G(p, t, x) \prod_{j=1}^n f_j(t, x)^{p_j} dx, \tag{21}$$

where G is the function

$$\begin{aligned} G(p, t, x) &:= \sum_{j=1}^n p_j (\mathbf{E}(\langle (A_j - A_j A_*^{-1} A_j) \mathbf{v}_j, \mathbf{v}_j \rangle) - \langle (A_j - A_j A_*^{-1} A_j) \mathbf{E}(\mathbf{v}_j), \mathbf{E}(\mathbf{v}_j) \rangle) \\ &\quad + \sum_{j=1}^n p_j \langle A_j \mathbf{E}(\mathbf{v}_j), \mathbf{E}(\mathbf{v}_j) \rangle - \sum_{1 \leq j, j' \leq n} p_j p_{j'} \langle A_*^{-1} A_j \mathbf{E}(\mathbf{v}_j), A_{j'} \mathbf{E}(\mathbf{v}_{j'}) \rangle \\ &= \sum_{j=1}^n p_j \mathbf{E}(\langle (A_j - A_j A_*^{-1} A_j)(\mathbf{v}_j - \mathbf{E}(\mathbf{v}_j)), (\mathbf{v}_j - \mathbf{E}(\mathbf{v}_j)) \rangle) \\ &\quad + \sum_{j=1}^n p_j \langle A_j (\mathbf{E}(\mathbf{v}_j) - \mathbf{E}(\bar{\mathbf{v}})), (\mathbf{E}(\mathbf{v}_j) - \mathbf{E}(\bar{\mathbf{v}})) \rangle. \end{aligned}$$

Note that G makes sense now not just for $p \in \mathbf{N}^n$, but for all $p \in (0, \infty)^n$. On the other hand, by the chain rule, we have

$$Q'_p(t) = 2\pi \int_{\mathbf{R}^d} \left(\sum_{k=1}^n p_k \mathbf{E}(\langle A_k \mathbf{v}_k, x - t \mathbf{v}_k \rangle) \right) \prod_{j=1}^n f_j(t, x)^{p_j} dx. \tag{22}$$

Now we observe that, since the adjugate (or cofactor) matrix $\text{adj}(A_*) := \det(A_*) A_*^{-1}$ is polynomial in p , $\det(A_*) G(p, t, x)$ is also polynomial in p . Hence, multiplying both (21) and (22) by $\det(A_*)$, and using Lemma 8.2 of the appendix (along with the hypothesis that A_* is nonsingular), we may deduce that (21) in fact holds for all $p \in (0, \infty)^n$. Now, if we choose p so that $A_* \succcurlyeq_{\text{pd}} A_j$ holds for all $1 \leq j \leq n$, then the inner product

$$\langle (A_j - A_j A_*^{-1} A_j) \cdot, \cdot \rangle$$

is positive semi-definite, and hence $G(p, t, x)$ is manifestly non-negative for all t and x . This proves Proposition 3.1. \square

Remark 3.6. The second formula for $G(p, t, x)$ above may be re-expressed as

$$G(p, t, x) = \sum_{j=1}^n p_j \mathbf{E}(\langle (A_j - A_j A_*^{-1} A_j)(\mathbf{v}_j - \mathbf{E}(\mathbf{v}_j)), (\mathbf{v}_j - \mathbf{E}(\mathbf{v}_j)) \rangle) + \sum_{j=1}^n p_j \langle A_j \mathbf{E}(\mathbf{v}_j), \mathbf{E}(\mathbf{v}_j) \rangle - \langle A_* \mathbf{E}(\bar{\mathbf{v}}), \mathbf{E}(\bar{\mathbf{v}}) \rangle.$$

While in this formulation it is not obvious at a glance that $G(p, t, x) \geq 0$, we can make a “centre of mass” change in the preceding argument to deduce this. Indeed, if we subtract $v_0 = v_0(t, x)$ from each $v_{j,k}$ in the definition (19) of δ , then the value of δ remains unchanged. If we now choose $v_0 = \mathbf{E}(\bar{\mathbf{v}})$, the last term in our expression for $G(p, t, x)$ vanishes, whence G is non-negative as before. This type of “Galilean invariance” will also be used crucially in the next section to obtain a similar positivity.

While this proof of Proposition 3.1 is a little more laboured than the one that we have presented in [4] (see also [11]), it nevertheless paves the way for the arguments of the next section where the (constant) matrices A_j are replaced by *random* matrices \mathbf{A}_j and are thus subject to the expectation operator \mathbf{E} . In that context, we are able to prove a suitable variant of the formula for G presented in this remark.

Remark 3.7. If $\sum_{j=1}^n p_j A_j > A_l$ for all l , then we can expect there to be room for a stronger estimate to hold. This will also play an important role in the next section, where we will use it to handle error terms arising in our analysis.

4. The perturbed Loomis–Whitney case

In this section, we prove a perturbed version of Proposition 3.1 of the previous section. Theorem 1.15 will then follow as a special case. Although Theorem 1.15 is our main goal, working at this increased level of generality has the advantage of providing more general (multilinear) k -plane transform estimates at all levels of multilinearity. We shall discuss these further applications briefly in the next section.

If A is a real symmetric $d \times d$ matrix, we use $\|A\|$ to denote the operator norm of A (one could also use other norms here, such as the Hilbert–Schmidt norm, as we are not tracking the dependence of constants on d).

PROPOSITION 4.1. *Let $d, n \geq 1$, $\varepsilon > 0$ and M_1, \dots, M_n be positive semi-definite real symmetric $d \times d$ matrices. Also, suppose that $\bigcap_{j=1}^n \ker M_j = \{0\}$ and $p = (p_1, \dots, p_n) \in (0, \infty)^n$ is such that*

$$p_1 M_1 + \dots + p_n M_n >_{\text{pd}} M_j \tag{23}$$

for all $1 \leq j \leq n$. For each $1 \leq j \leq n$, let Ω_j be a collection of pairs (A_j, v_j) , where $v_j \in \mathbf{R}^d$, $A_j \succ_{\text{pd}} 0$ and $\|A_j^{1/2} - M_j^{1/2}\| \leq \varepsilon$, and let μ_j be a finite compactly supported positive Borel measure on Ω_j . For $t \geq 0$, $x \in \mathbf{R}^d$, and $1 \leq j \leq n$, let $f_j(t, x)$ denote the non-negative quantity

$$f_j(t, x) := \int_{\Omega_j} e^{-\pi \langle A_j(x-v_j t), (x-v_j t) \rangle} d\mu_j(A_j, v_j).$$

Then if ε is sufficiently small depending on p and the M_j 's, we have the approximate monotonicity formula

$$\int_{\mathbf{R}^d} \prod_{j=1}^n f_j(t, x)^{p_j} dx \leq (1 + O(\varepsilon)) \int_{\mathbf{R}^d} \prod_{j=1}^n f_j(0, x)^{p_j} dx$$

for all $t \geq 0$, where we use $O(X)$ to denote a quantity bounded by CX for some constant $C > 0$ depending only on p and the M_j 's.

COROLLARY 4.2. *If p is such that (23) holds, then if ε is small enough depending on p and the M_j 's, we have the inequality*

$$\int_{\mathbf{R}^d} \prod_{j=1}^n f_j(1, x)^{p_j} dx \leq (1 + O(\varepsilon)) \left(\frac{1}{\det(M_*)} \right)^{1/2} \prod_{j=1}^n \|\mu_j\|^{p_j},$$

where $M_* = p_1 M_1 + \dots + p_n M_n$.

To prove the corollary, it is enough to show that there is a $c > 0$ (depending only on p and the M_j 's) such that

$$\prod_{j=1}^n f_j(0, x)^{p_j} \leq e^{-\pi \langle (1-c\varepsilon)M_*, x, x \rangle} \prod_{j=1}^n \|\mu_j\|^{p_j} \tag{24}$$

for all $x \in \mathbf{R}^d$. To this end, we first observe that for any $\lambda > 0$ we have the identity

$$\begin{aligned} \prod_{j=1}^n f_j(0, x)^{p_j} &= \prod_{j=1}^n \left(\int_{\Omega_j} e^{-\pi \langle A_j x, x \rangle} d\mu_j(A_j, v_j) \right)^{p_j} \\ &= e^{-\pi \langle (1-\lambda\varepsilon \sum_{k=1}^n p_k)M_*, x, x \rangle} \prod_{j=1}^n \left(\int_{\Omega_j} e^{-\pi \langle (A_j - M_j + \lambda\varepsilon M_*)x, x \rangle} d\mu_j(A_j, v_j) \right)^{p_j}. \end{aligned}$$

Now, since $M_* \succ_{\text{pd}} 0$, there exists a constant $c' > 0$, such that $M_* \succ_{\text{pd}} c'I$, where I denotes the identity matrix. Using this and (26) below, we may choose $\lambda > 0$ (depending only on p and the M_j 's) such that $A_j - M_j + \lambda\varepsilon M_*$ is positive definite, and hence $e^{-\pi \langle (A_j - M_j + \lambda\varepsilon M_*)x, x \rangle} \leq 1$ for all x . Setting $c = \lambda \sum_{k=1}^n p_k$ completes the proof of (24), and hence the corollary.

Remark 4.3. When the directions of the tubes in \mathbf{T}_j are sufficiently close to e_j , Theorem 1.15 follows from Corollary 4.2, on setting $n=d$, $M_j=A_j^0$, $p_j=p>1/(d-1)$ and the measures μ_j to be appropriate sums of Dirac masses. By affine invariance and the triangle inequality, one can handle any linearly independent direction sets, although of course the constants may now get significantly larger than 1.

Proof of Proposition 4.1. From the estimate $\|A_j^{1/2} - M_j^{1/2}\| \leq \varepsilon$, we have

$$A_j^{1/2} = M_j^{1/2} + O(\varepsilon) \tag{25}$$

and thus

$$A_j = M_j + O(\varepsilon). \tag{26}$$

Again, we begin by considering the case when p is an integer. Let Q_p denote the quantity

$$Q_p(t) := \int_{\mathbf{R}^d} \prod_{j=1}^n f_j(t, x)^{p_j} dx;$$

we can expand this as

$$Q_p(t) = \int_{\mathbf{R}^d} \int_{\Omega_1^{p_1}} \dots \int_{\Omega_n^{p_n}} e^{-\pi \sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_{j,k}(x-v_{j,k}t), (x-v_{j,k}t) \rangle} \prod_{j=1}^n \prod_{k=1}^{p_j} d\mu_j(A_{j,k}, v_{j,k}) dx.$$

It turns out that this quantity will not be easy for us to study directly (mainly because of the quantity A_*^{-1} which will appear in the derivative of Q_p). Instead, we shall consider the modified quantity $\tilde{Q}_p(t)$ defined by

$$\begin{aligned} \tilde{Q}_p(t) := & \int_{\mathbf{R}^d} \int_{\Omega_1^{p_1}} \dots \int_{\Omega_n^{p_n}} \det(A_*) e^{-\pi \sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_{j,k}(x-v_{j,k}t), (x-v_{j,k}t) \rangle} \\ & \times \prod_{j=1}^n \prod_{k=1}^{p_j} d\mu_j(A_{j,k}, v_{j,k}) dx, \end{aligned} \tag{27}$$

where $A_* := \sum_{j=1}^n \sum_{k=1}^{p_j} A_{j,k}$ is a positive definite matrix: the point is that the determinant $\det(A_*)$ will eventually be used to convert A_*^{-1} into a quantity which is a *polynomial* in the $A_{j,k}$'s; i.e. $\det(A_*)A_*^{-1} = \text{adj}(A_*)$, the adjugate matrix of A_* .

Let us now see why the weight $\det(A_*)$ is mostly harmless. From (26), we have $A_* = M_* + O(\varepsilon)$, where $M_* = p_1 M_1 + \dots + p_n M_n$, and so in particular we have

$$\det(A_*) = \det(M_*) + O(\varepsilon).$$

Thus, for ε small enough, $\det(A_*)$ is comparable to the positive constant $\det(M_*)$. Since all the terms in $Q_p(t)$ and $\tilde{Q}_p(t)$ are non-negative, we have thus established the relation

$$Q_p(t) = (1 + O(\varepsilon)) \det(M_*)^{-1} \tilde{Q}_p(t) \tag{28}$$

for integer p_1, \dots, p_n . However, for applications we need this relation for noninteger p_1, \dots, p_n . There is an obvious difficulty in doing so, namely that $\tilde{Q}_p(t)$ is not even defined for $p \notin \mathbf{N}^n$. However, this can be fixed by performing some manipulations (similar to ones considered previously) to rewrite $\tilde{Q}_p(t)$ as an expression which makes sense for arbitrary $p \in (0, \infty)^n$.

To this end, for each j we let $(\mathbf{A}_j, \mathbf{v}_j)$ be random variables (as before) associated to the probability measure

$$\frac{e^{-\pi \langle A_j(x-v_j t), (x-v_j t) \rangle} d\mu_j(A_j, v_j)}{f_j(t, x)}$$

and let $(\mathbf{A}_{j,1}, \mathbf{v}_{j,1}), \dots, (\mathbf{A}_{j,p_j}, \mathbf{v}_{j,p_j})$ be p_j independent samples of these random variables (with $(\mathbf{A}_{j,k}, \mathbf{v}_{j,k})$ being independent of $(\mathbf{A}_{j',k'}, \mathbf{v}_{j',k'})$ when $(j, k) \neq (j', k')$). Then, we can rewrite (27) as

$$\tilde{Q}_p(t) = \int_{\mathbf{R}^d} \mathbf{E}(\det(\mathbf{A}_*)) \prod_{j=1}^n f_j(t, x)^{p_j} dx,$$

where $\mathbf{A}_* := \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{A}_{j,k}$. If we write $\mathbf{R}_{j,k} := M_j - \mathbf{A}_{j,k}$, then the random variables $\mathbf{R}_{j,k}$ are $O(\varepsilon)$ (by (26)). Observe that for fixed j , all the $\mathbf{R}_{j,k}$'s have the same distribution as some fixed random variable \mathbf{R}_j , and we can write

$$\det(\mathbf{A}_*) = \det(M_*) + P((\mathbf{R}_{j,k})_{1 \leq j \leq n; 1 \leq k \leq p_j}),$$

where P is a polynomial in the coefficients of the $\mathbf{R}_{j,k}$'s which has no constant term (i.e. $P(0)=0$). Thus, we have

$$\tilde{Q}_p(t) = \det(M_*)Q_p(t) + \int_{\mathbf{R}^d} \mathbf{E}(P(\mathbf{R}_{j,k})_{1 \leq j \leq n; 1 \leq k \leq p_j}) \prod_{j=1}^n f_j(t, x)^{p_j} dx.$$

Now, by taking advantage of independence and symmetry of the random variables $\mathbf{R}_{j,k}$, we can write $\mathbf{E}(P(\mathbf{R}_{j,k}))$ as a polynomial combination of p and of the (tensor-valued) moments $\mathbf{E}(\mathbf{R}_j^{\otimes m})$, for some finite number of values of m ($m=1, \dots, M$, with M depending only on n, d and of course the M_j 's); thus, we have

$$\tilde{Q}_p(t) = \det(M_*)Q_p(t) + \int_{\mathbf{R}^d} \tilde{P}(p, \mathbf{E}(\mathbf{R}_j^{\otimes m})_{1 \leq j \leq n; 1 \leq m \leq M}) \prod_{j=1}^n f_j(t, x)^{p_j} dx \tag{29}$$

for some polynomial \tilde{P} depending only on n, d and the M_j 's. Since P has no constant term, it is easy to see that $\tilde{P}(p, \cdot)$ also has no constant term, i.e. $\tilde{P}(p, 0)=0$ (this can also be seen by considering the $\varepsilon=0$ case). We remark that this polynomial can be

computed explicitly using the Lagrange interpolation formula, although we make no use of this here. The right-hand side of (29) makes sense for any $p_1, \dots, p_n > 0$, not necessarily integers, and so we shall adopt it as our *definition* of \tilde{Q}_p in general. Since $\mathbf{R}_j = O(\varepsilon)$ and \tilde{P} has no constant term, we observe that $\tilde{P}(p, \mathbf{E}(\mathbf{R}_j^{\otimes m})_{1 \leq j \leq n; 1 \leq m \leq M}) = O(\varepsilon)$, whence we obtain (28) for all $p_1, \dots, p_n > 0$ (not just the integers), though of course the implicit constants in the O notation will certainly depend (polynomially) on p .

In order to prove the proposition, it thus suffices by (28) to show that the quantity $\tilde{Q}_p(t)$ is nondecreasing in time for sufficiently small ε . Again, we begin by working with $p \in \mathbf{N}^n$, so that we may use (27). Now, we differentiate $\tilde{Q}_p(t)$. As before, we can complete the square and write

$$\sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_{j,k}(x - v_{j,k}t), (x - v_{j,k}t) \rangle = \langle A_*(x - \bar{v}t), (x - \bar{v}t) \rangle + \delta t^2,$$

where $\bar{v} := A_*^{-1} \sum_{j=1}^n \sum_{k=1}^{p_j} A_{j,k} v_{j,k}$ is the weighted average velocity, and δ is the weighted variance

$$\delta := \sum_{j=1}^n \sum_{k=1}^{p_j} \langle A_{j,k} v_{j,k}, v_{j,k} \rangle - \langle A_* \bar{v}, \bar{v} \rangle. \quad (30)$$

Arguing as in the previous section, we thus have

$$\tilde{Q}'_p(t) = -2\pi t \int_{\mathbf{R}^d} \int_{\Omega_1^{p_1}} \dots \int_{\Omega_n^{p_n}} \det(A_*) \delta \prod_{j=1}^n \prod_{k=1}^{p_j} e^{-\pi \langle A_{j,k}(x - v_{j,k}t), (x - v_{j,k}t) \rangle} d\mu_j(A_{j,k}, v_{j,k}) dx.$$

Recalling the random variables $\mathbf{A}_{j,k}$ and $\mathbf{v}_{j,k}$, we thus have

$$\tilde{Q}'_p(t) = -2\pi t \int_{\mathbf{R}^d} \mathbf{E}(\det(\mathbf{A}_*) \delta) \prod_{j=1}^n f_j(t, x)^{p_j} dx, \quad (31)$$

where δ is now considered a function of the $\mathbf{A}_{j,k}$'s and $\mathbf{v}_{j,k}$'s.

Let us rewrite δ slightly by inserting the definition of $\bar{\mathbf{v}}$ and using the self-adjointness of A_*^{-1} , to obtain

$$\delta = \sum_{j=1}^n \sum_{k=1}^{p_j} \langle \mathbf{A}_{j,k} \mathbf{v}_{j,k}, \mathbf{v}_{j,k} \rangle - \left\langle \mathbf{A}_*^{-1} \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{A}_{j,k} \mathbf{v}_{j,k}, \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} \mathbf{A}_{j',k'} \mathbf{v}_{j',k'} \right\rangle.$$

We now take advantage of a certain ‘‘Galilean invariance’’ of the problem. We introduce an arbitrary (deterministic) vector field $v_0(t, x)$ which we are at liberty to select later, and observe that the above expression is unchanged if we replace each $\mathbf{v}_{j,k}$ by $\mathbf{v}_{j,k} - v_0$:

$$\begin{aligned} \delta &= \sum_{j=1}^n \sum_{k=1}^{p_j} \langle \mathbf{A}_{j,k} (\mathbf{v}_{j,k} - v_0), (\mathbf{v}_{j,k} - v_0) \rangle \\ &\quad - \left\langle \mathbf{A}_*^{-1} \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{A}_{j,k} (\mathbf{v}_{j,k} - v_0), \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} \mathbf{A}_{j',k'} (\mathbf{v}_{j',k'} - v_0) \right\rangle. \end{aligned}$$

If we multiply this quantity by $\det(\mathbf{A}_*)$, then we obtain a polynomial:

$$\det(\mathbf{A}_*)\delta = \sum_{j=1}^n \sum_{k=1}^{p_j} \det(\mathbf{A}_*) \langle \mathbf{A}_{j,k}(\mathbf{v}_{j,k} - v_0), (\mathbf{v}_{j,k} - v_0) \rangle - \left\langle \text{adj}(\mathbf{A}_*) \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{A}_{j,k}(\mathbf{v}_{j,k} - v_0), \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} \mathbf{A}_{j',k'}(\mathbf{v}_{j',k'} - v_0) \right\rangle, \tag{32}$$

where $\text{adj}(\mathbf{A}_*)$ is the adjugate matrix of \mathbf{A}_* , which is a polynomial in the coefficients of \mathbf{A}_* , and is thus a polynomial in the coefficients of the $\mathbf{A}_{j,k}$'s. To make this expression more elliptic, we introduce the matrices $\mathbf{B}_{j,k} := \mathbf{A}_{j,k}^{1/2}$ and the vectors

$$\mathbf{w}_{j,k} := \mathbf{B}_{j,k}(\mathbf{v}_{j,k} - v_0),$$

and write (32) as

$$\det(\mathbf{A}_*)\delta = \sum_{j=1}^n \sum_{k=1}^{p_j} \det(\mathbf{A}_*) \|\mathbf{w}_{j,k}\|^2 - \left\langle \text{adj}(\mathbf{A}_*) \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{B}_{j,k} \mathbf{w}_{j,k}, \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} \mathbf{B}_{j',k'} \mathbf{w}_{j',k'} \right\rangle. \tag{33}$$

Note that $\mathbf{A}_* = \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{B}_{j,k}^2$ is a polynomial in the $\mathbf{B}_{j,k}$'s, so the right-hand side of (33) is a polynomial in the $\mathbf{B}_{j,k}$'s and $\mathbf{w}_{j,k}$'s. We can then take expectations, taking advantage of the independence and symmetry of the random variables $(\mathbf{B}_{j,k}, \mathbf{w}_{j,k})$, and obtain a formula of the form

$$\mathbf{E}(\det(\mathbf{A}_*)\delta) = S(p, (\mathbf{E}((\mathbf{B}_j, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) \tag{34}$$

for some polynomial S , and where the power of moments M depends only on d (in fact, it is $2d+2$, which is the degree of the polynomial in (33)). Here of course $(\mathbf{B}_j, \mathbf{w}_j)$ represents any random variable with the same distribution as the $(\mathbf{B}_{j,k}, \mathbf{w}_{j,k})$'s. Also note that, as the right-hand side of (33) is purely quadratic in the $\mathbf{w}_{j,k}$'s, the expression (34) must be purely quadratic in the \mathbf{w}_j 's. Just as before, one may of course use the Lagrange interpolation formula here to write down an explicit expression for S .

Inserting this formula back into (31), we obtain

$$\tilde{Q}'_p(t) = -2\pi t \int_{\mathbf{R}^d} S(p, (\mathbf{E}((\mathbf{B}_j, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) \prod_{j=1}^n f_j(t, x)^{p_j} dx \tag{35}$$

when $p \in \mathbf{N}^n$. However, differentiating with respect to t through the integral in (29) leads to an alternative expression for $\tilde{Q}'_p(t)$ of the form

$$\tilde{Q}'_p(t) = 2\pi \int_{\mathbf{R}^d} K(p, t, x) \prod_{j=1}^n f_j(t, x)^{p_j} dx,$$

valid for all $p \in (0, \infty)^n$, where K is polynomial in p and of polynomial growth in x . Hence, by Lemma 8.2 of the appendix, identity (35) must also be true for arbitrary $p \in (0, \infty)^n$. Thus, to conclude the proof, it will suffice to show that

$$S(p, (\mathbf{E}((\mathbf{B}_j, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) \geq 0 \tag{36}$$

for any p satisfying the hypotheses of the theorem, if ε is sufficiently small depending on p and the M_j 's.

From (25), we know that $\mathbf{B}_j = M_j^{1/2} + O(\varepsilon)$, and in particular we have $\mathbf{B}_j = O(1)$. Since S is purely quadratic in the \mathbf{w}_j 's, we thus have

$$\begin{aligned} & S(p, (\mathbf{E}((\mathbf{B}_j, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) \\ &= S(p, (\mathbf{E}((M_j^{1/2}, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) - O\left(\varepsilon \sum_{j=1}^n \mathbf{E}(\|\mathbf{w}_j\|^2)\right), \end{aligned} \tag{37}$$

as we can use the Cauchy–Schwarz inequality to control any cross terms such as $\mathbf{E}(\mathbf{w}_j)^{\otimes 2}$ or $\mathbf{E}(\mathbf{w}_j) \otimes \mathbf{E}(\mathbf{w}_k)$. Before going ahead with the proof of Proposition 4.1, we need the following lemma.

LEMMA 4.4. *For arbitrary $p_1, \dots, p_n > 0$,*

$$\begin{aligned} S(p, (\mathbf{E}((M_j^{1/2}, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) &= \det M_* \left(\sum_{j=1}^n p_j \|\mathbf{E}(\mathbf{w}_j)\|^2 \right. \\ &+ \sum_{j=1}^n p_j \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2})(\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)), (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)) \rangle) \\ &\left. - \left\| M_*^{-1/2} \sum_{j=1}^n p_j M_j^{1/2} \mathbf{E}(\mathbf{w}_j) \right\|^2 \right). \end{aligned} \tag{38}$$

Proof. Since the desired identity is polynomial in p_1, \dots, p_n , it suffices to restrict our attention to $p_1, \dots, p_n \in \mathbf{N}$. From (34) and (33) (with \mathbf{B}_j replaced by $M_j^{1/2}$) we have

$$\begin{aligned} & S(p, (\mathbf{E}((M_j^{1/2}, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) \\ &= \mathbf{E} \left(\sum_{j=1}^n \sum_{k=1}^{p_j} \det(M_*) \|\mathbf{w}_{j,k}\|^2 - \left\langle \text{adj}(M_*) \sum_{j=1}^n \sum_{k=1}^{p_j} M_j^{1/2} \mathbf{w}_{j,k}, \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} M_{j'}^{1/2} \mathbf{w}_{j',k'} \right\rangle \right). \end{aligned}$$

Writing $\text{adj}(M_*) = \det(M_*) M_*^{-1}$, we thus have

$$\begin{aligned} & S(p, (\mathbf{E}((M_j^{1/2}, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) = \det(M_*) \sum_{j=1}^n \sum_{k=1}^{p_j} \mathbf{E}(\|\mathbf{w}_{j,k}\|^2) \\ & - \det(M_*) \mathbf{E} \left(\left\langle M_*^{-1} \sum_{j=1}^n \sum_{k=1}^{p_j} M_j^{1/2} \mathbf{w}_{j,k}, \sum_{j'=1}^n \sum_{k'=1}^{p_{j'}} M_{j'}^{1/2} \mathbf{w}_{j',k'} \right\rangle \right), \end{aligned}$$

which, by independence and symmetry, we can rewrite as

$$\begin{aligned} \det(M_*) \sum_{j=1}^n p_j \mathbf{E}(\|\mathbf{w}_j\|^2) &- \det(M_*) \sum_{j=1}^n p_j \mathbf{E}(\langle M_*^{-1} M_j^{1/2} \mathbf{w}_j, M_j^{1/2} \mathbf{w}_j \rangle) \\ &- \det(M_*) \sum_{j=1}^n p_j (p_j - 1) \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_j^{1/2} \mathbf{E}(\mathbf{w}_j) \rangle \\ &- 2 \det(M_*) \sum_{1 \leq j < j' \leq n} p_j p_{j'} \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_{j'}^{1/2} \mathbf{E}(\mathbf{w}_{j'}) \rangle. \end{aligned}$$

Now, we argue as in the previous section and write

$$\begin{aligned} \sum_{j=1}^n p_j \mathbf{E}(\|\mathbf{w}_j\|^2) &- \sum_{j=1}^n p_j \mathbf{E}(\langle M_*^{-1} M_j^{1/2} \mathbf{w}_j, M_j^{1/2} \mathbf{w}_j \rangle) \\ &= \sum_{j=1}^n p_j \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) \mathbf{w}_j, \mathbf{w}_j \rangle), \end{aligned}$$

and

$$\begin{aligned} \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) \mathbf{w}_j, \mathbf{w}_j \rangle) &= \langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) \mathbf{E}(\mathbf{w}_j), \mathbf{E}(\mathbf{w}_j) \rangle \\ &+ \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)), (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)) \rangle). \end{aligned}$$

Hence

$$\begin{aligned} S(p, (\mathbf{E}((M_j^{1/2}, \mathbf{w}_j)^{\otimes m}))_{1 \leq j \leq n; 1 \leq m \leq M}) &= \det(M_*) \left(\sum_{j=1}^n p_j \langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) \mathbf{E}(\mathbf{w}_j), \mathbf{E}(\mathbf{w}_j) \rangle \right. \\ &+ \sum_{j=1}^n p_j \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)), (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)) \rangle) \\ &- \sum_{j=1}^n p_j (p_j - 1) \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_j^{1/2} \mathbf{E}(\mathbf{w}_j) \rangle \\ &\left. - 2 \sum_{1 \leq j < j' \leq n} p_j p_{j'} \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_{j'}^{1/2} \mathbf{E}(\mathbf{w}_{j'}) \rangle \right). \end{aligned}$$

We can rearrange the right-hand side as

$$\begin{aligned} \det(M_*) \left(\sum_{j=1}^n p_j \|\mathbf{E}(\mathbf{w}_j)\|^2 + \sum_{j=1}^n p_j \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2}) (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)), (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)) \rangle) \right. \\ \left. - \sum_{j=1}^n p_j^2 \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_j^{1/2} \mathbf{E}(\mathbf{w}_j) \rangle \right. \\ \left. - 2 \sum_{1 \leq j < j' \leq n} p_j p_{j'} \langle M_*^{-1} M_j^{1/2} \mathbf{E}(\mathbf{w}_j), M_{j'}^{1/2} \mathbf{E}(\mathbf{w}_{j'}) \rangle \right) \end{aligned}$$

which can be rearranged further as

$$\det(M_*) \left(\sum_{j=1}^n p_j \|\mathbf{E}(\mathbf{w}_j)\|^2 + \sum_{j=1}^n p_j \mathbf{E}(\langle (I - M_j^{1/2} M_*^{-1} M_j^{1/2})(\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)), (\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)) \rangle) - \left\| M_*^{-1/2} \sum_{j=1}^n p_j M_j^{1/2} \mathbf{E}(\mathbf{w}_j) \right\|^2 \right).$$

This concludes the proof of the lemma. □

The last term in (38) has an unfavourable sign. However, we can at last select our vector field v_0 in order to remove this term. More precisely, we define v_0 by the formula

$$v_0 := \left(\mathbf{E} \left(\sum_{j=1}^n p_j M_j^{1/2} \mathbf{A}_j^{1/2} \right) \right)^{-1} \sum_{j=1}^n p_j \mathbf{E}(M_j^{1/2} \mathbf{A}_j^{1/2} \mathbf{v}_j),$$

which is well defined for ε sufficiently small since $\mathbf{A}_j = M_j + O(\varepsilon)$, and thus

$$\sum_{j=1}^n p_j M_j^{1/2} \mathbf{A}_j^{1/2}$$

is close to the invertible matrix M_* . With this choice of v_0 , we compute that

$$\sum_{j=1}^n p_j M_j^{1/2} \mathbf{E}(\mathbf{w}_j) = 0,$$

and so the last term in (38) vanishes.

Recalling (37) and (38), and using the fact that there exists a positive constant c , depending only on p and the M_j 's, such that $I - M_j^{1/2} M_*^{-1} M_j^{1/2} \geq_{\text{pd}} cI$ for all $1 \leq j \leq n$, we see that it suffices to show that

$$\sum_{j=1}^n p_j \|\mathbf{E}(\mathbf{w}_j)\|^2 + c \sum_{j=1}^n p_j \mathbf{E}(\|\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)\|^2) - O\left(\varepsilon \sum_{j=1}^n \mathbf{E}(\|\mathbf{w}_j\|^2)\right) \geq 0.$$

However, we have $\|\mathbf{w}_j\|^2 \leq 2\|\mathbf{w}_j - \mathbf{E}(\mathbf{w}_j)\|^2 + 2\|\mathbf{E}(\mathbf{w}_j)\|^2$, and so the claim now follows if ε is sufficiently small depending on p and the M_j 's. This completes the proof of Proposition 4.1. □

When each p_j is an integer, the quantity $Q_p(t)$ is monotone without further hypotheses on the matrices A_j , since δ defined in (30) is manifestly non-negative. This in particular includes the 2-dimensional bilinear Keakeya situation $n=d=2$ and $p_1=p_2=1$. (While we do not in this paper establish monotonicity of $Q_p(t)$ under the hypotheses of Proposition 4.1, neither do we rule it out.) The trilinear endpoint Keakeya situation in three dimensions corresponds to $n=d=3$, $p_1=p_2=p_3=\frac{1}{2}$ and each matrix A_j having rank 2. Interestingly, in this case $Q_p(t)$ is *not* monotone decreasing, as the following lemma will demonstrate.

LEMMA 4.5. *Let $d \geq 3$. For $j=1, \dots, d$, suppose A_j is a non-negative definite $d \times d$ matrix of rank $d-1$ such that $\ker A_1, \dots, \ker A_d$ span. Suppose furthermore that*

$$\frac{1}{d-1}(A_1 + \dots + A_d) \geq A_1, \dots, A_d.$$

Then, A_d is uniquely determined by A_1, \dots, A_{d-1} .

Proof. According to [4, Proposition 3.6], under the hypotheses of the lemma there exists an invertible $d \times d$ matrix D and rank $d-1$ projections P_j such that $A_j = D^* P_j D$ for all j and

$$\frac{1}{d-1}(P_1 + \dots + P_d) = I_d.$$

We claim that, for $j \neq k$, $\ker P_j$ and $\ker P_k$ are orthogonal. To see this, let x_j be a unit vector in $\ker P_j$ and note that

$$x_1 = \frac{1}{d-1}(P_1 + \dots + P_d)x_1 = \frac{1}{d-1}(P_2 + \dots + P_d)x_1.$$

So

$$\left\langle \frac{1}{d-1}(P_2 + \dots + P_d)x_1, x_1 \right\rangle = 1.$$

On the other hand, $\langle P_j x_1, x_1 \rangle \leq 1$ for all j , since P_j is a projection. Thus $\langle P_j x_1, x_1 \rangle = 1$, and $x_1 = P_j x_1$ for all $j \neq 1$. Similarly, $P_j x_i = x_i$ for all $i \neq j$. Hence, for $i \neq j$, $\langle x_i, x_j \rangle = \langle P_j x_i, x_j \rangle = \langle x_i, P_j x_j \rangle = 0$. Thus, the kernels of the P_j 's are mutually orthogonal. By choosing a suitable orthonormal basis (and possibly changing D), we may assume therefore that $P_j = A_j^0$, the j th Loomis–Whitney matrix.

Now, let us suppose that \tilde{D} is another $d \times d$ invertible matrix and that $\tilde{A}_j = \tilde{D}^* A_j^0 \tilde{D}$. Then, the statement $A_i = \tilde{A}_i$ is the same as $A_i^0 = R^* A_i^0 R$ with $R = \tilde{D} D^{-1}$; i.e. R leaves both $\ker A_i^0$ and its orthogonal complement $\text{im} A_i^0$ invariant, and acts as an isometry on the latter. If now $A_1 = \tilde{A}_1, \dots, A_{d-1} = \tilde{A}_{d-1}$, then this means that R acts as an isometry on every coordinate hyperplane except possibly $\{x : x_d = 0\}$, and the standard basis vectors e_1, \dots, e_{d-1} are eigenvectors of R . This forces the matrix of R with respect to the standard basis to be diagonal with entries ± 1 and thus $A_d = \tilde{A}_d$ too. \square

PROPOSITION 4.6. *For $j=1, \dots, d$, let W_j be a set of non-negative definite $d \times d$ matrices of rank $d-1$ such that, for all $A_j \in W_j$, $\ker A_1, \dots, \ker A_d$ span. Let μ_j be a positive Borel measure on $W_j \times \mathbf{R}^d$. Let $p = (1/(d-1), \dots, 1/(d-1))$. If $Q_p(1) \leq Q_p(0)$ for all such positive measures μ_j , then either $d=2$ or each W_j is a singleton, and*

$$\frac{1}{d-1}(A_1 + \dots + A_d) \geq A_1, \dots, A_d.$$

Remark 4.7. In the latter case of Proposition 4.6 ($d \geq 3$), the family $\{A_j\}_{1 \leq j \leq d}$ is an affine image of the Loomis–Whitney matrices $\{A_j^0\}_{1 \leq j \leq d}$, i.e. for some invertible D , $A_j = D^* A_j^0 D$, as the proof of Lemma 4.5 shows.

Proof of Proposition 4.6. Let $\mu_j^\#$ be arbitrary positive finite Borel measures on \mathbf{R}^d , and let $A_j \in W_j$. With

$$f_j(t, x) := \int_{\mathbf{R}^d} e^{-\pi \langle A_j(x - v_j t), (x - v_j t) \rangle} d\mu_j^\#(v_j)$$

and

$$Q_p^\#(t) := \int_{\mathbf{R}^d} f_1(t, x)^{1/(d-1)} \dots f_d(t, x)^{1/(d-1)} dx,$$

setting $\mu_j := \mu_j^\# \otimes \delta_0(A_j)$ we now have $Q_p^\#(1) \leq Q_p^\#(0)$. By [4, Proposition 3.6], this forces

$$\frac{1}{d-1} (A_1 + \dots + A_d) \geq A_1, \dots, A_d.$$

When $d \geq 3$, Lemma 4.5 shows that any $d-1$ of the A_j 's determine the remaining one. Thus, no W_j may contain more than one point. □

Remark 4.8. The same argument applies if instead of Q_p we consider \tilde{Q}_p . Consequently the crucial single-signedness of S in (36) fails in this setting at the endpoint $p = (1/(d-1), \dots, 1/(d-1))$.

5. Lower levels of multilinearity

As we remarked in the introduction, if $n < d$ then the conjectured exponents for n -linear restriction type problems depend on the curvature properties of the submanifolds in question; with flatter surfaces expected to enjoy fewer restriction estimates. Our methods here cannot easily take advantage of such curvature hypotheses, though, and we will instead address the issue of establishing n -linear restriction estimates in \mathbf{R}^d which assume only transversality properties rather than curvature properties. In such a case we can establish quite sharp estimates.

A similar situation occurs when considering n -linear Kakeya estimates in \mathbf{R}^d . The analogue of “curvature” would be some sort of direction separation condition (or perhaps a mixed Lebesgue norm condition) on the tubes in a given family. Here we will only consider estimates in which the tubes are counted by cardinality rather than in mixed norms, and transversality is assumed rather than direction separation.

We begin by discussing the n -linear Kakeya situation. Fix $3 \leq n \leq d$ (the case $n=1$ turns out to be void, and the $n=2$ case standard, see e.g. [30]). Suppose $\mathbf{T}_1, \dots, \mathbf{T}_n$ are

families of δ -tubes in \mathbf{R}^d . Suppose further that for each $1 \leq j \leq n$, the tubes in \mathbf{T}_j have long sides pointing in directions belonging to some sufficiently small *fixed* neighbourhood of the j th standard basis vector e_j in \mathbf{S}^{d-1} . (Again, the vectors e_1, \dots, e_n may be replaced by any fixed set of n linearly independent vectors in \mathbf{R}^d here, as affine invariance considerations reveal.)

THEOREM 5.1. *If $n/(n-1) < q \leq \infty$, then there exists a constant C , independent of δ and the families of tubes $\mathbf{T}_1, \dots, \mathbf{T}_n$, such that*

$$\left\| \prod_{j=1}^n \left(\sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right) \right\|_{L^{q/n}(\mathbf{R}^d)} \leq C \prod_{j=1}^n (\delta^{d/q} \#\mathbf{T}_j). \tag{39}$$

Remark 5.2. One can conjecture the same result to hold at the endpoint $q = n/(n-1)$; note for instance that the estimate is easily verifiable at this endpoint when $n=2$. For $q < n/(n-1)$, the estimate is false, as can be seen by choosing each family \mathbf{T}_j to be (essentially) a partition of a δ -neighbourhood of the unit cube in $\text{span}\{e_1, \dots, e_n\} \subset \mathbf{R}^d$ into parallel tubes oriented in the direction e_j . It is likely that the estimates can be improved if some direction separation condition is imposed on each of the \mathbf{T}_j 's, but we do not pursue this matter here.

In order to prove Theorem 5.1, we first apply a rescaling (in the spirit of (9)) to reduce inequality (39) to an equivalent statement for families of tubes of width 1, and arbitrary length. Then, we simply dominate the characteristic functions of these dilated tubes by appropriate gaussians, and appeal to Corollary 4.2, setting $p_j = p > 1/(n-1)$ and $M_j = A_j^0$, where A_j^0 is the j th Loomis–Whitney matrix in \mathbf{R}^d , and $1 \leq j \leq n$.

The corresponding n -linear restriction inequalities (with the familiar ε -loss in the localisation parameter R) may now be obtained by the bootstrapping argument from §2, thus obtaining an estimate of the form

$$\left\| \prod_{j=1}^n \mathcal{E}_j g_j \right\|_{L^{q/n}(B(0,R))} \leq CR^\varepsilon \prod_{j=1}^n \|g_j\|_{L^p(U_j)} \tag{40}$$

when $q \geq 2n/(n-1)$ and $p' \leq q(n-1)/n$. What is perhaps particularly curious is the fact that for $n < d$, the standard Rademacher-function argument *does not* allow the optimal n -linear Kakeya inequalities to be obtained from the corresponding optimal n -linear restriction inequalities. To avoid repetition of the arguments in §2, we omit the details.

Remark 5.3. The ε -loss in (40) should be removable (thus allowing R to be sent to infinity); certainly this is possible in the $n=2$ case, see for instance [29] for this standard and useful estimate. The conditions $q \geq 2n/(n-1)$ and $p' \leq q(n-1)/n$ can be verified to

be sharp (for example, by considering the Loomis–Whitney case when the maps Σ_j are linear), but can be improved when the Σ_j 's have additional curvature properties (again, see for instance [29] for a survey).

Remark 5.4. One may also obtain nontrivial multilinear estimates for k -plane transforms from Corollary 4.2, by choosing the matrices M_j to be appropriate projections onto $(d-k)$ -dimensional subspaces of \mathbf{R}^d . For example, if $M_j = I - A_j^0$ and $n = d$, Proposition 4.1 implies certain multilinear analogues of the Radon transform estimates of Oberlin and Stein [24].⁽⁵⁾ We leave the details of these implications to the interested reader.

6. Variable-coefficient extensions

More general (diffeomorphism-invariant) families of oscillatory integral operators, of which the extension operators are examples, were first considered by Hörmander in [17]. Hörmander conjectured that, under certain natural nondegeneracy conditions on the associated phase function (see [26]), such operators would satisfy $L^p \rightarrow L^q$ estimates in agreement with the classical restriction conjecture. It is now well known that this conjecture is in general false; see Bourgain [9]. In this section, we consider the validity of such generalisations of the *multilinear* restriction problem discussed in the introduction, and obtain almost optimal results in this setting.

Let $\Phi: \mathbf{R}^{d-1} \times \mathbf{R}^d \rightarrow \mathbf{R}$ be a smooth phase function, $\lambda > 0$ and $\psi: \mathbf{R}^{d-1} \times \mathbf{R}^d \rightarrow \mathbf{R}$ be a compactly supported smooth cut-off function. We define the operator S_λ by

$$S_\lambda g(\xi) := \int_{\mathbf{R}^{d-1}} e^{i\lambda\Phi(x,\xi)} \psi(x,\xi) g(x) dx,$$

and the vector field $X(\Phi)$ by

$$X(\Phi) := \bigwedge_{k=1}^{d-1} \frac{\partial}{\partial x_k} \nabla_\xi \Phi.$$

Remark 6.1. When the phase Φ takes the form $\Phi(x,\xi) = x \cdot \Sigma(\xi)$, then the operator S_λ is essentially (up to rescaling and cutoffs) an extension operator \mathcal{E} , and $X(\Phi)$ is essentially the vector field Y .

Now, we suppose that $S_\lambda^{(1)}, \dots, S_\lambda^{(d)}$ are such operators associated to phase functions Φ_1, \dots, Φ_d , and cut-off functions ψ_1, \dots, ψ_d .

⁽⁵⁾ It turns out that when all the A_j 's have rank 1, a nonperturbative linear analysis is rather straightforward. We shall return to such matters at a later date.

Our generalisation of the multilinear transversality condition, which we will impose from this point on, will be that for some constant $\nu > 0$,

$$\det(X(\Phi_1)(x^{(1)}, \xi), \dots, X(\Phi_d)(x^{(d)}, \xi)) > \nu \tag{41}$$

for all $(x^{(1)}, \xi) \in \text{supp}(\psi_1), \dots, (x^{(d)}, \xi) \in \text{supp}(\psi_d)$. In addition to this, for each multi-index $\beta \in \mathbf{N}^{d-1}$ let us suppose that for some constant $A_\beta \geq 0$,

$$\|\partial_x^\beta \Phi_j(x, \cdot)\|_{C_\xi^2(\mathbf{R}^d)} \leq A_\beta \quad \text{for all } 1 \leq j \leq d \text{ and } x \in \mathbf{R}^{d-1}. \tag{42}$$

THEOREM 6.2. *If (41) and (42) hold, then for each $\varepsilon > 0$, $q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$, there is a constant $C > 0$, depending only on $\varepsilon, p, q, d, \nu$ and finitely many of the A_β 's, for which*

$$\left\| \prod_{j=1}^d S_\lambda^{(j)} g_j \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C_\varepsilon \lambda^\varepsilon \prod_{j=1}^d \lambda^{-d/q} \|g_j\|_{L^p(\mathbf{R}^{d-1})}$$

for all $g_1, \dots, g_d \in L^p(\mathbf{R}^{d-1})$ and $\lambda > 0$.

Remark 6.3. As with Theorem 1.16, it is possible that the above inequality continues to hold with $\varepsilon = 0$, although we have been unable to prove this; for instance, the number of A_β 's that we need in our argument goes to infinity as $\varepsilon \rightarrow 0$, though this may well be unnecessary. It has been shown recently (in [5], by very different techniques), that if $\Phi_j(x, \xi) = x \cdot \Gamma_j(\xi)$, where $\Gamma_j: \mathbf{R}^d \rightarrow \mathbf{R}^{d-1}$ are smooth submersions, then one may indeed set $\varepsilon = 0$ in the conclusion of Theorem 6.2. In this particular example, $S_\lambda^{(j)} g_j(\xi)$ is essentially $\widehat{g}_j \circ \Gamma_j(\lambda \xi)$, and so at the sharp endpoint ($p = 2$ and $q = 2d/(d-1)$), by Plancherel's theorem one may see this example as a nonlinear generalisation of the Loomis–Whitney inequality (5).

Remark 6.4. The implicit conditions $X(\Phi_1) \neq 0, \dots, X(\Phi_d) \neq 0$ are of course equivalent to the statement that the matrices $\partial^2 \Phi_1 / \partial x \partial \xi, \dots, \partial^2 \Phi_d / \partial x \partial \xi$ all have full rank $d-1$. We point out that in general, given (41), one cannot expect further nondegeneracy assumptions on the phase functions Φ_j to lead to improvements in the claimed range of the exponents here. Again, this is very much in contrast with what happens at lower levels of multilinearity.

We now come to the corresponding variable-coefficient multilinear Kakeya-type problem. For a discussion of the original linear setting, see Wisewell [32].

For each $1 \leq j \leq d$, let \mathbf{T}_j denote a collection of subsets of \mathbf{R}^d of the form

$$\{\xi \in \mathbf{R}^d : |\nabla_x \Phi_j(a, \xi) - \omega| \leq \delta \text{ and } (a, \xi) \in \text{supp}(\psi_j)\},$$

where $a, \omega \in \mathbf{R}^{d-1}$. It is important for us to observe that the conditions (41) and (42) (with $|\beta| \leq 1$) imply that these sets contain, and are contained in, $O(\delta)$ -neighbourhoods of smooth curves in \mathbf{R}^d . For this reason it is convenient to extend the use of our tube notation and terminology from the previous sections. The implicit constants in the O -notation here depend on d, ν and the A_β 's with $|\beta| \leq 1$ (these quantities appear in (42)).

THEOREM 6.5. *If (41) and (42) hold, then for each $\varepsilon > 0$ and $q \geq d/(d-1)$ there exists a constant $C > 0$, depending only on ε, q, d, ν and finitely many of the A_β 's,⁽⁶⁾ such that*

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C_\varepsilon \delta^{-\varepsilon} \prod_{j=1}^d (\delta^{d/q} \#\mathbf{T}_j)$$

for all collections $\mathbf{T}_1, \dots, \mathbf{T}_d$ and $\delta > 0$.

Remark 6.6. Again, it seems possible that the above inequality continues to hold with $\varepsilon = 0$. Notice that in the special case described in Remark 6.3, the corresponding $T_j \in \mathbf{T}_j$ are simply unions of fibres of the submersions Γ_j .

The proofs of both Theorems 6.2 and 6.5 follow bootstrapping arguments closely related to that of Bourgain used in §2.⁽⁷⁾ For these we need some further notation.

For $\alpha > 0, q \geq 2d/(d-1)$ and $p' \leq q(d-1)/d$, let

$$\mathcal{R}_c^*(p \times \dots \times p \rightarrow q; \alpha)$$

denote the multilinear oscillatory integral estimate

$$\left\| \prod_{j=1}^d S_\lambda^{(j)} g_j \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \lambda^\alpha \prod_{j=1}^d \lambda^{-d/q} \|g_j\|_{L^2(\mathbf{R}^{d-1})},$$

for some constant C (depending only on α, p, q, d, ν and finitely many of the A_β 's), all $g_1, \dots, g_d \in L^2(\mathbf{R}^{d-1})$ and $\lambda > 0$.

Similarly, for each $\alpha > 0$ and $q \geq d/(d-1)$ let

$$\mathcal{K}_c^*(1 \times \dots \times 1 \rightarrow q; \alpha)$$

denote the multilinear ‘‘curvy’’ Keakeya estimate

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \delta^{-\alpha} \prod_{j=1}^d (\delta^{d/q} \#\mathbf{T}_j), \tag{43}$$

for a similarly uniform constant C , all $\delta > 0$ and all families $\mathbf{T}_1, \dots, \mathbf{T}_d$. (Of course the hypotheses (41) and (42) are assumed implicitly here.)

⁽⁶⁾ In fact, one only really sees the A_β 's with $|\beta| = 1$ here; this is easily seen from our proof.

⁽⁷⁾ Variants of such bootstrapping arguments have been considered previously by both Wolff and the third author, although not in a multilinear setting.

Remark 6.7. Inequality (43) is easily seen to be equivalent to the superficially stronger

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \lambda_{T_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)} \leq C \delta^{-\alpha} \prod_{j=1}^d \left(\delta^{d/q} \sum_{T_j \in \mathbf{T}_j} \lambda_{T_j} \right), \quad (44)$$

uniformly in the non-negative constants λ_{T_j} ($T_j \in \mathbf{T}_j$, $1 \leq j \leq d$).

Given that the estimate $\mathcal{K}^*(1 \times \dots \times 1 \rightarrow q; 0)$ holds for all $q > d/(d-1)$ (Theorem 1.15), we may reduce the proof of Theorem 6.5 to a repeated application of the following result. Note that, by Hölder’s inequality, it suffices to treat $q > d/(d-1)$.

PROPOSITION 6.8. *For each $\alpha, \varepsilon > 0$ and $q > d/(d-1)$,*

$$\mathcal{K}_c^*(1 \times \dots \times 1 \rightarrow q; \alpha) \text{ and } \mathcal{K}^*(1 \times \dots \times 1 \rightarrow q; \varepsilon) \implies \mathcal{K}_c^*(1 \times \dots \times 1 \rightarrow q; \tfrac{1}{2}\alpha + \tfrac{1}{2}\varepsilon).$$

Given Theorem 6.5 (and Hölder’s inequality), we may similarly reduce the proof of Theorem 6.2 to a bootstrapping argument.⁽⁸⁾

PROPOSITION 6.9. *For each $\alpha, \varepsilon, \varepsilon_0 > 0$ and $q > 2d/(d-1)$,*

$$\mathcal{R}_c^*(2 \times \dots \times 2 \rightarrow q; \alpha) \text{ and } \mathcal{K}_c^*(1 \times \dots \times 1 \rightarrow \tfrac{1}{2}q; \varepsilon) \implies \mathcal{R}_c^*(2 \times \dots \times 2 \rightarrow q; \tfrac{1}{2}\alpha + \tfrac{1}{4}\varepsilon + \varepsilon_0).$$

Proof of Proposition 6.8. Let $\{B\}$ be a tiling of \mathbf{R}^d by cubes of side $\sqrt{\delta}$, and write

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)}^{q/d} = \sum_B \left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j \cap B} \right\|_{L^{q/d}(\mathbf{R}^d)}^{q/d}.$$

Now, by the smoothness of the Φ_j ’s, each $T_j \cap B$ is contained in a rectangular tube of dimensions $O(\delta) \times \dots \times O(\delta) \times O(\sqrt{\delta})$, and furthermore (upon rescaling) the d families of these “rectangular” tubes $\{T_j \cap B\}_{T_j \in \mathbf{T}_j}$ ($1 \leq j \leq d$) have the transversality property required by the hypothesis $\mathcal{K}^*(1 \times \dots \times 1 \rightarrow q; \varepsilon)$. Hence

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j \cap B} \right\|_{L^{q/d}(\mathbf{R}^d)} \lesssim \delta^{-\varepsilon/2} \prod_{j=1}^d \delta^{d/q} \#\{T_j \in \mathbf{T}_j : T_j \cap B \neq \emptyset\},$$

uniformly in B . We note that the implicit constants in the O notation above depend only on d, ν and the A_β ’s with $|\beta|=1$.⁽⁹⁾ We next observe the elementary fact that

$$\#\{T_j \in \mathbf{T}_j : T_j \cap B \neq \emptyset\} \lesssim \sum_{T_j \in \mathbf{T}_j} \chi_{T_j + B(0, c\sqrt{\delta})}(\xi_B),$$

⁽⁸⁾ In principle, one ought to be able to prove Theorem 6.2 directly using a suitable variant of Theorem 1.16 along with Theorem 6.5. We do not pursue this matter here.

⁽⁹⁾ Here we are using the C_ξ^2 -control of the phase functions Φ_j to guarantee that the tubes are “locally straight” in the claimed way.

uniformly in $\xi_B \in B$, where c is a sufficiently large constant, again depending only on d, ν and the A_β 's with $|\beta|=1$. Now, each T_j is given by

$$T_j = \{\xi \in \mathbf{R}^d : |\nabla_x \Phi_j(a_j, \xi) - \omega_j| \leq \delta \text{ and } (a_j, \xi) \in \text{supp}(\psi_j)\}$$

for some $a_j, \omega_j \in \mathbf{R}^{d-1}$, and so, by our geometric interpretation of such tubes, we see that $T_j + B(0, c\sqrt{\delta}) \subset \tilde{T}_j$, where

$$\tilde{T}_j := \{\xi \in \mathbf{R}^d : |\nabla_x \Phi_j(a_j, \xi) - \omega_j| \lesssim \sqrt{\delta} \text{ and } (a_j, \xi) \in \text{supp}(\psi_j) + B(0, O(\sqrt{\delta}))\},$$

yielding

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)} \lesssim \delta^{-\varepsilon/2} \left(\sum_B \left(\prod_{j=1}^d \delta^{d/q} \sum_{\tilde{T}_j} \chi_{\tilde{T}_j}(\xi_B) \right)^{q/d} \right)^{d/q},$$

uniformly in the choice of ξ_B . Hence, upon averaging, we obtain

$$\left\| \prod_{j=1}^d \sum_{T_j \in \mathbf{T}_j} \chi_{T_j} \right\|_{L^{q/d}(\mathbf{R}^d)} \lesssim \delta^{-\varepsilon/2 + d^2/2q} \left\| \prod_{j=1}^d \sum_{\tilde{T}_j: T_j \in \mathbf{T}_j} \chi_{\tilde{T}_j} \right\|_{L^{q/d}(\mathbf{R}^d)},$$

which, by the remaining hypothesis⁽¹⁰⁾ $\mathcal{K}_c^*(1 \times \dots \times 1 \times q; \alpha)$, is

$$\lesssim \delta^{-\varepsilon/2 - \alpha/2} \prod_{j=1}^d \delta^{d/q} \#\mathbf{T}_j,$$

completing the proof of the proposition. □

Remark 6.10. Our approach to Proposition 6.9 is somewhat different technically from that of Proposition 2.1. This is largely due to our desire to avoid formulating a potentially cumbersome generalisation of Lemma 2.2. The drawback of the resulting (slightly cruder) argument, which is largely aesthetic, is the additional ε -loss present in the statement, and the role played by the high order derivatives in (42).

Proof of Proposition 6.9. The argument that we give here has essentially one additional ingredient to that of Proposition 6.8—a “wavepacket” decomposition, and thus one is forced to deal with the additional technicalities associated with the uncertainty principle; i.e. the fact that wavepackets are not genuinely supported on tubes, but rather

⁽¹⁰⁾ There is an extremely minor issue here, which is that the tubes \tilde{T}_j are defined with (a_j, ξ) ranging in $\text{supp}(\psi_j) + B(0, O(\sqrt{\delta}))$ rather than $\text{supp}(\psi_j)$. But this negligible enlargement of the support can be dealt with by modifying ψ_j very slightly, and checking that the various bounds on the geometry do not change very much. We omit the details.

decay rapidly away from them. Let $\{Q\}$ be a tiling of \mathbf{R}^{d-1} by cubes Q of side $\lambda^{-1/2}$, and for such a Q , let x_Q be its centre. We now decompose each g_j into local Fourier series at an appropriate scale. Let $\{\chi_Q\}$ be a (smooth) partition of unity adapted to the tiling $\{Q\}$. For uniformity purposes, let us suppose that $\chi_Q(x) = \chi(\lambda^{1/2}(x - x_Q))$ for some smooth compactly supported function χ . Now, for each g_j , we may write

$$g_j = \sum_Q \sum_{l \in \lambda^{-1/2} \mathbf{Z}^{d-1}} a_{Q,l}^{(j)} e_{Q,l},$$

where $e_{Q,l}$ is the modulated cap

$$e_{Q,l}(x) := \chi_Q(x) e^{il \cdot x},$$

and the $a_{Q,l}^{(j)}$'s are complex numbers. By linearity of $S_\lambda^{(j)}$,

$$S_\lambda^{(j)} g_j = \sum_{Q,l} a_{Q,l}^{(j)} S_\lambda^{(j)} e_{Q,l}.$$

Now, we localise the $S_\lambda^{(j)} e_{Q,l}$ to tubes. Let $\eta := q\varepsilon_0/d^2$. For each $l \in \lambda^{-1/2} \mathbf{Z}^{d-1}$ and $1 \leq j \leq d$, let $R_{Q,l}^{(j)}$ be the curved tube

$$R_{Q,l}^{(j)} := \{\xi \in \mathbf{R}^d : |\nabla_x \Phi_j(x_Q, \xi) - l| \leq \lambda^{-1/2+\eta} \text{ and } (x_Q, \xi) \in \text{supp}(\psi_j)\}.$$

By a standard repeated integration by parts argument, we have that for each $M \in \mathbf{N}$,

$$|S_\lambda^{(j)} e_{Q,l}(\xi)| \lesssim \lambda^{-M\eta - (d-1)/2}, \tag{45}$$

for all $\xi \in \mathbf{R}^d \setminus R_{Q,l}^{(j)}$. Naturally the implicit constants here depend on d and the smoothness bounds A_β with $|\beta| \leq M$.

We now tile \mathbf{R}^d by cubes B of side $\lambda^{-1/2}$. The idea is to use the oscillatory integral estimate in our hypothesis on each B , and then use the curvy Kakeya estimate to reassemble them. Let \mathcal{P} denote the set of nonempty subsets P of the set of integers $\{1, \dots, d\}$, and for each $P \in \mathcal{P}$ let P^c denote the complement of P in $\{1, \dots, d\}$. Now, by the triangle inequality,

$$\left\| \prod_{j=1}^d S_\lambda^{(j)} g_j \right\|_{L^{q/d}(\mathbf{R}^d)} = \left(\sum_B \left\| \prod_{j=1}^d S_\lambda^{(j)} g_j \right\|_{L^{q/d}(B)}^{q/d} \right)^{d/q} \leq I + \sum_{P \in \mathcal{P}} I_P,$$

where

$$I := \left(\sum_B \left\| \prod_{j=1}^d \left(\sum_{Q,l: R_{Q,l}^{(j)} \cap B \neq \emptyset} a_{Q,l}^{(j)} S_\lambda^{(j)} e_{Q,l} \right) \right\|_{L^{q/d}(B)}^{q/d} \right)^{d/q}$$

and I_P is given by

$$I_P^{q/d} := \sum_B \left\| \left(\prod_{j \in P} \sum_{Q,l: R_{Q,l}^{(j)} \cap B = \emptyset} a_{Q,l}^{(j)} S_\lambda^{(j)} e_{Q,l} \right) \left(\prod_{k \in P^c} \sum_{Q,l: R_{Q,l}^{(k)} \cap B \neq \emptyset} a_{Q,l}^{(k)} S_\lambda^{(k)} e_{Q,l} \right) \right\|_{L^{q/d}(B)}^{q/d}.$$

We first estimate the principal term I . By rescaling the hypothesis $\mathcal{R}_c^*(2 \times \dots \times 2 \rightarrow q; \alpha)$, and observing the scale-invariance⁽¹¹⁾ of conditions (41) and (42), we have that

$$I \lesssim \lambda^{\alpha/2+d^2/q} \left(\sum_B \prod_{j=1}^d \left\| \sum_{Q,l: R_{Q,l}^{(j)} \cap B \neq \emptyset} a_{Q,l}^{(j)} e_{Q,l} \right\|_2^{q/d} \right)^{d/q},$$

which, by the almost orthogonality of the $e_{Q,l}$'s, is further bounded by

$$\begin{aligned} & \lambda^{\alpha/2-d^2/q} \left(\sum_B \prod_{j=1}^d \left(\sum_Q \lambda^{-(d-1)/2} \sum_{l: R_{Q,l}^{(j)} \cap B \neq \emptyset} |a_{Q,l}^{(j)}|^2 \right)^{q/2d} \right)^{d/q} \\ & \lesssim \lambda^{\alpha/2-d^2/q-d(d-1)/4} \left(\sum_B \prod_{j=1}^d \left(\sum_{Q,l} |a_{Q,l}^{(j)}|^2 \chi_{R_{Q,l}^{(j)}}(\xi_B) \right)^{q/2d} \right)^{d/q}, \end{aligned}$$

uniformly in $\xi_B \in B$. Strictly speaking, the tubes in this last line should be replaced by slightly dilated versions of themselves; however, we shall gloss over this detail. On averaging and applying the remaining hypothesis $\mathcal{K}_c^*(1 \times \dots \times 1 \rightarrow \frac{1}{2}q; \varepsilon)$ (in the equivalent form (44)) to the above expression, we obtain

$$\begin{aligned} I & \lesssim \lambda^{\alpha/2-d^2/q-d(d-1)/4} \left(\lambda^{-d/2} \int_{\mathbf{R}^d} \left(\prod_{j=1}^d \sum_{Q,l} |a_{Q,l}^{(j)}|^2 \chi_{R_{Q,l}^{(j)}}(x) \right)^{q/2d} dx \right)^{d/q} \\ & \lesssim \lambda^{\alpha/2-d^2/2q-d(d-1)/4+\varepsilon(1-2\eta)/4-(1-2\eta)d^2/2q} \left(\prod_{j=1}^d \sum_{Q,l} |a_{Q,l}^{(j)}|^2 \right)^{1/2} \\ & \lesssim \lambda^{\alpha/2+\varepsilon/4+\varepsilon_0} \prod_{j=1}^d \lambda^{-d/q} \|g_j\|_2. \end{aligned}$$

In the last step, we have used Plancherel's theorem and the fact that $\eta = q\varepsilon_0/d^2$.

It is enough now to show that for each $P \in \mathcal{P}$ and $N > 0$ we have the error estimates

$$I_P \lesssim \lambda^{-N\eta} \prod_{j=1}^d \|g_j\|_2.$$

However, this is an elementary consequence of Hölder's inequality and the decay estimate (45). This proves Proposition 6.9. □

⁽¹¹⁾ There is a minor technical issue here. The terms in condition (42) containing zero ξ -derivatives actually fail to be invariant in the appropriate way. However, this may be easily rectified by subtracting off harmless affine factors from the phases Φ_j , and absorbing them into the functions g_j .

7. An application to the joints problem

We now give an application of the multilinear Kakeya estimate (Theorem 1.15) to a discrete geometry problem, namely the “joints” problem studied in [12], [25] and [16].

Let us recall the setup for this problem. Consider a collection L of n lines in \mathbf{R}^3 . Define a *joint* to be a point in \mathbf{R}^3 which is contained in at least one triple (l, l', l'') of concurrent lines in L which are not coplanar. (Note that a single joint may arise from multiple triples, but in such cases we only count those joints once.) The joints problem is to determine, for each fixed n , the maximum number of joints one can attain for a configuration L of n lines. This problem was observed to be formally related to the Kakeya problem in [33]; in this paper we establish for the first time a rigorous connection between the two problems.⁽¹²⁾

An easy lattice construction (where the lines are parallel to the coordinate axes and have two of the coordinates fixed to be integers between 1 and \sqrt{n}) shows that one can have at least $\sim n^{3/2}$ joints. In the other direction, one trivially observes that each line in L can contain at most n joints, and hence we have an upper bound of n^2 for the total number of joints. There has been some progress in improving the upper bound; the most recent result in [12] shows that the number of joints is at most $O(n^{112/69} \log^{6/23} n) \leq O(n^{1.6232})$. It is tentatively conjectured that the lower bound of $n^{3/2}$ is essentially sharp up to logarithms.

It turns out that the multilinear Kakeya estimate in Theorem 1.15 can support this conjecture, provided that the joints are sufficiently transverse. For any $0 < \theta \leq 1$, let us say that three concurrent lines l, l' and l'' are θ -*transverse* if the parallelepiped generated by the unit vectors parallel to l, l' , and l'' (henceforth referred to as the *directions* of l, l' and l'') has volume at least θ . Let us define a θ -*transverse joint* to be a point in \mathbf{R}^3 which is contained in at least one triple (l, l', l'') of θ -transverse concurrent lines in L . Note that every joint is θ -transverse for some θ .

THEOREM 7.1. *For any $0 < \theta \leq 1$, the number of θ -transverse joints is*

$$O_\varepsilon(n^{3/2+\varepsilon} \theta^{-1/2-\varepsilon})$$

for any $\varepsilon > 0$, where the subscripting of O by ε means that the implied constant can depend on ε .

This theorem suggests that the hard case of the joints problem arises when considering nearly-coplanar joints, with different joints being approximately coplanar in different

⁽¹²⁾ We note that this problem has an equally natural formulation in higher dimensions, where similar arguments apply.

orientations. This resembles the experience in [18], when the “plany” case of the Kakeya problem was by far the most difficult to handle.

Proof. We first establish this conjecture in the case $\theta \sim 1$. We cover the unit sphere \mathbf{S}^2 by $O(1)$ finitely overlapping caps C_1, \dots, C_k of width $\theta/1000$. Observe that if l, l' and l'' are a θ -transverse collection of lines, then the directions of l, l' and l'' will lie in three distinct caps $C_i, C_{i'}$ and $C_{i''}$, which are transverse in the sense of (1). Since the number of such triples of caps is $O(1)$, it thus suffices to show that

$$\#\{p \in \mathbf{R}^3 : p \in l, l', l'' \text{ for some } l \in L_i, l' \in L_{i'} \text{ and } l'' \in L_{i''}\} = O(n^{3/2}) \tag{46}$$

for each such transverse triple $(C_i, C_{i'}, C_{i''})$, where L_i is the collection of lines in L with directions in C_i .

By rescaling, we may assume that all the joints are contained in the ball of radius $1/1000$ centred at the origin. Let $\delta > 0$ be a small parameter (eventually it will go to zero), and for each line $l \in L$, let T_l denote the $\delta \times \dots \times \delta \times 1$ tube with axis l and centre equal to the closest point of l to the origin. Let \mathbf{T}_i denote the collection of all the tubes T_l associated to lines l in L_i , and similarly define $\mathbf{T}_{i'}$ and $\mathbf{T}_{i''}$. From elementary geometry, we see that if p is an element of the set in (46), then we have

$$\sum_{T_i \in \mathbf{T}_i} \chi_{T_i}(x) \geq 1$$

whenever $|x - p| < c\delta$, where $c > 0$ is a small absolute constant depending on the transversality constant of $(C_i, C_{i'}, C_{i''})$. Similarly for $\mathbf{T}_{i'}$ and $\mathbf{T}_{i''}$. Since the number of joints is finite, we see that for δ sufficiently small, the balls $\{x \in \mathbf{R}^3 : |x - p| < c\delta\}$ will be disjoint. We conclude that

$$\left\| \left(\sum_{T_i \in \mathbf{T}_i} \chi_{T_i} \right) \left(\sum_{T_{i'} \in \mathbf{T}_{i'}} \chi_{T_{i'}} \right) \left(\sum_{T_{i''} \in \mathbf{T}_{i''}} \chi_{T_{i''}} \right) \right\|_{L^{q/3}(\mathbf{R}^3)} \geq c_q N^{3/q} \delta^{9/q}$$

for any $\frac{3}{2} < q \leq \infty$, where N denotes the left-hand side of (46) and $c_q > 0$ is a constant depending only on c and q . Applying Theorem 1.15, we obtain

$$N^{3/q} \delta^{9/q} \leq C_q (\delta^{3/q} \#L_i) (\delta^{3/q} \#L_{i'}) (\delta^{3/q} \#L_{i''}),$$

or in other words

$$N \leq C_q^{q/3} (\#L_i \#L_{i'} \#L_{i''})^{q/3}. \tag{47}$$

Since $\#L_i, \#L_{i'}, \#L_{i''} \leq n$ and q can be arbitrarily close to $\frac{3}{2}$, the claim follows.

Now, we handle the case when θ is much smaller than 1, using some (slightly inefficient) trilinear variants of the bilinear rescaling arguments employed in [30].

Suppose that (l, l', l'') are θ -transverse. Each pair of lines in (l, l', l'') determines an angle; without loss of generality, we may take l and l' to subtend the largest angle. Calling this angle α , we see from elementary geometry that $\theta^{1/2} \lesssim \alpha \lesssim 1$, and that l'' makes an angle of at least $\gtrsim \theta/\alpha$ and at most α with respect to the plane spanned by l and l' . To exploit this, let us say that (l, l', l'') are (α, β) -transverse for some $0 < \beta \lesssim \alpha \lesssim 1$ if l and l' make an angle of $\sim \alpha$ and l'' makes an angle of $\sim \beta$ with respect to the plane spanned by l and l' . Define an (α, β) -transverse joint similarly. A simple dyadic decomposition argument (giving up some harmless factors of $\log 1/\theta$) then shows that it suffices to prove that the number of (α, β) -transverse joints is $O_\varepsilon(n^{3/2+\varepsilon}(\alpha\beta)^{-1/2+\varepsilon})$ for every $\varepsilon > 0$. In fact, we will prove the sharper bound of $O_\varepsilon(n^{3/2+\varepsilon}(\beta/\alpha)^{-1/2+\varepsilon})$.

Let us first handle the case when $\alpha \sim 1$, so that l and l' make an angle of ~ 1 . By symmetry, we may also assume that l'' makes a smaller angle with l' than it does with l , so l and l'' also make an angle of ~ 1 . By a decomposition of the sphere into $O(1)$ pieces, we can then assume that there exist transverse subsets S and S' of the sphere such that the direction of l lies in S , and the directions of l' and l'' lie in S' . (Note that S' may be somewhat larger than S .)

Let $\omega_1, \dots, \omega_K$ be a maximal β -separated set of directions on the sphere, thus $K = O(1/\beta^2)$. For each direction ω_k , let L_k denote the family of lines with direction in S which make an angle of $\pi/2 - O(\beta)$ with ω_k , thus they are nearly orthogonal to ω_k . Define L'_k similarly, but with S replaced by S' . From elementary geometry, we see that if (l, l', l'') are (α, β) -transverse, then there exists k such that $l \in L_k$ and $l', l'' \in L'_k$. Thus, the number of (α, β) -transverse joints can be bounded by

$$\sum_{k=1}^K \#\{p : p \in l, l', l'' \text{ for some } (\alpha, \beta)\text{-transverse } l \in L_k \text{ and } l', l'' \in L'_k\}.$$

Next, observe from elementary geometry that if (l, l', l'') are (α, β) -transverse in $L_k \cup L'_k$, then, after applying a dilation by $1/\beta$ in the ω_k direction, the resulting lines become c -transverse for some $c \sim 1$ (here we are using the hypothesis that $\alpha \sim 1$). Applying (47), we conclude that

$$\#\{p : p \in l, l', l'' \text{ for some } (\alpha, \beta)\text{-transverse } l \in L_k \text{ and } l', l'' \in L'_k\} = O_\varepsilon(n^\varepsilon)(\#L_k)^{1/2} \#L'_k,$$

so it suffices to establish the bound

$$\sum_{k=1}^K (\#L_k)^{1/2} \#L'_k \leq C n^{3/2} \beta^{-1/2}.$$

Now, observe from transversality of S and S' that if l has direction in S and l' has direction in S' , then there are at most $O(1)$ values of k for which $l \in L_k$ and $l' \in L'_k$. This

leads to the bound

$$\sum_{k=1}^K \#L_k \#L'_k \leq Cn^2.$$

On the other hand, observe that every line l' belongs to at most $O(1/\beta)$ families L'_k . This leads to the bound

$$\sum_{k=1}^K \#L'_k \leq Cn\beta^{-1}.$$

The claim now follows from the Cauchy–Schwarz inequality.

Finally, we handle the case when α is very small, using the bilinear rescaling argument from [30]. Let $\tilde{\omega}_1, \dots, \tilde{\omega}_{\tilde{K}}$ be a maximal α -separated set of directions of the sphere, and for each $\tilde{\omega}_k$ let \tilde{L}_k be all the lines in L which make an angle of $O(\alpha)$ with $\tilde{\omega}_k$. Observe that if (l, l', l'') are (α, β) -transverse, then there exists k such that all of l, l' and l'' lie in \tilde{L}_k . Thus, we can bound the total number of (α, β) -joints in this case by

$$\sum_{k=1}^{\tilde{K}} \#\{p : p \in l, l', l'' \text{ for some } (\alpha, \beta)\text{-transverse } l, l', l'' \in \tilde{L}_k\}.$$

Next, observe from elementary geometry that if $l, l', l'' \in \tilde{L}_k$ are (α, β) -transverse, then if we dilate l, l', l'' in the directions orthogonal to $\tilde{\omega}_k$ by $1/\alpha$, then the resulting triple of lines becomes $(1, \beta/\alpha)$ -transverse. Since we have already established the desired bound in the $\alpha \sim 1$ case, we conclude that

$$\#\{p : p \in l, l', l'' \text{ for some } (\alpha, \beta)\text{-transverse } l, l', l'' \in \tilde{L}_k\} \leq O_\varepsilon(n^\varepsilon)(\#\tilde{L}_k)^{3/2+\varepsilon}(\beta/\alpha)^{-1/2-\varepsilon}$$

and so it will suffice to show that

$$\sum_{k=1}^{\tilde{K}} (\#\tilde{L}_k)^{3/2+\varepsilon} \leq Cn^{3/2+\varepsilon}.$$

Using the crude bound $\#\tilde{L}_k \leq n$, it suffices to show that

$$\sum_{k=1}^{\tilde{K}} \#\tilde{L}_k \leq Cn.$$

But it is clear that each line $l \in L$ can belong to at most $O(1)$ families \tilde{L}_k , and the claim follows. □

Remark 7.2. If one had the endpoint $q=d/(d-1)$ in Conjecture 1.8, then one could remove the ε -losses from the n exponent, and possibly from the θ exponent as well. The deterioration of the bound as $\theta \rightarrow 0$ is closely related to the reason that the multilinear Kakeya estimate is currently unable to imply any corresponding linear Kakeya estimate. Thus, a removal of this θ -dependence in the joints estimate may lead to a new *linear* Kakeya estimate.

Remark 7.3. One can also phrase the joints problem for other families of curves than lines, in the spirit of §6. If one could remove the loss of $\delta^{-\varepsilon}$ in Theorem 6.5, one could obtain a result similar to Theorem 7.1 in this setting, but, as Theorem 6.5 stands, one would only obtain a rather unaesthetic result in which certain “entropy numbers” of the joints are controlled. We omit the details.

8. Appendix: A polynomial extrapolation lemma

The main aim of this paper (the contents of §3 and §4) is to obtain monotonicity *formulae* for spatial L^p -norms of certain multilinear expressions. As we have seen, this can be done quite explicitly for integer values of the exponent p , and in such a way that the identities obtained *make sense* at least for noninteger p . The pay-off of having proved such precise *identities* for $p \in \mathbf{N}$ is that we may use a density argument (e.g. using the Weierstrass approximation theorem) to deduce that they must also hold for $p \notin \mathbf{N}$. This is very much analogous to the classical result that a compactly supported probability distribution is determined uniquely by its moments.

Remark 8.1. As we noted in §3, there is a satisfactory way of avoiding this “integer p first” approach to the unperturbed situation (Proposition 3.1). This involves finding an appropriate function of divergence form which differs from the integrand in (22) by a manifestly non-negative quantity. See [4] and [11]. In principle, one could take a similar approach to Proposition 4.1, although as yet it seems quite unclear how to directly exhibit an appropriate divergence term.

LEMMA 8.2. *Let $f_1, \dots, f_n: \mathbf{R}^d \rightarrow \mathbf{R}$ be non-negative bounded measurable functions for which the product $f_1 \dots f_n$ is rapidly decreasing. Suppose that $G_1, G_2: \mathbf{R}^n \times \mathbf{R}^d \rightarrow \mathbf{R}$ are polynomial in their first n variables $p = (p_1, \dots, p_n)$, with coefficients which are measurable, and of polynomial growth in their last d ones. If the identity*

$$\int_{\mathbf{R}^d} G_1(p, x) f_1(x)^{p_1} \dots f_n(x)^{p_n} dx = \int_{\mathbf{R}^d} G_2(p, x) f_1(x)^{p_1} \dots f_n(x)^{p_n} dx$$

holds for all $p \in \mathbf{N}^n$, then it holds for all $p \in (0, \infty)^n$.

Proof. By linearity, we may assume that $G_2 = 0$ and rename G_1 as G . Write

$$G(p, x) = \sum_{|\alpha| \leq N} p^\alpha w_\alpha(x),$$

where w_α is measurable and of polynomial growth. We may further assume that $\|f_j\|_\infty = 1$ for all j . Since the function

$$p \mapsto \int_{\mathbf{R}^d} G(p, x) f_1(x)^{p_1} \dots f_n(x)^{p_n} dx$$

is an analytic function of each $p_j > 0$, it suffices to prove the result when $p_j \geq N$ for all j . Let $\phi_p(t) := \phi_p(t_1, \dots, t_n) = t_1^{p_1} \dots t_n^{p_n}$ for $t \in [0, 1]^n$. Note that $\phi_p(t) = 0$ if any $t_j = 0$. Observe that we may write

$$G(p, x) f^p(x) = \sum_{|\alpha| \leq N} \tilde{w}_\alpha(x) \phi_p^{(\alpha)}(f(x)),$$

where $\tilde{w}_\alpha(x) := r_\alpha(x) \prod_{j: \alpha_j \neq 0} f_j(x)$ and $\phi_p^{(\alpha)}$ denotes differentiation of order α , and where r_α is of polynomial growth. In particular, if each $\alpha_j \geq 1$, then $\tilde{w}_\alpha \in L^1$. By hypothesis, if ϕ is any polynomial which vanishes on the coordinate axes, then

$$\int_{\mathbf{R}^d} \sum_{|\alpha| \leq N} \tilde{w}_\alpha(x) \phi^{(\alpha)}(f(x)) dx = 0. \tag{48}$$

We wish to show that the same still holds for ϕ replaced by ϕ_p when each $p_j \geq N$. For such ϕ_p 's (which belong to the class $\mathcal{C} := \{\psi \in C^N([0, 1]^n) : \psi(t) = 0 \text{ whenever some } t_j = 0\}$) we can approximate it to within any given ε by a polynomial ϕ of class \mathcal{C} in the norm $\|\psi\|_* := \max\{\|\psi^{(\alpha)}\|_\infty : \alpha_j \geq 1 \text{ for all } j \text{ and } |\alpha| \leq N\}$. So

$$\int_{\mathbf{R}^d} \sum_{|\alpha| \leq N} \tilde{w}_\alpha(x) \phi_p^{(\alpha)}(f(x)) dx = \int_{\mathbf{R}^d} \sum_{|\alpha| \leq N} \tilde{w}_\alpha(x) [\phi_p^{(\alpha)} - \phi^{(\alpha)}](f(x)) dx + 0.$$

When α is such that each $\alpha_j \geq 1$, we can dominate its contribution to the right-hand side by $\int_{\mathbf{R}^d} |\tilde{w}_\alpha(x)| dx \|\phi_p - \phi\|_*$ which is as small as we like, since \tilde{w}_α is in L^1 .

When some of the α_j 's are zero, say $\alpha_1, \dots, \alpha_k = 0$ and $\alpha_{k+1}, \dots, \alpha_n \neq 0$, we set $\tilde{\alpha} = (1, \dots, 1, \alpha_{k+1}, \dots, \alpha_n)$ and write

$$\phi_p^{(\alpha)}(t) - \phi^{(\alpha)}(t) = \int_0^{t_1} \dots \int_0^{t_k} [\phi_p^{(\tilde{\alpha})} - \phi^{(\tilde{\alpha})}](s_1, \dots, s_k, t_{k+1}, \dots, t_n) ds_1 \dots ds_k$$

Thus, $[\phi_p^\alpha - \phi^\alpha](f(x)) \leq f_1(x) \dots f_k(x) \|\phi_p - \phi\|_*$, and so for these α we get

$$\int_{\mathbf{R}^d} |\tilde{w}_\alpha(x)| \cdot |[\phi_p^\alpha - \phi^\alpha](f(x))| dx \leq \int_{\mathbf{R}^d} f_1(x) \dots f_k(x) f_{k+1}(x) \dots f_n(x) |r_\alpha(x)| dx \|\phi_p - \phi\|_*,$$

which is likewise as small as we like, since r_α is of polynomial growth and the product $f_1 \dots f_n$ is rapidly decreasing.

Thus, formula (48) continues to hold for ϕ_p and we are finished. □

References

- [1] BARCELÓ, J. A., BENNETT, J. M. & CARBERY, A., A multilinear extension inequality in \mathbf{R}^n . *Bull. London Math. Soc.*, 36 (2004), 407–412.
- [2] BECKNER, W., CARBERY, A., SEMMES, S. & SORIA, F., A note on restriction of the Fourier transform to spheres. *Bull. London Math. Soc.*, 21 (1989), 394–398.
- [3] BENNETT, J., A trilinear restriction problem for the paraboloid in \mathbf{R}^3 . *Electron. Res. Announc. Amer. Math. Soc.*, 10 (2004), 97–102.
- [4] BENNETT, J., CARBERY, A., CHRIST, M. & TAO, T., The Brascamp–Lieb inequalities: finiteness, structure and extremals. To appear in *Geom. Funct. Anal.*
- [5] BENNETT, J., CARBERY, A. & WRIGHT, J., A non-linear generalisation of the Loomis–Whitney inequality and applications. *Math. Res. Lett.*, 12 (2005), 443–457.
- [6] BERGH, J. & LÖFSTRÖM, J., *Interpolation Spaces. An Introduction*. Springer, Berlin, 1976.
- [7] BLEI, R. C., Fractional Cartesian products of sets. *Ann. Inst. Fourier (Grenoble)*, 29:2 (1979), 79–105.
- [8] BOURGAIN, J., Besicovitch type maximal operators and applications to Fourier analysis. *Geom. Funct. Anal.*, 1 (1991), 147–187.
- [9] — L^p -estimates for oscillatory integrals in several variables. *Geom. Funct. Anal.*, 1 (1991), 321–374.
- [10] — Fourier transform restriction phenomena for certain lattice subsets and applications to nonlinear evolution equations. I, II. *Geom. Funct. Anal.*, 3 (1993), 107–156, 209–262.
- [11] CARLEN, E. A., LIEB, E. H. & LOSS, M., A sharp analog of Young’s inequality on S^N and related entropy inequalities. *J. Geom. Anal.*, 14 (2004), 487–520.
- [12] CHAZELLE, B., EDELSBRUNNER, H., GUIBAS, L. J., POLLACK, R., SEIDEL, R., SHARIR, M. & SNOEYINK, J., Counting and cutting cycles of lines and rods in space. *Comput. Geom.*, 1 (1992), 305–323.
- [13] CÓRDOBA, A., Multipliers of $\mathcal{F}(L^p)$, in *Euclidean Harmonic Analysis* (Proc. Sem., Univ. Maryland, College Park, MD, 1979), Lecture Notes in Math., 779, pp. 162–177. Springer, Berlin, 1980.
- [14] ERDOĞAN, M. B., A bilinear Fourier extension theorem and applications to the distance set problem. *Int. Math. Res. Not.*, 2005:23 (2005), 1411–1425.
- [15] FEFFERMAN, C., The multiplier problem for the ball. *Ann. of Math.*, 94 (1971), 330–336.
- [16] FELDMAN, S. & SHARIR, M., An improved bound for joints in arrangements of lines in space. *Discrete Comput. Geom.*, 33 (2005), 307–320.
- [17] HÖRMANDER, L., Oscillatory integrals and multipliers on FL^p . *Ark. Mat.*, 11 (1973), 1–11.
- [18] KATZ, N. H., LABA, I. & TAO, T., An improved bound on the Minkowski dimension of Besicovitch sets in \mathbf{R}^3 . *Ann. of Math.*, 152 (2000), 383–446.
- [19] KLAINERMAN, S. & MACHEDON, M., Space-time estimates for null forms and the local existence theorem. *Comm. Pure Appl. Math.*, 46 (1993), 1221–1268.
- [20] LEE, S., Improved bounds for Bochner–Riesz and maximal Bochner–Riesz operators. *Duke Math. J.*, 122 (2004), 205–232.
- [21] LIEB, E. H., Gaussian kernels have only Gaussian maximizers. *Invent. Math.*, 102 (1990), 179–208.
- [22] LOOMIS, L. H. & WHITNEY, H., An inequality related to the isoperimetric inequality. *Bull. Amer. Math. Soc.*, 55 (1949), 961–962.
- [23] MOYUA, A., VARGAS, A. & VEGA, L., Restriction theorems and maximal operators related to oscillatory integrals in \mathbf{R}^3 . *Duke Math. J.*, 96 (1999), 547–574.
- [24] OBERLIN, D. M. & STEIN, E. M., Mapping properties of the Radon transform. *Indiana Univ. Math. J.*, 31 (1982), 641–650.

- [25] SHARIR, M., On joints in arrangements of lines in space and related problems. *J. Combin. Theory Ser. A*, 67 (1994), 89–99.
- [26] STEIN, E. M., *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*. Princeton Mathematical Series, 43. Princeton University Press, Princeton, NJ, 1993.
- [27] TAO, T., Endpoint bilinear restriction theorems for the cone, and some sharp null form estimates. *Math. Z.*, 238 (2001), 215–268.
- [28] — A sharp bilinear restrictions estimate for paraboloids. *Geom. Funct. Anal.*, 13 (2003), 1359–1384.
- [29] — Recent progress on the restriction conjecture. [arXiv:math.CA/0311181](https://arxiv.org/abs/math/0311181).
- [30] TAO, T., VARGAS, A. & VEGA, L., A bilinear approach to the restriction and Keakeya conjectures. *J. Amer. Math. Soc.*, 11 (1998), 967–1000.
- [31] VARGAS, A., Restriction theorems for a surface with negative curvature. *Math. Z.*, 249 (2005), 97–111.
- [32] WISEWELL, L., Keakeya sets of curves. *Geom. Funct. Anal.*, 15 (2005), 1319–1362.
- [33] WOLFF, T. H., Recent work connected with the Keakeya problem, in *Prospects in Mathematics* (Princeton, NJ, 1996), pp. 129–162. Amer. Math. Soc., Providence, RI, 1999.
- [34] — A sharp bilinear cone restriction estimate. *Ann. of Math.*, 153 (2001), 661–698.
- [35] — *Lectures on Harmonic Analysis*. University Lecture Series, 29. Amer. Math. Soc., Providence, RI, 2003.

JONATHAN BENNETT
School of Mathematics
University of Birmingham
Birmingham, B15 2TT
U.K.
J.Bennett@bham.ac.uk

ANTHONY CARBERY
School of Mathematics
University of Edinburgh
Edinburgh, EH9 3JZ
U.K.
A.Carbery@ed.ac.uk

TERENCE TAO
Department of Mathematics
University of California
Los Angeles, CA 90095-1555
U.S.A.
tao@math.ucla.edu

Received September 30, 2005