

# The multilinear restriction estimate: almost optimality and localization

IOAN BEJENARU

The first result in this paper provides a very general  $\epsilon$ -removal argument for the multilinear restriction estimate. The second result provides a refinement of the multilinear restriction estimate in the case when some terms have appropriate localization properties; this generalizes a prior result of the author in [1].

<b>1</b>	<b>Introduction</b>	<b>599</b>
<b>2</b>	<b>Notation, basic results and reductions</b>	<b>606</b>
<b>3</b>	<b>The <math>\epsilon</math>-removal result</b>	<b>612</b>
<b>4</b>	<b>Proof of Theorem 1.3</b>	<b>619</b>
	<b>References</b>	<b>628</b>

## 1. Introduction

For  $n \geq 2$ , let  $U \subset \mathbb{R}^{n-1}$  be an open, bounded and connected neighborhood of the origin and let  $\Sigma : U \rightarrow \mathbb{R}^n$  be a smooth parametrization of an  $n - 1$ -dimensional submanifold of  $\mathbb{R}^n$  (hypersurface), which we denote by  $S = \Sigma(U)$ . By a smooth parametrization we mean that  $\Sigma$  satisfies

$$(1.1) \quad \|\partial^\alpha \Sigma\|_{L^\infty(U)} \lesssim_\alpha 1,$$

for  $|\alpha| \leq N$  for some large  $N$ . We say that  $S$  is a smooth hypersurface if it admits a parametrization satisfying (1.1). To such a parametrization of  $S$

we associate the extension operator  $\mathcal{E}$  defined by

$$\mathcal{E}f(x) = \int_U e^{ix \cdot \Sigma(\xi)} f(\xi) d\xi,$$

where  $f \in L^1(U)$  and  $x \in \mathbb{R}^n$ . This operator is closely related to a more intrinsic formulation of the extension operator:

$$\tilde{\mathcal{E}}g(x) = \int_S e^{ix \cdot \omega} g(\omega) d\sigma_S(\omega),$$

where  $g \in L^1(S; d\sigma_S)$  and  $x \in \mathbb{R}^n$ . Indeed, using the parametrization  $\Sigma$  as above we obtain  $\omega = \Sigma(\xi)$  and  $d\sigma_S(\omega) = \left| \frac{\partial \Sigma}{\partial \xi_1} \wedge \cdots \wedge \frac{\partial \Sigma}{\partial \xi_{n-1}} \right| d\xi$ , thus with  $f(\xi) = g(\Sigma(\xi)) \left| \frac{\partial \Sigma}{\partial \xi_1} \wedge \cdots \wedge \frac{\partial \Sigma}{\partial \xi_{n-1}} \right|$ , the two formulations are equivalent. We will use good parameterizations in the sense that  $\left| \frac{\partial \Sigma}{\partial \xi_1} \wedge \cdots \wedge \frac{\partial \Sigma}{\partial \xi_{n-1}} \right| \approx 1$  throughout the domain and, given that all results are in terms of  $L^p$  norms, the equivalence is carried out at the levels of results as well.

In order to avoid unnecessary technical issues, throughout this paper we assume that, in the definition of  $\mathcal{E}f$ , the support of  $f$  is a compact subset of  $U$ ; in other words  $f$  is supported away from the boundary of  $U$ . A similar assumption will be in place for the corresponding  $g$  that appears in the definition of  $\tilde{\mathcal{E}}g$ .

Given  $k$  smooth, compact hypersurfaces  $S_i \subset \mathbb{R}^n$ ,  $i = 1, \dots, k$ , where  $1 \leq k \leq n$ , we consider the following  $k$ -linear restriction estimate inequality

$$(1.2) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(\mathbb{R}^n)} \lesssim \Pi_{i=1}^k \|f_i\|_{L^2(U_i)}.$$

In a more compact format this estimate is abbreviated as follows:

$$(1.3) \quad \mathcal{R}^*(2 \times \cdots \times 2 \rightarrow p).$$

A natural condition to impose on the hypersurfaces is the standard transversality condition: there exists  $\nu > 0$  such that

$$(1.4) \quad |N_1(\zeta_1) \wedge \cdots \wedge N_k(\zeta_k)| \geq \nu,$$

for any  $\zeta_i \in S_i$ ,  $i \in \{1, \dots, k\}$ ; here  $N_i(\zeta_i)$  is the unit normal at  $\zeta_i$  to  $S_i$  and it is clear that the choice of orientation is not important.

There are two main conjectures regarding the optimal exponent in (1.3). For the generic case, when only the transversality condition is assumed, the optimal conjectured exponent for (1.3) is  $p = \frac{2}{k-1}$ . There is also the non-generic case when, in addition to the transversality condition, one assumes

also appropriate curvature conditions on the hypersurfaces  $S_i$  and in this case the optimal conjectured exponent in (1.3) is  $p = \frac{2(n+k)}{k(n+k-2)}$  (note that this value is  $< \frac{2}{k-1}$ ). This paper considers problems related to the generic case, but the results obtained here will have implications for the non-generic case.

In [5] Bennett, Carbery and Tao established the near-optimal version of the conjectured result in the generic case: for any  $\epsilon > 0$  and for any  $R > 0$  and  $x_0 \in \mathbb{R}^n$  the following holds true

$$(1.5) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(B_R(x_0))} \lesssim_\epsilon R^\epsilon \Pi_{i=1}^k \|f_i\|_{L^2(U_i)},$$

where  $B_R(x_0)$  is the ball of radius  $R$  centered at  $x_0$ . In a more compact form, this results is saying that  $\mathcal{R}^*(2 \times \cdots \times 2 \rightarrow p, \epsilon)$  holds true.

Closely related to this estimate, there is the multilinear Kakeya estimate. The multilinear Kakeya estimate follows from the multilinear restriction estimate by using a standard Rademacher-function argument, see [5] for details. Vice-versa, one can obtain the multilinear restriction estimate from the multilinear Kakeya estimate using a more delicate argument which incurs losses. In [5] these losses are of type  $R^\epsilon$ . We do not formalize the multilinear Kakeya estimate as it plays no role in our analysis; however it is important that we mention it for reference purposes.

An alternative and shorter proof for the near-optimal multilinear Kakeya estimate was provided by Guth in [9]. An alternative and shorter proof for the multilinear restriction estimate that bypasses the use of the multilinear Kakeya estimate was provided by the author in [1].

Removing the factor of  $R^\epsilon$  in these estimates seems to be a very difficult problem, even in the non-endpoint case. A major breakthrough was made by Guth in [8] where he proves the end-point case for the multilinear Kakeya estimate with no loss in  $R$ . In this paper Guth employs algebraic topology tools and initiates the use of the polynomial partitioning in the restriction theory that proved to be a very powerful tool, see [10–12, 17].

The result of Guth in [8] does not remove the factor of  $R^\epsilon$  for the multilinear restriction theory, not even in the non-endpoint case. However, in [4] Bennett uses Guth's result in [8] to improve the loss in (1.5) from  $R^\epsilon$  to  $(\log R)^\kappa$ , for some large  $\kappa$ .

Bourgain and Guth provide in [7] an  $\epsilon$ -removal type argument that establishes (1.3) for  $p > \frac{2}{k-1}$ , with no loss in  $R$ , in the case of hypersurfaces with non-degenerate curvature. Their proof is a modification of the  $\epsilon$ -removal argument of Tao in [15] in a linear context. In a very recent paper, see [14],

Tao removes the factor  $R^\epsilon$  in the non-endpoint case for the multilinear restriction estimate in a very general setup for the hypersurfaces involved - a certain amount of regularity is required, but nothing else. This is a very involved result that further refines the heat flow method from [5]. Tao's recent result essentially leaves only the optimal problem (1.3) with  $p = \frac{2}{k-1}$  as an open problem.

Our first result in this paper provides an  $\epsilon$ -removal argument for general hypersurfaces which have the property that the Fourier transform of the surface measure displays some decay. Precisely we assume that there exists some  $\alpha > 0$  such that the following decay property holds true

$$(1.6) \quad |\mathcal{E}\psi| \lesssim C(\psi)(1 + |x|)^{-\alpha}$$

for any smooth  $\psi$  supported within  $U$ .

**Theorem 1.1.** *Assume that the hypersurfaces  $S_i, i = 1, \dots, k$  are as above, satisfy the transversality condition 1.4 and the decay condition 1.6 for some  $\alpha > 0$ . If  $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow p, \epsilon)$  holds true for some  $p \geq \frac{2}{k}$ , then  $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow q)$  holds true for any  $q \geq p + (n + p + 1)\frac{C}{\log \frac{1}{\epsilon}}$ , where  $C$  is any constant satisfying  $C > \min(2, n - 1)$ .*

The first thing to notice in the above statement is that our threshold for the exponent  $p$  is  $\frac{2}{k}$  which is lower than the conjectured one of  $\frac{2}{k-1}$ . This improvement may seem artificial or vacuous at best, given that the conjectured optimal exponent is  $\frac{2}{k-1}$ . However  $\frac{2}{k-1}$  is the threshold for the generic case and it is conjectured that under appropriate curvature hypothesis on the hypersurfaces  $S_i$  the threshold in (1.3) can be lowered to exponents  $p < \frac{2}{k-1}$ ; in fact for some special class of hypersurfaces the author established the result up to, but excluding, the end-point  $p = \frac{2(n+k)}{k(n+k-2)}$ , see [2, 3].

Similar results have been obtained by Bourgain [6] and Tao and Vargas [16], both in the context of the bilinear restriction estimate. When taking into account the extended range of  $p$ , our result is new in the context of multilinear restriction estimate where the multilinearity is trilinear or higher in order; however it is not new when  $k = n$ , simply because the conjectured value of  $p$  is the same in the generic case as in the non-generic case.

Next we introduce the type of a hypersurface (an analytic condition that essentially guarantees that (1.6) holds true) and consider also the case when (1.6) holds only for some hypersurfaces, while the others are subsets of hyperplanes. Following [13] (see chapter 8, section 3.2), we define the  $l$ -type of a hypersurface  $S = \Sigma(U)$ . We fix  $x_0 \in U$  and assume that there

exists  $l \in \mathbb{N}$  such that for every unit vector  $\eta$  there exists  $\alpha$  with  $|\alpha| \leq l$  for which  $|\partial^\alpha(\Sigma \cdot \eta)(x_0)| \neq 0$ . The smallest such  $l$  is called the type of  $S$  at  $x_0$ . Assuming that  $U$  is a subset of compact set where  $\Sigma$  is defined, we define the type of  $S$  to be the maximum of the types of the  $x_0 \in U$ . In the real-analytic class, the condition that  $S$  has a finite type is equivalent to  $S$  not being a subset of any affine hyperplane. Outside this class, things can get more complicated.

As a direct consequence or by simple modifications of the argument in Theorem 1.1 we obtain the following result.

**Corollary 1.2.** *i) If all  $S_i, i = 1, \dots, k$ , are of finite type, then the result of Theorem 1.1 holds true.*

*ii) Assume that there exists  $0 \leq k_1 < k$  such that  $S_i$  are of  $l_i$ -type, where  $l_i < \infty, \forall i = 1, \dots, k_1$ , while  $S_i, i = k_1 + 1, \dots, k$  are subsets of hyperplanes. If  $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow p, \epsilon)$  holds true for some  $p \geq \frac{2}{k-1}$ , then  $\mathcal{R}^*(2 \times \dots \times 2 \rightarrow q)$  holds true for any  $q \geq p + (n + p + 1) \frac{C}{\log \frac{1}{\epsilon}}$ , where  $C$  is any constant satisfying  $C > \min(2, n - 1)$ .*

*iii) If all  $S_i, i = 1, \dots, k$  are real-analytic, then the result in ii) above holds true.*

The noticeable difference here is that we increased the threshold for  $p$  from  $\frac{2}{k}$  to  $\frac{2}{k-1}$ . This is not unnatural, given that if one of the hypersurfaces is a subset of a hyperplane, then the generic estimate in  $L^{\frac{2}{k-1}}$  is the best that can be expected.

When comparing the result above with the ones described earlier for the generic case, we see that our result is more general than the one of Bourgain-Guth [7] given that we do not impose any non-degeneracy conditions on the curvature. But it is not as strong as Tao's result in [14] given that we restrict ourselves to finite type or analytic setting, while Tao's result requires only a finite amount of differentiability. We note however that the arguments in this paper are significantly simpler than the ones employed by Tao in [14]; but we should also note that our arguments rely heavily on the work of Tao [15] where the  $\epsilon$ -removal argument is carried in the linear setup.

The second result of this paper is about the generic multilinear restriction estimate in the case when appropriate localization properties occur. In a nutshell, the localization properties are in small neighborhoods of submanifolds some of the  $S_i$ ; first we need to make this rigorous. The localization property can be stated in terms of a parametrization, which corresponds to

a localization property for  $f$  used in  $\mathcal{E}f$ , or in terms of the intrinsic formulation, which corresponds to a localization property for  $g$  used in  $\mathcal{E}g$ . We will state both and point out that they are in some sense equivalent.

By a submanifold  $S'_i$  of  $S_i$  we mean  $S'_i = \Sigma_i(M_i)$  where  $M_i \subset U_i \subset \mathbb{R}^{n-1}$  is a submanifold in  $\mathbb{R}^{n-1}$  that is given by  $\Sigma'_i : U'_i \rightarrow U_i$ , where  $U'_i$  is a bounded, open, connected neighborhood of the origin in  $\mathbb{R}^{m_i}$  and  $\Sigma'_i$  is smooth in the sense of (1.1);  $m_i$  is the dimension of the submanifold  $S'_i$ . For  $0 < \epsilon \ll 1$ , the  $\epsilon$ -neighborhood of  $M_i$  is defined by  $B_\epsilon(M_i) = \cup_{\xi \in M_i} B_\epsilon(\xi)$ . We say that  $f$  is supported in the  $\epsilon$ -neighborhood of  $M_i$  if it is supported in  $B_\epsilon(M_i)$ .

With  $S'_i = \Sigma_i(M_i)$ , we define the  $\epsilon$ -neighborhood of  $S'_i$  in  $\mathbb{R}^n$  in a similar manner  $B_\epsilon(S'_i) = \cup_{\xi \in S'_i} B_\epsilon(\xi)$ . We say that  $g : S_i \rightarrow \mathbb{C}$  is supported in the  $\epsilon$ -neighborhood of  $S'_i$  if it is supported in  $B_\epsilon(S'_i) \cap S_i$ .

In what follows by  $N_\zeta$  we denote the normal space to the corresponding submanifold within  $\mathbb{R}^n$ :  $N_\zeta S'_i$  is the normal space to  $S'_i \subset \mathbb{R}^n$  at  $\xi \in S'_i$ .

Recalling that the two formulations of the restriction operator are related via  $f(\xi) = g(\Sigma(\xi)) | \frac{\partial \Sigma}{\partial \xi_1} \wedge \dots \wedge \frac{\partial \Sigma}{\partial \xi_{n-1}} |$ , and using that  $\Sigma$  is a local diffeomorphism, it follows that a localization of  $f$  in the  $\epsilon_1$ -neighborhood of  $M_i$  gives a localization of  $g$  in the  $\epsilon_2$ -neighborhood of  $S'_i = \Sigma(M_i)$  and vice versa, where  $\epsilon_1 \approx \epsilon_2$ .

We introduce one more notation. If  $V_1, \dots, V_k$  are  $d_i$ -dimensional planes, then by  $|V_1 \wedge \dots \wedge V_k|$  we mean the quantity  $|v_{1,1} \wedge \dots \wedge v_{1,d_1} \wedge \dots \wedge v_{k,1} \wedge \dots \wedge v_{k,d_k}|$  where  $v_{i,1}, \dots, v_{i,d_i}$  is an orthonormal basis in  $V_i, i = 1, \dots, k$ ; it is easily seen that the defined quantity is independent of the choices of orthonormal systems.

The small support condition is the following:

**Condition 1.** Assume that we are given submanifolds  $S'_i \subset S_i, i = 1, \dots, k$ , of codimension  $c_i$ , with the property that there exists  $\nu > 0$  such that

$$(1.7) \quad |N_{\zeta_1} S'_1 \wedge \dots \wedge N_{\zeta_k} S'_k| \geq \nu$$

for all choices  $\zeta_i \in S'_i$ . Given  $g_i \in L^2(S_i), \forall i = 1, \dots, k$ , we assume that  $\text{supp} g_i \subset B_{\mu_i}(S'_i) \cap S_i$ , where  $0 < \mu_i \ll 1$ .

We make the observation that the codimension of  $S'_i$  is relative to  $S_i$ , thus the codimension of  $S'_i$  in  $\mathbb{R}^n$  is  $c_i + 1$ . In particular, the normal space  $N_{\zeta_i} S'_i$  has dimension  $c_i + 1$ . As a consequence, the total number of directions in which localization is provided cannot exceed  $n - k$ , that is  $c_1 + \dots + c_k \leq n - k$ ; this follows from (1.7) and the fact that the dimension of each  $N_{\zeta_i} S'_i$  is

$c_i + 1$ . We also note that if  $c_i = 0$ , then the localization property is vacuous and we simply work with  $S'_i = S_i$ ; at the same time  $N_{\zeta_i} S_i$  is just the normal to  $S_i$  at  $\zeta_i$ .

If the assumptions above are imposed in the generic multilinear estimate, we obtain the following result.

**Theorem 1.3.** *Assume  $S_i, i = 1, \dots, k$  are smooth. In addition, assume that  $g_1, \dots, g_k$  satisfy Condition 1. Then for any  $\epsilon > 0$ , there is  $C(\epsilon)$  such that for every  $R > 0$  the following holds true*

$$(1.8) \quad \|\Pi_{i=1}^k \tilde{\mathcal{E}}_i g_i\|_{L^{\frac{2}{k-1}}(B(0,R))} \leq C(\epsilon) \Pi_{i=1}^k \mu_i^{\frac{c_i}{2}} R^\epsilon \Pi_{i=1}^k \|g_i\|_{L^2(S_i)}.$$

A more restrictive version of the above result was established in [1]. One of the limitations there was that only one factor had localization properties. But the more important limitation is that the localization in [1] was around a flat submanifold (when using a specific projection-type parametrization). The more general statement in Theorem 1.3 will play a crucial role in a forthcoming paper of the author regarding the multilinear restriction estimate when curvature properties are taken into consideration; it is also likely that this new result will be useful in other applications.

The paper is organized as follows. In section 2 we introduce our most common notations, recall some basic results and perform some reductions which will be useful in the latter sections. In the following two sections we prove Theorems 1.1 and 1.3, respectively.

In proving Theorem 1.1 we adapt to the multilinear setup the arguments used by Tao in [15], where an  $\epsilon$ -removal argument has been carried out in the linear case. In addition, the analysis involves tracking the directions in which each wave  $\mathcal{E}_i f_i$  propagates and the components that are essential in the application of the near-optimal multilinear restriction estimate - this is very important in the final summation process.

In proving Theorem 1.3, we use an induction on scale argument, similar to the one from [1] used for proving a weaker version of this result. Dealing with the localization properties during the induction process requires a complete change in the approach (relative to [1]) that is more robust and general.

## 2. Notation, basic results and reductions

In this section we introduce some notation and some reductions of the setup described in the introduction with the aim of simplifying the level of technicalities involved in our arguments.

If  $v_1, \dots, v_m$  are vectors in  $\mathbb{R}^n$ , by  $\text{span}(v_1, \dots, v_m)$  we mean the subspace of  $\mathbb{R}^n$  spanned by  $v_1, \dots, v_m$ . We use the standard notation  $A \lesssim B$ , meaning  $A \leq CB$  for some universal  $C$  which is independent of variables used in this paper; in particular it will be independent of  $\delta$  and  $R$  that appear in the main proof. By  $A \lesssim_N B$  we mean  $A \leq C(N)B$  and indicate that  $C$  depends on  $N$ .

Our results involve estimates in  $L^p(S)$ ,  $S \subset \mathbb{R}^n$  with  $0 < p \leq 1$ . We recall the standard estimate for superpositions of functions in  $L^p$  for  $0 < p \leq 1$ :

$$(2.1) \quad \left\| \sum_{\alpha} f_{\alpha} \right\|_{L^p}^p \leq \sum_{\alpha} \|f_{\alpha}\|_{L^p}^p.$$

We make the convention that the hypersurfaces involved have very small diameter in the sense that if  $S = \Sigma(U)$  is a parametrization then  $U \subset \mathbb{R}^{n-1}$  has small diameter in the classical sense. This assumption does not affect our analysis: we can break each hypersurface in finitely many pieces of small diameter, run the arguments with the above setup and then sum up the result on these pieces.

Since the estimates are not affected by translating the hypersurfaces  $S_i$ , we can assume that they all pass through the origin, and moreover that  $0 \in S'_i$  for all  $i = 1, \dots, k$ . For each  $i = 1, \dots, k$ , we let  $N_i = N_0 S_i$  be the  $c_i$ -plane that is normal to  $S_i$  at 0.  $N_i$  is transversal to  $S'_i$  at 0 and because  $U'_i$  has small diameter,  $N_i$  is transversal to  $S'_i$  at every  $\zeta_i \in S'_i$ .

Next we proceed with reducing the problem to the case when the hypersurfaces or submanifolds involved have appropriate parameterizations; we do this for the more complicated setup of Theorem 1.3, and note that a simplified version covers the one needed in Theorem 1.1.

From (1.7), we know that

$$(2.2) \quad |N_1 \wedge \dots \wedge N_k| \geq \nu.$$

Our next goal is to prove that there exists a non-degenerate linear transformation  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  that allows us to assume that, under this transformation of the  $\xi$ -space, the system  $N_1, \dots, N_k$  is an orthogonal system in the sense that if  $v_i \in N_i, v_j \in N_j$  with  $i \neq j$ , then  $v_i \perp v_j$ .



$S'_i$  is transformed into  $AS'_i$ ,  $T_0S'_i$  is transformed into  $AT_0S'_i$ , but  $AN_i$  is not necessarily the normal to  $AS'_i$  at  $\xi = 0 \in S'_i$ ! Indeed, the normal condition is  $\langle n, v \rangle = 0, \forall v \in T_0S_i, n \in N_i$  and there is no guarantee that  $\langle An, Av \rangle = 0, \forall v \in T_0S_i, n \in N_i$  (unless  $A$  is orthonormal transformation which would not solve the problem we seek to fix). If we wanted  $e_{i,1}, \dots, e_{i,c_i+1}$  to be the normal vectors to  $AS_i$ , then we need to impose  $\langle e_{i,j}, Av \rangle = 0, \forall v \in T_0S_i$  which is equivalent to  $\langle A^*e_{i,j}, v \rangle = 0, \forall v \in T_0S_i$ , therefore if  $n_{i,j}$  was an orthonormal base for  $N_i$  it suffices to impose

$$A^*e_{i,j} = n_{i,j}, \forall i = 1, \dots, k, \forall j = 1, \dots, c_i + 1.$$

Given an orthonormal system  $\{e_{i,j}\}_{i=1, \dots, k; j=1, \dots, c_i+1}$ , we can solve for  $A$  satisfying these equations; from (2.2), it follows that we can choose  $A$  to obey good bounds, that is  $\|A\| + \|A^{-1}\| \lesssim \nu^{-1}$ .

The effect of  $A$  on the other properties in Condition 1 is fairly simple:  $S'_i$  becomes  $AS'_i$  which is a submanifold of  $AS_i$ , and the effect on the neighborhood can be quantified as follows:

$$AB_\epsilon(S'_i) \subset B_{c\epsilon}(AS'_i),$$

for some  $c$  depending on  $A$ .

The change of coordinates induced by  $A$  will be reflected in the norms  $\|g_i\|_{L^2(S_i)}$  and  $\|\prod_{i=1}^k \tilde{\mathcal{E}}_i g_i\|_{L^{\frac{2}{k-1}}(B(0,R))}$  via the usual determinants of the corresponding Jacobian transformation - these values are both bounded from below and above in terms of  $A$ . Thus we can skip the use of  $A$  and its related effects discussed above throughout the rest of the argument and simply work with the original objects with the additional assumption that the set of vectors  $N_1, \dots, N_k$  is an orthogonal base in  $R^n$ .

For the purpose of Theorem 1.1, a similar but simpler reduction is in place, the only difference being that there  $N_i$  is simply the 1-dimensional normal space to  $S_i$  at 0, so we could work directly with (unit) vectors as opposed to spaces.

Next, we arrange a bit the variables in our problem. Each  $N_i$  contains  $n_i = n_i(0)$ , the unit normal to  $S_i$  at 0. This particular vector plays an important role in the overall analysis, given that  $S_i$  are the ambient hypersurfaces. Next in each  $N_i, i = 1, \dots, k$  we pick an orthonormal basis that contains  $n_i$ , that is  $\{n_i\} \cup \{e_j^i\}_{j=1, \dots, c_i}$ . If we relabel  $\cup_{i=1}^k \{e_j^i\}_{j=1, \dots, c_i} = \{e_{k+1}, \dots, e_l\}$ , then we can further complete this system of vectors with unit vectors  $e_{l+1}, \dots, e_n$  with the property that the set of vectors  $n_1, \dots, n_k, e_{k+1}, \dots, e_n$  is an orthonormal system. Then let  $\xi = (\xi_1, \dots, \xi_n)$  where  $\xi_i$

is the coordinate in the direction of  $n_i$ ,  $i = 1, \dots, k$ , and  $\xi_i$  is the coordinate in the direction of  $e_i$  for  $i = k + 1, \dots, n$ ; we also use the notation  $\xi'_i = (\xi_1, \dots, \hat{\xi}_i, \dots, \xi_n)$  (skip the  $\xi_i$  coordinate). We let  $x = (x_1, \dots, x_n)$  be the dual system of coordinate and  $x'_i = (x_1, \dots, \hat{x}_i, \dots, x_n)$ .

For each  $i = 1, \dots, k$ , we let  $\mathcal{H}_i = \{x \in \mathbb{R}^n : x_i = 0\}$  be the hyperplane passing through the origin with normal  $n_i$ ; we also let  $\hat{\mathcal{H}}_i = \{\xi \in \mathbb{R}^n : \xi_i = 0\}$ . We denote by  $\pi_i : \mathbb{R}^n \rightarrow \mathcal{H}_i$  the projection along  $n_i$  onto  $\mathcal{H}_i$ ; similarly we denote by  $\hat{\pi}_i : \mathbb{R}^n \rightarrow \hat{\mathcal{H}}_i$  the projection along  $n_i$  onto  $\hat{\mathcal{H}}_i$ .

For technical purposes it is preferably to work with graph-like parametrizations of  $S_i = \Sigma_i(U_i)$ , that is with  $\Sigma_i : U_i \subset \hat{\mathcal{H}}_i \rightarrow \mathbb{R}^n$  being of the form  $\Sigma_i(x) = \{\xi \in \mathbb{R}^n : \xi_i = \varphi_i(\xi'_i)\}$ ,  $\forall \xi'_i \in U_i$  for some smooth injective map  $\varphi_i : U_i \rightarrow \mathbb{R}$ . Such parametrizations always exists under the assumption that the diameter of  $U_i$  (coming from the original parametrization) is sufficiently small (which is the case): one can simply project  $S_i$  onto  $\mathcal{H}_i \cong \mathbb{R}^{n-1}$  along  $n_i$  to obtain  $\Sigma_i^{-1} : S_i \rightarrow \hat{\mathcal{H}}_i$ .

We let  $f_i(\xi'_i) = g_i(\Sigma_i(\xi'_i)) \cdot \frac{d\sigma_{S_i}(\Sigma_i(\xi'_i))}{d\xi'_i}$ ; the relevance of  $f_i$  comes from the fairly straightforward identity

$$\tilde{\mathcal{E}}_i g_i = \mathcal{E}_i f_i.$$

The above parametrization and newly constructed function  $f_i$  have two additional properties. First,  $\Sigma_i^{-1}(S'_i)$  is a submanifold of  $\mathbb{R}^{n-1}$  whose co-dimension equals the co-dimension of  $S'_i$  relative to  $S_i$ . Second, if  $g_i$  has the properties in Condition 1, then  $f_i$  is supported in  $B_{c\epsilon}(\Sigma_i^{-1}(S'_i))$ . In what follows we set  $M_i = \Sigma_i^{-1}(S'_i) \subset \hat{\mathcal{H}}_i$  - this is the manifold whose specific neighborhood contains the support of  $f_i$ .

At one point in the argument we will be using more convenient versions of the  $\epsilon$ -neighborhood of a submanifold. As before, we can assume that  $U'_i$  is of small diameter; then  $S'_i$  has a graph type parametrization in the sense that, after an orthonormal change of coordinates, we have  $\Sigma'_i(\xi') = (\xi', \varphi_{i,m_i+1}(\xi'), \dots, \varphi_{i,n-1}(\xi'))$ ,  $\xi' = (\xi_1, \dots, \xi_{m_i}) \in U'_i$ . Thus a standard neighborhood can be traded in this context for the simpler version  $\tilde{B}_\epsilon(M_i) = (\xi', \varphi_{i,m_i+1}(\xi') + t_{m_i+1}, \dots, \varphi_{i,n-1}(\xi') + t_{n-1})$ ,  $\xi' \in U'_i + B_\epsilon(0)$ ,  $|t_j| < \epsilon, \forall j = m_i + 1, \dots, n - 1$ ; here we are implicitly assuming the fact  $\Sigma'_i$  extends in a neighborhood of size  $\epsilon$  of  $U'_i$ .

The above setup reduction works the same for Theorem 1.1 by simply ignoring the consideration involving the manifold  $S'_i$  and its neighborhood.

With the above notations we have

$$(2.3) \quad \mathcal{E}_i f_i(x) = \int_{U_i} e^{i(x'_i \xi'_i + x_i \varphi_i(\xi'_i))} f_i(\xi'_i) d\xi'_i.$$

The Fourier transform of a Schwartz function  $f : \mathbb{R}^n \rightarrow \mathbb{C}$  is defined by

$$\mathcal{F}(f)(\xi) = \hat{f}(\xi) = \int e^{-ix \cdot \xi} f(x) dx.$$

The inverse Fourier transform is defined by

$$\mathcal{F}^{-1}(g)(x) = \check{g}(x) = \frac{1}{(2\pi)^n} \int e^{ix \cdot \xi} g(\xi) d\xi.$$

The standard Fourier inversion formula is in place  $\mathcal{F}^{-1}\mathcal{F}f = f$  and  $\mathcal{F}\mathcal{F}^{-1}g = g$ . These definitions are then extended to distributions, in particular to  $L^p(\mathbb{R}^n)$  spaces, in the usual manner.

In a similar manner (just that we work in  $n - 1$  dimensions), we denote by  $\mathcal{F}_j : \mathcal{H}_j \rightarrow \hat{\mathcal{H}}_j$  the Fourier transform,  $x'_j \rightarrow \xi'_j$ , and by  $\mathcal{F}_j^{-1}$  the inverse Fourier transform,  $\xi'_j \rightarrow x'_j$ . Obviously,  $\mathcal{F}_j, \mathcal{F}_j^{-1}$  act on the variables  $x'_j, \xi'_j$  respectively.

With these notations in place we make a few more observations. We denote by  $\mathcal{E}_j f_j(x'_j, 0) = \mathcal{E}_j f_j|_{x_j=0}$  and note that

$$(2.4) \quad \mathcal{E}_j f_j(x'_j, 0) = (2\pi)^{n-1} \mathcal{F}_j^{-1} f_j.$$

With the parametrizations introduced earlier for each  $S_i$ , it follows that

$$(2.5) \quad \mathcal{E}_j f_j = e^{ix_j \varphi_j(D'_j)} \check{f}_j,$$

where the symbol of  $e^{ix_j \varphi_j(D'_j)}$  is  $e^{ix_j \varphi_j(\xi'_j)}$ .

Another observation is that the operator  $\mathcal{E}_j f_j(\cdot, x_j) : L^2(\mathcal{H}_j) \rightarrow L^2(\mathcal{H}_j)$  is an  $L^2$  isometry with respect to  $x_j$ , that is

$$(2.6) \quad \|\mathcal{E}_j f_j(\cdot, x_j)\|_{L^2} = \|\mathcal{E}_j f_j(\cdot, 0)\|_{L^2} = (2\pi)^{\frac{n-1}{2}} \|f_j\|_{L^2}, \quad \forall x_j \in \mathbb{R}.$$

For the operators  $\mathcal{E}_j$  we highlight a commutator estimate which will be used in our proof. We define the operator  $\nabla \varphi_j(\frac{D'_j}{i})$  to be multiplier with symbol  $\nabla \varphi_j(\xi')$ . For any  $x \in \mathbb{R}^n$  and  $c \in \mathcal{H}_j$ , it holds true that

$$(2.7) \quad (x'_j - c - x_j \nabla \varphi_j(\frac{D'_j}{i}))^N \mathcal{E}_j f(x) = \mathcal{E}_j(\mathcal{F}_j((x'_j - c)^N \mathcal{F}_j^{-1} f))(x), \quad \forall N \in \mathbb{N}.$$

Next we prepare some geometric elements that are needed in the proof. We let  $\mathcal{L} := \{z_1 n_1 + \dots + z_k n_k + z_{k+1} e_{k+1} + \dots + z_n e_n : (z_1, \dots, z_n) \in \mathbb{Z}^n\}$  be the standard lattice in  $\mathbb{R}^n$  generated by the vectors  $n_1, \dots, n_k, e_{k+1}, \dots, e_n$ .

In each  $\mathcal{H}_i, i = 1, \dots, k$  we construct the induced lattice  $\mathcal{L}(\mathcal{H}_i) = \pi_i(\mathcal{L})$ . Recall that for each  $i = 1, \dots, k$ , it holds true that  $\{n_1, \dots, n_k, e_{k+1}, \dots, e_l\} \setminus \{n_i\} \subset \mathcal{H}_i$ . We split  $\mathcal{H}_i = \mathcal{H}'_i \oplus \mathcal{H}''_i$ , where  $\mathcal{H}''_i$  is spanned by the set of vectors  $\{e_1^i, \dots, e_{c_i}^i\}$ , while  $\mathcal{H}'_i$  is spanned by

$$\{n_1, \dots, n_k, e_{k+1}, \dots, e_l\} \setminus \{N_i, e_1^i, \dots, e_{c_i}^i\}.$$

Correspondingly, this produces a split of  $\mathcal{L}(\mathcal{H}_i)$  as follows  $\mathcal{L}(\mathcal{H}_i) = \mathcal{L}(\mathcal{H}'_i) \oplus \mathcal{L}(\mathcal{H}''_i)$ .

Given  $r > 0$  we define  $\mathcal{C}(r)$  be the set of of cubes of size  $r$  in  $\mathbb{R}^n$  that are centered at points in the lattice  $r\mathcal{L}$ ; a cube in  $\mathcal{C}(r)$  has the following form

$$q(\mathbf{j}) := [r(j_1 - \frac{1}{2}), r(j_1 + \frac{1}{2})] \times \dots \times [r(j_n - \frac{1}{2}), r(j_n + \frac{1}{2})]$$

where  $\mathbf{j} = (j_1, \dots, j_n) \in \mathbb{Z}^n$ . For such a cube we define

$$c(q) = r\mathbf{j} = (rj_1, \dots, rj_n) \in r\mathcal{L}$$

to be its center; vice-versa, if  $c \in r\mathcal{L}$  then we define  $q(c)$  to be the cube whose center  $c(q) = c$ . Then, for each  $i = 1, \dots, n$ , we let  $\mathcal{CH}_i(r) = \pi_i\mathcal{C}(r)$  be the set of cubes of size  $r$  in the hyperplane  $\mathcal{H}_i$ . Finally, given two cubes  $q, q' \in \mathcal{C}(r)$  (or  $\mathcal{CH}_i(r)$  or  $\mathcal{CH}'_i(r), \mathcal{CH}''_i(r)$ ) we define  $d(q, q')$  to be the distance between them when considered as subsets of the underlying space, be it  $\mathbb{R}^n$  or  $\mathcal{H}_i$  or  $\mathcal{H}'_i/\mathcal{H}''_i$ .

Let  $\chi_0^n : \mathbb{R}^n \rightarrow [0, +\infty)$  be a Schwartz function, normalized in  $L^1$ , that is  $\|\chi_0^n\|_{L^1} = 1$ , and with Fourier transform supported on the unit ball. For each  $q \in \mathcal{C}(r)$ , define  $\chi_q : \mathbb{R}^n \rightarrow \mathbb{R}$  by

$$\chi_q(x) = \chi_0^n\left(\frac{x - c(q)}{r}\right)$$

Notice that  $\hat{\chi}_q$  is supported in  $B(0, r^{-1})$ . By the Poisson summation formula and the properties of  $\chi_0^n$ ,

$$(2.8) \quad \sum_{q \in \mathcal{C}(r)} \chi_q = 1.$$

Using the properties of  $\chi_q$ , a direct exercise shows that for each  $N \in \mathbb{N}$ , the following holds true

$$(2.9) \quad \sum_{q \in \mathcal{C}(r)} \left\| \left\langle \frac{x - c(q)}{r} \right\rangle^N \chi_q g \right\|_{L^2}^2 \lesssim_N \|g\|_{L^2}^2$$

for any  $g \in L^2(\mathbb{R}^n)$ .

We introduce similar entities in each  $\mathcal{H}_i, i = 1, \dots, k$ . For each  $q \in \mathcal{CH}_i(r)$ , define  $\chi_q : \mathcal{H}_i \rightarrow \mathbb{R}$  by  $\chi_q(x) = \chi_0^{n-1}(\frac{x-c(q)}{r})$  and note that  $\mathcal{F}_i\chi_q$  is supported in  $B(0, r^{-1})$ . By the Poisson summation formula and the properties of  $\chi_0^{n-1}$ ,

$$(2.10) \quad \sum_{q \in \mathcal{CH}_i(r)} \chi_q = 1.$$

In a similar way to (2.9), the following holds true

$$(2.11) \quad \sum_{q \in \mathcal{CH}_i(r)} \|\langle \frac{x - c(q)}{r} \rangle^N \chi_q g\|_{L^2}^2 \lesssim_N \|g\|_{L^2}^2$$

for any  $g \in L^2(\mathcal{H}_i)$ . Here, the variable  $x$  is the argument of  $g$  and belongs to  $\mathcal{H}_i$ .

We recall from [1] the following discrete version of the continuous Loomis-Whitney inequality:

$$(2.12) \quad \|\Pi_{i=1}^k g_i(\pi_i(z))\|_{l^{\frac{2}{k-1}}(\mathcal{L})} \lesssim \Pi_{i=1}^k \|g_i\|_{l^2(\mathcal{L}(\mathcal{H}_i))}.$$

We need a further refinement of this estimate.

For a function  $g : \mathcal{L}(\mathcal{H}_i) \rightarrow \mathbb{C}$ , we define the space  $l^2 l^\infty(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i))$  to be the space of functions whose norm

$$\|g\|_{l^2 l^\infty(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i))} = \left( \sum_{z' \in \mathcal{L}(\mathcal{H}'_i)} \sup_{z'' \in \mathcal{L}(\mathcal{H}''_i)} |g(z', z'')|^2 \right)^{\frac{1}{2}},$$

is finite. With this notation in place we have the following result:

**Lemma 2.1.** *Assume  $g_i \in l^2 l^\infty(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i)), i = 1, \dots, k$ . Then the following holds true*

$$(2.13) \quad \|\Pi_{i=1}^k g_i(\pi_i(z))\|_{l^{\frac{2}{k-1}}(\mathcal{L})} \lesssim \Pi_{i=1}^k \|g_i\|_{l^2 l^\infty(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i))}.$$

*Proof.* For  $z \in \mathcal{L}$  we write  $z = (z', z'', z''')$  where  $z' = (z_1, \dots, z_k)$  collects the coordinates in the directions of  $n_1, \dots, n_k$ ,  $z''$  collects the coordinates in the directions of  $e_{k+1}, \dots, e_l$  and  $z'''$  collects the coordinates in the directions of  $e_{l+1}, \dots, e_n$ . In this proof, we need to further refine the lattices  $\mathcal{H}_i$ . We let  $\tilde{\mathcal{H}}''_i$  be the projection of  $\mathcal{H}_i$  onto the subspace generated by the vectors  $\{e_{k+1}, \dots, e_l\} \setminus \{e_1^i, \dots, e_{c_i}^i\}$ .

We fix  $z'$  and  $z'''$ , let  $\{z'\} \times \mathcal{L}'' \times \{z'''\}$  be the sub-lattice of  $\mathcal{L}$  obtained by fixing  $z'$  and  $z'''$ . Then we use Hölder's inequality to obtain

$$\begin{aligned} & \|\Pi_{i=1}^k g_i(\pi_i(z', \cdot, z'''))\|_{l^{\frac{2}{k-1}}(\{z'\} \times \mathcal{L}'' \times \{z'''\})} \\ & \lesssim \Pi_{i=1}^k \|g_i(\pi_i(z', \cdot, z'''))\|_{l^2 l^\infty(\mathcal{L}(\tilde{\mathcal{H}}_i'') \times \mathcal{L}(\mathcal{H}_i''))}. \end{aligned}$$

In justifying this estimate, we have also used the fact that  $l^2 l^\infty(\mathcal{L}(\tilde{\mathcal{H}}_i'') \times \mathcal{L}(\mathcal{H}_i''))$  is the strongest norm that combines an  $l^2$  norm in the variables from  $\mathcal{L}(\tilde{\mathcal{H}}_i'')$  and an  $l^\infty$  norm in the variables from  $\mathcal{L}(\mathcal{H}_i'')$ ; for instance it controls  $l^\infty l^2(\mathcal{L}(\mathcal{H}_i'') \times \mathcal{L}(\tilde{\mathcal{H}}_i''))$ .

Then we apply (2.12) with respect to the variable  $z'$  to obtain

$$\begin{aligned} & \|\Pi_{i=1}^k g_i(\pi_i(\cdot, \cdot, z'''))\|_{l^{\frac{2}{k-1}}(\mathcal{L}' \times \mathcal{L}'' \times \{z'''\})} \\ & \lesssim \Pi_{i=1}^k \|g_i(\pi_i(z', \cdot, z'''))\|_{l^2 l^\infty(\mathcal{L}(\tilde{\mathcal{H}}_i'') \times \mathcal{L}(\mathcal{H}_i''))}. \end{aligned}$$

Applying Hölder with respect to the  $z'''$  variable gives

$$\|\Pi_{i=1}^k g_i \circ \pi_i\|_{l^{\frac{2}{k}}_{z'''} l^{\frac{2}{k-1}}_{z', z''}} \lesssim \Pi_{i=1}^k \|g_i \circ \pi_i\|_{l^2 l^\infty(\mathcal{L}(\mathcal{H}_i'') \times \mathcal{L}(\mathcal{H}_i''))}.$$

To conclude with (2.13) we only need to use the simple inequality  $\|a\|_{l^{\frac{2}{k-1}}} \lesssim \|a\|_{l^{\frac{2}{k}}}$  for any sequence  $a \in \mathbb{Z}^m$  and in any dimension  $m$ .  $\square$

### 3. The $\epsilon$ -removal result

In this Section we provide a proof of Theorem 1.1, a theorem that provides an  $\epsilon$ -removal technique for multilinear estimates. The closest type of result in this direction is the work of Tao and Vargas in [16] in the context of bilinear restriction estimates. Unfortunately we cannot adapt that approach to our context, since, among other issues, it relies on duality type arguments; in our particular context we work with  $L^p$  quasi-norms with  $p < 1$  and use of duality is not an option. Instead we follow a previous  $\epsilon$ -removal argument of Tao in [15] in a linear context and use techniques developed by the author in [1] to adapt it in the context of multilinear estimates.

Our first result is of a technical nature and its use will become clear once we proceed with the proof of Theorem 1.1. The basic idea is that we need to use the near-optimal result

$$\|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(B_R)} \lesssim R^\epsilon \Pi_{i=1}^k \|f_i\|_{L^2}$$

in a more optimal form that essentially collects the mass of each wave  $\mathcal{E}_i f_i$  from  $B_R(x_i)$ . The starting point is that the mass of  $f_i$  can be recovered in multiple ways based on (2.6); as a consequence we can rewrite the above as

$$\|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(B_R)} \lesssim R^\epsilon \Pi_{i=1}^k \|\mathcal{E}_i f_i|_{x_i=c_{Q,i}}\|_{L^2(\mathcal{H}_i)}.$$

On the other hand, information coming from cubes  $q \times \{x_i = c_{Q,i}\}, q \in \mathcal{CH}_i(R)$  that are at distance  $\gg R$  from  $\pi_i Q \times \{x_i = c_{Q,i}\}, q \in \mathcal{CH}_i(R)$  should not impact much the estimate, due to the finite speed of propagation. The following result encodes this heuristics.

**Lemma 3.1.** *Let  $Q \in \mathcal{C}(R)$ . Then the following holds true*

$$(3.1) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(Q)} \lesssim R^\epsilon \Pi_{i=1}^k \left( \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i Q, q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{E}_i f_i|_{x_i=c_{Q,i}} \right\|_{L^2(q)}^2 \right)^{\frac{1}{2}}.$$

There is one technicality that we skip in the above statement in order to keep it simple. The occurrence of  $\chi_q$  in various places has the effect of modifying the support of entities involving  $f_i$  by a factor of  $R^{-1}$ . This would require us to assume that the original near-optimal result holds for each  $f_i$  supported in  $U_i + B(0, R^{-1})$  in order to claim the above for each  $f_i$  supported in  $U_i$ . The practical effect in the final argument here is that in order to obtain the  $\epsilon$ -removal result under the hypothesis that each  $f_i$  is supported in  $U_i$ , we need to know the near-optimal result for each  $f_i$  supported in  $U_i + B(0, c)$  for some small  $0 < c \ll 1$ .

*Proof.* We just sketch the main steps in this argument, as the argument is very similar to the one used for proving the more difficult (4.3); in the present context the  $\delta$  that appears in (4.3) should be essentially  $\approx 1$ .

The first observation is that, without restricting the generality of the argument we can assume that  $Q$  is centered at the origin, that is  $c_Q = 0$ . The above then becomes

$$(3.2) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(Q)} \lesssim R^\epsilon \Pi_{i=1}^k \left( \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{c(q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{E}_i f_i|_{x_i=0} \right\|_{L^2(q)}^2 \right)^{\frac{1}{2}}.$$

We establish the improvement one term at the time, starting with  $\mathcal{E}_1 f_1$ . Since  $\varphi_1(0) = 0$  and  $\nabla\varphi_1(0) = 0$ , we work with the basic multiplier  $x'_1$  and (2.7) to justify the following identity

$$\begin{aligned} x'_1 \mathcal{E}_1 \mathcal{F}_1(\chi_q \mathcal{F}_1^{-1} f_1) &= (x'_1 - x_1 \nabla\varphi_1(\frac{D'}{i})) \mathcal{E}_1 \mathcal{F}_1(\chi_q \mathcal{F}_1^{-1} f_1) \\ &\quad + x_1 \nabla\varphi_1(\frac{D'}{i}) \mathcal{E}_1 \mathcal{F}_1(\chi_q \mathcal{F}_1^{-1} f_1) \\ &= \mathcal{E}_1(\mathcal{F}_1(x'_1 \chi_q \mathcal{F}_1^{-1} f_1))(x) \\ &\quad + x_1 \nabla\varphi_1(\frac{D'}{i}) \mathcal{E}_1(\mathcal{F}_1(\chi_q \mathcal{F}_1^{-1} f_1)). \end{aligned}$$

Arguing in a similar manner to the way we do in the prof of (4.3) we derive (3.2); the details are left as an exercise. One important observation is that (2.4) provides the relation between  $\mathcal{E}_1 f_1|_{x_1=0}$  and  $\mathcal{F}_1 f_1$  that allows us to switch from using one of them to the other one.  $\square$

We now proceed with the proof of Theorem 1.1. Without restricting the generality of the argument, we can assume that  $\|f_i\|_{L^2} = 1, i = 1, \dots, k$ . We do this by providing a weak type estimate, that is if

$$E(\lambda) = \{x \in \mathbb{R}^n : |\Pi_{i=1}^k \mathcal{E}_i f_i| \geq \lambda\},$$

it suffices to prove  $|E(\lambda)| \lesssim \lambda^{-q}$  for some  $q$  satisfying  $q < p + (n+p+1) \frac{D}{\log \frac{1}{c}}$ .

The first observation is that if  $x \in E(\lambda)$ , then for some universal  $c \approx 1$ ,  $B_{c\lambda}(x) \subset E(\frac{\lambda}{2})$ . This follows from the simple computation

$$\begin{aligned} |\Pi_{i=1}^k \mathcal{E}_i f_i(x) - \Pi_{i=1}^k \mathcal{E}_i f_i(x_0)| &\leq k C_1^{k-1} \max_i |\mathcal{E}_i f_i(x) - \mathcal{E}_i f_i(x_0)| \\ &\leq k C_1^{k-1} C_2 |x - x_0| \end{aligned}$$

where

$$C_1 = \max_i \|\mathcal{E}_i f_i\|_{L^\infty} \lesssim 1, \quad C_2 = \max_i \|\nabla \mathcal{E}_i f_i\|_{L^\infty} \lesssim 1.$$

Thus if we choose  $c = \min(1, \frac{1}{2}(k C_1^{k-1} C_2)^{-1}) \approx 1$ , it follows that

$$|\Pi_{i=1}^k \mathcal{E}_i f_i(x) - \Pi_{i=1}^k \mathcal{E}_i f_i(x_0)| \leq \frac{\lambda}{2}$$

and the conclusion follows.

As a consequence, if we let  $F = \cup_{x \in E(\lambda)} B_{c\lambda}(x)$  we obtain that  $F \subset E(\frac{\lambda}{2})$ ; we let  $G = \cup_{x \in E(\lambda)} B_1(x)$  and note that  $|G| \lesssim \lambda^{-n} |F|$ .



Next we use the following result due to Tao [15]. We say that a collection  $\{B_R(x_i)\}_{i=1}^N$  is sparse if the centers  $x_i$  are  $R^C N^C$  separated.

**Lemma 3.2 (Lemma 3.3, [15]).** *Suppose  $E$  is the union of  $c$ -cubes and  $N \geq 1$ . Then there exists  $O(N|E|^{\frac{1}{N}})$  sparse collections of balls that cover  $G$ , such that the balls in each collection have radius  $1 \lesssim R \lesssim |E|^{C^N}$ .*

It is easy to see that the above statement can be made in terms of cubes, rather than balls, and this is how we use it below. Thus the above lemma gives us  $O(N|G|^{\frac{1}{N}})$  sparse collections  $O_l$  of cubes that cover the set  $G$  (and hence the set  $F$ ), such that the cubes in each collection  $O_l$  have radius/side-length  $1 \lesssim R_l \lesssim |E|^{C^N}$ . We fix a sparse collection of cubes  $O_l$ ; for each cube  $Q \in O_l$  we apply the near-optimal result in its refined version (3.1):

$$\begin{aligned} & \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(Q)} \\ & \lesssim R_l^\epsilon \Pi_{i=1}^k \left( \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i Q, q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{E}_i f_i|_{x_i=c_{Q,i}} \right\|_{L^2(q)}^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Based on Lemma 3.3 part ii) below, we can conclude that each of the terms on the right hand-side has the  $l^2_{Q \in O_l}$  norm bounded by  $\|f_i\|_{L^2}$ ; thus their product has  $l^{\frac{2}{k}}_{Q \in O_l}$  norm bounded by  $\Pi_{i=1}^k \|f_i\|_{L^2}$  and since  $p \geq \frac{2}{k}$  it follows that

$$\|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(O_l)} \lesssim R_l^\epsilon \Pi_{i=1}^k \|f_i\|_{L^2} = R_l^\epsilon.$$

Since  $F \subset E(\frac{\lambda}{2})$ , it follows that

$$\begin{aligned} |F| & \lesssim \lambda^{-p} \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(F)}^p \lesssim \lambda^{-p} \sum_l \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(O_l)}^p \\ & \lesssim \lambda^{-p} \sum_l R_l^{p\epsilon} \lesssim \lambda^{-p} N |G|^{\frac{1}{N}} (|G|^{C^N})^{p\epsilon} \\ & = \lambda^{-p} N |G|^{\frac{1}{N} + p\epsilon C^N} \lesssim \lambda^{-p} N (\lambda^{-n} |F|)^{\frac{1}{N} + p\epsilon C^N}. \end{aligned}$$

We let  $\beta = \frac{1}{N} + p\epsilon C^N$  and derive the following estimate

$$|F| \lesssim N \lambda^{-\frac{p+n\beta}{1-\beta}}.$$

For  $\epsilon \ll e^{-C} \leq 1$ , we make the choice  $N = C^{-1} \log \frac{1}{\epsilon}$ ; then it follows  $\frac{C}{\log \frac{1}{\epsilon}} \leq \beta \leq \frac{2C}{\log \frac{1}{\epsilon}} \ll 1$  and the following holds true  $\frac{p+n\beta}{1-\beta} < p + (n+p+1)\beta \leq p +$

$(n + p + 1) \frac{2C}{\log \frac{1}{\epsilon}}$ . Thus our argument for Theorem 1.1 is complete provided that we establish the following result.

**Lemma 3.3.** *i) Assume that  $S = \Sigma(U)$  is a smooth hypersurface that is parametrized by  $\xi_n = \varphi(\xi'_n)$ . In addition, we assume that there exists  $\alpha > 0$  such that*

$$(3.3) \quad |\mathcal{E}\psi| \lesssim C(\psi)(1 + |x|)^{-\alpha}$$

for any smooth  $\psi$  supported within  $U$ .

Given a  $C$ -sparse family of cubes  $\{Q_j\}_{j=1}^N$ , for some  $C > \frac{\min(2, n-1)}{\alpha}$ , the following holds true

$$(3.4) \quad \sum_{j=1}^N \|\mathcal{E}f|_{x_n=c_n(Q_j)}\|_{L^2(\pi_n(Q_j))}^2 \lesssim \|f\|_{L^2}^2,$$

for any  $f$  supported within  $U$ .

ii) Under similar hypothesis as in i), the following holds true:

$$(3.5) \quad \sum_{j=1}^N \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i Q_j, q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{E}_i f_i|_{x_i=c_{Q,i}} \right\|_{L^2(q)}^2 \lesssim \|f_i\|_{L^2}^2.$$

*Proof.* The estimate above can be restated as follows. If  $g = \check{f}$ , then from (2.5), it follows that (3.4) is equivalent to

$$(3.6) \quad \sum_{j=1}^N \|\chi_{\pi_n Q_j} e^{ic_n(Q_j)\varphi(D)} g\|_{L^2(\pi_n(Q_j))}^2 \lesssim \|g\|_{L^2}^2,$$

for any  $g$  that is supported within  $U$ . We let  $T : L^2(\mathbb{R}^n) \rightarrow l^2 L^2(\mathbb{R}^n)$  be defined by

$$Tf = (\chi_{\pi_n Q_j} e^{ic_n(Q_j)\varphi(D)} g)_{j=1}^N.$$

Then (3.4) is equivalent to proving that  $\|T\| \lesssim 1$ , where by the operator norm we mean the norm of  $T$  acting from  $L^2(\mathbb{R}^n) \rightarrow l^2 L^2(\mathbb{R}^n)$ . Its adjoint operator  $T^* : l^2 L^2(\mathbb{R}^n) \rightarrow L^2(\mathbb{R}^n)$  is given by

$$T^*F = \sum_{j=1}^N e^{-ic_n(Q_j)\varphi(D)} (\chi_{\pi_n Q_j} F_j).$$

Thus, it suffices to establish that the operator  $TT^* : l^2L^2(\mathbb{R}^n) \rightarrow l^2L^2(\mathbb{R}^n)$  has a good bound on its norm, that is  $\|TT^*\| \lesssim 1$   $TT^*$  is given by

$$TT^*F = (\chi_{\pi_n Q_k} e^{-ic_n(Q_k)\varphi(D)} \sum_{j=1}^N e^{ic_n(Q_j)\varphi(D)} \chi_{\pi_n Q_j} F_j)_{k=1}^N.$$

In the above we make the additional observation that each  $F_j$  has its Fourier transform  $\hat{F}_j$  is supported in a small neighborhood (of size  $O(R^{-1})$ ) of  $U$ ; this is a consequence of the original localization property of the function  $g$ , namely that  $\hat{g}$  was supported in  $U$ . Formally we could have rigorously kept track of this by inserting an additional operator  $\chi_U(D)$  (where  $\chi_U$  is the characteristic function of the set  $U$  in the definition of  $T$ , but this would have complicated the exposition without bringing any useful insight into the argument.

We claim the following estimate:

$$(3.7) \quad \begin{aligned} & \|\chi_{\pi_n Q_k} e^{-i(c_n(Q_k) - c_n(Q_j))\varphi(D)} \chi_{\pi_n Q_j} F_j\|_{L^2} \\ & \lesssim \min(1, R^{n-1} \langle c(Q_k) - c(Q_j) \rangle^{-\alpha}) \|F_j\|_{L^2}. \end{aligned}$$

Using the sparseness of the set of cubes with  $C > \frac{\min(2, n-1)}{\alpha}$ , it follows that  $TT^*$  is bounded and we can conclude the proof of (3.4).

We now prove (3.7). Using the full translation invariance of the estimate, (3.7) is equivalent to

$$(3.8) \quad \|\chi_{Q(c'_n)} e^{-ic_n\varphi(D)} \chi_{Q(0)} F_j\|_{L^2} \lesssim \min(1, R^{n-1} \langle c \rangle^{-\alpha}) \|F_j\|_{L^2}.$$

To get the bound with constant 1, we simply use the fact that all operators above are bounded on  $L^2$ . To get the second bound we start with the explicit formula

$$e^{-ic_n\varphi(D)} \chi_{Q(0)} F_j = K(\cdot, c_n) * (\chi_{Q(0)} F_j),$$

with

$$K(x'_n, x_n) = \int e^{i(x'_n \xi'_n + x_n \varphi(\xi'_n))} \chi(\xi'_n) d\xi'_n$$

where  $\chi(\xi'_n)$  keeps track of the localization property of each  $F_j$ . From (3.3) it follows that  $K$  satisfies the decay properties:

$$|K(x)| \lesssim (1 + |x|)^{-\alpha}.$$

Therefore

$$\|K\|_{L^1(Q_R(c'_n))} \lesssim R^{n-1} (1 + \|c\|)^{-\alpha}.$$

From this we obtain the (3.8) with constant  $R^{n-1}(1 + \|c\|)^{-\alpha}$ . □

We end this section with the proof of Corollary 1.2. From [13] (see chapter 8, section 3.2, Theorem 2) it follows that if  $S$  has  $l$ -type, then  $|\mathcal{E}\psi(x)| \lesssim (1 + |x|)^{-\frac{1}{l}}$ . Then i) follows from Theorem 1.1.

For ii), the basic idea is that for the hypersurfaces  $S_1, \dots, S_{k_1}$  we can still use parts of the analysis above, while for the  $S_{k_1+1}, \dots, S_k$  planar hypersurfaces we will make use of the known trivial estimates.

The argument follows the same steps as the one used in the proof of Theorem 1.1. When we arrive at the use of the near-optimal result in its refined version (3.1):

$$\begin{aligned} & \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^p(Q)} \\ & \lesssim R_i^\epsilon \Pi_{i=1}^k \left( \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i Q, q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{E}_i f_i|_{x_i=c_{Q,i}} \right\|_{L^2(q)}^2 \right)^{\frac{1}{2}}, \end{aligned}$$

we cannot make use of Lemma 3.3 for all hypersurfaces involved, thus we cannot recover the estimate in  $l_{Q \in O_i}^{\frac{2}{k}}$  bounded by  $\Pi_{i=1}^k \|f_i\|_{L^2}$ . What we can do instead, is to recover an estimate in  $l_{Q \in O_i}^{\frac{2}{k_1}}$  for the product of the first  $k_1$  terms on the right-hand side with a bound of  $\lesssim \Pi_{i=1}^{k_1} \|f_i\|_{L^2}$ . For the other terms we rely on an exact calculus. As we described in the section 2, we have chosen the coordinates such that  $n_{k_1+1}, \dots, n_k$ , the normals at the hyperplanes where  $S_{k_1+1}, \dots, S_k$  respectively lie, belong to the set of coordinate directions. As a consequence,  $\mathcal{E}_i f_i|_{x_i=c_{Q,i}} = \mathcal{E}_i f_i|_{x_i=0}$ . Thus the remaining term to be estimated is

$$\Pi_{i=k_1+1}^k \left( \sum_{q \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i Q, q)}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q)}{R} \right\rangle^N \chi_q \mathcal{F}_i^{-1} f_i|_{x_i=0} \right\|_{L^2(q)}^2 \right)^{\frac{1}{2}}.$$

It is an easy exercise to see that we obtain an  $l^{\frac{2}{k-k_1-1}}$  estimate for the above term with a bound of  $\lesssim \Pi_{i=k_1+1}^k \|f_i\|_{L^2}$  - morally this is at the level of Loomis-Whitney inequality. Alternatively, each term above has  $l^2 l^\infty(R\mathcal{L}(\mathcal{H}_i) \times R\text{span}(N_i))$  with respect to  $Q \in \mathcal{C}(R)$  and with a bound of  $\lesssim \|f_i\|_{L^2}$ . This leads to the estimate  $l^{\frac{2}{k-k_1-1}}$  as above using similar but simpler arguments to the ones in Lemma 2.1.

Finally iii) follows from ii).

### 4. Proof of Theorem 1.3

The proof uses an induction of scales argument. When moving from smaller scales to the larger ones, it is very useful to have tight control of the leaks of mass between various spatial regions for each individual wave  $\mathcal{E}_i f_i$ . This is achieved by splitting each domain  $U_i$  into smaller pieces of diameter  $\leq \delta$ , for some  $0 < \delta \ll 1$ . This, in turn, splits the surfaces  $S_i = \Sigma_i(U_i)$  in the corresponding pieces. It suffices to prove the multilinear estimate for each  $S_i$  being replaced by one of its pieces, since then we can sum up the estimates for all possible combinations of pieces using (2.1) and generate the original estimate at a cost of picking a factor that is a power of  $\delta^{-1}$ . In the end of the argument,  $\delta$  will be chosen in terms of absolute constants and  $\epsilon$ , but not  $R$ , and the power of  $\delta^{-1}$  will be absorbed into  $C(\epsilon)$ . This idea originates from the work of Guth in [9] and the author later used it in [1–3].

Once we have decided to work with these pieces from each  $S_i$ , we run the argument in section 2: we translate the pieces so that  $0 \in S_i$ , and redefine all the entities there.

The proof of Theorem 1.3 relies on estimating  $\prod_{i=1}^k \mathcal{E}_i f_i$  on cubes on the physical side and analyze how the estimate behaves as the size of the cube goes to infinity by using an inductive type argument with respect to the size of the cube. As we move from one spatial scale to a larger one, we will have to tolerate slightly larger Fourier support in the argument. But this accumulation is in the form of a convergent geometric series, therefore the only harm it does is imposing an additional technical layer in the argument.

**Definition 4.1.** *Given  $R \geq \delta^{-2}$  we define  $A(R)$  to be the best constant for which the estimate*

$$(4.1) \quad \left\| \prod_{i=1}^k \mathcal{E}_i f_i \right\|_{L^{\frac{2}{k-1}}(Q)} \leq A(R, \delta, \mu) \prod_{i=1}^k \|f_i\|_{L^2}$$

*holds true for all cubes  $Q \in \mathcal{C}(R)$ , under the assumption that  $f_i$  is supported in the set*

$$B_{\mu_i+10R^{-1}}(M_i) \cap B(0, \delta), \quad \forall i = 1, \dots, k.$$

In the above we have used the notation  $\mu = (\mu_1, \dots, \mu_k)$ . The induction starts from  $R \geq \delta^{-2}$  in order to be able to propagate the support hypothesis. We also tacitly assume that  $\mu_i \ll \delta^{-2}, \forall i = 1, \dots, k$ , or else the final gain of  $\mu_i^{\frac{c_i}{2}}$  is indistinguishable from powers of  $\delta^{-1}$  which will naturally contribute to the final constant  $C(\epsilon)$  that appears in the min estimate (1.8).

We provide an estimate inside any cube  $Q \in \mathcal{C}(\delta^{-1}R)$  based on prior information on estimates inside cubes  $q \in \mathcal{C}(R) \cap Q$ . Without restricting the generality of the argument, we assume that  $Q$  is centered at the origin and recall that each  $q \in \mathcal{C}(R) \cap Q$  has its center in  $R\mathcal{L}$ . When such a  $q$  is projected using  $\pi_i$  onto  $\mathcal{H}_i$  one obtains  $\pi_i q \in \mathcal{CH}_i(R)$ .

Each  $q \in \mathcal{C}(R) \cap Q$  has size  $R$  and the induction hypothesis is the following:

$$(4.2) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \leq A(R, \delta, \mu) \Pi_{i=1}^k \|f_i\|_{L^2}.$$

We strengthen this as follows:

$$(4.3) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \lesssim A(R, \delta, \mu) \cdot \Pi_{i=1}^k \left( \sum_{q' \in \mathcal{CH}_i(R)} \left\langle \frac{d(\pi_i q, q')}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_i - c(q')}{R} \right\rangle^N \chi_{q'} \mathcal{F}_i^{-1} f_i \right\|_{L^2}^2 \right)^{\frac{1}{2}}.$$

The basic idea in (4.3) is the following: if  $q' \neq \pi_i q$ , then  $\mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1)$  has off-diagonal type contribution outside  $q' \times [-\delta^{-1}R, \delta^{-1}R]$  (the interval stands for the  $i$ 'th slot), thus it has off-diagonal type contribution to the left-hand side of (4.3). We now turn to the details and fix  $i = 1$  and  $q' \in \mathcal{CH}_1(R)$ . With  $x = (x_1, x'_1)$  we have

$$\begin{aligned} & \| (x'_1 - c(q') - x_1 \nabla \varphi_1(\xi'_0)) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i \|_{L^{\frac{2}{k-1}}(q)} \\ & \leq \| (x'_1 - c(q') - x_1 \nabla \varphi_1(\frac{D'}{i})) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i \|_{L^{\frac{2}{k-1}}(q)} \\ & \quad + \| x_1 (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\frac{D'}{i})) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i \|_{L^{\frac{2}{k-1}}(q)} \\ & = \| \mathcal{E}_1 \mathcal{F}_1((x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i \|_{L^{\frac{2}{k-1}}(q)} \\ & \quad + \| x_1 \mathcal{E}_1 \mathcal{F}_1((\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i \|_{L^{\frac{2}{k-1}}(q)} \\ & \leq A(R, \delta, \mu) \| (x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} \\ & \quad + A(R, \delta, \mu) \delta^{-1} R \| (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} \Pi_{i=2}^k \| f_i \|_{L^2} \\ & \lesssim A(R, \delta, \mu) (\| (x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2} + R \| \chi_{q'} \mathcal{F}_1^{-1} f_1 \|_{L^2}) \Pi_{i=2}^k \| f_i \|_{L^2} \\ & \lesssim RA(R, \delta, \mu) \left\| \left\langle \frac{x'_1 - c(q')}{R} \right\rangle \chi_{q'} \mathcal{F}_1^{-1} f_1 \right\|_{L^2} \Pi_{i=2}^k \| f_i \|_{L^2}. \end{aligned}$$

We have used the following: (2.7) in justifying the equality between the terms on the second and fourth line, the induction hypothesis and the

fact that inside  $Q$  we have  $|x_1| \lesssim \delta^{-1}R$  to justify the inequality in the sixth line. Note that it is in the above use of the induction estimate for  $\mathcal{E}_1\mathcal{F}_1((x'_1 - c(q'))\chi_{q'}\mathcal{F}_1^{-1}f_1)$  that we need to tolerate the relaxed support of  $f_1$ .  $f_1$  is supported in the set  $B_{\mu_1+10(\delta^{-1}R)^{-1}}(M_1) \cap B_1(0, \delta)$ , and this support is impacted by the convolution with  $\mathcal{F}_1((x'_1 - c(q'))\chi_{q'})$  which is supported in  $B_1(0, R^{-1})$ ; the sum set of the two supports is a subset of  $B_{\mu_1+10(\delta^{-1}R)^{-1}}(M_1) \cap B_1(0, \delta)$ ; thus we can invoke the induction hypothesis at scale  $R$ .

For any  $q \in \mathcal{C}(R) \cap Q$  and  $x' \in \pi_{N_1}(q)$ , it holds that  $\langle \frac{x' - c(q') - x_1 \nabla \varphi_1(\xi'_0)}{R} \rangle \approx \langle \frac{d(\pi_{N_1}(q), q')}{R} \rangle$ . This is justified by the fact that  $|x_1| \lesssim \delta^{-1}R$  and  $|\nabla \varphi_1(\xi'_0)| \leq \delta$ , therefore the contribution of  $|x_1 \nabla \varphi_1(\xi'_0)| \leq R$  is negligible. From this and the previous set of estimates, we conclude that

$$\begin{aligned} & \|\mathcal{E}_1\mathcal{F}_1(\chi_{q'}\mathcal{F}_1^{-1}f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ & \lesssim A(R, \delta, \mu) \langle \frac{d(\pi_{N_1}(q), q')}{R} \rangle^{-1} \|\langle \frac{x'_1 - c(q')}{R} \rangle \chi_{q'}\mathcal{F}_1^{-1}f_1\|_{L^2} \Pi_{i=2}^k \|f_i\|_{L^2}. \end{aligned}$$

We claim that we can extend the argument above to prove

$$\begin{aligned} (4.4) \quad & \|\mathcal{E}_1\mathcal{F}_1(\chi_{q'}\mathcal{F}_1^{-1}f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ & \lesssim_N A(R, \delta, \mu) \langle \frac{d(\pi_{N_1}(q), q')}{R} \rangle^{-N} \|\langle \frac{x'_1 - c(q')}{R} \rangle^N \chi_{q'}\mathcal{F}_1^{-1}f_1\|_{L^2} \Pi_{i=2}^k \|f_i\|_{L^2}, \end{aligned}$$

for all  $N \geq 1$ . In repeating the argument above, we need to start with the higher order terms:

$$\|(x'_1 - c(q') - x_1 \nabla \varphi_1(\xi'_0))^N \mathcal{E}_1\mathcal{F}_1(\chi_{q'}\mathcal{F}_1^{-1}f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)}.$$

If  $N \geq 2$  some more, but harmless, terms appear due to the lack of commutativity between the symbols of the operators that are being used. We do this for  $N = 2$ ; the other cases are treated in a similar fashion. We use the following operator decomposition

$$\begin{aligned} x'_1 - c(q') - x_1 \nabla \varphi_1(\xi'_0) &= A + B, \\ A &= x'_1 - c(q') - x_1 \nabla \varphi_1\left(\frac{D'}{i}\right), \\ B &= x_1(\nabla \varphi_1(\xi'_0) - \nabla \varphi_1\left(\frac{D'}{i}\right)), \end{aligned}$$

based on which we have

$$(x'_1 - c(q') - x_1 \nabla \varphi_1(\xi'_0))^2 = (A + B)^2 = A^2 + B^2 + 2BA + [A, B].$$

Each component is treated as follows. The contribution of the term with  $A^2$  is estimated as above and using (2.7). The contribution of the term with  $B^2$  is estimated as above and using the fact that

$$B^2 = x_1^2 (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\frac{D'}{i}))^2,$$

which is a consequence of the commutativity of  $x_1$  and  $\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\frac{D'}{i})$ . The contribution of the  $BA$  term is estimated as follows:

$$\begin{aligned} & \|BA \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &= \|x_1 (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\frac{D'}{i})) \mathcal{E}_1 \mathcal{F}_1((x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &\lesssim \delta^{-1} R \|\mathcal{E}_1 (\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \mathcal{F}_1((x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &\lesssim \delta^{-1} R \|(\nabla \varphi_1(\xi'_0) - \nabla \varphi_1(\xi')) \mathcal{F}_1((x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1)\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim R \|\mathcal{F}_1((x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1)\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim R \|(x'_1 - c(q')) \chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim R^2 \|\langle \frac{x'_1 - c(q')}{R} \rangle \chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2}. \end{aligned}$$

Finally since  $[A, B] = -\frac{x_1}{i} \Delta \varphi_1(\frac{D'}{i})$ , we estimate its contribution as follows

$$\begin{aligned} & \|[A, B] \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &= \|x_1 \Delta \varphi_1(\frac{D'}{i}) \mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &\lesssim \delta^{-1} R \|\mathcal{E}_1 (\Delta \varphi_1(\xi')) \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \Pi_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)} \\ &\lesssim \delta^{-1} R \|\Delta \varphi_1(\xi') \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1)\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim \delta^{-1} R \|\mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1)\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim \delta^{-1} R \|\chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2} \\ &\lesssim R^2 \|\chi_{q'} \mathcal{F}_1^{-1} f_1\|_{L^2} \cdot \Pi_{i=2}^k \|f_i\|_{L^2}, \end{aligned}$$

where, in passing to the last line, we have used that  $\delta^{-1} \ll R$ . Based on all the estimates above, we obtain (4.4) for  $N = 2$ . The above argument contains all the ingredients necessary to obtain (4.4) for  $N \geq 3$ .



Using (2.1), (2.8) and (4.4), we obtain

$$\begin{aligned} & \|\mathcal{E}_1 f_1 \cdot \prod_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)}^{\frac{2}{k-1}} \\ & \leq A(R, \delta, \mu)^{\frac{2}{k-1}} \sum_{q' \in \mathcal{CH}_1(R)} \|\mathcal{E}_1 \mathcal{F}_1(\chi_{q'} \mathcal{F}_1^{-1} f_1) \cdot \prod_{i=2}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(q)}^{\frac{2}{k-1}} \\ & \lesssim_N A(R, \delta, \mu)^{\frac{2}{k-1}} \\ & \quad \cdot \left( \sum_{q' \in \mathcal{CH}_1(R)} \left\langle \frac{d(\pi_{N_1} q, q')}{R} \right\rangle^{-N \cdot \frac{2}{k-1}} \left\| \left\langle \frac{x' - c(q')}{R} \right\rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1 \right\|_{L^2}^{\frac{2}{k-1}} \right) \\ & \quad \cdot \prod_{i=2}^k \|f_i\|_{L^2}^{\frac{2}{k-1}} \\ & \lesssim_N A(R, \delta, \mu)^{\frac{2}{k-1}} \prod_{i=2}^k \|f_i\|_{L^2}^{\frac{2}{k-1}} \\ & \quad \cdot \left( \sum_{q' \in \mathcal{CH}_1(R)} \left\langle \frac{d(\pi_{N_1} q, q')}{R} \right\rangle^{-(2N - (k-1)^2)} \left\| \left\langle \frac{x' - c(q')}{R} \right\rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1 \right\|_{L^2}^2 \right)^{\frac{1}{k-1}}. \end{aligned}$$

In justifying the last inequality, we have used the simple estimate for sequences

$$\|a_i \cdot b_i\|_{l_i^{\frac{2}{k-1}}} \lesssim \|a_i\|_{l_i^2} \|b_i\|_{l_i^{\frac{2}{k-2}}},$$

together with the straightforward estimate

$$\left\| \left\langle \frac{d(\pi_{N_1} q, q')}{R} \right\rangle^{-(\frac{k-1}{2})^2} \right\|_{l_{q'}^{\frac{2}{k-2}}} \lesssim 1.$$

Note that the previous inequality is (4.3) with the improvement for  $f_1$ . By repeating the procedure for all other terms  $f_2, \dots, f_k$  to conclude with (4.3).

Using (4.3) we invoke the discrete Loomis-Whitney inequality in (2.12) to conclude the argument. For  $i = 1, \dots, k$ , we define the functions  $g_i : \mathcal{L}(\mathcal{H}_i) \rightarrow \mathbb{R}$  by

$$g_i(\mathbf{j}) = \left( \sum_{q' \in \mathcal{CH}_i(R)} \left\langle \frac{d(q(\mathbf{j}), q')}{R} \right\rangle^{-(N - 2(k-1)^2)} \left\| \left\langle \frac{x' - c(q')}{R} \right\rangle^N \chi_{q'} \mathcal{F}_i^{-1} f_i \right\|_{L^2}^2 \right)^{\frac{1}{2}},$$

$\mathbf{j} \in \mathcal{L}(\mathcal{H}_i).$

For  $N$  large enough (depending only on  $n$ ), we claim the following estimate:

$$(4.5) \quad \|g_i\|_{l^{2l^\infty}(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i))} \lesssim C_i(\mu_i, \delta, R) \|f_i\|_{L^2}, \quad i = 1, \dots, k,$$

where  $C_i(\mu_i, \delta, R) = \min(1, (R\mu_i + 10\delta)^{\frac{c_i}{2}})$ . From (2.12) we conclude

$$(4.6) \quad A(\mu, \delta, \delta^{-1}R) \lesssim A(\mu, \delta, R) \prod_{i=1}^k C_i(\mu_i, \delta, R).$$

We will return to this estimate at the end of the proof and show how it is used to close the induction.

For now, we continue with the argument for (4.5). The bound with the choice of 1 in  $\min(1, (R\mu_i + 10\delta)^{\frac{c_i}{2}})$  is easily obtained from the stronger inequality:

$$\|g_i\|_{L^2(\mathcal{L}(\mathcal{H}'_i) \times \mathcal{L}(\mathcal{H}''_i))} \lesssim \|f_i\|_{L^2}, \quad i = 1, \dots, k.$$

The bound with the choice of  $(R\mu_i + 10\delta)^{\frac{c_i}{2}}$  in  $\min(1, (R\mu_i + 10\delta)^{\frac{c_i}{2}})$  is addressed below. We prove this claim in the case  $i = 1$ , the other choices of  $i$  being analogous; note that it suffices to prove

$$(4.7) \quad \|\mathcal{F}_1^{-1} f_1\|_{L^2(q(\mathbf{j}))} \|f_1\|_{L^\infty(\mathcal{L}(\mathcal{H}'_1) \times \mathcal{L}(\mathcal{H}''_1))} \lesssim (R\mu_1 + 10\delta)^{\frac{c_1}{2}} \|f_1\|_{L^2},$$

where in the above the variable  $\mathbf{j}$  runs in  $\mathcal{L}(\mathcal{H}_1)$ . In the context  $i = 1$  we ignore the variable along  $n_1$ . The variable in  $\mathcal{H}_1$  that was originally labeled by  $x'_1$  will be denoted by  $x$  and is further split as follows  $x = (x', x'') \in \mathbb{R}^{n-c_1-1} \times \mathbb{R}^{c_1}$ , with the dual Fourier variables being  $\xi = (\xi', \xi'')$ .  $f_1$  is localized in  $B_{\mu_1+10\delta R^{-1}}(M_1)$  where  $M_1$  is a  $n - c_1 - 1$ -dimensional manifold. As we discussed in Section 2, we can parametrize it as follows:  $(\xi', \Phi(\xi'))$  with  $\xi' \in \mathbb{R}^{n-c_1-1}$  and  $\Phi(\xi') = (\tilde{\varphi}_1(\xi'), \dots, \tilde{\varphi}_{c_1}(\xi'))$ ; this is true at least locally, but given that we work with  $f_1$  supported in  $B_1(0, 20\delta^2)$  with  $\delta \ll 1$ , we can obtain the parametrization in the whole support of  $f_1$ .

The scope of what follows next is to "flatten" the manifold  $M_1$ , since this allows us to exploit the localization in an easier way. We write

$$\begin{aligned} \mathcal{F}_1^{-1} f_1(x) &= \int e^{ix \cdot \xi} f_1(\xi) d\xi \\ &= \int e^{i(x' \cdot \xi' + x'' \cdot (\Phi(\xi') + \xi''))} f_1(\xi', \Phi(\xi') + \xi'') d\xi' d\xi'' \\ &= \int \left( e^{ix'' \cdot \xi''} \int e^{i(x' \cdot \xi' + x'' \cdot \Phi(\xi'))} h_1(\xi', \xi'') d\xi' \right) d\xi'' \\ &= \int e^{ix'' \cdot \xi''} \mathcal{E}_\Phi h_1(\cdot, \xi'') d\xi'' \end{aligned}$$

where

$$\mathcal{E}_\Phi h(x) = \int e^{i(x' \cdot \xi' + x'' \cdot \Phi(\xi'))} h(\xi') d\xi', \quad h_1(\xi', \xi'') = f_1(\xi', \Phi(\xi') + \xi'').$$

From its definition we see that  $h_1$  is supported in a set where  $|\xi'| \leq \delta + 10R^{-1}$  and  $\|\xi''\|_\infty \leq \mu_1 + 10\delta R^{-1}$ . A straightforward computation shows that  $\|h_1\|_{L^2} = \|f_1\|_{L^2}$ .

We note that for any  $q$  cube of size  $R$  the following estimate holds true:

$$(4.8) \quad \|\mathcal{E}_\Phi h\|_{L^2(q)}^2 \lesssim R^{\frac{c_1}{2}} \|h\|_{L^2},$$

where  $h \in L^2_{\xi'}$  depends only on the variable  $\xi'$ . This simply follows from the fact that  $\mathcal{E}_\Phi$  is an isometry on  $L^2_{x'}$  and by integrating with respect to the  $x''$  variable in an interval of size  $R$  (in all directions contained in  $x''$ ).

Using (4.8) we obtain

$$(4.9) \quad \begin{aligned} \|\mathcal{F}_1^{-1} f_1\|_{L^2(q)} &\leq R^{\frac{c_1}{2}} \int \|h_1(\cdot, \xi'')\|_{L^2_{\xi'}} d\xi'' \\ &\leq R^{\frac{c_1}{2}} (\prod_{j=1}^{c_1} (\mu_1 + 10\delta R^{-1}))^{\frac{1}{2}} \|h_1\|_{L^2} \\ &= (R\mu_1 + 10\delta)^{\frac{c_1}{2}} \|f_1\|_{L^2}. \end{aligned}$$

The constant obtained here has the correct numerology, but the estimate misses some additional localization that is necessary in closing the argument for (4.7). To prove (4.7), it suffices to prove the following more localized version:

$$(4.10) \quad \|\mathcal{F}_1^{-1} f_1\|_{L^2(q)}^2 \lesssim R^{\frac{c_1}{2}} \cdot \int \left( \sum_{q' \in \mathcal{C}\mathcal{H}'_1} \left\langle \frac{d(\pi'q, q')}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x' - c(q')}{R} \right\rangle^N \chi_{q'}(\mathcal{F}'_1)^{-1} h_1(\cdot, \xi'') \right\|_{L^2} \right) d\xi'',$$

where  $\pi'$  is the projection onto  $\mathcal{H}'_1$  and  $N \in \mathbb{N}$  and  $(\mathcal{F}'_1)^{-1}$  is the inverse Fourier transform with taken with respect to the variable  $x'$  only. Indeed, if we show (4.10), then

$$\begin{aligned} \|\|\mathcal{F}_1^{-1} f_1\|_{L^2(q(\mathbf{j}))}\|_{l^2l^\infty(\mathcal{L}(\mathcal{H}'_1) \times \mathcal{L}(\mathcal{H}''_1))} &\lesssim \int R^{\frac{c_1}{2}} \|h_1(\cdot, \xi'')\|_{L^2} d\xi'' \\ &\lesssim (R\mu_1 + 10\delta)^{\frac{c_1}{2}} \|h_1\|_{L^2} \\ &= (R\mu_1 + 10\delta)^{\frac{c_1}{2}} \|f_1\|_{L^2}, \end{aligned}$$

which establishes (4.7); in the above we have used that  $h_1$  is supported in a region with  $\|\xi''\|_\infty \leq \mu_1 + 10\delta R^{-1}$ .

It is also clear that (4.10) follows from the following estimate

$$(4.11) \quad \|\mathcal{E}_\Phi h\|_{L^2(q)}^2 \lesssim R^{\frac{c_1}{2}} \cdot \left( \sum_{q' \in \mathcal{CH}'_1} \left\langle \frac{d(\pi'q, q')}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x' - c(q')}{R} \right\rangle^N \chi_{q'}(\mathcal{F}'_1)^{-1} h \right\|_{L^2} \right),$$

where  $h \in L^2_\xi$ , depends only on the variable  $\xi'$ ;  $h$  inherits the properties that  $h_1$  has with respect to the  $\xi'$  variable, mainly the support in a ball of size  $\delta + 10R^{-1}$ . This is an improvement over the use of the isometry property in deriving (4.8) and it is analogous to how (4.3) improves upon (4.2). The analogy carries on to how we establish (4.11): we start from (4.8) and use the multiplier  $x' - c(q') - x''\nabla\Phi(\xi'_0)$  (where  $\xi'_0$  belongs to the support of  $h$ ) to run a similar argument as for the derivation of (4.3); the details are left to the interested reader.

We now return to (4.6)

$$A(\mu, \delta, \delta^{-1}R) \leq CA(\mu, \delta, R)\prod_{i=1}^k C_i(\mu_i, \delta, R)$$

and show how it implies the result in Theorem 1.3. Recall that  $C$  is independent of  $\delta, R$  and  $\mu$  and that  $C_i(\mu_i, \delta, R) = \min(1, (R\mu_j^i + 10\delta))^{\frac{c_i}{2}}$ . The localization in Theorem 1.3 and the one in Definition 4.1 are not quite the same, but they match provided that  $R \geq \max\{\mu_1^{-1}, \dots, \mu_k^{-1}\}$ ; we first provide the argument for this case.

For any  $R \geq \max\{\mu_1^{-1}, \dots, \mu_k^{-1}\}$ , we iterate (4.6) to obtain

$$A(\mu, \delta, R) \leq C^N A(\mu, \delta, \delta^N R)\prod_{m=1}^N \prod_{i=1}^k C_i(\mu_i, \delta, \delta^m R).$$

We pick  $N$  such that  $\delta^{-1} \leq \delta^N R \leq \delta^{-2}$ . We quantify the effect of some  $\mu_i$  for some  $i \in \{1, \dots, k\}$  in the expression above as follows:

$$\prod_{m=1}^N \min(1, (\delta^m R\mu_i + 10\delta)^{\frac{c_i}{2}}).$$

In the above we gain factors of  $\delta$  for as long as  $\delta^m R\mu_i \leq 10\delta$ ; the accumulating powers of  $10^{\frac{c_i}{2}}$  are added to the factor  $C^N$  above. Thus it follows that

$$(4.12) \quad \prod_{m=1}^N \min(1, (\delta^m R\mu_i + 10\delta)^{\frac{c_i}{2}}) \leq C_0^N \mu_i^{\frac{c_i}{2}},$$

where  $C_0$  is independent of  $\delta, R, \mu$ ; here we have used the fact that we have chosen  $R$  large enough. Thus we conclude with:

$$A(\mu, \delta, R) \leq C^N A(\mu, \delta, \delta^N R) \prod_{i=1}^k \mu_i^{\frac{c_i}{2}},$$

for any  $R \geq \max\{\mu_1^{-1}, \dots, \mu_k^{-1}\}$  and where  $N$  is chosen such that  $\delta^{-1} \leq \delta^N R \leq \delta^{-2}$ . From this we obtain

$$A(\mu, \delta, R) \leq C^N \max_{r \leq \delta^{-2}} A(\mu, \delta, r) \prod_{i=1}^k \mu_i^{\frac{c_i}{2}}.$$

From the uniform pointwise bound

$$\|\Pi_{i=1}^{n+1} \mathcal{E}_i f_i\|_{L^\infty} \lesssim \Pi_{i=1}^{n+1} \|\mathcal{E}_i f_i\|_{L^\infty} \lesssim \Pi_{i=1}^{n+1} \|f_i\|_{L^2}$$

it follows that  $\max_{r \leq \delta^{-2}} A(\mu, \delta, r) \lesssim \delta^{-10}$ . For  $R \in [\delta^{-N}, \delta^{-N-1}]$ , and  $N$  large enough so that  $R \geq \max\{\mu_1^{-1}, \dots, \mu_k^{-1}\}$ , the above implies

$$A(\mu, \delta, R) \leq C^N C(\delta) \leq R^\epsilon C(\delta)$$

provided that  $C^N \leq \delta^{-N\epsilon}$ . Therefore choosing  $\delta = C^{-\frac{1}{\epsilon}}$  leads to the desired result.

We now consider the case  $R \leq \max\{\mu_1^{-1}, \dots, \mu_k^{-1}\}$ . In this case, Definition 4.1 requires localization properties at larger scales since  $R^{-1} \geq \min\{\mu_1, \dots, \mu_k\}$ ; this is not a problem at all. However, when we run the argument above, we will not capture the full factor  $\prod_{i=1}^k \mu_i^{\frac{c_i}{2}}$ , and this is reasonable given that we work with weaker localization properties.

For simplicity, let us assume  $\mu_1 \leq \mu_2 \leq \dots \leq \mu_k$  (which does not restrict the generality of the argument) and look at the particular case  $\mu_1 \leq R^{-1} \leq \mu_2$  - the other cases are treated in a similar fashion. Then we run the above argument and note that the factor  $\prod_{i=1}^k \mu_i^{\frac{c_i}{2}}$  needs to be replaced by

$$R^{-\frac{c_1}{2}} \prod_{i=2}^k \mu_i^{\frac{c_i}{2}},$$

which is a simple consequence of re-quantifying the estimate (4.12). Thus, relatively to (1.8), we are missing a factor of  $(R\mu_1)^{\frac{c_1}{2}}$ , which needs to be recovered. As of now we obtain that the inequality

$$(4.13) \quad \|\Pi_{i=1}^k \mathcal{E}_i f_i\|_{L^{\frac{2}{k-1}}(Q)} \leq C(\epsilon) R^\epsilon R^{-\frac{c_1}{2}} \prod_{i=2}^k \mu_i^{\frac{c_i}{2}} \prod_{i=1}^k \|f_i\|_{L^2}$$

holds true for all cubes  $Q \in \mathcal{C}(R)$ , under the assumption that  $f_i$  is supported in  $B_{\mu_i+10R^{-1}}(M_i) \cap B(0, \delta), \forall i = 1, \dots, k$ . As before we assume that  $Q$  is centered at the origin. To improve this, we take advantage of the fact that, in the original estimate that we seek to establish,  $f_1$  has better localization properties: it is supported in  $B_{\mu_1}(M_1)$ . (4.3) allows us to replace  $\|f_1\|_{L^2}$  in (4.13) by

$$\left( \sum_{q' \in \mathcal{CH}_1(R)} \left\langle \frac{d(\pi_1 Q, q')}{R} \right\rangle^{-(2N-n^2)} \left\| \left\langle \frac{x'_1 - c(q')}{R} \right\rangle^N \chi_{q'} \mathcal{F}_1^{-1} f_1 \right\|_{L^2}^2 \right)^{\frac{1}{2}}$$

Morally this implies that, for the estimate at scale  $R$ , the contribution from  $f_1$  that really matters is  $\|\mathcal{F}_1^{-1} f_1\|_{L^2(q)}$ , where  $q = \pi_1 Q \in \mathcal{CH}_1(R)$ , and the rest of the terms come with appropriate decay. On the other hand, a similar argument to the one used for (4.9) gives that for any  $q \in \mathcal{CH}_1(R)$

$$\|\mathcal{F}_1^{-1} f_1\|_{L^2(q)} \lesssim (R\mu_1)^{\frac{c_1}{2}} \|f_1\|_{L^2};$$

here we use that  $f_1$  has better localization properties. This brings in the correction we needed in (4.13); we leave the details as an exercise to the interested reader.

### Acknowledgement

Part of this work was supported by an NSF grant, DMS-1900603. The author thanks Jonathan Bennett for in-depth help with the current literature on the subject. The author thanks Pavel Zorin-Kranich for bringing to our attention that some additional details would have been useful in our original work on the subject in [1]; the corresponding details have been added in the current paper in Section 4 in proving (4.3) for the case  $N = 2$ . The author thanks to the anonymous referee for the careful reading of the paper and the useful recommendations.

### References

[1] I. Bejenaru, *The multilinear restriction estimate: a short proof and a refinement*, Math. Res. Lett. **24** (2017), no. 6, 1585–1603.  
 [2] ———, *The optimal trilinear restriction estimate for a class of hypersurfaces with curvature*, Adv. Math. **307** (2017) 1151–1183.

- [3] ———, *Optimal multilinear restriction estimates for a class of hypersurfaces with curvature*, *Anal. PDE* **12** (2019), no. 4, 1115–1148.
- [4] J. Bennett, *Aspects of multilinear harmonic analysis related to transversality*, in *Harmonic analysis and partial differential equations*, Vol. 612 of *Contemp. Math.*, 1–28, Amer. Math. Soc., Providence, RI (2014).
- [5] J. Bennett, A. Carbery, and T. Tao, *On the multilinear restriction and Keakeya conjectures*, *Acta Math.* **196** (2006), no. 2, 261–302.
- [6] J. Bourgain, *Estimates for cone multipliers*, in *Geometric aspects of functional analysis (Israel, 1992–1994)*, Vol. 77 of *Oper. Theory Adv. Appl.*, 41–60, Birkhäuser, Basel (1995).
- [7] J. Bourgain and L. Guth, *Bounds on oscillatory integral operators based on multilinear estimates*, *Geom. Funct. Anal.* **21** (2011), no. 6, 1239–1295.
- [8] L. Guth, *The endpoint case of the Bennett-Carbery-Tao multilinear Keakeya conjecture*, *Acta Math.* **205** (2010), no. 2, 263–286.
- [9] ———, *A short proof of the multilinear Keakeya inequality*, *Math. Proc. Cambridge Philos. Soc.* **158** (2015), no. 1, 147–153.
- [10] ———, *A restriction estimate using polynomial partitioning*, *J. Amer. Math. Soc.* **29** (2016), no. 2, 371–413.
- [11] ———, *Restriction estimates using polynomial partitioning II*, *Acta Math.* **221** (2018), no. 1, 81–142.
- [12] J. Hickman and K. M. Rogers, *Improved Fourier restriction estimates in higher dimensions*, [arXiv:1807.10940](https://arxiv.org/abs/1807.10940).
- [13] E. M. Stein, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Vol. 43 of *Princeton Mathematical Series*, Princeton University Press, Princeton, NJ (1993), ISBN 0-691-03216-5. With the assistance of Timothy S. Murphy, *Monographs in Harmonic Analysis*, III.
- [14] T. Tao, *Sharp bounds for multilinear curved Keakeya, restriction and oscillatory integral estimates away from the endpoint*, *Mathematika* **66** (2020), no. 2, 517–576.
- [15] ———, *The Bochner-Riesz conjecture implies the restriction conjecture*, *Duke Math. J.* **96** (1999), no. 2, 363–375.

- [16] T. Tao and A. Vargas, *A bilinear approach to cone multipliers. I. Restriction estimates*, *Geom. Funct. Anal.* **10** (2000), no. 1, 185–215.
- [17] H. Wang, *A restriction estimate in  $\mathbf{R}^3$  using brooms*, *Duke Math. J.* **171** (2022), no. 8, 1749–1822.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, SAN DIEGO  
LA JOLLA, CA 92093-0112, USA  
*E-mail address:* `ibejenaru@math.ucsd.edu`

RECEIVED JANUARY 2, 2020

ACCEPTED JUNE 23, 2020