

# LOCALISED VARIANTS OF MULTILINEAR RESTRICTION

DAVID BELTRAN, JENNIFER DUNCAN, AND JONATHAN HICKMAN

**Abstract:** We revisit certain localised variants of the Bennett–Carbery–Tao multilinear restriction theorem, recently proved by Bejenaru. We give a new proof of Bejenaru’s theorem, relating the estimates to the theory of Keakeya–Brascamp–Lieb inequalities. Moreover, the new proof allows for a substantial generalisation, exploiting the full power of the Keakeya–Brascamp–Lieb theory.

**2020 Mathematics Subject Classification:** 42B20.

**Key words:** multilinear Fourier restriction, multilinear Keakeya, Brascamp–Lieb inequalities.

## 1. Introduction

**1.1. Background.** For  $1 \leq d \leq n$ , let  $\Sigma: U \rightarrow \mathbb{R}^n$  be a parametrisation of a smooth  $d$ -dimensional submanifold  $S$  of  $\mathbb{R}^n$ . That is,  $U \subset \mathbb{R}^d$  is an open, connected neighbourhood of the origin,  $\Sigma: U \rightarrow \mathbb{R}^n$  is a smooth, injective, regular map, and  $S = \Sigma(U)$ . In particular,  $\Sigma \in C^\infty(U)$  and  $\bigwedge_{j=1}^d \frac{\partial \Sigma}{\partial u_j}(u) \neq 0$  for all  $u \in U$ . By an abuse of notation, we shall often simply refer to  $S$ , with the tacit understanding that our analysis depends on a choice of parametrisation.

Given  $a \in C_c^\infty(U)$  satisfying  $0 \leq a(u) \leq 1$  for all  $u \in U$ , we define the *extension operator*  $E_S$  associated to  $S$  (or, more precisely,  $\Sigma: U \rightarrow \mathbb{R}^n$ ) by

$$(1.1) \quad E_S f(x) := \int_U e^{ix \cdot \Sigma(u)} f(u) a(u) du, \quad f \in L^1(S), x \in \mathbb{R}^n.$$

Here and below, again by an abuse of notation, we write  $L^p(S)$  in place of  $L^p(U)$  for  $1 \leq p \leq \infty$ . We refer to  $a$  as the *amplitude* of  $E_S$  and let  $\rho(E_S) := \text{diam supp } a > 0$ .

For  $\xi = \Sigma(u) \in S$ , where  $u \in U$ , the *tangent space*  $T_\xi S$  is the  $d$ -dimensional subspace of  $\mathbb{R}^n$  spanned by the vectors  $\frac{\partial \Sigma}{\partial u_1}(u), \dots, \frac{\partial \Sigma}{\partial u_d}(u)$ , and the *normal space*  $N_\xi S$  is the orthogonal complement of  $T_\xi S$  in  $\mathbb{R}^n$ . For  $1 \leq j \leq k$ , let  $V_j$  be a vector subspace of  $\mathbb{R}^n$  of dimension  $d_j$  and  $\{v_{j,1}, \dots, v_{j,d_j}\}$  be a choice of orthonormal basis of  $V_j$ . We define

$$|V_1 \wedge \dots \wedge V_k| := \left| \bigwedge_{j=1}^k \bigwedge_{l=1}^{d_j} v_{j,l} \right|,$$

noting that this quantity is independent of the choice of orthonormal bases.

**Definition 1.1.** Fix  $2 \leq k \leq n$  and for  $1 \leq j \leq k$  let  $S = \Sigma_j(U_j)$  be a smooth hypersurface in  $\mathbb{R}^n$ . Further, let  $\mathbf{q}_k := (q_k)_{j=1}^k = (q_k, \dots, q_k)$ , where  $q_k := \frac{2}{k-1}$ . We say  $\mathcal{E} = ((S_j)_{j=1}^k, \mathbf{q}_k)$  is a *transverse ensemble* if

$$N_{\mathcal{E}} := \left| \bigwedge_{j=1}^k N_{\xi_j} S_j \right| > 0, \quad \text{where } \xi_j := \Sigma_j(0) \in S_j \text{ for } 1 \leq j \leq k.$$

A fundamental result in modern Fourier analysis is the celebrated Bennett–Carbery–Tao theorem. Here and below,  $Q_R := [-R, R]^n$  for  $R > 0$ .

**Theorem 1.2** (Bennett–Carbery–Tao [9, §5]). *Let  $2 \leq k \leq n$  and  $\mathcal{E} = ((S_j)_{j=1}^k, \mathbf{q}_k)$  be a transverse ensemble in  $\mathbb{R}^n$ . There exists  $\rho_{\mathcal{E}} > 0$  such that for  $\max_j \rho(E_{S_j}) < \rho_{\mathcal{E}}$  the following holds. For all  $\varepsilon > 0$ , there exists a constant  $C_{\mathcal{E}, \varepsilon} \geq 1$  such that*

$$(1.2) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_j} f_j|^{q_k} \leq C_{\mathcal{E}, \varepsilon} R^\varepsilon \prod_{j=1}^k \|f_j\|_{L^2(S_j)}^{q_k}$$

holds for all  $R \geq 1$  and  $f_j \in L^2(S_j)$ ,  $1 \leq j \leq k$ .

Here we investigate localised variants of Theorem 1.2 where some of the functions  $f_j$  are assumed to be supported in thin sets. Under such hypotheses, we can hope to gain in the size of the constant appearing on the right-hand side of (1.2). This phenomenon was observed in recent papers of Bejenaru [1, 3], and we begin by describing the existing results.

Let  $S = \Sigma(U)$  be a  $d$ -dimensional submanifold of  $\mathbb{R}^n$ , as above. For  $1 \leq d' \leq d$ , we say  $S' \subseteq S$  is a  $d'$ -dimensional submanifold of  $S$  (or a *codimension  $d - d'$  submanifold of  $S$* ) if  $S' = \Sigma(M)$  for  $M = \gamma(U')$  a  $d'$ -dimensional submanifold of  $\mathbb{R}^d$  with parametrisation  $\gamma: U' \rightarrow U$ , where we always assume that  $\gamma(U')$  is compactly contained in  $U$  and  $\gamma(0) = 0$ . In this case, we write  $\text{codim}(S'|S) := d - d'$ . Note that  $S'$  is also a  $d'$ -dimensional submanifold of  $\mathbb{R}^n$ , with  $S' = \Sigma'(U')$  for  $\Sigma' := \Sigma \circ \gamma$  satisfying

$$(1.3) \quad \Sigma(0) = \Sigma'(0).$$

Continuing to abuse notation, given  $\mu > 0$  we say  $f \in L^2(S)$  is *supported in  $\mathcal{N}_\mu S'$* , or  $\text{supp } f \subseteq \mathcal{N}_\mu S'$ , if  $f$  is essentially supported in  $\mathcal{N}_\mu M$ , where  $\mathcal{N}_\mu M \subseteq \mathbb{R}^d$  is the (open Euclidean)  $\mu$ -neighbourhood of  $M$ .

**Definition 1.3.** Fix  $2 \leq k \leq n$  and for  $1 \leq j \leq k$  let  $S_j$  be a smooth hypersurface and  $S'_j = \Sigma'_j(U'_j)$  be a smooth submanifold of  $S_j$ . Further, let  $\mathbf{q}_k := (q_k)_{j=1}^k = (q_k, \dots, q_k)$ , where  $q_k := \frac{2}{k-1}$ . We say  $\mathcal{E} = ((S_j, S'_j)_{j=1}^k, \mathbf{q}_k)$  is a *transverse ensemble* in  $\mathbb{R}^n$  if

$$(1.4) \quad N_{\mathcal{E}} := \left| \bigwedge_{j=1}^k N_{\xi'_j} S'_j \right| > 0, \quad \text{where } \xi'_j := \Sigma'_j(0) \in S'_j \text{ for } 1 \leq j \leq k.$$

We say  $\mathcal{E}$  has *codimension  $(m_j)_{j=1}^k$* , where  $m_j := \text{codim}(S'_j|S_j)$  for  $1 \leq j \leq k$ .

With the above definition, the local version of Theorem 1.2 reads as follows.

**Theorem 1.4** (Bejenaru [3]). *Let  $2 \leq k \leq n$  and  $\mathcal{E} = ((S_j, S'_j)_{j=1}^k, \mathbf{q}_k)$  be a transverse ensemble in  $\mathbb{R}^n$  of codimension  $(m_j)_{j=1}^k$ . There exists a constant  $\rho_{\mathcal{E}} > 0$  such that for  $\max_j \rho(E_{S_j}) < \rho_{\mathcal{E}}$  the following holds. For all  $\varepsilon > 0$ , there exists a constant  $C_{\mathcal{E}, \varepsilon} \geq 1$  such that for any  $0 < \mu_j < 1$  the inequality*

$$(1.5) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_j} f_j|^{q_k} \leq C_{\mathcal{E}, \varepsilon} R^\varepsilon \prod_{j=1}^k \mu_j^{m_j q_k / 2} \prod_{j=1}^k \|f_j\|_{L^2(S_j)}^{q_k}$$

holds for all  $R \geq 1$  and  $f_j \in L^2(S_j)$  satisfying  $\text{supp } f_j \subseteq \mathcal{N}_{\mu_j} S'_j$ .

An earlier (and more restricted) version of this result appeared in [1]. We remark that it is possible to take  $m_j = 0$  in Theorem 1.4; in particular, if  $m_1 = \dots = m_k = 0$ , then we recover Theorem 1.2. Note that in this case the support condition is vacuous.

We also note that the constant  $C_{\mathcal{E},\varepsilon}$  blows up as the value of  $N_{\mathcal{E}}$  in (1.4) goes to 0. Thus, the new content of Theorem 1.4 over Theorem 1.2 comes when  $0 < \mu_j \ll 1$  and  $N_{\mathcal{E}} \sim 1$  (or small depending on some quantity unrelated to our choice of  $\mu_j$  in applications).

Theorem 1.4 has recently found two distinct applications:

- In [2], Theorem 1.4 is combined with a variant of Guth's polynomial partitioning method [11] to establish a sharp  $(n - 1)$ -linear restriction estimate under a curvature hypothesis (see also [16]).
- More recently, in [4], a special case of Theorem 1.4 (under a low regularity hypothesis) was used to prove  $L^p \rightarrow L^q$  estimates for the maximal function associated to dilates of a helix in  $\mathbb{R}^3$ .

One aim of this note is to give a new proof of Theorem 1.4. For this, we follow the standard induction-on-scales framework of [6, 9], which perhaps provides a more contextualised approach than that of [1, 3]. In particular, we clarify the relationship between (1.5) and Kakeya-type inequalities. In [1, p. 1588] the author raises the question of whether (1.5) has a multilinear Kakeya counterpart. Here we show that the corresponding geometric estimates are Kakeya–Brascamp–Lieb inequalities of the type studied in, for instance, [6, 19] and their discretised/regularised variants, as studied in [14, 20]. Moreover, the new approach leads to our main result (Theorem 1.6), which is a substantial generalisation of Theorem 1.4, making full use of the Kakeya–Brascamp–Lieb theory.

**1.2. Regularised Brascamp–Lieb inequalities.** In order to state the main result, we briefly recall some elements of Brascamp–Lieb theory. For  $2 \leq k \leq n$  and  $1 \leq n_j \leq n$  for  $1 \leq j \leq k$ , a *Brascamp–Lieb datum* is a pair  $(\mathbf{L}, \mathbf{p}) := ((L_j)_{j=1}^k, (p_j)_{j=1}^k)$  where  $L_j: \mathbb{R}^n \rightarrow \mathbb{R}^{n_j}$  are linear surjective maps and  $p_j \in (0, 1]$ . We let  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p})$  denote the best constant  $C > 0$  for which the inequality

$$(1.6) \quad \int_{\mathbb{R}^n} \prod_{j=1}^k (f_j \circ L_j)^{p_j} \leq C \prod_{j=1}^k \left( \int_{\mathbb{R}^{n_j}} f_j \right)^{p_j}$$

holds for all non-negative functions  $f_j \in L^1(\mathbb{R}^{n_j})$  that are constant on cubes in the unit cube lattice  $\mathcal{Q}^{n_j} := [0, 1]^{n_j} + \mathbb{Z}^{n_j}$ . The inequalities (1.6) are typically referred to as *regularised* Brascamp–Lieb inequalities and were first introduced in [8]. Moreover, in [8] it was shown that

$$(1.7) \quad \text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}) < \infty \iff \dim(V) \leq \sum_{j=1}^k p_j \dim(L_j V) \quad \text{for all } V \leq \mathbb{R}^n.$$

Here the notation  $V \leq \mathbb{R}^n$  denotes that  $V$  is a linear subspace of  $\mathbb{R}^n$ . An advantage of regularised Brascamp–Lieb inequalities is that the characterisation (1.7) for the finiteness of  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p})$  does not involve the scaling condition  $\sum_{j=1}^k p_j n_j = n$ , which is necessary in the standard Brascamp–Lieb theory [7].

**1.3. Main result.** Our main theorem generalises Theorem 1.4 in three directions. First, rather than just considering hypersurfaces  $S_j$ , we allow the  $S_j$  to be of arbitrary dimension. Secondly, rather than considering a single submanifold  $S'_j \subseteq S_j$  for each  $1 \leq j \leq k$ , we consider nested families. The nested setup naturally arises in recent polynomial partitioning approaches to multilinear restriction and Kakeya; see [10, 12, 13, 18]. Finally, rather than work with the transversality condition (1.4), we formulate our transversality hypothesis in terms of regularised Brascamp–Lieb constants, as defined above. This Brascamp–Lieb formulation greatly relaxes the constraints on the permissible codimensions.

We first explain the nested setup. Suppose  $S_1$  is a  $d_1$ -dimensional submanifold of a  $d_0$ -dimensional submanifold  $S_0 = \Sigma_0(U_0)$  of  $\mathbb{R}^n$ , as in §1.1. Recall that  $S_1$  is also a submanifold of  $\mathbb{R}^n$  with parametrisation  $\Sigma_1 := \Sigma_0 \circ \gamma_1$  for some map  $\gamma_1: U_1 \rightarrow U_0$ , where  $U_1 \subseteq \mathbb{R}^{d_1}$ . It therefore makes sense to consider a  $d_2$ -dimensional submanifold  $S_2$  of  $S_1$ , which itself is a submanifold of  $\mathbb{R}^n$  with parametrisation  $\Sigma_2 := \Sigma_1 \circ \gamma_2 = \Sigma_0 \circ \gamma_1 \circ \gamma_2$  for some map  $\gamma_2: U_2 \rightarrow U_1$ , where  $U_2 \subseteq \mathbb{R}^{d_2}$ . Continuing in this way, we can construct a whole nested sequence of submanifolds.

To describe the setup precisely, we say  $\mathcal{S} = (S_0, (S_\ell)_{\ell=1}^r)$  is a *nested family of submanifolds of  $\mathbb{R}^n$*  if  $S_0 = \Sigma_0(U_0) \subseteq \mathbb{R}^n$  is a submanifold of  $\mathbb{R}^n$  and  $S_\ell$  is a submanifold of  $S_{\ell-1}$  for  $1 \leq \ell \leq r$ . Thus, if  $d_\ell = \dim(S_\ell)$  for  $1 \leq \ell \leq r$ , then we have a family of open neighbourhoods  $U_\ell \subseteq \mathbb{R}^{d_\ell}$  of the origin and parametrisations  $\gamma_\ell: U_\ell \rightarrow U_{\ell-1}$  with  $\gamma_\ell(U_\ell)$  compactly contained in  $U_{\ell-1}$  and  $\gamma_\ell(0) = 0$  such that

$$(1.8) \quad S_\ell = \Sigma_\ell(U_\ell), \quad \text{where } \Sigma_\ell := \Sigma_0 \circ \sigma_\ell \text{ for } \sigma_\ell := \gamma_1 \circ \cdots \circ \gamma_\ell.$$

In this case, each set  $M_\ell := \sigma_\ell(U_\ell)$  is a  $d_\ell$ -dimensional submanifold of  $\mathbb{R}^{d_0}$  and  $S_\ell = \Sigma_0(M_\ell)$ . Given  $\mu = (\mu_\ell)_{\ell=1}^r \subset (0, 1]$ , we shall say  $f \in L^2(S)$  is *supported in  $\mathcal{N}_\mu \mathcal{S}$* , or  $\text{supp } f \subseteq \mathcal{N}_\mu \mathcal{S}$ , if  $\text{supp } f \subseteq \mathcal{N}_{\mu_\ell} S_\ell$  for all  $1 \leq \ell \leq r$ , in the sense defined above: that is,  $\text{supp } f \subseteq \mathcal{N}_{\mu_\ell} M_\ell$ , where  $\mathcal{N}_{\mu_\ell} M_\ell \subseteq \mathbb{R}^{d_\ell}$  is the (open Euclidean)  $\mu_\ell$ -neighbourhood of  $M_\ell$ . We also assume that  $d_r < \cdots < d_1 < d_0$  since this is the configuration of interest for our main theorem; in particular,  $r \leq n$ .

**Definition 1.5.** Fix  $2 \leq k \leq n$  and suppose  $\mathcal{E} = ((\mathcal{S}_j)_{j=1}^k, \mathbf{q})$ , where the  $\mathcal{S}_j = (S_{j,0}, (S_{j,\ell})_{\ell=1}^{r_j})$  are nested families of submanifolds and  $\mathbf{q} = (q_j)_{j=1}^k \subset (0, 2]$  is a sequence of exponents. Define the Brascamp–Lieb datum

$$(1.9) \quad (\mathbf{L}(\mathcal{E}), \mathbf{p}) = ((L_j(\mathcal{E}))_{j=1}^k, (p_j)_{j=1}^k)$$

by setting

$$L_j(\mathcal{E}) := \partial_s \Sigma_{j,r_j}(0)^\top: \mathbb{R}^n \rightarrow \mathbb{R}^{d_{j,r_j}} \quad \text{and} \quad p_j := q_j/2 \quad \text{for } 1 \leq j \leq k,$$

where  $d_{j,r_j} := \dim(S_{j,r_j})$  and  $S_{j,r_j} = \Sigma_{j,r_j}(U_{j,r_j})$ , with  $\Sigma_{j,r_j}$  parametrisations constructed as in (1.8). Here and in the sequel we use the symbol  $\partial_s$  to denote the Jacobian with respect to the variable  $s$ , and given a matrix  $A$ , we denote its transpose by  $A^\top$ .

(a) We say  $\mathcal{E}$  is a *transverse ensemble* in  $\mathbb{R}^n$  if

$$\text{BL}_{\text{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p}) < \infty.$$

(b) We say  $\mathcal{E}$  has *codimension  $(m(j, \ell))_{j,\ell}$* , where  $m(j, \ell) := \text{codim}(S_{j,\ell} | S_{j,\ell-1})$  for  $1 \leq \ell \leq r_j$  and  $1 \leq j \leq k$ .

(c) Finally, we say  $(\mu_j)_{j=1}^k$ , where  $\mu_j = (\mu_{j,\ell})_{\ell=1}^{r_j} \subset (0, 1]$  is a *compatible family of scales* for  $\mathcal{E}$  if  $0 < \mu_{j,1} \leq \cdots \leq \mu_{j,r_j} < 1$  for  $1 \leq j \leq k$ .

With the above definitions, the main theorem reads as follows.

**Theorem 1.6.** Let  $2 \leq k \leq n$  and  $\mathcal{E} = ((\mathcal{S}_j)_{j=1}^k, \mathbf{q})$  be a transverse ensemble in  $\mathbb{R}^n$  of codimension  $(m(j, \ell))_{j,\ell}$ , with  $\mathcal{S}_j = (S_j, (S_{j,\ell})_{\ell=1}^{r_j})$ . There exists  $\rho_{\mathcal{E}} > 0$  such that for  $\max_j \rho(E_{S_j}) < \rho_{\mathcal{E}}$  the following holds. For all  $\varepsilon > 0$ , there exists a constant  $C_{\mathcal{E},\varepsilon} \geq 1$  such that if  $(\mu_j)_{j=1}^k$  is any compatible family of scales for  $\mathcal{E}$ , then the inequality

$$(1.10) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_j} f_j|^{q_j} \leq C_{\mathcal{E},\varepsilon} R^\varepsilon \prod_{j=1}^k \prod_{\ell=1}^{r_j} \mu_{j,\ell}^{m(j,\ell)q_j/2} \prod_{j=1}^k \|f_j\|_{L^2(S_j)}^{q_j}$$

holds for all  $R \geq 1$  and all  $f_j \in L^2(S_j)$  satisfying  $\text{supp } f_j \subseteq \mathcal{N}_{\mu_j} \mathcal{S}_j$ .

Recall that we are tacitly working with a fixed choice of parametrisations for the  $S_j$  and  $S_{j,\ell}$ , and the constant  $C_\varepsilon$  will depend (amongst other things) on this choice. If  $\Sigma_j: U_j \rightarrow \mathbb{R}^n$  and  $\Sigma_{j,r_j}: U_{j,r_j} \rightarrow \mathbb{R}^n$  are our fixed parametrisations for  $S_j$  and  $S_{j,r_j}$  as submanifolds of  $\mathbb{R}^n$ , then, as discussed in (1.8) and (1.3), we always assume  $\Sigma_{j,r_j} := \Sigma_j \circ \gamma_{j,1} \circ \cdots \circ \gamma_{j,r_j}$  for intermediate parametrisations  $\gamma_{j,\ell}$ , and  $\Sigma_{j,r_j}(0) = \Sigma_j(0)$ .

*Remark 1.7.* To contextualise the result, we make the following remarks.

(a) Let  $2 \leq k < n$ . A prototypical application of our theorem is a  $k$ -linear estimate for the  $(n-1)$ -dimensional paraboloid in which one of the functions is localised to an anisotropic box. More precisely, let  $U := (-1, 1)^{n-1}$  and  $S := \Sigma(U)$ , where  $\Sigma(u) := (u, |u|^2)$ . Let  $0 < \delta \ll 1$  and  $B_\delta \subseteq \mathbb{R}^{n-1}$  be the  $1 \times \cdots \times 1 \times \delta \times \delta^2 \times \cdots \times \delta^{n-k}$  box centred at the origin pointing along the directions of the standard basis in  $\mathbb{R}^{n-1}$ . Let  $u_1, \dots, u_{k-1} \in U$  satisfy

$$(1.11) \quad \left| \bigwedge_{j=1}^{k-1} \frac{(2u_j, -1)}{|(2u_j, -1)|} \wedge e_k \wedge \cdots \wedge e_n \right| \geq 1/2.$$

If  $U_1, \dots, U_{k-1}$  are sufficiently small neighbourhoods of the points  $u_1, \dots, u_{k-1}$  and  $U_k$  is a sufficiently small neighbourhood of the origin, then for all  $\varepsilon > 0$  there exists a constant  $C_\varepsilon \geq 1$  such that

$$(1.12) \quad \int_{Q_R} \prod_{j=1}^k |E_S f_j|^{\frac{2}{k-1}} \leq C_\varepsilon R^\varepsilon \delta^{\frac{(n-k)(n-k+1)}{2(k-1)}} \prod_{j=1}^k \|f_j\|_{L^2(U_j)}^{\frac{2}{k-1}}$$

for all  $R \geq 1$  and all  $f_j \in L^2(U_j)$  for  $1 \leq j \leq k$  with  $\text{supp } f_k \subseteq B_\delta$ .

To see how this follows from Theorem 1.6, define  $U_{k,\ell} := (-1, 1)^{n-1-\ell}$  and  $\gamma_{k,\ell}: U_{k,\ell} \rightarrow U_{k,\ell-1}$  by  $\gamma_{k,\ell}(s) := (s, 0)$  for  $1 \leq \ell \leq n-k$ , with  $U_{k,0} := U$ . It is then clear that  $\text{supp } f_k \subseteq \mathcal{N}_{\mu_k} \mathcal{S}_k$ , where  $\mu_k = (\delta^{n-k+1-\ell})_{\ell=1}^{n-k}$  and  $\mathcal{S}_k = (S, (S_{k,\ell})_{\ell=1}^{n-k})$  for  $S_{k,\ell} = \Sigma_{k,\ell}(U_{k,\ell})$  and  $\Sigma_{k,\ell} := \Sigma \circ \gamma_{k,1} \circ \cdots \circ \gamma_{k,\ell}$ . For all  $1 \leq j \leq k-1$ , one can centre  $U_j$  at the origin by a change of variables, giving rise to a neighbourhood  $U_j^*$  and an extension operator  $E_{S_j}$  with  $S_j = \Sigma_j(U_j^*)$  and  $\Sigma_j(u) = (u + u_j, |u + u_j|^2)$ . Considering the ensemble  $\mathcal{E} = ((\mathcal{S}_j)_{j=1}^k, \mathbf{q}_k)$ , where  $\mathcal{S}_j = (S_j)$  for  $1 \leq j \leq k-1$  and  $\mathbf{q}_k = (q_k)_{j=1}^k = (q_k, \dots, q_k)$  for  $q_k = \frac{2}{k-1}$ , we have that  $\ker L_j(\mathcal{E}) = (2u_j, -1)^\top$  for all  $1 \leq j \leq k-1$  and  $\ker L_k(\mathcal{E}) = \text{span}\{e_k, \dots, e_n\}$ . One can show that (1.11) implies that the condition (1.7) holds with  $p_j = \frac{1}{k-1}$  (see, for instance, Corollary 2.3), and thus  $\mathcal{E}$  is a transverse ensemble. Consequently, the claimed multilinear inequality (1.12) follows from Theorem 1.6.

(b) Taking  $\mathcal{S}_j = (S_j, S'_j)$  for  $S_j \subset \mathbb{R}^n$  a hypersurface and  $S'_j$  a submanifold of  $S_j$ , for  $1 \leq j \leq k$ , and  $\mathbf{p}_k = (p_k)_{j=1}^k = (p_k, \dots, p_k)$  for  $p_k = \frac{1}{k-1}$ , we recover the localised multilinear estimates of Theorem 1.4. Indeed, this follows from the observation

$$\left| \bigwedge_{j=1}^k N_{\xi'_j} S'_j \right| > 0 \iff \text{BL}_{\text{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p}_k) < \infty,$$

where the  $\xi'_j$  are as defined in (1.4); see §2 for details.

(c) Suppose  $r_j = 0$  for all  $1 \leq j \leq k$ , so that  $\mathcal{S}_j = (S_j)$  and there is no nesting. If we further suppose that the scaling condition  $\sum_{j=1}^k p_j \dim(S_j) = n$  is satisfied, then Theorem 1.6 amounts to the restriction-Brascamp–Lieb estimates of [6, Theorem 1.3]. If we drop the scaling condition, then the corresponding estimates are not documented

in the literature, but can be deduced from the Keakeya–Brascamp–Lieb estimates in [14, 20] (see Theorem 2.6 below) by the argument in [6, 9].

(d) As discussed in §2.3, the hypothesis  $\text{BL}_{\text{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p}) < \infty$  of Theorem 1.6 implies

$$(1.13) \quad \text{BL}_{\text{reg}}((\partial_s \Sigma_j(0)^\top)_{j=1}^k, \mathbf{p}) < \infty.$$

By the argument of [6] (cf. item (c) above), the condition (1.13) implies the multilinear estimates

$$\int_{Q_R} \prod_{j=1}^k |E_{S_j} f_j|^{q_j} \leq C_{\mathcal{E}, \varepsilon} R^\varepsilon \prod_{j=1}^k \|f_j\|_{L^2(S_j)}^{q_j}.$$

The relevance of our theorem is therefore the improvement given by the factor  $\prod_{j=1}^k \prod_{\ell=1}^{r_j} \mu_{j,\ell}^{m(j,\ell)q_j/2}$  under the hypotheses  $\text{supp } f_j \subseteq \mathcal{N}_{\mu_j} S_j$ . Since the constant  $C_{\mathcal{E}, \varepsilon}$  blows up as the left-hand side of (1.13) goes to  $\infty$ , one should think of this quantity as being  $O(1)$  (or small depending on some other admissible parameters unrelated to our choice of  $\mu_j$ ) in order to take advantage of the  $\mu_j$ -gains in the inequality (1.10).

(e) Finally, we note that the definition of transverse ensemble only requires the finiteness of the Brascamp–Lieb constant for the datum associated with the innermost submanifolds in the ensemble. However, this condition also implies the finiteness of the Brascamp–Lieb constant for data associated with any intermediate submanifolds; see Proposition 2.4. This plays a key rôle in our argument and comes into play most prominently in §4.2 and §4.3. In this sense, the nesting structure is an essential geometric feature of our proof.

**1.4. Proof sketch.** We describe the main ideas in the simpler setting of Theorem 1.4. We remark that, due to the local nature of the desired estimates, we may assume without loss of generality that  $\max_j \mu_j < \mu_{\mathcal{E}}$  for some  $\mu_{\mathcal{E}} > 0$  depending only on  $\mathcal{E}$ . Under the hypotheses of Theorem 1.4, it is possible to bound

$$(1.14) \quad |E_{S_j} f_j(x)| \leq \int_{[-\mu_j, \mu_j]^{m_j}} |E_{S'_j(\boldsymbol{\eta}_j)} f_{j, \boldsymbol{\eta}_j}(x)| \, d\boldsymbol{\eta}_j,$$

where each  $E_{S'_j(\boldsymbol{\eta}_j)}$  is an extension operator associated to a submanifold  $S'_j(\boldsymbol{\eta}_j)$  of  $S_j$  of codimension  $m_j$ ; see §3.2 for details. Moreover, (1.4) and continuity imply that, for  $\mu_j$  sufficiently small, the submanifolds  $(S'_j(\boldsymbol{\eta}_j))_{\ell=1}^k$  satisfy certain transversality conditions uniformly over all choices of  $\boldsymbol{\eta}_j \in [-\mu_j, \mu_j]^{m_j}$ . For each index  $j$ , we plug the bound (1.14) into the left-hand side of (1.5). If one were then able to switch the order of the  $L_x^{q_k}$ -norm and the  $L_{\boldsymbol{\eta}_j}^1$ -norms, for instance using Minkowski’s integral inequality, then the multilinear restriction estimates of [6] (or certain regularised variants) would apply directly to give, together with an application of the Cauchy–Schwarz inequality, Theorem 1.4. However, Minkowski’s integral inequality requires  $q_k \geq 1$ , which is only valid for  $k = 2$  or  $k = 3$ . See [4, Appendix A] for an example of this strategy in the  $k = 3$  case.

A similar issue with the exponent  $q_k < 1$  arises when passing between the equivalent ‘neighbourhood’ and ‘extension’ formulations of the Bennett–Carbery–Tao restriction theorem in [9]. One strategy to deal with this is to use a vector-valued extension of the multilinear restriction estimates, derived from the Marcinkiewicz–Zygmund theorem; see [17, Appendix] for details. The present situation is somewhat more complicated than that in [9, 17], since the operators  $E_{S'_j(\boldsymbol{\eta}_j)}$  typically have a non-trivial dependence on  $\boldsymbol{\eta}_j$  (in the context of [9, 17], the analogous extension operators are independent of  $\boldsymbol{\eta}_j$ ). This means that the Marcinkiewicz–Zygmund theorem cannot be applied directly to obtain the desired vector-valued extension.

The above strategy does work, however, when  $\mu_j < R^{-1}$ . Since the  $x$ -integration is localised to  $Q_R$ , in this case uncertainty principle considerations ensure that  $E_{S'_j(\eta_j)} f_{j,\eta_j}$  is essentially constant in  $\eta_j \in [-\mu_j, \mu_j]^{m_j}$ . This allows for an application of the Marcinkiewicz–Zygmund theorem as in [17, Appendix]; see Lemma 4.1 for details. On the other hand, when  $\mu_j \geq R^{-1}$  it is possible to apply Kekeya–Brascamp–Lieb inequalities to reduce the scale  $R$ , following an induction-on-scales procedure similar to that used in [6, 9] (but adapted to take into account the localisation of the supports of the  $f_j$ ). Combining these two observations, one may formulate a refined induction-on-scales procedure, which takes into account the relative sizes of the  $\mu_j$  and  $R^{-1}$ ; see Lemma 4.2. As in [6, 9], the induction-on-scales procedure yields the desired estimate.

**1.5. Notational conventions.** We shall say a constant is *admissible* if it depends either only on a choice of dimension  $n$  or a choice of ensemble  $\mathcal{E}$ . In particular, any admissible constant is independent of parameters such as  $R \geq 1$  or the  $\mu_{j,\ell} > 0$ . We will frequently let  $C$  denote a choice of admissible constant, whose value may change from line to line. Given a list of objects  $L$  and real numbers  $A, B \geq 0$ , we write  $A \lesssim_L B$  or  $B \gtrsim_L A$  to indicate  $A \leq C_L B$  for some constant  $C_L$  which depends only on items in the list  $L$  and either a choice of dimension  $n$  or a choice of ensemble  $\mathcal{E}$ . We write  $A \sim_L B$  to indicate  $A \lesssim_L B$  and  $B \lesssim_L A$ . Given  $d \in \mathbb{N}$  we let  $I_d$  denote the  $d \times d$  identity matrix, and given  $m, n \in \mathbb{N}$ , we denote the transpose of a matrix  $M \in \text{Mat}(\mathbb{R}, n \times m)$  by  $M^\top$ , and the inverse of its transpose by  $M^{-\top} \equiv (M^\top)^{-1}$ .

## 2. Kekeya–Brascamp–Lieb theory

**2.1. Regularised Brascamp–Lieb revisited.** Let  $2 \leq k \leq n$ ,  $1 \leq n_j \leq n$  for  $1 \leq j \leq k$ , and  $(\mathbf{L}, \mathbf{p})$  a Brascamp–Lieb datum as in §1.2. Given  $R \geq 1$  and  $0 \leq \lambda < R$  dyadic, we let  $\text{BL}_{\lambda,R}(\mathbf{L}, \mathbf{p})$  denote the best constant  $C > 0$  for which the associated generalised Brascamp–Lieb inequality

$$(2.1) \quad \int_{Q_R} \prod_{j=1}^k (f_j \circ L_j)^{p_j} \leq C \prod_{j=1}^k \left( \int_{\mathbb{R}^{n_j}} f_j \right)^{p_j}$$

holds for all non-negative functions  $f_j \in L^1(\mathbb{R}^{n_j})$  that are constant on cubes of side length  $\lambda$  in the fixed dyadic grid  $[0, \lambda)^{n_j} + \lambda \mathbb{Z}^{n_j}$ . Inequalities of this form were first introduced in [7] and can be seen as a generalisation of the classical scale-invariant Brascamp–Lieb inequalities, which correspond to taking  $\lambda \rightarrow 0$ ,  $R \rightarrow \infty$ . Regarding the latter, we let  $\text{BL}(\mathbf{L}, \mathbf{p})$  denote the classical Brascamp–Lieb constant, which is the best constant  $C > 0$  for which (2.1) holds for all  $R \geq 1$  and all non-negative functions  $f_j \in L^1(\mathbb{R}^{n_j})$  (without any local constancy hypothesis).

The case  $\lambda = 1$  of the above is often referred to as *regularised*. Note that if  $0 < \lambda \leq R < \infty$ , one can always reduce to the case  $\lambda = 1$  by the scaling relation

$$\text{BL}_{\lambda,R}(\mathbf{L}, \mathbf{p}) = \lambda^{n - \sum_{j=1}^k p_j n_j} \text{BL}_{1, R\lambda^{-1}}(\mathbf{L}, \mathbf{p}).$$

Maldague ([14, Theorem 1.1]) recently showed that

$$(2.2) \quad \text{BL}_{1,R}(\mathbf{L}, \mathbf{p}) \sim_{\mathbf{L}, \mathbf{p}} R^{\alpha(\mathbf{L}, \mathbf{p})} \quad \text{for} \quad \alpha(\mathbf{L}, \mathbf{p}) := \sup_{V \leq \mathbb{R}^n} \left( \dim(V) - \sum_{j=1}^k p_j \dim(L_j V) \right).$$

Note that the constant  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p})$  introduced in §1.2 satisfies

$$\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}) = \sup_{R > 1} \text{BL}_{1,R}(\mathbf{L}, \mathbf{p});$$

in particular, one can recover (1.7) from the characterisation (2.2) for  $\text{BL}_{1,R}(\mathbf{L}, \mathbf{p})$ .

We remark that the classical Brascamp–Lieb constant  $\text{BL}(\mathbf{L}, \mathbf{p})$  is finite if and only if  $\alpha(\mathbf{L}, \mathbf{p}) = 0$  and  $\sum_{j=1}^k p_j n_j = n$ ; see [7]. In this case, we in fact have

$$\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}) = \text{BL}(\mathbf{L}, \mathbf{p})$$

as a consequence of a simple limiting argument.

**2.2. Regularised Brascamp–Lieb and transversality conditions.** Here we compare the hypotheses of Theorem 1.6 and Theorem 1.4 and, in particular, address the comments made in Remark 1.7(b).

**Proposition 2.1.** *Let  $\mathbf{p}_k = (p_k)_{j=1}^k = (p_k, \dots, p_k)$ , where  $p_k := q_k/2 = \frac{1}{k-1}$ , and suppose that  $L_j: \mathbb{R}^n \rightarrow \mathbb{R}^{n_j}$ ,  $1 \leq j \leq k$ , are linear surjections. Then the following are equivalent:*

$$(a) \text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}_k) < \infty;$$

$$(b) \left| \bigwedge_{j=1}^k \ker(L_j) \right| > 0.$$

*Proof:* For  $V \leq \mathbb{R}^n$ , we first note that

$$(2.3) \quad \dim(L_j(V)) = \dim(V) - \dim(V \cap \ker(L_j)),$$

by the rank-nullity theorem applied to the restriction  $L_j|_V$ ,  $1 \leq j \leq k$ . We define the subspace  $W := \ker(L_1) + \dots + \ker(L_k)$ .

(a)  $\Rightarrow$  (b). Assume (b) fails. Applying (2.3) to  $W$  as defined above, we deduce that

$$(2.4) \quad \dim(L_j(W)) = \dim(W) - \dim(\ker(L_j)),$$

since  $\ker(L_j) \leq W$  for all  $1 \leq j \leq k$ . The failure of (b) implies that  $\dim(W) < \sum_{j=1}^k \dim(\ker(L_j))$  and applying this to (2.4) yields

$$\sum_{j=1}^k \dim(L_j(W)) = k \dim(W) - \sum_{j=1}^k \dim(\ker(L_j)) < (k-1) \dim(W).$$

This means that  $W$  violates the finiteness characterisation (1.7) for the datum  $(\mathbf{L}, \mathbf{p}_k)$ , implying that (a) fails.

(b)  $\Rightarrow$  (a). Assume that (b) holds, so that  $W = \bigoplus_{j=1}^k \ker(L_j)$ . Given  $V \leq \mathbb{R}^n$ , we have  $\bigoplus_{j=1}^k (V \cap \ker(L_j)) \leq V \cap W$  and therefore

$$\sum_{j=1}^k \dim(V \cap \ker(L_j)) = \dim\left(\bigoplus_{j=1}^k V \cap \ker(L_j)\right) \leq \dim(V \cap W).$$

Combining the above inequality with (2.3), we obtain

$$\begin{aligned} \sum_{j=1}^k \dim(L_j V) &= \sum_{j=1}^k (\dim(V) - \dim(V \cap \ker(L_j))) \\ &\geq k \dim(V) - \dim(V \cap W) \\ &\geq (k-1) \dim(V). \end{aligned}$$

Since  $V \leq \mathbb{R}^n$  was arbitrary, the characterisation (1.7) implies that (a) holds.  $\square$



*Remark 2.2.* Suppose the maps  $L_j$  are orthogonal projections. By invoking [5, Proposition 1.2], one may deduce that

$$\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}, \mathbf{p}_k) \sim \left| \bigwedge_{j=1}^k \ker(L_j) \right|^{-\frac{1}{k-1}},$$

which is a quantitative variant of Proposition 2.1. Since in this paper we only need the finiteness of such transversality quantities, we leave the proof of this more quantitative result to the interested reader.

**Corollary 2.3.** *Let  $2 \leq k \leq n$ ,  $\mathcal{E} = ((S_j, S'_j)_{j=1}^k, \mathbf{q}_k)$  be a transverse ensemble in  $\mathbb{R}^n$ . Then*

$$(2.5) \quad \mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p}_k) < \infty \iff \left| \bigwedge_{j=1}^k N_{\xi'_j} S'_j \right| > 0.$$

Here the  $\xi'_j$  are as defined in (1.4). In particular, Theorem 1.6 implies Theorem 1.4.

*Proof:* For each  $1 \leq j \leq k$ , let  $\Sigma_j: U_j \rightarrow \mathbb{R}^n$  be the parametrisation of  $S_j$ , and let  $M_j \subset U_j$  be a submanifold of  $\mathbb{R}^{\dim(S_j)}$ , parametrised by  $\gamma_j: U'_j \rightarrow U_j$ , such that  $\Sigma_j(M_j) = S'_j$ . Define  $\Sigma'_j := \Sigma_j \circ \gamma_j: U'_j \rightarrow \mathbb{R}^n$ . Then, by the definition of  $\mathbf{L}(\mathcal{E})$  we have that  $L_j(\mathcal{E}) = \partial_s \Sigma'_j(0)^\top$ . Since  $\ker(\partial_s \Sigma'_j(0)^\top) = N_{\xi'_j} S'_j$ , we then see that the equivalence (2.5) is merely a rephrasing of Proposition 2.1 for this particular choice of Brascamp–Lieb datum.  $\square$

**2.3. Brascamp–Lieb constants and subensembles.** Fix  $2 \leq k \leq n$  and let  $\mathcal{E} = ((S_j)_{j=1}^k, \mathbf{q})$  be a transverse ensemble in the sense of Definition 1.5, with  $S_j = (S_j, (S_{j,\ell})_{\ell=1}^{r_j})$  and  $\mathbf{q} = (q_j)_{j=1}^k \subset (0, 2]$ . We say that  $\mathcal{F} = ((\tilde{S}_j)_{j=1}^k, \mathbf{q})$  is a *subensemble* of  $\mathcal{E}$  if  $\tilde{S}_j = (S_j, (S_{j,\ell})_{\ell=1}^{\tilde{r}_j})$  with  $\tilde{r}_j \leq r_j$ . Recall that we allow values  $\tilde{r}_j = 0$ , in which case  $\tilde{S}_j$  consists only of  $S_j$ .

**Proposition 2.4.** *Let  $\mathcal{E}$  be a transverse ensemble in  $\mathbb{R}^n$ . Then any subensemble  $\mathcal{F}$  of  $\mathcal{E}$  is also a transverse ensemble.*

*Proof:* It suffices to verify that  $\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{F}), \mathbf{p}) < \infty$  with  $\mathbf{L}(\mathcal{F})$  as in (1.9). In view of (1.8), for each  $1 \leq j \leq k$  we have that

$$\Sigma_{j,r_j} = \Sigma_{j,\tilde{r}_j} \circ \gamma_{j,\tilde{r}_j,r_j} \quad \text{for} \quad \gamma_{j,\tilde{r}_j,r_j} := \gamma_{j,\tilde{r}_j+1} \circ \cdots \circ \gamma_{j,r_j};$$

here, for the case  $\tilde{r}_j = r_j$ , we simply have  $\gamma_{j,r_j,r_j} := \mathrm{Id}$ . Thus,

$$L_j(\mathcal{E}) = \partial_s \Sigma_{j,r_j}(0)^\top = (\partial_s \gamma_{j,\tilde{r}_j,r_j}(0))^\top (\partial_s \Sigma_{j,\tilde{r}_j}(0))^\top = (\partial_s \gamma_{j,\tilde{r}_j,r_j}(0))^\top L_j(\mathcal{F}),$$

where we have used that  $\gamma_{j,\ell}(0) = 0$  for all  $1 \leq \ell \leq r_j$ . It is then clear that  $\dim(L_j(\mathcal{E})V) \leq \dim(L_j(\mathcal{F})V)$  for any  $V \leq \mathbb{R}^n$  and, in view of the characterisation (1.7) for  $\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}, \mathbf{p})$ , we conclude that the finiteness of  $\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p})$  implies that of  $\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{F}), \mathbf{p})$ .  $\square$

*Remark 2.5.* Similarly to Remark 2.2, if the linear surjections  $L_j(\mathcal{E})$  and  $L_j(\mathcal{F})$  are orthogonal projections for  $1 \leq j \leq k$ , then one can show that  $\mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{F}), \mathbf{p}) \lesssim \mathrm{BL}_{\mathrm{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p})$ , where the implicit constant does not depend on the choice of ensemble. This refined quantitative result is not required for our purposes, and so we leave the proof to the interested reader.

**2.4. Kakeya–Brascamp–Lieb inequalities.** Here we are interested in a class of *Kakeya–Brascamp–Lieb inequalities*, which can be thought of as perturbed variants of the regularised Brascamp–Lieb inequalities introduced above. The following result is due to Maldague [14].

**Theorem 2.6** (Kakeya–Brascamp–Lieb [14, Theorem 1.2]). *Let  $(\mathbf{L}, \mathbf{p})$  be a Brascamp–Lieb datum such that  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}) < \infty$ . Then there exists  $\nu > 0$  such that the following holds. For all  $\varepsilon > 0$ , there exists a constant  $C_\varepsilon \geq 1$  such that the inequality*

$$\int_{Q_R} \prod_{j=1}^k \left| \sum_{T_j \in \mathbb{T}_j} c_{T_j} \chi_{T_j} \right|^{p_j} \leq C_\varepsilon R^\varepsilon \lambda^{n-\varepsilon} \prod_{j=1}^k \left( \sum_{T_j \in \mathbb{T}_j} |c_{T_j}| \right)^{p_j}$$

*holds for all  $R \geq 1$  and  $0 < \lambda < R$ , whenever the  $\mathbb{T}_j$  are countable collections of infinite slabs of width  $\lambda$  whose core  $n_j$ -dimensional planes are, modulo translations, within a distance  $\nu > 0$  (in the Grassmannian sense) from the fixed subspace  $X_j := \ker L_j$ , and  $(c_{T_j})_{T_j \in \mathbb{T}_j} \in \ell^1(\mathbb{T}_j)$  for  $1 \leq j \leq k$ .*

We let  $|x|_\infty := \max_{1 \leq \ell \leq m} |x_\ell|$  for  $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ , and we say  $T \subseteq \mathbb{R}^n$  is an infinite slab of width  $r > 0$  if

$$T = \{x \in \mathbb{R}^n : |Lx - v|_\infty < r\}$$

for some linear surjective map  $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$  and  $v \in \mathbb{R}^m$ . In this case, the affine subspace  $V := \{x \in \mathbb{R}^n : Lx = v\}$  is called the *core plane* of  $T$ . We note that endpoint variants of Theorem 2.6 can be found in [19, 20]. The case in which  $\dim \ker L_j = 1$  for all  $1 \leq j \leq k$  corresponds to the  $k$ -linear Kakeya inequality from [9, Theorem 5.1].

We remark that there are some differences between the statement of Theorem 2.6 and that of [14, Theorem 1.2].

- Theorem 2.6 is a rescaled version of [14, Theorem 1.2]. The parameter  $\delta$  in [14, Theorem 1.2] corresponds to  $\lambda/R$  in Theorem 2.6.
- The coefficients  $(c_{T_j})_{T_j \in \mathbb{T}_j}$  are taken to all be 1 in [14, Theorem 1.2]. However, the result for general coefficients follows from a standard approximation argument; see [9].
- Theorem 2.6 is stated in terms of  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p})$  whilst [14, Theorem 1.2] is stated in terms of the exponent

$$\tilde{\alpha}(\mathbf{L}, \mathbf{p}) := \sup_{V \leq \mathbb{R}^n} \left( \dim(V) - \sum_{j=1}^k p_j \dim(\pi_j V) \right),$$

where  $\pi_j: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is the orthogonal projection onto the subspaces  $(X_j)^\perp$  for  $1 \leq j \leq n$ . To relate the former to the latter, one observes that the hypothesis  $\text{BL}_{\text{reg}}(\mathbf{L}, \mathbf{p}) < \infty$  implies  $\tilde{\alpha}(\mathbf{L}, \mathbf{p}) = 0$ . Indeed,  $\ker \pi_j = X_j = \ker L_j$  and, given any  $V \leq \mathbb{R}^n$ , it follows from rank-nullity as in (2.3) that  $\dim \pi_j(V) = \dim L_j(V)$  for all  $1 \leq j \leq n$ . Consequently,  $\tilde{\alpha}(\mathbf{L}, \mathbf{p}) = \alpha(\mathbf{L}, \mathbf{p})$  and the desired implication follows from the characterisation (2.2).

### 3. Fourier extension preliminaries

**3.1. Parametrising the neighbourhoods.** Let  $\mathcal{S} = (S_0, (S_\ell)_{\ell=1}^r)$  be a nested family of submanifolds of  $\mathbb{R}^n$ , with  $S_0 = \Sigma_0(U_0)$  a  $d_0$ -dimensional submanifold of  $\mathbb{R}^n$ . As in §1.3, we have a family of open neighbourhoods  $U_\ell \subseteq \mathbb{R}^{d_\ell}$  of the origin, with  $d_\ell = \dim(S_\ell)$  satisfying  $d_r < \dots < d_1 < d_0$ , and parametrisations  $\gamma_\ell: U_\ell \rightarrow U_{\ell-1}$  for  $1 \leq \ell \leq r$  satisfying  $\gamma_\ell(0) = 0$  such that

$$S_\ell = \Sigma_\ell(U_\ell), \quad \text{where } \Sigma_\ell := \Sigma_0 \circ \sigma_\ell \text{ for } \sigma_\ell := \gamma_1 \circ \dots \circ \gamma_\ell.$$

Each set  $M_\ell := \sigma_\ell(U_\ell)$  is a  $d_\ell$ -dimensional submanifold of  $\mathbb{R}^{d_0}$  and  $S_\ell = \Sigma_0(M_\ell)$ . We define

$$m_\ell := \text{codim}(S_\ell|S_{\ell-1}) = d_{\ell-1} - d_\ell \quad \text{and} \quad c_\ell := \text{codim}(S_\ell|S_0) = d_0 - d_\ell$$

for  $1 \leq \ell \leq r$ . Often we drop the subscript 0 and write  $S := S_0$ ,  $d := d_0$ ,  $\Sigma := \Sigma_0$ , and  $U := U_0$ .

We begin with some (unfortunately rather technical) definitions which allow us to work inductively with the nested framework of Theorem 1.6. With the above setup, we may define

$$M_{k,\ell} := \gamma_{k,\ell}(U_\ell) \subseteq U_k, \quad \text{where } \gamma_{k,\ell} := \gamma_{k+1} \circ \cdots \circ \gamma_\ell \text{ for } k+1 \leq \ell \leq r,$$

so that each  $M_{k,\ell}$  is a smooth  $d_\ell$ -dimensional submanifold of  $\mathbb{R}^{d_k}$ . Recall that the support conditions in Theorem 1.6 are defined in terms of neighbourhoods of the submanifolds  $M_\ell$ , and note that  $M_\ell = M_{0,\ell}$  with the above notation. In order to argue inductively, we shall work more generally with neighbourhoods of the  $M_{k,\ell}$  for  $0 \leq k < \ell$ .

To describe the neighbourhoods of the  $M_{k,\ell}$ , we introduce a family of parametrising maps  $\Phi_{k,\ell}$ . Given  $1 \leq \ell \leq r$ , we let  $G_\ell: U_\ell \rightarrow \text{Mat}(\mathbb{R}, d_{\ell-1} \times m_\ell)$  denote a  $(d_{\ell-1} \times m_\ell)$ -matrix-valued function such that the columns of  $G_\ell(s_\ell)$  form an orthogonal basis for  $N_{\gamma_\ell(s_\ell)} M_{\ell-1,\ell} \subset U_{\ell-1}$ . Let  $\bar{\mu}_0 > 0$  be an admissible parameter, chosen sufficiently small for the purposes of the forthcoming argument. Given  $1 \leq k < \ell \leq r$ , throughout the following we write

$$V_{k,\ell} := \prod_{t=k+1}^{\ell} (-\bar{\mu}_0, \bar{\mu}_0)^{m_t} \quad \text{and} \quad \boldsymbol{\eta}_{k,\ell} = (\eta_{k+1}, \dots, \eta_\ell) \text{ for } \boldsymbol{\eta}_{k,\ell} \in V_{k,\ell}.$$

For fixed  $1 \leq \ell \leq r$ , we recursively define maps  $\Phi_{k,\ell}: U_\ell \times V_{k,\ell} \rightarrow U_k$  for  $0 \leq k \leq \ell$ , where we interpret  $U_\ell \times V_{\ell,\ell}$  as  $U_\ell$ , by setting  $\Phi_{\ell,\ell}: U_\ell \rightarrow U_\ell$  to be the identity map and then defining

$$(3.1) \quad \Phi_{k-1,\ell}(s_\ell; \boldsymbol{\eta}_{k-1,\ell}) := \gamma_k \circ \Phi_{k,\ell}(s_\ell; \boldsymbol{\eta}_{k,\ell}) + G_k \circ \Phi_{k,\ell}(s_\ell; \boldsymbol{\eta}_{k,\ell}) \eta_k^\top$$

for  $1 \leq k \leq \ell$ , where here  $\boldsymbol{\eta}_{k-1,\ell} = (\eta_k, \boldsymbol{\eta}_{k,\ell})$  and  $\boldsymbol{\eta}_{\ell,\ell}$  is a null variable. Since each set  $\gamma_k(U_k)$  is, by hypothesis, compactly contained in  $U_{k-1}$ , provided  $\bar{\mu}_0 > 0$  is chosen sufficiently small, the maps  $\Phi_{k,\ell}$  are well defined and indeed map into  $U_k$ . By a simple induction argument,

$$(3.2) \quad \Phi_{k,\ell}(s_\ell; \mathbf{0}) = \gamma_{k,\ell}(s_\ell),$$

where  $\gamma_{\ell,\ell} := \text{Id}_{U_\ell}$ .

We may equivalently write (3.1) as

$$(3.3) \quad \Phi_{k-1,\ell}(s_\ell; \boldsymbol{\eta}_{k-1,\ell}) = \Phi_{k-1,k}(\Phi_{k,\ell}(s_\ell; \boldsymbol{\eta}_{k,\ell}); \eta_k).$$

Iterating the above observation, we shall derive the following formula.

**Lemma 3.1.** *For all  $0 \leq k \leq \ell \leq r$ ,  $s_r \in U_r$ , and  $\boldsymbol{\eta}_{k,r} \in V_{k,r}$  we have*

$$(3.4) \quad \Phi_{k,r}(s_r; \boldsymbol{\eta}_{k,r}) = \Phi_{k,\ell}(\Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r}); \boldsymbol{\eta}_{k,\ell}),$$

where we have written  $\boldsymbol{\eta}_{k,r} = (\boldsymbol{\eta}_{k,\ell}, \boldsymbol{\eta}_{\ell,r})$ .

*Proof:* The case when  $k = \ell = r$  is trivial. Now assume, by way of induction, that (3.4) holds for some  $1 \leq k \leq r$  and all  $\ell \in \mathbb{N}$  such that  $k \leq \ell \leq r$ .

We shall show that

$$(3.5) \quad \Phi_{k-1,r}(s_r; \boldsymbol{\eta}_{k-1,r}) = \Phi_{k-1,\ell}(\Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r}); \boldsymbol{\eta}_{k-1,\ell})$$

holds for all  $k-1 \leq \ell \leq r$ . If  $\ell = k-1$ , then this result is a trivial consequence of the definitions, so we assume  $k \leq \ell \leq r$ . By (3.3), we deduce that

$$(3.6) \quad \Phi_{k-1,r}(s_r; \boldsymbol{\eta}_{k-1,r}) = \Phi_{k-1,k}(\Phi_{k,r}(s_r; \boldsymbol{\eta}_{k,r}); \boldsymbol{\eta}_k).$$

Our induction hypothesis then implies that

$$(3.7) \quad \Phi_{k-1,r}(s_r; \boldsymbol{\eta}_{k-1,r}) = \Phi_{k-1,k}(\Phi_{k,\ell}(\Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r}); \boldsymbol{\eta}_{k,\ell}); \boldsymbol{\eta}_k).$$

Applying (3.3) again, here with  $s_\ell = \Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r})$ , then yields

$$(3.8) \quad \Phi_{k-1,k}(\Phi_{k,\ell}(\Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r}); \boldsymbol{\eta}_{k,\ell}); \boldsymbol{\eta}_k) = \Phi_{k-1,\ell}(\Phi_{\ell,r}(s_r; \boldsymbol{\eta}_{\ell,r}); \boldsymbol{\eta}_{k-1,\ell}).$$

Combining (3.6), (3.7), and (3.8), we obtain (3.5), which closes the induction.  $\square$

For each  $1 \leq k \leq r$ , the derivative  $\partial_{s,\eta_k} \Phi_{k-1,k}(s; \eta_k)|_{\eta_k=0} = [\partial_s \gamma_k(s) G_k(s)]$  has full rank over all of  $U_k$ . By the inverse function theorem, provided  $\bar{\mu}_o > 0$  is sufficiently small, there exist open neighbourhoods of the origin  $W_k \subseteq U_k$ ,  $1 \leq k \leq r$ , such that the restricted maps  $\Phi_{k-1,k}: W_k \times (-\bar{\mu}_o, \bar{\mu}_o)^{m_k} \rightarrow W_{k-1}$  are diffeomorphisms. Note that the sets  $W_j$  appear in *both* the domain and codomain of these restricted mappings. This ‘compatibility’ can be achieved by repeatedly pruning the sets arising from direct application of the inverse function theorem; we leave the details to the dedicated reader. By iterating (3.3), the same is true of the restrictions

$$(3.9) \quad \Phi_{k,\ell}: W_\ell \times V_{k,\ell} \rightarrow W_k, \quad 0 \leq k \leq \ell \leq r.$$

We later refer to these restricted mappings as the *diffeomorphism chain* for  $\mathcal{S}$ .

Let  $C_o \geq 1$  be an admissible constant, chosen sufficiently large for the forthcoming purposes of the argument, and define  $\mu_o := C_o^{-9/8} \bar{\mu}_o$ . We later refer to this quantity as the *threshold width* for  $\mathcal{S}$ . Given  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_\ell)$  with  $\max_{1 \leq \ell \leq r} \mu_\ell \leq C_o \mu_o$ , define the set

$$(3.10) \quad \Omega_{k,\ell}(\boldsymbol{\mu}) := \Phi_{k,\ell}(W_\ell \times P_{k,\ell}(\boldsymbol{\mu})) \subseteq W_k, \quad \text{where} \quad P_{k,\ell}(\boldsymbol{\mu}) := \prod_{t=k+1}^{\ell} [-C_o^{1/8} \mu_t, C_o^{1/8} \mu_t]^{m_t}.$$

The following lemma tells us that the  $\Phi_{k,\ell}$  can indeed be thought of as parametrising the (anisotropic) neighbourhoods of the submanifolds  $M_{k,\ell}$ .

**Lemma 3.2.** *Provided  $C_o \geq 1$  is chosen sufficiently large,*

$$(3.11) \quad \bigcap_{t=k+1}^{\ell} \mathcal{N}_{\mu_t} M_{k,t} \cap W_k \subseteq \Omega_{k,\ell}(\boldsymbol{\mu}) \subseteq \bigcap_{t=k+1}^{\ell} \mathcal{N}_{C_o \mu_t} M_{k,t} \cap W_k$$

*holds for  $1 \leq k \leq \ell \leq r$  and  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_\ell)$  such that  $0 < \mu_1 \leq \dots \leq \mu_r \leq C_o \mu_o$ .*

When  $k = \ell$ , the left- and right-hand sides of (3.11) are understood to equal  $W_\ell$ . For  $\rho > 0$ , it is convenient to write

$$(3.12) \quad \mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; \rho) := \bigcap_{t=k+1}^{\ell} \mathcal{N}_{\rho \mu_t} M_{k,t} \cap W_k,$$

so that the left- and right-hand sets in (3.11) correspond to  $\mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; 1)$  and  $\mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; C_o)$ .

*Proof of Lemma 3.2:* For the purpose of this proof, we shall rescale  $\mu$  and temporarily redefine  $P_{k,\ell}(\mu) := \prod_{t=k+1}^{\ell} [-\mu_t, \mu_t]^{m_t}$ . Let  $C \geq 1$  be a fixed constant, chosen to satisfy the forthcoming requirements of the proof. Fixing  $1 \leq \ell \leq r$ , we shall show that, under the modified definition,

$$(3.13) \quad \mathfrak{N}_{k,\ell}(\mu; \rho_k^{-1}) \subseteq \Omega_{k,\ell}(\mu) \subseteq \mathfrak{N}_{k,\ell}(\mu; \rho_k) \quad \text{for } \rho_k := C^{\ell-k}$$

holds for all  $1 \leq k \leq \ell$ , using induction on  $k$ . Once this is established, we may simply take  $C_o \geq C^{8n}$  and replace  $\mu$  with  $C_o^{1/8} \mu$  to conclude the desired result. The claim holds vacuously for  $k = \ell$ , which acts as the base case. Now let  $2 \leq k \leq \ell$  and suppose (3.13) holds for this value of  $k$ . It suffices to show

$$(3.14) \quad \mathfrak{N}_{k-1,\ell}(\mu; C^{-1}\rho_k^{-1}) \subseteq \Omega_{k-1,\ell}(\mu) \subseteq \mathfrak{N}_{k-1,\ell}(\mu; C\rho_k).$$

For  $\rho > 0$ , define the auxiliary sets

$$\tilde{\mathfrak{N}}_{k-1,\ell}(\mu; \rho) := M_{k-1,k} \cap \bigcap_{t=k+1}^{\ell} \mathcal{N}_{\rho\mu_t} M_{k-1,t} \cap W_{k-1}.$$

The proof of (3.14) can be reduced to showing

$$(3.15) \quad \tilde{\mathfrak{N}}_{k-1,\ell}(\mu; 2C^{-1}\rho_k^{-1}) \subseteq \gamma_k(\Omega_{k,\ell}(\mu)) \subseteq \tilde{\mathfrak{N}}_{k-1,\ell}(\mu; 2^{-1}C\rho_k).$$

We establish this reduction in two stages.

Assuming the first inclusion in (3.15), we show the first inclusion in (3.14). Let  $x \in \mathfrak{N}_{k-1,\ell}(\mu; C^{-1}\rho_k^{-1})$  and note, since  $x \in W_{k-1}$ , that there exists  $s \in W_{\ell}$  and  $\eta_{k-1,\ell} = (\eta_k, \eta_{k,\ell}) \in V_{k-1,\ell}$  such that

$$(3.16) \quad x = \Phi_{k-1,\ell}(s; \eta_{k-1,\ell}) = \gamma_k \circ \Phi_{k,\ell}(s; \eta_{k,\ell}) + G_k \circ \Phi_{k,\ell}(s; \eta_{k,\ell}) \eta_k^{\top},$$

where we have used the recursive definition (3.1). Observe that:

- $x_0 := \gamma_k \circ \Phi_{k,\ell}(s; \eta_{k,\ell}) \in M_{k-1,k}$  and  $x_0 = \Phi_{k-1,\ell}(s; (0, \eta_{k,\ell})) \in W_{k-1}$ ;
- $x \in \mathcal{N}_{C^{-1}\rho_k^{-1}\mu_t} M_{k-1,t}$  for  $k \leq t \leq \ell$ ;
- $x - x_0 = G_k \circ \Phi_{k,\ell}(s; \eta_{k,\ell}) \eta_k^{\top} \in N_{x_0} M_{k-1,k}$ .

We therefore conclude that  $|x - x_0| = |\eta_k| \leq C^{-1}\rho_k^{-1}\mu_k$ . By the ordering  $\mu_1 \leq \dots \leq \mu_r$  and the triangle inequality,  $x_0 \in \mathcal{N}_{2C^{-1}\rho_k^{-1}\mu_t} M_{k-1,t}$  for each  $k \leq t \leq \ell$ , and hence  $x_0 \in \tilde{\mathfrak{N}}_{k-1,\ell}(\mu; 2C^{-1}\rho_k^{-1})$ . The first inclusion in (3.15) therefore implies that  $x_0 \in \gamma_k(\Omega_{k,\ell}(\mu))$ . By the definition of  $x_0$  and the injectivity of  $\gamma_k$ , we therefore deduce that  $\Phi_{k,\ell}(s; \eta_{k,\ell}) \in \Omega_{k,\ell}(\mu)$ . By the injectivity of  $\Phi_{k,\ell}$  and the definition of  $\Omega_{k,\ell}(\mu)$ , we further deduce that  $\eta_{k,\ell} \in P_{k,\ell}(\mu)$ . Since we have already shown  $|\eta_k| \leq \mu_k$ , we conclude that  $\eta_{k-1,\ell} \in P_{k-1,\ell}(\mu)$ . Thus, (3.16) implies  $x \in \Omega_{k-1,\ell}(\mu)$ , as required.

Assuming the second inclusion in (3.15), we show the second inclusion in (3.14). Let  $x \in \Omega_{k-1,\ell}(\mu)$ , so that  $x = \Phi_{k-1,\ell}(s; \eta_{k-1,\ell})$  for some  $s \in W_{\ell}$  and  $\eta_{k-1,\ell} \in P_{k-1,\ell}(\mu)$ . For  $x_0 := \gamma_k \circ \Phi_{k,\ell}(s; \eta_{k,\ell})$  as before, we automatically have  $|x - x_0| = |\eta_k| \leq \mu_k$ . Furthermore,  $x_0 \in \gamma_k(\Omega_{k,\ell}(\mu))$  and so (3.15) implies that  $x_0 \in \tilde{\mathfrak{N}}_{k-1,\ell}(\mu; 2^{-1}C\rho_k)$ . By the ordering of  $\mu$  and the triangle inequality, provided  $C$  is chosen sufficiently large we have  $x \in \mathfrak{N}_{k-1,\ell}(\mu; C\rho_k)$ , as required.

It remains to verify (3.15). Let  $x \in \tilde{\mathfrak{N}}_{k-1,\ell}(\mu; 2C^{-1}\rho_k^{-1})$ . Since  $x \in M_{k-1,k} \cap W_{k-1}$ , there exists some  $s_k \in W_k$  such that  $x = \gamma_k(s_k)$ . On the other hand, fixing  $k+1 \leq t \leq \ell$ , there exists some  $\tilde{x} \in M_{k-1,t}$  such that  $|x - \tilde{x}| < 2C^{-1}\rho_k^{-1}\mu_t$ . Moreover,  $\tilde{x} = \gamma_k(\tilde{s}_k)$  for some  $\tilde{s}_k \in M_{k,t}$ . We therefore have  $|\gamma_k(s_k) - \gamma_k(\tilde{s}_k)| < 2C^{-1}\rho_k^{-1}\mu_t$ . Since the

parametrisations  $\gamma_k$  are regular, provided that  $C \geq 1$  is chosen sufficiently large, it follows that  $|s_k - \tilde{s}_k| < \rho_k^{-1} \mu_t$ . This implies that  $s_k \in \mathcal{N}_{\rho_k^{-1} \mu_t} M_{k,t}$ . Since this holds for all  $k+1 \leq t \leq \ell$ , we have shown

$$\tilde{\mathfrak{N}}_{k-1,\ell}(\boldsymbol{\mu}; 2C^{-1} \rho_k^{-1}) \subseteq \gamma_k(\mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; \rho_k^{-1})) \subseteq \gamma_k(\Omega_{k,\ell}(\boldsymbol{\mu})),$$

where the last inequality is due to the induction hypothesis (3.13).

Now suppose  $x \in \gamma_k(\Omega_{k,\ell}(\boldsymbol{\mu}))$ . Using the induction hypothesis (3.13) again, there exists some  $s_k \in \mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; C\rho_{k+1}) = \mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; \rho_k)$  such that  $x = \gamma_k(s_k)$ . Fixing  $k+1 \leq t \leq \ell$ , it follows that  $s_k \in \mathcal{N}_{\rho_k \mu_t} M_{k,t} \cap W_k$  and so there exists  $\tilde{s}_k \in M_{k,t}$  such that  $|s_k - \tilde{s}_k| < \rho_k \mu_t$ . By the mean value theorem, provided  $C \geq 1$  is sufficiently large,  $|\gamma_k(s_k) - \gamma_k(\tilde{s}_k)| < 2^{-1} C \rho_k \mu_t$ . Since  $\gamma_k(\tilde{s}_k) \in M_{k-1,t} \cap W_{k-1}$ , it therefore follows that  $x \in \mathcal{N}_{2^{-1} C \rho_k \mu_t} M_{k-1,t} \cap W_{k-1}$ . Since this holds for all  $k+1 \leq t \leq \ell$ , we have shown

$$\gamma_k(\Omega_{k,\ell}(\boldsymbol{\mu})) \subseteq \gamma_k(\mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; \rho_k)) \subseteq \tilde{\mathfrak{N}}_{k-1,\ell}(\boldsymbol{\mu}; 2^{-1} C \rho_k),$$

which concludes the proof of (3.15).  $\square$

Define  $\mathfrak{N}_{k,\ell}(\boldsymbol{\mu}) := \mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; 1)$  and  $\mathfrak{N}_{k,\ell}^+(\boldsymbol{\mu}) := \mathfrak{N}_{k,\ell}(\boldsymbol{\mu}; C_\circ)$ , where  $C_\circ$  is the constant appearing in the statement of Lemma 3.2.

**Corollary 3.3.** *Let  $0 \leq k \leq \ell \leq r$ , and define the sets*

$$\Omega_{k,\ell,r}(\boldsymbol{\mu}) := \Phi_{k,\ell}(\mathfrak{N}_{\ell,r}(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})) \quad \text{and} \quad \Omega_{k,\ell,r}^+(\boldsymbol{\mu}) := \Phi_{k,\ell}(\mathfrak{N}_{\ell,r}^+(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})).$$

*Then*

$$(3.17) \quad \mathfrak{N}_{k,r}(\boldsymbol{\mu}) \subseteq \Omega_{k,\ell,r}^+(\boldsymbol{\mu}) \quad \text{and} \quad \Omega_{k,\ell,r}(\boldsymbol{\mu}) \subseteq \mathfrak{N}_{k,r}^+(\boldsymbol{\mu})$$

*for all  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_\ell)$  such that  $0 < \mu_1 \leq \dots \leq \mu_r \leq C_\circ \mu_\circ$ .*

*Proof:* The formula (3.4) implies that

$$\Omega_{k,r}(\boldsymbol{\mu}) = \Phi_{k,\ell}(\Omega_{\ell,r}(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})).$$

The set inclusions (3.17) then follow from several applications of Lemma 3.2. In particular,

$$\mathfrak{N}_{k,r}(\boldsymbol{\mu}) \subseteq \Omega_{k,r}(\boldsymbol{\mu}) = \Phi_{k,\ell}(\Omega_{\ell,r}(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})) \subseteq \Phi_{k,\ell}(\mathfrak{N}_{\ell,r}^+(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})) = \Omega_{k,\ell,r}^+(\boldsymbol{\mu}),$$

and similarly

$$\Omega_{k,\ell,r}(\boldsymbol{\mu}) = \Phi_{k,\ell}(\mathfrak{N}_{\ell,r}(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})) \subseteq \Phi_{k,\ell}(\Omega_{\ell,r}(\boldsymbol{\mu}) \times P_{k,\ell}(\boldsymbol{\mu})) = \Omega_{k,r}(\boldsymbol{\mu}) \subseteq \mathfrak{N}_{k,r}^+(\boldsymbol{\mu}),$$

as required.  $\square$

**3.2. Slice formula.** Continuing with the setup from §3.1, let  $E_S$  be an extension operator associated to  $S$  with amplitude  $a \in C_c^\infty(U)$  chosen to have support in the open set  $W := W_0$ . Suppose  $f \in L^2(S)$  satisfies  $\text{supp } f \subseteq \mathcal{N}_\mu \mathcal{S}$  so that, by Lemma 3.2, we have  $\text{supp}(f \cdot a) \subseteq \mathfrak{N}_{0,\ell}(\boldsymbol{\mu}) \subseteq \Omega_{0,\ell}(\boldsymbol{\mu})$  for  $1 \leq \ell \leq r$ . Applying the change of variables

$$\Phi_\ell := \Phi_{0,\ell}: W_\ell \times V_{0,\ell} \rightarrow W,$$

and setting  $P_\ell(\boldsymbol{\mu}) := P_{0,\ell}(\boldsymbol{\mu})$ , we obtain

$$(3.18) \quad E_S f(x) = C \int_{P_\ell(\boldsymbol{\mu})} e^{ix \cdot \Sigma \circ \Phi_\ell(0; \boldsymbol{\eta})} E_{S_\ell(\boldsymbol{\eta})} f_{\ell, \boldsymbol{\eta}}(x) d\boldsymbol{\eta},$$

where  $f_{\ell, \boldsymbol{\eta}}(s) := f \circ \Phi_\ell(s; \boldsymbol{\eta})$  and  $E_{S_\ell(\boldsymbol{\eta})}$  is an extension operator associated to the  $d_\ell$ -dimensional submanifold  $S_\ell(\boldsymbol{\eta}) := \Sigma_{\ell, \boldsymbol{\eta}}(W_\ell)$  for

$$\Sigma_{\ell, \boldsymbol{\eta}}(s) := \Sigma \circ \Phi_\ell(s; \boldsymbol{\eta}) - \Sigma \circ \Phi_\ell(0; \boldsymbol{\eta}) \quad \text{for } s \in W_\ell \text{ and } \boldsymbol{\eta} \in P_\ell(\boldsymbol{\mu}).$$

In particular,

$$(3.19) \quad E_{S_\ell(\boldsymbol{\eta})}g(x) = \int_{W_\ell} e^{ix \cdot \Sigma_{\ell, \boldsymbol{\eta}}(s)} g(s) a_{\ell, \boldsymbol{\eta}}(s) ds$$

for  $a_{\ell, \boldsymbol{\eta}}(s) := a \circ \Phi_\ell(s; \boldsymbol{\eta}) C^{-1} J_{\boldsymbol{\eta}}(s)$ , where  $J_{\boldsymbol{\eta}}(s)$  denotes the Jacobian factor arising from the change of variables and  $C \geq 1$  is an admissible constant, chosen to ensure  $0 \leq a_{\ell, \boldsymbol{\eta}}(s) \leq 1$  for all  $s \in W_\ell$ .

By Corollary 3.3, we have  $\text{supp}(f \cdot a) \subseteq \mathfrak{N}_{0, \ell}(\boldsymbol{\mu}) \subseteq \Omega_{0, \ell, r}^+(\boldsymbol{\mu})$ , and so

$$(3.20) \quad \text{supp}(f_{\ell, \boldsymbol{\eta}} \cdot a_{\ell, \boldsymbol{\eta}}) \subseteq \mathfrak{N}_{\ell, r}^+(\boldsymbol{\mu}) \quad \text{for each } \boldsymbol{\eta} \in P_\ell(\boldsymbol{\mu}).$$

We refer to (3.18) as the *slice formula* for  $E_S f$ . We observe the following properties of the slice formula:

- $S_\ell(0) = \Sigma \circ \sigma_\ell(W_\ell) \subseteq S_\ell$  is a codimension 0 submanifold of  $S_\ell$ .
- Since  $f \in L^2(S)$ , we have that  $f_{\ell, \boldsymbol{\eta}} \in L^2(S_\ell)$  for almost every  $\boldsymbol{\eta} \in P_\ell(\boldsymbol{\mu})$ . Furthermore, by (3.20), we have

$$(3.21) \quad \text{supp}(f_{\ell, \boldsymbol{\eta}} \cdot a_{\ell, \boldsymbol{\eta}}) \subseteq \mathcal{N}_{\min\{C \circ \mu_k, \mu_o\}} S_k \quad \text{for all } \ell + 1 \leq k \leq r,$$

where the minimum is guaranteed provided  $\rho(E_S)$  is sufficiently small.

Here we are using the support (abuse of) notation introduced in §1.1.

**3.3. Local constancy of the slices.** Inequalities such as (1.5) are local in the sense that the left-hand  $L^{q_k}$ -norm is localised to the cube  $Q_R := [-R, R]^n$ . Consequently, by the uncertainty principle, we should expect our functions to be locally constant at scale  $R^{-1}$  in frequency space. In particular, continuing with the setup of the previous subsection, it should be possible to remove the dependence on  $\boldsymbol{\eta}$  in the  $S_\ell(\boldsymbol{\eta})$  when  $\boldsymbol{\eta} \in P_\ell(\boldsymbol{\mu})$  for  $\boldsymbol{\mu} = (\mu_i)_{i=1}^r$  with  $0 < \mu_1 \leq \dots \leq \mu_\ell < R^{-1}$ .

To implement the locally constant property rigorously, we go back to our formula for  $E_{S_\ell(\boldsymbol{\eta})}$  from (3.19) and suppose  $\mu_\ell = \max_{i \leq \ell} \mu_i < \min\{R^{-1}, C \circ \mu_o\}$ . The phase is given by

$$x \cdot \Sigma_{\ell, \boldsymbol{\eta}}(s) = x \cdot \Sigma_{\ell, 0}(s) + x \cdot \mathcal{E}(s; \boldsymbol{\eta}),$$

where, by the mean value theorem,  $\mathcal{E}(s; \boldsymbol{\eta})$  satisfies  $|\mathcal{E}(s; \boldsymbol{\eta})| \lesssim |\boldsymbol{\eta}|_\infty \lesssim R^{-1}$  for all  $s \in \text{supp } a_{\ell, \boldsymbol{\eta}}$  and all  $\boldsymbol{\eta} \in P_\ell(\boldsymbol{\mu})$ . Thus, for  $x \in Q_R$ , the function  $e^{ix \cdot \mathcal{E}(s; \boldsymbol{\eta})}$  is essentially non-oscillatory and can therefore be removed. More precisely, letting  $C \geq 1$  denote a large, admissible constant, by power series expansion

$$e^{ix \cdot \mathcal{E}(s; \boldsymbol{\eta})} = \sum_{\alpha \in \mathbb{N}_0^n} \prod_{j=1}^n \frac{(iCR^{-1}x_j)^{\alpha_j}}{\alpha_j!} (C^{-1}R\mathcal{E}_j(s; \boldsymbol{\eta}))^{\alpha_j}.$$

For  $x \in Q_R$ , we may therefore write

$$(3.22) \quad E_{S_\ell(\boldsymbol{\eta})}g(x) = \sum_{\alpha \in \mathbb{N}_0^n} B_\alpha(x) E_{S_\ell}[a_{\boldsymbol{\eta}}^\alpha g](x),$$

where, provided  $C \geq 1$  is chosen sufficiently large,

$$a_{\boldsymbol{\eta}}^\alpha(s) := a_{\ell, \boldsymbol{\eta}}(s) \prod_{j=1}^n (C^{-1}R\mathcal{E}_j(s; \boldsymbol{\eta}))^{\alpha_j}$$

satisfies  $|a_{\boldsymbol{\eta}}^\alpha(s)| \leq 1$ , and

$$B_\alpha(x) := \prod_{j=1}^n \frac{(iCR^{-1}x_j)^{\alpha_j}}{\alpha_j!} \quad \text{and} \quad b_\alpha := \prod_{j=1}^n \frac{C^{\alpha_j}}{\alpha_j!}$$

satisfy  $|B_\alpha(x)| \leq b_\alpha$ , so that  $\sum_{\alpha \in \mathbb{N}_0^n} |B_\alpha(t)| \leq e^{nC}$ . Here the extension operator  $E_{S_\ell}$  is defined as in (1.1), with an implicit amplitude  $a$  chosen so that (3.22) holds. In particular, we choose this amplitude such that it is identically 1 on the support of  $a_{\ell, \eta}$  and with support contained in  $W_\ell$ . Thus, by Cauchy–Schwarz,

$$(3.23) \quad |E_{S_\ell(\eta)}g(x)|^2 \lesssim \sum_{\alpha \in \mathbb{N}_0^n} b_\alpha |E_{S_\ell}[a_\eta^\alpha g](x)|^2, \quad x \in Q_R.$$

Since the sequence  $b_\alpha$  decreases rapidly, (3.23) effectively bounds  $|E_{S_\ell(\eta)}g(x)|$  by  $|E_{S_\ell}g(x)|$ , and therefore is a rigorous interpretation of the locally constant property.

**3.4. Wavepacket decomposition.** Here we construct a variant of the classical wavepacket decomposition adapted to our nested geometry.

**The derivative of  $\Phi_\ell$ .** We first compute the differential  $D\Phi_\ell|_{(s;0)}$ , which is a  $d \times d$  matrix. By first applying the chain rule to the recursive formula (3.1), we have

$$(3.24) \quad D\Phi_{k-1, \ell}|_{(s;0_{k-1, \ell})} \begin{pmatrix} v_\ell \\ \mathbf{w}_{k-1, \ell} \end{pmatrix} = D\gamma_k|_{\gamma_{k, \ell}(s)} D\Phi_{k, \ell}|_{(s;0_{k, \ell})} \begin{pmatrix} v_\ell \\ \mathbf{w}_{k, \ell} \end{pmatrix} + G_k(\gamma_{k, \ell}(s))w_k$$

for all  $1 \leq k \leq \ell \leq r$ ,  $s \in U_\ell$ ,  $v_\ell \in \mathbb{R}^{d_\ell}$ , and  $\mathbf{w}_{k-1, \ell} = (w_k, \mathbf{w}_{k, \ell}) \in \mathbb{R}^{m_k} \times \mathbb{R}^{c_\ell - c_k}$ . Here we have used (3.2).

Let  $\partial_s \sigma_\ell(s)$  denote the Jacobian matrix of  $\sigma_\ell$  at  $s$  and  $B_{k, \ell}(s) \in \text{Mat}(\mathbb{R}, d \times m_k)$  denote the composition of matrices

$$\begin{aligned} B_{k, \ell}(s) &:= D\gamma_1|_{\gamma_{1, \ell}(s)} \circ \cdots \circ D\gamma_{k-1}|_{\gamma_{k-1, \ell}(s)} \circ G_k(\gamma_{k, \ell}(s)) \\ &= \partial_s \sigma_{k-1}(\gamma_{k-1, \ell}(s)) G_k(\gamma_{k, \ell}(s)), \end{aligned}$$

for  $1 \leq k \leq \ell$ , where here  $B_{1, \ell}(s) := G_1 \circ \gamma_{1, \ell}(s)$  and the second equality is due to the chain rule. Furthermore, let  $\mathbf{B}_\ell(s) \in \text{Mat}(\mathbb{R}, d \times c_\ell)$  denote the block matrix

$$(3.25) \quad \mathbf{B}_\ell(s) := \begin{bmatrix} B_{1, \ell}(s) & B_{2, \ell}(s) & \cdots & B_{\ell, \ell}(s) \end{bmatrix}.$$

With these definitions, repeated application of (3.24) yields

$$(3.26) \quad D\Phi_\ell|_{(s;0)} \begin{pmatrix} v_\ell \\ \mathbf{w}_\ell \end{pmatrix} = \partial_s \sigma_\ell(s) v_\ell + \sum_{k=1}^{\ell} B_{k, \ell}(s) w_k = \partial_s \sigma_\ell(s) v_\ell + \mathbf{B}_\ell(s) \mathbf{w}_\ell$$

for  $s \in U_\ell$ ,  $v_\ell \in \mathbb{R}^{d_\ell}$ , and  $\mathbf{w}_\ell \in \mathbb{R}^{d-d_\ell}$ . Let  $b_{j, \ell}(s)$  denote the  $j$ th column of  $\mathbf{B}_\ell(s)$  and  $V_{k, \ell}(s) := \text{span}\{b_{1, \ell}(s), \dots, b_{c_k, \ell}(s)\}$  for  $1 \leq k \leq \ell$ . By the non-degeneracy of the parametrisations,

$$(3.27) \quad \left| \bigwedge_{j=1}^{c_\ell} b_{j, \ell}(s) \right| \gtrsim 1, \quad |b_{j, \ell}(s)| \lesssim 1, \quad |V_{k, \ell}(s) \wedge T_{\sigma_\ell(s)} M_k| \gtrsim 1$$

for all  $1 \leq k \leq \ell \leq r$ ,  $1 \leq j \leq c_\ell$ , and  $s \in U_\ell$ .

**Anisotropic decomposition.** We continue with the nested family  $\mathcal{S} = (S, (S_\ell)_{\ell=1}^r)$  from §3.1. Let  $R \geq 1$  and suppose  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_r)$  with  $R^{-1} \leq \mu_1 \leq \dots \leq \mu_r \leq \mu_\circ$ . Note that the upper bound is slightly stronger than that appearing in the setup in previous subsections, whilst the lower bound corresponds to the regime in which we do not have access to the local constancy property (3.23). Set

$$\mathcal{L}_{R^{1/2}, \boldsymbol{\mu}} := \{1 \leq \ell \leq r : \mu_\ell > R^{-1/2}\} \quad \text{and} \quad \ell_{R^{1/2}, \boldsymbol{\mu}} := \begin{cases} \min \mathcal{L}_{R^{1/2}, \boldsymbol{\mu}} - 1 & \text{if } \mathcal{L}_{R^{1/2}, \boldsymbol{\mu}} \neq \emptyset, \\ r & \text{otherwise,} \end{cases}$$



so that if  $\mathcal{L}_{R^{1/2}, \mu} \neq \emptyset$ , then  $\mathcal{L}_{R^{1/2}, \mu} = \{\ell_{R^{1/2}, \mu} + 1, \dots, r\}$ . For ease of notation in this section, we set  $\ell_\star := \ell_{R^{1/2}, \mu}$ . We define a  $d \times d$  diagonal matrix

$$D_{R^{1/2}, \mu} := \begin{bmatrix} \mu_1 I_{m_1} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mu_{\ell_\star} I_{m_{\ell_\star}} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & R^{-1/2} I_{d_\star} \end{bmatrix},$$

where  $d_\star := d - \sum_{\ell=1}^{\ell_\star} m_\ell = \dim(S_{\ell_\star})$ . Finally, for  $\mathbf{B}_\ell(s)$  as in (3.25), define the  $d \times d$  matrices

$$(3.28) \quad \Lambda_\ell(s) := \begin{bmatrix} \mathbf{B}_\ell(s) & \frac{\partial \sigma_\ell}{\partial s}(s) \end{bmatrix}, \quad s \in W_\ell,$$

which, by (3.27), are invertible for  $1 \leq \ell \leq r$ , and let

$$(3.29) \quad \Lambda_{R, \mu}(s) := \begin{cases} C_o^{1/2} \Lambda_{\ell_\star}(s) \circ D_{R^{1/2}, \mu} & \text{if } 1 \leq \ell_\star \leq r, \\ C_o^{1/2} R^{-1/2} I_d & \text{if } \ell_\star = 0, \end{cases}$$

noting that  $D_{R^{1/2}, \mu} = R^{-1/2} I_d$  in the  $\ell_\star = 0$  case.

Define a grid of points

$$\Gamma_{R, \mu}(\mathcal{S}) := R^{-1/2} \mathbb{Z}^{d_\star} \cap \mathfrak{N}_{\ell_\star, r}(\mu; 2C_o),$$

where  $\mathfrak{N}_{\ell_\star, r}(\mu; 2C_o)$  is as defined in (3.12) and  $C_o \geq 1$  is an admissible constant, chosen large enough so that the conclusion of Lemma 3.2 holds. We form a covering of

$$\mathfrak{N}_r(\mu) := \mathfrak{N}_{0, r}(\mu) = \mathcal{N}_{\mu_1} M_1 \cap \cdots \cap \mathcal{N}_{\mu_r} M_r \cap W$$

by parallelepipeds. Let  $\Theta_{R, \mu}(\mathcal{S})$  denote the family of all parallelepipeds of the form

$$(3.30) \quad \theta = u_\theta + \Lambda_\theta([-1, 1]^d), \quad \text{where } u_\theta := \sigma_{\ell_\star}(s_\theta) \text{ and } \Lambda_\theta := \Lambda_{R, \mu}(s_\theta)$$

for  $s_\theta \in \Gamma_{R, \mu}(\mathcal{S})$ . Here  $\sigma_0 := \text{Id}$ .

**Lemma 3.4.** *Provided  $C_o \geq 1$  is sufficiently large, the following properties hold:*

- (a) *The collection  $\Theta_{R, \mu}(\mathcal{S})$  forms a cover of  $\mathfrak{N}_r(\mu)$ .*
- (b) *The scaled parallelepipeds  $\{4 \cdot \theta : \theta \in \Theta_{R, \mu}(\mathcal{S})\}$  are finitely overlapping.*
- (c)  *$4 \cdot \theta \cap W \subseteq \mathfrak{N}_r(C_o^2 \mu)$  for all  $\theta \in \Theta_{R, \mu}(\mathcal{S})$ .*

*Proof:* If  $\ell_\star = 0$ , then the  $\Theta_{R, \mu}(\mathcal{S})$  is a standard covering by  $R^{1/2}$  cubes, and the desired properties are immediate. We therefore assume  $1 \leq \ell_\star \leq r$ .

(a) By Corollary 3.3, it suffices to show that  $\Theta_{R, \mu}(\mathcal{S})$  forms a cover of  $\Omega_{0, \ell_\star, r}^+(\mu)$ . Fix  $x \in \Omega_{0, \ell_\star, r}^+(\mu)$  so that

$$x = \Phi_{0, \ell_\star}(s; \eta) = \Phi_{\ell_\star}(s; \eta) \quad \text{for some } s \in \mathfrak{N}_{\ell_\star, r}^+(\mu) \text{ and } \eta \in P_{0, \ell_\star}(\mu) = P_{\ell_\star}(\mu).$$

Since  $\mathfrak{N}_{\ell_\star, r}^+(\mu) = \mathfrak{N}_{\ell_\star, r}(\mu; C_o)$ , there exists some  $s_\theta \in \Gamma_{R, \mu}(\mathcal{S})$  such that  $|s - s_\theta|_\infty < R^{-1/2}$ . The problem is therefore reduced to showing  $\Phi_{\ell_\star}(s; \eta) \in \theta$ , provided  $C_o \geq 1$  is sufficiently large.

For  $u_\theta := \sigma_{\ell_\star}(s_\theta) = \Phi_{\ell_\star}(s_\theta; \mathbf{0})$ , write

$$\Phi_{\ell_\star}(s; \eta) - u_\theta = \mathbf{B}_{\ell_\star}(s_\theta) \eta + \partial_s \sigma_{\ell_\star}(s_\theta)(s - s_\theta) + \mathcal{E}_\theta(s; \eta),$$

where

$$(3.31) \quad \mathcal{E}_\theta(s; \eta) := \Phi_{\ell_\star}(s; \eta) - \sigma_{\ell_\star}(s_\theta) - \partial_s \sigma_{\ell_\star}(s_\theta)(s - s_\theta) - \mathbf{B}_{\ell_\star}(s_\theta) \eta.$$

By combining the definitions (3.28), (3.29), and (3.30), we have

$$\Phi_{\ell_*}(s; \boldsymbol{\eta}) - u_\theta = \Lambda_\theta(\boldsymbol{\xi} + \boldsymbol{\zeta}), \quad \text{where} \quad \boldsymbol{\zeta} := (\Lambda_\theta)^{-1} \mathcal{E}_\theta(s; \boldsymbol{\eta})$$

and

$$\boldsymbol{\xi} := C_o^{-1/2} (D_{R^{1/2}, \boldsymbol{\mu}})^{-1} \begin{bmatrix} \boldsymbol{\eta} \\ s - s_\theta \end{bmatrix} \in [-1/2, 1/2]^d;$$

here we are assuming that  $C_o^{3/8} \geq 2$  and using the definition of  $P_{\ell_*}(\boldsymbol{\mu})$ . Since  $\mu_k \leq R^{-1/2}$  for  $1 \leq k \leq \ell_*$ , by (3.26), (3.2), and Taylor's theorem, we have

$$(3.32) \quad |\mathcal{E}_\theta(s; \boldsymbol{\eta})| \lesssim |s - s_\theta|_\infty^2 + |\boldsymbol{\eta}|_\infty^2 \lesssim C_o^{1/4} R^{-1}.$$

Thus, since we also have  $\mu_k \geq R^{-1}$  for  $1 \leq k \leq \ell_*$ , we can ensure that  $\boldsymbol{\zeta} \in [-1/2, 1/2]^d$  provided that  $C_o \geq 1$  is sufficiently large.

Combining the above observations,  $\Phi_{\ell_*}(s; \boldsymbol{\eta}) = u_\theta + \Lambda_\theta(\boldsymbol{\xi} + \boldsymbol{\zeta})$ , where  $\boldsymbol{\xi} + \boldsymbol{\zeta} \in [-1, 1]^d$  and so  $\Phi_{\ell_*}(s; \boldsymbol{\eta}) \in \theta$ , as required.

(b) The finite overlap property is immediate. Indeed, if  $\theta_1, \theta_2 \in \Theta_{R, \boldsymbol{\mu}}(\mathcal{S})$  are such that  $4 \cdot \theta_1 \cap 4 \cdot \theta_2 \neq \emptyset$ , then it follows that the corresponding centres  $u_{\theta_i} = \sigma_{\ell_*}(s_{\theta_i})$  satisfy  $|u_{\theta_1} - u_{\theta_2}|_\infty \lesssim R^{-1/2}$ . Since  $\sigma_{\ell_*}$  maps diffeomorphically onto its image, we conclude that  $|s_{\theta_1} - s_{\theta_2}|_\infty \lesssim R^{-1/2}$ . But for  $\theta_1$  fixed, this is only possible for  $O(1)$  choices of  $\theta_2$  since the  $s_{\theta_2}$  lie on  $R^{-1/2}$  separated points.

(c) Fix  $\theta \in \Theta_{R, \boldsymbol{\mu}}(\mathcal{S})$  with centre  $u_\theta = \sigma_{\ell_*}(s_\theta)$  for some  $s_\theta \in \Gamma_{R, \boldsymbol{\mu}}(\mathcal{S})$ .

Let  $\ell_* + 1 \leq t \leq r$  so that  $s_\theta \in \mathcal{N}_{2C_o\mu_t} M_{\ell_*, t}$ . Thus there exists some  $s_{\theta, t} \in M_{\ell_*, t}$  with  $|s_\theta - s_{\theta, t}| < 2C_o\mu_t$ . Consequently,  $|u_\theta - u_{\theta, t}| \lesssim C_o\mu_t$ , where  $u_{\theta, t} := \sigma_{\ell_*}(s_{\theta, t}) \in M_t$  and so  $\text{dist}(u_\theta, M_t) \lesssim C_o\mu_t$ . On the other hand, it is clear from the definition that  $4 \cdot \theta$  lies in a ball of radius  $O(C_o^{1/2} R^{-1/2})$  centred at  $u_\theta$ . Since  $\mu_t > R^{-1/2}$ , we conclude that

$$(3.33) \quad 4 \cdot \theta \subseteq \bigcap_{t=\ell_*+1}^r \mathcal{N}_{C_o^2\mu_t} M_t,$$

provided  $C_o \geq 1$  is chosen sufficiently large.

We now consider  $1 \leq k \leq \ell_*$ . Suppose  $u \in 4 \cdot \theta \cap W$ , so that, by the definitions (3.28), (3.29), and (3.30), we can write

$$u = \mathbf{B}_{\ell_*}(s_\theta) \boldsymbol{\eta} + \partial_s \sigma_{\ell_*}(s_\theta)(s - s_\theta) + \sigma_{\ell_*}(s_\theta)$$

for some  $\boldsymbol{\eta} \in P_{\ell_*}(4C_o^{3/8} \boldsymbol{\mu})$  and  $s - s_\theta \in [-4C_o^{1/2} R^{-1/2}, 4C_o^{1/2} R^{-1/2}]^{d_*}$ . Thus,

$$u = \Phi_{\ell_*}(s; \boldsymbol{\eta}) - \mathcal{E}_\theta(s; \boldsymbol{\eta}),$$

where  $\mathcal{E}_\theta(s; \boldsymbol{\eta})$  is as defined in (3.31). Arguing as in (3.32), we have  $|\mathcal{E}_\theta(s; \boldsymbol{\eta})| \lesssim C_o R^{-1}$ .

Since  $\Phi_{\ell_*}: W_{\ell_*} \times V_{\ell_*} \rightarrow W$  is a bijection, we can write  $u = \Phi_{\ell_*}(\tilde{s}; \tilde{\boldsymbol{\eta}})$  for some  $\tilde{s} \in W_{\ell_*}$  and  $\tilde{\boldsymbol{\eta}} \in V_{\ell_*} := V_{0, \ell_*}$ . Taking inverses,

$$(\tilde{s} - s; \tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}) = \Phi_{\ell_*}^{-1}(\Phi_{\ell_*}(s; \boldsymbol{\eta}) - \mathcal{E}_\theta(s; \boldsymbol{\eta})) - \Phi_{\ell_*}^{-1}(\Phi_{\ell_*}(s; \boldsymbol{\eta})),$$

and therefore, by the mean value theorem,  $|\tilde{\boldsymbol{\eta}} - \boldsymbol{\eta}|_\infty \lesssim C_o R^{-1}$ . Since  $\mu_k \geq R^{-1}$  for  $1 \leq k \leq r$  and  $\boldsymbol{\eta} \in P_{\ell_*}(4C_o^{3/8} \boldsymbol{\mu})$ , we have  $\tilde{\boldsymbol{\eta}} \in P_{\ell_*}(C_o \boldsymbol{\mu})$ , provided  $C_o \geq 1$  is chosen sufficiently large. By the definition (3.10), we have  $u = \Phi_{\ell_*}(\tilde{s}, \tilde{\boldsymbol{\eta}}) \in \Omega_{0, \ell_*}(C_o \boldsymbol{\mu})$ . Our hypotheses ensure that  $C_o \mu_1 \leq \dots \leq C_o \mu_r \leq C_o \mu_o$ , and so we may apply Lemma 3.2 to deduce that  $u \in \mathfrak{N}_{\ell_*}^+(C_o \boldsymbol{\mu}) = \mathfrak{N}_{\ell_*}(C_o^2 \boldsymbol{\mu})$ . Thus,

$$(3.34) \quad 4 \cdot \theta \cap W \subseteq \mathfrak{N}_{\ell_*}(C_o^2 \boldsymbol{\mu})$$

and combining (3.33) and (3.34) concludes the proof.  $\square$

**Definition of the wave packets.** Let  $\psi \in C_c^\infty(\mathbb{R}^d)$  satisfy  $0 \leq \psi \leq 1$ ;  $\psi(u) = 1$  if  $|u|_\infty \leq 1$  and  $\psi(u) = 0$  if  $|u|_\infty \geq 2$ . Define  $\tilde{\psi}(u) := \psi(u/2)$ , so that  $\psi\tilde{\psi} = \psi$ . Given  $\theta \in \Theta_{R,\mu}(\mathcal{S})$ , let

$$\psi_\theta(u) := \psi(\Lambda_\theta^{-1}(u - u_\theta)) \quad \text{and} \quad \tilde{\psi}_\theta(u) := \tilde{\psi}(\Lambda_\theta^{-1}(u - u_\theta)).$$

Thus,  $\psi_\theta(u) = 1$  if  $u \in \theta$ ,  $\text{supp } \psi_\theta \subseteq 2 \cdot \theta$ ,  $\text{supp } \tilde{\psi}_\theta \subseteq 4 \cdot \theta$ , and  $\psi_\theta \tilde{\psi}_\theta = \psi_\theta$ . By Lemma 3.4(a) and (b), we have

$$1 \leq \sum_{\theta \in \Theta_{R,\mu}(\mathcal{S})} \psi_\theta(u) \lesssim 1 \quad \text{for all } u \in \mathfrak{N}_r(\mu).$$

Now suppose  $f \in L^2(\mathcal{S})$  is smooth and satisfies  $\text{supp } f \subseteq \mathcal{N}_\mu \mathcal{S}$ . This abuse of notation translates as  $f \in C_c^\infty(U)$  with  $\text{supp } f \subseteq \mathfrak{N}_r(\mu)$ , provided we also assume  $\text{supp } f \subseteq W$ . Under these conditions, we may write

$$f = \sum_{\theta \in \Theta_{R,\mu}(\mathcal{S})} f_\theta \psi_\theta, \quad \text{where } f_\theta \in C_c^\infty(U) \text{ satisfy } |f_\theta(u)| \lesssim |f(u)| \text{ for all } u \in U.$$

Letting  $\Lambda_\theta^{-\top}(\mathbb{Z}^d) := \{\Lambda_\theta^{-\top} m : m \in \mathbb{Z}^d\}$ , define

$$\mathcal{T}_{R,\mu}(\mathcal{S}) := \Theta_{R,\mu}(\mathcal{S}) \times \Lambda_\theta^{-\top}(\mathbb{Z}^d) \quad \text{and} \quad \mathcal{T}_{R,\mu}(\mathcal{S}; \theta) := \{\theta\} \times \Lambda_\theta^{-\top}(\mathbb{Z}^d)$$

for all  $\theta \in \Theta_{R,\mu}(\mathcal{S})$ . By rescaling and applying a Fourier series decomposition,

$$f(u) = \sum_{(\theta,v) \in \mathcal{T}_{R,\mu}(\mathcal{S})} f_{\theta,v}(u) \quad \text{and} \quad (f_\theta \psi_\theta)(u) = \sum_{(\theta,v) \in \mathcal{T}_{R,\mu}(\mathcal{S}; \theta)} f_{\theta,v}(u)$$

for each  $\theta \in \Theta_{R,\mu}(\mathcal{S})$ , where

$$f_{\theta,v}(u) := (2\pi)^{-d} |\det \Lambda_\theta|^{-1} (f_\theta \psi_\theta)^\wedge(v) e^{iv \cdot u} \tilde{\psi}_\theta(u).$$

Here  $(f_\theta \psi_\theta)^\wedge(v) = \int_{\mathbb{R}^d} e^{-iv \cdot u} (f_\theta \psi_\theta)(u) du$  is the Fourier transform of  $f_\theta \psi_\theta$  evaluated at  $v$ .

**Basic properties of the wave packets.** We identify three key properties of the above decomposition:

*Orthogonality.* Parseval's identity for Fourier series and Lemma 3.4(b) give

$$\left\| \sum_{(\theta,v) \in \mathcal{T}} f_{\theta,v} \right\|_{L^2(\mathbb{R}^d)}^2 \lesssim \sum_{(\theta,v) \in \mathcal{T}} \|f_{\theta,v}\|_{L^2(\mathbb{R}^d)}^2 \lesssim \|f\|_{L^2(\mathbb{R}^d)}^2$$

for any collection  $\mathcal{T} \subseteq \mathcal{T}_{R,\mu}(\mathcal{S})$ .

*Spatial localisation.* Recall  $\Sigma_{\ell_*} := \Sigma \circ \sigma_{\ell_*}$  and  $S_{\ell_*} = \Sigma_{\ell_*}(U_{\ell_*})$ . Given  $\varepsilon > 0$ , let  $\varepsilon_\circ := \varepsilon/(100n)$  and define the slabs

$$(3.35) \quad T_{\theta,v} := \{x \in \mathbb{R}^n : |(\partial_s \Sigma_{\ell_*})(s_\theta)^\top x + v_\theta^*|_\infty < R^{1/2+\varepsilon_\circ}\}, \quad v_\theta^* := (\partial_s \sigma_{\ell_*})(s_\theta)^\top v$$

for all  $(\theta, v) \in \mathcal{T}_{R,\mu}(\mathcal{S})$ . Let  $u_\theta$  be as defined in (3.30). Observe that, by the chain rule formula

$$(\partial_s \Sigma_{\ell_*})(s_\theta) = (\partial_u \Sigma)(u_\theta) \circ (\partial_s \sigma_{\ell_*})(s_\theta)$$

and the definition of  $\Lambda_\theta$ , we have  $\tilde{T}_{\theta,v} \subseteq T_{\theta,v}$ , where

$$(3.36) \quad \tilde{T}_{\theta,v} := \{x \in Q_R : |\Lambda_\theta^\top((\partial_u \Sigma)(u_\theta)^\top x + v)|_\infty < R^{\varepsilon_\circ}\}.$$

Denote the collection of all slabs of the form (3.35) by

$$(3.37) \quad \mathbf{T}_{R,\mu}(\mathcal{S}) := \{T_{\theta,v} : (\theta, v) \in \mathcal{T}_{R,\mu}(\mathcal{S})\}$$

and, for each  $T = T_{\theta,v} \in \mathbf{T}_{R,\mu}(\mathcal{S})$ , write  $f_T := f_{\theta,v}$ . A standard integration-by-parts argument then shows that

$$|E_S f_T(x)| \lesssim_{N,\varepsilon} |E_S f_T(x)| \chi_T(x) + R^{-N} \|f\|_{L^2(U)}$$

for all  $x \in Q_R$  and all  $N \in \mathbb{N}_0$ , where  $E_S$  has amplitude  $a$  supported in the open set  $W$ . Indeed, we have a stronger estimate with  $\chi_T$  replaced with  $\chi_{\tilde{T}}$  for  $\tilde{T}$  as defined in (3.36), but, for our purposes, we only require localisation to the larger slab  $T$ .

*Preservation of support.* With the above setup, for  $C_\circ^2 \mu = (C_\circ^2 \mu_\ell)_{\ell=1}^r$ , we have

$$(3.38) \quad \text{supp } f_T \cap W \subseteq \mathfrak{N}_r(C_\circ^2 \mu) \quad \text{for all } T \in \mathbf{T}_{R,\mu}(\mathcal{S}).$$

Indeed, this follows from Lemma 3.4(c) after unravelling the notation and noting that  $\text{supp } \tilde{\psi}_\theta \subseteq 4 \cdot \theta$ . The localisation (3.38) is the key advantage of the above *anisotropic* wave packet decomposition, as opposed to the regular wave packet decomposition used, for instance, in [9, 6], and is crucial for the forthcoming induction-on-scales argument.

#### 4. The recursive scheme

Throughout this section, we fix  $2 \leq k \leq n$  and

$$\mathcal{E} = ((\mathcal{S}_j)_{j=1}^k, \mathbf{q}) \quad \text{with} \quad \mathcal{S}_j = (S_j, (S_{j,\ell})_{\ell=1}^{r_j})$$

a transverse ensemble in  $\mathbb{R}^n$  of codimension  $(m(j, \ell))_{j,\ell}$ . We apply the observations of §3 to each of the nested families of submanifolds  $\mathcal{S}_j$ . In particular, for  $1 \leq j \leq k$  we let  $\Phi_{k,\ell}^{(j)}: W_\ell^{(j)} \times V_{k,\ell}^{(j)} \rightarrow W_k^{(j)}$  denote the diffeomorphism chain and  $\mu_\circ^{(j)} > 0$  be the threshold width for  $\mathcal{S}_j$ , as defined in (3.9) and on page 144, respectively. We let  $\mu_\circ := \min\{\mu_\circ^{(1)}, \dots, \mu_\circ^{(k)}\} > 0$ .

**4.1. A reformulation.** Given  $R \geq 1$  and  $\mu = (\mu_j)_{j=1}^k$  with  $\mu_j := (\mu_{j,\ell})_{\ell=1}^{r_j}$  a compatible family of scales for  $\mathcal{E}$ , set

$$\mathcal{L}_{R,\mu}(j) := \{1 \leq \ell \leq r_j : \mu_{j,\ell} > R^{-1}\}$$

and

$$\ell_{R,\mu}(j) := \begin{cases} \min \mathcal{L}_{R,\mu}(j) - 1 & \text{if } \mathcal{L}_{R,\mu}(j) \neq \emptyset, \\ r_j & \text{otherwise,} \end{cases}$$

so that if  $\mathcal{L}_{R,\mu}(j) \neq \emptyset$ , then  $\mathcal{L}_{R,\mu}(j) = \{\ell_{R,\mu}(j) + 1, \dots, r_j\}$ . We then set

$$(4.1) \quad S_{j,R} := S_{j,\ell_{R,\mu}(j)} \quad \text{and} \quad U_{j,R} := U_{j,\ell_{R,\mu}(j)}, \quad 1 \leq j \leq k,$$

where  $S_{j,0} := S_j$  and  $U_{j,0} := U_j$ ; note that  $S_{j,R}$  and  $U_{j,R}$  also depend on  $\mu$ , but we suppress this for notational simplicity. Consider also the restricted nested family of submanifolds  $\mathcal{S}_{j,R} := (S_{j,R}, (S_{j,\ell})_{\ell \in \mathcal{L}_{R,\mu}(j)})$  for  $1 \leq j \leq k$ .

We let  $\rho_\circ > 0$  be a constant, chosen sufficiently small for the purposes of the forthcoming argument. In particular, we assume  $B(0, \rho_\circ) \subseteq W_0^{(j)}$  for  $1 \leq j \leq k$ . For all  $R \geq 1$  and  $\mu$  as above, let  $\mathfrak{R}(R; \mu)$  denote the smallest constant  $C \geq 1$  for which the inequality

$$(4.2) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \leq C C_R(\mu) \prod_{j=1}^k \|f_j\|_{L^2(U_{j,R})}^{q_j}$$

holds for

$$(4.3) \quad C_R(\mu) := \prod_{j=1}^k \prod_{\ell \in \mathcal{L}_{R,\mu}(j)} (\mu_{j,\ell})^{m(j,\ell)q_j/2}$$

whenever  $\rho(E_{S_{j,R}}) < \rho_\circ^{n-\ell_{R,\mu}(j)}$  for  $1 \leq j \leq k$  and the  $f_j \in L^2(S_{j,R})$  are smooth, compactly supported and satisfy

$$\text{supp } f_j \subseteq \mathcal{N}_{\mu_{j,\ell}} S_{j,\ell} \quad \text{for all } \ell \in \mathcal{L}_{R,\mu}(j).$$

Since  $f_j \in L^2(S_{j,R})$ , the reference manifolds here are  $S_{j,R}$ , and we understand  $\mathcal{N}_{\mu_{j,\ell}}(S_{j,\ell})$  as the neighbourhood in  $\mathbb{R}^{\dim(S_{j,R})}$  of the submanifold  $\gamma_{j,\ell_{R,\mu}(j)+1} \circ \cdots \circ \gamma_{j,\ell}(U_{j,\ell})$ . If  $\mathcal{L}_{R,\mu}(j) = \emptyset$ , then the product over  $\ell \in \mathcal{L}_{R,\mu}(j)$  in (4.3) is interpreted as equal to 1. We also define

$$\mathfrak{R}(R) := \sup_{\mu, \mu_{j,\ell} \leq \mu_\circ} \mathfrak{R}(R; \mu),$$

where the supremum is taken over all compatible families of scales  $\mu = (\mu_j)_{j=1}^k$  for  $\mathcal{E}$  with  $\mu_j = (\mu_{j,\ell})_{\ell=1}^{r_j}$  satisfying  $0 < \mu_{j,\ell} \leq \mu_\circ$  for  $1 \leq \ell \leq r_j$ ,  $1 \leq j \leq k$ .

Our goal is to show that, for all  $\varepsilon > 0$ , the bound  $\mathfrak{R}(R) \lesssim_\varepsilon R^\varepsilon$  holds for all  $R \geq 1$ . From this we shall deduce Theorem 1.6; see §4.3. We further reformulate our goal as follows. Let  $\mu$  be a compatible family of scales for  $\mathcal{E}$ . We first observe a trivial estimate for  $\mathfrak{R}(R; \mu)$ , by bounding the left-hand side of (4.2) in terms of the  $L^\infty$ -norms of the  $|E_{S_{j,R}} f_j|$  and applying a Riemann–Lebesgue-type argument to arrive at an expression involving the  $L^1$  norms of the  $f_j$ . We can then pass to  $L^2$  norms using Cauchy–Schwarz, taking advantage of the localisation of the  $\text{supp } f_j$ , to deduce that

$$(4.4) \quad \mathfrak{R}(R; \mu) \lesssim R^n \quad \text{for all } \mu \text{ compatible with } \mathcal{E}, \text{ and so } \mathfrak{R}(R) \lesssim R^n.$$

Thus, the *upper exponent*

$$(4.5) \quad \eta_{\text{exp}} := \limsup_{R \rightarrow \infty} \frac{\log \mathfrak{R}(R)}{\log R}$$

is a well-defined real number which satisfies  $\eta_{\text{exp}} \leq n$ . To prove Theorem 1.6, it shall suffice to show  $\eta_{\text{exp}} \leq 0$ .

**4.2. Uncertainty principle reduction.** As a manifestation of the uncertainty principle,  $\mathfrak{R}(R; \mu)$  automatically bounds a wider variety of multilinear expressions. In what follows, we let  $C_\circ \geq 1$  be an admissible constant, chosen sufficiently large for the purposes of the forthcoming argument and so that the conclusions of §3 hold. Let  $\mu = (\mu_j)_{j=1}^k$  be a compatible family of scales for  $\mathcal{E}$ , with  $\mu_j = (\mu_{j,\ell})_{\ell=1}^{r_j}$  and  $0 < \mu_{j,\ell} \leq \mu_\circ$  for  $1 \leq \ell \leq r_j$ ,  $1 \leq j \leq k$ . We define  $\tilde{\mu} = (\tilde{\mu}_j)_{j=1}^k$  with  $\tilde{\mu}_j = (\tilde{\mu}_{j,\ell})_{\ell=1}^{r_j}$  for  $\tilde{\mu}_{j,\ell} := \min\{C_\circ^2 \mu_{j,\ell}, \mu_\circ\}$ ,  $1 \leq \ell \leq r_j$ ,  $1 \leq j \leq k$ .

**Lemma 4.1.** *Let  $R \geq 1$  and  $\mathcal{L}_{R,\tilde{\mu}}(j) \subseteq \mathcal{L}^*(j) \subseteq \{1, \dots, r_j\}$  for  $1 \leq j \leq k$ . Suppose that whenever  $\ell \in \mathcal{L}^*(j)$  and  $\ell \leq \ell' \leq r_j$ , we have  $\ell' \in \mathcal{L}^*(j)$ . Define*

$$\ell^*(j) := \begin{cases} \min \mathcal{L}^*(j) - 1 & \text{if } \mathcal{L}^*(j) \neq \emptyset, \\ r_j & \text{otherwise,} \end{cases}$$

so that if  $\mathcal{L}^*(j) \neq \emptyset$ , then  $\mathcal{L}^*(j) := \{\ell^*(j) + 1, \dots, r_j\}$ . Set

$$S_j^* := S_{j,\ell^*(j)} \quad \text{and} \quad U_j^* := U_{j,\ell^*(j)} \quad \text{for } 1 \leq j \leq k.$$

If  $\rho(E_{S_j^*}) < \rho_\circ^{n-\ell^*(j)}$  for  $1 \leq j \leq k$ , then

$$(4.6) \quad \int_{Q^R} \prod_{j=1}^k |E_{S_j^*} f_j|^{q_j} \lesssim \mathfrak{R}(R; \tilde{\mu}) \prod_{j=1}^k \prod_{\ell \in \mathcal{L}^*(j)} (\mu_{j,\ell})^{m(j,\ell)q_j/2} \prod_{j=1}^k \|f_j\|_{L^2(U_j^*)}^{q_j}$$

holds whenever the  $f_j \in L^2(S_j^*)$  are smooth, compactly supported, and

$$(4.7) \quad \text{supp } f_j \subseteq \mathcal{N}_{\mu_{j,\ell}} S_{j,\ell} \quad \text{for all } \ell \in \mathcal{L}^*(j).$$

*Proof:* Let  $\mathcal{J}_R := \{1 \leq j \leq k : \ell^*(j) = \ell_{R,\tilde{\mu}}(j)\}$  and  $\mathcal{J}^* := \{1, \dots, k\} \setminus \mathcal{J}_R$ , so that  $S_j^* = S_{j,R}$  for all  $j \in \mathcal{J}_R$  and  $\ell^*(j) < \ell_{R,\tilde{\mu}}(j)$  for  $j \in \mathcal{J}^*$ . Provided  $\rho_o > 0$  is chosen sufficiently small, we can apply the slice formula (3.18) to the pairs  $S_j^*$  and  $S_{j,R}(\boldsymbol{\eta})$  for  $j \in \mathcal{J}^*$  and we have, by Cauchy–Schwarz,

$$(4.8) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_j^*} f_j|^{q_j} \lesssim \prod_{j \in \mathcal{J}^*} \prod_{\ell \in \mathbf{m}_j} (\mu_{j,\ell})^{m(j,\ell)q_j/2} I(R; \boldsymbol{\mu}),$$

where  $\mathbf{m}_j := \{\ell^*(j) + 1, \dots, \ell_{R,\tilde{\mu}}(j)\}$  and

$$I(R; \boldsymbol{\mu}) := \int_{Q_R} \prod_{j \in \mathcal{J}_R} |E_{S_{j,R}} f_j(x)|^{q_j} \prod_{j \in \mathcal{J}^*} \left( \int_{P_j(\boldsymbol{\mu}_j)} |E_{S_{j,R}(\boldsymbol{\eta})} f_{j,\boldsymbol{\eta}}(x)|^2 d\boldsymbol{\eta} \right)^{q_j/2} dx,$$

for

$$P_j(\boldsymbol{\mu}_j) := \prod_{\ell \in \mathbf{m}_j} [-C_o^{1/8} \mu_{j,\ell}, C_o^{1/8} \mu_{j,\ell}]^{m(j,\ell)} \quad \text{for } j \in \mathcal{J}^*.$$

Here, for an appropriate map  $\Phi_j^*$  satisfying  $\Phi_j^*(0; \mathbf{0}) = 0$ , we have

$$f_{j,\boldsymbol{\eta}}(s) := f_j \circ \Phi_j^*(s; \boldsymbol{\eta}) \quad \text{for } j \in \mathcal{J}^*.$$

Explicitly,

$$\Phi_j^* := \Phi_{\ell^*(j), \ell_{R,\tilde{\mu}}(j)}^{(j)} : W_{\ell_{R,\tilde{\mu}}(j)}^{(j)} \times V_{\ell^*(j), \ell_{R,\tilde{\mu}}(j)}^{(j)} \rightarrow W_{\ell^*(j)}^{(j)}.$$

Observe that  $\mathcal{L}^*(j) = \mathcal{L}_{R,\tilde{\mu}}(j) \cup \mathbf{m}_j$  for  $1 \leq j \leq k$ , where the union is disjoint and  $\mathbf{m}_j = \emptyset$  whenever  $j \in \mathcal{J}_R$ .

Fix  $j \in \mathcal{J}^*$  so that  $\ell^*(j) < \ell_{R,\tilde{\mu}}(j)$ . By the definition and the ordering of the  $(\mu_{j,\ell})_{\ell=1}^{r_j^j}$ , we have

$$\{1 \leq \ell \leq r_j : \tilde{\mu}_{j,\ell} \leq R^{-1}\} = \{1, \dots, \ell_{R,\tilde{\mu}}(j)\}$$

and so  $\mu_{j,\ell} \leq \tilde{\mu}_{j,\ell} \leq R^{-1}$  for all  $\ell \in \mathbf{m}_j$ . Consequently, given  $g \in C_c^\infty(U_{j,R})$ , from (3.23) we have

$$(4.9) \quad |E_{S_{j,R}(\boldsymbol{\eta})} g(x)|^2 \lesssim \sum_{\alpha[j] \in \mathbb{N}_0^n} b_{\alpha[j]} |E_{S_{j,R}}[a_{j,\boldsymbol{\eta}}^{\alpha[j]} g](x)|^2, \quad \boldsymbol{\eta} \in P_j(\boldsymbol{\mu}_j), x \in Q_R,$$

where the  $a_{j,\boldsymbol{\eta}}^{\alpha[j]} \in C^\infty(U_{j,R})$  satisfy  $|a_{j,\boldsymbol{\eta}}^{\alpha[j]}(s)| \leq 1$  and  $b_{\alpha[j]} := \prod_{i=1}^n \frac{C^{\alpha[j]_i}}{\alpha[j]_i!}$  for an admissible constant  $C \geq 1$ .

For each  $1 \leq j \leq k$ , define the Hilbert space  $H_j := L^2(P_j(\boldsymbol{\mu}_j))$  and consider the  $H_j$ -valued operator  $\mathbf{E}_{S_{j,R}}$  given by the mapping

$$\mathbf{E}_{S_{j,R}} \mathbf{g}_j(x) := (E_{S_{j,R}} g_{j,\boldsymbol{\eta}}(x))_{\boldsymbol{\eta}}, \quad \boldsymbol{\eta} \in P_j(\boldsymbol{\mu}_j), x \in \mathbb{R}^n,$$

where  $\mathbf{g}_j := (g_{j,\boldsymbol{\eta}})_{\boldsymbol{\eta}} \in L^2(U_{j,R}; H_j)$ . Observe that, by Fubini's theorem, for any  $x \in \mathbb{R}^n$  and  $h_j \in H_j$ , we have

$$\langle \mathbf{E}_{S_{j,R}} \mathbf{g}_j(x), h_j \rangle_{H_j} = E_{S_{j,R}}(\langle \mathbf{g}_j(\cdot), h_j \rangle_{H_j})(x).$$

Hence  $\mathbf{E}_{S_{j,R}}$  coincides with the  $H_j$ -valued extension of  $E_{S_{j,R}}$ , as described in the appendix. Letting  $\kappa := \#\mathcal{J}^*$ , we write  $\boldsymbol{\alpha} = (\alpha[j])_{j \in \mathcal{J}^*}$  for  $\boldsymbol{\alpha} \in (\mathbb{N}_0^n)^\kappa$ ,  $\mathbf{f}_j^\alpha := (f_{j,\boldsymbol{\eta}}^\alpha)_{\boldsymbol{\eta}}$  with  $f_{j,\boldsymbol{\eta}}^\alpha := a_{j,\boldsymbol{\eta}}^{\alpha[j]} f_{j,\boldsymbol{\eta}}$ . Using this notation, and applying (4.9), we have

$$(4.10) \quad \begin{aligned} I(R; \boldsymbol{\mu}) &\lesssim \int_{Q_R} \prod_{j \in \mathcal{J}_R} |E_{S_{j,R}} f_j(x)|^{q_j} \prod_{j \in \mathcal{J}^*} \left( \sum_{\alpha[j] \in \mathbb{N}_0^n} b_{\alpha[j]} \|\mathbf{E}_{S_{j,R}} \mathbf{f}_j^\alpha(x)\|_{H_j}^2 \right)^{q_j/2} dx \\ &\lesssim \sum_{\boldsymbol{\alpha} \in (\mathbb{N}_0^n)^\kappa} \prod_{j \in \mathcal{J}^*} b_{\alpha[j]}^{q_j/2} \int_{Q_R} \prod_{j \in \mathcal{J}_R} |E_{S_{j,R}} f_j(x)|^{q_j} \prod_{j \in \mathcal{J}^*} \|\mathbf{E}_{S_{j,R}} \mathbf{f}_j^\alpha(x)\|_{H_j}^{q_j} dx. \end{aligned}$$

Here the last inequality follows by the nesting of the  $\ell^{2/q_j}$  and  $\ell^1$  norms for  $2/q_j \geq 1$ .

If  $j \in \mathcal{J}^*$ , then, provided  $\rho_o > 0$  is chosen sufficiently small, we have  $\rho(E_{S_{j,R}}) < \rho_o^{-1}\rho(E_{S_j^*})$ . Our hypotheses therefore ensure

$$\rho(E_{S_{j,R}}) < \rho_o^{-1}\rho_o^{n-\ell^*(j)} \leq \rho_o^{n-\ell_{R,\tilde{\mu}}(j)} \quad \text{for } j \in \mathcal{J}^*,$$

since  $0 < \rho_o < 1$  and  $\ell_{R,\tilde{\mu}}(j) \geq \ell^*(j) + 1$ . Note also that  $\mathcal{L}_{R,\tilde{\mu}}(j) \subseteq \mathcal{L}^*(j)$  and that, by (3.21), the functions  $f_{j,\eta}$  are in  $L^2(S_{j,R})$  for almost every  $\eta \in P_j(\mu_j)$  and are supported in  $\mathcal{N}_{\tilde{\mu}_{j,\ell}}S_{j,\ell}$  for all  $\ell \in \mathcal{L}_{R,\tilde{\mu}}(j)$ .

In light of the preceding discussion, we can apply the inequality (4.2) featured in the definition of  $\mathfrak{R}(R; \tilde{\mu})$  to the extension operators  $E_{S_{j,R}}$  and the functions  $(f_j)_{j \in \mathcal{J}_R}$ ,  $(f_{j,\eta_j}^\alpha)_{j \in \mathcal{J}^*}$  as defined above. Moreover, the inequality (4.2) can be lifted to a vector-valued variant using a multilinear extension of the Marcinkiewicz-Zygmund theorem (see Proposition A.1). Combining these observations,

$$\begin{aligned} (4.11) \quad & \int_{Q_R} \prod_{j \in \mathcal{J}_R} |E_{S_{j,R}} f_j(x)|^{q_j} \prod_{j \in \mathcal{J}^*} \|\mathbf{E}_{S_{j,R}} \mathbf{f}_j^\alpha(x)\|_{H_j}^{q_j} dx \\ & \lesssim \mathfrak{R}(R; \tilde{\mu}) \prod_{j=1}^k \prod_{\ell \in \mathcal{L}_{R,\tilde{\mu}}(j)} (\mu_{j,\ell})^{m(j,\ell)q_j/2} \prod_{j \in \mathcal{J}_R} \|f_j\|_2^{q_j} \prod_{j \in \mathcal{J}^*} \|\mathbf{f}_j^\alpha\|_{H_j}^{q_j}. \end{aligned}$$

Here we have used the fact that  $\tilde{\mu}_{j,\ell} \leq C_o^2 \mu_{j,\ell}$  for all  $1 \leq j \leq k$  and  $\ell \in \mathcal{L}_{R,\tilde{\mu}}(j)$ .

By the pointwise bound for the  $a_{j,\eta}^{[j]}$  and a change of variables,

$$\|\mathbf{f}_j^\alpha\|_{H_j} = \left( \int_{P_j(\mu_j)} \|f_{j,\eta_j}^\alpha\|_2^2 d\eta_j \right)^{1/2} \lesssim \|f_j\|_2 \quad \text{for } j \in \mathcal{J}^*.$$

Combining this with (4.11), we obtain

$$\begin{aligned} (4.12) \quad & \int_{Q_R} \prod_{j \in \mathcal{J}_R} |E_{S_{j,R}} f_j(x)|^{q_j} \prod_{j \in \mathcal{J}^*} \|\mathbf{E}_{S_{j,R}} \mathbf{f}_j^\alpha(x)\|_{H_j}^{q_j} dx \\ & \lesssim \mathfrak{R}(R; \tilde{\mu}) \prod_{j=1}^k \prod_{\ell \in \mathcal{L}_{R,\tilde{\mu}}(j)} (\mu_{j,\ell})^{m(j,\ell)q_j/2} \prod_{j=1}^k \|f_j\|_2^{q_j}. \end{aligned}$$

We apply (4.12) to bound the right-hand side of (4.10), noting that the sequence  $(b_\alpha^{q_j/2})_{\alpha \in \mathbb{N}_0^n}$  is (absolutely) summable. Combining the resulting inequality with (4.8), we deduce the desired bound.  $\square$

**4.3. Recursive step.** Let  $p_j := q_j/2$  and  $1 \leq j \leq k$ . For  $R \geq 1$ , let  $\mathfrak{R}(R)$  be the smallest constant  $C \geq 1$  for which the inequality

$$\int_{Q_R} \prod_{j=1}^k \left| \sum_{T_j \in \mathbf{T}_j} c_{T_j} \chi_{T_j} \right|^{p_j} \leq C R^{n/2} \prod_{j=1}^k \left( \sum_{T_j \in \mathbf{T}_j} |c_{T_j}| \right)^{p_j}$$

holds whenever, for each  $1 \leq j \leq k$ , there exists some  $1 \leq \ell_j \leq r_j$  such that  $\mathbf{T}_j$  is a countable set of slabs of the form

$$(4.13) \quad T_j = \{x \in \mathbb{R}^n : |(\partial_s \Sigma_{j,\ell_j})(s)^\top x - v| < R^{1/2}\},$$

where  $s, v \in \mathbb{R}^{d'_j}$  with  $|s|_\infty < \rho_o$  for  $d'_j := \dim(S_{j,\ell_j})$  and  $(c_{T_j})_{T_j \in \mathbf{T}_j} \in \ell^1(\mathbf{T}_j)$ .

Let  $\mu = (\mu_j)_{j=1}^k$  be a compatible family of scales for  $\mathcal{E}$ , with  $\mu_j = (\mu_{j,\ell})_{\ell=1}^{r_j}$  and  $0 < \mu_{j,\ell} \leq \mu_o$  for  $1 \leq \ell \leq r_j$ ,  $1 \leq j \leq k$ . Recall the constant  $C_o \geq 1$  and the definition of  $\tilde{\mu}$  from §4.2. Let  $\tilde{\tilde{\mu}}$  denote the compatible family of scales which result from applying the operation  $\mu \mapsto \tilde{\mu}$  twice to  $\mu$ . Provided  $C_o \geq 1$  is chosen sufficiently large, the following holds.

**Lemma 4.2.** *For all  $\varepsilon > 0$ ,  $R \geq C_\circ^8$ , and  $\boldsymbol{\mu}$  compatible with  $\mathcal{E}$ , we have*

$$\mathfrak{R}(R; \boldsymbol{\mu}) \lesssim_\varepsilon R^\varepsilon \mathfrak{R}(R^{1/2}; \tilde{\boldsymbol{\mu}}) \mathfrak{R}(R).$$

Once Lemma 4.2 is established, Theorem 1.6 may be deduced as follows.

*Proof of Theorem 1.6:* We work with the fixed ensemble  $\mathcal{E}$  introduced at the beginning of the section. Without loss of generality, we may assume that our compatible family of scales  $\boldsymbol{\mu} = (\boldsymbol{\mu}_j)_{j=1}^k$  with  $\boldsymbol{\mu}_j = (\mu_{j,\ell})_{\ell=1}^{r_j}$  satisfies  $0 < \mu_{j,\ell} \leq \mu_\circ$  for  $1 \leq \ell \leq r_j$ ,  $1 \leq j \leq k$ . Indeed, otherwise we can pass to a (strict) subensemble  $\mathcal{F}$  of  $\mathcal{E}$  (consisting precisely of those  $S_{j,\ell}$  for which  $0 < \mu_{j,\ell} \leq \mu_\circ$  holds) and the analysis is simpler. This reduction can be formalised using an induction on the number of submanifolds in the ensemble.

In view of (4.4), we can assume  $R$  is sufficiently large, namely  $R \geq C_\circ^8$ . By the hypothesis  $\text{BL}_{\text{reg}}(\mathbf{L}(\mathcal{E}), \mathbf{p}) < \infty$  and Proposition 2.4, we have that  $\text{BL}_{\text{reg}}(\mathbf{L}(\mathcal{F}), \mathbf{p}) < \infty$  for all subensembles  $\mathcal{F}$  of  $\mathcal{E}$ . For  $\varepsilon > 0$ , let  $\nu_{\mathcal{F}} > 0$  and  $C_{\mathcal{F},\varepsilon} \geq 1$  be the constants appearing in Theorem 2.6 for each datum  $(\mathbf{L}(\mathcal{F}), \mathbf{p})$ , and set

$$\begin{aligned} \bar{\nu}_{\mathcal{E}} &:= \min\{\nu_{\mathcal{F}} : \mathcal{F} \text{ subensemble of } \mathcal{E}\} > 0, \\ \bar{C}_{\mathcal{E},\varepsilon} &:= \max\{C_{\mathcal{F},\varepsilon} : \mathcal{F} \text{ subensemble of } \mathcal{E}\} < \infty. \end{aligned}$$

Provided  $\rho_\circ > 0$  is sufficiently small, the slabs (4.13) have core planes which are, modulo translations, within distance  $\bar{\nu}_{\mathcal{E}}$  of the fixed subspaces  $\ker \partial_s \Sigma_{j,\ell_j}(0)^\top$ . Since, by (3.2), any  $(\partial_s \Sigma_{j,\ell_j}(0)^\top)_{j=1}^k = ((\partial_s \sigma_{\ell_j}(0))^\top (\partial_s \Sigma_j(0))^\top)_{j=1}^k$  corresponds to  $\mathbf{L}(\mathcal{F})$  for some subensemble  $\mathcal{F}$  of  $\mathcal{E}$ , it then follows from Theorem 2.6, taking  $\lambda = R^{1/2}$ , that for every  $\varepsilon > 0$  we have  $\mathfrak{R}(R) \leq \bar{C}_{\mathcal{E},\varepsilon} R^\varepsilon$  for all  $R \geq 1$ . Combining this with Lemma 4.2, for all  $\varepsilon > 0$  we have

$$(4.14) \quad \mathfrak{R}(R; \boldsymbol{\mu}) \lesssim_\varepsilon R^\varepsilon \mathfrak{R}(R^{1/2}; \tilde{\boldsymbol{\mu}}) \quad \text{for all } R \geq C_\circ^8.$$

Recall the definition of the upper exponent  $\eta_{\text{exp}}$  from (4.5), which, by (4.4), is a well-defined real number. The inequality (4.14) implies that  $\eta_{\text{exp}} \leq \eta_{\text{exp}}/2$  and so  $\eta_{\text{exp}} \leq 0$ . Thus, for all  $\varepsilon > 0$ , we have  $\mathfrak{R}(R) \lesssim_\varepsilon R^\varepsilon$ . Let  $\mathcal{L}^*(j) = \{1, \dots, r_j\}$ , so that  $\ell^*(j) = 0$ , for  $1 \leq j \leq k$ . Applying Lemma 4.1 with these  $\mathcal{L}^*(j)$  together with the above estimate for  $\mathfrak{R}(R)$ , the desired bound follows with  $\rho_{\mathcal{E}} := \rho_\circ^n$ .  $\square$

We turn to the proof of Lemma 4.2, which is a variant of an argument from [6, 9].

*Proof of Lemma 4.2:* Let  $\varepsilon > 0$  and, recalling (3.37), set

$$\mathbf{T}_j := \mathbf{T}_{R,\boldsymbol{\mu}}(S_{j,R}), \quad 1 \leq j \leq k.$$

Using the notation in (4.1), the slabs in  $\mathbf{T}_j$  have width  $R^{1/2+\varepsilon_\circ}$ , where  $\varepsilon_\circ := \varepsilon/(100n)$ , and their core planes are normal to the tangent plane to  $S_{j,R^{1/2}} = \Sigma_{j,R^{1/2}}(U_{j,R^{1/2}})$  at  $\Sigma_{j,R^{1/2}}(s)$  for some  $s \in U_{j,R^{1/2}}$ : that is, they are of the type (4.13). Provided  $\rho_\circ > 0$  is sufficiently small,

$$|s|_\infty \leq \rho_\circ^{-1} \rho(E_{S_{j,R}}) < \rho_\circ^{n-\ell_{R,\boldsymbol{\mu}}(j)-1} \leq \rho_\circ, \quad \text{since } \ell_{R,\boldsymbol{\mu}}(j) + 1 \leq r \leq n-1.$$

This observation will allow for an application of the bound  $\mathfrak{R}(R)$ .

Let  $\mathcal{Q}_{R^{1/2}}$  denote a cover of  $Q_R := [-R, R]^n$  by essentially disjoint cubes of side length  $R^{1/2}$ , so that

$$(4.15) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \leq \sum_{Q \in \mathcal{Q}_{R^{1/2}}} \int_Q \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j}.$$



Fix  $Q \in \mathcal{Q}_{R^{1/2}}$ . Provided  $\rho_o > 0$  is chosen sufficiently small, we can perform the wave packet decomposition from §3.4. By the spatial localisation property,

$$|E_{S_{j,R}} f_j(x)| \lesssim_{N,\varepsilon} |E_{S_{j,R}} f_{j,Q}(x)| + R^{-N} \|f_j\|_2 \quad \text{for all } N \in \mathbb{N}_0 \text{ and } x \in Q,$$

where

$$f_{j,Q} := \sum_{T_j \in \mathbf{T}_j[Q]} f_{j,T} \quad \text{for } \mathbf{T}_j[Q] := \{T \in \mathbf{T}_j : T \cap Q \neq \emptyset\}.$$

Thus,

$$\int_Q \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \lesssim_\varepsilon \int_Q \prod_{j=1}^k |E_{S_{j,R}} f_{j,Q}|^{q_j} + R^{-100n} \prod_{j=1}^k \|f_j\|_2^{q_j}.$$

For each  $T \in \mathbf{T}_j$ , we have  $\text{supp } f_{j,T} \subseteq B(0, \rho_o + 4 \cdot R^{-1/2})$ . Thus, for large  $R$ , we can guarantee that  $\text{supp } f_{j,Q} \subseteq W_0^{(j)}$ . By the preservation of support property (3.38), provided  $C_o \geq 1$  is chosen sufficiently large, each  $f_{j,Q}$  satisfies  $\text{supp } f_{j,Q} \subseteq \mathcal{N}_{C_o^2 \mu_j} \mathcal{S}_j$ . Moreover, by the support properties of our amplitude functions, we may assume without loss of generality that  $\text{supp } f_{j,Q} \subseteq \mathcal{N}_{\tilde{\mu}_j} \mathcal{S}_j$  for  $\tilde{\mu}_j$  as defined in §4.2.

Define  $\mathcal{L}^*(j) := \mathcal{L}_{R,\mu}(j)$  and, since  $\mu_{j,\ell} \geq C_o^{-4} \tilde{\mu}_{j,\ell} \geq R^{-1/2} \tilde{\mu}_{j,\ell}$ , note that

$$\mathcal{L}_{R^{1/2}, \tilde{\mu}}(j) \subseteq \mathcal{L}^*(j) \subseteq \{1, \dots, r_j\} \quad \text{for } 1 \leq j \leq k.$$

From the observations of the previous paragraph, the  $f_{j,Q}$  satisfy the hypothesis (4.7) of Lemma 4.1 with  $\mu$  replaced with  $\tilde{\mu}$ . We may therefore apply Lemma 4.1 to deduce that

$$(4.16) \quad \int_Q \prod_{j=1}^k |E_{S_{j,R}} f_{j,Q}|^{q_j} \lesssim \Re(R^{1/2}; \tilde{\mu}) \mathbf{C}_R(\mu) \prod_{j=1}^k \|f_{j,Q}\|_2^{q_j},$$

where  $\mathbf{C}_R(\mu)$  is as defined in (4.3). Observe that, for our choice of  $\mathcal{L}^*(j)$  and  $\mu$  replaced with  $\tilde{\mu}$ , the factors on the right-hand side of (4.6) satisfy

$$\prod_{\ell \in \mathcal{L}^*(j)} (\tilde{\mu}_{j,\ell})^{m(j,\ell)q_j/2} \lesssim \prod_{\ell \in \mathcal{L}_{R,\mu}(j)} (\mu_{j,\ell})^{m(j,\ell)q_j/2},$$

justifying the appearance of  $\mathbf{C}_R(\mu)$  in (4.16). Here we have also used the translation invariance of the estimates to recentre  $Q$  at the origin.

Taking the sum over  $\mathcal{Q}_{R^{1/2}}$  on both sides of (4.16) and combining the resulting estimate with (4.15), we deduce that

$$(4.17) \quad \int_{\mathcal{Q}_R} \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \lesssim_\varepsilon \Re(R^{1/2}; \tilde{\mu}) \mathbf{C}_R(\mu) \sum_{Q \in \mathcal{Q}_{R^{1/2}}} \prod_{j=1}^k \|f_{j,Q}\|_2^{q_j}$$

holds up to the inclusion of a rapidly decaying error term. It remains to estimate the right-hand sum in (4.17).

By the orthogonality properties of the wave packets,

$$\prod_{j=1}^k \|f_{j,Q}\|_2^2 \lesssim \prod_{j=1}^k \left( \sum_{T_j \in \mathbf{T}_j[Q]} \|f_{j,T_j}\|_2^2 \right).$$

Furthermore, if  $T_j \in \mathbf{T}_j[Q]$ , then it follows that  $Q \subset T_j^\circ$ , where  $T_j^\circ$  denotes the slab with the same core plane as  $T_j$  but with width scaled by a factor of 2. Consequently,

$$\prod_{j=1}^k \|f_{j,Q}\|_2^{q_j} \lesssim \frac{1}{|Q|} \int_Q \prod_{j=1}^k \left| \sum_{T_j \in \mathbf{T}_j} \|f_{j,T_j}\|_2^2 \cdot \chi_{T_j^\circ}(x) \right|^{q_j/2} dx.$$

Summing this inequality over all  $Q \in \mathcal{Q}_{R^{1/2}}$  and combining the resulting estimate with (4.17), we obtain

$$(4.18) \quad \int_{Q_R} \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \lesssim_\varepsilon \mathfrak{R}(R^{1/2}; \tilde{\boldsymbol{\mu}}) \mathbf{C}_R(\boldsymbol{\mu}) R^{-n/2} \int_{Q_R} \prod_{j=1}^k \left| \sum_{T_j \in \mathbf{T}_j} \|f_{j,T_j}\|_2^2 \cdot \chi_{T_j^\circ} \right|^{p_j},$$

up to the inclusion of a rapidly decaying error term, where we recall  $p_j := q_j/2$  for  $1 \leq j \leq k$ .

By the definition of the constant  $\mathfrak{R}(R)$  and the orthogonality of the wave packets,

$$(4.19) \quad \begin{aligned} \int_{Q_R} \prod_{j=1}^k \left| \sum_{T_j \in \mathbf{T}_j} \|f_{j,T_j}\|_2^2 \cdot \chi_{T_j^\circ} \right|^{p_j} &\lesssim_\varepsilon \mathfrak{R}(R) R^{n/2+\varepsilon} \prod_{j=1}^k \left( \sum_{T_j \in \mathbf{T}_j} \|f_{j,T_j}\|_2^2 \right)^{p_j} \\ &\lesssim \mathfrak{R}(R) R^{n/2+\varepsilon} \prod_{j=1}^k \|f_j\|_2^{q_j}. \end{aligned}$$

Note that the slabs featured in the above displays (whose geometry coincides with that of the slabs defined in (3.35)) are slightly wider (by a factor of  $R^{\varepsilon_\circ}$ , where  $\varepsilon_\circ := \varepsilon/(100n)$ ) than those appearing in the definition of  $\mathfrak{R}(R)$ . This discrepancy can be dealt with using a simple covering argument, which accounts for the additional  $R^\varepsilon$  factor on the right-hand side of the above inequality.

Combining (4.19) with (4.18), we obtain

$$\int_{Q_R} \prod_{j=1}^k |E_{S_{j,R}} f_j|^{q_j} \lesssim_\varepsilon R^\varepsilon \mathfrak{R}(R^{1/2}; \tilde{\boldsymbol{\mu}}) \mathfrak{R}(R) \mathbf{C}_R(\boldsymbol{\mu}) \prod_{j=1}^k \|f_j\|_2^{q_j}$$

and therefore, by definition,  $\mathfrak{R}(R; \boldsymbol{\mu}) \lesssim_\varepsilon R^\varepsilon \mathfrak{R}(R^{1/2}; \tilde{\boldsymbol{\mu}}) \mathfrak{R}(R)$ , as required.  $\square$

## Appendix A. A multilinear Marcinkiewicz–Zygmund theorem

Let  $H$  be a separable complex Hilbert space and  $\{e_a\}_{a \in \mathcal{A}}$  be a choice of orthonormal basis for  $H$ , where  $\mathcal{A}$  is a countable set. Given any  $g \in H$ , we may write

$$g = \sum_{a \in \mathcal{A}} \langle g, e_a \rangle e_a, \quad \text{where} \quad \|g\|_H = \left( \sum_{a \in \mathcal{A}} |\langle g, e_a \rangle|^2 \right)^{1/2}.$$

Let  $(\Omega, \Sigma, \mu)$  and  $(\Omega', \Sigma', \mu')$  be measure spaces. We say that  $\mathbf{f}: \Omega' \rightarrow H$  is  $\Sigma'$ -measurable if  $\langle \mathbf{f}, g \rangle$  is a  $\Sigma'$ -measurable function for all  $g \in H$ . Let  $T$  be a linear mapping sending  $\Sigma'$ -measurable functions on  $\Omega'$  to  $\Sigma$ -measurable functions on  $\Omega$ . We define the *vector-valued extension*  $\mathbf{T}$  to be the operator mapping  $H$ -valued measurable functions on  $\Omega'$  to  $H$ -valued measurable functions on  $\Omega$  given by

$$\langle \mathbf{T}\mathbf{f}(x), e_a \rangle := T(\langle \mathbf{f}, e_a \rangle)(x) \quad \text{for all } \mathbf{f}: \Omega' \rightarrow H \text{ measurable and all } a \in \mathcal{A}.$$

This uniquely defines the operator  $\mathbf{T}$ , which is linear.

The classical Marcinkiewicz–Zygmund inequality [15] for separable Hilbert spaces has the following multilinear counterpart.

**Proposition A.1** (Multilinear Marcinkiewicz–Zygmund theorem). *Let  $(\Omega, \Sigma, \mu)$  and  $(\Omega_j, \Sigma_j, \mu_j)$  for  $1 \leq j \leq m$  be  $\sigma$ -finite measure spaces. For  $1 \leq j \leq m$ , let  $T_j$  be linear mappings sending  $\Sigma_j$ -measurable functions on  $\Omega_j$  to  $\Sigma$ -measurable functions on  $\Omega$ . Let  $0 < p_j, q_j < \infty$  for  $1 \leq j \leq k$  and further suppose that there exists some  $M \geq 0$  such that*

$$\int_{\Omega} \prod_{j=1}^k |T_j f_j(x)|^{q_j} d\mu(x) \leq M \prod_{j=1}^k \|f_j\|_{L^{p_j}(\Omega_j, \mu_j)}^{q_j}$$

*holds for all  $\Sigma_j$ -measurable functions  $f_j$ ,  $1 \leq j \leq k$ .*

*Let  $H_j$  be separable complex Hilbert spaces and let  $\mathbf{T}_j$  denote the vector-valued extensions of  $T_j$ ,  $1 \leq j \leq k$ . Then the inequality*

$$\int_{\Omega} \prod_{j=1}^k \|\mathbf{T}_j \mathbf{f}_j(x)\|_{H_j}^{q_j} d\mu(x) \lesssim_{p,q} M \prod_{j=1}^k \|\mathbf{f}_j\|_{H_j}^{q_j} \| \mathbf{f}_j \|_{L^{p_j}(\Omega_j, \mu_j)}$$

*holds for all  $\Sigma_j$ -measurable functions  $\mathbf{f}_j: \Omega_j \rightarrow H_j$ ,  $1 \leq j \leq k$ .*

If the Hilbert spaces  $H_j$  are finite-dimensional, the result follows from a standard Khintchine’s inequality argument and the Fubini–Tonelli theorem. The extension to separable spaces follows from Fatou’s lemma. We omit the details.

**Acknowledgements.** We thank the referees for valuable suggestions. D. Beltran is supported by the grants RYC2020-029151-I and PID2022-140977NA-I00 funded by MICIU/AEI/10.13039/501100011033, “ESF investing in your future”, and FEDER, UE. J. Duncan is supported the Juan de la Cierva grant JDC2023-050459-I. J. Hickman is supported by the New Investigator award UKRI097.

## References

- [1] I. BEJENARU, The multilinear restriction estimate: a short proof and a refinement, *Math. Res. Lett.* **24**(6) (2017), 1585–1603. DOI: 10.4310/MRL.2017.v24.n6.a1.
- [2] I. BEJENARU, The almost optimal multilinear restriction estimate for hypersurfaces with curvature: the case of  $n - 1$  hypersurfaces in  $\mathbb{R}^n$ , *Int. Math. Res. Not. IMRN* **2022**(20) (2022), 16363–16404. DOI: 10.1093/imrn/rnab208.
- [3] I. BEJENARU, The multilinear restriction estimate: almost optimality and localization, *Math. Res. Lett.* **29**(3) (2022), 599–630. DOI: 10.4310/mrl.2022.v29.n3.a1.
- [4] D. BELTRAN, J. DUNCAN, AND J. HICKMAN, Off-diagonal estimates for the helical maximal function, *Proc. Lond. Math. Soc. (3)* **128**(4) (2024), Paper No. e12594, 36 pp. DOI: 10.1112/plms.12594.
- [5] J. BENNETT AND N. BEZ, Some nonlinear Brascamp–Lieb inequalities and applications to harmonic analysis, *J. Funct. Anal.* **259**(10) (2010), 2520–2556. DOI: 10.1016/j.jfa.2010.07.015.
- [6] J. BENNETT, N. BEZ, T. C. FLOCK, AND S. LEE, Stability of the Brascamp–Lieb constant and applications, *Amer. J. Math.* **140**(2) (2018), 543–569. DOI: 10.1353/ajm.2018.0013.
- [7] J. BENNETT, A. CARBERY, M. CHRIST, AND T. TAO, The Brascamp–Lieb inequalities: finiteness, structure and extremals, *Geom. Funct. Anal.* **17**(5) (2008), 1343–1415. DOI: 10.1007/s00039-007-0619-6.
- [8] J. BENNETT, A. CARBERY, M. CHRIST, AND T. TAO, Finite bounds for Hölder–Brascamp–Lieb multilinear inequalities, *Math. Res. Lett.* **17**(4) (2010), 647–666. DOI: 10.4310/MRL.2010.v17.n4.a6.
- [9] J. BENNETT, A. CARBERY, AND T. TAO, On the multilinear restriction and Kakeya conjectures, *Acta Math.* **196**(2) (2006), 261–302. DOI: 10.1007/s11511-006-0006-4.
- [10] S. GUO, H. WANG, AND R. ZHANG, A dichotomy for Hörmander-type oscillatory integral operators, *Invent. Math.* **238**(2) (2024), 503–584. DOI: 10.1007/s00222-024-01288-8.
- [11] L. GUTH, A restriction estimate using polynomial partitioning, *J. Amer. Math. Soc.* **29**(2) (2016), 371–413. DOI: 10.1090/jams827.

- [12] J. HICKMAN, K. M. ROGERS, AND R. ZHANG, Improved bounds for the Kakeya maximal conjecture in higher dimensions, *Amer. J. Math.* **144**(6) (2022), 1511–1560. DOI: 10.1353/ajm.2022.0037.
- [13] J. HICKMAN AND J. ZAHL, A note on Fourier restriction and nested polynomial Wolff axioms, *J. Anal. Math.* **152**(1) (2024), 19–52. DOI: 10.1007/s11854-023-0289-9.
- [14] D. MALDAGUE, Regularized Brascamp–Lieb inequalities and an application, *Q. J. Math.* **73**(1) (2022), 311–331. DOI: 10.1093/qmath/haab032.
- [15] J. MARCINKIEWICZ AND A. ZYGMUND, Quelques inégalités pour les opérations linéaires, *Fund. Math.* **32** (1939), 115–121. DOI: 10.4064/fm-32-1-115-121.
- [16] C. OH, An improved bilinear restriction estimate for the paraboloid in  $\mathbb{R}^3$ , *Math. Z.* **303**(4) (2023), Paper no. 88, 23 pp. DOI: 10.1007/s00209-023-03237-2.
- [17] T. TAO, Sharp bounds for multilinear curved Kakeya, restriction and oscillatory integral estimates away from the endpoint, *Mathematika* **66**(2) (2020), 517–576. DOI: 10.1112/mtk.12029.
- [18] J. ZAHL, New Kakeya estimates using Gromov’s algebraic lemma, *Adv. Math.* **380** (2021), Paper no. 107596, 42 pp. DOI: 10.1016/j.aim.2021.107596.
- [19] R. ZHANG, The endpoint perturbed Brascamp–Lieb inequalities with examples, *Anal. PDE* **11**(3) (2018), 555–581. DOI: 10.2140/apde.2018.11.555.
- [20] P. ZORIN-KRANICH, Kakeya–Brascamp–Lieb inequalities, *Collect. Math.* **71**(3) (2020), 471–492. DOI: 10.1007/s13348-019-00273-2.

David Beltran

Departament d’Anàlisi Matemàtica, Universitat de València, Dr. Moliner 50, 46100 Burjassot, Spain  
*E-mail address:* david.beltran@uv.es  
 ORCID: 0000-0001-6711-9741

Jennifer Duncan

Departamento de Matemáticas, Facultad de Ciencias, Módulo 17, Universidad Autónoma de Madrid, Ciudad Universitaria de Cantoblanco, 28049-Madrid, Spain  
*E-mail address:* jennifer.duncan@uam.es  
 ORCID: 0009-0002-0673-8818

Jonathan Hickman

University of Edinburgh, School of Mathematics, James Clerk Maxwell Building, The King’s Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK  
*E-mail address:* jonathan.hickman@ed.ac.uk  
 ORCID: 0000-0002-8193-0925

Received on April 22, 2024.

Accepted on December 20, 2024.