



A bilinear Rubio de Francia inequality for arbitrary rectangles

Frederic Bernicot, Marco Vitturi

► To cite this version:

Frederic Bernicot, Marco Vitturi. A bilinear Rubio de Francia inequality for arbitrary rectangles. International Mathematics Research Notices, 2020. hal-01858507

HAL Id: hal-01858507

<https://hal.science/hal-01858507>

Submitted on 20 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A BILINEAR RUBIO DE FRANCIA INEQUALITY FOR ARBITRARY RECTANGLES

FRÉDÉRIC BERNICOT AND MARCO VITTURI

ABSTRACT. Let \mathcal{R} be a collection of disjoint dyadic rectangles R , let π_R denote the non-smooth bilinear projection onto R

$$\pi_R(f, g)(x) := \iint \mathbf{1}_R(\xi, \eta) \hat{f}(\xi) \hat{g}(\eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta$$

and let $r > 2$. We show that the bilinear Rubio de Francia operator associated to \mathcal{R} given by

$$f, g \mapsto \left(\sum_{R \in \mathcal{R}} |\pi_R(f, g)|^r \right)^{1/r}$$

is $L^p \times L^q \rightarrow L^s$ bounded whenever $1/p + 1/q = 1/s$, $r' < p, q < r$. This extends from squares to rectangles a previous result by the same authors in [7], and as a corollary extends in the same way a previous result from [2] for smooth projections, albeit in a reduced range.

1. INTRODUCTION

In this paper we present an improvement over the results of Benea and the first author in [2] and the results of [7] by the same authors. The setup is as follows. Let \mathcal{D} denote the standard dyadic grid, that is the collection of intervals of the form $[2^k n, 2^k(n+1)]$ for $n, k \in \mathbb{Z}$. Let then $\mathcal{R} \subset \mathcal{D} \times \mathcal{D}$ be an arbitrary collection of disjoint rectangles $R = R_1 \times R_2$ with R_1, R_2 dyadic, and let π_R denote the non-smooth bilinear frequency projection on R , that is

$$\pi_R(f, g)(x) := \iint \hat{f}(\xi) \hat{g}(\eta) \mathbf{1}_R(\xi, \eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta.$$

We will consider the bilinear operators associated to the collection \mathcal{R} given by

$$T_{\mathcal{R}}^r(f, g)(x) := \left(\sum_{R \in \mathcal{R}} |\pi_R(f, g)(x)|^r \right)^{1/r}$$

for $r > 2$.

Before stating our results we provide some context. Linear variants of the operator $T_{\mathcal{R}}^r$ are known in the literature as Rubio de Francia operators. Indeed, given a collection of arbitrary disjoint intervals $\mathcal{I} = \{I_n, n \in \mathbb{Z}\}$, Rubio de Francia proved in [19] that the operator

$$\text{RdF}_{\mathcal{I}}^r f(x) := \left(\sum_n \left| \int \hat{f}(\xi) \mathbf{1}_{I_n}(\xi) e^{2\pi i \xi x} d\xi \right|^r \right)^{1/r}$$

is $L^p \rightarrow L^p$ bounded for $r' < p < \infty$ when $r \geq 2$, this last condition being necessary (a consequence of Khinchin's inequality). Notice that when $r = 2$ then $\text{RdF}_{\mathcal{I}}^2$ is also bounded at the endpoint $r' = 2$, this endpoint being just a consequence of Plancherel's identity, but when $r > 2$ the condition $r' < p$ is sharp (see [11]). Rubio de Francia's result is a generalization of the well known Littlewood-Paley inequalities ($I_n = [2^n, 2^{n+1}]$), extending an earlier result of Carleson [9] and (independently) Cordoba [10] for the collection $\mathcal{I} = \{[n, n+1], n \in \mathbb{Z}\}$; as such, it can be thought of as a statement about orthogonality for arbitrary frequency intervals.

The result has also been extended and reproved by different means in a number of papers, see [8, 12, 21, 22, 14, 3].

Bilinear operators of square-function type associated to collections of subsets of the frequency plane $\widehat{\mathbb{R}^2}$ have previously been considered as well. Perhaps the first such operator to have appeared in the literature is the square function

$$f, g \mapsto \left(\sum_{n \in \mathbb{Z}} \left| \iint \widehat{f}(\xi) \widehat{g}(\eta) \chi(\xi - \eta - n) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^2 \right)^{1/2},$$

where χ is a smooth bump function supported in $[-1/2, 1/2]$. Here the collection of subsets of $\widehat{\mathbb{R}^2}$ is evidently given by the strips $S_n := \{(\xi, \eta) \text{ s.t. } |\xi - \eta - n| < 1/2\}$ for $n \in \mathbb{Z}$, and the bilinear frequency projections are given by a smooth bilinear multiplier. This square function was introduced by Lacey in [13], in which he proved it is $L^p \times L^q \mapsto L^2$ bounded for $2 \leq p, q \leq \infty$ satisfying $1/p + 1/q = 1/2$. This range was later extended in [15], [6] to show that the operator is $L^p \times L^q \mapsto L^s$ bounded for $1/p + 1/q = 1/s$, $1 \leq s \leq 2$ and $2 \leq p, q \leq \infty$, this last condition being sharp. The operator admits a non-smooth variant given by

$$f, g \mapsto \left(\sum_{n \in \mathbb{Z}} \left| \iint \widehat{f}(\xi) \widehat{g}(\eta) \mathbf{1}_{[-1/2, 1/2]}(\xi - \eta - n) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^2 \right)^{1/2},$$

where the collection of subsets is again given by the strips S_n but now the bilinear frequency projections are given by a non-smooth multiplier - specifically, each projection is essentially a modulated Bilinear Hilbert Transform. The non-smoothness makes the operator inherently harder to bound. The first author proved in [5] that the operator is $L^p \times L^q \rightarrow L^s$ bounded for $2 < p, q < \infty$, $1 < s < 2$ and $1/p + 1/q = 1/s$. It should be remarked that the geometric regularity of the strips (in particular the fact that they have all the same width and separation from their neighbours) is fundamental to the proof, the case of arbitrary disjoint strips being an interesting open problem.

Another example of a bilinear square-function associated to subsets of $\widehat{\mathbb{R}^2}$ can be found in [4], where the authors considered the operator

$$f, g \mapsto \left(\sum_{n \in \mathbb{Z}} \left| \iint_{a_n < \xi < \eta < a_{n+1}} \widehat{f}(\xi) \widehat{g}(\eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^r \right)^{1/r}$$

for $r \geq 1$, where $(a_n)_{n \in \mathbb{Z}}$ is a strictly increasing subsequence of reals (a smoother version of this operator was also originally considered in [1]). They can be thought of as bilinear Rubio de Francia operators for iterated Fourier integrals, and arise naturally in the study of the stability of solutions to AKNS systems of differential equations (see [16]). Here we can see that the subsets of the frequency space $\widehat{\mathbb{R}^2}$ consist of disjoint right triangles whose hypotenuses are aligned along the $\xi = \eta$ diagonal, and the bilinear multipliers are non-smooth. In [4] it is proven that when $r \geq 2$ the operator is $L^p \times L^q \rightarrow L^s$ bounded in the same range in which the Bilinear Hilbert Transform is bounded, and when $1 \leq r < 2$ the operator is still bounded but in a range depending on r .

The final example of a bilinear operator of the above kind we provide is the closest to the operator $T_{\mathcal{Q}}^r$. Consider a collection Ω of disjoint *squares* ω with sides parallel to the axes in $\widehat{\mathbb{R}^2}$. Let χ denote a smooth bump function supported in $[-1/2, 1/2] \times [-1/2, 1/2]$ and let χ_ω be the rescaling of χ with support ω , namely

$$\chi_\omega(\xi, \eta) = \chi\left(\frac{\xi - c(\omega_1)}{|\omega_1|}, \frac{\eta - c(\omega_2)}{|\omega_2|}\right),$$

$c(I)$ denoting the center of the interval I . For $r \geq 2$ fixed, one can associate to the collection Ω the bilinear Rubio de Francia operator

$$f, g \mapsto S_\Omega^r(f, g)(x) := \left(\sum_{\omega \in \Omega} \left| \iint \hat{f}(\xi) \hat{g}(\eta) \chi_\omega(\xi, \eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^r \right)^{1/r}.$$

Thus the sets are now squares with sides parallel to the axes and the frequency projections are smooth. In [2] by Benea and the first author the following is proven.

Theorem 1.1 ([2]). *Let $r > 2$ and let Ω be a collection of disjoint squares. Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

and

$$r' < p, q < \infty, \quad r'/2 < s < r.$$

Then the operator S_Ω^r is $L^p \times L^q \rightarrow L^s$ bounded with constant independent of Ω , that is for every $f \in L^p$, $g \in L^q$ the inequality

$$\left\| \left(\sum_{\omega \in \Omega} \left| \iint \hat{f}(\xi) \hat{g}(\eta) \chi_\omega(\xi, \eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^r \right)^{1/r} \right\|_{L^s} \lesssim_{p,q} \|f\|_{L^p} \|g\|_{L^q}$$

holds.

The condition $p, q > r'$ is necessary, as can be seen by considering the collection of squares given by $\Omega = \{[n, n+1] \times [m, m+1] \text{ s.t. } m, n \in \mathbb{Z}\}$ (indeed, in this case the bilinear operator factorizes into linear smooth Rubio de Francia operators and the necessity follows from the linear case as by [11]). The condition $1/p + 1/q = 1/s$ is equally necessary, as can be seen from a simple rescaling argument. In [7] we extended the above result for squares to the case of non-smooth frequency projections, albeit in a smaller range.

Theorem 1.2 ([7]). *Let $r > 2$ and let $\Omega \subset \mathcal{D} \times \mathcal{D}$ be a collection of disjoint squares with dyadic sides. Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

and

$$r' < p, q < r, \quad r'/2 < s < r/2.$$

Then the operator T_Ω^r is $L^p \times L^q \rightarrow L^s$ bounded with constant independent of Ω , that is for every $f \in L^p$, $g \in L^q$ the inequality

$$\left\| \left(\sum_{\omega \in \Omega} \left| \iint \hat{f}(\xi) \hat{g}(\eta) \mathbb{1}_\omega(\xi, \eta) e^{2\pi i(\xi + \eta)x} d\xi d\eta \right|^r \right)^{1/r} \right\|_{L^s} \lesssim_{p,q} \|f\|_{L^p} \|g\|_{L^q}$$

holds.

Remark 1.3. First of all, notice that the condition $r'/2 < s < r/2$ is redundant in the above statement: it's a consequence of the other two conditions on p, q, s . Secondly, the larger range in Theorem 1.1 was obtained by means of a localisation argument (originally introduced in [4]) which is unavailable in the non-smooth case.

It should be remarked that the proofs of Theorems 1.1 and 1.2 both rely essentially on the sets in Ω being squares.

In this paper we will extend the result above to the case where the collection consists of arbitrary dyadic *rectangles* instead, again taking the frequency projections to be non-smooth. Precisely, our main result can be stated as follows.

Theorem 1.4. *Let $r > 2$ and let $\mathcal{R} \subset \mathcal{D} \times \mathcal{D}$ be a collection of disjoint dyadic rectangles in $\widehat{\mathbb{R}^2}$. Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

and such that

$$r' < p, q < r, \quad r'/2 < s < r/2.$$

Then the operator $T_{\mathcal{R}}^r$ is $L^p \times L^q \rightarrow L^s$ bounded with constant independent of \mathcal{R} , that is for every $f \in L^p$, $g \in L^q$ the inequality

$$\left\| \left(\sum_{R \in \mathcal{R}} |\pi_R(f, g)(x)|^r \right)^{1/r} \right\|_{L^s} \lesssim_{p,q} \|f\|_{L^p} \|g\|_{L^q}$$

holds.

A standard argument shows that we have as a corollary that the same result holds for smooth frequency projections instead, but this time with the added benefit of allowing *arbitrary* disjoint rectangles, that is with sides not necessarily dyadic in any way (but still parallel to the axes, of course).

Corollary 1.5. *Let $r > 2$ and let \mathcal{R} be a collection of disjoint arbitrary rectangles in $\widehat{\mathbb{R}^2}$ with sides parallel to the axes. Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

and such that

$$r' < p, q < r, \quad r'/2 < s < r/2.$$

Then the operator $S_{\mathcal{R}}^r$ is $L^p \times L^q \rightarrow L^s$ bounded with constant independent of \mathcal{R} , that is for every $f \in L^p$, $g \in L^q$ the inequality

$$\left\| \left(\sum_{R \in \mathcal{R}} \left| \iint \widehat{f}(\xi) \widehat{g}(\eta) \chi_R(\xi, \eta) e^{2\pi i(\xi+\eta)x} d\xi d\eta \right|^r \right)^{1/r} \right\|_{L^s} \lesssim_{p,q} \|f\|_{L^p} \|g\|_{L^q}$$

holds, where χ_R denotes

$$\chi_R(\xi, \eta) = \chi\left(\frac{\xi - c(R_1)}{|R_1|}, \frac{\eta - c(R_2)}{|R_2|}\right).$$

The proof of the corollary is elementary but has nevertheless been included in Appendix A for completeness.

The study of the boundedness of $T_{\mathcal{R}}^r(f, g)$ is reduced by duality to the study of the boundedness of the trilinear form

$$\widetilde{\Lambda}_{\mathcal{R}}^r(f, g, h) := \langle T_{\mathcal{R}}^r(f, g), h \rangle, \quad (1)$$

which can then be further reduced to the study of the boundedness of the trilinear form

$$\Lambda_{\mathcal{R}}^r(f, g, \mathbf{h}) := \int_{\mathbb{R}} \sum_{R \in \mathcal{R}} \pi_R(f, g)(x) h_R(x) dx, \quad (2)$$

where $\mathbf{h} = \{h_R\}_{R \in \mathcal{R}}$ is a vector valued function, specifically taking values in $\ell^{r'}(\mathcal{R})$. Observe moreover that the frequency projection $\pi_R(f, g)$ factorizes as $\pi_{R_1} f \cdot \pi_{R_2} g$, where $R = R_1 \times R_2$ and π_I denotes the multiplier given by $\widehat{\pi_I f} = \mathbb{1}_I \widehat{f}$.

We will prove the boundedness result stated below for the trilinear form $\Lambda_{\mathcal{R}}^r$; Theorem 1.4 then follows from the above remarks.

Theorem 1.6. *Let $r > 2$ and let $\mathcal{R} \subset \mathcal{D} \times \mathcal{D}$ be a collection of disjoint dyadic rectangles in $\widehat{\mathbb{R}^2}$. Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

and such that

$$r' < p, q < r, \quad r'/2 < s < r/2.$$

Then for every $f \in L^p$, $g \in L^q$ and $\mathbf{h} \in L^{s'}(\ell^{r'})$ we have

$$\Lambda_{\mathcal{R}}^r(f, g, \mathbf{h}) \lesssim_{r,p,q} \|f\|_{L^p} \|g\|_{L^q} \|\mathbf{h}\|_{L^{s'}(\ell^{r'})}. \quad (3)$$

Let $\text{ec}(R)$ denote the eccentricity of the rectangle R , defined as

$$\text{ec}(R) := \frac{|R_2|}{|R_1|}.$$

We split the collection \mathcal{R} into two subcollections, according to whether the eccentricity is high or low:

$$\mathcal{R}_{\text{high}} := \{R \in \mathcal{R} \text{ s.t. } \text{ec}(R) > 1\},$$

$$\mathcal{R}_{\text{low}} := \{R \in \mathcal{R} \text{ s.t. } \text{ec}(R) \leq 1\};$$

of course, the rectangles of eccentricity 1 are just squares and as such we could remove them from the collection since they give rise to a bounded operator, by Theorem 1.2. We then split the trilinear form into $\Lambda_{\mathcal{R}}^r = \Lambda_{\mathcal{R}_{\text{high}}}^r + \Lambda_{\mathcal{R}_{\text{low}}}^r$. We will prove Theorem 1.6 separately for $\mathcal{R}_{\text{high}}, \mathcal{R}_{\text{low}}$, using a time-frequency analysis of the trilinear form, and since the proof will be symmetric in the two cases we will concentrate exclusively on $\mathcal{R}_{\text{high}}$ (see however Remark 7.4 for further details on this symmetry).

Let us briefly explain the structure of the proof by analogy with the linear case of Rubio de Francia operators $\text{RdF}_{\mathcal{I}}^r$, where $\mathcal{I} = \{I_n, n \in \mathbb{Z}\}$ is a given collection of disjoint intervals. Recall that the inequality $\|\text{RdF}_{\mathcal{I}}^2 f\|_{L^2} \leq \|f\|_{L^2}$ is a trivial consequence of Plancherel's inequality; it can be rephrased in a vector-valued flavour as $\|(\pi_{I_n} f)_{n \in \mathbb{Z}}\|_{L^2(\ell^2)} \leq \|f\|_{L^2}$. On the other hand, observe that pointwise $|\pi_{I_n} f(x)| \leq 2\mathcal{C}f(x)$ independently of n , where \mathcal{C} denotes the Carleson operator

$$\mathcal{C}f(x) := \sup_{N \in \mathbb{R}} \left| \int_{-\infty}^N \widehat{f}(\xi) e^{2\pi i \xi x} d\xi \right|,$$

and therefore we have by the Carleson-Hunt theorem that $\|(\pi_{I_n} f)_{n \in \mathbb{Z}}\|_{L^p(\ell^\infty)} \lesssim_p \|f\|_{L^p}$ for any $1 < p < \infty$. By complex interpolation it follows immediately that

$$\|\text{RdF}_{\mathcal{I}}^r f\|_{L^p} \lesssim_{r,p} \|f\|_{L^p}$$

for $r' < p < r$. As stated before, the operator is also bounded in the range $p \geq r > 2$, but this is a consequence of the boundedness of $\text{RdF}_{\mathcal{I}}^2$ in $p \geq 2$, and the proof of this fact requires methods other than simple interpolation.

In the proof of Theorem 1.4 we have adopted a similar approach. Indeed, when $r = \infty$ the bilinear operator becomes

$$T_{\mathcal{R}}^\infty(f, g)(x) = \sup_{R \in \mathcal{R}} |\pi_R(f, g)(x)|,$$

and this is easily seen to be bounded pointwise by $\mathcal{C}f \cdot \mathcal{C}g$, in parallel to the linear case, which in turn yields the full range of boundedness for $T_{\mathcal{R}}^\infty$. Now, the natural bilinear analogue of the $L^2 \rightarrow L^2$ estimate would be $L^2 \times L^2 \rightarrow L^1$, and thanks to an interpolation result of Silva from [20] (see Lemma 8.6) Theorem 1.6 would follow from such an estimate for $r = 2$. However, there is no equivalent of Plancherel's theorem in the bilinear world, and this estimate is not straightforward. We currently

do not know if such an estimate holds, the $r = 2$ case having resisted treatment in general even in the simpler case of squares and smooth frequency projections. As a replacement, we will prove preliminary $L^p \times L^q \rightarrow L^s$ estimates for $T_{\mathcal{R}}^r$ for each $r > 2$, in a range that gets arbitrarily close to $L^2 \times L^2 \rightarrow L^1$ as $r \rightarrow 2$. This will be enough to allow us to run the interpolation argument to conclude boundedness in the range stated in Theorem 1.4.

Remark 1.7. The overall proof structure explained above is the same the authors have adopted in the proof of Theorem 1.2 in [7], which in turn was inspired by the previous work for smooth squares as in [2] and drew insights from [1, 3], in particular the idea of using non-local operators in the stopping-time arguments. The main difference here lies in the fact that we now have to deal with the simultaneous presence of two distinct scales in each $\pi_R(f, g)$ term, namely those given by $|R_1|$ and $|R_2|$, which breaks the symmetry in how the contributions of f and g are treated. See particularly sections §2, §3 for how this is achieved.

The rest of the paper is structured as follows. In section §2, under the hypothesis that all rectangles have high eccentricity, we will reduce the problem to that of bounding a discretized trilinear sum with additional structure of certain (shifted) averages of f, g, \mathbf{h} . We will introduce a notion of time-frequency tiles adapted to the particular situation. Sections §3, §4, §5 and §6 develop the time-frequency analysis tools needed to prove the main result. In particular, in section §3 we will introduce some structured collections of tiles referred to as \mathbf{n} -columns¹, where \mathbf{n} denotes a shifting parameter, for which we will be able to bound the discretized trilinear sum explicitly. This process will give rise to sizes for f and \mathbf{h} , but importantly not for g . Indeed, we will not try to optimize our estimates in g , and rather prove in section §7 some weaker estimates for the discretized trilinear sum in a partial range contained in $p \in (2, r)$ and $q = r$. In section §8 we will improve these estimates by interpolating the partial results with the case $r = \infty$, thus concluding the proof of Theorem 1.4.

Acknowledgements. Both authors are supported by ERC project FAnFArE no. 637510. The authors are also very grateful to Cristina Benea for many useful comments/discussions and in particular for having shared with us a preprint of [3].

2. REDUCTION TO A DISCRETIZED MODEL SUM

Recall that we are considering the collection $\mathcal{R}_{\text{high}}$ of rectangles of eccentricity bigger than 1. We start by performing a standard discretization procedure on the trilinear form $\Lambda_{\mathcal{R}_{\text{high}}}^r$.

We have (using Radon duality, with $d\sigma$ the induced Lebesgue measure on the plane $\xi_1 + \xi_2 + \xi_3 = 0$)

$$\begin{aligned} \Lambda_{\mathcal{R}_{\text{high}}}^r(f, g, \mathbf{h}) &= \int_{\mathbb{R}} \sum_{R \in \mathcal{R}_{\text{high}}} \pi_{R_1} f(x) \pi_{R_2} g(x) h_R(x) dx \\ &= \sum_{R \in \mathcal{R}_{\text{high}}} \int_{\xi_1 + \xi_2 + \xi_3 = 0} \widehat{f}(\xi_1) \mathbb{1}_{R_1}(\xi_1) \widehat{g}(\xi_2) \mathbb{1}_{R_2}(\xi_2) \widehat{h_R}(\xi_3) d\sigma(\xi_1, \xi_2, \xi_3) \\ &= \sum_{R \in \mathcal{R}_{\text{high}}} \int_{\xi_1 + \xi_2 + \xi_3 = 0} \widehat{f}(\xi_1) \mathbb{1}_{R_1}(\xi_1) \widehat{g}(\xi_2) \mathbb{1}_{R_2}(\xi_2) \widehat{h_R}(\xi_3) \chi_{R_3}(\xi_3) d\sigma \\ &= \sum_{R \in \mathcal{R}_{\text{high}}} \int_{\mathbb{R}} f * \check{\mathbb{1}}_{R_1}(x) g * \check{\mathbb{1}}_{R_2}(x) h_R * \check{\chi}_{R_3}(x) dx, \end{aligned}$$

¹These play the analogous rôle played by trees, in classical time-frequency analysis terminology.

where we have denoted $R_3 := 2(-R_1 - R_2)$ and χ_{R_3} is a C^∞ bump function adapted to R_3 and identically equal to 1 on $-R_1 - R_2$. Notice that $|R_3| \sim |R_2|$ by the assumption that R is of high eccentricity. Now, since the functions $f * \mathbf{1}_{R_j}$ are morally roughly constant in modulus at scale $|R_j|^{-1}$, we do the following change of variables with respect to the smallest time-scale involved, the one given by $|R_2|^{-1}$ (indeed, we are assuming $|R_2| > |R_1|$):

$$\begin{aligned} & \sum_{R \in \mathcal{R}_{\text{high}}} \int_{\mathbb{R}} f * \mathbf{1}_{R_1}(x) g * \mathbf{1}_{R_2}(x) h_R * \check{\chi}_{R_3}(x) dx \\ &= \sum_{R \in \mathcal{R}_{\text{high}}} |R_2|^{-1} \int_{\mathbb{R}} f * \mathbf{1}_{R_1}(|R_2|^{-1}y) g * \mathbf{1}_{R_2}(|R_2|^{-1}y) h_R * \check{\chi}_{R_3}(|R_2|^{-1}y) dy \\ &= \sum_{R \in \mathcal{R}_{\text{high}}} \sum_{n \in \mathbb{Z}} |R_2|^{-1} \int_0^1 \pi_{R_1} f(|R_2|^{-1}(n+z)) \pi_{R_2} g(|R_2|^{-1}(n+z)) \\ & \quad \cdot h_R * \check{\chi}_{R_3}(|R_2|^{-1}(n+z)) dz, \end{aligned}$$

where we have gone back to writing $\pi_\omega f$ for $f * \mathbf{1}_\omega$ in the last line.

Remark 2.1. Observe that $|\pi_{R_2} g|$ is morally roughly constant at scale $|R_2|^{-1}$, thus morally roughly constant on the intervals $|R_2|^{-1}(n + [0, 1])$ (so is $|h_R * \check{\chi}_{R_3}|$ since $|R_3| \sim |R_2|$ for $R \in \mathcal{R}_{\text{high}}$). However, $|\pi_{R_1} f|$ is morally roughly constant at the (generally) larger scale $|R_1|^{-1}$. The presence of two different simultaneous scales is the major source of difficulty in the analysis and it is here that the difference with the proof of Theorem 1.2 will rely.

In light of the above remark, we want to take advantage of the fact that $|\pi_{R_1} f|$ is morally roughly constant at a larger scale in space to reduce to a trilinear form with additional structure. In order to explain this reduction properly we will make use of two types of tiles (mirroring the existence of two simultaneous scales in each $\pi_R(f, g)$), the first of which we introduce now.

Definition 2.2. A *small tile* ρ is a pair of the form

$$(\omega_2 \times I_\rho, \omega_3 \times I_\rho)$$

where I_ρ, ω_2 are dyadic intervals such that there exists a rectangle $R \in \mathcal{R}$ with $\omega_2 = R_2, \omega_3 = R_3$ (recall $R_3 := 2(-R_1 - R_2)$) and $|\omega_2||I_\rho| = 1$ (hence $|\omega_3||I_\rho| \sim 1$ as well; ω_3 is in general not dyadic). Given a small tile ρ we denote by $R(\rho)$ this unique rectangle R . The collection of all small tiles is denoted by \mathbb{S} .

Given $R \in \mathcal{R}_{\text{high}}$ and $n \in \mathbb{Z}$ there exists a unique small tile $\rho = \rho(R, n)$ such that $R(\rho) = R$ and $I_\rho = |R_2|^{-1}[n, n+1]$, and viceversa. Because of this we can rewrite the trilinear form as

$$\Lambda_{\mathcal{R}_{\text{high}}}^r(f, g, \mathbf{h}) = \sum_{\rho \in \mathbb{S}} \int_{I_\rho} \pi_{R_1(\rho)} f(x) \pi_{R_2(\rho)} g(x) h_{R(\rho)} * \check{\chi}_{R_3(\rho)}(x) dx;$$

by Hölder's inequality this is controlled by

$$\sum_{\rho \in \mathbb{S}} \|\pi_{R_1(\rho)} f\|_{L^2(I_\rho)} \|\pi_{R_2(\rho)} g\|_{L^2(I_\rho)} \|h_{R(\rho)} * \check{\chi}_{R_3(\rho)}\|_{L^\infty(I_\rho)}.$$

Now we will use the fact that $|\pi_{R_1} f|$ is roughly constant on many I_ρ 's. Let ρ be fixed and let n be such that $I_\rho = |R_2|^{-1}[n, n+1]$. There exists a unique $k \in \mathbb{Z}$ such that $n = k \cdot \text{ec}(R) + \ell$ with $0 \leq \ell < \text{ec}(R)$, and thus we can associate uniquely to ρ the interval $I = |R_1|^{-1}[k, k+1]$. This is a dyadic interval of length $|R_1|^{-1}$ and it has the property that $I_\rho \subset I$.

Given the interval ω let χ_ω denote a smooth function such that $\chi_\omega(\xi) = 1$ for $\xi \in \omega$ and $\chi_\omega(\xi) = 0$ for $\xi \notin 2\omega$. By definition of the frequency projections we then have $\tilde{\chi}_{R_1} * (\pi_{R_1} f) = \pi_{R_1} f$, and thus we can bound for every $N > 0$

$$|\pi_{R_1(\rho)} f(x)| \lesssim_N \int |\pi_{R_1(\rho)} f(y)| \frac{1}{|R_1|^{-1}} \left(1 + \frac{|x-y|}{|R_1|^{-1}}\right)^{-N} dy.$$

Therefore we see that we can bound (by Minkowski's inequality)

$$\begin{aligned} \frac{\|\pi_{R_1(\rho)} f\|_{L^2(I_\rho)}}{|I_\rho|^{1/2}} &\lesssim \int |\pi_{R_1(\rho)} f(y)| \sum_{\mathbf{n} \in \mathbb{Z}} (1 + |\mathbf{n}|)^{-N} \frac{\mathbb{1}_{I^\mathbf{n}}(y)}{|I|} dy \\ &= \sum_{\mathbf{n} \in \mathbb{Z}} (1 + |\mathbf{n}|)^{-N} \int_{I^\mathbf{n}} |\pi_{R_1(\rho)} f|, \end{aligned}$$

where I is as above (in particular $I \supset I_\rho$) and

$$I^\mathbf{n} := I + \mathbf{n}|I|,$$

that is $I^\mathbf{n}$ denotes the interval I shifted by \mathbf{n} times its length.

With this in mind, we now define the second type of tiles.

Definition 2.3. Given a rectangle $R \in \mathcal{R}$ and a dyadic interval I , we define the (possibly empty) collection of small tiles $\mathbb{S}_{R,I}^\mathbf{n}$ to be

$$\mathbb{S}_{R,I}^\mathbf{n} := \{\rho = (\omega_2 \times I_\rho, \omega_3 \times I_\rho) \text{ small tile : } \omega_2 = R_2, \omega_3 = R_3, I_\rho \subseteq I^\mathbf{n}\}.$$

A *super tile* P is a pair of the form

$$(R_1 \times I, \mathbb{S}_{R,I}^\mathbf{n})$$

where $R \in \mathcal{R}$ and I is a dyadic interval such that $|R_1||I| = 1$ (notice that the pair (R, I) completely determines the super tile).

Given P as above, we let $P_1 := R_1 \times I$ and call it simply a *tile*, $I_P := I$, $R(P) := R$ (and analogously $R_j(P) := R_j$ for $j \in \{1, 2, 3\}$) and $\mathbb{S}_P^\mathbf{n} := \mathbb{S}_{R,I}^\mathbf{n}$. Finally, if \mathbb{P} is a collection of super tiles, we let $\mathcal{R}(\mathbb{P}) := \{R(P) : P \in \mathbb{P}\}$.

Now fix a parameter r_0 such that $r > r_0 > 2$. This will remain fixed throughout sections §3 - §7 (ultimately we will take r_0 to depend on r ; one could take for example $r_0 = (r + 2)/2$). Using the fact that $\int_I |\pi_{R_1} f| \leq (\int_I |\pi_{R_1} f|^{r_0})^{1/r_0}$ and the above remarks we can finally bound

$$|\Lambda_{\mathcal{R}_{\text{high}}}^r(f, g, \mathbf{h})| \lesssim_N \sum_{\mathbf{n} \in \mathbb{Z}} (1 + |\mathbf{n}|)^N \Lambda_{\mathbb{P}}^\mathbf{n}(f, g, \mathbf{h}),$$

where $\Lambda_{\mathbb{P}}^\mathbf{n}$ denote the *shifted* trilinear forms

$$\begin{aligned} \Lambda_{\mathbb{P}}^\mathbf{n}(f, g, \mathbf{h}) &:= \sum_{P \in \mathbb{P}} \left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} \left[\sum_{\rho \in \mathbb{S}_P^\mathbf{n}} |I_\rho|^{1/2} \|\pi_{R_2(P)} g\|_{L^2(I_\rho)} \right. \\ &\quad \left. \times \|h_{R(P)} * \tilde{\chi}_{R_3(P)}\|_{L^\infty(I_\rho)} \right]. \end{aligned} \tag{4}$$

To bound $\Lambda_{\mathcal{R}_{\text{high}}}^r(f, g, \mathbf{h})$ it will therefore suffice to bound the trilinear forms $\Lambda_{\mathbb{P}}^\mathbf{n}(f, g, \mathbf{h})$ in the same range with a constant that is at most polynomial in \mathbf{n} (indeed, we will be able to show it is at most logarithmic).

In the following sections we will consider \mathbf{n} fixed and study the trilinear forms $\Lambda_{\mathbb{P}}^\mathbf{n}$.

Remark 2.4. It might seem that we are introducing a gratuitous inefficiency in our argument by replacing the L^1 -average $\int_I |\pi_{R_1} f|$ with the larger $(\int_I |\pi_{R_1} f|^{r_0})^{1/r_0}$, but this will not be the case. This is due to our use of interpolation in the proof of Theorem 1.6: for the argument to carry through we will only need to prove boundedness of $\Lambda_{\mathcal{R}_{\text{high}}}^r$ in any range sufficiently close to the $L^2 \times L^2 \rightarrow L^1$ estimate,

for all given $r > 2$ (see Section 8). Using L^1 -averages would not enlarge the range obtained through interpolation, at least not with our argument.

3. \mathbf{n} -COLUMNS AND SIZES

We will now introduce some special collections of super tiles for which we will be able to bound the (shifted) trilinear form explicitly. In the process, we will come up with quantities that will be used as sizes in order to perform a time-frequency analysis of the general trilinear form.

First we define an order relation on the super tiles.

Definition 3.1. Let P, P' be super tiles and let \mathbf{n} be fixed. Then we say that $P <_{\mathbf{n}} P'$ if

$$R_1(P) \supset R_1(P'), \\ I_P^{\mathbf{n}} \subset I_{P'}^{\mathbf{n}}.$$

We say that $P \leq_{\mathbf{n}} P'$ if $P <_{\mathbf{n}} P'$ or $P = P'$.

Definition 3.2. An \mathbf{n} -column \mathcal{C} with top a super tile P_{top} is a collection of super tiles such that for every $P \in \mathcal{C}$

$$P \leq_{\mathbf{n}} P_{\text{top}}.$$

We denote $\text{Top}(\mathcal{C}) = P_{\text{top}}$.

Remark 3.3. Notice the rôle of \mathbf{n} in the above definitions. The difference between the above \mathbf{n} -columns and the columns originally defined in [2] resides in the fact that there is now the shifting parameter \mathbf{n} to be taken into account. Thus, unlike in [2], the tiles P_1 for $P \in \mathcal{C}$, that is the sets $R_1(P) \times I_P$, do not form an overlapping tree - unless $\mathbf{n} = 0$. See figure 1 for further clarification.

The following lemma regarding the structure of an \mathbf{n} -column is essentially obvious once one unpacks the definitions; it will find its application in the discussion that follows.

Lemma 3.4. Let \mathcal{C} be a column with top P_{top} . Then for every super tile $P \in \mathcal{C}$ we have

$$\forall \rho \in \mathbb{S}_P^{\mathbf{n}}, \quad I_\rho \subset I_{P_{\text{top}}}^{\mathbf{n}}.$$

Moreover, if $R, R' \in \mathcal{R}(\mathcal{C})$ are distinct, then we have $R_2 \cap R'_2 = \emptyset$.

In other words, the small tiles belonging to an \mathbf{n} -column agree with the time-frequency portrait of a column as in [2]. The second property in the lemma is due to the disjointness of the rectangles in \mathcal{R} . See figure 1 for a pictorial representation of an \mathbf{n} -column.

Let then \mathcal{C} be an \mathbf{n} -column and consider the trilinear form $\Lambda_{\mathcal{C}}^{\mathbf{n}}(f, g, \mathbf{h})$. First of all, we fix a super tile P in \mathcal{C} and look at the inner sum in the small tiles ρ in (4): we have by Hölder inequality

$$\begin{aligned} & \sum_{\rho \in \mathbb{S}_P^{\mathbf{n}}} |I_\rho|^{1/2} \|\pi_{R_2(P)} g\|_{L^2(I_\rho)} \|h_{R(P)} * \widetilde{\chi_{R_3(P)}}\|_{L^\infty(I_\rho)} \\ & \leq \left(\sum_{\rho \in \mathbb{S}_P^{\mathbf{n}}} |I_\rho|^{1-r/2} \|\pi_{R_2(P)} g\|_{L^2(I_\rho)}^r \right)^{1/r} \left(\sum_{\rho \in \mathbb{S}_P^{\mathbf{n}}} |I_\rho| \|h_{R(P)} * \widetilde{\chi_{R_3(P)}}\|_{L^\infty(I_\rho)}^{r'} \right)^{1/r'}. \end{aligned}$$

For the term in g , by Lemma 3.4 we can bound by Hölder's inequality

$$\sum_{\rho \in \mathbb{S}_P^{\mathbf{n}}} |I_\rho|^{1-r/2} \|\pi_{R_2(P)} g\|_{L^2(I_\rho)}^r = \sum_{\rho \in \mathbb{S}_P^{\mathbf{n}}} |I_\rho| \left(\int_{I_\rho} |\pi_{R_2(P)} g|^2 \right)^{r/2} \leq \int_{I_P^{\mathbf{n}}} |\pi_{R_2(P)} g|^r,$$

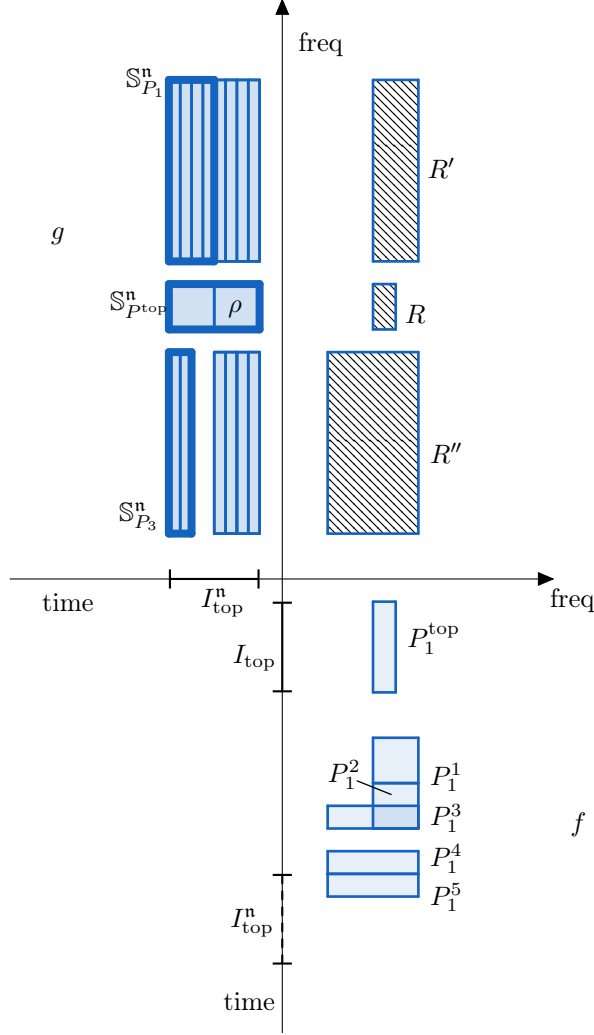


FIGURE 1. A composite pictorial representation of an n -column \mathcal{C} for $n = 3$. Quadrant I represents the frequency plane, quadrant II represents the time-frequency picture of g , quadrant IV represents the (rotated) time-frequency picture of f . The column consists of supertiles $\{P^{top}, P^1, \dots, P^5\}$ and has top P^{top} , whose time interval is I_{top} . The dashed rectangles in quadrant I are the frequency rectangles: $\mathcal{R}(\mathcal{C}) = \{R, R', R''\}$. The rectangles in quadrant II are small tiles (each has area 1). The thicker rectangles in quadrant II represent the union of all small tiles in \mathbb{S}_P^n for $P = P^{top}, P^1, P^3$. The rectangles in quadrant IV are the tiles $P_1^{top}, P_1^1, \dots, P_1^5$ as labeled (each has area 1). Observe that $R(P^{top}) = R$, $R(P^1) = R(P^2) = R'$, $R(P^3) = R(P^4) = R(P^5) = R''$. Notice also that for any $j = 1, \dots, 5$ we have $I_{P_1^j}^n \subset I_{top}^n$, as per definition.

because $r > 2$.

For the term in \mathbf{h} instead, we see that we have for any small tile ρ and any large

$M > 0$

$$\begin{aligned}
|I_\rho| \|h_{R(P)} * \widetilde{\chi_{R_3(P)}}\|_{L^\infty(I_\rho)}^{r'} &= |I_\rho| \sup_{y \in I_\rho} |h_R * \widetilde{\chi_{R_3}}(y)|^{r'} \\
&\leq |I_\rho| \left(\sup_{y \in I_\rho} \int |h_R(z)| |\widetilde{\chi_{R_3}}(y-z)| dz \right)^{r'} \\
&\lesssim |I_\rho| \left(\sup_{y \in I_\rho} \int |h_R(z)| \left(1 + \frac{|y-z|}{|I_\rho|}\right)^{-M} \frac{dz}{|I_\rho|} \right)^{r'} \\
&\leq |I_\rho| \left(\int |h_R(z)| \sup_{y \in I_\rho} \left(1 + \frac{|y-z|}{|I_\rho|}\right)^{-M} \frac{dz}{|I_\rho|} \right)^{r'} \\
&\lesssim \int |h_R(z)|^{r'} \Phi_{I_\rho}(z) dz,
\end{aligned}$$

where Φ_I denotes a rapidly decaying function concentrated in I and equal to 1 there. In particular, Φ_I decays like $(1 + \text{dist}(z, I)/|I|)^{-N}$ for some large N . Therefore, summing over $\rho \in \mathbb{S}_P^n$, we have

$$\sum_{\rho \in \mathbb{S}_P^n} |I_\rho| \|h_{R(P)} * \widetilde{\chi_{R_3(P)}}\|_{L^\infty(I_\rho)}^{r'} \lesssim \int |h_{R(P)}(z)|^{r'} \Phi_{I_P^n}(z) dz.$$

With these preliminary estimates at hand, we proceed to estimate $\Lambda_{\mathcal{C}}^n(f, g, \mathbf{h})$. Write I_{top} in place of $I_{\text{Top}(C)}$ for shortness. We take the supremum in the terms in f and use Hölder's inequality with exponents (r, r') once again, this time in the super tiles, thus obtaining

$$\begin{aligned}
\Lambda_{\mathcal{C}}^n(f, g, \mathbf{h}) &\lesssim \sup_{P \in \mathcal{C}} \left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} \left[\frac{1}{|I_{\text{top}}|} \sum_{P \in \mathcal{C}} \int_{I_P^n} |\pi_{R_2(P)} g|^r \right]^{1/r} \\
&\quad \times \left(\frac{1}{|I_{\text{top}}|} \sum_{P \in \mathcal{C}} \int |h_{R(P)}|^{r'} \Phi_{I_P^n} \right)^{1/r'} |I_{\text{top}}|.
\end{aligned}$$

We then define sizes for f and \mathbf{h} according to the above estimate.

Definition 3.5 (Sizes). Let \mathbb{P} be a collection of super tiles, and let $f \in L_{\text{loc}}^1(\mathbb{R})$, $\mathbf{h} \in L_{\text{loc}}^1(\ell^{r'})$. Then we define

$$\text{Size}_f^n(\mathbb{P}) := \sup_{P \in \mathbb{P}} \left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0},$$

and

$$\text{Size}_{\mathbf{h}}^n(\mathbb{P}) := \sup_{\substack{\mathcal{C} \text{ n-column,} \\ \mathcal{C} \subset \mathbb{P}}} \left(\frac{1}{|I_{\text{Top}(\mathcal{C})|} \sum_{P \in \mathcal{C}} \int |h_{R(P)}|^{r'} \Phi_{I_P^n} \right)^{1/r'}.$$

Notice that, strictly speaking, Size_f^n does not depend on the shifting parameter \mathbf{n} . However, the associated energy will and thus we keep the index \mathbf{n} as a reminder. We can then summarise the above discussion in the following proposition.

Proposition 3.6. *Let \mathcal{C} be an \mathbf{n} -column. Then we can bound*

$$\Lambda_{\mathcal{C}}^n(f, g, \mathbf{h}) \lesssim \left[\frac{1}{|I_{\text{Top}(\mathcal{C})|} \sum_{P \in \mathcal{C}} \int_{I_P^n} |\pi_{R_2(P)} g|^r \right]^{1/r} \text{Size}_f^n(\mathcal{C}) \text{Size}_{\mathbf{h}}^n(\mathcal{C}) |I_{\text{Top}(\mathcal{C})|}.$$

As explained in the overview of the proof, we will not introduce a size for g . Control of the contribution of g will rather be achieved through the following lemma.

Lemma 3.7. *If \mathcal{C} is an \mathbf{n} -column, then we have*

$$\frac{1}{|I_{\text{Top}(\mathcal{C})|} \sum_{P \in \mathcal{C}} \int_{I_P^n} |\pi_{R_2(P)} g|^r \lesssim \int_{I_{\text{Top}(\mathcal{C})}^n} (\text{Var}^r \mathcal{C} g(x))^r dx,$$

where $\text{Var}^r \mathcal{C}$ is the Variational Carleson operator given by

$$\text{Var}^r \mathcal{C} g(x) := \sup_N \sup_{\xi_1 < \dots < \xi_N} \left(\sum_{j=1}^{N-1} |\pi_{[\xi_j, \xi_{j+1}]} g(x)| \right)^{1/r}.$$

Proof. By the definition of \mathbf{n} -column, we have that $I_P^n \subset I_{\text{Top}(\mathcal{C})}^n$ for all $P \in \mathcal{C}$, and moreover if $P, P' \in \mathcal{C}$ are such that $R_2(P) \neq R_2(P')$ we necessarily have $R_2(P) \cap R_2(P') = \emptyset$ (see Lemma 3.4). In other words, the “shifted tiles” $R_2(P) \times I_P^n$ are all disjoint and contained in the strip $\mathbb{R} \times I_{\text{Top}(\mathcal{C})}^n$. If we then rewrite

$$\frac{1}{|I_{\text{Top}(\mathcal{C})}|} \sum_{P \in \mathcal{C}} \int_{I_P^n} |\pi_{R_2(P)} g|^r = \frac{1}{|I_{\text{Top}(\mathcal{C})}|} \int_{I_{\text{Top}(\mathcal{C})}^n} \sum_{P \in \mathcal{C}} |\pi_{R_2(P)} g(x)|^r \mathbb{1}_{I_P^n}(x) dx$$

we see that we can bound the integrand pointwise by $(\text{Var}^r \mathcal{C} g(x))^r$, and we are done. \square

Remark 3.8. The above lemma relies crucially on the structure of the \mathbf{n} -columns to hold; thus we can see that the lemma provides motivation for our definition.

Finally, we conclude this section with the observation that the sizes introduced are controlled by maximal averages, which will allow us to perform the time-frequency analysis.

Lemma 3.9. *Let \mathbb{P} be a collection of super tiles. Then we have*

$$\text{Size}_f^n(\mathbb{P}) \lesssim \sup_{P \in \mathbb{P}} \left(\int_{I_P} |\mathcal{C} f|^{r_0} \right)^{1/r_0},$$

where \mathcal{C} is the Carleson operator.

Proof. Obvious. \square

Lemma 3.10. *Let \mathbb{P} be a collection of super tiles. Then we have*

$$\text{Size}_h^n(\mathbb{P}) \lesssim \sup_{P \in \mathbb{P}} \left(\frac{1}{|I_P|} \int \|\mathbf{h}\|_{\ell^{r'}}^{r'} \Phi_{I_P^n} \right)^{1/r'}.$$

Proof. The claim follows quickly from the definition of an \mathbf{n} -column and the rapid decay of the functions $\Phi_{I_P^n}$. \square

4. ENERGIES

In this section we will introduce the remaining quantities that we need in order to enable the time-frequency analysis of $\Lambda_{\mathbb{P}}^n$. It will be necessary to introduce a special notion of disjointness for \mathbf{n} -columns (quite similar to that in [2]) in order to ensure good control of the quantities we are going to introduce, which is as follows.

Definition 4.1. Let \mathfrak{C} be a collection of \mathbf{n} -columns. We say that the \mathbf{n} -columns in \mathfrak{C} are *mutually disjoint* if

- i) for every $\mathcal{C}, \mathcal{C}' \in \mathfrak{C}$ we have $\mathcal{C} \cap \mathcal{C}' = \emptyset$ (that is, the \mathbf{n} -columns are disjoint as sets of super tiles);
- ii) any two tops are not comparable under the \leq_n relation, or equivalently the sets $R_1(\text{Top}(\mathcal{C})) \times I_{\text{Top}(\mathcal{C})}^n$ for $\mathcal{C} \in \mathfrak{C}$ are pairwise disjoint.

Notice the presence of the shifting parameter \mathbf{n} in the above definition. It is not the tiles $\text{Top}(\mathcal{C})_1$ that are assumed to be disjoint sets in time-frequency, but rather their shifted versions.

With this notion of mutual disjointness we can now define the energies.

Definition 4.2 (Energy of f). Let \mathbb{P} be a collection of super tiles, and let $f \in L^1_{\text{loc}}(\mathbb{R})$. We define

$$\text{Energy}_f^{\mathbf{n}}(\mathbb{P}) := \sup_{n \in \mathbb{Z}} \sup_{\mathfrak{C}} 2^n \left(\sum_{\mathcal{C} \in \mathfrak{C}} |I_{\text{Top}(\mathcal{C})}| \right)^{1/r_0},$$

where the inner supremum ranges over all collections \mathfrak{C} of mutually disjoint \mathbf{n} -columns \mathcal{C} such that for each $\mathcal{C} \in \mathfrak{C}$ and each $P \in \mathcal{C}$

$$\left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} \geq 2^n.$$

Definition 4.3 (Energy of \mathbf{h}). Let \mathbb{P} be a collection of super tiles, and let $\mathbf{h} \in L^1_{\text{loc}}(\ell^{r'})$. We define

$$\text{Energy}_{\mathbf{h}}^{\mathbf{n}}(\mathbb{P}) := \sup_{n \in \mathbb{Z}} \sup_{\mathfrak{C}} 2^n \left(\sum_{\mathcal{C} \in \mathfrak{C}} |I_{\text{Top}(\mathcal{C})}| \right)^{1/r'},$$

where the inner supremum ranges over all collections \mathfrak{C} of mutually disjoint \mathbf{n} -columns \mathcal{C} such that for each $\mathcal{C} \in \mathfrak{C}$

$$\left(\frac{1}{|I_{\text{Top}(\mathcal{C})}|} \sum_{P \in \mathcal{C}} \int |h_{R(P)}|^{r'} \Phi_{I_P^{\mathbf{n}}} \right)^{1/r'} \geq 2^n.$$

Remark 4.4. Observe that $\text{Energy}_f^{\mathbf{n}}$ is an L^{r_0} quantity, while $\text{Energy}_{\mathbf{h}}^{\mathbf{n}}$ is an $L^{r'}$ quantity. The fact that $r' < 2 < r_0$ will be fundamental in summing up the energies coming from a decomposition of \mathbb{P} into \mathbf{n} -columns (see Proposition 6.1).

We now show how to control the energies in terms of L^p norms of the functions f, \mathbf{h} . We start with the latter.

Lemma 4.5. *Let \mathbb{P} be a collection of super tiles and let $\mathbf{h} \in L^{r'}(\ell^{r'})$. Then we have*

$$\text{Energy}_{\mathbf{h}}^{\mathbf{n}}(\mathbb{P}) \lesssim \|\mathbf{h}\|_{L^{r'}(\ell^{r'})}.$$

Proof. This lemma is essentially the same as Proposition 2.10 in [7].

Let $n \in \mathbb{Z}$ and \mathfrak{C} be a collection of mutually disjoint \mathbf{n} -columns such that the pair realizes the supremum in the definition of $\text{Energy}_{\mathbf{h}}^{\mathbf{n}}$ within a factor of 2, that is

$$\text{Energy}_{\mathbf{h}}^{\mathbf{n}}(\mathbb{P})^{r'} \sim 2^{r'n} \sum_{\mathcal{C} \in \mathfrak{C}} |I_{\text{Top}(\mathcal{C})}|.$$

Then the right hand side is dominated by

$$\sum_{\mathcal{C} \in \mathfrak{C}} \sum_{P \in \mathcal{C}} \int |h_{R(P)}|^{r'} \Phi_{I_P^{\mathbf{n}}},$$

which in turn is dominated by

$$\sum_{R \in \mathcal{R}} \int |h_R|^{r'} \left(\sum_{\mathcal{C} \in \mathfrak{C}} \sum_{\substack{P \in \mathcal{C}: \\ R(P)=R}} \Phi_{I_P^{\mathbf{n}}} \right).$$

Observe that as P ranges over the super tiles such that $R(P) = R$, the intervals $I_P^{\mathbf{n}}$ are all disjoint; as the \mathbf{n} -columns are mutually disjoint we then have that

$$\sum_{\mathcal{C} \in \mathfrak{C}} \sum_{\substack{P \in \mathcal{C}: \\ R(P)=R}} \Phi_{I_P^{\mathbf{n}}} \lesssim 1$$

and the conclusion follows. \square

Next we will show how to control $\text{Energy}_f^{\mathbf{n}}$ in terms of $\|f\|_{L^{r_0}}$. Here the presence of the shift parameter will produce an unavoidable logarithmic loss in \mathbf{n} in the inequality, which will however be more than acceptable for our purposes.

Lemma 4.6. *Let \mathbb{P} be a collection of super tiles and let $f \in L^{r_0}(\mathbb{R})$. Then*

$$\text{Energy}_f^n(\mathbb{P}) \lesssim \log^+(\mathbf{n}) \|f\|_{L^{r_0}},$$

where $\log^+(x) = \log(2 + |x|)$.

Proof. Let $n \in \mathbb{Z}$ and \mathfrak{C} be a collection (which we can assume to be finite) of \mathbf{n} -columns such that they realize the supremum in the definition of Energy_f^n within a factor of 2, that is

$$\text{Energy}_f^n(\mathbb{P})^{r_0} \sim 2^{r_0 n} \sum_{\mathcal{C} \in \mathfrak{C}} |I_{\text{Top}(\mathcal{C})}|.$$

We'd like to argue by pointwise control by a Variational Carleson operator as in Lemma 3.7, but we can't do so at this stage because, by the nature of the definition of mutually disjoint \mathbf{n} -columns, the tops $\text{Top}(\mathcal{C})_1$ are not disjoint in general. However, we will show that up to introducing a loss of $\log^+(\mathbf{n})$ in the estimate, we can reduce to the disjoint case. To do so it is convenient to introduce a second ordering on the super tiles (distinct from the previously defined order $<_{\mathbf{n}}$).

Definition 4.7. Let P, P' be super tiles. We say that $P < P'$ if

$$\begin{aligned} R_1(P) \supsetneq R_1(P'), \\ I_P \subsetneq I_{P'}. \end{aligned}$$

Thus, the difference between the $<_{\mathbf{n}}$ order relation and the newly introduced $<$ is that the former involves the shifting parameter \mathbf{n} , as the subscript suggests, while the latter doesn't. Now, let

$$\mathbb{P}_{\text{tops}} = \{\text{Top}(\mathcal{C}) \text{ s.t. } \mathcal{C} \in \mathfrak{C}\}$$

and let $\mathbb{P}_{\text{tops}}^{\max}$ denote the subcollection of super tiles that are maximal with respect to the ordering $<$ just introduced. Notice that the tiles in \mathbb{P}_{tops} are all distinct by assumption, which has the following consequence: if one fixes $P_0 \in \mathbb{P}_{\text{tops}}^{\max}$ then trivially

$$\sum_{\substack{P \in \mathbb{P}_{\text{tops}} : P < P_0, \\ |I_P| \gtrsim \mathbf{n}^{-1} |I_{P_0}|}} |I_P| \lesssim \log^+(\mathbf{n}) |I_{P_0}|.$$

As for the tiles $P \in \mathbb{P}_{\text{tops}}$ such that $P < P_0$ but $|I_P| \ll \mathbf{n}^{-1} |I_{P_0}|$, observe that the latter implies that $I_P^n \subset 3I_{P_0}$. Since the \mathbf{n} -columns were assumed to be mutually disjoint, the shifted tiles $R_1(P) \times I_P^n$ are all disjoint, and therefore

$$\sum_{\substack{P \in \mathbb{P}_{\text{tops}} : P < P_0, \\ |I_P| \ll \mathbf{n}^{-1} |I_{P_0}|}} |I_P^n| \lesssim |I_{P_0}|.$$

We have thus shown that

$$\sum_{P \in \mathbb{P}_{\text{tops}}} |I_P| \lesssim \log^+(\mathbf{n}) \sum_{P_0 \in \mathbb{P}_{\text{tops}}^{\max}} |I_{P_0}|.$$

Since we have by assumption

$$\left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} \geq 2^n$$

for all $P \in \mathbb{P}_{\text{tops}}$, we are therefore able to bound

$$\text{Energy}_f^n(\mathbb{P})^{r_0} \lesssim \log^+(\mathbf{n}) \int \sum_{P_0 \in \mathbb{P}_{\text{tops}}^{\max}} |\pi_{R_1(P_0)} f|^{r_0} \mathbf{1}_{I_{P_0}};$$

but by maximality the super tiles in $\mathbb{P}_{\text{tops}}^{\max}$ are all disjoint (as tiles), and therefore we can bound pointwise (as in Lemma 3.7)

$$\sum_{P_0 \in \mathbb{P}_{\text{tops}}^{\max}} |\pi_{R_1(P_0)} f(x)|^{r_0} \mathbf{1}_{I_{P_0}}(x) \leq (\text{Var}^{r_0} \mathcal{C} f(x))^{r_0}.$$

Thus we have proven

$$\text{Energy}_f^n(\mathbb{P})^{r_0} \lesssim \log^+(\mathbf{n}) \int (\text{Var}^{r_0} \mathcal{C} f(x))^{r_0} dx,$$

and the lemma follows (even with the smaller constant $O(\log^+(\mathbf{n})^{1/r_0})$) from the fact that $r_0 > 2$ and hence $\text{Var}^{r_0} \mathcal{C}$ is $L^{r_0} \rightarrow L^{r_0}$ bounded (see [18]). \square

Remark 4.8. The argument in the above lemma motivates our choice of introducing the parameter $r_0 > 2$. Indeed, the Variational Carleson operator $\text{Var}^q \mathcal{C}$ is unbounded when $q \leq 2$, as shown in [18].

5. DECOMPOSITION LEMMAS

Now that all the relevant quantities are in place, we will establish two decomposition lemmas that will allow us to partition every collection of super tiles \mathbb{P} into structured subcollections which have controlled size and energy. The results in this section are classical and are based on simple stopping time arguments.

Lemma 5.1 (Decomposition lemma for f). *Let \mathbb{P} be a collection of super tiles, $f \in L_{\text{loc}}^1(\mathbb{R})$, and let $n \in \mathbb{Z}$ be such that*

$$\text{Size}_f^n(\mathbb{P}) \leq 2^{-n} \text{Energy}_f^n(\mathbb{P}).$$

Then we can decompose \mathbb{P} into $\mathbb{P}_{\text{high}} \sqcup \mathbb{P}_{\text{low}}$ so that

$$\text{Size}_f^n(\mathbb{P}_{\text{low}}) \leq \frac{1}{2} 2^{-n} \text{Energy}_f^n(\mathbb{P}),$$

and \mathbb{P}_{high} can be partitioned into a collection \mathfrak{C} of mutually disjoint \mathbf{n} -columns such that

$$\sum_{\mathcal{C} \in \mathfrak{C}} |I_{\text{Top}(\mathcal{C})}| \lesssim 2^{r_0 n}.$$

Proof. We can assume the collection \mathbb{P} is finite, for simplicity. Let $\mathbb{P}_{\text{stock}}$ be initialized to

$$\mathbb{P}_{\text{stock}} := \left\{ P \in \mathbb{P} \text{ s.t. } \left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} > 2^{-n-1} \text{Energy}_f^n(\mathbb{P}) \right\},$$

and let \mathfrak{C} be initialized to $\mathfrak{C} := \emptyset$. We set right away

$$\mathbb{P}_{\text{low}} := \mathbb{P} \setminus \mathbb{P}_{\text{stock}},$$

which will not be changed throughout the algorithm. The size property is then immediate from the definition of (the initial state of) $\mathbb{P}_{\text{stock}}$. As for the organization of $\mathbb{P}_{\text{high}} := \mathbb{P} \setminus \mathbb{P}_{\text{low}}$ into \mathbf{n} -columns, we proceed as follows. Let P_{max} be the maximal super tile in $\mathbb{P}_{\text{stock}}$ (with respect to $<_{\mathbf{n}}$) such that $\inf I_{P_{\text{max}}}^{\mathbf{n}}$ is minimum and $\sup R_1(P_{\text{max}})$ is maximum (any other total order will do, of course). Then we let \mathcal{C} be the maximal \mathbf{n} -column (with respect to $<_{\mathbf{n}}$) in $\mathbb{P}_{\text{stock}}$ with top P_{max} , update \mathfrak{C} to be $\mathfrak{C} \cup \{\mathcal{C}\}$ and update $\mathbb{P}_{\text{stock}}$ to be $\mathbb{P}_{\text{stock}} \setminus \bigcup_{P \in \mathcal{C}} \{P\}$. Repeat the process until $\mathbb{P}_{\text{stock}}$ is empty and the algorithm stops. Then we see that by maximality the \mathbf{n} -columns in \mathfrak{C} are mutually disjoint, and as for the bound on $\sum_{\mathcal{C} \in \mathfrak{C}} |I_{\mathcal{C}}|$ notice that we have for each \mathcal{C} and for each $P \in \mathcal{C}$

$$\left(\int_{I_P} |\pi_{R_1(P)} f|^{r_0} \right)^{1/r_0} > 2^{-n-1} \text{Energy}_f^n(\mathbb{P})$$

and therefore just by definition of energy (and its monotonicity)

$$2^{-n-1} \text{Energy}_f^n(\mathbb{P}) \left(\sum_{C \in \mathfrak{C}} |I_C| \right)^{1/r_0} \lesssim \text{Energy}_f^n(\mathbb{P}_{\text{high}}) \leq \text{Energy}_f^n(\mathbb{P}),$$

which proves the claim. \square

Next we have

Lemma 5.2 (Decomposition lemma for \mathbf{h}). *Let \mathbb{P} be a collection of super tiles, $\gamma = 2^{r_0/r'}$, $\mathbf{h} \in L_{\text{loc}}^1(\ell^{r'})$, and let $n \in \mathbb{Z}$ be such that*

$$\text{Size}_{\mathbf{h}}^n(\mathbb{P}) \leq \gamma^{-n} \text{Energy}_{\mathbf{h}}^n(\mathbb{P}).$$

Then we can decompose \mathbb{P} into $\mathbb{P}_{\text{high}} \sqcup \mathbb{P}_{\text{low}}$ so that

$$\text{Size}_{\mathbf{h}}^n(\mathbb{P}_{\text{low}}) \leq \gamma^{-n-1} \text{Energy}_{\mathbf{h}}^n(\mathbb{P}),$$

and \mathbb{P}_{high} can be partitioned into a collection \mathfrak{C} of mutually disjoint \mathbf{n} -columns such that

$$\sum_{C \in \mathfrak{C}} |I_{\text{Top}(C)}| \lesssim \gamma^{r'n} = 2^{r_0 n}.$$

We have chosen to introduce the constant γ in order to make the global decomposition lemma below more readily apparent. The proof of the decomposition lemma for \mathbf{h} is quite similar to the one for f and is thus omitted.

Finally, we can combine the two decomposition lemmas above iteratively in order to produce a global decomposition of the collection \mathbb{P} as follows.

Lemma 5.3. *Let \mathbb{P} be a collection of super tiles. Then there exists a decomposition $\mathbb{P} = \bigsqcup_n \mathbb{P}_n$ with the properties:*

- i) $\text{Size}_f^n(\mathbb{P}_n) \lesssim \min(2^{-n} \text{Energy}_f^n(\mathbb{P}), \text{Size}_f^n(\mathbb{P}))$,
- ii) $\text{Size}_{\mathbf{h}}^n(\mathbb{P}_n) \lesssim \min(2^{-r_0 n/r'} \text{Energy}_{\mathbf{h}}^n(\mathbb{P}), \text{Size}_{\mathbf{h}}^n(\mathbb{P}))$,
- iii) \mathbb{P}_n is organized into a collection \mathfrak{C}_n of mutually disjoint \mathbf{n} -columns,
- iv) $\sum_{C \in \mathfrak{C}_n} |I_C| \lesssim 2^{r_0 n}$.

Furthermore, the collection \mathbb{P}_n is empty if n is such that

$$2^{-n} \gtrsim \frac{\text{Size}_f^n(\mathbb{P})}{\text{Energy}_f^n(\mathbb{P})} \quad \text{and} \quad 2^{-r_0 n/r'} \gtrsim \frac{\text{Size}_{\mathbf{h}}^n(\mathbb{P})}{\text{Energy}_{\mathbf{h}}^n(\mathbb{P})}.$$

Proof. Initialize $\mathbb{P}_{\text{stock}} := \mathbb{P}$ and apply iteratively the decomposition Lemmas 5.1 and 5.2, in the order given by whichever of the quantities

$$\frac{\text{Size}_f^n(\mathbb{P}_{\text{stock}})}{\text{Energy}_f^n(\mathbb{P})}, \quad \left(\frac{\text{Size}_{\mathbf{h}}^n(\mathbb{P}_{\text{stock}})}{\text{Energy}_{\mathbf{h}}^n(\mathbb{P})} \right)^{r'/2}$$

is largest, putting the resulting \mathbb{P}_{high} collection in the corresponding \mathbb{P}_n and updating $\mathbb{P}_{\text{stock}}$ at the end of each step to be the collection $(\mathbb{P}_{\text{stock}})_{\text{low}}$ resulting from the last application of a decomposition lemma. We omit the details. \square

6. GENERIC ESTIMATE

In this section we will combine the estimate for a given \mathbf{n} -column in Proposition 3.6 with the global decomposition obtained in Section 5 in order to obtain a global estimate for the trilinear form $\Lambda_{\mathbb{P}}^{\mathbf{n}}$ in terms of the sizes and energies of the collection \mathbb{P} . More precisely, we obtain the following (recall that $r > r_0 > 2$).

Proposition 6.1 (Generic estimate). *Let $\sigma = (r - r_0)/r > 0$. Let \mathbb{P} be a collection of super tiles, and denote for shortness*

$$\begin{aligned} \text{Size}_f^n(\mathbb{P}) &=: S_f, & \text{Energy}_f^n(\mathbb{P}) &=: E_f, \\ \text{Size}_{\mathbf{h}}^n(\mathbb{P}) &=: S_{\mathbf{h}}, & \text{Energy}_{\mathbf{h}}^n(\mathbb{P}) &=: E_{\mathbf{h}}. \end{aligned}$$

Then for every $0 \leq \theta_1, \theta_2 \leq 1$ such that $\theta_1 + \theta_2 = 1$ we have

$$|\Lambda_{\mathbb{P}}^{\mathbf{n}}(f, g, \mathbf{h})| \lesssim \sup_{P \in \mathbb{P}} \left[\frac{1}{|I_P^{\mathbf{n}}|} \int_{I_P^{\mathbf{n}}} (\text{Var}^r \mathcal{C}g)^r \right]^{1/r} \times S_f^{\sigma\theta_1} E_f^{1-\sigma\theta_1} S_{\mathbf{h}}^{r'\sigma\theta_2/r_0} E_{\mathbf{h}}^{1-r'\sigma\theta_2/r_0}. \quad (5)$$

Proof. Apply the global decomposition lemma (Lemma 5.3) to the collection \mathbb{P} , which yields the subcollections \mathbb{P}_n and their partitions \mathfrak{C}_n into \mathbf{n} -columns. For each such subcollection we then have by Proposition 3.6 and Lemma 5.3 that we can bound

$$\begin{aligned} |\Lambda_{\mathbb{P}_n}^{\mathbf{n}}(f, g, \mathbf{h})| &\leq \sum_{\mathcal{C} \in \mathfrak{C}_n} |\Lambda_{\mathcal{C}}^{\mathbf{n}}(f, g, \mathbf{h})| \\ &\lesssim \sum_{\mathcal{C} \in \mathfrak{C}_n} \left[\frac{1}{|I_{\text{Top}(\mathcal{C})}|} \sum_{P \in \mathcal{C}} \int_{I_P^{\mathbf{n}}} |\pi_{R_2(P)} g|^r \right]^{1/r} \text{Size}_f^{\mathbf{n}}(\mathcal{C}) \text{Size}_{\mathbf{h}}^{\mathbf{n}}(\mathcal{C}) |I_{\text{Top}(\mathcal{C})}| \\ &\lesssim \sum_{\mathcal{C} \in \mathfrak{C}_n} \left[\frac{1}{|I_{\text{Top}(\mathcal{C})}|} \sum_{P \in \mathcal{C}} \int_{I_P^{\mathbf{n}}} |\pi_{R_2(P)} g|^r \right]^{1/r} \\ &\quad \times \min(2^{-n} E_f, S_f) \min(2^{-r_0 n/r'} E_{\mathbf{h}}, S_{\mathbf{h}}) |I_{\text{Top}(\mathcal{C})}|. \end{aligned}$$

By Lemma 3.7 the term $\frac{1}{|I_{\text{Top}(\mathcal{C})}|} \sum_{P \in \mathcal{C}} \int_{I_P^{\mathbf{n}}} |\pi_{R_2(P)} g|^r$ can be replaced by

$$\sup_{P \in \mathbb{P}} \frac{1}{|I_P^{\mathbf{n}}|} \int_{I_P^{\mathbf{n}}} (\text{Var}^r \mathcal{C}g)^r,$$

and therefore we obtain by property (iv) of Lemma 5.3 that

$$\begin{aligned} |\Lambda_{\mathbb{P}_n}^{\mathbf{n}}(f, g, \mathbf{h})| &\lesssim \sup_{P \in \mathbb{P}} \left[\frac{1}{|I_P^{\mathbf{n}}|} \int_{I_P^{\mathbf{n}}} (\text{Var}^r \mathcal{C}g)^r \right]^{1/r} \\ &\quad \times \min(2^{-n} E_f, S_f) \min(2^{-r_0 n/r'} E_{\mathbf{h}}, S_{\mathbf{h}}) 2^{r_0 n}. \end{aligned}$$

It thus suffices to show that upon summing in n we obtain a quantity that is bounded exactly by the right hand side of (5). This requires a tedious but easy case by case analysis.

Assume then that

$$\frac{S_f}{E_f} < \left(\frac{S_{\mathbf{h}}}{E_{\mathbf{h}}} \right)^{r'/r_0},$$

the other cases being similar. We have to distinguish two situations:

1) **case** $2^{-n} \leq \frac{S_f}{E_f} < \left(\frac{S_{\mathbf{h}}}{E_{\mathbf{h}}} \right)^{r'/2}$: in this case we have to bound the sum

$$\sum_{n : 2^{-n} \leq S_f(E_f)^{-1}} 2^{-n} E_f 2^{-r_0 n/r'} E_{\mathbf{h}} 2^{r_0 n} = E_f E_{\mathbf{h}} \sum_{n : 2^{-n} \leq S_f(E_f)^{-1}} 2^{-n(1-r_0/r)};$$

since $1 - r_0/r = \sigma$ this sum is readily evaluated to be bounded by

$$\begin{aligned} E_f E_{\mathbf{h}} \left(\frac{S_f}{E_f} \right)^{\sigma} &\leq E_f E_{\mathbf{h}} \left(\frac{S_f}{E_f} \right)^{\sigma\theta_1} \left(\frac{S_{\mathbf{h}}}{E_{\mathbf{h}}} \right)^{r'\sigma\theta_2/r_0} \\ &= S_f^{\sigma\theta_1} E_f^{1-\sigma\theta_1} S_{\mathbf{h}}^{r'\sigma\theta_2/r_0} E_{\mathbf{h}}^{1-r'\sigma\theta_2/r_0} \end{aligned}$$

as desired.

2) **case** $\frac{S_f}{E_f} < 2^{-n} \leq \left(\frac{S_{\mathbf{h}}}{E_{\mathbf{h}}}\right)^{r'/r_0}$: in this case we have

$$\begin{aligned} & \sum_{n : S_f(E_f)^{-1} < 2^{-n} \leq S_{\mathbf{h}}^{r'/r_0} E_{\mathbf{h}}^{-r'/r_0}} S_f 2^{-r_0 n/r'} E_{\mathbf{h}} 2^{r_0 n} \\ &= S_f E_{\mathbf{h}} \sum_{n : S_f(E_f)^{-1} < 2^{-n} \leq S_{\mathbf{h}}^{r'/r_0} E_{\mathbf{h}}^{-r'/r_0}} 2^{r_0 n/r} \\ &\lesssim S_f E_{\mathbf{h}} \left(\frac{S_f}{E_f}\right)^{-r_0/r} = E_f E_{\mathbf{h}} \left(\frac{S_f}{E_f}\right)^{\sigma}, \end{aligned}$$

which we have already seen is acceptable in the previous case.

This concludes the proof. \square

7. PRELIMINARY RESULT WHEN $p = r$ OR $q = r$

In this section we will prove a preliminary result that, when combined with the interpolation technique in Section §8 will yield Theorem 1.6. In particular, we will prove restricted weak type estimates for our trilinear forms with either $p = r$ (for rectangles of low eccentricity) or $q = r$ (for rectangles of high eccentricity).

We begin by stating rigorously the results that will be proven in this section. There are two such results, one for each of the collections $\mathcal{R}_{\text{high}}$, \mathcal{R}_{low} defined in Section §2.

Let A be a measurable set. We say that a measurable set A' is a *major subset* of A if $A' \subset A$ and $|A'| > \frac{1}{2}|A|$.

Proposition 7.1 (Rectangles of high eccentricity). *Let $r > r_0 > 2$ and let $\mathcal{R}_{\text{high}}$ be a collection of disjoint rectangles such that for every $R \in \mathcal{R}_{\text{high}}$ it holds that $\text{ec}(R) > 1$. Let p be such that*

$$\frac{1}{100} \frac{1}{r_0} + \frac{99}{100} \frac{1}{r} \leq \frac{1}{p} \leq \frac{1}{r_0}$$

and s be such that

$$\frac{1}{p} + \frac{1}{r} = \frac{1}{s}.$$

Then, for any F, G, H finite measurable subsets of \mathbb{R} and for any measurable functions f, g such that

$$|f| \leq \mathbf{1}_F, \quad |g| \leq \mathbf{1}_G,$$

there exists a major subset $H' \subset H$ such that for any measurable vector-valued function $\mathbf{h} = (h_R)_{R \in \mathcal{R}_{\text{high}}}$ that satisfies

$$\left(\sum_{R \in \mathcal{R}_{\text{high}}} |h_R|^{r'} \right)^{1/r'} \leq \mathbf{1}_{H'}$$

it holds that

$$|\Lambda_{\mathcal{R}_{\text{high}}}^r(f, g, \mathbf{h})| \lesssim_{p,r} |F|^{1/p} |G|^{1/r} |H|^{1/s'}.$$

The choice of major subset may depend on r, r_0, f, g, H but not on p, s .

Observe that the range of exponents is restricted to a subset of the line $q = r$.

Remark 7.2. The rather clumsy appearance of the above statement is due to some technicalities. Some comments are in order.

Firstly, there is the fact that in our argument the major subset H' will end up depending on f, g rather than on the sets F, G only - the latter being the standard in typical time-frequency analysis arguments. This is because we will control the contributions of f and g by non-local non-positive averages (see Lemma 3.9 and Lemma 3.7) and thus we will not be allowed to replace f and g by $\mathbf{1}_F, \mathbf{1}_G$ in the

definition of the exceptional set that we are going to remove from H (see proof below). However, this will not be a problem for the interpolation argument in the quasi-Banach range (that is, when $s < 1$), as will be clear from Lemma 8.6 and its proof.

Secondly, the range of boundedness in the statement looks unnatural and arbitrary. This is due to the fact that, in order to perform interpolation between estimates for different values of r , we need the major subset to be independent of p, s . If we gave up this requirement, the methods employed in the proof would yield the more natural range $2 < p < r$, but then we would not be able to ensure that the choice of major subset is indeed independent of p . In view of a more transparent proof, we have preferred the above formulation.

Similarly, we have

Proposition 7.3 (Rectangles of low eccentricity). *Let $r > r_0 > 2$ and let \mathcal{R}_{low} be a collection of disjoint rectangles such that for every $R \in \mathcal{R}_{\text{low}}$ it holds $\text{ec}(R) < 1$. Let q be such that*

$$\frac{1}{100} \frac{1}{r_0} + \frac{99}{100} \frac{1}{r} \leq \frac{1}{q} \leq \frac{1}{r_0}$$

and s be such that

$$\frac{1}{r} + \frac{1}{q} = \frac{1}{s}.$$

Then, for any F, G, H finite measurable subsets of \mathbb{R} and for any functions f, g such that

$$|f| \leq \mathbb{1}_F, \quad |g| \leq \mathbb{1}_G,$$

there exists a major subset $H' \subset H$ such that for any measurable vector-valued function $\mathbf{h} = (h_R)_{R \in \mathcal{R}_{\text{low}}}$ that satisfies

$$\left(\sum_{R \in \mathcal{R}_{\text{low}}} |h_R|^{r'} \right)^{1/r'} \leq \mathbb{1}_{H'}$$

it holds that

$$|\Lambda_{\mathcal{R}_{\text{low}}}^r(f, g, \mathbf{h})| \lesssim_{q,r} |F|^{1/r} |G|^{1/q} |H|^{1/s'}.$$

The choice of major subset H' may depend on r, r_0, f, g, H but not on q, s .

The range of exponents this time is restricted to a subset of the line $p = r$ instead. The same remarks as in 7.2 apply.

Although the ranges in the two above propositions don't intersect, upon interpolation with the $r = \infty$ case they will each yield the same range of estimates, allowing us to sum their two contributions. See next section for details.

We now proceed with the proof. We will prove the first of the two propositions above and then comment in Remark 7.4 on the modifications one has to make to the argument in order to adapt it to the second one.

Proof of Prop. 7.1. The proof follows a standard argument originating from [17] (although implicitly present in previous work), together with ideas from [1, 4], in particular the idea of using non-local operators for the stopping-time argument below.

It will suffice to prove, for data as given in the statement, that for any collection of super tiles \mathbb{P} it is

$$\Lambda_{\mathbb{P}}^{\mathbf{n}}(f, g, \mathbf{h}) \lesssim_{p,q,r} [\log^+(\mathbf{n})]^{O(1)} |F|^{1/p} |G|^{1/r} |H|^{1/s'} \quad (6)$$

for $\frac{1}{100r_0} + \frac{99}{100r} \leq \frac{1}{p} \leq \frac{1}{r_0}$ and $1/p + 1/r + 1/s' = 1$. Indeed, since the constant is only poly-logarithmic in the shifting parameter \mathbf{n} and since by Section §2 the

trilinear form $\Lambda_{\mathcal{R}_{\text{high}}}^r$ is dominated by

$$\sum_{\mathbf{n} \in \mathbb{Z}} (1 + |\mathbf{n}|)^{-N} \Lambda_{\mathbb{P}}^{\mathbf{n}}(f, g, \mathbf{h})$$

for an arbitrarily large $N > 0$, the above implies the same bounds for $\Lambda_{\mathcal{R}_{\text{high}}}^r$ itself. Given sets F, G, H and functions f, g such that

$$|f| \leq \mathbb{1}_F, \quad |g| \leq \mathbb{1}_G,$$

we set out to define an exceptional set E (independent of p, s but depending on f, g) which will be removed from H . It will be useful to introduce the following maximal operators: for any $\mathbf{m} \in \mathbb{Z}$, define $\mathcal{M}^{\mathbf{m}}$ to be the shifted maximal function

$$\mathcal{M}^{\mathbf{m}} f(x) := \sup_{\substack{I \text{ dyadic,} \\ x \in I}} \frac{1}{|I|} \int_{I^{\mathbf{m}}} |f(y)| dy.$$

It is well known that the shifted maximal function $\mathcal{M}^{\mathbf{m}}$ is $L^p \rightarrow L^p$ bounded for $1 < p \leq \infty$ and also $L^1 \rightarrow L^{1,\infty}$ bounded, with constant at most $O(\log^+(\mathbf{m}))$ (see for example §4.3.4 of [16]). Now we can define the exceptional set E to be

$$\begin{aligned} E := & \bigcup_{\mathbf{m} \in \mathbb{Z}} \left\{ x \in \mathbb{R} \text{ s.t. } \mathcal{M}^{-\mathbf{n}+\mathbf{m}}((\mathcal{C}f)^{100r_0})(x) \gtrsim \log^+(\mathbf{m}-\mathbf{n}) \langle \mathbf{m} \rangle^2 \frac{|F|}{|H|} \right\} \\ & \cup \bigcup_{\mathbf{m}' \in \mathbb{Z}} \left\{ x \in \mathbb{R} \text{ s.t. } \mathcal{M}^{\mathbf{m}'}((\text{Var}^r \mathcal{C}g)^r)(x) \gtrsim \log^+(\mathbf{m}') \langle \mathbf{m}' \rangle^2 \frac{|G|}{|H|} \right\} \end{aligned}$$

with implicit constants to be chosen below, where $\langle \mathbf{m} \rangle := (1 + |\mathbf{m}|^2)^{1/2}$. Define then the set $H' := H \setminus E$; we claim that H' is a major subset of H . Indeed, for a given $\mathbf{m} \in \mathbb{Z}$ we see by the $L^1 \rightarrow L^{1,\infty}$ boundedness of $\mathcal{M}^{-\mathbf{n}+\mathbf{m}}$ that

$$\begin{aligned} & \left| \left\{ x \in \mathbb{R} \text{ s.t. } \mathcal{M}^{-\mathbf{n}+\mathbf{m}}((\mathcal{C}f)^{100r_0})(x) \gtrsim \log^+(\mathbf{m}-\mathbf{n}) \langle \mathbf{m} \rangle^2 \frac{|F|}{|H|} \right\} \right| \\ & \lesssim \log^+(\mathbf{m}-\mathbf{n}) \frac{\int |\mathcal{C}f|^{100r_0} dx}{\log^+(\mathbf{m}-\mathbf{n}) \langle \mathbf{m} \rangle^2 \frac{|F|}{|H|}} \\ & \lesssim \langle \mathbf{m} \rangle^{-2} |H| \frac{\int |f|^{100r_0}}{|F|} \leq \langle \mathbf{m} \rangle^{-2} |H|, \end{aligned}$$

where in the last line we have used the fact that the Carleson operator \mathcal{C} is $L^{100r_0} \rightarrow L^{100r_0}$ bounded. Since this is summable in \mathbf{m} , we see that for a suitable choice of implicit constants the contribution to $|E|$ of these sets is at most $\ll |H|$. Similarly, for a given $\mathbf{m}' \in \mathbb{Z}$ we see by the $L^1 \rightarrow L^{1,\infty}$ boundedness of $\mathcal{M}^{\mathbf{m}'}$ and by the $L^r \rightarrow L^r$ boundedness of $\text{Var}^r \mathcal{C}g$ that

$$\left| \left\{ x \in \mathbb{R} \text{ s.t. } \mathcal{M}^{\mathbf{m}'}((\text{Var}^r \mathcal{C}g)^r)(x) \gtrsim \log^+(\mathbf{m}') \langle \mathbf{m}' \rangle^2 \frac{|G|}{|H|} \right\} \right| \ll \langle \mathbf{m}' \rangle^{-2} |H|,$$

which is again summable in \mathbf{m}' and thus we conclude that $|E| \ll |H|$, which proves the claim. Notice H' depends on f, g, r, r_0 and H but not on p .

Now, we partition the collection \mathbb{P} into

$$\mathbb{P}_{\text{small}} := \{P \in \mathbb{P} \text{ s.t. } I_P^{\mathbf{n}} \not\subset E\},$$

$$\mathbb{P}_{\text{large}} := \mathbb{P} \setminus \mathbb{P}_{\text{small}},$$

and will estimate the trilinear forms $\Lambda_{\mathbb{P}_{\text{small}}}^{\mathbf{n}}$ and $\Lambda_{\mathbb{P}_{\text{large}}}^{\mathbf{n}}$ separately. Notice the presence of the shifting parameter \mathbf{n} in the definitions.

We start with $\mathbb{P}_{\text{small}}$. Given $P \in \mathbb{P}_{\text{small}}$ we observe that since $I_P^{\mathbf{n}} \not\subset E$ there exists

an $x \in I_P^n$ such that $\mathcal{M}^{-n}((\mathcal{C}f)^{100r_0})(x) \lesssim \log^+(n)|F|/|H|$, and thus a fortiori it must be that

$$\frac{1}{|I_P|} \int_{(I_P^n)^{-n}} |\mathcal{C}f|^{100r_0} dx \lesssim \log^+(n) \frac{|F|}{|H|};$$

but $(I_P^n)^{-n} = I_P$, and therefore we must have by Lemma 3.9 that

$$\text{Size}_f^n(\mathbb{P}_{\text{small}}) \lesssim \log^+(n) \left(\frac{|F|}{|H|} \right)^{1/100r_0} \quad (7)$$

(we are not keeping track of the optimal powers of $\log^+(n)$ because they will be inconsequential to us). Similarly, we see that for a given $P \in \mathbb{P}_{\text{small}}$ there must be an $x \in I_P^n$ such that $\mathcal{M}^0((\text{Var}^r \mathcal{C}g)^r)(x) \lesssim |G|/|H|$, and therefore for all such P we have

$$\frac{1}{|I_P|} \int_{I_P^n} (\text{Var}^r \mathcal{C}g)^r dx \lesssim \frac{|G|}{|H|}. \quad (8)$$

Moreover, by Lemma 3.10 and the hypothesis that $\|\mathbf{h}\|_{\ell^{r'}} \leq \mathbf{1}_{H'}$ we have the trivial bound

$$\text{Size}_{\mathbf{h}}^n(\mathbb{P}) \lesssim 1 \quad (9)$$

for any arbitrary collection of super tiles \mathbb{P} .

Recall that $\sigma = (r - r_0)/r$. Combining (7), (8) and (9) with the generic estimate of Proposition 6.1 and the energy estimates of Lemmas 4.6, 4.5, we obtain after a little algebra that for any $0 \leq \theta_1, \theta_2 \leq 1$, $\theta_1 + \theta_2 = 1$,

$$\begin{aligned} |\Lambda_{\mathbb{P}_{\text{small}}}^n(f, g, \mathbf{h})| &\lesssim \log^+(n)^{O(1)} \left(\frac{|G|}{|H|} \right)^{1/r} \left(\frac{|F|}{|H|} \right)^{\sigma\theta_1/100r_0} \\ &\quad \times (|F|^{1/r_0})^{1-\sigma\theta_1} \cdot \mathbf{1} \cdot |H|^{1/r' - \sigma\theta_2/r_0} \\ &= \log^+(n)^{O(1)} |F|^{1/r_0 - (99/100)\sigma\theta_1/r_0} |G|^{1/r} \\ &\quad \times |H|^{1/r' - 1/r - \sigma\theta_2/r_0 - \sigma\theta_1/100r_0}. \end{aligned} \quad (10)$$

Since

$$\frac{1}{r_0} - \frac{99}{100} \frac{\sigma}{r_0} = \frac{1}{100r_0} + \frac{99}{100r_0},$$

this yields precisely the desired estimates (6) as θ_1 ranges over $[0, 1]$.

Now we are left with showing that (6) holds for $\Lambda_{\mathbb{P}_{\text{large}}}^n$ as well. In order to do so, we decompose $\mathbb{P}_{\text{large}}$ into $\bigsqcup_{d \in \mathbb{N}} \mathbb{P}_d$ where

$$\mathbb{P}_d := \left\{ P \in \mathbb{P}_{\text{large}} \text{ s.t. } 2^d \leq 1 + \frac{\text{dist}(I_P^n, E^c)}{|I_P|} < 2^{d+1} \right\};$$

it then suffices to prove that the contribution of $\Lambda_{\mathbb{P}_d}^n$ is summable in d and that the sum is bounded by the right hand side of (10).

Let then d be fixed and observe that if $P \in \mathbb{P}_d$ then $2^{d+1}I_P^n \not\subset E$ and $2^d I_P^n \subset E$. This means that for some integer \mathfrak{l} such that $|\mathfrak{l}| \sim 2^d$ the shifted interval $(I_P^n)^{\mathfrak{l}} = I_P^{n+\mathfrak{l}}$ is not contained in E . This implies, by definition of E , that for some $x \in I_P^{n+\mathfrak{l}}$ we have $\mathcal{M}^{-n-\mathfrak{l}}(\mathcal{C}f^{100r_0})(x) \lesssim \log^+(n+\mathfrak{l}) \langle \mathfrak{l} \rangle^2 |F|/|H|$, which in turn implies as before

$$\text{Size}_f^n(\mathbb{P}_d) \lesssim \left(\sup_{|\mathfrak{l}| \sim 2^d} \log^+(n+\mathfrak{l}) \right) 2^{2d/100r_0} \left(\frac{|F|}{|H|} \right)^{100r_0}. \quad (11)$$

Similarly, given $P \in \mathbb{P}_d$ we see that there must be an integer \mathfrak{l} such that $|\mathfrak{l}| \sim 2^d$ and the shifted interval $(I_P^n)^{\mathfrak{l}} = I_P^{n+\mathfrak{l}}$ contains a point x such that $\mathcal{M}^{-\mathfrak{l}}((\text{Var}^r \mathcal{C}g)^r)(x) \lesssim \log^+(\mathfrak{l}) \langle \mathfrak{l} \rangle^2 |G|/|H|$; this in turn implies the estimate

$$\sup_{P \in \mathbb{P}_d} \frac{1}{|I_P|} \int_{I_P^n} (\text{Var}^r \mathcal{C}g)^r dx \lesssim d 2^{2d} \frac{|G|}{|H|}. \quad (12)$$

Now, estimates (11), (12) are worse than the corresponding ones (7), (8); however, we now have an improved estimate for \mathbf{h} . Indeed, if $P \in \mathbb{P}_d$, we must have by definition of H' that $\text{dist}(I_P^n, H') \gtrsim 2^d |I_P|$; then, by the rapid decay of the functions $\Phi_{I_P^n}$, for any large $M > 0$ we can bound (again by Lemma 3.10)

$$\text{Size}_{\mathbf{h}}^n(\mathbb{P}_d) \lesssim_M 2^{-dM}. \quad (13)$$

Combining estimates (11), (12) and (13) with Proposition 6.1 and the energy Lemmas 4.6, 4.5 as done before, we obtain for any $0 \leq \theta_1, \theta_2 \leq 1$, $\theta_1 + \theta_2 = 1$, that the estimate

$$\begin{aligned} |\Lambda_{\mathbb{P}_d}^n(f, g, \mathbf{h})| &\lesssim \left(\sup_{|\mathbf{l}| \sim 2^d} \log^+(\mathbf{n} + \mathbf{l}) \right)^{O(1)} d 2^{-dM'} |F|^{1/r_0 - (99/100)\sigma\theta_1/r_0} \\ &\quad \times |G|^{1/r} |H|^{1/r' - 1/r - \sigma\theta_2/r_0 - \sigma\theta_1/100r_0} \end{aligned}$$

holds, where $M' > M/2$, provided M was chosen sufficiently large. This is summable in $d > 0$, and since

$$\sum_{d>0} \sum_{\mathbf{l} : |\mathbf{l}| \sim 2^d} \log^+(\mathbf{n} + \mathbf{l})^{O(1)} d 2^{-dM'} \lesssim (\log^+(\mathbf{n}))^{O(1)}$$

we see that $|\Lambda_{\mathbb{P}_{\text{large}}}^n(f, g, \mathbf{h})|$ is controlled by the right hand side of (10) as desired, thus concluding the proof. \square

Remark 7.4. We conclude the section with a comment on how Proposition 7.3 is proven. In the previous sections we have concentrated on the collection $\mathcal{R}_{\text{high}}$ of rectangles with large eccentricity (i.e. $|R_2| > |R_1|$). However, it is clear that the analysis is completely symmetric in the opposite case of \mathcal{R}_{low} , in which the rôles of functions f and g are simply swapped. In particular, one has to re-define super tiles reversing f and g and use \mathbf{n} -rows instead of \mathbf{n} -columns (adapting the definitions is a trivial exercise); then one performs a global decomposition of the collection of super tiles into uniformly controlled \mathbf{n} -rows. All lemmas and propositions from previous sections then hold with f replaced by g and viceversa, without substantial changes.

8. PROOF OF THE MAIN THEOREM

In this section we will finally complete the proof of Theorem 1.6. This will be achieved by interpolating the results of Section §7 for generic $r > 2$ with the trivial case of $r = \infty$, as in the authors' related work [7].

First of all, we state (and prove) the $r = \infty$ case of Theorem 1.6.

Theorem 8.1. *Let \mathcal{R} be a collection of disjoint rectangles. Then for all $1 < p, q < \infty$ and*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

it holds that for any $f \in L^p, g \in L^q$

$$\|S_{\mathcal{R}}^{\infty}(f, g)\|_{L^s} \lesssim_{p,q} \|f\|_{L^p} \|g\|_{L^q}.$$

Proof. Notice that

$$S_{\mathcal{R}}^{\infty}(f, g)(x) = \sup_{R \in \mathcal{R}} |\pi_{R_1} f(x) \cdot \pi_{R_2} g(x)|$$

is dominated pointwise by $\mathcal{C}f \cdot \mathcal{C}g$. The result is then a trivial consequence of Hölder's inequality and the Carleson-Hunt theorem. \square

Remark 8.2. We record here for future reference that Theorem 8.1 implies the $L^p \times L^r \times L^{s'}(\ell^1) \rightarrow \mathbb{R}$ boundedness of the trilinear form $\Lambda_{\mathcal{R}}^{\infty}(f, g, \mathbf{h})$ as given in (2).

In order to combine this trivial result with Propositions 7.1, 7.3, we will need a multilinear interpolation argument for vector-valued operators. This argument is originally due to Silva, who introduced it in [20], and is based upon complex interpolation. We introduce the following definition that extends the usual notion of generalized restricted weak type inequalities to the vector-valued setting we need.

Definition 8.3. Let Λ be a trilinear form, let $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ be such that $0 \leq \alpha_1, \alpha_2 \leq 1$, $\alpha_3 \leq 1$, $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and let $t \geq 1$. We say that Λ is of *generalized restricted weak type* $(\alpha; t)$ if for every measurable subsets F, G, H of \mathbb{R} of finite measure and for every functions f, g such that

$$|f| \leq \mathbf{1}_F, \quad |g| \leq \mathbf{1}_G,$$

there exists a major subset H' of H such that for any vector-valued measurable function $\mathbf{h} = (h_k)_k$ that satisfies

$$\left(\sum_k |h_k|^t \right)^{1/t} \leq \mathbf{1}_{H'},$$

the inequality

$$|\Lambda(f, g, \mathbf{h})| \lesssim |F|^{\alpha_1} |G|^{\alpha_2} |H|^{\alpha_3}$$

holds true.

Remark 8.4. The difference between Definition 8.3 and the classical definition of generalized restricted weak type (see for example Chapter 3 in [23]) is two-fold. Firstly, they differ in the presence of the additional parameter t which specifies the space where the vector-valued function \mathbf{h} takes values; secondly, the major subset H' is here allowed to depend on f, g instead of F, G only.

Remark 8.5. Using Definition 8.3 we can rephrase Proposition 7.1 as stating that, under the same hypotheses, the trilinear form $\Lambda_{\mathcal{R}_{\text{high}}}^r$ is of generalized restricted weak type

$$\left(\left(\frac{1}{p}, \frac{1}{r}, \frac{1}{s'} \right); r' \right)$$

for all p, s such that $\frac{1}{100r_0} + \frac{99}{100r} \leq \frac{1}{p} \leq \frac{1}{r_0}$ and $\frac{1}{p} + \frac{1}{r} + \frac{1}{s'} = 1$.

Similarly, Proposition 7.3 states that the trilinear form $\Lambda_{\mathcal{R}_{\text{low}}}^r$ is of generalized restricted weak type

$$\left(\left(\frac{1}{r}, \frac{1}{q}, \frac{1}{s'} \right); r' \right)$$

for all q such that $\frac{1}{100r_0} + \frac{99}{100r} \leq \frac{1}{q} \leq \frac{1}{r_0}$ and $\frac{1}{r} + \frac{1}{q} + \frac{1}{s'} = 1$.

Finally, Theorem 8.1 above implies that $\Lambda_{\mathcal{R}}^\infty$ is of generalized restricted weak type

$$\left(\left(\frac{1}{p}, \frac{1}{q}, \frac{1}{s'} \right); 1 \right)$$

for all $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} + \frac{1}{s'} = 1$.

The interpolation argument can now be stated as follows.

Lemma 8.6 ([20]). *Let Λ be a trilinear form of generalized restricted weak type $(\alpha; t_0)$ and $(\beta; t_1)$, with the property that the major subset is the same for $(\alpha; t_0)$ and $(\beta; t_1)$. Then for all θ such that $0 < \theta < 1$, with α^θ given by*

$$\alpha_j^\theta = (1 - \theta)\alpha_j + \theta\beta_j, \quad j = 1, 2, 3$$

and t_θ given by

$$\frac{1}{t_\theta} = \frac{1 - \theta}{t_0} + \frac{\theta}{t_1},$$

it holds that Λ is of generalized restricted weak type $(\alpha^\theta; t_\theta)$.

One can also bound the constant of the interpolated inequality explicitly in terms of the data, but we will not be interested in doing so here.

Proof. The lemma is a particular case of a more general interpolation lemma of Silva (specifically Lemma 4.3 in [20]). We sketch the proof here for the reader's convenience.

We argue by complex interpolation and assume $t_0, t_1 < \infty$. Let $F, G, H, f, g, H', \theta$ be given and let \mathbf{h} be such that

$$\left(\sum_k |h_k|^{t_\theta} \right)^{1/t_\theta} \leq \mathbf{1}_{H'}.$$

For $z \in \mathbb{C}$ with $\operatorname{Re} z \in [0, 1]$ define \mathbf{h}^z by

$$h_k^z(x) := |h_k(x)|^{t(z)}$$

for every k , where

$$t(z) := (1 - z) \frac{t_\theta}{r_0} + z \frac{t_\theta}{r_1}.$$

When $\operatorname{Re} z = 0$ we have $|h_k^z|^{t_0} = |h_k|^{t_\theta}$, and when $\operatorname{Re} z = 1$ we have $|h_k^z|^{t_1} = |h_k|^{t_\theta}$; hence by assumption we have for $\operatorname{Re} z = 0$

$$|\Lambda(f, g, \mathbf{h}^z)| \lesssim |F|^{\alpha_1} |G|^{\alpha_2} |H|^{\alpha_3},$$

and for $\operatorname{Re} z = 1$ we have

$$|\Lambda(f, g, \mathbf{h}^z)| \lesssim |F|^{\beta_1} |G|^{\beta_2} |H|^{\beta_3}.$$

Since the function $\Psi(z) := \Lambda(f, g, \mathbf{h}^z)$ is easily seen to be holomorphic in the open strip $S = \{z \in \mathbb{C} \text{ s.t. } 0 < \operatorname{Re} z < 1\}$, continuous in its closure and bounded, we can apply to it Hadamard's three-lines-lemma and conclude that since $\mathbf{h}^{\theta + i0} = \mathbf{h}$ we have

$$|\Lambda(f, g, \mathbf{h})| \lesssim |F|^{\alpha_1^\theta} |G|^{\alpha_2^\theta} |H|^{\alpha_3^\theta},$$

as desired. \square

We are now ready to prove Theorem 1.6. It will be a straightforward consequence of the interpolation Lemma 8.6, but some extra care is needed to make sure all hypotheses are verified. The resulting proof is elementary but technical and painful to read; in order to remedy this fact, we have included in Figure 2 a picture that illustrates the proof geometrically.

Proof of Th. 1.6. First of all, as noticed in Remark 8.5, the trilinear form $\Lambda_{\mathcal{R}}^\infty$ is of generalized restricted weak type $\left(\left(\frac{1}{p_0}, \frac{1}{q_0}, \frac{1}{s'_0} \right); 1 \right)$ for $1 < p_0, q_0 < \infty$, $1/p_0 + 1/q_0 + 1/s'_0 = 1$ and for any collection \mathcal{R} . Secondly, set $r_0 = r_0(r) := (r + 2)/2$ in the statement of Proposition 7.1 in order to remove the dependence on r_0 ; write t_1 in place of r' , so that $r_0(r) = (3t_1 - 2)/(2t_1 - 2)$. Then we have that the trilinear form $\Lambda_{\mathcal{R}_{\text{high}}}^r$ is of generalized restricted weak type $\left(\left(\frac{1}{p_1}, \frac{1}{t'_1}, \frac{1}{s'_1} \right); t_1 \right)$ for all $1 < t_1 < 2$ and all p_1, s_1 such that

$$\frac{t_1 - 1}{50(3t_1 - 2)} + \frac{99}{100} \frac{t_1 - 1}{t_1} \leq \frac{1}{p_1} \leq \frac{2t_1 - 2}{3t_1 - 2} \quad (14)$$

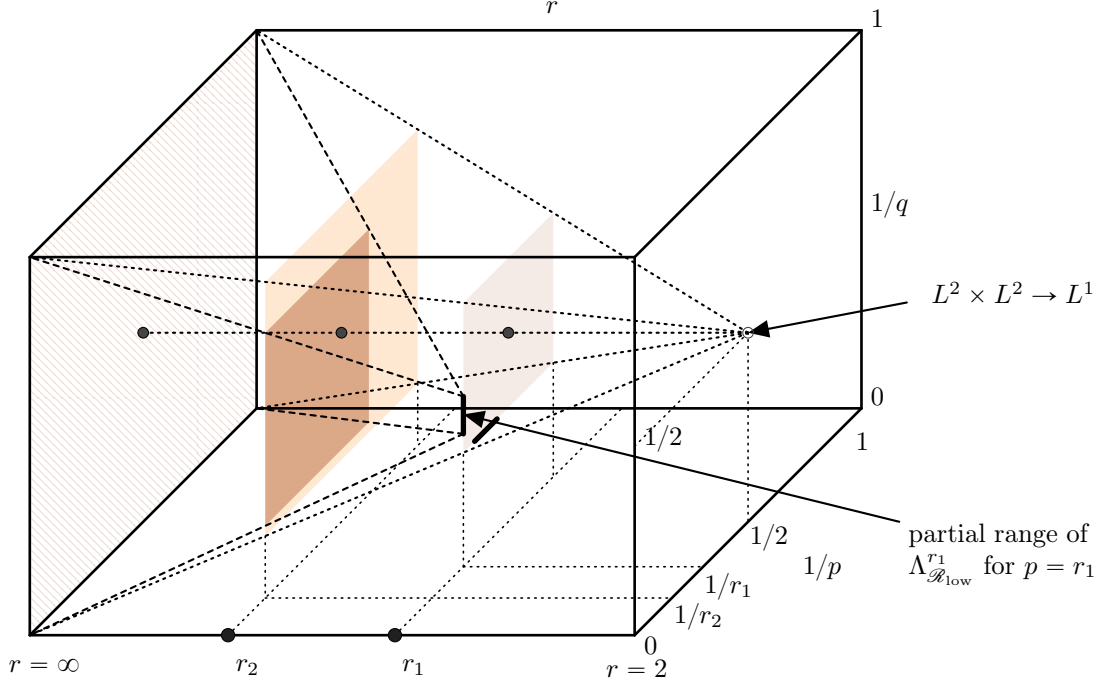


FIGURE 2. The interpolation argument in geometric form. The horizontal coordinate represents the exponent r , the other coordinates represent the (inverse) exponents $1/p, 1/q$. A point at coordinates $(r, 1/p, 1/q)$ represents the $L^p \times L^q \times L^{s'}(\ell^{r'}) \rightarrow \mathbb{R}$ estimate for the trilinear form $\Lambda_{\mathcal{R}}^r$, where $1/s = 1/p + 1/q$. In particular, the horizontal line through the axis of the parallelepiped represents (dualized) $L^2 \times L^2 \rightarrow L^1$ estimates. At $r = \infty$ the range of boundedness for $\Lambda_{\mathcal{R}}^\infty$ is the whole $[0, 1] \times [0, 1]$ (hatched) square in the $1/p-1/q$ plane, as given by Theorem 8.1. At exponent r_1 intermediate between 2 and ∞ , Proposition 7.1 gives for $\Lambda_{\mathcal{R}_{low}}^{r_1}$ the range of boundedness represented by the thickened interval above, as per label. The other thickened interval represents instead the range of boundedness for $\Lambda_{\mathcal{R}_{high}}^{r_1}$ given by Proposition 7.3. Interpolation between exponents r_1 and ∞ can be interpreted geometrically in the above picture as taking the convex hull of the corresponding endpoint ranges. This gives then for an intermediate exponent r_2 a range of boundedness (for $\Lambda_{\mathcal{R}_{low}}^{r_2}$) represented by the darker shaded area at $r = r_2$, a subset of the $r'_2 < p, q < r_2$ range represented by the lighter shaded area. As r_1 tends to 2, the darker shaded area fills the lighter one. Interpolation with the other thickened segment fills the same area in the limit.

and $1/p_1 + 1/t'_1 + 1/s'_1 = 1$. Fix then some other $r > 2$. We claim that for any p, q such that $r' < p, q < r$ we can choose $p_0, q_0, p_1, t_1, \theta$ so that

$$\begin{aligned} \frac{1}{p} &= \frac{1-\theta}{p_0} + \frac{\theta}{p_1}, \\ \frac{1}{q} &= \frac{1-\theta}{q_0} + \frac{\theta}{t'_1}, \\ \frac{1}{r'} &= \frac{1-\theta}{1} + \frac{\theta}{t_1}, \end{aligned} \tag{15}$$

where $1 < p_0, q_0 < \infty$, t_1 and p_1 satisfy condition (14), and $0 < \theta < 1$. By the above discussion, a direct application of Lemma 8.6 will then show that $\Lambda_{\mathcal{R}_{high}}^r$ is

of generalized restricted weak type $\left(\left(\frac{1}{p}, \frac{1}{q}, \frac{1}{s'}\right); r'\right)$. Notice that the major subset can be taken to be the one given by the value t_1 , since for $t_0 = 1$ the major subset is the set itself.

That we can find such quantities as above follows from the limiting case obtained by taking $t_1 = p_1 = 2$, that is assuming we have the estimate $\left(\left(\frac{1}{2}, \frac{1}{2}, 0\right); 1\right)$ (notice that for such a choice of parameters we don't know if the trilinear form is of generalized restricted weak type, but at the moment we are only concerned with showing the existence of solutions to system (15)). Indeed, in this case it's easy to solve the above system of equations and inequalities: one has $\frac{\theta}{2} = \frac{1}{r}$ and therefore $\frac{1}{p_0} = \left(\frac{1}{p} - \frac{1}{r}\right)\left(1 - \frac{2}{r}\right)^{-1}$ and $\frac{1}{q_0} = \left(\frac{1}{q} - \frac{1}{r}\right)\left(1 - \frac{2}{r}\right)^{-1}$. In the general case of $t_1 < 2$ (for which we know the trilinear form is of generalized restricted weak type), by choosing t_1 sufficiently close to 2 we can solve the system by a perturbation of the solution for the limiting case (by continuity), because the range where $\Lambda_{\mathcal{R}_{\text{high}}}^r$ is of generalized restricted weak type according to Proposition 7.1 is arbitrarily close to the $\left(\left(\frac{1}{2}, \frac{1}{2}, 0\right); 1\right)$ estimate.

Finally, observe that if $\Lambda_{\mathcal{R}_{\text{high}}}^r$ is of generalized restricted weak type $\left(\left(\frac{1}{p}, \frac{1}{q}, \frac{1}{s'}\right); r'\right)$ then $\tilde{\Lambda}_{\mathcal{R}_{\text{high}}}^r$ (as given by (1)) is of generalized restricted weak type $\left(\frac{1}{p}, \frac{1}{q}, \frac{1}{s'}\right)$ in the classical sense, and therefore we can conclude the strong $L^p \times L^q \times L^{s'} \rightarrow \mathbb{R}$ boundedness of $\tilde{\Lambda}_{\mathcal{R}_{\text{high}}}^r$ by classical multilinear interpolation. This is equivalent to the $L^p \times L^q \times L^{s'}(\ell^{r'}) \rightarrow \mathbb{R}$ boundedness of the trilinear form $\Lambda_{\mathcal{R}_{\text{high}}}^r$, of course.

The argument for $\Lambda_{\mathcal{R}_{\text{low}}}^r$ is essentially identical, and one can easily verify that it yields the same range of estimates. Thus we can conclude estimate (3) by summing up the separate contributions given by collections $\mathcal{R}_{\text{high}}, \mathcal{R}_{\text{low}}$. \square

APPENDIX A

In this appendix we show how to prove Corollary 1.5. What we want to show in particular is the boundedness of the trilinear form

$$\Xi_{\chi, \mathcal{R}}^r(f, g, \mathbf{h}) := \sum_{R \in \mathcal{R}} \int_{\mathbb{R}} \iint_{\mathbb{R}^2} \hat{f}(\xi) \hat{g}(\eta) \chi_R(\xi, \eta) e^{2\pi i(\xi + \eta)x} h_R(x) d\xi d\eta dx, \quad (16)$$

where χ_R is the smooth compactly supported bump function χ rescaled to fit the rectangle R . The boundedness will also be uniform in χ if χ is taken in a bounded subset² of $C_c^\infty(\mathbb{R}^2)$. We will perform a series of reductions and then show how the corollary is an immediate consequence of the following variant of Theorem 1.6, where dyadic rectangles are replaced by well-separated ones. First let us make the notion of well-separated rectangles rigorous.

Definition A.1. Let \mathcal{R} be a collection of rectangles. We say that \mathcal{R} is K -separated, with $K \geq 1$, if for any $R, R' \in \mathcal{R}$ such that $R \neq R'$ we have $KR \cap KR' = \emptyset$.

With this definition we can state the variant mentioned before.

Theorem (Variant of Theorem 1.6). *Let $r > 2$ and let \mathcal{R} be a 3-separated collection of rectangles in \mathbb{R}^2 . Let p, q, s be such that*

$$\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$$

²In particular, we ask that χ has support inside $B(2)$ and that $\|\partial_\xi^\alpha \partial_\eta^\beta \chi\|_{L^\infty} \leq C_{\alpha, \beta}$ for some absolute constants $C_{\alpha, \beta} > 0$ and all indices α, β .

and such that

$$r' < p, q < r, \quad r'/2 < s < r/2.$$

Then for every $f \in L^p$, $g \in L^q$ and $\mathbf{h} \in L^{s'}(\ell^{r'})$ we have

$$\Lambda_{\mathcal{R}}^r(f, g, \mathbf{h}) \lesssim_{r,p,q} \|f\|_{L^p} \|g\|_{L^q} \|\mathbf{h}\|_{L^{s'}(\ell^{r'})}. \quad (17)$$

Proof. We need only observe that if the rectangles are 3-separated, then for each rectangle $R \in \mathcal{R}$ there exists a rectangle \tilde{R} with $R \subset \tilde{R} \subset 3R$ that belongs to one of the dyadic (shifted) grids $\mathcal{D} \times \mathcal{D}'$, where $\mathcal{D}, \mathcal{D}'$ are either of the collections³

$$\begin{aligned} \mathcal{D}_0 &:= \{[2^k \ell, 2^k(\ell + 1)] : k, \ell \in \mathbb{Z}\}, \\ \mathcal{D}_1 &:= \{[2^k(\ell + \frac{(-1)^k}{3}), 2^k(\ell + 1 + \frac{(-1)^k}{3})] : k, \ell \in \mathbb{Z}\}, \\ \mathcal{D}_2 &:= \{[2^k(\ell - \frac{(-1)^k}{3}), 2^k(\ell + 1 - \frac{(-1)^k}{3})] : k, \ell \in \mathbb{Z}\}. \end{aligned}$$

Then one repeats the proof of Theorem 1.6 for each such product of grids using rectangles \tilde{R} . The details are left to the reader. \square

Now we claim that Corollary 1.5 follows for arbitrary disjoint rectangles if it holds for, say, 12-separated collections of rectangles. Indeed, let \mathcal{W} denote a Whitney decomposition of $[-1/2, 1/2]$ chosen so that $I \in \mathcal{W} \Rightarrow 16I \subset [-1/2, 1/2]$. Thus we have a decomposition of the square $[-1/2, 1/2]^2$ given by the rectangles in $\mathcal{W} \times \mathcal{W}$, where elements can be indexed with $(j, k) \in \mathbb{N}^2$ in such a way that the element corresponding to (j, k) has horizontal side of length $\lesssim \delta^j$ and vertical side of length $\lesssim \delta^k$ for some absolute constant $0 < \delta < 1$. We consider then a smooth partition of unity $\{\phi_{j,k}\}$ associated to the decomposition $\mathcal{W} \times \mathcal{W}$, that is $\mathbb{1}_{[-1/2, 1/2]^2} = \sum_{j,k} \phi_{j,k}$, which we can take so that each $\phi_{j,k}(\xi, \eta)$ factorizes as a tensor product $\phi_j \otimes \phi_k$ of functions of one variable and so that each $\phi_{j,k}$ is supported in the $4/3$ -enlargement of the rectangle of $\mathcal{W} \times \mathcal{W}$ indexed by (j, k) .

For a given rectangle R , we denote by Ψ_R the affine map that sends R to $[-1/2, 1/2]^2$ preserving orientation (thus $\chi_R = \chi \circ \Psi_R$); if the rectangle is the $4/3$ -enlargement of the one in $\mathcal{W} \times \mathcal{W}$ indexed by (j, k) we write Ψ_{jk} for this map. We decompose χ using the partition of unity above: we let $c_{jk} := \|\chi \phi_{jk}\|_\infty$ and $\chi_{jk} := c_{jk}^{-1}(\chi \cdot \phi_{jk}) \circ \Psi_{jk}^{-1}$, so that

$$\chi = \sum_{j,k} c_{jk} \chi_{jk} \circ \Psi_{jk}$$

and each χ_{jk} is L^∞ normalized. It's easy to see that $c_{jk} \lesssim \min(\gamma^j, \gamma^k)$ for some $0 < \gamma < 1$. Moreover, each χ_{jk} is adapted to $[-1/2, 1/2]^2$, in the sense that χ_{jk} is supported in $[-1/2, 1/2]^2$ and $\|\partial_\xi^\alpha \partial_\eta^\beta \chi_{jk}\|_{L^\infty} \lesssim_{\alpha,\beta} 1$.

Using this decomposition of χ we decompose the trilinear form $\Xi_{\chi, \mathcal{R}}^r$ as follows. Notice that $(\chi \cdot \phi_{jk}) \circ \Psi_R = c_{jk} \chi_{jk} \circ (\Psi_{jk} \circ \Psi_R)$, so if we define the rectangle R_{jk} to be $R_{jk} = (\Psi_{jk} \circ \Psi_R)^{-1}([-1/2, 1/2]^2)$ we can see that $\{3/4 R_{jk}\}_{jk}$ is a Whitney decomposition of R which is an affine image of $\mathcal{W} \times \mathcal{W}$. We define the collections \mathcal{R}_{jk} to be $\{R_{jk} \text{ s.t. } R \in \mathcal{R}\}$. Therefore we can write

$$\Xi_{\chi, \mathcal{R}}^r = \sum_{j,k} c_{jk} \Xi_{\chi_{jk}, \mathcal{R}_{jk}}^r,$$

and the boundedness of $\Xi_{\chi, \mathcal{R}}^r$ follows from the (uniform in j, k) boundedness of the $\Xi_{\chi_{jk}, \mathcal{R}_{jk}}^r$ thanks to the summability of c_{jk} . It's easy to verify from the construction above that the rectangles in \mathcal{R}_{jk} are indeed 12-separated, and hence it suffices to prove boundedness for 12-separated collections.

³Notice that \mathcal{D}_0 is the standard dyadic grid and $\mathcal{D}_1, \mathcal{D}_2$ also satisfy the combinatorial property $I, I' \in \mathcal{D}_j$ and $I \cap I' \neq \emptyset \Rightarrow I \subseteq I'$ or $I' \subseteq I$.

Now, fix χ and take \mathcal{R} to be 12-separated. We can assume that $\chi(\xi, \eta) = \chi_1(\xi)\chi_2(\eta)$ because we can always reduce to this case by a windowed Fourier series expansion⁴ (the coefficients will be summable thanks to the fact that χ is smooth). Letting $\phi_j(x) := \chi'_j(x)$ for $j = 1, 2$ we can write

$$\begin{aligned} & \Xi_{\chi, \mathcal{R}}^r(f, g, \mathbf{h}) \\ &= \sum_{R \in \mathcal{R}} \int_{\mathbb{R}} \iint_{\mathbb{R}^2} \int_{-\infty}^{\Psi_{R_2}(\eta)} \int_{-\infty}^{\Psi_{R_1}(\xi)} \hat{f}(\xi) \hat{g}(\eta) \mathbf{1}_R(\xi, \eta) \phi_1(\zeta) \phi_2(\theta) e^{2\pi i(\xi+\eta)x} h_R(x) d\zeta d\theta d\xi d\eta dx. \end{aligned}$$

We want to use Fubini to take the integration in $d\zeta d\theta$ out, so for ζ, θ fixed and $R \in \mathcal{R}$ we let $R_{\zeta, \theta} := R \cap ([\Psi_{R_1}^{-1}(\zeta), +\infty) \times [\Psi_{R_2}^{-1}(\theta), +\infty))$ and let $\mathcal{R}_{\zeta, \theta} := \{R_{\zeta, \theta} \text{ s.t. } R \in \mathcal{R}\}$. We see that by Fubini we can bound

$$\begin{aligned} & |\Xi_{\chi, \mathcal{R}}^r(f, g, \mathbf{h})| \\ & \leq \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\phi_1(\zeta)| |\phi_2(\theta)| \left| \sum_{R \in \mathcal{R}} \int_{\mathbb{R}} \iint_{\mathbb{R}^2} \hat{f}(\xi) \hat{g}(\eta) \mathbf{1}_{R_{\zeta, \theta}}(\xi, \eta) e^{2\pi i(\xi+\eta)x} h_R(x) d\xi d\eta dx \right| d\zeta d\theta \\ & = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\phi_1(\zeta)| |\phi_2(\theta)| |\Lambda_{\mathcal{R}_{\zeta, \theta}}^r(f, g, \mathbf{h})| d\zeta d\theta. \end{aligned}$$

To conclude, notice that $\mathcal{R}_{\zeta, \theta}$ is certainly a 3-separated collection of rectangles and therefore the variant of Theorem 1.6 above applies, giving a bound independent of ζ, θ . Since $\phi_1, \phi_2 \in L^1$, integrating in $d\zeta d\theta$ concludes the argument.

REFERENCES

- [1] Cristina Benea, *Vector-valued Extensions for Singular Bilinear Operators and Applications*, Ph.D. thesis, Cornell University, <https://ecommons.cornell.edu/handle/1813/40903>, 2015.
- [2] Cristina Benea and Frédéric Bernicot, *A bilinear Rubio de Francia inequality for arbitrary squares*, Forum Math. Sigma **4** (2016), e26, 34. MR 3569060
- [3] Cristina Benea and Camil Muscalu, *Rubio de Francia theorems revisited: linear and bilinear case*, preprint.
- [4] ———, *Multiple vector-valued inequalities via the helicoidal method*, Anal. PDE **9** (2016), no. 8, 1931–1988. MR 3599522
- [5] Frédéric Bernicot, *L^p estimates for non-smooth bilinear Littlewood-Paley square functions on \mathbb{R}* , Math. Ann. **351** (2011), no. 1, 1–49. MR 2824844
- [6] Frédéric Bernicot and Saurabh Shrivastava, *Boundedness of smooth bilinear square functions and applications to some bilinear pseudo-differential operators*, Indiana Univ. Math. J. **60** (2011), no. 1, 233–268. MR 2952417
- [7] Frédéric Bernicot and Marco Vitturi, *Bilinear Rubio de Francia inequalities for collections of non-smooth squares*, Publicacions Matemàtiques, to appear.
- [8] J. Bourgain, *On square functions on the trigonometric system*, Bull. Soc. Math. Belg. Sér. B **37** (1985), no. 1, 20–26.
- [9] Lennart Carleson, *On the Littlewood-Paley theorem*, Inst. Mittag-Leffler report (1967).
- [10] Antonio Córdoba, *Some remarks on the Littlewood-Paley theory*, Rend. Circ. Mat. Palermo (2) (1981), no. suppl. 1, 75–80. MR 639467
- [11] Michael Cowling and Terence Tao, *Some light on Littlewood-Paley theory*, Math. Ann. **321** (2001), no. 4, 885–888. MR 1872532
- [12] Jean-Lin Journé, *Calderón-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana **1** (1985), no. 3, 55–91. MR 836284
- [13] Michael T. Lacey, *On bilinear Littlewood-Paley square functions*, Publ. Mat. **40** (1996), no. 2, 387–396. MR 1425626
- [14] ———, *Issues related to Rubio de Francia’s Littlewood-Paley inequality*, New York Journal of Mathematics. NYJM Monographs, vol. 2, State University of New York, University at Albany, Albany, NY, 2007. MR 2293255

⁴This has the effect of reducing the separation constant slightly.

- [15] Parasar Mohanty and Saurabh Shrivastava, *A note on the bilinear Littlewood-Paley square function*, Proc. Amer. Math. Soc. **138** (2010), no. 6, 2095–2098. MR 2596047
- [16] Camil Muscalu and Wilhelm Schlag, *Classical and multilinear harmonic analysis. Vol. II*, Cambridge Studies in Advanced Mathematics, vol. 138, Cambridge University Press, Cambridge, 2013. MR 3052499
- [17] Camil Muscalu, Terence Tao, and Christoph Thiele, *L^p estimates for the biest. II. The Fourier case*, Math. Ann. **329** (2004), no. 3, 427–461. MR 2127985
- [18] Richard Oberlin, Andreas Seeger, Terence Tao, Christoph Thiele, and James Wright, *A variation norm Carleson theorem*, J. Eur. Math. Soc. (JEMS) **14** (2012), no. 2, 421–464. MR 2881301
- [19] José L. Rubio de Francia, *A Littlewood-Paley inequality for arbitrary intervals*, Rev. Mat. Iberoamericana **1** (1985), no. 2, 1–14. MR 850681
- [20] Prabath Silva, *Vector-valued inequalities for families of bilinear Hilbert transforms and applications to bi-parameter problems*, J. Lond. Math. Soc. (2) **90** (2014), no. 3, 695–724. MR 3291796
- [21] Per Sjölin, *A note on Littlewood-Paley decompositions with arbitrary intervals*, J. Approx. Theory **48** (1986), no. 3, 328–334. MR 864755
- [22] Fernando Soria, *A note on a Littlewood-Paley inequality for arbitrary intervals in \mathbf{R}^2* , J. London Math. Soc. (2) **36** (1987), no. 1, 137–142. MR 897682
- [23] Christoph Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics, vol. 105, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2006. MR 2199086

CNRS - UNIVERSITÉ DE NANTES, LABORATOIRE JEAN LERAY, 2, RUE DE LA HOUSSINIÈRE
44322 NANTES CEDEX 3, FRANCE

E-mail address: frederic.bernicot@univ-nantes.fr and marco.vitturi@univ-nantes.fr