Bilinear Rubio de Francia inequalities for collections of non-smooth squares

Frederic Bernicot, Marco Vitturi

BILINEAR RUBIO DE FRANCIA INEQUALITIES FOR COLLECTIONS OF NON-SMOOTH SQUARES

FRÉDÉRIC BERNICOT AND MARCO VITTURI

Abstract. Let Ω be a collection of disjoint dyadic squares ω, let π_ω denote the non-smooth bilinear projection onto ω

$$\pi_\omega(f,g)(x) := \int\int \mathbb{1}_\omega(\xi,\eta)\widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\mathrm{d}\xi\mathrm{d}\eta$$

and let $r>2$. We show that the bilinear Rubio de Francia operator

$$\Big(\sum_{\omega\in\Omega}|\pi_\omega(f,g)|^r\Big)^{1/r}$$

is $L^p\times L^q\to L^s$ bounded with constant at most $O_\varepsilon(\#\Omega^\varepsilon)$ for any $\varepsilon>0$ whenever $1/p+1/q=1/s$, $r'<p,q<r$, $r'/2<s<r/2$.

1. Introduction

Classical Littlewood-Paley theory on the real line is a staple of linear harmonic analysis and has proven vastly important in its development. It encodes a principle of orthogonality in L^p spaces even when $p\neq 2$ for dyadically separated frequencies, and can thus be seen as a substitute for Plancherel's identity; this usually allows one to decouple the action of a multiplier on each dyadic frequency and deal with them separately. Generalizations of the linear Littlewood-Paley inequalities were first considered by Carleson in [8] (later reproved in a different way by Cordoba in [9]) for the special case where one replaces the Littlewood-Paley dyadic intervals $[2^k,2^{k+1}], k\in\mathbb{Z}$ by the intervals $[n,n+1], n\in\mathbb{Z}$. Later, Rubio de Francia in [18] extended Carleson's result to arbitrary collections of disjoint intervals. In particular, he proved the following: let $\mathcal{I}=\{I_j\}_j$ be a collection of disjoint intervals and define the Rubio de Francia square function

$$\mathrm{RdF}^2_{\mathcal{I}}f(x) := \Big(\sum_j |\pi_{I_j}f(x)|^2\Big)^{1/2},$$

where π_I is the frequency projection operator defined by

$$\widehat{\pi_I f}(\xi) := \mathbb{1}_I(\xi)\widehat{f}(\xi);$$

then for all $2\leqslant p<\infty$ it holds that for all $f\in L^p(\mathbb{R})$

$$\|\mathrm{RdF}^2_{\mathcal{I}}f\|_{L^p(\mathbb{R})}\lesssim_p \|f\|_{L^p(\mathbb{R})} \tag{1}$$

(with constant independent of $\mathcal{I}$). The inequality is false in general for $p<2$, as was known since [8] - this corresponds to a failure of orthogonality in L^p spaces for small p's. More in general, by the same methods one can prove for a generic[1] $r>2$ that the Rubio de Francia r-function

$$\mathrm{RdF}^r_{\mathcal{I}}f(x) := \Big(\sum_j |\pi_{I_j}f(x)|^r\Big)^{1/r}$$

[1]The condition $r\geqslant 2$ is necessary, as can be seen for example by considering the collection of Littlewood-Paley intervals.

is bounded on L^p for all $r' < p < \infty$ (the lowerbound being sharp; see [10] for a proof). Known proofs of (1) (see [12], [20], [21], [14]) rely on an interpolation between the trivial L^2 case and (a substitute for) the L^∞ endpoint (or dually between L^2 and H^1, as in [6]). See also [3] about an alternative proof for such inequalities as well as for a bilinear generalization, involving a collection of paraproducts-type operators. Higher dimensional versions of the inequalities have been first shown in [12].
A natural question is whether similar orthogonality principles exist in the bilinear setting and to what extent. That is, given bilinear multiplier operators T_j with disjoint frequency supports in the frequency plane $\widehat{\mathbb{R}^2}$, under what conditions does it hold that, say, the square function

$$\Big(\sum_j |T_j(f,g)|^2\Big)^{1/2}$$

is bounded from $L^p \times L^q$ to L^s? Some results are known for special collections of supports. Perhaps the first one is to be found in Lacey's [13], where he proves the $L^p \times L^q \to L^2$ boundedness of the bilinear square function

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

for $p, q \geqslant 2$ such that $1/p + 1/q = 1/2$ (later extended to any $1/p + 1/q = 1/s$ in [15],[5]), where χ is a C^∞ function that is identically 1 in $[-1/2, 1/2]$ and vanishes outside $[-1, 1]$. Thus here the frequency supports consist of (smoothened) diagonal strips of roughly unit width and unit separation. This was later extended by the first author in [4] to the case of non-smooth diagonal strips, that is where one replaces the smooth function χ above by the non-smooth $\mathbb{1}_{[-1/2,1/2]}$. The discontinuity at the boundary of the strip makes the analysis inherently more complicated (the same phenomenon that arises in the study of, for example, the Bilinear Hilbert transform).

In this paper we are interested in bilinear operators built out of bilinear projections whose frequency supports consist of squares in the frequency plane $\widehat{\mathbb{R}^2}$. Here the reference we have in mind is [2] by Benea and the first author, in which the following bilinear versions of Rubio de Francia r-functions are considered: let Ω be a collection of disjoint squares in $\widehat{\mathbb{R}^2}$ and let r be fixed, then define the operator

$$S^r_\Omega(f,g)(x) := \Big(\sum_{\omega\in\Omega} \Big|\int\int \chi_\omega(\xi,\eta)\widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\mathrm{d}\xi\mathrm{d}\eta\Big|\Big)^{1/r},$$

where χ_ω is a C^∞ function that is identically 1 on $\frac{1}{2}\omega$ and vanishes outside ω. In [2] the authors prove the following theorem:

Theorem 1.1 ([2])**.** *Let Ω be a collection of disjoint squares in $\widehat{\mathbb{R}^2}$ and let $r > 2$. Then*

$$\|S^r_\Omega(f,g)\|_{L^s(\mathbb{R})} \lesssim_{p,q} \|f\|_{L^p}\|g\|_{L^q} \tag{2}$$

for all p, q, s such that $1/p + 1/q = 1/s$, $r' < p, q < \infty$, $r'/2 < s < r$. In particular, the constant is independent of Ω.

This result is to be thought of as a bilinear orthogonality principle for collections of (smoothened) frequency squares in the same way as the Rubio de Francia theorem is for the linear case. Observe however that the square function case $r = 2$ is not covered by the theorem - its boundedness is currently an open problem. We remark that the condition $r' < p, q$ is necessary (to see why it suffices to consider a collection of squares like the one given in Example 1 below).

Our interest here is to extend the results of [2] to the case where the smooth characteristic function χ_ω above is replaced by the non-smooth characteristic function $\mathbb{1}_\omega$. In particular, let Ω be a collection of disjoint squares $\omega = \omega_1 \times \omega_2$ in $\widehat{\mathbb{R}^2}$ and denote by π_ω the non-smooth bilinear frequency projection onto the square ω, that is

$$\pi_\omega(f,g)(x) := \int\int \mathbb{1}_\omega(\xi,\eta)\widehat{f}(\xi)\widehat{g}(\eta)e^{2\pi i(\xi+\eta)x}\mathrm{d}\xi\mathrm{d}\eta,$$

which in particular factorizes as $\pi_\omega = \pi_{\omega_1} \otimes \pi_{\omega_2}$. We are interested in the bilinear operator

$$f,g \mapsto T^r_\Omega(f,g)(x) := \Big(\sum_{\omega\in\Omega} |\pi_\omega(f,g)(x)|^r\Big)^{1/r}, \qquad r \geqslant 2,$$

and specifically in proving bounds of the form

$$\|T^r_\Omega(f,g)\|_{L^s} \leqslant C\|f\|_{L^p}\|g\|_{L^q}; \tag{3}$$

we denote by $C_{p,q,s,\Omega}$ the best constant C such that the above inequality holds for all $f \in L^p, g \in L^q$ (we consider r fixed). The usual scaling argument shows that a necessary condition is that the exponents p, q, s satisfy Hölder's relationship, that is it must be

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

(and therefore $C_{p,q,s,\Omega} = C_{p,q,\Omega}$).
We consider some examples in order to get acquainted with the problem at hand.

Example 1. Let $r \geqslant 2$. Suppose Ω_{line} consists of an arbitrary number of disjoint squares that all intersect a given vertical line, that is there exists a frequency ξ_0 such that for every $\omega \in \Omega_{\text{line}}$ we have $\xi_0 \in \omega_1$. Observe that the frequency intervals ω_2 must be all disjoint. We can bound pointwise

$$T^r_{\Omega_{\text{line}}}(f,g)(x) \leqslant \Big(\sum_{\omega\in\Omega_{\text{line}}} |\pi_{\omega_2}g(x)|^r\Big)^{1/r} \cdot \sup_{\omega\in\Omega_{\text{line}}} |\pi_{\omega_1}f(x)| \leqslant \mathrm{RdF}^r(g)(x)\cdot\mathscr{C}f(x),$$

where $\mathscr{C}$ denotes the Carleson operator, which is bounded on L^p for all $1 < p < \infty$ (by the Carleson-Hunt theorem, [7], [11]), and therefore we get that for $p > 1$ and $q > r'$ (or $q \geqslant 2$ if $r \geqslant 2$) we can estimate for this particular collection

$$\|T^r_{\Omega_{\text{line}}}(f,g)(x)\|_{L^s} \lesssim_{p,q} \|f\|_{L^p}\|g\|_{L^q},$$

or in other words $C_{p,q,\Omega_{\text{line}}} \lesssim_{p,q} 1$ in the stated range.

Example 2. Let $r > 2$ be fixed and consider now a collection of $N^{1/2} \times N^{1/2}$ points in $\widehat{\mathbb{R}^2}$ arranged in a rectangular grid with large spacing, and suppose that each point labeled by (i,j) is the center of a square ω^{ij} and furthermore that the squares are all disjoint (their sidelengths can be all distinct). We let $\Omega_{\text{grid}} := \{\omega^{ij}\}_{i,j\leqslant N^{1/2}}$ and we try to bound $T^r_{\Omega_{\text{grid}}}$ in some range. Observe that since a priori

$$|\pi_\omega(f,g)(x)| \leqslant \mathscr{C}f(x)\cdot\mathscr{C}g(x)$$

we always have the trivial bound

$$\|T^r_{\Omega_{\text{grid}}}(f,g)\|_{L^s} \lesssim_{p,q} N^{1/r}\|f\|_{L^p}\|g\|_{L^q}$$

for $p, q > 1$. We can beat this trivial bound of $C_{p,q,\Omega_{\text{grid}}} \lesssim_{p,q} N^{1/r}$ by the following argument: since for a fixed i the squares ω^{ij} are such that ω^{ij}_1 all contain a same

frequency as in the example above, we can bound pointwise

$$\begin{aligned}\Big(\sum_{\omega\in\Omega_{\rm grid}}|\pi_\omega(f,g)(x)|^r\Big)^{1/r}=&\Big(\sum_{i\leqslant N^{1/2}}\sum_{j\leqslant N^{1/2}}|\pi_{\omega_1^{ij}}f(x)\cdot\pi_{\omega_2^{ij}}g(x)|^r\Big)^{1/r}\\ \leqslant&\mathscr{C}f(x)\cdot\Big(\sum_{i\leqslant N^{1/2}}\sum_{j\leqslant N^{1/2}}|\pi_{\omega_2^{ij}}g(x)|^r\Big)^{1/r}\\ \leqslant&\mathscr{C}f(x)\cdot\Big(\sum_{i\leqslant N^{1/2}}(\mathrm{Var}_r\mathscr{C}g(x))^r\Big)^{1/r}\\ =&N^{1/2r}\mathscr{C}f(x)\mathrm{Var}_r\mathscr{C}g(x),\end{aligned}$$

where $\mathrm{Var}_r\mathscr{C}$ is the Variational Carleson operator

$$\mathrm{Var}_r\mathscr{C}f(x):=\sup_M\sup_{\xi_1<\cdots<\xi_M}\Big(\sum_{j=1}^{M-1}|\pi_{[\xi_j,\xi_{j+1}]}f(x)|^r\Big)^{1/r}.$$

It is known from [17] that this operator is $L^p\to L^p$ bounded for $r'<p<\infty$ if $r>2$, as is the case, and therefore we get for the range $p>1, q>r'$ an improvement in the dependence of the constant on the cardinality of Ω (specifically, $C_{p,q,\Omega_{\rm grid}}\lesssim_{p,q}N^{1/2r}$ instead of $N^{1/r}$).

It is natural to conjecture that for some range of exponents (possibly as large as[2] $p,q>r'$, like in [2]) one should have $C_{p,q,\Omega}\lesssim_{p,q}1$ for every admissible Ω, or in other words that inequality (3) should hold with constant $C_{p,q,\Omega}$ independent of Ω, and specifically independent of its cardinality $\#\Omega$. Simple pointwise arguments like the one given in Example 2 are unlikely to give such a result. However, by combining similar observations with the time-frequency analysis of [2] and some further ideas from [1], [3] (as for example the consideration in the time-frequency analysis of an exceptional subset built from non-local operators), we are able to reduce the dependence of the constant $C_{p,q,\Omega}$ to be at most of logarithmic type in $\#\Omega$ and otherwise independent of the specific collection[3]. More precisely, we show that

Theorem 1.2. *Let $r>2$ be fixed. Then for all p,q,s such that*

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

and

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

it holds that for every arbitrary finite collection Ω of disjoint dyadic squares in $\widehat{\mathbb{R}^2}$ and for every $\varepsilon>0$ the estimate

$$\Big\|\Big(\sum_{\omega\in\Omega}|\pi_\omega(f,g)(x)|^r\Big)^{1/r}\Big\|_{L^s}\lesssim_{\varepsilon,p,q,r}\#\Omega^\varepsilon\|f\|_{L^p}\|g\|_{L^q},\tag{4}$$

holds true for every $f\in L^p, g\in L^q$.

Remark 1.3. In Theorem 1.1 (from [2]) above, the statement encompasses arbitrary non-dyadic squares; this is because of the flexibility provided by the smoothness of the χ_ω functions. However, in the non-smooth case things are not as simple. One can replace the assumption that the squares are dyadic with a well-separation

[2]The $p,q>r'$ range is achieved by product-like collections of rectangles, that is collections of the form $\Omega=\{I\times J\text{ s.t. }I\in\mathcal{I},J\in\mathcal{J}\}$, where $\mathcal{I},\mathcal{J}$ are collections of disjoint intervals; this can be readily seen by a factorization of the operator and an application of Rubio de Francia's theorem.

[3]See Lemma 2.7 and Proposition 2.12 for the sources of this logarithmic-type loss in the argument.

assumption: namely, Theorem 1.2 still holds if we assume that Ω is a finite collection of arbitrary squares such that $4\omega \cap 4\omega' = \varnothing$ whenever $\omega \neq \omega'$. In the linear case it's always possible to reduce to a well-separated case (see [18]) by means of classical Littlewood-Paley theory, but in the bilinear case such tools are not currently available.

Remark 1.4. The condition $r > 2$ is a shortcoming we inherit from [2]. However, for the $r = 2$ case, one can deduce from the above theorem and Hölder's inequality that T^2_Ω is $L^2 \times L^2 \to L^1$ bounded with constant at most $O_\varepsilon(\#\Omega^\varepsilon)$ for any $\varepsilon > 0$. Indeed, we can bound pointwise $T^2_\Omega(f,g) \leqslant \#\Omega^\varepsilon T^r_\Omega(f,g)$, where $1/2 = 1/r + \varepsilon$, and conclude using Theorem 1.2 for exponents $p = q = 2$.

This result can be thought of as evidence in favor of the natural conjecture stated above. Observe the range of boundedness provided by Theorem 1.2 is smaller than the corresponding one in Theorem 1.1 above. We explain the reason why in Remark 2.22. Figure 1 in section §2.7 provides a graphical illustration of the range obtained in Theorem 1.2.
The proof of Theorem 1.2 is presented in section §2, and is split into a number of steps. The result is obtained by interpolation between a boundedness result for T^∞_Ω (a trivial consequence of the Carleson-Hunt theorem) and a partial boundedness result for T^r_Ω when r is close to 2. The latter is obtained by adapting the time-frequency methods of [2] to our setup, but using non-local operators to construct the exceptional set as in [1], [3]. The necessary preliminaries are carried out in sections §2.1 - §2.6. The proof is concluded in §2.7, where the particular interpolation result we will use (Lemma 2.21) is also presented. Finally, we present a simple application in §3.

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. Proof of Theorem 1.2

We let in the following $N := \#\Omega$. We can reduce the problem by linearization of the ℓ^r norm and duality to the following: given $f \in L^p, g \in L^q, h \in L^{s'}$ define the trilinear form

$$\Lambda_r(f,g,h) := \int_{\mathbb{R}} \sum_{\omega\in\Omega} \pi_{\omega_1} f(x) \pi_{\omega_2} g(x) h_\omega(x) \mathrm{d}x, \tag{5}$$

where $h_\omega(x) := h(x)\epsilon_\omega(x)$ and $\{\varepsilon_\omega(x)\}_{\omega\in\Omega}$ satisfies $\|\{\epsilon_\omega(x)\}_\omega\|_{\ell^{r'}} \leqslant 1$ for every $x \in \mathbb{R}$; then it suffices to prove that

$$|\Lambda_r(f,g,h)| \lesssim_{\varepsilon,p,q,r} N^\varepsilon \|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^{s'}}$$

uniformly in $\{\varepsilon_\omega(x)\}_{\omega\in\Omega}$. Thus we can further reduce the problem to that of bounding the trilinear form

$$\Lambda(f,g,\mathbf{h}) := \int_{\mathbb{R}} \sum_{\omega\in\Omega} \pi_{\omega_1} f(x) \pi_{\omega_2} g(x) h_\omega(x) \mathrm{d}x, \tag{6}$$

where $\mathbf{h} = \{h_\omega\}_{\omega\in\Omega}$ is a generic element of $L^{s'}(\ell^{r'})$.

2.1. Discretization of the trilinear form.

We perform the usual discretization procedure on the trilinear form Λ, except this time we will not resolve the singularities using Whitney cubes. We have (using Radon duality, with $\mathrm{d}\sigma$ the induced

Lebesgue measure on the plane $\xi_1+\xi_2+\xi_3=0$)

$$\begin{aligned}\Lambda(f,g,\mathbf{h}) &= \int_{\mathbb{R}} \sum_{\omega\in\Omega} f*\widetilde{\mathbb{1}_{\omega_1}}(x) g*\widetilde{\mathbb{1}_{\omega_2}}(x) h_\omega(x)\mathrm{d}x \\ &= \sum_{\omega\in\Omega}\int_{\xi_1+\xi_2+\xi_3=0} \widehat{f}(\xi_1)\mathbb{1}_{\omega_1}(\xi_1)\widehat{g}(\xi_2)\mathbb{1}_{\omega_1}(\xi_2)\widehat{h_\omega}(\xi_3)\mathrm{d}\sigma(\xi_1,\xi_2,\xi_3) \\ &= \sum_{\omega\in\Omega}\int_{\xi_1+\xi_2+\xi_3=0} \widehat{f}(\xi_1)\mathbb{1}_{\omega_1}(\xi_1)\widehat{g}(\xi_2)\mathbb{1}_{\omega_1}(\xi_2)\widehat{h_\omega}(\xi_3)\chi_{\omega_3}(\xi_3)\mathrm{d}\sigma(\xi_1,\xi_2,\xi_3) \\ &= \sum_{\omega\in\Omega}\int_{\mathbb{R}} f*\widetilde{\mathbb{1}_{\omega_1}}(x) g*\widetilde{\mathbb{1}_{\omega_2}}(x) h_\omega*\widetilde{\chi_{\omega_3}}(x)\mathrm{d}x,\end{aligned}$$

where we have denoted $\omega_3:=2(-\omega_1-\omega_2)$ and χ_{ω_3} is a smoothed out characteristic function, identically equal to 1 on $-\omega_1-\omega_2$ and identically vanishing outside ω_3. Now, although the kernels decay very slowly, the functions $f*\widetilde{\mathbb{1}_{\omega_j}}$ are morally still roughly constant in modulus at scale $|\omega_j|^{-1}=:|\omega|^{-1}$, and therefore it makes sense to do the following changes of variable:

$$\begin{aligned}&\sum_{\omega\in\Omega}\int_{\mathbb{R}} f*\widetilde{\mathbb{1}_{\omega_1}}(x) g*\widetilde{\mathbb{1}_{\omega_2}}(x) h_\omega*\widetilde{\chi_{\omega_3}}(x)\mathrm{d}x \\ &\quad= \sum_{\omega\in\Omega}|\omega|^{-1}\int_{\mathbb{R}} f*\widetilde{\mathbb{1}_{\omega_1}}(|\omega|^{-1}y) g*\widetilde{\mathbb{1}_{\omega_2}}(|\omega|^{-1}y) h_\omega*\widetilde{\chi_{\omega_3}}(|\omega|^{-1}y)\mathrm{d}y \\ &\quad= \sum_{\omega\in\Omega}\sum_{n\in\mathbb{Z}}|\omega|^{-1}\int_0^1 f*\widetilde{\mathbb{1}_{\omega_1}}(|\omega|^{-1}(n+z)) g*\widetilde{\mathbb{1}_{\omega_2}}(|\omega|^{-1}(n+z)) \\ &\qquad\qquad\cdot h_\omega*\widetilde{\chi_{\omega_3}}(|\omega|^{-1}(n+z))\mathrm{d}z.\end{aligned}$$

In classical time-frequency analysis one rewrites the above form as an average over z of discrete sums of coefficients, each given by an inner product against suitably defined wavepackets associated to tiles in the time-frequency plane, and then proceeds to bound the discrete sums uniformly in z; the approach we will take however is different and will involve allowing only a single scale for each square ω, roughly speaking - a choice reflected in our definition of tri-tiles given below. This will allow us to do a time-frequency analysis of the trilinear form Λ free from wavepackets (although wavepackets are intrinsically present in some strong results that we will use off-the-shelf). Define then the tri-tiles as follows:

Definition 2.1. A tri-tile P is a triple of sets of the form

$$P=(P_1,P_2,P_3)=(I\times\omega_1, I\times\omega_2, I\times\omega_3)$$

where $\omega=\omega_1\times\omega_2\in\Omega$, $\omega_3=2(-\omega_1-\omega_2)$ as before and I is a dyadic interval of length[4] $|\omega|^{-1}$. Sets P_j for $j=1,2,3$ are referred to as tiles. Given a tri-tile P we denote by I_P the interval I above; we also denote by $\omega(P)=\omega_1(P)\times\omega_2(P)$ the frequency square associated to the tri-tile P. Finally, given a collection of tiles $\mathbb{P}$ we denote by $\Omega(\mathbb{P})$ the collection of frequency squares on which $\mathbb{P}$ is supported, namely

$$\Omega(\mathbb{P}):=\{\omega\in\Omega \text{ s.t. } \omega=\omega(P) \text{ for some } P\in\mathbb{P}\}.$$

Using Hölder's inequality on each summand above, we have

$$\begin{aligned}&\Big||\omega|^{-1}\int_0^1 f*\widetilde{\mathbb{1}_{\omega_1}}(|\omega|^{-1}(n+z)) g*\widetilde{\mathbb{1}_{\omega_2}}(|\omega|^{-1}(n+z)) h_\omega*\widetilde{\chi_{\omega_3}}(|\omega|^{-1}(n+z))\mathrm{d}z\Big| \\ &\qquad\leqslant \|f*\widetilde{\mathbb{1}_{\omega_1}}\|_{L^2(I_P)}\|g*\widetilde{\mathbb{1}_{\omega_2}}\|_{L^2(I_P)}\|h_\omega*\widetilde{\chi_{\omega_3}}\|_{L^\infty(I_P)},\end{aligned}$$

[4]So that it's always $|I||\omega_j|\sim 1$ for $j=1,2,3$.

where the tri-tile P is given by

$$I_P = [|\omega|^{-1}n, |\omega|^{-1}(n+1)]$$

and

$$P = (I_P \times \omega_1, I_P \times \omega_2, I_P \times \omega_3).$$

We introduce the shorthand notation

$$\begin{aligned} f(P_1) &:= \|f * \widetilde{\mathbb{1}_{\omega_1}}\|_{L^2(I_P)}, \\ g(P_2) &:= \|g * \widetilde{\mathbb{1}_{\omega_2}}\|_{L^2(I_P)}, \\ \mathbf{h}(P_3) &:= \|h_{\omega(P)} * \widetilde{\chi_{\omega_3}}\|_{L^\infty(I_P)}. \end{aligned}$$

We have therefore that, if $\mathbb{P}$ denotes the collection of all possible tri-tiles (obtained by letting ω range in Ω and $n \in \mathbb{Z}$, in the above notation), the trilinear form Λ is bounded by the discretized sum

$$|\Lambda(f, g, \mathbf{h})| \leqslant \sum_{P \in \mathbb{P}} f(P_1) g(P_2) \mathbf{h}(P_3) =: \Lambda_{\mathbb{P}}(f, g, \mathbf{h}).$$

The reason for this unusual choice of coefficients will become clear later in light of Lemma 2.7 and Proposition 2.12 below (see particularly Remark 2.13). In the rest of the section we will concentrate on bounding the discretized sum.

2.2. **Columns and rows.** We introduce here some structured collections of tri-tiles, originating from [2], that will be fundamental to our analysis of the trilinear form Λ. They are to be thought of as the analogue for our setup of trees, in the language of classical time-frequency analysis.

Definition 2.2. A collection of tri-tiles $\mathcal{C}$ is a *column* if there exists a tri-tile $T \in \mathcal{C}$, referred to as the top of $\mathcal{C}$, such that for every $P \in \mathcal{C}$

$$\omega_1(P) \supseteq \omega_1(T)$$

and

$$I_P \subseteq I_T.$$

Analogously, a collection of tri-tiles $\mathcal{R}$ is a *row* if there exists a tri-tile $T \in \mathcal{R}$, referred to as the top of $\mathcal{R}$, such that for every $P \in \mathcal{R}$

$$\omega_2(P) \supseteq \omega_2(T)$$

and

$$I_P \subseteq I_T.$$

Given a column or row $\mathcal{T}$ we will use $\mathrm{Top}(\mathcal{T})$ to denote its top.

Remark 2.3. Observe that if $\mathcal{C}$ is a column then the collection of tiles $\{P_1 \text{ s.t. } P \in \mathcal{C}\}$ is overlapping, while the collection of tiles $\{P_2 \text{ s.t. } P \in \mathcal{C}\}$ is lacunary (because the frequency squares ω are disjoint). The reverse holds for a row. This will be important later on.

We show below that when $\mathcal{C}$ is a column we can give a good bound on $\Lambda_{\mathcal{C}}$ (and similarly for rows). In particular, we argue almost exactly as in [2] and bound the

discretized sum restricted to $\mathcal{C}$ as follows:

$$\begin{aligned}\sum_{P\in\mathcal{C}} f(P_1)g(P_2)\mathbf{h}(P_3) &\leqslant \Big(\sup_{P\in\mathcal{C}}\frac{f(P_1)}{|I_P|^{1/2}}\Big)\sum_{P\in\mathcal{C}} g(P_2)\mathbf{h}(P_3)|I_P|^{1/2}\\ &= \Big(\sup_{P\in\mathcal{C}}\frac{f(P_1)}{|I_P|^{1/2}}\Big)\sum_{P\in\mathcal{C}} g(P_2)|I_P|^{-(1/2-1/r)}\mathbf{h}(P_3)|I_P|^{1-1/r}\\ &\leqslant \Big(\sup_{P\in\mathcal{C}}\frac{f(P_1)}{|I_P|^{1/2}}\Big)\Big(\sum_{P\in\mathcal{C}} g(P_2)^r|I_P|^{-(r-2)/2}\Big)^{1/r}\\ &\quad\times\Big(\sum_{P\in\mathcal{C}}\mathbf{h}(P_3)^{r'}|I_P|\Big)^{1/r'}.\end{aligned}$$

Then, for the term in g we bound

$$\begin{aligned}\Big(\sum_{P\in\mathcal{C}} g(P_2)^r|I_P|^{-(r-2)/2}\Big)^{1/r} &= \Big(\sum_{P\in\mathcal{C}} g(P_2)^2\cdot\big(\frac{g(P_2)}{|I_P|^{1/2}}\big)^{r-2}\Big)^{1/r}\\ &\leqslant \Big(\sup_{P\in\mathcal{C}}\frac{g(P_2)}{|I_P|^{1/2}}\Big)^{(r-2)/r}\Big(\sum_{P\in\mathcal{C}} g(P_2)^2\Big)^{1/r}\end{aligned}$$

(notice we have introduced the same type of quantity that controls the contribution of f in here). As for the term in $\mathbf{h}$, we observe that

$$\begin{aligned}|I_P|\mathbf{h}(P_3)^{r'} = |I_P|\sup_{y\in I_P}|h_\omega * \widetilde{\chi_{\omega_3}}(y)|^{r'} &\leqslant |I_P|\big(\sup_{y\in I_P}\int |h_\omega(z)||\widetilde{\chi_{\omega_3}}(y-z)|\mathrm{d}z\big)^{r'}\\ &\lesssim |I_P|\Big(\sup_{y\in I_P}\int |h_\omega(z)|\big(1+\frac{|y-z|}{|I_P|}\big)^{-M}\frac{\mathrm{d}z}{|I_P|}\Big)^{r'}\\ &\leqslant |I_P|\Big(\int |h_\omega(z)|\sup_{y\in I_P}\big(1+\frac{|y-z|}{|I_P|}\big)^{-M}\frac{\mathrm{d}z}{|I_P|}\Big)^{r'}\\ &\lesssim \int |h_\omega(z)|^{r'}\Phi_{I_P}(z)\mathrm{d}z,\end{aligned}$$

where $M>0$ is a large number and Φ_I denotes some rapidly decaying function concentrated in the interval I. Now observe that for each fixed ω the tiles P which have ω as their frequency support have space support of fixed size $|I_P|=|\omega|^{-1}$, hence the intervals I_P are all disjoint. Define then for an interval I of length greater or equal to $|\omega|^{-1}$ the function

$$\Phi_I^\omega(x) := \sum_{\substack{J \text{ dyadic s.t.} J\subseteq I,\\ |J||\omega|=1}} \Phi_J(x);$$

notice Φ_I^ω is essentially ~ 1 inside I and decays like $(1+|\omega|\mathrm{dist}(I,x))^{-M+1}$ outside of it (see remark 2.11 for why we need to introduce such functions). We can thus bound

$$\begin{aligned}\sum_{P\in\mathcal{C}}\mathbf{h}(P_3)^{r'}|I_P| &\lesssim \sum_{P\in\mathcal{C}}\int |h_{\omega(P)}(z)|^{r'}\Phi_{I_P}(z)\mathrm{d}z\\ &\lesssim \sum_\omega\int |h_\omega|^{r'}\Phi^\omega_{I_\mathcal{C}}\mathrm{d}z.\end{aligned}$$

To summarize, we introduce sizes:

Definition 2.4 (Sizes)**.** For any collection of tri-tiles $\mathbb{P}$ define

$$\mathrm{Size}_f^1(\mathbb{P}) := \sup_{P\in\mathbb{P}} \frac{f(P_1)}{|I_P|^{1/2}} = \sup_{P\in\mathbb{P}} \Big(\fint_{I_P} |\pi_{\omega_1(P)} f|^2 \mathrm{d}x \Big)^{1/2},$$
$$\mathrm{Size}_g^2(\mathbb{P}) := \sup_{P\in\mathbb{P}} \frac{g(P_2)}{|I_P|^{1/2}} = \sup_{P\in\mathbb{P}} \Big(\fint_{I_P} |\pi_{\omega_2(P)} g|^2 \mathrm{d}x \Big)^{1/2},$$
$$\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}) := \sup_{\mathcal{T}\subset\mathbb{P}} \Big(\frac{1}{|I_{\mathcal{T}}|} \sum_{\omega\in\mathcal{T}} \int |h_\omega|^{r'} \Phi_{I_{\mathcal{T}}}^{\omega} \mathrm{d}z \Big)^{1/r'},$$

where the last supremum is taken over sub-collections $\mathcal{T}$ of $\mathbb{P}$ which are either rows or columns.

With this notation, what has been shown in this section can be summarized as

Proposition 2.5. *Let $\mathcal{C}$ be a column of tri-tiles, then*

$$|\Lambda_{\mathcal{C}}(f,g,\mathbf{h})| \lesssim \Big(\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 \Big)^{1/r} \mathrm{Size}_f^1(\mathcal{C}) \big[\mathrm{Size}_g^2(\mathcal{C})\big]^{(r-2)/r} \mathrm{Size}_{\mathbf{h}}^3(\mathcal{C}) |I_{\mathcal{C}}|,$$

and similarly, if $\mathcal{R}$ is a row of tri-tiles, we have

$$|\Lambda_{\mathcal{R}}(f,g,\mathbf{h})| \lesssim \Big(\frac{1}{|I_{\mathcal{R}}|} \sum_{P\in\mathcal{R}} f(P_1)^2 \Big)^{1/r} \big[\mathrm{Size}_f^1(\mathcal{R})\big]^{(r-2)/r} \mathrm{Size}_g^2(\mathcal{R}) \mathrm{Size}_{\mathbf{h}}^3(\mathcal{R}) |I_{\mathcal{R}}|.$$

2.3. **Size bounds.** We have the following immediate bounds for the sizes introduced above:

Proposition 2.6. *Let $j = 1, 2$ and let $\mathbb{P}$ be a collection of tri-tiles, then*

$$\mathrm{Size}_f^j(\mathbb{P}) \lesssim \sup_{P\in\mathbb{P}} \Big(\fint_{I_P} |\mathscr{C} f|^2 \mathrm{d}x \Big)^{1/2},$$

where $\mathscr{C}$ is the Carleson operator.

Proof. Obvious. □

We do not state an analogous proposition for Size^3 since this size is already in a convenient form.
Later on we will also need the following simple bound (notice the appearance of a logarithmic type loss in the constant):

Lemma 2.7. *Let $\#\Omega = N$. Let $\mathcal{C}$ be a column of tri-tiles and let $\varepsilon > 0$. Then*

$$\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 \leqslant N^{\varepsilon} \fint_{I_{\mathcal{C}}} |\mathrm{Var}_{q_0} \mathscr{C} g|^2 \mathrm{d}x,$$

where $q_0 = q_0(\varepsilon) > 2$ is given by

$$\frac{1}{2} = \frac{1}{q_0} + \frac{\varepsilon}{2}.$$

Clearly, an analogous statement holds for rows.

Proof. Observe that by unwrapping the definitions we have

$$\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 = \fint_{I_{\mathcal{C}}} \sum_{P\in\mathcal{C}} |\pi_{\omega_2(P)} g(x)|^2 \mathbb{1}_{I_P}(x) \mathrm{d}x,$$

thus by Hölder's inequality we can bound the above by

$$\fint_{I_{\mathcal{C}}} \Big(\sum_{P\in\mathcal{C}} |\pi_{\omega_2(P)} g(x)|^{q_0} \mathbb{1}_{I_P}(x) \Big)^{2/q_0} \Big(\sum_{P\in\mathcal{C}} \mathbb{1}_{I_P}(x) \Big)^{\varepsilon} \mathrm{d}x.$$

Since $\mathcal{C}$ is a column, the tiles P_2 are disjoint in the time-frequency space, and therefore we have the trivial bound

$$\| \sum_{P\in\mathcal{C}} \mathbb{1}_{I_P} \|_{L^\infty} \leqslant N;$$

for the same reason, given x, the frequency intervals $\omega_2(P)$ such that $x \in I_P$ are disjoint, so that we can bound pointwise

$$\Big(\sum_{P\in\mathcal{C}} |\pi_{\omega_2(P)} g(x)|^{q_0} \mathbb{1}_{I_P}(x) \Big)^{1/q_0} \leqslant \mathrm{Var}_{q_0} \mathscr{C}(g)(x),$$

and the claim follows. □

2.4. **Energies and energy estimates.** In this subsection we introduce the energies that will allow us to run a time-frequency argument for the trilinear form Λ.

Definition 2.8 (Energies)**.** We denote

$$\mathrm{Energy}^1_f(\mathbb{P}) := \sup_{n\in\mathbb{Z}} \sup_{\mathfrak{C}} 2^n \Big(\sum_{\mathcal{C}\in\mathfrak{C}} |I_{\mathsf{Top}(\mathcal{C})}| \Big)^{1/2},$$

where the inner supremum runs over the collections $\mathfrak{C}$ of disjoint columns in $\mathbb{P}$ such that for any column $\mathcal{C} \in \mathfrak{C}$ it is

$$\frac{f(\mathsf{Top}(\mathcal{C})_1)}{|I_{\mathsf{Top}(\mathcal{C})}|^{1/2}} \geqslant 2^n.$$

Define analogously $\mathrm{Energy}^2_g(\mathbb{P})$ with respect to rows of tri-tiles in the obvious way. Finally, we denote

$$\mathrm{Energy}^3_{\mathbf{h}}(\mathbb{P}) := \sup_{n\in\mathbb{Z}} \sup_{\mathfrak{T}} 2^n \Big(\sum_{\mathcal{T}\in\mathfrak{T}} |I_{\mathsf{Top}(\mathcal{T})}| \Big)^{1/r'},$$

where the inner supremum runs over the collections $\mathfrak{T}$ of disjoint rows and columns in $\mathbb{P}$ such that for every $\mathcal{T} \in \mathfrak{T}$ it is

$$\Big(\frac{1}{|I_{\mathsf{Top}(\mathcal{T})}|} \int \sum_{\omega\in\mathcal{T}} |h_\omega|^{r'} \Phi^\omega_{I_{\mathsf{Top}(\mathcal{T})}} \,\mathrm{d}x \Big)^{1/r'} \geqslant 2^n.$$

Remark 2.9. Our definition of Energy^1 is slightly different from the corresponding one of [2] (in particular it's somewhat relaxed) because in our arguments we won't have to resort to Bessel-type inequalities.

We must show that these quantities are well behaved in order for the machinery of time-frequency analysis to work. In particular, we ought to show that the energies can be controlled in terms of L^p norms of the functions. This is what we do next. First of all, we have the simple

Proposition 2.10. *For any collection of tri-tiles $\mathbb{P}$ and for any $\mathbf{h} \in L^{r'}(\ell^{r'})$ we have*

$$\mathrm{Energy}^3_{\mathbf{h}}(\mathbb{P}) \lesssim \|\mathbf{h}\|_{L^{r'}(\ell^{r'})}.$$

Proof. We may assume for simplicity that the collection $\mathbb{P}$ is finite, since our argument will not depend on its cardinality. Let $n \in \mathbb{Z}$ and $\mathfrak{T}$ be a collection of disjoint maximal rows and columns that realize the supremum in the definition of Energy^3, that is

$$\mathrm{Energy}^3_{\mathbf{h}}(\mathbb{P})^{r'} = 2^{r'n} \sum_{\mathcal{T}\in\mathfrak{T}} |I_{\mathsf{Top}(\mathcal{T})}|.$$

By definition we then have

$$\begin{aligned}\mathrm{Energy}^3_{\mathbf{h}}(\mathbb{P})^{r'} &\lesssim \sum_{\mathcal{T}\in\mathfrak{T}}\sum_{\omega\in\mathcal{T}}\int |h_\omega|^{r'}\Phi^\omega_{I_{\mathsf{Top}(\mathcal{T})}}\mathrm{d}x\\ &= \sum_{\omega\in\Omega}\int |h_\omega|^{r'}\sum_{\substack{\mathcal{T}\in\mathfrak{T}:\\ \mathcal{T}\ni\omega}}\Phi^\omega_{I_{\mathsf{Top}(\mathcal{T})}}\mathrm{d}x;\end{aligned}$$

but the collection $\mathfrak{T}$ is maximal with respect to inclusion and therefore for a fixed ω the intervals $I_{\mathsf{Top}(\mathcal{T})}$ are disjoint, hence by the rapid decay of the functions Φ^ω we have

$$\sum_{\substack{\mathcal{T}\in\mathfrak{T}:\\ \mathcal{T}\ni\omega}}\Phi^\omega_{I_{\mathsf{Top}(\mathcal{T})}}\lesssim 1,$$

and this concludes the proof. □

Remark 2.11. Two things should be noticed. Firstly, that in the last lines of the above proof we crucially needed the decay of the functions Φ^ω_I away from I to be controlled by $|\omega|^{-1}$ rather than by the larger $|I|$, which justifies their introduction and the subsequent definition of Size^3.
Secondly, the maximality we appealed to above means the following: given a column (or row) $\mathcal{C}$ such that

$$\frac{1}{|I_{\mathrm{Top}(\mathcal{C})}|}\int\sum_{\omega\in\Omega(\mathcal{C})}|h_\omega|^{r'}\Phi^\omega_{I_{\mathrm{Top}(\mathcal{C})}}\mathrm{d}x\geqslant 2^{r'n},$$

we can enlarge $\mathcal{C}$ by adjoining all tri-tiles $P\in\mathbb{P}$ such that $\omega(P)\in\Omega(\mathcal{C})$ and $I_P\subseteq I_{\mathrm{Top}(\mathcal{C})}$, and doing so will not change the left hand side of the inequality at all. In other words, the only information $\mathrm{Size}^3_{\mathbf{h}}$ is sensitive to is the space support of columns or rows $\mathcal{T}$ and the squares $\omega\in\Omega(\mathcal{T})$.

Next we look at a bound for Energy^j for $j=1,2$. Here our method necessarily introduces an unfortunate logarithmic loss (already encountered in Lemma 2.7; the proofs indeed rely on the same idea).

Proposition 2.12. *Let Ω be a collection of disjoint dyadic squares with $\#\Omega=N$. Then for any $\varepsilon>0$, for any collection of tri-tiles $\mathbb{P}$ and for any $f\in L^2$ we have that*

$$\mathrm{Energy}^j_f(\mathbb{P})\lesssim_\varepsilon N^\varepsilon\|f\|_{L^2}\qquad j=1,2.$$

Proof. We let $j=1$ in here, the proof for $j=2$ being identical. Let $n\in\mathbb{Z}$ and $\mathfrak{C}$ be a collection of disjoint columns that realize the supremum in the definition of $\mathrm{Energy}^1_f(\mathbb{P})$ within a factor of 2, that is

$$\mathrm{Energy}^1_f(\mathbb{P})^2\sim 2^{2n}\sum_{\mathcal{C}\in\mathfrak{C}}|I_{\mathsf{Top}(\mathcal{C})}|.$$

Then by definition of energy this implies

$$\mathrm{Energy}^1_f(\mathbb{P})^2\lesssim\sum_{\mathcal{C}\in\mathfrak{C}}\|\pi_{\omega_1(\mathcal{C})}f\|^2_{L^2(I_{\mathcal{C}})},$$

where we have abused notation by writing $\mathcal{C}$ in place of $\mathsf{Top}(\mathcal{C})$ to ease readability. We rewrite the latter quantity as

$$\int\sum_{\mathcal{C}\in\mathfrak{C}}|\pi_{\omega_1(\mathcal{C})}f(x)|^2\mathbb{1}_{I_{\mathcal{C}}}(x)\mathrm{d}x;$$

observe that we have the trivial bound

$$\Big\|\sum_{\mathcal{C}\in\mathfrak{C}}\mathbb{1}_{I_{\mathcal{C}}}\Big\|_{L^\infty}\leqslant\#\Omega=N.$$

Let then $q_0(\varepsilon) = q_0 > 2$ be such that

$$\frac{1}{2} = \frac{1}{q_0} + \varepsilon$$

and bound by Hölder's inequality

$$\int \sum_{\mathcal{C}\in\mathfrak{C}} |\pi_{\omega_1(\mathcal{C})} f(x)|^2 \mathbb{1}_{I_{\mathcal{C}}}(x) \mathrm{d}x \leqslant N^{2\varepsilon} \int \Big(\sum_{\mathcal{C}\in\mathfrak{C}} |\pi_{\omega_1(\mathcal{C})} f(x)|^{q_0} \mathbb{1}_{I_{\mathcal{C}}}(x) \Big)^{2/q_0} \mathrm{d}x.$$

Since by definition of Energy^1 the tops $\mathsf{Top}(\mathcal{C})$ are disjoint as tiles, we have that pointwise

$$\Big(\sum_{\mathcal{C}\in\mathfrak{C}} |\pi_{\omega_1(\mathcal{C})} f(x)|^{q_0} \mathbb{1}_{I_{\mathcal{C}}}(x) \Big)^{1/q_0} \leqslant \mathrm{Var}_{q_0} \mathscr{C} f(x),$$

where $\mathrm{Var}_{q_0}\mathscr{C}$ denotes the q_0-variational Carleson operator. We know from [17] that this operator is $L^p \to L^p$ bounded for $p > q_0'$, and therefore our quantity above is bounded by

$$N^{2\varepsilon} \int |\mathrm{Var}_{q_0} \mathscr{C} f|^2 \mathrm{d}x \lesssim_\varepsilon N^{2\varepsilon} \|f\|_{L^2}^2,$$

which finishes the proof. □

Remark 2.13. Lemma 2.7 and the above proposition are the reason for our choice of working with the coefficients $f(P_1), g(P_2), \mathbf{h}(P_3)$. Indeed, it is their form and precise localization (that is, the L^2 norms don't involve weights supported everywhere on $\mathbb{R}$) that allow us to introduce pointwise estimates of the relevant quantities in terms of Variational Carleson operators.

2.5. **Decomposition lemmas.** The decomposition lemma for Size^1 is well known and perhaps immediate. An identical result holds for Size^2 by replacing columns with rows.

Lemma 2.14 (Decomposition lemma for Size^1)**.** *Let $\mathbb{P}$ be a collection of tiles and let n be such that*

$$\mathrm{Size}_f^1(\mathbb{P}) \leqslant 2^{-n} \mathrm{Energy}_f^1(\mathbb{P}).$$

Then we can decompose $\mathbb{P} = \mathbb{P}_{\mathrm{low}} \sqcup \mathbb{P}_{\mathrm{high}}$ such that

$$\mathrm{Size}_f^1(\mathbb{P}_{\mathrm{low}}) \leqslant 2^{-n-1} \mathrm{Energy}_f^1(\mathbb{P})$$

and $\mathbb{P}_{\mathrm{high}}$ can be organized into a collection $\mathfrak{C}$ of mutually disjoint columns $\mathcal{C}$ such that

$$\sum_{\mathcal{C}\in\mathfrak{C}} |I_{\mathcal{C}}| \lesssim 2^{2n}.$$

Proof. The proof is well known. We select the maximal tiles P in $\mathbb{P}$ such that

$$\frac{f(P_1)}{|I_P|^{1/2}} > 2^{-n-1} \mathrm{Energy}_f^1(\mathbb{P}),$$

and start from the leftmost, higher one, denoted $P_{\max}$. Let $\mathbb{P}_{\mathrm{stock}} := \mathbb{P}$ and $\mathfrak{C} := \varnothing$ at the start and let $\mathcal{C}_1$ be the maximal column in $\mathbb{P}_{\mathrm{stock}}$ with top $P_{\max}$; then update $\mathfrak{C}$ to $\mathfrak{C} \cup \{\mathcal{C}_1\}$, and update $\mathbb{P}_{\mathrm{stock}}$ to $\mathbb{P}_{\mathrm{stock}} \backslash \bigcup_{P\in\mathcal{C}_1} \{P\}$. Repeat the process until the algorithm stops. Define $\mathbb{P}_{\mathrm{low}} := \mathbb{P}_{\mathrm{stock}}$ and $\mathbb{P}_{\mathrm{high}} := \bigcup_j \bigcup_{P\in\mathcal{C}_j} \{P\}$. Then the size property is immediate, and as for the bound on the measure of the tops notice that we have for each j

$$\frac{f(\mathsf{Top}(\mathcal{C}_j)_1)}{|I_{\mathsf{Top}(\mathcal{C}_j)}|^{1/2}} > 2^{-n-1} \mathrm{Energy}_f^1(\mathbb{P})$$

and therefore by definition of energy (and its monotonicity)

$$2^{-n-1}\mathrm{Energy}_f^1(\mathbb{P})\Big(\sum_j |I_{\mathsf{Top}(\mathcal{C}_j)}|\Big)^{1/2} \lesssim \mathrm{Energy}_f^1(\mathbb{P}_{\mathrm{high}}) \leqslant \mathrm{Energy}_f^1(\mathbb{P}),$$

which proves the claim. □

The decomposition lemma for Size^3 is entirely similar (we have replaced the constant 2 with γ in view of its application to the proof of Lemma 2.16 below).

Lemma 2.15 (Decomposition lemma for Size^3)**.** *Let $\gamma = 2^{2/r'}$. Let $\mathbb{P}$ be a collection of tiles and let n be such that*

$$\mathrm{Size}_h^3(\mathbb{P}) \leqslant \gamma^{-n}\mathrm{Energy}_h^3(\mathbb{P}).$$

Then we can decompose $\mathbb{P} = \mathbb{P}_{\mathrm{low}} \sqcup \mathbb{P}_{\mathrm{high}}$ such that

$$\mathrm{Size}_h^3(\mathbb{P}_{\mathrm{low}}) \leqslant \gamma^{-n-1}\mathrm{Energy}_h^3(\mathbb{P})$$

and $\mathbb{P}_{\mathrm{high}}$ can be organized into a collection $\mathfrak{T}$ of disjoint columns and rows $\mathcal{T}$ such that

$$\sum_{\mathcal{T}\in\mathfrak{T}} |I_{\mathcal{T}}| \lesssim \gamma^{r'n} = 2^{2n}.$$

The proof is essentially identical to the one given above for Size^1, and is thus omitted.
Finally, by applying the decomposition lemmas simultaneously and then iterating one can achieve a global decomposition of a given collection $\mathbb{P}$ with good control of the sizes of the sub-collections. In particular

Lemma 2.16 (global decomposition)**.** *Let $\mathbb{P}$ be a collection of tri-tiles. Then there exists a partition $\mathbb{P} = \bigsqcup_n (\mathbb{P}_n^{\mathrm{col}} \sqcup \mathbb{P}_n^{\mathrm{row}})$ with the properties:*

i) $\mathrm{Size}_f^1(\mathbb{P}_n^{\mathrm{col,row}}) \lesssim \min(2^{-n}\mathrm{Energy}_f^1(\mathbb{P}), \mathrm{Size}_f^1(\mathbb{P}))$,
ii) $\mathrm{Size}_g^2(\mathbb{P}_n^{\mathrm{col,row}}) \lesssim \min(2^{-n}\mathrm{Energy}_g^2(\mathbb{P}), \mathrm{Size}_g^2(\mathbb{P}))$,
iii) $\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}_n^{\mathrm{col,row}}) \lesssim \min(2^{-2n/r'}\mathrm{Energy}_{\mathbf{h}}^3(\mathbb{P}), \mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}))$,
iv) $\mathbb{P}_n^{\mathrm{col}}$ *is organized into a collection $\mathfrak{C}_n$ of disjoint columns,*
v) $\mathbb{P}_n^{\mathrm{row}}$ *is organized into a collection $\mathfrak{R}_n$ of disjoint rows,*
vi) $\sum_{\mathcal{C}\in\mathfrak{C}_n} |I_{\mathcal{C}}| \lesssim 2^{2n}$,
vii) $\sum_{\mathcal{R}\in\mathfrak{R}_n} |I_{\mathcal{R}}| \lesssim 2^{2n}$.

The collection $\mathbb{P}_n^{\mathrm{col}}$ is empty if n is such that

$$2^{-n} \gtrsim \frac{\mathrm{Size}_f^1(\mathbb{P})}{\mathrm{Energy}_f^1(\mathbb{P})} \quad \text{and} \quad 2^{-2n/r'} \gtrsim \frac{\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P})}{\mathrm{Energy}_{\mathbf{h}}^3(\mathbb{P})},$$

and similarly the collection $\mathbb{P}_n^{\mathrm{row}}$ is empty if n is such that

$$2^{-n} \gtrsim \frac{\mathrm{Size}_g^2(\mathbb{P})}{\mathrm{Energy}_g^2(\mathbb{P})} \quad \text{and} \quad 2^{-2n/r'} \gtrsim \frac{\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P})}{\mathrm{Energy}_{\mathbf{h}}^3(\mathbb{P})}.$$

Proof. Initialize $\mathbb{P}_{\mathrm{stock}} := \mathbb{P}$ and apply iteratively the decomposition Lemmas 2.14 and 2.15, in the order given by whichever of the quantities

$$\frac{\mathrm{Size}_f^1(\mathbb{P}_{\mathrm{stock}})}{\mathrm{Energy}_f^1(\mathbb{P})}, \quad \frac{\mathrm{Size}_g^2(\mathbb{P}_{\mathrm{stock}})}{\mathrm{Energy}_g^2(\mathbb{P})}, \quad \Big(\frac{\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}_{\mathrm{stock}})}{\mathrm{Energy}_{\mathbf{h}}^3(\mathbb{P})}\Big)^{r'/2}$$

is largest, sorting columns and rows into the current $\mathbb{P}^{\mathrm{col}}, \mathbb{P}^{\mathrm{row}}$ respectively and updating $\mathbb{P}_{\mathrm{stock}}$ at the end of each step to be the collection $(\mathbb{P}_{\mathrm{stock}})_{\mathrm{low}}$ resulting from the last application of a decomposition lemma. We omit the details. □

2.6. **General estimate for $\Lambda_{\mathbb{P}}$.** In this section we prove the following general estimate, which will then be the main ingredient in the proof of Proposition 2.19 in §2.7.

Lemma 2.17. *Let Ω, Λ be as above, with $\#\Omega = N$, let $\mathbb{P}$ be a collection of tri-tiles and let $\varepsilon > 0$. Let $\sigma = \frac{r-2}{r}$ and denote for shortness*

$$\begin{aligned} \mathrm{Size}_f^1(\mathbb{P}) =: \mathcal{S}_1, \qquad & \mathrm{Energy}_f^1(\mathbb{P}) =: \mathcal{E}_1, \\ \mathrm{Size}_g^2(\mathbb{P}) =: \mathcal{S}_2, \qquad & \mathrm{Energy}_g^2(\mathbb{P}) =: \mathcal{E}_2, \\ \mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}) =: \mathcal{S}_3, \qquad & \mathrm{Energy}_{\mathbf{h}}^3(\mathbb{P}) =: \mathcal{E}_3. \end{aligned}$$

Then if we let $q_0(\varepsilon) = q_0 > 2$ be given by

$$\frac{1}{2} = \frac{1}{q_0} + \frac{\varepsilon}{2}$$

we have

$$\begin{aligned} |\Lambda_{\mathbb{P}}(f,g,\mathbf{h})| \lesssim & N^{\varepsilon} \sup_{P\in\mathbb{P}} \Big[\fint_{I_P} |\mathrm{Var}_{q_0} \mathscr{C} g|^2 \Big]^{(1-\sigma)/2} \\ & \qquad \times \mathcal{S}_1^{2\sigma\theta_1} \mathcal{E}_1^{1-2\sigma\theta_1} \mathcal{S}_2^{2\sigma\theta_2} \mathcal{E}_2^{\sigma-2\sigma\theta_2} \mathcal{S}_3^{r'\sigma\theta_3} \mathcal{E}_3^{1-r'\sigma\theta_3} \\ & + N^{\varepsilon} \sup_{P\in\mathbb{P}} \Big[\fint_{I_P} |\mathrm{Var}_{q_0} \mathscr{C} f|^2 \Big]^{(1-\sigma)/2} \\ & \qquad \times \mathcal{S}_1^{2\sigma\xi_1} \mathcal{E}_1^{\sigma-2\sigma\xi_1} \mathcal{S}_2^{2\sigma\xi_2} \mathcal{E}_2^{1-2\sigma\xi_2} \mathcal{S}_3^{r'\sigma\xi_3} \mathcal{E}_3^{1-r'\sigma\xi_3} \end{aligned} \tag{7}$$

for any θ_j, ξ_j such that $\theta_1 + \theta_2 + \theta_3 = 1$ and respectively $\xi_1 + \xi_2 + \xi_3 = 1$, and

$$\begin{aligned} & 0 \leqslant \theta_1 \leqslant \min(1, (2\sigma)^{-1}), && 0 \leqslant \xi_1 \leqslant \frac{1}{2}, \\ & 0 \leqslant \theta_2 \leqslant \frac{1}{2}, && 0 \leqslant \xi_2 \leqslant \min(1, (2\sigma)^{-1}), \\ & 0 < \theta_3 \leqslant 1, && 0 < \xi_3 \leqslant 1. \end{aligned}$$

Proof. Apply the global decomposition lemma (Lemma 2.16) to the collection $\mathbb{P}$, thus obtaining a partition $\mathbb{P} = \bigsqcup_n \mathbb{P}_n^{\mathrm{col}} \sqcup \mathbb{P}_n^{\mathrm{row}}$. It suffices to consider the collections $\mathbb{P}_n^{\mathrm{col}}$ (which correspond to the first term in (7)), the proof for the collections $\mathbb{P}_n^{\mathrm{row}}$ being entirely analogous. Since $\mathbb{P}_n^{\mathrm{col}}$ is organized into a collection $\mathfrak{C}_n$ of disjoint columns, using Proposition 2.5 we can bound

$$\begin{aligned} |\Lambda_{\mathbb{P}_n^{\mathrm{col}}}(f,g,\mathbf{h})| & \leqslant \sum_{\mathcal{C}\in\mathfrak{C}_n} |\Lambda_{\mathcal{C}}(f,g,\mathbf{h})| \\ & \lesssim \sum_{\mathcal{C}\in\mathfrak{C}_n} \Big(\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 \Big)^{1/r} \mathrm{Size}_f^1(\mathcal{C}) \big[\mathrm{Size}_g^2(\mathcal{C})\big]^{\sigma} \mathrm{Size}_{\mathbf{h}}^3(\mathcal{C}) |I_{\mathcal{C}}| \\ & \lesssim \sum_{\mathcal{C}\in\mathfrak{C}_n} \Big(\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 \Big)^{1/r} \min(2^{-n}\mathcal{E}_1, \mathcal{S}_1) \big[\min(2^{-n}\mathcal{E}_2, \mathcal{S}_2)\big]^{\sigma} \\ & \qquad\qquad \times \min(2^{-2n/r'}\mathcal{E}_3, \mathcal{S}_3) |I_{\mathcal{C}}|; \end{aligned}$$

by Lemma 2.7 term $\Big(\frac{1}{|I_{\mathcal{C}}|} \sum_{P\in\mathcal{C}} g(P_2)^2 \Big)^{1/r}$ can be replaced with

$$N^{\varepsilon} \sup_{P\in\mathbb{P}} \Big[\fint_{I_P} |\mathrm{Var}_{q_0} \mathscr{C} g|^2 \Big]^{1/r},$$

which then factors out of the sum (notice $(1-\sigma)/2 = 1/r$). By definition of $\mathbb{P}_n^{\mathrm{col}}$, what remains is controlled by

$$\min(2^{-n}\mathcal{E}_1, \mathcal{S}_1) \big[\min(2^{-n}\mathcal{E}_2, \mathcal{S}_2)\big]^{\sigma} \min(2^{-2n/r'}\mathcal{E}_3, \mathcal{S}_3) 2^{2n},$$

and therefore it suffices to show that the sum over n of all these contributions is controlled by the corresponding product of sizes and energies in the first term of the right hand side of (7). This requires a tedious but easy case by case analysis. Assume that

$$\frac{\mathcal{S}_1}{\mathcal{E}_1} < \frac{\mathcal{S}_2}{\mathcal{E}_2} < \Big(\frac{\mathcal{S}_3}{\mathcal{E}_3}\Big)^{r'/2}, \tag{8}$$

the other cases being similar and thus omitted. We have

1) **case** $2^{-n} \leqslant \frac{\mathcal{S}_1}{\mathcal{E}_1} < \frac{\mathcal{S}_2}{\mathcal{E}_2} < \Big(\frac{\mathcal{S}_3}{\mathcal{E}_3}\Big)^{r'/2}$: in this case the sum we have to bound becomes

$$\sum_{n\,:\,2^{-n}\leqslant \mathcal{S}_1\mathcal{E}_1^{-1}} 2^{-n}\mathcal{E}_1 2^{-n\sigma}\mathcal{E}_2^{\sigma} 2^{-2n/r'}\mathcal{E}_3 2^{2n}$$
$$= \mathcal{E}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3 \sum_{n\,:\,2^{-n}\leqslant \mathcal{S}_1\mathcal{E}_1^{-1}} 2^{-n(\sigma+2/r'-1)},$$

and since $\sigma + 2/r' - 1 = 2\sigma$ the above evaluates to

$$\mathcal{S}_1^{2\sigma}\mathcal{E}_1^{1-2\sigma}\mathcal{E}_2^{\sigma}\mathcal{E}_3,$$

which by assumption (8) is clearly controlled by the desired

$$\mathcal{S}_1^{2\sigma\theta_1}\mathcal{E}_1^{1-2\sigma\theta_1}\mathcal{S}_2^{2\sigma\theta_2}\mathcal{E}_2^{\sigma-2\sigma\theta_2}\mathcal{S}_3^{r'\sigma\theta_3}\mathcal{E}_1^{1-r'\sigma\theta_3}.$$

2) **case** $\frac{\mathcal{S}_1}{\mathcal{E}_1} < 2^{-n} \leqslant \frac{\mathcal{S}_2}{\mathcal{E}_2} < \Big(\frac{\mathcal{S}_3}{\mathcal{E}_3}\Big)^{r'/2}$: in this case the sum becomes

$$\sum_{n\,:\,\mathcal{S}_1\mathcal{E}_1^{-1}<2^{-n}\leqslant \mathcal{S}_2\mathcal{E}_2^{-1}} \mathcal{S}_1 2^{-n\sigma}\mathcal{E}_2^{\sigma} 2^{-2n/r'}\mathcal{E}_3 2^{2n}$$
$$= \mathcal{S}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3 \sum_{n\,:\,\mathcal{S}_1\mathcal{E}_1^{-1}<2^{-n}\leqslant \mathcal{S}_2\mathcal{E}_2^{-1}} 2^{-n(\sigma+2/r'-2)},$$

and $\sigma + 2/r' - 2 = 2\sigma - 1$. Thus we have further sub-cases:

i) **subcase** $2\sigma - 1 < 0$: in this case the sum is controlled by

$$\mathcal{S}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{2\sigma-1} = \mathcal{S}_1^{2\sigma}\mathcal{E}_1^{1-2\sigma}\mathcal{E}_2^{\sigma}\mathcal{E}_3,$$

which we have already established is fine;

ii) **subcase** $2\sigma - 1 > 0$: in this case the sum is controlled by

$$\mathcal{S}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma-1} = \mathcal{E}_1\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{2\sigma\theta_1}\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{1-2\sigma\theta_1}\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma-1},$$

and since by assumption $1 - 2\sigma\theta_1 \geqslant 0$ we can further bound this by

$$\mathcal{E}_1\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{2\sigma\theta_1}\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{1-2\sigma\theta_1}\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma-1} = \mathcal{E}_1\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{2\sigma\theta_1}\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma\theta_2+2\sigma\theta_3},$$

which is clearly controlled by the desired quantity;

iii) **subcase** $2\sigma - 1 = 0$: in this case the sum is controlled by

$$\mathcal{S}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3\log\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\cdot\frac{\mathcal{E}_1}{\mathcal{S}_1}\Big) \lesssim \mathcal{S}_1\mathcal{E}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\cdot\frac{\mathcal{E}_1}{\mathcal{S}_1}\Big)^{2\sigma(\theta_2+\theta_3)}$$
$$= \mathcal{E}_1\Big(\frac{\mathcal{S}_1}{\mathcal{E}_1}\Big)^{1-2\sigma(\theta_2+\theta_3)}\mathcal{E}_2^{\sigma}\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma(\theta_2+\theta_3)}\mathcal{E}_3,$$

which is again the desired quantity since for this value of σ it is $1 - 2\sigma(\theta_2 + \theta_3) = 2\sigma\theta_1$.

3) **case** $\frac{\mathcal{S}_1}{\mathcal{E}_1} < \frac{\mathcal{S}_2}{\mathcal{E}_2} < 2^{-n} \leqslant \left(\frac{\mathcal{S}_3}{\mathcal{E}_3}\right)^{r'/2}$: in this case the sum becomes

$$\sum_{n\,:\,\mathcal{S}_2\mathcal{E}_2^{-1}<2^{-n}\leqslant\mathcal{S}_3^{r'/2}\mathcal{E}_2^{-r'/2}} \mathcal{S}_1\mathcal{S}_2^{\sigma}2^{-2n/r'}\mathcal{E}_3 2^{2n}$$
$$=\mathcal{S}_1\mathcal{S}_2^{\sigma}\mathcal{E}_3 \sum_{n\,:\,\mathcal{S}_2\mathcal{E}_2^{-1}<2^{-n}\leqslant\mathcal{S}_3^{r'/2}\mathcal{E}_2^{-r'/2}} 2^{2n/r};$$

since $2/r = 1-\sigma$, this is controlled by

$$\mathcal{S}_1\mathcal{S}_2^{\sigma}\mathcal{E}_3\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{\sigma-1} = \mathcal{S}_1\Big(\frac{\mathcal{S}_2}{\mathcal{E}_2}\Big)^{2\sigma-1}\mathcal{E}_2^{\sigma}\mathcal{E}_3,$$

which we have encountered in the previous case and is therefore fine too.

Thus the proof is concluded. □

2.7. **Proof of the main theorem.** We are now ready to prove the main theorem (Theorem 1.2). It will be obtained by interpolation between the two extreme situations, namely $r = \infty$ and r close to 2.
For the first case, we only use the Carleson operator which is bounded on all L^p spaces for $p \in (1,\infty)$ to deduce the following:

Proposition 2.18. *The bilinear operator T_Ω^∞ given by*

$$T_\Omega^\infty(f,g)(x) := \sup_{\omega\in\Omega}|\pi_\omega(f,g)(x)|$$

is bounded from $L^p \times L^q$ to L^s for all $1 < p,q < \infty$, where $1/p + 1/q = 1/s$.

Proof. As observed in Example 2, the operator T_Ω^∞ is bounded pointwise by

$$T_\Omega^\infty(f,g)(x) \leqslant \mathscr{C}f(x)\cdot\mathscr{C}g(x),$$

and the result then follows from the Carleson-Hunt theorem. □

For the second case, we will prove the following proposition, whose statement is identical to that of Theorem 1.2 except for the smaller range of p,q (namely $p,q>2$ here, and hence $s>1$ too).

Proposition 2.19. *Let $r > 2$ be fixed* [5]*. Then for all p,q,s such that*

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

and

$$2 < p,q < r, \quad 1 < s < r'/2$$

it holds that for every arbitrary finite collection Ω of disjoint squares in $\widehat{\mathbb{R}^2}$ and for every $\varepsilon > 0$ the estimate

$$\Big\|\Big(\sum_{\omega\in\Omega}|\pi_\omega(f,g)(x)|^r\Big)^{1/r}\Big\|_{L^s} \lesssim_{\varepsilon,p,q} \#\Omega^{\varepsilon}\|f\|_{L^p}\|g\|_{L^q}, \tag{9}$$

holds true for every $f\in L^p, g\in L^q$.

Theorem 1.2 follows from Proposition 2.18 and 2.19 by multilinear interpolation of vector-valued operators. More precisely, it will follow from a straightforward application of the next lemma (originating from [19]), which we state after a definition.

[5]For interpolation purposes, r should be thought of as being very close to 2.

Definition 2.20. Let $\Lambda(f,g,\mathbf{h})$ be a trilinear form and let $(\alpha_1,\alpha_2,\alpha_3;r)$ be such that $0 \leqslant \alpha_1,\alpha_2 \leqslant 1$, $\alpha_3 \leqslant 1$, $\alpha_1+\alpha_2+\alpha_3 = 1$ and $r \geqslant 1$. We say that Λ is of *generalized restricted weak type* $(\alpha_1,\alpha_2,\alpha_3;r)$ if for every measurable subsets F,G,H of $\mathbb{R}$ of finite measure there exists a subset $H' \subseteq H$, called *major subset*, such that $|H'| \sim |H|$ and for all functions $f,g,\mathbf{h}$ such that

$$|f| \leqslant \mathbb{1}_F, \quad |g| \leqslant \mathbb{1}_G, \quad \Big(\sum_k |h_k|^r\Big)^{1/r} \leqslant \mathbb{1}_{H'}$$

the inequality

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

holds true.

Lemma 2.21 ([19])**.** *Let Λ be a trilinear form of generalized restricted weak type $(\alpha_1,\alpha_2,\alpha_3;r_0)$ and $(\beta_1,\beta_2,\beta_3;r_1)$, with the property that the major subset doesn't depend on the α's or β's. Then for all θ such that $0<\theta<1$, with*

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

and

$$\frac{1}{r_\theta} = \frac{1-\theta}{r_0} + \frac{\theta}{r_1},$$

it holds that Λ is of generalized restricted weak type $(\alpha_1^\theta,\alpha_2^\theta,\alpha_3^\theta,r_\theta)$.

Proof. The lemma is a particular case of a more general interpolation lemma originating from [19] (specifically Lemma 4.3). We sketch the proof here for the reader's convenience.
We argue by complex interpolation. Let F,G,H,H',f,g,θ be given and let $\mathbf{h}$ be such that

$$\Big(\sum_k |h_k|^{r_\theta}\Big)^{1/r_\theta} \leqslant \mathbb{1}_{H'}.$$

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

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

for every k, where

$$r(z) = (1-z)\frac{r_\theta}{r_0} + z\frac{r_\theta}{r_1}.$$

When $\mathrm{Re}z = 0$ we have $|h_k^z|^{r_0} = |h_k|^{r_\theta}$, and when $\mathrm{Re}z = 1$ we have $|h_k^z|^{r_1} = |h_k|^{r_\theta}$; hence by assumption we have for $\mathrm{Re}z = 0$

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

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

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

Since the function $\Phi(z) := \Lambda(f,g,\mathbf{h}^z)$ is easily seen to be holomorphic in the open strip $S = \{z \in \mathbb{C} \text{ s.t. } 0 < \mathrm{Re}z < 1\}$, 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. □

Theorem 1.2 follows by taking $r_1 = \infty$ and r_0 sufficiently close to 2 and applying Lemma 2.21 above to the trilinear form Λ in (6). The hypotheses are verified by Propositions 2.18 and 2.19, and we thus get that for a given $r > 2$ the trilinear form Λ in (6) is of generalized restricted weak type $(\alpha_1,\alpha_2,\alpha_3;r)$ for all $1/r < \alpha_1,\alpha_2 < 1/r'$; hence the trilinear form Λ_r in (5) is of generalized restricted weak type (in the classical sense) $(\alpha_1,\alpha_2,\alpha_3)$ for the same range of α's. Finally, the strong type

estimates for T^r_Ω follow by classical multilinear interpolation.

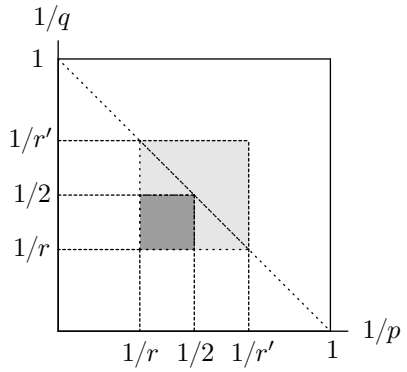

FIGURE 1. The darker square corresponds to the p, q range given by Proposition 2.19; interpolation with Proposition 2.18 extends the range to that corresponding to the additional lighter area.

We end the proof of Theorem 1.2 by proving the last remaining proposition.

Proof of Proposition 2.19. The proof follows a standard argument originating from [16] (although implicitly present in previous work).
By multilinear interpolation, it suffices to prove restricted weak type estimates, that is it suffices to prove that if F, G, H are measurable subsets of $\mathbb{R}$ of finite measure, then there exists a subset H' of H such that $|H'| \sim |H|$ and if

$$|f| \leqslant \mathbb{1}_F, \quad |g| \leqslant \mathbb{1}_G, \quad \Big(\sum_{\omega \in \Omega} |h_\omega|^{r'} \Big)^{1/r'} \leqslant \mathbb{1}_{H'}$$

then it holds that for any collection of tri-tiles $\mathbb{P}$ it is

$$|\Lambda_{\mathbb{P}}(f, g, \mathbf{h})| \lesssim_{\varepsilon,p,q} N^{\varepsilon} |F|^{1/p} |G|^{1/q} |H|^{1/s'} \tag{10}$$

for any $\varepsilon > 0$.
Given sets F, G, H and functions f, g as above, we fix two large numbers $\mathfrak{p}, \mathfrak{q} > 2$, and we define the exceptional set E to be

$$\begin{aligned} E := &\Big\{ x \in \mathbb{R} \text{ s.t. } M(|\mathscr{C} f|^{\mathfrak{p}})(x) \gtrsim \frac{|F|}{|H|} \Big\} \\ &\cup \Big\{ x \in \mathbb{R} \text{ s.t. } M(|\mathscr{C} g|^{\mathfrak{q}})(x) \gtrsim \frac{|G|}{|H|} \Big\} \\ &\cup \Big\{ x \in \mathbb{R} \text{ s.t. } M(|\mathrm{Var}_{q_0} \mathscr{C} f|^2)(x) \gtrsim \frac{|F|}{|H|} \Big\} \\ &\cup \Big\{ x \in \mathbb{R} \text{ s.t. } M(|\mathrm{Var}_{q_0} \mathscr{C} g|^2)(x) \gtrsim \frac{|G|}{|H|} \Big\}, \end{aligned}$$

where M is the dyadic Hardy-Littlewood maximal function. Define $H' := H \backslash E$; we claim that if we choose the implicit constants in the definition of E to be large enough, we have $|H'| \sim |H|$. Indeed, this follows from the $L^1 \to L^{1,\infty}$ boundedness of M and the boundedness of the relevant operators for the given exponents, for example

$$\begin{aligned} \Big| \Big\{ x \in \mathbb{R} \text{ s.t. } M(|\mathscr{C} f|^{\mathfrak{p}})(x) \gtrsim \frac{|F|}{|H|} \Big\} \Big| &\lesssim \frac{\| \mathscr{C} f^{\mathfrak{p}} \|_{L^1}}{|F|} |H| \\ &= \frac{\| \mathscr{C} f \|^{\mathfrak{p}}_{L^{\mathfrak{p}}}}{|F|} |H| \lesssim_{\mathfrak{p}} \frac{\| f \|^{\mathfrak{p}}_{L^{\mathfrak{p}}}}{|F|} |H| \ll |H|, \end{aligned}$$

where we have used the Carleson-Hunt theorem in the second to last inequality. The same holds for the other terms in the definition of E, where in particular one also has to invoke the $L^2 \to L^2$ boundedness of $\mathrm{Var}_{q_0}\mathscr{C}$ proven in [17].
Now, we partition the collection $\mathbb{P}$ into

$$\begin{aligned}\mathbb{P}_{\mathrm{small}} :=& \{P \in \mathbb{P} \text{ s.t. } I_P \not\subset E\},\\ \mathbb{P}_{\mathrm{large}} :=& \mathbb{P}\backslash\mathbb{P}_{\mathrm{small}},\end{aligned}$$

and will estimate separately the trilinear forms $\Lambda_{\mathbb{P}_{\mathrm{small}}}$ and $\Lambda_{\mathbb{P}_{\mathrm{large}}}$.
We start with $\mathbb{P}_{\mathrm{small}}$. Since $\mathfrak{p} > 2$ we have

$$\Big(\fint_{I_P} |\mathscr{C}f|^2\Big)^{1/2} \leqslant \Big(\fint_{I_P} |\mathscr{C}f|^{\mathfrak{p}}\Big)^{1/\mathfrak{p}},$$

so given $P \in \mathbb{P}_{\mathrm{small}}$ we observe that since $I_P \not\subset E$ we must have (see Proposition 2.6)

$$\mathrm{Size}_f^1(\mathbb{P}_{\mathrm{small}}) \lesssim_{\mathfrak{p}} \Big(\frac{|F|}{|H|}\Big)^{1/\mathfrak{p}}.$$

Similarly, we see that

$$\begin{aligned}\sup_{P\in\mathbb{P}_{\mathrm{small}}} \fint_{I_P} |\mathrm{Var}_{q_0}\mathscr{C}f|^2 &\lesssim_{\varepsilon} \frac{|F|}{|H|},\\ \mathrm{Size}_g^2(\mathbb{P}_{\mathrm{small}}) &\lesssim_{\mathfrak{q}} \Big(\frac{|G|}{|H|}\Big)^{1/\mathfrak{q}},\\ \sup_{P\in\mathbb{P}_{\mathrm{small}}} \fint_{I_P} |\mathrm{Var}_{q_0}\mathscr{C}g|^2 &\lesssim_{\varepsilon} \frac{|G|}{|H|};\end{aligned}$$

moreover, we have trivially

$$\mathrm{Size}_{\mathbf{h}}^3(\mathbb{P}_{\mathrm{small}}) \lesssim 1.$$

Combining this information with the general estimate in Lemma 2.17 (for which we set $\theta_j = \xi_j$ for $j = 1, 2, 3$, thus forcing the condition $0 \leqslant \theta_1, \xi_1, \theta_2, \xi_2 \leqslant 1/2$) and the energy estimates in Propositions 2.10, 2.12, we obtain after some algebra[6]

$$\begin{aligned}|\Lambda_{\mathbb{P}_{\mathrm{small}}}(f,g,\mathbf{h})| \lesssim_{\varepsilon,\mathfrak{p},\mathfrak{q}} N^{O(\varepsilon)}&\Big[\frac{|G|}{|H|}\Big]^{(1-\sigma)/2}\Big(\frac{|F|}{|H|}\Big)^{2\sigma\theta_1/\mathfrak{p}}|F|^{1/2-\sigma\theta_1}\\ &\times\Big(\frac{|G|}{|H|}\Big)^{2\sigma\theta_2/\mathfrak{q}}|G|^{\sigma/2-\sigma\theta_2}\cdot 1\cdot|H|^{1/r'-\sigma\theta_3}\\ +\, N^{O(\varepsilon)}&\Big[\frac{|F|}{|H|}\Big]^{(1-\sigma)/2}\Big(\frac{|F|}{|H|}\Big)^{2\sigma\theta_1/\mathfrak{p}}|F|^{\sigma/2-\sigma\theta_1}\\ &\times\Big(\frac{|G|}{|H|}\Big)^{2\sigma\theta_2/\mathfrak{q}}|G|^{1/2-\sigma\theta_2}\cdot 1\cdot|H|^{1/r'-\sigma\theta_3}\\ =N^{O(\varepsilon)}&|F|^{1/2-\sigma(1-2/\mathfrak{p})\theta_1}|G|^{1/2-\sigma(1-2/\mathfrak{q})\theta_2}|H|^{\sigma-\sigma\theta_3-2\sigma\theta_1/\mathfrak{p}-2\sigma\theta_2/\mathfrak{q}}.\end{aligned} \tag{11}$$

By choosing $\mathfrak{p}, \mathfrak{q}$ large enough, we obtain (10) for any choice of exponents in the stated range[7].
Now we are left with showing that (11) holds for $\Lambda_{\mathbb{P}_{\mathrm{large}}}$ as well. In order to do so, we decompose $\mathbb{P}_{\mathrm{large}}$ into $\bigsqcup_{d\in\mathbb{N}} \mathbb{P}_d$ where

$$\mathbb{P}_d := \Big\{P \in \mathbb{P}_{\mathrm{large}} \text{ s.t. } 1 + \frac{\mathrm{dist}(I_P, E^c)}{|I_P|} \sim 2^d\Big\};$$

[6]Notice $1/r' - 1/r = \sigma$.

[7]In order to prove Theorem 1.2 by interpolation we don't need the full range of exponents provided by Proposition 2.19; it suffices to take $\mathfrak{p}, \mathfrak{q}$ large but fixed for all r, so that the hypotheses of the interpolation lemma 2.21 apply, to conclude Theorem 1.2.

it then suffices to prove that the contribution of $\Lambda_{\mathbb{P}_d}$ is summable in d and the sum is bounded by (11). Let then d be fixed and observe that if $P \in \mathbb{P}_d$ then $2^{d+O(1)} I_P \not\subset E$, thus as seen above we must have[8]

$$\fint_{2^{d+O(1)} I_P} |\mathscr{C} f|^{\mathfrak{p}} \lesssim \frac{|F|}{|H|},$$

and therefore

$$\fint_{I_P} |\mathscr{C} f|^{\mathfrak{p}} \lesssim 2^d \frac{|F|}{|H|},$$

and hence by Proposition 2.6 and Hölder's inequality

$$\mathrm{Size}^1_f(\mathbb{P}_d) \lesssim_{\mathfrak{p}} 2^{d/\mathfrak{p}} \Big(\frac{|F|}{|H|}\Big)^{1/\mathfrak{p}}.$$

Similarly we have

$$\sup_{P \in \mathbb{P}_d} \fint_{I_P} |\mathrm{Var}_{q_0} \mathscr{C} f|^2 \lesssim_{\varepsilon} 2^d \frac{|F|}{|H|},$$

$$\mathrm{Size}^2_g(\mathbb{P}_d) \lesssim_{\mathfrak{q}} 2^{d/\mathfrak{q}} \Big(\frac{|G|}{|H|}\Big)^{1/\mathfrak{q}},$$

$$\sup_{P \in \mathbb{P}_d} \fint_{I_P} |\mathrm{Var}_{q_0} \mathscr{C} g|^2 \lesssim_{\varepsilon} 2^d \frac{|G|}{|H|}.$$

However, for $\mathrm{Size}^3_{\mathbf{h}}$ we now have a better estimate, namely for any $P \in \mathbb{P}_d$ it must be[9]

$$\frac{1}{|I_P|} \int \sum_{\omega \in \Omega} |h_\omega|^{r'} \Phi^{\omega}_{I_P} \lesssim 2^{-dM}$$

for a large $M > 0$ of our choice thanks to the fast decay of the functions $\Phi^{\omega}_{I_P}$, and this estimate in turn implies the bound

$$\mathrm{Size}^3_{\mathbf{h}}(\mathbb{P}_d) \lesssim 2^{-dM/r'}.$$

If we apply the general estimate of Proposition 2.17 to $\mathbb{P}_d$ as done before we then get

$$|\Lambda_{\mathbb{P}_d}(f,g,\mathbf{h})| \lesssim_{\varepsilon,\mathfrak{p},\mathfrak{q}} 2^{-M'd} N^{O(\varepsilon)} |F|^{1/2-\sigma(1-2/\mathfrak{p})\theta_1} \times |G|^{1/2-\sigma(1-2/\mathfrak{q})\theta_2} |H|^{\sigma-\sigma\theta_3-2\sigma\theta_1/\mathfrak{p}-2\sigma\theta_2/\mathfrak{q}}$$

for some large $M' > 0$ depending on M, r. As this is summable in d, the proof is concluded. □

Remark 2.22. We comment here on why, aside from the logarithmic loss, we cannot recover the same range for Λ as in Theorem 1.1 ([2]), in which the bilinear frequency projections onto the ω's are taken to be smooth. If one uses the appropriate version of the above argument in that context, the range obtained is symmetric with respect to 2, that is one gets estimates for all $r' < p, q < r$ directly, without the need to appeal to interpolation results like Lemma 2.21. The reason behind this is two-fold: firstly there's the fact that in that case all sizes satisfy $\mathrm{Size}^j(\mathbb{P}) \lesssim 1$ a priori for $j = 1, 2, 3$ (and with this information alone one already obtains the range $2 < p, q < r$); and secondly the sizes are controlled by L^1-averages instead of L^2-averages as in our case (see Proposition 2.6). Thus in the smooth case of [2] one can effectively bound $\mathrm{Size}^1_f(\mathbb{P}_{\mathrm{small}}) \lesssim \min(1, |F|/|H|)$ and similarly for Size^2, which then yields the wider range described above. Our use of the $\mathfrak{p}, \mathfrak{q}$ powers essentially amounts to a substitute for the condition $\mathrm{Size}^j(\mathbb{P}) \lesssim 1$, hence the smaller range.

[8] Technically, one needs to find a dyadic interval I that contains $2^{d+O(1)} I_P$ and has comparable length to claim so, but this is always possible.

[9] Here for convenience we are writing Φ^{ω}_I for Φ_I even when $|I| < |\omega|^{-1}$.

Finally, the full range $p, q > r'$ in [2] is obtained by a further argument involving the localization of sizes and energies; alternatively, one can obtain it by considering the formal adjoints of the bilinear operator. In the non-smooth case both approaches fail: our sizes and energies don't localize well, since we are controlling them with non-local operators; and the formal adjoints cannot be simply reduced to the original operator, so that the analysis developed in here doesn't extend to them automatically.

3. Application to bilinear multipliers

Let Ω be a collection of dyadic frequency squares, not necessarily finite and not necessarily disjoint, and let $\mathbf{a} = \{a_\omega\}_{\omega\in\Omega}$ be a sequence of complex coefficients; form then the bilinear multiplier T given by

$$T_{\mathbf{a}}(f,g)(x) := \sum_{\omega\in\Omega} a_\omega \pi_\omega(f,g)(x).$$

We are interested in finding conditions on Ω and $\{a_\omega\}_{\omega\in\Omega}$ which ensure the $L^p \times L^q \to L^s$ boundedness of T in some range of exponents p, q, s.
Consider the following situation: assume that for some $\beta \in (0,2)$ we have $\|\mathbf{a}\|_{\ell^\beta} < \infty$, and moreover the coefficients a_ω satisfy the Carleson Condition

$$\sum_{\substack{\omega'\in\Omega,\\ \omega'\subset\omega}} |a_{\omega'}|^\beta \leqslant C|a_\omega|^\beta, \qquad \forall \omega \in \Omega. \tag{12}$$

Then we argue that the bilinear multiplier $T_{\mathbf{a}}$ is bounded from $L^p \times L^q$ into L^s with $1/p + 1/q = 1/s$ for $\beta < p, q < \beta'$, where β' is replaced by ∞ if $\beta \leqslant 1$. Indeed, we partition the collection Ω as follows: let $n \in \mathbb{N}$ and define the sub-collection

$$\Omega_n := \{\omega \in \Omega \text{ s.t. } |a_\omega| \sim 2^{-n}\|\mathbf{a}\|_{\ell^\beta}\};$$

then clearly

$$\#\Omega_n \lesssim 2^{\beta n} \tag{13}$$

and moreover every collection Ω_n is the union of $O(1)$ collections of disjoint dyadic squares. This last fact is due to the Carleson Condition, since for every $\omega_0 \in \Omega_n$ it must be by definition

$$C|a_{\omega_0}|^\beta \geqslant \sum_{\substack{\omega\in\Omega_n,\\ \omega\subset\omega_0}} |a_\omega|^\beta \sim |a_{\omega_0}|^\beta \#\{\omega \in \Omega_n \text{ s.t. } \omega \subset \omega_0\};$$

thus if we do a generational decomposition of Ω_n (starting from the collection of maximal elements with respect to inclusion and so on), we will encounter at most $O(1)$ generations, which proves the claim.
Assume henceforth for the sake of clarity that for each n the collection Ω_n consists of disjoint dyadic squares only. If we take $r \in (2, \beta')$ we can bound

$$\begin{aligned}\Big|\sum_{\omega\in\Omega} a_\omega \pi_\omega(f,g)(x)\Big| &\leqslant \sum_{n\in\mathbb{N}} \Big(\sum_{\omega\in\Omega_n} |a_\omega|^{r'}\Big)^{1/r'} \Big(\sum_{\omega\in\Omega_n} |\pi_\omega(f,g)(x)|^r\Big)^{1/r} \\ &\sim \sum_{n\in\mathbb{N}} 2^{-n}\|\mathbf{a}\|_{\ell^\beta} \#\Omega_n^{1/r'} T^r_{\Omega_n}(f,g)(x).\end{aligned}$$

By Theorem 1.2 and triangle inequality we then have that

$$\|T_{\mathbf{a}}(f,g)\|_{L^s} \lesssim_{\varepsilon,p,q} \|\mathbf{a}\|_{\ell^\beta} \|f\|_{L^p}\|g\|_{L^q} \sum_{n\in\mathbb{N}} 2^{-n} \#\Omega_n^{1/r'+\varepsilon},$$

but by (13) the sum is bounded by

$$\sum_{n\in\mathbb{N}} 2^{-n} 2^{\beta(1/r'+\varepsilon)n} \lesssim_\varepsilon 1$$

for a sufficiently small $\varepsilon > 0$, thanks to our choice of r.

Remark 3.1. The Carleson condition (12) is introduced to enforce the fact that the collections Ω_n are made of essentially disjoint squares, and in particular they can be decomposed into at most $O(1)$ collections of disjoint squares. But actually, if we had that for some $\delta < 1$ each Ω_n can be decomposed into at most $O(\#\Omega_n^\delta)$ collections of disjoint squares, we could still bound the multiplier in a (smaller) range.

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] 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

[5] 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

[6] J. Bourgain, *On the behavior of the constant in the Littlewood-Paley inequality*, Geometric aspects of functional analysis (1987–88), Lecture Notes in Math., vol. 1376, Springer, Berlin, 1989, pp. 202–208. MR 1008724

[7] Lennart Carleson, *On convergence and growth of partial sums of Fourier series*, Acta Math. **116** (1966), 135–157. MR 0199631

[8] ———, *On the Littlewood-Paley theorem*, Inst. Mittag-Leffler report (1967).

[9] Antonio Córdoba, *Some remarks on the Littlewood-Paley theory*, Rend. Circ. Mat. Palermo (2) (1981), no. suppl. 1, 75–80. MR 639467

[10] Michael Cowling and Terence Tao, *Some light on Littlewood-Paley theory*, Math. Ann. **321** (2001), no. 4, 885–888. MR 1872532

[11] Richard A. Hunt, *On the convergence of Fourier series*, Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967), Southern Illinois Univ. Press, Carbondale, Ill., 1968, pp. 235–255. MR 0238019

[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

[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, 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

[17] 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

[18] José L. Rubio de Francia, *A Littlewood-Paley inequality for arbitrary intervals*, Rev. Mat. Iberoamericana **1** (1985), no. 2, 1–14. MR 850681

[19] 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

[20] Per Sjölin, *A note on Littlewood-Paley decompositions with arbitrary intervals*, J. Approx. Theory **48** (1986), no. 3, 328–334. MR 864755

[21] 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

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