ANALYSIS & PDE

Volume 9 No. 8 2016

CRISTINA BENEA AND CAMIL MUSCALU

MULTIPLE VECTOR-VALUED INEQUALITIES VIA THE HELICOIDAL METHOD

msp

ANALYSIS AND PDE
Vol. 9, No. 8, 2016
dx.doi.org/10.2140/apde.2016.9.1931

MULTIPLE VECTOR-VALUED INEQUALITIES VIA THE HELICOIDAL METHOD

CRISTINA BENEA AND CAMIL MUSCALU

We develop a new method of proving vector-valued estimates in harmonic analysis, which we call "the helicoidal method". As a consequence of it, we are able to give affirmative answers to several questions that have been circulating for some time. In particular, we show that the tensor product $\mathrm{BHT} \otimes \Pi$ between the bilinear Hilbert transform BHT and a paraproduct Π satisfies the same L^p estimates as the BHT itself, solving completely a problem introduced by Muscalu et al. (*Acta Math.* **193**:2 (2004), 269–296). Then, we prove that for "locally L^2 exponents" the corresponding vector-valued $\overrightarrow{\mathrm{BHT}}$ satisfies (again) the same L^p estimates as the BHT itself. Before the present work there was not even a single example of such exponents.

Finally, we prove a biparameter Leibniz rule in mixed norm L^p spaces, answering a question of Kenig in nonlinear dispersive PDE.

1. Introduction

Vector-valued estimates for classical Calderón–Zygmund operators are known from the work of Burkholder [1983], Benedek, Calderón and Panzone [Benedek et al. 1962], Rubio de Francia, Ruiz and Torrea [Rubio de Francia et al. 1986], to mention a few. A customary way of proving such vector-valued estimates is through weighted norm inequalities and extrapolation, as explained in [García-Cuerva and Rubio de Francia 1985]. Initially, the vector-valued approach unified the existing theory for maximal operators, square functions, and singular integrals. Later on, the setting was generalized to Banach spaces which have the *unconditional martingale difference* property, and it was shown by Bourgain [1986] that this is in fact a necessary condition for this theory.

For bilinear operators, however, the theory is far from being fully understood, even in the scalar case. In this paper, we study vector-valued estimates for the bilinear Hilbert transform and for paraproducts. Our initial motivation was an AKNS system-related problem, which can be reduced to understanding a Rubio de Francia operator for iterated Fourier integrals. Because of the specific nature of this question, our general approach is concrete, rather than abstract. As much as possible, the present article aims to be self-contained.

Central to time-frequency analysis is the bilinear Hilbert transform operator, defined by

$$\mathrm{BHT}(f,g)(x) = \mathrm{p.v.} \int_{\mathbb{R}} f(x-t)g(x+t)\,\frac{dt}{t}.$$

Muscalu is a member of the Simion Stoilow Institute of Mathematics of the Romanian Academy.

MSC2010: 42A45, 42B15, 42B25, 42B37.

Keywords: vector-valued estimates for singular and multilinear operators, tensor products in mixed norms, Leibniz rule, AKNS systems.

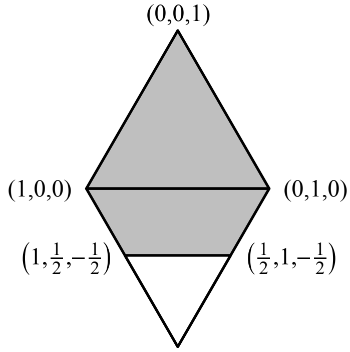

Figure 1. Range for BHT operator.

This operator was first introduced by Calderón, in connection with his work on the Cauchy integral on Lipschitz curves. L^p estimates for BHT were proved nearly thirty years later, by M. Lacey and C. Thiele, without establishing the optimality of the range.

Theorem 1 [Lacey and Thiele 1999]. BHT *is a bounded bilinear operator from* $L^p \times L^q$ *into* L^s *for any* $1 < p, q \le \infty,\ 0 < s < \infty$, *satisfying* $\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$ *and* $\frac{2}{3} < s < \infty$.

The range of the operator Range(BHT) consists of the set of triples (p, q, s) satisfying the conditions above. The question that remains open is whether the bilinear Hilbert transform is bounded also for $s \in \left(\frac{1}{2}, \frac{2}{3}\right]$. The Hölder-type condition $\frac{1}{p}+\frac{1}{q}=\frac{1}{s}$ reflects the scaling invariance of the operator, and it can be reformulated as $\frac{1}{p}+\frac{1}{q}+\frac{1}{s'}=1$, where s' is the conjugate exponent of s. Thus $(p, q, s) \in$ Range(BHT) if $\left(\frac{1}{p}, \frac{1}{q}, \frac{1}{s'}\right)$ lies in the plane $\{(x, y, z) \in \mathbb{R}^3 \mid x + y + z = 1\}$, and is contained inside the convex hull of the points

$$(0, 0, 1), \quad (1, 0, 0), \quad \left(1, \tfrac{1}{2}, -\tfrac{1}{2}\right), \quad \left(\tfrac{1}{2}, 1, -\tfrac{1}{2}\right), \quad (0, 1, 0)$$

(see Figure 1). Regarded as a bilinear multiplier operator, BHT becomes equivalent to

$$(f, g) \mapsto \int_{\xi<\eta} \hat{f}(\xi)\, \hat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi\, d\eta. \tag{1}$$

The method of the proof, which breaks down when $\frac{1}{p} + \frac{1}{q} \ge \frac{3}{2}$, consists of approximating BHT by a *model operator* obtained through a Whitney decomposition of the frequency region $\{\xi < \eta\}$. In essence, this model operator is a superposition of "almost orthogonal" objects of a lower complexity, called *discretized paraproducts*.

Paraproducts play an important role on their own, especially in the analysis of PDE. A *paraproduct* is an expression of the form

$$(f, g) \mapsto \int_{\mathbb{R}} \int_{\mathbb{R}} f(x-t)\, g(x-s)\, k(s, t)\, ds\, dt, \tag{2}$$

where $k(s,t)$ is a Calderón–Zygmund kernel in the plane $\mathbb{R}^2$. Alternatively, a paraproduct can be regarded as a bilinear multiplier operator

$$(f,g) \mapsto \int_{\mathbb{R}^2} m(\xi,\eta)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi i x(\xi+\eta)}\,d\xi\,d\eta,$$

where m is a classical Marcinkiewicz–Mikhlin–Hörmander multiplier in two variables, sufficiently smooth away from the origin. The singularity of the multiplier m consists of one point: $(\xi,\eta)=(0,0)$. On the other hand, we can see from (1) that the BHT multiplier is singular along the line $\xi=\eta$.

We have the following result on paraproducts:

Theorem 2 [Meyer and Coifman 1997]. *Any bilinear multiplier operator associated to a symbol $m(\xi,\eta)$ satisfying $|\partial^\alpha m(\xi,\eta)| \lesssim |(\xi,\eta)|^{-\alpha}$ for sufficiently many multi-indices α, maps $L^p(\mathbb{R})\times L^q(\mathbb{R})$ into $L^s(\mathbb{R})$ provided that $1<p,q\le\infty$, $\frac{1}{2}<s<\infty$, and $\frac{1}{p}+\frac{1}{q}=\frac{1}{s}$.*

Following the presentation in [Muscalu and Schlag 2013], any bilinear operator of this form can be essentially written as a finite sum of paraproducts of the form

$$(f,g)\mapsto \sum_k \big((f*\psi_k)\cdot(g*\psi_k)\big)*\varphi_k(x)=\sum_k P_k(Q_kf\cdot Q_kg), \tag{I}$$

$$(f,g)\mapsto \sum_k \big((f*\varphi_k)\cdot(g*\psi_k)\big)*\psi_k(x)=\sum_k Q_k(P_kf\cdot Q_kg), \tag{II}$$

$$(f,g)\mapsto \sum_k \big((f*\psi_k)\cdot(g*\varphi_k)\big)*\psi_k(x)=\sum_k Q_k(Q_kf\cdot P_kg). \tag{III}$$

From now on, a *paraproduct* will designate any of the expressions (I), (II) or (III), and will be denoted by $\Pi(f,g)$. Here $\psi_k(x)=2^k\psi(2^kx)$, $\varphi_k(x)=2^k\varphi(2^kx)$, $\hat{\varphi}(\xi)\equiv 1$ on $\left[-\frac{1}{2},\frac{1}{2}\right]$ and is supported on $[-1,1]$ and $\hat{\psi}(\xi)=\hat{\varphi}(\xi/2)-\hat{\varphi}(\xi)$. The $\{Q_k\}_k$ represent Littlewood–Paley projections onto the frequency $|\xi|\sim 2^k$, while $\{P_k\}_k$ are convolution operators associated with dyadic dilations of a nice bump function of integral 1.

A classical application of Theorem 2 is the Leibniz rule

$$\|D^\alpha(f\cdot g)\|_s \lesssim \|D^\alpha f\|_{p_1}\|g\|_{q_1}+\|f\|_{p_2}\|D^\alpha g\|_{q_2}, \tag{3}$$

which holds for any $\alpha>0$, as long as $\frac{1}{p_i}+\frac{1}{q_i}=\frac{1}{s}$, $1<p_i,q_i\le\infty$, and $1/(1+\alpha)<s<\infty$. In particular, if $s\ge 1$, which is the case in most applications, the Leibniz rule holds for any $\alpha>0$.

For functions on $\mathbb{R}^2$, with (fractional) partial derivatives in both variables, a corresponding Leibniz rule is

$$\begin{aligned}\big\|D_1^\alpha D_2^\beta(f\cdot g)\big\|_s &\\ \lesssim \|D_1^\alpha D_2^\beta f\|_{p_1}\|g\|_{q_1}+\|f\|_{p_2}\|D_1^\alpha D_2^\beta g\|_{q_2}&+\|D_1^\alpha f\|_{p_3}\|D_2^\beta g\|_{q_3}+\|D_2^\beta f\|_{p_4}\|D_1^\alpha g\|_{q_4}.\end{aligned} \tag{4}$$

The proof of the above inequality relies on discrete biparameter paraproducts $\Pi\otimes\Pi$, which are expressions of the form

$$\sum_{k,l}\big((f*(\varphi_k\otimes\psi_l))\cdot(g*(\psi_k\otimes\varphi_l))\big)*\psi_k\otimes\psi_l(x,y). \tag{5}$$

Muscalu, Pipher, Thiele, and Tao proved the following theorem:

Theorem 3 [Muscalu et al. 2004a]. $\Pi \otimes \Pi$ *is a bounded operator from* $L^p(\mathbb{R}^2) \times L^q(\mathbb{R}^2)$ *into* $L^s(\mathbb{R}^2)$ *provided that* $1 < p, q \le \infty$, $\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$, *and* $0 < s < \infty$.

This further implies that (4) is true whenever

$$\frac{1}{p_i} + \frac{1}{q_i} = \frac{1}{s}, \quad 1 < p_i, q_i \le \infty, \quad \text{and} \quad \max\Big(\frac{1}{1+\alpha}, \frac{1}{1+\beta}\Big) < r < \infty.$$

If $r \ge 1$ the last condition is redundant, so (4) holds for any $\alpha, \beta > 0$.

Related to this, Carlos Kenig asked the following question, which has been circulating for some time:

Question 1. *Assuming that* $1 \le s_1, s_2 < \infty$, *and* $\alpha, \beta > 0$, *is there a Leibniz rule for mixed norm* L^p *spaces of the form*

$$\big\| D_1^\alpha D_2^\beta (f \cdot g) \big\|_{L_x^{s_1} L_y^{s_2}} \lesssim \| D_1^\alpha D_2^\beta f \|_{L_x^{p_1} L_y^{p_2}} \| g \|_{L_x^{q_1} L_y^{q_2}} + \| f \|_{L_x^{p_3} L_y^{p_4}} \| D_1^\alpha D_2^\beta g \|_{L_x^{q_3} L_y^{q_4}} + \| D_1^\alpha f \|_{L_x^{p_5} L_y^{p_6}} \| D_2^\beta g \|_{L_x^{q_5} L_y^{q_6}} + \| D_2^\beta f \|_{L_x^{p_7} L_y^{p_8}} \| D_1^\alpha g \|_{L_x^{q_7} L_y^{q_8}} ?$$

Here the mixed norms are defined by

$$\| f \|_{L_x^p L_y^q} := \big\| \| f \|_{L_y^q} \big\|_{L_x^p} := \bigg(\int_{\mathbb{R}} \Big(\int_{\mathbb{R}} |f(x,y)|^q \, dy \Big)^{\frac{p}{q}} dx \bigg)^{\frac{1}{p}}. \tag{6}$$

A result of a similar type appeared in [Kenig et al. 1993], as an important tool in establishing local well-posedness for the generalized Korteweg–de Vries equation. This is a dispersive, nonlinear equation given by

$$\begin{cases} \dfrac{\partial u}{\partial t} + \dfrac{\partial^3 u}{\partial x^3} + u^k \dfrac{\partial u}{\partial x} = 0, & t, x \in \mathbb{R},\ k \in \mathbb{Z}^+, \\ u(x,0) = u_0(x). \end{cases} \tag{7}$$

In order to prove existence, the authors use the contraction principle, but to be able to do so, they need to construct a suitable Banach space. The norm of the Banach space involves mixed L^p norms of fractional derivatives in the first variable D_1^α, and the Leibniz rule employed in this paper is

$$\big\| D_1^\alpha (f \cdot g) - f \cdot D_1^\alpha g - D_1^\alpha f \cdot g \big\|_{L_x^p L_t^q} \lesssim C \| D_1^{\alpha_1} f \|_{L_x^{p_1} L_t^{q_1}} \| D_1^{\alpha_2} g \|_{L_x^{p_2} L_t^{q_2}}. \tag{8}$$

Here $\alpha \in (0,1)$, $\alpha_1 + \alpha_2 = \alpha$ and $\frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{p}$, $\frac{1}{q_1} + \frac{1}{q_2} = \frac{1}{q}$. Also, $p, p_1, p_2, q, q_1, q_2 \in (1, \infty)$, but one can allow $q_1 = \infty$ if $\alpha_1 = 0$.

The fractional derivatives appear as a consequence of the smoothness requirement on the initial data: u_0 is assumed to be in some Sobolev space $H^\alpha(\mathbb{R})$, where α depends on the value of k in (7).

Question 1 is an extension of (8), and we managed to provide an answer by proving estimates for $\Pi \otimes \Pi$ in L^p spaces with mixed norms.

Biparameter bilinear operators were first studied in [Journé 1985], where he introduced a new way of generalizing Calderón–Zygmund operators on product spaces. More exactly, in that work he proved that "bicommutators of Calderón–Coifman-type" are bounded, which translates to "$\Pi \otimes \Pi$ maps $L^2(\mathbb{R}^2) \times L^\infty(\mathbb{R}^2)$ into $L^2(\mathbb{R}^2)$". The full range of estimates for $\Pi \otimes \Pi$ was established in [Muscalu et al. 2004a], where was also noticed that $\mathrm{BHT} \otimes \mathrm{BHT}$ does not satisfy any L^p estimates. What remained undecided for some time was the following question:

Question 2. *Does the tensor product* $\mathrm{BHT} \otimes \Pi$ *satisfy any* L^p *estimates? Would it be possible to prove it satisfies the same estimates as the* BHT *itself?*

Some significant progress in answering this question was made by Silva [2014]. It was showed that $\mathrm{BHT} \otimes \Pi$ maps $L^p \times L^q$ into L^s under the constraints that $\frac{1}{p} + \frac{2}{q} < 2$ and $\frac{1}{q} + \frac{2}{p} < 2$. Our helicoidal method allows us to remove these restrictions, proving in this way that $\mathrm{BHT} \otimes \Pi$ satisfies indeed the same L^p estimates as BHT.

As it turned out, the study of Question 1 and Question 2 is related to proving (sometimes multiple) vector-valued inequalities for Π and BHT. Let $\vec{r} = (r_1, r_2, r)$ be a tuple so that $1 < r_1, r_2 \le \infty,\ 1 \le r < \infty$ and $\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r}$. We say that an inequality of the type

$$\Big\| \Big(\sum_k \big|\mathrm{BHT}(f_k, g_k)\big|^r \Big)^{\frac{1}{r}} \Big\|_s \lesssim \Big\| \Big(\sum_k |f_k|^{r_1} \Big)^{r_1} \Big\|_p \Big\| \Big(\sum_k |g_k|^{r_2} \Big)^{r_2} \Big\|_q \tag{9}$$

represents L^p estimates for vector-valued BHT, corresponding to the exponent $\vec{r}$; in short, we have L^p estimates for $\overrightarrow{\mathrm{BHT}}_{\vec{r}}$.

Some L^p estimates for vector-valued BHT have been proved recently by Silva [2014], provided $r \in \left(\frac{4}{3}, 4\right)$. UMD-valued extensions for the quartile operator (the Fourier–Walsh analogue of BHT) were studied by Hytönen, Lacey and Parissis [Hytönen et al. 2013]. Their results, transferred to the L^p setting, hold under the same constraint that $r \in \left(\frac{4}{3}, 4\right)$. Moreover, through this method it is impossible to obtain vector-valued extensions when L^1 or L^∞ spaces are involved, as these are not UMD spaces. A similar abstract approach was taken in [Di Plinio and Ou 2015], where Banach-valued estimates for paraproducts were proved.

In spite of these results, some important questions remained unsettled:

Question 3. *Are there any exponents* $\vec{r}$ *as before for which the corresponding vector-valued* $\overrightarrow{\mathrm{BHT}}_{\vec{r}}$ *satisfy the same* L^p *estimates as the* BHT *itself?*

As the question suggests, until the present work, there was not even a single example of such an exponent. We show that whenever $\vec{r}$ is in the "local ℓ^2 range" $\left(\text{that is, } 0 \le \frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'} \le \frac{1}{2}\right)$, $\overrightarrow{\mathrm{BHT}}_{\vec{r}}$ satisfies the same L^p estimates as the BHT operator. Moreover, whenever $2 \le p, q \le \infty$, we show L^p estimates exist for any exponent $\vec{r} = (r_1, r_2, r)$.

To summarize, the main task of the present work is to give affirmative answers to Question 1, Question 2, and Question 3 described above. In what follows, we will present our main results, sometimes in a more general setting.

Theorem 4. *For any* $\alpha, \beta > 0$,

$$\begin{aligned} \| D_1^\alpha D_2^\beta (f \cdot g) \|_{L_x^{s_1} L_y^{s_2}} \lesssim \| D_1^\alpha D_2^\beta f \|_{L_x^{p_1} L_y^{p_2}} \| g \|_{L_x^{q_1} L_y^{q_2}} + \| f \|_{L_x^{p_3} L_y^{p_4}} \| D_1^\alpha D_2^\beta g \|_{L_x^{q_3} L_y^{q_4}} \\ + \| D_1^\alpha f \|_{L_x^{p_5} L_y^{p_6}} \| D_2^\beta g \|_{L_x^{q_5} L_y^{q_6}} + \| D_2^\beta f \|_{L_x^{p_7} L_y^{p_8}} \| D_1^\alpha g \|_{L_x^{q_7} L_y^{q_8}} \end{aligned}$$

whenever $1 < p_j, q_j \le \infty,\ \frac{1}{2} < s_1 < \infty,\ 1 \le s_2 < \infty$, *with* $\frac{1}{1+\alpha} < s_1 < \infty$, *and the indices satisfy the natural Hölder-type conditions.*

This answers Question 1 in the affirmative. Of course, one may wonder if Theorem 4 holds in arbitrary dimensions. As the careful reader will notice, our methods allow for such a generalization, with the outer-most Lebesgue exponent possibly less than 1, if all the indices p_i, q_i involved are *strictly* between 1 and ∞. However, in applications L^∞ norms appear, so it will be of interest to have a more general theorem for $1 < p_i, q_i \le \infty$. Although we cannot obtain this result in this paper due to some delicate technical issues, we plan to return to this problem sometime in the future.

An n-dimensional version of a Leibniz rule was presented in [Torres and Ward 2015] for indices that are again strictly between 1 and ∞:

$$\|D_2^\beta(f\cdot g)\|_{L_x^{s_1}L_y^{s_2}(\mathbb{R}\times\mathbb{R}^n)} \lesssim \|D_2^\beta f\|_{L_x^{p_1}L_y^{p_2}(\mathbb{R}\times\mathbb{R}^n)}\|g\|_{L_x^{q_1}L_y^{q_2}(\mathbb{R}\times\mathbb{R}^n)} + \|f\|_{L_x^{p_1}L_y^{p_2}(\mathbb{R}\times\mathbb{R}^n)}\|D_2^\beta g\|_{L_x^{q_1}L_y^{q_2}(\mathbb{R}\times\mathbb{R}^n)}.$$

This can be regarded as an n-dimensional generalization of (8), and it is simpler than our variant of the Leibniz rule because it doesn't require a multiparameter analysis.

Our Theorem 4 is a consequence, modulo technical but "classical" complications, of the following result:

Theorem 5 (mixed norm estimates for paraproducts on the bidisc). *Let* $1 < p_j, q_j \le \infty$, $\frac{1}{2} < s_1 < \infty$, $1 \le s_2 < \infty$, *so that* $\frac{1}{p_j} + \frac{1}{q_j} = \frac{1}{s_j}$, $1 \le j \le 2$. *Then*

$$\big\|\Pi\otimes\Pi(f,g)\big\|_{L_x^{s_1}L_y^{s_2}} \lesssim \|f\|_{L_x^{p_1}L_y^{p_2}}\|g\|_{L_x^{q_1}L_y^{q_2}}.$$

The above theorem provides L^p estimates for $\Pi\otimes\Pi$ in mixed norm L^p spaces. Through our methods, we can also recover the results from [Muscalu et al. 2006a], stating that $\Pi\otimes\cdots\otimes\Pi$ maps $L^p(\mathbb{R}^n)\times L^q(\mathbb{R}^n)$ into $L^s(\mathbb{R}^n)$ whenever $1 < p, q \le \infty$, $\frac{1}{2} < s < \infty$ and $\frac{1}{p}+\frac{1}{q}=\frac{1}{s}$. Moreover, we answer Question 2 by proving that $\mathrm{BHT}\otimes\Pi$ and $\mathrm{BHT}\otimes\Pi^{\otimes n}$ satisfy the same L^p estimates as BHT:

Theorem 6. *For any* p, q, r *with* $\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$, *with* $1 < p, q \le \infty$ *and* $\frac{2}{3} < r < \infty$,

$$\big\|\mathrm{BHT}\otimes\Pi\otimes\cdots\otimes\Pi(f,g)\big\|_{L^r(\mathbb{R}^{n+1})} \lesssim \|f\|_{L^p(\mathbb{R}^{n+1})}\|g\|_{L^q(\mathbb{R}^{n+1})}.$$

The same is true for $\Pi\otimes\cdots\otimes\Pi\otimes\mathrm{BHT}\otimes\Pi\otimes\cdots\otimes\Pi$.

For $n \ge 2$, no such results were known previously, and furthermore, a new approach was necessary for $n \ge 3$. This will be explained later in part (3) of the Remark on page 1939.

Some mixed norm L^p estimates for $\Pi^{\otimes d_1}\otimes\mathrm{BHT}\otimes\Pi^{\otimes d_2}$ can also be proved (see Section 5.1). For $\Pi\otimes\mathrm{BHT}$, they are similar to [Di Plinio and Ou 2015] in the case $n = 1$. We recently learned that in [loc. cit.] mixed norm estimates for $\Pi\otimes\Pi$, close to our Theorem 5, are also obtained.

In proving the results mentioned above, multiple vector-valued extensions for BHT and Π play a very important role. Given a totally σ-finite measure space $(\mathcal{W}, \Sigma, \mu)$, and $f, g : \mathbb{R}\times\mathcal{W}\to\mathbb{C}$, we define

$$\mathrm{BHT}(f,g)(x,w) := \text{p.v.}\int_{\mathbb{R}} f(x-t,w)g(x+t,w)\,\frac{dt}{t}.$$

Note that for a fixed value $w\in\mathcal{W}$, we have $\mathrm{BHT}(f,g)(x,w) = \mathrm{BHT}(f_w,g_w)(x)$, where $f_w(x) = f(x,w)$.

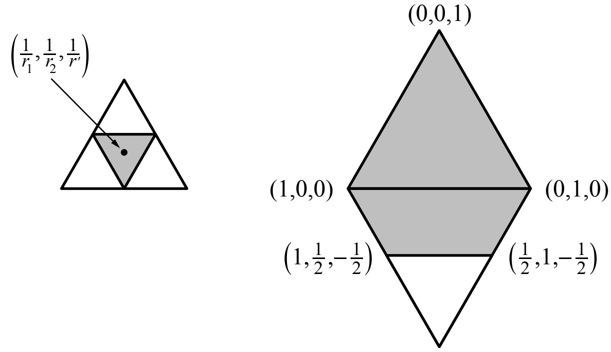

Figure 2. Range for vector-valued BHT when $\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'} \le \frac{1}{2}$.

Theorem 7. *For any triple* (r_1, r_2, r) *with* $1 < r_1, r_2 \le \infty,\ 1 \le r < \infty$ *and so that* $\frac{1}{r_1} + \frac{1}{r_2} = \frac{1}{r}$, *there exists a nonempty set* $\mathcal{D}_{r_1,r_2,r}$ *of triples* (p, q, s) *satisfying* $\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$ *for which*

$$\mathrm{BHT} : L^p(\mathbb{R}; L^{r_1}(\mathcal{W}, \mu)) \times L^q(\mathbb{R}; L^{r_2}(\mathcal{W}, \mu)) \to L^s(\mathbb{R}; L^r(\mathcal{W}, \mu)).$$

This means that there exists a constant C so that

$$\big\| \|\mathrm{BHT}(f, g)\|_{L^r(\mathcal{W},\mu)} \big\|_{L^s(\mathbb{R})} \le C \big\| \|f\|_{L^{r_1}(\mathcal{W},\mu)} \big\|_{L^p(\mathbb{R})} \big\| \|g\|_{L^{r_2}(\mathcal{W},\mu)} \big\|_{L^q(\mathbb{R})}.$$

Depending on the values of r_1, r_2, r', *we can give an explicit characterization of* $\mathcal{D}_{r_1,r_2,r}$, *as follows*:

(i) *If* $\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'} \le \frac{1}{2}$, *then* $\mathcal{D}_{r_1,r_2,r} = \mathrm{Range(BHT)}$.

(ii) *If* $\frac{1}{r_2}, \frac{1}{r'} \le \frac{1}{2}$ *and* $\frac{1}{r_1} > \frac{1}{2}$, *then* $\mathcal{D}_{r_1,r_2,r}$ *corresponds to the tuples* $(p, q, s) \in \mathrm{Range(BHT)}$ *for which* $0 \le \frac{1}{q} < \frac{3}{2} - \frac{1}{r_1}$.

(iii) *If* $\frac{1}{r_1}, \frac{1}{r'} \le \frac{1}{2}$ *and* $\frac{1}{r_2} > \frac{1}{2}$, *then the range of exponents is similar to the one in* (ii), *with the roles of* r_1 *and* r_2 *interchanged. That is,* $\mathcal{D}_{r_1,r_2,r}$ *consists of tuples* $(p, q, s) \in \mathrm{Range(BHT)}$ *for which* $0 \le \frac{1}{p} < \frac{3}{2} - \frac{1}{r_2}$.

(iv) *If* $\frac{1}{r_1}, \frac{1}{r_2} \le \frac{1}{2}$ *and* $\frac{1}{r'} > \frac{1}{2}$, *then* $\mathcal{D}_{r_1,r_2,r}$ *corresponds to the tuples* $(p, q, s) \in \mathrm{Range(BHT)}$ *for which* $0 \le \frac{1}{p}, \frac{1}{q} < \frac{1}{2} + \frac{1}{r}$ *and* $-\frac{1}{r} < \frac{1}{s'} < 1$.

See Figures 2–4 for the ranges of BHT in the cases above.

We emphasize that whenever (p, q, s) are such that $0 \le \frac{1}{p}, \frac{1}{q} \le \frac{1}{2}$ (and consequently $1 \le s < \infty$), vector-valued estimates exist for any tuple (r_1, r_2, r). These are the first examples of tuples (p, q, s) which allow for any $\overrightarrow{\mathrm{BHT}}_{\vec{r}}$ extension.

Theorem 7 can be further generalized to multiple vector-valued inequalities. For an n-tuple $P = (p_1, \dots, p_n)$, the mixed L^P norm on the product space

$$(\mathcal{W}, \Sigma, \mu) = \left(\prod_{j=1}^{n} \mathcal{W}_j, \prod_{j=1}^{n} \Sigma_j, \prod_{j=1}^{n} \mu_j \right)$$

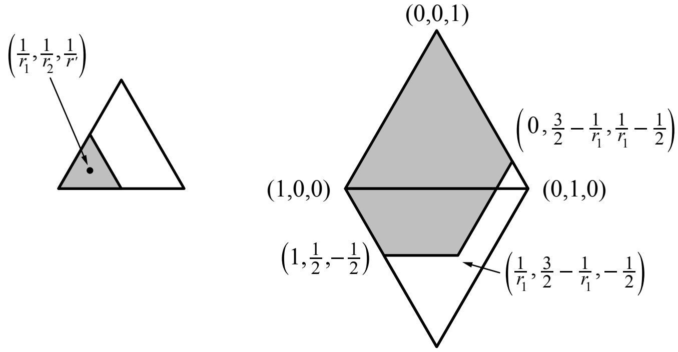

Figure 3. Range for vector-valued BHT when $\frac{1}{r_1} > \frac{1}{2}$.

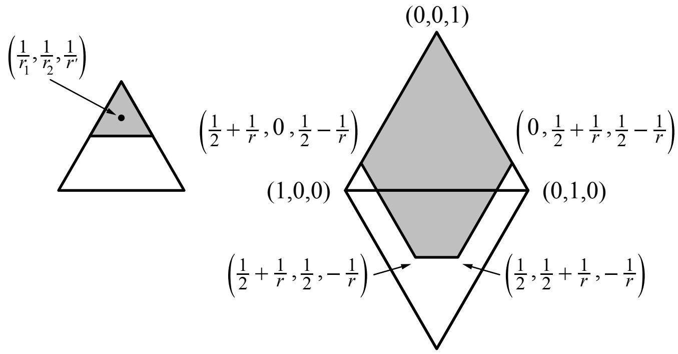

Figure 4. Range for vector-valued BHT when $\frac{1}{r'} > \frac{1}{2}$.

is defined as

$$\|f\|_P := \left(\int_{\mathcal{W}_1} \cdots \left(\int_{\mathcal{W}_n} \left|f(w_1,\dots,w_n)\right|^{p_n} d\mu_n(w_n)\right)^{\frac{p_{n-1}}{p_n}} \cdots d\mu_1(w_1)\right)^{\frac{1}{p_1}}.$$

Consider the tuples $R_1 = (r_1^1,\dots,r_1^n)$, $R_2 = (r_2^1,\dots,r_2^n)$ and $R = (r^1,\dots,r^n)$ satisfying for every $1 \le j \le n$,

$$1 < r_1^j, r_2^j \le \infty, \quad 1 \le r^j < \infty, \quad \frac{1}{r_1^j} + \frac{1}{r_2^j} = \frac{1}{r^j}$$

(from now on, this will be written as $1 < R_1, R_2 \le \infty$, $1 \le R < \infty$, and $\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{R}$). Then we have the following multiple vector-valued result:

Theorem 8. *Let R_1, R_2 and R be as above. If the tuples R_1, R_2, R satisfy the condition $(r_1^j, r_2^j, r^j) \in \mathcal{D}_{r_1^{j+1}, r_2^{j+1}, r^{j+1}}$ for every $1 \le j \le n-1$, then there exists a set $\mathcal{D}_{R_1,R_2,R}$ of triples (p,q,s) for which*

$$\text{BHT}: L^p(\mathbb{R}; L^{R_1}(\mathcal{W},\mu)) \times L^q(\mathbb{R}; L^{R_2}(\mathcal{W},\mu)) \to L^s(\mathbb{R}; L^R(\mathcal{W},\mu)).$$

In addition, $\mathcal{D}_{R_1,R_2,R} = \mathcal{D}_{r_1^1, r_2^1, r^1}$.

Remark. (1) The vector spaces $L^r(\mathcal{W}_j, \Sigma_j, \mu_j)$ can be both discrete ℓ^r spaces or the Euclidean $L^r(\mathbb{R})$ spaces. For our applications, they are going to be either of these.

(2) If the exponents $R_1 = (r_1^1, \dots, r_1^n)$, $R_2 = (r_2^1, \dots, r_2^n)$ and $R = (r^1, \dots, r^n)$ are in the "local L^2" range, then the multiple vector-valued inequalities hold for any $(p, q, s) \in \text{Range(BHT)}$. As particular cases, we mention

$$\text{BHT} : L^p(\ell^2(\ell^\infty)) \times L^q(\ell^\infty(\ell^2)) \to L^s(\ell^2(\ell^2)),$$
$$\text{BHT} : L^p(\ell^2(\ell^\infty)) \times L^q(\ell^2(\ell^2)) \to L^s(\ell^1(\ell^2))$$

for any $(p, q, s) \in \text{Range(BHT)}$.

Also, for proving an equivalent of Theorem 6 in mixed norm spaces, we need the more complex version

$$\text{BHT} : L_x^{p_1}(L_y^{p_2}(\ell^\infty(\ell^2))) \times L_x^{q_1}(L_y^{q_2}(\ell^2(\ell^2))) \to L_x^{s_1}(L_y^{s_2}(\ell^2(\ell^1))).$$

(3) As mentioned earlier, multiple vector-valued estimates for BHT play an important role in estimating $\text{BHT} \otimes \Pi^{\otimes^n}$. In the case $n = 1$, one can obtain estimates for $\text{BHT} \otimes \Pi$ in the Banach range by using duality and vector-valued inequalities of the type

$$\text{BHT} : L^p(\ell^2) \times L^q(\ell^\infty) \to L^s(\ell^2) \quad \text{and} \quad \text{BHT} : L^p(\ell^\infty) \times L^q(\ell^2) \to L^s(\ell^2).$$

However, ℓ^1-valued estimates cannot be avoided for $n \geq 3$, for example, if $\Pi \otimes \Pi \otimes \Pi$ has the form

$$\Pi \otimes \Pi \otimes \Pi(f, g)(x, y, z) = \sum_{k,l,m} Q_k^1 Q_l^2 P_m^3 \big(P_k^1 Q_l^2 Q_m^3 f \cdot Q_k^1 P_l^2 Q_m^3\big)(x, y, z).$$

This is in part the novelty of our approach in Theorem 6, and it contrasts with the situation of classical Calderón–Zygmund operators, where ℓ^1-valued estimates cannot be expected.

(4) The optimality of the range in Theorem 7 or that in Theorem 8 remains without answer, for now. Since we use in our proofs the model operator for BHT, the obstructions appearing are similar to those in [Lacey and Thiele 1999]. These are described in the constraint $C(r_1, r_2, r')$ on page 1954.

Equally important are multiple vector-valued inequalities for paraproducts, as they are essential in proving Theorem 4.

Theorem 9. *For any tuples* $R_1 = (r_1^1, \dots, r_1^n)$, $R_2 = (r_2^1, \dots, r_2^n)$ *and* $R = (r^1, \dots, r^n)$ *satisfying componentwise* $1 < R_1, R_2 \leq \infty$, $1 \leq R < \infty$, *and* $\frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{R}$,

$$\Pi : L^p(\mathbb{R}; L^{R_1}(\mathcal{W}, \mu)) \times L^q(\mathbb{R}; L^{R_2}(\mathcal{W}, \mu)) \to L^s(\mathbb{R}; L^R(\mathcal{W}, \mu)),$$

provided $1 < p, q \leq \infty$, $\frac{1}{2} < s < \infty$, *and* $\frac{1}{p} + \frac{1}{q} = \frac{1}{s}$.

In other words, vector-valued estimates for paraproducts exist within the same range as that of scalar paraproducts. This is also the case with classical Calderón–Zygmund operators.

Original motivation. We now describe the previously mentioned Rubio de Francia operator for iterated Fourier integrals, and the context where it appeared. AKNS systems are systems of differential equations of the form

$$u' = i\lambda Du + Au, \tag{10}$$

where $u = [u_1, \dots, u_n]^t$ is a vector-valued function defined on $\mathbb{R}$, D is a diagonal $n \times n$ matrix with real and distinct entries $d_1, d_2, \dots, d_n$, and $A = (a_{jk}(\cdot))_{j,k=1}^n$ is a matrix-valued function defined on $\mathbb{R}$ and such that $a_{jj} \equiv 0$ for all $1 \le j \le n$.

Then one would like to prove that the solutions u_j^λ (which depend on λ as well) are bounded "for all times"; that is,

$$\|u_j^\lambda\|_\infty < \infty \quad \text{for a.e. } \lambda \text{ and all } 1 \le j \le n. \tag{11}$$

We want to have such an estimate under the weakest possible assumptions, so we only require the entries of the potential matrix A to be integrable in some L^p spaces:

$$a_{jk}(\cdot) \in L^{p_{jk}}(\mathbb{R}) \quad \text{for all } 1 \le j, k \le n,\ j \ne k.$$

In the case of an upper triangular matrix A, whose entries are functions $g_k \in L^{p_k}$, the solutions $u_j(t)$ at a fixed time t are a finite sum of expressions of the form

$$C \int_{x_1 < \cdots < x_m < t} g_1(x_1) \cdots g_m(x_m) e^{i\lambda(\alpha_1 x_1 + \cdots + \alpha_m x_m)}\, dx_1 \cdots dx_m.$$

Here $m \le n$ and $\alpha_k \ne 0$ for all k, as a consequence of $d_1 \ne \cdots \ne d_n$. Hence the problem (11) reduces to estimating

$$\widetilde{C}_m^\alpha(g_1, g_2, \dots, g_m)(\lambda) := \sup_t \left| \int_{x_1 < \cdots < x_m < t} g_1(x_1) \cdots g_m(x_m) e^{i\lambda(\alpha_1 x_1 + \cdots + \alpha_m x_m)}\, dx_1 \cdots dx_m \right|.$$

It was proved by Christ and Kiselv [2001a; 2001b] that $\widetilde{C}_m^\alpha$ is a bounded operator:

$$\big\|\widetilde{C}_m^\alpha(g_1, \dots, g_m)\big\|_{s_m} \lesssim \prod_{k=1}^m \|g_k\|_{p_k}$$

for all $1 \le p_k < 2$ such that $\frac{1}{s_m} = \frac{1}{p_1'} + \cdots + \frac{1}{p_m'}$.

On the other hand, if the entries of the matrix A are L^2 functions, the previous expression becomes equivalent to

$$\sup_t \left| \int_{x_1 < \cdots < x_m < t} \hat{f}_1(x_1) \cdots \hat{f}_m(x_m) e^{i\lambda(\alpha_1 x_1 + \cdots + \alpha_m x_m)}\, dx_1 \cdots dx_m \right|, \tag{12}$$

denoted $C_m^\alpha(f_1, \dots, f_m)(\lambda)$. For $m = 1$, this is exactly the Carleson operator, while $m = 2$ corresponds to the bi-Carleson operator of [Muscalu et al. 2006b], both of which are known to be bounded operators (with the remark that for the bi-Carleson, the α_k need to satisfy some nondegeneracy condition):

$$\|C_2^\alpha(h_1, h_2)\|_{s_2} \lesssim \|h_1\|_{p_1} \|h_2\|_{p_2}$$

for $1 < p_1, p_2 \le \infty$, $\frac{1}{s_2} = \frac{1}{p_1} + \frac{1}{p_2}$, and $\frac{2}{3} < s_2 < \infty$.

Moreover, if instead of considering the sup in the expression (12), we look at the limiting behavior $\lim_{t\to\infty} u_j(t)$, then we encounter iterated Fourier integrals, for example, the BHT operator as seen in (1), or the bi-est operator of [Muscalu et al. 2004b]:

$$\int_{\xi_1<\xi_2<\xi_3} \hat{f}_1(\xi_1)\,\hat{f}_2(\xi_2)\,\hat{f}_3(\xi_3)\,e^{2\pi i x(\xi_1+\xi_2+\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3.$$

Now we consider the following mixed problem: The matrix A is the sum of a lower triangular matrix with entries $\hat{f}_k \in L^2$, and an upper triangular matrix with entries $g_k \in L^{p_k}$, where $1 \le p_k < 2$. Using Picard iteration, the solutions $u_j(t)$ can be expressed as a series of terms of the form

$$C\int_R \hat{f}_{11}(\xi_{11})\cdots\hat{f}_{1m_1}(\xi_{1m_1})g_{21}(x_{21})\cdots g_{2n_2}(x_{2n_2})\cdots\hat{f}_{l1}(\xi_{l1})\cdots\hat{f}_{lm_l}(\xi_{lm_l})\,dx\,d\xi,$$

where $R = \{\xi_{11} < \cdots < \xi_{1m_1} < x_{21} < \cdots < x_{2n_2} < \cdots < \xi_{l1} < \cdots < \xi_{lm_l} < t\}$.

The simplest of these operators, where the sup is dropped, is given by

$$M(f_1, f_2, g)(\xi) = \int_{x_1<x_2<x_3} \hat{f}_1(x_1)\hat{f}_2(x_2)g(x_3)e^{2\pi i\xi(x_1+x_2+x_3)}\,dx_1\,dx_2\,dx_3, \tag{13}$$

where $f_1 \in L^{p_1}$, $f_2 \in L^{p_2}$, $1 < p_1, p_2 < \infty$, and $g \in L^p$ with $1 < p < 2$. The techniques from [Christ and Kiselev 1998; 2001a; 2001b], akin to those used by Paley [1931], are based on a dyadic filtration associated to one of the functions. This involves a structure on $\mathbb{R}$ similar to that of the dyadic mesh: on every level of the filtration, one has a partition of $\mathbb{R}$, and passing to the next level of the filtration means refining the previous partition. We want to use g in order to obtain this structure and for simplicity we assume $\|g\|_p = 1$. Define the function

$$\varphi(x) = \int_{-\infty}^{x} |g(y)|^p\,dy.$$

Its image is the unit interval [0, 1], and the filtration will consist of preimages through φ of the collection $\mathcal{D}$ of dyadic intervals in [0, 1]. Because φ is increasing, whenever $x_2 < x_3$ we have $0 \le \varphi(x_2) \le \varphi(x_3) \le 1$. Hence there exists a unique dyadic interval $\omega \subset [0, 1]$ such that $\varphi(x_2)$ is contained in the left half of ω, which we denote ω_L, while $\varphi(x_3)$ is contained in the right half ω_R. To simplify notation, we identify $\varphi^{-1}(\omega)$ with ω.

Then the operator M can be written as

$$\begin{aligned}\sum_{\omega\in\mathcal{D}}\int_{\substack{x_1<x_2\\ x_2\in\omega_L, x_3\in\omega_R}} &\hat{f}_1(x_1)\hat{f}_2(x_2)g(x_3)e^{2\pi i\xi(x_1+x_2+x_3)}\,dx_1\,dx_2\,dx_3\\ &= \sum_{\omega}\int_{\substack{x_1<x_2\\ x_1,x_2\in\omega_L, x_3\in\omega_R}} \hat{f}_1(x_1)\hat{f}_2(x_2)g(x_3)e^{2\pi i\xi(x_1+x_2+x_3)}\,dx_1\,dx_2\,dx_3 \qquad (14)\\ &\quad+\sum_{\omega}\int_{\substack{x_1<L(\omega_L)\\ x_2\in\omega_L, x_3\in\omega_R}} \hat{f}_1(x_1)\hat{f}_2(x_2)g(x_3)e^{2\pi i\xi(x_1+x_2+x_3)}\,dx_1\,dx_2\,dx_3. \qquad (15)\end{aligned}$$

Here $L(\omega_L)$ denotes the left endpoint of the interval ω_L. We call the operators in (14) and (15) M_1 and M_2 respectively. The first term M_1 accounts for the occurrence of arbitrary intervals (they are in fact

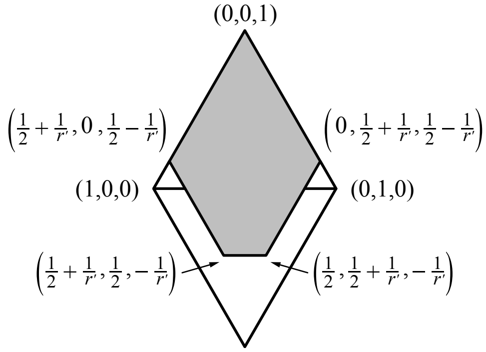

Figure 5. Range for T_r operator for $1 \le r \le 2$.

$\varphi^{-1}(\omega_L)$), and this combined with Hölder's inequality motivates the operator

$$T_r(f,g)(x) = \left(\sum_{k=1}^{N} \left| \int_{a_k<\xi_1<\xi_2<b_k} \hat{f}(\xi_1)\,\hat{g}(\xi_2)\,e^{2\pi i x(\xi_1+\xi_2)}\,d\xi_1\,d\xi_2 \right|^r \right)^{\frac{1}{r}}. \tag{16}$$

We have the following result:

Theorem 10. *If* $1 \le r \le 2$, *then*

$$\|T_r(f,g)\|_s \lesssim \|f\|_p\,\|g\|_q$$

whenever $\frac{1}{p}+\frac{1}{q}=\frac{1}{s}$, *and* p,q,s *satisfy*

$$0 \le \frac{1}{p}, \frac{1}{q} < \frac{1}{2}+\frac{1}{r}, \qquad -\frac{1}{r'} < \frac{1}{s'} < 1.$$

On the other hand, if $r \ge 2$, *then* T_r *is a bounded operator with the same range as the* BHT *operator; see Figure 5.*

In Section 7 we will show how both M_1 and M_2 are bounded operators:

Theorem 11. *The operators* M_1 *and* M_2 *satisfy the following*:

$$M_1 : L^{p_1} \times L^{p_2} \times L^p \to L^q \quad \text{provided } 1<p<2 \text{ and } \frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p'}=\frac{1}{q},$$

while

$$M_2 : L^{p_1} \times L^{p_2} \times L^p \to L^q \quad \text{provided } 1<p<2,\ \frac{1}{p_2}+\frac{1}{p'}<1 \text{ and } \frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p'}=\frac{1}{q}.$$

Hence $M = M_1+M_2$ *is a bounded operator from* $L^{p_1}\times L^{p_2}\times L^p \to L^q$ *provided* $1<p<2$, $\frac{1}{p_2}+\frac{1}{p'}<1$ *and* $\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p'}=\frac{1}{q}$.

However, as Robert Kesler [2015] noticed, the boundedness of the operator M can also be proved by making use of a vector-valued extension for the "linear" operator $\mathrm{BHT}(f_1,\cdot)$. The constraint for the exponents is given by $\frac{1}{p_2}+\frac{1}{p'}<1$. So even if M splits as $M=M_1+M_2$ and the range of M_1 is larger, one gets the same range for M through both methods.

Because the intervals $\{[a_k, b_k]\}_k$ are disjoint and arbitrary, we refer to T_r as a *bilinear Rubio de Francia operator for iterated Fourier integrals*. Recall that *Rubio de Francia's square function* is the operator

$$f \mapsto \mathrm{RF}(f)(x) := \left(\sum_{k=1}^{N} \left| \int_{I_k} \hat{f}(\xi) e^{2\pi i \xi x} \, d\xi \right|^2 \right)^{\frac{1}{2}} = \left(\sum_{k=1}^{N} |P_{I_k} f(x)|^2 \right)^{\frac{1}{2}},$$

where $\{I_k = [a_k, b_k]\}_{1 \le k \le N}$ is a family of disjoint intervals, and $P_I(f)$ denotes the Fourier projection of f onto the interval I. Using vector-valued singular integrals theory, Rubio de Francia [1985] proved the boundedness of the RF operator on L^p for $p \ge 2$. Interpolating this result with estimates for Carleson's operator [1966], one gets more generally that the operator

$$\mathrm{RF}_\nu(f)(x) := \left(\sum_{k=1}^{N} |P_{I_k} f(x)|^\nu \right)^{\frac{1}{\nu}}$$

is bounded on L^p, as long as $\frac{1}{p} + \frac{1}{\nu} < 1$.

In the particular case of a *lacunary* family of intervals (that is, $I_k = [2^{k-1}, 2^k]$ and $k \in \mathbb{Z}$), the above operator corresponds to a Littlewood–Paley square function with sharp cutoffs, which is bounded on $L^p(\mathbb{R})$ for any $1 < p < \infty$. Even more, the L^p norm of the square function is comparable to the L^p norm of the initial function:

$$C_p^{-1} \|f\|_p \le \left\| \left(\sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \mathbf{1}_{\{2^{k-1} \le \xi < 2^k\}} \hat{f}(\xi) e^{2\pi i x \xi} \, d\xi \right|^2 \right)^{\frac{1}{2}} \right\|_p \le C_p \|f\|_p.$$

Rubio de Francia's theorem addresses the boundedness of a square function associated to an arbitrary family of intervals, and in this sense it is optimal: in the case $\nu = 2$, the condition $p \ge 2$ is necessary, while for $\nu > 2$, we need the strict inequality $\nu > p'$.

Returning to our operator T_r, note that it can also be regarded as a vector-valued bilinear Hilbert transform

$$T_r(f, g)(x) = \left(\sum_k \left| \mathrm{BHT}(P_{I_k} f, P_{I_k} g)(x) \right|^r \right)^{\frac{1}{r}},$$

because the multiplier of the BHT operator is equivalent to $\mathbf{1}_{\{\xi_1 < \xi_2\}}$, as seen in (1).

Using solely Khintchine's inequality, it was proved in [Grafakos and Li 2006] that

$$\left\| \left(\sum_k \left| \mathrm{BHT}(f_k, g_k) \right|^2 \right)^{\frac{1}{2}} \right\|_s \lesssim \left\| \left(\sum_k |f_k|^2 \right)^{\frac{1}{2}} \right\|_p \left\| \left(\sum_k |g_k|^2 \right)^{\frac{1}{2}} \right\|_q.$$

This implies the boundedness of T_r for $r \ge 2$, $p, q \ge 2$. But this is a very limited range, and in order to obtain estimates in the case $p < 2$ or $q < 2$, one needs the full power of vector-valued extensions.

We note that our estimates for the operator T_r are sharp, in the sense that the same estimates are satisfied by

$$(f, g) \mapsto \left(\sum_k \left| P_{I_k} f(x) \cdot P_{I_k} g(x) \right|^r \right)^{\frac{1}{r}}. \tag{17}$$

In (17), $\mathrm{BHT}(P_{I_k} f, P_{I_k} g)$ is replaced by the product of the functions $P_{I_k} f \cdot P_{I_k} g$. In general, the best one can hope for a bilinear Fourier multiplier operator is that it satisfies the same L^p estimates as the product $(f, g) \mapsto f \cdot g$, and this is the case for T_r.

Moreover, in the special case of lacunary dyadic intervals, for any $1 \le r < \infty$, we have that

$$(f, g) \mapsto \Big(\sum_k \Big| \int_{2^k < \xi < \eta < 2^{k+1}} \hat{f}(\xi)\, \hat{g}(\eta)\, e^{2\pi i x(\xi+\eta)}\, d\xi\, d\eta \Big|^r \Big)^{\frac{1}{r}}$$

is a bounded operator from $L^p \times L^q$ to L^s for any $(p, q, s) \in \mathrm{Range(BHT)}$. The cases $p = \infty$ and $q = \infty$ cannot be obtained directly, but follow by duality.

Our initial proof of Theorem 10 did not involve vector-valued bilinear Hilbert transform operators, but it was built around localizations of BHT, in conjunction with several stopping times. Afterwards we realized that this method is suitable for other general situations, which eventually led to the development of the helicoidal method. This applies to paraproducts, BHT, the Carleson operator, the Rubio de Francia operator, etc. In the study of the T_r operator, the stopping times were dictated by level sets of linear Rubio de Francia operators: $\mathrm{RF}_{r_1}(f)$ and $\mathrm{RF}_{r_2}(g)$. For the vector-valued BHT, the three stopping times that are used for estimating the trilinear form are dictated by level sets of

$$\Big(\sum_k |f_k|^{r_1} \Big)^{\frac{1}{r_1}}, \quad \Big(\sum_k |g_k|^{r_2} \Big)^{\frac{1}{r_2}} \quad \text{and} \quad \Big(\sum_k |h_k|^{r'} \Big)^{\frac{1}{r'}}.$$

The method of the proof is described in more detail in Section 2.5.

Lastly, we want to point out an interesting connection with another open problem in time-frequency analysis: the boundedness of the Hilbert transform along vector fields. More exactly, if $v : \mathbb{R}^2 \to \mathbb{R}^2$ is a nonvanishing measurable vector field, then one defines the Hilbert transform along v as

$$H_v f(x, y) = \text{p.v.} \int_{\mathbb{R}} f\big((x, y) - t \cdot v(x, y)\big) \frac{dt}{t}.$$

It was conjectured by Stein that H_v is a bounded operator on L^2 whenever v is Lipschitz. Some partial results in this direction are known in the case of a one-variable vector field. M. Bateman and C. Thiele [2013] proved the L^p boundedness of H_v for $\frac{3}{2} < p < \infty$ and provided that $v(x, y) = v(x, 0)$.

The proof makes use of the Littlewood–Paley square function in the second variable and restrictions to certain fixed sets G and H, together with single annulus estimates for H_v from [Bateman 2013]. In the special case when $f(x, y) = g(x)h(y)$, estimates for the variational Carleson from [Oberlin et al. 2012] yield the same result whenever $p > \frac{4}{3}$. It is still not known if this can be extended to general functions $f(x, y)$, or whether one can push the lower bound for p below $\frac{4}{3}$.

Silva [2014] uses ideas similar to the ones described above, obtaining in this way vector-valued extensions for BHT whenever $\frac{4}{3} < r < 4$. Our methods allow us to prove that vector-valued extensions exist for any $1 \le r < \infty$ (in fact, for any triple (r_1, r_2, r)). It would be interesting to understand whether the localization argument that we are employing can be transferred to the study of the Hilbert transform along vector fields.

Besides having sharp estimates for the local version of the operator, the structure of the intervals chosen through the triple stopping time can play a role in itself. The collections of intervals constitute a maximal covering for the level sets of certain maximal operators, and for that reason, they form a *sparse* collection of intervals (in the sense of [Lerner 2013]). From here, weighted estimates can be deduced, and a similar approach was carried out in [Culiuc et al. 2016].

The rest of the paper is organized as follows: in Section 2 we recall some definitions and results regarding multilinear operators. The helicoidal method is described in detail in Section 2.5. Multiple vector-valued extensions for BHT are presented in Section 3, and those for paraproducts in Section 4. Following in Section 5 are the estimates for $\mathrm{BHT} \otimes \Pi^{\otimes^n}$. The Leibniz rules are a modification of mixed norm L^p estimates for $\Pi \otimes \Pi$ and are discussed in Section 6. The Rubio de Francia theorem for iterated Fourier integrals and its application to the AKNS system problem appear in Section 7.

2. Some classical results on the bilinear Hilbert transform

In this paper we use Chapter 6 of [Muscalu and Schlag 2013] as a black box, but we recall a few definitions and results to ease the reading of the presentation. Essential here are the notions of *size* and *energy*, which are quantities associated to certain subsets of the phase-frequency space.

Notation. For any interval $I \subset \mathbb{R}$, define

$$\tilde{\chi}_I(x) := \left(1 + \frac{\operatorname{dist}(x, I)}{|I|}\right)^{-100}.$$

The mesh of dyadic intervals is denoted by $\mathcal{D}$.

Definition 12. A *tile* is a rectangle $P = I_P \times \omega_P$ with the property that $I_P, \omega_P \in \mathcal{D}$ or ω_P is in a shifted variant of $\mathcal{D}$. We define a *tritile* to be a tuple $P = (P_1, P_2, P_3)$ where each P_i is a tile as defined above and the spatial intervals are the same: $I_{P_i} = I_P$ for all $1 \le i \le 3$.

Definition 13 (order relation). Given two tiles P and P', we say $P' < P$ if $I_{P'} \subsetneq I_P$ and $\omega_P \subset 3\omega_{P'}$, and $P' \le P$ if $P' < P$ or $P' = P$. Also, $P' \lesssim P$ if $I_{P'} \subset I_P$ and $\omega_P \subseteq 100\omega_{P'}$, and $P' \lesssim' P$ if $P' \lesssim P$ but $P' \nleq P$.

Definition 14. A collection $\mathbb{P}$ of tritiles is said to have *rank 1* if for any $P, P' \in \mathbb{P}$ the following conditions are satisfied:

- If the tritiles are distinct, i.e., $P \neq P'$, then $P'_j \neq P_j$ for all $1 \le j \le 3$.
- If $\omega_{P_{j_0}} = \omega_{P'_{j_0}}$ for some j_0, then $\omega_{P_j} = \omega_{P'_j}$ for all $1 \le j \le 3$.
- If $P'_{j_0} \le P_{j_0}$ for some j_0, then $P'_j \lesssim P_j$ for all $1 \le j \le 3$.
- If in addition to $P'_{j_0} \le P_{j_0}$ one also assumes $|I_{P'}| \ll |I_P|$, then $P'_j \lesssim' P_j$ for all $j \neq j_0$.

Definition 15. Let $\mathbb{P}$ be a sparse rank 1 collection of tritiles, and let $1 \le j \le 3$. A subcollection T of $\mathbb{P}$ is called a *j-tree* if and only if there exists a tritile P_T (called the *top* of the tree) such that $P_j \le P_{T,j}$ for all $P \in T$. We write I_T for I_{P_T} and ω_{T_j} for $\omega_{P_T,j}$ and we say T is a *tree* if it is a j-tree for some $1 \le j \le 3$.

Definition 16. Let $1 \le i \le 3$. A finite sequence of trees $T_1, \dots, T_M$ is said to be a chain of strongly i-disjoint trees if and only if

(i) $P_i \neq P'_i$ for every $P \in T_{l_1}$ and $P' \in T_{l_2}$, with $l_1 \neq l_2$;

(ii) whenever $P \in T_{l_1}$ and $P' \in T_{l_2}$ with $l_1 \neq l_2$ are such that $2\omega_{P_i} \cap 2\omega_{P'_i} \neq \varnothing$, then if $|\omega_{P_i}| < |\omega_{P'_i}|$, one has $I_{P'} \cap I_{T_{l_1}} = \varnothing$, and if $|\omega_{P'_i}| < |\omega_{P_i}|$, one has $I_P \cap I_{T_{l_2}} = \varnothing$.

(iii) whenever $P \in T_{l_1}$ and $P' \in T_{l_2}$ with $l_1 < l_2$ are such that $2\omega_{P_i} \cap 2\omega_{P'_i} \neq \varnothing$ and $|\omega_{P_i}| = |\omega_{P'_i}|$, then $I_{P'} \cap I_{T_{l_1}} = \varnothing$.

Definition 17. Let P be a tile. A *wave packet* on P is a smooth function ϕ_P which has Fourier support inside $\frac{9}{10}\omega_P$ and is L^2-adapted to I_P in the sense that

$$|\phi_P^{(l)}(x)| \le C_{l,M} \frac{1}{|I_P|^{\frac{1}{2}+l}} \left(1 + \frac{\operatorname{dist}(x, I_P)}{|I_P|}\right)^{-M} \tag{18}$$

for sufficiently many derivatives l and any $M > 0$.

2.1. ***Model operator for* BHT.** A discretized model operator for BHT is given by

$$\mathrm{BHT}_{\mathbb{P}}(f, g)(x) = \sum_{P \in \mathbb{P}} \frac{1}{|I_P|^{\frac{1}{2}}} \langle f, \phi_{P_1}^1 \rangle \langle g, \phi_{P_2}^2 \rangle \phi_{P_3}^3(x), \tag{19}$$

where the family $\mathbb{P}$ of tritiles is sparse and has rank 1, while $(\phi_{P_j}^j)_{P \in \mathbb{P}}$ are wave packets associated to the tiles P_j. In some sense, the bilinear Hilbert transform is the canonical example of such an operator. Above we also included the definitions of *trees* and *chains of strongly disjoint trees* because they are essential in understanding such singular bilinear operators.

The model operator from (19) was introduced in [Lacey and Thiele 1999], and the bilinear Hilbert transform itself can be represented as an average of such shifted model operators. The detailed reduction can be found in [Muscalu and Schlag 2013, Chapter 6]. As a consequence, the boundedness of the bilinear Hilbert transform within Range(BHT) can be deduced from similar estimates for the model operator. Similarly, estimates for vector-valued and for the localized bilinear Hilbert transform will follow once we prove their equivalents for the model operator, and we will not insist on the exact distinction between the two.

It is worth mentioning however, that the model operator fails to be bounded for $s \le \frac{2}{3}$, leaving undecided the boundedness of the bilinear Hilbert transform itself for $\frac{1}{2} < s \le \frac{2}{3}$.

Bilinear operators are often studied with the use of the associated trilinear form. In the case of the (model operator for the) BHT operator, the trilinear form is given by

$$\Lambda_{\mathrm{BHT};\mathbb{P}}(f, g, h) = \sum_{P \in \mathbb{P}} \frac{1}{|I_P|^{\frac{1}{2}}} \langle f, \phi_{P_1}^1 \rangle \langle g, \phi_{P_2}^2 \rangle \langle h, \phi_{P_3}^3 \rangle. \tag{20}$$

Definition 18. If $\mathbb{P}$ is a collection of tritiles and I_0 is a dyadic interval, we denote by $\mathbb{P}(I_0)$ the tiles P in $\mathbb{P}$ whose spatial interval I_P is contained in I_0:

$$\mathbb{P}(I_0) := \{P \in \mathbb{P} : I_P \subseteq I_0\}.$$

Definition 19. Let $\mathbb{P}$ be a finite collection of tritiles, let $j \in \{1,2,3\}$, and let f be an arbitrary function. We define the *size* of the sequence $\langle f, \phi_{P_j}^j \rangle_P$ by

$$\operatorname{size}\big(\langle f, \phi_{P_j}^j \rangle_P\big) := \sup_{T \subseteq \mathbb{P}} \bigg(\frac{1}{|I_T|} \sum_{P \in T} |\langle f, \phi_{P_j}^j \rangle|^2 \bigg)^{\frac{1}{2}}, \tag{21}$$

where T ranges over all trees in $\mathbb{P}$ that are i-trees for some $i \neq j$.

Lemma 20 [Muscalu and Schlag 2013, Lemma 6.13]. *Let $j \in \{1,2,3\}$ and let E be a set of finite measure. Then for every $|f| \leq \mathbf{1}_E$ one has*

$$\operatorname{size}\big(\langle f, \phi_{P_j}^j \rangle_P\big) \lesssim \sup_{P \in \mathbb{P}} \frac{1}{|I_P|} \int_E \tilde{\chi}_{I_P}^M \, dx$$

for all $M > 0$, with implicit constants depending on M.

Thanks to Lemma 20, which is a consequence of the John–Nirenberg inequality, we can work with the simpler "sizes"

$$\operatorname{size} f \sim \sup_{P \in \mathbb{P}} \frac{1}{|I_P|} \int_{\mathbb{R}} |f| \cdot \tilde{\chi}_{I_P}^M \, dx,$$

where M is some large number to be chosen later.

We will also need a size that behaves well with respect to localization. In the formula above we consider the supremum over the spacial intervals I_P of the collection $\mathbb{P}$. In our proofs, we will need to compare $\operatorname{size}_{\mathbb{P}(I_0)} f$ and $(1/|I_0|) \int_{\mathbb{R}} |f| \cdot \tilde{\chi}_{I_0} \, dx$, so the following definition is natural:

Definition 21. If I_0 is a fixed dyadic interval, then we define

$$\widetilde{\operatorname{size}}_{\mathbb{P}(I_0)} f := \sup_{\substack{J \subseteq 3I_0 \\ \exists P \in \mathbb{P}(I_0), I_P \subseteq J}} \frac{1}{|J|} \int_{\mathbb{R}} |f| \cdot \tilde{\chi}_J^M \, dx. \tag{22}$$

We note that for any function f,

$$\operatorname{size}_{\mathbb{P}(I_0)} f \leq \widetilde{\operatorname{size}}_{\mathbb{P}(I_0)} f.$$

Definition 22. Let $\mathbb{P}$ be a finite collection of tritiles, $j \in \{1,2,3\}$ and let f be a fixed function. We define the energy of the sequence $\langle f, \phi_{P_j}^j \rangle_P$ by

$$\operatorname{energy}\big(\langle f, \phi_{P_j}^j \rangle_P\big) := \sup_{n \in \mathbb{Z}} 2^n \sup_{\mathbb{T}} \bigg(\sum_{T \in \mathbb{T}} |I_T| \bigg)^{\frac{1}{2}}, \tag{23}$$

where $\mathbb{T}$ ranges over all chains of strongly j-disjoint trees in $\mathbb{P}$ (which are i-trees for some $i \neq j$) having the property that

$$\bigg(\sum_{P \in T} |\langle f, \phi_{P_j}^j \rangle|^2 \bigg)^{\frac{1}{2}} \geq 2^n |I_T|^{\frac{1}{2}}$$

for all $T \in \mathbb{T}$ and such that

$$\bigg(\sum_{P \in T'} |\langle f, \phi_{P_j}^j \rangle|^2 \bigg)^{\frac{1}{2}} \leq 2^{n+1} |I_{T'}|^{\frac{1}{2}}$$

for all subtrees $T' \subseteq T \in \mathbb{T}$.

We have the following estimates for the trilinear form and energy:

Proposition 23 [Muscalu and Schlag 2013, Proposition 6.12]. *Let $\mathbb{P}$ be a finite collection of tritiles. Then*

$$\Lambda_{\mathrm{BHT};\mathbb{P}}(f_1, f_2, f_3) \lesssim \prod_{j=1}^{3} \big(\mathrm{size}(\langle f_j, \phi^j_{P_j}\rangle_P)\big)^{\theta_j} \big(\mathrm{energy}(\langle f_j, \phi^j_{P_j}\rangle_P)\big)^{1-\theta_j}$$

for any $0 \le \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$; the implicit constants depend on the θ_j but are independent of the other parameters.

Lemma 24 [Muscalu and Schlag 2013, Lemma 6.14]. *Let $j \in \{1, 2, 3\}$ and $f \in L^2(\mathbb{R})$. Then*

$$\mathrm{energy}\big(\langle f, \phi^j_{P_j}\rangle_P\big) \lesssim \|f\|_2.$$

However, for our specific problem we need more accurate estimates for the localized trilinear form. This will follow in Sections 2.4 and 3.1.

2.2. *Interpolation.* Since this is a fundamental tool in harmonic analysis, we recall a few facts about interpolation methods. We adapt the results from [Thiele 2006] and emphasize how the constants change through interpolation. In our applications, we need to keep track of the constants. Many of the proofs in the following sections are iterative, and the operatorial norm obtained after interpolation becomes a "size" on the subsequent step of the induction. We recall a few definitions and results, but we will be mainly using their generalization to Banach spaces.

Definition 25. For a subset $E \subset \mathbb{R}$ of finite measure, define

$$X(E) = \{f : |f| \le \mathbf{1}_E \text{ a.e.}\}.$$

We will denote by V the linear span of all $X(E)$, which plays an important role because it is a dense subspace of all L^p spaces for $1 \le p < \infty$.

Definition 26. A tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ is called *admissible* if for all $1 \le i \le n$,

$$-\infty < \alpha_i < 1 \quad \text{and} \quad \alpha_1 + \dots + \alpha_n = 1,$$

and there is at most one index j_0 so that $\alpha_{j_0} < 0$. We call an index *good* if $\alpha_i > 0$ and *bad* if $\alpha_i \le 0$.

Definition 27. A multilinear form $\Lambda : V \times \dots \times V \to \mathbb{C}$ is of restricted type $\alpha = (\alpha_1, \dots, \alpha_n)$ with $0 \le \alpha_i \le 1$ if there exists a constant C (possibly depending on α) such that for each tuple $E = (E_1, \dots, E_n)$ of measurable subsets of $\mathbb{R}$ and for each tuple $f = (f_1, \dots, f_n)$ with $f_j \in X(E_j)$, we have

$$\big|\Lambda(f_1, \dots, f_n)\big| \le C \prod_j |E_j|^{\alpha_j}.$$

Theorem 28 (similar to [Thiele 2006, Theorem 3.2]). *Let $\beta = (\beta_1, \dots, \beta_n)$ be a tuple of real numbers such that $\sum_j \beta_j = 1$ and $\beta_j > 0$ for all j. Assume Λ is of restricted type α for all α in a neighborhood of β satisfying $\sum_j \alpha_j = 1$, with constant $C(\alpha)$ depending continuously on α. Then Λ is of strong type β with constant $C(\beta)$:*

$$\big|\Lambda(f_1, \dots, f_n)\big| \le C(\beta) \prod_{j=1}^{n} \|f_j\|_{\frac{1}{\beta_j}} \quad \textit{for all } f_j \in V.$$

For multilinear operators, it often happens that the target space is an L^p space with $0 < p < 1$. This is not a Banach space, but we can conclude the desired outcome by interpolating weak-L^q estimates for q in a neighborhood of p. Additionally, $L^{q,\infty}$ norms are dualized in the following way:

Lemma 29 [Muscalu and Schlag 2013, Lemma 2.5]. *Let* $0 < r \le 1$, *and* $A > 0$. *Then the following statements are equivalent*:

(i) $\|f\|_{r,\infty} \le A$.

(ii) *For every set* E *with* $0 < |E| < \infty$, *there exists a major subset* $E' \subseteq E$ (*i.e.*, $|E'| \ge |E|/2$) *so that* $|\langle f, \mathbf{1}_{E'}\rangle| \lesssim A|E|^{\frac{1}{r'}}$, *where* $\frac{1}{r} + \frac{1}{r'} = 1$. (*Note that for* $r \ne 1$, *we have* r' *is a negative number.*)

Definition 30. Let α be an n-tuple of real numbers and assume $\alpha_j \le 1$ for all j. An n-linear form Λ is called of *generalized restricted type* α if there is a constant C (possibly depending on α) such that for all tuples $E = (E_1, \dots, E_n)$, there is an index j_0 and a major subset $E'_{j_0} \subseteq E_{j_0}$ so that for all tuples $f = (f_1, \dots, f_n)$ with $f_j \in X(E_j)$ for $j \ne j_0$ and $f_{j_0} \in X(E'_{j_0})$,

$$\big|\Lambda(f_1, \dots, f_n)\big| \le C \prod_{j=1}^{n} |E_j|^{\alpha_j}. \tag{24}$$

If a tuple $\alpha = (\alpha_1, \dots, \alpha_n)$ is good, then generalized restricted-type estimates coincide with restricted-type estimates:

Proposition 31 (similar to [Thiele 2006, Lemma 3.6]). *If* $\alpha = (\alpha_1, \dots, \alpha_n)$ *is a good tuple*, *and* Λ *is of generalized restricted type* α *with constant* $C(\alpha)$ *and the major subset corresponds to the index* j_0, *then* Λ *is of restricted type* α *with constant* $C(\alpha)/(1-2^{-j_0})$.

Theorem 32 [Thiele 2006, Theorem 3.8]. *Assume*

$$\Lambda = \big\langle T(f_1, \dots, f_{n-1}), f_n\big\rangle$$

is of generalized restricted type β, *where* $\sum_j \beta_j = 1$. *Assume* $\beta_k > 0$ *for* $1 \le k \le n-1$ *and* $\beta_n \le 0$. *Assume* Λ *is also of generalized restricted type* α *with constant* $C(\alpha)$ (*continuously depending on* α) *for all* α *in a neighborhood of* β *satisfying* $\sum_j \alpha_j = 1$. *Then the multilinear operator* T *satisfies*

$$\big\|T(f_1, \dots, f_{n-1})\big\|_{\frac{1}{(1-\beta_n)}} \le C(\beta) \prod_{j=1}^{n-1} \|f_j\|_{\frac{1}{\beta_j}}. \tag{25}$$

2.3. ***Interpolation for Banach-valued functions.*** The Banach space interpolation theory is very similar to the scalar version, the difference consisting in replacing the norm $|\cdot|$ on $\mathbb{C}$ by $\|\cdot\|_X$ on a Banach space X.

We say that $F \in L^p(\mathbb{R}; X)$ provided

$$\|F\|_{L^p(\mathbb{R};X)} := \left(\int_{\mathbb{R}} \|F(x)\|_X^p \, dx\right)^{\frac{1}{p}} < \infty.$$

The question of integrability of $F(x)$ is reduced to the Lebesgue integrability of $x \mapsto \|F(x)\|_X$. The set of vector-valued step functions is dense in $L^p(\mathbb{R}; X)$ and for this reason, similarly to the scalar case, it will be enough to deal with function in

$$\{F : \|F(x)\|_X \le \mathbf{1}_E(x) \text{ a.e. } E \subset \mathbb{R} \text{ subset of finite measure}\}.$$

The linear span of such sets will be denoted V_X.

The multilinear form associated with an operator is obtained through dualization. More exactly,

$$\|F\|_{L^p(\mathbb{R};X)} := \sup_{\|G\|_{L^{p'}(\mathbb{R};X^*)} \le 1} \left| \int_{\mathbb{R}} \langle G(x), F(x) \rangle \, dx \right|$$

whenever $1 \le p < \infty$.

We will deal with a vector-valued multilinear (or multisublinear) operator of the form

$$\vec{T} : L^{p_1}(\mathbb{R}; X_1) \times \cdots \times L^{p_{n-1}}(\mathbb{R}; X_{n-1}) \to L^{p_n}(\mathbb{R}; X_n).$$

The multilinear form associated with this operator, $\Lambda : V_{X_1} \times \cdots \times V_{X_{n-1}} \times V_{X_n^*} \to \mathbb{C}$, is given by

$$\Lambda(F_1, \ldots, F_{n-1}, F_n) = \int_{\mathbb{R}} \big\langle \vec{T}(F_1, \ldots, F_{n-1})(x), F_n(x) \big\rangle \, dx.$$

The definitions and proofs from the scalar case are adaptable to the vector-valued situation. For completeness, we present them here, adapting the equivalent statements from [Thiele 2006].

Definition 33. A tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ is called admissible if $\alpha_1 + \cdots + \alpha_n = 1$, $\alpha_1, \ldots, \alpha_n < 1$ and for at most one index j_0 we have $\alpha_{j_0} < 0$.

A multisublinear form Λ as above is of restricted type $\alpha = (\alpha_1, \ldots, \alpha_n)$ for a good admissible tuple α if there exists a constant C so that for each tuple $E = (E_1, \ldots, E_n)$ of measurable subsets of $\mathbb{R}$, and for each tuple $F = (F_1, \ldots, F_n)$ with $\|F_j\|_X \le \mathbf{1}_{E_j}$, we have

$$\big|\Lambda(F_1, \ldots, F_n)\big| \le C |E_1|^{\alpha_1} \cdots |E_n|^{\alpha_n}.$$

Proposition 34 (equivalent of [Thiele 2006, Theorem 3.2]). *Let $\beta = (\beta_1, \ldots, \beta_n)$ be an admissible tuple of real numbers such that $\beta_j > 0$ for all j. Assume that Λ is of restricted type α for all admissible tuples α in a neighborhood of β. Then there is a constant C such that for all $F_j \in V_{X_j}$,*

$$\big|\Lambda(F_1, \ldots, F_n)\big| \le C \|F_1\|_{L^{1/\beta_1}(\mathbb{R};X_1)} \cdots \|F_n\|_{L^{1/\beta_n}(\mathbb{R};X_n)}.$$

Definition 35. Let α be an admissible tuple; the n-sublinear form Λ is of *generalized restricted type* α if there is a constant C such that for all tuples $E = (E_1, \ldots, E_n)$ there is an index j_0 and a major subset E'_{j_0} of E_{j_0}(that is, $|E'_{j_0}| \ge |E_{j_0}|/2$) such that for all tuples $F = (F_1, \ldots, F_n)$ with $\|F_j\|_{X_j} \le \mathbf{1}_{E_j}$ for $j \ne j_0$, and $\|F_{j_0}\|_{X_{j_0}} \le \mathbf{1}_{E'_{j_0}}$, we have

$$\big|\Lambda(F_1, \ldots, F_n)\big| \le C \prod_j |E_j|^{\alpha_j}.$$

Proposition 36. *If Λ is of generalized restricted type $\alpha = (\alpha_1, \ldots, \alpha_n)$, and $\alpha_j > 0$ for all j, then Λ is of restricted type α.*

On the other hand, if one of the indices α_j is ≤ 0, the generalized restricted-type implies only weak-L^p estimates. This works in the case when the multisublinear form is given by

$$\Lambda(F_1, \dots, F_n) = \int_{\mathbb{R}} \big\langle \vec{T}(F_1, \dots, F_{n-1})(x), F_n(x) \big\rangle \, dx, \tag{26}$$

and corresponds to an operator $\vec{T}$ defined on $V_{X_1} \times \cdots \times V_{X_{n-1}}$ and taking values in V_{X_n}.

Proposition 37. *Let Λ be a multisublinear form as in* (26), *and $\alpha = (\alpha_1, \dots, \alpha_n)$ an admissible tuple with $\alpha_n \le 0$. Assuming that Λ is of generalized restricted type α, we have*

$$\lambda \big| \{ x : \| \vec{T}(F_1, \dots, F_{n-1})(x) \|_{X_n} > \lambda \} \big|^{\frac{1}{1-\alpha_n}} \le A \prod_{j=1}^{n-1} |E_j|^{\alpha_j}$$

for all tuples $F = (F_1, \dots, F_{n-1})$ with $\| f_j \|_{X_j} \le \mathbf{1}_{E_j}$.

Proposition 38. *Assume Λ is of generalized restricted type β, where β is an admissible tuple with $\beta_n \le 0$. Assume Λ is also of generalized restricted type α for all admissible tuples α in a neighborhood of β. Then $\vec{T}$ satisfies*

$$\big\| \vec{T}(F_1, \dots, F_{n-1}) \big\|_{L^{1/(1-\beta_n)}(\mathbb{R}; X_n)} \le C \prod_{j=1}^{n-1} \| F_j \|_{L^{1/\beta_j}(\mathbb{R}; X_j)}. \tag{27}$$

The proofs of the last two propositions follow exactly the same ideas as those corresponding to the scalar case, with very minor differences.

2.4. ***A few technical lemmas.*** In this section, we present a few results that will be useful later on for estimating a trilinear form associated to a collection $\mathbb{P}$ of tritiles well-localized in space: $I_P \subset I_0$ for all $P \in \mathbb{P}$.

Lemma 39. *If I_0 is a fixed dyadic interval, $k \in \mathbb{Z}^+$, and f is a function such that*

$$2^{k-1} \le \frac{\operatorname{dist}(\operatorname{supp} f, I_0)}{|I_0|} \le 2^k,$$

then

$$\operatorname{energy}_{\mathbb{P}(I_0)} f \lesssim 2^{Mk} \| f \|_2.$$

Proof. Following Definition 22, there exists a collection $\mathbb{T}$ of j-disjoint trees $T \in \mathbb{T} \subseteq \mathbb{P}(I_0)$, so that

$$(\operatorname{energy}_{\mathbb{P}(I_0)} f)^2 \sim \sum_{T \in \mathbb{T}} \sum_{P \in T} \big| \langle f, \phi_{P_j} \rangle \big|^2.$$

We define $\mathcal{T} := \bigcup_{T \in \mathbb{T}} \bigcup_{P \in T} P$, the collection of all tiles in $\mathbb{T}$, and estimate the right-hand side of the expression above:

$$\sum_{T \in \mathbb{T}} \sum_{P \in T} \big| \langle f, \phi_{P_j} \rangle \big|^2 \lesssim \sum_{m \ge 0} \sum_{\substack{I \subseteq I_0 \\ |I| = 2^{-m} |I_0|}} \sum_{\substack{P \in \mathcal{T} \\ I_P = I}} \big| \langle f, \phi_{P_j} \rangle \big|^2.$$

The collection of tiles $P \in \mathcal{T}$ with $I_P = I$ for a fixed interval I are all disjoint in frequency. In fact, since they are of the same scale, they are translations of some fixed tile and hence

$$\sum_{\substack{P\in\mathcal{T}\\ I_P=I}} \left|\langle f, \phi_{P_j}\rangle\right|^2 \lesssim \int_{\mathbb{R}} |f(x)|^2 \left(1+\frac{\operatorname{dist}(x,I)}{|I|}\right)^{-2M} dx.$$

This implies

$$\begin{aligned}\sum_{T\in\mathbb{T}}\sum_{P\in T}\left|\langle f, \phi_{P_j}\rangle\right|^2 &\lesssim \sum_{m\geq 0}\sum_{\substack{I\subseteq I_0\\ |I|=2^{-m}|I_0|}} \int_{\mathbb{R}} |f(x)|^2 \left(1+\frac{\operatorname{dist}(x,I)}{|I|}\right)^{-2M} dx\\ &\lesssim \sum_{m\geq 0}\sum_{\substack{I\subseteq I_0\\ |I|=2^{-m}|I_0|}} \|f\|_2^2\, 2^{-2kM}\left(\frac{|I_0|}{|I|}\right)^{-2M}\\ &\lesssim \|f\|_2^2\, 2^{-2kM}\sum_{m\geq 0} 2^{-mM}\\ &\lesssim \|f\|_2^2\, 2^{-2kM}. \end{aligned}$$

□

On the other hand, if f is supported inside $5I_0$, we know from Lemma 24, that $\text{energy}_{\mathbb{P}(I_0)} f \lesssim \|f\|_2$.

Since the collection $\mathbb{P}(I_0)$ is localized in space on the interval I_0, we have the following estimate for the trilinear form $\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}$:

Lemma 40 (refinement of [Muscalu and Schlag 2013, Proposition 6.12]). *The trilinear form* $\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}$ *satisfies*

$$\begin{aligned}&\left|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h)\right|\\ &\qquad\lesssim (\text{size}_{\mathbb{P}(I_0)} f)^{\theta_1}(\text{size}_{\mathbb{P}(I_0)} g)^{\theta_2}(\text{size}_{\mathbb{P}(I_0)} h)^{\theta_3}\,\|f\cdot\tilde{\chi}_{I_0}\|_2^{1-\theta_1}\,\|g\cdot\tilde{\chi}_{I_0}\|_2^{1-\theta_2}\,\|h\cdot\tilde{\chi}_{I_0}\|_2^{1-\theta_3} \end{aligned} \tag{28}$$

for any $0\leq \theta_1,\theta_2,\theta_3<1$, *with* $\theta_1+\theta_2+\theta_3=1$; *the implicit constants depend on the* θ_j, *but are independent of the other parameters.*

Proof. For any $l\geq 1$, we define $\mathcal{I}_l := 2^{l+1}I_0\setminus 2^l I_0$, and $\mathcal{I}_0 := 2I_0$. In this way, for any $x\in\mathcal{I}_l$, $1+\operatorname{dist}(x,I_0)/|I_0|\sim 2^l$.

We will be using the following decompositions:

$$f := \sum_{k_1\geq 0} f_{k_1} := \sum_{k_1\geq 0} f\cdot\mathbf{1}_{\mathcal{I}_{k_1}}, \tag{29}$$

and similarly,

$$g := \sum_{k_2\geq 0} g_{k_2} := \sum_{k_2\geq 0} g\cdot\mathbf{1}_{\mathcal{I}_{k_2}},\quad h := \sum_{k_3\geq 0} h_{k_3} := \sum_{k_3\geq 0} h\cdot\mathbf{1}_{\mathcal{I}_{k_3}}.$$

From Proposition 23, the trilinear form can be estimated by

$$\begin{aligned}\left|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h)\right| &\lesssim \sum_{k_1,k_2,k_3} \left|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}(f_{k_1},g_{k_2},h_{k_3})\right| \\ &\lesssim \sum_{k_1,k_2,k_3} (\mathrm{size}_{\mathbb{P}(I_0)}\, f_{k_1})^{\theta_1} (\mathrm{size}_{\mathbb{P}(I_0)}\, g_{k_2})^{\theta_2} (\mathrm{size}_{\mathbb{P}(I_0)}\, h_{k_3})^{\theta_3} \\ &\qquad (\mathrm{energy}_{\mathbb{P}(I_0)}\, f_{k_1})^{1-\theta_1} (\mathrm{energy}_{\mathbb{P}(I_0)}\, g_{k_2})^{1-\theta_2} (\mathrm{energy}_{\mathbb{P}(I_0)}\, h_{k_3})^{1-\theta_3}.\end{aligned}$$

We will only employ the extra decay in the energy; for the size, we have simply

$$\mathrm{size}_{\mathbb{P}(I_0)}\, f_{k_1} \lesssim \mathrm{size}_{\mathbb{P}(I_0)}\, f$$

uniformly in k_1.

On the other hand, since f_{k_1} is supported on $\mathcal{I}_{k_1}$, Lemma 39 implies

$$\mathrm{energy}_{\mathbb{P}(I_0)}\, f_{k_1} \lesssim 2^{-k_1 M} \| f_{k_1} \|_2.$$

Hence we obtain

$$\begin{aligned}\left|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h)\right| &\lesssim (\mathrm{size}_{\mathbb{P}(I_0)}\, f)^{\theta_1} (\mathrm{size}_{\mathbb{P}(I_0)}\, g)^{\theta_2} (\mathrm{size}_{\mathbb{P}(I_0)}\, h)^{\theta_3} \\ &\quad \cdot \sum_{k_1,k_2,k_3} (2^{-k_1 M} \| f_{k_1} \|_2)^{1-\theta_1} (2^{-k_2 M} \| g_{k_2} \|_2)^{1-\theta_2} (2^{-k_1 M} \| h_{k_3} \|_2)^{1-\theta_3}.\end{aligned}$$

The expressions in the last line are summable, via Hölder's inequality; more exactly, since $\theta_j < 1$,

$$\begin{aligned}\sum_{k_1 \geq 0} 2^{-k_1 M \frac{1-\theta_1}{2}} \big(2^{-k_1 \frac{M}{2(1-\theta_1)}} \| f_{k_1} \|_2\big)^{1-\theta_1} &\lesssim \Big(\sum_{k_1} 2^{-k_1 M \frac{1-\theta_1}{1+\theta_1}}\Big)^{\frac{1+\theta_1}{2}} \Big(\sum_{k_1} 2^{-k_1 \frac{M}{1-\theta_1}} \| f_{k_1} \|_2^2\Big)^{\frac{1-\theta_1}{2}} \\ &\lesssim \| f \cdot \tilde{\chi}_{I_0} \|_2^{1-\theta_1}\end{aligned}$$

for M sufficiently large. Note the implicit constants will depend on θ_1 only. This proves inequality (28). □

2.5. ***The helicoidal method.*** With the intention of bringing to light the ideas behind our proofs, we present the main strategy in a simplified setting. Unfortunately, we cannot avoid the specific terminology, but one should think of the *sizes* as being averages, while the *energies* are L^2 quantities that reflect orthogonality. For estimating the norms $\|\mathrm{BHT}(f,g)\|_s$, we use interpolation results for the trilinear form $\Lambda_{\mathrm{BHT}}(f,g,h) = \langle \mathrm{BHT}(f,g), h\rangle$. In what follows, $\Lambda_{I_0}(f,g,h)$ denotes a space localization of $\Lambda_{\mathrm{BHT}}(f,g,h)$ to the fixed interval I_0. More specifically, it is the form associated to a model operator of BHT as in (19), where the spatial intervals of the tiles lie inside the fixed dyadic interval I_0. Similarly, $\Lambda^n_{I_0}(f,g,h)$ denotes a space localization of the corresponding trilinear form in the multiple vector-valued setting.

The helicoidal method is an iterated induction procedure suitable for proving vector-valued estimates for linear and multilinear operators. We describe the main ideas in the case of the BHT operator, and later on we will indicate the equivalent statements for paraproducts and the Carleson operator. At the heart of our argument lies the following induction statement:

Induction statement. Let $n \geq 0$. We fix I_0 a dyadic interval, and F, G, H' subsets of $\mathbb{R}$ of finite measure. Let $R_1 = (r_1^1, \dots, r_1^n)$, $R_2 = (r_2^1, \dots, r_2^n)$ and $R' = ((r')^1, \dots, (r')^n)$ be n-tuples so that $\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R'} = 1$, while f, g and h are vector-valued functions satisfying

$$\|f(x)\|_{L^{R_1}(\mathcal{W},\mu)} \leq \mathbf{1}_F(x), \quad \|g(x)\|_{L^{R_2}(\mathcal{W},\mu)} \leq \mathbf{1}_G(x) \quad \text{and} \quad \|h(x)\|_{L^{R'}(\mathcal{W},\mu)} \leq \mathbf{1}_{H'}(x).$$

Then we have the following estimate $\mathcal{P}(n)$ for the trilinear form $\Lambda_{I_0}^n$:

$$\left|\Lambda_{I_0}^n(f, g, h)\right| \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{\frac{1}{2}+\frac{\theta_1}{2}-\epsilon} (\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{\frac{1}{2}+\frac{\theta_2}{2}-\epsilon} (\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1}{2}+\frac{\theta_3}{2}-\epsilon} |I_0|$$

for every $0 \leq \theta_1, \theta_2, \theta_3 < 1$, $\theta_1 + \theta_2 + \theta_3 = 1$, satisfying an extra condition $C(R_1, R_2, R')$.

In the local L^2 case the condition $C(R_1, R_2, R')$ is satisfied automatically: that is, the $\mathcal{P}(n)$ statement is true for all $0 \leq \theta_1, \theta_2, \theta_3$ as above. This condition is the main obstruction in obtaining for $\overrightarrow{\text{BHT}}_{\vec{r}}$ the same range of L^p estimates as that of the scalar BHT; in (37) we point out the source of this constraint. Now we present the proofs of the induction statements $\mathcal{P}(0)$ and $\mathcal{P}(n) \Rightarrow \mathcal{P}(n+1)$. Also, for the reader's convenience, we include the $\mathcal{P}(0) \Rightarrow \mathcal{P}(1)$ step.

As we will see later on, the fact that $\mathcal{P}(n)$ implies our Theorems 7 and 8 is based on a standard triple stopping time argument, involving the above localized sizes.

***Check* $\mathcal{P}$(0):** This is the scalar BHT case, with $|f| \leq \mathbf{1}_F$, $|g| \leq \mathbf{1}_G$ and $|h| \leq \mathbf{1}_{H'}$. This situation is well understood, and we have from Proposition 23:

$$|\Lambda_{I_0}(f, g, h)| \lesssim (\widetilde{\text{size}}_{I_0} f)^{\theta_1} (\widetilde{\text{size}}_{I_0} g)^{\theta_2} (\widetilde{\text{size}}_{I_0} h)^{\theta_3} (\text{energy}_{I_0} f)^{1-\theta_1} (\text{energy}_{I_0} g)^{1-\theta_2} (\text{energy}_{I_0} h)^{1-\theta_3}$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$ such that $\theta_1 + \theta_2 + \theta_3 = 1$.

Since we are considering a localized model of BHT, where all the tiles have their spatial intervals I_P lying in I_0, one can refine Lemma 20 by replacing $\text{energy}_{I_0} f$ with $\|f \cdot \tilde{\chi}_{I_0}\|_2$. Noticing that

$$\|f \cdot \tilde{\chi}_{I_0}\|_2 \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{\frac{1}{2}} |I_0|^{\frac{1}{2}}$$

and $|I_0|^{\frac{1-\theta_1}{2}} |I_0|^{\frac{1-\theta_3}{2}} |I_0|^{\frac{1-\theta_3}{2}} = |I_0|$, we obtain the desired $\mathcal{P}(0)$.

***Check* $\mathcal{P}$(0) $\Rightarrow$ $\mathcal{P}$(1).** Assume that

$$\Big(\sum_k |f_k|^{r_1}\Big)^{\frac{1}{r_1}} \leq \mathbf{1}_F, \quad \Big(\sum_k |g_k|^{r_2}\Big)^{\frac{1}{r_2}} \leq \mathbf{1}_G \quad \text{and} \quad \Big(\sum_k |h_k|^{r'}\Big)^{\frac{1}{r'}} \leq \mathbf{1}_{H'}. \tag{30}$$

Given that we know $\mathcal{P}(0)$, we will prove $\mathcal{P}(1)$, given by

$$\Big|\sum_k \Lambda_{I_0}(f_k, g_k, h_k)\Big| \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{\frac{1}{2}+\frac{\theta_1}{2}-\epsilon} (\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{\frac{1}{2}+\frac{\theta_2}{2}-\epsilon} \big(\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'}\big)^{\frac{1}{2}+\frac{\theta_3}{2}-\epsilon} |I_0|$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$, $\theta_1 + \theta_2 + \theta_3 = 1$, satisfying the constraint $C(r_1, r_2, r')$, given by

$$\frac{1+\theta_1}{2} - \frac{1}{r_1} > 0, \quad \frac{1+\theta_2}{2} - \frac{1}{r_2} > 0, \quad \frac{1+\theta_3}{2} - \frac{1}{r'} > 0.$$

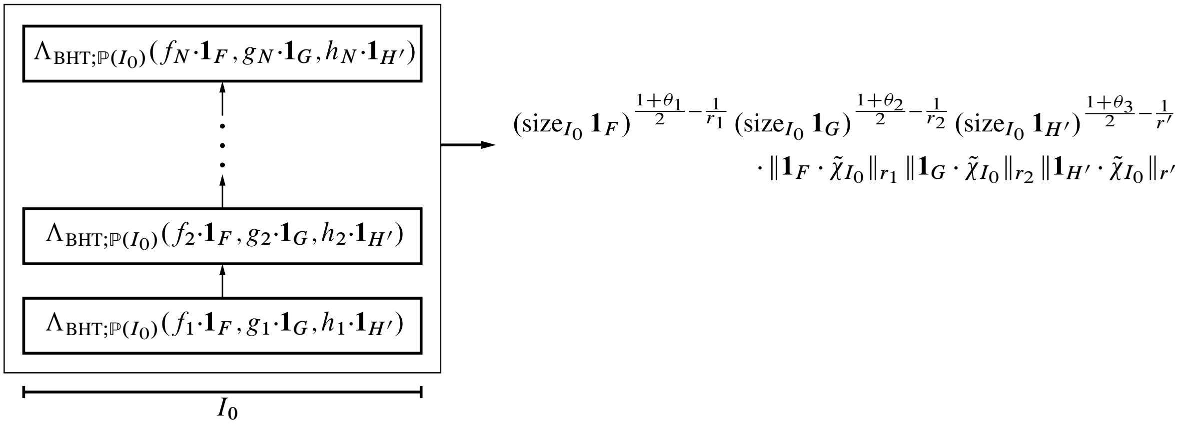

Figure 6. Output of the localization process.

Here an intermediate step is necessary in order to get a finer estimate for each $\Lambda_{I_0}(f_k, g_k, h_k)$. That is, we need to prove

$$\Lambda_{I_0}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'}) \lesssim \|\Lambda_{I_0}\| \, \|f_k \cdot \tilde{\chi}_{I_0}\|_{r_1} \|g_k \cdot \tilde{\chi}_{I_0}\|_{r_2} \|h_k \cdot \tilde{\chi}_{I_0}\|_{r'}, \tag{31}$$

where the operatorial norm is given by

$$\|\Lambda_{I_0}\| = (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}-\epsilon} (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}-\epsilon} (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon}.$$

Once we have such an estimate, we sum in k, use Hölder's inequality and (30) to further estimate (31) by

$$\|\Lambda_{I_0}\| \frac{\|\mathbf{1}_F \cdot \tilde{\chi}_{I_0}\|_{r_1}}{|I_0|^{\frac{1}{r_1}}} \frac{\|\mathbf{1}_G \cdot \tilde{\chi}_{I_0}\|_{r_2}}{|I_0|^{\frac{1}{r_2}}} \frac{\|\mathbf{1}_{H'} \cdot \tilde{\chi}_{I_0}\|_{r'}}{|I_0|^{\frac{1}{r'}}} |I_0|.$$

This is illustrated in Figure 6 and it proves $\mathcal{P}(1)$.

The proof of (31) is a slight modification of the proof of the boundedness of the bilinear Hilbert transform. Using interpolation methods, we can assume that $|f_k| \le \mathbf{1}_{E_1}$, $|g_k| \le \mathbf{1}_{E_2}$, $|h_k| \le \mathbf{1}_{E_3}$. So we need to show

$$\Lambda_{I_0}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'}) \lesssim \|\Lambda_{I_0}\| \, |E_1|^{\alpha_1} |E_2|^{\alpha_2} |E_3|^{\alpha_3},$$

where $(\alpha_1, \alpha_2, \alpha_3)$ is an admissible tuple arbitrarily close to $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$. In order to get the desired expression for $\|\Lambda_{I_0}\|$, we need another stopping time inside I_0. This is illustrated in Figure 7.

Let $I \subseteq I_0$ be a subinterval of I_0. Now we use $\mathcal{P}(0)$ as follows:

$$\begin{aligned}
&\left|\Lambda_I(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'})\right| \\
&\qquad \lesssim (\widetilde{\mathrm{size}}_I(\mathbf{1}_F \cdot \mathbf{1}_{E_1}))^{\frac{1+\theta_1}{2}-\epsilon} (\widetilde{\mathrm{size}}_I(\mathbf{1}_G \cdot \mathbf{1}_{E_2}))^{\frac{1+\theta_2}{2}-\epsilon} (\widetilde{\mathrm{size}}_I(\mathbf{1}_{H'} \cdot \mathbf{1}_{E_3}))^{\frac{1+\theta_3}{2}-\epsilon} |I| \\
&\qquad \lesssim (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_F)^{\frac{1+\theta_1}{2}-\alpha_1-\epsilon} (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_G)^{\frac{1+\theta_2}{2}-\alpha_2-\epsilon} (\widetilde{\mathrm{size}}_{I_0} \mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\alpha_3-\epsilon} \\
&\qquad\qquad \cdot (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_1})^{\alpha_1} (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_2})^{\alpha_2} (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_3})^{\alpha_3} |I|.
\end{aligned}$$

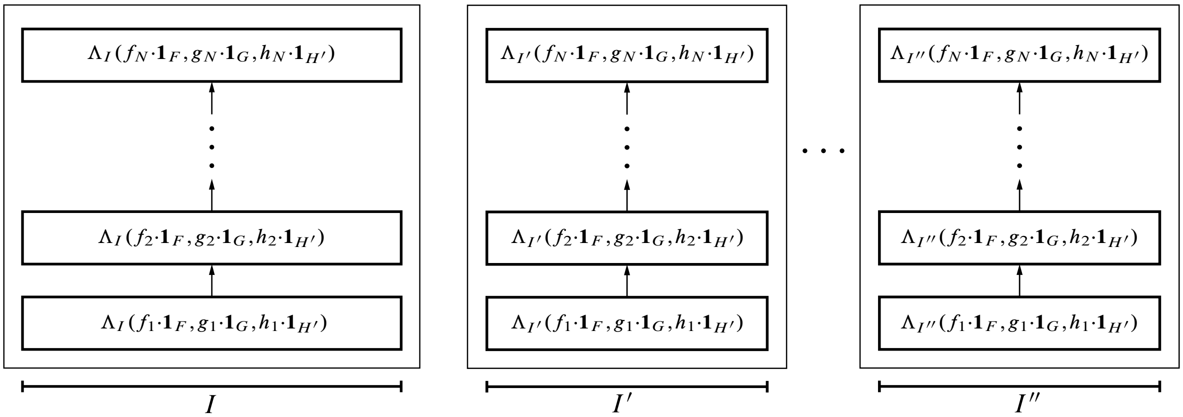

Figure 7. Extra stopping time.

In order to obtain the last inequality, we have to make sure that the exponents

$$\frac{1+\theta_1}{2}-\alpha_1-\epsilon,\quad \frac{1+\theta_2}{2}-\alpha_2-\epsilon,\quad \frac{1+\theta_3}{2}-\alpha_3-\epsilon$$

are all positive, which is always the case in the local L^2 situation. Since $(\alpha_1,\alpha_2,\alpha_3)$ are arbitrarily close to $\left(\frac{1}{r_1},\frac{1}{r_2},\frac{1}{r'}\right)$, this is the origin of the constraint $C(r_1,r_2,r')$ on page 1954.

Summing over the intervals I given by the alluded to triple stopping time over the corresponding averages, we recover $|E_1|^{\alpha_1}|E_2|^{\alpha_2}|E_3|^{\alpha_3}$. We note that the operatorial norm given by interpolation is

$$(\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}-\tilde\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}-\tilde\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\tilde\epsilon},$$

where $\tilde\epsilon$ is slightly larger than the initial ϵ, but the difference between the two is irrelevant.

Check $\mathcal{P}(n)\Rightarrow\mathcal{P}(n+1)$. Lastly, we present the general induction step, in the case of iterated ℓ^p spaces. We have multi-indices $\vec r_1=(r_1^1,\dots,r_1^n)$, $\vec r_2=(r_2^1,\dots,r_2^n)$, $\vec{r'}=((r')^1,\dots,(r')^n)$, and $\|f\|_{\vec r_1}\le\mathbf{1}_F$, $\|g\|_{\vec r_2}\le\mathbf{1}_G$, $\|h\|_{\vec{r'}}\le\mathbf{1}_{H'}$. Then $\mathcal{P}(n)$ is equivalent to

$$\begin{aligned}\left|\Lambda^n_{I_0}(f,g,h)\right| &= \left|\int_{\mathbb{R}}\sum_{\vec l}\text{BHT}_{\mathbb{P}(I_0)}(f_{\vec l},g_{\vec l})(x)\cdot h_{\vec l}(x)\,dx\right|\\ &\lesssim(\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{\frac12+\frac{\theta_1}{2}-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{\frac12+\frac{\theta_2}{2}-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{\frac12+\frac{\theta_3}{2}-\epsilon}|I_0|,\end{aligned}\tag{32}$$

whenever I_0 is a dyadic interval. For $\mathcal{P}(n+1)$ we consider $n+1$ iterated ℓ^p spaces, given by the multi-indices: $\vec R_1=(r_1,\vec r_1)$, $\vec R_2=(r_2,\vec r_2)$ and $\vec R'=(r',\vec{r'})$, while f, g and h are vector-valued functions satisfying

$$\|f\|_{\vec R_1}:=\left(\sum_k\|f_k\|_{\vec r_1}^{r_1}\right)^{\frac{1}{r_1}}\le\mathbf{1}_F,\quad \|g\|_{\vec R_2}:=\left(\sum_k\|g_k\|_{\vec r_2}^{r_2}\right)^{\frac{1}{r_2}}\le\mathbf{1}_G,\quad \|h\|_{\vec R'}:=\left(\sum_k\|h_k\|_{\vec{r'}}^{r'}\right)^{\frac{1}{r'}}\le\mathbf{1}_{H'}.\tag{33}$$

We want a result similar to (32), so we need to estimate

$$\Lambda_{I_0}^{n+1}(f,g,h) := \int_{\mathbb{R}} \sum_k \sum_{\vec{l}} \mathrm{BHT}_{\mathbb{P}(I_0)}(f_{k,\vec{l}}, g_{k,\vec{l}})(x) \cdot h_{k,\vec{l}}(x)\, dx = \sum_k \Lambda_{I_0}^n(f_k, g_k, h_k).$$

We can't directly apply $\mathcal{P}(n)$, and instead we will need the following result, similar to (31):

$$\left|\Lambda_{I_0}^n(f_k, g_k, h_k)\right| \lesssim \|\Lambda_{I_0}^n\| \, \|f_k \cdot \tilde{\chi}_{I_0}\|_{r_1} \|g_k \cdot \tilde{\chi}_{I_0}\|_{r_2} \|h_k \cdot \tilde{\chi}_{I_0}\|_{r'}, \tag{34}$$

where $\|\Lambda_{I_0}^n\| = (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}-\epsilon} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}-\epsilon} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon}$. Once we have such a result, $\mathcal{P}(n+1)$ follows easily by Hölder, exactly as before.

We will prove (34) by using restricted-type interpolation. Instead of estimating the trilinear form $\Lambda_{I_0}^n$, we will deal with

$$\Lambda_{I_0}^{n,F,G,H'}(f_k, g_k, h_k) := \Lambda_{I_0}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'}). \tag{35}$$

This is natural since condition (33) implies that the functions f_k are supported on F, and similarly the functions g_k are supported on G and h_k on H'. By interpolation theory, we can assume that

$$\|f_k\|_{\vec{r}_1} \le \mathbf{1}_{E_1}, \quad \|g_k\|_{\vec{r}_2} \le \mathbf{1}_{E_2}, \quad \text{and} \quad \|h_k\|_{\vec{r'}} \le \mathbf{1}_{E_3},$$

and it suffices to prove

$$\left|\Lambda_{I_0}^{n,F,G,H'}(f_k, g_k, h_k)\right| \lesssim \|\Lambda_{I_0}^n\| \, |E_1|^{\alpha_1} |E_2|^{\alpha_2} |E_3|^{\alpha_3} \tag{36}$$

for $(\alpha_1, \alpha_2, \alpha_3)$ in a small neighborhood of $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$. Similarly to the case $\mathcal{P}(0) \Rightarrow \mathcal{P}(1)$, we will have a stopping time inside I_0, so in fact we need to estimate $\Lambda_I^{n,F,G,H'}(f_k, g_k, h_k)$ for some $I \subseteq I_0$. It is here that we use hypothesis $\mathcal{P}(n)$:

$$\left|\Lambda_I^{n,F,G,H'}(f_k, g_k, h_k)\right| = \left|\Lambda_I^n(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'})\right|,$$

with $\|f_k \cdot \mathbf{1}_F\|_{\vec{r}_1} \le \mathbf{1}_{F \cap E_1}$, $\|g_k \cdot \mathbf{1}_G\|_{\vec{r}_2} \le \mathbf{1}_{G \cap E_2}$ and $\|h_k \cdot \mathbf{1}_{H'}\|_{\vec{r'}} \le \mathbf{1}_{H' \cap E_3}$. More precisely,

$$\begin{aligned}
&\left|\Lambda_I^{n,F,G,H'}(f_k, g_k, h_k)\right| \\
&\quad \lesssim (\widetilde{\mathrm{size}}_I(\mathbf{1}_F \cdot \mathbf{1}_{E_1}))^{\frac{1}{2}+\frac{\theta_1}{2}-\epsilon} (\widetilde{\mathrm{size}}_I(\mathbf{1}_G \cdot \mathbf{1}_{E_2}))^{\frac{1}{2}+\frac{\theta_2}{2}-\epsilon} (\widetilde{\mathrm{size}}_I(\mathbf{1}_{H'} \cdot \mathbf{1}_{E_3}))^{\frac{1}{2}+\frac{\theta_3}{2}-\epsilon} |I| \\
&\quad \lesssim (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_F)^{\frac{1}{2}+\frac{\theta_1}{2}-\alpha_1-\epsilon} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_G)^{\frac{1}{2}+\frac{\theta_2}{2}-\alpha_2-\epsilon} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1}{2}+\frac{\theta_3}{2}-\alpha_3-\epsilon} \\
&\qquad \cdot (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_1})^{\alpha_1} (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_2})^{\alpha_2} (\widetilde{\mathrm{size}}_I \mathbf{1}_{E_3})^{\alpha_3} |I|
\end{aligned}$$

for $(\alpha_1, \alpha_2, \alpha_3)$ in a neighborhood of $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$. Due to the stopping time, which is performed with respect to the three sizes, we know the expressions $(\widetilde{\mathrm{size}}_I \mathbf{1}_{E_1})^{\alpha_1}$ add up to $|E_1|^{\alpha_1}$ and it is similar for the sizes of $\mathbf{1}_{E_2}$ and $\mathbf{1}_{E_3}$. Interpolating, we get the desired (36). From the above equation, we can see why the operatorial norm has the form

$$\|\Lambda_{I_0}^n\| = (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}-\tilde{\epsilon}} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}-\tilde{\epsilon}} (\widetilde{\mathrm{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\tilde{\epsilon}}.$$

The $\tilde{\epsilon}$ (which is a slight modification on the ϵ in the $\mathcal{P}(n)$ statement), appears as an interpolation error; moreover, the conditions

$$\frac{1+\theta_1}{2}-\frac{1}{r_1}>0, \quad \frac{1+\theta_2}{2}-\frac{1}{r_2}>0, \quad \frac{1+\theta_3}{2}-\frac{1}{r'}>0 \tag{37}$$

are necessary, and they imply the constraint $C(R_1, R_2, R')$. This ends the proof of the induction step.

The same method applies in the case of paraproducts. The difference here is that the energies are L^1 quantities, and for that reason we don't have any extra assumptions; the range of the multiple vector-valued extensions is the same as that of the paraproducts. The model operator for paraproducts Π corresponds to a "rank 0" family of tritiles; that is, once we know the spatial interval I_P, there is no other degree of freedom and the frequency intervals are $\left[1/|I_P|, 2/|I_P|\right]$ or $\left[0, 1/|I_P|\right]$. The exact definitions will be introduced in Section 4.

Induction statement (paraproducts case). Under the same assumptions as in the induction statement on page 1954, the localized trilinear form for paraproducts satisfies $\mathcal{P}(n)$, given by

$$\left|\Lambda^n_{I_0}(f, g, h)\right| \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{1-\epsilon}\,(\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{1-\epsilon}\,(\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{1-\epsilon}\,|I_0|,$$

provided

$$\|f(x)\|_{L^{R_1}(\mathcal{W},\mu)} \le \mathbf{1}_F(x), \quad \|g(x)\|_{L^{R_2}(\mathcal{W},\mu)} \le \mathbf{1}_G(x) \quad \text{and} \quad \|h(x)\|_{L^{R'}(\mathcal{W},\mu)} \le \mathbf{1}_{H'}(x).$$

Finally, we want to point out that the helicoidal method applies equally in the case of (sub)linear operators. One last example is that of the Carleson operator

$$C_{\mathbb{R}} f(x) = \sup_N \left| \int_{\xi<N} \hat{f}(\xi) e^{2\pi i x\xi}\, d\xi \right|$$

for which UMD-valued extensions are already known from the work of Hytönen and Lacey [2013].

Demeter and Silva [2015] gave an alternative proof for ℓ^2-valued inequalities for the Carleson operator. In fact, they present a new principle, built around ideas from [Bateman and Thiele 2013], for dealing with ℓ^2-valued inequalities for sublinear operators which are not of Calderón–Zygmund type.

We do not present all the details here, but the essential statement for proving multiple vector-valued inequalities for the Carleson operator, using the helicoidal method, is the following:

Induction statement (Carleson operator). Under the same assumptions as in the induction statement on page 1954, the localized bilinear form for the discretized Carleson operator satisfies $\mathcal{P}(n)$, given by

$$\left|\Lambda^n_{C(I_0)}(f, g)\right| \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_F)^{1-\epsilon}\,(\widetilde{\text{size}}_{I_0}\mathbf{1}_G)^{1-\epsilon}\,|I_0|,$$

provided that

$$\|f(x)\|_{L^{R_1}(\mathcal{W},\mu)} \le \mathbf{1}_F(x) \quad \text{and} \quad \|g(x)\|_{L^{R_2}(\mathcal{W},\mu)} \le \mathbf{1}_G(x).$$

Comparing the main statements of the above three examples, we can see from the exponents of the sizes that the range of L^p estimates for the vector-valued Carleson operator and for the vector-valued paraproduct Π will coincide with the range of the scalar operator. However, for BHT things are more complicated.

3. Multiple vector-valued estimates for BHT

In this section we describe the detailed proof of our Theorems 7 and 8.

3.1. ***Estimates for localized* BHT.** Here we assume that F, G and H' are fixed subsets of $\mathbb{R}$ of finite measure and I_0 is a fixed dyadic interval. We are interested in finding estimates for the bilinear operator

$$\mathrm{BHT}_{I_0}^{F,G,H'}(f,g)(x) := \sum_{P \in \mathbb{P}(I_0)} \frac{1}{|I_P|^{\frac{1}{2}}} \langle f \cdot \mathbf{1}_F, \phi_{P_1}^1 \rangle \langle g \cdot \mathbf{1}_G, \phi_{P_2}^2 \rangle \phi_{P_3}^3(x) \mathbf{1}_{H'}(x).$$

In doing so, we first study the associated trilinear form

$$\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}^{F,G,H'}(f,g,h) := \sum_{P \in \mathbb{P}(I_0)} \frac{1}{|I_P|^{\frac{1}{2}}} \langle f \cdot \mathbf{1}_F, \phi_{P_1}^1 \rangle \langle g \cdot \mathbf{1}_G, \phi_{P_2}^2 \rangle \langle h \cdot \mathbf{1}_{H'}, \phi_{P_3}^3 \rangle.$$

While this operator satisfies the same estimates as the bilinear Hilbert transform, the localization to the sets F, G and H', and the restriction to the tiles in $\mathbb{P}(I_0)$ will bring some extra decay. First we prove a result in the "local L^2 case", when $\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'} < \frac{1}{2}$. In this situation the proof is simpler, because we are employing "energies", which are L^2 expressions, and they can easily be related to L^{r_i} averages when $r_i \geq 2$.

Proposition 41 (the case $r_1, r_2, r' > 2$). *Let $\mathbb{P}$ be a family of tritiles, I_0 a dyadic interval and $F, G, H' \subset \mathbb{R}$ sets of finite measure. Then one can find positive numbers a_1, a_2 and a_3 so that*

$$\left|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}^{F,G,H'}(f,g,h)\right| \lesssim (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_F)^{a_1} (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_G)^{a_2} (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_{H'})^{a_3} \|f \cdot \tilde{\chi}_{I_0}\|_{r_1} \|g \cdot \tilde{\chi}_{I_0}\|_{r_2} \|h \cdot \tilde{\chi}_{I_0}\|_{r'}. \quad (38)$$

We can choose $a_j = 1 - \frac{2}{r_j} - \epsilon > 0$ for a very small $\epsilon > 0$.

Proof. In this case we are proving restricted-type estimates by applying directly Proposition 23: let E_1, E_2, E_3 be sets of finite measure, and $|f| \leq \mathbf{1}_{E_1},\ |g| \leq \mathbf{1}_{E_2},\ |h| \leq \mathbf{1}_{E_3}$. We have

$$\Lambda_{\mathrm{BHT}}(f \cdot \mathbf{1}_F, g \cdot \mathbf{1}_G, h \cdot \mathbf{1}_{H'}) \lesssim (\mathrm{size}_{\mathbb{P}(I_0)}(f \cdot \mathbf{1}_F))^{\theta_1} (\mathrm{size}_{\mathbb{P}(I_0)}(g \cdot \mathbf{1}_G))^{\theta_2} (\mathrm{size}_{\mathbb{P}(I_0)}(h \cdot \mathbf{1}_{H'}))^{\theta_3} \cdot (\mathrm{energy}(f \cdot \mathbf{1}_F))^{1-\theta_1} (\mathrm{energy}(g \cdot \mathbf{1}_G))^{1-\theta_2} (\mathrm{energy}(h \cdot \mathbf{1}_{H'}))^{1-\theta_3} \quad (39)$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$ such that $\theta_1 + \theta_2 + \theta_3 = 1$. Recall that the sizes can be estimated by

$$\mathrm{size}_{\mathbb{P}(I_0)}(f \cdot \mathbf{1}_F) \lesssim \sup_{P \in \mathbb{P}(I_0)} \frac{1}{|I_P|} \int \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_{I_P}^M \, dx,$$

where M can be chosen as large as we wish. Then we observe that if E_1 is supported away from I_0, the sizes will decay fast, giving the desired $\|f \cdot \tilde{\chi}_{I_0}\|_{r_1}$ on the right-hand side. It is similar for E_2 and E_3. For this reason, we can assume that the sets E_1, E_2, E_3 are supported on $5I_0$ and then we will need to show only that

$$|\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}(f \cdot \mathbf{1}_F, g \cdot \mathbf{1}_G, h \cdot \mathbf{1}_{H'})| \lesssim (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_F)^{a_1} (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_G)^{a_2} (\mathrm{size}_{\mathbb{P}(I_0)} \mathbf{1}_{H'})^{a_3} \|f\|_{r_1} \|g\|_{r_2} \|h\|_{r'}.$$

We are using the energies precisely for estimating the norms of f, g and h, so the sizes are playing the role of a constant here. As we have seen in Lemma 24, the energies are bounded by L^2 norms, so from (39), we have

$$\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h) \lesssim (\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_F)^{\theta_1}(\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_G)^{\theta_2}(\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_{H'})^{\theta_3}\,|E_1|^{\frac{1-\theta_1}{2}}\,|E_2|^{\frac{1-\theta_2}{2}}\,|E_3|^{\frac{1-\theta_3}{2}}.$$

By varying θ_1, θ_2 and θ_3, we see that these restricted-type estimates are true in a very small neighborhood of $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$, and the interpolation, Theorem 28, yields strong-type estimates. Note that the constant in this case is

$$(\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_F)^{\theta_1}(\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_G)^{\theta_2}(\mathrm{size}_{\mathbb{P}(I_0)}\,\mathbf{1}_{H'})^{\theta_3},$$

which depends on the functions $\mathbf{1}_F, \mathbf{1}_G, \mathbf{1}_{H'}$, the fixed interval I_0, the values of θ_1, θ_2, and θ_3, but not on the functions f, g, h. □

Now we deal with the general Banach triangle case, where $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$ is an admissible tuple satisfying

$$0 < \frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'} < 1.$$

The proof is going to be more complicated because we will need to use the sizes as well for reconstructing the norms of f, g, h. In addition, we will also need to use the sizes of $\mathbf{1}_F, \mathbf{1}_G$ and $\mathbf{1}_{H'}$ later on.

Proposition 42. *Let F, G and H' be as above and let $\mathbb{P}(I_0)$ be a family of tritiles localized to the dyadic interval I_0. Then there exist positive numbers a_1, a_2 and a_3 so that*

$$\left|\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h)\right| \lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{a_1}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{a_2}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{a_3}\|f\cdot\tilde{\chi}_{I_0}\|_{r_1}\|g\cdot\tilde{\chi}_{I_0}\|_{r_2}\|h\cdot\tilde{\chi}_{I_0}\|_{r'}, \quad (40)$$

where $\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r'}=1$. In fact, for $\epsilon>0$ small enough,

$$a_1 = \frac{1+\theta_1}{2} - \frac{1}{r_1} - \epsilon, \quad a_2 = \frac{1+\theta_2}{2} - \frac{1}{r_2} - \epsilon, \quad a_3 = \frac{1+\theta_3}{2} - \frac{1}{r'} - \epsilon, \quad (41)$$

where $\theta_1, \theta_2, \theta_3$ are so that $0 \le \theta_1, \theta_2, \theta_3 < 1$, $\theta_1+\theta_2+\theta_3=1$, and the expressions in (41) *are positive.*

Proof. In this case, we will use the interpolation, Theorem 32, and for this reason we cannot obtain directly the expression in the right-hand side of (40), which represents *localized L^p* norms. However, as we will see soon, it will be enough to prove that $\Lambda_{\mathrm{BHT};\mathbb{P}(I_0)}$ is of generalized restricted type $\alpha=(\alpha_1,\alpha_2,\alpha_3)$ for α in a small neighborhood of $\left(\frac{1}{r_1}, \frac{1}{r_2}, \frac{1}{r'}\right)$. Then the result in (40) will be a consequence of the fast decay of the wave packets away from I_0.

We start with sets of finite measure E_1, E_2, E_3 and define $\widetilde{\Omega}$ to be the exceptional set

$$\widetilde{\Omega} := \left\{x : \mathcal{M}(\mathbf{1}_{E_1}) > C\frac{|E_1|}{|E_3|}\right\} \cup \left\{x : \mathcal{M}(\mathbf{1}_{E_2}) > C\frac{|E_2|}{|E_3|}\right\}.$$

Let $E_3' := E_3 \setminus \widetilde{\Omega}$. We want to prove that (40) holds for any functions f, g, h so that $|f| \le \mathbf{1}_{E_1}$, $|g| \le \mathbf{1}_{E_2}$, and $|h| \le \mathbf{1}_{E_3'}$. For simplicity, we assume that $1+\mathrm{dist}(I_P, \widetilde{\Omega}^c)/|I_P| \sim 2^d$ for every tile $P \in \mathbb{P}(I_0)$.

Equivalently, we could decompose the collection of tiles into subcollections for which this property holds for all $d \geq 0$. In the end, however, the estimate (40) will be independent of such a decomposition.

With the above assumption, for every $P \in \mathbb{P}(I_0)$, we have

$$\frac{1}{|I_P|}\int_{\mathbb{R}} \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_{I_P}^M \, dx \lesssim 2^d \frac{|E_1|}{|E_3|} \quad \text{and} \quad \frac{1}{|I_P|}\int_{\mathbb{R}} \mathbf{1}_{E_2} \cdot \mathbf{1}_G \cdot \tilde{\chi}_{I_P}^M \, dx \lesssim 2^d \frac{|E_2|}{|E_3|}.$$

This is important because now we can perform a stopping time which will allow us to estimate the "sizes" of the functions $\mathbf{1}_{E_j}$. For each of the functions $\mathbf{1}_F \cdot \mathbf{1}_{E_1}$, $\mathbf{1}_G \cdot \mathbf{1}_{E_2}$ and $\mathbf{1}_{H'} \cdot \mathbf{1}_{E_3'}$, we will be looking for maximal dyadic intervals J which are maximizers for

$$\sup_{\substack{J \subseteq I_0 \\ \exists P \in \mathbb{P}(I_0),\, I_P \subseteq J}} \frac{1}{|J|}\int_{\mathbb{R}} \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_J^M \, dx. \tag{42}$$

This is the reason we introduced the new size in Definition 21.

The selection of the intervals and tiles is described in more detail in Section 3.2, so here we only sketch this process.

We start with the largest possible value $2^{-l_1} \lesssim 2^d |E_1|/|E_2|$ and define $\mathcal{I}_{l_1}$ to be the collection of maximal dyadic intervals I with the property that it contains some $I_P \in \mathbb{P}(I_0)$ which is not contained in any of the intervals previously selected, and I also has the property that

$$2^{-l_1-1} \leq \frac{1}{|I|}\int_{\mathbb{R}} \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_I^M \, dx \leq 2^{-l_1}.$$

Then for each $I \in \mathcal{I}_{l_1}$ we find the relevant tiles P with $I_P \subseteq I$, and move them into $\mathbb{P}(I)$. Afterwards we restart the algorithm for the collection $\mathbb{P}(I_0) \setminus \bigcup_{I \in \mathcal{I}_{l_1}} \mathbb{P}(I)$.

The algorithm continues by decreasing 2^{-l_1} until all tiles in $\mathbb{P}(I_0)$ are exhausted. In this way, for any l_1 and any $I \in \mathcal{I}_{l_1}$, we have $\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F) \sim 2^{-l_1}$. Similarly we define the collections of dyadic intervals $\mathcal{I}_{l_2}$ associated with the functions $\mathbf{1}_{E_2} \cdot \mathbf{1}_G$ as long as $2^{-l_2} \lesssim 2^d |E_2|/|E_3|$.

For the third component, the collections $\mathcal{I}_{l_3}$ are nonempty as long as $2^{-n_3} \lesssim 2^{-\tilde{M}d}$, and in that case, for any $I \in \mathcal{I}_{l_3}$, we have $\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{H'} \cdot \mathbf{1}_{E_3'}) \sim 2^{-n_3}$. The extra decay is due to the fact that E_3' is actually supported on $\tilde{\Omega}^c$.

Given l_1, l_2, l_3 as above, we define $\mathcal{J}^{l_1,l_2,l_3} := \mathcal{I}_{l_1} \cap \mathcal{I}_{l_2} \cap \mathcal{I}_{l_3}$. This is also going to be a collection of dyadic intervals, and any tile in $\mathbb{P}(I_0)$ will be contained in some $\mathbb{P}(I)$, with $I \in \mathcal{J}^{l_1,l_2,l_3}$. In fact, these collections depend on the parameter d as well, which controls the distance from the exceptional set. We have

$$\mathbb{P}(I_0) = \bigcup_d \bigcup_{l_1,l_2,l_3} \bigcup_{I \in \mathcal{J}_d^{l_1,l_2,l_3}} \mathbb{P}(I),$$

but we suppress the dependency on d in the notation. Thus

$$\Lambda_{\text{BHT};\mathbb{P}(I_0)}^{F,G,H'}(f,g,h) = \sum_{l_1,l_2,l_3} \sum_{I \in \mathcal{J}^{l_1,l_2,l_3}} \Lambda_{\text{BHT};\mathbb{P}(I)}^{F,G,H'}(f,g,h). \tag{43}$$

Every $\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I)}(f,g,h)$ is going to be estimated by Lemma 40:

$$\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I)}(f,g,h) \lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1}\cdot\mathbf{1}_F))^{\theta_1}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_2}\cdot\mathbf{1}_G))^{\theta_2}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_3'}\cdot\mathbf{1}_{H'}))^{\theta_3} \cdot \|\mathbf{1}_{E_1}\cdot\mathbf{1}_F\cdot\tilde{\chi}_I\|_2^{1-\theta_1}\|\mathbf{1}_{E_2}\cdot\mathbf{1}_G\cdot\tilde{\chi}_I\|_2^{1-\theta_2}\|\mathbf{1}_{E_3'}\cdot\mathbf{1}_{H'}\cdot\tilde{\chi}_I\|_2^{1-\theta_3}.$$

For the particular function $\mathbf{1}_{E_1}\cdot\mathbf{1}_F$ and an interval $I\in\mathcal{J}^{l_1,l_2,l_3}$, we have

$$\Big(\int_{\mathbb{R}}\mathbf{1}_{E_1}\cdot\mathbf{1}_F\cdot\tilde{\chi}_I^M\,dx\Big)^{\frac{1}{2}} \lesssim 2^{-\frac{l_1}{2}}|I|^{\frac{1}{2}} \lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1}\cdot\mathbf{1}_F))^{\frac{1}{2}}|I|^{\frac{1}{2}}.$$

In this way, as long as

$$\frac{1+\theta_1}{2}-\frac{1}{r_1}>0,\quad \frac{1+\theta_2}{2}-\frac{1}{r_2}>0,\quad \frac{1+\theta_3}{2}-\frac{1}{r'}>0, \tag{44}$$

we can estimate $\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I_0)}(f,g,h)$ as

$$\begin{aligned}\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I_0)}&(f,g,h)\\ &\lesssim \sum_{l_1,l_2,l_3}\sum_{I\in\mathcal{J}^{l_1,l_2,l_3}}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1}\cdot\mathbf{1}_F))^{\theta_1}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_2}\cdot\mathbf{1}_G))^{\theta_2}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_3'}\cdot\mathbf{1}_{H'}))^{\theta_3}\\ &\qquad\Big(\frac{1}{|I|}\int_{\mathbb{R}}\mathbf{1}_{E_1}\cdot\mathbf{1}_F\cdot\tilde{\chi}_I^M\,dx\Big)^{\frac{1-\theta_1}{2}}\Big(\frac{1}{|I|}\int_{\mathbb{R}}\mathbf{1}_{E_2}\cdot\mathbf{1}_G\cdot\tilde{\chi}_I^M\,dx\Big)^{\frac{1-\theta_2}{2}}\Big(\frac{1}{|I|}\int_{\mathbb{R}}\mathbf{1}_{E_3'}\cdot\mathbf{1}_{H'}\cdot\tilde{\chi}_I^M\,dx\Big)^{\frac{1-\theta_3}{2}}|I|\\ &\lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon}\\ &\qquad\cdot\sum_{l_1,l_2,l_3}\sum_{I\in\mathcal{J}^{l_1,l_2,l_3}}2^{-\frac{l_1}{r_1}}2^{-\frac{l_2}{r_2}}2^{-l_3(\frac{1}{r'}+\epsilon)}|I|.\end{aligned} \tag{45}$$

The quantity

$$(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon}$$

is going to represent the operatorial norm $\|\Lambda^{F,G,H'}_{\mathrm{BHT};\mathbb{P}(I_0)}\|$ associated to the trilinear form $\Lambda^{F,G,H}_{\mathrm{BHT};\mathbb{P}(I_0)}$, as seen in (40).

We are left with estimating $\sum_{I\in\mathcal{J}^{l_1,l_2,l_3}}|I|$, which can be realized in three different ways; for example,

$$\sum_{I\in\mathcal{J}^{l_1,l_2,l_3}}|I| \le \sum_{I\in\mathcal{J}_{l_1}}|I| = \Big\|\sum_{I\in\mathcal{J}_{l_1}}\mathbf{1}_I\Big\|_{1,\infty} \lesssim \Big\|\sum_{I\in\mathcal{J}_{l_1}}2^{l_1}\mathcal{M}(\mathbf{1}_{E_1})\cdot\mathbf{1}_I\Big\|_{1,\infty} \lesssim 2^{n_1}|E_1|.$$

For this reason, whenever $0\le\alpha_j\le 1$, with $\alpha_1+\alpha_2+\alpha_3=1$, we have

$$\sum_{I\in\mathcal{J}^{l_1,l_2,l_3}}|I| \lesssim (2^{l_1}|E_1|)^{\alpha_1}(2^{l_2}|E_2|)^{\alpha_2}(2^{l_3}|E_3'|)^{\alpha_3}.$$

This yields

$$\sum_{l_1,l_2,l_3}\sum_{I\in\mathcal{I}^{l_1,l_2,l_3}} 2^{-\frac{l_1}{r_1}}2^{-\frac{l_2}{r_2}}2^{-l_3(\frac{1}{r'}+\epsilon)}|I|$$
$$\lesssim \sum_{l_1,l_2,l_3} 2^{-l_1(\frac{1}{r_1}-\alpha_1)}2^{-l_2(\frac{1}{r_2}-\alpha_2)}2^{-l_3(\frac{1}{r'}+\epsilon-\alpha_1)}|E_1|^{\alpha_1}|E_2|^{\alpha_2}|E_3|^{\alpha_3}$$
$$\lesssim \left(2^d\frac{|E_1|}{|E_3|}\right)^{\frac{1}{r_1}-\alpha_1}\left(2^d\frac{|E_2|}{|E_3|}\right)^{\frac{1}{r_2}-\alpha_2}(2^{-\widetilde{M}d})^{(\frac{1}{r'}+\epsilon-\alpha_3)}|E_1|^{\alpha_1}|E_2|^{\alpha_2}|E_3|^{\alpha_3}$$
$$\lesssim 2^{-100d}|E_1|^{\frac{1}{r_1}}|E_2|^{\frac{1}{r_2}}|E_3|^{\frac{1}{r'}}.$$

Summing over d, this proves (40) in the particular case of characteristic functions. Upon interpolating, we lose an ϵ-power of $\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F$ and $\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G$ respectively, to get

$$\left|\Lambda^{F,G,H'}_{BHT;\mathbb{P}(I_0)}(f,g,h)\right|\lesssim(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{a_1}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{a_2}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{a_3}\|f\cdot\tilde{\chi}_{I_0}\|_{r_1}\|g\cdot\tilde{\chi}_{I_0}\|_{r_2}\|h\cdot\tilde{\chi}_{I_0}\|_{r'}.$$

We note that the "weights" $\tilde{\chi}_{I_0}$ will not affect the interpolation process; once we have an inequality that holds for characteristic functions of finite sets, interpolation implies a similar result in full generality.

The exponents a_1, a_2 and a_3 can be described as

$$a_1=\frac{1+\theta_1}{2}-\frac{1}{r_1}-\epsilon,\quad a_2=\frac{1+\theta_2}{2}-\frac{1}{r_2}-\epsilon,\quad a_3=\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon$$

for some sufficiently small ϵ, and for $0\le\theta_1,\theta_2,\theta_3<1$, satisfying $\theta_1+\theta_2+\theta_3=1$, that will be chosen later. □

Corollary 43 (the case $r=1$). *Let $1<r_1,r_2<\infty$ be such that $\frac{1}{r_1}+\frac{1}{r_2}=1$, and θ_1,θ_2 satisfy $\frac{1}{2}(1+\theta_1)>\frac{1}{r_1}$ and $\frac{1}{2}(1+\theta_2)>\frac{1}{r_2}$. Then*

$$\|\text{BHT}^{F,G,H'}_{\mathbb{P}(I_0)}(f,g)\|_1$$
$$\lesssim(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{r_1}-\epsilon}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r_2}-\epsilon}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\epsilon}\|f\cdot\tilde{\chi}_{I_0}\|_{r_1}\|g\cdot\tilde{\chi}_{I_0}\|_{r_2}.$$

Proof. A careful inspection of (45) shows that one can choose any triple $(\beta_1,\beta_2,\beta_3)$ with $\beta_1+\beta_2+\beta_3=1$, even with $\beta_3\le 0$, in the place of $\left(\frac{1}{r_1},\frac{1}{r_2},\frac{1}{r'}\right)$. In this case we get

$$\left|\Lambda^{F,G,H'}_{\text{BHT};\mathbb{P}(I_0)}(f,g,h)\right|$$
$$\lesssim(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\beta_1}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\beta_2}(\widetilde{\text{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\epsilon}|E_1|^{\beta_1}|E_2|^{\beta_2}|E_3|^{\beta_3}.$$

The restrictions are that $\beta_j<\frac{1}{2}(1+\theta_j)$, which works well for very small or negative values of β_3. Interpolating between tuples $(\beta_1,\beta_2,\beta_3)$ that lie in a small open neighborhood of $\left(\frac{1}{r_1},\frac{1}{r_2},0\right)$, we get the conclusion. In this case, the interpolation is used for estimating the L^1 norm of the operator, and not the trilinear form $\Lambda^{F,G,H'}_{\text{BHT};\mathbb{P}(I_0)}$. □

3.2. ***Proof of Theorem 7.*** Recall that the vector-valued BHT is defined by

$$\mathrm{BHT}(f,g)(x,w)=\int_{\mathbb{R}} f(x-t,w)g(x+t,w)\,\frac{dt}{t}=\mathrm{BHT}(f_w,g_w)(x).$$

Then the trilinear form associated with it is

$$\Lambda_{\overrightarrow{\mathrm{BHT}}}(f,g,h)=\int_{\mathbb{R}}\int_{\mathcal{W}}\mathrm{BHT}(f,g)(x,w)h(x,w)\,d\mu(w)\,dx.$$

First we prove generalized restricted-type estimates for $\Lambda_{\overrightarrow{\mathrm{BHT}}}(f,g,h)$, and the general result will follow from the vector-valued interpolation result presented in Proposition 38. Let F, G and H be sets of finite measure. In what follows, we will construct a major subset $H'\subseteq H$ and show

$$|\Lambda_{\overrightarrow{\mathrm{BHT}};\mathbb{P}}(f,g,h)|\lesssim |F|^{\alpha_1}|G|^{\alpha_2}|H|^{\alpha_3} \tag{46}$$

whenever $\|f(x,\cdot)\|_{L^{r_1}(\mathcal{W},\mu)}\le \mathbf{1}_F(x)$, $\|g(x,\cdot)\|_{L^{r_2}(\mathcal{W},\mu)}\le \mathbf{1}_G(x)$ and $\|h(x,\cdot)\|_{L^{r'}(\mathcal{W},\mu)}\le \mathbf{1}_{H'}(x)$. For simplicity, assume $|H|=1$. The exceptional set is defined as

$$\Omega:=\big\{x:\mathcal{M}(\mathbf{1}_F)>C|F|\big\}\cup\big\{x:\mathcal{M}(\mathbf{1}_G)>C|G|\big\}.$$

Because of the $L^1\to L^{1,\infty}$ boundedness of the maximal operator, for a constant C large enough, we have $|\Omega|\ll 1$.

We partition the collection of tritiles according to the scaled distance from the exceptional set

$$\mathbb{P}^d=\left\{P\in\mathbb{P}:1+\frac{\mathrm{dist}(I_P,\Omega^c)}{|I_P|}\sim 2^d\right\}$$

and we will prove estimates equivalent to (46) for the family $\mathbb{P}^d$, with an extra 2^{-10d} decay:

$$\big|\Lambda_{\overrightarrow{\mathrm{BHT}};\mathbb{P}^d}(f,g,h)\big|\lesssim 2^{-10d}|F|^{\frac{1}{p}}|G|^{\frac{1}{q}}|H|^{\frac{1}{s'}}. \tag{47}$$

We suppress the d-dependency for the moment, but all the subcollections $\mathcal{I}_j^{n_j}$ and $\mathcal{I}^{n_1,n_2,n_3}$ will actually depend on this parameter. At the very end we sum in d, and use interpolation, so that the final estimate depends only on the fixed interval I_0, and the fixed sets F,G,H'.

Now we construct a collection $\{\mathcal{I}_1^{n_1}\}_{n_1\ge\bar{n}_1}$ of relevant dyadic intervals, according to the concentration of $\mathbf{1}_F$:

- Start with $\bar{n}_1$ such that $2^{-\bar{n}_1}\sim 2^d|F|$ and let $\mathbb{P}'_{\bar{n}_1-1}=\mathbb{P}$ (here $\mathbb{P}'_{n_1}$ will play the role of *stock*, or the collection of available tiles).
- Define $\mathcal{I}_1^{\bar{n}_1}$ to be the collection of maximal dyadic intervals I with the property that there exists at least one tile $P\in\mathbb{P}'_{\bar{n}_1}$ with $I_P\subseteq I$ and

$$\frac{1}{|I|}\int \mathbf{1}_F\cdot\tilde{\chi}_I^M\,dx\sim 2^{-\bar{n}_1}. \tag{48}$$

- For every such interval I, let $\mathbb{P}_{\bar{n}_1}(I)$ be the collection of tiles $P\in\mathbb{P}'_{\bar{n}_1}$ with the property that $I_P\subseteq I$.
- Set $\mathbb{P}'_{\bar{n}_1}=\mathbb{P}\setminus\bigcup_{I\in\mathcal{I}_1^{\bar{n}_1}}\mathbb{P}_{\bar{n}_1}(I)$.

- Repeat the procedure for all $n_1 \geq \bar{n}_1$. Let $\mathcal{J}_1^{n_1}$ denote the collection of maximal dyadic intervals which contain a time interval I_P for some $P \in \mathbb{P}'_{n_1-1}$ (which was not selected previously) and such that

$$2^{-n_1-1} \leq \frac{1}{|I|} \int \mathbf{1}_F \cdot \tilde{\chi}_I^M \, dx < 2^{-n_1}.$$

- As before, $\mathbb{P}_{n_1}(I) := \{P \in \mathbb{P}'_{n_1} : I_P \subseteq I\}$.
- Set $\mathbb{P}'_{n_1} = \mathbb{P}_{n_1-1} \setminus \bigcup_{I \in \mathcal{J}_1^{n_1}} \mathbb{P}_{n_1}(I)$ and notice that after a finite number of steps, $\mathbb{P}'_{n_1} = \varnothing$.
- Note that we always have $2^{-n_1} \lesssim 2^d |F|$.

For d sufficiently large, the intervals I_P for $P \in \mathbb{P}^d$ are going to be essentially disjoint and the intervals $I \in \mathcal{J}_1^{n_1}$ can be selected in an easier way, but this is not the case, for example, when $d = 0$, which corresponds to $I_P \cap \Omega^c \neq \varnothing$. However, for every n_1, the intervals in $\mathcal{J}_1^{n_1}$ are going to be disjoint and this is going to be used later in the proof.

Similarly, $\mathcal{J}_2^{n_2}$ denotes the collection of maximal dyadic intervals I containing at least some $I_P \subseteq I$ for some $P \in \mathbb{P}^d$, and

$$\frac{1}{|I|} \int \mathbf{1}_G \cdot \tilde{\chi}_I^M \, dx \sim 2^{-n_2} \lesssim 2^d |G|.$$

For $\mathbf{1}_{H'}$, let $\mathcal{J}_3^{n_3}$ be the collection of maximal dyadic intervals I containing at least some I_P for some $P \in \mathbb{P}^d$ and such that

$$\frac{1}{|I|} \int \mathbf{1}_{H'} \cdot \tilde{\chi}_I^M \, dx \sim 2^{-n_3} \lesssim 2^{-Md}.$$

We define $\mathcal{J}^{n_1,n_2,n_3} := \mathcal{J}_1^{n_1} \cap \mathcal{J}_2^{n_2} \cap \mathcal{J}_3^{n_3}$, and we further partition $\mathbb{P}^d$ as $\mathbb{P}^d = \bigcup_{n_1,n_2,n_3} \bigcup_{I \in \mathcal{J}^{n_1,n_2,n_3}} \mathbb{P}(I)$.

For $I \in \mathcal{J}_1^{n_1}$, we have $\widetilde{\text{size}}_{\mathbb{P}_{n_1}(I)} \mathbf{1}_F \sim 2^{-n_1}$. When we consider the intersection I' of different intervals in $\mathcal{J}_1^{n_1}, \mathcal{J}_2^{n_2}$ and $\mathcal{J}_3^{n_3}$, all we can say is that $\widetilde{\text{size}}_{\mathbb{P}(I')} \mathbf{1}_F \lesssim 2^{-n_1}$. This fact is the technical obstruction in obtaining vector-valued BHT estimates for any p, q, s in the whole range of BHT.

In a similar way, the relation $(1/|I|) \int_{\mathbb{R}} \mathbf{1}_F \cdot \tilde{\chi}_I^M \, dx \sim 2^{-n_1}$ for $I \in \mathcal{J}_1^{n_1}$ becomes for an interval $I' \in \mathcal{J}_1^{n_1} \cap \mathcal{J}_2^{n_2} \cap \mathcal{J}_3^{n_3}$ an inequality: $(1/|I'|) \int_{\mathbb{R}} \mathbf{1}_F \cdot \tilde{\chi}_{I'}^M \, dx \lesssim 2^{-n_1}$.

The trilinear form in (47) becomes

$$\begin{aligned}
&\sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{J}^{n_1,n_2,n_3}} \Lambda_{\overrightarrow{\text{BHT}};\mathbb{P}(I)}(f, g, h) \\
&\qquad = \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{J}^{n_1,n_2,n_3}} \int_{\mathbb{R}} \int_{\mathcal{W}} \text{BHT}_{\mathbb{P}(I)}(f_w, g_w)(x) \cdot h_w(x) \, d\mu(w) \, dx \\
&\qquad = \int_{\mathcal{W}} \Big(\sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{J}^{n_1,n_2,n_3}} \int_{\mathbb{R}} \text{BHT}_{\mathbb{P}(I)}(f_w \cdot \mathbf{1}_F, g_w \cdot \mathbf{1}_G)(x) \cdot \mathbf{1}_{H'}(x) \cdot h_w(x) \, dx \Big) \, d\mu(w).
\end{aligned}$$

Note that the functions f_w are supported on F, the g_w on G and the h_w on H', for a.e. w. We can apply the localization, Proposition 42, to get

$$\begin{aligned}
&\big|\Lambda_{\text{BHT};\mathbb{P}(I)}^{F,G,H'}(f_w, g_w, h_w)\big| \\
&\qquad \lesssim (\widetilde{\text{size}}_{\mathbb{P}(I)} \mathbf{1}_F)^{a_1} (\widetilde{\text{size}}_{\mathbb{P}(I)} \mathbf{1}_G)^{a_2} (\widetilde{\text{size}}_{\mathbb{P}(I)} \mathbf{1}_{H'})^{a_3} \, \| f_w \cdot \tilde{\chi}_I \|_{r_1} \, \| g_w \cdot \tilde{\chi}_I \|_{r_2} \, \| h_w \cdot \tilde{\chi}_I \|_{r'},
\end{aligned}$$

where $\frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r'} = 1$.

Recall the expressions for a_j from (41):

$$a_1 = \frac{1+\theta_1}{2} - \frac{1}{r_1} - \epsilon, \quad a_2 = \frac{1+\theta_2}{2} - \frac{1}{r_2} - \epsilon, \quad a_3 = \frac{1+\theta_3}{2} - \frac{1}{r'} - \epsilon,$$

where the only conditions we have on θ_1, θ_2 and θ_3 are that $\theta_1+\theta_2+\theta_3 = 1$ and $a_j > 0$. Using Hölder's inequality, the initial trilinear form can be estimated by

$$\begin{aligned}
&\sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} \int_{\mathcal{W}} \left| \Lambda_{\mathrm{BHT};\mathbb{P}(I)}(f_w, g_w, h_w) \right| \\
&\quad \lesssim \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_F)^{a_1} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_G)^{a_2} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_{H'})^{a_3} \\
&\qquad \left(\int_{\mathcal{W}} \| f_w \cdot \tilde{\chi}_I \|_{r_1}^{r_1} \, d\mu(w) \right)^{\frac{1}{r_1}} \left(\int_{\mathcal{W}} \| g_w \cdot \tilde{\chi}_I \|_{r_2}^{r_2} \, d\mu(w) \right)^{\frac{1}{r_2}} \left(\int_{\mathcal{W}} \| h_w \cdot \tilde{\chi}_I \|_{r'}^{r'} \, d\mu(w) \right)^{\frac{1}{r'}} \\
&\quad \lesssim \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_F)^{a_1} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_G)^{a_2} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)} \mathbf{1}_{H'})^{a_3} \\
&\qquad \cdot \frac{\| \mathbf{1}_F \cdot \tilde{\chi}_I \|_{r_1}}{|I|^{\frac{1}{r_1}}} \frac{\| \mathbf{1}_G \cdot \tilde{\chi}_I \|_{r_2}}{|I|^{\frac{1}{r_2}}} \frac{\| \mathbf{1}_{H'} \cdot \tilde{\chi}_I \|_{r'}}{|I|^{\frac{1}{r'}}} |I| \\
&\quad \lesssim \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} 2^{-\frac{n_1}{p}} 2^{-\frac{n_2}{q}} 2^{-n_3(a_3 + \frac{1}{r'})} |I|.
\end{aligned}$$

In the last inequality we need to assume $\frac{1}{p} \le a_1 + \frac{1}{r_1} = \frac{1}{2}(1+\theta_1)$ and similarly $\frac{1}{q} \le \frac{1}{2}(1+\theta_2)$. We will be summing $|I|$ when $I \in \mathcal{I}^{n_1,n_2,n_3}$. Note that

$$\sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} |I| \le \sum_{I \in \mathcal{I}_1^{n_1}} |I| = \Big\| \sum_{I \in \mathcal{I}_1^{n_1}} \mathbf{1}_I \Big\|_{1,\infty} \lesssim \Big\| \sum_{I \in \mathcal{I}_1^{n_1}} 2^{n_1} \mathcal{M}(\mathbf{1}_F) \cdot \mathbf{1}_I \Big\|_{1,\infty} \lesssim 2^{n_1} |F|.$$

Similarly, $\sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} |I| \lesssim 2^{n_2}|G|$ and $\sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} |I| \lesssim 2^{n_3}|H|$ and interpolating these three inequalities we get

$$\sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} |I| \lesssim (2^{n_1}|F|)^{\gamma_1} (2^{n_2}|G|)^{\gamma_2} (2^{n_3}|H|)^{\gamma_3},$$

where $0 \le \gamma_j \le 1$ and $\gamma_1+\gamma_2+\gamma_3 = 1$. Finally,

$$\begin{aligned}
\Big| \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} \Lambda_{\overrightarrow{\mathrm{BHT}};\mathbb{P}(I)}(f,g,h) \Big| &\lesssim \sum_{n_1,n_2,n_3} 2^{-\frac{n_1}{p}} 2^{-\frac{n_2}{q}} 2^{-n_3 \frac{1+\theta_3}{2}} (2^{n_1}|F|)^{\gamma_1} (2^{n_2}|G|)^{\gamma_2} (2^{n_3}|H|)^{\gamma_3} \\
&\lesssim \sum_{n_1,n_2,n_3} 2^{-n_1(\frac{1}{p}-\gamma_1)} 2^{-n_2(\frac{1}{q}-\gamma_2)} 2^{-n_3(\frac{1+\theta_3}{2}-\gamma_3)} |F|^{\gamma_1} |G|^{\gamma_2}.
\end{aligned}$$

The above series converges if we can pick γ_j such that

$$\frac{1}{p} > \gamma_1, \quad \frac{1}{q} > \gamma_2 \quad \text{and} \quad \frac{1+\theta_3}{2} > \gamma_3.$$

This will be possible as long as

$$\frac{1}{p}+\frac{1}{q}+\frac{1+\theta_3}{2}>1. \tag{49}$$

If the above conditions are satisfied, we get generalized restricted-type estimates

$$|\Lambda_{\overrightarrow{\mathrm{BHT}}}(f,g,h)| \lesssim |F|^{\frac{1}{p}}|G|^{\frac{1}{q}}.$$

There are four distinct cases:

(i) $\frac{1}{r_1},\frac{1}{r_2},\frac{1}{r'}\le\frac{1}{2}$. In this case, if we pick $\theta_1=\theta_2\sim 0$ and $\theta_3\sim 1$, all the conditions hold and the range of L^p estimates for $\overrightarrow{\mathrm{BHT}}_{\vec{r}}$ is going to be the convex hull of the points

$$(0,0,1),\quad (1,0,0),\quad \big(1,\tfrac{1}{2},-\tfrac{1}{2}\big),\quad \big(\tfrac{1}{2},1,-\tfrac{1}{2}\big),\quad (0,1,0).$$

That is, we get the same range as that of the BHT operator: $p,q>1$, $s>\frac{2}{3}$ and $\frac{1}{p}+\frac{1}{q}=\frac{1}{s}$.

(ii) $\frac{1}{r_2},\frac{1}{r'}\le\frac{1}{2}$ and $\frac{1}{r_1}>\frac{1}{2}$. For the condition $\frac{1}{2}(1+\theta_1)-\frac{1}{r_1}>0$ to hold, we have to choose $\theta_1>\frac{2}{r_1}-1$ and this will imply that the range of the operator, described as a region in the hyperplane $\beta_1+\beta_2+\beta_3=1$, is the convex hull of the points

$$(0,0,1),\quad (1,0,0),\quad \big(1,\tfrac{1}{2},-\tfrac{1}{2}\big),\quad \big(\tfrac{1}{r_1},\tfrac{3}{2}-\tfrac{1}{r_1},-\tfrac{1}{2}\big),\quad \big(0,\tfrac{3}{2}-\tfrac{1}{r_1},\tfrac{1}{r_1}-\tfrac{1}{2}\big).$$

(iii) $\frac{1}{r_1},\frac{1}{r'}\le\frac{1}{2}$ and $\frac{1}{r_2}>\frac{1}{2}$. Similarly to the previous case, the range of the operator is the convex hull of

$$(0,0,1),\quad (0,1,0),\quad \big(1,\tfrac{1}{2},-\tfrac{1}{2}\big),\quad \big(\tfrac{3}{2}-\tfrac{1}{r_2},\tfrac{1}{r_2},-\tfrac{1}{2}\big),\quad \big(\tfrac{3}{2}-\tfrac{1}{r_2},0,\tfrac{1}{r_2}-\tfrac{1}{2}\big).$$

(iv) $\frac{1}{r_1},\frac{1}{r_2}\le\frac{1}{2}$ and $\frac{1}{r'}>\frac{1}{2}$. The range is the convex hull of

$$(0,0,1),\quad \big(\tfrac{1}{2}+\tfrac{1}{r},0,\tfrac{1}{2}-\tfrac{1}{r}\big),\quad \big(\tfrac{1}{2}+\tfrac{1}{r},\tfrac{1}{2},-\tfrac{1}{r}\big),\quad \big(\tfrac{1}{2},\tfrac{1}{2}+\tfrac{1}{r},-\tfrac{1}{r}\big),\quad \big(0,\tfrac{1}{2}+\tfrac{1}{r},\tfrac{1}{2}-\tfrac{1}{r}\big).$$

3.3. ***The cases $r=1$ or $r_i=\infty$.*** The proof is similar to the one in the previous Section 3.2. We first consider the case $r=1$. Because the dual space of $L^1(\mathcal{W},\mu)$ is $L^\infty(\mathcal{W},\mu)$, the functions appearing in the trilinear form satisfy

$$\|f(x,\cdot)\|_{L^{r_1}(\mathcal{W},\mu)}\le \mathbf{1}_F(x),\quad \|g(x,\cdot)\|_{L^{r_2}(\mathcal{W},\mu)}\le \mathbf{1}_G(x),\quad \|h(x,\cdot)\|_{L^{\infty}(\mathcal{W},\mu)}\le \mathbf{1}_{H'}.$$

All the details are identical to the case $r>1$; the restrictions are given by only two inequalities:

$$\frac{1+\theta_1}{2}>\frac{1}{r_1},\qquad \frac{1+\theta_2}{2}>\frac{1}{r_2}.$$

In the case $r_1=r_2=2$ and $r=1$, these are automatically satisfied and $\mathcal{D}_{r_1,r_2,r}=\mathrm{Range(BHT)}$.

When $r_1=\infty$, we use the fact that the adjoint $\mathrm{BHT}^{*,1}$ of BHT is a bilinear operator of the same kind, which is bounded from $L^r\times L^{r'}\to L^1$; more precisely,

$$\Lambda_{\mathrm{BHT}}(f_w,g_w,h_w)=\int_{\mathbb{R}}\mathrm{BHT}(f_w,g_w)(x)\cdot h_w(x)\,dx=\int_{\mathbb{R}} f_w(x)\cdot \mathrm{BHT}^{*,1}(g_w,h_w)(x)\,dx.$$

In proving the boundedness of vector-valued BHT via interpolation, we assume

$$\|f(x,\cdot)\|_{L^{\infty}(\mathcal{W},\mu)}\le \mathbf{1}_F(x),\quad \|g(x,\cdot)\|_{L^{r}(\mathcal{W},\mu)}\le \mathbf{1}_G(x),\quad \|h(x,\cdot)\|_{L^{r'}(\mathcal{W},\mu)}\le \mathbf{1}_{H'}.$$

Then

$$\begin{aligned}&\left|\Lambda_{\mathrm{BHT};\mathbb{P}(I)}(f_w, g_w, h_w)\right|\\ &\qquad\le \left\|\mathrm{BHT}^{*,1}_{\mathbb{P}(I)}(g_w\cdot\mathbf{1}_G, h_w\cdot\mathbf{1}_{H'})\cdot\mathbf{1}_F\right\|_1\\ &\qquad\lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\epsilon}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{r}-\epsilon}(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{r'}-\epsilon}\|g_w\cdot\tilde{\chi}_I\|_r\,\|h_w\cdot\tilde{\chi}_I\|_{r'}.\end{aligned}$$

The rest follows as before. Note that in the case $(\infty, 2, 2)$ we have no constraints on p, q, and s except those coming from the original BHT operator itself: indeed, for $\theta_2, \theta_3 > 0$, we have

$$\frac{1+\theta_2}{2}-\frac{1}{2}>0,\qquad \frac{1+\theta_3}{2}-\frac{1}{2}>0.$$

3.4. ***Iterated $L^p(\mathcal{W},\mu)$ spaces estimates for* BHT.** Previously, we proved that for any tuple (r_1, r_2, r) with $\frac{1}{r_1}+\frac{1}{r_2}=\frac{1}{r}$, $1\le r<\infty$, and $1<r_1, r_2\le\infty$, we have

$$\mathrm{BHT}: L^p(\mathbb{R}; L^{r_1}(\mathcal{W},\mu))\times L^q(\mathbb{R}; L^{r_2}(\mathcal{W},\mu))\to L^s(\mathbb{R}; L^r(\mathcal{W},\mu))$$

whenever p, q, r are in a certain range $\mathcal{D}_{r_1,r_2,r}$, which can be described in a precise manner. The general ideas for proving multiple vector-valued estimates for BHT (as presented in Theorem 8) via the helicoidal method were described in the Introduction. In this section, we present in more detail the proof in the case of two iterated spaces $\ell^s(\ell^r)$ in order to simplify the notation. First, we prove the following localized vector-valued result:

Proposition 44.

$$\left\|\Big(\sum_{k=1}^N\left|\mathrm{BHT}_{\mathbb{P}(I_0)}(f_k\cdot\mathbf{1}_F, g_k\cdot\mathbf{1}_G)\right|^r\Big)^{\frac{1}{r}}\cdot\mathbf{1}_{H'}\right\|_s\le\widetilde{C}\left\|\Big(\sum_{k=1}^N|f_k|^{r_1}\Big)^{\frac{1}{r_1}}\cdot\tilde{\chi}_{I_0}\right\|_p\left\|\Big(\sum_{k=1}^N|g_k|^{r_2}\Big)^{\frac{1}{r_2}}\cdot\tilde{\chi}_{I_0}\right\|_q,$$

where $\widetilde{C}=(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{p}-\epsilon}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{q}-\epsilon}(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{s'}-\epsilon}$.

Proof. This is going to be a refinement of the proof of Theorem 7 from the previous section. In constructing the collection of intervals $\mathcal{I}_j^{n_j}$, we note that we only need to select intervals I that are already contained in I_0, because all the tiles in $\mathbb{P}(I_0)$ are such that $I_P\subseteq I_0$.

As before, we prove generalized restricted-type estimates, and we assume that the functions have the properties

$$\Big(\sum_k|f_k|^{r_1}\Big)^{\frac{1}{r_1}}\le\mathbf{1}_{E_1},\quad\Big(\sum_k|f_k|^{r_2}\Big)^{\frac{1}{r_2}}\le\mathbf{1}_{E_2},\quad\Big(\sum_k|h_k|^{r'}\Big)^{\frac{1}{r'}}\le\mathbf{1}_{E_3'}.$$

The exceptional set is defined by

$$\widetilde{\Omega}=\left\{\mathcal{M}(\mathbf{1}_{E_1})>C\frac{|E_1|}{|E_3|}\right\}\cup\left\{\mathcal{M}(\mathbf{1}_{E_2})>C\frac{|E_2|}{|E_3|}\right\},$$

and we assume the tiles to be such that $1+\mathrm{dist}(I_P,\widetilde{\Omega}^c)/|I_P|\sim 2^d$.

For intervals $I \in \mathcal{I}_1^{n_1}$, we have

$$\frac{1}{|I|}\int_{\mathbb{R}} \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_I^M \, dx \sim \widetilde{\text{size}}_{\mathbb{P}_{n_1}(I)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F) \sim 2^{-n_1} \leq 2^d \frac{|E_1|}{|E_3|}.$$

When we consider intervals $I \in \mathcal{I}_1^{n_1} \cap \mathcal{I}_2^{n_2} \cap \mathcal{I}_3^{n_3}$, the above approximations become inequalities. We also need to point out that

$$\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F) \leq \widetilde{\text{size}}_{\mathbb{P}(I_0)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F) \quad \text{and} \quad \frac{1}{|I|}\int_{\mathbb{R}} \mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_I^M \, dx \leq \widetilde{\text{size}}_{\mathbb{P}(I_0)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F).$$

Now we add the trilinear forms in order to obtain generalized restricted-type estimates:

$$\begin{aligned}
&\sum_k \left|\Lambda_{\text{BHT};\mathbb{P}(I_0)}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'})\right| \\
&\leq \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} \sum_k \left|\Lambda_{\text{BHT};\mathbb{P}(I_0 \cap I)}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'})\right| \\
&\lesssim \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} (\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F))^{\frac{1+\theta_1}{2} - \frac{1}{r_1} - \epsilon} (\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_2} \cdot \mathbf{1}_G))^{\frac{1+\theta_2}{2} - \frac{1}{r_2} - \epsilon} \\
&\qquad \cdot (\widetilde{\text{size}}_{\mathbb{P}(I)}(\mathbf{1}_{E_3'} \cdot \mathbf{1}_{H'}))^{\frac{1+\theta_3}{2} - \frac{1}{r'} - \epsilon} \frac{\|\mathbf{1}_{E_1} \cdot \mathbf{1}_F \cdot \tilde{\chi}_I\|_{r_1}}{|I|^{\frac{1}{r_1}}} \frac{\|\mathbf{1}_{E_2} \cdot \mathbf{1}_G \cdot \tilde{\chi}_I\|_{r_2}}{|I|^{\frac{1}{r_2}}} \frac{\|\mathbf{1}_{E_3'} \cdot \mathbf{1}_{H'} \cdot \tilde{\chi}_I\|_{r'}}{|I|^{\frac{1}{r'}}} |I|.
\end{aligned}$$

Using the *modified* sizes from Definition 21, this implies

$$\begin{aligned}
&\sum_k \left|\Lambda_{\text{BHT};\mathbb{P}(I_0)}(f_k \cdot \mathbf{1}_F, g_k \cdot \mathbf{1}_G, h_k \cdot \mathbf{1}_{H'})\right| \\
&\quad \lesssim (\widetilde{\text{size}}_{\mathbb{P}(I_0)}(\mathbf{1}_{E_1} \cdot \mathbf{1}_F))^{\frac{1+\theta_1}{2} - \frac{1}{p} - \epsilon} (\widetilde{\text{size}}_{\mathbb{P}(I_0)}(\mathbf{1}_{E_2} \cdot \mathbf{1}_G))^{\frac{1+\theta_2}{2} - \frac{1}{q} - \epsilon} (\widetilde{\text{size}}_{\mathbb{P}(I_0)}(\mathbf{1}_{E_3'} \cdot \mathbf{1}_{H'}))^{\frac{1+\theta_3}{2} - \frac{1}{s'} - \epsilon} \\
&\qquad \cdot \sum_{n_1,n_2,n_3} \sum_{I \in \mathcal{I}^{n_1,n_2,n_3}} 2^{-\frac{n_1}{p}} 2^{-\frac{n_2}{q}} 2^{-n_3(\frac{1}{s'}+\epsilon)} |I|.
\end{aligned}$$

The last part adds up to something $\lesssim 2^{-\tilde{M}d} |E_1|^{\frac{1}{p}} |E_2|^{\frac{1}{q}} |E_3|^{\frac{1}{s'}}$, which is precisely what we were aiming in the beginning.

The cases when one of r_1, r_2 or $r' = \infty$ follow in a similar manner. □

The above proposition is an intermediate step in the proof of L^p estimates for $\overrightarrow{\text{BHT}}_{\vec{R}}$, in the case of two iterated vector spaces, which is presented below.

Proposition 45.

$$\left\| \Big(\sum_l \Big(\sum_k \left|\text{BHT}(f_{kl}, g_{kl})\right|^r\Big)^{\frac{s}{r}}\Big)^{\frac{1}{s}} \right\|_t \leq C \left\| \Big(\sum_l \Big(\sum_k |f_{kl}|^{r_1}\Big)^{\frac{s_1}{r_1}}\Big)^{\frac{1}{s_1}} \right\|_p \left\| \Big(\sum_l \Big(\sum_k |g_{kl}|^{r_r}\Big)^{\frac{s_2}{r_2}}\Big)^{\frac{1}{s_2}} \right\|_q.$$

Proof. Once again, we use generalized restricted-type interpolation; F, G, H are sets of finite measure, with $|H| = 1$. The exceptional set is defined as usual, and $H' = H \setminus \Omega$. The sequences of functions will be such that

$$\Big(\sum_l \Big(\sum_k |f_{kl}|^{r_1}\Big)^{\frac{s_1}{r_1}}\Big)^{\frac{1}{s_1}} \le \mathbf{1}_F, \quad \Big(\sum_l \Big(\sum_k |g_{kl}|^{r_2}\Big)^{\frac{s_2}{r_2}}\Big)^{\frac{1}{s_2}} \le \mathbf{1}_G, \quad \Big(\sum_l \Big(\sum_k |h_{kl}|^{r'}\Big)^{\frac{s'}{r'}}\Big)^{\frac{1}{s'}} \le \mathbf{1}_{H'}.$$

The collections $\mathcal{I}_j^{n_j}$ are going to be chosen in the same way as in the proof of Theorem 7, depending on the sizes and averages of the characteristic functions $\mathbf{1}_F, \mathbf{1}_G, \mathbf{1}_{H'}$. Proposition 44 yields the following:

$$\sum_k \big|\Lambda_{\mathrm{BHT};\mathbb{P}(I)}(f_{kl}, g_{kl}, h_{kl})\big|$$
$$\lesssim (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{s_1}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{s_2}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{s'}-\epsilon} \tag{50}$$
$$\cdot \Big\| \Big(\sum_k |f_{kl}|^{r_1}\Big)^{\frac{1}{r_1}} \tilde{\chi}_I \Big\|_{s_1} \Big\| \Big(\sum_k |g_{kl}|^{r_2}\Big)^{\frac{1}{r_2}} \tilde{\chi}_I \Big\|_{s_2} \Big\| \Big(\sum_k |h_{kl}|^{r'}\Big)^{\frac{1}{r'}} \tilde{\chi}_I \Big\|_{s'}. \tag{51}$$

Then we sum (51) over l as well, and apply Hölder on the triple (s_1, s_2, s'). In this way, we recover $\|\mathbf{1}_F \cdot \tilde{\chi}_I\|_{s_1}$, and the corresponding quantities for the second and third entries. We have

$$\Big|\sum_{k,l} \Lambda_{\mathrm{BHT}}(f_{kl}, g_{kl}, h_{kl})\Big|$$
$$\lesssim \sum_{n_1,n_2,n_3} \sum_{\mathcal{I}^{n_1,n_2,n_3}} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{s_1}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{s_2}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{s'}-\epsilon}$$
$$\cdot \frac{\|\mathbf{1}_F\,\tilde{\chi}_I\|_{s_1}}{|I|^{\frac{1}{s_1}}} \frac{\|\mathbf{1}_G\,\tilde{\chi}_I\|_{s_2}}{|I|^{\frac{1}{s_2}}} \frac{\|\mathbf{1}_{H'}\,\tilde{\chi}_I\|_{s'}}{|I|^{\frac{1}{s'}}} |I|$$
$$\lesssim \sum_{n_1,n_2,n_3} \sum_{\mathcal{I}^{n_1,n_2,n_3}} (\text{“}\widetilde{\mathrm{size}}\text{”}_{\mathbb{P}(I)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{p}-\epsilon} (\text{“}\widetilde{\mathrm{size}}\text{”}_{\mathbb{P}(I)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{q}-\epsilon} (\text{“}\widetilde{\mathrm{size}}\text{”}_{\mathbb{P}(I)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{t'}-\epsilon}$$
$$\cdot 2^{-\frac{n_1}{p}} 2^{-\frac{n_2}{q}} 2^{-n_3(\frac{-1}{t'}+\epsilon)} |I|.$$

Remark. The “sizes” appearing in the line above are not exactly the ones from Definition 19, but the modified ones from Definition 21 . Note that

$$\max\Big(\widetilde{\mathrm{size}}_{\mathbb{P}(I)}\mathbf{1}_F, \frac{1}{|I|}\int_{\mathbb{R}} \mathbf{1}_F \cdot \tilde{\chi}_I^M\, dx\Big) \le \text{“}\widetilde{\mathrm{size}}\text{”}_{\mathbb{P}(I)}\mathbf{1}_F.$$

This is the step where we can prove also the localized version of the statement in Proposition 45. Assuming all the tiles are sitting above an interval I_0, we can obtain the same result with operatorial norm

$$(\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_F)^{\frac{1+\theta_1}{2}-\frac{1}{p}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_G)^{\frac{1+\theta_2}{2}-\frac{1}{q}-\epsilon} (\widetilde{\mathrm{size}}_{\mathbb{P}(I_0)}\mathbf{1}_{H'})^{\frac{1+\theta_3}{2}-\frac{1}{t'}-\epsilon}.$$

The rest of the proof is identical to the simpler vector case of Theorem 7; the quantities on the left-hand side add up to $|F|^{\frac{1}{p}} |G|^{\frac{1}{q}}$, provided

$$\frac{1+\theta_1}{2} > \frac{1}{p}, \quad \frac{1+\theta_2}{2} > \frac{1}{q}, \quad \frac{1+\theta_3}{2} > \frac{1}{s'}.$$

□

4. Similar results for paraproducts: proof of Theorem 9

The paraproduct case is similar to BHT, even though the bilinear Hilbert transform is a much more complicated object. The extra difficulties are hidden in Proposition 23, but we will see from the proof of the vector-valued extensions that the complexity of the paraproduct case is comparable to the "local L^2" case for BHT. In both situations, we recover the maximal range for vector-valued estimates.

We will be working with the *discretized paraproduct* of the functions f and g, which is defined by

$$\Pi(f,g)(x) = \sum_{I \in \mathcal{I}} \frac{1}{|I|^{\frac{1}{2}}} \langle f, \phi_I^1 \rangle \langle g, \phi_I^2 \rangle \phi_I^3(x).$$

Here $\mathcal{I}$ is a family of dyadic intervals, and the wave packets $\{\phi_I^j\}_{I \in \mathcal{I}}$ are so that two of the families are *lacunary* (ϕ_I^j is a wave packet on $I \times [1/|I|, 2/|I|]$), and the third one is *nonlacunary* ($\phi_I^{j_0}$ is a wave packet on $I \times [0, 1/|I|]$). Again, we present the case of ℓ^p spaces for simplicity. The operator we are interested in is

$$\vec{\Pi}_r(f,g) := \Big(\sum_{k=1}^{N} \big| \Pi(f_k, g_k) \big|^r \Big)^{\frac{1}{r}}.$$

Remark. We could alternatively look at operators of the form

$$(f,g) \mapsto \Big(\sum_{k=1}^{N} \big| \Pi_k(f_k, g_k) \big|^r \Big)^{\frac{1}{r}},$$

where each paraproduct Π_k is associated to a family $\mathcal{I}_k$ of dyadic intervals. The Π_k don't need to be precisely the same, but they display a similar behavior. Similarly, for $\overrightarrow{\mathrm{BHT}}$ we could have a "perturbation" BHT_w for each $w \in \mathcal{W}$, and the method of the proof applies in that case as well.

4.1. ***A few results about paraproducts.*** The concepts of *sizes* and *energies* are similar to the corresponding ones for the bilinear Hilbert transform; we don't need to organize the tiles into trees because the family of tiles is of rank 0. We recall some definitions below.

Definition 46. Let $\mathcal{I}$ be a family of dyadic intervals. For any $1 \leq j \leq 3$, we define

$$\mathrm{size}_{\mathcal{I}}\big(\langle f, \phi_I^j \rangle_{I \in \mathcal{I}}\big) = \sup_{I \in \mathcal{I}} \frac{|\langle f, \phi_I^j \rangle|}{|I|^{\frac{1}{2}}} \quad \text{if } (\phi_I^j)_I \text{ is nonlacunary}$$

and

$$\mathrm{size}_{\mathcal{I}}\big(\langle f, \phi_I^j \rangle_{I \in \mathcal{I}}\big) = \sup_{I_0 \in \mathcal{I}} \frac{1}{|I_0|^{\frac{1}{2}}} \Big\| \Big(\sum_{\substack{I \in \mathcal{I} \\ I \subseteq I_0}} \frac{|\langle f, \phi_I^j \rangle|^2}{|I|} \cdot \mathbf{1}_I \Big)^{\frac{1}{2}} \Big\|_{1,\infty} \quad \text{if } (\phi_I^j)_I \text{ is lacunary.}$$

Similarly to the BHT case, energy is defined as

$$\mathrm{energy}_{\mathcal{I}}^j\big(\langle f, \phi_I^j \rangle_{I \in \mathcal{I}}\big) := \sup_{n \in \mathbb{Z}} 2^n \sup_{\mathbb{D}} \Big(\sum_{I \in \mathbb{D}} |I| \Big),$$

where $\mathbb{D}$ ranges over all collections of disjoint intervals I_0 with the property that

$$\frac{|\langle f, \phi^j_{I_0}\rangle|}{|I_0|^{\frac{1}{2}}} \geq 2^n \quad \text{if } (\phi^j_I)_I \text{ is nonlacunary}$$

and

$$\frac{1}{|I_0|}\bigg\|\bigg(\sum_{\substack{I\in\mathcal{I}\\ I\subseteq I_0}} \frac{|\langle f, \phi^j_I\rangle|^2}{|I|}\cdot \mathbf{1}_I\bigg)^{\frac{1}{2}}\bigg\|_{1,\infty} \geq 2^n \quad \text{if } (\phi^j_I)_I \text{ is lacunary.}$$

We have estimates similar to Lemmas 20 and 24. However, because we don't need to use orthogonality of trees, the energy becomes an L^1 quantity.

Lemma 47 [Muscalu and Schlag 2013, Lemma 2.13]. *If F is an L^1 function and $1 \leq j \leq 3$, then*

$$\text{size}^j_{\mathcal{I}}\big(\langle F, \phi^j_I\rangle_{I\in\mathcal{I}}\big) \lesssim \sup_{I\in\mathcal{I}} \frac{1}{|I|}\int_{\mathbb{R}} |F|\tilde{\chi}^M_I\, dx$$

for $M > 0$, with implicit constants depending on M.

Lemma 48 [Muscalu and Schlag 2013, Lemma 2.14]. *If F is an L^1 function and $1 \leq j \leq 3$, then*

$$\text{energy}^j_{\mathcal{I}}\big(\langle F, \phi^j_I\rangle_{I\in\mathcal{I}}\big) \lesssim \|F\|_1.$$

Proposition 49 [Muscalu and Schlag 2013, Proposition 2.12]. *Given a paraproduct Π associated with a family $\mathcal{I}$ of intervals,*

$$\begin{aligned}\big|\Lambda_\Pi(f_1, f_2, f_3)\big| &= \bigg|\sum_{I\in\mathcal{I}} \frac{1}{|I|^{\frac{1}{2}}}\langle f_1, \phi^1_I\rangle\langle f_2, \phi^2_I\rangle\langle f_3, \phi^3_I\rangle\bigg| \\ &\lesssim \prod_{j=1}^{3}\big(\text{size}^{(j)}_{\mathcal{I}}((\langle f_j, \phi^j_I\rangle_{I\in\mathcal{I}}))\big)^{1-\theta_j}\big(\text{energy}^{(j)}_{\mathcal{I}}((\langle f_j, \phi^j_I\rangle_{I\in\mathcal{I}}))\big)^{\theta_j}\end{aligned}$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$ such that $\theta_1 + \theta_2 + \theta_3 = 1$, where the implicit constant depends on $\theta_1, \theta_2, \theta_3$ only.

While the above proposition is the main ingredient, we need "localized" estimates. If I_0 is some fixed dyadic interval, then we define

$$\Pi(I_0)(f, g)(x) = \sum_{\substack{I\in\mathcal{I}\\ I\subseteq I_0}} \frac{1}{|I|^{\frac{1}{2}}}\langle f, \phi^1_I\rangle\langle g, \phi^2_I\rangle\phi^3_I(x).$$

Here again we need some localization results which play the role of Proposition 42 and Corollary 43 from the BHT case.

The trilinear form associated to the localized paraproduct is given by

$$\Lambda^{F,G,H'}_{\Pi(I_0)}(f, g, h) := \Lambda_{\Pi(I_0)}(f\cdot\mathbf{1}_F, g\cdot\mathbf{1}_G, h\cdot\mathbf{1}_{H'}).$$

Proposition 50. *Let I_0 be a fixed dyadic interval and $F, G, H' \subset \mathbb{R}$ sets of finite measure. Then there exist some positive numbers $0 \leq a_1, a_2, a_3 < 1$ so that*

$$\left|\Lambda^{F,G,H'}_{\Pi(I_0)}(f,g,h)\right| \lesssim (\widetilde{\text{size}}_{\mathcal{I}(I_0)}\mathbf{1}_F)^{a_1}\,(\widetilde{\text{size}}_{\mathcal{I}(I_0)}\mathbf{1}_G)^{a_2}\,(\widetilde{\text{size}}_{\mathcal{I}(I_0)}\mathbf{1}_{H'})^{a_3}\,\|f\cdot\tilde{\chi}_{I_0}\|_{r_1}\,\|g\cdot\tilde{\chi}_{I_0}\|_{r_2}\,\|h\cdot\tilde{\chi}_{I_0}\|_{r'}$$

whenever $\frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r'}=1$, and $1 < r_1, r_2, r' < \infty$. Here $a_j = 1 - \frac{1}{r_j} - \epsilon$.

Proof. The idea of the proof is very similar to that of Proposition 41. Restricted-type estimates are proved by performing a triple stopping time and then the result follows by interpolation. We leave the routine details to the reader. □

The case $r = 1$ is obtained through interpolation of restricted-type estimates only. This comes in contrast with the $r = 1$ case for BHT, where generalized restricted-type interpolation is necessary. More exactly, for the BHT operator, in order to conclude estimates for $\left(\frac{1}{r_1}, \frac{1}{r_2}, 0\right)$, one needs to interpolate between good ($\beta_i > 0$) and bad ($\beta_3 < 0$) tuples $\beta = (\beta_1, \beta_2, \beta_3)$.

Proposition 51. *If H' is a fixed set of finite measure,*

$$\left|\Lambda_{\Pi(I_0)}(f,g,\mathbf{1}_{H'})\right| \lesssim \widetilde{\text{size}}_{\mathcal{I}(I_0)}\mathbf{1}_{H'}\|f\cdot\tilde{\chi}_{I_0}\|_p\,\|g\cdot\tilde{\chi}_{I_0}\|_q \tag{52}$$

whenever $\frac{1}{p}+\frac{1}{q}=1$, and $1 < p, q < \infty$.

Proof. In this case $\Lambda_{\Pi(I_0)}(f,g,\mathbf{1}_{H'})$ becomes a bilinear form with respect to the first two entries. Because of the decay of $\tilde{\chi}_{I_0}$, it will be sufficient to prove the proposition in the case supp $f, g \subseteq 5I_0$. By Theorem 28, it will be enough to show restricted-type estimates for the bilinear form

$$(f,g) \mapsto \Lambda_{\Pi(I_0)}(f,g,\mathbf{1}_{H'}).$$

Let F and G be sets of finite measure and $|f| \leq \mathbf{1}_F$ and $|g| \leq \mathbf{1}_G$. Using Proposition 49 with $\theta_3 = 0$ and estimating $\widetilde{\text{size}}_{\mathcal{I}(I_0)}f \lesssim 1$ and $\widetilde{\text{size}}_{\mathcal{I}(I_0)}g \lesssim 1$, we get

$$\left|\Lambda_{\Pi(I_0)}(f,g,\mathbf{1}_{H'})\right| \lesssim \widetilde{\text{size}}_{\mathcal{I}(I_0)}\mathbf{1}_{H'}|F|^{\theta_1}|G|^{\theta_2},$$

where $\theta_1 + \theta_2 = 1$ and $0 < \theta_1, \theta_2 < 1$. This proves restricted-type estimates in a small neighborhood of $\left(\frac{1}{p}, \frac{1}{q}\right)$. □

4.2. ***Proof of Theorem 8: a particular case.*** We will be using vector-valued interpolation theorems, as usual. Hence, we fix sets of finite measure F, G and H and we assume $|H| = 1$. Let $f = \{f_k\}_k$ and $g = \{g_k\}_k$, with $(\sum_k |f_k|^{r_1})^{\frac{1}{r_1}} \leq \mathbf{1}_F$ and $(\sum_k |g_k|^{r_2})^{\frac{1}{r_2}} \leq \mathbf{1}_G$.

The exceptional set will be

$$\tilde{\Omega} := \left\{x : \mathcal{M}(\mathbf{1}_F)(x) > C|F|\right\} \cup \left\{x : \mathcal{M}(\mathbf{1}_G)(x) > C|G|\right\}$$

and $H' = H \setminus \tilde{\Omega}$. We have a sequence of functions $\{h_k\}_k$ with $\left(\sum_k |h_k|^{r'}\right)^{\frac{1}{r'}} \leq \mathbf{1}_{H'}$.

For every $d \geq 0$,

$$\mathcal{I}^d := \left\{I \in \mathcal{I} : 1 + \frac{\text{dist}(I, \Omega^c)}{|I|} \sim 2^d\right\}.$$

When estimating paraproducts associated to the collection $\mathcal{I}^d$, we get an extra 2^{-10d} decay and thus the d-dependency of the paraproducts can be assumed to be implicit. As before, for each of the sets F, G and H' we define collections of disjoint maximal intervals $\mathcal{J}_1^{n_1}$, $\mathcal{J}_2^{n_2}$ and $\mathcal{J}_3^{n_3}$ respectively. For example, if $I \in \mathcal{J}_1^{n_1}$, then

$$2^{-n_1-1} \le \frac{1}{|I|} \int_{\mathbb{R}} \mathbf{1}_F \cdot \tilde{\chi}_I \, dx \le 2^{-n_1} \lesssim |F|.$$

Returning to the operator $\vec{\Pi}_r$, we have for the associated multilinear form

$$\Big| \sum_k \Lambda_\Pi(f_k, g_k, h_k) \Big| \le \sum_{n_1,n_2,n_3} \sum_{I_0 \in \mathcal{J}^{n_1,n_2,n_3}} \sum_k \big| \Lambda_{\Pi(I_0)}(f_k, g_k, h_k) \big|.$$

Now we use the localization results of Proposition 50 to estimate the above expression by

$$\begin{aligned} &\sum_{n_1,n_2,n_3} \sum_{I_0 \in \mathcal{J}^{n_1,n_2,n_3}} \sum_{k=1}^{n} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_F)^{b_1} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_G)^{b_2} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_{H'})^{b_3} \\ &\qquad\qquad \cdot \| f_k \cdot \tilde{\chi}_{I_0} \|_{r_1} \| g_k \cdot \tilde{\chi}_{I_0} \|_{r_2} \| h_k \cdot \tilde{\chi}_{I_0} \|_{r'} \\ &\lesssim \sum_{n_1,n_2,n_3} \sum_{I_0 \in \mathcal{J}^{n_1,n_2,n_3}} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_F)^{b_1} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_G)^{b_2} (\widetilde{\text{size}}_{\mathcal{I}(I_0)} \mathbf{1}_{H'})^{b_3} \\ &\qquad\qquad \frac{\| \mathbf{1}_F \cdot \tilde{\chi}_{I_0} \|_{r_1}}{|I_0|^{\frac{1}{r_1}}} \frac{\| \mathbf{1}_G \cdot \tilde{\chi}_{I_0} \|_{r_2}}{|I_0|^{\frac{1}{r_2}}} \frac{\| \mathbf{1}_{H'} \cdot \tilde{\chi}_{I_0} \|_{r'}}{|I_0|^{\frac{1}{r'}}} |I_0|. \end{aligned}$$

Here we choose some $0 \le b_j \le a_j$, which we can do because the sizes are subunitary. Whenever $0 \le \gamma_j \le 1$ are so that $\gamma_1 + \gamma_2 + \gamma_3 = 1$,

$$\sum_{I_0 \in \mathcal{J}^{n_1,n_2,n_3}} |I_0| \lesssim (2^{n_1}|F|)^{\gamma_1} (2^{n_2}|G|)^{\gamma_2} (2^{n_3}|H|)^{\gamma_3}.$$

Adding all the pieces together we have

$$\begin{aligned} \Big| \sum_k \Lambda_\Pi(f_k, g_k, h_k) \Big| &\lesssim \sum_{n_1,n_2,n_3} 2^{-n_1(b_1 + \frac{1}{p} - \gamma_1)} 2^{-n_2(b_2 + \frac{1}{q} - \gamma_2)} 2^{-n_3(b_3 + \frac{1}{r'} - \gamma_3)} |F|^{\gamma_1} |G|^{\gamma_2} \\ &\lesssim |F|^{\frac{1}{p}} |G|^{\frac{1}{q}}. \end{aligned}$$

Of course, the last inequality is true provided we can choose $\gamma_1, \gamma_2, \gamma_3$ so that the series converges. Choosing the θ_j and α_j carefully, one can prove that the restricted weak-type estimates hold arbitrarily close to the points

$$(0, 0, 1), \quad (1, 0, 0), \quad (0, 1, 0), \quad (1, 1, -1).$$

Then the general result follows by interpolation.

Remark. With a few adjustments, the proof is valid in the case $r = 1$ as well.

5. Tensor products $\mathrm{BHT} \otimes \Pi^{\otimes^n}$

In this section, we will prove the boundedness of the tensor product

$$\mathrm{BHT} \otimes \Pi^{\otimes n} = \mathrm{BHT} \otimes \Pi \otimes \cdots \otimes \Pi : L^p(\mathbb{R}^{n+1}) \times L^q(\mathbb{R}^{n+1}) \to L^r(\mathbb{R}^{n+1})$$

whenever $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$, with $\frac{2}{3} < r < \infty$, $1 \le p, q < \infty$.

If $T_1 : L^p(\mathbb{R}^{n_1}) \times L^q(\mathbb{R}^{n_1}) \to L^r(\mathbb{R}^{n_1})$ and $T_2 : L^p(\mathbb{R}^{n_2}) \times L^q(\mathbb{R}^{n_2}) \to L^r(\mathbb{R}^{n_2})$ are two bilinear operators, then the tensor product

$$T_1 \otimes T_2 : L^p(\mathbb{R}^{n_1+n_2}) \times L^q(\mathbb{R}^{n_1+n_2}) \to L^r(\mathbb{R}^{n_1+n_2})$$

will act as T_1 in the first variable and as T_2 in the second variable. In our case, the operators are given by singular multipliers, and in this situation we can give a characterization of the tensor product. Assume

$$T_1(f,g)(x) = \int_{\mathbb{R}^{2n_1}} \hat{f}(\xi_1)\,\hat{g}(\xi_2)\,m_1(\xi_1,\xi_2)\,e^{2\pi i x(\xi_1+\xi_2)}\,d\xi_1\,d\xi_2$$

and similarly

$$T_2(f,g)(y) = \int_{\mathbb{R}^{2n_2}} \hat{f}(\eta_1)\,\hat{g}(\eta_2)\,m_2(\eta_1,\eta_2)\,e^{2\pi i y(\eta_1+\eta_2)}\,d\eta_1\,d\eta_2.$$

Then the multiplier of the tensor product is precisely $m_1(\xi_1,\xi_2) \cdot m_2(\eta_1,\eta_2)$:

$$\begin{aligned} &T_1 \otimes T_2(f,g)(x,y) \\ &\quad = \int \hat{f}(\xi_1,\eta_1)\,\hat{g}(\xi_2,\eta_2)\,m_1(\xi_1,\xi_2)\,m_2(\eta_1,\eta_2)\,e^{2\pi i x(\xi_1+\xi_2)} e^{2\pi i y(\eta_1+\eta_2)}\,d\xi_1\,d\xi_2\,d\eta_1\,d\eta_2. \end{aligned}$$

The multiplier associated with BHT is $\operatorname{sgn}(\xi_1 - \xi_2)$, while the multiplier of a paraproduct of two functions on the real line is a classical Marcinkiewicz–Mikhlin–Hörmander multiplier $m(\xi_1,\xi_2)$, smooth away from the origin, satisfying the condition $|\partial^\alpha m(\xi)| \lesssim |\xi|^{-|\alpha|}$ for sufficiently many multi-indices α. The decay in m and a Fourier series decomposition allows one to approximate the multiplier by a finite number of sums of the form

$$\sum_k \hat{\varphi}_k(\xi_1)\,\hat{\psi}_k(\xi_2)\,\hat{\psi}_k(\xi_1+\xi_2), \quad \sum_k \hat{\psi}_k(\xi_1)\,\hat{\varphi}_k(\xi_2)\,\hat{\psi}_k(\xi_1+\xi_2) \quad \text{or} \quad \sum_k \hat{\psi}_k(\xi_1)\,\hat{\psi}_k(\xi_2)\,\hat{\varphi}_k(\xi_1+\xi_2).$$

Recall that Q_k is the Littlewood–Paley projection onto $\{|\xi| \sim 2^k\}$ (which is really the convolution with $\psi_k(\cdot)$), and P_k is the projection onto $\{|\xi| \le 2^k\}$, corresponding to the convolution with φ_k. Then we can regard paraproducts as being expressions of the form

$$\sum_k Q_k(P_k f \cdot Q_k g)(x,y), \quad \sum_k Q_k(Q_k f \cdot P_k g)(x,y) \quad \text{or} \quad \sum_k P_k(Q_k f \cdot Q_k g)(x,y). \tag{53}$$

It is important in the following proofs that the outermost functions $\hat{\varphi}_k(\xi_1+\xi_2)$ and $\hat{\psi}_k(\xi_1+\xi_2)$ are identically equal to 1 on the supports of $\hat{\psi}_k(\xi_1) \cdot \hat{\psi}_k(\xi_2)$ and $\hat{\psi}_k(\xi_1) \cdot \hat{\varphi}_k(\xi_2)$ respectively. This can always be achieved with the price of an extra decomposition.

Proposition 52. *Let $T_m : L^p(\mathbb{R}^n) \times L^q(\mathbb{R}^n) \to L^r(\mathbb{R}^n)$ be a bilinear operator with smooth symbol m, and $\Pi : L^p(\mathbb{R}) \times L^q(\mathbb{R}) \to L^r(\mathbb{R})$ a paraproduct as described above.*

(1) *If Π is given by $\sum_k Q_k(P_k f \cdot Q_k g)(x, y)$, then*

$$(T_m \otimes \Pi)(f, g)(x, y) = \sum_k Q_k^2\big(T_m(P_k^y f, Q_k^y g)\big)(x) = \sum_k T_m(P_k^y f, Q_k^y g)(x).$$

(2) *If Π is given by $\sum_k P_k(Q_k f \cdot Q_k g)(x, y)$, then*

$$(T_m \otimes \Pi)(f, g)(x, y) = \sum_k P_k^2\big(T_m(Q_k^y f, Q_k^y g)\big)(x) = \sum_k T_m(Q_k^y f, Q_k^y g)(x).$$

Here we need to explain the notation: Q_k^2 denotes the projection onto $|\xi_2| \sim 2^k$ in the second variable, and $P_k^y f$ is a function of x only, with the variable y fixed. The exact formulas are

$$P_k^y f(x) = \int_{\mathbb{R}} \varphi_k(s) f(x, y-s)\, ds, \quad P_k^2 f(x, y) = \int_{\mathbb{R}} \varphi_k(s) f(x, y-s)\, ds,$$
$$Q_k^y f(x) = \int_{\mathbb{R}} \psi_k(s) f(x, y-s)\, ds, \quad Q_k^2 f(x, y) = \int_{\mathbb{R}} \psi_k(s) f(x, y-s)\, ds.$$

Proof. The proof is a series of direct computations, and we only present the case (1):

$$\begin{aligned}
&(T_m \otimes \Pi)(f, g)(x, y) \\
&\quad = \int_{\mathbb{R}^{2n+2}} \hat{f}(\xi_1, \eta_1)\hat{g}(\xi_2, \eta_2) m(\xi_1, \xi_2) \\
&\qquad\qquad \Big(\sum_k \hat{\varphi}_k(\eta_1)\hat{\psi}_k(\eta_2)\hat{\psi}_k(\eta_1+\eta_2)\Big) e^{2\pi i x(\xi_1+\xi_2)} e^{2\pi i y(\eta_1+\eta_2)}\, d\xi\, d\eta \\
&\quad = \sum_k \int_{\mathbb{R}^{2n+2}} \hat{f}(\xi_1, \eta_1)\hat{g}(\xi_2, \eta_2) m(\xi_1, \xi_2)\hat{\varphi}_k(\eta_1)\hat{\psi}_k(\eta_2) \\
&\qquad\qquad \Big(\int_{\mathbb{R}} \psi_k(s) e^{-2\pi i s(\eta_1+\eta_2)}\, ds\Big) e^{2\pi i x(\xi_1+\xi_2)} e^{2\pi i y(\eta_1+\eta_2)}\, d\xi\, d\eta \\
&\quad = \sum_k \int_{\mathbb{R}} \psi_k(s)\big(T_m(P_k^{y-s} f, Q_k^{y-s} g)(x)\big)\, ds \\
&\quad = \sum_k Q_k^2 T_m(P_k^y f, Q_k^y g)(x). \qquad \square
\end{aligned}$$

A final ingredient that we will need in the proof of Theorem 6 is the following lemma, which appears in [Ruan 2010]:

Lemma 53. *Let $f \in \mathcal{S}(\mathbb{R}^n)$, and $1 \le l \le n$, and $\{i_1, \dots, i_l\} \subset \{1, \dots, n\}$. Then*

$$\|f\|_{L^p} \lesssim \Big\| \Big(\sum_{k_1,\dots,k_l} \big|Q_{k_1}^{i_1} \cdots Q_{k_l}^{i_l} f\big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^p}$$

for any $0 < p < \infty$.

Lemma 53 above states that the L^p norm of f is bounded by the L^p norm of a square function associated with the variables $x_{i_1}, \dots, x_{i_l}$, even when $0 < p \le 1$. In the case $p > 1$, it is well known that the two norms are equivalent. When $p < 1$, the proof makes use of multiparameter Hardy spaces.

5.1. ***Proof of Theorem 6.*** We start with the proof in the case $\mathrm{BHT} \otimes \Pi$, in order to make the presentation clear.

(a) Assume that $\Pi(f, g) = \sum_k Q_k(P_k f \cdot Q_k g)$. Then Proposition 52 implies that $\mathrm{BHT}\otimes\Pi(f, g)(x, y) = \sum_k Q_k^2 \mathrm{BHT}(P_k^y f, Q_k^y g)(x)$. Lemma 53 yields

$$\|\mathrm{BHT} \otimes \Pi\|_{L^s(\mathbb{R}^2)} \lesssim \Big\| \Big(\sum_k \big| Q_k^2 \mathrm{BHT}(P_k^y f, Q_k^y g) \big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^s(\mathbb{R}^2)}.$$

For the paraproducts that we are considering, $Q_k(P_k f \cdot Q_k g)(y) = P_k f(y) \cdot Q_k g(y)$, so we need to estimate

$$\Big\| \Big(\sum_k \big| \mathrm{BHT}(P_k^y f, Q_k^y g) \big|^2 \Big)^{\frac{1}{2}} \Big\|_{L^s(\mathbb{R}^2)}.$$

We first estimate the L^s norm of $x \mapsto \big(\sum_k |\mathrm{BHT}(P_k^y f, Q_k^y g)(x)|^2\big)^{\frac{1}{2}}$, and Fubini will imply the desired result for $\mathrm{BHT} \otimes \Pi$. Here we use the vector-valued extension for the bilinear Hilbert transform

$$\mathrm{BHT} : L^p(\ell^\infty) \times L^q(\ell^2) \to L^s(\ell^2),$$

which holds whenever $(p, q, s) \in \mathrm{Range(BHT)}$. More exactly,

$$\begin{aligned}
\|\mathrm{BHT} \otimes \Pi\|_{L^s(\mathbb{R}^2)} &\lesssim \Big\| \Big\| \Big(\sum_k \big| \mathrm{BHT}(P_k^y f, Q_k^y g)(x) \big|^2 \Big)^{\frac{1}{2}} \Big\|_{L_x^s} \Big\|_{L_y^s} \\
&\lesssim \Big\| \big\| \sup_k |P_k^y f| \big\|_{L_x^p} \Big\| \Big(\sum_k |Q_k^y g|^2 \Big)^{\frac{1}{2}} \Big\|_{L_x^q} \Big\|_{L_y^s} \\
&\lesssim \Big\| \big\| \sup_k |P_k^y f| \big\|_{L_x^p} \Big\|_{L_y^p} \Big\| \Big\| \Big(\sum_k |Q_k^y g|^2 \Big)^{\frac{1}{2}} \Big\|_{L_x^q} \Big\|_{L_y^q} \\
&\lesssim \|f\|_p \|g\|_q.
\end{aligned}$$

To get the conclusion, we are using Fubini again, and the boundedness of the maximal and square function operators.

(b) The case $\Pi(f, g) = \sum_k P_k(Q_k f, Q_k g)$ is more direct, but the ideas are similar. The functions φ in the paraproduct definition are such that $\Pi(f, g) = \sum_k (Q_k f \cdot Q_k g)$, so we have

$$\mathrm{BHT} \otimes \Pi(f, g)(x, y) = \sum_k \mathrm{BHT}(Q_k^y f, Q_k^y g)(x).$$

Now we use the vector-valued extension $\mathrm{BHT} : L^p(\ell^2) \times L^q(\ell^2) \to L^s(\ell^1)$ (which is well-defined for any $(p, q, s) \in \mathrm{Range(BHT)}$) together with Fubini and the boundedness of the square function to get

$$
\begin{aligned}
\|\mathrm{BHT} \otimes \Pi\|_{L^s(\mathbb{R}^2)} &\lesssim \Big\| \Big\| \sum_k \big|\mathrm{BHT}(Q_k^y f, Q_k^y g)(x)\big| \Big\|_{L_x^s} \Big\|_{L_y^s} \\
&\lesssim \Big\| \Big\| \Big(\sum_k |Q_k^y f|^2\Big)^{\frac{1}{2}} \Big\|_{L_x^p} \Big\| \Big(\sum_k |Q_k^y g|^2\Big)^{\frac{1}{2}} \Big\|_{L_x^q} \Big\|_{L_y^s} \\
&\lesssim \|f\|_p \|g\|_q.
\end{aligned}
$$

The general case of Theorem 6 is similar, but slightly more technical. We present it below for completeness. The paraproducts can be of three types, as seen in (53). This generates a partition of $\{1, \dots, n\}$ into three subsets of indices $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$ so that if $k \in \mathcal{I}_1$, then

$$\Pi(f, g)(y) = \sum_k Q_k(P_k f \cdot Q_k g)(y),$$

and similarly for $\mathcal{I}_2$ and $\mathcal{I}_3$.

Because the projections on different coordinates commute, i.e., $Q_k^i P_l^j = P_l^j Q_k^i$ and $Q_k^i Q_l^j = Q_l^j Q_k^i$, we can assume

$$\mathcal{I}_1 = \{1, \dots, l\}, \quad \mathcal{I}_2 = \{l+1, \dots, l+d\}, \quad \mathcal{I}_3 = \{l+d+1, \dots, n\}.$$

Of course, we allow the possibility that one or even two of these sets of indices are empty. With this assumption, Proposition 52 applied iteratively yields

$$
\begin{aligned}
&\mathrm{BHT} \otimes \Pi \otimes \cdots \otimes \Pi(f, g)(x, y_1, \dots, y_n) \\
&\quad = \sum_{k_1, \dots, k_n} Q_{k_1}^1 \cdots Q_{k_l}^l Q_{k_{l+1}}^{l+1} \cdots Q_{k_{l+d}}^{l+d} P_{k_{l+d+1}}^{l+d+1} \cdots P_{k_n}^n \circ \\
&\qquad\qquad \mathrm{BHT}\big(P_{k_1}^{y_1} \cdots P_{k_l}^{y_l} Q_{k_{l+1}}^{y_{l+1}} \cdots Q_{k_n}^{y_n} f, Q_{k_1}^{y_1} \cdots Q_{k_l}^{y_l} P_{k_{l+1}} \cdots P_{k_{l+d}}^{y_{l+d}} Q_{k_{l+d+1}}^{y_{l+d+1}} \cdots Q_{k_n}^{y_n} g\big)(x).
\end{aligned}
$$

The outer-most expressions $Q_{k_1}^1 \cdots Q_{k_l}^l Q_{k_{l+1}}^{l+1} \cdots Q_{k_{l+d}}^{l+d} P_{k_{l+d+1}}^{l+d+1} \cdots P_{k_n}^n$ are extremely important. Expressions of the type P_k will be associated with ℓ^1 norms, and the Q_k with ℓ^2 norms and square functions. Here we want to apply Lemma 53, so we need to deal with the Q_k functions first. Once we do this, we can estimate the L^r norm of $\mathrm{BHT} \otimes \Pi \otimes \cdots \Pi(f, g)$ by

$$
\begin{aligned}
&\Big\| \Big(\sum_{k_1, \dots, k_{l+d}} \Big| \sum_{k_{l+d+1}, \dots, k_n} P_{k_{l+d+1}}^{l+d+1} \cdots P_{k_n}^n \mathrm{BHT}\big(P_{k_1}^{y_1} \cdots Q_{k_{l+1}}^{y_{l+1}} \cdots f, Q_{k_1}^{y_1} \cdots P_{k_{l+1}}^{y_{l+1}} \cdots Q_{k_{l+d+1}}^{y_{l+d+1}} \cdots g\big) \Big|^2 \Big)^{\frac{1}{2}} \Big\|_r \\
&\qquad \lesssim \Big\| \Big(\sum_{k_1, \dots, k_{l+d}} \Big| \sum_{k_{l+d+1}, \dots, k_n} \big|\mathrm{BHT}\big(P_{k_1}^{y_1} \cdots Q_{k_{l+1}}^{y_{l+1}} \cdots f, Q_{k_1}^{y_1} \cdots P_{k_{l+1}}^{y_{l+1}} \cdots Q_{k_{l+d+1}}^{y_{l+d+1}} \cdots g\big)\big| \Big|^2 \Big)^{\frac{1}{2}} \Big\|_r \\
&\qquad \lesssim \|f\|_p \|g\|_q.
\end{aligned}
$$

For the last part we used the following vector-valued estimates for the BHT:

$$L^p\big(\underbrace{\ell^\infty(\cdots(\ell^\infty}_{l}(\underbrace{\ell^2(\cdots(\ell^2}_{d}(\underbrace{\ell^2(\cdots(\ell^2}_{n-l-d}))\cdots\big)\times L^q\big(\underbrace{\ell^2(\cdots(\ell^2}_{l}(\underbrace{\ell^\infty(\cdots(\ell^\infty}_{d}(\underbrace{\ell^2(\cdots(\ell^2}_{n-l-d}))\cdots\big)$$
$$\mapsto L^s\big(\underbrace{\ell^2(\cdots(\ell^2}_{l}(\underbrace{\ell^2(\cdots(\ell^2}_{d}(\underbrace{\ell^1(\cdots(\ell^1}_{n-l-d}))\cdots\big)$$

together with the boundedness of the maximal operator and square function.

Similarly, we can obtain estimates for $\Pi^{\otimes^{d_1}}\otimes \mathrm{BHT}\otimes \Pi^{\otimes^{d_2}}$ within the same range as that of BHT. Some partial results in mixed norm L^p spaces can be obtained too, but the general case, for arbitrary values of d_1 and d_2 remains open. We present a few particular cases that illustrate the main ideas, without being too technical.

(i) Here, we prove mixed norm L^p estimates for $\Pi_1\otimes \mathrm{BHT}\otimes\Pi_3$, where $\Pi_1=\sum_k Q_k^1(P_k^1\cdot Q_k^1)$, $\Pi_3=\sum_l Q_l^3(Q_l^3\cdot P_l^3)$, and the exponents p_j, q_j are in $[2,\infty)$. We note that

$$\Pi_1\otimes \mathrm{BHT}\otimes\Pi_3(f,g)(x,y,z)=\sum_{k,l} Q_k^1 Q_l^3\mathrm{BHT}(P_k^x Q_l^z f, Q_k^x P_l^z g)(y),$$

and we want to estimate the above expression in the space $\|\cdot\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}$. The key observation is that whenever $1<s_2,s_3<\infty$,

$$\Big\|\sum_{k,l} Q_k^1 Q_l^3 F(x,y,z)\Big\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}\lesssim\Big\|\Big(\sum_{k,l}\big|Q_k^1Q_l^3F(x,y,z)\big|^2\Big)^{\frac12}\Big\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}, \tag{54}$$

which is a Banach-valued equivalent of Lemma 53. This result, for $s_1>1$, can be found in [Fernandez 1987; Rubio de Francia et al. 1986], and it follows from the boundedness of Calderón–Zygmund operators (the dual of the square function is such an operator) on L^p spaces with mixed norms. The proof in the case $s_1\le 1$ is a Banach space adaptation of the proof of Lemma 53. Given the special properties of the Q_k^1 and Q_l^3 operators, we obtain

$$\big\|\Pi_1\otimes \mathrm{BHT}\otimes\Pi_3(f,g)\big\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}\lesssim\Big\|\Big(\sum_{k,l}\big|\mathrm{BHT}(P_k^xQ_l^zf,Q_k^xP_l^zg)(y)\big|^2\Big)^{\frac12}\Big\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}.$$

The multiple vector-valued estimates

$$\mathrm{BHT}: L_y^{p_2}(L_z^{p_3}(\ell^\infty(\ell^2))))\times L_y^{q_2}(L_z^{q_3}(\ell^2(\ell^\infty))))\to L_y^{s_2}(L_z^{s_3}(\ell^2(\ell^2)))),$$

which exist in the local L^2 case at least, together with Hölder's inequality imply

$$\begin{aligned}\big\|\Pi_1\otimes \mathrm{BHT}\otimes\Pi_3(f,g)\big\|_{L_x^{s_1}L_y^{s_2}L_z^{s_3}}&\\ &\lesssim\Big\|\sup_k\Big(\sum_l|P_k^xQ_l^zf(y)|^2\Big)^{\frac12}\Big\|_{L_x^{p_1}L_y^{p_2}L_z^{p_3}}\Big\|\Big(\sum_k\big|\sup_l|Q_k^xP_l^zg(y)|\big|^2\Big)^{\frac12}\Big\|_{L_x^{q_1}L_y^{q_2}L_z^{q_3}}\\ &\lesssim\|f\|_{L_x^{p_1}L_y^{p_2}L_z^{p_3}}\|g\|_{L_x^{q_1}L_y^{q_2}L_z^{q_3}}.\end{aligned}$$

The last inequality follows again from Banach-valued extensions of convolution operators. Since our proof makes use of multiple vector-valued estimates for BHT, we cannot obtain mixed norm L^p estimates for all the exponents in the Banach range. From the above example, one can see that besides the constraints imposed by the square functions and maximal operators, we also need $(p_3, q_3, s_3) \in \mathcal{D}_{p_2,q_2,s_2}$.

(ii) If $d_1 = 0$ and $d_2 = 1$, we have

$$\mathrm{BHT} \otimes \Pi : L_x^{p_1} L_y^{p_2} \times L_x^{q_1} L_y^{q_2} \to L_x^{s_1} L_y^{s_2}$$

whenever $1 < p_2, q_2, s_2 < \infty$, $1 < p_1, q_1 \le \infty$, $\frac{2}{3} < s_1 < \infty$ and $(p_2, q_2, s_2) \in \mathcal{D}_{p_1,q_1,s_1}$.

(iii) If $d_1 = 1$ and $d_2 = 0$, we have

$$\Pi \otimes \mathrm{BHT} : L_x^{p_1} L_y^{p_2} \times L_x^{q_1} L_y^{q_2} \to L_x^{s_1} L_y^{s_2}$$

whenever $1 < p_2, q_2, s_2 < \infty$, $1 < p_1, q_1 \le \infty$, $\frac{1}{2} < s_1 < \infty$. Since the "target" spaces (that is, inner spaces in the mixed norms) are strictly between 1 and ∞, the outer L^∞ cases (that is, $p_1 = \infty$ or $q_1 = \infty$) follow easily from similar estimates on the adjoints.

We note that mixed norm estimates for $\Pi \otimes \mathrm{BHT}$ appear also in [Di Plinio and Ou 2015], where all the inner spaces involved are L^p spaces with $1 < p < \infty$ (in our notation, that means $1 < p_2, q_2, s_2 < \infty$).

6. Leibniz rules: Theorem 4

Now we present some ideas behind the proof of Theorem 4. Littlewood–Paley projections play an important role when dealing with derivatives:

$$\begin{aligned} D_1^\alpha D_2^\beta (f \cdot g)(x,y) &= \sum_{k,l} \big[(f * \varphi_k \otimes \varphi_l) \cdot (g * \psi_k \otimes \psi_l)\big] * (D_1^\alpha \psi_k \otimes D_2^\beta \psi_l)(x,y) \\ &= \sum_{k,l} \big[(f * \varphi_k \otimes \varphi_l) \cdot (g * \psi_k \otimes \psi_l)\big] * (2^{k\alpha} \tilde{\psi}_k \otimes 2^{l\beta} \tilde{\psi}_l)(x,y), \end{aligned}$$

where

$$\widehat{\tilde{\psi}}_k(\xi) = \frac{|\xi|^\alpha}{2^{k\alpha}} \hat{\psi}_k(\xi) \quad \text{and} \quad \widehat{\tilde{\psi}}_l(\eta) = \frac{|\eta|^\beta}{2^{l\beta}} \hat{\psi}_l(\eta).$$

Then one can move the $2^{k\alpha}$ inside, and couple it with the ψ_k because $2^{k\alpha} \psi_k(x) = D^\alpha \tilde{\tilde{\psi}}_k(x)$. Here

$$\widehat{\tilde{\tilde{\psi}}}_k(\xi) = \frac{2^{k\alpha}}{|\xi|^\alpha} \widehat{\psi}_k(\xi).$$

In this way, we obtain $D_1^\alpha D_2^\beta (f \cdot g) = \tilde{\Pi} \otimes \tilde{\Pi}(f, D_1^\alpha D_2^\beta g) +$ eight other similar terms. We can estimate $\Pi \otimes \Pi$ in L^p spaces with mixed norms, as long as the "outside" functions $\widehat{\psi}_k$ and $\widehat{\varphi}_k$ are constantly equal to 1 on $2^{k-2} \le |\xi| \le 2^{k+2}$ and $|\xi| \le 2^{k+2}$ respectively. The operators $\tilde{\Pi}$ are slightly different, but using Fourier series we can write $\tilde{\Pi}(F, G)$ as

$$(F, G) \mapsto \sum_{n \in \mathbb{Z}} c_n \sum_{k,l} \big[F * (\varphi_k \otimes \varphi_l) \cdot G * (\tilde{\tilde{\psi}}_k \otimes \tilde{\tilde{\psi}}_l)\big] * \psi_k \otimes \tilde{\psi}_{l,n}(x,y).$$

Here the coefficients satisfy $|c_n| \lesssim n^{-M}$, and $\psi_{k,n}(x) = \psi_k(x + 2^{-k} n)$. Now notice that the right-hand side above becomes

$$\sum_n c_n \sum_l Q_l^2 \widetilde{\Pi}(P_{l,n}^y F, \tilde{\tilde{Q}}_{l,n}^y G)(x),$$

which is a superposition of $\Pi \otimes \Pi$ operators.

The proof of the Leibniz rule follows from

(1) (multiple) vector-valued estimates for the paraproduct

$$\widetilde{\Pi}(f, g) = \sum_l \big[(f * \varphi_l) \cdot (g * \tilde{\tilde{\psi}}_l)\big] * \tilde{\psi}_l,$$

(2) the boundedness of the shifted maximal and square functions:

$$\big\| \sup_l |f * \varphi_{l,n}| \big\|_p \lesssim \log\langle n \rangle \, \|f\|_p, \qquad \bigg\| \Big(\sum_l |f * \tilde{\tilde{\psi}}_{l,n}|^2 \Big)^{\frac{1}{2}} \bigg\|_p \lesssim \log\langle n \rangle \, \|f\|_p.$$

Returning to the Leibniz rules, we have for $s_1, s_2 \geq 1$,

$$\begin{aligned}
\big\| \|D_1^\alpha D_2^\beta (f, g)\|_{L_y^{s_2}} \big\|_{L_x^{s_1}} &\leq \sum_n |c_n| \bigg\| \Big\| \sum_l Q_l^2 \widetilde{\Pi}(P_{l,n}^y F, \tilde{\tilde{Q}}_{l,n}^y G) \Big\|_{L_y^{s_2}} \bigg\|_{L_x^{s_1}} \\
&\lesssim \sum_n |c_n| \bigg\| \Big\| \Big(\sum_l \big| \widetilde{\Pi}^{\beta_1, \beta_2}(P_{l,n}^y F, \tilde{\tilde{Q}}_{l,n}^y G) \big|^2 \Big)^{\frac{1}{2}} \Big\|_{L_y^{s_2}} \bigg\|_{L_x^{s_1}} \\
&\lesssim \sum_n |c_n| \Big\| \big\| \sup_l |P_{l,n}^y F| \big\|_{L_y^{p_2}} \Big\|_{L_x^{p_1}} \bigg\| \Big\| \Big(\sum_l |\tilde{\tilde{Q}}_{l,n}^y G|^2 \Big)^{\frac{1}{2}} \Big\|_{L_y^{q_2}} \bigg\|_{L_x^{q_1}} \\
&\lesssim \|f\|_{L_x^{p_1} L_y^{p_2}} \|D_1^\alpha D_2^\beta g\|_{L_x^{q_1} L_y^{q_2}}.
\end{aligned}$$

Here we used the vector-valued estimates

$$\widetilde{\Pi} : L_x^{p_1}(L_y^{p_2}(\ell^\infty)) \times L_x^{q_1}(L_y^{q_2}(\ell^2)) \to L_x^{s_1}(L_y^{s_2}(\ell^2)),$$

as well as the boundedness of the square function and maximal operator. We note that the square function is in the y-variable, and for that reason at first we cannot allow $p_2 = \infty$ or $q_2 = \infty$. However, this obstruction can be removed by using duality.

The same proof works in the case $\frac{1}{2} < s_1 < 1$, if $1 < p_2, q_2 < \infty$. In this case, we use the subadditivity of $\| \cdot \|_{s_1}^{s_1}$. The case $\frac{1}{2} < s_1 < 1$ and $p_2 = \infty$ requires a slightly different reasoning, and can be deduced from the corresponding mixed norm estimates for $\Pi \otimes \Pi$. This will be presented at the end of this section.

A slightly more difficult case of the Leibniz rule is when one of the last components is a φ-type function:

$$\begin{aligned}
D_1^\alpha D_2^\beta (f \cdot g)(x, y) &= \sum_{k,l} \big[(f * \psi_k \otimes \varphi_l) \cdot (g * \psi_k \otimes \psi_l)\big] * (D_1^\alpha \varphi_k \otimes D_2^\beta \psi_l)(x, y) \\
&= \sum_{k,l} \big[(f * \psi_k \otimes \varphi_l) \cdot (g * \psi_k \otimes \psi_l)\big] * (2^{k\alpha} \tilde{\varphi}_k \otimes 2^{l\beta} \tilde{\psi}_l)(x, y).
\end{aligned}$$

In this case

$$\hat{\tilde{\varphi}}_k(\xi) = \frac{|\xi|^\alpha}{2^{k\alpha}} \hat{\varphi}_k(\xi),$$

but $\tilde{\varphi}$ doesn't behave as nicely as $\tilde{\psi}$; since $\hat{\tilde{\varphi}}$ is not smooth at the origin, the decay in $\tilde{\varphi}$ is much slower:

$$|\tilde{\varphi}(x)| \leq \frac{1}{(1+|x|)^{1+\alpha}}.$$

We use a Fourier series decomposition of $\hat{\tilde{\varphi}}_k$ on its support

$$\hat{\tilde{\varphi}}_k(\xi) = \sum_{n \in \mathbb{Z}} c_n e^{\frac{2\pi i n \xi}{2^k}} \cdot \hat{\varphi}_k(\xi), \quad \text{where } c_n = \frac{1}{2^k} \int_{\mathbb{R}} \hat{\tilde{\varphi}}_k(\xi) e^{-\frac{2\pi i n \xi}{2^k}} \, d\xi.$$

In this case we only have $|c_n| \leq 1/(1+|n|)^{1+\alpha}$, but this is enough for the coefficients to sum up, if $s_1 > 1/(1+\alpha)$. Since $s_2 \geq 1$, we will not have a similar issue when doing the decomposition in the second variable.

Following the same line of ideas, the problem reduces to estimating

$$\sum_n c_n \sum_k P_k^1 \tilde{\Pi}\big(\tilde{\tilde{Q}}_{k,n}^x F, Q_{k,n}^x G\big)(y),$$

and it would imply "mixed square functions" estimates of the form

$$\Big\| \Big(\sum_n |Q_{k,n}^x G|^2 \Big)^{\frac{1}{2}} \Big\|_{L_x^{q_1} L_y^{q_2}}.$$

This is bounded as long as $1 < q_1, q_2 < \infty$, and in order to recover the case $p_i = \infty$ or $q_i = \infty$ we want to make sure that the square functions are in the innermost variable, which is y. So we need a decomposition of $\tilde{\psi}_l$, as before. Also, we will need vector-valued estimates for the "generalized paraproduct"

$$(f, g) \mapsto \sum_k (f * \psi_k \cdot g * \psi_k) * \tilde{\varphi}_k,$$

where the last component $\tilde{\varphi}$ has slow decay. The vector spaces involved are $(\ell^2, \ell^\infty, \ell^2)$ or (ℓ^2, ℓ^2, ℓ^1), and such estimates can be proved using ideas similar to those in Section 4, modulo standard technical difficulties, as discussed in [Muscalu and Schlag 2013].

We now present the proof of the mixed norm estimates for the biparameter paraproducts:

Proof of Theorem 5. Since the other cases are very similar, we can assume that Π_y, the paraproduct acting on the variable y, is of the form

$$\Pi_y(\cdot\,,\cdot) = \sum_k Q_k\big(P_k(\cdot), Q_k(\cdot)\big).$$

Then we can write $\Pi \otimes \Pi$ as $\Pi \otimes \Pi(f, g)(x, y) = \sum_k Q_k^2 \Pi(P_k^y, Q_k^y)(x)$. Then we have

$$\Big\| \Big\| \sum_k Q_k^2 \Pi(P_k^y, Q_k^y)(x) \Big\|_{L_y^{s_2}} \Big\|_{L_x^{s_1}} \lesssim \Big\| \Big\| \Big(\sum_k \big| \Pi(P_k^y, Q_k^y)(x) \big|^2 \Big)^{\frac{1}{2}} \Big\|_{L_y^{s_2}} \Big\|_{L_x^{s_1}}$$
$$\lesssim \Big\| \big\| \sup_k |P_k^y f(x)| \big\|_{L_y^{p_2}} \Big\|_{L_x^{p_1}} \Big\| \Big\| \Big(\sum_k |Q_k^y g(x)|^2 \Big)^{\frac{1}{2}} \Big\|_{L_y^{q_2}} \Big\|_{L_x^{q_1}}.$$

In the above inequality we used the multiple vector-valued estimate

$$\Pi_x : L_x^{p_1}(L_y^{p_2}(\ell^\infty)) \times L_x^{q_1}(L_y^{q_2}(\ell^2)) \to L_x^{s_1}(L_y^{s_2}(\ell^2)),$$

which is a consequence of Theorem 9.

Now we focus on the case $p_2 = \infty$, $1 < q_2 = q < \infty$, since $q_2 = \infty$ is symmetric. We want to prove that

$$\Pi \otimes \Pi : L_x^{p_1} L_y^\infty \times L_x^{q_1} L_y^q \to L_x^{s_1} L_y^q,$$

by using Banach-valued restricted-type interpolation. That is, for any sets of finite measure F, G, H, we can find a major subset $H' \subseteq H$, and we will prove that

$$\left| \int_{\mathbb{R}^2} \Pi \otimes \Pi(f, g)(x, y) h(x, y)\, dx\, dy \right| \lesssim |F|^{\alpha_1} |G|^{\alpha_2} |H|^{\alpha_3} \tag{55}$$

for any functions f, g and h satisfying

$$\|f(x,\cdot)\|_{L_y^\infty} \le \mathbf{1}_F(x), \quad \|g(x,\cdot)\|_{L_y^q} \le \mathbf{1}_G(x), \quad \|h(x,\cdot)\|_{L_y^{q'}} \le \mathbf{1}_{H'}(x),$$

and $(\alpha_1, \alpha_2, \alpha_3)$ any tuple satisfying $\alpha_1 + \alpha_2 + \alpha_3 = 1$, situated in the neighborhood of $\big(\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{p'}\big)$.

A triple stopping time similar to the one appearing in the proof of Theorem 7 will allow us to recover any exterior $L_x^{p_j}$ norms, while the interior norms are fixed: L_y^∞, L_y^q, L_y^q.

We will consider *localizations* of the paraproduct acting on the x-variable. More exactly, the following estimate, the proof of which is a combination of Proposition 50 and L^p estimates for $\Pi \otimes \Pi$, is key:

If I_0 is a fixed dyadic interval, then $\Pi_{I_0}^{F,G,H'} \otimes \Pi : L_x^\infty L_y^\infty \times L_x^q L_y^q \to L_x^q L_y^q$ with operatorial norm

$$\big\| \Pi_{I_0}^{F,G,H'} \otimes \Pi \big\|_{L_x^\infty L_y^\infty \times L_x^q L_y^q \to L_x^q L_y^q} = \big\| (\Pi_{I_0}^{F,G,H'} \otimes \Pi)^{*,1} \big\|_{L_x^{q'} L_y^{q'} \times L_x^q L_y^q \to L_x^1 L_y^1}.$$

The latter is bounded above by

$$\big\| (\Pi_{I_0}^{F,G,H'} \otimes \Pi)^{*,1} \big\|_{L_x^{q'} L_y^{q'} \times L_x^q L_y^q \to L_x^1 L_y^1} \lesssim (\widetilde{\text{size}}_{I_0} \mathbf{1}_{H'})^{\frac{1}{q} - \epsilon} (\widetilde{\text{size}}_{I_0} \mathbf{1}_G)^{\frac{1}{q'} - \epsilon} (\widetilde{\text{size}}_{I_0} \mathbf{1}_F)^{1-\epsilon},$$

which is a consequence of the localized multiple vector-valued estimates that always appear in the iterative step of the helicoidal method.

More exactly, we have

$$\Big|\Pi_{I_0}^{F,G,H} \otimes \Pi(f,g)(x,y)h(x,y)\,dx\,dy\Big| \\ \lesssim (\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{\frac{1}{q}-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{G})^{\frac{1}{q'}-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{F})^{1-\epsilon} \\ \big\|\|h(x,\cdot)\|_{L_y^{q'}}\cdot\tilde{\chi}_{I_0}\big\|_{L_x^{q'}}\big\|\|g(x,\cdot)\|_{L_y^{q}}\cdot\tilde{\chi}_{I_0}\big\|_{L_x^{q}}\|f(\cdot,\cdot)\|_{L_x^\infty L_y^\infty}.$$

This implies, after performing the usual stopping times, that

$$\left|\int_{\mathbb{R}^2}(\Pi\otimes\Pi)(f,g)(x,y)h(x,y)\,dx\,dy\right| \lesssim \sum_{n_1,n_2,n_3}\sum_{I_0}\left|\int_{\mathbb{R}^2}(\Pi_{I_0}^{F,G,H'}\otimes\Pi)(f,g)(x,y)h(x,y)\,dx\,dy\right| \\ \lesssim \sum_{n_1,n_2,n_3}\sum_{I_0}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{F})^{1-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{G})^{1-\epsilon}(\widetilde{\text{size}}_{I_0}\mathbf{1}_{H'})^{1-\epsilon}|I_0|.$$

From here, the desired L^p estimates follow almost immediately. □

7. Rubio de Francia theorem for iterated Fourier integrals

We end by answering the initial question that motivated the study of vector-valued BHT. More exactly, we prove Theorem 10, which is a consequence of Theorem 7, with r_1, r_2 chosen carefully so that $\frac{1}{r_1}+\frac{1}{r_2}=\frac{1}{r}$.

Proof of Theorem 10. We start with the case $r \geq 2$; this follows from Theorem 7:

$$\left\|\Big(\sum_k |\text{BHT}(P_{I_k}f, P_{I_k}g)(x)|^2\Big)^{\frac{1}{2}}\right\|_s \lesssim \left\|\Big(\sum_k |P_{I_k}f|^{r_1}\Big)^{\frac{1}{r_1}}\right\|_p \left\|\Big(\sum_k |P_{I_k}g|^{r_2}\Big)^{\frac{1}{r_2}}\right\|_q \tag{56}$$

for any $1 < p, q < \infty$, $\frac{2}{3} < s < \infty$.

This is implied by Rubio de Francia's theorem, if one can find r_1 and r_2 with $\frac{1}{r_1}+\frac{1}{r_2}=\frac{1}{2}$ and

$$\tfrac{1}{p} < \tfrac{1}{r_1'}, \quad \tfrac{1}{q} < \tfrac{1}{r_2'}.$$

This is possible as long as $\frac{1}{s}=\frac{1}{p}+\frac{1}{q}<\frac{1}{r_1'}+\frac{1}{r_2'}=\frac{3}{2}$, which coincides with the condition that we have for the range of BHT.

The case $1 \leq r < 2$ is similar; for p, q, and s as above, one needs to find r_1 and $r_2 \geq 2$ so that

$$2-\tfrac{1}{r}=\tfrac{1}{r_1'}+\tfrac{1}{r_2'}>\tfrac{1}{p}+\tfrac{1}{q}.$$

Note that $\frac{1}{p}<\frac{1}{r_1'}=1-\frac{1}{r}+\frac{1}{r_2}\leq\frac{1}{r'}+\frac{1}{2}$, and similarly for q. Because of this restriction, the operator T_r is bounded as long as admissible triple $\big(\frac{1}{p},\frac{1}{q},\frac{1}{s'}\big)$ is in the convex hull of the points

$$(0,0,1), \quad \big(\tfrac{1}{2}+\tfrac{1}{r'},\tfrac{1}{2},-\tfrac{1}{r'}\big), \quad \big(\tfrac{1}{2},\tfrac{1}{2}+\tfrac{1}{r'},-\tfrac{1}{r'}\big), \quad \big(\tfrac{1}{2}+\tfrac{1}{r'},0,\tfrac{1}{2}-\tfrac{1}{r'}\big), \quad \big(0,\tfrac{1}{2}+\tfrac{1}{r'},\tfrac{1}{2}-\tfrac{1}{r'}\big). \quad □$$

Remark. An alternative way of proving the boundedness of T_r within the range mentioned in Theorem 10 is by interpolating between

$$L^{p_1} \times L^{q_1} \to L^{s_1}(\ell^2) \quad \text{with } p_1, q_1, s_1 \text{ in the range of the BHT operator, and} \tag{57}$$

$$L^{p_2} \times L^{q_2} \to L^{s_2}(\ell^1) \quad \text{with } p_2, q_2 > 1,\ s_2 \geq 1. \tag{58}$$

7.1. ***Boundedness of operators M_1 and M_2.*** In what follows we prove the boundedness of operators M_1 and M_2 presented in (14) and (15):

$$M_1(f_1, f_2, g)(\xi) = \sum_{\omega} \int_{\substack{x_1 < x_2 \\ x_1, x_2 \in \omega_L, x_3 \in \omega_R}} \hat{f}_1(x_1) \hat{f}_2(x_2) g(x_3) e^{2\pi i \xi (x_1 + x_2 + x_3)}\, dx_1\, dx_2\, dx_3$$

and

$$M_2(f_1, f_2, g)(\xi) = \sum_{\omega} \int_{\substack{x_1 < L(\omega_L) \\ x_2 \in \omega_L, x_3 \in \omega_R}} \hat{f}_1(x_1) \hat{f}_2(x_2) g(x_3) e^{2\pi i \xi (x_1 + x_2 + x_3)}\, dx_1\, dx_2\, dx_3.$$

For both operators, we are going to use the triangle inequality in L^r, the target space for operators M_1 and M_2. However, if $r < 1$, this inequality is not available anymore for the quasinorm $\|\cdot\|_r$ and instead we use the triangle inequality for $\|\cdot\|_r^r$. This is the only difference between the Banach and quasi-Banach case, and for simplicity we assume $r \geq 1$. Also, as previously stated, we assume $\|g\|_p = 1$.

Proposition 54. *Let $1 < p < 2$ and $\frac{1}{r} = \frac{1}{s} + \frac{1}{p'} = \frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p'}$. Then*

$$\big\| M_1(f_1, f_2, g) \big\|_r \lesssim \|f_1\|_{p_1} \|f_2\|_{p_2} \|g\|_p.$$

Proof. Recall that $\omega \in \mathcal{D}$ is the mesh of dyadic intervals contained in $[0, 1]$, and we identify them with their preimage: $\omega \sim \varphi^{-1}(\omega)$. We rewrite M_1 as

$$M_1(f_1, f_2, g)(\xi) = \sum_{\omega} \mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)(\xi) \cdot \widehat{g \cdot \mathbf{1}_{\omega_R}}(\xi).$$

Then

$$\begin{aligned}
\big\| M_1(f_1, f_2, g) \big\|_r &\lesssim \sum_{k \geq 0} \Big\| \sum_{|\omega| = 2^{-k}} \mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2) \cdot \widehat{g \cdot \mathbf{1}_{\omega_R}} \Big\|_r \\
&\lesssim \sum_{k \geq 0} \Big\| \Big(\sum_{|\omega| = 2^{-k}} \big|\mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)\big|^p \Big)^{\frac{1}{p}} \Big(\sum_{|\omega| = 2^{-k}} \big|\widehat{g \cdot \mathbf{1}_{\omega_R}}\big|^{p'} \Big)^{\frac{1}{p'}} \Big\|_r \\
&\lesssim \sum_{k \geq 0} \Big\| \Big(\sum_{|\omega| = 2^{-k}} \big|\mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)\big|^p \Big)^{\frac{1}{p}} \Big\|_s \Big(\sum_{|\omega| = 2^{-k}} \|\widehat{g \cdot \mathbf{1}_{\omega_R}}\|_{p'}^{p'} \Big)^{\frac{1}{p'}}.
\end{aligned}$$

We estimate $\|\widehat{g \cdot \mathbf{1}_{\omega_R}}\|_{p'} \lesssim \|g \cdot \mathbf{1}_{\omega_R}\|_p = 2^{-\frac{k}{p}}$ using the Hausdorff–Young theorem. Also, there are 2^k dyadic intervals of length 2^{-k} in $[0, 1]$ and because of this

$$\big\| M_1(f_1, f_2, g) \big\|_r \lesssim \sum_{k \geq 0} 2^{-k(\frac{1}{p} - \frac{1}{p'})} \Big\| \Big(\sum_{|\omega| = 2^{-k}} |\mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)|^p \Big)^{\frac{1}{p}} \Big\|_s.$$

If we estimate the last term using the operator T_p directly, we will not obtain the full range stated above, as there will appear extra constraints of the type

$$\frac{1}{p_1}+\frac{1}{p}<\frac{3}{2}, \quad \frac{1}{p_2}+\frac{1}{p}<\frac{3}{2}.$$

Instead, using Hölder and the fact that $1<p<2$, we have

$$\big\|\mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)\big\|_{\ell^p(\omega)} \leq \big\|\mathrm{BHT}(P_{\omega_L} f_1, P_{\omega_L} f_2)\big\|_{\ell^2(\omega)} 2^{k\left(\frac{1}{p}-\frac{1}{2}\right)}.$$

Using the boundedness of T_2, we have $\|M_1(f_1,f_2,g)\|_r \lesssim \sum_{k\geq 0} 2^{-k\left(\frac{1}{2}-\frac{1}{p'}\right)}\|f_1\|_{p_1}\|f_2\|_{p_2}$. □

Proposition 55. *Let* $1<p<2$ *and* $\frac{1}{r}=\frac{1}{s}+\frac{1}{p'}=\frac{1}{p_1}+\frac{1}{p_2}+\frac{1}{p'}$. *Then*

$$\big\|M_2(f_1,f_2,g)\big\|_r \lesssim \|f_1\|_{p_1}\|f_2\|_{p_2}\|g\|_p,$$

provided $\frac{1}{p_2}+\frac{1}{p'}<1$.

Proof. First, we remark that

$$\big|M_2(f_1,f_2,g)(\xi)\big| \leq \sum_{\omega} |Cf_1(\xi)|\,|P_{\omega_L} f_2(\xi)|\,|\widehat{g\omega_R}(\xi)|,$$

where C is the Carleson operator, bounded on L^p whenever $1<p<\infty$. From here on the estimates are similar to those in Proposition 54, but instead of the bilinear operator $T_r(f,g)$ we will have to use the more restrictive Rubio de Francia operator RF_ν:

$$\begin{aligned}
\big\|M_2(f_1,f_2,g)\big\|_r &\leq \sum_{k\geq 0}\bigg\|Cf_1\Big(\sum_{|\omega|=2^{-k}}|P_{\omega_L} f_2|^p\Big)^{\frac{1}{p}}\Big(\sum_{|\omega|=2^{-k}}|\widehat{g\cdot\mathbf{1}_{\omega_R}}|^{p'}\Big)^{\frac{1}{p'}}\bigg\|_r \\
&\leq \sum_{k\geq 0}\|Cf_1\|_{p_1}\bigg\|\Big(\sum_{|\omega|=2^{-k}}|P_{\omega_L} f_2|^p\Big)^{\frac{1}{p}}\bigg\|_{p_2}\Big(\sum_{|\omega|=2^{-k}}\|\widehat{g\cdot\mathbf{1}_{\omega_R}}\|_{p'}^{p'}\Big)^{\frac{1}{p'}} \\
&\leq \sum_{k\geq 0}2^{k\left(\frac{1}{p}-\frac{1}{\nu}\right)}\|Cf_1\|_{p_1}\bigg\|\Big(\sum_{|\omega|=2^{-k}}|P_{\omega_L} f_2|^\nu\Big)^{\frac{1}{\nu}}\bigg\|_{p_2}\Big(\sum_{|\omega|=2^{-k}}\|\widehat{g\cdot\mathbf{1}_{\omega_R}}\|_{p'}^{p'}\Big)^{\frac{1}{p'}} \\
&\leq \sum_{k\geq 0}2^{-k\left(\frac{1}{\nu}-\frac{1}{p'}\right)}\|f_1\|_{p_1}\|\mathrm{RF}_\nu(f_2)\|_{p_2}.
\end{aligned}$$

If $p_2\geq 2$,we can take $\nu=2$ and there are no other restrictions. In the case $p_2<2$, Rubio de Francia requires $\frac{1}{\nu}+\frac{1}{p_2}<1$. This and the condition $\frac{1}{\nu}-\frac{1}{p'}>0$ (so that the geometric series above is finite) can be summarized as $\frac{1}{p_2}+\frac{1}{p'}<1$. □

Acknowledgements

Benea was partially supported by NSF grant DMS 1500262 and ERC project FAnFArE no. 637510; Muscalu was partially supported by NSF grant DMS 1500262.

Note added in proof

We recently improved Theorems 4 and 5, allowing for the exponent s_2 to be < 1. This is a consequence of new multiple quasi-Banach valued inequalities for Π. In [Benea and Muscalu 2016], we also prove multiple quasi-Banach valued inequalities for the bilinear Hilbert transform operator, extending also Theorem 7.

References

[Bateman 2013] M. Bateman, "Single annulus L^p estimates for Hilbert transforms along vector fields", *Rev. Mat. Iberoam.* **29**:3 (2013), 1021–1069. MR 3090145 Zbl 1283.42019

[Bateman and Thiele 2013] M. Bateman and C. Thiele, "L^p estimates for the Hilbert transforms along a one-variable vector field", *Anal. PDE* **6**:7 (2013), 1577–1600. MR 3148061 Zbl 1285.42014

[Benea and Muscalu 2016] C. Benea and C. Muscalu, "Quasi-Banach valued inequalities via the helicoidal method", preprint, 2016. arXiv 1609.01090

[Benedek et al. 1962] A. Benedek, A.-P. Calderón, and R. Panzone, "Convolution operators on Banach space valued functions", *Proc. Nat. Acad. Sci. U.S.A.* **48** (1962), 356–365. MR 0133653 Zbl 0103.33402

[Bourgain 1986] J. Bourgain, "Vector-valued singular integrals and the H^1-BMO duality", pp. 1–19 in *Probability theory and harmonic analysis* (Cleveland, Ohio, 1983), Monogr. Textbooks Pure Appl. Math. **98**, Dekker, New York, 1986. MR 830227 Zbl 0602.42015

[Burkholder 1983] D. L. Burkholder, "A geometric condition that implies the existence of certain singular integrals of Banach-space-valued functions", pp. 270–286 in *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I* (Chicago, IL, 1981), Wadsworth, Belmont, CA, 1983. MR 730072

[Carleson 1966] L. Carleson, "On convergence and growth of partial sums of Fourier series", *Acta Math.* **116** (1966), 135–157. MR 0199631 Zbl 0144.06402

[Christ and Kiselev 1998] M. Christ and A. Kiselev, "Absolutely continuous spectrum for one-dimensional Schrödinger operators with slowly decaying potentials: some optimal results", *J. Amer. Math. Soc.* **11**:4 (1998), 771–797. MR 1621882 Zbl 0899.34051

[Christ and Kiselev 2001a] M. Christ and A. Kiselev, "Maximal functions associated to filtrations", *J. Funct. Anal.* **179**:2 (2001), 409–425. MR 1809116 Zbl 0974.47025

[Christ and Kiselev 2001b] M. Christ and A. Kiselev, "WKB asymptotic behavior of almost all generalized eigenfunctions for one-dimensional Schrödinger operators with slowly decaying potentials", *J. Funct. Anal.* **179**:2 (2001), 426–447. MR 1809117 Zbl 0985.34078

[Culiuc et al. 2016] A. Culiuc, F. Di Plinio, and Y. Ou, "Domination of multilinear singular integrals by positive sparse forms", 2016. arXiv 1603.05317

[Demeter and Silva 2015] C. Demeter and P. Silva, "Some new light on a few classical results", *Colloq. Math.* **140**:1 (2015), 129–147. MR 3352435 Zbl 1325.42010

[Di Plinio and Ou 2015] F. Di Plinio and Y. Ou, "Banach-valued multilinear singular integrals", preprint, 2015. arXiv 1506.05827

[Fernandez 1987] D. L. Fernandez, "Vector-valued singular integral operators on L^p-spaces with mixed norms and applications", *Pacific J. Math.* **129**:2 (1987), 257–275. MR 909030 Zbl 0634.42014

[García-Cuerva and Rubio de Francia 1985] J. García-Cuerva and J. L. Rubio de Francia, *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies **116**, North-Holland, Amsterdam, 1985. MR 807149 Zbl 0578.46046

[Grafakos and Li 2006] L. Grafakos and X. Li, "The disc as a bilinear multiplier", *Amer. J. Math.* **128**:1 (2006), 91–119. MR 2197068 Zbl 1143.42015

[Hytönen and Lacey 2013] T. P. Hytönen and M. T. Lacey, "Pointwise convergence of vector-valued Fourier series", *Math. Ann.* **357**:4 (2013), 1329–1361. MR 3124934 Zbl 1285.42005

[Hytönen et al. 2013] T. P. Hytönen, M. T. Lacey, and I. Parissis, "The vector valued quartile operator", *Collect. Math.* **64**:3 (2013), 427–454. MR 3084406 Zbl 1288.42008

[Journé 1985] J.-L. Journé, "Calderón–Zygmund operators on product spaces", *Rev. Mat. Iberoamericana* **1**:3 (1985), 55–91. MR 836284 Zbl 0634.42015

[Kenig et al. 1993] C. E. Kenig, G. Ponce, and L. Vega, "Well-posedness and scattering results for the generalized Korteweg–de Vries equation via the contraction principle", *Comm. Pure Appl. Math.* **46**:4 (1993), 527–620. MR 1211741 Zbl 0808.35128

[Kesler 2015] R. Kesler, "Mixed estimates for degenerate multi-linear operators associated to simplexes", *J. Math. Anal. Appl.* **424**:1 (2015), 344–360. MR 3286565 Zbl 1305.47031

[Lacey and Thiele 1999] M. Lacey and C. Thiele, "On Calderón's conjecture", *Ann. of Math.* (2) **149**:2 (1999), 475–496. MR 1689336 Zbl 0934.42012

[Lerner 2013] A. K. Lerner, "A simple proof of the A_2 conjecture", *Int. Math. Res. Not.* **2013**:14 (2013), 3159–3170. MR 3085756 Zbl 1318.42018

[Meyer and Coifman 1997] Y. Meyer and R. Coifman, *Wavelets: Calderón–Zygmund and multilinear operators*, Cambridge Studies in Advanced Mathematics **48**, Cambridge University Press, 1997. MR 1456993 Zbl 0916.42023

[Muscalu and Schlag 2013] C. Muscalu and W. Schlag, *Classical and multilinear harmonic analysis. Vol. II*, Cambridge Studies in Advanced Mathematics **138**, Cambridge University Press, 2013. MR 3052499 Zbl 1282.42001

[Muscalu et al. 2004a] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, "Bi-parameter paraproducts", *Acta Math.* **193**:2 (2004), 269–296. MR 2134868 Zbl 1087.42016

[Muscalu et al. 2004b] C. Muscalu, T. Tao, and C. Thiele, "L^p estimates for the biest, II: The Fourier case", *Math. Ann.* **329**:3 (2004), 427–461. MR 2127985 Zbl 1073.42010

[Muscalu et al. 2006a] C. Muscalu, J. Pipher, T. Tao, and C. Thiele, "Multi-parameter paraproducts", *Rev. Mat. Iberoam.* **22**:3 (2006), 963–976. MR 2320408 Zbl 1114.42005

[Muscalu et al. 2006b] C. Muscalu, T. Tao, and C. Thiele, "The bi-Carleson operator", *Geom. Funct. Anal.* **16**:1 (2006), 230–277. MR 2221256 Zbl 1130.42004

[Oberlin et al. 2012] R. Oberlin, A. Seeger, T. Tao, C. Thiele, and J. Wright, "A variation norm Carleson theorem", *J. Eur. Math. Soc.* (*JEMS*) **14**:2 (2012), 421–464. MR 2881301 Zbl 1246.42016

[Paley 1931] R. E. A. C. Paley, "Some theorems on orthonormal functions", *Studia Math.* **3** (1931), 226–238. Zbl 0003.35201

[Ruan 2010] Z. Ruan, "Multi-parameter Hardy spaces via discrete Littlewood–Paley theory", *Anal. Theory Appl.* **26**:2 (2010), 122–139. MR 2653706 Zbl 1224.42039

[Rubio de Francia 1985] J. L. Rubio de Francia, "A Littlewood–Paley inequality for arbitrary intervals", *Rev. Mat. Iberoamericana* **1**:2 (1985), 1–14. MR 850681 Zbl 0611.42005

[Rubio de Francia et al. 1986] J. L. Rubio de Francia, F. J. Ruiz, and J. L. Torrea, "Calderón–Zygmund theory for operator-valued kernels", *Adv. in Math.* **62**:1 (1986), 7–48. MR 859252 Zbl 0627.42008

[Silva 2014] P. Silva, "Vector-valued inequalities for families of bilinear Hilbert transforms and applications to bi-parameter problems", *J. Lond. Math. Soc.* (2) **90**:3 (2014), 695–724. MR 3291796 Zbl 1311.42021

[Thiele 2006] C. Thiele, *Wave packet analysis*, CBMS Regional Conference Series in Mathematics **105**, American Mathematical Society, Providence, RI, 2006. MR 2199086 Zbl 1101.42001

[Torres and Ward 2015] R. H. Torres and E. L. Ward, "Leibniz's rule, sampling and wavelets on mixed Lebesgue spaces", *J. Fourier Anal. Appl.* **21**:5 (2015), 1053–1076. MR 3393695 Zbl 06493292

Received 20 Jan 2016. Revised 12 Jun 2016. Accepted 12 Jul 2016.

CRISTINA BENEA: cristina.benea@univ-nantes.fr
Université de Nantes, Laboratoire de Mathématiques, Nantes 44322, France

CAMIL MUSCALU: camil@math.cornell.edu
Department of Mathematics, Cornell University, Ithaca, NY 14853, United States

Analysis & PDE

msp.org/apde

See inside back cover or msp.org/apde for submission instructions.

The subscription price for 2016 is US $235/year for the electronic version, and $430/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to MSP.

Analysis & PDE (ISSN 1948-206X electronic, 2157-5045 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840, is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

APDE peer review and production are managed by EditFlow® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ANALYSIS & PDE

Volume 9 No. 8 2016

2157-5045(2016)9:8;1-4