Acta Math., 193 (2004), 269–296

Bi-parameter paraproducts

by

CAMIL MUSCALU

Cornell University
Ithaca, NY, U.S.A.

JILL PIPHER

Brown University
Providence, RI, U.S.A.

TERENCE TAO and CHRISTOPH THIELE

University of California
Los Angeles, CA, U.S.A.

University of California
Los Angeles, CA, U.S.A.

1. Introduction

Let $f\in\mathcal{S}(\mathbf{R}^2)$ be a Schwartz function in the plane. A well-known inequality in elliptic partial differential equations says that

$$\left\|\frac{\partial^2 f}{\partial x_1\partial x_2}\right\|_p\lesssim\|\Delta f\|_p \tag{1}$$

for $1<p<\infty$, where

$$\Delta=\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}$$

is the Laplace operator.

To prove (1) one just has to observe that

$$\frac{\partial^2 f}{\partial x_1\partial x_2}=cR_1R_2\Delta f,$$

where

$$R_jf(x)=\int_{\mathbf{R}^2}\frac{\xi_j}{|\xi|}\hat{f}(\xi)e^{2\pi ix\xi}\,d\xi,\quad j=1,2,$$

are the Riesz transforms, and they are bounded linear operators on $L^p(\mathbf{R}^2)$ [18].

An estimate of a similar flavour in non-linear partial differential equations is the following inequality of Kato and Ponce [9]. If $f,g\in\mathcal{S}(\mathbf{R}^2)$ and $\widehat{\mathcal{D}^\alpha f}(\xi):=|\xi|^\alpha\hat{f}(\xi)$, $\alpha>0$, is the homogeneous derivative, then

$$\|\mathcal{D}^\alpha(fg)\|_r\lesssim\|\mathcal{D}^\alpha f\|_p\|g\|_q+\|f\|_p\|\mathcal{D}^\alpha g\|_q \tag{2}$$

for $1<p,q\leqslant\infty$, $1/r=1/p+1/q$ and $0<r<\infty$.

Heuristically, if f oscillates more rapidly than g, then g is essentially constant with respect to f, and so $\mathcal{D}^\alpha(fg)$ behaves like $(\mathcal{D}^\alpha f)g$. Similarly, if g oscillates more rapidly than f, then one expects $\mathcal{D}^\alpha(fg)$ to be like $f(\mathcal{D}^\alpha g)$, and this is why there are two terms on the right-hand side of (2). In order to make this argument rigorous, one needs to recall the classical Coifman–Meyer theorem [7], [11], [13]. Let m be a bounded function on $\mathbf{R}^4$, smooth away from the origin and satisfying

$$|\partial^\beta m(\gamma)|\lesssim\frac{1}{|\gamma|^{|\beta|}} \tag{3}$$

for sufficiently many β. Denote by $T_m(f,g)$ the bilinear operator defined by

$$T_m(f,g)(x)=\int_{\mathbf{R}^4} m(\xi,\eta)\hat{f}(\xi)\hat{g}(\eta)e^{2\pi ix(\xi+\eta)}\,d\xi\,d\eta. \tag{4}$$

Then, T_m maps $L^p\times L^q\to L^r$ as long as $1<p,q\leqslant\infty$, $1/r=1/p+1/q$ and $0<r<\infty$.

This operator takes care of the inequality (2) in essentially the same way in which the Riesz transforms take care of (1). The details will be presented later on in the appendix (see also [9]).

But sometimes (see [10]) in non-linear partial differential equations one faces the situation when a partial differential operator such as

$$(\mathcal{D}_1^\alpha\mathcal{D}_2^\beta f)\widehat{\ }(\xi_1,\xi_2):=|\xi_1|^\alpha|\xi_2|^\beta\hat{f}(\xi_1,\xi_2),\quad \alpha,\beta>0,$$

acts on a nonlinear expression such as the product of two functions. It is therefore natural to ask if there is an inequality analogous to (2) for these operators. The obvious candidate, according to the same heuristics, is the inequality

$$\|\mathcal{D}_1^\alpha\mathcal{D}_2^\beta(fg)\|_r\lesssim\|\mathcal{D}_1^\alpha\mathcal{D}_2^\beta f\|_p\|g\|_q+\|f\|_p\|\mathcal{D}_1^\alpha\mathcal{D}_2^\beta g\|_q+\|\mathcal{D}_1^\alpha f\|_p\|\mathcal{D}_2^\beta g\|_q+\|\mathcal{D}_1^\alpha g\|_p\|\mathcal{D}_2^\beta f\|_q. \tag{5}$$

If one tries to prove it, one realizes that one needs to understand bilinear operators whose symbols satisfy estimates of the form

$$|\partial_{\xi_1}^{\alpha_1}\partial_{\xi_2}^{\alpha_2}\partial_{\eta_1}^{\beta_1}\partial_{\eta_2}^{\beta_2}m(\xi,\eta)|\lesssim\frac{1}{|(\xi_1,\eta_1)|^{\alpha_1+\beta_1}}\,\frac{1}{|(\xi_2,\eta_2)|^{\alpha_2+\beta_2}}. \tag{6}$$

Clearly, the class of symbols verifying (6) is strictly wider than the class of symbols satisfying (3). These new m's behave as if they were products of two homogeneous symbols of type (3), one of the variables (ξ_1,η_1) and the other of the variables (ξ_2,η_2).

The main task of the present paper is to prove L^p-estimates for such operators in this more delicate product setting. Our main theorem is the following result:

THEOREM 1.1. *If m is a symbol in $\mathbf{R}^4$ satisfying* (6), *then the bilinear operator T_m defined by* (4) *maps $L^p \times L^q \to L^r$ as long as $1<p,q\leqslant\infty$, $1/r=1/p+1/q$ and $0<r<\infty$.*

It will be clear from the proof of the theorem that the n-linear analogue of this result is also true (see §8 for a precise statement). Particular cases of this theorem have been considered by Journé (see [8] and also [3]), who proved that in the situation of tensor products of two generic paraproducts, one has $L^2 \times L^\infty \to L^2$ estimates. Our approach is different from his and is based on arguments with a strong geometric structure. The reader will notice that part of the difficulties of the general case comes from the fact that there is no analogue of the classical Calderón–Zygmund decomposition in this bi-parameter framework, and so the standard argument [7], [11], [13] used to prove such estimates has to be changed.

The paper is organized as follows. In the next section, we discretize our operator and reduce it to a bi-parameter general paraproduct. In the third section we present a new proof of the classical one-parameter case. This technique will be very helpful to handle an error term later on in §6. §§4, 5 and 6 are devoted to the proof of our main theorem (Theorem 1.1). §7 contains a counterexample to the boundedness of the double bilinear Hilbert transform, and then, the paper ends with some further comments and open questions. In the appendix we explain how Theorem 1.1 implies inequality (5).

Acknowledgement. We would like to express our thanks to Carlos Kenig for valuable conversations and to the referees for their suggestions, which improved the presentation of the paper.

The first two authors were partially supported by NSF grants. The third author is a Clay Prize Fellow and is partially supported by a Packard Foundation grant. The fourth author was partially supported by the NSF grants DMS 9985572 and DMS 9970469.

2. Reduction to bi-parameter paraproducts

In order to understand the operator T_m, the plan is to carve it into smaller pieces well adapted to its bi-parameter structure. First, by writing the characteristic functions of the planes (ξ_1,η_1) and (ξ_2,η_2) as finite sums of smoothed versions of characteristic functions of cones of the form $\{(\xi,\eta):|\xi|\leqslant C|\eta|\}$ or $\{(\xi,\eta):|\xi|\geqslant C|\eta|\}$, we decompose our operator into a finite sum of several parts. Since all the operators obtained in this decomposition can be treated in the same way, we will discuss in detail only one of them, which will be carefully defined below (in fact, as the reader will notice, the only difference between any arbitrary case and the one we will explain here is that the functions MM, SS, MS and SM defined in §2 have to be moved around).

Let ϕ and ψ be two Schwartz bumps on $[0,1]$, symmetric with respect to the origin and such that $\operatorname{supp}\hat{\phi}\subseteq[-\frac{1}{4},\frac{1}{4}]$ and $\operatorname{supp}\hat{\psi}\subseteq[\frac{3}{4},\frac{5}{4}]$. Recall the translation and dilation operators τ_h and D^p_λ given by

$$\tau_h f(x)=f(x-h),$$
$$D^p_\lambda f(x)=\lambda^{-1/p}f(\lambda^{-1}x),$$

and then define

$$C'(\xi_1,\eta_1)=\int_{\mathbf{R}} D^\infty_{2^{k'}}\hat{\phi}(\xi_1)\,D^\infty_{2^{k'}}\hat{\psi}(\eta_1)\,dk'$$

and

$$C''(\xi_2,\eta_2)=\int_{\mathbf{R}} D^\infty_{2^{k''}}\hat{\psi}(\xi_2)\,D^\infty_{2^{k''}}\hat{\phi}(\eta_2)\,dk''.$$

As we said, we will now study the operator whose symbol is $mC'C''$. It can be written as

$$\begin{aligned}T_{mC'C''}(f_1,f_2)(x)&=\int_{\mathbf{R}^6} m(\xi,\eta)\,D^\infty_{2^{k'}}\hat{\phi}(\xi_1)\,D^\infty_{2^{k'}}\hat{\psi}(\eta_1)\,D^\infty_{2^{k''}}\hat{\psi}(\xi_2)\,D^\infty_{2^{k''}}\hat{\phi}(\eta_2)\\&\qquad\times\hat{f}_1(\xi_1,\xi_2)\,\hat{f}_2(\eta_1,\eta_2)\,e^{2\pi i x(\xi+\eta)}\,d\xi\,d\eta\,dk'\,dk''\\&=\int_{\mathbf{R}^6} m(\xi,\eta)\,\widehat{\Phi}_{1,k',k''}(\xi_1,\xi_2)\,\widehat{\Phi}_{2,k',k''}(\eta_1,\eta_2)\\&\qquad\times\hat{f}_1(\xi_1,\xi_2)\,\hat{f}_2(\eta_1,\eta_2)\,e^{2\pi i x(\xi+\eta)}\,d\xi\,d\eta\,dk'\,dk''\\&=\int_{\mathbf{R}^6} m(\xi,\eta)\,(f_1*\Phi_{1,k',k''})\widehat{\ }(\xi)\,(f_2*\Phi_{2,k',k''})\widehat{\ }(\eta)\,e^{2\pi i x(\xi+\eta)}\,d\xi\,d\eta\,dk'\,dk'',\end{aligned}$$

where $\Phi_{1,k',k''}:=D^1_{2^{-k'}}\phi\otimes D^1_{2^{-k''}}\psi$ and $\Phi_{2,k',k''}:=D^1_{2^{-k'}}\psi\otimes D^1_{2^{-k''}}\phi$.

In particular, the trilinear form

$$\Lambda_{mC'C''}(f_1,f_2,f_3):=\int_{\mathbf{R}^2} T_{mC'C''}(f_1,f_2)(x)\,f_3(x)\,dx$$

associated to it can be written as

$$\begin{aligned}\int_{\xi+\eta+\gamma=0} & m_{k',k''}(\xi,\eta,\gamma)\,(f_1*\Phi_{1,k',k''})\widehat{\ }(\xi)\\&\times(f_2*\Phi_{2,k',k''})\widehat{\ }(\eta)\,(f_3*\Phi_{3,k',k''})\widehat{\ }(\gamma)\,d\xi\,d\eta\,d\gamma\,dk'\,dk'',\end{aligned}\tag{7}$$

where $\Phi_{3,k',k''}:=D^1_{2^{-k'}}\psi'\otimes D^1_{2^{-k''}}\psi'$ and ψ' is again a Schwartz function such that $\operatorname{supp}\hat{\psi}'\subseteq[-\frac{7}{4},-\frac{1}{4}]$ and $\hat{\psi}'=1$ on $[-\frac{3}{2},-\frac{1}{2}]$, while $m_{k',k''}(\xi,\eta,\gamma)=m(\xi,\eta)\,\lambda_{k',k''}(\xi,\eta,\gamma)$, where $\lambda_{k',k''}(\xi,\eta,\gamma)$ is a smooth function supported on

$$2\operatorname{supp}(\widehat{\Phi}_{1,k',k''}(\xi)\,\widehat{\Phi}_{2,k',k''}(\eta)\,\widehat{\Phi}_{3,k',k''}(\gamma)),$$

which equals 1 on $\operatorname{supp}(\widehat{\Phi}_{1,k',k''}(\xi)\,\widehat{\Phi}_{2,k',k''}(\eta)\,\widehat{\Phi}_{3,k',k''}(\gamma))$.

Then, we write (7) as

$$\begin{aligned}
&\int_{\mathbf{R}^{10}} \widetilde{m}_{k',k''}((n_1',n_1''),(n_2',n_2''),(n_3',n_3''))\\
&\quad\times\prod_{j=1}^{3}(f_j*\Phi_{j,k',k''})((x',x'')-(n_j',n_j''))\,dn_j'\,dn_j''\,dx'\,dx''\,dk'\,dk''\\
&\quad=\int_{\mathbf{R}^{10}} 2^{-4k'}2^{-4k''}\widetilde{m}_{k',k''}((2^{-k'}n_1',2^{-k''}n_1''),(2^{-k'}n_2',2^{-k''}n_2''),(2^{-k'}n_3',2^{-k''}n_3''))\\
&\qquad\times\prod_{j=1}^{3}(f_j*\Phi_{j,k',k''})((2^{-k'}x',2^{-k''}x'')-(2^{-k'}n_j',2^{-k''}n_j''))\,dn_j'\,dn_j''\,dx'\,dx''\,dk'\,dk''\\
&\quad=\int_{\mathbf{R}^{10}} 2^{-3k'}2^{-3k''}\widetilde{m}_{k',k''}((2^{-k'}n_1',2^{-k''}n_1''),(2^{-k'}n_2',2^{-k''}n_2''),(2^{-k'}n_3',2^{-k''}n_3''))\\
&\qquad\times 2^{k'/2}2^{k''/2}\prod_{j=1}^{3}\langle f_j,\Phi_{j,\vec{k},\vec{x},\vec{n}_j}\rangle\,d\vec{n}_j\,d\vec{x}\,d\vec{k},
\end{aligned}$$

where

$$\Phi_{j,\vec{k},\vec{x},\vec{n}_j} := 2^{-k'/2}2^{-k''/2}\tau_{(2^{-k'}x',2^{-k''}x'')-(2^{-k'}n_j',2^{-k''}n_j'')}\Phi_{j,k',k''}.$$

Notice that our functions $\Phi_{j,\vec{k},\vec{x},\vec{n}_j}$ are now $L^2(\mathbf{R}^2)$-normalized. The above expression can be discretized as

$$\sum_{(\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l})\in\mathbf{Z}^{10}} \Lambda_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}(f_1,f_2,f_3), \tag{8}$$

where

$$\begin{aligned}
\Lambda_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}(f_1,f_2,f_3) := \int_{[0,1]^{10}} & 2^{-3(k'+\varkappa')}2^{-3(k''+\varkappa'')}\widetilde{m}_{k'+\varkappa',k''+\varkappa''}(\dots)\\
&\times 2^{(k'+\varkappa')/2}2^{(k''+\varkappa'')/2}\prod_{j=1}^{3}\langle f_j,\Phi_{j,\vec{k}+\vec{\varkappa},\vec{l}+\vec{\lambda},\vec{n}_j+\vec{\nu}_j}\rangle\,d\vec{\nu}_j\,d\vec{\varkappa}\,d\vec{\lambda}.
\end{aligned}$$

Consequently, the operator $T_{mC'C''}(f_1,f_2)$ splits as

$$T_{mC'C''}(f_1,f_2) = \sum_{(\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l})\in\mathbf{Z}^{10}} T_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}(f_1,f_2),$$

where $T_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}$ is the operator whose trilinear form is $\Lambda_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}$. Clearly, by Fatou's lemma it is enough to prove estimates for the operator

$$\sum_{\substack{(\vec{n}_1,\vec{n}_2,\vec{n}_3)\in\mathbf{Z}^6\\ |\vec{k}|,|\vec{l}|<N}} T_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}(f_1,f_2), \tag{9}$$

as long as they are independent of the constant N. Now fix a large constant N and write (9) as

$$\sum_{(\vec{n}_1,\vec{n}_2,\vec{n}_3)\in\mathbf{Z}^6}\Big(\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}(f_1,f_2)\Big). \tag{10}$$

We also observe that by using (6) and integrating by parts several times, we have

$$\begin{aligned}|2^{-3k'}2^{-3k''}\widetilde{m}_{k',k''}((2^{-k'}n_1',2^{-k''}n_1''),(2^{-k'}n_2',2^{-k''}n_2''),(2^{-k'}n_3',2^{-k''}n_3''))|\\ \lesssim\prod_{j=1}^{3}\frac{1}{(1+|\vec{n}_j|)^M}\end{aligned} \tag{11}$$

for M arbitrarily large.

We are going to prove explicitly that the operator

$$\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{0},\vec{0},\vec{0},\vec{k},\vec{l}}(f_1,f_2)=:\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{k},\vec{l}}(f_1,f_2) \tag{12}$$

satisfies the required estimates. It will be clear from the proof and (11) that the same arguments give

$$\Big\|\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{n}_1,\vec{n}_2,\vec{n}_3,\vec{k},\vec{l}}\Big\|_{L^p\times L^q\to L^r}\lesssim\prod_{j=1}^{3}\frac{1}{(1+|\vec{n}_j|)^{100}}\Big\|\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{k},\vec{l}}\Big\|_{L^p\times L^q\to L^r} \tag{13}$$

for any $(\vec{n}_1,\vec{n}_2,\vec{n}_3)\in\mathbf{Z}^6$. Together with (10) this would prove our desired estimates. It is therefore enough to deal with

$$\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{k},\vec{l}}(f_1,f_2).$$

Fix now p and q, two numbers bigger than 1 and very close to 1. Let also f_1 and f_2 be such that $\|f_1\|_p=\|f_2\|_q=1$. We will show that

$$\Big\|\sum_{|\vec{k}|,|\vec{l}|<N} T_{\vec{k},\vec{l}}(f_1,f_2)\Big\|_{r,\infty}\lesssim 1, \tag{14}$$

where $1/r=1/p+1/q$.

Using Lemma 5.4 in [1] and scaling invariance, it is enough to show that for every set $E_3\subseteq\mathbf{R}^2$, $|E_3|=1$, one can find a subset $E_3'\subseteq E_3$ with $|E_3'|\sim 1$ and such that

$$\Big|\sum_{|\vec{k}|,|\vec{l}|<N}\Lambda_{\vec{k},\vec{l}}(f_1,f_2,f_3)\Big|\lesssim 1, \tag{15}$$

where $f_3 := \chi_{E_3'}$. If this is true, then by using the symmetry of our form, the symmetry of our arguments plus multilinear interpolation as in [14], we would complete the proof.

In order to construct the set E_3' we need to define the *maximal-square function* and the *square-maximal function* as follows.

For $(x', x'') \in \mathbf{R}^2$ let

$$\mathrm{MS}(f_1)(x', x'') := \sup_{k',l'} \frac{1}{2^{-k_1'/2}} \left(\sum_{k'',l''} \sup_{\vec{\varkappa},\vec{\lambda},\vec{\nu}_1} \frac{|\langle f_1, \Phi_{1,\vec{k}+\vec{\varkappa},\vec{l}+\vec{\lambda},\vec{\nu}_1} \rangle|^2}{2^{-k''}} 1_{I_{k'',l''}}(x'') \right)^{1/2} 1_{I_{k',l'}}(x')$$

and

$$\mathrm{SM}(f_2)(x', x'') := \left(\sum_{k',l'} \frac{\left(\sup_{k'',l''} \sup_{\vec{\varkappa},\vec{\lambda},\vec{\nu}_2} \frac{|\langle f_2, \Phi_{2,\vec{k}+\vec{\varkappa},\vec{l}+\vec{\lambda},\vec{\nu}_2} \rangle|}{2^{-k''/2}} 1_{I_{k'',l''}}(x'') \right)^2}{2^{-k'}} 1_{I_{k',l'}}(x') \right)^{1/2}.$$

Then, we also define the *double square function*,

$$\mathrm{SS}(f_3)(x', x'') := \left(\sum_{k',l',k'',l''} \sup_{\vec{\varkappa},\vec{\lambda},\vec{\nu}_3} \frac{|\langle f_3, \Phi_{3,\vec{k}+\vec{\varkappa},\vec{l}+\vec{\lambda},\vec{\nu}_3} \rangle|^2}{2^{-k'}2^{-k''}} 1_{I_{k',l'}}(x') 1_{I_{k'',l''}}(x'') \right)^{1/2},$$

where in general, $I_{k,l}$ is the dyadic interval $2^{-k}[l, l+1]$. Finally, we recall the bi-parameter Hardy–Littlewood maximal function

$$\mathrm{MM}(g)(x', x'') := \sup_{(x',x'') \in R} \frac{1}{|R|} \int_R |g(y', y'')| \, dy' \, dy'',$$

where R ranges over all rectangles in the plane whose sides are parallel to the coordinate axes.

The reader should not worry too much about the presence of the suprema over $\varkappa$, λ, ν_1, ν_2 and ν_3 in the above definitions. They need to be there for some technical reasons, but their appearance is completely harmless from the point of view of the boundedness of the corresponding operators.

It is well known that both the bi-parameter maximal function MM and the double square function SS map $L^p(\mathbf{R}^2)$ into $L^p(\mathbf{R}^2)$ whenever $1<p<\infty$, see [2].

Similarly, it is not difficult to observe, by using Fubini's theorem and the Fefferman–Stein inequality [5], that the operators MS and SM are also bounded on $L^p(\mathbf{R}^2)$ if $1<p<\infty$ (first, one treats the SM-function iteratively, as we said, and then one simply observes that the MS-function is pointwise smaller than SM).

We then set

$$\begin{aligned} \Omega_0 = \{x \in \mathbf{R}^2 : \mathrm{MS}(f_1)(x) > C\} &\cup \{x \in \mathbf{R}^2 : \mathrm{SM}(f_2)(x) > C\} \\ &\cup \{x \in \mathbf{R}^2 : \mathrm{MM}(f_1)(x) > C\} \\ &\cup \{x \in \mathbf{R}^2 : \mathrm{MM}(f_2)(x) > C\}. \end{aligned}$$

Also, define

$$\Omega = \{x \in \mathbf{R}^2 : \mathrm{MM}(1_{\Omega_0})(x) > \tfrac{1}{100}\} \tag{16}$$

and finally

$$\widetilde{\Omega} = \{x \in \mathbf{R}^2 : \mathrm{MM}(1_{\Omega})(x) > \tfrac{1}{2}\}.$$

Clearly, we have $|\widetilde{\Omega}| < \frac{1}{2}$, if C is a big enough constant, which we fix from now on. Then, we define $E_3' := E_3 \setminus \widetilde{\Omega} = E_3 \cap \widetilde{\Omega}^c$ and observe that $|E_3'| \sim 1$.

Since the form $\sum_{|\vec{k}|,|\vec{l}|<N} \Lambda_{\vec{k},\vec{l}}(f_1, f_2, f_3)$ is an average of some other forms depending on the parameters $(\vec{\varkappa}, \vec{\lambda}, \vec{\nu}_1, \vec{\nu}_2, \vec{\nu}_3) \in [0,1]^{10}$, it is enough to prove our inequality (15) for each of them, uniformly with respect to $(\vec{\varkappa}, \vec{\lambda}, \vec{\nu}_1, \vec{\nu}_2, \vec{\nu}_3)$. We will do this in the particular case when all these parameters are zero, but the same argument works in general. In this case, we prefer to change our notation and write the corresponding form as

$$\begin{aligned} \Lambda_{\vec{\mathbf{P}}}(f_1, f_2, f_3) &= \int_{\mathbf{R}^2} \Pi_{\vec{\mathbf{P}}}(f_1, f_2)(x) f_3(x)\, dx \\ &= \sum_{\vec{P} \in \vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} \langle f_1, \Phi_{\vec{P}_1} \rangle \langle f_2, \Phi_{\vec{P}_2} \rangle \langle f_3, \Phi_{\vec{P}_3} \rangle, \end{aligned} \tag{17}$$

where the $\vec{P}$'s are bi-parameter tiles corresponding to the indices k', l', k'', l''. More precisely, we have

$$\begin{aligned} \vec{P}_1 &= (P_1', P_1'') = \left(2^{-k'}[l', l'+1] \times 2^{k'}\left[-\tfrac{1}{4}, \tfrac{1}{4}\right], 2^{-k''}[l'', l''+1] \times 2^{k''}\left[\tfrac{3}{4}, \tfrac{5}{4}\right]\right), \\ \vec{P}_2 &= (P_2', P_2'') = \left(2^{-k'}[l', l'+1] \times 2^{k'}\left[\tfrac{3}{4}, \tfrac{5}{4}\right], 2^{-k''}[l'', l''+1] \times 2^{k''}\left[-\tfrac{1}{4}, \tfrac{1}{4}\right]\right), \\ \vec{P}_3 &= (P_3', P_3'') = \left(2^{-k'}[l', l'+1] \times 2^{k'}\left[-\tfrac{7}{4}, -\tfrac{1}{4}\right], 2^{-k''}[l'', l''+1] \times 2^{k''}\left[-\tfrac{7}{4}, -\tfrac{1}{4}\right]\right) \end{aligned}$$

and $|I_{\vec{P}}| := |I_{\vec{P}_1}| = |I_{\vec{P}_2}| = |I_{\vec{P}_3}| = 2^{-k'} 2^{-k''}$.

$\vec{\mathbf{P}}$ will be a finite set of such bi-parameter tiles. Note that $\vec{P}_1$, $\vec{P}_2$ and $\vec{P}_3$ are the bi-parameter Heisenberg boxes of the L^2-normalized wave packets $\Phi_{\vec{P}_1}$, $\Phi_{\vec{P}_2}$ and $\Phi_{\vec{P}_3}$, respectively. These new functions $\Phi_{\vec{P}_j}$ are just the old functions $\Phi_{j,\vec{k},\vec{l}}$ previously defined, for $j=1,2,3$. We therefore need to show the inequality

$$\sum_{\vec{P} \in \vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f_1, \Phi_{\vec{P}_1} \rangle| \, |\langle f_2, \Phi_{\vec{P}_2} \rangle| \, |\langle f_3, \Phi_{\vec{P}_3} \rangle| \lesssim 1 \tag{18}$$

in order to finish the proof. This will be our main goal in the next sections.

At the end of this section we would like to observe that it is very easy to obtain the desired estimates when all the indices are strictly between 1 and ∞. To see this, let

$f_1 \in L^p$, $f_2 \in L^q$ and $f_3 \in L^r$, where $1<p,q,r<\infty$ with $1/p+1/q+1/r=1$. Then,

$$\begin{aligned}\left|\int_{\mathbf{R}^2} \Pi_{\vec{\mathbf{P}}}(f_1,f_2)(x) f_3(x)\,dx\right| &\lesssim \sum_{\vec{P}\in\vec{\mathbf{P}}} \frac{1}{|I_{\vec{P}}|^{1/2}} |\langle f_1, \Phi_{\vec{P}_1}\rangle|\,|\langle f_2, \Phi_{\vec{P}_2}\rangle|\,|\langle f_3, \Phi_{\vec{P}_3}\rangle| \\ &= \int_{\mathbf{R}^2} \sum_{\vec{P}\in\vec{\mathbf{P}}} \frac{|\langle f_1, \Phi_{\vec{P}_1}\rangle|}{|I_{\vec{P}}|^{1/2}} \frac{|\langle f_2, \Phi_{\vec{P}_2}\rangle|}{|I_{\vec{P}}|^{1/2}} \frac{|\langle f_3, \Phi_{\vec{P}_3}\rangle|}{|I_{\vec{P}}|^{1/2}} \chi_{I_{\vec{P}}}(x)\,dx \\ &\lesssim \int_{\mathbf{R}^2} \mathrm{MS}(f_1)(x)\,\mathrm{SM}(f_2)(x)\,\mathrm{SS}(f_3)(x)\,dx \\ &\lesssim \|\mathrm{MS}(f_1)\|_p \|\mathrm{SM}(f_2)\|_q \|\mathrm{SS}(f_3)\|_r \\ &\lesssim \|f_1\|_p \|f_2\|_q \|f_3\|_r.\end{aligned}$$

3. Proof of the one-parameter case

In the particular case when $\vec{\mathbf{P}}=\mathbf{P}'\times\mathbf{P}''$ and all the functions f_j are functions of tensor product type (i.e. $f_j=f_j'\otimes f_j''$, $j=1,2,3$), our bi-parameter paraproduct splits as

$$\Lambda_{\vec{\mathbf{P}}}(f_1,f_2,f_3) = \Lambda_{\mathbf{P}'}(f_1',f_2',f_3')\,\Lambda_{\mathbf{P}''}(f_1'',f_2'',f_3'').$$

In this section, we describe an argument which proves L^p-estimates for these one-parameter paraproducts $\Lambda_{\mathbf{P}'}$ and $\Lambda_{\mathbf{P}''}$. On one hand, this method will be very useful for us in §6, and on the other hand, it provides a new proof of the classical Coifman–Meyer theorem. A sketch of it in a simplified "Walsh framework" has been presented in the expository paper [1].

If I is an interval on the real line, we denote by $\widetilde{\chi}_I(x)$ the function

$$\widetilde{\chi}_I(x) = \left(1+\frac{\operatorname{dist}(x,I)}{|I|}\right)^{-M},$$

where $M>0$ is a big and fixed constant. For simplicity of notation we will suppress the "primes" and write (for instance) $\Lambda_{\mathbf{P}'}(f_1',f_2',f_3')$ simply as

$$\Lambda_{\mathbf{P}}(f_1,f_2,f_3) = \sum_{P\in\mathbf{P}} \frac{1}{|I_P|^{1/2}} \langle f_1, \Phi_{P_1}\rangle \langle f_2, \Phi_{P_2}\rangle \langle f_3, \Phi_{P_3}\rangle. \tag{19}$$

Notice that in this case, as P runs inside the finite set $\mathbf{P}$, the frequency supports $\operatorname{supp}\widehat{\Phi}_{P_j}$, $j=2,3$, lie inside some intervals which are essentially lacunarily disjoint, while the frequency intervals $\operatorname{supp}\widehat{\Phi}_{P_1}$ are all intersecting each other.

In order to deal with the expression (19) we need to introduce some definitions.

Definition 3.1. Let $\mathbf{P}$ be a finite set of tiles as before. For $j=1$ we define

$$\operatorname{size}_{\mathbf{P}}(f_j) := \sup_{P\in\mathbf{P}} \frac{|\langle f_j, \Phi_{P_j}\rangle|}{|I_P|^{1/2}},$$

and for $j=2,3$ we set

$$\operatorname{size}_{\mathbf{P}}(f_j) := \sup_{P\in\mathbf{P}} \frac{1}{|I_P|} \left\| \left(\sum_{I_{P'}\subseteq I_P} \frac{|\langle f_j, \Phi_{P'_j}\rangle|^2}{|I_{P'}|} 1_{I_{P'}} \right)^{1/2} \right\|_{1,\infty}.$$

Also, for $j=1,2,3$ we define

$$\operatorname{energy}_{\mathbf{P}}(f_j) := \sup_{\mathcal{D}\subseteq\mathbf{P}} \left\| \sum_{P\in\mathcal{D}} \frac{\langle |f_j|, \tilde{\chi}_{I_P}\rangle}{|I_P|} 1_{I_P} \right\|_{1,\infty},$$

where $\mathcal{D}$ ranges over all subsets of $\mathbf{P}$ such that the intervals $\{I_P : P\in\mathcal{D}\}$ are disjoint.

The following John–Nirenberg-type inequality holds in this context (see [14]).

LEMMA 3.2. *Let $\mathbf{P}$ be a finite collection of tiles as before and $j=2,3$. Then*

$$\operatorname{size}_{\mathbf{P}}(f_j) \sim \sup_{P\in\mathbf{P}} \left(\frac{1}{|I_P|} \sum_{I_{P'}\subseteq I_P} |\langle f_j, \Phi_{P'_j}\rangle|^2 \right)^{1/2}.$$

We will also need the following lemma (see also [14]).

LEMMA 3.3. *Let $\mathbf{P}$ be a finite collection of tiles and $j=2,3$. Then, we have*

$$\left\| \left(\sum_{I_{P'}\subseteq I_P} \frac{|\langle f, \Phi_{P'_j}\rangle|^2}{|I_{P'}|} 1_{I_{P'}} \right)^{1/2} \right\|_{1,\infty} \lesssim \|f\tilde{\chi}_{I_P}\|_1.$$

The following proposition will be very helpful.

PROPOSITION 3.4. *Let $j=1,2,3$, $\mathbf{P}'$ be a subset of $\mathbf{P}$, $n\in\mathbf{Z}$, and suppose that*

$$\operatorname{size}_{\mathbf{P}'}(f_j) \leqslant 2^{-n} \operatorname{energy}_{\mathbf{P}}(f_j).$$

Then, we may decompose $\mathbf{P}'=\mathbf{P}''\cup\mathbf{P}'''$ so that

$$\operatorname{size}_{\mathbf{P}''}(f_j) \leqslant 2^{-n-1} \operatorname{energy}_{\mathbf{P}}(f_j) \tag{20}$$

and so that $\mathbf{P}'''$ can be written as a disjoint union of subsets $T\in\mathbf{T}$ such that for every $T\in\mathbf{T}$, there exists an interval I_T (corresponding to a certain tile) having the property that every $P\in T$ has $I_P\subseteq I_T$, and also such that

$$\sum_{T\in\mathbf{T}} |I_T| \lesssim 2^n. \tag{21}$$

Proof. The idea is to remove large subsets of $\mathbf{P}'$ one by one, placing them in $\mathbf{P}'''$ until (20) is satisfied.

Case 1: $j=1$. Pick a tile $P\in\mathbf{P}'$ such that $|I_P|$ is as big as possible and such that

$$\frac{|\langle f_j,\Phi_{P_j}\rangle|}{|I_P|^{1/2}}>2^{-n-1}\operatorname{energy}_{\mathbf{P}}(f_j).$$

Then, collect all the tiles $P'\in\mathbf{P}'$ such that $I_{P'}\subseteq I_P$ in a set called T, and place T in $\mathbf{P}'''$. Define $I_T:=I_P$. Then look at the remaining tiles in $\mathbf{P}'\setminus T$ and repeat the procedure. Since there are finitely many tiles, the procedure ends after finitely many steps producing the subsets $T\in\mathbf{T}$. Clearly, (20) is now satisfied, and it remains to show (21). To see this, one can write

$$\sum_{T\in\mathbf{T}}|I_T|=\biggl\|\sum_{T\in\mathbf{T}}1_{I_T}\biggr\|_1=\biggl\|\sum_{T\in\mathbf{T}}1_{I_T}\biggr\|_{1,\infty},$$

since by construction, our intervals I_T are disjoint. Then, the right-hand side of the above equality is smaller than

$$2^n\operatorname{energy}_{\mathbf{P}}(f_j)^{-1}\biggl\|\sum_{T\in\mathbf{T}}\frac{\langle|f_j|,\widetilde{\chi}_{I_T}\rangle}{|I_T|}1_{I_T}\biggr\|_{1,\infty}\lesssim 2^n.$$

Case 2: $j=2,3$. The algorithm is very similar. Pick again a tile $P\in\mathbf{P}'$ such that $|I_P|$ is as big as possible and such that

$$\frac{1}{|I_P|}\biggl\|\biggl(\sum_{I_{P'}\subseteq I_P}\frac{|\langle f_j,\Phi_{P'_j}\rangle|^2}{|I_{P'}|}1_{I_{P'}}\biggr)^{1/2}\biggr\|_{1,\infty}>2^{-n-1}\operatorname{energy}_{\mathbf{P}}(f_j).$$

Then, as before, collect all the tiles $P'\in\mathbf{P}'$ such that $I_{P'}\subseteq I_P$ in a set named T, and place this T in $\mathbf{P}'''$. Define, as in Case 1, $I_T:=I_P$. Then look at the remaining tiles $\mathbf{P}'\setminus T$ and repeat the procedure, which of course ends after finitely many steps. Inequality (20) is now clear, and it only remains to understand (21).

Since the intervals I_T are disjoint by construction, we can write

$$\begin{aligned}\sum_{T\in\mathbf{T}}|I_T|&=\biggl\|\sum_{T\in\mathbf{T}}1_{I_T}\biggr\|_1=\biggl\|\sum_{T\in\mathbf{T}}1_{I_T}\biggr\|_{1,\infty}\\&\lesssim 2^n\operatorname{energy}_{\mathbf{P}}(f_j)^{-1}\biggl\|\sum_{T\in\mathbf{T}}\frac{1}{|I_T|}\biggl\|\biggl(\sum_{I_{P'}\subseteq I_T}\frac{|\langle f_j,\Phi_{P'_j}\rangle|^2}{|I_{P'}|}1_{I_{P'}}\biggr)^{1/2}\biggr\|_{1,\infty}1_{I_T}\biggr\|_{1,\infty}\\&\lesssim 2^n\operatorname{energy}_{\mathbf{P}}(f_j)^{-1}\biggl\|\sum_{T\in\mathbf{T}}\frac{\langle|f_j|,\widetilde{\chi}_{I_T}\rangle}{|I_T|}1_{I_T}\biggr\|_{1,\infty}\\&\lesssim 2^n,\end{aligned}$$

by using Lemma 3.3, and this ends the proof. □

By iterating the above lemma, we immediately obtain the following consequence:

COROLLARY 3.5. *Let $j=1,2,3$. There exists a partition*

$$\mathbf{P}=\bigcup_{n\in\mathbf{Z}}\mathbf{P}_n$$

such that for every $n\in\mathbf{Z}$ we have

$$\operatorname{size}_{\mathbf{P}_n}(f_j)\leqslant\min\{2^{-n}\operatorname{energy}_{\mathbf{P}}(f_j),\operatorname{size}_{\mathbf{P}}(f_j)\}.$$

Also, we may write each $\mathbf{P}_n$ as a disjoint union of subsets $T\in\mathbf{T}_n$ as before, such that

$$\sum_{T\in\mathbf{T}_n}|I_T|\lesssim 2^n.$$

We now prove the following proposition.

PROPOSITION 3.6. *Let $\mathbf{P}$ be a set as before. Then,*

$$\sum_{P\in\mathbf{P}}\frac{1}{|I_P|^{1/2}}|\langle f_1,\Phi_{P_1}\rangle|\,|\langle f_2,\Phi_{P_2}\rangle|\,|\langle f_3,\Phi_{P_3}\rangle|\lesssim\prod_{j=1}^{3}\operatorname{size}_{\mathbf{P}}(f_j)^{1-\theta_j}\operatorname{energy}_{\mathbf{P}}(f_j)^{\theta_j}\tag{22}$$

for any $0\leqslant\theta_1,\theta_2,\theta_3<1$ such that $\theta_1+\theta_2+\theta_3=1$, with the implicit constant depending on θ_j, $j=1,2,3$.

Proof. During this proof, we will write for simplicity

$$S_j:=\operatorname{size}_{\mathbf{P}}(f_j)\quad\text{and}\quad E_j:=\operatorname{energy}_{\mathbf{P}}(f_j)$$

for $j=1,2,3$. If we apply Corollary 3.5 to the functions f_j/E_j, $j=1,2,3$, we obtain a decomposition

$$\mathbf{P}=\bigcup_{n\in\mathbf{Z}}\mathbf{P}_n^j$$

such that each $\mathbf{P}_n^j$ can be written as a union of subsets in $\mathbf{T}_n^j$ with the properties described in Corollary 3.5. In particular, one can write the left-hand side of our described inequality (22) as

$$E_1E_2E_3\sum_{n_1,n_2,n_3}\sum_{T\in\mathbf{T}^{n_1,n_2,n_3}}\sum_{P\in T}\frac{1}{|I_P|^{1/2}}\left|\left\langle\frac{f_1}{E_1},\Phi_{P_1}\right\rangle\right|\left|\left\langle\frac{f_2}{E_2},\Phi_{P_2}\right\rangle\right|\left|\left\langle\frac{f_3}{E_3},\Phi_{P_3}\right\rangle\right|,\tag{23}$$

where $\mathbf{T}^{n_1,n_2,n_3}:=\mathbf{T}_{n_1}^1\cap\mathbf{T}_{n_2}^2\cap\mathbf{T}_{n_3}^3$. By using Hölder's inequality on every $T\in\mathbf{T}^{n_1,n_2,n_3}$ together with Lemma 3.2, one can estimate the sum in (23) by

$$E_1E_2E_3\sum_{n_1,n_2,n_3}2^{-n_1}2^{-n_2}2^{-n_3}\sum_{T\in\mathbf{T}^{n_1,n_2,n_3}}|I_T|,\tag{24}$$

where (according to the same Corollary 3.5) the summation goes over those $n_1, n_2, n_3 \in \mathbf{Z}$ satisfying

$$2^{-n_j} \lesssim \frac{S_j}{E_j}. \tag{25}$$

On the other hand, Corollary 3.5 allows us to estimate the inner sum in (24) in three different ways, namely

$$\sum_{T \in \mathbf{T}^{n_1,n_2,n_3}} |I_T| \lesssim 2^{n_j}, \quad j = 1,2,3,$$

and so, in particular, we can also write

$$\sum_{T \in \mathbf{T}^{n_1,n_2,n_3}} |I_T| \lesssim 2^{n_1\theta_1} 2^{n_2\theta_2} 2^{n_3\theta_3} \tag{26}$$

whenever $0 \leqslant \theta_1, \theta_2, \theta_3 < 1$ with $\theta_1 + \theta_2 + \theta_3 = 1$. Using (26) and (25), one can estimate (24) further by

$$\begin{aligned} E_1 E_2 E_3 \sum_{n_1,n_2,n_3} 2^{-n_1(1-\theta_1)} 2^{-n_2(1-\theta_2)} 2^{-n_3(1-\theta_3)} &\lesssim E_1 E_2 E_3 \left(\frac{S_1}{E_1}\right)^{1-\theta_1} \left(\frac{S_2}{E_2}\right)^{1-\theta_2} \left(\frac{S_2}{E_2}\right)^{1-\theta_3} \\ &= \prod_{j=1}^{3} S_j^{1-\theta_j} \prod_{j=1}^{3} E_j^{\theta_j}, \end{aligned}$$

which ends the proof. □

Using this Proposition 3.6, one can prove the L^p-boundedness of one-parameter paraproducts, as follows. We just need to show that they map $L^1 \times L^1 \to L^{1/2,\infty}$, because then, by interpolation and symmetry one can deduce that they map $L^p \times L^q \to L^r$ as long as $1 < p, q \leqslant \infty$, $0 < r < \infty$ and $1/p + 1/q = 1/r$.

Let $f_1, f_2 \in L^1$ be such that $\|f_1\|_1 = \|f_2\|_1 = 1$. As before, it is enough to show that given $E_3 \subseteq \mathbf{R}$, $|E_3| = 1$, one can find a subset $E_3' \subseteq E_3$ with $|E_3'| \sim 1$ and

$$\sum_{P \in \mathbf{P}} \frac{1}{|I_P|^{1/2}} |\langle f_1, \Phi_{P_1} \rangle| \, |\langle f_2, \Phi_{P_2} \rangle| \, |\langle f_3, \Phi_{P_3} \rangle| \lesssim 1, \tag{27}$$

where $f_3 := \chi_{E_3'}$. For we define the set U by

$$U := \{x \in \mathbf{R} : M(f_1)(x) > C\} \cup \{x \in \mathbf{R} : M(f_2)(x) > C\},$$

where $M(f)$ is the Hardy–Littlewood maximal operator of f. Clearly, we have $|U| < \frac{1}{2}$ if $C > 0$ is big enough. We define our set $E_3' := E_3 \cap U^c$ and remark that $|E_3'| \sim 1$.

Then, we write

$$\mathbf{P} = \bigcup_{d \geqslant 0} \mathbf{P}_d,$$

where

$$\mathbf{P}_d := \left\{ P \in \mathbf{P} : \frac{\operatorname{dist}(I_P, U^c)}{|I_P|} \sim 2^d \right\}.$$

After that, by using Lemma 3.3, we observe that $\operatorname{size}_{\mathbf{P}_d}(f_j) \lesssim 2^d$ for $j=1,2$, while $\operatorname{size}_{\mathbf{P}_d}(f_3) \lesssim 2^{-Nd}$ for an arbitrarily big number $N>0$. We also observe that

$$\operatorname{energy}_{\mathbf{P}_d}(f_j) \lesssim \|M(f_j)\|_{1,\infty} \lesssim \|f_j\|_1 = 1.$$

By applying Proposition 3.6 in the particular case $\theta_1=\theta_2=\theta_3=\frac{1}{3}$, we get that the left-hand side of (27) can be majorized by

$$\sum_{d\geqslant 0} \sum_{P\in\mathbf{P}_d} \frac{1}{|I_P|^{1/2}} |\langle f_1, \Phi_{P_1}\rangle| \, |\langle f_2, \Phi_{P_2}\rangle| \, |\langle f_3, \Phi_{P_3}\rangle| \lesssim \sum_{d\geqslant 0} 2^{2d/3} 2^{2d/3} 2^{-2Nd/3} \lesssim 1$$

as wanted, and this finishes the proof of the one-parameter case.

The reader should compare this Proposition 3.6 with the corresponding Proposition 6.5 in [16]. Our present "lacunary setting" allows for an L^1-type definition of the "energies" (instead of L^2-type as in [16]), and this is why we can obtain the full range of estimates this time.

4. Proof of Theorem 1.1

We reduced our proof to showing (18). Clearly, this inequality is the bi-parameter analogue of the inequality (27) above. Unfortunately, the technique just described in §3, so useful when estimating (27), cannot handle our sum in (18) this time. In fact, we do not know if there exists a satisfactory bi-parameter analogue of Proposition 3.6, and this is where some of the main new difficulties are coming from. Hence, we have to proceed differently.

We split the left-hand side of that inequality into two parts,

$$\sum_{\vec{P}} = \sum_{I_{\vec{P}}\cap\Omega^c\neq\varnothing} + \sum_{I_{\vec{P}}\cap\Omega^c=\varnothing} =: \mathrm{I}+\mathrm{II}, \tag{28}$$

where Ω is the set defined in (16).

5. Estimates for the term I

We first estimate the term I. The argument goes as follows.

Since $I_{\vec{P}}\cap\Omega^c\neq\varnothing$, it follows that $|I_{\vec{P}}\cap\Omega_0|/|I_{\vec{P}}|<\frac{1}{100}$, or equivalently,

$$|I_{\vec{P}}\cap\Omega_0^c| > \tfrac{99}{100}|I_{\vec{P}}|.$$

We are now going to describe three decomposition procedures, one for each function f_1, f_2 and f_3. Later on, we will combine them, in order to handle our sum.

First, define

$$\Omega_1 = \left\{ x \in \mathbf{R}^2 : \mathrm{MS}(f_1)(x) > \frac{C}{2^1} \right\}$$

and set

$$\mathbf{T}_1 = \{ \vec{P} \in \vec{\mathbf{P}} : |I_{\vec{P}} \cap \Omega_1| > \tfrac{1}{100} |I_{\vec{P}}| \},$$

then define

$$\Omega_2 = \left\{ x \in \mathbf{R}^2 : \mathrm{MS}(f_1)(x) > \frac{C}{2^2} \right\}$$

and set

$$\mathbf{T}_2 = \{ \vec{P} \in \vec{\mathbf{P}} \setminus \mathbf{T}_1 : |I_{\vec{P}} \cap \Omega_2| > \tfrac{1}{100} |I_{\vec{P}}| \},$$

and so on. The constant $C>0$ is the one in the definition of the set E_3' in §2. Since there are finitely many tiles, this algorithm ends after a while, producing the sets $\{\Omega_n\}$ and $\{\mathbf{T}_n\}$ such that $\vec{\mathbf{P}} = \bigcup_n \mathbf{T}_n$.

Independently, define

$$\Omega_1' = \left\{ x \in \mathbf{R}^2 : \mathrm{SM}(f_2)(x) > \frac{C}{2^1} \right\}$$

and set

$$\mathbf{T}_1' = \{ \vec{P} \in \vec{\mathbf{P}} : |I_{\vec{P}} \cap \Omega_1'| > \tfrac{1}{100} |I_{\vec{P}}| \},$$

then define

$$\Omega_2' = \left\{ x \in \mathbf{R}^2 : \mathrm{SM}(f_2)(x) > \frac{C}{2^2} \right\}$$

and set

$$\mathbf{T}_2' = \{ \vec{P} \in \vec{\mathbf{P}} \setminus \mathbf{T}_1' : |I_{\vec{P}} \cap \Omega_2'| > \tfrac{1}{100} |I_{\vec{P}}| \},$$

and so on, producing the sets $\{\Omega_n'\}$ and $\{\mathbf{T}_n'\}$ such that $\vec{\mathbf{P}} = \bigcup_n \mathbf{T}_n'$. We would like to have such a decomposition available for the function f_3 also. To do this, we first need to construct the analogue of the set Ω_0 for it. Pick $N>0$, a big enough integer such that for every $\vec{P} \in \vec{\mathbf{P}}$ we have $|I_{\vec{P}} \cap \Omega''^{c}_{-N}| > \frac{99}{100} |I_{\vec{P}}|$, where we defined

$$\Omega''_{-N} = \{ x \in \mathbf{R}^2 : \mathrm{SS}(f_3)(x) > C 2^N \}.$$

Then, similarly to the previous algorithms, we define

$$\Omega''_{-N+1} = \left\{ x \in \mathbf{R}^2 : \mathrm{SS}(f_2)(x) > \frac{C 2^N}{2^1} \right\}$$

and set

$$\mathbf{T}''_{-N+1} = \{\vec{P} \in \vec{\mathbf{P}} : |I_{\vec{P}} \cap \Omega''_{-N+1}| > \tfrac{1}{100}|I_{\vec{P}}|\},$$

then define

$$\Omega''_{-N+2} = \left\{ x \in \mathbf{R}^2 : \mathrm{SS}(f_3)(x) > \frac{C2^N}{2^2} \right\}$$

and set

$$\mathbf{T}''_{-N+2} = \{\vec{P} \in \vec{\mathbf{P}} \setminus \mathbf{T}''_{-N+1} : |I_{\vec{P}} \cap \Omega''_{-N+2}| > \tfrac{1}{100}|I_{\vec{P}}|\},$$

and so on, constructing the sets $\{\Omega''_n\}$ and $\{\mathbf{T}''_n\}$ such that $\vec{\mathbf{P}} = \bigcup_n \mathbf{T}''_n$.

Then we write the term I as

$$\sum_{\substack{n_1,n_2>0\\ n_3>-N}} \sum_{\vec{P}\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_{\vec{P}}|^{3/2}} |\langle f_1, \Phi_{\vec{P}_1}\rangle| \, |\langle f_2, \Phi_{\vec{P}_2}\rangle| \, |\langle f_3, \Phi_{\vec{P}_3}\rangle| \, |I_{\vec{P}}|, \tag{29}$$

where $\mathbf{T}_{n_1,n_2,n_3} := \mathbf{T}_{n_1} \cap \mathbf{T}'_{n_2} \cap \mathbf{T}''_{n_3}$. Now, if $\vec{P}$ belongs to $\mathbf{T}_{n_1,n_2,n_3}$, this means in particular that $\vec{P}$ has not been selected at the previous n_1-1, n_2-1 and n_3-1 steps, respectively, which means that $|I_{\vec{P}} \cap \Omega_{n_1-1}| < \frac{1}{100}|I_{\vec{P}}|$, $|I_{\vec{P}} \cap \Omega'_{n_2-1}| < \frac{1}{100}|I_{\vec{P}}|$ and $|I_{\vec{P}} \cap \Omega''_{n_3-1}| < \frac{1}{100}|I_{\vec{P}}|$, or equivalently, $|I_{\vec{P}} \cap \Omega^c_{n_1-1}| > \frac{99}{100}|I_{\vec{P}}|$, $|I_{\vec{P}} \cap \Omega'^c_{n_2-1}| > \frac{99}{100}|I_{\vec{P}}|$ and $|I_{\vec{P}} \cap \Omega''^c_{n_3-1}| > \frac{99}{100}|I_{\vec{P}}|$. But this implies that

$$|I_{\vec{P}} \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}| > \tfrac{97}{100}|I_{\vec{P}}|. \tag{30}$$

In particular, using (30), the term in (29) is smaller than

$$\begin{aligned}
&\sum_{\substack{n_1,n_2>0\\ n_3>-N}} \sum_{\vec{P}\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_{\vec{P}}|^{3/2}} |\langle f_1, \Phi_{\vec{P}_1}\rangle| \, |\langle f_2, \Phi_{\vec{P}_2}\rangle| \, |\langle f_3, \Phi_{\vec{P}_3}\rangle| \, |I_{\vec{P}} \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}| \\
&= \sum_{\substack{n_1,n_2>0\\ n_3>-N}} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}} \sum_{\vec{P}\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_{\vec{P}}|^{3/2}} |\langle f_1, \Phi_{\vec{P}_1}\rangle| \, |\langle f_2, \Phi_{\vec{P}_2}\rangle| \, |\langle f_3, \Phi_{\vec{P}_3}\rangle| \chi_{I_{\vec{P}}}(x)\, dx \\
&\lesssim \sum_{\substack{n_1,n_2>0\\ n_3>-N}} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}\cap\Omega_{\mathbf{T}_{n_1,n_2,n_3}}} \mathrm{MS}(f_1)(x)\, \mathrm{SM}(f_2)(x)\, \mathrm{SS}(f_3)(x)\, dx \qquad (31) \\
&\lesssim \sum_{\substack{n_1,n_2>0\\ n_3>-N}} 2^{-n_1} 2^{-n_2} 2^{-n_3} |\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|,
\end{aligned}$$

where

$$\Omega_{\mathbf{T}_{n_1,n_2,n_3}} := \bigcup_{\vec{P}\in\mathbf{T}_{n_1,n_2,n_3}} I_{\vec{P}}.$$

On the other hand, we can write

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \leqslant |\Omega_{\mathbf{T}_{n_1}}| \leqslant \left|\left\{x\in\mathbf{R}^2 : \mathrm{MM}(\chi_{\Omega_{n_1}})(x) > \tfrac{1}{100}\right\}\right|$$
$$\lesssim |\Omega_{n_1}| = \left|\left\{x\in\mathbf{R}^2 : \mathrm{MS}(f_1)(x) > \frac{C}{2^{n_1}}\right\}\right| \lesssim 2^{n_1 p}.$$

Similarly, we have

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_2 q},$$

and also

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_2 \alpha}$$

for every $\alpha>1$. Here we used the fact that all the operators SM, MS, SS and MM are bounded on L^s as long as $1<s<\infty$, and also that $|E_3'|\sim 1$. In particular, it follows that

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_1 p\theta_1} 2^{n_2 q\theta_2} 2^{n_3 \alpha\theta_3} \tag{32}$$

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

Now we split the sum in (31) into

$$\sum_{\substack{n_1,n_2>0\\ n_3>0}} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| + \sum_{\substack{n_1,n_2>0\\ 0>n_3>-N}} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|. \tag{33}$$

To estimate the first term in (33) we use the inequality (32) in the particular case $\theta_1=\theta_2=\frac{1}{2}$ and $\theta_3=0$, while to estimate the second term we use (32) for θ_j, $j=1,2,3$, such that $1-p\theta_1>0$, $1-q\theta_2>0$ and $\alpha\theta_3-1>0$. With these choices, the sum in (33) is $O(1)$. This ends the discussion of the term I.

6. Estimates for the term II

It remains to estimate the term II in (28). The sum now runs over those tiles having the property that $I_{\vec{P}}\subseteq\Omega$. For every such $\vec{P}$ there exists a maximal dyadic rectangle R such that $I_{\vec{P}}\subseteq R\subseteq\Omega$. We collect all such distinct maximal rectangles into a set called $R_{\max}$. For an integer $d\geqslant 1$, we denote by $R_{\max}^d$ the set of all $R\in R_{\max}$ such that $2^d R\subseteq\widetilde{\Omega}$ and d is maximal with this property.

By using Journé's lemma [8](1) in the form presented in [6], we have that for every $\varepsilon>0$,

$$\sum_{R\in R_{\max}^d} |R| \lesssim 2^{\varepsilon d}|\Omega|. \tag{34}$$

(1) The use of Journé's lemma in estimating this error term can be replaced by a simpler argument, which works in the multiparameter setting as well. Therefore, the proof presented in this paper readily extends to three or more parameters. The details will appear elsewhere.

Our initial sum in the term II is now smaller than

$$\sum_{d\geqslant 1}\sum_{R\in R^d_{\max}}\sum_{I_{\vec{P}}\subseteq R\cap\Omega}\frac{1}{|I_{\vec{P}}|^{1/2}}|\langle f_1,\Phi_{\vec{P}_1}\rangle|\,|\langle f_2,\Phi_{\vec{P}_2}\rangle|\,|\langle f_3,\Phi_{\vec{P}_3}\rangle|. \tag{35}$$

We claim that for every $R\in R^d_{\max}$ we have

$$\sum_{I_{\vec{P}}\subseteq R\cap\Omega}\frac{1}{|I_{\vec{P}}|^{1/2}}|\langle f_1,\Phi_{\vec{P}_1}\rangle|\,|\langle f_2,\Phi_{\vec{P}_2}\rangle|\,|\langle f_3,\Phi_{\vec{P}_3}\rangle|\lesssim 2^{-Nd}|R| \tag{36}$$

for any number $N>0$. If (36) is true, then by combining it with (34), we can estimate (35) by

$$\sum_{d\geqslant 1}\sum_{R\in R^d_{\max}}2^{-Nd}|R|=\sum_{d\geqslant 1}2^{-Nd}\sum_{R\in R^d_{\max}}|R|\lesssim\sum_{d\geqslant 1}2^{-Nd}2^{\varepsilon d}\lesssim 1,$$

which would complete the proof.

It remains to prove (36). Fix $R:=I\times J$ in $R^d_{\max}$. Since $2^dR:=\tilde{I}\times\tilde{J}\subseteq\widetilde{\Omega}$, it follows that $2^dR\cap E_3'=\varnothing$, and so $\chi_{E_3'}=\chi_{E_3'}\chi_{(\tilde{I}\times\tilde{J})^c}$. Now we write

$$\chi_{(\tilde{I}\times\tilde{J})^c}=\chi_{\tilde{I}^c}+\chi_{\tilde{J}^c}-\chi_{\tilde{I}^c}\cdot\chi_{\tilde{J}^c}.$$

As a consequence, the left-hand side in (36) splits into three sums. Since all are similar, we will treat only the first one.

Recall that every $I_{\vec{P}}$ is of the form $I_{\vec{P}}=I_{P'}\times I_{P''}$, and let us denote by $\mathcal{L}$ the set

$$\mathcal{L}:=\{I_{P'}:I_{\vec{P}}\subseteq R\}.$$

Then split

$$\mathcal{L}=\bigcup_{d_1\geqslant 0}\mathcal{L}_{d_1},$$

where

$$\mathcal{L}_{d_1}:=\left\{K'\in\mathcal{L}:\frac{|I|}{|K'|}\sim 2^{d_1}\right\},$$

and observe that

$$\sum_{K'\in\mathcal{L}_{d_1}}|K'|\lesssim|I|. \tag{37}$$

Then, we can majorize the left-hand side of (36) by

$$\begin{aligned}&\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}\sum_{\substack{I_{\vec{P}}\subseteq R\\ I_{P'}=K'}}\frac{1}{|I_{\vec{P}}|^{1/2}}|\langle f_1,\Phi_{\vec{P}_1}\rangle|\,|\langle f_2,\Phi_{\vec{P}_2}\rangle|\,|\langle f_3,\Phi_{\vec{P}_3}\rangle|\\ &\qquad=\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}\sum_{\substack{I_{\vec{P}}\subseteq R\\ I_{P'}=K'}}|I_{P'}|\frac{1}{|I_{P''}|^{1/2}}\\ &\qquad\qquad\times\left|\left\langle\frac{\langle f_1,\Phi_{P_1'}\rangle}{|I_{P'}|^{1/2}},\Phi_{P_1''}\right\rangle\right|\left|\left\langle\frac{\langle f_2,\Phi_{P_2'}\rangle}{|I_{P'}|^{1/2}},\Phi_{P_2''}\right\rangle\right|\left|\left\langle\frac{\langle f_3,\Phi_{P_3'}\rangle}{|I_{P'}|^{1/2}},\Phi_{P_3''}\right\rangle\right|,\end{aligned}$$

where we redefined $f_3:=\chi_{E_3'}\chi_{\tilde{J}^c}$.

Let us observe that if $\vec{P}$ is such that $I_{P'}=K'$, then the one-parameter tiles P_j', $j=1,2,3$, are fixed, and we will denote for simplicity $\Phi_{P_j'}=:\Phi_{K'}^j$. We also write

$$\mathbf{P}_{K'}:=\{P'': I_{\vec{P}}\subseteq R \text{ and } I_{P'}=K'\}.$$

With this notation, we rewrite our sum as

$$\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{P''\in\mathbf{P}_{K'}}\frac{1}{|I_{P''}|^{1/2}}\prod_{j=1}^{3}\left|\left\langle\frac{\langle f_j,\Phi_{K'}^j\rangle}{|K'|^{1/2}},\Phi_{P_j''}\right\rangle\right|. \tag{38}$$

Next we split $\mathbf{P}_{K'}$ as

$$\mathbf{P}_{K'}=\bigcup_{d_2\geqslant 0}\mathbf{P}_{K'}^{d_2},$$

where

$$\mathbf{P}_{K'}^{d_2}:=\left\{P''\in\mathbf{P}_{K'}:\frac{|J|}{|I_{P''}|}\sim 2^{d_2}\right\}.$$

As a consequence, (38) splits as

$$\begin{aligned}
&\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{d_2\geqslant 0}\sum_{P''\in\mathbf{P}_{K'}^{d_2}}\frac{1}{|I_{P''}|^{1/2}}\prod_{j=1}^{3}\left|\left\langle\frac{\langle f_j,\Phi_{K'}^j\rangle}{|K'|^{1/2}},\Phi_{P_j''}\right\rangle\right|\\
&\quad=\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{P''\in\bigcup_{d_2<d_1}\mathbf{P}_{K'}^{d_2}}\frac{1}{|I_{P''}|^{1/2}}\prod_{j=1}^{3}\left|\left\langle\frac{\langle f_j,\Phi_{K'}^j\rangle}{|K'|^{1/2}},\Phi_{P_j''}\right\rangle\right| \qquad (39)\\
&\quad\quad+\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{P''\in\bigcup_{d_2\geqslant d_1}\mathbf{P}_{K'}^{d_2}}\frac{1}{|I_{P''}|^{1/2}}\prod_{j=1}^{3}\left|\left\langle\frac{\langle f_j,\Phi_{K'}^j\rangle}{|K'|^{1/2}},\Phi_{P_j''}\right\rangle\right|.
\end{aligned}$$

To estimate the first term on the right-hand side of (39) we observe that

$$\operatorname{size}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_1,\Phi_{K'}^1\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_1+d},$$

$$\operatorname{size}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_2,\Phi_{K'}^2\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_1+d},$$

$$\operatorname{size}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_3,\Phi_{K'}^3\rangle}{|K'|^{1/2}}\right)\lesssim 2^{-N(d_1+d)},$$

where N is as big as we want. Similarly, we have

$$\operatorname{energy}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_1,\Phi_{K'}^1\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_1+d}|J|,$$

$$\operatorname{energy}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_2,\Phi_{K'}^2\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_1+d}|J|,$$

$$\operatorname{energy}_{\bigcup_{d_2\leqslant d_1}\mathbf{P}_{K'}^{d_2}}\left(\frac{\langle f_3,\Phi_{K'}^3\rangle}{|K'|^{1/2}}\right)\lesssim 2^{-N(d_1+d)}|J|.$$

Using these inequalities and applying Proposition 3.6, we can majorize that first term by

$$\begin{aligned}\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|2^{d_1+d}2^{d_1+d}2^{-N(d_1+d)}|J| &= 2^{-(N-2)d}|J|\sum_{d_1\geqslant 0}2^{-(N-2)d_1}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\\ &\lesssim 2^{-(N-2)d}|J|\sum_{d_1\geqslant 0}2^{-(N-2)d_1}|I|\\ &\lesssim 2^{-(N-2)d}|I|\,|J|\\ &= 2^{-(N-2)d}|R|,\end{aligned}\tag{40}$$

also by using (37). Then, to handle the second term on the right-hand side of (39), we decompose

$$\bigcup_{d_2\geqslant d_1}\mathbf{P}_{K'}^{d_2} = \bigcup_{d_3}\mathbf{P}_{K',d_3},\tag{41}$$

where $\mathbf{P}_{K',d_3}$ is the collection of all tiles $P''\in\bigcup_{d_2\geqslant d_1}\mathbf{P}_{K'}^{d_2}$ so that $2^{d_3}(K'\times I_{P''})\subseteq\widetilde{\Omega}$ and d_3 is maximal with this property.

It is not difficult to observe that in fact we have the constraint $d_1+d\leqslant d_3$. Taking this into account, the second term can be written as

$$\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{d_3\geqslant d_1+d}\sum_{P''\in\mathbf{P}_{K',d_3}}\frac{1}{|I_{P''}|^{1/2}}\prod_{j=1}^{3}\left|\left\langle\frac{\langle f_j,\Phi_{K'}^j\rangle}{|K'|^{1/2}},\Phi_{P_j''}\right\rangle\right|.\tag{42}$$

Now we estimate as before the sizes and energies as follows:

$$\operatorname{size}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_1,\Phi_{K'}^1\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_3},$$

$$\operatorname{size}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_2,\Phi_{K'}^2\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_3},$$

$$\operatorname{size}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_3,\Phi_{K'}^3\rangle}{|K'|^{1/2}}\right)\lesssim 2^{-Nd_3},$$

where, as usual, N is as big as we want. Similarly, we have

$$\operatorname{energy}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_1,\Phi_{K'}^1\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_3}|J|,$$

$$\operatorname{energy}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_2,\Phi_{K'}^2\rangle}{|K'|^{1/2}}\right)\lesssim 2^{d_3}|J|,$$

$$\operatorname{energy}_{\mathbf{P}_{K',d_3}}\left(\frac{\langle f_3,\Phi_{K'}^3\rangle}{|K'|^{1/2}}\right)\lesssim 2^{-Nd_3}|J|.$$

Using all these estimates, the term (42) is seen to be smaller than

$$\sum_{d_1\geqslant 0}\sum_{K'\in\mathcal{L}_{d_1}}|K'|\sum_{d_1+d\leqslant d_3}2^{d_3}2^{d_3}2^{-Nd_3}|J| = |J|\sum_{d_1\geqslant 0}2^{-(N-2)(d_1+d)}\sum_{K'\in\mathcal{L}_{d_1}}|K'| \\ \lesssim |I|\,|J|2^{-(N-2)d} = 2^{(N-2)d}|R|, \tag{43}$$

by using (37), and this completes the proof.

7. Counterexamples

The next step in understanding this bi-parameter multilinear framework is to consider more singular multipliers. The most natural candidate is the double bilinear Hilbert transform, defined by

$$\begin{aligned} B_d(f,g)(x,y) &= \int_{\mathbf{R}^2} f(x-t_1,y-t_2)\,g(x+t_1,y+t_2)\,\frac{dt_1}{t_1}\,\frac{dt_2}{t_2} \\ &= \int_{\mathbf{R}^4}\operatorname{sgn}(\xi_1-\xi_2)\operatorname{sgn}(\eta_1-\eta_2)\hat{f}(\xi_1,\eta_1)\,\hat{g}(\xi_2,\eta_2) \\ &\qquad\times e^{2\pi i(x,y)\cdot((\xi_1,\eta_1)+(\xi_2,\eta_2))}\,d\xi\,d\eta. \end{aligned} \tag{44}$$

It is the bi-parameter analogue of the bilinear Hilbert transform studied in [12] and given by

$$B(f_1,f_2)(x) = \int_{\mathbf{R}} f_1(x-t)\,f_2(x+t)\,\frac{dt}{t} = \int_{\mathbf{R}^2}\operatorname{sgn}(\xi-\eta)\,\hat{f}(\xi)\,\hat{g}(\eta)\,e^{2\pi i x(\xi+\eta)}\,d\xi\,d\eta. \tag{45}$$

This time, the functions f_1 and f_2 are defined on the real line. It is known (see [12]) that B satisfies many L^p-estimates.

However, regarding B_d we have the following theorem:

THEOREM 7.1. *The double bilinear Hilbert transform B_d defined by* (44) *does not satisfy any L^p-estimates.*

Proof. It is based on the following simple observation. Let $f(x,y)=g(x,y)=e^{ixy}$. Since

$$(x-t_1)(y-t_2)+(x+t_1)(y+t_2) = 2xy+2t_1t_2,$$

one can formally write

$$\begin{aligned} B(e^{ixy},e^{ixy})(x,y) &= e^{2ixy}\int_{\mathbf{R}^2} e^{2it_1t_2}\,\frac{dt_1}{t_1}\,\frac{dt_2}{t_2} = 4e^{2ixy}\int_0^\infty\!\!\int_0^\infty \frac{\sin(t_1t_2)}{t_1t_2}\,dt_1\,dt_2 \\ &= 4e^{2ixy}\int_0^\infty\left(\int_0^\infty \frac{\sin(t_1t_2)}{t_2}\,dt_2\right)\frac{dt_1}{t_1} = 4e^{2ixy}\,\frac{\pi}{2}\int_0^\infty \frac{dt}{t}. \end{aligned}$$

To obtain a quantitative version of this, we need the following lemma:

LEMMA 7.2. *There are two universal constants $C_1, C_2>0$ such that*

$$\left|\int_0^N\int_0^N \frac{\sin(xy)}{xy}\,dx\,dy\right| \geqslant C_1 \log N \tag{46}$$

as long as $N>C_2$.

Proof. Since $\int_0^\infty (\sin t/t)\,dt=\frac{1}{2}\pi$, there is a constant $C>0$ such that

$$\int_0^x \frac{\sin t}{t}\,dt \in \left[\frac{\pi}{4}, \frac{3\pi}{4}\right] \tag{47}$$

whenever $x>C$. Then,

$$\begin{aligned}\int_0^N\int_0^N \frac{\sin(xy)}{xy}\,dx\,dy &= \int_0^N\left(\int_0^N \frac{\sin(xy)}{y}\,dy\right)\frac{dx}{x}\\ &= \int_0^N\left(\int_0^{Nx} \frac{\sin t}{t}\,dt\right)\frac{dx}{x}\\ &= \int_0^{C/N}\left(\int_0^{Nx} \frac{\sin t}{t}\,dt\right)\frac{dx}{x}+\int_{C/N}^N\left(\int_0^{Nx} \frac{\sin t}{t}\,dt\right)\frac{dx}{x}\\ &= \int_0^{C}\left(\int_0^{x} \frac{\sin t}{t}\,dt\right)\frac{dx}{x}+\int_{C/N}^N\left(\int_0^{Nx} \frac{\sin t}{t}\,dt\right)\frac{dx}{x}.\end{aligned} \tag{48}$$

Since the function $x\mapsto(1/x)\int_0^x(\sin t/t)\,dt$ is continuous on $[0,C]$, it follows that the first term in (48) is actually $O(1)$. To estimate the second term in (48), we observe that since $x>C/N$, it follows that $Nx>C$, and so, by using (47) we can write

$$\int_{C/N}^N\left(\int_0^{Nx} \frac{\sin t}{t}\,dt\right)\frac{dx}{x} \geqslant \frac{\pi}{4}\int_{C/N}^N \frac{dx}{x} = \frac{\pi}{4}(2\log N-\log C),$$

and this ends the proof of the lemma, if N is big enough. □

Now, coming back to the proof of the theorem, we define

$$f_N(x,y)=g_N(x,y)=e^{ixy}\chi_{[-N,N]}(x)\chi_{[-N,N]}(y)$$

and observe that

$$|B_d(f_N,g_N)(x,y)| \geqslant C\left|\int_0^{N/10}\int_0^{N/10}\frac{\sin(zt)}{zt}\,dz\,dt\right|+O(1)\geqslant C\log N+O(1)$$

as long as $x,y\in\left[-\frac{1}{1000}N, \frac{1}{1000}N\right]$. This pointwise estimate precludes that we have $\|B_d(f_N,g_N)\|_r\leqslant C\|f_N\|_p\|g_N\|_q$ uniformly in N. □

At the end of this section, we would like to observe that, in the same manner, one can disprove the boundedness of the following operator considered in [15]. Let V be the trilinear operator V defined by

$$V(f,g,h)(x)=\int_{\xi_1<\xi_2<\xi_3}\hat f(\xi_1)\hat g(\xi_2)\hat h(\xi_3)e^{2\pi i x(\xi_1-\xi_2+\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3. \tag{49}$$

The following theorem holds (see [15]):

THEOREM 7.3. *The trilinear operator V constructed above does not map*

$$L^2 \times L^2 \times L^2 \to L^{2/3,\infty}.$$

Proof. First, by a simple change of variables one can reduce the study of V to the study of V_1 defined by

$$V_1(f,g,h)(x) = \int_{\xi_1 < -\xi_2 < \xi_3} \hat{f}(\xi_1)\,\hat{g}(\xi_2)\,\hat{h}(\xi_3)\,e^{2\pi i x(\xi_1+\xi_2+\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3. \tag{50}$$

Also, we observe that the behaviour of V_1 is similar to the behaviour of V_2 defined by

$$V_2(f,g,h)(x) = \int_{\mathbf{R}^3} \operatorname{sgn}(\xi_1+\xi_2)\operatorname{sgn}(\xi_2+\xi_3)\,\hat{f}(\xi_1)\,\hat{g}(\xi_2)\,\hat{h}(\xi_3)\,e^{2\pi i x(\xi_1+\xi_2+\xi_3)}\,d\xi_1\,d\xi_2\,d\xi_3, \tag{51}$$

since the difference between V_1 and V_2 is a sum of simpler bounded operators.

But then, V_2 can be rewritten as

$$V_2(f,g,h)(x) = \int_{\mathbf{R}^2} f(x-t_1)\,g(x-t_1-t_2)\,h(x-t_2)\,\frac{dt_1}{t_1}\,\frac{dt_2}{t_2}.$$

The counterexample is based on the following observation, similar to the one before. Consider $f(x)=h(x)=e^{ix^2}$ and $g(x)=e^{-ix^2}$. Because

$$(x-t_1)^2-(x-t_1-t_2)^2+(x-t_2)^2 = x^2+2t_1t_2,$$

we can again formally write

$$V_2(e^{ix^2}, e^{-ix^2}, e^{ix^2})(x) = e^{ix^2}\int_{\mathbf{R}^2} e^{2it_1t_2}\,\frac{dt_1}{t_1}\,\frac{dt_2}{t_2} = 4e^{ix^2}\,\frac{\pi}{4}\int_0^\infty \frac{dt}{t}.$$

To quantify this, we define $f_N(x)=h_N(x)=e^{ix^2}\chi_{[-N,N]}(x)$ and $g_N(x)=e^{-ix^2}\chi_{[-N,N]}(x)$, and observe as before that

$$|V_2(f_N,g_N,h_N)(x)| \geqslant C\left|\int_0^{N/10}\int_0^{N/10} \frac{\sin(xy)}{xy}\,dx\,dy\right| + O(1)$$

if $x\in\left[-\frac{1}{1000}N, \frac{1}{1000}N\right]$, and this, as we have seen, contradicts the boundedness of the operator. □

8. Further remarks

First of all, we would like to remark that Theorem 1.1 has a straightforward generalization to the case of n-linear operators, for $n \geqslant 1$.

Let $m \in L^\infty(\mathbf{R}^{2n})$ be a symbol satisfying the bi-parameter Marcinkiewicz–Hörmander–Mikhlin condition

$$|\partial_\xi^\alpha \partial_\eta^\beta m(\xi,\eta)| \lesssim \frac{1}{|\xi|^{|\alpha|}} \frac{1}{|\eta|^{|\beta|}} \tag{52}$$

for many multiindices α and β. Then, for Schwartz functions $f_1, ..., f_n$ in $\mathbf{R}^2$, define the operator T_m by

$$T_m(f_1, ..., f_n)(x) := \int_{\mathbf{R}^{2n}} m(\xi,\eta) \hat{f}_1(\xi_1,\eta_1) ... \hat{f}_n(\xi_n,\eta_n) e^{2\pi i x \cdot ((\xi_1,\eta_1)+...+(\xi_n,\eta_n))} \, d\xi \, d\eta. \tag{53}$$

We thus record the following result:

THEOREM 8.1. *The bi-parameter n-linear operator T_m maps $L^{p_1} \times ... \times L^{p_n} \to L^p$ as long as $1 < p_1, ..., p_n \leqslant \infty$, $1/p_1 + ... + 1/p_n = 1/p$ and $0 < p < \infty$.*

Here, when such an $(n+1)$-tuple $(p_1, ..., p_n, p)$ has the property that $0<p<1$ and $p_j = \infty$ for some $1 \leqslant j \leqslant n$, then, for some technical reasons (see [14]), by L^∞ one actually means L_c^∞, the space of bounded measurable functions with compact support.

On the other hand, one can ask what is happening if one is interested in more singular multipliers. Suppose that Γ_1 and Γ_2 are subspaces in $\mathbf{R}^n$, and consider operators T_m defined by (53) where m satisfies

$$|\partial_\xi^\alpha \partial_\eta^\beta m(\xi,\eta)| \lesssim \frac{1}{\operatorname{dist}(\xi,\Gamma_1)^{|\alpha|}} \frac{1}{|\operatorname{dist}(\eta,\Gamma_2)^{|\beta|}}. \tag{54}$$

Our theorem says that if $\dim \Gamma_1 = \dim \Gamma_2 = 0$, then we have many L^p-estimates available. On the other hand, the previous counterexamples show that when $\dim \Gamma_1 = \dim \Gamma_2 = 1$, then we do not have any L^p-estimates. But it is of course natural to ask the following question:

QUESTION 8.2. *Let $\dim \Gamma_1 = 0$ and $\dim \Gamma_2 = 1$ with Γ_2 non-degenerate in the sense of* [14]. *If m is a multiplier satisfying* (54), *does the corresponding T_m satisfy any L^p-estimates?*

9. Appendix: differentiating paraproducts

In this section we describe how the Kato–Ponce inequality (2) can be reduced to the Coifman–Meyer theorem, and also how the more general inequality (5) can be reduced to our Theorem 1.1.

The argument is standard and is based on some "calculus with paraproducts". We include it here for the reader's convenience.

In what follows, we will define generic classes of paraproducts. First we consider the sets Φ and Ψ given by

$$\Phi := \{\phi \in \mathcal{S}(\mathbf{R}) : \operatorname{supp} \hat{\phi} \subseteq [-1,1]\},$$
$$\Psi := \{\psi \in \mathcal{S}(\mathbf{R}) : \operatorname{supp} \hat{\psi} \subseteq [1,2]\}.$$

The intervals $[-1,1]$ and $[1,2]$ are not important. What is important is the fact that Φ consists of Schwartz functions whose Fourier support is compact and contains the origin, and Ψ consists of Schwartz functions whose Fourier support is compact and does not contain the origin. Then, for various $\phi \in \Phi$ and $\psi, \psi', \psi'' \in \Psi$, we define the paraproducts Π_j, $j=0,1,2,3$, as

$$\Pi_0(f,g)(x) := \int_{\mathbf{R}} ((f * D^1_{2^k}\psi)(g * D^1_{2^k}\psi')) * D^1_{2^k}\psi''(x)\,dk, \tag{55}$$

$$\Pi_1(f,g)(x) := \int_{\mathbf{R}} ((f * D^1_{2^k}\phi)(g * D^1_{2^k}\psi)) * D^1_{2^k}\psi'(x)\,dk, \tag{56}$$

$$\Pi_2(f,g)(x) := \int_{\mathbf{R}} ((f * D^1_{2^k}\psi)(g * D^1_{2^k}\phi)) * D^1_{2^k}\psi'(x)\,dk, \tag{57}$$

$$\Pi_3(f,g)(x) := \int_{\mathbf{R}} ((f * D^1_{2^k}\psi)(g * D^1_{2^k}\psi')) * D^1_{2^k}\phi(x)\,dk. \tag{58}$$

All these paraproducts are bilinear operators for which the Coifman–Meyer theorem applies. For instance, one can rewrite $\Pi_0(f,g)$ as

$$\Pi_0(f,g)(x) = \int_{\mathbf{R}^2} m(\xi_1,\xi_2)\hat{f}(\xi_1)\hat{g}(\xi_2)e^{2\pi i x(\xi_1+\xi_2)}\,d\xi_1\,d\xi_2,$$

where the symbol $m(\xi_1,\xi_2)$ is given by

$$m(\xi_1,\xi_2) = \int_{\mathbf{R}} (D^\infty_{2^{-k}}\hat{\psi})(\xi_1)(D^\infty_{2^{-k}}\hat{\psi}')(\xi_2)(D^\infty_{2^{-k}}\hat{\psi}'')(-\xi_1-\xi_2)\,dk$$

and satisfies the Marcinkiewicz–Hörmander–Mikhlin condition.

The reduction relies on the following simple observation:

PROPOSITION 9.1. *Let $\alpha>0$. Then, for every paraproduct Π_1 there exists a paraproduct Π_1' so that*

$$\mathcal{D}^\alpha \Pi_1(f,g) = \Pi_1'(f, \mathcal{D}^\alpha g) \tag{59}$$

for all Schwartz functions f and g on $\mathbf{R}$.

Proof. It is based on the equalities

$$\begin{aligned}
\mathcal{D}^{\alpha}\Pi_1(f,g) &= \int_{\mathbf{R}} ((f*D^1_{2^k}\phi)(g*D^1_{2^k}\psi))*\mathcal{D}^{\alpha}(D^1_{2^k}\psi')\,dk \\
&= \int_{\mathbf{R}} ((f*D^1_{2^k}\phi)(g*D^1_{2^k}\psi))*2^{-k\alpha}D^1_{2^k}(\mathcal{D}^{\alpha}\psi')\,dk \\
&= \int_{\mathbf{R}} ((f*D^1_{2^k}\phi)(g*2^{-k\alpha}D^1_{2^k}\psi))*D^1_{2^k}(\mathcal{D}^{\alpha}\psi')\,dk \\
&= \int_{\mathbf{R}} ((f*D^1_{2^k}\phi)(g*\mathcal{D}^{\alpha}(D^1_{2^k}(\mathcal{D}^{-\alpha}\psi))))*D^1_{2^k}(\mathcal{D}^{\alpha}\psi')\,dk \\
&= \int_{\mathbf{R}} ((f*D^1_{2^k}\phi)(\mathcal{D}^{\alpha}g*D^1_{2^k}(\mathcal{D}^{-\alpha}\psi)))*D^1_{2^k}(\mathcal{D}^{\alpha}\psi')\,dk \\
&=: \Pi_1'(f,\mathcal{D}^{\alpha}g),
\end{aligned}$$

where $\mathcal{D}^{-\alpha}\psi$ is the Schwartz function whose Fourier transform is given by $\widehat{\mathcal{D}^{-\alpha}\psi}(\xi)=|\xi|^{-\alpha}\hat{\psi}(\xi)$, which is well defined since $\psi\in\Psi$. $\square$

Clearly, one has similar identities for all the other types of paraproducts Π_j, $j\neq 1$. To prove the Kato–Ponce inequality, one just has to realize that every product of two functions f and g on $\mathbf{R}$ can be written as a sum of such paraproducts,

$$fg=\sum_{j=0}^{3}\Pi_j(f,g),$$

and then, after using the above Proposition 9.1, to apply the Coifman–Meyer theorem.

A similar treatment is available in the bi-parameter case too. Here, one has to handle bi-parameter paraproducts $\Pi_{i,j}$ for $i,j=0,1,2,3$, formally defined by $\Pi_{i,j}:=\Pi_i\otimes\Pi_j$.

One first observes the following extension of Proposition 9.1:

PROPOSITION 9.2. *Let $\alpha,\beta>0$. Then, for every paraproduct $\Pi_{1,2}$ there exists a paraproduct $\Pi'_{1,2}$ so that*

$$\mathcal{D}_1^{\alpha}\mathcal{D}_2^{\beta}\Pi_{1,2}(f,g)=\Pi'_{1,2}(\mathcal{D}_2^{\beta}f,\mathcal{D}_1^{\alpha}g) \tag{60}$$

for all Schwartz functions f and g on $\mathbf{R}^2$.

As before, there are similar equalities for the remaining paraproducts $\Pi_{i,j}$ when $(i,j)\neq(1,2)$. Since every product of two functions f and g on $\mathbf{R}^2$ can be written as

$$fg=\sum_{i,j=0}^{3}\Pi_{i,j}(f,g),$$

everything follows from Theorem 1.1. In fact, the above argument proves that an even more general inequality holds, namely

$$\|\mathcal{D}_1^\alpha \mathcal{D}_2^\beta (fg)\|_r \lesssim \|\mathcal{D}_1^\alpha \mathcal{D}_2^\beta f\|_{p_1} \|g\|_{q_1} + \|f\|_{p_2} \|\mathcal{D}_1^\alpha \mathcal{D}_2^\beta g\|_{q_2} + \|\mathcal{D}_1^\alpha f\|_{p_3} \|\mathcal{D}_2^\beta g\|_{q_3} + \|\mathcal{D}_1^\alpha g\|_{p_4} \|\mathcal{D}_2^\beta f\|_{q_4} \tag{61}$$

whenever $1<p_j, q_j \leqslant \infty$, $1/p_j+1/q_j=1/r$ for $j=1,2,3,4$ and $0<r<\infty$.

References

[1] Auscher, P., Hofmann, S., Muscalu, C., Tao, T. & Thiele, C., Carleson measures, trees, extrapolation, and $T(b)$ theorems. *Publ. Mat.*, 46 (2002), 257–325.

[2] Chang, S.-Y. A. & Fefferman, R., Some recent developments in Fourier analysis and H^p-theory on product domains. *Bull. Amer. Math. Soc.*, 12 (1985), 1–43.

[3] Christ, M. & Journé, J.-L., Polynomial growth estimates for multilinear singular integral operators. *Acta Math.*, 159 (1987), 51-80.

[4] Fefferman, C., On the divergence of multiple Fourier series. *Bull. Amer. Math. Soc.*, 77 (1971), 191–195.

[5] Fefferman, C. & Stein, E. M., Some maximal inequalities. *Amer. J. Math.*, 93 (1971), 107–115.

[6] Ferguson, S. & Lacey, M., A characterization of product BMO by commutators. *Acta Math.*, 189 (2002), 143–160.

[7] Grafakos, L. & Torres, R. H., Multilinear Calderón–Zygmund theory. *Adv. Math.*, 165 (2002), 124–164.

[8] Journé, J.-L., Calderón–Zygmund operators on product spaces. *Rev. Mat. Iberoamericana*, 1 (1985), 55–91.

[9] Kato, T. & Ponce, G., Commutator estimates and the Euler and Navier–Stokes equations. *Comm. Pure Appl. Math.*, 41 (1988), 891–907.

[10] Kenig, C. E., On the local and global well-posedness theory for the KP-I equation. Preprint, 2003.

[11] Kenig, C. E. & Stein, E. M., Multilinear estimates and fractional integration. *Math. Res. Lett.*, 6 (1999), 1–15.

[12] Lacey, M. & Thiele, C., On Calderón's conjecture. *Ann. of Math.*, 149 (1999), 475–496.

[13] Meyer, Y. & Coifman, R. R., *Ondelettes et opérateurs*, III: *Opérateurs multilinéaires.* Hermann, Paris, 1991.

[14] Muscalu, C., Tao, T. & Thiele, C., Multilinear operators given by singular multipliers. *J. Amer. Math. Soc.*, 15 (2002), 469–496.

[15] — A counterexample to a multilinear endpoint question of Christ and Kiselev. *Math. Res. Lett.*, 10 (2003), 237–246.

[16] — L^p estimates for the "biest", II: The Fourier case. *Math. Ann.*, 329 (2004), 427–461.

[17] Pipher, J., Journé's covering lemma and its extension to higher dimensions. *Duke Math. J.*, 53 (1986), 683-690.

[18] Stein, E. M., *Harmonic Analysis*: *Real Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton Math. Ser., 43. Princeton Univ. Press, Princeton, NJ, 1993.

CAMIL MUSCALU
School of Mathematics
Institute for Advanced Study
Princeton, NJ 08540
U.S.A.
camil@math.ias.edu

JILL PIPHER
Department of Mathematics
Brown University
Providence, RI 02912
U.S.A.
jpipher@math.brown.edu

TERENCE TAO
Department of Mathematics
University of California
Los Angeles, CA 90095
U.S.A.
tao@math.ucla.edu

CHRISTOPH THIELE
Department of Mathematics
University of California
Los Angeles, CA 90095
U.S.A.
thiele@math.ucla.edu

Received November 4, 2003