MULTI-PARAMETER PARAPRODUCTS

CAMIL MUSCALU, JILL PIPHER, TERENCE TAO, AND CHRISTOPH THIELE

Abstract. We prove that the classical Coifman-Meyer theorem holds on any polydisc $\mathbf{T}^d$ of arbitrary dimension $d \geq 1$.

1. Introduction

This article is a continuation of our previous paper [7]. For $n \geq 1$ let $m(= m(\tau))$ in $L^\infty(\mathbb{R}^n)$ be a bounded function, smooth away from the origin and satisfying

$$|\partial^\alpha m(\tau)| \lesssim \frac{1}{|\tau|^{|\alpha|}} \tag{1}$$

for sufficiently many multi-indices α [1]. Denote by $T_m^{(1)}$ the n-linear operator defined by

$$T_m^{(1)}(f_1, ..., f_n)(x) = \int_{\mathbb{R}^n} m(\tau)\widehat{f_1}(\tau_1)...\widehat{f_n}(\tau_n)e^{2\pi i x(\tau_1+...+\tau_n)}\, d\tau \tag{2}$$

where $f_1, ..., f_n$ are Schwartz functions on the real line $\mathbb{R}$. The following statement of Coifman and Meyer is a classical theorem in Analysis [2], [6], [4].

Theorem 1.1. *$T_m^{(1)}$ maps $L^{p_1} \times ... \times L^{p_n} \to L^p$ boundedly, as long as $1 < p_1, ..., p_n \leq \infty$, $\frac{1}{p_1} + ... + \frac{1}{p_n} = \frac{1}{p}$ and $0 < p < \infty$.*

In [7] we considered the bi-parameter analogue of $T_m^{(1)}$ defined as follows. Let $m(= m(\gamma, \eta))$ in $L^\infty(\mathbb{R}^{2n})$ be a bounded function, smooth away from the subspaces $\{\gamma = 0\} \cup \{\eta = 0\}$ and satisfying

$$|\partial_\gamma^\alpha \partial_\eta^\beta m(\gamma, \eta)| \lesssim \frac{1}{|\gamma|^{|\alpha|}} \frac{1}{|\eta|^{|\beta|}} \tag{3}$$

for sufficiently multi-indices α and β. Denote by $T_m^{(2)}$ the n-linear operator defined by

$$T_m^{(2)}(f_1, ..., f_n)(x) = \int_{\mathbb{R}^{2n}} m(\gamma, \eta)\widehat{f_1}(\gamma_1, \eta)...\widehat{f_n}(\gamma_n, \eta_n)e^{2\pi i x[(\gamma_1,\eta_1)+...+(\gamma_n,\eta_n)]}\, d\gamma d\eta \tag{4}$$

where $f_1, ..., f_n$ are Schwartz functions on the plane $\mathbb{R}^2$. The following theorem has been proven in [7].

Theorem 1.2. *$T_m^{(2)}$ maps $L^{p_1} \times ... \times L^{p_n} \to L^p$ boundedly, as long as $1 < p_1, ..., p_n \leq \infty$, $\frac{1}{p_1} + ... + \frac{1}{p_n} = \frac{1}{p}$ and $0 < p < \infty$.*

[1] $A \lesssim B$ means that there exists an universal constant $C > 0$ so that $A \leq CB$.

The main goal of the present paper is to generalize Theorem 1.2 to the d-parameter setting, for any $d \geq 1$.

In general, if $\xi_1 = (\xi_1^i)_{i=1}^d, ..., \xi_n = (\xi_n^i)_{i=1}^d$ are n generic vectors in $\mathbb{R}^d$, they naturally generate the following d vectors in $\mathbb{R}^n$ which we will denote by $\overline{\xi_1} = (\xi_j^1)_{j=1}^n, ..., \overline{\xi_d} = (\xi_j^d)_{j=1}^n$. As before, let $m(= m(\xi) = m(\overline{\xi}))$ in $L^\infty(\mathbb{R}^{dn})$ be a bounded symbol, smooth away from the subspaces $\{\overline{\xi_1} = 0\} \cup ... \cup \{\overline{\xi_d} = 0\}$ and satisfying

$$|\partial_{\overline{\xi_1}}^{\alpha_1} ... \partial_{\overline{\xi_d}}^{\alpha_d} m(\overline{\xi})| \lesssim \prod_{i=1}^d \frac{1}{|\overline{\xi_i}|^{|\alpha_i|}} \tag{5}$$

for sufficiently many multi-indices $\alpha_1, ..., \alpha_d$. Denote by $T_m^{(d)}$ the n-linear operator defined by

$$T_m^{(d)}(f_1, ..., f_n)(x) = \int_{\mathbb{R}^{dn}} m(\xi) \widehat{f_1}(\xi_1) ... \widehat{f_n}(\xi_n) e^{2\pi i x(\xi_1 + ... + \xi_n)} \, d\xi \tag{6}$$

where $f_1, ..., f_n$ are Schwartz functions on $\mathbb{R}^d$. The main theorem of the article is the following.

Theorem 1.3. *$T_m^{(d)}$ maps $L^{p_1} \times ... \times L^{p_n} \to L^p$ boundedly, as long as $1 < p_1, ..., p_n \leq \infty$, $\frac{1}{p_1} + ... + \frac{1}{p_n} = \frac{1}{p}$ and $0 < p < \infty$.*

Traditionally, [2], [6], [4] an estimate as the one in Theorem 1.1 is proved by using the $T(1)$ theorem of David and Journé [8] together with the Calderón-Zygmund decomposition. In particular, the theory of BMO functions and Carleson measures is involved.

On the other hand, it is well known [1], [5] that in the multi-parameter setting all these results and concepts are much more delicate (BMO, John-Nirenberg inequality) or inexistent (Calderón-Zygmund theory). To overcome these difficulties, in [7] we had to develop a completely new approach to prove Theorem 1.2. This approach relied on the one dimensional BMO theory and also on Journé's lemma [5]. At that time, we didn't know how to extend our argument and prove the general d-parameter case.

The novelty of the present paper is that it simplifies the method introduced in [7] and this simplification works equally well in all dimensions. Surprisingly, it turned out that one doesn't need to know anything about BMO, Carleson measures or Journé's lemma in order to prove the estimates in Theorem 1.3.

We shall rely on our previous paper [7] and for the reader's convenience we chose to present the argument in the same bi-linear bi-parameter setting (so both n and d will be equal to 2). However, it will be clear from the proof that its extension to the n-linear d-parameter case is straightforward.

The paper is organized as follows. In Section 2 we recall the discretization procedure from [7] which reduces the study of our operator to the study of some general multi-parameter paraproducts. In Section 3 we present the proof of our main theorem, Theorem 1.3 and in the Appendix we give a proof of Lemma 3.1 which plays an important role in our simplified construction.

Acknowledgements C.Muscalu and J.Pipher were partially supported by NSF Grants. T.Tao is a Clay Prize Fellow and is partially supported by a Packard Foundation Grant.

C.Thiele was partially supported by the NSF Grants 9985572 and DMS 9970469. Two of the authors (C.Muscalu and C.Thiele) are particularly grateful to Centro di Ricerca Matematica Ennio de Giorgi in Pisa for their warm hospitality during their visit in June 2004.

2. Discrete paraproducts

As we promised, assume throughout the paper that $n = d = 2$. In this case, our operator $T_m^{(d)}$ can be written as

$$T_m^{(2)}(f,g)(x) = \int_{\mathbb{R}^4} m(\gamma,\eta)\widehat{f}(\gamma_1,\eta_1)\widehat{g}(\gamma_2,\eta_2)e^{2\pi i x[(\gamma_1,\eta_1)+(\gamma_2,\eta_2)]}\,d\gamma d\eta. \tag{7}$$

In [7] Section 1, we decomposed the operator $T_m^{(2)}$ into smaller pieces, well adapted to its bi-parameter structure. This allowed us to reduce its analysis to the analysis of some simpler discretized dyadic paraproducts. We will recall their definitions below.

An interval I on the real line $\mathbb{R}$ is called dyadic if it is of the form $I = 2^k[n, n+1]$ for some $k, n \in \mathbf{Z}$. If $\lambda, t \in [0,1]$ are two parameters and I is as above, we denote by $I_{\lambda,t}$ the interval $I_{\lambda,t} = 2^{k+\lambda}[n+t, n+t+1]$.

Definition 2.1. *For $J \subseteq \mathbb{R}$ an arbitrary interval, we say that a smooth function Φ_J is a bumb adapted to J, if and only if the following inequalities hold*

$$|\Phi_J^{(l)}(x)| \leq C_{l,\alpha}\frac{1}{|J|^l}\frac{1}{\left(1+\frac{\mathrm{dist}(x,J)}{|J|}\right)^{\alpha}}, \tag{8}$$

for every integer $\alpha \in \mathbb{N}$ and for sufficiently many derivatives $l \in \mathbb{N}$. If Φ_J is a bump adapted to J, we say that $|J|^{-1/2}\Phi_J$ is an L^2 - normalized bump adapted to J.

For $\lambda, t_1, t_2, t_3 \in [0,1]$ and $j \in \{1,2,3\}$ we define the discretized dyadic paraproduct $\Pi^j_{\lambda,t_1,t_2,t_3}$ of "type j" by

$$\Pi^j_{\lambda,t_1,t_2,t_3}(f,g) = \sum_{I\in\mathcal{D}} \frac{1}{|I|^{1/2}}\langle f, \Phi^1_{I_{\lambda,t_1}}\rangle\langle g, \Phi^2_{I_{\lambda,t_2}}\rangle\Phi^3_{I_{\lambda,t_3}}, \tag{9}$$

where f, g are complex-valued measurable functions on $\mathbb{R}$ and $\Phi^i_{I_{\lambda,t_i}}$ are L^2-normalized bumps adapted to I_{λ,t_i} with the additional property that $\int_{\mathbb{R}} \Phi^i_{I_{\lambda,t_i}}(x)dx = 0$ for $i \neq j$, $i = 1,2,3$. $\mathcal{D}$ is an arbitrary finite set of dyadic intervals and by $\langle\cdot,\cdot\rangle$ we denoted the complex scalar product.

Similarly, for $\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0,1]^2$ and $\vec{j} \in \{1,2,3\}^2$, we define the discretized dyadic bi-parameter paraproduct of "type $\vec{j}$

$$\Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}} = \Pi^{j'}_{\lambda',t_1',t_2',t_3'} \otimes \Pi^{j''}_{\lambda'',t_1'',t_2'',t_3''}$$

by

$$\Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g) = \sum_{R\in\vec{\mathcal{D}}} \frac{1}{|R|^{1/2}}\langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_1}}}\rangle\langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}}\rangle\Phi^3_{R_{\vec{\lambda},\vec{t_3}}}, \tag{10}$$

where this time f, g are complex-valued measurable functions on $\mathbb{R}^2$, $R = I \times J$ are dyadic rectangles and $\Phi^i_{R_{\vec{\lambda},\vec{t_i}}}$ are given by

$$\Phi^i_{R_{\vec{\lambda},\vec{t_i}}} = \Phi^i_{I_{\lambda',t'_i}} \otimes \Phi^i_{J_{\lambda'',t''_i}}$$

for $i = 1, 2, 3$. In particular, if $i \neq j'$ then $\int_{\mathbb{R}} \Phi^i_{I_{\lambda',t'_i}}(x)dx = 0$ and if $i \neq j''$ then $\int_{\mathbb{R}} \Phi^i_{J_{\lambda'',t''_i}}(x)dx = 0$. $\vec{\mathcal{D}}$ is an arbitrary finite collection of dyadic rectangles.

We will also denote by $\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f, g, h)$ the trilinear form given by

$$\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f, g, h) = \int_{\mathbb{R}^2} \Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f, g)(x, y)h(x, y)dxdy. \tag{11}$$

In [7] we showed that Theorem 1.2 can be reduced to the following Proposition.

Proposition 2.2. *Fix $\vec{j} \in \{1, 2, 3\}^2$ and let $1 < p, q < \infty$ be two numbers arbitrarily close to 1. Let also $f \in L^p$, $\|f\|_p = 1$, $g \in L^q$, $\|g\|_q = 1$ and $E \subseteq \mathbb{R}^2$, $|E| = 1$. Then, there exists a subset $E' \subseteq E$ with $|E'| \sim 1$[2] such that*

$$\left|\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f, g, h)\right| \lesssim 1 \tag{12}$$

uniformly in the parameters $\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0, 1]^2$, where $h := \chi_{E'}$.

It is therefore enough to prove the above Proposition 2.2, in order to complete the proof of our main Theorem 1.2. Since all the cases are similar, we assume as in [7] that $\vec{j} = (1, 2)$.

To construct the desired set E', we need to recall the "maximal-square", "square-maximal" and "square-square" functions considered in [7].

For $(x, y) \in \mathbb{R}^2$ define

$$MS(f)(x, y) = \sup_I \frac{1}{|I|^{1/2}} \left(\sum_{J:R=I\times J\in\vec{\mathcal{D}}} \sup_{\vec{\lambda},\vec{t_1}} \frac{|\langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_1}}}\rangle|^2}{|J|} \chi_J(y) \right) \chi_I(x), \tag{13}$$

$$SM(g)(x, y) = \left(\sum_I \frac{\sup_{J:R=I\times J\in\vec{\mathcal{D}}} \sup_{\vec{\lambda},\vec{t_2}} \frac{|\langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}}\rangle|^2}{|J|} \chi_J(y)}{|I|} \chi_I(x) \right)^{1/2} \tag{14}$$

and

$$SS(h)(x, y) = \left(\sum_{R\in\vec{\mathcal{D}}} \sup_{\vec{\lambda},\vec{t_3}} \frac{|\langle h, \Phi^3_{R_{\vec{\lambda},\vec{t_3}}}\rangle|^2}{|R|} \chi_R(x, y) \right)^{1/2}. \tag{15}$$

Then, we also recall (see [8]) the bi-parameter Hardy-Littlewood maximal function

[2] $A \sim B$ means that $A \lesssim B$ and $B \lesssim A$

$$MM(F)(x,y) = \sup_{(x,y)\in I\times J} \frac{1}{|I||J|}\int_{I\times J} |F(x',y')|dx'dy'. \tag{16}$$

The following simple estimates explain the appearance of these functions. In particular, we will see that our desired bounds in Theorem 1.2 can be easily obtained as long as all the indices involved are strictly between 1 and ∞.

We start by recalling the following basic inequality, [7]. If Π^1 ia a one-parameter paraproduct of "type 1" given by

$$\Pi^1(f_1,f_2) = \sum_I \frac{1}{|I|^{1/2}}\langle f_1,\Phi_I^1\rangle\langle f_2,\Phi_I^2\rangle\Phi_I^3 \tag{17}$$

then we can write

$$\begin{aligned}\left|\Lambda^1(f_1,f_2,f_3)\right| &= \left|\int_{\mathbb{R}} \Pi^1(f_1,f_2)(x)f_3(x)dx\right| \\ &\lesssim \sum_I \frac{1}{|I|^{1/2}}|\langle f_1,\Phi_I^1\rangle||\langle f_2,\Phi_I^2\rangle||\langle f_3,\Phi_I^3\rangle| \\ &= \int_{\mathbb{R}}\left(\sum_I \frac{|\langle f_1,\Phi_I^1\rangle|}{|I|^{1/2}}\frac{|\langle f_2,\Phi_I^2\rangle|}{|I|^{1/2}}\frac{|\langle f_3,\Phi_I^3\rangle|}{|I|^{1/2}}\chi_I(x)\right)dx \\ &\lesssim \int_{\mathbb{R}} M(f_1)(x)S(f_2)(x)S(f_3)(x)dx \end{aligned} \tag{18}$$

where M denotes the Hardy-Littlewood maximal function and S is the square function of Littlewood and Paley. In particular, we easily see that $\Pi^1 : L^p \times L^q \to L^r$ for any $1 < p,q,r < \infty$ satisfying $1/p + 1/q = 1/r$. Analogous estimates hold for any other type of paraproducts Π^j for $j = 1,2,3$.

Similarly, for the bi-parameter paraproduct $\Pi^{(1,2)}$ of "type $(1,2)$" formally defined by $\Pi^{(1,2)} = \Pi^1 \otimes \Pi^2$ one obtains the inequalities

$$\begin{aligned}\left|\Lambda^{(1,2)}(f_1,f_2,f_3)\right| &= \left|\int_{\mathbb{R}} \Pi^{(1,2)}(f_1,f_2)(x,y)f_3(x,y)dxdy\right| \\ &\lesssim \cdots \lesssim \int_{\mathbb{R}^2} MS(f_1)(x,y)SM(f_2)(x,y)SS(f_3)(x,y)dxdy, \end{aligned} \tag{19}$$

and analogous estimates hold for any other type of paraproducts $\Pi^{\vec{j}}$ for $\vec{j} \in \{1,2,3\}^2$. It is important that all these MS, SM and SS functions are bounded on L^p for any $1 < p < \infty$. We recall the proof of this fact here (see [7]). We start with $SM(f_2)(x,y)$. It can be written as

$$SM(f_2)(x,y) = \left(\sum_{\tilde{I}} \frac{\sup_{\tilde{J}} \frac{|\langle f_2,\Phi_{\tilde{I}}^2\otimes\Phi_{\tilde{J}}^2\rangle|^2}{|\tilde{J}|}\chi_{\tilde{J}}(y)}{|\tilde{I}|}\chi_{\tilde{I}}(x)\right)^{1/2} \tag{20}$$

$$\lesssim \left(\sum_{\tilde{I}} M(\frac{\langle f_2, \Phi^2_{\tilde{I}} \rangle}{|\tilde{I}|^{1/2}})^2(y) \chi_{\tilde{I}}(x) \right)^{1/2}$$

where $\tilde{I}$ and $\tilde{J}$ are the intervals where the corresponding supremums over $\vec{\lambda}, \vec{t_2} \in [0,1]^2$ in (14) are attained.

In particular, by using Fefferman-Stein [3] and Littlewood-Paley [8] inequalities, we have

$$\|SM(f_2)\|_p \lesssim \| \left(\sum_{\tilde{I}} M(\frac{\langle f_2, \Phi^2_{\tilde{I}} \rangle}{|\tilde{I}|^{1/2}})^2(y) \chi_{\tilde{I}}(x) \right)^{1/2} \|_p \tag{21}$$

$$\lesssim \| \left(\sum_{\tilde{I}} \frac{|\langle f_2, \Phi^2_{\tilde{I}} \rangle|^2}{|\tilde{I}|}(y) \chi_{\tilde{I}}(x) \right)^{1/2} \|_p \lesssim \|f_2\|_p$$

for any $1 < p < \infty$. Then, we observe that the MS function is pointwise smaller than a certain SM type function and hence bounded on L^p, while the SS function is a classical double square function and its boundedness on L^p spaces is well known, [1]. As a consequence, it follows as before that $\Pi^{(1,2)} : L^p \times L^q \to L^r$ as long as $1 < p, q, r < \infty$ with $1/p + 1/q = 1/r$.

3. Proof of Proposition 2.2

It remains to prove Proposition 2.2. First, we state the following Lemma.

Lemma 3.1. *Let $J \subseteq \mathbb{R}$ be an arbitrary interval. Then, every bump function ϕ_J adapted to J can be written as*

$$\phi_J = \sum_{k \in \mathbb{N}} 2^{-1000k} \phi^k_J \tag{22}$$

where for each $k \in \mathbb{N}$, ϕ^k_J is also a bump adapted to J but with the additional property that $\text{supp}(\phi^k_J) \subseteq 2^k J$ [3]*. Moreover, if we assume $\int_R \phi_J(x)dx = 0$ then all the functions ϕ^k_J can be chosen so that $\int_{\mathbb{R}} \phi^k_J(x)dx = 0$ for every $k \in \mathbb{N}$.*

The proof of this Lemma will be presented later on in the Appendix. It is the main new ingredient which allows us to simplify our previous argument in [7]. Using it, we can decompose our trilinear form in (10) as

$$\Lambda^{\vec{j}}_{\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3}}(f, g, h) = \sum_{\vec{k} \in \mathbb{N}^2} 2^{-1000|\vec{k}|} \sum_{R \in \vec{\mathcal{D}}} \frac{1}{|R|^{1/2}} \langle f, \Phi^1_{R_{\vec{\lambda}, \vec{t_1}}} \rangle \langle g, \Phi^2_{R_{\vec{\lambda}, \vec{t_2}}} \rangle \langle h, \Phi^{3,\vec{k}}_{R_{\vec{\lambda}, \vec{t_3}}} \rangle, \tag{23}$$

[3] $2^k J$ is the interval having the same center as J and whose length is $2^k|J|$.

where the new functions $\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}$ have basically the same structure as the old $\Phi^{3}_{R_{\vec{\lambda},\vec{t_3}}}$ but they also have the additional property that $\mathrm{supp}(\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}) \subseteq 2^{\vec{k}} R_{\vec{\lambda},\vec{t_3}}$. We denoted by $2^{\vec{k}} R_{\vec{\lambda},\vec{t_3}} := 2^{k_1} I_{\lambda',t_3'} \times 2^{k_2} J_{\lambda'',t_3''}$, $\vec{k} = (k_1, k_3)$ and $|\vec{k}| = k_1 + k_2$.

Fix now f, g, E, p, q as in Proposition 2.2. For each $\vec{k} \in \mathbb{N}^2$ define

$$\Omega_{-5|\vec{k}|} = \{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > C2^{5|\vec{k}|}\} \cup \{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > C2^{5|\vec{k}|}\}. \tag{24}$$

Also, define

$$\tilde{\Omega}_{-5|\vec{k}|} = \{(x,y) \in \mathbb{R}^2 : MM(\chi_{\Omega_{-5|\vec{k}|}})(x,y) > \frac{1}{100}\} \tag{25}$$

and then

$$\tilde{\tilde{\Omega}}_{-5|\vec{k}|} = \{(x,y) \in \mathbb{R}^2 : MM(\chi_{\tilde{\Omega}_{-5|\vec{k}|}})(x,y) > \frac{1}{2^{|\vec{k}|}}\}. \tag{26}$$

Finally, we denote by

$$\Omega = \bigcup_{\vec{k} \in \mathbb{N}^2} \tilde{\tilde{\Omega}}_{-5|\vec{k}|}.$$

It is clear that $|\Omega| < 1/2$ if C is a big enough constant, which we fix from now on. Then, define $E' := E \setminus \Omega$ and observe that $|E'| \sim 1$. We now want to show that the corresponding expression in (12) is $O(1)$ uniformly in the parameters $\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0,1]^2$. Since our argument will not depend on these parameters, we can assume for simplicity that they are all zero and in this case we will write Φ^i_R instead of $\Phi^i_{R_{\vec{\lambda},\vec{t_i}}}$ for $i = 1, 2$ and $\Phi^{3,\vec{k}}_R$ instead of $\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}$.

Fix then $\vec{k} \in \mathbb{N}^2$ and look at the corresponding inner sum in (23). We split it into two parts as follows. Part I sums over those rectangles R with the property that

$$R \cap \tilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset \tag{27}$$

while Part II sums over those rectangles with the property that

$$R \cap \tilde{\Omega}^c_{-5|\vec{k}|} = \emptyset. \tag{28}$$

We observe that Part II is identically equal to zero, because if $R \cap \tilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset$ then $R \subseteq \tilde{\Omega}_{-5|\vec{k}|}$ and in particular this implies that $2^{\vec{k}} R \subseteq \tilde{\tilde{\Omega}}_{-5|\vec{k}|}$ which is a set disjoint from E'. It is therefore enough to estimate Part I only. This can be done by using the technique developed in [7].

Since $R \cap \tilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset$, it follows that $\frac{|R \cap \Omega_{-5|\vec{k}|}|}{|R|} < \frac{1}{100}$ or equivalently, $|R \cap \Omega^c_{-5|\vec{k}|}| > \frac{99}{100}|R|$.

We are now going to describe three decomposition procedures, one for each function f, g, h. Later on, we will combine them, in order to handle our sum.

First, define

$$\Omega_{-5|\vec{k}|+1} = \{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > \frac{C2^{5|\vec{k}|}}{2^1}\}$$

and set

$$\mathbf{T}_{-5|\vec{k}|+1} = \{R \in \vec{\mathcal{D}} : |R \cap \Omega_{-5|\vec{k}|+1}| > \frac{1}{100}|R|\},$$

then define

$$\Omega_{-5|\vec{k}|+2} = \{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > \frac{C2^{5|\vec{k}|}}{2^2}\}$$

and set

$$\mathbf{T}_{-5|\vec{k}|+2} = \{R \in \vec{\mathcal{D}} \setminus \mathbf{T}_{-5|\vec{k}|+1} : |R \cap \Omega_{-5|\vec{k}|+2}| > \frac{1}{100}|R|\},$$

and so on. The constant $C > 0$ is the one in the definition of the set E' above. Since there are finitely many rectangles, this algorithm ends after a while, producing the sets $\{\Omega_n\}$ and $\{\mathbf{T}_n\}$ such that $\vec{\mathcal{D}} = \cup_n \mathbf{T}_n$.

Independently, define

$$\Omega'_{-5|\vec{k}|+1} = \{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > \frac{C2^{5|\vec{k}|}}{2^1}\}$$

and set

$$\mathbf{T}'_{-5|\vec{k}|+1} = \{R \in \vec{\mathcal{D}} : |R \cap \Omega'_{-5|\vec{k}|+1}| > \frac{1}{100}|R|\},$$

then define

$$\Omega'_{-5|\vec{k}|+2} = \{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > \frac{C2^{5|\vec{k}|}}{2^2}\}$$

and set

$$\mathbf{T}'_{-5|\vec{k}|+2} = \{R \in \vec{\mathcal{D}} \setminus \mathbf{T}'_{-5|\vec{k}|+1} : |R \cap \Omega'_{-5|\vec{k}|+2}| > \frac{1}{100}|R|\},$$

and so on, producing the sets $\{\Omega'_n\}$ and $\{\mathbf{T}'_n\}$ such that $\vec{\mathcal{D}} = \cup_n \mathbf{T}'_n$. We would like to have such a decomposition available for the function h also. To do this, we first need to construct the analogue of the set $\Omega_{-5|\vec{k}|}$, for it. Pick $N > 0$ a big enough integer such that for every $R \in \vec{\mathcal{D}}$ we have $|R \cap \Omega''^c_{-N}| > \frac{99}{100}|R|$ where we defined

$$\Omega''_{-N} = \{(x,y) \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x,y) > C2^N\}.$$

Here $SS^{\vec{k}}$ denotes the same "square-square" function defined in (15) but with the functions $\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}$ instead of $\Phi^3_{R_{\vec{\lambda},\vec{t_3}}}$ Then, similarly to the previous algorithms, we define

$$\Omega''_{-N+1} = \{(x,y) \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x) > \frac{C2^N}{2^1}\}$$

and set

$$\mathbf{T}''_{-N+1} = \{R \in \vec{\mathcal{D}} : |R \cap \Omega''_{-N+1}| > \frac{1}{100}|R|\},$$

then define

$$\Omega''_{-N+2} = \{x \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x) > \frac{C2^N}{2^2}\}$$

and set

$$\mathbf{T}''_{-N+2} = \{R \in \vec{\mathcal{D}} \setminus \mathbf{T}''_{-N+1} : |R \cap \Omega''_{-N+2}| > \frac{1}{100}|R|\},$$

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

Then we write Part I as

$$\sum_{n_1,n_2>-5|\vec{k}|,n_3>-N} \sum_{R\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f, \Phi^1_R\rangle||\langle g, \Phi^2_R\rangle||\langle h, \Phi^{3,\vec{k}}_R\rangle||R|, \tag{29}$$

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

$$|R \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}| > \frac{97}{100}|R|. \tag{30}$$

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

$$\sum_{n_1,n_2>-5|\vec{k}|,n_3>-N} \sum_{R\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f, \Phi^1_R\rangle||\langle g, \Phi^2_R\rangle||\langle h, \Phi^{3,\vec{k}}_R\rangle||R \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}| =$$

$$\sum_{n_1,n_2>-5|\vec{k}|,n_3>-N} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}} \sum_{R\in\mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f, \Phi^1_R\rangle||\langle g, \Phi^2_R\rangle||\langle h, \Phi^{3,\vec{k}}_R\rangle|\chi_R(x,y)\,dxdy$$

$$\lesssim \sum_{n_1,n_2>-5|\vec{k}|,n_3>-N} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}\cap\Omega_{\mathbf{T}_{n_1,n_2,n_3}}} MS(f)(x,y)SM(g)(x,y)SS^{\vec{k}}(h)(x,y)\,dxdy$$

$$\lesssim \sum_{n_1,n_2>-5|\vec{k}|,n_3>-N} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|, \tag{31}$$

where

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

On the other hand we can write

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \leq |\Omega_{\mathbf{T}_{n_1}}| \leq |\{(x,y)\in\mathbb{R}^2 : MM(\chi_{\Omega_{n_1}})(x,y) > \frac{1}{100}\}|$$

$$\lesssim |\Omega_{n_1}| = |\{(x,y)\in\mathbb{R}^2 : MS(f)(x,y) > \frac{C}{2^{n_1}}\}| \lesssim 2^{n_1 p}.$$

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, $SS^{\vec{k}}$, MM are bounded on L^s (independently of $\vec{k}$) as long as $1 < s < \infty$ and also that $|E'| \sim 1$. 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 $0 \leq \theta_1, \theta_2, \theta_3 < 1$, such that $\theta_1 + \theta_2 + \theta_3 = 1$.

Now we split the sum in (31) into

$$\sum_{n_1,n_2>-5|\vec{k}|,n_3>0} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| + \sum_{n_1,n_2>-5|\vec{k}|,0>n_3>-N} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|. \tag{33}$$

To estimate the first term in (33) we use the inequality (32) in the particular case $\theta_1 = \theta_2 = 1/2$, $\theta_3 = 0$, 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 $O(2^{10|\vec{k}|})$ and this makes the expression in (23) to be $O(1)$, after summing over $\vec{k}\in\mathbb{N}^2$.

This completes our proof.

It is now clear that our argument works equally well in all dimensions. In the general case, exactly as in [7] Section 1, one first reduces the study of the operator $T_m^{(d)}$ to the study of generic d-parameter dyadic paraproducts $\Pi^{\vec{j}}$ for $\vec{j} = (j_1, ..., j_d) \in \{1,2,3\}^d$ formally defined by $\Pi^{\vec{j}} = \Pi^{j_1}\otimes\cdots\otimes\Pi^{j_d}$. Then, one observes as before, by using the one-parameter theory and Fefferman-Stein inequality, that all the corresponding "square and maximal" type functions which naturally appear in inequalities analogous to (18), (19) are bounded in L^p for $1 < p < \infty$. Having all these ingredients, the argument used in Section 3 works similarly. Finally, the n-linear case follows in the same way. The details are left to the reader.

4. Appendix: proof of Lemma 3.1

In this section we prove Lemma 3.1. Fix $J \subseteq \mathbb{R}$ an interval and let ϕ_J be a bump function adapted to J.

Consider ψ a smooth function such that $\text{supp}(\psi) \subseteq [-1/2, 1/2]$ and $\psi = 1$ on $[-1/4, 1/4]$. If $I \subseteq \mathbb{R}$ is a generic interval with center x_I, we denote by ψ_I the function defined by

$$\psi_I(x) = \psi(\frac{x - x_I}{|I|}). \tag{34}$$

Since

$$1 = \psi_J + (\psi_{2J} - \psi_J) + (\psi_{2^2 J} - \psi_{2J}) + ...$$

it follows that

$$\begin{aligned}\phi_J &= \phi_J \cdot \psi_J + \sum_{k=1}^{\infty} \phi_J \cdot (\psi_{2^k J} - \psi_{2^{k-1} J}) \\ &= \phi_J \cdot \psi_J + \sum_{k=1}^{\infty} 2^{-1000k} \cdot [2^{1000k} \phi_J \cdot (\psi_{2^k J} - \psi_{2^{k-1} J})] \\ &:= \sum_{k=0}^{\infty} 2^{-1000k} \phi_J^k\end{aligned}$$

and it is easy to see that all the ϕ_J^k functions are bumps adapted to J, having the property that $\text{supp}(\phi_J^k) \subseteq 2^k J$.

Suppose now that in addition we have $\int_{\mathbb{R}} \phi_J(x)dx = 0$. This time, we write

$$\begin{aligned}\phi_J &= \phi_J \cdot \psi_J + \phi_J \cdot (1 - \psi_J) \\ &= \left[\phi_J \cdot \psi_J - \left(\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} \cdot \int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right) \cdot \psi_J\right] \\ &+ \left[\left(\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} \cdot \int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right) \cdot \psi_J + \phi_J(1 - \psi_J)\right] \\ &:= \phi_J^0 + R_J^0.\end{aligned}$$

Clearly, by construction we have that $\int_{\mathbb{R}} \phi_J^0(x)dx = 0$ and therefore $\int_{\mathbb{R}} R_J^0(x)dx = 0$. Moreover, ϕ_J^0 is a bump adapted to the interval J having the property that $\text{supp}(\phi_J^0) \subseteq J$. On the other hand, since

$$\begin{aligned}\left|\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} \cdot \int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right| &= \left|\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} \cdot \int_{\mathbb{R}} \phi_J(x)(1 - \psi_J(x))dx\right| \\ &\lesssim 2^{-1000}\end{aligned} \tag{35}$$

it follows that $\|R_J^0\|_\infty \lesssim 2^{-1000}$.

Then, we perform a similar decomposition for the "rest function" R_J^0, but this time we localize it on the larger interval $2J$. We have

$$
\begin{aligned}
R_J^0 &= R_J^0 \cdot \psi_{2J} + R_J^0 \cdot (1 - \psi_{2J}) \\
&= \left[R_J^0 \cdot \psi_{2J} - \left(\frac{1}{\int_{\mathbb{R}} \psi_{2J}(x)dx} \cdot \int_{\mathbb{R}} R_J^0(x)\psi_{2J}(x)dx \right) \cdot \psi_{2J} \right] \\
&+ \left[\left(\frac{1}{\int_{\mathbb{R}} \psi_{2J}(x)dx} \cdot \int_{\mathbb{R}} R_J^0(x)\psi_{2J}(x)dx \right) \cdot \psi_{2J} + R_J^0 \cdot (1 - \psi_{2J}) \right] \\
&:= 2^{-1000}\phi_J^1 + R_J^1.
\end{aligned}
$$

As before, we observe that $\int_{\mathbb{R}} \phi_J^1(x)dx = 0$ and also $\int_{\mathbb{R}} R_J^1(x)dx = 0$. Moreover, ϕ_J^1 is a bump adapted to J whose support lies in $2J$ and $\|R_J^1\|_\infty \lesssim 2^{-1000\cdot 2}$. Iterating this procedure N times, we obtain the decomposition

$$
\phi_J = \sum_{k=0}^{N} 2^{-1000k}\phi_J^k + R_J^N \tag{36}
$$

where all the functions ϕ_J^k are bumps adapted to J with $\int_{\mathbb{R}} \phi_J^k(x)dx = 0$ and $\mathrm{supp}(\phi_J^k) \subseteq 2^k J$, while $\|R_J^N\|_\infty \lesssim 2^{-1000N}$.

This completes the proof of the Lemma.

References

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

[2] Coifman R., Meyer Y., *Ondelettes et operateurs III. Operateurs multilineaires*, Hermann, Paris, [1991].

[3] Fefferman C., Stein E., *Some maximal inequalities*, Amer. J. Math., vol. 93, 107-115, [1971].

[4] Grafakos L., Torres R., *Multilinear Calderón-Zygmund theory*, Adv. Math., vol. 165, 124-164, [2002].

[5] Journé J., *Calderon-Zygmund operators on product spaces*, Rev. Mat. Iberoamericana, vol. 1 55-91, [1985].

[6] Kenig C., Stein E., *Multilinear estimates and fractional integration*, Math. Res. Lett., vol. 6, 1-15, [1999].

[7] Muscalu C., Pipher J, Tao T., Thiele C., *Bi-parameter paraproducts*, submitted, [2003].

[8] Stein, E., *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press, Princeton, [1993]

Department of Mathematics, Cornell University, Ithaca, NY 14853
E-mail address: `camil@math.cornell.edu`

Department of Mathematics, Brown University, Providence, RI 02912
E-mail address: `jpipher@math.brown.edu`

Department of Mathematics, UCLA, Los Angeles, CA 90095
E-mail address: `tao@math.ucla.edu`

Department of Mathematics, UCLA, Los Angeles, CA 90095
E-mail address: `thiele@math.ucla.edu`