A T(1) THEOREM ON PRODUCT SPACES

SANDRA POTT AND PACO VILLARROYA

ABSTRACT. We prove a new T(1) theorem for multiparameter singular integrals

1. Introduction

1.1. **Historical introduction.** In 1984 G. David and J.L. Journé (see [8]) published their celebrated $T(1)$ theorem, a result that characterizes the L^2-boundedness of non-convolution integral operators with a Calderón-Zygmund kernel. In their theorem, the necessary and sufficient conditions for boundedness are expressed by the behaviour of the operator when acting over particular families of funcions: the belonging to BMO of properly defined $T(1)$, $T^*(1)$ functions and the so-called weak boundedness property, which is the fulfillment of L^2 bounds when duality is tested over bump functions with the same space localization.

Since then, many other proofs of this fundamental result in the theory of singular integration have appeared, while it has also been extended to a large variety of settings. Actually, only one year later Journé [20] established the extension to product spaces when he proved an analogous result of L^2-boundedness for multiparameter singular integrals. Those are operators whose class of kernels is homogeneous with respect to non-isotropic dilations of the form $\rho_{\delta_1,\ldots,\delta_n}(x_1,\ldots,x_n) = (\delta_1 x_1,\ldots,\delta_n x_n)$ for $x_i \in \mathbb{R}^{d_i}$ and $\delta_i > 0$, where the number of parameters of the problem coincides with the quantity of independent dilations. The simplest examples of such operators are convolution type operators like the multiple Hilbert transform defined in $\mathbb{R}^n$ by

$$H_1 \cdots H_n(f) = \text{p.v.} f * \frac{1}{x_1 \ldots x_n}$$

or the multiple Riesz transforms defined in $\prod_{i=1}^n \mathbb{R}^{d_i}$ by

$$R_{j_1} \cdots R_{j_n}(f) = f * (\frac{\pi_{j_1}(x_1)}{|x_1|^{d_1+1}} \cdots \frac{\pi_{j_n}(x_n)}{|x_n|^{d_n+1}})$$

where π_{j_i} is the orthogonal projection from $\mathbb{R}^{d_i}$ into $\mathbb{R}$ that "keeps" the j_i-coordinate.

A direct application of Fubini's theorem shows that the multiple Hilbert transform is bounded in all spaces $L^p(\mathbb{R}^n)$ for $1 < p < \infty$. However, the situation is not so simple for more general multiparameter singular integrals, specially if they are of non-convolution

Date: August 11, 2011.
Key words and phrases. Multiparametric singular integral operators, time-frequency analysis.
The second author is supported by Marie-Curie Intra European Fellowship PIEF-GA-2008-220266 and Spanish MEC project MTM2008-04594.

type. These multiparameter operators, even in the simplest cases, are very different from their classical counterparts mainly because the singularities of their kernels lie not only at the origin as in the case of standard Calderón-Zygmund kernels, but instead, they are spread over larger subspaces. For example, in the case of multiple Hilbert transform the set of singularities is the union of the coordinate axes $x_i = 0$. As a consequence, these operators are not in general weak type on $L^1(\mathbb{R}^n)$ and moreover, the strong maximal operator does not control their boundedness properties.

The main motivation to extend the theory of singular integration to operators that commute with multiparameter families of dilations comes from their close relationship with multiplier operators in $\mathbb{R}^n$. Namely, in the same way the classical linear Hilbert transform is closely related to the Fourier partial sum operator $S_N(f)(x) = \sum_{|k|<N} \hat{f}(k)e^{2\pi ikx}$, different multiparameter singular integrals are related to different Fourier partial sum operators in several variables. In particular, the rectangular partial sums operator defined in $\mathbb{R}^n$ by

$$S_{N_1,\dots,N_n}(f)(x_1,\dots,x_n) = \sum_{j=1}^{n} \sum_{|k_j|<N_j} \widehat{f}(k_1,\dots,k_n)e^{2\pi ik_jx_j}$$

is controlled by the multiple Hilbert transform. In section 3 we apply our main result to extend boundedness of product multiplier operators to the non-convolution setting.

We highlight the fact that although the issue about multiparametric singular integrals was intesively studied more than twenty years ago, the field has experienced recently a renewed interest as it can be seen from the recent papers [1], [11], [24], [25], [26].

1.2. **On Journé's theorem.** Journé's result is the first attempt to characterize L^2 boundedness of non-convolution multiparameter singular integral operators. As stated before, many of the classical techniques, like for example a proper Calderón-Zygmund decomposition and the control of singular integrals by means of maximal functions, are no longer available in the multiparameter setting. So, the method Journé chose to overcome such difficulty was the use of vector valued Calderón-Zygmund theory. In order to state his theorem in a simplified form, we require some notation.

Let Δ be the diagonal in $\mathbb{R}^2$ and B be a Banach space. A continuous function $K : \mathbb{R}^2 \setminus \Delta \to B$ is called a vector valued standard Calderón-Zygmund kernel, if for some $0 < \delta \leq 1$ and some constant $C > 0$ we have

$$\|K(x,t)\|_B \leq C|x-t|^{-1}$$

$$\|K(x,t) - K(x',t')\|_B \leq C(|x-x'| + |t-t'|)^{\delta}|x-t|^{-1-\delta}$$

whenever $|x-x'| + |t-t'| \leq |x-t|/2$. In this context, $|K|$ usually denotes the best constant in both inequalities.

Definition 1.1. *A continuous linear mapping T from $C_0^\infty(\mathbb{R})\otimes C_0^\infty(\mathbb{R})$ into its algebraic dual is called a singular integral operator if there are $K^1, K^2 : \mathbb{R}^2\setminus\Delta \to \mathcal{L}(L^2(\mathbb{R}), L^2(\mathbb{R}))$*

vector valued C-Z kernels such that for $f_1, f_2, g_1, g_2 \in C_0^\infty(\mathbb{R})$ *we have*

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f_1(t_1)g_1(x_1)\langle K^1(x_1,t_1)f_2, g_2\rangle dt_1 dx_1$$

whenever $\operatorname{supp} f_1 \cap \operatorname{supp} g_1 = \emptyset$ *and symmetrically for* K^2*.*

The definition of weak boundedness property makes use of the restricted operators: given $f_i, g_i \in C_0^\infty(\mathbb{R})$ for $i = 1, 2$, let $\langle T^i(f_i), g_i\rangle : C_0^\infty(\mathbb{R}) \to C_0^\infty(\mathbb{R})'$ defined by

$$\langle\langle T^1(f_2), g_2\rangle f_1, g_1\rangle = \langle\langle T^2(f_1), g_1\rangle f_2, g_2\rangle = \langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle$$

Notice that the kernel of T^1 for example is precisely $\langle K^1(x_1,t_1)(f_2), g_2\rangle$.

Then, a singular integral operator T is said to satisfy the weak boundedness property if for any bounded subset A of $C_0^\infty(\mathbb{R})$ there is a constant $C > 0$, that may depend on A, such that for any $f, g \in A$ we have that

$$\|\langle T^i(f_{x,R}), g_{x,R}\rangle\|_{CZ} := \|\langle T^i(f_{x,R}), g_{x,R}\rangle\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} + |K^i| \le C$$

where $f_{x,R}(y) = R^{-1/2}f(R^{-1}(y-x))$ and the same for $g_{x,R}$.

Finally, also associated with T we can define its partial adjoints as the adjoint operators with respect to each variable, that is, the operator given by

$$\langle T_1(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \langle T(g_1 \otimes f_2), f_1 \otimes g_2\rangle$$

and analogously for T_2. Notice that $T_2 = T_1^*$.

With all these definitions we can state Journé's result:

Theorem 1.2. *Let* T *be a singular integral operator on* $\mathbb{R}\times\mathbb{R}$ *as described in definition 1.1 satisfying the weak boundedness property and* $T(1), T^*(1), T_1(1), T_1^*(1) \in \mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)$*. Then* T *extends boundedly on* $L^2(\mathbb{R}^2)$*.*

We would like to stress here how restrictive these conditions are, in particular the definitions of singular integral operator and of the weak boundedness property. When written in the language of vector valued Calderón-Zygmund theory they look quite simple, but a more detailed description reveals all their complexity. The sufficient hypotheses for T to be bounded on $L^2(\mathbb{R}^2)$ are the following ones:

a) The K^i are vector valued C-Z kernels. This condition implies that $K^1(x_1,t_1)$ are C-Z operators bounded on $L^2(\mathbb{R})$ and that, moreover, their C-Z norms defined by $\|K^1(x_1,t_1)\|_{CZ} := \|K^1(x_1,t_1)\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} + |K^1_{x_1,t_1}|$ satisfy

$$\|K^1(x_1,t_1)\|_{CZ} \le C|x_i - t_i|^{-1}$$

$$\|K^1(x_1,t_1) - K^1(x_1',t_1')\|_{CZ} \le C(|x_1-x_1'| + |t_1-t_1'|)^\delta |x_1-t_1|^{-1-\delta}$$

whenever $|x_1 - x_1'| + |t_1 - t_1'| \le |x_1 - t_1|/2$ and the same for $K^2(x_2,t_2)$.

b) Weak boundedness property. This condition implies that $\langle T^1(f_{s,R}), g_{s,R}\rangle$ are also C-Z operators bounded on $L^2(\mathbb{R})$ and moreover, their C-Z norms defined as $\|\langle T^1(f_{s,R}), g_{s,R}\rangle\|_{CZ} := \|\langle T^1(f_{s,R}), g_{s,R}\rangle\|_{L^2(\mathbb{R})\to L^2(\mathbb{R})} + |K^1|$ satisfy

$$\|\langle T^1(f_{s,R}), g_{s,R}\rangle\|_{CZ} \le C$$

and the same for T^2.

c) $T(1), T^*(1), \tilde{T}_1(1), \tilde{T}_1^*(1) \in \mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)$, the latter space being much more complex that its one variable counterpart.

So, in order to conclude that the product operator is bounded, Journé's theorem assumes that "some parts" of the operator, in particular the vector valued kernels and the restricted operators, are known to be bounded a priori. This is quite a different situation from the original $T(1)$ theorem in which nothing is assumed to be bounded a priori. However, for the same before mentioned reasons, the use of vector valued theory was also adopted by other authors in later developments of singular integration in product spaces (see [14], [15], [16], [20] and [21]).

Our purpose in the present paper is to state and prove a new $T(1)$ theorem for product spaces in which any hypothesis related to operators which need to be bounded a priori disappear. Therefore, we give up with the use of vector valued Calderón-Zygmund theory and instead, we seek other sufficient hypotheses for L^2-boundedness which are much closer to the spirit of the classical $T(1)$ theorem of David and Journé: conditions related with scalar decay estimates of the kernel and with the behaviour of the operator over special families of functions. To get such new hypotheses, we combine the three classical conditions (kernel estimates, weak boundedness condition and $T(1) \in$ BMO) appropriately to their separated action over different parameters to generate a range of new mixed conditions. For example, in the bi-parameter case we consider new properties by combining two classical ones, namely kernel decay estimates in one parameter and weak boundedness property in the other parameter, to get what we call the mixed 'kernel'-'weak boundedness' condition. As a result, we obtain nine different conditions that cover all possible combinations. This procedure better preserves the symmetry given by the product structure of the kernels and therefore, it is better suited for the general multi-parameter situation.

Moreover, in lemma 8.1 we obtain a decomposition of the operators under consideration which shows that the quantity and the statement of our conditions are the right ones in the sense that they describe entirely their boundedness properties. We would like to highlight the role played by some of those conditions that give sense to a new class of paraproducts which do not appear in previous developments of the theory. We plan a deeper study of such operators in forthcoming papers.

The main advantage of our approach is that, at least in principle, the result can be applied to a larger family of operators since in our hypotheses no operator is ever assumed to be bounded a priori. Actually, none of the examples treated in section 3 are under the scope of Journé's theorem. Moreover, those new conditions should, again in principle, be easier to be tested since there is no need to calculate operators norms.

As a minor advantage of our result we mention the fact that, due to the use of vector-valued C-Z theory, Journé's result needs to impose the following condition on the kernel $\int_{|x-y|>2^k|x-x'|} |K(x,y) - K(x',y)| dy \leq C 2^{-k}$, which is slighlty less general that the classical $\int_{|x-y|>2|x-x'|} |K(x,y) - K(x',y)| dy \leq C$. In our case, some standard arguments allow to apply our theorem to operators whose kernels satisfy the latter condition.

On the other hand, the prize to pay by adopting this new point of view is a larger number of hypotheses, growing rapidly with the number of parameters. In the case of n-parameter operators we have to deal with 3^n hypotheses to ensure that the operator is bounded. However, although Journé's theorem only requires three conditions and so its statement remains as concise as in the uni-parameter case, when the number of parameters grows, these three hypotheses need to be applied iteratively. Then, one might also say that the number of conditions also increases exponentially. From this perspective, the vector valued formulation turns out to be a clever way to encode the complicated structure of the problem and when one unfolds all the information, the complexity always grows accordingly.

Finally, it has to be said that either Journé's theorem and our result exhibit a common weak point: the given sufficient conditions for boundedness of product singular integrals are not necessary. This was first shown by Journé (see the same paper [20]) when he constructed an example of a bounded operator for which the partial adjoint $T_1(1)$ is not in $\mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)$. The problem is that either in his theorem and in ours, the stated conditions imply not only boundedness of T but also of T_1 (and so in such case $T_1(1)$ will have to be in $\mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)$). The underlying reason for this is that the partial adjoint of a bounded operator on $L^2(\mathbb{R}^2)$ is not necessarily bounded. Or in the language of operator spaces, taking adjoints is not a completely bounded map.

The paper is organized as follows. In section 2 we define all the sufficient hypotheses for L^2-boundedness of biparameter singular integral operators and state our $T(1)$ theorem. We also state without proof the analogous results for multiparametric operators in several variables. In section 3 we apply our main result to prove boundedness of non-convolution operators previoulsy studied by R. Fefferman and E. Stein in the convolution setting.

We start the proof of our result in section 4 by the rigourous definition of the functions $T(1)$ and $T(\phi_I \otimes 1)$. In section 5 we obtain an appropriate estimate for the rate of decay of the action of the operator over bump functions when special cancellation properties are assumed. Sections 6 and 7 focus on the proof of L^2 boundedness and the extension to L^p spaces respectively, both of them under the special cancellation hypotheses. The latter case makes use of new bi-parameter modified square functions whose boundedness is a direct consequence of analogous uni-parameter modified square functions. Proof of boundedness of these new square functions is provided in an appendix at the end of the paper. Finally, in section 8 we construct the necessary paraproducts to prove the result in the general case, that is, in absence of the cancellation assumptions.

In a sequel of the present paper, we plan to deal the endpoint case of boundedness from $L^\infty(\mathbb{R}^2)$ into $\mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)$.

We would like to thank Anthony Carbery and Jim Wright for valuable conversations and helpful comments. We would also like to acknowledge the School of Mathematics of the University of Edinburgh for the stimulating research environment provided under which great part of this work was carried out.

2. Definitions and statement of the main theorem

Definition 2.1. *Let Δ be the diagonal in $\mathbb{R}^2$. A function $K : (\mathbb{R}^2 \setminus \Delta) \times (\mathbb{R}^2 \setminus \Delta) \to \mathbb{R}$ is called a product Calderón-Zygmund kernel, if for some $0 < \delta \leq 1$ and some constant $C > 0$ we have*

$$|K(x,t)| \leq C \prod_{i=1,2} \frac{1}{|x_i - t_i|}$$

$$|K(x,t) - K((x_1,x_2'),(t_1,t_2')) - K((x_1',x_2),(t_1',t_2)) + K(x',t')| \leq C \prod_{i=1,2} \frac{(|x_i - x_i'| + |t_i - t_i'|)^\delta}{|x_i - t_i|^{1+\delta}}$$

whenever $2(|x_i - x_i'| + |t_i - t_i'|) \leq |x_i - t_i|$.

Remark 2.1. *The second hypothesis is satisfied if we assume the stronger smoothness condition*

$$|\partial_{t_1}\partial_{t_2}K(x,t)| + |\partial_{t_1}\partial_{x_2}K(x,t)| + |\partial_{x_1}\partial_{t_2}K(x,t)| + |\partial_{x_1}\partial_{x_2}K(x,t)| \leq C \prod_{i=1,2} |x_i - t_i|^{-1-\delta}$$

(notice that the derivatives $\partial_{t_1}\partial_{x_1}K(x,t)$ and $\partial_{t_2}\partial_{x_2}K(x,t)$ do not appear in this condition). This is due to the trivial inequality

$$\begin{aligned}
&|K(x,t) - K((x_1,x_2'),(t_1,t_2')) - K((x_1',x_2),(t_1',t_2)) + K(x',t')| \\
&\leq |K(x,t) - K((x_1,x_2),(t_1,t_2')) - K((x_1,x_2),(t_1',t_2)) + K(x,t')| \\
&+|K((x_1,x_2),(t_1,t_2')) - K((x_1,x_2'),(t_1,t_2')) - K((x_1,x_2),(t_1',t_2')) + K((x_1,x_2'),(t_1',t_2')) \\
&+|K((x_1,x_2),(t_1',t_2)) - K((x_1,x_2),(t_1',t_2')) - K((x_1',x_2),(t_1',t_2)) + K((x_1',x_2),(t_1',t_2')) \\
&+|K(x,t') - K((x_1,x_2'),(t_1',t_2')) - K((x_1',x_2),(t_1',t_2')) + K(x',t')|
\end{aligned}$$

Moreover, as in the classical setting, the same result can be achieved assuming the weaker integral conditions

$$\int_{2|x_i - x_i'| \leq |x_i - t_i|} |K(x,t) - K((x_1,x_2'),t) - K((x_1',x_2),t) + K(x',t)| dt_1 dt_2 \leq C$$

and

$$\int_{2|t_i - t_i'| \leq |x_i - t_i|} |K(x,t) - K(x,(t_1,t_2')) - K(x,(t_1',t_2)) + K(x,t')| dx_1 dx_2 \leq C$$

Given a bilinear form $\Lambda : \mathcal{S}(\mathbb{R}^2) \times \mathcal{S}(\mathbb{R}^2) \to \mathbb{C}$, we define associated linear operators T, adjoint bilinear forms Λ_i, and restricted linear forms Λ^i, in the following way:

Definition 2.2. *(Dual operators). Given a bilinear form Λ, we define linear operators T, T^* through duality:*

$$\langle T(f), g\rangle = \langle f, T^*(g)\rangle = \Lambda(f,g)$$

Definition 2.3. *(Adjoint bilinear forms). We define the adjoint bilinear forms Λ_i such that for $f = f_1 \otimes f_2$, $g = g_1 \otimes g_2$ functions of tensor product type, we have*

$$\Lambda_1(f,g) = \Lambda(g_1 \otimes f_2, f_1 \otimes g_2), \quad \Lambda_2(f,g) = \Lambda(f_1 \otimes g_2, g_1 \otimes f_2)$$

and then extended by linearity and continuity.

These new bilinear forms are also associated with linear operators T_1, T_2 via duality

$$\langle T_i(f), g\rangle = \langle f, T_i^*(g)\rangle = \Lambda_i(f,g)$$

which in the case of tensor products, $f = f_1 \otimes f_2$, $g = g_1 \otimes g_2$, satisfy

$$\langle T_1(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \Lambda_1(f,g) = \Lambda(g_1 \otimes f_2, f_1 \otimes g_2) = \langle T(g_1 \otimes f_2), f_1 \otimes g_2\rangle$$

Notice that $T_2 = T_1^*$ and $T_2^* = T_1$.

From now we will sometimes denote $\Lambda_0 = \Lambda$ and T_0 associated to Λ_0.

Definition 2.4. *(Restricted bilinear forms). We define the restricted bilinear forms by*

$$\langle \Lambda^1(f_2, g_2) f_1, g_1\rangle = \langle \Lambda^2(f_1, g_1) f_2, g_2\rangle = \Lambda(f_1 \otimes f_2, g_1 \otimes g_2)$$

We will call restricted operators T^i to the linear operators associated with the restricted bilinear form Λ^i through duality $\Lambda^i(f_j, g_j) = \langle T^i(f_j), g_j\rangle$.

Notice that the kernels of the forms Λ^i depend on the variables of the functions f_j, g_j and so we will often write $\Lambda^i_{t_j,x_j}$. The same holds for the restricted operators. Also notice that most ot the times we use subindexes to denote the partial adjoint operators or forms while we use superindexes to denote the restricted ones.

Definition 2.5. *A bilinear form $\Lambda : \mathcal{S}(\mathbb{R}^2) \times \mathcal{S}(\mathbb{R}^2) \to \mathbb{C}$ is said to be associated with a product Calderón-Zygmund kernel K if it satisfies the following three integral representations:*

1. *for all Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^2)$ such that $f(\cdot, t_2), g(\cdot, x_2)$ and $f(t_1, \cdot), g(x_1, \cdot)$ have respectively disjoint supports, we have*
$$\Lambda(f,g) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f(t)g(x)K(x,t)\,dx\,dt$$
2. *for all Schwartz functions $f_1, f_2, g_1, g_2 \in \mathcal{S}(\mathbb{R})$ such that f_1 and g_1 have disjoint supports, we have*
$$\Lambda(f,g) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f_1(t_1)g_1(x_1)\Lambda^1(f_2, g_2)\,dt_1\,dx_1$$
3. *analogous representation for Λ^2.*

If the form is continuous on $\mathcal{S}(\mathbb{R}^2) \times \mathcal{S}(\mathbb{R}^2)$ then it will be called a bilinear Calderón-Zygmund form.

With a small abuse of notation, we will say that a bilinear form is bounded on $L^p(\mathbb{R}^2)$ if there is a constant $C > 0$ such that $|\Lambda(f,g)| \leq C\|f\|_{L^p(\mathbb{R}^2)}\|g\|_{L^{p'}(\mathbb{R}^2)}$ for all $f, g \in \mathcal{S}(\mathbb{R}^2)$.

As mentioned in the introduction, boundedness of the bilinear form Λ implies boundedness of the dual linear operators T and T^*, but it does not imply boundedness of any of the adjoint bilinear forms Λ_i nor their corresponding associated adjoint operators T_i, T_i^*. In other words, boundedness of Λ on $L^2(\mathbb{R}^2)$

$$\Big|\Lambda(\sum_n f_1^{n_1}\otimes f_2^{n_2},\sum_m g_1^{m_1}\otimes g_2^{m_2})\Big|\leq C\Big\|\sum_n f_1^{n_1}\otimes f_2^{n_2}\Big\|_2\Big\|\sum_m g_1^{m_1}\otimes g_2^{m_2}\Big\|_2$$

implies boundedness of Λ_1 only on $L^2(\mathbb{R})\hat{\otimes}L^2(\mathbb{R})$:

$$\Big|\Lambda_1(\sum_n f_1^{n_1}\otimes f_2^{n_2},\sum_m g_1^{m_1}\otimes g_2^{m_2})\Big|=\Big|\sum_{n,m}\Lambda(g_1^{m_1}\otimes f_2^{n_2},f_1^{n_1}\otimes g_2^{m_2})\Big|$$

$$\leq C\sum_{n,m}\|g_1^{m_1}\|_2\|f_2^{n_2}\|_2\|f_1^{n_1}\|_2\|g_2^{m_2}\|_2=C\sum_n\|f_1^{n_1}\otimes f_2^{n_2}\|_2\sum_m\|g_1^{m_1}\otimes g_2^{m_2}\|_2$$

Definition 2.6. *For every interval $I\subset\mathbb{R}$ we denote its centre by $c(I)$ its and length by $|I|$. Then, a $L^2(\mathbb{R})$-normalized bump function adapted to I with constant $C>0$ and order $N\in\mathbb{N}$, is a Schwartz function ϕ such that*

$$|\phi^{(n)}(x)|\leq C|I|^{-1/2-n}(1+|I|^{-1}|x-c(I)|)^{-N},\quad 0\leq n\leq N$$

We will call standard cube to the cube in $\mathbb{R}^d$ of measure one centered at the origin. A bump function ϕ adapted to the standard cube of order N is a Schwartz function satisfying

$$|\partial^\alpha\phi(x)|\leq C(1+|x|)^{-N},\quad 0\leq|\alpha|\leq N$$

A bump function ϕ is said to be adapted to a box R in $\mathbb{R}^d$, if for any affine linear transformation A transforming the standard cube into the box R, the function

$$|\det(A)|^{1/2}\phi(Ax)$$

is a bump function adapted to the standard cube. This definition does not depend on the finite choice of the linear transformation.

Observe that all these bump functions are normalized to be uniformly bounded in $L^2(\mathbb{R}^d)$. The order of the bump functions will always be denoted by N, even though its value might change from line to line. It is also worthy to say that we usually reserve the greek letter ϕ, φ for general bump functions while we reserve the use of ψ to denote bump functions with mean zero.

Definition 2.7. *We say that a bilinear form Λ satisfies the weak boundedness condition, if for any rectangle R and every pair ϕ_R,φ_R of L^2-normalized bump functions adapted to R with constant C, we have*

$$|\Lambda(\phi_R,\varphi_R)|\leq C$$

Definition 2.8. *We say that a bilinear form Λ satisfies the mixed weak boundedness-Calderón Zygmund condition, if for any interval I and every pair ϕ_I,φ_I of L^2-normalized bump functions adapted to I with constant C, we have*

$$|\Lambda^i_{t_j,x_j}(\phi_I,\varphi_I)|\leq C|t_j-y_j|^{-1}$$

$$|(\Lambda^i_{t_j,x_j} - \Lambda^i_{t'_j,x'_j})(\phi_I,\varphi_I)| \leq C(|x_j - x'_j| + |t_j - t'_j|)^\delta |t_j - x_j|^{-(1+\delta)}$$

whenever $2(|x_j - x'_j| + |t_j - t'_j|) < |t_j - x_j|$ *for all* $i, j \in \{1, 2\}$.

Obviously, the second condition is implied by the smoothness condition

$$|\partial_{t_j}\Lambda^i_{t_j,x_j}(\phi_I,\varphi_I)| + |\partial_{x_j}\Lambda^i_{t_j,x_j}(\phi_I,\varphi_I)| \leq C|t_j - x_j|^{-(1+\delta)}$$

Finally, we notice that, in order to simplify notation, from now being the space product BMO, that is, the dual of $H^1(\mathbb{R}^2)$ will be simply denote by $\mathrm{BMO}(\mathbb{R}^2)$.

We can now state our main result,

Theorem 2.9. *(bi-parameter T(1) theorem). Let* Λ *be a bilinear Calderón-Zygmund form satisfying the mixed WB-CZ conditions. Then, the following are equivalent:*

(1) Λ_i *are bounded bilinear forms on* $L^2(\mathbb{R}^2)$ *for all* $i = 0, 1, 2$,
(2) Λ *satisfies the weak boundedness condition and the special cancellation conditions:*
  a) $T(1), T^*(1), T_1(1), T_1^*(1) \in \mathrm{BMO}(\mathbb{R}^2)$,
  b) $\langle T(\phi_I \otimes 1), \varphi_I \otimes \cdot\rangle, \langle T(1 \otimes \phi_I), \cdot \otimes \varphi_I\rangle, \langle T^*(\phi_I \otimes 1), \varphi_I \otimes \cdot\rangle, \langle T^*(1 \otimes \phi_I), \cdot \otimes \varphi_I\rangle \in \mathrm{BMO}(\mathbb{R})$ *for all* ϕ_I, φ_I *bump functions adapted to* I *with norms uniformly bounded in* I.

We remark that boundedness of those operators T_i and T_j for $i \neq j$ are not equivalent. A way to show this is by considering Carleson's example that proves $\mathrm{BMO}_{\mathrm{rec}}(\mathbb{R}^2) \subsetneq \mathrm{BMO}(\mathbb{R}^2)$. In [2], he described a recursive process to construct a sequence of functions such that $\|b_n\|_{\mathrm{BMO}_{\mathrm{rec}}(\mathbb{R}^2)} = 1$ while $\|b_n\|_{\mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)} \geq C_n$ where $(C_n)_{n\in\mathbb{N}}$ was an unbounded sequence of positive numbers. Then, we can consider paraproducts associated with that sequence of functions

$$T_n(f) = \sum_I \langle b_n, \psi_R\rangle\langle f, \psi_R^2\rangle\psi_R$$

in such a way that $\|T_n\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \approx \|b_n\|_{\mathrm{BMO}_{\mathrm{prod}}(\mathbb{R}^2)} \geq C_n$ while $\|T_n^*\|_{L^2(\mathbb{R}^2)\to L^2(\mathbb{R}^2)} \approx \|b_n\|_{\mathrm{BMO}_{\mathrm{rect}}(\mathbb{R}^2)} \leq C$. This shows again that none of the conditions $T_1(1) \in \mathrm{BMO}(\mathbb{R}^2)$, $T_1^*(1) \in \mathrm{BMO}(\mathbb{R}^2)$ are necessary for boundedness of T.

We end this section by stating the analogous result in the multiparameter case. We simplify the notation as much as possible.

Let $m \leq n$ and $n_1, \ldots, n_m$ such that $n = \sum_{i=1}^m n_i$. Let $K : \prod_{i=1}^m (\mathbb{R}^{n_i} \setminus \Delta_{n_i}) \to \mathbb{R}$ be such that

$$|K(x,t)| \leq C \prod_{i=1}^m |x_i - t_i|^{-n_i}$$

$$|\nabla_{t_{i_1}} \cdots \nabla_{t_{i_m}} K(x,t)| \leq C \prod_{i=1}^m |x_i - t_i|^{-(n_i+\delta)}$$

where $x_i, t_i \in \mathbb{R}^{n_i}$, $0 < \delta \leq 1$.

Definition 2.10. *(Restricted bilinear forms). Let $N_1, N_2 \subset \{1, \dots, m\}$ such that $N_1 \cup N_2 = \{1, \dots, m\}$ disjointly. Given a bilinear form Λ, we define the restricted bilinear forms by*

$$\langle \Lambda^{N_1}(\otimes_{j\in N_1} f_j, \otimes_{j\in N_1} g_j) \otimes_{j\in N_2} f_j, \otimes_{j\in N_2} g_j\rangle = \Lambda(f,g)$$

for $f = \otimes_{i=1}^m f_i$, $g = \otimes_{i=1}^m g_i$ with $f_i, g_i \in \mathcal{S}(\mathbb{R}^{n_i})$, and then extended by linearity and continuity.

We will call restricted operators to the linear operators associated with the restricted bilinear forms by duality.

Notice that the kernels of the forms Λ^{N_1} depend on the variables of the functions $\otimes_{j\in N_2} f_j$, $\otimes_{j\in N_2} g_j$ and so we can write $\Lambda^{N_1}_{t_j,x_j}$.

Definition 2.11. *A bilinear form $\Lambda : \mathcal{S}(\mathbb{R}^n)\times\mathcal{S}(\mathbb{R}^n) \to \mathbb{C}$ is said to be associated with a product Calderón-Zygmund kernel K if it satisfies the following integral representations:*

(1) *for all Schwartz functions $f, g \in \mathcal{S}(\mathbb{R}^n)$ such that $f(\cdot, t_2), g(\cdot, x_2)$ and $f(t_1, \cdot), g(x_1, \cdot)$ have respectively disjoint supports, we have*

$$\Lambda(f,g) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} f(t)g(x)K(x,t)\,dx\,dt$$

(2) *for every $N_1, N_2 \subset \mathbb{N}$ such that $N_1 \cup N_2 = \{1, \dots, m\}$ disjointly and for all Schwartz functions $f_i, g_i \in \mathcal{S}(\mathbb{R})$ such that f_j and g_j with $j \in N_2$ have disjoint supports, we have the integral representation*

$$\Lambda(f,g) = \int_{\mathbb{R}^{n_{N_2}}}\int_{\mathbb{R}^{n_{N_2}}} \otimes_{j\in N_2} f_j(t_j) \otimes_{j\in N_2} g_j(x_j) \Lambda_{N_1}(\otimes_{j\in N_1} f_j, \otimes_{j\in N_1} g_j)\,dt_j\,dx_j$$

with $n_{N_2} = \sum_{j\in N_2} n_j$.

Definition 2.12. *We say that a bilinear form Λ satisfies the weak boundedness condition if for any box $R \subset \mathbb{R}^n$ and every ϕ_R, φ_R $L^2(\mathbb{R}^n)$-normalized bump functions adapted to R with constant C we have*

$$|\Lambda(\phi_R, \varphi_R)| \le C$$

Definition 2.13. *Let $WB, CZ \subset \{1, \dots, m\}$ such that $WB \cup CZ = \{1, \dots, m\}$ disjointly. We say that a bilinear form Λ satisfies the mixed weak boundedness-Calderón Zygmund condition if for any $R = \prod_{i\in WB} R_i$ and every ϕ_R, φ_R bump functions $L^2(\mathbb{R}^{\sum_{i\in WB} n_i})$-normalized and adapted to R with constant $C > 0$, we have*

$$|\Lambda^{WB}_{t_{i_j},x_{i_j}}(\phi_R, \varphi_R)| \le C \prod_{j\in CZ} |t_{i_j} - y_{i_j}|^{-1}$$

$$|(\Lambda^{WB}_{t_{i_j},x_{i_j}} - \Lambda^{i}_{t'_{i_j},x'_{i_j}})(\phi_R, \varphi_R)| \le C \prod_{j\in CZ} (|t_{i_j} - t'_{i_j}| + |x_{i_j} - x'_{i_j}|)^{\delta} |t_{i_j} - x_{i_j}|^{-(n_j+\delta)}$$

Theorem 2.14. *(m-parameter $T(1)$ theorem). Let Λ be a bilinear Calderón-Zygmund form with associated kernel K satisfying the mixed WB-CZ conditions.*

Then, the following are equivalent:

(1) Λ_i *are bounded bilinear forms on* $L^2(\mathbb{R}^n)$ *for all* i,

(2) Λ *satisfies the weak boundedness condition and the following sequence of special cancellation conditions: for every* $k \in \{n_1, n_1+n_2, \ldots, n\}$ *and all bump functions* $\phi_{R_i}, \varphi_{R_i} \in \mathcal{S}(\mathbb{R}^{n_i})$ *both adapted to* $R_i \subset \mathbb{R}^{n_i}$ *we have*

$$\langle T(1 \otimes \ldots \otimes 1 \otimes \Phi_{R^{n-k}}), \tilde{\Phi}_{R^{n-k}} \rangle \in \mathrm{BMO}(\mathbb{R}^k)$$

where $\Phi_{R^{n-k}} = \phi_{R_1} \otimes \ldots \otimes \phi_{R_{n-k}}$ *and* $\tilde{\Phi}_{R^{n-k}} = \varphi_{R_1} \otimes \ldots \otimes \varphi_{R_{n-k}}$ *with norms uniformly bounded when varying over the boxes* $R^{n-k} = \prod_{i=1}^{n-k} R_i$. *The same condition applies for all possible permutations of the entries in* $1 \otimes \ldots \otimes 1 \otimes \Phi_{R^{n-k}}$ *(in total* $\binom{n}{n-k} 2^k$ *conditions).*

3. Application

We give now an example of how our results can be applied to the study of boundedness of operators defined by product kernels.

In [17], R. Fefferman and E. Stein explain that in some boundary-value problems, in particular in the $\bar{\partial}$-Neuman problem, one faces convolution operators defined in $\mathbb{R}^{n+1}$ with kernels like

$$K_k(t, t_{n+1}) = \frac{t_k}{(|t|^2 + t_{n+1}^2)^{(n+1)/2}} \frac{1}{|t|^2 + it_{n+1}}$$

with $t \in \mathbb{R}^n$ and $t_{n+1} \in \mathbb{R}$, which are product of two kernels with different types of homogeneity. With this motivation in mind they prove the following

Theorem 3.1. *Let* K *be a kernel defined in* $\mathbb{R}^n \times \mathbb{R}^m$ *by* $K(t) = K_1(t)K_2(t)$ *such that* K_1 *is homogeneous of degree* $-n$ *with respect the family of dilations* $t \to (\delta t_1, \delta^a t_2)$ *for all* $\delta > 0$ *and fixed* $a > 0$ *while* K_2 *is homogeneous of degree* $-m$ *with respect the family of dilations* $t \to (\delta^b t_1, \delta t_2)$ *for all* $\delta > 0$ *and fixed* $b > 0$.

It is also assumed that $K_1(t_1, 0)$ *has mean zero on the unit sphere of* $\mathbb{R}^n$, $K_2(0, t_2)$ *has mean zero on the unit sphere of* $\mathbb{R}^m$ *and*

$$\Big| \int_{\alpha_1 < |t_1| < \beta_1, \alpha_2 < |t_2| < \beta_2} K(t) dt \Big| \leq A$$

for all $0 < \alpha_i < \beta_i$. *Then, for all* $1 < p < \infty$,

$$\|K * f\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m)} \leq C_p \|f\|_{L^p(\mathbb{R}^n \times \mathbb{R}^m)}$$

where the constant C_p *depends on* A *and* p.

In their paper, Theorem 3.1 appears as a corollary of the following more general result:

Theorem 3.2. *Let* $K : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ *be an integrable function that satisfies*

(1) *the kernel conditions: for* $t = (t_1, t_2)$, $h = (h_1, h_2)$,
  - (a) $|K(t)| \leq A|t_1|^{-n}|t_2|^{-m}$
  - (b) $|K(t_1 + h_1, t_2) - K(t)| \leq A|h|^{\delta_1}|t_1|^{-n-\delta_1}|t_2|^{-m}$ *whenever* $2|h_1| < |t_1|$

(c) $|K(t_1,t_2+h_2)-K(t)| \leq A|h|^{\delta_2}|t_1|^{-n}|t_2|^{-m-\delta_2}$ *whenever* $2|h|<|t_2|$
(d) $|K(t+h)-K(t_1+h_1,t_2)-K(t_1,t_2+h_2)+K(t)| \leq A|h_1|^{\delta_1}|h_2|^{\delta_2}|t_1|^{-n-\delta_1}|t_2|^{-m-\delta_2}$ *whenever* $2|h_1|<|t_1|$, $2|h_2|<|t_2|$

(2) *the cancellation condition:* $\left|\int_{\alpha_i<|t_i|<\beta_i} K(t)dt\right| \leq A$ *for all* $0<\alpha_i<\beta_i$

(3) *the mixed kernel-cancellation conditions:*
(a) *if* $K_1(t_1)=\int_{\alpha_2<|t_2|<\beta_2} K(t_1,t_2)dt_2$ *then*
(i) $|K_1(t_1)| \leq A|t_1|^{-n}$
(ii) $|K_1(t_1+h_1)-K_1(t_1)| \leq A|h|^{\delta_1}|t_1|^{-n-\delta_1}$ *whenever* $2|h_1|<|t_1|$
(b) *similar conditions for* $K_2(t_2)=\int_{\alpha_1<|t_1|<\beta_1} K(t_1,t_2)dt_1$.

Then, for all $1<p<\infty$,

$$\|f*K\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \leq A_p\|f\|L^p(\mathbb{R}^n\times\mathbb{R}^m)$$

with A_p *depending only on* A *and* p.

It is not difficult to see that that conditions of Theorem 3.2 imply the hypotheses of Theorem 2.9. Conditions $(1)-a)$ and $(1)-d)$ imply K is a product Calderón-Zygmund standard kernel while (2) implies that the convolution operator T with kernel K satisfies weak boundedness condition and $T(1), T^*(1), T_1(1), T_1^*(1) \in \mathrm{BMO}(\mathbb{R}^n\times\mathbb{R}^m)$. On the other hand, the mixed type hypotheses of Theorem 2.9, that is, the mixed WB-CZ condition and mixed T(1)-CZ conditions follow from $(1)-b)$, $(1)-c)$; and (3) respectively.

Even more, Theorem 2.9 allow us two extend Theorem 3.2 to the case of non-convolution kernels, a result that is stated below.

Definition 3.3. *Let* $K(x,t)$ *with* $x,t\in\mathbb{R}^n\times\mathbb{R}^m$ *be an integrable function that satisfies*

(1) *the kernel conditions:*
(a) $|K(x,t)| \leq A|t_1|^{-n}|t_2|^{-m}$
(b) $|K((x_1',x_2),(t_1',t_2))-K(x,t)| \leq A(|x_1-x_1'|+|t_1-t_1'|)^{\delta_1}|t_1|^{-n-\delta_1}|t_2|^{-m}$ *whenever* $2(|x_1-x_1'|+|t_1-t_1'|)<|t_1|$
(c) $|K((x_1,x_2'),(t_1,t_2'))-K(x,t)| \leq A(|x_2-x_2'|+|t_2-t_2'|)^{\delta_2}|t_1|^{-n}|t_2|^{-m-\delta_2}$ *whenever* $2(|x_2-x_2'|+|t_2-t_2'|)|<|t_2|$
(d) $|K(x',t')-K((x_1',x_2),(t_1',t_2))-K((x_1,x_2'),(t_1,t_2'))+K(x,t)|$ $\leq A(|x_1-x_1'|+|t_1-t_1'|)^{\delta_1}(|x_2-x_2'|+|t_2-t_2'|)^{\delta_2}|t_1|^{-n-\delta_1}|t_2|^{-m-\delta_2}$ *whenever* $2(|x_i-x_i'|+|t_i-t_i'|)<|t_i|$ *for* $i=1,2$

(2) *the cancellation condition:* $\left|\int_{\alpha_1<|t_1|<\beta_1,\alpha_2<|t_2|<\beta_2} K(x,t)dt\right| \leq A$

(3) *the mixed kernel-cancellation conditions:*
(a) *if* $K_1(x,t_1)=\int_{\alpha_2<|t_2|<\beta_2} K(x,t_1,t_2)dt_2$ *then*
(i) $|K_1(x,t_1)| \leq A|t_1|^{-n}$
(ii) $|K_1((x_1',x_2),t_1')-K_1(x,t_1)| \leq A(|x_1-x_1'|+|t_1-t_1'|)^{\delta_1}|t_1|^{-n-\delta_1}$ *whenever* $2(|x_1-x_1'|+|t_1-t_1'|)<|t_1|$
(b) *similar conditions for* $K_2(x,t_2)=\int_{\alpha_1<|t_1|<\beta_1} K(x,t_1,t_2)dt_1$.

Definition 3.4. *We say that an operator T is associated with K if*

$$T(f)(x) = \int_{\mathbb{R}^n\times\mathbb{R}^m} f(x-t)K(x,t)dt$$

whenever $x \notin \sup(f)$.

Then, we have

Theorem 3.5. *Let T be an operator associated with K satisfying all the conditions of definition 3.3. Then, For all $1 < p < \infty$,*

$$\|T(f)\|_{L^p(\mathbb{R}^n\times\mathbb{R}^m)} \le A_p\|f\|L^p(\mathbb{R}^n\times\mathbb{R}^m)$$

with A_p depending only on A and p.

Sketch of proof. To give a flavour of the ideas involved on dealing with non convolution kernels in the product seeting, we outline how hypotheses of Theorem 3.5 imply the ones in Theorem 2.9. In particular, we partially show how the mixed weak boundedness Calderón-Zygmund condition $|\langle T^1_{x_1,t_1}(\phi_I),\varphi_I\rangle| \le C|x_1-t_1|^{-n}$ and the cancellation property $\langle T(\phi_I\otimes 1),\varphi_I\otimes\cdot\rangle \in \mathrm{BMO}(\mathbb{R})$ are checked.

Let I be a fixed interval. We consider $\phi_I = |I|^{-1/2}\sum_k a_k\chi_{I_k}$ to be an approximation by step functions of a general bump function adapted to I, where the intervals I_k are pairwise disjoint and of the same lenght $|I_k| < \epsilon$ as small as we want. We consider a similar description for φ_I. Then,

$$\langle T^1_{x_1,t_1}(\phi_I),\varphi_I\rangle = \lim_{\epsilon\to 0}|I|^{-1}\sum_{k,j\in\mathbb{Z}} a_kb_j\int_{|x_2-t_2|>\epsilon}\chi_{I_k}(x_2)\chi_{I_j}(t_2)K(x,x-t)dt_2dx_2$$

so we just need to bound

$$|I|^{-1}\sum_{k,j\in\mathbb{Z}} a_kb_jT_{k,j}$$

where $T_{k,j}$ denotes the integral in the sum, independently of $\epsilon > 0$ and $|I|$.

When $k = j$,

$$T_{k,k} = \int_{\substack{|x_2-c(I_k)|<|I_k|/2\\ |t_2-c(I_k)|<|I_k|/2\\ |x_2-t_2|>\epsilon}} K(x,x-t)dt_2dx_2 = \int_{\substack{|x_2|<|I_k|/2\\ |t_2|<|I_k|/2\\ |x_2-t_2|>\epsilon}} K(x_1,x_2-c(I_k),x-t)dt_2dx_2$$

$$= \int_{|x_2|<|I_k|/2}\int_{\substack{|x_2-t_2|<|I_k|/2\\ |t_2|>\epsilon}} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2dx_2$$

$$= \int_{0<x_2<|I_k|/2}\Big(\int_{\substack{|x_2-t_2|<|I_k|/2\\ |t_2|>\epsilon}} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2$$

$$+\int_{\substack{|-x_2-t_2|<|I_k|/2\\ |t_2|>\epsilon}} K(x_1,-x_2-c(I_k),x_1-t_1,t_2)dt_2\Big)dx_2$$

By addition and substraction the inner sum in last expression equals

(3.1)

$$\int_{\substack{|x_2-t_2|<|I_k|/2\\|t_2|>\epsilon}} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2+\int_{\substack{|-x_2-t_2|<|I_k|/2\\|t_2|>\epsilon}} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2$$

$$+\int_{\substack{|-x_2-t_2|<|I_k|/2\\|t_2|>\epsilon}} \big(K(x_1,-x_2-c(I_k),x_1-t_1,t_2)-K(x_1,x_2-c(I_k),x_1-t_1,t_2)\big)dt_2$$

The first two terms can be rewritten with a symmetric domain of integration as

$$\int_{|x_2-|I_k|/2|<|t_2|<|x_2+|I_k|/2|} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2$$

$$+2\int_{\epsilon<|t_2|<|x_2-|I_k|/2|} K(x_1,x_2-c(I_k),x_1-t_1,t_2)dt_2$$

where the second integral is zero if $|x_2-|I_k|/2|\leq\epsilon$. Then, by the hypothesis 3.a.i) with $\alpha_2=|x_2-|I_k|/2|$, $\beta_2=|x_2+|I_k|/2|$ for the first integral and $\alpha_2=\epsilon$, $\beta_2=|x_2-|I_k|/2|$ for the second one, we can bound them by $3A|x_1-t_1|^{-n}$.

Meanwhile, the last term in expression (3.1) can be treated by condition 1.c) and bounded by

$$\int_{\substack{|-x_2-t_2|<|I_k|/2\\|t_2|>\epsilon}} 2|x_2|^{\delta_2}|x_1-t_1|^{-n}|t_2|^{-m-\delta_2}dt_2\leq C|x_2|^{\delta_2}|x_1-t_1|^{-n}$$

With both things we get

$$\Big|\int_{\substack{|x_2-c(I_k)|<|I_k|/2\\|t_2-c(I_k)|<|I_k|/2\\|x_2-t_2|>\epsilon}} K(x-t)dt_2dx_2\Big|\leq\int_{0<x_2<|I_k|/2} 3A(1+|x_2|^{\delta_2})|x_1-t_1|^{-n}dx_2$$

$$\leq 3A|x_1-t_1|^{-n}(1+|I_k|^{\delta})|I_k|/2\leq 3A|x_1-t_1|^{-n}(1+\epsilon)|I_k|/2$$

Finally, since we may assume $|b_k|\leq 1$ and $\sum_{k\in\mathbb{Z}}|a_k||I_k|\leq C|I|$, we end this case with the following bounds:

$$|I|^{-1}\sum_{k\in\mathbb{Z}}|a_k||b_k|T_{k,k}|\leq C|I|^{-1}\sum_{k\in\mathbb{Z}}|a_k||b_k|CA|x_1-t_1|^{-n}|I_k|\leq CA|x_1-t_1|^{-n}$$

The case $k\neq j$ is technically more complex since we need to consider several terms together in order to get the same kind of symmetry in the domain of integration. Despite this, the same type of ideas apply: the kernel decay estimates allow to obtain a similar result and prove this way the mixed WB-CZ condition.

On the other hand, let ψ_J an atom. Then,

$$\langle T(\phi_I\otimes 1),\varphi_I\otimes\psi_J\rangle$$

$$=\lim_{\lambda\to\infty}\lim_{\epsilon\to 0}\iint_{|x_2-t_2|>\epsilon}\chi_{\lambda I}(x_2)\psi_J(t_2)|I|^{-1}\sum_{k,j\in\mathbb{Z}}a_kb_j\iint_{|x_1-t_1|>\epsilon}\chi_{I_k}(x_1)\chi_{I_j}(t_1)K(x,x-t)dtdx$$

$$= \lim_{\lambda\to\infty}\lim_{\epsilon\to 0}\iint_{|x_2-t_2|>\epsilon} \chi_{\lambda I}(x_2)\psi_J(t_2)|I|^{-1}\sum_{k,j\in\mathbb{Z}} a_k b_j T_{k,j}(x_2,t_2)dx_2dt_2$$

and we bound the last expression independently of $\lambda > 0$. By using the mean zero of ψ_J this is equal to

$$\lim_{\epsilon\to 0}\iint_{|x_2-t_2|>\epsilon} \chi_{\lambda I}(x_2)\psi_J(t_2)\Big(|I|^{-1}\sum_{k,j\in\mathbb{Z}} a_k b_j T_{k,j}(x_2,t_2)dx_2 - |I|^{-1}\sum_{k,j\in\mathbb{Z}} a_k b_j T_{k,j}(x_2,0)\Big)dx_2dt_2$$

now by a similar argument as before but using the smoothness condition instead the decay we can bound by

$$\lim_{\epsilon\to 0}\int |\psi_J(t_2)|\int_{|x_2-t_2|>\epsilon} |\chi_{\lambda I}(x_2)|CA|x_2-t_2|^{-(n+\delta)}dx_2dt_2 \le CA\|\psi_J\|_{L^1(\mathbb{R})}$$

4. Definition of $T(1)$, $\langle T(\phi_I\otimes 1),\varphi_I\rangle$ and $\langle T(\phi_I\otimes 1),\varphi_J\rangle$

In this section we give a rigorous definition of $T(1)$, $\langle T(\phi_I\otimes 1),\varphi_I\rangle$ and $\langle T(\phi_I\otimes 1),\varphi_J\rangle$ as distributions modulo constants. The approach is similar to the uni-parametric case and so we will follow some of the arguments in [28].

We start with the technical lemma that gives sense to $T(1)$ (and also the partial adjoints $T_i(1)$). The condition $T(1)\in \mathrm{BMO}(\mathbb{R}^2)$ means that the following inequality

$$|\langle T(1), f\rangle| \le C\|f\|_{H^1(\mathbb{R}^2)}$$

holds for all f that belong to a dense subset of $H^1(\mathbb{R}^2)$. In our case, such dense subset will be the family of Schwartz functions f compactly supported with mean zero in each variable, meaning $\int_{\mathbb{R}} f(x,t)dx = \int_{\mathbb{R}} f(x,t)dt = 0$. Then, in order to give a proper sense to the left hand side of previous inequality we use the following lemma:

Lemma 4.1. *Let $\Phi\in\mathcal{S}(\mathbb{R}^2)$ such that $\Phi(x)=1$ for $\|x\|\le 1$ and $\Phi(x)=0$ for $\|x\|>2$. Let Λ be a Calderón-Zygmund bilinear form with associated kernel K. Let S be a rectangle and $f\in\mathcal{S}(\mathbb{R}^2)$ with compact support in S and mean zero in each variable. Then, the limit*

$$L(f) = \lim_{k_1,k_2\to\infty}\Lambda(T_{c(S)}D_{2^{k_1}|S_1|,2^{k_2}|S_2|}\Phi, f)$$

exists. Moreover, we have the error bound

$$|L(f) - \Lambda(T_{c(S)}D_{2^{k_1}|S_1|,2^{k_2}|S_2|}\Phi, f)| \le C2^{-\delta(k_1+k_2)}\|f\|_{L^1(\mathbb{R}^2)}$$

where δ is the parameter in the Calderon-Zygmund property of the kernel K and the constant depends only on Φ and Λ.

Proof. For simplicity of notation we shall assume that S is centered at the origin.

For $k\in\mathbb{N}^2$ with $k_i\ge 1$, we set $\psi_k = D_{2^{k_1}|S_1|,2^{k_2}|S_2|}\Phi - D_{2^{k_1-1}|S_1|,2^{k_2-1}|S_2|}\Phi$ and we estimate $|\Lambda(\psi_k, f)|$.

Since the supports of ψ_k and f are disjoint, we use the kernel representation

$$\Lambda(\psi_k, f) = \int \psi_k(t)f(x)K(x,t)\,dtdx$$

Due to the support of f we may restrict the domain of integration to $|x_i|<|S_i|/2$ while due to the support of ψ_k we have $2^{k_i-1}|S_i|<|t_i|<2^{k_i+1}|S_i|$.

Using the mean zero of f in each x_i variable we rewrite the above integral as

$$\int \psi_k(t)f(x)(K(x,t)-K((x_1,0),t)-K((0,x_2),t)+K(0,t))\,dtdx$$

Since $2|x_i|<|S_i|<|t_i|$ we use the properties of product C-Z kernel to estimate the last display by

$$C\int |\psi_k(t)||f(x)|\frac{|x_1|^\delta}{|t_1|^{1+\delta}}\frac{|x_2|^\delta}{|t_2|^{1+\delta}}\,dtdx$$

and using again the restriction on the variables we estimate now by

$$C\int_{|t_i|<2^k|S_i|}|f(x)|\frac{1}{2^{k_1(1+\delta)}}\frac{1}{|S_1|}\frac{1}{2^{k_2(1+\delta)}}\frac{1}{|S_2|}\,dtdx\leq C2^{-k_1\delta}2^{-k_2\delta}\|f\|_{L^1(\mathbb{R}^2)}$$

This estimate proves that the sequence $(\Lambda(D_{2^{k_1}|S_1|,2^{k_2}|S_2|}\Phi,f))_{k>0}$ is Cauchy and so the existence of the limit $L(f)$.

Now the explicit rate of convergence stated in the lemma follows by summing geometric series: for every $k\in\mathbb{N}^2$, and every $0<\epsilon<2^{-(k_1+k_2)\delta}\|f\|_{L^1(\mathbb{R}^2)}$ let $m\in\mathbb{N}^2$ be with modulus big enough so that $2^{-(m_1+m_2)\delta}\|f\|_{L^1(\mathbb{R}^2)}$; then,

$$|L(f)-\Lambda(D_{2^{k_1}|S_1|,2^{k_2}|S_2|}\Phi,f)|\leq|L(f)-\Lambda(D_{2^{m_1}|S_1|,2^{m_2}|S_2|}\Phi,f)|+\sum_{k_1'=k_1}^{m_1}\sum_{k_2'=k_2}^{m_2}|\Lambda(\psi_{k'},f)|$$

$$\leq\epsilon+C\sum_{k_1'=k_1}^{m_1}\sum_{k_2'=k_2}^{m_2}2^{-(k_1'+k_2')\delta}\|f\|_{L^1(\mathbb{R}^2)}\leq C2^{-(k_1+k_2)\delta}\|f\|_{L^1(\mathbb{R}^2)}$$

and the proof is finished.

It can be easily proved that the definition of $T(1)$ is independent of the translation selected proving that L is invariant under scaling and translation. Moreover, it can also be shown that the definition is independent of the chosen cutoff function Φ.

We notice that, since we have only worked with smooth atoms, strictly speaking we haven't finished the definition of $T(1)$. To do it rigorously, we should prove that the sequence $(T(D_{2^{k_1},2^{k_2}}\Phi))_{k\in\mathbb{Z}^2}$ is uniformly bounded in $\mathrm{BMO}(\mathbb{R}^2)$. Then, using that the unit ball of the dual of Banach space is weak*-compact, we can extract a subsequence of previous sequence which converges to $L(f)$ for functions f in $C^\infty(\mathbb{R}^2)$ with compact support. Finally, since these functions are dense in $H^1(\mathbb{R}^2)$, we can deduce that previous functional can properly been extended to all $H^1(\mathbb{R}^2)$ and that $T(1\otimes 1)$ is the unique limit in $\mathrm{BMO}(\mathbb{R}^2)$ of the previous sequence. We will not get into any further detail about this.

We move now to the definition of $\langle T(\phi_I\otimes 1),\varphi_I\otimes\cdot\rangle$ following the previous schedule. The condition $\langle T(\phi_I\otimes 1),\varphi_I\otimes\cdot\rangle\in\mathrm{BMO}(\mathbb{R})$ means the fulfillment of the following inequality

$$|\langle T(\phi_I \otimes 1), \varphi_I \otimes f\rangle| \leq C\|f\|_{H^1(\mathbb{R})}$$

for all f that belong to a dense subset of $H^1(\mathbb{R})$. In this case, such dense subset will be the family of Schwartz functions f compactly supported with mean zero. Then, in order to give a proper sense to the left hand side of previous inequality we use the following lemma:

Lemma 4.2. *Let $\Phi \in \mathcal{S}(\mathbb{R})$ such that $\Phi(x) = 1$ for $|x| \leq 1$ and $\Phi(x) = 0$ for $|x| \geq 2$. Let S be a rectangle and ϕ_{S_1}, φ_{S_1} be two L^2-normalized bump functions adapted to S_1. Let $f \in \mathcal{S}(\mathbb{R})$ be supported in S_2 with mean zero. Then, the limit*

$$L_{\phi_{S_1},\varphi_{S_1}}(f) = \lim_{k\to\infty} \Lambda(\phi_{S_1} \otimes T_{c(S_2)} D_{2^k|S_2|}\Phi, \varphi_{S_1} \otimes f)$$

exists. Moreover, we have the error bound

$$|L_{\phi_{S_1},\varphi_{S_1}}(f) - \Lambda(\phi_{S_1} \otimes T_{c(S)} D_{2^k|S_2|}\Phi, \varphi_{S_1} \otimes f)| \leq C 2^{-\delta k}\|f\|_{L^1(\mathbb{R})}$$

where δ is the parameter in the Calderon-Zygmund property of the kernel K and C depends only on Φ and Λ.

Proof. We mimic the proof of previous lemma and for simplicity of notation we assume that S_2 is centered at the origin. For $k \geq 1$, we set $\psi_k = D_{2^k|S_2|}\Phi - D_{2^{k-1}|S_2|}\Phi$. We will estimate $|\Lambda(\phi_{S_1} \otimes \psi_k, \varphi_{S_1} \otimes f)|$.

Since the supports of ψ_k and f are disjoint we use the kernel representation of the restricted operator $T^2_{t_2,x_2}$

$$\Lambda(\phi_{S_1} \otimes \psi_k, \psi_{S_1} \otimes \psi_{S_2}) = \int \psi_k(t_2) f(x_2) \langle T^2_{t_2,x_2}(\phi_{S_1}), \varphi_{S_1}\rangle \, dt_2 dx_2$$

Due to the supports of f and ψ_k we have $|x_2| < |S_2|/2$ and $2^{k-1}|S_2| < |t_2| < 2^{k+1}|S_2|$ respectively. Using the mean zero of f we write the above integral as

$$\int \psi_k(t_2) f(x_2) \langle (T^2_{t_2,x_2} - T^2_{t_2,0})(\phi_{S_1}), \varphi_{S_1}\rangle \, dt_2 dx_2$$

Since $2|x_2| < |S_2| < |t_2|$, by the mixed WB-CZ properties we can estimate the last expression by

$$\int |\psi_k(t_2)||f(x_2)| C \frac{|x_2|^\delta}{|t_2|^{1+\delta}} \, dt_2 dx_2$$

and finally due to the restriction on the variables we can estimate by

$$C \int_{|t_2| < 2^k|S_2|} |f(x_2)| \frac{1}{2^{k(1+\delta)}} \frac{1}{|S_2|} \, dt_2 dx_2 \leq C 2^{-k\delta}\|f\|_{L^1(\mathbb{R})}$$

As before this estimate is summable in k, which proves that the sequence $(\Lambda(\phi_{S_1} \otimes D_{2^k|S_2|}\Phi, \varphi_{S_1} \otimes f))_{k>0}$ is Cauchy and so, the existence of the limit $L(f)$. The explicit rate of convergence stated in the lemma follows again by summing a geometric series.

Notice that the functional $L_{\phi_{S_1},\varphi_{S_1}}$ may be indistinctly denoted by $\Lambda(\phi_{S_1}\otimes 1,\varphi_{S_1}\otimes\cdot)$, $\langle T(\phi_{S_1}\otimes 1),\varphi_{S_1}\otimes\cdot\rangle$, $\langle\langle T^1(1),\cdot\rangle\phi_{S_1},\phi_{S_1}\rangle$ or $\langle T^2(\phi_{S_1}),\varphi_{S_1}\rangle 1$ since

$$\begin{aligned}\Lambda(\phi_{S_1}\otimes 1,\varphi_{S_1}\otimes f) &= \langle T(\phi_{S_1}\otimes 1),\varphi_{S_1}\otimes f\rangle\\ &= \langle\langle T^1(1),f\rangle\phi_{S_1},\varphi_{S_1}\rangle\\ &= \langle\langle T^2(\phi_{S_1}),\varphi_{S_1}\rangle 1,f\rangle\end{aligned}$$

Notice that this way the condition $L_{\phi_{S_1},\varphi_{S_1}}\equiv 0$ turns into $\langle T^2(\phi_{S_1}),\varphi_{S_1}\rangle\equiv 0$ for all ϕ_{S_1},φ_{S_1} adapted to S_1. On the other hand, the condition $L_{\phi_{S_1},\varphi_{S_1}}\in \mathrm{BMO}(\mathbb{R})$ turns into $\langle T^2(\phi_{S_1}),\varphi_{S_1}\rangle 1\in\mathrm{BMO}(\mathbb{R})$ or $\langle T(\phi_{S_1}\otimes 1),\varphi_{S_1}\rangle_{x_1}(x_2)\in\mathrm{BMO}(\mathbb{R})$ for all ϕ_{S_1},φ_{S_1} adapted to S_1.

Finally we define $\langle T(1\otimes\phi_I),\cdot\otimes\psi_J\rangle$ when ϕ_I,ψ_J have disjoint support and ψ_J has mean zero. We follow similar schedule as before by mixing the two previous cases. The condition $\langle T(1\otimes\phi_I),\cdot\otimes\psi_J\rangle\in\mathrm{BMO}(\mathbb{R})$ means that

$$|\langle T(1\otimes\phi_I),f\otimes\psi_J\rangle|\leq C\|f\|_{H^1(\mathbb{R})}$$

for all f that belong to a dense subset of $H^1(\mathbb{R})$. Again, the dense subset will be the family of Schwartz functions f compactly supported with mean zero. Then, in order to give a proper sense to the left hand side of previous inequality we use the following lemma:

Lemma 4.3. *Let $\Phi\in\mathcal{S}(\mathbb{R})$ such that $\Phi(x)=1$ for $|x|\leq 1$ and $\Phi(x)=0$ for $|x|\geq 2$. Let ϕ_{R_2}, ψ_{S_2} be L^2-normalized bump functions adapted and supported to the dyadic intervals R_2, S_2 respectively, such that $|R_2|\geq|S_2|$, $|R_2|<\mathrm{diam}(R_2,S_2)$ and ψ_{S_2} has mean zero.*

Let $f\in\mathcal{S}(\mathbb{R})$ be supported in a dyadic interval S_1 with mean zero. Then, the limit

$$L(f)=\lim_{k\to\infty}\Lambda(T_{c(S_1)}D_{2^k|S_1|}\Phi\otimes\phi_{R_2},f\otimes\psi_{S_2})$$

exists. Moreover, we have the error bound

$$|L(f)-\Lambda(T_{c(S_1)}D_{2^k|S_1|}\Phi\otimes\phi_{R_2},f\otimes\psi_{S_2})|\leq C2^{-\delta k}\Big(\frac{|S_2|}{|R_2|}\Big)^{1/2+\delta}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-(1+\delta)}\|f\|_{L^1(\mathbb{R})}$$

where δ is the parameter in the Calderon-Zygmund property of the kernel K and C depends only on Φ and Λ.

Proof. Again for simplicity of notation we assume that S_1 is centered at the origin.

For $k\geq 1$, we set $\psi_k=D_{2^k|S_1|}\Phi-D_{2^{k-1}|S_1|}\Phi$. We will estimate $|\Lambda(\psi_k\otimes\phi_{R_2},f\otimes\psi_{S_2})|$.

Since the supports of ψ_k and f and the supports of ϕ_{R_2} and ψ_{S_2} and are respectively disjoint we use the kernel representation

$$\Lambda(\psi_k\otimes\phi_{R_2},f\otimes\psi_{S_2})=\int\psi_k(t_1)\phi_{R_2}(t_2)f(x_1)\psi_{S_2}(x_2)K(x,t)dtdx$$

Due to the supports of the functions ψ_k and f we may restrict the domain of integration to $2^{k-1}|S_1| < |t_1| < 2^{k+1}|S_1|$, $|x_1| < |S_1|/2$ while, by hypothesis, we have $|t_2 - c(S_2)| > \mathrm{diam}(R_2 \cup S_2)$, $|x_2 - c(S_2)| < |S_2|/2$.

Using the mean zero of f and ψ_{S_2} we write the above integral as

$$\int \psi_k(t_1)\phi_{R_2}(t_2)f(x_1)\psi_{S_2}(x_2)(K(x,t)-K((x_1,c(S_2)),t)-K((0,x_2),t)+K((0,c(S_2)),t)dtdx$$

Since $2|x_1| < |S_1| < |t_1|$ and $2|x_2 - c(S_2)| < |S_2| < |R_2| < \mathrm{diam}(R_2 \cup S_2) < |t_2 - c(S_2)|$, by the kernel properties we can estimate the last expression by

$$\int_{\substack{|x_1|<|S_1|/2\\ |t_1|<2^k|S_1|}} \int_{\substack{|x_2-c(S_2)|<|S_2|/2\\ |x_2-t_2|>\mathrm{diam}(R_2\cup S_2)}} |\psi_k(t_1)||\phi_{R_2}(t_2)||f(x_1)||\psi_{S_2}(x_2)|C\frac{|x_1|^\delta}{|t_1|^{1+\delta}}\frac{|x_2-c(S_2)|^\delta}{|t_2|^{1+\delta}}\,dtdx$$

which, due to the restriction on the variables, we can estimate by

$$C\int_{|t_1|<2^k|S_1|}|f(x_1)|\frac{1}{2^{k(1+\delta)}}\frac{1}{|S_1|}\,dt_1dx_1\frac{|S_2|^\delta}{\mathrm{diam}(R_2\cup S_2)^{1+\delta}}\int|\phi_{R_2}(t_2)||\psi_{S_2}(x_2)|dt_2dx_2$$

$$= C2^{-k\delta}\|f\|_{L^1(\mathbb{R})}\frac{|S_2|^\delta}{\mathrm{diam}(R_2\cup S_2)^{1+\delta}}\|\phi_{R_2}\|_{L^1(\mathbb{R})}\|\psi_{S_2}\|_{L^1(\mathbb{R})}$$

$$\leq C2^{-k\delta}\|f\|_{L^1(\mathbb{R})}\frac{|S_2|^\delta}{\mathrm{diam}(R_2\cup S_2)^{1+\delta}}|R_2|^{1/2}|S_2|^{1/2}$$

$$\leq C2^{-k\delta}\|f\|_{L^1(\mathbb{R})}\Big(\frac{|S_2|}{|R_2|}\Big)^{1/2+\delta}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-(1+\delta)}$$

As before this estimate is summable in k, which proves that the sequence $(\Lambda(D_{2^k|S_1|}\Phi\otimes \phi_{R_2}, f\otimes\psi_{S_2}))_{k>0}$ is Cauchy and so, the existence of the limit $L(f)$. The explicit rate of convergence stated in the Lemma follows again by summing a geometric series.

5. Λ applied to bump functions

In this section we study the action of Λ on bump functions to obtain good estimates of the dual pair in terms of the space and frequency (or scale) localization of the bump functions.

Before starting we state and prove two lemmata about localization properties of bump functions. Both results will be frequently used in proposition 5.3, the main result of this section. In particular, Lemma 5.1 will be used when we apply weak boundedness condition away from the origin while Lemma 5.2 will be mostly used when we need to use the cancellation condition $T(1) = 0$ and weak boundedness close to the origin.

Lemma 5.1. *Let I, J be two intervals such that $|I| \geq |J|$. Let $0 < \theta < 1$, $\lambda = (|J|^{-1}\mathrm{diam}(I\cup J))^\theta \geq 1$ and λJ the interval with same center as J and lenght $\lambda|J|$.*

Let $\Phi_{\lambda J}$ be the usual L^∞-normalized function adapted to λJ. Let ϕ_J be a L^2-normalized bump function adapted to J with constant C and order N.

Then, $\phi_J(1-\Phi_{\lambda J})$ is a L^2-normalized bump function adapted to I with constant

$$C\Big(\frac{|J|}{|I|}\Big)^{(\theta N-1)/2}(|I|^{-1}\mathrm{diam}(I\cup J))^{-\theta N/2}$$

and order $\theta N/4$.

Proof. We first study the decay of $\phi_J(x)(1-\Phi_{\lambda J})(x)$. Because of the support of $1-\Phi_{\lambda J}$ we have $|x-c(J)|\geq \lambda|J|$ and so

$$|J|^{-1}|x-c(J)|\geq \lambda=(|J|^{-1}\mathrm{diam}(I\cup J))^{\theta}>(|I|^{-1}\mathrm{diam}(I\cup J))^{\theta}$$

This implies

$$|I|^{-1}|x-c(I)|\leq |I|^{-1}(|x-c(J)|+|c(I)-c(J)|)$$
$$\leq |I|^{-1}|x-c(J)|+|I|^{-1}\mathrm{diam}(I\cup J)\leq |J|^{-1}|x-c(J)|+(|J|^{-1}|x-c(J)|)^{1/\theta}$$
$$\leq 2(|J|^{-1}|x-c(J)|)^{1/\theta}$$

where the last inequality holds because $|J|^{-1}|x-c(J)|\geq\lambda\geq 1$ and $\theta\leq 1$. Then,

$$1+|I|^{-1}|x-c(I)|\leq 1+2(|J|^{-1}|x-c(J)|)^{1/\theta}$$
$$\leq 2(1+(|J|^{-1}|x-c(J)|)^{1/\theta})\leq 2(1+|J|^{-1}|x-c(J)|)^{1/\theta}$$

With all this together with the inequalities $|x-c(J)|\geq\lambda|J|$ and $1+\lambda\geq 2$, we have

$$|\phi_J(x)(1-\Phi_{\lambda J}(x))|\leq C|J|^{-1/2}\frac{1}{(1+|J|^{-1}|x-c(J)|)^N}$$
$$\leq C|J|^{-1/2}\frac{1}{(1+\lambda)^{3N/4}}\frac{1}{(1+|J|^{-1}|x-c(J)|)^{N/4}}$$
$$\leq C\Big(\frac{|J|}{|I|}\Big)^{-1/2}\lambda^{-N/2}\frac{1}{(1+\lambda)^{N/4}}|I|^{-1/2}\frac{2^{\theta N/4}}{(1+|I|^{-1}|x-c(I)|)^{\theta N/4}}$$
$$\leq C\Big(\frac{|J|}{|I|}\Big)^{-1/2}(|J|^{-1}\mathrm{diam}(I\cup J))^{-\theta N/2}\frac{2^{\theta N/4}}{(1+\lambda)^{N/4}}|I|^{-1/2}\frac{1}{(1+|I|^{-1}|x-c(I)|)^{\theta N/4}}$$
$$\leq C\Big(\frac{|J|}{|I|}\Big)^{\theta N/2-1/2}(|I|^{-1}\mathrm{diam}(I\cup J))^{-\theta N/2}|I|^{-1/2}\frac{1}{(1+|I|^{-1}|x-c(I)|)^{\theta N/4}}$$

This proves the result for the decay of the bump function. The same ideas prove the corresponding decay for the derivatives of the bump function.

Lemma 5.2. *Let I,J be two dyadic intervals such that $|I|\geq|J|$ and J is centered at the origin. Let $\tilde{I}$ the interval centered at the origin with $|\tilde{I}|=|I|$. Let $\Phi_{\tilde{I}}$ be the usual L^∞-normalized function supported in $\tilde{I}$ and ϕ_I be a L^2-normalized bump function adapted to I with constant C and order N.*

Then, $\phi_I(0)\Phi_{\tilde{I}}$ is a L^∞-normalized bump function adapted to I with constant

$$C|I|^{-1/2}(|I|^{-1}\mathrm{diam}(I\cup J))^{-N/2}$$

and order $N/2$.

Proof. Since ϕ_I is L^2-adapted to I, we have by definition

$$|\Phi_{\tilde{I}}(x)\phi_I(0)| \leq C|I|^{-1/2}\frac{1}{(1+|I|^{-1}|c(I)|)^N}$$

Now, J is centered at the origin and so we have $|I|+|c(I)| = |I|+|c(I)-c(J)| \geq \mathrm{diam}(I\cup J)$. This implies $1+|I|^{-1}|c(I)| \geq |I|^{-1}\mathrm{diam}(I\cup J)$.

On the other hand, because of the support of $\Phi_{\tilde{I}}$ we have $|x| \leq |I|/2$. Then, $|x-c(I)| \leq |I|/2+|c(I)|$ and so $1+|I|^{-1}|x-c(I)| \leq 3/2+|I|^{-1}|c(I)| \leq 3/2(1+|I|^{-1}|c(I)|)$.

Finally, since I is dyadic, we have that $|c(I)| \geq |I|/2$ and so, also $1+|I|^{-1}|c(I)| > 3/2$.

With these three inequalities, we have

$$|\Phi_{\tilde{I}}(x)\phi_I(0)| \leq C|I|^{-1/2}\frac{1}{(1+|I|^{-1}|c(I)|)^{N/2}}\frac{1}{(1+|I|^{-1}|c(I)|)^{N/4}}\frac{1}{(1+|I|^{-1}|c(I)|)^{N/4}}$$

$$\leq C|I|^{-1/2}(|I|^{-1}\mathrm{diam}(I\cup J)^{-N/2}\frac{1}{(1+|I|^{-1}|c(I)|)^{N/4}}\frac{(3/2)^{N/4}}{(1+|I|^{-1}|x-c(I)|)^{N/4}}$$

$$\leq C|I|^{-1/2}(|I|^{-1}\mathrm{diam}(I\cup J)^{-N/2}\frac{1}{(1+|I|^{-1}|x-c(I)|)^{N/4}}$$

Again, the same ideas also prove the corresponding decay for the derivatives of the bump function.

Now we state and prove the technical lemma that describes the action of the operator when it satisfies the special cancellation properties.

Proposition 5.3. *(Bump lemma) Let K be a product Calderon-Zygmund kernel with parameter δ. Let Λ be a bilinear Calderon-Zygmund form with associated kernel K which satisfies the mixed WB-CZ conditions.*

Assume that Λ also satisfies the weak boundedness condition and the special cancellation conditions $\Lambda(1\otimes 1,\psi_S)=0$ for all $\psi\in\mathcal{S}(\mathbb{R}^2)$ with mean zero and $\Lambda(f_1\otimes 1, g_1\otimes\psi) = \Lambda(1\otimes f_2,\psi\otimes g_2)=0$ for all $f_i,g_i,\psi\in\mathcal{S}(\mathbb{R})$ with ψ of mean zero.

Let R, S be rectangles such that $|R_i|\geq|S_i|$ for $i=1,2$. Let ϕ_1 be a bump function L^2-adapted to R and ψ_2 a bump function L^2-adapted to S with mean zero.

Then, for any $0<\delta'<\delta$

$$|\Lambda(\phi_1,\psi_2)| \leq C_{\delta'}\left(\frac{|S|}{|R|}\right)^{1/2+\delta'}\prod_{i=1}^{2}\left(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i)\right)^{-(1+\delta')}$$

Notice that with some abuse of notation, whenever we use this estimate we will simply write δ instead of δ'.

By symmetry on the arguments one can prove the following

Corollary 5.4. *Let Λ a bilinear form that satisfies all the requested previous properties and the following special cancellation conditions:*

$$\Lambda(1\otimes 1,\psi)=\Lambda(\psi,1\otimes 1)=\Lambda(\psi_1\otimes 1,1\otimes\psi_2)=\Lambda(1\otimes\psi_2,\psi_1\otimes 1)=0$$

for all $\psi \in \mathcal{S}(\mathbb{R}^2)$, $\psi_i \in \mathcal{S}(\mathbb{R})$ *with mean zero; and*

$$\Lambda(f_1 \otimes 1, g_1 \otimes \psi) = \Lambda(1 \otimes f_2, \psi \otimes g_2) = \Lambda(\psi \otimes f_2, 1 \otimes g_2) = \Lambda(f_1 \otimes \psi, g_1 \otimes 1) = 0$$

for all smooth functions $f_i, g_i, \psi \in \mathcal{S}(\mathbb{R})$ *with* ψ *of mean zero.*

Let R, S *be rectangles and let* ψ_1, ψ_2 *be bump functions* L^2*-adapted to* R *and* S *respectively with mean zero. Then, for any* $0 < \delta' < \delta$,

(5.1)
$$|\Lambda(\psi_1, \psi_2)| \leq C \prod_{i=1}^{2} \left(\frac{\min(|R_i|, |S_i|)}{\max(|R_i|, |S_i|)} \right)^{1/2+\delta'} \left(\max(|R_i|, |S_i|)^{-1} \mathrm{diam}(R_i \cup S_i)\right)^{-(1+\delta')}$$

Proof of Proposition 5.3. For simplicity of notation we shall assume that S_i are both centered at the origin. For each rectangle R and $\lambda \in \mathbb{R}^2$, we denote by λR the dilated rectangle $(\lambda_1 R_1) \times (\lambda_2 R_2)$ that shares the same centre as R and has measure $|\lambda||R|$.

Let Φ_{R_i} be the usual L^∞-normalized function adapted to the interval R_i and let $\Phi_R = \Phi_{R_1} \otimes \Phi_{R_2}$.

We denote $\psi(t, x) = \phi_1(t)\psi_2(x)$ and truncate the function as follows.

We start by splitting ψ in the x_i variables iteratively, first in x_1 and later in x_2. Let $\lambda_i = (|S_i|^{-1} \mathrm{diam}(R_i \cup S_i))^\epsilon$, with $0 < \epsilon < 1$. Then, $\psi = \psi_{in} + \psi_{out}$ where

$$\psi_{in}(t, x) = (\psi(t, x) - c_1(t, x_2))\Phi_{\lambda_1 S_1}(x_1)$$

and

$$\psi_{out}(t, x) = \psi(t, x)(1 - \Phi_{\lambda_1 S_1}(x_1)) + c_1(t, x_2)\Phi_{\lambda_1 S_1}(x_1)$$

with $c_1(t, x_2)$ chosen so that both ψ_{in} and ψ_{out} have mean zero in the variable x_1. Notice that both ψ_{in} and ψ_{out} have mean zero in the variable x_2. Now, $\psi_{in} = \psi_{in,in} + \psi_{in,out}$, where

$$\psi_{in,in}(t, x) = (\psi_{in}(t, x) - c_2(t, x_1))\Phi_{\lambda_2 S_2}(x_2) \tag{5.2}$$

and

$$\psi_{in,out}(t, x) = \psi_{in}(t, x)(1 - \Phi_{\lambda_2 S_2}(x_2)) + c_2(t, x_1)\Phi_{\lambda_2 S_2}(x_2)$$

with $c_2(t, x_1)$ chosen so that both $\psi_{in,in}$ and $\psi_{in,out}$ have mean zero in the x_2 variable. Meanwhile $\psi_{out} = \psi_{out,in} + \psi_{out,out}$, where

$$\psi_{out,in}(t, x) = (\psi_{out}(t, x) - c_3(t, x_1))\Phi_{\lambda_2 S_2}(x_2) \tag{5.3}$$

and

$$\psi_{out,out}(t, x) = \psi_{out}(t, x)(1 - \Phi_{\lambda_2 S_2}(x_2)) + c_3(t, x_1)\Phi_{\lambda_2 S_2}(x_2)$$

where $c_3(t, x_1)$ chosen so that both $\psi_{out,in}$ and $\psi_{out,out}$ have mean zero in the x_2 variable. Notice that for example

$$c_2(t, x_1) = -c|S_2|^{-1} \int \psi_{in}(t, x)(1 - \Phi_{\lambda_2 S_2}(x_2))dx_2 \tag{5.4}$$

We see now that the four functions $\psi_{out,in}$, $\psi_{out,out}$,... have mean zero in each variable x_i. This is obvious in the variable x_2, since c_2 and c_3 have been chosen to accomplish this. Moreover, we know that ψ_{in} and ψ_{out} have mean zero in the variable x_1. Because of this, we have for each x_2, t:

$$\int \psi_{in,in}(t,x)dx_1\Phi_{\lambda_2 S_2}(x_2) = -\int c_2(t,x)dx_1\Phi_{\lambda_2 S_2}(x_2) \tag{5.5}$$

$$= -c|S_2|^{-1}\int\int \psi_{in}(t,x)dx_1(1-\Phi_{\lambda_2 S_2}(x_2))\Phi_{\lambda_2 S_2}(x_2)dx_2 = 0$$

An analogous argument also proves mean zero of $\psi_{out,out}$ in each variable x_i. Mean zero of both $\psi_{in,in}$, $\psi_{out,out}$ imply the same for $\psi_{in,out}$ and $\psi_{out,in}$.

Now we split the four functions in the t_i variables. For ψ_{out} we only decompose the first two terms to get $\psi_{out,in}(t,x) = \psi_{1,2}(t,x) + \psi_{1,3}(t,x)$, $\psi_{in,out}(t,x) = \psi_{2,1}(t,x) + \psi_{3,1}(t,x)$ and $\psi_{out,out}(t,x) = \psi_{1,1}(t,x)$ where

$$\psi_{1,2}(t,x) = \psi_{out,in}(t,x)\Phi_{\mu_2\tilde{R}_2}(t_2), \qquad \psi_{1,3}(t,x) = \psi_{out,in}(t,x)(1-\Phi_{\mu_2\tilde{R}_2}(t_2))$$

$$\psi_{2,1}(t,x) = \psi_{in,out}(t,x)\Phi_{\mu_1\tilde{R}_1}(t_1), \qquad \psi_{3,1}(t,x) = \psi_{in,out}(t,x)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))$$

$$\psi_{1,1}(t,x) = \psi_{out,out}(t,x)$$

with $\tilde{R}_i$ the translate of R_i centered at the origin and $\mu_i = |R_i|^{-1}\text{diam}(R_i \cup S_i)$. The reason for notation $\psi_{i,j}$ will become clear later.

Finally, for $\psi_{in,in}$ we repeat the first type of decomposition to get the following four terms:

$$\begin{aligned}
\psi_{in}(t,x) &= \psi_{in,in}(x,t)\Phi_{\mu\tilde{R}}(t) \\
&+ \psi_{in,in}(x,t)\Phi_{\mu_1\tilde{R}_1}(t_1)(1-\Phi_{\mu_2\tilde{R}_2}(t_2)) \\
&+ \psi_{in,in}(x,t)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))\Phi_{\mu_2\tilde{R}_2}(t_2) \\
&+ \psi_{in,in}(x,t)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))(1-\Phi_{\mu_2\tilde{R}_2}(t_2)) \\
&= \psi_{2,2}(t,x) + \psi_{2,3}(t,x) + \psi_{3,2}(t,x) + \psi_{3,3}(t,x)
\end{aligned}$$

A carefull look at all these terms reveals that they can be described by

$$\begin{aligned}
\psi_{1,2}(t,x) &= \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))\Phi_{\lambda_2 S_2}(x_2)\Phi_{\mu_2\tilde{R}_2}(t_2) + c_1^1 + c_3^1 \\
\psi_{1,3}(t,x) &= \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))\Phi_{\lambda_2 S_2}(x_2)(1-\Phi_{\mu_2\tilde{R}_2}(t_2)) + c_1^2 + c_3^2 \\
\psi_{2,1}(t,x) &= \psi(t,x)\Phi_{\lambda_1 S_1}(x_1)(1-\Phi_{\lambda_2 S_2}(x_2))\Phi_{\mu_1\tilde{R}_1}(t_1) + c_1^3 + c_2^3 \\
\psi_{3,1}(t,x) &= \psi(t,x)\Phi_{\lambda_1 S_1}(x_1)(1-\Phi_{\lambda_2 S_2}(x_2))(1-\Phi_{\mu_1\tilde{R}_1}(t_1)) + c_1^4 + c_2^4 \\
\psi_{1,1}(t,x) &= \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))(1-\Phi_{\lambda_2 S_2}(x_2)) + c_1^5 + c_3^5 \\
\psi_{2,2}(t,x) &= \psi(t,x)\Phi_{\lambda S}(x)\Phi_{\mu\tilde{R}}(t) + c_1^6 + c_2^6 \\
\psi_{2,3}(t,x) &= \psi(t,x)\Phi_{\lambda S}(x)\Phi_{\mu_1\tilde{R}_1}(t_1)(1-\Phi_{\mu_2\tilde{R}_2}(t_2)) + c_1^7 + c_2^7 \\
\psi_{3,2}(t,x) &= \psi(t,x)\Phi_{\lambda S}(x)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))\Phi_{\mu_2\tilde{R}_2}(t_2) + c_1^8 + c_2^8 \\
\psi_{3,3}(t,x) &= \psi(t,x)\Phi_{\lambda S}(x)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))(1-\Phi_{\mu_2\tilde{R}_2}(t_2)) + c_1^9 + c_3^9
\end{aligned}$$

where the functions c_j^i are error terms that ensure that all functions $\psi_{i,j}$ have mean zero in the variables x_1, x_2. We notice that $c_1^i = c_1^i(t, x_2)$, $c_2^i = c_2^i(t, x_1)$ and $c_3^i = c_3^i(t, x_2)$. At the end we will prove that the functions c_j^i are small and have the right support to allow us to assume that the main terms have the stated zero averages. We will call denote the main terms again by $\psi_{i,j}$. Also notice that they are of tensor product type. Moreover, with a small abuse of notation, we will write the action of the dual pair over $\psi_{i,j}$ as $\Lambda(\psi_{i,j})$.

We call (1) the use of *weak boundedness condition* away from the origin, the mean zero in the variables x_1, x_2 and rate of decay of ψ; (2) the use of *the special cancellation condition* $T^i(1) \equiv 0$, the weak boundedness condition close to the origin and the mean zero of ψ in the x_i variable; and (3) the use of *the integral representation, the properties of the Calderón-Zygmund kernel* and the mean zero of ψ in the variable x_i. We call $(i) \times (j)$ the combined use of (i) in the variables t_1, x_1 and (j) in the variables t_2, x_2. Then, we plan to bound $\Lambda(\psi)$ dealing each term $\Lambda(\psi_{i,j})$ by means of $(i) \times (j)$.

a) We start with $\psi_{1,1}(t,x) = \psi(t,x)(1 - \Phi_{\lambda_1 S_1}(x_1))(1 - \Phi_{\lambda_2 S_2}(x_2))$ with mean zero in variables x_1, x_2, for which we will prove the decay by using weak boundedness in these variables.

We know that ψ is adapted to $R \times S$ and so, by lemma 5.1, $\psi_{1,1}$ is adapted to $R \times R$ with a gain in the constant of at least

$$C \left(\frac{|S|}{|R|} \right)^{\epsilon N} \prod_{i=1,2} \left(|R_i|^{-1} \mathrm{diam}(R_i \cup S_i) \right)^{-\epsilon N}$$

Then, by the weak boundedness condition we have

$$|\Lambda(\psi_{1,1})| \leq C \left(\frac{|S|}{|R|} \right)^{\epsilon N} \prod_{i=1,2} \left(|R_i|^{-1} \mathrm{diam}(R_i \cup S_i) \right)^{-\epsilon N}$$

b) To bound, $\psi_{2,2}(t,x) = \psi(t,x)\Phi_{\lambda S}(x)\Phi_{\mu \tilde{R}}(t)$ we first argue the fact that we are allowed to make the extra assumption $\psi_{2,2}(0,x) = 0$ for any x.

The assumption comes from the subtitution by

$$\tilde{\psi}_{2,2}(t,x) = \psi_{2,2}(t,x) - D_{|R_1|,|R_2|}\Phi(t)\psi_{2,2}(0,x)$$

and we just need to prove that the substracted term satisfies the same bounds we want to prove. By lemma 5.2, the function $D_{|R_1|,|R_2|}\Phi(t)\psi_{2,2}(0,x)$ is $L^\infty \times L^2$-adapted to $R \times S$ just like ψ, with a gain of constant of at least

$$C|R|^{-1/2} \prod_{i=1,2} (|R_i|^{-1} \mathrm{diam}(R_i \cup S_i))^{-N}$$

Then, using the special cancellation condition $\Lambda(1\otimes 1,\psi_{S_1}\otimes\psi_{S_2})=0$ and the explicit estimate of Lemma 4.1, we see that

$$|\Lambda(D_{|R_1|,|R_2|}\Phi(t),\psi_{2,2}(0,x))| = |\Lambda(D_{\frac{|R_1|}{|S_1|}|S_1|,\frac{|R_2|}{|S_2|}|S_2|}\Phi(t),\psi_{2,2}(0,x)) - \Lambda(1\otimes 1,\psi_{2,2}(0,x))|$$

$$\leq C|R|^{-1/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}\left(\frac{|R_1|}{|S_1|}\right)^{-\delta}\left(\frac{|R_2|}{|S_2|}\right)^{-\delta}\|\psi_{2,2}(0,\cdot)\|_1$$

$$\leq C|R|^{-1/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}\left(\frac{|S|}{|R|}\right)^{\delta}|S|^{1/2}$$

$$= C\left(\frac{|S|}{|R|}\right)^{\delta+1/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}$$

and the right hand side is no larger than the desired bound.

Now with this assumption, we can prove $\psi_{2,2}$ is adapted to $S\times S$ with constant $C|R|^{-3/2}|S|\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i,S_i))^{-N}$. In order to do so we prove that, in $S\times S$, $\psi_{2,2}$ is bounded by this quantity, while the work for the derivatives can be done in a similar way. By the extra assumption,

$$|\psi_{2,2}(t,x)| = \Big|\int_0^{t_1}\int_0^{t_2}\partial_{t_1}\partial_{t_2}\psi_{2,2}(t,x)\,dt\Big|$$

$$\leq C|t_1||t_2|\|\partial_{t_1}\partial_{t_2}\psi_{2,2}(\cdot,x)\|_\infty \leq C|S|\|\partial_{t_1}\partial_{t_2}\psi_{2,2}(\cdot,x)\|_\infty$$

and by the definition of a bump function

$$\|\partial_{t_1}\partial_{t_2}\psi_{2,2}(\cdot,x)\|_\infty \leq C|R|^{-3/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}D_{|S_1|,|S_2|}\Phi(t)\phi(x)$$

where ϕ is an L^2-normalized bump function adapted to S. This shows the bound

$$|\psi_{2,2}| \leq C|S||R|^{-3/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}D_{|S_1|,|S_2|}\Phi(t)\phi(x)$$

Finally then, by the weak boundedness property of Λ we get

$$|\Lambda(\psi_{2,2})| \leq C|S||R|^{-3/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}|S|^{1/2}$$

$$= C\left(\frac{|S|}{|R|}\right)^{3/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}$$

c) Now we consider $\psi_{3,3}(t,x)=\psi(t,x)\Phi_{\lambda S}(x)(1-\Phi_{\mu_1\tilde{R}_1}(t_1))(1-\Phi_{\mu_2\tilde{R}_2}(t_2))$ which will be bounded by the integral representation and the properties of the CZ kernel.

On the support of the $\psi_{3,3}$, we have that $|t_i|>C^{-1}\mu_i|R_i|=C^{-1}\mathrm{diam}(R_i\cup S_i)$ while $|x_i|\leq C\lambda_i|S_i|=C|S_i|^{1-\epsilon}\mathrm{diam}(R_i\cup S_i)^{\epsilon}$. Since $|S_i|^{-1}\mathrm{diam}(R_i\cup S_i)\geq 1$, the previous two inequalities imply $|x_i| < |t_i|$ and so, the support of $\psi_{3,3}$ is disjoint with the diagonal. This allow us to use the Calderón-Zygmund kernel representation

$$\Lambda(\psi_{3,3}) = \int \psi_{3,3}(t,x)K(x,t)\,dtdx$$

Now using the mean zero of $\psi_{3,3}$ in the variable x, the above integral equals

$$\int \psi_R(t)\psi_S(x)(K(x,t) - K((x_1,0),t) - K((0,x_2),t) + K(0,t))\,dtdx$$

Moreover, since $2|x_i| < |t_i|$ we can use the property of a product CZ kernel and estimate the last expression by

$$C\int |\psi_{3,3}(t,x)|\frac{|x_1|^\delta}{|t_1|^{1+\delta}}\frac{|x_2|^\delta}{|t_2|^{1+\delta}}\,dtdx$$

Finally, since $|t_i| > C^{-1}\text{diam}(R_i,S_i)$ and $|x_i| < C|S_i|^{1-\epsilon}\text{diam}(R_i,S_i)^\epsilon$, we estimate by

$$C\prod_{i=1,2}|S_i|^{(1-\epsilon)\delta}\text{diam}(R_i,S_i)^{\epsilon\delta}\text{diam}(R_i,S_i)^{-1-\delta}\|\psi_{3,3}\|_1$$

$$\leq C\prod_{i=1,2}|S_i|^{(1-\epsilon)\delta}\text{diam}(R_i,S_i)^{-1-\delta+\epsilon\delta}|R_i|^{1/2}|S_i|^{1/2}$$

$$= C\prod_{i=1,2}(|S_i|/|R_i|)^{1/2+(1-\epsilon)\delta}(\text{diam}(R_i\cup S_i)/|R_i|)^{-1-(1-\epsilon)\delta}$$

This proves the desired estimate for $\Lambda(\psi_{3,3})$.

d) Once finished the three "pure" cases, we move to the proof of the "mixed" ones. We start with $\psi_{1,2}(t,x) = \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))\Phi_{\lambda_2 S_2}(x_2)\Phi_{\mu_2\tilde{R}_2}(t_2)$ which will be bounded by the use of weak boundedness in t_1, x_1 and the special cancellation properties in t_2, x_2.

First, we impose the extra assumption $\psi_{1,2}((t_1,0),x) = 0$ for any t_1, x. The assumption is possible by the substitution

$$\tilde{\psi}_{1,2}(t,x) = \psi_{1,2}(t,x) - D_{|R_2|}\Phi(t_2)\psi_{1,2}((t_1,0),x)$$

and we first need to show that the substracted term also satisfies the stated bounds.

Due to the constant t_2 variable and the decay of ψ_1 away from S_1, we can apply Lemma 5.2 in the variable t_2 and Lemma 5.1 in the variable x_1, to deduce that the function $D_{|R_2|}\Phi(t_2)\psi_{1,2}((t_1,0),x)$ is adapted to $(R_1\times R_2)\times(R_1\times S_2)$ with a double gain of constant of at least

$$|R_2|^{-1/2}(|R_2|^{-1}\text{diam}(R_2,S_2))^{-N}\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N}(|R_1|^{-1}\text{diam}(R_1,S_1))^{-\epsilon N}$$

Now, we make use of the special cancellation condition $\Lambda(\phi_{R_1}\otimes 1, \varphi_{R_1}\otimes\psi_{S_2}) = 0$ for all bump functions $\phi_{R_1}, \varphi_{R_1}$ and all bump functions ψ_{S_2} of mean zero. Since $\psi_{1,2}((t_1,0),x)$ is adapted to $R_1\times(R_1\times S_2)$ and $\psi_{1,2}$ is of tensor product type we can write $\psi_{1,2}((t_1,0),x) = \phi_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{S_2}(x_2)$ and then use the explicit estimate of Lemma 4.2 to get

$$|\Lambda(\phi_{R_1}(t_1)\Phi_{|R_2|}(t_2), \psi_{R_1}(x_1)\psi_{S_2}(x_2))|$$
$$= |\Lambda(\phi_{R_1}(t_1)D_{\frac{|R_2|}{|S_2|}|S_2|}\Phi(t_2), \psi_{R_1}(x_1)\psi_{S_2}(x_2)) - \Lambda(\phi_{R_1}\otimes 1, \psi_{R_1}\otimes\psi_{S_2})|$$
$$\leq C|R_2|^{-1/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\left(\frac{|R_2|}{|S_2|}\right)^{-\delta}\|\psi_{S_2}\|_1$$
$$\leq C|R_2|^{-1/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\left(\frac{|S_2|}{|R_2|}\right)^{\delta}|S_2|^{1/2}$$
$$= C\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\left(\frac{|S_2|}{|R_2|}\right)^{\delta+1/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}$$
$$\leq C\left(\frac{|S|}{|R|}\right)^{\delta+1/2}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-N}$$

which is no larger than the desired bound.

This proves that we can make the assumption $\psi_{1,2}((t_1,0),x) = 0$. Now we prove that $\psi_{1,2}$ is adapted to $(R_1\times S_2)\times(R_1\times S_2)$ with constant

$$C\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}\left(|R_1|^{-1}\mathrm{diam}(R_1,S_1)\right)^{-\epsilon N}|R_2|^{-3/2}|S_2|(|R_2|^{-1}\mathrm{diam}(R_2,S_2))^{-N} \tag{5.6}$$

We only prove that, in $(R_1\times S_2)\times(R_1\times S_2)$, the function $|\psi_{1,2}(t,x)|$ is bounded appropriately. As before, with the extra assumption we have

$$|\psi_{1,2}(t,x)| = \left|\int_0^{t_2}\partial_{t_2}\psi_{1,2}(t,x)\,dt\right| \leq C|t_2|\|\partial_{t_2}\psi_{1,2}(t_1,\cdot,x)\|_\infty \leq C|S_2|\|\partial_{t_2}\psi_{1,2}(t_1,\cdot,x)\|_\infty$$

and by definition of a bump function

$$\|\partial_{t_2}\psi_{1,2}(t_1,\cdot,x)\|_\infty \leq C|R_2|^{-3/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}D_{|S_2|}\Phi(t_2)\phi(t_1,x)$$

where, ϕ is a L^2-normalized bump function adapted to $R_1\times(S_1\times S_2)$. This implies that $\psi_{1,2}$ can be bounded by

$$C|S_2||R_2|^{-3/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}D_{|S_2|}\Phi(t_2)\phi(t_1,x)$$

Moreover, by the factor $1-\Phi_{\lambda_1 S_1}$ of $\psi_{1,2}$ and Lemma 5.1, we can assume ϕ to be adapted to $R_1\times(R_1\times S_2)$ with a gain of constant of at most $C(|S_1|/|R_1|)^{\epsilon N}\left(|R_1|^{-1}\mathrm{diam}(R_1,S_1)\right)^{-\epsilon N}$. proving the bound (5.6). Similar work for the derivatives shows the claim that $\psi_{1,2}$ is adapted to $(R_1\times S_2)\times(R_1\times S_2)$ with the constant stated in (5.6).

Then, by the weak boundedness property of Λ we finally get

$$|\Lambda(\psi_{1,2})| \leq C\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}\left(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1)\right)^{-\epsilon N}\frac{|S_2|}{|R_2|^{3/2}}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}|S_2|^{1/2}$$
$$= C\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}\left(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1)\right)^{-\epsilon N}\left(\frac{|S_2|}{|R_2|}\right)^{3/2}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-N}$$

Notice that the factor $|S_2|^{1/2}$ at the end of first inequality comes from the L^∞ normalization of $D_{|S_2|}\Phi$.

e) We now consider $\psi_{1,3}(t,x) = \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))\Phi_{\lambda_2 S_2}(x_2)(1-\Phi_{\mu_2 \tilde{R}_2}(t_2))$ which will be work out by the use of weak boundedness and the kernel representation.

By lemma 5.1 in the variable x_1, we have that $\psi_{1,3}$ is adapted to $(R_1\times R_2)\times(R_1\times S_2)$ with a gain in the constant of at least

$$C\left(\frac{|S_1|}{|R_1|}\right)^{\epsilon N}\left(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1)\right)^{-\epsilon N}$$

This way we can assume $\psi_{1,3}(t,x) = \phi_{R_1}(t_1)\phi_{R_2}(t_2)\varphi_{R_1}(x_1)\psi_{S_2}(x_2)$ with ψ_{S_2} of zero mean. Moreover, on the support of ψ_2 we have that $|t_2| > C^{-1}\mathrm{diam}(S_2\cup R_2)$ while $|x_2| < C|S_2|^{1-\epsilon}\mathrm{diam}(R_2\cup S_2)^{\epsilon}$. This implies $|x_2| < |t_2|$ and so by the integral representation of the restricted operator $T^2_{t_2,x_2}$, we have

$$\Lambda(\psi_{1,3}) = \int \phi_{R_2}(t_2)\psi_{S_2}(x_2)\langle T^2_{t_2,x_2}(\phi_{R_1}),\varphi_{R_1}\rangle\, dt_2dx_2$$

Using the mean zero of ψ_{S_2} we obtain for the above integral

$$\int \phi_{R_2}(t_2)\psi_{S_2}(x_2)\langle (T^2_{t_2,x_2}-T^2_{t_2,0})(\phi_{R_1}),\varphi_{R_1}\rangle\, dt_2dx_2$$

Since $2|x_2| < |t_2|$, by the mixed WB-CZ property of Λ and the gain in the constant, we estimate the integral by

$$\int |\psi_{R_2}(t_2)||\psi_{S_2}(x_2)|C\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\frac{|x_2|^\delta}{|t_2|^{1+\delta}}\, dt_2dx_2 \tag{5.7}$$

which, using the restriction on the variables, we bound by

$$C\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}$$

$$|S_2|^{(1-\epsilon)\delta}\mathrm{diam}(R_2,S_2)^{\epsilon\delta}\mathrm{diam}(S_2\cup R_2)^{-(1+\delta)}\|\psi_{R_2}\psi_{S_2}\|_{L^1(\mathbb{R}^2)}$$

$$\leq C\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N-1/2}\prod_{i=1,2}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}$$

$$|S_2|^{(1-\epsilon)\delta}\mathrm{diam}(R_2\cup S_2)^{\epsilon\delta}\mathrm{diam}(S_2\cup R_2)^{-(1+\delta)}|R_2|^{1/2}|S_2|^{1/2}$$

$$= C\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\Big(\frac{|S_2|}{|R_2|}\Big)^{(1/2+(1-\epsilon)\delta)}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-(1+(1-\epsilon)\delta)}$$

f) The last term we work with is $\psi_{2,3}(t,x) = \psi(t,x)\Phi_{\lambda S}(x)\Phi_{\mu_1\tilde{R}_1}(t_1)(1-\Phi_{\mu_2\tilde{R}_2}(t_2))$ by using the special cancellation and the kernel representation.

As before, we first impose the extra assumption $\psi_{2,3}((0,t_2),x) = 0$ for any $t_2\in\mathbb{R}$, $x\in\mathbb{R}^2$. The assumption comes the substitution

$$\tilde{\psi}_{2,3}(t,x) = \psi_{2,3}(t,x) - D_{|R_1|}\Phi(t_1)\psi_{2,3}((0,t_2),x)$$

and we again need to prove that the substracted term satisfies the stated bounds.

We know that $\psi_{2,3}$ is adapted to $R\times S$. But due to the constant coordinate in the variable t_1 we have by Lemma 5.2 applied in the variable t_1 that the function $D_{|R_1|}\Phi(t_1)\psi_{2,3}((0,t_2),x)$ is also adapted to $R\times S$ with a gain of constant of at least

$$|R_1|^{-1/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}$$

Then, we can then write $D_{|R_1|}\Phi(t_1)\psi_{2,3}((0,t_2),x)=\phi_{S_1}(0)D_{|R_1|}\Phi(t_1)\phi_{R_2}(t_2)\psi_{S_1}(x_1)\psi_{S_2}(x_2)$ where ϕ_{R_1} is a bump function adapted to R_1 and ϕ_{R_2}, ψ_{S_2} are bump functions supported and adapted to R_2, S_2 respectively, such that ψ_{S_2} has mean zero and $R_2\cap S_2=\emptyset$.

This way, by the special cancellation condition $\Lambda(1\otimes\phi_{R_2},\psi_{S_1}\otimes\psi_{S_2})=0$, the explicit estimate of Lemma 4.3 and the gain in the constant, we have that

$$|\Lambda(\phi_{S_1}(0)D_{|R_1|}\Phi\otimes\phi_{R_2}\otimes,\psi_{S_1}\otimes\psi_{S_2})|$$

$$=|\Lambda(\phi_{S_1}(0)D_{\frac{|R_1|}{|S_1|}|S_1|}\Phi\otimes\phi_{R_2},\psi_{S_1}\otimes\psi_{S_2})-\Lambda(1\otimes\phi_{R_2},\psi_{S_1}\otimes\psi_{S_2})|$$

$$\leq C|R_1|^{-1/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}\left(\frac{|R_1|}{|S_1|}\right)^{-\delta}\left(\frac{|S_2|}{|R_2|}\right)^{1/2+\delta'}\Big(\frac{\text{diam}(R_2\cup S_2)}{\max(|R_2|,|S_2|)}\Big)^{1+\delta'}\|\psi_{S_1}\|_1$$

$$\leq C|R_1|^{-1/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}\left(\frac{|S_1|}{|R_1|}\right)^{\delta}\left(\frac{|S_2|}{|R_2|}\right)^{1/2+\delta'}\Big(\frac{\text{diam}(R_2\cup S_2)}{\max(|R_2|,|S_2|)}\Big)^{1+\delta'}|S_1|^{1/2}$$

$$=C\left(\frac{|S_1|}{|R_1|}\right)^{\delta+1/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}\Big(\frac{|S_2|}{|R_2|}\Big)^{(1/2+\delta')}(|R_2|^{-1}\text{diam}(R_2\cup S_2))^{-(1+\delta')}$$

which is no larger than the desired bound.

Now with the assumption $\psi_{2,3}((0,t_2),x)=0$, we can show that $\psi_{2,3}$ is adapted to $(S_1\times R_2)\times(S_1\times S_2)$ with constant $C|R_1|^{-3/2}|S|(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}$. We only prove that, $\psi_{2,3}$ satisfies the necessary bound in $(R_1\times S_2)\times(S_1\times S_2)$. This follows from

$$|\psi_{2,3}(t,x)|=\Big|\int_0^{t_1}\partial_{t_1}\psi_{2,3}(t,x)\,dt\Big|\leq C|t_1|\|\partial_{t_1}\psi_{2,3}((\cdot,t_2),x)\|_\infty\leq C|S_1|\|\partial_{t_1}\psi_{2,3}((\cdot,t_2),x)\|_\infty$$

and by the definition of a bump function

$$\|\partial_{t_1}\psi_{2,3}((\cdot,t_2),x)\|_\infty\leq C|R_1|^{-3/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}D_{|S_1|}\Phi(t_1)\phi(t_2,x)$$

where ϕ is an L^2-normalized bump function adapted to $R_2\times S$. This implies that

$$|\psi_{2,3}(t,x)|\leq C|S_1||R_1|^{-3/2}(|R_1|^{-1}\text{diam}(R_1\cup S_1))^{-N}D_{|S_1|}\Phi(t_1)\phi(t_2,x)$$

Analogous estimates work for the derivatives.

This way we can assume $\psi_{2,3}(t,x)=\phi_{S_1}(t_1)\phi_{R_2}(t_2)\psi_{S_1}(x_1)\psi_{S_2}(x_2)$ with bump functions adapted to the corresponding intervals and ψ_{S_i} of zero mean. Moreover, in the support of $\psi_{2,3}$ we have that $|t_2|>C^{-1}\text{diam}(S_2\cup R_2)$ while $|x_2|<C|S_2|^{1-\epsilon}\text{diam}(R_2\cup$

$S_2)^\epsilon$. This implies $|x_2| < |t_2|$ and so by the integral representation of the restricted operator $T^2_{t_2,x_2}$ we have

$$\Lambda(\psi_{2,3}) = \int \phi_{R_2}(t_2)\psi_{S_2}(x_2)\langle T^2_{t_2,x_2}(\phi_{S_1}), \psi_{S_1}\rangle \, dt_2 dx_2$$

Using the mean zero of ψ_{S_2} we obtain for the above integral

$$\int \phi_{R_2}(t_2)\psi_{S_2}(x_2)\langle (T^2_{t_2,x_2} - T^2_{t_2,0})(\phi_{S_1}), \psi_{S_1}\rangle \, dt_2 dx_2$$

Since $2|x_2| < |t_2|$, by the mixed WB-CZ property we can bound this by

$$\int |\psi_{R_2}(t_2)||\psi_{S_2}(x_2)|C|S_1||R_1|^{-3/2}(|R_1|^{-1}\mathrm{diam}(R_1 \cup S_1))^{-N}\frac{|x_2|^\delta}{|t_2|^{1+\delta}}\, dt_2 dx_2$$

and using the restriction on the variables we can estimate as we did from (5.7) by

$$C\Big(\frac{|S_1|}{|R_1|}\Big)^{3/2}(|R_1|^{-1}\mathrm{diam}(R_1 \cup S_1))^{-N}\Big(\frac{|S_2|}{|R_2|}\Big)^{(1/2+(1-\epsilon)\delta)}(|R_2|^{-1}\mathrm{diam}(R_2 \cup S_2))^{-(1+(1-\epsilon)\delta)}$$

We end the proof of Proposition 5.3 by dealing with the error terms. We only check that the factors $c_1^5(t,x_2) = c_1(t,x_2)\Phi_{\lambda_1 S_1}(x_1)(1-\Phi_{\lambda_2 S_2}(x_2))$ and $c_3^5(t,x_1) = c_3(t,x_1)\Phi_{\lambda_2 S_2}(x_2)$ are small enough being analogous all the other ones. In a similar way we obtained equality (5.4), we now have

$$c_3(t,x_1) = -c|S_2|^{-1}\int \psi_{out}(t,x)(1-\Phi_{\lambda_2 S_2}(x_2))dx_2$$

and so $\int c_3(t,x_1)dx_1 = 0$. Since also $\int \psi_{1,1}(t,x)dx_1 = 0$ we have

$$c_1(t,x_2)(1-\Phi_{\lambda_2 S_2}(x_2))\int \Phi_{\lambda_1 S_1}(x_1)dx_1 = -(1-\Phi_{\lambda_2 S_2}(x_2))\int \psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))dx_1$$

and thus

$$|c_1(t,x_2)|\lambda_1|S_1| \leq \int_{|x_1|>\lambda_1|S_1|}|\psi(t,x)|dx_1$$

$$\leq C|S_1|^{-1/2}\int_{|x_1|>\lambda_1|S_1|}(1+|S_1|^{-1}|x_1|)^{-N}dx_1 \;\phi(t,x_2)$$

$$\leq C{\lambda_1}^{-N}|S_1|^{1/2}\,\phi(t,x_2)$$

that is,

$$|c_1(t,x_2)| \leq C{\lambda_1}^{-N}|S_1|^{-1/2}\,\phi(t,x_2)$$

where ϕ is a L^2-normalized bump function adapted to $R \times S_2$. With this we have that

$$|c_1(t,x_2)\Phi_{\lambda_1 S_1}(x_1)| \leq C\lambda_1^{-N}\phi(t,x_2)|S_1|^{-1/2}\Phi_{\lambda_1 S_1}(x_1)$$

with $|S_1|^{-1/2}\Phi_{\lambda_1 S_1}$ a bump function L^2-adapted to S_1. By lemma 5.1, this shows that $|c_1(t,x_2)\Phi_{\lambda_1 S_1}(x_1)(1-\Phi_{\lambda_2 S_2}(x_2))|$ is adapted to $R\times R$ with a gain of constant of

$$C\lambda_1^{-N}\Big(\frac{|S_2|}{|R_2|}\Big)^{\epsilon N}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-\epsilon N}$$

and by the definition of λ_i this equals

$$C(|S_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}\Big(\frac{|S_2|}{|R_2|}\Big)^{\epsilon N}(|R_2|^{-1}\mathrm{diam}(R_2\cup S_2))^{-\epsilon N}$$

$$=C\Big(\frac{|S|}{|R|}\Big)^{\epsilon N}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-\epsilon N}$$

which is smaller than the required bounds.

Symmetrically we have that since also $\int\psi_{1,1}(t,x)dx_2=0$,

$$c_3(t,x_1)\int\Phi_{\lambda_2 S_2}(x_2)dx_2=-(1-\Phi_{\lambda_1 S_1}(x_1))\int\psi(t,x)(1-\Phi_{\lambda_2 S_2}(x_2))dx_2$$

$$-\Phi_{\lambda_1 S_1}(x_1)\int c_1(t,x_2)(1-\Phi_{\lambda_2 S_2}(x_2))dx_2$$

The first term in the right hand side can be treated in a similar way we did before and obtain

$$|(1-\Phi_{\lambda_1 S_1}(x_1))\int\psi(t,x)(1-\Phi_{\lambda_2 S_2}(x_2))dx_2|\leq C{\lambda_2}^{-N}|S_2|^{1/2}\,(1-\Phi_{\lambda_1 S_1}(x_1))\phi(t,x_1)$$

For the second term, we use the definition of $c_1(t,x_2)=-c|S_1|^{-1}\int\psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))dx_1$ to bound by

$$\Big|\frac{\Phi_{\lambda_1 S_1}(x_1)}{\lambda_1|S_1|}\int\int\psi(t,x)(1-\Phi_{\lambda_1 S_1}(x_1))dx_1(1-\Phi_{\lambda_2 S_2}(x_2))dx_2\Big|$$

$$\leq\frac{\Phi_{\lambda_1 S_1}(x_1)}{\lambda_1|S_1|}\int_{|x_i|>\lambda_i|S_i|}|\psi(t,x)|dx_1dx_2$$

$$\leq C\frac{\Phi_{\lambda_1 S_1}(x_1)}{\lambda_1|S_1|}|S_1|^{-1/2}|S_2|^{-1/2}\int_{|x_i|>\lambda_i|S_i|}(1+|S_2|^{-1}|x_2|)^{-N}(1+|S_2|^{-1}|x_2|)^{-N}dx_1dx_2\ \phi(t)$$

$$\leq C\frac{\Phi_{\lambda_1 S_1}(x_1)}{\lambda_1|S_1|}{\lambda_1}^{-N}{\lambda_2}^{-N}|S_1|^{1/2}|S_2|^{1/2}\,\phi(t)$$

where ϕ is a L^2-normalized bump function adapted to R.

Both things together imply

$$|c_3(t,x_1)|\lambda_2|S_2|\leq C{\lambda_2}^{-N}|S_2|^{1/2}\,(1-\Phi_{\lambda_1 S_1}(x_1))\phi(t,x_1)+C\frac{\Phi_{\lambda_1 S_1}(x_1)}{\lambda_1|S_1|}{\lambda_1}^{-N}{\lambda_2}^{-N}|S_1|^{1/2}|S_2|^{1/2}\,\phi(t)$$

and so

$$|c_3(t,x_1)|\leq C{\lambda_2}^{-N}|S_2|^{-1/2}\,\phi(t,x_1)+C{\lambda_1}^{-N}{\lambda_2}^{-N}|S_1|^{-1/2}|S_2|^{-1/2}\,\Phi_{\lambda_1 S_1}(x_1)\phi(t)$$

With this we have that

$$|c_3(t,x_1)\Phi_{\lambda_2 S_2}(x_2)| \le C\lambda_2^{-N}\phi(t,x_1)(1-\Phi_{\lambda_1 S_1}(x_1))|S_2|^{-1/2}\Phi_{\lambda_2 S_2}(x_2)$$
$$+C{\lambda_1}^{-N}{\lambda_2}^{-N}|S_1|^{-1/2}\Phi_{\lambda_1 S_1}(x_1)|S_2|^{-1/2}\Phi_{\lambda_2 S_2}(x_2)\phi(t)$$

with $|S_i|^{-1/2}\Phi_{\lambda_i S_i}$ a bump function L^2-adapted to S_i. This shows that the function is adapted to $R\times R$ with constant

$$C\lambda_2^{-N}\Big(\frac{|S_1|}{|R_1|}\Big)^{\epsilon N}(|R_1|^{-1}\mathrm{diam}(R_1\cup S_1))^{-\epsilon N}+C\lambda_1^{-N}\lambda_2^{-N}$$

where we have used Lemma 5.1 for the first term, which is analogous to the previous case. The second one is also all right since by definition of λ_i we have

$$C\lambda_1^{-N}\lambda_2^{-N} = C\prod_{i=1,2}(|S_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-\epsilon N}$$
$$= C\Big(\frac{|S|}{|R|}\Big)^{\epsilon N}\prod_{i=1,2}(|R_i|^{-1}\mathrm{diam}(R_i\cup S_i))^{-\epsilon N}$$

6. Proof of the main result

Theorem 6.1. *(L^2 boundedness). Let Λ be a bilinear Calderón-Zygmund form satisfying the mixed WB-CZ condition.*

We also assume that Λ satisfies the weak boundedness condition, and the special cancellation conditions

$$\Lambda(1,\psi_R\otimes\psi_S)=\Lambda(\psi_R\otimes\psi_S,1)=\Lambda(\psi_R\otimes 1,1\otimes\psi_S)=\Lambda(1\otimes\psi_R,\psi_S\otimes 1)=0$$

for all bump functions ψ_R,ψ_S adapted to intervals R, S with mean zero and

$$\Lambda(\phi_R\otimes 1,\varphi_R\otimes\psi_S)=\Lambda(1\otimes\phi_S,\psi_R\otimes\varphi_S)=\Lambda(\psi_R\otimes\phi_S,1\otimes\varphi_S)=\Lambda(\phi_R\otimes\psi_S,\varphi_R\otimes 1)=0$$

for all bump functions $\phi_R,\phi_S,\varphi_R,\varphi_S$ and all bump functions ψ_R,ψ_S with mean zero.

Then $\Lambda_0,\Lambda_1,\Lambda_2$ are bounded bilinear forms on L^2.

Proof. Because of the symmetry on the hypothesis it is clear that we only need to prove the result for Λ. We decompose the frequency plane in the standard way to obtain first a Littlewood-Paley decompositon and later a wavelet decomposition.

Let $\phi\in\mathcal{S}(\mathbb{R})$ be an even function such that $\widehat{\phi}$ is supported in $\{\xi\in\mathbb{R}:|\xi|\le 2\}$ and equals 1 on $\{\xi\in\mathbb{R}:|\xi|\le 1\}$. Let ψ be the function $\psi(x)=\phi(x)-\phi(x/2)$. Then $\hat\psi$ is supported on the annulus $\{\xi\in\mathbb{R}:2^{-1}\le|\xi|\le 2\}$ and moreover $\sum_{k\in\mathbb{Z}}\widehat{\psi}(\xi/2^k)\approx 1$, for all $\xi\neq 0$. We define the Littlewood-Paley projection operators in $\mathbb{R}$ given by $P_k(f)=f*D^1_{2^{-k}}\psi$ and $P_{\le k}(f)=f*D^1_{2^{-k}}\phi$. We observe that $\lim_{k\to\infty}P_{\le k}(f)=f$ while $\lim_{k\to\infty}P_{\le -k}(f)=0$ where in both cases the convergence is understood in the topology of $\mathcal{S}(\mathbb{R})$.

We consider now their counterparts in the biparameter case: for $k\in\mathbb{Z}^2$,

$$P_k(f)=f*(D^1_{2^{-k_1}}\psi\otimes D^1_{2^{-k_2}}\psi)\qquad P_{\le k_1,k_2}(f)=f*(D^1_{2^{-k_1}}\phi\otimes D^1_{2^{-k_2}}\psi)$$
$$P_{k_1,\le k_2}(f)=f*(D^1_{2^{-k_1}}\psi\otimes D^1_{2^{-k_2}}\phi)\qquad P_{\le k}(f)=f*(D^1_{2^{-k_1}}\phi\otimes D^1_{2^{-k_2}}\phi)$$

which satisfy $\lim_{k\to\infty} P_{\leq k}(f) = f$ in the topology of $\mathcal{S}(\mathbb{R}^2)$ while the other three operators tend to zero in the same sense.

For $N \in \mathbb{N}$, let Λ_N be the bilinear form given by

$$\Lambda_N(f,g) = \sum_{|k_i|,|j_i|\leq N} \Lambda(P_j f, P_k g)$$

where $k, j \in \mathbb{Z}^2$. We see that for all $f, g \in \mathcal{S}(\mathbb{R}^2)$ we have $\Lambda(f,g) = \lim_{N\to\infty} \Lambda_N(f,g)$: unfolding the sum in Λ_N, we have

$$\begin{aligned}\Lambda_N(f,g) &= \Lambda(P_{\leq(N,N)}f, P_{\leq(N,N)}g) - \Lambda(P_{\leq(-(N-1),-(N-1))}f, P_{\leq(N,N)}g) \\ &\quad -\Lambda(P_{\leq(N,N)}f, P_{\leq(-(N-1),-(N-1))}g) + \Lambda(P_{\leq-(N-1,N-1)}f, P_{\leq-(N-1,N-1)}g)\end{aligned}$$

and by the continuity of Λ we have that the first term tends to $\Lambda(f,g)$ while the other three tend to zero.

Let us now consider the family of intervals $\omega_{k_i} = [-2^{k_i+1}, -2^{k_i-1}] \cup [2^{k_i-1}, 2^{k_i+1}]$. Since $P_k(f)$ has Fourier support in $\omega_k = \omega_{k_1} \times \omega_{k_2}$, we have by Shannon's sampling theorem that

$$P_k(f) = \sum_R \langle f, \psi_{R,\omega_k}\rangle \psi_{R,\omega_k}$$

where the sum runs over all dyadic rectangles $R = R_1 \times R_2$ such that $|R_i| = |\omega_{k_i}|^{-1}$ and the convergence is understood in the topology of $\mathcal{S}(\mathbb{R}^2)$. Moreover, the functions ψ_{R,ω_k} satisfy that $\psi_{R,\omega_k} = \psi_{R_1,\omega_{k_1}} \otimes \psi_{R_2,\omega_{k_2}}$ where $\psi_{R_i,\omega_{k_i}}$ are Schwartz functions such that $\operatorname{supp} \widehat{\psi_{R_i,\omega_{k_i}}} \subset \omega_{k_i}$ and $e^{-2\pi i c(\omega_{k_i})}\psi_{R_i,\omega_{k_i}}$ are bump functions adapted to R_i. From now we drop the index ω_k in the notation of ψ_R.

Then by continuity of Λ in $\mathcal{S}(\mathbb{R}^2)$, we can write

$$\Lambda(f,g) = \sum_{k,j} \Lambda(P_j f, P_k g) = \sum_{R,S} \langle f, \psi_R\rangle\langle g, \psi_S\rangle \Lambda(\psi_R, \psi_S)$$

where now the sums run over the whole family of dyadic rectangles in $\mathbb{R}^2$. From now we work to obtain bounds of the last expression when the sum runs over finite families of dyadic rectangles in such way that the bounds are independent of the particular families of rectangles. Because of the rate of decay of Corollary 5.4, we parametrize the sums according to eccentricities and relative positions of the rectangles:

$$\sum_{R,S} \langle f, \psi_R\rangle\langle g, \psi_S\rangle \Lambda(\psi_R, \psi_S) = \sum_{i=1,2}\sum_{e_i\in\mathbb{Z}}\sum_{m_i\in\mathbb{N}}\sum_{R}\sum_{S\,\in\, R_{e,m}} \langle f, \psi_R\rangle\langle g, \psi_S\rangle \Lambda(\psi_R, \psi_S)$$

where for fixed eccentricities e_i, relative distances m_i and every given rectangle R, we define the family

$$R_{e,m} = \{S : |R_i| = 2^{e_i}|S_i|, m_i \leq \max(|R_i|, |S_i|)^{-1}\operatorname{diam}(R_i \cup S_i) < m_i + 1 \ \text{ for } i = 1, 2\}$$

Notice that by symmetry the product family $\{(R,S) : S \in R_{e,m}\}$ can be also parameterized as $\{(R,S) : R \in S_{-e,m}\}$ with analogous definition for $S_{-e,m}$.

We denote by $\sum_P$ the three first sums over parameters. By Lemma 5.3 and Cauchy's inequality, we bound the previous quantity by

$$\sum_P \sum_R \sum_{S\,\in\, R_{e,m}} |\Lambda(\psi_R, \psi_S)||\langle f, \psi_R\rangle||\langle g, \psi_S\rangle|$$

$$\lesssim \sum_P \sum_R \sum_{S\,\in\, R_{e,m}} 2^{-(|e_1|+|e_2|)(1/2+\delta)}(m_1 m_2)^{-(1+\delta)}|\langle f, \psi_R\rangle||\langle g, \psi_S\rangle|$$

$$\leq \sum_P 2^{-|e_1+e_2|(1/2+\delta)}(m_1 m_2)^{-(1+\delta)}\Big(\sum_R \sum_{S\in R_{e,m}} |\langle f, \psi_R\rangle|^2\Big)^{1/2}\Big(\sum_S \sum_{R\in S_{-e,m}} |\langle g, \psi_S\rangle|^2\Big)^{1/2}$$

Now, for every fixed R_i and each $m_i \in \mathbb{N}$ there are $2^{\max(e_i,0)}$ dyadic intervals S_i such that $|R_i| = 2^{e_i}|S_i|$ and $m_i \leq \max(|R_i|, |S_i|)^{-1}\mathrm{diam}(R_i \cup S_i) < m_i+1$. This implies that the cardinal of R_{e_i,m_i} is $2^{\max(e_1,0)}2^{\max(e_2,0)}$. For the same reason, the cardinal of $S_{-e,m}$ is $2^{\max(-e_1,0)}2^{\max(-e_2,0)} = 2^{-\min(e_1,0)}2^{-\min(e_2,0)}$. Then previous expression coincides with

$$\sum_P 2^{-|e_1+e_2|(1/2+\delta)}(m_1 m_2)^{-(1+\delta)} \Big(2^{\max(e_1,0)}2^{\max(e_2,0)}\sum_R |\langle f, \psi_R\rangle|^2\Big)^{1/2}$$
$$\Big(2^{-\min(e_1,0)}2^{-\min(e_2,0)}\sum_S |\langle g, \psi_S\rangle|^2\Big)^{1/2}$$

$$\leq \prod_{i=1,2} \sum_{e_i\in\mathbb{Z}} 2^{-|e_i|(1/2+\delta)}2^{\max(e_i,0)/2}2^{-\min(e_i,0)/2} \sum_{m_i\in\mathbb{N}} m_i^{-(1+\delta)}\|f\|_2\|g\|_2$$
$$= \Big(\sum_{e\in\mathbb{Z}} 2^{-|e|\delta}\sum_{m\in\mathbb{N}} m^{-(1+\delta)}\Big)^2\|f\|_2\|g\|_2$$

since $2^{\max(e_i,0)}2^{-\min(e_i,0)} = 2^{|e_i|}$.

7. Extension to L^p spaces

As said in the introduction, the weak L^1 estimates are no longer true in the multi-parameter case. So, in order to prove L^p bounds we cannot apply to our operator the classical method of interpolating between L^2 and the weak L^1 estimates. Instead, we follow the steps of the previous proof and perform again a decomposition of the dual pair which will be controlled by multi-parameter square functions whose L^p boundedness follows from weak L^1 bounds in the uni-parameter case.

Definition 7.1. *Given a $L^2(\mathbb{R}^2)$-normalized basis $(\psi_R)_R$, we define the double square function by*

$$SS(f) = \Big(\sum_R \frac{|\langle f, \psi_R\rangle|^2}{|R|}\chi_R\Big)^{1/2}$$

where the sum runs over all dyadic rectangles R in $\mathbb{R}^2$.

See [6] and specially [17] for a proof of boundedness of SS on $L^p(\mathbb{R}^2)$ with $1 < p < \infty$.

We also need to consider the following modified double square function

Definition 7.2. *Let $k \in \mathbb{Z}^2$, $n \in \mathbb{N}^2$. For every dyadic rectangle, R, we select a unique dyadic rectangle S such that $|R_i| = 2^{k_i}|S_i|$ and $n_i \leq \frac{\mathrm{diam}(R_i \cup S_i)}{\max(|R_i|,|S_i|)} < n_i + 1$. Then, with such a choice, we define*

$$SS_{k,n}(f)(x) = \Big(\sum_R \frac{\langle f, \psi_R\rangle^2}{|S|}\chi_S(x)\Big)^{1/2}$$

Obviously what we are defining here is a family of operators depending on the particular choice of the rectangles S. However, we will see that their bounds are independent of this particular choice. Moreover, we notice that the choice does not depend on the point x. The double square function corresponds to the values $k_i = 0$, $n_i = 1$.

We state in the proposition below boundedness of this modified square function. Its proof follows directly from the analogous result in the uni-parameter case and so, for the sake of completeness, at the end of the paper we include an appendix in which a proof of this result in the uni-parameter case can be found (see Proposition 9.3).

Proposition 7.3. *For every $1 < p < \infty$,*

$$\|SS_{k,n}(f)\|_{L^p(\mathbb{R}^2)} \leq C_p \prod_{i=1,2} (2^{-k_i \,\mathrm{sign}(\frac{2}{p}-1)} \log(n_i) + 1)^{|\frac{2}{p}-1|} \|f\|_{L^p(\mathbb{R}^2)}$$

Proof. Given $k \in \mathbb{Z}^2$ and $n \in \mathbb{N}^2$, let $TT_{k,n}$ be the operator defined by

$$TT_{k,n}(f)(x) = \sum_S \langle f, \psi_R\rangle \psi_S(x)$$

where the relationship between R and S is the same one given in the definition of the modified square function. Now we see that the double square function of $TT_{k,n}(f)$ coincides with $SS_{k,n}(f)$:

$$SS(TT_{k,n}(f))(x) = \Big(\sum_S \frac{\langle f, \psi_R\rangle^2}{|S|}\chi_S(x)\Big)^{1/2} = \Big(\sum_R \frac{\langle f, \psi_R\rangle^2}{|S|}\chi_S(x)\Big)^{1/2} = SS_{k,n}(f)(x)$$

and so

$$\|SS_{k,n}(f)\|_{L^p(\mathbb{R}^2)} = \|SS(TT_{k,n}(f))\|_{L^p(\mathbb{R}^2)} \approx \|TT_{k,n}(f)\|_{L^p(\mathbb{R}^2)}$$

Moreover, by linearity

$$TT_{k,n}(f)(x) = \sum_S \langle f, \psi_R\rangle \psi_S(x) = \sum_{S_1} \langle \sum_{S_2} \langle f, \psi_{R_2}\rangle \psi_{S_2}(x_2), \psi_{R_1}\rangle \psi_{S_1}(x_1)$$
$$= T_{k_1,n_1}(T_{k_2,n_2}(f)(\cdot, x_2))(x_1)$$

where $T_{k_i,n_i}(g)$ is defined in the obvious way

$$T_{k_i,n_i}(g)(x_i) = \sum_{R_i} \langle g, \psi_{R_i}\rangle \psi_{S_i}(x_i)$$

while $T_{k_2,n_2}(f)(x_1,x_2) = T_{k_2,n_2}(f_{x_1})(x_2)$ and $f_{y_1}(y_2) = f(y_1,y_2)$.

Proposition 9.3 gives us the boundedness result in the uni-parameter case,

$$\|T_{k_i,n_i}(f)\|_{L^p(\mathbb{R})} \leq C_p \|S_{k_i,n_i}(f)\|_{L^p(\mathbb{R})} \leq C_p(2^{-k_i s}n_i + 1)^{|\frac{2}{p}-1|}\|f\|_{L^p(\mathbb{R})}$$

where $s = \text{sign}(\frac{2}{p} - 1)$. Then, we have

$$\begin{aligned}
\|TT_{k,n}(f)\|_{L^p(\mathbb{R}^2)} &= \Big(\int_{\mathbb{R}} \|T_{k_1,n_1}(T_{k_2,n_2}(f)(\cdot,x_2))\|^p_{L^p(\mathbb{R})} dx_2\Big)^{1/p} \\
&\leq \Big(\int_{\mathbb{R}} C_p^p (2^{-k_1 s}n_1 + 1)^{|\frac{2}{p}-1|p} \|T_{k_2,n_2}(f)(\cdot,x_2)\|^p_{L^p(\mathbb{R})} dx_2\Big)^{1/p} \\
&= C_p(2^{-k_1 s}n_1 + 1)^{|\frac{2}{p}-1|}\Big(\int_{\mathbb{R}} \|T_{k_2,n_2}(f)(x_1,\cdot)\|^p_{L^p(\mathbb{R})} dx_1\Big)^{1/p} \\
&= C_p(2^{-k_1 s}n_1 + 1)^{|\frac{2}{p}-1|}\Big(\int_{\mathbb{R}} \|T_{k_2,n_2}(f_{x_1})\|^p_{L^p(\mathbb{R})} dx_1\Big)^{1/p} \\
&\leq C_p(2^{-k_1 s}n_1 + 1)^{|\frac{2}{p}-1|}\Big(\int_{\mathbb{R}} C_p^p (2^{-k_2 s}n_2 + 1)^{|\frac{2}{p}-1|p} \|f_{x_1}\|^p_{L^p(\mathbb{R})} dx_1\Big)^{1/p} \\
&= C_p(2^{-k_1 s}n_1 + 1)^{|\frac{2}{p}-1|}(2^{-k_2 s}n_2 + 1)^{|\frac{2}{p}-1|}\|f\|_{L^p(\mathbb{R}^2)}
\end{aligned}$$

Now we turn to the main result of this section.

Theorem 7.4. *(L^p boundedness). Let Λ be a bilinear Calderón-Zygmund form satisfying the mixed WB-CZ condition.*

We also assume that Λ satisfies the weak boundedness condition, and the special cancellation conditions

$$\Lambda(1, \psi_R \otimes \psi_S) = \Lambda(\psi_R \otimes \psi_S, 1) = \Lambda(\psi_R \otimes 1, 1 \otimes \psi_S) = \Lambda(1 \otimes \psi_R, \psi_S \otimes 1) = 0$$

for all bump functions ψ_R, ψ_S adapted to intervals R, S with mean zero and

$$\Lambda(\phi_R \otimes 1, \varphi_R \otimes \psi_S) = \Lambda(1 \otimes \phi_S, \psi_R \otimes \varphi_S) = \Lambda(\psi_R \otimes \phi_S, 1 \otimes \varphi_S) = \Lambda(\phi_R \otimes \psi_S, \varphi_R \otimes 1) = 0$$

for all bump functions ϕ_R, φ_S and all bump functions ψ_R with mean zero.

Then $\Lambda_0, \Lambda_1, \Lambda_2$ are bounded bilinear forms on L^p.

Proof. Again, we prove the result only for Λ. As in previous theorem, we use a L^2-normalized wavelet expansion of the functions appearing in the dual pair, we parametrize the terms accordingly with eccentricity and relative distances in exactly the same way and we apply Lemma 5.4 to obtain

$$|\Lambda(f,g)| \leq \sum_{R,S} |\Lambda(\psi_R, \psi_S)||\langle f, \psi_R\rangle||\langle g, \psi_S\rangle|$$

$$\leq C \sum_{i=1,2}\sum_{e_i \in \mathbb{Z}}\sum_{m_i \in \mathbb{N}} 2^{-|e_1+e_2|(1/2+\delta)}(m_1 m_2)^{-(1+\delta)} \sum_{(R,S)\,\in\, P_{e,m}} |\langle f, \psi_R\rangle||\langle g, \psi_S\rangle|$$

where $P_{e,m}$ is the set of pairs of dyadic rectangles (R,S) such that $|R_i| = 2^{e_i}|S_i|$ and $n_i \leq \frac{\text{diam}(R_i \cup S_i)}{\max(|R_i|,|S_i|)} < n_i + 1$. Notice that $(R,S) \in P_{e,m}$ if and only if $(S,R) \in P_{-e,m}$.

Now, we denote by $K = K_1 \times K_2$ the rectangle minimum, that is, such that $K_i = R_i$ if $|R_i| \leq |S_i|$ and $K_i = S_i$ otherwise. This way, the inner sum can be rewritten as

$$\sum_{(R,S)\,\in\, P_{e,m}} \int_{\mathbb{R}^2} \frac{|\langle f, \psi_R\rangle|}{|K|^{1/2}} \frac{|\langle g, \psi_S\rangle|}{|K|^{1/2}} \chi_K(x) dx$$

$$\leq \int_{\mathbb{R}^2} \Big(\sum_{(R,S)\,\in\, P_{e,m}} \frac{|\langle f, \psi_R\rangle|^2}{|K|} \chi_K(x) \Big)^{1/2} \Big(\sum_{(S,R)\,\in\, P_{-e,m}} \frac{|\langle g, \psi_S\rangle|^2}{|K|} \chi_K(x) \Big)^{1/2} dx \tag{7.1}$$

In order to build up the modified square functions, we denote by $k, k' \in \mathbb{Z}^2$ the scale parameters $k_i = \max(e_i, 0)$, $k'_i = -\min(e_i, 0)$ and by $n, n' \in \mathbb{Z}^2$ the translation parameters $n_i = m_i$, $n'_i = 1$ if $e_i \geq 0$ while $n_i = 1$, $n'_i = m_i$ if $e_i \leq 0$. Notice that $2^{\max(e_i,0)} 2^{-\min(e_i,0)} = 2^{|e_i|}$ and $n_i n'_i = m_i$.

We show how to bound the first factor. By the choice of K we have that $|R_i|, |S_i| \geq |K_i|$. If $k_i \geq 0$ then $K_i = R_i$ and there is nothing to show. So, we may assume $k_i \leq 0$ and $K = S$ in which case

$$\tilde{S}_{k,n}(f)(x) = \Big(\sum_{(R,S)\in\mathcal{P}_{k,n}} \frac{\langle f, \psi_R\rangle^2}{|S|} \chi_S(x) \Big)^{1/2} = \Big(\sum_R \langle f, \psi_R\rangle^2 \sum_{\substack{S \\ (R,S)\,\in\,\mathcal{P}_{k,n}}} \frac{\chi_S(x)}{|S|} \Big)^{1/2}$$

$$= \Big(\sum_R \langle f, \psi_R\rangle^2 \frac{\chi_{\tilde{S}}(x)}{2^{-k}|\tilde{S}|} \Big)^{1/2} = 2^{(k_1+k_2)/2} SS_{0,n}(f)(x)$$

where $\tilde{S}$ is the dyadic rectangle such that $|\tilde{S}_i| = |R_i|$ and $|R_i|^{-1}\mathrm{diam}(\tilde{S}_i \cup R_i) = n_i$.

This implies that expression 7.1 is equal to

$$\int_{\mathbb{R}^2} 2^{\max(e_1,0)/2} 2^{\max(e_2,0)/2} SS_{0,n}(f)(x) 2^{-\min(e_1,0)/2} 2^{-\min(e_2,0)/2} SS_{0,n'}(g)(x) dx$$

$$\leq 2^{(|e_1|+|e_2|)/2} \|SS_{0,n}(f)\|_{L^p(\mathbb{R}^2)} \|SS_{0,n'}(g)\|_{L^{p'}(\mathbb{R}^2)}$$

According to the boundedness of the modified double square functions given by Proposition 7.3, this can be bounded by

$$C_p\, 2^{(|e_1|+|e_2|)/2} \prod_{i=1,2} (\log(n_i) + 1)^{|\frac{2}{p}-1|} (\log(n'_i) + 1)^{|\frac{2}{p}-1|} \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)}$$

$$= C_p\, 2^{(|e_1|+|e_2|)/2} \prod_{i=1,2} (\log(m_i) + 1)^{|\frac{2}{p}-1|} \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)}$$

$$\leq C_p\, 2^{(|e_1|+|e_2|)/2} \prod_{i=1,2} m_i^{\epsilon|\frac{2}{p}-1|} \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)}$$

Then, putting everything back together, we have

$$|\Lambda(f,g)| \leq C_p \sum_{i=1,2} \sum_{e_i\in\mathbb{Z}} \sum_{m_i\in\mathbb{N}} 2^{-(|e_1|+|e_2|)(1/2+\delta)} (m_1 m_2)^{-(1+\delta)} 2^{(|e_1|+|e_2|)/2} (m_1 m_2)^{\epsilon|\frac{2}{p}-1|}$$

$$= C_p \prod_{i=1,2} \sum_{e_i \in \mathbb{Z}} 2^{-|e_i|\delta} \sum_{m_i \in \mathbb{N}} m_i^{-(1+\delta-\epsilon|\frac{2}{p}-1|)} \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)} \leq C_p \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)}$$

as long as $\epsilon|\frac{2}{p} - 1| < 1 + \delta$.

8. The general case: different types of paraproducts

We devote this last section to the extension of the previous theorems to the general case, that is, the proof of boundedness for singular integral operators that do not satisfy the special cancellation properties. As in the classical case, this is done by constructing appropriate paraproducts. But in the multiparametric case, the process is be more involved not only because we need more paraproducts (three different types in total) but also because these paraproducts can not be independent each other.

In particular, let b_i with $i = 1, \ldots, 4$, be four functions in $\mathrm{BMO}(\mathbb{R}^2)$ and b_i with $i = 5, \ldots, 8$, be four functions in $\mathrm{BMO}(\mathbb{R})$. Let also Λ be a bilinear form satisfying the hypotheses of Theorem 2.9 such that $\Lambda(1 \otimes 1, \cdot) = b_1$ $\Lambda(\cdot, 1 \otimes 1) = b_2$, $\Lambda(\cdot \otimes 1, 1 \otimes \cdot) = b_3$, $\Lambda(1 \otimes \cdot, \cdot \otimes 1) = b_4$, $\Lambda(1 \otimes \cdot, \cdot \otimes \cdot) = b_5$ and so on.

In order to prove boundedness of Λ, we construct eight bilinear forms Λ_i organized in three different groups in such a way that their associated linear operators are bounded and moreover they recover the functions b_i, in the sense that for example $\Lambda_{b_1}(1 \otimes 1, \cdot) = b_1$, while the bilinear form vanishes in all other possible cases, namely, $\Lambda_{b_1}(\cdot, 1 \otimes 1) = \Lambda_{b_1}(\cdot \otimes 1, 1 \otimes \cdot) = 0$ and so on. As we will see, the last type four paraproducts will be construct not only using the functions b_i but also certain values of the previously constructed paraproducts evaluated over the function 1.

This way the bilinear form $\Lambda_T = \Lambda - \sum_i \Lambda_{b_i}$ satisfies the eight special cancellation hypotheses of Corollary 5.4 and so, by applying the corollary, we deduce that Λ_T is bounded. Moreover, since every Λ_i is also bounded by construction, we finally obtain boundedness of the initial form Λ.

Before we start with the construction of paraproducts, we present a lemma that in some way shows that the sufficient conditions we have used in the main theorem are the right ones, while it also justifies the paraproducts we will define later on. For the sake of simplicity, we write the proof only for operators that preserve the space support, since then the error terms are zero and then the expression can be written by means of the Haar basis.

Let $(h_I)_I$ the Haar basis in $\mathbb{R}$ defined by $h_I = |I|^{-1/2}\chi_{I_l} - |I|^{-1/2}\chi_{I_r}$ where I_l and I_r are the children intervals of I. Let $(h_R)_R$ the Haar basis in $\mathbb{R}^2$ defined by $h_R = h_{R_1} \otimes h_{R_2}$.

Lemma 8.1. *Let $T : C_0^\infty(\mathbb{R}^2) \to \mathbb{C}$ be a linear mapping continuous with respect to the topology in $C_0^\infty(\mathbb{R}^2)$, such that* $\operatorname{supp} T(f) \subset \operatorname{supp} f$. *Then*

$$\langle T(f), g \rangle = \sum_R \langle f, h_R \rangle \langle g, h_R \rangle \langle T(h_R), h_R \rangle$$

$$+ \Big\langle \sum_R \langle f, h_R^2 \rangle \langle g, h_R \rangle h_R \, , \, T(1) \Big\rangle + \Big\langle \sum_R \langle f, h_R \rangle \langle g, h_R^2 \rangle h_R \, , \, T^*(1) \Big\rangle$$

$$+\Big\langle \sum_R \langle f, h_{R_1}h^2_{R_2}\rangle\langle g, h^2_{R_1}h_{R_2}\rangle h_R\,,\, T_1(1)\Big\rangle + \Big\langle \sum_R \langle f, h^2_{R_1}h_{R_2}\rangle\langle g, h_{R_1}h^2_{R_2}\rangle h_R\,,\, T_1^*(1)\Big\rangle$$

$$+\sum_R \langle f, h_{R_1}h^2_{R_2}\rangle\langle g, h_{R_1}h_{R_2}\rangle\langle T(h_{R_1}\otimes 1), h_R\rangle + \sum_R \langle f, h_{R_1}h_{R_2}\rangle\langle g, h_{R_1}h^2_{R_2}\rangle\langle T^*(h_{R_1}\otimes 1), h_R\rangle$$

$$+\sum_R \langle f, h^2_{R_1}h_{R_2}\rangle\langle g, h_{R_1}h_{R_2}\rangle\langle T(1\otimes h_{R_2}), h_R\rangle + \sum_R \langle f, h_{R_1}h_{R_2}\rangle\langle g, h^2_{R_1}h_{R_2}\rangle\langle T^*(1\otimes h_{R_2}), h_R\rangle$$

Remark 8.1. *The formula for more general operators includes some error terms whose contribution is smaller than the one described in previous statement.*

Let $(\psi_R)_R$ be a wavelet basis in $\mathbb{R}^2$ and for every rectagle R, let ψ^2_R be a bump function L^1-adapted to R and mean one. Then, such general formula can be stated in the following way

$$\begin{aligned}\langle T(f), g\rangle &= \sum_{n\in\mathbb{Z}^2}\Big(\sum_R \langle f, \psi_R\rangle\langle g, \psi_{R^n}\rangle\langle T(\psi_R), \psi_{R^n}\rangle \\ &+ \Big\langle \sum_R \langle f, \psi^2_R\rangle\langle g, h_{R^n}\rangle\psi_R\,,\, T(1)\Big\rangle + \ldots \\ &+ \sum_R \langle f, \psi_{R_1}\psi^2_{R_2}\rangle\langle g, \psi_{R_1^n}\psi_{R_2^n}\rangle\langle T(\psi_{R_1}\otimes 1), \psi_{R^n}\rangle + \ldots\Big)\end{aligned}$$

where $R_i^n = R_i + n_i|R_i|$. The leading term is associated with $n = 0$, which is the one appearing in the statement of the lemma.

Proof. Since $\operatorname{supp} T(h_R) \subset \operatorname{supp} h_R = R$ we have

$$\langle T(f), g\rangle = \sum_{R\cap S\neq\emptyset} f_R g_S\langle T(h_R), h_S\rangle$$

where $f_R = \langle f, h_R\rangle$ and the same for the function g.

Now, given two dyadic rectangles R, S such that $R\cap S \neq \emptyset$ there are only nine different possibilities, namely,

1) $R = S$, which leads to $\langle T(h_R), h_R\rangle$
2) $R \subset S$, which gives $T^*(1)$
3) $S \subset R$, which analogously gives $T(1)$
4) $R < S$, meaning $R_1 \subset S_1$ and $R_2 \subset S_2$, which leads to $T_1(1) = T_2^*(1)$
5) $S < R$, meaning $R_1 = S_1$ and $R_2 \subset S_2$, which leads to $T_1^*(1) = T_2(1)$
6) $R_1 = S_1$ and $S_2 \subset R_2$, which leads to $\langle T(h_{R_1}\otimes 1), h_{R_1}\otimes h_{R_2}\rangle$
7) $R_1 = S_1$ and $R_2 \subset S_2$, which leads to $\langle T(h_{R_1}\otimes h_{R_2}), h_{R_1}\otimes 1\rangle$
8) $S_1 \subset R_1$ and $R_2 = S_2$, which leads to $\langle T(1\otimes h_{R_2}), h_{S_1}\otimes h_{R_2}\rangle$
9) $R_1 \subset S_1$ and $R_2 = S_2$, which leads to $\langle T(h_{R_1}\otimes h_{R_2}), 1\otimes h_{R_2}\rangle$

Then the decomposition of $\langle T(f), g\rangle$ is obtained as follows: from 1) we get directly the first term

$$\sum_R f_R g_R\langle T(h_R), h_R\rangle$$

From 2) and 3) we get the two following terms (we only write the second one)

$$\sum_R \sum_{R\subset S} f_R g_S \langle T(h_R), h_S\rangle = \sum_R f_R \langle T(h_R), \big(\sum_{R\subset S} g_S\, h_S\big)\chi_R)\rangle$$

$$= \sum_R f_R \langle T(h_R), m_R(g)\rangle = \sum_R f_R\, m_R(g)\langle T(h_R), 1\rangle = \sum_R f_R\, m_R(g)\langle h_R, T^*(1)\rangle$$

$$= \Big\langle \sum_R f_R\, m_R(g) h_R, T^*(1)\Big\rangle$$

where we write $m_R(g) = \big(\sum_{R\subset S} g_S\, h_S\big)\chi_R = \langle g, h_R^2\rangle = |R|^{-1}\int_R g(x)dx$.

On the other hand, from 4) and 5) and using the partial adjoints T_i we get the two following ones (we only write the fourth one)

$$\sum_R \sum_{R<S} f_R g_S \langle T(h_R), h_S\rangle = \sum_R \sum_{R<S} f_R g_S \langle T_2(h_{R_1}\otimes h_{S_2}), h_{S_1}\otimes f_{R_2}\rangle$$

$$= \sum_{R_1,S_2} \Big\langle T_2(h_{R_1}\otimes h_{S_2}), \big(\sum_{R_1\subset S_1, S_2\subset R_2} f_R g_S h_{S_1}\otimes h_{R_2}\big)\chi_{R_1\times S_2}\Big\rangle$$

$$= \sum_{R_1,S_2} \Big\langle T_2(h_{R_1}\otimes h_{S_2}), \big(\sum_{R_1\subset S_1} (g_{S_2})_{S_1} h_{S_1}\big) \otimes \big(\sum_{S_2\subset R_2} (f_{R_1})_{R_2} h_{R_2}\big)\chi_{R_1\times S_2}\Big\rangle$$

$$= \sum_{R_1,S_2} \langle T_2(h_{R_1}\otimes h_{S_2}), m_{R_1}(g_{S_2}) m_{S_2}(f_{R_1})\rangle = \sum_{R_1,S_2} m_{R_1}(g_{S_2}) m_{S_2}(f_{R_1})\langle T_2(h_{R_1}\otimes h_{S_2}), 1\rangle$$

$$= \langle \sum_{R_1,S_2} m_{S_2}(f_{R_1}) m_{R_1}(g_{S_2}) h_{R_1}\otimes h_{S_2}, T_2^*(1)\rangle = \langle \sum_R m_{R_2}(f_{R_1}) m_{R_1}(g_{R_2}) h_R, T_1(1)\rangle$$

From 6) 7), 8) and 9) we get the remaining terms (we only write the sixth one)

$$\sum_S \sum_{R_1=S_1, S_2\subset R_2} f_R g_S \langle T(h_R), h_S\rangle = \sum_S \sum_{R_1=S_1, S_2\subset R_2} f_R g_S \langle h_R, T^*(h_S)\rangle$$

$$= \sum_S g_S \Big\langle \big(\sum_{S_2\subset R_2} f_{S_1\times R_2} h_{S_1\times R_2}\big)\chi_{S_2}, T^*(h_S)\Big\rangle = \sum_S g_S \Big\langle \big(\sum_{S_2\subset R_2} \langle f, h_{S_1\times R_2}\rangle h_{S_1}\otimes h_{R_2}\big)\chi_{S_2}, T^*(h_S)\Big\rangle$$

$$= \sum_S g_S \Big\langle h_{S_1}\otimes \big(\sum_{S_2\subset R_2} \langle f, h_{S_1\times R_2}\rangle h_{R_2}\big)\chi_{S_2}, T^*(h_S)\Big\rangle = \sum_S g_S \langle h_{S_1}\otimes m_{S_2}(f_{S_1}), T^*(h_S)\rangle$$

$$= \sum_S m_{S_2}(f_{S_1}) g_S \langle h_{S_1}\otimes 1, T^*(h_S)\rangle = \sum_S m_{S_2}(f_{S_1}) g_S \langle T(h_{S_1}\otimes 1), h_S\rangle$$

We start now with the construction of paraproducts. We need up to eight of such operators but by symmetry it will be enough to show only three of them. In particular, we construct the paraproducs associated $T(1)$, $T_1(1)$ and $\langle T(\psi_{R_1}\otimes 1), \psi_{S_1}\rangle$.

Lemma 8.2. *(Classical paraproducts). Given a function b in* $\mathrm{BMO}(\mathbb{R}^2)$*, there exists a bounded bilinear form Λ_b^1 such that $\Lambda_b^1(1\otimes 1, \cdot\otimes\cdot) = b$, $\Lambda_b^1(\cdot\otimes\cdot, 1\otimes 1) = \Lambda_b^1(\cdot\otimes 1, 1\otimes\cdot) = \Lambda_b^1(1\otimes\cdot, \cdot\otimes 1) = 0$, $\Lambda_b^1(\cdot\otimes\cdot, 1\otimes\cdot) = \Lambda_b^1(\cdot\otimes\cdot, \cdot\otimes 1) = 0$.*

Proof. Let $(\psi_I)_I$ be a wavelets basis on $L^2(\mathbb{R})$. Let $(\psi_R)_R$ be the wavelets basis on $L^2(\mathbb{R}^2)$ defined by $\psi_R = \psi_{R_1} \otimes \psi_{R_2}$. We denote by ψ_I^2 a bump function adapted to I such that $\widehat{\psi_I^2}$ has compact support in a set of measure comparable with $|I|^{-1}$ with center the origin (?). Let finally $\psi_R^2 = \psi_{R_1}^2 \otimes \psi_{R_2}^2$.

We define the bilinear form

$$\Lambda_b^1(f,g) = \sum_R \langle b, \psi_R\rangle \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle$$

which, to simplify notation, we will just denote by Λ_b during the proof of the lemma.

At least formally, Λ_b satisfies

$$\Lambda_b(1,g) = \langle g, b\rangle$$
$$\Lambda_b(f,1) = \Lambda_b(f_1 \otimes 1, 1 \otimes g_2) = \Lambda_b(1 \otimes f_2, g_1 \otimes 1) = 0$$

For the proof of their boundedness we proceed by using the duality $H^1(\mathbb{R}^2) - \mathrm{BMO}(\mathbb{R}^2)$. Since

$$\Lambda_b(f,g) = \sum_R \langle b, \psi_R\rangle \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle = \Big\langle b, \sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R \Big\rangle$$

we have

$$|\Lambda_b(f,g)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \| \sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R\|_{H^1(\mathbb{R}^2)}$$

Just assuming the sum is finite, we get that $\sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R \in H^1(\mathbb{R}^2)$ and then

$$\| \sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R\|_{H^1(\mathbb{R}^2)} \approx \|S(\sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R)\|_{L^1(\mathbb{R}^2)}$$

with implicit constants independent of the number of terms in the sum. Now

$$\begin{aligned} S\Big(\sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R\Big)^2 \\ &= \sum_R \langle f, \psi_{R_1}^2 \otimes \psi_{R_2}^2\rangle^2 \langle g, \psi_{R_1} \otimes \psi_{R_2}\rangle^2 \frac{\chi_{R_1}}{|R_1|} \otimes \frac{\chi_{R_2}}{|R_2|} \\ &\leq \sup_R |\langle f, \psi_{R_1}^2 \otimes \psi_{R_2}^2\rangle|^2 \ \sum_R \langle g, \psi_{R_1} \otimes \psi_{R_2}\rangle^2 \frac{\chi_{R_1}}{|R_1|} \otimes \frac{\chi_{R_2}}{|R_2|} \\ &= (M \otimes M)(f)^2 (S \otimes S)(g)^2 \end{aligned}$$

where $M \otimes M$ and $S \otimes S$ are defined by the two previous expressions and are known to be bounded operators on $L^p(\mathbb{R}^2)$. Then finally

$$\begin{aligned} \|S(\sum_R \langle f, \psi_R^2\rangle \langle g, \psi_R\rangle \psi_R)\|_{L^1(\mathbb{R}^2)} \\ \leq \|(M \otimes M)(f)\|_{L^p(\mathbb{R}^2)} \|(S \otimes S)(g)\|_{L^{p'}(\mathbb{R}^2)} \leq C\|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)} \end{aligned}$$

To end, we still need to prove that this family of operators also belong to the class of operators for which the theory applies. In particular, we show that they have integral representations like the ones stated in definition 2.5 with kernels satisfying the definition of a product Calderon-Zygmund kernel 2.1. We also prove the operator satisfies the weak boundedness Calderón-Zygmund condition stated in 2.8. From

$$\Lambda_b(f,g)=\sum_R\langle b,\psi_R\rangle\langle f,\psi_R^2\rangle\langle g,\psi_R\rangle=\int f(t)g(x)\sum_R\langle b,\psi_R\rangle\psi_R^2(t)\psi_R(x)dtdx$$

we obtain the integral representation regardless disjointness of supports of the argument funtions. Moreover,

$$K(x,t)=\Big\langle b,\sum_R\psi_R^2(t)\psi_R(x)\psi_R\Big\rangle$$

and we check the two properties of a product C-Z kernel:

$$|K(x,t)|\leq\|b\|_{\mathrm{BMO}(\mathbb{R}^2)}\Big\|\sum_R\psi_R^2(t)\psi_R(x)\psi_R\Big\|_{H^1(\mathbb{R}^2)}$$

As before, the H^1-norm is equivalent to

$$\Big\|S\Big(\sum_R\psi_R^2(t)\psi_R(x)\psi_R\Big)\Big\|_{L^1(\mathbb{R}^2)}=\int\Big(\sum_R\psi_R^2(t)^2\psi_R(x)^2\frac{\chi_R(y)}{|R|}\Big)^{1/2}dy$$

$$=\prod_{i=1,2}\int\Big(\sum_{R_i}\psi_{R_i}^2(t_i)^2\psi_{R_i}(x_i)^2\frac{\chi_{R_i}(y_i)}{|R_i|}\Big)^{1/2}dy_i$$

Let I_{x_i,t_i} be the smallest dyadic interval such that $x_i,t_i\in I_{x_i,t_i}$ and let $(I_k)_{k\geq 0}$ the family of dyadic intervals such that $I_{x_i,t_i}\subset I_k$ with $|I_k|=2^k|I_{x_i,t_i}|$. Moreover, since $|\psi_{R_i}^2|\leq|R_i|^{-1}\chi_{R_i}$ and $|\psi_{R_i}|\leq|R_i|^{-1/2}\chi_{R_i}$, both integrals previously displayed are bounded by

$$\sum_{k\geq 0}\int_{I_{k+1}\backslash I_k}\Big(\sum_{j\geq k}\frac{1}{|I_j|^2}\frac{1}{|I_j|}\frac{1}{|I_j|}\Big)^{1/2}dy=\sum_{k\geq 0}\Big(\sum_{j\geq k}\frac{1}{|I_j|^4}\Big)^{1/2}|I_{k+1}\backslash I_k|$$

$$=\sum_{k\geq 0}\Big(\sum_{j\geq k}\frac{1}{2^{4j}|I_{x_i,t_i}|^4}\Big)^{1/2}2^k|I_{x_i,t_i}|\lesssim\frac{1}{|I_{x_i,t_i}|}\sum_{k\geq 0}\frac{1}{2^{2k}}2^k\lesssim\frac{1}{|x_i-t_i|}$$

ending the first condition. For the second one, we prove that

$$|\partial_{t_1}\partial_{t_2}K(x,t)|+|\partial_{t_1}\partial_{x_2}K(x,t)|+|\partial_{x_1}\partial_{t_2}K(x,t)|+|\partial_{x_1}\partial_{x_2}K(x,t)|\leq C\prod_{i=1,2}|x_i-t_i|^{-2}$$

For expository reasons we deal only with the second term

$$\partial_{t_1}\partial_{x_2}K(x,t)=\Big\langle b,\sum_R\partial_{t_1}\psi_{R_1}^2(t_1)\psi_{R_2}^2(t_2)\psi_{R_1}(x_1)\partial_{x_2}\psi_{R_2}(x_2)\psi_R\Big\rangle$$

The four possible terms are not really symmetric since the averaging function $\psi_{R_i}^2$ only appear in the t_i variables. So, sometimes the derivatives hit an averaging function while some other times they do not. However, it is the presence of derivatives of wavelets what produces the final estimates, regardless whether it is $\partial_{t_i}\psi^2_{R_i}(t_i)$ or $\partial_{x_i}\psi_{R_i}(x_i)$. Actually, in all cases the derivatives increase by one the degree of the powers in $|R|$ involved and so, they have the same impact in all four terms. Let's see this point. As before

$$|\partial_{t_1}\partial_{x_2}K(x,t)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)}\Big\| \sum_{R_1} \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_2}(t_2)\psi_{R_1}(x_1)\partial_{x_2}\psi_{R_2}(x_2)\psi_R\Big\|_{H^1(\mathbb{R})}$$

with the H^1-norm equivalent to

$$\Big\|S\Big(\sum_R \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_2}(t_2)\psi_{R_1}(x_1)\partial_{x_2}\psi_{R_2}(x_2)\psi_R\Big)\Big\|_{L^1(\mathbb{R}^2)}$$

$$= \int_{\mathbb{R}^2}\Big(\sum_R (\partial_{t_1}\psi^2_{R_1}(t_1))^2\psi_{R_2}(t_2)^2\psi_{R_1}(x_1)^2(\partial_{x_2}\psi_{R_2}(x_2))^2\frac{\chi_R(y)}{|R|}\Big)^{1/2}dy$$

$$= \int \Big(\sum_{R_1}(\partial_{t_1}\psi^2_{R_1}(t_1))^2\psi_{R_1}(x_1)^2\frac{\chi_{R_1}(y_1)}{|R_1|}\Big)^{1/2}dy_1 \int\Big(\sum_{R_2}\psi^2_{R_2}(t_2)^2(\partial_{x_2}\psi_{R_2}(x_2))^2\frac{\chi_{R_2}(y_2)}{|R_2|}\Big)^{1/2}dy_2$$

Let I_{x_1,t_1} be the smallest dyadic interval such that $x_1, t_1 \in I_{x_1,t_1}$ and let $(I_k)_{k\geq 0}$ the family of dyadic intervals such that $I_{x_1,t_1} \subset I_k$ with $|I_k| = 2^k|I_{x_1,t_1}|$. Since $|\partial_{t_1}\psi^2_{R_1}(t_1)| \leq |R_1|^{-2}\chi_{R_1}(t_1)$ and $|\partial_{t_2}\psi_{R_2}(t_2)| \leq |R_2|^{-3/2}\chi_{R_2}(t_2)$ then, the first previous integrals can be bounded by

$$\sum_{k\geq 0}\int_{I_{k+1}\backslash I_k}\Big(\sum_{j\geq k}\frac{1}{|I_j|^4}\frac{1}{|I_j|}\frac{1}{|I_j|}\Big)^{1/2}dy = \sum_{k\geq 0}\Big(\sum_{j\geq k}\frac{1}{|I_j|^6}\Big)^{1/2}|I_{k+1}\backslash I_k|$$

$$= \sum_{k\geq 0}\Big(\sum_{j\geq k}\frac{1}{2^{6j}|I_{x_i,t_i}|^6}\Big)^{1/2}2^k|I_{x_i,t_i}| \lesssim \frac{1}{|I_{x_i,t_i}|^2}\sum_{k\geq 0}\frac{1}{2^{3k}}2^k \lesssim \frac{1}{|x_i-t_i|^2}$$

while the second one is bounded by

$$\sum_{k\geq 0}\int_{I_{k+1}\backslash I_k}\Big(\sum_{j\geq k}\frac{1}{|I_j|^2}\frac{1}{|I_j|^3}\frac{1}{|I_j|}\Big)^{1/2}dy = \sum_{k\geq 0}\Big(\sum_{j\geq k}\frac{1}{|I_j|^6}\Big)^{1/2}|I_{k+1}\backslash I_k| \lesssim \frac{1}{|x_i-t_i|^2}$$

This way,

$$|\partial_{t_1}\partial_{x_2}K(x,t)| \lesssim \|b\|_{\mathrm{BMO}(\mathbb{R}^2)}\frac{1}{|x_1-t_1|^2}\frac{1}{|x_2-t_2|^2}$$

On the other hand, we also have

$$\Lambda_b(f_1\otimes f_2, g_1\otimes g_2) = \int f_1(t_1)g_1(x_1)\sum_R\langle b,\psi_R\rangle\langle f_2,\psi^2_{R_2}\rangle\langle g_2,\psi_{R_2}\rangle\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)dt_1dx_1$$

and we obtain the integral representation regardless disjointness of supports of the argument funtions. Moreover,

$$\Lambda^2_{x_1,t_1}(f_2,g_2) = \Big\langle b, \sum_R\langle f_2,\psi^2_{R_2}\rangle\langle g_2,\psi_{R_2}\rangle\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R\Big\rangle$$

and we check the two properties of a weakboundedness-CZ condition: for all bump functions f_2, g_2 which are L^2-adapted to the same interval,

$$|\Lambda^2_{x_1,t_1}(f_2,g_2)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \Big\| \sum_R \langle f_2, \psi^2_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R \Big\|_{H^1(\mathbb{R}^2)}$$

As before, the H^1-norm is equivalent to

$$\Big\| S\Big(\sum_R \langle f_2, \psi^2_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R \Big) \Big\|_{L^1(\mathbb{R}^2)}$$

$$= \int \Big(\sum_R \langle f_2, \psi^2_{R_2}\rangle^2 \langle g_2, \psi_{R_2}\rangle^2 \psi^2_{R_1}(t_1)^2 \psi_{R_1}(x_1)^2 \frac{\chi_R(y)}{|R|} \Big)^{1/2} dy$$

$$= \int \Big(\sum_{R_1} \psi^2_{R_1}(t_1)^2 \psi_{R_1}(x_1)^2 \frac{\chi_{R_1}(y_1)}{|R_1|} \Big)^{1/2} dy_1 \int \Big(\sum_{R_2} \langle f_2, \psi^2_{R_2}\rangle^2 \langle g_2, \psi_{R_2}\rangle^2 \frac{\chi_{R_2}(y_2)}{|R_2|} \Big)^{1/2} dy_2$$

$$\lesssim \frac{1}{|x_1 - t_1|} \int M(f_2)(y_2) S(g_2)(y_2) dy_2$$

$$\leq C \|f_2\|_{L^2(\mathbb{R})} \|g_2\|_{L^2(\mathbb{R})} \frac{1}{|x_1 - t_1|} \leq C \frac{1}{|x_1 - t_1|}$$

Finally, we need to prove the analog estimates for $(\Lambda^2_{x_1,t_1} - \Lambda^2_{x_1',t_1'})(f_2,g_2)$ which will be deduced from $|\partial_{t_1}\Lambda^2_{x_1,t_1}(f_2,g_2)| + |\partial_{x_1}\Lambda^2_{x_1,t_1}(f_2,g_2)| \leq C|x_1 - t_1|^{-2}$. By symmetry, we work only with one of such terms:

$$\partial_{t_1}\Lambda^2_{x_1,t_1}(f_2,g_2) = \Big\langle b, \sum_R \langle f_2, \psi^2_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R \Big\rangle$$

and therefore, as before,

$$|\partial_{t_1}\Lambda_{x_1,t_1}(f_2,g_2)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \Big\| \sum_R \langle f_2, \psi^2_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R \Big\|_{H^1(\mathbb{R}^2)}$$

being the H^1-norm is equivalent to

$$\Big\| S\Big(\sum_R \langle f_2, \psi^2_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_R \Big) \Big\|_{L^1(\mathbb{R}^2)}$$

$$= \int \Big(\sum_{R_1} (\partial_{t_1}\psi^2_{R_1}(t_1))^2 \psi_{R_1}(x_1)^2 \frac{\chi_{R_1}(y_1)}{|R_1|} \Big)^{1/2} dy_1 \int \Big(\sum_{R_2} \langle f_2, \psi^2_{R_2}\rangle^2 \langle g_2, \psi_{R_2}\rangle^2 \frac{\chi_{R_2}(y_2)}{|R_2|} \Big)^{1/2} dy_2$$

$$\lesssim \frac{1}{|x_1 - t_1|^2} \int M(f_2)(y_2) S(g_2)(y_2) dy_2 \leq C \frac{1}{|x_1 - t_1|^2}$$

We also need to prove endpoint estimates for this paraproduct.

$$\Lambda_b(1 \otimes f_1, \psi \otimes g_2) = \sum_R \langle b, \psi_R\rangle \langle f_2, \psi_{R_2}\rangle \langle \psi, \psi_{R_1}\rangle \langle g_2, \psi_{R_2}\rangle$$

$$= \Big\langle b, \sum_R \langle f_2, \psi_{R_2}\rangle \langle \psi, \psi_{R_1}\rangle \langle g_2, \psi_{R_2}\rangle \psi_{R_1} \otimes \psi_{R_2} \Big\rangle$$

$$= \Big\langle b, (\sum_{R_1} \langle \psi, \psi_{R_1}\rangle \psi_1) \otimes (\sum_{R_2} \langle f_2, \psi_{R_2}\rangle \langle g_2, \psi_{R_2}\rangle \psi_{R_2})\Big\rangle$$

$$= \langle b, \psi \otimes \prod\nolimits_{f_2}(g_2)\rangle$$

and so

$$\Lambda_b(1\otimes f_1, \psi\otimes g_2) \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \|\psi \otimes \prod\nolimits_{f_2}(g_2)\|_{H^1(\mathbb{R}^2)}$$
$$\approx \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \|S(\psi \otimes \prod\nolimits_{f_2}(g_2))\|_{L^1(\mathbb{R}^2)}$$
$$= \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \|S(\psi)\|_{L^1(\mathbb{R})} \|\prod\nolimits_{f_2}(g_2)\|_{L^1(\mathbb{R})}$$
$$\leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \|\psi\|_{H^1(\mathbb{R})} \|M(f_2)S(g_2)\|_{L^1(\mathbb{R})}$$
$$\leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \|\psi\|_{H^1(\mathbb{R})} \|f_2\|_{L^p(\mathbb{R})} \|g_2\|_{L^{p'}(\mathbb{R})}$$

which proves that $\Lambda_b(1\otimes f_1, \cdot\otimes g_2) \in \mathrm{BMO}(\mathbb{R})$.

We notice that we cannot demmand these paraproducts to satisfy that $\Lambda_b(1\otimes f_2, \psi\otimes g_2) = 0$ for all smooth functions f_2, g_2 and all bump functions ψ with mean zero. The same thing will happen with the following class of paraproducts which will force us to deal the third class of paraproducts in a special manner.

We continue with the so called mixed paraproduct, that is, the one associated with $T_1(1)$.

Lemma 8.3. *(Mixed paraproducts). Given a function b in* $\mathrm{BMO}(\mathbb{R}^2)$*, there exists a bounded bilinear form Λ_b^2 such that $\Lambda_b^2(1\otimes\cdot, \cdot\otimes 1) = b$, $\Lambda_b^2(1\otimes 1, \cdot\otimes\cdot) = \Lambda_b^2(\cdot\otimes\cdot, 1\otimes 1) = \Lambda_b^2(\cdot\otimes 1, 1\otimes\cdot) = 0$.*

Proof. Using the same basis as in previous lemma, we define

$$\Lambda_b^2(f,g) = \sum_R \langle b, \psi_R\rangle \langle f, \psi_{R_1}^2 \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1}\otimes \psi_{R_2}^2\rangle$$

which in this proof we will just denote by Λ_b.

At least formally, Λ_b satisfies

$$\Lambda_b(1\otimes f_2, g_1\otimes 1) = \langle b, f_2\otimes g_1\rangle$$
$$\Lambda_b(1,g) = \Lambda_b(f,1) = \Lambda_b(f_1\otimes 1, 1\otimes g_2) = 0$$

For the proof of their boundedness we proceed as before.

$$\Lambda_b(f,g) = \sum_R \langle b, \psi_R\rangle \langle f, \psi_{R_1}^2 \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1}\otimes \psi_{R_2}^2\rangle$$

$$= \Big\langle b, \sum_R \langle f, \psi_{R_1}^2 \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1}\otimes \psi_{R_2}^2\rangle \psi_R\Big\rangle$$

and then, using the duality $H^1(\mathbb{R}^2) - \mathrm{BMO}(\mathbb{R}^2)$ we have

$$|\Lambda_b(f,g)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R}^2)} \| \sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle \psi_R \|_{H^1(\mathbb{R}^2)}$$

Just assuming the sum is finite, we get that $\sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle \psi_R \in H^1(\mathbb{R}^2)$ and then

$$\| \sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle \psi_R \|_{H^1(\mathbb{R}^2)} \approx \|S(\sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle \psi_R)\|_{L^1(\mathbb{R}^2)}$$

with implicit constants independent of the number of terms in the sum. Since

$$\begin{aligned} S\Big(\sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle \psi_R\Big)^2 (x,y) \\ = \sum_R \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle^2 \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle^2 \frac{\chi_{R_1}}{|R_1|}(x) \frac{\chi_{R_2}}{|R_2|}(y) \\ \leq \Big(\sup_{R_1} \sum_{R_2} \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle^2 \chi_{R_1}(x) \frac{\chi_{R_2}}{|R_2|}(y)\Big)\Big(\sum_{R_1} \sup_{R_2} \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle^2 \frac{\chi_{R_1}}{|R_1|}(x) \chi_{R_2}(y)\Big) \\ = M_1(S_2(f))^2(x,y) S_1(M_2(g))^2(x,y) \end{aligned}$$

where the given expressions are not a composition of operators but just notation. We first prove that those operators are bounded on $L^p(\mathbb{R}^2)$. We do so by applying Fefferman-Stein's inequality to $S_i M_j$: by denoting $g_y(x) = g(x,y)$, we have

$$\begin{aligned} \Big\| \Big(\sum_{R_1} \sup_{R_2} \langle g, \psi_{R_1} \otimes \psi^2_{R_2}\rangle^2 \chi_{R_2} \frac{\chi_{R_1}}{|R_1|}\Big)^{1/2} \Big\|_{L^{p'}(\mathbb{R}^2)} \leq \Big\| \Big(\sum_{R_1} M(\langle g, \psi_{R_1}\rangle)^2 \frac{\chi_{R_1}}{|R_1|}\Big)^{1/2} \Big\|_{L^{p'}(\mathbb{R}^2)} \\ = \Big(\int_{\mathbb{R}} \Big\| \Big(\sum_{R_1} M(\langle g, \psi_{R_1}\rangle)^2(y) \frac{\chi_{R_1}}{|R_1|}\Big)^{1/2} \Big\|^{p'}_{L^{p'}(\mathbb{R})} dy\Big)^{1/p'} \\ \leq C\Big(\int_{\mathbb{R}} \Big\| \Big(\sum_{R_1} \langle g_y, \psi_{R_1}\rangle^2 \frac{\chi_{R_1}}{|R_1|}\Big)^{1/2} \Big\|^{p'}_{L^{p'}(\mathbb{R})} dy\Big)^{1/p'} \\ = C\Big(\int_{\mathbb{R}} \|S(g_y)\|^{p'}_{L^{p'}(\mathbb{R})} dy\Big)^{1/p'} \leq C\Big(\int_{\mathbb{R}} \|g_y\|^{p'}_{L^{p'}(\mathbb{R})} dy\Big)^{1/p'} = C\|g\|_{L^{p'}(\mathbb{R}^2)} \end{aligned}$$

The other operator is easier because of the pointwise inequality $M_i S_j \leq S_j M_i$.

Then, we finally get

$$\begin{aligned} \|S(\sum_R \langle f, \psi^2_R\rangle \langle g, \psi_R\rangle \psi_R)\|_{L^1(\mathbb{R}^2)} \\ \leq \|(M_1S_2)(f)\|_{L^p(\mathbb{R}^2)} \|(S_1M_2)(g)\|_{L^{p'}(\mathbb{R}^2)} \leq C\|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)} \end{aligned}$$

To prove that this operators belong to the class of operators with a product Calderón-Zygmund kernel satisfying the WB-CZ condition, we apply the same reasoning as in the case of classical paraproducts. We do not write the details.

Now we construct the last class of paraproducts, the ones associated with the terms $\langle T(1 \otimes \cdot), \cdot \otimes \cdot\rangle$.

Lemma 8.4. *(Third type of paraproducts). Given a function b in* $\mathrm{BMO}(\mathbb{R})$*, there exists a bounded bilinear form Λ_b^3 such that* $\Lambda_b^3(1\otimes 1, \cdot\otimes\cdot) = \Lambda_b^3(\cdot\otimes\cdot, 1\otimes 1) = \Lambda_b^3(\cdot\otimes 1, 1\otimes\cdot) = \Lambda_b^3(1\otimes\cdot, \cdot\otimes 1) = 0$, $\Lambda_b^3(1\otimes\cdot, \cdot\otimes\cdot) = b$ *and* $\Lambda_b^3(\cdot\otimes 1, \cdot\otimes\cdot) = \Lambda_b^3(\cdot\otimes\cdot, 1\otimes\cdot) = \Lambda_b^3(\cdot\otimes\cdot, \cdot\otimes 1) = 0$.

Proof. Let $(\psi_I)_I$ be a wavelet basis in $L^2(\mathbb{R})$ such that every ψ_I is a bump function adapted to a dyadic interval I with constant $C > 0$.

We define

$$\Lambda_b^3(f, g) = \sum_R \langle b, \psi_{R_1}\rangle\langle f, \psi_{R_1}^2 \otimes \psi_{R_2}\rangle\langle g, \psi_{R_1} \otimes \psi_{R_2}\rangle$$

which along the proof of its properties we will just denote by Λ_b.

At least formally, Λ_b satisfies

$$\Lambda_b(1, g) = \Lambda_b(f, 1) = \Lambda_b(f_1 \otimes 1, 1 \otimes g_2) = \Lambda_b(1 \otimes f_2, g_1 \otimes 1) = 0$$
$$\Lambda_b(f_1 \otimes 1, g_1 \otimes g_2) = \Lambda_b(f_1 \otimes f_2, 1 \otimes g_2) = \Lambda_b(f_1 \otimes f_2, g_1 \otimes 1) = 0$$

being the proof trivial in all cases. It also trivially satisfies that for every $f_2, g_2 \in \mathcal{S}(\mathbb{R})$

$$\Lambda_b(1 \otimes f_2, \psi \otimes g_2) = \sum_R \langle b, \psi_{R_1}\rangle\langle f_2, \psi_{R_2}\rangle\langle \psi, \psi_{R_1}\rangle\langle g_2, \psi_{R_2}\rangle$$
$$= \sum_R \langle b \otimes f_2, \psi_R\rangle\langle \psi \otimes g_2, \psi_R\rangle = \langle b \otimes f_2, \psi \otimes g_2\rangle$$

We prove now boundedness of Λ_b. Notice that

$$\Lambda_b(f, g) = \sum_R \langle b, \psi_{R_1}\rangle\langle\langle f, \psi_{R_2}\rangle, \psi_{R_1}^2\rangle\langle\langle g, \psi_{R_2}\rangle, \psi_{R_1}\rangle = \sum_{R_2} \Big\langle \prod\nolimits_b(\langle f, \psi_{R_2}\rangle), \langle g, \psi_{R_2}\rangle \Big\rangle$$

Then, by boundedness of paraproducts we have

$$|\Lambda_b(f, g)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R})} \sum_{R_2} \|\langle f, \psi_{R_2}\rangle\|_p \|\langle g, \psi_{R_2}\rangle\|_{p'}$$

Now, the sum can be rewritten as

$$\sum_{R_2} \int_{\mathbb{R}} \|\langle f, \psi_{R_2}\rangle\|_p \|\langle g, \psi_{R_2}\rangle\|_{p'} \frac{\chi_{R_2}(x_2)}{|R_2|} dx_2$$

$$\leq \int_{\mathbb{R}} \Big(\sum_{R_2} \|\langle f, \psi_{R_2}\rangle\|_p^2 \frac{\chi_{R_2}(x_2)}{|R_2|} \Big)^{1/2} \Big(\sum_{R_2} \|\langle g, \psi_{R_2}\rangle\|_{p'} \frac{\chi_{R_2}(x_2)}{|R_2|} \Big)^{1/2} dx_2 \tag{8.1}$$

If we denote by

$$S^p(f) = \Big(\sum_{R_2} \|\langle f, \psi_{R_2}\rangle\|_p^2 \frac{\chi_{R_2}}{|R_2|} \Big)^{1/2}$$

the vector-valued square function, we have that expression 8.1 is equal to

$$\int_{\mathbb{R}^2} S^p(f)(x) S^{p'}(g)(x) dx \leq \|S^p(f)\|_{L^p(\mathbb{R})} \|S^{p'}(g)\|_{L^{p'}(\mathbb{R})}$$
$$\leq C_p \|f\|_{L^p(\mathbb{R}^2)} \|g\|_{L^{p'}(\mathbb{R}^2)}$$

ending the proof of boundedness.

We prove now that this family of operators also belong to the class of operators for which the theory applies. For this, we just need to show that they satisfy the integral representation stated in 2.5 with a kernel satisfying the definition 2.1 of a product Calderon-Zygmund kernel. From

$$\Lambda_b(f,g) = \sum_R \langle b, \psi_{R_1}\rangle \langle f, \psi^2_{R_1} \otimes \psi_{R_2}\rangle \langle g, \psi_{R_1} \otimes \psi_{R_2}\rangle$$

$$= \int f(t)g(x) \sum_R \langle b, \psi_{R_1}\rangle \psi^2_{R_1}(t_1)\psi_{R_2}(t_2)\psi_{R_1}(x_1)\psi_{R_2}(x_2)dtdx$$

we directly obtain the integral representation regardless disjointness of supports of the argument funtions. Moreover, this time the kernel is of tensor product type since

$$K(x,t) = \Big\langle b, \sum_{R_1} \psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{R_1}\Big\rangle \sum_{R_2} \psi_{R_2}(t_2)\psi_{R_2}(x_2)$$

and we check the two properties of a product C-Z kernel. As we have seen before, the first factor can be bounded by

$$\|b\|_{\mathrm{BMO}(\mathbb{R})} \Big\| \sum_{R_1} \psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{R_1}\Big\|_{H^1(\mathbb{R})} \lesssim \|b\|_{\mathrm{BMO}(\mathbb{R})} \frac{1}{|x_1 - t_1|}$$

For the second factor we reason as follows. Let I_{x_2,t_2} be the smallest dyadic interval such that $x_2, t_2 \in I_{x_2,t_2}$ and let $(I_k)_{k\geq 0}$ the family of dyadic intervals such that $I_{x_2,t_2} \subset I_k$ with $|I_k| = 2^k |I_{x_2,t_2}|$. Moreover, since $|\psi_{R_2}| \leq |R_2|^{-1/2}\chi_{R_2}$, we have

$$\sum_{R_2} |\psi_{R_2}(t_2)|\,|\psi_{R_2}(x_2)| \leq \sum_{k\geq 0} \frac{1}{|I_k|^{1/2}} \frac{1}{|I_k|^{1/2}}$$

$$= \sum_{k\geq 0} \frac{1}{2^k |I_{x_2,t_2}|} \lesssim \frac{1}{|I_{x_2,t_2}|} \lesssim \frac{1}{|x_2 - t_2|}$$

For the second condition, we prove that

$$|\partial_{t_1}\partial_{t_2}K(x,t)| + |\partial_{t_1}\partial_{x_2}K(x,t)| + |\partial_{x_1}\partial_{t_2}K(x,t)| + |\partial_{x_1}\partial_{x_2}K(x,t)| \leq C \prod_{i=1,2} |x_i - t_i|^{-2}$$

But, by symmetry we deal only with the first term

$$\partial_{t_1}\partial_{t_2}K(x,t) = \Big\langle b, \sum_{R_1} \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{R_1}\Big\rangle \sum_{R_2} \partial_{t_2}\psi_{R_2}(t_2)\psi_{R_2}(x_2)$$

Then, as before

$$|\partial_{t_1}\partial_{t_2}K(x,t)| \leq \|b\|_{\mathrm{BMO}(\mathbb{R})} \Big\| \sum_{R_1} \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{R_1}\Big\|_{H^1(\mathbb{R})} \sum_{R_2} |\partial_{t_1}\psi_{R_2}(t_2)|\,|\psi_{S_2}(x_2)|$$

with the $h^1(\mathbb{R})$ norm equivalent to

$$\Big\| S\Big(\sum_{R_1} \partial_{t_1}\psi^2_{R_1}(t_1)\psi_{R_1}(x_1)\psi_{R_1}\Big)\Big\|_{L^1(\mathbb{R}^2)} \lesssim \frac{1}{|x_1-t_1|^2}$$

For the second factor, we reason similarly as before. Now, let I_{x_2,t_2} be the smallest dyadic interval such that $x_2, t_2 \in I_{x_2,t_2}$ and let $(I_k)_{k\geq 0}$ the family of dyadic intervals such that $I_{x_2,t_2} \subset I_k$ with $|I_k| = 2^k|I_{x_2,t_2}|$. Moreover, since $|\psi_{R_2}| \leq |R_2|^{-1/2}\chi_{R_2}$, $|\partial_{t_2}\psi^2_{R_2}| \leq |R_1|^{-2}\chi_{R_2}$ we have

$$\sum_{R_2} |\partial_{t_1}\psi_{R_2}(t_2)|\,|\psi_{S_2}(x_2)| \leq \sum_{k\geq 0} \frac{1}{|I_k|^2}\frac{1}{|I_k|^{1/2}}$$

$$= \sum_{k\geq 0} \frac{1}{2^{3k/2}|I_{x_2,t_2}|^{3/2}} \lesssim \frac{1}{|I_{x_2,t_2}|^{3/2}} \lesssim \frac{1}{|x_2-t_2|^{3/2}}$$

proving finally

$$|\partial_{t_1}\partial_{t_2}K(x,t)| \lesssim \|b\|_{\mathrm{BMO}(\mathbb{R})}\frac{1}{|x_1-t_1|^2}\frac{1}{|x_2-t_2|^{3/2}}$$

The other properties are proven in a similar way.

To apply previously constructed operators to the problem of reduction to the special cancellation we proceed as follows. We first consider the functions defined by $\langle b_1, \psi\rangle = \Lambda(1\otimes 1, \psi)$, $\langle b_2, \psi\rangle = \Lambda(\psi, 1\otimes 1)$ and $\langle b_3, \psi\rangle = \Lambda(\psi_1\otimes 1, 1\otimes\psi_2)$, $\langle b_4, \psi\rangle = \Lambda(1\otimes\psi_2, \psi_1\otimes 1)$. By hypothesis, all of them are functions in $\mathrm{BMO}(\mathbb{R}^2)$ and so we can construct the paraproducts $\Lambda^1_{b_i}$ and $\Lambda^2_{b_j}$ for $i = 1,2$ and $j = 3,4$ respectively.

Now we define the bilinear form

$$\tilde{\Lambda} = \Lambda - \sum_{i=1,2} \Lambda^1_{b_i} - \sum_{j=3,4} \Lambda^2_{b_j}$$

which clearly satisfies the first cancellation conditions

$$\tilde{\Lambda}(1\otimes 1, \psi) = \tilde{\Lambda}(\psi, 1\otimes 1) = \tilde{\Lambda}(\psi_1\otimes 1, 1\otimes\psi_2) = \tilde{\Lambda}(1\otimes\psi_2, \psi_1\otimes 1) = 0$$

but not the remaining ones in the bump lemma 5.1 or its subsequent corollary.

We have seen in lemma 8.2 that $\Lambda^1_{b_i}(1\otimes f_2, \cdot\otimes g_2) \in \mathrm{BMO}(\mathbb{R})$ and in a similar way $\Lambda^2_{b_i}(1\otimes f_2, \cdot\otimes g_2) \in \mathrm{BMO}(\mathbb{R})$. Therefore, once we prove that also $\Lambda(1\otimes f_2, \psi\otimes g_2) \in \mathrm{BMO}(\mathbb{R})$ we will have that $\tilde{\Lambda}(1\otimes f_2, \psi\otimes g_2) \in \mathrm{BMO}(\mathbb{R})$ and so, we can define the function

$$\langle \tilde{b}_5\otimes f_2, \psi\otimes g_2\rangle = \tilde{\Lambda}(1\otimes f_2, \psi\otimes g_2) \tag{8.2}$$

for every ψ of mean zero. With such function we construct the third type of paraproducts $\Lambda^3_{b_5}$. We repeat the procedure three more times by taking into account the different permutations of the argument functions in the right hand part of equality (8.2). Finally, we define

$$\Lambda_T = \tilde{\Lambda} - \sum_{k=5}^{8} \Lambda^3_{\tilde{b}_i}$$

which clearly satisfies all the required cancellation conditions.

To prove that $\Lambda(1 \otimes f_2, \psi \otimes g_2) \in \mathrm{BMO}(\mathbb{R})$ we need first the following lemma

Lemma 8.5. *Let Λ be a bilinear Calderón-Zygmund form with associated kernel K and satisfying the mixed WB-CZ conditions.*

We also assume that Λ satisfies the weak boundedness condition and the special cancellation conditions:

$$\langle T(\phi_I \otimes 1), \varphi_I \otimes \cdot \rangle, \langle T(1 \otimes \phi_I), \cdot \otimes \varphi_I \rangle, \langle T^*(\phi_I \otimes 1), \varphi_I \otimes \cdot \rangle, \langle T^*(1 \otimes \phi_I), \cdot \otimes \varphi_I \rangle \in \mathrm{BMO}(\mathbb{R})$$

for all ϕ_I, φ_I bump functions adapted to I with norms uniformly bounded in I.

Then,

$$\langle T(\phi_I \otimes 1), \psi_J \otimes \cdot \rangle, \langle T(1 \otimes \phi_I), \cdot \otimes \psi_J \rangle, \langle T^*(\phi_I \otimes 1), \psi_J \otimes \cdot \rangle, \langle T^*(1 \otimes \phi_I), \cdot \otimes \psi_J \rangle \in \mathrm{BMO}(\mathbb{R})$$

for all ϕ_I, ψ_J bump functions supported and adapted to I and J respectively such that $I \cap J = \emptyset$ and ψ_J has mean zero, with norms satisfying

$$\|\langle T(\phi_I \otimes 1), \psi_J \otimes \cdot \rangle\|_{\mathrm{BMO}(\mathbb{R})} \lesssim \Big(\frac{\min(|I|,|J|)}{\max(|I|,|J|)}\Big)^{1/2+\delta} \Big(\frac{\mathrm{diam}(I \cup J)}{\max(|I|,|J|)}\Big)^{-(1+\delta)}$$

Proof. We assume that ϕ_I and ψ_J are supported in I and J respectively. The general case might need some extra decomposition of the argument functions involved in the same way we did in the proof of the bump lemma 5.1, but we will not develope the details here. We also assume that $|J| \leq |I|$ and ψ_J has mean zero.

If $|I|^{-1}\mathrm{diam}(I,J) \leq 2$ then ϕ_I and ψ_J are both adapted to I with the same constant and thus, by hypothesis $(WB \otimes T(1))$

$$|\Lambda(1 \otimes \phi_I, f \otimes \psi_J)| \leq C\|f\|_{H^1}$$

for any atom f.

If $|I|^{-1}\mathrm{diam}(I,J) > 2$, we reason as follows. Let f be an atom supported in K. Since ϕ_I and ψ_J have disjoint support, we have the following integral representation

$$\Lambda(1 \otimes \phi_I, f \otimes \psi_J) = \int_{\mathbb{R}^2} \phi_I(t_2)\psi_J(x_2)\Lambda_{x_2,t_2}(1,f)dt_2dx_2$$

Now, let Φ a bump function L^∞-adapted and supported in K. We denote $c_J = c(J)$ and $c_K = c(K)$. Then,

$$\Lambda(1 \otimes \phi_I, f \otimes \psi_J) = \int_{\mathbb{R}^2} \phi_I(t_2)\psi_J(x_2)\Lambda_{x_2,t_2}(\Phi,f)dt_2dx_2$$

$$+ \int_{\mathbb{R}^2} \phi_I(t_2)\psi_J(x_2)\Lambda_{x_2,t_2}(1-\Phi,f)dt_2dx_2$$

We use the mean zero of ψ_J to rewrite the first term in the following way

$$\int_{\mathbb{R}^2} \phi_I(t_2)\psi_J(x_2)(\Lambda_{x_2,t_2}(\Phi, f) - \Lambda_{c_J,t_2}(\Phi, f))dt_2dx_2$$

Now, we have $|x_2-t_2| \geq \text{diam}(I, J) > 2|I| \geq 2|J| \geq 2|x_2-c_J|$. Then, since $|K|^{-1/2}\Phi$ and $|K|^{1/2}f$ are L^2-adapted in the same interval, by the WB, we can bound previuos expression by

$$\int_{|x_2-t_2|>\text{diam}(I,J)} |\phi_I(t_1)||\psi_J(x_1)||(\Lambda_{x_2,t_2} - \Lambda_{c_J,t_2})(|K|^{-1/2}\Phi, |K|^{1/2}f)|dt_2dx_2$$

$$\lesssim \int_{|x_2-t_2|>\text{diam}(I,J)} |\phi_I(t_1)|\psi_J(x_2)\frac{|x_2 - c_J|^\delta}{|x_2 - t_2|^{1+\delta}}dt_2dx_2$$

$$\leq \|\phi_I\|_1\|\psi_J\|_1|J|^\delta\text{diam}(I, J)^{-(1+\delta)} \lesssim |I|^{1/2}|J|^{1/2}||J|^\delta\text{diam}(I, J)^{-(1+\delta)}$$

$$= \Big(\frac{|J|}{|I|}\Big)^{1/2+\delta}(|I|^{-1}\text{diam}(I, J))^{-(1+\delta)}$$

To deal with the second term, we notice that $\phi_I \otimes (1 - \Phi)$ and $\psi_J \otimes f$ have disjoint support and so we can use the integral representation

$$\int_{\mathbb{R}^4}(1 - \Phi)(t_1)f(x_1)\phi_I(t_2)\psi_J(x_2)K(x, t)dtdx$$

Now, because of the mean zero of both ψ_J and f we can rewrite the integral as

$$\int_{\mathbb{R}^4}(1-\Phi)(t_1)f(x_1)\phi_I(t_2)\psi_J(x_2)(K(x,t)-K((c_K, x_2), t)-K((x_1, c_J), t)+K((c_J, c_K), t))dtdx$$

Notice that $2|x_1 - c_K| \leq 2|K| < |x_1 - t_1|$ and $2|x_2 - c_J| < |x_2 - t_2|$. Then, by the property of product C-Z kernel, we can bound by

$$\int_{\mathbb{R}^4} |(1 - \Phi)(t_1)||f(x_1)||\phi_I(t_2)||\psi_J(x_2)|\frac{|x_1 - c_K|^\delta}{|x_1 - t_1|^{1+\delta}}\frac{|x_2 - c_J|^\delta}{|x_2 - t_2|^{1+\delta}}dtdx$$

$$\lesssim |J|^\delta\text{diam}(I, J)^{-(1+\delta)}\|\phi_I\|_1\|\psi_J\|_1|K|^\delta \int_{\mathbb{R}} |f(x_1)| \int_{|x_1-t_1|>|K|} \frac{1}{|x_1 - t_1|^{1+\delta}}dt_1dx_1$$

$$\lesssim |J|^\delta\text{diam}(I, J)^{-(1+\delta)}|I|^{1/2}|J|^{1/2}|K|^\delta\|f\|_1\frac{1}{|K|^\delta}$$

$$= \Big(\frac{|J|}{|I|}\Big)^{1/2+\delta}(|I|^{-1}\text{diam}(I, J))^{-(1+\delta)}\|f\|_1$$

which ends the proof of this lemma.

Now, with the help of previous lemma and the hypotheses... we can prove $\Lambda(1 \otimes f_2, \psi \otimes g_2) \in \text{BMO}(\mathbb{R})$. The procedure is similar to the one developed to prove L^p boundedness of the linear form under the special cancellation properties.

With the three previous lemmata and the other five symmetrical statements which come from all the possible permutations of the argument functions, we can finally prove boundedness of product singular integrals in the general case and finish this way the proof of theorem 2.9 and also theorem 7.4.

9. Appendix

In the proof of the extension to L^p spaces (see theorem 7.4), we used some bi-parameter modified square functions whose boundedness properties are a direct consequence of their uni-parameter counterparts. Now, in this appendix, we prove boundedness of such uni-parameter modified square functions.

Definition 9.1. *Given $k \in \mathbb{Z}$, $n \in \mathbb{N}$, we consider the following operator*

$$\tilde{S}_{k,n}(f)(x) = \Big(\sum_{I,J \in \mathcal{P}_{k,n}} \frac{\langle f, \psi_I \rangle^2}{|J|} \chi_J(x) \Big)^{1/2}$$

where $\mathcal{P}_{k,n}$ is the family of pairs of dyadic intervals (I, J) satisfying $|I| = 2^k|J|$ and $n \leq \frac{\mathrm{diam}(I \cup J)}{\max(|I|,|J|)} < n + 1$.

We will prove bounds of such operators by means of the following modified square functions:

Definition 9.2. *Given $k \in \mathbb{Z}$, $n \in \mathbb{N}$, we consider the following variant of square function*

$$S_{k,n}(f)(x) = \Big(\sum_{I} \frac{\langle f, \psi_I \rangle^2}{|J|} \chi_J(x) \Big)^{1/2}$$

where I and J are two dyadic intervals satisfying $|I| = 2^k|J|$ and $n \leq \frac{\mathrm{diam}(I \cup J)}{\max(|I|,|J|)} < n+1$, chosen in such a way that for every interval I there is a unique interval J.

This way we actually define a family of operators that depends on the particular choice of intervals but whose bounds do not depend on such choice, as we will soon prove. Notice that the particular choice does not depend on the point x.

We see now the the reason why this modified square function helps to control boundedness of the previous ones. Fixed a dyadic interval I, we denote $I_{k,m}$ the family of dyadic intervals J such that $|I| = 2^k|J|$ and $n \leq \max(|I|, |J|)^{-1}\mathrm{diam}(I \cup J) < n + 1$. We also denote by $\tilde{I}$ the dyadic interval such that $|\tilde{I}| = |I|$ and $|I|^{-1}\mathrm{diam}(I \cup \tilde{I}) = n$. Then, for all $k \geq 0$

$$\tilde{S}_{k,n}(f)(x) = \Big(\sum_{I,J \in \mathcal{K}_{k,n}} \frac{\langle f, \psi_I \rangle^2}{|J|} \chi_J(x) \Big)^{1/2}$$

$$= \Big(\sum_{I} \langle f, \psi_I \rangle^2 \sum_{J \in I_{k,n}} \frac{\chi_J(x)}{|J|} \Big)^{1/2} = \Big(\sum_{I} \langle f, \psi_I \rangle^2 \frac{\chi_{\tilde{I}}(x)}{2^{-k}|\tilde{I}|} \Big)^{1/2} = 2^{k/2} S_{0,n}(f)(x)$$

Meanwhile, when $k \leq 0$ we have

$$\tilde{S}_{k,n}(f)(x) = \Big(\sum_{I,J \in \mathcal{K}_{k,n}} \frac{\langle f, \psi_I \rangle^2}{|J|} \chi_J(x) \Big)^{1/2}$$

$$= \Big(\sum_J \Big(\sum_{I\in I_{k,n}} \langle f,\psi_I\rangle^2\Big)\frac{\chi_J(x)}{|J|}\Big)^{1/2} \leq \Big(\sum_J 2^k\langle f,\psi_{I_J}\rangle^2\frac{\chi_J(x)}{|J|}\Big)^{1/2} = 2^{k/2}S_{k,n}(f)(x)$$

Proposition 9.3. *For every $1<p<\infty$, we have that if $k\geq 0$*

$$\|S_{k,n}f\|_p \leq C_p(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_p$$

while if $k\leq 0$

$$\|S_{k,n}f\|_p \leq C_p(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}+\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_p$$

with constants C_p independent of f, k and n.

Remark 9.1. *Before starting with the proof, we notice that a carefull read of it reveals that, by means of vector-valued interpolation,the result also holds for vector-valued modified square function with values in a Banach space X with the UMD property of the form*

$$S^X_{k,n}(f)(x) = \Big(\sum_I \frac{\|\langle f,\psi_I\rangle\|_X^2}{|J|}\chi_J(x)\Big)^{1/2}$$

for which every $1<p<\infty$, we have that if $k\geq 0$

$$\|S_{k,n}f\|_{L^p(X)} \leq C_{p,X}(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_{L^p(X)}$$

while if $k\leq 0$

$$\|S_{k,n}f\|_{L^p(X)} \leq C_{p,X}(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}+\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_{L^p(X)}$$

Then, in particular for $X=L^p(\mathbb{R})$ we get for $k\geq 0$

$$\|S_{k,n}f\|_{L^p(\mathbb{R}^2)} \leq C_p(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_{L^p(\mathbb{R}^2)}$$

while if $k\leq 0$

$$\|S_{k,n}f\|_{L^p(\mathbb{R}^2)} \leq C_p(2^{-k\,\mathrm{sign}(\frac{2}{p}-1)}+\log(n)+1)^{|\frac{2}{p}-1|}\|f\|_{L^p(\mathbb{R}^2)}$$

The estimate for $p=2$ is a trivial consequence of Plancherel's inequality. To extend the result to other exponents p we plan to use interpolation and duality. So, we first prove the following weak L^1 type estimate whose proof comes from a slight modification of the one appearing in [27].

Proposition 9.4. *If f is integrable and $\lambda>0$, then we have*

$$\|\{x : S_{k,n}f(x)>\lambda\}\| \leq C(2^{-k}n+1)\|f\|_1\lambda^{-1}$$

with a constant C independent of f and λ.

Proof. Consider the collection $\mathcal{I}$ of maximal dyadic intervals I with respect to set inclusion such that

$$|I|^{-1}\int_I |f(x)|dx > \lambda$$

Let E be the union of all $I \in \mathcal{I}$, the set that contains all intervals where f has large average. The intervals in $\mathcal{I}$ are pairwise disjoint and so

$$|E| \leq \sum_{I\in\mathcal{I}} |I| \leq \lambda^{-1} \sum_{I\in\mathcal{I}} \int_I |f(x)|dx \leq \|f\|_1 \lambda^{-1}$$

We take a classical Calderon-Zymund decomposition $f = g + b$ given by

$$g = \sum_I m_I(f)\chi_I + f\chi_{E^c} \qquad b = \sum_I f_I$$

with $m_I(f) = |I|^{-1} \int_I f$ and $f_I = (f - m_I(f))\chi_I$.

We see that g is essentially bounded by 2λ. Outside E, this follows by Lebesgue's differentiation theorem. To prove this inside E, it sufices to consider each interval $I \in \mathcal{I}$ separately. Let I be such an interval and $\tilde{I}$ its parent interval. Then by maximality of I we have

$$\int_I f(x)dx \leq \int_I |f|dx \leq \int_{\tilde{I}} |f|dx \leq \lambda|\tilde{I}| = 2\lambda|I|$$

Moreover it is also clear that

$$\int |g| \leq \int |f|$$

Because of the L^2 boundedness of $S_{k,n}$ we have

$$\|S_{k,n}g\|_2^2 \leq C\|g\|_2^2 \leq C \int |g|\lambda dx \leq C\lambda\|f\|_1$$

and so

$$\{S_{k,n}g > \lambda/2\} \leq C\|S_{k,n}g\|_2 \lambda^{-2} \leq C\|f\|_1\lambda^{-1}$$

We plan to prove the same estimate for $b = \sum_I f_I$. To do so, we define $\tilde{E}$ as the union of all $3I$ with $I \in \mathcal{I}$. We also define F as the union of all J such that the corresponding I satisfies $I \subset I'$ for some $I' \in \mathcal{I}$ and $\tilde{F}$ as the union of all $3J$ with $J \subset F$. Then,

$$|\{S_{k,n}b > \lambda/2\}| \leq |\tilde{F}| + \lambda^{-1}\|S_{k,n}b\|_{L^1(\mathbb{R}\setminus\tilde{F})}$$

Now we measure $\tilde{F}$ by means of a geometric argument that distinguishes between large and small scales. Since $|\tilde{F}| \leq 3|F|$ and

$$F = \bigcup_{I'\in\mathcal{I}} \cup\{J : I \subset I'\}$$

we fix now $I' \in \mathcal{I}$.

We also separate between $k \geq 0$ and $k \leq 0$ since the separation in scales is slightly different. We first assume $k \geq 0$ for which we separate into two different scales: smaller and larger than $\log(n)$.

The family of dyadic intervals $I \subset I'$ such that $|I| = 2^{-r}|I'|$ with $0 \leq r \leq \log(n)$ has the property that the corresponding intervals J satisfy

$$\mathrm{diam}(I' \cup J) \leq \mathrm{diam}(I \cup J) \geq n|I| = n2^{-r}|I'| > |I'|$$

and so the intervals J are disjoint with I'. Moreover, their union measures at most $2^{-k}|I'|$ as we see: the intervals J are pairwise disjoint and so for every $0\leq r\leq \log(n)$ we have

$$|\bigcup\{J:I\subset I',|I|=2^{-r}|I'|\}|=\sum_{\substack{J:I\subset I'\\ |I|=2^{-r}|I'|}}|J|=2^{-k}\sum_{\substack{I:I\subset I'\\ |I|=2^{-r}|I'|}}|I|=2^{-k}|I'|$$

Then

$$|\bigcup_{r=0}^{\log(n)}\bigcup\{J:I\subset I',|I|=2^{-r}|I'|\}|\leq 2^{-k}\log(n)|I'|$$

On the other hand for smaller scales, that is, intervals $I\subset I'$ such that $|I|=2^{-r}|I'|$ with $r>\log(n)$, we have that the corresponding intervals J satisfy

$$\mathrm{diam}(I'\cup J)\leq |I'|/2+|c(I')-c(J)|+|J|/2\leq |I'|/2+|c(I')-c(I)|+|c(I)-c(J)|+|J|/2$$
$$\leq |I'|+\mathrm{diam}(I\cup J)\leq |I'|+(n+1)|I|<|I'|+2n2^{-r}|I'|<3|I'|$$

Then the intervals J are included in $3I'$ and so

$$|\bigcup_{r>\log(n)}\bigcup\{J:I\subset I',|I|=2^{-r}|I'|\}|\leq C|I'|$$

Both things show that

$$|\cup\{J:I\subset I'\}|\leq C(2^{-k}\log(n)+1)|I'|$$

and therefore

$$|\tilde{F}|\lesssim 3\sum_{I'\in\mathcal{I}}|\cup\{J:I\subset I'\}|$$
$$\leq C(2^{-k}\log(n)+1)\sum_{I\in\mathcal{I}}|I'|\leq C(2^{-k}\log(n)+1)\|f\|_1\lambda^{-1}$$

When $k\leq 0$, the computations are similar but we separate into three different scales: smaller than $-k$, between $-k$ and $-k+\log(n)$, and larger than $-k+\log(n)$.

The subfamily of dyadic intervals $I\subset I'$ such that $|I|=2^{-r}|I'|$ with $0\leq r\leq -k$ has the property that the corresponding intervals J satisfy

$$\mathrm{diam}(I'\cup J)\geq \mathrm{diam}(I\cup J)\geq n|J|=n2^{-k}|I|=n2^{-k-r}|I'|>|I'|$$

and so they are disjoint with I'. Moreover, their union measures at most $2^{-k}|I'|2^{-r}$ as we see: for all intervals I considered, their correspoding intervals J satisfy $|J|=2^{-k-r}|I'|>|I'|$ and so there is a unique interval J corresponding with the different I. This way for every $0\leq r\leq -k$ we have

$$|\bigcup\{J:I\subset I',|I|=2^{-r}|I'|\}|\leq |J|=2^{-k}|I|=2^{-k-r}|I'|$$

and then summing a geometric series we have

$$|\bigcup_{r=0}^{-k}\bigcup\{J:I\subset I',|I|=2^{-r}|I'|\}|\leq C2^{-k}|I'|$$

On the other hand, the subfamily of dyadic intervals $I \subset I'$ such that $|I| = 2^{-r}|I'|$ with $-k \leq r \leq -k + \log(n)$ has the property that the corresponding intervals J satisfy $\mathrm{diam}(I \cup J) \geq n2^{-k-r}|I'| > |I'|$ and so they are still disjoint with I'. Moreover, their union measures at most $|I'|$ as we see: now $|J| = 2^{-k-r}|I'| < |I'|$ and on varying I there are $|I'|/|J| = 2^{-k-r}$ different disjoint intervals J whose union measures exactly $|I'|$. Then for every $-k \leq r \leq k + \log(n)$ we have

$$|\bigcup\{J : I \subset I', |I| = 2^{-r}|I'|\}| \leq |I'|$$

For different r, the intervals J are contained in a different translation of I'. Then the intervals J are pairwise disjoint and so

$$|\bigcup_{r=-k}^{-k+\log(n)} \bigcup\{J : I \subset I', |I| = 2^{-r}|I'|\}| \leq \log(n)|I'|$$

Finally, for smaller scales, that is, intervals $I \subset I'$ such that $|I| = 2^{-r}|I'|$ with $r > -k + \log(n)$, we have that the corresponding intervals J satisfy

$$\mathrm{diam}(I' \cup J) \leq |I'| + \mathrm{diam}(I \cup J) \leq (n+1)|J| < |I'| + 2n2^{-k-r}|I'| < 3|I'|$$

and so they are included in $3I'$ and so

$$|\bigcup_{r>-k+\log(n)} \bigcup\{J : I \subset I', |I| = 2^{-r}|I'|\}| \leq C|I'|$$

The three bounds together show that

$$|\cup\{J : I \subset I'\}| \leq C(2^{-k} + \log(n) + 1)|I'|$$

and so

$$|\tilde{F}| \lesssim 3 \sum_{I' \in \mathcal{I}} |\cup\{J : I \subset I'\}|$$

$$\leq C(2^{-k} + \log(n) + 1) \sum_{I \in \mathcal{I}} |I'| \leq C(2^{-k} + \log(n) + 1)\|f\|_1 \lambda^{-1}$$

The following step of the proof is to show

$$\|S_{k,n} b\|_{L^1(\mathbb{R}\setminus\tilde{F})} \leq C2^{-k/2}\|f\|_1$$

and, by sublinearity, it suffices to prove

$$\|S_{k,n} f_{I'}\|_{L^1(\mathbb{R}\setminus\tilde{F})} \leq C2^{-k/2}\lambda|I'|$$

for each $I' \in \mathcal{I}$. This in turn follows from

$$\left\| \left(\sum_{J \not\subset \tilde{F}} \frac{|\langle f_{I'}, \psi_I\rangle|^2}{|J|} \chi_J \right)^{1/2} \right\|_1 \leq \left\| \sum_{J \not\subset \tilde{F}} \frac{|\langle f_{I'}, \psi_I\rangle|}{|J|^{1/2}} \chi_J \right\|_1$$

$$\leq \sum_{J \not\subset \tilde{F}} |\langle f_{I'}, \psi_I\rangle||J|^{1/2} \leq 2^{-k/2} \sum_{I \not\subset \tilde{E}} |\langle f_{I'}, \psi_I\rangle||I|^{1/2}$$

$$\leq C2^{-k/2}\|f_{I'}\|_1 \leq C2^{-k/2}\lambda|I'|$$

where we have used that $J \not\subset \tilde{F}$ implies by definition of $\tilde{F}$ that the corresponding interval I satisfies $I \not\subset \tilde{E}$. Moreover, the last inequality follows from lemma 9.5 below.

Finally then,

$$|\{S_{k,n}b > \lambda/2\}| \leq C(2^{-k}n+1)\|f\|_1\lambda^{-1} + C2^{-k/2}\|f\|_1\lambda^{-1}$$

which ends the proof since $2^{-k/2} \leq 2^{-k} + \log(n) + 1 \leq 2^{-k}\log(n) + 1$ for all $k \in \mathbb{Z}$ and $n \in \mathbb{N}$.

The following lemma is the technical result needed to prove proposition 9.4. For the sake of completedness we include its proof although is exactly the same one that can be found in [27].

Lemma 9.5. *Let I' be some interval and f be an integrable function supported in I' with mean zero. For each dyadic interval I let ϕ_I be a bump function adapted to I. Then*

$$\sum_{I:I\not\subset 3I'} |\langle f,\phi_I\rangle||I|^{1/2} \leq C\|f\|_1$$

Here $3I'$ denotes the interval that shares the center with I' and is of length $3|I'|$.

Proof. We first consider the sum over all dyadic intervals I such that $I \not\subset 3I'$ with $|I| < |I'|$. Let c be the midpoint between $c(I)$ and $c(I')$. By symmetry we may assume that $\operatorname{supp} f \subset (-\infty, c)$ and then,

$$\begin{aligned}|\langle f,\phi_I\rangle| &\leq \|f\|_{L^1(-\infty,c)}\|\phi_I\|_{L^\infty(-\infty,c)} + \|f\|_{L^\infty(c,\infty)}\|\phi_I\|_{L^1(c,\infty)} \\ &\leq \|f\|_1 C|I|^{-1/2}\Big(1+\frac{|c(I)-c(I')|}{|I|}\Big)^{-N}\end{aligned}$$

which gives

$$|\langle f,\phi_I\rangle||I|^{1/2} \leq C\|f\|_1\Big(1+\frac{|c(I)-c(I')|}{|I|}\Big)^{-N}$$

For any two integers $k > 0$ and $m > 0$ there are at most two intervals I such that $|I'|/|I| = 2^k$ and the integer part of $1 + \frac{|c(I)-c(I')|}{|I|}$ is m. If $m < 2^k$, there are no such intervals which satisfy $I \not\subset 3I'$. Thus we can estimate

$$\sum_{\substack{I\,:\,|I|<|I'| \\ I\not\subset 3I'}} |\langle f,\phi_I\rangle||I|^{1/2} \leq C\|f\|_1\sum_{k>0}\sum_{m\geq 2^k} m^{-N} \leq C\|f\|_1$$

We now consider the sum over all dyadic I with $I \not\subset 3I'$ and $|I| \geq |I'|$. Let D denote the operator of differentiation and D^{-1} the antiderivative operator

$$D^{-1}f(x) = \int_{-\infty}^{x} f(y)dy$$

Notice that because of the mean zero of f, the support of $D^{-1}f$ is also included in I'. Then, by partial integration and the fact that $|I|D\phi_I$ is a bump function adapted to

I, we have

$$|\langle f,\phi_I\rangle| = |I|^{-1}|\langle D^{-1}f, |I|D\phi_I\rangle|$$
$$\leq |I|^{-1}(\|D^{-1}f\|_{L^1(-\infty,c)}\||I|D\phi_I\|_{L^\infty(-\infty,c)} + \|D^{-1}f\|_{L^\infty(c,\infty)}\||I|D\phi_I\|_{L^1(c,\infty)})$$
$$\leq |I|^{-1}\|D^{-1}f\|_1 C|I|^{-1/2}\Big(1+\frac{|c(I)-c(I')|}{|I|}\Big)^{1-N}$$

Now, from $\|D^{-1}f\|_1 \leq |I'|\|f\|_1$ we obtain

$$|\langle f,\phi_I\rangle||I|^{1/2} \leq C\|f\|_1\frac{|I'|}{|I|}\Big(1+\frac{|c(I)-c(I')|}{|I|}\Big)^{1-N}$$

For any two integers $k\geq 0$ and $m>0$, there are at most two intervals such that $|I|/|I'|=2^k$ and the integer part of $1+\frac{|c(I)-c(I')|}{|I|}$ is m. Thus we can estimate

$$\sum_{\substack{I\,:\,|I|\geq|I'|\\ I\not\subset 3I'}}|\langle f,\phi_I\rangle||I|^{1/2}\leq C\|f\|_1\sum_{k\geq 0}\sum_{m\geq 1}2^{-k}m^{1-N}\leq C\|f\|_1$$

ending the proof of this lemma.

Once the weak L^1 type inequality is proved, by interpolation we obtain for $1<p\leq 2$

$$\|S_{k,n}(f)\|_p \leq C_p(2^{-k}n+1)^{\frac{2}{p}-1}\|f\|_p$$

In order to obtain boundedness for the remaining exponents $2\leq p<\infty$ we consider the following martingale operator

$$T_{k,n}(f)(x)=\sum_J\langle f,\psi_I\rangle\psi_J(x)$$

where I and J are given by the same relationship that in the definition of $S_{k,n}$. This operator trivially satisfies $T^*_{k,n}(f)=T_{-k,n}(f)$. Actually, the implicit index j does not change and so we may also write $T^*_{k,n,j}(f)=T_{-k,n,j}(f)$. Moreover, we have that the classical square function of $T_{k,n}(f)$ coincides with $S_{k,n}(f)$:

$$S(T_{k,n}(f))(x)=\Big(\sum_J\frac{\langle f,\psi_I\rangle^2}{|J|}\chi_J(x)\Big)^{1/2}=\Big(\sum_I\frac{\langle f,\psi_I\rangle^2}{|J|}\chi_J(x)\Big)^{1/2}=S_{k,n}(f)(x)$$

So, by properties of classical square function and the previous case with $1<p'\leq 2$, we have

$$\|S_{k,n}\|_{L^p\to L^p}=\|S(T_{k,n})\|_{L^p\to L^p}\approx\|T_{k,n}\|_{L^p\to L^p}$$
$$=\|T^*_{k,n}\|_{L^{p'}\to L^{p'}}=\|T_{-k,n}\|_{L^{p'}\to L^{p'}}\approx\|S_{-k,n}\|_{L^{p'}\to L^{p'}}$$
$$\leq C_{p'}(2^k\log(n)+1)^{(\frac{2}{p'}-1)}=C_p(2^k\log(n)+1)^{|\frac{2}{p}-1|}$$

or

$$\|S_{k,n}\|_{L^p\to L^p}\approx\|S_{-k,n}\|_{L^{p'}\to L^{p'}}$$
$$\leq C_{p'}(2^k+\log(n)+1)^{(\frac{2}{p'}-1)}=C_p(2^k+\log(n)+1)^{|\frac{2}{p}-1|}$$

This ends the proof.

References

[1] Bernicot F., *Fiber-wise Calderón-Zygmund decomposition and application to a bi-dimensional paraproduct*, preprint

[2] Carleson L., *A counterexample for measures bounded on H^p spaces for the bidisk*, Mittag-Leffler Rep., 7, pp. 177-202 [1974]

[3] Coifman R.R., Meyer Y. *Commutateurs d' integrales singulières et opérateurs multilinéaires*, Ann. Inst. Fourier (Grenoble),bf 28, pp.177-202 [1978]

[4] Coifman R.R., Meyer Y. *Fourier analysis of multilinear convolutions, Calderón's theorem and analysis of Lipschitz curves*, Euclidean harmonic analysis (Proc. Sem. Univ. Maryland, College Park, Md.), pp.104-122, Lecture Notes in Math., 779 [1979]

[5] Coifman R.R. and Meyer Y. *Ondelettes and opérateurs III, Operatéurs multilinéaires*, Actualités Mathematiques, Hermman, Paris [1991]

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

[7] Christ M., *Lectures on Singular Integral Operators*, Cbms Regional Conference Series in Mathematics, 77 Amer. Math. Soc., Providence, RI, 1990.

[8] David G. and Journé J.L. *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math., Vol. 120, pp. 371-397, [1984]

[9] David G., Journé J.L. and Semmes S. *Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation*, Rev. Mat. Iberoamer. 1:1-56 [1985]

[10] Demeter C., Thiele C., *On the two dimensional Bilinear Hilbert Transform*, Amer. J. Math. 132 (2010), no. 1, 201–256

[11] Do Y., Muscalu C., Thiele C. *Variational estimates for paraproducts* to appear in Rev. Mat. Iberoamericana

[12] Duoandikoetxea J., *Análisis de Fourier.* , Addison-Wesley/Univ. Aut. de Madrid, [1995]

[13] Fefferman R., *Singular Integrals on Product H^p Spaces*, Revista Matemática Iberoamericana, Vol. 1, n. 2, pp. 25-31, [1985]

[14] Fefferman R., *Fourier Analysis in Several Parameters*, Revista Matemática Iberoamericana, Vol. 2, n. 1-2, pp. 89-98, [1986]

[15] Fefferman R., *Calderón-Zygmund theory for product domains: H^p spaces*, Proc. Natl. Acad. Sci., Vol. 83, pp. 840-843, [1986]

[16] Fefferman R., *Harmonic Analysis on product spaces*, Annals of Mathematics, Vol. 7, pp. 109-130, [1987]

[17] Fefferman R., Stein E. *Singular Integrals on Product Spaces*, Adv. in Math., Vol. 45, n. 2, pp. 117-143, [1982]

[18] Ferguson S., Lacey M., *A characterization of product* BMO *by commutators*, Acta Math., Vol. 189, pp. 143-160, [2002]

[19] Grafakos L., Torres, R., *Multilinear Calderón-Zygmund theory* Adv. in Math., Vol. 165, pp. 124-164, [2002]

[20] Journé J. L., *Calderón-Zygmund Operators on Product Spaces*, Revista Matemática Iberoamericana, Vol **1**, n. 3 [1985], 55-91

[21] Journé J. L., *A covering lemma for product spaces*, Proc. Amer. Math. Soc., Vol. 96, n. 4 [1986], p. 593-598

[22] Journé J. L., *Two problems of Calderón-Zygmund theory on product spaces*, Annales de l'institut Fourier, Vol **38**, n. 1 [1988], p. 111-132

[23] Meyer Y., *Wavelets and operators*, Cambridge Studies In Advanced Mathematics, 37, Cambridge University Press

[24] Muscalu C., Pipher J., Tao T. and Thiele C., *Bi-parameter paraproducts*, Acta Math. 193 (2004), 269-296

[25] Muscalu C., Pipher J., Tao T. and Thiele C., *Multi-parameter paraproducts*, Revista Mat. Iber. 22 (2006), 963,976
[26] Lacey M., Metcalfe J., *Paraproducts in one and several parameters*, Forum Math. 19 (2007), 325–351
[27] Thiele C., *Wave packet analysis*, CBMS **105**, [2006].
[28] Stein E. M., *Harmonic Analysis: real-variable methods, orthogonality and oscilatory integrals*, Princeton Univ. Press, [1993].

Centre for Mathematical Sciences, University of Lund, Lund, Sweden
E-mail address: s.pott@maths.lth.se

Department of Mathematics, University of Glasgow, Glasgow, UK
E-mail address: paco.villarroya@maths.gla.uk