SOME NEW WEIGHTED ESTIMATES ON PRODUCT SPACES

KANGWEI LI, HENRI MARTIKAINEN, AND EMIL VUORINEN

ABSTRACT. We complete our theory of weighted $L^p(w_1) \times L^q(w_2) \to L^r(w_1^{1/r}w_2^{1/q})$ estimates for bilinear bi-parameter Calderón–Zygmund operators under the assumption that $w_1 \in A_p$ and $w_2 \in A_q$ are bi-parameter weights. This is done by lifting a previous restriction on the class of singular integrals by extending a classical result of Muckenhoupt and Wheeden to the product BMO setting.

1. INTRODUCTION

Singular integral operators (SIOs) are operators of the form

$$T f(x) = \int_{\mathbb{R}^n} K(x, y) f(y) \, dy.$$  

They include many important linear transformations that arise in the analysis connecting geometric measure theory, partial differential equations, harmonic analysis and functional analysis. Classical one-parameter kernels are singular when $x = y$. Product space theory (multi-parameter theory), on the other hand, is concerned with kernels whose singularities are more spread out. To get an idea, for $x, y \in \mathbb{C} = \mathbb{R} \times \mathbb{R}$, compare the one-parameter Beurling kernel $1/(x - y)^2$ with the bi-parameter kernel $1/[(x_1 - y_1)(x_2 - y_2)]$ – the product of Hilbert kernels in both coordinate directions.

In the product setting the multilinear theory of SIOs – where a singular integral acts on a tuple of functions $(f_1, \ldots, f_m)$ – is historically significantly more limited than in the one-parameter setting. We refer to our recent work [14] for a thorough background on the subject, its significance and for new developments. In [14] we developed a wide general theory of bilinear bi-parameter singular integrals $T$. This includes general Calderón–Zygmund style principles: simpler estimates in the Banach range ($r > 1$), imply boundedness in the full bilinear range $L^p \times L^q \to L^r$, $1/p + 1/q = 1/r$, $1 < p, q \leq \infty$, $1/2 < r < \infty$, weighted estimates, mixed-norm estimates, and so on.

This was the first time such completely general principles could be formulated. However, our weighted estimates (see Section 2 for the definition of $A_p$ weights) still had the restriction that we needed the cancellation $T(1, 1) = 0$, and the same for the adjoints and partial adjoints. In this paper we remove this final restriction. Weighted estimates have significant independent interest, but they are also of fundamental use in obtaining other estimates, like vector-valued and mixed-norm estimates. This is due to the very
powerful bilinear extrapolation results – see e.g. \[3, 5, 11, 12, 17\]. In the product setting this viewpoint is particularly useful as many of the classical one-parameter tools are crudely missing. This is how we e.g. obtained the mixed-norm estimates in \[14\], and these estimates thus had the same restriction as the weighted estimates. With the understanding that a Calderón–Zygmund operator (CZO) is an SIO satisfying natural $T1$ type assumptions, our improved result now reads:

1.1. **Theorem.** Let $T$ be a bilinear bi-parameter CZO as defined in \[14\]. Then we have the weighted estimate

$$\|T(f, g)\|_{L^r(w)} \leq C([w_1]_{A_p}, [w_2]_{A_q}) \|f\|_{L^p(w_1)} \|g\|_{L^q(w_2)}$$

for all $1 < p, q < \infty$ and $1/2 < r < \infty$ with $1/r = 1/p + 1/q$, and for all bi-parameter weights $w_1 \in A_p(\mathbb{R}^n \times \mathbb{R}^m)$, $w_2 \in A_q(\mathbb{R}^n \times \mathbb{R}^m)$ with $w = w_1^{r/p} w_2^{r/q}$. In the unweighted case we also have the mixed-norm estimates

$$\|T(f, g)\|_{L^r(\mathbb{R}^n; L^2(\mathbb{R}^m))} \lesssim \|f\|_{L^{p_1}(\mathbb{R}^n; L^{p_2}(\mathbb{R}^m))} \|g\|_{L^{q_1}(\mathbb{R}^n; L^{q_2}(\mathbb{R}^m))}$$

for all $1 < p_i, q_i \leq \infty$ and $1/2 < r_i < \infty$ with $1/p_i + 1/q_i = 1/r_i$, except that if $r_2 < 1$ we have to assume $\infty \not\in \{p_1, q_1\}$.

Bilinear weights pose a problem with duality: notice e.g. that if $w_1, w_2 \in A_1$ then $w := w_1^{1/2} w_2^{1/2} \in A_4$, while we need to work in $L^2(w)$. This is a relevant problem and often makes bilinear bi-parameter weighted estimates different and harder than in the linear case. For the linear estimates see Fefferman–Stein \[9\] and Fefferman \[6, 7\], and the much more recent Holmes–Petermichl–Wick \[10\] that is rooted on the modern dyadic–probabilistic methods \[15\]. In particular, already some linear paraproduct estimates depend on suitable $H^1$-BMO type duality arguments, and for this reason we could not previously handle weighted estimates for certain model operators. This led to the restriction on the class of CZOs. We now remove this restriction by developing some new theory for the product BMO space of Chang and Fefferman \[11, 2\]. Theorem \[3, 2\] below is an extension of a classical result of Muckenhoupt and Wheeden \[16\] to the product BMO setting – it says that certain weighted product BMO spaces are actually the same as the unweighted product BMO space. This gives a useful way to construct objects in some genuinely weighted product BMO spaces by starting with an object in the unweighted product BMO. We prove this result of independent interest in its full generality, and apply a special case of it to prove our bilinear bi-parameter weighted estimates.

2. **Notations and preliminaries**

**Basic notation.** Throughout this paper $A \lesssim B$ means that $A \leq CB$ with some constant $C$ that we deem unimportant to track at that point. In particular, we often do not track the dependence on the weight constants. We write $A \sim B$ if $A \lesssim B \lesssim A$.

We work in the bi-parameter setting in the product space $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$. We write $x = (x_1, x_2)$ with $x_1 \in \mathbb{R}^n$ and $x_2 \in \mathbb{R}^m$ and for $f : \mathbb{R}^{n+m} \to \mathbb{C}$ and $h : \mathbb{R}^n \to \mathbb{C}$, we define

$$\langle f, h \rangle_1(x_2) := \int_{\mathbb{R}^n} f(x_1, x_2) h(x_1) \, dx_1.$$
Haar functions. We denote a dyadic grid in $\mathbb{R}^n$ by $D^n$ and a dyadic grid in $\mathbb{R}^m$ by $D^m$. We often write $D = D^n \times D^m$ for the related dyadic rectangles.

For an interval $I \subset \mathbb{R}$ we denote by $I_1$ and $I_2$ the left and right halves of $I$, respectively. We define $h_{I_1}^0 = |I|^{-1/2}1_I$ and $h_{I_2}^1 = |I|^{-1/2}(1_{I_1} - 1_{I_2})$. Let now $Q = I_1 \times \cdots \times I_n \in D^n$, and define the Haar function $h_Q^\eta, \eta = (\eta_1, \ldots, \eta_n) \in \{0, 1\}^n$, by setting

$$h_Q^\eta = h_{I_1}^{\eta_1} \otimes \cdots \otimes h_{I_n}^{\eta_n}.$$ 

If $\eta \neq 0$ the Haar function is cancellative: $\int h_Q^\eta = 0$. We exploit notation by suppressing the presence of $\eta$, and write $h_Q$ for some $h_Q^\eta, \eta \neq 0$. If $R = I \times J \in D = D^n \times D^m$, we define $h_R = h_I \otimes h_J$.

Weights. A weight $w(x_1, x_2)$ (i.e. a locally integrable a.e. positive function) belongs to the bi-parameter $A_p(\mathbb{R}^n \times \mathbb{R}^m), 1 < p < \infty$, if

$$[w]_{A_p(\mathbb{R}^n \times \mathbb{R}^m)} := \sup_R \frac{1}{|R|} \int_R w \left( \frac{1}{|R|} \int_R w^{1-p'} \right)^{p-1} < \infty,$$

where the supremum is taken over $R = I \times J$, where $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$ are cubes with sides parallel to the axes (we simply call such $R$ rectangles). We have

$$[w]_{A_p(\mathbb{R}^n \times \mathbb{R}^m)} < \infty \text{ iff } \max \left( \text{ess sup}_{x_1 \in \mathbb{R}^n} [w(x_1, \cdot)]_{A_p(\mathbb{R}^m)}, \text{ess sup}_{x_2 \in \mathbb{R}^m} [w(\cdot, x_2)]_{A_p(\mathbb{R}^n)} \right) < \infty,$$

and that $\max \left( \text{ess sup}_{x_1 \in \mathbb{R}^n} [w(x_1, \cdot)]_{A_p(\mathbb{R}^m)}, \text{ess sup}_{x_2 \in \mathbb{R}^m} [w(\cdot, x_2)]_{A_p(\mathbb{R}^n)} \right) \leq [w]_{A_p(\mathbb{R}^n \times \mathbb{R}^m)}$, while the constant $[w]_{A_p}$ is dominated by the maximum to some power. We say $w \in A_\infty(\mathbb{R}^n \times \mathbb{R}^m)$ if

$$[w]_{A_\infty(\mathbb{R}^n \times \mathbb{R}^m)} := \sup_R \frac{1}{|R|} \int_R w \exp \left( \frac{1}{|R|} \int_R \log w^{-1} \right) < \infty.$$

It is well-known (see e.g. [4] Section 7) that

$$A_\infty(\mathbb{R}^n \times \mathbb{R}^m) = \bigcup_{1 < p < \infty} A_p(\mathbb{R}^n \times \mathbb{R}^m).$$

We do not have any explicit use for the $A_\infty$ constant. The $w \in A_\infty$ assumption can always be replaced with the explicit assumption $w \in A_s$ for some $s \in (1, \infty)$, and then estimating everything with a dependence on $[w]_{A_s}$.

Of course, $A_p(\mathbb{R}^n)$ is defined similarly as $A_p(\mathbb{R}^n \times \mathbb{R}^m)$ – just take the supremum over cubes $Q$. A modern reference for the basic theory of bi-parameter weights is e.g. [10].

2.1. Square functions and maximal functions. Given $f : \mathbb{R}^{n+m} \to \mathbb{C}$ and $g : \mathbb{R}^n \to \mathbb{C}$ we denote the dyadic maximal functions by

$$M_D f := \sup_{I \in D} \frac{1}{|I|} \langle |g| \rangle_I \text{ and } M_D^w f := \sup_{R \in D} \frac{1}{|R|} \langle |f| \rangle_R,$$

where $\langle f \rangle_A = |A|^{-1} \int_A f$. We also set $M_{D^n} f(x_1, x_2) = M_{D^n} (f(\cdot, x_2))(x_1)$. The operator $M_{D^m}^w$ is defined similarly. The weighted maximal function is defined by

$$M_D^w f := \sup_{R \in D} \frac{1}{|R|} \langle |f| \rangle_R^w,$$

where $\langle |f| \rangle_R^w = w(R)^{-1} \int_R |f|w$. We require the following very nice result of Fefferman [8]. For a modern reference see [13] Proposition B.1. Notice that in the bi-parameter setting this result is very non-trivial as $w$ is not of tensor form.
2.1. Lemma. Let \( w \in A_\infty(\mathbb{R}^n \times \mathbb{R}^m) \). Then for all \( 1 < p \leq \infty \) we have
\[
\| M^n_D f \|_{L^p(w)} \lesssim \| f \|_{L^p(w)}.
\]

Now define the square functions
\[
S_D f = \left( \sum_{R \in D} |(f, h_R)|^{2} \frac{1}{|R|} \right)^{1/2}, \quad S_D^1 f = \left( \sum_{I \in D^n} \frac{1}{|I|} \otimes |(f, h_I)|^2 \right)^{1/2}
\]
and define \( S_{D,m}^1 f \) analogously. Define also
\[
S_{D,M}^1 f = \left( \sum_{I \in D^n} \frac{1}{|I|} \otimes [M_{D,m}(f, h_I)]^2 \right)^{1/2} \quad \text{and} \quad S_{D,M}^2 f = \left( \sum_{J \in D^m} [M_{D,n}(f, h_J)]^2 \otimes \frac{1}{|J|} \right)^{1/2}.
\]

We record the following most basic weighted estimates, which are used repeatedly (sometimes implicitly) below.

2.2. Lemma. For \( p \in (1, \infty) \) and \( w \in A_p = A_p(\mathbb{R}^n \times \mathbb{R}^m) \) we have the weighted square function estimates
\[
\| f \|_{L^p(w)} \sim \| S_D f \|_{L^p(w)} \sim \| S_{D,n}^1 f \|_{L^p(w)} \sim \| S_{D,m}^2 f \|_{L^p(w)}.
\]

Moreover, for \( p, s \in (1, \infty) \) we have the Fefferman–Stein inequality
\[
\| \left( \sum_j |Mf_j|^s \right)^{1/s} \|_{L^p(w)} \lesssim \| \left( \sum_j |f_j|^s \right)^{1/s} \|_{L^p(w)}.
\]

Here \( M \) can e.g. be \( M_{D,n}^1 \) or \( M_D \). Finally, we have
\[
\| S_{D,m}^1 f \|_{L^p(w)} + \| S_{D,m}^2 f \|_{L^p(w)} \lesssim \| f \|_{L^p(w)}.
\]

See [13, Lemma 2.1] for an indication on how to prove this standard result.

Of key importance to us is the following lower square function estimate valid for \( A_\infty \) weights.

2.3. Lemma. There holds
\[
\| f \|_{L^p(w)} \lesssim \| S_{D,n}^1 f \|_{L^p(w)}
\]
and
\[
\| f \|_{L^p(w)} \lesssim \| S_D f \|_{L^p(w)}
\]
for all \( p \in (0, \infty) \) and bi-parameter weights \( w \in A_\infty \).

See [14, Section 6] for an explanation of this well-known inequality. This is important to us as the weight \( w = w_1^{r/p} w_2^{r/q} \) in Theorem [14] is at least \( A_\infty \) – it is in fact \( A_{2r} \).

3. WEIGHTED BMO SPACES

Let \( D = D^n \times D^m \) be a lattice of dyadic rectangles, \( A = (a_R)_{R \in D} \) be a sequence of scalars and \( \Omega \subset \mathbb{R}^{n+m} \). We define
\[
S_A(x) = \left( \sum_{R \in D} |a_R|^2 \frac{1}{|R|} \right)^{1/2} \quad \text{and} \quad S_{A,\Omega}(x) = \left( \sum_{R \in D : R \subseteq \Omega} |a_R|^2 \frac{1}{|R|} \right)^{1/2}.
\]
Let \( p \in (0, \infty) \). Define
\[
\| A \|_{\text{BMO}_{\text{prod}}(p)} = \sup_{\Omega} \frac{1}{|\Omega|^{1/2}} \| S_{A,\Omega} \|_{L^p},
\]
where $\Omega$ is open and $0 < |\Omega| < \infty$. There are many possibilities how to define a weighted version. The following is not the ‘correct’ definition for many things. Nonetheless, it will be of key use to us. Thus, let $w \in A_\infty$ and set
\[
\|A\|_{\text{BMO}_{\prod,w}\,(p)} = \sup_{\Omega} \frac{1}{w(\Omega)^{1/2}} \|S_{A,\Omega}\|_{L^p(w)}.
\]
We set $\|A\|_{\text{BMO}_{\prod}} = \|A\|_{\text{BMO}_{\prod}(2)}$ and $\|A\|_{\text{BMO}_{\prod,w}} = \|A\|_{\text{BMO}_{\prod,w}(2)}$. The weight turns out not to play a role here - that is, we have $\|A\|_{\text{BMO}_{\prod}} = \|A\|_{\text{BMO}_{\prod,w}}$ for all bi-parameter weights $w \in A_\infty$. To prove this we need the bi-parameter John-Nirenberg. The unweighted version $\|A\|_{\text{BMO}_{\prod}} \sim \|A\|_{\text{BMO}_{\prod}(p)}$ is well-known. However, we need to know it in the following form, which is a priori stronger. The proof is similar, though, but requires the very non-trivial Lemma 2.1.

3.1. Proposition. For all bi-parameter weights $w \in A_\infty$ we have
\[
\|A\|_{\text{BMO}_{\prod,w}} \sim \|A\|_{\text{BMO}_{\prod,w}(p)}, \quad 0 < p < \infty.
\]

Proof. By Hölder’s inequality we only need to prove $\|A\|_{\text{BMO}_{\prod,w}(q)} \lesssim \|A\|_{\text{BMO}_{\prod,w}(p)}$ for $p < q$ and $q > 2$. We may assume $a_R \neq 0$ for only finitely many $R \in D$. We fix $\Omega$ and, for a large enough $N > 0$, denote $E := \{S_{A,\Omega} > N\|A\|_{\text{BMO}_{\prod,w}(p)}\}$. We now have
\[
w(E) \leq (N\|A\|_{\text{BMO}_{\prod,w}(p)})^{-p}\|S_{A,\Omega}\|_{L^p(w)}^p \leq N^{-p}w(\Omega).
\]

Split $D = R_1 \cup R_2$, where
\[
R_1 := \{R: w(E \cap R) > w(R)/2\}, \quad R_2 := \{R: w(E \cap R) \leq w(R)/2\}.
\]

Notice that clearly for all $R \in R_1$ we have
\[
R \subset \{M^w_D(1_E) > 1/2\} =: \tilde{E}.
\]

Since $w \in A_\infty$, by Lemma 2.1 we have
\[
w(\tilde{E}) \lesssim \|M^w_D(1_E)\|_{L^2(w)}^2 \lesssim w(E).
\]

We now fix $N$ so that we always have $w(\tilde{E}) \leq w(\Omega)/2^q$, and then notice that
\[
\left\|\left(\sum_{R \in \tilde{E}} |a_R|^2 \frac{1_{R}}{|R|}\right)^{1/2}\right\|_{L^q(w)} \leq \|S_{A,\tilde{E}}\|_{L^q(w)}
\]
\[
\leq \|A\|_{\text{BMO}_{\prod,w}(q)}w(\tilde{E})^{1/2} \leq \frac{1}{2}\|A\|_{\text{BMO}_{\prod,w}(q)}w(\Omega)^{1/2}.
\]

This is absorbable, so we now move on to consider the sum, where $R \in R_2$. As $q > 2$ we may calculate
\[
\left\|\left(\sum_{R \in \tilde{E}, R \subset \Omega} |a_R|^2 \frac{1_{R}}{|R|}\right)^{1/2}\right\|_{L^q(w)} \leq 2 \sup_{\|g\|_{L^{(q/2)}(w)} = 1} \sum_{R \in \tilde{E}, R \subset \Omega} |a_R|^2 (gw)_R
\]
\[
\leq 2 \sup_{\|g\|_{L^{(q/2)}(w)} = 1} \|1_E \cdot S_{A,\Omega} M^w_D g\|_{L^1(w)}
\]
\[
\leq 2 \sup_{\|g\|_{L^{(q/2)}(w)} = 1} \|1_E \cdot S_{A,\Omega}\|_{L^q(w)}^2 \|M^w_D g\|_{L^{(q/2)}(w)}.
\]
Then by Lemma 2.2 we have
\[ \|A\|_{\mathrm{BMO}_{\prod},w(p)}^2 w(\Omega)^{2/q}, \]
where we used Lemma 2.1 in the last step. The proof is done as we have shown that
\[ \|A\|_{\mathrm{BMO}_{\prod},w(q)} \leq \frac{1}{2} \|A\|_{\mathrm{BMO}_{\prod},w(q)} + C\|A\|_{\mathrm{BMO}_{\prod},w(p)}. \]

3.2. Theorem. For all bi-parameter weights \( w \in A_\infty \) we have
\[ \|A\|_{\mathrm{BMO}_{\prod}} \sim \|A\|_{\mathrm{BMO}_{\prod},w}. \]

Proof. Fix \( w \in A_\infty \). Then there exists \( s > 2 \) so that \( w \in A_s \). We first prove \( \|A\|_{\mathrm{BMO}_{\prod},w} \lesssim \|A\|_{\mathrm{BMO}_{\prod}} \). Define the linear bi-parameter paraproduct
\[ \Pi f = \Pi_A f = \sum_{R \in D} a_R \langle f \rangle_R R h_R. \]
It is well-known (see e.g. [10]) that
\[ \|\Pi f\|_{L^s(w)} \lesssim \|A\|_{\mathrm{BMO}_{\prod}} \|f\|_{L^s(w)}. \]

Then by Lemma 2.2 we have
\[ \left\| \left( \sum_{R \in D} |a_R|^2 \langle f \rangle_R^2 \frac{1}{|R|} \right)^{\frac{1}{2}} \right\|_{L^s(w)} = \|S_D(\Pi f)\|_{L^s(w)} \lesssim \|A\|_{\mathrm{BMO}_{\prod}} \|f\|_{L^s(w)}. \]
Testing with \( f = 1_\Omega \) we get
\[ \|S_{A,\Omega}\|_{L^s(w)} \leq \left\| \left( \sum_{R \in D} |a_R|^2 \langle 1_\Omega \rangle_R^2 \frac{1}{|R|} \right)^{\frac{1}{2}} \right\|_{L^s(w)} \lesssim \|A\|_{\mathrm{BMO}_{\prod},w(\Omega)^{\frac{1}{2}}}. \]
This means that \( \|A\|_{\mathrm{BMO}_{\prod},w(\Omega)^{\frac{1}{2}}} \lesssim \|A\|_{\mathrm{BMO}_{\prod}} \). By Proposition 3.1 we conclude that \( \|A\|_{\mathrm{BMO}_{\prod},w} \lesssim \|A\|_{\mathrm{BMO}_{\prod}} \).

It remains to prove \( \|A\|_{\mathrm{BMO}_{\prod}} \lesssim \|A\|_{\mathrm{BMO}_{\prod},w} \). For \( 0 \leq f \in L^s(w) \) and \( 0 \leq g \in L^{(s/2)'(w)} \), we have
\[ \sum_{R \in D} |a_R|^2 \langle f \rangle_R R \langle g \rangle_R^2 \leq \int_0^\infty \sum_{R \in D} \left( \langle f \rangle_R \langle g \rangle_R > t \right) |a_R|^2 \langle f \rangle_R R dt \]
\[ \leq \int_0^\infty \sum_{R \in D} \left( \langle f \rangle_R \langle g \rangle_R > t \right) |a_R|^2 \langle w \rangle_R dt \]
\[ \leq \|A\|_{\mathrm{BMO}_{\prod},w} \int_0^\infty w(\{ (M_D f)^2 M^w_D g > t \}) dt \]
\[ = \|A\|_{\mathrm{BMO}_{\prod},w} \| (M_D f)^2 M^w_D g \|_{L^1(w)} \]
\[ \leq \|A\|_{\mathrm{BMO}_{\prod},w} \| M_D f \|_{L^s(w)}^2 \| M^w_D g \|_{L^{(s/2)'}(w)} \]
\[ \lesssim \|A\|_{\mathrm{BMO}_{\prod},w} \| f \|_{L^s(w)}^2 \| g \|_{L^{(s/2)'}(w)} \].
where we have used Lemma 2.1 in the last step. Testing the above inequality with \( f = w^{-1/2} 1_{\Omega} \), \( g = w^{1/2} 1_{\Omega} \) we get
\[
\sum_{R \in \mathcal{D}} |a_R|^2 \langle w^{-1/2} \rangle_R^2 \langle w^{1/2} \rangle_R \lesssim \|A\|_{\text{BMO}_{\text{prod},w}}^2 |\Omega|.
\]

We conclude the proof by noticing that \( 1 \leq \langle w^{-1/2} \rangle_R^2 \langle w^{1/2} \rangle_R \).

3.3. Remark. In the one-parameter case, the equivalence between BMO and BMO\(w\), where \( w \in A_\infty \), is due to Muckenhoupt and Wheeden [16].

Finally, we define the actual weighted product BMO by setting
\[
\|A\|_{\text{BMO}_{\text{prod}}(w)} = \sup_{\Omega} \frac{1}{w(\Omega)^{1/2}} \left\| \sum_{R \in \mathcal{D}} |a_R|^2 \frac{1}{w(R)} \right\|_{L^2(\text{d}x)}^{1/2} = \sup_{\Omega} \left( \frac{1}{w(\Omega)} \sum_{R \in \mathcal{D}} |a_R|^2 \right)^{1/2}.
\]

The previous theorem is of independent interest, but also yields the following key lemma.

3.4. Lemma. If \( A \in \text{BMO}_{\text{prod}} \) define \( A_w = (a_R R)_{R \in \mathcal{D}} \) for \( w \in A_\infty \). Then we have
\[
\|A_w\|_{\text{BMO}_{\text{prod}}(w)} \sim \|A\|_{\text{BMO}_{\text{prod}}}.
\]

Proof. Notice that
\[
\|A_w\|_{\text{BMO}_{\text{prod}}(w)} = \|A\|_{\text{BMO}_{\text{prod},w}} \sim \|A\|_{\text{BMO}_{\text{prod}}}.
\]

Here the first equality is obvious and the second estimate is Theorem 3.2.

3.5. Corollary. For sequences of scalars \( A = (a_R) \) and \( B = (b_R) \) we have
\[
\sum_{R \in \mathcal{D}} |a_R| R b_R \lesssim \|A\|_{\text{BMO}_{\text{prod}}} \|S_B\|_{L^1(w)}
\]
whenever \( w \in A_\infty \).

Proof. Follows from the known, see e.g. [10], weighted \( H^1\)-BMO duality
\[
\sum_{R \in \mathcal{D}} |a_R| b_R \lesssim \|A\|_{\text{BMO}_{\text{prod}}(w)} \|S_B\|_{L^1(w)}
\]
and Lemma 3.4.

4. Proof of Theorem 1.1

A bilinear bi-parameter full paraproduct on a grid \( \mathcal{D} = \mathcal{D}^n \times \mathcal{D}^m \) has the form
\[
\Pi(f_1, f_2) = \Pi(f_1, f_2) = \sum_{R=1 \times I \in \mathcal{D}} a_R \langle f_1, h_I \otimes \frac{1}{|J|} \rangle \langle f_2 \rangle_R \frac{1}{|I|} \otimes h_J,
\]
where \( \|A\|_{\text{BMO}_{\text{prod}}} \leq 1 \). What is important is that there are actually nine different types of full paraproducts – the full paraproduct above corresponds to the tuples \( (h_I, \frac{1}{|I|}, \frac{1}{|J|}) \) and \( (\frac{1}{|J|}, \frac{1}{|I|}, h_J) \), but the \( h_J \) can be in any of the three slots and so can the \( h_I \).

It follows from [14] that to prove Theorem 1.1 it suffices to prove the following weighted estimate for the full paraproducts.
4.2. Proposition. Let $1 < p, q < \infty$ and $1/2 < r < \infty$ satisfy $1/p + 1/q = 1/r$, $w_1 \in A_p$ and $w_2 \in A_q$ be bi-parameter weights, and set $w := w_1^{r/p}w_2^{r/q}$. Then we have

$$\|\Pi(f_1, f_2)\|_{L^r(w)} \lesssim \|f_1\|_{L^p(w_1)}\|f_2\|_{L^q(w_2)}.$$  

Proof. Case 1. Suppose that there is a full average over $R = I \times J$ at least in $f_1$ or $f_2$. In such cases the bilinear paraproduct estimate decouples reducing to linear estimates. For example, suppose II has the form (4.1). Then using the weighted lower square function estimate, Lemma 2.2, and the basic Lemma 2.2 we have

$$\|\Pi(f_1, f_2)\|_{L^r(w)} \lesssim \left\| \left( \sum_I \left( \sum_J |a_R| \langle f_1, h_I \otimes \frac{1}{|I|} \rangle \langle f_2, 1_{J^c} \rangle \right)^2 \right)^{1/2} \right\|_{L^r(w)}$$

$$\leq \|M_Df_2\|_{L^r(w)} \left( \sum_I \left( \sum_J |a_R| \langle f_1, h_I \otimes \frac{1}{|I|} \rangle \langle f_2, 1_{J^c} \rangle \right)^2 \right)^{1/2} \|\Pi(f_1, f_2)\|_{L^r(w)}$$

$$\lesssim \|f_2\|_{L^q(w_2)} \|S^2_{D, M}h\|_{L^p(w_1)}.$$  

where

$$h = \sum_{R=I \times J} |a_R| \langle f_1, h_I \otimes \frac{1}{|I|} \rangle \frac{1}{|I|} \otimes h_J.$$  

This is just a standard linear bi-parameter paraproduct, and thus satisfies the weighted estimate $\|h\|_{L^p(w_1)} \lesssim \|f_1\|_{L^p(w_1)}$ (see e.g. [10]). Thus, we are done by Lemma 2.2.

Case 2. Out of the remaining cases we choose the symmetry

$$\Pi(f_1, f_2) = \sum_{R=I \times J} a_R \left( f_1, h_I \otimes \frac{1}{|I|} \right) \left( f_2, \frac{1}{|J|} \otimes h_J \right) \frac{1}{|R|}.$$  

Equipped with our current tools we can prove the desired estimate directly for any $p_0, q_0 \in (1, \infty)$ and $r_0 \in [1, \infty)$ satisfying $1/r_0 = 1/p_0 + 1/q_0$. By bilinear extrapolation [3, 5] it is enough to prove the estimate with only one fixed tuple, so this is certainly enough to get the claimed full range. For example, in the case $r_0 = 1$ we get

$$\|\Pi(f_1, f_2)\|_{L^1(w)} \leq \sum_{R=I \times J} |a_R| \langle w \rangle_R \left| \left( f_1, h_I \otimes \frac{1}{|I|} \right) \left( f_2, \frac{1}{|I|} \otimes h_J \right) \right|$$

$$\lesssim \left\| \left( \sum_R \left( \langle f_1, h_I \otimes \frac{1}{|I|} \rangle \langle f_2, \frac{1}{|I|} \otimes h_J \rangle \right)^2 \frac{1}{|I|} \right)^{1/2} \right\|_{L^1(w)}$$

$$\lesssim \|S^2_{D, M}f_1\|_{L^{p_0}(w_1)} \|S^2_{D, M}f_2\|_{L^{q_0}(w_2)} \lesssim \|f_1\|_{L^{p_0}(w_1)} \|f_2\|_{L^{q_0}(w_2)},$$  

where we have used Lemma 2.2 in the last step and Corollary 3.5 in the beginning.  

4.3. Remark. The advantage of the case $r_0 = 1$ in Case 2 above is that then $w \in A_2$, so that proving the required estimate $\|A\|_{BMO_{\prod, w}} \lesssim \|A\|_{BMO_{\prod, w}}$ does not require the John-Nirenberg inequality and thus not even Lemma 2.1. However, we still note that the case $r_0 > 1$ could be done with a similar calculation, but it requires bounding $|\langle \Pi(f_1, f_2), f_3w \rangle|$.
for \( f_3 \in L^\infty(w) \) with
\[
\sum_{R=1 \times J} |a_R| \langle w \rangle_R \left\langle f_1, h_I \otimes \frac{1}{|J|} \right\rangle \left\| f_2, \frac{1}{|I|} \otimes h_J \right\| \langle f_3 \rangle_R^w,
\]
using the fuller strength of Corollary 3.5 and also Lemma 2.1.

REFERENCES

[1] S.-Y. A. Chang, R. Fefferman, A continuous version of duality of \( H^1 \) with BMO on the Bidisc, Ann. of Math. 112 (1980) 179–201.
[2] S.-Y. A. Chang, R. Fefferman, Some recent developments in Fourier analysis and \( H^p \) theory on product domains, Bull. Amer. Math. Soc. 12 (1985) 1–43.
[3] J. Duoandikoetxea, Extrapolation of weights revisited: New proofs and sharp bounds, J. Funct. Anal. 260 (2011) 1886–1901.
[4] J. Duoandikoetxea, F. Martín-Reyes, S. Ombrosi, On the \( A_\infty \) conditions for general bases. Math. Z. 282 (2016) 955–972.
[5] L. Grafakos, J.M. Martell, Extrapolation of weighted norm inequalities for multivariable operators and applications, J. Geom. Anal. 14 (2004) 19–46.
[6] R. Fefferman, Harmonic analysis on product spaces, Ann. of Math. 126 (1987) 109–130.
[7] R. Fefferman, \( A^p \) weights and singular integrals, Amer. J. Math. 110 (1988) 975–987.
[8] R. Fefferman, Strong differentiation with respect to measures. Amer. J. Math. 103 (1981) 33–40.
[9] R. Fefferman, E. Stein, Singular integrals on product spaces, Adv. Math. 45 (1982) 117–143.
[10] I. Holmes, S. Petermichl, B. Wick, Weighted little bmo and two-weight inequalities for Journé commutators. Anal. PDE 11 (2018) 1693–1740.
[11] K. Li, J. M. Martell, S. Ombrosi, Extrapolation for multilinear Muckenhoupt classes and applications to the bilinear Hilbert transform, preprint, arXiv:1802.03338 2018.
[12] K. Li, J.M. Martell, H. Martikainen, S. Ombrosi, E. Vuorinen, End-point estimates, extrapolation for multilinear Muckenhoupt classes, and applications, preprint arXiv:1902.04951 2019.
[13] K. Li, H. Martikainen, E. Vuorinen, Bloom type inequality for bi-parameter singular integrals: efficient proof and iterated commutators, Int. Math. Res. Not. IMRN (2019), rnz072, https://doi.org/10.1093/imrn/rnz072
[14] K. Li, H. Martikainen, E. Vuorinen, Bilinear Calderón–Zygmund theory on product spaces, to appear in J. Math. Pures Appl., available at arXiv:1712.08133 2018.
[15] H. Martikainen, Representation of bi-parameter singular integrals by dyadic operators, Adv. Math. 229 (2012) 1734–1761.
[16] B. Muckenhoupt and R. L. Wheeden, Weighted bounded mean oscillation and the Hilbert transform, Studia Math. 54 (1975/76) 221–237.
[17] B. Nieraeth, Quantitative estimates and extrapolation for multilinear weight classes, Math. Ann. 375 (2019) 453–507.

(K.L.) CENTER FOR APPLIED MATHEMATICS, TIANJIN UNIVERSITY, WEIJIN ROAD 92, 300072 TIANJIN, CHINA
E-mail address: kangwei.nku@gmail.com

(H.M.) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND
E-mail address: henri.martikainen@helsinki.fi

(E.V.) DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, P.O.B. 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND
E-mail address: j.e.vuorin@gmail.com