# MATRIX WEIGHTED NORM INEQUALITIES FOR COMMUTATORS AND PARAPRODUCTS WITH MATRIX SYMBOLS 

JOSHUA ISRALOWITZ, HYUN KYOUNG KWON, AND SANDRA POTT


#### Abstract

Let $B$ be a locally integrable matrix function, $W$ a matrix $\mathrm{A}_{p}$ weight with $1<p<\infty$, and $T$ be any of the Riesz transforms. We will characterize the boundedness of the commutator $[T, B]$ on $L^{p}(W)$ in terms of the membership of $B$ in a natural matrix weighted BMO space. To do this, we will characterize the boundedness of dyadic paraproducts on $L^{p}(W)$ via a new matrix weighted Carleson embedding theorem. Finally, we will use some of the ideas from these proofs to (among other things) obtain quantitative weighted norm inequalities for these operators and also use them to prove sharp $L^{2}$ bounds for the Christ/Goldberg matrix weighted maximal function associated with matrix $\mathrm{A}_{2}$ weights.


## 1. Introduction

1.1. Motivation. The $L^{p}$ boundedness of commutators between functions and Calderón-Zygmund operators (or CZOs for short) is a classical subject that was first studied in [9] and has numerous applications to PDEs, operator theory, and complex analysis (see [9, 32] for a small sampling of these applications). Although numerous authors have subsequently used or extended the boundedness results in [9], and although weighted norm inequalities for CZOs have been extensively studied for the past 40 years or so (starting with the seminal work [16]), less attention has been paid towards weighted norm inequalities for commutators. It is well known, however, that the commutator $[T, b]$ is bounded on $L^{p}(w)$ (where $T$ is a CZO and $w$ is an $\mathrm{A}_{p}$ weight) if $b$ is in the classical John-Nirenberg BMO space. Furthermore, it is well known that the boundedness of $[T, b]$ on $L^{p}(w)$ implies that $b \in \mathrm{BMO}$ when $T$ is one of the Riesz transforms (see [1, 6] for example. Also see

[^0]the interesting preprints [14, 15] for a modern discussion and extensions of the results in [6]).

On the other hand, it is well known that proving matrix weighted norm inequalities for even CZOs is a very difficult task, and because of this, matrix weighted norm inequalities for certain CZOs have only recently been investigated (see [33, 34] for specific details of these difficulties). In particular, if $n$ and $d$ are natural numbers and if $W: \mathbb{R}^{d} \rightarrow$ $\mathcal{M}_{n}(\mathbb{C})$ is positive definite a. e. (where as usual $\mathcal{M}_{n}(\mathbb{C})$ is the algebra of $n \times n$ matrices with complex scalar entries), then define $L^{p}(W)$ for $1<p<\infty$ to be the space of measurable functions $\vec{f}: \mathbb{R}^{d} \rightarrow \mathbb{C}^{n}$ where

$$
\|\vec{f}\|_{L^{p}(W)}^{p}=\int_{\mathbb{R}^{d}}\left|W^{\frac{1}{p}}(x) \vec{f}(x)\right|^{p} d x<\infty .
$$

It was proved by F. Nazarov and S. Treil, M. Goldberg, and A. Volberg, respectively in [13, 27, 34 that certain CZOs acting componentwise on $\mathbb{C}^{n}$ valued functions are bounded on $L^{p}(W)$ when $1<p<\infty$ if $W$ is a matrix $\mathrm{A}_{p}$ weight, which means that

$$
\begin{equation*}
\|W\|_{\mathrm{A}_{p}}:=\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left(\frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x) W^{-\frac{1}{p}}(t)\right\|^{p^{\prime}} d t\right)^{\frac{p}{p^{\prime}}} d x<\infty \tag{1.1}
\end{equation*}
$$

where $p^{\prime}$ is the conjugate exponent of $p$.
Despite this, virtually nothing has been studied regarding matrix weighted norm inequalities for operators related to CZOs that themselves have matrix kernels (in the case of CZOs) or matrix symbols (in the case of operators such as commutators, paraproducts, or Haar multipliers). The purpose of this paper is to initiate such a study, and in particular, we will characterize matrix weighted norm inequalities for commutators $[T, B]$ when $W$ is a matrix $\mathrm{A}_{p}$ weight, $B$ is a locally integrable matrix function, and $T$ is any of the Riesz transforms (see also the first author's preprint [19] where the matrix weighted boundedness of certain matrix kernelled CZOs is investigated).
1.2. Reducing operators. We will need to briefly discuss a very important reformulation of the matrix $\mathrm{A}_{p}$ condition before we state our main results. Given any norm $\rho$ on $\mathbb{C}^{n}$, the classical "John's ellipsoid theorem" (see [7]) says that there exists a reducing operator $V$ (i.e. a positive definite $n \times n$ matrix $V$ ) where

$$
\rho(\vec{e}) \leq|V \vec{e}| \leq \sqrt{n} \rho(\vec{e})
$$

for all $\vec{e} \in \mathbb{C}^{n}$. Given a matrix weight $W$, a cube $I$, and some $1<$ $p<\infty$, let $V_{I}=V_{I}(W, p)$ and $V_{I}^{\prime}=V_{I}^{\prime}(W, p)$ be reducing operators corresponding to the norms
$\rho_{W, I, p}(\vec{e}):=\left(\frac{1}{|I|} \int_{I}\left|W^{\frac{1}{p}}(x) \vec{e}\right|^{p} d x\right)^{\frac{1}{p}}$ and $\rho_{W, I, p}^{*}(\vec{e}):=\left(\frac{1}{|I|} \int_{I}\left|W^{-\frac{1}{p}}(x) \vec{e}\right|^{p^{\prime}} d x\right)^{\frac{1}{p^{\prime}}}$
on $\mathbb{C}^{n}$, respectively. Note that while these reducing operators are not necessarily unique, the precise reducing operator being used will not be important. It will be important later, however, to notice that by definition we can take $V_{I}\left(W^{1-p^{\prime}}, p^{\prime}\right)=V_{I}^{\prime}(W, p)$ and $V_{I}^{\prime}\left(W^{1-p^{\prime}}, p^{\prime}\right)=$ $V_{I}(W, p)$.

Using these reducing operators and the equivalence of the canonical matrix norm and trace norm on $\mathcal{M}_{n}(\mathbb{C})$, we have for a matrix $\mathrm{A}_{p}$ weight $W$ that

$$
\|W\|_{\mathrm{A}_{p}}=\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left(\frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x) W^{-\frac{1}{p}}(t)\right\|^{p^{\prime}} d t\right)^{\frac{p}{p^{p}}} d x \approx \sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}}\left\|V_{I} V_{I}^{\prime}\right\|^{p}
$$

which also immediately gives us that $W$ is a matrix $\mathrm{A}_{p}$ weight if and only if $W^{1-p^{\prime}}$ is a matrix $\mathrm{A}_{p^{\prime}}$ weight. Furthermore, it is not difficult to see (using Hölder's inequality and some elementary arguments involving norms and dual norms, see [13] p. 4) that

$$
\begin{equation*}
\left|V_{I}^{\prime} V_{I} \vec{e}\right| \geq|\vec{e}| \tag{1.2}
\end{equation*}
$$

for any matrix (not necessarily matrix $\mathrm{A}_{p}$ ) weight $W$, any cube $I$, any $1<p<\infty$, and any $\vec{e} \in \mathbb{C}^{n}$.

Also when $p=2$, a very simple and direct computation shows that we may take $V_{I}=\left(m_{I} W\right)^{\frac{1}{2}}$ and $V_{I}^{\prime}=\left(m_{I}\left(W^{-1}\right)\right)^{\frac{1}{2}}$ where $m_{I} W$ is the average of $W$ on $I$. In particular, the matrix $\mathrm{A}_{2}$ condition takes on a particularly simple form that is very similar to the scalar $\mathrm{A}_{2}$ condition. Similarly when $W(x)=w(x) \operatorname{Id}_{n \times n}$ for a scalar $\mathrm{A}_{p}$ weight $w$ we can take $V_{I}=\left(m_{I} w\right)^{\frac{1}{p}} \operatorname{Id}_{n \times n}$ and $V_{I}^{\prime}=\left(m_{I} w^{1-p^{\prime}}\right)^{\frac{1}{p^{\prime}}} \operatorname{Id}_{n \times n}$.

Lastly, it will be useful later in the paper to examine the relationship between $V_{I}$ and $V_{\tilde{I}}$ where $I, \tilde{I}$ are cubes with $I \subseteq \tilde{I}$ and comparable side-lengths. In particular, for any $\vec{e} \in \mathbb{C}^{n}$ we have

$$
\begin{equation*}
\left|V_{I} \vec{e}\right|^{p} \approx \frac{1}{|I|} \int_{I}\left|W^{\frac{1}{p}}(x) \vec{e}\right|^{p} d x \lesssim \frac{1}{|\tilde{I}|} \int_{\tilde{I}}\left|W^{\frac{1}{p}}(x) \vec{e}\right|^{p} d x \lesssim\left|V_{\tilde{I}} \vec{e}\right|^{p} \tag{1.3}
\end{equation*}
$$

and a similar computation shows that

$$
\left|V_{I}^{\prime} \vec{e}\right|^{p^{\prime}} \lesssim\left|V_{\tilde{I}}^{\prime} \vec{e}\right|^{p^{\prime}}
$$

or equivalently

$$
\left|\left(V_{\tilde{I}}^{\prime}\right)^{-1} \vec{e}\right|^{p^{\prime}} \lesssim\left|\left(V_{I}^{\prime}\right)^{-1} \vec{e}\right|^{p^{\prime}}
$$

On the other hand, if $W$ is a matrix $\mathrm{A}_{p}$ weight then the inequality above combined with the the $\mathrm{A}_{p}$ condition gives us that

$$
\begin{equation*}
\left|V_{\tilde{I}} \vec{e}\right|^{p} \lesssim\left\|V_{\tilde{I}} V_{\tilde{I}}^{\prime}\right\|^{p}\left|\left(V_{\tilde{I}}^{\prime}\right)^{-1} \vec{e}\right|^{p} \leq\|W\|_{\mathrm{A}_{p}}\left|\left(V_{I}^{\prime}\right)^{-1} \vec{e}\right|^{p} \leq\|W\|_{\mathrm{A}_{p}}\left|V_{I} \overrightarrow{e^{p}}\right|^{p} \tag{1.4}
\end{equation*}
$$

where the last line follows from (1.2).
1.3. Notation and main results. Now if $1<p<\infty$ and $W$ is a matrix $\mathrm{A}_{p}$ weight, then let $\mathrm{BMO}_{W}^{p}$ be the space of locally integrable functions $B: \mathbb{R}^{d} \rightarrow \mathcal{M}_{n}(\mathbb{C})$ where

$$
\begin{cases}\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I} B\right) V_{I}^{-1}\right\|^{p} d x<\infty & : \text { if } 2 \leq p<\infty \\ \sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{p}}(x)\left(B^{*}(x)-m_{I} B^{*}\right)\left(V_{I}^{\prime}\right)^{-1}\right\|^{p^{\prime}} d x<\infty & : \text { if } 1<p \leq 2\end{cases}
$$

Also, given a dyadic grid $\mathscr{D}$, we will let $\mathrm{BMO}_{W, \mathscr{D}}^{p}$ denote the space of locally integrable $n \times n$ functions satisfying the condition above but where the supremum is taken over all $I \in \mathscr{D}$. Note that these two conditions should be thought of as dual to each other (in a precise sense that will be explained later in this introduction.) The main result of this paper is the following

Theorem 1.1. Let $1<p<\infty$. If $W$ is a matrix $A_{p}$ weight and $T$ is any of the Riesz transforms, then $[T, B]$ is bounded on $L^{p}(W)$ if and only if $B \in B M O_{W}^{p}$.

As is well known, the study of such commutators is often reduced to the study of paraproducts, and this is the approach we will take for proving Theorem 1.1. Before we define our paraproducts, let us review some definitions and notation regarding Haar functions in several variables. Following the notation in [25], for any dyadic grid in $\mathbb{R}$ and any interval in this grid, let

$$
h_{I}^{1}=|I|^{-\frac{1}{2}} \chi_{I}(x), \quad h_{I}^{0}(x)=|I|^{-\frac{1}{2}}\left(\chi_{I_{\ell}}(x)-\chi_{I_{r}}(x)\right) .
$$

Now given any dyadic grid $\mathscr{D}$ in $\mathbb{R}^{d}$, any cube $I=I_{1} \times \cdots \times I_{d}$, and any $\varepsilon \in\{0,1\}^{d}$, let $h_{I}^{\varepsilon}=\Pi_{i=1}^{d} h_{I_{i}}^{\varepsilon}$. It is then easily seen that $\left\{h_{I}^{\varepsilon}\right\}_{I \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}$ where $\operatorname{Sig}_{d}=\{0,1\}^{d} \backslash\{\overrightarrow{1}\}$ is an orthonormal basis for $L^{2}\left(\mathbb{R}^{d}\right)$. We will say $h_{I}^{\varepsilon}$ is "cancellative" if $\varepsilon \neq \overrightarrow{1}$ since in this case $\int_{I} h_{I}^{\varepsilon}=0$.

Now given a locally integrable function $B: \mathbb{R}^{d} \rightarrow \mathcal{M}_{n}(\mathbb{C})$, define the dyadic paraproduct $\pi_{B}$ with respect to a dyadic grid $\mathscr{D}$ by

$$
\begin{equation*}
\pi_{B} \vec{f}=\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}} B_{I}^{\varepsilon}\left(m_{I} \vec{f}\right) h_{I}^{\varepsilon} \tag{1.5}
\end{equation*}
$$

where $B_{I}^{\varepsilon}$ is the matrix of Haar coefficients of the entries of $B$ with respect to $I$ and $\varepsilon$, and $m_{I} \vec{f}$ is the vector of averages of the entries of $\vec{f}$. The proof of Theorem [1.1 will be largely based on the following

Theorem 1.2. Let $1<p<\infty$ and let $\mathscr{D}$ be a dyadic grid. If $W$ is a matrix $A_{p}$ weight then the paraproduct $\pi_{B}$ with respect to $\mathscr{D}$ is bounded on $L^{p}(W)$ if and only if $B \in B M O_{W, \mathscr{D}}^{p}$.

Note that the proofs of Theorems 1.1 and 1.2 actually give us quantitative bounds when $p=2$. In particular, we will prove that

$$
\begin{equation*}
\left\|\pi_{B}\right\|_{L^{2}(W) \rightarrow L^{2}(W)} \lesssim\left(\log \|W\|_{\mathrm{A}_{2}}\right)^{\frac{1}{2}}\|W\|_{\mathrm{A}_{2}}^{\frac{3}{2}}\|B\|_{*}^{\frac{1}{2}} \tag{1.6}
\end{equation*}
$$

where $\|B\|_{*}$ is the canonical supremum in condition (b) of Theorem 1.3 below. Moreover, we will prove that

$$
\begin{align*}
\|[T, B]\|_{L^{2}(W) \rightarrow L^{2}(W)} & \lesssim\|Q\|_{L^{2}(W) \rightarrow L^{2}(W)} \max \left\{\left\|\pi_{B}\right\|_{L^{2}(W) \rightarrow L^{2}(W)},\left\|\pi_{B^{*}}\right\|_{L^{2}\left(W^{-1}\right) \rightarrow L^{2}\left(W^{-1}\right)}\right. \\
& \left.+\|W\|_{\mathrm{A}_{2}}^{\frac{3}{2}} \log \|W\|_{\mathrm{A}_{2}}\|B\|_{*}^{\frac{1}{2}}\right\} \tag{1.7}
\end{align*}
$$

where $T$ is any of the Riesz transforms and $Q$ is a first order Haar shift (see Section 3.1 for the definition.) It would be very interesting to know if any similar commutator bounds for general scalar CZOs are true, and this will be explored in a forthcoming paper by the first and third authors.

Besides being extremely important for proving results regarding commutators (see [14, 15, 25] for example), note that paraproducts are central to the study of CZOs themselves since they allow one to decompose an arbitrary CZO $T$ as $T=\pi_{T 1}+\pi_{T^{*} 1}^{*}+R$ where $R$ is cancellative in the sense that $R 1=R^{*} 1=0$. In fact, the first author in [19] will employ Theorem 1.2 to prove a T1 theorem regarding the matrix weighted boundedness of certain matrix kernelled CZOs.

The proof of Theorem 1.2 will easily follow from the following matrix weighted Carleson embedding theorem (see the next section for details), which is obviously of independent interest itself. Here, for a dyadic grid $\mathscr{D}$ and $J \in \mathscr{D}$, we define $\mathscr{D}(J)=\{I \in \mathscr{D}: I \subseteq J\}$.

Theorem 1.3. Let $1<p<\infty$ and let $\mathscr{D}$ be a dyadic grid. If $W$ is a matrix $A_{p}$ weight and $A:=\left\{A_{I}^{\varepsilon}\right\}_{I \in \mathscr{D}, \varepsilon \in S i g_{d}}$ is a sequence of matrices, then the following are equivalent
(a) The operator $\Pi_{A}$ defined by

$$
\Pi_{A} \vec{f}:=\sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{D}} V_{I} A_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}} \vec{f}\right) h_{I}^{\varepsilon}
$$

is bounded on $L^{p}\left(\mathbb{R}^{d} ; \mathbb{C}^{n}\right)$.
(b)

$$
\sup _{J \in \mathscr{D}} \frac{1}{|J|} \sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{D}(J)}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|^{2}<\infty .
$$

(c) There exists $C>0$ independent of $J \in \mathscr{D}$ such that

$$
\frac{1}{|J|} \sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{D}(J)}\left(A_{I}^{\varepsilon}\right)^{*} V_{I}^{2} A_{I}^{\varepsilon}<C V_{J}^{2}
$$

if $2 \leq p<\infty$, and

$$
\frac{1}{|J|} \sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{D}(J)} A_{I}^{\varepsilon}\left(V_{I}^{\prime}\right)^{2}\left(A_{I}^{\varepsilon}\right)^{*}<C\left(V_{J}^{\prime}\right)^{2}
$$

if $1<p \leq 2$.
Furthermore, the operator norm in (a) and the canonical supremums in (b) and (c) are equivalent in the sense that they are independent of the sequence $A$. Finally, a matrix function $B \in B M O_{W, \mathscr{D}}^{p}$ if and only if the sequence of Haar coefficients of $B$ satisfies any of the above equivalent conditions.

Note that the constants in the equivalence between the operator norm in (a) and the canonical supremums in $(b)$ and $(c)$ of course depend on the $\mathrm{A}_{p}$ characteristic of $W$, and throughout the proof we will track precisely the nature of this dependence (modulo constants involved in the matrix weighted Triebel-Lizorkin imbedding theorem when $p \neq 2$, since in this case efficient bounds are not known, see Section 2.1 for more details). Also, note that (as to be expected), we have the following relationship between $\mathrm{BMO}_{W}^{p}$ and $\mathrm{BMO}_{W, \mathscr{D}}^{p}$

Proposition 1.4. There exists dyadic grids $\mathscr{D}^{t}$ for $t=1, \ldots, 2^{d}$ where

$$
B M O_{W}^{p}=\bigcup_{t=1}^{2^{d}} B M O_{W, \mathscr{D}}^{p}
$$

Note that while the two different cases for different $p$ in the definition of $\mathrm{BMO}_{W}^{p}$ might seem awkward, it will turn out that Theorem 1.1 and duality will together prove the following

Corollary 1.5. If $1<p<\infty$ and $W$ is a matrix $A_{p}$ weight, then $B \in B M O_{W}^{p}$ if and only if both

$$
\begin{equation*}
\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I} B\right) V_{I}^{-1}\right\|^{p} d x<\infty \tag{1.8}
\end{equation*}
$$

and the dual condition

$$
\begin{equation*}
\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{p}}(x)\left(B^{*}(x)-m_{I} B^{*}\right)\left(V_{I}^{\prime}\right)^{-1}\right\|^{p^{\prime}} d x<\infty \tag{1.9}
\end{equation*}
$$

are true.
Finally, to prove Theorem 1.1 we will need to characterize matrix weighted norm inequalities for Haar multipliers. More precisely we will prove the following

Proposition 1.6. Let $1<p<\infty$ and let $W$ be a matrix $A_{p}$ weight. If $\mathscr{D}$ is any dyadic grid and $A:=\left\{A_{I}^{\varepsilon}\right\}_{I \in \mathscr{D}, \varepsilon \in S i g_{d}}$ is a sequence of matrices, then the Haar multiplier

$$
T_{A} \vec{f}:=\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in S i g_{d}} A_{I}^{\varepsilon} \vec{f}_{I}^{\varepsilon} h_{I}^{\varepsilon}
$$

is bounded on $L^{p}(W)$ if and only if

$$
\sup _{I \in \mathscr{D}, \varepsilon \in S i g_{d}}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|<\infty .
$$

1.4. The scalar setting. Let us now make a few comments about these results in the scalar seting. First, it is very easy to see that a scalar function $b$ is in $\mathrm{BMO}_{W}^{p}$ if and only if $b \in \mathrm{BMO}$ when $W$ is a matrix $\mathrm{A}_{p}$ weight of the form $W(x)=w(x) \operatorname{Id}_{n \times n}$ for a scalar $\mathrm{A}_{p}$ weight $w$. In particular, for each dyadic grid $\mathscr{D}$, condition (a) in Theorem 1.3 is trivially equivalent to $b \in \mathrm{BMO}_{\mathscr{D}}$, which given Proposition 1.4 clearly proves the claim. Furthermore, it is well known that $\pi_{b}$ is bounded on $L^{p}(w)$ if and only if $b \in \mathrm{BMO}$ (when $w$ is a scalar $\mathrm{A}_{p}$ weight, see [2]).

Moreover, note that when $p=2$, a careful tracking of the $\|W\|_{\mathrm{A}_{2}}$ characteristic contribution from the implication $(c) \Rightarrow$ (a) in Theorem 1.3 gives us the following (after replacing $\vec{f}$ with $W^{\frac{1}{2}} \vec{f}$ and replacing $A_{I}^{\varepsilon}$ with $\left.\left(m_{I} W\right)^{-\frac{1}{2}} A_{I}^{\varepsilon}\right)$

Corollary 1.7. If $W$ is a matrix $A_{2}$ weight, $\mathscr{D}$ is a dyadic grid, and $\left\{A_{I}^{\varepsilon}\right\}$ is any sequence of $n \times n$ matrices satisfying

$$
\sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{O}(J)}\left(A_{I}^{\varepsilon}\right)^{*} A_{I}^{\varepsilon}<C \int_{J} W(x) d x
$$

for all $J \in \mathscr{D}$ (where $C$ is independent of $J$ ) then

$$
\sum_{\varepsilon \in S i g_{d}} \sum_{I \in \mathscr{O}}\left|A_{I}^{\varepsilon}\left(m_{I} \vec{f}\right)\right|^{2} \lesssim C^{\frac{1}{2}}\|W\|_{A_{2}}^{3}\|\vec{f}\|_{L^{2}(W)}^{2}
$$

Interestingly, note that Corollary 1.7 in the scalar $d=1$ setting appears as Lemma 5.7 in 29 for scalar $\mathrm{A}_{\infty}$ weights and was implicitly used in sharp form with quadratic $\|W\|_{\mathrm{A}_{2}}$ characteristic (versus cubic above) by O. Beznosova in [2] (see (2.4) and (2.5) in [2]) to prove sharp weighted norm inequalities for scalar paraproducts.

Also, a similar $p=2$ matrix weighted Carleson embedding theorem for positive semidefinite sequences was proved in [5] using virtually the same argument as the one used to prove Theorem 1.3. Additionally, note that a version of Corollary 1.7 for positive semidefinite sequences that does not require $W$ to be a matrix $\mathrm{A}_{2}$ weight was very recently proved in [11]. While this result is obviously of great potential for proving sharp matrix $\mathrm{A}_{2}$ results, it is not clear whether one can prove Corollary 1.7 using the results in [11.
1.5. Outline of paper. We will now briefly outline the contents of the paper. In Section 2 we will prove Theorems 1.3, 1.2, and also prove (1.6). In Section 3 we will prove Proposition 1.6 and use this in conjunction with Theorem 1.2 to prove Theorem 1.1 and (1.7). Additionally, we will give short proofs of Proposition 1.4 and Corollary 1.5 in Section 3. Finally, in the last section we will provide very explicit "counterexamples" to Proposition 1.6, Theorem [1.2, and Theorem 1.1 in the sense that, as one would expect, none of these results are true for arbitrary matrix valued symbols and matrix $\mathrm{A}_{2}$ weights.

We will also prove some simple yet nonetheless interesting results involving quantitative matrix weighted norm inequalities for objects related to maximal functions. In particular, we will prove sharp $L^{2}$ estimates for the Christ/Goldberg matrix weighted maximal function from [8], prove weak type estimates for "the" universal $p=2$ matrix weighted maximal function for not necessarily matrix $\mathrm{A}_{2}$ weights, and give a simple "maximal function" proof of the matrix weighted bounds
for sparse operators from [4]. While these results are not needed to prove any of our main results, their proofs require some ideas utilized in this paper and clearly complement (1.6) and (1.7)

Finally, we will remark that well after this paper was written, the first author in [20] has proved that in fact $B \in \mathrm{BMO}_{W}^{p} \Leftrightarrow(1.8) \Leftrightarrow(1.9)$, and that a similar result holds for $\mathrm{BMO}_{W, \mathscr{D}}^{p}$. Furthermore, it is very interesting to note that most of the techniques in this paper are in fact "two weight" techniques in that slight modifications to them allow for extensions of Theorems 1.2 and 1.1 to the $L^{p}(U) \rightarrow L^{p}(W)$ setting where $U, W$ are matrix $\mathrm{A}_{p}$ weights, see [20] for more details.

## 2. Proof of Theorem 1.2

The main goal of this section is to prove Theorem 1.2 via Theorem 1.3. Before we do either, however, we will need to discuss some preliminary results.
2.1. Matrix weighted Littlewood-Paley theory. We will now need the "matrix weighted Triebel-Lizorkin imbedding theorem" from [27, [34], which say that if $W$ is a matrix $\mathrm{A}_{p}$ weight then

$$
\begin{equation*}
\|\vec{f}\|_{L^{p}(W)}^{p} \approx \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} \vec{f}_{I}^{\varepsilon}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \tag{2.1}
\end{equation*}
$$

where $\vec{f}_{I}^{\varepsilon}$ is the vector of Haar coefficients of the components of $\vec{f}$. Note that these were only proved in the $d=1$ setting in [27,34], though a very simple proof that works for $\mathbb{R}^{d}$ was given by the first author in [18]. Furthermore, note that when $p=2$ the above "Littlewood-Paley expression" reduces to a matrix weighted dyadic square function, and in this setting it is known that one has the quantitative bounds (see [3] for $d=1$ and [12] for $d>1$ )

$$
\begin{equation*}
\left(\sum_{I \in \mathscr{O}} \sum_{\varepsilon \in \operatorname{Sig}_{d}}\left|\left(m_{I} W\right)^{\frac{1}{2}} \vec{f}_{I}^{\varepsilon}\right|^{2}\right)^{\frac{1}{2}} \lesssim\|W\|_{\mathrm{A}_{2}}\left(\log \|W\|_{\mathrm{A}_{2}}\right)^{\frac{1}{2}}\|\vec{f}\|_{L^{2}(W)} \tag{2.2}
\end{equation*}
$$

and

$$
\begin{equation*}
\|\vec{f}\|_{L^{2}(W)} \lesssim\|W\|_{\mathrm{A}_{2}}^{\frac{1}{2}}\left(\log \|W\|_{\mathrm{A}_{2}}\right)^{\frac{1}{2}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}}\left|\left(m_{I} W\right)^{\frac{1}{2}} \vec{f}_{I}^{\varepsilon}\right|^{2}\right)^{\frac{1}{2}} \tag{2.3}
\end{equation*}
$$

Unfortunately, while one can attempt to track the matrix $\mathrm{A}_{p}$ dependence in (2.1) from the arguments in either [27], [33], or [18], when $p \neq 2$, it is very unlikely that any of the arguments in these papers provide efficient bounds similar to the ones in (2.2) or (2.3) when $p=2$. With this in mind, it will be implicit that all inequalities involving (2.1) when $p \neq 2$ involve matrix $\mathrm{A}_{p}$ dependence and we will not further comment on this.
2.2. Preliminary lemmas. Before we prove Theorem 1.3 we will need the following three preliminary results, the first of which is from [27], p. 49 .

Lemma 2.1. Suppose that $A$ is an $n \times n$ matrix where $|A \vec{e}| \geq|\vec{e}|$ for any $\vec{e} \in \mathbb{C}^{n}$. If $|\operatorname{det} A| \leq \delta$ for some $\delta \geq 0$, then $\|A\| \leq \delta$ where $\|\cdot\|$ is the canonical matrix norm on $\mathcal{M}_{n}(\mathbb{C})$.

Lemma 2.2. If $W$ is a matrix $A_{p}$ weight then

$$
\left|V_{I}^{\prime} \vec{e}\right| \approx\left|m_{I}\left(W^{-\frac{1}{p}}\right) \vec{e}\right|
$$

for any $\vec{e} \in \mathbb{C}^{n}$. In particular,

$$
\left|m_{I}\left(W^{-\frac{1}{p}}\right) \vec{e}\right| \leq\left|V_{I}^{\prime} \vec{e}\right| \leq\|W\|_{A_{p}}^{\frac{n}{p}}\left|m_{I}\left(W^{-\frac{1}{p}}\right) \vec{e}\right|
$$

Proof. First we show that

$$
\left\|V_{I}^{\prime}\left(m_{I}\left(W^{-\frac{1}{p}}\right)\right)^{-1}\right\| \leq\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}
$$

which will prove half of the lemma. Furthermore, note that the proof of this inequality will in fact also complete the other half of the proof. Since $W$ is a matrix $A_{p}$ weight, Jensen's inequality gives us that

$$
\begin{equation*}
\exp \left[\frac{1}{|I|} \int_{I} \log \left|W^{\frac{1}{p}}(x) \vec{e}\right| d x\right] \leq\|W\|_{\mathrm{A}_{p}}^{\frac{1}{p}}\left|\left(V_{I}^{\prime}\right)^{-1} \vec{e}\right| \tag{2.4}
\end{equation*}
$$

We now prove that $\operatorname{det} V_{I}^{\prime}\left(m_{I}\left(W^{-\frac{1}{p}}\right)^{-1}\right) \leq\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}$ by using some arguments in the proof of Proposition 2.2 in [34]. First, as was commented in [34], $\operatorname{det} Q \leq \prod_{i=1}^{n}\left|Q \vec{e}_{i}\right|$ for any orthonormal basis $\left\{\vec{e}_{i}\right\}_{i=1}^{n}$ of $\mathbb{C}^{n}$ and any positive definite $Q$. Now for fixed $I$ let $\left\{\vec{e}_{i}\right\}_{i=1}^{n}$ be an orthornormal basis of $\mathbb{C}^{n}$ consisting of eigenvectors of $\left(V_{I}^{\prime}\right)^{-1}$. Applying (2.4) to each $\vec{e}_{i}$, taking logarithms, summing, and using the above
inequality gives us that

$$
\begin{aligned}
\frac{1}{|I|} \int_{I} \log \operatorname{det} W^{\frac{1}{p}}(x) d x & \leq \sum_{i=1}^{n} \frac{1}{|I|} \int_{I} \log \left|W^{\frac{1}{p}}(x) \vec{e}_{i}\right| d x \\
& \leq \log \|W\|_{{A_{p}}_{p}}^{\frac{n}{p}}+\log \Pi_{i=1}^{n}\left|\left(V_{I}^{\prime}\right)^{-1} \vec{e}_{i}\right| \\
& =\log \|W\|_{A_{p}}^{\frac{n}{p}}+\log \operatorname{det}\left(V_{I}^{\prime}\right)^{-1}
\end{aligned}
$$

or equivalently

$$
\log \operatorname{det} V_{I}^{\prime} \leq \log \|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}+\frac{1}{|I|} \int_{I} \log \operatorname{det} W^{-\frac{1}{p}}(x) d x
$$

so that

$$
\operatorname{det} V_{I}^{\prime} \leq\|W\|_{A_{p}}^{\frac{n}{p}} \exp \left(m_{I} \log \operatorname{det}\left(W^{-\frac{1}{p}}\right)\right)
$$

for any $I \in \mathscr{D}$. Combining this with the matrix Jensen inequality (Lemma 7.2 in [27]) we have that

$$
\operatorname{det} V_{I}^{\prime} \leq\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}} \operatorname{det} m_{I}\left(W^{-\frac{1}{p}}\right)
$$

so that det $V_{I}^{\prime}\left(m_{I}\left(W^{-\frac{1}{p}}\right)^{-1}\right) \leq\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}$.
Moreover, note that for any $\vec{e} \in \mathbb{C}^{n}$ we have

$$
\left|m_{I}\left(W^{-\frac{1}{p}}\right) \vec{e}\right| \leq \frac{1}{|I|} \int_{I}\left|W^{-\frac{1}{p}}(x) \vec{e}\right| d x \leq\left(\frac{1}{|I|} \int_{I}\left|W^{-\frac{1}{p}}(x) \vec{e}\right|^{p^{\prime}} d x\right)^{\frac{1}{p^{\prime}}} \leq\left|V_{I}^{\prime} \vec{e}\right|
$$

which means that

$$
\left|V_{I}^{\prime}\left(m_{I}\left(W^{-\frac{1}{p}}\right)^{-1}\right) \vec{e}\right| \geq|\vec{e}|
$$

for any $\vec{e} \in \mathbb{C}^{n}$. The proof now follows immediately from Lemma 2.1,

Finally, we will need the following $\mathbb{R}^{d}$ version of the classical dyadic Carleson Lemma from [29]. Note that the proof is almost identical to the proof of Lemma 5.3 in [29] and will therefore be omitted.

Proposition 2.3. Let $\left\{\lambda_{I}^{\varepsilon}: I \in \mathscr{D}, \varepsilon \in\right.$ Sig $\left._{d}\right\}$ be a"Carleson sequence" of positive numbers in the sense that

$$
\sup _{J \in \mathscr{D}} \frac{1}{|J|} \sum_{I \in \mathscr{D}(J)} \sum_{\epsilon \in S i g_{d}} \lambda_{I}^{\varepsilon} \leq C<\infty
$$

Then for any positive sequence of real numbers $\left\{a_{I}\right\}$, we have that

$$
\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in S i g_{d}} a_{I} \lambda_{I}^{\varepsilon} \leq C \int_{\mathbb{R}^{d}} a^{*}(x) d x
$$

where $a^{*}(x)=\sup _{I \ni x} a_{I}$.
2.3. Stopping time. Let us now review the surprisingly useful stopping time from [18], which is a matrix $p \neq 2$ adaption of the stopping time from [23, 30]. Assume that $W$ is a matrix $\mathrm{A}_{p}$ weight. For any cube $I \in \mathscr{D}$, let $\mathscr{J}(I)$ be the collection of maximal $J \in \mathscr{D}(I)$ such that

$$
\begin{equation*}
\left\|V_{J} V_{I}^{-1}\right\|^{p}>\lambda_{1} \text { or }\left\|V_{J}^{-1} V_{I}\right\|^{p^{\prime}}>\lambda_{2} \tag{2.5}
\end{equation*}
$$

for some $\lambda_{1}, \lambda_{2}>1$ to be specified later. Also, let $\mathscr{F}(I)$ be the collection of dyadic subcubes of $I$ not contained in any cube $J \in \mathscr{J}(I)$, so that clearly $J \in \mathscr{F}(J)$ for any $J \in \mathscr{D}(I)$.

Let $\mathscr{J}^{0}(I):=\{I\}$ and inductively define $\mathscr{J}^{j}(I)$ and $\mathscr{F}^{j}(I)$ for $j \geq 1$ by $\mathscr{J}^{j}(I):=\bigcup_{J \in \mathscr{J}^{j-1}(I)} \mathscr{J}(J)$ and $\mathscr{F}^{j}(I):=\bigcup_{J \in \mathscr{L}^{j-1}(I)} \mathscr{F}(J)$. Clearly the cubes in $\mathscr{J}^{j}(I)$ for $j>0$ are pairwise disjoint. Furthermore, since $J \in \mathscr{F}(J)$ for any $J \in \mathscr{D}(I)$, we have that $\mathscr{D}(I)=$ $\bigcup_{j=0}^{\infty} \mathscr{F}^{j}(I)$. We will slightly abuse notation and write $\bigcup \mathscr{J}(I)$ for the set $\bigcup_{J \in \mathscr{J}(I)} J$ and write $|\bigcup \mathscr{J}(I)|$ for $\left|\bigcup_{J \in \mathscr{J}(I)} J\right|$. We will now show that $\mathscr{J}$ is a decaying stopping time in the sense of [23]. Note that while the easy proof is from [18, we will include the details since we will need to precisely track the $\mathrm{A}_{p}$ characteristic contribution.

Lemma 2.4. Let $1<p<\infty$ and let $W$ be a matrix $A_{p}$ weight. For $\lambda_{1}, \lambda_{2}>1$ large enough, we have that $\left|\bigcup \mathscr{J}^{j}(I)\right| \leq 2^{-j}|I|$ for every $I \in \mathscr{D}$.

Proof. By iteration, it is enough to prove the lemma for $j=1$. For $I \in \mathscr{D}$, let $\mathscr{G}(I)$ denote the collection of maximal $J \in \mathscr{D}(I)$ such that the first inequality (but not necessarily the second inequality) in (2.5) holds. Then by maximality and elementary arguments involving the definition of $V_{I}$ and the equivalence between the matrix and trace norm for positive matrices, we have that

$$
\left|\bigcup_{J \in \mathscr{G}(I)} J\right|=\sum_{J \in \mathscr{G}(I)}|J| \lesssim \frac{1}{\lambda_{1}} \sum_{J \in \mathscr{G}(I)} \int_{J}\left\|W^{\frac{1}{p}}(y) V_{I}^{-1}\right\|^{p} d y \leq \frac{C_{1}|I|}{\lambda_{1}}
$$

for some $C_{1}>0$ only depending on $n$ and $d$.
On the other hand, let $I \in \mathscr{D}$, let $\widetilde{\mathscr{G}}(I)$ denote the collection of maximal $J \in \mathscr{D}(I)$ such that the second inequality (but not necessarily the first inequality) in (2.5) holds. Then by the matrix $\mathrm{A}_{p}$ condition
we have

$$
\left|\bigcup_{J \in \widetilde{\mathscr{G}}(I)} J\right| \leq \frac{C_{2}}{\lambda_{2}} \sum_{J \in \widetilde{\mathscr{G}}(I)} \int_{J}\left\|W^{-\frac{1}{p}}(y) V_{I}\right\|^{p^{\prime}} d y \leq \frac{C_{2}^{\prime}\|W\|_{A_{p}}^{\frac{p^{\prime}}{p}}}{\lambda_{2}}|I|
$$

for some $C_{2}^{\prime}$ only depending on $n$ and $d$. The proof is now completed by setting $\lambda_{1}=4 C_{1}$ and $\lambda_{2}=4 C_{2}^{\prime}\|W\|_{A_{p}}^{\frac{p^{\prime}}{p}}$.

While we will not have a need to discuss matrix $\mathrm{A}_{p, \infty}$ weights in detail in this paper, note that in fact Lemma 3.1 in [34] immediately gives us that Lemma 2.4 holds for matrix $\mathrm{A}_{p, \infty}$ weights (with a different $\lambda_{2}$ of course.)
2.4. Proofs. We now prove Theorem 1.3

Proof of Theorem 1.3. (b) $\Rightarrow$ (a): By dyadic Littlewood-Paley theory, we need to show that

$$
\begin{align*}
& \int_{\mathbb{R}^{d}}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}} \frac{\left|V_{I} A_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}} \vec{f}\right)\right|^{2}}{|I|} \chi_{I}(t)\right)^{\frac{p}{2}} d t \\
& \quad \leq \int_{\mathbb{R}^{d}}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{\mathscr { D }}} \frac{\left(\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\| m_{I}\left|V_{I} W^{-\frac{1}{p}} \vec{f}\right|\right)^{2}}{|I|} \chi_{I}(t)\right)^{\frac{p}{2}} d t  \tag{2.6}\\
& \quad \lesssim\|\vec{f}\|_{L^{p}}^{p}
\end{align*}
$$

for any $\vec{f} \in L^{p}\left(\mathbb{R}^{d} ; \mathbb{C}^{n}\right)$.
Now let

$$
\tilde{A}=\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\| h_{I}^{\varepsilon}
$$

and let

$$
M_{W}^{\prime} \vec{f}(x)=\sup _{\mathscr{D} \ni I \ni x} m_{I}\left|V_{I} W^{-\frac{1}{p}} \vec{f}\right|
$$

Clearly for any $\mathscr{D} \ni I \ni x$ we have that

$$
m_{I}\left|V_{I} W^{-\frac{1}{p}} \vec{f}\right| \leq m_{I} M_{W}^{\prime} \vec{f}
$$

so that again by dyadic Littlewood-Paley theory

$$
\begin{aligned}
(2.6) & \leq \int_{\mathbb{R}^{d}}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}} \frac{\left(\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\| m_{I}\left(M_{W}^{\prime} \vec{f}\right)\right)^{2}}{|I|} \chi_{I}(t)\right)^{\frac{p}{2}} d t \\
& \lesssim\left\|\pi_{\tilde{A}} M_{W}^{\prime} \vec{f}\right\|_{L^{p}}^{p} \\
& \lesssim\|A\|_{*}^{p}\left\|M_{W}^{\prime} \vec{f}\right\|_{L^{p}}^{p}
\end{aligned}
$$

where $\|A\|_{*}$ is the canonical supremum from condition (b) and $\pi_{\tilde{A}}$ is the scalar dyadic paraproduct with respect to the function $\tilde{A}$.

However, it is easy to see that

$$
\left\|M_{W}^{\prime}\right\|_{L^{p} \rightarrow p} \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{1}{p-1}}
$$

by using some simple ideas from [13]. Namely, it is well known (see [13], p. 6) that $\left|W^{-\frac{1}{p}}(x) \vec{e}\right|^{p^{\prime}}$ is a scalar $\mathrm{A}_{p^{\prime}}$ weight for any matrix $\mathrm{A}_{p}$ weight $W$ and $\vec{e} \in \mathbb{C}^{n}$ with $\mathrm{A}_{p^{\prime}}$ characteristic less than or equal to $\|W\|_{\mathrm{A}_{p}}^{-\frac{p^{\prime}}{p}}$. Thus, by the scalar reverse Hölder inequality, the matrix $\mathrm{A}_{p}$ condition, and the equivalence of the operator and trace norms, we have for $\epsilon \approx\|W\|_{\mathrm{A}_{p}}^{-\frac{p^{\prime}}{p}}$ small enough that

$$
\begin{aligned}
\left(\frac{1}{|I|} \int_{I}\left\|V_{I} W^{-\frac{1}{p}}(y)\right\|^{p^{\prime}+p^{\prime} \epsilon} d y\right)^{\frac{1}{p^{\prime}+p^{\prime} \epsilon}} & =\left(\frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{p}}(y) V_{I}\right\|^{p^{\prime}+p^{\prime} \epsilon} d y\right)^{\frac{1}{p^{\prime}+p^{\prime} \epsilon}} \\
& \lesssim\left(\frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{p}}(y) V_{I}\right\|^{p^{\prime}} d y\right)^{\frac{1}{p^{\prime}}} \\
& \lesssim\|W\|_{A_{p}}^{\frac{1}{p}}
\end{aligned}
$$

Therefore, we have

$$
\begin{aligned}
m_{I}\left|V_{I} W^{-\frac{1}{p}} \vec{f}\right| & =\frac{1}{|I|} \int_{I}\left|V_{I} W^{-\frac{1}{p}}(y) \vec{f}(y)\right| d y \\
& \leq\left(\frac{1}{|I|} \int_{I}\left\|V_{I} W^{-\frac{1}{p}}(y)\right\|^{p^{\prime}+p^{\prime} \epsilon} d y\right)^{\frac{1}{p^{\prime}+p^{\prime} \epsilon}}\left(\frac{1}{|I|} \int_{I}|\vec{f}(y)|^{\frac{p+p \epsilon}{1+p \epsilon}} d y\right)^{\frac{1+p \epsilon}{p+p \epsilon}} \\
& \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{1}{p}}\left(\frac{1}{|I|} \int_{I}|\vec{f}(y)|^{\frac{p+p \epsilon}{1+p \epsilon}} d y\right)^{\frac{1+p \epsilon}{p+p \epsilon}}
\end{aligned}
$$

Which means that

$$
M_{W}^{\prime} \vec{f}(x) \leq\|W\|_{\mathrm{A}_{p}}^{\frac{1}{p}}\left(M|\vec{f}|^{\frac{p+p \epsilon}{1+p \epsilon}}(x)\right)^{\frac{1+p \epsilon}{p+p \epsilon}}
$$

(where $M$ is the ordinary unweighted maximal function) so that by the standard $L^{1+\delta}$ bounds for the ordinary maximal function when $\delta>0$ is small, we get

$$
\left\|M_{W}^{\prime}\right\|_{L^{p} \rightarrow L^{p}} \lesssim \epsilon^{\frac{1}{p}}\|W\|_{A_{p}}^{\frac{1}{p}} \lesssim\|W\|_{A_{p}}^{\frac{1}{p}+\frac{p^{\prime}}{p^{2}}}=\|W\|_{A_{p}}^{\frac{1}{p-1}}
$$

which completes the proof that $(\mathrm{b}) \Rightarrow(\mathrm{a})$.
(a) $\Rightarrow(\mathrm{b}):$ Fixing $J \in \mathscr{D}$, plugging in the test functions $\vec{f}:=\chi_{J} \vec{e}_{i}$ into $\Pi_{A}$ for any orthonormal basis $\left\{\vec{e}_{i}\right\}_{i=1}^{n}$ of $\mathbb{C}^{n}$, and using (a) combined with dyadic Littlewood-Paley theory and the equivalence of the matrix norm and the Hilbert-Schmidt norm gives us that

$$
\begin{aligned}
\left\|\Pi_{A}\right\|_{L^{p} \rightarrow L^{p}}^{p}|J| & \gtrsim \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left\|V_{I} A_{I}^{\varepsilon} m_{I}\left(\chi_{J} W^{-\frac{1}{p}}\right)\right\|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \geq \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}(J)} \frac{\left\|V_{I} A_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}}\right)\right\|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x
\end{aligned}
$$

which in conjunction with Lemma 2.2 says that

$$
\begin{aligned}
\sup _{J \in \mathscr{D}} \frac{1}{|J|} & \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)} \frac{\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \lesssim \sup _{J \in \mathscr{D}} \frac{1}{|J|} \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)} \frac{\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{\prime}\right\|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \lesssim\|W\|_{A_{p}}^{n} \sup _{J \in \mathscr{O}} \frac{1}{|J|} \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)} \frac{\left\|V_{I} A_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}}\right)\right\|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x .
\end{aligned}
$$

Condition (b) now follows immediately from $(i i) \Longleftrightarrow(i)$ of Theorem 3.1 in [28], which (after a trivial relabeling) says that for any nonnegative sequence $\left\{a_{I}\right\}_{I \in \mathscr{D}}$ of real numbers we have that

$$
\sup _{J \in \mathscr{D}}\left(\frac{1}{|J|} \sum_{I \in \mathscr{D}(J)} a_{I}^{2}\right)^{\frac{1}{2}} \approx \sup _{J \in \mathscr{D}}\left(\frac{1}{|J|} \int_{J}\left(\sum_{I \in \mathscr{D}(J)} \frac{a_{I}^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x\right)^{\frac{1}{p}}
$$

We now prove that (c) $\Rightarrow(\mathrm{b})$ and (a) $\Rightarrow(\mathrm{c})$ for the case $2 \leq p<\infty$.
(c) $\Rightarrow$ (b) when $2 \leq p<\infty$ : Note that condition (c) is equivalent to

$$
\frac{1}{|K|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(K)}\left\|V_{K}^{-1}\left(A_{I}^{\varepsilon}\right)^{*} V_{I}^{2} A_{I}^{\varepsilon} V_{K}^{-1}\right\| \lesssim 1
$$

for any $K \in \mathscr{D}$. Fix $J \in \mathscr{D}$ and for each $j \in \mathbb{N}$ let $\mathscr{J}^{j}(J)$ and $\mathscr{F}^{j}(J)$ be defined as they previously where $\lambda_{1} \approx 1$, and $\lambda_{2} \approx\|W\|_{A_{p}}^{\frac{p^{\prime}}{p}}$ are picked so that Lemma 2.4 is true. Then (2.5) tells us that

$$
\begin{aligned}
\frac{1}{|J|} & \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{\mathscr { O }}(J)}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|^{2} \\
& \leq \frac{1}{|J|} \sum_{j=1}^{\infty} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{K \in \mathscr{\mathscr { G }}^{j-1}(J)} \sum_{I \in \mathscr{F}(K)}\left\|V_{I}^{-1} V_{K}\right\|\left\|V_{K}^{-1}\left(A_{I}^{\varepsilon}\right)^{*} V_{I}^{2} A_{I}^{\varepsilon} V_{K}^{-1}\right\|\left\|V_{K} V_{I}^{-1}\right\| \\
& \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{2}{p}} \frac{1}{|J|} \sum_{j=1}^{\infty} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{K \in \mathcal{J}^{j-1}(J)} \sum_{I \in \mathscr{D}(K)}\left\|V_{K}^{-1}\left(A_{I}^{\varepsilon}\right)^{*} V_{I}^{2} A_{I}^{\varepsilon} V_{K}^{-1}\right\| \\
& \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{2}{p}} \frac{1}{|J|} \sum_{j=1}^{\infty} \sum_{K \in \mathscr{\mathscr { C }}^{j-1}(J)}|K| \\
& \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{2}{p}} \sum_{j=1}^{\infty} 2^{-j} \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{2}{p}} .
\end{aligned}
$$

(a) $\Rightarrow$ (c) when $2 \leq p<\infty$ : Fix $J \in \mathscr{D}$ and $\vec{e} \in \mathbb{C}^{n}$. If $\vec{f}=W^{\frac{1}{p}} \chi_{J} \vec{e}$, then condition (a), the definition of $V_{J}$, and Hölder's inequality give us that

$$
\begin{aligned}
|J| \mid V_{J} \overrightarrow{e^{p}}\left\|\Pi_{A}\right\|_{L^{p} \rightarrow L^{p}}^{p} & \gtrsim \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} A_{I}^{\varepsilon} m_{I}\left(\chi_{J} \vec{e}\right)\right|^{2}}{|I|} \chi_{I}(t)\right)^{\frac{p}{2}} d t \\
& \geq|J|\left[\frac{1}{|J|} \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}(J)} \frac{\left|V_{I} A_{I}^{\varepsilon} \vec{e}\right|^{2}}{|I|} \chi_{I}(t)\right)^{\frac{p}{2}} d t\right] \\
& \geq|J|\left[\frac{1}{|J|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)}\left|V_{I} A_{I}^{\varepsilon} \vec{e}\right|^{2}\right]^{\frac{p}{2}}
\end{aligned}
$$

which proves (c), and in fact shows that (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) when $2 \leq$ $p<\infty$. We will now complete the proof when $1<p \leq 2$.
(b) $\Rightarrow$ (c) when $1<p \leq 2$ : To avoid confusion in the subsequent arguments, we will write $V_{I}=V_{I}(W, p)$ to indicate which $W$ and $p$ the $V_{I}$ at hand is referring to. As mentioned before, it is easy to see that $W$ is a matrix $\mathrm{A}_{p}$ weight if and only if $W^{1-p^{\prime}}$ is a matrix $\mathrm{A}_{p^{\prime}}$ weight and that we may take $V_{I}\left(W^{1-p^{\prime}}, p^{\prime}\right)=V_{I}^{\prime}(W, p)$ and $V_{I}^{\prime}\left(W^{1-p^{\prime}}, p^{\prime}\right)=$ $V_{I}(W, p)$. Now if $(\mathrm{b})$ is true, then the two equalities above give us that

$$
\begin{aligned}
& \sup _{J \in \mathscr{D}} \frac{1}{|J|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}(J)}\left\|V_{I}\left(W^{1-p^{\prime}}, p^{\prime}\right)\left(A_{I}^{\varepsilon}\right)^{*} V_{I}^{\prime}\left(W^{1-p^{\prime}}, p^{\prime}\right)\right\|^{2} \\
& =\sup _{J \in \mathscr{D}} \frac{1}{|J|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)}\left\|V_{I}(W, p) A_{I}^{\varepsilon} V_{I}^{\prime}(W, p)\right\|^{2}<\infty .
\end{aligned}
$$

However, repeating word for word the proofs of $(\mathrm{b}) \Rightarrow(\mathrm{a}) \Rightarrow$ (c) for the case $2 \leq p<\infty$ (where $W^{1-p^{\prime}}$ replaces $W$ and $A^{*}:=\left\{\left(A_{I}^{\varepsilon}\right)^{*}\right.$ : $\left.I \in \mathscr{D}, \varepsilon \in \in \operatorname{Sig}_{d}\right\}$ replaces the sequence $A$ ) gives us that there exists $C>0$ where

$$
\frac{1}{|J|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}(J)} A_{I}^{\varepsilon}\left(V_{I}\left(W^{1-p^{\prime}}, p^{\prime}\right)\right)^{2}\left(A_{I}^{\varepsilon}\right)^{*}<C\left(V_{J}\left(W^{1-p^{\prime}}, p^{\prime}\right)\right)^{2}
$$

which proves (c) when $1<p \leq 2$.
(c) $\Rightarrow$ (b) when $1<p \leq 2$ : This follows immediately by again replacing $W$ with $W^{1-p^{\prime}}$, replacing $A$ with $A^{*}:=\left\{\left(A_{I}^{\varepsilon}\right)^{*}\right\}_{I \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}$, and using the proof of $(\mathrm{c}) \Rightarrow$ (b) when $2 \leq p<\infty$. Since (a) $\Leftrightarrow$ (b) was shown for all $1<p<\infty$, we therefore have that (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) for all $1<p<\infty$. The proof is now completed.

We can now prove 1.2

Proof of Theorem 1.2. Note that (2.1), standard dyadic LittlewoodPaley theory, and the definition of $\pi_{B}$ gives us that

$$
\left\|\pi_{B} W^{-\frac{1}{p}} \vec{f}\right\|_{L^{p}(W)}^{p} \approx \int_{\mathbb{R}^{d}}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}} \frac{\left|V_{I} B_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}} \vec{f}\right)\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \approx\left\|\Pi_{B} \vec{f}\right\|_{L^{p}}^{p}
$$

where $\Pi_{B}$ is the operator in $(a)$ of 1.3 with respect to the Haar coefficient sequence $\left\{B_{I}^{\varepsilon}\right\}_{I \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}$. Thus, the $L^{p}(W)$ boundedness of $\pi_{B}$ is equivalent to the $L^{p}$ boundedness of $\Pi_{B}$. Thanks to Theorem 1.3, the proof will be completed by showing that $B \in \mathrm{BMO}_{W}^{p}$ if and only if $\Pi_{B}$ is bounded on $L^{p}$.

To that end, if again $\left\{\vec{e}_{i}\right\}_{i=1}^{n}$ is any orthonormal basis of $\mathbb{C}^{n}$, then $\Pi_{B}$ being bounded on $L^{p}$ in conjunction with (2.1) gives us that

$$
\begin{aligned}
\sup _{J \in \mathscr{D}} \frac{1}{|J|} & \int_{J}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{J} B\right) V_{J}^{-1}\right\|^{p} d x \\
& \approx \sum_{i=1}^{n} \sup _{J \in \mathscr{D}} \frac{1}{|J|} \int_{\mathbb{R}^{d}}\left|W^{\frac{1}{p}}(x) \chi_{J}(x)\left(B(x)-m_{J} B\right) V_{J}^{-1} \vec{e}_{i}\right|^{p} d x \\
& \approx \sum_{i=1}^{n} \sup _{J \in \mathscr{D}} \frac{1}{|J|} \int_{\mathbb{R}^{d}}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)} \frac{\left|V_{I} B_{I}^{\varepsilon} V_{J}^{-1} \vec{e}_{i}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \leq \sum_{i=1}^{n} \sup _{J \in \mathscr{D}} \frac{1}{|J|} \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{O}(J)} \sum_{\epsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} B_{I}^{\varepsilon} m_{I}\left(W^{-\frac{1}{p}}\left\{\chi_{J} W^{\frac{1}{p}} V_{J}^{-1} \vec{e}_{i}\right\}\right)\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \lesssim \sum_{i=1}^{n} \sup _{J \in \mathscr{D}}|J|^{-1}\left\|\Pi_{B} \chi_{J} W^{\frac{1}{p}} V_{J}^{-1} \vec{e}_{i}\right\|_{L^{p}}^{p} \\
& \lesssim\left\|\Pi_{B}\right\|_{L^{p}}^{p}
\end{aligned}
$$

by the definition of $V_{J}$, which means that the first condition of the definition of $\mathrm{BMO}_{W}^{p}$ is true for all $1<p<\infty$. Similarly, the validity of the second condition of the definition of $\mathrm{BMO}_{W}^{p}$ for all $1<p<\infty$ if $\Pi_{B}$ is bounded follows by the above arguments in conjunction with the arguments used to prove $(b) \Rightarrow(c)$ when $1<p \leq 2$ (that is, taking adjoints in condition (b) and using the fact that $W$ is a matrix $\mathrm{A}_{p}$ weight if and only if $W^{1-p^{\prime}}$ is a matrix $\mathrm{A}_{p^{\prime}}$ weight.

Now if $2 \leq p<\infty$ and $B \in \mathrm{BMO}_{W}^{p}$ then as before (2.1) gives us that for any $\vec{e} \in \mathbb{C}^{n}$

$$
\begin{aligned}
\sup _{J \in \mathscr{D}} \frac{1}{|J|} & \int_{J}\left|W^{\frac{1}{p}}(x)\left(B(x)-m_{J} B\right) V_{J}^{-1} \vec{e}\right|^{p} d x \\
& \approx \sup _{J \in \mathscr{D}} \frac{1}{|J|} \int_{J}\left(\sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{O}(J)} \frac{\left|V_{I} B_{I}^{\varepsilon} V_{J}^{-1} \vec{e}\right|^{2}}{|I|} \chi_{I}(x) d x\right)^{\frac{p}{2}} d x \\
& \geq \sup _{J \in \mathscr{D}}\left(\frac{1}{|J|} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \sum_{I \in \mathscr{D}(J)}\left|V_{I} B_{I}^{\varepsilon} V_{J}^{-1} \vec{e}\right|^{2}\right)^{\frac{p}{2}}
\end{aligned}
$$

since $\frac{p}{2} \geq 1$, which says that condition (c) (when $2 \leq p<\infty$ ) is true if $B \in \mathrm{BMO}_{W}^{p}$. The same argument using (2.1) for the space $L^{p^{\prime}}\left(W^{1-p^{\prime}}\right)$
also shows that condition (c) (when $1<p \leq 2$ ) is true if $B \in \mathrm{BMO}_{W}^{p}$ since in this case $p^{\prime} \geq 2$. The proof is now completed.

We now prove (1.6)

Proof of (1.6). By (2.3) and the proof of (b) $\Rightarrow$ (a) we have

$$
\left\|\pi_{B} W^{-\frac{1}{2}} \vec{f}\right\|_{L^{2}(W)} \lesssim\|W\|_{\mathrm{A}_{2}}^{\frac{1}{2}}\left(\log \|W\|_{\mathrm{A}_{2}}\right)^{\frac{1}{2}}\left\|\Pi_{B} \vec{f}\right\|_{L^{2}} \lesssim\|W\|_{\mathrm{A}_{2}}^{\frac{3}{2}}\left(\log \|W\|_{\mathrm{A}_{2}}\right)^{\frac{1}{2}}\|B\|_{*}^{\frac{1}{2}} .
$$

We will end this section with an important comment. Note that the proof of $(\mathrm{b}) \Rightarrow(\mathrm{c})$ when $1<p \leq 2$ immediately says that $B \in \mathrm{BMO}_{W}^{p}$ if and only if $B^{*} \in \mathrm{BMO}_{W^{1-p^{\prime}}}^{p^{\prime}}$. Thus, we have $B \in \mathrm{BMO}_{W}^{p}$ if and only if $\pi_{B^{*}}$ is bounded on $L^{p}\left(W^{1-p^{\prime}}\right)$ if and only if $\left(\pi_{B^{*}}\right)^{*}$ is bounded on $L^{p}(W)$. We will record this as a corollary since we will need this fact when we prove sufficiency in the proof of Theorem 1.1.

Corollary 2.5. If $W$ is a matrix $A_{p}$ weight then $B \in B M O_{W}^{p}$ if and only if $\left(\pi_{B^{*}}\right)^{*}$ is bounded on $L^{p}(W)$.

## 3. Proof of Theorem 1.1

3.1. Preliminaries. Before we prove Theorem 1.1 we will need some preliminary results, including Proposition 1.6 .

Proof of Proposition 1.6. If $M=\sup _{I \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{\prime}\right\|<\infty$, then two applications of (2.1) and (1.3) give us that

$$
\begin{aligned}
\left\|T_{A} \vec{f}\right\|_{L^{p}(W)}^{p} & \approx \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} A_{I}^{\varepsilon} \vec{f}_{I}^{\varepsilon}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \leq \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|^{2}\left|V_{I} \vec{f}_{I}^{\varepsilon}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \leq M^{p} \int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{O}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} \vec{f}_{I}^{\varepsilon}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \approx M^{p}\|\vec{f}\|_{L^{p}(W)}^{p} .
\end{aligned}
$$

For the other direction, let $\ell(J)$ denote the side-length of the cube $J$. Fix some $J_{0} \in \mathscr{D}$ and $\varepsilon^{\prime} \in \operatorname{Sig}_{d}$, and let $J_{0}^{\prime} \in \mathscr{D}\left(J_{0}\right)$ with $\ell\left(J_{0}^{\prime}\right)=\frac{1}{2} \ell\left(J_{0}\right)$.

Again by (2.1) we have that

$$
\begin{equation*}
\int_{\mathbb{R}^{d}}\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \frac{\left|V_{I} A_{I}^{\varepsilon}\left(W^{-\frac{1}{p}} \vec{f}\right)_{I}^{\varepsilon}\right|^{2}}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \lesssim\|\vec{f}\|_{L^{p}}^{p} \tag{3.1}
\end{equation*}
$$

Plugging $\vec{f}:=\chi_{J_{0}} \vec{e}$ for any $\vec{e} \in \mathbb{C}^{n}$ into (3.1) and noticing that

$$
\left(W^{-\frac{1}{p}} \chi_{J_{0}^{\prime}} \vec{e}\right)_{J_{0}}^{\varepsilon^{\prime}}= \pm 2^{-\frac{d}{2}}\left|J_{0}\right|^{\frac{1}{2}} m_{J_{0}^{\prime}}\left(W^{-\frac{1}{p}} \vec{e}\right)
$$

gives us (in conjunction with Lemma (2.2) that

$$
\left\|V_{J_{0}} A_{J_{0}}^{\varepsilon^{\prime}} V_{J_{0}^{\prime}}^{\prime}\right\| \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}\left\|V_{J_{0}} A_{J_{0}}^{\varepsilon^{\prime}} m_{J_{0}^{\prime}}\left(W^{-\frac{1}{p}}\right)\right\| \lesssim\|W\|_{\mathrm{A}_{p}}^{\frac{n}{p}}\left\|T_{A}\right\|_{L^{p} \rightarrow L^{p}}
$$

Using the definition of $V_{J_{0}^{\prime}}^{\prime}$ and summing over all of the $2^{d}$ first generation children $J_{0}^{\prime}$ of $J_{0}$ in conjunction with (1.3) finally (after taking the supremum over $J_{0} \in \mathscr{D}$ ) gives us that

$$
\sup _{J \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}\left\|V_{J} A_{J}^{\varepsilon} V_{J}^{-1}\right\| \lesssim \sup _{J \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}\left\|V_{J} A_{J}^{\varepsilon} V_{J}^{\prime}\right\|<\|W\|_{A_{p}}^{\frac{n}{p}}\left\|T_{A}\right\|_{L^{p} \rightarrow L^{p}}
$$

as desired.
To prove both necessity in Theorem 1.1 and Proposition 1.4 we will need the following lemma. While the simple proof is very similar to the proof of Lemma 6.2 in [21], we will nevertheless provide the details.

Lemma 3.1. Let $B$ be a locally integrable $\mathcal{M}_{n}(\mathbb{C})$ valued function on $\mathbb{R}^{d}$, W a matrix $A_{p}$ weight, and $Q$ a cube. Then

$$
\begin{aligned}
& \left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q} B\right) V_{Q}^{-1}\right|^{p} d x\right)^{\frac{1}{p}} \\
& \quad \leq\left(1+\|W\|_{A_{p}}^{\frac{1}{p}}\right) \inf _{A \in \mathcal{M}_{n}(\mathbb{C})}\left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(x)(B(x)-A) V_{Q}^{-1}\right|^{p} d x\right)^{\frac{1}{p}} .
\end{aligned}
$$

Proof. By the triangle inequality,

$$
\begin{align*}
\left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q} B\right) V_{Q}^{-1}\right|^{p} d x\right)^{\frac{1}{p}} & \leq\left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(x)(B(x)-A) V_{Q}^{-1}\right|^{p} d x\right)^{\frac{1}{p}} \\
& +\left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(x)\left(A-m_{Q} B\right) V_{Q}^{-1}\right|^{p} d x\right)^{\frac{1}{p}} \tag{3.2}
\end{align*}
$$

However,

$$
\begin{aligned}
\left\lvert\, W^{\frac{1}{p}}(x)\right. & \left.\left(A-m_{Q} B\right) V_{Q}^{-1}\right|^{p} \\
& =\left|\frac{1}{|Q|} \int_{Q} W^{\frac{1}{p}}(x)(B(y)-A) V_{Q}^{-1} d y\right|^{p} \\
& =\left|\frac{1}{|Q|} \int_{Q}\left(W^{\frac{1}{p}}(x) W^{-\frac{1}{p}}(y)\right) W^{\frac{1}{p}}(y)(B(y)-A) V_{Q}^{-1} d y\right|^{p} \\
& \leq\left(\frac{1}{|Q|} \int_{Q}\left\|W^{\frac{1}{p}}(x) W^{-\frac{1}{p}}(y)\right\|^{p^{\prime}} d y\right)^{\frac{p}{p}}\left(\frac{1}{|Q|} \int_{Q}\left|W^{\frac{1}{p}}(y)(B(y)-A) V_{Q}^{-1}\right|^{p} d y\right) .
\end{aligned}
$$

Plugging this into (3.2) and using (1.1) completes the proof.

The proof strategy for sufficiency in Theorem 1.1 will follow the simple arguments in [25], though of course more care must be taken in our situation due to noncommutativity. As in [25] the starting point is the fact that any of the Riesz transforms are in the $L^{2}$ SOT convex hull of the so called "first order Haar shifts" (or for short, "Haar shifts") which are defined by

$$
Q_{\sigma} h_{I}^{\varepsilon}=h_{\sigma(I)}^{\sigma(\varepsilon)}
$$

and where (slightly abusing notation in the obvious way) $\sigma: \mathscr{D} \times \operatorname{Sig}_{d} \rightarrow$ $\mathscr{D} \times \operatorname{Sig}_{d}$ satisfies $2 \ell(\sigma(I))=\ell(I)$ and $\sigma(I) \subseteq I$ for each $I \in \mathscr{D}$ (see [17] for the definition of general Haar shifts, which are used to study general CZOs). Fixing $\sigma$ and letting $Q=Q_{\sigma}$, it is then enough to get an $L^{p}(W)$ bound on each $[B, Q]$. Before we do this, however, we will need to prove that first of all $Q$ is bounded on $L^{p}(W)$ if $W$ is a matrix $\mathrm{A}_{p}$ weight. Note that this was in fact done for $p=2$ in [12].

Proposition 3.2. Each of the Haar shifts $Q$ are bounded on $L^{p}(W)$ if $W$ is a matrix $A_{p}$ weight.

Proof. The proof follows easily from two applications of (2.1). In particular, note that

$$
Q \vec{f}=\sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} \sum_{I^{\prime} \in \mathscr{D}} \vec{f}_{I^{\prime}}^{\varepsilon^{\prime}} h_{\sigma\left(I^{\prime}\right)}^{\sigma\left(\varepsilon^{\prime}\right)}
$$

If $\tilde{I}$ denotes the parent of $I \in \mathscr{D}$ then (2.1) in conjunction with (1.3) gives us that

$$
\begin{aligned}
\|Q \vec{f}\|_{L^{p}(W)}^{p} & \lesssim \int_{\mathbb{R}^{d}}\left(\sum_{I \in \sigma(\mathscr{D})} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} \frac{\left|V_{I} \vec{f}_{\tilde{I}}^{\prime}\right|}{|I|} \chi_{I}(x)\right)^{\frac{p}{2}} d x \\
& \lesssim \int_{\mathbb{R}^{d}}\left(\sum_{I \in \sigma(\mathscr{D})} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} \frac{\left|V_{\tilde{I}} \overrightarrow{\tilde{f}}_{\tilde{I}}^{\prime}\right|}{|\tilde{I}|} \chi_{\tilde{I}}(x)\right)^{\frac{p}{2}} d x \\
& \lesssim\|\vec{f}\|_{L^{p}(W)}^{p} .
\end{aligned}
$$

3.2. Proof of sufficiency. We now prove sufficiency in Theorem 1.1

Proof of Sufficiency. First write

$$
B=\sum_{I^{\prime} \in \mathscr{O}} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} h_{I^{\prime}}^{\varepsilon^{\prime}}, \quad \vec{f}=\sum_{I \in \mathscr{O} \in \operatorname{Sig}_{d}} \sum_{I} \vec{f}_{I}^{\varepsilon} h_{I}^{\varepsilon}
$$

so that

$$
\begin{aligned}
{[B, Q] \vec{f} } & =\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}}\left(B \vec{f}_{I}^{\varepsilon} Q h_{I}^{\varepsilon}-Q\left(B h_{I}^{\varepsilon}\right) \vec{f}_{I}^{\varepsilon}\right) \\
& =\sum_{I, I^{\prime} \in \mathscr{D},, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} \sum_{I^{\prime}}\left(B_{I^{\prime}}^{\varepsilon^{\prime}} h^{\varepsilon^{\prime}}\left(Q h_{I}^{\varepsilon}\right) \vec{f}_{I}^{\varepsilon}-B_{I^{\prime}}^{\varepsilon^{\prime}}\left(Q h_{I^{\prime}}^{\varepsilon^{\prime}} h_{I}^{\varepsilon}\right) \vec{f}_{I}^{\varepsilon}\right) \\
& =\sum_{I, I^{\prime} \in \mathscr{D},, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} \sum_{I^{\prime}}^{\varepsilon^{\prime}}\left(\left[h_{I^{\prime}}^{\varepsilon^{\prime}}, Q\right] h_{I}^{\varepsilon}\right) \vec{f}_{I}^{\varepsilon} .
\end{aligned}
$$

Clearly there is no contribution if $I \cap I^{\prime}=\emptyset$ and otherwise we have that

$$
\left[h_{I^{\prime}}^{\varepsilon^{\prime}}, Q\right] h_{I}^{\varepsilon}= \begin{cases}0 & I \subsetneq I^{\prime}  \tag{3.3}\\ \pm|I|^{-1 / 2} h_{\sigma(I)}^{\sigma(\varepsilon)}-Q\left(h_{I}^{\epsilon^{\prime}} h_{I}^{\epsilon}\right) & I=I^{\prime} \\ h_{\sigma(I)}^{\varepsilon^{\prime}} h_{\sigma(I)}^{\sigma(\varepsilon)} \pm|I|^{-1 / 2} h_{\sigma^{2}(I)}^{\sigma\left(\varepsilon^{\prime}\right)} & I^{\prime}=\sigma(I) \\ h_{I^{\prime}}^{\varepsilon^{\prime}} Q\left(h_{I}^{\varepsilon}\right)-Q\left(h_{I}^{\varepsilon} h_{I^{\prime}}^{\varepsilon^{\prime}}\right) & I^{\prime} \subsetneq I \text { and } I^{\prime} \neq \sigma(I)\end{cases}
$$

Note that we can disregard sign changes thanks to the unconditionality of Theorem 1.3, (2.1), and Proposition 1.6, and we will not comment
on this further in the proof. When $I=I^{\prime}$ we need to bound the two sums

$$
\begin{equation*}
\sum_{I \in \mathscr{D}} \sum_{\varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I}^{\varepsilon^{\prime}} \vec{f}_{I}^{\varepsilon}|I|^{-1 / 2} h_{\sigma(I)}^{\sigma(\varepsilon)} \text { and } Q\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I}^{\varepsilon^{\prime}} \vec{f}_{I}^{\varepsilon}|I|^{-1 / 2} h_{I}^{\psi_{\varepsilon^{\prime}}(\varepsilon)}\right) \tag{3.4}
\end{equation*}
$$

where $\psi_{\varepsilon^{\prime}}(\varepsilon)$ is the signature defined by

$$
h_{I}^{\psi_{\varepsilon^{\prime}}(\varepsilon)}=|I|^{\frac{1}{2}} h_{I}^{\varepsilon} h_{I}^{\varepsilon^{\prime}}
$$

which is cancellative if and only if $\varepsilon \neq \varepsilon^{\prime}$.
However, if $B \in \mathrm{BMO}_{W}^{p}$ then condition (b) in Theorem 1.3 tells us that for $\epsilon, \epsilon^{\prime}$ fixed and $\tilde{J}$ being the parent of $J \in \mathscr{D}$

$$
\sup _{J \in \sigma(\mathscr{D})}\left\|V_{J}\left(|\tilde{J}|^{-\frac{1}{2}} B_{\tilde{J}}^{\epsilon^{\prime}}\right) V_{\tilde{J}}^{-1}\right\| \lesssim \sup _{J \in \sigma(\mathscr{D})}\left\|V_{\tilde{J}}\left(|\tilde{J}|^{-\frac{1}{2}} B_{\tilde{J}}^{\epsilon^{\prime}}\right) V_{\tilde{J}}^{-1}\right\|<\infty
$$

so that the first sum in (3.4) can be estimated in a manner that is very similar to the proof of sufficiency in Proposition 1.6 (that is, using (2.1) twice).

Note that the second sum of (3.4) when $\varepsilon \neq \varepsilon^{\prime}$ is also "Haar multiplier like" and can be estimated in exactly the same way as the first sum in (3.4). On the other hand, when $\epsilon=\epsilon^{\prime}$ the second sum of (3.4) becomes

$$
Q\left(\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} B_{I}^{\varepsilon} \vec{f}_{I}^{\varepsilon} \frac{\chi_{I}}{|I|}\right)=Q\left(\pi_{B^{*}}\right)^{*} \vec{f} .
$$

However, by Corollary [2.5 we have that $B \in \mathrm{BMO}_{W}^{p}$ if and only if $\left(\pi_{B^{*}}\right)^{*}$ is bounded on $L^{p}(W)$.

We now look at the case when $I^{\prime}=\sigma(I)$ which clearly gives us two sums corresponding to the two terms in (3.3). For the first term, we obtain the sum

$$
\begin{aligned}
\sum_{I \in \mathscr{D}} \sum_{\varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{\sigma(I)}^{\varepsilon^{\prime}} h_{\sigma(I)}^{\varepsilon^{\prime}} h_{\sigma(I)}^{\sigma(\varepsilon)} \vec{f}_{I}^{\varepsilon} & =\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} B_{\sigma(I)}^{\sigma(\varepsilon)} \vec{f}_{I}^{\vec{\varepsilon}} \frac{\chi_{\sigma(I)}}{|\sigma(I)|} \\
& +\sum_{I \in \mathscr{D}} \sum_{\substack{\varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d} \\
\varepsilon^{\prime} \neq \sigma(\varepsilon)}}|I|^{-\frac{1}{2}} B_{\sigma(I)}^{\varepsilon^{\prime}} h_{\sigma(I)}^{\psi_{\varepsilon^{\prime}}(\sigma(\varepsilon))} \vec{f}_{I}^{\varepsilon}
\end{aligned}
$$

However, a simple computation gives us

$$
\sum_{I \in \mathscr{D}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} B_{\sigma(I)}^{\sigma(\varepsilon)} \vec{f}_{I}^{\varepsilon} \frac{\chi_{\sigma(I)}}{|\sigma(I)|}=\left(\pi_{B^{*}}\right)^{*} Q \vec{f}
$$

which is bounded on $L^{p}(W)$. Also, the second sum is again "Haar multiplier like" and can be estimated in easily in a manner that is similar to the proof of sufficiency for Proposition 1.6,

Furthermore, for the second sum in the two terms when $I^{\prime}=\sigma(I)$, we need to bound

$$
\sum_{I \in \mathscr{D} \varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{\sigma(I)}^{\varepsilon^{\prime}}|I|^{-1 / 2} h_{\sigma^{2}(I)}^{\sigma\left(\varepsilon^{\prime}\right)} \vec{f}_{I}^{\varepsilon}
$$

which yet again is "Haar multiplier like" and can be estimated in a manner that is similar to the proof of sufficiency for Proposition 1.6

To finally finish the proof of sufficiency we bound the triangular terms. First, if $I \supsetneq I^{\prime}$ then obviously $h_{I}^{\varepsilon}$ is constant on $I^{\prime}$. Thus,

$$
\begin{aligned}
\sum_{I^{\prime} \in \mathscr{D}} \sum_{I \supsetneq I^{\prime}} \sum_{\varepsilon, \varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} Q\left(h_{I}^{\varepsilon} h_{I^{\prime}}^{\varepsilon^{\prime}}\right) \vec{f}_{I}^{\vec{\varepsilon}} & =\sum_{I^{\prime} \in \mathscr{D}} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} Q\left(h_{I^{\prime}}^{\varepsilon^{\prime}}\right) \sum_{I \supsetneq I^{\prime}} \sum_{\varepsilon \in \operatorname{Sig}_{d}} \vec{f}_{I}^{\varepsilon} h_{I}^{\varepsilon} \\
& =\sum_{I^{\prime} \in \mathscr{D}} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} Q h_{I^{\prime}}^{\varepsilon^{\prime}} m_{I^{\prime}} \vec{f} \\
& =Q \pi_{B} \vec{f} .
\end{aligned}
$$

Now clearly $h_{I^{\prime}}^{\varepsilon^{\prime}} Q\left(h_{I}^{\varepsilon}\right)=0$ if $I^{\prime} \cap \sigma(I)=\emptyset$. Furthermore, since $I \supsetneq I^{\prime}$ and $I^{\prime} \neq \sigma(I)$, we must have $\sigma(I) \supsetneq I^{\prime}$ so that

$$
\begin{aligned}
\sum_{I^{\prime} \in \mathscr{O}} \sum_{I \supsetneq I^{\prime}} \sum_{\varepsilon, \varepsilon^{\prime} \in \mathrm{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} h_{I^{\prime}}^{\varepsilon^{\prime}} Q\left(h_{I}^{\varepsilon}\right) \vec{f}_{I}^{\varepsilon} & =\sum_{I^{\prime} \in \mathscr{D}} \sum_{\varepsilon^{\prime} \in \mathrm{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} h_{I^{\prime}}^{\varepsilon^{\prime}} \sum_{I: \sigma(I) \supsetneq I^{\prime}} \sum_{\varepsilon \in \mathrm{Sig}_{d}} h_{\sigma(I)}^{\sigma(\varepsilon)} \vec{f}_{I}^{\varepsilon} \\
& =\sum_{I^{\prime} \in \mathscr{O}} \sum_{\varepsilon^{\prime} \in \operatorname{Sig}_{d}} B_{I^{\prime}}^{\varepsilon^{\prime}} h_{I^{\prime}}^{\varepsilon^{\prime}} m_{I^{\prime}}(Q \vec{f}) \\
& =\pi_{B} Q \vec{f}
\end{aligned}
$$

which is obviously bounded on $L^{p}(W)$. The proof of sufficiency is now complete.

We now prove (1.7).

Proof of (1.7). Note that the simple arguments used to prove sufficiency in Proposition 1.6 combined with (2.2) and (2.3) shows that (as
was noticed in [3, [2])

$$
\left\|T_{A}\right\|_{L^{2}(W) \rightarrow L^{2}(W)} \lesssim\|W\|_{\mathrm{A}_{2}}^{\frac{3}{2}} \log \|W\|_{\mathrm{A}_{2}}\left(\sup _{I \in \mathscr{D}, \varepsilon \in \operatorname{Sig}_{d}}\left\|V_{I} A_{I}^{\varepsilon} V_{I}^{-1}\right\|\right)
$$

Also a careful reading of the proof of sufficiency in Theorem 1.1 reveals that

$$
\begin{aligned}
\|[T, B]\|_{L^{2}(W) \rightarrow L^{2}(W)} & \lesssim\|Q\|_{L^{2}(W) \rightarrow L^{2}(W)} \max \left\{\left\|\pi_{B}\right\|_{L^{2}(W) \rightarrow L^{2}(W)},\left\|\pi_{B^{*}}\right\|_{L^{2}\left(W^{-1}\right) \rightarrow L^{2}\left(W^{-1}\right)}\right\} \\
& +\|W\|_{\mathrm{A}_{2}}^{\frac{3}{2}} \log \|W\|_{\mathrm{A}_{2}}\|B\|_{*}^{\frac{1}{2}}
\end{aligned}
$$

where again $\|B\|_{*}$ is the canonical supremum in condition (b) of Theorem 1.3
3.3. Proof of necessity. For the proof of necessity in Theorem 1.1 we will use some simple ideas from [22]. We will in fact prove the following more general result for commutators with kernels considered in [22].

Theorem 3.3. Let $K: \mathbb{R}^{d} \backslash\{0\} \rightarrow \mathbb{R}$ be not identically zero, be homogenous of degree $-d$, have mean zero over the unit sphere $\partial \mathbb{B}_{d}$, and satisfy $K \in C^{\infty}\left(\partial \mathbb{B}_{d}\right)$ (so in particular $K$ could be any of the Riesz kernels). If $T$ is the (convolution) CZO associated to $K$, then we have that $[T, B]$ being bounded on $L^{p}(W)$ implies that $B \in B M O_{W}^{p}$.

Proof. First note that it is enough to prove that $B$ satisfies the first condition in the definition of $\mathrm{BMO}_{W}^{p}$ when $2 \leq p<\infty$ and $[T, B]$ is bounded on $L^{p}(W)$ since

$$
\left(W^{\frac{1}{p}}[T, B] W^{-\frac{1}{p}}\right)^{*}=-W^{-\frac{1}{p}}\left[T, B^{*}\right] W^{\frac{1}{p}}
$$

which will allow us to immediately conclude that the second condition in the definition of $\mathrm{BMO}_{W}^{p}$ is true when $1 \leq p<2$ as $W^{1-p^{\prime}}$ is a matrix $\mathrm{A}_{p^{\prime}}$ weight and $2 \leq p^{\prime}<\infty$. Now by assumption, there exists $z_{0} \neq 0$ and $\delta>0$ where $\frac{1}{K(x)}$ is smooth on $\left|x-z_{0}\right|<\sqrt{d} \delta$, and thus can be expressed as an absolutely convergent Fourier series

$$
\frac{1}{K(x)}=\sum a_{n} e^{i v_{n} \cdot x}
$$

for $\left|x-z_{0}\right|<\sqrt{d} \delta$ (where the exact nature of the vectors $v_{n}$ is irrelevant.) Set $z_{1}=\delta^{-1} z_{0}$. Thus, if $\left|x-z_{1}\right|<\sqrt{d}$, then we have by homogeneity

$$
\frac{1}{K(x)}=\frac{\delta^{-d}}{K(\delta x)}=\delta^{-d} \sum a_{n} e^{i v_{n} \cdot(\delta x)}
$$

Now for any cube $Q=Q\left(x_{0}, r\right)$ of side length $r$ and center $x_{0}$, let $y_{0}=x_{0}-r z_{1}$ and $Q^{\prime}=Q\left(y_{0}, r\right)$ so that $x \in Q$ and $y \in Q^{\prime}$ implies that

$$
\left|\frac{x-y}{r}-z_{1}\right| \leq\left|\frac{x-x_{0}}{r}\right|+\left|\frac{y-y_{0}}{r}\right| \leq \sqrt{d} .
$$

Let

$$
S_{Q}(x)=\chi_{Q}(x) \frac{V_{Q}^{-1}\left(B^{*}(x)-m_{Q^{\prime}} B^{*}\right) W^{\frac{1}{p}}(x)}{\left\|V_{Q}^{-1}\left(B^{*}(x)-m_{Q^{\prime}} B^{*}\right) W^{\frac{1}{p}}(x)\right\|}
$$

so that

$$
\begin{align*}
\frac{1}{r^{d}} & \left\|\int_{\mathbb{R}^{d}} W^{\frac{1}{p}}(x)(B(x)-B(y)) V_{Q}^{-1} \frac{r^{d} K(x-y)}{K\left(\frac{x-y}{r}\right)} S_{Q}(x) \chi_{Q^{\prime}}(y) d y\right\|  \tag{3.5}\\
& =\chi_{Q}(x) \frac{1}{r^{d}}\left\|\int_{Q^{\prime}} W^{\frac{1}{p}}(x)(B(x)-B(y)) V_{Q}^{-1} \frac{\left(W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right)^{*}}{\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right\|} d y\right\| \\
& =\chi_{Q}(x)\left\|\frac{W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\left(W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right)^{*}}{\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right\|}\right\| \\
& =\chi_{Q}(x)\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right\| .
\end{align*}
$$

However,

$$
\begin{aligned}
(3.5) & \leq \sum\left|a_{n}\right|\left\|W^{\frac{1}{p}}(x)\left(\int_{\mathbb{R}^{d}}(B(x)-B(y)) K(x-y) e^{-i \frac{\delta}{r} v_{n} \cdot y} V_{Q}^{-1} \chi_{Q^{\prime}}(y) d y\right) S_{Q}(x) e^{i \frac{\delta}{r} v_{n} \cdot x}\right\| \\
& =\sum\left|a_{n}\right|\left\|W^{\frac{1}{p}}(x)\left([T, B] g_{n}\right)(x) f_{n}(x)\right\| \\
& \leq \sum\left|a_{n}\right|\left\|W^{\frac{1}{p}}(x)[T, B] g_{n}(x)\right\|
\end{aligned}
$$

where

$$
g_{n}(y)=e^{-i \frac{\delta}{r} v_{n} \cdot y} V_{Q}^{-1} \chi_{Q^{\prime}}(y), \quad f_{n}(x)=S_{Q}(x) e^{i \frac{\delta}{r} v_{n} \cdot x}
$$

and where the last inequality follows from the fact that $\left\|f_{n}(x)\right\| \leq 1$ for a.e. $x \in \mathbb{R}^{d}$.

But as $\left|x_{0}-y_{0}\right|=r \delta^{-1} z_{0}$, we can pick some $C>1$ only depending on $K$ where $\tilde{Q}=Q\left(x_{0}, C r\right)$ satisfies $Q \cup Q^{\prime} \subseteq \tilde{Q}$. Combining this with
the previous estimates, we have from the absolute summability of the $a_{n}^{\prime} s$ and the boundedness of $[T, B]$ that

$$
\begin{aligned}
\left(\int_{Q}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{Q^{\prime}} B\right) V_{Q}^{-1}\right\|^{p} d x\right)^{\frac{1}{p}} & \leq \sum\left|a_{n}\right|\left\|W^{\frac{1}{p}}[T, B] g_{n}\right\|_{L^{p}} \\
& \leq \sum\left|a_{n}\right|\left\|W^{\frac{1}{p}}[T, B] W^{-\frac{1}{p}}\left(W^{\frac{1}{p}} g_{n}\right)\right\|_{L^{p}} \\
& \leq \sup _{n}\left\|W^{\frac{1}{p}} g_{n}\right\|_{L^{p}} \\
& \leq\left\|\chi_{Q^{\prime}} W^{\frac{1}{p}} V_{Q}^{-1}\right\|_{L^{p}} \\
& \lesssim\|W\|_{A_{p}}^{\frac{1}{p}}\left\|\chi_{\tilde{Q}} W^{\frac{1}{p}} V_{\tilde{Q}}^{-1}\right\|_{L^{p}} \\
& \lesssim|Q|^{\frac{1}{p}}\|W\|_{A_{p}}^{\frac{1}{p}}
\end{aligned}
$$

where the second to last inequality is due to (1.4). The proof is now complete thanks to Lemma 3.1.

Lastly in this section we will give quick proofs of Corollary 1.5 and Proposition 1.4 .

Proof of Corollary 1.5. If $B \in \mathrm{BMO}_{W}^{p}$ then from the proof of Theorem 3.3 we have

$$
\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I} B\right) V_{I}^{-1}\right\|^{p} d x<\infty .
$$

On the other hand, again since

$$
\left(W^{\frac{1}{p}}[T, B] W^{-\frac{1}{p}}\right)^{*}=-W^{-\frac{1}{p}}\left[T, B^{*}\right] W^{\frac{1}{p}}
$$

we can again use the proof of Theorem 3.3] to get that the dual condition

$$
\sup _{\substack{I \subset \mathbb{R}^{d} \\ I \text { is a cube }}} \frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{p}}(x)\left(B^{*}(x)-m_{I} B^{*}\right)\left(V_{I}^{\prime}\right)^{-1}\right\|^{p^{\prime}} d x<\infty
$$

is true since $W^{1-p^{\prime}}$ is a matrix $\mathrm{A}_{p^{\prime}}$ weight. The proof is now complete.

Proof of Proposition 1.4. It is well known (see [24]) that if $\mathscr{D}^{t}=$ $\left\{2^{-k}\left([0,1)^{d}+m+(-1)^{k} t\right): k \in \mathbb{Z}, m \in \mathbb{Z}^{d}\right\}$, then for any cube $I$ there exists $1 \leq t \leq 2^{d}$ and $I_{t} \in \mathscr{D}^{t}$ such that $I \subset I_{t}$ and $\ell\left(I_{t}\right) \leq 6 \ell(Q)$. Thus, Lemma 3.1 gives us

$$
\begin{aligned}
\frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I} B\right) V_{I}^{-1}\right\|^{p} d x & \lesssim \frac{1}{|I|} \int_{I}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I_{t}} B\right) V_{I}^{-1}\right\|^{p} d x \\
& \lesssim \frac{1}{\left|I_{t}\right|} \int_{I_{t}}\left\|W^{\frac{1}{p}}(x)\left(B(x)-m_{I_{t}} B\right) V_{I_{t}}^{-1}\right\|^{p} d x
\end{aligned}
$$

which completes the proof.

## 4. Counterexamples and other quantitative estimates

In this last section we will produce the counterexamples mentioned in the introduction and additionally prove quantitative matrix weighted bounds for maximal functions and sparse operators.
4.1. Counterexamples. For the rest of this section let

$$
A:=\left(\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right) \text { and } W:=\left(\begin{array}{cc}
|x|^{\alpha} & 0 \\
0 & |x|^{-\alpha}
\end{array}\right)
$$

for $x \in \mathbb{R}$, where $0<\alpha<1$ so that $W$ is trivially a matrix $\mathrm{A}_{2}$ weight on $\mathbb{R}$ since $W$ is diagonal. Also in this section let $\mathscr{D}$ be the standard dyadic grid on $\mathbb{R}$.

Proposition 4.1. There exists a sequence $\left\{A_{I}\right\}_{I \in \mathscr{D}}$ where $T_{A}$ is not bounded on $L^{2}(W)$.

Proof. We will in fact prove that there exists a constant sequence $\left\{A_{I}\right\}_{I \in \mathscr{D}}$ with the above property, and in particular let $\left\{A_{I}\right\}_{I \in \mathscr{D}}$ be the constant sequence $A_{I}=A$. For $I_{N}=\left[0,2^{-N}\right)$ we have

$$
\begin{aligned}
\lim _{N \rightarrow \infty}\left\|\left(m_{I_{N}} W\right)^{\frac{1}{2}} A\left(m_{I_{N}} W\right)^{-\frac{1}{2}}\right\| & \geq \lim _{N \rightarrow \infty}\left\|\left(m_{I_{N}} W\right)^{\frac{1}{2}} A\left(m_{I_{N}} W^{-1}\right)^{\frac{1}{2}}\right\| \\
& =\lim _{N \rightarrow \infty} \frac{2^{\alpha N}}{(1-\alpha)^{2}}=\infty
\end{aligned}
$$

An application of Proposition 1.6 now says that $T_{A}$ is not bounded on $L^{2}(W)$.

We will now show that $B \in \mathrm{BMO}$ and $W$ being a matrix $\mathrm{A}_{2}$ weight is not sufficient for $\pi_{B}$ with respect to $\mathscr{D}$ to be bounded on $L^{2}(W)$.

Proposition 4.2. There exists a matrix function $B$ with scalar $B M O$ entries where $\pi_{B}$ is not bounded on $L^{2}(W)$, and consequently $[T, B]$ is not bounded on $L^{2}(W)$ where $T$ is any of the Riesz transforms.

Proof. Let $B(x):=(\log |x|) A$ and let $J_{N}=\left[2^{-N-1}, 2^{-N}\right)$ for $N \in \mathbb{N}$. Now assume that $N \in \mathbb{N}$ is in fact large enough where

$$
\frac{1}{\left|J_{N}\right|} \sum_{I \in \mathscr{D}\left(J_{N}\right)}\left|b_{I}\right|^{2}>\frac{1}{2}\|b\|_{\mathrm{BMO}} .
$$

Let $\vec{f}_{N}:=\chi_{J_{N}} W^{-\frac{1}{2}} \vec{e}$ where

$$
\vec{e}:=\binom{1}{0} .
$$

By (2.1) we have

$$
\begin{aligned}
\left\|\pi_{B} W^{-\frac{1}{2}} \vec{f}_{N}\right\|_{L^{2}(W)}^{2} & \gtrsim \sum_{I \in \mathscr{D}\left(J_{N}\right)}\left|\left(m_{I} W\right)^{\frac{1}{2}} B_{I} m_{I}\left(W^{-1}\right) A \vec{e}\right|^{2} \\
& =\sum_{I \in \mathscr{D}\left(J_{N}\right)}\left|b_{I}\right|^{2}\left|\left(m_{I} W\right)^{\frac{1}{2}} A m_{I}\left(W^{-1}\right) \vec{e}\right|^{2} \\
& \gtrsim \frac{\|b\|_{\mathrm{BMO}}}{2} 2^{3 \alpha N} .
\end{aligned}
$$

However,

$$
\begin{aligned}
\left\|\vec{f}_{N}\right\|_{L^{2}}^{2} & =\int_{J_{N}}\left|W^{-\frac{1}{2}}(t) \vec{e}\right|^{2} d t \\
& \approx 2^{\alpha N}
\end{aligned}
$$

which shows that $\pi_{B}$ can not be bounded on $L^{2}\left(\mathbb{R} ; \mathbb{C}^{2}\right)$.
4.2. Maximal function and sparse operator bounds. We will end this paper with some quantitative weighted norm inequalities that were mentioned earlier in the paper. Now let

$$
M_{W}^{\prime} \vec{f}(x)=\sup _{I \ni x} \frac{1}{|I|} \int_{I}\left|\left(m_{I}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y
$$

and

$$
M_{W} \vec{f}(x)=\sup _{I \ni x} \frac{1}{|I|} \int_{I}\left|W^{\frac{1}{2}}(x) W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y
$$

where the supremum is over dyadic cubes $I$ taken from some fixed dyadic grid. Note that the proofs of the next three results are slight modifications to the corresponding ones in [8] (which is where $M_{W}$ and a slight variation of $M_{W}^{\prime}$ were first defined).

Lemma 4.3. $M_{W}^{\prime}$ is bounded on $L^{2}$ and in particular

$$
\left\|M_{W}^{\prime}\right\|_{L^{2} \rightarrow L^{2}}^{2} \lesssim\|W\|_{A_{2}}
$$

Proof. By the (scalar) reverse Hölder inequality, as before, we can pick $\epsilon \approx\|W\|_{\mathrm{A}_{2}}^{-1}$ where

$$
\left(\frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{2}}(y)\left(m_{I}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\right\|^{2+\epsilon} d y\right)^{\frac{1}{2+\epsilon}} \lesssim\left(\frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{2}}(y)\left(m_{I}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\right\|^{2} d y\right)^{\frac{1}{2}} \lesssim 1
$$

Thus by Hölder's inequality we have

$$
\begin{aligned}
M_{W}^{\prime} \vec{f}(x) & \leq \sup _{I \ni x}\left(\frac{1}{|I|} \int_{I}\left\|W^{-\frac{1}{2}}(y)\left(m_{I}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\right\|^{2+\epsilon} d y\right)^{\frac{1}{2+\epsilon}}\left(\frac{1}{|I|} \int_{I}|\vec{f}(y)|^{\frac{2+\epsilon}{1+\epsilon}} d y\right)^{\frac{1+\epsilon}{2+\epsilon}} \\
& \lesssim\left(M\left(|\vec{f}|^{\frac{2+\epsilon}{1+\epsilon}}\right)(x)\right)^{\frac{1+\epsilon}{2+\epsilon}}
\end{aligned}
$$

where $M$ is the standard maximal function with respect to cubes.
Finally, as before, the usual $L^{1+\delta} \rightarrow L^{1+\delta}$ maximal function bound given by the Marcinkewicz interpolation theorem gives us that

$$
\int_{\mathbb{R}^{d}}\left|M_{W}^{\prime} \vec{f}(x)\right|^{2} d x \leq \int_{\mathbb{R}^{d}}\left(M\left(|\vec{f}|^{\frac{2+\epsilon}{1+\epsilon}}\right)(x)\right)^{\frac{2+2 \epsilon}{2+\epsilon}} d x \lesssim \epsilon^{-1}\|\vec{f}\|_{L^{2}}^{2}
$$

which completes the proof as $\epsilon^{-1} \approx\|W\|_{\mathrm{A}_{2}}$.
Lemma 4.4. If $Q$ is a cube and

$$
N_{Q}(x)=\sup _{x \in R \subseteq Q}\left\|W^{\frac{1}{2}}(x)\left(m_{R}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|
$$

then

$$
\int_{Q}\left(N_{Q}(x)\right)^{2} d x \lesssim|Q|\|W\|_{A_{2}}
$$

Proof. We truncate $W$ as in [3] p. 1733. More precisely, write

$$
W(x)=\sum_{j=1}^{n} \lambda_{j}(x) P_{E_{j}(x)}
$$

where the $\lambda_{j}(x)$ 's are the eigenvalues of $W(x)$ with corresponding eigenspaces $E_{j}(x)$ and $P_{E_{j}(x)}$ is the orthogonal projection onto $E_{j}(x)$. Now for $n \in \mathbb{N}$, let $E_{1}^{n}(x), E_{2}^{n}(x)$, and $E_{3}^{n}(x)$ be the span of the eigenspaces corresponding to the eigenvalues $\lambda_{j}(x) \leq n^{-1}, n^{-1}<$ $\lambda_{j}(x)<n$, and $\lambda_{j}(x) \geq n$, respectively. Finally, define the truncation $W_{n}$ as

$$
W_{n}(x)=n^{-1} P_{E_{1}^{n}(x)}+P_{E_{2}^{n}(x)} W(x) P_{E_{2}^{n}(x)}+n P_{E_{3}^{n}(x)}
$$

WEIGHTED INEQUALITIES, COMMUTATORS, AND PARAPRODUCTS 31
It is then easy to see that $W_{n} \rightarrow W$ and $W_{n}^{-1} \rightarrow W^{-1}$ pointwise a.e., $\left\|W_{n}\right\|_{\mathrm{A}_{2}} \lesssim\|W\|_{\mathrm{A}_{2}}$ for each $n$, and $W_{n}, W_{n}^{-1} \leq n \mathrm{Id}_{d \times d}$ (see [3]). If

$$
N_{Q}^{n}(x)=\sup _{x \in R \subseteq Q}\left\|W_{n}^{\frac{1}{2}}(x)\left(m_{R}\left(W_{n}^{-1}\right)\right)^{\frac{1}{2}}\right\|
$$

then $\left\|\left(m_{R}\left(W_{n}^{-1}\right)\right)^{-\frac{1}{2}}-\left(m_{R}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\right\| \rightarrow 0$ as $n \rightarrow \infty$ by the dominated convergence theorem since clearly

$$
\left\|W_{n}^{-1}-W^{-1}\right\| \leq\left\|W_{n}^{-1}\right\|+\left\|W^{-1}\right\| \leq 2 \max \left\{1,\left\|W^{-1}\right\|\right\} .
$$

Thus, we have that

$$
\int_{Q}\left(N_{Q}(x)\right)^{2} d x \leq \int_{Q}\left(\liminf _{n \rightarrow \infty} N_{Q}^{n}(x)\right)^{2} d x \leq \liminf _{n \rightarrow \infty} \int_{Q}\left(N_{Q}^{n}(x)\right)^{2} d x .
$$

Obviously $N_{Q}^{n}(x) \leq n^{2}$ so trivially there exists $B_{n}$ such that

$$
\int_{Q}\left(N_{Q}^{n}(x)\right)^{2} d x \lesssim B_{n}|Q|
$$

Putting this all together, it is enough to show that $B \lesssim\|W\|_{\mathrm{A}_{2}}$ if we assume that

$$
\int_{Q}\left(N_{Q}(x)\right)^{2} d x \lesssim B|Q|
$$

(or in other words we show in fact that $B_{n} \lesssim\|W\|_{\mathrm{A}_{2}}$.) To that end, let $\left\{R_{j}\right\}$ be maximal subcubes of $Q$ satisfying

$$
\left\|\left(m_{Q}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\left(m_{R_{j}}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|>C
$$

for some large $C$ independent of $W$ to be determined.

Note that if $x \in Q \backslash \cup_{j} R_{j}$ then for any dyadic cube $x \in R \subset Q$ we have

$$
\begin{aligned}
\left\|W^{\frac{1}{2}}(x)\left(m_{R}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\| & \leq\left\|W^{\frac{1}{2}}(x)\left(m_{Q}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|\left\|\left(m_{Q}\left(W^{-1}\right)\right)^{-\frac{1}{2}}\left(m_{R}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\| \\
& \leq C\left\|W^{\frac{1}{2}}(x)\left(m_{Q}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\| .
\end{aligned}
$$

so that

$$
\int_{Q \backslash \cup_{j} R_{j}}\left(N_{Q}(x)\right)^{2} d x \leq C \int_{Q}\left\|W^{\frac{1}{2}}(x)\left(m_{Q}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|^{2} d x \leq C\|W\|_{\mathrm{A}_{2}}|Q|
$$

On the other hand,

$$
\begin{aligned}
C^{2} \sum_{j}\left|R_{j}\right| & \leq \sum_{j}\left|R_{j}\right|\left\|\left(m_{Q} W^{-1}\right)^{-\frac{1}{2}}\left(m_{R_{j}}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|^{2} \\
& \lesssim \sum_{j} \int_{R_{j}}\left\|W^{-\frac{1}{2}}(x)\left(m_{Q} W^{-1}\right)^{-\frac{1}{2}}\right\|^{2} d x \lesssim|Q| .
\end{aligned}
$$

Thus for $C$ large enough independent of $W$ we have $\sum_{j}\left|R_{j}\right| \leq \frac{1}{2}|Q|$.
Clearly by the definition of $R_{j}$ and their maximality we can assume for each $x \in R_{j}$ that $N_{Q}(x)=N_{R_{j}}(x)$ since otherwise $N_{Q}(x) \lesssim$ $C\left\|W^{\frac{1}{2}}(x)\left(m_{Q}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|$. Thus without loss of generality

$$
\int_{\cup_{j} R_{j}}\left(N_{Q}(x)\right)^{2} d x=\sum_{j} \int_{R_{j}}\left(N_{R_{j}}(x)\right)^{2} d x \leq B \sum_{j}\left|R_{j}\right| \leq \frac{1}{2} B|Q| .
$$

Finally this implies that there exists $C$ independent of $W$ where $B \leq$ $\frac{1}{2} B+C\|W\|_{\mathrm{A}_{2}}$ which completes the proof.

Theorem 4.5. $M_{W}$ is bounded on $L^{2}$ and in fact

$$
\left\|M_{W}\right\|_{L^{2} \rightarrow L^{2}} \lesssim\|W\|_{A_{2}} .
$$

Proof. For each $x \in \mathbb{R}^{d}$ pick (and fix) some dyadic cube $R_{x}$ such that

$$
\begin{aligned}
\frac{1}{2} M_{W}(\vec{f})(x) & \leq \frac{1}{\left|R_{x}\right|} \int_{R_{x}}\left|W^{\frac{1}{2}}(x) W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y \\
& \leq\left\|W^{\frac{1}{2}}(x)\left(m_{R_{x}}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|\left(\frac{1}{\left|R_{x}\right|} \int_{R_{x}}\left|\left(m_{R_{x}}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y\right)
\end{aligned}
$$

For $x \in \mathbb{R}^{d}$ pick $j \in \mathbb{Z}$ where

$$
\begin{equation*}
2^{j} \leq \int_{R_{x}}\left|\left(m_{R_{x}}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y<2^{j+1} \tag{4.2}
\end{equation*}
$$

and let $S_{j}$ be the collection of all cubes $R=R_{x}$ for all $x \in \mathbb{R}^{d}$ that are maximal and satisfy (4.2) (note that the Cauchy Schwarz inequality implies that such a maximal cube exists). We therefore have that for every $x \in \mathbb{R}^{d}, x \in R_{x} \subseteq S$ for some $S \in S_{j}$ where $j=j_{x} \in \mathbb{Z}$. Note that if $R_{x}$ satisfies (4.2) for $j \in \mathbb{Z}$ then
$\frac{1}{\left|R_{x}\right|} \int_{R_{x}}\left|\left(m_{R_{x}}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y \leq \frac{2}{|S|} \int_{S}\left|\left(m_{S}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y$
since otherwise trivially (4.2) is violated.

Now if $x \in \mathbb{R}^{d}$ then pick $j=j_{x}$ as before and pick $S \in S_{j}$ with $R_{x} \subseteq S \in S_{j}$ so

$$
\begin{aligned}
M_{W}(\vec{f})(x) & \leq 2\left\|W^{\frac{1}{2}}(x)\left(m_{R_{x}}\left(W^{-1}\right)\right)^{\frac{1}{2}}\right\|\left(\frac{1}{\left|R_{x}\right|} \int_{R_{x}}\left|\left(m_{R_{x}}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y\right) \\
& \leq 4 N_{S}(x)\left(\frac{2}{|S|} \int_{S}\left|\left(m_{S}\left(W^{-1}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y\right) \\
& \leq 42^{j+1} N_{S}(x)
\end{aligned}
$$

so that finally the previous two lemmas give us that

$$
\begin{aligned}
\int_{\mathbb{R}^{d}}\left|M_{W} \vec{f}(x)\right|^{2} d x & \lesssim \sum_{j \in \mathbb{Z}, S \in S_{j}} 2^{2 j} \int_{S}\left(N_{S}(x)\right)^{2} d x \\
& \lesssim\|W\|_{\mathrm{A}_{2}} \sum_{j \in \mathbb{Z}} 2^{2 j}\left|\bigsqcup_{S \in S_{j}} S\right| \\
& \leq\|W\|_{\mathrm{A}_{2}} \sum_{j \in \mathbb{Z}} 2^{2 j}\left|\left\{x: M_{W}^{\prime} \vec{f}(x)>2^{j}\right\}\right| \\
& \approx\|W\|_{\mathrm{A}_{2}}\left\|M_{W}^{\prime} \vec{f}\right\|_{L^{2}}^{2} \\
& \lesssim\|W\|_{\mathrm{A}_{2}}^{2}\|\vec{f}\|_{L^{2}}^{2}
\end{aligned}
$$

which completes the proof.

Interestingly, note that Lemma 4.3 is sharp with respect to $\|W\|_{\mathrm{A}_{2}}$ since otherwise we could get a better $\|W\|_{\mathrm{A}_{2}}$ bound for $\left\|M_{W}\right\|_{L^{2} \rightarrow L^{2}}$ (which is known to be sharp in the scalar setting, and thus the matrix setting). Despite this, the following simple result says that we can legitimately consider $M_{W}^{\prime}$ to be "the" universal $p=2$ matrix weighted maximal function corresponding to $W^{-1}$.

Proposition 4.6. $M_{W}^{\prime}$ is weak $(2,2)$ for any (not necessarily $A_{2}$ ) matrix weight $W$.

Proof. Let $\lambda>0$ and let $\left\{I_{j}\right\}$ be the collection of maximal dyadic cubes such that

$$
\frac{1}{\left|I_{j}\right|} \int_{I_{j}}\left|\left(m_{I_{j}} W^{-1}\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y>\lambda
$$

so as usual $\left\{x: M_{W}^{\prime} \vec{f}(x)>\lambda\right\}=\bigsqcup_{j} I_{j}$. Then

$$
\begin{aligned}
\sum_{j}\left|I_{j}\right| & =\sum_{j} \frac{\left|I_{j}\right|^{2}}{\left|I_{j}\right|} \leq \frac{1}{\lambda^{2}} \sum_{j} \frac{1}{\left|I_{j}\right|}\left(\int_{I_{j}}\left|\left(m_{I_{j}} W^{-1}\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y) \vec{f}(y)\right| d y\right)^{2} \\
& \leq \frac{1}{\lambda^{2}} \sum_{j}\left(\frac{1}{\left|I_{j}\right|} \int_{I_{j}}\left\|\left(m_{I_{j}} W^{-1}\right)^{-\frac{1}{2}} W^{-\frac{1}{2}}(y)\right\|^{2} d y\right)\left(\int_{I_{j}}|\vec{f}(y)|^{2} d y\right) \\
& \lesssim \frac{1}{\lambda^{2}} \sum_{j} \int_{I_{j}}|\vec{f}(y)|^{2} d y \leq \frac{\|\vec{f}\|_{L^{2}}^{2}}{\lambda^{2}} .
\end{aligned}
$$

Note that very similar maximal functions can be defined when $p \neq 2$ and similar weighted norm inequalities can be proved for these maximal functions (see [21] for proofs, where in fact fractional matrix weighted maximal functions are studied in detail and are applied to the study of matrix weighted norm inequalities for fractional integral operators and related matrix weighted Poincare and Sobolev inequalities.)

Lastly we will give a very simple "maximal function" proof of the matrix weighted norm inequalities from [4] for sparse operators, which provides a simpler proof that avoids the Carleson embedding theorem and that is similar to the by now classical proof from [10]. Also, note that in the scalar weighted case, the $L^{2}(W)$ bound of sparse operators (see the definition below) has linear $\|w\|_{\mathrm{A}_{2}}$ dependence (see [10] for the very easy maximal function proof in the scalar setting).

Finally, despite it's relative ease, note that this proof clearly highlights one of the severe challenges in using maximal functions (in any way, shape, or form) to prove sharp matrix weighted norm inequalities: the absence of $L^{2}$ bounds independent of $W$ for the universal matrix weighted maximal function makes scalar arguments much less efficient in the matrix weighted setting.

Let $\mathcal{G} \subset \mathscr{D}$ be a "sparse" collection in the sense that for any $I \in \mathcal{G}$ we have

$$
\sum_{J \in \operatorname{ch}_{\mathcal{G}}(I)}|J| \leq \frac{1}{2}|I|
$$

where $\operatorname{ch}_{\mathcal{G}}(I)$ are the children of $I$ that are also members of $\mathcal{G}$. Furthermore, define the sparse operator $S=S_{\mathcal{G}}$ by

$$
S \vec{f}=\sum_{I \in \mathcal{G}} m_{I} \vec{f} \chi_{I} .
$$

Proposition 4.7. If $W$ is a matrix $A_{2}$ weight and $S$ is a sparse operator then

$$
\|S\|_{L^{2}(W) \rightarrow L^{2}(W)} \lesssim\|W\|_{A_{2}}^{\frac{3}{2}} .
$$

Proof. Let $\vec{f}, \vec{g} \in L^{2}$. For any $I \in \mathcal{G}$ let $E_{I}$ be defined by

$$
E_{I}=I \backslash\left(\bigcup_{J \in \mathrm{ch}_{\mathcal{G}}(I)} J\right)
$$

so that clearly $\left\{E_{I}\right\}_{I \in \mathcal{G}}$ is a disjoint collection of measurable sets satisfying (by the sparseness condition) $2\left|E_{I}\right| \geq|I|$. Then

$$
\begin{aligned}
& \left|\left\langle W^{\frac{1}{2}} S\left(W^{-\frac{1}{2}} \vec{f}\right), \vec{g}\right\rangle_{L^{2}}\right| \\
& =\left|\left\langle S\left(W^{-\frac{1}{2}} \vec{f}\right), W^{\frac{1}{2}} \vec{g}\right\rangle_{L^{2}}\right| \\
& \leq \sum_{I \in \mathcal{G}}|I|\left|\left\langle m_{I}\left(W^{-\frac{1}{2}} \vec{f}\right), m_{I}\left(W^{\frac{1}{2}} \vec{g}\right)\right\rangle_{\mathbb{C}^{n}}\right| \\
& \leq\|W\|_{\mathrm{A}_{2}}^{\frac{1}{2}} \sum_{I \in \mathcal{G}}|I|\left(m_{I}\left|\left(m_{I}\left(W^{-\frac{1}{2}}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}} \vec{f}\right|\right)\left(m_{I}\left|\left(m_{I} W^{\frac{1}{2}}\right)^{-\frac{1}{2}}\left(W^{\frac{1}{2}} \vec{g}\right)\right|\right) \\
& \leq 2\|W\|_{A_{2}}^{\frac{1}{2}} \sum_{I \in \mathcal{G}}\left|E_{I}\right|\left(m_{I}\left|\left(m_{I}\left(W^{-\frac{1}{2}}\right)\right)^{-\frac{1}{2}} W^{-\frac{1}{2}} \vec{f}\right|\right)\left(m_{I}\left|\left(m_{I} W^{\frac{1}{2}}\right)^{-\frac{1}{2}}\left(W^{\frac{1}{2}} \vec{g}\right)\right|\right) \\
& \leq 2\|W\|_{\mathrm{A}_{2}}^{\frac{1}{2}} \sum_{I \in \mathcal{G}} \int_{E_{I}} M_{W}^{\prime} \vec{f}(x) M_{W^{-1}}^{\prime} \vec{g}(x) d x \\
& \lesssim\|W\|_{A_{2}}^{\frac{3}{2}}\|\vec{f}\|_{L^{2}}\|\vec{g}\|_{L^{2}} .
\end{aligned}
$$

## 5. Acknowlegements

The first author would like to thank Kelly Bickel and Brett Wick for their interesting discussions regarding sharp matrix weighted norm inequalities.

## References

[1] J. Alvarez, R. Bagby, and D. Kurtz, 'Weighted estimates for commutators of linear operators', Studia Math. 104 (1993), 195-209.
[2] O. Beznosova, 'Linear bound for the dyadic paraproduct on weighted Lebesgue space $L_{2}(w), '$ J. Funct. Anal. 255 (2008), 994-1007.
[3] K. Bickel, S. Petermichl, and B. Wick, 'Bounds for the Hilbert Transform with Matrix $A_{2}$ Weights,' J. Funct. Anal. 270 (2016), 17191743.
[4] K. Bickel and B. Wick, 'A Study of the Matrix Carleson Embedding Theorem with Applications to Sparse Operators,' Preprint available at http://arxiv.org/abs/1503.06493.
[5] K. Bickel and B. Wick, 'Well-Localized Operators on Matrix Weighted $L^{2}$ Spaces,' Preprint available at http://arxiv.org/abs/1407.3819.
[6] S. Bloom, 'A commutator theorem and weighted BMO,' Trans. Amer. Math. Soc. 292 (1985), 103-122
[7] M. Bownik, 'Inverse volume inequalities for matrix weights,' Indiana Univ. Math. J. 50 (2001), 383-410.
[8] M. Christ and M. Goldberg, 'Vector $\mathrm{A}_{2}$ weights and a Hardy-Littlewood maximal function,' Trans. Amer. Math. Soc. 353 (2001), 1995-2002.
[9] R. Coifmann, R. Rochberg and G. Weiss, 'Factorization theorems for Hardy spaces in several variables,' Ann. of Math. 103 (1976), 611-635.
[10] D. Cruz-Uribe SFO, J. Martell and C. Perez, 'Sharp weighted estimates for approximating dyadic operators,' Electron. Res. Announc. Math. Sci. 17 (2010) 12-19.
[11] A. Culiuc and S. Triel, 'The Carleson Embedding Theorem with matrix weights,' Preprint available at http://arxiv.org/abs/1508.01716.
[12] A. Culiuc and B. Wick, 'A proof of the boundedness of the Riesz and AhlforsBeurling transforms in matrix weighted spaces,' Preprint.
[13] M. Goldberg, 'Matrix $A_{p}$ weights via maximal functions,' Pacific J. Math. 211 (2003), 201-220.
[14] I. Holmes, M. Lacey, and B. Wick, 'Bloom's Inequality: Commutators in a Two-Weight Setting,' Preprint available at http://arxiv.org/abs/1505.07947.
[15] I. Holmes, M. Lacey, and B. Wick, 'Commutators in the Two-Weight Setting, Preprint available at http://arxiv.org/abs/1506.05747.
[16] R. Hunt, B. Muckenhoupt, and R. Wheeden, 'Weighted norm inequalities for the conjugate function and Hilbert transform,' Trans. Amer. Math. Soc. 176 (2003), 227-251.
[17] Hytönen, T., 'Representation of singular integrals by dyadic operators, and the $A_{2}$ theorem,' Preprint available at http://arxiv.org/abs/1108.5119v1
[18] J. Isralowitz, 'Matrix weighted Triebel-Lizorkin bounds: a simple stopping time proof,' Preprint available at http://arxiv.org/abs/1507.06700.
[19] J. Isralowitz, 'A Matrix weighted T1 theorem for matrix kernelled CZOs,' Preprint available at http://arxiv.org/abs/1508.02474.
[20] J. Isralowitz, 'Boundedness of commutators and $\mathrm{H}^{1}$-BMO duality in the two matrix weighted setting,' Preprint available at http://arxiv.org/abs/1511.02926.
[21] J. Isralowitz and K. Moen, 'Matrix weighted Poincare inequalities and applications to degenerate elliptic systems,' Preprint available at http://arxiv.org/abs/1601.00111.
[22] S. Janson, 'Mean oscillation and commutators of singular integral operators,' Ark. Mat. 16 (1978), 263-270.
[23] N. H. Katz and M. C. Pereyra, 'Haar multipliers, paraproducts, and weighted inequalities,' Appl. Numer. Harmon. Anal., Analysis of divergence (1997), $145-$ 170.
[24] A. Lerner, A and F, Nazarov, 'Intuitive dyadic calculus: the basics,' preprint avilable at arxiv.org/abs/1508.05639
[25] M. Lacey, S. Petermichl, J. Pipher and B. Wick, 'Iterated Riesz commutators: a simple proof of boundedness, Contemp. Math. 505 (2010), 171-178
[26] B. Muckenhoupt and R. Wheeden, 'Weighted bounded mean oscillation and the Hilbert transform,' Studia Math. 54 (1975/76), 221-237
[27] F. Nazarov and S. Treil, 'The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis,' Algebra i Analiz 8 (1996) 32-162.
[28] F. Nazarov, S. Treil, and A. Volberg, 'The $T b$-theorem on non-homogeneous spaces,' Acta Math. 190 (2003), 151-239.
[29] M. C. Pereyra, 'Lecture notes on dyadic harmonic analysis,' Contemp. Math. 289 (2000), 1-60.
[30] S. Pott, 'A sufficient condition for the boundedness of operator weighted martingale transforms and Hilbert transform,' Studia Mathematica, 182 (2), 2007 pp. 99-111.
[31] S. Roudenko, 'Matrix-weighted Besov spaces,' Trans. Amer. Math. Soc., 355 (2003), 273-314.
[32] M. Taylor, 'Tools for PDE. Pseudodifferential operators, paradifferential operators, and layer potentials,' Mathematical Surveys and Monographs (2000), $\mathrm{x}+257 \mathrm{pp}$.
[33] S. Treil and A. Volberg, 'Wavelets and the angle between past and future,' $J$. Funct. Anal. 143 (1997), 269-308.
[34] A. Volberg, 'Matrix Ap weights via S-functions, J. Amer. Math. Soc. 10 (1997), 445-466.
(Joshua Isralowitz) Department of Mathematics and Statistics, SUNY Albany, 1400 Washington Ave., Albany, NY, 12222

E-mail address, Joshua Isralowitz: jisralowitz@albany.edu
(Hyun Kyoung Kwon) Department of Mathematics, The University of Alabama, P. O. Box 870350, Tuscaloosa, AL, 35487

E-mail address, Hyun Kyoung Kwon: hkwon@ua.edu
(Sandra Pott) Centre for Mathematical Sciences, Lund University, P.O. Box 118, S-221 00 Lund, Sweden

E-mail address, Sandra Pott: sandra@maths.lth.se


[^0]:    2010 Mathematics Subject Classification. 42B20.

