SHARP MATRIX WEIGHTED WEAK AND STRONG TYPE INEQUALITIES FOR THE DYADIC SQUARE FUNCTION

JOSHUA ISRALOWITZ

Abstract. In this paper we refine the recent sparse domination of the integrated $p = 2$ matrix weighted dyadic square function by T. Hytönen, S. Petermichl, and A. Volberg to prove a pointwise sparse domination of general matrix weighted dyadic square functions. We then use this to prove quantitative two matrix weighted estimates for the matrix weighted dyadic square function and along the way prove quantitative two matrix weighted estimates for the matrix weighted maximal function. In particular, we prove sharp two matrix weighted weak and strong type inequalities for matrix weighted dyadic square functions when $1 < p \leq 2$ and prove sharp two matrix weighted weak type inequalities for the matrix weighted maximal function.

1. Introduction

Let $U$ an a.e. positive definite $n \times n$ matrix valued function on $\mathbb{R}^d$ (that is, a matrix weight), and for a measurable $\mathbb{C}^n$ valued function $\vec{f}$ on $\mathbb{R}^d$ define

$$\|\vec{f}\|_{L^p(U)} = \left(\int_{\mathbb{R}^d} |U^{\frac{p}{2}}(x)\vec{f}(x)|^p \, dx\right)^{\frac{1}{p}}.$$ 

We will say that a pair of matrix weights $U, V$ is matrix $A_p$ if

$$[U, V]_{A_p} := \sup_{I \subseteq \mathbb{R}^d} \int_I \left(\int_I \|V^{-\frac{p}{2}}(y)U^{\frac{p}{2}}(x)\|^{p'} \, dy\right)^{\frac{p}{p'}} \, dx < \infty$$

where $\bar{f}_I$ refers to the unweighted average. Clearly this is a condition that reduces to the classical Muckenhoupt two weight $A_p$ condition in the scalar setting (when $n = 1$). If $U = V$ then we will say $U$ is a matrix $A_p$ weight if $[U]_{A_p} := [U, U]_{A_p} < \infty$. While it is known that most “classical” operators from harmonic analysis (such as the maximal function, Calderón-Zygmund operators, paraproducts, martingale transforms, square functions, etc.) are bounded on $L^p(U)$ for matrix $A_p$ weights, it extremely difficult to determine the sharp dependence of such operators on $[U]_{A_p}$.

In fact, the only two such operators where sharp one weighted matrix weighted norm inequalities for $p = 2$ are known are for the dyadic square function, which was recently proved in [10] and the maximal function, which was proved in [11] by slightly modifying the ideas in [4]. Furthermore, among these two operators, sharp one matrix weighted $A_p$ bounds for $1 < p < \infty$ are only known for the maximal function, which were proved in [12] by slightly modifying the ideas in [8]. Moreover, and perhaps more interestingly, there are no known sharp weak type matrix weighted norm inequalities even for $p = 2$.

The purpose of this paper is to prove sharp strong type and weak type matrix weighted norm inequalities for the dyadic square function in the range $1 < p \leq 2$. 
providing the first sharp \( p \neq 2 \) estimates for a singular operator in the matrix weighted setting. We will then use a sort of linearization argument from \[4,8\] to prove sharp weak type matrix weighted norm inequalities for the maximal function in the full range \( 1 < p < \infty \) (see \[18\] for a more traditional linearization argument for matrix weighted maximal functions, which unfortunately does not seem to be of use when considering weak type matrix weighted norm inequalities).

We now precisely state the results of this paper. First, for matrix weights \( U \) and \( V \), define the (matrix weighted) maximal function \( M_{U,V,p} \) (originally defined in \[4\] when \( p = 2 \) and in general in \[8\] in the one weighted setting) by

\[
M_{U,V,p}(\vec{f})(x) = \sup_{I \text{ is a cube}} \int_I |U^{\frac{1}{p}}(x)V^{-\frac{1}{p}}(y)\vec{f}(y)| dy.
\]

Now let \( \mathcal{D} \) be a dyadic grid and let \( \{h^k_J\} \) for \( k = 1, \ldots, 2^d - 1 \) and \( J \in \mathcal{D} \) be any Haar system on \( \mathbb{R}^d \), meaning that \( \{h^k_J\} \) is an orthonormal system of \( L^2(\mathbb{R}^d) \) with \( h^k_J \) supported on \( J \), \( \int_J h^k_J(x) dx = 0 \), and each \( h^k_J \) is constant on dyadic subcubes of \( J \). Also, for a function \( \vec{f} : \mathbb{R}^d \to \mathbb{C}^n \) let

\[
\vec{f}^k_J = \int_J \vec{f}(x)h^k_J(x) dx
\]

and define the matrix weighted dyadic square function \( S_{U,V,p}\vec{f} \) by

\[
S_{U,V,p}\vec{f}(x) = \left( \sum_{J,k} \left| \frac{U^{\frac{1}{p}}(x)(V^{-\frac{1}{p}}\vec{f})^k_J}{|J|} \right|^{\frac{p}{2}} 1_{J(x)} \right)^{\frac{2}{p}}.
\]

For notational ease we will omit the dependence of \( h^k_J \) on \( k \) and presume all sums involving Haar functions are taken over \( k = 1, \ldots, 2^d - 1 \).

Note that these operators are natural substitutes for the ordinary maximal function and dyadic square function in the sense that if \( M \) is the ordinary maximal function on scalar valued functions then

\[
\|Mf\|_{L^p(u) \to L^p(v)} = \|M_{u,v,p}f\|_{L^p \to L^p}
\]

and the same is true for the dyadic square function. To state our first result we need the following definition. We say that a matrix weight \( U \) is matrix \( A_p^{wk} \) if

\[
[U]_{A_p^{wk}} := \sup_{\vec{e} \in \mathbb{C}^n} \left\|U^{\frac{1}{p}}\vec{e} \right\|_{A_\infty} < \infty.
\]

It is easy to show (see \[8\] for example) that a matrix \( A_p \) weight is also a matrix \( A_p^{wk} \) weight with \( [U]_{A_p^{wk}} \leq [U]_{A_p} \) and clearly in the scalar setting we have \( [u]_{A_p^{wk}} = [u]_{A_\infty} \). Our first result is the following.

**Theorem 1.1.** If \( U, V \) is a pair of matrix \( A_p \) weights, \( V^{-\frac{1}{p}} \) is a matrix \( A_p^{wk} \) weight, and \( 1 < p \leq 2 \) then the sharp estimate

\[
\|S_{U,V,p}\|_{L^p \to L^p} \lesssim [U, V]_{A_p^{\frac{1}{p}}}^{\frac{1}{p}}[V^{-\frac{1}{p}}]_{A_p^{wk}}^{\frac{1}{p}}
\]

holds.
Furthermore, if $2 < p < \infty$ we have

$$\|S_{U,V,p}\|_{L^p \rightarrow L^p} \lesssim [U,V]^\frac{1}{p} [V^{-\frac{p}{p'}}]^{\frac{1}{p}} [A_p]_{\tilde{A}_p^w}^{1 - \frac{1}{p}}$$

Note that this was proved when $p = 2$ in \cite{10} in the one weighted case and that sharpness when $1 < p \leq 2$ follows from the well known sharpness in the scalar setting (see \cite{9, 13}). Also, note that while it is unlikely that Theorem 1.1 is sharp when $p > 2$, it is a natural bound and in fact we will recover from the proof of Theorem 1.1 the current best mixed matrix weighted $A_p - A_\infty$ bound for a positive sparse operators $S_{U,V}$ from \cite{5} (and thus the current best bound for CZOs via the sparse convex body domination theorem from \cite{17}), namely

$$\|S_{U,V}\|_{L^p \rightarrow L^p} \lesssim [U,V]^{\frac{1}{p}} [V^{-\frac{p}{p'}}]^{\frac{1}{p}} [A_p]_{\tilde{A}_p^w}^{1 - \frac{1}{p}} [A_p]_{\tilde{A}_p^w}.$$

Concerning weak type estimates we have the following

**Theorem 1.2.** If $U,V$ is a pair of matrix $A_p$ weights and $1 < p \leq 2$ then

$$\|S_{U,V,p}\|_{L^p \rightarrow L^p, \infty} \lesssim [U,V]^{\frac{1}{p}}$$

and furthermore, if $1 < p < \infty$ and $\Phi$ is any function where

$$\|S_{U,p}\|_{L^p \rightarrow L^p, \infty} \lesssim \Phi([U]_{A_p})$$

then $\Phi(t) \gtrsim t^{\max\left(\frac{1}{p} - \frac{1}{p}, \frac{1}{p} - \frac{1}{p}\right)}$.

As mentioned above, we will obtain the following as a corollary of the proof of Theorem 1.2

**Theorem 1.3.** If $U,V$ is a pair of matrix $A_p$ weights and $1 < p < \infty$ then we have the sharp estimate

$$\|M_{U,V,p}\|_{L^p \rightarrow L^p, \infty} \lesssim [U,V]^{\frac{1}{p}}.$$
In the third section we will prove Theorem 1.1 by “matrixizing” some of the ideas in [6] to prove a matrix weighted Carleson embedding type theorem. Of particular novelty here is that we will use a matrix weighted “stopping moment” decomposition, which to the author’s knowledge is the first time such an argument in the matrix weighted setting has appeared. Note that a similar matrix weighted parallel corona decomposition argument should be possible to construct (which in fact was used to prove a sharp version of Theorem 1.1 in the scalar $p > 2$ setting in [13]). In the last section, we will prove Theorems 1.2 and 1.3 by modifying the ideas from [9, 15]. Particularly, as we are estimating the unweighted measure of level sets when proving Theorems 1.2 and 1.3 some of the ideas in [9, 15] are not applicable. Furthermore, as was mentioned before, we will prove Theorem 1.3 as a corollary of the proof of Theorem 1.2 via the local boundedness of a two weight “Christ-Goldberg $N −$ function” (see [8]) and a linearization similar to that in [8]. Note that more traditional linearization arguments are not applicable because of the presence of $U^\dagger(x)$ in the definition of $M_{U, V, p}$. Furthermore, the elementary linearization argument from [18] is not applicable either because of the absence of a weak type bilinear embedding theorem. In the final short section we will prove the sharpness of Theorem 1.2 when $1 < p \leq 2$, the sharpness of Theorem 1.3 prove the quantitative lower bound estimate in Theorem 1.2 and compare this to known scalar situation. In particular, we will briefly discuss the question of whether this lower bound is sharp or not.

We will end this paper with two important points. First, as of the date of writing this paper, it is unknown whether the Rubio de Francia extrapolation theorem holds. Namely, it is not known whether the boundedness of an operator $T$ on $L^2(U)$ for all matrix $A_2$ weights $U$ implies the boundedness of $T$ on $L^p(U)$ for all matrix $A_p$ weights $U$ and all $1 < p < \infty$. Thus, unlike in the scalar setting, sharp estimates (or even just boundedness) of operators for $p \neq 2$ do not at this moment follow from sharp estimates of operators for $p = 2$. Second, it is perhaps more natural to ask for sharp $L^1 \to L^{1, \infty}$ matrix weighted estimates with respect to matrix $A_1$ weights than it is to ask for sharp $L^p \to L^{p, \infty}$ matrix weighted estimates with respect to matrix $A_p$ weights for $p > 1$. With this in mind, sharp matrix weighted $A_1$ estimates will be pursued in joint work with David Cruz Uribe, Kabe Moen, Sandra Pott, and Israel-Pablo Rivera Rios.

2. Sparse domination of square functions

Before we state the main result of this section we will need to introduce some definitions and notation. First we will introduce the concept of a reducing operator, whose importance was emphasized in [8] and which has since shown to be vital in the theory of matrix weighted norm inequalities. Namely, for a matrix weight $U$, a cube $I$, and $e \in \mathbb{C}^n$ there exists positive definite matrices $\mathcal{U}_I, \mathcal{U}_I'$ where

$$
|\mathcal{U}_I e| \simeq \left( \int_I |U^\dagger(x)e|^p \, dx \right)^{\frac{1}{p}}, \quad |\mathcal{U}_I' e| \simeq \left( \int_I |U^{-\dagger}(x)e|^p' \, dx \right)^{\frac{1}{p'}}
$$

where the implicit constant depends only on $n$. In particular, it is easy to see that

$$
[U, V]_{A_p} \simeq \sup_{I \subseteq \mathbb{R}^d} \|\mathcal{U}_I V_I\|^p.
$$

4 J. ISRALOWITZ

$S_{U, V, p}f(x)$. In the third section we will prove Theorem 1.1 by “matrixizing” some of the ideas in [6] to prove a matrix weighted Carleson embedding type theorem. Of particular novelty here is that we will use a matrix weighted “stopping moment” decomposition, which to the author’s knowledge is the first time such an argument in the matrix weighted setting has appeared. Note that a similar matrix weighted parallel corona decomposition argument should be possible to construct (which in fact was used to prove a sharp version of Theorem 1.1 in the scalar $p > 2$ setting in [13]). In the last section, we will prove Theorems 1.2 and 1.3 by modifying the ideas from [9, 15]. Particularly, as we are estimating the unweighted measure of level sets when proving Theorems 1.2 and 1.3 some of the ideas in [9, 15] are not applicable. Furthermore, as was mentioned before, we will prove Theorem 1.3 as a corollary of the proof of Theorem 1.2 via the local boundedness of a two weight “Christ-Goldberg $N −$ function” (see [18]) and a linearization similar to that in [8]. Note that more traditional linearization arguments are not applicable because of the presence of $U^\dagger(x)$ in the definition of $M_{U, V, p}$. Furthermore, the elementary linearization argument from [18] is not applicable either because of the absence of a weak type bilinear embedding theorem. In the final short section we will prove the sharpness of Theorem 1.2 when $1 < p \leq 2$, the sharpness of Theorem 1.3 prove the quantitative lower bound estimate in Theorem 1.2 and compare this to known scalar situation. In particular, we will briefly discuss the question of whether this lower bound is sharp or not.

We will end this paper with two important points. First, as of the date of writing this paper, it is unknown whether the Rubio de Francia extrapolation theorem holds. Namely, it is not known whether the boundedness of an operator $T$ on $L^2(U)$ for all matrix $A_2$ weights $U$ implies the boundedness of $T$ on $L^p(U)$ for all matrix $A_p$ weights $U$ and all $1 < p < \infty$. Thus, unlike in the scalar setting, sharp estimates (or even just boundedness) of operators for $p \neq 2$ do not at this moment follow from sharp estimates of operators for $p = 2$. Second, it is perhaps more natural to ask for sharp $L^1 \to L^{1, \infty}$ matrix weighted estimates with respect to matrix $A_1$ weights than it is to ask for sharp $L^p \to L^{p, \infty}$ matrix weighted estimates with respect to matrix $A_p$ weights for $p > 1$. With this in mind, sharp matrix weighted $A_1$ estimates will be pursued in joint work with David Cruz Uribe, Kabe Moen, Sandra Pott, and Israel-Pablo Rivera Rios.

2. Sparse domination of square functions

Before we state the main result of this section we will need to introduce some definitions and notation. First we will introduce the concept of a reducing operator, whose importance was emphasized in [8] and which has since shown to be vital in the theory of matrix weighted norm inequalities. Namely, for a matrix weight $U$, a cube $I$, and $e \in \mathbb{C}^n$ there exists positive definite matrices $\mathcal{U}_I, \mathcal{U}_I'$ where

$$
|\mathcal{U}_I e| \simeq \left( \int_I |U^\dagger(x)e|^p \, dx \right)^{\frac{1}{p}}, \quad |\mathcal{U}_I' e| \simeq \left( \int_I |U^{-\dagger}(x)e|^p' \, dx \right)^{\frac{1}{p'}}
$$

where the implicit constant depends only on $n$. In particular, it is easy to see that

$$
[U, V]_{A_p} \simeq \sup_{I \subseteq \mathbb{R}^d} \|\mathcal{U}_I V_I\|^p.
$$

4 J. ISRALOWITZ
Now let \( \{\tilde{e}_j\}_{j=1}^n \) be any orthonormal basis of \( \mathbb{C}^n \). We will then use the following simple estimate without further mention throughout the rest of the paper: If \( A \) is any \( n \times n \) matrix then
\[
\|U_1 A\|^p \approx \sum_{j=1}^n |U_1 A \tilde{e}_j|^p \approx \sum_{j=1}^n \int_I |U_1^p (x) A \tilde{e}_j|^p \, dx \approx \int_I \|U_1^p (x) A\|^p \, dx.
\]

Let \( \mathcal{D} \) be a dyadic grid. A collection \( \mathcal{L} \) of dyadic cubes in \( \mathcal{D} \) is sparse if
\[
\bigcup_{L \in \mathcal{L}, J \subseteq L} |L| \leq \frac{1}{2} |J|.
\]
See [17] or [14] for more properties of sparse collections.

Given a sparse collection \( \mathcal{L} \), define the “sparse positive operator” \( \tilde{S} = \tilde{S}_{U,V,p,\mathcal{L}} \) by
\[
\tilde{S}\bar{f}(x) := \left( \sum_{L \in \mathcal{L}} \left\langle \mathcal{L} \mathcal{L} \right| U_L V^{-\frac{1}{2}} \tilde{f} \right\rangle_L^2 \|U_\mathcal{L}^p (x) U_L^{-1}\|_L^2 1_L(x) \right)^{\frac{1}{2}}
\]
where \( \langle \rangle_L \) denotes the unweighted average over \( L \). Furthermore, for any \( J \in \mathcal{L} \) define the localized sparse positive operator \( \tilde{S}_J \) by
\[
\tilde{S}_J \bar{f}(x) := \left( \sum_{L \in \mathcal{L}, L \subseteq J} \left\langle \mathcal{L} \mathcal{L} \right| U_L V^{-\frac{1}{2}} \tilde{f} \right\rangle_L^2 \|U_\mathcal{L}^p (x) U_L^{-1}\|_L^2 1_L(x) \right)^{\frac{1}{2}}
\]
and similarly define \( S_J \bar{f} \) by
\[
S_J \bar{f}(x) = \left( \sum_{L \in \mathcal{L}, L \subseteq J} \frac{|U_\mathcal{L}^p (x) (V^{-\frac{1}{2}} \tilde{f})_L|^2}{|L|} 1_L(x) \right)^{\frac{1}{2}}
\]

The main result of this section is the following.

**Theorem 2.1.** If \( J \in \mathcal{D} \) then there exists a sparse collection \( \mathcal{L} \) of dyadic subcubes of \( J \) where \( S_J \bar{f}(x) \lesssim \tilde{S}_J \bar{f}(x) \) with implicit constant independent of \( x, J, U, \) and \( V \).

**Proof.** Let \( S \) denote the unweighted dyadic square function with respect to \( \mathcal{D} \). Obviously it is enough to show that \( S_J(V^{-\frac{1}{2}} \tilde{f})(x) \lesssim \tilde{S}_J(V^{-\frac{1}{2}} \tilde{f})(x) \). Let \( \mathcal{J}(J) \) be the maximal dyadic subcubes \( L \) of \( J \) where
\[
\sum_{J \supseteq L \supseteq J} \frac{|U_\mathcal{L} \tilde{f}|^2}{|L|} > \lambda \left\langle \mathcal{L} \bar{f} \right\rangle_J^2
\]
We claim that \( \sum_{L \in \mathcal{J}(J)} |L| \leq \frac{1}{2} |J| \) for large enough \( \lambda \). For that matter, if \( x \in L \in \mathcal{J}(J) \) then
\[
S(U_\mathcal{L} \tilde{f})(x) = \sum_{L \in \mathcal{L}} \frac{|U_\mathcal{L} \tilde{f}(x)|^2}{|L|} 1_L(x) \geq \sum_{J \supseteq L \supseteq J} \frac{|(U_\mathcal{L} \tilde{f})_L|^2}{|L|} \geq \lambda \left\langle U_\mathcal{L} \tilde{f} \right\rangle_J^2
\]
so that

\[
\sum_{L \in \mathcal{J}(J)} |L| = \left| \bigcup_{L \in \mathcal{J}(J)} L \right| \leq |\{ x : S(UJ_f)(x) \geq \lambda \frac{\langle UJ_f \rangle}{J} \}| \leq \frac{1}{\lambda^2} |J|
\]

which clearly proves the claim.

Let \( \mathcal{F}(J) \) denote the collection of all \( L \in \mathcal{J}(J) \) such that \( L \not\subseteq Q \) for any \( Q \in \mathcal{J}(J) \). Furthermore, abusing notation slightly, we will denote \( \cup_{Q \in \mathcal{J}(J)} Q \) by \( \cup \mathcal{J}(J) \). Then if \( x \in J \) we have

\[
\sum_{L \in \mathcal{J}(J)} \frac{\left| U^\perp(x)f_L \right|^2 1_L(x)}{|L|} = \sum_{L \in \mathcal{F}(J)} \frac{\left| U^\perp(x)f_L \right|^2 1_{L \cap (\cup \mathcal{J}(J))}(x)}{|L|} + \sum_{L \in \mathcal{F}(J)} \frac{\left| U^\perp(x)f_L \right|^2 1_{L \cup \mathcal{J}(J)}(x)}{|L|}
\]

\[
+ \sum_{Q \in \mathcal{J}(J)} \sum_{L \in \mathcal{J}(Q)} \frac{\left| U^\perp(x)f_L \right|^2 1_L(x)}{|L|}
\]

\[
= A_1(x) + A_2(x) + \sum_{Q \in \mathcal{J}(J)} \sum_{L \in \mathcal{J}(Q)} \frac{\left| U^\perp(x)f_L \right|^2 1_L(x)}{|L|}
\]

We estimate \( A_1(x) \) as follows. First, if \( x \in I \) for some \( I \in \mathcal{J}(J) \) and \( x \in L \in \mathcal{J}(J) \) then again by definition of \( \mathcal{J}(J) \) we have \( J \supseteq L \supseteq I \) so that

\[
A_1(x) \leq \left| U^\perp(x)U^{-1}_J \right|^2 1_J(x) \sum_{J \supseteq I \supseteq I} \frac{\left| U_Jf_L \right|^2}{|L|}
\]

\[
\leq \lambda \left| U^\perp(x)U^{-1}_J \right|^2 1_J(x) \frac{\langle U_Jf \rangle}{J}^2.
\]

We now estimate \( A_2(x) \). If \( x \not\in \cup \mathcal{J}(J) \) then we can pick a sequence of nested dyadic cubes \( \{L_\alpha^x\} = \{L \in \mathcal{J}(J) : x \in L\} = \{L \in \mathcal{J}(J) : x \in L\} \). However, if

\[
\sum_k \frac{\left| U_Jf \right|^2}{|L_k^x|} > \lambda \frac{\langle U_Jf \rangle}{J}^2
\]

then obviously for some \( k' \) we must have

\[
\sum_{J \supseteq L \supseteq L_\alpha^x} \frac{\left| U_Jf \right|^2}{|L|} > \lambda \frac{\langle U_Jf \rangle}{J}^2
\]

which means \( x \in L_\alpha^x \subseteq I \) for some \( I \in \mathcal{J}(J) \). Thus,
maximal 
L
λ > 1 to be determined momentarily. Given □
Theorem 2.1
Finally set
J

Putting this together, we get
∑
L ∈ J
|U(x)f|^{2} \frac{1}{|L|} \sum_{J \in J} \frac{|U(x)f_{L}|^{2}}{|L|}
\leq \lambda \|U(x)f_{J}\|^{2} \frac{1}{|J|} \sum_{J \in J} \frac{|U(x)f_{L}|^{2}}{|L|}.

Finally set \( J_{0}(J) = \{ J \} \) and for \( k \in \mathbb{N} \) set \( J_{k}(J) = \{ L \in J(Q) : Q \in J_{k-1}(J) \} \).
If \( J = \bigcup_{k} J_{k}(J) \) then \( J \) is sparse, which by iteration completes the proof of Theorem 2.1.

3. Proof of Theorem 1.1

In this section we will prove Theorem 1.1 by utilizing Theorem 2.1. Again fix \( \lambda > 1 \) to be determined momentarily. Given \( Q \in \mathcal{D} \) let \( \mathcal{G}(Q) \) denote the set of maximal \( L \in \mathcal{D}(J) \) such that
\[
\int_{Q} |U(x)f| > \lambda \int_{J} |U(x)f|.
\]
We now prove that for \( \lambda > 0 \) large enough we have that
\[
(3.1) \quad \sum_{L \in \mathcal{G}(J)} |L| \leq \frac{1}{4} |J|
\]
To handle the difference in reducing operators, let \( \mathcal{J}(J) \) be the collection of maximal dyadic subcubes \( L \) of \( J \) satisfying
\[
\|U_{L}U^{-1}_{J}\| > \lambda.
\]
Then by the disjointness of the cubes in \( \mathcal{J}(J) \), as usual,
\[
\sum_{L \in \mathcal{J}(J)} |L| \lesssim \frac{1}{(\lambda)^{p}} \sum_{L \in \mathcal{J}(J)} \int_{L} \left\| U(x)U^{-1}_{J}\right\|^{p} \lesssim \frac{1}{(\lambda)^{p}} \int_{J} \left\| U(x)U^{-1}_{J}\right\|^{p} \lesssim \frac{1}{(\lambda)^{p}} |J|.
\]
Letting \( \mathcal{F}(J) \) denote the collection of all \( L \in \mathcal{D}(J) \) that are not contained in any \( L' \in \mathcal{J}(J) \), we get that for \( \lambda' > 0 \) large enough
\[
\sum_{L \in \mathcal{G}(J) \cap \mathcal{F}(J)} |L| \leq \sum_{L \in \mathcal{G}(J) \cap \mathcal{F}(J)} |L| + \sum_{L' \in \mathcal{J}(J) \cap \mathcal{G}(J) \cap \mathcal{F}(J)} |L| \leq \sum_{L \in \mathcal{G}(J) \cap \mathcal{F}(J)} |L| + \frac{1}{4} |J|
\]
where in the last inequality we have used the disjointness of the sets in both \( \mathcal{G}(J) \) and \( \mathcal{J}(J) \).
However,
\[
\sum_{L \in G(J) \cap F(J)} |L| \leq \frac{1}{\lambda} \sum_{L \in G(J) \cap F(J)} \left| \int_{U_L} \frac{|U_L V^{-\frac{r}{p}} \mathbf{f}|}{|U_J V^{-\frac{r}{p}} \mathbf{f}|} \right| \\
\leq \frac{\lambda'}{\lambda} \sum_{L \in G(J) \cap F(J)} \left| \int_{U_L} \frac{|U_L V^{-\frac{r}{p}} \mathbf{f}|}{|U_J V^{-\frac{r}{p}} \mathbf{f}|} \right| \\
\leq \frac{\lambda'}{\lambda} |J|
\]

where in the second equality we have used the fact that \( L \in F(J) \) so that \( |U_L V^{-\frac{r}{p}} \mathbf{f}| \leq \lambda' \). This completes the claim by setting \( \lambda > 8 \lambda' \).

Now for fixed \( N \in \mathbb{N} \) let
\[ G_0 = \{ J \in \mathcal{D} : |J| = 2^N \} \]
and inductively define
\[ G_{k+1} = \{ L \in \mathcal{D} : L \in G(J) \text{ for some } J \in G_k \} \]
If \( E(J) \) denotes the collection of all \( L' \in \mathcal{D}(J) \) that are not contained in any \( L \in G(J) \) and \( G \) is the union
\[ G = \bigcup_{k=0}^{\infty} G_k \]
then we clearly have
\[ \bigcup_{J \in G} E(J) = \{ J \in \mathcal{D} : |J| \leq 2^N \} \]
for \( \lambda > 0 \) large enough since \( J \in E(J) \) for any \( J \in \mathcal{D} \). Also clearly an iteration of (3.1) gives us that for any \( Q \in \mathcal{D} \)
\[
\sum_{L \in G, L \subseteq J} |L| \leq \frac{1}{2} |J|.
\]

We now state and prove a Carleson embedding type theorem for the type of operator used in the previous section, which will easily show Theorem 1.1. Given nonnegative measurable functions \( \{a_L(x)\}_{L \in \mathcal{D}} \) and \( r > 0 \), define \( \tilde{S}_a \mathbf{f} = \tilde{S}_{a,U,V,r,p} \mathbf{f} \) by
\[
\tilde{S}_a \mathbf{f}(x) = \left( \sum_{L \in \mathcal{D}} a_L(x) \left( \left| U_L V^{-\frac{r}{p}} \mathbf{f} \right|_{L^r} \right) \right)^{\frac{1}{r}}.
\]

Theorem 3.1. Let \( 1 < p \leq r \), let \( V^{-\frac{r}{p}} \) be a matrix \( A^w_p \) weight and let
\[
\|A\|_* = \sup_{J \in \mathcal{D}} \int_J \left( \sum_{L \in \mathcal{D}(J)} a_L(x) 1_L(x) \right)^{\frac{1}{r}} dx.
\]
then
\[
\|S_a\|_{L^p \rightarrow L^p} \lesssim [U,V]^\frac{1}{p} [V^{-\frac{r}{p}}]^\frac{1}{r} \|A\|_*^{\frac{1}{r}}.
\]
Proof. Let
\[ F_J(x) = \left( \sum_{L \in E(J)} a_L(x) \left\langle \left| \mathcal{U}_L V^{-\frac{1}{p}} f \right|_L^r \right\rangle \right)^{\frac{1}{r}}. \]
We first get a bound for \( \| F_J \|_{L^p} \). Note that
\[ F_J(x) = \left( \sum_{L \in E(J)} a_L(x) \left\langle \left| \mathcal{U}_L V^{-\frac{1}{p}} f \right|_L^r \right\rangle \right)^{\frac{1}{r}} \]
\[ \leq \left( \sum_{L \in E(J)} a_L(x) \left\langle \mathcal{U}_L V^{-\frac{1}{p}} f \right|_L^r \right)^{\frac{1}{r}} \]
since \( L \in E(J) \). Thus,
\[ \| F_J \|_{L^p} \lesssim \left( \sum_{L \in E(J)} a_L(x) \right) \left\langle \mathcal{U}_L V^{-\frac{1}{p}} f \right|_L^r \right)^{\frac{1}{r}}. \]
However, if
\[ \tilde{S}_{a,N} f(x) = \left( \sum_{|L| \leq 2N} a_L(x) \left\langle \mathcal{U}_L V^{-\frac{1}{p}} f \right|_L^r \right)^{\frac{1}{r}} \]
then (3.2) gives us that
\[ \tilde{S}_{a,N} f(x) = \left( \sum_{J \in G} F_J(x) \right)^{\frac{1}{r}}. \]
Then using the fact that \( p \leq r \),
\[ \| \tilde{S}_{a,N} f \|_{L^p} \lesssim \left( \sum_{J \in G} \| F_J \|_{L^p} \right)^{\frac{1}{r}} \]
\[ \leq \left( \sum_{Q \in G} \| F_Q \|_{L^p} \right)^{\frac{1}{r}} \]
\[ = \left( \sum_{Q \in G} \| F_Q \|_{L^p} \right)^{\frac{1}{r}} \]
\[ \lesssim \| A \|_{A_w^p} \left( \sum_{J \in G} |J| \left\langle \mathcal{U}_J V^{-\frac{1}{p}} f \right|_J^p \right)^{\frac{1}{p}}. \]
By the sharp reverse Hölder inequality for \( A_w^p \) weights, we can pick \( \epsilon \approx |V|^{-\frac{1}{p}} [A_w^p]_{A_w^p}^{-1} \) small enough where
Moreover, by the (unweighted) dyadic Carleson embedding theorem for the exponent \( \frac{p}{p-\epsilon} > 1 \) and (3.4) we get

\[
\tilde{S}_{a,N} \lesssim \|A\|_{\mathcal{R}}^p \left( \sum_{J \in G} |J| \left\langle |\tilde{f}|^{p-\epsilon} \right\rangle_{J} \right)^{\frac{1}{p}}
\]

\[
\lesssim \|A\|_{\mathcal{R}}^p \left( \int_{\mathbb{R}^d} (|\tilde{f}(x)|^{p-\epsilon})^\frac{1}{p-\epsilon} \, dx \right)^{\frac{1}{p}}
\]

\[
= \|A\|_{\mathcal{R}}^p [U, V]_{A_p}^\frac{1}{p} \|\tilde{f}\|_{L^p}.
\]

Letting \( N \to \infty \) completes the proof. \( \square \)

Finally, to see how this proves 1.1 when \( 1 < p \leq 2 \), set \( r = 2 \) and let \( \mathcal{L} \) be a sparse collection. Then set

\[
a_L(x) = \begin{cases} \|U^\frac{1}{p'}(x)\mathcal{L}^{-1}\|_L^2 & L \in \mathcal{L} \\ 0 & L \notin \mathcal{L} \end{cases}
\]

Then since \( \frac{p}{2} \leq 1 \)

\[
\sup_{J \in \mathcal{G}} \int J \left( \sum_{L \in \mathcal{D}(J)} \|U^\frac{1}{p'}(x)\mathcal{L}^{-1}\|_L^2 1_L(x) \right)^{\frac{2}{p}} \leq \sup_{J \in \mathcal{G}} \frac{1}{|J|} \sum_{L \in \mathcal{D}(J) \cap \mathcal{L}} \int J \|U^\frac{1}{p'}(x)\mathcal{L}^{-1}\|_L^p \, dx
\]

\[
\lesssim \sup_{J \in \mathcal{G}} \frac{1}{|J|} \sum_{L \in \mathcal{D}(J) \cap \mathcal{L}} |L| \leq \infty.
\]

Theorem 1.1 when \( 1 < p \leq 2 \) then easily follows from Theorem 2.1. To prove Theorem 1.1 when \( p > 2 \), we argue as in [3] and use a routine duality argument. In fact, we will prove a slightly stronger result. Given a sparse collection \( \mathcal{Z} \) and \( r > 0 \), define the \( \tilde{S}_r \) by

\[
\tilde{S}_r \tilde{f}(x) := \left( \sum_{L \in \mathcal{Z}} \left\langle |\mathcal{L} V^{-\frac{1}{r}} \tilde{f}| \right\rangle_{L} \|U^\frac{1}{p'}(x)\mathcal{L}^{-1}\|_L^r 1_L(x) \right)^{\frac{1}{r}}.
\]
Assume $p > r$ so that $\frac{p}{r} > 1$. Then
\[
\|\tilde{S}_r f\|_{L^p \to L^p} = \left\| \sum_{L \in \mathcal{L}} \left[ \mathcal{U}_L V^{-\frac{1}{r}} \right]^r \| \mathcal{U}_L^{-1} \|^r 1_L \right\|_{L^p}^{\frac{1}{r}} \\
\lesssim \sup_{\|g\|_{L^p} \leq 1} \left( \sum_{L \in \mathcal{L}} |L| \left[ \mathcal{U}_L V^{-\frac{1}{r}} \right]^r \| \mathcal{U}_L^{-1} \|^r g \right)_L^{\frac{1}{r}}.
\]

However, as in the $1 < p \leq 2$ case, by the sharp reverse Hölder inequality for $A_\infty$ weights we can pick $\epsilon_1 \approx [V^{\frac{1}{r}}]_{A_p}^{-1}$ and $\epsilon_2 \approx [U]_{A_p}$ where
\[
\left[ U, V \right]_{A_p}^{\frac{1}{r}} \left( \sum_{L \in \mathcal{L}} |L| \left[ \mathcal{U}_L V^{-\frac{1}{r}} \right]^r \| \mathcal{U}_L^{-1} \|^r g \right)_L^{\frac{1}{r}} \approx [U, V]_{A_p} \left[ f^{p-1} \right]_{L^p} \left[ g^{\frac{p-2}{p}} \right]_{L^{p^*}}^{\frac{1}{p}}
\]
and
\[
\left[ U, V \right]_{A_p}^{\frac{1}{r}} \left( |f|^{p-1} \right)^{\frac{p-2}{p}}_{L} \left( |g|^{\frac{p-2}{p}} \right)_{L}^{\frac{1}{p}} \lesssim [U, V]_{A_p} \left[ f^{p-1} \right]_{L^p} \left[ g^{\frac{p-2}{p}} \right]_{L^{p^*}}^{\frac{1}{p}}
\]
If as usual
\[
E_L = L \setminus \bigcup_{L' \in \mathcal{L}} L'
\]
then the sets $\{E_L : L \in \mathcal{L}\}$ are disjoint and $|L| \leq 2|E_L|$. We then have
\[
[U, V]_{A_p}^{\frac{1}{r}} \left( \sum_{L \in \mathcal{L}} |L| \left[ f^{p-1} \right]_{L}^{\frac{p-2}{p}} \left( |g|^{\frac{p-2}{p}} \right)_{L}^{\frac{1}{p}} \right)^{\frac{1}{r}} \lesssim [U, V]_{A_p} \left[ f^{p-1} \right]_{L^p} \left[ g^{\frac{p-2}{p}} \right]_{L^{p^*}}^{\frac{1}{p}}
\]
where $M_d$ is the dyadic maximal function. This completes the proof.

Notice that when $r = 1$ and $p > 1$ we get that
\[
\|\tilde{S}_1\|_{L^p \to L^p} \lesssim [U, V]_{A_p}^{\frac{1}{r}} [V^{\frac{1}{r}}]_{A_p}^{-1} [U]_{A_p}^{\frac{1}{p}}
\]
which as mentioned before coincides with the best known $A_p - A_\infty$ bound for sparse operators, since a sparse operator $S_{U, V}$ defined by
\[
S_{U, V} \hat{f}(x) = \sum_{L \in \mathcal{L}} \left[ U \right]_{L}^\frac{1}{r} (x) \left( V^{\frac{1}{r}} \hat{f} \right)_L 1_L(x)
\]
can be trivially dominated by $S_1$.

4. PROOF OF THEOREMS 1.2 AND 1.3

For a sparse collection $\mathcal{L}$ let
\[
\tilde{S}_r \hat{f}(x) := \left( \sum_{L \in \mathcal{L}} \left[ \mathcal{U}_L V^{-\frac{1}{r}} \hat{f} \right]^r L \right)^{\frac{1}{r}} \left[ U \right]_{L}^\frac{1}{r} (x) \left( V^{\frac{1}{r}} \hat{f} \right)_L 1_L(x)
\]
We now prove the following, which in conjunction with Theorem 2.1 proves Theorem 1.2.
Theorem 4.1. If $\mathcal{L}$ is sparse and $1 < p \leq r$ then
\[ \| \tilde{S} \|_{L^p \rightarrow L^p, \infty} \lesssim [U, V]_{A_p}^{\frac{1}{2}}. \]

Proof. By rescaling, it is enough to prove that
\[ \left| \left\{ x \in \mathbb{R}^d : \tilde{S} \tilde{f}(x) > 1 \right\} \right| \lesssim [U, V]_{A_p} \| f \|_{L^p} \]

Let $a = 2^{(d+1)}$ and further sparsify $\mathcal{L}$ so that
\[ \left| \bigcup_{L \subseteq Q, L \not\subseteq \mathcal{L}} L \right| \leq \frac{1}{2a^2} |Q| \]
(see p. 26 in [14])

First, by Hölder’s inequality we have
\[ \langle \mathcal{U}_L V^{-\frac{1}{r}} \tilde{f} \rangle_L^r \leq \langle \mathcal{U}_L V^{-\frac{1}{r}} \| f \|_L^r \rangle_L^{\frac{1}{r}} \langle f \|_L^p \rangle_L \]
\[ \lesssim [U, V]_{A_p} \langle f \|_L^p \rangle_L \]
and thus it is enough to prove that
\begin{equation}
\left\{ x \in \mathbb{R}^d : \sum_{L \in \mathcal{L}} \left( \langle f \|_L^p \rangle_L \mathcal{U}_L^{-1} \right)^r 1_L(x) > 1 \right\} \lesssim \| f \|_{L^p} \tag{4.1}
\end{equation}

For that matter, let $\hat{S}^k$ for $k \in \mathbb{Z}$ be the collection
\[ \{ L \in \mathcal{L} : a^k \leq \langle f \|_L^p \rangle_L < a^{k+1} \}. \]

Furthermore, let $S^k$ be the collection of maximal $L \in \mathcal{L}$ satisfying $\langle f \|_L^p \rangle_L > a^k$ for $k \in \mathbb{Z}$ and let $\mathcal{S} = \bigcup_{k \in \mathbb{N}} S^k$. Then
\[ \left\{ x \in \mathbb{R}^d : \sum_{L \in \mathcal{L}} \left( \langle f \|_L^p \rangle_L \mathcal{U}_L^{-1} \right)^r 1_L(x) > 1 \right\} \]
\[ \left\{ x \in \mathbb{R}^d : \sum_{k \in \mathbb{Z}} \sum_{L \in \mathcal{S}^k} \langle f \|_L^p \rangle_L \mathcal{U}_L^{-1} \right\} := I. \]

Clearly since $\frac{p}{r} \leq 1$ we have
Using sparsity, however, gives us that
\[
\sum_{k \in \mathbb{Z}} \sum_{L \in \mathcal{S}^k} a^k |L| = \sum_{k \in \mathbb{Z}} a^k \sum_{Q \in S_k} \sum_{L \in Q} |L| \\
\leq \sum_{k \in \mathbb{Z}} \sum_{Q \in S_k} a^k |Q| \\
\leq \sum_{k \in \mathbb{Z}} \sum_{Q \in S_k} |Q| \left\langle |\hat{f}|^p \right\rangle_Q \\
= \sum_{Q \in \mathcal{S}} \int_Q |\hat{f}(x)|^p \, dx
\]
However, if
\[
E_Q = Q \setminus \bigcup_{Q' \in S \atop Q' \subseteq Q} Q' = Q \setminus \bigcup_{Q' \in S^{k+1} \atop Q' \subseteq Q} Q' 
\]
for $Q \in S^k$ then the sets $E_Q$ for $Q \in S$ are pairwise disjoint. Moreover, for $Q \in S^k$, if $\tilde{Q}$ is the double of $Q$ then maximality implies
\[
\left\langle |\hat{f}|^p \right\rangle_{\tilde{Q}} \leq 2^d \left\langle |\hat{f}|^p \right\rangle_Q \leq 2^d a^k \leq a^{k+1}
\]
so that
\[
\left\langle |\hat{f}|^p 1_{E_Q} \right\rangle_Q = \frac{1}{|Q|} \int_Q |\hat{f}(x)|^p \, dx - \frac{1}{|Q|} \sum_{Q' \in S^{k+1} \setminus Q} |Q'| \int_{Q'} |\hat{f}(x)|^p \, dx \\
= \frac{1}{|Q|} \int_Q |\hat{f}(x)|^p \, dx - \sum_{Q' \in S^{k+1} \setminus Q} |Q'| \frac{1}{|Q'|} \int_{Q'} |\hat{f}(x)|^p \, dx \\
\geq \frac{1}{|Q|} \int_Q |\hat{f}(x)|^p \, dx - \frac{1}{2a^2} (a^{k+2}) \\
\geq \frac{1}{2} \left\langle |\hat{f}|^p \right\rangle_Q
\]
Finally, this means
\[
\sum_{Q \in S} \int_Q |f(x)|^p \, dx \leq 2 \sum_{Q \in S} \int_{E_Q} |f(x)|^p \, dx \leq \|f\|_{L^p}^p
\]

by the disjointness of the sets \( \{E_Q\}_{Q \in S} \).

Before we prove Theorem 1.3 we will need a crucial lemma. Note that a very similar one weighted result was proved in [4][8] (or more accurately outlined in [4][8] and more completely proved in [2][12], where the latter corrected a slight error in the proof from [2]).

**Lemma 4.2.** If \( Q \) is a cube and \( N_Q(x) := \sup_{x \in R \supseteq Q} \left| U^\pm(x)V_R' \right| \) where the supremum is taken over all dyadic \( R \) such that \( x \in R \supseteq Q \) then

\[
\int_Q (N_Q(x))^p \, dx \lesssim [U, V]_{A_p} |Q|.
\]

**Proof.** Let \( Q \) be any dyadic cube and for \( m \in \mathbb{N} \) let

\[
N_{Q,m}(x) = \sup_{x \in R \supseteq Q \atop \ell(R) > 2^{-m}} \left| U^\pm(x)V_R' \right|
\]

where \( \ell(R) \) denotes the side length of the cube \( R \).

Let \( \{R_j\} \) be maximal dyadic subcubes of \( Q \) satisfying

\[
\|V_{R_j}(V_Q')^{-1}\| > C
\]

and \( \ell(R_j) > 2^{-m} \) for some large \( C > 1 \) independent of \( U \) and \( V \) to be determined. Then as before

\[
C' \sum_j |R_j| \leq \sum_j |R_j| \|V_{R_j}(V_Q')^{-1}\|_{V'} \lesssim \sum_j \int_{R_j} \|V^{-\pm}(x)(V_Q')^{-1}\|_{V'} \, dx \lesssim |Q|
\]

Thus for \( C \) large enough independent of \( U \) and \( V \) we have \( \sum_j |R_j| \leq \frac{1}{2}|Q| \) and each \( R_j \) (if any even exist) satisfies \( R_j \subseteq Q \).

On the other hand if \( x \in Q \setminus \bigcup_j R_j \) then for any dyadic cube \( R \subseteq Q \) containing \( x \) with \( \ell(R) > 2^{-m} \) we have

\[
\|U^\pm(x)V_R'\| \leq \|U^\pm(x)V_Q'\| \|(V_Q')^{-1}V_R'\| \leq C \|U^\pm(x)V_Q'\|
\]

so that

\[
\int_{Q \setminus \bigcup_j R_j} (N_{Q,m}(x))^p \, dx \leq C' \int_Q \|U^\pm(x)V_Q'\|^p \, dx \\
\leq C' [U, V]_{A_p} |Q|.
\]

If \( x \in R_j \) and \( N_{Q,m}(x) \neq N_{R_j,m}(x) \) then by maximality and the arguments above we have \( N_{Q,m}(x) \leq C \|U^\pm(x)V_Q'\| \). So if \( R = R_j \) and

\[
F_R = \{x \in R : N_{Q,m}(x) \neq N_{R,m}(x)\}
\]
then arguing as above gives us that
\[ \int_{F_R} (N_{Q,m}(x))^p \, dx \leq C^p |U,V| A_p |Q|. \]

Setting \( D_1 = \{ R_j \} \), \( \tilde{C} = 2C^p |U,V| A_p \), and combining what is above gives us that
\[ \int_Q (N_{Q,m}(x))^p \, dx = \left( \int_Q \bigcup_{R \in D_1} F_R \right) + \sum_{R \in D_1} \int_{F_R} (N_{Q,m}(x))^p \, dx \]
\leq \tilde{C} |Q| + \sum_{R \in D_1} \int_{F_R} (N_{R,m}(x))^p \, dx \]
(4.2)
where \( D_1 \) is a (possibly empty) disjoint collection of dyadic subcubes strictly contained in \( Q \) and satisfying
\[ \sum_{R \in D_1} |R| \leq 2^{-1} |Q|. \]

We now proceed inductively. Clearly the Lemma is proved if \( D_1 = \emptyset \). Otherwise, for each \( \tilde{R} \in D_1 \), let \( R = R_{\tilde{R},j} \) be maximal dyadic subcubes of \( \tilde{R} \) satisfying
\[ \| (V'_{\tilde{R}})^{-1} V'_{\tilde{R}} \| > C \]
and \( \ell(R) > 2^{-m} \). Furthermore, let \( D_2 = \{ R_{\tilde{R},j} : \tilde{R} \in D_1 \} \). Then by (4.2) we have
\[ \int_Q (N_{Q,m}(x))^p \, dx \leq \tilde{C} |Q| + \sum_{R \in D_1} \int_{F_R} (N_{R,m}(x))^p \, dx \]
\leq \tilde{C} |Q| + \tilde{C} \sum_{R \in D_1} |R| + \sum_{R \in D_2} \int_{F_R} (N_{R,m}(x))^p \, dx \]
\leq \tilde{C} |Q| + \frac{\tilde{C}}{2} |Q| + \sum_{R \in D_2} \int_{F_R} (N_{R,m}(x))^p \, dx \]
(4.3)
where
\[ \sum_{R \in D_2} |R| \leq 2^{-2} |Q| \]
and each dyadic cube of \( D_2 \) is a strict subset of a dyadic cube in \( D_1 \) (where again the lemma is proved if \( D_2 = \emptyset \)). Continuing like this, we obtain classes \( \{ D_k \} \) where each dyadic cube in \( D_k \) is a strict subset of a dyadic cube in \( D_{k-1} \) and
\[ \int_Q (N_{Q,m}(x))^p \, dx \leq (2 - 2^{-k}) \tilde{C} |Q| + \sum_{R \in D_k} \int_{F_R} (N_{R,m}(x))^p \, dx. \]
Let \( M = \log_2(\ell(Q)) \) then by definition \( D_k \) for \( k \geq m + M \) is empty which gives us that
\[ \int_Q (N_{Q,m}(x))^p \, dx \leq (2 - 2^{-m-M}) \tilde{C} |Q|. \]
The monotone convergence theorem now completes the proof.

We are now in a position to prove Theorem 1.3. Unlike in the scalar setting, however, the presence of \( U^+(x) \) in the definition of \( M_{U,V,p} f(x) \) makes the domination of \( M_{U,V,p} f(x) \) by a sparse type operator more difficult. To do this we will use the arguments from [4,8], which should be thought of as a linearization replacement.
Furthermore, as is customary, we will prove Theorem 1.3 for the dyadic maximal function with respect to an arbitrary dyadic grid \( D \), which by considering a finite number of translates of the standard dyadic grid, will prove Theorem 1.3 (see [14]).

First, as before, let \( a = 2^{d+1} \) and let \( \hat{S}^k \) for \( k \in \mathbb{Z} \) be the collection

\[
\{ L \in D : a^k \leq \left\langle |\vec{f}|^p \right\rangle_L < a^{k+1} \}.
\]

Furthermore, let \( S^k \) be the collection of maximal \( L \in D \) satisfying \( \left\langle |\vec{f}|^p \right\rangle_L > a^k \) for \( k \in \mathbb{Z} \) and let \( S = \bigcup_{k \in \mathbb{N}} S^k \).

It is then easy to see (and very standard) that \( S \) is sparse. Thus, again, we can further sparsify and assume that

\[
\left| \bigcup_{L,Q} L \subseteq Q \right| \leq \frac{1}{2a^2 |Q|}.
\]

Now, for \( x \in \mathbb{R}^d \), pick a dyadic cube \( R_x \ni x \) where

\[
M_{U,V,p} \bar{f}(x) \leq 2 \int_{R_x} \left| U^{\frac{1}{2}}(x)V^{-\frac{1}{2}}(y) \right| dy
\]
\[
\leq 2 \left| U^{\frac{1}{2}}(x)V_{R_x} \right| \int_{R_x} \left| (V'_{R_x})^{-1}V^{\frac{1}{2}}(y) \vec{f}(y) \right| dy
\]
\[
\leq 2 \left| U^{\frac{1}{2}}(x)V_{R_x} \right| \left\langle \left| V^{\frac{1}{2}}(V'_{R_x})^{-1} \right|^p \right\rangle_{R_x} \left\langle \left| \vec{f} \right|^p \right\rangle_{R_x},
\]
\[
\lesssim 2 \left| U^{\frac{1}{2}}(x)V_{R_x} \right| \left\langle \left| \vec{f} \right|^p \right\rangle_{R_x}^{\frac{1}{p}}.
\]

Pick some \( k \in \mathbb{Z} \) such that \( R_x \in \hat{S}^k \) and pick maximal \( Q \in S^k \) with \( R_x \subseteq Q \). Then we can put everything together and estimate

\[
M_{U,V,p} \bar{f}(x) \lesssim \left( \sum_{k \in \mathbb{Z}} \sum_{Q \in S^k} a^k (N_Q(x))^p 1_Q(x) \right)^{\frac{1}{p}}.
\]

Thus, Lemma 3.2 gives us that

\[
\left| \{ x \in \mathbb{R}^d : |M_{U,V,p} \bar{f}(x)| > 1 \} \right| \leq \left\{ x \in \mathbb{R}^d : \sum_{k \in \mathbb{Z}} \sum_{Q \in S^k} a^k (N_Q(x))^p 1_Q(x) > 1 \right\}
\]
\[
\leq \sum_{k \in \mathbb{Z}} \sum_{Q \in S^k} a^k \int_Q (N_Q(x))^p \, dx
\]
\[
\lesssim [U,V]_{A_p} \sum_{Q \in S} \int_Q |\vec{f}(x)|^p \, dx
\]

Arguing as in the end of the proof of Theorem 4.1 then completes the proof. \( \square \)
5. Sharpness of Theorems 1.2 and 1.3

In this short section we prove the sharpness of Theorems 1.2 and 1.3 by slightly modifying examples from [3] and [15, 24]. As before

\[ Sf(x) = \left( \sum_{J \in \mathcal{D}} \frac{1_J(x)}{|J|} |f_J|^2 \right)^{\frac{1}{2}} \]

denote the classical dyadic square function defined on scalar valued functions, where here \( \mathcal{D} \) is the standard dyadic lattice and as before

\[ f_J = \int_J f \, dt. \]

Let \( w(x) = |x|^{(1-\delta)(p-1)} \) for \( 0 < \delta < 1 \) and let \( f(x) = x^{\delta-1} 1_{[0,1)}(x) \) \( \text{sgn} \ h_{[0,1)} \). Then \( \|f\|_{L^p(w)} \approx \delta^{-\frac{1}{p}} \) while \( \|w\|_{A_p} \approx \delta^{\frac{1}{p}}. \) Further,

\[ Sf(x) \geq \frac{1_{[0,1)}(x)}{|J|} |f_J| = 1_{[0,1)}(x)\delta^{-1}. \]

Thus,

\[ \|w^\frac{1}{p} Sf\|_{L^p(w)} \geq \sup_{t>0} t |\{ x \in [0,1) : \delta^{-1} x^\frac{(\delta-1)}{p} > t \}| \geq \frac{1}{2} \delta^{-1} |\{ x \in [0,1) : \delta^{-1} x^\frac{(\delta-1)}{p} > \frac{1}{2} \delta^{-1} \}| \approx \delta^{-1}. \]

Finally, this means that if \( \|w^\frac{1}{p} S(w^{-\frac{1}{p}}) f\|_{L^p(w)} \leq N \|f\|_{L^p}, \) or equivalently \( \|w^\frac{1}{p} Sf\|_{L^p(w)} \leq N \|f\|_{L^p}, \) then using \( w \) and \( f \) as above gives us that

\[ N \geq \frac{\|w^\frac{1}{p} Sf\|_{L^p(w)}}{\|f\|_{L^p}} \approx \delta^{-1+\frac{1}{p}} \approx \|w\|_{A_p}^{\frac{1}{p}}. \]

Clearly the same example but with \( f(x) = x^{\delta-1} 1_{[0,1)}(x) \) gives that

\[ \sup_{f \neq 0} \frac{\|S_{w, r, p} f\|_{L^p}}{\|f\|_{L^p}} \geq \|w\|_{A_p}^{\frac{1}{p}} \]

for all \( r > 0 \) and \( p > 1. \) Note that almost the same example shows the sharpness of Theorem 1.3.

To finish the lower bound claim in Theorem 1.1 assume \( \Phi \) is a power function and \( \|w^\frac{1}{p} S(w^{-\frac{1}{p}})\|_{L^p(w) \rightarrow L^p(w)} \lesssim \Phi(\|w\|_{A_p}). \) We show that \( \Phi(t) \gtrsim t^{\frac{1}{2} - \frac{1}{p}} \) by modifying the example in [3][15]. Thinking of \( w^\frac{1}{p} S(w^{-\frac{1}{p}}) \) as a map from \( L^p(\ell^2) \) to \( L^{p, \infty}(\ell^2) \) the dual takes \( L^{p', 1}(\ell^2) \) to \( L^p(\ell^2) \) so that

\[ \left\| w^{-\frac{1}{p}} \sum_Q |Q|^\frac{1}{2} \langle w^\frac{1}{p} a_Q \rangle_Q h_Q \right\|_{L^{p', 1}} \lesssim \Phi(\|w\|_{A_p}) \left\| \sum_Q a_Q^2 \right\|_{L^{p', 1}} \]

which combined with Khintchine’s inequality means that

\[ \left\| w^{-\frac{1}{p}} \left( \sum_Q \langle w^\frac{1}{p} a_Q \rangle_Q^2 1_Q \right) \right\|_{L^{p', 1}} \lesssim \Phi(\|w\|_{A_p}) \left\| \sum_Q a_Q^2 \right\|_{L^{p', 1}} \]

where here \( \{a_Q\} \) is a sequence of positive measurable functions. For \( \delta > 0 \) set \( w(x) = |x|^{\delta-1} \) so that \( \|w\|_{A_p} \approx \delta^{-1}. \)
For \( k \in \mathbb{N} \) define
\[
a_{[0,2^{-k})}(x) := a_k(x) = x^{\frac{k-1}{p}} \sum_{j=k+1}^{\infty} 2^{-\delta(j-k)} 1_{[2^{-j},2^{-j+1})}(x)
\]
so that \( \sum_k (a_k(x))^2 \lesssim x^{\frac{k-1}{p}} 1_{[0,1)}(x) \) and so easy estimates give us that
\[
\left\| \left( \sum_{k=1}^{\infty} (a_k(x))^2 \right)^{\frac{1}{2}} \right\|_{L^{p',1}} \lesssim \delta^{-1}.
\]
Furthermore,
\[
\left( a_k \cdot w^{\frac{1}{p}} \right)_{[0,2^{-k})} \approx \delta^1 \sum_{j=k+1}^{\infty} 2^{-j(\frac{\delta-1}{p})} 2^{-\delta(j-k)} \approx \delta^{-\frac{1}{2}} 2^{k(1-\delta)}
\]
so that if \( 2^{-\ell} \leq x < 2^{-\ell+1} \) then
\[
\sum_{k=1}^{\infty} \left( a_k \cdot w^{\frac{1}{p}} \right)_{[0,2^{-k})} 1_{[0,2^{-k})}(x) \approx \delta^{-\ell} \sum_{k=1}^{\infty} 2^{2k(1-\delta)} \approx \delta^{-\ell} 2^{2\ell(1-\delta)} \approx \delta^{-1} x^{2(1-\delta)}.
\]
It follows that
\[
\int_0^1 \left( \sum_{k=1}^{\infty} \left( a_k \cdot w^{\frac{1}{p}} \right)_{[0,2^{-k})} 1_{[0,2^{-k})} \right) w^{-\frac{1}{p'}} dx \approx \delta^{-\frac{\ell}{p'}} \int_0^1 x^{p'(\delta-1)} x^{-\frac{p'(\delta-1)}{p'}} dx = \delta^{-\frac{\ell}{p'}}
\]
which finally means that \( \delta^{-\frac{1}{2}} 2^{k(1-\delta)} \lesssim \delta^{-1} \Phi(||w||_{A_p}) \leq \delta^{-1} \Phi(c\delta^{-1}) \), and hence \( \Phi(t) \gtrsim t^{(\frac{1}{2}-\frac{1}{p'})} \).

Finally, it is instructive to quickly compare this example to the one in [15]. Namely, let \( w \) be as above and let \( \sigma(t) = w^{-\frac{1}{p'}}(t) \). Suppose now that \( ||S||_{L^p(w) \rightarrow L^p(w)} \leq \Phi(||w||_{A_p}) \). Then arguing as above, we get
\[
\left\| \left( \sum_Q (w \cdot a_Q)^2 1_Q \right)^{\frac{1}{2}} \right\|_{L^{p',1}(\sigma)} \lesssim \Phi(||w||_{A_p}) \left( \sum_Q a_Q^2 \right)^{\frac{1}{2}}.
\]
Let
\[
a_{[0,2^{-k})}(x) := a_k(x) = \sum_{j=k+1}^{\infty} 2^{-\delta(j-k)} 1_{[2^{-j},2^{-j+1})}(x)
\]
so again \( \sum_k (a_k(x))^2 \lesssim 1_{[0,1)}(x) \) which means that
\[
\left\| \left( \sum_{k=1}^{\infty} (a_k(x))^2 \right)^{\frac{1}{2}} \right\|_{L^{p',1}(w)} \lesssim w([0,1])^{\frac{1}{p'}}.
\]
Furthermore,
\[
\int_0^1 \left( \sum_{k=1}^{\infty} \left( a_k \cdot w \right)_{[0,2^{-k})} 1_{[0,2^{-k})} \right) w^{\frac{1}{p'}} \sigma \approx \delta^{-\frac{1}{p'}} 2^k \lesssim \delta^{-\frac{1}{p'}} w([0,1]).
\]
which finally means that \( \delta^{-\frac{1}{2}} \lesssim \Phi(||w||_{A_p}) \leq \Phi(c\delta^{-1}) \). Clearly contrasting these two examples shows the dramatic effect of tucking in the weight and working with
unweighted $L^{p,\infty}$ vs working with weighted $L^{p,\infty}$, and it is unclear whether this means that our example exhibits sharp behavior or whether our example is simply inefficient.

References

[1] J. T. Stuart, ‘Mathematics applied in fluid motion’, Quart. Appl. Math. 56 (1998) 787–796.
[2] A. Aaen, ‘Singular Integral Operators on Matrix-weighted $L^p$ Spaces’, Thesis available at https://projekter.aau.dk/projekter/files/17607609/speciale.pdf.
[3] Buckley, S., ‘Estimates for operator norms on weighted spaces and reverse Jensen inequalities.’ Trans. Amer. Math. Soc. 340 (1993), no. 1, 253–272.
[4] Christ, M., Goldberg, M., ‘Vector $A_2$ weights and a Hardy-Littlewood maximal function.’ Trans. Amer. Math. Soc. 353 (2001), no. 5, 1995–2002.
[5] D. Cruz-Uribe, J., Isralowitz, J., and Moen, K. ‘Two weight bump conditions for matrix weights.’ Integral Equations Operator Theory 90 (2018), no. 3, Art. 36, 31 pp.
[6] Culic, A., ‘A note on two weight bounds for the generalized Hardy-Littlewood Maximal operator,’ preprint available at https://arxiv.org/abs/1506.07125.
[7] D. Duoandikoetxea, J. ‘Fourier analysis.’ Translated and revised from the 1995 Spanish original by David Cruz-Uribe. Graduate Studies in Mathematics, 29. American Mathematical Society, Providence, RI, 2001. xviii+222 pp.
[8] Goldberg, M. ‘Matrix $A_p$ weights via maximal functions.’ Pacific J. Math. 211 (2003), no. 2, 201–220.
[9] Hytönen, T and Li, K. ‘Weak and strong $A_p$-$A_{\infty}$ estimates for square functions and related operators.’ J. Funct. Anal. 146 (2018), no. 6, 2497–2507.
[10] Hytönen, T., Petermichl, S., and Volberg, A. ‘The sharp square function estimate with matrix weight,’ preprint available at https://arxiv.org/abs/1702.04569.
[11] Isralowitz, J., Kwon, H. K., Pott, S. ‘Matrix weighted norm inequalities for commutators and paraproducts with matrix symbols.’ J. Lond. Math. Soc. (2) 96 (2017), no. 1, 243–270.
[12] Isralowitz, J., and Moen, K. ‘Matrix weighted Poincaré inequalities and applications to degenerate elliptic systems.’ To appear in the Indiana Journal of Mathematics, https://arxiv.org/abs/1601.00111.
[13] Lacey, M. and Li, K. ‘On $A_p$-$A_{\infty}$ type estimates for square functions.’ Math. Z. 284 (2016), no. 3 - 4, 1211–1222.
[14] Lerner, A., and Nazarov, F. ‘Intuitive dyadic calculus: the basics’, preprint available at https://arxiv.org/abs/1508.05639.
[15] Lacey, M. and Scurry, J. ‘Weighted Weak Type Estimates for Square Functions,’ preprint available at https://arxiv.org/abs/1211.4219.
[16] Muckenhoupt, B. and Wheeden, R. ‘Some weighted weak-type inequalities for the Hardy-Littlewood maximal function and the Hilbert transform.’ Indiana Univ. Math. J. 26 (1977), no. 5, 801–816.
[17] Nazarov, F., Petermichl, S., Treil, S., and Volberg, A. ‘Convex body domination and weighted estimates with matrix weights.’ Adv. Math. 318 (2017), 279–306.
[18] Petermichl, S., Pott, S., Reguera, M. ‘A matrix weighted bilinear Carleson Lemma and Maximal Function,’ https://arxiv.org/abs/1811.05838.

Department of Mathematics, University at Albany, 1400 Washington Ave., Albany, NY 12159, USA
E-mail address: jisralowitz@albany.edu