SHARP BOUNDS AND $T_1$ THEOREM FOR CALDERÓN-ZYGMDUN OPERATORS WITH MATRIX KERNEL ON MATRIX WEIGHTED SPACES

SANDRA POTT AND ANDREI STOICA

Abstract. For a matrix $A_2$ weight $W$ on $\mathbb{R}^p$, we introduce a new notion of $W$-Calderón-Zygmund matrix kernels, following earlier work in [11]. We state and prove a $T_1$ theorem for such operators and give a representation theorem in terms of dyadic $W$-Haar shifts and paraproducts, in the spirit of [7]. Finally, by means of a Bellman function argument, we give sharp bounds for such operators in terms of bounds for weighted matrix martingale transforms and paraproducts.

1. Introduction

One of the most important topics in Harmonic Analysis is the study of weighted norm inequalities for Calderón-Zygmund operators. This goes back to the 1970’s with the now classical works of R.A. Hunt, B. Muckenhoupt and R.L. Wheeden [10] and R.R. Coifman and C. Fefferman [2]. While the equivalence between the boundedness of Calderón-Zygmund singular integral operators on the weighted space $L^p(w)$ and the membership of the scalar weight $w$ to the so-called $A_p$ class has been shown in the above papers, the relation between the operator norm of a Calderón-Zygmund operator and the $A_p$ characteristic of the weight, $[w]_{A_p}$, remained open for quite some time. In the case $p = 2$, the so-called $A_2$ conjecture (now theorem) stated that this dependence is linear in $[w]_{A_2}$. It took roughly the first decade of the 2000’s to go from the J. Wittwer's proof of the conjecture for the martingale and Hilbert transform on the $A_2$ conjecture for the dyadic martingale transforms (see [21]) to the proof for general Calderón-Zygmund operators by T. Hytönen [7]. Important contributions have been brought by S. Hukovic, S. Treil and A. Volberg [13], who proved the conjecture for the dyadic square function, by S. Petermichl and A. Volberg [19], who showed the linear dependence for the Beurling-Ahlfors transform, and by S. Petermichl [16], who proved it for the Hilbert and Riesz transforms.

There are several ways to extend the classical (scalar) theory of Calderón-Zygmund operators. One of these extensions assumes that the functions are $C^d$-valued, but the kernel of the Calderón-Zygmund operators is still scalar-valued. Inspired by applications to multivariate stationary stochastic processes, an analogue notion of the Muckenhoupt $A_p$ weights, the matrix $A_p$ weights, have been introduced by S. Treil and A. Volberg (see [21]). M. Goldberg [3], F. Nazarov and S. Treil [13] and A. Volberg [22], showed that certain Calderón-Zygmund operators are bounded on $L^p(W)$ with $1 < p < \infty$ if $W$ is a matrix $A_p$ weight. However, a corresponding $A_2$ conjecture in the matrix setting is still open and currently attracting much interest. In [1], K. Bickel, S. Petermichl and B. Wick modified a scalar argument to show that the dependence of the norm of the martingale and Hilbert transform on the $A_2$ characteristic of the weight $W$ is at most $[W]_{A_2}^{3/2} \log [W]_{A_2}$. The authors of the current paper proved in [18] that the dependence of the norm of all Calderón-Zygmund operators with cancellation on $[W]_{A_2}$ is the same as the one for the matrix martingale transform, thus reducing the $A_2$ conjecture for all such operators to the proof of the linear bound for the latter operator. In a recent paper, F. Nazarov, S. Petermichl, S. Treil and A. Volberg [15] showed that for all Calderón-Zygmund operators the dependence is no worse than $[W]_{A_2}^{3/2}$. The first sharp results in this direction have recently been obtained by T. Hytönen, S. Petermichl and A. Volberg [10], and by Isralowitz, Kwon, and the first author in [12]. Here, the linear (in $[W]_{A_2}$) upper bound was proved for the matrix-weighted square function $S_W$, which can be seen as an average of the

2010 Mathematics Subject Classification. 42B20, 60G46, 46B09, 46B28, 26B25.
Key words and phrases. Calderón-Zygmund operator, matrix $A_2$ weights, weighted $L^2$ spaces, martingale transform, Bellman function, dyadic Haar shift.
matrix martingale transforms we consider, and for the matrix-weighted maximal function $M'_W$, respectively.

The theory can be generalised even further by taking Calderón-Zygmund operators with matrix-valued kernel. These operators appear naturally in geometric function theory, multivariate prediction theory or in the study of Toeplitz operators. On a matrix-weighted space, such an operator associated to a matrix kernel can now no longer be considered on its own, but has to be considered together with the matrix weight.

For the case of dyadic paraproducts, this was done in [12], where the correct version of matrix weighted BMO spaces was introduced, and boundedness of the dyadic paraproducts $\Pi_B$ on $L^p(W)$ was characterised by means of a matrix weighted Carleson embedding theorem (the necessary definitions and results are given in the following section). Building on this work and inspired by the proof of the scalar $A_2$ theorem of T. Hytönen, J. Isralowitz [11] then introduced a notion of $W$-Calderón-Zygmund operators and proved a matrix weighted $T1$ theorem for these operators on $L^p(W)$, $1 < p < \infty$.

In the current paper, we only consider the case $p = 2$, and we state and prove a $T1$-Theorem on $L^2(W)$. Compared to the results in [11], our aim is two-fold. First, and most importantly, we give sharp bounds in the $T1$ theorem, in terms of bounds for matrix martingale transforms and bounds for matrix paraproducts on matrix-weighted spaces. The key for this result is the notion of $W$-dyadic Haar shifts, which we introduce, and sharp bounds for these shifts (Theorem 2.2). The proof of the sharp bound relies on a Bellman function argument for matrix weights in [18], which was originally inspired by Treil’s work in the scalar case [20]. The sharp bounds for matrix-weighted martingale transforms and paraproducts are conjectured to be linear in the $A_2$-characteristic, but this has so far been out of reach. However, the bounds we can prove come close to the best known bounds for scalar kernels and matrix weights, which were recently proved by [15] by means of convex Lerner-type operators.

Note that these results do not easily give rise to to sharp bounds for $p \neq 2$, since no extrapolation method is known in the matrix setting, and we do not have any generalisation of our Bellman function approach for $p \neq 2$.

Our second aim is to give a more general definition of $W$-Calderón-Zygmund kernel than [11], which has in particular local decay and smoothness conditions. Moreover, our $T1$ theorem works with a natural weak boundedness condition which is easily seen to be necessary.

The paper is organised as follows. In Section 2 we state the necessary definitions, some background, and the main results, Theorems 2.6 and 2.2. In Section 3 we prove the $T1$ theorem 2.6 up the bound from Theorem 2.2. Finally, in Section 4 we prove Theorem 2.2 with a Bellman function argument.

2. DEFINITIONS AND STATEMENT OF THE MAIN RESULTS

2.1. Matrix $A_2$ and $A_\infty$ weights. For $p, d \geq 1$, the non-weighted Lebesgue space $L^2(\mathbb{R}^p)$ consists of all measurable functions $f : \mathbb{R}^p \to \mathbb{C}^d$ with norm

$$\|f\|_{L^2(\mathbb{R}^p)} := \left( \int_{\mathbb{R}^p} \|f(t)\|_{\mathbb{C}^d}^2 \, dt \right)^{1/2} < \infty.$$  

We will also use the space $C^1_{\text{c}}(\mathbb{R}^p)$ of compactly supported, continuously differentiable functions $f : \mathbb{R}^p \to \mathbb{C}^d$.

Let $\mathcal{M}_d(\mathbb{C})$ be the space of $d \times d$ complex matrices. A matrix weight on $\mathbb{R}^p$ is a measurable locally integrable function $W : \mathbb{R}^p \to \mathcal{M}_d(\mathbb{C})$ whose values are almost everywhere positive definite. We define $L^2(W)$ to be the space of measurable functions $f : \mathbb{R}^p \to \mathbb{C}^d$ with norm

$$\|f\|_{L^2(W)}^2 = \int_{\mathbb{R}^p} \|W^{1/2}(t)f(t)\|_{\mathbb{C}^d}^2 \, dt = \int_{\mathbb{R}^p} \langle W(t)f(t), f(t) \rangle \, dt < \infty.$$  

It is well-known that the dual of $L^2(W)$ can be identified with $L^2(W^{-1})$, where the duality between these two spaces is given by the standard inner product.
We say that a matrix weight $W$ satisfies the matrix $A_2$ Muckenhoupt condition, if
\begin{equation}
[W]_{A_2} := \sup_I \left\| \left( \frac{1}{|I|} \int_I W(t) \, dt \right)^{1/2} \left( \frac{1}{|I|} \int_I W^{-1}(t) \, dt \right)^{1/2} \right\| < \infty,
\end{equation}
where the supremum is taken over all cubes $I \subset \mathbb{R}^d$, and $\| \cdot \|$ denotes the norm of the matrix acting on $\mathbb{C}^d$. The number $[W]_{A_2}$ is called the $A_2$ characteristic of the weight $W$. We say that a matrix weight $W$ satisfies the dyadic matrix Muckenhoupt condition $A_d^d$ on $\mathbb{R}^d$ or $\mathbb{R}$, if (1) is satisfied, but with the supremum being now taken only over dyadic cubes or intervals, respectively (see [21]).

Furthermore, we say that a matrix weight $W$ satisfies the weak matrix $A_\infty$ condition, if the scalar weights $\langle We, e \rangle$ are $A_\infty$-weights in the scalar sense, uniformly for all $e \in \mathbb{C}^d$. We define the weak matrix $A_\infty$ characteristic $[W]_{A_\infty}$ by
\begin{equation}
[W]_{A_\infty} := \sup_{e \in \mathbb{C}^d, e \neq 0} \left[ \langle We, e \rangle \right]_{A_\infty,F_W}
\end{equation}
where $[\cdot]_{A_\infty,F_W}$ denotes the Fuji-Wilson version of the scalar $A_\infty$-constant, see e.g. [23].

To our knowledge, so far sharp bounds in terms of matrix $A_2$ and weak matrix $A_\infty$ characteristic, which correspond to the known sharp bounds in the scalar case, are known for only two of the important operators in this setting. Both will play an important role in the following. One is the matrix dyadic weighted square function,
\begin{equation}
S_W f(t) = \left( \sum_{\varepsilon \in \{0,1\}^d \setminus \{0\}} \sum_{I \in \mathcal{D}_0} \frac{X_I(t)}{|I|} \left\| \langle W \rangle_I^{1/2} \langle f, h_I^*_1 \rangle \right\|_2^2 \right)^{1/2}
\end{equation}
(see also [17], [1]). For the notation on dyadic cubes and Haar coefficients, see Subsection 2.2 below. The sharp bound
\begin{equation}
\|S_W\|_{L^2(W) \rightarrow L^2} \leq C_{p,d}[W]_{A_2}^{1/2}[W]_{A_\infty}^{1/2}
\end{equation}
was very recently proved in [10]. A natural conjecture, which currently remains open, is the following lower bound corresponding to the scalar weight case,
\begin{equation}
\|f\|_{L^2(W)} \leq C_{p,d}[W]_{A_\infty}^{1/2} \|S_W f\|_{L^2(W)}.
\end{equation}
The best known lower bound appears in [1], following an argument from [21]. It is
\begin{equation}
\|f\|_{L^2(W)} \leq C_{p,d}[W]_{A_2}^{1/2} \log([W]_{A_2})^{1/2} \|S_W f\|_{L^2(W)}.
\end{equation}

The second operator, for which a sharp bound is known, is the matrix-weighted maximal function $M'_W$,
\begin{equation}
M'_W f(t) = \sup_{t \in I, I \text{ cube}} \frac{1}{|I|} \int_I \left\| \langle W \rangle_I^{1/2} f(x) \right\|_2 \, dx.
\end{equation}
It was shown in [12] that the bound
\begin{equation}
\|M'_W\|_{L^2(W) \rightarrow L^2} \leq C_{p,d}[W]_{A_2}
\end{equation}
holds, which is the best possible in terms of the $A_2$-characteristic. However, one easily sees that the proof in [12] gives actually a slightly better bound, namely
\begin{equation}
\|M'_W\|_{L^2(W) \rightarrow L^2} \leq C_{p,d}[W]_{A_2}^{1/2}[W]_{A_\infty}^{1/2}.
\end{equation}
This can be seen by noticing that the $\epsilon$ chosen in the proof of (5) for use in the reverse Hölder property can actually be chosen as $[W]_{A_\infty}^{-1}$, according to the sharp reverse Hölder result for $A_\infty$-weights proved in [9].
2.2. **Dyadic setting.** The standard system of dyadic cubes in \( \mathbb{R}^p \) is

\[
D^0 := \bigcup_{j \in \mathbb{Z}} D_j^0, \quad \text{where} \quad D_j^0 := \{2^{-j}([0,1)^p + k) : k \in \mathbb{Z}^p\}.
\]

Given a sequence \( \omega = \{\omega_i\}_{i \in \mathbb{Z}} \subset \{0,1\}^p \), the dyadic system \( D^\omega \) in \( \mathbb{R}^p \) is defined by \( D^\omega := \{ I + \omega : I \in D^0 \} \), where the translated dyadic cube \( I + \omega \) is defined as

\[
I + \omega := I + \sum_{i:2^{-\ell(I)} < \ell(I)} 2^{-i} \omega_i.
\]

When the particular choice of \( \omega \) is not important, we will use the notation \( D \) for a generic dyadic system. We equip the set \( \Omega := \{0,1\}^p \) with the canonical product probability measure \( \mathbb{P}_\Omega \) which makes the coordinates \( \omega_i \) independent and identically distributed: the probability of each coordinate \( \omega_i \) taking any of the values in \( \{0,1\} \) is \( 2^{-p} \). We denote by \( \mathbb{E}_\Omega \) the expectation over the random variables \( \omega_i, i \in \mathbb{Z} \).

Let us introduce a few useful notations. For a cube \( I \in D \), let \( \ell(I) \) and \( |I| \) denote its side length and volume, respectively. Let

\[
\mathcal{D}_n(I) := \{ J \in D : J \subseteq I, \ell(J) = 2^{-n} \ell(I) \}
\]

be the collection of \( n \)-th generation children of \( I \). For any dyadic cube \( I \in D \), we will denote its parent by \( \hat{I} \).

Any system of dyadic cubes \( D \) has an associated function system, the Haar functions. When \( p = 1 \), any dyadic interval \( I \) has two Haar functions associated with it:

\[
h^0_I = |I|^{-1/2} \chi_I, \quad h^1_I = |I|^{-1/2} (\chi_{I^+} - \chi_{I^-}),
\]

where \( \chi_I \) is the characteristic function of the interval \( I \), and \( I^+ \) and \( I^- \) are the left and right children of \( I \), respectively.

If \( p > 1 \), the Haar functions associated to a cube \( I = I_1 \times \cdots \times I_p \) are

\[
h^\varepsilon_I(x) = h^\varepsilon_{I_1 \times \cdots \times I_p}(x_1, \cdots, x_p) = \prod_{i=1}^p h^\varepsilon_{I_i}(x_i),
\]

where \( \varepsilon \in \{0,1\}^p \).

It is well known that the Haar functions \( h^\varepsilon_I \), with \( I \in D \) and \( \varepsilon \in \{0,1\}^p \setminus \{0\} \), form an orthogonal basis of \( L^2(\mathbb{R}^p) \). Hence any function \( f \in L^2(\mathbb{R}^p) \) admits the orthogonal expansion

\[
f = \sum_{I \in D} \sum_{\varepsilon \in \{0,1\}^p \setminus \{0\}} \langle f, h^\varepsilon_I \rangle h^\varepsilon_I.
\]

We denote the average of a locally integrable function \( f \) on the cube \( I \) by \( \langle f \rangle_I := |I|^{-1} \int_I f(t) \, dt \).

2.3. **Matrix martingale transforms and dyadic \( W \)-Haar shifts.** Let \( W \) be a matrix weight. For a sequence of \( d \times d \) matrices \( \sigma = \{\sigma_I\}_{I \in D} \), we introduce the notation

\[
\|\sigma\|_{\infty,W} := \max_{I \in D} \left\| \mathbb{E}_{\Omega} \left[ \frac{1}{|I|^{1/2}} W^\ell(I)^{1/2} \sigma_I \right] \right\|_{L^2(W)}.
\]

For a sequence \( \sigma \) such that \( \|\sigma\|_{\infty,W} < \infty \), we define the martingale transform operator \( T_\sigma \) by

\[
T_\sigma f = \sum_{Q \in D} \sum_{\varepsilon \in \{0,1\}^p \setminus \{0\}} \langle f, h^\varepsilon_Q \rangle h^\varepsilon_Q.
\]

The condition \( \|\sigma\|_{\infty,W} < \infty \) is equivalent to the boundedness of \( T_\sigma \) on \( L^2(W) \) (see, e.g., Proposition 1.6 in \( [22] \) and Theorem 5.2 in \( [1] \) and for an explicit statement; it is also contained in \( [21] \)). The martingale transform is considered a good model for Calderón-Zygmund singular integral operators.

We define the function \( N : [1, \infty) \to [1, \infty) \) by

\[
N(X) = \sup \|T_\sigma\|_{L^2(W) \to L^2(W)},
\]
where the supremum is taken over all $d \times d$ matrix $A_d^2$ weights $W$ with $|W|_{A_d^2} \leq X$ and $\|\sigma\|_{\infty,W} \leq 1$. It was shown in \cite{1} that
\begin{equation}
N(X) \lesssim (\log X)X^{3/2}.
\end{equation}
From the estimates \cite{2} and \cite{3} mentioned above, one obtains immediately the slightly improved bond
\begin{equation}
\|T\sigma\|_{L^2(W)\rightarrow L^2(W)} \leq C_{p,d} \log(|W|_{A_d^2})^{1/2}[W]_{A_d^2}[W]_{A_\infty}^{1/2}
\end{equation}
and thus $N(X) \lesssim (\log X)^{1/2}X^{3/2}$. The linear bound $N(X) \lesssim X$ is conjectured, but this seems out of reach at the moment. If the conjectured lower bound for the matrix weighted dyadic square function \cite{4} holds, \cite{5} could be improved to
\begin{equation}
\|T\sigma\|_{L^2(W)\rightarrow L^2(W)} \leq C_{p,d}[W]_{A_d^2}^{1/2}[W]_{A_\infty}^{1/2} [W^{-1}]_{A_\infty}^{1/2},
\end{equation}
which is the best currently known bound for scalar kernels and matrix weights, even in the case of scalar Lerner operators \cite{6}.

We will now go one step further and allow off-diagonal terms of the operator in the Haar expansion:

**Definition 2.1.** A (cancellative) dyadic W-Haar shift on $\mathbb{R}^p$ of parameters $(m,n)$, with $m,n \in \mathbb{N} \cup \{0\}$, is an operator of the form
\begin{equation}
Sf = \sum_{L \in D} \sum_{e \in \{0,1\}^p \setminus \{0\}} \sum_{J \in D_m(L)} A_{f,J,e,e'}^{l,J,e,e'}(f,h_J^e)L^2(\mathbb{R}^p)h_J^e',
\end{equation}
where $A_{f,J,e,e'}^{l,J,e,e'} \in \mathcal{M}_d(\mathbb{C})$ such that $\left\| (W)_L^{1/2} A_{f,J,e,e'}^{l,J,e,e'}(W)_L^{-1/2} \right\| \leq \sqrt{|I|/|J|} = 2^{-\left(\frac{m+n}{2}\right)}$, and $f$ is any locally integrable function. The number $k := \max\{m,n\} + 1$ is called the complexity of the W- Haar shift.

For $0 \leq t \leq k - 1$ we introduce the notation $\mathcal{L}_t := \{ I \in D : \ell(I) = 2^{t+kq}, q \in \mathbb{Z} \}$, and define the slice $S_t$ of the shift $S$ by
\begin{equation}
S_t f = \sum_{L \in \mathcal{L}_t} \sum_{e \in \{0,1\}^p \setminus \{0\}} \sum_{J \in D_m(L)} A_{f,J,e,e'}^{l,J,e,e'}(f,h_J^e)L^2(\mathbb{R}^p)h_J^e',
\end{equation}
where $S_t$ also acts on locally integrable functions. We can thus decompose $S$ as $S = \sum_{t=0}^{k-1} S_t$. These slices $S_t$ can be seen as martingale transforms when we are moving $k$ units of time at once, which we will exploit in the proof of Theorem 2.2 below in Section 3.

**Theorem 2.2.** Let $W$ be a matrix $A_d^2$ weight and $S$ be a dyadic W- Haar shift on $\mathbb{R}^p$ of complexity $k$. Then
\begin{equation}
\|S\|_{L^2(W)\rightarrow L^2(W)} \leq C_p \cdot kp \delta^d N(|W|_{A_d^2}),
\end{equation}
where $C_p$ is a constant depending only on $p$.

2.4. Matrix BMO space and paraproducts. We now introduce the appropriate notion of BMO space for our $T_1$ theorem:

**Definition 2.3 (\cite{12}).** If $W$ is a matrix $A_2$ weight, we say that a locally integrable function $B : \mathbb{R}^p \rightarrow \mathcal{M}_d(\mathbb{C})$ belongs to $\text{BMO}_W(\mathbb{R}^p)$ if
\begin{equation}
\sup_{I \text{ cube}} \frac{1}{|I|} \int_I \| W^{1/2}(x)(B(x) - (B)_I)(W)_I^{-1/2} \|^2 dx < \infty.
\end{equation}
The corresponding dyadic space $\text{BMO}_W^d$ is obtained by taking the supremum only over dyadic cubes $I \in D$.

It was proved in \cite{12}, Theorem 1.3 and Corollary 1.5, that for a matrix $A_2$ weight $W$, this $\text{BMO}_W^d$ norm is equivalent to the norm given by the square root of
\begin{equation}
\sup_{J \in D} \frac{1}{|J|} \sum_{I \subseteq J} \sum_{e \in \{0,1\}^p \setminus \{0\}} \left\| (W)_I^{1/2} (B_I^e)^* B_I^{e'} (W)_I^{-1/2} \right\|,
\end{equation}
where $B_I^\varepsilon$ is the matrix of Haar coefficients of the entries of $B$ with respect to $h_I^\varepsilon$. In case that $B$ is a scalar-valued function, this coincides with the usual scalar dyadic BMO norm. We will therefore consider the norm given by (9) as the natural norm on $\text{BMO}_d^0$.

We will now introduce the dyadic paraproducts with matrix coefficients and their adjoints.

**Definition 2.4.** If $B: \mathbb{R}^p \to \mathcal{M}_d(\mathbb{C})$ and $f : \mathbb{R}^p \to \mathbb{C}^d$ are locally integrable functions, then the dyadic paraproduct $\Pi_B$ with respect to a dyadic grid $\mathcal{D}$ is defined as

$$
\Pi_B f = \sum_{I \in \mathcal{D}} \sum_{\varepsilon \in \{0,1\}^p \setminus \{0\}} B_I^\varepsilon(f) h_I^\varepsilon.
$$

The formal adjoint of $\Pi_B$ is the operator $\Pi_B^* = (\Pi_B)^*$ given by

$$
\Pi_B^* f = \sum_{I \in \mathcal{D}} \sum_{\varepsilon \in \{0,1\}^p \setminus \{0\}} (B^*)_I^\varepsilon(f, h_I^\varepsilon)_{L^2(\mathbb{R}^p)} \chi_I.
$$

It was shown in [12] that if $W$ is a matrix $A_2$ weight, then $\Pi_B$ is bounded on $L^2(W)$ if and only if $B \in \text{BMO}_d^0$. For $X \geq 1$, we thus write $C_{\Pi}(X)$ for the smallest constant such that

$$
\|\Pi_B\|_{L^2(W) \to L^2(W)} \leq C_{\Pi}(X) \|B\|_{\text{BMO}_d^0}
$$

for all $B \in \text{BMO}_d^0$ and all matrix $A_2$ weights $W$ with $[W]_{A_2} \leq X$. Here as in the following, $\text{BMO}_d^0$ is equipped with the norm given by (9). It was shown in the proof of Theorem 1.3, (b) $\Rightarrow$ (a) in [12] that

$$
\|S_W \Pi_B\|_{L^2(W) \to L^2(\mathbb{R}^p)} \leq C_{p,d} \|M'_W\|_{L^2(W) \to L^2(\mathbb{R}^p)} \|B\|_{\text{BMO}_d^0} \leq \tilde{C}_{p,d} [W]_{A_2}^{1/2} [W]_{A_\infty}^{1/2} \|B\|_{\text{BMO}_d^0},
$$

and this is sharp, by comparison with the scalar case. Hence, using the lower bound in (4), we obtain

$$
\|\Pi_B\|_{L^2(W) \to L^2(W)} \leq \tilde{C}_{p,d} \log([W]_{A_2})^{1/2} [W]_{A_2} [W]_{A_\infty} \|B\|_{\text{BMO}_d^0}.
$$

and thus

$$
C_{\Pi}(X) \lesssim \log(X)^{1/2} X^{3/2}.
$$

We are now ready to consider Calderón-Zygmund operators with matrix kernels adapted to a weight $W$. Since the different values of the signature $\varepsilon$ do not play an important role, we will usually omit the sum in $\varepsilon$ from now on.

2.5. **Matrix-weighted Calderón-Zygmund operators.**

**Definition 2.5.** Let $\Delta = \{(x, x) : x \in \mathbb{R}^p\}$ be the diagonal of $\mathbb{R}^p \times \mathbb{R}^p$ and let $W$ be a matrix weight. We say that a function $K : \mathbb{R}^p \times \mathbb{R}^p \setminus \Delta \to \mathcal{M}_d(\mathbb{C})$ is a standard $W$-Calderón-Zygmund kernel, if there exists $\delta > 0$ and constants $C_0, C_\delta$ such that

$$
\|W_I^{1/2} K(x, y) (W_I)^{-1/2}\| \leq \frac{C_0}{|x-y|^p},
$$

$$
\|W_I^{1/2} (K(x, y) - K(x', y)) (W_I)^{-1/2}\| + \|W_I^{-1/2} (K^*(y, x) - K^*(y, x')) (W_I^{-1})^{1/2}\| \leq C_\delta \frac{|x-x'|^\delta}{|x-y|^{p+\delta}},
$$

for all cubes $I \subset \mathbb{R}^p$ and all points $x, x' \in I, y \in \mathbb{R}^p$ with $|x-y| > 2|x-x'|$.

The notion of $W$-Calderón-Zygmund kernels, albeit with more restrictive conditions, was first introduced by J. Isralowitz (see [11]).

Given a $W$-Calderón-Zygmund kernel $K$, an operator $T$, defined on the class of step functions (which is dense in $L^2(\mathbb{R}^p)$), is called a $W$-Calderón-Zygmund operator on $\mathbb{R}^p$ associated with $K$, if its satisfies the kernel representation

$$
Tf(x) = \int_{\mathbb{R}^p} K(x, y) f(y) \, dy, \quad x \notin \text{supp } f.
$$
Note that for a matrix $A_2$ weight $W$, it follows immediately from the definition that $K(x,y)$ is a standard $W$-Calderón-Zygmund kernel if and only if the adjoint kernel $K^*(y,x)$ is a $W^{-1}$-Calderón-Zygmund kernel. Moreover, the $W$-Calderón-Zygmund operator $T$ is associated to $K(x,y)$, if and only if the $W^{-1}$-Calderón-Zygmund operator $T^*$ is associated to $K^*(y,x)$.

Generally, for suitable scalar functions $f$ and $g$, we will write $\langle Tf, g \rangle$ for the $d \times d$ matrix with entries

$$\langle Tf, g \rangle_{ij} = \langle Tf_{e_i}, g_{e_i} \rangle = \langle Tf, g_{e_i} \otimes e_j \rangle_{S_2 \otimes L^2(\mathbb{R}^p)}.$$  

Furthermore, we want to say that $T$ satisfies the $W$-weak boundedness property, if there exists a constant $C_{WBP} > 0$ such that

$$\|\langle W \rangle_{I}^{1/2} T \langle W \rangle_{I}^{-1/2} \chi_{I_1} \chi_{I_2} \| \leq C_{WBP} |I|,$$

for all cubes $I$ and all first-generation dyadic children $I_i$ of $I$.

Let $T$ be a $W$-Calderón-Zygmund operator as above which satisfies the $W$-weak boundedness property (13). Even though $T$ does not formally act on the constant function 1, we can define the Haar coefficients of $T^1$ (and, similarly, of $T^*1$) in the following way: We have

$$\langle T^1, h_I \rangle = \langle 1, T^* h_I \rangle = \int_{\mathbb{R}^p} (T^* h_I)^*(x) \, dx$$

$$= \int_{|I|} (T^* h_I)^*(x) \, dx \quad + \int_{|I|^{c}} (T^* h_I)^*(x) \, dx$$

$$= \int_{|I|} (T^* h_I)^*(x) \, dx \quad + \int_{|I|^{c}} K(y,x)h_I(y) \, dy \, dx$$

$$= \int_{|I|} (T^* h_I)^*(x) \, dx \quad + \langle W \rangle_{I}^{-1/2} \int_{|I|} \langle W \rangle_{I}^{1/2} (K(y,x) - K(c_I, x)) \langle W \rangle_{I}^{-1/2} h_I(y) \, dy \, dx \langle W \rangle_{I}^{1/2},$$

where $c_I$ is the centre of $I$. The first integral is well-defined by repeated use of the $W$-weak boundedness property (13). For the second part, we have

$$\|\langle W \rangle_{I}^{-1/2} \| \langle W \rangle_{I}^{1/2} \| \int_{|I|} \|\langle W \rangle_{I}^{1/2} (K(y,x) - K(c_I, x)) \| |I|^{-1/2} \| |I|^{-1/2} \| dy \, dx$$

$$\leq \|\langle W \rangle_{I}^{-1/2} \| \langle W \rangle_{I}^{1/2} \| |I|^{-1/2} C_{\delta} \int_{|I|^{c}} \int_{|x|} \langle W \rangle_{I}^{1/2} |y - c_I|^{4} |I|^{-1/2} \| |I|^{-1/2} \| dy \, dx$$

$$\leq C_{\delta} \|\langle W \rangle_{I}^{-1/2} \| \langle W \rangle_{I}^{1/2} \| |I|^{-1/2} \int_{|I|} \langle W \rangle_{I}^{1/2} |x| \| |x|^{1/2} |I|^{-1/2} \| dx |I| < \infty.$$  

Here, $\langle T^1, h_I \rangle$ is the matrix with the entries

$$\langle T^1 e_i, h_I e_i \rangle = \langle T^1, h_I e_i \otimes e_j \rangle_{S_2 \otimes L^2(\mathbb{R}^p)} := \int_{\mathbb{R}^p} \text{tr} (T^1(x)(e_j \otimes h_I(x) e_i)) \, dx,$$

where $\langle \cdot, \cdot \rangle_{S_2}$ denotes the Hilbert-Schmidt inner product on the space of $d \times d$ matrices, and $\{e_1, \ldots, e_d\}$ is the standard orthonormal basis of $\mathbb{C}^d$.

Since the Haar coefficients $\langle T^1, h_I \rangle$ and $\langle T^*1, h_I \rangle$ are well-defined, we can thus give meaning to the operators $\Pi_{T^1}$ and $\Pi_{T^*1}$, respectively.

We can now state the main result of the paper.

**Theorem 2.6.** Let $W$ be a $d \times d$ matrix $A_2$ weight on $\mathbb{R}^p$. Let $T$ be a $W$-Calderón-Zygmund operator on $\mathbb{R}^p$ associated to the matrix kernel $K$. Then the following are equivalent:

1. $T$ satisfies the $W$-weak boundedness property, $T^1 \in BMO_W$, and $T^*1 \in BMO_{W^{-1}}$.  
2. $T : L^2(W) \rightarrow L^2(W)$ is bounded. 
3. $T^* : L^2(W^{-1}) \rightarrow L^2(W^{-1})$ is bounded.

In particular, we have in this case the representation

$$\langle Tf, g \rangle = C\mathbb{E}_{1} \sum_{n,m} \tau(m,n) \langle S_{m,n}^w f, g \rangle + \mathbb{E}_{1} (\Pi_{T^1} f, g) + \mathbb{E}_{1} (\Pi_{T^*1}^* f, g)$$

for all $f \in C^1_{c}(\mathbb{R}^p, \mathbb{C}^d) \cap L^2(W)$ and all $g \in C^1_{c}(\mathbb{R}^p, \mathbb{C}^d) \cap L^2(W^{-1})$. Here, the $S_{m,n}^w$ are cancellative dyadic $W$-Haar shifts with respect to the grid $D^p$, $\tau(m,n) \leq 2^{-\delta(m+n)/4}(\max(m,n)+1)$, and $C$ is
We note that this probability is independent of the cube. The proof of our main result is based on the following random expansion of an operator $h$ of Haar functions (Hytönen [8], Hänninen, Hytönen [4]).

Proposition 2.10. Proposition 2.10 which is the best currently known bound for matrix weights and scalar kernels, see [15].

Remark 2.8. If the conjectured lower bound [3] holds for the weighted square function $S_W$, then this can be improved to

$$\|T\|_{L^2(W)\to L^2(W)} \leq C(1 + \|T\|_{BMO} + \|T^*\|_{BMO}) |W|^{3/2},$$

which is the best currently known bound for matrix weights and scalar kernels, see [15].

2.6. Good cubes, bad cubes, and the Representation Theorem. In the proof of our main result we will need the notion of "good" cubes, which was introduced in [14].

Definition 2.9. Let us fix a large parameter $r \in \mathbb{N}$. We say that a cube $I \in \mathcal{D}_\omega$ is bad, if there exists $J \in \mathcal{D}_\omega$ such that $\ell(I) \geq 2^r \ell(I)$ and

$$\text{dist}(I, \partial J) \leq \ell(I)^\gamma \ell(J)^{1-\gamma},$$

where $\gamma = \frac{\delta}{\pi(r+\pi)}$. A cube $I \in \mathcal{D}_\omega$ is called good, if it is not bad.

Note that our choice of $\gamma$ differs by a factor 2 from the usual one.

As was shown in [3], we can fix $r$ large enough such that

$$\pi_{\text{bad}} := \mathbb{P}_\Omega(\{\omega : I + \omega \text{ is bad}\}) < 1.$$  

We note that this probability is independent of the cube $I \in \mathcal{D}_\omega$. We also define $\pi_{\text{good}} := 1 - \pi_{\text{bad}}$.

The proof of our main result is based on the following random expansion of an operator $T$ in terms of Haar functions $h_I$, where the bad cubes are discarded.

Proposition 2.10 (Hytönen [3], Hänninen, Hytönen [3]). For all $f, g \in C^1_0(\mathbb{R}^p, \mathbb{C}^d)$, we have the expansion

$$\langle Tf, g \rangle_{L^2(\mathbb{R}^p)} = \frac{1}{\pi_{\text{good}}} \mathbb{E}_\Omega \sum_{I, J \in \mathcal{D}_\omega} \langle \langle Th_I, h_J \rangle \langle f, h_I \rangle_{L^2(\mathbb{R}^p)}, \langle g, h_J \rangle_{L^2(\mathbb{R}^p)} \rangle_{\mathbb{C}^d},$$

where

$$\text{smaller}\{I, J\} := \begin{cases} I, & \text{if } \ell(I) \leq \ell(J), \\ J, & \text{otherwise}. \end{cases}$$

Remark 2.11. This follows as in the proof of [3], Prop. 3.5 (see also [4], Cor. 6.3). These papers use slightly different conditions on $T$, but we only need here that the inner product $\langle Tf, g \rangle$ can be expanded in the Haar basis for each of the dyadic grids $\mathcal{D}^\omega$, and this is for example ensured by $f, g \in C^1_0(\mathbb{R}^p)$, a (large) a priori bound on the norms of $W, W^{-1}$, and the conditions (1) in the $T_1$ Theorem 2.7, see Lemma 2.7.1 below. We should mention that this version is a particular case of Corollary 6.3 in [3].

We only need the result for $C^d$-valued functions instead of functions taking values in an arbitrary Banach space $E$, whereas our kernels are matrix-valued (they are operator-valued in [4]). Therefore, the Rademacher $R$-bounds reduce to uniform bounds in our case.

Let us first mention how Proposition 2.10 can be used to prove the main result. For the moment, we fix $\omega \in \Omega$ and focus on the sum inside $\mathbb{E}_\Omega$; for notational ease, we also drop the index $\omega$.

Following [11], we extract the paraproducts by considering the operator $\tilde{T} := T - \Pi_{T_1} - \Pi_{(T_1)^{-1}}$.

We will now show how to identify the sum involving $\tilde{T}$ as a sum of dyadic $W$-shifts. In order to do that, the sum is rearranged according to the minimal common dyadic ancestor of $I$ and $J$, taking values in an arbitrary Banach space $E$, whereas our kernels are matrix-valued (they are operator-valued in [4]). Therefore, the Rademacher $R$-bounds reduce to uniform bounds in our case.
which is denoted by $I \vee J$ (if $I \subseteq J$, then $I \vee J = J$; if $I \cap J = \emptyset$, Lemma 3.7 in [8] shows the existence of a common dyadic ancestor).

Splitting the sum according to which of the cubes $I$ and $J$ has smaller side length (and hence is good), then rearranging the sum according to which cube $L$ is the minimal common dyadic ancestor $I \vee J$, and what the size of $I$ and $J$ relative to $L$ is, we obtain

$$
\sum_{I,J \text{ smaller } \{I,J\} \text{ is good}} = \sum_{m \geq n} \sum_{L} \sum_{I,J : I \vee J = L \text{ is good}} + \sum_{m < n} \sum_{L} \sum_{I,J : I \vee J = L \text{ is good}}.
$$

If we write

$$
\sum_{L} \sum_{I,J : I \vee J = L \text{ is good}} \langle \langle T h_I, h_J \rangle \rangle f, g \rangle_{L^2(\mathbb{R}^p)} = \langle \langle S_{mn} f, g \rangle \rangle_{L^2(\mathbb{R}^p)},
$$

we get that

$$
\langle T f, g \rangle_{L^2(\mathbb{R}^p)} = \frac{1}{\pi_{\text{good}}} \mathbb{E}_{m,n} \sum_{S_{mn}} \langle S_{mn} f, g \rangle_{L^2(\mathbb{R}^p)} + \frac{1}{\pi_{\text{good}}} \mathbb{E}_{m,n} \langle (\Pi_{T^*} + (\Pi_{T^*})^*) f, g \rangle_{L^2(\mathbb{R}^p)}.
$$

3. The proof of Theorem 2.6

The equivalence of (2) and (3) in Theorem 2.2 is immediate. We will first consider the sufficiency direction (1) $\Rightarrow$ (2).

Since we already have the necessary bounds for the paraproducts by (11) and (12), it remains to study the dyadic shifts $S_{mn}$ that appear in the expansion (17) of the operator. We will show in this section that up to a constant $C$ depending only on $C_0, C_3$ in the $W$-Calderón-Zygmund kernel conditions, the $W$-weak boundedness constant $C_{WBC}$, and the dyadic shifts $S_{mn}, m$ are of the form $2^{-\delta(m+n)/4} S_{mn}$, where $S_{mn}$ is a dyadic $W$-Haar shifts in the sense of Definition 2.1. Theorem 2.2 then yields the necessary bounds for these shifts, which guarantee convergence, thereby proving the representation formula (14), and moreover prove Corollary 2.7.

Before we proceed, let us introduce some more useful notations. We first fix a dyadic lattice $\mathcal{D}$ in $\mathbb{R}^p$; all dyadic operators will be considered with respect to this grid $\mathcal{D}$. As before, let $\bar{T} := T - \Pi_{T^*} - \Pi_{T^*}^*].$ For $I, J \in \mathcal{D}$, define the matrix $\bar{T}_{I,J}$ as $\bar{T}_{I,J} = \langle \bar{T} h_I, h_J \rangle$ (also define $\bar{T}_{I,J}$ in a similar way). Moreover, for any fixed dyadic cube $L$, let $\sqrt[L]{\bar{T}} := \langle W \rangle^{1/2} \bar{T} \langle W \rangle^{1/2}$ and $\sqrt[L]{\bar{T}}_{I,J} := \langle W \rangle^{1/2} \bar{T}_{I,J} \langle W \rangle^{1/2}$ (and similarly define $T^L$ and $T^L_{I,J}$).

In the following lemma we prove that the dyadic shifts $S_{mn}$ are $W$-Haar shifts. We are only considering the case $m \geq n$ (which means $\ell(I) = \ell(J)$), since the case $m < n$ can be treated similarly by duality. A version of this lemma was stated and proved in [11], and the proof here runs along the same lines. However, we work with different notions of $W$-weak boundedness and $W$-Calderón-Zygmund kernels, which changes some arguments. For clarity, we give the whole proof here.

**Lemma 3.1.** Let $m, n \geq 0, n \leq m$.

Let $W$ be a matrix $A_2$ weight on $\mathbb{R}^p$ and let $T$ be a $W$-Calderón-Zygmund operator on $\mathbb{R}^p$, which satisfies the $W$-weak boundedness condition in Theorem 2.6.

Fix a cube $L \in \mathcal{D}$. If $I$ and $J$ are two cubes such that $L$ is their smallest common dyadic ancestor, $\ell(I) = 2^{-m} \ell(L), \ell(J) = 2^{-n} \ell(L)$, and $I$ is good, then

$$
\|\bar{T}_{I,J}\| \leq C \frac{\ell(I) \ell(J)}{D(I,J)^{p+1}} \leq C \sqrt{\ell(I) \ell(J)} \frac{m+n}{|L|} 2^{-\frac{(m+n)\ell}{4}},
$$

where the long distance $D(I, J)$ is defined as $D(I, J) := \ell(I) + \ell(J) + \text{dist}(I, J)$, and $C, \tilde{C}$ are constants depending only on $p,d$, the smoothness and decay constants $C_{\delta}, C_0$ of the $W$-Calderón-Zygmund kernel, and the $W$-weak boundedness constant $C_{WBC}$. 


Proof. Before we proceed, note that $\Pi_T h_I$ is contained in the span of the Haar functions $h_K$, $K \subseteq I$, with $d \times d$ matrix coefficients, and $\Pi_{(T-1)^*} h_I$ is a multiple of $\frac{h}{|T|}$ with a matrix coefficient. In particular, the support of both $\Pi_T h_I$ and $\Pi_{(T-1)^*} h_I$ is contained in $I$.

Following [8] and [11], we decompose the set $\Gamma := \{(I, J) \in D \times D : \ell(I) \leq \ell(J)\}$ as
\[
\{(I, J) \in D \times D : \ell(I) \leq \ell(J)\} = \{(I, J) \in D \times D : I \subseteq J\} \cup \{(I, J) \in D \times D : I = J\}
\]
\[
\cup \{(I, J) \in D \times D : \text{dist}(I, J) > \ell(I)^\gamma \ell(J)^{-\gamma}\}
\]
\[
\cup \{(I, J) \in D \times D : I \cap J = \emptyset, \text{dist}(I, J) \leq \ell(I)^\gamma \ell(J)^{-\gamma}\}
\]
\[
= : \Gamma_{\text{in}} \cup \Gamma_{\text{out}} \cup \Gamma_{\text{near}}.
\]

We will now estimate $\|\tilde{T}_{I,J}^L\|$ for $(I, J)$ in each of these sets.

Case 1: $(I, J) \in \Gamma_{\text{in}}$

It is not difficult to check that $(\tilde{T}^L)^* 1 = 0$, in the sense that $\int_{\mathbb{R}^n} \tilde{T}^L h_I(x) dx = 0$ for each cube $I$. Indeed, we have
\[
\langle h_I, (\tilde{T}^L)^* \rangle = \langle \tilde{T}^L h_I, 1 \rangle = \langle \langle W \rangle_L^{1/2} (T - \Pi_T - \Pi_{(T-1)^*}) \langle W \rangle_L^{-1/2} h_I, 1 \rangle
\]
\[
= \langle T^L h_I, 1 \rangle - \langle \langle W \rangle_L^{1/2} \Pi_T \langle W \rangle_L^{-1/2} h_I, 1 \rangle - \langle \langle W \rangle_L^{1/2} \Pi_{(T-1)^*} \langle W \rangle_L^{-1/2} h_I, 1 \rangle.
\]

Using the definitions of the paraproduct and its adjoint, we get that $\langle \langle W \rangle_L^{1/2} \Pi_T \langle W \rangle_L^{-1/2} h_I, 1 \rangle = 0$, since $\langle h_K, 1 \rangle_{L^2(\mathbb{R}^n)} = 0$ for any $K \in D$, and
\[
\langle \langle W \rangle_L^{1/2} \Pi_{(T-1)^*} \langle W \rangle_L^{-1/2} h_I, 1 \rangle = \langle \langle W \rangle_L^{1/2} \langle (T^*)^* h_I \rangle \langle W \rangle_L^{-1/2} \chi_I, 1 \rangle
\]
\[
= \langle W \rangle_L^{1/2} \langle h_I, T^{*1} \rangle \langle W \rangle_L^{-1/2}
\]
\[
= \langle W \rangle_L^{1/2} \langle T h_I, 1 \rangle \langle W \rangle_L^{-1/2}.
\]

The last term is equal to
\[
\langle \langle W \rangle_L^{1/2} T h_I \langle W \rangle_L^{-1/2}, 1 \rangle = \langle \langle W \rangle_L^{1/2} T \langle W \rangle_L^{-1/2} h_I, 1 \rangle,
\]
which is exactly $\langle T^L h_I, 1 \rangle$. Therefore $\int_{\mathbb{R}^n} \tilde{T}^L h_I(x) dx = \langle \tilde{T}^L h_I, 1 \rangle = 0$.

Let $J_1$ be the child of $J$ that contains $I$. Then
\[
\langle \tilde{T}^L h_I, h_J \rangle = \langle \tilde{T}^L h_I, \chi_{J_1} h_J \rangle + \langle \tilde{T}^L h_I, \chi_{J_1^c} h_J \rangle
\]
\[
= \langle \tilde{T}^L h_I, \chi_{J_1} h_J \rangle + \langle h_J, J_1 \rangle \langle \tilde{T}^L h_I, \chi_{J_1} \rangle
\]
\[
= \langle \tilde{T}^L h_I, \chi_{J_1} (h_J - \langle h_J, J_1 \rangle) \rangle + \langle h_J, J_1 \rangle \langle \tilde{T}^L h_I, 1 \rangle,
\]

hence $\|\tilde{T}_{I,J}^L\| = \|\tilde{T}^L h_I, h_J\| \leq 2 |J_1|^{-1/2} \int_{J_1} \|\tilde{T}^L h_I(x)\| dx$.

To estimate the last integral, we first notice that $\tilde{T}^L h_I(x) = T^L h_I(x)$ when $x \in J_1$ (this follows from the remark at the beginning of the proof, since $\langle W \rangle_L^{1/2} (\Pi_T + \Pi_{(T-1)^*}) \langle W \rangle_L^{-1/2} h_I$ is supported on $I$).
If $\ell(I) \leq \ell(J) \leq 2^p \ell(I)$, then
\[
\left\| \tilde{T}_{I,J}^L \right\| \leq 2|J|^{-1/2} \int_{I} \left\| \tilde{T}^L h_I(x) \right\| dx = 2|J|^{-1/2} \int_{I} \left\| T^L h_I(x) \right\| dx \\
\leq 2|J|^{-1/2} \left( \int_{(3J)^c} \left\| \int_{I} \langle W_{I} \rangle_{L}^{1/2} K(x,y) \langle W \rangle_{L}^{-1/2} h_I(y) dy \right\| dx \right. \\
\left. + \int_{(3J)^c} \left\| \int_{I} \langle W_{I} \rangle_{L}^{1/2} (K(x,y) - K(x,c_I)) \langle W \rangle_{L}^{-1/2} h_I(y) dy \right\| dx \right)
\leq |J|^{-1/2} \left( C_0 \int_{(3J)^c} \int \frac{1}{|x-y|^p} dx dy |I|^{-1/2} + C_\delta \int_{(3J)^c} \frac{\ell(I)^\delta}{\text{dist}(x,I)^{p+\delta}} dx \right)
\leq |J|^{-1/2} \left( C_0 |I|^{1/2} + C_\delta \int_{(l(I),l(J))^{1-\gamma}} \frac{\ell(I)^\delta}{\ell(I)^{1/2} \ell(J)^{1/2}} dx \right)
\leq \left( \frac{|I|}{|J|} \right)^{1/2} \frac{\ell(I)^\delta}{\ell(I)^{1/2} \ell(J)^{1/2}}
\leq \frac{\ell(I)^{1+p}}{\ell(I)^{1/2} \ell(J)^{1/2}} \approx \frac{\ell(I)^{1+p}}{D(I,J)^{p+\delta}}.
\]

where the last inequality is true since $\gamma \leq \frac{1}{2}$.

Case 2: $(I,J) \in \Gamma_{\text{equal}}$

By the remark at the beginning of the proof, we have $\tilde{T}_{I,I} = T_{I,I}$, and thus also $\tilde{T}_{I,J}^L = T_{I,J}^L$.

Note that by minimality of $L$, $L = I$.

To emphasize that the two Haar functions appearing in the definition of the matrix $T_{I,J}^L$ are not the same (even though $I = J$ in this case), we briefly reintroduce the superscripts $\varepsilon, \varepsilon'$.

If $\{I_i\}_{i=1}^{2^p}$ are the dyadic children of $I$, then using the kernel representation and the weak boundedness property we obtain
\[
\left\| \tilde{T}_{I,I}^L \right\| = \left\| T_{I,I}^L \right\| = \left\| \langle T^L h_I^\varepsilon, h_J^\varepsilon' \rangle \right\| \leq \sum_{i,j=1}^{2^p} \left\| \langle h_I^\varepsilon, h_J^\varepsilon' \rangle \right\| \left\langle h_J^\varepsilon', \chi_{I_i} \right\rangle \\
\leq \sum_{i \neq j} |I|^{-1} \left\| \int_{I_i} \int_{I_j} \langle W_{I_i} \rangle_{L}^{1/2} K(x,y) \langle W \rangle_{L}^{-1/2} \chi_{I_i}(y) \chi_{I_i}(x) dy dx \right\| + \sum_{i=1}^{2^p} |I|^{-1} \left\| \langle T^L \chi_{I_i}, \chi_{I_i} \rangle \right\| \\
\leq C_0 |I|^{-1} \int_{I_i} \int \frac{1}{|x-y|^p} dx dy + \sum_{i=1}^{2^p} |I|^{-1} \left\| \langle T^L \chi_{I_i}, \chi_{I_i} \rangle_{L^2(\mathbb{R}^p), L^2(\mathbb{R}^p)} \right\| \\
\leq C_0 + \sum_{i=1}^{2^p} |I|^{-1} C_{WBP}|I_i| = C_0 + 2^p C_{WBP}.
\]

Case 3: $(I,J) \in \Gamma_{\text{out}}$
As before, the remark at the beginning of the proof shows that $\tilde{T}^L_{I,J} = T^L_{I,J}$ if $I$ and $J$ are disjoint (which is obviously the case here).

If $c_l$ is the centre of $I$, the decay property of $K$ and the cancellation of $h_I$ allow us to estimate

\[
\left\|\tilde{T}^L_{I,J}\right\| = \left\| T^L_{I,J} \right\| = \left\| \langle T^L h_I, h_J \rangle \right\| = \left\| \int_I \int_I \langle W_{L}^{1/2}(K(x,y) - K(x, c_I))(W_{L}^{1/2} h_I(y))h_J(x) \rangle dy dx \right\|
\]

\[
\leq \int_I \int_I \left\| W_{L}^{1/2}(K(x,y) - K(x, c_I)) \right\| \left\| h_I(y) \right\| \left\| h_J(x) \right\| dy dx
\]

\[
\leq |I|^{-1/2} |J|^{-1/2} \int_I \int_I \frac{|y - c_I|^\delta}{|x - y|^{p+\delta}} dy dx
\]

\[
\leq |I|^{-1/2} |J|^{-1/2} \left( \frac{\ell(I)}{2} \right)^\delta \frac{1}{\text{dist}(I, J)^{p+\delta}} |I| |J| \leq \frac{\ell(I)^\delta}{\text{dist}(I, J)^{p+\delta}} \ell(I)^{1/2} \ell(J)^{1/2}.
\]

If $\text{dist}(I, J) > \ell(J)$, then $D(I, J) = \ell(I) + \ell(J) + \text{dist}(I, J) < 3\ell(I, J)$, thus

\[
\frac{\ell(I)^\delta}{\text{dist}(I, J)^{p+\delta}} \ell(I)^{1/2} \ell(J)^{1/2} \leq \frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{D(I, J)^{p+\delta}},
\]

since $\ell(I) \leq \ell(J)$.

On the other hand, if $\text{dist}(I, J) \leq \ell(J)$, then $D(I, J) = \ell(I) + \ell(J) + \text{dist}(I, J) < 3\ell(J)$, so

\[
\frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{D(I, J)^{p+\delta}} \leq \frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{(\ell(I) \gamma^{(1-\gamma)} \ell(J)^1)^{p+\delta}} \leq \frac{\ell(I)^{1/2} \ell(J)^{1/2}}{\ell(I)^{p+\delta}} \leq \frac{\ell(I)^{1/2} \ell(J)^{1/2}}{D(I, J)^{p+\delta}},
\]

where in the above inequality we have used that $\gamma(p + \delta) = \frac{\delta}{4}$.

Case 4: $(I, J) \in \Gamma_{\text{near}}$

As in the previous case, the disjointness of $I$ and $J$ implies that $\tilde{T}^L_{I,J} = T^L_{I,J}$. Since $\text{dist}(I, J) \leq \ell(I) \gamma \ell(J)^{1-\gamma} \leq \ell(J)$, it follows that $I \subseteq 5J \setminus J$. Using the kernel representation of the operator $T$, we have

\[
\left\| \tilde{T}^L_{I,J} \right\| = \left\| T^L_{I,J} \right\| = \left\| \langle T^{L} h_I, h_J \rangle \right\| = \left\| \int_I \int_I \langle W^{1/2}_L K(x,y)(W^{1/2}_L h_I(y))h_J(x) \rangle dy dx \right\|
\]

\[
\leq C_0 |I|^{-1/2} |J|^{-1/2} \int_I \int_I \frac{1}{|x - y|^{p}} dy dx \leq C_0 |I|^{-1/2} |J|^{-1/2} \int_I \int_{J \setminus J} \frac{1}{|x - y|^{p}} dy dx
\]

\[
\leq C_0 |J|^{-1/2} |J|^{-1/2} = C_0,
\]

where the last inequality is true, since the goodness of $I$ ensures that $\ell(J) \leq 2^\gamma \ell(I)$. This completes the proof of the first inequality in the statement of the lemma.

To prove the second inequality, we notice that if $n \geq 1$, the minimality of $L$ implies that $I$ and $J$ are disjoint. Let $L_I$ be the child of $L$ that contains $I$. Since $L$ is the smallest common dyadic ancestor of $I$ and $J$, we have $\text{dist}(I, J) \geq \text{dist}(I, \partial L_I)$. Again by minimality and goodness of $I$,

\[
D(I, J) = \ell(I) + \ell(J) + \text{dist}(I, J) \geq \ell(I)^\gamma \ell(L)^{1-\gamma}
\]

and thus

\[
\frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{D(I, J)^{p+\delta}} \leq \frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{(\ell(I)^{\gamma(1-\gamma)} \ell(L)^{1-\gamma})^{(p+\delta)}} \leq \frac{\sqrt{|I|} \sqrt{|J|} \ell(I)^{1/2} \ell(J)^{1/2}}{|L|^{(p+\delta)/4}} \leq \frac{\sqrt{|I|} \sqrt{|J|} 2^{-(m+n)\delta/2}}{|L|^{(p+\delta)/4}}
\]

where we use $\gamma(p + \delta) = \frac{\delta}{4}$.

If $n = 0$, that is $J = L$, we have $D(I, J) \geq \ell(L)$. Then

\[
\left\| \tilde{T}^L_{I,J} \right\| \leq C \frac{\ell(I)^{p+\delta} \ell(J)^{p+\delta}}{\ell(L)^{p+\delta}} = C \frac{2^{-(m+n)\delta}}{|L|^{(p+\delta)/4}} \leq C \frac{\sqrt{|I|} \sqrt{|J|} 2^{-(m+n)\delta}}{|L|^{(p+\delta)/4}}
\]

which concludes the proof of the lemma.

$\square$
This proves that the operators $S_{mn}^\omega$ appearing in (15) are appropriate scalar multiples of W-Haar shifts. Together with Theorem 2.2 and the estimates (11) and (12), this gives the sufficiency direction (1) $\Rightarrow$ (2) and the claimed bounds. We now prove necessity, namely (2) $\Rightarrow$ (1).

First, we show that an $L^2(W)$-bounded W-Calderón-Zygmund operator satisfies the W-weak boundedness property. If $I$ is a cube and $I_1$ is a child of $I$, then

$$\left\| \langle W \rangle_{I_1}^{1/2} T(W)^{-1/2} \chi_{I_1} \chi_{I_1} \right\|$$

$$= \left\| \langle W \rangle_{I_1}^{1/2} TW^{-1/2} W^{1/2} \langle W \rangle_{I_1}^{1/2} \chi_{I_1} \chi_{I_1} \right\|$$

$$\leq \|W^{1/2} \langle \chi_{I_1} \rangle \|_{L^2(R^p)} \|W^{1/2} \langle W \rangle_{I_1}^{1/2} \chi_{I_1} \chi_{I_1} \|_{L^2(R^p)}$$

$$\leq \|T\|_{L^2(W) \to L^2(W)} C[I_1^{1/2} \|W\|_{A_2}^{1/2} \leq C[I_1^{1/2} \|W\|_{A_2}^{1/2} |I|].$$

To show that $T \in BMO_W$ if $T : L^2(W) \to L^2(W)$ is bounded, we first introduce the operator

$$\tilde{P}_I f = \langle f, h_I \rangle h_I + \sum_{J \subseteq I} \langle h_I, h_J \rangle h_J,$$

where $I_J$ is the child of $I$ containing $J$. Here $f$ is a locally integrable $C^2$-valued function, but we will also allow the operator $\tilde{P}_I$ to act on locally integrable $M_d(C)$-valued functions. Note that since $W$ is $A_2$, $W^{1/2} \tilde{P}_I W^{-1/2}$ is bounded on $L^2(R^p)$, with bound independent of $I$. We then have

$$\|W^{1/2} \tilde{P}_I T(W)^{-1/2} h_I\|_{L^2(R^p)} = \|W^{1/2} \tilde{P}_I W^{1/2} TW^{-1/2} W^{1/2} (W)^{-1/2} h_I\|_{L^2(R^p)}$$

$$\leq \|W^{1/2} \tilde{P}_I TW^{-1/2} \|_{L^2(R^p) \to L^2(R^p)} \|W^{1/2} \langle W \rangle_{I_1}^{1/2} h_I\|_{L^2(R^p)}$$

$$\leq C[I_1^{1/2} \|W\|_{A_2}^{1/2} \|T\|_{L^2(W) \to L^2(W)}].$$

On the other hand, we can write

$$W^{1/2} \tilde{P}_I T(W)^{-1/2} h_I = W^{1/2} \langle T(W)^{-1/2} h_I, h_I \rangle h_I + W^{1/2} \sum_{J \subseteq I} \langle h_I, h_J \rangle \langle T(W)^{-1/2} h_I, h_J \rangle h_J.$$

Using the same splitting as in the proof of the first case ($\Gamma_{in}$) of Lemma 3.1, the terms of the previous sum can be expressed as

$$W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle h_I \rangle_{I_1} \langle T^I (\chi_{I_1} (h_I - \langle h_I \rangle_{I_1})), h_J \rangle h_J + W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle (h_I)_{I_1} \rangle^2 \langle T^I 1, h_J \rangle h_J.$$ 

It then follows that

$$W^{1/2} \tilde{P}_I T(W)^{-1/2} h_I = W^{1/2} \langle T^I h_I, h_I \rangle h_I + \sum_{J \subseteq I} W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle h_I \rangle_{I_1} \langle T^I (\chi_{I_1} (h_I - \langle h_I \rangle_{I_1})), h_J \rangle h_J$$

$$+ \sum_{J \subseteq I} W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle T^I 1, h_J \rangle h_J.$$ 

Since the Haar functions form an unconditional basis in $L^2(W)$, we have

$$\left\|W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle T^I h_I, h_I \rangle h_I + \sum_{J \subseteq I} W^{1/2} \langle W \rangle_{I_1}^{-1/2} \langle h_I \rangle_{I_1} \langle T^I (\chi_{I_1} (h_I - \langle h_I \rangle_{I_1})), h_J \rangle h_J \right\|_{L^2(R^p)}^2$$

$$\leq C[I_1^{1/2} \|W\|_{A_2}^{1/2} \|T\|_{L^2(W) \to L^2(W)} + \frac{1}{|I|} \sum_{J \subseteq I} \|\langle T^I (\chi_{I_1} (h_I - \langle h_I \rangle_{I_1})), h_J \rangle h_J \|_{L^2(R^p)}^2$$

$$\lesssim C[I_1^{1/2} \|W\|_{A_2}^{1/2} \|\Gamma_{in} \|_{\Gamma_{in}} \sum_{J \subseteq I} \|\langle T^I (\chi_{I_1} (h_I - \langle h_I \rangle_{I_1})), h_J \rangle h_J \|_{L^2(R^p)}^2.$$

Here we have used the estimates in the $\Gamma_{in}$ and $\Gamma_{equal}$ cases from Lemma 3.1.
Altogether, we obtain that \[ \frac{1}{|I|^{1/2}} \sum_{J \subseteq I} W^{1/2} (W^{-1/2})^I \langle T^1, h_J \rangle h_J \] is bounded in \( L^2(\mathbb{R}^p) \) with a bound independent of \( I \), hence also the \[ \frac{1}{|I|^{1/2}} \sum_{J \subseteq I} W^{1/2} (W^{-1/2})^I \langle T^1, h_J \rangle h_J \] are uniformly bounded in \( L^2(\mathbb{R}^p) \). But this last sum is equal to

\[
\frac{1}{|I|^{1/2}} \sum_{J \subseteq I} W^{1/2} \langle T_{11} \rangle h_J = \frac{1}{|I|^{1/2}} \langle T_{11}, h_J \rangle \chi_I (W^{-1/2}),
\]

and the uniform \( L^2(\mathbb{R}^p) \)-boundedness in \( I \) of these functions is exactly the condition from Definition 2.3. Therefore \( T_1 \in BMO_W \).

\( T^* 1 \in BMO_{W^{-1}} \) follows immediately by the same argument, since \( T^* : L^2(W^{-1}) \to L^2(W^{-1}) \) is bounded by (3). This finishes the proof of Theorem 2.6 up to proving Theorem 2.2. \( \square \)

4. The proof of Theorem 2.2

We now to the proof of the sharp bound for dyadic \( W \)-Haar shifts, Theorem 2.2. Following the approach in \[20\], one can show that it is enough to consider only dyadic \( W \)-Haar shifts on a dyadic system in \( \mathbb{R} \). This reduction is obtained by arranging the dyadic cubes in an appropriate way on the real line (for more details in the matrix-weighted case, see also \[19\]).

Let \( W \) be a \( d \times d \) matrix \( A_2 \) weight on \( \mathbb{R} \). For each \( I \in \mathcal{D} \), choose an orthonormal basis of eigenvectors \( B_I = \{ e_{I,1}, e_{I,2}, \ldots, e_{I,d} \} \) of \( \langle W \rangle_I \) with associated eigenvalues \( \{ \lambda_{I,1}, \lambda_{I,2}, \ldots, \lambda_{I,d} \} \), and let \( P_I^j, 1 \leq i \leq d \), be the corresponding orthogonal projection onto the span of \( e_{I,i} \).

Since \( S \) is a \( W \)-Haar shift operator of complexity \( k \), it has the form

\[
Sf = \sum_{L \in \mathcal{D}} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} A_L^{I,J} \langle f, h_I \rangle h_J,
\]

where \( \| A_L^{I,J} \| := \| (W_L^{1/2} A_L^{I,J} (W_L^{-1/2}) \leq \sqrt{|I|/|J|} = 2^{-(m+n)p/2} \).
Let \( f \in L^2(W) \), \( g \in L^2(W^{-1}) \) and \( 0 \leq t \leq k - 1 \) be fixed. For the slice \( S_t \), we can write

\[
\langle S_t f, g \rangle_{L^2(W), L^2(W^{-1})} = \left\langle \sum_{L \in \mathcal{L}_t} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} A^L_{I,J}(f, h_I)h_J, \sum_{l \in \mathcal{D}} \langle g, h_I \rangle h_I \right\rangle_{L^2(W), L^2(W^{-1})} \\
= \sum_{L \in \mathcal{L}_t} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} \langle A^L_{I,J}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \langle h_J, h_J \rangle_{L^2(\mathbb{R}), L^2(\mathbb{R})} \\
= \sum_{L \in \mathcal{L}_t} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} \langle A^L_{I,J}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \\
= \sum_{L \in \mathcal{L}_t} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} \langle A^L_{I,J}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \\
= \sum_{d, L \in \mathcal{L}_t} \sum_{i=1}^d \sum_{j=1}^d \langle \tilde{A}^L_{i,j}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \\
= \sum_{d, L \in \mathcal{L}_t} \sum_{i=1}^d \sum_{j=1}^d \langle \tilde{A}^L_{i,j}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \\
\leq C \sum_{d, L \in \mathcal{L}_t} \sum_{i=1}^d \sum_{j=1}^d \langle \tilde{A}^L_{i,j}(f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d} \\
= C \sum_{d, L \in \mathcal{L}_t} \sum_{i=1}^d \sum_{j=1}^d \langle \langle \lambda^{-1/2}_{L,i} e^j_L \otimes \lambda^{1/2}_{L,i} e^j_L \rangle (f, h_I), \langle g, h_I \rangle \rangle_{\mathcal{C}_d}.
\]

Since \( \langle W \rangle_{L}^{1/2} (\lambda^{-1/2}_{L,j} e^j_L \otimes \lambda^{1/2}_{L,i} e^j_L) \langle W \rangle_{L}^{-1/2} = e^j_L \otimes e^i_L \), we have

\[
\left\| \left\{ \lambda^{-1/2}_{L,j} e^j_L \otimes \lambda^{1/2}_{L,i} e^j_L \right\} \right\|_{L \in \mathcal{D}, \infty, W} = \sup_{L \in \mathcal{D}} \langle W \rangle_{L}^{1/2} (\lambda^{-1/2}_{L,j} e^j_L \otimes \lambda^{1/2}_{L,i} e^j_L) \langle W \rangle_{L}^{-1/2} = \sup_{L \in \mathcal{D}} \| e^j_L \otimes e^i_L \| = 1.
\]

It follows that
\[ \langle S_j f, g \rangle_{L^2(W), L^2(W^{-1})} \]

\[ \leq C \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \sum_{L \in \mathcal{L}_j} \sum_{L \in \mathcal{L}_j} \sum_{I \in \mathcal{D}_m(L)} \sum_{J \in \mathcal{D}_n(L)} \left| \langle \sigma_L(f, h_I), (g, h_J) \rangle_{\mathcal{C}^d} \right| \]

\[ \leq C d^2 \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \sum_{L \in \mathcal{L}_j} \sum_{L \in \mathcal{L}_j} \left| \frac{|I|^{1/2}}{2^{k-n}} \right| \left( \sum_{P \in \mathcal{D}_m(L)} \sigma_L(p) - \langle f \rangle_L \right) + \sum_{P \in \mathcal{D}_n(L)} \sigma_L(p) - \langle f \rangle_L \right)

\[ + \sum_{Q \in \mathcal{D}_n(L)} (\langle g \rangle_Q - \langle g \rangle_L) + \sum_{Q \in \mathcal{D}_n(L)} (\langle g \rangle_Q - \langle g \rangle_L) \right)_{\mathcal{C}^d} \cdot \]

We therefore have

\[ \left| \langle S_j f, g \rangle_{L^2(W), L^2(W^{-1})} \right| \]

\[ \leq C \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \sum_{L \in \mathcal{L}_j} \sum_{P, Q \in \mathcal{D}_m(L)} \left| \langle \sigma_L \left( \frac{p - f}{2^k} \right), \frac{(g_Q - g_L)}{2^k} \right)_{\mathcal{C}^d} \right| \]

Using the definition of the martingale transform operator \( T_\sigma \), we can write

\[ \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \sum_{L \in \mathcal{D}} \left| \langle \sigma_L (f_{I^+} - f_{I^-}), (g_{I^+} - g_{I^-}) \rangle_{\mathcal{C}^d} \right| \cdot |I| \]

\[ \leq 4 \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \left| \langle \sigma_L (f - \langle f \rangle_L), (g - \langle g \rangle_L) \rangle_{\mathcal{C}^d} \right| \]

\[ = 4 \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \left| \langle T_\sigma f, g \rangle_{L^2(W), L^2(W^{-1})} \right| \]

\[ \leq 4 \sup_{\sigma = \{\sigma_L\} \in \mathcal{D}} \left| \langle T_\sigma f, g \rangle_{L^2(W), L^2(W^{-1})} \right| \]

\[ \leq 4N(\|W\|_{A^2_d}) \|f\|_{L^2(W)} \|g\|_{L^2(W^{-1})} \]

The left hand side of this chain of inequalities is what motivates the following definition of the Bellman function associated to our problem. Notice that this Bellman function is exactly the same as the one that appears in [15].

Let \( X > 1 \), fix a dyadic interval \( I_0 \) and for \( f, g \in \mathbb{C}^d, f \in \mathbb{R}, U \in \mathcal{M}_d(\mathbb{C}), g \in \mathbb{C}^d, G \in \mathbb{R}, V \in \mathcal{M}_d(\mathbb{C}) \) satisfying

\[ \|U\|_{\mathcal{M}_d(\mathbb{C})} \|V\|_{\mathcal{M}_d(\mathbb{C})} \leq \mathbf{1}, \]

define the function \( B_X = B_X^{I_0} : \mathbb{C}^d \times \mathbb{R} \times \mathcal{M}_d(\mathbb{C}) \times \mathbb{C}^d \times \mathbb{R} \times \mathcal{M}_d(\mathbb{C}) \) by

\[ B_X(f, F, U, g, G, V) := \|I_0\|^{-1} \sup_{L \subset I_0} \left| \langle \sigma_L (f_{I^+} - f_{I^-}), (g_{I^+} - g_{I^-}) \rangle_{\mathcal{C}^d} \right| \cdot |I| \]

where the supremum is taken over all functions \( f, g : \mathbb{R} \to \mathbb{C}^d \), all matrix \( A^2_d \) weights \( W \) on \( I_0 \), and all sequences of \( d \times d \) matrices \( \sigma = \{\sigma_L\} \in \mathcal{D} \) such that

\[ \langle f \rangle_{I_0} = f \in \mathbb{C}^d, \quad \langle g \rangle_{I_0} = g \in \mathbb{C}^d, \quad \langle W^{-1/2}g \rangle_{I_0} = G \in \mathbb{R}, \]

\[ \sup_{L \subset I_0} \|W_{I_0}^{-1/2}(W^{-1})_{I_0} \|_{\mathcal{C}^d} \leq X, \quad \langle W \rangle_{I_0} = U, \quad \langle W^{-1} \rangle_{I_0} = V, \quad \|\sigma\|_{\mathcal{M}_d(\mathbb{C})} \leq 1. \]

The Bellman function \( B_X \) has the following properties:
Here, the supremum is taken over all dyadic intervals $\sigma$. Let $\sigma$ be a dyadic interval such that $\sigma = (f_-, U_-, g_-, V_-)$ in $\mathfrak{D}_X$ such that $A = (A_+ + A_-)/2$. For all such tuples, we have the following concavity condition:

$$B_X(A) \leq \frac{B_X(A_+) + B_X(A_-)}{2} + \sup_{\|\tau\|_U \leq 1} |\langle \tau(f_- - f_+), g_- - g_+ \rangle|.$$ 

Here, the supremum is taken over all $d \times d$ matrices $\tau$ with $\|\tau\|_U := \|U^1/2 \tau U^{-1/2}\| \leq 1$.

More details about these properties can be found in [18].

We can now state the main tool for the proof of Theorem 2.2.

**Lemma 4.1.** Let $X > 1$ and $B_X$ be a function satisfying properties (i)-(iii) from above. Fix $k \geq 1$ and a dyadic interval $I_0$. For all $I \in \mathfrak{D}_n(I_0)$, $0 \leq n \leq k$, let the points $A_I = (f_I, F_I, U_I, g_I, V_I) \in \mathfrak{D}_X = \text{Dom} B_X$ be given. Assume that the points $A_I$ satisfy the dyadic martingale dynamics, i.e. $A = (A_+ + A_-)/2$, where $I^+$ and $I^-$ are the children of $I$. Let $\sigma_{I_0}$ be a $d \times d$ matrix such that $\|U_{I_0}^{1/2} \sigma_{I_0} U_{I_0}^{-1/2}\| \leq 1$. For $K, L \in \mathfrak{D}_k(I_0)$, we define the coefficients $\lambda_{KL}$ by

$$\lambda_{KL} := \left\langle \sigma_{I_0} \left( \frac{f_K - f_{I_0}}{2^k}, \left( \frac{g_K - g_{I_0}}{2^k} \right) \right) \right\rangle_{C^d}.$$ 

Then

$$\sum_{K, L \in \mathfrak{D}_k(I_0)} |\lambda_{KL}| \leq c \cdot d \left( B_X(A_{I_0}) - 2^{-k} \sum_{I \in \mathfrak{D}_k(I_0)} B_X(A_I) \right),$$ 

where $c$ is a positive absolute constant and $X' = \frac{100}{9} X$.

The proof of this lemma follows exactly as in [18], the only difference being the use of the matrix $J_{I_0}$ instead of the projection $P_{I_0}$.

We are now ready to finish the proof of Theorem 2.2.

Recall that for all slices $S_I$ of $S$ we have

$$\left| \langle S_I f, g \rangle_{L^2(W), L^2(W^{-1})} \right| \leq C d^2 \sup_{\|\sigma\|_{C^d} \leq 1} \sum_{L \in \mathfrak{L}_t} |L| \sum_{P, Q \in \mathfrak{D}_k(L)} \left| \left( \frac{\langle f, g \rangle - \langle f, L \rangle_{C^d}}{2^k}, \frac{\langle g, L \rangle_{C^d}}{2^k} \right) \right|_{C^d}.$$ 

Let $X := |W| A_0^2$; fix $0 \leq t \leq k - 1$ and for all $I \in \mathfrak{L}_t$ define

$$A_I := \left( \langle f, I \rangle, \|W^{1/2} f\|_{C^d}^2 I, \langle g, I \rangle, \|W^{1/2} g\|_{C^d}^2 I, \langle W^{-1/2} f, W^{-1/2} g \rangle_{C^d} I, \langle W^{-1/2} |f - f|, W^{-1/2} |g - g| \rangle_{C^d} I \right).$$ 

Notice that all these points are in Dom $B_X = \mathfrak{D}_X$. Lemma 4.1 says that

$$\left| \sum_{P, Q \in \mathfrak{D}_k(L)} \left| \left( \frac{\langle f, g \rangle - \langle f, L \rangle_{C^d}}{2^k}, \frac{\langle g, L \rangle_{C^d}}{2^k} \right) \right|_{C^d} \right| \leq c \cdot d \left( \sum_{I \in \mathfrak{D}_k(L)} |B_X(A_I)| - \sum_{I \in \mathfrak{D}_k(L)} |B_X(A_I)| \right).$$
for all \( L \in \mathcal{L} \) and all \( d \times d \) matrices \( \sigma_L \) such that \( \| (W)^{1/2} \sigma_L (W)^{-1/2} \| \leq 1 \). We write this estimate for each \( I \in \mathcal{D}_k (L) \) and then iterate the procedure \( \ell \) times to obtain

\[
\sum_{I \in \mathcal{L}_k} |I| \sum_{P,Q \in \mathcal{D}_k (I)} \left| \sigma_L \left( \frac{(f)_P - (f)_I}{2^k} \right), \frac{(g)_Q - (g)_I}{2^k} \right| \leq c \cdot d \left( |I| |B^{X'}(A_L) - \sum_{I \in \mathcal{D}_k (L)} |I| B^{X'}(A_I) | \right)
\]

\[
\leq c \cdot d N(X') \| W^{1/2} f \|_{C^d}^{1/2} \| W^{-1/2} g \|_{C^d}^{1/2}
\]

\[
\leq c \cdot d N(X') \| f_{\chi_L} \|_{L^2(W)} \| g_{\chi_L} \|_{L^2(W^{-1})},
\]

where the second inequality follows from property (ii) of the Bellman function.

Letting \( \ell \to \infty \), we have

\[
\sum_{I \in \mathcal{L}_k} |I| \sum_{P,Q \in \mathcal{D}_k (I)} \left| \sigma_L \left( \frac{(f)_P - (f)_I}{2^k} \right), \frac{(g)_Q - (g)_I}{2^k} \right| \leq c \cdot d N(X') \| f_{\chi_L} \|_{L^2(W)} \| g_{\chi_L} \|_{L^2(W^{-1})}.
\]

We now cover the real line with intervals \( L \in \mathcal{L}_k \) of length \( 2^M \) and apply the last inequality to each \( L \) to obtain that

\[
\sup_{\sigma = \{ \sigma_I \}_{I \in \mathcal{D}_k}, \| \sigma \|_{\infty, W} \leq 1} \sum_{I \in \mathcal{L}_k} |I| \sum_{P,Q \in \mathcal{D}_k (I)} \left| \sigma_L \left( \frac{(f)_P - (f)_I}{2^k} \right), \frac{(g)_Q - (g)_I}{2^k} \right| \leq c \cdot d N(X) \| f \|_{L^2(W)} \| g \|_{L^2(W^{-1})}.
\]

When \( M \to \infty \), we get that the norm of \( S_I \) is bounded by \( c \cdot d^3 N(X) \). Since \( S \) was decomposed into \( k \) slices, it follows that the operator norm of \( S \) is bounded by \( c \cdot k d^3 N([W]_{A_2^d}) \), and therefore the proof of Theorem 2.2 (and therefore also of the sufficiency in Theorem 2.6) is complete. \( \square \)

REFERENCES

[1] K. Bickel, S. Petermichl, and B. Wick, Bounds for the Hilbert transform with matrix \( A_2 \) weights, J. Funct. Anal. 270 (2016), no. 5, 1719-1743.
[2] R. R. Coifman and C. Fefferman, Weighted norm inequalities for maximal functions and singular integrals, Studia Math. 51 (1974), 241-250.
[3] M. Goldberg, Matrix \( A_p \) weights via maximal functions, Pacific J. Math. 211 (2003), no. 2, 201-220.
[4] T. S. Hämmänen and T. P. Hytönen, Operator-valued dyadic shifts and the \( T(1) \) theorem, Monatsh. Math. 180 (2016), no. 2, 213 – 253.
[5] S. Hukovic, S. Treil, and A. Volberg.
[6] R. A. Hunt, B. Muckenhoupt, and R. L. Wheeden, Weighted norm inequalities for the conjugate function and Hilbert transform, Trans. Amer. Math. Soc. 176 (1973), 227 – 251.
[7] T. P. Hytönen, The sharp weighted bound for general Calderón-Zygmund operators, Ann. of Math. (2) 175 (2012), no. 3, 1473-1506.
[8] . . , Representation of singular integrals by dyadic operators, and the \( A_2 \) theorem, Lecture notes of an intensive course at Universidad de Sevilla, Summer 2011 Preprint, arXiv:1108.5119 (2011).
[9] T. P. Hytönen and C. Pérez, Sharp weighted bounds involving \( A_\infty \), Analysis and PDE 4 (2013), 777 – 818.
[10] T. Hytönen, S. Petermichl, and A. Volberg, The sharp square function estimate with matrix weight Preprint, arXiv:1702.04569 (2017).
[11] J. Isralowitz, A matrix weighted \( T_1 \) theorem for matrix kernelised CZOs and a matrix weighted John-Nirenberg theorem Preprint, arXiv:1508.02474 (2015).
[12] J. Isralowitz, H.-K. Kwon, and S. Pott, Matrix-weighted norm inequalities for commutators and paraproducts with matrix symbols, to appear in J. London Math. Soc. Preprint, arXiv:1507.04032 (2015).
[13] 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, St. Petersburg Math. J. 8 (1997), no. 5, 721–824.
[14] F. Nazarov, S. Treil, and A. Volberg, The \( T_b \)-theorem on non-homogeneous spaces, Acta Mathematica 190 (2003), no. 2, 151 - 239.
[15] F. Nazarov, S. Petermichl, S. Treil, and A. Volberg, *Convex body domination and weighted estimates with matrix weights* arXiv:1701.01907 [math.CA] (2017).
[16] S. Petermichl, *The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic*, Amer. J. Math. 129 (2007), no. 5, 1355-1375.
[17] S. Petermichl and S. Pott, *An estimate for weighted Hilbert transform via square functions*, Trans, Amer. Math. Soc. 354 (2002), no. 4, 1699-1703.
[18] S. Pott and A. Stoica, *Bounds for Calderón-Zygmund operators with matrix A_2 weights*, accepted by Bull. Sci. Math. arXiv:1508.06408 [math.CA] (2015).
[19] S. Petermichl and A. Volberg, *Heating of the Ahlfors-Beurling operator: weakly quasiregular maps on the plane are quasiregular*, Duke Math. J. 112 (2002), no. 2, 281-305.
[20] S. Treil, *Sharp A_2 estimates of Haar shifts via Bellman function Preprint, arXiv:1105.2252* (2011).
[21] S. Treil and A. Volberg, *Wavelets and the Angle between Past and Future*, J. Funct. Anal. 143 (1997), no. 2, 269-308.
[22] A. Volberg, *Matrix A_p weights via S-functions*, J. Amer. Math. Soc. 10 (1997), no. 2, 445-466.
[23] M. Wilson, *Weighted Littlewood-Paley theory and exponential-square integrability*, Springer Lecture Notes in Mathematics, vol. 1924, Berlin, 2008.
[24] J. Wittwer, *A sharp estimate on the norm of the martingale transform*, Math. Res. Lett. 7 (2000), no. 1, 1-12.

Centre for Mathematical Sciences, University of Lund, P.O. Box 118, SE-221 00 Lund, Sweden
E-mail address: sandra@maths.lth.se

Centre for Mathematical Sciences, University of Lund, P.O. Box 118, SE-221 00 Lund, Sweden
E-mail address: andrei@maths.lth.se