A matrix weighted bilinear Carleson lemma and maximal function

Reguera, Maria Carmen; Pott, Sandra; Petermichl, Stefanie

Document Version
Peer reviewed version

A MATRIX WEIGHTED BILINEAR CARLESON LEMMA AND MAXIMAL FUNCTION

STEFANIE PETERMICHL, SANDRA POTT, AND MARIA CARMEN REGUERA

ABSTRACT. We prove a bilinear Carleson embedding theorem with matrix weight and scalar measure. In the scalar case, this becomes exactly the well known weighted bilinear Carleson embedding theorem. Although only allowing scalar Carleson measures, it is to date the only extension to the bilinear setting of the recent Carleson embedding theorem by Culiuc and Treil that features a matrix Carleson measure and a matrix weight.

It is well known that a Carleson embedding theorem implies a Doob's maximal inequality and this holds true in the matrix weighted setting with an appropriately defined maximal operator. It is also known that a dimensional growth must occur in the Carleson embedding theorem with matrix Carleson measure, even with trivial weight. We give a definition of a maximal type function whose norm in the matrix weighted setting does not grow with dimension.

1. INTRODUCTION

The Carleson embedding theorem first appeared in L. Carleson's solution of the free interpolation problem [3] and later was used in his celebrated proof of the Corona Theorem [4]. It is a classical theorem in harmonic analysis with many applications to PDE. It states that a Carleson measure μ gives rise to a bounded embedding $L^2(\mathbb{T}) \to L^2(\mathbb{D}, \mu)$. In other words, it states that L^2 boundedness holds if and only if it holds when testing on certain elementary functions.

In this note we are concerned with matrix analogues. An unweighted Carleson lemma with matrix Carleson measure holds trivially, derived from the scalar case. However, the dimensional growth was under investigation in the 90s and it was observed that at least a logarithmic growth with dimension, i.e. the size of the arising matrices, must occur. The sharp dimensional estimate was found in work of Nazarov, Pisier, Treil and Volberg [11] by means of a clever trick with Bellman functions, using elementary complex analysis (Zweikonstantensatz).

Recently Culiuc and Treil [5] added an observation to this idea to obtain the matrix weighted version for an embedding theorem with matrix Carleson measures, without any restrictions on the weight other than it being a matrix weight. Previous versions assumed the matrix A_2 condition on the weight, a strong assumption, which did not appear in the scalar case.

The result of Culiuc and Treil implies a Doob's inequality in the matrix weighted setting. Since even the unweighted Carleson embedding theorem with matrix measure induces a dimensional growth with the size of the matrices, this growth is inherited by the thus obtained Doob's inequality for the maximal operator defined in Definition 2.4. We are interested in a version of a maximal type operator that does not have growth with dimension, see Definition 2.6 and Theorem 2.7. This

operator has certain nice properties, but it is not a classical maximal operator. The estimate we provide is via Bellman function.

A weighted bilinear version of the Carleson embedding lemma, taking two functions as arguments, was an important tool in the early days of sharp weighted theory or two–weighted theory of singular operators in the scalar case. The bilinear Carleson embedding theorem (BET) states that a Carleson measure gives rise to a bilinear estimate.

Motivated by the so–called matrix A_2 conjecture, which asks for the exact growth of the matrix weighted norm of the Hilbert transform acting on vector functions in terms of the matrix A_2-constant, we aim for a bilinear version of the matrix weighted Carleson lemma. Boundedness of matrix weighted Hilbert transform was first studied by Treil and Volberg in [20]. We do not succeed entirely, in the sense that we have to assume that our Carleson measure has scalar entries, see Theorem 2.1. The restriction to scalar Carleson sequence is serious, but it is not very surprising, given the array of difficulties encountered in the task of finding various sharp weighted estimates in the matrix case. As of today, the only optimal estimate for a singular operator is for the square function with matrix weight by Hytönen, Petermichl and Volberg [7] and Treil [19], and a matrix-weighted maximal function [8]. The estimate for the Hilbert transform is still open, with the best to date estimate missing the sharp conjecture by a half power of the characteristic, see Nazarov, Petermichl, Treil and Volberg [10]. Despite the many advances made in the scalar weighted theory, matrix weights do not seem to respond well to most of the tools known to us today.

The first proof of a version of the scalar bilinear Carleson lemma is found in [17] and was rather complicated, using tools and ideas from [13], featuring a Bellman function and three conditions on the measure sequence rather than one. It was understood for some time by the experts that two of the arising conditions were actually redundant. Two of the conditions stated in [17] are implied by the third, as can be seen by a simple Bellman function argument. Indeed, we prove that this implication remains correct in the case of matrix weights and scalar Carleson measure, see Theorem 2.3. Despite this, we believe that the original proof of the bilinear Carleson lemma still presents obstacles in the matrix case, even with scalar coefficients, such as considered here. The argument we apply in this note has the flavour of level sets found in [18].

We hope that the result on BET can be improved - we give some indications of possible questions at the end of this note. Either improvement could be used towards improving the bound on the Hilbert transform with matrix weight. Even though Carleson type theorems with matrix weight and scalar measure do appear to be useful in getting estimates for singular operators, for example in [20], the constants increase by at least a logarithmic factor. In addition, we do not know if our version of BET can be used directly for such estimates. It seems to us at least some improvement would be required, see the end of this note.

We also point out open questions arising on the dimensional dependence at the end of this note.

2. Notation and results

Let us say $Q_0 = [0,1]$ is endowed with a dyadic filtration generated by dyadic intervals $\mathcal{D} = \mathcal{D}(Q_0)$. We index the time by non–negative integers $k \geq 0$ so that

$\mathcal{F}_0$ has the atom $\mathcal{D}_0 = \{Q_0\}$ of size $|Q_0| = 1^0$ and using atoms $\mathcal{D}_k = \{Q \in \mathcal{D}(Q_0) : |Q| = 2^{-k}\}$ we define the generated σ–algebra at time k: $\mathcal{F}_k = \sigma(\mathcal{D}_k)$.

We call a $d \times d$ matrix–valued function W a weight if W is entry–wise locally integrable and if $W(x)$ is positive semidefinite almost everywhere. One defines $L^2(\mathbb{C}^d; W)$ to be the set of vector functions with

$$\|f\|^2_{L^2(\mathbb{C}^d;W)} = \int_{Q_0} \|W^{1/2}(x)f(x)\|^2_{\mathbb{C}^d}\mathrm{d}x = \int_{Q_0} \langle W(x)f(x), f(x)\rangle_{\mathbb{C}^d}\mathrm{d}x < \infty.$$

Let us denote by $\langle \cdot \rangle_Q$ the average of a scalar, vector or matrix function over Q and let us denote by $\cdot(Q)$ the integral over Q.

The precise statement of the bilinear matrix weighted Carleson lemma we prove is the following.

Theorem 2.1 (BET). *Let (α_Q) be a sequence of non-negative scalars. Then for vector functions f, g supported in Q_0,*

$$\frac{1}{|K|}\sum_{Q \in \mathcal{D}(K)} \alpha_Q \le A \quad \forall K \in \mathcal{D}(Q_0)$$

$$\implies \exists B = B(A,d): \sum_{Q \in \mathcal{D}(Q_0)} \alpha_Q |\langle \langle W \rangle_Q^{-1} \langle W^{1/2}f \rangle_Q, \langle W^{-1} \rangle_Q^{-1} \langle W^{-1/2}g \rangle_Q \rangle_{\mathbb{C}^d}|$$

$$\le B\|f\|_{L^2(\mathbb{C}^d)}\|g\|_{L^2(\mathbb{C}^d)} \quad \forall f, g.$$

The best estimate for the arising constant we can give to date is of order d^2.

We recall the recent result [5] where the Carleson sequence is allowed to be matrix valued, but only one function f is being used:

Theorem 2.2 (CET). *(Culiuc–Treil) Let W be an invertible selfadjoint matrix weight of size d and (A_Q) a sequence of positive semidefinite matrices of size d. Then for a vector function f supported in Q_0,*

$$\frac{1}{|K|}\sum_{Q \in \mathcal{D}(K)} \langle W \rangle_Q A_Q \langle W \rangle_Q \le A \langle W \rangle_K \quad \forall K \in \mathcal{D}(Q_0)$$

$$\implies \exists B = B(A,d): \sum_{Q \in \mathcal{D}(Q_0)} \|A_Q^{1/2}\langle W^{1/2}f \rangle_Q\|^2_{\mathbb{C}^d}$$

$$\le B\|f\|^2_{L^2(\mathbb{C}^d)} \quad \forall f.$$

It is known that the constant is at least of order $\log^2(d)$, since this is the confirmed blow-up when $W = \mathrm{Id}$, see [11] and at most of order d^2, as proved in [5].

Related to Theorem 2.1 is the following implication, which we prove via Bellman functions.

Theorem 2.3. *Let (α_Q) be a non–negative sequence and W be a matrix weight such that W and W^{-1} are summable on all dyadic intervals. Then*

$$\frac{1}{|K|}\sum_{Q \in \mathcal{D}(K)} \alpha_Q \lessdot 1 \quad \forall K \in \mathcal{D}(Q_0)$$

$$\implies \frac{1}{|K|}\sum_{Q \in \mathcal{D}(K)} \alpha_Q \langle W^{-1} \rangle_Q^{-1} \lessdot 4\langle W \rangle_K \quad \forall K \in \mathcal{D}(Q_0),$$

where the second inequality is in the sense of operators.

Its motivation and historic meaning are discussed in Section 4.

Here is the usual definition of the maximal function with matrix measure, with the supremum over the dyadic intervals:

Definition 2.4. Let
$$M^W f(x) = \sup_{Q \in \mathcal{D}(Q_0), x \in Q} \|W^{1/2}(x)\langle W \rangle_Q^{-1} \langle W^{1/2} f \rangle_Q\|_{\mathbb{C}^d}.$$

Definition 2.4 is one way of extending the martingale maximal function to the matrix weighted setting, see Section 3 for the motivation. Indeed, using Theorem 2.2 one can show

Theorem 2.5. *There exists $C(d)$ so that for all matrix weights there holds*
$$\|M^W\|_{L^2(\mathbb{C}^d) \to L^2(\mathbb{C})} \le C(d).$$

Here, $C(d)$ at most of order d. In turn, this estimate is used to obtain Theorem 2.1. The implied constants thus obtained for the estimate in Theorem 2.5 are no better than the constants obtained in Theorem 2.2. Consider the following alternative below. See Section 3 for the motivation of this object.

Definition 2.6. For time $k \ge 0$ and $x \in Q \in \mathcal{D}_k$, let
$$M_k^W f(x) = \|\langle W \rangle_Q^{1/2} \langle W \rangle_{K_k(Q)}^{-1} \langle W^{1/2} f \rangle_{K_k(Q)}\|_{\mathbb{C}^d}$$

with $K_k : \mathcal{D}_k \to \cup_{i=0}^k \mathcal{D}_i; Q \mapsto K_k(Q)$ defined inductively: let $K_0(Q_0) = Q_0$ and for $k > 0$ let

$$K_k(Q) = Q \text{ if } \|\langle W \rangle_Q^{-1/2} \langle W^{1/2} f \rangle_Q\|_{\mathbb{C}^d} > \|\langle W \rangle_Q^{1/2} \langle W \rangle_{K_{k-1}(\hat{Q})}^{-1} \langle W^{1/2} f \rangle_{K_{k-1}(\hat{Q})}\|_{\mathbb{C}^d}$$

and $K_k(Q) = K_{k-1}(\hat{Q})$ else. Here, $\hat{Q}$ denotes the parent of Q.

Here is our result, which carries the same norm estimate as seen in Doob's inequality and in particular does not depend on dimension.

Theorem 2.7. *For all $k \ge 0$ there holds*
$$\|M_k^W\|_{L^2(\mathbb{C}^d) \to L^2(\mathbb{C})} \le 2.$$

3. MAXIMAL FUNCTION

In this section we prove that the matrix weighted maximal function is bounded without additional assumptions on the matrix weight. This line of argument is well known to the experts in the area and rests on the strength of the matrix weighted Carleson lemma by Culiuc and Treil [5]. This estimate inherits the dimensional growth that occurs in the embedding theorem. We then proceed with the definition of a maximal function via an adapted sequence. We show that for this definition, the dimensional growth does not occur. The arguments rests on the Bellman function argument in [12].

3.1. **Usual Maximal function.** Definition 2.4 is motivated as follows. Since maximal norm quantities have scalar values, the matrix weight is included in the definition. Let the expression

$$(M_w f)(x) = \sup_{Q:x \in Q} |\langle f \rangle_{Q,w}| = \sup_{Q:x \in Q} |w(Q)^{-1} \int_Q fw| = \sup_{Q:x \in Q} |\langle w \rangle_Q^{-1} \langle fw \rangle_Q|$$

denote the classical dyadic maximal function with measure w for non–negative weight w and scalar valued f. Observe that since w is non–negative, the suprema in $M_w f(x)$ and $M^w f(x)$ are attained at the same interval $I \ni x$ and

$$\|M^w f(x)\|_{L^2} = \|M_w f(x)\|_{L^2(w)}.$$

As a consequence of Doob's theorem, we have the estimate

$$\|M^w f(x)\|_{L^2} = \|M_w f\|_{L^2(w)} \leqslant 2\|f\|_{L^2(w)} = 2\|f w^{1/2}\|_{L^2}.$$

We turn to the proof of Theorem 2.5 via linearization, a well known argument.

Proof. For any $Q \in \mathcal{D}(Q_0)$ define E_Q to be the subset of $x \in Q$ so that Q is the maximal cube with respect to the collection that attains the maximum in the definition of M^W. So

$$M^W f(x) = \sum_{Q \in \mathcal{D}(Q_0)} \|W^{1/2}(x)\langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q\|_{\mathbb{C}^d} \chi_{E_Q}(x).$$

Now by the disjointness of the E_Q

$$
\begin{aligned}
\|M^W f\|_{L^2(\mathbb{C})}^2 &= \int_{Q_0} \sum_{Q \in \mathcal{D}(Q_0)} \|W^{1/2}(x)\langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q\|_{\mathbb{C}^d}^2 \chi_{E_Q}(x) \mathrm{d}x \\
&= \sum_{Q \in \mathcal{D}(Q_0)} \int_{E_Q} \langle W(x)\langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q, \langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q\rangle_{\mathbb{C}^d} \mathrm{d}x \\
&= \sum_{Q \in \mathcal{D}(Q_0)} \langle\langle W\rangle_{E_Q} |E_Q|\langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q, \langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q\rangle_{\mathbb{C}^d}.
\end{aligned}
$$

To proceed with the estimate, we first observe that

$$\frac{1}{|K|} \sum_{Q \in \mathcal{D}(K)} \langle W\rangle_Q \langle W\rangle_Q^{-1}\langle W\rangle_{E_Q} |E_Q|\langle W\rangle_Q^{-1}\langle W\rangle_Q \leqslant \langle W\rangle_K$$

in the sense of operators, which is satisfied thanks to the disjointness of the E_Q. Now we use Theorem 2.2 with

$$A_Q = \langle W\rangle_Q^{-1}\langle W\rangle_{E_Q} |E_Q|\langle W\rangle_Q^{-1}.$$

The conclusion of this theorem gives us the desired estimate:

$$
\begin{aligned}
\|M^W f\|_{L^2(\mathbb{C})}^2 &= \sum_{Q \in \mathcal{D}(Q_0)} \langle\langle W\rangle_Q^{-1}\langle W\rangle_{E_Q} |E_Q|\langle W\rangle_Q^{-1}\langle W^{1/2} f\rangle_Q, \langle W^{1/2} f\rangle_Q\rangle_{\mathbb{C}^d} \\
&= \sum_{Q \in \mathcal{D}(Q_0)} \langle A_Q \langle W^{1/2} f\rangle_Q, \langle W^{1/2} f\rangle_Q\rangle_{\mathbb{C}^d} \\
&\leq B(1, d)\|f\|_{L^2(\mathbb{C}^d)}^2.
\end{aligned}
$$

$\square$

Observe that the implied constant is the same as in the conclusion of Theorem 2.2.

3.2. Maximal function with poor memory.
Let us now motivate and explain the different maximal function from Definition 2.6.

Notice that the classical scalar maximal function is an adapted process if the supremum is restricted to early times: adapted means that the expression

$$(M_w f)_k(x) = \sup_{Q:x\in Q, |Q|\geqslant 2^{-k}} |\langle f\rangle_{Q,w}|$$

has the property $(M_w f)_k \in \mathcal{F}_k$. This notation means the function $(M_w f)_k$ is measurable in the filtration $\mathcal{F}_k$ for all k, which is so since it is piecewise constant on atoms $\mathcal{D}_k$. If we denote by $K_k(x)$ the interval containing x so that

$$(M_w f)_k(x) = |\langle f\rangle_{K_k(x),w}|,$$

then $(M_w f)_{k+1}(x) = \max\{|\langle f\rangle_{K_k(x),w}|, |\langle f\rangle_{Q,w}|\}$ where $|Q| = 2^{-k-1}$ and $x \in Q$. It therefore suffices to 'memorize' just one past interval, so 'poor memory' is enough in the scalar case. These properties motivated our definition for the matrix weighted case.

The definition above for the matrix weighted maximal operator M^W is not adapted when truncating because of the pointwise multiplication by $W(x)^{1/2}$. We now review the inductive definition of the adapted M_k^W for $k \geqslant 0$ with 'poor memory' in this setting. In this case 'poor memory' may be 'bad memory'.

For time $k \geq 0$ and $x \in Q \in \mathcal{D}_k$, the process we defined is

$$M_k^W f(x) = \|\langle W\rangle_Q^{1/2} \langle W\rangle_{K_k(Q)}^{-1} \langle W^{1/2}f\rangle_{K_k(Q)}\|_{\mathbb{C}^d}$$

for dyadic intervals $K_k(Q)$ defined inductively with $K_0(Q_0) = Q_0$ so that

$$M_0^W f(x) = \|\langle W\rangle_{Q_0}^{-1/2}\langle W^{1/2}f\rangle_{Q_0}\|_{\mathbb{C}^d} = \|\langle W\rangle_{Q_0}^{1/2}\langle W\rangle_{Q_0}^{-1}\langle W^{1/2}f\rangle_{Q_0}\|_{\mathbb{C}^d}.$$

Assuming $k \geqslant 1$ and $K_{k-1}(Q)$ chosen for all dyadic intervals $Q \in \mathcal{D}_{k-1}$, choose $K_k(Q) \in \{K_{k-1}(\hat{Q}), Q\}$ for $Q \in \mathcal{D}_k$ by

$$K_k(Q) = Q \text{ if } \|\langle W\rangle_Q^{-1/2}\langle W^{1/2}f\rangle_Q\|_{\mathbb{C}^d} > \|\langle W\rangle_Q^{1/2}\langle W\rangle_{K_{k-1}(\hat{Q})}^{-1}\langle W^{1/2}f\rangle_{K_{k-1}(\hat{Q})}\|_{\mathbb{C}^d}$$

and $K_k(Q) = K_{k-1}(\hat{Q})$ otherwise. Hence

$$M_k^W f(x) = \|\langle W\rangle_Q^{1/2}\langle W\rangle_{K_k(Q)}^{-1}\langle W^{1/2}f\rangle_{K_k(Q)}\|_{\mathbb{C}^d}$$

$$= \max\left\{\|\langle W\rangle_Q^{1/2}\langle W\rangle_{K_{k-1}(\hat{Q})}^{-1}\langle W^{1/2}f\rangle_{K_{k-1}(\hat{Q})}\|_{\mathbb{C}^d}, \|\langle W\rangle_Q^{-1/2}\langle W^{1/2}f\rangle_Q\|_{\mathbb{C}^d}\right\}.$$

Observe that the time index k on the function K can be omitted without confusion. Notice that the sequence $(M_k^W f)_{k\geq 0}$ is adapted. As motivated by the scalar case, we compete the value using the interval retained from the previous step against the one from the finest filtration corresponding to k.

Let us discuss differences between $M_\infty^W f$ and M^W, say in the case when W and f are dyadic step functions and measurable in some $\mathcal{F}_k$. We observe that generally $M_\infty^W f$ is larger. If Q is an atom in $\mathcal{F}_k$ and $x \in Q$ then $W(x) = \langle W\rangle_Q$ and for $x \in Q$ we have

$$
\begin{aligned}
M_\infty^W f(x) &= \sup_{J:x\in J} \|\langle W\rangle_Q^{1/2}\langle W\rangle_J^{-1}\langle W^{1/2}f\rangle_J\|_{\mathbb{C}^d} \\
&\geqslant \sup_{J:Q\subseteq J} \|\langle W\rangle_Q^{1/2}\langle W\rangle_J^{-1}\langle W^{1/2}f\rangle_J\|_{\mathbb{C}^d} \\
&\geqslant M_k^W f(x).
\end{aligned}
$$

Indeed, the supremum taken in $M_k^W f(x)$ is a maximum over only two competitors. To illustrate, let $x \in Q \subset \hat{Q}$ with $Q \in \mathcal{D}_k$. In the passage from $k-1$ to k, the first factor changes from $\langle W \rangle_{\hat{Q}}^{1/2}$ to $\langle W \rangle_Q^{1/2}$ which may mean that $K_{k-1}(\hat{Q})$ is not a sensible choice for a competitor at all if one tried to estimate a 'true' maximal function at time k. However in the scalar case, one obtains an approximating sequence for $M^w f$.

We proceed with the proof of Theorem 2.7.

Proof. We modify the Bellman function argument from [12]. It shows a dimensionless bound for M_k^W uniformly in k. Let

$$B(F, f, L, W) = 4F - 2\langle W^{-1/2} f, L \rangle - 2\langle L, W^{-1/2} f \rangle + 2\langle L, L \rangle.$$

We think of the variables having the following roles:

$$W_Q = \langle W \rangle_Q, f_Q = \langle W^{1/2} f \rangle_Q, F_Q = \langle \|f\|_{\mathbb{C}^d}^2 \rangle_Q \text{ and}$$

$$L_Q = \langle W \rangle_Q^{1/2} \langle W \rangle_{K(Q)}^{-1} \langle W^{1/2} f \rangle_{K(Q)}$$

with interval $K(Q) \in \{K(\hat{Q}), Q\}$ inductively chosen for maximizing the norm $\|\langle W \rangle_Q^{1/2} \langle W \rangle_{K(Q)}^{-1} \langle W^{1/2} f \rangle_{K(Q)} \|_{\mathbb{C}^d}$ as described above.

Domain of B. The domain is given by $W_Q \geqslant 0$ and

(3.1) $$F_Q \geqslant \|W_Q^{-1/2} f_Q\|_{\mathbb{C}^d}^2.$$

To see this, recall the fact

$$\langle W_Q^{-1} f_Q, f_Q \rangle_{\mathbb{C}^d} = \sup_{e \neq 0} \frac{|\langle f_Q, e \rangle_{\mathbb{C}^d}|^2}{\langle W_Q e, e \rangle_{\mathbb{C}^d}},$$

for example from [5]. Then observe that since W is selfadjoint,

$$|\langle \langle W^{1/2} f, e \rangle_{\mathbb{C}^d} \rangle_Q|^2 \leq \langle \|f\|_{\mathbb{C}^d}^2 \rangle_Q \langle \langle We, e \rangle_{\mathbb{C}^d} \rangle_Q = F_Q \langle W_Q e, e \rangle_{\mathbb{C}^d}.$$

Since Q itself is an admissible choice for $K(Q) \in \{K(\hat{Q}); Q\}$ we require furthermore that

(3.2) $$\|L_Q\|_{\mathbb{C}^d} \geqslant \|W_Q^{-1/2} f_Q\|_{\mathbb{C}^d}.$$

Range of B. Notice that $L_{Q_0} = W_{Q_0}^{-1/2} f_{Q_0}$, since Q_0 is the only competitor. Therefore we have

$$B(F_{Q_0}, f_{Q_0}, L_{Q_0}, W_{Q_0}) = 4F_{Q_0} - 2\|W_{Q_0}^{-1/2} f_{Q_0}\|_{\mathbb{C}^d}^2 \leqslant 4F_{Q_0}.$$

We have also

$$B(F_Q, f_Q, L_Q, W_Q) \geqslant \|L_Q\|_{\mathbb{C}^d}^2$$

since

$$\begin{aligned} & B(F_Q, f_Q, L_Q, W_Q) \\ \geqslant \; & 4\langle W_Q^{-1/2} f_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} - 2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} \\ & -2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d} \\ = \; & \langle 2W_Q^{-1/2} f_Q - L_Q, 2W_Q^{-1/2} f_Q - L_Q \rangle_{\mathbb{C}^d} + \langle L_Q, L_Q \rangle_{\mathbb{C}^d} \\ \geqslant \; & \langle L_Q, L_Q \rangle_{\mathbb{C}^d}. \end{aligned}$$

Dynamics of B. If we denote by $Q_\pm$ the left and right halves of Q, then $W_{Q_\pm}$ and $F_{Q_\pm}$ and $f_{Q_\pm}$ are the corresponding averages so that these variables are of martingale type. But

$$L_{Q_\pm} = \begin{cases} W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q & \text{if} \quad \|W_{Q_\pm}^{-1/2} f_{Q_\pm}\|_{\mathbb{C}^d} \leqslant \|W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d} \\ W_{Q_\pm}^{-1/2} f_{Q_\pm} & \text{if} \quad \|W_{Q_\pm}^{-1/2} f_{Q_\pm}\|_{\mathbb{C}^d} > \|W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d} \end{cases}.$$

We claim that

$$B(F_Q, f_Q, L_Q, W_Q) \geqslant \frac{1}{2} B(F_{Q_+}, f_{Q_+}, L_{Q_+}, W_{Q_+}) + \frac{1}{2} B(F_{Q_-}, f_{Q_-}, L_{Q_-}, W_{Q_-}).$$

Case 1. No changes in $K(Q) \to K(Q_\pm)$, thus $K(Q_\pm) = K(Q)$. This implies

$$(3.3) \qquad \|W_{Q_\pm}^{-1/2} f_{Q_\pm}\|_{\mathbb{C}^d} \leqslant \|W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d}$$

and

$$L_{Q_\pm} = W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q.$$

Observe

$$\langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} + \langle W_{Q_-} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$= 2\langle W_Q W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} = 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$

so

$$B(F_Q, f_Q, L_Q, W_Q)$$
$$= 4F_Q - 2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} - 2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$
$$= 4F_Q - \langle f_{Q_+}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle f_{Q_-}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$\quad - \langle W_Q^{-1/2} L_Q, f_{Q_+} \rangle_{\mathbb{C}^d} - \langle W_Q^{-1/2} L_Q, f_{Q_-} \rangle_{\mathbb{C}^d}$$
$$\quad + \langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} + \langle W_{Q_-} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$= 2F_{Q_+} - \langle W_{Q_+}^{-1/2} f_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} - \langle L_{Q_+}, W_{Q_+}^{-1/2} f_{Q_+} \rangle_{\mathbb{C}^d} + \langle L_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d}$$
$$\quad + 2F_{Q_-} - \langle W_{Q_-}^{-1/2} f_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d} - \langle L_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} + \langle L_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d}$$
$$= \frac{1}{2} B(F_{Q_+}, f_{Q_+}, L_{Q_+}, W_{Q_+}) + \frac{1}{2} B(F_{Q_-}, f_{Q_-}, L_{Q_-}, W_{Q_-}).$$

We notice that in this case property (3.3) was not used.

Case 2. Two changes in $K(Q) \to K(Q_\pm)$ so $K(Q_\pm) = Q_\pm$. Let us thus assume

$$L_{Q_\pm} = W_{Q_\pm}^{-1/2} f_{Q_\pm}.$$

Due to the definition of $L_{Q_\pm}$ as a maximum of norms, we know that

$$(3.4) \qquad \|W_{Q_\pm}^{-1/2} f_{Q_\pm}\|_{\mathbb{C}^d}^2 \geqslant \|W_{Q_\pm}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d}^2$$

and therefore

$$(3.5) \quad \|L_{Q_+}\|_{\mathbb{C}^d}^2 + \|L_{Q_-}\|_{\mathbb{C}^d}^2 \quad \geqslant \quad \|W_{Q_+}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d}^2 + \|W_{Q_-}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d}^2$$
$$= \quad 2\|L_Q\|_{\mathbb{C}^d}^2 \geqslant 2\|W_Q^{-1/2} f_Q\|_{\mathbb{C}^d}^2,$$

where we used property (3.4) and the domain condition (3.2). Observe that

$$2\langle W_Q^{-1/2} f_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d}$$
$$-2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} - 2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$
$$= 2\langle W_Q^{-1/2} f_Q - L_Q, 2W_Q^{-1/2} f_Q - L_Q \rangle_{\mathbb{C}^d} \geqslant 0,$$

so using property (3.5)

$$-2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} - 2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$
$$\geqslant -2\|W_Q^{-1/2} f_Q\|_{\mathbb{C}^d}^2 \geqslant -\langle L_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} - \langle L_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d}.$$

We will now show the dynamics inequality for this case:

$$B(F_Q, f_Q, L_Q, W_Q)$$
$$= 4F_Q - 2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} - 2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$
$$\geqslant 2F_{Q_+} + 2F_{Q_-} - \langle L_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} - \langle L_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d}$$
$$= 2F_{Q_+} - \langle W_{Q_+}^{-1/2} f_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} - \langle L_{Q_+}, W_{Q_+}^{-1/2} f_{Q_+} \rangle_{\mathbb{C}^d} + \langle L_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d}$$
$$+ 2F_{Q_-} - \langle W_{Q_-}^{-1/2} f_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d} - \langle L_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} + \langle L_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d}$$
$$= \frac{1}{2} B(F_{Q_+}, f_{Q_+}, L_{Q_+}, W_{Q_+}) + \frac{1}{2} B(F_{Q_-}, f_{Q_-}, L_{Q_-}, W_{Q_-}).$$

Case 3. Mixed case, one change in $K(Q) \to K(Q_\pm)$ with $K(Q_+) = K(Q)$ and $K(Q_-) = K(Q)$. The other mixed case is symmetric. Let us thus assume

$$L_{Q_-} = W_{Q_-}^{-1/2} f_{Q_-} \text{ and } L_{Q_+} = W_{Q_+}^{1/2} W_Q^{-1/2} L_Q.$$

This gives

$$\|W_{Q_-}^{-1/2} f_{Q_-}\|_{\mathbb{C}^d} > \|W_{Q_-}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d} \text{ and } \|W_{Q_+}^{-1/2} f_{Q_+}\|_{\mathbb{C}^d} \leqslant \|W_{Q_+}^{1/2} W_Q^{-1/2} L_Q\|_{\mathbb{C}^d}.$$

First notice

$$0 \leqslant \langle W_{Q_-}^{-1/2} f_{Q_-} - W_{Q_-}^{1/2} W_Q^{-1/2} L_Q, W_{Q_-}^{-1/2} f_{Q_-} - W_{Q_-}^{1/2} W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$= \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} + \langle W_{Q_-}^{1/2} W_Q^{-1/2} L_Q, W_{Q_-}^{1/2} W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$- \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{1/2} W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_-}^{1/2} W_Q^{-1/2} L_Q, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d}$$
$$= \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} + \langle W_{Q_-} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}$$
$$- \langle f_{Q_-}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_Q^{-1/2} L_Q, f_{Q_-} \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d}$$
$$- \langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_-} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d}.$$

Thus

$$2\langle L_Q, L_Q \rangle_{\mathbb{C}^d} \geqslant \langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d}$$
$$+ \langle f_{Q_-}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} + \langle W_Q^{-1/2} L_Q, f_{Q_-} \rangle_{\mathbb{C}^d}$$

and therefore

$$
\begin{aligned}
&B(F_Q, f_Q, L_Q, W_Q) \\
={}& 4F_Q - 2\langle W_Q^{-1/2} f_Q, L_Q \rangle_{\mathbb{C}^d} - 2\langle L_Q, W_Q^{-1/2} f_Q \rangle_{\mathbb{C}^d} + 2\langle L_Q, L_Q \rangle_{\mathbb{C}^d} \\
\geqslant{}& 4F_Q - \langle f_{Q_+}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle f_{Q_-}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} \\
& -\langle W_Q^{-1/2} L_Q, f_{Q_+} \rangle_{\mathbb{C}^d} - \langle W_Q^{-1/2} L_Q, f_{Q_-} \rangle_{\mathbb{C}^d} \\
& +\langle f_{Q_-}, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} + \langle W_Q^{-1/2} L_Q, f_{Q_-} \rangle_{\mathbb{C}^d} \\
& +\langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} \\
={}& 2F_{Q_+} + 2F_{Q_-} \\
& -\langle W_{Q_+}^{-1/2} f_{Q_+}, W_{Q_+}^{1/2} W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_+}^{1/2} W_Q^{-1/2} L_Q, W_{Q_+}^{-1/2} f_{Q_+} \rangle_{\mathbb{C}^d} \\
& +\langle W_{Q_+} W_Q^{-1/2} L_Q, W_Q^{-1/2} L_Q \rangle_{\mathbb{C}^d} - \langle W_{Q_-}^{-1/2} f_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} \\
={}& 2F_{Q_+} - \langle W_{Q_+}^{-1/2} f_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} - \langle L_{Q_+}, W_{Q_+}^{-1/2} f_{Q_+} \rangle_{\mathbb{C}^d} + \langle L_{Q_+}, L_{Q_+} \rangle_{\mathbb{C}^d} \\
& +2F_{Q_-} - \langle W_{Q_-}^{-1/2} f_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d} - \langle L_{Q_-}, W_{Q_-}^{-1/2} f_{Q_-} \rangle_{\mathbb{C}^d} + \langle L_{Q_-}, L_{Q_-} \rangle_{\mathbb{C}^d} \\
={}& \frac{1}{2} B(F_{Q_+}, f_{Q_+}, L_{Q_+}, W_{Q_+}) + \frac{1}{2} B(F_{Q_-}, f_{Q_-}, L_{Q_-}, W_{Q_-}).
\end{aligned}
$$

Using the dynamics and size properties we obtain the inequality estimating the norm of M_k^W uniformly in k:

$$
\begin{aligned}
\sum_{J \in \mathcal{D}_k(Q_0)} \| L_J \|_{\mathbb{C}^d}^2 |J| & \leqslant \sum_{J \in \mathcal{D}_k(Q_0)} B(F_J, f_J, L_J, W_J)|J| \\
& \leqslant B(F_{Q_0}, f_{Q_0}, L_{Q_0}, W_{Q_0})|Q_0| \\
& \leqslant 4F_{Q_0}|Q_0|.
\end{aligned}
$$

$\square$

4. Embeddings

In this section we turn to the BET and the motivation and proof of the reduction estimate Theorem 2.3.

4.1. **Bilinear Embedding Theorem (BET).** We turn to the proof of Theorem 2.1.

Proof. Let $\mu(\mathcal{K}) = \sum_{Q \in \mathcal{K}} \alpha_Q$ for any collection $\mathcal{K}$ of dyadic cubes. Let F be any nonnegative function defined on the dyadic cubes. Then let $\{F(Q) > \lambda\}$ denote the collection of cubes Q such that $F(Q) > \lambda$. It follows that

$$
\int_0^\infty \mu(\{F(Q) > \lambda\}) \mathrm{d}\lambda = \sum_{Q \in \mathcal{D}(Q_0)} F(Q)\alpha_Q,
$$

which is the classical fact on Choquet integrals.

Let us define

$$
F(Q) = |\langle \langle W \rangle_Q^{-1} \langle W^{1/2} f \rangle_Q, \langle W^{-1} \rangle_Q^{-1} \langle W^{-1/2} g \rangle_Q \rangle_{\mathbb{C}^d}|
$$

and for any $\lambda > 0$, let $\mathcal{J}_\lambda$ denote the collection of maximal dyadic intervals for which $F(Q) > \lambda$. Hence

$$\sum_{Q \in \mathcal{J}_\lambda} \alpha_Q \leqslant \sum_{Q \in \mathcal{J}_\lambda} \sum_{Q' \in \mathcal{D}(Q)} \alpha_{Q'} \leqslant \sum_{Q \in \mathcal{J}_\lambda} |Q|.$$

Now let

$$\Phi(x) = M^W f(x) M^{W^{-1}} g(x).$$

Observe that with

$$A_Q^W f(x) := W^{1/2}(x) \langle W \rangle_Q^{-1} \langle W^{1/2} f \rangle_Q$$

and

$$A_Q^{W^{-1}} g(x) = W^{-1/2}(x) \langle W^{-1} \rangle_Q^{-1} \langle W^{-1/2} g \rangle_Q,$$

Cauchy Schwarz yields for all x and all Q

$$F(Q) \leqslant \|A_Q^W f(x)\|_{\mathbb{C}^d} \|A_Q^{W^{-1}} g(x)\|_{\mathbb{C}^d}.$$

Now if $x \in Q$ with $F(Q) > \lambda$, then $\Phi(x) > \lambda$. Hence

$$\sum_{Q \in \mathcal{J}_\lambda} |Q| \leqslant |\{x \in \mathbb{R} : \Phi(x) > \lambda\}|.$$

Integrating with respect to $d\lambda$ gives

$$\sum_{Q \in \mathcal{D}(Q_0)} \alpha_Q F(Q) \leqslant \int_0^\infty |\{x \in \mathbb{R} : \Phi(x) > \lambda\}| \mathrm{d}\lambda = \int_{Q_0} M^W f(x) M^{W^{-1}} g(x) \mathrm{d}x$$

Using the above estimate of the maximal function in Theorem 2.5 and an application of Cauchy Schwarz finishes the estimate. $\qquad\square$

4.2. **Altered Carleson condition.** As mentioned earlier, our BET reduces to the classical version of the scalar case. The first proof of the scalar theorem was by the Bellman method, using a rather cleverly built function, that in a way allows to deduce the bilinear version BET from the linear one CET, see [13], [17] and [16]. This is exactly what we do here, too, although with completely different methods, that are applicable to the matrix weighted case. The original scalar version of BET, first formulated in [13], seemed to require three Carleson conditions, namely

$$\frac{1}{|K|} \sum_{Q \in \mathcal{D}(K)} \frac{\alpha_Q}{\langle w \rangle_Q} \leqslant C \langle w^{-1} \rangle_K \quad \forall K \in \mathcal{D}(Q_0),$$

$$\frac{1}{|K|} \sum_{Q \in \mathcal{D}(K)} \frac{\alpha_Q}{\langle w^{-1} \rangle_Q} \leqslant C \langle w \rangle_K \quad \forall K \in \mathcal{D}(Q_0),$$

$$\frac{1}{|K|} \sum_{Q \in \mathcal{D}(K)} \alpha_Q \leqslant C \quad \forall K \in \mathcal{D}(Q_0).$$

As mentioned in the introduction, it turns out that the first two conditions can be removed in the scalar case through the use of an embedding–like Bellman function, thus only retaining one Carleson condition, the analog of what we used here. This is implicit in [6] and similar considerations also appeared in [1]. This reduction also holds in the case of a scalar Carleson measure and a matrix weight, as stated in Theorem 2.3. We turn to its proof.

Proof. Consider the matrix valued Bellman function of positive matrix variables U, V and scalar variable m

$$B(U, V, m) = U - (m+1)^{-1}V^{-1}.$$

The variables we have in mind are matrix valued $U_K = \langle W \rangle_K$, $V_K = \langle W^{-1} \rangle_K$ and scalar valued $m_K = |K|^{-1} \sum_{Q \in \mathcal{D}(K)} \alpha_Q$.

Domain of B. This function has the domain conditions

$$\text{Id} \leqslant V^{1/2} U V^{1/2} \text{ and } 0 \leqslant m \leqslant 1.$$

Indeed, the first condition is equivalent to $U \geqslant V^{-1}$ and is implied by Lemma 3.2 in [20]: if W is a matrix weight such that W and W^{-1} are summable on Q, then for all vectors e,

$$\langle \langle W \rangle_Q e, e \rangle \geqslant \langle \langle W^{-1} \rangle_Q^{-1} e, e \rangle.$$

Range of B. We have the size estimate

$$(4.1) \qquad\qquad 0 \leqslant B(U, V, m) \leqslant U.$$

Indeed, $0 \leqslant U - V^{-1} \leqslant U - (m+1)^{-1}V^{-1} \leqslant U$.

Dynamics of B. The function B is concave: Dropping the linear dependence on U, its Hessian acting on the matrix difference ΔV and scalar Δm is

$$-2V^{-1}\Delta V V^{-1}\Delta V V^{-1}(m+1)^{-1} - 2V^{-1}\Delta V V^{-1}(m+1)^{-2}\Delta m - 2(m+1)^{-3}V^{-1}(\Delta m)^2.$$

Observe that

$$V^{-1}\Delta V V^{-1}\Delta V V^{-1}(m+1)^{-1} \geqslant 0,$$

and

$$(m+1)^{-3}V^{-1}(\Delta m)^2 \geqslant 0.$$

Now add these two positive terms, then factor $-(m+1)^{-1}$ and reverse the sign:

$$V^{-1}\Delta V V^{-1}\Delta V V^{-1} + 2V^{-1}\Delta V V^{-1}(m+1)^{-1}\Delta m + (m+1)^{-2}V^{-1}(\Delta m)^2$$
$$= (\Delta m(m+1)^{-1}\,\text{Id} + V^{-1}\Delta V)V^{-1}(\Delta V V^{-1} + \text{Id}(m+1)^{-1}\Delta m) \geqslant 0.$$

This inequality proves concavity of B, which in turn implies mid–point concavity

$$(4.2) \qquad B(U, V, m)$$
$$\geqslant \frac{1}{2}B(U + \Delta U, V + \Delta U, m + \Delta m) + \frac{1}{2}B(U - \Delta U, V - \Delta V, m - \Delta m).$$

We have also

$$(4.3) \qquad\qquad (\partial B/\partial m)(U, V, m) = (m+1)^{-2}V^{-1} \geqslant \frac{1}{4}V^{-1}.$$

With martingale matrix variables $U_K = \langle W \rangle_K$, $V_K = \langle W^{-1} \rangle_K$ and scalar variable $m_K = |K|^{-1} \sum_{Q \in \mathcal{D}(K)} \alpha_Q$ in the domain, we see that

$$(4.4) \qquad\qquad m_K - |K|^{-1}\alpha_K = \frac{1}{2}(m_{K_-} + m_{K_+}).$$

The usual argument gives the estimate of the operator sum: Fix $K \in \mathcal{D}(Q_0)$ and estimate

$$
\begin{aligned}
|K|\langle W \rangle_K &= |K|U_K \\
&\geqslant |K|B(U_K, V_K, m_K) \\
&= |K|B(U_K, V_K, m_K) - |K|B(U_K, V_K, m_K - |K|^{-1}\alpha_K) \\
&\quad + |K|B(U_K, V_K, m_K - |K|^{-1}\alpha_K) \\
&\geqslant \frac{1}{4}V_K^{-1}\alpha_K + |K|B(U_K, V_K, m_K - |K|^{-1}\alpha_K) \\
&\geqslant \frac{1}{4}V_K^{-1}\alpha_K + |K_-|B(U_{K_-}, V_{K_-}, m_{K_-}) + |K_+|B(U_{K_+}, V_{K_+}, m_{K_+}).
\end{aligned}
$$

In these inequalities we used the size estimate (4.1), the intermediate value theorem together with the derivative estimate (4.3) and the dynamics of the martingale variables and (4.4), and the concavity of B in the form of mid-point concavity equation (4.2). Iterating this argument gives the desired estimate. □

5. Questions

Our theorems leave room for improvement in several directions.

5.1. A_2 conjecture.

Currently, the matrix A_2 conjecture for the Hilbert transform is under investigation. This question asks for the growth of the function Φ so that

$$
\|Hf\|_{L^2(\mathbb{C}^d;W)} \leq C\Phi([W]_{A_2})\|f\|_{L^2(\mathbb{C}^d;W)},
$$

where the Hilbert transform H is applied to the vector function f componentwise and the matrix A_2 characteristic of Treil and Volberg [20] is defined as

$$
[W]_{A_2} = \sup_Q \|\langle W \rangle_Q^{1/2} \langle W^{-1} \rangle_Q^{1/2}\|^2.
$$

In the scalar case the best growth is linear, and this is optimal, [16]. In the original text on the matrix case [20] that laid the ground for these investigations, the dependence on $[W]_{A_2}$ was not tracked. The first quantified result with better estimates is [2], which obtains the power $3/2$ with an additional logarithmic term. In the text [10] the logarithmic term was dropped and the best known estimate of power $3/2$ stands at current time. To improve this power further, one may aim at some improvements in the embedding theorems we demonstrated in this note:

- Even within the framework of scalar α_Q in BET it would be useful to be able to get an estimate of the (larger) sum

$$
\sum_{Q \in \mathcal{D}(Q_0)} \alpha_Q \|\langle W \rangle_Q^{-1} \langle W^{1/2}f \rangle_Q\|_{\mathbb{C}^d} \|\langle W^{-1} \rangle_Q^{-1} \langle W^{-1/2}g \rangle_Q\|_{\mathbb{C}^d}
$$

  as this may improve the matrix weighted estimate for certain classical operators. Staying with scalar coefficients is unlikely to press the constants to the desired linear estimate in the matrix A_2 constant, but one may hope for just an extra logarithmic term instead of the extra half power in current estimates.

- Does Theorem 2.3 hold with a matrix Carleson measure? From the required size conditions a natural choice for a Bellman function is

$$U - V^{-1/2}(M + \mathrm{Id})^{-1}V^{-1/2}.$$

  It is not clear to us if, for example, the convexity in V is as required. Indeed, to obtain derivatives of square roots, one may use the spectral resolution formula.
- Find a version of BET with matrix coefficients α_Q.

5.2. **Dimensional growth.** In the 90s the dimensional growth of matrix versions of typical scalar results was under investigation. There was a series of results on the Carleson embedding theorem with matrix measure and the papaproducts or Hankel operators with matrix symbol [9], [15], [11]. Some of the recent results indicate that dimensional growth may be a finer indicator of non–commutativity than the growth with the A_2 constant.

- Find the dimensional growth of the matrix weighted CET Theorem 2.2. When the weight is the identity, then the dimensional growth is $\log^2(d)$ and this is sharp, see [11]. It is not clear if the upper estimate is attainable in the presence of a non–trivial matrix weight. The best to date estimate is of order d^2. It appears that the matrix weight induces additional non–commutativity that causes the clever argument in [11] to fail.
- What is the dimensional growth of the matrix maximal function of Definition 2.4?

References

[1] O. Beznosova, *Linear bound for the dyadic paraproduct on weighted Lebesgue space $L_2(w)$*, J. Funct. Anal. **255** (2008), no. 4, 994–1007.

[2] K. Bickel, S. Petermichl, and B. Wick, *Bounds for the Hilbert transform with matrix A_2 weights*, J. Funct. Anal. **250** (2016), no. 5, 1719–1743.

[3] L. Carleson, *An interpolation problem for bounded analytic functions*, Amer. J. Math. **80** (1958), 921–930.

[4] _____, *Interpolations by bounded analytic functions and the corona problem*, Ann. Math. **76** (1962), 547–559.

[5] A. Culiuc and S. Treil, *The Carleson embedding theorem with matrix weights*, To appear in IMRN.

[6] K. Domelevo and S. Petermichl, *Differential subordination under change of law*, Ann. Prob. **47** (2019), no. 2, 896–925.

[7] T. Hytönen, S. Petermichl, and A. Volberg, *The sharp square function estimate with matrix weight*, Discrete Analysis **2** (2019), DOI 10.19086/da.7597.

[8] J. Isralowitz, H.-K. Kwon, and S.Pott, *Matrix-weighted norm inequalities for commutators and paraproducts with matrix symbols*, J. London Math. Soc. **96** (2017), no. 1, 243–270.

[9] N. Katz, *Matrix valued paraproducts*, J. of Fourier Anal. Appl. **300** (1997), 913–921.

[10] F. Nazarov, S. Petermichl, Sergei Treil, and A. Volberg, *Convex body domination and weighted estimates with matrix weights*, Adv. Math. **318** (2017), 279–306.

[11] F. Nazarov, G. Pisier, S. Treil, and A. Volberg, *Sharp estimates in vector Carleson imbedding theorem and for vector paraproducts*, J. Reine Angew. Math. **542** (2002), 147–171.

[12] F. Nazarov and S. Treil, *The hunt for a Bellman functions: applications to estimates of singular integral operators and to other classical problems in harmonic analysis*, Algebra i Analysis **8** (1997), no. 5, 32–162.

[13] F. Nazarov, S. Treil, and A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc. **12** (1999), no. 4, 909–928.

[14] _____, *Two weight inequalities for individual Haar multipliers and other well localized operators*, Math. Res. Lett. **15** (2008), no. 3, 583–597.

[15] S. Petermichl, *Dyadic Shifts and a Logarithmic Estimate for Hankel Operators with Matrix Symbol*, Comptes Rendus Acad. Sci. Paris **330** (2000), no. 1, 455–460.

[16] _____, *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 J. Wittwer, *A sharp estimate for the weighted Hilbert transform via Bellman functions*, Michigan Math. J. **50** (2002), no. 1, 71–87.

[18] C. Thiele, S. Treil, and A. Volberg, *Weighted martingale multipliers in the non-homogeneous setting and outer measure spaces*, Adv. Math. **285** (2015), 1155–1188.

[19] S. Treil, *Mixed $A_2 - A_\infty$ estimates of the non-homogeneous vector square function with matrix weights*, Submitted.

[20] S. Treil and A. Volberg, *Wavelets and the angle between past and future*, J. Funct. Anal. **143** (1997), no. 2, 269–308.