Math. Ann. (2016) 365:1111–1135
DOI 10.1007/s00208-015-1320-y

Mathematische Annalen

CrossMark

A pointwise estimate for positive dyadic shifts and some applications

José M. Conde-Alonso[1] · **Guillermo Rey**[2]

Received: 27 April 2015 / Revised: 28 September 2015 / Published online: 27 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract We prove a (sharp) pointwise estimate for positive dyadic shifts of complexity m which is linear in the complexity. This can be used to give a pointwise estimate for Calderón-Zygmund operators and to answer a question originally posed by Lerner. Several applications to weighted estimates for both multilinear Calderón-Zygmund operators and square functions are discussed.

1 Introduction

One particularly useful way to study many important operators in Harmonic Analysis is that of decomposing them into sums of simpler dyadic operators. An example of a recent striking result using this strategy is the proof of the sharp weighted estimate for the Hilbert transform by Petermichl [23]. This was a key step towards the full A_2 theorem for general Calderón-Zygmund operators, finally proven by Hytönen in [9]. Of course there are many instances of this useful technique, but we will not try to give a thorough historical overview here.

J. M. Conde-Alonso was partially supported by the ERC StG-256997-CZOSQP, the Spanish Grant MTM2010-16518 and by ICMAT Severo Ochoa Grant SEV-2011-0087 (Spain).

✉ Guillermo Rey
reyguill@math.msu.edu

José M. Conde-Alonso
jose.conde@icmat.es

1 Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, C/ Nicolás Cabrera 13-15, 28049 Madrid, Spain

2 Department of Mathematics, Michigan State University, East Lansing, MI 48824-1027, USA

The proof in [9] was a *tour de force* which was the culmination of many previous partial efforts by others, see [9] and the references therein. Hytönen did not only prove the A_2 theorem, but he also showed that general Calderón-Zygmund operators could be represented as averages of certain simpler "Haar shifts" in the spirit of [23]. The sharp weighted bound then followed from the corresponding one for these simpler operators.

Later, Lerner gave a simplification of the A_2 theorem in [15] which avoided the use of most of the complicated machinery in [9]; it mainly relied on a general pointwise estimate for functions in terms of positive dyadic operators which had already been proven in [14]. The weighted result for the positive dyadic shifts that this contribution reduced the problem to had already been shown before in [12], see also [3] and [4]. More precisely, the proof of Lerner (essentially) gave the following pointwise estimate for general Calderón-Zygmund operators T: for every dyadic cube Q

$$|Tf(x)| \lesssim \sum_{m=0}^{\infty} 2^{-\delta m} \mathcal{A}_{\mathcal{S}}^m |f|(x) \quad \text{for a.e. } x \in Q, \tag{1.1}$$

where $\delta > 0$ depends on the operator T, $\mathcal{S}$ are collections of dyadic cubes (belonging to same dyadic grid for each fixed $\mathcal{S}$) which depend on f, T and m, and $\mathcal{A}_{\mathcal{S}}^m$ are positive dyadic operators defined by

$$\mathcal{A}_{\mathcal{S}}^m f(x) = \sum_{Q \in \mathcal{S}} \langle f \rangle_{Q^{(m)}} \mathbb{1}_Q(x),$$

where $Q^{(m)}$ denotes the mth dyadic parent of Q. Moreover, the collections $\mathcal{S}$ in (1.1) are *sparse* in the usual sense: given $0 < \eta < 1$, we say that a collection of cubes $\mathcal{S}$ belonging to the same dyadic grid is η-sparse if for all cubes $Q \in \mathcal{S}$ there exist measurable subsets $E(Q) \subset Q$ with $|E(Q)| \geq \eta |Q|$ and $E(Q) \cap E(Q') = \emptyset$ unless $Q = Q'$. A collection is called simply sparse if it is $\frac{1}{2}$-sparse.

From this pointwise estimate Lerner continues the proof by showing that bounding the operator norm of each $\mathcal{A}_{\mathcal{S}}^m$ can be reduced to just estimating the operator norm of $\mathcal{A}_{\mathcal{S}'}^0$ in the same space for all possible sparse collections $\mathcal{S}'$. More precisely, he shows that

$$\|\mathcal{A}_{\mathcal{S}}^m f\|_{\mathbb{X}} \lesssim (m+1) \sup_{\mathscr{D}, \mathcal{S}'} \|\mathcal{A}_{\mathcal{S}'}^0 f\|_{\mathbb{X}}, \tag{1.2}$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collection $\mathcal{S}' \subset \mathscr{D}$, and where $\mathbb{X}$ is any Banach function space, in the sense of [1, Chapter 1].

It is at this point where the duality of $\mathbb{X}$ is needed in the argument; the operators $\mathcal{A}_{\mathcal{S}}^m$ do not lend themselves to Lerner's pointwise formula, while their adjoints do. Consequently, the question of what to do when no duality is present was left open. Our main result answers this question by proving a stronger (though localized) statement: the operators $\mathcal{A}_{\mathcal{S}}^m$ are actually pointwise bounded by positive dyadic 0-shifts:

Theorem A *Let P be a cube and $\mathcal{S}$ a sparse collection of dyadic subcubes Q such that $Q^{(m)} \subseteq P$, then for all nonnegative integrable functions f on P there exists another sparse collection $\mathcal{S}'$ of dyadic subcubes of P such that*

$$\mathcal{A}^m_{\mathcal{S}} f(x) \lesssim (m+1)\mathcal{A}^0_{\mathcal{S}'} f(x) \quad \forall x \in P \tag{1.3}$$

In fact, we prove Theorem A in a slightly more general setting: first, the statement is proven for a certain natural multilinear generalization of the operators $\mathcal{A}^m_{\mathcal{S}}$. Second, the sparse collection $\mathcal{S}$ is replaced by a more general Carleson sequence. The relevant details are given in the next section.

The novelty in our approach is two-fold: we directly attack the pointwise estimate for the operators $\mathcal{A}^m$, instead of bounding their norm in various spaces. Also, in proving the pointwise bound we develop an algorithm that constructively selects those cubes which will form the family $\mathcal{S}'$. This algorithm has "memory" in a certain sense: each iteration takes into account the previous steps, a feature which is crucial in our method to ensure that $\mathcal{S}$ is sparse.

As a corollary of Theorem A, we find an analogue of (1.1) for Calderón-Zygmund operators with more general moduli of continuity (see the next section for the precise definition). In particular, we obtain the following pointwise estimate for Calderón-Zygmund operators:

Corollary A.1 *If P is a dyadic cube, f is an integrable function supported on P and T is a Calderón-Zygmund operator whose kernel has modulus of continuity ω, then*

$$|Tf(x)| \lesssim \sum_{m=0}^{\infty} \omega(2^{-m})(m+1)\mathcal{A}^0_{\mathcal{S}_m}|f|(x) \quad \textit{for a.e. } x \in P, \tag{1.4}$$

where $\mathcal{S}_m$ are sparse collections belonging to at most 3^d different dyadic grids.

Moreover, if we know that ω satisfies the logarithmic Dini condition:

$$\int_0^1 \omega(t)\left(1+\log\left(\frac{1}{t}\right)\right)\frac{dt}{t} < \infty, \tag{1.5}$$

then we can find sparse collections $\{\mathcal{S}'_1, \ldots, \mathcal{S}'_{3^d}\}$, belonging to possibly different dyadic grids, such that

$$|Tf(x)| \lesssim \sum_{i=1}^{3^d} \mathcal{A}^0_{\mathcal{S}'_i}|f|(x) \quad \textit{for a.e. } x \in P. \tag{1.6}$$

The factor m in (1.2) precluded a naive adaptation of the proof in [16] to an A_2 theorem with the usual Dini condition:

$$\int_0^1 \omega(t)\frac{dt}{t} < \infty, \tag{1.7}$$

since the sum

$$\sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \simeq \int_0^1 \omega(t)\left(1+\log\frac{1}{t}\right)\frac{dt}{t} \tag{1.8}$$

could diverge for some moduli ω satisfying only (1.7). Moreover, it was shown in [8] that the weak-type $(1, 1)$ norm of the adjoints of the operators $\mathcal{A}_{\mathcal{S}}^m$ was at least linear in m, even in the unweighted case, so using duality prevented an extension of this type. However, although our argument does not quite give an A_2 theorem for Calderón-Zygmund operators satisfying the Dini condition (we still need (1.8) to be finite), our proof avoids the use of duality and the study of the adjoint operators $(\mathcal{A}_{\mathcal{S}}^m)^*$. It thus removes at least one of the obstructions to possible proofs of the A_2 theorem with the Dini condition which follow this strategy. Hence, removing the linear factor of m in Theorem A remains as an open problem.

Apart from being interesting in its own right, a bound for Calderón-Zygmund operators by these sums of positive 0-shifts in cases where there is no duality has interesting applications, some of which we describe later. Before, let us state a second corollary to Theorem A.1:

Corollary A.2 *Let $\|\cdot\|_{\mathbb{X}}$ be a function quasi-norm (see Sect. 2) and T a Calderón-Zygmund operator satisfying the logarithmic Dini condition, then*

$$\|Tf\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}_{\mathcal{S}}^0 |f|\|_{\mathbb{X}}, \tag{1.9}$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collections $\mathcal{S} \subset \mathscr{D}$.

We now describe two immediate applications of our result. First we can continue the program, initiated in [5] and extended in [21], which aims to extend the sharp weighted estimates for Calderón-Zygmund operators to their multilinear analogues (as in [6]). In particular we obtain

Theorem B *Let T be a multilinear Calderón-Zygmund operator. Suppose $1 < p_1, \ldots, p_k < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$ and $\vec{w} \in A_{\vec{P}}$. Then*

$$\|T\vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [\vec{w}]_{A_{\vec{P}}}^{\max\left(1, \frac{p_1'}{p}, \ldots, \frac{p_k'}{p}\right)} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)}. \tag{1.10}$$

The same theorem was proven in [21] but with the additional hypothesis that p had to be at least 1. The proof of this theorem is an application of the result in [21] which proved the same estimate (without the condition $p \geq 1$) but for a multilinear analogue of the operators $\mathcal{A}_{\mathcal{S}}^m$, together with Theorem A. In fact, we will need a multilinear version of Theorem A which we state and prove in the next section.

Our second application is a sharp aperture weighted estimate for square functions which extends a result in [17]. In particular:

Theorem C *Let $\alpha > 0$, then the square function $S_{\alpha,\psi}$ for the cone in $\mathbb{R}_+^{d+1}$ of apperture α and the standard kernel ψ satisfies*

$$\|S_{\alpha,\psi} f\|_{L^{p,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_p}^{1/p} \|f\|_{L^p(\mathbb{R}^d,w)} \quad \textit{for } 1 < p < 2$$

and

$$\|S_{\alpha,\psi} f\|_{L^{2,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_2}^{1/2}(1+\log[w]_{A_2})\|f\|_{L^2(\mathbb{R}^d,w)}. \tag{1.11}$$

An analogous result was shown in [17] for $2 < p < 3$:

$$\|S_{\alpha,\psi} f\|_{L^{p,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_p}^{1/2}(1+\log[w]_{A_p})\|f\|_{L^p(\mathbb{R}^d,w)}.$$

The proof relies on the use of Lerner's pointwise formula and previous results by Lacey and Scurry [13]. However, in [17] the requirement of $p > 2$ was necessary for the same reason why the proof of the multilinear weighted estimates required $p \geq 1$ (a certain space had no satisfactory duality properties). Theorem A can be used in almost the same way as with the weighted multilinear estimates to prove Theorem C. Indeed, the proofs in [13,17] reduce the problem to estimating certain discrete positive operators which can be seen to be particular instances of the positive multilinear m-shifts used in the proof of Theorem B.

As was noted in [13], estimate (1.11) can be seen as an analogue of the result in [19] stablishing the endpoint weighted weak-type estimate for Calderón-Zygmund operators

$$\|Tf\|_{L^{1,\infty}(w)} \lesssim [w]_{A_1}(1+\log[w]_{A_1})\|f\|_{L^1(w)}.$$

See also [22] for a similar estimate from below and more information on the sharpness of this estimate, known as the weak A_1 conjecture. In this direction, it seems reasonable that Lacey and Scurry's proof in [13] could be adapted to the multilinear setting, however we will not pursue this problem here.

Finally, as a third application of our results, it is possible to give a more direct proof of the result in [10] for the q-variation of Calderón-Zygmund operators satisfying the logarithmic Dini condition by using the pointwise estimate analogous to (1.1) in [10] and then applying Theorem A. However, we will not pursue this argumentation either.

Shortly before uploading this preprint, Andrei Lerner kindly communicated to the authors that he, jointly with Fedor Nazarov, had independently proven a theorem very similar to Corollary A.1 [18]. Though the hypothesis are the same, their result differs from the one in this note in that we give a localized pointwise estimate while their pointwise estimate is valid for all of $\mathbb{R}^d$. However, our result seems to be as powerful in the applications.[1]

2 Pointwise domination

The goal of this section is the proof of Theorem A and its consequences as stated in the introduction. We will prove the result in the level of generality of multilinear operators. Given a cube P_0 on $\mathbb{R}^d$, we will denote by $\mathscr{D}(P_0)$ the dyadic lattice obtained by successive dyadic subdivisions of P_0. By a dyadic grid we will denote any dyadic

[1] Since we uploaded this document to *arXiv*, two other articles have appeared: [7,11], in which similar estimates are obtained.

lattice composed of cubes with sides parallel to the axis. A k-linear positive dyadic shift of complexity m is an operator of the form

$$\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) = \mathcal{A}^m_{P_0,\alpha}(f_1, f_2, \ldots, f_k)(x) := \sum_{\substack{Q\in\mathscr{D}(P_0)\\ Q^{(m)}\subseteq P_0}} \alpha_Q \left(\prod_{i=1}^k \langle f_i\rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x).$$

As a first step towards the proof of Theorem A, it is convenient to separate the scales of (or *slice*) $\mathcal{A}^m_{P_0,\alpha}$ as follows:

$$\begin{aligned}\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) &= \sum_{n=0}^{m-1}\sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm+n}(P_0)} \alpha_Q \left(\prod_{i=1}^k \langle f_i\rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x)\\ &=: \sum_{n=0}^{m-1} \mathcal{A}^{m,n}_{P_0,\alpha}\vec{f}(x).\end{aligned}$$

Note that $\mathscr{D}_k(P_0)$ denotes the kth generation of the lattice $\mathscr{D}(P_0)$. Now we rewrite $\mathcal{A}^{m;n}_{P_0,\alpha}$ as a sum of disjointly supported operators of the form $\mathcal{A}^{m;0}_{P,\alpha}$. Indeed,

$$\begin{aligned}\mathcal{A}^{m;n}_{P_0,\alpha}\vec{f}(x) &= \sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm+n}(P_0)} \alpha_Q \left(\prod_{i=1}^k \langle f_i\rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x)\\ &= \sum_{P\in\mathscr{D}_n(P_0)}\sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm}(P)} \alpha_Q \left(\prod_{i=1}^k \langle f_i\rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x)\\ &= \sum_{P\in\mathscr{D}_n(P_0)} \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x),\end{aligned}$$

which leads to the expression

$$\mathcal{A}^m_{\alpha,P_0}\vec{f}(x) = \sum_{n=0}^{m-1}\sum_{P\in\mathscr{D}_n(P_0)} \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x).$$

We say that a sequence $\{\alpha_Q\}_{Q\in\mathscr{D}(P_0)}$ is Carleson if its Carleson constant $\|\alpha\|_{\mathrm{Car}(P_0)} < \infty$, where

$$\|\alpha\|_{\mathrm{Car}(P_0)} = \sup_{P\in\mathscr{D}(P_0)} \frac{1}{|P|}\sum_{Q\in\mathscr{D}(P)} \alpha_Q|Q|.$$

The following intermediate step is the key to our approach:

Proposition 2.1 *Let $m \geq 1$ and α be a Carleson sequence. For integrable functions $f_1, \dots, f_k \geq 0$ on P_0 there exists a sparse collection $\mathcal{S}$ of cubes in $\mathscr{D}(P_0)$ such that*

$$\mathcal{A}^{m;0}_{P_0,\alpha} \vec{f}(x) \leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q \in \mathcal{S}} \left(\prod_{i=1}^{k} \langle f_i \rangle_Q \right) \mathbb{1}_Q(x),$$

where C_1 only depends on k and d, and in particular is independent of m.

To prove Proposition 2.1 we will proceed in three steps: we will first construct the collection $\mathcal{S}$, then show that we have the required pointwise bound, and finally that $\mathcal{S}$ is sparse. By homogeneity, we will assume that $\|\alpha\|_{\mathrm{Car}(P_0)} = 1$. Also, we will assume that the sequence α is finite, but our constants will be independent of the number of elements in the sequence.

Let $\Delta_{P_0} = 0$ and, for each $Q \in \mathscr{D}_{mj}(P_0)$ with $j \geq 0$, define the sequence $\{\gamma_Q\}_Q$ by

$$\gamma_Q = \max_{R \in \mathscr{D}_m(Q)} \alpha_R.$$

For each $Q \in \mathscr{D}_{mj}(P_0)$ with $j \geq 0$, we will inductively define the quantities Δ_Q and β_Q as follows:

$$\beta_Q = \begin{cases} 0 & \text{if } \Delta_Q - \left(\prod_{i=1}^{k} \langle f_i \rangle_Q \right) \gamma_Q \geq 0 \\ 2^{2(k+1)} C_W & \text{otherwise,} \end{cases}$$

where C_W is the boundedness constant of the unweighted endpoint weak-type of the operators $\mathcal{A}^m$ proved in Theorem 4.1 in the Appendix. Also, for every $R \in \mathscr{D}_m(Q)$ we define

$$\Delta_R = \Delta_Q + (\beta_Q - \alpha_R) \left(\prod_{i=1}^{k} \langle f_i \rangle_Q \right).$$

Note that the definition only applies to cubes in $\mathscr{D}_{mj}(P_0)$ for some j. For all other cubes in $\mathscr{D}_{P_0}$, we set $\beta_Q = \Delta_Q = 0$. The collection $\mathcal{S}$ consists of those cubes $Q \in \mathscr{D}(P_0)$ for which $\beta_Q \neq 0$. Note that, since $2^{2(k+1)} C_W > 1 = \|\alpha\|_{\mathrm{Car}(P_0)} \geq \alpha_R$ for all R and by the definition of γ_Q, we must have $\Delta_Q \geq 0$ for all Q. This can be easily seen by induction.

Remark 2.2 We are trying to construct a sparse operator of complexity 0 which dominates $\mathcal{A}^{m;0}_{P_0,\alpha}$. One way to achieve this is to let $\mathcal{S}$ be the collection of all dyadic subcubes of P_0, but of course this does not yield a sparse collection. A better way would be to let $\mathcal{S}$ consist of all dyadic cubes in P_0 for which at least one of its mth generation children R satisfies $\alpha_R > 0$; unfortunately this yields a collection $\mathcal{S}$ which is not sparse, and in fact it can be seen that the Carleson sequence β associated with this collection can have a Carleson norm $\|\beta\|_{\mathrm{Car}(P_0)}$ which grows exponentially in m.

The main problem with this approach is that, when the time comes to decide whether a cube should be in $\mathcal{S}$ or not, we do not take into account which cubes have been selected in the previous steps. Note that whenever we add a cube Q to $\mathcal{S}$ we are not only "helping" to dominate the portion of $\mathcal{A}^{m;0}_{P_0,\alpha}$ coming from Q, but also what may come from any of its descendants.

One can account for this by having the algorithm use a sort of "memory" to, essentially, keep track of how many cubes in $\mathcal{S}$ (appropriately weighted with the averages of $\vec{f}$) lie above any particular cube. This is the purpose of Δ_Q. This can also be seen as the stopping time algorithm which selects a cube whenever the previously selected cubes do not provide enough height to dominate the operator until that point.

Lemma 2.3 *We have the pointwise bound*

$$\mathcal{A}^{m;0}_{P_0,\alpha}\vec{f}(x) \leq \sum_{Q\in\mathscr{D}(P_0)} \beta_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_Q\right) \mathbb{1}_Q(x). \tag{2.1}$$

Proof We will prove by induction the following claim: if $P \in \mathscr{D}_{jm}(P_0)$ for some $j \geq 0$, then

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) \leq \Delta_P + \sum_{Q\in\mathscr{D}(P)} \beta_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_Q\right) \mathbb{1}_Q(x). \tag{2.2}$$

Note that, when $P = P_0$, this is exactly (2.1). Since α is finite, there is a smallest $j_0 \in \mathbb{N}$ such that $\alpha_Q = 0$ for all cubes $Q \in \mathscr{D}_{\geq j_0 m}(P_0)$.[2] Let Q be any cube in $\mathscr{D}_{j_0 m}(P_0)$, we obviously have

$$\mathcal{A}^{m;0}_{Q,\alpha}\vec{f} \equiv 0 \quad \text{in } Q.$$

Since $\Delta_Q \geq 0$, the claim (2.2) is trivial for $P \in \mathscr{D}_{j_0 m}(P_0)$. Now, assume by induction that we have proved (2.2) for all cubes $P \in \mathscr{D}_{jm}(P_0)$ with $1 \leq j_1 \leq j$ and let P be any cube in $\mathscr{D}_{(j_1-1)m}(P_0)$. By definition,

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) = \sum_{Q\in\mathscr{D}_m(P)} \left(\alpha_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_P\right) \mathbb{1}_Q(x) + \mathcal{A}^{m;0}_{Q,\alpha}\vec{f}(x)\right).$$

Let $x \in Q \in \mathscr{D}_m(P)$, then by the induction hypothesis and the definition of Δ_Q:

$$\begin{aligned}\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) &\leq \alpha_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_P\right) + \Delta_Q + \sum_{R\in\mathscr{D}(Q)} \beta_R \left(\prod_{i=1}^{k}\langle f_i\rangle_R\right) \mathbb{1}_R(x) \\ &= \alpha_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_P\right) + \Delta_P + (\beta_P - \alpha_Q)\left(\prod_{i=1}^{k}\langle f_i\rangle_P\right)\end{aligned}$$

[2] We use $\mathscr{D}_{\geq k}(P)$ to denote those cubes Q in $\mathscr{D}(P)$ of generation at least k, so $|Q| \leq 2^{-dk}|P|$.

$$
\begin{aligned}
&+\sum_{R\in\mathscr{D}(Q)}\beta_R\left(\prod_{i=1}^k\langle f_i\rangle_R\right)\mathbb{1}_R(x)\\
&=\Delta_P+\beta_P\left(\prod_{i=1}^k\langle f_i\rangle_P\right)+\sum_{R\in\mathscr{D}(Q)}\beta_R\left(\prod_{i=1}^k\langle f_i\rangle_R\right)\mathbb{1}_R(x)\\
&=\Delta_P+\sum_{R\in\mathscr{D}(P)}\beta_R\left(\prod_{i=1}^k\langle f_i\rangle_R\right)\mathbb{1}_R(x),
\end{aligned}
$$

which is what we wanted to show. □

Lemma 2.4 *The collection $\mathcal{S}$ is sparse.*

Proof Let $P\in\mathcal{S}$, we have to show that the set

$$F:=\bigcup_{Q\subsetneq P,\,Q\in\mathcal{S}}Q$$

satisfies $|F|\leq\frac{1}{2}|P|$. To this end, let $\mathcal{R}$ be the collection of maximal (strict) subcubes of P which are in $\mathcal{S}$, Note that for all $R\in\mathcal{R}$ we have $R\in\mathscr{D}_{N_R m}(P)$ for some $N_R\geq 1$. We thus have

$$F=\bigsqcup_{R\in\mathcal{R}}R.$$

By maximality, for all $R\in\mathcal{R}$ and dyadic cubes Q with $R\subsetneq Q\subsetneq P$ we have $\beta_Q=0$. For all $R\in\mathcal{R}$ and $1\leq j\leq N_R$ we now claim that

$$\Delta_{R^{((N_R-j)m)}}\geq\beta_P\left(\prod_{i=1}^k\langle f_i\rangle_P\right)-\sum_{\nu=1}^{j}\alpha_{R^{((N_R-\nu)m)}}\left(\prod_{i=1}^k\langle f_i\rangle_{R^{((N_R-\nu+1)m)}}\right).\qquad(2.3)$$

Indeed, one can prove this by induction on j. If $j=1$ then by definition we have

$$
\begin{aligned}
\Delta_{R^{((N_R-1)m)}}&=\Delta_P+(\beta_P-\alpha_{R^{((N_R-1)m)}})\left(\prod_{i=1}^k\langle f_i\rangle_P\right)\\
&\geq\beta_P\left(\prod_{i=1}^k\langle f_i\rangle_P\right)-\alpha_{R^{((N_R-1)m)}}\left(\prod_{i=1}^k\langle f_i\rangle_P\right),
\end{aligned}
$$

since $\Delta_P\geq 0$.

To prove the induction step, observe that (by the induction hypothesis) for $j>1$

$$\begin{aligned}\Delta_{R((N_R-j)m)} &= \Delta_{R((N_R-j+1)m)} + (\beta_{R((N_R-j+1)m)} - \alpha_{R((N_R-j)m)})\left(\prod_{i=1}^k \langle f_i\rangle_{R((N_R-j+1)m)}\right)\\ &= \Delta_{R((N_R-j+1)m)} - \alpha_{R((N_R-j)m)}\left(\prod_{i=1}^k \langle f_i\rangle_{R((N_R-j+1)m)}\right)\\ &\geq \beta_P\left(\prod_{i=1}^k \langle f_i\rangle_P\right) - \sum_{\nu=1}^{j} \alpha_{R((N_R-\nu)m)}\left(\prod_{i=1}^k \langle f_i\rangle_{R((N_R-\nu+1)m)}\right).\end{aligned}$$

From (2.3) with $j = N_R$, we have (since the terms are nonnegative)

$$\Delta_R \geq \beta_P\left(\prod_{i=1}^k \langle f_i\rangle_P\right) - \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x)$$

for all $x \in R$. Since $\beta_R \neq 0$, we must have

$$\left(\prod_{i=1}^k \langle f_i\rangle_R\right)\gamma_R - \Delta_R > 0,$$

i.e.:

$$\left(\prod_{i=1}^k \langle f_i\rangle_R\right)\gamma_R + \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) > 2^{2(k+1)}C_W\left(\prod_{i=1}^k \langle f_i\rangle_P\right)$$

for all $x \in R$. Let $\mathcal{G}_P\vec{f} = \sum_{R\in\mathcal{R}} \gamma_R \left(\prod_{i=1}^k \langle f_i\rangle_R\right)\mathbb{1}_R$, then for all $x \in R$ we have

$$\mathcal{G}_P f(x) + \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) > 2^{2(k+1)}C_W\left(\prod_{i=1}^k \langle f_i\rangle_P\right),$$

hence

$$\begin{aligned}|F| &\leq \left|\left\{x \in P : \mathcal{G}_P\vec{f}(x) + \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) > 2^{2(k+1)}C_W\left(\prod_{i=1}^k \langle f_i\rangle_P\right)\right\}\right|\\ &\leq \frac{\left\|\mathcal{G}_P + \mathcal{A}^{m;0}_{P,\alpha}\right\|^{1/k}_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}}{\left(2^{2(k+1)}C_W\left(\prod_{i=1}^k \langle f_i\rangle_P\right)\right)^{1/k}}\left(\prod_{i=1}^k \|f_i\|_{L^1(P)}\right)^{1/k}\\ &= \frac{\left\|\mathcal{G}_P + \mathcal{A}^{m;0}_{P,\alpha}\right\|^{1/k}_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}}{(2^{2(k+1)}C_W)^{1/k}}|P|\end{aligned}$$

Let us compute the operator norm $\|\mathcal{G}_P\|_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}$. Observe that, since $\gamma_Q \leq 1$ for all Q, the operator $\mathcal{G}$ is pointwise bounded by the multi-linear projection

$$\mathcal{P}_P\vec{f}(x) = \sum_{R\in\mathcal{R}}\left(\prod_{i=1}^k \langle f_i\rangle_R\right)\mathbb{1}_R(x) = \prod_{i=1}^k\left(\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R(x)\right).$$

For each $1 \leq i \leq k$, we have $\|\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R\|_{L_1(P)} \leq \|f_i\|_{L_1(P)}$. Therefore, by Hölder's inequality we get

$$\|\mathcal{P}_P\vec{f}\|_{L_{1/k,\infty}(P)} \leq \prod_{i=1}^k\left\|\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R\right\|_{L_1(P)} \leq \prod_{i=1}^k\|f_i\|_{L_1(P)}.$$

On the other hand we have

$$\|\mathcal{A}^{m;0}_{P,\alpha}\vec{f}\|_{L^{1/k,\infty}(P)} \leq C_W\prod_{i=1}^k\|f_i\|_{L^1(P)}$$

by Theorem 4.1. Combining these estimates we get

$$\|\mathcal{G}_P + \mathcal{A}^{m;0}_{P,\alpha}\|_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)} \leq 2^{k+1}(1+C_W) \leq 2^{k+2}C_W$$

and the result follows. □

From Lemmas 2.3 and 2.4 Proposition 2.1 follows at once. The proof shows that one can actually take $C_1 = 2^{2+k(7+d(2k-1))}$. We are now ready to finish the proof of Theorem A, which we state here in full generality:

Theorem 2.5 *Let α be a Carleson sequence and let P_0 be a dyadic cube. For every k-tuple of nonnegative integrable functions $f_1,\ldots,f_k$ on P there exists a sparse collection $\mathcal{S}$ of cubes in $\mathscr{D}(P)$ such that*

$$\mathcal{A}^m_{P,\alpha}\vec{f}(x) \leq C_2\sum_{Q\in\mathcal{S}}\left(\prod_{i=1}^k\langle f_i\rangle_Q\right)\mathbb{1}_Q(x).$$

Proof If $m = 0$ we can just apply Proposition 2.1 after noting that $\mathcal{A}^0_{P_0,\alpha}$ can be written as $\mathcal{A}^{1;0}_{P_0,\beta}$, where

$$\beta_Q = \alpha_{Q^{(1)}}.$$

One easily sees that $\|\alpha\|_{\mathrm{Car}(P_0)} = \|\beta\|_{\mathrm{Car}(P_0)}$. Hence, we may assume that $m \geq 1$. Recall the expression

$$\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) = \sum_{n=0}^{m-1} \sum_{P \in \mathscr{D}_n(P_0)} \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x).$$

from the beginning of the section. By Proposition 2.1, for each $0 \leq n \leq m-1$ and each $P \in \mathscr{D}_n(P_0)$ we can find a sparse collection of cubes $\mathcal{S}^n_P \subset \mathscr{D}(P)$ such that

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) \leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q \in \mathcal{S}^n_P} \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x).$$

Observe that the collection $\mathcal{S}^n = \cup_{P \in \mathscr{D}_n(P_0)} \mathcal{S}^n_P$ is also sparse, so

$$\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) \leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathcal{S}^n} \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x). \tag{2.4}$$

For $0 \leq n \leq m-1$ define

$$\mu^n_Q = \begin{cases} 1 & \text{if } Q \in \mathcal{S}^n \\ 0 & \text{otherwise.} \end{cases}$$

Since the collections $\mathcal{S}^n$ are sparse, the sequences μ^n are Carleson sequences with $\|\mu^n\|_{\mathrm{Car}(P_0)} \leq 2$, therefore the sequence

$$\mu_Q := \sum_{n=0}^{m-1} \mu^n_Q$$

is also Carleson with $\|\mu\|_{\mathrm{Car}(P_0)} \leq 2m$.

With this we can continue the argument using estimate (2.4) and the case $m = 0$:

$$\begin{aligned}
\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) &\leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathcal{S}^n} \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathscr{D}(P_0)} \mu^n_Q \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q \in \mathscr{D}(P_0)} \mu_Q \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \mathcal{A}^0_{P_0,\mu}\vec{f}(x) \\
&\leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} C_1 2m \sum_{Q \in \mathcal{S}} \left(\prod_{i=1}^k \langle f_i \rangle_Q\right) \mathbb{1}_Q(x),
\end{aligned}$$

which yields the result with $C_2 = 2C_1^2$. □

Remark 2.6 The above procedure does not rely on any specific property of the Lebesgue measure. In fact, Theorem A also holds when we replace all averages—both in complexity 0 and complexity m operators—by averages with respect to any other locally finite Borel measure, because the proof is unaffected.

We now detail how to use Theorem A to derive the multilinear version of Corollaries A.1 and A.2. For us, a multilinear Calderón-Zygmund operator will be an operator T satisfying

$$T(f_1,\dots,f_k)=\int_{\mathbb{R}^{dk}}K(x,y_1,\dots,y_k)f_1(y_1)\dots f_k(y_k)dy_1\,\dots\,dy_k$$

for all $x\notin\cap_{i=1}^k \operatorname{supp} f_i$ for appropriate f_i. Also we will require that T extends to a bounded operator from $L^{q_1}\times\cdots L^{q_k}$ to L^q where

$$\frac{1}{q}=\frac{1}{q_1}+\cdots+\frac{1}{q_k},$$

and that it satisfies the size estimate

$$|K(y_0,\dots,y_k)|\le\frac{A}{\left(\sum_{i,j=0}^k|y_i-y_j|\right)^{kd}}.$$

ω will be the modulus of continuity of the kernel of the operator i.e. a positive nondecreasing continuous and doubling function that satisfies

$$\begin{aligned}&|K(y_0,\dots,y_j,\dots,y_k)-K(y_0,\dots,y_j',\dots,y_k)|\\&\le C\omega\left(\frac{|y_j-y_j'|}{\sum_{i,j=0}^k|y_i-y_j|}\right)\frac{1}{\left(\sum_{i,j=0}^k|y_i-y_j|\right)^{kd}}\end{aligned}$$

for all $0\le j\le k$, whenever $|y_j-y_j'|\le\frac{1}{2}\max_{0\le i\le k}|y_j-y_i|$. We can now prove Corollary A.1:

Proof of Corollary A.1 Fix a measurable f, and a cube $Q_0\subset\mathbb{R}^d$. Our starting point is the formula

$$|T\vec{f}(x)-m_{T\vec{f}}(Q_0)|\lesssim\sum_{Q\in\mathcal{S}}\sum_{m=0}^{\infty}\omega(2^{-m})\prod_{i=1}^{m}\langle|f_i|\rangle_{2^mQ}\mathbb{1}_Q(x),$$

which holds for a sparse subcollection $\mathcal{S}\subset\mathscr{D}(Q_0)$ (see [5,10], we are implicitly using a slight improvement of Lerner's formula which can be found in [8, Theorem 2.3]). Here $m_f(Q)$ denotes the median of a measurable function f over a cube Q (see [16] for the precise definition), which satisfies

$$|m_f(Q)|\lesssim\frac{\|f\|_{L^{1,\infty}(Q)}}{|Q|}.$$

Hence we can just write

$$|T\vec{f}(x)| \lesssim \sum_{m=0}^{\infty} \omega(2^{-m}) \prod_{i=1}^{m} \langle |f_i| \rangle_{2^m Q} \mathbb{1}_Q(x), \tag{2.5}$$

By an elaboration of $\sum_{Q\in\mathcal{S}}$ the well-known one-third trick, it was proven in [10] that there exist dyadic systems $\{\mathscr{D}^\rho\}_{\rho\in\{0,1/3,2/3\}^d}$ such that for every cube Q in $\mathbb{R}^d$ and every $m \geq 1$, there exists $\rho \in \{0, 1/3, 2/3\}^d$ and $R_{Q,m} \in \mathscr{D}^\rho$ such that

$$Q \subset R_{Q,m},\ 2^m Q \subset Q^{(m)},\ 3\ell(Q) < \ell(R_{Q,m}) \leq 6\ell(Q).$$

Also, we may assume that for each $\rho \in \{0, 1/3, 2/3\}^d$ there exists a cube $P(\rho)$ such that $Q_0 \subset P(\rho) \subset c_d P(\rho)$ for some dimensional constant c_d. Using this, we can further write (2.5) as

$$|T\vec{f}(x)| \lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{Q\in\mathcal{S} R_{Q,m}\in\mathscr{D}^\rho} \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{R_{Q,m}^{(m)}} \right) \mathbb{1}_{R_Q}.$$

Let $\mathcal{F}_m^\rho = \{R_{Q,m} : R_Q \in \mathscr{D}^\rho\} \subset \mathscr{D}(P(\rho))$. Then, we can estimate

$$|T\vec{f}(x)| \lesssim 6^d \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{R\in\mathcal{F}_m^\rho} \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{R^{(m)}} \right) \mathbb{1}_R,$$

since at most 6^d cubes Q in $\mathscr{D}$ are mapped to the same cube $R_{Q,m}$. Define the sequence

$$\alpha_Q^\rho = \begin{cases} 1 & \text{if } Q \in \mathcal{F}_m^\rho \\ 0 & \text{otherwise.} \end{cases}$$

The collections $\mathcal{F}_m^\rho$ are $2^{-1} \cdot 6^{-d}$-sparse, and hence Carleson with constant $2 \cdot 6^d$. In order to apply Theorem A, for each fixed $\rho \in \left\{0, \frac{1}{3}, \frac{2}{3}\right\}^d$, $m \geq 0$, we now split the sum as follows:

$$\begin{aligned} \sum_{Q\in\mathscr{D}^\rho} \alpha_Q^\rho \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) &= \sum_{Q\in\mathscr{D}_{\geq m}(P(\rho))} \alpha_Q^\rho \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) \\ &\quad + \sum_{\ell=1}^{\infty} \sum_{Q\in\mathscr{D}_{m-\ell}(P(\rho))} \alpha_Q^\rho \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) \\ &= \mathrm{I} + \mathrm{II}. \end{aligned}$$

Now, since f_i is supported on $Q_0 \subset P(\rho)$ for $1 \leq i \leq k$ and all $\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$, we claim that II $\leq$ I. Indeed, compute

$$\sum_{\ell=1}^{\infty} \sum_{Q \in \mathscr{D}_{m-\ell}(P(\rho))} \alpha_Q^\rho \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x) \leq \sum_{\ell=1}^{\infty} \sum_{Q \in \mathscr{D}_{m-\ell}(P(\rho))} \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x)$$
$$= \sum_{\ell=1}^{\infty} \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{P(\rho)^{(\ell)}}\right).$$

Now observe that, by the support condition on the tuple $\vec{f}$,

$$\prod_{i=1}^{k} \langle |f_i| \rangle_{P(\rho)^{(\ell)}} = 2^{-dk\ell} \prod_{i=1}^{k} \langle |f_i| \rangle_{P(\rho)},$$

which is enough to prove the claim. Therefore, we only need to work in the localized cubes $P(\rho)$, $\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$. Therefore, we can obtain the first assertion of Corollary A.1 applying Theorem A:

$$|T\vec{f}(x)| \lesssim \sum_{\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{Q \in \mathscr{D}^\rho,\ Q \subset P(\rho)^{(m)}} \alpha_Q^\rho \left(\prod_{i=1}^{k} \langle |f_i| \rangle_{Q^{(m)}}\right) \mathbb{1}_Q(x)$$
$$\lesssim \sum_{\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \sum_{Q \in \mathcal{S}_{m,\vec{f}}} \left(\prod_{i=1}^{k} \langle |f_i| \rangle_Q\right) \mathbb{1}_Q$$
$$= \sum_{\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \mathcal{A}_{\mathcal{S}_{m,\vec{f}}} \vec{f}(x),$$

for sparse collections $\mathcal{S}_{m,\vec{f}}$ that may depend both on m and $\vec{f}$ (and which are subfamilies of $\mathscr{D}(P(\rho))$ for each value of ρ). Now, reorganizing the sum above we obtain

$$|T\vec{f}(x)| \lesssim \sum_{\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} \sum_{\mathcal{S}_{m,\vec{f}} \subset \mathscr{D}^\rho} \omega(2^{-m})(m+1) \mathcal{A}_{\mathcal{S}_{m,\vec{f}}} \vec{f}(x)$$
$$=: \sum_{\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d} \mathcal{A}_\rho \vec{f}(x).$$

Now, by the logarithmic Dini condition, each of the operators $\mathcal{A}_\rho$ is bounded above by some absolute constant times a 0-shift whose associated sequence is 1-Carleson (and localized in $P(\rho)$) to which we can apply again Theorem A. Therefore, we obtain

$$|T\vec{f}(x)| \lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \mathcal{A}_{\mathcal{S}_\rho}\vec{f}(x),$$

for some sparse families $\mathcal{S}_\rho \subset \mathscr{D}^\rho$ which depend on $\vec{f}$. □

We now introduce the notion of function quasi-norm. We say that $\|\cdot\|_{\mathbb{X}}$, defined on the set of measurable functions, is a function quasi-norm if:

(P1): There exists a constant $C > 0$ such that

$$\|f+g\|_{\mathbb{X}} \le C\left(\|f\|_{\mathbb{X}} + \|g\|_{\mathbb{X}}\right),$$

(P2): $\|\lambda f\|_{\mathbb{X}} = |\lambda|\|f\|_{\mathbb{X}}$ for all $\lambda \in \mathbb{C}$.
(P3): If $|f(x)| \le |g(x)|$ almost-everywhere then $\|f\|_{\mathbb{X}} \le \|g\|_{\mathbb{X}}$.
(P4): $\|\liminf_{n\to\infty} f_n\|_{\mathbb{X}} \le \liminf_{n\to\infty}\|f_n\|_{\mathbb{X}}$

Fix some dyadic system $\mathscr{D}$ such that there exists an increasing sequence of dyadic cubes $\{P_\ell\}_\ell \subset \mathscr{D}$ whose union is the whole space $\mathbb{R}^d$, and denote $\mathbb{1}_{P_\ell}\vec{f} = (\mathbb{1}_{P_\ell} f_1, \ldots, \mathbb{1}_{P_\ell} f_k)$. Now, taking into account properties (P1) and (P3), if we take quasi-norms in the second assertion of Corollary A.1, we have

$$\|\mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell}\vec{f})\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}_{\mathcal{S}}(\mathbb{1}_{P_\ell}\vec{f})\|_{\mathbb{X}}\ \forall \ell.$$

On the one hand, since $\vec{f}$ is integrable, $T(\mathbb{1}_{P_\ell}\vec{f})$ converges pointwise to $T(\vec{f})$. Therefore, we have

$$\mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell}\vec{f}) \to T(\vec{f})$$

pointwise. Finally, we apply property (P4) and we get

$$\|T\vec{f}\|_{\mathbb{X}} = \left\|\liminf_{\ell} \mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell}\vec{f})\right\|_{\mathbb{X}} \le \liminf_{\ell}\left\|\mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell}\vec{f})\right\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}}\left\|\mathcal{A}_{\mathcal{S}}\vec{f}\right\|_{\mathbb{X}}.$$

This is exactly Corollary A.2.

Remark 2.7 We note that the dependence on m in the pointwise estimate of shifts of complexity m must be at least linear in m. To see this, let us work in dimension one and fix a large integer m. For any interval $I = [a, b)$ let I_j be the j-th interval of $\mathscr{D}_m(I)$:

$$I_j = a + |I|[j2^{-m}, (j+1)2^{-m}).$$

Define a tower over an interval I to be the collection of intervals

$$\mathcal{T}_I = \left\{[a, a+2^{-k}|I|) : k \in \mathbb{N}\right\}.$$

The collection of intervals $\mathcal{S} = \bigcup_{J\in\mathscr{D}_m(I)} \mathcal{T}_J$ is a sparse collection. Now consider a function f on I which is defined by

$$f(x) = \begin{cases} 0 & \text{if } x \in I_j \text{ with } j \text{ even}, \\ 2 & \text{otherwise}. \end{cases}$$

Denote $\operatorname{gen}(J) = \log_2(\ell(I)\ell(J)^{-1})$ for cubes $J \in \mathscr{D}(I)$. Observe that for any dyadic interval $J \subseteq I$ with $\operatorname{gen}(J) \leq m - 1$ we have

$$\langle f \rangle_J = 1.$$

Consider now the action of $\mathcal{A}_{\mathcal{S}}^m$ on f. If $x \in (I_j)_0$ with j even then

$$\mathcal{A}_{\mathcal{S}}^m f(x) = m.$$

In order to construct a collection $\mathcal{S}'$ of intervals in I for which we have

$$\mathcal{A}_{\mathcal{S}}^m f(x) \leq C \mathcal{A}_{\mathcal{S}'}^0 f(x),$$

we would need to select every interval $J \subset I$ with $\operatorname{gen}(J) \geq m - 1$. Indeed, let $I^k(x)$ be the interval in $\mathscr{D}_k(I)$ which contains x and let α_J be 1 if $J \in \mathcal{S}'$ and 0 otherwise. Then

$$C \mathcal{A}_{\mathcal{S}'}^0 f(x) = C \sum_{k=0}^{m-1} \alpha_{I^k(x)} \geq m$$

for all $x \in (I_j)_0$ with j even. This implies that at least m/C of these intervals must be in $\mathcal{S}'$. But this implies that the height

$$\sum_{J \in \mathcal{S}'} \alpha_J \mathbb{1}_J(x) \geq m/C$$

on half of the interval I, which contradicts the hypothesis of $\mathcal{S}'$ being sparse if m is large enough.

3 Applications

We are now ready to fully state and prove the applications of the pointwise bound as stated in the introduction. We begin with the multilinear sharp weighted estimates:

3.1 Multilinear A_2 theorem

We need some more definitions first. These were introduced in [20].

Definition 3.1 ($A_{\vec{P}}$ *weights*) Let $\vec{P} = (p_1, \ldots, p_k)$ with $1 \leq p_1, \ldots, p_k < \infty$ and $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$. Given $\vec{w} = (w_1, \ldots, w_k)$, set

$$v_{\vec{w}} = \prod_{i=1}^{k} w_i^{p/p_i}.$$

We say that $\vec{w}$ satisfies the k-linear $A_{\vec{P}}$ condition if

$$[\vec{w}]_{A_{\vec{P}}} := \sup_Q \left(\frac{1}{|Q|}\int_Q v_{\vec{w}}\right) \prod_{i=1}^{k} \left(\frac{1}{|Q|}\int_Q w_i^{1-p_i'}\right)^{p/p_i}.$$

We call $[\vec{w}]_{A_{\vec{P}}}$ the $A_{\vec{P}}$ constant of $\vec{w}$. As usual, if $p_i = 1$ then we interpret $\frac{1}{|Q|}\int_Q w_i^{1-p_i'}$ to be $(\operatorname{essinf}_Q w_i)^{-1}$.

The following theorem was proved in [21]:

Theorem 3.2 *Suppose* $1 < p_1, \dots, p_k < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$ *and* $\vec{w} \in A_{\vec{P}}$. *Then*

$$\|\mathcal{A}_S \vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [w]_{A_{\vec{P}}}^{\max\left(1, \frac{p_1'}{p}, \dots, \frac{p_k'}{p}\right)} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)},$$

whenever $\mathcal{S}$ *is sparse.*

We can now use Corollary A.2 to extend the above result to general k-linear Calderón-Zygmund operators:

Theorem 3.3 *Under the conditions of Theorem 3.2, for any* k*-linear Calderón-Zygmund operator* T*, we have*

$$\|T \vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [\vec{w}]_{A_{\vec{P}}}^{\max\left(1, \frac{p_1'}{p}, \dots, \frac{p_k'}{p}\right)} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)}.$$

Proof We just need to apply Corollary A.2 with $\|\cdot\|_{\mathbb{X}} := \|\cdot\|_{L^p(v_{\vec{w}})}$, which clearly is a function quasi-norm. The assumption of $\vec{f}$ being integrable is a qualitative one and can be trivially removed by the usual density arguments. □

3.2 Sharp aperture weighted Littlewood-Paley theorem

Here we follow Lerner [17], the reader can find a nice introduction and some references there. We begin with some definitions:

Let $\psi \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \psi(x)\,dx = 0$ satisfy

$$|\psi(x)| \lesssim \frac{1}{(1+|x|)^{d+\epsilon}} \tag{3.1}$$

$$\int_{\mathbb{R}^d} |\psi(x+h) - \psi(x)|\,dx \lesssim |h|^{\epsilon}. \tag{3.2}$$

We will denote the upper half-space $\mathbb{R}^d \times \mathbb{R}$ by $\mathbb{R}^{d+1}_+$ and the α-cone at x by

$$\Gamma_\alpha(x) = \left\{(y,t) \in \mathbb{R}^{d+1}_+ : |y-x| \le \alpha t\right\}.$$

Let ψ_t be the dilation of ψ which preserves the L^1 norm, i.e.: $\psi_t(x) = t^{-d}\psi(x/t)$, then we can define the square function $S_{\alpha,\psi} f$ by

$$S_{\alpha,\psi} f(x) = \left(\int_{\Gamma_\alpha(x)} |(f * \psi_t)(y)|^2 \frac{dy\,dt}{t^{d+1}}\right)^{1/2}.$$

We will also need a regularized version. Let Φ be a Schwartz function such that

$$\mathbb{1}_{B(0,1)}(x) \le \Phi(x) \le \mathbb{1}_{B(0,2)}(x).$$

We define the regularized square function $\widetilde{S}_{\alpha,\psi}$ by

$$\widetilde{S}_{\alpha,\psi} f(x) = \left(\int_{\mathbb{R}^{d+1}_+} \Phi\left(\frac{x-y}{t\alpha}\right) |(f * \psi_t)(y)|^2 \frac{dy\,dt}{t^{d+1}}\right)^{1/2}.$$

The regularized version can be used instead of $S_{\alpha,\psi}$ in most cases since we have

$$S_{\alpha,\psi} f(x) \le \widetilde{S}_{\alpha,\psi} f(x) \le S_{2\alpha,\psi} f(x).$$

It was proved in [17] that

$$|(\widetilde{S}_{\alpha,\psi} f(x))^2 - (m_{Q_0}(\widetilde{S}_{\alpha,\psi} f)^2)| \lesssim \alpha^{2d} \sum_{m=0}^{\infty} 2^{-\delta m} \sum_{Q \in \mathcal{S}} \langle |f| \rangle^2_{2^m Q} \mathbb{1}_Q(x)$$

By the same Theorem A in its bilinear formulation (with $f_1 = f_2 = f$), the last expression can be bounded, up to a constant, by an expression of the form

$$\alpha^{2d} \sum_{\rho \in \{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} 2^{-\delta m}(m+1) \sum_{Q \in \mathcal{S}^{\rho,m}} \langle |f| \rangle^2_Q \mathbb{1}_Q(x).$$

As in [17], we know (a priori) that $m_{Q_0}(\widetilde{S}_{\alpha,\psi} f) \to 0$ as $|Q| \to \infty$ so by the triangle inequality and Fatou's Lemma we can ignore that term (or by arguing as we did in the previous section). Finally, arguing as in the proof of Corollaries A.1 and A.2, we arrive at

$$\|\widetilde{S}_{\alpha,\psi} f\|_{L^{p,\infty}(w)} \lesssim \alpha^d \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}^0_{\mathcal{S}}(f,f)^{1/2}\|_{L^{p,\infty}(w)},$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collections $\mathcal{S} \subset \mathscr{D}$. To finish the argument we recall the following result, which was shown in [13]:

$$\|\mathcal{A}^0_{\mathcal{S}}(f, f)^{1/2}\|_{L^{p,\infty}(w)} \lesssim [w]_{A_p}^{\max(\frac{1}{2},\frac{1}{p})} \Phi_p([w]_{A_p})\|f\|_{L^p(w)} \tag{3.3}$$

for $1 < p < 3$, where

$$\Phi_p(t) = \begin{cases} 1 & \text{if } 1 < p < 2 \\ 1 + \log t & \text{if } 2 \le p < 3. \end{cases}$$

We are thus able to extend Lerner's estimate to $1 < p \le 2$, obtaining

$$\|S_{\alpha,\psi} f\|_{L^{p,\infty}(w)} \lesssim \alpha^d [w]_{A_p}^{1/p} \|f\|_{L^p(w)} \quad \text{for } 1 < p < 2$$

and

$$\|S_{\alpha,\psi} f\|_{L^{2,\infty}(w)} \lesssim \alpha^d [w]_{A_2}^{1/2}(1 + \log[w]_{A_2})\|f\|_{L^2(w)}.$$

Acknowledgments The authors wish to thank Javier Parcet, Ignacio Uriarte-Tuero and Alexander Volberg for insightful discussions, and Andrei Lerner and Fedor Nazarov for sharing with us the details of their construction.

Appendix: The weak-type estimate for multilinear *m*-shifts

Here we prove the weak-type estimate for k-linear m-shifts needed in Sect. 2. Notice that the only important point of the estimates below is the independenceof the constants from the parameter m. The proof could be more or less standard by now, but the authors have not been able to find it elsewhere. Therefore we include it for completeness.

Theorem 4.1

$$\sup_{\lambda>0} \lambda \left| \left\{ x \in P_0 : \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) > \lambda \right\} \right|^k \le C_W \|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^k \|f_i\|_{L^1(P_0)}, \tag{4.1}$$

where $C_W > 0$ only depends on k and d, and in particular is independent of m.

We will essentially follow Grafakos-Torres [6,9]. We first prove an L^2 bound and then apply a Calderón-Zygmund decomposition. For the L^2 bound we will use a multilinear Carleson embedding theorem by Chen and Damián [2], from which we only need the unweighted result:

$$\left(\sum_{Q \in \mathscr{D}(P_0)} \alpha_Q \left(\prod_{i=1}^k \langle f_i \rangle_Q \right)^p \right)^{\frac{1}{p}} \le \|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^k p_i' \|f_i\|_{L^{p_i}(P_0)} \tag{4.2}$$

whenever

$$\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}.$$

Now we can prove

Proposition 4.2

$$\|\mathcal{A}^m_{P_0,\alpha}\vec{f}\|_{L^2(P_0)} \leq 4\|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^{k} \|f_i\|_{L^{2k}(P_0)}$$

Proof We begin by using duality and homogeneity to reduce to showing

$$\int_{P_0} g(x)\mathcal{A}^m_{P_0,\alpha}\vec{f}(x)\,dx \leq 4$$

assuming that $\|f_i\|_{L^{2k}(P_0)} = \|g\|_{L^2(P_0)} = \|\alpha\|_{\mathrm{Car}(P_0)} = 1$ and $g \geq 0$. By definition and Cauchy-Schwarz, this is equivalent to

$$\left(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \left(\prod_{i=1}^{k}\langle f_i\rangle_{Q^{(m)}}\right)^2 |Q|\right)^{1/2} \left(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle g\rangle_Q^2 |Q|\right)^{1/2}.$$

The second term can be estimated, using (4.2) in the linear case, by

$$\left(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle g\rangle_Q^2 |Q|\right)^{1/2} \leq 2.$$

For the first term observe that the sequence β_Q defined by

$$\beta_Q = \frac{1}{2^{dm}} \sum_{R\in\mathscr{D}_m(Q)} \alpha_R$$

is a Carleson sequence adapted to P_0 of the same constant. Indeed:

$$\begin{aligned}
\frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)} \beta_R|R| &= \frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)} |R|\frac{1}{2^{dm}}\sum_{T\in\mathscr{D}_m(R)} \alpha_T \\
&= \frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)}\sum_{T\in\mathscr{D}_m(R)} \alpha_T|T| \\
&= \frac{1}{|Q|}\sum_{R\in\mathscr{D}_{\geq m}(Q)} \alpha_R|R| \\
&\leq \|\alpha\|_{\mathrm{Car}(I)} \\
&= 1.
\end{aligned}$$

Therefore, we can write the first term as

$$\left(\sum_{Q\in\mathscr{D}(P_0)} \beta_Q \left(\sum_{i=1}^{k} \langle f_i\rangle_Q\right)^2 |Q|\right)^{1/2},$$

which can also be estimated by (4.2) as follows:

$$\left(\sum_{Q\in\mathscr{D}(P_0)} \beta_Q \left(\sum_{i=1}^{k} \langle f_i\rangle_Q\right)^2 |Q|\right)^{1/2} \leq \left(\frac{2k}{2k-1}\right)^k \leq 2.$$

Combining both terms we arrive at

$$\int_{P_0} g(x)\mathcal{A}^m_{P_0,\alpha}\vec{f}(x)\,dx \leq 4$$

which is what we wanted. □

Now we can prove Theorem 4.1.

Proof By homogeneity we can assume $\|\alpha\|_{\mathrm{Car}(P_0)} = \|f_i\|_{L^1(P_0)} = 1$. We now follow the classical scheme which uses the L^2 bound and a standard Calderón-Zygmund decomposition, see for example Grafakos-Torres [6]. However, we need to be careful with the dependence on m, so we will adapt the proof in [9] to our operators.

Assume without loss of generality that $f_i \geq 0$. Define

$$\Omega_i = \left\{x \in P_0 : \mathcal{M}^d f_i(x) > \lambda^{1/k}\right\}.$$

If $\langle f_i\rangle_{P_0} > \lambda^{1/k}$ then by the homogeneity assumption

$$|P_0| < \lambda^{-1/k}$$

and the estimate follows. Therefore, we can assume $\langle f_i\rangle_{P_0} \leq \lambda^{1/k}$ for all $1 \leq i \leq k$ and hence we can write Ω_i as a union the cubes in a collection $\mathcal{R}_i$ consisting of pairwise disjoint dyadic (strict) subcubes of P_0 with the property

$$\langle f_i\rangle_R > \lambda^{1/k} \quad \text{and} \quad \langle f_i\rangle_{R^{(1)}} \leq \lambda^{1/k}.$$

For each $1 \leq i \leq k$ let $b_i = \sum_{R\in\mathcal{R}_i} b_i^R$, where

$$b_i^R(x) := (f_i(x) - \langle f_i\rangle_R)\,\mathbb{1}_R(x).$$

We now let $g_i = f_i - b_i$.

Observe that we have

$$|g_i(x)| \leq 2^d \lambda^{1/k},$$

as well as

$$|\Omega_i| = \sum_{R \in \mathcal{R}_i} |R| \leq \lambda^{-1/k}.$$

Define $\Omega = \cup_{i=1}^k \Omega_i$, then we have

$$\begin{aligned} |\{x \in P_0 : \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) > \lambda\}| &\leq |\Omega| + |\{x \in P_0 \backslash \Omega : \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) > \lambda\}| \\ &\leq k\lambda^{-1/k} + |\{x \in P_0 \backslash \Omega : \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) > \lambda\}|. \end{aligned} \tag{4.3}$$

To estimate the second term observe that

$$\begin{aligned} \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) &= \mathcal{A}^m_{P_0,\alpha}(\vec{g} + \vec{b})(x) \\ &= \mathcal{A}^m_{P_0,\alpha} \vec{g}(x) + \sum_{j=1}^{2^k-1} \mathcal{A}^m_{P_0,\alpha}(h^j_1, \ldots, h^j_k)(x), \end{aligned}$$

where the functions h^j_i are either g_i or b_i and, furthermore, for each $1 \leq j \leq 2^k - 1$ there is at least one $1 \leq i \leq k$ such that $h^j_i = b_i$. Fix j and let i_j be such that $h^j_{i_j} = b_{i_j}$, then

$$\begin{aligned} &\mathcal{A}^m_{P_0}(h^j_1, h^j_2, \ldots, h^j_{i_j}, \ldots, h^j_k)(x) \\ &= \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \left(\prod_{i=1}^k \langle h^j_i \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) \\ &= \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle b_{i_j} \rangle_{Q^{(m)}} \left(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h^j_i \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) \\ &= \sum_{R \in \mathcal{R}_{i_j}} \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle b^R_{i_j} \rangle_{Q^{(m)}} \left(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h^j_i \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x) \\ &= \sum_{R \in \mathcal{R}_{i_j}} \sum_{Q \in \mathscr{D}_{>m}(R)} \alpha_Q \langle b^R_{i_j} \rangle_{Q^{(m)}} \left(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h^j_i \rangle_{Q^{(m)}} \right) \mathbb{1}_Q(x). \end{aligned}$$

So we deduce that $\mathcal{A}^m_{P_0,\alpha}(h^j_1, \ldots, h^j_k)(x) = 0$ for all $x \notin \Omega_{i_j}$. With this fact we can see that the second term in (4.3) is actually identical to

$$|\{x \in P_0 \backslash \Omega : \mathcal{A}^m_{P_0,\alpha} \vec{g}(x) > \lambda\}|.$$

Now we can use the L^2 bound as follows:

$$\begin{aligned}
|\{x \in P_0 \backslash \Omega : \mathcal{A}^m_{P_0,\alpha} \vec{g}(x) > \lambda\}| &\leq \frac{1}{\lambda^2} \|\mathcal{A}^m_{P_0,\alpha} \vec{g}\|^2_{L^2(P_0)} \\
&\leq \frac{16}{\lambda^2} \prod_{i=1}^{k} \|g_i\|^2_{L^{2k}(P_0)} \\
&\leq \frac{16}{\lambda^2} \prod_{i=1}^{k} \left(2^d \lambda^{1/k}\right)^{\frac{2k-1}{k}} \|g_i\|^{1/k}_{L^1(P_0)} \\
&= \frac{16}{\lambda^2} 2^{d(2k-1)} \lambda^{2-1/k} \\
&= 2^{4+d(2k-1)} \lambda^{-1/k}.
\end{aligned}$$

Putting both estimates together we arrive at

$$|\{x \in P_0 : \mathcal{A}^m_{P_0,\alpha} \vec{f}(x) > \lambda\}| \leq 2^{5+d(2k-1)} \lambda^{-1/k}$$

which yields the result with $C_W = 2^{k(5+d(2k-1))}$. □

References

1. Bennett, C., Sharpley, R.: Interpolation of operators. Pure and Applied Mathematics, vol. 129. Academic Press Inc., Boston (1988)
2. Chen, W., Damián, W.: Weighted estimates for the multisublinear maximal function. Rend. Circ. Mat. Palermo (2), **62**(3), 379–391 (2013)
3. Cruz-Uribe, D., Martell, J.M., Pérez, C.: Sharp weighted estimates for approximating dyadic operators. Electron. Res. Announc. Math. Sci. **17**, 12–19 (2010)
4. Cruz-Uribe, D., Martell, J.M., Pérez, C.: Sharp weighted estimates for classical operators. Adv. Math. **229**(1), 408–441 (2012)
5. Damián, W., Lerner, A.K., Pérez, C.: Sharp weighted bounds for multilinear maximal functions and Calderón-Zygmund operators (2012). arXiv:1211.5115
6. Grafakos, L., Torres, R.H.: Multilinear Calderón-Zygmund theory. Adv. Math. **165**(1), 124–164 (2002)
7. Hänninen, T.S.: Remark on dyadic pointwise domination and median oscillation decomposition (2015,ArXiv e-prints)
8. Hytönen, T.: A_2 theorem: remarks and complements (2012, preprint)
9. Hytönen, T.P.: The sharp weighted bound for general Calderón-Zygmund operators. Ann. of Math. (2). **175**(3), 1473–1506 (2012)
10. Hytönen, T.P., Lacey, M.T., Pérez, C.: Sharp weighted bounds for the q-variation of singular integrals. Bull. Lond. Math. Soc. **45**(3), 529–540 (2013)
11. Lacey, M.T.: Weighted weak type estimates for square functions (2015, ArXiv e-prints)
12. Lacey, M.T., Sawyer, E.T., Uriarte-Tuero, I.: Two weight inequalities for discrete positive operators (2009, ArXiv e-prints)
13. Lacey, M.T., Scurry, J.: Weighted weak type estimates for square functions (2012, ArXiv e-prints)
14. Lerner, A.K.: A pointwise estimate for the local sharp maximal function with applications to singular integrals. Bull. Lond. Math. Soc. **42**(5), 843–856 (2010)
15. Lerner, A.K.: On an estimate of Calderón-Zygmund operators by dyadic positive operators. J. Anal. Math. **121**, 141–161 (2013)
16. Lerner, A.K.: A simple proof of the A_2 conjecture. Int. Math. Res. Not. IMRN **90**(14), 3159–3170 (2013)

17. Lerner, A.K.: On sharp aperture-weighted estimates for square functions. J. Fourier Anal. Appl. **20**(4), 784–800 (2014)
18. Lerner, A.K., Nazarov, F.: Intuitive dyadic calculus: the basics (2015). arXiv:1508.05639
19. Lerner, A.K., Ombrosi, S., Pérez, C.: A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden. Math. Res. Lett. **16**(1), 149–156 (2009)
20. Lerner, A.K., Ombrosi, S., Pérez, C., Torres, R.H., Trujillo-González, R.: New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory. Adv. Math. **220**(4), 1222–1264 (2009)
21. Li, K., Moen, K., Sun, W.: The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators. J. Fourier Anal. Appl. **20**(4), 751–765 (2014)
22. Nazarov, F., Reznikov, A., Vasyunin, V., Volberg, A.: A_1 conjecture: weak norm estimates of weighted singular integral operators and Bellman functions (2013, Preprint)
23. Petermichl, S.: The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. Am. J. Math. **129**(5), 1355–1375 (2007)