# WEIGHTED ESTIMATES
# FOR THE MULTILINEAR MAXIMAL OPERATOR

ADAM OSĘKOWSKI

ABSTRACT. The paper contains the study of the weighted $L^{p_1} \times L^{p_2} \times \ldots \times L^{p_m} \to L^p$ estimates for the multilinear maximal operator, in the context of abstract probability spaces equipped with a tree-like structure. Using the Bellman function method, we identify the associated optimal constants in the symmetric case $p_1 = p_2 = \ldots = p_m$, and a tight constant for remaining choices of exponents.

## 1. INTRODUCTION

The purpose of this paper is to study a class of weighted inequalities for the multilinear maximal operator, an important object in harmonic analysis. We will be particularly interested in the size of the constants involved. To present the results from an appropriate perspective, let us start with the necessary background. Fix a dimension $N$ and suppose that $\mathscr{D}$ is the standard dyadic lattice in $\mathbb{R}^N$. Let $M$ be the dyadic maximal operator, acting on locally integrable functions $f$ on $\mathbb{R}^N$ by

$$Mf(x) = \sup_{Q \in \mathscr{D}} \langle |f| \rangle_Q \chi_Q.$$

Here the symbol $\langle \varphi \rangle_Q$ stands for the average of $\varphi$ over $Q$, calculated with respect to the Lebesgue measure: $\langle \varphi \rangle_Q = \frac{1}{|Q|} \int_Q \varphi$. Maximal operators play a prominent role in many areas of mathematics, and from the viewpoint of applications, it is often of interest to study the boundedness properties of these objects, treated as operators on various function spaces. A fundamental example is the sharp estimate

$$(1.1) \qquad \|Mf\|_{L^p} \le p' \|f\|_{L^p}, \qquad 1 < p \le \infty$$

(here and below, $p'$ denotes the conjugate exponent to $p$, given by $p' = p/(p-1)$). One can investigate numerous generalizations of this result, our motivation comes from an extension to the weighted theory. In what follows, the word 'weight' refers to a positive and locally integrable function on $\mathbb{R}^N$. Any weight $w$ gives rise to the corresponding Borel measure, also denoted by $w$, and given by $w(A) = \int_A w \mathrm{d}x$. One can also introduce the associated weighted $L^p$ spaces, defined as the classes of all (equivalence classes of) functions $f : \mathbb{R}^N \to \mathbb{R}$ for which

$$\|f\|_{L^p(w)} := \left( \int_{\mathbb{R}^N} |f|^p w \mathrm{d}x \right)^{1/p} < \infty, \qquad 0 < p < \infty.$$

Now one can study the following problem related to (1.1): given $1 < p < \infty$, characterize those weights $w$, for which the dyadic maximal operator $M$ is bounded as an operator on $L^p(w)$. This problem was solved by Muckenhoupt [8] in the

seventies: the boundedness is true if and only if $w$ satisfies the so-called dyadic $A_p$ condition (or $w$ belongs to the dyadic $A_p$ class). The latter means that the quantity

$$[w]_{A_p} := \sup_{Q \in \mathscr{D}} \langle w \rangle_Q \langle \sigma \rangle_Q^{p-1},$$

called the $A_p$ characteristic of $w$, is finite. Here $\sigma = w^{1-p'}$ is the dual weight to $w$. There is a stronger, quantitative version of this result, established by Buckley at the beginning of the nineties. Namely, the problem is to find, for any fixed $1 < p < \infty$, the least exponent $\beta(p)$ such that

(1.2) $$\|M\|_{L^p(w) \to L^p(w)} \leq \kappa_p [w]_{A_p}^{\beta(p)},$$

where $\kappa_p$ depends only on $p$. In other words, one is interested in the extraction of the optimal dependence of $\|M\|_{L^p(w) \to L^p(w)}$ on the $A_p$ characteristic $[w]_{A_p}$. The aforementioned result of Buckley [1] asserts that the optimal choice is given by $\beta(p) = (p-1)^{-1}$. This statement can be further improved: we need some additional notation for the further discussion. Given $1 < p < \infty$ and $c \geq 1$, let $d = d(p, c)$ be the unique positive root of the equation

(1.3) $$c(1 + d)(p - 1 - d)^{p-1} = (p - 1)^{p-1}.$$

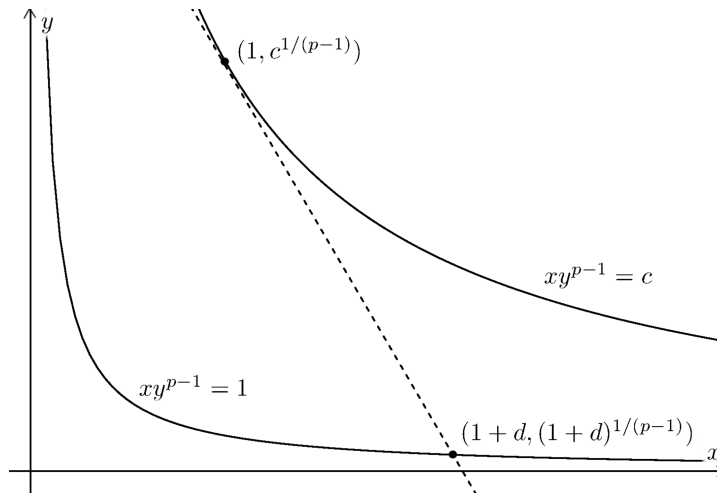This parameter has a nice geometric interpretation: see Figure 1 below.



FIGURE 1. The tangent to the curve $xy^{p-1} = c$ at the point $(1, c^{1/(p-1)})$ intersects the curve $xy^{p-1} = 1$ at $(1+d, (1+d)^{1/(p-1)})$.

The paper [9] contains the proof of the following extension of Buckley's result.

**Theorem 1.1.** *For any $1 < p < \infty$ we have the inequality*

$$\|M\|_{L^p(w) \to L^p(w)} \leq \frac{p}{p - 1 - d(p, [w]_{A_p})}.$$

*The constant on the right is the best possible: for any $c \geq 1$ and $\varepsilon > 0$ there is an $A_p$ weight $w$ satisfying $[w]_{A_p} \leq c$ such that*

$$\|M\|_{L^p(w) \to L^p(w)} \geq \frac{p}{p - 1 - d(p, c)} - \varepsilon.$$

In our considerations below, we will be interested in the multilinear analogues of the above theorem. Suppose that $m \geq 1$ is a fixed integer. Following [5], the $m$-linear dyadic maximal operator $\mathcal{M}$ acts on vectors $\vec{f} = (f_1, f_2, \ldots, f_m)$ of locally integrable functions on $\mathbb{R}^N$ by the formula

$$\mathcal{M}\vec{f} = \sup_{Q \in \mathscr{D}} \prod_{j=1}^{m} \langle |f_j| \rangle_Q \chi_Q.$$

A straightforward combination of (1.1) and the Hölder inequality implies that if $\vec{P} = (p_1, p_2, \ldots, p_m)$ is a sequence of exponents satisfying $1 < p_1, p_2, \ldots, p_m \leq \infty$ and $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_m}$, then we have the sharp bound

$$\|\mathcal{M}\vec{f}\|_{L^p} \leq \prod_{j=1}^{m} p_j' \|f_j\|_{L^{p_j}}.$$

One can ask about the weighted version of this estimate, motivated by the above discussion for $M$. The precise answer was given by Lerner et. al. in [5]. Given a vector $\vec{w} = (w_1, w_2, \ldots, w_m)$ of weights on $\mathbb{R}^N$, we set

$$v_{\vec{w}} = \prod_{j=1}^{m} w_j^{p/p_j}$$

and say that $\vec{w}$ satisfies the dyadic multilinear condition $A_{\vec{P}}$, if

$$[\vec{w}]_{A_{\vec{P}}} := \sup_{Q \in \mathscr{D}} \langle v_{\vec{w}} \rangle_Q \prod_{j=1}^{m} \langle \sigma_j \rangle_Q^{p/p_j'} < \infty.$$

Here, as before, $\sigma_j = w_j^{1-p_j'}$ is the dual weight to $w_j$, $j = 1, 2, \ldots, m$. Note that for $m = 1$, the above requirement reduces to the classical Muckenhoupt's condition $A_p$. Actually, the connection to the one-dimensional setting goes deeper: as proved in [5], we have $\vec{w} \in A_{\vec{P}}$ if and only if $\sigma_j \in A_{mp_j'}$ for all $j$ and $v_{\vec{w}} \in A_{mp}$. The multilinear $A_{\vec{P}}$ condition provides us with the answer to the above question: $\mathcal{M}$ is bounded as an operator from $L^{p_1}(w_1) \times L^{p_2}(w_2) \times \ldots \times L^{p_m}(w_m)$ to $L^p(v_{\vec{w}})$ if and only if $\vec{w} \in A_{\vec{P}}$. In [3], the authors established the following version of (1.2):

$$(1.4) \qquad \|\mathcal{M}\vec{f}\|_{L^p(v_{\vec{w}})} \leq \kappa_{\vec{P}} [\vec{w}]_{A_{\vec{P}}}^{\beta(\vec{P})} \prod_{j=1}^{m} \|f_j\|_{L^{p_j}(w_j)},$$

with $\beta(\vec{P})$ satisfying $\frac{m}{mp-1} \leq \beta(\vec{P}) \leq \frac{1}{p}\left(1 + \sum_{j=1}^{m} \frac{1}{p_j - 1}\right)$. Furthermore, in the special case $p_1 = p_2 = \ldots = p_m = mp$, they showed that the estimate holds with $\beta(\vec{P}) = m/(mp - 1) = p_1'/p$, which is optimal. The question about the best exponent for an arbitrary vector $\vec{P}$ was answered in [6]: the optimal choice is $\beta(\vec{P}) = \max\{p_1'/p, p_2'/p, \ldots, p_m'/p\}$.

The purpose of this paper is to present a different approach to the estimate (1.4), which will yield the further information about the size of the constant $\kappa_{\vec{P}}$. In addition, it will allow us to identify the best constant in the special case $p_1 = p_2 = \ldots = p_m = mp$. Actually, we will work in the more general context of probability spaces equipped with a tree-like structure. Here is the precise definition (cf. [7]).

**Definition 1.2.** Suppose that $(X, \mu)$ is a nonatomic probability space. A set $\mathcal{T}$ of measurable subsets of $X$ will be called a tree if the following conditions are satisfied:

(i) $X \in \mathcal{T}$ and for every $Q \in \mathcal{T}$ we have $\mu(Q) > 0$.

(ii) For every $Q \in \mathcal{T}$ there is a finite subset $C(Q) \subset \mathcal{T}$ containing at least two elements such that
  (a) the elements of $C(Q)$ are pairwise disjoint subsets of $Q$,
  (b) $Q = \bigcup C(Q)$.

(iii) $\mathcal{T} = \bigcup_{n \geq 0} \mathcal{T}^n$, where $\mathcal{T}^0 = \{X\}$ and $T^{n+1} = \bigcup_{Q \in \mathcal{T}^n} C(Q)$.

(iv) We have $\lim_{n \to \infty} \sup_{Q \in \mathcal{T}^n} \mu(Q) = 0$.

An important example, which links this definition with the preceding considerations, is the cube $X = [0,1)^n$ endowed with Lebesgue measure and the tree of its dyadic subcubes. Any probability space equipped with a tree gives rise to the corresponding multilinear maximal operators $\mathcal{M} = \mathcal{M}_{X,\mathcal{T}}$ and Muckenhoupt's classes $A_{\vec{P}} = A_{\vec{P}}(X,\mathcal{T})$: the definitions are word-by-word the same, one only needs to change the base space $\mathbb{R}^N$ to $X$ and replace the dyadic lattice $\mathscr{D}$ by the tree $\mathcal{T}$.

Our main results are gathered in two statements below.

**Theorem 1.3.** *Let $(X,\mu)$ be an arbitrary nonatomic probability space with a tree structure $\mathcal{T}$. Fix $m^{-1} < p < \infty$ and put $p_1 = p_2 = \ldots = p_m = mp$. Then for any vector $\vec{w} \in A_{\vec{P}}(X,\mathcal{T})$ we have the inequality*
(1.5)
$$\|\mathcal{M}_{X,\mathcal{T}}\|_{L^{p_1}(w_1) \times L^{p_2}(w_2) \times \ldots \times L^{p_m}(w_m) \to L^p(v_{\vec{w}})} \leq \left( \frac{mp}{mp - 1 - d(mp, [w]_{A_{\vec{P}}})} \right)^m.$$

*The constant on the right is the best possible: for any $c \geq 1$ and $\varepsilon > 0$ there is a weight $\vec{w}$ satisfying $[\vec{w}]_{A_{\vec{P}}} \leq c$ such that*

$$\|\mathcal{M}_{X,\mathcal{T}}\|_{L^{p_1}(w_1) \times L^{p_2}(w_2) \times \ldots \times L^{p_m}(w_m) \to L^p(v_{\vec{w}})} > \left( \frac{mp}{mp - 1 - d(mp, [w]_{A_{\vec{P}}})} - \varepsilon \right)^m.$$

**Theorem 1.4.** *Let $(X,\mu)$ be an arbitrary nonatomic probability space with a tree structure $\mathcal{T}$. Suppose that $\vec{P} = (p_1, p_2, \ldots, p_m)$ with $1 < p_1, p_2, \ldots, p_m < \infty$ and let $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_m}$. Then the optimal constant $K_{\vec{P},c}$ in the estimate*

(1.6) $$\|\mathcal{M}_{X,\mathcal{T}}\|_{L^{p_1}(w_1) \times L^{p_2}(w_2) \times \ldots \times L^{p_m}(w_m) \to L^p(v_{\vec{w}})} \leq K_{\vec{P},[w]_{A_{\vec{P}}}} \prod_{j=1}^{m} p_j'$$

*satisfies*

$$([w]_{A_{\vec{P}}}(1 + d(q,c)))^{p_*'/p} \leq K_{\vec{P},[w]_{A_{\vec{P}}}} \leq ([w]_{A_{\vec{P}}}(1 + d(p_*,c)))^{p_*'/p},$$

*where $p_* = \min\{p_1, p_2, \ldots, p_m\}$ and $q = p/p_*' + 1$.*

In particular, in (1.6) we recover the optimal dependence on the characteristic $[\vec{w}]_{A_{\vec{P}}}$. Interestingly, the lower and the upper bound for $K_{\vec{P},[w]_{A_{\vec{P}}}}$ involves the multiplicative constant of the same type: $(1 + d(\cdot, c))^{p_*'/p}$ (unfortunately, these constants can be quite distant if $p_*$ is close to 1). We would also like to emphasize that the estimate (1.5) is sharp for each individual probability space. Finally, let us mention that standard translation and scaling arguments allow to extend the above results to the non-probabilistic, dyadic context studied at the beginning.

Our reasoning will rest on a multilinear version of Carleson embedding theorem and a tight 'testing' estimate (2.4) below. To show the latter bound, we will exploit

the so-called Bellman function method, a powerful tool for establishing inequalities, used widely in probability and analysis. Quite interestingly, the function we invent involves as many as $m + 2$ variables, and still, as we will see, the calculations are quite quick and do not require elaborate analysis.

The paper is organized as follows. In the next section, we present the proof of the estimates (1.5) and (1.6). Section 3 is devoted to the lower bounds for the constants involved in these estimates.

## 2. Proof of (1.5) and (1.6)

Throughout, $\vec{P} = (p_1, p_2, \ldots, p_m)$ is a vector of exponents belonging to $(1, \infty)$ and the parameter $p$ is given by $\frac{1}{p} = \frac{1}{p_1} + \frac{1}{p_2} + \ldots + \frac{1}{p_m}$. By symmetry, we may and will assume that $p_1$ is the smallest exponent.

Let us briefly describe the idea behind our approach. We will exploit the following multilinear version of Carleson embedding theorem established in [2] (see also [4]). Here and below, the symbol $\langle f \rangle_{Q,\sigma}$ stands for the weighted average $\frac{1}{\sigma(Q)} \int_Q f \mathrm{d}\mu$.

**Theorem 2.1.** *Suppose that a sequence $(a_Q)_{Q \in \mathcal{T}}$ of nonnegative numbers satisfies the following condition: for any $R \in \mathcal{T}$ we have the estimate*

$$(2.1) \qquad \sum_{Q \subseteq R} a_Q \leq C \int_R \prod_{j=1}^m \sigma_j^{p/p_j} \, d\mu.$$

*Then for any vector $\vec{f} = (f_1, f_2, \ldots, f_m)$ such that $f_j \in L^{p_j}(\sigma_j)$ for all $j$, we have*

$$(2.2) \qquad \left( \sum_{Q \in \mathcal{T}} a_Q \prod_{j=1}^m \langle f_j \rangle_{Q,\sigma_j}^p \right)^{1/p} \leq C^{1/p} \prod_{j=1}^m p_j' \|f_j\|_{L^{p_j}(\sigma_j)}.$$

It is well known (see [6], for example) that the appropriate choice of the sequence $(a_Q)_{Q \in \mathcal{T}}$ transforms the estimate (2.2) into (1.5) and (1.6). Thus the problem reduces to the identification of the appropriate constant $C$ in (2.1). To handle this, we will apply the Bellman function method. We introduce the auxiliary constants

$$\alpha = 1 + \frac{1}{d(p_1, c)} \qquad \text{and} \qquad C = \frac{c(1 + d(p_1, c))(p_1 - 1)}{p_1 - 1 - d(p_1, c)} = (c(1 + d(p_1, c)))^{p_1'},$$

where the last equality follows from (1.3). Consider the Bellman function $B = B_{\vec{P},c} : \mathbb{R}_+^{m+2} \to \mathbb{R}$ given by

$$B(s_1, s_2, \ldots, s_m, t, u) = \alpha \left( t^p u + c(p_1 - 1) \prod_{j=1}^m s_j^{p/p_j} - c p_1 t^{p/p_1} \prod_{j=1}^m s_j^{p/p_j - p/p_1} \right).$$

Sometimes, for the sake of brevity, it will be convenient to write $B(s, t, u)$ instead of $B(s_1, s_2, \ldots, s_m, t, u)$. The function $B$ is the key ingredient of the proof of the weighted estimates (1.5) and (1.6). Let us study several crucial properties of this object.

**Lemma 2.2.** *(i) For any fixed $t$ and $u$, the function*

$$(s_1, s_2, \ldots, s_m) \to B(s_1, s_2, \ldots, s_m, t, u)$$

*is concave.*

(ii) For any fixed $s$ and $u$, the function $t \mapsto B(s,t,u)$ is nonincreasing on the interval $\left(0, \left(c\prod_{j=1}^{m} s_j^{p/p_j - p/p_1} u^{-1}\right)^{p_1'/p}\right)$.

(iii) If $u \geq \prod_{j=1}^{m} s_j^{-p/p_j'}$, then we have the majorization

(2.3) $$B(s,t,u) \geq t^p u - C s_1^{p/p_1} s_2^{p/p_2} \ldots s_m^{p/p_m}.$$

*Proof.* (i) If $\gamma_1, \gamma_2, \ldots, \gamma_m$ is a sequence of positive numbers summing up to 1, then the function $(s_1, s_2, \ldots, s_m) \mapsto s_1^{\gamma_1} s_2^{\gamma_2} \ldots s_m^{\gamma_m}$ is concave. Furthermore, if $\lambda_1, \lambda_2, \ldots, \lambda_m$ are nonpositive numbers, then the function $(s_1, s_2, \ldots, s_m) \mapsto s_1^{\lambda_1} s_2^{\lambda_2} \ldots s_m^{\lambda_m}$ is convex. Both these facts can be easily proved by the induction on $m$, and they immediately yield the desired claim.

(ii) This is straightforward: we have

$$B_t(s,t,u) = \alpha p t^{p/p_1 - 1} \left(t^{p/p_1'} u - c \prod_{j=1}^{m} s_j^{p/p_j - p/p_1}\right).$$

(iii) The assertion is equivalent to

$$(\alpha - 1)t^p u + (C + \alpha c(p_1 - 1)) \prod_{j=1}^{m} s_j^{p/p_j} \geq \alpha c p_1 t^{p/p_1} \prod_{j=1}^{m} s_j^{p/p_j - p/p_1}.$$

The left-hand side is an increasing function of $u$, so it is enough to prove the estimate for $u = \prod_{j=1}^{m} s_j^{-p/p_j'}$. Divide both sides by $(\alpha - 1)p_1 \prod_{j=1}^{m} s_j^{p/p_j}$ to obtain the equivalent bound

$$\frac{1}{p_1} \cdot \left(\frac{t}{s_1 s_2 \ldots s_m}\right)^p + \frac{1}{p_1'} \cdot \frac{C + \alpha c(p_1 - 1)}{(\alpha - 1)(p_1 - 1)} \geq \frac{\alpha c}{\alpha - 1} \left(\frac{t}{s_1 s_2 \ldots s_m}\right)^{p/p_1}.$$

This will follow immediately from Young's inequality (with exponents $p_1$ and $p_1'$), as soon as we show that

$$\frac{C + \alpha c(p_1 - 1)}{(\alpha - 1)(p_1 - 1)} = \left(\frac{\alpha c}{\alpha - 1}\right)^{p_1'}.$$

Plugging the formulas for $\alpha$ and $C$, we transform the above identity into

$$\frac{p_1 - 1}{p_1 - 1 - d(p_1, c)} = (c(1 + d(p_1, c)))^{1/(p_1 - 1)},$$

which holds true by the very definition (1.3) of the parameter $d(p_1, c)$. $\qquad \square$

The following result will imply the validity of (2.1).

**Theorem 2.3.** *Suppose that $(X, \mu)$ is a probability space with a tree structure $\mathcal{T}$. Fix $c \geq 1$ and let $\vec{w}$ be a vector of weights on $X$, satisfying $[\vec{w}]_{A_{\vec{P}}} \leq c$. Then for any $R \in \mathcal{T}$, the vector $\vec{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_m)$ of dual weights satisfies*

(2.4) $$\int_R \mathcal{M}(\vec{\sigma}\chi_R)^p v_{\vec{w}} d\mu \leq C \int_R \prod_{j=1}^{m} \sigma_j^{p/p_j} d\mu.$$

*Proof.* It is convenient to split the reasoning into a few intermediate parts.

*Step 1. An auxiliary notation.* Fix $w$ and $R$ as in the statement and let $n_0$ be the unique integer such that $R \in \mathcal{T}^{n_0}$. We introduce the auxiliary functional sequences $(x_n)_{n \geq n_0}$, $(y_n)_{n \geq n_0}$ and $(z_n)_{n \geq n_0}$ as follows: for any $n \geq n_0$ and $\omega \in \Omega$,

$$x_n(\omega) = (x_{n1}(\omega), x_{n2}(\omega), \ldots, x_{nm}(\omega)) = \left( \langle \sigma_1 \rangle_{Q_n(\omega)}, \langle \sigma_2 \rangle_{Q_n(\omega)}, \ldots, \langle \sigma_m \rangle_{Q_n(\omega)} \right)$$

$$y_n(\omega) = \max_{n_0 \leq k \leq n} \left( x_{k1}(\omega) x_{k2}(\omega) \ldots x_{km}(\omega) \right),$$

$$z_n(\omega) = \langle v_{\vec{w}} \rangle_{Q_n(\omega)},$$

where $Q_n(\omega)$ is the unique element of $\mathcal{T}^n$ which contains $\omega$. There is a nice probabilistic interpretation of these sequences: $(x_n)_{n \geq N}$ and $(z_n)_{n \geq N}$ are martingales induced by the filtration $(\mathcal{T}^n)_{n \geq N}$, with the terminal variables equal to $(\sigma_1, \sigma_2, \ldots, \sigma_m)$ and $v_{\vec{w}}$, respectively; in addition, $(y_n)_{n \geq N}$ is the maximal process associated with the sequence of products $(x_{k1} x_{k2} \ldots x_{km})_{k \geq n_0}$. Observe that by Lebesgue's differentiation theorem (or rather, by Doob's martingale convergence theorem), we have $x_{nj} \to \sigma_j$, $y_n \to \mathcal{M}(\vec{\sigma} \chi_R)$ and $z_n \to v_{\vec{w}}$ almost surely as $n \to \infty$.

*Step 2. Monotonicity.* Now we will show that the sequences $(x_n)_{n \geq n_0}$, $(y_n)_{n \geq n_0}$ and $(z_n)_{n \geq n_0}$ combine nicely with the Bellman function $B$ defined previously. More specifically, we will prove that the sequence

$$\left( \int_Q B(x_n, y_n, z_n) \mathrm{d}\mu \right)_{n \geq n_0}$$

is nonincreasing. To see this, fix an integer $n \geq n_0$, let $Q$ be an arbitrary element of $\mathcal{T}^n$ and let $Q_1, Q_2, \ldots, Q_\ell$ be the collection of all children of $Q$ in $\mathcal{T}^{n+1}$. The functions $x_n$, $y_n$, $z_n$ are constant on $Q$, while $x_{n+1}$, $y_{n+1}$ and $z_{n+1}$ are constant on each $Q_j$. It is easy to check that these constant values satisfy

$$(2.5) \qquad \langle x_n \rangle_Q = \sum_{j=1}^{\ell} \frac{|Q_j|}{|Q|} \langle x_{n+1} \rangle_{Q_j}, \qquad \langle z_n \rangle_Q = \sum_{j=1}^{\ell} \frac{|Q_j|}{|Q|} \langle z_{n+1} \rangle_{Q_j}.$$

(This is nothing but the martingale property, expressed in analytic terms). Next, observe that for any $j$ we have

$$(2.6) \qquad \int_{Q_j} B(x_{n+1}, y_{n+1}, z_{n+1}) \mathrm{d}\mu \leq \int_{Q_j} B(x_{n+1}, y_n, z_{n+1}) \mathrm{d}\mu.$$

Indeed, if $y_{n+1} = y_n$ on $Q_j$, then there is nothing to prove; on the other hand, if $y_{n+1} \neq y_n$ on $Q_j$, then by the definition of the sequence $y$ we must have $y_n < y_{n+1} = x_{(n+1)1} x_{(n+1)2} \ldots x_{(n+1)m}$. The latter product is equal to

$$\langle \sigma_1 \rangle_{Q_j} \langle \sigma_2 \rangle_{Q_j} \ldots \langle \sigma_m \rangle_{Q_j} \leq \left( c \prod_{k=1}^{m} \langle \sigma_k \rangle_{Q_j}^{p/p_k - p/p_1} \langle v_{\vec{w}} \rangle_{Q_j}^{-1} \right)^{p_1'/p}$$

$$= \left( c \prod_{k=1}^{m} x_{(n+1)k}^{p/p_k - p/p_1} \cdot z_{n+1}^{-1} \right)^{p_1'/p},$$

where the estimate follows from the condition $A_{\vec{P}}$. By Lemma 2.2 (ii), this implies $B(x_{n+1}, y_{n+1}, z_{n+1}) \leq B(x_{n+1}, y_n, z_{n+1})$ on $Q_j$ and hence (2.6) follows. Summing

over $j$, we thus obtain

$$\int_Q B(x_{n+1}, y_{n+1}, z_{n+1}) \mathrm{d}\mu \leq \int_Q B(x_{n+1}, y_n, z_{n+1}) \mathrm{d}\mu.$$

Now, by the very definition of $B$, the second identity in (2.5) and the fact that $y_n$ is constant on $Q$, the right-hand side above is equal to $\int_Q B(x_{n+1}, y_n, z_n) \mathrm{d}\mu$. Finally, by the first part of Lemma 2.2 and the first identity in (2.5), we have

$$\int_Q B(x_{n+1}, y_n, z_n) \mathrm{d}\mu \leq \int_Q B(x_n, y_n, z_n) \mathrm{d}\mu.$$

Hence, summing over all $Q \in \mathcal{T}^n$ contained in $R$, we obtain the aforementioned monotonicity.

*Step 3. Completion of the proof.* By the previous step, we obtain that for any $n \geq n_0$ we have

$$(2.7) \qquad \int_R B(x_n, y_n, z_n) \mathrm{d}\mu \leq \int_R B(x_{n_0}, y_{n_0}, z_{n_0}) \mathrm{d}\mu.$$

Let us inspect the expression on the right. By the very definition of $y$, we have $y_{n_0} = x_{n_0 1} x_{n_0 2} \ldots x_{n_0 m}$; furthermore, by the $A_{\vec{P}}$ condition, we obtain $z_{n_0} \leq c \prod_{j=1}^m x_{n_0 j}^{-p/p_j'}$. Consequently,

$$B(x_{n_0}, y_{n_0}, z_{n_0}) \leq \alpha \left( c \prod_{j=1}^m x_{n_0 j}^{p/p_j} + c(p_1 - 1) \prod_{j=1}^m x_{n_0 j}^{p/p_j} - cp_1 \prod_{j=1}^m x_{n_0 j}^{p/p_j} \right) = 0.$$

To handle the left-hand side of (2.7), we apply the third part of Lemma 2.2. As the result, we get

$$\int_R y_n^p z_n \mathrm{d}\mu \leq C \int_R \prod_{j=1}^m x_{nj}^{p/p_j} \mathrm{d}\mu.$$

It remains to let $n \to \infty$ and carry out an appropriate limiting procedure. Recall that we have the almost sure convergence $x_{nj} \to \sigma_j$, $y_n \to \mathcal{M}(\vec{\sigma}\chi_R)$ and $z_n \to v_{\vec{w}}$. Therefore, the left-hand side above can be handled by Fatou's lemma:

$$\int_R \mathcal{M}(\vec{\sigma}\chi_R)^p v_{\vec{w}} \mathrm{d}\mu \leq \liminf_{n \to \infty} \int_R y_n^p z_n \mathrm{d}\mu.$$

To deal with the right hand side, we will apply Lebesgue's dominated convergence theorem: we have $\prod_{j=1}^m x_{nj}^{p/p_j} \to \prod_{j=1}^m \sigma_j^{p/p_j}$ almost surely, so all we need is a suitable majorant. The inclusion $\vec{w} \in A_{\vec{P}}$ implies $\sigma_j \in A_{mp_j'}$ (cf. [5]), which gives that $\sigma_j$, and hence also the maximal function $M\sigma_j$, belongs to $L^r$ for some $r > 1$. Thus, by the Hölder inequality, the majorant $\prod_{j=1}^m \left( \sup_{n \geq 0} x_{nj} \right)^{p/p_j}$ is integrable and the assertion follows. $\square$

We are ready for the proof of our main estimates.

*Proof of* (1.5) *and* (1.6). Fix a vector $\vec{w} \in A_{\vec{P}}(X, \mathcal{T})$ and an arbitrary sequence $\vec{f} = (f_1, f_2, \ldots, f_m)$ of functions on $X$ such that $f_j \in L^{p_j}(w_j)$ for each $j$. Clearly, we may assume that $f_j > 0$ for all $j$. Let $\varepsilon > 0$ be a fixed parameter. By the

definition of the multilinear maximal operator, for any $\omega \in X$ there is a set $\mathcal{Q}(\omega)$ containing $\omega$ such that

$$\mathcal{M}\vec{f}(\omega) \leq (1+\varepsilon)\langle f_1\rangle_{\mathcal{Q}(\omega)}\langle f_2\rangle_{\mathcal{Q}(\omega)}\ldots\langle f_m\rangle_{\mathcal{Q}(\omega)}.$$

Of course, such a set $\mathcal{Q}(\omega)$ need not be unique: to avoid ambiguity, we take $\mathcal{Q}(\omega)$ belonging to $\mathcal{T}^n$ with $n$ as small as possible. Now, for any $Q \in \mathcal{T}$, we define $E(Q) = \{\omega \in X : \mathcal{Q}(\omega) = Q\}$ and let

$$a_Q = \prod_{j=1}^{m}\langle \sigma_j\rangle_Q^p \cdot v_{\vec{w}}(E(Q)).$$

By the very definition, we have $E(Q) \subseteq Q$ and the sets $E(Q)$ corresponding to different $Q$'s are disjoint. Furthermore, we have $a_Q \leq (\mathcal{M}(\vec{\sigma}\chi_Q)(\omega))^p v_{\vec{w}}(E(Q))$ for $\omega \in Q$. Consequently, for any $R \in \mathcal{T}$ we get

$$\sum_{Q \subseteq R} a_R \leq \int_R (\mathcal{M}(\vec{\sigma}\chi_Q))^p v_{\vec{w}}\mathrm{d}\mu \leq C \int_R \prod_{j=1}^{m}\sigma_j^{p/p_j}\mathrm{d}\mu,$$

where the second inequality is due to (2.4). Thus, by the Carleson embedding theorem applied to the vector $\overrightarrow{f\sigma^{-1}} = (f_1\sigma_1^{-1}, f_2\sigma_2^{-1}, \ldots, f_m\sigma_m^{-1})$, we obtain

$$\left(\sum_{Q \in \mathcal{T}} a_Q \prod_{j=1}^{m}\langle f_j\sigma_j^{-1}\rangle_{Q,\sigma_j}^p\right)^{1/p} \leq C^{1/p}\prod_{j=1}^{m}p_j'\|f_j\sigma_j^{-1}\|_{L^{p_j}(\sigma_j)} = C^{1/p}\prod_{j=1}^{m}p_j'\|f_j\|_{L^{p_j}(w_j)}.$$

Now, by the definition of $a_Q$ and the identity $\langle f_j\sigma_j^{-1}\rangle_{Q,\sigma_j} = \langle f_j\rangle_Q\langle \sigma_j\rangle_Q^{-1}$, we get

$$\sum_{Q \in \mathcal{T}} a_Q \prod_{j=1}^{m}\langle f_j\sigma_j^{-1}\rangle_{Q,\sigma_j}^p = \sum_{Q \in \mathcal{T}} v_{\vec{w}}(E(Q))\prod_{j=1}^{m}\langle f_j\rangle_Q^p \geq (1+\varepsilon)^{-1}\int_X (\mathcal{M}\vec{f})^p v_{\vec{w}}\mathrm{d}\mu.$$

Since $\varepsilon$ was arbitrary, the desired estimates follow. Let us remark that in the symmetric case $p_1 = p_2 = \ldots = p_m$, the obtained constant is

$$C^{1/p}\prod_{j=1}^{m}p_j' = (c(1+d(mp,c)))^{m/(mp-1)} \cdot \left(\frac{mp}{mp-1}\right)^m$$

$$= \left(\frac{mp-1}{mp-1-d(mp,c)}\right)^m \cdot \left(\frac{mp}{mp-1}\right)^m = \left(\frac{mp}{mp-1-d(mp,c)}\right)^m,$$

as we have announced in the statement of Theorem 1.3. $\qquad\square$

## 3. On the lower bound for the constant

3.1. **Sharpness for $p_1 = p_2 = \ldots = p_m$.** First we will show that in the symmetric case the constant we have obtained is the best possible. Fix an integer $m$, an exponent $p \in (m^{-1}, \infty)$ and a constant $c \in [1, \infty)$. By the result of [9], for any $\varepsilon > 0$ there is a weight $w \in A_{mp}$ with $[w]_{A_{mp}} = c$ and a function $f \in L^{mp}(w)$ such that the (one-dimensional) maximal function $M$ satisfies

$$\|Mf\|_{L^{mp}(w)} > \left(\frac{mp}{mp-1-d(mp,c)} - \varepsilon\right)\|f\|_{L^{mp}(w)}.$$

We consider the vectors $\vec{f} = (f, f, \ldots, f)$ and $\vec{w} = (w, w, \ldots, w)$, each consisting of $m$ coordinates. Then $v_{\vec{w}} = \prod_{j=1}^m w^{p/mp} = w$, so for any $Q \in \mathcal{T}$,

$$\langle v_{\vec{w}} \rangle_Q \prod_{j=1}^m \langle w_j^{1-p_j'} \rangle_Q^{p/p_j'} = \langle w \rangle_Q \langle w^{1-(mp)'} \rangle_Q^{mp-1}.$$

In particular, this implies $[\vec{w}]_{A_{\vec{p}}} = [w]_{A_{mp}} = c$. Furthermore,

$$\|\mathcal{M}\vec{f}\|_{L^p(v_{\vec{w}})} = \|Mf\|_{L^{mp}(w)}^m > \left( \frac{mp}{mp - 1 - d(mp, c)} - \varepsilon \right)^m \|f\|_{L^{mp}(w)}^m$$

$$= \left( \frac{mp}{mp - 1 - d(mp, c)} - \varepsilon \right)^m \prod_{j=1}^m \|f\|_{L^{mp}(w)},$$

which establishes the desired sharpness.

3.2. **The lower bound in the asymmetric case.** Here the calculations will be more involved. It is convenient to split the construction into a few parts. We may and do assume that $c > 1$: when $c = 1$, then the vector $\vec{w}$ consists of constant weights and the claim follows easily from the unweighted theory.

*Step 1. Auxiliary geometrical facts and parameters.* Pick $\tilde{c} \in (1, c)$ and set $q = p/p_1' + 1$. There are two lines passing through the point $K = (1, \tilde{c}^{1/(q-1)})$ which are tangent to the curve $xy^{p-1} = c$; we take the line $\ell$ for which the $x$-coordinate of the tangency point is smaller than 1. This line intersects the curve $xy^{q-1} = 1$ at two points: pick the point $L$ with bigger $x$-coordinate and denote this coordinate by $1 + d(\tilde{c})$. Furthermore, the line $\ell$ intersects the curve $xy^{q-1} = \tilde{c}$ at two points: one of them is $K$, while the second, denoted by $M$, is of the form $\left(1 - \delta, (\tilde{c}(1-\delta))^{1/(1-q)}\right)$. See Figure 2 below.
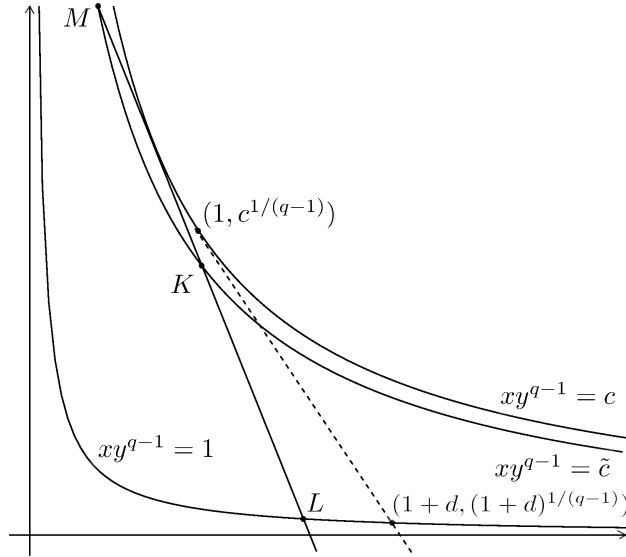


FIGURE 2. The crucial three points $K = (1, \tilde{c}^{1/(q-1)})$, $L = \left(1 + d(\tilde{c}), (1 + d(\tilde{c}))^{1/(1-q)}\right)$ and $M = \left(1 - \delta, (\tilde{c}(1-\delta))^{1/(1-q)}\right)$.

The points $K$, $L$, $M$ are colinear: some simple algebra transforms this into the equality

$$(3.1) \qquad (\tilde{c}(1 + d(\tilde{c})))^{1/(q-1)}\big(d(\tilde{c}) + \delta - d(\tilde{c})(1 - \delta)^{1/(1-q)}\big) = \delta,$$

which will be useful later.

*Step 2. Construction.* Now, recall the following simple measure-theoretic fact, which can be found in [7].

**Lemma 3.1.** *For every $Q \in \mathcal{T}$ and every $\beta \in (0,1)$ there is a subfamily $F(Q) \subset \mathcal{T}$ consisting of pairwise disjoint subsets of $Q$ such that*

$$\mu\left(\bigcup_{R \in F(Q)} R\right) = \sum_{R \in F(Q)} \mu(R) = \beta\mu(Q).$$

We use this fact recursively and construct an appropriate sequence $A_0 \supset A_1 \supset A_2 \supset \ldots$ of subsets of $X$. The starting point is the choice $A_0 = X$. To describe the inductive step, assume we have constructed $A_n$, which is a union of pairwise almost disjoint elements of $\mathcal{T}$, called the *atoms* of $A_n$. Of course, this condition is satisfied for $n = 0$: we have $A_0 = X \in \mathcal{T}$. Then, for each atom $Q$ of $A_n$, we apply the above lemma with $\beta = d(\tilde{c})/(d(\tilde{c}) + \delta)$ and get the corresponding subfamily $F(Q)$. Put $A_{n+1} = \bigcup_Q \bigcup_{Q' \in F(Q)} Q'$, the first union taken over all atoms $Q$ of $A_n$. Directly from the definition, this set is a union of the family $\{F(Q) : Q \text{ an atom of } A_n\}$, which consists of pairwise disjoint elements of $\mathcal{T}$. We call these elements the atoms of $A_{n+1}$ and conclude the description of the induction step.

As an immediate consequence of the above construction, we see that if $Q$ is an atom of $A_k$, then for any $n \geq k$ we have

$$\mu(Q \cap A_n) = \mu(Q)\left(\frac{d(\tilde{c})}{d(\tilde{c}) + \delta}\right)^{n-k}$$

and hence

$$(3.2) \qquad \mu(Q \cap (A_n \setminus A_{n+1})) = \mu(Q)\left(\frac{d(\tilde{c})}{d(\tilde{c}) + \delta}\right)^{n-k}\frac{\delta}{d(\tilde{c}) + \delta}.$$

Now, introduce the weights $w_1, w_2, \ldots, w_m$ on $X$ by the formula

$$w_1^{p/p_1} = \sum_{n=0}^{\infty} \chi_{A_n \setminus A_{n+1}}(1 - \delta)^n$$

and $w_2 = w_3 = \ldots = w_m \equiv 1$. Furthermore, for $j = 1, 2, \ldots, m$, we let

$$f_j = \sum_{n=0}^{\infty} \chi_{A_n \setminus A_{n+1}}(1 + a_j\delta)^n,$$

where $a_1 = (p_1 d(\tilde{c}))^{-1} + p^{-1} - \varepsilon/p_1$ and, for $j \geq 2$, $a_j = (p_j d(\tilde{c}))^{-1} - \varepsilon/p_j$. Here $\varepsilon$ is a fixed positive parameter (which will be sent to zero at the end of the proof).

*Step 3. We have $[w]_{A_{\vec{P}}} \leq c$.* The equality $w_2 = w_3 = \ldots = w_m = 1$ implies that $v_{\vec{w}} = w_1^{p/p_1}$ and $v_{\vec{w}}^{-p_1'/p} = w_1^{1-p_1'}$. Hence $[w]_{A_{\vec{P}}} \leq c$ is equivalent to showing that $v_{\vec{w}} \in A_q$. This in turn amounts to saying that for any $Q \in \mathcal{T}$,

$$\langle v_{\vec{w}}\rangle_Q \langle v_{\vec{w}}^{1/(1-q)}\rangle_Q^{q-1} \leq c.$$

To prove this, we use (3.2) to obtain that for each atom $Q$ of $A_k$ we have

$$(3.3) \qquad \langle v_{\vec{w}} \rangle_Q = \sum_{n=k}^{\infty} \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^{n-k} (1-\delta)^n \cdot \frac{\delta}{d(\tilde{c}) + \delta} = \frac{(1-\delta)^k}{1 + d(\tilde{c})}$$

and

$$\begin{aligned}
\langle v_{\vec{w}}^{1/(1-q)} \rangle_Q &= \sum_{n=k}^{\infty} \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^{n-k} (1-\delta)^{n/(1-q)} \cdot \frac{\delta}{d(\tilde{c}) + \delta} \\
&= \frac{\delta}{d(\tilde{c}) + \delta} (1-\delta)^{k/(1-q)} \cdot \left( 1 - \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} (1-\delta)^{1/(1-q)} \right)^{-1} \\
&= \frac{c^{1/(q-1)}(1-\delta)^{k/(1-q)}}{(1 + d(\tilde{c}))^{1/(1-q)}},
\end{aligned}$$

where in the last line we have used (3.1). Suppose that $R$ is an arbitrary element of $\mathcal{T}$. Then there is an integer $k$ such that $R \subseteq A_{k-1}$ and $R \nsubseteq A_k$. We have

$$\begin{aligned}
\langle v_{\vec{w}} \rangle_R &= \frac{1}{\mu(R)} \int_{R \setminus A_k} v_{\vec{w}} \mathrm{d}\mu + \frac{1}{\mu(R)} \int_{R \cap A_k} v_{\vec{w}} \mathrm{d}\mu \\
&= \frac{1}{\mu(R)} \int_{R \setminus A_k} (1-\delta)^{k-1} \mathrm{d}\mu + \frac{1}{\mu(R)} \int_{R \cap A_k} v_{\vec{w}} \mathrm{d}\mu.
\end{aligned}$$

By (3.3), applied to each atom $Q$ of $A_k$ contained in $R$, we get

$$\int_{R \cap A_k} v_{\vec{w}} \mathrm{d}\mu = \frac{\mu(R \cap A_k)(1-\delta)^k}{1 + d(\tilde{c})}$$

and hence, setting $\eta := \mu(R \cap A_k)/\mu(R)$, we rewrite the preceding equality as

$$\langle v_{\vec{w}} \rangle_R = (1-\eta)(1-\delta)^{k-1} + \eta \cdot \frac{(1-\delta)^k}{1 + d(\tilde{c})}.$$

Similarly, we obtain

$$\langle v_{\vec{w}}^{1/(1-q)} \rangle_R = (1-\eta)(1-\delta)^{(k-1)/(1-q)} + \eta \cdot \frac{c^{1/(q-1)}(1-\delta)^{k/(1-q)}}{(1 + d(\tilde{c}))^{1/(1-q)}},$$

which implies

$$\begin{aligned}
&\langle v_{\vec{w}} \rangle_R \langle v_{\vec{w}}^{1/(1-q)} \rangle_R^{q-1} \\
&= \left( \eta(1-\delta) + (1-\eta)(1 + d(\tilde{c})) \right) \left( \eta(1-\delta)^{\frac{1}{1-q}} + (1-\eta)(1 + d(\tilde{c}))^{\frac{1}{1-q}} \right)^{q-1}.
\end{aligned}$$

This number does not exceed $c$. Indeed, the right-hand side can be rewritten as

$$(\eta M_x + (1-\eta)L_x)(\eta M_y + (1-\eta)L_y)^{q-1},$$

where $M_x$, $M_y$ and $L_x$, $L_y$ are the coordinates of the points $M$ and $L$ (see Figure 2). As $\eta$ ranges from 0 to 1, the point $\eta M + (1-\eta)L$ runs over the line segment $ML$ which is entirely contained in $\{(x,y) : xy^{q-1} \le c\}$. Since $R$ was arbitrary, we obtain the desired $A_{\vec{P}}$ condition: $[w]_{A_{\vec{P}}} \le c$.

   *Step 4. Completion of the proof.* In the same manner as above, one verifies that if $Q$ is an atom of $A_k$, then

$$\langle f_j \rangle_Q = \sum_{n \ge k} (1 + a_j \delta)^n \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^{n-k} \frac{\delta}{d(\tilde{c}) + \delta} = \frac{(1 + a_j \delta)^k}{1 - a_j d(\tilde{c})}.$$

Note that the ratio of the above geometric series, is equal to

$$\frac{(1 + a_j\delta)d(\tilde{c})}{d(\tilde{c}) + \delta} = 1 + \frac{(a_j d(\tilde{c}) - 1)\delta}{d(\tilde{c}) + \delta} < 1 + \frac{(p_j^{-1} - 1)\delta}{d(\tilde{c}) + \delta} < 1,$$

so the series is convergent. Hence, on $A_k$ (and hence also on $A_k \setminus A_{k-1}$) we have

$$\mathcal{M}\vec{f} \geq \prod_{j=1}^{m} \langle f_j \rangle_Q = \prod_{j=1}^{m} \frac{(1 + a_j\delta)^k}{1 - a_j d(\tilde{c})}$$

and therefore

$$\|\mathcal{M}\vec{f}\|_{L^p(v_{\vec{w}})} \geq \left( \sum_{n \geq 0} \left( \prod_{j=1}^{m} \frac{(1 + a_j\delta)^n}{1 - a_j d(\tilde{c})} \right)^p (1 - \delta)^n \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^n \frac{\delta}{d(\tilde{c}) + \delta} \right)^{1/p}.$$

The ratio of the geometric series in the parentheses is given by

$$\prod_{j=1}^{m} (1 + a_j\delta)^p (1 - \delta) \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)$$

$$= 1 + \left[ p(a_1 + a_2 + \ldots + a_m) - 1 - (d(\tilde{c}))^{-1} \right] \delta + o(\delta)$$

$$= 1 - \varepsilon\delta + o(\delta)$$

as $\delta \to 0$. Thus the series is convergent and we obtain

$$\|\mathcal{M}\vec{f}\|_{L^p(v_{\vec{w}})} \geq (d(\tilde{c})\varepsilon)^{-1/p} \prod_{j=1}^{m} (1 - a_j d(\tilde{c}))^{-1} + O(\delta).$$

On the other hand, an analogous computation shows that

$$\|f_1\|_{L^{p_1}(w_1)} = \left( \sum_{n \geq 0} (1 + a_1\delta)^{np_1} (1 - \delta)^{np_1/p} \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^n \frac{\delta}{d(\tilde{c}) + \delta} \right)^{1/p_1}$$

$$= \left( p_1 d(\tilde{c}) \left( (p_1 d(\tilde{c}))^{-1} + p^{-1} - a_1 \right) \right)^{-1/p_1} + O(\delta) = (d(\tilde{c})\varepsilon)^{-1/p_1} + O(\delta)$$

and for $j \geq 2$,

$$\|f_j\|_{L^{p_j}(w_j)} = \left( \sum_{n \geq 0} (1 + a_j\delta)^{np_j} \left( \frac{d(\tilde{c})}{d(\tilde{c}) + \delta} \right)^n \frac{\delta}{d(\tilde{c}) + \delta} \right)^{1/p_j}$$

$$= (1 - a_j p_j d(\tilde{c}))^{-1/p_j} + O(\delta) = (d(\tilde{c})\varepsilon)^{-1/p_j} + O(\delta).$$

Putting the above facts together (and noting that $d(\tilde{c}) \to d(q, c)$ as $\delta \to 0$), we get

$$\limsup_{\delta \to 0} \frac{\|\mathcal{M}\vec{f}\|_{L^p(v_{\vec{w}})}}{\prod_{j=1}^{m} \|f_j\|_{L^{p_j}(w_j)}} \geq \prod_{j=1}^{m} (1 - a_j d(q, c))^{-1}.$$

Finally, we check that the constant on the right converges, as $\varepsilon \to 0$, to

$$\left( 1 - \frac{p_1' d(q, c)}{p} \right)^{-1} \prod_{j=1}^{m} p_j' = (c(1 + d(q, c)))^{p_1'/p} \prod_{j=1}^{m} p_j'.$$

This yields the desired lower bound for the optimal constant.

## Acknowledgments

## References

[1] S. M. Buckley, *Estimates for operator norms on weighted spaces and reverse Jensen inequalities*, Trans. Amer. Math. Soc. **340** (1993), 253–272.

[2] W. Chen, W. Damián, *Weighted estimates for the multisublinear maximal function*, Rend. Circ. Mat. Palermo **62** (2013), 379–391.

[3] W. Damián, A. K. Lerner and C. Pérez, *Sharp weighted bounds for multilinear maximal functions and Calderón-Zygmund operators*, J. Fourier Anal. Appl. **21** (2015), no. 1, 161–181.

[4] T. Hytönen and C. Pérez, *Sharp weighted bounds involving $A_\infty$*, Anal. PDE **6** (2013), 777–818.

[5] A. K. Lerner, S. Ombrosi, C. Pérez, R. H. Torres, R. Trujillo-González, *New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory*, Adv. Math. **220** (2009), 1222–1264.

[6] K. Li, K. Moen and W. Sun, *The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators*, J. Fourier Anal. Appl. **20** (2014), no. 4, 751–765.

[7] A. D. Melas, *The Bellman functions of dyadic-like maximal operators and related inequalities*, Adv. Math. 192 (2005), 310–340.

[8] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Amer. Math. Soc **165** (1972), 207–226.

[9] A. Osękowski, *Best constants in Muckenhoupt's inequality*, Ann. Acad. Sci. Fenn. Math. **42** (2017), 889–904.

Faculty of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland

*Email address*: A.Osekowski@mimuw.edu.pl