Monatsh Math (2018) 186:321–336
https://doi.org/10.1007/s00605-018-1184-9

Weighted maximal inequalities for the Haar system

Adam Osękowski[1]

Received: 29 October 2016 / Accepted: 9 April 2018 / Published online: 19 April 2018

Abstract The paper contains the study of weighted maximal L^p-inequalities for the Haar system, with the optimal dependence on the characteristics of the weights involved. The proofs exploit certain special functions, enjoying appropriate size conditions and concavity.

Keywords Maximal · Dyadic · Bellman function · Best constants

Mathematics Subject Classification Primary 42B25; Secondary 46E30, 60G42

1 Introduction

The purpose of this paper is to study certain class of weighted inequalities for the Haar system. Let $h = (h_n)_{n\geq 0}$ be the collection of Haar functions on [0, 1), given by

$$
\begin{aligned}
h_0 &= \chi_{[0,1)}, & h_1 &= \chi_{[0,1/2)} - \chi_{[1/2,1)},\\
h_2 &= \chi_{[0,1/4)} - \chi_{[1/4,1/2)}, & h_3 &= \chi_{[1/2,3/4)} - \chi_{[3/4,1)},\\
h_4 &= \chi_{[0,1/8)} - \chi_{[1/8,1/4)}, & h_5 &= \chi_{[1/4,3/8)} - \chi_{[3/8,1/2)},\\
h_6 &= \chi_{[1/2,5/8)} - \chi_{[5/8,3/4)}, & h_7 &= \chi_{[3/4,7/8)} - \chi_{[7/8,1)}
\end{aligned}
$$

Communicated by A. Constantin.

Research supported by the NCN Grant DEC-2014/14/E/ST1/00532.

✉ Adam Osękowski
ados@mimuw.edu.pl

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

and so on. A classical result of Schauder [16] asserts that the Haar system forms a basis of $L^p = L^p(0,1)$, $1 \leq p < \infty$ (with the underlying Lebesgue measure). Marcinkiewicz showed in [8] that if $1 < p < \infty$, then this basis is unconditional: there is a finite positive constant β_p depending only on p with the property that if n is a nonnegative integer, $a_0, a_1, \ldots, a_n$ are real numbers and $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n$ is a sequence of signs, then

$$\left\| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right\|_{L^p} \leq \beta_p \left\| \sum_{k=0}^{n} a_k h_k \right\|_{L^p}. \tag{1.1}$$

One can study a weighted version of this estimate. Here and below, the word "weight" will refer to a positive and integrable function on $[0,1)$. Given $1 < p < \infty$, we say that a weight w satisfies dyadic Muckenhoupt's condition A_p (shorter: w is a dyadic A_p weight) if

$$[w]_{A_p} := \sup_I \left(\frac{1}{|I|} \int_I w \right) \left(\frac{1}{|I|} \int_I w^{-1/(p-1)} \right)^{p-1} < \infty.$$

Here the supremum is taken over all dyadic subintervals I of $[0,1)$ (i.e., all I of the form $[a2^{-k}, (a+1)2^{-k})$, for some $k \geq 0$ and $a \in \{0, 1, \ldots, 2^k - 1\}$). There are versions of this definition for $p \in \{1, \infty\}$: w belongs to the dyadic A_∞ class if

$$[w]_{A_\infty} := \sup_I \left(\frac{1}{|I|} \int_I w \right) \exp \left(\frac{1}{|I|} \int_I \log(1/w) \right) < \infty,$$

and w is a dyadic A_1 weight if

$$[w]_{A_1} = \sup_I \left(\frac{1}{|I|} \int_I w \right) \Big/ \operatorname*{essinf}_I w < \infty.$$

In both conditions, the suprema are taken over all dyadic subintervals I of $[0,1)$. One easily verifies that $[w]_{A_p} \leq [w]_{A_q}$ if $q \leq p$ and hence the classes A_p grow as p increases.

It follows from the work [10] of Nazarov, Treil and Volberg and the extrapolation theorem of Rubio de Francia [15] that if $1 < p < \infty$ and w is an A_p weight, then there is a constant $C_{p,w}$ depending only on the parameters indicated such that

$$\left\| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right\|_{L^p(w)} \leq C_{p,w} \left\| \sum_{k=0}^{n} a_k h_k \right\|_{L^p(w)}.$$

There is a natural and interesting question concerning the optimal dependence of $C_{p,w}$ on the A_p-characteristics $[w]_{A_p}$. More precisely, the problem is to find, for each $1 < p < \infty$, an optimal exponent $\alpha = \alpha(p)$ such that $C_{p,w} \leq C_p [w]_{A_p}^{\alpha}$, where C_p does not depend on w. This type of question was first studied by Buckley [1] in the context of weighted estimates for maximal operators. Wittwer [22] showed that $\alpha(2) = 1$ which, by the sharp version of the extrapolation theorem of Rubio de Francia, established by Dragičević et al. [6] (see also Duoandikoetxea [7]), yielded the optimal dependence:

$$\left\| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right\|_{L^p(w)} \leq C_p [w]_{A_p}^{\max\{1,1/(p-1)\}} \left\| \sum_{k=0}^{n} a_k h_k \right\|_{L^p(w)} . \tag{1.2}$$

Our first result is the following maximal version of the estimate above. Throughout the paper, C_p denotes the optimal value of the constant appearing in (1.2).

Theorem 1.1 *Let* $1 < p < \infty$. *If* w *is a dyadic* A_p *weight,* N *is a nonnegative integer,* $a_0, a_1, \ldots, a_N$ *are real numbers and* $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_N$ *is a sequence of signs, then*

$$\left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right| \right\|_{L^p(w)} \leq 2^{1+1/p} C_p [w]_{A_p}^{\max\{1,1/(p-1)\}} \left\| \sum_{k=0}^{N} a_k h_k \right\|_{L^p(w)} . \tag{1.3}$$

The exponent $\max\{1, 1/(p-1)\}$ *is the best possible.*

Clearly, only the validity of (1.3) is an issue, the optimality of the exponent follows at once from the fact that the above bound is stronger than (1.2).

In the case $p = 1$, the inequality (1.2) fails even in the unweighted setting, but one can study the substitute in which the maximal function appears on the right. Such an estimate allows much wider class of weights. A simple modification of the argument of Burkholder [4] and Coifman [5] shows that if $1 \leq q < \infty$ and w satisfies the dyadic condition A_∞, then we have the bound

$$\left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right| \right\|_{L^q(w)} \leq c_{q,w} \left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} a_k h_k \right| \right\|_{L^q(w)} \tag{1.4}$$

for $N = 0, 1, 2, \ldots$, with some $c_{q,w} < \infty$ depending only on the parameters indicated. Indeed, the aforementioned paper of Burkholder gives an appropriate unweighted good-lambda inequality involving the functions $\max_{0\leq n\leq N} \left|\sum_{k=0}^{n} a_k h_k\right|$ and $\max_{0\leq n\leq N} \left|\sum_{k=0}^{n} \varepsilon_k a_k h_k\right|$ (see (8.13) in [4]), which is then transformed into the context of A_∞ weights by the argument of Coifman. This weighted good-lambda estimate yields in turn the above L^q-inequality (1.4) by standard integration. Since $A_p \subset A_\infty$ for all p, we see that in particular the L^q-inequality holds true for A_p weights. Our principal goal is to extract the optimal dependence of $c_{q,w}$ on $[w]_{A_p}$. Here is the precise statement.

Theorem 1.2 *For any parameters* $1 \leq p, q < \infty$, *there is a constant* $C_{p,q}$ *depending only on* p *and* q *which has the following property. If* w *is a dyadic* A_p *weight,* N *is a nonnegative integer,* $a_0, a_1, \ldots, a_N$ *are real numbers and* $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_N$ *is a sequence of signs, then*

$$\left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right| \right\|_{L^q(w)} \leq C_{p,q} [w]_{A_p} \left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} a_k h_k \right| \right\|_{L^q(w)} . \tag{1.5}$$

The linear dependence on the A_p *characteristics is optimal for each* p.

Since $A_\infty = \bigcup_{1\leq p<\infty} A_p$, this gives us an alternative proof of (1.4) for A_∞ weights. There is a natural question whether the dependence of $c_{q,w}$ on $[w]_{A_\infty}$ is also linear. We have been unable to answer it, though some information on $C_{p,q}$ indicate that this might not be the case. More precisely, our proof will establish (1.5) with

$$C_{p,q} = 2^{1/q} \cdot 6 \cdot \inf_r \left\{ C_r \left(\frac{r}{q} + 3^{-r} \right)^{1/q} \right\}, \tag{1.6}$$

where C_r is the best constant in (1.2) and the infimum is taken over all r satisfying $r \geq \max\{p, 2\}$ and $r > q$. Let us provide a more explicit formula for $C_{p,q}$. As shown in [2], we have the estimate $C_2 \leq 1109$. Now, it follows from the extrapolation theorems of Duoandikoetxea [7] and the sharp weighted bounds for the dyadic maximal operator established in [12], that if $r \geq 2$, then

$$C_r \leq \left(\sqrt{8}re\right)^{(r-2)/(r-1)} C_2 \leq \sqrt{8}reC_2 < 8527r.$$

Modulo the constant factor, this inequality can be reversed. Burkholder [1] proved that the optimal choice for β_p in (1.1) is $\max\{p-1, 1/(p-1)\}$; this yields $C_r \geq r-1 \geq r/2$ for $r \geq 2$. Consequently, we obtain that

$$C_{p,q} \sim 2^{1/q} \inf_r \left\{ r \left(\frac{r}{q} + 3^{-r} \right)^{1/q} \right\}$$

(where '$\sim$' means that the ratio of both sides is bounded from below and from above by universal constants). It is easy to see that the expression in the parentheses is an increasing function of r (on the interval $[\max\{p, q, 2\}, \infty)$), so the infimum is attained for the choice $r = \max\{p, q, 2\}$. Note that if q is fixed and p goes to infinity, then the constant $C_{p,q}$ is of order $O(p^{1+1/q})$; this explosion suggests that the inequality (1.5) in the limit case $p = \infty$ might not hold (i.e., the dependence of the constant on $[w]_{A_\infty}$ might not be linear).

A few words about the proof and the organization of the paper are in order. Our approach will rest on the Bellman function method: we will deduce the validity of (1.3) and (1.5) from the existence of certain special functions, enjoying appropriate size conditions and concavity. The approach originates from the theory of stochastic optimal control, and its fruitful connection with probability and harmonic analysis was firstly observed by Burkholder in [3], during the study of the sharp version of (1.1). Following the seminal work [3], Burkholder and others applied the method in many semimartingale estimates (see the monograph [11] for details). A decisive step towards wider applications of the technique in harmonic analysis was made by Nazarov, Treil and Volberg [9,10], who put the approach in a more modern and universal form. Since then, the method has been applied in numerous problems arising in various areas of mathematics (cf. e.g. [13,14,17–21] and consult references therein).

The Bellman function proof presented in this paper is quite unusual. We start Sect. 2 with a standard statement that a successful treatment of the estimates (1.3) and (1.5) requires the construction of a certain function of six variables. However, instead of providing an explicit formula for such an object (which is a typical ingredient of a proof), we propose an abstract two-step reasoning. Namely, first we decrease the dimension of the problem, by showing that finding appropriate functions of *four* variables is sufficient to deduce the desired estimates. Then, in Sect. 3, we provide these special four-dimensional objects. Again, we do not present explicit formulas (which might be quite complicated, and the analysis of their properties could be delicate). Instead, we manage to get rid of almost all technicalities and extract the *existence* of these objects from the validity of the inequality (1.2).

The final part of the paper is devoted to the optimality of the exponents, which is demonstrated by constructing appropriate examples.

2 On the method of proof

Throughout this section, $1 < p < \infty$, $1 \leq q < \infty$ and $c \geq 1$ are given and fixed parameters. Introduce the "hyperbolic" set

$$\mathcal{D}_{p,c} = \left\{ (\mathbf{w}, \mathbf{v}) \in (0, \infty)^2 : 1 \leq \mathbf{w}\mathbf{v}^{p-1} \leq c \right\}.$$

This object arises naturally in the analysis of A_p weights. Next, introduce another domain $Dom_{p,c} = \mathbb{R} \times \mathbb{R} \times (0, \infty) \times [0, \infty) \times \mathcal{D}_{p,c}$, pick a function $B : Dom_{p,c} \to \mathbb{R}$ and consider the following set of requirements.

(i) For any $\mathbf{x} \in \mathbb{R} \backslash \{0\}$ and any $(\mathbf{w}, \mathbf{v}) \in \mathcal{D}_{p,c}$,

$$B(\mathbf{x}, \pm \mathbf{x}, |\mathbf{x}|, (\pm \mathbf{x}) \vee 0, \mathbf{w}, \mathbf{v}) \leq 0. \tag{2.1}$$

(Here and below, $a \vee b = \max\{a, b\}$.)

(ii) For any $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) \in Dom_{p,c}$ we have

$$B(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) = B(\mathbf{x}, \mathbf{y}, |\mathbf{x}| \vee \mathbf{z}, \mathbf{y} \vee \mathbf{u}, \mathbf{w}, \mathbf{v}). \tag{2.2}$$

(iii) For any $(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) \in Dom_{p,c}$ we have

$$B(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) \geq (\mathbf{y} \vee \mathbf{u})^q \mathbf{w} - \left| \varphi(\mathbf{x}, |\mathbf{x}| \vee \mathbf{z}) \right|^q \mathbf{w}, \tag{2.3}$$

where $\varphi : \mathbb{R} \times (0, \infty) \to \mathbb{R}$ is some fixed continuous function depending only on p, q and c.

(iv) The function B is midpoint concave in the following sense. Suppose that the points $(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$, $(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm) \in \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c}$ satisfy the conditions $|\mathbf{x}| \leq \mathbf{z}$, $\mathbf{y} \leq \mathbf{u}$,

$$(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) = \frac{(\mathbf{x}_+, \mathbf{y}_+, \mathbf{w}_+, \mathbf{v}_+) + (\mathbf{x}_-, \mathbf{y}_-, \mathbf{w}_-, \mathbf{v}_-)}{2}$$

and $|\mathbf{x}_+ - \mathbf{x}_-| = |\mathbf{y}_+ - \mathbf{y}_-|$. Then

$$B(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) \geq \frac{B(\mathbf{x}_+, \mathbf{y}_+, \mathbf{z}, \mathbf{u}, \mathbf{v}_+, \mathbf{w}_+) + B(\mathbf{x}_-, \mathbf{y}_-, \mathbf{z}, \mathbf{u}, \mathbf{v}_-, \mathbf{w}_-)}{2}. \tag{2.4}$$

The statement below presents the connection between the list of the above conditions with the validity of certain maximal inequalities.

Theorem 2.1 *If there exists a function B satisfying the conditions (i)–(iv), then we have the estimate*

$$\left\| \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} \varepsilon_k a_k h_k \right| \right\|_{L^q(w)} \leq 2^{1/q} \left\| \varphi\left(\sum_{k=0}^{N} a_k h_k, \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} a_k h_k \right| \right) \right\|_{L^q(w)}$$

for all $N \geq 0$, all sequences $a_0, a_1, \ldots, a_N \in \mathbb{R}$, $\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ and all dyadic A_p weighs w satisfying $[w]_{A_p} \leq c$.

Remark 2.2 In the considerations below, we will take $\varphi(\mathbf{x}, \mathbf{z}) = K\mathbf{x}$ or $\varphi(\mathbf{x}, \mathbf{z}) = K\mathbf{z}$, for some constant K: then the assertion corresponds to the estimates (1.3) or (1.5), respectively.

Proof It is convenient to split the reasoning into three parts.

Step 1 Some reductions and notation. By standard limiting arguments (Fatou's lemma and Lebesgue's dominated convergence theorem), we may and do assume that $a_0 \neq 0$ (recall that φ is assumed to be continuous). Note that it is enough to show the one-sided estimates

$$\left\| \max_{0\leq n\leq N} \left(\sum_{k=0}^{n} \varepsilon_k a_k h_k \right)_+ \right\|_{L^q(w)} \leq \left\| \varphi\left(\sum_{k=0}^{N} a_k h_k, \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} a_k h_k \right| \right) \right\|_{L^q(w)},$$

$$\left\| \max_{0\leq n\leq N} \left(\sum_{k=0}^{n} (-\varepsilon_k) a_k h_k \right)_+ \right\|_{L^q(w)} \leq \left\| \varphi\left(\sum_{k=0}^{N} a_k h_k, \max_{0\leq n\leq N} \left| \sum_{k=0}^{n} a_k h_k \right| \right) \right\|_{L^q(w)}$$

(where $a_+ = \max\{a, 0\}$), since if we rise their sides to power q and add them, we obtain an estimate which is stronger than the assertion. Furthermore, switching from $(\varepsilon_k)_{k\geq 0}$ to $(-\varepsilon_k)_{k\geq 0}$, we see that it is enough to focus on the first bound. For $n \geq 0$, introduce the notation

$$f_n = \sum_{k=0}^{n} a_k h_k, \quad g_n = \sum_{k=0}^{n} \varepsilon_k a_k h_k, \quad |f|_n^* = \max_{0\leq k\leq n} |f_k|, \quad g_n^* = \max_{0\leq k\leq n} (g_k)_+$$

and let w_n, v_n denote the projections of w and $w^{-1/(p-1)}$ on the space generated by the first $n+1$ Haar functions. That is, if $w = \sum_{k=0}^{\infty} b_k h_k$ and $w^{-1/(p-1)} = \sum_{k=0}^{\infty} c_k h_k$, then $w_n = \sum_{k=0}^{n} b_k h_k$ and $v_n = \sum_{k=0}^{n} c_k h_k$. Since $[w]_{A_p} \leq c$, one easily checks that for any n the pair (w_n, v_n) takes values in the set $\mathcal{D}_{p,c}$.

Step 2 Monotonicity property. The main part of the proof is to show that for $0 \le n \le N-1$ we have

$$\int_0^1 B\left(f_n, g_n, |f_n|^*, g_n^*, w_n, v_n\right) \mathrm{d}s \geq \int_0^1 B\left(f_{n+1}, g_{n+1}, |f_{n+1}|^*, g_{n+1}^*, w_{n+1}, v_{n+1}\right) \mathrm{d}s.$$

Note that the integrands are well-defined: as have already observed above, the pairs (w, v) take values in $\mathcal{D}_{p,c}$ and the assumption $a_0 \neq 0$, imposed at the beginning, guarantees that $|f_n|^* \geq |f_0| > 0$.

To check the above estimate, let I be the support of h_{n+1}. The functions $(f_n, g_n, |f_n|^*, g_n^*, w_n, v_n)$ and $(f_{n+1}, g_{n+1}, |f_{n+1}|^*, g_{n+1}^*, w_{n+1}, v_{n+1})$ coincide outside I, so it is enough to show that

$$\int_I B\left(f_n, g_n, |f_n|^*, g_n^*, w_n, v_n\right) \mathrm{d}s \geq \int_I B\left(f_{n+1}, g_{n+1}, |f_{n+1}|^*, g_{n+1}^*, w_{n+1}, v_{n+1}\right) \mathrm{d}s = \int_I B\left(f_{n+1}, g_{n+1}, |f_n|^*, g_n^*, w_{n+1}, v_{n+1}\right) \mathrm{d}s,$$

where in the latter passage we have exploited (2.2) and the trivial identities $|f_{n+1}^*| = |f_{n+1}| \vee |f_n|^*$, $g_{n+1}^* = g_{n+1} \vee g_n^*$. However, $f_n, g_n, |f_n|^*, g_n^*, w_n$ and v_n are constant on I; denote the appropriate values by $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}$, $\mathbf{u}$, $\mathbf{w}$ and $\mathbf{v}$, respectively. Then, on I, we have $f_{n+1} = \mathbf{x} + a_{n+1}h_{n+1}$, $g_{n+1} = \mathbf{y} + \varepsilon_{n+1}a_{n+1}h_{n+1}$, $w_{n+1} = \mathbf{w} + b_{n+1}h_{n+1}$ and $v_{n+1} = \mathbf{v} + c_{n+1}h_{n+1}$. We see that these four functions, restricted to I, take values in two-point sets: there are two points $\mathbf{x}_\pm$ with $\mathbf{x} = (\mathbf{x}_- + \mathbf{x}_+)/2$ such that $f_{n+1} \in \{\mathbf{x}_-, \mathbf{x}_+\}$; there are two points $\mathbf{y}_\pm$ with $\mathbf{y} = (\mathbf{y}_- + \mathbf{y}_+)/2$ such that $g_{n+1} \in \{\mathbf{y}_-, \mathbf{y}_+\}$, and similarly for w_{n+1} and v_{n+1}. Plugging this observation above, we see that the monotonicity property reduces to the condition (2.4).

Step 3 Completion of the proof. Now, applying (2.3), we get

$$\begin{aligned}
&\left\| \max_{0\le n\le N} \left(\sum_{k=0}^{n} \varepsilon_k a_k h_k\right)_+ \right\|_{L^q(w)}^q - \left\| \varphi\left(\sum_{k=0}^{N} a_k h_k, \max_{0\le n\le N} \left|\sum_{k=0}^{n} a_k h_k\right|\right) \right\|_{L^q(w)}^q \\
&= \int_0^1 \left((g_N^*)^q - \varphi(f_N, |f_N|^*)^q\right) w \mathrm{d}s \\
&= \int_0^1 \left((g_N^*)^q - \varphi(f_N, |f_N|^*)^q\right) w_N \mathrm{d}s \\
&\le \int_0^1 B(f_N, g_N, |f_N|^*, g_N^*, w_N, v_N) \mathrm{d}s \\
&\le \int_0^1 B(f_0, g_0, |f_0|^*, g_0^*, w_0, v_0) \mathrm{d}s
\end{aligned}$$

and it remains to note that the latter integrand is nonpositive, due to (2.1). □

Therefore, we have reduced the problem of showing (1.5) to the construction of an appropriate function of six variables. This seems to be a difficult task; the following theorem allows to decrease the number of variables to four.

Theorem 2.3 *Let $r \geq q$ and $L > 0$ be fixed. Suppose that $U : \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c} \to \mathbb{R}$ satisfies the majorizations*

$$U(\mathbf{x}, \pm\mathbf{x}, \mathbf{w}, \mathbf{v}) \leq 0 \qquad \text{for all } \mathbf{x} \in \mathbb{R},\ \mathbf{w},\ \mathbf{v} \in \mathcal{D}_{p,c}, \tag{2.5}$$
$$U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) \geq |\mathbf{y}|^r \mathbf{w} - L^r |\mathbf{x}|^r \mathbf{w} \qquad \text{for all } (\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) \in \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c}, \tag{2.6}$$
$$U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) \geq U(\mathbf{x}, 0, \mathbf{w}, \mathbf{v}) \qquad \text{for all } (\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) \in \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c}, \tag{2.7}$$

and the following concavity inequality. If the points $(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v})$, $(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm) \in \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c}$ satisfy the conditions

$$(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) = \frac{(\mathbf{x}_+, \mathbf{y}_+, \mathbf{w}_+, \mathbf{v}_+) + (\mathbf{x}_-, \mathbf{y}_-, \mathbf{w}_-, \mathbf{v}_-)}{2}$$

and $|\,\mathbf{x}_+ - \mathbf{x}_-| = |\,\mathbf{y}_+ - \mathbf{y}_-|$, then

$$U(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) \geq \frac{U(\mathbf{x}_+, \mathbf{y}_+, \mathbf{w}_+, \mathbf{v}_+) + U(\mathbf{x}_-, \mathbf{y}_-, \mathbf{w}_-, \mathbf{v}_-)}{2}. \tag{2.8}$$

Then the function

$$\begin{aligned} B&(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}, \mathbf{w}, \mathbf{v}) \\ &= 2^{q-1} \cdot \frac{U(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) + U(\mathbf{x},\ (\mathbf{u}-\mathbf{y}) \vee 0,\ \mathbf{w},\ \mathbf{v})}{(3L)^{r-q}(|\mathbf{x}| \vee \mathbf{z})^{r-q}} \\ &\quad - \frac{r-q}{q} \cdot (6L)^q (|\mathbf{x}| \vee \mathbf{z})^q \mathbf{w} \end{aligned}$$

satisfies the conditions (i)–(iv) with

$$\varphi(\mathbf{x}, \mathbf{z}) = \begin{cases} 2L\mathbf{x} & \text{if } r = q, \\ 6L \cdot \left(r/q + 3^{-r}\right)^{1/q} \mathbf{z} & \text{if } r > q. \end{cases}$$

Proof We start with (2.1). Observe that $a \vee 0 - a = (-a) \vee 0$, so by (2.5) and (2.7),

$$\begin{aligned} B(\mathbf{x}, \pm\mathbf{x}, |\mathbf{x}|, (\pm\mathbf{x}) \vee 0, \mathbf{w}, \mathbf{v}) &\leq 2^{q-1} \cdot \frac{U(\mathbf{x}, \pm\mathbf{x}, \mathbf{w}, \mathbf{v}) + U(\mathbf{x}, (\mp\mathbf{x}) \vee 0, \mathbf{w}, \mathbf{v})}{(3L)^{r-q}|\mathbf{x}|^{r-q}} \\ &\leq 2^{q-1} \cdot \frac{U(\mathbf{x}, \pm\mathbf{x}, \mathbf{w}, \mathbf{v}) + U(\mathbf{x}, \mp\mathbf{x}, \mathbf{w}, \mathbf{v})}{(3L)^{r-q}|\mathbf{x}|^{r-q}} \leq 0. \end{aligned}$$

The condition (2.2) is evident. To show (2.3), suppose first that $r = q$. Then, directly from (2.6) and the elementary estimate $(a+b)^r \le 2^{r-1}(a^r + b^r)$,

$$\begin{aligned} B(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{w},\mathbf{v}) &\ge 2^{r-1}\big(|\mathbf{y}|^r\mathbf{w} + ((\mathbf{u}-\mathbf{y})\vee 0)^r\mathbf{w} - 2L^r|\mathbf{x}|^r\mathbf{w}\big)\\ &\ge (\mathbf{y}\vee\mathbf{u})^r\mathbf{w} - (2L)^r|\mathbf{x}|^r\mathbf{w}. \end{aligned}$$

In the case $r > q$, note that for any nonnegative numbers A_1, A_2 we have the estimate $A_1^r A_2 + A_2 \ge A_1^q A_2$. If we plug $A_1 = (((\mathbf{u}-\mathbf{y})\vee 0)/(3L(|\mathbf{x}|\vee\mathbf{z}))$ and $A_2 = \big(3L(|\mathbf{x}|\vee\mathbf{z})\big)^q\mathbf{w}$, we get

$$\frac{((\mathbf{u}-\mathbf{y})\vee 0)^r\mathbf{w}}{(3L)^{r-q}(|\mathbf{x}|\vee\mathbf{z})^{r-q}} + \big(3L(|\mathbf{x}|\vee\mathbf{z})\big)^q\mathbf{w} \ge \big((\mathbf{u}-\mathbf{y})\vee 0\big)^q\mathbf{w}.$$

Similarly, one shows that

$$\frac{|\mathbf{y}|^r\mathbf{w}}{(3L)^{r-q}(|\mathbf{x}|\vee\mathbf{z})^{r-q}} + \big(3L(|\mathbf{x}|\vee\mathbf{z})\big)^q\mathbf{w} \ge |\mathbf{y}|^q\mathbf{w}.$$

Multiply these inequalities throughout by 2^{q-1}, add them and apply the elementary estimate $(a+b)^q \le 2^{q-1}(a^q + b^q)$ to obtain

$$2^{q-1}\cdot\frac{|\mathbf{y}|^r\mathbf{w} + ((\mathbf{u}-\mathbf{y})\vee 0)^r\mathbf{w}}{(3L)^{r-q}(|\mathbf{x}|\vee\mathbf{z})^{r-q}} \ge (\mathbf{y}\vee\mathbf{u})^q\mathbf{w} - \big(6L(|\mathbf{x}|\vee\mathbf{z})\big)^q\mathbf{w}.$$

Consequently, (2.6) gives

$$\begin{aligned} &B(\mathbf{x},\mathbf{y},\mathbf{z},\mathbf{u},\mathbf{w},\mathbf{v})\\ &\ge 2^{q-1}\cdot\frac{|\mathbf{y}|^r\mathbf{w} + ((\mathbf{u}-\mathbf{y})\vee 0)^r\mathbf{w} - 2L^r|\mathbf{x}|^r\mathbf{w}}{(3L)^{r-q}(|\mathbf{x}|\vee\mathbf{z})^{r-q}} - \frac{r-q}{q}(6L)^q(|\mathbf{x}|\vee\mathbf{z})^q\mathbf{w}\\ &\ge (\mathbf{y}\vee\mathbf{u})^q\mathbf{w} - \big(r/q + 3^{-r}\big)\,(6L)^q(|\mathbf{x}|\vee\mathbf{z})^q\mathbf{w}. \end{aligned}$$

It remains to verify the concavity inequality (2.4). Fix parameters $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$ as in the statement of (iv); by symmetry, we may assume that $|\mathbf{x}_+| \ge |\mathbf{x}_-|$. Observe first that by (2.6),

$$\begin{aligned} &\frac{\partial}{\partial\mathbf{s}}\left(2^{q-1}\cdot\frac{U(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) + U(\mathbf{x},\ (\mathbf{u}-\mathbf{y})\vee 0,\ \mathbf{w},\ \mathbf{v})}{(3L)^{r-q}\mathbf{s}^{r-q}} - \frac{r-q}{q}\cdot(6L)^q\mathbf{s}^q\mathbf{w}\right)\\ &= -2^{q-1}(r-q)\frac{U(\mathbf{x},\ \mathbf{y},\ \mathbf{w},\ \mathbf{v}) + U(\mathbf{x},\ (\mathbf{u}-\mathbf{y})\vee 0,\ \mathbf{w},\ \mathbf{v})}{(3L)^{r-q}\mathbf{s}^{r-q+1}}\\ &\quad -(r-q)\cdot(6L)^q\mathbf{s}^{q-1}\mathbf{w}\\ &\le (r-q)(6L)^q\mathbf{s}^{q-1}\mathbf{w}\cdot\left(\left(\frac{|\mathbf{x}|}{3\mathbf{s}}\right)^r - 1\right), \end{aligned}$$

which is nonpositive for $|\mathbf{x}| \le 3\mathbf{s}$. This calculation will allow us to change appropriately the terms $|\mathbf{x}|\vee\mathbf{z}$ in the formula for B, sometimes with values smaller than $|\mathbf{x}|$.

The first consequence is that we may assume that $|\mathbf{x}_-| \le \mathbf{z}$. Indeed, if both $|\mathbf{x}_+|$, $|\mathbf{x}_-|$ are larger than $\mathbf{z}$, then replacing $\mathbf{z}$ by $|\mathbf{x}_-|$ does not change the right-hand side of (2.4) and does not increase the left-hand side, making the inequality stronger. Our next step is to note that since $|\mathbf{x}_-| \le \mathbf{z}$, we have

$$|\mathbf{x}_+| \le |\mathbf{x}_+ - \mathbf{x}| + |\mathbf{x}| = |\mathbf{x} - \mathbf{x}_-| + |\mathbf{x}| \le 2|\mathbf{x}| + |\mathbf{x}_-| \le 3\mathbf{z}$$

and hence, by the above bound for the partial derivative $\partial/\partial \mathbf{s}$, we conclude that

$$\begin{aligned} &B(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{z}, \mathbf{u}, \mathbf{w}_\pm, \mathbf{v}_\pm) \\ &\quad \le 2^{q-1} \cdot \frac{U(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm) + U(\mathbf{x}_\pm, (\mathbf{u} - \mathbf{y}_\pm) \vee 0, \mathbf{w}_\pm, \mathbf{v}_\pm)}{(3L)^{r-q}\mathbf{z}^{r-q}} \\ &\qquad - \frac{r-q}{q}(6L)^q \mathbf{z}^q \mathbf{w}. \end{aligned} \tag{2.9}$$

Now, we obviously have $\mathbf{z}^q\mathbf{w} = (\mathbf{z}^q\mathbf{w}_+ + \mathbf{z}^q\mathbf{w}_-)/2$ and, by the midpoint concavity of U assumed in the statement of the theorem, we know that

$$U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) \ge \frac{U(\mathbf{x}_+, \mathbf{y}_+, \mathbf{w}_+, \mathbf{v}_+) + U(\mathbf{x}_-, \mathbf{y}_-, \mathbf{w}_-, \mathbf{v}_-)}{2}.$$

Finally, the inequality (2.7) implies

$$U(\mathbf{x}_\pm, (\mathbf{u} - \mathbf{y}_\pm) \vee 0, \mathbf{w}_\pm, \mathbf{v}_\pm) \le U(\mathbf{x}_\pm, \mathbf{u} - \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm)$$

and hence, applying the concavity of U again,

$$U(\mathbf{x}, \mathbf{u} - \mathbf{y}, \mathbf{w}, \mathbf{v}) \ge \frac{U(\mathbf{x}_+, (\mathbf{u} - \mathbf{y}_+) \vee 0, \mathbf{w}_+, \mathbf{v}_+) + U(\mathbf{x}_-, (\mathbf{u} - \mathbf{y}_-) \vee 0, \mathbf{w}_-, \mathbf{v}_-)}{2}.$$

Combining these observations with (2.9) establishes the desired estimate (2.4). □

3 An abstract Bellman function of four variables

As we have seen in the previous section, having constructed an appropriate special function immediately gives us a desired inequality for the Haar system. A well-known fact in the general Bellman function theory is that this implication can be reversed: the validity of a given estimate implies the existence of the corresponding abstract Bellman function. Our argumentation depends heavily on this phenomenon: the four dimensional U we search for will be extracted from the estimate (1.2).

To state things precisely, we fix, throughout this section, the parameters $1 < p < \infty$ and $1 \le q < \infty$. Pick $c \ge 1$ and take $r \ge p$. We have $[w]_{A_r} \le [w]_{A_p}$ and hence (1.2) implies

$$\left\| \sum_{k=0}^n \varepsilon_k a_k h_k \right\|_{L^r(w)} \le C_r [w]_{A_p}^{\max\{1,1/(r-1)\}} \left\| \sum_{k=0}^n a_k h_k \right\|_{L^r(w)}, \tag{3.1}$$

for $n = 0, 1, 2, \ldots$. Define $U = U_{p,r,c} : \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c} \to \mathbb{R}$ by the formula

$$U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) = \sup \left\{ \left\| \mathbf{y} + \sum_{k=1}^{n} \varepsilon_k a_k h_k \right\|_{L^r(w)}^{r} - C_r^r c^{\max\{r, r/(r-1)\}} \left\| \mathbf{x} + \sum_{k=1}^{n} a_k h_k \right\|_{L^r(w)}^{r} \right\},$$

where the supremum is taken over all n, all sequences $a_1, a_2, \ldots, a_n$ of real numbers, all sequences $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ of signs and all dyadic A_p weights w satisfying $[w]_{A_p} \leq c$, $\int_0^1 w = \mathbf{w}$ and $\int_0^1 w^{-1/(p-1)} = \mathbf{v}$. To see that the definition makes sense (the supremum is taken over a nonempty set), we need the following.

Lemma 3.1 *For any* $(\mathbf{w}, \mathbf{v}) \in \mathcal{D}_{p,c}$ *there is a dyadic* A_p *weight* w *on* $[0, 1)$ *with* $[w]_{A_p} \leq c$, $\int_0^1 w = \mathbf{w}$ *and* $\int_0^1 w^{-1/(p-1)} = \mathbf{v}$.

Proof It is easy to show, using a Darboux property, that there are two points $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ lying at the lower boundary of $\mathcal{D}_{p,c}$ (i.e., satisfying $x_1 y_1^{p-1} = x_2 y_2^{p-1} = 1$) such that $(\mathbf{w}, \mathbf{v})$ is the middle of the line segment $P_1 P_2$. Define w on $[0, 1)$ by setting

$$w(s) = \begin{cases} x_1 & \text{if } s < 1/2, \\ x_2 & \text{if } s \geq 1/2. \end{cases}$$

Then $\int_0^1 w = (x_1 + x_2)/2 = \mathbf{w}$ and $\int_0^1 w^{-1/(p-1)} = (x_1^{-1/(p-1)} + x_2^{-1/(p-1)})/2 = (y_1 + y_2)/2 = \mathbf{v}$. It remains to verify that $[w]_{A_p} \leq c$. By the above calculation,

$$\left(\int_0^1 w \right) \left(\int_0^1 w^{-1/(p-1)} \right)^{p-1} = \mathbf{w}\mathbf{v}^{p-1} \leq c,$$

and if I is an arbitrary dyadic, proper subinterval of $[0, 1)$, then w is constant on I, so

$$\left(\frac{1}{|I|} \int_I w \right) \left(\frac{1}{|I|} \int_I w^{-1/(p-1)} \right)^{p-1} = 1 \leq c.$$

Hence w has all the desired properties. □

Now we will show that the abstract function U above possesses all the properties required in Theorem 2.3.

Lemma 3.2 *The function* U *satisfies* (2.5), (2.6) *with* $L = C_r c^{\max\{1, 1/(r-1)\}}$, (2.7) *and the midpoint concavity* (2.8).

Proof The first condition follows directly from (3.1): all the expressions appearing under the supremum defining $U(\mathbf{x}, \pm\mathbf{x}, \mathbf{w}, \mathbf{v})$ are nonpositive. The majorization (2.6) is obtained by considering the sequence $a_1 = a_2 = \cdots = 0$ in the definition of

$U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$. To check (2.7), we will prove that $U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) = U(\mathbf{x}, -\mathbf{y}, \mathbf{w}, \mathbf{v})$ and that the function $\mathbf{y} \mapsto U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$ is convex. Both these facts are simple. The first of them follows by switching from (ε_k) to $(-\varepsilon_k)$ in the definition of $U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$; then it is clear that the suprema defining $U(\mathbf{x}, \pm\mathbf{y}, \mathbf{w}, \mathbf{v})$ are the same. To prove the convexity of $U(\mathbf{x}, \cdot, \mathbf{w}, \mathbf{v})$, pick $\alpha \in (0, 1)$, two real numbers $\mathbf{y}_1, \mathbf{y}_2$ and set $\mathbf{y} = \alpha\mathbf{y}_1 + (1-\alpha)\mathbf{y}_2$. If n is a nonnegative integer, $a_1, a_2, \ldots, a_n$ is a sequence of real numbers and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ is an arbitrary sequence of signs, then the definition of U and the convexity of the function $t \mapsto |t|^r$ implies that

$$\begin{aligned}
&\left\|\mathbf{y} + \sum_{k=1}^n \varepsilon_k a_k h_k\right\|_{L^r(w)}^r - C_r^r c^{\max\{r, r/(r-1)\}} \left\|\mathbf{x} + \sum_{k=1}^n a_k h_k\right\|_{L^r(w)}^r \\
&\quad \le \alpha \left\|\mathbf{y}_1 + \sum_{k=1}^n \varepsilon_k a_k h_k\right\|_{L^r(w)}^r + (1-\alpha)\left\|\mathbf{y}_2 + \sum_{k=1}^n \varepsilon_k a_k h_k\right\|_{L^r(w)}^r \\
&\qquad - C_r^r c^{\max\{r, r/(r-1)\}} \left\|\mathbf{x} + \sum_{k=1}^n a_k h_k\right\|_{L^r(w)}^r \\
&\quad \le \alpha U(\mathbf{x}, \mathbf{y}_1, \mathbf{w}, \mathbf{v}) + (1-\alpha) U(\mathbf{x}, \mathbf{y}_2, \mathbf{w}, \mathbf{v}).
\end{aligned}$$

Therefore, taking the supremum over all n and all sequences (a_k), (ε_k) gives the convexity of $U(\mathbf{x}, \cdot, \mathbf{w}, \mathbf{v})$, and hence (2.7) is established.

It remains to show (2.8). Fix points $(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$, $(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm) \in \mathbb{R} \times \mathbb{R} \times \mathcal{D}_{p,c}$ as in the statement. Let $a_1^\pm, a_2^\pm, \ldots, a_n^\pm, \varepsilon_1^+, \varepsilon_2^+, \ldots, \varepsilon_n^\pm, w^\pm$ be as in the definition of $U(\mathbf{x}_\pm, \mathbf{y}_\pm, \mathbf{w}_\pm, \mathbf{v}_\pm)$ (we may assume that the lengths of the sequences $a_1^+, a_2^+, \ldots$ and $a_1^-, a_2^-, \ldots$ are the same, adding some zeros if necessary). Let us splice these objects using the following procedure: consider the function $f : [0, 1) \to \mathbb{R}$ given by

$$f(s) = \begin{cases} \mathbf{x}_+ + \sum_{k=1}^n a_k^+ h_k(2s) & \text{if } s < 1/2, \\ \mathbf{x}_- + \sum_{k=1}^n a_k^- h_k(2s-1) & \text{if } s \ge 1/2. \end{cases}$$

Using the self-similarity of the Haar system, we see that

$$f = \frac{\mathbf{x}_1 + \mathbf{x}_2}{2} + \sum_{k=1}^N a_k h_k = \mathbf{x} + \sum_{k=1}^N a_k h_k,$$

for some N and some sequence $(a_k)_{k=1}^N$: we have $a_1 = \mathbf{x}_+ - \mathbf{x} = (\mathbf{x}_+ - \mathbf{x}_-)/2$ and all remaining terms a_j are either zero or belong to the set $\{a_1^\pm, a_2^\pm, \ldots, a_n^\pm\}$. We can do the same splicing procedure with the functions $\mathbf{y}_\pm + \sum_{k=1}^n \varepsilon_k^\pm a_k^\pm h_k$, arriving at the function $\mathbf{y} + \sum_{k=1}^N \varepsilon_k a_k h_k$, where N and a_k are the same as above, and $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ take values in $\{-1, 1\}$ (here we have used the assumption $|\mathbf{x}_+ - \mathbf{x}_-| = |\mathbf{y}_+ - \mathbf{y}_-|$: it implies that $\varepsilon_1 \in \{-1, 1\}$). Similarly, we "glue" the weights w^+ and w^- into one weight in $[0, 1)$, setting

$$w(s) = \begin{cases} w^+(2s) & \text{if } s < 1/2, \\ w^-(2s-1) & \text{if } s \geq 1/2. \end{cases}$$

This new weight satisfies $\int_0^1 w = \int_0^{1/2} w^+(2s)\mathrm{d}s + \int_{1/2}^1 w^-(2s-1)\mathrm{d}s = (\mathbf{w}_+ + \mathbf{w}_-)/2$ and, analogously, $\int_0^1 w^{-1/(p-1)} = \mathbf{v}$. Furthermore, we have $[w]_{A_p} \leq c$. Indeed, first note that

$$\left(\int_0^1 w\right)\left(\int_0^1 w^{-1/(p-1)}\right)^{p-1} = \mathbf{w}\mathbf{v}^{p-1} \leq c.$$

Next, if I is a dyadic subinterval of $[0, 1/2)$, then

$$\begin{aligned} &\left(\frac{1}{|I|}\int_I w\right)\left(\frac{1}{|I|}\int_I w^{-1/(p-1)}\right)^{p-1} \\ &\quad = \left(\frac{1}{|2I|}\int_{2I} w^+\right)\left(\frac{1}{|2I|}\int_{2I} (w^+)^{-1/(p-1)}\right)^{p-1} \leq c, \end{aligned}$$

since $[w^+]_{A_p} \leq c$; the case when I is a dyadic subinterval of $[1/2, 1)$ is dealt with similarly. Thus, by the very definition of $U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v})$, we have

$$\begin{aligned} &U(\mathbf{x}, \mathbf{y}, \mathbf{w}, \mathbf{v}) \\ &\geq \left\| \mathbf{y} + \sum_{k=1}^N \varepsilon_k a_k h_k \right\|_{L^r(w)}^r - C_r^r c^{\max\{r, r/(r-1)\}} \left\| \mathbf{x} + \sum_{k=1}^N a_k h_k \right\|_{L^r(w)}^r \\ &= \frac{1}{2}\left[\left\| \mathbf{y}_+ + \sum_{k=1}^n \varepsilon_k^+ a_k^+ h_k \right\|_{L^r(w^+)}^r - C_r^r c^{\max\{r, r/(r-1)\}} \left\| \mathbf{x}_+ + \sum_{k=1}^n a_k^+ h_k \right\|_{L^r(w^+)}^r \right. \\ &\quad \left. + \left\| \mathbf{y}_- + \sum_{k=1}^n \varepsilon_k^- a_k^- h_k \right\|_{L^r(w^-)}^r - C_r^r c^{\max\{r, r/(r-1)\}} \left\| \mathbf{x}_- + \sum_{k=1}^n a_k^- h_k \right\|_{L^r(w^-)}^r \right]. \end{aligned}$$

Since $a_1^\pm, a_2^\pm, \ldots, \varepsilon_1^\pm, \varepsilon_2^\pm, \ldots, w^\pm$ were arbitrary, the inequality (2.8) follows. □

We are ready to establish the inequalities announced in the introductory section.

Proof of (1.3) Fix $1 < p < \infty$, a weight w and let $c = [w]_{A_p}$. Apply Lemma 3.2 with $r = p$ to obtain an appropriate function U with the majorizing constant $C_p c^{\max\{1, 1/(p-1)\}}$. Plug this function into Theorem 2.3 with $q = p$ to get the Bellman function B with the majorizing function $\varphi(\mathbf{x}, \mathbf{z}) = 2C_p c^{\max\{1, 1/(p-1)\}}\mathbf{x}$. This function, used in Theorem 2.1, yields the assertion with the desired constant $2^{1/p} \cdot 2C_p c^{\max\{1, 1/(p-1)\}}$. □

Proof of (1.5) Fix $1 \leq p < \infty$ and $1 \leq q < \infty$, take an A_p weight w and set $c = [w]_{A_p}$. Suppose further that $r \geq p$ and $r > q$. Then Lemma 3.2, applied with this value of r, yields an appropriate function U with the majorizing constant

$C_r c^{\max\{1,1/(r-1)\}}$. This function can be used in Theorem 2.3 to obtain the Bellman function B with the majorizing function $\varphi(\mathbf{x},\mathbf{z}) = 6C_r c^{\max\{1,1/(r-1)\}} \cdot \left(r/q + 3^{-r}\right)^{1/q} \mathbf{z}$. This Bellman function, used in Theorem 2.1, yields the estimate (1.5) with the constant $2^{1/q} \cdot 6C_r \cdot \left(r/q + 3^{-r}\right)^{1/q} c^{\max\{1,1/(r-1)\}}$. To sharpen the dependence of the constant on the characteristic $c = [w]_{A_p}$, we impose the additional assumption $r \geq 2$. This leads us precisely to the claim, with the constant $C_{p,q}$ given by (1.6). □

4 Sharpness of the exponent in (1.5)

Since $[w]_{A_p} \leq [w]_{A_1}$ for any $p \geq 1$, it is enough to show the optimality of the exponent for $p = 1$. For the sake of clarity, we split the reasoning into three parts.

Step 1 Construction. Let N be a positive integer. Define the coefficients $a_0, a_1, \ldots, a_{2^N}$ by

$$\sum_{k=0}^{2^N} a_k h_k := h_0 - 2h_1 + 2h_2 - 2h_4 + 2h_8 - \cdots + 2 \cdot (-1)^{N-1} h_{2^N}$$

and the signs $\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{2^N}$ by requiring that

$$\sum_{k=0}^{2^N} \varepsilon_k a_k h_k := h_0 + 2h_1 + 2h_2 + 2h_4 + 2h_8 + \cdots + 2h_{2^N}.$$

Observe that $|\sum_{k=0}^{2^N} a_k h_k| \leq 3$ almost everywhere (one easily checks the identity $|\sum_{k=0}^{2^N} a_k h_k| = \chi_{[0,2^{-N-1})} + 3\chi_{[2^{-N-1},1)}$) and hence also

$$\max_{0 \leq n \leq 2^N} \left| \sum_{k=0}^{n} a_k h_k \right| \leq 3 \qquad \text{almost everywhere.} \tag{4.1}$$

In addition, we see that

$$\left. \sum_{k=0}^{2^N} \varepsilon_k a_k h_k \right|_{[0,2^{-N-1})} = 1 + 2(N+1) \geq 2(N+1). \tag{4.2}$$

Next, set $a = N/(N+1)$ and consider the weight w on $[0, 1)$, given by

$$\begin{aligned} w &= h_0 + ah_1 + (1+a)ah_2 + (1+a)^2 ah_4 + \cdots + (1+a)^N ah_{2^N} \\ &= (1+a)^{N+1} \chi_{[0,2^{-N-1})} + \sum_{k=0}^{N} (1+a)^k (1-a) \chi_{[2^{-k-1},2^{-k})}. \end{aligned}$$

Step 2 Verifying an A_1 condition. We will prove that w is an A_1 weight satisfying $[w]_{A_1} = (1-a)^{-1} = N + 1$. To this end, fix a dyadic interval $I \subseteq [0, 1)$. If

$|I| \le 2^{-N-1}$, then w is constant on I and hence $\frac{1}{|I|}\int_I w/\operatorname{essinf} w = 1 \le (1-a)^{-1}$. So, suppose that the length of I is at least 2^{-N}; then there is a nonnegative integer $m \le N$ and $k \in \{0, 1, 2, \ldots, 2^m - 1\}$ such that $I = [k\cdot 2^{-m}, (k+1)\cdot 2^{-m})$. If $k = 1$, then w is constant on I and hence $\frac{1}{|I|}\int_I w/\operatorname{essinf}_I w = 1 \le (1-a)^{-1}$, as previously. If $k = 0$, then $\operatorname{essinf}_I w = (1+a)^m(1-a)$ and

$$\frac{1}{|I|}\int_I w = 2^m\left[(1+a)^{N+1}2^{-N-1} + \sum_{k=m}^{N}(1+a)^k(1-a)2^{-k-1}\right] = (1+a)^m,$$

so $\frac{1}{|I|}\int_I w/\operatorname{essinf}_I w = (1-a)^{-1}$. Finally, if $k \ge 2$, then there is a unique ℓ such that $I \subset [2^{-\ell-1}, 2^{-\ell})$. Since w is constant on this larger interval and nondecreasing on $[0, 1)$, we get $\operatorname{essinf}_I w = \operatorname{essinf}_{[0,2^{-\ell})} w$ and

$$\frac{1}{|I|}\int_I w = \frac{1}{\left|[2^{-\ell-1}, 2^{-\ell})\right|}\int_{[2^{-\ell-1},2^{-\ell})} w \le \frac{1}{\left|[0, 2^{-\ell})\right|}\int_{[0,2^{-\ell})} w,$$

so the estimate $\frac{1}{|I|}\int_I w/\operatorname{essinf}_I w \le (1-a)^{-1}$ follows from the case $k = 0$ considered above (replace k by ℓ there).

Step 3 Completion of the proof. By the elementary bound $e^x \ge 1 + x$, we get

$$\left(\frac{1+a}{2}\right)^{N+1} = \left(1 + \frac{1-a}{1+a}\right)^{-N-1} \ge \exp\left(-(N+1)\frac{1-a}{1+a}\right) \ge e^{-1}.$$

Therefore, if we fix $q \ge 1$ and apply (4.1), (4.2), we obtain

$$\begin{aligned}
\left\|\max_{0\le n\le 2^N}\left|\sum_{k=0}^{n}\varepsilon_k a_k h_k\right|\right\|_{L^q(w)} &\ge \left\|\max_{0\le n\le 2^N}\left|\sum_{k=0}^{n}\varepsilon_k a_k h_k\right|\right\|_{L^1(w)}\\
&\ge \left\|\sum_{k=0}^{2^N}\varepsilon_k a_k h_k\right\|_{L^1(w)}\\
&\ge 2^{-N-1}\cdot 2(N+1)\cdot(1+a)^{N+1}\\
&= [w]_{A_1}\cdot 2(N+1)(1-a)\cdot\left(\frac{1+a}{2}\right)^{N+1}\\
&\ge [w]_{A_1}\cdot 2e^{-1}\\
&\ge \frac{2}{3}e^{-1}[w]_{A_1}\left\|\max_{0\le n\le 2^N}\left|\sum_{k=0}^{n}a_k h_k\right|\right\|_{L^q(w)}.
\end{aligned}$$

The proof is complete.

Acknowledgements The author would like to thank an anonymous referee for the very careful reading of the first version of the paper and many helpful suggestions which greatly improved the presentation.

References

1. Buckley, S.M.: Estimates for operator norms on weighted spaces and reverse Jensen inequalities. Trans. Am. Math. Soc. **340**, 253–272 (1993)
2. Bañuelos, R., Brzozowski, M., Osękowski, A.: Weighted inequalities for differentially subordinated martingales (**in preparation**)
3. Burkholder, D.L.: Boundary value problems and sharp inequalities for martingale transforms. Ann. Probab. **12**, 647–702 (1984)
4. Burkholder, D.L.: Martingales and fourier analysis in banach spaces probability and analysis (Varenna, 1985). Lecture Notes in Math, vol. 1206, pp. 61–108. Springer, Berlin (1986)
5. Coifman, R.R.: Distribution function inequalities for singular integrals. Proc. Nat. Acad. Sci. USA **69**, 2838–2839 (1972)
6. Dragičević, O., Grafakos, L., Pereyra, M.C., Petermichl, S.: Extrapolation and sharp norm estimates for classical operators on weighted Lebegue spaces. Publ. Math. **49**(1), 73–91 (2005)
7. Duoandikoetxea, J.: Extrapolation of weights revisited: new proofs and sharp bounds. J. Funct. Anal. **260**(6), 1886–1901 (2011)
8. Marcinkiewicz, J.: Quelques théoremes sur les séries orthogonales. Ann. Soc. Pol. Math. **16**, 84–96 (1937)
9. Nazarov, F.L., Treil, S.R.: The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. St. Petersb. Math. J. **8**, 721–824 (1997)
10. Nazarov, F.L., Treil, S.R., Volberg, A.: The Bellman functions and two-weight inequalities for Haar multipliers. J. Am. Math. Soc. **12**, 909–928 (1999)
11. Osękowski, A.: Sharp martingale and semimartingale inequalities. In: Wojtaszczyk, P. (ed.) Monografie Matematyczne, vol. 72. Birkhäuser (2012)
12. Osękowski, A.: Best constants in Muckenhoupt's inequality. Ann. Acad. Sci. Fenn. Math. **42**, 889–904 (2017)
13. 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)
14. Petermichl, S., Pott, S.: An estimate for weighted Hilbert transform via square functions. Trans. Am. Math. Soc. **354**, 1699–1703 (2002)
15. de Francia, R.: J.L.: Factorization theory and A_p weights. Am. J. Math. **106**(3), 533–547 (1984)
16. Schauder, J.: Eine Eigenschaft des Haarschen Orthogonalsystems. Math. Z. **28**, 317–320 (1928)
17. Slavin, L., Stokolos, A., Vasyunin, V.: Monge–Ampère equations and Bellman functions: the dyadic maximal operator. C. R. Acad. Sci. Paris Ser. I **346**, 585–588 (2008)
18. Slavin, L., Vasyunin, V.: Sharp results in the integral-form John–Nirenberg inequality. Trans. Am. Math. Soc. **363**, 4135–4169 (2011)
19. Slavin, L., Volberg, A.: Bellman function and the H^1-BMO duality, harmonic analysis, partial differential equations, and related topics. Contemp. Math. **428**, 113–126 (2007)
20. Vasyunin, V.: The exact constant in the inverse Hölder inequality for Muckenhoupt weights. Algebra i Analiz 15 (2003), 73–117; translation in St. Petersburg Math. J. **15**, 49–79 (2004) (**Russian**)
21. Vasyunin, V., Volberg, A.: Monge–Ampére equation and Bellman optimization of Carleson embedding theorems. Am. Math. Soc. Transl. Ser. **2**(226), 195–238 (2009)
22. Wittwer, J.: A sharp bound for the martingale transform. Math. Res. Lett. **7**, 1–12 (2000)