A WEIGHTED MAXIMAL INEQUALITY FOR DIFFERENTIALLY SUBORDINATE MARTINGALES

RODRIGO BAÑUELOS AND ADAM OSĘKOWSKI

ABSTRACT. The paper contains the proof of a weighted Fefferman-Stein inequality in a probabilistic setting. Suppose that $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ are martingales such that g is differentially subordinate to f and let $w = (w_n)_{n\geq 0}$ be a weight, i.e., a nonnegative, uniformly integrable martingale. Denoting by $Mf = \sup_{n\geq 0} |f_n|$, $Mw = \sup_{n\geq 0} w_n$ the maximal functions of f and w, we prove the weighted inequality

$$||g||_{L^1(w)} \leq C||Mf||_{L^1(Mw)},$$

where $C = 3+\sqrt{2}+4\ln 2 = 7.186802\ldots$. The proof rests on the existence of a special function enjoying appropriate majorization and concavity.

1. INTRODUCTION

Suppose that w is a weight, i.e., a nonnegative, locally integrable function on $\mathbb{R}^d$, and let M stand for the Hardy-Littlewood maximal operator. In 1971, Fefferman and Stein [5] proved the existence of a finite constant c_d, depending only on the dimension, such that

$$w\left(\left\{x \in \mathbb{R}^d : Mf(x) \geq 1\right\}\right) \leq c_d||f||_{L^1(Mw)}.$$

Here and below, we use the standard notation $w(E) = \int_E w(x)\mathrm{d}x$ and $||f||_{L^p(w)} = \left(\int_{\mathbb{R}^d} |f(x)|^p w(x)\mathrm{d}x\right)^{1/p}$, $0 < p < \infty$. This result led to the following natural conjecture, formulated by Muckenhoupt and Wheeden in the seventies. Namely, for any Calderón-Zygmund singular integral operator T, there is a constant $c_{T,d}$, depending only on T and d, such that

$$w\left(\left\{x \in \mathbb{R}^d : |Tf(x)| \geq 1\right\}\right) \leq c_{T,d}||f||_{L^1(Mw)}. \tag{1.1}$$

This problem remained open for a long time, and finally, a few years ago, it was proved to be false: see the counterexamples for the Hilbert transform provided by Reguera, Thiele, Nazarov, Reznikov, Vasyunin and Volberg in [10, 18, 19].

This leads to the following natural problem: what should be done to the right-hand side of (1.1) so that the inequality holds true? It follows from the results of Lerner, Ombrosi and Perez [8, 9] that this will be the case if we replace f by Mf. Actually, after this modification we have even the stronger estimate

$$||Tf||_{L^1(w)} \leq c_{T,d}||Mf||_{L^1(Mw)}. \tag{1.2}$$

One of the motivations for the results of this paper comes from the dyadic counterpart of this inequality. Let $h = (h_n)_{n\geq 0}$ stand for the usual Haar system on $[0, 1)$: $h_0 = \chi_{[0,1)}$,

2010 *Mathematics Subject Classification.* Primary: 60G44. Secondary: 40B25.
Key words and phrases. martingale, differential subordination, weight.
R. Bañuelos is supported in part by NSF Grant # 0603701-DMS.
A. Osękowski is supported in part by the NCN grant DEC-2014/14/E/ST1/00532.

$h_1 = \chi_{[0,1/2)} - \chi_{[1/2,1)}$, $h_2 = \chi_{[0,1/4)} - \chi_{[1/4,1/2)}$, and so on. For a given integrable function $f = \sum_{n=0}^{\infty} a_n h_n$ on $[0,1)$, let its maximal function $M_d f$ be given by $\sup_{N\geq 0}|f_N|$, where $f_N = \sum_{n=0}^{N} a_n h_n$ is the projection of f onto the space generated by the first $N+1$ Haar functions. For a given sequence $\varepsilon = (\varepsilon_n)_{n\geq 0}$ of numbers belonging to $[-1,1]$, we define the associated Haar multiplier $T = T_\varepsilon$ by $T\left(\sum_{n=0}^{\infty} a_n h_n\right) = \sum_{n=0}^{\infty} \varepsilon_n a_n h_n$. Finally, let w be a nonnegative function on $[0,1)$. In this context, the inequality (1.2) becomes

$$\int_0^1 |T_\varepsilon f| w \mathrm{d}x \leq C \int_0^1 M_d f M_d w \mathrm{d}x. \tag{1.3}$$

In the paper we take a closer look at this inequality. Actually, it will be more convenient for us to study the estimate in the more general, probabilistic setup. Suppose that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space filtered by $(\mathcal{F}_n)_{n\geq 0}$, a nondecreasing family of sub-σ-algebras of $\mathcal{F}$. Let $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ be two adapted real-valued martingales with difference sequences $df = (df_n)_{n\geq 0}$, $dg = (dg_n)_{n\geq 0}$ given by the equalities

$$df_0 = f_0, \qquad df_n = f_n - f_{n-1}, \qquad n = 1, 2, \ldots,$$

with a similar definition for dg. The maximal function of f is given by $f^* = \sup_{n\geq 0}|f_n|$. Following Burkholder [4], we say that $(g_n)_{n\geq 0}$ is differentially subordinate to $(f_n)_{n\geq 0}$, if for any nonnegative integer n we have $|dg_n| \leq |df_n|$ with probability 1. For example, this domination holds true in the above context of Haar multipliers: the martingale $\left(\sum_{k=0}^{n} \varepsilon_k a_k h_k\right)_{n\geq 0}$ is differentially subordinate to $\left(\sum_{k=0}^{n} a_k h_k\right)_{n\geq 0}$. Finally, let w be a weight, i.e., a nonnegative, integrable random variable; this variable gives rise to the martingale $(w_n)_{n\geq 0}$ given by $w_n = \mathbb{E}(w|\mathcal{F}_n)$, $n = 0, 1, 2, \ldots$. We will establish the following statement.

Theorem 1.1. *Suppose that $(f_n)_{n\geq 0}$, $(g_n)_{n\geq 0}$ are martingales such that $(g_n)_{n\geq 0}$ is differentially subordinate to $(f_n)_{n\geq 0}$. Then for any weight w we have the inequality*

$$\sup_{n\geq 0} ||g_n||_{L^1(w)} \leq C||f^*||_{L^1(w^*)}, \tag{1.4}$$

where $C = 3 + \sqrt{2} + 4\ln 2 = 7.186802\ldots$.

Obviously, the above theorem implies (1.3).

There is a natural question whether the weight w^* on the right can be decreased. A natural candidate is the smaller weight $M_r w := ((w^r)^*)^{1/r}$ for some fixed $r < 1$. We will show that after the replacement the inequality fails, even in the context of Haar multipliers.

Theorem 1.2. *For any constants $c > 0$ and $r \in (0,1)$, there is a weight w, a function f and a Haar multiplier T_ε on $[0,1)$ such that*

$$||T_\varepsilon f||_{L^1(w)} > c||M_d f||_{L^1(M_r w)}.$$

Though the constant C does not seem to be sharp in (1.4), we strongly believe that it is not far from the optimal value. The best constant in this inequality in the unweighted setting (i.e., for $w \equiv 1$) is equal to $2.536\ldots$ and was identified by Burkholder [4] in the context of stochastic integrals (see also [16]). His proof rests on the construction of a certain special function of three variables, enjoying appropriate size and concavity properties. This type of argument, originating from the theory of optimal control, is called the Bellman function method and has turned out to be an extremely efficient tool in probability and analysis [1, 2, 3, 11, 12, 13, 14, 20, 22]. Our approach will use this technique and will exploit a certain special function of five variables. The special function leading to the estimate (1.4) is introduced and studied in the next section. Section 3 contains the proofs of Theorems 1.1

and 1.2. In the final part of the paper we come back to the analytic setting, and apply the martingale inequalities to obtain maximal bounds for dyadic shifts (which, in turn, yield corresponding results for Calderón-Zygmund operators).

2. A SPECIAL FUNCTION

Let $\beta = 2 + \sqrt{2} + 4\ln 2$ and let $C = \beta + 1$ be the constant of Theorem 1.1. Consider the domain

$$\mathcal{D} = \mathbb{R} \times \mathbb{R} \times (0, \infty) \times [0, \infty) \times (0, \infty)$$

and let $B : \mathcal{D} \to \mathbb{R}$ be given by

$$B(x, y, z, u, v) = (y^2 + z^2)^{1/2}u - x^2z^{-1}v - \beta zv + 4zv \ln\left(uv^{-1} + 1\right).$$

Some steps which have led us to the discovery of this function, are described in Remark 2.1 below. We will need the following properties of this object.

Lemma 2.1. *(i) For any $x,\ y \in \mathbb{R}$ satisfying $|y| \leq |x|$ and any $u > 0$,*

$$B(x, y, |x|, u, u) \leq 0. \tag{2.1}$$

(ii) For any $(x, y, z, u, v) \in \mathcal{D}$ such that $|x| \leq z$ and $u \leq v$ we have

$$B(x, y, z, u, v) \geq |y|u - Czv. \tag{2.2}$$

(iii) For any $(x, y, z, u, v) \in \mathcal{D}$ such that $|x| \leq z$ and $u \leq v$ we have

$$\begin{aligned} &|B_x(x, y, z, u, v)| \leq 2v, \\ &|B_y(x, y, z, u, v)| \leq v, \\ &|B_u(x, y, z, u, v)| \leq |y| + 3z \end{aligned} \tag{2.3}$$

and

$$|B(x, y, z, u, v)| \leq |y|v + Czv. \tag{2.4}$$

Proof. The proof is very straightforward and requires nothing but some simple manipulations. We will only show (2.1) and (2.2), leaving the proof of the remaining properties to the reader. Note that

$$B(x, y, |x|, u, u) \leq (2x^2)^{1/2}u - |x|u - \beta|x|u + 4|x|u \ln 2 = -2|x|u \leq 0.$$

To prove (2.2), it suffices to observe that $(y^2 + z^2)^{1/2}u \geq |y|u$ and

$$-x^2z^{-1}v - \beta zv + 4zv \ln(uv^{-1} + 1) \geq -zv - \beta zv = -Czv. \quad \square$$

The main property of B is the following condition, which can be regarded as a concavity-type property.

Lemma 2.2. *For any $(x, y, z, u, v) \in \mathcal{D}$ and $d,\ h,\ k \in \mathbb{R}$ such that $|x| \leq z$, $u \leq v$, $u + d \geq 0$ and $|k| \leq |h|$, we have*

$$\begin{aligned} &B(x + h, y + k, (x + h) \vee z, u + d, (u + d) \vee v) \\ &\leq B(x, y, z, u, v) + B_x(x, y, z, u, v)h + B_y(x, y, z, u, v)k + B_u(x, y, z, u, v)d. \end{aligned} \tag{2.5}$$

Proof. By continuity, we may assume that $|x| < z$ and $u < v$. It is convenient to split the reasoning into four separate parts.

Case I: $|x + h| \leq z$, $u + d \leq v$. Consider the continuous function $G = G_{x,y,z,u,v,d,h,k} : [0, 1] \to \mathbb{R}$ given by

$$G(t) = B(x + th, y + tk, z, u + td, v). \tag{2.6}$$

The inequality (2.5) will follow if we prove that $G(1) \leq G(0) + G'(0)$, and hence we will be done if we show that G is concave on $[0,1)$. It is easy to compute that for $t \in (0,1)$,

$$\begin{aligned} G''(t) &= \frac{z^2k^2(u+td)}{((y+tk)^2+z^2)^{3/2}} + \frac{2(y+tk)kd}{((y+tk)^2+z^2)^{1/2}} - \frac{2vh^2}{z} - \frac{4zd^2}{v((u+td)/v+1)^2} \\ &\leq \frac{k^2(u+td)}{z} + \frac{2(y+tk)kd}{((y+tk)^2+z^2)^{1/2}} - \frac{2vh^2}{z} - \frac{zd^2}{v} \\ &\leq -\frac{h^2v}{z} + 2|h||d| - \frac{zd^2}{v} \leq 0. \end{aligned}$$

Hence (2.5) is established.

Case II: $|x+h| \geq z$, $u+d \leq v$. The estimate reads

$$\begin{aligned} &((y+k)^2+(x+h)^2)^{1/2}(u+d) - (\beta+1)|x+h|v + 4|x+h|v\ln((u+d)v^{-1}+1) \\ &\leq (y^2+z^2)^{1/2}(u+d) - x^2z^{-1}v - \beta zv + 4zv\ln(uv^{-1}+1) - 2xz^{-1}vh \\ &\quad + y(y^2+z^2)^{-1/2}uk + 4zv(u+v)^{-1}d. \end{aligned}$$

Changing the signs of x and h if necessary, we may assume that $h > 0$ (and then $|x+h| = x+h$). Put all the terms on the left-hand side and note that the obtained expression $E(x,y,z,u,v,h,k,d)$, considered as a function of k, is convex; hence it is enough to establish the estimate for extremal values of k, i.e., for $k = \pm h$. We will assume that $k = h$, for the other possibility the argumentation is analogous. By Case I, the above estimate holds in the limit case $|x+h| = z$; thus, it is enough to show that the expression $E(x,y,z,u,v,h,h,d)$ is a nonincreasing function of h (when the remaining parameters are fixed). To his end, we compute that the partial derivative of E with respect to h equals

$$\begin{aligned} &\frac{(y+h)+(x+h)}{((y+h)^2+(x+h)^2)^{1/2}}(u+d) - (\beta+1)v + 4v\ln((u+d)v^{-1}+1) \\ &\quad + 2xz^{-1}v - \frac{yu}{(y^2+z^2)^{1/2}} \\ &\leq 2^{1/2}v - (\beta+1)v + 4v\ln 2 + 2v + v = 0. \end{aligned}$$

Case III: $|x+h| \leq z$, $u+d > v$. Note that the second assumption implies $d > 0$. Let $t_0 \in (0,1)$ be the number determined by the condition $u + t_0 d = v$ and let G be given be the formula (2.6). By the reasoning appearing in Case I above, G is concave on $[0,t_0]$ and therefore

$$\begin{aligned} &B_x(x,y,z,u,v)h + B_y(x,y,z,u,v)k + B_u(x,y,z,u,v)d \\ &= G(0) + G'(0) \geq G(t_0) + G'_-(t_0)(1-t_0). \end{aligned}$$

Hence it is enough to show that

$$B(x+h,y+k,z,u+d,u+d) \leq G(t_0) + G'_-(t_0)(1-t_0). \tag{2.7}$$

This inequality can be rewritten in the form

$$\begin{aligned} &B(\tilde{x}+\tilde{h},\tilde{y}+\tilde{k},(\tilde{x}+\tilde{h})\vee z,\tilde{u}+\tilde{d},(\tilde{u}+\tilde{d})\vee v) \\ &\quad \leq B(\tilde{x},\tilde{y},z,\tilde{u},v) + B_x(\tilde{x},\tilde{y},z,\tilde{u},v)\tilde{h} + B_y(\tilde{x},\tilde{y},z,\tilde{u},v)\tilde{k} + B_u(\tilde{x},\tilde{y},z,\tilde{u},v)\tilde{d}, \end{aligned}$$

i.e., it is precisely (2.5), where $\tilde{x} = x + t_0h$, $\tilde{y} = y + t_0k$, $\tilde{u} = u + t_0d = v$, $\tilde{d} = d(1-t_0)$, $\tilde{h} = (1-t_0)h$ and $\tilde{k} = (1-t_0)k$. Plugging the formula for B and its partial derivatives, one easily checks that this inequality reads

$$
\begin{aligned}
&((\tilde{y}+\tilde{k})^2+z^2)^{1/2}(\tilde{u}+\tilde{d})-(\tilde{x}+\tilde{h})^2(\tilde{u}+\tilde{d})z^{-1}-\beta z(\tilde{u}+\tilde{d})+4z(\tilde{u}+\tilde{d})\ln 2\\
&\leq (\tilde{y}^2+z^2)^{1/2}(\tilde{u}+\tilde{d})-\frac{\tilde{x}^2\tilde{u}}{z}-\beta z\tilde{u}+4z\tilde{u}\ln 2+\frac{\tilde{y}\tilde{u}\tilde{k}}{(\tilde{y}^2+z^2)^{1/2}}-\frac{2\tilde{x}\tilde{u}\tilde{h}}{z}+2z\tilde{d}
\end{aligned}
$$

or

$$
\begin{aligned}
&\Big[((\tilde{y}+\tilde{k})^2+z^2)^{1/2}-(\tilde{y}^2+z^2)^{1/2}-\tilde{h}^2z^{-1}-\tilde{y}(\tilde{y}^2+z^2)^{-1/2}\tilde{k}\Big]\,\tilde{u}\\
&+\left[((\tilde{y}+\tilde{k})^2+z^2)^{1/2}-(\tilde{y}^2+z^2)^{1/2}-\frac{(\tilde{x}+\tilde{h})^2}{z}-\beta z+4z\ln 2-2z\right]\tilde{d}\leq 0.
\end{aligned}
$$

We will show that both expressions in the square brackets are non-positive. First, consider the function $\varphi(s)=((y+s)^2+z^2)^{1/2}$, $s\in\mathbb{R}$. Then by mean-value theorem,

$$
\begin{aligned}
&((\tilde{y}+\tilde{k})^2+z^2)^{1/2}-(\tilde{y}^2+z^2)^{1/2}-\tilde{y}(\tilde{y}^2+z^2)^{-1/2}\tilde{k}\\
&=\varphi(\tilde{k})-\varphi(0)-\varphi'(0)\tilde{k}=\varphi''(\xi)\tilde{k}^2/2,
\end{aligned}
$$

for some ξ lying between 0 and $\tilde{k}$. But $\varphi''(\xi)=z^2((y+\xi)^2+z^2)^{-3/2}\leq z^{-1}$ and hence the first expression in the square bracket above is non-positive (we use the inequality $\tilde{k}^2\leq\tilde{h}^2$ here). To handle the second expression, recall that $|\tilde{x}+\tilde{h}|=|x+h|\leq z$, so $|\tilde{k}|\leq|\tilde{h}|\leq|\tilde{x}+\tilde{h}|+|\tilde{x}|\leq 2z$ (the inequality $\tilde{x}\leq z$ follows from the estimates $|x|\leq z$ and $|x+h|\leq z$). Consequently,

$$
\begin{aligned}
&((\tilde{y}+\tilde{k})^2+z^2)^{1/2}-(\tilde{y}^2+z^2)^{1/2}-\frac{(\tilde{x}+\tilde{h})^2}{z}-\beta z+4z\ln 2-2z\\
&\leq|\tilde{k}|+z-(\beta-4\ln 2+2)z\leq 0.
\end{aligned}
$$

This proves the validity of (2.5), since both $\tilde{u}$ and $\tilde{d}$ are nonnegative.

Case IV: $|x+h|>z$, $u+d>v$. The inequality reads

$$
\begin{aligned}
&((y+k)^2+(x+h)^2)^{1/2}(u+d)-(3+\sqrt{2})|x+h|(u+d)\\
&\leq (y^2+z^2)^{1/2}(u+d)-x^2vz^{-1}-\beta zv-4zv\ln 2+y(y^2+z^2)^{-1/2}uk-2xz^{-1}vh.
\end{aligned}
$$

Put all the terms on the left-hand side and note that the obtained sum depends linearly on d. Thus it suffices to show that this sum decreases as d increases (since then the claim follows from Case II). This is equivalent to proving that

$$((y+k)^2+(x+h)^2)^{1/2}-(y^2+z^2)^{1/2}-(3+\sqrt{2})|x+h|\leq 0.$$

But, by the triangle inequality, $((y+k)^2+(x+h)^2)^{1/2}-(y^2+z^2)^{1/2}\leq(k^2+(|x+h|-z)^2)^{1/2}\leq|k|+(|x+h|-z)\leq|h|+|x+h|-z\leq 2|x+h|$. This shows the claim and completes the proof of the lemma. □

Remark 2.1. We briefly describe some of the (informal) steps which led us to the Bellman function B used above. One way to address this problem is to use the abstract approach, i.e., write the abstract formula for the Bellman function associated with (1.4) and try to solve the underlying partial differential equation (of Monge-Ampére type) and/or find the extremizers. This type of reasoning has turned out to be very efficient in a number of problems (see e.g. [12], [20], [22]). However, in our current case we have not been able to apply this approach successfully and we had to use a different path.

Our motivation comes from the unweighted setting. A natural starting point is to consider the special function constructed by Burkholder in [4] to establish the sharp version of the un-weighted inequality

$$\sup_{n\geq 0} ||g_n||_{L^1} \leq C||f^*||_{L^1}. \tag{2.8}$$

However, this function is very complicated and does not seem to extend nicely to the weighted case. Fortunately, the function corresponding to the version of (2.8) for continuous-time, continuous-path martingales, constructed by Osękowski in [15] is much simpler. It is given by

$$b(x,y,z) = \frac{y^2-x^2}{z} - z.$$

This is a function of three variables since no weights are involved. This object suggests that in the weighted realm, the candidate should be of the form

$$B(x,y,z,u,v) = \frac{y^2u - x^2v}{z} - \beta zv,$$

for some positive constant β to be specified. The bad news is that this object cannot possibly satisfy the key estimate (2.5). Indeed, the right-hand side depends linearly on k, while the left-hand side is of order $O(k^2)$ as $k \to \infty$. This indicates that the dependence on the variable y in the function B should be linear, which leads us to the choice

$$B(x,y,z,u,v) = |y|u - x^2z^{-1}v - \beta zv.$$

This function does not work either: it is not even of class C^1, so the right-hand side of (2.5) does not make sense. To smoothen the cusp at $y = 0$, some further thought and analysis leads to the function

$$B(x,y,z,u,v) = (y^2+z^2)^{1/2}u - x^2z^{-1}v - \beta zv.$$

Again, the crucial estimate (2.5) might fail, if d and v are large: both sides depend linearly on d (at least when $u + d \leq v$), but the coefficient in front of d on the left-hand side may be larger. To solve this issue, we need to add to B a term which depends on u in a concave manner. After a lot of experimentation, one arrives at the "correction" $4zv\ln(uv^{-1} + 1)$, which turns out to work just fine.

3. Proofs of main results

Proof of (1.4). Fix martingales $f = (f_n)_{n\geq 0}$, $g = (g_n)_{n\geq 0}$ and a weight $w = (w_n)_{n\geq 0}$ as in the statement. Clearly, we may and do assume that $\mathbb{E}MfMw < \infty$, since otherwise there is nothing to prove. Furthermore, we may assume that the weight w is strictly positive, by adding a small positive ε to w, and letting $\varepsilon \to 0$ at the very end.

The key part of the proof is to show that the sequence $(B(H_n))_{n\geq 0}$ is a supermartingale, where, for brevity, we have set $H_n = (f_n, g_n, M_nf, w_n, M_nw)$, $n = 0, 1, 2, \ldots$. To do this, take $n \geq 0$ and note that by (2.5),

$$\begin{aligned}
&B(H_{n+1})\\
&= B(f_n + df_{n+1}, g_n + dg_{n+1}, |f_n + df_{n+1}| \vee M_nf, w_n + dw_{n+1}, (w_n + dw_{n+1}) \vee M_nw)\\
&\leq B(H_n) + B_x(H_n)df_{n+1} + B_y(H_n)dg_{n+1} + B_u(H_n)dw_{n+1}.
\end{aligned}$$

Now, observe that both sides are integrable. This easily follows from the estimates (2.3), (2.4) and the assumption $\mathbb{E}MfMw < \infty$ we have imposed at the beginning.Thus, taking the conditional expectation with respect to $\mathcal{F}_n$ yields

$$\mathbb{E}\big[B(H_{n+1})|\mathcal{F}_n\big] \leq B(H_n),$$

since $B_x(H_n)$, $B_y(H_n)$, $B_u(H_n)$ are $\mathcal{F}_n$-measurable and df_{n+1}, dg_{n+1} and dw_{n+1} are martingale differences. This establishes the supermartingale property of the sequence $(B(H_n))_{n\geq 0}$ and hence, by (2.1) and (2.2), we see that

$$\mathbb{E}|g_n|w_n - C\mathbb{E}M_nfM_nw \leq \mathbb{E}B(H_n) \leq \mathbb{E}B(H_0) = \mathbb{E}B(f_0, g_0, |f_0|, w_0, w_0) \leq 0.$$

Since $w_n = \mathbb{E}(w|\mathcal{F}_n)$, the above estimate implies $\mathbb{E}|g_n|w \leq C\mathbb{E}MfMw$, and taking the supremum over all $n \geq 0$ completes the proof. □

Proof of Theorem 1.2. Fix $r < 1$ and a large positive integer N and consider the weight $w = 2^{N/r}\chi_{[0,2^{-N})}$. To compute $M_d(w^r)$, we easily check that the martingale $(v_n)_{n\geq 0} = (\mathbb{E}(w^r|\mathcal{F}_n))_{n\geq 0}$ is given as follows. We have $v_0 = \chi_{[0,1)}$, $v_{2^k} = v_{2^k+1} = \ldots = v_{2^{k+1}-1} = 2^{k+1}\chi_{[0,2^{-1-k})}$ for $k = 0, 1, \ldots, N-1$, and $v_n = w^r$ for $n \geq 2^N$. This implies

$$M_d(w^r) = \sum_{n=0}^{N-1} 2^n\chi_{[2^{-1-n},2^{-n})} + 2^N\chi_{[0,2^{-N})}. \tag{3.1}$$

Next, let $f, g : [0,1) \to \mathbb{R}$ be given by

$$f = \frac{1}{3}h_0 - \frac{2}{3}\sum_{k=0}^{N-1}(-1)^k h_{2^k} \quad \text{and} \quad g = \frac{1}{3}h_0 + \frac{2}{3}\sum_{k=0}^{N-1} h_{2^k}.$$

Clearly, $g = T_\varepsilon f$ for an appropriate choice $\varepsilon = (\varepsilon_n)_{n\geq 0}$ of signs. We easily check that f is bounded by 1; actually, we have $|f| \equiv 1$ on $[2^{-N}, 1)$ and $|f| = 1/3$ on $[0, 2^{-N})$. On the other hand, on $[0, 2^{-N})$ we have $g = 1/3 + N \cdot 2/3$. Combining these observations with (3.1), we derive that

$$\mathbb{E}gw \geq \left(\frac{1}{3} + \frac{2N}{3}\right) \cdot 2^{N(1/r-1)}$$

and

$$\begin{aligned}\mathbb{E}M_df\,(M_d(w^r))^{1/r} \leq \mathbb{E}(M_d(w^r))^{1/r} &= \sum_{n=0}^{N-1} 2^{n(1/r-1)-1} + 2^{N(1/r-1)} \\ &= \frac{2^{N(1/r-1)}-1}{2(2^{1/r-1}-1)} + 2^{N(1/r-1)} \leq 2^{N(1/r-1)}K_r,\end{aligned}$$

where K_r depends only on r. This clearly yields the assertion, since N was arbitrary. □

4. A maximal inequality for dyadic shifts

The martingale estimate studied in the preceding sections can be used to obtain a related result for a certain class of dyadic shifts which, in turn, leads to the corresponding bound for a class of Calderón-Zygmund singular integral operators. Our starting point is the following lemma. Here by $\sigma(A_1, A_2, \ldots, A_n)$ we denote the σ-algebra of subsets of $[0,1)$ generated by the sets A_1, A_2, …, A_n and $(h_n)_{n\geq 0}$ is the Haar system.

Lemma 4.1. *For any a_1, a_2, $a_3 \in \mathbb{R}$, let $f = a_1h_1 + a_2h_2 + a_3h_3$ and $g = a_1h_2$. Then there is a filtration $(\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3)$ satisfying the conditions $\mathcal{F}_0 = \{[0,1), \emptyset\}$ and $\mathcal{F}_3 = \sigma([0,1/4), [1/4,1/2), [1/2,3/4))$, such that the martingale $(g_n)_{n=0}^3$ associated with g is differentially subordinate to the martingale $(2f_n)_{n=0}^3$ associated with $2f$.*

Proof. By homogeneity, we may and do assume that $a_1 = 1$. We consider two cases.

The case $a_2 \in (-\infty, -3/2] \cup [0, 3/2)$. For such a_2, we take $\mathcal{F}_1 = \sigma([0, 1/4))$ and $\mathcal{F}_2 = \sigma([0, 1/4), [1/4, 1/2))$. We easily compute that $f_0 = g_0 = 0$,

$$
\begin{aligned}
f_1 &= (1 + a_2)\chi_{[0,1/4)} - \frac{1 + a_2}{3}\chi_{[1/4,1)},\\
f_2 &= (1 + a_2)\chi_{[0,1/4)} + (1 - a_2)\chi_{[1/4,1/2)} - \chi_{[1/2,1)},\\
f_3 &= f,\\
g_1 &= \chi_{[0,1/4)} - \frac{1}{3}\chi_{[1/4,1)},\\
g_2 &= g_3 = g.
\end{aligned}
$$

We are ready to verify the differential subordination. Clearly, $|dg_0| \leq |2df_0|$, since both differences vanish. Next, we have $df_1 = f_1$ and $dg_1 = g_1$, so the inequality $|dg_1| \leq 2|df_1|$ is equivalent to the estimate $1 \leq 2|1 + a_2|$, which holds due to the assumption on the range of a_2. Now, we derive that

$$
df_2 = \left(\frac{4}{3} - \frac{2a_2}{3}\right)\chi_{[1/4,1/2)} - \left(\frac{2}{3} - \frac{a_2}{3}\right)\chi_{[1/2,1)}, \qquad dg_2 = -\frac{2}{3}\chi_{[1/4,1/2)} + \frac{1}{3}\chi_{[1/2,1)}
$$

and the inequality $|dg_2| \leq |2df_2|$ can be rewritten as $\frac{2}{3} \leq 2|\frac{4}{3} - \frac{2a_2}{3}|$, or $|a_2 - 2| \geq 1/2$. The latter bound holds true because of the assumptions on a_2. Finally, we have $dg_3 = 0$, so the condition $|dg_3| \leq 2|df_3|$ is evident.

The case $a_2 \in [-3/2, 0) \cup [3/2, \infty)$. In this case we take $\mathcal{F}_1 = \sigma([1/4, 1/2))$ and $\mathcal{F}_2 = \sigma([0, 1/4), [1/4, 1/2))$. The remaining analysis is the same as in the preceding case: we have

$$
\begin{aligned}
f_1 &= (1 - a_2)\chi_{[1/4,1/2)} - \frac{1 - a_2}{3}\chi_{[0,1/4)\cup[1/2,1)},\\
f_2 &= (1 + a_2)\chi_{[0,1/4)} + (1 - a_2)\chi_{[1/4,1/2)} - \chi_{[1/2,1)},\\
f_3 &= f,\\
g_1 &= -\chi_{[1/4,1/2)} + \frac{1}{3}\chi_{[0,1/4)\cup[1/2,1)},\\
g_2 &= g_3 = g
\end{aligned}
$$

and the differential subordination follows from the estimates $2|1 - a_2| \geq 1$ and $|2 + a_2| \geq 1/2$, which are guaranteed by the assumption on the range of a_2. $\square$

Of course, the assertion of the above lemma remains valid if we change the definition of g to $a_1 h_3$. This, by the self-similarity of the Haar system, allows us to obtain an important corollary. For the sake of convenience, we will use a slightly different notation for the Haar system: given a dyadic interval I, we denote by I_- and I_+ its left and right halves, respectively, and set $h_I = |I|^{-1/2}(\chi_{I_-} - \chi_{I_+})$. For any dyadic interval I, choose arbitrarily one of its halves and denote it by $a(I)$ (it may happen that for some I's we choose the left half, for other I's - the right half). Define the associated "odd" and "even" dyadic shifts by

$$
S^o f(a) = \sum_I \langle f, h_I\rangle h_{a(I)}, \qquad S^e f(a) = \sum_I \langle f, h_I\rangle h_{a(I)},
$$

where $\langle f, h_I\rangle = \int_0^1 f h_I \mathrm{d}x$ is the scalar product of f and h_I in $L^2(0, 1)$, and the summations run over all dyadic intervals I such that $\log_2 |I|$ is odd/even, respectively.

Corollary 4.1. *Fix a function f of finite Haar expansion, an arbitrary "selection function" a and let $g = S^o f(a)$ or $g = S^e f(a)$. Then there is a finite filtration $(\mathcal{F}_n)_{n=0}^N$ of subsets of $[0,1)$ and two adapted martingales $(f_n)_{n=0}^N$, $(g_n)_{n=0}^N$ such that $f_N = f$, $g_N = g$ and $(g_n)_{n=0}^N$ is differentially subordinate to $(2f_n)_{n=0}^N$.*

Proof. We will show the claim for "odd" shifts only. Let us start with setting $\mathcal{F}_0 = \{\emptyset, \Omega, [0,1/2), [1/2,1)\}$. The remaining σ-algebras of the filtration are constructed in triples: $\{\mathcal{F}_1, \mathcal{F}_2, \mathcal{F}_3\}$, $\{\mathcal{F}_4, \mathcal{F}_5, \mathcal{F}_6\}$, and so on. To explain the construction of the triple $\{\mathcal{F}_{3n+1}, \mathcal{F}_{3n+2}, \mathcal{F}_{3n+3}\}$ for a given $n \geq 0$, consider the family $\mathcal{I}$ which consists of all dyadic intervals I with $|I| = 2^{-2n-1}$. If $\langle f, h_I\rangle = 0$ for all I with $|I| < 2^{-2n-1}$, we stop the construction (i.e., we set $\mathcal{F}_{3n+1} = \mathcal{F}_{3n+2} = \mathcal{F}_{3n+3} = \mathcal{F}_{3n}$. Otherwise, for each $I \in \mathcal{I}$, we let

$$\mathtt{f}^I := \langle f, h_I\rangle h_I + \langle f, h_{I_-}\rangle h_{I_-} + \langle f, h_{I_+}\rangle h_{I_+}, \qquad \mathtt{g}^I = \langle f, h_I\rangle h_{a(I)}$$

be the parts of expansions of f and $S^o f(a)$, corresponding to the interval I and its first generation. Now we apply the previous lemma conditionally on the interval I. This interval can be filtered with the use of three σ-algebras $\mathcal{F}_0(I) = \{\emptyset, I\}$, $\mathcal{F}_1(I)$, $\mathcal{F}_2(I)$, $\mathcal{F}_3(I) = \sigma(I_{--}, I_{-+}, I_{+-}, I_{++})$ such that the martingale corresponding to $\mathtt{g}^I$ is differentially subordinate to the martingale induced by $2\mathtt{f}^I$. The crucial observation is that the filtration $\mathcal{F}_i(I)$ affects only I and its first generation. Therefore if we set

$$\mathcal{F}_{3n+1} = \sigma(\mathcal{F}_1(J) : J \in \mathcal{I}),\ \mathcal{F}_{3n+2} = \sigma(\mathcal{F}_2(J) : J \in \mathcal{I}),\ \mathcal{F}_{3n+3} = \sigma(\mathcal{F}_3(J) : J \in \mathcal{I}),$$

then $\mathcal{F}_{3n+3} = \sigma(J : |J| = 2^{-2n-3})$. In other words, the σ-algebras are properly ordered (i.e., they indeed form a filtration). Furthermore, directly from the construction, we see that if $k = 3n + i$, $i = 1, 2, 3$, then on $I \in \mathcal{I}$ we have

$$|dg_k| = |d\mathtt{g}_i^I| \leq |2d\mathtt{f}_i^I| = |2df_k|,$$

which proves the desired differential subordination. □

Let S denotes the usual dyadic shift introduced by Petermichl [17]:

$$Sf = \sum_I \langle f, h_I\rangle (h_{I_-} - h_{I_+}),$$

where the summation runs over all dyadic subintervals of $[0,1)$. Note that this operator can be expressed as a combination of four shifts of the above type: two odd shifts

$$\sum_{I:\log_2 |I| \text{ odd}} \langle f, h_I\rangle h_{I_-} - \sum_{I:\log_2 |I| \text{ odd}} \langle f, h_I\rangle h_{I_+}$$

and two even shifts

$$\sum_{I:\log_2 |I| \text{ even}} \langle f, h_I\rangle h_{I_-} - \sum_{I:\log_2 |I| \text{ even}} \langle f, h_I\rangle h_{I_+}.$$

Combining this representation with Theorem 1.1 and Corollary 4.1, we get the following statement. In what follows, M is an uncentered maximal operator of Hardy and Littlewood.

Theorem 4.1. *Let C be the constant of Theorem 1.1. For any function f and any weight w on $[0,1)$ we have*

$$||S^o f||_{L^1(w)} \leq 2C||Mf||_{L^1(Mw)}, \qquad ||S^e f||_{L^1(w)} \leq 2C||Mf||_{L^1(Mw)}$$

and consequently,

$$||Sf||_{L^1(w)} \leq 8C||Mf||_{L^1(Mw)}.$$

Standard scaling argument shows that the above statement holds true, with unchanged constants, if we consider the Haar system (and the associated dyadic shifts) on $\mathbb{R}$. Such shifts, after rescaling, translation and averaging-type operations, lead to a Hilbert transform $\mathcal{H}$ on the real line, as Petermichl showed in [17] (see also the work of Hytönen [6], which shows explicitly that Hilbert transform is equal to $-8\ln 2/(\pi\sqrt{2}) = -1.2481\ldots$ times the average of such shifts). Thus, the above statement implies the version of (1.2) for this operator.

Corollary 4.2. *Let C be the constant of Theorem 1.1. For any f and any weight w on $\mathbb{R}$ we have*

$$||\mathcal{H}f||_{L^1(w)} \leq \frac{64C\ln 2}{\pi\sqrt{2}}||Mf||_{L^1(Mw)} \leq 72||Mf||_{L^1(Mw)}. \tag{4.1}$$

Using the method of rotation, the inequality for the Hilbert transform yields the same for the class of singular integral operators with odd kernels. For a given odd function $\Omega \in L^1(S^{d-1})$, consider the associated singular integral operator T_Ω given by

$$T_\Omega f(x) = \text{p.v.} \int_{\mathbb{R}^d} \frac{\Omega(y/|y|)}{|y|^d} f(x-y)\mathrm{d}y.$$

The Riesz transforms

$$R_j f(x) = \frac{\Gamma(\frac{d+1}{2})}{\pi^{(d+1)/2}}\text{p.v.} \int_{\mathbb{R}^d} \frac{y_j}{|y|^{d+1}} f(x-y)\,dy, \quad j = 1, \ldots, d,$$

are the classical examples of such singular integrals.

Though the result below is well-known by now, we believe that it is a nice application of the above probabilistic considerations which also yields explicit constants. In particular, it gives constants independent of the dimension d for Riesz transforms.

Theorem 4.2. *For any f and any weight w on $\mathbb{R}^d$ we have*

$$||T_\Omega f||_{L^1(w)} \leq 36\pi||\Omega||_{L^1(S^{d-1})}||Mf||_{L^1(Mw)}.$$

In particular, for the Riesz transforms we have

$$||R_j f||_{L^1(w)} \leq 72||Mf||_{L^1(Mw)}, \quad j = 1, \ldots, d.$$

Proof. We exploit the classical method of rotations. For $f : \mathbb{R}^d \to \mathbb{R}$, define the directional Hilbert transform in the direction θ by

$$\mathcal{H}_\theta f(x) = \frac{1}{\pi}\text{p.v.} \int_{\mathbb{R}} f(x - t\theta)\frac{\mathrm{d}t}{t}. \tag{4.2}$$

A straightforward combination of Fubini's theorem and (4.1) implies that

$$||\mathcal{H}_\theta f||_{L^1(w)} \leq 72||Mf||_{L^1(Mw)}.$$

However, it is well-known [21] that T_Ω is an average of directional Hilbert transforms:

$$T_\Omega f(x) = \frac{\pi}{2}\int_{S^{d-1}} \Omega(\theta)\mathcal{H}_\theta f(x)\mathrm{d}\theta.$$

Combining this representation with (4.2) yields the claim. □

Remark 4.3. There is a very interesting question whether there are some analogs of Lemma 4.1 and Corollary 4.1 which would enable the representation of general dyadic shifts via differentially subordinate martingales (see. e.g. [7] for the necessary definitions). We do not know the asnwer to this question. It the answer is affirmative, then Theorem

1.1 would give an alternative proof of the estimate (1.2) for general Calderón-Zygmund singular integrals.

ACKNOWLEDGMENTS

The authors would like to thank an anonymous referee for a very thorough reading of the first version of this article and many helpful comments that greatly improved the presentation.

REFERENCES

[1] R. Bañuelos and G. Wang, *Sharp inequalities for martingales with applications to the Beurling-Ahlfors and Riesz transformations*, Duke Math. J. **80** (1995), pp. 575–600.

[2] D. L. Burkholder, *Boundary value problems and sharp inequalities for martingale transforms*, Ann. Probab. **12** (1984), pp. 647–702.

[3] D. L. Burkholder, *Explorations in martingale theory and its applications*, École d'Ete de Probabilités de Saint-Flour XIX—1989, pp. 1–66, Lecture Notes in Math., **1464**, Springer, Berlin, 1991.

[4] D. L. Burkholder, *Sharp norm comparison of martingale maximal functions and stochastic integrals*, Proceedings of the Norbert Wiener Centenary Congress, 1994 (edited by V. Mandrekar and P. R. Masani), Proceedings of Symposia in Applied Mathematics **52** (1997), pp. 343–358.

[5] C. Fefferman and E.M. Stein, *Some maximal inequalities*, Amer. J. Math. **93** (1971), pp. 107–115.

[6] T. Hytönen, *On Petermichl's dyadic shift and the Hilbert transform*, C. R. Math. Acad. Sci. Paris **346** (2008), no. 21-22, pp. 1133–1136.

[7] T. Hytönen, C. Pérez, S. Treil and A. Volberg, *Sharp weighted estimates for dyadic shifts and the A_2 conjecture*, J. Reine Angew. Math. **687** (2014), pp. 43-Ű86.

[8] A. K. Lerner, S. Ombrosi and C. Pérez, *Sharp A_1 bounds for Calderón-Zygmund operators and the relationship with a problem of Muckenhoupt and Wheeden*, Int. Math. Res. Not. IMRN **6** (2008), Art. ID rnm161, 11 pp.

[9] A. K. Lerner, S. Ombrosi, C. Pérez, *A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden*, Math. Res. Lett. **16** (2009), pp. 149–156.

[10] F. L. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg, *Weak norm estimates of weighted singular operators and Bellman functions*. Manuscript (2010).

[11] F. L. Nazarov and S. R. Treil, *The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis,* St. Petersburg Math. J. **8** (1997), pp. 721–824.

[12] F. L. Nazarov, S. R. Treil and A. Volberg, *The Bellman functions and two-weight inequalities for Haar multipliers*, J. Amer. Math. Soc., **12** (1999), pp. 909–928.

[13] F. L. Nazarov, S. R. Treil and A. Volberg, *Bellman function in stochastic optimal control and harmonic analysis (how our Bellman function got its name),* Oper. Theory: Adv. Appl. **129** (2001), pp. 393–424.

[14] A. Osękowski, *Sharp martingale and semimartingale inequalities*, Monografie Matematyczne **72** (2012), Birkhäuser, 462 pp.

[15] A. Osękowski, *Maximal inequalities for continuous martingales and their differential subordinates*, Proc. Amer. Math. Soc. **139** (2011), pp. 721–734.

[16] A. Osękowski, *Maximal inequalities for martingales and their differential subordinates*, J. Theoret. Probab. **27** (2014), pp. 1–21.

[17] S. Petermichl, *Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol*, C. R. Acad. Sci. Paris Ser. I Math. **330** (2000), no. 6, 455–460 (English, with English and French summaries).

[18] M. C. Reguera, *On Muckenhoupt-Wheeden conjecture*, Adv. Math. 227 (2011), no. 4, pp. 1436–1450.

[19] M. C. Reguera and C. Thiele, *The Hilbert transform does not map $L^1(Mw)$ to $L^{1,\infty}(w)$*, Math. Res. Lett. **19** (2012), no. 1, pp. 1–7.

[20] L. Slavin, V. Vasyunin, *Sharp results in the integral-form John-Nirenberg inequality*, Trans. Amer. Math. Soc. **363** (2011), 4135–4169.

[21] E. M. Stein, *Singular integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.

[22] V. Vasyunin, A. Volberg, *Monge-Ampére equation and Bellman optimization of Carleson Embedding Theorems*, Amer. Math. Soc. Transl. (2), vol. 226, "Linear and Complex Analysis", 2009, pp. 195–238.

DEPARTMENT OF MATHEMATICS, PURDUE UNIVERSITY, WEST LAFAYETTE, IN 47907, USA
E-mail address: banuelos@math.purdue.edu

DEPARTMENT OF MATHEMATICS, INFORMATICS AND MECHANICS, UNIVERSITY OF WARSAW, BANACHA 2, 02-097 WARSAW, POLAND
E-mail address: ados@mimuw.edu.pl