BEST CONSTANTS IN THE WEAK TYPE INEQUALITIES FOR A MARTINGALE CONDITIONAL SQUARE FUNCTION

ADAM OSĘKOWSKI

ABSTRACT. We determine the optimal constants in the weak type (p, q) inequalities involving a martingale, its square and conditional square function. As an application, we present related bounds for predictable projections of adapted sequences of random variables and certain estimates related to Khintchine inequalities.

1. INTRODUCTION

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, filtered by $(\mathcal{F}_n)_{n\geq 0}$, a non-decreasing family of sub-σ-fields of $\mathcal{F}$. Throughout the paper, $f = (f_n)_{n\geq 0}$ will stand for a martingale adapted to $(\mathcal{F}_n)_{n\geq 0}$ and taking values in a certain separable Hilbert space $\mathcal{H}$ with norm $|\cdot|$ and scalar product $\langle\cdot,\cdot\rangle$. Let $df = (df_n)_{n\geq 0}$ denote the difference sequence of f, defined by the equations $f_n = \sum_{k=0}^n df_k$, $n = 0, 1, 2, \ldots$. Then $S(f)$, the *square function* of f, and $s(f)$, the *conditional square function* of f, are given by

$$S(f) = \left[\sum_{k=0}^{\infty} ||df_k||^2\right]^{1/2} \quad \text{and} \quad s(f) = \left[\sum_{k=0}^{\infty} \mathbb{E}(||df_k||^2|\mathcal{F}_{k-1})\right]^{1/2}.$$

Here and below, $\mathcal{F}_{-1}$ is assumed to be equal to $\mathcal{F}_0$. We will also use the notation

$$S_n(f) = \left[\sum_{k=0}^{n} ||df_k||^2\right]^{1/2} \quad \text{and} \quad s_n(f) = \left[\sum_{k=0}^{n} \mathbb{E}(||df_k||^2|\mathcal{F}_{k-1})\right]^{1/2},$$

and, furthermore, when $0 < p < \infty$, we will write

$$||f||_p = \sup_n ||f_n||_p = \sup_n (\mathbb{E}||f_n||^p)^{1/p},$$

$$||f||_{p,\infty} = \sup_n ||f_n||_{p,\infty} = \sup_n \sup_{\lambda>0} \lambda\big[\mathbb{P}(||f_n|| \geq \lambda)\big]^{1/p}.$$

and define the norms $||S(f)||_p$, $||S(f)||_{p,\infty}$ and $||s(f)||_p$, $||s(f)||_{p,\infty}$ analogously.

The purpose of this paper is to determine optimal constants in some inequalities involving f, $S(f)$ and $s(f)$. We begin with recalling related results from the literature. The sharp estimates

$$(1.1) \qquad ||s(f)||_p \leq \sqrt{\frac{p}{2}}||f||_p, \qquad ||s(f)||_p \leq \sqrt{\frac{p}{2}}||S(f)||_p,$$

2000 *Mathematics Subject Classification.* Primary: 60G42. Secondary: 60G44.

Key words and phrases. Martingale, square function, conditional square function, weak type inequality.

Partially supported by MNiSW Grant N N201 397437.

for $2 \leq p < \infty$, and

$$||f||_p \leq \sqrt{\frac{2}{p}}||s(f)||_p, \qquad ||S(f)||_p \leq \sqrt{\frac{2}{p}}||s(f)||_p, \tag{1.2}$$

for $0 < p \leq 2$, were established by [21], who also showed that the inequalities fail to hold for the remaining values of p, even if the martingales are assumed to be real-valued. However, [12] proved that if $1 < p < 2$ and f is real-valued, then, for some absolute C_p,

$$C_p^{-1}||f||_p \leq \inf\left\{ ||s(g)||_p + \left|\left| \left(\sum_{k=0}^\infty |dh_k|^p\right)^{1/p}\right|\right|_p \right\} \leq C_p||f||_p, \tag{1.3}$$

where the infimum runs over all possible decompositions of f as a sum $f = g + h$ of two martingales. This estimate can be regarded as a dual version of the Burkholder-Rosenthal inequality (see [2], [11], [12] for details): for $2 \leq p < \infty$,

$$C_p^{-1}||f||_p \leq \max\left\{ ||s(f)||_p\,, \left|\left| \left(\sum_{k=0}^\infty |df_k|^p\right)^{1/p}\right|\right|_p \right\} \leq C_p||f||_p. \tag{1.4}$$

By Burkholder-Davis-Gundy inequalities (see [2]), the above bounds are still valid (possibly with different C_p) when $||f||_p$ is replaced by $||S(f)||_p$. The best constants in the related weak type (p,p) estimates were found by [16]: for $0 < p \leq 2$,

$$\begin{aligned} ||f||_{p,\infty} &\leq \left(\Gamma\left(\frac{p}{2}+1\right)\right)^{-1/p}||s(f)||_p,\\ ||S(f)||_{p,\infty} &\leq \left(\Gamma\left(\frac{p}{2}+1\right)\right)^{-1/p}||s(f)||_p. \end{aligned} \tag{1.5}$$

On the other hand, if $2 \leq p < \infty$,

$$||s(f)||_{p,\infty} \leq \left(\frac{p}{2}\right)^{1/2-1/p}||S(f)||_p, \quad ||s(f)||_{p,\infty} \leq \left(\frac{p}{2}\right)^{1/2-1/p}||f||_p.$$

We will study such estimates in the case when the weak moments on the left and the strong moments on the right are of different order. For $0 < q \leq p < \infty$, let

$$K_{p,q} = \begin{cases} 1 & \text{if } q \leq 2 \leq p,\\ \left(\Gamma\left(\frac{p}{2}+1\right)\right)^{-1/p} & \text{if } 0 < q \leq p < 2,\\ \infty & \text{if } 2 < q \leq p < \infty, \end{cases}$$

$$L_{p,q} = \begin{cases} 1 & \text{if } q \leq 2 \leq p,\\ \infty & \text{if } 0 < q \leq p < 2,\\ \left(\frac{q}{2}\right)^{1/2-1/p}\left(\frac{p-q}{p-2}\right)^{(p/2-1)(1/q-1/p)} & \text{if } 2 < q \leq p < \infty, \end{cases}$$

The result of the paper can be stated as follows.

Theorem 1.1. *For any $0 < q \leq p < \infty$ and any $\mathcal{H}$-valued martingale f we have*

$$||f||_{q,\infty} \leq K_{p,q}||s(f)||_p, \qquad ||S(f)||_{q,\infty} \leq K_{p,q}||s(f)||_p. \tag{1.6}$$

and

$$||s(f)||_{q,\infty} \leq L_{p,q}||f||_p, \qquad ||s(f)||_{q,\infty} \leq L_{p,q}||S(f)||_p. \tag{1.7}$$

The constants $K_{p,q}$, $L_{p,q}$ are the best possible, even when $\mathcal{H} = \mathbb{R}$.

To complete the description of the optimal constants, note that if $q > p$, then the inequalities (1.6), (1.7) do not hold in general with any finite $K_{p,q}$, $L_{p,q}$, even in the real-valued case. This can be seen, for example, by considering a constant martingale $f_0 = f_1 = f_2 = \ldots$ with $f_0 \in L^p \setminus L^{q,\infty}$.

The estimates (1.6) and (1.7) will be established in the next section. Our approach is based on a certain version of Burkholder's method which relates the validity of a given martingale inequality to the existence of a certain special function (see e.g. [4], [16] and [21]). In Section 3 we construct examples to show the optimality of the constants $K_{p,q}$ and $L_{p,q}$. The final part of the paper contains some applications.

2. Proof of (1.6) and (1.7)

2.1. **Proof of** (1.6)**.** It suffices to establish the estimate for $q \leq 2$. If $p \geq 2$, then

$$||f||_{q,\infty} \leq ||f||_q \leq ||f||_2 = ||s(f)||_2 \leq ||s(f)||_p = K_{p,q}||s(f)||_p,$$

with a similar reasoning for $||S(f)||_{q,\infty} \leq K_{p,q}||s(f)||_p$. If $p < 2$, then by (1.5),

$$||f||_{q,\infty} \leq ||f||_{p,\infty} \leq \left(\Gamma\left(\frac{p}{2}+1\right)\right)^{-1/p} ||s(f)||_p = K_{p,q}||s(f)||_p$$

and the analogous argumentation gives the second bound in (1.6).

2.2. **Proof of** (1.7)**.** As previously, we may focus on the case $2 < q \leq p$, since for other choices of p and q the bound is trivial. The key ingredient of the proof is the following related estimate; we shall see later how to deduce (1.7) from it.

Theorem 2.1. *Suppose that $p > 2$ and $\lambda \in [0, 1-2/p]$. Then for any $\mathcal{H}$-valued martingale f we have*

$$\begin{aligned} \mathbb{P}(s(f) \geq 1) &\leq \left(\frac{p}{2}\right)^{p/2-1} (1-\lambda)^{p/2}||f||_p^p + \left(\frac{p}{p-2}\right)^{p/2-1} \lambda^{p/2}, \\ \mathbb{P}(s(f) \geq 1) &\leq \left(\frac{p}{2}\right)^{p/2-1} (1-\lambda)^{p/2}||S(f)||_p^p + \left(\frac{p}{p-2}\right)^{p/2-1} \lambda^{p/2}. \end{aligned} \tag{2.1}$$

Both estimates are sharp, even in the real-valued setting.

Here by sharpness we mean that neither of the constants $(p/2)^{p/2-1}(1-\lambda)^{p/2}$, $(p/(p-2))^{p/2-1}\lambda^{p/2}$ can be replaced by a smaller number. Fix p and λ as in the statement and put $r = p/2$. Let $\gamma = \gamma_{p,\lambda} : [1-(r(1-\lambda))^{-1}, 1) \to \mathbb{R}$ be given by

$$\gamma(y) = (r(1-\lambda))^{-r/(r-1)}(1-y)^{-1/(r-1)}$$

and introduce the subsets D_1, D_2, …, D_5 of $[0,\infty) \times [0,\infty)$ by setting

$$\begin{aligned} D_1 &= \left\{(x,y) : y \leq 1-\big(r(1-\lambda)\big)^{-1},\, x \leq y(r-1)^{-1} + \lambda\big((1-\lambda)(r-1)\big)^{-1}\right\}, \\ D_2 &= \left\{(x,y) : 1-\big(r(1-\lambda)\big)^{-1} < y < 1, \quad y-1+\gamma(y) < x < \gamma(y)\right\}, \\ D_3 &= \left\{(x,y) : 1-\big(r(1-\lambda)\big)^{-1} < y < 1, \quad x \leq y-1+\gamma(y)\right\}, \\ D_4 &= [0,\infty) \times [1,\infty), \\ D_5 &= [0,\infty) \times [0,\infty) \setminus \big(D_1 \cup D_2 \cup D_3 \cup D_4\big). \end{aligned}$$

See Figure 1 below.

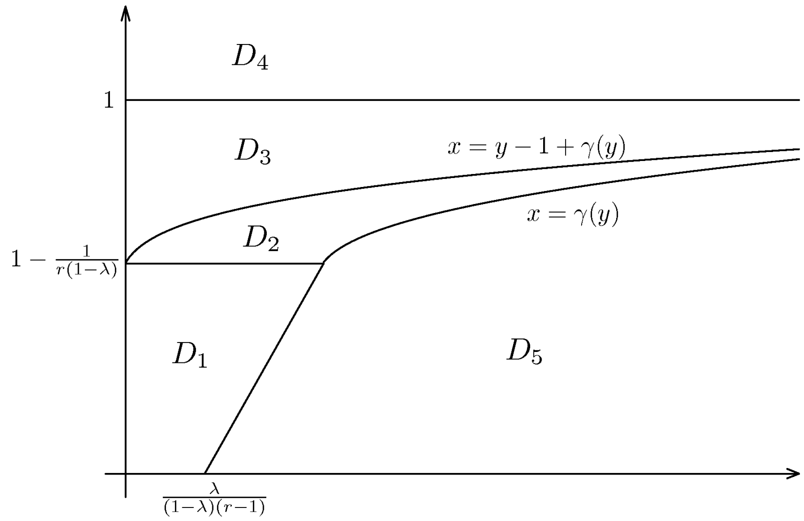

FIGURE 1. The regions D_1–D_5.

Now, define $U = U_{p,\lambda} : [0, \infty) \times [0, \infty) \to \mathbb{R}$ by

$$U(x,y) = \begin{cases} \big(r((1-\lambda)y+\lambda)/(r-1)\big)^{r-1}\big((1-\lambda)(y-rx)+\lambda\big) & \text{if } (x,y) \in D_1, \\ \big(r(1-y)\big)^{-1}\big[(r-1)\gamma(y) - rx\big] & \text{if } (x,y) \in D_2, \\ 1 - r^{r-1}(1-\lambda)^r(1+x-y)^r & \text{if } (x,y) \in D_3, \\ 1 - r^{r-1}(1-\lambda)^r x^r & \text{if } (x,y) \in D_4, \\ -r^{r-1}(1-\lambda)^r x^r & \text{if } (x,y) \in D_5 \end{cases}$$

and let $V = V_{p,\lambda} : [0, \infty) \times [0, \infty) \to \mathbb{R}$ be given by $V(x, y) = 1_{\{y \geq 1\}} - r^r(1-\lambda)^r x^r$.

Let us study the key properties of the functions we have just introduced.

Lemma 2.2. *(i) The function U is continuous on $[0, \infty) \times [0, \infty)$. Furthermore, U_x is continuous on $(0, \infty) \times (0, \infty)$ and U_y is continuous on $(0, \infty) \times ((0, \infty) \setminus \{1\})$.*

(ii) For any fixed y, the function $U(\cdot, y)$ is concave on $[0, \infty)$.

(iii) For any $y \geq 0$, the function $t \mapsto U(t, y + t)$ is nonincreasing on $[0, \infty)$.

(iv) We have the majorization $U \geq V$.

Proof. (i) One easily checks the continuity of U. Next, denoting by A^o the interior of a set A, we derive that

$$U_x(x,y) = \begin{cases} -r^r(r-1)^{1-r}(1-\lambda)\big((1-\lambda)y+\lambda\big)^{r-1} & \text{if } (x,y) \in D_1^o, \\ -(1-y)^{-1} & \text{if } (x,y) \in D_2^o, \\ -r^r(1-\lambda)^r(1+x-y)^{r-1} & \text{if } (x,y) \in D_3^o, \\ -r^r(1-\lambda)^r x^{r-1} & \text{if } (x,y) \in D_4^o \cup D_5^o, \end{cases}$$

$$U_y(x,y) = \begin{cases} \frac{r^r(1-\lambda)^2}{r-1}\left(\frac{(1-\lambda)y+\lambda}{r-1}\right)^{r-2}\left[y - (r-1)x + \frac{\lambda}{1-\lambda}\right] & \text{if } (x,y) \in D_1^o, \\ (1-y)^{-2}\big(\gamma(y) - x\big) & \text{if } (x,y) \in D_2^o, \\ r^r(1-\lambda)^r(1+x-y)^{r-1} & \text{if } (x,y) \in D_3^o, \\ 0 & \text{if } (x,y) \in D_4^o \cup D_5^o, \end{cases}$$

and it remains to verify that the derivatives match appropriately at the common boundaries of the sets D_i. The details are left to the reader.

(ii) The concavity with respect to x is evident if we restrict ourselves to the interiors of D_i. Since U_x is continuous, the property follows.

(iii) Directly from the above formulas for the partial derivatives, we have that

$$U_x(x,y)+U_y(x,y)=\begin{cases} -r^r(1-\lambda)^2\left(\frac{(1-\lambda)y+\lambda}{r-1}\right)^{r-2}x & \text{if } (x,y)\in D_1^o,\\ (1-y)^{-2}(y-1+\gamma(y)-x) & \text{if } (x,y)\in D_2^o,\\ 0 & \text{if } (x,y)\in D_3^o,\\ -r^r(1-\lambda)^r x^{r-1} & \text{if } (x,y)\in D_4^o\cup D_5^o \end{cases}$$

and all the expressions are easily seen to be nonpositive.

(iv) Fix $x>0$ and consider the function $y\mapsto U(x,y)-V(x,y)$. It is easy to see that U_y is nonnegative on its domain, so it suffices to prove that $U(x,0)-V(x,0)\geq 0$ and $U(x,1)-V(x,1)\geq 0$. In both cases we obtain the equality. □

Proof of (2.1). It suffices to focus on the first estimate. To deduce the second one, we proceed as follows. For any $\mathcal{H}$-valued martingale f, consider a sequence $F=(F_n)_{n\geq 0}$, taking values in $\ell^2(\mathcal{H})$, given by the formula

$$F_n=(df_0,df_1,df_2,\ldots,df_n,0,0,\ldots),\qquad n=0,\,1,\,2,\ldots.$$

Then F is also a martingale and satisfies $|F_n|=S_n(f)$ and $s(F)=s(f)$ almost surely. Thus, the second bound in (2.1) is a consequence of the first one.

With no loss of generality, we may assume that $||f||_p<\infty$; otherwise, there is nothing to prove. This assumption will guarantee the integrability of the random variables appearing below. The key part of the proof is to show that the sequence $(U(|f_n|^2,s_n^2(f)))_{n=0}^\infty$ is a supermartingale. To do this, fix $n\geq 0$ and write

$$\mathbb{E}\big[U(|f_{n+1}|^2,s_{n+1}^2(f))|\mathcal{F}_n\big]=\mathbb{E}\big[U(|f_n|^2+2\langle f_n,df_{n+1}\rangle+|df_{n+1}|^2,s_{n+1}^2(f))|\mathcal{F}_n\big].$$

However,

$$\begin{aligned}&|f_n|^2+2\langle f_n,df_{n+1}\rangle+|df_{n+1}|^2\\ &\quad=\Big\{|f_n|^2+\mathbb{E}(|df_{n+1}|^2|\mathcal{F}_n)\Big\}+\Big\{2\langle f_n,df_{n+1}\rangle+|df_{n+1}|^2-\mathbb{E}(|df_{n+1}|^2|\mathcal{F}_n)\Big\}\end{aligned}$$

and the expression in the second parentheses has zero expectation with respect to $\mathcal{F}_n$. Thus, using properties (ii) and (iii) of Lemma 2.2, we see that

$$\begin{aligned}&\mathbb{E}\big[U(|f_{n+1}|^2,s_{n+1}^2(f))|\mathcal{F}_n\big]\\ &\qquad\leq\mathbb{E}\Big[U\Big(|f_n|^2+\mathbb{E}(|df_{n+1}|^2|\mathcal{F}_n),s_n^2(f)+\mathbb{E}(|df_{n+1}|^2|\mathcal{F}_n)\Big)|\mathcal{F}_n\Big]\\ &\qquad\leq\mathbb{E}\big[U\big(|f_n|^2,s_n^2(f)\big)|\mathcal{F}_n\big]=U\big(|f_n|^2,s_n^2(f)\big),\end{aligned}$$

which is the desired supermartingale property. Combining this with the majorization $U\geq V$, we get that for any n,

$$\mathbb{E}V(|f_n|^2,s_n^2(f))\leq\mathbb{E}U(|f_n|^2,s_n^2(f))\leq\mathbb{E}U(|f_0|^2,s_0^2(f)).$$

By property (iii), $U(|f_0|^2,s_0^2(f))=U(|f_0|^2,|f_0|^2)\leq U(0,0)$ almost surely. Plugging this into the preceding estimate and using the definition of V, we obtain

$$\mathbb{P}(s_n(f)\geq 1)\leq\left(\frac{p}{2}\right)^{p/2-1}(1-\lambda)^{p/2}||f||_p^p+\left(\frac{p}{p-2}\right)^{p/2-1}\lambda^{p/2}.$$

Next, fix $\varepsilon > 0$ and note that $\{s(f) \geq 1\} \subseteq \bigcup_{n\geq 0}^{\infty} \{s_n(f) \geq 1-\varepsilon\}$. The events on the right are nondecreasing (with respect to n), so applying the above estimate to the martingale $f/(1-\varepsilon)$, we get

$$\begin{aligned}\mathbb{P}\left(s(f) \geq 1\right) &\leq \lim_{n\to\infty} \mathbb{P}\left(s_n(f) \geq 1-\varepsilon\right)\\ &\leq \left(\frac{p}{2}\right)^{p/2-1} \frac{(1-\lambda)^{p/2}}{(1-\varepsilon)^p}||f||_p^p + \left(\frac{p}{p-2}\right)^{p/2-1} \lambda^{p/2}.\end{aligned}$$

It suffices to use the fact that ε was arbitrary. □

Proof of (1.7). Again, we focus on the estimate involving f; the second inequality, between $S(f)$ and $s(f)$, is obtained similarly or by the use of the $\ell^2(\mathcal{H})$-valued martingale F as above. We will prove that

$$\mathbb{P}(s(f) \geq 1)^{1/q} \leq L_{p,q}||f||_p, \tag{2.2}$$

which yields the claim by homogenization. If $||f||_p \geq 1$, this is obvious (a little calculation shows that $L_{p,q} \geq 1$), so we may assume that $||f||_p < 1$. It is easy to check that the right-hand side of (2.1), as a function of λ, attains its minimum at

$$\lambda = \frac{(p-2)||f||_p^{2p/(p-2)}}{2+(p-2)||f||_p^{2p/(p-2)}},$$

which lies in $[0, 1-2p^{-1}]$ (since $||f||_p \leq 1$). Plugging this value of λ in (2.1) gives

$$\mathbb{P}(s(f) \geq 1) \leq \frac{p^{p/2-1}||f||_p^{p-q}}{(2+(p-2)||f||_p^{2p/(p-2)})^{p/2-1}} \cdot ||f||_p^q.$$

A standard analysis shows that the maximum of the function G, given by

$$G(t) = \frac{p^{p/2-1}t^{p-q}}{(2+(p-2)^{2p/(p-2)})^{p/2-1}}, \qquad t \geq 0,$$

is equal to $L_{p,q}^q$; therefore, we obtain

$$\mathbb{P}(s(f) \geq 1) \leq G(||f||_p)||f||_p^q \leq L_{p,q}^q||f||_p^q,$$

which is (2.2). □

3. Sharpness

3.1. **Sharpness of** (1.6)**.** If $q \leq 2 \leq p$, then equality holds in both estimates for the martingale $f_0 = f_1 = f_2 = \ldots \equiv 1$. If $2 < q \leq p$, then neither of the inequalities holds with a finite constant. Indeed, take f such that $f_0 \equiv 0$,

$$\mathbb{P}(f_1 = -1) = \mathbb{P}(f_1 = 1) = \kappa = 1 - \mathbb{P}(f_1 = 0)$$

for some $\kappa \in (0, 1/2)$, and set $df_2 = df_3 = \ldots \equiv 0$. Then $s(f) = (\mathbb{E}|f_1|^2)^{1/2} = (2\kappa)^{1/2}$ almost surely (here and in all the examples below, we consider the natural filtration of f). Therefore,

$$||f||_{q,\infty} \geq (\mathbb{P}(|f_1| \geq 1))^{1/q} = (2\kappa)^{1/q} = (2\kappa)^{1/q-1/2}||s(f)||_p$$

with the same bound for $||S(f)||_{q,\infty}$. Since the constant in front of $||s(f)||_p$ explodes as $\kappa \to 0$, we see that $K_{p,q} = \infty$ is the best. Finally, when $0 < q \leq p < 2$, fix $\delta \in (0, 1)$ and let $(X_n)_{n=0}^{\infty}$ be a sequence of independent random variables such that

$$\mathbb{P}(X_n = 1) = \delta = 1 - \mathbb{P}(X_n = 0).$$

Furthermore, let $(\varepsilon_n)_{n\geq 0}$ be a sequence of independent Rademacher variables, independent also of (X_n). Introduce the stopping time $\tau=\inf\{n: X_n=1\}$ and set $df_n=\varepsilon_n X_n 1_{\{\tau\geq n\}}$, $n=0,1,2,\ldots$. Then f is a martingale such that $|f_n|\uparrow |f_\infty|\equiv 1$ and $S(f)\equiv 1$ almost surely, so in particular $||f||_{q,\infty}=||S(f)||_{q,\infty}=1$. Furthermore, as $\mathbb{E}(df_n^2|\mathcal{F}_{n-1})=1_{\{\tau\geq n\}}\mathbb{E}X_n=\delta 1_{\{\tau\geq n\}}$, we have $s^2(f)=\delta(\tau+1)$. But τ has geometric distribution, so for $0<p\leq 2$,

$$||s(f)||_p^p=\left|\left|s^2(f)\right|\right|_{p/2}^{p/2}=\delta\sum_{n=1}^\infty(\delta n)^{p/2}(1-\delta)^{n-1}$$

and we see that the right hand side, by choosing δ sufficiently small, can be made arbitrarily close to $\int_0^\infty t^{p/2}e^{-t}dt=\Gamma(p/2+1)$. This completes the proof.

3.2. **Sharpness of** (2.1)**.** It suffices to show the sharpness for λ lying in the interior of $[0,1-2/p]$, the case $\lambda\in\{0,1-2/p\}$ follows easily from the limiting procedure. Fix a positive integer N and let $\delta>0$ be determined by $N\delta=1-2(p(1-\lambda))^{-1}$. Consider a sequence $X_1, X_2, \ldots, X_{N+1}$ of independent random variables such that

$$p_n:=\mathbb{P}\left(X_n=\frac{2n\delta}{p-2}+\frac{2\lambda}{(1-\lambda)(p-2)}\right)=1-\mathbb{P}(X_n=0)=\frac{(p-2)(1-\lambda)\delta}{2(n(1-\lambda)\delta+\lambda)}$$

for $n=1,2,\ldots,N$, and put $X_{N+1}\equiv 1-\frac{2}{p(1-\lambda)}$. Next, let $\varepsilon_1,\varepsilon_2,\ldots,\varepsilon_{N+1}$ be independent Rademacher variables, independent also from the sequence $(X_n)_{n=1}^{N+1}$. Let $\tau=\inf\{n: X_n\neq 0\}$ and define $f=(f_n)_{n=0}^{N+1}$ by $f_0\equiv 0$ and $df_n=\varepsilon_n\sqrt{X_n}1_{\{\tau\geq n\}}$, $n=1,2,\ldots,N+1$. We easily compute that for each $n=1,2,\ldots,N$ we have $\mathbb{E}(|df_n|^2|\mathcal{F}_{n-1})=\delta$ on $\{\tau\geq n\}$. To gain some intuition about f, let us look at the behavior of $(f^2,s^2(f))$. This pair starts from $(0,0)$ and moves vertically for a number of steps. Then at time τ it leaves the y-axis and jumps to the point

$$\left(\frac{2\tau\delta}{p-2}+\frac{2\lambda}{(1-\lambda)(p-2)},\tau\delta\right)$$

if $\tau\leq N$, or to $(2(p(1-\lambda))^{-1},1)$ if $\tau=N+1$, where it stays forever. Derive that

$$\begin{aligned}\mathbb{P}(s(f)\geq 1)&=\mathbb{P}(\tau=N+1)=\mathbb{P}(X_n=0\quad\text{for } n=1,2,\ldots,N)\\&=\prod_{n=1}^N\left(1-\frac{(p-2)(1-\lambda)\delta}{2(n(1-\lambda)\delta+\lambda)}\right)\\&=\exp\left(-\sum_{n=1}^N\frac{(p-2)(1-\lambda)\delta}{2(n(1-\lambda)\delta+\lambda)}\right)\cdot\kappa_1\\&=\exp\left(-\frac{p-2}{2}\int_\lambda^{N(1-\lambda)\delta+\lambda}x^{-1}\mathrm{d}x\right)\cdot\kappa_2\\&=\exp\left(-\frac{p-2}{2}\int_\lambda^{1-2/p}x^{-1}\mathrm{d}x\right)\kappa_2=\left(\frac{\lambda p}{p-2}\right)^{p/2-1}\cdot\kappa_2,\end{aligned}$$

where $\kappa_i=\kappa_i(\delta,\lambda)$ are error terms converging to 1 as $\delta\to 0$, $i=1,2$. Arguing similarly and using elementary bound $1-x\leq e^{-x}$, we obtain the estimate

$$\mathbb{P}\left(|f_{N+1}|^2=\frac{2n\delta}{p-2}+\frac{2\lambda}{(1-\lambda)(p-2)}\right)\leq\frac{(p-2)\delta}{2(n\delta+\lambda/(1-\lambda))}\left[\frac{\lambda+(1-\lambda)\delta}{n(1-\lambda)\delta+\lambda}\right]^{p/2-1}$$

for $n = 1, 2, \ldots, N$ (the event under the probability is just $\{\tau = n\}$), and

$$\mathbb{P}\left(|f_{N+1}|^2 = \frac{2}{p(1-\lambda)}\right) \leq \left[\frac{p(\lambda + (1-\lambda)\delta)}{p-2}\right]^{p/2-1}.$$

Note that this time no error terms κ_i are involved. Consequently,

$$\begin{aligned}||f||_p^p &= ||\, |f_{N+1}|^2\, ||_{p/2}^{p/2}\\ &\leq \sum_{n=1}^N \left(\frac{2n\delta}{p-2} + \frac{2\lambda}{(1-\lambda)(p-2)}\right)^{p/2} \cdot \frac{(p-2)\delta}{2(n\delta + \lambda/(1-\lambda))}\left[\frac{\lambda + (1-\lambda)\delta}{n(1-\lambda)\delta + \lambda}\right]^{p/2-1}\\ &\quad + \left(\frac{2}{p(1-\lambda)}\right)^{p/2} \cdot \left[\frac{p(\lambda + (1-\lambda)\delta)}{p-2}\right]^{p/2-1}.\end{aligned}$$

It is easy to see that all the terms under the sum are equal and hence

$$\begin{aligned}||f||_p^p &\leq N \cdot \frac{(p-2)(1-\lambda)\delta}{2\lambda}\left(\frac{2(\lambda + (1-\lambda)\delta)}{(p-2)(1-\lambda)}\right)^{p/2} + \left[\frac{2\lambda}{(1-\lambda)(p-2)}\right]^{p/2}\frac{p-2}{p\lambda}\\ &= \left(\frac{2\lambda}{(1-\lambda)(p-2)}\right)^{p/2-1}.\end{aligned}$$

Now plug the above expressions for $\mathbb{P}(s(f) \geq 1)$ and $||f||_p^p$ into (2.1) and let $\delta \to 0$: in the limit both sides become equal. Finally, observe that $S_n(f) = |f_n|$ for each n, which implies that the second estimate in (2.1) is also sharp.

3.3. **Sharpness of** (1.7)**.** If $q \leq 2 \leq p$, both sides are equal for the martingale $f_0 = f_1 = f_2 = \ldots \equiv 1$. If $0 < q \leq p < 2$, no finite $L_{p,q}$ suffices: take $f_0 = 0$ and let $f_1 = f_2 = \ldots$ be a mean-zero random variable which belongs to $L^p \setminus L^2$. Then $s(f) = \infty$ and $||f||_p = ||S(f)||_p < \infty$, which enforces $L_{p,q} = \infty$. Finally, for $2 < q \leq p$, take the example from the previous section, with $\lambda = 1 - q/p$. Then

$$\frac{(\mathbb{P}(s(f) \geq 1))^{1/q}}{||f||_p} = \frac{(\mathbb{P}(s(f) \geq 1))^{1/q}}{||S(f)||_p} \geq \frac{\left(\left(\frac{p-q}{p-2}\right)^{p/2-1}\kappa_2\right)^{1/q}}{\left(\left(\frac{2(p-q)}{q(p-2)}\right)^{p/2-1}\right)^{1/p}} = L_{p,q} \cdot \kappa_2^{1/q}$$

and letting $\delta \to 0$ we get the optimality of $L_{p,q}$. The proof is complete.

4. Further remarks and applications

4.1. **On the control of** $||f||_p$ **over** $\mathbb{P}(s(f) \geq 1)$**.** Our starting point is the following problem for the distribution function of $s(f)$ (a version of it, concerning martingale transforms, was studied in depth by [3] and [5]. See also [4]). Suppose that $t \in [0,1]$ is a fixed number and consider the class of those $\mathcal{H}$-valued martingales f, which satisfy $\mathbb{P}(s(f) \geq 1) \geq t$. How small can $||f||_p$ be? The precise answer is contained in the following statement.

Theorem 4.1. *If* $\mathbb{P}(s(f) \geq 1) \geq t$, *then*

$$||f||_p \geq \begin{cases} 0 & \text{if } 0 < p < 2,\\ \left(\frac{2t^{2/(p-2)}}{p-(p-2)t^{2/(p-2)}}\right)^{1/2-1/p} & \text{if } p \geq 2.\end{cases} \tag{4.1}$$

The constant on the right is the best possible. The same statement holds true if we replace $||f||_p$ *by* $||S(f)||_p$.

Proof. The estimate (4.1) is obvious for $0 < p < 2$, so it suffices to prove it for $p \geq 2$. By (1.7), we have

$$||f||_p \geq L_{p,q}^{-1}\,\mathbb{P}(s(f) \geq 1)^{1/q} \geq L_{p,q}^{-1} t^{1/q}.$$

The latter expression, considered as a function of q, attains its maximum for $q = p - t^{2/(p-2)}(p-2)$, and the extremal value is precisely the right-hand side of (4.1). To see that the bound is optimal, it suffices to take a closer look at the examples presented in Section 3 above. We leave the details to the reader. □

There is a dual result, which can be proved by analogous reasoning, applied to (1.6).

Theorem 4.2. *Suppose that f is an $\mathcal{H}$-valued martingale such that $\mathbb{P}(|f_n| \geq 1) \geq t$ for some n. Then*

$$||s(f)||_p \geq \begin{cases} \left(\frac{t}{\Gamma(p/2+1)}\right)^{1/p} & \text{if } 0 < p \leq 2, \\ 0 & \text{if } p > 2. \end{cases}$$

The constant on the right is the best possible. The same statement holds true if we replace the assumption by $\mathbb{P}(S(f) \geq 1) \geq t$.

4.2. **Sharp bounds for nonnegative random variables and their predictable projections.** There is an interesting connection between the square function inequalities studied above and the estimates for sums of nonnegative random variables. To be more precise, suppose that $(e_n)_{n\geq 0}$ is an adapted sequence of nonnegative and integrable random variables and let $(\mathbb{E}(e_n|\mathcal{F}_{n-1}))_{n\geq 0}$ be the sequence of corresponding predictable projections (we set $\mathcal{F}_{-1} = \mathcal{F}_0$). The problem of comparing various norms of these two sequences, with the emphasis on the sizes of the constants involved, has gained a considerable interest in the literature and has been applied in several areas of mathematics, including probability theory, harmonic analysis and mathematical finance. See e.g. [1], [6], [7], [8], [14], [15], [17], [20] and [21].

Let us relate this subject with the inequalities of the previous sections. We say that a sequence $(e_n)_{n\geq 0}$ is simple, if it is finite and each term takes only a finite number of values. In the statement below, the filtration is to vary as well as the probability space.

Theorem 4.3. *Fix $V : [0,\infty) \times [0,\infty) \to \mathbb{R}$. The following statements are equivalent.*

(i) For any adapted simple sequence $(e_n)_{n\geq 0}$ of nonnegative random variables we have

$$\mathbb{E}V\left(\sum_{n=0}^{\infty} e_n, \sum_{n=0}^{\infty} \mathbb{E}(e_n|\mathcal{F}_{n-1})\right) \leq 0. \tag{4.2}$$

(ii) For any adapted, simple, real-valued martingale f we have

$$\mathbb{E}V(S^2(f), s^2(f)) \leq 0. \tag{4.3}$$

Proof. To get (i)⇒(ii), apply the substitution $e_n = |df_n|^2$, $n = 0, 1, 2, \ldots$. To obtain (ii)⇒(i), use the martingale with the difference sequence $(\varepsilon_n\sqrt{e_n})_{n\geq 0}$, where ε_0, ε_1, ε_2, … are independent, adapted Rademacher variables, independent also from $(e_n)_{n\geq 0}$ (it may be necessary to enlarge the probability space and the filtration to get such a sequence). □

Now, observe that the estimate $\mathbb{P}(S(f) \geq 1) \leq (\Gamma(p/2+1))^{-1}||s(f)||_p^p$ (valid for $p \leq 2$) and the second inequality in (2.1) (valid for $p \geq 2$) are of the form (4.3). Therefore, using the above theorem and repeating the reasoning from Sections 2 and 3, we obtain the following statement.

Theorem 4.4. *Let $(e_n)_{n\geq 0}$ be a sequence of adapted, nonnegative and integrable random variables. Then for any $0 < q \leq p < \infty$ we have the sharp inequalities*

$$\left|\left|\sum_{n=0}^{\infty} e_n\right|\right|_{q,\infty} \leq K_{2p,2q}^2 \left|\left|\sum_{n=0}^{\infty} \mathbb{E}(e_n|\mathcal{F}_{n-1})\right|\right|_p ,$$

$$\left|\left|\sum_{n=0}^{\infty} \mathbb{E}(e_n|\mathcal{F}_{n-1})\right|\right|_{q,\infty} \leq L_{2p,2q}^2 \left|\left|\sum_{n=0}^{\infty} e_n\right|\right|_p .$$

It is also easy to obtain the corresponding versions of Theorems 4.1 and 4.2. We omit the straightforward details.

4.3. **A bound related to Khintchine's inequality.** Khintchine's inequality (see [13]) plays a fundamental role in the both commutative and non-commutative probability theory, geometry of Banach spaces, harmonic analysis and many other areas of mathematics. The problem of determining the optimal (or almost optimal) constants in this classical estimate and its various versions has interested many mathematicians (see e.g. [9], [10], [18] and [19]). The estimates of the previous sections can be used to obtain the following result in this direction. Suppose that $(a_n)_{n\geq 0}$ is a predictable sequence of random variables, taking values in a Hilbert space $\mathcal{H}$, and let ε_0, ε_1, ε_2, ... be an adapted sequence of independent Rademacher variables, independent also from $(a_n)_{n\geq 0}$. Then the process $f = (\sum_{k=0}^n a_k\varepsilon_k)_{n\geq 0}$ is a martingale and its conditional square function equals $s(f) = \left(\sum_{k=0}^{\infty} |a_k|^2\right)^{1/2}$. Applying inequalities (1.6) and (1.7), we get the following.

Theorem 4.5. *Let $(a_n)_{n\geq 0}$ and $(\varepsilon_n)_{n\geq 0}$ be as above. Then for $0 < q \leq p < \infty$ we have*

$$\left|\left|\sum_{k=0}^{\infty} a_k\varepsilon_k\right|\right|_{q,\infty} \leq K_{p,q} \left|\left|\left(\sum_{k=0}^{\infty} |a_k|^2\right)^{1/2}\right|\right|_p ,$$

$$\left|\left|\left(\sum_{k=0}^{\infty} |a_k|^2\right)^{1/2}\right|\right|_{q,\infty} \leq L_{p,q} \left|\left|\sum_{k=0}^{\infty} a_k\varepsilon_k\right|\right|_p .$$

Similarly, one can state the appropriate versions of Theorems 4.1 and 4.2. We omit the formulation, leaving it to the interested reader.

Acknowledgment

The author would like to thank the Referee for the careful reading of the first version of the manuscript and helpful suggestions. The research was partially supported by Polish Ministry of Science and Higher Education (MNiSW) Grant N N201 397437.

References

[1] Bourgain, J., 1983. Embedding L^1 into L^1/H^1. Trans. Amer. Math. Soc., 689–702.

[2] Burkholder, D. L., 1973. Distribution function inequalities for martingales. Ann. Probab. 1, 19–42.

[3] Burkholder, D. L., 1984. Boundary value problems and sharp inequalities for martingale transforms. Ann. Probab. 12, 647–702.

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

[5] Choi, K. P., 1988. Some sharp inequalities for martingale transforms, Trans. Amer. Math. Soc. 307, 279–300.

[6] Delbaen, F. and Schachermayer, W., 1994. A general version of the fundamental theorem of asset pricing. Math. Ann. 300, 463–520.

[7] Delbaen, F. and Schachermayer, W., 1995. An inequality for the predictable projection of an adapted process. Séminaire de Probabilités XXIX, Lecture Notes in Math. 1613, 17–24.

[8] Garsia, A.M., 1973. Martingale inequalities: Seminar notes on recent progress. Mathematics Lecture Notes Series, W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam.

[9] Haagerup, U., 1982. The best constants in the Khintchine inequality. Studia Math. 70, 231-283.

[10] Haagerup, U. and Musat, M., 2007. On the best constants in noncommutative Khintchine-type inequalities. J. Funct. Anal. 250, 588-624.

[11] Hitczenko, P., 1990. Best constants in martingale version of Rosenthal's inequality. Ann. Probab. 18, 1656–1668.

[12] Junge, M. and Xu, Q., 2003. Noncommutative Burkholder/Rosenthal inequalities. Ann. Probab. 31, 948–995.

[13] Khintchine, A., 1923. Über dyadische Brüche. Math Z., 18, 109–116.

[14] Lépingle, D., 1978. Une inégalité de martingales. Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg), pp. 134–137, Lecture Notes in Math., 649, Springer, Berlin.

[15] Osękowski, A., 2009. Sharp norm comparison of the maxima of a sequence and its predictable projection, Statist. Probab. Lett. 79, 1784–1788.

[16] Osękowski, A., 2010. Sharp inequalities for sums of nonnegative random variables and for a martingale conditional square function. ALEA Lat. Am. J. Probab. Math. Stat. 7, 243–256.

[17] Stein, E.M., 1970. Topics in harmonic analysis. Ann. of Math. Studies 63, Princeton Univ. Press.

[18] Szarek, S. J., 1976. On the best constants in the Khinchin inequality. Studia Math. 58, 197-208.

[19] Tomaszewski, B., 1987. A simple and elementary proof of the Khintchine inequality with the best constant. Bull. Sci. Math. 111, 103-109.

[20] Yor, M., 1978. Inégalités entre processus minces el applications, CRAS Paris 286, Serie A, 799–801.

[21] Wang, G. 1991. Sharp inequalities for the conditional square function of a martingale, Ann. Probab. 19, 1679–1688.

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

E-mail address: ados@mimuw.edu.pl