A NOTE ON BURKHOLDER-ROSENTHAL INEQUALITY

ADAM OSEKOWSKI

ABSTRACT. Let $d f$ be a Hilbert-space-valued martingale difference sequence. The paper is devoted to a new, elementary proof of the estimate

$$\left\| \sum_{k=0}^{\infty} d f_k \right\|_p \leq C_p \left\{ \left( \sum_{k=0}^{\infty} \mathbb{E}(|d f_k|^2 |\mathcal{F}_{k-1}) \right)^{1/2} + \left( \sum_{k=0}^{\infty} |d f_k|^p \right)^{1/p} \right\},$$

with $C_p = O(p/\ln p)$ as $p \to \infty$.

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, filtered by $(\mathcal{F}_n)_{n \geq 0}$, a nondecreasing family of sub-$\sigma$-algebras of $\mathcal{F}$. Assume that $f$ is an adapted martingale, taking values in a certain separable Hilbert space $H$ with norm $|\cdot|$ and scalar product $\langle \cdot, \cdot \rangle$. Then $d f = (d f_n)_{n \geq 0}$, the difference sequence of $f$, is given by $d f_0 = f_0$ and $d f_n = f_n - f_{n-1}$, $n \geq 1$. We define the conditional square function of $f$ by

$$s(f) = \left[ \sum_{k=0}^{\infty} \mathbb{E}(|d f_k|^2 |\mathcal{F}_{k-1}) \right]^{1/2},$$

(here and below, $\mathcal{F}_{-1} = \mathcal{F}_0$) and use the notation

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

for the truncated conditional square function of $f$.

The purpose of this note is to investigate Burkholder-Rosenthal inequality

$$(1.1) \quad \|f\|_p \leq c_p \left( \|s(f)\|_p + \left\| \sum_{k=0}^{\infty} |d f_k|^p \right\|^{1/p}_p \right)$$

where $p \geq 2$ and $c_p$ is a constant depending only on $p$. The special case in which the martingale $f$ is a sum of independent mean-zero random variables forms an important extension of Khinchine inequality and was studied by Rosenthal in the 60’s. The proof from [11] gives the constant $c_p$ which grows exponentially in $p$ as $p \to \infty$. Johnson, Schechtman and Zinn [4] refined the reasoning and showed that the optimal order of $c_p$ as $p \to \infty$ (still in the independent case) is $p/\ln p$. Applying difficult isoperimetric techniques, Talagrand [12] extended this statement to the case of independent Banach-space-valued random variables. Using hypercontractivity
methods, Kwapień and Szulga [7] gave a completely elementary proof of Talagrand’s result.

The inequality (1.1) for general real martingales (and some \( c_p \)) was established by Burkholder in [1]. The validity of this estimate with \( c_p = O(p/\ln p) \) was proved by Hitczenko [5] (see also [6]). This result was further generalized to vector-valued setting by Pinelis [10]. Consult also Nagaev [8] for a yet another approach.

The purpose of this paper is to present a new and elementary proof of (1.1) with \( c_p = O(p/\ln p) \). Precisely, we will establish the following statement.

**Theorem 1.1.** If \( f \) is a Hilbert-space-valued martingale, then for \( p \geq 4 \) we have

\[
\|f\|_p \leq C_p \left( \|s(f)\|_p^p + \left( \sum_{k=0}^{\infty} |d_k|^p \right)^{1/p} \right)^{1/p},
\]

where

\[
C_p = 2\sqrt{2} \left( \frac{p}{4} + 1 \right)^{1/p} \left( 1 + \frac{p}{\ln(p/2)} \right).
\]

In fact, using Davis’ decomposition, we will be able to prove a slightly stronger estimate: see (2.10) and Remark 2.5 below.

A few words about the proof are in order. Hitczenko [5], [6] and Pinelis [10] apply the extrapolation method (good \( \lambda \)-inequality) of Burkholder and Gundy, combined with appropriate version of Prokhorov “arcsinh” estimate for martingales. Nagaev [8] first establishes a certain exponential bound for the tail of \( f \) and deduces Burkholder-Rosenthal estimate using a standard integration argument. Our approach is entirely different and exploits the properties of a certain special function; this type of proof can be regarded as an application of Burkholder’s method (see [2] and [9] for more on the subject).

**2. Proof of Theorem 1.1**

The starting point is the following technical estimate proved by Kwapień and Szulga [7].

**Lemma 2.1.** Let \( p \geq 4 \) and put

\[
\eta = \eta(p) := \frac{\ln(p/2)/p}{1 + \ln(p/2)/p}.
\]

Then for any \( t \geq 0 \) we have

\[
(1 + t\eta)^p - pt\eta \leq 1 + \left( \frac{p}{2} - 1 \right) t^2 + t^p.
\]

We shall require the following vector-valued version of this bound. From now on, we assume that \( p \geq 4 \) and that \( \sigma = \sigma(p) = \eta(p)/\sqrt{2} \).

**Lemma 2.2.** For any \( y, d \in \mathcal{H} \) we have

\[
|y + \sqrt{2}\sigma d|^p - p|y|^{p-2} \langle y, \sqrt{2}\sigma d \rangle \leq |y|^p + \frac{p}{2}|y|^{p-2}|d|^2 + |d|^p.
\]

**Proof.** The left-hand side can be rewritten in the form \( F(y, \sqrt{2}\sigma d) \), where

\[
F(s) = ||y|^2 + 2\sigma^2|d|^2 + 2s|p/2 - p|y|^{p-2}s, \quad s \in \mathbb{R}.
\]

Now keep \( |y| \) and \( |d| \) fixed; since the function \( F \) is convex, it suffices to prove the estimate for \( \langle y, \sqrt{2}\sigma d \rangle = \pm \sqrt{2}\sigma |y||d| \), i.e. in the case when \( y \) and \( d \) are linearly
dependent. If $\langle y, \sqrt{2}\sigma d \rangle = \sqrt{2}\sigma |y||d|$, then (2.2) follows directly from (2.1); on the other hand, if $\langle y, \sqrt{2}\sigma d \rangle = -\sqrt{2}\sigma |y||d|$, we have

$$|y + \sqrt{2}\sigma d|^p - p|y|^{p-2}\langle y, \sqrt{2}\sigma d \rangle = ||y| - \sqrt{2}\sigma |d||^p + p\sqrt{2}\sigma |y|^{p-1}|d|$$

$$\leq ||y| + \sqrt{2}\sigma |d||^p - p\sqrt{2}\sigma |y|^{p-1}|d|,$$

so the claim again follows from (2.1).

The key ingredient of the proof is the special function $U : [0, \infty) \times \mathcal{H} \times [0, \infty) \to \mathbb{R}$, given by

$$U(x, y, z) = \begin{cases} (|y|^2 - x^2)^{p/2} - cx^p - z & \text{if } |y| \geq \sqrt{2}x, \\ |y|^p - (2p/2 - 1 + c)x^p - z & \text{if } |y| < \sqrt{2}x, \end{cases}$$

where

$$c = p 2^{p/2-2} + 1.$$

Let us list some properties of this function.

**Lemma 2.3.** (i) For any $(x, y, z) \in [0, \infty) \times \mathcal{H} \times [0, \infty)$ we have

$$U(x, y, z) = \min \left\{ ||y|^2 - x^2||^{p/2} - cx^p - z, |y|^p - (2p/2 - 1 + c)x^p - z \right\}.$$ (2.3)

(ii) For any $x \geq 0$ and $y \in \mathcal{H}$ we have

$$U(x, y, 0) = \begin{cases} (|y|^2 - x^2)^{p/2} - cx^p - z & \text{if } |y| \geq \sqrt{2}x, \\ |y|^p - (2p/2 - 1 + c)x^p - z & \text{if } |y| < \sqrt{2}x, \end{cases}.$$ (2.4)

(iii) For all $(x, y, z) \in [0, \infty) \times \mathcal{H} \times [0, \infty)$ we have

$$U(x, y, z) \geq 2^{-p/2} \left[ |y|^p - \sigma p C_p(x^p + z) \right].$$ (2.5)

**Proof.** (i) For fixed $x$, $z \geq 0$, the function

$$F(s) = s^p - (2p/2 - 1 + c)x^p - z - (|s|^2 - x^2)^{p/2} - cx^p - z,$$

vanishes at $s = \sqrt{2}x$ and is strictly increasing:

$$F'(s) = p s^{p-2} - |s|^2 - x^2 (p-2)/2 \sgn (|s|^2 - x^2).$$

This yields (2.3).

(ii) This is obvious, since $\sigma \leq 1$.

(iii) Using the definitions of $C_p$ and $\sigma$, we see that we must prove the bound

$$U(x, y, z) \geq 2^{-p/2} \left[ |y|^p - \left( \frac{p}{4} + 1 \right) 2^p (x^p + z) \right].$$

Now, for $|y| < \sqrt{2}x$, we have

$$U(x, y, z) = |y|^p - \left( \frac{p}{4} + 1 \right) 2^{p/2} x^p - z \geq 2^{-p/2} \left[ |y|^p - \left( \frac{p}{4} + 1 \right) 2^p (x^p + z) \right].$$

On the other hand, if $|y| \geq \sqrt{2}x$, then $|y|^2 - x^2 \geq |y|^2/2$ and hence

$$U(x, y, z) \geq 2^{-p/2} \left[ |y|^p - 2p/2 cx^p - 2p^2/2 \right],$$

so the majorization is clear.

We turn to the key property of the function $U$.

**Lemma 2.4.** For any $x, z \geq 0$, $y \in \mathcal{H}$ and any $\mathcal{H}$-valued, mean-zero random variable $d$ with $||d||_p < \infty$ we have

$$\mathbb{E} U \left( \sqrt{x^2 + \mathbb{E} |d|^2}, y + \sigma d, z + |d|^p \right) \leq U(x, y, z).$$ (2.6)
Proof. We consider three cases separately.

1° The case \(|y|^2 \leq 2x^2\). By (2.3), we have

\[
\mathbb{E} \left( \sqrt{x^2 + \mathbb{E}|d|^2}, y + \sigma d, z + |d|^p \right)
\leq \mathbb{E}|y + \sigma d|^p - (2^{p/2} - 1 + c)(x^2 + \mathbb{E}|d|^2)^{p/2} - z - \mathbb{E}|d|^p
\]

= \mathbb{E}\{ |y + \sigma d|^p - p|y|^{p-2}(y, \sigma d) - |d|^p \} - (2^{p/2} - 1 + c)(x^2 + \mathbb{E}|d|^2)^{p/2} - z.

By (2.2), the expression in the parentheses does not exceed \(|y|^p + p|y|^{p-2}|d|^2/2\); furthermore, we have

\[
(2^{p/2} - 1 + c)(x^2 + \mathbb{E}|d|^2)^{p/2} \geq (2^{p/2} - 1 + c) \left( x^p + \frac{p}{2} x^{p-2} \mathbb{E}|d|^2 \right)
\]

\[
\geq (2^{p/2} - 1 + c)x^p + \frac{p}{2} x^{p-2} \mathbb{E}|d|^2
\]

\[
\geq (2^{p/2} - 1 + c)x^p + \frac{p}{2} |y|^{p-2} \mathbb{E}|d|^2.
\]

Combining these estimates, we obtain

\[
\mathbb{E} \left( \sqrt{x^2 + \mathbb{E}|d|^2}, y + \sigma d, z + |d|^p \right) \leq |y|^p - (2^{p/2} - 1 + c)x^p - z,
\]

which is precisely the desired bound.

2° The case \(2x^2 < |y|^2 \leq 2(x^2 + \mathbb{E}|d|^2)\). We start as previously: by (2.3) and then (2.2),

\[
\mathbb{E} \left( \sqrt{x^2 + \mathbb{E}|d|^2}, y + \sigma d, z + |d|^p \right)
\leq |y|^p + \frac{p}{2} |y|^{p-2} \mathbb{E}|d|^2 - (2^{p/2} - 1 + c)(x^2 + \mathbb{E}|d|^2)^{p/2} - z.
\]

The latter expression decreases as \(\mathbb{E}|d|^2\) increases; indeed, the function

\[
F(s) = |y|^p + \frac{p}{2} |y|^{p-2}s - (2^{p/2} - 1 + c)(x^2 + s)^{p/2} - z, \quad s \geq \frac{|y|^2}{2} - x^2,
\]

satisfies

\[
F'(s) \leq \frac{p}{2} \left[ |y|^{p-2} - 2^{p/2-1}(x^2 + s)^{p/2-1} \right] \leq 0.
\]

In consequence, we have

\[
\mathbb{E} \left( \sqrt{x^2 + \mathbb{E}|d|^2}, y + \sigma d, z + |d|^p \right)
\leq F \left( \frac{|y|^2}{2} - x^2 \right)
\leq \frac{p}{2} |y|^{p-2} \left( \frac{|y|^2}{2} - x^2 \right) - (c - 1) \left( \frac{y^2}{2} \right)^{p/2} - z
\leq \frac{p}{2} |y|^{p-2}x^2 - z
\leq \left( \frac{|y|^2}{2} \right)^{p/2} \left( x^2 - \left( \frac{p}{2} + 2^{1-p/2} \right) |y|^{p-2}x^2 \right.
\leq \left( \frac{|y|^2}{2} \right)^{p/2} - c \left( \frac{|y|^2}{2} \right)^{p/2-1} x^2 - z
\leq (|y|^2 - x^2)^{p/2} - cx^p - z;
\]
and we are done.

3' The case $|y|^2 > 2(x^2 + E|d|^2)$. Here the reasoning is a bit more complicated. First we show the pointwise estimate

\begin{equation}
|y + \sigma d|^2 - x^2 - E|d|^2|^{p/2} - p|y|^2 - x^2 - E|d|^2|^{p/2-1}(y, \sigma d)
\leq \left( |y|^2 - x^2 - E|d|^2 \right)^{1/2} + \sqrt{2}\sigma|d|
\right)^p - p|y|^2 - x^2 - E|d|^2|^{(p-1)/2}\sqrt{2}\sigma|d|.
\end{equation}

In fact, we will establish a slightly stronger inequality:

\[
|y + \sigma d|^2 - x^2 - E|d|^2|^{p/2} - p|y|^2 - x^2 - E|d|^2|^{p/2-1}(y, \sigma d)
\leq \left( |y|^2 - x^2 - E|d|^2 + 2\sqrt{2}|y|^2 - x^2 - E|d|^2|^{1/2}\sigma|d| \right)^{p/2}
- p|y|^2 - x^2 - E|d|^2|^{(p-1)/2}\sqrt{2}\sigma|d|.
\]

To do this, divide throughout by $||y|^2 - x^2 - E|d|^2|^{p/2}$ and substitute

\[
A^2 = \frac{|y|^2 - x^2 - E|d|^2 + \sigma^2|d|^2}{|y|^2 - x^2 - E|d|^2}, \quad Y = \frac{y}{||y|^2 - x^2 - E|d|^2|^{1/2}}
\]

and

\[
D = \frac{d}{||y|^2 - x^2 - E|d|^2|^{1/2}}.
\]

The estimate becomes

\begin{equation}
|A^2 + 2\langle Y, \sigma D \rangle|^{p/2} - p(Y, \sigma D) \leq |A^2 + 2\sqrt{2}\sigma|D||^p - p\sqrt{2}\sigma|D|.
\end{equation}

However, the reasoning presented in the proof of (2.2) gives

\[
|A^2 + 2\langle Y, \sigma D \rangle|^{p/2} - p(Y, \sigma D) \leq (A^2 + 2\sigma|Y||D|)^{p/2} - p\sigma|Y||D|.
\]

It suffices to use the bounds $|Y| \leq \sqrt{2}$ and $A^2 \geq 1$ to obtain (2.8), because the function $s \mapsto (A^2 + 2s)^{p/2} - ps$ is increasing on $[0, \infty)$. Thus (2.7) follows. We turn to (2.6): applying (2.3), we get

\[
EU \left( \sqrt{x^2 + E|d|^2}, y + \sigma d, z + |d|^p \right)
\leq E \left[ |y + \sigma d|^2 - x^2 - E|d|^2|^{p/2} - c(x^2 + E|d|^2)^{p/2} - z - E|d|^p \right]
\leq E \left[ \left( |y + \sigma d|^2 - x^2 - E|d|^2 \right)^{1/2} + \sqrt{2}\sigma|d| \right)^p
\leq \left( |y|^2 - x^2 - E|d|^2 \right)^{1/2} + \sqrt{2}\sigma|d|
\right)^p - p|y|^2 - x^2 - E|d|^2|^{(p-1)/2}\sqrt{2}\sigma|d|\right)^p
- p|y|^2 - x^2 - E|d|^2\right)^{p/2-1}(y, \sigma d)
\leq \left( |y|^2 - x^2 - E|d|^2 + 2\sqrt{2}|y|^2 - x^2 - E|d|^2\right)^{p/2}
- p|y|^2 - x^2 - E|d|^2\right)^{(p-1)/2}\sqrt{2}\sigma|d|.
\]

Now we apply (2.2) (in the real case) to obtain

\[
EU \left( \sqrt{x^2 + E|d|^2}, y + \sigma d, z + |d|^p \right)
\leq \left( |y|^2 - x^2 - E|d|^2 \right)^{p/2} + \frac{p}{2} |y|^2 - x^2 - E|d|^2\right)^{p/2-1}\sigma|d|^2 + c|x|^p - z
\leq |y|^2 - x^2|^{p/2} - cx^p - z = U(x, y, z).
\]

This completes the proof. \qed
Proof of (1.2). It suffices to prove that for any nonnegative integer $n$,

$$
E|f_n|^p \leq C_p E \left( s_n^p(f) + \sum_{k=0}^n |df_k|^p \right).
$$

Of course, we may assume that $df_0, df_1, \ldots, df_n$ (and hence also $f_n$) belong to $L^p$, since otherwise there is nothing to prove. The key observation is that the process

$$
\left( U\left( s_n(f), \sigma f_n, \sum_{k=0}^n |df_k|^p \right) \right)_{n \geq 0}
$$

is a supermartingale with respect to $(\mathcal{F}_n)_{n \geq 0}$. Indeed, the integrability follows from the above assumption on $df$; furthermore, for any $n \geq 0$ we have

$$
E \left[ U \left( s_{n+1}(f), \sigma f_{n+1}, \sum_{k=0}^{n+1} |df_k|^p \right) \bigg| \mathcal{F}_n \right] = E \left[ U \left( \sqrt{s_n^2(f) + E(|df_{n+1}|^2|\mathcal{F}_n)|}, \sigma f_n + \sigma df_{n+1}, \sum_{k=0}^n |df_k|^p + |df_{n+1}|^p \right) \bigg| \mathcal{F}_n \right],
$$

which does not exceed $U(s_n(f), \sigma f_n, \sum_{k=0}^n |df_k|^p)$, by (2.6) applied conditionally with respect to $\mathcal{F}_{n-1}$. Next, we have $U(s_0(f), \sigma f_0, |df_0|^p) \leq 0$, in view of (2.4). Combining these two facts with (2.5) yields the claim:

$$
E \left[ |f_n|^p - C_p \left( s_n^p(f) + \sum_{k=0}^n |df_k|^p \right) \right] \leq \frac{2p/2}{\sigma_p} E \left( s_n(f), \sigma f_n, \sum_{k=0}^n |df_k|^p \right) \leq 0. \quad \Box
$$

Remark 2.5. Using Davis’ decomposition (see e.g. Davis [3] or Burkholder [1]), one can deduce a slightly stronger form of (1.2). Namely, for all $f$ as in the statement of Theorem 1.1 and $p \geq 4$ we have

$$
\|f\|_p \leq 2C_p \left( \|s(f)\|_p^p + \|df^*\|_p^p \right)^{1/p},
$$

where $df^* = \sup_{n \geq 0} |df_n|$. Indeed, fix a martingale $f$ and consider the random variables $d_{n-1}' = df_n 1_{|df_n| \leq 2df^*_{n-1}}$, $d_n'' = df_n 1_{|df_n| > 2df^*_{n-1}}$. Here, as usual, $df^*_n \equiv 0$ and $df^*_n = \max_{0 \leq k \leq n} |df_k|$. Note that on the set $\{ |df_n| \geq 2df^*_{n-1} \}$ we have

$$
(2p - 1)|d_{n-1}'|^p + (2df^*_{n-1})^p \leq (2|df_n|)^p \leq (2df^*_n)^p,
$$

which implies

$$
\sum_{k=0}^n |d_k'|^p \leq \frac{2p}{2p - 1} (df^*_n)^p, \quad n = 0, 1, 2, \ldots.
$$

Next, observe that for any $n$,

$$
E \left[ \sum_{k=0}^n |d_k|^p \right] \leq \sum_{k=0}^n E(2|df^*_k|^{2} (2df^*_k)^{p-2})
$$

$$
= \sum_{k=0}^n E(|df_k|^2|\mathcal{F}_{k-1})(2df^*_k)^{p-2}
$$

$$
\leq E s_n^p(f)(2df^*_n)^{p-2}
$$

$$
\leq \frac{2}{p} \|s_n(f)\|_p^p + \frac{p - 2}{p} \|2df^*_n\|_p^p.
$$
where in the last line we have exploited Young’s inequality. Combining the above estimates for the sums of \( d'_n \) and \( d''_n \) we get

\[
\mathbb{E} \sum_{k=0}^{n} |df_k|^p \leq \frac{2}{p} \left( \frac{1}{2p - 1} + \frac{p - 2}{p} \right) ||s_n(f)||_p^p + 2^p ||df^*_n||_p^p
\]

Plugging this into (2.9) and using the fact that \( n \) is an arbitrary nonnegative integer, we obtain (2.10).

References


Department of Mathematics, Informatics and Mechanics, University of Warsaw, Banacha 2, 02-097 Warsaw, Poland
E-mail address: ados@mimuw.edu.pl