# SHARP INEQUALITIES FOR SUMS OF NONNEGATIVE RANDOM VARIABLES AND FOR A MARTINGALE CONDITIONAL SQUARE FUNCTION

ADAM OSĘKOWSKI

Abstract. In the paper we prove weak-type and Φ-inequalities for the conditional square function of a martingale. A related estimates for the sums of nonnegative random variables and sums of their predictable projections are established.

## 1. Introduction

Let $(\Omega,\mathcal{F},\mathbb{P})=([0,1],\mathcal{B}([0,1]),|\cdot|)$ be a probability space and $(\mathcal{F}_n)$ be a filtration, that is, a nondecreasing sequence of sub-$\sigma$-fields of $\mathcal{F}$. Throughout the paper, $f=(f_n)$ will be a martingale adapted to $(\mathcal{F}_n)$ and taking values in a certain separable Banach space $(B,||\cdot||)$. Let $df=(df_n)$ be the difference sequence of $f$, defined by $f_n=\sum_{k=0}^n df_k$, $n=0,\,1,\,\ldots$. Then $S(f)$, the *square function* of $f$, and $s(f)$, the *conditional square function* of $f$, are given by

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

where $\mathcal{F}_{-1}=\mathcal{F}_0$. 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, we will write

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

and

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

where $0<p\leq\infty$.

The purpose of this paper is to provide some sharp estimates involving the sizes of a martingale, its square and conditionally square function. Let us start with

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

*Key words and phrases.* Martingale, conditional square function, smooth Banach space, convex Banach space, weak type inequality, Φ-inequality.

Partially supported by Foundation for Polish Science.

related results from the literature. The sharp inequalities

$$\tag{1.1} \begin{aligned} ||s(f)||_p &\leq \sqrt{\frac{p}{2}}||f||_p, \\ ||s(f)||_p &\leq \sqrt{\frac{p}{2}}||S(f)||_p, \end{aligned}$$

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

$$\tag{1.2} \begin{aligned} ||f||_p &\leq \sqrt{\frac{2}{p}}||s(f)||_p, \\ ||S(f)||_p &\leq \sqrt{\frac{2}{p}}||s(f)||_p, \end{aligned}$$

for $0 < p \leq 2$, were established by Wang [5] in the case when $B$ is a Hilbert space. Furthermore, the inequalities fail to hold for the remaining values of $p$; for example, the inequality $||s(f)||_p \leq C_p||f||_p$, with $0 < p < 2$, is not valid in general with any finite $C_p$, even for real-valued martingales. However, the following estimate was established by Junge and Xu [4]: if $1 < p < 2$ and $f$ is real-valued, then, for some absolute $C_p$,

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

where the infimum runs over all possible decompositions of $f$ as a sum $f = g + h$ of two martingales. This estimate can be seen as dual to the Burkholder-Rosenthal inequality (see [1], [3], [4] for details)

$$\tag{1.4} C_p^{-1}||f||_p \leq \max\left\{ ||s(f)||_p, \left|\left| \left( \sum_{k=0}^{\infty} |df_k|^p \right)^{1/p} \right|\right|_p \right\} \leq C_p||f||_p,\ 2 \leq p < \infty.$$

In view of Burkholder-Davis-Gundy inequalities (consult [1]), bounds similar to (1.3) and (1.4) are valid when $||f||_p$ is replaced by $||S(f)||_p$.

In the present paper we will study related inequalities, putting particular emphasis on the size of the constants. First, we determine the optimal constants in the following weak-type estimate. It is convenient to formulate the result in the case $0 < p \leq 2$ and $2 \leq p < \infty$ separately. Recall that a martingale $f$ is *conditionally symmetric* if for any $n \geq 0$, the conditional distributions of $df_n$ and $-df_n$ given $\mathcal{F}_{n-1}$ coincide. In the results below, we also use the notion of convexity and smoothness of a Banach space; for the definition, see Section 2 below.

**Theorem 1.1.** *Suppose $p$ belongs to the interval $(0, 2]$ and $f$ is a $B$-valued martingale.*

*(i) We have*

$$\tag{1.5} ||S(f)||_{p,\infty} \leq (\Gamma(p/2 + 1))^{-1/p}||s(f)||_p.$$

*(ii) If $B$ is a Hilbert space, then*

$$\tag{1.6} ||f||_{p,\infty} \leq (\Gamma(p/2 + 1))^{-1/p}||s(f)||_p.$$

*(iii) If $B$ is $(2, \alpha)$-smooth and $f$ is conditionally symmetric, then*

$$\tag{1.7} ||f||_{p,\infty} \leq (\Gamma(p/2 + 1))^{-1/p} \cdot \alpha||s(f)||_p.$$

*The inequalities* (1.5) *and* (1.6) *are sharp, even if* $B = \mathbb{R}$.

For $2 \le p < \infty$ we have the following.

**Theorem 1.2.** *Suppose $p$ belongs to the interval $[2,\infty)$ and $f$ is a $B$-valued martingale.*

*(i) We have*

$$||s(f)||_{p,\infty} \le \left(\frac{p}{2}\right)^{1/2-1/p} ||S(f)||_p. \tag{1.8}$$

*(ii) If $B$ is a Hilbert space, then*

$$||s(f)||_{p,\infty} \le \left(\frac{p}{2}\right)^{1/2-1/p} ||f||_p. \tag{1.9}$$

*(iii) If $B$ is $(2,\alpha)$-convex and $f$ is conditionally symmetric, then*

$$||s(f)||_{p,\infty} \le \left(\frac{p}{2}\right)^{1/2-1/p} \cdot \frac{1}{\alpha}||f||_p. \tag{1.10}$$

*The inequalities* (1.8) *and* (1.9) *are sharp, even if* $B = \mathbb{R}$.

The following sharp estimate for the tail of $s(f)$ can be regarded as a version of Theorem 1.2 in the case $p = \infty$.

**Theorem 1.3.** *Let $f$ be a $B$-valued martingale.*

*(i) If $||S(f)||_\infty \le 1$, then for any $\lambda > 0$ we have*

$$\mathbb{P}(s(f) \ge \lambda) \le \min(e^{1-\lambda^2}, 1). \tag{1.11}$$

*(ii) If $B$ is a Hilbert space and $||f||_\infty \le 1$, then for any $\lambda > 0$ we have*

$$\mathbb{P}(s(f) \ge \lambda) \le \min(e^{1-\lambda^2}, 1). \tag{1.12}$$

*(iii) If $B$ is a $(2,\alpha)$-smooth and $||f||_\infty \le 1$, then for any $\lambda > 0$ we have*

$$\mathbb{P}(s(f) \ge \lambda) \le \min(e^{1-\lambda^2/\alpha^2}, 1). \tag{1.13}$$

*The inequalities* (1.11) *and* (1.12) *are sharp, even if* $B = \mathbb{R}$.

The theorem above can be used to provide the proof of the following well known exponential inequality (see e.g. Garsia [2]). If $f$ is a Hilbert-space-valued martingale bounded by 1, and we integrate (1.12), we get, for any $\beta < 1$,

$$\mathbb{E}\exp\left(\beta s^2(f)\right) \le \frac{(2-\beta)e^{1-\beta}}{1-\beta}.$$

However, the bound on the right is not optimal. We determine it in the next result.

**Theorem 1.4.** *Suppose $\Phi : [0,\infty) \to \mathbb{R}$ is an increasing convex function and $f$ is a $B$-valued martingale.*

*(i) If $||S(f)||_\infty \le 1$, then*

$$\mathbb{E}\Phi(s^2(f)) \le \int_0^\infty \Phi(t)e^{-t}dt. \tag{1.14}$$

*(ii) If $B$ is a Hilbert space and $||f||_\infty \le 1$, then*

$$\mathbb{E}\Phi(s^2(f)) \le \int_0^\infty \Phi(t)e^{-t}dt. \tag{1.15}$$

*(iii) If $B$ is $(2,\alpha)$-smooth and $f$ is conditionally symmetric with $||f||_\infty \leq 1$, then*

$$\mathbb{E}\Phi(s^2(f)) \leq \int_0^\infty \Phi(t\alpha)e^{-t}dt. \tag{1.16}$$

*The inequalities* (1.14) *and* (1.15) *are sharp, even if $B=\mathbb{R}$.*

The estimates (1.1) and (1.2) are accompanied by the inequalities for the sums of nonnegative random variables (in fact we have the equivalence, see Theorem 2.1 below): if $(e_n)_{n=0}^\infty$ is an adapted sequence of nonnegative random variables and $(\mathbb{E}(e_n|\mathcal{F}_{n-1}))_{n=0}^\infty$ stands for its predictable projection, then

$$||\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_{k-1})||_p \leq p||\sum_{k=0}^\infty e_k||_p, \qquad \text{for } 1\leq p<\infty,$$

and

$$p||\sum_{k=0}^\infty e_k||_p \leq ||\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_{k-1})||_p, \qquad \text{for } 0<p\leq 1.$$

Furthermore, both inequalities are sharp (for details, see [5]). In the present paper we have a similar situation: the inequalities formulated in Theorems 1.1, 1.2, 1.3 and 1.4 have their analogues for the sums of nonnegative random variables. Precisely, we have the following.

**Theorem 1.5.** *Let $(e_n)$ be an adapted sequence of nonnegative random variables.*

*(i) If $0<p\leq 1$, then*

$$\left|\left|\sum_{k=0}^\infty e_k\right|\right|_{p,\infty} \leq (\Gamma(p+1))^{-1/p}\left|\left|\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_k)\right|\right|_p \tag{1.17}$$

*and the constant $(\Gamma(p+1))^{-1/p}$ is the best possible.*

*(ii) If $1\leq p<\infty$ then*

$$\left|\left|\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_{k-1})\right|\right|_{p,\infty} \leq p^{1-1/p}\left|\left|\sum_{k=0}^\infty e_k\right|\right|_p \tag{1.18}$$

*and the constant $p^{1-1/p}$ is the best possible.*

*(iii) If $||\sum_{k=0}^\infty e_k||_\infty \leq 1$, then for any $\lambda>0$ we have*

$$\mathbb{P}(\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_{k-1}) \geq \lambda) \leq \min(e^{1-\lambda},1) \tag{1.19}$$

*and the bound on the right is the best possible.*

*(iv) Suppose that $\Phi:[0,\infty)\to\mathbb{R}$ is a convex and nondecreasing function. If $||\sum_{k=0}^\infty e_k||_\infty \leq 1$, then*

$$\mathbb{E}\Phi\left(\sum_{k=0}^\infty \mathbb{E}(e_k|\mathcal{F}_{k-1})\right) \leq \int_0^\infty \Phi(t)e^{-t}dt \tag{1.20}$$

*and the bound on the right is the best possible.*

A few words about the organization of the paper. In the next section we present the method which allows us to obtain the announced estimates. Then, in Section 3, we make use of the technique and provide the proofs of the inequalities. In the final section we prove that the constants in some of the estimates above can not be replaced by smaller ones.

## 2. On the method of proof

Let us first recall the notions of smoothness and convexity of Banach spaces. We say that a Banach space $B$ is $(2,\alpha)$-smooth, if for any $x$, $y \in B$ we have

$$||x+y||^2 + ||x-y||^2 \leq 2||x||^2 + 2\alpha||y||^2.$$

We say that $B$ is $(2,\alpha)$-convex, if for any $x$, $y \in B$ we have

$$||x+y||^2 + ||x-y||^2 \geq 2||x||^2 + 2\alpha||y||^2.$$

To give some examples, $L^p$ spaces are $(2,\sqrt{p-1})$-convex for $p \geq 2$ and $(2, 1/\sqrt{p-1})$-smooth for $1 < p \leq 2$. Any Hilbert space is $(2,1)$-smooth and $(2,1)$-convex.

We start with the following easy fact. Let $I$ be a subinterval of $[0,\infty)$, containing 0. Recall that the martingale is *simple* if there is $N$ such that $df_N = df_{N+1} = \ldots = 0$ almost surely and for any $n \geq 0$, the variable $f_n$ is simple, i.e. takes only a finite number of values.

**Theorem 2.1.** *Let $V : I \times [0,\infty) \to \mathbb{R}$ be fixed. The following two statements are equivalent.*

*(i) For any adapted finite sequence $(e_n)$ (that is, satisfying $0 \equiv e_N = e_{N+1} = \ldots$ for some $N$) of simple nonnegative random variables such that $\sum_{n=0}^{\infty} e_n \in I$ almost surely, we have*

$$\mathbb{E}V\left(\sum_{n=0}^{\infty} e_n, \sum_{n=0}^{\infty} \mathbb{E}(e_n|\mathcal{F}_{n-1})\right) \leq 0. \tag{2.1}$$

*(ii) For any simple real valued martingale $f$ satisfying $S^2(f) \in I$ almost surely, we have*

$$\mathbb{E}V(S^2(f), s^2(f)) \leq 0. \tag{2.2}$$

*Proof.* The proof is just a matter of the substitution

$$e_n = |df_n|^2, \qquad n = 0, 1, 2, \ldots. \tag{2.3}$$

One only needs to observe that, given $(e_n)$ as in (i), there exists a martingale $f$ for which (2.3) is valid. For example, take a sequence $(\varepsilon_n)$ of independent Rademacher variables (also independent of $(e_n)$) and consider a conditionally symmetric martingale $f$ defined by $df_n = \varepsilon_n \sqrt{e_n}$, $n = 0, 1, 2, \ldots$. □

Now we turn to the description of the technique we will use to establish the estimates formulated in the Introduction. The method converts the problem of proving a given inequality to the problem of the construction of a certain special function.

**Theorem 2.2.** *Let $I$ be a subinterval of $[0,\infty)$ such that $0 \in I$ and suppose that $U$, $V$ are functions from $I \times [0,\infty)$ to $\mathbb{R}$ satisfying*

$$V(x,y) \leq U(x,y), \quad x \in I,\ y \geq 0, \tag{2.4}$$

$$U(\cdot\,, y):\ x \mapsto U(x,y),\ x \in I, \ \textit{is concave for any } y \geq 0 \tag{2.5}$$

*and*

$$U(x+d, y+d) \leq U(x,y), \quad \textit{for } d \geq 0,\ y \geq 0 \textit{ and } x,\ x+d \in I. \tag{2.6}$$

*Then we have the following.*

*(i) If $(e_k)_{k=0}^\infty$ is a sequence of simple nonnegative variables satisfying the condition $\sum_{k=1}^\infty e_k \in I$ almost surely, then for any nonnegative integer $n$,*

$$\mathbb{E}V\left(\sum_{k=0}^n e_k, \sum_{k=0}^n \mathbb{E}(e_k|\mathcal{F}_{k-1})\right) \leq U(0,0). \tag{2.7}$$

*(ii) If $f$ is a simple $B$-valued martingale such that $S(f) \in I$ almost surely, then for any nonnegative integer $n$,*

$$\mathbb{E}V\left(S_n^2(f), s_n^2(f)\right) \leq U(0,0). \tag{2.8}$$

*(iii) If $B$ is a Hilbert space and $f$ is a simple $B$-valued martingale such that $||f|| \in I$ almost surely, then for any nonnegative integer $n$,*

$$\mathbb{E}V\left(||f_n||^2, s_n^2(f)\right) \leq U(0,0). \tag{2.9}$$

*(iv) If $U(\cdot, y)$ is nondecreasing for any $y$, $B$ is $(2,\alpha)$-smooth and $f$ is a simple $B$-valued conditionally symmetric martingale such that $||f|| \in I$ almost surely, then for any nonnegative integer $n$,*

$$\mathbb{E}V\left(\alpha^{-1}||f_n||^2, s_n^2(f)\right) \leq U(0,0). \tag{2.10}$$

*(v) If $U(\cdot, y)$ is nonincreasing for any $y$, $B$ is $(2,\alpha)$-convex and $f$ is a simple $B$-valued conditionally symmetric martingale such that $||f|| \in I$ almost surely, then for any nonnegative integer $n$,*

$$\mathbb{E}V\left(\alpha^{-1}||f_n||^2, s_n^2(f)\right) \leq U(0,0). \tag{2.11}$$

*Proof.* By (2.4), it suffices to show the assertions with $V$ replaced by $U$. Observe that the assumption on the simplicity of $(e_n)$ and $f$ guarantee the integrability of all the variables appearing in the above estimates.

(i) Denote $F_{-1} = G_{-1} \equiv 0$,

$$F_n = \sum_{k=0}^n e_k \quad \text{and} \quad G_n = \sum_{k=0}^n \mathbb{E}(e_k|\mathcal{F}_{k-1}), \quad n = 0,\, 1,\, 2,\, \ldots.$$

Since for any $n \geq 0$ the variable $G_n$ is $\mathcal{F}_{n-1}$-measurable, we have, by (2.5) and conditional Jensen's inequality,

$$\mathbb{E}U(F_n, G_n) = \mathbb{E}\big[\mathbb{E}\big(U(F_n, G_n)|\mathcal{F}_{n-1}\big)\big] \leq \mathbb{E}\big[\mathbb{E}U\big(\mathbb{E}(F_n|\mathcal{F}_{n-1}), G_n\big)\big].$$

The process $(U(\mathbb{E}(F_n|\mathcal{F}_{n-1}), G_n))_{n=0}^\infty$ is a supermartingale with respect to $(\mathcal{F}_{n-1})$. Indeed, for $n \geq 0$,

$$\begin{aligned}
&\mathbb{E}[U(\mathbb{E}(F_{n+1}|\mathcal{F}_n), G_{n+1})|\mathcal{F}_{n-1}] \\
&= \mathbb{E}\big[U\big(\mathbb{E}(F_n|\mathcal{F}_n) + \mathbb{E}(e_{n+1}|\mathcal{F}_n), G_n + \mathbb{E}(e_{n+1}|\mathcal{F}_n)\big)\big|\mathcal{F}_{n-1}\big] \\
&\leq \mathbb{E}[U(\mathbb{E}(F_n|\mathcal{F}_n), G_n)|\mathcal{F}_{n-1}] \\
&\leq U(\mathbb{E}(F_n|\mathcal{F}_{n-1}), G_n),
\end{aligned} \tag{2.12}$$

where in the first estimate we have used (2.6) and the second one follows from (2.5) and the conditional Jensen's inequality. Thus

$$\mathbb{E}U(F_n, G_n) \leq \mathbb{E}U(\mathbb{E}(F_0|\mathcal{F}_{-1}), G_0) = \mathbb{E}U(\mathbb{E}(e_0|\mathcal{F}_{-1}), \mathbb{E}(e_0|\mathcal{F}_{-1})) \leq U(0,0),$$

in view of (2.6), and the estimate follows.

(ii) This is a consequence of Theorem 2.1.

(iii) We repeat the proof of (i), word by word, this time with the processes $F_n = ||f_n||^2$ and $G_n = s_n^2(f)$, $n = 0,\, 1,\, 2,\, \ldots$. The only fact we need is that if $B$

is a Hilbert space, then $\mathbb{E}(F_{n+1}|\mathcal{F}_n) = F_n + \mathbb{E}(||df_{n+1}||^2|\mathcal{F}_n)$; therefore, (2.12) is valid, with $e_{n+1}$ replaced by $||df_{n+1}||^2$, and the claim follows.

(iv), (v) We will only show (iv), the arguments leading to (v) are similar. Adding the variable $f_{-1} \equiv 0$ if necessary, we may assume that $f$ starts from 0. Clearly, we will be done if we show that the process $(U(\alpha^{-1}||f_n||^2, s_n^2(f)))$ is a supermartingale (with respect to the filtration $(\mathcal{F}_{n-1})$). To this end, fix $n \geq 1$ and note that, by the conditional symmetry of $f$,

$$\begin{aligned} &\mathbb{E}[U(\alpha^{-1}||f_n||^2, s_n^2(f))|\mathcal{F}_{n-1}] \\ &\quad = \frac{1}{2}\mathbb{E}[U(\alpha^{-1}||f_{n-1} + df_n||^2, s_n^2(f)) + U(\alpha^{-1}||f_{n-1} - df_n||^2, s_n^2(f))|\mathcal{F}_{n-1}]. \end{aligned} \tag{2.13}$$

Now, by the concavity and monotonicity of $U$, we infer that there is a nonnegative variable $A = A(\alpha^{-1}||f_{n-1}||^2 + ||df_n||^2, s_n^2(f))$ such that

$$\begin{aligned} \mathbb{E}[U(\alpha^{-1}||f_{n-1} \pm df_n||^2, s_n^2(f))|\mathcal{F}_{n-1}] &\leq U(\alpha^{-1}||f_{n-1}||^2 + ||df_n||^2, s_n^2(f)) \\ &\quad + A\left[\alpha^{-1}||f_{n-1} \pm df_n||^2 - (\alpha^{-1}||f_{n-1}||^2 + ||df_n||^2)\right]. \end{aligned}$$

Using the $(2,\alpha)$-smoothness property of $B$,

$$A[\alpha^{-1}||f_{n-1} + df_n||^2 + \alpha^{-1}||f_{n-1} - df_n||^2 - 2(\alpha^{-1}||f_{n-1}||^2 + ||df_n||^2)] \leq 0,$$

so, by (2.13),

$$\begin{aligned} \mathbb{E}[U(\alpha^{-1}||f_n||^2, s_n^2(f))|\mathcal{F}_{n-1}] &\leq \mathbb{E}[U(\alpha^{-1}||f_{n-1}||^2 + ||df_n||^2, s_n^2(f))|\mathcal{F}_{n-1}] \\ &\leq \mathbb{E}[U(\alpha^{-1}||f_{n-1}||^2, s_{n-1}^2(f))|\mathcal{F}_{n-1}] \\ &= U(\alpha^{-1}||f_{n-1}||^2, s_{n-1}^2(f)), \end{aligned}$$

where in the second inequality we have exploited (2.6). The proof is complete. □

## 3. Proofs of the inequalities (1.5) – (1.20)

### 3.1. **Weak type estimates.**

*Proof of the inequality* (1.17). Using standard approximation arguments, we may restrict ourselves to finite sequences $(e_n)$ of simple nonnegative random variables. We must show that for any $\lambda > 0$ and $n = 0, 1, 2, \ldots$,

$$\lambda^p \mathbb{P}(\sum_{k=0}^n e_k \geq \lambda) \leq \Gamma(p+1)^{-1}\mathbb{E}\left(\sum_{k=0}^n \mathbb{E}(e_k|\mathcal{F}_{k-1})\right)^p.$$

By homogeneity, we may and will assume $\lambda = 1$; then the inequality can be written in the form

$$\mathbb{E}\left[1_{\{\sum_{k=0}^n e_k \geq 1\}} - \Gamma(p+1)^{-1}\left(\sum_{k=0}^n \mathbb{E}(e_k|\mathcal{F}_{k-1})\right)^p\right] \leq 0.$$

Now we will introduce the special functions $U_p$, $V_p : [0,\infty) \times [0,\infty) \to \mathbb{R}$ for which the above inequality is of the form (2.7). Let

$$V_p(x,y) = 1_{\{x \geq 1\}} - \Gamma(p+1)^{-1}y^p$$

and

$$U_p(x,y) = \begin{cases} 1 - \Gamma(p+1)^{-1}\left[(1-x)e^y \int_y^\infty t^p e^{-t}dt + xy^p\right] & \text{if } x < 1, \\ 1 - \Gamma(p+1)^{-1}y^p & \text{if } x \geq 1. \end{cases}$$

Observe that $U_p(\cdot,y)$ is nondecreasing for any fixed $y$. Therefore, by part (iv) of Theorem 2.2, the proof will be complete if we show that the functions satisfy (2.4), (2.5) and (2.6). To establish the majorization $V_p \le U_p$, observe first that we have equality if $x \ge 1$. If $x < 1$, then the inequality can be written in the equivalent form

$$(1-x)\left(e^y\int_y^\infty t^pe^{-t}dt - y^p\right) \le \Gamma(p+1).$$

This holds true, since $1-x\le 1$, the function $H:[0,\infty)\to[0,\infty)$, given by

$$H(y) := e^y\int_y^\infty t^pe^{-t}dt - y^p = pe^y\int_y^\infty t^{p-1}e^{-t}dt$$

is nonincreasing: $H'(y) = p(p-1)e^y\int_y^\infty t^{p-2}e^{-t}dt < 0$, and $H(0)=\Gamma(p+1)$. The condition (2.5) is guaranteed by the fact that $U_p(\cdot,y)$ continuous, linear and increasing on $[0,1]$, and constant on $[1,\infty)$. Finally, to check (2.6), it suffices to show that $U_{px}(x,y)+U_{py}(x,y)\le 0$ for $x\neq 1$, $y>0$. This is clear for $x>1$, while for $x<1$ we have that

$$U_{px}(x,y)+U_{py}(x,y) = Cp(p-1)xe^y\int_y^\infty t^{p-2}e^{-t}dt \le 0. \quad \square$$

*Proof of* (1.5), (1.6) *and* (1.7). We may assume that the martingales we deal with are simple. Now taking $U=U_{p/2}$ and $V=V_{p/2}$ (with $U_p$ and $V_p$ defined above) turns (1.5), (1.6) and (1.7) into (2.8), (2.9) and (2.11), respectively - and the latter three estimates are valid in view of Theorem 2.2. $\square$

*Proof of the inequality* (1.18). As previously, we will deduce the estimate from Theorem 2.2 applied to appropriate $U$ and $V$. Let $\gamma:[1-1/p,1)\to\mathbb{R}$ be given by

$$\gamma(y) = \frac{1}{p}(p(1-y))^{-1/(p-1)}.$$

Consider the following subsets of $[0,\infty)\times[0,\infty)$:

$$\begin{aligned}
&D_1 = [0,\infty)\times[0,1-1/p],\\
&D_2 = \{(x,y): x>\gamma(y)+y-1,\ 1-1/p<y<1\},\\
&D_3 = ([0,\infty)\times[0,1))\setminus(D_1\cup D_2),\\
&D_4 = [0,\infty)\times[1,\infty).
\end{aligned}$$

Define $U_p$, $V_p:[0,\infty)\times[0,\infty)\to\mathbb{R}$ by

$$U_p(x,y) = \begin{cases} p^{p-1}(\frac{y}{p-1})^{p-1}[y-px] & \text{on } D_1,\\ \frac{1}{p(1-y)}[(p-1)p^{-p/(p-1)}(1-y)^{-1/(p-1)}-px] & \text{on } D_2,\\ 1-p^{p-1}(1+x-y)^p & \text{on } D_3,\\ 1-p^{p-1}x^p & \text{on } D_4\end{cases}$$

and $V_p(x,y) = 1_{[1,\infty)}(y) - p^{p-1}x^p$.

Arguing as previously, it suffices to show that for $n=0,\,1,\,2,\,\ldots$,

$$\mathbb{E}\left[1_{\{\sum_{k=0}^n\mathbb{E}(e_k|\mathcal{F}_{k-1})\ge 1\}} - p^{p-1}\left(\sum_{k=0}^n e_k\right)^p\right] \le 0.$$

This is precisely the estimate (2.7), for the above choice of the functions $U_p$ and $V_p$. One easily verifies that $U_p(\cdot,y)$ is nonincreasing for fixed $y$. Therefore, to complete the proof, in view of Theorem 2.2, it suffices to check that $U_p$, $V_p$ satisfy the conditions (2.4), (2.5) and (2.6). To prove (2.4), observe that if $(x, y) \in D_1$, then

$$U_{px}(x,y) - V_{px}(x,y) = p^p \left[ x^{p-1} - \left( \frac{y}{p-1} \right)^{p-1} \right],$$

which implies that

$$U_p(x,y) - V_p(x,y) \geq U_p(y/(p-1), y) - V_p(y/(p-1), y) = 0.$$

If $(x, y) \in D_2$, then

$$U_{px}(x,y) - V_{px}(x,y) = -\frac{1}{1-y} + p^p x^{p-1}.$$

Setting

$$x_0 = \frac{1}{p}[p(1-y)]^{-1/(p-1)},$$

we see that $(x_0, y) \in D_2$ and $U_p(x,y) - V_p(x,y) \geq U_p(x_0, y) - V_p(x_0, y) = 0$. If $(x, y)$ belongs to $D_3$, then $U_{px}(x,y) - V_{px}(x,y) = p^p[x^{p-1} - (1 + x - y)^{p-1}] < 0$, so

$$U_p(x,y) - V_p(x,y) \geq U_p(\gamma(y) + y - 1, y) - V_p(\gamma(y) + y - 1, y),$$

and the right hand side is nonnegative, as $(\gamma(y) + y - 1, y)$ belongs to the closure of $D_2$, where we have already established the majorization. We complete the proof of (2.4) noting that we have $U_p = V_p$ on $D_4$. The condition (2.5) is apparent. To establish (2.6) it suffices to prove that $U_{px} + U_{py} \leq 0$ in the interiors $D_i^\circ$ of the sets $D_i$, $i = 1, 2, 3, 4$. The direct calculation shows that $U_{px}(x,y) + U_{py}(x,y)$ equals

$$\begin{cases} -p^p(p-1)^{2-p} x y^{p-2} & \text{if } (x,y) \in D_1^o, \\ (1-y)^{-2}[-x + y - 1 + \gamma(y)] & \text{if } (x,y) \in D_2^o, \\ 0 & \text{if } (x,y) \in D_3^o, \\ -p^p x^{p-1} & \text{if } (x,y) \in D_4^o \end{cases}$$

and it is evident that all the expressions are nonpositive on the corresponding sets. □

*Proof of* (1.8), (1.9) *and* (1.10). Theorem 2.2 applied to $U_{p/2}$, $V_{p/2}$ gives us the desired estimates under the additional assumption that the martingales are simple. The general case follows by standard approximation. □

In the case $p \geq 2$, we have the following extension of (1.10) to the case of general martingales.

**Corollary 3.1.** *Let $p \geq 2$. If $B$ is $(2, \alpha)$-convex and $f$ is a $B$-valued martingale, then*

$$||s(f)||_{p,\infty} \leq 2(p/2)^{1/2-1/p} \alpha^{-1} ||f||_p. \tag{3.1}$$

*Proof.* This can be proved using standard decoupling techniques. Consider the probability space $(\Omega \times \Omega, \mathcal{F} \otimes \mathcal{F}, \mathbb{P} \otimes \mathbb{P})$ and let $f'$, $f''$ be independent copies of $f$, given by $f'_n(\omega, \omega') = f''_n(\omega', \omega) = f_n(\omega')$. Introduce the product filtration $(\mathcal{F}'_n) = (\mathcal{F}_n \otimes \mathcal{F}_n)$. Then $h = f' - f''$ is a conditionally symmetric martingale with respect to $(\mathcal{F}'_n)$ and it follows from the conditional Jensen's inequality that

$$\mathbb{E}(||dh_n||^2 | \mathcal{F}'_{n-1}) \geq \mathbb{E}(||df_n||^2 | \mathcal{F}_{n-1}).$$

Therefore

$$||s^2(f)||_{p,\infty} \le ||s^2(h)||_{p,\infty} \le p^{1-1/p}\alpha^{-1}||h||_p \le 2p^{1/2-1/p}\alpha^{-1}||f||_p. \quad \square$$

3.2. **The bounded case.** Here we will study the versions of the estimates above in the case $p = \infty$. We will only focus on the inequalities for the finite sums of nonnegative simple random variables; as we have already seen, this easily extends to the general sequences. Furthermore, the martingale versions follow immediately.

*Proof of* (1.19). For $\lambda \le 1$ the inequality is trivial, so we may assume that $\lambda > 1$. Let $U$, $V : [0,1] \times [0,\infty) \to \mathbb{R}$ be given by $V(x,y) = 1_{\{y \ge \lambda\}}$ and

$$U(x,y) = \begin{cases} (1-x)\exp(y+1-\lambda), & \text{if } y \le \lambda - 1, \\ (1-x)(\lambda-y)^{-1}, & \text{if } \lambda - 1 < y < x + \lambda - 1, \\ 1, & \text{if } y \ge x + \lambda - 1. \end{cases}$$

It is straightforward to verify the conditions (2.4) and (2.5). Furthermore, (2.6) follows from the continuity of $U$ on $[0,1) \times \mathbb{R}$ and the fact that

$$U_x(x,y) + U_y(x,y) = \begin{cases} -x\exp(y+1-\lambda), & \text{if } y < \lambda - 1, \\ (1-x-\lambda+y)(\lambda-y)^{-2} & \text{if } \lambda - 1 < y < x + \lambda - 1, \\ 0, & \text{if } y > x + \lambda - 1 \end{cases}$$

is nonpositive. Hence, by Theorem 2.2, we have, for any $n$,

$$\mathbb{P}(\sum_{k=0}^{n} \mathbb{E}(e_k|\mathcal{F}_{k-1}) \ge \lambda) \le e^{1-\lambda}.$$

This yields (1.19). $\square$

*Proof of* (1.20). With no loss of generality we may assume that $\Phi$ is of class $C^1$. The functions $U$, $V : [0,1] \times [0,\infty) \to \mathbb{R}$ corresponding to our problem are given by $V(x,y) = \Phi(y)$ and

$$U(x,y) = x\Phi(y) + (1-x)e^y \int_y^\infty e^{-z}\Phi(z)dz.$$

Since $\Phi$ is nondecreasing, we have

$$e^y \int_y^\infty e^{-z}\Phi(z)dz \ge e^y \int_y^\infty e^{-z}\Phi(y)dz = \Phi(y)$$

and (2.4) follows. For a fixed $y$, the function $U(\cdot, y)$ is linear, so (2.5) is satisfied. Finally, to check (2.6), observe that if $x$, $y$, $d$ are as assumed, then $U_x(x,y)+U_y(x,y)$ equals

$$x\left[\Phi'(y) + \Phi(y) - e^y \int_y^\infty e^{-z}\Phi(z)dz\right] = x\left[\Phi'(y) - e^y \int_y^\infty e^{-z}\Phi'(z)dz\right] \le 0,$$

where we have used integration by parts and the fact that $\Phi'$ is nondecreasing.

Hence, by Theorem 2.2, for any $n$,

$$\mathbb{E}\Phi\left(\sum_{k=0}^{n} \mathbb{E}(e_k|\mathcal{F}_{k-1})\right) \le \int_0^\infty \Phi(t)e^{-t}dt,$$

which is what we need. $\square$

We conclude this section presenting the result on $\Phi$-inequalities for concave $\Phi$. The fact is well known and very easy to prove.

**Theorem 3.2.** *Suppose $\Phi : [0,\infty) \to \mathbb{R}$ is a concave function.*

*(i) For any sequence $(e_n)_{n=0}^\infty$ of nonnegative random variables, which satisfies $||\sum_{n=0}^\infty e_n||_\infty \leq 1$, we have*

$$\mathbb{E}\Phi\left(\sum_{n=0}^\infty \mathbb{E}(e_n|\mathcal{F}_n)\right) \leq \Phi(1) \tag{3.2}$$

*and the inequality is sharp.*

*(ii) For any $B$-valued martingale $f$ satisfying $||S(f)||_\infty \leq 1$ we have*

$$\mathbb{E}\Phi(s^2(f)) \leq \Phi(1) \tag{3.3}$$

*and the bound on the right is the best possible, even if $B = \mathbb{R}$.*

*(iii) If $B$ is a Hilbert space, then for any $B$-valued martingale $f$ satisfying $||f||_\infty \leq 1$, we have*

$$\mathbb{E}\Phi(s^2(f)) \leq \Phi(1) \tag{3.4}$$

*and the bound on the right is the best possible, even if $B = \mathbb{R}$.*

*(iv) If $B$ is $(2,\alpha)$-convex Banach space, then for any $B$-valued conditionally martingale $f$ satisfying $||f||_\infty \leq 1$, we have*

$$\mathbb{E}\Phi(s^2(f)) \leq \Phi(\alpha^{-1}). \tag{3.5}$$

*Proof.* This is straightforward: the estimates follow from Jensen's inequality and the fact that if $f$ is conditionally symmetric, then $\alpha\mathbb{E}s^2(f) \leq \lim_{n\to\infty} \mathbb{E}||f_n||^2 \leq 1$. To see that the bounds in (3.2), (3.3), (3.4) are optimal, consider constant sequences $e_0 = e_1 = e_2 = \ldots \equiv 1$, $f_0 = f_1 = f_2 = \ldots \equiv 1$. □

## 4. Sharpness

In this section we deal with the optimality of the constants appearing in the inequalities established above. We will focus only on the martingale estimates; in view of Theorem 2.1, this will show the sharpness of their analogues for the sums of nonnegative random variables.

4.1. **Sharpness of** (1.5)**,** (1.6)**.** Let $\delta \in (0,1)$ be fixed and let $(X_n)_{n=0}^\infty$ be a sequence of independent random variables sharing the same distribution given by

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

Furthermore, let $(\varepsilon_n)$ be a sequence of independent Rademacher variables, independent also of $(X_n)$. Introduce the stopping time $\tau = \inf\{n : X_n = 1\}$, set $df_n = \varepsilon_n X_n 1_{\{\tau\geq n\}}$, $n = 0, 1, 2, \ldots$ and let $(\mathcal{F}_n)$ be the natural filtration of $f$. Then $f$ is a martingale (which is even conditionally symmetric), for which $|f_n| \uparrow |f_\infty| \equiv 1$ and $S(f) \equiv 1$ almost surely; hence $||f||_{p,\infty} = ||s(f)||_{p,\infty} = 1$. 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 $s^2(f) = \delta(\tau+1)$. Since $\tau$ has geometric distribution, we have, 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 $\delta$ sufficiently small, can be made arbitrarily close to $\int_0^\infty t^{p/2}e^{-t}dt = \Gamma(p/2+1)$. This implies that the constant in (1.5) and (1.6) is indeed the best possible.

4.2. **Sharpness of** (1.8) **and** (1.9)**.** If $p=2$, then the constant martingale $f_0 = f_1 = f_2 = \ldots \equiv 1$ gives equality in (1.8) and (1.9). Suppose then, that $p>2$. Let $\delta \in (0, 1-2/p)$, $N$ be a positive integer and set

$$r = \left(\frac{p-2}{p\delta}\right)^{1/N} > 1. \tag{4.1}$$

Furthermore, assume that $N$ is large enough to guarantee that $q := (r-1)(p-2)/(2r) < 1$. Consider a sequence $(X_n)_{n=0}^N$ of independent random variables such that

$$\mathbb{P}\left(X_n = \frac{2r^n\delta}{p-2}\right) = q = 1 - \mathbb{P}(X_n = 0), \quad n = 0,\,1,\,2,\,\ldots,$$

and $X_N \equiv 2/p$. Let $(\varepsilon_n)_{n=0}^N$ be a sequence of independent Rademacher variables, independent also of $(X_n)$. Set $\tau = \inf\{n : X_n \neq 0\}$ and let $df_n = \varepsilon_n\sqrt{X_n}1_{\{\tau\geq n\}}$, $n=0,\,1,\,2,\,\ldots,\,N$. We easily see that $f$ is a conditionally symmetric martingale satisfying $|f_N| = S_N(f) = \sqrt{X_\tau}$. Therefore

$$\begin{aligned} ||f||_p^p = ||S(f)||_p^p &= \sum_{n=0}^{N-1}\left(\frac{2r^n\delta}{p-2}\right)^{p/2}(1-q)^n q + \left(\frac{2}{p}\right)^{p/2}(1-q)^N \\ &= \left(\frac{2\delta}{p-2}\right)^{p/2} q\frac{1-(r^{p/2}(1-q))^N}{1-r^{p/2}(1-q)} + \left(\frac{2}{p}\right)^{p/2}(1-q)^N. \end{aligned} \tag{4.2}$$

On the other hand, it can be easily verified that $\mathbb{E}(df_n^2|\mathcal{F}_{n-1}) = (r-1)r^{n-1}\delta 1_{\{\tau\geq n\}}$ for $n<N$ and $\mathbb{E}(df_N^2|\mathcal{F}_{N-1}) = 2/p1_{\{\tau=N\}}$, so

$$s^2(f) = \sum_{n=0}^{\tau\wedge(N-1)}(r-1)r^{n-1}\delta + \frac{2}{p}1_{\{\tau=N\}} = \frac{\delta}{r}(r^{(\tau+1)\wedge N}-1) + \frac{2}{p}1_{\{\tau=N\}}.$$

On the set $\{\tau=N\}$ we have, by (4.1),

$$s^2(f) = \frac{p-2}{pr} - \frac{\delta}{r} + \frac{2}{p} \geq \frac{1-\delta}{r},$$

so

$$\mathbb{P}\left(s(f) \geq \left(\frac{1-\delta}{r}\right)^{1/2}\right) \geq \mathbb{P}(\tau=N) = (1-q)^N.$$

Hence

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

Now we will let $N\to\infty$ (so $r$ tends to 1). We have

$$\lim_{N\to\infty}\frac{q}{1-r^{p/2}(1-q)} = \lim_{r\to 1}\left(\frac{1-r^{p/2}}{(r-1)(p-2)}\cdot 2r + r^{p/2}\right)^{-1} = -\frac{p-2}{2}$$

and, by (4.1),

$$\lim_{N\to\infty}(1-q)^N = \lim_{r\to 1}\left(1-\frac{(r-1)(p-2)}{2r}\right)^{\log\frac{p-2}{p\delta}/\log r} = \left(\frac{p\delta}{p-2}\right)^{(p-2)/2}.$$

Therefore, again using (4.1),

$$\begin{aligned}\lim_{N\to\infty}\frac{||f||_p^p}{(1-q)^N} &= \left(\frac{2\delta}{p-2}\right)^{p/2}\cdot\left(-\frac{p-2}{2}\right)\cdot\left(\frac{p-2}{p\delta}\right)^{(p-2)/2}\\ &\quad+\left(\frac{2}{p-2}\right)^{p/2}\cdot\frac{p-2}{2}\cdot\left(\frac{p-2}{p}\right)^{p/2}+\left(\frac{2}{p}\right)^{p/2}\\ &= -\delta\left(\frac{2}{p}\right)^{(p-2)/2}+\left(\frac{2}{p}\right)^{p/2-1}.\end{aligned}$$

Now if $\delta$ is taken sufficiently small, we see that for any $\kappa>0$ the ratios $||f||_p^p/||s(f)||_p$ and $||S(f)||_p^p/||s(f)||_p$ can be made smaller than $(2/p)^{p/2-1}+\kappa$. This shows that the constant $(p/2)^{1/2-1/p}$ is indeed the best possible in (1.8) and (1.9).

4.3. **Sharpness of** (1.11) **and** (1.12)**.** If $\lambda\leq 1$ then we have equalities if we take $e_0=e_1=e_2=\ldots\equiv 1$ and $f_0=f_1=f_2=\ldots\equiv 1$. Suppose that $\lambda>1$, take a positive integer $N$ and set $\delta=(\lambda^2-1)/N$. The example is similar to the one used in Section 5.1. Let $(X_n)_{n=0}^N$ be a sequence of independent random variables such that

$$\mathbb{P}(X_n=1)=\delta=1-\mathbb{P}(X_n=0),\qquad n=0,\,1,\,2,\,\ldots,\,N-1$$

and $X_N\equiv 1$. Finally, let $\tau=\inf\{n:X_n=1\}$, $(\varepsilon_n)$ be a sequence of independent Rademacher variables and $df_n=\varepsilon_nX_n1_{\{\tau\geq n\}}$, $n=0,\,1,\,2,\,\ldots,\,N$, $df_n\equiv 0$ for $n>N$.

We easily check that $||f||_\infty=||S(f)||=1$ (in fact, for any $0\leq n\leq N-1$, $S_n(f)$, $|f_n|\in\{0,1\}$ and $S_N(f)=|f_N|=1$ with probability 1). Moreover, we see that $\mathbb{E}(df_n^2|\mathcal{F}_{n-1})=\delta 1_{\{\tau\geq n\}}$ almost surely for $n<N$ and hence $s^2(f)=(\tau+1)\delta\leq\lambda-1<\lambda$ on $\tau<N$; on the other hand, as $\mathbb{E}(df_N^2|\mathcal{F}_{N-1})=1$ with probability 1, we have $s(f)=\lambda$ on $\{\tau=N\}$ and hence

$$\mathbb{P}(s(f)\geq\lambda)=(1-\delta)^N.$$

It suffices to note that the right hand side converges to $e^{1-\lambda^2}$ as $N\to\infty$. Therefore (1.11) and (1.12) are sharp.

4.4. **Sharpness of** (1.14) **and** (1.15)**.** For $\delta\in(0,1)$, let $f$ be a martingale as in Subsection 5.1. We have $||f||_\infty=||S(f)||_\infty=1$,

$$\mathbb{E}\Phi(s^2(f))=\delta\sum_{n=1}^{\infty}\Phi(\delta n)(1-\delta)^n,$$

which, if $n$ is chosen sufficiently large, can be made arbitrarily close to $\int_0^\infty\Phi(t)e^{-t}dt$. This shows the bounds in (1.14) and (1.15) are optimal.

## References


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

[2] A. M. Garsia, *Martingale Inequalities: Seminar Notes on Recent Progress*, Benjamin, Reading, Massachusetts, 1973.

[3] P. Hitczenko, *Best constants in martingale version of Rosenthal's inequality*, Ann. Probab. **18** (1990), no. 4, 1656–1668.

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

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

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