ISRAEL JOURNAL OF MATHEMATICS 84 (1993), 161–178

DOMINATION INEQUALITY FOR MARTINGALE TRANSFORMS OF A RADEMACHER SEQUENCE

BY

PAWEŁ HITCZENKO*

Department of Mathematics
North Carolina State University, Raleigh, NC 27695-8205, USA

ABSTRACT

Let $f_n = \sum_{k=1}^n v_k r_k$, $n = 1,\ldots,$ be a martingale transform of a Rademacher sequence (r_n) and let (r'_n) be an independent copy of (r_n). The main result of this paper states that there exists an absolute constant K such that for all p, $1 \le p < \infty$, the following inequality is true:

$$\left\| \sum v_k r_k \right\|_p \le K \left\| \sum v_k r'_k \right\|_p .$$

In order to prove this result, we obtain some inequalities which may be of independent interest. In particular, we show that for every sequence of scalars (a_n) one has

$$\left\| \sum a_k r_k \right\|_p \approx K_{1,2}((a_n), \sqrt{p}),$$

where $K_{1,2}((a_n), \sqrt{p}) = K((a_n), \sqrt{p}; \ell_1, \ell_2)$ is the K-interpolation norm between ℓ_1 and ℓ_2. We also derive a new exponential inequality for martingale transforms of a Rademacher sequence.

1. Introduction

This paper concerns the L_p-norm inequalities for certain martingale transforms. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let $(\mathcal{F}_n)$ be an increasing sequence of sub-σ-algebras of $\mathcal{F}$. A sequence (X_n) of random variables is said to be $(\mathcal{F}_n)$-predictable (or just predictable, if there is no risk of confusion) provided for each $n \ge 1$, X_n is

* This research was supported in part by an NSF grant and an FRPD grant at NCSU.
Received August 24, 1992 and in revised form January 12, 1993

$\mathcal{F}_{n-1}$-measurable random variable. All equalities or inequalities between random variables are assumed to hold almost surely. For any sequence (X_n) of random variables we shall write $X^* = \sup_n |X_n|$ and $X_n^* = \max_{1\le k\le n} |X_k|$. The L_p-norm, $1 \le p \le \infty$, of a random variable X is denoted by $\|X\|_p$, and $I(A)$ or I_A will denote the indicator function of a set A. Let (ξ_n) be a sequence of independent, mean-zero random variables and set $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$, $n = 1, \dots$. If (v_n) is a sequence predictable relative to $(\mathcal{F}_n)$ then a sequence (f_n) defined by

$$f_n = \sum_{k=1}^{n} v_k \xi_k, \qquad n = 1, \dots$$

is a martingale and is referred to as martingale transform of the sequence (ξ_k). Denote by (ξ_n') an independent copy of a sequence (ξ_n). It is known that for each p, $1 \le p < \infty$, there exist constants C_p and K_p depending only on p such that for all predictable sequences (v_n) of random variables the following two-sided estimate holds:

$$(*) \qquad C_p^{-1} \Big\| \sum_{k=1}^{n} v_k \xi_k' \Big\|_p \le \Big\| \sum_{k=1}^{n} v_k \xi_k \Big\|_p \le K_p \Big\| \sum_{k=1}^{n} v_k \xi_k' \Big\|_p, \quad n = 1, \dots$$

Several proofs of $(*)$ as well as of more general results are available (cf. e.g. Zinn [26], Hitczenko [10, 13] or McConnell [23]). The proofs given in the last two papers show that both constants C_p and K_p are bounded above by $O(p)$ for large p. It is not hard to see that the above bound on C_p is optimal i.e. that there are sequences (ξ_n) and (v_n) for which:

$$\Big\| \sum_{k=1}^{n} v_k \xi_k' \Big\|_p \approx O(p) \Big\| \sum_{k=1}^{n} v_k \xi_k \Big\|_p.$$

Indeed, if (r_n) is a Rademacher sequence (i.e. a sequence of i.i.d. random variables such that $P(r_n = 1) = P(r_n = -1) = 1/2$) and $v_k = I(\tau \ge k)$, where $\tau = \inf\{n \in \mathrm{N}: r_n = r_{n-1}\}$, then $\|\sum v_k r_k\|_p \le 1$. On the other hand, denoting by $[x]$ the integer part of a number x, we have that

$$\begin{aligned} \Big\| \sum v_k r_k' \Big\|_p^p \ge & \sum_{k=1}^{\infty} P(\tau = k) E\Big| \sum_{j=1}^{k} r_j' \Big|^p \ge P(\tau = [p]) E\Big| \sum_{j=1}^{[p]} r_j' \Big|^p \\ \ge & P(\tau = [p]) [p]^p P(r_j' = 1, j = 1, \dots, [p]) \approx (p/4)^p. \end{aligned}$$

In this paper we focus our attention on the right inequality in $(*)$, which is more interesting from the point of view of applications. Our main result asserts that, if $\xi_k = r_k$ then the constant K_p may actually be chosen to be ***independent*** of p. This should be compared with a result of Klass [16, Theorem 3.1] who showed that if (ξ_k) is a sequence of independent random variables and $v_k = I(\tau \geq k)$, where τ is a stopping time with respect to (ξ_k), then $(*)$ holds with constant K_p independent of p.

Our proof relies on martingale methods, and we believe that some aspects of our approach may be of independent interest. In particular, we obtain a two-sided inequality for the L_p-norm of the sum $\sum a_k r_k$ in terms of $K((a_n), \sqrt{p}; \ell_1, \ell_2)$; the K-interpolation norm of a sequence (a_n). We also establish a new exponential inequality for martingale transforms of a Rademacher sequence.

In view of the main result of this paper one may be tempted to ask whether the following statement is true: There exists a constant C such that for all sequences (v_n) predictable with respect to the natural filtration generated by a Rademacher sequence (r_n) and all $t > 0$ one has

$$P(|\sum v_k r_k| \geq Ct) \leq CP(|\sum v_k r'_k| \geq t),$$

where (r'_n) is an independent copy of (r_n) (such inequality was actually conjectured by Kwapień and Woyczyński [19]). A negative answer to this question was given recently by Talagrand (personal communication). We shall present his example and would like to thank him for permitting us to include it in this paper.

The paper is organized as follows. The next section contains some preliminary lemmas. Sections 3 and 4 contain the main ingredients needed for the proof. In the former we prove an approximation for the L_p-norm of Rademacher averages, while in the latter we obtain an exponential inequality for martingale transforms of (r_n). The main result is stated in Section 5. It may be easily generalized to martingale transforms satisfying certain regularity condition. This extension is not completely satisfactory, since the constant K depends on the regularity constant of the original martingale. Lastly, in Section 6, we present Talagrand's example.

2. Preliminaries

As we mentioned in the introduction, we shall use martingale methods. We begin by recalling briefly necessary terminology and we refer the reader to the paper

Burkholder and Gundy [6] for more details.

Given a martingale $f = (f_n)$ with difference sequence $d = (d_n)$ and two stopping times ν and μ such that $\nu \geq \mu$, for $n = 1, \ldots,$ we let

$$^{\mu}f_n^{\nu} = \sum_{k=1}^{n} I(\mu < k \leq \nu) d_k.$$

Then the sequence $^{\mu}f^{\nu} = (^{\mu}f_n^{\nu})$ is also a martingale (referred to as f started at μ and stopped at ν). In particular, $f^n = (f_0, \ldots, f_{n-1}, f_n, f_n, \ldots)$. Let $(\mathcal{F}_n)$ be an increasing sequence of σ-algebras and assume that $\mathcal{N}$ is a collection of martingales relative to $(\mathcal{F}_n)$ which is closed under starting and stopping (i.e., $f \in \mathcal{N}$ implies $^{\mu}f^{\nu} \in \mathcal{N}$ for all stopping times ν and μ satisfying $\nu \geq \mu$). We shall consider an operator T defined on a collection $\mathcal{N}$ with values in the class of all nonnegative random variables on the same probability space. We shall assume that T satisfies the following conditions (cf. Burkholder and Gundy [6]):

(B1) T is quasi-linear, i.e., $T(f+g) \leq \gamma(T(f) + T(g))$, for some nonnegative γ and all martingales $f, g \in \mathcal{N}$.

(B2) T is local, i.e., $T(f) = 0$ on the set $\{s(f) = 0\}, f \in \mathcal{N}$ where

$$s(f) = \Big(\sum E(d_k^2 | \mathcal{F}_{k-1})\Big)^{1/2}$$

is the conditional square function of f.

(B3) T is symmetric, i.e., $T(f) = T(-f)$, for all martingales $f \in \mathcal{N}$.

If $\gamma = 1$ then T is sublinear. An operator T is called **measurable** (resp. **predictable**) if, for $n = 1, \ldots,$ $T(f^n)$ is $\mathcal{F}_n$-measurable (resp. $\mathcal{F}_{n-1}$-measurable) random variable. For example, the square function $S(f)$ defined by $S(f) = (\sum d_k^2)^{1/2}$ is a measurable operator while the conditional square function is a predictable operator on the collection of all martingales relative to $(\mathcal{F}_n)$.

Our first lemma gives a sufficient condition for a comparison of the L_p-norms of random variables and can be found in Burkholder [4, Lemma 7.1]:

LEMMA 2.1: *Fix $0 < p < \infty$. Suppose that X and Y are nonnegative random variables and that there exist positive numbers δ, $\beta > 1 + \delta$ and $\epsilon < 1/\beta^p$ such that for all positive λ's one has*

$$P(X \geq \beta\lambda, Y < \delta\lambda) \leq \epsilon P(X \geq \lambda).$$

Then

$$\|X\|_p^p \leq \frac{(\beta/\delta)^p}{1 - \beta^p \epsilon} \|Y\|_p^p.$$

The above lemma in a more general form was originally used to prove some martingale inequalities (cf. Burkholder [4], Burkholder and Gundy [6] and Burkholder, Davis and Gundy [5]. It was realized later (see e.g. Bañuelos [1] or Hitczenko [11, 12]) that it could also be used to give rather precise information on the growth rate of the constants involved in some of those inequalities. The next lemma hints at the way we are going to use that idea.

LEMMA 2.2: *Fix $1 \le p < \infty$. Let $\mathcal{N}$ be any class of martingales which is closed under starting and stopping and let T be a predictable operator on $\mathcal{N}$ satisfying conditions* (B1)–(B3). *Suppose that there exist positive constants c_0 and κ such that, for all mean-zero martingales $f \in \mathcal{N}$ the following inequality holds:*

$$P(|f_n| > c\|T^*(f^n)\|_\infty) \le 2\exp\left\{-\kappa c^2 p\right\}, \quad c \ge c_0$$

(recall that, according to our notation, $T^*(f^n) = \max_{1\le k\le n} T(f^k)$). Then, there exists an absolute constant $\vartheta = \vartheta(\kappa, \gamma)$ such that, for all $\delta > 0$ and $\beta > 1 + \delta(1 + c_0)$, we have that

$$P(f^* \ge \beta\lambda, T^*(f) \vee d^* < \delta\lambda) \le 2\exp\left\{-\frac{\vartheta(\beta - 1 - \delta)^2 p}{\delta^2}\right\} P(f^* \ge \lambda).$$

Consequently, if additionally, $\|d^*\|_p \le \alpha\|T^*(f)\|_p$, for some absolute constant α then

$$\|f^*\|_p \le C\|T^*(f)\|_p,$$

where C is an absolute constant.

Proof: Our proof follows usual argument; we let

$$\nu = \inf\{n \in \mathrm{N}: |f_n| \ge \beta\lambda\},$$
$$\mu = \inf\{n \in \mathrm{N}: |f_n| \ge \lambda\},$$
$$\tau = \inf\{n \in \mathrm{N}: T(f^{n+1}) \vee d_n^* \ge \delta\lambda\}.$$

Then, on the set $\{\nu < \infty, \tau = \infty\}$ we have

$$\Big|\sum_{k=\mu+1}^{\nu\wedge\tau} d_k\Big| \ge \Big|\sum_{k=1}^{\nu\wedge\tau} d_k\Big| - \Big|\sum_{k=1}^{\mu-1} d_k\Big| - |d_\mu| \ge (\beta - 1 - \delta)\lambda.$$

Moreover, by computation similar to that in Hitczenko [12, proof of Theorem 3.1] or Burkholder and Gundy [6, bottom half of page 256] we infer that

$\|T^*({}^\mu f^{\nu\wedge\tau})\|_\infty \le 2\gamma^2\delta\lambda$. Therefore, as long as $(\beta-1-\delta)/2\gamma^2\delta \ge c_0$, we have

$$\begin{aligned}P(f^* \ge \beta\lambda, T^*(f)\vee d^* < \delta\lambda,)&\\ =&P(\nu<\infty,\ \tau=\infty)\\ \le&P(|\sum_{k=\mu}^{\nu\wedge\tau} d_k| > (\beta-1-\delta)\lambda)\\ \le&P(|\sum_{k=\mu}^{\nu\wedge\tau} d_k| > \frac{\beta-1-\delta}{2\gamma^2\delta}\|T^*({}^\mu f^{\nu\wedge\tau})\|_\infty)\\ =&EP(|\sum_{k=\mu}^{\nu\wedge\tau} d_k| > \frac{\beta-1-\delta}{2\gamma^2\delta}\|T^*({}^\mu f^{\nu\wedge\tau})\|_\infty|\mathcal{F}_\mu).\end{aligned}$$

Note that, conditionally on $\mathcal{F}_\mu$, ${}^\mu f^{\nu\wedge\tau}$ is a mean-zero martingale. Therefore, it follows from our assumption (applied to the conditional measure $P(\,\cdot\,|\mathcal{F}_\mu)$) that the conditional probability above does not exceed $2\exp\left\{-\frac{\kappa(\beta-1-\delta)^2p}{(2\gamma^2\delta)^2}\right\}$, if $\mu<\infty$, and is zero if $\mu=\infty$. Consequently, the last expectation does not exceed

$$2E\exp\left\{-\frac{\kappa(\beta-1-\delta)^2p}{\delta^2}\right\}I(\mu<\infty) = 2\exp\left\{-\frac{\kappa(\beta-1-\delta)^2p}{\delta^2}\right\}P(f^*>\lambda).$$

This completes the proof of the first statement. The second statement is now a consequence of Lemma 2.1. ∎

We shall now define an operator on martingale transforms of Rademacher sequence. Assume for convenience that (r_n) and (r'_n) are defined on different probability spaces, $(\Omega,\mathcal{F},P)$ and $(\Omega',\mathcal{F}',P')$, respectively. The symbol $E'(\,\cdot\,)$ will be used to denote the integration with respect to the measure P'. Let $\mathcal{N}$ be a class of all martingale transforms of a sequence (r_n), i.e. $\mathcal{N}$ consists of all martingales $f=(f_n)$ of the form:

$$f_n = \sum_{k=1}^n v_k r_k, \quad n=1,\dots,$$

where (v_n) is a predictable sequence. For $1\le p<\infty$ we define an operator on $\mathcal{N}$ by the formula:

$$T_p(f) = (E'|\sum v_k r'_k|^p)^{1/p},$$

where $f_n = \sum_{k=1}^n v_k r_k$. Observe that T_p is a predictable operator satisfying the conditions (B1)–(B3) above, and that $\mathcal{N}$ is a subclass of the class of all martingales which is closed under starting and stopping (see Remark 3.2 (ii) in Hitczenko [12]). Also note that, by Fubini's theorem

$$\|T_p(f^n)\|_p = \Big\| \sum_{k=1}^{n} v_k r'_k \Big\|_p,$$

and by Levy's inequality we have $T_p^*(f^n) \le 2T_p(f^n)$. Thus, in view of Lemma 2.2, we need to establish an appropriate upper bound on the quantity

$$P(|f_n| \ge t\|T_p(f^n)\|_\infty), \quad t > 0.$$

In order to accomplish that goal we shall proceed in two steps. Firstly, we shall need an information on the size of the L_∞-norm of $T_p(f^n)$. To this end it suffices to have a precise approximation for the L_p-norm of a sum $\sum a_k r_k$, where (a_k) is a sequence of real numbers. Unfortunately, the hitherto known two-sided estimates for $\|\sum a_k r_k\|_p$ lack the required precision, since an upper and lower bounds usually involve constants depending on p, typically of different order of the magnitude. This difficulty will be overcome in the next section, where we shall establish a rather exact approximation for $\|\sum a_k r_k\|_p$. Once this is done, we shall easily obtain an appropriate exponential bound on the tail probabilities of $\sum v_k r_k$. This will be done in the fourth section of the present paper.

3. L_p-norms of Rademacher averages

In this section we shall establish a precise approximation for the expressions of the form:

$$\Big\| \sum_{k=1}^{n} a_k r_k \Big\|_p, \quad 1 \le p < \infty,$$

where (a_k) is a sequence of reals. Our approach will follow a method used by Montgomery-Smith [24]. Before we state the result we need to introduce some more notation. For a sequence $a = (a_k)$ of real numbers we let $|a|_p = (\sum |a_k|^p)^{1/p}$, $1 \le p < \infty$. Let $a \in \ell_2$ be a sequence of real numbers and consider the interpolation norm

$$K(a,t;\ell_1,\ell_2) = K_{1,2}(a,t) = \inf\{|a'|_1 + t|a''|_2 \colon a', a'' \in \ell_2, a' + a'' = a\}.$$

We shall also use the following norm on ℓ_2: For $t \in \mathrm{N}$ and $a \in \ell_2$ we let

$$\|a\|_{P(t)} = \sup \left\{ \sum_{m=1}^{t} \left(\sum_{k \in B_m} a_k^2 \right)^{1/2} \right\},$$

where the supremum is taken over all pairwise disjoint subsets, $B_1, \ldots, B_t$ of positive integers. Those two norms are related by the following inequality established in Montgomery-Smith [24]:

LEMMA 3.1: *For all $a \in \ell_2$ and t such that $t^2 \in \mathbf{N}$ we have*

$$\|a\|_{P(t^2)} \leq K_{1,2}(a,t) \leq \sqrt{2}\|a\|_{P(t^2)}.$$

We shall use this lemma to obtain the following two sided estimate on the L_p-norm of Rademacher averages:

THEOREM 3.2: *Let (r_k) be a Rademacher sequence. Then, there exists a constant $c > 0$ such that, for all p, $1 \leq p < \infty$, and all $a \in \ell_2$ the following inequality is true:*

$$cK_{1,2}(a, \sqrt{p}) \leq \left\| \sum a_k r_k \right\|_p \leq K_{1,2}(a, \sqrt{p}).$$

Proof: The right hand side inequality is a direct consequence of the Khintchine's inequality with the best constant (cf. Haagerup [9]). Given an $\epsilon > 0$ pick a' and a'' such that

$$|a'|_1 + \sqrt{p}|a''|_2 \leq K_{1,2}(a, \sqrt{p}) + \epsilon.$$

Then we have that

$$\begin{aligned} \left\| \sum a_k r_k \right\|_p \leq & \left\| \sum a'_k r_k \right\|_p + \left\| \sum a''_k r_k \right\|_p \\ \leq & \sum |a'_k| + \sqrt{p} \left(\sum (a''_k)^2 \right)^{1/2} \\ \leq & K_{1,2}(a, \sqrt{p}) + \epsilon, \end{aligned}$$

and the inequality follows by letting $\epsilon \to 0$.

To prove the left hand side inequality we first observe that without loss of generality p can be assumed to be a positive integer. This is a consequence of the monotonicity of $K_{1,2}(a,t)$ as a function of t and the following hypercontraction inequality for Rademacher functions due to Borell [3]:

$$\left\| \sum a_k r_k \right\|_p \leq \sqrt{\frac{p-1}{q-1}} \left\| \sum a_k r_k \right\|_q, \quad 1 < q < p < \infty.$$

Fix $0 < \delta < 1$ and take $\epsilon > 0$ such that $(1+\epsilon)\delta < 1$. By Lemma 3.1 there exist pairwise disjoint subsets $B_1, \ldots, B_p$ of $\mathbf{N}$ such that

$$\|a\|_{P(p)} \leq (1+\epsilon) \left\{ \sum_{m=1}^{p} \left(\sum_{k \in B_m} a_k^2 \right)^{1/2} \right\}.$$

Then

$$\Big\|\sum a_k r_k\Big\|_p^p = \Big\|\sum_{m=1}^{p}\sum_{k\in B_m} a_k r_k\Big\|_p^p$$

$$\geq \left(\sum_{m=1}^{p}(1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right)^p$$

$$\times P\left(\bigcap_{m=1}^{p}\left\{\sum_{k\in B_m} a_k r_k \geq (1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right\}\right)$$

$$= \left(\sum_{m=1}^{p}(1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right)^p$$

$$\times \prod_{m=1}^{p} P\left(\sum_{k\in B_m} a_k r_k \geq (1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right).$$

By the Paley-Zygmund inequality (cf. e.g. Kahane [15, p. 24]) we infer that

$$P\left(\sum_{k\in B_m} a_k r_k \geq (1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right) \geq \frac{1}{6}\left(1-\delta^2(1+\epsilon)^2\right)^2.$$

Therefore,

$$\Big\|\sum a_k r_k\Big\|_p^p \geq \left(\frac{1}{6}\left(1-\delta^2(1+\epsilon)^2\right)^2\right)^p \left(\sum_{m=1}^{p}(1+\epsilon)\delta\left(\sum_{k\in B_m} a_k^2\right)^{1/2}\right)^p,$$

so that

$$\Big\|\sum a_k r_k\Big\|_p \geq \frac{1}{6}\left(1-\delta^2(1+\epsilon)^2\right)^2(1+\epsilon)\delta\sum_{m=1}^{p}\left(\sum_{k\in B_m} a_k^2\right)^{1/2}$$

$$\geq \frac{\delta}{6\sqrt{2}}(1-\delta^2(1+\epsilon)^2)^2 K_{1,2}(a,\sqrt{p}).$$

The left hand side inequality follows now by letting $\epsilon \to 0$. This completes the proof. ∎

4. Weak type estimate for T_p

In this section we shall prove the following exponential inequality:

THEOREM 4.1: *Let f be a mean-zero martingale, where $f_n = \sum_{k=1}^n v_k r_k$, $n = 1, \ldots$, is a transform of a Rademacher sequence. Then, there exist constants $c_0 > 0$ and $\kappa > 0$ such that for every $c > c_0$ and for every positive number t the following inequality is true:*

$$P(|f_n| \geq c\|K_{1,2}(v,t)\|_\infty) \leq 2\exp\left\{-\kappa c^2 t^2\right\}.$$

Before we pass to the proof let us observe, that Theorems 3.1 and 4.1 immediately yield the following

COROLLARY 4.2: *Let $1 \leq p < \infty$. Under the assumptions of Theorem 4.1, there exist constants $c_0 > 0$ and $\kappa > 0$ such that for every $c > c_0$ we have that*

$$P(|f_n| \geq c\|T_p(f^n)\|_\infty) \leq 2\exp\left\{-\kappa c^2 p\right\}, \quad n = 1, \ldots.$$

Proof of Theorem 4.1: According to Holmstedt's formula (cf. Holmstedt [14, Theorem 4.1] or Bergh and Löfström [2, Theorem 3.6.1]), for every sequence $a \in \ell_2$ and all $t > 0$, the interpolation norm $K_{1,2}(a,t)$ is equivalent (with constant not depending on a or t) to the quantity

$$\sum_{k=1}^{[t^2]} a^{(k)} + t\left(\sum_{k>[t^2]} (a^{(k)})^2\right)^{1/2}.$$

Here, $a^{(k)}$ denotes the kth largest element of the sequence $(|a_j|)$. We shall also write $a_n^{(k)}$ for the kth largest element of the sequence $(|a_1|, \ldots, |a_n|)$, and we adopt the convention that $a_n^{(k)} = 0$ if $k > n$. Let $c_{1,2} > 0$ be any constant for which the inequality

$$K_{1,2}(a,t) \geq c_{1,2}\left\{\sum_{k=1}^{[t^2]} a^{(k)} + t\left(\sum_{k>[t^2]} (a^{(k)})^2\right)^{1/2}\right\}$$

is true for all sequences $a \in \ell_2$ and all $t > 0$. Fix a positive real number t and put $s = [t^2]$. We shall show that the conclusion of our theorem holds with $c_0 = 2/c_{1,2}$ and $\kappa = c_{1,2}^2/8$. Given a predictable sequence of random variables (v_n) let us write

$$v_n = u_n + w_n,$$

where

$$u_n = \operatorname{sgn}(v_n)(|v_n| - v_{n-1}^{(s)})^+ = \operatorname{sgn}(v_n)(|v_n| - v_{n-1}^{(s)})I(|v_n| > v_{n-1}^{(s)}),$$

and

$$w_n = \operatorname{sgn}(v_n)\left(|v_n| \wedge v_{n-1}^{(s)}\right).$$

Since (v_n) is predictable, so are (u_n) and (w_n) and we have that

$$\sum_{k=1}^{\infty} |u_k| = \sum_{k=1}^{s} v^{(k)}$$

and

$$\sum_{k=1}^{\infty} w_k^2 \le \sum_{k>s} (v^{(k)})^2$$

(the latter follows from the fact that $w^{(r)} \le v^{(r+s)}$, $r = 1, \ldots$). Since

$$\begin{aligned}
\|\sum_{k=1}^{s} v^{(k)} + t(\sum_{k>s}(v^{(k)})^2)^{1/2}\|_\infty \ge& \max\left\{\|\sum_{k=1}^{s} v^{(k)}\|_\infty,\ t\|(\sum_{k>s}(v^{(k)})^2)^{1/2}\|_\infty\right\} \\
\ge& \frac{1}{2}\left\{\|\sum_{k=1}^{s} v^{(k)}\|_\infty + t\|(\sum_{k>s}(v^{(k)})^2)^{1/2}\|_\infty\right\} \\
\ge& \frac{1}{2}\left\{\|\sum u_k\|_\infty + t\|(\sum w_k^2)^{1/2}\|_\infty\right\}
\end{aligned}$$

we get that

$$\begin{aligned}
P(|\sum v_k r_k| \ge& c\|K_{1,2}(v,t)\|_\infty) \\
\le& P(|\sum v_k r_k| \ge c c_{1,2}\|\sum_{k=1}^{s} v^{(k)} + t(\sum_{k>s}(v^{(k)})^2)^{1/2}\|_\infty) \\
\le& P(|\sum u_k r_k| \ge \frac{c c_{1,2}}{2}\|\sum u_k\|_\infty) \\
&+ P(|\sum w_k r_k| \ge \frac{c c_{1,2}}{2} t\|(\sum w_k^2)^{1/2}\|_\infty) \\
\le& 0 + 2\exp\left\{-\frac{c^2 c_{1,2}^2 t^2}{8}\right\},
\end{aligned}$$

where the last inequality follows by the subgaussian inequality for conditionally symmetric martingales (cf. e.g. Chang, Wilson and Wolff [7, Theorem 3.1] or Hitczenko [12, Lemma 4.3]), and the fact that $cc_{1,2}/2 > 1$. This completes the proof. ∎

5. Domination inequality

The inequalities proved in the preceding two sections immediately imply the following result:

THEOREM 5.1: *There exists an absolute constant K such that, for every p, $1 \leq p < \infty$, and every martingale f with difference sequence of the form: $d_n = v_n r_n$, $n = 1, \ldots$ we have*

$$\|f_n^*\|_p \leq K \Big\| \sum_{k=1}^{n} v_k r_k' \Big\|_p, \quad n = 1, \ldots .$$

Remark: The above theorem remains true if Rademacher sequence is replaced by a sequence (γ_k) of i.i.d. standard Gaussian variables. This follows from the fact that

$$(E' |\sum v_k \gamma_k'|^p)^{1/p} = \|\gamma_1\|_p (\sum v_k^2)^{1/2} \approx \sqrt{p} (\sum v_k^2)^{1/2}$$

and from the estimate

$$P(|\sum v_k \gamma_k| \geq t) \leq 2 \exp \left\{ \frac{-t^2}{2 \|\sum v_k^2\|_\infty} \right\}. \quad \blacksquare$$

Theorem 5.1 can be easily extended to martingale transforms of sequences satisfying some regularity conditions (although with constant K depending on "regularity constant" of that sequence). Following Gundy [8], we say that a sequence (ξ_k) is L^∞-regular if ξ_k's are uniformly bounded and there exists a $\delta > 0$ such that, for each $k = 1, \ldots$,

$$E|\xi_k| \geq \delta \|\xi_k\|_\infty.$$

With the above definition we have:

THEOREM 5.2: *Let (ξ_n) be L^∞-regular sequence of independent mean-zero random variables. Then, there exists a constant K depending only on δ such that, for every predictable sequence (v_k) we have that*

$$\Big\| \sum v_k \xi_k \Big\|_p \leq K \Big\| \sum v_k \xi_k' \Big\|_p, \quad 1 \leq p < \infty.$$

Proof: Let us observe first, that if ξ and ξ' are i.i.d. mean-zero random variables such that $\|\xi\|_1 \geq \delta \|\xi\|_\infty$ then

$$\|\xi - \xi'\|_1 \geq \|\xi\|_1 \geq \delta \|\xi\|_\infty \geq \frac{\delta}{2} \|\xi - \xi'\|_\infty.$$

Moreover

$$\|\sum v_k\xi_k\|_p \le \|\sum v_k\xi_k'\|_p + \|\sum v_k(\xi_k - \xi_k')\|_p.$$

Therefore, we can and do assume in the proof that ξ_k's are symmetric. We shall now use an observation due to Wang [25, the bottom half of p. 399] (cf. also Hitczenko [12, proof of Lemma 4.3]). Define a sequence of σ-algebras $(\mathcal{A}_n)$ by letting $\mathcal{A}_n = \sigma(\mathcal{F}_n, |\xi_{n+1}|)$. It follows from the symmetry that, relative to $(\mathcal{A}_n)$, $(v_n\xi_n)$ is a martingale difference sequence equidistributed with $(u_n\epsilon_n)$, where (ϵ_n) is a Rademacher sequence and $(u_n) = (v_n|\xi_n|)$ is predictable sequence. Let (ϵ_n') be a Rademacher sequence independent of all other random variables under consideration. We shall write

$$T_p^{\xi}(f) = \left(E'|\sum v_k\xi_k'|^p\right)^{1/p}$$

and

$$T_p^{\epsilon}(f) = \left(E'|\sum v_k\epsilon_k'|^p\right)^{1/p}.$$

With this notation, denoting by E_ϵ' and E_ξ' integration with respect to (ϵ_n') and (ξ_n'), respectively we have

$$\begin{aligned} T_p^{\xi}(f) &= \left(E_\xi'|\sum v_k\xi_k'|^p\right)^{1/p} = \left(E_\epsilon' E_\xi'|\sum v_k|\xi_k'|\epsilon_k'|^p\right)^{1/p} \\ &\ge \left(E_\epsilon'|\sum v_k\|\xi_k\|_1\epsilon_k'|^p\right)^{1/p} \ge \delta\left(E_\epsilon'|\sum v_k\|\xi_k\|_\infty\epsilon_k'|^p\right)^{1/p} \\ &\ge \delta\left(E_\epsilon'|\sum v_k|\xi_k|\epsilon_k'|^p\right)^{1/p} = \delta T_p^{\epsilon}(f), \end{aligned}$$

where the last inequality follows by the contraction principle (cf. e.g. Ledoux and Talagrand [22, Theorem 4.4]). Therefore,

$$\|T_p^{\xi}(f^n)\|_\infty \ge \delta\|T_p^{\epsilon}(f^n)\|_\infty,$$

so that, for each $c > 0$ we have

$$P(f^* \ge c\|T_p^{\xi}(f)\|_\infty) \le P(f^* \ge c\delta\|T_p^{\epsilon}(f)\|_\infty).$$

The result now follows from Lemma 2.2 and Corollary 4.2. ∎

6. An example

It is natural to ask whether the results of the preceding section (Theorem 5.1 in particular) can be strengthened to domination of tail probabilities. The answer to this question turns out to be negative. We would like to conclude by presenting an example, which was communicated to us by M. Talagrand and which is included here with his kind permission.

THEOREM 6.1: *There is no constant C which makes the inequality:*

$$P(\left|\sum v_j r_j\right| \geq Ct) \leq CP(\left|\sum v_j r'_j\right| \geq t)$$

true.

Proof: Given an integer k let N_1 be an integer which will be specified later and let $N_2, \ldots N_k$ be defined by:

$$N_i - N_{i-1} = 2^{-(i-1)} N_1, \quad i = 2, \ldots, k.$$

Put

$$\Omega_1 = \{r_1 = \cdots = r_{N_1} = 1\},$$

and then

$$\Omega_i = \Omega_{i-1} \cap \{r_{N_{i-1}+1} = \cdots = r_{N_i}\} \quad i = 2, \ldots, k.$$

Define a predictable sequence of random variables (v_i) by the formulas:

$$v_1 = \cdots = v_{N_1} = 1$$
$$v_{N_1+1} = \cdots = v_{N_2} = 2I_{\Omega_1}$$
$$\cdots\cdots\cdots\cdots$$
$$v_{N_{k-1}+1} = \cdots = v_{N_k} = 2^{k-1} I_{\Omega_{k-1}}.$$

Then for $t = N_1$ we have:

$$P(\left|\sum_{j=1}^{N_k} v_j r_j\right| \geq kt) = 2P(r_j = 1,\ j = 1, \ldots, N_k) = 2 \cdot 2^{-N_1(1+2^{-1}+\cdots+2^{-(k-1)})}.$$

We shall now estimate

$$P\Big(\left|\sum_{j=1}^{N_k} v_j r'_j\right| \geq 4t\Big)$$

from above.

$$
\begin{aligned}
P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t\Big) &= P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t, \Omega_1^c\Big) \\
&\quad + P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t, \Omega_1 \setminus \Omega_2\Big) + \cdots \\
&\quad + P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t, \Omega_{k-1} \setminus \Omega_k\Big) \\
&\quad + P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t, \Omega_k\Big) \\
= P\Big(\Big|\sum_{j=1}^{N_1} r'_j\Big| \geq 4t\Big)(1 - P(\Omega_1)) & \\
&\quad + P\Big(\Big|\sum_{j=1}^{N_1} r'_j + 2 \cdot \sum_{j=N_1+1}^{N_2} r'_j\Big| \geq 4t\Big) P(\Omega_1 \setminus \Omega_2) + \cdots \\
&\quad + P\Big(\Big|\sum_{j=1}^{N_1} r'_j + \cdots + 2^{k-1} \cdot \sum_{j=N_{k-1}+1}^{N_k} r'_j\Big| \geq 4t\Big) P(\Omega_{k-1}).
\end{aligned}
$$

Since $4t = 4N_1 > N_1 + 2(N_2 - N_1) + 4(N_3 - N_2)$, the first three terms in the last sum are zero, and the ith term can be estimated as follows:

$$P(\Omega_{i-1} \setminus \Omega_i) \leq P(\Omega_{i-1}) = 2^{-N_{i-1}} = 2^{-N_1(1+\cdots+2^{-(i-2)})}.$$

By subgaussian inequality for Rademacher functions, one has:

$$
\begin{aligned}
P\Big(\Big|\sum_{j=1}^{N_1} r'_j + \cdots + 2^{i-1} \cdot \sum_{j=N_{i-1}+1}^{N_i} r'_j\Big| \geq 4t\Big) & \\
\leq 2\exp\left\{ \frac{-16t^2}{2E(\sum_{j=1}^{N_1} r'_j + \cdots + 2^{i-1} \sum_{j=N_{i-1}+1}^{N_i} r'_j)^2} \right\}.
\end{aligned}
$$

But

$$E\Big(\sum_{j=1}^{N_1} r'_j + \cdots + 2^{i-1} \cdot \sum_{j=N_{i-1}+1}^{N_i} r'_j\Big)^2 = N_1(1 + \cdots + 2^{i-1}) \leq 2^i N_1,$$

so that the exponent above is less or equal than

$$\exp\{-8N_1/2^i\} \leq 2^{-8N_1/2^i}.$$

Combining the above estimates we obtain that:

$$\begin{aligned}
P\Big(\Big|\sum_{j=1}^{N_1} r'_j + \cdots + 2^{i-1}\cdot \sum_{j=N_{i-1}+1}^{N_i} r'_j\Big| \geq 4t\Big)P(\Omega_{i-1})
&\leq 2\cdot 2^{-8N_1/2^i}\cdot 2^{-N_1(1+\cdots+2^{-(i-2)})}\\
&= 2\cdot 2^{-N_1(1+\cdots+2^{-(k-1)})}\cdot 2^{-8N_1/2^i}\cdot 2^{N_1(2^{-(i-1)}+\cdots+2^{-(k-1)})}\\
&\leq P\Big(\Big|\sum_{j=1}^{N_k} v_j r_j\Big| \geq kt\Big)\cdot 2^{-8N_1/2^i}\cdot 2^{N_1/2^{i-2}}\\
&= 2^{-N_1/2^{i-2}} P\Big(\Big|\sum_{j=1}^{N_k} v_j r_j\Big| \geq kt\Big).
\end{aligned}$$

Therefore, summing over i we get that

$$\begin{aligned}
P\Big(\Big|\sum_{j=1}^{N_k} v_j r'_j\Big| \geq 4t\Big) &\leq \sum_{i=4}^{k} 2^{-N_1/2^{i-2}} P(\Big|\sum_{j=1}^{N_k} v_j r_j\Big| \geq kt)\\
&\leq k\cdot 2^{-N_1/2^{k-2}} P\Big(\Big|\sum_{j=1}^{N_k} v_j r_j\Big| \geq kt\Big).
\end{aligned}$$

This completes the proof since N_1 can be made arbitrarily large. ∎

Remark: It is probably worth noticing that the above example shows a little bit more, namely that v_k's may be chosen to be of the form $v_k = a_k I(\tau \geq k)$, $k = 1, \ldots$, where (a_k) is a nonrandom sequence and τ is a stopping time; one just puts

$$a_j = 2^{i-1} \quad \text{for} \quad N_{i-1} < j \leq N_i,\ i = 1, \ldots, k;$$

and $\tau = \inf\{n\colon r_n \neq r_{n-1}\}$. This shows that moment inequality proved by Klass [16, Theorem 3.1] for randomly stopped sums of independent random variables can not be strengthened to distributional statement, even for linear combinations of Rademacher functions. ∎

ACKNOWLEDGEMENT: It is a pleasure to thank Michel Talagrand who generously permitted me to include his example (Theorem 6.1) in this paper. I would also like to thank the referee for pointing out several omissions.

References

[1] R. Bañuelos, *A sharp good-λ inequality with an application to Riesz transforms,* Michigan Math. J. **35** (1988), 117–125.

[2] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction,* Springer, Berlin, Heidelberg, 1976.

[3] C. Borell, *On the integrability of Banach valued Walsh polynomials,* Séminaire de Probabilités XIII, Lecture Notes in Mathematics **721**, Springer, Berlin, Heidelberg, 1979, pp. 1–3.

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

[5] D. L. Burkholder, B. Davis and R. F. Gundy, *Integral inequalities for convex functions of operators on martingales,* in *Proceedings of the 6th Berkeley Symposium on Mathematical Statistics and Probability,* Vol. 2, University of California Press, Berkeley, 1972, pp. 223–240.

[6] D. L. Burkholder and R. F. Gundy, *Extrapolation and interpolation of quasilinear operators on martingales,* Acta Math. **124** (1970), 250–304.

[7] S. Chang, M. Wilson and J. Wolff, *Some weighted norm inequalities concerning the Schrödinger operators,* Comment. Math. Helv. **60** (1985), 217–246.

[8] R. F. Gundy, *On the class of $L \log L$ martingales, and singular integrals,* Studia Math. **33** (1969), 109–118.

[9] U. Haagerup, *Best constants in the Khintchine's inequality,* Studia Math. **70** (1982), 231–283.

[10] P. Hitczenko, *Comparison of moments for tangent sequences of random variables,* Probab. Theory Related Fields **78** (1988), 223–230.

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

[12] P. Hitczenko, *Upper bounds for the L_p-norms of martingales,* Probab. Theory Related Fields. **86** (1990), 225–238.

[13] P. Hitczenko, *On the behavior of the constant in a decoupling inequality for martingales,* Proc. Amer. Math. Soc., to appear.

[14] T. Holmstedt, *Interpolation of quasi-normed spaces,* Math. Scand. **26** (1970), 177–199.

[15] J-P. Kahane, *Some Random Series of Functions,* Heath Mathematical Monographs, Lexington, Massachusetts, 1968.

[16] M. J. Klass, *A best possible improvement of Wald's equation,* Ann. Probab. **16** (1988), 840–853.

[17] M. J. Klass, *Uniform lower bounds for randomly stopped sums of Banach space valued random sums,* Ann. Probab. **18** (1990), 790–809.

[18] S. Kwapień, *Decoupling inequalities for polynomial chaos,* Ann. Probab. **15** (1987), 1062–1071.

[19] S. Kwapień and W. A. Woyczyński, *Semimartingale integrals via decoupling inequalities and tangent processes,* Case Western Reserve University Preprint, 86–56, 1986.

[20] S. Kwapień and W. A. Woyczyński, *Tangent sequences of random variables: basic inequalities and their applications,* Proc. Conf. on Almost Everywhere Convergence in Probability and Ergodic Theory Columbus, Ohio, June 11–14, 1988 (G. A. Edgar and L. Sucheston, eds.), Academic Press, New York, 1989, pp. 237–265.

[21] S. Kwapień and W. A. Woyczyński, *Random Series and Stochastic Integrals. Single and Multiple,* Birkhäuser, Boston, 1992.

[22] M. Ledoux and M. Talagrand, *Probability in Banach Spaces,* Springer, Berlin, Heidelberg, 1991.

[23] T. R. McConnell, *Decoupling and integration in UMD Banach space,* Probab. Math. Statist. **10** (1990), 283–295.

[24] S. J. Montgomery-Smith, *The distribution of Rademacher sums,* Proc. Amer. Math. Soc. **109** (1990), 517–522.

[25] G. Wang, *Sharp square function inequalities for conditionally symmetric martingales,* Trans. Amer. Math. Soc. **328** (1991), 393–420.

[26] J. Zinn, *Comparison of martingale differences,* in *Probability in Banach Spaces V* (A. Beck, R. Dudley, M. Hahn, J. Kuelbs and M. Marcus, eds.), Lecture Notes in Mathematics, **1153**, Springer, Berlin, Heidelberg, 1985, pp. 453–457.