The Annals of Probability
1985, Vol. 13, No. 1, 234–253

BEST CONSTANTS IN MOMENT INEQUALITIES FOR LINEAR COMBINATIONS OF INDEPENDENT AND EXCHANGEABLE RANDOM VARIABLES

By W. B. Johnson[1, 3]
The Ohio State University and Texas A&M University

G. Schechtman[1, 3]
The Weizmann Institute of Science, The Ohio State University and Texas A&M University

J. Zinn[2]
Texas A&M University

In [4] Rosenthal proved the following generalization of Khintchine's inequality:

$$\text{(B)}\quad \begin{cases} \max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\} \\ \qquad\qquad \le \|\sum_{i=1}^n X_i\|_p \le B_p \max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=i}^n \|X_i\|_p^p)^{1/p}\} \\ \text{for all independent symmetric random variables } X_1, X_2, \cdots, \text{ with} \\ \text{finite } p\text{th moment, } 2 < p < \infty. \end{cases}$$

Rosenthal's proof of (B) as well as later proofs of more general results by Burkholder [1] yielded only exponential of p estimates for the growth rate of B_p as $p \to \infty$. The main result of this paper is that the actual growth rate of B_p as $p \to \infty$ is $p/\operatorname{Log} p$, as compared with a growth rate of $\sqrt{p}$ in Khintchine's inequality.

1. Introduction. In what follows the expression $\|X\|_s$ for a random variable X and for $0 < s < \infty$ always denotes $(E|X|^s)^{1/s}$. $\|X\|_\infty$ denotes ess sup $|X|$.

In [4] Rosenthal proved the following fundamental generalization of Khintchine's inequality:

$$\text{(B)}\quad \begin{cases} \max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\} \\ \qquad\qquad \le \|\sum_{i=1}^n X_i\|_p \le B_p \max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\} \\ \text{for all independent symmetric random variables } X_1, X_2, \cdots, \text{ with finite} \\ p\text{th moment, } 2 < p < \infty. \end{cases}$$

Rosenthal's proof of (B) as well as later proofs of more general results by Burkholder [1] yielded only exponential of p estimates for the growth rate of B_p as $p \to \infty$. The main result of this paper is that the actual growth rate of B_p as

Received May 1983; revised June 1984.

[1] Supported in part by NSF MCS-79-03042.

[2] Supported in part by NSF MCS-82-13743.

[3] Part of the research of these authors was done while they were in residence at the Workshop on Banach Space Theory at the University of Iowa in July, 1981.

AMS 1980 *subject classification.* 60E15, 60G50, 60G42.

Key words and phrases. Rosenthal X_p-inequality, exchangeable.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.* www.jstor.org

$p \to \infty$ is $p/\text{Log}\, p$. This is perhaps a little surprising since the growth rate in Khintchine's inequality is $\sqrt{p}$ (see [2] for the best constants).

We also show that the constants A_p, C_p, and D_p in the following inequalities grow like $p/\text{Log}\, p$ as $p \to \infty$ ((A) is contained in [4] while (C) and (D) were proved in Chapter 1 of [3]).

$$\text{(A)}\begin{cases} \max\{\|\sum_{i=1}^n X_i\|_1, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\} \\ \qquad \le \|\sum_{i=1}^n X_i\|_p \le A_p \max\{\|\sum_{i=1}^n X_i\|_1, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\} \\ \text{for all independent nonnegative random variables } X_1, X_2, \cdots, X_n \text{ with} \\ \text{finite } p\text{th moment, } 1 \le p < \infty. \end{cases}$$

$$\text{(C)}\begin{cases} \max\{Cn^{-1/2}\sum_{i=1}^n a_i, (\sum_{i=1}^n a_i^p)^{1/p}\} \le \|\sum_{i=1}^n a_i Y_i\|_p \\ \qquad \le C_p \max\{Cn^{-1/2}\sum_{i=1}^n a_i, (\sum_{i=1}^n a_i^p)^{1/p}\} \\ \qquad\qquad (\text{where } C \equiv \|\sum_{i=1}^n Y_i\|_p) \\ \text{for all nonnegative exchangeable random variables } Y_1, Y_2, \cdots, Y_n \text{ with} \\ EY_i^p = 1, 1 \le p < \infty \text{ and all nonnegative scalars } a_1, a_2, \cdots, a_n. \end{cases}$$

$$\text{(D)}\begin{cases} \max\{Cn^{-1/2}(\sum_{i=1}^n |a_i|^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\} \\ \qquad \le \|\sum_{i=1}^n a_i Y_i\|_p \\ \qquad \le D_p \max\{Cn^{-1/2}(\sum_{i=1}^n |a_i|^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\} \\ \qquad\qquad (\text{where } C \equiv \|(\sum_{i=1}^n |Y_i|^2)^{1/2}\|_p) \\ \text{for all symmetrically exchangeable random variables } Y_1, Y_2, \cdots, Y_n \\ \text{with } EY_i^p = 1;\ 2 \le p < \infty \text{ and all scalars } a_1, a_2, \cdots, a_n. \end{cases}$$

We recall that $(Y_1, \cdots, Y_n)$ is *exchangeable* (respectively, *symmetrically exchangeable*) provided $(Y_{\pi(1)}, \cdots, Y_{\pi(n)})$ (respectively, $(\varepsilon_1 Y_{\pi(1)}, \cdots, \varepsilon_n Y_{\pi(n)})$) has the same joint distribution as $(Y_1, \cdots, Y_n)$ for every permutation π of $\{1, \cdots, n\}$ (and every choice $\varepsilon_i = \pm 1$).

To prove (A) and (C) with the best order of A_p and C_p we incorporate an elementary unbalanced scalar inequality (Lemma 2.1) into the arguments in [3] and [4]. The further arguments of [3] and [4] then yields that

$$B_p, D_p \le Kp/\sqrt{\text{Log}\, p}$$

for some absolute constant K. To check that B_p and D_p are of order $p/\text{Log}\, p$ we apply (A) and (C) to the "peaky parts" or large values of the individual random variables and use an exponential integrability argument on the truncated variables. The main technical tool is Prokhorov's "arcsinh" inequality ([5], Theorem 5.22(ii)) (to prove (D) we need to generalize this inequality to the appropriate martingale setting). We also use a formalization of the relationship, already known in special cases, between an exponential norm of a random variable and the absolute pth moments of the variable.

Since the discussed inequalities may have some applications in statistical work, we have tried to obtain estimates on numerical constants. For example, we check in Theorem 2.5 and Proposition 2.9 (see also Proposition 4.3) that for all $1 \le p < \infty$

$$p/e \operatorname{Log} p \le A_p \le 2p/\operatorname{Log} p$$

and in Corollary 2.6 that for all $2 \le p < \infty$,

$$B_p \le p/\sqrt{\operatorname{Log} p}.$$

This last inequality is easier to obtain and better, for real-life values of p, than the estimate

$$B_p \le 7.35p/\operatorname{Log} p$$

obtained in Theorem 4.1.

The paper is arranged as follows: Section 2 contains proofs of the easier inequalities (A) and (C). Section 3 contains preliminary material needed for Section 4. In particular, we prove some relations between pth moments and exponential moments of random variables as well as a martingale version of Prokhorov's "arcsinh" inequality. Section 4 is devoted to inequalities (B) and (D). It ends with a fifth inequality of a similar type.

Section 2.

LEMMA 2.1. *Let a, b, c, $s > 0$. Then*

$$(a+b)^s \le \max\{(1+1/cs)^s a^s, (cs+1)^s b^s\} \le \max\{e^{1/c} a^s, (cs+1)^s b^s\}.$$

PROOF. Assume, without loss of generality, that $b = 1$. The inequalities are obvious if $a \le cs$; otherwise

$$(a+1)^s a^{-s} \le (1+1/cs)^s \le e^{1/c}. \quad \square$$

REMARK. The choice $c = s^{-1}$ in the first inequality gives the commonly used inequality

$$(a+b)^s \le 2^s \max\{a^s, b^s\}.$$

Here and through the whole paper we use the following:

CONVENTION. $\operatorname{Log} t \equiv \max\{1, \ln t\}$, where "ln" is the natural logarithm function.

PROPOSITION 2.2. *Let $X_1, \cdots, X_n$ be nonnegative random variables with finite rth moment for some $1 \le r < \infty$. Then*

$$E\left(\sum_{i=1}^n X_i\right)^r \le \max\left\{K \frac{r}{\operatorname{Log} r} \sum_{j=1}^n E\left(\sum_{i=1; i\ne j}^n X_i\right)^{r-1} X_j, \left(K \frac{r}{\operatorname{Log} r}\right)^r \sum_{j=1}^n EX_j^r\right\}.$$

where K is an absolute constant. (In fact, $K \le 2$.)

PROOF. Applying Lemma 2.1 with $s = r - 1$, we obtain for every $c > 0$:

$$E(\sum_{i=1}^n X_i)^r = \sum_{j=1}^n E(\sum_{i=1}^n X_i)^{r-1}X_j$$
$$\le e^{1/c} \sum_{j=1}^n E(\sum_{i=1;i\ne j}^n X_i)^{r-1}X_j + [c(r-1)+1]^{r-1} \sum_{j=1}^n EX_j^r$$
$$\le \max\{2e^{1/c} \sum_{j=1}^n E(\sum_{i=1;i\ne j}^n X_i)^{r-1}X_j,\ 2[c(r-1)+1]^{r-1} \sum_{j=1}^n EX_j^r\}.$$

The conclusion follows by setting

$$e^{1/c} = \begin{cases} r/\text{Log } r, & \text{if } r \ge 2 \\ 2, & \text{if } 1 \le r \le 2. \end{cases} \quad \square$$

That the above choice for c yields that $K \le 2$ in the range $r \ge 2$ is an obvious consequence of the following (the proof of which is left to the reader):

SUBLEMMA 2.3. *For all* $r \ge 2$,

$$\frac{r-1}{\ln[r/\text{Log } r]} + 1 \le \frac{2r}{\text{Log } r}.$$

To check that $K \le 2$ in Proposition 2.2 where $r \le 2$, we use the inequality $(a+b)^s \le (2a)^s + (2b)^s$ in the argument for Proposition 2.2. (This is just Rosenthal's original argument [4].) This gives

$$E(\sum_{i=1}^n X_i)^r \le \max\{2^r \sum_{j=1}^n E(\sum_{i=1;i\ne j}^n X_i)^{r-1}X_j,\ 2^r \sum_{j=1}^n EX_j^r\}.$$

Since $2^r \le 2r = 2(r/\text{Log } r)$ for $1 \le r \le 2$, we get that $K \le 2$ in the range $1 \le r \le 2$.

THEOREM 2.5. *If* $X_1, \cdots, X_n$ *are nonnegative, independent random variables with finite* r*th moments* $(1 \le r < \infty)$, *then*

$$\max\{\| \sum_{i=1}^n X_i \|_1, (\sum_{i=1}^n \| X_i \|_r^r)^{1/r}\}$$
$$\le \| \sum_{i=1}^n X_i \|_r \le K\,(r/\text{Log } r)\max\{\| \sum_{i=1}^n X_i \|_1, (\sum_{i=1}^n \| X_i \|_r^r)^{1/r}\}$$

where $K \le 2$ *is an absolute constant.*

PROOF. The left inequality is evident, so we prove the right one. Since the X_i's are independent, we have

$$\sum_{j=1}^n E(\sum_{i=1;i\ne j}^n X_i)^{r-1}X_j$$
$$= \sum_{j=1}^n E(\sum_{i=1;i\ne j}^n X_i)^{r-1}(EX_j) \le E(\sum_{i=1}^n X_i)^{r-1}(\sum_{j=1}^n EX_j).$$

So from Proposition 2.2 we have

$$E(\sum_{i=1}^n X_i)^r \le \max\{K(r/\text{Log } r)E(\sum_{i=1}^n X_i)^{r-1} \sum_{i=1}^n EX_i,\ (K(r/\text{Log } r))^r \sum_{i=1}^n EX_i^r\}.$$

If this maximum in the preceding expression occurs in the second term we get

$$\| \sum_{i=1}^n X_i \|_r \le K(r/\text{Log } r)(\sum_{i=1}^n EX_i^r)^{1/r}.$$

Otherwise we have

$$E(\sum_{i=1}^n X_i)^r \le K(r/\mathrm{Log}\, r)E(\sum_{i=1}^n X_i)^{r-1} \sum_{i=1}^n EX_i$$
$$\le K(r/\mathrm{Log}\, r)[E(\sum_{i=1}^n X_i)^r]^{(r-1)/r} \sum_{i=1}^n EX_i$$

so that

$$[E(\sum_{i=1}^n X_i)^r]^{1/r} \le K(r/\mathrm{Log}\, r) \sum_{i=1}^n EX_i. \quad \square$$

The following corollary is a weaker form of Theorem 4.1 below.

COROLLARY 2.6. *Suppose that $X_1, \cdots, X_n$ are independent, symmetric random variables with finite pth moments* $(2 \le p < \infty)$. *Then*

$$\max\{\| \sum_{i=1}^n X_i \|_2, (\sum_{i=1}^n \| X_i \|_p^p)^{1/p}\}$$
$$\le \| \sum_{i=1}^n X_i \|_p \le (p/(\mathrm{Log}\, p)^{1/2})\max\{\| \sum_{i=1}^n X_i \|_2, (\sum_{i=1}^n \| X_i \|_p^p)^{1/p}\}.$$

PROOF. The left inequality is known and easy; cf. e.g. [4]. The right one follows just as in [4]: For any choice $\varepsilon_i = \pm 1$ we have by the symmetry and independence of the X_i's that

$$\| \sum_{i=1}^n X_i \|_p^p = \| \sum_{i=1}^n \varepsilon_i X_i \|_p^p.$$

Averaging over all choices $\varepsilon_i = \pm 1$ and using Khintchine's inequality in L_p (with constant $\| N(0, 1) \|_p \le (p/2)^{1/2}$; cf. [2]) we obtain

$$\| \sum_{i=1}^n X_i \|_p \le (p/2)^{1/2} \| (\sum_{i=1}^n X_i^2)^{1/2} \|_p = (p/2)^{1/2} \| \sum_{i=1}^n X_i^2 \|_{p/2}^{1/2}.$$

Since $p/2 \ge 1$, we get from Theorem 2.5

$$\| \sum_{i=1}^n X_i^2 \|_{p/2} \le 2 \frac{p/2}{\mathrm{Log}(p/2)} \max\{\| \sum_{i=1}^n X_i^2 \|_1, [\sum_{i=1}^n \| X_i^2 \|_{p/2}^{p/2}]^{2/p}\}$$
$$\le \frac{2p}{\mathrm{Log}\, p} \max\{\| \sum_{i=1}^n X_i \|_2^2, (\sum_{i=1}^n \| X_i \|_p^p)^{2/p}\}.$$

Thus

$$\| \sum_{i=1}^n X_i \|_p \le (p/2)^{1/2}(2p/\mathrm{Log}\, p)^{1/2}\max\{\| \sum_{i=1}^n X_i \|_2, (\sum_{i=1}^n \| X_i \|_p^p)^{1/p}\}. \quad \square$$

THEOREM 2.7. *Let $Y_1, \cdots, Y_n$ be a sequence of nonnegative exchangeable random variables with $\| Y_i \|_r = 1$ $(1 \le r < \infty)$ and set $C = \| \sum_{i=1}^n Y_i \|_r$. Then for all nonnegative scalars $\{a_i\}_{i=1}^n$,*

$$(*) \quad \max\{(C/n) \sum_{i=1}^n a_i, (\sum_{i=1}^n a_i^r)^{1/r}\}$$
$$\le \| \sum_{i=1}^n a_i Y_i \|_r \le (6r/\mathrm{Log}\, r)\max\{(C/n) \sum_{i=1}^n a_i, (\sum_{i=1}^n a_i^r)^{1/r}\}.$$

PROOF. We first note that

$$\{\tilde{Y}_i\}_{i=1}^n = \{CY_i(\sum_{j=1}^n Y_j)^{-1}\}_{i=1}^n$$

is exchangeable with respect to

$$d\nu = C^{-r}(\textstyle\sum_{i=1}^n Y_i)^r \, dP.$$

Furthermore $\sum_{i=1}^n \tilde{Y}_i \equiv C$ and for any $\{b_i\}_{i=1}^n$

$$\int |\textstyle\sum_{i=1}^n b_i \tilde{Y}_i|^r \, d\nu = \int |\textstyle\sum_{i=1}^n b_i Y_i|^r \, dP.$$

Hence without loss of generality we may assume that $\sum_{i=1}^n Y_i \equiv C$. With this normalization we can rewrite (*) in a form which looks more like the inequality in Theorem 2.5:

$$(**)\qquad \begin{aligned} \max\{\|\textstyle\sum_{i=1}^n a_i Y_i\|_1, (\sum_{i=1}^n a_i^r)^{1/r}\} &\le \|\textstyle\sum_{i=1}^n a_i Y_i\|_r \\ &\le (6r/\operatorname{Log} r)\{\|\textstyle\sum_{i=1}^n a_i Y_i\|_1, (\sum_{i=1}^n a_i^r)^{1/r}\}. \end{aligned}$$

The left inequality is now evident, so we prove the right inequality.

Set $m = [n/2] + 1$ and suppose for a moment that

$$\|\textstyle\sum_{i=1}^m a_i Y_i\|_r > (2r/\operatorname{Log} r)(\sum_{i=1}^m a_i^r)^{1/r}.$$

Proposition 2.2 then implies that

$$E(\textstyle\sum_{i=1}^m a_i Y_i)^r \le (2r/(\operatorname{Log} r)) \sum_{j=1}^m a_j E(\sum_{i=1; i\ne j}^m a_i Y_i)^{r-1} Y_j.$$

From the exchangeability of $(Y_i)_{i=1}^m$, we get for each $m < k \le n$ and all $1 \le j \le m$ that

$$E(\textstyle\sum_{i=1;i\ne j}^m a_i Y_i)^{r-1} Y_j = E(\sum_{i=1;i\ne j}^m a_i Y_i)^{r-1} Y_k$$

and hence by averaging over the set $\{j, m+1, \cdots, n\}$ (which has cardinality at least $n/2$) that

$$\begin{aligned} E(\textstyle\sum_{i=1;i\ne j}^m a_i Y_i)^{r-1} Y_j &\le (2/n) E(\textstyle\sum_{i=1;i\ne j}^m a_i Y_i)^{r-1}[Y_j + \textstyle\sum_{k=m+1}^n Y_k] \\ &\le (2C/n) E(\textstyle\sum_{i=1}^m a_i Y_i)^{r-1}. \end{aligned}$$

Thus,

$$\begin{aligned} E(\textstyle\sum_{i=1}^m a_i Y_i)^r &\le \frac{2C}{n} \frac{2r}{\operatorname{Log} r} \textstyle\sum_{i=1}^m a_i E(\sum_{j=1}^m a_j Y_j)^{r-1} \\ &\le \frac{4C}{n} \frac{r}{\operatorname{Log} r} (\textstyle\sum_{i=1}^n a_i) \|\textstyle\sum_{i=1}^m a_i Y_i\|_r^{r-1} \end{aligned}$$

so that

$$\|\textstyle\sum_{i=1}^m a_i Y_i\|_r \le \frac{4C}{n} \frac{r}{\operatorname{Log} r} (\sum_{i=1}^m a_i).$$

Dropping now the assumption on $\|\sum_{i=1}^m a_i Y_i\|_r$, we have:

$$\|\textstyle\sum_{i=1}^m a_i Y_i\|_r \le \frac{r}{\operatorname{Log} r} \max\left\{\frac{4C}{n} (\sum_{i=1}^m a_i), 2(\sum_{i=1}^m a_i^r)^{1/r}\right\}.$$

Similarly we have:

$$\| \textstyle\sum_{i=m+1}^{n} a_i Y_i \|_r \le \frac{r}{\operatorname{Log} r} \max\left\{\frac{4C}{n} \sum_{i=m+1}^{n} a_i,\ 2(\sum_{i=m+1}^{n} a_i^r)^{1/r}\right\}.$$

Thus:

$$\| \textstyle\sum_{i=1}^{n} a_i Y_i \|_r \le \frac{r}{\operatorname{Log} r} \max\left\{\frac{4C}{n} \sum_{i=1}^{n} a_i,\ 2(\sum_{i=1}^{m} a_i^r)^{1/r} + 2(\sum_{i=m+1}^{n} a_i^r)^{1/r},\ \frac{4C}{n} \sum_{i=1}^{n} a_i + 2(\sum_{i=1}^{n} a_i^r)^{1/r}\right\}$$

$$\le \frac{6r}{\operatorname{Log} r} \max\left\{\frac{C}{n} \textstyle\sum_{i=1}^{n} a_i,\ (\sum_{i=1}^{n} a_i^r)^{1/r}\right\}. \quad \square$$

The deduction of Corollary 2.6 from Theorem 2.5 now gives:

COROLLARY 2.8. *Let $Y_1, \cdots, Y_n$ be a symmetrically exchangeable sequence of random variables with $\| Y_i \|_p = 1$ $(2 \le p < \infty)$ and set $C = \| (\sum_{i=1}^{n} Y_i^2)^{1/2} \|_p$. Then for all scalars $\{a_i\}_{i=1}^{n}$,*

$$\max\left\{\frac{C}{\sqrt{n}} (\textstyle\sum_{i=1}^{n} a_i^2)^{1/2},\ (\sum_{i=1}^{n} | a_i |^p)^{1/p}\right\}$$

$$\le \| \textstyle\sum_{i=1}^{n} a_i Y_i \|_p \le \frac{3\, p}{\sqrt{\operatorname{Log} p}} \max\left\{\frac{C}{\sqrt{n}} (\sum_{i=1}^{n} a_i^2)^{1/2},\ (\sum_{i=1}^{n} | a_i |^p)^{1/p}\right\}.$$

We conclude this section with an example showing that, up to a universal constant, $p/\operatorname{Log} p$ is best possible in Theorems 2.5 and 2.7.

PROPOSITION 2.9. *For $1 \le p < \infty$ let $C = C_p$ be the smallest constant so that for all nonnegative, independent, identically distributed random variables and all $m = 1, 2, \cdots$ we have:*

$$\| S_m \|_p \le C \max\{m \| X_1 \|_1,\ m^{1/p} \| X_1 \|_p\} \quad (\textit{where } S_m \equiv (\textstyle\sum_{i=1}^{m} X_i)).$$

Then $C \ge p/(e \operatorname{Log} p)$ for all $1 \le p < \infty$.

PROOF. Since $p/\operatorname{Log} p = p$ for $p \le e$, the statement is obvious when $p \le e$, so we assume $p \ge e$.

For $e \le p \le 6$ we consider the $\{0, 1\}$-valued i.i.d. sequence $(X_i)_{i=1}^{\infty}$ with $P[X_1 = 1] = ½$. Then

$$2 \| X_1 \|_1 = 1 = 2^{1/p} \| X_1 \|_p\ .$$

while

$$\| S_2 \|_p \ge \| S_2 \|_e = \left(\frac{2^e + 2}{4}\right)^{1/e} \ge 1.32 \ge 1.23\left(\approx \frac{6}{e \operatorname{Log} 6}\right) \ge \frac{p}{e \operatorname{Log} p}.$$

For $6 \le p < \infty$, we consider the $\{0, 1)$-valued i.i.d. sequence $(X_i)_{i=1}^{\infty}$, where

$P[X_i = 1] = \operatorname{Log} p/p$. Then

$$\| S_m \|_p \geq mP[S_m = m]^{1/p} = m\left(\frac{\operatorname{Log} p}{p}\right)^{m/p},$$

so

$$m\left(\frac{\operatorname{Log} p}{p}\right)^{m/p} \leq C \max\left\{m \frac{\operatorname{Log} p}{p}, \left(m \frac{\operatorname{Log} p}{p}\right)^{1/p}\right\}$$

or

$$C \geq \min\left\{\frac{p}{\operatorname{Log} p}\left(\frac{\operatorname{Log} p}{p}\right)^{m/p}, m^{1-1/p}\left(\frac{\operatorname{Log} p}{p}\right)^{(m-1)/p}\right\}.$$

Choosing m so that $m - 1 \leq p/\operatorname{Log} p < m$, we see that

$$C \geq \min\left\{\frac{p}{\operatorname{Log} p}\left(\frac{\operatorname{Log} p}{p}\right)^{1/\operatorname{Log} p+1/p}, \left(\frac{p}{\operatorname{Log} p}\right)^{1-1/p}\left(\frac{\operatorname{Log} p}{p}\right)^{1/\operatorname{Log} p}\right\}$$

$$= \frac{p}{\operatorname{Log} p}\left(\frac{\operatorname{Log} p}{p}\right)^{1/\operatorname{Log} p+1/p}$$

$$\geq \frac{p}{\operatorname{Log} p} \frac{(\operatorname{Log} p)^{1/\operatorname{Log} p}}{p^{1/\operatorname{Log} p} p^{1/p}} = \frac{p}{e \operatorname{Log} p} \frac{(\operatorname{Log} p)^{1/\operatorname{Log} p}}{p^{1/p}}.$$

Now the function $f(x) = x^{1/x}$ is increasing for $1 \leq x \leq e$ and decreasing for $e \leq x < \infty$, so for $p \geq e^e$,

$$\frac{(\operatorname{Log} p)^{1/\operatorname{Log} p}}{p^{1/p}} \geq 1$$

and

$$\min_{6 \leq p \leq e^e} \frac{(\operatorname{Log} p)^{1/\operatorname{Log} p}}{p^{1/p}} = \frac{(\operatorname{Log} 6)^{1/\operatorname{Log} 6}}{6^{1/6}} > 1. \quad \square$$

REMARK. To see that $r/\operatorname{Log} r$ is the best possible rate of growth in Theorem 2.7, consider (for the Y_i's) the random variables $\{X_i\}_{i=1}^n$ from the proof of Proposition 2.9 ($6 \leq p < \infty$). Take

$$p = r \qquad s = [r/\operatorname{Log} r], \quad n \text{ large}$$

and

$$a_i = \begin{cases} 1 & i = 1, 2, \cdots, s \\ 0 & i > s. \end{cases}$$

Then from the above proof (for $n \geq s$) $\| \sum_{i=1}^n a_i X_i \|_r \doteq \| \sum_{i=1}^s X_i \|_r \geq d(r/\log r)$ for some positive constant d.

On the other hand,

$$\frac{C}{n} = \frac{\| \sum_{i=1}^n X_i \|_r}{n} \to \| EX_1 \|_r = \| X_1 \|_1$$

as $n \to \infty$. Hence,

$$\max\left\{\frac{C}{n} \sum_{i=1}^n a_i, (\sum_{i=1}^n a_i^r)^{1/r}\right\} \to \max\left\{\frac{\operatorname{Log} r}{r}\left[\frac{r}{\operatorname{Log} r}\right], \left[\frac{r}{\operatorname{Log} r}\right]^{1/r}\right\}$$

which is "essentially" a constant.

Section 3. This section contains some inequalities needed for the proofs of Theorems 4.1 and 4.2 below. We begin with a martingale version of the "arc sinh" inequality of Prokhorov.

Let $\mathscr{F}_1 \subseteq \mathscr{F}_2 \subseteq \cdots \subseteq \mathscr{F}_n \subseteq \mathscr{F}$ be an increasing sequence of σ-fields on some probability space $(\Omega, \mathscr{F}, P)$. We denote by $E_i(\cdot) = E(\cdot \mid \mathscr{F}_i)$, $i = 1, \cdots, n$ the conditional expectation operators with respect to this sequence. We also put $E_0 = E$ the expectation operator.

PROPOSITION 3.1. *Let $(d_i)_{i=1}^n$ be a mean zero bounded martingale difference sequence with respect to $(\mathscr{F}_i)_{i=1}^n$; i.e., $E_{i-1}d_i = 0$, $i = 1, \cdots, n$. Let $\delta_i = \| d_i \|_\infty$, $M = \max_{1\le i\le n} \delta_i$ and $\eta_i^2 = \| E_{i-1}d_i^2 \|_\infty$, $i = 1, \cdots, n$. Then, for each $t \ge 0$,*

$$P\{|\sum_{i=1}^n d_i| \ge t\} \le 2 \exp\left\{-\frac{t}{2M} \operatorname{arc\,sinh}\left(\frac{Mt}{2\sum_{i=1}^n \eta_i^2}\right)\right\}.$$

PROOF. Using the two numerical inequalities $e^x - x - 1 \le e^x + e^{-x} - 2$ and $\cosh x - 1 \le \frac{1}{2} x \sinh x$ which hold for all x, we get for all $\lambda > 0$ and all $i = 1, \cdots, n$

$$\begin{aligned} &E_{i-1}e^{\lambda d_i} - 1 \\ (1)\qquad &= E_{i-1}(e^{\lambda d_i} - \lambda d_i - 1) \le E_{i-1}(\lambda d_i \sinh \lambda d_i) \\ &= E_{i-1}(\lambda |d_i| \sinh\lambda |d_i|) = E_{i-1}\left(\lambda^2 d_i^2 \frac{\sinh \lambda |d_i|}{\lambda |d_i|}\right) \le \frac{\lambda\eta_i^2}{M} \sinh \lambda M. \end{aligned}$$

Since $x \le e^{x-1}$ for all x, we get from (1)

$$\begin{aligned} E \exp(\lambda \textstyle\sum_{i=1}^n d_i) &= EE_{n-1}\exp(\lambda \textstyle\sum_{i=1}^n d_i) \\ &= E \exp(\lambda \textstyle\sum_{i=1}^{n-1} d_i)E_{n-1}\exp(\lambda d_n) \\ &\le E \exp(\lambda \textstyle\sum_{i=1}^{n-1} d_i)\exp(E_{n-1}e^{\lambda d_n} - 1) \\ &\le E \exp(\lambda \textstyle\sum_{i=1}^{n-1} d_i)\exp(\lambda\eta_n^2/M)\sinh \lambda M. \end{aligned}$$

Iterating the above argument, we obtain

$$(2)\qquad E \exp(\lambda \sum_{i=1}^n d_i) \le \exp\left(\lambda \sum_{i=1}^n \eta_i^2 \frac{\sinh \lambda M}{M}\right).$$

Therefore, for all $\lambda, t > 0$ we have

$$\begin{aligned} P\{\textstyle\sum_{i=1}^n d_i \ge t\} &= P\{\exp(\lambda \textstyle\sum_{i=1}^n d_i - \lambda t) \ge 1\} \\ &\le E \exp(\lambda \textstyle\sum_{i=1}^n d_i - \lambda t) \le \exp \lambda\left(\textstyle\sum_{i=1}^n \eta_i^2 \dfrac{\sinh \lambda M}{M} - t\right). \end{aligned}$$

Let $\lambda_0 = (1/M)\text{arc sinh}(Mt/(2\sum_{i=1}^n \eta_i^2))$. Then $\sum_{i=1}^n \eta_i^2(\sinh \lambda_0 M/M) = t/2$ so that

$$P\{\textstyle\sum_{i=1}^n d_i \geq t\} \leq \exp\left(-\frac{\lambda_0 t}{2}\right) = \exp\left\{\frac{-t}{2M} \text{arc sinh} \frac{Mt}{2\sum_{i=1}^n \eta_i^2}\right\}$$

and

$$P\{|\textstyle\sum_{i=1}^n d_i| \geq t\} \leq 2\exp\left\{\frac{-t}{2M} \text{arc sinh} \frac{Mt}{2\sum_{i=1}^n \eta_i^2}\right\}. \quad \square$$

When specialized to mean zero, independent random variables, Theorem 3.1 becomes Prokhorov's inequality:

COROLLARY 3.2. *Let $\{X_j\}_{j=1}^n$ be independent, mean zero random variables, $S = \sum_{j=1}^n X_j$, $N = \sup_j |X_j|$ and assume that $\|N\|_\infty < \infty$. Then*

$$P(S \geq t) \leq \exp\left(\frac{-t}{2\|N\|_\infty} \text{arc sinh} \frac{\|N\|_\infty t}{2\|S\|_2^2}\right)$$

and

$$P(|S| \geq t) \leq 2\exp\left(\frac{-t}{2\|N\|_\infty} \text{arc sinh} \frac{\|N\|_\infty t}{2\|S\|_2^2}\right).$$

The next three results give estimates for the pth norm of a random variable in terms of some exponential norms. We begin with some notation.

Fix $0 < \gamma < \infty$. We denote by C_γ the class of all functions with the following two properties:

(a) $\psi: (-\infty, \infty) \to [0, \infty)$
$\psi(0) = 0$; ψ is continuous and even.

(b) $\dfrac{\psi(t)}{t^\gamma}$ is increasing on $[0, \infty)$.

For a random variable X and $\psi \in C_\gamma$ we denote

$$\|X\|_\psi = \inf\{\lambda > 0; E\exp(\psi(X/\lambda)) \leq e\}.$$

Note that if ψ is convex $\|\cdot\|_\psi$ is actually a norm.

LEMMA 3.3. *Let $0 < \gamma < \infty$ and $\psi \in C_\gamma$. Then*

$$\left(\frac{t}{\psi^{-1}(p/\gamma)}\right)^p \leq \exp(\psi(t))$$

for all $t \geq 0$, $p > 0$.

PROOF. We first consider the case $\gamma = 1$. If $t \leq \psi^{-1}(p)$, then $(t/(\psi^{-1}(p)))^p \leq 1 \leq \exp(\psi(t))$. If $t > \psi^{-1}(p)$, then, by (b) of the definition of C_1

$$\frac{p}{\psi^{-1}(p)} = \frac{\psi(\psi^{-1}(p))}{\psi^{-1}(p)} \leq \frac{\psi(t)}{t}.$$

Since $(x/p)^p \le (ex/p)^p \le e^x$ for all $x \ge 0$,

$$\left(\frac{t}{\psi^{-1}(p)}\right)^p \le \left(\frac{\psi(t)}{p}\right)^p \le \exp(\psi(t)).$$

For general γ we consider $\phi(t) = \psi(|t|^{1/\gamma})$ and note that $\phi \in C_1$. □

PROPOSITION 3.4. *Let $0 < \gamma < \infty$, $\psi \in C_\gamma$, ρ, λ, $t_0 > 0$ and X a random variable which satisfies*

$$P(|X| \ge t) \le \rho \exp(-\psi(t/\lambda)) \quad \textit{for} \quad t \ge t_0 > 0.$$

Then,

$$\|X\|_\psi \le \lambda \max\left\{\left(\frac{2\rho + e}{e}\right)^{1/\gamma}, \frac{t_0}{\lambda\psi^{-1}(\ln e/2)}\right\}.$$

PROOF. We first assume $\lambda = 1$. For any $A \ge 1$, $\alpha > 1$

$$\begin{aligned} &E \exp \psi\left(\frac{X}{\alpha}\right) \\ &\le A + \int_A^\infty P\left(\exp\left(\psi\left(\frac{X}{\alpha}\right)\right) \ge s\right) ds \\ &= A + \int_A^\infty P(|X| \ge \alpha\psi^{-1}(\ln s))\, ds \le A + \rho \int_A^\infty \exp(-\psi(\alpha\psi^{-1}(\ln s)))\, ds \end{aligned}$$

if $\alpha\psi^{-1}(\ln A) \ge t_0$. From (b) of the definition of C_γ we get

$$E \exp \psi\left(\frac{X}{\alpha}\right) \le A + \rho \int_A^\infty \exp(-\alpha^\gamma \ln s)\, ds \le A + \frac{\rho}{(\alpha^\gamma - 1)A^{\alpha^\gamma - 1}}.$$

Now take

$$\alpha = \max\left\{\left(\frac{2\rho + e}{e}\right)^{1/\gamma}, \frac{t_0}{\psi^{-1}(\ln(e/2))}\right\} \quad \text{and} \quad A = \exp \psi\left(\frac{t_0}{\alpha}\right).$$

Then $A \le e/2$ and

$$\frac{\rho}{(\alpha^\gamma - 1)A^{\alpha^\gamma - 1}} \le \frac{\rho}{\alpha^\gamma - 1} \le \frac{e}{2}.$$

Hence

$$\|X\|_\psi \le \max\left\{\left(\frac{2\rho + e}{e}\right)^{1/\gamma}, \frac{t_0}{\psi^{-1}(\ln(e/2))}\right\}.$$

For the general case consider X/λ. □

The next corollary hints at the way we are going to apply Prokhorov's inequality.

COROLLARY 3.5. *Set* $\psi(t) = t\ln(1+t)$. *If for all* $u \geq 0$ *and some* $M, K > 0$ *we have*

$$P\{|Y| \geq u\} \leq 2\exp\left(-\frac{u}{2M}\operatorname{arc\,sinh}\frac{uM}{2K^2}\right)$$

then

$$\|Y\|_\psi \leq 2((4+e)/e)\max\{M, K\}.$$

PROOF. Since $\operatorname{arc\,sinh} s \geq \ln(1+s)$ for all $s \geq 0$, we get, putting $L = \max\{M, K\}$

$$P\{|Y| \geq u\} \leq 2\exp\left(-\frac{u}{2M}\ln\left(1+\frac{uM}{2K^2}\right)\right) \leq 2\exp\left(-\frac{u}{2L}\ln\left(1+\frac{uL}{2K^2}\right)\right)$$

$$\leq 2\exp\left(-\frac{u}{2L}\ln\left(1+\frac{u}{2L}\right)\right) = 2\exp\left(-\psi\left(\frac{u}{2L}\right)\right).$$

Applying Proposition 3.4, we get

$$\|Y\|_\psi \leq 2L\frac{4+e}{e} = 2\frac{4+e}{e}\max\{M, K\}. \quad \square$$

The last proposition of this section gives a relation between the pth norms and the exponential norms. Only the left hand side of the following inequality is going to be used in the next section.

PROPOSITION 3.6. *Let* $0 < \gamma,\ r < \infty$ *and* $\psi \in C_\gamma$. *Then for any random variable* X

$$e^{-1/r}\sup_{p\geq r}\frac{\|X\|_p}{\psi^{-1}(p/\gamma)} \leq \|X\|_\psi$$

$$\leq e^{1/\gamma}\max\left\{\left(\frac{2+e}{e}\right)^{1/\gamma}, \frac{\psi^{-1}(r/\gamma)}{\psi^{-1}(\ln e/2)}\right\}\sup_{p\geq r}\frac{\|X\|_p}{\psi^{-1}(p/\gamma)}.$$

PROOF. By Lemma 3.3, if $\lambda > \|X\|_\psi$

$$E\left(\frac{|X|/\lambda}{\psi^{-1}(p/\gamma)}\right)^p \leq E\exp\psi\left(\frac{X}{\lambda}\right) \leq e.$$

Hence,

$$\|X\|_p \leq e^{1/p}\psi^{-1}(p/\gamma)\|X\|_\psi$$

and the left-hand side inequality follows.

Assume now $\sup_{p\geq r}(\|X\|_p/\psi^{-1}(p/\gamma)) \leq 1$. Then for any $t \geq 0$,

$$P(|X| \geq t) \leq \left(\frac{\|X\|_p}{t}\right)^p \leq \left(\frac{\psi^{-1}(p/\gamma)}{t}\right)^p.$$

If t is such that $\gamma\psi(t/e^{1/\gamma}) \geq r$, then for $p = \gamma\psi(t/e^{1/\gamma})$, we get

$$P(|X| \geq t) \leq \exp\{-\psi(t/e^{1/\gamma})\}$$

for all $t \geq e^{1/\gamma}\,\psi^{-1}(r/\gamma)$. Now apply Proposition 3.4. □

Section 4.

THEOREM 4.1. *Let $X_1, \cdots, X_n$ be independent, symmetric random variables with finite pth moment for some $p > 2$. Then*

$$\max\{\|\textstyle\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\}$$
$$\leq \|\textstyle\sum_{i=1}^n X_i\|_p \leq K(p/\mathrm{Log}\, p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\}$$

where $K \leq 7.35$ is an absolute constant.

PROOF. The left-hand side inequality is easy; cf. [4]. For the right-hand side inequality set

$$A_i = \{|X_i| \geq \|\textstyle\sum_{j=1}^n X_j\|_2\} \quad i = 1, \cdots, n$$

then

$$\|\textstyle\sum_{i=1}^n X_i\|_p \leq \|\sum_{i=1}^n X_i I_{A_i^c}\|_p + \|\sum_{i=1}^n X_i I_{A_i}\|_p, \tag{1}$$

and the summands in each of the two terms on the right are independent. Now apply Proposition 3.8 and Corollaries 3.2 and 3.5, with $\psi(t) = t\ln(1 + t)$ to get

$$\|\textstyle\sum_{i=1}^n X_i I_{A_i^c}\|_p \leq e^{1/p}\psi^{-1}(p)\,\|\sum_{i=1}^n X_i I_{A_i^c}\|_\psi$$
$$\leq 2((4 + e)/e)e^{1/p}\psi^{-1}(p)\max\{\max_{1\leq i\leq n}\|X_i I_{A_i^c}\|_\infty, \|\textstyle\sum_{i=1}^n X_i I_{A_i^c}\|$$
$$\leq 2((4 + e)/e)e^{1/p}\psi^{-1}(p)\,\|\textstyle\sum_{i=1}^n X_i\|_2.$$

(Note that $X_i I_{A_i^c}$ are symmetric and thus orthogonal so that $\|\sum_{i=1}^n X_i I_{A_i^c}\|_2 \leq \|\sum_{i=1}^n X_i\|_2$.)

To estimate the second term on the right of (1) we use Theorem 2.5

$$\|\textstyle\sum_{i=1}^n X_i I_{A_i}\|_p \leq \|\sum_{i=1}^n |X_i| I_{A_i}\|_p$$
$$\leq (2p/\mathrm{Log}\, p)\mathit{max}\{\|\textstyle\sum_{i=1}^n |X_i| I_{A_i}\|_1, (\sum_{i=1}^n \|X_i I_{A_i}\|_p^p)^{1/p}\}.$$

But,

$$\|\textstyle\sum_{i=1}^n |X_i| I_{A_i}\|_1 \leq \|\sum_{i=1}^n |X_i|\,|X_i|/\|\sum_{i=1}^n X_i\|_2\|_1 = \|\sum_{i=1}^n X_i\|_2.$$

Hence,

$$\|\textstyle\sum_{i=1}^n X_i I_{A_i}\|_p \leq (2p/\mathrm{Log}\, p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\}$$

and

$$\|\textstyle\sum_{i=1}^n X_i\|_p \leq K_p(p/\mathrm{Log}\, p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p^p)^{1/p}\}$$

with

$$K_p = 2\,\frac{4+e}{e}\,e^{1/p}\psi^{-1}(p)\,\frac{\operatorname{Log} p}{p} + 2.$$

As $p \to \infty$, $K_p \to 2(4 + e)/e + 2 \le 7$, so this proves the theorem with some absolute K. To show that $K \le 7.35$ for all $p \ge 2$, notice that by Corollary 2.6, $K \le 7.35$ as long as $\sqrt{\operatorname{Log} p} \le 7.35$, in particular if $p \le e^{54}$.

For $p > e^{54}$ we first check that

$$(*) \qquad \exp(e^{1/54})(1.08)2((4+e)/e) < 5.35$$

so that it is enough to check that

$$\psi^{-1}(p) \le 1.08(p/\operatorname{Log} p) \quad \text{for} \quad p \ge e^{54},$$

or

$$p \le \frac{1.08p}{\operatorname{Log} p}\ln\left(1 + \frac{1.08p}{\operatorname{Log} p}\right) \qquad p \ge e^{54}.$$

But, as can be easily checked, even the stronger inequality

$$\ln p \le 1.08\ln(p/\ln p) \qquad p \ge e^{54}$$

holds. □

REMARK. A simple symmetrization argument shows that Theorem 4.1 remains true if we replace the symmetry assumption by merely requiring that $EX_i = 0$, $i = 1, \cdots, n$ and changing the bound on K to $2 \cdot 7.35 = 14.7$.

We now turn to the symmetrically exchangeable case.

THEOREM 4.2. *Let $Y_1, \cdots, Y_n$ be a symmetrically exchangeable sequence of random variables with $\| Y_i \|_p = 1$ for some $2 \le p < \infty$. Set $C = \| (\sum_{i=1}^n Y_i^2)^{1/2} \|_p$; then, for all scalars $a_1, \cdots, a_n$,*

$$\max\{(C/\sqrt{n})(\textstyle\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\}$$
$$\le \| \textstyle\sum_{i=1}^n a_i Y_i \|_p \le K(p/\operatorname{Log} p)\max\{(C/\sqrt{n})(\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\}$$

where $K \le 21.2$ is an absolute constant.

PROOF. As in the proof of Theorem 2.7, we may assume without loss of generality that $\sum_{i=1}^n Y_i^2 = C^2$. With this normalization

$$\| \textstyle\sum_{i=1}^n a_i Y_i \|_2 = (C/\sqrt{n})(\sum_{i=1}^n a_i^2)^{1/2}$$

and the left-hand side inequality follows easily (cf. [3]).

For $i = 1, \cdots, n$, let

$$A_i = \{ | a_i Y_i | \ge \| \textstyle\sum_{j=1}^n a_j Y_j \|_2 \}$$

and set $d_i = a_i Y_i I_{A_i^c}$. Notice that since the Y_i's are symmetrically exchangeable

$\{d_i\}_{i=1}^n$ is a martingale difference sequence relative to $\{\sigma(Y_i, \cdots, Y_i)\}_{i=1}^n$ with $Ed_1 = 0$ and

$$\| d_i \|_\infty \le \| \textstyle\sum_{j=1}^n a_j Y_j \|_2 \quad i = 1, \cdots, n.$$

In order to apply Proposition 3.1 it is enough to evaluate $\| E_{i-1} d_i^2 \|_\infty$, where E_j is the conditional expectation operator with respect to $\sigma(Y_1, \cdots, Y_j)$, $j = 1, \cdots, n$. By the exchangeability of $Y_1, \cdots, Y_n$ we have that

$$E_{i-1} Y_i^2 = E_{i-1} Y_k^2 \quad \text{for} \quad i \le k \le n.$$

Thus,

$$\| E_{i-1} Y_i^2 \|_\infty = \frac{1}{n-i+1} \| \textstyle\sum_{k=i}^n E_{i-1} Y_k^2 \|_\infty \le \frac{1}{n-i+1} \sum_{k=1}^n Y_k^2 = \frac{C^2}{n-i+1}.$$

Setting $m = [(n+1)/2]$ we get

$$\| E_{i-1} Y_i^2 \|_\infty \le C^2/m \quad \text{for } 1 \le i \le m$$

and

$$\| E_{i-1} d_i^2 \|_\infty \le a_i^2 C^2/m \quad \text{for} \quad 1 \le i \le m.$$

Proposition 3.1 now yields

$$P\{ | \textstyle\sum_{i=1}^m d_i | \ge t \} \le 2 \exp\left(\frac{-t}{2 \| \sum_{i=1}^n a_i Y_i \|_2} \operatorname{arc\,sinh} \frac{\| \sum_{i=1}^n a_i Y_i \|_2 t}{2D} \right)$$

where $D = (C^2 \sum_{i=1}^m a_i^2)/m$, so that by Corollary 3.5,

$$\| \textstyle\sum_{i=1}^m d_i \|_\psi \le 2\left(\frac{4+e}{e}\right) \max\left\{ \| \sum_{i=1}^n a_i Y_i \|_2, \frac{C}{\sqrt{m}} (\sum_{i=1}^m a_i^2)^{1/2} \right\}$$

$$\le 2^{3/2} \left(\frac{4+e}{e}\right) \frac{C}{\sqrt{n}} (\textstyle\sum_{i=1}^n a_i^2)^{1/2}.$$

Since a similar inequality holds for $\| \sum_{i=m+1}^n d_i \|_\psi$, we conclude from Proposition 3.6 that

$$\text{(2)} \quad \| \textstyle\sum_{i=1}^n a_i Y_i I_{A_i^c} \|_p = \| \sum_{i=1}^n d_i \|_p \le e^{1/p} \psi^{-1}(p) 2^{5/2} \left(\frac{4+e}{e}\right) \frac{C}{\sqrt{n}} (\sum_{i=1}^n a_i^2)^{1/2}.$$

We turn now to the estimate of $\| \sum_{i=1}^n a_i Y_i I_{A_i} \|_p$. Set again $m = [(n+1)/2]$; then by Proposition 2.2

$$\begin{aligned} &E | \textstyle\sum_{i=1}^m a_i Y_i I_{A_i} |^p \\ &\quad \le E(\textstyle\sum_{i=1}^m | a_i Y_i I_{A_i} |)^p \\ \text{(3)} \quad &\quad \le \max\left\{ \frac{2p}{\operatorname{Log} p} \textstyle\sum_{j=1}^m E[(\sum_{i=1; i\ne j}^m | a_i Y_i I_{A_i} |)^{p-1} | a_j Y_j I_{A_j} |], \right. \\ &\qquad\qquad \left. \left(\frac{2p}{\operatorname{Log} p}\right)^p \textstyle\sum_{i=1}^m | a_i |^p E | Y_i I_{A_i} |^p \right\}. \end{aligned}$$

Assume for a minute that the maximum in (3) occurs in the first term. Then by the exchangeability of $Y_1, \cdots, Y_n$ we can write

$$E(\textstyle\sum_{i=1}^m |a_i Y_i I_{A_i}|)^p$$

$$\le \frac{2p}{\operatorname{Log} p} \textstyle\sum_{j=1}^m |a_j| E\left[(\sum_{i=1; i\ne j}^m |a_i Y_i I_{A_i}|)^{p-1} \frac{2}{n} (|Y_j| I_{A_j} + \sum_{k=m+1}^n |Y_k| I_{[|a_j Y_k| \ge \|\sum_{s=1}^n a_s Y_s\|_2]})\right]$$

$$\le \frac{4p}{n \operatorname{Log} p} \textstyle\sum_{j=1}^m |a_j| (E(\sum_{i=1}^m |a_i Y_i I_{A_i}|)^p)^{(p-1)/p} (E(\sum_{k=1}^n |Y_k| I_{[|a_j Y_k| \ge \|\sum_{s=1}^n a_s Y_s\|_2]})^p)^{1/p}.$$

Thus by dividing by the first expectation on the right and then applying Hölder's inequality to the remaining integrand, we have

$$\| \textstyle\sum_{i=1}^m |a_i Y_i I_{A_i}| \|_p$$

$$\le \frac{4p}{n \operatorname{Log} p} \textstyle\sum_{j=1}^m |a_j| (E(\sum_{k=1}^n Y_k^2 \sum_{\ell=1}^n I_{[|a_j Y_\ell| \ge \|\sum_{s=1}^n a_s Y_s\|_2]})^{p/2})^{1/p}.$$

Now,

$$I_{[|a_j Y_k| \ge \|\sum_{s=1}^n a_s Y_s\|_2]} \le \frac{a_j^2 Y_k^2}{\|\sum_{s=1}^n a_s Y_s\|_2^2}.$$

Thus,

$$\| \textstyle\sum_{i=1}^m |a_i Y_i I_{A_i}| \|_p$$

$$\le \frac{4p}{n \operatorname{Log} p} C \textstyle\sum_{j=1}^m |a_j| \left(E\left(a_j^2 \sum_{k=1}^n \frac{Y_k^2}{\|\sum_{s=1}^n a_s Y_s\|_2^2}\right)^{p/2}\right)^{1/p}$$

$$\le \frac{4p}{n \operatorname{Log} p} C^2 \textstyle\sum_{j=1}^m a_j^2 \frac{1}{\|\sum_{i=1}^n a_i Y_i\|_2} = \frac{4p}{\operatorname{Log} p} \frac{C}{\sqrt{n}} \sum_{j=1}^m \frac{a_j^2}{(\sum_{j=1}^n a_j^2)^{1/2}}.$$

Dropping now the assumption that the maximum in (3) occurs in the first term, we get

$$\| \textstyle\sum_{i=1}^m a_i Y_i I_{A_i} \|_p \le \frac{p}{\operatorname{Log} p} \max\left\{\frac{4C}{\sqrt{n}} \frac{\sum_{i=1}^m a_i^2}{(\sum_{i=1}^n a_i^2)^{1/2}}, 2(\sum_{i=1}^m |a_i|^p)^{1/p}\right\}.$$

Since a similar estimate holds for $\| \sum_{i=m+1}^n a_i Y_i I_{A_i} \|_p$, we get, by checking cases,

$$\| \textstyle\sum_{i=1}^n a_i Y_i I_{A_i} \|_p$$

$$\le \frac{p}{\operatorname{Log} p} \max\left\{\frac{4C}{\sqrt{n}} (\textstyle\sum_{i=1}^n a_i^2)^{1/2}, 4(\sum_{i=1}^n |a_i|^p)^{1/p}, \frac{4C}{\sqrt{n}} (\sum_{i=1}^n a_i^2)^{1/2} + 2(\sum_{i=1}^n |a_i|^p)^{1/p}\right\}$$

$$\le \frac{6p}{\operatorname{Log} p} \max\left\{\frac{C}{\sqrt{n}} (\textstyle\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\right\}.$$

Combining this with (2) we have

$$\| \textstyle\sum_{i=1}^n a_i Y_i \|_p \le K_p \frac{p}{\operatorname{Log} p} \max\left\{\frac{C}{\sqrt{n}} (\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p}\right\}$$

with

$$K_p = e^{1/p}\psi^{-1}(p)2^{5/2}(4 + e)/e + 6.$$

This shows that the theorem holds with some absolute constant K (≤ 20.2 for large p). To see that $K \leq 21.2$ we use the estimate (*) at the latter part of the proof of Theorem 4.1 to show that $K_p \leq 21.2$ for $p \geq e^{54}$. For $p \leq e^{54}$ we use Corollary 2.8 to get

$$\sqrt{3}\,\frac{p}{\sqrt{\operatorname{Log} p}} \leq \sqrt{3}\,\sqrt{54}\,\frac{p}{\operatorname{Log} p} \leq 21.2\,\frac{p}{\operatorname{Log} p}. \quad \square$$

Next we give an example showing that, up to a universal constant, $p/\operatorname{Log} p$ is the best possible constant in Theorems 4.1 and 4.2.

PROPOSITION 4.3. *For $2 \leq p < \infty$ let $C = C_p$ be the smallest constant so that for all symmetric, independent, identically distributed random variables $X_1, X_2, \cdots$, and all $m = 1, 2, \cdots$ we have*

$$\| S_m \|_p \leq C \max\{m^{1/2} \| X_1 \|_2, m^{1/p} \| X_1 \|_p\}.$$

Then $C \geq p/(2^{1/2}e \operatorname{Log} p)$.

PROOF. For $2 \leq p \leq e$ this is clear. For $e \leq p \leq 6$ consider the Rademacher sequence $P(X_i = 1) = P(X_i = -1) = 1/2$. Then

$$\| X_1 \|_2 = \| X_1 \|_p = 1 \quad \text{and} \quad \| S_2 \|_p \geq \| S_2 \|_e = 2^{1-1/e}.$$

So

$$C \geq 2^{1-1/e}\min\{2^{-1/2}, 2^{-1/p}\} \geq 1$$

while $p/(2^{1/2}e \operatorname{Log} p) \leq 1$. For $p \geq 6$ we consider the i.i.d. $\{-1, 0, 1\}$-valued sequence $X_1, X_2, \cdots$ where

$$P(X_i = 1) = P(X_i = -1) = \operatorname{Log} p/p.$$

Then

$$\| S_m \|_p \geq mP\{S_m = m\}^{1/p} = m(\operatorname{Log} p/p)^{m/p}.$$

So

$$m\left(\frac{\operatorname{Log} p}{p}\right)^{m/p} \leq C \max\left\{m^{1/2}2^{1/2}\left(\frac{\operatorname{Log} p}{p}\right)^{1/2}, m^{1/p}2^{1/p}\left(\frac{\operatorname{Log} p}{p}\right)^{1/p}\right\}$$

or

$$C \geq \min\left\{2^{-1/2}m^{1/2}\left(\frac{p}{\operatorname{Log} p}\right)^{1/2-m/p}, 2^{-1/p}m^{1-1/p}\left(\frac{p}{\operatorname{Log} p}\right)^{1/p-m/p}\right\}.$$

Choosing m so that

$$m - 1 \leq p/\operatorname{Log} p < m$$

we get

$$C \geq \min\left\{2^{-1/2}\frac{p}{\operatorname{Log} p}\left(\frac{p}{\operatorname{Log} p}\right)^{-m/p}, 2^{-1/p}\frac{p}{\operatorname{Log} p}\left(\frac{p}{\operatorname{Log} p}\right)^{-m/p}\right\}$$

$$\geq 2^{-1/2}\frac{p}{\operatorname{Log} p}\left(\frac{\operatorname{Log} p}{p}\right)^{1/\operatorname{Log} p+1/p}$$

and the conclusion follows from the computation at the end of the proof of Proposition 2.9. □

We conclude this section with an inequality similar in nature to the main results of this paper. Recall that a sequence $Y_1, \cdots, Y_n$ in L_p is said to be K-unconditional if

$$\| \sum_{i=1}^n a_i Y_i \|_p \leq K \| \sum_{i=1}^n \varepsilon_i a_i Y_i \|_p$$

for all scalars $a_1, \cdots, a_n$ and *all* signs $\varepsilon_1, \cdots, \varepsilon_n$.

THEOREM 4.4. *For all $K, L \geq 1, p \geq 2$, if $X_1, \cdots, X_n$ is K-unconditional in L_p and $X_1^2 - EX_1^2, \cdots, X_n^2 - EX_n^2$ is L-unconditional in $L_{p/2}$ then,*

$$\| \sum_{i=1}^n X_j \|_p \leq 2KLp \max\{(\sum_{j=1}^n \| X_j \|_p^p)^{1/p}, (\sum_{j=1}^n \| X_j \|_2^2)^{1/2}\}.$$

PROOF. From the version of Khintchine's inequality which is found in [2] we have

$$(E | \sum a_j \varepsilon_j^* |^r)^{1/r} \leq ((r \vee 2)/2)^{1/2} (\sum a_j^2)^{1/2},$$

where $\{\varepsilon_j^*\}$ are i.i.d. with $P(\varepsilon_i^* = 1) = P(\varepsilon_i^* = -1) = ½$. Hence,

$$\| \sum_{j=1}^n X_j \|_p \leq K(E_\varepsilon \| \sum_{j=1}^n \varepsilon_j^* X_j \|_p^p)^{1/p} \leq K(p/2)^{1/2} \| (\sum_{j=1}^n X_j^2)^{1/2} \|_p,$$

where E_ε is the expectation with respect to $\{\varepsilon_j^*\}$. Applying this to $X_j^2 - EX_j^2$ with $p/2$ we get

$$\begin{aligned}
&\| (\sum_{j=1}^n X_j^2)^{1/2} \|_p \\
&= \| \sum_{j=1}^n X_j^2 \|_{p/2}^{1/2} \leq \| \sum_{j=1}^n (X_j^2 - EX_j^2) \|_{p/2}^{1/2} + (\sum_{j=1}^n EX_j^2)^{1/2} \\
&\leq L^{1/2}(p/2)^{1/4} \| (\sum_{j=1}^n (X_j^2 - EX_j^2)^2)^{1/2} \|_{p/2}^{1/2} + (\sum_{j=1}^n EX_j^2)^{1/2} \qquad (4)\\
&\leq L^{1/2}(p/2)^{1/4} \| (\sum_{j=1}^n X_j^4)^{1/2} \|_{p/2}^{1/2} + L^{1/2}(p/2)^{1/4} (\sum_{j=1}^n (EX_j^2)^2)^{1/4} + (\sum_{j=1}^n EX_j^2)^{1/2} \\
&\leq 2\max\{L^{1/2}(p/2)^{1/4} \| (\sum_{j=1}^n X_j^4)^{1/2} \|_{p/2}^{1/2}, (L^{1/2}(p/2)^{1/4} + 1)(\sum_{j=1}^n EX_j^2)^{1/2}\}.
\end{aligned}$$

For $p > 4$ we write $4 = \theta p + (1 - \theta)2$ i.e.; $\theta = 2/(p - 2)$. Then, by Holder's inequality

$$\sum_{j=1}^n X_j^4 \leq (\sum_{j=1}^n | X_j |^p)^\theta (\sum_{j=1}^n X_j^2)^{1-\theta}.$$

Hence

$$\begin{aligned}
E(\sum_{j=1}^n X_j^4)^{p/4} &\leq E[(\sum_{j=1}^n | X_j |^p)^{\theta p/4} (\sum_{j=1}^n X_j^2)^{(1-\theta)p/4}] \\
&\leq (E \sum_{j=1}^n | X_j |^p)^{\theta p/4} (E(\sum_{j=1}^n X_j^2)^{p/2})^{(1-\theta)/2}.
\end{aligned}$$

Thus,

$$\| (\textstyle\sum_{j=1}^{n} X_j^4)^{1/2} \|_{p/2}^{1/2} \le (E \sum_{j=1}^{n} | X_j |^p)^{\theta/4} (E(\sum_{j=1}^{n} X_j^2)^{p/2})^{(1-\theta)/2p}.$$

If the maximum in (4) is attained in the first term then

$$\| (\textstyle\sum_{j=1}^{n} X_j^2)^{1/2} \|_p^{(1+\theta)/2} \le 2L^{1/2}(p/2)^{1/4}(E \sum_{j=1}^{n} | X_j |^p)^{\theta/4}$$

and

$$\| (\textstyle\sum_{j=1}^{n} X_j^2)^{1/2} \|_p \le (2L^{1/2}(p/2)^{1/4})^{2(p-2)/p}(E \sum_{j=1}^{n} | X_j |^p)^{1/p}.$$

Thus, dropping the last assumption

$$\| (\textstyle\sum_{j=1}^{n} X_j^2)^{1/2} \|_p \le \max\{(2L^{1/2}(p/2)^{1/4})^{(2(p-2))/p}(E \sum_{j=1}^{n} | X_j |^p)^{1/p}, 2(L^{1/2}(p/2)^{1/4} + 1)(\sum_{j=1}^{n} EX_j^2)^{1/2}\}$$

and

$$\begin{aligned} \| \textstyle\sum_{j=1}^{n} X_j \|_p &\le p \max\{K(2L^{1/2})^{2(p-2)/p}p^{1-1/p}(\textstyle\sum_{j=1}^{n} \| X_j \|_p^p)^{1/p}, \\ &\qquad (2KL^{1/2}(p/2)^{3/4} + 2K(p/2)^{1/2})(\textstyle\sum_{j=1}^{n} \| X_j \|_2^2)^{1/2}\} \\ &\le 2KLp \max\{(\textstyle\sum_{j=1}^{n} \| X_j \|_p^p)^{1/p}, (\sum_{j=1}^{n} \| X_j \|_2^2)^{1/2}\}, \quad \text{since} \quad K, L \ge 1. \end{aligned}$$

If $p < 4$,

$$E(\textstyle\sum_{j=1}^{n} X_j^4)^{p/4} \le E \sum_{j=1}^{n} | X_j |^p$$

and we get

$$\| \textstyle\sum_{j=1}^{n} X_j \|_p \le 2K(p/2)^{1/2}\max\{L^{1/2}(p/2)^{1/4}(\sum_{j=1}^{n} \| X_j \|_p^p)^{1/p}, (L^{1/2}(p/2)^{1/4} + 1)(\sum_{j=1}^{n} \| X_j \|_2^2)^{1/2}\}. \quad \square$$

REMARK. A similar result may be obtained if we replace the assumption

$$X_1^2 - EX_1^2, \cdots, X_n^2 - EX_n^2 \quad L\text{-unconditional}$$

by

$$X_1^2 - X_1'^2, \cdots, X_n^2 - X_n'^2 \quad L\text{-unconditional},$$

where $\{X_1', \cdots, X_n'\}$ is an independent copy of $\{X_1, \cdots, X_n\}$.

REFERENCES

[1] BURKHOLDER, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–43.

[2] HAAGERUP, U. (1982). The best constants in the Khintchine inequality. *Studia math.* **70** 232–283.

[3] JOHNSON, W. B., MAUREY, B., SCHECHTMAN, G. and TZAFRIRI, L. (1979). Symmetric structures in Banach spaces. *Memoirs of the A.M.S.* **217**.

[4] ROSENTHAL, H. P. (1970). On the subspaces of $L^p(p > 2)$ spanned by sequences of independent random variables. *Israel J. Math.* **8** 273–303.
[5] STOUT, W. R. (1974). *Almost Sure Convergence.* Academic, New York.

W. B. JOHNSON AND J. ZINN
DEPARTMENT OF MATHEMATICS
TEXAS A&M UNIVERSITY
COLLEGE STATION, TEXAS 77843-3368

G. SCHECHTMAN
DEPARTMENT OF THEORETICAL MATHEMATICS
THE WEIZMANN INSTITUTE OF SCIENCE
REHOVOT, ISRAEL