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:

\[
\begin{align*}
\max \left\{ \| \sum_{i=1}^n x_i \|_2, \ (\sum_{i=1}^n \| x_i \|_p)^{1/p} \right\} \\
\leq \| \sum_{i=1}^n x_i \|_p \leq B_p \max \left\{ \| \sum_{i=1}^n x_i \|_2, \ (\sum_{i=1}^n \| x_i \|_p)^{1/p} \right\}
\end{align*}
\]

for all independent symmetric random variables \(x_1, x_2, \ldots\), with
finite \(p\)th moment, \(2 < p < \infty\).

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 \(\sqrt{p \log p}\), 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 vari-
able \(x\) 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:

\[
\begin{align*}
\max \left\{ \| \sum_{i=1}^n x_i \|_2, \ (\sum_{i=1}^n \| x_i \|_p)^{1/p} \right\} \\
\leq \| \sum_{i=1}^n x_i \|_p \leq B_p \max \left\{ \| \sum_{i=1}^n x_i \|_2, \ (\sum_{i=1}^n \| x_i \|_p)^{1/p} \right\}
\end{align*}
\]

for all independent symmetric random variables \(x_1, x_2, \ldots\), with finite
\(p\)th moment, \(2 < p < \infty\).

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.
$p \to \infty$ is $p/\log p$. 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 $p/\log p$ as $p \to \infty$ ((A) is contained in [4] while (C) and (D) were proved in Chapter 1 of [3]).

\[
\begin{align*}
\max \{ & \| \sum_{i=1}^{\infty} X_i \|_1, (\sum_{i=1}^{\infty} \| X_i \|_p^{1/p}) \} \\
\leq & \| \sum_{i=1}^{\infty} X_i \|_p \leq A_p \max \{ \| \sum_{i=1}^{\infty} X_i \|_1, (\sum_{i=1}^{\infty} \| X_i \|_p^{1/p}) \}
\end{align*}
\]
for all independent nonnegative random variables $X_1, X_2, \ldots, X_n$ with finite $p$th moment, $1 \leq p < \infty$.

\[
\begin{align*}
\max \{ & Cn^{-1/2} \sum_{i=1}^{n} a_i, (\sum_{i=1}^{n} a_i^p)^{1/p} \} \leq \| \sum_{i=1}^{n} a_i Y_i \|_p \\
\leq & C_p \max \{ Cn^{-1/2} \sum_{i=1}^{n} a_i, (\sum_{i=1}^{n} a_i^p)^{1/p} \}
\end{align*}
\]
(\text{where } C = \| \sum_{i=1}^{n} Y_i \|_p)

for all nonnegative exchangeable random variables $Y_1, Y_2, \ldots, Y_n$ with $\E Y_i^p = 1$, $1 \leq p < \infty$ and all nonnegative scalars $a_1, a_2, \ldots, a_n$.

\[
\begin{align*}
\max \{ & Cn^{-1/2}(\sum_{i=1}^{n} | a_i |^2)^{1/2}, (\sum_{i=1}^{n} | a_i |^p)^{1/p} \} \\
\leq & \| \sum_{i=1}^{n} a_i Y_i \|_p \\
\leq & D_p \max \{ Cn^{-1/2}(\sum_{i=1}^{n} | a_i |^2)^{1/2}, (\sum_{i=1}^{n} | a_i |^p)^{1/p} \}
\end{align*}
\]
(\text{where } C = \| (\sum_{i=1}^{n} | Y_i |^2)^{1/2} \|_p)

for all symmetrically exchangeable random variables $Y_1, Y_2, \ldots, Y_n$ with $\E Y_i^p = 1$; $2 \leq p < \infty$ and all scalars $a_1, a_2, \ldots, a_n$.

We recall that $(Y_1, \ldots, Y_n)$ is exchangeable (respectively, symmetrically exchangeable) provided $(Y_{\pi(1)}, \ldots, Y_{\pi(n)})$ (respectively, $(\epsilon_1 Y_{\pi(1)}, \ldots, \epsilon_n Y_{\pi(n)})$) has the same joint distribution as $(Y_1, \ldots, Y_n)$ for every permutation $\pi$ of $\{1, \ldots, n\}$ (and every choice $\epsilon_i = \pm 1$).

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

\[
B_p, D_p \leq Kp/\sqrt{\log p}
\]
for some absolute constant $K$. To check that $B_p$ and $D_p$ are of order $p/\log p$ 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 $p$th 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 $1 \leq p < \infty$

$$p/e \log p \leq A_p \leq 2p/\log p$$

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

$$B_p \leq p/\sqrt{\log p}.$$ 

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

$$B_p \leq 7.35p/\log p$$

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 $p$th 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

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

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

$$(a + 1)s^{-t} \leq (1 + 1/cs)^s \leq e^{1/c}.$$  

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

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

Here and through the whole paper we use the following:

**Convention.** $\log t = \max\{1, \ln t\}$, where "ln" is the natural logarithm function.

**Proposition 2.2.** Let $X_1, \ldots, X_n$ be nonnegative random variables with finite $r$th moment for some $1 \leq r < \infty$. Then

$$E(\sum_{i=1}^n X_i)^r \leq \max\left\{K \frac{r}{\log r} \sum_{j=1}^n E(\sum_{i=1; i \neq j}^n X_i)^{r-1} X_j, \left(K \frac{r}{\log r}\right)^r \sum_{j=1}^n EX_j\right\},$$

where $K$ is an absolute constant. (In fact, $K \leq 2.$)
PROOF. Applying Lemma 2.1 with \( s = r - 1 \), we obtain for every \( c > 0 \):
\[
E(\sum_{i=1}^{n} X_i)^r = \sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} X_j
\]
\[
\leq e^{1/c} \sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} X_j + [c(r - 1) + 1]^{-1} \sum_{j=1}^{n} EX_j
\]
\[
\leq \max\{2e^{1/c} \sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} X_j, 2[c(r - 1) + 1]^{-1} \sum_{j=1}^{n} EX_j\}.
\]

The conclusion follows by setting
\[
e^{1/c} = \begin{cases} \frac{r}{\log r}, & \text{if } r \geq 2 \\ 2, & \text{if } 1 \leq r \leq 2. \end{cases}
\]

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

**Sublemma 2.3.** For all \( r \geq 2 \),
\[
\frac{r - 1}{\ln(r/\log r)} + 1 \leq \frac{2r}{\log r}.
\]

To check that \( K \leq 2 \) in Proposition 2.2 where \( r \leq 2 \), we use the inequality \((a + b)^r \leq (2a)^r + (2b)^r\) in the argument for Proposition 2.2. (This is just Rosenthal’s original argument [4].) This gives
\[
E(\sum_{i=1}^{n} X_i)^r \leq \max\{2^r \sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} X_j, 2^r \sum_{j=1}^{n} EX_j\}.
\]
Since \( 2^r \leq 2r = 2(r/\log r) \) for \( 1 \leq r \leq 2 \), we get that \( K \leq 2 \) in the range \( 1 \leq r \leq 2 \).

**Theorem 2.5.** If \( X_1, \ldots, X_n \) are nonnegative, independent random variables with finite \( r \)th moments \((1 \leq r < \infty)\), then
\[
\max\{\|\sum_{i=1}^{n} X_i\|_1, (\sum_{i=1}^{n} X_i^r)^{1/r}\}
\]
\[
\leq \|\sum_{i=1}^{n} X_i\|_r \leq K (r/\log r) \max\{\|\sum_{i=1}^{n} X_i\|_1, (\sum_{i=1}^{n} X_i^r)^{1/r}\}
\]
where \( K \leq 2 \) 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
\[
\sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} X_j
\]
\[
= \sum_{j=1}^{n} E(\sum_{i=1; i \neq j}^{n} X_i)^{r-1} (EX_j) \leq E(\sum_{i=1}^{n} X_i)^{r-1} (\sum_{j=1}^{n} EX_j).
\]
So from Proposition 2.2 we have
\[
E(\sum_{i=1}^{n} X_i)^r \leq \max\{K(r/\log r) E(\sum_{i=1}^{n} X_i)^{r-1} \sum_{i=1}^{n} EX_i, (K(r/\log r))^r \sum_{i=1}^{n} EX_i\}.
\]
If this maximum in the preceding expression occurs in the second term we get
\[
\|\sum_{i=1}^{n} X_i\|_r \leq K(r/\log r)(\sum_{i=1}^{n} EX_i)^{1/r}.
\]
Otherwise we have
\[ E(\sum_{i=1}^{n} X_i)' \leq K(r/\log r) E(\sum_{i=1}^{n} X_i)^{\gamma - 1} \sum_{i=1}^{n} E X_i \]
\[ \leq K(r/\log r)[E(\sum_{i=1}^{n} X_i)^{\gamma}]^{(\gamma - 1)/\gamma} \sum_{i=1}^{n} E X_i \]
so that
\[ [E(\sum_{i=1}^{n} X_i)^{\gamma}]^{1/\gamma} \leq K(r/\log r) \sum_{i=1}^{n} E X_i. \quad \Box \]

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

**Corollary 2.6.** Suppose that \( X_1, \ldots, X_n \) are independent, symmetric random variables with finite \( p \)th moments \((2 \leq p < \infty)\). Then
\[
\max\{\| \sum_{i=1}^{n} X_i \|_2, (\sum_{i=1}^{n} \| X_i \|_p^{1/p})\}
\leq \| \sum_{i=1}^{n} X_i \|_p \leq (p/(\log p))^{1/2}[\max\{\| \sum_{i=1}^{n} X_i \|_2, (\sum_{i=1}^{n} \| X_i \|_p^{1/p})\}.
\]

**Proof.** The left inequality is known and easy; cf. e.g. [4]. The right one follows just as in [4]: For any choice \( \epsilon_i = \pm 1 \) we have by the symmetry and independence of the \( X_i \)'s that
\[
\| \sum_{i=1}^{n} X_i \|_p = \| \sum_{i=1}^{n} \epsilon_i X_i \|_p.
\]

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

Since \( p/2 \geq 1 \), we get from Theorem 2.5
\[
\| \sum_{i=1}^{n} X_i^2 \|_p^{1/2} \leq 2 \frac{p/2}{\log p} \max\{\| \sum_{i=1}^{n} X_i^2 \|_1, [\sum_{i=1}^{n} \| X_i^2 \|_p^{2/p}]^{2/p}\}
\leq \frac{2p}{\log p} \max\{\| \sum_{i=1}^{n} X_i \|_2, (\sum_{i=1}^{n} \| X_i \|_p^{2/p})\}.
\]

Thus
\[
\| \sum_{i=1}^{n} X_i \|_p \leq (p/2)^{1/2}(2p/\log p)^{1/2}\max\{\| \sum_{i=1}^{n} X_i \|_2, (\sum_{i=1}^{n} \| X_i \|_p^{2/p})\}. \quad \Box
\]

**Theorem 2.7.** Let \( Y_1, \ldots, Y_n \) be a sequence of nonnegative exchangeable random variables with \( \| Y_i \|_r = 1 \) \((1 \leq r < \infty)\) and set \( C = \| \sum_{i=1}^{n} Y_i \|_r \). Then for all nonnegative scalars \( |a_i|_{r-1} \),
\[
\max\{|C/n| \sum_{i=1}^{n} a_i, (\sum_{i=1}^{n} a_i^{1/r})^{1/r}\}
\leq \| \sum_{i=1}^{n} a_i Y_i \|_r \leq (6r/\log r)\max\{|C/n| \sum_{i=1}^{n} a_i, (\sum_{i=1}^{n} a_i^{1/r})^{1/r}\}.
\]

**Proof.** We first note that
\[
|Y_i|_{r-1} = |CY_i(\sum_{i=1}^{n} Y_j)^{-1}|_{r-1}.
\]

\[
\sum_{i=1}^{n} a_i Y_i \leq (\sum_{i=1}^{n} a_i^{1/r})^{1/r}\sum_{i=1}^{n} Y_i \leq (6r/\log r)\max\{|C/n| \sum_{i=1}^{n} a_i, (\sum_{i=1}^{n} a_i^{1/r})^{1/r}\}.
\]
is exchangeable with respect to
\[ d\nu = C^{-r}(\sum_{i=1}^{n} Y_i)^r \, dP. \]

Furthermore \( \sum_{i=1}^{n} \bar{Y}_i = C \) and for any \( |b_i|_{1} \)
\[ \int |\sum_{i=1}^{n} b_i\bar{Y}_i|^r \, d\nu = \int |\sum_{i=1}^{n} b_iY_i|^r \, dP. \]

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

\[
\max\{\| \sum_{i=1}^{n} a_iY_i \|_1, (\sum_{i=1}^{n} a_i^r)^{1/r} \} \leq \left( \frac{6r}{\log r} \right) \| \sum_{i=1}^{n} a_iY_i \|_r, (\sum_{i=1}^{n} a_i^r)^{1/r} \}.
\]

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

Set \( m = \lceil n/2 \rceil + 1 \) and suppose for a moment that
\[ \| \sum_{i=1}^{m} a_iY_i \|_r > (2r/\log r)(\sum_{i=1}^{m} a_i^r)^{1/r}. \]

Proposition 2.2 then implies that
\[ E(\sum_{i=1}^{m} a_iY_i)^r \leq (2r/(\log r)) \sum_{i=1}^{m} a_iE(\sum_{i=1, i \neq j}^{m} a_iY_i)^{r-1}Y_j. \]

From the exchangeability of \( (Y_i)_{i=1}^{n} \), we get for each \( m < k \leq n \) and all \( 1 \leq j \leq m \) that
\[ E(\sum_{i=1, i \neq j}^{m} a_iY_i)^{r-1}Y_j = E(\sum_{i=1, i \neq j}^{m} a_iY_i)^{r-1}Y_k \]
and hence by averaging over the set \( \{ j, m+1, \ldots, n \} \) (which has cardinality at least \( n/2 \)) that
\[
E(\sum_{i=1, i \neq j}^{m} a_iY_i)^{r-1}Y_j \leq (2/n)E(\sum_{i=1, i \neq j}^{m} a_iY_i)^{r-1}[Y_j + \sum_{k=m+1}^{n} Y_k]
\]
\[ \leq (2C/n)E(\sum_{i=1}^{m} a_iY_i)^{r-1}. \]

Thus,
\[
E(\sum_{i=1}^{m} a_iY_i)^r \leq \frac{2C}{n} \frac{2r}{\log r} \sum_{i=1}^{m} a_iE(\sum_{i=1}^{m} a_iY_i)^{r-1}
\]
\[ \leq \frac{4C}{n} \frac{r}{\log r} \left( \sum_{i=1}^{n} a_i \right) \| \sum_{i=1}^{m} a_iY_i \|_r^{r-1} \]
so that
\[ \| \sum_{i=1}^{m} a_iY_i \|_r \leq \frac{4C}{n} \frac{r}{\log r} \left( \sum_{i=1}^{n} a_i \right). \]

Dropping now the assumption on \( \| \sum_{i=1}^{m} a_iY_i \|_r \), we have:
\[ \| \sum_{i=1}^{m} a_iY_i \|_r \leq \frac{r}{\log r} \max \left\{ \frac{4C}{n} \left( \sum_{i=1}^{m} a_i \right), 2(\sum_{i=1}^{m} a_i)^{1/r} \right\}. \]
Similarly we have:

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

Thus:

$$\| \sum_{i=1}^n a_i Y_i \|_r \leq \frac{r}{\log r} \max \left\{ \frac{4C}{\sqrt{n}} \sum_{i=1}^n a_i, 2(\sum_{i=1}^n a_i)^{1/r} + 2(\sum_{i=m+1}^n a_i)^{1/r}, 4C \sum_{i=1}^n a_i \right\} \leq \frac{6r}{\log r} \max \left\{ \frac{C}{\sqrt{n}} \sum_{i=1}^n a_i, (\sum_{i=1}^n a_i)^{1/r} \right\}. \quad \square$$

The deduction of Corollary 2.6 from Theorem 2.5 now gives:

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

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

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

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

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

Then $C \geq p/(e \log p)$ for all $1 \leq p < \infty$.

**Proof.** Since $p/\log p = p$ for $p \leq e$, the statement is obvious when $p \leq e$, so we assume $p \geq e$.

For $e \leq p \leq 6$ we consider the $\{0, 1\}$-valued i.i.d. sequence $(X_i)_{i=1}^m$ with $P[X_i = 1] = \frac{1}{2}$. Then

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

while

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

For $6 \leq p < \infty$, we consider the $\{0, 1\}$-valued i.i.d. sequence $(X_i)_{i=1}^m$, where
\( P[X_1 = 1] = \log p/p. \) Then
\[
\| S_m \|_p \geq m P[S_m = m]^{1/p} = m \left( \frac{\log p}{p} \right)^{m/p},
\]
so
\[
m \left( \frac{\log p}{p} \right)^{m/p} \leq C \max \left\{ m \frac{\log p}{p}, \left( m \frac{\log p}{p} \right)^{1/p} \right\}.
\]
or
\[
C \geq \min \left\{ \frac{p}{\log p} \left( \frac{\log p}{p} \right)^{1/\log p + 1/p}, \left( \frac{p}{\log p} \right)^{1-1/p} \left( \frac{\log p}{p} \right)^{(m-1)/p} \right\}.
\]
Choosing \( m \) so that \( m - 1 \leq p/\log p < m \), we see that
\[
C \geq \min \left\{ \frac{p}{\log p} \left( \frac{\log p}{p} \right)^{1/\log p + 1/p}, \left( \frac{p}{\log p} \right)^{1-1/p} \left( \frac{\log p}{p} \right)^{1/\log p} \right\} = \frac{p}{\log p} \left( \frac{\log p}{p} \right)^{1/\log p + 1/p} \geq \frac{p}{e \log p} \left( \frac{\log p}{p} \right)^{1/\log p}.
\]
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 \( p \geq e^e \),
\[
\left( \frac{\log p}{p} \right)^{1/\log p} \geq 1
\]
and
\[
\min_{6 < s \leq e} \frac{\left( \frac{\log p}{p} \right)^{1/\log p}}{p^{1/p}} = \frac{(\log 6)^{1/\log 6}}{6^{1/6}} > 1. \quad \square
\]

**Remark.** To see that \( r/\log r \) 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
\[
p = r \quad s = [r/\log r], \quad n \text{ large}
\]
and
\[
a_i = \begin{cases} 1 & i = 1, 2, \ldots, s \\ 0 & i > s. \end{cases}
\]
Then from the above proof (for \( n \geq s \)) \( \| \sum_{i=1}^n a_i X_i \|_r = \| \sum_{i=1}^s X_i \|_r \geq d(r/\log r) \) for some positive constant \( d \).

On the other hand,
\[
\frac{C}{n} = \frac{\| \sum_{i=1}^n X_i \|_r}{n} \rightarrow \| EX_i \|_r = \| X_1 \|_1
\]
as $n \to \infty$. Hence,

$$\max \left\{ \frac{C}{n} \sum_{i=1}^{n} a_i, \left( \sum_{i=1}^{n} a_i^2 \right)^{1/2} \right\} \to \max \left\{ \frac{\log r}{r}, \left( \frac{r}{\log r} \right)^{1/2} \right\}$$

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 $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ be an increasing sequence of $\sigma$-fields on some probability space $(\Omega, \mathcal{F}, P)$. We denote by $E_i(\cdot) = E(\cdot \mid \mathcal{F}_i)$, $i = 1, \ldots, n$ 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 $(\mathcal{F}_i)_{i=1}^{n}$; i.e., $E_i d_i = 0$, $i = 1, \ldots, n$. Let $\delta_i = \|d_i\|_{\infty}$, $M = \max_{1 \leq i \leq n} \delta_i$ and $\eta_i = \|E_{i-1} d_i\|_{\infty}$, $i = 1, \ldots, n$. Then, for each $t \geq 0$,

$$P\{|\sum_{i=1}^{n} d_i| \geq t\} \leq 2 \exp \left( -\frac{t}{2M} \arcsinh \left( \frac{M}{2 \sum_{i=1}^{n} \eta_i^2} \frac{Mt}{\eta_i^2} \right) \right).$$

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

$$E_{i-1} e^{\lambda d_i} - 1 = E_{i-1}(e^{\lambda d_i} - \lambda d_i - 1) \leq E_{i-1}(\lambda d_i \sinh \lambda d_i).$$

(1) 

$$= E_{i-1}(\lambda |d_i| \sinh \lambda |d_i|) = E_{i-1} \left( \lambda^2 d_i^2 \frac{\sinh \lambda |d_i|}{\lambda |d_i|} \right) \leq \frac{\lambda \eta_i^2}{M} \sinh \lambda M.$$

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

$$E \exp(\lambda \sum_{i=1}^{n} d_i) = E E_{n-1} \exp(\lambda \sum_{i=1}^{n} d_i)$$

$$= E \exp(\lambda \sum_{i=1}^{n-1} d_i) E_{n-1} \exp(\lambda d_n)$$

$$\leq E \exp(\lambda \sum_{i=1}^{n-1} d_i) \exp\left(E_{n-1} e^{\lambda d_n} - 1\right)$$

$$\leq E \exp(\lambda \sum_{i=1}^{n-1} d_i) \exp(\lambda \eta_n^2 / M) \sinh \lambda M.$$

Iterating the above argument, we obtain

(2) 

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

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

$$P\{\sum_{i=1}^{n} d_i \geq t\} = P\{|\exp(\lambda \sum_{i=1}^{n} d_i) - \lambda t| \geq 1\}$$

$$\leq E \exp(\lambda \sum_{i=1}^{n} d_i - \lambda t) \leq \exp \left( \sum_{i=1}^{n} \eta_i^2 \frac{\sinh \lambda M}{M} - t \right).$$
Let $\lambda_0 = (1/M) \arcsinh(Mt/(2 \sum_{i=1}^{n} \eta_i^2)).$ Then $\sum_{i=1}^{n} \eta_i^2(\sinh \lambda_0 M/M) = t/2$ so that

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

and

$$P(\| \sum_{i=1}^{n} d_i \| \geq t) \leq 2 \exp \left( -\frac{t}{2M} \arcsinh \frac{Mt}{2 \sum_{i=1}^{n} \eta_i^2} \right). \square$$

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

**Corollary 3.2.** Let $\{X_i\}_{i=1}^{n}$ be independent, mean zero random variables, $S = \sum_{i=1}^{n} X_i,$ $N = \sup_{i} |X_i|$ and assume that $\| N \|_{\infty} < \infty.$ Then

$$P(S \geq t) \leq \exp \left( -\frac{t}{2 \| N \|_{\infty}} \arcsinh \frac{\| N \|_{\infty} t}{2 \| S \|_{2}} \right)$$

and

$$P(\| S \| \geq t) \leq 2 \exp \left( -\frac{t}{2 \| N \|_{\infty}} \arcsinh \frac{\| N \|_{\infty} t}{2 \| S \|_{2}} \right).$$

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

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

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

(b) $\frac{\psi(t)}{t^\gamma}$ 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 $\psi$ 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 \((x/p)^p \leq (e^x/p)^p \leq e^x\) for all \(x \geq 0\),
\[
\left( \frac{t}{\psi^{-1}(p)} \right)^p \leq \left( \frac{\psi(t)}{p} \right)^p \leq \exp(\psi(t)).
\]

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

**Proposition 3.4.** Let \(0 < \gamma < \infty, \psi \in C_\gamma, \rho, \lambda, t_0 > 0\) and \(X\) a random variable which satisfies
\[
P(|X| \geq t) \leq \rho \exp(-\psi(t/\lambda)) \quad \text{for} \quad t \geq t_0 > 0.
\]
Then,
\[
\|X\|_\psi \leq \lambda \max \left\{ \left( \frac{2\rho + e}{e} \right)^{1/\gamma}, \frac{t_0}{\lambda \psi^{-1}(\ln e/2)} \right\}.
\]

**Proof.** We first assume \(\lambda = 1\). For any \(A \geq 1, \alpha > 1\)
\[
E \exp \left( \frac{X}{\alpha} \right) \leq A + \int_A^\infty P \left( \exp \left( \frac{X}{\alpha} \right) \geq s \right) ds
\]
\[
= A + \int_A^\infty P(|X| \geq \alpha \psi^{-1}(\ln s)) ds \leq A + \rho \int_A^\infty \exp(-\psi(\alpha \psi^{-1}(\ln s))) ds
\]
if \(\alpha \psi^{-1}(\ln A) \geq t_0\). From (b) of the definition of \(C_\gamma\), we get
\[
E \exp \left( \frac{X}{\alpha} \right) \leq A + \rho \int_A^\infty \exp(-\alpha \gamma \ln s) ds \leq A + \frac{\rho}{(\alpha \gamma - 1) A^{\alpha \gamma - 1}}.
\]
Now take
\[
\alpha = \max \left\{ \left( \frac{2\rho + e}{e} \right)^{1/\gamma}, \frac{t_0}{\psi^{-1}(\ln e/2)} \right\} \quad \text{and} \quad A = \exp \left( \frac{t_0}{\alpha} \right).
\]
Then \(A \leq e/2\) and
\[
\frac{\rho}{(\alpha \gamma - 1) A^{\alpha \gamma - 1}} \leq \frac{\rho}{\alpha \gamma - 1} \leq \frac{e}{2}.
\]
Hence
\[
\|X\|_\psi \leq \max \left\{ \left( \frac{2\rho + e}{e} \right)^{1/\gamma}, \frac{t_0}{\psi^{-1}(\ln e/2)} \right\}.
\]

For the general case consider \(X/\lambda\). □

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

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

then

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

**Proof.** Since $\ln(1 + s) \geq \ln(1 + s)$ for all $s \geq 0$, we get, putting $L = \max |M, K|

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

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

Applying Proposition 3.4, we get

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

The last proposition of this section gives a relation between the $p$th 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 $X$

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

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

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

$$E \left( \frac{|X|}{\lambda} \right)^p \leq E \exp \left( \frac{X}{\lambda} \right) \leq e.$$

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 $t \geq 0$,

$$P(|X| \geq t) \leq \left( \frac{\|X\|_p}{t} \right)^p \leq \left( \psi^{-1}(p/\gamma) \right)^p.$$
If \( t \) is such that \( \gamma \psi(t/e^{1/\gamma}) \geq r \), then for \( p = \gamma \psi(t/e^{1/\gamma}) \), we get

\[
P(\{X_i \geq t\} \leq \exp(-\psi(t/e^{1/\gamma}))
\]

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

Section 4.

**Theorem 4.1.** Let \( X_1, \ldots, X_n \) be independent, symmetric random variables with finite \( p \)th moment for some \( p > 2 \). Then

\[
\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p)^{1/p}\}
\leq \|\sum_{i=1}^n X_i\|_p \leq K(p/\log p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p)^{1/p}\}
\]

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

\[A_i = \{X_i \geq \sum_{j=1}^n X_j\} \quad i = 1, \ldots, n\]

then

\[\|\sum_{i=1}^n X_i\|_p \leq \|\sum_{i=1}^n X_i I_{A_i}\|_p + \|\sum_{i=1}^n X_i I_{\bar{A_i}}\|_p, \]

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

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

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

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

\[
\|\sum_{i=1}^n X_i I_{A_i}\|_p \leq \|\sum_{i=1}^n X_i I_{A_i}\|_p \leq (2p/\log p)\max\{\|\sum_{i=1}^n X_i | I_{A_i}\|_1, (\sum_{i=1}^n \|X_i I_{A_i}\|_p)^{1/p}\}
\]

But,

\[
\|\sum_{i=1}^n X_i | I_{A_i}\|_1 \leq \|\sum_{i=1}^n X_i | X_i\|_1/\|\sum_{i=1}^n X_i\|_2 = \|\sum_{i=1}^n X_i\|_2.
\]

Hence,

\[
\|\sum_{i=1}^n X_i I_{A_i}\|_p \leq (2p/\log p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p)^{1/p}\}
\]

and

\[
\|\sum_{i=1}^n X_i\|_p \leq K_p(p/\log p)\max\{\|\sum_{i=1}^n X_i\|_2, (\sum_{i=1}^n \|X_i\|_p)^{1/p}\}
\]
with
\[ K_p = 2 \frac{4 + e}{e} e^{1/p} \psi^{-1}(p) \frac{\log p}{p} + 2. \]

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

For \( p > e^{54} \) we first check that
\[ \exp(e^{54})(1.08)2((4 + e)/e) < 5.35 \]
so that it is enough to check that
\[ \psi^{-1}(p) \leq 1.08(p/\log p) \quad \text{for} \quad p \geq e^{54}, \]
or
\[ p \leq \frac{1.08p}{\log p} \ln \left( 1 + \frac{1.08p}{\log p} \right) \quad p \geq e^{54}. \]

But, as can be easily checked, even the stronger inequality
\[ \ln p \leq 1.08 \ln(p/\ln p) \quad p \geq e^{54} \]
holds. \( \Box \)

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 \( 2 \leq p < \infty \). Set \( C = \| (\sum_{i=1}^n Y_i^2)^{1/2} \|_p \); then, for all scalars \( a_1, \cdots, a_n \),
\[ \max\{ (C/\sqrt{n})(\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p} \} \]
\[ \leq \| \sum_{i=1}^n a_i Y_i \|_p \leq K(p/\log p) \max\{ (C/\sqrt{n})(\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n |a_i|^p)^{1/p} \} \]
where \( K \leq 21.2 \) 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
\[ \| \sum_{i=1}^n a_i Y_i \|_2 = (C/\sqrt{n})(\sum_{i=1}^n a_i^2)^{1/2} \]
and the left-hand side inequality follows easily (cf. [3]).

For \( i = 1, \cdots, n \), let
\[ A_i = \{ |a_i Y_i| \geq \| \sum_{j=1}^n a_j Y_j \|_2 \} \]
and set \( d_i = a_i Y_i I_{A_i} \). Notice that since the \( Y_i \)'s are symmetrically exchangeable
\( \{d_i\}_{i=1}^{n} \) is a martingale difference sequence relative to \( \{\sigma(Y_i, \cdots, Y_j)\}_{j=1}^{n} \) with \( Ed_{i} = 0 \) and

\[
\| d_{i} \|_{\infty} \leq \| \sum_{j=1}^{n} a_{j} Y_{j} \|_{2} \quad i = 1, \cdots, n.
\]

In order to apply Proposition 3.1 it is enough to evaluate \( \| E_{i-1}d_{i} \|_{\infty} \), 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

\[
E_{i-1} Y_{i}^{2} = E_{i-1} Y_{k}^{2} \quad \text{for} \quad i \leq k \leq n.
\]

Thus,

\[
\| E_{i-1} Y_{i}^{2} \|_{\infty} = \frac{1}{n-i+1} \| \sum_{k=1}^{n} E_{i-1} Y_{k}^{2} \|_{\infty} \leq \frac{1}{n-i+1} \sum_{k=1}^{n} Y_{k}^{2} = \frac{C^{2}}{n-i+1}.
\]

Setting \( m = \lceil (n+1)/2 \rceil \) we get

\[
\| E_{i-1} Y_{i}^{2} \|_{\infty} \leq C^{2}/m \quad \text{for} \quad 1 \leq i \leq m
\]

and

\[
\| E_{i-1}d_{i} \|_{\infty} \leq a_{i}^{2} C^{2}/m \quad \text{for} \quad 1 \leq i \leq m.
\]

Proposition 3.1 now yields

\[
P(\| \sum_{i=1}^{n} d_{i} \|_{1} \geq t) \leq 2 \exp\left(\frac{-t}{2 \| \sum_{i=1}^{n} a_{i} Y_{i} \|_{2} t} \right) \text{arc sinh} \frac{\| \sum_{i=1}^{n} a_{i} Y_{i} \|_{2} t}{2D}
\]

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

\[
\| \sum_{i=1}^{n} d_{i} \|_{\psi} \leq 2 \left(\frac{4 + e}{e}\right) \max \left\{ \| \sum_{i=1}^{n} a_{i} Y_{i} \|_{2}, \frac{C}{\sqrt{m}} (\sum_{i=1}^{n} a_{i}^{2})^{1/2} \right\}
\]

\[
\leq 2^{3/2} \left(\frac{4 + e}{e}\right) \frac{C}{\sqrt{n}} (\sum_{i=1}^{n} a_{i}^{2})^{1/2}.
\]

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

\[
(2) \quad \| \sum_{i=1}^{n} a_{i} Y_{i} I_{A_{i}} \|_{p} = \| \sum_{i=1}^{n} d_{i} \|_{p} \leq e^{1/p} \psi^{-1}(p) 2^{5/2} \left(\frac{4 + e}{e}\right) \frac{C}{\sqrt{n}} (\sum_{i=1}^{n} a_{i}^{2})^{1/2}.
\]

We turn now to the estimate of \( \| \sum_{i=1}^{n} a_{i} Y_{i} I_{A_{i}} \|_{p}. \) Set again \( m = \lceil (n+1)/2 \rceil \); then by Proposition 2.2

\[
E(\sum_{i=1}^{n} a_{i} Y_{i} I_{A_{i}})^{p} \leq E(\sum_{i=1}^{n} a_{i} Y_{i} I_{A_{i}})^{p}
\]

\[
\leq \max \left\{ \frac{2p}{\log p} \sum_{i=1}^{n} E[(\sum_{j=1}^{m} a_{j} | a_{i} Y_{i} I_{A_{i}} |)^{p-1} | a_{i} Y_{i} I_{A_{i}} |], \right\}

\[
\left(\frac{2p}{\log p}\right)^{p} \sum_{i=1}^{n} a_{i}^{p} E[| Y_{i} I_{A_{i}} |^{p}].
\]
Assume for a minute that the maximum in (3) occurs in the first term. Then by the exchangeability of \( Y_1, \ldots, Y_n \) we can write
\[
E(\Sigma_{i=1}^n | a_i Y_i | I_{A_i})^p \\
\leq \frac{2p}{\log p} \Sigma_{i=1}^n | a_j | E\left( \sum_{i,j=1}^n | a_i Y_i I_{A_i} | \right)^{p-1} \left( \frac{1}{n} \right) \sum_{i=1}^n Y_i | I_{(a_i Y_i | I_{(a_i Y_i | I_{a_i Y_i a_i})})}^{p-1} \right) \sum_{i=1}^n | a_j | \left( E(\sum_{k=1}^n Y_k^2 \sum_{i=1}^n I_{(a_i Y_i | I_{a_i Y_i a_i})})^{p/2} \right)^{1/p}.
\]
Thus by dividing by the first expectation on the right and then applying Hölder's inequality to the remaining integrand, we have
\[
\| \Sigma_{i=1}^n | a_j Y_i I_{A_i} | |_p \\
\leq \frac{4p}{n \log p} \Sigma_{i=1}^n | a_j | \left( E(\sum_{k=1}^n Y_k^2 \sum_{i=1}^n I_{(a_i Y_i | I_{a_i Y_i a_i})})^{p/2} \right)^{1/p}.
\]
Now,
\[
I_{(a_i Y_i | I_{a_i Y_i a_i})} \leq \frac{a_i^2 Y_k^2}{\| \Sigma_{i=1}^n | a_i Y_i | |_2}.
\]
Thus,
\[
\| \Sigma_{i=1}^n | a_j Y_i I_{A_i} | |_p \\
\leq \frac{4p}{n \log p} C \Sigma_{i=1}^n | a_j | \left( E(\sum_{k=1}^n Y_k^2 \sum_{i=1}^n a_i Y_i | |_2)^{p/2} \right)^{1/p} \\
\leq \frac{4p}{n \log p} C^2 \Sigma_{i=1}^n | a_j | \frac{1}{\| \Sigma_{i=1}^n | a_i Y_i | |_2} = \frac{4p}{n \log p} \frac{C}{\sqrt{n}} \Sigma_{i=1}^n \frac{a_i^2}{(\sum_{i=1}^n a_i^2)^{1/2}}.
\]
Dropping now the assumption that the maximum in (3) occurs in the first term, we get
\[
\| \Sigma_{i=1}^n | a_i Y_i I_{A_i} | |_p = \frac{p}{\log p} \max \left\{ \frac{4C}{\sqrt{n}} \frac{\sum_{i=1}^n a_i^2}{(\sum_{i=1}^n a_i^2)^{1/2}}, \frac{2(\sum_{i=1}^n | a_i |)^{1/p}} \right\}.
\]
Since a similar estimate holds for \( \sum_{i=m+1}^n | a_i Y_i I_{A_i} | |_p \), we get, by checking cases,
\[
\| \Sigma_{i=1}^n | a_i Y_i I_{A_i} | |_p \\
\leq \frac{p}{\log p} \max \left\{ \frac{4C}{\sqrt{n}} (\sum_{i=1}^n a_i^2)^{1/2}, \frac{2(\sum_{i=1}^n | a_i |)^{1/p}} \right\}.
\]
Combining this with (2) we have
\[
\| \Sigma_{i=1}^n | a_i Y_i | |_p \leq K_p \frac{p}{\log p} \max \left\{ \frac{C}{\sqrt{n}} (\sum_{i=1}^n a_i^2)^{1/2}, (\sum_{i=1}^n | a_i |)^{1/p} \right\}.
\]
with

\[ K_p = e^{1/p} \sqrt[3]{p} 2^{5/2} (4 + e)/e + 6. \]

This shows that the theorem holds with some absolute constant \( K \) (\( \leq 20.2 \) 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

\[ \sqrt{3} \frac{p}{\sqrt[3]{\log p}} \leq \sqrt{3} \frac{p}{\log p} \leq 21.2 \frac{p}{\log p}. \]

Next we give an example showing that, up to a universal constant, \( p/\log p \) 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 \( X_1, X_2, \ldots \), and all \( m = 1, 2, \ldots \) we have

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

Then \( C \geq p/(2^{1/2} \log p) \).

**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 \( p/(2^{1/2} \log p) \leq 1 \). For \( p \geq 6 \) we consider the i.i.d. \([-1, 0, 1]\)-valued sequence \( X_1, X_2, \ldots \) where

\[ P(X_i = 1) = P(X_i = -1) = \log p/p. \]

Then

\[ \| S_m \|_p \geq m P[S_m = m]^{1/p} = m (\log p/p)^{m/p}. \]

So

\[ m \left( \frac{\log p}{p} \right)^{m/p} \leq C \max\{m^{1/2} \left( \frac{\log p}{p} \right)^{1/2}, m^{1/p} 2^{1/p} \left( \frac{\log p}{p} \right)^{1/p} \} \]

or

\[ C \geq \min\left\{ 2^{-1/2} m^{1/2} \left( \frac{p}{\log p} \right)^{1/2-m/p}, 2^{-1/p} m^{1-1/p} \left( \frac{p}{\log p} \right)^{1/p-m/p} \right\}. \]

Choosing \( m \) so that

\[ m - 1 \leq p/\log p < m \]
we get
\[
C \geq \min \left\{ 2^{-1/2} \frac{p}{\log p} \left( \frac{p}{\log p} \right)^{-m/p}, \ 2^{-1/2} \frac{p}{\log p} \left( \frac{p}{\log p} \right)^{-m/p} \right\}
\]
\[
\geq 2^{-1/2} \frac{p}{\log p} \left( \frac{\log p}{p} \right)^{1/\log p + 1/p}
\]
and the conclusion follows from the computation at the end of the proof of Proposition 2.9. \(\square\)

We conclude this section with an inequality similar in nature to the main results of this paper. Recall that a sequence \(Y_1, \ldots, Y_n\) in \(L_p\) is said to be \(K\)-unconditional if
\[
\| \sum_{i=1}^n a_i Y_i \|_p \leq K \| \sum_{i=1}^n \epsilon_i a_i Y_i \|_p
\]
for all scalars \(a_1, \ldots, a_n\) and all signs \(\epsilon_1, \ldots, \epsilon_n\).

**Theorem 4.4.** For all \(K, L \geq 1, p \geq 2,\) if \(X_1, \ldots, X_n\) is \(K\)-unconditional in \(L_p\) and \(X_1^2 - EX_1^2, \ldots, X_n^2 - EX_n^2\) is \(L\)-unconditional in \(L_{p/2}\) then,
\[
\| \sum_{i=1}^n X_i \|_p \leq 2K \max \{ (\sum_{i=1}^n \| X_i \|_p^{p/2} (\sum_{i=1}^n \| X_i \|_p^{p/2})^{1/2}\}
\]

**Proof.** From the version of Khintchine’s inequality which is found in [2] we have
\[
(\mathbb{E} | \sum a_i \epsilon_i \frac{1}{r})^{1/r} \leq (r \vee 2)/2 \| (\sum a_i^2) \|^{1/2},
\]
where \(|\epsilon_i^r|\) are i.i.d. with \(P(\epsilon_i^r = 1) = P(\epsilon_i^r = -1) = \frac{1}{2}.\) Hence,
\[
\| \sum_{i=1}^n X_i \|_p \leq K(\mathbb{E} \| \sum_{i=1}^n \epsilon_i X_i \|_p^{p/2})^{1/p} \leq K(p/2)^{1/2} \| (\sum_{i=1}^n X_i^2) \|^{1/2} \|_p,
\]
where \(\mathbb{E}\) is the expectation with respect to \(\{\epsilon_i\}.\) Applying this to \(X_i^2 - EX_i^2\) with \(p/2\) we get
\[
\| (\sum_{i=1}^n X_i^2) \|^{1/2} \|_p
\]
\[
= \| \sum_{i=1}^n X_i^2 \|^{1/2} \leq \| \sum_{i=1}^n (X_i^2 - EX_i^2) \|^{1/2} + (\sum_{i=1}^n EX_i^2)^{1/2}
\]
\[
\leq L^{1/2}(p/2)^{1/4} \| (\sum_{i=1}^n (X_i^2 - EX_i^2)^2) \|^{1/2} + (\sum_{i=1}^n EX_i^2)^{1/2}
\]
\[
\leq L^{1/2}(p/2)^{1/4} \| (\sum_{i=1}^n X_i^2) \|^{1/2} + L^{1/2}(p/2)^{1/4}(\sum_{i=1}^n (EX_i^2)^2)^{1/4} + (\sum_{i=1}^n EX_i^2)^{1/2}
\]
\[
\leq 2 \max[L^{1/2}(p/2)^{1/4} \| (\sum_{i=1}^n X_i^2) \|^{1/2}, (L^{1/2}(p/2)^{1/4} + 1)(\sum_{i=1}^n EX_i^2)^{1/2}]
\]
For \(p > 4\) we write 4 = \(\theta p + (1 - \theta)2\) i.e.; \(\theta = 2/(p - 2).\) Then, by Holder’s inequality
\[
\sum_{i=1}^n X_i^2 \leq (\sum_{i=1}^n |X_i|^p)^{p/(p - 2)}(\sum_{i=1}^n X_i^2)^{1-p/2}
\]
Hence
\[
E(\sum_{i=1}^n X_i^2)^{p/4} \leq E[(\sum_{i=1}^n |X_i|^p)^{p/(p - 2)}(\sum_{i=1}^n X_i^2)^{1-p/2}]
\]
\[
\leq (E \sum_{i=1}^n |X_i|^p)^{p/(p - 2)}(E(\sum_{i=1}^n X_i^2)^{p/2}(1-p/2))
\]
Thus,
\[
\| \sum_{j=1}^{n} X_j^{1/2} \|_{p}^{1/2} \leq (E \sum_{j=1}^{n} |X_j|^{p})^{1/4} (E(\sum_{j=1}^{n} X_j^2)^{p/2} )^{(1-\theta)/2p}.
\]
If the maximum in (4) is attained in the first term then
\[
\| \sum_{j=1}^{n} X_j^{1/2} \|_{p}^{(1+\theta)/2} \leq 2L^{1/2} (p/2)^{1/4} (E \sum_{j=1}^{n} |X_j|^{p})^{1/4}
\]
and
\[
\| \sum_{j=1}^{n} X_j^{1/2} \|_{p} \leq (2L^{1/2} (p/2)^{1/4})^{2(p-2)/p} (E \sum_{j=1}^{n} |X_j|^{p})^{1/p}.
\]
Thus, dropping the last assumption
\[
\| \sum_{j=1}^{n} X_j^{1/2} \|_{p} \leq \max \{ (2L^{1/2} (p/2)^{1/4})^{2(p-2)/p} (E \sum_{j=1}^{n} |X_j|^{p})^{1/p}, 2L^{1/2} (p/2)^{1/4} + 1 \} (\sum_{j=1}^{n} EX_j^{1/2})
\]
and
\[
\| \sum_{j=1}^{n} X_j \|_{p} \leq p \max \{ K(2L^{1/2})^{2(p-2)/p} (\sum_{j=1}^{n} |X_j|^{p})^{1/p}, 2Kp^{1/2} (\sum_{j=1}^{n} |X_j|^{1/2})^{1/2} \}
\]
\[
\leq 2KLp \max \{ (\sum_{j=1}^{n} |X_j|^{p})^{1/p}, (\sum_{j=1}^{n} |X_j|^{1/2})^{1/2} \}, \quad \text{since} \quad K, L \geq 1.
\]
If \( p < 4 \),
\[
E(\sum_{j=1}^{n} X_j^{1/2})^{p/4} \leq E \sum_{j=1}^{n} |X_j|^{p}
\]
and we get
\[
\| \sum_{j=1}^{n} X_j \|_{p} \leq 2K(p/2)^{1/2} \max \{ L^{1/2} (p/2)^{1/4} (\sum_{j=1}^{n} |X_j|^{p})^{1/p}, (L^{1/2} (p/2)^{1/4} + 1) (\sum_{j=1}^{n} |X_j|^{1/2})^{1/2} \}.
\]

**Remark.** A similar result may be obtained if we replace the assumption
\[
X_1^2 - EX_1^2, \ldots, X_n^2 - EX_n^2 \quad L\text{-unconditional}
\]
by
\[
X_1^2 - X_1'^2, \ldots, X_n^2 - X_n'^2 \quad L\text{-unconditional},
\]
where \( \{X_1', \ldots, X_n'\} \) is an independent copy of \( \{X_1, \ldots, X_n\} \).

**References**


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