The Annals of Probability
2000, Vol. 28, No. 2, 851–862

AN IMPROVEMENT OF HOFFMANN–JØRGENSEN'S INEQUALITY

BY MICHAEL J. KLASS[1] AND KRZYSZTOF NOWICKI[2]

University of California, Berkeley and Lund University

Let B be a Banach space and $\mathscr{F}$ any family of bounded linear functionals on B of norm at most one. For $x \in B$ set $\|x\| = \sup_{\Lambda\in\mathscr{F}} \Lambda(x)$ ($\|\cdot\|$ is at least a seminorm on B). We give probability estimates for the tail probability of $S_n^* = \max_{1\le k\le n} \|\sum_{j=1}^k X_j\|$ where $\{X_i\}_{i=1}^n$ are independent symmetric Banach space valued random elements. Our method is based on approximating the probability that S_n^* exceeds a threshold defined in terms of $\sum_{j=1}^k Y^{(j)}$, where $Y^{(r)}$ denotes the rth largest term of $\{\|X_i\|\}_{i=1}^n$. Using these tail estimates, essentially all the known results concerning the order of magnitude or finiteness of quantities such as $E\Phi(\|S_n\|)$ and $E\Phi(S_n^*)$ follow (for any fixed $1 \le n \le \infty$). Included in this paper are uniform $\mathscr{L}^p$ bounds of S_n^* which are within a factor of 4 for all $p \ge 1$ and within a factor of 2 in the limit as $p \to \infty$.

1. Introduction. Let $X_1, X_2, \dots$ be independent symmetric random elements taking values in a Banach space $(B, \|\cdot\|)$. Suppose that $\|X_j\|_\infty \le 1$ for all $j \ge 1$. Let $S_n = \sum_{j=1}^n X_j$ and $S_n^* = \max_{1\le k\le n} \|\sum_{j=1}^k X_j\|$. Hoffmann-Jørgensen (1974) introduced a technique by which one can prove that

$$P(S_n^* \ge 2a+1) \le 2P^2(S_n^* \ge a).$$

If this technique is iterated one may show that, for any integer $k \ge 1$,

$$P(S_n^* \ge ka+k-1) \le 2^{k-1}P^k(S_n^* \ge a).$$

As a means of establishing that tail probabilities of S_n^* decrease at least geometrically fast if the summands are uniformly bounded symmetric independent variates, his approach is both elegant and sufficient. Needless to say, questions have since arisen which require more precise information on the rate of decay of this tail probability.

Exponential bounds using martingale methods were introduced by Yurinskii (1974) and further developed by de Acosta (1980); de Acosta's bounds require knowledge of the order of magnitude of $\sum_{j=1}^n E\|X_j\|^2$. It is possible that Talagrand (1988, 1989) introduced his isoperimetric methods to repair this deficiency.

In this paper we return to Hoffmann-Jørgensen's approach in the attempt to find accurate and yet straightforward tail probability estimates. His idea

Received May 1998; revised October 1999.
[1]Supported in part by NSF Grants DMS-96-26236 and DMS-99-72417.
[2]Supported by Swedish Council for Research in the Humanities and Social Sciences (HSFR).
AMS 1991 *subject classifications*. Primary 60E15, 60G50.
Key words and phrases. Tail probability inequalities, Hoffmann–Jørgensen's inequality, exponential inequalities, Banach space valued random variables.

can continue to bear fruit since it has not yet incorporated the fact that as the sums reach higher and higher levels, more and more summands have been used, and so ever fewer remain to assist in future growth. By keeping track of the probabilities of a large family of events whose interactions collaborate to produce an S_n^* which exceeds a threshold, we can reconfigure that information to upper-bound the probability of exceeding higher thresholds.

We establish the following theorem.

THEOREM 1.1. *Let $X_1, X_2, \ldots, X_n$ be independent random elements taking values in a Banach space $(B, \|\cdot\|)$ and suppose that $P(S_n^* \geq 1) \equiv \lambda < 1$, where*

$$S_n^* = \max_{1 \leq k \leq n} \|X_1 + \cdots + X_k\|.$$

Let $Y^{(1)} = \max_{1 \leq j \leq n} \|X_j\|$ and let

$$\tau_1 = \text{1st } 1 \leq j \leq n \colon \|X_j\| = Y^{(1)}.$$

Having defined $(Y^{(1)}, \tau_1), (Y^{(2)}, \tau_2), \ldots, (Y^{(k-1)}, \tau_{k-1})$, let

$$Y^{(k)} = \max_{1 \leq j \leq n,\ j \notin \{\tau_1, \ldots, \tau_{k-1}\}} \|X_j\|$$

and

$$\tau_k = \text{1st} \colon 1 \leq j \leq n,\ j \notin \{\tau_1, \ldots, \tau_{k-1}\} \text{ and } \|X_{\tau_k}\| = Y^{(k)}.$$

Then

(a) *If the X_j are nonnegative,*

$$P\left(S_n \geq k + \sum_{j=1}^{k-1} Y^{(j)}\right) \leq \frac{1}{k!}\left[n(1-(1-\lambda)^{1/n})\right]^k \leq \frac{1}{k!}\left(\ln \frac{1}{1-\lambda}\right)^k. \tag{1.1}$$

(b) *If the X_j are symmetric,*

$$P\left(S_n^* \geq k + \sum_{j=1}^{k-1} Y^{(j)}\right) \leq \frac{1}{2k!}\left[2n(1-(1-\lambda)^{1/n})\right]^k \leq \frac{2^{k-1}}{k!}\left(\ln \frac{1}{1-\lambda}\right)^k. \tag{1.2}$$

REMARK 1.1. The idea of using quantities such as $Y^{(j)}$ as part of the exceedence level is due (we believe) to Talagrand (1989). Montgomery-Smith (1990) made independent use of a related idea.

REMARK 1.2. These results apply not only to norms, but to the supremum of any fixed family of bounded linear functionals on B. Thus, on the real line, letting $\mathscr{F} = \{\lambda_1, \lambda_0\}$ where $\lambda_1(x) \equiv x$ and $\lambda_0(x) \equiv 0$, and setting $\|x\| \equiv \max_{i=0,1} \lambda_i(x) = x^+$, the results apply to $\max_{1 \leq k \leq n} S_k^+$.

2. Proof of Theorem 1.1. For each $0 \le i < j \le n$ let $S^*_{(i,j]} = \max_{i<k\le j} \|X_{i+1} + \cdots + X_k\|$ and

$$A_{ij} = \{S^*_{(i,j-1]} < 1, S^*_{(i,j]} \ge 1\}.$$

For $i = 0$ let $S^*_{(0,j]} = S^*_j$ and $A_{0j} = A_j$. Let $T_0 = 0$. Having defined $T_0, T_1, T_2, \dots, T_i$, let

$$T_{i+1} = \begin{cases} \text{1st } j \in (T_i, n]\colon A_{T_i\, j} & \text{occurs,} \\ \infty, & \text{if no such } j \text{ exists.} \end{cases}$$

Then let $L_n + 1 = \text{1st } 1 \le i \le n+1\colon T_i = \infty$. By set inclusion,

$$P\left(S^*_n \ge k + \sum_{j=1}^{k-1} Y^{(j)}\right) \le P(L_n \ge k). \tag{2.1}$$

For $0 \le i < j \le n$, let $q_{ij} = P(A_{ij})$, and put $q_{ij} = 0$ for $j \le i$. The $(n+1)\times(n+1)$ matrix $Q = (q_{ij})$ contains all the information necessary to compute the exact distribution of L_n. In fact, $P(L_n \ge k)$ equals the sum of the zeroth (top) row entries in the matrix Q^k.

Computing $P(L_n \ge k)$ directly from its probabilistic definition,

$$\begin{aligned} P(L_n \ge k) = P(T_k \le n) &= \sum_{1 \le i_1 < \cdots < i_k \le n} P(T_1 = i_1, \dots, T_k = i_k) \\ &= \sum_{i \le i_1 < \cdots < i_k \le n} q_{0i_1} q_{i_1 i_2} \dots q_{i_{k-1} i_k} \quad \text{(by independence).} \end{aligned} \tag{2.2}$$

To upperbound $P(L_n \ge k)$ economically, we need to find some means of replacing the $(n(n+1)/2)$ unknowns q_{ij} which comprise it by quantities more obviously restricted by our single constraint $\sum_{j=1}^n q_{0j} = \lambda$.

To this end we record the following relationships between the ith row sums of the initial segments of the q_{ij} and the corresponding interval row sum for rows $0 \le i_0 < i$. Let

$$\gamma = \begin{cases} 1, & \text{in nonnegative case,} \\ 2, & \text{in symmetric case.} \end{cases}$$

Then, for any $0 \le i_0 < i < j \le n$,

$$\begin{gathered} \sum_{m=i+1}^{j} q_{im} \le \frac{\gamma}{1 - \lambda_{i_0 i}} \sum_{m=i+1}^{j} q_{i_0 m}, \\ \text{(or equivalently)} \\ P(S^*_{(i,j)} \ge 1) \le \frac{\gamma}{1 - \lambda_{i_0 i}} P(S^*_{(i_0, i]} < 1, S^*_{(i_0, j]} \ge 1), \end{gathered} \tag{2.3}$$

where $\lambda_{i_0 i} = q_{i_0 1} + \cdots + q_{i_0 i} (= q_{i_0 i_0 + 1} + \cdots + q_{i_0 i})$.

We only need to prove (2.3) for $i_0 = 0$. Fix any $0 \le i < j \le n$, let $\lambda_{0i} = \lambda_i$ and let

$$\tau = \begin{cases} \text{1st} & m \in (i, j]: S^*_{(i,m]} \ge 1, \\ \infty, & \text{if no such } m \le j \text{ exists.} \end{cases}$$

Set $S_{(i,j]} = \sum_{i<k\le j} X_k$ and $S_j = S_{(0,j]}$. In the *nonnegative* case,

$$\begin{aligned}
&(1-\lambda_i)P(S^*_{(i,j]} \ge 1) \\
&\quad = P(S^*_i < 1) \sum_{k=i+1}^{j} P(\tau = k) \\
&\quad = \sum_{k=i+1}^{j} P(0 \le S_i < 1, \tau = k) \text{ (by independence and since } S_i = S^*_i \ge 0) \\
&\quad = P(0 \le S_i < 1, \tau < \infty) \le P(0 \le S_i < 1,\ S_j \ge 1) = P(S^*_i < 1,\ S^*_j \ge 1)
\end{aligned}$$

which proves the nonnegative part of (2.3).

As for the symmetric case, first observe that for any x, y in B,

$$\max\{\|x+y\|, \|x-y\|\} \ge \|x\|. \tag{2.4}$$

Then

$$\begin{aligned}
&\left(\frac{1-\lambda_i}{2}\right)P(S^*_{(i,j]} \ge 1) \\
&\quad = \frac{1}{2}P(S^*_i < 1) \sum_{k=i+1}^{j} P(\tau = k) \\
&\quad = \frac{1}{2}\sum_{k=i+1}^{j} P(S^*_i < 1, \tau = k) \text{ (by independence)} \\
&\quad = \frac{1}{2}\sum_{k=i+1}^{j} P(S^*_i < 1, \tau = k, \{\|(S_k - S_i) + S_i\| \ge \|S_k - S_i\|\} \\
&\qquad\quad \cup \{\|(S_k - S_i) - S_i\| \ge \|S_k - S_i\|\}) \quad \text{[by (2.4)]} \\
&\quad \le \frac{1}{2}\sum_{k=i+1}^{j} P(S^*_i < 1, \tau = k, \|(S_k - S_i) + S_i\| \ge \|S_k - S_i\|) \\
&\qquad + \frac{1}{2}\sum_{k=i+1}^{j} P\big(S^*_i < 1, \tau = k, \|(S_k - S_i) - S_i\| \ge \|S_k - S_i\|\big) \\
&\quad = \sum_{k=i+1}^{j} P(S^*_i < 1, \tau = k, \|(S_k - S_i) + S_i\| \ge \|S_k - S_i\|) \text{ (by symmetry)} \\
&\quad \le \sum_{k=i+1}^{j} P(S^*_i < 1, \|S_k\| \ge 1, \tau = k) \le P(S^*_i < 1, S^*_j \ge 1),
\end{aligned}$$

which proves (2.3) in the symmetric case.

Put $q_m \equiv q_{0m}$. Upper-bounding,

$$\begin{aligned}
&P(L_n \geq k)\\
&\quad = \sum_{1\leq i_1<i_2<\cdots<i_{k-1}} q_{0i_1}q_{i_1i_2}\cdots q_{i_{k-2}i_{k-1}} \sum_{i_{k-1}<i_k\leq n} q_{i_{k-1}i_k}\\
&\quad \leq \sum_{1<i_1<i_2<\cdots<i_{k-1}} q_{1i_1}q_{i_1i_2}\cdots q_{i_{k-2}i_{k-1}}\gamma \sum_{i_{k-1}<i_k\leq n} \frac{q_{i_k}}{1-\lambda_{i_k-1}} \text{ [by (2.3)]}\\
&\quad \leq \gamma \sum_{1<i_1<i_2<\cdots<i_{k-1}} q_{0i_1}q_{i_1i_2}\cdots q_{i_{k-2}i_{k-1}} \sum_{i_{k-1}<i_k\leq n} \frac{q_{i_k}}{1-\lambda_{i_k-1}} \text{ (since } \lambda_j \text{ increases).}
\end{aligned}$$

Similarly if we freeze all i_j, for $1 \leq j \leq k$ except for $j = k-1$, the crucial factor which must be upper-bounded is

$$\begin{aligned}
\sum_{\{i_{k-1}:\ i_{k-2}<i_{k-1}<i_k\}} q_{i_{k-2}i_{k-1}} &\leq \gamma \sum_{\{i_{k-1}:\ i_{k-2}<i_{k-1}<i_k\}} \frac{q_{i_{k-1}}}{1-\lambda_{i_{k-2}}}\\
&\leq \gamma \sum_{\{i_{k-1}:\ i_{k-2}<i_{k-1}<i_k\}} \frac{q_{i_{k-1}}}{1-\lambda_{i_{k-1}-1}}.
\end{aligned}$$

Iterating this we obtain

$$\begin{aligned}
P(L_n \geq k) &\leq \gamma^{k-1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} q_{i_1} \prod_{j=2}^{k} \frac{q_{i_j}}{1-\lambda_{i_j-1}}\\
&\leq \frac{\gamma^{k-1}}{k!} \sum_{1\leq i_1, i_2, \ldots, i_k\leq n} \prod_{j=1}^{n} \frac{q_{i_j}}{1-\lambda_{i_j-1}}\\
&= \frac{\gamma^{k-1}}{k!}\left(\sum_{i=1}^{n} \frac{q_i}{1-\lambda_{i-1}}\right)^k \text{ (since } \lambda_0 = 0)\\
&= \frac{\gamma^{k-1}}{k!}\left(\sum_{i=1}^{n} \frac{\lambda_i-\lambda_{i-1}}{1-\lambda_{i-1}}\right)^k \equiv \frac{\gamma^{k-1}}{k!}\big(g(\lambda_1, \ldots, \lambda_n)\big)^k.
\end{aligned} \tag{2.5}$$

By the continuity of $g(\cdot)$ and compactness, there exists $0 \leq \lambda_1^* \leq \cdots \leq \lambda_n^* \leq \lambda$ such that

$$g(\lambda_1^*, \ldots, \lambda_n^*) = \sup_{0\leq y_1\leq\cdots\leq y_n\leq\lambda} g(y_1, \ldots, y_n). \tag{2.6}$$

Clearly, $\lambda_n^* = \lambda$. To identify $\lambda_1^*, \ldots, \lambda_{n-1}^*$ fix any $1 \leq j_0 < n$. Then

$$\begin{aligned}
g(\lambda_1^*, \ldots, \lambda_n^*) = &\sum_{\{i:1\leq i\leq n,\, i\notin\{j_0,\, j_0+1\}\}} \frac{\lambda_i^*-\lambda_{i-1}^*}{1-\lambda_{i-1}^*}\\
&+ \sup_{\lambda_{j_0-1}^*\leq y\leq\lambda_{j_0+1}^*} \left(\frac{y-\lambda_{j_0-1}^*}{1-\lambda_{j_0-1}^*} + \frac{\lambda_{j_0+1}^*-y}{1-y}\right)\\
\equiv\ & g_{j_0}(\lambda_1^*, \ldots, \lambda_n^*) + \sup_{\lambda_{j_0-1}^*\leq y\leq\lambda_{j_0+1}^*} h\big(\lambda_{j_0-1}^*, \lambda_{j_0+1}^*, y\big).
\end{aligned} \tag{2.7}$$

Since, for $0 \leq a \leq b < 1$,

$$h(a, b, y) = \frac{y-a}{1-a} - \frac{1-b}{1-y} + 1$$

is clearly concave in y, it assumes a maximum at $y = 1 - \sqrt{(1-a)(1-b)}$. Hence

$$1 - \lambda_{j_0}^* = \sqrt{\left(1 - \lambda_{j_0-1}^*\right)\left(1 - \lambda_{j_0+1}^*\right)},$$

which indeed places $\lambda_{j_0}^*$ between $\lambda_{j_0-1}^*$ and $\lambda_{j_0+1}^*$. Equivalently, we have just learned that

$$r_{j_0} \equiv \frac{1 - \lambda_{j_0}^*}{1 - \lambda_{j_0-1}^*} = \frac{1 - \lambda_{j_0+1}^*}{1 - \lambda_{j_0}^*} \equiv r_{j_0+1} \tag{2.8}$$

for $1 \le j_0 \le n$. Therefore, r_j is some constant r^*. Since $\lambda_0^* = 0$,

$$1 - \lambda_j^* = \prod_{i=1}^{j} \frac{1 - \lambda_i^*}{1 - \lambda_{i-1}^*} = (r^*)^j$$

for $1 \le j \le n$. Putting $j = n$ we obtain

$$r^* = (1 - \lambda)^{1/n}. \tag{2.9}$$

Hence,

$$g\left(\lambda_1^*, \dots, \lambda_n^*\right) = \sum_{i=1}^{n} \frac{(r^*)^{i-1} - (r^*)^i}{(r^*)^{i-1}} = n(1 - r^*) = n\left(1 - (1-\lambda)^{1/n}\right).$$

By the r.h.s. of (2.6), it is obvious that $g(\lambda_1^*, \dots, \lambda_n^*)$ is nondecreasing in n. Letting $n \to \infty$ we may conclude that

$$n\left(1 - (1-\lambda)^{1/n}\right) \le -\ln(1-\lambda). \qquad \square$$

We would like to formally compare probability bounds for random variables having independent nonnegative summands with those having independent symmetric summands which take values in an arbitrary Banach space. For real-valued random variables it is harder for sums of symmetric variates to exceed higher and higher levels than for sums of nonnegative variates. One expects that the same feature carries over to Banach spaces. If so, any universal upper bound of the l.h.s. of (1.1) should also upper-bound the l.h.s. of (1.2) despite the fact that these inequalities are to hold relative to the value of λ. In particular, this leads us to suspect that (1.2) holds without the factor of 2^{k-1} in its r.h.s. Paradoxically, all we can establish (see Theorem 2.1 below) is the reverse bound, namely, that the upper bounds in the nonnegative case cannot in general exceed those for the symmetric case in Banach space.

Let

$$\mathscr{B}_k^+(\lambda) = \sup \sup_{n \ge 1} P\left(S_n \ge k + \sum_{j=1}^{k-1} Y^{(j)}\right), \tag{2.10}$$

where the leftmost supremum is taken over all independent nonnegative random variables $X_1, \dots, X_n$ such that $P(S_n \ge 1) \le \lambda$ and

$$\mathscr{B}_k(\lambda) = \sup_{(B, \|\cdot\|)} \sup \sup_{n \ge 1} P\left(S_n^* \ge k + \sum_{j=1}^{k-1} Y^{(j)}\right), \tag{2.11}$$

where the middle supremum is taken over all independent symmetric random elements $X_1, \dots, X_n$ taking values in the Banach space $(B, \|\cdot\|)$ and satisfying $P(S_n^* \geq 1) \leq \lambda$. We conjecture that

$$\mathscr{R}_k^+(\lambda) = \mathscr{B}_k^+(\lambda) \quad \text{for } k \geq 1. \tag{2.12}$$

Somewhat suprisingly, all we can prove is the reverse of the inequality that seems most natural.

THEOREM 2.1. *Under the above conditions,*

$$\mathscr{R}_k^+(\lambda) \leq \mathscr{B}_k(\lambda) \quad \textit{for all } 0 < \lambda < 1 \textit{ and } k \geq 1. \tag{2.13}$$

PROOF. For $k = 1$ there is nothing to prove. Hence we assume that $k \geq 2$. Our proof is trivial. To obtain a contradiction, suppose there exists an $n \geq 1, k \geq 2, 0 < \lambda < 1, \varepsilon > 0$, and independent nonnegative random variables $X_1, \dots, X_n$ such that $P(\sum_{j=1}^n X_j \geq 1) \leq \lambda$ and $P(\sum_{j=1}^n X_j \geq k + \sum_{j=1}^{k-1} Y^{(j)}) \geq \varepsilon + \mathscr{B}_k(\lambda)$. Let $\{\varepsilon_{jk}\}$ be i.i.d. Rademacher random variables, independent of $\{X_j\}$. Let $\overline{X}_j = (X_j\varepsilon_{j1}, X_j\varepsilon_{j2}, X_j\varepsilon_{j3}, \dots)$. The $\{\overline{X}_j\}$ are independent symmetric random elements taking values in the Banach space l^∞. With probability 1, $\|\overline{S}_n^*\|_\infty = \sum_{j=1}^n X_j$. Therefore,

$$P(\|\overline{S}_n^*\|_\infty \geq 1) = P\biggl(\sum_{j=1}^n X_j \geq 1\biggr) \leq \lambda.$$

Moreover, using the obvious definitions of $Y^{(j)}$ and $\overline{Y}^{(j)}$,

$$\mathscr{B}_k(\lambda) \geq P\biggl(\|\overline{S}_n^*\|_\infty \geq k + \sum_{j=1}^{k-1} \overline{Y}_j^{(j)}\biggr) = P\biggl(\sum_{j=1}^n X_j \geq k + \sum_{j=1}^{k-1} Y^{(j)}\biggr),$$

which gives a contradiction. $\square$

3. Application to $\mathscr{L}^p$ norms of sums of symmetric or nonnegative variates. Theorem 1.1 shows that the maximum of normed partial sums of independent, symmetric or nonnegative suitably truncated Banach space-valued random elements has a rapidly decaying tail probability. Applying this result leads to uniformly good $\mathscr{L}^p$ bounds, as we now illustrate.

THEOREM 3.1. *Let $X_1, X_2, \dots, X_n$ be independent symmetric random elements taking values in a Banach space $(B, \|\cdot\|)$ and set $\gamma = 2$; or else let $X_1, X_2, \dots, X_n$ be independent nonnegative real-valued random variates and set $\gamma = 1$. Let $S_n^* = \max_{1 \leq k \leq n} \|\sum_{j=1}^k X_j\|$. Then, for any fixed $0 < p \leq 1$,*

$$\begin{aligned} &2^{-1-(\gamma-1)/p}\Bigl[\max\bigl\{\tfrac{1}{2}(q_p')^p, \exp(-2^{-p})(t_p)^p\bigr\}\Bigr]^{1/p} \\ &\qquad \leq \|S_n^*\|_p \leq v_p'\bigl(g_\gamma(p) + 2^{-p}\bigr)^{1/p} \end{aligned} \tag{3.1}$$

and for any fixed $p > 1$,

$$\frac{v'_p}{2^{1+1/p}} \le \|S_n^*\|_p \le v'_p\Big((g_\gamma(p))^{1/p} + \frac{1}{2}\|1 + \mathscr{N}_{2^{-p}}\|_p\Big) \equiv v'_p h_\gamma(p), \tag{3.2}$$

where $\mathscr{N}_\lambda \sim$ Poisson(λ),

$$t_p = \sup\left\{t \ge 0: \sum_{j=1}^n E\|X_j\|^p I(\|X_j\| > t) \ge \left(\frac{t}{2}\right)^p\right\},$$

$$q'_p = \sup\left\{q \ge 0: P\left(\max_{1\le k\le n}\left\|\sum_{j=1}^k X_j I(\|X_j\| \le t_p)\right\| > q\right) \ge 2^{-p-1}\right\},$$

$$v'_p = \max\{q'_p, t_p\}$$

and

$$g_\gamma(p) = 1 + \frac{1}{\gamma}\sum_{k=1}^\infty \frac{(2k+1)^p - (2k-1)^p}{k!}\big(-\gamma\ln(1-2^{-p-1})\big)^k.$$

Moreover,

$$\frac{1}{2} \le \liminf_{p\to\infty} \frac{\|S_n^*\|_p}{v'_p} \le \limsup_{p\to\infty} \frac{\|S_n^*\|_p}{v'_p} \le 2. \tag{3.3}$$

PROOF. Set

$$X'_j = X_j I(\|X_j\| \le t_p), \qquad X''_j = X_j I(\|X_j\| > t_p),$$

$$S^*_{n,1} = \max_{1\le k\le n}\left\|\sum_{j=1}^k X'_j\right\|, \qquad S^*_{n,2} = \max_{1\le k\le n}\left\|\sum_{j=1}^k X''_j\right\|.$$

To avoid difficulties caused by atoms located at t_p, introduce zero–one valued random variables $\delta_1, \dots, \delta_n$ independent of each other and of $\{X_i\}$ such that

$$\sum_{j=1}^n E\|X'''_j\|^p = \left(\frac{t_p}{2}\right)^p,$$

where

$$X'''_j = X_j I(\|X_j\| > t_p) + \delta_j X_j I(\|X_j\| = t_p).$$

This is possible since

$$\sum_{j=1}^n E\|X_j\|^p I(\|X_j\| > t_p) \le \left(\frac{t_p}{2}\right)^p \le \sum_{j=1}^n E\|X_j\|^p I(\|X_j\| \ge t_p).$$

Let

$$\gamma_* = \begin{cases} 1, & \text{if } p \ge 1 \text{ or if the } \{X_j\} \text{ are nonnegative and } 0 < p < 1,\\ \frac{1}{2}, & \text{if } 0 < p < 1 \text{ and the } \{X_j\} \text{ are symmetric.}\end{cases}$$

We claim that

$$E\|S_n^*\|^p \geq \gamma_* \max\left\{E\|S_{n,1}^*\|^p, E\left\|\sum_{j=1}^n X_j'''\right\|^p\right\}. \tag{3.4}$$

Equation (3.4) follows trivially for the X_j which are nonnegative. Hence assume that the X_j are symmetric. If $p \geq 1$ the result follows by an appropriate conditional application of Jensen's inequality. If $0 < p < 1$, it follows because certain variables are conditionally symmetric.

Lower-bounding the first term,

$$E\|S_{n,1}^*\|^p \geq E(q_p')^p I(\|S_{n,1}^*\| \geq q_p') \geq (q_p')^p 2^{-p-1}.$$

To lower-bound the other quantity, first reorder the indices if necessary, to obtain

$$P(\|X_n'''\| \neq 0) = \max_{1\leq j\leq n} P(\|X_j'''\| \neq 0).$$

Then let

$$\tau = \begin{cases} 1\text{st} & j \in [1,n]\colon X_j''' \neq 0, \\ \infty, & \text{if no such } j \text{ exists.} \end{cases}$$

By construction of t_p, $\sum_{j=1}^n P(\|X_j'''\| \neq 0) \leq 2^{-p}$ and consequently,

$$\begin{aligned} E\left\|\sum_{j=1}^n X_j'''\right\|^p &\geq E\|X_\tau'''\|^p I(\tau \leq n) \\ &= \sum_{j=1}^n E\|X_j'''\|^p I(\tau \geq j) = \sum_{j=1}^n E\|X_j'''\|^p P(\tau \geq j) \\ &\geq P(\tau \geq n) \sum_{j=1}^n E\|X_j'''\|^p = P(\tau \geq n)\left(\frac{t_p}{2}\right)^p \\ &\geq \left(1 - \frac{1}{n2^p}\right)^{n-1}\left(\frac{t_p}{2}\right)^p \text{ [by Klass (1981)]} \\ &> \exp(-2^{-p})\left(\frac{t_p}{2}\right)^p. \end{aligned}$$

Hence the lower bounds in (3.1) and (3.2) hold.

The upper bounds are based on expressing $\|S_n^*\|$ in terms of $\|S_{n,1}^*\|_p$ and $\|S_{n,2}^*\|_p$ and then upper-bounding each of these two. For $0 < p \leq 1$, $\|x+y\|^p \leq \|x\|^p + \|y\|^p$. Hence,

$$E \max_{1\leq k\leq n}\left\|\sum_{j=1}^k X_j\right\|^p \leq E \max_{1\leq k\leq n}\left\|\sum_{j=1}^k X_j'\right\|^p + E \max_{1\leq k\leq n}\left\|\sum_{j=1}^k X_j''\right\|^p.$$

Trivially,

$$E \max_{1\le k\le n} \left\| \sum_{j=1}^{k} X_j'' \right\|^p \le \sum_{j=1}^{n} E\|X_j''\|^p \le \left(\frac{t_p}{2}\right)^p.$$

For $p > 1$, by Minkowski's inequality,

$$\|S_n^*\|_p \le \|S_{n,1}^*\|_p + \|S_{n,2}^*\|_p.$$

By Lemmas 1.1 and 3.3 of Klass (1981) we have, for $p > 1$,

$$\|S_{n,2}^*\|_p^p \le E(1+\mathscr{N}_{2^{-p}})^p \sum_{j=1}^{n} E\|X_j''\|^p \le 2^{-p} E\big(1+\mathscr{N}_{2^{-p}}\big)^p \big(t_p\big)^p.$$

The upper bound of $\|S_{n,1}^*\|_p$ for all $p > 0$ is next on the agenda. Since, by construction,

$$P(S_{n,1}^* > q_p') \le 2^{-p-1},$$

using (1.2) we conclude that

$$P\big(S_{n,1}^* > (2k-1)v_p'\big) \le P\big(S_{n,1}^* > kq_p' + (k-1)t_p\big) \le \gamma^{k-1}\frac{1}{k!}\big(\ln(1-2^{-p-1})\big)^k.$$

Therefore,

$$\begin{aligned}
E(S_{n,1}^*)^p &= \int_0^\infty P\big((S_{n,1}^*)^p \ge y\big)\,dy \\
&\le \int_0^{(v_p')^p} P\big(S_{n,1}^* \ge y^{1/p}\big)\,dy + \sum_{k=1}^{\infty} \int_{((2k-1)v_p')^p}^{((2k+1)v_p')^p} P\big(S_{n,1}^* \ge y^{1/p}\big)\,dy \\
&\le (v_p')^p + (v_p')^p \sum_{k=1}^{\infty} \big[(2k+1)^p - (2k-1)^p\big] \frac{\gamma^{k-1}}{k!} \big(-\ln(1-2^{-p-1})\big)^k \\
&\le (v_p')^p \left(1 + \frac{1}{\gamma} \sum_{k=1}^{\infty} \frac{(2k+1)^p - (2k-1)^p}{k!} \big(-\gamma \ln(1-2^{-p-1})\big)^k\right) \\
&\equiv (v_p')^p g_\gamma(p).
\end{aligned}$$

This completes the r.h.s. of (3.1) and (3.2). Finally, we consider $p \to \infty$. To approximate $(g_\gamma(p))^{1/p}$ as $p \to \infty$, observe that only the $k = 1$ term of the series counts since for all $p \ge 1$ the successive terms in

$$1 + \frac{1}{\gamma} \sum_{k=1}^{\infty} \frac{(2k+1)^p - (2k-1)^p}{k!} \big(-\gamma \ln(1-2^{-p-1})\big)^k$$

decrease by a factor bounded above by 5/6 for all $k \ge 1$, and the $k = 1$ term tends to ∞ as $p \to \infty$. Hence

$$\lim_{p\to\infty} \big(g_\gamma(p)\big)^{1/p} = 1.5.$$

Similarly,

$$\lim_{p\to\infty} \|1+\mathscr{N}_{2^{-p}}\|_p = 1$$

and so $\lim_{p\to\infty} h_\gamma(p) = 2$. $\square$

REMARK 3.1. By direct calculation $h_2(1) = 91/36$ and

$$h_2(2) = \frac{\sqrt{29}}{8} + \sqrt{1 + \frac{512}{49}\ln\frac{8}{7}} = 2.22081.$$

Computer calculations due to Jaimyoung Kwon also show that $h_2(p) \le h_2(1)$ for all $p \ge 1$ and that $h_2(p) \le h_2(2)$ for $p \ge 2$. Clearly, we also have $h_1(p) \le h_2(p)$.

REMARK 3.2. Variants of our approximations could be created by slightly changing the probability in the definition of q'_p or the factor in t_p or by constructing somewhat different quantities altogether. For example, instead of q'_p one could employ

$$q''_p = \sup\left\{q \ge 0 \colon P\Big(\max_{1\le k\le n}\|S_k\| \ge q\Big) \ge 2^{-p-2}\right\}.$$

Alternatively, define

$$\underline{t}_p = \sup\left\{t \ge 0 \colon \sum_{j=1}^n P(\|X_j\| > t) \ge 2^{-p}\right\},$$

and let

$$\underline{q}'_p = \sup\left\{q \ge 0 \colon P\left(\max_{1\le k\le n}\left\|\sum_{j=1}^k X_j I(\|X_j\| \le \underline{t}_p)\right\| \ge q\right) \ge 2^{-p-1}\right\}.$$

Since

$$\begin{aligned}\sum_{j=1}^n E\|X''_j\|^p &\le \sum_{j=1}^n E\|X_j\|^p I\big(\|X_j\| > \underline{t}_p\big)\\ &\le \sum_{j=1}^n E(t_p)^p I(\underline{t}_p < \|X_j\| \le t_p) + \sum_{j=1}^n E\|X_j\|^p I(\|X_j\| > t_p)\\ &\le (t_p)^p \sum_{j=1}^n P(\|X_j\| > \underline{t}_p) + \left(\frac{t_p}{2}\right)^p \le 2\left(\frac{t_p}{2}\right)^p,\end{aligned}$$

one could use $\max\{\underline{q}'_p, t_p\}$ to approximate $\|S_n^*\|_p$.

REMARK 3.3. Previously, the magnitude of $\|S_n\|_p$ has been compared to quantities such as

$$m_{n,1} \equiv \max\left\{\|S_n\|_1, \left|\max_{1\le k\le n}\|X_k\|\right|_p\right\}.$$

It was shown that

$$c(p) \le \frac{\|S_n\|_p}{m_{n,1}} \le C(p),$$

where $c(p)$ is about 1 for all $p \geq 1$ and $C(p)$ is about $p/(1+\ln p)$. It appeared that the norm of S_n could not be approximated uniformly in $p \geq 1$.

In the real-valued case Latała (1997) suprisingly demonstrated the existence of a constructable constant q_p satisfying

$$\sum_{j=1}^{n} \ln E\left(1+\frac{X_j}{q_p}\right)^p = p$$

such that

$$\frac{e-1}{2e^2} \leq \frac{\|S_n\|_p}{(q_p)^p} \leq e$$

for $p \geq 1$ and $X_j \geq 0$ or for $p \geq 2$ and X_j symmetric. By interpreting q_p as deriving from the maximum of two effects, we have been able to find an alternative v'_p to q_p which not only applies to the real-valued case, but extends to Banach space. Theorem 3.1 compares v'_p with $\|S_n^*\|_p$. To relate this to Latała's work, note that $\|S_n\|_p = \|S_n^*\|_p$ in the nonnegative case and $\|S_n\|_p \leq \|S_n^*\|_p \leq 2^{1/p}\|S_n\|_p$ in the symmetric case. Thus, for $p \geq 1$, our results translate into

$$\frac{v'_p}{2^{1+2/p}} \leq \|S_n\|_p \leq h_\gamma(p) v'_p.$$

Latała also showed that his results even extend to $0 < p < 1$, as do ours. The proportional improvement we obtained tends to ∞ as $p \to 0$. We must admit, however, that our quantity q'_p from which v'_p derives is not necessarily very computable, except perhaps in the real-valued or even Hilbert space case.

REFERENCES

DE ACOSTA, A. (1980). Strong exponential integrability of sums of independent B-valued random vectors. *Probab. Math. Statist.* **1** 133–150.

HOFFMANN-JØRGENSEN, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **52** 159–186.

KLASS, M. J. (1981). A method of approximating expectations of functions of sum of independent random variables. *Ann. Probab.* **9** 413–428.

LATAŁA, R. (1997). Estimation of moments of sums of independent real random variables. *Ann. Probab.* **25** 1502–1513.

MONTGOMERY-SMITH, S. (1990). Distributions of Rademacher sums. *Proc. Amer. Math. Soc.* **109** 517–522.

TALAGRAND, M. (1988). An isoperimetric theorem on the cube and the Khintchine–Kahane inequalities. *Proc. Amer. Math. Soc.* **104** 905–909.

TALAGRAND, M. (1989). Isoperimetry and integrability of the sum of independent Banach-space valued random variables. *Ann. Probab.* **17** 1546–1570.

YURINSKII, V. V. (1974). Exponential bounds for large deviations. *Theory Probab. Appl.* **19** 154–155.

UNIVERSITY OF CALIFORNIA
DEPARTMENT OF STATISTICS AND MATHEMATICS
367 EVANS HALL #3860
BERKELEY, CALIFORNIA 94720-3860

LUND UNIVERSITY
DEPARTMENT OF STATISTICS
BOX 743
S-22007 LUND
SWEDEN