The Annals of Probability
1991, Vol. 19, No. 2, 1206–1226

MAJORIZATION, EXPONENTIAL INEQUALITIES AND ALMOST SURE BEHAVIOR OF VECTOR-VALUED RANDOM VARIABLES[1]

BY ERICH BERGER

Universität Göttingen

In this paper we describe a general device that allows us to deduce various kinds of theorems (moment estimates, exponential inequalities, strong law of large numbers, stability results, bounded law of the iterated logarithm) for partial sums of independent vector-valued random variables from related results for partial sums of independent real-valued random variables. The concept of majorization will play a key role in our considerations.

1. Introduction and main results. The random variables (r.v.'s) occurring in this paper will always be assumed to be defined on a common (sufficiently rich) probability space $(\Omega, \mathscr{A}, P)$. $I(A)$ stands for the indicator function of a set A. The distribution of an r.v. X (taking values in an arbitrary measurable space) will be denoted by $P \circ X^{-1}$. Almost sure (a.s.) convergence and convergence in probability of a sequence of r.v.'s will be indicated by $\to_{\text{a.s.}}$ and $\to_P$, respectively.

As usual, $\mathbb{R} := $ set of all real numbers, $\mathbb{R}^+ := \{x \in \mathbb{R}: x > 0\}$ and $\mathbb{N} :=$ set of all positive integers. $\mathbb{R}$ will always be assumed to be equipped with its Borel σ-field.

A real-valued r.v. R is called *Bernoulli* if $P\{R = 1\} = P\{R = -1\} = 1/2$.

It is the aim of this article to describe a general device that allows us to deduce various kinds of theorems (such as exponential inequalities and strong laws of large numbers) for partial sums of independent vector-valued r.v.'s from related results for partial sums of independent real-valued r.v.'s. The natural framework for the methods of this paper is seminormed measurable vector spaces.

DEFINITION 1.1 [cf. Kuelbs and Zinn (1979), page 75]. Let $\mathbb{F}$ be a real vector space, let $\mathscr{F}$ be a σ-field of subsets of $\mathbb{F}$ and let $\|\cdot\|$ be a seminorm on $\mathbb{F}$. The triple $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ is called a *seminormed measurable vector space* if the following conditions are satisfied:

(1.1) addition in $\mathbb{F}$ and scalar multiplication are measurable operations ($\mathbb{F} \times \mathbb{F} \to \mathbb{F}$ and $\mathbb{R} \times \mathbb{F} \to \mathbb{F}$, respectively);

(1.2) the function $x \mapsto \|x\|$, $x \in \mathbb{F}$, is measurable.

Received June 1988; revised May 1990.

[1]This paper was developed in connection with a project partially supported by Deutsche Forschungsgemeinschaft Grant Kr-765/3-1.

AMS 1980 *subject classifications*. Primary 60B12, 60F10, 60F15; secondary 60E15, 60G50.

Key words and phrases. Majorization, moment inequalities, exponential inequalities, strong law of large numbers, bounded law of the iterated logarithm.

A wide class of seminormed measurable vector spaces can be constructed by using the following observation. Let $\mathbb{F}$ be a real vector space, let $\mathscr{F}$ be the σ-field generated by a nonvoid set A of linear functionals on $\mathbb{F}$ and let $B \subset A$ with $B \neq \emptyset$ be such that $\|x\| := \sup\{|g(x)|: g \in B\} < \infty$ for all $x \in \mathbb{F}$. If the function $x \mapsto \|x\|$, $x \in \mathbb{F}$, is $\mathscr{F}$-measurable, then $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ is a seminormed measurable vector space.

The following examples of seminormed measurable vector spaces $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ can be obtained as special cases of the just mentioned construction:

(a) $(\mathbb{F}, \|\cdot\|)$ is a real separable normed linear space, and $\mathscr{F}$ is the Borel σ-field of $\mathbb{F}$ (recall the Itô–Nisio theorem);
(b) $\mathbb{F}$ is the Skorohod space $D[0, 1]$, $\mathscr{F}$ is the σ-field generated by Skorohod's J_1-topology and $\|\cdot\|$ is the supremum norm on $D[0, 1]$ [see, e.g., Billingsley (1968), page 109 ff.].

To state our first theorem, it is useful to introduce the following notion of majorization.

Definition 1.2. Let X and Y be real-valued r.v.'s such that $E|X| < \infty$ and $E|Y| < \infty$. Then X is said to be *majorized* by Y (or $X \prec Y$) if and only if for all convex functions f: $\mathbb{R} \to \mathbb{R}$:

$$Ef(X) \leq Ef(Y). \tag{1.3}$$

By the convexity of f and the integrability of X and Y, the expectations in (1.3) are well defined but may be equal to ∞.

Remark 1.1. An important characterization of majorization goes back to Hardy, Littlewood and Pólya (1929) and Karamata (1932) [cf. Marshall and Olkin (1979), page 449]: Let X and Y be as in Definition 1.2; then $X \prec Y$ if and only if $EX = EY$ and $E(X-t)^+ \leq E(Y-t)^+$ for all $t \in \mathbb{R}$.

Now let $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ be a seminormed measurable vector space, let $n \in \mathbb{N}$, let $X_1, \dots, X_n$ be independent $\mathbb{F}$-valued r.v.'s such that $E\|X_j\| < \infty$ for all $j \in \{1, \dots, n\}$, and write

$$S_n := \sum_{j=1}^n X_j.$$

For $j \in \{1, \dots, n\}$, let

$$V_j := \|X_j\| + E\|X_j\|,$$

let $V' := (V_1', \dots, V_n')$ be an independent copy of $V := (V_1, \dots, V_n)$ and let $R_1, \dots, R_n$ be independent Bernoulli r.v.'s such that $(R_1, \dots, R_n)$ and (V, V') are independent. Also define

$$T_n := \sum_{j=1}^n R_j(V_j + V_j') \quad \text{and} \quad T_n^* := \sum_{j=1}^n R_j\|X_j\|.$$

The key result of this paper is the following theorem.

THEOREM 1.1 (Majorization of vector-valued r.v.'s). *Under the above hypotheses, one has*

$$\|S_n\| - E\|S_n\| \prec T_n \prec 4T_n^*. \tag{1.4}$$

The *proof* of Theorem 1.1 will be carried out in Sections 2 and 3.

We shall now apply Theorem 1.1 to the problem of deriving moment inequalities, exponential estimates and stability results for partial sums of independent vector-valued r.v.'s.

A first consequence of Theorem 1.1 is that moment inequalities for T_n (or T_n^*) such as the upper half of the Marcinkiewicz–Zygmund (1937) inequality or the Rosenthal inequality [Rosenthal (1972), Theorem 1.1] can immediately be translated into moment inequalities for $\|S_n\| - E\|S_n\|$. As a typical example in this direction, we mention de Acosta's (1981) vector analog of the Marcinkiewicz–Zygmund inequality.

THEOREM 1.2 (de Acosta). *The notation is as in Theorem* 1.1. *Let* $p \in [1, \infty)$, *and suppose that* $E\|X_j\|^p < \infty$ *for all* $j \in \{1, \dots, n\}$. *Then there is a constant* $A \in \mathbb{R}^+$ *(depending only on p) such that*

$$E\big|\,\|S_n\| - E\|S_n\|\,\big|^p \le AE\left\{\left(\sum_{j=1}^{n} \|X_j\|^2\right)^{p/2}\right\}. \tag{1.5}$$

In order to establish (1.5), de Acosta required the hard-to-prove martingale analog of the Marcinkiewicz–Zygmund inequality [cf. Burkholder (1973)]. Using Theorem 1.1, the proof of Theorem 1.2 becomes more elementary in that it works with the original result for partial sums of independent real-valued r.v.'s.

Our second application of Theorem 1.1 concerns exponential inequalities for vector-valued r.v.'s.

THEOREM 1.3. *The notation is as in Theorem* 1.1. *If*

$$\alpha := \sup\Big\{t \in \mathbb{R}\colon E\big(\exp(t\|X_j\|)\big) < \infty \quad \text{for all } j \in \{1, \dots, n\}\Big\} > 0, \tag{1.6}$$

then, for $\theta \in \{-1, 1\}$ *and any* $x \in \mathbb{R}^+$,

$$P\big\{\theta(\|S_n\| - E\|S_n\|) \ge x\big\} \le \inf\big\{e^{-tx}E(\exp(tT_n))\colon t \in [0, \alpha)\big\} \tag{1.7}$$

and

$$P\big\{\theta(\|S_n\| - E\|S_n\|) \ge x\big\} \le \inf\big\{e^{-tx}E(\exp(4tT_n^*))\colon 4t \in [0, \alpha)\big\}. \tag{1.8}$$

PROOF. Fix $x \in \mathbb{R}^+$. For $t \in [0, \alpha)$, we have

$$P\big\{\theta(\|S_n\| - E\|S_n\|) \ge x\big\} \le \exp(-tx)E\big(\exp(t\theta(\|S_n\| - E\|S_n\|))\big). \tag{1.9}$$

According to Theorem 1.1,

$$E\big(\exp(s(\|S_n\| - E\|S_n\|))\big) \le E(\exp(sT_n)) \quad \text{for all } s \in (-\alpha, \alpha) \tag{1.10}$$

and

$$(1.11) \qquad E(\exp(sT_n)) \le E(\exp(4sT_n^*)) \quad \text{for all } s \in \mathbb{R} \text{ with } 4|s| < \alpha.$$

Combining (1.9)–(1.11) (and taking into account that the r.v.'s T_n and T_n^* are symmetric), we obtain (1.7) and (1.8). □

Regarded as functions of x, the expressions on the right-hand side of (1.7) and (1.8) are Chernoff functions of partial sums of independent real-valued r.v.'s. This means that any exponential inequality for such partial sums based on estimating the Chernoff function (and this is the usual method for deriving exponential inequalities) can be carried over to an exponential inequality for vector-valued r.v.'s. For a detailed survey of exponential inequalities for real-valued r.v.'s, the reader is referred to Nagaev's (1979) article.

A reduction argument similar to that expressed in Theorem 1.3 was earlier utilized by Yurinskii (1974) who traced the problem of finding exponential inequalities for partial sums of independent vector-valued r.v.'s to a corresponding problem for real-valued martingales. His point of departure is the representation (3.12) of $\|S_n\| - E\|S_n\|$ as a partial sum of martingale differences. [(3.12) will also play a crucial role in our proof of Theorem 1.1.] Following his line of reasoning, several well-known exponential inequalities for real-valued r.v.'s such as Bernstein's inequality [Yurinskii (1974)], the Kolmogorov upper exponential bound [Kuelbs (1977)] and Bennett's inequality [de Acosta (1980)] have been carried over to vector-valued r.v.'s. However, there exist important and interesting exponential inequalities for partial sums of independent real-valued r.v.'s that cannot be extended to the vector space setting by adapting Yurinskii's arguments in a straightforward manner. An example of this type is the exponential estimate underlying Nagaev's (1972) theorem about necessary and sufficient conditions for the strong law of large numbers (SLLN) for independent real-valued r.v.'s. As we shall see, it is the vector analog of this last-mentioned exponential inequality that will turn out to be of decisive importance for proving Theorems 1.4 and 1.5 dealing with stability results for vector-valued r.v.'s.

To formulate Theorems 1.4 and 1.5, we begin by introducing some notation. As above, let $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ be a seminormed measurable vector space. Moreover, let $\{X_j;\ j \in \mathbb{N}\}$ be a sequence of independent $\mathbb{F}$-valued r.v.'s, let $\{R_j;\ j \in \mathbb{N}\}$ be a sequence of independent Bernoulli r.v.'s which is also independent of $\{X_j;\ j \in \mathbb{N}\}$ and let $\{a_j;\ j \in \mathbb{N}\}$ be a nondecreasing sequence in $\mathbb{R}^+$ with $\sup\{a_j\colon j \in \mathbb{N}\} = \infty$. For $n \in \mathbb{N}$, we write $S_n := \sum_{j=1}^n X_j$.

THEOREM 1.4. *The notation is as above. Suppose*

$$(1.12) \qquad a_n^{-1} \sum_{j=1}^{n} R_j \|X_j\| \to_{a.s.} 0.$$

Then

$$(1.13) \qquad a_n^{-1}\|S_n\| \to_{a.s.} 0 \quad \textit{if and only if} \quad a_n^{-1}\|S_n\| \to_P 0.$$

THEOREM 1.5. *The notation is as above. Suppose*

$$\limsup_{n\to\infty} a_n^{-1} \sum_{j=1}^{n} R_j \|X_j\| < \infty \quad a.s. \tag{1.14}$$

Then

$$\limsup_{n\to\infty} a_n^{-1} \|S_n\| < \infty \quad a.s. \tag{1.15}$$

if and only if

(1.16) *the sequence* $\{a_n^{-1}\|S_n\|;\ n \in \mathbb{N}\}$ *is bounded in probability.*

The *proofs* of Theorems 1.4 and 1.5 are deferred to Section 4.

Beginning with Kuelbs' (1977) paper, there have been published various special cases of Theorems 1.4 and 1.5 stating that the conclusion of Theorem 1.4 (Theorem 1.5, resp.) holds if (1.12) [(1.14), resp.] are replaced by certain classical criteria for the SLLN. In this way there have been established vector analogs of the upper half of the Hartman–Wintner (1941) law of the iterated logarithm [Kuelbs (1977)], of Prohorov's (1959) SLLN [Kuelbs and Zinn (1979)], of a Marcinkiewicz–Zygmund (1937) SLLN [de Acosta (1981)] and of iterated logarithm type results by Klass (1976, 1977) [Kuelbs and Zinn (1983), Theorem 3 and Corollaries 1–5]. Compared to these results, Theorems 1.4 and 1.5 assert, roughly speaking, that under "classical" conditions the SLLN in the vector space setting is always equivalent to the weak law.

Before concluding this section, let us add some general remarks on the scope and the limitations of the methods of this paper. An interesting aspect is that our results have a sort of universal character; any theorem of a certain type for partial sums of independent real-valued r.v.'s has in a sense a natural analog for vector-valued r.v.'s. Moreover, disregarding the constants implicitly or explicitly involved, the results also have optimality properties as far as the assumptions cannot in general be weakened by conditions that depend on the r.v.'s X_j only via $\|X_j\|$. On the other hand, it is well known that exact characterizations for the strong law of large numbers and sharp exponential bounds must depend on the r.v.'s X_j in a more complicated way, for example, via weak moments. (In particular, conditions (1.12) [and (1.14)] are not necessary in order that the conclusion of Theorem 1.4 (Theorem 1.5, resp.) be true.) Such results are not accessible by the methods of this paper. For further information in this direction, the reader is referred to Alexander (1984) and Ledoux and Talagrand (1988, 1990).

Though primarily designed to prove Theorems 1.1, 1.4 and 1.5, Sections 2–4 also contain several results of independent interest, namely Theorem 3.1 (a majorization result for real-valued martingales), Theorem 4.2 [a somewhat extended version of Nagaev's (1972) theorem about necessary and sufficient conditions for the SLLN] and Theorem 4.3 (a general stability result for vector-valued r.v.'s).

2. Generation and preservation of majorization. This section contains some lemmas that are required for proving Theorem 3.1 which in turn leads to Theorem 1.1. There is an extensive literature about majorization and related concepts [see, e.g., Marshall and Olkin (1979)], but we shall only need the few facts that are collected below.

Throughout, $\mathscr{C}$ denotes the class of all convex functions $\mathbb{R} \to \mathbb{R}$, and $\mathscr{C}^* \subset \mathscr{C}$ the subclass of all convex functions f with the property that there exist positive real constants a, b such that $|f(x)| \le a + b|x|$ for all $x \in \mathbb{R}$.

In order to avoid trivial complications caused by the fact that the expectations in (1.3) need not be finite, it is often convenient to use the following equivalent condition for X to be majorized by Y.

LEMMA 2.1. *Let X and Y be real-valued r.v.'s with $E|X| < \infty$ and $E|Y| < \infty$. Then $X \prec Y$ if and only if*

$$Eg(X) \le Eg(Y) \quad \text{for all } g \in \mathscr{C}^*. \tag{2.1}$$

PROOF. Let $f \in \mathscr{C}$. We have to show that (2.1) implies that $Ef(X) \le Ef(Y)$. It is not difficult to see that there is a nondecreasing sequence $\{g_n;\ n \in \mathbb{N}\}$ in $\mathscr{C}^*$ such that $f = \lim_{n\to\infty} g_n$ [cf. Roberts and Varberg (1973), Section 11]. By (2.1), $Eg_n(X) \le Eg_n(Y)$ for all $n \in \mathbb{N}$. Since $g_n \le f$ for all $n \in \mathbb{N}$, this entails that $\sup_{n\in\mathbb{N}} Eg_n(X) \le Ef(Y)$, which together with the monotone convergence theorem [note that $E|g_1(X)| < \infty$] leads to the desired result. □

LEMMA 2.2. *Let $\mathscr{A}_0$ be a sub-σ-field of $\mathscr{A}$, and let X and Y be real-valued r.v.'s such that $E|X| < \infty$, $E|Y| < \infty$ and*

$$E(Y|\mathscr{A}_0) = 0 \quad a.s. \tag{2.2}$$

Suppose

$$X \text{ and } Y \text{ are conditionally independent given } \mathscr{A}_0. \tag{2.3}$$

Then

$$E(g(X)|\mathscr{A}_0) \le E(g(X+Y)|\mathscr{A}_0) \quad a.s.\ \text{for any } g \in \mathscr{C}^*.$$

PROOF. Let $\mathscr{A}_1 \subset \mathscr{A}$ be the smallest σ-field containing $\mathscr{A}_0$ and the σ-field generated by X. Then $E(Y|\mathscr{A}_1) = 0$ a.s. [by (2.2) and (2.3)]. It follows that for any function $g \in \mathscr{C}^*$,

$$\begin{aligned} E(g(X)|\mathscr{A}_0) &= E(g(E(X+Y|\mathscr{A}_1))|\mathscr{A}_0) \\ &\le E(E(g(X+Y)|\mathscr{A}_1)|\mathscr{A}_0) \quad \text{(by Jensen's inequality)} \\ &= E(g(X+Y)|\mathscr{A}_0) \quad \text{a.s.} \end{aligned}$$

□

LEMMA 2.3. *Let X and X' be real-valued r.v.'s having the same distribution. Suppose $E|X| < \infty$. Then $X + X' \prec 2X$.*

PROOF. For any function $f \in \mathscr{C}$, we have

$$Ef(X+X') \leq \tfrac{1}{2}Ef(2X) + \tfrac{1}{2}Ef(2X') = Ef(2X). \qquad \square$$

LEMMA 2.4. *Let X be a real-valued r.v. with $E|X| < \infty$. Then*

$$X + EX \prec 2X. \tag{2.4}$$

PROOF. Let X' be an independent copy of X. Then, for any $g \in \mathscr{C}^*$,

$$\begin{aligned} Eg(X+EX) &= Eg(E(X+X'|X)) \\ &\leq Eg(X+X') \quad \text{(by Jensen's inequality)} \\ &\leq Eg(2X) \qquad \text{(by Lemma 2.3)}, \end{aligned}$$

which together with Lemma 2.1 gives (2.4). □

LEMMA 2.5 (Majorization by randomized bounds). *Let $\mathscr{A}_0$ be a sub-σ-field of $\mathscr{A}$, and let X, X', Y, Y' be real-valued r.v.'s. Suppose*:

(2.5) *(X, Y) and (X', Y') are conditionally independent given $\mathscr{A}_0$, and $P((X,Y) \in A|\mathscr{A}_0) = P((X',Y') \in A|\mathscr{A}_0)$ a.s. for any Borel set $A \subset \mathbb{R}^2$;*

(2.6) *(Y, Y') and $\mathscr{A}_0$ are independent;*

(2.7) *$|X| \leq Y$ a.s. and $EY < \infty$.*

Then

$$E(g(X-X')|\mathscr{A}_0) \leq \tfrac{1}{2}Eg(Y+Y') + \tfrac{1}{2}Eg(-(Y+Y')) \quad a.s. \tag{2.8}$$

for any $g \in \mathscr{C}^*$.

PROOF. Let the r.v. $\Gamma\colon \Omega \to \mathbb{R}$ be defined by

$$\Gamma(\omega) := \begin{cases} \dfrac{1}{2} & \text{if } (Y+Y')(\omega) = 0, \\ \dfrac{1}{2} + \dfrac{1}{2}\dfrac{(X-X')(\omega)}{(Y+Y')(\omega)} & \text{otherwise.} \end{cases}$$

Then

$$0 \leq \Gamma \leq 1 \quad \text{a.s.}$$

and

$$X - X' = \Gamma(Y+Y') - (1-\Gamma)(Y+Y') \quad \text{a.s.}$$

This entails that for any function $g \in \mathscr{C}^*$,

$$g(X-X') \leq \Gamma g(Y+Y') + (1-\Gamma)g(-(Y+Y')) \quad \text{a.s.} \tag{2.9}$$

Now let $\mathscr{A}_1$ be the smallest σ-field containing $\mathscr{A}_0$ and the σ-field generated by $Y + Y'$. Combining (2.9) and the fact that $E(\Gamma|\mathscr{A}_1) = 1/2$ a.s. [by (2.5)], we find that

$$E(g(X - X')|\mathscr{A}_1) \le \tfrac{1}{2}g(Y + Y') + \tfrac{1}{2}g(-(Y + Y')) \quad \text{a.s.},$$

which together with (2.6) leads to (2.8). □

LEMMA 2.6 (Majorization by independent sequences). *Let $n \in \mathbb{N} \setminus \{1\}$, and let $Y_1, \dots, Y_n, Z_1, \dots, Z_n$ be real-valued r.v.'s. Suppose*:

(2.10) *$E|Y_k| < \infty$ and $E|Z_k| < \infty$ for all $k \in \{1, \dots, n\}$;*

(2.11) *the r.v.'s $Z_1, \dots, Z_n$ are independent;*

(2.12) *$Eg(Y_1) \le Eg(Z_1)$ and $E(g(Y_l)|Y_1, \dots, Y_{l-1}) \le Eg(Z_l)$ a.s. for any $g \in \mathscr{C}^*$ and all $l \in \{2, \dots, n\}$.*

Then

$$\sum_{k=1}^{n} Y_k \prec \sum_{k=1}^{n} Z_k. \tag{2.13}$$

PROOF. In view of Lemma 2.1, it is enough to show that

$$Eg\left(\sum_{k=1}^{n} Y_k\right) \le Eg\left(\sum_{k=1}^{n} Z_k\right) \quad \text{for all } g \in \mathscr{C}^*. \tag{2.14}$$

To prove (2.14), let $Q_1 := P \circ Y_1^{-1}$ and let $Q_l(\cdot|y_1, \dots, y_{l-1})$ be a regular conditional distribution for Y_l given $Y_1 = y_1, \dots, Y_{l-1} = y_{l-1}$ ($y_1, \dots, y_{l-1} \in \mathbb{R}$; $l \in \{2, \dots, n\}$). Also define $Q_k^* := P \circ Z_k^{-1}$, $k = 1, \dots, n$, and let $h \in \mathscr{C}^*$ and $m \in \{2, \dots, n\}$ be arbitrarily fixed. By (2.10) and (2.12), there is a $P \circ (Y_1, \dots, Y_{m-1})^{-1}$-null set $N \subset \mathbb{R}^{m-1}$ such that for all $(y_1, \dots, y_{m-1}) \in \mathbb{R}^{m-1} \setminus N$ and for all rational numbers t,

$$\int |y_m| Q_m(dy_m|y_1, \dots, y_{m-1}) < \infty \tag{2.15}$$

and

$$\int h(t + y_m) Q_m(dy_m|y_1, \dots, y_{m-1}) \le Eh(t + Z_m). \tag{2.16}$$

Recalling the definition of $\mathscr{C}^*$ and taking into account that h (being a convex function $\mathbb{R} \to \mathbb{R}$) is continuous, it follows from (2.10), (2.15), (2.16) and Lebesgue's dominated convergence theorem that (2.16) must actually hold for all $t \in \mathbb{R}$ [provided $(y_1, \dots, y_{m-1}) \in \mathbb{R}^{m-1} \setminus N$]. Therefore

$$\begin{aligned} &\int \cdots \int h(y_1 + \cdots + y_m) Q_m(dy_m|y_1, \dots, y_{m-1}) \cdots Q_1(dy_1) \\ &\quad \le \int \cdots \int h(y_1 + \cdots + y_m) Q_m^*(dy_m) \\ &\qquad \times Q_{m-1}(dy_{m-1}|y_1, \dots, y_{m-2}) \cdots Q_1(dy_1). \end{aligned} \tag{2.17}$$

Leaving $y_1, \ldots, y_{m-2}$ fixed, the function

$$y_{m-1} \mapsto \int h(y_1 + \cdots + y_{m-1} + y_m) Q_m^*(dy_m)$$

obviously belongs to $\mathscr{C}^*$, so that the proof of (2.14) can now be accomplished by a straightforward induction based on (2.11), (2.12) and (2.17). □

3. Majorization of martingales and partial sums of independent vector-valued r.v.'s. Given a real-valued r.v. X, we use the symbol $\tilde{X}$ to denote an r.v. which has the same distribution as $R(X + X')$, where X' is an independent copy of X and R is a Bernoulli r.v. such that R and (X, X') are independent.

As we shall see, Theorem 1.1 is an almost immediate consequence of the following theorem.

THEOREM 3.1 (Majorization of martingales). *Let $n \in \mathbb{N}$, and let $U_1, \ldots, U_n, V_1, \ldots, V_n$ be real-valued r.v.'s. Put $\mathscr{A}_0 := \{\varnothing, \Omega\}$, and let $\mathscr{A}_k$, $k \in \{1, \ldots, n\}$, be the σ-field generated by $U_1, \ldots, U_k, V_1, \ldots, V_k$. Suppose*:

(3.1) $\quad E|U_k| < \infty$ *and* $E|V_k| < \infty$ *for all* $k \in \{1, \ldots, n\}$;

(3.2) $\quad E(U_k|\mathscr{A}_{k-1}) = 0$ *a.s. for all* $k \in \{1, \ldots, n\}$;

(3.3) $\quad$ *for each* $k \in \{1, \ldots, n\}$, V_k *and* $\mathscr{A}_{k-1}$ *are independent*;

(3.4) $\quad |U_k| \le V_k$ *a.s. for all* $k \in \{1, \ldots, n\}$.

Then

$$\sum_{j=1}^{n} U_j \prec \sum_{j=1}^{n} \tilde{V}_j, \tag{3.5}$$

where $\tilde{V}_1, \ldots, \tilde{V}_n$ are independent.

PROOF. Let $U := (U_1, \ldots, U_n)$ and $V := (V_1, \ldots, V_n)$. We may (and do) assume that the underlying probability space $(\Omega, \mathscr{A}, P)$ is so rich that we can find $\mathbb{R}^n$-valued r.v.'s $U' = (U_1', \ldots, U_n')$, $V' = (V_1', \ldots, V_n')$ and independent Bernoulli r.v.'s $R_1, \ldots, R_n$ such that $(R_1, \ldots, R_n)$ and (U, V, U', V') are independent and such that the joint distribution of (U, V, U', V') has the following property (from which it is uniquely determined): For each $k \in \{1, \ldots, n\}$,

(3.6) $\quad (U_k, V_k)$ and (U_k', V_k') are conditionally independent given $\mathscr{G}_{k-1}$, and $P((U_k, V_k) \in A|\mathscr{G}_{k-1}) = P((U_k', V_k') \in A|\mathscr{G}_{k-1}) = P((U_k, V_k) \in A|\mathscr{A}_{k-1})$ a.s. for any Borel set $A \subset \mathbb{R}^2$,

where $\mathscr{G}_0 := \{\varnothing, \Omega\}$ and $\mathscr{G}_j$, $j \in \{1, \ldots, n\}$, denotes the σ-field generated by $U_1, \ldots, U_j, V_1, \ldots, V_j, U_1', \ldots, U_j', V_1', \ldots, V_j'$. It is readily checked that (for each $k \in \{1, \ldots, n\}$):

(3.7) $\quad (V_k, V_k')$ and $\mathscr{G}_{k-1}$ are independent,

(3.8) $\quad E(U_k'|\mathscr{G}_{k-1}) = 0 \quad$ a.s.

Now let $k \in \{1, \dots, n\}$ and $g \in \mathscr{C}^*$ be arbitrarily fixed. In view of (3.1), (3.6) and (3.8), Lemma 2.2 yields

$$E(g(U_k)|\mathscr{G}_{k-1}) \le E(g(U_k - U_k')|\mathscr{G}_{k-1}) \quad \text{a.s.} \tag{3.9}$$

Moreover, using (3.1), (3.4), (3.6) and (3.7), it follows from Lemma 2.5 that

$$\begin{aligned} E(g(U_k - U_k')|\mathscr{G}_{k-1}) &\le \tfrac{1}{2}Eg(V_k + V_k') + \tfrac{1}{2}Eg(-(V_k + V_k')) \\ &= Eg(R_k(V_k + V_k')) \quad \text{a.s.} \end{aligned} \tag{3.10}$$

Taking into account that $g \in \mathscr{C}^*$ and $k \in \{1, \dots, n\}$ in (3.9) and (3.10) are arbitrary, (3.5) can now be derived from Lemma 2.6. □

REMARK 3.1. At first glance, the conditions about the joint distribution of the random vectors $(U_1, \dots, U_n)$ and $(V_1, \dots, V_n)$ in Theorem 3.1 look somewhat artificial. A more natural (but essentially equivalent) formulation of Theorem 3.1 can be given by using the concept of stochastic ordering [cf. Marshall and Olkin (1979), page 485, Proposition C.1].

PROOF OF THEOREM 1.1. Let $\mathscr{B}_0 := \{\varnothing, \Omega\}$, and let $\mathscr{B}_k$, $k \in \{1, \dots, n\}$, be the σ-field generated by $X_1, \dots, X_k$. Following Yurinskii (1974), we introduce the r.v.'s

$$U_k := E(\|S_n\| \,|\, \mathscr{B}_k) - E(\|S_n\| \,|\, \mathscr{B}_{k-1}), \qquad k \in \{1, \dots, n\}. \tag{3.11}$$

Then

$$\|S_n\| - E\|S_n\| = \sum_{j=1}^{n} U_j \quad \text{a.s.,} \tag{3.12}$$

$$E(U_k|\mathscr{B}_{k-1}) = 0 \quad \text{a.s.} \tag{3.13}$$

and [cf. Yurinskii (1974)]

$$|U_k| \le \|X_k\| + E\|X_k\| \quad \text{a.s.} \tag{3.14}$$

(for all $k \in \{1, \dots, n\}$). Hence Theorem 3.1 shows that $\|S_n\| - E\|S_n\| \prec T_n$. Finally, the relation $T_n \prec 4T_n^*$ can easily be obtained by combining Lemmas 2.3, 2.4 and 2.6. □

4. Almost sure behavior of partial sums of independent vector-valued r.v.'s. Throughout this section, $(\mathbb{F}, \mathscr{F}, \|\cdot\|)$ is a seminormed measurable vector space.

The principal purpose of this section is to prove Theorems 1.4 and 1.5. To this end, we shall first demonstrate that (1.12) and (1.14) are equivalent to certain analytic conditions. This is achieved by invoking a somewhat extended version (see Section 4.1) of Nagaev's (1972) theorem on necessary and sufficient conditions for the SLLN. Section 4.2 contains three technical lemmas. In Section 4.3 we shall combine Theorem 1.3 and the exponential inequality (Proposition 4.1) underlying Nagaev's just mentioned SLLN to establish a general stability result for $\mathbb{F}$-valued r.v.'s (Theorem 4.3). Together with Theorem 4.2, this result leads to Theorems 1.4 and 1.5 (see Section 4.4).

4.1. *The Nagaev condition.* We now proceed to describe the already announced extension of Nagaev's (1972) theorem. This generalization (Theorem 4.2) is basically a combination of results by Nagaev (1972), Kruglov (1974) and Volodin and Nagaev (1977). However, in order to obtain a version of Theorem 4.2(ii) that can be regarded as an exact analog of part (i), we have to add some further arguments.

The following symbols are used throughout this section:

(4.1) $\{\xi_j;\ j \in \mathbb{N}\}$ is a sequence of independent symmetric real-valued r.v.'s;

$$\zeta_n := \sum_{j=1}^{n} \xi_j \quad \text{for } n \in \mathbb{N}; \tag{4.2}$$

(4.3) $\{a_j;\ j \in \mathbb{N}\}$ is a nondecreasing sequence in $\mathbb{R}^+$ with $\sup\{a_j\colon j \in \mathbb{N}\} = \infty$;

(4.4) $\{\nu(j);\ j \in \mathbb{N}\}$ is a strictly increasing sequence in $\mathbb{N}$;

$$\Lambda := \{l \in \mathbb{N} \setminus \{1\}\colon \nu(l) - \nu(l-1) \geq 2\}; \tag{4.5}$$

$$J(l) := (\nu(l-1), \nu(l)] \cap \mathbb{N} \quad \text{for } l \in \mathbb{N} \setminus \{1\}; \tag{4.6}$$

(4.7) c, d, δ, ρ are positive real numbers with $1 < c \leq d$ and $\rho \leq 1$;

$$\eta_{k,\delta} := \xi_k I(\{|\xi_k| \leq \delta a_k\}) \quad \text{for } k \in \mathbb{N}; \tag{4.8}$$

$$C_{k,\delta}(t) := \log E(\exp(t\eta_{k,\delta})) \quad \text{for } k \in \mathbb{N} \text{ and } t \in \mathbb{R}. \tag{4.9}$$

DEFINITION 4.1. The sequence $\{\nu(j);\ j \in \mathbb{N}\}$ is said to have the property $\Pi(c, d, \rho, \{a_j;\ j \in \mathbb{N}\})$ if and only if

$$a_{\nu(k+r)} \geq \rho c^r a_{\nu(k)} \quad \text{for all } k, r \in \mathbb{N} \tag{4.10}$$

and

$$a_{\nu(l)} \leq d a_{\nu(l-1)+1} \quad \text{for all } l \in \Lambda. \tag{4.11}$$

This definition is motivated by the validity of the following theorem in conjunction with Lemma 4.1.

THEOREM 4.1. *Suppose the sequence $\{\nu(j);\ j \in \mathbb{N}\}$ has the property $\Pi(c, d, \rho, \{a_j;\ j \in \mathbb{N}\})$. Then*

$$a_n^{-1}\zeta_n \to_{a.s.} 0 \tag{4.12}$$

if and only if

$$\sum_{l=2}^{\infty} P\{\zeta_{\nu(l)} - \zeta_{\nu(l-1)} \geq \varepsilon a_{\nu(l)}\} < \infty \quad \textit{for any } \varepsilon \in \mathbb{R}^+. \tag{4.13}$$

PROOF. The proof of Theorem 3.4.1 in Stout (1974) can be carried over almost verbatim [see also Volodin and Nagaev (1977), page 812], and hence we omit the details. □

LEMMA 4.1. *Given c and* $\{a_j;\ j \in \mathbb{N}\}$, *it is always possible to find a strictly increasing sequence* $\{\nu(j);\ j \in \mathbb{N}\}$ *in* $\mathbb{N}$ *having the property* $\Pi(c, c, c^{-1}, \{a_j; j \in \mathbb{N}\})$.

PROOF. Cf. Volodin and Nagaev (1977), page 811. □

To examine the convergence properties of the series in (4.13), it is natural to introduce the truncated r.v.'s $\eta_{j,\delta}$ [cf. (4.8)], to show that the effect of the truncation is negligible and to estimate the probabilities

$$P\left\{\sum_{j=\nu(l)+1}^{\nu(l+1)} \eta_{j,\delta} \geq \varepsilon a_{\nu(l+1)}\right\}, \qquad \varepsilon \in \mathbb{R}^+,\ l \in \mathbb{N}, \tag{4.14}$$

by means of a suitable exponential inequality. Using the exponential estimate (4.17), this approach actually leads to necessary and sufficient conditions for the SLLN (see Theorem 4.2).

PROPOSITION 4.1 [Essentially due to Nagaev (1972)]. *Let* $Z_1, \ldots, Z_n$ *be independent real-valued r.v.'s such that*

$$\alpha := \sup\Big\{t \in \mathbb{R}: E\big(\exp(t|Z_j|)\big) < \infty \quad \text{for all } j \in \{1, \ldots, n\}\Big\} > 0, \tag{4.15}$$

and write $V_n := \sum_{j=1}^n Z_j$. *Also define*

$$C_n(t) := \sum_{j=1}^{n} \log E\big(\exp(tZ_j)\big) \quad \text{for } t \in (-\alpha, \alpha). \tag{4.16}$$

Then, for any pair $(t, x) \in (0, \alpha) \times \mathbb{R}^+$ *with* $x \geq C_n'(t)$,

$$P\{V_n \geq x\} \leq e^{-tx} E(\exp(tV_n)) \leq \exp(-t(x - C_n'(t))). \tag{4.17}$$

The proof is very short, and so we include it for completeness.

PROOF. The first inequality in (4.17) is obvious from Markov's inequality; to prove the second one, we begin by observing that $E(\exp(tV_n)) = \exp(C_n(t))$ (since $Z_1, \ldots, Z_n$ are independent). The inequality $C_n'' \geq 0$ [see, e.g., Feller (1969), page 2] shows that C_n' is nondecreasing. Hence and from the fact that $C_n(0) = 0$, it follows that

$$C_n(t) = \int_0^t C_n'(s)\, ds \leq tC_n'(t),$$

that is, $\exp(C_n(t)) \leq \exp(tC_n'(t))$, as desired. □

DEFINITION 4.2. Let $\varepsilon \in \mathbb{R}^+$. We say that the sequence $\{\xi_j;\ j \in \mathbb{N}\}$ satisfies the *Nagaev condition* $N(\delta, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$ if and only if the following relations are fulfilled:

$$\sum_{k=1}^{\infty} P\{|\xi_k| > \delta a_k\} < \infty; \tag{4.18}$$

(4.19) there is a sequence $\{t_k;\ k \in \mathbb{N}\}$ in $\mathbb{R}^+$ such that $\sum_{j \in J(l)} C'_{j,\delta}(t_l) \le \varepsilon a_{\nu(l)}$ for all $l \in \Lambda$ and $\sum_{l \in \Lambda} \exp(-\varepsilon a_{\nu(l)} t_l) < \infty$.

REMARK 4.1. As already mentioned in the proof of Proposition 4.1, the function $t \mapsto \sum_{j \in J(l)} C'_{j,\delta}(t)$ is nondecreasing (and continuous), and so (4.19) is equivalent to the original slightly different formulation of this condition occurring, for example, in the paper by Volodin and Nagaev (1977).

Now suppose

(4.20) the sequence $\{\nu(j);\ j \in \mathbb{N}\}$ has the property $\Pi(c, d, \rho, \{a_j;\ j \in \mathbb{N}\})$

(recall Lemma 4.1), and consider the following statements (4.21)–(4.24):

(4.21) $a_n^{-1}\zeta_n \to_{\text{a.s.}} 0$;

(4.22) for each $\varepsilon \in \mathbb{R}^+$, $\{\xi_j;\ j \in \mathbb{N}\}$ satisfies the Nagaev condition $N(\varepsilon, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$;

(4.23) $\limsup_{n \to \infty} a_n^{-1}\zeta_n < \infty$ a.s.;

(4.24) there exist $\delta, \varepsilon \in \mathbb{R}^+$ such that $\{\xi_j;\ j \in \mathbb{N}\}$ satisfies the Nagaev condition $N(\delta, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$.

THEOREM 4.2 [Essentially due to Nagaev (1972); see also Kruglov (1974) and Volodin and Nagaev (1977)]. *If* (4.20) *is fulfilled, then*

(i) (4.21) $\Leftrightarrow$ (4.22),
(ii) (4.23) $\Leftrightarrow$ (4.24).

PROOF. (i) In the proof of Theorem 2 in Volodin and Nagaev (1977), it is only required that the sequence $\{n_r;\ r \in \mathbb{N}\}$ of integers satisfies the conditions imposed on the sequence $\{\nu(j);\ j \in \mathbb{N}\}$ in (4.20), i.e., the specific construction of the sequence $\{n_r;\ r \in \mathbb{N}\}$ is not really needed (see also Theorem 4.1). Taking Remark 4.1 into account, (i) is equivalent to the thus obtained generalization of Theorem 2 of Volodin and Nagaev.

(ii) We begin by demonstrating that (4.23) implies (4.24). Suppose (4.23) is fulfilled. According to the Kolmogorov zero–one law, there is a $\delta \in \mathbb{R}^+$ such that

$$\limsup_{n\to\infty} a_n^{-1}|\zeta_n| \le \delta/3 \quad \text{a.s.}$$

(Recall that the ξ_j's are symmetric.) This entails that

$$\limsup_{n\to\infty} a_n^{-1}|\xi_n| \le 2\delta/3 \quad \text{a.s.},$$

so that

$$\sum_{k=1}^{\infty} P\{|\xi_k| > \delta a_k\} < \infty \tag{4.25}$$

(by the Borel–Cantelli lemma). Now let

$$\xi_j^* := \eta_{j,\delta} \quad \text{for } j \in \mathbb{N} \text{ [cf. (4.8)]}. \tag{4.26}$$

By (4.23), (4.25) and the Borel–Cantelli lemma,

$$\limsup_{n\to\infty} a_n^{-1} \sum_{j=1}^{n} \xi_j^* < \infty. \tag{4.27}$$

Suppose (4.24) does not hold. Then there is no $\varepsilon \in \mathbb{R}^+$ such that $\{\xi_j;\ j \in \mathbb{N}\}$ satisfies the Nagaev condition $N(\delta, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$. Letting

$$t_l^*(\varepsilon) := \sup\left\{ s \in \mathbb{R}^+ \colon \sum_{j\in J(l)} C'_{j,\delta}(s) \le \varepsilon a_{\nu(l)} \right\} \tag{4.28}$$

for $l \in \Lambda$ and $\varepsilon \in \mathbb{R}^+$ [note that $C'_{j,\delta}(0) = E\xi_j^* = 0$ for all $j \in \mathbb{N}$], this means that for any $\varepsilon \in \mathbb{R}^+$,

$$\sum_{l\in\{k\in\Lambda\colon t_k^*(\varepsilon)<\infty\}} \exp\big(-\varepsilon a_{\nu(l)} t_l^*(\varepsilon)\big) = \infty. \tag{4.29}$$

Consequently, there also exists a nondecreasing sequence $\{\varepsilon_j;\ j \in \mathbb{N}\}$ in $\mathbb{R}^+$ satisfying

$$\sup\{\varepsilon_j\colon j \in \mathbb{N}\} = \infty \tag{4.30}$$

and

$$\varepsilon_{\nu(l)} \le d\varepsilon_{\nu(l-1)+1} \quad \text{for all } l \in \Lambda \tag{4.31}$$

such that

$$\sum_{l\in\{k\in\Lambda\colon t_k^*(\varepsilon_{\nu(k)})<\infty\}} \exp\big(-\varepsilon_{\nu(l)} a_{\nu(l)} t_l^*(\varepsilon_{\nu(l)})\big) = \infty. \tag{4.32}$$

By (4.20) and (4.31),

(4.33) the sequence $\{\nu(j);\ j \in \mathbb{N}\}$ has the property $\Pi(c, d^2, \rho, \{\varepsilon_j a_j;\ j \in \mathbb{N}\})$.

Since $|\xi_k^*| \le \delta a_k$ for all $k \in \mathbb{N}$ and

$$\lim_{n\to\infty} (\varepsilon_n a_n)^{-1} \sum_{j=1}^{n} \xi_j^* = 0 \quad \text{a.s.} \tag{4.34}$$

[recall (4.27)], it follows from (4.26), (4.33) and part (i) that the series in (4.32) must be convergent. This fact is the desired contradiction.

We now turn to the converse implication (4.24) $\Rightarrow$ (4.23). Let $\varepsilon, \delta \in \mathbb{R}^+$ be such that $\{\xi_j;\ j \in \mathbb{N}\}$ satisfies the Nagaev condition $N(\delta, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$. For $p \in \mathbb{N}$ and $q \in \mathbb{N} \setminus \{1\}$, define

$$\zeta_{p,\delta} := \sum_{j=1}^{p} \eta_{j,\delta} \quad \text{[cf. (4.8)]}$$

and

$$M_{q,\delta} := a_{\nu(q)}^{-1} \max_{k \in J(q)} |\zeta_{k,\delta} - \zeta_{\nu(q-1),\delta}|.$$

Then, for any pair (l, n) with $l \in \mathbb{N} \setminus \{1\}$ and $n \in J(l)$,

$$\begin{aligned} a_n^{-1}|\zeta_{n,\delta} - \zeta_{\nu(1),\delta}| &\le a_n^{-1} \sum_{k=2}^{l} a_{\nu(k)} M_{k,\delta} \\ &\le d \sum_{k=2}^{l} a_{\nu(k)} a_{\nu(l)}^{-1} M_{k,\delta} \\ &\le d\rho^{-1} \sum_{k=2}^{l} c^{-(l-k)} M_{k,\delta} \end{aligned} \tag{4.35}$$

[by (4.20)]. Now let $\{t_k;\ k \in \mathbb{N}\}$ be chosen according to (4.19). Then, for $k \in \Lambda$,

$$\begin{aligned} &P\{M_{k,\delta} \ge 2\varepsilon\} \\ &\quad\le 4P\{\zeta_{\nu(k),\delta} - \zeta_{\nu(k-1),\delta} \ge 2\varepsilon a_{\nu(k)}\} \quad \text{(by Lévy's inequality)} \\ &\quad\le 4\exp(-\varepsilon a_{\nu(k)} t_k) \quad \text{[using (4.19) and Proposition 4.1]}. \end{aligned} \tag{4.36}$$

Combining (4.18), (4.19), (4.35), (4.36), the Borel–Cantelli lemma and the Toeplitz lemma, we get

$$\limsup_{n \to \infty} a_n^{-1}\zeta_n \le d\rho^{-1}c(c-1)^{-1} \max(2\varepsilon, \delta) < \infty \quad \text{a.s.,} \tag{4.37}$$

as desired. □

4.2. *Some inequalities for vector-valued r.v.'s.* Let $n \in \mathbb{N}$, and let $Y_1, \dots, Y_n$ be independent $\mathbb{F}$-valued r.v.'s. Write $T_0 := 0$ and $T_k := \sum_{j=1}^{k} Y_j$ (for $k \in \{1, \dots, n\}$). Moreover, let $\alpha \in \mathbb{R}^+$ and $\beta \in [0, 1)$ be such that

$$P\{\|T_n - T_k\| > \alpha\} \le \beta \quad \text{for all } k \in \{1, \dots, n\}. \tag{4.38}$$

Essentially the same argument as for real-valued r.v.'s [see, e.g., Chung (1974), page 120] also yields:

LEMMA 4.2 (Ottaviani's inequality). *If* (4.38) *holds, then*

$$P\Big\{\max_{1 \le k \le n} \|T_k\| > \alpha + x\Big\} \le (1 - \beta)^{-1} P\{\|T_n\| > x\} \quad \textit{for all } x \in \mathbb{R}^+. \tag{4.39}$$

The next lemma is a variant of a result by Hoffmann-Jørgensen (1974).

LEMMA 4.3. *Let* $N_n := \max\{\|Y_j\|: 1 \le j \le n\}$. *If* (4.38) *holds, then*

$$P\{\|T_n\| > 2\alpha + x + y\} \le P\{N_n > x\} + \beta(1-\beta)^{-1}P\{\|T_n\| > y\} \tag{4.40}$$

for all $x, y \in \mathbb{R}^+$.

PROOF [Partly adapted from Hoffmann-Jørgensen (1974), page 164]. For $\omega \in \Omega$, let $\sigma(\omega) := \min\{k \in \{1,\dots,n\}: \|T_k(\omega)\| > \alpha + y\}$ or $:= n+1$ according as $\max\{\|T_j(\omega)\|: 1 \le j \le n\} > \alpha + y$ or not. Then

$$P\{\|T_n\| > 2\alpha + x + y\} = \sum_{k=1}^{n} P\{\|T_n\| > 2\alpha + x + y, \sigma = k\}. \tag{4.41}$$

If $\sigma(\omega) = k$ $(\le n)$ and $\|T_n(\omega)\| > 2\alpha + x + y$, then $\|T_{k-1}(\omega)\| \le \alpha + y$, and so

$$\begin{aligned}\|T_n(\omega) - T_k(\omega)\| &\ge \|T_n(\omega)\| - \|T_{k-1}(\omega)\| - \|Y_k(\omega)\| \\ &> 2\alpha + x + y - (\alpha + y) - N_n(\omega) = \alpha + x - N_n(\omega).\end{aligned}$$

It follows that (for $k \in \{1,\dots,n\}$)

$$\begin{aligned}P\{\|T_n\| &> 2\alpha + x + y, \sigma = k\} \\ &\le P\{\|T_n - T_k\| > \alpha + x - N_n, \sigma = k\} \\ &\le P\{N_n > x, \sigma = k\} + P\{\|T_n - T_k\| > \alpha, \sigma = k\} \\ &= P\{N_n > x, \sigma = k\} + P\{\|T_n - T_k\| > \alpha\}P\{\sigma = k\}.\end{aligned}$$

Hence and from (4.41) we conclude that

$$\begin{aligned}P\{\|T_n\| &> 2\alpha + x + y\} \\ &\le P\{N_n > x\} + \sum_{k=1}^{n} P\{\sigma = k\}P\{\|T_n - T_k\| > \alpha\} \\ &\le P\{N_n > x\} + \beta P\Big\{\max_{1\le k\le n} \|T_k\| > \alpha + y\Big\} \\ &\le P\{N_n > x\} + \beta(1-\beta)^{-1}P\{\|T_n\| > y\} \quad \text{(by Lemma 4.2)}. \qquad \square\end{aligned}$$

In the case of symmetric r.v.'s, the following lemma is closely related to results that are implicit in papers by Kuelbs (1977) and Kuelbs and Zinn (1979).

LEMMA 4.4. *Let* $N_n := \max\{\|Y_j\|: 1 \le j \le n\}$, *and suppose that*

$$E\|Y_k\| < \infty \quad \text{for all } k \in \{1,\dots,n\}. \tag{4.42}$$

If (4.38) *holds for some* $\beta \in [0, 1/5)$, *then*

$$E\|T_n\| \le \alpha + \left(1 - \frac{4\beta}{1-\beta}\right)^{-1}\left(3\alpha + 4\int_\alpha^\infty P\{N_n > x\}\,dx\right). \tag{4.43}$$

PROOF. We have

$$(4.44)\qquad E\|T_n\| = \int_0^\infty P\{\|T_n\| > x\}\,dx \le \alpha + \int_\alpha^\infty P\{\|T_n\| > x\}\,dx.$$

Now

$$\begin{aligned}\int_\alpha^\infty P\{\|T_n\| > x\}\,dx &= \int_\alpha^{4\alpha} + \int_{4\alpha}^\infty \le 3\alpha + 4\int_\alpha^\infty P\{\|T_n\| > 4x\}\,dx\\ &\le 3\alpha + 4\int_\alpha^\infty P\{N_n > x\}\,dx\\ &\quad + 4\beta(1-\beta)^{-1}\int_\alpha^\infty P\{\|T_n\| > x\}\,dx\end{aligned}$$

(by Lemma 4.3). Therefore

$$(4.45)\quad \int_\alpha^\infty P\{\|T_n\| > x\}\,dx \le \left(1 - \frac{4\beta}{1-\beta}\right)^{-1}\left(3\alpha + 4\int_\alpha^\infty P\{N_n > x\}\,dx\right),$$

which together with (4.44) leads to (4.43). □

4.3. *A general stability result for vector-valued r.v.'s.* We now return to the setting underlying Theorems 1.4 and 1.5, that is,

(4.46) $\{X_j;\ j \in \mathbb{N}\}$ is a sequence of independent $\mathbb{F}$-valued r.v.'s, $S_n := \sum_{j=1}^n X_j$ for $n \in \mathbb{N}$, and $\{R_j;\ j \in \mathbb{N}\}$ is a sequence of independent Bernoulli r.v.'s which is also independent of $\{X_j;\ j \in \mathbb{N}\}$.

Moreover,

(4.47) $\{a_j;\ j \in \mathbb{N}\}$, $\{\nu(j);\ j \in \mathbb{N}\}$, Λ, $J(l)$ $(l \in \mathbb{N} \setminus \{1\})$, c, d and ρ are as in (4.3)–(4.7).

In the sequel, it is always assumed that

(4.48) the sequence $\{\nu(j);\ j \in \mathbb{N}\}$ has the property $\Pi\big(c, d, \rho, \{a_j;\ j \in \mathbb{N}\}\big)$

(recall Lemma 4.1).

The next theorem gives an upper bound for $\limsup_{n\to\infty} a_n^{-1}\|S_n\|$ under the assumption that the sequence $\{R_j\|X_j\|;\ j \in \mathbb{N}\}$ satisfies a Nagaev type condition. Combining this result, Theorem 4.2, and Lemmas 4.5 and 4.6, we obtain Theorems 1.4 and 1.5 (see Section 4.4).

In conjunction with Lemmas 4.5 and 4.6, Theorem 4.3 can also be utilized as a convenient starting point (note that its proof does not depend on Theorem 4.2) for deriving various sufficient criteria for the SLLN that are easier to work with than (4.49) and (4.50). [For further details, see Berger (1991).]

THEOREM 4.3. *Let $\delta, \varepsilon, \gamma \in \mathbb{R}^+$. Suppose* (4.48) *is fulfilled, and in addition*:

(4.49) *the sequence $\{R_j\|X_j\|; j \in \mathbb{N}\}$ satisfies the Nagaev condition $N(\delta, \varepsilon, \{\nu(j); j \in \mathbb{N}\}, \{a_j; j \in \mathbb{N}\})$*,

$$(4.50)\qquad \limsup_{n\to\infty} a_n^{-1} E\left\| \sum_{j=1}^{n} X_j I\big(\{\|X_j\| \le \delta a_j\}\big)\right\| \le \gamma.$$

Then

$$(4.51)\qquad \limsup_{n\to\infty} a_n^{-1}\|S_n\| \le d\rho^{-1} c(c-1)^{-1}\max(8\varepsilon + 6\gamma, \delta) \quad a.s.$$

PROOF. For $p \in \mathbb{N}$ and $q \in \mathbb{N} \setminus \{1\}$, define

$$T_{p,\delta} := \sum_{j=1}^{p} X_j I\big(\{\|X_j\| \le \delta a_j\}\big)$$

and

$$M_{q,\delta} := a_{\nu(q)}^{-1} \max_{k \in J(q)} \|T_{k,\delta} - T_{\nu(q-1),\delta}\|.$$

Using (4.48) and arguing as in (4.35), we find that

$$(4.52)\qquad a_n^{-1}\|T_{n,\delta} - T_{\nu(1),\delta}\| \le d\rho^{-1} \sum_{k=2}^{l} c^{-(l-k)} M_{k,\delta}$$

for any pair (l, n) with $l \in \mathbb{N} \setminus \{1\}$ and $n \in J(l)$. Now let $\gamma^* \in (\gamma, \infty)$ be arbitrarily fixed, let

$$\tilde{\eta}_{j,\delta} := R_j\|X_j\| I\big(\{\|X_j\| \le \delta a_j\}\big) \quad \text{for } j \in \mathbb{N},$$
$$\tilde{C}_{j,\delta}(t) := \log E\big(\exp(t\tilde{\eta}_{j,\delta})\big) \quad \text{for } j \in \mathbb{N} \text{ and } t \in \mathbb{R},$$

and let the sequence $\{t_k;\ k \in \mathbb{N}\}$ in $\mathbb{R}^+$ be chosen in such a way that

$$(4.53)\qquad \sum_{j \in J(l)} \tilde{C}'_{j,\delta}(t_l) \le \varepsilon a_{\nu(l)} \quad \text{for } l \in \Lambda$$

and

$$(4.54)\qquad \sum_{l \in \Lambda} \exp\big(-\varepsilon a_{\nu(l)} t_l\big) < \infty$$

[recall (4.49) and Definition 4.2]. By (4.50), there is a $p_0 \in \mathbb{N}$ such that

$$(4.55)\qquad E\|T_{p,\delta}\| \le \gamma^* a_p \quad \text{for } p \ge p_0.$$

For $k \in \Lambda$ with $\nu(k-1) \ge p_0$ and $n \in [\nu(k-1), \nu(k)] \cap \mathbb{N}$, we have

$$(4.56)\qquad P\big\{\|T_{\nu(k),\delta} - T_{n,\delta}\| > 4\gamma^* a_{\nu(k)}\big\} \le \tfrac{1}{2}$$

[by (4.55) and Markov's inequality] and

$$\begin{aligned}
P\{M_{k,\delta} &\geq 8\varepsilon + 6\gamma^*\} \\
&\leq 2P\{\|T_{\nu(k),\delta} - T_{\nu(k-1),\delta}\| \geq 2(4\varepsilon + \gamma^*)a_{\nu(k)}\} \qquad \text{[by (4.56) and Lemma 4.2]} \\
&\leq 2P\{|\,\|T_{\nu(k),\delta} - T_{\nu(k-1),\delta}\| \\
&\qquad - E\|T_{\nu(k),\delta} - T_{\nu(k-1),\delta}\|\,| \geq 8\varepsilon a_{\nu(k)}\} \qquad \text{[by (4.55)]} \\
&\leq 4\exp(-2\varepsilon a_{\nu(k)}t_k)E\left(\exp\left(t_k \sum_{j\in J(k)} \tilde{\eta}_{j,\delta}\right)\right) \qquad \text{(by Theorem 1.3)} \\
&\leq 4\exp\left(-t_k\left(2\varepsilon a_{\nu(k)} - \sum_{j\in J(k)} \tilde{C}'_{j,\delta}(t_k)\right)\right) \\
&\qquad \text{[by (4.53) and Proposition 4.1]} \\
&\leq 4\exp(-\varepsilon a_{\nu(k)}t_k) \quad \text{[by (4.53)]}.
\end{aligned} \tag{4.57}$$

(4.49) also yields

$$\sum_{j=1}^{\infty} P\{\|X_j\| > \delta a_j\} < \infty. \tag{4.58}$$

Combining (4.52), (4.57), (4.54), (4.58), the Borel–Cantelli lemma and the Toeplitz lemma, we get

$$\limsup_{n\to\infty} a_n^{-1}\|S_n\| \leq d\rho^{-1}c(c-1)^{-1}\max(8\varepsilon + 6\gamma^*, \delta) \quad \text{a.s.}$$

Since $\gamma^* \in (\gamma, \infty)$ is arbitrary, this implies (4.51). □

4.4. *Conclusion of the proof of Theorems* 1.4 *and* 1.5. The notation and assumptions are as introduced at the beginning of the preceding section. By virtue of Lemma 4.1 and Theorem 4.2, to establish Theorems 1.4 and 1.5, it suffices to prove the following two theorems.

THEOREM 1.4′. *The assertion of Theorem* 1.4 *holds if instead of* (1.12) *it is assumed that*

(4.59) *for each* $\varepsilon \in \mathbb{R}^+$, *the sequence* $\{R_j\|X_j\|;\ j \in \mathbb{N}\}$ *satisfies the Nagaev condition* $N(\varepsilon, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$.

THEOREM 1.5′. *The assertion of Theorem* 1.5 *holds if instead of* (1.14) *it is assumed that*

(4.60) *there exist* $\delta, \varepsilon \in \mathbb{R}^+$ *such that the sequence* $\{R_j\|X_j\|;\ j \in \mathbb{N}\}$ *satisfies the Nagaev condition* $N(\delta, \varepsilon, \{\nu(j);\ j \in \mathbb{N}\}, \{a_j;\ j \in \mathbb{N}\})$.

Using Theorem 4.3 (and recalling Definition 4.2), we see that Theorems 1.4′ and 1.5′ will be proven once the following two lemmas are verified. Both lemmas being easy consequences of Lemma 4.4, the proofs are left to the reader.

LEMMA 4.5. *The notation is as in* (4.46) *and* (4.47). *Suppose*

$$a_n^{-1}\|S_n\| \to_P 0 \tag{4.61}$$

and

$$\sum_{k=1}^{\infty} P\{\|X_k\| > \delta a_k\} < \infty \quad \text{for some } \delta \in \mathbb{R}^+. \tag{4.62}$$

Then

$$\limsup_{n\to\infty} a_n^{-1} E\left\| \sum_{j=1}^{n} X_j I\big(\{\|X_j\| \le \delta a_j\}\big)\right\| \le 8\delta. \tag{4.63}$$

LEMMA 4.6. *If condition* (4.61) *in Lemma* 4.5 *is weakened to*

(4.64) *the sequence* $\{a_n^{-1}\|S_n\|;\ n \in \mathbb{N}\}$ *is bounded in probability,*

then

$$\limsup_{n\to\infty} a_n^{-1} E\left\| \sum_{j=1}^{n} X_j I\big(\{\|X_j\| \le \delta a_j\}\big)\right\| < \infty. \tag{4.65}$$

REFERENCES

ALEXANDER, K. S. (1984). Probability inequalities for empirical processes and a law of the iterated logarithm. *Ann. Probab.* **12** 1041–1067.

BERGER, E. (1991). A note on the strong law of large numbers for partial sums of independent random vectors. In *Almost Everywhere Convergence in Probability and Ergodic Theory II* (A. Bellow and R. Jones, eds.). Academic, New York. To appear.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

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

CHUNG, K. L. (1974). *A Course in Probability Theory*, 2nd ed. Academic, New York.

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

DE ACOSTA, A. (1981). Inequalities for B-valued random vectors with applications to the strong law of large numbers. *Ann. Probab.* **9** 157–161.

FELLER, W. (1969). Limit theorems for probabilities of large deviations. *Z. Wahrsch. Verw. Gebiete* **14** 1–20.

HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1929). Some simple inequalities satisfied by convex functions. *Messenger Math.* **58** 145–152.

HARTMAN, P. and WINTNER, A. (1941). On the law of the iterated logarithm. *Amer. J. Math.* **63** 169–176.

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

KARAMATA, J. (1932). Sur une inégalité rélative aux fonctions convexes. *Publ. Math. Univ. Belgrade* **1** 145–148.

KLASS, M. (1976). Toward a universal law of the iterated logarithm. I. *Z. Wahrsch. Verw. Gebiete* **36** 165–178.

KLASS, M. (1977). Toward a universal law of the iterated logarithm. II. *Z. Wahrsch. Verw. Gebiete* **39** 151–165.

KRUGLOV, V. M. (1974). On the behavior of sums of independent random variables. *Theory Probab. Appl.* **19** 374–379.

KUELBS, J. (1977). Kolmogorov's law of the iterated logarithm for Banach space valued random variables. *Illinois J. Math.* **21** 784–800.

KUELBS, J. and ZINN, J. (1979). Some stability results for vector valued random variables. *Ann. Probab.* **7** 75–84.

KUELBS, J. and ZINN, J. (1983). Some results on LIL behavior. *Ann. Probab.* **11** 506–557.

LEDOUX, M. and TALAGRAND, M. (1988). Characterization of the law of the iterated logarithm in Banach spaces. *Ann. Probab.* **16** 1242–1264.

LEDOUX, M. and TALAGRAND, M. (1990). Some applications of isoperimetric methods to strong limit theorems for sums of independent random variables. *Ann. Probab.* **18** 754–789.

MARCINKIEWICZ, J. and ZYGMUND, A. (1937). Sur les fonctions indépendantes. *Fund. Math.* **29** 60–90.

MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications.* Academic, New York.

NAGAEV, S. V. (1972). On necessary and sufficient conditions for the strong law of large numbers. *Theory Probab. Appl.* **17** 573–581.

NAGAEV, S. V. (1979). Large deviations of sums of independent random variables. *Ann. Probab.* **7** 745–789.

PROHOROV, YU. V. (1959). Some remarks on the strong law of large numbers. *Theory Probab. Appl.* **4** 204–208.

ROBERTS, A. W. and VARBERG, D. E. (1973). *Convex Functions.* Academic, New York.

ROSENTHAL, H. (1972). On the span in L^p of sequences of independent random variables. In *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **2** 149–167. Univ. California Press, Berkeley.

STOUT, W. (1974). *Almost Sure Convergence.* Academic, New York.

VOLODIN, N. A. and NAGAEV, S. V. (1977). A remark on the strong law of large numbers. *Theory Probab. Appl.* **22** 810–813.

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

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
UNIVERSITÄT GÖTTINGEN
LOTZESTRASSE 13
D-3400 GÖTTINGEN
GERMANY