The Annals of Probability
1989, Vol. 17, No. 4, 1546–1570

ISOPERIMETRY AND INTEGRABILITY OF THE SUM OF INDEPENDENT BANACH-SPACE VALUED RANDOM VARIABLES[1]

BY MICHEL TALAGRAND

University of Paris VI and The Ohio State University

We develop a new method to study the tails of a sum of independent mean zero Banach-space valued random variables $(X_i)_{i \le N}$. It relies on a new isoperimetric inequality for subsets of a product of probability spaces. In particular, we prove that for $p \ge 1$,

$$\Big\| \sum_{i \le N} X_i \Big\|_p \le \frac{Kp}{1 + \log p} \left(\Big\| \sum_{i \le N} X_i \Big\|_1 + \| \max_{i \le N} \|X_i\| \|_p \right),$$

where K is a universal constant. Other optimal inequalities for exponential moments are obtained.

1. Introduction. Consider a sequence $(X_i)_{i \ge 1}$ of independent random variables (r.v.) valued in a Banach space B. Throughout the paper, we assume that these variables are integrable, with $EX_i = 0$. In the case where $S_N = \sum_{i=1}^N X_i$ converges a.s. to a r.v. S, the integrability of S has been studied by a number of authors. A landmark result in that direction is the following inequality, due to Hoffmann-Jørgensen [6]. For each $p \ge 1$, there is a constant $C(p)$ such that

$$\Big\| \sum_{i \le N} X_i \Big\|_p \le C(p) \left(\Big\| \sum_{i \le N} X_i \Big\|_1 + \| \max_{i \le N} \|X_i\| \|_p \right).$$

The usual proof of this result yields a constant $C(p)$ that has exponential growth in p. One of the main results of this paper is that the actual order of growth of $C(p)$ is $p/\log p$.

THEOREM 1. *For some universal constant K, and all $p \ge 1$, we have*

$$\Big\| \sum_{i \le N} X_i \Big\|_p \le \frac{Kp}{1 + \log p} \left(\Big\| \sum_{i \le N} X_i \Big\|_1 + \| \max_{i \le N} \|X_i\| \|_p \right).$$

In the case of real-valued r.v. a result of a similar nature had been obtained by Johnson, Schechtman and Zinn [8]. In that case, the order $p/\log p$ is already optimal.

When $\sup_i \|X_i\|_\infty < \infty$, and when S_N converges a.s. to some limit S, a consequence of Theorem 1 is that $E \exp(\alpha \|S\| \log^+ \|S\|) < \infty$ for some $\alpha > 0$. This was previously known in cotype 2 spaces [2]. In general, it was only known that

Received March 1987; revised January 1989.
[1]This work was partially supported by an NSF grant.
AMS 1980 *subject classifications*. Primary 60B11; secondary 60E15, 60G50, 28A35.
Key words and phrases. Isoperimetric inequality, Hoffmann-Jørgensen inequality, Poisson measure, exponential inequalities.

$E \exp \lambda\|S\| \leq \infty$ for all $\lambda > 0$. We can actually prove the following, in any Banach space.

THEOREM 2. *Assume that the sequence* (X_i) *is symmetric, that* $\sup_i \|X_i\|_\infty \leq C$ *and* $\sup_N E\|S_N\| < \infty$. *Let* $W = \sup_N \|S_N\|$. *Then for each* $\alpha > 3$, *we have*

$$E\big(\exp C^{-1} W(\log^+ W - \alpha \log\log(e + W))\big) < \infty.$$

There is no reason to believe that the number 3 is sharp, or even the order of the term $\log\log(e + W)$. Theorem 2 is based on an explicit (but complicated) bound for the tails $P(\|S_N\| > t)$. An application of Theorem 2 is given to the integrability of Poisson measures.

To state our next results we set, for $\alpha > 0$,

$$\|X\|_{\Phi_\alpha} = \inf\left\{c > 0;\ E\exp\left\|\frac{X}{c}\right\|^\alpha \leq 2\right\}.$$

For $\alpha \geq 1$, $\|\cdot\|_{\Phi_\alpha}$ is a norm, and for $\alpha < 1$ it is equivalent to a norm. The following results appear to be new even for real-valued r.v. The next theorem improves on Theorem 3 of [6].

THEOREM 3. *For* $0 \leq \alpha \leq 1$, *we have*

$$\left\|\sum_{i\leq N} X_i\right\|_{\Phi_\alpha} \leq K(\alpha)\left(\left\|\sum_{i\leq N} X_i\right\|_1 + \left\|\max_{i\leq N}\|X_i\|\right\|_{\Phi_\alpha}\right),$$

where $K(\alpha)$ *depends on* α *only.*

The next theorem improves on [3], Theorems 5.2 and 5.3.

THEOREM 4. *Consider* $1 < \alpha \leq 2$, *and* β *such that* $1/\alpha + 1/\beta = 1$. *Then*

$$\left\|\sum_{i\leq N} X_i\right\|_{\Phi_\alpha} \leq K(\alpha)\left(\left\|\sum_{i\leq N} X_i\right\|_1 + \left(\sum_{i\leq N}\|X_i\|^\beta_{\Phi_\alpha}\right)^{1/\beta}\right),$$

where $K(\alpha)$ *depends on* α *only. Moreover,* $K(\alpha)$ *is bounded in any interval* $[1 + \varepsilon, 2]$.

We recall that for a sequence $(a_i)_{i\geq 1}$, and $p \geq 1$ we set

$$\|(a_i)\|_{p\infty} = \left(\sup\left\{t\,\mathrm{card}\left\{i;\ |a_i| \geq t^{-1/p}\right\};\ t > 0\right\}\right)^{1/p}.$$

The following rather general exponential inequality was inspired by the results of [4].

THEOREM 5. *Consider numbers* $\alpha, r, s > 0$, *with* $1/r + 1/s = 1$.

(a) *Set* $m = E\|\sum_{i \le N} X_i\|_1$, $M = \|(\|X_i\|_{\Phi_\alpha})\|_{r,\infty}$. *Assume that* $\alpha > s \ge 1$, $s \le 2$. *Then there exists* $K > 0$, *depending on* α, r *only such that for all* $t \ge Km$, *we have*

$$P\left(\left\|\sum_{i \le N} X_i\right\| > t\right) \le K \exp - \frac{t^2}{Km^2}$$

if $M^s t^{2-s} \le eKm^2$ *and*

$$P\left(\left\|\sum_{i \le n} X_i\right\| \ge t\right) \le K \exp - \left(\frac{t}{KM}\right)^s \left(\log\left(\min\left(\frac{M^s t^{2-s}}{Km^2}, \frac{t}{Km}\right)\right)\right)^{1-s/\alpha}$$

if $M^s t^{2-s} \ge eKm^2$.

(b) *If* $\alpha > s > 2$, *there exists a constant* K *depending on* α, s *only such that for all* $t > Km$, *we have*

$$P\left(\left\|\sum_{i \le N} X_i\right\| > t\right) \le K \exp - \left(\frac{t}{KM}\right)^s.$$

Previous research on the topic is based on what one might call rather classical probabilistic ideas, like martingale theory or converse Kolmogorov inequalities. These ideas do not seem to work here, and our approach, that we outline now, is of a rather different nature. A simple idea is to consider a Bernoulli sequence $(\varepsilon_i)_{i \le N}$ independent of X_i, and to study first the conditional expectation $E_\varepsilon\|\sum_{i \le N}\varepsilon_i X_i\|$ of $\|\sum_{i \le N}\varepsilon_i X_i\|$ given $(X_i)_{i \le N}$. The idea is that the integrability properties of $\|\sum_{i \le N}\varepsilon_i X_i\|$ are closely related to those of $E_\varepsilon\|\sum_{i \le N}\varepsilon_i X_i\|$. This is a well-known fact. It will be convenient to use the following recent result (also of an isoperimetric nature) established in [13].

THEOREM 6. *Consider a sequence* $(x_i)_{i \le N}$ *in a Banach space* B, *and set*

$$\sigma^2 = \sup\left\{\sum_{i \le N} x^*(x_i)^2 : x^* \in B^*, \|x^*\| \le 1\right\}.$$

Then, if M *is a median of* $\|\sum_{i \le N}\varepsilon_i x_i\|$, *we have, for* $t > 0$,

$$P\left(\left|\left\|\sum_{i \le N} \varepsilon_i x_i\right\| - M\right| \ge t\right) \le 4 \exp - t^2/8\sigma^2.$$

In particular,

$$P\left(\left\|\sum_{i \le N} \varepsilon_i x_i\right\| \ge 2E\left\|\sum_{i \le N} \varepsilon_i x_i\right\| + t\right) \le 4 \exp - t^2/8\sigma^2.$$

The study of $E_\varepsilon\|\sum_{i \le N}\varepsilon_i X_i\|$ is made easier by the observation that $E_\varepsilon\|\sum_{i \in I}\varepsilon_i X_i\|$ is an increasing function of the subset I of $\{1, \dots, N\}$. Let $m = E\|\sum_{i \le N}\varepsilon_i X_i\|$. Consider the set $A = \{E_\varepsilon\|\sum_{i \le N}\varepsilon_i X_i\| \le 2m\}$, so $P(A) \ge \frac{1}{2}$. For ω

in Ω, $x^1, \dots, x^q$ in A fixed and $l \le q$, we consider

$$I_l = \{i \le N;\ X_i(\omega) = X_i(x^l)\},$$

so we have

$$E_\varepsilon \Big\| \sum_{i \in I_l} \varepsilon_i X_i(\omega) \Big\| = E_\varepsilon \Big\| \sum_{i \in I_l} \varepsilon_i X_i(x^l) \Big\| \le E_\varepsilon \Big\| \sum_{i \le N} \varepsilon_i X_i(x^l) \Big\| \le 2m.$$

If we set $I = \bigcup_{l \le q} I_l$, we have

$$E_\varepsilon \Big\| \sum_{i \in I} \varepsilon_i X_i(\omega) \Big\| \le 2qm,$$

so we have the bound

$$E_\varepsilon \Big\| \sum_{i \le N} \varepsilon_i X_i(\omega) \Big\| \le 2qm + \sum_{i \notin I} \|X_i(\omega)\|. \tag{1.1}$$

The natural idea is that, when q becomes large, for each ω outside a small exceptional set, one can find $x^1, \dots, x^q$ in A such that $k = N - \operatorname{card} I$ is small, so that the estimate (1.1) is useful when one bounds $\sum_{i \notin I} \|X_i(\omega)\|$ by the sum of the k largest terms of the sequence $(\|X_i(\omega)\|)_{I \le N}$. The key point of this method is to obtain an estimate of the size of the exceptional set. Consider a probability space (Ω, Σ, μ), and $N \ge 1$. Denote by P the product measure μ^N on Ω^N. (For simplicity of notation, the Nth power of a measure μ is denoted by μ^N instead of $\mu^{\otimes N}$.) To a subset A of Ω^N, we associate

$$H(A, q, k) = \left\{\omega \in \Omega^N;\ \exists\, x^1, \dots, x^q \in A, \operatorname{card}\left\{i \le N;\ \omega_i \notin \{x_i^1, \dots, x_i^q\}\right\} \le k\right\}.$$

The main result of this work is as follows.

Theorem 7. *For some universal constant K, we have*

$$P_*(H(A, q, k)) \ge 1 - \left(K\left(\frac{1}{k}\log\frac{1}{P(A)} + \frac{1}{q}\right)\right)^k. \tag{$*$}$$

The use of inner probability is needed since $H(A, q, k)$ need not be measurable. We should note that the value of N is irrelevant in $(*)$, so one can deduce from the theorem a similar statement for infinite products. Our proof of this theorem relies on rearrangements. We reduce step by step the problem to a simpler one; after six such reductions, we arrive at a situation that is explicit enough to permit computation. There is a most interesting comparison in [11] of rearrangement methods versus martingale difference sequence (m.d.s.) methods. It seems that overall arrangement methods, while more delicate, are also more powerful, and our present results seems to be an occurrence of this phenomenon. Indeed, we do not see how Theorem 7 (or even Theorem 1) could be proved using m.d.s. Theorem 7 will typically be used for $P(A) \ge \frac{1}{2}$, $k \ge q$, in which case it becomes

$$(1.2) \qquad P_*(H(A, q, k)) \le 1 - (K_0/q)^k = 1 - \exp(-k \log(q/K_0)).$$

Throughout the paper K_0 will keep the meaning of the constant of (1.2).

The presence of the factor $\log q$ is a crucial fact. The estimate of Theorem 7 is all that is needed for the application to Theorems 1 to 5. It might, however, be of interest to know that exact solution of the associated isoperimetric problem, i.e., for $a > 0$, to compute $\inf\{P_*(H(A, q, k));\ P(A) \ge a\}$. Our method does not seem to allow this. As explained after the proof of Theorem 7, it, however, allows one to get an estimate of the best possible order.

At the time of this writing, it seems that Theorem 7, when used as described, captures the exact size of the tails of $\|\sum_{i \le N} X_i\|$ in all the situations. We refer the reader to [10] for its application to strong laws.

We will prove Theorems 1 to 5 in Section 2; and the more delicate Theorem 7 in Section 3.

2. Tails of $\|\sum_{i \le N} X_i\|$. Throughout the paper, K will denote a universal constant, that may vary from line to line.

We now describe a convenient (and usual) setting. Consider two rich enough probability spaces (Ω, Σ, μ) and (Ω', Σ', P') (say copies of $[0,1]$ with Lebesgue measure). Set $P = \mu^N$. Our basic probability space is $(\Omega^N \times \Omega', \Sigma^N \otimes \Sigma', \mathrm{Pr})$ where $\mathrm{Pr} = P \otimes P'$. The variables X_i are defined on Ω^N, in such a way that for $\omega = (\omega_i)_{i \le N}$, $X_i(\omega)$ depends only on ω_i. We consider the sequence ε_i defined on Ω'. So, the sum $\sum_{i \le N} \varepsilon_i X_i$ has value $\sum_{i \le N} \varepsilon_i(\omega') X_i(\omega)$ at the point $(\omega, \omega') \in \Omega^N \times \Omega'$. Conditional expectation with respect to ω' is denoted by E_ε. The usual symmetrization argument shows that to prove Theorems 1 to 5 one can assume that the X_i are symmetric, i.e., that X_i has the same distribution as $-X_i$. So it is enough to prove these theorems when X_i is replaced by $\varepsilon_i X_i$.

We start by describing the basic estimate, that was outlined in the Introduction. Throughout this section, we denote by $Y^{(r)}$ the rth largest term of the sequence $(\|X_i\|)_{i \le N}$ (ties being broken by the index). We set $m = E\|\sum_{i \le N} \varepsilon_i X_i\|$, and we consider the set

$$A = \left\{\omega \in \Omega^N;\ E_\varepsilon \left\| \sum_{i \le N} \varepsilon_i X_i(\omega) \right\| \le 2m \right\},$$

so that $P(A) \ge \frac{1}{2}$.

For $x^* \in B^*$, $\|x^*\| \le 1$, $\omega \in A$, we have $E_\varepsilon|\sum_{i \le N} \varepsilon_i x^*(X_i(\omega))| \le 2m$, so that by Khintchine's inequality we have

$$(2.1) \qquad \sum_{i \le N} x^*(X_i(\omega))^2 = E_\varepsilon \left| \sum_{i \le N} \varepsilon_i x^*(X_i(\omega)) \right|^2 \le 8m^2.$$

(The use of the best constant in Khintchine's inequality is essentially irrelevant.) Consider now $q, k > 0$ and $\omega \in H(A, q, k)$. By definition of $H(A, q, k)$, there exists $x^1, \ldots, x^q \in A$ such that $\mathrm{card}\{i \le N;\ \omega_i \notin \{x_i^1, \ldots, x_i^q\}\} \le k$. So, one can find disjoint sets $I_l \subset \{i \le N;\ \omega_i = x_i^l\}$ and J, with $\mathrm{card}\, J \le k$, such that

$\{1, \dots, N\} = I \cup J$, where $I = \bigcup_{l \le q} I_l$. We have

$$E_\varepsilon \Big\| \sum_{i \in I_l} \varepsilon_i X_i(\omega) \Big\| = E_\varepsilon \Big\| \sum_{i \in I_l} \varepsilon_i X_i(x^l) \Big\| \le 2m,$$

so that

$$E_\varepsilon \Big\| \sum_{i \in I} \varepsilon_i X_i(\omega) \Big\| \le \sum_{l \le q} E_\varepsilon \Big\| \sum_{i \in I_l} \varepsilon_i X_i(\omega) \Big\| \le 2mq. \tag{2.2}$$

From (2.1), we have, for all $x^* \in B^*$, $x^* \le 1$,

$$\sum_{i \in I_l} x^*(X_i(\omega))^2 = \sum_{i \in I_l} x^*(X_i(x^l))^2 \le 8m^2, \tag{2.3}$$

so that

$$\sum_{i \in I} x^*(X_i(\omega))^2 \le 8qm^2.$$

By Theorem 6, and (2.2) for $u > 0$, we have

$$P_\varepsilon\left(\Big\| \sum_{i \in I} \varepsilon_i X_i(\omega) \Big\| \ge 4qm + u \right) \le 4 \exp - \frac{u^2}{64qm^2}. \tag{2.4}$$

We have

$$\begin{aligned} \Big\| \sum_{i \le N} \varepsilon_i X_i(\omega) \Big\| &\le \Big\| \sum_{i \in I} \varepsilon_i X_i(\omega) \Big\| + \sum_{i \in J} \|X_i(\omega)\| \\ &\le \Big\| \sum_{i \in I} \varepsilon_i X_i(\omega) \Big\| + \sum_{r \le k} Y^{(r)}(\omega). \end{aligned}$$

It follows from (2.3) that

$$P_\varepsilon\left(\Big\| \sum_{i \le N} \varepsilon_i X_i(\omega) \Big\| \ge 4qm + \sum_{r \le k} Y^{(r)}(\omega) + u \right) \le 4 \exp - \frac{u^2}{64qm^2}.$$

This is true whenever $\omega \in H(A, q, k)$. From (1.2), we deduce by Fubini's theorem the following.

THE BASIC ESTIMATE. For $k \ge q$, we have

$$\begin{aligned} P\left(\Big\| \sum_{i \le N} \varepsilon_i X_i(\omega) \Big\| \ge 4qm + u + u' \right) &\le 4 \exp - \frac{u^2}{64qm^2} \\ &\quad + \left(\frac{K_0}{q} \right)^k + P\left(\sum_{r \le k} Y^{(r)} > u' \right). \end{aligned} \tag{2.5}$$

REMARK. The inequality (2.3) is rather brutal. When we have information about $\sup_{\|x^*\| \le 1} \sum_{i \le N} Ex^*(X_i)^2$, a better method is available. While this more clever device is essential in the proof of strong laws, as in [10], it is of little relevance here.

We now collect two simple facts.

LEMMA 8. *If* $P(\max_{i\le N}\|X_i\| \ge t) \le \frac{1}{2}$, *then*

$$\sum_{i\le N} P(\|X_i\| \le t) \le 2P\Big(\max_{i\le N}\|X_i\| \ge t\Big).$$

PROOF. We note that for $0 \le x \le \frac{1}{2}$, we have $\exp - 2x \le 1 - x \le \exp - x$, so that

$$\begin{aligned}\exp - 2P\Big(\max_{i\le N}\|X_i\| \le t\Big) &\le 1 - P\Big(\max_{i\le N}\|X_i\| \ge t\Big)\\ &= \prod_{i\le N}(1 - P(\|X_i\| \ge t))\\ &\le \exp - \sum_{i\le N} P(\|X_i\| \ge t).\end{aligned}$$

□

LEMMA 9. $P(Y^{(r)} \ge t) \le (\Sigma_{i\le N}P(\|X_i\| \ge t))^r/r!$.

PROOF. $P(Y^{(r)} \ge t) \le \Sigma\Pi_{j\le r}P(\|X_{i_j}\| \ge t)$, where the summation is over all choices of indexes $1 \le i_1 < i_2 < \cdots < i_r \le N$. Now

$$\begin{aligned}\sum\prod_{j\le l} P\Big(\|X_{i_j}\| > t\Big) &= \frac{1}{r!}\sum_{\text{distinct } i_1,\dots,i_r}\prod_{j\le r} P\Big(\|X_{i_j}\| \ge t\Big)\\ &\le \frac{1}{r!}\sum_{\text{all } i_1,\dots,i_r}\prod_{j\le r} P\Big(\|X_{i_j}\| \ge t\Big)\\ &= \frac{1}{r!}\Big(\sum_{i\le N} P(\|X_i\| \ge t)\Big)^r.\end{aligned}$$

□

PROOF OF THEOREM 1. By homogeneity we can assume $m \le 1$, $\|Y^{(1)}\|_p \le 1$, where $Y^{(1)} = \max_{i\le N}\|X_i\|$. In particular, $P(Y^{(1)} \ge t) \le t^{-p}$, so from Lemmas 8 and 9 we have $P(Y^{(r)} \ge t) \le 2^r t^{-pr}$ for $t \ge 2$. Let us fix $t \ge 4$. We have $P(Y^{(2)} \ge t^{2/3}) \le 4t^{-4p/3}$. Let s be the smallest integer such that $2^{-s} \le t^{-2}$. We have

$$P(Y^{(s)} \ge 4) \le (2\cdot 4^{-p})^s \le 2^{-ps} \le t^{-2p} \le t^{-4p/3}.$$

Hence if we consider the event

$$D = \{Y^{(1)} \le t;\ Y^{(2)} \le t^{2/3},\ Y^{(s)} \le 4\},$$

we have $P(D^c) \le P(Y^{(1)} \ge t) + 5t^{-4p/3}$. Let k be the smallest integer $\ge t$. On D, we have

$$\sum_{r\le k} Y^{(r)} \le t + st^{2/3} + 4(k-1) \le Ct$$

for some constant C, since $2^{-s+1} \ge t^{-2}$, and hence $s \le 1 + K\log t$. Denoting by

$[x]$ the integer part of x, we use (2.5) with $u' = Ct$, $u = t$, $q = [\sqrt{t}]$ to get

$$P\left(\left\|\sum_{i \le N} \varepsilon_i X_i\right\| \ge (C+2)t\right) \le 4\exp - t^{3/2}/K + e^{-t\log(\sqrt{t}/K)}$$
$$+ P\left(\max_{i \le N}\|X_i\| \ge t\right) + Kt^{-4p/3}.$$

Then Theorem 1 follows by a standard computation from the identity

$$\|Z\|_p^p = \int_0^\infty pt^{p-1}P(|Z| > t)\,dt$$

for any r.v. Z. □

We now turn to the proof of Theorem 2. It is an obvious consequence of Lévy's inequality

$$P\left(\max_{k \le N}\left\|\sum_{i \le k} X_i\right\| > t\right) \le 2P\left(\left\|\sum_{i \le N} X_i\right\| > t\right)$$

and of the bound given in the next theorem.

Theorem 10. *There exists a universal constant L with the following property. For each sequence $(X_i)_{i \le N}$ of symmetric bounded r.v. valued in a Banach space B, let us set $m = E\|\sum_{i \le N} X_i\|$, $M = \max_{i \le N}\|X_i\|_\infty$. Then whenever t is large enough that*

$$t \ge Lm, \qquad t \ge LM, \qquad t \ge L\frac{m^2}{M}\left(\log\frac{t}{m}\right)^3, \tag{2.6}$$

we have

$$P\left(\left\|\sum_{i \le N} X_i\right\| \ge t\right) \le \exp - \frac{t}{M}\left(\log\frac{t}{m} - 2\log\log\frac{t}{m} - \frac{2M}{t}\log\frac{t}{m} - L - \log^+\frac{m}{M}\right). \tag{2.7}$$

Proof. The proof will rely on (2.5), and a careful choice of the parameters involved. The point of Theorem 10 is actually to exemplify how accurate (2.5) can be. The precise possible behavior of $\log P(\|\sum_{i \le N} X_i\| \ge t)$ is not known.

Case 1. $m \le M$. We set

$$t_2 = t\left(\log\frac{t}{m}\right)^{-3}, \qquad t_3 = t\left(\frac{m}{M}\right)^{1/2}\left(\log\frac{t}{m}\right)^{-1}, \qquad t_1 = t - t_2 - t_3.$$

We set $k = [t_1/M]$, $q = [t_2/64m]$. We have $q \le k$ and $t_2 \ge 128m$ provided $t \ge Lm$ and L is large enough. We use (2.5) with $u' = t_1$, $u = t_3$, observing that

$P(\Sigma_{r \le k} Y^{(r)} \ge t_1) = 0$ since $kM \le t_1$ by definition of k. We get

$$P\left(\left\| \sum_{i \le N} \varepsilon_i X_i \right\| \ge t\right) \le 4 \exp\left(- \frac{t}{M} \log \frac{t}{m} \right) + \left(\frac{K_0}{q} \right)^k.$$

Since $t_2/64m \ge 2$, we have $q \ge t_2/128m$ and hence

$$\left(\frac{K_0}{q} \right)^k = \exp - k \log \frac{q}{K_0} \le \exp - \frac{1}{M}(t - t_2 - t_3 - M) \log\left(\frac{t_2}{4Km} \right),$$

from which (2.7) follows by computation.

CASE 2. $m \ge M$. We set

$$t_2 = \frac{M}{m} t \left(\log \frac{t}{m} \right)^{-3}, \qquad t_3 = \left(\log \frac{t}{m} \right)^{-1}, \qquad t_1 = t - t_2 - t_3$$

and define $k = [t_1/M]$, $q = [t_2/2m]$. If $t \ge 4(m^2/M)(\log(t/m))^3$, then $t_2/2m \ge 2$ and we proceed as above. □

Using Theorem 2, the proof of the next corollary is identical to that of Corollary 3.3 of [2].

COROLLARY 11. *Let μ be a Lévy measure on a Banach space. Assume that μ is supported by the ball of center* 0 *and radius r. Then, for each $\alpha > 3$ and each* $\tau > 0$,

$$\int \exp\left(r^{-1} \|x\| (\log^+ \|x\| - \alpha \log \log(e + \|x\|)) \right) d(c_t \operatorname{Pois} \mu)(x) < \infty.$$

We now turn to the proof of Theorems 3 and 4. The proof will rely on the following specialization of (2.5): Fixing $q \ge eK_0$, for any $k \ge q$, $t \ge 2qm$, $a > 0$, we have

$$P\left(\left\| \sum_{i \le N} \varepsilon_i X_i \right\| \ge (3 + a)t\right) \le 4 \exp - \frac{t^2}{Kqm^2} + \exp - k + P\left(\sum_{r \le k} Y^{(r)} \ge at \right). \tag{2.8}$$

The difficulty in these proofs will be the control of the last term. In these proofs, we will denote by $K(\alpha)$ a constant depending on α only, that may vary from line to line.

PROOF OF THEOREM 3. We can assume $\|\Sigma_{i \le N} X_i\|_1 \le 1$, $\|\max_{i \le N} \|X_i\| \|_{\Phi_\alpha} \le 1$. Hence $P(\max_{i \le N} \|X_i\| \ge t) \le 2 \exp - t^\alpha$ so that, by Lemma 8, we have $\Sigma_{i \le N} P(\|X_i\| \ge t) \le 4 \exp - t^\alpha$ for $t \ge 2^{1/\alpha}$. To explain the difficulty, we first complete the much easier case $\alpha < 1$. In that case, Lemma 9 shows that $P(Y^{(r)} \ge t) \le (4 \exp - t^\alpha)^r$. For $t \ge K(\alpha)$, we have $4 \exp - t^\alpha \le \exp - (t/2)^\alpha$ so that $P(Y^{(r)} \geq t) \leq \exp - (r^{1/\alpha}t/2)^\alpha$, and hence $P(Y^{(r)} \geq tr^{-1/\alpha}) \leq \exp - (t/2)^\alpha$, provided $t \geq r^{1/\alpha}K(\alpha)$. For $\alpha < 1$ (but not for $\alpha = 1$) the sequence $\sum_{r \geq 1} r^{-1/\alpha}$ is summable. Let S_α be its sum. We thus have

$$P\Big(\sum_{r \leq K(\alpha)t^\alpha} Y^{(r)} \geq tS_\alpha\Big) \leq K(\alpha)t^\alpha \exp - (t/2)^\alpha.$$

If we use (2.8) with $a = S_\alpha$, k of order $K(\alpha)t^\alpha$ we obtain

$$P\Big(\Big\|\sum_{i \leq N} X_i\Big\| \geq K(\alpha)t\Big) \leq K(\alpha)\exp - t^\alpha/K(\alpha)$$

and hence $\|\sum_{i \leq N} X_i\|_{\Phi_\alpha} \leq K(\alpha)$ by standard computations. When $\alpha = 1$ we will also deduce the result in the same manner from (2.8). For this, it is enough to prove that if $\sum_{i \leq N} P(\|X_i\| \geq t) \leq 4\exp - t$ for $t \geq 2$, then, for some universal constant c, we have

$$t \geq c \Rightarrow P\Big(\sum_{r \leq t} Y^{(r)} \geq 22t\Big) \leq \exp - t.$$

We denote by n the largest integer such that $2^n \leq t$.

Suppose that $\sum_{r \leq t} Y^{(r)} \geq 22t$. Let p be the largest integer $\leq n$ such that $Y^{(2^p)} \geq 2$. Then $\sum_{r < 2^{p+1}} Y^{(r)} \geq 20t$. Also $\sum_{r < 2^{p+1}} Y^{(r)} \leq \sum_{0 \leq l \leq p} 2^l Y^{(2^l)}$. Hence we find for $0 \leq l \leq p$ numbers $1 \leq n(l) \leq n + 4$ such that if we set $a_l = 2^{n(l)}$, we have $a_l \leq Y^{(2^l)}$ and that $\sum_{0 \leq l \leq p} 2^l a_l \geq 10t$. Since $Y^{(2^l)} \geq a_l$, we can find disjoint subsets $(I_l)_{0 \leq l \leq p}$ of $\{1, \ldots, N\}$, such that card $I_0 = 1$, card $I_l \geq 2^{l-1}$ for $l \geq 1$ and $\|X_i\| \geq a_l$ if $i \in I_l$. It follows that, if p and $(a_l)_{l=1}^p$ are fixed

$$P\big(\forall\, l \leq p,\, Y^{(2^l)} \geq a_l\big) \leq \sum \prod_{l \leq p} \prod_{i \in I_l} P(\|X_i\| \geq a_l),$$

where the summation is taken over all choices $I_0, \ldots, I_p$ of subsets of $\{1, \ldots, N\}$ such that card $I_l = \max(1, 2^{l-1})$. Hence we have

$$\begin{aligned} P\big(\forall\, l \leq p,\, Y^{(2^l)} \geq a_l\big) &\leq \prod_{l \leq p} \Big[\sum_{i \leq N} P(\|X_i\| \geq a_l)/2^{l-1}!\Big]^{2^{l-1}} \\ &\leq 4^{2^n} \exp\Big(-\sum_{l \leq p} 2^{l-1} a_l\Big) \qquad (2.9) \\ &\leq 4^{2^n} \exp - 5t. \end{aligned}$$

There are at most $(n + 5)^{(n+1)}$ possible choices for the sequence $(a_l)_{0 \leq l \leq p}$, so that for t large enough

$$\begin{aligned} P\Big(\sum_{r \leq t} Y^{(r)} \geq at\Big) &\leq (n + 5)^{(n+1)} 4^{2^n} \exp - 5t \\ &\leq e^{4t} \exp - 5t \leq \exp - t, \end{aligned}$$

since for t large enough $(n + 5)^{(n+1)} \leq 4^{2^n}$, and $4^{2^{n+1}} \leq e^{2^{n+2}} \leq e^{4t}$. □

We now investigate by what kind of Orlicz function ψ one can replace Φ_1 in Theorem 3, and more specifically, we investigate how fast these functions can grow. Interestingly enough, it turns out that these functions can grow faster than exponentially, although in a rather slow way.

PROPOSITION 12. *Consider an Orlicz function ψ such that for x large enough $\Psi(x) = \exp(x\xi(x))$, where ξ is nondecreasing.*

(a) *If for some K and for all i.i.d. sequences $(X_i)_{i\le N}$ of symmetric real-valued r.v. we have*

$$\Big\| \sum_{i\le N} X_i \Big\|_{\psi} \le K\bigg(\Big\| \sum_{i\le N} X_i \Big\|_1 + \|\max_{i\le N}|X_i|\,\|_{\psi}\bigg), \tag{2.10}$$

then for u large enough we have

$$\xi(e^u) \le L\xi(u) \tag{2.11}$$

for some constant L.

(b) *Conversely, if* (2.11) *holds for all u large enough, then for each sequence $(X_i)_{i\le N}$ of independent r.v. valued in a Banach space, we have*

$$\Big\| \sum_{i\le N} X_i \Big\|_{\psi} \le K\bigg(\Big\| \sum_{i\le N} X_i \Big\|_1 + \|\max_{i\le N}|X_i|\,\|_{\psi}\bigg).$$

PROOF. (a) Let $u, N > 0$, such that $N\psi(u) \ge 1$, and consider a r.v. X such that $P(X = \pm u) = (2N\psi(u))^{-1}$, $P(X = 0) = 1 - (N\psi(u))^{-1}$. Consider an i.i.d. sequence $(X_i)_{i\le N}$ distributed like X. Then $\max_{i\le N}|X_i| \le u$, $P(\max_{i\le N}|X_i| \ne 0) \le \psi(u)^{-1}$, so that $\|\max_{i\le N}|X_i|\,\|_{\psi} \le 1$. Also $\|\sum_{i\le N} X_i\|_1 \le u/\psi(u)$, and we take u large enough that this is less than or equal to 1. Hence, by (2.10), we have $\|\sum_{i\le N} X_i\|_{\psi} \le 2K$. We observe that

$$P\bigg(\sum_{i\le N} X_i = Nu \bigg) = (2N\psi(u))^{-N},$$

so that

$$(2N\psi(u))^{-N}\psi\left(\frac{Nu}{2K}\right) \le E\left(\psi\left(\frac{1}{2K}\sum_{i\le N} X_i\right)\right) \le 1.$$

Thus we have $\psi(Nu/2K) \le (2N\psi(u))^N$. We remember that $\psi(u) = \exp(u\xi(u))$, so that, taking logarithms,

$$\frac{Nu}{2K}\xi\left(\frac{Nu}{2K}\right) \le N\log 2N + Nu\xi(u).$$

Consider now $u \ge 8K$. We can take N such that $2Ke^u u^{-1} \le e^u/4 \le N \le e^u/2$, so that $2N \le e^u$ and $e^u \le Nu/2K$. It follows that

$$\frac{ue^u}{8K}\xi(e^u) \le e^u u + ue^u\xi(u).$$

This implies (2.11).

(b) The proof follows the proof of Theorem 3 in the case $\alpha = 1$; and we keep the same notation. We denote universal con-stants by $C_1, C_2, \dots$. It is enough to prove that for $t \geq t_0$ if $\sum_{i \leq N} P(\|X_i\| \geq t) \leq 2\exp - t\xi(t)$, then we have

$$t \geq C_1 \Rightarrow P\left(\sum_{r \leq t\xi(t)} Y^{(r)} \geq C_2 t \right) \leq \exp - t\xi(t).$$

We proceed as in the proof of Theorem 3, denoting now by n the largest integer for which $2^n \leq t\xi(t)$. We observe that $(2^{l-1})! \geq e^{2^{l-1}(l\log 2 - C_3)^+}$, so that, by Lemma 9,

$$P\left(\forall\, l \leq p,\, Y^{(2^l)} \geq a_l\right) \leq 4^{2^n} \exp\left(- \sum_{l \leq p} \left(2^{l-1} a_l \xi(a_l) + 2^{l-1}(l\log 2 - C_3)^+\right)\right).$$

We now show the following, where L is the constant of (2.11). If t (hence n) is large enough

$$\sum_{l \leq p} 2^{l-1} a_l \geq 10L^2 t \Rightarrow \sum_{l \leq p} \left(2^{l-1} a_l \xi(a_l) + 2^{l-1}(l\log 2 - C_3)^+\right) \geq 5t\xi(t). \tag{2.12}$$

We can and do assume $L \geq 100$.

Case 1. $5L2^{n-p} \geq n$. Then we have $\sum 2^{l-1} a_l \leq 5Lt$, where the summation is over those $l \leq p$ for which $a_l \leq 5Lt2^{-p}$. Thus $\sum 2^{l-1} a_l \geq 5L^2 t$, where the summation is over those $l \leq p$ for which $a_l \geq 5Lt2^{-p}$. Now

$$5Lt2^{-p} \geq tn2^{-n} \geq n/\xi(t).$$

Since (2.12) implies that $\xi(t)$ grows very slowly, and in particular that $\xi(t) \leq \log t$ for t large enough, for t large enough, we have $n/\xi(t) \geq \log(\log t)$. Thus $\xi(5Lt2^{-p}) \geq L^{-2}\xi(t)$, and (2.12) follows.

Case 2. $5L2^{n-p} < n$. Then we have $p \geq n/2$ for n large enough; so, for n large enough, we have

$$2^{p-1}(p\log 2 - C_3)^+ \geq 2^{p-3} n \log 2 \geq 2^{n+1}\left(\frac{5L\log 2}{16}\right) \geq 5t\xi(t)$$

and this proves (2.12).

The conclusion then follows as in the case of Theorem 3. □

Proof of Theorem 4. We set $d_i = \|X_i\|_{\phi_\alpha}$ so that

$$P(\|X_i\| \geq t) \leq 2\exp - (t/d_i)^\alpha.$$

We can assume $\|\sum_{i \leq N} X_i\|_1 \leq 1$, $\sum_{i \leq N} d_i^\beta \leq 1$. There is no loss of generality to assume the sequence $(d_i)_{i \leq N}$ nonincreasing. Hence $\sum_{2^i \leq N} 2^i d_{2^i}^\beta \leq 2$. We can find a sequence $\gamma_i \geq 2^i d_{2^i}^\beta$ such that $\gamma_i \geq \gamma_{i+1}/\sqrt{2} \geq \gamma_i/2$ for $i \geq 1$ and $\sum \gamma_i \leq 20$ (e.g., $\gamma_i = \sum_{j \geq 1} \delta_j 2^{-|j-i|/2}$ where $\delta_j = 2^j d_{2^j}^\beta$). Let $c_i = (2^{-i}\gamma_i)^{1/\beta}$, so that $c_{i+1}^\alpha \leq c_i^\alpha 2^{-\alpha/2\beta}$, $c_i^\alpha \leq c_{i+1}^\alpha 2^{3\alpha/2\beta}$. Then $\sum 2^i c_i^\beta \leq 20$ and $d_j \leq c_i$ for $j \geq 2^i$. We observe that $\sum_{l \geq 2} 4c_l (4\log 4^{l+1})^{1/\alpha} \leq C(\alpha) < \infty$.

As in the proof of Theorem 3, using (2.5) it is enough to show that for some constants a, c, depending on α only, we have

$$t \geq c \Rightarrow P\Big(\sum_{r \leq t^\alpha} Y^{(r)} \geq at\Big) \leq \exp - t^\alpha.$$

Since $P(Y^{(1)} \geq t) \leq 2\exp - t^\alpha$, it is actually enough to find s, a, c such that

$$t \geq c \Rightarrow P\Big(\sum_{2^s \leq r \leq t^\alpha} Y^{(r)} \geq at\Big) \leq \exp - t^\alpha.$$

Fix t, and denote by n the largest integer such that $2^n \leq t^\alpha$. Suppose that $\sum_{r \leq t^\alpha} Y^{(r)} \geq at$ so that $\sum_{0 \leq l \leq n} 2^l Y^{(2^l)} \geq at$.
Let

$$L = \left\{s \leq l \leq n;\ Y^{(2^l)} \geq 2c_l\left(4\log 4^{l+1}\right)^{1/\alpha}\right\}.$$

Then $\sum_{l \in L} 2^l Y^{(2^l)} \geq at - C(\alpha) \geq at/2$ for t large enough. For $l \in L$, we can find a number a_l, of the form 2^m ($m \in \mathbb{Z}$) such that $Y^{(2^l)} \geq a_l$ and $a_l \geq c_l(4\log 4^{l+1})^{1/\alpha}$ and $\sum_{l \in L} 2^l a_l \geq at/4$. We can find disjoint subsets $(J_l)_{l \in L}$ of $\{1, \dots, N\}$ such that $\operatorname{card} J_l \geq 2^{l-1}$ and $\|X_i\| \geq a_l$ for $i \in J_l$. Set $I_l = J_l \setminus \{1, \dots, 2^{l-2}\}$. We have

$$P\left(\forall\, l \in L,\ Y^{(2^l)} \geq a_l\right) \leq \sum \prod_{l \in L} \prod_{i \in I_l} P(\|X_i\| \geq a_l),$$

where the summation is taken over all possible choices of $(I_l)_{l \in L}$. Hence

$$P\left(\forall\, l \in L, Y^{(2^l)} \geq a_l\right) \leq \prod_{l \in L}\Bigg(\sum_{2^{l-2} < i \leq N} P(\|X_i\| \geq a_l)\Bigg)^{2^{l-2}}.$$

We have $P(\|X_i\| \geq a_l) \leq 2\exp - (a_l/d_i)^\alpha$ so

$$\sum_{2^{l-2} < i \leq N} P(\|X_i\| \geq a_l) \leq \sum_{j \geq l-2} 2^{j+1} \exp - \left(a_l/c_j\right)^\alpha.$$

Since $l \in L$, we have

$$\left(a_l/c_l\right)^\alpha \geq \left(a_l/c_l\right)^\alpha/2 + \log 4^{j+2}$$

and since $c_{j+1}^\alpha \leq c_j^\alpha 2^{-\alpha/2\beta}$, $c_j^\alpha \leq c_{j+1}^\alpha 2^{3\alpha/2\beta}$ we clearly have

$$\left(a_l/c_j\right)^\alpha \geq \left(a_l/c_l\right)^\alpha/4 + \log 4^{j+2}$$

for $j \geq l - 2$, provided s has been taken large enough. It follows that

$$\sum_{j \geq l-2} 2^{j+1} \exp - \left(\frac{a_l}{c_j}\right)^\alpha \leq \sum_{j \geq l-2} 2^{-j-1} \exp - \frac{1}{4}\left(\frac{a_l}{c_l}\right)^\alpha$$

$$\leq \exp - \frac{1}{4}\left(\frac{a_l}{c_l}\right)^\alpha.$$

Hence we have

$$P\Big(\forall\, l \le n,\, Y^{(2^l)} \ge a_l\Big) \le \exp -\frac{1}{8}\sum_{l\in L} 2^l\bigg(\frac{a_l}{d_l}\bigg)^{\alpha}.$$

By Holder's inequality, we have

$$\frac{at}{4} \le \sum_{l\in L} 2^l a_l \le \left(\sum_{l\in L} 2^l\bigg(\frac{a^l}{d_l}\bigg)^{\alpha}\right)^{1/\alpha}\bigg(\sum_{l\in L} 2^l c_l^{\beta}\bigg)^{1/\beta},$$

so that $\sum_{l\in L}2^l(a_l/d_l)^{\alpha} \ge t^{\alpha}a^{\alpha}/80$, and, if we take $a = 1280$, we have

$$P\Big(\forall\, l \le n,\, Y^{(2^l)} \ge a_l\Big) \le \exp - 2t^{\alpha}.$$

The number of possible choices for the sequence $(a_l)_{l\le n}$ is $\le \exp t^{\alpha}$ for c large enough, and this concludes the proof. □

Proof of Theorem 5. During this proof, K will denote a constant depending only on α and r, that may vary in each occurrence.

(a) There is no loss of generality to assume that the sequence $\|X_i\|_{\Phi_\alpha}$ decreases, so that $\|X_i\|_{\Phi_\alpha} \le Mi^{-1/r}$. If $Y^{(i)} \ge t$, then at least $i - [i/2]$ of the numbers $\|X_j\|$, $j > i/2$, are $> t$. By Lemma 9, we have

$$\begin{aligned}P(Y^{(i)} \ge t) &\le \bigg(\sum_{j>i/2} P\big(\|X_j\| \ge t\big)\bigg)^{i-[i/2]}\Big/(i - [i/2])!\\ &\le \left(\sum_{j>i/2} 2\exp - \bigg(\frac{tj^{1/r}}{M}\bigg)^{\alpha}\right)^{i-[i/2]}\Big/(i - [i/2])!.\end{aligned}$$

For $t \ge KMi^{-1/r}$ we see easily that

$$P(Y^{(i)} \ge t) \le \exp - \bigg(\frac{t\,i^{1/r+1/\alpha}}{KM}\bigg)^{\alpha}$$

or

$$P(Y^{(i)} \ge ti^{-1/r-1/\alpha}) \le \exp - \bigg(\frac{t}{KM}\bigg)^{\alpha},$$

provided $t \ge Ki^{1/\alpha}M$. We observe that $\sum_{i\le k} i^{-1/r-1/\alpha} \le Kk^{1/s-1/\alpha}$. Thus

$$P\bigg(\sum_{i\le k} Y^{(i)} \ge tk^{1/s-1/\alpha}\bigg) \le k\exp - \bigg(\frac{t}{KM}\bigg)^{\alpha},$$

provided $t \ge Kk^{1/\alpha}M$, and

$$P\bigg(\sum_{i\le k} Y^{(i)} \ge t\bigg) \ge k\exp - \bigg(\frac{t}{KMk^{1/s-1/\alpha}}\bigg)^{\alpha},$$

provided $t \ge Kk^{1/s}M$. Then Theorem 5 follows from (2.5), by choosing $u = u' = t$,

$k \sim (t/M)^s(\log q)^{-s/\alpha}$ and

$$q = \min\left[\frac{t}{Km}, \frac{t^{2-s}M^s}{m^2}\left(\log\frac{t^{s-2}M^s}{m^2}\right)^{-1+s/\alpha}\right] \quad \text{if } \frac{t^{2-s}M^2}{m^2} \ge eK,$$

$$= eK \quad \text{otherwise.}$$

(b) Taking k of order $(t/M)^s$, one first shows as above that

$$P\left(\sum_{i\le k} Y^{(i)} \ge t\right) \le K\exp-\left(\frac{t}{KM}\right)^s,$$

$$P\left(\sum_{i>k} Y^{(i)2} \ge Kt^2k^{-1}\right) \le K\exp-\left(\frac{t}{KM}\right)^s.$$

Denote by J the set of indexes for which $\|X_i\|$ is one of $Y^{(i)}$, $i \le k$. We observe that

$$\left\|\sum_{i\in I}\varepsilon_i X_i(\omega)\right\| \le \sum_{i\le k} Y^{(i)}(\omega) + \left\|\sum_{i\in I\setminus J}\varepsilon_i X_i(\omega)\right\|.$$

We now use the notation of (2.4), and let $\omega \in A$. When $\sum_{i\le k} Y^{(i)} \le t$, $\sum_{i>k} Y^{(i)2} \le Kt^2k^{-1}$, we have, by Theorem 6,

$$P_\varepsilon\left(\left\|\sum_{i\in I}\varepsilon_i X_i(\omega)\right\| \ge 4qm + 2t\right) \le 4\exp-\frac{t^2}{Kt^2k^{-1}}$$

$$\le 4\exp-\left(\frac{t}{KM}\right)^s.$$

Thus, arguing as in the proof of (2.5), we get

$$P\left(\left\|\sum_{i\le N}\varepsilon_i X_i(\omega)\right\| \ge 4qm + 3t\right) \le \left(\frac{K_0}{q}\right)^k + K\exp-\left(\frac{t}{KM}\right)^s,$$

from which the result follows by fixing $q \ge 2K_0$. □

3. Proof of Theorem 7.

STEP 1. We reduce the proof to the case where $\Omega = [0,1]$, μ is Lebesgue's measure λ and A is compact. The argument is a routine technicality. This step is needed only to be able to state Theorem 4 for general measure spaces, while the case $\Omega = [0,1]$, $\mu = \lambda$ is sufficient for applications. The reader who is satisfied with that special case should hence go directly to Step 2. We can assume that Σ_0 is the completion of a countably generated σ-algebra Σ. There exists a measurable map ϕ from Ω to $[0,1]$, such that $\Sigma_0 = \{\phi^{-1}(B);\ B \text{ Borel}\}$. Denote by ψ the product map from Ω^N to $[0,1]^N$ and set $\nu = \phi(\mu)$, $Q = \nu^N$, so $Q = \psi(P)$. We note that $\nu^*(\phi(\Omega)) = 1$. It is enough to prove $(*)$ when $A \in \Sigma_0^N$. Then we have $A = \psi^{-1}(B)$ for some Borel set B of $[0,1]^N$, and $P(A) = Q(B)$.

Our first task is to construct a Borel subset $B' \subset B$, such that $Q(B') = Q(B)$, with the following property: Whenever $x \in B'$ and $I = \{i \le N;\ x_i \in \phi(\Omega)\}$, there exists $y \in B$, with $y_i = x_i$ whenever i belongs to I, and $y_i \in \phi(\Omega)$ for $i \le N$. For each partition $\{1, \dots, N\} = I \cup J$, we can view $([0,1]^N, Q)$ as a product $[(0,1]^I \times [0,1]^J, \nu^I \otimes \nu^J)$. Denote by $C_{I,J}$ the set $D \times [0,1]^J$, where

$$D = \left\{u \in [0,1]^I;\ \nu^J\left(\left\{v \in [0,1]^J;\ (u,v) \in B\right\}\right) = 0\right\}.$$

By Fubini's theorem, we have $Q(C_{I,J}) = 0$. Let $B' = B \setminus UC_{I,J}$, where the union is taken over all possible partitions of $\{1, \dots, N\}$. For x in B', let $I = \{i\colon x_i \in \phi(\Omega)\}$, $u = (x_i)_{i \in I}$. By definition of B', we have

$$\nu^J\left(\left\{v \in [0,1]^J;\ (u,v) \in B\right\}\right) > 0.$$

Also, we have $(\nu^J)^*(\phi(\Omega)^J) = 1$. So, there is v in $\phi(\Omega)^J$ such that $y = (u,v)$ belongs to B. This is what we wanted to achieve. It is now routine to check that

$$\psi^{-1}(H(B', q, k)) \subset H(A, q, k),$$

and hence $P(H(A,q,k)) \ge Q(H(B',q,k))$. So, to prove inequality $(*)$ for A, it is enough to prove it for B'. Hence we have reduced the problem to the case where $\Omega = [0,1]$.

To reduce to the case where A is compact, one notices that $H(A,q,k)$ is an increasing function of A, and that $P(A) = \sup\{P(K);\ K \subset A,\ K \text{ compact}\}$. To reduce to the case where $\mu = \lambda$, we consider a measurable map $\phi\colon [0,1] \to [0,1]$ such that $\nu = \phi(\lambda)$. If ψ denotes the product map, it is easy to see that for each subset A of $[0,1]^N$, we have

$$\psi\left(H\left(\psi^{-1}(A), q, k\right)\right) \subset H(A, q, k).$$

This shows that it is enough to prove $(*)$ for $\psi^{-1}(A)$ instead of A, and λ instead of ν.

STEP 2. We reduce to the case where $\Omega = [0,1]$, $\mu = \lambda$, A is compact and moreover satisfies the condition

$$\text{(H)} \qquad \forall\, x \in A, \forall\, y \in [0,1]^N, \forall\, i \le N,\ y_i \ge x_i \Rightarrow y \in A.$$

The proof relies on a simple rearrangement. Let A be any compact subset of $[0,1]^N$. Fix $1 \le j \le N$; to simplify the notation, we assume $j = N$. For $z \in [0,1]^{N-1}$, we define $t(z) \in [0,1]$ by the relation

$$\lambda([t(z), 1]) = \lambda(\{u \in [0,1];\ (z,u) \in A\}),$$

and we define

$$T(A) = \bigcup_{z \in [0,1]^{N-1}} \{z\} \times [t(z), 1].$$

Since $t(z)$ is an u.s.c. function of z, $T(A)$ is compact. The essential step of the proof will be to show that

$$(3.1) \qquad P(H(T(A), q, k)) \le P(H(A, q, k)),$$

where $P = \lambda^N$. Once this is done, we notice that if for some subset J of $\{1,\dots,N\}$, A has the property

$$(H(J)): \forall\, x \in A, \forall\, y \in [0,1]^N, \forall\, i \notin J,\ y_i = x_i, \forall\, i \in J,\ y_j \geq x_j \Rightarrow y \in A,$$

then $T(A)$ has $H(J \cup \{N\})$. So, if we successively apply the transformation T to each coordinate, we transform A into a compact subset of $[0,1]$ that satisfies (H) and has the same measure; and (3.1) shows that we have not increased $P(H(A,q,k))$. To prove (3.1), we fix z in $[0,1]^{N-1}$. We set

$$U = \{u \in [0,1];\ (z,u) \in H(A,q,k)\},$$
$$V = \{u \in [0,1];\ (z,u) \in H(T(A),q,k)\}.$$

It is enough to prove that $\lambda(V) \leq \lambda(U)$. We use the notation $(z,u)_N = u$, $(z,u)_i = z_i$ for $i < N$.

CASE 1. The following holds:

$$(3.2)\quad \exists\, x^1,\dots,x^q \in A, \qquad \operatorname{card}\{i \leq N-1;\ z_i \notin \{x_i^1,\dots,x_i^q\} \leq k-1\}.$$

In that case we have $U = V = [0,1]$.

CASE 2. (3.2) fails. Let $v = \inf V$. We have $\lambda(V) \leq 1 - v$, so it is enough to show that $\lambda(U) \geq 1 - v$. Since $H(T(A),q,k)$ is compact, $v \in V$, so there exist $x^1,\dots,x^q$ in $T(A)$ such that

$$\operatorname{card}\{i \leq N;\ (z,v)_i \notin \{x_i^1,\dots,x_i^q\}\} \leq k.$$

By definition of $T(A)$, there exist $y^1,\dots,y^q$ in A with $y_i^l = x_i^l$ for $l \leq q$, $i \leq N-1$. Since (3.2) fails,

$$\operatorname{card}\{i \leq N-1; z_i \notin \{x_i^1,\dots,x_i^q\}\} = \operatorname{card}\{i \leq N-1; z_i \notin \{y_i^1,\dots,y_i^q\}\} \geq k.$$

So we have $v \in \{x_N^1,\dots,x_N^q\}$, say $v = x_N^1$. Let $y = (\ x_i^1)_{i \leq N-1}$. If $(y,t) \in A$, then

$$\operatorname{card}\{i \leq N;\ (z,t)_i \notin \{(y,t)_i, x_i^2,\dots,x_i^q\}\}$$
$$= \operatorname{card}\{i \leq N-1;\ z_i \notin \{x_i^1,\dots,x_i^q\}\} \leq k,$$

so we have $(z,t) \in (A,q,k)$. Since $(y,v) = x^1 \in T(A)$, the definition of $T(A)$ shows that $\lambda(\{t;\ (y,t) \in A \geq 1 - v\})$, so $\lambda(U) \geq 1 - v$. The step is complete.

STEP 3. We show that it is enough to consider the case where $\Omega = [0,1]$, $\mu = \lambda$, A is compact, satisfies (H) and is measurable for the algebra $\otimes_{i \leq N}\mathscr{B}_i$, where $\mathscr{B}_i$ is the Σ-algebra generated by $[0,\alpha_i[, [\alpha_i,1]$ for some α_i in $[0,1]$. The reduction is done "one component at a time." If A satisfies (H), is compact and is measurable for $\otimes_{i \leq N}\mathscr{B}_i$, where $\mathscr{B}_N$ consists of all the Borel sets, we find a compact subset B of $[0,1]^N$ that satisfies (H), such that $P(B) = P(A)$, $P(H(B,q,k)) \leq P(H(A,q,k))$ and that B is measurable for $\otimes_{i \leq N}\mathscr{B}_i'$, where $\mathscr{B}_i' = \mathscr{B}_i$ for $i < N$, while $\mathscr{B}_N'$ is generated by $[0,\alpha[, [\alpha,1]$ for some α in $[0,1]$. The proof will use the following elementary lemma.

LEMMA 13. *Let f, g be two measurable functions from $[0,1]$ to $[0,1]$. Let $a = \int f\,d\lambda$, $b = \int g\,d\lambda$. Then we find s, t, α in $[0,1]$ such that $f(s) \le a \le f(t)$ and*

$$a = \alpha f(s) + (1-\alpha) f(t), \qquad \alpha g(s) + (1-\alpha) g(t) \le b.$$

PROOF. Set

$$U = \{s \in [0,1];\ f(s) < a\}, \qquad V = \{t \in [0,1];\ f(t) > a\}.$$

Suppose that the conclusion fails. For s in U, t in V, we have

$$\frac{g(t)-b}{f(t)-a} > \frac{b-g(s)}{a-f(s)}, \tag{3.3}$$

so for some number c we have

$$\frac{g(t)-b}{f(t)-a} \ge c \ge \frac{b-g(s)}{a-f(s)},$$

whenever $s \in U$, $t \in V$, i.e., $g(x) - b \ge c(f(x) - a)$ whenever $x \in U \cup V$. Also we must have $g(x) > b$ whenever $f(x) = a$, so actually $g(x) - b - c(f(x) - a) \ge 0$ for all x. Since the integral of this function is 0, we have $g(x) - b = c(f(x) - a)$ a.e. Since either $\lambda(U) > 0$, $\lambda(V) > 0$ or $\lambda(U) = 0$, $\lambda(V) = 0$, this contradicts either (3.3) or the fact that $g(x) > b$ when $f(x) = a$. The proof is complete. □

For t in $[0,1]$, we set

$$A_t = \{x \in [0,1]^{N-1};\ (x,t) \in A\}.$$

Since A satisfies (H), so does A_t. Also $A_t \subset A_u$ for $t \le u$. We set

$$H(A,q,k)_t = \{y \in [0,1]^{N-1};\ (y,t) \in H(A,q,k)\}.$$

For subsets D_1, D_2 of $[0,1]^{N-1}$, we set

$$G(D_1, D_2, q, k) = \Big\{y \in [0,1]^{N-1};\ \exists\, x^1,\dots,x^{q-1} \in D_1,\ \exists\, x^q \in D_2,$$
$$\operatorname{card}\big\{i \le N-1;\ y_i \notin \{x_i^1,\dots,x_i^q\}\big\} \le k\Big\}.$$

LEMMA 14. *We have*

$$H(A,q,k)_t = H(A_1,q,k-1) \cup G(A_1, A_t, q, k).$$

In particular, $H(A,q,k)_t$ depends only on A_1 and A_t.

PROOF. Let $(y,t) \in H(A,q,k)_t$. By definition, there is $x^1,\dots,x^q \in A$ such that

$$\operatorname{card}\big\{i \le N;\ (y,t)_i \notin \{x_i^1,\dots,x_i^q\}\big\} \le k.$$

CASE 1. $t \notin \{x_N^1, \dots, x_N^q\}$. Then

$$\operatorname{card}\left\{i \le N-1;\ y_i \notin \left\{x_i^1, \dots, x_i^q\right\}\right\} \le k-1.$$

Since for $l \le q$ we have $(x_i^l)_{i \le N-1} \in A_{x_N^l} \subset A_1$, we have $y \in H(A_1, q, k-1)$.

CASE 2. $t \in \{x_N^1, \dots, x_N^q\}$, say $t = x_N^q$. Then

$$y \in \Big\{\gamma \in [0,1]^{N-1};\ \exists\, x^1, \dots, x^q \in A,\ x_N^q = t,$$
$$\operatorname{card}\left\{i \le N-1;\ \gamma_i \notin \left\{x_i^1, \dots, x_i^q\right\}\right\} \le k\Big\} \subset G(A_1, A_t, q, k).$$

We have shown that

$$H(A,q,k)_t \subset H(A_1, q, k-1) \cup G(A_1, A_t, q, k).$$

The reverse inclusion is obvious, and the lemma is proved. □

We set $f(t) = \lambda^{N-1}(A_t)$, so $a = \int f\, d\lambda = P(A)$. We set

$$g(t) = \lambda^{N-1}(G(A_1, A_t, q, k) \setminus H(A_1, q, k-1)),$$

so, if $b = \int g\, d\lambda$, we have by Lemma 14:

$$P(H(A,q,k)) = \int_0^1 \lambda^{N-1}(H(A,q,k)_t)\, dt = \lambda^{N-1}(H(A_1, q, k-1)) + b. \tag{3.4}$$

According to Lemma 13 there exist s, t, α in $[0,1]$, $s \le t$, such that $\alpha f(s) + (1-\alpha) f(t) = a$ and $\alpha g(s) + (1-\alpha) g(t) \le b$. We set $B = [0, \alpha[\times A_s \cup [\alpha, 1] \times A_t$, so we have $B_u = A_s$ for $0 \le u < \alpha$ and $B_u = A_t$ for $\alpha \le u \le 1$. Since $A_s \subset A_t$, B is compact; since $A_s \subset A_t$ and both A_s, A_t satisfy (H), so does B. Also

$$P(B) = \alpha \lambda^{N-1}(A_s) + (1-\alpha)\lambda^{N-1}(A_t) = \alpha f(s) + (1-\alpha) f(t) = a = P(A).$$

For $0 \le u < \alpha$, we have, by Lemma 14,

$$H(B,q,k)_u = H(B_1, q, k-1) \cup G(B_1, B_u, q, k)$$
$$\subset H(A_1, q, k-1) \cup G(A_1, A_s, q, k),$$

so

$$\lambda^{N-1}(H(B,q,k)_u) \le \lambda^{N-1}(H(A_1, q, k-1)) + g(s).$$

For $\alpha \le u \le 1$, in a similar way we have

$$\lambda^{N-1}(H(B,q,k)_u) \le \lambda^{N-1}(H(A_1, q, k-1)) + g(t).$$

It follows that

$$\begin{aligned} P(H(B,q,k)) &= \int_0^1 \lambda^{N-1}(H(B,q,k)_u)\, du \\ &\le \lambda^{N-1}(H(A_1, q, k-1)) + \alpha g(s) + (1-\alpha) g(t) \\ &\le \lambda^{N-1}(H(A_1, q, k-1)) + b \\ &= P(H(A,q,k)). \end{aligned}$$

This concludes this step.

STEP 4. We reduce now the problem to the case where $\Omega = \{0,1\}^N$, $\mu = \alpha\delta_0 + (1-\alpha)\delta_1$, where $\alpha > 0$ is small enough to satisfy $\alpha q \le 1$ and $1 - (1-\alpha)^q \ge \alpha q/2$, and where A satisfies (H) (where the condition $y \in [0,1]$ is replaced by $y \in \{0,1\}$).

Consider a sequence $(\alpha_i)_{i \le N}$, $0 \le \alpha_i \le 1$, and the measure $Q = \otimes_{i \le N} \eta_i$ on $\{0,1\}^N$, where $\eta_i = \alpha_i \delta_0 + (1-\alpha_i)\delta_1$. The result of Step 3 can be interpreted as follows:

To prove $(*)$, it is enough to prove that

$$Q(H(A,q,k)) \ge 1 - \big(K((1/k)\log(1/Q(A)) + 1/q)\big)^k, \tag{3.5}$$

whenever $A \subset \{0,1\}^N$ satisfies (H), and for any choice of the α_i. The quantities $Q(A)$ and $Q(H(A,q,k))$ are continuous functions of the numbers α_i; so it is enough to prove (3.5) in the case where $1 - \alpha_i = (1-\alpha)^{n_i}$, for some $n_i \in \mathbb{N}$, and $\alpha > 0$, α being small enough that $\alpha q \le 1$, $1 - (1-\alpha)^q \ge \alpha q/2$. Consider disjoint sets $(J_i)_{i \le N}$, with card $J_i = n_i$, and let $J = \bigcup_{i \le N} J_i$. We consider the map ϕ from $\{0,1\}^J$ to $\{0,1\}^N$ given, for x in $\{0,1\}^J$, by $\phi(x) = (\inf_{j \in J_i} x_j)_{i \le N}$. If $\mu = \alpha\delta_0 + (1-\alpha)\delta_1$ and $P = \mu^J$, the condition $1 - \alpha_i = (1-\alpha)^{n_i}$ implies that $\phi(P) = Q$. Let $B = \phi^{-1}(A)$, so $P(B) = Q(A)$. Since ϕ is increasing when both $\{0,1\}^J$ and $\{0,1\}^N$ are provided with the product order, we see that B satisfies (H) since A does. To conclude, it is enough to show that $\phi(H(B,q,k)) \subset H(A,q,k)$ [so that $Q(H(A,q,k)) \ge P(H(B,q,k))$]. We first show that we have

$$H(A,q,k) = \Big\{ y \in \{0,1\}^N;\ \exists\, x^1,\dots,x^q \in A, \\ \operatorname{card}\Big\{ i \le N;\ y_i < \inf_{l \le q} x_i^l \Big\} \le k \Big\}. \tag{3.6}$$

Indeed, the inclusion $\subset$ is clear. For the other inclusion, take $y \in \{0,1\}^N$ and $x^1,\dots,x^q \in A$ with $\operatorname{card}\{i \le N;\ \{y_i < \inf_{l \le q} x_i^l\} \le k\}$. Define z^l by $z_i^l = \max(x_i^l, y_i)$ for $i \le N$, $l \le q$, so that $z^l \in A$ from (H), and $\operatorname{card}\{i \le N;\ y_i \notin \{z_i^1,\dots,z_i^q\}\} \le k$ so that $y \in H(A,q,k)$. Now let y in $H(B,q,k)$; there exist $x^1,\dots,x^q \in B$ and $I \subset J$, with card $I \le k$, such that $y_i \in \{x_i^1,\dots,x_i^q\}$ for $i \notin I$. Let $L = \{i \le N;\ J_i \cap I \ne \varnothing\}$, so card $L \le k$. Fix $i \le N$ with $i \notin L$, so $J_i \cap I = \varnothing$. Let $j \in J_i$ with $y_j = \inf_{p \in J_i} y_p$. We have $j \notin I$, so $y_j = x_j^p$ for some $p \le q$. So we have

$$\phi(y)_i = y_j = x_j^p \ge \phi(x^p)_i \ge \inf_{l \le q} \phi(x^l)_i.$$

This shows that

$$\operatorname{card}\Big\{ i \le N;\ \phi(y)_i < \inf_{l \le q} \phi(x^l)_i \Big\} \le \operatorname{card} L \le k.$$

Since $\phi(x^l) \in A$, (3.6) shows that $\phi(y) \in H(A,q,k)$. This concludes this step.

STEP 5. In this step, we are going to replace the problem by a simpler one; while this is adequate to prove $(*)$, there is no way by this method to solve exactly the underlying isoperimetric problem, i.e., to compute $\inf\{P(H(A,q,k)); P(A)=\alpha\}$.

Let $\Omega=\{0,1\}$, $0<\alpha<1$, such that $\alpha q<1$, $\beta=1-(1-\alpha)^q\geq \alpha q/2$. Let $\mu=\alpha\delta_0+(1-\alpha)\delta_1$, $\nu=\beta\delta_0+(1-\beta)\delta_1$, $P=\mu^N$, $Q=\nu^N$. For a set $B\subset\{0,1\}^N$, let

$$H(B,k)=H(B,1,k)=\left\{\omega\in\{0,1\}^N; \exists\, x\in B, \operatorname{card}\{i\leq N;\ \omega_i<x_i\}\leq k\right\}.$$

We show that to prove ($*$), it is enough to show that for any subset B of $\{0,1\}^N$, we have

$$(**)\qquad P(H(B,k))\geq 1-\big((K/q)((1/k)\log(1/Q(B))+1)\big)^k.$$

Consider a set $A\subset\{0,1\}^N$ that satisfies (H). Let $B=\{(\inf_{l\leq 1}x_i^l)_i;\ x^1,\dots,x^q\in A\}$. Since A satisfies (H), (3.6) shows that $H(B,k)=H(A,q,k)$. Also, if we denote by ϕ the map from $(\{0,1\}^N)^q$ to $\{0,1\}^N$ given by $\phi(x^1,\dots,x^q)=(\inf_{l\leq q}x_i^l)_{i\leq N}$, we have $B=\phi(A^q)$. The relation between β and α shows that $\phi(P^q)=Q$, so $Q(B)\geq(P(A))^q$; Step 4 then shows that ($*$) follows from ($**$). It should be noted that it would be enough to prove ($**$) when B moreover satisfies (H); but condition (H) will not be useful in the rest of our proof.

STEP 6. Each element in $\{0,1\}^N$ is a sequence of 0's and 1's. We say that a set $B\subset\{0,1\}^N$ is *right-hereditary* whenever it has the following property. If $x\in B$ is a sequence where $x_{i_1},\dots,x_{i_p}$, $p\leq N$, are the coordinates equal to 0 ($i_1<\cdots<i_p$) and when $y\in\{0,1\}^N$ is a sequence where the coordinates equal to 0 are $y_{j_1},\dots,y_{j_p}$, $j_1<\cdots<j_p$, and when $j_1\geq i_1,\dots,j_p\geq i_p$, then $y\in B$. In this step, we reduce the proof of ($**$) to the case where B is right-hereditary.

Let us fix $1\leq s<t\leq N$. For x in $\{0,1\}^N$, we denote by $\bar{x}$ the element obtained from x by exchanging the coordinates of rank s and t. For x in B, we define $T(x)=T_{s,t}(x)$ as follows. We have $T(x)=x$, unless $x_s=0$, $x_t=1$, $\bar{x}\notin B$, in which case $T(x)=\bar{x}$. We set $T(B)=\{T(x);\ x\in B\}$. It is clear that T is one to one; since $T(x)$ has the same number of coordinates equal to 0 as x, we have $Q(T(B))=Q(B)$. The essential point to prove is that

$$(3.7)\qquad P(H(T(B),k))\leq P(H(B,k)).$$

Once (3.7) is proved, we conclude as follows. For a subset B of $\{0,1\}^N$, define $Z(B)=\sum_{i\leq N, x\in B}ix_i$. Clearly, $Z(T(B))\leq Z(B)$, and $Z(T(B))=Z(B)$ if and only if $B=T(B)$. We then apply transformations $T_{s,t}$ to the original set B until we obtain a set C for which $Z(C)$ is as small as possible. We have $Q(C)=Q(B)$, $P(H(C,k))\leq P(H(B,k))$ from (3.7), and we have $T_{s,t}(C)=C$ whenever $1\leq s<t\leq N$. It is then very simple to see that C is right-hereditary. For example, one can show by induction over k that if $x\in C$ is a sequence such that $x_{i_1},\dots,x_{i_p}$ are the coordinates equal to 0 ($i_1<\cdots<i_p$), and if $y\in\{0,1\}^N$ is a sequence such that $y_{j_1},\dots,y_{j_p}$ are the coördinates equal to 0 ($j_1<\cdots<j_p$), and if $j_1\geq i_1,\dots,j_p\geq i_p$, $\sum_{l\leq p}j_l-i_l\leq k$, then $y\in C$. We should also mention that it does not seem to be true in general that $P(H(T(B),q,k))\leq P(H(B,q,k))$; this was the reason for transforming the problem in Step 5.

To prove (3.7), we note that the transformation $\omega \to \bar{\omega}$ preserves P, and we show that it maps $H(T(B), k) \setminus H(B, k)$ into $H(B, k) \setminus H(T(B), k)$. Let $\omega \in H(T(B), k) \setminus H(B, k)$, so we have

$$\exists\, x \in T(B), \qquad \operatorname{card}\{i \le N;\ \omega_i < x_i\} \le k, \tag{3.8}$$

$$\forall\, y \in B, \qquad \operatorname{card}\{i \le N;\ \omega_i < y_i\} > k. \tag{3.9}$$

Since $x \in T(B)$, we have $x = T(z)$ for some z in B. From (3.9) we have $x \notin B$, so $x = \bar{z}$. The definition of T shows that $x_s = 1$, $x_t = 0$. From (3.8) and the fact that $\operatorname{card}\{i \le N;\ \omega_i < z_i\} > k$, we see by looking at all cases that we must have $\omega_s = 1$, $\omega_t = 0$. We have

$$\operatorname{card}\{i \le N;\ \omega_i < x_i\} = \operatorname{card}\{i \le N;\ \bar{\omega}_i < z_i\}.$$

This implies that $\bar{\omega} \in H(B, k)$. Suppose, if possible, that $\bar{\omega} \in H(T(B), k)$. Then for some $v \in T(B)$, we have $\operatorname{card}\{i \le N;\ \bar{\omega}_i < v_i\} \le k$. We have $v \in B$, for otherwise $v = \bar{w}$ for some w in B, and this would imply $\omega \in H(B, k)$. From (3.9), we have $\operatorname{card}\{i \le N;\ \omega_i < v_i\} > k$. Since $\omega_s = 1$, $\omega_t = 0$, the only possible case is $v_s = 0$, $v_t = 1$. Since $v \in T(B)$, we have $v = T(w)$ for some w in B; the only possibility is $v = w$, so $T(v) = v$; but this implies $\bar{v} \in B$, so $\operatorname{card}\{i < N,\ \omega_i < \bar{v}_i\} \le k$, which contradicts $\omega \notin H(B, k)$. The proof is complete.

STEP 7. We will prove $(**)$ when B is right-hereditary by a direct computation. We set $\theta = 24 + (4/k)\log[1/Q(B)]$.

LEMMA 15. *For $l \ge 1$, let $a_l = \theta kl/\alpha q$. Then there is x in B such that*

$$\operatorname{card}\{i \le a_l;\ x_i = 0\} \ge kl \tag{3.10}$$

whenever $a_l < N + 1$.

PROOF. We denote by l_0 the largest integer such that $a_{l_0} < N + 1$ (if $a_1 \ge N + 1$, there is nothing to prove). Denote by $B(M, \tau)$ the number of successes in a run of M Bernoulli trials with probability of success τ. It is enough to show that

$$\sum_{l \le l_0} P(B([a_l], \beta) < kl) < Q(B), \tag{3.11}$$

where $[t]$ is the integer part of t. We will use the following inequality, due to Chernoff (see [10], pages 15–16). For $0 < t < 1$,

$$P(B(M, \tau) \le tM) \le \left(\left(\frac{\tau}{t}\right)^t \left(\frac{1-\tau}{1-t}\right)^{1-t}\right)^M. \tag{3.12}$$

We note that

$$\left(\frac{1-\tau}{1-t}\right)^{1-t} = \left(1 - \frac{\tau - t}{1-t}\right)^{1-t} \le \exp -(\tau - t),$$

so

$$P(B(M,\tau) \le tM) \le \exp\left(-tM\left(\frac{\tau}{t} - 1 - \log\frac{\tau}{t}\right)\right).$$

For $x \ge 6$, we have $x - 1 - \log x \ge x/2$, so for $\tau/t \ge 6$, we have

$$P(B(M,\tau) \le tM) \le \exp(-\tau M/2).$$

We use this inequality with $M = [a_l]$, $t = kl/M$, $\tau = \alpha q/2$. Since $\alpha q \le 1$, we have $a_l \ge 2$, so $[a_l] \ge a_l/2$, so

$$\frac{\tau}{t} \ge \frac{\alpha q M}{2kl} \ge \frac{\alpha q a_l}{4kl} \ge \frac{\theta}{4}.$$

Since $\theta \ge 24$, we have $\tau/t \ge 6$, so since $\tau M \ge \alpha q a_l/2$, we have

$$\begin{aligned} P(B([a_l],\beta) \le kl) &\le P(B([a_l],\alpha q/2) \le kl) \le \exp(-\alpha q a_l/4) \\ &= \exp(-kl\theta/4) \le \exp(-l(\log(1/Q(B)) + 6k)), \end{aligned}$$

which implies (3.11), and proves the lemma. □

The basic remark is now as follows. If $\omega \notin H(B,k)$, then for some $p \le N$ we have

$$\text{card}\{i \le p;\ \omega_i = 0\} \ge k + \text{card}\{i \le p,\ x_i = 0\}. \tag{3.13}$$

Indeed, suppose that ω is such that (3.13) fails for all $p \le N$. Denote by i_p (resp. j_p) the pth index for which $x_i = 0$ (resp. $\omega_i = 0$). Then, for each p such that i_p is defined, we have

$$\text{card}\{i \le i_p;\ \omega_i = 0\} < k + \text{card}\{i \le i_p;\ x_i = 0\} = k + p,$$

which shows that $j_{p+k} > i_p$. Since B is right-hereditary, and $x \in B$, we can find y in B with $y_{j_l} = 0$ whenever $l > k$. We can have $\omega_i < y_i$ only when $\omega_i = 0$, $y_i = 1$, so i has to be one of the indexes j_l for $l \le k$, and so $\text{card}\{i \le N; \omega_i < y_i\} \le k$, so $\omega \in H(B,k)$. From (3.10) and (3.13) it follows that if $\omega \notin H(B,k)$, we have

$$\exists\, l \le l_0, \qquad \text{card}\{i \le a_l,\ \omega_i = 0\} \ge kl.$$

Indeed, if p is as in (3.13), let m be the largest integer with $a_m \le p$, and let $l = m+1$ (if $m = l_0$, we set $a_l = N$; if $a_1 \ge p$, we set $l = 1$). We have

$$\begin{aligned} \text{card}\{i \le a_l;\ \omega_i = 0\} &\ge \text{card}\{i \le p;\ \omega_i = 0\} \\ &\ge k + \text{card}\{i \le p;\ x_i = 0\} \\ &\ge k + \text{card}\{i \le a_m;\ x_i = 0\} \\ &\ge k(m+1) = kl. \end{aligned}$$

This shows that

$$P(H(B,k)) \ge 1 - \sum_{l \ge 1} P(B([a_l], 1-\alpha) \le [a_l] - kl).$$

We again use (3.12); but now we note that

$$\left(\frac{\tau}{t}\right)^t = \left(1 + \frac{\tau - t}{t}\right)^t \leq \exp(\tau - t) \leq \exp(1 - t),$$

so

$$P(B(M,\tau) \leq tM) \leq \left(e\frac{(1-\tau)}{1-t}\right)^{M(1-t)}.$$

Using this with $M = [a_l]$, $t = 1 - kl/M$, $\tau = 1 - \alpha$, we have

$$P(B([a_l], 1-\alpha) \leq [a_l] - kl) \leq \left(\frac{eM\alpha}{kl}\right)^{kl} \leq \left(\frac{e\theta}{q}\right)^{kl}.$$

Since for $0 < u \leq 1$, we have $u/(1-u) \leq 2u$, this gives $P(H(B,k)) \geq 1 - 2(e\theta/q)^k \geq 1 - (2e\theta/q)^k$ and completes the proof. □

REMARK. At the expense of some complications, it is possible to improve the computations of the last step [the idea now being to define $a_l = \theta kl/(\alpha q \log q)$ for a suitable θ]. In particular, in the case $P(A) \geq \frac{1}{2}$, the bound of Theorem 7 can be replaced by

$$P_*(H(A,q,k)) \geq 1 - \left(K\left(\frac{1}{k} + \frac{1}{q\log q}\right)\right)^k. \tag{3.14}$$

The gain of the factor $\log q$ is irrelevant for the applications presented here or in [10]. It should be noted that (3.14) is sharp, at least for $k \geq q \log q$ (we have not checked this for smaller k; note however that the values of $k \geq q \log q$ are those used for our applications). Indeed, consider the case where $\Omega = \{0,1\}$, and $\mu(\{0\}) = 1 - r/N$, $\mu(\{1\}) = r/N$. Consider $r \geq 1$, and let

$$A = \left\{x = (x_i)_{i \leq N};\ \sum_{i \leq N} x_i \leq r\right\}.$$

Clearly,

$$H(A,q,k) = \left\{x = (x_i)_{i \leq N};\ \sum_{i \leq N} x_i \leq rq + k\right\}.$$

Now by elementary computations, for k of order $rq \log q$, and N large enough

$$P(H^c(A,q,k)) \geq \left(1 - \frac{r}{N}\right)^{N-rq-k-1}\left(\frac{r}{N}\right)^{rq+k+1}\binom{N}{rq+k+1}$$
$$\geq \left(\frac{1}{Kq\log q}\right)^k.$$

Acknowledgments. I am indebted to A. de Acosta who mentioned to me the question of the integrability of $\exp \alpha\|X\|\log^+\|X\|$ when $\sup\|X_i\|_\infty < \infty$; this was the starting point of this research. Thanks are also due to M. Ledoux for some useful comments.

REFERENCES

[1] BOREL, C. (1975). The Brunn–Minkowski inequality in Gauss spaces. *Invent. Math.* **30** 207–216.

[2] DE ACOSTA, A. (1980a). Exponential moments of vector valued random series and triangular arrays. *Ann. Probab.* **8** 381–389.

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

[4] HEINKEL, B. (1986). Some exponential inequalities with applications to the central limit theorem in $C[0,1]$. Unpublished.

[5] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–23.

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

[7] JAIN, N. and MARCUS, M. (1975). Integrability of infinite sums of independent vector-valued random variables. *Trans. Amer. Math. Soc.* **212** 1–36.

[8] JOHNSON, W. B., SCHECHTMAN, G. and ZINN, J. (1985). Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. *Ann. Probab.* **13** 234–253.

[9] KUELBS, J. (1978). Some exponential moments of sums of independent random variables. *Trans. Amer. Math. Soc.* **240** 145–162.

[10] LEDOUX, M. and TALAGRAND, M. (1989). Some applications of isoperimetric methods to strong limit theorems for sums of independent random variables. *Ann. Probab.* To appear.

[11] MILMAN, V. D. and SCHECHTMAN, G. (1986). *Asymptotic Theory of Finite Dimensional Normed Spaces. Lecture Notes in Math.* **1200**. Springer, New York.

[12] PISIER, G. (1986). Probabilistic methods in the geometry of Banach spaces. *Probability and Analysis (Varenna, 1985). Lecture Notes in Math.* **1206** 167–241. Springer, New York.

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

EQUIPE D'ANALYSE, TOUR 46
U.A. AU C.N.R.S. NO. 754
UNIVERSITÉ PARIS VI
75230 PARIS CEDEX 05
FRANCE

DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
231 WEST 18TH AVENUE
COLUMBUS, OHIO 43210