The Annals of Probability
2002, Vol. 30, No. 3, 1223–1237

REMARKS ON DEVIATION INEQUALITIES FOR FUNCTIONS OF INFINITELY DIVISIBLE RANDOM VECTORS[1]

BY CHRISTIAN HOUDRÉ

Georgia Institute of Technology

We obtain deviation inequalities for some classes of functions of infinitely divisible random vectors having finite exponential moments.

1. Introduction and statement of results. It is by now well known that the concentration of measure phenomenon, and in particular its functional form, is of great importance in probability theory and its applications. At the methodological level various tools have been developed to obtain this type of results: isoperimetric techniques, inductive methods, functional inequality techniques, transportation and information theoretic methods. In particular, every single one of the methods stated above recovers the classical fact (motivating many of the others) that Lipschitz functions of independent Gaussian random variables have near Gaussian tails (we refer to the sets of notes of Ledoux [12, 13] for precise credit and references for this large body of work). More recently, the Gaussian deviation has also been recovered (at the polynomial level) using a covariance representation [1]; and this method might be the simplest to date to obtain this result. Since this representation has a version for infinitely divisible (i.d.) random vectors, it is natural to try to explore its deviation consequences. This is initiated below.

Let $X \sim ID(b, 0, \nu)$, that is, let X be a d-dimensional infinitely divisible random vector without Gaussian component. Its characteristic function is given by

$$\varphi(t) = \exp\left\{ i\langle t, b\rangle + \int_{\mathbb{R}^d} \left(e^{i\langle t,u\rangle} - 1 - i\langle t, u\rangle \mathbf{1}_{\|u\|\le 1}\right)\nu(du)\right\}, \tag{1.1}$$

where $t, b \in \mathbb{R}^d$ and where ν (the Lévy measure) is a positive measure without atom at the origin on $\mathcal{B}(\mathbb{R}^d)$ (the Borel σ-field of $\mathbb{R}^d$) such that $\int_{\mathbb{R}^d}(\|u\|^2 \wedge 1) \times \nu(du) < +\infty$ (throughout, $\langle \cdot, \cdot\rangle$ and $\|\cdot\|$ are, respectively, the Euclidean inner product and norm in $\mathbb{R}^d$, and we also assume that $\nu \not\equiv 0$). As is well known, the behavior of the Lévy measure ν controls the integrability property of the vector X and any feasible light or heavy tail is possible for i.d. vectors (we refer to Sato [17] for a good general introduction to infinitely divisible distributions and Lévy processes). Here, we only take X without Gaussian component but it is rather easy to get at once results for the general case.

Received August 2000; revised September 2001.

[1]Supported in part by NSF Grant DMS-98-03239.

AMS 2000 subject classifications. Primary 62G07, 62C20; secondary 60G70, 41A25.

Key words and phrases. Infinite divisible random vectors, deviation inequalities, concentration of measure phenomenon, Lipschitz functions.

Let us now state a summary of some of the results we intend to prove. Again, let $X \sim ID(b, 0, \nu)$, and let $f : \mathbb{R}^d \to \mathbb{R}$, where $\mathbb{R}^d$ is equipped with the Euclidean norm. The function f is said to be Lipschitz (with constant α) if

$$|f(x+u) - f(x)| \le \alpha \|u\|, \tag{1.2}$$

for every $x \in \mathbb{R}^d$ and $u \in \mathbb{R}^d$. Below, the function f need not be defined on the whole of $\mathbb{R}^d$ but just on a subset of $\mathbb{R}^d$ containing $R_X + S_\nu$ and R_X, where R_X is the range of X and S_ν the support of ν, for example, if X is a Poisson random variable, a Lipschitz function (with constant α) is then defined on $\mathbb{N}$ and such that

$$|f(n+1) - f(n)| \le \alpha, \tag{1.3}$$

for every $n \in \mathbb{N}$. However, since we do not wish to distinguish between the various ranges and supports, and also distinguish between continuous and discrete random variables, it is in the sense of (1.2) that Lipschitz will be understood in the rest of these notes. Moreover, a Lipschitz function f defined on a subset of $\mathbb{R}^d$ can be extended (without increasing its Lipschitz seminorm) to a function $\tilde{f}$ which is Lipschitz on all of $\mathbb{R}^d$ (see Theorem 6.1.1 in Dudley [6]). Now, as integrands, f and $\tilde{f}$ are the same (as seen from the various proofs below). Finally, it is clear that although we mainly consider the Lipschitz property with respect to the Euclidean norm, other norms could have equally been used.

We are now ready to present our first general deviation result.

THEOREM 1. *Let $X \sim ID(b, 0, \nu)$ be such that $Ee^{t\|X\|} < +\infty$, for some $t > 0$. Then, for any Lipschitz function f (with constant α),*

$$P\big(f(X) - Ef(X) \ge x\big) \le \exp\left(-\int_0^x h^{-1}(s)\, ds\right), \tag{1.4}$$

for all $0 < x < h((\frac{M}{\alpha})^-)$, where $M = \sup\{t \ge 0 : Ee^{t\|X\|} < +\infty\}$ and where h^{-1} is the inverse of $h(s) = \int_{\mathbb{R}^d} \alpha\|u\|(e^{\alpha s\|u\|} - 1)\nu(du)$, $0 < s < \frac{M}{\alpha}$.

It is clear that applying (1.4) to $-f$, gives left tail estimates; that is,

$$P\big(f(X) - Ef(X) \le x\big) \le \exp\left(-\int_x^0 h^{-1}(-s)\, ds\right), \tag{1.5}$$

for all $-h((\frac{M}{\alpha})^-) < x < 0$, and that two sided tails follow from (1.4) and (1.5). [Throughout, $h(a^-)$ denotes the left-hand limit of h at the point a.]

In essence, the previous results tell us that under a Cramér condition, the tail behavior of $f(X)$ is controlled by the tail behavior of X. The deviation inequality (1.4) is sharp. Indeed, let X be a Poisson random variable with mean λ,

$X \sim \mathcal{P}(\lambda)$ and let $\alpha = 1$. Then, $h(s) = \lambda \int_{\mathbb{R}} (|u|e^{s|u|} - |u|)\delta_1(du) = \lambda(e^s - 1)$, and $h^{-1}(s) = \log(1 + \frac{s}{\lambda})$. Hence, (1.4) becomes

$$\begin{aligned} P\big(f(X) - Ef(X) \geq x\big) & \\ &\leq \exp\biggl(-\int_0^x \log\Bigl(1 + \frac{s}{\lambda}\Bigr)\,ds\biggr) \\ &= \exp\biggl(-\Bigl\{x \log\Bigl(1 + \frac{x}{\lambda}\Bigr) - x + \lambda \log(\lambda + x) - \lambda \log \lambda\Bigr\}\biggr) \\ &= e^x \exp\biggl(-(x + \lambda)\log\Bigl(1 + \frac{x}{\lambda}\Bigr)\biggr), \end{aligned} \tag{1.6}$$

for all $x > 0$, since $h(M^-) = +\infty$ in that case. In particular, if $f(X) = X$ we get

$$P(X - \lambda \geq x) \leq e^x \exp\biggl(-(x + \lambda)\log\Bigl(1 + \frac{x}{\lambda}\Bigr)\biggr). \tag{1.7}$$

Already this particular case of (1.4) recovers some recent results of Bobkov and Ledoux [3], and (1.7) provides rather precise asymptotics. Indeed, for λ and n integers, $P(X \geq \lambda + n) = \int_0^\lambda e^{-t} t^{\lambda+n-1}\,dt / \Gamma(\lambda + n)$. Now, $\int_0^\lambda e^{-t} t^{\lambda+n-1}\,dt \sim (\lambda + n)^{-1} e^{-\lambda} \lambda^{\lambda+n}$, as $n \to +\infty$, while standard estimates on the Gamma function give $\Gamma(x) = \sqrt{2\pi} e^{-x} x^{x-1/2}(1 + \frac{1}{12x} + \frac{1}{288x^2} + \cdots)$, as $x \to +\infty$ (e.g., see [14]). Thus, $P(X \geq \lambda + n) \sim \frac{1}{\sqrt{2\pi(\lambda+n)}} e^n e^{-(\lambda+n)\log(1+n/\lambda)}$ as $n \to +\infty$, which is, up to an inverse fractional polynomial factor, exactly (1.7). Unfortunately this polynomial factor is not recovered here, in contrast to the Gaussian situation [1].

Other types of tail behavior are possible. Indeed, still assuming $d = 1$, let X be a geometric random variable with parameter p; that is, $\mu(k) = pq^k$, $k = 0, 1, 2, \ldots,$ where as usual $q = 1 - p$. Then ν is concentrated on the positive integers with $\nu(k) = \frac{q^k}{k}$, $k = 1, 2, \ldots.$ Since, $Ee^{tX} = \frac{p}{1 - qe^t}$, for $t < -\log q$, we can apply Theorem 1. Hence, for any $f : \mathbb{N} \to \mathbb{R}$ such that $|f(n + 1) - f(n)| \leq 1$,

$$P\big(f(X) - Ef(X) \geq x\big) \leq \exp\biggl(\Bigl(x + \frac{q}{p}\Bigr)\log\Bigl(\frac{q + pqx}{q + px}\Bigr) + \log(1 + px)\biggr), \tag{1.8}$$

for all $x > 0$, since again $h(M^-) = +\infty$. Clearly [using $\log(1 + x) \leq x, x \geq 0$] (1.8) provides rather precise tail estimates, since for $x \geq 0$,

$$P(X - EX \geq x) = \exp\biggl(\Bigl(\Bigl[x + \frac{q}{p}\Bigr] + 1\Bigr)\log q\biggr), \tag{1.9}$$

whenever $x + \frac{q}{p}$ is not a positive integer (with easily modifications when $x + \frac{q}{p}$ is a positive integer). In complete similarity with these results, it is easy to see that if X is negative binomial, then up to a multiplicative polynomial factor, $f(X) - Ef(X)$ has negative binomial tails. It can also be checked that if X is a Gamma random variable, then again $f(X) - Ef(X)$ has near Gamma tails. If X is an extreme distribution given by $P(X \leq x) = e^{-e^{-x}}$, $x \in \mathbb{R}$, then X is infinitely divisible and the left tails of $f(X)-Ef(X)$ are of order $e^{-e^{-x}}$. If X is a compound Poisson random variable, that is, $X=\sum_{n=0}^{N} Z_n$, where $Z_0=0$, and where the Z_n, $n\geq 1$, are i.i.d. random variables with law μ, independent of the Poisson random variable N with mean λ, then X is infinitely divisible with Lévy measure $\nu=\lambda\mu$. Thus, if $Ee^{t|Z_1|}<+\infty$ for some $t>0$, it follows that $P(f(X)-Ef(X)\geq x)\leq \exp(\min_{0<t<M}(H(t)-tx))$, where $H(t)=\lambda E(e^{t|Z_1|}-t|Z_1|-1)$, and where $M=\sup\{t>0: Ee^{t|Z_1|}<+\infty\}$.

It is clearly not our purpose to list here every particular case of infinitely divisible random vector to which Theorem 1 applies, but many examples can be found. The previous examples just illustrate the potential sharpness of the estimate (1.4).

For ν with bounded support, the hypotheses of the previous theorem are satisfied and (1.4) can also be made more precise.

COROLLARY 1. *Let ν have bounded support with $R=\inf\{\rho>0:\nu(\{x:\|x\|>\rho\})=0\}$, and let $V^2=\int_{\mathbb{R}^d}\|u\|^2\nu(du)$. Then, for any Lipschitz function f (with constant α),*

$$(1.10)\quad P\big(f(X)-Ef(X)\geq x\big)\leq e^{(x/\alpha R)}\exp\biggl(-\biggl(\frac{x}{\alpha R}+\frac{V^2}{R^2}\biggr)\log\biggl(1+\frac{Rx}{\alpha V^2}\biggr)\biggr),$$

for all $x>0$.

A direct consequence of (1.10) is the strong exponential integrability of the functional $f-Ef$. More precisely, under the hypothesis of Corollary 1 (with $\alpha=1$),

$$(1.11)\qquad Ee^{t|f-Ef|\log|f-Ef|}<+\infty,$$

for any $t<\frac{1}{R}$. In case $f(X)=\|X\|$, a result similar to (1.11) is known and due to Sato (see Theorem 26.1 in [17]). For norms of Banach space valued random variables, it is due to De Acosta [5] and Talagrand [19]. The representation (2.1) on which the proof of Theorem 1 is based, is also valid for Banach space valued random variables (see [9]), and so Theorem 1 should have a version in that setting as well. Actually, a stronger result due to Rosiński [16] asserts that (for norms of Banach-space valued random variables)

$$E\exp\big(R^{-1}\|X\|\log^{+} tR^{-1}\|X\|\big)<+\infty,$$

for all $t>0$ such that $tp_0\leq\frac{1}{e}$, where $p_0=\nu(\{x:\|x\|=R\})$.

Sato also showed that for infinitely divisible (real) vectors X with boundedly supported Lévy measure, and for any $0<t<\frac{1}{R}$,

$$P(\|X\|\geq x)=o(e^{-tx\log x}),\qquad x\to+\infty.$$

Thus, (1.10) also generalizes the above.

Note that in the multivariate Poisson case, ν consists of point masses, of arbitrary sizes $0 \le \lambda_i < +\infty$, distributed on the elements d_i of $\{0,1\}^d$. Hence in (1.10), $R = \sqrt{d}$ and $V^2 = \sum_{i=1}^{2^d} \lambda_i \|d_i\|^2$.

Let us now present another corollary which also provides tail inequalities under Bernstein-type moment assumptions.

COROLLARY 2. *Let $X \sim ID(b, 0, \nu)$ be such that $E\|X\|^2 < +\infty$, with moreover,*

$$\int_{\mathbb{R}^d} \|u\|^k \nu(du) \le \frac{C^{k-2}k!}{2} \int_{\mathbb{R}^d} \|u\|^2 \nu(du), \qquad k \ge 3,$$

for some $C > 0$. Then for any Lipschitz function f (with constant α),

$$P\big(f(X) - Ef(X) \ge x\big) \le \exp\biggl(-\frac{1}{8}\min\biggl(\frac{2x^2}{\alpha^2 V^2}, \frac{x}{\alpha C}\biggr)\biggr),$$

where $V^2 = \int_{\mathbb{R}^d} \|u\|^2 \nu(du)$.

Theorem 1 has a drawback; it requires a knowledge of the Lévy measure ν of the vector X in order to compute h to then identify its inverse. However, very little is known on the Lévy measure of d-dimensional infinitely divisible vectors (with dependent components). Again, we refer to Sato [17] for up-to-date information on the topic. Sometimes an explicit knowledge of the Lévy measure can be bypassed. Indeed, in the proof of Theorem 1 (and, say, for $\alpha = 1$), we wish to find $\min_{0<t<M} e^{-tx+H(t)}$, where

$$H(t) = \int_{\mathbb{R}^d} (e^{t\|u\|} - 1 - t\|u\|)\nu(du) = K(t) - \int_{\|u\|>1} t\|u\|\nu(du) - \|b\|t,$$

and so $K(t) = \int_{\mathbb{R}^d} (e^{t\|u\|} - 1 - t\|u\|\mathbf{1}_{\|u\|\le 1})\nu(du) + \|b\|t$. In some instances, $e^{K(t)}$ can be expressed via the distribution of X (see Theorem 21.9 in [17]), but this representation might not always be that useful for our purposes. When some knowledge of the Lévy measure is available, Theorem 1 is sharp. This can be verified for the density $ce^{-\|x\|}$ (c normalizing constant), which is infinitely divisible (see Takano [18] or [17]), and whose Lévy measure is explicitely known.

Other versions of Theorem 1 are possible. For example, let $X \sim ID(b, 0, \nu)$ and let $f : \mathbb{R}^d \to \mathbb{R}$, be such that the representation (2.1) holds, with also:

(i) $A = \|\int_{\mathbb{R}^d} (f(X+u) - f(X))^2 \nu(du)\|_{L^\infty(X)}^{1/2} < +\infty$.

(ii) $k(t) = \|\int_{\mathbb{R}^d} (e^{t(f(X+u)-f(X))} - 1)^2 \nu(du)\|_{L^\infty(X)}^{1/2} < +\infty$, for all $t \in (0, T)$, $T > 0$.

For k integrable, it should be clear from the proof of Theorem 1 that

$$P\big(f(X) - Ef(X) \ge x\big) \le \exp\biggl(\min_{0<t<T}\biggl(\int_0^t Ak(s)\,ds - tx\biggr)\biggr), \tag{1.12}$$

which is however more untractable than (1.4). This last estimate also applies to a different class of functions, but nevertheless recovers Corollary 1. Indeed, for ν with bounded support, f Lipschitz (with constant one) and using the notation of Corollary 1, we have

$$\begin{aligned}
k(t) &\le \left(\int_{\mathbb{R}^d} (e^{t\|u\|}-1)^2 \nu(du)\right)^{1/2} \\
&= \left(\int_{\|u\|\le R} \|u\|^2 \left(\sum_{k=1}^{\infty} \frac{t^k \|u\|^{k-1}}{k!}\right)^2 \nu(du)\right)^{1/2} \\
&\le \left(\int_{\|u\|\le R} \|u\|^2 \left(\sum_{k=1}^{\infty} \frac{t^k R^{k-1}}{k!}\right)^2 \nu(du)\right)^{1/2} \\
&= V\left(\frac{e^{tR}-1}{R}\right).
\end{aligned}$$

Thus, since $A \le V$, (1.12) gives Corollary 1.

Actually, it is easily seen from Theorem 1 and (1.12), that what is needed to get similar results "in greater generality" is an upper bound on

$$\left\| \int_{\mathbb{R}^d} \big(f(X+u)-f(X)\big)\big(e^{t(f(Y+u)-f(Y))}-1\big)\nu(du) \right\|_{L^\infty(X,Y)},$$

but this leads to tail expressions even more untractable than (1.12).

To finish this introduction, let us deal with the case of infinite divisible vectors with independent components (and, for simplicity of notation, identically distributed components). In that case, the Lévy measure ν is concentrated on the axes (see [17], page 67); that is,

$$\nu(dx_1,\dots,dx_d) = \sum_{i=1}^{d} \delta_0(dx_1)\cdots\delta_0(dx_{i-1})\tilde{\nu}(dx_i)\delta_0(dx_{i+1})\cdots\delta_0(dx_d).$$

Identifying $u_i \in \mathbb{R}$, with $(0,\dots,0,u_i,0,\dots,0) \in \mathbb{R}^d$, $i = 1,\dots,d$, Theorem 1 becomes Theorem 2.

THEOREM 2. *Let $X \sim ID(b,0,\nu)$ with i.i.d. components, be such that $Ee^{t\|X\|} < +\infty$, for some $t > 0$. Let $f:\mathbb{R}^d \to \mathbb{R}$ be such that $|f(x+u_i)-f(x)| \le \beta|u_i|$, for all $x \in \mathbb{R}^d$, $u_i \in \mathbb{R}$, and $i = 1,\dots,d$. Then*

$$P\big(f(X)-Ef(X) \ge x\big) \le \exp\left(-\int_0^x h^{-1}(s)\,ds\right), \tag{1.13}$$

for all $0 < x < h((\frac{M}{\beta})^-)$, where $M = \sup\{t \ge 0,\ Ee^{t\|X\|} < +\infty\}$ and where h^{-1} is the inverse of $h(s) = \beta d \int_{\mathbb{R}} |u|(e^{\beta s|u|}-1)\tilde{\nu}(du)$, $0 < s < \frac{M}{\beta}$.

The tail estimate (1.13) is not always optimal. A case at hand is the Laplace distribution. Indeed, let $X_1, \dots, X_d$, be i.i.d. random variables with density $2^{-1}e^{-|x|}$. Then, the Lévy measure of X_i has density $\frac{e^{-|u|}}{|u|}$, $u \in \mathbb{R}$, $u \neq 0$. Now

$$
\begin{aligned}
H(t) &= \int_{\mathbb{R}^d} (e^{\beta t\|u\|} - \beta t\|u\| - 1)\nu(du) \\
&= d \int_{\mathbb{R}\setminus\{0\}} \sum_{k=2}^{\infty} \frac{(\beta t)^k |u|^k}{k!} \frac{e^{-|u|}}{|u|} \, du \\
&= 2d \int_0^{\infty} \sum_{k=2}^{\infty} \frac{(\beta t u)^k}{k!} \frac{e^{-u}}{u} \, du \\
&= -2d\bigl(\log(1 - \beta t) + \beta t\bigr).
\end{aligned}
$$

Hence $\min_{0<t<1/\beta}(H(t) - tx)$ is attained at $t = \frac{1}{\beta}(\frac{x}{x+2\beta d})$ and the corresponding tail estimate is

$$
\exp\biggl(2d \log\biggl(1 + \frac{x}{2\beta d}\biggr) - \frac{x}{\beta}\biggr). \tag{1.14}
$$

Now for x large, (and up to absolute constants) the exponent in (1.14) is of order $-\frac{x}{\beta}$, while for x small it is of order $\frac{-x^2}{d\beta^2}$. Now an influential estimate of Talagrand [20] tells us that if f is such that $|f(x + u) - f(x)|^2 \leq \alpha^2 \sum_{i=1}^{d} |u_i|^2$ and $|f(x + u) - f(x)| \leq \beta \sum_{i=1}^{d} |u_i|$, then $-\min(\frac{x^2}{\alpha^2}, \frac{x}{\beta})$ is the right order. With our approach, the estimate for x small is worse than $\frac{-x^2}{\alpha^2}$, while the same for x large. One reason for this might be the fact that we know since the work of Bobkov and Ledoux [2] that a Poincaré inequality leads to the estimate $-\min(\frac{x^2}{\alpha^2}, \frac{x}{\beta})$, while, in our case, we have just used the Lipschitz structure. In the absence of a Poincaré inequality, (1.13) might be optimal. It should, however be noted that for linear sums, for example, $f(x_1, \dots, x_d) = \frac{1}{d} \sum_{i=1}^{d} |x_i|$, both Talagrand's inequality and (1.14) give the same order $e^{-cd \min(x, x^2)}$, where c is an absolute constant. It would be interesting to know if modifications of our methods can give Talagrand's inequality. A simple modification such as replacing the "uniform" β by β_i, $i = 1, \dots, d$, will still come short of this goal.

Following a referee's suggestion, let us explicitely link, in a classical manner, the results obtained above for functions to concentration results for sets. Let d_2 denote the Euclidean distance on $\mathbb{R}^d \times \mathbb{R}^d$, let $A \subset \mathbb{R}^d$ be a Borel set of measure $\mu(A) \geq \frac{1}{2}$, where μ is the law of X, and let $f_r(x) = \min(d_2(x, A), r)$, $r > 0$. Clearly, f_r is Lipschitz with constant at most one and $Ef_r \leq \frac{r}{2}$, where E is expectation with respect to μ. Let also A^r be the open r-neighborhood of A with respect to the Euclidean distance; that is, $A^r = \{x \in \mathbb{R}^d : d_2(x, A) < r\}$, then applying (1.4) to f_r, we obtain the concentration inequality,

$$1 - \mu(A^r) \le \mu\{f_r - Ef_r \ge r/2\} \le \exp\left(-\int_0^{r/2} h^{-1}(s)\,ds\right),$$

where X, h and $r/2$, satisfy the conditions of Theorem 1. Since $A^r = A + rB_2$, where B_2 is the open Euclidean unit ball in $\mathbb{R}^d$, this last inequality is just

$$\mu(A + rB_2) \ge 1 - \exp\left(-\int_0^{r/2} h^{-1}(s)\,ds\right). \tag{1.15}$$

As is well known, an inequality such as (1.15) also (essentially) implies (1.4). Let us present the version of Theorem 2 for sets. In the independent case, by the form of the Lévy measure, B_2 is replaced by B_1 the open ℓ_1-unit ball of $\mathbb{R}^d$. Thus, if the product measure μ^d is the law of X (and so μ is the law of any of the i.i.d. components of X), (1.15) becomes

$$\mu^d(A + rB_1) \ge 1 - \exp\left(-\int_0^{r/2} h^{-1}(s)\,ds\right),$$

where h is now as in Theorem 2. In particular if $\mu = \eta$, where η is the one-dimensional Laplace distribution, we get for all $r > 0$, and since $\beta = 1$,

$$\eta^d(A + rB_1) \ge 1 - \exp\left(-c\min\left(\frac{r^2}{d}, r\right)\right),$$

where c is some absolute constant and since, as explained above, the exponent in (1.14) is equivalent to $-\min(\frac{x^2}{d}, x)$. In turn, it is easily seen that this last inequality is equivalent to

$$\eta^d(A + \sqrt{rd}\,B_1 + rB_1) \ge 1 - e^{-cr}, \tag{1.16}$$

where c is an(other) absolute constant. It is again clear that (1.16) is weaker than Talagrand's

$$\eta^d(A + \sqrt{r}\,B_2 + rB_1) \ge 1 - e^{-cr}. \tag{1.17}$$

Using an idea of Pisier ([15], page 181), Talagrand [20] shows that (1.17) recovers (and improves some aspects of) the Gaussian concentration. The same arguments show that (1.16) also recovers the Gaussian concentration, but only for r large enough.

Let us now return to the functional forms of deviation inequalities. In the independent case, a result similar to Corollary 1 also follows from (1.13), but this also leads to dimension dependent (on d) estimates. However, as shown in the next corollary, the method of Theorem 2 can be modified sometimes to obtain dimension independent estimates. First, we have:

THEOREM 3. *Let $X \sim ID(b, 0, \nu)$ with i.i.d. components, be such that $\int_{\mathbb{R}} e^{t|u|}\tilde{\nu}(du) < +\infty$, for some $t > 0$. Let $1 \le p, q, r \le +\infty$, $r \ne +\infty$, with $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, and let $f : \mathbb{R}^d \to \mathbb{R}$ be such that*

$$\left\| \sum_{i=1}^{d} \int_{\mathbb{R}} |f(X + u_i) - f(X)|^p \tilde{\nu}(du_i) \right\|_{L^\infty(X)}^{1/p} = \alpha < +\infty, \tag{1.18}$$

$$\left\| \sum_{i=1}^{d} \int_{\mathbb{R}} |f(X + u_i) - f(X)|^q \tilde{\nu}(du_i) \right\|_{L^\infty(X)}^{1/q} = \beta < +\infty, \tag{1.19}$$

$$\max_{1 \le i \le d} \left(\frac{\| f(X + u_i) - f(X) \|_{L^\infty(X)}}{|u_i|} \right) \le \gamma < +\infty, \tag{1.20}$$

for all $u_i \in \mathbb{R}$, $u_i \ne 0$. Then

$$P\big(f(X) - Ef(X) \ge x\big) \le \exp\left(- \int_0^x h^{-1}(s)\, ds \right), \tag{1.21}$$

for all $0 < x < h((\frac{M}{\gamma r})^-)$, where h^{-1} is the inverse of $h(s) = \alpha\beta s (\int_{\mathbb{R}^d} e^{\gamma r s \|u\|} \times \nu(du))^{1/r}$, $0 < s < \frac{M}{\gamma r}$, and where $M = \sup\{t \ge 0 : \int_{\mathbb{R}} e^{t|u|}\tilde{\nu}(u) < +\infty\}$.

COROLLARY 3. *Let ν have bounded support with $R = \inf\{\rho > 0 : \nu(\{x : \|x\| > \rho\}) = 0\}$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be such that*

$$\left\| \sum_{i=1}^{d} \int_{|u_i| \le R} |f(X + u_i) - f(X)|^p \tilde{\nu}(du_i) \right\|_{L^\infty(X)}^{1/p} = \alpha < +\infty, \tag{1.22}$$

$$\left\| \sum_{i=1}^{d} \int_{|u_i| \le R} |f(X + u_i) - f(X)|^q \tilde{\nu}(du_i) \right\|_{L^\infty(X)}^{1/q} = \beta < +\infty, \tag{1.23}$$

$$\| f(X + u_i) - f(X) \|_{L^\infty(X)} \le \gamma |u_i|, \tag{1.24}$$

for all $u_i \in \mathbb{R}$, $i = 1, \dots, d$, and where $\frac{1}{p} + \frac{1}{q} = 1$. Then,

$$P\big(f(X) - Ef(X) \ge x\big) \le \exp\left(- \frac{x}{4\gamma R} \log\left(1 + \frac{\gamma R x}{2\alpha\beta}\right)\right), \tag{1.25}$$

for all $x > 0$.

Again, when X is a vector of independent identically distributed Poisson random variables and $p = q = 2$, Corollary 3 is known (see [3]), with actually better absolute constants. Extensions, to various infinite dimensional settings, for example, Wiener functionals, Poisson functionals, Bernoulli process, Riemannian path space, of some of the results presented in the present paper can also be found in [10]. Moreover, concentration results for stable vectors are obtained in [8].

2. Proofs. The following covariance representation is crucial to our approach.

LEMMA 1. *Let $X \sim ID(b, 0, \nu)$ be such that $E\|X\|^2 < +\infty$, and let $f, g : \mathbb{R}^d \to \mathbb{R}$ be Lipschitz. Then*

$$\begin{aligned} &Ef(X)g(X) - Ef(X)Eg(X) \\ &\quad = \int_0^1 E_z \int_{\mathbb{R}^d} \big(f(X+u) - f(X)\big)\big(g(Y+u) - g(Y)\big)\nu(du)\,dz, \end{aligned} \tag{2.1}$$

where the expectation E_z is with respect to the $\mathbb{R}^{2d}$ i.d. vector with parameter (b, b) and Lévy measure $\nu_z = z\nu_1 + (1-z)\nu_0$, $0 \le z \le 1$. The measures ν_0 and ν_1 are given by $\nu_0(du, dv) = \nu(du)\delta_0(dv) + \delta_0(du)\nu(dv)$, and $\nu_1(du, dv)$ is the measure ν supported on the main diagonal of $\mathbb{R}^{2d}$, $(u, v) \in \mathbb{R}^{2d}$.

We refer to [9] for a simple proof of this representation in case $f, g \in C_c^\infty$; that is, f and g are compactly supported infinitely differentiable functions and without the finite second moment assumption (it is easily verified at the level of characteristic functions, hence for trigonometric polynomials and by density extended to C_c^∞). Above, the passage from $f, g \in C_c^\infty$ to Lipschitz ones can be done as follows: approximate f Lipschitz by bounded Lipschitz functions f_n where $f_n = f$ if $|f| \le n$, $f_n = n$ if $f \ge n$ and $f_n = -n$ if $f \le -n$, and similarly for g. Next apply to bounded Lipschitz functions the standard mollification argument, that is, convolve them with p_ε, $\varepsilon > 0$, where for $x \in \mathbb{R}^d$, $p_\varepsilon(x) = \frac{1}{\varepsilon^d} p(\frac{x}{\varepsilon})$, and where $p(x) = c\exp(\frac{1}{\|x\|^2 - 1})\mathbf{1}_{\|x\|<1}$ (c normalizing constant). Finally, to pass from C^∞ to C_c^∞, just multiply by a compactly supported C^∞ cutoff function.

The representation (2.1) is proved in [9] for $X \sim ID(b, \Sigma, \nu)$ and, following Chen [4], for functions which are partially differentiable, bounded on bounded subsets of $\mathbb{R}^d$, having also bounded gradient on bounded subsets of $\mathbb{R}^d$ and such that the set of discontinuity of the gradient has probability zero. The reader will also find in [7, 1] further information, references and applications of this representation.

PROOF OF THEOREM 1. Without loss of generality, let $\alpha = 1$. Let $C = \{t \ge 0 : Ee^{t\|X\|} < +\infty\}$, clearly C is convex (an interval) not reduced to $\{0\}$ and moreover,

$$\begin{aligned} C &= \left\{t \ge 0 : \int_{\|u\|>1} e^{t\|u\|}\nu(du) < +\infty\right\} \\ &= \left\{t \ge 0 : \int_{\|u\|>1} (e^{t\|u\|} - t\|u\| - 1)\nu(du) < +\infty\right\}, \end{aligned}$$

where the first equality follows from Theorem 25.3 in [17], while the second is easily verified. Now let us apply the representation formula (2.1) to the function f

which is bounded, Lipschitz (with constant at most 1) with $Ef = 0$ and to $g = e^{tf}$, $t \in C$, $t < M$. Thus,

$$\begin{aligned} Efe^{tf} &= \int_0^1 E_z \int_{\mathbb{R}^d} \big(f(X+u) - f(X)\big)(e^{tf(Y+u)} - e^{tf(Y)})\nu(du)\,dz \\ &\le \int_0^1 E_z e^{tf(Y)} \int_{\mathbb{R}^d} |f(X+u) - f(X)|\,|e^{t(f(Y+u)-f(Y))} - 1|\nu(du)\,dz \\ &\le \int_0^1 E_z e^{tf(Y)} \int_{\mathbb{R}^d} \|u\|(e^{t\|u\|} - 1)\nu(du)\,dz \\ &= \int_{\mathbb{R}^d} \|u\|(e^{t\|u\|} - 1)\nu(du) E e^{tf}, \end{aligned}$$

since the marginal of (X, Y) is X. Hence setting $L(t) = Ee^{tf}$, we have

$$\frac{L'(t)}{L(t)} \le \int_{\mathbb{R}^d} \|u\|(e^{t\|u\|} - 1)\nu(du).$$

Thus,

$$\int_0^t \frac{L'(s)}{L(s)}\,ds \le \int_0^t \int_{\mathbb{R}^d} \|u\|(e^{s\|u\|} - 1)\nu(du)\,ds,$$

that is,

$$Ee^{t(f-Ef)} \le \exp\left(\int_{\mathbb{R}^d} (e^{t\|u\|} - t\|u\| - 1)\nu(du)\right), \tag{2.2}$$

for any bounded Lipschitz function f. To remove the boundedness assumption, let f_n where $f_n = f$ if $|f| \le n$, $f_n = n$ if $f \ge n$ and $f_n = -n$ if $f \le -n$. The f_n are bounded Lipschitz functions (with constant at most 1 since $|f_n(x) - f_n(y)| \le |f(x) - f(y)|$) converging pointwise to f with also Ef_n converging to Ef. By Fatou's lemma, (2.2) holds for any Lipschitz function. [This exponential estimate is sharp since it becomes equality for $X \sim \mathcal{P}(\lambda)$, and $f(X) = X$.] Now, let

$$H(t) = \int_{\mathbb{R}^d} (e^{t\|u\|} - t\|u\| - 1)\nu(du).$$

By the assumptions on $Ee^{t\|X\|}$, H is infinitely differentiable on $(0, M)$ with

$$H'(t) = h(t) = \int_{\mathbb{R}^d} \|u\|(e^{t\|u\|} - 1)\nu(du) > 0,$$

$$H''(t) = \int_{\mathbb{R}^d} \|u\|^2 e^{t\|u\|}\nu(du) > 0,$$

for all $0 < t < M$. Thus, for any f,

$$P\big(f(X) - Ef(X) \ge x\big) \le e^{H(t) - tx},$$

and we now wish to minimize in t. This will be done using the standard Cramér method in large deviation. Indeed, for any $0 < x < h(M^-)$, $\min_{0<t<M}(H(t) - tx) = H(h^{-1}(x)) - xh^{-1}(x)$, since $h(t) - x$ changes sign at $t = h^{-1}(x)$. Hence, since $H(0) = h(0) = h^{-1}(0) = 0$,

$$\begin{aligned} H\big(h^{-1}(x)\big) &= \int_0^{h^{-1}(x)} h(s)\,ds \\ &= \int_0^x s\,dh^{-1}(s) \\ &= xh^{-1}(x) - \int_0^x h^{-1}(s)\,ds. \end{aligned}$$

Thus, $\min_{0<t<M}(H(t) - tx) = -\int_0^x h^{-1}(s)\,ds$, for all $0 < x < h(M^-)$, and the proof is complete. $\square$

PROOF OF COROLLARY 1. Again, without loss of generality, let $\alpha = 1$. Since $\operatorname{supp}\nu \subset B_2(0, R)$, $Ee^{t\|X\|} < +\infty$, for all $t > 0$, and we have

$$\begin{aligned} h(s) &= \int_{\|u\|\le R} (\|u\|e^{s\|u\|} - \|u\|)\nu(du) \\ &= \int_{\|u\|\le R} \|u\|^2\left(\sum_{k=1}^{\infty} \frac{s^k\|u\|^{k-1}}{k!}\right)\nu(du) \\ &\le \int_{\|u\|\le R} \|u\|^2 \sum_{k=1}^{\infty} \frac{s^k R^{k-1}}{k!}\nu(du) \\ &= V^2\left(\frac{e^{sR} - 1}{R}\right). \end{aligned}$$

Hence,

$$\begin{aligned} P\big(f(X) - Ef(X) \ge x\big) &\le \exp\left(-\int_0^x \frac{1}{R}\log\left(1 + \frac{Rs}{V^2}\right)ds\right) \\ &= \exp\left(-\frac{R}{x}\log\left(1 + \frac{Rx}{V^2}\right) + \frac{x}{R}\right. \\ &\qquad \left. - \frac{V^2}{R^2}\log(V^2 + Rx) + \frac{V^2\log V^2}{R^2}\right) \\ &= e^{x/R}\exp\left(-\left(\frac{x}{R} + \frac{V^2}{R^2}\right)\log\left(1 + \frac{Rx}{V^2}\right)\right), \end{aligned}$$

for all $x > 0$. $\square$

PROOF OF COROLLARY 2. Let $\alpha = 1$. First note that $E\|X\|^2 < +\infty$ is equivalent to $V^2 = \int_{\mathbb{R}^d} \|u\|^2\nu(du) < +\infty$. Hence, under the hypothesis of the corollary, $\int_{\mathbb{R}^d}(e^{t\|u\|}-t\|u\|-1)\nu(du)\le\frac{V^2t^2}{2(1-Ct)}, 0<t<\frac{1}{C}$, and thus $Ee^{t\|X\|}<+\infty, 0<t<\frac{1}{C}$. Now we just need to find

$$\min_{0<t<1/C}\left(-tx+\frac{V^2t^2}{2(1-Ct)}\right). \tag{2.3}$$

Solving a quadratic, the above minimum is attained at $t=\frac{1}{C}(1-\frac{V}{\sqrt{V^2+2Cx}})$. Pluging this value in (2.3) we get for minimal value,

$$\begin{aligned}-\frac{1}{2C^2}\left(\sqrt{V^2+2Cx}-V\right)^2&=-\frac{V^2}{2C^2}\left(\sqrt{1+\frac{2Cx}{V^2}}-1\right)^2\\&\le-\frac{V^2}{16C^2}\min\left(\frac{4C^2x^2}{V^4},\frac{2Cx}{V^2}\right)\\&=-\frac{1}{8}\min\left(\frac{2x^2}{V^2},\frac{x}{C}\right).\end{aligned}$$

This last estimate proves the corollary and also provides sharp asymptotics since

$$\frac{-V^2}{2C^2}\left(\sqrt{1+\frac{2Cx}{V^2}}-1\right)^2\ge\frac{-V^2}{8C^2}\min\left(\frac{4C^2x^2}{V^4},\frac{2Cx}{V^2}\right). \qquad\square$$

PROOF OF THEOREM 2. Proceeding as in Theorem 1, for f bounded,

$$\begin{aligned}Efe^{tf}&=\int_0^1E_z\int_{\mathbb{R}^d}\big(f(X+u)-f(X)\big)(e^{tf(Y+u)}-e^{tf(Y)})\nu(du)\,dz\\&\le\int_0^1E_ze^{tf(Y)}\sum_{i=1}^d\int_{\mathbb{R}}|f(X+u_i)-f(X)|(e^{t|f(Y+u_i)-f(Y)|}-1)\tilde{\nu}(du_i)\,dz\\&\le\int_0^1E_ze^{tf(Y)}\sum_{i=1}^d\int_{\mathbb{R}}\beta|u_i|(e^{\beta t|u_i|}-1)\tilde{\nu}(du_i)\,dz\\&=\int_{\mathbb{R}^d}\beta\|u\|(e^{\beta t\|u\|}-1)\nu(du)Ee^{tf}.\end{aligned}$$

The rest of the proof is similar to the proof of Theorem 1; let us just note that $\{t\ge0:Ee^{t\|X\|}<+\infty\}=\{t\ge0:\int_{|u|>1}e^{t|u|}\tilde{\nu}(du)<+\infty\}$, since the components of X are i.i.d. $\square$

PROOF OF THEOREM 3. The proof also proceeds as in Theorem 1. Applying the covariance representation to f bounded Lipschitz such that $Ef=0$ and to $g=e^{tf}$, we see that Efe^{tf} is dominated by

$$\begin{aligned}&\int_0^1E_z\int_{\mathbb{R}^d}\big(f(X+u)-f(X)\big)(e^{tf(Y+u)}-e^{tf(Y)})\nu(du)\,dz\\&\quad\le\int_0^1E_ze^{tf(Y)}\sum_{i=1}^d\int_{\mathbb{R}}|f(X+u_i)-f(X)|(e^{|tf(Y+u_i)-tf(Y)|}-1)\tilde{\nu}(du)\,dz\end{aligned}$$

$$\leq \int_0^1 E_z t e^{tf(Y)} \sum_{i=1}^d \int_{\mathbb{R}} |f(X+u_i) - f(X)|\,|f(Y+u_i) - f(Y)| \times t e^{\gamma t |u_i|} \tilde{\nu}(du_i)\, dz$$

$$= \int_0^1 E_z e^{tf(Y)} \int_{\mathbb{R}^d} |f(X+u) - f(X)|\,|f(Y+u) - f(Y)| t e^{\gamma t \|u\|} \nu(du)\, dz$$

$$\leq \alpha\beta t \left(\int_{\mathbb{R}^d} e^{\gamma r t \|u\|} \nu(du) \right)^{1/r} E e^{tf},$$

where we have used $e^x - 1 \leq x e^x$, $x \geq 0$ and Hölder's inequality. Now, setting $h(t) = \alpha\beta t (\int_{\mathbb{R}^d} e^{\gamma r t \|u\|} \nu(du))^{1/r}$ and proceeding as in the proof of Theorem 1 gives the result. □

PROOF OF COROLLARY 3. Proceeding as in the proof of Theorem 3, we get $h(s) \leq \alpha\beta s e^{\gamma R s}$, for all $s \geq 0$. Instead of working with h^{-1}, it is easier in that case to go back directly to the minimization problem. We want to estimate

$$m = \min_{0 < t < +\infty} \left(-tx + \frac{\alpha\beta}{\gamma^2 R^2} (\gamma R t e^{\gamma R t} - e^{\gamma R t} + 1) \right). \tag{2.4}$$

Now, if $x \leq \frac{2\alpha\beta}{\gamma R}$, choose $t = \frac{x}{2\alpha\beta}$, hence in (2.4), we get

$$\begin{aligned} m &\leq -tx + \frac{\alpha\beta t^2}{2} e^{\gamma R t} \\ &\leq -\frac{x^2}{2\alpha\beta} + \frac{x^2 e}{8\alpha\beta} \\ &\leq -\frac{x^2}{8\alpha\beta}. \end{aligned} \tag{2.5}$$

If $x > \frac{2\alpha\beta}{\gamma R}$, choose $t = \frac{1}{\gamma R} \log \frac{\gamma R x}{2\alpha\beta}$. Then,

$$\begin{aligned} m &\leq -tx + \frac{\alpha\beta}{\gamma^2 R^2} (\gamma R t e^{\gamma R t}) \\ &\leq -\frac{x}{2\gamma R} \log \frac{\gamma R x}{2\alpha\beta}. \end{aligned} \tag{2.6}$$

Finally, combining (2.5) and (2.6) gives (1.25). □

REFERENCES

[1] BOBKOV, S. G., GÖTZE, F. and HOUDRÉ, C. (2001). On Gaussian and Bernoulli covariance representations. *Bernoulli* **7** 439–451.

[2] BOBKOV, S. G. and LEDOUX, M. (1997). Poincaré inequalities and Talagrand's concentration phenomenon for the exponential measure. *Probab. Theory Related Fields* **107** 383–400.

[3] BOBKOV, S. G. and LEDOUX, M. (1998). On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.* **156** 347–365.

[4] CHEN, L. H. Y. (1985). Poincaré-type inequalities via stochastic integrals. *Z. Wahrsch. Verw. Gebiete* **69** 251–277.

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

[6] DUDLEY, R. (1989). *Real Analysis and Probability*. Wadsworth Brooks/Cole, Pacific Grove, CA.

[7] HOUDRÉ, C. (1998). Comparison and deviation from a representation formula. In *Stochastic Processes and Related Topics*: *A Volume in Memory of Stamatis Cambanis, 1943–1995* (I. Karatzas, B. Rajput and M. Taqqu, eds.) 207–218. Birkhäuser, Boston.

[8] HOODRÉ, C. and MARCHAL, P. (2002). On the concentration of measure phenomenon for stable and related random vectors. Preprint.

[9] HOUDRÉ, C., PÉREZ-ABREU, V. and SURGAILIS, D. (1998). Interpolation, correlation identities and inequalities for infinitely divisible variables. *J. Fourier Anal. Appl.* **4** 651–668.

[10] HOUDRÉ, C. and PRIVAULT, N. (2001). Concentration and deviation inequalities in infinite dimensions: An approach via covariance representations. *Bernoulli*. To appear.

[11] HOUDRÉ, C. and PRIVAULT, N. (2001). A deviation inequality on Riemannian path space. Preprint.

[12] LEDOUX, M. (1996). Isoperimetry and Gaussian analysis. *Ecole d'été de Probabilités de St. Flour 1994. Lecture Notes in Math.* **1648** 165–294. Springer, Berlin.

[13] LEDOUX, M. (1999). Concentration of measure and logarithmic Sobolev inequalities. *Séminaire de Probabilités XXXIII. Lecture Notes in Math.* **1709** 120–216. Springer, Berlin.

[14] OLVER, F. (1974). *Asymptotics and Special Functions.* Academic Press, New York.

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

[16] ROSIŃSKI, J. (1995). Remarks on strong exponential integrability of vector-valued random series and triangular arrays. *Ann. Probab.* **23** 464–473.

[17] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions.* Cambridge Univ. Press.

[18] TAKANO, K. (1988). On the Lévy representation of the characteristics function of the probability distribution $Ce^{-|x|}\,dx$. *Bull. Fac. Sci. Ibaraki Univ.* **20** 61–65.

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

[20] TALAGRAND, M. (1991). A new isoperimetric inequality for product measure, and the concentration of measure phenomenon. Israel Seminar (GAFA). *Lecture Notes in Math.* **1469** 91–124. Springer, Berlin.

SCHOOL OF MATHEMATICS
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332
E-MAIL: houdre@math.gatech.edu