Journal of Mathematical Sciences, Vol. 199, No. 2, May, 2014

BOUNDS ON THE MAXIMUM OF THE DENSITY FOR SUMS OF INDEPENDENT RANDOM VARIABLES

S. G. Bobkov[*] and G. P. Chistyakov[†]

UDC 519.2

Sublinear bounds on the maximum of the density for sums of independent random variables are given in terms of the maxima of the densities of summands. Bibliography: 18 titles.

Given a random vector X in Euclidean space $\mathbf{R}^d$ with density p, let

$$M(X) = M(p) = \operatorname{ess\,sup}_x p(x).$$

In other cases, when the distribution X is not absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}^d$, put $M(X) = \infty$.

The aim of this paper is to attract reader's attention to a general property of the functional M which regulates its possible behavior on sums of independent variables or vectors. From our point of view, this property is useful even in the one-dimensional case.

Theorem 1. *For any independent random vectors $X_1, \dots, X_n$ in $\mathbf{R}^d$,*

$$M^{-\frac{2}{d}}(X_1 + \cdots + X_n) \geq \frac{1}{e} \sum_{k=1}^n M^{-\frac{2}{d}}(X_k). \tag{1}$$

Therefore, for increasing sums, the value $M^{-\frac{2}{d}}$ grows linearly or faster than linearly with respect to $M^{-\frac{2}{d}}(X_k)$.

We can draw an obvious analogy between (1) and a variety of other famous inequalities for sums of independent random vectors, usually having the form

$$L(X_1 + \cdots + X_n) \geq \sum_{k=1}^n L(X_k). \tag{2}$$

For example, (2) is true for

$$L(X) = \exp\left[\frac{2}{d}\, h(X)\right], \tag{3}$$

where $h(X) = -\int p(x) \log p(x)\, dx$ is the Shannon entropy. In this case, we come to the so-called "entropy power inequality," an informational variant of the Brunn–Minkowski inequality from convex geometry, which has the same form (see [1–3]). As another example, we point out the work of Stam [4], who obtained inequality (2) for the functional $L(X) = 1/I(X)$, where $I(X)$ is the Fischer information. In both cases, (2) turns into equality on Gaussian distributions with proportional covariance matrices. Regarding the functional $L = 1/M^2$ (for $d = 1$), our main reason for its study were questions of densities behavior in the Erdős–Kac limit theorem for the maximum of increasing sums of independent variables. It is interesting that no assumptions on moments should be made in these inequalities.

The following two statements directly follow from Theorem 1. Since the functional $M^{-\frac{2}{d}}$ is homogeneous of degree 2, i.e.,

$$M^{-\frac{2}{d}}(\lambda X) = \lambda^2 M^{-\frac{2}{d}}(X), \qquad \lambda \in \mathbf{R},$$

we get the following result.

Corollary 2. *Assume that independent vectors X_k in $\mathbf{R}^d$ satisfy the estimate $M(X_k) \leq M$, $\quad 1 \leq k \leq n$. Then*

$$M(a_1 X_1 + \cdots + a_n X_n) \leq e^{d/2}\, M \tag{4}$$

for any $a_k \in \mathbf{R}$ such that $a_1^2 + \cdots + a_n^2 = 1$.

[*]University of Minnesota, Minneapolis, MN, USA, e-mail: bobkov@math.umn.edu.

[†]University of Bielefeld, Bielefeld, Germany, e-mail: chistyak@math.uni-bielefeld.de.

Translated from *Zapiski Nauchnykh Seminarov POMI*, Vol. 408, 2012, pp. 62–73. Original article submitted October 15, 2012.

An example of two independent random variables X_1 and X_2 with the uniform distribution on the interval $(0,1)$ shows that the constant $\frac{1}{e}$ in inequality (1) and, respectively, $e^{d/2}$ in (4) cannot be entirely taken away. It would be interesting to find the best constant in these inequalities or to describe extremal distributions. As we show below, for $n = 2$, the best constant in (1) is $\frac{1}{2}$.

Corollary 3. *Assume that the series $\sum_n X_n$ comprised of independent random vectors in $\mathbf{R}^d$ converges in probability. Then*

$$\sum_n M^{-\frac{2}{d}}(X_n) < \infty.$$

For $d = 1$, the necessary condition is

$$\sum_{n=1}^{\infty} \frac{1}{M^2(X_n)} < \infty. \tag{5}$$

Necessary and sufficient conditions for convergence of a series of independent random variables are well known (e.g., see [5, 6]). For uniformly bounded random variables X_n ($|X_n| \le C$ a.s. for all n with a constant C) with zero expectation such a condition is $\sum_n \mathrm{Var}(X_n) < \infty$. In this case, (5) obviously follows from the known lower-bound estimate:

$$M^2(X)\,\mathrm{Var}(X) \ge \frac{1}{12}$$

(where equality is attained at the uniform distribution on any finite interval; e.g., see [7]). However, the general case is not so obvious.

Now we pass to a variant of Theorem 1 for any (not necessarily bounded) densities. Since the convolution of unbounded densities can be unbounded, it is natural to approximate it with some bounded density, e.g., measuring distances in L^1 metric (i.e., by full variation for the corresponding distributions). We cannot use maximums of the original densities for estimation of the maximum of the approximating density. It turns out that other functionals can be used, such as density quantiles, if the density is considered as a random variable in $\mathbf{R}^d$ with a measure that has the same density. Here is a statement of this kind.

Corollary 4. *Assume that independent random vectors X_k in $\mathbf{R}^d$, $k = 1, \dots, n$, have densities p_k and let m_k be the medians of the random variables $p_k(X_k)$. Then the density p of the sum $X_1 + \dots + X_n$ can be approximated by a bounded density $\widetilde{p}$ so that*

$$\int_{\mathbf{R}^d} |\widetilde{p}(x) - p(x)|\, dx = \frac{1}{2^{n-1}} \tag{6}$$

and

$$M(\widetilde{p}) \le \frac{C_d}{n^{\frac{d}{2}+1}} \sum_{k=1}^{n} m_k$$

with some constant C_d depending only on d.

An equivalent statement is the following: The density p of the normalized sum $(X_1 + \dots + X_n)/\sqrt{n}$ can be approximated by $\widetilde{p}$ so that condition (6) holds, and also

$$M(\widetilde{p}) \le \frac{C_d}{n} \sum_{k=1}^{n} m_k.$$

In the case of identically distributed summands, the right-hand side expression equals $C_d\, m$, where m is the median of the random variable $p_1(X_1)$; hence, this estimate does not depend on n.

The density $\widetilde{p}$ in Corollary 4 can be constructed canonically in some sense, as is shown below. Furthermore, instead of the medians of the random variables $p_k(X_k)$ we can take their quantiles of any given degree with some changes in formulation.

Now we proceed to proofs. Inequality (1) can be obtained using the Young (or Hausdorff–Young) inequality, written with exact constants. Lieb [8] used such an approach to derive (2) for the entropy functional (3). However, in contrast to (2), inequality (1) can hardly be reduced to the case of two summands using induction on n, because it leads to fastly decreasing constants depending on n. For this reason, our departure point is the exact Young inequality for a set of more than two functions.

Let L^ν be the space of all functions u on $\mathbf{R}^d$ with the finite norm

$$\|u\|_\nu = \bigg(\int\limits_{\mathbf{R}^d} |u(x)|^\nu\,dx\bigg)^{1/\nu}, \qquad 1 \le \nu \le \infty.$$

In particular, $\|u\|_\infty = \operatorname{ess\,sup}_x |u(x)|$. Let $\nu' = \frac{\nu}{\nu-1}$ be the adjacent exponent, so that $\frac{1}{\nu} + \frac{1}{\nu'} = 1$. The next important result belongs to Beckner [9] and Brascamp and Lieb [10], see also [11]. We state it below as a lemma, according to [10, Theorem 4]. Let

$$A_\nu = \nu^{1/\nu}(\nu')^{-1/\nu'} \text{ and } A_1 = A_\infty = 1.$$

Lemma 5. *Assume that functions $u_k \in L^{\nu_k}$, $1 \le k \le n$, are given and that*

$$\sum_{k=1}^n \frac{1}{\nu_k'} = \frac{1}{\nu'}, \qquad 1 \le \nu_k, \nu \le \infty.$$

Then the convolution $u = u_1 * \dots * u_n$ belongs to L^ν and has the norm

$$\|u\|_\nu \le \ A\,\|u_1\|_{\nu_1} \dots \|u_n\|_{\nu_n},$$

where

$$A = (A_{\nu_1} \dots A_{\nu_n} A_{\nu'})^{d/2}.$$

Proof of Theorem 1. Without loss of generality, assume that all the X_k have bounded densities p_k. Let $M_k = M(X_k)$. Due to the homogeneity of inequality (1), we may assume that

$$\sum_{k=1}^n M_k^{-\frac{2}{d}} = 1. \tag{7}$$

Let $t_k > 0$ be arbitrary numbers such that $t_1 + \dots + t_n = 1$. We apply Lemma 5 to the functions $u_k = p_k$ choosing ν_k such that

$$\frac{1}{\nu_k'} = t_k, \quad 1 \le k \le n, \ \nu' = 1, \text{ and } \nu = \infty.$$

All the conditions of Lemma 5 hold true, and for the density p of an arbitrary vector $S_n = X_1 + \dots + X_n$ we get the inequality

$$\|p\|_\infty \le A\,\|p_1\|_{\nu_1} \dots \|p_n\|_{\nu_n} \tag{8}$$

with constant $A = (A_{\nu_1} \dots A_{\nu_n})^{d/2}$. To estimate the right-hand side of this inequality, we notice that

$$\begin{aligned} \|p_k\|_{\nu_k} \ &= \ \bigg(\int\limits_{\mathbf{R}^d} p_k(x) \cdot p_k(x)^{\nu_k - 1}\,dx\bigg)^{1/\nu_k} \\ &\le \ \bigg(\int\limits_{\mathbf{R}^d} p_k(x) \cdot M_k^{\nu_k - 1}\,dx\bigg)^{1/\nu_k} \ = \ M_k^{\frac{\nu_k - 1}{\nu_k}} \ = \ M_k^{t_k}. \end{aligned}$$

Setting $s_k = 1 - t_k$, we write the definition of the constants A_ν for $\nu = \nu_k$ in the form

$$A_{\nu_k} = \frac{(\frac{1}{\nu_k'})^{1/\nu_k'}}{(\frac{1}{\nu_k})^{1/\nu_k}} = \frac{t_k^{t_k}}{s_k^{s_k}}.$$

Therefore, from (8) we derive the inequality

$$M(S_n) \ \le \ \bigg(\prod_{k=1}^n \frac{t_k^{t_k}}{s_k^{s_k}}\bigg)^{d/2} M_1^{t_1} \cdots M_n^{t_n},$$

or, what is the same,

$$\log\big(M(S_n)^{-\frac{2}{d}}\big) \ \ge \ \sum_{k=1}^n t_k \log\big(M_k^{-\frac{2}{d}}\big) + \sum_{k=1}^n s_k \log s_k - \sum_{k=1}^n t_k \log t_k. \tag{9}$$

In the next step, we optimize this inequality over the values t_k within the simplex

$$\Delta_n = \{\, t = (t_1, \dots, t_n) :\ t_k \ge 0,\ t_1 + \dots + t_n = 1\}.$$

We can apply (9) with $t_k = M_k^{-\frac{2}{d}}$ without a big loss, which is justified by assumption (7). In this case, inequality (9) is essentially simpler:

$$\log\big(M(S_n)^{-\frac{2}{d}}\big) \ \geq\ \psi_n(t) \ \equiv\ \sum_{k=1}^{n} s_k \log s_k. \tag{10}$$

Therefore, it is sufficient to estimate the lower bound of the right-hand side of (10) uniformly over all $t \in \Delta_n$ for $n \geq 2$.

The function ψ_n is obviously convex, and it turns into ψ_{n-1} on the boundary of the simplex. Thus, we consider ψ_n as a function of $n-1$ variables $t_1, \dots, t_{n-1}$ in the domain $t_k > 0$, $t_1 + \dots + t_{n-1} < 1$, assuming that $t_n = 1 - (t_1 + \dots + t_{n-1})$. We have

$$\frac{\partial \psi_n(t)}{\partial t_k} = -\log s_k + \log s_n = 0, \qquad 1 \leq k \leq n-1,$$

at the minimum point, which is possible if and only if all t_k are equal for $k \leq n-1$. Since ψ_n is invariant under permutations of coordinates, $t_k = \frac{1}{n}$ at the minimum point, i.e., $s_k = 1 - \frac{1}{n}$. Therefore, from (10) we obtain the estimate

$$\log\big(M^{-\frac{2}{d}}(S_n)\big) \ \geq\ \inf_{n\geq 2}\ \inf_{t\in\Delta_n}\ \psi_n(t) \ =\ \inf_{n\geq 2}\ (n-1)\log\big(1 - \frac{1}{n}\big).$$

It remains to notice that $(1 - \frac{1}{n})^{n-1} > \frac{1}{e}$. Theorem 1 is proved. $\square$

Remark. For $n = 2$ we have $\inf_{t\in\Delta_n} \psi_n(t) = \psi_2(\frac{1}{2}, \frac{1}{2}) = -\log 2$; hence, Theorem 1 can be specified in the case of two summands: For any independent random vectors X and Y in $\mathbf{R}^d$,

$$M^{-\frac{2}{d}}(X+Y) \geq \frac{1}{2}\,\big(M^{-\frac{2}{d}}(X) + M^{-\frac{2}{d}}(Y)\big).$$

This inequality is optimal in the sense that equality is attained, in fact, when X and Y are uniformly distributed in the cube $[0,1]^d$. However, this case is obvious since we have a stronger elementary estimate:

$$M^{-\frac{2}{d}}(X+Y) \geq \max\big\{M^{-\frac{2}{d}}(X), M^{-\frac{2}{d}}(Y)\big\}.$$

Proof of Corollary 3. We assume that a random vector X_{n_0} has bounded density for some n_0; otherwise, $M(X_n) = \infty$ for all n and there is nothing to prove.

Let $S = \sum_{n=1}^{\infty} X_n$, where the series converges in probability (hence, with probability one). Consider the partial sums $S_n = \sum_{k=1}^{n} X_k$ and let $R_n = \sum_{k=n+1}^{\infty} X_k$, so that $S = S_n + R_n$. For $n \geq n_0$, the random vectors S_n (and so S) have absolutely continuous distributions. In addition, $0 < M(S) \leq M(S_n)$ because the density maximum cannot increase due to convolution multiplication. Applying Theorem 1, we obtain the estimates

$$M^{-\frac{2}{d}}(S) \geq M^{-\frac{2}{d}}(S_n) \geq \frac{1}{e}\sum_{k=1}^{n} M^{-\frac{2}{d}}(X_k),$$

and, therefore, reach the required conclusion. $\square$

Proof of Corollary 4. Quantile generalization. Fix a value $0 < \delta < 1$. Let m_k be quantiles of the random variables $p_k(X_k)$ of degree δ, i.e., any numbers that satisfy the inequalities

$$\int\limits_{p_k(x)<m_k} p_k(x)\,dx \leq \delta \leq \int\limits_{p_k(x)\leq m_k} p_k(x)\,dx.$$

For any k we divide $\mathbf{R}^d$ into two measurable parts $A_k \subset \{x : p_k(x) \leq m_k\}$ and $B_k \subset \{x : p_k(x) \geq m_k\}$ so that

$$\int\limits_{A_k} p_k(x)\,dx = \delta \quad \text{and} \quad \int\limits_{B_k} p_k(x)\,dx = 1 - \delta.$$

We get the representation

$$p_k(x) = \delta p_{k0}(x) + (1-\delta)p_{k1}(x),$$

where p_{k0} and p_{k1} are defined as normalized restrictions of the density p_k on the sets A_k and B_k, respectively, with $M(p_{k0}) \leq m_k$. Assuming that

$$q_\varepsilon = \big(p_{10}^{\varepsilon_1} * p_{11}^{1-\varepsilon_1}\big) * \dots * \big(p_{n0}^{\varepsilon_n} * p_{n1}^{1-\varepsilon_n}\big), \qquad \varepsilon_k \in \{0,1\},$$

we obtain a representation for the convolution:

$$p = p_1 * \cdots * p_n = \sum_{\varepsilon} \delta^{n(\varepsilon)}(1-\delta)^{n-n(\varepsilon)}\, q_\varepsilon, \tag{11}$$

where the summation is performed over all possible sequences $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ of zeros and ones, using the notation

$$n(\varepsilon) = \varepsilon_1 + \cdots + \varepsilon_n$$

for the number of ones in a sequence ε.

We remove from expression (11) the summand $(1-\delta)^n\, p_{11} * \cdots * p_{n1}$ corresponding to the sequence $\varepsilon = (0, \dots, 0)$. Notice that convolution can be an unbounded function only in the case of this sequence. Thus, we can take the density

$$\widetilde{p} = \frac{1}{1-(1-\delta)^n} \sum_{n(\varepsilon)\ge 1} q_\varepsilon \tag{12}$$

as a canonical approximation for p, where the normalizing constant satisfies the condition $\int \widetilde{p}(x)\,dx = 1$. By construction,

$$\int_{\mathbf{R}^d} |\widetilde{p}(x) - p(x)|\, dx = 2\,(1-\delta)^n,$$

which corresponds to condition (6) in the case of $\delta = \frac{1}{2}$.

To estimate the maximum of the density $\widetilde{p}$, we divide the sum in (12) into two parts. To begin with, notice that the density maximum can only increase due to removing convolution factors $p_{k1}^{1-\varepsilon_k}$ from the density q_ε. Moreover, applying Theorem 1 to p_{k0}, we get the estimates

$$M(q_\varepsilon) \le M\big(p_{10}^{\varepsilon_1} * \cdots * p_{n0}^{\varepsilon_n}\big) \le \Big(\sum_{k=1}^n \varepsilon_k m_k^{-\frac{2}{d}}\Big)^{-\frac{d}{2}} = n(\varepsilon)^{-\frac{d}{2}} \Big(\frac{1}{n(\varepsilon)} \sum_{k=1}^n \varepsilon_k m_k^{-\frac{2}{d}}\Big)^{-\frac{d}{2}}.$$

There are $n(\varepsilon)$ summands in the last sum. Using the monotonicity of the functions $\alpha \to (\mathbf{E}\,\xi^\alpha)^{1/\alpha}$, the right-hand side expression can be estimated by the value

$$n(\varepsilon)^{-\frac{d}{2}}\, \frac{1}{n(\varepsilon)} \sum_{k=1}^n \varepsilon_k m_k \le \Big(\frac{2}{\delta n}\Big)^{\frac{d}{2}+1} \sum_{k=1}^n m_k$$

under the assumption that $n(\varepsilon) \ge \frac{\delta n}{2}$. For the values $1 \le n(\varepsilon) < \frac{\delta n}{2}$ we use the rough estimate

$$M(q_\varepsilon) \le \min_{k:\, \varepsilon_k = 1} M(p_{k0}) \le \min_{k:\, \varepsilon_k=1} m_k \le \sum_{k=1}^n m_k.$$

Combining both estimates, we obtain the inequality

$$M(\widetilde{p}) \le \frac{1}{1-(1-\delta)^n} \bigg(\Big(\frac{2}{\delta n}\Big)^{\frac{d}{2}+1} + \sum_{1 \le n(\varepsilon) < \frac{\delta n}{2}} \delta^{n(\varepsilon)}(1-\delta)^{n-n(\varepsilon)}\bigg) \sum_{k=1}^n m_k.$$

Now we apply a well-known inequality for probabilities of deviations of sums of Bernoulli random variables ξ_k taking values 0 and 1 with probabilities $\mathbf{P}\{\xi_k = 0\} = \delta$ and $\mathbf{P}\{\xi_k = 1\} = 1-\delta$. Specifically (e.g., see [12]),

$$\mathbf{P}\Big\{\frac{1}{\sqrt{n}} \sum_{k=1}^n (\xi_k - \delta) \le -r\Big\} \le e^{-2r^2}, \quad r \ge 0.$$

In particular,

$$\sum_{n(\varepsilon) \le \frac{\delta n}{2}} \delta^{n(\varepsilon)}(1-\delta)^{n-n(\varepsilon)} = \mathbf{P}\Big\{\sum_{k=1}^n \xi_k \le \frac{\delta n}{2}\Big\} \le e^{-n\delta^2/2}.$$

Since $1-(1-\delta)^n \ge \delta$, it follows that

$$M(\widetilde{p}) \le \frac{1}{\delta}\bigg(\Big(\frac{2}{\delta n}\Big)^{\frac{d}{2}+1} + e^{-n\delta^2/2}\bigg) \sum_{k=1}^n m_k.$$

It remains to estimate the factor at $\sum_{k=1}^n m_k$. Making the substitution $n = \frac{2x}{\delta^2}$ and considering as x an arbitrary positive number, we see that

$$\sup_n \left(\frac{\delta n}{2}\right)^{\frac{d}{2}+1} e^{-n\delta^2/2} \le \delta^{-(\frac{d}{2}+1)} \sup_{x>0} x^{\frac{d}{2}+1} e^{-x} = \left(\frac{d+2}{2e\delta}\right)^{\frac{d}{2}+1}.$$

Therefore,

$$\begin{aligned} M(\widetilde{p}) &\le \frac{1}{\delta}\left(\frac{2}{\delta n}\right)^{\frac{d}{2}+1}\left(1+\left(\frac{d+2}{2e\delta}\right)^{\frac{d}{2}+1}\right) \\ &\le \frac{1}{\delta}\left(\frac{1}{\delta^2 n}\right)^{\frac{d}{2}+1}\left(2^{\frac{d}{2}+1}+\left(\frac{d+2}{e}\right)^{\frac{d}{2}+1}\right) < \frac{1}{\delta}\left(\frac{4d}{\delta^2 n}\right)^{\frac{d}{2}+1}, \end{aligned}$$

and we obtain an extended variant of Corollary 4. □

Corollary 6. *Assume that independent random vectors X_k in $\mathbf{R}^d$, $k = 1, \dots, n$, have densities p_k, and m_k are the quantiles of the random variables $p_k(X_k)$ of degree $0 < \delta < 1$. Then the density p of the sum $X_1 + \dots + X_n$ can be approximated by a bounded density $\widetilde{p}$ so that*

$$\int_{\mathbf{R}^d} |\widetilde{p}(x) - p(x)|\, dx = 2\,(1-\delta)^n,$$

and also

$$M(\widetilde{p}) \le \frac{C_d(\delta)}{n^{\frac{d}{2}+1}} \sum_{k=1}^n m_k$$

with the constant $C_d(\delta) = \frac{1}{\delta}\left(\frac{4d}{\delta^2}\right)^{\frac{d}{2}+1}$.

In particular, for $\delta \ge \frac{1}{2}$ we obtain the estimate $C_d(\delta) \le 2\,(16\,d)^{\frac{d}{2}+1}$, independent of δ. The case where the value δ is sufficiently close to 1 is, in fact, specified by some applications. Finally, we notice that letting $\delta \to 1$, $\widetilde{p} = p$ in the limit, and Corollary 6 brings us back to a weaker case of Theorem 1.

Remark. After delivering this paper to press, we found out a work of B. A. Rogozin [13], where the following delicate theorem is proved (as an extension and development of results obtained in [14]). Assume that $S_n = X_1 + \dots + X_n$ is a sum of independent random variables with fixed finite $M_k = M(X_k)$. Then the value $M(S_n)$ is maximized in the case where every X_k is uniformly distributed on an interval of length $1/M_k$. Therefore,

$$M(S_n) \le M(S'_n) = M(X'_1 + \dots + X'_n),$$

where X'_k are independent and uniformly distributed on $(-\frac{1}{2M_k}, \frac{1}{2M_k})$.

This result can be used for specifying Theorem 1 in the case of $d = 1$ after estimation of $M(S'_n)$ by the variance $\mathrm{Var}(S'_n)$. Due to the Hensley conjecture and Busemann–Petty problem, this problem was studied by Ball in [15, 16], where he proved the following. Assume that independent random variables $\xi_1, \dots, \xi_n$ are uniformly distributed on the interval $(-\frac{1}{2}, \frac{1}{2})$ and let $S'_n = a_1\xi_1 + \dots + a_n\xi_n$ with $a_1^2 + \dots + a_n^2 = 1$. Then

$$1 \le M(S'_n) \le \sqrt{2}.$$

Combining the right-hand side inequality with the Rogozin theorem, we get the estimate

$$\frac{1}{M^2(S_n)} \ge \frac{1}{2} \sum_{k=1}^n \frac{1}{M^2(X_k)},$$

where the constant $\frac{1}{2}$ appears to be the best.

Note that in dimension 1, inequality (1) up to an absolute factor also follows from some estimates for the concentration function, see [17–18].

The first author's work is supported by the NSF and Simons Foundation. The second author's work is supported by SFB-701 of the Bielefeld University.

Translated by K. S. Pilyugin.

REFERENCES

1. A. Dembo, T. M. Cover, and J. A. Thomas, "Information-theoretic inequalities," *IEEE Trans. Inform. Theory*, **37**, 1501–1518 (1991).
2. O. Johnson, *Information Theory and the Central Limit Theorem*, Imperial College Press, London (2004).
3. S. G. Bobkov and M. Madiman, "Reverse Brunn–Minkowski and reverse entropy power inequalities for convex measures," *J. Funct. Anal.*, **262**, 3309–3339 (2012).
4. A. J. Stam, "Some inequalities satisfied by the quantities of information of Fisher and Shannon," *Information Control*, **2**, 101–112 (1959).
5. V. V. Petrov, *Sums of Independent Random Values* [in Russian], Moscow (1972).
6. A. N. Shiryaev, *Probability* [in Russian], Moscow (1980).
7. S. G. Bobkov, G. P. Chistyakov, and F. Götze, "Bounds for characteristic functions in terms of quantiles and entropy," *Electron. Commun. Probab.*, **17**, (2012).
8. E. H. Lieb, "Proof of an entropy conjecture of Wehrl," *Comm. Math. Phys.*, **62**, 35–41 (1978).
9. W. Beckner, "Inequalities in Fourier analysis," *Ann. Math.*, **2**, 159–182 (1975).
10. H. J. Brascamp and E. H. Lieb, "Best constants in Young's inequality, its converse, and its generalization to more than three functions," *Advances Math.*, **20**, 151–173 (1976).
11. F. Barthe, "Optimal Young's inequality and its converse: a simple proof," *Geom. Funct. Anal.*, **8**, 234–242 (1998).
12. S. G. Bobkov, C. Houdre, and P. Tetali, "The subgaussian constant and concentration inequalities," *Israel J. Math.*, **156**, 255–283 (2006).
13. B. A. Rogozin, "An estimate for the maximum of the convolution for bounded densities," *Teor. Veroyatn. Primen.*, **32**, 53–61 (1987).
14. Yu. V. Prokhorov, "Extremal problems in limit theorems," in: *Proc. Sixth All-Union Conf. Theory Prob. and Math. Statist.* (Vilnius, 1960), Vilnius (1962), pp. 77–84.
15. K. Ball, "Cube slicing in R^n," *Proc. Amer. Math. Soc.*, **97**, 465–473 (1986).
16. K. Ball, "Some remarks on the geometry of convex sets," *Lect. Notes Math.*, **1317**, 224–231 (1988).
17. L. P. Postnikova and A. A. Yudin, "A sharpened form of an inequality for the concentration function," *Teor. Veroyatn. Primen.*, **23**, 376–379 (1978).
18. A. L. Miroshnikov and B. A. Rogozin, "Inequalities for the concentration function," *Teor. Veroyatn. Primen.*, **25**, 178–183 (1980).