ESTIMATES FOR THE CONCENTRATION FUNCTIONS IN THE LITTLEWOOD–OFFORD PROBLEM

YULIA S. ELISEEVA, FRIEDRICH GÖTZE, AND ANDREI YU. ZAITSEV

ABSTRACT. Let $X, X_1, \ldots, X_n$ be independent identically distributed random variables. In this paper we study the behavior of the concentration functions of the weighted sums $\sum_{k=1}^{n} a_k X_k$ with respect to the arithmetic structure of coefficients $a_k$. Such concentration results recently became important in connection with investigations about singular values of random matrices. In this paper we formulate and prove some refinements of a result of Vershynin (R. Vershynin, Invertibility of symmetric random matrices, arXiv:1102.0300. (2011). To appear in Random Structures and Algorithms).

1. Introduction

This paper is an extended and modified version of preprint [6].

Let $X, X_1, \ldots, X_n$ be independent identically distributed (i.i.d.) random variables with common distribution $F = \mathcal{L}(X)$. The Lévy concentration function of a random variable $X$ is defined by the equality

$$Q(F, \lambda) = \sup_{x \in \mathbb{R}} F\{[x, x + \lambda]\}, \quad \lambda > 0.$$ 

Let $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$, $a \neq 0$. In this paper we study the behavior of the concentration functions of the weighted sums $S_a = \sum_{k=1}^{n} a_k X_k$ with respect to the arithmetic structure of coefficients $a_k$. Refined concentration results for these weighted sums play an important role in the study of singular values of random matrices (see, for instance, Nguyen and Vu [19], Rudelson and Vershynin [22, 23], Tao and Vu [24, 25], Vershynin [26]). In this context the problem is referred to as the Littlewood–Offord problem (see also [7, 13, 16]).

In the sequel, let $F_a$ denote the distribution of the sum $S_a$, and let $G$ be the distribution of the symmetrized random variable $\tilde{X} = X_1 - X_2$. Let

$$M(\tau) = \tau^{-2} \int_{|x| \leq \tau} x^2 G(dx) + \int_{|x| > \tau} G(dx) = E \min \{\tilde{X}^2 / \tau^2, 1\}, \quad \tau > 0. \quad (1)$$

The symbol $c$ will be used for absolute positive constants. Note that $c$ can be different in different (or even in the same) formulas. We will write $A \ll B$ if $A \leq cB$. Also we will write

1991 Mathematics Subject Classification. Primary 60F05; secondary 60E15, 60G50.

Key words and phrases. concentration functions, inequalities, the Littlewood–Offord problem, sums of independent random variables.
$A \preceq B$ if $A \ll B$ and $B \ll A$. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ we will denote $\|x\|^2 = x_1^2 + \cdots + x_n^2$ and $\|x\|_\infty = \max_j |x_j|$.

The elementary properties of concentration functions are well studied (see, for instance, [2, 14, 20]). In particular, it is obvious that
\[ Q(F, \mu) \leq (1 + \lfloor \mu/\lambda \rfloor) Q(F, \lambda) \]
for any $\mu, \lambda > 0$, where $\lfloor x \rfloor$ is the integer part of a number $x$. Hence, \[ Q(F, c\lambda) \preceq Q(F, \lambda) \quad (2) \]
and \[ \text{if } Q(F, \lambda) \ll B, \text{ then } Q(F, \mu) \ll B (1 + \mu/\lambda). \quad (3) \]

The problem of estimating the concentration function of weighted sums $S_a$ under different restrictions on the vector $a \in \mathbb{R}^n$ and distributions of summands has been studied in [10, 19, 23, 22, 25, 26]. Eliseeva and Zaitsev [5] (see also [4]) obtained some improvements of the results [10, 23]. In this paper we formulate and prove similar refinements of a result of Vershynin [26].

Note that a relation between the rate of decay of the concentration function and the arithmetic structure of distributions of independent random variables was discovered for arbitrary distributions of summands in a paper of Arak [1] (see also [2, 27]). Much later, similar relations was found in [19, 22, 23, 24, 25, 26] in the particular case of distributions involved in the Littlewood–Offord problem. The authors of the present paper are going to devote a separate publication to compare the results of aforementioned papers.

Let $\log_+ (x) = \max \{0, \log x\}$. The result of Vershynin [26], related to the Littlewood–Offord problem, is formulated as follows.

**Proposition 1.** Let $X, X_1, \ldots, X_n$ be i.i.d. random variables and $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ with $\|a\| = 1$. Assume that there exist positive numbers $\tau, p, K, L, D$ such that $Q(L(X), \tau) \leq 1 - p$, $E |X| \leq K$, and
\[ \|ta - m\| \geq L \sqrt{\log_+ (t/L)} \quad \text{for all } m \in \mathbb{Z}^n \text{ and } t \in (0, D]. \quad (4) \]
If $L^2 \geq 1/p$, then
\[ Q(F_a, \frac{1}{D}) \leq \frac{CL}{D}, \quad (5) \]
where the quantity $C$ depends on $\tau, p, K$ only.

**Corollary 1.** Let the conditions of Proposition 1 be satisfied. Then, for any $\varepsilon \geq 0$,
\[ Q(F_a, \varepsilon) \ll CL \left( \varepsilon + \frac{1}{D} \right). \quad (6) \]

It is clear that if
\[ 0 < D \leq D(a) = D_L(a) = \inf \left\{ t > 0 : \text{dist}(ta, \mathbb{Z}^n) < L \sqrt{\log_+ (t/L)} \right\}, \quad (7) \]
where
\[ \text{dist} \left( ta, Z^n \right) = \min_{m \in Z^n} \| ta - m \| = \left( \sum_{k=1}^{n} \min_{m_k \in Z} | ta_k - m_k |^2 \right)^{1/2}, \]
then condition (1) holds. In paper [26] the quantity \( D(a) \) is called the least common denominator of the vector \( a \in \mathbb{R}^n \) (see also [22, 23] for similar definitions).

Note that for \( |t| \leq 1/2 \|a\|_{\infty} \) we have
\[ \left( \text{dist}(ta, Z^n) \right)^2 = \sum_{k=1}^{n} |ta_k|^2 = \|a\|^2 t^2 = t^2. \] (8)

By definition, \( D(a) > L \). Moreover, equality (8) implies that \( D(a) \geq 1/2 \|a\|_{\infty} \) (see [26], Lemma 6.2).

Note that just the statement of Corollary 1 with \( D = D(a) \) is presented as the corresponding concentration result for the Littlewood–Offord problem in [26]. The formulation of Proposition 1 is more natural than the statement of Corollary 1. Furthermore, Proposition 1 implies Corollary 1 using relations (3) and (7). Minimal \( L \), for which Proposition 1 holds, depends on \( a \) and \( D \). Moreover, generally speaking, it can be essentially larger than \( p^{-1/2} \).

In the formulation of Proposition 1 w.l.o.g. we can replace assumption (4) by the following:
\[ \| ta - m \| \geq f_L(t) \quad \text{for all } m \in Z^n \text{ and } t \in \left[ \frac{1}{2 \|a\|_{\infty}}, D \right], \] (9)
where
\[ f_L(t) = \begin{cases} t/6, & \text{for } 0 < t < eL, \\ L \sqrt{\log(t/L)}, & \text{for } t \geq eL. \end{cases} \] (10)

Note that equality (8) justifies why the assumption \( t \geq 1/2 \|a\|_{\infty} \) in condition (9) is natural. For \( 0 < t < 1/2 \|a\|_{\infty} \), inequality (9) is satisfied automatically.

Formally, condition (9) can be more restrictive than condition (4). However, if condition (4) is satisfied, but condition (9) is not satisfied, then inequality (6) holds for trivial reasons.

Indeed, if \( t \geq eL \), then condition (9) for such a \( t \) follows from assumption (4). If \( 0 < t < eL \) and there exists an \( m \in Z^n \) such that \( \| ta - m \| < t/6 \), then, denoting \( k = \lfloor eL/t \rfloor + 1 \), we have \( tk \geq eL \) and
\[ \| tk - m \| < tk/6 \leq 2eL/6 < L \leq L^{\sqrt{\log_+ (tk/L)}}. \]

Since \( km \in Z^n \), we have \( D \leq D(a) \leq tk < 6L \) and the required inequality (3) is a consequence of \( Q(F_a, 1/D) \leq 1 \).

Note that there may be a situation such that condition (9) is satisfied, but condition (4) is not satisfied, for some \( t \) from the interval \( L < t < eL \). Then the estimates for the concentration functions in Proposition 1 and Corollary 1 still hold. This follows from Theorem 1 of this paper.
The above argument justifies that it would be reasonable to define the alternative least common denominator as
\[ D^*(a) = \inf \left\{ t > 0 : \text{dist}(ta, Z^n) < f_L(t\|a\|) \right\}. \tag{11} \]
This definition will be also used below in the case when \( \|a\| \neq 1 \). Obviously,
\[ D^*(\lambda a) = D^*(a)/\lambda, \quad \text{for any } \lambda > 0, \tag{12} \]
and equality (8) implies again that \( D^*(a) \geq 1/2 \|a\|_\infty \).

Now we formulate the main result of this paper.

**Theorem 1.** Let \( X, X_1, \ldots, X_n \) be i.i.d. random variables. Let \( a = (a_1, \ldots, a_n) \in \mathbb{R}^n \) with \( \|a\| = 1 \). Assume that condition (9) is satisfied. If \( L^2 \geq 1/M(1) \), where the quantity \( M(1) \) is defined by formula (1), then
\[ Q \left( F_a, \frac{1}{D} \right) \ll \frac{1}{D\sqrt{M(1)}}. \tag{13} \]

Let us reformulate Theorem 1 for arbitrary \( a \), without assuming that \( \|a\| = 1 \).

**Corollary 2.** Let the conditions of Theorem 1 be satisfied with condition (9) replaced by the condition
\[ \|ta - m\| \geq f_L(t\|a\|) \quad \text{for all } m \in Z^n \text{ and } t \in \left[ \frac{1}{2\|a\|_\infty}, D \right], \tag{14} \]
and without the assumption \( \|a\| = 1 \). If \( L^2 \geq 1/M(1) \), then
\[ Q \left( F_a, \frac{1}{D} \right) \ll \frac{1}{\|a\|D\sqrt{M(1)}}. \tag{15} \]

The proofs of our Theorem 1 and Corollary 2 are similar to the proof of the main results of [5]. They are in a sense more natural than the proofs of Vershynin [26], since they do not use unnecessary assumptions like \( E|X| \leq K \). This is achieved by an application of relation (16). Our proof differs from the arguments used in [10, 23, 26]. We apply the methods developed by Essén [9] (see the proof of Lemma 4 of Chapter II in [20]).

Applying Corollary 2 to the random variables \( X_k/\tau, \tau > 0 \), we obtain the following result.

**Corollary 3.** Let \( V_{a,\tau} = L \left( \sum_{k=1}^n a_kX_k/\tau \right), \tau > 0 \). Then, under the conditions of Corollary 2 with the condition \( L^2 \geq 1/M(1) \) replaced by the condition \( L^2 \geq 1/M(\tau) \), we have
\[ Q \left( V_{a,\tau}, \frac{1}{D} \right) = Q \left( F_a, \frac{\tau}{D} \right) \ll \frac{1}{\|a\|D\sqrt{M(\tau)}}. \tag{16} \]
In particular, if \( \|a\| = 1 \), then
\[ Q \left( F_a, \frac{\tau}{D} \right) \ll \frac{1}{D\sqrt{M(\tau)}}. \tag{17} \]
Estimates for the Concentration Functions

For the proof of Corollary 3, it suffices to use Corollary 2 and relation (1). It is evident that
\[ M(\tau) \gg 1 - Q(G, \tau) \geq 1 - Q(F, \tau) \geq p, \]
where \( p \) is introduced in Proposition 1. Note that \( M(\tau) \) can be essentially larger than \( L \) under the conditions of Corollary 3. Moreover, there is an unnecessary assumption \( \mathbb{E} |X| \leq K \) in the formulation of Proposition 1. Finally, the dependence of constants on the distribution \( L(Y) \) is stated explicitly, in inequalities (13), (15), (17) constants are absolute as opposed to inequalities (5), (6), where the value \( C \) depends on \( \tau, p, K \) not explicitly. An improvement of Corollary 1 is given below in Theorem 2.

We recall now the well-known Kolmogorov–Rogozin inequality (see [2, 14, 20, 21]).

**Proposition 2.** Assume that \( Y_1, \ldots, Y_n \) are independent random variables with the distributions \( W_k = L(Y_k) \). Let \( \lambda_1, \ldots, \lambda_n \) be positive numbers such that \( \lambda_k \leq \lambda \), for \( k = 1, \ldots, n \). Then
\[
Q(L\left(\sum_{k=1}^{n} Y_k\right), \lambda) \ll \lambda \left(\sum_{k=1}^{n} \lambda_k^2 \left(1 - Q(W_k, \lambda_k)\right)\right)^{-1/2}.
\]

Esséen [9] (see [20], Theorem 3 of Chapter III) has improved this result. He has shown that the following statement is true.

**Proposition 3.** Under the conditions of Proposition 2 we have
\[
Q(L\left(\sum_{k=1}^{n} Y_k\right), \lambda) \ll \lambda \left(\sum_{k=1}^{n} \lambda_k^2 M_k(\lambda_k)\right)^{-1/2},
\]
where \( M_k(\tau) = \mathbb{E} \min \{ \wedge_k^2 / \tau^2, 1 \} \).

Further improvements of (18) and (19) may be found in [11, 12, 15, 17, 18, 28].

It is clear that Theorem 1 is related to Proposition 1 in a similar way as Esséen’s inequality (19) is related to the Kolmogorov–Rogozin inequality (18). Moreover, in inequalities (5) and (6), the dependence of \( C \) on \( \tau, p \) and \( K \) is not written out explicitly.

If we consider a special case, where \( D = 1/2 \|a\|_\infty \), then no restrictions on the arithmetic structure of the vector \( a \) are made, and Corollary 3 implies the bound
\[
Q(F_a, \|a\|_\infty \tau) \ll \frac{\|a\|_\infty}{\|a\| \sqrt{M(\tau)}}.
\]
This result follows from Esséen’s inequality (19) applied to the sum of non-identically distributed random variables \( Y_k = a_k X_k \) with \( \lambda_k = a_k \tau, \lambda = \|a\|_\infty \tau \). For \( a_1 = \cdots = a_n = n^{-1/2} \), inequality (20) turns into the well-known particular case of Proposition 3:
\[
Q(F^{*n}, \tau) \ll \frac{1}{\sqrt{n M(\tau)}}.
\]
Inequality (21) implies as well the Kolmogorov–Rogozin inequality for i.i.d. random variables:

\[ Q(F^{*n}, \tau) \ll \frac{1}{\sqrt{n} (1 - Q(F, \tau))}. \]

Inequality (20) is not able to yield a bound of better order than \( O(n^{-1/2}) \), since the right-hand side of (20) is at least \( n^{-1/2} \). The results stated above are more interesting if \( D \) is essentially larger than \( 1/2 \|a\|_\infty \). In this case one can expect the estimates of better order than \( O(n^{-1/2}) \). Just such estimates of \( Q(F_a, \lambda) \) are required to study the distributions of eigenvalues of random matrices.

For \( 0 < D < 1/2 \|a\|_\infty \), the inequality

\[ Q(F_a, \frac{\tau}{D}) \ll \frac{1}{\|a\|D \sqrt{M(\tau)}} \]  

holds assuming the conditions of Corollary 3 too. In this case it follows from (3) and (20).

Under the conditions of Corollary 3, there exist many possibilities to represent a fixed \( \varepsilon \) as \( \varepsilon = \tau/D \) for an application of inequality (16). Therefore, for a fixed \( \varepsilon = \tau/D \) we can try to minimize the right-hand side of inequality (16) choosing optimal \( \tau \) and \( D \). This is possible, and the optimal bound is given in the following Theorem 2.

**Theorem 2.** Let the conditions of Corollary 2 for \( D \leq D^*(a) \) be satisfied except the condition \( L \geq 1/M(1) \). Let \( L > 1/P \), where

\[ P = \mathbb{P}(\bar{X} \neq 0) = \lim_{\tau \to 0} M(\tau). \]

Then there exists a \( \tau_0 \) such that \( L^2 = 1/M(\tau_0) \). Moreover, the bound

\[ Q(F_a, \varepsilon) \ll \frac{1}{\|a\|D^*(a) \sqrt{M(\varepsilon D^*(a))}} \]  

is valid for \( 0 < \varepsilon \leq \varepsilon_0 = \tau_0/D^*(a) \). Furthermore, for \( \varepsilon \geq \varepsilon_0 \), the bound

\[ Q(F_a, \varepsilon) \ll \frac{\varepsilon L}{\varepsilon_0 \|a\|D^*(a)} \]  

holds.

In the statement of Theorem 2, the quantity \( \varepsilon \) can be arbitrarily small. If \( \varepsilon \) tends to zero and \( L^2 > 1/P \), we obtain

\[ Q(F_a, 0) \ll \frac{1}{\|a\|D^*(a) \sqrt{P}} \]  

Applying inequalities (23)–(25), we should take into account that \( \|a\|D^*(a) = D^*(a/\|a\|) \) by virtue of (12).
Theorem 2 follows easily from Corollary 3. Indeed, denoting \( \varepsilon = \tau / D \), we can rewrite inequality (16) as

\[
Q(F_a, \varepsilon) \ll \frac{1}{\|a\| D \sqrt{M(\varepsilon D)}}.
\]  

(26)

Inequality (26) holds if \( L^2 \geq 1/M(\varepsilon D) \) and \( 0 \leq D \leq D^*(a) \). If \( L^2 \geq 1/M(\varepsilon D^*(a)) \), then the choice \( D = D^*(a) \) is optimal in inequality (26) since

\[
D^2 M(\varepsilon D) = E \min \{ \tilde{X}^2 / \varepsilon^2, D^2 \}
\]

is increasing when \( D \) increases. For the same reason, if \( L^2 < 1/M(\varepsilon D^*(a)) \), the optimal choice of \( D \) in inequality (26) is given by the solution \( D_0(\varepsilon) \) of the equation \( L^2 = 1/M(\varepsilon D) \). This solution exists and is unique if \( L^2 > 1/P \), since the function \( M(\tau) \) is continuous and strictly decreasing if \( M(\tau) < P \). Moreover, \( M(\tau) \to 0 \) as \( \tau \to \infty \). In this case inequality (26) turns into

\[
Q(F_a, \varepsilon) \ll \frac{L}{\|a\| D_0(\varepsilon)}.
\]  

(27)

Furthermore, choosing \( \tau_0 \) to be the solution of the equation \( L^2 = 1/M(\tau) \), we see that inequality (24) is valid for \( 0 < \varepsilon \leq \varepsilon_0 = \tau_0 / D^*(a) \). It is clear that \( D_0(\varepsilon_0) = D^*(a) \). Moreover, for \( \varepsilon \geq \varepsilon_0 \), we have

\[
M(\varepsilon D_0(\varepsilon)) = M(\varepsilon_0 D_0(\varepsilon_0)) = L^{-2}
\]

and, hence, \( \varepsilon D_0(\varepsilon) = \varepsilon_0 D_0(\varepsilon_0) \). Therefore, for \( \varepsilon \geq \varepsilon_0 \), inequality (24) holds. The right-hand side of the inequality (24) with \( \|a\| = 1 \) admits the following representations

\[
\frac{\varepsilon L}{\varepsilon_0 D^*(a)} = \frac{L}{D_0(\varepsilon)} = \frac{1}{D_0(\varepsilon) \sqrt{M(\varepsilon D_0(\varepsilon))}}.
\]

Obviously, inequality (24) could be derived from (26) with \( \varepsilon = \varepsilon_0 \) by an application of inequality (3). On the other hand, for \( 0 < \varepsilon_1 < \varepsilon \leq \varepsilon_0 \), we could apply inequality (3) to inequality (23) obtaining the bound

\[
Q(F_a, \varepsilon) \ll \frac{\varepsilon L}{\varepsilon_1 D^*(a)} \ll \frac{\varepsilon}{\varepsilon_1 \|a\| D^*(a) \sqrt{M(\varepsilon_1 D^*(a))}}.
\]  

(28)

However, inequality (28) is weaker than inequality (23) since, evidently,

\[
\varepsilon^2 M(\varepsilon \mu) = E \min \{ \tilde{X}^2 / \mu^2, \varepsilon^2 \} \geq E \min \{ \tilde{X}^2 / \mu^2, \varepsilon_1^2 \} = \varepsilon_1^2 M(\varepsilon_1 \mu),
\]

(29)

for any \( \mu > 0 \).

Theorem 2 is an essential improvement of Corollary 1. In particular, in contrast with inequality (6) of Corollary 1 for small \( \varepsilon \), the right-hand side of inequality (23) of Theorem 2 may be decreasing as \( \varepsilon \) decreases. Moreover, we have just shown that an application of inequality (3) would lead to a loss of precision. Recall that just an application of inequality (3) allows us to derive Corollary 1 from Proposition 1.
Consider a simple example. Let $X$ be the random variable taking values 0 and 1 with probabilities
\[ P\{X = 1\} = 1 - P\{X = 0\} = p > 0. \] (30)
Then
\[ P\{\tilde{X} = \pm 1\} = p(1 - p), \quad P\{\tilde{X} = 0\} = 1 - 2p(1 - p), \] (31)
and the function $M(\tau)$ has the form
\[ M(\tau) = \begin{cases} 2p(1 - p), & \text{for } 0 < \tau < 1, \\ \frac{2p(1 - p)}{\tau^2}, & \text{for } \tau \geq 1. \end{cases} \] (32)
Assume for simplicity that $\|a\| = 1$. If $L^2 > 1/2p(1 - p)$, then $\tau_0 = L\sqrt{2p(1 - p)}$ and, for $\varepsilon \geq \varepsilon_0 = L\sqrt{2p(1 - p)}/D^*(a)$, we have the bound
\[ Q(F_a, \varepsilon) \ll \frac{\varepsilon}{\sqrt{p(1 - p)}}. \] (33)
The same bound (33) follows from inequality (23) of Theorem 2 for $1/D^*(a) \leq \varepsilon \leq \varepsilon_0$. For $0 < \varepsilon \leq 1/D^*(a)$, inequality (23) implies the bound
\[ Q(F_a, \varepsilon) \ll \frac{1}{D^*(a)\sqrt{p(1 - p)}}. \] (34)
Thus,
\[ Q(F_a, \varepsilon) \ll \min \left\{ \frac{1}{\sqrt{p(1 - p)}}(\varepsilon + \frac{1}{D^*(a)}) \right\}, \quad \text{for all } \varepsilon \geq 0. \] (35)
Inequality (35) cannot be essentially improved. Consider, for example,
\[ a = (s^{-1/2}, \ldots, s^{-1/2}, 0, \ldots, 0) \] (36)
with the first $s \leq n$ coordinates equal to $s^{-1/2}$ and the last $n - s$ coordinates equal to zero. In this case $D^*(a) \asymp s^{1/2}$, the random variable $s^{1/2}S_a$ has binomial distribution with parameters $s$ and $p$, and it is well-known that
\[ Q(F_a, \varepsilon) \gg \min \left\{ \frac{1}{\sqrt{p(1 - p)}}(\varepsilon + \frac{1}{\sqrt{s}}) \right\}, \quad \text{for all } \varepsilon \geq 0. \] (37)
Comparing the bounds (35) and (37), we see that Theorem 2 provides the optimal order of $Q(F_a, \varepsilon)$ for all possible values of $\varepsilon$. Moreover, the corresponding constant depend on $p$ optimally.

It would seem that the last example may be reduced to the trivial case $n = s$. This is not quite right. It is clear that the value $Q(F_a, 1)$ does not change significantly after a small perturbation of the vector $a$ (defined in (36)), if the absolute values of the last $n - s$ coordinates of vector $a$ are small but nonzero. Moreover, the order of smallness of the last $n - s$ coordinates can be chosen in such a way that inequalities (35) and (37) are satisfied with $\varepsilon \gg s^{-1}$ and $D^*(a) \asymp s^{1/2}$. 
For the sake of completeness, we give below a short proof of inequality (37). It is easy to see that
\[
\text{Var} S_a = p(1 - p).
\]
The random variable \( S_a \) takes values which are multiples of \( s^{-1/2} \). Therefore, if \( s p(1 - p) \leq 1 \), then inequality (38) implies that \( Q(F_a, 0) \approx 1 \) and inequality (37) is trivially valid.

Assume now \( s p(1 - p) > 1 \). If \( 0 < \varepsilon \leq 4 \sqrt{p(1 - p)} \), then, using (3) and (38), we obtain
\[
3/4 \leq Q(F_a, 4 \sqrt{p(1 - p)}) \ll \varepsilon^{-1} \sqrt{p(1 - p)} Q(F_a, \varepsilon),
\]
and, hence,
\[
Q(F_a, \varepsilon) \gg \frac{\varepsilon}{\sqrt{p(1 - p)}}.
\]
It is clear that (2), (3) and (40) imply that \( Q(F_a, \varepsilon) \approx 1 \), for \( \varepsilon \geq 4 \sqrt{p(1 - p)} \). Applying inequality (40) for \( \varepsilon = s^{-1/2} \) and using the lattice structure of the support of distribution \( F_a \), we conclude that, for \( 0 \leq \varepsilon < s^{-1/2} \),
\[
Q(F_a, \varepsilon) \geq Q(F_a, 0) \gg \frac{1}{\sqrt{s p(1 - p)}}.
\]
Thus, inequalities (2), (3), (40) and (41) imply (37).

The results of this paper are formulated for a fixed \( L \). It is clear that in their application we should try to choose the optimal \( L \), which satisfies the conditions and minimizes the right-hand sides of inequalities for the concentration functions. Recall that the least common denominator \( D^*(a) \) depends on \( L \).

The quantity \( \tau_0 = \varepsilon_0 D^*(a) \) (which is the solution of the equation \( L^2 = 1/M(\tau) \)) may be interpreted as a quantity depending on \( L \) and on the distribution \( L(X) \). Moreover, comparing the bounds (33) and (24) for relatively large values of \( \varepsilon \), we see that \( \tau_0 \to \infty \) as \( L \to \infty \). Therefore, the factor \( L/\tau_0 \) is much smaller than \( L \) for large values of \( L \). In particular, in the example above we have \( \tau_0 = L \sqrt{2 p(1 - p)} \).

Another example would be a symmetric stable distribution with parameter \( \alpha \), \( 0 < \alpha < 2 \). In this case the characteristic function \( \hat{F}(t) = \mathbb{E} \exp(itX) \) has the form \( \hat{F}(t) = \exp(-c|t|^\alpha) \). It could be shown that then \( \tau_0 \) behaves as \( L^{2/\alpha} \) as \( L \to \infty \).

Inequality (33) can be rewritten in the form
\[
Q(F_a, \varepsilon) \ll \frac{\varepsilon}{\sigma}, \quad \text{for} \ \varepsilon \geq \varepsilon_0,
\]
where \( \sigma^2 = \text{Var } X \). It is clear that a similar situation occurs for any random variable \( X \) with finite variance.

In particular, inequality (12) is obviously satisfied for all \( \varepsilon \geq 0 \), if \( \|a\| = 1 \) and \( X \) has a Gaussian distribution with \( \text{Var } X = \sigma^2 \). Moreover, the order of the right-hand side of the
inequality is optimal. In this particular case, for any \( \tau > 0 \) the relation
\[
\frac{1}{\sqrt{M(\tau)}} \asymp 1 + \frac{\tau}{\sigma},
\]
holds. The use of this inequality and of Theorem 2 with \( \|a\| = 1 \) implies easily that
\[
Q(F_a, \varepsilon) \ll \frac{\varepsilon}{\sigma}, \quad \text{for} \quad \varepsilon \geq \frac{\sigma}{D^*(a)}.
\]
Inequality (43) provides the correct dependence of the concentration function on \( \sigma \) for \( \sigma/D^*(a) \leq \varepsilon \leq \sigma \). It is impossible to obtain estimates of such order from inequality (6). Estimate (43) cannot be deduced from Theorem 2 for small \( \varepsilon \), since in Theorem 2 the distribution \( F = \mathcal{L}(X) \) is arbitrary and the concentration function \( Q(F_a, \varepsilon) \) may be not tending to zero as \( \varepsilon \to 0 \) (see (37)).

2. Proofs

We will use the classical Esséen inequalities ([8], see also [14] and [20]):
\[
Q(F, \lambda) \ll \lambda \int_0^{\lambda^{-1}} |\hat{F}(t)| \, dt, \quad \lambda > 0,
\]
where \( \hat{F}(t) \) is the corresponding characteristic function. In the general case, \( Q(F, \lambda) \) cannot be estimated from below by the right hand side of inequality (44). However, if we assume additionally that the distribution \( F \) is symmetric and its characteristic function is non-negative for all \( t \in \mathbb{R} \), then we have the lower bound:
\[
Q(F, \lambda) \gg \lambda \int_0^{\lambda^{-1}} \hat{F}(t) \, dt
\]
and, therefore,
\[
Q(F, \lambda) \asymp \lambda \int_0^{\lambda^{-1}} \hat{F}(t) \, dt
\]
(see [2], Lemma 1.5 of Chapter II). The use of relation (46) allows us to simplify the arguments of [10, 23, 26] which were applied to the Littlewood–Offord problem (see also [4, 5]).

Proof of Theorem 1. Let \( r \) be a fixed number satisfying \( 1 < r \leq \sqrt{2} \). Represent the distribution \( G = \mathcal{L}(\tilde{X}) \) as a mixture
\[
G = qE + \sum_{j=0}^{\infty} p_j G_j,
\]
where \( q = \mathbf{P}(\tilde{X} = 0) \),
\[
p_j = \mathbf{P}(\tilde{X} \in A_j), \quad j = 0, 1, 2, \ldots,
\]
$A_0 = \{x : |x| > 1\}, \ A_j = \{x : r^{-j} < |x| \leq r^{-j+1}\}, \ E$ is probability measure concentrated in zero, $G_j$ are probability measures defined for $p_j > 0$ by the formula

$$G_j \{X\} = G\{X \cap A_j\}/p_j,$$

for any Borel set $X$. In fact, $G_j$ is the conditional distribution of $\tilde{X}$ provided that $\tilde{X} \in A_j$. If $p_j = 0$, then we can take as $G_j$ arbitrary measures.

For $z \in \mathbb{R}$, $\gamma > 0$, introduce the distribution $H_{z,\gamma}$, with the characteristic function

$$\hat{H}_{z,\gamma}(t) = \exp\left(-\frac{\gamma}{2} \sum_{k=1}^{n} \left(1 - \cos(2a_kzt)\right)\right). \quad (47)$$

It is clear that $H_{z,\gamma}$ is a symmetric infinitely divisible distribution. Therefore, its characteristic function is positive for all $t \in \mathbb{R}$.

For the characteristic function $\hat{F}(t) = \mathbb{E}\exp(itX)$, we have

$$|\hat{F}(t)|^2 = \mathbb{E}\exp(it\tilde{X}) = \mathbb{E}\cos(t\tilde{X}),$$

where $\tilde{X} = X_1 - X_2$ is the corresponding symmetrized random variable. Hence,

$$|\hat{F}(t)| \leq \exp\left(-\frac{1}{2} (1 - |\hat{F}(t)|^2)\right) = \exp\left(-\frac{1}{2} \mathbb{E}\left(1 - \cos(t\tilde{X})\right)\right). \quad (48)$$

According to (44) and (48), we have

$$Q(F_a, 1/D) = Q(F_{2a}, 2/D) \leq 2Q(F_{2a}, 1/D)$$

$$\ll \frac{1}{D} \int_{0}^{D} |\hat{F}_{2a}(t)| \ dt$$

$$\ll \frac{1}{D} \int_{0}^{D} \exp\left(-\frac{1}{2} \sum_{k=1}^{n} \mathbb{E}\left(1 - \cos(2a_k t\tilde{X})\right)\right) \ dt = I. \quad (49)$$

It is evident that

$$\sum_{k=1}^{n} \mathbb{E}(1 - \cos(2a_k t\tilde{X})) = \sum_{k=1}^{n} \int_{-\infty}^{\infty} (1 - \cos(2a_k tx)) G\{dx\}$$

$$= \sum_{k=1}^{n} \sum_{j=0}^{\infty} \int_{-\infty}^{\infty} (1 - \cos(2a_k tx)) p_j G_j\{dx\}$$

$$= \sum_{j=0}^{\infty} \sum_{k=1}^{n} \int_{-\infty}^{\infty} (1 - \cos(2a_k tx)) p_j G_j\{dx\}. $$
We denote $\beta_j = r^{-2j}p_j$, $\beta = \sum_{j=0}^\infty \beta_j$, $\mu_j = \beta_j/\beta$, $j = 0, 1, 2, \ldots$. It is clear that $\sum_{j=0}^\infty \mu_j = 1$ and $p_j/\mu_j = r^{2j}\beta$ (for $p_j > 0$).

Let us estimate the quantity $\beta$:

$$
\beta = \sum_{j=0}^\infty \beta_j = \sum_{j=0}^\infty r^{-2j}p_j = P\{|\widetilde{X}| > 1\} + \sum_{j=1}^\infty r^{-2j}P\{r^{-j} < |\widetilde{X}| \leq r^{-j+1}\}
$$

$$
\geq \int G\{dx\} + \sum_{j=1}^\infty \int_{r^{-j}<|x|\leq r^{-j+1}} \frac{x^2}{r^2} G\{dx\}
$$

$$
\geq \frac{1}{r^2} \int_{|x|>1} G\{dx\} + \frac{1}{r^2} \int_{|x|\leq1} x^2 G\{dx\} = \frac{1}{r^2} M(1).
$$

Since $1 < r \leq \sqrt{2}$, this implies

$$
\beta \geq \frac{1}{2} M(1). \quad (50)
$$

Inequality (50) and condition $L^2 \geq 1/M(1)$ give the bound

$$
L^2 \beta \geq \frac{1}{2}. \quad (51)
$$

We now proceed similarly to the proof of a result of Esséen [9] (see [20], Lemma 4 of Chapter II). Using the Hölder inequality, it is easy to see that

$$
I \leq \prod_{j=0}^\infty I_j^{\mu_j}, \quad (52)
$$

where

$$
I_j = \frac{1}{D} \int_0^D \exp \left( - \frac{p_j}{2 \mu_j} \sum_{k=1}^n \int_{-\infty}^\infty (1 - \cos(2a_ktx)) G_j\{dx\} \right) dt
$$

$$
= \frac{1}{D} \int_0^D \exp \left( - \frac{1}{2} r^{2j} \beta \sum_{k=1}^n \int_{A_j} (1 - \cos(2a_ktx)) G_j\{dx\} \right) dt
$$

if $p_j > 0$, and $I_j = 1$ if $p_j = 0$. 
Applying Jensen's inequality to the exponential in the integral (see \[20\], p. 49), we obtain

\[
I_j \leq \frac{1}{D} \int_0^D \int_{A_j} \exp \left( -\frac{1}{2} r^{2j} \beta \sum_{k=1}^n (1 - \cos(2a_ktx)) \right) G_j \{dx\} dt
\]

\[
= \frac{1}{D} \int_0^D \int_{A_j} \exp \left( -\frac{1}{2} r^{2j} \beta \sum_{k=1}^n (1 - \cos(2a_ktx)) \right) dt G_j \{dx\}
\]

\[
\leq \sup_{z \in A_j} \frac{1}{D} \int_0^D \hat{H}_{z,1}^{r^{2j} \beta} (t) dt.
\]

(53)

Let us estimate the characteristic function \(\hat{H}_{\pi,1}(t)\) for \(|t| \leq D\). We can proceed in the same way as the authors of \[10, 23, 26\]. It is evident that

\[1 - \cos x \geq \frac{2x^2}{\pi^2}, \quad \text{for } |x| \leq \pi.\]

For an arbitrary \(x\), this implies that

\[1 - \cos x \geq 2 \pi^{-2} \min_{m \in \mathbb{Z}} |x - 2\pi m|^2.\]

Substituting this inequality into (47), we obtain

\[
\hat{H}_{\pi,1}(t) \leq \exp \left( -\frac{1}{\pi^2} \sum_{k=1}^n \min_{m_k \in \mathbb{Z}} \left| 2\pi t a_k - 2\pi m_k \right|^2 \right)
\]

\[
= \exp \left( - 4 \sum_{k=1}^n \min_{m_k \in \mathbb{Z}} |t a_k - m_k|^2 \right)
\]

\[
= \exp \left( - 4 \left( \text{dist}(ta, \mathbb{Z}^n) \right)^2 \right).
\]

(54)

Using (8), we see that, for \(|t| \leq 1/2 \|a\|_\infty\), inequality (54) turns into

\[
\hat{H}_{\pi,1}(t) \leq \exp(-4 t^2).
\]

(55)

Now we can use relations (9), (54) and (55) to estimate the integrals \(I_j\). First we consider the case \(j = 1, 2, \ldots\). Note that the characteristic functions \(\hat{H}_{z,\gamma}(t)\) satisfy the equalities

\[
\hat{H}_{z,\gamma}(t) = \hat{H}_{y,\gamma}(zt/y) \quad \text{and} \quad \hat{H}_{z,\gamma}(t) = \hat{H}_{z,1}^\gamma(t).
\]

(56)

The first equality (56) implies that

\[
\text{if } H_{z,\gamma} = \mathcal{L}(\xi), \quad \text{then } H_{y,\gamma} = \mathcal{L}(y \xi/z).
\]

(57)

For \(z \in A_j\) we have \(r^{-j} < |z| \leq r^{-j+1} < \pi\). Hence, for \(|t| \leq D\), we have \(|zt/\pi| < D\). Therefore, using the properties (56) with \(y = \pi\) and aforementioned estimates (9), (54)
and (55), we obtain, for \( z \in A_j \) and for \( z = \pi \),

\[
\hat{H}_{z,1}(t) \leq \exp \left( -4 f_2^2(zt/\pi) \right)
= \begin{cases} 
\exp \left( -(zt/\pi)^2/9 \right), & \text{for } 0 < t \leq cL\pi/z, \\
\exp \left( -4 L^2 \log(zt/L\pi) \right), & \text{for } t > cL\pi/z.
\end{cases}
\]

Hence,

\[
\sup_{z \in A_j} \int_0^D \hat{H}_{z,1}^{2\beta}(t) \, dt \leq \int_0^D \exp \left( -t^2\beta/9\pi^2 \right) \, dt + \int_{r^{j-1}L\pi e}^\infty \left( \frac{r^j L\pi}{t} \right)^{4r^j \beta L^2} \, dt \ll \frac{1}{\sqrt{\beta}}.
\]

In the last inequality we used inequality (51).

Consider now the case \( j = 0 \). The relation (57) yields, for \( z > 0, \gamma > 0, \)

\[
Q(H_{z,\gamma}, 1/D) = Q(H_{1,\gamma}, 1/Dz).
\]

Thus, according to (2), (46), (56) and (59), we obtain

\[
\sup_{z \in A_0} \frac{1}{D} \int_0^D \hat{H}_{z,1}^{\beta}(t) \, dt = \sup_{z>1} \frac{1}{D} \int_0^D \hat{H}_{z,\beta}(t) \, dt \asymp \sup_{z>1} Q(H_{z,\beta}, 1/D) \leq Q(H_{1,\beta}, 1/D) \leq Q(H_{\pi,\beta}, 1/D) \asymp \frac{1}{D} \int_0^D \hat{H}_{\pi,\beta}(t) \, dt = \frac{1}{D} \int_0^D \hat{H}_{\pi,1}^{\beta}(t) \, dt.
\]

Using the bounds (9), (54) and (55) for the characteristic function \( \hat{H}_{\pi,1}(t) \) and taking into account inequality (51), we have:

\[
\int_0^D \hat{H}_{\pi,1}^{\beta}(t) \, dt \leq \int_0^D \exp(-t^2\beta/9) \, dt + \int_{L e}^\infty \left( \frac{L}{t} \right)^{4\beta L^2} \, dt \ll \frac{1}{\sqrt{\beta}}.
\]

According to (53), (58), (60) and (61), we obtained the same estimate

\[
I_j \ll \frac{1}{D \sqrt{\beta}}
\]

for all integrals \( I_j \) with \( p_j \neq 0 \). In view of \( \sum_{j=0}^{\infty} \mu_j = 1 \), from (52) and (62) it follows that

\[
I \leq \prod_{j=0}^{\infty} I_j^{\mu_j} \ll \frac{1}{D \sqrt{\beta}}
\]
Using (49), (50) and (63), we complete the proof. □

Now we will deduce Corollary 2 from Theorem 1.

Proof of Corollary 2. We denote $b = a/\|a\| \in \mathbb{R}^n$. Then the equality $Q(F_a, \lambda) = Q(F_b, \lambda/\|a\|)$, for all $\lambda \geq 0$, holds. The vector $b$ satisfies the conditions of Theorem 1 which hold for the vector $a$ when replacing $D$ by $D\|a\|$. Indeed, $\|ub - m\| \geq f_L(u)$ for $u \in \left[\frac{1}{2\|b\|_\infty}, D\|a\|\right]$ and for all $m \in \mathbb{Z}^n$. This follows from condition (9) of Theorem 1, if we denote $u = t\|a\|$. It remains to apply Theorem 1 to the vector $b$. □

Acknowledgements. The first and the third authors are supported by grants RFBR 10-01-00242 and 13-01-00256. The second and the third authors are supported by the SFB 701 in Bielefeld. The third author is supported by grant RFBR 11-01-12104 and by the Program of Fundamental Researches of Russian Academy of Sciences “Modern Problems of Fundamental Mathematics”.

REFERENCES

[1] Arak T.V. On the convergence rate in Kolmogorov’s uniform limit theorem. I. Theory Probab. Appl., 1981, v. 26, 225–245.
[2] Arak T.V., Zaitsev A.Yu. Uniform limit theorems for sums of independent random variables. Proc. Steklov Inst. Math., 1986, v. 174, 1–216.
[3] Bretagnolle J. Sur l’inégalité de concentration de Doeblin–Lévy, Rogozin–Kesten. In: Parametric and semiparametric models with applications to reliability, survival analysis, and quality of life, Stat. Ind. Technol., Birkhäuser Boston, Boston, MA, 2004, pp. 533–551.
[4] Eliseeva Yu.S. Multivariate estimates for the concentration functions of weighted sums of independent identically distributed random variables. Zap. Nauchn. Sem. POMI 2013, v. 412, 121–137.
[5] Eliseeva Yu.S., Zaitsev A.Yu. Estimates for the concentration functions of weighted sums of independent random variables. Theory Probab. Appl., 2012, v. 57, 767–777.
[6] Eliseeva Yu.S., Götzte F., Zaitsev A.Yu. Estimates for the concentration functions in the Littlewood–Offord problem. arXiv:1203.6763, 2012.
[7] Erdős P. On a lemma of Littlewood and Offord. Bull. Amer. Math. Soc., 1945, v. 51, 898–902.
[8] Esséen C.-G., On the Kolmogorov–Rogozin inequality for the concentration function. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete., 1966, v. 5, 210–216.
[9] Esséen C.-G., On the concentration function of a sum of independent random variables. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete., 1968, v. 9, 290–308.
[10] Friedland O., Sodin S. Bounds on the concentration function in terms of Diophantine approximation. C. R. Math. Acad. Sci. Paris, 2007, v. 345, 513–518.
[11] Götzte F., Zaitsev A.Yu. Estimates for the rapid decay of concentration functions of $n$-fold convolutions. J. Theoret. Probab., 1998, v. 11, no. 3, 715–731.
[12] Götzte F., Zaitsev A.Yu. A multiplicative inequality for concentration functions of $n$-fold convolutions. in: High dimensional probability, II (Seattle, WA, 1999), Progr. Probab., v. 47, Birkhäuser Boston, Boston, MA, 2000, pp. 39–47.
[13] Halász G. Estimates for the concentration function of combinatorial number theory and probability, Periodica Mathematica Hungarica, 1977, v. 8, 197–211.
[14] Hengartner W., Theodorescu R. Concentration function. Academic Press, New York, 1973.
[15] Kesten H., A sharper form of the Doeblin–Levy–Kolmogorov–Rogozin inequality for concentration functions. Math. Scand., 1969, v. 25, 133–144.

[16] Littlewood J.E., Offord A.C., On the number of real roots of a random algebraic equation. Rec. Math. [Mat. Sbornik] N.S., 1943, v. 12, 277–286.

[17] Miroshnikov A.L., Rogozin B.A., Inequalities for the concentration functions. Theory Probab. Appl., 1980, v. 25, 178–183.

[18] Nagaev S.V., Hodžhabagyan S.S. On the estimate for the concentration function of sums of independent random variables. Theory Probab. Appl., 1996, v. 41, 655–665.

[19] Nguyen Hoi, Vu Van. Optimal inverse Littlewood–Offord theorems. Adv. Math., 2011, v. 226, 5298–5319.

[20] Petrov V.V. Sums of independent random variables. Nauka, Moscow, 1972.

[21] Rogozin B.A. On the increase of dispersion of sums of independent random variables. Theory Probab. Appl., 1961, v. 6, 106–108.

[22] Rudelson M., Vershynin R. The Littlewood–Offord problem and invertibility of random matrices. Adv. Math., 2008, v. 218, 600–633.

[23] Rudelson M., Vershynin R. Smallest singular value of a random rectangular matrix. Comm. Pure Appl. Math., 2009, v. 62, 1707–1739.

[24] Tao T., Vu Van. Inverse Littlewood–Offord theorems and the condition number of random discrete matrices. Ann. of Math. (2), 2009, v. 169, 595–632.

[25] Tao T., Vu Van. From the Littlewood–Offord problem to the circular law: universality of the spectral distribution of random matrices. Bull. Amer. Math. Soc. (N.S.), 2009, v. 46, 377–396.

[26] Vershynin R. Invertibility of symmetric random matrices. [arXiv:1102.0300] (2011). To appear in Random Structures and Algorithms.

[27] Zaitsev A.Yu. Using of the concentration function for the estimation of the uniform distance. Zap. Nauchn. Sem. LOMI, 1982, v. 119, 93–107.

[28] Zaitsev A.Yu. On the decrease rate of the concentration function of convolutions of probability distributions. Vestn. St-Peterbg. Univ., Ser. I, Mat. Mekh. Astron., 2011, no. 2, 29–33.

E-mail address: pochta106@yandex.ru

ST. PETERSBURG STATE UNIVERSITY

E-mail address: goetze@math.uni-bielefeld.de

FAKULTÄT FÜR MATHEMATIK,
UNIVERSITÄT BIELEFELD, POSTFACH 100131,
D-33501 BIELEFELD, GERMANY

E-mail address: zaitsev@pdmi.ras.ru

ST. PETERSBURG DEPARTMENT OF STEKLOV MATHEMATICAL INSTITUTE
FONTKA 27, ST. PETERSBURG 191023, RUSSIA
AND ST. PETERSBURG STATE UNIVERSITY