BOUNDS FOR THE RATE OF STRONG APPROXIMATION IN THE MULTIDIMENSIONAL INVARIANCE PRINCIPLE

F. GÖTZE[1] AND A. YU. ZAITSEV[1,2]

University of Bielefeld[1]
St. Petersburg Department of Steklov Mathematical Institute[2]

May 2007

ABSTRACT. The aim of this paper is to derive consequences of the result of Zaitsev (2000, 2001). We establish bounds for the rate of strong Gaussian approximation of sums of independent $\mathbf{R}^d$-valued random vectors ξ_j having finite moments $\mathbf{E}\,\|\xi_j\|^\gamma$, $\gamma \geqslant 2$. A multidimensional version of the results of Sakhanenko (1985) is obtained.

1. Introduction

The aim of this paper is to derive consequences of a result of Zaitsev (2000–2001) providing optimal bounds for the strong Gaussian approximation of sums of independent $\mathbf{R}^d$-valued random vectors with finite exponential moments (see Theorem 3 below). Theorem 3 may be considered as a multidimensional generalization of well-known results of Komlós, Major and Tusnády (KMT) (1975–76) and Sakhanenko (1984), who generalized and essentially sharpened KMT results in the case of non-identically distributed $\mathbf{R}^1$-valued random variables. Then, Sakhanenko (1985) used the main result of Sakhanenko (1984) to obtain optimal bounds for the strong Gaussian approximation of sums of independent random variables ξ_j having finite moments of the form $\mathbf{E}\,H\,(|\xi_j|)$, where $H\,(x)$ is a monotone function growing not

---

1991 *Mathematics Subject Classification.* Primary 62F17; secondary 60F15.

Key words and phrases. multidimensional invariance principle, strong approximation, sums of independent random vectors.

[1] Research supported by the SFB 701, University of Bielefeld

[2] Research supported by by Russian Foundation of Basic Research (RFBR) Grant 05-01-00911, by RFBR-DFG Grant 04-01–04000, by the Grant of leading scientific scools NSh 4222.2006.1 and by a program of fundamental researches of the RAS "Modern problems of fundamental mathematics".

Typeset by $\mathcal{A}_{\mathcal{M}}\mathcal{S}$-TEX

slower than x^2 and not faster than e^{cx}. The most fine and complete results were obtained in the case $H\,(x) = x^\gamma$, $\gamma \geqslant 2$. In this paper, basing on the results of Zaitsev (2000–2001), we provide a new attempt to obtain a multidimentional generalization of the results of Sakhanenko (1985), restricting ourselves on the case $H(x) = x^\gamma$. The first one was realized in Zaitsev (2006) (see Theorem 2 below), but the corresponding inequality was established not for all possible values of the argument z.

The problem of strong approximation of sums of independent random vectors is more delicate than that of estimating the closeness of distributions. It is required to construct on a probability space a sequence of independent random vectors $X_1, \ldots, X_n$ (with given distributions) and a corresponding sequence of independent Gaussian random vectors $Y_1, \ldots, Y_n$ so that the quantity

$$\Delta(X,Y) = \max_{1\leqslant k\leqslant n} \Big\| \sum_{i=1}^{k} X_i - \sum_{i=1}^{k} Y_i \Big\| \tag{1.1}$$

would be so small as possible with sufficiently large probability. The estimation of the rate of strong approximation in the invariance principle may be reduced to this problem.

The rate of strong approximation in the one-dimensional invariance principle was studied by many authors (see, for example, Prokhorov (1956), Skorokhod (1961), Borovkov (1973), Csörgő and Révész (1975), KMT (1975–76), Borovkov and Sakhanenko (1980) and the bibliography in the monograph of Csörgő and Révész (1981) and in the papers of Borovkov and Sakhanenko (1981), Csörgő and Hall (1984), Major (1978) and Shao (1995)). One should mention the recent results of Sakhanenko (2006), who sharpened the inequalities of Sakhanenko (1985) expressing the bounds in terms of truncated moments.

Multidimensional estimates in the strong invariance principle can be found in Gorodetskii (1975), Berkes and Philipp (1979), Philipp (1979), Berger (1982), Einmahl (1987a,b), Sakhanenko (2000). The most of these papers contain the results for sufficiently small values of $\gamma \geqslant 2$ or bounds for the Prokhorov distance and for similar quantities. The Prokhorov distance estimation (coupled with the well-known Strassen–Dudley theorem) allows to get the statements about the strong approximation for concrete values of the argument z (see (1.3), (1.7) and (1.16)–(1.18)). This means that the probability space involved in the results depends on this z.

Einmahl (1989) obtained multidimensional results about the optimal rate of the strong Gaussian approximation of infinite sequences of sums of i.i.d. random vectors with $\mathbf{E}\,H\,(\|\xi_1\|) < \infty$. In particular, he established the optimal rate in the case $H\,(x) = x^\gamma$, $\gamma \geqslant 2$. Similarly to KMT (1975–76) and Sakhanenko (1985), the proof in Einmahl (1989) was based on the result for the vectors with finite exponential moments. The latter contained an extra logarithmic factor in the corresponding inequalities. Zaitsev (1998, 2000–2001) removed this unnecessary logarithmic factor from the result of Einmahl (1989) and obtained multidimensional analogues of the KMT and Sakhanenko results for the strong Gaussian approximation of sums of independent $\mathbf{R}^d$-valued random vectors with finite exponential moments.

Below we need some notation. The distribution of a random vector ξ will be denoted by $\mathcal{L}(\xi)$. The corresponding covariance operator will be denoted by $\operatorname{cov}\xi$ or $\operatorname{cov}V$, if $V=\mathcal{L}(\xi)$. We denote by c universal constants which might be different in different places of the text. We write $\log^* b=\max\{1,\log b\}$, for $b>0$.

The aim of the present paper is to obtain multidimensional analogues of the following result of Sakhanenko.

Theorem 1 (Sakhanenko (1985)). *Let* $\xi_1,\dots,\xi_n$ *be independent random variables with* $\mathbf{E}\,\xi_j=0$, $j=1,\dots,n$. *Let* $\gamma\geqslant 2$ *and*

$$L_\gamma=\sum_{j=1}^n \mathbf{E}\,|\xi_j|^\gamma<\infty.$$

Then one can construct on a probability space a sequence of independent random variables $X_1,\dots,X_n$ *and a corresponding sequence of independent Gaussian random variables* $Y_1,\dots,Y_n$ *so that* $\mathcal{L}(X_j)=\mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j=0$, $\operatorname{Var}Y_j=\operatorname{Var}X_j$, $j=1,\dots,n$,

$$\mathbf{E}\,\big(\Delta(X,Y)\big)^\gamma\leqslant c\,\gamma^{2\gamma}\,L_\gamma, \tag{1.2}$$

and, hence, for all $z>0$,

$$\mathbf{P}\,\big(\Delta(X,Y)\geqslant z\big)\leqslant c\,\gamma^{2\gamma}\,L_\gamma\,z^{-\gamma}, \tag{1.3}$$

where c *is a universal constant. Moreover, there exists an absolute positive constant* C *such that*

$$\mathbf{P}\,\big(\Delta(X,Y)\geqslant C\,z\big)\leqslant c\,\gamma^{2\gamma b}\,b^{\gamma b}\,\big(L_\gamma\,z^{-\gamma}\big)^b+2\sum_{j=1}^n\mathbf{P}\,\big(\|\xi_j\|\geqslant z/b\big), \tag{1.4}$$

for any $b\geqslant 1$ *and* $z\geqslant 0$.

It should be mentioned that, in Sakhanenko (1985), one can find more general results. In particular, there was considered the case, where the random variable ξ_j have finite moments of the form $\mathbf{E}\,H\,(|\xi_j|)$, where $H\,(x)$ is a monotone function growing not slower than x^2 and not faster than e^{cx}. The results of Sakhanenko (1985) are very precise and general. They imply a lot of other approximation results (see Shao (1995)). In particular, in Sakhanenko (1985), it is observed that the inequality (1.2) implies the well-known Rosenthal (1970, 1972) inequality (see Lemma 3 below).

After the natural normalization, we see that (1.2) is equivalent to

$$\mathbf{E}\left(\Delta(X,Y)/\sigma\right)^{\gamma} \leqslant c\gamma^{2\gamma}\, L_{\gamma}/\sigma^{\gamma},$$

where $\sigma^2 = \operatorname{Var}\big(\sum_{j=1}^{n}\xi_j\big)$. It is clear that L_γ/σ^γ, $2<\gamma\leqslant 3$, is the well-known Lyapunov fraction involved in the Lyapunov and Esséen bounds for the Kolmogorov distance in the CLT.

The following Theorem 2 was proved in a recent publication of Zaitsev (2006).

Theorem 2 (Zaitsev (2006))**.** *Suppose that* $\alpha>0$ *and* $\xi_1,\dots,\xi_n$ *are independent* $\mathbf{R}^d$*-valued random vectors with* $\mathbf{E}\,\xi_j=0$, $j=1,\dots,n$. *Let* $\gamma\geqslant 2$ *and*

$$L_{\gamma}=\sum_{j=1}^{n}\mathbf{E}\,\|\xi_j\|^{\gamma}<\infty. \tag{1.5}$$

Assume that there exist a positive integer s *and a strictly increasing sequence of non-negative integers* $m_0=0$, $m_1,\dots,m_s=n$ *satisfying the following conditions. Let*

$$\zeta_k=\xi_{m_{k-1}+1}+\cdots+\xi_{m_k},\quad \operatorname{cov}\zeta_k=\mathbb{B}_k,\qquad k=1,\dots,s,$$

and assume that, for all $v\in\mathbf{R}^d$ *and* $k=1,\dots,s$,

$$r^2L_{\gamma}^{2/\gamma}\,\|v\|^2\leqslant\big\langle\mathbb{B}_kv,v\big\rangle\leqslant C\,r^2L_{\gamma}^{2/\gamma}\,\|v\|^2\,, \tag{1.6}$$

with some constant $r\geqslant 2e$ *and* $C\geqslant 1$. *Then one can construct on a probability space a sequence of independent random vectors* $X_1,\dots,X_n$ *and a corresponding sequence of independent Gaussian random vectors* $Y_1,\dots,Y_n$ *so that* $\mathcal{L}(X_j)=\mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j=0$, $\operatorname{cov}Y_j=\operatorname{cov}X_j$, $j=1,\dots,n$, *and, for* $z>a_1\,d^{15/2+\alpha}\,\log^* d\cdot r\,L_\gamma^{1/\gamma}\,\log^* s$,

$$\mathbf{P}\left(\Delta(X,Y)\geqslant 5\,z\right)\leqslant 2\,L_\gamma\,z^{-\gamma}+\exp\Big(-\frac{a_2\,z}{r\,L_\gamma^{1/\gamma}d^{9/2}\log^* d}\Big), \tag{1.7}$$

where a_1, a_2 *are positive quantities depending only on* α *and* C.

Theorem 2 provides the multidimensional analogue of a weakened version of the inequality (1.3). Roughly speaking, in the inequality (1.3) of Theorem 1 the statement is the same as in Theorem 2 but without the exponential term in (1.7), without the condition (1.6) and without any restrictions on $z>0$. If $\operatorname{Var}\big(\sum_{j=1}^{n}\xi_j\big)\geqslant 32\,L_\gamma^{2/\gamma}$, then the condition (1.6) is always satisfied in the one-dimensional situation since $\operatorname{Var}\xi_j\leqslant\big(\mathbf{E}\,|\xi_j|^\gamma\big)^{2/\gamma}\leqslant L_\gamma^{2/\gamma}$, $j=1,\dots,n$, and we can find $m_1,\dots,m_s$ so that (1.6) is valid with $e^{-1}r=C=2$. Note that, for $z\leqslant L_\gamma^{1/\gamma}$, the inequality (1.7) is satisfied trivially. It is evident that, in the case $d=1$, $e^{-1}r=C=2$, the exponential term may be easily estimated by $c\,\gamma^\gamma L_\gamma\,z^{-\gamma}$, for $z > L_\gamma^{1/\gamma}$, with an absolute constant c. In the degenerate case, where $\operatorname{Var}\bigl(\sum_{j=1}^n \xi_j\bigr) \leqslant 32\,L_\gamma^{2/\gamma}$, the required bounds are valid for any constructing the random vectors $\bigl\{X_j\bigr\}$ and $\bigl\{Y_j\bigr\}$ on the same probability space (see Proposition 1 below).

Thus, the only difference with Sakhanenko's inequality (1.3) is in the logarithmic factor $\log^* s$ in the restriction on z. It is clear that $\log^* s$ behaves as $\log^* n$ in the case of i.i.d. $\xi_1,\dots,\xi_n$ and as the logarithm of variances of coordinates of sums $\xi_1+\dots+\xi_n$ in the general case. Therefore, it is natural to try to establish a multidimensional analogue of the inequality (1.3) which is valid without any restrictions on z as well as those of the inequalities (1.2) and (1.4). Under some additional conditions, this is realized in Theorem 4 below, which provides a partial solution of the problem.

Note that one can find in Zaitsev (2006) the corresponding versions of the inequality (1.7) for the vectors with finite moments of the form $\mathbf{E}\,H\left(\|\xi_j\|\right)$, where $H\left(x\right)$ is a monotone function growing not slower than x^2 and not faster than e^{cx}. Unfortunately, in these results there are similar logarithmic factors $\log^* s$ in the restrictions on z. Such functions were also considered in Sakhanenko (1985).

The main tool for the proof of Theorems 2 and 4 is the following Theorem 3. It was proved by Zaitsev (2000–2001). In this theorem, the most important case, where the summands have finite exponential moments, was considered. The main result of Sakhanenko (1984) was successively generalized to the multidimensional case.

Let $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$, $d \in \mathbf{N}$, denote classes of d-dimensional distributions, introduced in Zaitsev (1986), see as well Zaitsev (1998, 2000–2001, 2002). The class $\mathcal{A}_d(\tau)$ (with a fixed $\tau \geqslant 0$) consists of d-dimensional distributions V for which the function

$$\varphi(z) = \varphi(V,z) = \log \int_{\mathbf{R}^d} e^{\langle z,x\rangle} V\{dx\} \qquad (\varphi(0)=0) \tag{1.8}$$

is defined and analytic for $\|z\|\,\tau < 1$, $z \in \mathbf{C}^d$, and

$$\bigl|d_u d_v^2\,\varphi(z)\bigr| \leqslant \|u\|\tau\,\bigl\langle \mathbb{D}\,v, v\bigr\rangle, \tag{1.9}$$

for all $u, v \in \mathbf{R}^d$ and $\|z\|\,\tau < 1$, where $\mathbb{D} = \operatorname{cov} V$, and $d_u\varphi$ is the derivative of the function φ in direction u. Below we shall use simplest properties of the classes $\mathcal{A}_d(\tau)$ which are listed in Zaitsev (2002).

Theorem 3 (Zaitsev (2000–2001)). *Suppose that* $\alpha > 0$, $\tau \geqslant 1$, *and* $\xi_1,\dots,\xi_n$ *are independent* $\mathbf{R}^d$*-valued random vectors with* $\mathbf{E}\,\xi_j = 0$, $j = 1,\dots,n$. *Assume that there exist a positive integer* s *and a strictly increasing sequence of non-negative integers* $m_0 = 0$, $m_1,\dots,m_s = n$ *satisfying the following conditions. Let*

$$\zeta_k = \xi_{m_{k-1}+1} + \dots + \xi_{m_k}, \quad \operatorname{cov}\zeta_k = \mathbb{B}_k, \quad \mathcal{L}(\zeta_k) \in \mathcal{A}_d(\tau), \quad k = 1,\dots,s,$$

and suppose that, for all $v \in \mathbf{R}^d$,

$$C_1 \|v\|^2 \leqslant \langle \mathbb{B}_k v, v \rangle \leqslant C_2 \|v\|^2, \quad k = 1, \dots, s, \tag{1.10}$$

with some positive constants C_1 *and* C_2*. Then one can construct on a probability space a sequence of independent random vectors* $X_1, \dots, X_n$ *and a corresponding sequence of independent Gaussian random vectors* $Y_1, \dots, Y_n$ *so that* $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\, Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, \dots, n$, *and*

$$\mathbf{E} \exp\Big(\frac{a_3\, \Delta(X, Y)}{\tau d^{9/2} \log^* d}\Big) \leqslant \exp\big(a_4\, d^{3+\alpha} \log^*(s/\tau^2)\big),$$

where a_3, a_4 *are positive quantities depending on* α, C_1, C_2 *only.*

The i.i.d. version of Theorem 3 was proved by Zaitsev (1998), where he removed an additional logarithmic factor from a result of Einmahl (1989) and obtained a multidimensional analogue of the KMT result. Theorem 3 provides an adequate multidimensional generalization of the main result of Sakhanenko (1984). In Theorem 3, the random vectors are, generally speaking, non-identically distributed. However, they have to satisfy the condition (1.10) on the covariance operators. The condition (1.10) generates the conditions (1.6) and (1.12). A plausible conjecture is that the condition (1.10) could be weakened in the statement of Theorem 3. The same may be said about the conditions (1.6) and (1.12) in our results. At least, no similar restrictions are imposed in the paper of Zaitsev (1986) (which, however, is not related to the invariance principle). But the proof of this conjecture would require some new ideas and methods.

The main result of this paper is the following Theorem 4.

Theorem 4. *Suppose that* $\alpha > 0$, *and* $\xi_1, \dots, \xi_n$ *are independent* $\mathbf{R}^d$*-valued random vectors with* $\mathbf{E}\, \xi_j = 0$, $j = 1, \dots, n$*. Let* $\gamma \geqslant 2$ *and the quantity* L_γ *be defined by* (1.5)*. Assume that there exist a positive integer* s *and a strictly increasing sequence of non-negative integers* $m_0 = 0, m_1, \dots, m_s = n$ *satisfying the following conditions. Let*

$$\zeta_k = \xi_{m_{k-1}+1} + \dots + \xi_{m_k}, \quad \operatorname{cov} \zeta_k = \mathbb{B}_k, \qquad k = 1, \dots, s, \tag{1.11}$$

and assume that, for all $v \in \mathbf{R}^d$ *and* $k = 1, \dots, s$,

$$w^2 \|v\|^2 \leqslant \langle \mathbb{B}_k v, v \rangle \leqslant C_1\, w^2 \|v\|^2, \tag{1.12}$$

where

$$w = C_2\, L_\gamma^{1/\gamma} / \log^* s, \tag{1.13}$$

with some positive constants C_1 *and* C_2. *Suppose that the quantities*

$$\lambda_{k,\gamma} = \sum_{j=m_{k-1}+1}^{m_k} \mathbf{E}\,\|\xi_j\|^\gamma, \qquad k = 1,\dots,s, \tag{1.14}$$

satisfy, for some $0 < \varepsilon < 1$,

$$C_3\, d^{\gamma/2}\, s^\varepsilon\, (\log^* s)^{\gamma+3} \max_{1\leqslant k\leqslant s} \lambda_{k,\gamma} \leqslant L_\gamma, \tag{1.15}$$

with a positive constant C_3. *Then one can construct on a probability space a sequence of independent random vectors* $X_1,\dots,X_n$ *and a corresponding sequence of independent Gaussian random vectors* $Y_1,\dots,Y_n$ *so that* $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1,\dots,n$, *and, for any* $z > a_5\,\varepsilon^{-1}\, L_\gamma^{1/\gamma}\, d^{21/2+\alpha} \log^* d$ *and any* $b \geqslant 1$,

$$\begin{aligned}\mathbf{P}\left(\Delta(X,Y) \geqslant 9\,z\right) &\leqslant 10\, d^{1+\gamma b/2}\, (2b)^{b(\gamma-1)}\, z^{-\gamma b} \sum_{k=1}^{s} \lambda_{k,\gamma}^b + \exp\bigl(-a_6\, z^2/d\, L_\gamma^{2/\gamma}\bigr) \\ &\quad + \exp\Bigl(-\frac{a_7\,\varepsilon\, z\, \log^* s}{L_\gamma^{1/\gamma} d^{15/2} \log^* d}\Bigr) + 5\,d \sum_{j=1}^{n} \mathbf{P}\left(\|\xi_j\| \geqslant z/2b\sqrt{d}\right),\end{aligned} \tag{1.16}$$

$$\begin{aligned}\mathbf{P}\left(\Delta(X,Y) \geqslant 9\,z\right) &\leqslant d^{1+\gamma/2}\, \left(a_8\, b^4\, \varepsilon^{-1}\right)^{\gamma b} \left(L_\gamma\, z^{-\gamma}\right)^b \\ &\quad + 5\,d \sum_{j=1}^{n} \mathbf{P}\left(\|\xi_j\| \geqslant z/2b\sqrt{d}\right),\end{aligned} \tag{1.17}$$

$$\mathbf{P}\left(\Delta(X,Y) \geqslant z\right) \leqslant a_9\, d^{1+\gamma/2}\, L_\gamma\, (\varepsilon z)^{-\gamma}, \tag{1.18}$$

and

$$\mathbf{E}\left(\Delta(X,Y)\right)^\gamma \leqslant a_{10}\left(\varepsilon^{-1}\, d^{21/2+\alpha} \log^* d\right)^\gamma L_\gamma, \tag{1.19}$$

where a_5–a_{10} *are positive quantities depending on* α, γ, C_1, C_2 *and* C_3 *only.*

Note that, for $z \leqslant a_5\,\varepsilon^{-1}\, L_\gamma^{1/\gamma}\, d^{21/2+\alpha} \log^* d$, we have, in the conditions of Theorem 4,

$$\mathbf{P}\left(\Delta(X,Y) \geqslant z\right) \leqslant 1 \leqslant \left(a_5\,\varepsilon^{-1}\, d^{21/2+\alpha} \log^* d\right)^\gamma L_\gamma\, z^{-\gamma}.$$

On the other hand, the inequality (1.19) implies that

$$\mathbf{P}\left(\Delta(X,Y) \geqslant z\right) \leqslant a_{10}\left(\varepsilon^{-1}\, d^{21/2+\alpha} \log^* d\right)^\gamma L_\gamma\, z^{-\gamma}, \tag{1.20}$$

for all $z > 0$. Thus, in Theorem 4, the inequality of type (1.3) is established, in fact, for all possible values of z. It is impossible, however, to derive from (1.19) the inequality (1.18) which is sharper with respect to the dependence of constants on the dimension d. Note that in this paper we do not try to optimize the dependence of constants on the dimension. It is important here that it is a power-type dependence. It is also clear that (1.19) does not follows from (1.20) or from (1.18). Moreover, the inequality (1.17) is sometimes stronger than (1.18) (for example, if the last sum in (1.17) is equal to zero). The inequality (1.16) is sharper than (1.17) and (1.18). On the other hand, (1.17) and (1.18) are formulated simpler than (1.16).

The condition (1.15) is rather cumbersome, but it is trivially satisfied if $\xi_1,\dots,\xi_n$ are i.i.d. random vectors, and n is sufficiently large. Indeed, it is clear that in this case $m_k \asymp k\,n^{1/\gamma}/\log^* n$, $\lambda_{k,\gamma} \asymp n^{1/\gamma}/\log^* n$, $k = 1,\dots,s$, $s \asymp n^{1-1/\gamma}\log^* n$ and the condition (1.15) turns into

$$C_3\, d^{\gamma/2}\, n^{\varepsilon(1-1/\gamma)}(\log^* n)^{\gamma+2+\varepsilon} \leqslant A\, n^{1-1/\gamma}$$

with some A which does not depend on n. Moreover, in the general case

$$L_\gamma = \sum_{k=1}^{s} \lambda_{k,\gamma} \tag{1.21}$$

and (1.15) has to be satisfied if the quantities $\lambda_{k,\gamma}$ have a regular behaviour.

Choosing, for example, $C_3 = d^{-\gamma/2}$, we could remove the dependence on d in the condition (1.15). But, of course, this would imply that a_5–a_{10} would depend on d too.

Let $2 \leqslant \beta < \gamma$. Then, in the conditions of Theorem 4,

$$\bigl(\mathbf{E}\,\bigl(\Delta(X,Y)\bigr)^{\beta}\bigr)^{1/\beta} \leqslant \bigl(\mathbf{E}\,\bigl(\Delta(X,Y)\bigr)^{\gamma}\bigr)^{1/\gamma} \leqslant a_{10}^{1/\gamma}\,\varepsilon^{-1}\,d^{21/2+\alpha}\,\log^* d\cdot L_\gamma^{1/\gamma}.$$

This bound could be better than the bound which could be obtained if we apply Theorem 4 with $\gamma = \beta$. For example, in the i.i.d. case, $L_\gamma^{1/\gamma} = \bigl(n\,\mathbf{E}\,\|\xi_1\|^{\gamma}\bigr)^{1/\gamma} \asymp n^{1/\gamma}$, while $L_\beta^{1/\beta} = \bigl(n\,\mathbf{E}\,\|\xi_1\|^{\beta}\bigr)^{1/\beta} \asymp n^{1/\beta}$. In the general case, for a fixed β, one should try to minimize $L_\gamma^{1/\gamma}$, $\gamma \geqslant \beta$. Note, however, that the conditions (1.12) and (1.15) are different for different γ.

The conditions (1.6) and (1.12) imply that the minimal eigenvalue of $\operatorname{cov}\bigl(\sum_{j=1}^n \xi_j\bigr)$ is bounded from below by $r^2 L_\gamma^{2/\gamma}$ and w^2 respectively. Nevertheless, if the eigenvalues of $\operatorname{cov}\bigl(\sum_{j=1}^n \xi_j\bigr)$ are small, less than $C_4^2 L_\gamma^{2/\gamma}$, the required inequalities are valid for any constructing the random vectors $\{X_j\}$ and $\{Y_j\}$ on the same probability space.

Proposition 1. *Suppose that* $\xi_1,\dots,\xi_n$ *are independent* $\mathbf{R}^d$*-valued random vectors with* $\mathbf{E}\,\xi_j=0$, $j=1,\dots,n$. *Let* $\gamma\geqslant 2$ *and the quantity* L_γ *be defined by* (1.5). *Let* σ^2 *be the maximal eigenvalue of* $\operatorname{cov}\bigl(\sum_{j=1}^n\xi_j\bigr)$. *Assume that*

$$\sigma\leqslant C_4\,L_\gamma^{1/\gamma} \tag{1.22}$$

with some positive constant C_4. *Then, for any construction on a probability space of a sequence of independent random vectors* $X_1,\dots,X_n$ *and a corresponding sequence of independent Gaussian random vectors* $Y_1,\dots,Y_n$ *such that* $\mathcal{L}(X_j)=\mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j=0$, $\operatorname{cov}Y_j=\operatorname{cov}X_j$, $j=1,\dots,n$, *the following inequalities are valid:*

$$\begin{aligned}\mathbf{P}\bigl(\Delta(X,Y)\geqslant 4z\bigr)&\leqslant d^{1+\gamma/2}\,\bigl(a_{11}\,b\bigr)^{b(\gamma-1)}\,\bigl(L_\gamma\,z^{-\gamma}\bigr)^b\\&\quad+3\,d\sum_{j=1}^n\mathbf{P}\bigl(\|\xi_j\|\geqslant z/2b\sqrt{d}\bigr),\qquad\text{for all } z>0,\quad b\geqslant 1,\end{aligned}\tag{1.23}$$

$$\mathbf{E}\,\bigl(\Delta(X,Y)\bigr)^\gamma\leqslant a_{12}\,d^{1+\gamma/2}\,L_\gamma, \tag{1.24}$$

and, hence,

$$\mathbf{P}\bigl(\Delta(X,Y)\geqslant z\bigr)\leqslant a_{12}\,d^{1+\gamma/2}\,L_\gamma\,z^{-\gamma}, \tag{1.25}$$

where a_{11} *and* a_{12} *are positive quantities depending on* γ *and* C_4 *only.*

The proof of Proposition 1 is based, in fact, on the trivial inequality $\Delta(X,Y)\leqslant\Delta(X,0)+\Delta(0,Y)$. It remains to estimate $\Delta(X,0)$ and $\Delta(0,Y)$.

The proof of Theorem 4 is a modified version of the corresponding proofs from Sakhanenko (1985) and Zaitsev (2006). The main difference is that we truncate not the initial summands ξ_j but the sums ζ_k over blocks of summands ξ_j and then show that the distributions of truncated vectors belong to the class $\mathcal{A}_d(\tau)$ with τ which is essentially less than the level of truncation u. A similar idea was used by Einmahl (1989). The complete proof of Sakhanenko (1985) is yet not generalized due to technical difficulties occuring in the multidimensional case.

Below we shall need the following Lemmas 1–3.

Lemma 1. *Let* $X_1,\dots,X_n$ *be independent random vectors and* $S_k=X_1+\dots+X_k$, $k=1,\dots,n$. *Then*

$$\mathbf{P}\bigl\{\max_{1\leqslant k\leqslant n}\|S_k\|\geqslant 3t\bigr\}\leqslant 3\max_{1\leqslant k\leqslant n}\mathbf{P}\bigl\{\|S_k\|\geqslant t\bigr\},\qquad t\geqslant 0. \tag{1.26}$$

Lemma 1 is a version of the Ottaviani inequality, see de la Peña and Giné (1999, Proposition 1.1.2). In the form (1.26), this inequality can be found in Kwapień and Woyczynski (1992), and (with 4 instead of 3 (twice)) in Etemadi (1985).

Lemma 2. *Let* $\gamma \geqslant 2$, $b \geqslant 1$ *and* $S_n = \psi_1 + \dots + \psi_n$ *be the sum of independent random variables* $\psi_j \in \mathbf{R}^1$ *with* $\mathbf{E}\,\psi_j = 0$, $\mathbf{E}\,|\psi_j|^\gamma < \infty$, $j = 1, \dots, n$, *and let* $\operatorname{Var} S_n = \sigma^2$. *Then*

$$\mathbf{P}\left(\bigl|S_n\bigr| \geqslant x\right) \leqslant 2\,\exp\left(-a_{13}\,x^2/\sigma^2\right) + 2\,(2b)^{b(\gamma-1)}\Bigl(\sum_{j=1}^{n} \frac{\mathbf{E}\,|\psi_j|^\gamma}{x^\gamma}\Bigr)^b + \sum_{j=1}^{n} \mathbf{P}\left(\bigl|\psi_j\bigr| \geqslant x/2b\right), \tag{1.27}$$

for all $x > 0$, *where* a_{13} *is a positive quantity depending on* γ *only.*

Lemma 2 follows easily from the main result of Fuk and Nagaev (1971), see also Nagaev (1979, Corollary 1.7).

It is easy to see that, by the Chebyshev inequality, the inequality (1.27) with $b = 1$ implies that

$$\mathbf{P}\left(\bigl|S_n\bigr| \geqslant x\right) \leqslant 2\,\exp\left(-a_{13}\,x^2/\sigma^2\right) + 2^{\gamma+1} \sum_{j=1}^{n} \frac{\mathbf{E}\,|\psi_j|^\gamma}{x^\gamma}, \tag{1.28}$$

for all $x > 0$, in the conditions of Lemma 2.

Lemma 3. *Let* $\xi_1, \dots, \xi_n$ *denote independent random vectors which have mean zero and assume values in* $\mathbf{R}^d$. *Then*

$$\mathbf{E}\Bigl\|\sum_{j=1}^{n} \xi_j\Bigr\|^\gamma \leqslant a_{14}\Bigl(\sum_{j=1}^{n} \mathbf{E}\,\|\xi_j\|^\gamma + \Bigl(\sum_{j=1}^{n} \mathbf{E}\,\|\xi_j\|^2\Bigr)^{\gamma/2}\Bigr), \quad 0 < \gamma < \infty,$$

with a_{14} *depending on* γ *only.*

This Rosenthal type inequality easily follows from a result of de Acosta (1981) and the fact that the $\mathbf{R}^d$ is a type 2 space, that is, $\mathbf{E}\Bigl\|\sum\limits_{j=1}^{n} \xi_j\Bigr\|^2 \leqslant c \sum\limits_{j=1}^{n} \mathbf{E}\,\|\xi_j\|^2$.

2. Proofs

For any random vector ξ we shall denote by $\xi^{[u]}$ its truncation on the level $u > 0$:

$$\xi^{[u]} = \begin{cases} \xi, & \text{if } \|\xi\| \leqslant u, \\ 0, & \text{otherwise.} \end{cases} \tag{2.1}$$

Set

$$\xi^{(u)} = \xi - \xi^{[u]}, \quad \xi^{\{u\}} = \xi^{[u]} - \mathbf{E}\,\xi^{[u]}. \tag{2.2}$$

Assuming $\mathbf{E}\,\|\xi\|^2<\infty$ and $\mathbf{E}\,\xi=0$, we have $\xi=\xi^{(u)}+\xi^{[u]}$,

$$\mathbf{E}\,\xi^{(u)}+\mathbf{E}\,\xi^{[u]}=\mathbf{E}\,\xi=0,$$

and, hence,

$$\big\|\mathbf{E}\,\xi^{(u)}\big\|=\big\|\mathbf{E}\,\xi^{[u]}\big\|\leqslant\mathbf{E}\,\big\|\xi^{(u)}\big\|.$$

Moreover, for any $v\in\mathbf{R}^d$,

$$\big|\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle\big|=\big|\mathbf{E}\,\big\langle\,\xi^{[u]},v\big\rangle\big|\leqslant\|v\|\,\mathbf{E}\,\big\|\xi^{(u)}\big\|. \tag{2.3}$$

Furthermore,

$$\begin{aligned}\operatorname{Var}\big\langle\,\xi,v\big\rangle&=\mathbf{E}\,\big\langle\,\xi,v\big\rangle^2=\mathbf{E}\,\big\langle\,\xi^{[u]},v\big\rangle^2+\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2\\&\geqslant\operatorname{Var}\big\langle\,\xi^{[u]},v\big\rangle+\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2.\end{aligned}\tag{2.4}$$

By (2.3), (2.4) and by the Cauchy–Bunyakovskii inequality, we have

$$\begin{aligned}\operatorname{Var}\big\langle\,\xi^{[u]},v\big\rangle&=\mathbf{E}\,\big\langle\,\xi^{[u]},v\big\rangle^2-\big(\mathbf{E}\,\big\langle\,\xi^{[u]},v\big\rangle\big)^2\\&=\operatorname{Var}\big\langle\,\xi,v\big\rangle-\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2-\big(\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle\big)^2\\&\geqslant\operatorname{Var}\big\langle\,\xi,v\big\rangle-2\,\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2.\end{aligned}\tag{2.5}$$

Inequalities (2.4) and (2.5) imply

$$2\,\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2\geqslant\operatorname{Var}\big\langle\,\xi,v\big\rangle-\operatorname{Var}\big\langle\,\xi^{[u]},v\big\rangle\geqslant\mathbf{E}\,\big\langle\,\xi^{(u)},v\big\rangle^2. \tag{2.6}$$

Proof of Theorem 4. The symbols $c,c_2,c_1,\dots$ will be used for absolute positive constants. The letter c can denote different constants when we do not need to fix their numerical values. The same notation will be used for positive quantities which depend only on α, γ, C_1, C_2 and C_3 involved in the conditions of Theorem 4. This means, in fact, that we assume α, γ, C_1, C_2 and C_3 to be absolute positive constants. We shall write $A\ll B$, if there exists a c such that $A\leqslant cB$.

Let $\xi_1,\dots,\xi_n$ and $\zeta_1,\dots,\zeta_s$ be random vectors satisfying the conditions of Theorem 4. For $u>0$, consider independent Gaussian random vectors $\eta_1^{\{u\}},\dots,\eta_n^{\{u\}}$ with

$$\mathbf{E}\,\eta_k^{\{u\}}=\mathbf{E}\,\zeta_k^{\{u\}}=0,\quad\operatorname{cov}\eta_k^{\{u\}}=\operatorname{cov}\zeta_k^{\{u\}}=\operatorname{cov}\zeta_k^{[u]}\overset{\mathrm{def}}{=}\mathbb{B}_k^{\{u\}},\qquad k=1,\dots,s.$$

According to (1.11) and (2.6), for any $v\in\mathbf{R}^d$, we have

$$2\,\mathbf{E}\,\big\langle\,\zeta_k^{(u)},v\big\rangle^2\geqslant\big\langle\mathbb{B}_kv,v\big\rangle-\big\langle\mathbb{B}_k^{\{u\}}v,v\big\rangle\geqslant\mathbf{E}\,\big\langle\,\zeta_k^{(u)},v\big\rangle^2,\qquad k=1,\dots,s. \tag{2.7}$$

Thus, the operators

$$\mathbb{B}_k - \mathbb{B}_k^{\{u\}} \overset{\text{def}}{=} \mathbb{B}_k^{(u)} \tag{2.8}$$

are positive definite and may be treated as covariance operators. Therefore, we can construct the Gaussian vectors $\eta_k^{(u)}$ with

$$\mathbf{E}\,\eta_k^{(u)} = 0, \quad \operatorname{cov}\eta_k^{(u)} = \mathbb{B}_k^{(u)},$$

such that they are jointly independent and independent of the vectors $\eta_k^{\{u\}}$. Denote $\eta_k = \eta_k^{(u)} + \eta_k^{\{u\}}$. Clearly,

$$\mathbf{E}\,\eta_k = 0, \quad \operatorname{cov}\eta_k = \mathbb{B}_k.$$

Using the standard tool of strong approximation—Lemma A of Berkes and Philipp (1979), we can take ξ_j as X_j in the proof of Theorem 4, choosing below the joint distribution of $\{\zeta_k^{\{u\}}\}$ and $\{\eta_k^{\{u\}}\}$ in a special way. Moreover, we can assume that

$$\eta_k = \sum_{j=m_{k-1}+1}^{m_k} Y_j, \qquad k = 1, \dots, s, \tag{2.9}$$

where $Y_1, \dots, Y_n$ are the independent random vectors satisfying the conditions of Theorem 4.

Denoting by $\mathbb{B}^{(u)}$ the covariance operator of the sum $\sum_{k=1}^{s} \eta_k^{(u)}$, we have

$$\mathbb{B}^{(u)} = \sum_{k=1}^{s} \mathbb{B}_k^{(u)}. \tag{2.10}$$

Let the quantity $B^2(u)$ be defined by

$$B^2(u) \overset{\text{def}}{=} \sum_{k=1}^{s} \mathbf{E}\,\big\|\zeta_k^{(u)}\big\|^2. \tag{2.11}$$

By (2.7)–(2.11), for any $v \in \mathbf{R}^d$, the inequalities

$$\sum_{k=1}^{s} \mathbf{E}\,\big\langle \zeta_k^{(u)}, v\big\rangle^2 \leqslant \big\langle \mathbb{B}^{(u)} v, v\big\rangle \leqslant 2 \sum_{k=1}^{s} \mathbf{E}\,\big\langle \zeta_k^{(u)}, v\big\rangle^2 \leqslant 2\,B^2(u)\ \|v\|^2. \tag{2.12}$$

are valid. Let σ_u^2 be the maximal eigenvalue of the operator $\mathbb{B}^{(u)}$. By (2.12),

$$\sigma_u^2 = \max_{\|v\|=1} \big\langle \mathbb{B}^{(u)} v, v\big\rangle \leqslant 2\,B^2(u). \tag{2.13}$$

Let $Z, Z_1, \dots , Z_d$ be i.i.d. standard normal random variables. Using standard tools to obtain exponential inequalities, it is easy to show that there exists an absolute constant c such that, for any $x > c\sigma_u\sqrt{d}$,

$$\mathbf{P}\left(\Big\|\sum_{k=1}^{s}\eta_k^{(u)}\Big\| \geqslant x\right) \leqslant \mathbf{P}\left(\sigma_u^2\sum_{k=1}^{d}Z_j^2 \geqslant x^2\right) \leqslant \exp\left(-x^2/8\,\sigma_u^2\right). \tag{2.14}$$

It suffices to note that $\mathbf{E}\,e^{hZ^2} = (1-2h)^{-1/2}$, for $0 < h < 1/2$, and to apply the exponential Chebyshev inequality, choosing $h = 1/6$.

Replacing n by s in the definition (1.1) of $\Delta(\,\cdot\,,\,\cdot\,)$, introduce the notation

$$\Delta = \Delta(\zeta,\eta), \quad \Delta^{\{u\}} = \Delta\big(\zeta^{\{u\}},\eta^{\{u\}}\big), \quad \Delta_u = \Delta\big(\zeta - \mathbf{E}\,\zeta^{[u]},\eta^{\{u\}}\big). \tag{2.15}$$

It is easy to see that

$$\Delta \leqslant \Delta_u + \delta_u + A(u), \tag{2.16}$$

where the quantity $A(u)$ is defined by

$$A(u) \overset{\text{def}}{=} \max_{1\leqslant m\leqslant s}\Big\|\sum_{k=1}^{m}\mathbf{E}\,\zeta_k^{[u]}\Big\| = \max_{1\leqslant m\leqslant s}\Big\|\sum_{k=1}^{m}\mathbf{E}\,\zeta_i^{(u)}\Big\|, \tag{2.17}$$

and

$$\delta_u = \Delta(\eta,\eta^{\{u\}}). \tag{2.18}$$

Following Sakhanenko (1985), introduce the class $\mathcal{G}(u)$, $u \geqslant 0$, of positive functions $G(\,\cdot\,)$ such that the functions $G(x)$ and $x/\log G(x)$ are non-decreasing for $x > u$. For example, the functions $G(x) = x^\gamma$, $\gamma > 0$, and $\exp\big(\lambda x^\beta\big)$, $0 \leqslant \beta \leqslant 1$, $\lambda > 0$, belong to $\mathcal{G}(e)$ and $\mathcal{G}(0)$ respectively.

The formulation and proof of the following lemma almost coincide with the formulation and proof of Theorem 3 of Sakhanenko (1985). One can find it also in Zaitsev (2006) (with $\{\zeta_k\}$ replaced by $\{\xi_j\}$).

Lemma 4. *Let $u > 0$, $G(\,\cdot\,) \in \mathcal{G}(u)$ and*

$$F(x) \overset{\text{def}}{=} \sum_{k=1}^{s}\mathbf{P}\left(\|\zeta_k\| \geqslant x\right) \leqslant 1/G(x), \quad \textit{for all } x > u. \tag{2.19}$$

Then, for all $x_0 \geqslant 0$ and $x \geqslant z \geqslant u$,

$$\mathbf{P}\left(\Delta_u \geqslant x_0 + x\right) \leqslant \mathbf{P}\left(\Delta^{\{u\}} \geqslant x_0\right) + \mathbf{P}\left(\max_{1\leqslant k\leqslant s}\|\zeta_k\| \geqslant z\right) + e\,G^{1-x/z}(z)/G(u). \tag{2.20}$$

It is easy to see that, for $k=1,\dots,s,\ x\geqslant 0$,

$$F_k(x)\stackrel{\text{def}}{=}\mathbf{P}\left(\left\|\zeta_k\right\|\geqslant x\right)\leqslant T_k(x)\stackrel{\text{def}}{=}\max_{m_{k-1}+1\leqslant m\leqslant m_k}\mathbf{P}\Bigl(\Bigl\|\sum_{j=m_{k-1}+1}^{m}\xi_j\Bigr\|\geqslant x\Bigr). \tag{2.21}$$

It is clear that

$$F(x)=\sum_{k=1}^{s}F_k(x)\leqslant\sum_{k=1}^{s}T_k(x) \tag{2.22}$$

(see (2.19) and (2.21)).

We shall use the following Fuk–Nagaev type inequalities:

$$\begin{aligned}F_k(x)\leqslant T_k(x)\leqslant 2d\,\exp\left(-c_1x^2/dw^2\right)+2d\,(2b)^{b(\gamma-1)}\Bigl(d^{\gamma/2}\sum_{j=m_{k-1}+1}^{m_k}\frac{\mathbf{E}\,\|\xi_j\|^\gamma}{x^\gamma}\Bigr)^b&\\+d\sum_{j=m_{k-1}+1}^{m_k}\mathbf{P}\left(\left\|\xi_j\right\|\geqslant x/2b\sqrt{d}\right),\qquad k=1,\dots,s,&\end{aligned} \tag{2.23}$$

$$F_k(x)\leqslant T_k(x)\ll d\,\exp\left(-c_1x^2/dw^2\right)+d^{1+\gamma/2}\sum_{j=m_{k-1}+1}^{m_k}\frac{\mathbf{E}\,\|\xi_j\|^\gamma}{x^\gamma},\qquad k=1,\dots,s, \tag{2.24}$$

which holds for any $x>0$ and $b\geqslant 1$. They follow from (2.21) and from the corresponding one-dimensional inequalities (1.27) and (1.28). Constants depend here on γ and C_1, but recall that we treat γ and C_1 as universal constants. The inequality (2.24) is weaker than (2.23), but it is more convenient in some applications below.

It should be mentioned that the inequality (2.24) with constants which do not depend on the dimension d was proved for $\gamma=3$ by Ebralidze (1971), see Nagaev (1979, Theorem 4.3). Therefore, for $\gamma=3$, the dependence of constants on the dimension in the inequality (1.18) of Theorem 4 might be sharpened.

We shall choose

$$u=c_2\,d\,L_\gamma^{1/\gamma}, \tag{2.25}$$

where the constant c_2 will be so large as it will be required below.

Using (1.14) and (2.23)–(2.25), we get that, for $x\geqslant u$, $b\geqslant 1$,

$$\exp\left(-c_1x^2/2dw^2\right)\ll 1/2ds, \tag{2.26}$$

$$\begin{aligned}T_k(x)\leqslant\frac{1}{s}\,\exp\left(-c_1x^2/2dw^2\right)+2d\,(2b)^{b(\gamma-1)}\Bigl(d^{\gamma/2}\,\frac{\lambda_{k,\gamma}}{x^\gamma}\Bigr)^b&\\+d\sum_{j=m_{k-1}+1}^{m_k}\mathbf{P}\left(\left\|\xi_j\right\|\geqslant x/2b\sqrt{d}\right),\qquad k=1,\dots,s,&\end{aligned} \tag{2.27}$$

and

$$F_k(x) \leqslant T_k(x) \ll \frac{1}{s}\exp\bigl(-c_1 x^2/2dw^2\bigr) + d^{1+\gamma/2}\,\frac{\lambda_{k,\gamma}}{x^\gamma}\,, \qquad k=1,\dots,s, \tag{2.28}$$

if c_2 is large enough.

According to (1.13), (1.21), (2.19), (2.22), (2.25) and (2.28), we have

$$F(x) \ll \exp\bigl(-c_1 x^2/2dw^2\bigr) + d^{1+\gamma/2}\sum_{j=1}^{n}\frac{\mathbf{E}\,\|\xi_j\|^\gamma}{x^\gamma} \ll d^{1+\gamma/2}\,\frac{L_\gamma}{x^\gamma}\,, \tag{2.29}$$

for $x \geqslant u$, that is, there exists c_3 such that

$$F(x) \leqslant \frac{1}{G(x)}\,, \quad \text{where} \quad G(x) = \frac{x^\gamma}{c_3\,d^{1+\gamma/2}L_\gamma}\,, \tag{2.30}$$

for $x \geqslant u$, and

$$G(u) \geqslant e^\gamma, \tag{2.31}$$

if c_2 is large enough. We used here that $\gamma \geqslant 2$ and, hence, $1+\gamma/2 \leqslant \gamma$. In order to apply Lemma 4 we have to verify that $G(\,\cdot\,) \in \mathcal{G}(u)$. Let us show that the function $x/\log G(x)$ is non-decreasing for $x > u$. Indeed, for $x \geqslant y \geqslant u$, we have $G(y) \geqslant G(u) = e^\gamma$ and $\frac{G(x)}{x^\gamma} = \frac{G(y)}{y^\gamma}$. Denoting $t = x/y \geqslant 1$, $q = G(y)$, we get $q \geqslant e^\gamma$, $t^\gamma \leqslant e^{\gamma(t-1)} \leqslant q^{t-1}$ and $G(x) = qt^\gamma \leqslant q^t$. Hence, $\log G(x) \leqslant t\,\log q = \frac{x}{y}\log G(y)$ Thus, $G(\,\cdot\,) \in \mathcal{G}(u)$.

By (2.16) and (2.20), for all $x_0, x_1 \geqslant 0$ and $x \geqslant z \geqslant u$,

$$\begin{aligned}\mathbf{P}\bigl(\Delta \geqslant x_0+x_1+x+A(u)\bigr) &\leqslant \mathbf{P}\,\left(\Delta_u \geqslant x_0+x\right) + \mathbf{P}\bigl(\delta_u \geqslant x_1\bigr)\\ &\leqslant \mathbf{P}\bigl(\Delta^{\{u\}} \geqslant x_0\bigr) + \mathbf{P}\bigl(\max_{1\leqslant k\leqslant s}\|\zeta_k\| \geqslant z\bigr)\\ &\quad + e\,G^{1-x/z}(z)/G(u) + \mathbf{P}\bigl(\delta_u \geqslant x_1\bigr).\end{aligned} \tag{2.32}$$

Using (2.18) and the Lévy inequality (see de la Peña and Giné (1999, Theorem 1.1.1)) and taking into account that $\eta_k - \eta_k^{\{u\}} = \eta_k^{(u)}$, we obtain

$$\begin{aligned}\mathbf{P}\,\left(\delta_u \geqslant x_1\right) &= \mathbf{P}\Bigl(\max_{1\leqslant m\leqslant s}\Bigl\|\sum_{k=1}^{m}\Bigl(\eta_k - \eta_k^{\{u\}}\Bigr)\Bigr\| \geqslant x_1\Bigr)\\ &\leqslant 2\,\mathbf{P}\Bigl(\Bigl\|\sum_{k=1}^{s}\eta_k^{(u)}\Bigr\| \geqslant x_1\Bigr).\end{aligned} \tag{2.33}$$

By (2.13), (2.14) and (2.33), we have

$$\mathbf{P}\,\left(\delta_u \geqslant x_1\right) \leqslant 2\,\exp\bigl(-x_1^2/16\,B^2(u)\bigr)\,, \tag{2.34}$$

for any $x_1 > c\,B(u)\sqrt{d}$, where c is an absolute constant.

Integrating by parts and using the relations (1.14), (1.15), (1.21), (2.25), (2.26) and (2.28)–(2.31), we get rough bounds

$$\begin{aligned}\mathbf{E}\,\big\|\zeta_k^{(u)}\big\|^2 &= -\int_{\{x>u\}} x^2\,dF_k(x) = u^2\,F_k(u) + 2\int_{\{x>u\}} x\,F_k(x)\,dx\\ &\leqslant c u^2\,\Big(\frac{1}{s}\,\exp\big(-c_1\,u^2/2\,d\,w^2\big) + d^{1+\gamma/2}\,\frac{\lambda_{k,\gamma}}{u^\gamma}\Big)\\ &\quad + c\int_{\{x\geqslant u\}} x\,\Big(\frac{1}{s}\,\exp\big(-c_1\,x^2/2\,d\,w^2\big) + d^{1+\gamma/2}\,\frac{\lambda_{k,\gamma}}{x^\gamma}\Big)\,dx\\ &\leqslant \frac{w^2}{8s} + \frac{c\,d^{1+\gamma/2}\,\lambda_{k,\gamma}}{u^{\gamma-2}} \leqslant w^2/4 \leqslant u^2/4, \qquad k=1,\dots,s, \end{aligned}\tag{2.35}$$

and

$$\begin{aligned}B^2(u) &= -\int_{\{x>u\}} x^2\,dF(x) = u^2\,F(u) + 2\int_{\{x>u\}} x\,F(x)\,dx\\ &\leqslant u^2/2 + \frac{c\,d^{1+\gamma/2}\,L_\gamma}{u^{\gamma-2}} \leqslant u^2,\end{aligned}\tag{2.36}$$

if c_2 is large enough.

By (1.13), (2.1), (2.2), (2.35) and by the Chebyshev inequality,

$$\big\|\mathbf{E}\,\zeta_k^{(u)}\big\| \leqslant \mathbf{E}\,\big\|\zeta_k^{(u)}\big\|^2/u \leqslant w^2/u \leqslant w, \qquad k=1,\dots,s. \tag{2.37}$$

By (2.7) and (2.35), for any $v\in\mathbf{R}^d$,

$$\big\langle \mathbb{B}_k v, v\big\rangle - \big\langle \mathbb{B}_k^{\{u\}} v, v\big\rangle \leqslant 2\,\mathbf{E}\,\big\langle \zeta_k^{(u)}, v\big\rangle^2 \leqslant 2\,\mathbf{E}\,\big\|\zeta_k^{\{u\}}\big\|^2\,\|v\|^2 \leqslant w^2\,\|v\|^2/2. \tag{2.38}$$

Using (1.12), (2.7) and (2.38), it is not difficult to make sure that, for any $v\in\mathbf{R}^d$,

$$w^2\,\|v\|^2/2 \leqslant \big\langle \mathbb{B}_k^{\{u\}} v, v\big\rangle \leqslant C_1\,w^2\,\|v\|^2, \qquad k=1,\dots,s. \tag{2.39}$$

Arguing similarly to the proof of (2.14) and using (1.12), we obtain that there exists an absolute constant c such that, for any $x > cw\sqrt{dC_1}$, $k=1,\dots,s$,

$$\mathbf{P}\,\Big(\Big\|\sum_{j=m_{k-1}+1}^{m} Y_j\Big\| \geqslant x\Big) \leqslant \exp\big(-x^2/8C_1\,w^2\big), \quad m_{k-1}+1 \leqslant m \leqslant m_k. \tag{2.40}$$

Therefore, for any $x>u$, $k=1,\dots,s$,

$$\mathbf{P}\,\Big(\Big\|\sum_{j=m_{k-1}+1}^{m} Y_j\Big\| \geqslant x\Big) \leqslant \tfrac{1}{s}\,\exp\big(-x^2/16C_1\,w^2\big), \quad m_{k-1}+1 \leqslant m \leqslant m_k, \tag{2.41}$$

if c_2 in (2.25) is large enough.

Let

$$w_0 = c_4\, d\, w/\varepsilon, \tag{2.42}$$

where $0 < \varepsilon < 1$ and c_4 is large enough. Then, choosing $c_4 \geqslant 1$, we can ensure that $w_0 \geqslant d w \geqslant w$.

Below we shall estimate the expectation

$$\mathbf{E}\,\big\|\zeta_k^{\{u\}}\big\|^3 \exp\Big(\big\|\zeta_k^{\{u\}}\big\|/w_0\Big) = I_1 + I_2, \qquad k = 1, \dots, s, \tag{2.43}$$

where

$$I_1 = \mathbf{E}\,\big\|\zeta_k^{\{u\}}\big\|^3 \exp\Big(\big\|\zeta_k^{\{u\}}\big\|/w_0\Big)\ \mathbf{1}\big\{\big\|\zeta_k^{[u]}\big\| < w\big\} \tag{2.44}$$

and

$$I_2 = \mathbf{E}\,\big\|\zeta_k^{\{u\}}\big\|^3 \exp\Big(\big\|\zeta_k^{\{u\}}\big\|/w_0\Big)\ \mathbf{1}\big\{\big\|\zeta_k^{[u]}\big\| \geqslant w\big\}. \tag{2.45}$$

Notation $\mathbf{1}\{A\}$ is used here for the indicator function of the event A. Taking into account the relations (2.2), (2.37), (2.44) and $w_0 \geqslant w$, we obtain $\zeta_k^{\{u\}} = \zeta_k^{[u]} - \mathbf{E}\,\zeta_k^{[u]} = \zeta_k^{[u]} + \mathbf{E}\,\zeta_k^{(u)}$ and

$$I_1 \ll w^3. \tag{2.46}$$

Moreover, by (2.1), (2.37) and (2.45),

$$\begin{aligned} I_2 &\leqslant \mathbf{E}\,\big\|\zeta_k + \mathbf{E}\,\zeta_k^{(u)}\big\|^3 \exp\Big(\big\|\zeta_k + \mathbf{E}\,\zeta_k^{(u)}\big\|/w_0\Big)\ \mathbf{1}\big\{w \leqslant \|\zeta_k\| \leqslant u\big\} \\ &\ll \mathbf{E}\,\big(\|\zeta_k\|^3 + w^3\big) \exp\big(\|\zeta_k\|/w_0\big)\ \mathbf{1}\big\{w \leqslant \|\zeta_k\| \leqslant u\big\}. \end{aligned} \tag{2.47}$$

Integrating by parts and using the relations (2.21) and $w_0 \geqslant w$, we obtain

$$\begin{aligned} &\mathbf{E}\,\big(\|\zeta_k\|^3 + w^3\big) \exp\big(\|\zeta_k\|/w_0\big)\ \mathbf{1}\big\{w \leqslant \|\zeta_k\| \leqslant u\big\} \\ &\qquad = -\int\limits_{\{w \leqslant x \leqslant u\}} (x^3 + w^3)\, e^{x/w_0}\, dF_k(x) \\ &\qquad \leqslant 2 e\, w^3\, F_k(w) + \int\limits_{\{w \leqslant x \leqslant u\}} \Big(3x^2 + \frac{x^3 + w^3}{w_0}\Big)\, e^{x/w_0}\, F_k(x)\, dx. \end{aligned} \tag{2.48}$$

Furthermore, by (1.14) and (2.24), we have

$$\int\limits_{\{w \leqslant x \leqslant u\}} \Big(3x^2 + \frac{x^3 + w^3}{w_0}\Big)\, e^{x/w_0}\, F_k(x)\, dx \ll I_3 + I_4, \tag{2.49}$$

where

$$I_3 = d \int\limits_{\{w\leqslant x\leqslant u\}} \Big(3x^2 + \frac{x^3+w^3}{w_0}\Big)\, e^{x/w_0} \exp\big(-c_1 x^2/2dw^2\big)\, dx \tag{2.50}$$

and

$$I_4 = \int\limits_{\{w\leqslant x\leqslant u\}} \Big(3x^2 + \frac{x^3+w^3}{w_0}\Big)\, e^{x/w_0}\, d^{1+\gamma/2}\, \frac{\lambda_{k,\gamma}}{x^\gamma}\, dx. \tag{2.51}$$

It is easy to see that

$$I_3 \ll \int\limits_{\{w\leqslant x\}} \Big(3x^2 + \frac{x^3+w^3}{w_0}\Big) \exp\big(-c_1 x^2/4dw^2\big)\, dx \ll d^2\, w^3 \ll d\, w_0\, w^2, \tag{2.52}$$

if c_4 in (2.42) is sufficiently large. Furthermore, by (1.13), (1.15), (2.25), (2.42) and (2.51),

$$\begin{aligned} I_4 &\ll d^{1+\gamma/2}\Big(3u^2 + \frac{u^3+w^3}{w_0}\Big)\, e^{u/w_0}\, \lambda_{k,\gamma} \int\limits_{\{w\leqslant x\}} \frac{dx}{x^\gamma} \\ &\ll \frac{d^{2+\gamma/2}\lambda_{k,\gamma}\, s^{\varepsilon}\, (\log^* s)^{\gamma+3}}{L_\gamma}\, w_0\, w^2 \ll d^2\, w_0\, w^2, \end{aligned} \tag{2.53}$$

if c_4 is large enough. Collecting the bounds (2.43) and (2.46)–(2.53), we obtain

$$\mathbf{E}\, \big\|\zeta_k^{\{u\}}\big\|^3 \exp\Big(\big\|\zeta_k^{\{u\}}\big\|/w_0\Big) \ll d^2\, w_0\, w^2. \tag{2.54}$$

Using (2.54), we get, for $m = 3, 4, \dots$,

$$\begin{aligned} \mathbf{E}\, \big\|\zeta_k^{\{u\}}\big\|^m &\leqslant (m-3)!\, w_0^{m-3}\, \mathbf{E}\, \big\|\zeta_k^{\{u\}}\big\|^3 \exp\Big(\big\|\zeta_k^{\{u\}}\big\|/w_0\Big) \\ &\ll d^2\, m!\, w_0^{m-2}\, w^2. \end{aligned} \tag{2.55}$$

By (2.39), the inequality (2.55) is valid for $m = 2, 3, \dots$.

Let us prove that $\mathcal{L}\big(\zeta_k^{\{u\}}\big) \in \mathcal{A}_d(c_5\, d^2\, w_0)$, where the constant c_5 will be chosen to be so large as it will be necessary for the arguments below. Note that, in Zaitsev (2006), the trivial relation $\mathcal{L}\big(\zeta_k^{\{u\}}\big) \in \mathcal{A}_d(c\, u)$ with a universal constant c was used. To simplify the formulas, denote $\xi = \zeta_k^{\{u\}}$. Let $z \in \mathbb{C}^d$, $\|z\|\, c_5\, w_0 \leqslant 1$, $x = \operatorname{Re} z$, $y = \operatorname{Im} z$. Recall that $dw \leqslant w_0$. Choosing c_5 to be sufficiently large, and using (2.39), (2.42) and (2.55), we ensure the validity of the relations

$$\langle \mathbb{B}_k^{\{u\}} x, x\rangle + \langle \mathbb{B}_k^{\{u\}} y, y\rangle \leqslant C_1\, w^2\, \big(\|x\|^2 + \|y\|^2\big) = C_1\, w^2\, \|z\|^2 \leqslant 1, \tag{2.56}$$

$$
\begin{aligned}
\bigl|\,\mathbf{E}\,e^{\langle z,\xi\rangle}-1\bigr| &= \Bigl|\sum_{m=2}^{\infty}\frac{1}{m!}\,\mathbf{E}\,\bigl\langle z,\xi\bigr\rangle^{m}\Bigr| \\
&\leqslant \sum_{m=2}^{\infty}\frac{2^m}{m!}\Bigl(\,\mathbf{E}\,\bigl|\langle x,\xi\rangle\bigr|^{m}+\mathbf{E}\,\bigl|\langle y,\xi\rangle\bigr|^{m}\Bigr) \\
&\leqslant \sum_{m=2}^{\infty}2^m\Bigl(\bigl(\|x\|\,w_0\bigr)^{m-2}+\bigl(\|y\|\,w_0\bigr)^{m-2}\Bigr)\,d^2\,w^2\|z\|^2 \\
&\leqslant 8\,w_0^2\|z\|^2\leqslant\frac{1}{2}\,, \qquad (2.57)
\end{aligned}
$$

and, consequently,

$$
\frac{1}{2}\leqslant\bigl|\,\mathbf{E}\,e^{\langle z,\xi\rangle}\bigr|\leqslant\frac{3}{2}\,. \qquad (2.58)
$$

Choosing c_5 to be sufficiently large, expanding the exponent and using (2.55), we see that

$$
\mathbf{E}\,\bigl\langle\xi,v\bigr\rangle^2 e^{2|\langle\xi,x\rangle|}\ll d^2\,w^2\,\bigl\|v\bigr\|^2,\qquad \text{for all } v\in\mathbb{R}^d. \qquad (2.59)
$$

Differentiating the function $\varphi(V,z)$ with $V=\mathcal{L}\bigl(\zeta_k^{\{u\}}\bigr)$ (see (1.8)), we obtain

$$
\begin{aligned}
d_u d_v^2\,\varphi(V,z) = {} & \frac{\mathbf{E}\,\langle\xi,u\rangle\langle\xi,v\rangle^2 e^{\langle z,\xi\rangle}\,\mathbf{E}\,e^{\langle z,\xi\rangle}-\mathbf{E}\,\langle\xi,v\rangle^2 e^{\langle z,\xi\rangle}\,\mathbf{E}\,\langle\xi,u\rangle\,e^{\langle z,\xi\rangle}}{\bigl(\mathbf{E}\,e^{\langle z,\xi\rangle}\bigr)^2} \\
& -\frac{2\,\mathbf{E}\,\langle\xi,v\rangle\,e^{\langle z,\xi\rangle}}{\mathbf{E}\,e^{\langle z,\xi\rangle}}\cdot\frac{\mathbf{E}\,\langle\xi,u\rangle\langle\xi,v\rangle\,e^{\langle z,\xi\rangle}\,\mathbf{E}\,e^{\langle z,\xi\rangle}-\mathbf{E}\,\langle\xi,v\rangle\,e^{\langle z,\xi\rangle}\,\mathbf{E}\,\langle\xi,u\rangle\,e^{\langle z,\xi\rangle}}{\bigl(\mathbf{E}\,e^{\langle z,\xi\rangle}\bigr)^2}\,. \qquad (2.60)
\end{aligned}
$$

Using (2.39), (2.42), (2.55), (2.58)–(2.60), and the Hölder inequality, we have

$$
\begin{aligned}
\bigl|d_u d_v^2\,\varphi(V,z)\bigr| &\leqslant c\,\Bigl(\bigl(\,\mathbf{E}\,\langle\xi,u\rangle^2\langle\xi,v\rangle^2\,\mathbf{E}\,\langle\xi,v\rangle^2 e^{2\langle\xi,x\rangle}\bigr)^{1/2} \\
&\qquad +\bigl(\,\mathbf{E}\,\langle\xi,v\rangle^2+\mathbf{E}\,\langle\xi,v\rangle^2 e^{2\langle\xi,x\rangle}\bigr)\bigl(\,\mathbf{E}\,\langle\xi,u\rangle^2 e^{2\langle\xi,x\rangle}\bigr)^{1/2}\Bigr) \\
&\leqslant c\,\|u\|\,w_0\,d^2\,w^2\,\|v\|^2 \\
&\leqslant c_5\,d^2\,w_0\,\|u\|\langle\mathbb{B}_k^{\{u\}}v,v\rangle, \qquad (2.61)
\end{aligned}
$$

if c_5 is chosen to be large enough. This implies that $\mathcal{L}\bigl(\zeta_k^{\{u\}}\bigr)\in\mathcal{A}_d(c_5\,d^2\,w_0)$ (see (1.9)). Using (2.42) and the properties of the classes $\mathcal{A}_d(\tau)$, we see that

$$
\mathcal{L}\bigl(\zeta_k^{\{u\}}/w\bigr)\in\mathcal{A}_d(c_5\,d^2\,w_0/w)\subset\mathcal{A}_d(c_6\,d^3),
$$

with some $c_6\geqslant 1$.

By Theorem 3 applied (with $\tau=c_6\,d^3$) to the vectors $\zeta_k^{\{u\}}/w$, there exists a construction such that the joint distribution of $\bigl\{\zeta_k^{\{u\}}\bigr\}$ and $\bigl\{\eta_k^{\{u\}}\bigr\}$ satisfies

$$
\mathbf{E}\,\exp\left(\frac{c_7\,\Delta^{\{u\}}}{w\,d^{15/2}\log^* d}\right)=\mathbf{E}\,\exp\left(\frac{c_8\,\varepsilon\,\Delta^{\{u\}}\,\log^* s}{L_\gamma^{1/\gamma}\,d^{15/2}\log^* d}\right)\leqslant\exp\bigl(c_9\,d^{3+\alpha}\,\log^* s\bigr)\,, \qquad (2.62)
$$

where c_7, c_8, c_9 are positive quantities depending on α, γ, C_1, C_2 and C_3 only. Recall that we consider them as universal constants. The condition (1.10) of Theorem 3 is satisfied due to (2.39). By (2.62) and by the Chebyshev inequality,

$$\begin{aligned}\mathbf{P}\left(\Delta^{\{u\}} \geqslant x_0\right) &\leqslant \exp\left(c_9\, d^{3+\alpha}\, \log^* s - \frac{c_8\, \varepsilon\, x_0\, \log^* s}{L_\gamma^{1/\gamma} d^{15/2} \log^* d}\right)\\ &\leqslant \exp\left(-\frac{c_8\, \varepsilon\, x_0\, \log^* s}{2\, L_\gamma^{1/\gamma} d^{15/2} \log^* d}\right), \qquad (2.63)\end{aligned}$$

provided that $c_9\, d^{3+\alpha} < \dfrac{c_8\, \varepsilon\, x_0}{2\, L_\gamma^{1/\gamma} d^{15/2} \log^* d}$.

Using the relations (2.32), (2.34) and (2.63), we get, for $x_0 > 2\, L_\gamma^{1/\gamma} c_9\, c_8^{-1}\, \varepsilon^{-1} \cdot d^{21/2+\alpha} \log^* d$, $x_1 > c\, B(u)\sqrt{d}$ and $x \geqslant z \geqslant u$,

$$\begin{aligned}\mathbf{P}\left(\Delta \geqslant x_0 + x_1 + x + A(u)\right) &\leqslant 2\, \exp\bigl(-x_1^2/16\, B^2(u)\bigr)\\ &\quad + \exp\left(-\frac{c_8\, \varepsilon\, x_0\, \log^* s}{2\, L_\gamma^{1/\gamma} d^{15/2} \log^* d}\right)\\ &\quad + \mathbf{P}\left(\max_{1\leqslant k\leqslant s} \|\zeta_k\| \geqslant z\right)\\ &\quad + e\, G^{1-x/z}(z)/G(u). \qquad (2.64)\end{aligned}$$

We shall apply this inequality in the case, where $G(x)$ is defined in (2.30) and

$$u \leqslant z = x_0 = x_1 \leqslant x = 2z. \qquad (2.65)$$

By (2.1), (2.2), (2.11), (2.17), (2.30), (2.36) and (2.65), we have

$$A(u)/u \leqslant B^2(u)/u^2 \leqslant 1, \qquad (2.66)$$

$$\mathbf{P}\left(\max_{1\leqslant k\leqslant s} \|\zeta_k\| \geqslant z\right) \leqslant \sum_{k=1}^{s} \mathbf{P}\left(\|\zeta_k\| \geqslant z\right) = F(z) \ll d^{1+\gamma/2}\, \frac{L_\gamma}{z^\gamma}. \qquad (2.67)$$

Recall that $G(u) \geqslant e^\gamma$ (see (2.31)). Finally, by the relations (2.64)–(2.67),

$$\begin{aligned}\mathbf{P}\ (\Delta \geqslant 5z) &\leqslant \mathbf{P}\ (\Delta \geqslant 4z + A(u))\\ &= \mathbf{P}\ (\Delta \geqslant x_0 + x_1 + x + A(u))\\ &\leqslant 2\, \exp\left(-z^2/16\, u^2\right) + \exp\left(-\frac{c_{10}\, \varepsilon\, z\, \log^* s}{L_\gamma^{1/\gamma} d^{15/2} \log^* d}\right) + 2\, F(z), \qquad (2.68)\end{aligned}$$

provided that

$$z > z_0 \overset{\text{def}}{=} c_{11}\, \varepsilon^{-1}\, L_\gamma^{1/\gamma}\, d^{21/2+\alpha}\, \log^* d \geqslant d^9\, u \geqslant u, \qquad (2.69)$$

where c_{10}, c_{11} are positive quantities depending on α, γ, C_1, C_2 and C_3 only. The c_{11} will be chosen to be sufficiently large. Define

$$\Delta^* = \max_{1\leqslant k\leqslant s}\ \max_{m_{k-1}+1\leqslant m\leqslant m_k}\Big\|\sum_{j=m_{k-1}+1}^{m}\xi_j\Big\|, \tag{2.70}$$

$$\Delta^{**} = \max_{1\leqslant k\leqslant s}\ \max_{m_{k-1}+1\leqslant m\leqslant m_k}\Big\|\sum_{j=m_{k-1}+1}^{m}Y_j\Big\|, \tag{2.71}$$

Then, according to (1.11), (2.9), (2.15), (2.70) and (2.71), we have, for any $z\geqslant 0$,

$$\mathbf{P}\left(\Delta(\xi,Y)\geqslant 9z\right)\leqslant \mathbf{P}\left(\Delta\geqslant 5z\right)+\mathbf{P}\left(\Delta^*\geqslant 3z\right)+\mathbf{P}\left(\Delta^{**}\geqslant z\right). \tag{2.72}$$

To bound the second summand in (2.72), we apply Lemma 1. By this lemma, we derive

$$\begin{aligned}\mathbf{P}\left(\Delta^*\geqslant 3z\right)&\leqslant\sum_{k=1}^{s}\mathbf{P}\Big(\max_{m_{k-1}+1\leqslant m\leqslant m_k}\Big\|\sum_{j=m_{k-1}+1}^{m}\xi_j\Big\|\geqslant 3z\Big)\\&\leqslant 3\sum_{k=1}^{s}\max_{m_{k-1}+1\leqslant m\leqslant m_k}\mathbf{P}\Big(\Big\|\sum_{j=m_{k-1}+1}^{m}\xi_j\Big\|\geqslant z\Big)\\&=3\sum_{k=1}^{s}T_k(z),\end{aligned} \tag{2.73}$$

(see (2.21)). By (2.41) and by the Lévy inequality,

$$\begin{aligned}\mathbf{P}\left(\Delta^{**}\geqslant z\right)&\leqslant\sum_{k=1}^{s}\mathbf{P}\Big(\max_{m_{k-1}+1\leqslant m\leqslant m_k}\Big\|\sum_{j=m_{k-1}+1}^{m}Y_j\Big\|\geqslant z\Big)\\&\leqslant 2\sum_{k=1}^{s}\mathbf{P}\Big(\Big\|\sum_{j=m_{k-1}+1}^{m}Y_j\Big\|\geqslant z\Big)\\&\leqslant 2\,\exp\left(-z^2/16C_1w^2\right),\end{aligned} \tag{2.74}$$

for $z>u$. Collecting the bounds (2.22), (2.67), (2.68) and (2.72)–(2.74), we obtain

$$\begin{aligned}\mathbf{P}\left(\Delta(\xi,Y)\geqslant 9z\right)&\leqslant 5\sum_{k=1}^{s}T_k(z)+2\,\exp\left(-z^2/16C_1w^2\right)\\&\quad+2\,\exp\left(-z^2/16\,u^2\right)+\exp\Big(-\frac{c_{10}\,\varepsilon\,z\,\log^* s}{L_\gamma^{1/\gamma}d^{15/2}\log^* d}\Big),\end{aligned} \tag{2.75}$$

provided that $z > z_0$. The relations (1.13), (1.21), (2.25), (2.27) and (2.75) imply the inequality (1.16).

Choosing c_{11} to be large enough and using (2.69) and that $r^{\beta}\,\beta^{-\beta} \leqslant e^{r}$, we obtain the rough bound

$$\begin{aligned}\exp\Big(-\frac{c_{10}\,\varepsilon\,z}{L_\gamma^{1/\gamma} d^{15/2}\log^* d}\Big) &\leqslant \Big(\frac{2\,L_\gamma^{1/\gamma} d^{15/2}\log^* d\cdot\gamma\,b}{c_{10}\,\varepsilon\,z}\Big)^{\gamma b}\exp\Big(-\frac{c_{10}\,\varepsilon\,z}{2\,L_\gamma^{1/\gamma} d^{15/2}\log^* d}\Big)\\ &\leqslant \big(2\,c_{10}^{-1}\,\gamma\,b\,\varepsilon^{-1}\,d^{15/2}\,\log^* d\big)^{\gamma b}\,\big(L_\gamma\,z^{-\gamma}\big)^{b}\,e^{-d^3},\end{aligned}$$

for any $b \geqslant 1$ and $z > z_0$. Moreover, $(d^9)^{\gamma b} \leqslant (3\gamma b)^{3\gamma b}\,e^{d^3}$, and, hence, there exists c_{13} such that

$$\exp\Big(-\frac{c_{10}\,\varepsilon\,z}{L_\gamma^{1/\gamma} d^{15/2}\log^* d}\Big) \leqslant \big(c_{13}\,\gamma^4\,b^4\,\varepsilon^{-1}\big)^{\gamma b}\,\big(L_\gamma\,z^{-\gamma}\big)^{b}, \tag{2.76}$$

for any $b \geqslant 1$ and $z > z_0$. In a similar way one can estimate the other exponential terms in (2.75). Using (1.14), (1.15) and (1.21), we get

$$\sum_{k=1}^{s}\lambda_{k,\gamma}^{b} \leqslant L_\gamma\,\big(\max_{1\leqslant k\leqslant s}\lambda_{k,\gamma}\big)^{b-1} \leqslant L_\gamma^{b}/\big(C_3\,d^{\gamma/2}\,s^{\varepsilon}\,(\log^* s)^{\gamma+3}\big)^{b-1}, \quad \text{for } b \geqslant 1. \tag{2.77}$$

The relations (1.13), (2.25), (2.27) and (2.75)–(2.77) imply the inequality (1.17). The inequality (1.18) may be easily deduced from (1.17) with $b = 1$ estimating the last sum by the Chebyshev inequality.

Integrating by parts, we get

$$\mathbf{E}\,\big(\Delta(\xi,Y)/9\big)^{\gamma} = \gamma\int_0^\infty z^{\gamma-1}\,\mathbf{P}\,\big(\Delta(\xi,Y) \geqslant 9\,z\big)\,dz = \gamma\big(K_1+K_2\big), \tag{2.78}$$

where

$$K_1 = \int_0^{z_0} z^{\gamma-1}\,\mathbf{P}\,\big(\Delta(\xi,Y) \geqslant 9\,z\big)\,dz, \tag{2.79}$$

and

$$K_2 = \int_{z_0}^{\infty} z^{\gamma-1}\,\mathbf{P}\,\big(\Delta(\xi,Y) \geqslant 9\,z\big)\,dz. \tag{2.80}$$

It is easy to see that, by (2.69) and (2.79)

$$K_1 \leqslant \int_0^{z_0} z^{\gamma-1}\,dz = \frac{z_0^{\gamma}}{\gamma} \ll L_\gamma\,\big(\varepsilon^{-1}\,d^{21/2+\alpha}\,\log^* d\big)^{\gamma}. \tag{2.81}$$

Applying the inequality (1.17) with $b = 2$, we see that

$$K_2 \ll J_1 + J_2, \tag{2.82}$$

where

$$J_1 = d^{1+\gamma/2}\,\varepsilon^{-2\gamma}\,L_\gamma^2 \int_{z_0}^{\infty} \frac{dz}{z^{\gamma+1}}\,, \tag{2.83}$$

and

$$J_2 = d \int_{z_0}^{\infty} z^{\gamma-1} \sum_{j=1}^{n} \mathbf{P}\left(\|\xi_j\| \geqslant z/4\sqrt{d}\right) dz. \tag{2.84}$$

Using (2.69) and (2.83), we see that

$$J_1 \ll d^{1+\gamma/2}\,\varepsilon^{-2\gamma}\,L_\gamma^2\,z_0^{-\gamma} \ll \varepsilon^{-\gamma}\,L_\gamma. \tag{2.85}$$

Moreover,

$$\int_{z_0}^{\infty} z^{\gamma-1}\,\mathbf{P}\left(\|\xi_j\| \geqslant z/4\sqrt{d}\right) dz \ll \mathbf{E}\left(4\sqrt{d}\,\|\xi_j\|\right)^{\gamma} \ll d^{\gamma/2}\,\mathbf{E}\,\|\xi_j\|^{\gamma}, \tag{2.86}$$

Using (1.5), (2.84) and (2.86), we obtain

$$J_2 \ll d^{1+\gamma/2}\,L_\gamma. \tag{2.87}$$

Collecting the relations (2.78)–(2.82), (2.85) and (2.87), we obtain

$$\mathbf{E}\left(\Delta(\xi, Y)\right)^{\gamma} \ll L_\gamma\left(\varepsilon^{-1}\,d^{21/2+\alpha}\,\log^* d\right)^{\gamma}, \tag{2.88}$$

proving (1.19). The proof of Theorem 4 is thus completed. □

Proof of Proposition 1. Similarly to the proof of Theorem 4, we shall treat γ and C_4 as absolute positive constants.

Define

$$\Delta_1 = \max_{1\leqslant m\leqslant n} \Bigl\|\sum_{j=1}^{m} X_j\Bigr\|, \tag{2.89}$$

$$\Delta_2 = \max_{1\leqslant m\leqslant n} \Bigl\|\sum_{j=1}^{m} Y_j\Bigr\|, \tag{2.90}$$

Then, according to (1.1), (2.89) and (2.90), we have, for any $z \geqslant 0$,

$$\mathbf{P}\left(\Delta(X, Y) \geqslant 4z\right) \leqslant \mathbf{P}\left(\Delta_1 \geqslant 3z\right) + \mathbf{P}\left(\Delta_2 \geqslant z\right). \tag{2.91}$$

To bound the first summand, we apply Lemma 1. By this lemma, we derive

$$\mathbf{P}\left(\Delta_1 \geqslant 3z\right) \leqslant 3 \max_{1\leqslant m\leqslant n} \mathbf{P}\left(\Bigl\|\sum_{j=1}^{m} X_j\Bigr\| \geqslant z\right) \tag{2.92}$$

(see (2.89)). By (2.90) and by the Lévy inequality,

$$\mathbf{P}\left(\Delta_2 \geqslant z\right) \leqslant 2\,\mathbf{P}\Big(\Big\|\sum_{j=1}^{n} Y_j\Big\| \geqslant z\Big). \tag{2.93}$$

Applying Lemma 2 and using (1.5), we get, similarly to (2.23), that, for any $z > 0$ and $b \geqslant 1$,

$$\begin{aligned}\mathbf{P}\Big(\Big\|\sum_{j=1}^{m} X_j\Big\| \geqslant z\Big) &\leqslant 2d\,\exp\left(-c_{12}\,z^2/d\sigma^2\right) + 2\,(2b)^{b(\gamma-1)}\,d^{1+\gamma b/2}\,\frac{L_\gamma^b}{z^{b\gamma}}\\ &\qquad + d\sum_{j=1}^{n}\mathbf{P}\left(\left\|X_j\right\| \geqslant z/2b\sqrt{d}\right), \qquad m = 1,\dots,n. \end{aligned}\tag{2.94}$$

Integrating by parts and using (2.91), we get

$$\mathbf{E}\left(\Delta(X,Y)/4\right)^\gamma = \gamma\int_0^\infty z^{\gamma-1}\,\mathbf{P}\left(\Delta(X,Y) \geqslant 4z\right)dz = \gamma\left(M_1 + M_2 + M_3\right), \tag{2.95}$$

where

$$M_1 = \int_0^y z^{\gamma-1}\,\mathbf{P}\left(\Delta(X,Y) \geqslant 4z\right)dz, \tag{2.96}$$

$$M_2 = \int_y^\infty z^{\gamma-1}\,\mathbf{P}\left(\Delta_1 \geqslant 3z\right)dz, \tag{2.97}$$

and

$$M_3 = \int_y^\infty z^{\gamma-1}\,\mathbf{P}\left(\Delta_2 \geqslant z\right)dz. \tag{2.98}$$

with $y = c_{13}\sqrt{d}\,L_\gamma^{1/\gamma}$ and sufficiently large constant c_{13}. It is easy to see that

$$M_1 \leqslant \int_0^y z^{\gamma-1}\,dz = \frac{y^\gamma}{\gamma} \ll d^{\gamma/2}\,L_\gamma, \tag{2.99}$$

To estimate M_2, we shall apply the inequality (2.94) with $b = 2$. It is clear that

$$\int_y^\infty z^{\gamma-1}\,\exp\left(-c_{12}\,z^2/d\sigma^2\right)dz \ll d^{\gamma/2}\,\sigma^\gamma \leqslant d^{\gamma/2}\,L_\gamma, \tag{2.100}$$

(see (1.22)) and

$$\int_y^\infty z^{\gamma-1}\,\frac{L_\gamma^2}{z^{2\gamma}}\,dz \ll L_\gamma^2\,y^{-\gamma} \ll d^{-\gamma/2}\,L_\gamma. \tag{2.101}$$

Moreover,

$$\int_y^\infty z^{\gamma-1}\,\mathbf{P}\left(\|\xi_j\| \geqslant z/4\sqrt{d}\right)dz \ll \mathbf{E}\left(4\sqrt{d}\,\|\xi_j\|\right)^\gamma = 4^\gamma\,d^{\gamma/2}\,\mathbf{E}\,\|\xi_j\|^\gamma, \tag{2.102}$$

and, by (1.5), (2.92), (2.94), (2.97) and (2.100)–(2.102),

$$M_2 \ll d^{1+\gamma/2}\,L_\gamma. \tag{2.103}$$

Arguing similarly to the proof of (2.14) and using the definition of σ^2, we obtain that there exists an absolute constant c_{14} such that, for any $x > c_{14}\,\sigma\sqrt{d}$,

$$\mathbf{P}\left(\Big\|\sum_{j=1}^n Y_j\Big\| \geqslant z\right) \leqslant \exp\left(-z^2/8\sigma^2\right). \tag{2.104}$$

Choosing $c_{13} \geqslant c_{14}\,C_4$ and using (1.22), we have $y \geqslant c_{14}\,\sigma\sqrt{d}$. By (1.22), (2.93), (2.98) and (2.104), we get

$$M_3 \ll \sigma^\gamma \leqslant L_\gamma. \tag{2.105}$$

Collecting the relations (2.95), (2.99), (2.103) and (2.105), we obtain

$$\mathbf{E}\left(\Delta(X,Y)/4\right)^\gamma \ll d^{1+\gamma/2}\,L_\gamma. \tag{2.106}$$

proving (1.24). Note that a similar inequality for $\left\|\sum_{j=1}^n X_j - \sum_{j=1}^n Y_j\right\|$ instead of $\Delta(X,Y)$ might be easily derived from Lemma 3. The inequality (1.25) follows from (1.24) by the Chebyshev inequality.

By (1.22), there exists c_{15} such that

$$\exp\left(-c_{12}\,z^2/d\sigma^2\right) + \exp\left(-z^2/8\sigma^2\right) \leqslant (c_{15}\,\gamma\,b\,d)^{b\gamma/2}\,\frac{L_\gamma^b}{z^{b\gamma}}. \tag{2.107}$$

The inequality (1.23) follows from (2.91)–(2.94), (2.104) and (2.107). The proof of Proposition 1 is thus completed. □

REFERENCES

Berger, E., *Fast sichere Approximation von Partialsummen unabhängiger und stationärer ergodischer Folgen von Zufallsvectoren*, Dissertation (1982), Universität Göttingen.

Berkes, I. and Philipp, W., *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. **7** (1979), 29–54.

Borovkov, A. A., *On the rate of convergence in the invariance principle*, Theor. Probab. Appl. **18** (1973), no. 2, 207–225.

Borovkov, A. A. and Sakhanenko, A. I., *On the rate of convergence in invariance principle*, Lect. Notes Math. **1021** (1981), 59–66.

Borovkov, A. A. and Sakhanenko, A. I., *On estimates for the rate convergence in the invariance principle for Banach spaces*, Theor. Probab. Appl. **25** (1980), 721–731.

Csörgő, M. and Révész, P., *A new method to prove Strassen type laws of invariance principle.* I; II, Z. Wahrscheinlichkeitstheor. verw. Geb. **31** (1975), 255–259; 261–269.

Csörgő, M. and Révész, P., *Strong approximations in probability and statistics*, Academic Press, New York, 1981.

Csörgő, S. and Hall, P., *The Komlós–Major–Tusnády approximations and their applications*, Austral. J. Statist. **26** (1984), no. 2, 189–218.

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

de la Peña, V. H. and Giné, E., *Decoupling. From dependence to independence. Randomly stopped processes. U-statistics and processes. Martingales and beyond*, Probability and its Applications (New York). Springer-Verlag, New York, 1999.

Ebralidze, Š. S., *Inequalities for probabilities of large deviations in the multidimensional case*, Theor. Probab. Appl. **16** (1971), 737–741.

Einmahl, U., *A useful estimate in the multidimensional invariance principle*, Probab. Theor. Rel. Fields **76** (1987a), no. 1, 81–101.

Einmahl, U., *Strong invariance principles for partial sums of independent random vectors*, Ann. Probab. **15** (1987b), 1419–1440.

Einmahl, U., *Extensions of results of Komlós, Major and Tusnády to the multivariate case*, J. Multivar. Anal. **28** (1989), 20–68.

Etemadi, N., *On some classical results in probability theory*, Sankhyā, Ser. A **47** (1985), no. 2, 215–221.

Fuk, D. H. and Nagaev, S. V., *Probability inequalities for sums of independent random variables*, Theor. Probab. Appl. **16** (1971), 643–660.

Gorodetskii, V. V., *On the rate of convergence for the multidimensional invariance principle.*, Theor. Probab. Appl. **20** (1975), 631–638.

Kwapień, S. and Woyczynski, W., *Random series and stochastic integrals: single and multiple*, Birkhäuser, Boston, 1992.

Komlós, J., Major, P. and Tusnády, G., *An approximation of partial sums of independent RV'-s and the sample DF.* I; II, Z. Wahrscheinlichkeitstheor. verw. Geb. **32** (1975), 111–131; **34** (1976), 34–58.

Major, P., *On the invariance principle for sums of independent identically distributed random variables*, J. Multivar. Anal. **8** (1978), 487–517.

Nagaev, S. V., *Large deviations of sums of independent random variables*, Ann. Probab. **7** (1979), no. 5, 745–789.

Philipp, W., *Almost sure invariance principles for sums of B-valued random variables*, Lect. Notes in Math. **709** (1979), 171–193.

Prokhorov, Yu. V., *Convergence of random processes and limit theorem of probability theory*, Theor. Probab. Appl. **1** (1956), no. 2, 157–214.

Rosenthal, H. P., *On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables*, Israel J. Math. **8** (1970), 273–303.

Rosenthal, H. P., *On the span in L^p of sequences of independent random variables. II*, Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory, Univ. California Press, Berkeley, Calif., 1972, pp. 149–167.

Sakhanenko, A. I., *Rate of convergence in the invariance principles for variables with exponential moments that are not identically distributed*, In: Trudy Inst. Mat. SO AN SSSR, vol. 3, Nauka, Novosibirsk, 1984, pp. 4–49. (Russian)

Sakhanenko, A. I., *Estimates in the invariance principles*, In: Trudy Inst. Mat. SO AN SSSR, vol. 5, Nauka, Novosibirsk, 1985, pp. 27–44. (Russian)

Sakhanenko, A. I., *A new way to obtain estimates in the invariance principle*, High dimensional probability, II (Seattle, 1999), Progr. Probab., 47, Birkhäuser Boston, Boston, 2000, pp. 223–245.

Sakhanenko, A. I., *Estimates in the Invariance Principle in terms of truncated moments*, Siberian Math. J., vol. 47, 2006, pp. 1355–1371. (Russian)

Shao, Qi-Man, *Strong approximation theorems for independent random variables and their applications*, J. Multivar. Anal. **52** (1995), no. 1, 107–130.

Skorokhod, A. V., *Studies in the theory of random processes*, Univ. Kiev, Kiev, 1961 (Russian); Engl. transl., Addison–Wesley, Reading, Mass., 1965.

Zaitsev, A. Yu., *Estimates of the Lévy–Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments*, Theor. Probab. Appl. **31** (1986), no. 2, 203–220.

Zaitsev, A. Yu., *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments*, ESAIM : Probability and Statistics **2** (1998), 41–108.

Zaitsev, A. Yu., *Multidimensional version of the results of Sakhanenko in the invariance principle for vectors with finite exponential moments.* I; II; III, Theor. Probab. Appl. **45** (2000), 718–738; **46** (2001), 535–561; 744–769.

Zaitsev, A. Yu., *Estimates for the strong approximation in multidimensional Central Limit Theorem*, In: Proceedings of the International Congress of Mathematicians. Bejing 2002. Vol. III. Invited Lectures, 2002, pp. 107–116.

Zaitsev, A. Yu., *Estimates for the strong approximation in the multidimensional invariance principle*, Zapiski Nauchnykh Seminarov LOMI **339** (2006), 37–53 (Russian); English transl. in J. Math. Sci.

Friedrich Götze
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33501 Bielefeld 1
Germany
E-mail address: `goetze@math.uni-bielefeld.de`

Andrei Zaitsev
St. Petersburg Department of Steklov Mathematical Institute,
Fontanka 27, St. Petersburg 191023, Russia
E-mail address: `zaitsev@pdmi.ras.ru`