RATES OF APPROXIMATION IN THE MULTIDIMENSIONAL INVARIANCE PRINCIPLE FOR SUMS OF I.I.D. RANDOM VECTORS WITH FINITE MOMENTS

F. Götze[1] and A. Yu. Zaitsev[1,2]

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

Abstract. The aim of this paper is to derive some consequences of a result of Götze and Zaitsev [5]. We shall show that the i.i.d. case of this result implies a multidimensional version of some results of Sakhanenko [12]. We establish bounds for the rate of strong Gaussian approximation of sums of i.i.d. $\mathbf{R}^d$-valued random vectors ξ_j having finite moments $\mathbf{E}\,\|\xi_j\|^\gamma$, $\gamma > 2$.

1. Introduction

The aim of this paper is to derive some consequences of the main result of Götze and Zaitsev [5] (see Theorem 2 below). We shall show that the i.i.d. case of this result implies the multidimensional version of a result of Sakhanenko [12]. We shall obtain bounds for the rate of strong Gaussian approximation of sums of i.i.d. $\mathbf{R}^d$-valued random vectors ξ_j having finite moments $\mathbf{E}\,\|\xi_j\|^\gamma$, $\gamma > 2$.

We consider the following well-known problem. One has to construct on a probability space a sequence of independent random vectors $X_1, \dots, X_n$ (with given distributions) and a corresponding sequence of independent Gaussian random vectors $Y_1, \dots, Y_n$ so that the quantity

$$\Delta_n(X,Y) = \max_{1\leqslant k\leqslant n} \Big\| \sum_{j=1}^{k} X_j - \sum_{j=1}^{k} Y_j \Big\| \tag{1.1}$$

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 Russian Foundation of Basic Research (RFBR) Grant 09-01-12180, by RFBR-DFG Grant 09-01–91331, by the Grant of leading scientific scools NSh 638.2008.1 and by a program of fundamental researches of the RAS "Modern problems of fundamental mathematics".

would be as 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. We omit the detailed history of the problem referring the reader to Götze and Zaitsev [5] and Zaitsev [13].

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 $\operatorname{cov}\xi$. We write $\log^* b = \max\{1, \log b\}$, for $b > 0$. By $\lceil x\rceil$ we shall denote the integer part of a number x.

The aim of the present paper is to obtain multidimensional analogues of the following result of Sakhanenko in the case of i.i.d. summands.

Theorem 1 (Sakhanenko [12])**.** *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 $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{Var} Y_j = \operatorname{Var} X_j$, $j = 1, \dots, n$, and

$$\mathbf{E}\left(\Delta_n(X, Y)\right)^\gamma \leqslant c\,\gamma^{2\gamma}\,L_\gamma, \tag{1.2}$$

where c is an absolute constant.

It should be mentioned that, in Sakhanenko [12], somewhat more general results are proved. Sakhanenko [12] noted that the inequality (1.2) implies the well-known Rosenthal [10, 11] inequality (see Lemma 1 below).

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

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

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.

In this paper we shall derive Theorems 3 and 4 which are, in fact, rather immediate consequences of the following Theorem 2 which was proved in Götze and Zaitsev [5]. In Theorem 2, we considered the case of (generally speaking) non-i.i.d. summands. Theorem 3 is derived from Theorem 2 in the particular case, where the summands are identically distributed. Theorem 3 is a multidimensional version of Theorem 1 for i.i.d. summands.

Theorem 2. *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 let the quantity* L_γ *be defined by*

$$L_\gamma = \sum_{j=1}^{n} \mathbf{E}\,\|\xi_j\|^\gamma < \infty. \tag{1.3}$$

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*

$$\zeta_k = \xi_{m_{k-1}+1} + \dots + \xi_{m_k}, \quad \operatorname{cov}\zeta_k = \mathbb{B}_k, \qquad k = 1, \dots, s, \tag{1.4}$$

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

$$w^2\,\|v\|^2 \leqslant \big\langle \mathbb{B}_k v, v \big\rangle \leqslant C_1\, w^2\,\|v\|^2\,, \tag{1.5}$$

where

$$w = C_2\, L_\gamma^{1/\gamma} / \log^* s, \tag{1.6}$$

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

$$\lambda_{k,\gamma} = \sum_{j=m_{k-1}+1}^{m_k} \mathbf{E}\,\|\xi_j\|^\gamma, \qquad k = 1, \dots, s, \tag{1.7}$$

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

$$C_3\, d^{\gamma/2}\, s^\varepsilon\, (\log^* s)^{\gamma+3} \max_{1 \leqslant k \leqslant s} \lambda_{k,\gamma} \leqslant L_\gamma, \tag{1.8}$$

with a positive quantity 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*

$$\mathbf{E}\,\big(\Delta_n(X, Y)\big)^\gamma \leqslant a_1\,\big(\varepsilon^{-1}\, d^{21/2+\alpha}\, \log^* d\big)^\gamma\, L_\gamma, \tag{1.9}$$

where a_1 *is a positive quantity depending on* α, γ, C_1, C_2 *and* C_3 *only.*

Condition (1.8) is rather cumbersome, but we shall show that it is trivially satisfied if $\xi_1, \dots, \xi_n$ are i.i.d. random vectors, and n is sufficiently large.

The main results of this paper are the following Theorems 3 and 4.

Theorem 3. *Suppose that* $\alpha > 0$, *and* ξ *is a* $\mathbf{R}^d$*-valued random vector with* $\mathbf{E}\,\xi = 0$ *and* $\mathbf{E}\,\|\xi\|^\gamma < \infty$, *for some* $\gamma \geqslant 2$. *Let* $\sigma^2_{\max}$ *and* $\sigma^2_{\min}$ *be the maximal and minimal strictly positive eigenvalues of the covariance operator* $\operatorname{cov}\xi$ *respectively. Let* n *be an arbitrary positive integer. 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* $Y_1, \dots, Y_n$ *such that* $\mathcal{L}(X_j) = \mathcal{L}(\xi)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, \dots, n$, *and the following inequality is valid*:

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \leqslant a_2\,A\,\big(\sigma_{\max}/\sigma_{\min}\big)^\gamma\, n\,\mathbf{E}\,\|\xi\|^\gamma, \tag{1.10}$$

where

$$A = A(\gamma,\alpha,d) = \max\Big\{\big(d^{21/2+\alpha}\,(\log^* d)^2\big)^\gamma,\ d^{\frac{\gamma(\gamma+2)}{4}}\,(\log^* d)^{\frac{\gamma(\gamma+1)}{2}}\Big\}, \tag{1.11}$$

and a_2 *is a positive quantity depending on* γ *and* α *only.*

Theorem 4. *Let the conditions of Theorem* 3 *be satisfied. Then one can construct on a probability space a sequence of independent random vectors* $X_1, X_2, \dots$ *and a corresponding sequence of independent Gaussian random vectors* $Y_1, Y_2, \dots$ *such that* $\mathcal{L}(X_j) = \mathcal{L}(\xi)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, 2 \dots$, *and the following inequality is valid*:

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \leqslant a_3\,A\,\big(\sigma_{\max}/\sigma_{\min}\big)^\gamma\, n\,\mathbf{E}\,\|\xi\|^\gamma, \quad \text{for all } n, \tag{1.12}$$

where A *is defined in* (1.11) *and* a_3 *is a positive quantity depending on* γ *and* α *only.*

Applying the Chebyshev inequality, we see that, in the conditions of Theorem 4,

$$\mathbf{P}\,\big\{\,\Delta_n(X,Y) \geqslant x\,\big\} \leqslant a_3\,A\,\big(\sigma_{\max}/\sigma_{\min}\big)^\gamma\, n\,\mathbf{E}\,\|\xi\|^\gamma/x^\gamma, \tag{1.13}$$

for all $x > 0$ and all $n = 1, 2, \dots$. Clearly, the statement of Theorem 4 is stronger than (1.13). A construction for which (1.13) is valid for $d = 1$ for fixed n and $x = O\big(\sqrt{n\,\log n}\big)$ with a_3 depending on γ and $\mathcal{L}(\xi)$ was proposed by Komlós, Major, and Tusnády (KMT) [7], see also Borovkov [1] and Major [9] in the case where $2 < \gamma \leqslant 3$. Then Sakhanenko [12] proved Theorem 1 which provides the validity of one-dimensional version of inequality (1.13) for all x on the same probability space. Einmahl [4] obtained a multidimensional version of the KMT result without restrictions on the values of x.

Theorem 3 is formulated for fixed n. This means that the probability space depends on this n. However, an application of the result for fixed n allows to modify the construction in such a way that (1.12) is satisfied for all n simultaneously on the same probability space (Theorem 4). It suffices to use independent constructions which exist by Theorem 3 with fixed $n \asymp 2^m$, $m = 1, 2, \ldots$. A version of Theorem 1 which is valid for all n simultaneously on the same probability space was obtained by Lifshits [8] by an application of the Rosenthal inequality (see Lemma 1). We may derive Theorem 4 from Theorem 3 by repeating the arguments of Lifshits [8]. However, an application of an analogue of the Rosenthal inequality for sums of independent non-negative random variables (see Lemma 2) yields a simpler proof. We do not try to optimize the dependence of constants on the dimension d. It is important that it is a power-type dependence. Fixing α, say, $\alpha = 1/2$, we could rewrite inequality (1.12) in the form

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \leqslant a_4\big(\sigma_{\max}/\sigma_{\min}\big)^\gamma\, n\,\mathbf{E}\,\|\xi\|^\gamma, \quad \text{for all } n, \tag{1.14}$$

where a_4 is a positive quantity depending on γ and d only. In [5], it was noted that Theorem 2 implies the validity of (1.14) for sufficiently large fixed n. In this paper we show that there exists a construction such that (1.14) holds for all n simultaneously.

If one or more eigenvalues of the operator $\operatorname{cov}\xi$ are such that their sum is infinitesimally small in comparison with the rest of eigenvalues, then Theorems 3 and 4 could provide stronger results if one applies them to sums of coordinates with "large" variances. The other coordinates may be constructed in an arbitary way. The moments of their sums could be estimated by inequalities (1.15) and (1.16), by analogy with (2.16). The existence of construction is ensured by the well-known Berkes–Philipp lemma. Since inequalities (1.15) and (1.16) are valid for Hilbert-valued random vectors too, these arguments could be used for obtaining comprehensive results even in infinite dimensional situation.

For the proof we need the following Lemmas 1–3.

Lemma 1. *Let $\xi_1, \ldots, \xi_n$ be independent random vectors which have mean zero and assume values in $\mathbf{R}^d$. Then*

$$\mathbf{E}\,\Big\|\sum_{j=1}^n \xi_j\Big\|^\gamma \leqslant a_5\Big(\sum_{j=1}^n \mathbf{E}\,\|\xi_j\|^\gamma + \Big(\sum_{j=1}^n \mathbf{E}\,\|\xi_j\|^2\Big)^{\gamma/2}\Big), \quad \textit{for } \gamma \geqslant 2, \tag{1.15}$$

with a positive quantity a_5 depending on γ only.

This multidimensional version of the Rosenthal inequality follows easily from a result of de Acosta [2]. In the i.i.d. case, the second summand in the right-hand side of (1.15) grows faster than the first one as $n \to \infty$. Inequality (1.14) shows that this growth corresponds to the growth of moments of sums of Gaussian approximating vectors.

Lemma 2 below is proved by Rosenthal [10], see also Johnson, Schechtman and Zinn [6].

Lemma 2. *Let $\xi_1,\dots,\xi_n$ be independent random variables which are non-negative with probability one. Then*

$$\mathbf{E}\,\Big\|\sum_{j=1}^{n}\xi_j\Big\|^{\gamma}\leqslant a_6\Big(\sum_{j=1}^{n}\mathbf{E}\,\|\xi_j\|^{\gamma}+\Big(\sum_{j=1}^{n}\mathbf{E}\,\|\xi_j\|\Big)^{\gamma}\Big),\quad \textit{for}\ \ \gamma\geqslant 1,$$

with a_6 depending on γ only.

The following Lemma 3 is a particular case of Theorem 1.1.5 from the monograph of de la Peña and Giné [3].

Lemma 3. *Let $\xi_1,\dots,\xi_n$ be i.i.d. random vectors which assume values in $\mathbf{R}^d$. Then, for all $x\geqslant 0$,*

$$\mathbf{P}\Big\{\max_{1\leqslant k\leqslant n}\ \Big\|\sum_{j=1}^{k}\xi_j\Big\|>x\Big\}\leqslant 9\ \mathbf{P}\Big\{\Big\|\sum_{j=1}^{n}\xi_j\Big\|>x/30\Big\}.$$

Combined with the well-known equality

$$\mathbf{E}\ |\eta|^{\gamma}=\gamma\int_0^{\infty}x^{\gamma-1}\ \mathbf{P}\big\{\,|\eta|>x\,\big\}\,dx,\quad \gamma>0,$$

which is valid for any random variable η, Lemma 3 allows one to estimate the moments

$$\mathbf{E}\max_{1\leqslant k\leqslant n}\ \Big\|\sum_{j=1}^{k}\xi_j\Big\|^{\gamma}\leqslant a_7\,\mathbf{E}\,\Big\|\sum_{j=1}^{n}\xi_j\Big\|^{\gamma},\quad \gamma>0, \tag{1.16}$$

in the case of i.i.d. random vectors $\xi_1,\dots,\xi_n$, where a_7 is a quantity depending on γ only.

2. Proofs

The symbols $c,c_1,c_2,\dots$ will be used for absolute positive constants. The letter c may denote different constants if we are not interested in their numerical values. The same notation is used for positive quantities which depend only on α and γ involved in the conditions of Theorem 3. This means, in fact, that we consider α and γ as absolute positive constants. We shall write $A\ll B$, if there exists a c such that $A\leqslant cB$. We shall also use notation $A\asymp B$, if $A\ll B\ll A$.

Proof of Theorem 3. Without loss of generality, we assume that $\gamma>2$ and $\operatorname{cov}\xi=\mathbb{I}_d$. Indeed, if $\gamma=2$, then inequality (1.10) is evident. It is clear that we can assume that the operator $\mathbb{B}=\operatorname{cov}\xi$ is non-degenerate. If $\mathbb{B}\neq\mathbb{I}_d$, we can consider the vector $\mathbb{B}^{-1/2}\,\xi$ instead of ξ.

By the Lyapunov inequality,

$$1 \leqslant d = \mathbf{E}\,\|\xi\|^2 \leqslant (\,\mathbf{E}\,\|\xi\|^\gamma)^{2/\gamma}. \tag{2.1}$$

Let $\xi_1, \dots, \xi_n$ be independent copies of the vector ξ and let L_γ be defined in (1.3). Thus, $L_\gamma = n\,\mathbf{E}\,\|\xi\|^\gamma$. Set

$$y = \frac{n}{L_\gamma^{2/\gamma}} = n^{1-2/\gamma}\big(\,\mathbf{E}\,\|\xi\|^\gamma\big)^{-2/\gamma}. \tag{2.2}$$

Let us assume that

$$y \geqslant c_1\, d^{\gamma/2}\,(\log^* d)^{\gamma+1}, \tag{2.3}$$

where the quantity c_1 is as large as necessary for the validity of the arguments below. Denote

$$u = \frac{L_\gamma^{2/\gamma}}{\log^2 y} = \frac{n}{y\,\log^2 y} = \frac{n^{2/\gamma}(\,\mathbf{E}\,\|\xi\|^\gamma\,)^{2/\gamma}}{\log^2 y} \tag{2.4}$$

(see (2.2)). By (2.1)–(2.4), we have

$$1 \leqslant \frac{y^{2/(\gamma-2)}}{\log^2 y} \leqslant u \leqslant n/4, \tag{2.5}$$

if c_1 is large enough.

Let us choose an integer $s \geqslant 2$ and a strictly increasing sequence of non-negative integers $m_0 = 0, m_1, \dots, m_s = n$ such that $m_j - m_{j-1} = \lceil u \rceil$, for $j = 1, \dots, s-1$ and $\lceil u \rceil \leqslant m_s - m_{s-1} \leqslant 2\,\lceil u \rceil$. Let ζ_k, $\mathbb{B}_k$ and $\lambda_{k,\gamma}$ be defined in equalities (1.4) and (1.7).

Then

$$\mathbb{B}_k = (m_k - m_{k-1})\,\mathbb{I}_d, \qquad k = 1, \dots, s. \tag{2.6}$$

According to (2.3)–(2.5),

$$s \asymp \frac{n}{\lceil u \rceil} \asymp \frac{n}{u} = y\,\log^2 y \gg d^{\gamma/2}(\log^* d)^{\gamma+3} \tag{2.7}$$

and, hence,

$$\log y \asymp \log s. \tag{2.8}$$

Therefore, by (2.4), (2.7) and (2.8),

$$m_k - m_{k-1} \asymp \lceil u \rceil \asymp u \asymp \frac{n}{s} \asymp \frac{L_\gamma^{2/\gamma}}{\log^2 s}, \qquad k = 1, \dots, s. \tag{2.9}$$

Now, using (2.6) and (2.9), we can find $C_1 = c_2$ and $C_2 = c_3$ such that relations (1.5) and (1.6) are satisfied. Moreover, by (2.9),

$$\lambda_{k,\gamma} = (m_k - m_{k-1})\,\mathbf{E}\,\|\xi\|^\gamma \asymp \frac{n\,\mathbf{E}\,\|\xi\|^\gamma}{s} = \frac{L_\gamma}{s}, \qquad k = 1, \dots, s. \tag{2.10}$$

Set

$$\varepsilon = \frac{1}{3\,\log^* d} \leqslant \frac{1}{3}\,. \tag{2.11}$$

Let us show that

$$d^{\gamma/2}\, s^{\varepsilon-1}\, (\log^* s)^{\gamma+3} \ll 1. \tag{2.12}$$

If $s \geqslant d^\gamma$, then (2.12) is evident. If $s < d^\gamma$, then, by (2.7), (2.8) and (2.11), $\log d \asymp \log s \asymp \log y$, $s^\varepsilon \asymp 1$, and (2.12) follows easily from (2.7). Using (2.10) and (2.12), we obtain

$$d^{\gamma/2}\, s^{\varepsilon}\, (\log^* s)^{\gamma+3}\, L_\gamma^{-1} \max_{1\leqslant k\leqslant s} \lambda_{k,\gamma} \ll 1. \tag{2.13}$$

Condition (1.8) is now satisfied with $C_3 = c_4$. Taking into account (2.11) and applying the statement of Theorem 2, we see that 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 $Y_1, \dots, Y_n$ such that $\mathcal{L}(X_j) = \mathcal{L}(\xi)$, $\mathbf{E}\, Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j = \mathbb{I}_d$, $j = 1, \dots, n$, and

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \ll \big(d^{21/2+\alpha}\,(\log^* d)^2\big)^\gamma\, L_\gamma. \tag{2.14}$$

Consider now the case, where

$$y \leqslant c_1\, d^{\gamma/2}\, (\log^* d)^{\gamma+1}. \tag{2.15}$$

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 $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\, Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j = \mathbb{I}_d$, $j = 1, \dots, n$, the following inequality is valid

$$\mathbf{E} \max_{1\leqslant k\leqslant n} \Big\| \sum_{j=1}^k X_j \Big\|^\gamma \ll n\,\mathbf{E}\,\|\xi\|^\gamma + (n\,d)^{\gamma/2}\,,$$

$$\mathbf{E} \max_{1\leqslant k\leqslant n} \Big\| \sum_{j=1}^k Y_j \Big\|^\gamma \ll (n\,d)^{\gamma/2}\,,$$

and

$$\begin{aligned}\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma &\ll \mathbf{E} \max_{1\leqslant k\leqslant n} \Big\| \sum_{j=1}^k X_j \Big\|^\gamma + \mathbf{E} \max_{1\leqslant k\leqslant n} \Big\| \sum_{j=1}^k Y_j \Big\|^\gamma \\ &\ll n\,\mathbf{E}\,\|\xi\|^\gamma + (n\,d)^{\gamma/2} \\ &\ll d^{\frac{\gamma(\gamma+2)}{4}}\,(\log^* d)^{\frac{\gamma(\gamma+1)}{2}}\, n\,\mathbf{E}\,\|\xi\|^\gamma \end{aligned} \tag{2.16}$$

(we have used Lemma 1 and relations (1.1), (1.16), (2.1), (2.2) and (2.12)). Using (1.11), (2.14) and (2.16), we obtain the statement of Theorem 3. □

Proof of Theorem 4. Without loss of generality, we assume again that $\gamma > 2$ and $\operatorname{cov}\xi = \mathbb{I}_d$. By Theorem 3, there exists a construction such that the joint distribution of $\{X_j\}$ and $\{Y_j\}$ satisfies the relation

$$\mathbf{E}\,\delta_m^\gamma \ll 2^m A\,\mathbf{E}\,\|\xi\|^\gamma, \quad m = 1, 2, \dots, \tag{2.17}$$

where

$$\delta_m = \max_{2^{m-1} < k \leqslant 2^m} \Big\| \sum_{j=2^{m-1}+1}^{k} X_j - \sum_{j=2^{m-1}+1}^{k} Y_j \Big\|. \tag{2.18}$$

Of course, we can assume that the constructions described above are jointly independent for different m's. By the Lyapunov inequality,

$$\mathbf{E}\,\delta_m \leqslant \big(\mathbf{E}\,\delta_m^\gamma\big)^{1/\gamma}. \tag{2.19}$$

By Lemma 2 and by relations (2.17)–(2.19),

$$\begin{aligned}\mathbf{E}\,\big(\Delta_{2^k}(X,Y)\big)^\gamma \leqslant \mathbf{E}\Big(\sum_{m=1}^{k}\delta_m\Big)^\gamma &\ll A\Big(\sum_{m=1}^{k} 2^m\,\mathbf{E}\,\|\xi\|^\gamma + \Big(\sum_{m=1}^{k} 2^{m/\gamma}\Big)^\gamma \mathbf{E}\,\|\xi\|^\gamma\Big)\\ &\ll 2^k A\,\mathbf{E}\,\|\xi\|^\gamma,\end{aligned} \tag{2.20}$$

for $k = 1, 2, \dots$. Thus, the statement of Theorem 4 is proved for $n = 2^k$, $k = 1, 2, \dots$. If $2^{k-1} < n \leqslant 2^k$, then the statement of Theorem 4 follows from already proved inequality (2.20). The corresponding constant should be multiplied not more than by two. □

References

1. Borovkov, A. A., *On the rate of convergence in the invariance principle*, Theor. Probab. Appl. **18** (1973), no. 2, 207–225.
2. de Acosta, A., *Inequalities for B-valued random vectors with applications to the strong law of large numbers*, Ann. Prob. **9** (1981), 157–161.
3. 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.
4. Einmahl, U., *Extensions of results of Komlós, Major and Tusnády to the multivariate case*, J. Multivar. Anal. **28** (1989), 20–68.
5. Götze, F. and Zaitsev, A. Yu., *Bounds for the rate of strong approximation in the multidimensional invariance principle*, Theor. Probab. Appl. **53** (2008), no. 1, 100–123.
6. Johnson, W. B., Schechtman, G. and Zinn, J., *Best constants in moment inequalities for linear combinations of independent and exchangeable random variables.*, Ann. Probab. **13** (1985), 234–253.

7. 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.
8. Lifshits, M. A., *Lectures on the strong approximation*, St. Petersburg University Press, St. Petersburg,, 2007, pp. 1–32. (in Russian)
9. Major, P., *The approximation of partial sums of independent r.v.'s*, Z. Wahrscheinlichkeitstheor. verw. Geb. **35** (1976), 213–220.
10. 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.
11. 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.
12. Sakhanenko, A. I., *Estimates in the invariance principles*, In: Trudy Inst. Mat. SO AN SSSR, vol. 5, Nauka, Novosibirsk, 1985, pp. 27–44. (Russian)
13. Zaitsev, A. Yu., *Rate for the strong approximation in the multidimensional invariance principle*, Zapiski Nauchnykh Seminarov POMI **364** (2009), 141–156 (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