arXiv:1203.5695v1 [math.PR] 26 Mar 2012

ESTIMATES FOR THE RATE OF STRONG APPROXIMATION IN HILBERT SPACE

FRIEDRICH GÖTZE[1] AND ANDREI YU. ZAITSEV[1,2]

1. Introduction

The aim of this paper is to investigate, which infinite dimensional consequences follow from the main results of recently published paper of the authors [10] (see Theorems 2 and 3). We show that the finite dimensional Theorem 3 implies meaningful estimates for the rate of strong Gaussian approximation of sums of i.i.d. Hilbert space valued random vectors ξ_j with finite moments $\mathbf{E}\,\|\xi_j\|^\gamma$, $\gamma > 2$. We show that the rate of approximation depends substantially on the rate of decay of the sequence of eigenvalues of the covariance operator of summands.

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 denote $\log^* b = \max\{1, \log b\}$ for $b > 0$. We shall write $A \ll_t B$, if there exists a positive quantity $c(t)$ depending only on t and such that $A \leqslant c(t)\,B$. We shall also write $A \asymp_t B$, if $A \ll_t B \ll_t A$. The absence of lower indices means that the corresponding constants are absolute.

We consider the following well-known problem. Let $\xi_1, \ldots, \xi_n$ be independent random vectors with zero means and finite moments of second order. One has to construct on the same probability space a sequence of independent random vectors $X_1, \ldots, X_n$ and independent Gaussian random vectors $Y_1, \ldots, Y_n$ such that

$$\mathcal{L}(X_j) = \mathcal{L}(\xi_j), \quad \mathbf{E}\,Y_j = 0, \quad \operatorname{cov} Y_j = \operatorname{cov} X_j, \qquad j = 1, \ldots, n,$$

Date: FMarch 18, 2011.

1991 *Mathematics Subject Classification.* Primary 60F05.

Key words and phrases. Central Limit Theorem, strong approximation convergence rates, Hilbert space, invariance principle .

[1]Research supported by the SFB 701 in Bielefeld and by grant RFBR-DFG 09-01-91331.

[2]Research supported by grant 10-01-00242, by grant of leading Scientific scools 4472.2010.1 and by a program of fundamental researches of Russian Academy of Sciences "Modern problems of fundamental mathematics".

and the quantity

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

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 just to this problem. We omit the detailed history of the problem referring the reader to Götze and Zaitsev [9] and Zaitsev [21].

For brevity, instead of writting out the properties of the vectors $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ listed above we simply say that *there exists a construction* having additional properties which are mentioned explicitly in the text. As a rule, we consider the case where the vectors $\xi_1,\ldots,\xi_n$ are identically distributed with some random vector Z and, in conditions of Theorems, we mention just this vector.

In this paper, we obtain infinite dimensional analogues of the following result of Sakhanenko [18] in the case of i.i.d. summands.

Theorem 1. *Let $\xi_1,\xi_2,\ldots,\xi_n$ be independent random variable with $\mathbf{E}\,\xi_j = 0$, $j=1,\ldots,n$. Let $\gamma > 2$ and*

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

Then there exists a construction such that

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \ll_\gamma L_\gamma. \tag{2}$$

It should be mentioned that, in Sakhanenko [18], one can find more general results. In Sakhanenko [18], it is observed that inequality (2) implies the well-known Rosenthal inequality ([16], [17], see Lemma 1).

Upon the natural normalization, we see that (2) is equivalent to

$$\mathbf{E}\,\big(\Delta_n(X,Y)/\sigma\big)^\gamma \ll_\gamma L_\gamma/\sigma^\gamma,$$

where $\sigma^2 = \mathrm{Var}\big(\sum_{j=1}^n \xi_j\big)$. 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 prove Theorems 4 and 5 which are quite elementary consequences of Theorem 2, proved by the authors in [9] and [10]. In [9], we consider the case of independent and (in general) non-identically distributed summands. Theorem 2 shows what follows from the results of [9] in a particular case, where summands are identically distributed. Theorem 2 is a multidimensional version of Theorem 1 for identically distributed summands.

Denote by $\mathbf{H}$ the separable Hilbert space, which consists of all real sequences $x = (x_1, x_2, \dots)$, for which $\|x\|^2 = x_1^2 + x_2^2 + \dots < \infty$. Also put $\|x\|_\infty = \max_j |x_j|$, $x^{(d)} = (x_1, x_2, \dots, x_d) \in \mathbf{R}^d$ and

$$x^{[d]} = (0, 0, \dots, 0, x_d, x_{d+1}, \dots) \in \mathbf{H}.$$

The formulations of our results involve a random vector

$$Z = (Z_1, Z_2, \dots),$$

taking values in $\mathbf{H}$ or $\mathbf{R}^d$. Independent copies of the vector Z are to be constructed on the same probability space with a sequence of independent Gaussian random vectors. Without loss of generality, we assume, that the coordinates of the vector Z are uncorrelated, and

$$\sigma_1^2 \geqslant \sigma_2^2 \geqslant \dots \geqslant \sigma_m^2 \geqslant \dots, \quad \text{where} \quad \sigma_m^2 = \mathbf{E}\, Z_m^2, \qquad m = 1, 2, \dots, \tag{3}$$

and

$$\mathbb{D} = \operatorname{cov} Z, \quad \mathbb{D}_d = \operatorname{cov} Z^{(d)}, \quad B_d^2 = \sum_{m=d+1}^{\infty} \sigma_m^2 = \mathbf{E} \left\| Z^{[d]} \right\|^2. \tag{4}$$

In particular,

$$B_0^2 = \sum_{m=1}^{\infty} \sigma_m^2 = \mathbf{E} \left\| Z \right\|^2. \tag{5}$$

Moreover, in the formulations of our results, a parameter ψ satisfying

$$21/2 < \psi \leqslant 11. \tag{6}$$

is involved. In the sequel, Many constants below depend on ψ. In order to avoid this complication, one can simply take $\psi = 11$.

Theorem 2. *Let ψ satisfy* (6) *and let Z be an $\mathbf{R}^d$-valued random vector with $\sigma_d^2 > 0$, $\mathbf{E}\, Z = 0$ and $\mathbf{E} \|Z\|^\gamma < \infty$, for some $\gamma \geqslant 2$. Then there exists a construction such that*

$$\mathbf{E} \left(\Delta_n(X, Y) \right)^\gamma \ll_{\psi,\gamma} A \left(\sigma_1 / \sigma_d \right)^\gamma n\, \mathbf{E} \|Z\|^\gamma, \quad \textit{for all } n, \tag{7}$$

where

$$A = A(\gamma, \psi, d) = \max \left\{ d^{\psi\gamma},\ d^{\frac{\gamma(\gamma+2)}{4}} \left(\log^* d \right)^{\frac{\gamma(\gamma+1)}{2}} \right\}. \tag{8}$$

We need a slightly different version of the finite dimensional result. The following statement is proved in [10] while proving Theorem 2.

Theorem 3. *Let ψ satisfy* (6) *and let Z be an $\mathbf{R}^d$-valued random vector with $\sigma_d^2 > 0$, $\mathbf{E}\, Z = 0$ and $\mathbf{E}\, \|Z\|^\gamma < \infty$, for some $\gamma \geqslant 2$. There exists a positive quantity $c_1(\gamma)$ depending only on γ and such that if*

$$C(\gamma)\, d^{\gamma/2} (\log^* d)^{\gamma+1} \left(\mathbf{E} \left\|\mathbb{D}^{-1/2} Z\right\|^\gamma\right)^{2/\gamma} \leqslant n^{1-2/\gamma}, \tag{9}$$

for some positive integer n, then there exists a construction such that

$$\mathbf{E} \left(\Delta_n(\mathbb{D}^{-1/2} X, \mathbb{D}^{-1/2} Y)\right)^\gamma \ll_{\gamma,\psi} d^{\psi\gamma}\, n\, \mathbf{E}\, \|\mathbb{D}^{-1/2} Z\|^\gamma. \tag{10}$$

Remark 1. In [10], the statements of Theorems 2 and 3 involve, for $d^{\psi\gamma}$, the additional logarithmic factor $(\log^* d)^{2\gamma}$. We can easily eliminate it, observing that we allow the constants in (7) and (10) to depend on ψ satisfying (6).

Remark 2. If condition (9) is not satisfied, then the estimates in Theorem 2 are obtained not due to a successful approximation, but by estimating $\mathbf{E} \max_{1\leqslant s\leqslant n} \left\| \sum_{j=1}^s X_j \right\|^\gamma$ and $\mathbf{E} \max_{1\leqslant s\leqslant n} \left\| \sum_{j=1}^s Y_j \right\|^\gamma$ from above with the help of Lemma 1 and inequality (25). Thus, the presence of condition (9) in the formulation of Theorem 3 do not lead to a loss of information on the closeness of distributions in comparison with Theorem 2.

The main results of this paper are Theorems 4 and 5.

Theorem 4. *Let ψ satisfy* (6) *and let Z be an $\mathbf{H}$-valued random vector with $\mathbf{E}\, Z = 0$ and $\mathbf{E}\, \|Z\|^\gamma < \infty$, for some $\gamma > 2$. If, for some fixed positive integers d and n, the inequality*

$$C(\gamma)\, d^{\gamma/2} (\log^* d)^{\gamma+1} \left(\mathbf{E} \left\|\mathbb{D}_d^{-1/2} Z^{(d)}\right\|^\gamma\right)^{2/\gamma} \leqslant n^{1-2/\gamma} \tag{11}$$

is valid, where $C(\gamma)$ is defined in Theorem 3*, then there exists a construction such that*

$$\mathbf{E} \left(\Delta_n(X, Y)\right)^\gamma \ll_{\gamma,\psi} d^{\psi\gamma}\, n\, \sigma_1^\gamma\, \mathbf{E} \left\|\mathbb{D}_d^{-1/2} Z^{(d)}\right\|^\gamma + n\, \mathbf{E} \left\|Z^{[d]}\right\|^\gamma + (n\, B_d^2)^{\gamma/2}. \tag{12}$$

Theorem 5. *Let ψ satisfy* (6) *and let Z be a $\mathbf{H}$-valued random vector with $\mathbf{E}\, Z = 0$ and $\mathbf{E}\, \|Z\|^\gamma < \infty$, for some $\gamma > 2$. If, for some fixed positive integers d and n, the inequality*

$$C(\gamma)\, d^{\gamma/2} (\log^* d)^{\gamma+1} \left(\mathbf{E}\, \|Z\|^\gamma\right)^{2/\gamma} \leqslant n^{1-2/\gamma}\, \sigma_d^2, \tag{13}$$

is valid, where $C(\gamma)$ is defined in Theorem 3*, then there exists a construction such that*

$$\mathbf{E} \left(\Delta_n(X, Y)\right)^\gamma \ll_{\gamma,\psi} d^{\psi\gamma} \left(\sigma_1/\sigma_d\right)^\gamma n\, \mathbf{E}\, \|Z\|^\gamma + (n\, B_d^2)^{\gamma/2}. \tag{14}$$

Theorems 4 and 5 make it possible to obtain meaningful infinite dimensional estimates by a suitable choice of dimension d, for which the summands in the right-hand side of inequality (12) have approximately the same order in n. Theorem 5 is an elementary consequence of Theorem 4 and the inequality

$$\mathbf{E}\left\|\mathbb{D}_d^{-1/2}Z^{(d)}\right\|^\gamma \leqslant \sigma_d^{-\gamma}\,\mathbf{E}\left\|Z^{(d)}\right\|^\gamma \leqslant \sigma_d^{-\gamma}\,\mathbf{E}\,\|Z\|^\gamma. \tag{15}$$

In general, Theorem 4 is sharper than Theorem 5. Many distributions with a regular behavior of moments satisfy the relation

$$K = \sup_{1\leqslant d<\infty} d^{-\gamma/2}\,\mathbf{E}\,\|\mathbb{D}_d^{-1/2}Z^{(d)}\|^\gamma < \infty, \tag{16}$$

which may lead to a substantial improvement of the order of estimates. For instance, if the vector Z has independent coordinates Z_m, then, by Lemma 2 of Section 2,

$$\mathbf{E}\,\|\mathbb{D}_d^{-1/2}Z^{(d)}\|^\gamma = \mathbf{E}\left(\sum_{m=1}^d \frac{Z_m^2}{\sigma_m^2}\right)^{\gamma/2} \ll_\gamma d^{\gamma/2} + \sum_{m=1}^d \sigma_m^{-\gamma}\,\mathbf{E}\,|Z_m|^\gamma. \tag{17}$$

Hence, $K<\infty$, if the sequence of moments $\sigma_m^{-\gamma}\,\mathbf{E}\,|Z_m|^\gamma$ is bounded or grows not faster than $O(m^{(\gamma-2)/2})$. Observe that Lyapunov's inequality yields $\mathbf{E}\,\|\mathbb{D}_d^{-1/2}Z^{(d)}\|^\gamma \geqslant d^{\gamma/2}$.

On the other hand, in the general case, the application of (15) may not lead to a loss of precision, while the statement of Theorem 5 is simpler than that of Theorem 4. It involves only the moment $\mathbf{E}\,\|Z\|^\gamma$ and the eigenvalues of the covariance operator of $\mathbb{D}$ of the vector Z. An intermediate situation is possible, where inequality (16) is not valid, but the statement of Theorem 4 is still stronger than that of Theorem 5.

The proofs of Theorems 4 and 5 are based on the method of finite dimensional approximation, related to the method applied for estimating the accuracy of approximation in the CLT in infinite dimensional spaces (see., for instance, the survey [1]).

Applying Chebyshev's inequality, we see that, under the assumptions of Theorem 2, we have

$$\mathbf{P}\big\{\Delta_n(X,Y)\geqslant x\big\} \ll_{\gamma,d} \big(\sigma_1/\sigma_d\big)^\gamma\, n\,\mathbf{E}\,\|Z\|^\gamma/x^\gamma \tag{18}$$

for all $x>0$ and all $n=1,2,\ldots$. Clearly, the statement of Theorem 2 is stronger than (18). A construction for which (18) is valid for $d=1$ for fixed n and $x=O\big(\sqrt{n}\,\log n\big)$ with constants, depending on γ and $\mathcal{L}(Z)$ only, was proposed by Komlós, Major, and Tusnády (KMT) [13], see also Borovkov [3] and Major [14] in the case $2<\gamma\leqslant 3$. Then Sakhanenko [18] proved Theorem 1, which ensures the validity of the

one-dimensional version of inequality (18) for all x on the same probability space. Einmahl [8] obtained a multidimensional version of the KMT result without restrictions on the values of x.

Previously, the estimates for the rate of strong approximation in infinite dimensional spaces appeared, for example, in [4], [5], [19] and [20]. The closest to the subject of this paper is the following infinite dimensional result of Sakhanenko [20].

Theorem 6. *Let Z be an $\mathbf{H}$-valued random vector with $\mathbf{E}\,Z = 0$ and $\mathbf{E}\,\|Z\|^\gamma < \infty$, for some γ with $2 \leqslant \gamma \leqslant 3$. Then, for any fixed $x > 0$, there exists a construction such that*

$$\mathbf{P}\big\{\Delta_n^\infty(X,Y) \geqslant x\big\} \ll n\,\mathbf{E}\,\|Z\|^\gamma/x^\gamma \quad \textit{for all } n, \tag{19}$$

where

$$\Delta_n^\infty(X,Y) = \max_{1\leqslant s\leqslant n} \Big\| \sum_{j=1}^s X_j - \sum_{j=1}^s Y_j \Big\|_\infty. \tag{20}$$

Theorem 6 is formulated for fixed x. This means that the probability space depends on this x. Furthermore, in the statement of Theorem 6, the quantity $\Delta_n(X,Y)$ is replaced by $\Delta_n^\infty(X,Y)$, which is (in general) essentially smaller than $\Delta_n(X,Y)$. On the other hand, inequality (19) looks almost as inequality (18) for $2 \leqslant \gamma \leqslant 3$. We should note that Sakhanenko [20] obtained substantially more general results in comparison with Theorem 6. They are proved for non-identically distributed depending summands, forming, for example, infinite dimensional martingales.

The following theorem yields a lower bound under the assumptions of Theorems 4 and 5.

Theorem 7. *Let positive numbers σ_m^2, $m = 1, 2, \dots$, satisfy the relations*

$$\sigma_1^2 \geqslant \sigma_2^2 \geqslant \dots \geqslant \sigma_m^2 \geqslant \dots, \qquad \sum_{m=1}^\infty \sigma_m^2 < \infty. \tag{21}$$

Let n be a fixed positive integer, and $\lambda > 0$ with $\sigma_1^2 \leqslant \lambda^2$. Denote

$$k = \min\{m : n\,\sigma_m^2 < \lambda^2\} - 1. \tag{22}$$

Then there exists an $\mathbf{H}$-valued random vector $Z = (Z_1, Z_2, \dots)$, satisfying (3)–(5) and such that $\mathbf{E}\,\|Z\|^\gamma < \infty$, for all $\gamma \geqslant 0$, and for any construction we have the lower bound

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \gg_\gamma \mathbf{E}\,\big(\Delta_n(X^{(k)},Y^{(k)})\big)^\gamma + (nB_k^2)^{\gamma/2}. \tag{23}$$

Meanwhile, the first term in the right-hand side of (23) is assumed to be zero if $k = 0$.

Remark 3. Finding a lower bound for $\mathbf{E}\left(\Delta_n(X^{(k)}, Y^{(k)})\right)^\gamma$ is a separate problem. Note, however, that the vector Z from the proof of Theorem 7 satisfies the rough bound $\mathbf{E}\left(\Delta_n(X^{(k)}, Y^{(k)})\right)^\gamma \gg_\gamma (\lambda^2 k)^{\gamma/2}$, since it has a lattice distribution.

The presence of the quantity $(nB_k^2)^{\gamma/2}$ in the right-hand side of (23) confirms that the appearance of the summand $(nB_d^2)^{\gamma/2}$ in (12) and (14) is natural. It becomes clear when we compare inequality (23) with the intermediate inequality (30).

In Section 3, we consider Examples 1–4, showing, in particular, that for many distributions Theorem 5 yields estimates, which are stronger than the estimates of Theorem 6. Moreover, in Example 5, we verify that, if the sequence of eigenvalues σ_m^2 decreases slowly, then Theorems 4 and 5 provide estimates which are optimal in order.

2. Proofs

We shall 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 $\mathbf{H}$. Then*

$$\mathbf{E}\left\|\sum_{j=1}^n \xi_j\right\|^\gamma \ll_\gamma \sum_{j=1}^n \mathbf{E}\,\|\xi_j\|^\gamma + \Big(\sum_{j=1}^n \mathbf{E}\,\|\xi_j\|^2\Big)^{\gamma/2}, \quad \textit{for } \gamma \geqslant 2. \tag{24}$$

This multidimensional version of the Rosenthal inequality follows easily from a result of de Acosta [6]. In the i.i.d. case, the second summand in the right-hand side of (24) grows faster than the first term as $n \to \infty$. Theorems 1 and 2 show that this growth corresponds to the growth of moments of sums of Gaussian approximating vectors.

The next lemma is proved by Rosenthal [16], see also Johnson, Schechtman and Zinn [12].

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

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

The following Lemma 3 is proved by Montgomery-Smith [15]. It is a particular case of Theorem 1.1.5 from the monograph of de la Peña and Giné [7].

Lemma 3. *Let $\xi_1, \dots, \xi_n$ be i.i.d. random vectors with values in $\mathbf{H}$. Then*

$$\mathbf{P}\Big\{\max_{1\leqslant s\leqslant n}\Big\|\sum_{j=1}^{s}\xi_j\Big\| > x\Big\} \leqslant 9\,\mathbf{P}\Big\{\Big\|\sum_{j=1}^{n}\xi_j\Big\| > x/30\Big\} \quad \textit{for all } x \geqslant 0.$$

Coupled 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 us to estimate the moments

$$\mathbf{E}\,\max_{1\leqslant s\leqslant n}\Big\|\sum_{j=1}^{s}\xi_j\Big\|^\gamma \ll_\gamma \mathbf{E}\,\Big\|\sum_{j=1}^{n}\xi_j\Big\|^\gamma, \quad \gamma > 0, \tag{25}$$

in the case of i.i.d. random vectors $\xi_1, \dots, \xi_n$.

Proof of Theorem 4. It is not difficult to understand that for any construction we have

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \ll_\gamma \mathbf{E}\,\big(\Delta_n(X^{(d)},Y^{(d)})\big)^\gamma + \mathbf{E}\,\max_{1\leqslant s\leqslant n}\Big\|\sum_{j=1}^{s}X_j^{[d]}\Big\|^\gamma + \mathbf{E}\,\max_{1\leqslant s\leqslant n}\Big\|\sum_{j=1}^{s}Y_j^{[d]}\Big\|^\gamma. \tag{26}$$

Using (25) and (26), we obtain

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \ll_\gamma \mathbf{E}\,\big(\Delta_n(X^{(d)},Y^{(d)})\big)^\gamma + \mathbf{E}\,\Big\|\sum_{j=1}^{n}X_j^{[d]}\Big\|^\gamma + \mathbf{E}\,\Big\|\sum_{j=1}^{n}Y_j^{[d]}\Big\|^\gamma. \tag{27}$$

Lemma 1 together with $\mathcal{L}(X_j) = \mathcal{L}(Z)$ and $\operatorname{cov} Y_j = \operatorname{cov} X_j = \operatorname{cov} Z$ yields

$$\mathbf{E}\,\Big\|\sum_{j=1}^{n}X_j^{[d]}\Big\|^\gamma \ll_\gamma n\,\mathbf{E}\,\big\|Z^{[d]}\big\|^\gamma + \big(n\,\mathbf{E}\,\big\|Z^{[d]}\big\|^2\big)^{\gamma/2}, \tag{28}$$

and

$$\mathbf{E}\,\Big\|\sum_{j=1}^{n}Y_j^{[d]}\Big\|^\gamma \ll_\gamma \big(n\,\mathbf{E}\,\big\|Z^{[d]}\big\|^2\big)^{\gamma/2}. \tag{29}$$

Inequalities (27)–(29) imply that

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \ll_\gamma \mathbf{E}\,\big(\Delta_n(X^{(d)},Y^{(d)})\big)^\gamma + n\,\mathbf{E}\,\big\|Z^{[d]}\big\|^\gamma + (n\,B_d^2)^{\gamma/2}. \tag{30}$$

It is easy to show that condition (11) implies that the d-dimensional vector $Z^{(d)}$ satisfies condition (9) of Theorem 3. Applying that theorem, we see that from (10) and from the well-known Berkes–Philipp

Lemma [2] it follows, that there exists a construction such that

$$\mathbf{E}\left(\Delta_n(X^{(d)}, Y^{(d)})\right)^\gamma \ll_{\gamma,\psi} d^{\psi\gamma}\, n\, \sigma_1^\gamma\, \mathbf{E}\left\|\mathbb{D}_d^{-1/2} Z^{(d)}\right\|^\gamma. \tag{31}$$

Using (30) and (31), we obtain the statement of Theorem 4.

Proof of Theorem 7. Let Z_j (coordinates of the vector Z) be independent random variables, taking values $-\lambda$, 0 and λ with probabilities

$$\mathbf{P}\{Z_m = \pm\lambda\} = \sigma_m^2/2\lambda^2, \quad \mathbf{P}\{Z_m = 0\} = 1 - \sigma_m^2/\lambda^2, \quad m = 1, 2, \ldots. \tag{32}$$

With the help of Lemma 2 it is not difficult to show that $\mathbf{E}\,\|Z\|^\gamma < \infty$, for all $\gamma \geqslant 0$. Assume that we have constructed a sequence of independent random vectors $X_1, \ldots, X_n$ and a corresponding sequence of independent Gaussian random vectors $Y_1, \ldots, Y_n$ such that $\mathcal{L}(X_j) = \mathcal{L}(Z)$, $\mathbf{E}\, Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, \ldots, n$.

Then the coordinates of the vectors X_j (namely $\{X_{jm},\, j = 1, 2, \ldots, n,\, m = 1, 2, \ldots\}$) are jointly independent random variables with distributions $\mathcal{L}(Z_m)$, while the coordinates of the vectors Y_j (namely $\{Y_{jm},\, j = 1, 2, \ldots, n,\, m = 1, 2, \ldots\}$) are jointly independent Gaussian random variables with mean zero and variances σ_m^2. Set

$$S_{nm} = \sum_{j=1}^n X_{jm}, \quad T_{nm} = \sum_{j=1}^n Y_{jm}, \quad m = 1, 2, \ldots. \tag{33}$$

It is clear that $\operatorname{Var} S_{nm} = \operatorname{Var} T_{nm} = n\,\sigma_m^2$, for $m = 1, 2, \ldots$, and

$$\Delta_n(X, Y) \geqslant \max\left\{\Delta_n(X^{(k)}, Y^{(k)}),\ \Delta_n(X^{[k]}, Y^{[k]})\right\}. \tag{34}$$

Obviously,

$$\Delta_n(X^{[k]}, Y^{[k]}) \geqslant \Big\| \sum_{j=1}^n X_j^{[k]} - \sum_{j=1}^n Y_j^{[k]} \Big\|, \tag{35}$$

while

$$\Big\| \sum_{j=1}^n X_j^{[k]} - \sum_{j=1}^n Y_j^{[k]} \Big\|^2 = \sum_{m=k+1}^\infty |S_{nm} - T_{nm}|^2. \tag{36}$$

If $m > k$, then

$$|S_{nm} - T_{nm}| \geqslant \eta_{nm}, \quad \text{where } \eta_{nm} = |T_{nm}|\ \mathbf{1}\{|T_{nm}| \leqslant \lambda/2\}, \tag{37}$$

since the random variables S_{nm} take only values which are multiples of λ. Put

$$U_{nk} = \sum_{m=k+1}^\infty \eta_{nm}^2. \tag{38}$$

For fixed n, the set $\{\eta_{nm}\}$ is a collection of jointly independent random variables. According to (22), (33) and (37), for $m > k$,

$$\mathbf{E}\,(\eta_{nm}^2) \asymp n\,\sigma_m^2 \quad \text{and} \quad \operatorname{Var}(\eta_{nm}^2) \asymp n^2\sigma_m^4. \tag{39}$$

Denote $a = \mathbf{E}\,U_{nk}$ and $b = \operatorname{Var} U_{nk}$. Note that by relations (22), (37), (38) and (39),

$$a = \sum_{m=k+1}^{\infty} \mathbf{E}\,(\eta_{nm}^2) \asymp n\,B_k^2 \quad \text{and} \quad b = \sum_{m=k+1}^{\infty} \operatorname{Var}(\eta_{nm}^2) \ll a^2, \tag{40}$$

where the quantity B_k^2 is defined by formula (4). According to inequality (7.5) from Feller [11], p. 180,

$$\mathbf{P}\big\{U_{nk} - a < -t\big\} \leqslant \frac{b}{b+t^2} = 1 - \frac{t^2}{b+t^2}, \quad \text{for all } t \geqslant 0. \tag{41}$$

Applying (41) for $t = a/2$ and relations (40), it is easy to show that

$$\mathbf{P}\big\{U_{nk} \geqslant a/2\big\} \gg 1. \tag{42}$$

Therefore, relations (37), (38) and (42) yield

$$\mathbf{P}\Big\{\sum_{m=k+1}^{\infty} |S_{nm} - T_{nm}|^2 \geqslant a/2\Big\} \gg 1. \tag{43}$$

¿From (35), (36) and (43) we obtain

$$\mathbf{P}\big\{\big(\Delta_n(X^{[k]}, Y^{[k]})\big)^2 \geqslant a/2\big\} \gg 1 \tag{44}$$

and, hence,

$$\mathbf{E}\,\big(\Delta_n(X^{[k]}, Y^{[k]})\big)^\gamma \gg_\gamma a^{\gamma/2} \asymp_\gamma (nB_k^2)^{\gamma/2}. \tag{45}$$

Finally (34) and (45) imply the lower bound (23).

3. Examples

In Examples 1–5 we compare the estimates which follows from Theorem 4 when condition (16) is satisfied to bounds of Theorem 5 for concrete sequences of eigenvalues of the covariance operator of the vector Z.

Example 1. Let $\sigma_m^2 = \exp\{-\alpha\, m^\beta\}$, $m = 1, 2, \ldots$, where $\alpha, \beta > 0$. Assume that n is so large that

$$d = \max\{m : n^{2/\gamma}\,(\log^* n)^{2\psi/\beta}/\sigma_m^2 \ < n\sigma_m^2\} \geqslant 1. \tag{46}$$

Then it is clear that

$$d \asymp_{\alpha,\beta} (\log^* n)^{1/\beta} \tag{47}$$

and

$$\sigma_{d+1}^4 \leqslant n^{-1+2/\gamma}\,(\log^* n)^{2\psi/\beta} \leqslant \sigma_d^4. \tag{48}$$

Thus, for sufficiently large n, the right-hand side of inequality (14) admits the upper bound

$$d^{\psi\gamma}\big(\sigma_1/\sigma_d\big)^{\gamma}\, n\,\mathbf{E}\,\|Z\|^{\gamma} + n^{\gamma/2}B_d^{\gamma} \ll_{\alpha,\beta,\gamma} n^{(2+\gamma)/4}\,(\log^* n)^{\psi\gamma/2\beta}\,\mathbf{E}\,\|Z\|^{\gamma}. \tag{49}$$

Using relations (47) and (48), it is not difficult to verify that, for sufficiently large n, condition (13) is satisfied and, hence, the statement of Theorem 5 is valid with the estimate

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\alpha,\beta,\gamma,\psi} n^{(2+\gamma)/4}\,(\log^* n)^{\psi\gamma/2\beta}\,\mathbf{E}\,\|Z\|^{\gamma}. \tag{50}$$

The right-hand side of inequality (50) grows slower than $n^{\gamma/2}$ (the order of the trivial estimate which follows from Lemma 1 and inequality (25)). Therefore, inequality (50) is a meaningful estimate of the rate of approximation in the infinite dimensional invariance principle. In particular, using Lyapunov's inequality $\mathbf{E}\,\Delta^3 \leqslant (\mathbf{E}\,\Delta^{\gamma})^{3/\gamma}$, we obtain that, for $\gamma > 3$,

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{3} \ll_{\alpha,\beta,\gamma,\psi} n^{3(2+\gamma)/4\gamma}\,(\mathbf{E}\,\|Z\|^{\gamma})^{3/\gamma}. \tag{51}$$

For $\gamma > 6$, the order of inequality (51) with respect to n is better than the order of estimate (19).

Example 2. Suppose now, under the assumptions of Example 1, that (16) holds. Assume that n is so large that

$$d = \min\{m : nB_m^2 < 1\} \geqslant 1. \tag{52}$$

It is clear that then relation (47) is still satisfied. Thus, for sufficiently large n, the right-hand side of inequality (12) admits the upper bound

$$\begin{aligned} d^{\psi\gamma}\, n\,\sigma_1^{\gamma}\,\mathbf{E}\,\big\|\mathbb{D}_d^{-1/2}Z^{(d)}\big\|^{\gamma} \;+\; & n\,\mathbf{E}\,\big\|Z^{[d]}\big\|^{\gamma} + (n\,B_d^2)^{\gamma/2} \\ & \ll_{\alpha,\beta,\gamma,K} \quad n\,(\log^* n)^{(2\psi+1)\gamma/2\beta}\,\mathbf{E}\,\|Z\|^{\gamma}. \end{aligned} \tag{53}$$

Using relation (47), it is also not difficult to verify that, for sufficiently large n, condition (11) is satisfied and, hence, the statement of Theorem 4 is valid with the estimate

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\alpha,\beta,\gamma,\psi,K} n\,(\log^* n)^{(2\psi+1)\gamma/2\beta}\,\mathbf{E}\,\|Z\|^{\gamma}, \tag{54}$$

which is considerably stronger than (50) and is close to the finite-dimensioval estimate (7) of Theorem 2.

Example 3. Let $\sigma_m^2 = m^{-b}$, $m = 1, 2, \ldots$, where $b > 1$. Choose

$$d = \max\{m : n^{2/\gamma}\, m^{2\psi}/\sigma_m^2 \,< nm\sigma_m^2\}. \tag{55}$$

It is clear that then $d \geqslant 1$ and

$$d^{b-1} \asymp_b n^{r(\gamma-2)/\gamma}, \quad \text{where } r = \frac{b-1}{2b-1+2\psi}. \tag{56}$$

Therefore, if (13) is fulfilled, then the right-hand side of inequality (14) admits the upper bound

$$d^{\psi\gamma}\big(\sigma_1/\sigma_d\big)^{\gamma}\, n\,\mathbf{E}\,\|Z\|^{\gamma} + n^{\gamma/2}B_d^{\gamma} \ll_{b,\gamma} n^{(\gamma - r(\gamma-2))/2}\,\mathbf{E}\,\|Z\|^{\gamma}. \tag{57}$$

Using (56), it is not difficult to verify that, for sufficiently large n, condition (13) is satisfied provided that $\gamma < 2\,(b-1+2\psi)$. In this case the statement of Theorem 5 is valid with the estimate

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\psi,b,\gamma} n^{(\gamma - r(\gamma-2))/2}\,\mathbf{E}\,\|Z\|^{\gamma}. \tag{58}$$

Using Lyapunov's inequality by analogy with Example 1, we obtain that for $\gamma > 3$

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{3} \ll_{\psi,b,\gamma} n^{3(\gamma - r(\gamma-2))/2\gamma}\,(\mathbf{E}\,\|Z\|^{\gamma})^{3/\gamma}. \tag{59}$$

For $3r-1>0$ and $\gamma > 6r/(3r-1)$, the order of inequality (59) with respect to n is better than the order of estimate (19).

If condition (13) is not fulfilled for d defined by (55), one should decrease d choosing

$$d = \max\big\{m : C(\gamma)\, m^{\gamma/2}(\log^* m)^{\gamma+1}\,\big(\mathbf{E}\,\|Z\|^{\gamma}\big)^{2/\gamma} \leqslant n^{1-2/\gamma}\,\sigma_m^2\big\}. \tag{60}$$

It is clear that then, for sufficiently large n, we have

$$d^{b-1} \asymp_{b,\gamma,\lambda} n^{\delta(\gamma-2)/\gamma}(\log^* n)^{-\delta(\gamma+1)}, \qquad \delta = \frac{2(b-1)}{2b+\gamma},\ \lambda = \mathbf{E}\,\|Z\|^{\gamma}. \tag{61}$$

In this case the statement of Theorem 5 is valid with the estimate

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\psi,b,\gamma,\lambda} n^{(\gamma-\delta(\gamma-2))/2}\,(\log^* n)^{\delta\gamma(\gamma+1)/2}, \tag{62}$$

which must be weaker in order in comparison with (58). Thus, in the general case, for sufficiently large n, there exists a construction such that

$$\begin{aligned}&\mathbf{E}\,\big(\Delta_n(X,Y)\big)^{\gamma}\\ &\qquad \ll_{\psi,b,\gamma,\lambda} \max\big\{n^{(\gamma-\delta(\gamma-2))/2}\,(\log^* n)^{\delta\gamma(\gamma+1)/2},\, n^{(\gamma-r(\gamma-2))/2}\big\}.\end{aligned} \tag{63}$$

Example 4. Suppose now, under the assumptions of Example 3, that (16) holds. Choose

$$d = \min\{m : n^{2/\gamma}\, m^{2\psi+1} < nm\sigma_m^2\}. \tag{64}$$

It is clear that then $d \geqslant 1$ and

$$d^{b-1} \asymp_b n^{\rho(\gamma-2)/\gamma}, \quad \text{where } \rho = \frac{b-1}{b+2\psi}. \tag{65}$$

Therefore, if (11) is fulfilled, then the right-hand side of inequality (12) admits the upper bound

$$d^{\psi\gamma}\, n\, \sigma_1^{\gamma}\, \mathbf{E}\, \big\|\mathbb{D}_d^{-1/2} Z^{(d)}\big\|^{\gamma} + \; n\, \mathbf{E}\, \big\|Z^{[d]}\big\|^{\gamma} + (n\, B_d^2)^{\gamma/2} \\ \ll_{b,\gamma,K} n^{(\gamma-\rho(\gamma-2))/2}\, \mathbf{E}\, \|Z\|^{\gamma}. \tag{66}$$

Using (65), it is also not difficult to verify that, for sufficiently large n, condition (11) is satisfied provided that $\gamma < 2\,(b-1+2\psi)$. In this case the statement of Theorem 4 is valid with the estimate

$$\mathbf{E}\, \big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\psi,b,\gamma,K} n^{(\gamma-\rho(\gamma-2))/2}\, \mathbf{E}\, \|Z\|^{\gamma}. \tag{67}$$

If condition (11) is not fulfilled, then one should choose d for applying Theorem 4 not by formula (64), but by relation

$$d = \max\big\{m : C(\gamma)\, K^{2/\gamma}\, m^{1+\gamma/2} (\log^* m)^{\gamma+1} \leqslant n^{1-2/\gamma}\big\}. \tag{68}$$

It is clear that, for sufficiently large n,

$$d^{b-1} \asymp_{b,\gamma,K} n^{\mu(\gamma-2)/\gamma} (\log^* n)^{-\mu(\gamma+1)}, \quad \text{where } \mu = \frac{2(b-1)}{2+\gamma}. \tag{69}$$

In this case the statement of Theorem 4 is valid with the estimate

$$\mathbf{E}\, \big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\psi,b,\gamma,K} n^{(\gamma-\mu(\gamma-2))/2}\, (\log^* n)^{\mu\gamma(\gamma+1)/2}, \tag{70}$$

which is weaker in order in comparison with (67). In the general case, for sufficiently large n, there exists a construction such that

$$\mathbf{E}\, \big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\psi,b,\gamma,K} \\ \max\big\{n^{(\gamma-\mu(\gamma-2))/2}\, (\log^* n)^{\mu\gamma(\gamma+1)/2},\; n^{(\gamma-\rho(\gamma-2))/2}\big\}\, \mathbf{E}\, \|Z\|^{\gamma}. \tag{71}$$

Example 5. Let $\sigma_m^2 = 1/m\, (\log^* m)^{1+\tau}$, for $m = 1, 2, \ldots$, where $\tau > 0$. Denote by $\lceil x \rceil$ the integer part of a number x. Choose

$$d = \lceil n^{\varepsilon} \rceil, \quad \text{where } \varepsilon = \frac{\gamma-2}{\gamma\,(\gamma+22)}. \tag{72}$$

It is clear that then

$$B_d^2 \asymp_{\tau} \frac{1}{(\log^* d)^{\tau}} \asymp_{\gamma,\tau} \frac{1}{(\log^* n)^{\tau}}. \tag{73}$$

Using relations (72) and (73), it is not difficult to verify that, for sufficiently large n, condition (13) is satisfied and the statement of Theorem 5 is valid with the estimate

$$\mathbf{E}\, \big(\Delta_n(X,Y)\big)^{\gamma} \ll_{\gamma,\tau} (n/(\log^* n)^{\tau})^{\gamma/2}. \tag{74}$$

Let us compare the upper bounds obtained in Examples 1, 3 and 5 using Theorem 5 and the lower bound

$$\mathbf{E}\left(\Delta_n(X,Y)\right)^\gamma \gg_\gamma (nB_k^2)^{\gamma/2}, \tag{75}$$

which follows from (23).

In Example 1 the lower bound (75) is far from the upper bound (50). It is not difficult to calculate that, in Example 3, the positive integer k, defined by (22), satisfies the relations $k \asymp_{b,\lambda} n^{1/b}$, $B_k^2 \asymp_{b,\lambda} n^{(1-b)/b}$, while the lower bound (75) is of order $O\big(n^{\gamma/2b}\big)$. This shows, that the order of upper bounds should be expected to grow when γ increases. Notice that, for large values of γ and b, the order of upper bounds is close to $n^{\gamma/4}$.

For relatively small values of γ and b, the orders of estimates depend essentially on ψ, which is involved in the bounds due to sufficiently large powers of dimension d in the estimates of Theorems 2 and 3. Possible improvements of Theorems 2 and 3 have to improve the order of the upper bounds in Examples 1–4 and in Theorems 4 and 5.

In Example 5, it is easy to verify that $k \asymp_\tau n/(\log^* n)^{1+\tau}$. Thus, the upper and lower bounds are of the same order $O\big((n/(\log^* n)^\tau)^{\gamma/2}\big)$, and Theorem 5 provides the correct order for the rate of approximation. The same is true if the variances of coordinates σ_m^2 are decreasing slower than in Example 5. Therefore, the order of estimates could be made arbitrarily close to the trivial order $O\big(n^{\gamma/2}\big)$.

The authors are grateful to a referee for a series of useful remarks which enable us to improve exposition substantially.

References

[1] Bentkus, V., Götze, F., Paulauskas, V. and Račkauskas, A. The accuracy of Gaussian approximation in Banach spaces Preprint 90-100 SFB 343, Universität Bielefeld, 1990 (In Russian: Itogi Nauki i Tehniki, ser. Sovr. Probl. Matem., Moskva, VINITI, 1991, v. 81, 39–139).

[2] Berkes I., Philipp W. Approximation theorems for independent and weakly dependent random vectors. Ann. Probab. 1979. v. 7. no. 1, 29–54.

[3] Borovkov A.A. On the rate of convergence in the invariance principle. Theor. Probab. Appl., 1973, v. 18, no. 2, 207–225.

[4] Borovkov A.A., Sakhanenko, A.I. On the estimates for the rate of convergence in the invariance principle for Banach spaces. Theor. Probab. Appl., 1980, v. 25, no. 2, 734–744.

[5] Borovkov A.A. On the rate of convergence in the invariance principle in the Hilbert space. Theor. Probab. Appl., 1984, v. 29, 532–535.

[6] de Acosta A. Inequalities for B-valued random vectors with applications to the strong law of large numbers. Ann. Probab. 1981. v. 9, 157–161.

[7] de la Peña V. H., 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.
[8] Einmahl U. Extensions of results of Komlós, Major and Tusnády to the multivariate case. J. Multivar. Anal., 1989, v. 28, 20–68.
[9] Götze F., Zaitsev A.Yu. Bounds for the rate of strong approximation in the multidimensional invariance principle. Theory Probab. Appl., 2008, v. 53, no. 1, 100–123.
[10] Götze F., Zaitsev A.Yu. Rate of approximation in the multidimensional invariance principle for sums of independent identically distributed random vectors with finite moments. Zapiski Nauchnykh Seminarov POMI, 2009, v. 368, 110–121.
[11] Feller W. An introduction to probability theory and its applications. Vol. 2 (Russian tramslation), Mir, Moskva, 1984.
[12] Johnson W.B., Schechtman G., Zinn J. Best constants in moment inequalities for linear combinations of independent and exchangeable random variables. Ann. Probab., 1985, v. 13, 234–253.
[13] Komlós J., Major P., Tusnády G. An approximation of partial sums of independent RV'-s and the sample DF. I; II. Z. Wahrscheinlichkeitstheor. verw. Geb. 1975, v. 32, 111–131; 1976, v. 34, 34–58.
[14] Major P. The approximation of partial sums of independent r.v.'s. Z. Wahrscheinlichkeitstheor. verw. Geb. 1976, v. 35, 213–220.
[15] Montgomery-Smith S.J. Comparison of sums of independent identically distributed random vectors. Probab. Math. Statist., 1993, v. 14., no. 2, 281–285.
[16] Rosenthal H.P. On the subspaces of L_p $(p > 2)$ spanned by sequences of independent random variables. Israel J. Math., 1970, v. 8, 273–303.
[17] Rosenthal H.P. On the span in L^p of sequences of independent random variables. II. In: 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.
[18] Sakhanenko, A.I. Estimates in the invariance principles. In: Trudy Inst. Mat. SO AN SSSR, v.. 5, Nauka, Novosibirsk, 1985, pp. 27–44.
[19] Sakhanenko A. I. Simple method of obtaining estimates in the invariance principle. Lecture Notes Math., 1987,. v. 1299, 430–443.
[20] Sakhanenko A.I. A new way to obtain estimates in the invariance principle. In: High dimensional probability, II (Seattle, WA, 1999. Progr. Probab., v. 47, Birkhäuser Boston, Boston, MA, 2000, pp. 223–245.
[21] Zaitsev, A.Yu. Rate for the strong approximation in the multidimensional invariance principle. Zapiski Nauchnykh Seminarov POMI, 2009, v. 364, 141–156.

Fakultät für Mathematik,Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany
E-mail address: goetze@math.uni-bielefeld.de

St. Petersburg Department of Steklov Mathematical Institute,, Fontanka 27, St. Petersburg 191023, Russia
E-mail address: zaitsevpdmi.ras.ru