Strong approximation for sums of independent random vectors

Andrei Yu. Zaitsev[1]

Steklov Institute, St Petersbourg, RUSSIA.

2011, 12–13 April

Let $X, X_1, X_2, \ldots$ a sequence of i.i.d. random variables such that $\mathbf{E}\,|X|^2 < \infty$, $\mathbf{E}\,X = 0$. Denote

$$S_k = \sum_{j=1}^{k} X_j, \quad k = 1, 2, \ldots.$$

Let $\xi_n(t)$, $t \in [0, 1]$, be a random continuous broken line defined as

$$\xi_n(k/n) = S_k/\sqrt{n}, \quad k = 1, 2, \ldots, n,$$

and by linear interpolation for $t \in [(k-1)/n, k/n]$. By the Donsker–Prokhorov invariance principle, the distribution of $\xi_n(\,\cdot\,)$ converges weakly (as $n \to \infty$) to the distribution of the standard Brownian motion $W(t)$ in the space $C([0, 1])$.

The rate of strong approximation in the invariance principle for sums of independent random vectors is estimated in two different but closely connected situations. The estimation of the rate of strong approximation in the probability invariance principle may be reduced to these problems.

(A) One have to construct on a probability space a sequence of i.i.d. random vectors $X, X_1, X_2, \ldots$ (with given distribution, $\mathbf{E}\, \|X\|^2 < \infty$, $\mathbf{E}\, X = 0$) and a sequence of independent Gaussian random vectors $Y_1, Y_2, \ldots$ such that

$$\mathcal{L}(X_j) = \mathcal{L}(X), \quad \mathbf{E}\, Y_j = 0, \quad \operatorname{cov} Y_j = \operatorname{cov} X, \quad \text{for } j = 1, 2, \ldots, n, \tag{1}$$

and

$$\Big\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \Big\| = O(f(n)) \quad \text{or } o(f(n))$$

almost surely, for a sequence $f(n)$ tending to infinity as slow as possible.

$d = 1$: Strassen (1964, 1967), Breiman (1967),

Csörgő and Révész (1975),

Komlós, Major and Tusnády (KMT) (1975),

Major and (1976. 1978, 1979),

Einmahl (1987, 1989, 2009), Einmahl and Mason (1993).

$d > 1$: U. Einmahl (1987, 1989, 2009)

(B) One have to construct on a probability space a sequence of i.i.d. random vectors $X, X_1, \ldots, X_n$ (with given distributions) and a sequence of independent Gaussian random vectors $Y_1, \ldots, Y_n$ so that the quantity

$$\Delta_n(X, Y) = \max_{1 \le k \le n} \left\| \sum_{j=1}^{k} X_j - \sum_{j=1}^{k} Y_j \right\|$$

would be as small as possible with sufficiently large probability.

$d = 1$: Prokhorov (1956), Skorokhod (1961), Borovkov (1973), Csörgő and Révész (1975), KMT (1975), Sakhanenko (1984, 1985, 1989, 2006), Chatterjee (2008)

$d > 1$: Gorodetskii (1975), Berkes and Philipp (1979), Philipp (1979), Borovkov and Sakhanenko (1981), Berger (1982), Einmahl (1987, 1989), Sakhanenko (2000)

Strassen (1967) started the study of Problem (A) in the one-dimensional case. He has shown that there exists a construction such that

$$\left| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \right| = o(\sqrt{n \log\log n}) \quad \text{a.s.} \quad \text{as } n \to \infty, \tag{2}$$

assuming that there exists finite $\mathbf{E}\, X^2$ and $\mathbf{E}\, X = 0$ (see Philipp (1979) for a multidimensional version of this statement). Strassen (1967) has used in the construction Skorokhod's embedding (1961). Major (1978) proved that, for any sequence $\{a_n\}$ of real numbers such that $a_n \nearrow \infty$, there exists a one-dimensional distribution $\mathcal{L}(X)$ with $\mathbf{E}\, \mathbf{X}^\mathbf{2} < \infty$ and $\mathbf{E}\, X = 0$ such that, for any construction,

$$\mathbf{P}\left\{ \limsup_{n\to\infty} \ a_n\, (n \log\log n)^{-1/2} \left| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \right| = \infty \right\} = 1.$$

This confirms the optimality of Strassen's result (2).

Improvements and generalizations of relation (2) under additional assumptions concerning the existence of some moments were obtained in the papers of Strassen (1967), Breiman (1967), and Csörgő and Révész (1975). Optimal with respect to order one-dimensional results were obtained in the papers of KMT (1975) and Major (1976) by the method of dyadic approximation. In particular, it was shown that if $\mathbf{E}\,X = 0$ and if the vector X has finite exponential moment $\mathbf{E}\,e^{\lambda|X|}$ for some $\lambda > 0$, then there exists a construction such that

$$\left|\sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j\right| = O(\log n) \quad \text{a.s.} \quad \text{as} \quad n \to \infty. \tag{3}$$

The corresponding multidimensional statement was proved by Einmahl (1989) for sufficiently smooth distributions and by Zaitsev (1998) in the general case. From results of Bártfai (1966) it follows that the accuracy of approximation in (3) is the best possible: in (3), it is impossible to replace O large by o small if the distribution of the vector X is non-normal.

KMT (1975) has shown that if $\mathbf{E}\,X = 0$ and $\mathbf{E}\,|X|^\gamma < \infty$ for some $\gamma > 2$, then there exists a construction such that

$$\left| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \right| = o(n^{1/\gamma}) \quad \text{a.s.} \quad \text{as } n \to \infty. \tag{4}$$

The corresponding multidimensional statement was proved by Einmahl (1989). The case $2 < \gamma \le 3$ was investigated earlier by Berger (1982). It is well known that it is impossible to obtain statement (4) for large γ if one uses for the construction Skorokhod's embedding.

For $\delta > 0$ and $x_0 > 0$, introduce the class $\mathcal{H}(\delta, x_0)$ of nonnegative nondecreasing functions $H : [0, \infty) \to \mathbf{R}^1$ such that the functions $H(x)/x^{2+\delta}$ and $x/\log H(x)$ are nondecreasing for $x \geq x_0$. Denote

$$\mathcal{H} = \bigcup_{\delta>0,\ x_0>0} \mathcal{H}(\delta, x_0).$$

Examples of $H \in \mathcal{H}$:

$$H(x) = c\,x^\alpha,\ \alpha > 2; \qquad H(x) = \exp\{c\,x^\beta\},\ 0 < \beta \leq 1.$$

Theorem (Zaitsev (2009)). *Let $H \in \mathcal{H}$ and let X be a random vector with $\mathbf{E}\,X = 0$, and $\mathbf{E}\,H(\|X\|) < \infty$. Then there exists a construction such that*

$$\mathbf{P}\left(\limsup_{n\to\infty} \Big\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \Big\| / H^{-1}(n) \leq C\right) = 1, \tag{5}$$

where $C < \infty$ is a non-random quantity depending on d, $\mathcal{L}(X)$ and on the function $H(\cdot)$ only.

It is easy to see that (5) implies that

$$\left\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \right\| = O(H^{-1}(n)) \quad \text{a.s.} \quad \text{as } n \to \infty. \tag{6}$$

This statement generalizes to the multidimensional case the results of KMT (1975) and Major (1976).

In the one-dimensional case, KMT (1975) have proved the same statement for functions H from the class $\widetilde{\mathcal{H}}(\delta, x_0)$, $\delta > 0$, of nonnegative nondecreasing functions $H \in \mathcal{H}(\delta, x_0)$ such that the functions $H(x)/x^{3+\delta}$ are nondecreasing for $x \geq x_0$. Major (1976) extended the result to functions H from the class $\mathcal{H}(\delta, x_0)$, $\delta > 0$, such that the functions $H(x)/x^3$ are nonincreasing. Berger (1982) generalized the result of Major (1976) to the multidimensional case.

Einmahl (1989) proved the same statement for functions H from the class $\mathcal{H}^*(\delta, x_0)$, $\delta > 0$, of nonnegative nondecreasing functions H such that the functions $H(x)/x^{3+\delta}$ and $\sqrt{x}/\log H(x)$ are nondecreasing for $x \geq x_0$. Clearly, there exist functions belonging to $\mathcal{H}(\delta, x_0)$ and not belonging to $\mathcal{H}^*(\delta, x_0)$. For example, we may mention the functions $H(x) = \exp\bigl(\lambda x^\beta\bigr)$, $1/2 < \beta \leq 1$, $\lambda > 0$.

Breiman (1967) has shown that the statement of Theorem 1 is optimal in the following sense: if

$$\mathbf{E}\, H\,(\|X\|) = \infty \tag{7}$$

in the conditions of Theorem 1, then

$$\mathbf{P}\left\{\limsup_{n\to\infty} \Big\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \Big\| / H^{-1}(n) \geq 1/4 \right\} = 1 \tag{8}$$

for any construction of the i.i.d. X_j and i.i.d. Gaussian Y_j with the needed distributions on the same probability space.

Note that if the conditions of Theorem 1 are satisfied for $H(x) = |x|^\gamma$, then one can ensure the validity of relation (4), which is stronger than (6). On the other hand, Shao (1989) has shown that for the functions $H(x) = e^{x^\beta}$, $0 < \beta < 1$, it is impossible to replace O large by o small in relation (6), at least for distributions $\mathcal{L}(X)$ such that $\mathbf{E}\, e^{2\|X\|^\beta} = \infty$. In this case, (6) turns into

$$\sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j = O((\log n)^{1/\beta}) \quad \text{a.s.} \quad \text{as} \quad n \to \infty.$$

The question on description of the class of functions $H \in \mathcal{H}$ for which one can replace O large by o small in relation (6) remains open.

Sakhanenko (1985) obtained the following result.

Theorem. *Let $X_1, X_2, \dots, X_n$ be independent random variable with* $\mathbf{E}\, X_j = 0$*,* $j = 1, \dots, n$*. Let* $\gamma > 2$ *and*

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

Then there exists a construction such that

$$\mathbf{E}\, \big(\Delta_n(X, Y)\big)^\gamma \leq c\, \gamma^{2\gamma}\, L_\gamma, \tag{9}$$

where c is an absolute constant.

After the natural normalization, we see that (9) is equivalent to

$$\mathbf{E}\, \big(\Delta_n(X, Y)/\sigma\big)^\gamma \leq c\, \gamma^{2\gamma}\, L_\gamma/\sigma^\gamma,$$

where $\sigma^2 = \mathrm{Var}\big(\sum_{j=1}^{n} X_j\big)$. 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.

It should be mentioned that, in Sakhanenko (1985), it is observed that inequality (9) implies the well-known Rosenthal inequality (1972).

Lemma. *Let $X_1, \dots, X_n$ be independent random vectors which have mean zero and assume values in $\mathbf{R}^d$. Then*

$$\mathbf{E}\Big\|\sum_{j=1}^{n} X_j\Big\|^{\gamma} \le c(\gamma)\Big(\sum_{j=1}^{n}\mathbf{E}\,\|X_j\|^{\gamma} + \Big(\sum_{j=1}^{n}\mathbf{E}\,\|X_j\|^{2}\Big)^{\gamma/2}\Big), \quad \textit{for } \gamma \geqslant 2. \tag{10}$$

This multidimensional version of the Rosenthal inequality follows easily from a result of de Acosta (1981). In the i.i.d. case, the second summand in the right-hand side of (10) grows faster than the first one as $n \to \infty$. Sakhanenko's Theorem shows that this growth corresponds to the growth of moments of sums of Gaussian approximating vectors.

We formulate the results published in the papers of Zaitsev (1998, 2001, 2006, 2007, 2009) and Götze and Zaitsev (2008, 2009). They can be considered as multidimensional generalizations and improvements of some results of Komlós, Major and Tusnády (1975), Sakhanenko (1985) and Einmahl (1989).

Let $\mathcal{A}_d(\tau)$, $\tau \geq 0$, $d \in \mathbf{N}$, denote classes of d-dimensional distributions, introduced in Zaitsev (1986). The class $\mathcal{A}_d(\tau)$ (with a fixed $\tau \geq 0$) consists of d-dimensional distributions V for which the function

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

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

$$\left| d_u d_v^2 \,\varphi(z) \right| \leq \|u\| \tau \left\langle \mathbb{D}\, v, v \right\rangle,$$

for all $u, v \in \mathbf{R}^d$ and $\|z\|\,\tau < 1$, where $\mathbb{D} = \mathrm{cov}\, V$, and $d_u \varphi$ is the derivative of the function φ in direction u.

Below we consider simplest properties of the classes $\mathcal{A}_d(\tau)$.

In particular, for fixed τ, the class $\mathcal{A}_d(\tau)$ is closed with respect to convolution: if $F_1, F_2, \ldots, F_n \in \mathcal{A}_d(\tau)$, then $F_1 F_2 \cdots F_n \in \mathcal{A}_d(\tau)$.
If $\mathcal{L}(X) \in \mathcal{A}_d(\tau)$ and $a \in \mathbf{R}$, then $\mathcal{L}(a X) \in \mathcal{A}_d(|a|\, \tau)$.

As examples of distributions from $\mathcal{A}_d(c\tau)$ we can consider distributions concentrated on the ball $B_\tau = \{x \in \mathbf{R}^d : \|x\| \le \tau\}$ and infinitely divisible distributions with spectral measures concentrated on the ball B_τ. The same may be said about distributions satisfying multidimensional analogs of conditions of Bernstein's inequality.
The class $\mathcal{A}_d(0)$ coincides with the class of all d-dimensional Gaussian distributions.

The following inequality was proved in Zaitsev (1986) and can be considered as an estimate of stability of this characterization: if $F \in \mathcal{A}_d(\tau)$, $\tau > 0$, then

$$\pi\big(F, \Phi_F\big) \le c\, d^2 \tau \log^*(\tau^{-1}),$$

where the notation Φ_F is used for the Gaussian distribution whose mean and covariance operator are the same as those of a distribution F,

$$\pi(F, G) = \inf\big\{\varepsilon > 0 : F\{X\} \le G\{X^\varepsilon\} + \varepsilon, G\{X\} \le F\{X^\varepsilon\} + \varepsilon\big\}$$

is the Prokhorov distance. Here $X^\varepsilon = \Big\{y \in \mathbf{R}^d : \inf\limits_{x \in X} \|x - y\| < \varepsilon\Big\}$ is the ε-neighborhood of the set X. Moreover, in Zaitsev (1986) it was established that for all Borel sets X and all $\lambda > 0$

$$F\{X\} \le \Phi_F\{X^\lambda\} + c\, d^2 \exp\Big(-\frac{\lambda}{c\, d^2\, \tau}\Big),$$

$$\Phi_F\{X\} \le F\{X^\lambda\} + c\, d^2 \exp\Big(-\frac{\lambda}{c\, d^2\, \tau}\Big).$$

Theorem 1 (*Zaitsev* (*1998*)). *Suppose that* $\tau \geq 1$, *and* $X_1, \ldots, X_n$ *are random vectors with distributions* $\mathcal{L}(X_k) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\, X_k = 0$, $\operatorname{cov} X_k = \mathbb{I}_d$, $k = 1, \ldots, n$. *Then there exists a construction such that*

$$\mathbf{E}\exp\Big(\frac{c_1\,\Delta_n(X,Y)}{\tau\, d^3 \log^* d}\Big) \leq \exp\Big(c_2\, d^{5/2} \log^*\,(n/\tau^2)\Big),$$

where c_1, c_2 *are absolute positive constants.*

Theorem 1 allowed us to remove a logarithmic factor from the result of Einmahl (1989) and to obtain a multidimensional analog of the KMT result (1975) for vectors with finite exponential moments. Zaitsev (2001) generalized Theorem 2 to the case of non-i.i.d. summands with different covariance operators. In Zaitsev (1998), Theorem 2 was formulated and proved for a fixed n. This means that the probability space depends on this n. However, a repeated application of the result for fixed n allows one to get a construction of the vectors $\{X_j\}$ and $\{Y_j\}$ such that (1.11) is satisfied for all n simultaneously on the same probability space. It suffices to take independent constructions from the formulation of Theorem 2 with fixed $n = 2^{2^m}$, $m = 1, 2, \ldots$, just as was really done in the KMT paper (1975).

Let $\mathcal{H}$ be the class of non-negative non-decreasing continuous functions $H : [0, \infty) \to \mathbf{R}^1$ such that (for some $\delta > 0$ and $x_0 > 0$) the functions $H(x)/x^{2+\delta}$ and $x/\log H(x)$ are non-decreasing, for $x \geq x_0$. The distribution of a random vector ξ will be denoted below by $\mathcal{L}(\xi)$.

Examples: $\quad H(x) = c\,x^{\alpha}$, $\alpha > 2$; $\quad H(x) = \exp\{c\,x^{\beta}\}$, $0 < \beta \leq 1$.

We consider the rate of strong approximation assuming that, for some function $H \in \mathcal{H}$, $\mathbf{E}\,H\,(\|X_j\|) < \infty$, $j = 1, 2, \ldots, n$. The $X_1, \ldots, X_n$ will be generally speaking non-i.i.d., but, for the sake of simplicity, firstly we give the results in the case of i.i.d. $X_1, \ldots, X_n$ only.

Theorem 2 (*Zaitsev* (*2007*)). *Let $H \in \mathcal{H}$ and X be a random vector with* $\mathbf{E}\, X = 0$ *and* $\mathbf{E}\, H(\|X\|) < \infty$. *Then, for any $z > 0$ and $n \geq 1$, there exists a construction such that*

$$\mathbf{P}\left(\Delta_n(X, Y) > c_3\, z\right) \leq \frac{c_4\, n}{H(z)}, \tag{11}$$

where c_3 and c_4 are positive quantities depending only on $\mathcal{L}(X)$ and on the function $H(\,\cdot\,)$.

Theorem 3 (*Zaitsev* (*2009*)). *Let H and X satisfy the conditions of Theorem* 2. *Then there exists a construction such that* (1) *is satisfied for all $j = 1, 2, \ldots$, and*

$$\mathbf{P}\left(\limsup_{n\to\infty} \Big\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \Big\| \big/ H^{-1}(n) < C\right) = 1.$$

where $C < \infty$ is a non-random quantity depending only on d, $\mathcal{L}(X)$ and on the function $H(\,\cdot\,)$

Theorems 2 and 3 generalize to the multidimensional case the results of Komlós, Major and Tusnády (1975–1976). Einmahl (1989) proved the same statements for the functions H from the class $\widetilde{\mathcal{H}}$ of non-negative non-decreasing continuous functions H such that the functions $H(x)/x^{3+\delta}$ and $\sqrt{x}/\log H(x)$ are non-decreasing, for $x \geq x_0$. Clearly, there exists a lot of functions belonging to $\mathcal{H}$ and not belonging to $\widetilde{\mathcal{H}}$. For example, we may mention the functions $H(x) = \exp\big(\lambda x^{\beta}\big)$, $1/2 < \beta \leq 1$, $\lambda > 0$.

Theorem 4 (*Zaitsev* (*1998*)). *Let H and X satisfy the conditions of Theorem 2, and the function $x/\log(H(x)/L_H)$ be non-decreasing for $x > u$, where $L_H = n\,\mathbf{E}\,H(\|X\|)$ and*

$$u = C_1\,H^{-1}(C_2\,L_H), \tag{12}$$

with some constants $C_1 \geq 1$ and $C_2 \geq 1$, where $H^{-1}(\,\cdot\,)$ is the inverse function for H. Then, for any $n \geq 1$, there exists a construction such that

$$\mathbf{P}\left(\Delta_n(X, Y) > c_5\,z\right) \leq \frac{c_6\,n}{H(z)}, \tag{13}$$

for any $z > 0$, where c_5 and c_6 are positive quantities depending only on C_1, C_2, $\mathcal{L}(X_1)$ and on the function $H(\,\cdot\,)$.

The conditions of Theorem 4 are satisfied, for example, for the function $H \in \mathcal{H}$ such that the function $H(x)/x^{\gamma}$ is non-increasing for some $\gamma > 2$. Then, in the proof of Corollary 2 of Zaitsev (2006), it was shown that one can take $u = H^{-1}(e^{\gamma} L_H)$ in (12).

Another example is given by $H(x) = \exp\big(\lambda x^{\beta}\big)$, $\lambda > 0$, $0 < \beta < 1$. In this case one can take $u = (1-\beta)^{-1/\beta} H^{-1}(L_H)$ in (12).

Theorem 5 (*Götse and Zaitsev* (*2008, 2009*)). *Assume that $\gamma > 2$ and X is a random vector with* $\mathbf{E}\, X = 0$, $\mathbf{E}\, \|X\|^\gamma < \infty$ *and* $\operatorname{cov} X = \mathbb{I}$, *the identity operator. Then, for any $n \geq 1$, there exists a construction such that*

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \leq c_7\, A\, n\, \mathbf{E}\, \|X\|^\gamma\,, \tag{14}$$

where

$$A = A(\gamma, d) = \max\big\{d^{11\gamma}, d^{\frac{\gamma(\gamma+2)}{4}}\,(\log^* d)^{\frac{\gamma(\gamma+1)}{2}}\big\}, \tag{15}$$

and c_7 is a positive constant depending only on γ.

Corollary 1. *Let X satisfy the conditions of Theorem 5. Then there exists a construction such that* (1) *is satisfied for all $j = 1, 2, \ldots$, and*

$$\mathbf{E}\,\big(\Delta_n(X,Y)\big)^\gamma \leq c_8\, A\, n\, \mathbf{E}\, \|X\|^\gamma\,, \tag{16}$$

for all $n \geq 1$, where A is defined in (15) *and c_8 is a positive constant depending only on γ.*

Theorem 6 (*Götse and Zaitsev* (*2008*)). *Suppose that* $X_1, \dots, X_n$ *are independent* $\mathbf{R}^d$*-valued random vectors with* $\mathbf{E}\, X_j = 0$*,* $j = 1, \dots, n$*. Let* $\gamma \geq 2$ *and*

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

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*

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

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

$$w^2 \|v\|^2 \leq \left\langle \mathbb{B}_k v, v \right\rangle \leq C_1 w^2 \|v\|^2,$$

where

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

with some positive constants C_1 *and* C_2*.*

Suppose that the quantities

$$\lambda_{k,\gamma} = \sum_{j=m_{k-1}+1}^{m_k} \mathbf{E}\,\|X_j\|^{\gamma}, \qquad k = 1, \ldots, s,$$

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

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

with a positive constant C_3*. Then there exists a construction such that*

$$\mathrm{E}\big(\Delta_n(X,Y)\big)^{\gamma} \le c_9\, \big(\varepsilon^{-1}\, d^{11}\, \log^* d\big)^{\gamma}\, L_{\gamma},$$

where c_9 *is a positive quantity depending on* γ*,* C_1*,* C_2 *and* C_3 *only.*

Theorem 7. (*Zaitsev* (*2001*)). *Suppose that* $\tau \geq 1$, $D > 0$ *and* $X_1, \ldots, X_s$ *are independent random vectors with* $\mathrm{E}\, X_j = 0$, $j = 1, \ldots, s$. *Assume that there exists a strictly increasing sequence of non-negative integers* $m_0 = 0$, $m_1, \ldots,$ $m_l = s$ *satisfying the following conditions. Write*

$$\zeta_p = D\left(X_{m_{p-1}+1} + \cdots + X_{m_p}\right), \qquad p = 1, \ldots, n,$$

and suppose that (*for all* $p = 1, \ldots, n$) $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$ *and the covariance operators* $\operatorname{cov} \zeta_p = \mathbb{B}_p$ *satisfy for all* $u \in \mathbf{R}^d$

$$C_1^2 \, \|u\|^2 \leq \langle \mathbb{B}_p u, u \rangle \leq C_2^2 \, \|u\|^2$$

with some positive constants C_1 *and* C_2 *which do not depend on* p. *Then there exists a construction such that*

$$\mathrm{E} \exp\Big(\frac{c_{10} \, D \, \Delta_s(X, Y)}{\tau \, d^{9/2} \log^* d}\Big) \leq \exp\Big(c_{11} \, d^{7/2} \log^*(n/\tau^2)\Big),$$

where c_{10}, c_{11} *are positive quantities depending on* C_1, C_2 *only.*

References

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

Götze, F. and Zaitsev, A.Yu. (2008): Bounds for the rate of strong approximation in the multidimensional invariance principle. *Theor. Probab. Appl.*, *53*, 100–123.

Götze, F. and Zaitsev A.Yu. (2009): Rates of approximation in the multidimensional invariance principle for sums of i.i.d. random vectors with finite moments *Zapiski Nauchnyh Seminarov POMI*, *368*, 110–121.

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

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

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

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

Zaitsev, A.Yu. (2006): Estimates for the rate of strong approximation in the multidimensional invariance principle. *Zapiski Nauchnyh Seminarov POMI*, *339*, 37–53 (in Russian).

Zaitsev, A.Yu. (2007): Estimates for the rate of strong Gaussian approximation for the sums of i.i.d. multidimensional random vectors. *Zapiski Nauchnyh Seminarov POMI*, *351*, 141–157 (in Russian).

Zaitsev A.Yu. (2009): The rate of Gaussian strong approximation for the sums of i.i.d. multidimensional random vectors *Zapiski Nauchnyh Seminarov POMI*, *364*, 148–165 (in Russian).