On the strong gaussian approximation in multidimensional case
A. Yu. Zaïtsev

To cite this version:

HAL Id: hal-03649059
https://hal.science/hal-03649059
Submitted on 22 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
ON THE STRONG GAUSSIAN APPROXIMATION IN MULTIDIMENSIONAL CASE

A. YU. ZAITSEV

St. Petersburg Branch of Steklov Mathematical Institute

June 2000

Abstract. A multidimensional version of a result of Komlôs, Major and Tusnády for sums of independent random vectors with finite exponential moments is proved.

The aim of this paper is to present a proof of Corollary 1 from Zaitsev (1998a) (given therein without proof) which yields the rate of strong Gaussian approximation for sums of independent multidimensional random vectors with finite exponential moments (see Theorem 1). Moreover, in Theorem 2 we give a sharpening of Theorem 1.

Notation. We shall write $z \in \mathbb{R}^d$ (resp. $\mathbb{C}^d$), if $z = (z_1, \ldots, z_d) = z_1 e_1 + \cdots + z_d e_d$, where $z_j \in \mathbb{R}^1$ (resp. $\mathbb{C}^1$) and the $e_j$ are the basis vectors. The scalar product of vectors $x, y \in \mathbb{R}^d$ (or $\mathbb{C}^d$) is denoted by $\langle x, y \rangle = x_1 y_1 + \cdots + x_d y_d$. For $z \in \mathbb{R}^d$ (or $\mathbb{C}^d$) we shall use the Euclidean norm $||z|| = (\langle z, z \rangle)^{1/2}$ and the maximum norm $|z| = \max_{1 \leq j \leq d} |z_j|$. For $b > 0$ we denote $L(b) = \max \{1, \log b\}$. The distribution and the corresponding covariance operator of a random vector $\xi$ will be denoted by $\mathcal{L}(\xi)$ and $\text{cov} \xi$. The letter $I$ will be used for the identity operator. The symbols $c_1(\cdot), c_2(\cdot), \ldots$ will be used for positive quantities depending on the arguments only. The letter $c$ can denote different absolute positive constants when we do not need to fix their numerical values. We shall say that a sequence $\{Y_n\}$ of Gaussian vectors is corresponding to a sequence of random vectors $\{X_n\}$ if these sequences have coinciding means and covariances.

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

Key words and phrases. Multidimensional invariance principle, strong approximation, sums of independent random vectors, Central Limit Theorem.

Research supported by the Russian Foundation of Basic Research (RFBR), grant 99-01-00112, by grant INTAS 99-01317, by grant RFBR-DFG 99-01-04027 and by the SFB 343 in Bielefeld.

Typeset by A4T\LaTeX
A. YU. ZAITSEV

Denote by $A_d(\tau)$ a class of probability distributions depending on a parameter $\tau \geq 0$. The class $A_d(\tau)$ (introduced by Zaitsev (1986)) consists of $d$-dimensional distributions $F = L(\xi)$ for which the function

$$\varphi(z) = \varphi(F, z) = \log \int_{\mathbb{R}^d} e^{(z, x)} F(dx) \quad (\varphi(0) = 0)$$

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

$$|d_u d_v \varphi(z)| \leq \|u\| \tau \langle Dv, u \rangle$$

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

In Zaitsev (1986) one can find some properties of classes $A_d(\tau)$. As examples of distributions from $A_d(\tau)$ one can consider the distributions concentrated on the ball $B_\tau = \{x \in \mathbb{R}^d : \|x\| \leq \tau\}$, more general distributions satisfying Bernstein-type inequality conditions and infinitely divisible distributions with spectral measures concentrated on $B_\tau$ (see Zaitsev (1986, pp. 205-207)).

The following theorem is the main result of Zaitsev (1998a).

**Theorem A.** Suppose that $\tau \geq 1$ and $\xi_1, \ldots, \xi_n$ are random vectors with distributions $L(\xi_k) \in A_d(\tau)$, $E \xi_k = 0$, $\text{cov} \xi_k = I$, $k = 1, \ldots, n$. Then, for any $\alpha > 0$, one can construct on a probability space a sequence of independent random vectors $X_1, \ldots, X_n$ and a corresponding sequence of independent Gaussian random vectors $Y_1, \ldots, Y_n$ so that $L(X_k) = L(\xi_k), k = 1, \ldots, n$, and

$$E \exp \left( \frac{c_1(\alpha) \Delta(X, Y)}{\tau d^3 L(d)} \right) \leq \exp \left( c_2(\alpha) d^{\beta/4 + \alpha} L(n/\tau^2) \right),$$

where

$$\Delta(X, Y) = \max_{1 \leq k \leq n} \left| \sum_{i=1}^{k} X_i - \sum_{i=1}^{k} Y_i \right|$$

and $c_1(\alpha), c_2(\alpha)$ are positive quantities depending only on $\alpha$.

In a particular case, when $d = 1$ and all summands have a common variance, Theorem A is equivalent to the main result of Sakhanenko (1984) which is a generalization of the famous result of Komlós, Major and Tusnády (KMT) (1975-76) to the case of non-identically distributed random variables. A multidimensional generalization of the result of Sakhanenko (1984) can be found in Zaitsev (1999).

One can easily verify that if a vector $\xi$ has finite exponential moments $E e^{(h, \xi)}$, for $h \in V$, where $V \subset \mathbb{R}^d$ is some neighborhood of zero, then $F = L(\xi) \in A_d(c_3(F))$. Therefore, Theorem A can be considered as a generalization and refinement of the main result of KMT (1975-76). In particular, from Theorem A one can easily derive the following result, obtained by KMT (1975-76) in the one-dimensional case.
STRONG APPROXIMATION

Theorem 1. Let a random vector $\xi$ have finite exponential moments $\mathbb{E}e^{(h,\xi)}$, for $h \in V$, where $V \subset \mathbb{R}^d$ is a neighborhood of zero. Then one can construct on a probability space a sequence of independent random vectors $X_1, X_2, \ldots$ and a corresponding sequence of independent Gaussian random vectors $Y_1, Y_2, \ldots$ so that $\mathcal{L}(X_k) = \mathcal{L}(\xi)$, $k = 1, 2, \ldots$, and

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

As it is noted in KMT (1975-76), from the results of Bártfai (1966) it follows that the accuracy of approximation in (1) is the best possible for non-Gaussian vectors $\xi$. An analog of Theorem 1 was obtained by Einmahl (1989) under additional smoothness-type restrictions on the distributions $\mathcal{L}(\xi)$. Einmahl (1989, Theorem 10) has also proved an analog of Theorem 1, but only for sufficiently smooth distributions $\mathcal{L}(\xi)$. In general case the corresponding results of Einmahl (1989) contain an additional logarithmic factor.

Instead of Theorem 1 we shall prove the following stronger statement.

Theorem 2. Let $\tau \geq 1$. Assume that a random vector $\xi$ has the distribution such that $\mathcal{L}(\mathbb{D}^{-1/2} \xi) \in \mathcal{A}_d(\tau)$, where $\mathbb{D} = \text{cov} \xi$ is a reversible operator. Let $\sigma^2, \sigma > 0$, be the maximal eigenvalue of $\mathbb{D}$. Then, for any $\alpha > 0$, there exist a construction from Theorem 1 such that

$$P \left\{ \limsup_{n \to \infty} \frac{1}{\log n} \left\| \sum_{j=1}^{n} X_j - \sum_{j=1}^{n} Y_j \right\| \leq c_4(\alpha) \sigma \tau d^{3/4 + \alpha} L(d) \right\} = 1 \quad (2)$$

with some $c_4(\alpha)$ depending on $\alpha$ only.

In Theorem 2 the dependence of the constant which corresponds to $O(\log n)$ in (1) on the distribution $\mathcal{L}(\xi)$ is written out in an explicit form.

Remark. In Theorem 2 we consider the case $\tau \geq 1$. The case of small $\tau$ was considered by Götze and Zaitsev (2000). It is shown that under additional smoothness-type restrictions on the distribution $\mathcal{L}(\xi)$ the expression in the right-hand side of the inequality in (2) can be arbitrarily small if the parameter $\tau$ is small enough.

Proof of Theorem 2. Without loss of generality we assume that $\text{cov} \xi = \mathbb{I}$. In general situation one should consider (instead of $\xi$) the vector $\mathbb{D}^{-1/2} \xi$, with $\text{cov} \mathbb{D}^{-1/2} \xi = \mathbb{I}$. Then one should apply to the constructed vectors the operator $\mathbb{D}^{1/2}$ having the maximal eigenvalue $\sigma$.

Define $m_0, m_1, m_2, \ldots$ and $n_1, n_2, \ldots$ by

$$m_0 = 0, \quad m_s = 2^s, \quad n_s = m_s - m_{s-1}, \quad s = 1, 2, \ldots \quad (3)$$

It is easy to see that

$$\log n_s \leq \log m_s = 2^s \log 2, \quad s = 1, 2, \ldots \quad (4)$$
By Theorem A, for each $s = 1, 2, \ldots$ one can construct on a probability space a sequence of i.i.d. $X_1^{(s)}, \ldots, X_n^{(s)}$ and a sequence of i.i.d. Gaussian $Y_1^{(s)}, \ldots, Y_n^{(s)}$ so that $L(X_k^{(s)}) = L(\xi)$, $E Y_k^{(s)} = 0$, $\text{cov} Y_k^{(s)} = I$, and (using Chebyshev's inequality)

$$
P \{ c_1(\alpha) \Delta_s \geq \tau d^{\alpha/2} L(d) \left( c_2(\alpha) d^{\alpha/2} \log n_s + x \right) \} \leq e^{-x}, \quad x \geq 0,$$

where

$$
\Delta_s = \max_{1 \leq r \leq n_s} \left\| \sum_{k=1}^{r} X_k^{(s)} - \sum_{k=1}^{r} Y_k^{(s)} \right\|.
$$

It is clear that we can define all the vectors mentioned above on the same probability space so that the collections $\Xi_s = \{ X_1^{(s)}, \ldots, X_n^{(s)}; Y_1^{(s)}, \ldots, Y_n^{(s)} \}$, $s = 1, 2, \ldots$ are jointly independent. Then we define $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ by

$$
X_{m_s + 1 + k} = X_k^{(s)}, \quad Y_{m_s + 1 + k} = Y_k^{(s)},
$$

$k = 1, \ldots, n_s$, $s = 1, 2, \ldots$

In order to show that these sequences satisfy the assertion of Theorem 2, it remains to verify the equality (2).

Put

$$
c_5(\alpha) = \frac{c_2(\alpha) \log 2 + 1}{c_1(\alpha)}, \quad c_6(\alpha) = c_5(\alpha) \sum_{l=0}^{\infty} 2^{-l/2} = \frac{c_5(\alpha) \sqrt{2}}{\sqrt{2} - 1},
$$

and introduce the events

$$
A_l = \{ \omega : \Delta^{(l)} \geq 2^l c_6(\alpha) \tau d^{3/4+\alpha} L(d) \}, \quad l = 1, 2, \ldots,
$$

where

$$
\Delta^{(l)} = \max_{1 \leq r \leq m_l} \left\| \sum_{j=1}^{r} X_j - \sum_{j=1}^{r} Y_j \right\|.
$$

According to (6), (7) and (10), we have

$$
\Delta^{(l)} \leq \Delta_1 + \cdots + \Delta_l.
$$

Taking into account the relations (4), (8), (9), (11) and applying the inequality (5) with $x = 2^{(s+l)/2}$, we get

$$
P \{ A_l \} \leq \sum_{s=1}^{l} P \{ \Delta_s \geq 2^{(s+l)/2} c_6(\alpha) \tau d^{3/4+\alpha} L(d) \}
\leq \sum_{s=1}^{l} \exp \left( -2^{(s+l)/2} \right) \leq c \exp \left( -2^{l/2} \right).
$$
STRONG APPROXIMATION

The inequality (12) implies that
\[ \sum_{i=1}^{\infty} P \{ A_i \} < \infty, \]
Hence, by the Borel-Cantelli lemma with probability one a finite number of the events \( A_i \) occurs only. This implies the equality (2) with \( c_d(\alpha) = 2c_0(\alpha)/\log 2 \) (see (3), (9) and (10)).

Proof of Theorem 1. Without loss of generality we assume the operator \( \mathbb{D} = \text{cov} \xi \) to be reversible. Denote \( F = \mathcal{L}(\mathbb{D}^{-1/2} \xi) \). We have already mentioned that there exists \( \tau = c_3(F) \) such that
\[ F \in \mathcal{A}_d(\tau). \]
We can take \( \tau \geq 1 \) because if \( \tau < 1 \) in (13), then the relation (13) is trivially satisfied for \( \tau = 1 \). It remains to apply Theorem 2.

REFERENCES


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)


Andrei Yu. Zaitsev
St. Petersburg Branch of Steklov Mathematical Institute
Fontanka 27
St. Petersburg 191011
Russia
E-mail address: zaitsev@pdmi.ras.ru