A MULTIVARIATE VERSION OF KOLMOGOROV'S SECOND UNIFORM LIMIT THEOREM

FRIEDRICH GÖTZE, ANDREI YU. ZAITSEV, AND DMITRY ZAPOROZHETS

Abstract. The aim of the present work is to show that the results obtained earlier on the approximation of distributions of sums of independent summands by infinitely divisible laws may be transferred to the estimation of the closeness of distributions on convex polyhedra.

1. Introduction

Let us fix some $n \in \mathbf{N}$ and consider *independent* random variables

$$\xi_1, \dots, \xi_n \in \mathbb{R}^1,$$

which are not necessary identically distributed. The classical result of Kolmogorov [5] states that if n is large, then under quite general conditions the distribution of their sum is close in the *Lévy metric* to the class of all infinitely divisible distributions. To be more precise, let us first fix some notation.

For a random variable $\xi \in \mathbb{R}^1$ define its *concentration function* as

$$Q(\xi; \tau) := \sup_{x \in \mathbb{R}^1} \mathbf{P}[x \le \xi \le x + \tau].$$

The Lévy metric between two random variables $\xi, \xi' \in \mathbb{R}^1$ is defined as

$$L(\xi, \xi') := \inf\{\lambda > 0 : \mathbf{P}[\xi \le x] \le \mathbf{P}[\xi' \le x + \lambda] + \lambda, \\ \mathbf{P}[\xi' \le x] \le \mathbf{P}[\xi \le x + \lambda] + \lambda \text{ for all } x \in \mathbb{R}^1\}.$$

Denote by $\mathfrak{D}$ the class of all infinitely divisible random variables.

Kolmogorov [5] showed that there exists an *absolute* constant c such that for *arbitrary* independent random variables $\xi_1, \dots, \xi_n$ and for all $\tau > 0$,

$$\inf_{\eta \in \mathfrak{D}} L(\xi_1 + \dots + \xi_n, \eta) \le c \cdot \big(p^{1/5} + \tau^{1/2}(|\log \tau|^{1/4} + 1)\big), \tag{1}$$

1991 *Mathematics Subject Classification.* Primary 60F05; secondary 60E15, 60G50.

Key words and phrases. sums of independent random variables, closeness of successive convolutions, convex polyhedra, approximation, inequalities.

The authors were supported by SFB 1283. The first and second authors were supported by the SPbGU-DFG grant 6.65.37.2017 The second author was supported by grant RFBR 19-01-00356. The second and third authors were supported by the Program of Fundamental Researches of Russian Academy of Sciences "Fundamental Mathematics and its Applications", grant PRAS-18-01 and by the CRC 1283.

where

$$p := \max\{p_1, \dots, p_n\} \quad \text{and} \quad p_i := 1 - Q(\xi_i; \tau),\ i = 1, \dots, n. \tag{2}$$

Let us note that the original result in [5] has a slightly different form formulated for $\tau \in (0, 1/2]$, see Remark 2 below.

The restriction (2) on the distributions of summands is a non-asymptotic analogue of the classical limit constancy condition for the triangular scheme for independent random variables. The bound for the rate of approximations may be considered as a qualitative improvement of the classical Khinchin theorem for the set of infinitely divisible distributions being limit laws of the distributions of sums in a triangular scheme.

The Lévy distance metrizes the weak convergence of probability distributions on the real line. Therefore, Kolmogorov's inequality (1) proves Khinchin's theorem since weak convergence as $p \to 0$ and $\tau \to 0$ of distributions of sums $\xi_1 + \cdots + \xi_n$ to some distribution implies weak convergence to the same limit of distributions of some infinitely divisible variables.This limit is infinitely divisible as a limit of infinitely divisible distributions. However, Kolmogorov's inequality (1) provides good infinitely divisible approximations for fixed small p and τ even if the distributions of sums in the triangular scheme with $p \to 0$ and $\tau \to 0$ are not sequentially compact.

Kolmogorov's methods of proving (1) along with a combinatorial lemma of Sperner have been later used by Rogozin in [8, 9] to obtain the result which is now well known as Kolmogorov–Rogozin inequality: for some absolute constant c and all $\tau > 0$,

$$Q(\xi_1 + \cdots + \xi_n; \tau) \le \frac{c}{\sqrt{p_1 + \cdots + p_n}}.$$

Thereafter, subsequent improvements of the bound in the right-hand side of (1) have been obtained in [6, 4]. Finally, the optimal bound was derived in Zaitsev and Arak [13]:

$$\inf_{\eta \in \mathfrak{D}} L(\xi_1 + \cdots + \xi_n, \eta) \le c \cdot \big(p + \tau(|\log \tau| + 1)\big). \tag{3}$$

The authors also considered in [13] the Lévy–Prokhorov metric

$$\pi(\xi, \xi') := \inf\{\lambda \ge 0 : \mathbf{P}[\xi \in B] \le \mathbf{P}[\xi' \in B^\lambda] + \lambda, \\ \mathbf{P}[\xi' \in B] \le \mathbf{P}[\xi \in B^\lambda] + \lambda \text{ for all Borel sets } B\},$$

where

$$B^\lambda := \{x \in \mathbb{R}^1 : \inf_{y \in B} |x - y| < \lambda\},$$

and obtained the bound

$$\inf_{\eta \in \mathfrak{D}} \pi(\xi_1 + \cdots + \xi_n, \eta) \le c \cdot \big(p + \tau(|\log \tau| + 1)\big) + \sum_{i=1}^n p_1^2, \tag{4}$$

which is optimal as well. For a more detailed discussion of the subject we refer the reader to [1].

Estimates (3) and (4), although being optimal, have their drawbacks inasmuch they are for instance not invariant with respect to the scaling of the random variables. In particular, they become trival for large τ. Thus in [10] the following generalization of (3) and (4) hasb been suggested which lacks these disadvantages: for some absolute constants $c, \varepsilon > 0$ and all $\lambda, \tau > 0$,

$$\begin{aligned} \inf_{\eta \in \mathfrak{D}} L_\lambda(\xi_1 + \cdots + \xi_n, \eta) &\leq c \cdot \big(p + \exp(-\varepsilon \cdot \lambda/\tau)\big), \\ \inf_{\eta \in \mathfrak{D}} \pi_\lambda(\xi_1 + \cdots + \xi_n, \eta) &\leq c \cdot \big(p + \exp(-\varepsilon \cdot \lambda/\tau)\big) + \sum_{i=1}^{n} p_1^2, \end{aligned} \tag{5}$$

where $L_\lambda(\,\cdot\,,\,\cdot\,)$ and $\pi_\lambda(\,\cdot\,,\,\cdot\,)$ are the following refinements of the Lévy and Lévy–Prokhorov metrics:

$$\begin{aligned} L_\lambda(\xi, \xi') &:= \sup_{x \in \mathbb{R}^1} \max\{\mathbf{P}[\xi \leq x] - \mathbf{P}[\xi' \leq x + \lambda], \mathbf{P}[\xi' \leq x] - \mathbf{P}[\xi \leq x + \lambda]\}, \\ \pi_\lambda(\xi, \xi') &:= \sup_{\substack{\text{Borel} \\ \text{sets } B}} \max\{\mathbf{P}[\xi \in B] - \mathbf{P}[\xi' \in B^\lambda], \mathbf{P}[\xi' \in B] - \mathbf{P}[\xi \in B^\lambda]\}. \end{aligned}$$

The latter quantity was first considered in [2], while the former one – in [10]. Knowing $L_\lambda(\,\cdot\,,\,\cdot\,)$ and $\pi_\lambda(\,\cdot\,,\,\cdot\,)$ provides more information on the closeness of the distributions of random variables rather than just knowing $L(\,\cdot\,,\,\cdot\,)$ and $\pi(\,\cdot\,,\,\cdot\,)$. In particular, we have

$$\begin{aligned} L(\xi, \xi') &:= \inf\{\lambda > 0 : L_\lambda(\xi, \xi') < \lambda\}, \\ \pi(\xi, \xi') &:= \inf\{\lambda > 0 : \pi_\lambda(\xi, \xi') < \lambda\}. \end{aligned} \tag{6}$$

Remark 1. We would like to stress that (5) is indeed the generalization of (3) and (4): it is straightforward to check that (5) together with (6) implies (3) and (4), see [10] for details.

Remark 2. In [5], it was essentially proved that for all τ, λ such that $0 < 2\tau \leq \lambda$,

$$\inf_{\eta \in \mathfrak{D}} L_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot \left(p^{1/5} + \frac{\tau}{\lambda} \log^{1/2} \frac{\lambda}{\tau}\right),$$

which together with (6) implies (1) for $\tau \in (0, 1/2]$.

2. Higher Dimensions

The problem discussed in the previous section can be naturally extended to higher dimensions. Now let

$$\xi_1, \ldots, \xi_n \in \mathbb{R}^d$$

be independent d-dimensional random vectors. The concentration function of a random vector $\xi \in \mathbb{R}^d$ is defined as

$$Q(\xi; \tau) := \sup_{x \in \mathbb{R}^d} \mathbf{P}[|\xi - x| \leq \tau], \tag{7}$$

where for $d = 1$ this definition differs from the previous one by the scaling factor 2 in the argument. The definitions of $\pi(\,\cdot\,,\,\cdot\,)$ and $\pi_\lambda(\,\cdot\,,\,\cdot\,)$ in $\mathbb{R}^d$ stay without any changes with B^λ being defined as

$$B^\lambda := \{x \in \mathbb{R}^d : \inf_{y\in B} |x - y| < \lambda\}.$$

The situation with $L(\,\cdot\,,\,\cdot\,)$ and $L_\lambda(\,\cdot\,,\,\cdot\,)$ is more ambiguous. In [11], the following multidimensional versions of these quantities have been suggested. Let

$$\mathbf{1}_d := (1, \dots, 1) \in \mathbb{R}^d.$$

For random vectors $\xi, \xi' \in \mathbb{R}^d$ define

$$\begin{aligned} L(\xi, \xi') := \inf\{\lambda > 0 : \mathbf{P}[\xi \le x] \le \mathbf{P}[\xi' \le x + \lambda\mathbf{1}_d] + \lambda, \\ \mathbf{P}[\xi' \le x] \le \mathbf{P}[\xi \le x + \lambda\mathbf{1}_d] + \lambda \text{ for all } x \in \mathbb{R}^d\} \end{aligned} \tag{8}$$

and

$$L_\lambda(\xi, \xi') := \sup_{x\in\mathbb{R}^d} \max\{\mathbf{P}[\xi \le x] - \mathbf{P}[\xi' \le x + \lambda\mathbf{1}_d], \mathbf{P}[\xi' \le x] - \mathbf{P}[\xi \le x + \lambda\mathbf{1}_d]\}. \tag{9}$$

Consider some arbitrary independent random vectors $\xi_1, \dots, \xi_n \in \mathbb{R}^d$ and let $p, p_1, \dots, p_n$ be defined as in (2). It was shown in [11] that the following multidimensional version of (5) holds: for some constants $c_d, \varepsilon_d > 0$ depending on d only and all $\tau, \lambda > 0$ we have

$$\inf_{\eta\in\mathfrak{D}_d} L_\lambda(\xi_1 + \dots + \xi_n, \eta) \le c_d \cdot \big(p + \exp(-\varepsilon_d \cdot \lambda/\tau)\big), \tag{10}$$

$$\inf_{\eta\in\mathfrak{D}_d} \pi_\lambda(\xi_1 + \dots + \xi_n, \eta) \le c_d \cdot \big(p + \exp(-\varepsilon_d \cdot \lambda/\tau)\big) + \sum_{i=1}^n p_i^2, \tag{11}$$

where $\mathfrak{D}_d$ denotes the class of all d-dimensional infinitely divisible random vectors.

Remark 3. Note that (6) remains true in higher dimensions, too. Thus it is not hard to show (see [10] for details) that (10) together with (6) implies the multidimensional analogues of (3) and (4):

$$\inf_{\eta\in\mathfrak{D}_d} L(\xi_1 + \dots + \xi_n, \eta) \le c_d \cdot \big(p + \tau(|\log\tau| + 1)\big), \tag{12}$$

$$\inf_{\eta\in\mathfrak{D}_d} \pi(\xi_1 + \dots + \xi_n, \eta) \le c_d \cdot \big(p + \tau(|\log\tau| + 1)\big) + \sum_{i=1}^n p_i^2. \tag{13}$$

3. Main results

The definition of the multidimensional Lévy metric given in the previous section is not really natural: it heavily depends on the choice of the coordinate basis, while the concentration functions involved in the upper bounds in (10) do not. Our aim is to suggest a coordinate-free definition and to obtain a counterpart of (10) for it. Let us start with some notation.

For $m \in \mathbf{N}$ we denote by $\mathcal{P}_m$ the class of convex polyhedra $P \subset \mathbb{R}^d$ representable as

$$P = \big\{x \in \mathbb{R}^d : \langle x, t_j\rangle \le b_j,\ j = 1, \dots, m\big\}, \tag{14}$$

where $t_j \in \mathbb{R}^d$ with $|t_j| = 1$ and $b_j \in \mathbb{R}^1$, $j = 1, \dots, m$. For $P \in \mathcal{P}_m$ defined in (14) and $\lambda \geq 0$ let

$$P_\lambda = \big\{x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j + \lambda,\ j = 1, \dots, m\big\}.$$

By definition, P_λ is the intersection of closed λ-neighborhoods of half-spaces $\big\{x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j\big\}$, where $j = 1, \dots, m$. Let us stress that P_λ depends on representation (14) which might be not unique for a given polyhedron.

Let $e_1, \dots, e_d \in \mathbb{R}^d$ be the vectors of the standard Euclidean basis. If we consider the subclass $\mathcal{P}_d^* \subset \mathcal{P}_d$ of those polyhedra which are representable as

$$P = \big\{x \in \mathbb{R}^d : \langle x, e_j \rangle \leq b_j,\ j = 1, \dots, d\big\}, \tag{15}$$

then (8) and (9) are obviously equivalent to

$$\begin{aligned} L(\xi, \xi') := \inf\{\lambda \geq 0 : \mathbf{P}[\xi \in P] &\leq \mathbf{P}[\xi' \in P_\lambda] + \lambda, \\ \mathbf{P}[\xi' \in P] &\leq \mathbf{P}[\xi \in P_\lambda] + \lambda \text{ for all } P \in \mathcal{P}_d^*\} \end{aligned}$$

and

$$L_\lambda(\xi, \xi') := \sup_{P \in \mathcal{P}_d^*} \max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P_\lambda], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P_\lambda]\}. \tag{16}$$

This observation suggests the following coordinate-free definition for $L(\,\cdot\,,\,\cdot\,)$ and $L_\lambda(\,\cdot\,,\,\cdot\,)$ in $\mathbb{R}^d$.

Definition 3.1. For $m \in \mathbf{N}$ and for random vectors $\xi, \xi' \in \mathbb{R}^d$ define

$$\begin{aligned} L_m(\xi, \xi') := \inf\{\lambda \geq 0 : \mathbf{P}[\xi \in P] &\leq \mathbf{P}[\xi' \in P_\lambda] + \lambda, \\ \mathbf{P}[\xi' \in P] &\leq \mathbf{P}[\xi \in P_\lambda] + \lambda \text{ for all } P \in \mathcal{P}_m\} \end{aligned}$$

and

$$L_{\lambda,m}(\xi, \xi') := \sup_{P \in \mathcal{P}_m} \max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P_\lambda], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P_\lambda]\}.$$

It follows directly from the definition that for $m \geq d$,

$$L(\,\cdot\,,\,\cdot\,) \leq L_m(\,\cdot\,,\,\cdot\,) \quad \text{and} \quad L_\lambda(\,\cdot\,,\,\cdot\,) \leq L_{\lambda,m}(\,\cdot\,,\,\cdot\,). \tag{17}$$

Our first result provides upper bounds for $L_m(\,\cdot\,,\,\cdot\,)$ and $L_{\lambda,m}(\,\cdot\,,\,\cdot\,)$ similar to (10) and (12).

Consider again some arbitrary independent random vectors $\xi_1, \dots, \xi_n \in \mathbb{R}^d$ and let $p, p_1, \dots, p_n$ be defined as in (2).

Theorem 3.1. *For any $m \in \mathbf{N}$ there exist constants c_m, ε_m depending on m only such that*

$$\begin{aligned} \inf_{\eta \in \mathfrak{D}_d} L_m(\xi_1 + \dots + \xi_n, \eta) &\leq c_m \cdot \big(p + \tau(|\log \tau| + 1)\big), \\ \inf_{\eta \in \mathfrak{D}_d} L_{\lambda,m}(\xi_1 + \dots + \xi_n, \eta) &\leq c_m \cdot \big(p + \exp(-\varepsilon_m \cdot \lambda/\tau)\big), \end{aligned}$$

for all $\tau, \lambda > 0$.

Let us emphasize that the constants c_m, ε_m do not depend on dimension d.

Applying Theorem 3.1 for $\tau = 0$, and that

$$\mathbf{P}[\xi \in P] = \lim_{\lambda\to 0} \mathbf{P}[\xi \in P_\lambda], \qquad \text{for any } P \in \mathcal{P}_m,$$

we get the following Corollary 3.1.

Corollary 3.1 (Götze and Zaitsev [3], see also Zaitsev [12])**.** *Assume that the conditions of Theorem* 3.1 *are satisfied with $\tau = 0$. Then, for any $m \in \mathbf{N}$, there exist a constant c_m depending on m only such that*

$$\inf_{\eta\in\mathfrak{D}_d} \rho_m(\xi_1 + \dots + \xi_n, \eta) \le c_m p,$$

where

$$\rho_m(\xi, \xi') := \sup_{P\in\mathcal{P}_m} \left|\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P]\right|.$$

In order to prove Corollary 3.1 one should apply Theorem 3.1 as $\tau \to 0$ with $\lambda = \sqrt{\tau} \to 0$.

In the definition of $L_m(\,\cdot\,,\,\cdot\,)$ and $L_{\lambda,m}(\,\cdot\,,\,\cdot\,)$ we considered convex polyhedra P along with their approximations P_λ, while in the classical Lévy–Prokhorov metric *all* Borel sets B along with their neighborhoods B^λ are considered. Thus it is natural to examine the intermediate case: convex polyhedra P along with their neighborhoods P^λ.

Definition 3.2. For $m \in \mathbf{N}$ and for random vectors $\xi, \xi' \in \mathbb{R}^d$ define

$$\begin{aligned}\pi_m(\xi, \xi') := \inf\{\lambda \ge 0 : \mathbf{P}[\xi \in P] \le \mathbf{P}[\xi' \in P^\lambda] + \lambda,\\ \mathbf{P}[\xi' \in P] \le \mathbf{P}[\xi \in P^\lambda] + \lambda \text{ for all } P \in \mathcal{P}_m\}\end{aligned}$$

and

$$\pi_{\lambda,m}(\xi, \xi') := \sup_{P\in\mathcal{P}_m} \max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P^\lambda], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P^\lambda]\}.$$

Again, from the definition we have that

$$L(\,\cdot\,,\,\cdot\,) \le L_m(\,\cdot\,,\,\cdot\,) \le \pi_m(\,\cdot\,,\,\cdot\,) \le \pi(\,\cdot\,,\,\cdot\,) \quad \text{and} \quad L_\lambda(\,\cdot\,,\,\cdot\,) \le L_{\lambda,m}(\,\cdot\,,\,\cdot\,) \le \pi_{\lambda,m}(\,\cdot\,,\,\cdot\,) \le \pi_\lambda(\,\cdot\,,\,\cdot\,). \tag{18}$$

Therefore, (11) and (13) readily give upper bounds for $\pi_m(\,\cdot\,,\,\cdot\,)$ and $\pi_{\lambda,m}(\,\cdot\,,\,\cdot\,)$. However, as our next theorem shows, the restriction from the class of the Borel sets to the convex polyhedra allows us to remove the term $\sum_{i=1}^n p_i^2$ from these bounds.

Theorem 3.2. *For any $m \in \mathbf{N}$ there exist constants c_m, ε_m depending on m only such that*

$$\inf_{\eta\in\mathfrak{D}_d} L_m(\xi_1 + \dots + \xi_n, \eta) \le c_m \cdot \big(p + \tau(|\log\tau| + 1)\big),$$

$$\inf_{\eta\in\mathfrak{D}_d} L_{\lambda,m}(\xi_1 + \dots + \xi_n, \eta) \le c_m \cdot \big(p + \exp(-\varepsilon_m \cdot \lambda/\tau)\big),$$

for all $\tau, \lambda > 0$.

Again, the constants c_m, ε_m are dimension-free.

It readily follows from (18) that Theorem 3.2 implies Theorem 3.1. However, it will be convenient for us first to prove Theorem 3.1 and then to derive Theorem 3.2 from it.

We will derive Theorem 3.1 by means of constructing for any convex polyhedron from $\mathcal{P}_m$ some linear transformation from $\mathbb{R}^d$ to $\mathbb{R}^m$ so that the polyhedron turns out to be a pre-image of another convex polyhedron in $\mathbb{R}^m$ of the form (15), and then applying (10) and (12).

The proof of Theorem 3.2 is more involved. To derive it from Theorem 3.1, we will need to bound P_λ by $P^{c\lambda}$ for some $c > 1$. In general, this is not possible: as it is easily seen, for any $c > 1$ there exists $P \in \mathcal{P}_m$ such that $P_\lambda \not\subset P^{c\lambda}$. However, as noted above, P_λ depends on representation (14) which is not unique. Thus, as the next proposition shows, it will be possible to add to the right-hand side of (14) "not too many" half-spaces which do not affect P, but change P_λ so that $P_\lambda \subset P^{c\lambda}$.

Proposition 3.1. *Fix some $m \in \mathbf{N}$ and $\varepsilon > 0$. Then there exists a constant $c_{m,\varepsilon}$ depending on m, ε only such that for any polyhedron $P \in \mathcal{P}_m$ of the form* (14) *there exist $m_0 \le c_{m,\varepsilon}$ and $b_j, t_j \in \mathbb{R}^d$ with $|t_j| = 1$, $b_j \in \mathbb{R}^1$, $j = m+1, \dots, m_0$, such that*

$$P = \big\{x \in \mathbb{R}^d : \langle x, t_j\rangle \le b_j,\ j = 1, \dots, m_0\big\},$$

and for any $\lambda > 0$,

$$P_\lambda := \big\{x \in \mathbb{R}^d : \langle x, t_j\rangle \le b_j + \lambda,\ j = 1, \dots, m_0\big\} \subset P^{(1+\varepsilon)\lambda}. \tag{19}$$

The proof of Proposition 3.1 is postponed to Section 6. Now let us proceed with the proofs of the theorems.

4. Remark on the compound Poisson distributions

Let $\xi \in \mathbb{R}^d$ be a random vector and let $\xi^{(1)}, \xi^{(2)}, \dots$ be its independent copies. Denote by $e(\xi)$ a random vector in $\mathbb{R}^d$ distributed as

$$e(\xi) \stackrel{d}{=} \sum_{k=0}^{\infty} (\xi^{(1)} + \dots + \xi^{(k)}) \cdot \mathbf{1}\{\zeta = k\},$$

where ζ has the standard Poisson distribution and is independent of $\xi^{(1)}, \xi^{(2)}, \dots$. We say that $e(\xi)$ has the *compound Poisson distribution* with respect to ξ. Clearly, $e(\xi)$ is infinitely divisible. Every time we construct $e(\xi)$ by some random vector ξ, we tacitly assume that it is independent of everything else (including ξ).

Le Cam [7] has found that the compound Poisson distributions are good candidates for the infinitely divisible approximations of the sums of independent random vectors. This observation provided a new way of obtaining the Kolmogorov-type bounds like (3) and (4). Let us be more specific.

As above, fix some $\tau > 0$, and let

$$\xi_1, \dots, \xi_n \in \mathbb{R}^d$$

be independent d-dimensional random vectors and let the quantities $p, p_1, \dots, p_n$ be defined as in (2). From (7) and from the definition of p it follows that there exist $a_1', \dots, a_n' \in \mathbb{R}^d$ such that

$$\mathbf{P}[|\xi_i - a_i'| \le \tau] = 1 - p_i \ge 1 - p, \quad i = 1, \dots, n.$$

Generalizing the previous one-dimensional results, Zaitsev [11] proved the following statement.

Lemma 4.1. *Let $\alpha_i \in \mathbf{R}^1$, $X_i \in \mathbf{R}^d$, $i = 1, \dots, n$, be independent random variables and vectors such that for some $a_i' \in \mathbf{R}^d$*

$$\mathbf{P}[\alpha_i = 1] = 1 - \mathbf{P}[\alpha_i = 0] = p_i, \quad \mathbf{P}[|X_i - a_i'| \le \tau|] = 1. \quad i = 1, \dots, n. \tag{20}$$

Let

$$a_i = \mathbf{E}\, X_i, \quad \xi_i = (1 - \alpha_i) X_i + \alpha_i Y_i, \quad i = 1, \dots, n, \tag{21}$$

where $Y_i \in \mathbf{R}^d$ are some independent random vectors which are independent of $\{X_i, \alpha_i\}$. Then for the vector

$$\eta_0 := \sum_{i=1}^n \big[a_i + e(\xi_i - a_i)\big]. \tag{22}$$

and for some constants $c_d, \varepsilon_d > 0$, depending on d only it holds

$$L_\lambda(\xi_1 + \dots + \xi_n, \eta_0) \le c_d \cdot \big(p + \exp(-\varepsilon_d \cdot \lambda/\tau)\big), \tag{23}$$

$$\pi_\lambda(\xi_1 + \dots + \xi_n, \eta_0) \le c_d \cdot \big(p + \exp(-\varepsilon_d \cdot \lambda/\tau)\big) + \sum_{i=1}^n p_i^2,$$

which implies the bounds (10) *and* (11)*.*

5. Proof of Theorems 3.1 and 3.2

Proof of Theorem 3.1. Fix some polyhedron $P \in \mathcal{P}_m$:

$$P = \big\{x \in \mathbb{R}^d : \langle x, t_j \rangle \le b_j, \ j = 1, \dots, m\big\}.$$

Let $A : \mathbb{R}^d \to \mathbb{R}^m$ be a linear operator mapping as

$$x \mapsto y = \big(\langle x, t_1 \rangle, \dots, \langle x, t_m \rangle\big).$$

Let $e_1, \dots, e_m$ be the standard Euclidean basis in $\mathbb{R}^m$. Consider the polyhedron $\widetilde{P} \subset \mathbb{R}^m$ belonging to the class P_m^* (see (15)) defined as

$$\widetilde{P} = \big\{y \in \mathbb{R}^m : \langle y, e_j \rangle \le b_j, \ j = 1, \dots, m\big\}.$$

Since

$$\langle x, t_j \rangle = \langle x, A^* e_j \rangle = \langle Ax, e_j \rangle,$$

it follows that

$$P = A^{-1}(\widetilde{P}) \quad \text{and} \quad P_\lambda = A^{-1}(\widetilde{P}_\lambda).$$

Therefore, for any random vectors $\xi, \xi' \in \mathbb{R}^d$ we have

$$\max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P_\lambda], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P_\lambda]\}$$
$$= \max\{\mathbf{P}[A\xi \in \widetilde{P}] - \mathbf{P}[A\xi' \in \widetilde{P}_\lambda], \mathbf{P}[A\xi' \in \widetilde{P}] - \mathbf{P}[A\xi \in \widetilde{P}_\lambda]\} \le L_\lambda(A\xi, A\xi'),$$

where in the last step we used (16). Hence,

$$L_{\lambda,m}(\xi, \xi') \le L_\lambda(A\xi, A\xi'). \tag{24}$$

Recall that we consider independent random vectors $\xi_1, \dots, \xi_n \in \mathbb{R}^d$ and $p, p_1, \dots, p_n$ which are defined as in (2). Without loss of generality, we can assume that ξ_i are represented as in (21), where the distributions of X_i's coincide with the conditional distribution of ξ_i's provided that $|\xi_i - a_i'| \le \tau$ with some $a_i' \in \mathbf{R}^d$, which exist by (2). Moreover, $\mathbf{P}[|\xi_i - a_i'| > \tau] = p_i$, and the distributions of Y_i's coincide with the conditional distribution of ξ_i's provided that $|\xi_i - a_i'| > \tau$. Let $a_1, \dots, a_n$ and η_0 be defined in (21) and (22). Since $|t_j| = 1$ for $j = 1, \dots, n$, we have $\|A\| \le \sqrt{m}$. Using this fact gives

$$\mathbf{P}[|AX_i - Aa_i'| \le \sqrt{m}\tau] = 1, \quad i = 1, \dots, n.$$

Moreover,

$$\mathbf{E}\,[AX_i] = Aa_i, \quad i = 1, \dots, n. \tag{25}$$

Notice that

$$A\eta_0 = \sum_{i=1}^n \big[Aa_i + Ae(\xi_i - a_i)\big] = \sum_{i=1}^n \big[Aa_i + e(A\xi_i - Aa_i)\big].$$

Thus, the vectors $A\xi_1, \dots, A\xi_n$ satisfy all the conditions imposed on the vectors $\xi_1, \dots, \xi_n$ in Lemma 4.1 with a_i replaced by Aa_i, a_i' by Aa_i', and τ by $\tau\sqrt{m}$. Therefore, applying (23) to the vectors $A\xi_1, \dots, A\xi_n$ gives (for some constants $c_m, \varepsilon_m > 0$, depending on m only):

$$L_\lambda(A\xi_1 + \dots + A\xi_n, A\eta_0) \le c_m \cdot \big(p + \exp(-\varepsilon_m \cdot \lambda/\tau)\big),$$

which together with (24) implies

$$L_{\lambda,m}(\xi_1 + \dots + \xi_n, \eta_0) \le c_m \cdot \big(p + \exp(-\varepsilon_m \cdot \lambda/\tau)\big).$$

Recalling that η_0 is infinitely divisible finishes the proof of the second part of Theorem 3.1. The first part follows from the second one by standard reasoning, see Remarks 1, 3. □

Proof of Theorem 3.2. Fix some polyhedron $P \in \mathcal{P}_m$:

$$P = \big\{x \in \mathbb{R}^d : \langle x, t_j \rangle \le b_j,\ j = 1, \dots, m\big\}.$$

It follows from Proposition 3.1 that it is possible to represent P in the form

$$P = \big\{x \in \mathbb{R}^d : \langle x, t_j \rangle \le b_j,\ j = 1, \dots, m_0\big\}.$$

such that

$$P_{\lambda/2} \subset P^\lambda \quad \text{and} \quad m_0 \le N_m \in \mathbf{N},$$

where the constant N_m depends on m only. Thus for any random vectors ξ, ξ' we have

$$\max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P^\lambda], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P^\lambda]\}$$
$$\le \max\{\mathbf{P}[\xi \in P] - \mathbf{P}[\xi' \in P_{\lambda/2}], \mathbf{P}[\xi' \in P] - \mathbf{P}[\xi \in P_{\lambda/2}]\} \le L_{\lambda/2,N_m}(\xi,\xi').$$

Since this holds for any $P \in \mathcal{P}_m$ we arrive at

$$\pi_{\lambda,m}(\,\cdot\,,\,\cdot\,) \le L_{\lambda/2,N_m}(\,\cdot\,,\,\cdot\,).$$

Thus the second part of Theorem 3.2 follows from the second part of Theorem 3.1. The first part follows from the second one by standard reasoning, see Remarks 1, 3. □

6. Proof of Proposition 3.1

Proof. First let us construct $b_{m+1}, t_{m+1}, \dots, b_{m_0}, t_{m_0}$. Let

$$H_j := \{x \in \mathbf{R}^d : \langle x, t_j\rangle = b_j\}, \quad j = 1, \dots, m.$$

Denote by $\mathcal{F}_k(P)$ the collection of all k-dimensional faces of the polyhedron P. If $k < d - \min(d,m)$, then $\mathcal{F}_k(P)$ is empty. Therefore,

$$\mathcal{F}(P) := \bigcup_{k=d-\min(d,m)}^{d-1} \mathcal{F}_k(P)$$

is the collection of all proper faces of P. Fix some face $F \in \mathcal{F}(P)$ and let x_F be some point from its relative interior, relint$F$. Denote by T_F the tangent cone at the face F which is defined as

$$T_F = T_F(P) := \{x \in \mathbb{R}^d : x_F + \varepsilon x \in P \text{ for some } \varepsilon > 0\}.$$

Obviously, T_F does not depend on the choice of x_F. Also define by N_F the normal cone at the face F which is defined as the polar cone of T_F:

$$N_F = N_F(P) := \{x \in \mathbb{R}^d : \langle x, y\rangle \le 0 \text{ for all } y \in T_F\}.$$

It is known that

$$\dim N_F \le m. \tag{26}$$

Let $\delta > 0$ be some small enough constant to be fixed later. Since N_F intersected with the unit sphere $\mathbb{S}^{d-1}$ is compact, there exists a finite covering subset $\mathcal{S}_F \subset N_F \cap \mathbb{S}^{d-1}$ such that

$$\max_{v \in \mathcal{S}_F} \langle u, v\rangle \ge 1 - \delta \quad \text{for any} \quad u \in N_F \cap \mathbb{S}^{d-1}. \tag{27}$$

Moreover, it follows from (26) that $\mathcal{S}_F$ is contained in the $(\min(d,m)-1)$-dimensional unit sphere, so

$$\#\mathcal{S}_F \le c_{m,\delta}, \tag{28}$$

where $c_{m,\delta}$ depends on m, δ only.

Now define the desirable set $\{(b_j, t_j)\}_{j=m+1}^{m_0}$ to be a set of all pairs

$$(\langle x_F, v\rangle, v),$$

where v runs over all points of $\mathcal{S}_F$, while F runs over all faces of P. Again, let us emphasize that this definition does not depend on the choices of x_F's. Since any face is represented as an intersection of P with several hyperplanes from $H_1, \dots, H_m$, their number is at most 2^m, which together with (28) implies

$$m_0 \le m + 2^m c_{m,\delta}.$$

Now having defined $b_{m+1}.t_{m+1}, \dots, b_{m_0}, t_{m_0}$, let us prove (19). Fix some point $x_0 \in P_\lambda$. Let y_0 denote the Euclidean projection of x_0 onto P:

$$y_0 := \operatorname{argmin}\{\|x_0 - y\| : y \in P\}. \tag{29}$$

The task is to show that

$$\|x_0 - y_0\| \le (1+\varepsilon)\lambda. \tag{30}$$

We obviously may assume that $x_0 \notin P$ which implies $y_0 \in \partial P$. It is well-known that the boundary of a polyhedron can be represented as a union of the relative interiors of its proper faces (which do not intersect). Thus there is a unique face F of P such that

$$y_0 \in \mathrm{relint} F.$$

The latter together with (29) implies

$$x_0 - y_0 \in N_F.$$

Thus it follows from (27) applying to $(x_0 - y_0)/\|x_0 - y_0\|$ that for some $v \in \mathcal{S}_F$,

$$\langle x_0 - y_0, v\rangle \ge (1-\delta)\|x_0 - y_0\|.$$

At the other hand, since $x_0 \in P_\lambda$ and by the construction of $\{(b_j, t_j)\}_{j=m+1}^{m_0}$,

$$\langle x_0, v\rangle \le \langle y_0, v\rangle + \lambda.$$

Combining the last two inequalities gives

$$\|x_0 - y_0\| \le \frac{\lambda}{1-\delta}.$$

Choosing $\delta = 1 - \frac{1}{1+\varepsilon}$ implies (30), and the proposition follows. □

References

[1] T. V. Arak, A. Yu. Zaitsev, *Uniform limit theorems for sums of independent random variables.* Proc. Steklov Inst. Math., **174** (1988), 222 p.

[2] A. A. Borovkov, A. I. Sahanenko, *On the estimates of the rate of convergence in the invariance principle for Banach spaces.* Theory Probab. Appl. **25**, 4 (1980), 721–731.

[3] F. Götze, A. Yu. Zaitsev, *Estimates for the closeness of convolutions of probability distributions on convex polyhedra.* Zap. Nauchn. Semin. POMI **474** (2018), 108–117 (in Russian). English version arXiv:1812.07473.

[4] I. A. Ibragimov, E. L. Presman, *On the rate of approach of the distributions of sums of independent random variables to accompanying distributions.* Theory Probab. Appl., **18** (1973), 713–727.

[5] A. N. Kolmogorov, *Two uniform limit theorems for sums of independent random variables.* Theory Probab. Appl., **1** (1956), 384–394 (in Russian).

[6] A. N. Kolmogorov, *Approximation of distributions of sums of independent terms by infinitely divisible distributions*. Trudy Moskov. Matem. Ob. **12** (1963), 437–451 (in Russian). English version: *On the approximation of distributions of sums of independent summands by infinitely divisible distributions*. Sankhyā, Ser. A, **25** (1963), 159–174.

[7] L. Le Cam, *On the distribution of sums of independent random variables.* In: Bernoulli, Bayes, Laplace (anniversary volume), pp. 179–202. Berlin; Heidelberg; N.Y.: Springer. 1965.

[8] B. A. Rogozin, *An estimate for concentration functions.* Theory Probab. Appl. **6**, 1 (1961), 94–97.

[9] B. A. Rogozin, *On the increase of dispersion of sums of independent random variables.* Theory Probab. Appl. **6**, 1 (1961), 97–99.

[10] A. Yu. Zaitsev, *Some remarks regarding the approximation of distributions of sums of independent termss*, Zap. Nauchn. Semin. LOMI, **136** (1984), 48–57. (In Russian). English transl. in J. Soviet Math., **33**, 1 (1986), 728–733.

[11] A. Yu. Zaitsev, *Multivariate version of the second Kolmogorov's uniform limit theorem.* Theory Probab. Appl. **34**, 1 (1989), 108–128.

[12] A. Yu. Zaitsev, *Certain class of nonuniform estimates in multidimensional limit theorems.* Zap. Nauchn. Semin. POMI **184** (1990), 92–105 (in Russian). English transl. in J. Math. Sci. (N. Y.), **68**, 4 (1994), 459–468.

[13] A. Yu. Zaitsev, T. V. Arak, On the rate of convergence in the second Kolmogorov's uniform limit theorem, Theory Probab. Appl., **28**, 2 (1984), 351–374.

Email address: goetze@math.uni-bielefeld.de

Fakultät für Mathematik,
Universität Bielefeld, Postfach 100131,
D-33501 Bielefeld, Germany

Email address: zaitsev@pdmi.ras.ru

St. Petersburg Department of Steklov Mathematical Institute
Fontanka 27, St. Petersburg 191023, Russia
and St.Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg, 199034 Russia