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 $n \in \mathbb{N}$ and consider independent random variables

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

which are not necessarily 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

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

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

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

Denote by $\mathcal{D}$ 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 $\xi_1, \ldots, \xi_n$ and for all $\tau > 0$,

$$\inf_{\eta \in \mathcal{D}} L(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot (p^{1/5} + \tau^{1/2}(|\log \tau|^{1/4} + 1)), \quad (1)$$
where
\[
p := \max\{p_1, \ldots, p_n\} \quad \text{and} \quad p_i := 1 - Q(\xi_i; \tau), \quad i = 1, \ldots, n.
\] (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 \( \tau \) 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 \( \tau > 0 \),
\[
Q(\xi_1 + \cdots + \xi_n; \tau) \leq \frac{c}{\sqrt{p_1 + \cdots + p_n}}.
\]

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]:
\[
\inf_{\eta \in \mathcal{D}} L(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot (p + \tau(|\log \tau| + 1)). \tag{3}
\]

The authors also considered in [13] the Lévy–Prokhorov metric
\[
\pi(\xi, \xi') := \inf\{\lambda \geq 0 : P[\xi \in B] \leq P[\xi' \in B^\lambda] + \lambda, \quad P[\xi' \in B] \leq P[\xi \in B^\lambda] + \lambda \text{ for all Borel sets } B\},
\]
where
\[
B^\lambda := \{x \in \mathbb{R}^1 : \inf_{y \in B} |x - y| < \lambda\},
\]
and obtained the bound
\[
\inf_{\eta \in \mathcal{D}} \pi(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot (p + \tau(|\log \tau| + 1)) + \sum_{i=1}^n p_i^2, \tag{4}
\]
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 trivial for large $\tau$. Thus in [10] the following generalization of (3) and (4) has been suggested which lacks these disadvantages: for some absolute constants $c, \varepsilon > 0$ and all $\lambda, \tau > 0$,

$$\inf_{\eta \in D} L_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot \left( p + \exp(-\varepsilon \cdot \lambda / \tau) \right),$$

$$\inf_{\eta \in D} \pi_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c \cdot \left( p + \exp(-\varepsilon \cdot \lambda / \tau) \right) + \sum_{i=1}^{n} p_i^2,$$

(5)

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

$$L_\lambda(\xi, \xi') := \sup_{x \in \mathbb{R}} \max \{ P[\xi \leq x] - P[\xi' \leq x + \lambda], P[\xi \leq x + \lambda] - P[\xi' \leq x] \},$$

$$\pi_\lambda(\xi, \xi') := \sup_{\text{Borel sets } B} \max \{ P[\xi \in B] - P[\xi' \in B^\lambda], P[\xi' \in B^\lambda] - P[\xi \in B] \}. $$

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

$$L(\xi, \xi') := \inf \{ \lambda > 0 : L_\lambda(\xi, \xi') < \lambda \},$$

$$\pi(\xi, \xi') := \inf \{ \lambda > 0 : \pi_\lambda(\xi, \xi') < \lambda \}. $$

(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 $\tau, \lambda$ such that $0 < 2\tau \leq \lambda$,

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

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

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

(7)
4 F. GÖTZE, A.YU. ZAITSEV, AND D. ZAPOROZHETS

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 \( B^\lambda \) 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

\[
1_d := (1, \ldots, 1) \in \mathbb{R}^d.
\]

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

\[
L(\xi, \xi') := \inf\{ \lambda > 0 : P[\xi \leq x] \leq P[\xi' \leq x + \lambda 1_d] + \lambda, \quad P[\xi' \leq x] \leq P[\xi \leq x + \lambda 1_d] + \lambda \text{ for all } x \in \mathbb{R}^d \}
\]

and

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

Consider some arbitrary independent random vectors \( \xi_1, \ldots, \xi_n \in \mathbb{R}^d \) and let \( p, p_1, \ldots, p_n \) 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

\[
\inf_{\eta \in \mathbb{D}_d} L_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c_d \cdot \left( p + \exp(-\varepsilon_d \cdot \lambda/\tau) \right),
\]

\[
\inf_{\eta \in \mathbb{D}_d} \pi_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c_d \cdot \left( p + \exp(-\varepsilon_d \cdot \lambda/\tau) \right) + \sum_{i=1}^n p_i^2,
\]

where \( \mathbb{D}_d \) 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):

\[
\inf_{\eta \in \mathbb{D}_d} L(\xi_1 + \cdots + \xi_n, \eta) \leq c_d \cdot \left( p + \tau(|\log \tau| + 1) \right),
\]

\[
\inf_{\eta \in \mathbb{D}_d} \pi(\xi_1 + \cdots + \xi_n, \eta) \leq c_d \cdot \left( p + \tau(|\log \tau| + 1) \right) + \sum_{i=1}^n p_i^2.
\]

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 \( m \in \mathbb{N} \) we denote by \( \mathbb{P}_m \) the class of convex polyhedra \( P \subset \mathbb{R}^d \) representable as

\[
P = \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j, \quad j = 1, \ldots, m \},
\]

\[
\mathbb{P}_m := \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j, \quad j = 1, \ldots, m \},
\]

where \( B^\lambda \) denotes the class of all \( d \)-dimensional infinitely divisible random vectors.
where \( t_j \in \mathbb{R}^d \) with \( |t_j| = 1 \) and \( b_j \in \mathbb{R}^1, j = 1, \ldots, m \). For \( P \in \mathcal{P}_m \) defined in (14) and \( \lambda \geq 0 \) let
\[
P_\lambda = \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j + \lambda, \ j = 1, \ldots, m \}.
\]
By definition, \( P_\lambda \) is the intersection of closed \( \lambda \)-neighborhoods of half-spaces \( \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j \} \), where \( j = 1, \ldots, m \). Let us stress that \( P_\lambda \) depends on representation (14) which might be not unique for a given polyhedron.

Let \( e_1, \ldots, e_d \in \mathbb{R}^d \) 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
\[
P = \{ x \in \mathbb{R}^d : \langle x, e_j \rangle \leq b_j, \ j = 1, \ldots, d \},
\]
then (8) and (9) are obviously equivalent to
\[
L(\xi, \xi') := \inf \{ \lambda \geq 0 : \mathbb{P}[\xi \in P] \leq \mathbb{P}[\xi' \in P_\lambda] + \lambda, \mathbb{P}[\xi' \in P] \leq \mathbb{P}[\xi \in P_\lambda] + \lambda \text{ for all } P \in \mathcal{P}_d^* \}
\]
and
\[
L_\lambda(\xi, \xi') := \sup \max \{ \mathbb{P}[\xi \in P] - \mathbb{P}[\xi' \in P_\lambda], \mathbb{P}[\xi' \in P] - \mathbb{P}[\xi \in P_\lambda] \}.
\]
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 \( m \in \mathbb{N} \) and for random vectors \( \xi, \xi' \in \mathbb{R}^d \) define
\[
L_m(\xi, \xi') := \inf \{ \lambda \geq 0 : \mathbb{P}[\xi \in P] \leq \mathbb{P}[\xi' \in P_\lambda] + \lambda, \mathbb{P}[\xi' \in P] \leq \mathbb{P}[\xi \in P_\lambda] + \lambda \text{ for all } P \in \mathcal{P}_m \}
\]
and
\[
L_\lambda(\xi, \xi') := \sup \max \{ \mathbb{P}[\xi \in P] - \mathbb{P}[\xi' \in P_\lambda], \mathbb{P}[\xi' \in P] - \mathbb{P}[\xi \in P_\lambda] \}.
\]
It follows directly from the definition that for \( m \geq d \),
\[
L(\cdot, \cdot) \leq L_m(\cdot, \cdot) \quad \text{and} \quad L_\lambda(\cdot, \cdot) \leq L_\lambda(\cdot, \cdot).
\]
Both results provide upper bounds for \( L_m(\cdot, \cdot) \) and \( L_\lambda(\cdot, \cdot) \) similar to (10) and (12).

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

**Theorem 3.1.** For any \( m \in \mathbb{N} \) there exist constants \( c_m, \varepsilon_m \) depending on \( m \) only such that
\[
\inf_{\eta \in \mathbb{D}_d} L_m(\xi_1 + \cdots + \xi_n, \eta) \leq c_m \cdot (p + \tau(|\log \tau| + 1)),
\]
\[
\inf_{\eta \in \mathbb{D}_d} L_\lambda(\xi_1 + \cdots + \xi_n, \eta) \leq c_m \cdot (p + \exp(-\varepsilon_m \cdot \lambda/\tau)),
\]
for all \( \tau, \lambda > 0 \).
Let us emphasize that the constants $c_m, \varepsilon_m$ do not depend on dimension $d$.

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

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

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 $m \in \mathbb{N}$, there exist a constant $c_m$ depending on $m$ only such that

$$\inf_{\eta \in \mathcal{D}_d} \rho_m(\xi_1 + \cdots + \xi_n, \eta) \leq c_m p,$$

where

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

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 $P_\lambda$, while in the classical Lévy–Prokhorov metric all Borel sets $B$ along with their neighborhoods $B^\lambda$ are considered. Thus it is natural to examine the intermediate case: convex polyhedra $P$ along with their neighborhoods $P^\lambda$.

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

$$\pi_m(\xi, \xi') := \inf \{ \lambda \geq 0 : P[\xi \in P] \leq P[\xi' \in P^\lambda] + \lambda, \quad P[\xi' \in P] \leq P[\xi \in P^\lambda] + \lambda \text{ for all } P \in \mathcal{P}_m \}$$

and

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

Again, from the definition we have that

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

(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 $m \in \mathbb{N}$ there exist constants $c_m, \varepsilon_m$ depending on $m$ only such that

$$\inf_{\eta \in \mathcal{D}_d} L_m(\xi_1 + \cdots + \xi_n, \eta) \leq c_m \cdot (p + \tau(| \log \tau | + 1)), \quad \inf_{\eta \in \mathcal{D}_d} L_{\lambda,m}(\xi_1 + \cdots + \xi_n, \eta) \leq c_m \cdot (p + \exp(-\varepsilon_m \cdot \lambda/\tau)),$$

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


Again, the constants $c_m, \varepsilon_m$ 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 $P_m$ 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 $P_\lambda$ 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 $P \in P_m$ such that $P_\lambda \not\subset P_{c\lambda}$. However, as noted above, $P_\lambda$ 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 $P_\lambda$ so that $P_\lambda \subset P_{(1+\varepsilon)\lambda}$.

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

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

and for any $\lambda > 0$,

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

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

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

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

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

$$\xi_1, \ldots, \xi_n \in \mathbb{R}^d$$
be independent \(d\)-dimensional random vectors and let the quantities \(p, p_1, \ldots, p_n\) be defined as in (2). From (7) and from the definition of \(p\) it follows that there exist \(a'_1, \ldots, a'_n \in \mathbb{R}^d\) such that
\[
P[|\xi_i - a'_i| \leq \tau] = 1 - p_i \geq 1 - p, \quad i = 1, \ldots, n.
\]

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

**Lemma 4.1.** Let \(\alpha_i \in \mathbb{R}^1, X_i \in \mathbb{R}^d, i = 1, \ldots, n,\) be independent random variables and vectors such that for some \(a'_i \in \mathbb{R}^d\)
\[
P[\alpha_i = 1] = 1 - \frac{1}{p_i}, \quad \frac{1}{2} \leq p_i \leq 1, \quad i = 1, \ldots, n.
\]
Let
\[
a_i = E(X_i), \quad \xi_i = (1 - \alpha_i)X_i + \alpha_iY_i, \quad i = 1, \ldots, n,
\]
where \(Y_i \in \mathbb{R}^d\) are some independent random vectors which are independent of \(\{X_i, \alpha_i\}\). Then for the vector
\[
\eta_0 := n \sum_{i=1}^n [a_i + e(\xi_i - a_i)].
\]
and for some constants \(c_d, \varepsilon_d > 0\), depending on \(d\) only it holds
\[
L_\lambda(\xi_1 + \cdots + \xi_n, \eta_0) \leq c_d \cdot (p + \exp(-\varepsilon_d \cdot \lambda/\tau)), \quad \pi_\lambda(\xi_1 + \cdots + \xi_n, \eta_0) \leq c_d \cdot (p + \exp(-\varepsilon_d \cdot \lambda/\tau)) + \sum_{i=1}^n p_i^2,
\]
which implies the bounds (10) and (11).

**5. Proof of Theorems 3.1 and 3.2**

**Proof of Theorem 3.1.** Fix some polyhedron \(P \in \mathcal{P}_m:\)
\[
P = \{x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j, \quad j = 1, \ldots, m\}.
\]
Let \(A : \mathbb{R}^d \to \mathbb{R}^m\) be a linear operator mapping as
\[
x \mapsto y = (\langle x, t_1 \rangle, \ldots, \langle x, t_m \rangle).
\]
Let \(e_1, \ldots, e_m\) be the standard Euclidean basis in \(\mathbb{R}^m\). Consider the polyhedron \(\tilde{P} \subset \mathbb{R}^m\) belonging to the class \(\mathcal{P}_m\) (see (15)) defined as
\[
\tilde{P} = \{y \in \mathbb{R}^m : \langle y, e_j \rangle \leq b_j, \quad j = 1, \ldots, m\}.
\]
Since
\[
\langle x, t_j \rangle = \langle x, A^*e_j \rangle = \langle Ax, e_j \rangle,
\]
it follows that
\[
P = A^{-1}(\tilde{P}) \quad \text{and} \quad P_\lambda = A^{-1}(\tilde{P}_\lambda).
\]
Therefore, for any random vectors $\xi, \xi' \in \mathbb{R}^d$ we have
\[
\max\{\mathbb{P}[\xi \in P] - \mathbb{P}[\xi' \in P], \mathbb{P}[\xi' \in P] - \mathbb{P}[\xi \in P]\}
= \max\{\mathbb{P}[A\xi \in \tilde{P}] - \mathbb{P}[A\xi' \in \tilde{P}], \mathbb{P}[A\xi' \in \tilde{P}] - \mathbb{P}[A\xi \in \tilde{P}]\} \leq L_\lambda(A\xi, A\xi'),
\]
where in the last step we used (16). Hence,
\[
L_{\lambda,m}(\xi, \xi') \leq L_\lambda(A\xi, A\xi'). \tag{24}
\]
Recall that we consider independent random vectors $\xi_1, \ldots, \xi_n \in \mathbb{R}^d$ and $p, p_1, \ldots, p_n$ which are defined as in (2). Without loss of generality, we can assume that $\xi_i$ are represented as in (21), where the distributions of $X_i$’s coincide with the conditional distribution of $\xi_i$’s provided that $|\xi_i - a'_i| \leq \tau$ with some $a'_i \in \mathbb{R}^d$, which exist by (2). Moreover, $\mathbb{P}[|\xi_i - a'_i| > \tau] = p_i$, and the distributions of $Y_i$’s coincide with the conditional distribution of $\xi_i$’s provided that $|\xi_i - a'_i| > \tau$. Let $a_1, \ldots, a_n$ and $\eta_0$ be defined in (21) and (22). Since $|t_j| = 1$ for $j = 1, \ldots, n$, we have $\|A\| \leq \sqrt{m}$. Using this fact gives
\[
\mathbb{P}[^{\text{for some polyhedron } P \in \mathcal{P}_m,}
\begin{array}{ll}
\mathbb{P}[AX_i - Aa'_i| \leq \sqrt{m}\tau] = 1, & i = 1, \ldots, n.
\end{array}
\tag{25}
\]
Moreover,
\[
E[AX_i] = Aa_i, \quad i = 1, \ldots, n.
\]
Notice that
\[
A\eta_0 = \sum_{i=1}^n \left[Aa_i + Ae(\xi_i - a_i)\right] = \sum_{i=1}^n \left[Aa_i + c(A\xi_i - Aa_i)\right].
\]
Thus, the vectors $A\xi_1, \ldots, A\xi_n$ satisfy all the conditions imposed on the vectors $\xi_1, \ldots, \xi_n$ in Lemma 4.1 with $a_i$ replaced by $Aa_i$, $a'_i$ by $Aa'_i$, and $\tau$ by $\tau\sqrt{m}$. Therefore, applying (23) to the vectors $A\xi_1, \ldots, A\xi_n$ gives (for some constants $c_m, \varepsilon_m > 0$, depending on $m$ only):
\[
L_\lambda(A\xi_1 + \cdots + A\xi_n, A\eta_0) \leq c_m \cdot (p + \exp(-\varepsilon_m \cdot \lambda/\tau)),
\]
which together with (24) implies
\[
L_{\lambda,m}(\xi_1 + \cdots + \xi_n, \eta_0) \leq c_m \cdot (p + \exp(-\varepsilon_m \cdot \lambda/\tau)).
\]
Recalling that $\eta_0$ 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.

\begin{proof}[Proof of Theorem 3.2] Fix some polyhedron $P \in \mathcal{P}_m$:
\[
P = \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j, \quad j = 1, \ldots, m \}.
\]
It follows from Proposition 3.1 that it is possible to represent $P$ in the form
\[
P = \{ x \in \mathbb{R}^d : \langle x, t_j \rangle \leq b_j, \quad j = 1, \ldots, m_0 \}.
\]
such that
\[
P_{\lambda/2} \subset P^\lambda \quad \text{and} \quad m_0 \leq N_m \in \mathbb{N},
\]
where the constant $N_m$ depends on $m$ only. Thus for any random vectors $\xi, \xi'$ we have
\[
\max\{P[\xi \in P] - P[\xi' \in P^\lambda], P[\xi' \in P] - P[\xi \in P^\lambda]\} \\
\leq \max\{P[\xi \in P] - P[\xi' \in P_{\lambda/2}], P[\xi' \in P] - P[\xi \in P_{\lambda/2}]\} \leq L_{\lambda/2,N_m}(\xi, \xi').
\]
Since this holds for any $P \in \mathcal{P}_m$ we arrive at
\[
\pi_{\lambda,m}(\cdot, \cdot) \leq L_{\lambda/2,N_m}(\cdot, \cdot).
\]
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 $b_{m+1}, t_{m+1}, \ldots, b_{m_0}, t_{m_0}$. Let
\[
H_j := \{x \in \mathbb{R}^d : \langle x, t_j \rangle = b_j\}, \quad j = 1, \ldots, m.
\]
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, $\text{relint} F$. 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 $T_F$:
\[
N_F = N_F(P) := \{x \in \mathbb{R}^d : \langle x, y \rangle \leq 0 \text{ for all } y \in T_F\}.
\]
It is known that
\[
\dim N_F \leq m. \quad (26)
\]
Let $\delta > 0$ be some small enough constant to be fixed later. Since $N_F$ intersected with the unit sphere $S^{d-1}$ is compact, there exists a finite covering subset $\mathcal{S}_F \subset N_F \cap S^{d-1}$ such that
\[
\max_{v \in \mathcal{S}_F} \langle u, v \rangle \geq 1 - \delta \quad \text{for any } u \in N_F \cap S^{d-1}. \quad (27)
\]
Moreover, it follows from (26) that $\mathcal{S}_F$ is contained in the $(\min(d, m) - 1)$-dimensional unit sphere, so
\[
\# \mathcal{S}_F \leq c_{m,\delta}, \quad (28)
\]
where $c_{m,\delta}$ depends on $m, \delta$ 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 \( S_F \), 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 \( H_1, \ldots, H_m \), their number is at most \( 2^m \), which together with (28) implies

\[
m_0 \leq m + 2^m c_{m, \delta}.
\]

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

\[
y_0 := \arg\min \{ \| x_0 - y \| : y \in P \}.
\]

(29)

The task is to show that

\[
\| x_0 - y_0 \| \leq (1 + \varepsilon) \lambda.
\]

(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

\[
y_0 \in \text{relint} F.
\]

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 \( v \in S_F \),

\[
\langle x_0 - y_0, v \rangle \geq (1 - \delta) \| x_0 - y_0 \|.
\]

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

\[
\langle x_0, v \rangle \leq \langle y_0, v \rangle + \lambda.
\]

Combining the last two inequalities gives

\[
\| x_0 - y_0 \| \leq \frac{\lambda}{1 - \delta}.
\]

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

References


*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**