Math-Net.Ru

Общероссийский математический портал

A. Yu. Zaitsev, Multidimensional version of a result of Sakhanenko in the invariance principle for vectors with finite exponential moments. I, *Теория вероятн. и ее примен.*, 2000, том 45, выпуск 4, 718–738

DOI: <https://doi.org/10.4213/tvp500>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.197.217.227

26 апреля 2023 г., 11:54:55

© 2000 г.

ZAITSEV A. YU.*

MULTIDIMENSIONAL VERSION OF A RESULT OF SAKHANENKO IN THE INVARIANCE PRINCIPLE FOR VECTORS WITH FINITE EXPONENTIAL MOMENTS. I¹⁾

Получен многомерный вариант результата Саханенко [18] о гауссовской аппроксимации последовательности сумм независимых случайных векторов с конечными экспоненциальными моментами.

Ключевые слова и фразы: многомерный принцип инвариантности, сильная аппроксимация, суммы независимых разнораспределенных случайных векторов.

1. Introduction. The aim of this paper is to give a multidimensional generalization of a result of Sakhanenko [18] which is an extension and sharpening of the famous result of Komlós, Major and Tusnády (KMT) [12] to the case of nonidentically distributed random variables. For formulations of results we need the following notation.

Notation 1.1. We write $z \in \mathbf{R}^d$ (or $\mathbf{C}^d$), if $z = (z_1, \dots, z_d) = z_1 e_1 + \dots + z_d e_d$, where $z_j \in \mathbf{R}^1$ (or $\mathbf{C}^1$) and the e_j are the basis vectors. The scalar product of vectors $x, y \in \mathbf{R}^d$ (or $\mathbf{C}^d$) is denoted by $\langle x, y \rangle = x_1 \bar{y}_1 + \dots + x_d \bar{y}_d$. For $z \in \mathbf{R}^d$ (or $\mathbf{C}^d$) we 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 write $\ln^* b = \max\{1, \ln b\}$. The distribution and the corresponding covariance operator of a random vector ξ is denoted by $\mathcal{L}(\xi)$ and $\text{cov } \xi$ (or $\text{cov } F$, if $F = \mathcal{L}(\xi)$). For random vectors $X_1, \dots, X_s$ and $Y_1, \dots, Y_s$ we write

$$\Delta(X, Y) = \max_{1 \leq j \leq s} \left\| \sum_{i=1}^j X_i - \sum_{i=1}^j Y_i \right\|. \quad (1.1)$$

The symbol $\mathbb{I}_d: \mathbf{R}^d \rightarrow \mathbf{R}^d$ is used for the identity operator in $\mathbf{R}^d$. Introduce the projectors $\mathbb{P}_i: \mathbf{R}^d \rightarrow \mathbf{R}^1$ and $\bar{\mathbb{P}}_j: \mathbf{R}^d \rightarrow \mathbf{R}^j$, $i, j = 1, \dots, d$, by the relations $\mathbb{P}_i x = x_i$, $\bar{\mathbb{P}}_j x = (x_1, \dots, x_j)$, where $x = (x_1, \dots, x_d) \in \mathbf{R}^d$.

*С.-Петербургское отделение Математического института им. В. А. Стеклова РАН, ул. Фонтанка, 27, 191011 С.-Петербург, Россия; e-mail: zaitsev@pdmi.ras.ru

¹⁾ Research supported by the SFB 343 in Bielefeld, by the Russian Foundation of Basic Research, grant 96-01-00672, by grant INTAS-RFBR 95-0099 and by grant RFBR-DFG 96-01-00096-ge.

The rate for strong approximation in the one-dimensional invariance principle was studied by many authors (see, e.g., [15], [22], [3], [4] and the bibliography in [5], [6], [21]). Skorokhod [22] developed a method of construction of close sequences of sequential sums of independent random variables on the same probability space. For a long time the best rates of approximation were obtained by this method, known now as the Skorokhod embedding. However, Komlós, Major and Tusnády [12] elaborated a new, more powerful method of dyadic approximation. With the help of this method they obtained optimal rates of Gaussian approximation for sequences of independent identically distributed random variables.

Sakhanenko [18] generalized and essentially sharpened KMT results in the case of nonidentically distributed random variables. He considered the following class of one-dimensional distributions:

$$\mathcal{S}_1(\tau) = \left\{ \mathcal{L}(\xi): \mathbf{E}\xi = 0, \mathbf{E}|\xi|^3 \exp(\tau^{-1}|\xi|) \leq \tau \mathbf{E}|\xi|^2 \right\}. \quad (1.2)$$

His main result is formulated as follows.

Theorem 1.1 (see [18]). *Suppose that $\tau > 0$, and $\xi_1, \dots, \xi_s$ are independent random variables with $\mathcal{L}(\xi_j) \in \mathcal{S}_1(\tau)$, $j = 1, \dots, s$. Then one can construct on a probability space a sequence of independent random variables $X_1, \dots, X_s$ and a sequence of independent Gaussian random variables $Y_1, \dots, Y_s$ so that $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}Y_j = 0$, $\mathbf{E}Y_j^2 = \mathbf{E}X_j^2$, $j = 1, \dots, s$, and*

$$\mathbf{E} \exp \left(\frac{c_1 \Delta(X, Y)}{\tau} \right) \leq 1 + \frac{B}{\tau}, \quad (1.3)$$

where c_1 is an absolute constant and $B^2 = \mathbf{E}\xi_1^2 + \dots + \mathbf{E}\xi_s^2$.

Multidimensional estimates in the strong invariance principle can be found in [2], [14], [1], [7]–[9]. The author [29]–[31] removed an unnecessary logarithmic factor from the result of Einmahl [9] and obtained multidimensional analogs of KMT results. In [31, Theorem 1.3] the random vectors are, generally speaking, nonidentically distributed. However, they have the same covariance operator $\mathbb{I}_d$. Therefore, the problem of obtaining an adequate multidimensional generalization of the main result of Sakhanenko [18] remained open. This generalization is given in Theorem 1.2 below.

Definition 1.1 (see [26]). Denote by $\mathcal{A}_d(\tau)$, $\tau \geq 0$, the class of probability distributions which consists of d -dimensional distributions F for which the function

$$\varphi(z) = \varphi(F, z) = \ln \int_{\mathbf{R}^d} e^{\langle z, x \rangle} F\{dx\} \quad (\varphi(0) = 0) \quad (1.4)$$

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

$$|d_u d_v^2 \varphi(z)| \leq \|u\| \tau \langle \mathbb{D}v, v \rangle \quad (1.5)$$

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

In Section 2 we consider the properties of classes $\mathcal{A}_d(\tau)$ and describe the relation of $\mathcal{A}_d(\tau)$ to other classes of probability distributions.

For the formulation of Theorem 1.2 we need the following Condition $\mathbf{B}_2$ (which is in fact one of Conditions $\mathbf{B}_1$ – $\mathbf{B}_4$ imposed on the dyadic scheme, see Section 6 in Part II).

Condition $\mathbf{B}_2$. We say that the covariance operators $\mathbb{B}_1, \dots, \mathbb{B}_l$ satisfy Condition $\mathbf{B}_2$ if for all $u \in \mathbf{R}^d$ and for all $p = 1, \dots, l$

$$C_1^2 \|u\|^2 \leq \langle \mathbb{B}_p u, u \rangle \leq C_2^2 \|u\|^2 \quad (1.6)$$

with some constants C_1 and C_2 .

The following theorem is the main result of the paper.

Theorem 1.2. Let $\alpha > 0$, $\tau \geq 1$, $D > 0$ and let $\xi_1, \dots, \xi_s$ be independent random vectors with $\mathbf{E}\xi_j = 0$, $j = 1, \dots, s$. Assume that there exists a strictly increasing sequence of non-negative integers $m_0 = 0, m_1, \dots, m_l = s$ satisfying the following conditions. Write

$$\zeta_p = D (\xi_{m_{p-1}+1} + \xi_{m_{p-1}+2} + \dots + \xi_{m_p}), \quad p = 1, \dots, l, \quad (1.7)$$

and suppose that the covariance operators $\text{cov } \zeta_p = \mathbb{B}_p$ satisfy Condition $\mathbf{B}_2$ and $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$ (for all $p = 1, \dots, l$). Then one can construct on a probability space a sequence of independent random vectors $X_1, \dots, X_s$ and a sequence of independent Gaussian random vectors $Y_1, \dots, Y_s$ so that $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}Y_j = 0$, $\text{cov } Y_j = \text{cov } X_j$, $j = 1, \dots, s$, and

$$\mathbf{E} \exp \left(\frac{a_1 D \Delta(X, Y)}{\tau d^{9/2} \ln^* d} \right) \leq \exp \left(a_2 d^{3+\alpha} \ln^* \left(\frac{l}{\tau^2} \right) \right), \quad (1.8)$$

where a_1, a_2 are positive quantities depending only on α, C_1, C_2 . Moreover, if all moments of the third order of the vectors ζ_p are equal to zero, then this assertion is valid without the factor $\ln^* d$ in the denominator of the fraction in the left-hand side of (1.8).

The main result of [31] (see Theorem 1.3 therein) can be obtained if one takes in Theorem 1.2 and in Condition $\mathbf{B}_2$ $l = s$, $m_p = p$, $\zeta_p = \xi_p$, and $D = C_1 = C_2 = 1$ (hence, $\mathbb{B}_p = \mathbb{I}_d$, $p = 1, \dots, s$). Note that the dependence of constants on the dimension d in [31, Theorem 1.3] is somewhat better due to the fact that the proof of Theorem 1.2 is more complicated.

In Theorem 1.2 we give an explicit form of the dependence of constants on the distributions of summands and on the dimension d . It is clear that applying Theorem 1.2 one should minimize the parameter τ in (1.8). Moreover, one can try to optimize the choice of ζ_p and D in (1.7). For example, let the conditions of Theorem 1.2 be satisfied with $D = 1$, $\tau = 2^{k/2}$ and

$l = 2^r$, where $k, r, k \leq r$, are positive integers. Then one can enlarge the blocks of summands putting $\zeta_i^* = D^*(\zeta_{(i-1) \cdot 2^{k+1}} + \zeta_{(i-1) \cdot 2^{k+2}} + \dots + \zeta_{i \cdot 2^k})$, $i = 1, \dots, 2^{r-k}$, where $D^* = 2^{-k/2} = \tau^{-1}$. Write $l^* = l/\tau^2 = 2^{r-k}$, $\tau^* = 1$. Using simple properties of classes $\mathcal{A}_d(\tau)$ (see Section 2), we see that the conditions of Theorem 1.2 are satisfied with the change of ζ_p, D, l, τ by $\zeta_i^*, D^*, l^*, \tau^*$ respectively. An application of Theorem 1.2 gives the same inequality (1.8) as before. On the other hand, even if the conditions of Theorem 1.2 are satisfied for $\zeta_i^*, D^*, l^*, \tau^*$, they can be not valid for ζ_p, D, l, τ .

The conditions of Theorem 1.2 are expressed in terms of distributions of sums over blocks of summands. A simple sufficient condition for the validity of the relations $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$, $p = 1, \dots, l$, is given by the relations $\mathcal{L}(\xi_j) \in \mathcal{A}_d(\tau/D)$, $j = 1, \dots, s$ (see Section 2). On the other hand, the relations $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$, $p = 1, \dots, l$, do not imply that $\mathcal{L}(\xi_j) \in \mathcal{A}_d(\tau/D)$, $j = 1, \dots, s$.

The following Theorem 1.3 will be used in the proof of Theorem 1.2.

Theorem 1.3. Suppose that $\tau \geq 1$, $\alpha > 0$ and $\tilde{\xi}_1, \dots, \tilde{\xi}_n$ are random vectors with distributions $\mathcal{L}(\tilde{\xi}_k) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\tilde{\xi}_k = 0$. Let covariance operators $\text{cov } \tilde{\xi}_k = \mathbb{B}_k$, $k = 1, \dots, n$, satisfy Condition $\mathbf{B}_2$. Assume that there exist random vectors $\tilde{\eta}_k$ which have the same moments of the first three orders as the vectors $\tilde{\xi}_k$, $\mathcal{L}(\tilde{\eta}_k) \in \mathcal{A}_d(\tau)$ and $\mathbf{P}\{\|\tilde{\eta}_k\| \leq \lambda\} = 1$, $k = 1, \dots, n$, where $\lambda \geq \max\{C_2, 1\}\sqrt{d}$. Then one can construct on a probability space a sequence of independent random vectors $\tilde{X}_1, \dots, \tilde{X}_n$ and a sequence of independent Gaussian random vectors $Y_1, \dots, Y_n$ so that $\mathcal{L}(\tilde{X}_k) = \mathcal{L}(\tilde{\xi}_k)$, $\mathbf{E}Y_k = 0$, $\text{cov } Y_k = \text{cov } \tilde{X}_k$, $k = 1, \dots, n$, and

$$\mathbf{E} \exp \left(\frac{a_3 \Delta(\tilde{X}, Y)}{\lambda d^4 \tau^2} \right) \leq \exp(a_4 (\lambda d^{3/2} \tau)^{3/2+\alpha} \ln^* n), \quad (1.9)$$

where a_3, a_4 are positive quantities depending only on α, C_1, C_2 .

It is well known that the results similar to Theorem 1.2 imply many useful consequences. For example, one can derive estimates in the invariance principle in the case when the summands X_j satisfy less restrictive moment conditions (see, e.g., [21] and the bibliography therein). The author is going to devote to the consequences of Theorem 1.2 a separate publication.

The proof of Theorems 1.2 and 1.3 consists of several steps, many of which are close to corresponding steps from the proofs of the main results of [12], [18], [9], [29], [31]. In Sections 2 and 3 we introduce the necessary notation and formulate auxiliary results about some classes of probability distributions closely connected with classes $\mathcal{A}_d(\tau)$. In particular, we formulate key Lemmas 3.2 and 3.3 which were proved in [29]–[31] and contain the estimates for quantiles of conditional distributions of the last coordinates of the vectors which have sufficiently smooth distributions from $\mathcal{A}_d(\tau)$.

Section 4 is devoted to well-known facts about conditional distributions which are necessary for justifying the dyadic construction.

In Sections 5, 6–7 (Part II) we describe a modification of the well-known dyadic construction due to Komlós, Major, and Tusnády [12], one-dimensional case, and to [9], multidimensional case. The construction in the present paper differs from that of [29], [31] since now we have to work with random vectors having not coinciding covariance operators. Therefore we introduce in the construction new linear operators $A_{n,k}$, $D_{n,k}$, $\tilde{D}_{n,k}$, $J_{n,k}$, $\tilde{J}_{n,k}$ (see (5.5), (6.3) and (6.5)). This is important for applying Lemmas 3.2 and 3.3 because these lemmas can be applied only in the case when the random variable to be constructed is noncorrelated with random variables which are already constructed. In the case when all summands have the identity covariance operators I_d the operators $A_{n,k}$, $D_{n,k}$, $\tilde{D}_{n,k}$, $J_{n,k}$, $\tilde{J}_{n,k}$ «disappear» since then $A_{n,k} = A^\pm$ (see Remark 6.5), $D_{n,k} = \tilde{D}_{n,k} = 2^n I_d$, $J_{n,k} = \tilde{J}_{n,k} = I_d$. The sufficiently complicated formulas (6.5) turn into the well-known relation $\tilde{U}_{n,k} = \tilde{U}_{n,k} = U_{n-1,2k-1} - U_{n-1,2k}$ which was traditionally used in the dyadic constructions. We assume that the operators mentioned above satisfy Conditions B_1 – B_4 (see Section 6) which we call for the sake of brevity Condition **B** (cf. Condition B_2 above). The crucial role in the proof of Theorems 1.2 and 1.3 is played by Lemma 6.1. In the proof of this lemma the most important formula is (6.47) in which we use several times the relations of type $D_{n,k}^{-1} D_{n,k} = I_d$.

In Section 7 we introduce a simple transformation of probability distributions $\mathcal{L}(X)$ which is denoted by $\Psi(\mathcal{L}(X))$ (see Notation 7.2). Such a transformation was already used in [29], [31]. The distributions $\Psi(\mathcal{L}(X))$ have the same moments of the first three orders as $\mathcal{L}(X)$. Moreover, the distribution $\Psi(\mathcal{L}(X))$ is «better» than $\mathcal{L}(X)$ because it is sufficiently smooth. Then we introduce Conditions $A(\tau)$, $A^*(\tau)$, $C_\Psi(p)$, D_Φ , D_Ψ . These conditions describe the distributions of the vectors involved in the dyadic scheme (which will be applied several times in different situations). For example, Condition D_Φ means that we start the dyadic construction having the Gaussian vectors being constructed. Condition D_Ψ says that the initial vectors have bounded and smooth distributions $\Psi(\mathcal{L}(\tilde{\eta}_j))$ (see Condition B_1 and Remark 6.2). Condition $C_\Psi(p)$ describes the vectors which are to be constructed. It says that, in each block of summands of size 2^p , the first vector has smooth distribution $\Psi(\mathcal{L}(\tilde{\eta}_j))$ (other vectors have the required distributions). The presence of smooth distributions ensures the existence of smooth densities for distributions of sums over blocks and, hence, the possibility to apply the estimates for quantiles of conditional distributions from Lemmas 3.2 and 3.3. Conditions $A(\tau)$ and $A^*(\tau)$ yield that belong distributions of summands and sums over blocks belong to some subclasses of $\mathcal{A}_d(\tau)$.

In Section 8 we prove Theorems 8.1 and 8.2. Theorem 8.1 contains KMT-type estimate under Conditions **B**, D_Φ and $A^*(\tau)$ for sufficiently small τ . The last condition means that the distributions of sums over blocks

of summands are sufficiently smooth, belong to $\mathcal{A}_d(\tau)$ and, hence, are close to Gaussian ones. Namely, it is obvious that the class $\mathcal{A}_d(0)$ coincides with the class of all d -dimensional Gaussian distributions. The following inequality was proved in [26] and can be considered as an estimate of stability of this characterization: if $F \in \mathcal{A}_d(\tau)$, $\tau > 0$, then

$$\pi(F, \Phi(F)) \leq cd^2\tau \ln^*(\tau^{-1}), \quad (1.10)$$

where $\pi(\cdot, \cdot)$ is the Prokhorov distance and $\Phi(F)$ denotes the Gaussian distribution whose mean and covariance operator coincide with those of F . Moreover, in [26] it was established that for all d -dimensional Borel sets X and all $\lambda > 0$

$$\begin{aligned} F\{X\} &\leq \Phi(F)\{X^\lambda\} + cd^2 \exp\left(-\frac{\lambda}{cd^2\tau}\right), \\ \Phi(F)\{X\} &\leq F\{X^\lambda\} + cd^2 \exp\left(-\frac{\lambda}{cd^2\tau}\right), \end{aligned} \quad (1.11)$$

where $X^\lambda = \{y \in \mathbf{R}^d: \inf_{x \in X} \|x - y\| < \lambda\}$ is the λ -neighborhood of the set X . It is very essential (and important) that the inequalities (1.10), (1.11) are proved for all $\tau > 0$ and for arbitrary cov F , in contrast to Theorems 1.2 and 1.3, where $\tau \geq 1$ and covariance operators satisfy Condition $\mathbf{B}_2$. The question about the necessity of Condition $\mathbf{B}_2$ in Theorems 1.2 and 1.3 remains open. It is clear that the statement of Theorem 8.1 becomes stronger for small τ . Thus, the approximation is better in the case when summands have smooth distributions which are close to Gaussian ones. Theorem 8.1 generalize the main result of Götze and Zaitsev [11] which is proved for the case of identity covariance operators $\mathbf{I}_d$. That is, when $C_1 = C_2 = 1$. In [11] one can find simple examples in which the sufficiently complicated Condition $\mathbf{A}^*(\tau)$ is satisfied.

Theorem 8.2 yields a bound of the rate of strong approximation in the case when $\tau \geq 1$ and the dyadic scheme satisfies Conditions $\mathbf{A}(\tau)$, $\mathbf{B}$, $\mathbf{D}_\Phi$ and $\mathbf{C}_\Psi(0)$. The last condition means that all random vectors which are to be constructed have smooth distributions $\Psi(\mathcal{L}(\tilde{\eta}_j))$. Lemma 7.4 allows us to reduce the proof of Theorem 8.2 to an application of Theorem 8.1.

The scheme of the proof of Theorem 1.2 is very close to that of the main results of [18], [29], [31]. Theorem 1.2 is an immediate consequence of Theorem 1.3. In the proof of Theorem 1.3 we suppose that the Gaussian vectors $Y_1, \dots, Y_n$, $n = 2^N$, are already constructed and construct the independent vectors which are bounded with probability one, have sufficiently smooth distributions and the same moments of the first, second and third orders as the needed independent random vectors $\tilde{X}_1, \dots, \tilde{X}_n$. The rate of approximation is estimated in Theorem 8.2. Then we construct the vectors $\tilde{X}_1, \dots, \tilde{X}_n$ in several steps. After each step the number of $\tilde{X}_k$ which are not

constructed becomes smaller in 2^p times, where p is a suitably chosen positive integer. In each step we begin with already constructed vectors which are bounded with probability one and have sufficiently smooth distributions and the needed moments up to the third order. Then we construct the vectors such that, in each block of 2^p summands, only the first vector has the initial bounded smooth distribution. The rest $2^p - 1$ vectors have the needed distributions $\mathcal{L}(\tilde{\xi}_k)$. These $2^p - 1$ vectors from each block will be chosen as $\tilde{X}_k$ and will not be involved in the next steps of the procedure. The coincidence of third moments will allow us to use more precise estimates of the closeness of quantiles of conditional distributions contained in Lemma 3.3. Section 9 is devoted to the estimation of closeness of random vectors in the steps of the procedure described above. The proof of Theorem 1.3 is given in Section 10. Theorem 1.2 is derived from Theorem 1.3 in Section 11. At the end of Section 11 we show that Theorem 1.1 is an easy consequence of Theorem 1.2.

2. Notation and properties of classes $\mathcal{A}_d(\tau)$.

Notation 2.1. Let $\mathfrak{F}_d$ be the set of all d -dimensional probability distributions defined on the σ -algebra $\mathfrak{B}_d$ of Borel subsets of the Euclidean space $\mathbf{R}^d$. Let $\mathfrak{G}_d$ be the collection of all d -dimensional Gaussian distributions. The symbols $c, c_1, c_2, \dots$ will be used for absolute positive constants. The letter c can denote different constants when we do not need to fix their numerical values. The same notation will be used for positive quantities which depend only on C_1 and C_2 involved in Condition $\mathbf{B}_2$. This means, in fact, that we assume C_1 and C_2 to be absolute positive constants. By $\hat{F}(t)$, $t \in \mathbf{R}^d$, we denote the characteristic function of a distribution $F \in \mathfrak{F}_d$. The product of measures is understood as their convolution: $FG = F * G$. In one-dimensional case we denote by $\Phi_\sigma(\cdot)$ the distribution function of the normal law with mean zero and variance σ^2 .

Let $\tau \geq 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $\|h\|\tau < 1$, $h \in \mathbf{R}^d$. Then the conjugate distribution $\bar{F} = \bar{F}(h)$ is defined by

$$\bar{F}\{dx\} = (\mathbf{E}e^{(h, \xi)})^{-1} e^{(h, x)} F\{dx\} \quad (2.1)$$

(sometimes it is called the Cramér transform). It is clear that

$$\begin{aligned} \text{if } F_1, \dots, F_n \in \mathcal{A}_d(\tau), \quad \|h\|\tau < 1, \quad F = \prod_{j=1}^n F_j, \\ \text{then } \bar{F}(h) = \prod_{j=1}^n \bar{F}_j(h). \end{aligned} \quad (2.2)$$

Lemma 2.1 ([26, Lemma 2.1]). Suppose that $\tau \geq 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $h \in \mathbf{R}^d$, $\|h\|\tau < 1$, $\mathbb{D} = \text{cov } F$, $\mathbb{D}(h) = \text{cov } \bar{F}(h)$, $\mathbf{E}\xi = 0$. Then

a) for all $u, v \in \mathbf{R}^d$ the following relations are valid:

$$\langle \mathbb{D}(h) u, u \rangle = \langle \mathbb{D}u, u \rangle (1 + \theta \|h\| \tau), \quad (2.3)$$

$$\ln \mathbf{E} e^{i\langle h, \xi \rangle} = -\frac{1}{2} \langle \mathbb{D}h, h \rangle \left(1 + \frac{1}{3} \theta \|h\| \tau \right) \quad (2.4)$$

(here θ symbolizes different quantities not exceeding one in absolute value);

b) if $\|h\| \tau \leq \frac{1}{2}$, then $\overline{F}(h) \in \mathcal{A}_d(2\tau)$.

The classes $\mathcal{A}_d(\tau)$ have very convenient properties. Some of them were considered in [26], [30], [31]. In particular, if $0 \leq \tau_1 \leq \tau_2$, then $\mathcal{A}_d(\tau_1) \subset \mathcal{A}_d(\tau_2)$. Moreover, it is easy to see that, for fixed τ and d , the class $\mathcal{A}_d(\tau)$ is closed with respect to convolution: if $F_1, F_2 \in \mathcal{A}_d(\tau)$, then $F_1 F_2 \in \mathcal{A}_d(\tau)$.

The classes $\mathcal{A}_d(\tau)$ are closely connected with other naturally defined classes of multidimensional distributions. From the definition of $\mathcal{A}_d(\tau)$ it follows that if $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$ then the vector ξ has finite exponential moments $\mathbf{E} e^{i\langle h, \xi \rangle}$, $h \in \mathbf{R}^d$, $\|h\| \tau < 1$. This leads to exponential estimates for the tails of distributions (see, e.g., Lemma 2.6).

The condition $\mathcal{L}(\xi) \in \mathcal{A}_1(c\tau)$ is equivalent to Statulevičius' [23] conditions

$$|\gamma_m| \leq cm! \tau^{m-2} \gamma_2, \quad m = 3, 4, \dots, \quad (2.5)$$

on the rate of increasing of cumulants γ_m of the random variable ξ . This equivalence means that if one of these conditions is satisfied with parameter τ , then the second is valid with parameter $c\tau$. However, the condition $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$ differs from other multidimensional analogs of Statulevičius' conditions, considered by Rudzkis [17] and Saulis [19].

Another class of distributions $\tilde{\mathcal{A}}_d(\tau)$, $\tau \geq 0$, which is equivalent to $\mathcal{A}_d(\tau)$ is defined similarly to $\mathcal{A}_d(\tau)$, with the change of (1.5) by

$$|d_v^2 \varphi(z)| \leq 2 \langle \mathbb{D}v, v \rangle \quad (2.6)$$

for all $v \in \mathbf{R}^d$ and $z \in \mathbf{C}^d$, $\|z\| \tau < 1$. One can show that $\mathcal{A}_d(\tau) \subset \tilde{\mathcal{A}}_d(c\tau)$, $\tilde{\mathcal{A}}_d(\tau) \subset \mathcal{A}_d(c\tau)$. This fact will not be used in this paper. Its proof is elementary and, therefore, is omitted.

An analog of the inequalities (1.10), (1.11) was proved in [25], [27] for the convolutions of distributions from the class

$$\mathcal{B}_d(\tau) = \left\{ F = \mathcal{L}(\xi) \in \mathfrak{F}_d: \mathbf{E}\xi = 0, |\mathbf{E}\langle \xi, v \rangle|^2 \langle \xi, u \rangle^{m-2} \leq \frac{1}{2} m! \tau^{m-2} \right. \\ \left. \times \|u\|^{m-2} \mathbf{E}\langle \xi, v \rangle^2 \text{ for all } u, v \in \mathbf{R}^d, m = 3, 4, \dots \right\}. \quad (2.7)$$

It can be easily verified that all distributions with zero mean, which are concentrated on the ball $\{x \in \mathbf{R}^d: \|x\| \leq \tau\}$, belong to $\mathcal{B}_d(\tau)$. In the one-dimensional case the condition $F \in \mathcal{B}_1(\tau)$ coincides with the conditions of the well-known Bernstein inequality. As is shown in [25], the classes $\mathcal{B}_1(\tau)$

and $\mathcal{S}_1(\tau)$ are equivalent (see Lemma 2.3 for the d -dimensional analog of this statement). The following Lemma 2.2 shows the connection between classes $\mathcal{A}_d(\tau)$ and $\mathcal{B}_d(\tau)$.

Lemma 2.2 (see [26], [28]). *Let $B^2 = B^2(F)$ be the maximal eigenvalue of the covariance operator of a distribution F . Then there exist absolute constants c_2 and c_3 such that the following assertions are valid.*

- a) *If $F = \mathcal{L}(\xi) \in \mathcal{B}_d(\tau)$, then $B^2 \leq 12\tau^2$, $\mathbf{E}\xi = 0$ and $F \in \mathcal{A}_d(c_2\tau)$.*
- b) *If $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $B^2 \leq \tau^2$ and $\mathbf{E}\xi = 0$, then $F \in \mathcal{B}_d(c_3\tau)$.*

Introduce the following analog of the class $\mathcal{B}_d(\tau)$:

$$\mathcal{S}_d(\tau) = \left\{ F = \mathcal{L}(\xi) \in \mathfrak{F}_d: \mathbf{E}\xi = 0, \mathbf{E}|\langle \xi, v \rangle \langle \xi, u \rangle e^{i\langle \xi, u \rangle}| \leq \mathbf{E}\langle \xi, v \rangle^2 \right. \\ \left. \text{for all } u, v \in \mathbf{R}^d \text{ such that } \|u\| \leq \tau^{-1} \right\}. \quad (2.8)$$

Obviously, for $d = 1$, the class $\mathcal{S}_d(\tau)$ turns into the Sakhanenko class $\mathcal{S}_1(\tau)$ defined in (1.2). The following elementary Lemma 2.3 is contained in [25].

Lemma 2.3. *The classes $\mathcal{B}_d(\tau)$ and $\mathcal{S}_d(\tau)$ are equivalent in the following sense: $\mathcal{B}_d(\tau) \subset \mathcal{S}_d(7\tau)$ and $\mathcal{S}_d(\tau) \subset \mathcal{B}_d(\tau)$, for all $\tau \geq 0$.*

As examples of distributions from $\mathcal{A}_d(c\tau)$ one can consider infinitely divisible distributions with spectral measures concentrated on the ball $\{x \in \mathbf{R}^d: \|x\| \leq \tau\}$ (see [26]). One can find therein a more general condition which ensure an infinitely divisible distribution belongs to $\mathcal{A}_d(c\tau)$ (see [26, condition (1.8)]).

Assume, in particular, that $H = \mathcal{L}(\zeta) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\zeta = 0$, and B^2 is the maximal eigenvalue of the operator $\text{cov } \zeta$. Suppose that $\lambda > 0$ and an infinitely divisible distribution F is defined by its characteristic function

$$\widehat{F}(t) = \exp\left(\lambda \int_{\mathbf{R}^d} (e^{i\langle t, x \rangle} - 1) H\{dx\}\right), \quad t \in \mathbf{R}^d.$$

Then it can be verified that $F \in \mathcal{A}_d(c\tau_0)$, where $\tau_0 = \max\{\tau, B\}$.

The following elementary Lemmas 2.4–2.6 are proved in [31].

Lemma 2.4. *Let $\tau \geq 0$, $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $y \in \mathbf{R}^m$, $\alpha \in \mathbf{R}^1$, and let $\mathbf{A}: \mathbf{C}^d \rightarrow \mathbf{C}^m$ be a linear operator such that $\mathbf{A}(\mathbf{R}^d) \subset \mathbf{R}^m$. Let $\tilde{\xi} \in \mathbf{R}^k$ be the vector consisting of a subset of coordinates of the vector ξ . Then*

$$\mathcal{L}(\mathbf{A}\xi + y) \in \mathcal{A}_m(\|\mathbf{A}\|\tau), \quad \text{where } \|\mathbf{A}\| = \sup_{x \in \mathbf{R}^d, \|x\| \leq 1} \|\mathbf{A}x\|, \\ \mathcal{L}(\alpha\xi) \in \mathcal{A}_d(|\alpha|\tau), \quad \mathcal{L}(\tilde{\xi}) \in \mathcal{A}_k(\tau).$$

Lemma 2.5. *Suppose that $\tau \geq 0$, the vectors $\xi^{(k)}$, $k = 1, 2$, are independent, and $\mathcal{L}(\xi^{(k)}) \in \mathcal{A}_{d_k}(\tau)$. Let $\xi \in \mathbf{R}^{d_1+d_2}$ be the vector with the first d_1 coordinates coinciding with those of $\xi^{(1)}$ and with the last d_2 coordinates coinciding with those of $\xi^{(2)}$. Then $\mathcal{L}(\xi) \in \mathcal{A}_{d_1+d_2}(\tau)$.*

Lemma 2.6. Let $\tau \geq 0$ and let $X_1, \dots, X_n$ be independent random vectors with $\mathbf{E}X_i = 0$, $S_i = X_1 + \dots + X_i$, for $i = 1, \dots, n$. Let $h, t \in \mathbf{R}^1$, $0 \leq h\tau < 1$, $0 \leq t\tau < \frac{1}{2}$, $\xi = \max_{1 \leq i \leq n} |S_i|$. Assume that $\mathcal{L}(S_n) \in \mathcal{A}_d(\tau)$. Denote by B^2 the maximal eigenvalue of $\text{cov } S_n$. Then

$$\mathbf{E}e^{h|S_n|} \leq 2d e^{h^2 B^2}, \quad \mathbf{E}e^{t\xi} \leq 3d e^{4t^2 B^2}$$

and

$$\mathbf{P}\{\xi \geq x\} \leq 2d \max \left\{ \exp \left(-\frac{x^2}{4B^2} \right), \exp \left(-\frac{x}{4\tau} \right) \right\}, \quad x \geq 0.$$

3. Classes $\mathcal{A}_d^*(\tau, \delta, \rho)$, $\mathcal{A}_d^{}(\tau, \delta, \rho)$, $\overline{\mathcal{A}}_d(\tau, \rho)$, and $\mathcal{R}_d(\tau)$ and some of their properties.** In the sequel we shall use the following subclasses of $\mathcal{A}_d(\tau)$ containing distributions satisfying some special smoothness-type restrictions.

Definition 3.1. Let $\tau \geq 0$, $\delta > 0$, $\rho > 0$, $h \in \mathbf{R}^d$. Consider the conditions:

$$(2\pi)^{-d} \int_{\rho \|t\| \tau d \geq 1} |\hat{F}_h(t)| dt \leq \frac{\tau d^{3/2}}{\sigma (2\pi)^{d/2} (\det \mathbb{D})^{1/2}}, \quad (3.1)$$

$$(2\pi)^{-d} \int_{\rho \|t\| \tau d \geq 1} |\hat{F}_h(t)| dt \leq \frac{\tau^2 d^2}{\sigma^2 (2\pi)^{d/2} (\det \mathbb{D})^{1/2}}, \quad (3.2)$$

$$(2\pi)^{-d} \int_{\rho \|t\| \tau d \geq 1} |\langle t, v \rangle \hat{F}_h(t)| dt \leq \frac{\langle \mathbb{D}^{-1} v, v \rangle^{1/2}}{\delta (2\pi)^{d/2} (\det \mathbb{D})^{1/2}}, \quad \text{for all } v \in \mathbf{R}^d, \quad (3.3)$$

where $F_h = \overline{F}(h)$, $\mathbb{D} = \mathbb{D}(F) = \text{cov } F$ and $\sigma^2 = \sigma^2(F) > 0$ is the minimal eigenvalue of $\mathbb{D}$. Denote by $\mathcal{A}_d^*(\tau, \delta, \rho)$; $\mathcal{A}_d^{**}(\tau, \delta, \rho)$; and $\overline{\mathcal{A}}_d(\tau, \rho)$ the classes of distributions $F \in \mathcal{A}_d(\tau)$ such that the conditions (3.2) and (3.3); (3.1) and (3.3); and (3.1), respectively, are satisfied for $h \in \mathbf{R}^d$, $\|h\|\tau < 1$.

The classes just defined are closely connected, in particular,

$$\mathcal{A}_d^*(\tau, \delta, \rho) \subset \mathcal{A}_d^{**}(\tau, \delta, \rho), \quad \text{if } \frac{\tau d^{1/2}}{\sigma} \leq 1, \quad (3.4)$$

$$\mathcal{A}_d^{**}(\tau, \delta, \rho) \subset \overline{\mathcal{A}}_d(\tau, \rho). \quad (3.5)$$

It is easy to see that (if $F = \mathcal{L}(\xi)$)

$$\hat{F}_h(t) = (\mathbf{E}e^{\langle h, \xi \rangle})^{-1} \mathbf{E}e^{\langle h+it, \xi \rangle}. \quad (3.6)$$

Therefore, the conditions (3.1)–(3.3) play in this paper the role which is similar to that of [18, inequality (49), p. 9] or [9, inequality (1.5)]. Note,

however, that the condition (3.3) is needed for estimating the derivatives of densities which is necessary to obtain more precise bounds for the closeness of conditional distributions when the compared distribution have the same third moments (cf. Lemmas 3.2 and 3.3).

Lemma 3.1. *Let $\tau \geq 0$, $\delta > 0$, $\rho > 0$, $\alpha > 0$ and $F = \mathcal{L}(\xi) \in \mathcal{A}_d^*(\tau, \delta, \rho)$. Then $G = \mathcal{L}(\alpha\xi) \in \mathcal{A}_d^*(\alpha\tau, \delta, \rho)$.*

The proof of this lemma is elementary and therefore is omitted. The following Lemmas 3.2 and 3.3 were proved in [29]–[31].

Lemma 3.2. *Let the distribution of a random vector $\xi \in \mathbf{R}^d$ with $\mathbf{E}\xi = 0$ satisfy the condition $\mathcal{L}(\xi) \in \overline{\mathcal{A}}_d(\tau, 4)$, $\tau \geq 0$. Assume that the variance $\sigma^2 = \mathbf{E}\xi_d^2 > 0$ of the last coordinate ξ_d of the vector ξ is the minimal eigenvalue of $\text{cov } \xi$. Then there exist absolute positive constants $c_4, \dots, c_8$ such that the following assertions hold.*

a) *Let $d \geq 2$. Assume that ξ_d is not correlated with previous coordinates $\xi_1, \dots, \xi_{d-1}$ of the vector ξ . Define $\mathbb{B} = \text{cov } \bar{\mathbb{P}}_{d-1}\xi$ and denote by $F(z|x)$, $z \in \mathbf{R}^1$, the regular conditional distribution function (r.c.d.f.) of ξ_d for a given value of $\bar{\mathbb{P}}_{d-1}\xi = x \in \mathbf{R}^{d-1}$. Let $\mathcal{L}(\bar{\mathbb{P}}_{d-1}\xi) \in \overline{\mathcal{A}}_{d-1}(\tau, 4)$. Then there exists $y \in \mathbf{R}^1$ such that*

$$|y| \leq c_4 \tau \|\mathbb{B}^{-1/2}x\|^2 \leq c_4 \tau \frac{\|x\|^2}{\sigma^2}, \quad (3.7)$$

and

$$\Phi_\sigma(z - \gamma(z)) < F(z + y|x) < \Phi_\sigma(z + \gamma(z)), \quad (3.8)$$

provided that $\tau d^{3/2}/\sigma \leq c_5$, $\|\mathbb{B}^{-1/2}x\| \leq (c_6\sigma)/(d^{3/2}\tau)$, $|z| \leq c_7\sigma^2/(d\tau)$, where

$$\gamma(z) = c_8\tau \left(d^{3/2} + d\delta \left(1 + \frac{|z|}{\sigma} \right) + \frac{z^2}{\sigma^2} \right), \quad \delta = \|\mathbb{B}^{-1/2}x\|. \quad (3.9)$$

b) *The assertion a) remains valid for $d = 1$ with $F(z|x) = \mathbf{P}\{\xi < z\}$ and $y = \delta = 0$ without any restrictions on $\mathbb{B}$, $\bar{\mathbb{P}}_{d-1}\xi$ and x .*

Definition 3.2. For $\tau \geq 0$ denote by $\mathcal{R}_d(\tau)$ the collection of pairs of d -dimensional probability distributions (F_1, F_2) which have identical cumulants of second and third orders and $F_k \in \mathcal{A}_d(\tau)$, $k = 1, 2$.

Lemma 3.3. *Let the distributions of random vectors $\xi_1, \xi_2 \in \mathbf{R}^d$ with $\mathbf{E}\xi_1 = \mathbf{E}\xi_2 = 0$ satisfy the conditions $(\mathcal{L}(\xi_1), \mathcal{L}(\xi_2)) \in \mathcal{R}_d(\tau)$, and $\mathcal{L}(\xi_k) \in \mathcal{A}_d^*(\tau, 4, 4)$, for $k = 1, 2$, $\tau > 0$. Assume that the variance $\sigma^2 = \mathbf{E}\xi_{k,d}^2 > 0$ of the last coordinate $\xi_{k,d}$ of the vector ξ_k is the minimal eigenvalue of $\text{cov } \xi_1 = \text{cov } \xi_2$. Then there exist absolute positive constants $c_9, \dots, c_{14}$ such that the following assertions hold.*

a) *Let $d \geq 2$. Assume that $\xi_{k,d}$ is not correlated with previous coordinates $\xi_{k,1}, \dots, \xi_{k,d-1}$ of the vector ξ_k . Define $\mathbb{B} = \text{cov } \bar{\mathbb{P}}_{d-1}\xi_1 = \text{cov } \bar{\mathbb{P}}_{d-1}\xi_2$ and*

denote by $F_k(z | x_k)$, $z \in \mathbf{R}^1$, the regular conditional distribution function of $\xi_{k,d}$ for given values of $\bar{\mathbb{P}}_{d-1}\xi_k = x_k \in \mathbf{R}^{d-1}$. Let

$$\mathcal{L}(\bar{\mathbb{P}}_{d-1}\xi_k) \in \mathcal{A}_{d-1}^*(\tau, 4, 4), \quad \text{for } k = 1, 2.$$

Then there exist $y_1, y_2 \in \mathbf{R}^1$ such that

$$|y_k| \leq c_9 \tau \|\mathbb{B}^{-1/2} x_k\|^2 \leq c_9 \tau \frac{\|x_k\|^2}{\sigma^2}, \quad \text{for } k = 1, 2, \quad (3.10)$$

$$|y_1 - y_2| \leq c_{10} (\tau^2 \sigma^{-1} \Gamma^3 + \tau \Lambda \Gamma), \quad (3.11)$$

and

$$F_1(z - \kappa(z) + y_1 | x_1) < F_2(z + y_2 | x_2) < F_1(z + \kappa(z) + y_1 | x_1), \quad (3.12)$$

provided that $\tau d^{3/2}/\sigma \leq c_{11}$, $|z| \leq c_{12} \sigma^2/(d\tau)$, $\|\mathbb{B}^{-1/2} x_k\| \leq c_{13} \sigma/(d^{3/2} \tau)$, $k = 1, 2$, where

$$\kappa(z) = c_{14} \frac{\tau^2}{\sigma} \left(d^2 + \beta \left(d + \frac{|z|}{\sigma} \right) + \frac{|z|^3}{\sigma^3} \right), \quad \beta = \Gamma^2 + \sigma \tau^{-1} \Lambda, \quad (3.13)$$

$$\Lambda = \|\mathbb{B}^{-1/2} x_1 - \mathbb{B}^{-1/2} x_2\|, \quad \Gamma = \max_{k=1,2} \|\mathbb{B}^{-1/2} x_k\|. \quad (3.14)$$

b) The assertion a) remains valid for $d = 1$ with $F_k(z | x_k) = \mathbf{P}\{\xi_k < z\}$ and $y_1 = y_2 = \beta = 0$ without any restrictions on $\mathbb{B}$, $\bar{\mathbb{P}}_{d-1}\xi_k$ and x_k .

R e m a r k 3.1. In [29]–[31] the formulation of Lemma 3.2 is in some sense weaker, see [29], [30, Lemmas 6.1 and 8.1] or [31, Lemma 2.14]. In particular, instead of the conditions

$$\mathcal{L}(\xi) \in \bar{\mathcal{A}}_d(\tau, 4) \quad \text{and} \quad \mathcal{L}(\bar{\mathbb{P}}_{d-1}\xi) \in \bar{\mathcal{A}}_{d-1}(\tau, 4) \quad (3.15)$$

the stronger conditions

$$\mathcal{L}(\xi) \in \mathcal{A}_d^*(\tau, 4, 4) \quad \text{and} \quad \mathcal{L}(\bar{\mathbb{P}}_{d-1}\xi) \in \mathcal{A}_{d-1}^*(\tau, 4, 4) \quad (3.16)$$

are used. However, in the proof of (3.7) and (3.8) only the conditions (3.15) are applied. The conditions (3.16) are necessary for the investigation of quantiles of conditional distributions corresponding to random vectors having coinciding moments up to third order which has been done in [29]–[31] simultaneously with the proof of (3.7) and (3.8), cf. Lemma 3.3.

4. Conditional distributions. In this section we expose some well-known standard facts about conditional distributions (see, e.g., [10, Chapter I, § 3]). The results of this section are present here for the sake of completeness. They will be used to justify the dyadic construction in Section 5.

Lemma 4.1. Suppose that the joint distribution of $\xi \in \mathbf{R}^1$ and $\zeta \in \mathbf{R}^j$ is absolutely continuous with respect to the $(j+1)$ -dimensional Lebesgue measure. Let $G(y|z)$ be the regular conditional distribution function of ξ given $\zeta = z$. Then the random variable $G(\xi|\zeta)$ is independent of ζ and has uniform distribution on the interval $[0, 1]$.

Lemma 4.2. Suppose that the distribution of a random variable $\xi \in \mathbf{R}^1$ is absolutely continuous with respect to the one-dimensional Lebesgue measure. Let $G(y)$ be the distribution function of ξ . Then the random variable $G(\xi)$ has uniform distribution on the interval $[0, 1]$.

Lemma 4.3. Assume that a random variable η has uniform distribution on the interval $[0, 1]$ and is independent of a random vector $\zeta \in \mathbf{R}^j$. Let $G(y|z)$ be the regular conditional distribution function of a random variable ξ given $\zeta = z$ and $G^{-1}(x|z) = \sup\{y: G(y|z) \leq x\}$, $0 < x < 1$. Then the joint distribution $\mathcal{L}((G^{-1}(\eta|\zeta), \zeta))$ coincides with $\mathcal{L}((\xi, \zeta))$.

Lemma 4.4. Assume that a random variable η has uniform distribution on the interval $[0, 1]$. Let $G(y)$ be the distribution function of a random variable ξ and $G^{-1}(x) = \sup\{y: G(y) \leq x\}$, $0 < x < 1$. Then $\mathcal{L}(G^{-1}(\eta)) = \mathcal{L}(\xi)$.

5. Dyadic scheme. General description. Let N be a non-negative integer and $\{\xi_1, \dots, \xi_{2^N}\}$, $\{Y_1, \dots, Y_{2^N}\}$ two collections of independent random vectors. We shall always suppose that the random vectors $\xi_1, \dots, \xi_{2^N}$; $Y_1, \dots, Y_{2^N}$ have some known distributions from $\mathcal{A}_d(c\tau)$. Write $R^{(k)} = \mathcal{L}(\xi_k)$, $1 \leq k \leq 2^N$. Assume that the distributions of $Y_1, \dots, Y_{2^N}$ are absolutely continuous. Write

$$S_0^* = 0; \quad S_k^* = \sum_{l=1}^k \xi_l, \quad 1 \leq k \leq 2^N; \quad (5.1)$$

$$U_{m,k}^* = S_{k \cdot 2^m}^* - S_{(k-1) \cdot 2^m}^*, \quad 1 \leq k \leq 2^{N-m}, \quad 0 \leq m \leq N. \quad (5.2)$$

In particular, $U_{0,k}^* = \xi_k$, $U_{N,1}^* = S_{2^N}^* = \xi_1 + \dots + \xi_{2^N}$. In the sequel we shall use the term *block of summands* for a collection of summands with indices of the form $(k-1) \cdot 2^m + 1, \dots, k \cdot 2^m$, where $1 \leq k \leq 2^{N-m}$, $0 \leq m \leq N$. Thus, $U_{m,k}^*$ is the sum over a block containing 2^m summands. It is obvious that, for fixed m , $0 \leq m \leq N$, the random vectors $U_{m,k}^*$, $1 \leq k \leq 2^{N-m}$, are jointly independent. Introduce the vectors

$$\tilde{U}_{n,k}^* = (U_{n-1,2k-1}^*, U_{n-1,2k}^*) \in \mathbf{R}^{2d}, \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N, \quad (5.3)$$

with the first d coordinates coinciding with those of the vectors $U_{n-1,2k-1}^*$ and with the last d coordinates coinciding with those of the vectors $U_{n-1,2k}^*$. Note that

$$U_{n-1,2k-1}^* + U_{n-1,2k}^* = U_{n,k}^*, \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N. \quad (5.4)$$

Let $A_{n,k}: \mathbf{R}^{2d} \mapsto \mathbf{R}^{2d}$ be nondegenerate linear operators such that the first d coordinates of the vector

$$U_{n,k}^* = A_{n,k} \tilde{U}_{n,k}^* \in \mathbf{R}^{2d}, \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N, \quad (5.5)$$

coincide with those of the vectors $\mathcal{U}_{n,k}^* = J_{n,k} U_{n,k}^*$, where $J_{n,k}: \mathbf{R}^d \mapsto \mathbf{R}^d$ are also some nondegenerate linear operators. Write, for simplicity,

$$\mathcal{U}^* = \mathcal{U}_{N,1}^* = J_{N,1} U_{N,1}^*. \quad (5.6)$$

Below we describe a procedure of constructing the random vectors $\{U_{n,k}\}$ with $\mathcal{L}(\{U_{n,k}\}) = \mathcal{L}(\{U_{n,k}^*\})$, provided that the vectors $Y_1, \dots, Y_{2^N}$ are already constructed. To this end we shall use the so-called Rosenblatt quantile transformation (see [16] and [9]).

Denote by $F^{(1)}(x_1) = \mathbf{P}\{\mathbb{P}_1 \mathcal{U}^* < x_1\}$, $x_1 \in \mathbf{R}^1$, the distribution function of the first coordinate of the vector $\mathcal{U}^*$. Introduce the conditional distributions, denoting by $F^{(j)}(\cdot | x_1, \dots, x_{j-1})$, $2 \leq j \leq d$, r.c.d.f. of $\mathbb{P}_j \mathcal{U}^*$, given $\bar{\mathbb{P}}_{j-1} \mathcal{U}^* = (x_1, \dots, x_{j-1})$. Denote by $\tilde{F}_{n,k}^{(j)}(\cdot | x_1, \dots, x_{j-1})$ the r.c.d.f. of $\mathbb{P}_j U_{n,k}^*$, given $\bar{\mathbb{P}}_{j-1} U_{n,k}^* = (x_1, \dots, x_{j-1})$, $1 \leq k \leq 2^{N-n}$, $1 \leq n \leq N$, $d+1 \leq j \leq 2d$. Put

$$T_0 = 0, \quad T_k = \sum_{l=1}^k Y_l, \quad 1 \leq k \leq 2^N; \quad (5.7)$$

$$\begin{aligned} V_{m,k} &= (V_{m,k}^{(1)}, \dots, V_{m,k}^{(d)}) = T_{k2^m} - T_{(k-1)2^m}, \\ 1 \leq k \leq 2^{N-m}, \quad 0 \leq m \leq N; \\ \tilde{V}_{n,k} &= (V_{n-1,2k-1}, V_{n-1,2k}) = (\tilde{V}_{n,k}^{(1)}, \dots, \tilde{V}_{n,k}^{(2d)}) \in \mathbf{R}^{2d}, \\ 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N; \end{aligned} \quad (5.8)$$

and

$$\begin{aligned} V_{n,k} &= (V_{n,k}^{(1)}, \dots, V_{n,k}^{(2d)}) = A_{n,k} \tilde{V}_{n,k} \in \mathbf{R}^{2d}, \\ 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N. \end{aligned} \quad (5.9)$$

Note that, according to the definition of the operators $A_{n,k}$, we have (see (5.5))

$$\begin{aligned} \bar{\mathbb{P}}_d U_{n,k}^* &= J_{n,k} U_{n,k}^* = \mathcal{U}_{n,k}^*, \\ \bar{\mathbb{P}}_d V_{n,k} &= J_{n,k} V_{n,k} \stackrel{\text{def}}{=} \gamma_{n,k}, \end{aligned} \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N. \quad (5.10)$$

Moreover,

$$V_{n,k} = V_{n-1,2k-1} + V_{n-1,2k}, \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N, \quad (5.11)$$

$$V_{N,1} = Y_1 + \dots + Y_{2N}. \quad (5.12)$$

Set

$$\mathcal{V} = \mathcal{V}_{N,1} = \mathbb{J}_{N,1} V_{N,1} = (\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(d)}) \in \mathbf{R}^d. \quad (5.13)$$

Roughly speaking, the vectors $V_{m,k}$, $\tilde{V}_{n,k}$, $V_{n,k}$, $\mathcal{V}_{n,k}$, $\mathcal{V}$ are constructed from the vectors $Y_1, \dots, Y_{2N}$ by the same linear procedure which was used for constructing the vectors $U_{m,k}^*$, $\tilde{U}_{n,k}^*$, $U_{n,k}^*$, $\mathcal{U}_{n,k}^*$, $\mathcal{U}^*$ from the vectors $\xi_1, \dots, \xi_{2N}$. The absolute continuity of the distributions of $Y_1, \dots, Y_{2N}$ automatically implies the validity of the same property for distributions of the vectors $V_{m,k}$, $\tilde{V}_{n,k}$, $V_{n,k}$, $\mathcal{V}_{n,k}$, $\mathcal{V}$.

For the r.c.d.f. of these vectors we shall use the analogous notation. Denote by $G^{(1)}(x_1) = \mathbf{P}\{\mathbb{P}_1 \mathcal{V} < x_1\}$, $x_1 \in \mathbf{R}^1$, the distribution function of the first coordinate of the vector $\mathcal{V}$. Let $G^{(j)}(\cdot | x_1, \dots, x_{j-1})$, $2 \leq j \leq d$, be the r.c.d.f. of $\mathbb{P}_j \mathcal{V}$, given $\bar{\mathbb{P}}_{j-1} \mathcal{V} = (x_1, \dots, x_{j-1})$. Denote by $\tilde{G}_{n,k}^{(j)}(\cdot | x_1, \dots, x_{j-1})$ the r.c.d.f. of $\mathbb{P}_j V_{n,k}$, given $\bar{\mathbb{P}}_{j-1} V_{n,k} = (x_1, \dots, x_{j-1})$, $d+1 \leq j \leq 2d$, $1 \leq k \leq 2^{N-n}$, $1 \leq n \leq N$.

Define now the new collection of random vectors X_k as follows. At first we define

$$\mathcal{G}^{(1)} = G^{(1)}(\mathcal{V}^{(1)}), \quad \mathcal{U}^{(1)} = (F^{(1)})^{-1}(\mathcal{G}^{(1)}), \quad (5.14)$$

and, for $2 \leq j \leq d$,

$$\begin{aligned} \mathcal{G}^{(j)} &= G^{(j)}(\mathcal{V}^{(j)} | \mathcal{V}^{(1)}, \dots, \mathcal{V}^{(j-1)}), \\ \mathcal{U}^{(j)} &= (F^{(j)})^{-1}(\mathcal{G}^{(j)} | \mathcal{U}^{(1)}, \dots, \mathcal{U}^{(j-1)}) \end{aligned} \quad (5.15)$$

(here $(F^{(1)})^{-1}(x) = \sup\{y: F^{(1)}(y) \leq x\}$, $0 < x < 1$, and so on). If the distributions of the random vectors $\xi_1, \dots, \xi_{2N}$ are absolutely continuous too, then the formulas (5.14) and (5.15) can be rewritten in a more natural «symmetric» form (see [18, p. 30–31]):

$$\begin{aligned} F^{(1)}(\mathcal{U}^{(1)}) &= G^{(1)}(\mathcal{V}^{(1)}) \quad \text{and, for } 2 \leq j \leq d, \\ F^{(j)}(\mathcal{U}^{(j)} | \mathcal{U}^{(1)}, \dots, \mathcal{U}^{(j-1)}) &= G^{(j)}(\mathcal{V}^{(j)} | \mathcal{V}^{(1)}, \dots, \mathcal{V}^{(j-1)}). \end{aligned} \quad (5.16)$$

Define

$$\mathcal{U} = \mathcal{U}_{N,1} = (\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(d)}) \in \mathbf{R}^d, \quad U_{N,1} = \mathbb{J}_{N,1}^{-1} \mathcal{U}. \quad (5.17)$$

Introduce also the random vectors

$$\mathcal{U}^j = (\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(j)}), \quad \mathcal{V}^j = (\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(j)}), \quad j = 1, \dots, d, \quad (5.18)$$

consisting of the first j coordinates of the vectors $\mathcal{U}$, $\mathcal{V}$ respectively.

Suppose that the random vectors $U_{n,k} = (U_{n,k}^{(1)}, \dots, U_{n,k}^{(d)})$, $1 \leq k \leq 2^{N-n}$, corresponding to blocks containing each 2^n summands, are already constructed, for some n , $1 \leq n \leq N$. Now our aim is to construct the blocks containing each 2^{n-1} summands. To this end we define

$$\begin{aligned} \mathcal{U}_{n,k} &= \mathbb{J}_{n,k} U_{n,k} = (\mathcal{U}_{n,k}^{(1)}, \dots, \mathcal{U}_{n,k}^{(d)}); \\ \mathbf{U}_{n,k}^{(j)} &= \mathcal{U}_{n,k}^{(j)} = \mathbb{P}_j \mathbb{J}_{n,k} U_{n,k}, \quad 1 \leq j \leq d, \end{aligned} \quad (5.19)$$

and, for $d+1 \leq j \leq 2d$,

$$\begin{aligned} \mathcal{G}_{n,k}^{(j)} &= \tilde{G}_{n,k}^{(j)}(\mathbf{V}_{n,k}^{(j)} \mid \mathbf{V}_{n,k}^{(1)}, \dots, \mathbf{V}_{n,k}^{(j-1)}), \\ \mathbf{U}_{n,k}^{(j)} &= (\tilde{F}_{n,k}^{(j)})^{-1}(\mathcal{G}_{n,k}^{(j)} \mid \mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(j-1)}). \end{aligned} \quad (5.20)$$

It is clear that there exists the symmetric form of (5.20), similar to (5.16). Then we put

$$\begin{aligned} \mathbf{U}_{n,k} &= (\mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(2d)}) \in \mathbf{R}^{2d}, \\ \mathbf{U}_{n,k}^j &= (\mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(j)}) = \bar{\mathbb{P}}_j \mathbf{U}_{n,k} \in \mathbf{R}^j, \\ \mathbf{V}_{n,k}^j &= (\mathbf{V}_{n,k}^{(1)}, \dots, \mathbf{V}_{n,k}^{(j)}) = \bar{\mathbb{P}}_j \mathbf{V}_{n,k} \in \mathbf{R}^j, \\ \tilde{\mathbf{U}}_{n,k} &= (\tilde{\mathbf{U}}_{n,k}^{(1)}, \dots, \tilde{\mathbf{U}}_{n,k}^{(2d)}) = \mathbf{A}_{n,k}^{-1} \mathbf{U}_{n,k} \in \mathbf{R}^{2d}, \end{aligned} \quad j = 1, \dots, 2d, \quad (5.21)$$

and

$$\begin{aligned} U_{n-1,2k-1} &= \bar{\mathbb{P}}_d \tilde{\mathbf{U}}_{n,k} \in \mathbf{R}^d, \\ U_{n-1,2k} &= (\mathbb{P}_{d+1} \tilde{\mathbf{U}}_{n,k}, \dots, \mathbb{P}_{2d} \tilde{\mathbf{U}}_{n,k}) \in \mathbf{R}^d. \end{aligned} \quad (5.22)$$

Thus, we have constructed the random vectors $U_{n-1,k}$, $1 \leq k \leq 2^{N-n+1}$. After N steps we obtain the random vectors $U_{0,k}$, $1 \leq k \leq 2^N$. Now we set

$$X_k = U_{0,k}, \quad S_0 = 0, \quad S_k = \sum_{l=1}^k X_l, \quad 1 \leq k \leq 2^N. \quad (5.23)$$

Finally, according to (5.2)–(5.5) and (5.10), we have, for $1 \leq k \leq 2^{N-n}$, $1 \leq n \leq N$,

$$\begin{aligned} \tilde{\mathbf{U}}_{n,k} &= (U_{n-1,2k-1}, U_{n-1,2k}) \in \mathbf{R}^{2d}, \\ U_{n,k} &= U_{n-1,2k-1} + U_{n-1,2k} = S_k 2^n - S_{(k-1)2^n} \\ &= X_{(k-1)2^n+1} + \dots + X_{k2^n}, \\ \mathbf{U}_{n,k} &= \mathbf{A}_{n,k} \tilde{\mathbf{U}}_{n,k}, \quad \bar{\mathbb{P}}_d \mathbf{U}_{n,k} = \mathbb{J}_{n,k} U_{n,k} = \mathcal{U}_{n,k}, \end{aligned} \quad (5.24)$$

and

$$U_{N,1} = X_1 + \dots + X_{2^N} = S_{2^N} \quad (5.25)$$

(it is clear, that (5.24) follows from (5.19), (5.21)–(5.23) and from the definition of the operators $\mathbf{A}_{n,k}$).

Lemma 5.1. *The mutual distribution of the constructed vectors $U_{n,k}$ coincides with that of the vectors $U_{n,k}^*$. In particular, $X_k, k = 1, \dots, 2^N$, are independent with $\mathcal{L}(X_k) = \mathcal{L}(\xi_k) = R^{(k)}$.*

Proof. Let us first prove that $\mathcal{L}(U_{N,1}) = \mathcal{L}(U_{N,1}^*)$. According to (5.6) and (5.17), for this it suffices to show that $\mathcal{L}(\mathcal{U}) = \mathcal{L}(\mathcal{U}^*)$. We shall verify by induction that

$$\mathcal{L}(\bar{\mathbb{P}}_j \mathcal{U}) = \mathcal{L}(\bar{\mathbb{P}}_j \mathcal{U}^*), \quad j = 1, \dots, d. \quad (5.26)$$

For $j = 1$ this statement follows immediately from Lemmas 4.2 and 4.4. Suppose that the relation (5.26) is valid for $j = 1, \dots, s-1$, where $2 \leq s \leq d$. It suffices to prove (5.26) for $j = s$. By Lemma 4.1, the random variable $\mathcal{G}^{(s)}$ defined in (5.15) is independent of $(\gamma^{(1)}, \dots, \gamma^{(s-1)})$ and has uniform distribution on the interval $[0, 1]$. Furthermore, $\mathcal{G}^{(s)}$ is independent of $(\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(s-1)})$ too, since $\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(s-1)}$ are measurable functions of $(\gamma^{(1)}, \dots, \gamma^{(s-1)})$ (see (5.15)). To prove (5.26) for $j = s$, it remains to apply Lemma 4.3 to the second of equalities (5.15).

We shall prove by induction in n , that, for fixed n , $0 \leq n \leq N$, the random vectors $U_{n,k}, k = 1, \dots, 2^{N-n}$, are jointly independent and $\mathcal{L}(U_{n,k}) = \mathcal{L}(U_{n,k}^*)$. This implies the statement of the lemma, since, according to (5.4) and (5.24), all the vectors $U_{n,l}$ can be obtained from $\{U_{0,k}\}$ in the same way as $U_{n,l}^*$ can be obtained from $\{U_{0,k}^*\}$.

The case $n = N$ was already considered above. Suppose that the induction conjecture is satisfied for $n = m$, $1 \leq m \leq N$. It suffices to verify the validity of the same statement for $n = m-1$. To this end we shall prove that the random vectors $\mathbf{U}_{m,k}, k = 1, \dots, 2^{N-m}$, are jointly independent and $\mathcal{L}(\mathbf{U}_{m,k}) = \mathcal{L}(\mathbf{U}_{m,k}^*)$. This is sufficient since the operators $\mathbf{A}_{m,k}$ are nondegenerate (see (5.3), (5.5), (5.21), (5.22)).

Note that the random vectors $\mathbf{V}_{m,k}, k = 1, \dots, 2^{N-m}$, are jointly independent (see (5.7)–(5.9)). Introduce the vectors

$$\begin{aligned} \mathcal{H}_{m,k} &= (\mathbb{P}_1 \mathbb{J}_{m,k} V_{m,k}, \dots, \mathbb{P}_d \mathbb{J}_{m,k} V_{m,k}, \mathcal{G}_{m,k}^{(d+1)}, \dots, \mathcal{G}_{m,k}^{(2d)}) \\ &= (\mathbf{V}_{m,k}^{(1)}, \dots, \mathbf{V}_{m,k}^{(d)}, \mathcal{G}_{m,k}^{(d+1)}, \dots, \mathcal{G}_{m,k}^{(2d)}) \in \mathbf{R}^{2d} \end{aligned} \quad (5.27)$$

(see (5.10)). Each of the vectors $\mathcal{H}_{m,k}$ is a function of the corresponding $\mathbf{V}_{m,k}$ (see (5.20)). Therefore, the vectors $\mathcal{H}_{m,k}, k = 1, \dots, 2^{N-m}$, are jointly independent as well.

Now we shall show that, for fixed $k, k = 1, \dots, 2^{N-m}$, the random vector $\bar{\mathbb{P}}_d \mathbf{V}_{m,k} = \mathbb{J}_{m,k} V_{m,k}$ and the random variables $\mathcal{G}_{m,k}^{(d+1)}, \dots, \mathcal{G}_{m,k}^{(2d)}$ are independent in the aggregate. Moreover, we shall prove that $\mathcal{G}_{m,k}^{(j)}, j = d+1, \dots, 2d$, are uniformly distributed on the interval $[0, 1]$.

Indeed, by Lemma 4.1 and by (5.20), for $j = d+1, \dots, 2d$ the random variable $\mathcal{G}_{m,k}^{(j)}$ has uniform distribution on the interval $[0, 1]$ and is indepen-

dent of $\bar{\mathbb{P}}_{j-1} \mathbf{V}_{m,k}$. Furthermore, it is independent of $\bar{\mathbb{P}}_{j-1} \mathcal{H}_{m,k}$ because $\bar{\mathbb{P}}_{j-1} \mathcal{H}_{m,k}$ is a function of $\bar{\mathbb{P}}_{j-1} \mathbf{V}_{m,k}$ (see (5.20) and (5.27)). The needed independence property can be easily derived now by induction.

Thus, we have just proved that the jointly independent random variables $\mathcal{G}_{m,k}^{(j)}$, $k = 1, \dots, 2^{N-m}$, $j = d+1, \dots, 2d$, are uniformly distributed on the interval $[0, 1]$ and independent of the random vectors $\{\bar{\mathbb{P}}_d \mathbf{V}_{m,l} = \mathbb{J}_{m,l} V_{m,l}, l = 1, \dots, 2^{N-m}\}$. Moreover, they are independent of independent (by the induction conjecture) random vectors $U_{m,r}$, $r = 1, \dots, 2^{N-m}$, because all the vectors $U_{m,r}$ are functions of $\{V_{m,l} = \mathbb{J}_{m,l}^{-1} \bar{\mathbb{P}}_d \mathbf{V}_{m,l}, l = 1, \dots, 2^{N-m}\}$ (see (5.8), (5.9), (5.11)–(5.15), (5.17) and (5.19)–(5.22)).

Introduce, for $k = 1, \dots, 2^{N-m}$, the vectors

$$\begin{aligned} \mathcal{J}_{m,k} &= (\mathbb{P}_1 \mathbb{J}_{m,k} U_{m,k}, \dots, \mathbb{P}_d \mathbb{J}_{m,k} U_{m,k}, \mathcal{G}_{m,k}^{(d+1)}, \dots, \mathcal{G}_{m,k}^{(2d)}) \\ &= (\mathbf{U}_{m,k}^{(1)}, \dots, \mathbf{U}_{m,k}^{(d)}, \mathcal{G}_{m,k}^{(d+1)}, \dots, \mathcal{G}_{m,k}^{(2d)}) \in \mathbf{R}^{2d} \end{aligned} \quad (5.28)$$

(see (5.19)). The vectors $\mathcal{J}_{m,k}$, $k = 1, \dots, 2^{N-m}$, are independent in the aggregate. This follows from the mutual independence of

$$\{U_{m,r}, r = 1, \dots, 2^{N-m}; \mathcal{G}_{m,k}^{(j)}, k = 1, \dots, 2^{N-m}, j = d+1, \dots, 2d\}$$

established above. The random vectors $\mathbf{U}_{m,k}$, $k = 1, \dots, 2^{N-m}$, are jointly independent too, because each of $\mathbf{U}_{m,k}$ is a function of the corresponding $\mathcal{J}_{m,k}$ (see (5.20) and (5.28)).

According to (5.21) and (5.22), it remains to show that $\mathcal{L}(\mathbf{U}_{m,k}) = \mathcal{L}(\mathbf{U}_{m,k}^*)$, $k = 1, \dots, 2^{N-m}$. We shall verify by induction that

$$\mathcal{L}(\bar{\mathbb{P}}_j \mathbf{U}_{m,k}) = \mathcal{L}(\bar{\mathbb{P}}_j \mathbf{U}_{m,k}^*), \quad j = d, \dots, 2d. \quad (5.29)$$

For $j = d$ this statement easily follows from the induction conjecture with respect to m (see (5.19) and (5.21)). Suppose that the relation (5.29) is valid for $j = s-1$, where $d+1 \leq s \leq 2d$. It suffices to prove (5.29) for $j = s$. The random variable $\mathcal{G}_{m,k}^{(s)}$ is independent of $\bar{\mathbb{P}}_{s-1} \mathcal{J}_{m,k}$ and has uniform distribution on the interval $[0, 1]$. Furthermore, $\mathcal{G}_{m,k}^{(s)}$ is independent of $\bar{\mathbb{P}}_{s-1} \mathbf{U}_{m,k}$ too, since $\bar{\mathbb{P}}_{s-1} \mathbf{U}_{m,k}$ is a function of $\bar{\mathbb{P}}_{s-1} \mathcal{J}_{m,k}$ (see (5.20), (5.28)). To prove (5.29) for $j = s$, it remains to apply Lemma 4.3 to the second of equalities (5.20). Lemma 5.1 is proved.

If the vectors $X_1, \dots, X_{2N}$, $U_{n,k}$ are constructed as functions of the vectors $Y_1, \dots, Y_{2N}$ as is described above, we shall say that $X_1, \dots, X_{2N}$, $U_{n,k}$ are *quantile transformations of the vectors $Y_1, \dots, Y_{2N}$ associated with the operators $\mathbb{A}_{n,k}$* .

Remark 5.1. It is easy to see that, for some fixed m such that $0 \leq m \leq N$, the vectors $U_{n,k}$, $m \leq n \leq N$, $1 \leq k \leq 2^{N-n}$, are the quantile

transformations of the vectors $V_{m,1}, \dots, V_{m,2^{N-m}}$ associated with the operators $A_{n,k}$, $m \leq n \leq N$, $1 \leq k \leq 2^{N-n}$. This follows from the sequential procedure of constructing the sums over blocks of 2^n summands $U_{n,k}$.

Lemma 5.2. *Let $\alpha > 0$ and let the vectors $X_1, \dots, X_{2^N}$, $U_{n,k}$ be the quantile transformations of the vectors $Y_1, \dots, Y_{2^N}$ associated with the operators $A_{n,k}$. Then the vectors $\alpha X_1, \dots, \alpha X_{2^N}$, $\alpha U_{n,k}$ are the quantile transformations of the vectors $\alpha Y_1, \dots, \alpha Y_{2^N}$ associated as well with the operators $A_{n,k}$.*

We omit the elementary proof of this lemma.

For the formulation of Lemma 5.3 below we shall use the notation

$$U_{n,k}^- = U_{n-1,2k-1} - U_{n-1,2k}, \quad 1 \leq k \leq 2^{N-n}, \quad 1 \leq n \leq N. \quad (5.30)$$

U. Einmahl has informed the author that the shortest proof of Lemma 5.3 can be obtained with the help of a geometrical approach due to Massart [13, p. 275].

Lemma 5.3 (see [9, Lemma 5, p. 55]). *Let $1 \leq m = (2s+1)2^r \leq 2^N$, where s, r are non-negative integers. Then the following representation is valid:*

$$S_m = \frac{m}{2^N} S_{2^N} + \sum_{n=r+1}^N \gamma_n U_{n,l_{n,m}}^-, \quad (5.31)$$

where $\gamma_n = \gamma_n(m) \in [0, \frac{1}{2}]$ and the integers $l_{n,m}$, $n = 0, \dots, N$, are defined by the relations

$$(l_{n,m} - 1)2^n < m \leq l_{n,m}2^n. \quad (5.32)$$

Remark 5.2. The inequalities (5.32) give a formal definition of $l_{n,m}$. To understand better the mechanism of the dyadic scheme, one should remember another characterization of these numbers: $U_{n,l_{n,m}}$ is the sum over the block of 2^n summands which contains X_m , the last summand in the considered sum S_m .

Lemma 5.4. *Let $1 \leq m = (2s+1)2^r \leq 2^N$, where s, r are non-negative integers. Then the following representation is valid:*

$$S_m = \sum_{n=r}^N \beta_n U_{n,l_{n,m}}, \quad (5.33)$$

where real-valued coefficients $\beta_n = \beta_n(m)$ satisfy $|\beta_n| \leq \frac{3}{2}$ and the integers $l_{n,m}$ are defined by (5.32).

Proof. Using (5.24), (5.30) and Remark 5.2, we see that

$$U_{n-1,l_{n-1,m}} = \frac{1}{2} U_{n,l_{n,m}} - \frac{b_n}{2} U_{n,l_{n,m}}^-, \quad n = 1, \dots, N,$$

where the factors $b_n = b_n(m)$ can be equal either to 1 or to -1 . Hence,

$$U_{n,l_{n,m}}^- = b_n(U_{n,l_{n,m}} - 2U_{n-1,l_{n-1,m}}), \quad n = 1, \dots, N. \quad (5.34)$$

Note that $l_{N,m} = 1$ (see (5.32) or Remark 5.2). Substituting (5.34) in (5.31) and using (5.25), we get

$$S_m = \frac{m}{2^N} U_{N,1} + \sum_{n=r+1}^N \gamma_n b_n (U_{n,l_{n,m}} - 2U_{n-1,l_{n-1,m}}) = \sum_{n=r}^N \beta_n U_{n,l_{n,m}}$$

with

$$\beta_n = \begin{cases} 1, & \text{if } n = r = N, \\ 2^{-N} m + \gamma_N b_N, & \text{if } n = N, r \neq N, \\ \gamma_n b_n - 2\gamma_{n+1} b_{n+1}, & \text{if } r+1 \leq n \leq N-1, \\ -2\gamma_{r+1} b_{r+1}, & \text{if } n = r \neq N. \end{cases}$$

In order to verify that $|\beta_n| \leq \frac{3}{2}$, it remains to recall that $|b_j| = 1$, $0 \leq \gamma_j \leq \frac{1}{2}$, and $0 < m \leq 2^N$. Lemma 5.4 is proved.

REFERENCES

1. Berger E. Fast sichere Approximation von Partialsummen unabhängiger und Stationärer ergodischer Folgen von Zufallsvektoren. Dissertation. Universität Göttingen, 1982.
2. Berkes I., Philipp W. Approximation theorems for independent and weakly dependent random vectors. — Ann. Probab., 1979, v. 7, p. 29–54.
3. Боровков А. А. О скорости сходимости в принципе инвариантности. — Теория вероятн. и ее примен., 1973, т. 18, в. 2, с. 217–234.
4. Csörgő M., Révész P. A new method to prove Strassen type laws of invariance principle. I; II. — Z. Wahrscheinlichkeitstheor. verw. Geb., 1975, v. 31, p. 255–259; p. 261–269.
5. Csörgő M., Révész P. Strong Approximations in Probability and Statistics. New York: Academic Press, 1981.
6. Csörgő S., Hall P. The Komlós–Major–Tusnády approximations and their applications. — Austral. J. Statist., 1984, v. 26, № 2, p. 189–218.
7. Einmahl U. A useful estimate in the multidimensional invariance principle. — Probab. Theory Relat. Fields, 1987, v. 76, № 1, p. 81–101.
8. Einmahl U. Strong invariance principles for partial sums of independent random vectors. — Ann. Probab., 1987, v. 15, p. 1419–1440.
9. Einmahl U. Extensions of results of Komlós, Major and Tusnády to the multivariate case. — J. Multivariate Anal., 1989, v. 28, p. 20–68.
10. Гизман И. И., Скороход А. В. Теория случайных процессов. Т. I. М.: Наука, 1971, 664 с.
11. Götze F., Zaitsev A. Yu. Multidimensional Hungarian construction for vectors with almost Gaussian smooth distributions. Preprint 97-071 SFB 343. Bielefeld: Universität Bielefeld, 1997, 29 p.
12. Komlós J., Major P., Tusnády G. An approximation of partial sums of independent RV's and the sample DF. I; II. — Z. Wahrscheinlichkeitstheor. verw. Geb., 1975, v. 32, p. 111–131; 1976, v. 34, p. 34–58.
13. Massart P. Strong approximation for multivariate empirical and related processes, via KMT construction. — Ann. Probab., 1989, v. 17, № 1, p. 266–291.

14. *Philipp W.* Almost sure invariance principles for sums of B -valued random variables. — *Lecture Notes in Math.*, 1979, v. 709, p. 171–193.
15. *Прозоров Ю. В.* Сходимость случайных процессов и предельные теоремы теории вероятностей. — *Теория вероятн. и ее примен.*, 1956, т. 1, в. 2, с. 177–238.
16. *Rosenblatt M.* Remarks on a multivariate transformation. — *Ann. Math. Statist.*, 1952, v. 23, p. 470–472.
17. *Рудский Р.* Вероятности больших отклонений случайных векторов. — *Литов. матем. сб.*, 1983, т. 23, № 1, с. 195–204.
18. *Саханенко А. И.* Скорость сходимости в принципе инвариантности для разно-распределенных величин с экспоненциальными моментами. — *Труды СО АН СССР*, 1984, т. 3, с. 4–49.
19. *Саулис Л.* Большие отклонения для случайных векторов для некоторых классов множеств. — *Литов. матем. сб.*, 1983, т. 23, № 3, с. 142–154.
20. *Sazonov V. V.* Normal Approximation — Some Recent Advances. Berlin–Heidelberg: Springer-Verlag, 1981, 105 p.
21. *Shao Qi-Man.* Strong approximation theorems for independent random variables and their applications. — *J. Multivariate Anal.*, 1995, v. 52, № 1, p. 107–130.
22. *Скорозод А. В.* Исследования по теории случайных процессов. Киев: Киевский гос. ун-т, 1961, 216 с.
23. *Statulevičius V. A.* On large deviations. — *Z. Wahrscheinlichkeitstheor. verw. Geb.*, 1966, v. 6, № 2, p. 133–144.
24. *Yurinskii V. V.* On the error of the Gaussian approximation to the probability of a ball. Unpublished manuscript.
25. *Зайцев А. Ю.* О гауссовской аппроксимации сверток при выполнении многомерных аналогов условий типа неравенства Бернштейна. Препринт Р-9-84. Ленинград: ЛОМИ, 1984, 54 с.
26. *Зайцев А. Ю.* Оценки расстояния Леви–Прохорова в многомерной центральной предельной теореме для случайных векторов с конечными экспоненциальными моментами. — *Теория вероятн. и ее примен.*, 1986, т. 31, в. 2, с. 246–265.
27. *Zaitsev A. Yu.* On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein inequality conditions. — *Probab. Theory Relat. Fields*, 1987, v. 74, № 4, p. 535–566.
28. *Зайцев А. Ю.* О связи между двумя классами вероятностных распределений. — *Кольца и модули. Предельные теоремы теории вероятностей. Т. 2.* Ленинград: Ленинградский ун-т, 1988, с. 153–158.
29. *Zaitsev A. Yu.* Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. Preprint 95-055 SFB 343. Bielefeld: Universität Bielefeld, 1995, 112 p.
30. *Зайцев А. Ю.* Оценки квантилей гладких условных распределений и многомерный принцип инвариантности. — *Сиб. матем. журн.*, 1996, т. 37, № 4, с. 807–831.
31. *Zaitsev A. Yu.* Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. — *ESAIM, Probab. Statist.*, 1998, v. 2, p. 41–108.

Поступила в редакцию
7.IX.1998