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

A. Yu. Zaitsev

St. Petersburg Branch of Steklov Mathematical Institute
and University of Bielefeld

May 1998

Abstract. A multidimensional version of a result of Sakhanenko for the Gaussian approximation of sequences of successive sums of independent non-identically distributed random vectors with finite exponential moments is obtained.

1. Introduction

The aim of this paper is to give a multidimensional generalization of a result of Sakhanenko (1984) which is an extension and sharpening of the famous result of Komlós, Major and Tusnády (KMT) (1975–76) to the case of non-identically distributed random variables. For formulations of results we need the following notation.

Notation 1.1. We write $z \in \mathbb{R}^d$ (or $\mathbb{C}^d$), if $z = (z_1, \dots, z_d) = z_1 e_1 + \dots + z_d e_d$, where $z_j \in \mathbb{R}^1$ (or $\mathbb{C}^1$) and the e_j are the basis vectors. The scalar product of vectors $x, y \in \mathbb{R}^d$ (or $\mathbb{C}^d$) is denoted by $\langle x, y \rangle = x_1 \overline{y}_1 + \dots + x_d \overline{y}_d$. For $z \in \mathbb{R}^d$ (or $\mathbb{C}^d$) we use the Euclidean norm $\|z\| = \langle z, z \rangle^{1/2}$ and the maximum norm $|z| = \max\limits_{1 \leqslant j \leqslant k} |z_j|$. For $b > 0$ we write $\log^* b = \max \{1, \log b\}$. The distribution and the corresponding

1991 *Mathematics Subject Classification.* Primary 62F17; secondary 60F15.

Key words and phrases. Multidimensional invariance principle, strong approximation, sums of independent non-i.i.d. random vectors.

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.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

covariance operator of a random vector ξ is denoted by $\mathcal{L}(\xi)$ and $\operatorname{cov}\xi$ (or $\operatorname{cov} F$, if $F=\mathcal{L}(\xi)$). For random vectors $X_1,\dots,X_s$ and $Y_1,\dots,Y_s$ we write

$$\Delta(X,Y)=\max_{1\leqslant j\leqslant s}\Big\|\sum_{i=1}^{j}X_i-\sum_{i=1}^{j}Y_i\Big\|. \tag{1.1}$$

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

The rate for strong approximation in the one-dimensional invariance principle was studied by many authors (see, e.g., Prokhorov (1956), Skorokhod (1961), Borovkov (1973), Csörgő and Révész (1975) and the bibliography in Csörgő and Révész (1981), Csörgő and Hall (1984), Shao (1995)). Skorokhod (1961) 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 (KMT) (1975–76) 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 (1984) generalized and essentially sharpened KMT results in the case of non-identically distributed random variables. He considered the following class of one-dimensional distributions:

$$\mathcal{S}_1(\tau)=\Big\{\mathcal{L}(\xi):\mathbf{E}\,\xi=0,\ \mathbf{E}\,|\xi|^3\exp\big(\tau^{-1}|\xi|\big)\leqslant\tau\,\mathbf{E}\,|\xi|^2\Big\}. \tag{1.2}$$

His main result is formulated as follows.

Theorem 1.2 (Sakhanenko, 1984). *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 vectors* $X_1,\dots,X_s$ *and a sequence of independent Gaussian random vectors* $Y_1,\dots,Y_s$ *so that* $\mathcal{L}(X_j)=\mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j=0$, $\mathbf{E}\,Y_j^2=\mathbf{E}\,X_j^2$, $j=1,\dots,s$, *and*

$$\mathbf{E}\,\exp\big(c_1\,\Delta(X,Y)/\tau\big)\leqslant 1+B/\tau, \tag{1.3}$$

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 Berkes and Philipp (1979), Philipp (1979), Berger (1982), Einmahl (1987a,b; 1989). Zaitsev (1995, 1996, 1998) removed an unnecessary logarithmic factor from the result of Einmahl (1989) and obtained multidimensional analogs of KMT results. In Zaitsev (1998, Theorem 1.3) the random vectors are, generally speaking, non-identically 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 (1984) remained open. This generalization is given in Theorem 1.4 below.

Definition 1.3 (Zaitsev (1986))**.** Denote by $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$, the class of probability distributions which consists of d-dimensional distributions F for which the function

$$\varphi(z) = \varphi(F, z) = \log \int_{\mathbb{R}^d} e^{\langle z,x\rangle} F\{dx\} \qquad (\varphi(0) = 0) \tag{1.4}$$

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

$$\left|d_u d_v^2\, \varphi(z)\right| \leqslant \|u\|\,\tau \left\langle \mathbb{D}v, v\right\rangle \tag{1.5}$$

for all u, $v \in \mathbb{R}^d$ and $\|z\|\,\tau < 1$, where $\mathbb{D} = \operatorname{cov} F$, 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 1.4 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).

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 $u \in \mathbb{R}^d$ and for all $p = 1, \dots, l$

$$C_1^2\,\|u\|^2 \leqslant \left\langle \mathbb{B}_p u, u\right\rangle \leqslant C_2^2\,\|u\|^2 \tag{1.6}$$

with some constants C_1 and C_2.

The following theorem is the main result of the paper.

Theorem 1.4. *Let $\alpha > 0$, $\tau \geqslant 1$, $D > 0$ and $\xi_1, \dots, \xi_s$ 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*

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

and suppose that the covariance operators $\operatorname{cov}\zeta_p = \mathbb{B}_p$ 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 $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, \dots, s$, and

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

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 then this assertion is valid without the factor $\log^* d$ in the denominator of the fraction in the left-hand side of (1.8).

The main result of Zaitsev (1998, Theorem 1.3) can be obtained if one takes in Theorem 1.4 and in Condition $\mathbf{B}_2$ $l = s$, $m_p = p$, $\zeta_p = \xi_p$, and $D = C_1 = C_2 = 1$, (hence, $\mathbb{B}_p = \mathbb{I}_d$, $p = 1, \ldots, s$). Note that the dependence of constants on the dimension d in Zaitsev (1998, Theorem 1.3) is somewhat better due to the fact that the proof of Theorem 1.4 is more complicated.

In Theorem 1.4 we give an explicite form of the dependence of constants on the distributions of summands and on the dimension d. It is clear that applying Theorem 1.4 one should minimize the parameter τ in (1.8). Moreover, one can try to optimize the choice ζ_p and D in (1.7). For example, let the conditions of Theorem 1.4 be satisfied with $D = 1$, $\tau = 2^{k/2}$ and $l = 2^r$ where k, r, $k \leqslant r$, are positive integers. Then one can enlarge the blocks of summands putting

$$\zeta_i^* = D^* \left(\zeta_{(i-1)\cdot 2^k+1} + \zeta_{(i-1)\cdot 2^k+2} + \cdots + \zeta_{i\cdot 2^k}\right), \qquad i = 1, \ldots, 2^{r-k},$$

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 1.4 are satisfied with the change of ζ_p, D, l, τ by ζ_i^*, D^*, l^*, τ^* respectively. An application of Theorem 1.4 gives the same inequality (1.8) as before. On the other hand, even if the conditions of Theorem 1.4 are satisfied for ζ_i^*, D^*, l^*, τ^*, they can be not valid for ζ_p, D, l, τ.

The conditions of Theorem 1.4 are expressed in terms of distributions of sums over blocks of summands. A simple sufficient condition for the validity of the relations $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$, $p = 1, \ldots, l$, is given by the relations $\mathcal{L}(\xi_j) \in \mathcal{A}_d(\tau/D)$, $j = 1, \ldots, s$ (see Section 2). On the other hand, the relations $\mathcal{L}(\zeta_p) \in \mathcal{A}_d(\tau)$, $p = 1, \ldots, l$, do not imply that $\mathcal{L}(\xi_j) \in \mathcal{A}_d(\tau/D)$, $j = 1, \ldots, s$.

The following Theorem 1.5 will be used in the proof of Theorem 1.4.

Theorem 1.5. *Suppose that $\tau \geqslant 1$, $\alpha > 0$ and $\widetilde{\xi}_1, \ldots, \widetilde{\xi}_n$ are random vectors with distributions $\mathcal{L}(\widetilde{\xi}_k) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\,\widetilde{\xi}_k = 0$. Let covariance operators $\operatorname{cov}\widetilde{\xi}_k = \mathbb{B}_k$, $k = 1, \ldots, n$, satisfy Condition $\mathbf{B}_2$. Assume that there exist random vectors $\widetilde{\eta}_k$ which have the same moments of the first three orders as the vectors $\widetilde{\xi}_k$, $\mathcal{L}(\widetilde{\eta}_k) \in \mathcal{A}_d(\tau)$ and $\mathbf{P}\left\{\left\|\widetilde{\eta}_k\right\| \leqslant \lambda\right\} = 1$, $k = 1, \ldots, n$, where $\lambda \geqslant \max\{C_2, 1\}\sqrt{d}$. Then one can construct on a probability space a sequence of independent random vectors $\widetilde{X}_1, \ldots, \widetilde{X}_n$ and a sequence of independent Gaussian random vectors $Y_1, \ldots, Y_n$ so that $\mathcal{L}(\widetilde{X}_k) = \mathcal{L}(\widetilde{\xi}_k)$, $\mathbf{E}\,Y_k = 0$, $\operatorname{cov} Y_k = \operatorname{cov}\widetilde{X}_k$, $k = 1, \ldots, n$, and*

$$\mathbf{E}\,\exp\Big(\frac{a_3\,\Delta(\widetilde{X}, Y)}{\lambda\, d^4\, \tau^2}\Big) \leqslant \exp\Big(a_4\,(\lambda\, d^{3/2}\,\tau)^{3/2+\alpha}\log^* n\Big), \tag{1.9}$$

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 1.4 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., Shao (1995) and the bibliography therein). The author is going to devote to the consequences of Theorem 1.4 a separate publication.

The proof of Theorems 1.4 and 1.5 consists of several steps, many of which are close to corresponding steps from the proofs of the main results of KMT (1975–76), Sakhanenko (1984), Einmahl (1989) and Zaitsev (1995, 1998). In Section 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 3.3 and 3.5 which were proved in Zaitsev (1995, 1996, 1998) 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 5–7 we describe a modification of the well-known dyadic construction due to KMT (1975–76), one-dimensional case, and to Einmahl (1989), multidimensional case. The construction in the present paper differs from that of Zaitsev (1995, 1998) since now we have to work with random vectors having not coinciding covariance operators. Therefore we introduce in the construction new linear operators $\mathbb{A}_{n,k}$, $\mathbb{D}_{n,k}$, $\widetilde{\mathbb{D}}_{n,k}$, $\mathbb{J}_{n,k}$, $\widetilde{\mathbb{J}}_{n,k}$ (see (5.5), (6.3) and (6.5)). This is important for applying Lemmas 3.3 and 3.5 because these lemmas can be applied only in the case when the random variable to be constructed is non-correlated with random variables which are already constructed. In the case when all summands have the identity covariance operators $\mathbb{I}_d$ the operators $\mathbb{A}_{n,k}$, $\mathbb{D}_{n,k}$, $\widetilde{\mathbb{D}}_{n,k}$, $\mathbb{J}_{n,k}$, $\widetilde{\mathbb{J}}_{n,k}$ "disappear" since then $\mathbb{A}_{n,k}=\mathbb{A}^{\pm}$ (see Remark 6.5), $\mathbb{D}_{n,k}=\widetilde{\mathbb{D}}_{n,k}=2^n\,\mathbb{I}_d$, $\mathbb{J}_{n,k}=\widetilde{\mathbb{J}}_{n,k}=\mathbb{I}_d$. The sufficiently complicated formulas (6.5) turn into the well-known relation $\widetilde{\mathcal{U}}_{n,k}=\widetilde{U}_{n,k}=U_{n-1,2k-1}-U_{n-1,2k}$ which was traditionally used in the dyadic constructions. We assume that the operators mentioned above satisfy Conditions $\mathbf{B}_1$–$\mathbf{B}_4$ (see Section 6) which we call for the sake of brevity Condition $\mathbf{B}$ (cf. Condition $\mathbf{B}_2$ above). The crucial role in the proof of Theorems 1.4 and 1.5 is played by Lemma 6.7. In the proof of this lemma the most important formula is (6.47) in which we use several times the relations of type $\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n,k}=\mathbb{I}_d$.

In Section 7 we introduce a simple transformation of probability distributions $\mathcal{L}(X)$ which is denoted by $\Psi\big(\mathcal{L}(X)\big)$ (see Notation 7.4). Such a transformation was already used in Zaitsev (1995, 1998). The distributions $\Psi\big(\mathcal{L}(X)\big)$ have the same moments of the first three orders as $\mathcal{L}(X)$. Moreover, the distribution $\Psi\big(\mathcal{L}(X)\big)$ is "better" than $\mathcal{L}(X)$ because it is sufficiently smooth. Then we introduce Conditions $\mathbf{A}(\tau)$, $\mathbf{A}^*(\tau)$, $\mathbf{C}_\Psi(p)$, $\mathbf{D}_\Phi$, $\mathbf{D}_\Psi$. 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 $\mathbf{D}_\Phi$ means that we start the dyadic construction having the Gaussian vectors being constructed. Condition $\mathbf{D}_\Psi$ says that the initial vectors have bounded and smooth distributions $\Psi\big(\mathcal{L}(\widetilde{\eta}_j)\big)$ (see Condition $\mathbf{B}_1$ and Remark 6.2). Condition $\mathbf{C}_\Psi(p)$ 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 $\Psi\big(\mathcal{L}(\widetilde{\eta}_j)\big)$ (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 3.3 and 3.5. Conditions $\mathbf{A}(\tau)$ and $\mathbf{A}^*(\tau)$ yield belonging distributions of summands and sums over blocks to some subclasses of $\mathcal{A}_d(\tau)$.

In Section 8 we prove Theorems 8.1 and 8.4. Theorem 8.1 contained KMT-type estimate under Conditions $\mathbf{B}$, $\mathbf{D}_\Phi$ and $\mathbf{A}^*(\tau)$ 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 Zaitsev (1986) and can be considered as an estimate of stability of this characterization: if $F \in \mathcal{A}_d(\tau)$, $\tau > 0$, then

$$\pi\big(F, \Phi(F)\big) \leqslant c d^2 \tau \log^*(\tau^{-1}), \tag{1.10}$$

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 Zaitsev (1986) it was established that for all d-dimensional Borel sets X and all $\lambda > 0$

$$\begin{aligned} F\{X\} &\leqslant \Phi(F)\{X^\lambda\} + c d^2 \exp\Big(-\frac{\lambda}{c d^2 \tau}\Big), \\ \Phi(F)\{X\} &\leqslant F\{X^\lambda\} + c d^2 \exp\Big(-\frac{\lambda}{c d^2 \tau}\Big), \end{aligned} \tag{1.11}$$

where $X^\lambda = \Big\{ y \in \mathbb{R}^d : \inf\limits_{x \in X} \|x - y\| < \lambda \Big\}$ 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 $\operatorname{cov} F$, in contrast to Theorems 1.4 and 1.5, where $\tau \geqslant 1$ and covariance operators satisfy Condition $\mathbf{B}_2$. The question about the necessity of Condition $\mathbf{B}_2$ in Theorems 1.4 and 1.5 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 (1997) which is proved for the case of identity covariance operators $\mathbb{I}_d$. That is, when $C_1 = C_2 = 1$. In the preprint Götze and Zaitsev (1997) one can found simple examples in which the sufficiently complicated Condition $\mathbf{A}^*(\tau)$ is satisfied.

Theorem 8.4 yields a bound of the rate of strong approximation in the case when $\tau \geqslant 1$ 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 $\Psi\big(\mathcal{L}(\widetilde{\eta}_j)\big)$. Lemma 7.6 allows us to reduce the proof of Theorem 8.4 to an application of Theorem 8.1.

The scheme of the proof of Theorem 1.4 is very close to that of the main results of Sakhanenko (1984) and Zaitsev (1995, 1998). Theorem 1.4 is an immediate consequence of Theorem 1.5. In the proof of Theorem 1.5 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 $\widetilde{X}_1, \dots, \widetilde{X}_n$. The rate of approximation is estimated in Theorem 8.4. Then we construct the vectors $\widetilde{X}_1, \dots, \widetilde{X}_n$ in several steps. After each step the number of $\widetilde{X}_k$ 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 $\mathcal{L}(\widetilde{\xi}_k)$. These $2^p - 1$ vectors from each block will be chosen as $\widetilde{X}_k$ and will be not 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 3.5. Section 9 is devoted to the estimation of closeness of random vectors in the steps of the procedure described above. The proof of Theorem 1.5 is given in Section 10. Theorem 1.4 is derived from Theorem 1.5 in Section 11. At the end of Section 11 we show that Theorem 1.2 is an easy consequence of Theorem 1.4.

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 $\mathbb{R}^d$. 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 $\widehat{F}(t)$, $t \in \mathbb{R}^d$, 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 $\tau \geqslant 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $\|h\|\,\tau < 1$, $h \in \mathbb{R}^d$. Then the conjugate distribution $\overline{F} = \overline{F}(h)$ is defined by

$$\overline{F}\{dx\} = \Big(\mathbf{E}\, e^{\langle h,\xi\rangle}\Big)^{-1} e^{\langle h,x\rangle} F\{dx\} \tag{2.1}$$

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

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

Lemma 2.2 (Zaitsev 1986, Lemma 2.1)**.** *Suppose that* $\tau \geqslant 0$, $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $h \in \mathbb{R}^d$, $\|h\|\,\tau < 1$, $\mathbb{D} = \operatorname{cov} F$, $\mathbb{D}(h) = \operatorname{cov} \overline{F}(h)$. $\mathbf{E}\,\xi = 0$. *Then*

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

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

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

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

b) *If* $\|h\|\,\tau \leqslant 1/2$, *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 Zaitsev (1986, 1996, 1998). In particular, if $0 \leqslant \tau_1 \leqslant \tau_2$, 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 $\mathbf{E}\, e^{\langle h,\xi\rangle}$, $h \in \mathbb{R}^d$, $\|h\|\,\tau < 1$. This leads to exponential estimates for the tails of distributions (see, e.g., Lemma 2.7).

The condition $\mathcal{L}(\xi) \in \mathcal{A}_1(c\tau)$ is equivalent to Statulevičius' (1966) conditions

$$|\gamma_m| \leqslant c m!\, \tau^{m-2}\gamma_2, \qquad m = 3, 4, \dots, \tag{2.6}$$

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 (1983) and Saulis (1983).

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

$$\big|d_v^2\, \varphi(z)\big| \leqslant 2\,\big\langle \mathbb{D}v, v\big\rangle$$

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

An analog of the inequalities (1.10), (1.11) was proved in Zaitsev (1984, 1987) for the convolutions of distributions from the class

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

It can be easily verified that all distributions with zero mean, which are concentrated on the ball $\big\{x \in \mathbb{R}^d : \|x\| \leqslant \tau\big\}$, 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 Zaitsev (1984), the classes $\mathcal{B}_1(\tau)$ and $\mathcal{S}_1(\tau)$ are equivalent (see Lemma 2.4 for the d-dimensional analog of this statement). The following Lemma 2.3 shows the connection between classes $\mathcal{A}_d(\tau)$ and $\mathcal{B}_d(\tau)$.

Lemma 2.3 (Zaitsev 1986, 1988)**.** *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 $B^2 \leqslant 12\,\tau^2$, $\mathbf{E}\,\xi = 0$ and $F \in \mathcal{A}_d(c_2\,\tau)$.*
b) *If $F = \mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $B^2 \leqslant \tau^2$ and $\mathbf{E}\,\xi = 0$, then $F \in \mathcal{B}_d(c_3\,\tau)$.*

It is easy to check that if $\mathcal{L}(\xi) \in \mathcal{B}_d(\tau)$, then, for all $u, v \in \mathbb{R}^d$, $m = 3, 4\dots$,

$$\mathbf{E}\,\big\langle \xi, v\big\rangle^2\,\Big|\big\langle \xi, u\big\rangle\Big|^{m-2} \leqslant \frac{1}{\sqrt{3}}\,m!\,\tau^{m-2}\,\|u\|^{m-2}\,\mathbf{E}\,\big\langle \xi, v\big\rangle^2 \tag{2.8}$$

(it suffices to consider odd m only and to use the Hölder inequality).

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

$$\mathcal{S}_d(\tau) = \Big\{F = \mathcal{L}(\xi) \in \mathfrak{F}_d : \mathbf{E}\,\xi = 0,\ \mathbf{E}\,\Big|\big\langle \xi, v\big\rangle^2 \big\langle \xi, u\big\rangle\,e^{|\langle \xi, u\rangle|}\Big| \leqslant \mathbf{E}\,\big\langle \xi, v\big\rangle^2 \ \text{ for all } u, v \in \mathbb{R}^d \text{ such that } \|u\| \leqslant \tau^{-1}\Big\}. \tag{2.9}$$

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 Lemma 2.4 is contained in Zaitsev (1984). We give here its proof for the sake of completeness.

Lemma 2.4. *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 $\tau \geqslant 0$.*

Proof. Without loss of generality we assume that $\tau > 0$. Let $\mathcal{L}(\xi) \in \mathcal{B}_d(\tau)$. Then, by (2.8), for any $u, v \in \mathbb{R}^d$ such that $\|u\| \leqslant (7\tau)^{-1}$, we have

$$\begin{aligned}
\mathbf{E}\left|\left\langle \xi, v\right\rangle^2 \left\langle \xi, u\right\rangle e^{|\langle \xi, u\rangle|}\right| &\leqslant \sum_{m=0}^{\infty} \frac{1}{m!}\, \mathbf{E}\left\langle \xi, v\right\rangle^2 \left|\left\langle \xi, u\right\rangle\right|^{m+1} \\
&\leqslant \frac{1}{\sqrt{3}}\, \mathbf{E}\left\langle \xi, v\right\rangle^2 \sum_{m=0}^{\infty} (m+1)\,(m+2)\,(m+3)\,\tau^{m+1}\,\|u\|^{m+1} \\
&\leqslant \mathbf{E}\left\langle \xi, v\right\rangle^2,
\end{aligned}$$

Therefore, $\mathcal{L}(\xi) \in \mathcal{S}_d(7\tau)$.

Let now $\mathcal{L}(\xi) \in \mathcal{S}_d(\tau)$. Then for all $m = 3, 4 \ldots$, and for all $u, v \in \mathbb{R}^d$ such that $\|u\| = \tau^{-1}$ we obtain

$$\begin{aligned}
\left|\, \mathbf{E}\left\langle \xi, v\right\rangle^2 \left\langle \xi, u\right\rangle^{m-2} \right| &\leqslant (m-3)!\, \mathbf{E}\left|\left\langle \xi, v\right\rangle^2 \left\langle \xi, u\right\rangle e^{|\langle \xi, u\rangle|}\right| \\
&\leqslant (m-3)!\, \big(\, \|u\|\, \tau\big)^{m-2}\, \mathbf{E}\left\langle \xi, v\right\rangle^2 \\
&\leqslant \frac{1}{2}\, m!\, \big(\, \|u\|\, \tau\big)^{m-2}\, \mathbf{E}\left\langle \xi, v\right\rangle^2.
\end{aligned} \tag{2.10}$$

Obviously, the validity of the inequality

$$\left|\, \mathbf{E}\left\langle \xi, v\right\rangle^2 \left\langle \xi, u\right\rangle^{m-2} \right| \leqslant \frac{1}{2}\, m!\, \tau^{m-2}\, \|u\|^{m-2}\, \mathbf{E}\left\langle \xi, v\right\rangle^2, \quad \text{for all } u, v \in \mathbb{R}^d,$$

follows automatically from the same fact obtained in (2.10) under the restriction $\|u\| = \tau^{-1}$. Therefore, $\mathcal{S}_d(\tau) \subset \mathcal{B}_d(\tau)$. $\square$

As examples of distributions from $\mathcal{A}_d(c\tau)$ one can consider infinitely divisible distributions with spectral measures concentrated on the ball $\left\{x \in \mathbb{R}^d : \|x\| \leqslant \tau\right\}$ (see Zaitsev (1986)). One can find therein a more general condition which ensure belonging an infinitely divisible distribution to $\mathcal{A}_d(c\tau)$ (see Zaitsev (1986, 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 $\operatorname{cov}\zeta$. Suppose that $\lambda > 0$ and an infinitely divisible distribution F is defined by its characteristic function

$$\widehat{F}(t) = \exp\bigg(\lambda \int_{\mathbb{R}^d} \big(e^{i\langle t, x\rangle} - 1\big)\, H\{dx\}\bigg), \qquad t \in \mathbb{R}^d.$$

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

The following elementary Lemmas 2.5–2.7 are proved in Zaitsev (1998).

Lemma 2.5. *Let* $\tau \geqslant 0$, $\mathcal{L}(\xi) \in \mathcal{A}_d(\tau)$, $y \in \mathbb{R}^m$, $\alpha \in \mathbb{R}^1$, *and* $\mathbb{A} : \mathbb{C}^d \to \mathbb{C}^m$ *be a linear operator such that* $\mathbb{A}(\mathbb{R}^d) \subset \mathbb{R}^m$. *Let* $\widetilde{\xi} \in \mathbb{R}^k$ *be the vector consisting of a subset of coordinates of the vector* ξ. *Then*

$$\mathcal{L}(\mathbb{A}\xi + y) \in \mathcal{A}_m\big(\|\mathbb{A}\|\,\tau\big), \qquad \textit{where} \quad \|\mathbb{A}\| = \sup_{x\in\mathbb{R}^d,\ \|x\|\leqslant 1} \|\mathbb{A}x\|\,,$$

$$\mathcal{L}(\alpha\xi) \in \mathcal{A}_d(|\alpha|\,\tau), \qquad \mathcal{L}(\widetilde{\xi}) \in \mathcal{A}_k(\tau).$$

Lemma 2.6. *Suppose that* $\tau \geqslant 0$, *the vectors* $\xi^{(k)}$, $k = 1, 2$, *are independent, and* $\mathcal{L}\big(\xi^{(k)}\big) \in \mathcal{A}_{d_k}(\tau)$. *Let* $\xi \in \mathbb{R}^{d_1+d_2}$ *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 2.7. *Let* $\tau \geqslant 0$ *and* $X_1, \dots, X_n$ *be independent random vectors with* $\mathbf{E}\, X_i = 0$, $S_i = X_1 + \cdots + X_i$, *for* $i = 1, \dots, n$. *Let* $h, t \in \mathbb{R}^1$, $0 \leqslant h\tau < 1$, $0 \leqslant t\tau < {}^1\!/_2$, $\xi = \max\limits_{1\leqslant i\leqslant n} \big|S_i\big|$. *Assume that* $\mathcal{L}(S_n) \in \mathcal{A}_d(\tau)$. *Denote by* B^2 *the maximal eigenvalue of* $\operatorname{cov} S_n$. *Then*

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

and

$$\mathbf{P}\big\{\xi \geqslant x\big\} \leqslant 2\,d\,\max\Big\{\exp\big(-x^2/4\,B^2\big),\ \exp\big(-x/4\,\tau\big)\Big\}, \qquad x \geqslant 0.$$

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 $\tau \geqslant 0$, $\delta > 0$, $\rho > 0$, $h \in \mathbb{R}^d$. Consider the conditions:

$$(2\,\pi)^{-d} \int\limits_{\rho\|t\|\,\tau\, d\geqslant 1} \big|\widehat{F}_h(t)\big|\, dt \leqslant \frac{\tau\, d^{3/2}}{\sigma\,(2\,\pi)^{d/2}\,(\det \mathbb{D})^{1/2}}\,, \tag{3.1}$$

$$(2\,\pi)^{-d} \int\limits_{\rho\|t\|\,\tau\, d\geqslant 1} \big|\widehat{F}_h(t)\big|\, dt \leqslant \frac{\tau^2\, d^{2}}{\sigma^2\,(2\,\pi)^{d/2}\,(\det \mathbb{D})^{1/2}}\,, \tag{3.2}$$

$$(2\,\pi)^{-d} \int\limits_{\rho\|t\|\,\tau\, d\geqslant 1} \big|\langle t, v\rangle\, \widehat{F}_h(t)\big|\, dt \leqslant \frac{\langle \mathbb{D}^{-1} v, v\rangle^{1/2}}{\delta\,(2\,\pi)^{d/2}\,(\det \mathbb{D})^{1/2}}\,, \qquad \text{for all} \quad v \in \mathbb{R}^d, \tag{3.3}$$

where $F_h = \overline{F}(h)$, $\mathbb{D} = \mathbb{D}(F) = \operatorname{cov} F$ 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 $h \in \mathbb{R}^d$, $\|h\|\,\tau < 1$.

The classes just defined are closely connected, in particular,

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

and

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

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

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

Therefore, the conditions (3.1)–(3.3) play in this paper the role which is similar to that of Sakhanenko (1984, inequality (49), p. 9) or Einmahl (1989, 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 3.3 and 3.5).

Lemma 3.2. *Let* $\tau \geqslant 0$, $\delta > 0$, $\rho > 0$, $\alpha > 0$ *and* $F = \mathcal{L}(\xi) \in \mathcal{A}_d^*(\tau,\delta,\rho)$. *Then* $G = \mathcal{L}(\alpha\xi) \in \mathcal{A}_d^*(\alpha\tau,\delta,\rho)$.

Proof. We shall use the notation involved in the definition of classes $\mathcal{A}_d^*(\tau,\delta,\rho)$. It is evident that

$$\operatorname{cov} G = \mathbb{D}(G) = \alpha^2\, \mathbb{D}(F), \qquad \det \mathbb{D}(G) = \alpha^{2d}\, \mathbb{D}(F), \qquad \sigma^2(G) = \alpha^2\, \sigma^2(F). \tag{3.7}$$

Using (3.2), (3.6) and (3.7), we see that

$$\begin{aligned}\int\limits_{\rho\|u\|\,\alpha\tau\, d\geqslant 1} \big|\widehat{G}_h(u)\big|\, du &= \int\limits_{\rho\|u\|\,\alpha\tau\, d\geqslant 1} \big|\widehat{F}_{\alpha h}(\alpha u)\big|\, du = \alpha^{-d} \int\limits_{\rho\|t\|\,\tau\, d\geqslant 1} \big|\widehat{F}_{\alpha h}(t)\big|\, dt \\ &\leqslant \frac{(2\pi)^{d/2}\, \tau^2\, d^2}{\alpha^d\, \sigma^2(F)\,\big(\det \mathbb{D}(F)\big)^{1/2}} = \frac{(2\pi)^{d/2}\, (\alpha\tau)^2\, d^2}{\sigma^2(G)\,\big(\det \mathbb{D}(G)\big)^{1/2}},\end{aligned} \tag{3.8}$$

if $h \in \mathbb{R}^d$, $\|h\|\,\alpha\tau < 1$. Similarly, by (3.3), (3.6) and (3.7), we obtain, for the same h and for all $v \in \mathbb{R}^d$,

$$\begin{aligned}\int\limits_{\rho\|u\|\,\alpha\tau\, d\geqslant 1} \big|\big\langle u,v\big\rangle\, \widehat{G}_h(u)\big|\, du &= \int\limits_{\rho\|u\|\,\alpha\tau\, d\geqslant 1} \big|\big\langle u,v\big\rangle\, \widehat{F}_{\alpha h}(\alpha u)\big|\, du = \alpha^{-d-1} \int\limits_{\rho\|t\|\,\tau\, d\geqslant 1} \big|\big\langle t,v\big\rangle\, \widehat{F}_{\alpha h}(t)\big|\, dt \\ &\leqslant \frac{(2\pi)^{d/2}\, \alpha^{-1}\, \big\langle(\mathbb{D}(F))^{-1} v, v\big\rangle^{1/2}}{\delta\, \alpha^d\, \big(\det \mathbb{D}(F)\big)^{1/2}} = \frac{(2\pi)^{d/2}\, \big\langle(\mathbb{D}(G))^{-1} v, v\big\rangle^{1/2}}{\delta\, \big(\det \mathbb{D}(G)\big)^{1/2}}.\end{aligned} \tag{3.9}$$

The inequalities (3.8) and (3.9) provide the analogs of (3.2) and (3.3) for the distribution G with $\alpha\tau$ instead of τ. It remains to note that the relation $G \in \mathcal{A}_d(\alpha\tau)$ can be easily derived from $F \in \mathcal{A}_d(\tau)$ by an application of Lemma 2.5. $\square$

The following Lemmas 3.3 and 3.5 were proved in Zaitsev (1995, 1996, 1998).

Lemma 3.3. *Let the distribution of a random vector* $\xi \in \mathbb{R}^d$ *with* $\mathbf{E}\,\xi = 0$ *satisfy the condition* $\mathcal{L}(\xi) \in \overline{\mathcal{A}}_d(\tau, 4)$, $\tau \geqslant 0$. *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* $\operatorname{cov}\xi$. *Then there exist absolute positive constants* $c_4, \dots, c_8$ *such that the following assertions hold*:

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

$$|y| \leqslant c_4\,\tau\,\left\|\mathbb{B}^{-1/2}x\right\|^2 \leqslant c_4\,\tau\,\frac{\|x\|^2}{\sigma^2}, \tag{3.10}$$

and

$$\Phi_\sigma\bigl(z - \gamma(z)\bigr) < F(z + y\,|\,x) < \Phi_\sigma\bigl(z + \gamma(z)\bigr), \tag{3.11}$$

provided that $\dfrac{\tau\, d^{3/2}}{\sigma} \leqslant c_5$, $\left|\mathbb{B}^{-1/2}x\right| \leqslant \dfrac{c_6\,\sigma}{d^{3/2}\tau}$, $|z| \leqslant \dfrac{c_7\,\sigma^2}{d\tau}$, *where*

$$\gamma(z) = c_8\,\tau\left(d^{3/2} + d\delta\left(1 + \frac{|z|}{\sigma}\right) + \frac{z^2}{\sigma^2}\right), \qquad \delta = \left\|\mathbb{B}^{-1/2}x\right\|. \tag{3.12}$$

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

Definition 3.4. For $\tau \geqslant 0$ 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 3.5. *Let the distributions of random vectors* $\xi_1, \xi_2 \in \mathbb{R}^d$ *with* $\mathbf{E}\,\xi_1 = \mathbf{E}\,\xi_2 = 0$ *satisfy the conditions* $\bigl(\mathcal{L}(\xi_1), \mathcal{L}(\xi_2)\bigr) \in \mathcal{R}_d(\tau)$, *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* $\operatorname{cov}\xi_1 = \operatorname{cov}\xi_2$. *Then there exist absolute positive constants* $c_9, \dots, c_{14}$ *such that the following assertions hold*:

a) *Let* $d \geqslant 2$. *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* $\mathbb{B} = \operatorname{cov}\overline{\mathbb{P}}_{d-1}\xi_1 = \operatorname{cov}\overline{\mathbb{P}}_{d-1}\xi_2$ *and denote by* $F_k(z\,|\,x_k)$, $z \in \mathbb{R}^1$, *the r.c.d.f. of* $\xi_{k,d}$ *for given values of* $\overline{\mathbb{P}}_{d-1}\xi_k = x_k \in \mathbb{R}^{d-1}$. *Let* $\mathcal{L}(\overline{\mathbb{P}}_{d-1}\xi_k) \in \mathcal{A}_{d-1}^*(\tau, 4, 4)$, *for* $k = 1, 2$. *Then there exist* $y_1, y_2 \in \mathbb{R}^1$ *such that*

$$|y_k| \leqslant c_9\,\tau\,\left\|\mathbb{B}^{-1/2}x_k\right\|^2 \leqslant c_9\,\tau\,\frac{\|x_k\|^2}{\sigma^2}, \qquad \text{for} \quad k = 1, 2, \tag{3.13}$$

$$\bigl|y_1 - y_2\bigr| \leqslant c_{10}\,\bigl(\tau^2\sigma^{-1}\,\Gamma^3 + \tau\Lambda\Gamma\bigr), \tag{3.14}$$

and

$$F_1(z-\varkappa(z)+y_1\,|\,x_1)<F_2(z+y_2\,|\,x_2)<F_1(z+\varkappa(z)+y_1\,|\,x_1), \tag{3.15}$$

provided that $\dfrac{\tau\, d^{3/2}}{\sigma}\leqslant c_{11}$, $|z|\leqslant\dfrac{c_{12}\,\sigma^2}{d\,\tau}$, $\left|\mathbb{B}^{-1/2}x_k\right|\leqslant\dfrac{c_{13}\,\sigma}{d^{3/2}\tau}$, $k=1,2$, *where*

$$\varkappa(z)=c_{14}\,\frac{\tau^2}{\sigma}\,\Big(d^2+\beta\Big(d+\frac{|z|}{\sigma}\Big)+\frac{|z|^3}{\sigma^3}\Big),\qquad \beta=\Gamma^2+\sigma\,\tau^{-1}\Lambda, \tag{3.16}$$

$$\Lambda=\big\|\mathbb{B}^{-1/2}x_1-\mathbb{B}^{-1/2}x_2\big\|,\qquad \Gamma=\max_{k=1,2}\big\|\mathbb{B}^{-1/2}x_k\big\|. \tag{3.17}$$

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

Remark 3.6. In Zaitsev (1995, 1996, 1998) the formulation of Lemma 3.3 is in some sense weaker, see Zaitsev (1995, 1996, Lemmas 6.1 and 8.1) or (1998, Lemma 2.14). In particular, instead of the conditions

$$\mathcal{L}(\xi)\in\overline{\mathcal{A}}_d(\tau,4)\qquad\text{and}\qquad\mathcal{L}(\overline{\mathbb{P}}_{d-1}\xi)\in\overline{\mathcal{A}}_{d-1}(\tau,4) \tag{3.18}$$

the stronger conditions

$$\mathcal{L}(\xi)\in\mathcal{A}_d^*(\tau,4,4)\qquad\text{and}\qquad\mathcal{L}(\overline{\mathbb{P}}_{d-1}\xi)\in\mathcal{A}_{d-1}^*(\tau,4,4) \tag{3.19}$$

are used. However, in the proof of (3.10) and (3.11) only the conditions (3.18) are applied. The conditions (3.19) 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 Zaitsev (1995, 1996, 1998) simultaneously with the proof of (3.10) and (3.11), cf. Lemma 3.5.

4. Conditional distributions

In this section we expose some well-known standard facts about conditional distributions (see, e.g., Gihman and Skorokhod (1974, Chapter I, §3)). The results of this sections are present here for the sake of completeness. They will be used to justify the dyadic construction in Section 5.

Let $\big(\Omega,\mathfrak{S},\mathbf{P}\big)$ be a probability space, $\mathfrak{A}$ a σ-algebra of events contained in $\mathfrak{S}$ and ξ an arbitrary random variable with $\mathbf{E}\,|\xi|<\infty$. The conditional mathematical expectation of a random variable ξ given σ-algebra $\mathfrak{A}$ is defined as the random variable $\mathbf{E}\left\{\xi \,\middle|\, \mathfrak{A}\right\}$ which is measurable with respect to $\mathfrak{A}$ and satisfies for all $B \in \mathfrak{A}$ the equality

$$\int_B \mathbf{E}\left\{\xi \,\middle|\, \mathfrak{A}\right\} \mathbf{P}\{d\omega\} = \int_B \xi \, \mathbf{P}\{d\omega\} \tag{4.1}$$

The random variable $\mathbf{E}\left\{\xi \,\middle|\, \mathfrak{A}\right\}$ is unique (we identify the random variables which are equal to each other with probability 1).

The function $Q(B, \omega)$ defined on $\mathfrak{B}_1 \times \Omega$ is called the regular conditional distribution of a random variable ξ given σ-algebra $\mathfrak{A}$ if

a) for each fixed $B \in \mathfrak{B}_1$ the function $Q(B, \omega)$ is $\mathfrak{A}$-measurable;
b) there exists $A \in \mathfrak{A}$, with $\mathbf{P}\{A\} = 1$ such that, for each fixed $\omega \in A$, $Q(B, \omega)$ is a probability measure on $\mathfrak{B}_1$;
c) for all $B \in \mathfrak{B}_1$ and $A \in \mathfrak{A}$,

$$\int_A Q(B, \omega) \, \mathbf{P}\{d\omega\} = \mathbf{P}\Big\{\big\{\omega : \xi(\omega) \in B\big\} \cap A\Big\}. \tag{4.2}$$

It is well known that regular conditional distributions of one-dimensional random variables always exist. Recall that we denote by $\mathfrak{B}_j$ the σ-algebra of Borel subsets of the Euclidean space $\mathbb{R}^j$.

Let $\zeta = \big(\zeta_1, \zeta_2, \dots, \zeta_j\big)$ be a random vector with some distribution $\mathcal{L}(\zeta) \in \mathfrak{F}_j$. Then the conditional mathematical expectation $\mathbf{E}\left\{\xi \,\middle|\, \zeta\right\}$, of a random variable ξ given ζ is defined as $\mathbf{E}\left\{\xi \,\middle|\, \mathfrak{S}_\zeta\right\}$, where $\mathfrak{S}_\zeta$, is the σ-algebra generated by the mapping $\zeta : \big(\Omega, \mathfrak{S}\big) \mapsto \big(\mathbb{R}^j, \mathfrak{B}_j\big)$. More precisely,

$$\mathfrak{S}_\zeta = \Big\{S : S = \big\{\omega : \zeta(\omega) \in B\big\},\ B \in \mathfrak{B}_j\Big\}. \tag{4.3}$$

The conditional mathematical expectation $\mathbf{E}\left\{\xi \,\middle|\, \zeta\right\}$, can be written in the form $\mathbf{E}\left\{\xi \,\middle|\, \zeta\right\} = s(\zeta)$, where $s(\cdot)$ is a $\mathfrak{B}_j$-measurable function. The regular conditional distribution of ξ given ζ, is defined as the regular conditional distribution say, $Q(B, \omega)$, of ξ given σ-algebra $\mathfrak{S}_\zeta$. It can be shown that, for any fixed $B \in \mathfrak{B}_1$, there exists a $\mathfrak{B}_j$-measurable function $q(B, z)$, such that $Q(B, \omega) = q\big(B, \zeta(\omega)\big)$. There exist a set $C \in \mathfrak{B}_j$ with $\mathbf{P}\left\{\zeta \in C\right\} = 1$ such that, for each fixed $z \in C$, $q(B, z)$ is a probability measure on $\mathfrak{B}_1$. This measure is called the conditional distribution of ξ given $\zeta = z$. For any $\mathfrak{B}_{j+1}$-measurable non-negative function f the following equality is valid:

$$\mathbf{E}\, f(\xi, \zeta) = \int_{\mathbb{R}^j} \left(\int_{-\infty}^{\infty} f(x, z) \, q(dx,\, z)\right) F\{dz\}. \tag{4.4}$$

where $F = \mathcal{L}(\zeta)$. Moreover,

$$\mathbf{E}\left\{f(\xi, \zeta) \,\middle|\, \zeta\right\} = \int_{-\infty}^{\infty} f(x, \zeta) \, q(dx,\, \zeta). \tag{4.5}$$

The distribution function, which corresponds to the measure $q(\,\cdot\,, z)$, $z \in C$, and is defined by

$$G(y\,|\,z) = \int_{-\infty}^{\infty} \mathbf{I}\big\{\, x < y \,\big\}\, q(dx,\, z), \qquad y \in \mathbb{R}^1, \tag{4.6}$$

is called the regular conditional distribution function (r.c.d.f.) of ξ given $\zeta = z$. Here and below $\mathbf{I}\big\{\, A \,\big\}$ is the indicator function of an event A. For $z \in \mathbb{R}^j \setminus C$ the function $G(y\,|\,z)$ can be defined in an arbitrary way. For example, we can take

$$G(y\,|\,z) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}}\, \exp\Big(-\frac{x^2}{2}\Big)\, dx, \qquad z \in \mathbb{R}^j \setminus C. \tag{4.7}$$

The $G(y\,|\,z)$ is a $\mathfrak{B}_{j+1}$-measurable function of y and z. Therefore, $G(\xi\,|\,\zeta)$ can be considered as a random variable.

If ξ and ζ are independent, the r.c.d.f. $G(y\,|\,z)$ does not depend on z and coincides with $\mathbf{P}\big\{\, \xi < y \,\big\}$, the distribution function of ξ.

Assume that the joint distribution of ξ and ζ is absolutely continuous with respect to the $(j+1)$-dimensional Lebesgue measure. Let $p(x, z) \geqslant 0$, $x \in \mathbb{R}^1$, $z \in \mathbb{R}^j$, be the density of the distribution of the random vector $(\xi, \zeta) = \big(\xi, \zeta_1, \zeta_2, \ldots, \zeta_j\big) \in \mathbb{R}^{j+1}$. Let $p_\zeta(z) \geqslant 0$, $z \in \mathbb{R}^j$, be the density of the distribution of the random vector ζ. Then the function

$$p(x\,|\,z) = \frac{p(x, z)}{p_\zeta(z)}, \qquad z \in \mathbb{R}^j \setminus \big\{\, z : p_\zeta(z) = 0 \,\big\}, \tag{4.8}$$

is called the conditional density of the random variable ξ given $\zeta = z$. In this case

$$G(y\,|\,z) = \int_{-\infty}^{y} p(x\,|\,z)\, dx, \qquad y \in \mathbb{R}^1, \quad z \in \mathbb{R}^j \setminus \big\{\, z : p_\zeta(z) = 0 \,\big\}. \tag{4.9}$$

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

Proof. Let $F = \mathcal{L}(\zeta)$. According to (4.9), the r.c.d.f. $G(y\,|\,z)$ is continuous with respect to y for F-almost all z. Using (4.4), (4.6) and changing variables, we obtain

$$\begin{aligned}
\mathbf{P}\big\{\, G(\xi\,|\,\zeta) < y,\, \zeta \in D \,\big\} &= \int_{\mathbb{R}^j} \mathbf{I}\big\{\, z \in D \,\big\} \Big(\int_{-\infty}^{\infty} \mathbf{I}\big\{\, G(x\,|\,z) < y \,\big\}\, \partial_x G(x\,|\,z) \Big)\, F\{dz\} \\
&= \int_{\mathbb{R}^j} \mathbf{I}\big\{\, z \in D \,\big\} \Big(\int_{0}^{1} \mathbf{I}\big\{\, u < y \,\big\}\, du \Big)\, F\{dz\} \\
&= \mathbf{P}\big\{\, \eta < y \,\big\}\; \mathbf{P}\big\{\, \zeta \in D \,\big\},
\end{aligned} \tag{4.10}$$

for all $y \in \mathbb{R}^1$ and $D \in \mathfrak{B}_j$, where η is a random variable with uniform distribution on the interval $[0,\, 1]$. The equality (4.10) implies that the $(j+1)$-dimensional random vector $\big(G(\xi\,|\,\zeta), \zeta\big)$ has the needed joint distribution function and, hence, the needed joint distribution. $\square$

Lemma 4.2. *Suppose that the distribution of a random variable* $\xi \in \mathbb{R}^1$ *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* $[0, 1]$.

The proof of Lemma 4.2 is simpler than that of Lemma 4.1 and therefore is omitted.

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

Proof. Let $F = \mathcal{L}(\zeta)$ and $q(\,\cdot\,, z)$ the conditional distribution of ξ given $\zeta = z$. It is easy to see that $G^{-1}(x\,|\,z) < t$ iff $G(t\,|\,z) > x$. Using (4.4) and (4.6), we obtain

$$\begin{aligned}
\mathbf{P}\big\{ G^{-1}(\eta\,|\,\zeta) < t,\, \zeta \in D \big\} &= \int_{\mathbb{R}^j} \mathbf{I}\big\{ z \in D \big\} \bigg(\int_0^1 \mathbf{I}\big\{ G^{-1}(x\,|\,z) < t \big\}\, dx \bigg)\, F\{dz\} \\
&= \int_{\mathbb{R}^j} \mathbf{I}\big\{ z \in D \big\} \bigg(\int_0^1 \mathbf{I}\big\{ G(t\,|\,z) > x \big\}\, dx \bigg)\, F\{dz\} \\
&= \int_{\mathbb{R}^j} \mathbf{I}\big\{ z \in D \big\}\, G(t\,|\,z)\, F\{dz\} \\
&= \int_{\mathbb{R}^j} \mathbf{I}\big\{ z \in D \big\} \bigg(\int_{-\infty}^{\infty} \mathbf{I}\big\{ x < t \big\}\, q(dx,\, z) \bigg)\, F\{dz\} \\
&= \mathbf{P}\big\{ \xi < t,\, \zeta \in D \big\}, \qquad (4.11)
\end{aligned}$$

for all $t \in \mathbb{R}^1$ and $D \in \mathfrak{B}_j$. The equality (4.11) implies that the random vector $\big(G^{-1}(\eta\,|\,\zeta),\, \zeta\big)$ has the needed joint distribution function. $\square$

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* $G^{-1}(x) = \sup\big\{ y : G(y) \leqslant x \big\}$, $0 < x < 1$. *Then* $\mathcal{L}\big(G^{-1}(\eta)\big) = \mathcal{L}(\xi)$.

The proof of Lemma 4.4 is similar to that of Lemma 4.3 and therefore is omitted.

5. Dyadic scheme. General description

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

$$S_0^* = 0; \qquad S_k^* = \sum_{l=1}^{k} \xi_l, \quad 1 \leqslant k \leqslant 2^N; \tag{5.1}$$

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

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

$$\widetilde{\mathbf{U}}_{n,k}^* = \big(U_{n-1,2k-1}^*, U_{n-1,2k}^*\big) \in \mathbb{R}^{2d}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{5.3}$$

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

$$U_{n-1,2k-1}^* + U_{n-1,2k}^* = U_{n,k}^*, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{5.4}$$

Let $\mathbb{A}_{n,k} : \mathbb{R}^{2d} \mapsto \mathbb{R}^{2d}$ be non-degenerate linear operators such that the first d coordinates of the vector

$$\mathbf{U}_{n,k}^* = \mathbb{A}_{n,k}\, \widetilde{\mathbf{U}}_{n,k}^* \in \mathbb{R}^{2d}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{5.5}$$

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

$$\mathcal{U}^* = \mathcal{U}_{N,1}^* = \mathbb{J}_{N,1}\, U_{N,1}^*. \tag{5.6}$$

Below we describe a procedure of constructing the random vectors $\{U_{n,k}\}$ with $\mathcal{L}\big(\{U_{n,k}\}\big) = \mathcal{L}\big(\{U_{n,k}^*\}\big)$, provided that the vectors $Y_1, \ldots, Y_{2^N}$ are already constructed. To this end purpose we shall use the so-called Rosenblatt quantile transformation (see Rosenblatt (1952) and Einmahl (1989)).

Denote by $F^{(1)}(x_1) = \mathbf{P}\big\{\mathbb{P}_1 \mathcal{U}^* < x_1\big\}$, $x_1 \in \mathbb{R}^1$, the distribution function of the first coordinate of the vector $\mathcal{U}^*$. Introduce the conditional distributions, denoting by $F^{(j)}\big(\,\cdot\,\big|\, x_1, \ldots, x_{j-1}\big)$, $2 \leqslant j \leqslant d$, r.c.d.f. of $\mathbb{P}_j\, \mathcal{U}^*$, given $\overline{\mathbb{P}}_{j-1}\mathcal{U}^* = (x_1, \ldots, x_{j-1})$. Denote by $\widetilde{F}_{n,k}^{(j)}\big(\,\cdot\,\big|\, x_1, \ldots, x_{j-1}\big)$ the r.c.d.f. of $\mathbb{P}_j\, \mathbf{U}_{n,k}^*$, given

$\overline{\mathbb{P}}_{j-1}\,\mathbf{U}^*_{n,k} = (x_1, \dots, x_{j-1})$, $1 \leqslant k \leqslant 2^{N-n}$, $1 \leqslant n \leqslant N$, $d+1 \leqslant j \leqslant 2d$. Put

$$T_0 = 0, \qquad T_k = \sum_{l=1}^{k} Y_l, \qquad 1 \leqslant k \leqslant 2^N; \tag{5.7}$$

$$\begin{aligned} V_{m,k} &= \left(V^{(1)}_{m,k}, \dots, V^{(d)}_{m,k}\right) = T_{k\cdot 2^m} - T_{(k-1)\cdot 2^m}, \\ &\qquad\qquad\qquad\qquad 1 \leqslant k \leqslant 2^{N-m}, \quad 0 \leqslant m \leqslant N; \\ \widetilde{\mathbf{V}}_{n,k} &= \left(V_{n-1,2k-1},\, V_{n-1,2k}\right) = \left(\widetilde{\mathbf{V}}^{(1)}_{n,k}, \dots, \widetilde{\mathbf{V}}^{(2d)}_{n,k}\right) \in \mathbb{R}^{2d}, \\ &\qquad\qquad\qquad\qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N; \end{aligned} \tag{5.8}$$

and

$$\mathbf{V}_{n,k} = \left(\mathbf{V}^{(1)}_{n,k}, \dots, \mathbf{V}^{(2d)}_{n,k}\right) = \mathbb{A}_{n,k}\,\widetilde{\mathbf{V}}_{n,k} \in \mathbb{R}^{2d}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{5.9}$$

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

$$\left.\begin{aligned} \overline{\mathbb{P}}_d\,\mathbf{U}^*_{n,k} &= \mathbb{J}_{n,k}\,U^*_{n,k} = \mathcal{U}^*_{n,k}, \\ \overline{\mathbb{P}}_d\,\mathbf{V}_{n,k} &= \mathbb{J}_{n,k}\,V_{n,k} \overset{\text{def}}{=} \mathcal{V}_{n,k}, \end{aligned}\right. \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{5.10}$$

Moreover,

$$V_{n,k} = V_{n-1,2k-1} + V_{n-1,2k}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{5.11}$$

and

$$V_{N,1} = Y_1 + \dots + Y_{2^N}. \tag{5.12}$$

Set

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

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

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

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

$$\mathcal{G}^{(1)} = G^{(1)}\big(\mathcal{V}^{(1)}\big), \qquad \mathcal{U}^{(1)} = \Big(F^{(1)}\Big)^{-1}\big(\mathcal{G}^{(1)}\big), \tag{5.14}$$

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

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

(here $\Big(F^{(1)}\Big)^{-1}(x) = \sup\big\{y : F^{(1)}(y) \leqslant x\big\}$, $0 < x < 1$, and so on). If the distributions of the random vectors $\xi_1, \dots, \xi_{2^N}$ are absolutely continuous too, then the formulas (5.14) and (5.15) can be rewritten in a more natural "symmetric" form (see Sakhanenko (1984, pp. 30–31)):

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

Define

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

Introduce also the random vectors

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

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

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

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

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

$$\begin{aligned} \mathcal{G}_{n,k}^{(j)} &= \widetilde{G}_{n,k}^{(j)}\big(\mathbf{V}_{n,k}^{(j)} \,\big|\, \mathbf{V}_{n,k}^{(1)}, \dots, \mathbf{V}_{n,k}^{(j-1)}\big) \\ \mathbf{U}_{n,k}^{(j)} &= \Big(\widetilde{F}_{n,k}^{(j)}\Big)^{-1}\Big(\mathcal{G}_{n,k}^{(j)} \,\Big|\, \mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(j-1)}\Big). \end{aligned} \tag{5.20}$$

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

$$\begin{aligned}
\mathbf{U}_{n,k} &= \left(\mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(2d)}\right) \in \mathbb{R}^{2d}, \\
\mathbf{U}_{n,k}^{j} &= \left(\mathbf{U}_{n,k}^{(1)}, \dots, \mathbf{U}_{n,k}^{(j)}\right) = \overline{\mathbb{P}}_j\, \mathbf{U}_{n,k} \in \mathbb{R}^{j}, \\
\mathbf{V}_{n,k}^{j} &= \left(\mathbf{V}_{n,k}^{(1)}, \dots, \mathbf{V}_{n,k}^{(j)}\right) = \overline{\mathbb{P}}_j\, \mathbf{V}_{n,k} \in \mathbb{R}^{j}, \\
\widetilde{\mathbf{U}}_{n,k} &= \left(\widetilde{\mathbf{U}}_{n,k}^{(1)}, \dots, \widetilde{\mathbf{U}}_{n,k}^{(2d)}\right) = \mathbb{A}_{n,k}^{-1}\, \mathbf{U}_{n,k} \in \mathbb{R}^{2d},
\end{aligned} \qquad j = 1, \dots, 2\,d, \tag{5.21}$$

and

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

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

$$X_k = U_{0,k}, \qquad S_0 = 0, \quad S_k = \sum_{l=1}^{k} X_l, \qquad 1 \leqslant k \leqslant 2^N. \tag{5.23}$$

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

$$\begin{aligned}
\widetilde{\mathbf{U}}_{n,k} &= \left(U_{n-1,2k-1}, U_{n-1,2k}\right) \in \mathbb{R}^{2d} \\
U_{n,k} &= U_{n-1,2k-1} + U_{n-1,2k} = S_{k\cdot 2^n} - S_{(k-1)\cdot 2^n}, \\
&= X_{(k-1)\cdot 2^n + 1} + \dots + X_{k\cdot 2^n} \\
\mathbf{U}_{n,k} &= \mathbb{A}_{n,k}\, \widetilde{\mathbf{U}}_{n,k}, \\
\overline{\mathbb{P}}_d\, \mathbf{U}_{n,k} &= \mathbb{J}_{n,k} U_{n,k} = \mathcal{U}_{n,k},
\end{aligned} \tag{5.24}$$

and

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

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

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

$$\mathcal{L}\big(\overline{\mathbb{P}}_j\, \mathcal{U}\big) = \mathcal{L}\big(\overline{\mathbb{P}}_j\, \mathcal{U}^*\big), \qquad j = 1, \dots, d. \tag{5.26}$$

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 $2 \leqslant s \leqslant d$. 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 $\big(\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(s-1)}\big)$ and has uniform distribution on the interval $[0, 1]$. Furthermore, $\mathcal{G}^{(s)}$ is independent of $\big(\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(s-1)}\big)$ too, since $\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(s-1)}$ are measurable functions of $\big(\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(s-1)}\big)$ (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, $0 \leqslant n \leqslant N$, the random vectors $U_{n,k}$, $k = 1, \dots, 2^{N-n}$, are jointly independent and $\mathcal{L}\big(U_{n,k}\big) = \mathcal{L}\big(U^*_{n,k}\big)$. This imply the statement of the lemma, since, according to (5.4) and (5.24), all the vectors $U_{n,l}$ can be obtained from $\big\{U_{0,k}\big\}$ in the same way as $U^*_{n,l}$ can be obtained from $\big\{U^*_{0,k}\big\}$.

The case $n = N$ was already considered above. Suppose that the induction conjecture is satisfied for $n = m$, $1 \leqslant m \leqslant N$. 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 $\mathcal{L}\big(\mathbf{U}_{m,k}\big) = \mathcal{L}\big(\mathbf{U}^*_{m,k}\big)$. This is sufficient since the operators $\mathbb{A}_{m,k}$ are non-degenerate (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

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

(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 $\overline{\mathbb{P}}_d\mathbf{V}_{m,k} = \mathbb{J}_{m,k}V_{m,k}$ and the random variables $\mathcal{G}^{(d+1)}_{m,k}, \dots, \mathcal{G}^{(2d)}_{m,k}$ are independent in the aggregate. Moreover, we shall prove that $\mathcal{G}^{(j)}_{m,k}$, $j = d+1, \dots, 2d$, 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 $\mathcal{G}^{(j)}_{m,k}$ has uniform distribution on the interval $[0, 1]$ and is independent of $\overline{\mathbb{P}}_{j-1}\mathbf{V}_{m,k}$. Furthermore, it is independent of $\overline{\mathbb{P}}_{j-1}\mathcal{H}_{m,k}$ because $\overline{\mathbb{P}}_{j-1}\mathcal{H}_{m,k}$ is a function of $\overline{\mathbb{P}}_{j-1}\mathbf{V}_{m,k}$ (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 $\mathcal{G}^{(j)}_{m,k}$, $k = 1, \dots, 2^{N-m}$, $j = d+1, \dots, 2d$, are uniformly distributed on the interval $[0, 1]$ and independent of the random vectors $\big\{\overline{\mathbb{P}}_d\mathbf{V}_{m,l} = \mathbb{J}_{m,l}V_{m,l},\ l = 1, \dots, 2^{N-m}\big\}$. 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

$\{V_{m,l} = \mathbb{J}_{m,l}^{-1}\overline{\mathbb{P}}_d \mathbf{V}_{m,l},\ l = 1,\dots,2^{N-m}\}$ (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

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

(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

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

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

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

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 $d+1 \leqslant s \leqslant 2d$. It suffices to prove (5.29) for $j = s$. The random variable $\mathcal{G}_{m,k}^{(s)}$ is independent of $\overline{\mathbb{P}}_{s-1}\mathcal{J}_{m,k}$ and has uniform distribution on the interval $[0,1]$. Furthermore, $\mathcal{G}_{m,k}^{(s)}$ is independent of $\overline{\mathbb{P}}_{s-1}\mathbf{U}_{m,k}$ too, since $\overline{\mathbb{P}}_{s-1}\mathbf{U}_{m,k}$ is a function of $\overline{\mathbb{P}}_{s-1}\mathcal{J}_{m,k}$ (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). $\square$

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

Remark 5.2. It is easy to see that, for some fixed m such that $0 \leqslant m \leqslant N$, the vectors $U_{n,k}$, $m \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$, are the quantile transformations of the vectors $V_{m,1},\dots,V_{m,2^{N-m}}$ associated with the operators $\mathbb{A}_{n,k}$, $m \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$. This follows from the sequential procedure of constructing the sums over blocks of 2^n summands $U_{n,k}$.

Lemma 5.3. *Let* $\alpha > 0$ *and 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* $\mathbb{A}_{n,k}$. *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* $\mathbb{A}_{n,k}$.

Proof. Denote by $F_\alpha^{(1)}(x_1) = \mathbf{P}\left\{\mathbb{P}_1\,\mathcal{U}_\alpha < x_1\right\}$, $x_1\in\mathbb{R}^1$, the distribution function of the first coordinate of the vector $\mathcal{U}_\alpha = \alpha\,\mathcal{U} = \mathbb{J}_{N,1}\big(\alpha X_1+\dots+\alpha X_{2^N}\big)$. Consider the conditional distributions, denoting by $F_\alpha^{(j)}\big(\,\cdot\,\big|\,x_1,\dots,x_{j-1}\big)$, $2\leqslant j\leqslant d$, the r.c.d.f. of $\mathbb{P}_j\,\mathcal{U}_\alpha$, given $\overline{\mathbb{P}}_{j-1}\mathcal{U}_\alpha = (x_1,\dots,x_{j-1})$. Denote by $G_\alpha^{(1)}(x_1) = \mathbf{P}\left\{\mathbb{P}_1\,\mathcal{V}_\alpha < x_1\right\}$, $x_1\in\mathbb{R}^1$, the distribution function of the first coordinate of the vector $\mathcal{V}_\alpha = \alpha\,\mathcal{V} = \mathbb{J}_{N,1}\big(\alpha Y_1+\dots+\alpha Y_{2^N}\big)$. Let $G_\alpha^{(j)}\big(\,\cdot\,\big|\,x_1,\dots,x_{j-1}\big)$, $2\leqslant j\leqslant d$, be the r.c.d.f. of $\mathbb{P}_j\,\mathcal{V}_\alpha$, given $\overline{\mathbb{P}}_{j-1}\,\mathcal{V}_\alpha = (x_1,\dots,x_{j-1})$.

It is easy to see that $G_\alpha^{(1)}(x_1) = G^{(1)}(x_1/\alpha)$, for all $x_1\in\mathbb{R}^1$. Moreover, $G_\alpha^{(j)}\big(\,x_j\,\big|\,x_1,\dots,x_{j-1}\big) = G^{(j)}\big(\,x_j/\alpha\,\big|\,x_1/\alpha,\dots,x_{j-1}/\alpha\big)$, $2\leqslant j\leqslant d$, $x_j\in\mathbb{R}^1$. Furthermore, it is not difficult to verify that $\big(F_\alpha^{(1)}\big)^{-1}(y) = \alpha\big(F^{(1)}\big)^{-1}(y)$, and $\big(F_\alpha^{(j)}\big)^{-1}\big(\,y\,\big|\,x_1,\dots,x_{j-1}\big) = \alpha\big(F^{(j)}\big)^{-1}\big(\,y\,\big|\,x_1/\alpha,\dots,x_{j-1}/\alpha\big)$, $2\leqslant j\leqslant d$, for all $y\in\mathbb{R}^1$. Thus, we have

$$\big(F_\alpha^{(1)}\big)^{-1}\big(G_\alpha^{(1)}\big(\mathbb{P}_1\,\mathcal{V}_\alpha\big)\big) = \alpha\big(F^{(1)}\big)^{-1}\big(G^{(1)}\big(\mathbb{P}_1\,\mathcal{V}\big)\big) = \alpha\,\mathcal{U}^{(1)} = \mathbb{P}_1\,\mathcal{U}_\alpha, \tag{5.30}$$

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

$$\begin{aligned}\big(F_\alpha^{(j)}\big)^{-1}\Big(G_\alpha^{(j)}\big(\mathbb{P}_j\,\mathcal{V}_\alpha\,\big|\,\overline{\mathbb{P}}_{j-1}\,\mathcal{V}_\alpha\big)\,\Big|\,\overline{\mathbb{P}}_{j-1}\mathcal{U}_\alpha\Big) &= \alpha\big(F^{(j)}\big)^{-1}\Big(G^{(j)}\big(\mathbb{P}_j\,\mathcal{V}\,\big|\,\overline{\mathbb{P}}_{j-1}\,\mathcal{V}\big)\,\Big|\,\overline{\mathbb{P}}_{j-1}\mathcal{U}\Big)\\ &= \alpha\,\mathcal{U}^{(j)} = \mathbb{P}_j\,\mathcal{U}_\alpha.\end{aligned} \tag{5.31}$$

The relations (5.30) and (5.31) provide the required analogs of (5.14) and (5.15). The corresponding analogs of (5.20) can be proved in a similar way. It remains to note that linear equalities (such as, for example, (5.24)) are valid after multiplying the both sides of them by α. $\square$

For the formulation of the following below Lemma 5.4 we shall use the notation

$$U_{n,k}^- = U_{n-1,2k-1} - U_{n-1,2k}, \qquad 1\leqslant k\leqslant 2^{N-n}, \quad 1\leqslant n\leqslant N. \tag{5.32}$$

U. Einmahl has informed the author that the shortest proof of Lemma 5.4 can be obtained with the help of a geometrical approach due to Massart (1989, p. 275). We give this proof here for the sake of completeness.

Lemma 5.4 (Einmahl 1989, Lemma 5, p. 55). *Let* $1\leqslant m = (2s+1)\cdot 2^r\leqslant 2^N$, *where* s,r *are non-negative integers. Then the following representation is valid*:

$$S_m = \frac{m}{2^N}\,S_{2^N} + \sum_{n=r+1}^{N}\gamma_n\,U_{n,l_{n,m}}^-, \tag{5.33}$$

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

$$\big(l_{n,m} - 1\big) \cdot 2^n < m \leqslant l_{n,m} \cdot 2^n. \tag{5.34}$$

Remark 5.5. The inequalities (5.34) 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.

Proof of Lemma 5.4. Let $e_m = \big(e_m^{(1)}, \dots, e_m^{(2^N)}\big) \in \mathbb{R}^{2^N}$, $m = 1, \dots, 2^N$, be the standard orthonormal basis in $\mathbb{R}^{2^N}$:

$$e_m^{(j)} = \begin{cases} 1, & \text{if} \quad j = m, \\ 0, & \text{otherwise.} \end{cases} \tag{5.35}$$

Define the vectors $e_{n,k} = \big(e_{n,k}^{(1)}, \dots, e_{n,k}^{(2^N)}\big) \in \mathbb{R}^{2^N}$, $1 \leqslant k \leqslant 2^{N-n}$, $0 \leqslant n \leqslant N$, by

$$e_{n,k}^{(j)} = \begin{cases} 1, & \text{if} \quad (k-1) \cdot 2^n + 1 \leqslant j \leqslant k \cdot 2^n, \\ 0, & \text{otherwise.} \end{cases} \tag{5.36}$$

Obviously,

$$e_{n,k} = e_{n-1,2k-1} + e_{n-1,2k}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N. \tag{5.37}$$

Let, furthermore,

$$e_{n,k}^{-} = e_{n-1,2k-1} - e_{n-1,2k}, \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N, \tag{5.38}$$

and

$$I_m = e_1 + \dots + e_m, \qquad m = 1, \dots, 2^N. \tag{5.39}$$

It is easy to see that $e_m = e_{0,m}$, $m = 1, \dots, 2^N$, and

$$\begin{aligned} e_{n,k} &= e_{(k-1)\cdot 2^n + 1} + \dots + e_{k \cdot 2^n} \\ &= I_{k\cdot 2^n} - I_{(k-1)\cdot 2^n}, \end{aligned} \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad 0 \leqslant n \leqslant N. \tag{5.40}$$

For $x = \big(x_1, \dots, x_{2^N}\big) \in \mathbb{R}^{2^N}$ $(x_j \in \mathbb{R}^1)$ and $y = \big(y_1, \dots, y_{2^N}\big) \in (\mathbb{R}^d)^{2^N}$ $(y_j \in \mathbb{R}^d)$ define the "scalar product" $\big\langle x, y \big\rangle = x_1 y_1 + \dots + x_{2^N} y_{2^N}$. Using this notation, we can rewrite the formula (5.23) in the form

$$S_m = \big\langle I_m, X \big\rangle, \qquad m = 1, \dots, 2^N, \tag{5.41}$$

where
$$X=\bigl(X_1,\dots,X_{2^N}\bigr)\in(\mathbb{R}^d)^{2^N}. \tag{5.42}$$
According to (5.24), (5.32), (5.38), (5.40) and (5.41), we have
$$U_{n,k}=\bigl\langle e_{n,k},\,X\bigr\rangle,\qquad 1\leqslant k\leqslant 2^{N-n},\quad 0\leqslant n\leqslant N, \tag{5.43}$$
and
$$U^-_{n,k}=\bigl\langle e^-_{n,k},\,X\bigr\rangle,\qquad 1\leqslant k\leqslant 2^{N-n},\quad 1\leqslant n\leqslant N. \tag{5.44}$$
It can be easily verified that
$$\bigl\{e^-_{n,k},\ 1\leqslant k\leqslant 2^{N-n},\ 1\leqslant n\leqslant N\bigr\}\cup\bigl\{e_{N,1}\bigr\}\subset\mathbb{R}^{2^N} \tag{5.45}$$
is a collection of 2^N orthogonal non-zero vectors. Therefore, each vector $x\in\mathbb{R}^{2^N}$ can be represented in the form
$$x=\frac{\langle x,e_{N,1}\rangle\,e_{N,1}}{\|e_{N,1}\|^2}+\sum_{n=1}^{N}\sum_{k=1}^{2^{N-n}}\frac{\langle x,e^-_{n,k}\rangle\,e^-_{n,k}}{\|e^-_{n,k}\|^2}. \tag{5.46}$$
In particular,
$$I_m=\frac{\langle I_m,\,e_{N,1}\rangle\,e_{N,1}}{\|e_{N,1}\|^2}+\sum_{n=1}^{N}\sum_{k=1}^{2^{N-n}}\frac{\langle I_m,\,e^-_{n,k}\rangle\,e^-_{n,k}}{\|e^-_{n,k}\|^2},\qquad m=1,\dots,2^N. \tag{5.47}$$
Let $m=(2\,s+1)\cdot 2^r$, where s,r are non-negative integers. Then, obviously,
$$\bigl\langle I_m,\,e^-_{n,k}\bigr\rangle=0\qquad\text{for}\quad 1\leqslant k\leqslant 2^{N-n},\quad 0\leqslant n\leqslant r. \tag{5.48}$$
Moreover, for fixed n, $r+1\leqslant n\leqslant N$,
$$\bigl\langle I_m,\,e^-_{n,k}\bigr\rangle=0\qquad\text{for}\quad 1\leqslant k\leqslant 2^{N-n},\quad k\neq l_{n,m}, \tag{5.49}$$
and
$$0\leqslant\bigl\langle I_m,\,e^-_{n,l_{n,m}}\bigr\rangle\leqslant\frac12\,\|e^-_{n,l_{n,m}}\|^2=2^{n-1} \tag{5.50}$$
(see (5.37), (5.38), (5.43), (5.44) and Remark 5.5). Furthermore,
$$\frac{\langle I_m,e_{N,1}\rangle}{\|e_{N,1}\|^2}=\frac{m}{2^N}\quad\text{and}\quad\bigl\langle e_{N,1},\,X\bigr\rangle=S_{2^N}. \tag{5.51}$$

The equality (5.33) with $0\leqslant\gamma_n=\dfrac{\langle I_m,\,e^-_{n,l_{n,m}}\rangle}{\|e^-_{n,l_{n,m}}\|^2}\leqslant\dfrac12$ can be easily derived from (5.41), (5.44) and (5.47)–(5.51). $\square$

Lemma 5.6. *Let* $1 \leqslant m = (2s+1) \cdot 2^r \leqslant 2^N$, *where* s, r *are non-negative integers. Then the following representation is valid*:

$$S_m = \sum_{n=r}^{N} \beta_n U_{n,l_{n,m}}, \tag{5.52}$$

where real valued coefficients $\beta_n = \beta_n(m)$ *satisfy* $|\beta_n| \leqslant 3/2$ *and the integers* $l_{n,m}$ *are defined by* (5.34).

Proof. Using (5.24), (5.32) and Remark 5.5, we see that

$$U_{n-1,l_{n-1,m}} = \frac{1}{2} U_{n,l_{n,m}} - \frac{b_n}{2} U^-_{n,l_{n,m}}, \qquad n = 1, \ldots, N, \tag{5.53}$$

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

$$U^-_{n,l_{n,m}} = b_n \left(U_{n,l_{n,m}} - 2 U_{n-1,l_{n-1,m}} \right), \qquad n = 1, \ldots, N. \tag{5.54}$$

Note that $l_{N,m} = 1$ (see (5.34) or Remark 5.5). Substituting (5.54) in (5.33) and using (5.25), we get

$$S_m = \frac{m}{2^N} U_{N,1} + \sum_{n=r+1}^{N} \gamma_n b_n \left(U_{n,l_{n,m}} - 2 U_{n-1,l_{n-1,m}} \right) = \sum_{n=r}^{N} \beta_n U_{n,l_{n,m}} \tag{5.55}$$

with

$$\beta_n = \begin{cases} 1, & \text{if} \quad n = r = N, \\ 2^{-N} m + \gamma_N b_N, & \text{if} \quad n = N, \ r \neq N, \\ \gamma_n b_n - 2\gamma_{n+1} b_{n+1}, & \text{if} \quad r+1 \leqslant n \leqslant N-1, \\ -2\gamma_{r+1} b_{r+1}, & \text{if} \quad n = r \neq N. \end{cases} \tag{5.56}$$

In order to verify that $|\beta_n| \leqslant 3/2$, it remains to recall that $|b_j| = 1$, $0 \leqslant \gamma_j \leqslant {}^1\!/_2$, and $0 < m \leqslant 2^N$. $\square$

6. Dyadic scheme. Conditions B and their consequences

In the sequel we shall always suppose that the dyadic scheme described above satisfies the following additional assumptions $\mathbf{B}_1$–$\mathbf{B}_4$ which we call for the sake of brevity Condition $\mathbf{B}$ (cf. Condition $\mathbf{B}_2$ in Section 1).

$\mathbf{B}_1$. The vectors $\xi_1, \ldots, \xi_{2^N}$, $X_1, \ldots, X_{2^N}$ and $Y_1, \ldots, Y_{2^N}$ (involved in the dyadic scheme) have the same zero means and the same covariance operators as certain vectors $\widetilde{\xi}_1, \ldots, \widetilde{\xi}_{2^N}$:

$$\begin{aligned} \mathbf{E}\, \widetilde{\xi}_k &= \mathbf{E}\, \xi_k = \mathbf{E}\, X_k = \mathbf{E}\, Y_k = 0, \\ \operatorname{cov} \widetilde{\xi}_k &= \operatorname{cov} \xi_k = \operatorname{cov} X_k = \operatorname{cov} Y_k \overset{\text{def}}{=} \mathbb{B}_k, \end{aligned} \qquad k = 1, \ldots, 2^N. \tag{6.1}$$

Moreover, let $\widetilde{\eta}_k$ be auxiliary random vectors which have the same moments of the first three orders as the vectors $\widetilde{\xi}_k$, $k = 1, \dots, 2^N$.

B$_2$. The covariance operators $\mathbb{B}_k$ for all $u \in \mathbb{R}^d$ and for all $k = 1, \dots, 2^N$ satisfy the condition

$$C_1^2 \, \|u\|^2 \leqslant \big\langle \mathbb{B}_k u, u \big\rangle = \big\| \mathbb{B}_k^{1/2} u \big\|^2 \leqslant C_2^2 \, \|u\|^2 \tag{6.2}$$

with some absolute constants C_1 and C_2.

B$_3$. The first d coordinates of the vector $\mathbf{U}_{n,k} = \mathbb{A}_{n,k} \widetilde{\mathbf{U}}_{n,k}$, $1 \leqslant k \leqslant 2^{N-n}$, $1 \leqslant n \leqslant N$, are defined by the relation $\overline{\mathbb{P}}_d \, \mathbf{U}_{n,k} = \mathbb{J}_{n,k} U_{n,k}$, where

$$\mathbb{J}_{n,k} = 2^{n/2} \, \mathbb{D}_{n,k}^{-1/2}, \quad \mathbb{D}_{n,k} = \mathbb{B}_{(k-1)\cdot 2^n + 1} + \cdots + \mathbb{B}_{k \cdot 2^n} = \operatorname{cov} U_{n,k}. \tag{6.3}$$

B$_4$. The last d coordinates of the vector $\mathbf{U}_{n,k} = \mathbb{A}_{n,k} \widetilde{\mathbf{U}}_{n,k}$, $1 \leqslant k \leqslant 2^{N-n}$, $1 \leqslant n \leqslant N$, are defined by the relations

$$\mathbb{P}_j \, \mathbf{U}_{n,k} = \mathbb{P}_{j-d} \, \widetilde{\mathcal{U}}_{n,k}, \qquad j = d+1, \dots, 2d, \tag{6.4}$$

where

$$\begin{aligned} &\widetilde{\mathcal{U}}_{n,k} = \widetilde{\mathbb{J}}_{n,k} \widetilde{U}_{n,k}, \qquad \widetilde{\mathbb{J}}_{n,k} = 2^{n/2} \, \widetilde{\mathbb{D}}_{n,k}^{-1/2}, \qquad \widetilde{\mathbb{D}}_{n,k} = \operatorname{cov} \widetilde{U}_{n,k}, \\ &\widetilde{U}_{n,k} = 2 \, \mathbb{D}_{n-1,2k} \, \mathbb{D}_{n,k}^{-1} \, U_{n-1,2k-1} - 2 \, \mathbb{D}_{n-1,2k-1} \, \mathbb{D}_{n,k}^{-1} \, U_{n-1,2k}. \end{aligned} \tag{6.5}$$

Remark 6.1. It is convenient for us to suppose that C_1, C_2 are absolute constants. In this case all the quantities depending only on C_1, C_2 can be also considered as absolute constants.

Remark 6.2. The vectors $\widetilde{\xi}_1, \dots, \widetilde{\xi}_{2^N}$ play below the role of vectors the independent copies of which are to be constructed provided that we already have the independent Gaussian vectors $Y_1, \dots, Y_{2^N}$. However, in general case, one should apply the dyadic scheme several times and with distributions of ξ_k not necessarily coinciding with those of $\widetilde{\xi}_k$, see Conditions $\mathbf{C}_\Psi(p)$ below. Moreover, the vectors $Y_1, \dots, Y_{2^N}$ involved in the dyadic scheme may be non-Gaussian, see Condition $\mathbf{D}_\Psi$. The vectors $\widetilde{\eta}_1, \dots, \widetilde{\eta}_{2^N}$ will appear not always in the construction. They will be sometimes used as intermediate bounded vectors using which one can construct some smooth and bounded distributions denoted by $\Psi\big(\mathcal{L}(\widetilde{\eta}_1)\big), \dots, \Psi\big(\mathcal{L}(\widetilde{\eta}_{2^N})\big)$ and having the required moments of the first three orders (see (7.14) and Lemmas 7.3 and 11.2). In Sections 9 and 10, the vectors with distributions $\Psi\big(\mathcal{L}(\widetilde{\eta}_1)\big), \dots, \Psi\big(\mathcal{L}(\widetilde{\eta}_{2^N})\big)$ will be used as $Y_1, \dots, Y_{2^N}$.

Remark 6.3. The condition (6.1) automatically implies the coincidence of the corresponding first and second moments of the vectors $\mathbf{U} = \Big\{ U_{n,k}, \, \widetilde{\mathbf{U}}_{n,k}, \, \mathbf{U}_{n,k}, \, \mathcal{U}_{n,k}, \, \mathcal{U} \Big\}$

and $\mathbf{V} = \left\{ V_{n,k},\, \widetilde{\mathbf{V}}_{n,k},\, \mathbf{V}_{n,k},\, \mathcal{V}_{n,k},\, \mathcal{V} \right\}$ respectively, since the vectors $\mathbf{U}$ can be restored from independent vectors $X_1, \ldots, X_{2^N}$ by the same linear procedure which is used for reconstruction of the vectors $\mathbf{V}$ from independent vectors $Y_1, \ldots, Y_{2^N}$. From (6.1) it follows, in particular, that the mathematical expectations of all vectors $\mathbf{U}$ and $\mathbf{V}$ are equal to zero.

Remark 6.4. To define the operators $\mathbb{A}_{n,k}$, one would indicate a concrete way of constructing the coordinates of the vectors $\mathbf{U}^*_{n,k} = \mathbb{A}_{n,k}\, \widetilde{\mathbf{U}}^*_{n,k}$. Instead of this, according to Lemma 5.1 and in order to avoid the repetition of almost identical formulas (with asterisks and without them), we defined the operators $\mathbb{A}_{n,k}$ by the coordinates $\mathbb{P}_j\, \mathbf{U}_{n,k}$, $j = 1, \ldots, 2\,d$, of the vectors $\mathbf{U}_{n,k}$ (compare (5.3)–(5.5), (5.10) and (5.24)). By (5.10) and (6.4), we have

$$\mathbf{U}_{n,k} = \mathbb{A}_{n,k}\, \widetilde{\mathbf{U}}_{n,k} = \big(\mathcal{U}_{n,k},\, \widetilde{\mathcal{U}}_{n,k}\big) \in \mathbb{R}^{2d} \tag{6.6}$$

and the linear operators $\mathbb{A}_{n,k}$ is completely described by (5.10), (6.3) and (6.5). Note also that, according to (5.8)–(5.10), (5.24), and (6.3)–(6.5),

$$\mathbf{V}_{n,k} = \mathbb{A}_{n,k}\, \widetilde{\mathbf{V}}_{n,k} = \big(\mathcal{V}_{n,k},\, \widetilde{\mathcal{V}}_{n,k}\big) \in \mathbb{R}^{2d}, \tag{6.7}$$

where

$$\begin{gathered}
\widetilde{\mathbf{V}}_{n,k} = \big(V_{n-1,2k-1},\, V_{n-1,2k}\big), \\
\mathcal{V}_{n,k} = \mathbb{J}_{n,k}\, V_{n,k} = \overline{\mathbb{P}}_d\, \mathbf{V}_{n,k}, \qquad \widetilde{\mathcal{V}}_{n,k} = \widetilde{\mathbb{J}}_{n,k}\, \widetilde{V}_{n,k}, \\
\mathbb{P}_j\, \mathbf{V}_{n,k} = \mathbb{P}_{j-d}\, \widetilde{\mathcal{V}}_{n,k}, \qquad j = d+1, \ldots, 2\,d, \\
\widetilde{V}_{n,k} = 2\, \mathbb{D}_{n-1,2k}\, \mathbb{D}^{-1}_{n,k}\, V_{n-1,2k-1} - 2\, \mathbb{D}_{n-1,2k-1}\, \mathbb{D}^{-1}_{n,k}\, V_{n-1,2k}.
\end{gathered} \tag{6.8}$$

Remark 6.5. We could take $\mathbb{A}_{n,k} = \mathbb{A}^{\pm}$, where $\mathbb{A}^{\pm} : \mathbb{R}^{2d} \mapsto \mathbb{R}^{2d}$ are defined as follows: for $x = (x_1, x_2, \ldots, x_{2d}) \in \mathbb{R}^{2d}$,

$$\begin{aligned}
\mathbb{P}_j\, \mathbb{A}^{\pm}\, x &= x_j + x_{d+j}, \qquad j = 1, \ldots, d, \\
\mathbb{P}_j\, \mathbb{A}^{\pm}\, x &= x_j - x_{d+j}, \qquad j = d+1, \ldots, 2\,d.
\end{aligned} \tag{6.9}$$

However we have to consider here, generally speaking, other operators, since Lemmas 3.3 and 3.5 can be applied to conditional distributions described in Section 5 only when the operators $\mathbb{A}_{n,k}$ possess some additional properties. Namely, it is necessary that $\mathcal{U}^{(j)}$ and $\mathbf{U}^{(j)}_{n,k}$ from (5.15) and (5.20) do not correlate with $\mathcal{U}^{(1)}, \ldots, \mathcal{U}^{(j-1)}$ and $\mathbf{U}^{(1)}_{n,k}, \ldots, \mathbf{U}^{(j-1)}_{n,k}$ respectively (see (6.14) and (6.29) below). Note that if the vectors $\xi_1, \ldots, \xi_{2^N}$; $Y_1, \ldots, Y_{2^N}$ have covariance operators $\mathbb{I}_d$, the conditions just mentioned are satisfied for the operators $\mathbb{A}_{n,k} = \mathbb{A}^{\pm}$. It is not surprising, that in this case these operators was used earlier (see KMT (1975–76), Einmahl (1989), Zaitsev (1995)). It is clear that then $\widetilde{U}_{n,k} = U^-_{n,k}$ (cf. (5.32), (6.5)).

Remark 6.6. According to (5.17), (5.18), (5.21) and (6.6), under Condition **B** we have

$$\overline{\mathbb{P}}_d \mathbf{U}_{N,1} = \mathcal{U} \qquad \text{and} \qquad \mathbf{U}^{(j)}_{N,1} = \mathcal{U}^{(j)}, \quad \mathbf{U}^{j}_{N,1} = \mathcal{U}^{j}, \qquad j = 1, \dots, d. \tag{6.10}$$

Consider some additional properies of the operators $\mathbb{J}_{n,k}$, $\widetilde{\mathbb{J}}_{n,k}$ and $\mathbb{A}_{n,k}$ which follow from Condition **B**.

It is well known that if $\mathbb{D} = \operatorname{cov} \xi$ for some random vector ξ and $\mathbb{A} : \mathbb{R}^d \to \mathbb{R}^m$ is some linear operator, then

$$\operatorname{cov} \mathbb{A}\xi = \mathbb{A}\mathbb{D}\,\mathbb{A}^*, \tag{6.11}$$

where $\mathbb{A}^* : \mathbb{R}^m \to \mathbb{R}^d$ is the adjoint operator for $\mathbb{A}$. Indeed, for all $u, v \in \mathbb{R}^m$ we have $\mathbf{E}\left\langle \mathbb{A}\xi,\, u\right\rangle \left\langle \mathbb{A}\xi,\, v\right\rangle = \mathbf{E}\left\langle \xi,\, \mathbb{A}^* u\right\rangle \left\langle \xi,\, \mathbb{A}^* v\right\rangle = \left\langle \mathbb{D}\mathbb{A}^* u,\, \mathbb{A}^* v\right\rangle = \left\langle \mathbb{A}\mathbb{D}\,\mathbb{A}^* u,\, v\right\rangle$.

From (5.19), (6.3), (6.5) and (6.11) it is easy to deduce that

$$\operatorname{cov} \mathcal{U}_{n,k} = \operatorname{cov}\big(\mathbb{J}_{n,k}\, U_{n,k}\big) = \operatorname{cov} \widetilde{\mathcal{U}}_{n,k} = \operatorname{cov}\big(\widetilde{\mathbb{J}}_{n,k}\, \widetilde{U}_{n,k}\big) = 2^n\, \mathbb{I}_d, \tag{6.12}$$

for $1 \leqslant k \leqslant 2^{N-n}$, $1 \leqslant n \leqslant N$. The equality (6.12) implies, in particular, that the coordinates $\mathcal{U}^{(1)}, \mathcal{U}^{(2)}, \dots, \mathcal{U}^{(d)}$ of the vector $\mathcal{U} = \mathbb{J}_{N,1} U_{N,1}$ are non-correlated and

$$\operatorname{Var} \mathcal{U}^{(1)} = \operatorname{Var} \mathcal{U}^{(2)} = \cdots = \operatorname{Var} \mathcal{U}^{(d)} = 2^N. \tag{6.13}$$

Moreover,

$$\operatorname{cov}\, \mathcal{U}^j = 2^N\, \mathbb{I}_j, \qquad j = 1, \dots, d, \tag{6.14}$$

see (5.17), (5.18) and (6.12).

Let Ξ be an arbitrary subset of the set $\{1, 2, \dots, 2^N\}$. Write

$$U(\Xi) = \sum_{k \in \Xi} X_k, \qquad \mathbb{D}(\Xi) = \sum_{k \in \Xi} \mathbb{B}_k = \operatorname{cov} U(\Xi). \tag{6.15}$$

The relations (6.2) and (6.15) imply that

$$|\Xi|\, C_1^2\, \|u\|^2 \leqslant \left\langle \mathbb{D}(\Xi)\, u, u\right\rangle \leqslant |\Xi|\, C_2^2\, \|u\|^2, \qquad \text{for all} \quad u \in \mathbb{R}^d, \tag{6.16}$$

where $|\Xi|$ is the number of elements of the set Ξ. In particular, applying (6.16) to $u = e_j$, we get

$$|\Xi|\, C_1^2 \leqslant \left\langle \mathbb{D}(\Xi)\, e_j, e_j\right\rangle = \mathbf{E}\left\langle U(\Xi), e_j\right\rangle^2 \leqslant |\Xi|\, C_2^2, \qquad j = 1, \dots, d. \tag{6.17}$$

According to (5.24), (6.3) and (6.15), we have

$$U_{n,k} = U(\Xi), \quad \mathbb{D}_{n,k} = \mathbb{D}(\Xi) \qquad \text{for} \quad \Xi = \big\{(k-1)\cdot 2^n + 1, \dots, k \cdot 2^n\big\}. \tag{6.18}$$

The relations (6.15), (6.16) and (6.18) imply that

$$2^n C_1^2 \, \|u\|^2 \leqslant \big\langle \mathbb{D}_{n,k} u, u \big\rangle = \big\| \mathbb{D}_{n,k}^{1/2} u \big\|^2 \leqslant 2^n C_2^2 \, \|u\|^2, \tag{6.19}$$

for all $u \in \mathbb{R}^d$. Putting $u = \mathbb{D}_{n,k}^{-1/2} v$, we see that

$$2^{-n} C_2^{-2} \, \|v\|^2 \leqslant \big\langle \mathbb{D}_{n,k}^{-1} v, \, v \big\rangle = \big\| \mathbb{D}_{n,k}^{-1/2} v \big\|^2 \leqslant 2^{-n} C_1^{-2} \, \|v\|^2, \tag{6.20}$$

for all $v \in \mathbb{R}^d$. Therefore,

$$\big\| \mathbb{D}_{n,k}^{1/2} \big\|^2 = \big\| \mathbb{D}_{n,k} \big\| \leqslant 2^n C_2^2 \quad \text{and} \quad \big\| \mathbb{D}_{n,k}^{-1/2} \big\|^2 = \big\| \mathbb{D}_{n,k}^{-1} \big\| \leqslant 2^{-n} C_1^{-2}. \tag{6.21}$$

From (6.3), (6.19) and (6.20) it follows that

$$\begin{array}{c} C_2^{-2} \, \|v\|^2 \leqslant \big\| \mathbb{J}_{n,k} v \big\|^2 \leqslant C_1^{-2} \, \|v\|^2, \\ C_1^2 \, \|v\|^2 \leqslant \big\| \mathbb{J}_{n,k}^{-1} v \big\|^2 \leqslant C_2^2 \, \|v\|^2, \end{array} \qquad v \in \mathbb{R}^d, \tag{6.22}$$

and, hence,

$$\|\mathbb{J}_{n,k}\| \leqslant C_1^{-1} \quad \text{and} \quad \|\mathbb{J}_{n,k}^{-1}\| \leqslant C_2. \tag{6.23}$$

Furthermore,

$$\operatorname{cov} \mathbf{U}_{n,k} = 2^n \, \mathbb{I}_{2d}. \tag{6.24}$$

By virtue of (6.6) and (6.12), in order to verify (6.24) it suffices to show that the coordinates of $\mathcal{U}_{n,k}$ do not correlate with those of $\widetilde{\mathcal{U}}_{n,k}$. Indeed, according to Remark 6.3, these coordinates have zero means. Moreover, using (5.10), (5.24), (6.3), (6.5), the symmetricity of the operators $\mathbb{J}_{n,k}$, $\widetilde{\mathbb{J}}_{n,k}$, $\mathbb{D}_{n,k}$, $\mathbb{D}_{n-1,2k-1}$, $\mathbb{D}_{n-1,2k}$ and the fact that $U_{n-1,2k}$ is independent of $U_{n-1,2k-1}$, we have, for all $u, v \in \mathbb{R}^d$:

$$\begin{aligned} \frac{1}{2} \, \mathbf{E} \big\langle \widetilde{\mathcal{U}}_{n,k}, \, u \big\rangle \big\langle \mathcal{U}_{n,k}, \, v \big\rangle &= \frac{1}{2} \, \mathbf{E} \big\langle \widetilde{U}_{n,k}, \, \widetilde{\mathbb{J}}_{n,k} \, u \big\rangle \big\langle U_{n,k}, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &= \mathbf{E} \big\langle \mathbb{D}_{n-1,2k} \, \mathbb{D}_{n,k}^{-1} \, U_{n-1,2k-1}, \, \widetilde{\mathbb{J}}_{n,k} \, u \big\rangle \big\langle U_{n-1,2k-1}, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &\quad - \mathbf{E} \big\langle \mathbb{D}_{n-1,2k-1} \, \mathbb{D}_{n,k}^{-1} \, U_{n-1,2k}, \, \widetilde{\mathbb{J}}_{n,k} \, u \big\rangle \big\langle U_{n-1,2k}, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &= \mathbf{E} \big\langle U_{n-1,2k-1}, \, \mathbb{D}_{n,k}^{-1} \, \mathbb{D}_{n-1,2k} \, \widetilde{\mathbb{J}}_{n,k} \, u \big\rangle \big\langle U_{n-1,2k-1}, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &\quad - \mathbf{E} \big\langle U_{n-1,2k}, \, \mathbb{D}_{n,k}^{-1} \, \mathbb{D}_{n-1,2k-1} \, \widetilde{\mathbb{J}}_{n,k} \, u \big\rangle \big\langle U_{n-1,2k}, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &= \mathbf{E} \big\langle \mathbb{D}_{n-1,2k-1} \, \mathbb{D}_{n,k}^{-1} \, \mathbb{D}_{n-1,2k} \, \widetilde{\mathbb{J}}_{n,k} \, u, \, \mathbb{J}_{n,k} \, v \big\rangle \\ &\quad - \mathbf{E} \big\langle \mathbb{D}_{n-1,2k} \, \mathbb{D}_{n,k}^{-1} \, \mathbb{D}_{n-1,2k-1} \, \widetilde{\mathbb{J}}_{n,k} \, u, \, \mathbb{J}_{n,k} \, v \big\rangle = 0, \end{aligned}$$

since

$$\mathbb{D}_{n-1,2k}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k-1} = \mathbb{D}_{n-1,2k-1}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k}. \tag{6.25}$$

The relation (6.25) immediately follows from the equality

$$\mathbb{D}_{n,k} = \mathbb{D}_{n-1,2k-1} + \mathbb{D}_{n-1,2k} \tag{6.26}$$

implying that

$$\begin{aligned}\Big(\mathbb{D}_{n-1,2k}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k-1}\Big)^{-1} &= \mathbb{D}_{n-1,2k-1}^{-1}\,\mathbb{D}_{n,k}\,\mathbb{D}_{n-1,2k}^{-1} = \mathbb{D}_{n-1,2k-1}^{-1} + \mathbb{D}_{n-1,2k}^{-1}\\ &= \mathbb{D}_{n-1,2k}^{-1}\,\mathbb{D}_{n,k}\,\mathbb{D}_{n-1,2k-1}^{-1}\\ &= \Big(\mathbb{D}_{n-1,2k-1}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k}\Big)^{-1}.\end{aligned} \tag{6.27}$$

The relations (6.6) and (6.24) automatically guarantee that the operator $\mathbb{A}_{n,k}$ is non-degenerate.

Thus, the coordinates $\mathbf{U}_{n,k}^{(1)}, \mathbf{U}_{n,k}^{(2)}, \dots, \mathbf{U}_{n,k}^{(2d)}$ of the vector $\mathbf{U}_{n,k}$ are non-correlated and

$$\operatorname{Var}\mathbf{U}_{n,k}^{(1)} = \operatorname{Var}\mathbf{U}_{n,k}^{(2)} = \dots = \operatorname{Var}\mathbf{U}_{n,k}^{(2d)} = 2^n. \tag{6.28}$$

Moreover (see (5.21) and (6.24)),

$$\operatorname{cov}\mathbf{U}_{n,k}^{j} = 2^n\,\mathbb{I}_j, \qquad j = 1,\dots,2\,d. \tag{6.29}$$

By (6.5), (6.11), (6.25), (6.26) and by the independence of $U_{n-1,2k-1}$ and $U_{n-1,2k}$,

$$\begin{aligned}\widetilde{\mathbb{D}}_{n,k} &= 4\,\mathbb{D}_{n-1,2k}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k-1}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k}\\ &\quad + 4\,\mathbb{D}_{n-1,2k-1}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k-1}\\ &= 4\,\mathbb{D}_{n-1,2k}\,\mathbb{D}_{n,k}^{-1}\,\mathbb{D}_{n-1,2k-1}.\end{aligned} \tag{6.30}$$

and, hence,

$$\widetilde{\mathbb{D}}_{n,k}^{-1} = \frac{1}{4}\,\mathbb{D}_{n-1,2k-1}^{-1}\,\mathbb{D}_{n,k}\,\mathbb{D}_{n-1,2k}^{-1}. \tag{6.31}$$

The relations (6.5), (6.21), (6.30) and (6.31) imply that

$$2^n\big\|\widetilde{\mathbb{J}}_{n,k}^{-1}\big\|^2 = \big\|\widetilde{\mathbb{D}}_{n,k}^{1/2}\big\|^2 = \big\|\widetilde{\mathbb{D}}_{n,k}\big\| \leqslant 4\,\big\|\mathbb{D}_{n-1,2k}\big\|\big\|\mathbb{D}_{n,k}^{-1}\big\|\big\|\mathbb{D}_{n-1,2k-1}\big\| \leqslant 2^n\,C_2^4\,C_1^{-2} \tag{6.32}$$

and, similarly,

$$2^{-n}\big\|\widetilde{\mathbb{J}}_{n,k}\big\|^2 = \big\|\widetilde{\mathbb{D}}_{n,k}^{-1/2}\big\|^2 = \big\|\widetilde{\mathbb{D}}_{n,k}^{-1}\big\| \leqslant 2^{-n}\,C_2^2\,C_1^{-4}. \tag{6.33}$$

Note that the bounds (6.32) and (6.33) can be refined with the help of (6.20), (6.27), (6.30) and (6.31).

Define $\mathbb{M}_{n,k} = \operatorname{cov} \widetilde{\mathbf{U}}_{n,k}$ and write an arbitrary vector $u = \big(u_1, \dots, u_{2d}\big) \in \mathbb{R}^{2d}$ in the form $u = \big(u^{(1)}, u^{(2)}\big)$, where $u^{(1)} = \big(u_1, \dots, u_d\big)$, $u^{(2)} = \big(u_{d+1}, \dots, u_{2d}\big) \in \mathbb{R}^d$. Using (5.24), (6.3), we see that

$$\big\langle \mathbb{M}_{n,k}\, u,\, u \big\rangle = \big\langle \mathbb{D}_{n-1,2k-1}\, u^{(1)},\, u^{(1)} \big\rangle + \big\langle \mathbb{D}_{n-1,2k}\, u^{(2)},\, u^{(2)} \big\rangle. \tag{6.34}$$

Obviously, $\|u\|^2 = \big\| u^{(1)} \big\|^2 + \big\| u^{(2)} \big\|^2$. Therefore, from (6.19) and (6.34) it follows that

$$2^{n-1}\, C_1^2\, \|u\|^2 \leqslant \big\langle \mathbb{M}_{n,k}\, u, u \big\rangle = \big\| \mathbb{M}_{n,k}^{1/2}\, u \big\|^2 \leqslant 2^{n-1}\, C_2^2\, \|u\|^2 . \tag{6.35}$$

Furthermore, according to (6.6) and (6.24), for all $v \in \mathbb{R}^{2d}$ we have

$$2^n\, \|v\|^2 = \mathbf{E}\, \big\langle \mathbb{A}_{n,k}\, \widetilde{\mathbf{U}}_{n,k},\, v \big\rangle^2 = \mathbf{E}\, \big\langle \widetilde{\mathbf{U}}_{n,k},\, \mathbb{A}_{n,k}^*\, v \big\rangle^2 = \big\langle \mathbb{M}_{n,k}\, \mathbb{A}_{n,k}^*\, v,\, \mathbb{A}_{n,k}^*\, v \big\rangle. \tag{6.36}$$

From (6.35) and (6.36) it follows that

$$\begin{aligned} c_{15}^{-2}\, \|v\|^2 &\leqslant \big\| \mathbb{A}_{n,k}^*\, v \big\|^2 \leqslant c_{16}^{-2}\, \|v\|^2 , \\ c_{16}^{2}\, \|v\|^2 &\leqslant \big\| (\mathbb{A}_{n,k}^*)^{-1}\, v \big\|^2 \leqslant c_{15}^{2}\, \|v\|^2 , \end{aligned} \qquad v \in \mathbb{R}^{2d}, \tag{6.37}$$

with $c_{16} = C_1/\sqrt{2}$, $c_{15} = C_2/\sqrt{2}$. By the well-known properties of adjoint operators,

$$\big\| \mathbb{A}_{n,k}^* \big\| = \sup_{\|u\|=1,\ \|v\|=1} \big\langle \mathbb{A}_{n,k}\, u,\, v \big\rangle = \big\| \mathbb{A}_{n,k} \big\| \quad \text{and} \quad (\mathbb{A}_{n,k}^*)^{-1} = (\mathbb{A}_{n,k}^{-1})^*. \tag{6.38}$$

Using (6.38), we see that (6.37) implies

$$\big\| \mathbb{A}_{n,k} \big\| \leqslant c_{16}^{-1} \quad \text{and} \quad \big\| \mathbb{A}_{n,k}^{-1} \big\| \leqslant c_{15} \tag{6.39}$$

or, equivalently,

$$\begin{aligned} c_{15}^{-1}\, \|v\| &\leqslant \big\| \mathbb{A}_{n,k}\, v \big\| \leqslant c_{16}^{-1}\, \|v\| , \\ c_{16}\, \|v\| &\leqslant \big\| \mathbb{A}_{n,k}^{-1}\, v \big\| \leqslant c_{15}\, \|v\| , \end{aligned} \qquad v \in \mathbb{R}^{2d}. \tag{6.40}$$

Lemma 6.7. *Let* $1 \leqslant m = (2\,s+1) \cdot 2^r \leqslant 2^N$, *where* s, r *are non-negative integers. Assume that the dyadic scheme satisfies Condition* **B**. *Then there exist an absolute constant* $c_{17} \geqslant 1$ *such that the following representations are valid*:

$$S_m = \mathbb{K}\, U_{N,1} + \sum_{n=r+1}^{N} \mathbb{K}_n\, \widetilde{\mathcal{U}}_{n,l_{n,m}},$$

$$U_{n,l_{n,m}} = \mathbb{L}\, U_{N,1} + \sum_{s=n+1}^{N} \mathbb{L}_s\, \widetilde{\mathcal{U}}_{s,l_{s,m}}, \qquad n = r, \dots, N, \tag{6.41}$$

where linear operators $\mathbb{K}$, $\mathbb{K}_n : \mathbb{R}^d \to \mathbb{R}^d$ *depend on* m *and satisfy*

$$\max\left\{ \|\mathbb{K}\|,\, \|\mathbb{K}_{r+1}\|,\, \dots,\, \|\mathbb{K}_N\| \right\} \leqslant c_{17}, \tag{6.42}$$

linear operators $\mathbb{L}$, $\mathbb{L}_s : \mathbb{R}^d \to \mathbb{R}^d$ *depend on* m *and* n *and satisfy*

$$\|\mathbb{L}\| \leqslant c_{17} \cdot 2^{n-N}, \qquad \|\mathbb{L}_s\| \leqslant c_{17} \cdot 2^{n-s}, \quad s = n+1, \dots, N, \tag{6.43}$$

and the integers $l_{n,m}$ *are defined by* (5.34).

Proof. Below we shall change the subscript ${}_{n,l_{n,m}}$ by ${}_{n}$, $0 \leqslant n \leqslant N$. Then the equality (5.52) may be rewritten in the form

$$S_m = \sum_{n=r}^{N} \beta_n\, U_n = \sum_{n=r+1}^{N+1} \beta_{n-1}\, U_{n-1}, \qquad \text{where} \quad |\beta_{n-1}| \leqslant \frac{3}{2}\,. \tag{6.44}$$

The relations (5.24), (6.5) and (6.26) imply that, for $1 \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$,

$$\begin{aligned} U_{n-1,2k-1} &= \frac{1}{2}\, \widetilde{U}_{n,k} + \mathbb{D}_{n-1,2k-1}\, \mathbb{D}_{n,k}^{-1}\, U_{n,k}, \\ U_{n-1,2k} &= -\frac{1}{2}\, \widetilde{U}_{n,k} + \mathbb{D}_{n-1,2k}\, \mathbb{D}_{n,k}^{-1}\, U_{n,k}. \end{aligned} \tag{6.45}$$

Thus, taking into account Remark 5.5, we see that, by (6.45),

$$U_{n-1} = \frac{\varkappa_n}{2}\, \widetilde{U}_n + \mathbb{D}_{n-1}\, \mathbb{D}_n^{-1}\, U_n, \qquad n = r+1, \dots, N, \tag{6.46}$$

where the factors $\varkappa_n = \varkappa_n(m)$ can be equal either to 1 or to -1. Substituting sequentially the equality (6.46) in itself, we get

$$\begin{aligned} U_{n-1} &= \frac{\varkappa_n}{2}\, \widetilde{U}_n + \frac{\varkappa_{n+1}}{2}\, \mathbb{D}_{n-1}\, \mathbb{D}_n^{-1}\, \widetilde{U}_{n+1} + \mathbb{D}_{n-1}\, \mathbb{D}_{n+1}^{-1}\, U_{n+1} = \cdots \\ &= \sum_{s=n}^{N} \frac{\varkappa_s}{2}\, \mathbb{D}_{n-1}\, \mathbb{D}_{s-1}^{-1}\, \widetilde{U}_s + \mathbb{D}_{n-1}\, \mathbb{D}_N^{-1}\, U_N. \end{aligned} \tag{6.47}$$

Thus, we see that

$$U_n = \mathbb{L}\, U_N + \sum_{s=n+1}^{N} \mathbb{G}_s\, \widetilde{U}_s, \qquad n = r, \dots, N, \tag{6.48}$$

where the linear operators $\mathbb{L}$ and $\mathbb{G}_s$ (depending on m and n) are defined by

$$\mathbb{L} = \mathbb{D}_n\,\mathbb{D}_N^{-1}, \qquad \mathbb{G}_s = \frac{\varkappa_s}{2}\,\mathbb{D}_n\,\mathbb{D}_{s-1}^{-1}, \qquad s = n+1,\dots,N, \tag{6.49}$$

The equalities (6.44) and (6.47) imply now that

$$\begin{aligned} S_m &= \sum_{n=r+1}^{N+1} \beta_{n-1}\,\mathbb{D}_{n-1}\,\mathbb{D}_N^{-1}\,U_N + \sum_{n=r+1}^{N}\sum_{s=n}^{N} \frac{\beta_{n-1}\,\varkappa_s}{2}\,\mathbb{D}_{n-1}\,\mathbb{D}_{s-1}^{-1}\,\widetilde{U}_s \\ &= \mathbb{K}\,U_N + \sum_{s=r+1}^{N} \mathbb{C}_s\,\widetilde{U}_s, \end{aligned} \tag{6.50}$$

where the linear operators $\mathbb{K}$ and $\mathbb{C}_s$ (depending on m) are defined by

$$\mathbb{K} = \sum_{n=r+1}^{N+1} \beta_{n-1}\,\mathbb{D}_{n-1}\,\mathbb{D}_N^{-1}, \qquad \mathbb{C}_s = \sum_{n=r+1}^{s} \frac{\beta_{n-1}\,\varkappa_s}{2}\,\mathbb{D}_{n-1}\,\mathbb{D}_{s-1}^{-1} \tag{6.51}$$

and the coefficients β_{n-1} and $\varkappa_s$ satisfy

$$|\beta_{n-1}| \leqslant 3/2, \quad |\varkappa_s| = 1. \tag{6.52}$$

By (6.21), (6.49), (6.51) and (6.52), we have the bounds

$$\|\mathbb{L}\| \leqslant c\cdot 2^{n-N}, \qquad \|\mathbb{G}_s\| \leqslant c\cdot 2^{n-s}, \tag{6.53}$$

$$\|\mathbb{K}\| \leqslant c \sum_{n=r+1}^{N+1} 2^{n-N} \leqslant c \quad \text{and} \quad \|\mathbb{C}_s\| \leqslant c \sum_{n=r+1}^{s} 2^{n-s} \leqslant c. \tag{6.54}$$

It remains to note that $U_N = U_{N,1}$, $\widetilde{U}_s = \widetilde{\mathbb{J}}_s^{-1}\widetilde{\mathcal{U}}_s$ (see (6.5)) and define $\mathbb{K}_s = \mathbb{C}_s\,\widetilde{\mathbb{J}}_s^{-1}$ and $\mathbb{L}_s = \mathbb{G}_s\,\widetilde{\mathbb{J}}_s^{-1}$. Then the assertion of the lemma follows from (6.32), (6.48), (6.50), (6.53) and (6.54) provided that c_{17} is chosen to be sufficiently large. □

Remark 6.8. Since the vectors T_m, $V_{n,k}$ and $\mathcal{V}_{n,k}$ can be reconstructed from the vectors $Y_1,\dots,Y_{2^N}$ by means of the same linear procedure which connects the vectors S_m, $U_{n,k}$ and $\mathcal{U}_{n,k}$ with the vectors $X_1,\dots,X_{2^N}$, we see that in the conditions of Lemma 6.7

$$T_m = \mathbb{K}\,V_{N,1} + \sum_{n=r+1}^{N} \mathbb{K}_n\,\widetilde{\mathcal{V}}_{n,l_{n,m}}, \tag{6.55}$$

$$V_{n,l_{n,m}} = \mathbb{L}\,V_{N,1} + \sum_{s=n+1}^{N} \mathbb{L}_s\,\widetilde{\mathcal{V}}_{s,l_{s,m}}, \qquad n = r,\dots,N. \tag{6.56}$$

7. Dyadic scheme. Conditions $\mathbf{A}(\tau)$, $\mathbf{A}^*(\tau)$, $\mathbf{C}_\Psi(p)$, $\mathbf{D}_\Phi$, $\mathbf{D}_\Psi$ and their consequences

In order to introduce further conditions we need some notation.

Notation 7.1. Consider a one-dimensional symmetric distribution $H = \mathcal{L}(\eta)$ concentrated on the interval $[-1,1]$ and having a three times differentiable density $w(x)$. It is evident that $\mathbf{E}\,\eta = 0$ and such a distribution can be easily constructed. Below we shall assume that the distribution H with properties indicated above is fixed. For the sake of definiteness, we take

$$w(x) = \begin{cases} \dfrac{15015}{256}\, x^4 \left(1 - x^2\right)^4, & \text{if } -1 \leqslant x \leqslant 1, \\ 0, & \text{otherwise.} \end{cases}$$

Then $\mathbf{E}\,\eta^2 = {}^1\!/_3$. Introduce the corresponding conjugate distributions $H_h = \overline{H}(h)$, $h \in \mathbb{R}^1$ (see (2.1)).

Lemma 7.2. *There exist absolute positive constants $c_{18}, \dots, c_{21}$ such that the following relations are true*:

$$\sup_{|h|\leqslant c_{18}} \sup_{|t|\geqslant c_{19}} \left|\widehat{H}_h(t)\right| \leqslant e^{-c_{20}}, \tag{7.1}$$

$$\sup_{|h|\leqslant c_{18}} \left|\widehat{H}_h(t)\right| \leqslant \exp\Big(-\frac{1}{8}\,t^2\Big) \qquad \textit{for} \quad |t| \leqslant c_{19}, \tag{7.2}$$

$$\sup_{|h|\leqslant c_{18}} \left|\widehat{H}_h(t)\right| \leqslant \frac{1}{(1 + c_{21}\,t^2)}, \qquad \textit{for} \quad t \in \mathbb{R}^1. \tag{7.3}$$

Proof. Putting $w(h,x) = e^{hx}\,w(x)$, $h, x \in \mathbb{R}^1$, and integrating by parts, we see that (for $t \neq 0$)

$$\begin{aligned} \widehat{H}_h(t) &= \left(\mathbf{E}\,e^{h\eta}\right)^{-1} \int\limits_{|x|\leqslant 1} e^{itx}\, w(h,x)\, dx \\ &= -\left(\mathbf{E}\,e^{h\eta}\right)^{-1} \int\limits_{|x|\leqslant 1} t^{-2}\, e^{itx}\, \frac{\partial^2 w(h,x)}{\partial x^2}\, dx. \end{aligned} \tag{7.4}$$

Furthermore,

$$\sup_{x\in\mathbb{R}^1} \sup_{|h|\leqslant 1} \left| \frac{\partial^2 w(h,x)}{\partial x^2} \right| \leqslant c_{22} \tag{7.5}$$

(constants here and below are absolute, since the distribution H is fixed). From (7.4), (7.5) it follows that

$$\sup_{|h|\leqslant 1} \left|\widehat{H}_h(t)\right| \leqslant 2\,c_{22}\,t^{-2} \tag{7.6}$$

(we take into account that, by the Jensen inequality, $\mathbf{E}\, e^{h\eta} \geqslant e^{\mathbf{E}\, h\eta} = 1$). The inequality (7.6) implies that, for $|t| \geqslant c_{23} = 2\, c_{22}^{1/2}$,

$$\sup_{|h|\leqslant 1} \big|\widehat{H}_h(t)\big| \leqslant \Big(1 + \frac{t^2}{4\, c_{22}}\Big)^{-1} \tag{7.7}$$

and

$$\sup_{|h|\leqslant 1} \sup_{|t|\geqslant c_{23}} \big|\widehat{H}_h(t)\big| \leqslant 1/2. \tag{7.8}$$

Obviously, $H \in \mathcal{A}_1(c)$. Using Lemma 2.2(b) and the relations (2.3) and (2.4), we see that the relation (7.2) is valid, if $c_{18} \leqslant 1$ and $c_{19} \leqslant c_{23}$ are sufficiently small.

Since the distributions H_h are absolutely continuous, the relation $\big|\widehat{H}_h(t)\big| = 1$ can be valid for $t = 0$ only. Furthermore, the function $\big|\widehat{H}_h(t)\big|$ considered as a function of two variables h and t is continuous for all $h, t \in \mathbb{R}^1$. Therefore,

$$\sup_{|h|\leqslant c_{18}} \sup_{c_{19}\leqslant|t|\leqslant c_{23}} \big|\widehat{H}_h(t)\big| \leqslant c < 1. \tag{7.9}$$

The inequalities (7.8), (7.9) imply (7.1).

It is easy to see that

$$\max\Big\{ \exp\Big(- \frac{1}{8}\, t^2\Big),\, e^{-c_{20}} \Big\} \leqslant \frac{1}{(1 + c_{24}\, t^2)}, \qquad \text{for} \qquad |t| \leqslant c_{23}, \tag{7.10}$$

if c_{24} is sufficiently small. The inequality (7.3) follows now from (7.1), (7.2), (7.7) and (7.10) with $c_{21} = \min\big\{ c_{24},\, (4 c_{22})^{-1} \big\}$. Thus, Lemma 7.2 is proved. $\square$

Below we shall need the following useful fact.

Lemma 7.3. *Assume that* $0 < \alpha < 1$ *and a random variable* μ *has values* 0 *with probability* $1-\alpha^3$ *and* α^{-1} *with probability* α^3. *Let* $X, K \in \mathbb{R}^d$ *be random vectors with* $\mathbf{E}\, X = \mathbf{E}\, K = 0$, $\operatorname{cov} X = \mathbb{B}$, $\operatorname{cov} K = (1-\alpha)\, \mathbb{B}$, $\mathbf{E}\, \|X\|^3 < \infty$, $\mathbf{E}\, \|K\|^3 < \infty$. *Suppose also that* X, K, μ *are mutually independent and the distribution of* K *is symmetric. Then the random vector* $W = \mu X + K$ *has mean zero, covariance operator* $\mathbb{B}$ *and, for any* $u, v, w \in \mathbb{R}^d$, *we have*

$$\mathbf{E}\, \big\langle W, u\big\rangle \big\langle W, v\big\rangle \big\langle W, w\big\rangle = \mathbf{E}\, \big\langle X, u\big\rangle \big\langle X, v\big\rangle \big\langle X, w\big\rangle$$

(*in other words,* W *and* X *have the same moments of the first three orders*).

The proof of Lemma 7.3 is elementary and, therefore, is omitted. One should note that Lemma 7.3 is a simplified (and modified) version of (due to Yurinskii (1978)) Lemma 4 from Sazonov (1981, p. 85)[1].

---

[1]The author is grateful to B. A. Zalesskii who attracted the author's attention to the book Sazonov (1981).

Notation 7.4. Let a random vector Z have independent coordinates Z_j with $\mathcal{L}(Z_j) = H$, $j = 1, \dots, d$. Assume that the conditions of Lemma 7.3 are satisfied with $\alpha = {}^2\!/_3$, $\mathcal{L}(K) = \mathcal{L}(\mathbb{B}^{1/2} Z)$. Thus, we have

$$\operatorname{cov} Z = \frac{1}{3}\, \mathbb{I}_d, \qquad \operatorname{cov} K = \frac{1}{3}\, \mathbb{B}, \tag{7.11}$$

$$\mathbf{P}\big\{\, \|K\| \leqslant \|\mathbb{B}^{1/2}\|\, \sqrt{d} \,\big\} = 1, \tag{7.12}$$

$$\mathbf{P}\big\{\, \mu = {}^3\!/_2 \,\big\} = \frac{8}{27}\,, \qquad \mathbf{P}\big\{\, \mu = 0 \,\big\} = \frac{19}{27}\,. \tag{7.13}$$

In this case we shall use below the notation

$$\Psi\big(\mathcal{L}(X)\big) \overset{\text{def}}{=} \mathcal{L}(W) = \mathcal{L}(\mu X + K). \tag{7.14}$$

According to Lemma 7.3, the distributions $\mathcal{L}(X)$ and $\Psi\big(\mathcal{L}(X)\big)$ have coinciding moments of the first three orders. However, the distribution $\Psi\big(\mathcal{L}(X)\big)$ is "better" than $\mathcal{L}(X)$ because it is sufficiently smooth.

Lemma 7.5. *Let a random vector $X \in \mathbb{R}^d$ satisfy the conditions $\mathbf{E}\, X = 0$, $\mathcal{L}(X) \in \mathcal{A}_d(\tau)$, $\tau \geqslant 1$. Fix absolute positive constants C_1, C_2 and suppose that the operator $\mathbb{B} = \operatorname{cov} X$ for all $u \in \mathbb{R}^d$ satisfies the condition*

$$C_1^2\, \|u\|^2 \leqslant \langle \mathbb{B} u, u\rangle = \big\|\mathbb{B}^{1/2} u\big\|^2 \leqslant C_2^2\, \|u\|^2\,. \tag{7.15}$$

Assume that the distribution $\Psi\big(\mathcal{L}(X)\big)$ is defined by (7.14) *and $R_h = \overline{\Psi\big(\mathcal{L}(X)\big)}(h)$, $h \in \mathbb{R}^d$. Then there exist absolute positive constants $c_{25}, c_{26}, c_{27}, c_{28}$ such that*

$$\mathcal{L}(K) = \mathcal{L}(\mathbb{B}^{1/2} Z) \in \mathcal{A}_d(c_{25}), \qquad \mathcal{L}(W) = \Psi\big(\mathcal{L}(X)\big) \in \mathcal{A}_d(c_{25}\, \tau), \qquad c_{25} \geqslant 1 \tag{7.16}$$

and, for all $t \in \mathbb{R}^d$,

$$\sup_{\|h\|\tau \leqslant c_{26}} \big|\widehat{R}_h(t)\big| \leqslant \min\Big\{\, \max\big\{\, \exp\big(-c_{27}\, \|t\|^2\,\big),\, e^{-c_{20}} \big\},\, \frac{1}{(1 + c_{28}\, \|t\|^2)} \Big\}. \tag{7.17}$$

Proof. Using Lemma 2.3(b), we see that from the conditions of the lemma it follows that $\mathcal{L}(X) \in \mathcal{B}_d(c\tau)$ (recall that $\mathbf{E}\, X = 0$). Using (2.7), (7.13), we see that, for all $u, v \in \mathbb{R}^d$ and for all integer $m \geqslant 3$,

$$\begin{aligned} \Big|\mathbf{E}\, \big\langle \mu X, v\big\rangle^2 \big\langle \mu X, u\big\rangle^{m-2}\Big| &= \mathbf{E}\, \mu^m \Big|\mathbf{E}\, \big\langle X, v\big\rangle^2 \big\langle X, u\big\rangle^{m-2}\Big| \leqslant \frac{1}{2}\, m!\, \big(c\tau\, \|u\|\big)^{m-2} \mathbf{E}\, \big\langle X, v\big\rangle^2 \\ &\leqslant \frac{1}{2}\, m!\, \big(c\tau\, \|u\|\,\big)^{m-2}\, \mathbf{E}\, \big\langle \mu X, v\big\rangle^2. \end{aligned} \tag{7.18}$$

According to (7.18), (2.7), $\mathcal{L}(\mu X) \in \mathcal{B}_d(c\tau)$ (evidently, $\mathbf{E}\,\mu X = \mathbf{E}\,\mu\,\mathbf{E}\,X = 0$). By Lemma 2.3(a), this implies that

$$\mathcal{L}(\mu X) \in \mathcal{A}_d(c\tau). \tag{7.19}$$

Since $\mathbf{P}\big\{|Z_j| \leqslant 1\big\} = 1$, $j = 1, \ldots, d$, the relations (7.14), (7.15), (7.19) together with Lemmas 2.3 and 2.6 imply (7.16) (we used the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution and their monotonicity with respect to τ). It is clear that (7.15) implies that the operators $\mathbb{B}$ and $\mathbb{B}^{1/2}$ are non-degenerate and the distributions $\mathcal{L}(K)$ and, consequently, $\Psi\big(\mathcal{L}(X)\big)$ are absolutely continuous. Write $P_h = \overline{\mathcal{L}(Z)}(h)$, $h \in \mathbb{R}^d$. According to (2.2), we have

$$\widehat{P}_h(t) = \prod_{j=1}^{d} \widehat{H}_{h_j}(t_j), \qquad \text{for} \quad t \in \mathbb{R}^d.$$

Therefore, by Lemma 7.2, we obtain, for $t \in \mathbb{R}^d$,

$$\sup_{\|h\| \leqslant c_{18}} \big|\widehat{P}_h(t)\big| \leqslant \max\Big\{ \exp\big(-\|t\|^2/8\big),\, e^{-c_{20}} \Big\} \tag{7.20}$$

and

$$\sup_{\|h\| \leqslant c_{18}} \big|\widehat{P}_h(t)\big|\, dt \leqslant \frac{1}{(1 + c_{21}\,\|t\|^2)}. \tag{7.21}$$

According to (2.2), (3.6), (7.14) and (7.16), we have, for $\|h\|\,\tau < 1$,

$$\big|\widehat{R}_h(t)\big| \leqslant \Big(\mathbf{E}\, e^{\langle h, K\rangle}\Big)^{-1} \big|\,\mathbf{E}\, e^{\langle h+it, K\rangle}\big| = \big|\widehat{P}_\gamma(\mathbb{B}^{1/2} t)\big| \qquad \text{where} \quad \gamma = \mathbb{B}^{1/2} h. \tag{7.22}$$

Moreover, by (7.15), (7.20)–(7.22) and since $\tau \geqslant 1$, we conclude that the relation (7.17) hold. □

Suppose that the dyadic scheme satisfies Condition **B** and p is integer, $0 \leqslant p \leqslant N$. Introduce the following conditions on the distributions $R^{(k)} = \mathcal{L}(\xi_k)$:

$\mathbf{C}_\Psi(p)$. $\mathcal{L}(\xi_l) = \Psi\big(\mathcal{L}(\widetilde{\eta}_l)\big)$ for $l = (m-1)\cdot 2^p + 1$, $m = 1, \ldots, 2^{N-p}$,
$\mathcal{L}(\xi_l) = \mathcal{L}(\widetilde{\xi}_l)$ otherwise.

By Lemma 7.3 and by Condition $\mathbf{B}_1$, ξ_k and $\widetilde{\xi}_k$ have in the case $\mathbf{C}_\Psi(p)$ the same moments of the first three orders. Condition $\mathbf{C}_\Psi(0)$ means that $\mathcal{L}(\xi_k) = \Psi\big(\mathcal{L}(\widetilde{\eta}_k)\big)$, for all $k = 1, \ldots, 2^N$. Introduce also the conditions on the distributions of the vectors $Y_1, \ldots, Y_{2^N}$:

$\mathbf{D}_\Phi$. $\mathcal{L}(Y_k) = \Phi\big(\mathcal{L}(\widetilde{\xi}_k)\big)$, $k = 1, \ldots, 2^N$;
$\mathbf{D}_\Psi$. $\mathcal{L}(Y_k) = \Psi\big(\mathcal{L}(\widetilde{\eta}_k)\big)$, $k = 1, \ldots, 2^N$.

According to Lemma 7.3 and Condition $\mathbf{B}_1$, the random vectors Y_k and $\widetilde{\xi}_k$ have in the case $\mathbf{D}_\Psi$ the same moments of the first three orders.

Introduce, finally, the conditions of belonging the distributions involved in our scheme to the classes of type $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$:

$\mathbf{A}(\tau)$. $\mathcal{L}(\widetilde{\xi}_k) \in \mathcal{A}_d(\tau)$ and $\mathcal{L}(\widetilde{\eta}_k) \in \mathcal{A}_d(\tau)$, for $k = 1, \ldots, 2^N$;
$\mathbf{A}^*(\tau)$.

$$\mathcal{L}\big(\mathbf{U}^j_{n,k}\big) \in \overline{\mathcal{A}}_j(\tau, 4) \qquad \text{for} \quad 1 \leqslant k \leqslant 2^{N-n}, \quad 1 \leqslant n \leqslant N, \quad d \leqslant j \leqslant 2\,d, \tag{7.23}$$

and

$$\mathcal{L}\big(\mathbf{U}^j_{N,1}\big) \in \overline{\mathcal{A}}_j(\tau, 4) \qquad \text{for} \quad 1 \leqslant j \leqslant 2\,d. \tag{7.24}$$

Lemma 7.6. *Assume that Conditions $\mathbf{A}(\tau)$, $\mathbf{B}$ and $\mathbf{C}_\Psi(p)$ are satisfied with $\tau \geqslant 1$ and $0 \leqslant p \leqslant N$. Then there exist an absolute positive constant c_{29} such that if $\tau^2\, d^3\, (p+1)/2^{n-p} \leqslant c_{29}$, $1 \leqslant k \leqslant 2^{N-n}$ and $1 \leqslant n \leqslant N$, then*

$$F = \mathcal{L}(\mathbf{U}^j_{n,k}) \in \mathcal{A}^*_j(\,c_{30}\,\tau,\ 4,\ 4), \qquad j = 1, \ldots, 2\,d, \tag{7.25}$$

with $c_{30} = c_{16}^{-1} \max\big\{\,c_{25},\ c_{26}^{-1},\ 1\,\big\}$.

Corollary 7.7. *Let Conditions $\mathbf{A}(\tau)$, $\mathbf{B}$ and $\mathbf{C}_\Psi(p)$ be satisfied with $0 \leqslant p \leqslant N$. Then*

$$\mathcal{L}(\mathcal{U}^j) \in \mathcal{A}^*_j(c_{30}\,\tau,\ 4,\ 4), \qquad j = 1, \ldots, d,$$

if $\tau \geqslant 1$ and $\tau^2\, d^3\, (p+1)/2^{N-p} \leqslant c_{29}$.

Proof. It suffices to note that $\mathcal{U}^j = \mathbf{U}^j_{N,1}$, see Remark 6.6. It remains to apply Lemma 7.6 with $n = N$ and $k = 1$. $\square$

Proof of Lemma 7.6. Recall that $\mathbf{U}_{n,k} = \mathbb{A}_{n,k}\,\widetilde{\mathbf{U}}_{n,k}$ (see (5.24)), where $\mathbb{A}_{n,k}$ is a linear operator defined by the relations (5.10), (6.3), (6.5) and (6.6) and satisfying (6.37), (6.39) and (6.40). Moreover, $\widetilde{\mathbf{U}}_{n,k} = \big(U_{n-1,2k-1},\, U_{n-1,2k}\big) \in \mathbb{R}^{2d}$, where $U_{n-1,2k-1}$ and $U_{n-1,2k}$ are independent elements of $\mathbb{R}^d$. Therefore, the relation $\mathcal{L}(\mathbf{U}_{n,k}) \in \mathcal{A}_{2d}(c_{30}\,\tau)$ can be easily derived from Conditions $\mathbf{A}(\tau)$ and $\mathbf{C}_\Psi(p)$ with the help of Lemmas 2.5, 2.6, 5.1 and 7.5 (see (6.39) and (7.16)) if we take into account the completeness of classes $\mathcal{A}_d(\tau)$ with respect to convolution and their monotonicity with respect to τ. It is easy to see that $\mathbf{U}^j_{n,k} = \overline{\mathbb{P}}_j\,\mathbf{U}_{n,k}$, where the projector $\overline{\mathbb{P}}_j : \mathbb{R}^{2d} \to \mathbb{R}^j$ can be considered as a linear operator with $\|\overline{\mathbb{P}}_j\| = 1$ (see (5.21)). Applying Lemma 2.5 again, we obtain the relations $\mathcal{L}(\mathbf{U}^j_{n,k}) \in \mathcal{A}_j(c_{30}\,\tau)$, $1 \leqslant j \leqslant 2\,d$.

It remains to verify that if c_{29} is small enough, then for $h \in \mathbb{R}^j$, $\|h\|\, c_{30}\,\tau < 1$, and for all $v \in \mathbb{R}^j$ the following inequalities hold:

$$(2\pi)^{-j} \int\limits_T \bigl|\widehat{F}_h(t)\bigr|\, dt \leqslant \frac{c_{30}^2\,\tau^2\, j^2}{\sigma^2\,(2\pi)^{j/2}\,(\det \mathbb{D})^{1/2}}\,, \tag{7.26}$$

$$(2\pi)^{-j} \int\limits_T \bigl|\bigl\langle t, v\bigr\rangle\, \widehat{F}_h(t)\bigr|\, dt \leqslant \frac{\bigl\langle \mathbb{D}^{-1} v, v\bigr\rangle^{1/2}}{4\,(2\pi)^{j/2}\,(\det \mathbb{D})^{1/2}}\,, \tag{7.27}$$

where F_h denotes the conjugate distribution $F_h = \overline{F}(h)$ (see (2.1), (3.2) and (3.3)),

$$T = \Bigl\{ t \in \mathbb{R}^j : 4\, \|t\|\, c_{30}\,\tau\, j \geqslant 1 \Bigr\} \tag{7.28}$$

and σ^2 is the minimal eigenvalue of the operator $\mathbb{D} = \operatorname{cov} \mathbf{U}_{n,k}^j$. Note that, according to (6.28) and (6.29), we have

$$\mathbb{D} = 2^n\, \mathbb{I}_j, \quad \sigma^2 = 2^n, \quad \bigl\langle \mathbb{D}^{-1} v, v\bigr\rangle = 2^{-n}\, \|v\|^2, \quad \det \mathbb{D} = 2^{nj}. \tag{7.29}$$

Introduce 2^{n-1} random vectors

$$\mathbf{X}_r = \bigl(X_r,\, X_{2^{n-1}+r}\bigr) \in \mathbb{R}^{2d}, \qquad r = 2^n\,(k-1)+1, \ldots, 2^{n-1}\,(2\,k-1). \tag{7.30}$$

Obviously, these vectors are independent. According to (5.23), (5.24) and (7.30),

$$\widetilde{\mathbf{U}}_{n,k} = \bigl(U_{n-1,2k-1},\, U_{n-1,2k}\bigr) = \sum_{r=2^n\,(k-1)+1}^{2^{n-1}(2k-1)} \mathbf{X}_r. \tag{7.31}$$

Write $R_h^{(s)} = \overline{\mathcal{L}(X_s)}(h)$, for $s = 1, \ldots, 2^N$, $h \in \mathbb{R}^d$, and $M_h^{(r)} \overset{\text{def}}{=} \overline{\mathcal{L}(\mathbf{X}_r)}(h)$, $Q_h^{(r)} \overset{\text{def}}{=} \overline{\mathcal{L}(\mathbb{A}_{n,k}\, \mathbf{X}_r)}(h)$, for $r = 2^n\,(k-1)+1, \ldots, 2^{n-1}\,(2\,k-1)$, $h \in \mathbb{R}^{2d}$. As usually, we consider only such h for which these distributions exist. It is easy to see that, by virtue of (3.6), for all $t \in \mathbb{R}^{2d}$,

$$\begin{aligned} \widehat{Q}_h^{(r)}(t) &= \frac{\mathbf{E}\, \exp\bigl(\bigl\langle h + i\,t,\, \mathbb{A}_{n,k}\, \mathbf{X}_r\bigr\rangle\bigr)}{\mathbf{E}\, \exp\bigl(\bigl\langle h,\, \mathbb{A}_{n,k}\, \mathbf{X}_r\bigr\rangle\bigr)} \\ &= \frac{\mathbf{E}\, \exp\bigl(\bigl\langle \mathbb{A}_{n,k}^*\, h + i\, \mathbb{A}_{n,k}^*\, t, \mathbf{X}_r\bigr\rangle\bigr)}{\mathbf{E}\, \exp\bigl(\bigl\langle \mathbb{A}_{n,k}^*\, h, \mathbf{X}_r\bigr\rangle\bigr)} = \widehat{M}_{\mathbb{A}_{n,k}^*\, h}^{(r)}(\mathbb{A}_{n,k}^*\, t). \end{aligned} \tag{7.32}$$

By (2.2) and (7.31), we have (for $j = 2\,d$)

$$\bigl|\widehat{F}_h(t)\bigr| = \prod_{r=2^n\,(k-1)+1}^{2^{n-1}(2k-1)} \bigl|\widehat{Q}_h^{(r)}(t)\bigr|. \tag{7.33}$$

Write an arbitrary vector $t=\bigl(t_1,\dots,t_{2d}\bigr)\in\mathbb{R}^{2d}$ in the form $t=\bigl(t^{(1)},t^{(2)}\bigr)$, where $t^{(1)}=\bigl(t_1,\dots,t_d\bigr)$, $t^{(2)}=\bigl(t_{d+1},\dots,t_{2d}\bigr)\in\mathbb{R}^d$. Using (7.30), (3.6) and introducing a similar notation for $h=\bigl(h_1,\dots,h_{2d}\bigr)\in\mathbb{R}^{2d}$, it is easy to check that

$$\widehat{M}_h^{(r)}(t)=\widehat{R}_{h^{(1)}}^{(r)}\bigl(t^{(1)}\bigr)\,\widehat{R}_{h^{(2)}}^{(2^{n-1}+r)}\bigl(t^{(2)}\bigr). \tag{7.34}$$

Recall that the distributions $\mathcal{L}(\xi_s)=\mathcal{L}(X_s)$, $s=1,\dots,2^N$, satisfy Condition $\mathbf{C}_\Psi(p)$. Hence, $R_h^{(l)}=\overline{\Psi\bigl(\mathcal{L}(\widetilde{\eta}_l)\bigr)}(h)$ for $l=(m-1)\cdot 2^p+1$, $m=1,\dots,2^{N-p}$. Using Lemma 7.5 and Condition $\mathbf{A}(\tau)$, we see that the following estimate is valid, for such l and for all $t\in\mathbb{R}^d$:

$$\sup_{\|h\|\tau\leqslant c_{26}}\bigl|\widehat{R}_h^{(l)}(t)\bigr|\leqslant\min\Bigl\{\,\max\Bigl\{\,\exp\bigl(-c_{27}\,\|t\|^2\bigr),\,e^{-c_{20}}\Bigr\},\ \frac{1}{(1+c_{28}\,\|t\|^2)}\Bigr\}. \tag{7.35}$$

Choosing c_{29} to be sufficiently small and taking into account that $\tau\geqslant 1$, we ensure the validity of the inequalities

$$2^{n-p-2}\geqslant 2\,d+2\geqslant j+2 \tag{7.36}$$

and

$$\exp\bigl(-c_{20}\cdot 2^{n-p-3}\bigr)=\exp\bigl(-c\cdot 2^{(n-p)/2}\cdot 2^{(n-p)/2}\bigr)\leqslant\frac{1}{2^{n-p}\cdot 2^{(n-p)j/2}}\,. \tag{7.37}$$

Using (7.34)–(7.36), we get

$$\begin{aligned}\sup_{\|h\|\tau\leqslant c_{26}}\bigl|\widehat{M}_h^{(l)}(t)\bigr|&\leqslant\prod_{\mu=1}^{2}\min\Bigl\{\,\max\Bigl\{\,\exp\bigl(-c_{27}\,\|t^{(\mu)}\|^2\bigr),\,e^{-c_{20}}\Bigr\},\ \frac{1}{(1+c_{28}\,\|t^{(\mu)}\|^2\,)}\Bigr\}\\&\leqslant\min\Bigl\{\,\max\Bigl\{\,\exp\bigl(-c_{27}\,\|t\|^2\bigr),\,e^{-c_{20}}\Bigr\},\ \frac{1}{(1+c_{28}\,\|t\|^2)}\Bigr\}.\end{aligned} \tag{7.38}$$

for $l=(m-1)\cdot 2^p+1$, $m=2^{n-p}\,(k-1)+1,\dots,2^{n-p-1}\,(2\,k-1)$ and for all $t\in\mathbb{R}^{2d}$. Note that such l satisfy the inequality $2^n\,(k-1)+1\leqslant l\leqslant 2^{n-1}\,(2\,k-1)$ and the amount of such l equals to 2^{n-p-1}. Using (6.37), (7.32) and (7.38), we see that, for the same l and for all $t\in\mathbb{R}^{2d}$,

$$\sup_{\|h\|\tau\leqslant c_{31}}\bigl|\widehat{Q}_h^{(l)}(t)\bigr|\leqslant\min\Bigl\{\,\max\Bigl\{\,\exp\bigl(-c_{32}\,\|t\|^2\bigr),\,e^{-c_{20}}\Bigr\},\ \frac{1}{(1+c_{33}\,\|t\|^2)}\Bigr\} \tag{7.39}$$

with $c_{31}=c_{26}\,c_{16}$, $c_{32}=c_{27}\,c_{15}^{-1}$, $c_{33}=c_{28}\,c_{15}^{-1}$.

Let us show that the inequality (7.39) implies that, for $h \in \mathbb{R}^j$, $\|h\|\, c_{30}\,\tau < 1$, and for all $t \in \mathbb{R}^j$,

$$\big|\widehat{F}_h(t)\big| \leqslant \Big(\min\Big\{ \max\Big\{ \exp\big(-c_{32}\,\|t\|^2\big),\, e^{-c_{20}} \Big\},\ \frac{1}{(1+c_{33}\,\|t\|^2)} \Big\}\Big)^{2^{n-p-1}}. \quad (7.40)$$

It suffices to prove (7.40) for $j = 2d$ (for $1 \leqslant j < 2d$ one should apply (7.40) for $j = 2d$ and for $h \in \mathbb{R}^{2d}$, $\|h\|\, c_{30}\,\tau < 1$, $t \in \mathbb{R}^{2d}$ with $h_m = t_m = 0$, for $m = j+1, \dots, 2d$). The restriction on h is used here to justify the existence of conjugate distributions F_h only. It remains to note that $c_{30} \geqslant c_{31}^{-1} = c_{26}^{-1}\, c_{16}^{-1}$ and to use (7.33) and (7.39).

Represent the set T defined in (7.28) in the form

$$T = T^{(1)} \cup T^{(2)}, \qquad T^{(1)} = \Big\{ t \in \mathbb{R}^j : \|t\|^2 \geqslant c_{20}/c_{32} \Big\}, \qquad T^{(2)} = T \setminus T^{(1)}. \quad (7.41)$$

Below (in the proof of (7.26) and (7.27)) we assume that $\|h\|\, c_{30}\,\tau < 1$. According to (7.28), (7.40) and (7.41), for $t \in T^{(1)}$ we have

$$\big|\widehat{F}_h(t)\big| \leqslant \frac{1}{(1+c_{33}\,\|t\|^2)^{2^{n-p-2}}} \exp\big(-c_{20}\cdot 2^{n-p-2}\big). \quad (7.42)$$

and, for $t \in T^{(2)}$,

$$\big|\widehat{F}_h(t)\big| \leqslant \exp\Big(-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\, j^2}\Big) \exp\big(-c_{32}\cdot 2^{n-p-2}\,\|t\|^2\big). \quad (7.43)$$

Using (7.29), (7.36) and (7.42), we obtain

$$\begin{aligned}
(2\pi)^{-j} \int\limits_{T^{(1)}} \big(1+\|t\|^2\big)\, \big|\widehat{F}_h(t)\big|\, dt &\leqslant c\, \exp\big(-c_{20}\cdot 2^{n-p-2}\big) \int\limits_{\mathbb{R}^j} \frac{dt}{(1+c_{33}\,\|t\|^2)^j} \\
&\leqslant c\, \exp\big(-c_{20}\cdot 2^{n-p-2}\big) \Big(\int_{\mathbb{R}} \frac{dt}{1+c_{33}\,t^2} \Big)^j \\
&\leqslant \frac{2^{(n-p)j/2}}{2^p\cdot(2\pi)^{j/2}(\det \mathbb{D})^{1/2}} \exp\Big(c_{35}\,(p+1)\,j - c_{20}\cdot 2^{n-p-3} - c_{20}\cdot 2^{n-p-3}\Big). \quad (7.44)
\end{aligned}$$

Similarly, using (7.29) and (7.43), we see that

$$\begin{aligned}
(2\pi)^{-j} \int\limits_{T^{(2)}} \big|\widehat{F}_h(t)\big|\, dt &\leqslant \exp\Big(-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\, j^2}\Big) \int\limits_{\mathbb{R}^j} \exp\big(-c_{32}\cdot 2^{n-p-2}\,\|t\|^2\big)\, dt \\
&\leqslant \frac{1}{2^{(n-p)j/2}} \exp\Big(c\,j - \frac{c_{34}\cdot 2^{n-p}}{\tau^2\, j^2}\Big) \\
&\leqslant \frac{1}{2^p\cdot(2\pi)^{j/2}(\det \mathbb{D})^{1/2}} \exp\Big(c_{36}\,(p+1)\,j - \frac{c_{34}\cdot 2^{n-p}}{\tau^2\, j^2}\Big). \quad (7.45)
\end{aligned}$$

It is evident that, choosing c_{29} to be small enough, we can ensure the validity of the inequalities

$$\exp\Big(c_{36}\,(p+1)\,j-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\,j^2}\Big)\leqslant\frac{c_{30}^2\,\tau^2\,j^2}{2\cdot 2^{n-p}}\leqslant\frac{1}{32}\tag{7.46}$$

and

$$\exp\big(c_{35}\,(p+1)\,j-c_{20}\cdot 2^{n-p-3}\big)\leqslant\frac{c_{30}^2}{2\cdot 2^{n-p}}\leqslant\frac{c_{30}^2\,\tau^2\,j^2}{2\cdot 2^{n-p}}\leqslant\frac{1}{32}\,.\tag{7.47}$$

The relations (7.29), (7.37), (7.41) and (7.44)–(7.47) imply (7.26) and

$$(2\,\pi)^{-j}\int\limits_T\big|\widehat{F}_h(t)\big|\,dt\leqslant\frac{1}{16\,(2\,\pi)^{j/2}\,(\det\mathbb{D})^{1/2}}\,.\tag{7.48}$$

For the proof of (7.27) we apply the Hölder inequality and obtain that

$$\int\limits_T\big|\langle t,v\rangle\,\widehat{F}_h(t)\big|\,dt\leqslant\|v\|\,\Big(\int\limits_T\big|\widehat{F}_h(t)\big|\,dt\Big)^{1/2}\Big(\int\limits_T\|t\|^2\,\big|\widehat{F}_h(t)\big|\,dt\Big)^{1/2}.\tag{7.49}$$

The first integral in the right-hand side of (7.49) is estimated in (7.48). Let us estimate the second integral. For this, using (7.29) and (7.43), we obtain

$$\begin{aligned}(2\,\pi)^{-j}\int\limits_{T^{(2)}}\|t\|^2\,\big|\widehat{F}_h(t)\big|\,dt&\leqslant\int\limits_{\mathbb{R}^j}\|t\|^2\,\exp\Big(-c_{32}\cdot 2^{n-p-2}\,\|t\|^2-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\,j^2}\Big)\,dt\\&\leqslant\frac{1}{2^{n-p}\,.\,2^{(n-p)j/2}}\,\exp\Big(c\,j-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\,j^2}\Big)\\&\leqslant\frac{1}{2^n\,(2\,\pi)^{j/2}(\det\mathbb{D})^{1/2}}\,\exp\Big(c\,(p+1)\,j-\frac{c_{34}\cdot 2^{n-p}}{\tau^2\,j^2}\Big).\end{aligned}\tag{7.50}$$

According to (7.37), (7.41), (7.44) and (7.50), we have

$$(2\,\pi)^{-j}\int\limits_T\|t\|^2\,\big|\widehat{F}_h(t)\big|\,dt\leqslant\frac{1}{2^n\,(2\,\pi)^{j/2}(\det\mathbb{D})^{1/2}}\,\exp\Big(c_{37}\,(p+1)\,j-\frac{c_{38}\cdot 2^{n-p}}{\tau^2\,j^2}\Big).\tag{7.51}$$

It is evident that, choosing c_{29} to be small enough, we can ensure the validity of the inequality

$$\exp\Big(c_{37}\,(p+1)\,j-\frac{c_{38}\cdot 2^{n-p}}{\tau^2\,j^2}\Big)\leqslant 1.\tag{7.52}$$

The relations (7.29), (7.48), (7.49), (7.51) and (7.52) imply (7.27). This completes the proof of the lemma. □

8. Estimation of the rate of approximation in constructing the vectors with needed moments up to the third order

Theorem 8.1. *Let Conditions* $\mathbf{A}^*(\tau)$, $\mathbf{B}$, $\mathbf{D}_\Phi$ *be satisfied for the dyadic scheme described in Sections* 5–7. *Then there exist absolute positive constants* c_{39}, c_{40}, c_{41} *such that*

$$\mathbf{E}\,\exp\Big(\frac{c_{39}\,\Delta(X,Y)}{d^{5/2}\,\tau}\Big)\leqslant \exp\Big(c_{40}\,(N+1)\,\log^* d\Big) \tag{8.1}$$

provided that $\tau\,d^{3/2}\leqslant c_{41}$, *where* $\Delta(X,Y)$ *is defined in* (1.1) *with* $s=2^N$.

Below (up to the formulation of Theorem 8.4) we shall prove Theorem 8.1. Therefore, we assume that its conditions are fulfilled.

Lemma 8.2. *There exist absolute positive constants* c_{41}, c_{42} *such that*

a) *If* $\tau\,d^{3/2}/2^{N/2}\leqslant c_{41}$, *then*

$$\big\|U_{N,1}-V_{N,1}\big\|\leqslant c_{42}\,\tau\,\Big(d^2+d^{3/2}\cdot 2^{-N}\,\big\|U_{N,1}\big\|^2\Big) \tag{8.2}$$

provided that $\big\|U_{N,1}\big\|\leqslant \dfrac{c_{41}\cdot 2^N}{d^{3/2}\,\tau}$;

b) *If* $1\leqslant n\leqslant N$, $1\leqslant k\leqslant 2^{N-n}$, $\tau\,d^{3/2}/2^{n/2}\leqslant c_{41}$, *then*

$$\big\|\widetilde{\mathcal{U}}_{n,k}-\widetilde{\mathcal{V}}_{n,k}\big\|\leqslant c_{42}\,\tau\,\Big(\,d^2+d^{3/2}\cdot 2^{-n}\,\big\|\mathbf{U}_{n,k}\big\|^2\Big) \tag{8.3}$$

provided that $\big\|\mathbf{U}_{n,k}\big\|\leqslant \dfrac{c_{41}\cdot 2^n}{d^{3/2}\,\tau}$.

In the proof of this lemma we need the following auxiliary assertion which is useful for the application of Lemmas 3.3 and 3.5 to conditional distributions involved in the dyadic scheme.

Lemma 8.3. *Let a continuous distribution function* $F(\cdot)$ *and an arbitrary distribution function* $G(\cdot)$ *satisfy for* $z\in B\in\mathfrak{B}_1$ *the inequality*

$$G\big(z-f(z)+w_1\big)<F(z+w_2)<G\big(z+f(z)+w_1\big)$$

with some $f:B\to\mathbb{R}^1$ *and* $w_1, w_2\in\mathbb{R}^1$. *Let* $\eta\in\mathbb{R}^1$, $0<G(\eta)<1$ *and* $\xi=F^{-1}\big(G(\eta)\big)$, *where* $F^{-1}(x)=\sup\big\{u:F(u)\leqslant x\big\}$, $0<x<1$. *Then*

$$|\,\xi-\eta\,|<f(\xi-w_2)+|\,w_2-w_1\,|,\qquad \text{if}\quad \xi-w_2\in B. \tag{8.4}$$

Proof. Put $\zeta=\xi-w_2$. The continuity of F implies that $F\big(F^{-1}(x)\big)\equiv x$, for $0<x<1$. Therefore,

$$\begin{aligned}\zeta\in B &\Longrightarrow G\big(\zeta-f(\zeta)+w_1\big)<F(\xi)=G(\eta)\\ &\Longrightarrow \zeta-f(\zeta)+w_1<\eta \Longrightarrow \xi-\eta<f(\zeta)+w_2-w_1\end{aligned}$$

and

$$\begin{aligned}\zeta \in B &\implies G(\eta) = F(\xi) < G\big(\zeta + f(\zeta) + w_1\big)\\ &\implies \eta < \zeta + f(\zeta) + w_1 \implies \eta - \xi < f(\zeta) + w_1 - w_2.\end{aligned}$$

This completes the proof of the lemma. □

Proof of Lemma 8.2. Let us prove at first the statement a). By condition **B** (see (6.13) and (6.14)), the coordinates $\mathcal{U}^{(1)}, \dots, \mathcal{U}^{(d)}$ of the vectors $\mathcal{U} = \mathbb{J}_{N,1} U_{N,1}$ and $\mathcal{U}^j = \overline{\mathbb{P}}_j(\mathcal{U})$, $j = 1, \dots, d$, are non-correlated and have mean zero and the same variance 2^N. According to Condition $\mathbf{D}_\Phi$, this implies that the distribution functions $G^{(j)}$ defined in the dyadic scheme are Gaussian with mean zero and variance 2^N. Moreover, by Remarks 6.3 and 6.6, we have $\mathcal{L}(\mathcal{V}^j) = \Phi\big(\mathcal{L}(\mathcal{U}^j)\big)$ and $\mathcal{U}^j = \mathbf{U}^j_{N,1}$. Thus, using Condition $\mathbf{A}^*(\tau)$, we see that the distributions $\mathcal{L}(\mathcal{U}^j)$, $\mathcal{L}(\mathcal{V}^j)$, satisfy in the j-dimensional situation the conditions of Lemma 3.3 with $\operatorname{cov}\xi = \operatorname{cov}\mathcal{U}^j = 2^N\,\mathbb{I}_j$, $\mathbb{B} = \operatorname{cov}\mathcal{U}^{j-1} = 2^N\,\mathbb{I}_{j-1}$ (the last equality for $j \geqslant 2$).

Taking into account (5.14), (5.15) and applying Lemmas 3.3, 8.3 in the j-dimensional case, $j = 1, \dots, d$, we obtain that

$$\big|\mathcal{U}^{(1)} - \mathcal{V}^{(1)}\big| \leqslant c\tau\left(1 + \frac{|\mathcal{U}^{(1)}|^2}{2^N}\right), \tag{8.5}$$

provided that $\dfrac{\tau}{2^{N/2}} \leqslant c_{43}$, $\big|\mathcal{U}^{(1)}\big| \leqslant \dfrac{c_{43}\cdot 2^N}{\tau}$. Moreover,

$$\big|\mathcal{U}^{(j)} - \mathcal{V}^{(j)}\big| \leqslant c\tau\left(j^{3/2} + j\,\frac{\|\mathcal{U}^{j-1}\|}{2^{N/2}}\left(1 + \frac{|\mathcal{U}^{(j)} - y_j|}{2^{N/2}}\right) + \frac{|\mathcal{U}^{(j)} - y_j|^2}{2^N}\right) + |y_j|, \tag{8.6}$$

provided that

$$\frac{\tau j^{3/2}}{2^{N/2}} \leqslant c_{43}, \quad \big|\mathcal{U}^{(j)} - y_j\big| \leqslant \frac{c_{43}\cdot 2^N}{j\,\tau}, \quad \frac{|\mathcal{U}^{j-1}|}{2^{N/2}} \leqslant \frac{c_{43}\cdot 2^{N/2}}{j^{3/2}\tau}, \qquad 2 \leqslant j \leqslant d, \tag{8.7}$$

where

$$|y_j| \leqslant c_4\,\tau\,\frac{\|\mathcal{U}^{j-1}\|^2}{2^N}, \qquad 2 \leqslant j \leqslant d. \tag{8.8}$$

Obviously,

$$\big|\mathcal{U}^{(1)}\big| \leqslant \max\big\{\|\mathcal{U}^{j-1}\|,\ |\mathcal{U}^{(j)}|\big\} \leqslant \|\mathcal{U}^j\| \leqslant \|\mathcal{U}\|, \qquad 2 \leqslant j \leqslant d, \tag{8.9}$$

see (5.17) and (5.18). Using (8.5), (8.6), (8.8) and (8.9), we see that one can choose c_{44} to be so large and c_{45} to be so small that

$$\big|\mathcal{U}^{(j)} - \mathcal{V}^{(j)}\big| \leqslant c_{44}\,\tau\left(d^{3/2} + d\,\frac{\|\mathcal{U}\|^2}{2^N}\right), \tag{8.10}$$

provided that

$$\frac{\tau\, d^{3/2}}{2^{N/2}} \leqslant c_{45}, \qquad \bigl\|\mathcal{U}\bigr\| \leqslant \frac{c_{45}\cdot 2^N}{d^{3/2}\tau}, \qquad 1 \leqslant j \leqslant d. \tag{8.11}$$

Taking into account (5.13), (5.17) and (6.23), we see that

$$\bigl\|\mathcal{U}\bigr\| \leqslant C_1^{-1}\bigl\|U_{N,1}\bigr\|, \qquad \bigl\|U_{N,1}-V_{N,1}\bigr\| \leqslant C_2\bigl\|\mathcal{U}-\mathcal{V}\bigr\|. \tag{8.12}$$

The assertion a) follows from (8.10)–(8.12) if we choose c_{41} to be sufficiently small and c_{42} sufficiently large.

Let us prove now statement b). By condition **B** (see (6.28) and (6.29)), the coordinates $\mathbf{U}_{n,k}^{(1)}, \ldots, \mathbf{U}_{n,k}^{(2d)}$ of the vectors $\mathbf{U}_{n,k} = \mathbb{A}_{n,k}\,\widetilde{\mathbf{U}}_{n,k}$ and $\mathbf{U}_{n,k}^{j} = \overline{\mathbb{P}}_j(\mathbf{U}_{n,k})$, $j=1,\ldots,2d$, are non-correlated and have the same variance 2^n. By Condition $\mathbf{D}_\Phi$, this implies that the distribution functions $\widetilde{G}_{n,k}^{(j)}$, $j=d+1,\ldots,2d$, defined in the dyadic scheme are Gaussian with mean zero and variance 2^n. Moreover, by Remark 6.3, $\mathcal{L}(\mathbf{V}_{n,k}^{j}) = \Phi\bigl(\mathcal{L}(\mathbf{U}_{n,k}^{j})\bigr)$, $j=1,\ldots,2d$. Thus, using Condition $\mathbf{A}^*(\tau)$, we see that the distributions $\mathcal{L}(\mathbf{U}_{n,k}^{j}), \mathcal{L}(\mathbf{V}_{n,k}^{j})$, $j=d+1,\ldots,2d$, satisfy in the j-dimensional situation the conditions of Lemma 3.3 with $\operatorname{cov}\xi = \operatorname{cov}\mathbf{U}_{n,k}^{j} = 2^n\,\mathbb{I}_j$, $\mathbb{B} = \operatorname{cov}\mathbf{U}_{n,k}^{j-1} = 2^n\,\mathbb{I}_{j-1}$.

Taking into account (5.20) and applying Lemmas 3.3, 8.3 in the j-dimensional case, $j=d+1,\ldots,2d$, we obtain that

$$\bigl|\mathbf{U}_{n,k}^{(j)} - \mathbf{V}_{n,k}^{(j)}\bigr| \leqslant c\tau\left(j^{3/2} + j\,\frac{\|\mathbf{U}_{n,k}^{j-1}\|}{2^{n/2}}\left(1 + \frac{|\mathbf{U}_{n,k}^{(j)} - y_j|}{2^{n/2}}\right) + \frac{|\mathbf{U}_{n,k}^{(j)} - y_j|^2}{2^n}\right) + |y_j|, \tag{8.13}$$

provided that

$$\frac{\tau\, j^{3/2}}{2^{n/2}} \leqslant c_{46}, \quad \bigl|\mathbf{U}_{n,k}^{(j)} - y_j\bigr| \leqslant \frac{c_{46}\cdot 2^n}{j\,\tau}, \quad \frac{|\mathbf{U}_{n,k}^{j-1}|}{2^{n/2}} \leqslant \frac{c_{46}\cdot 2^{n/2}}{j^{3/2}\tau}, \tag{8.14}$$

where

$$|y_j| \leqslant c_4\,\tau\,\frac{\|\mathbf{U}_{n,k}^{j-1}\|^2}{2^n}, \qquad d+1 \leqslant j \leqslant 2d. \tag{8.15}$$

Obviously,

$$\max\Bigl\{\bigl\|\mathbf{U}_{n,k}^{j-1}\bigr\|, \bigl|\mathbf{U}_{n,k}^{(j)}\bigr|\Bigr\} \leqslant \bigl\|\mathbf{U}_{n,k}^{j}\bigr\| \leqslant \bigl\|\mathbf{U}_{n,k}\bigr\|, \qquad d+1 \leqslant j \leqslant 2d, \tag{8.16}$$

see (5.17) and (5.18). Using (8.13), (8.15), and (8.16), we see that one can choose c_{47} to be so large and c_{48} to be so small that

$$\bigl|\mathbf{U}_{n,k}^{(j)} - \mathbf{V}_{n,k}^{(j)}\bigr| \leqslant c_{47}\,\tau\left(d^{3/2} + d\,\frac{\|\mathbf{U}_{n,k}\|^2}{2^n}\right), \tag{8.17}$$

provided that

$$\frac{\tau\, d^{3/2}}{2^{n/2}} \leqslant c_{48}, \qquad \big\| \mathbf{U}_{n,k} \big\| \leqslant \frac{c_{48} \cdot 2^n}{d^{3/2} \tau}, \qquad d+1 \leqslant j \leqslant 2\,d. \tag{8.18}$$

The assertion b) follows from (6.4)–(6.8), (8.17) and (8.18) provided that we choose c_{41} to be sufficiently small and c_{42} sufficiently large. □

Proof of Theorem 8.1. Write

$$\Delta = \Delta(X, Y) = \max_{1 \leqslant k \leqslant 2^N} \big\| S_k - T_k \big\|, \tag{8.19}$$

$$c_{49} = \min\big\{ c_{41},\, c_{16}\, c_{41} \sqrt{2},\, C_2^2 \big\}, \qquad y \overset{\text{def}}{=} \frac{c_{49}}{d^{3/2}\, \tau} \leqslant \frac{C_2^2}{\tau}, \tag{8.20}$$

Let us prove that there exists a c_{50} such that

$$\mathbf{P}\Big\{ \Delta \geqslant x \Big\} \leqslant (15\, d)^{N+1} \exp\Big(- \frac{x}{c_{50}\, d^{5/2}\, \tau} \Big), \qquad x > 0, \tag{8.21}$$

provided that $\tau\, d^{3/2} \leqslant c_{41}$. Fix some $x > 0$, and choose the integer M satisfying the conditions

$$x < 8\, y \cdot 2^M \leqslant 2\, x. \tag{8.22}$$

Write

$$P = \max\big\{ 0,\, M \big\}, \qquad L = \min\big\{ N,\, P \big\} \tag{8.23}$$

and

$$\Delta_1 = \max_{0 \leqslant k < 2^{N-L}} \; \max_{1 \leqslant s \leqslant 2^L} \big\| S_{k \cdot 2^L + s} - S_{k \cdot 2^L} \big\|, \tag{8.24}$$

$$\Delta_2 = \max_{0 \leqslant k < 2^{N-L}} \; \max_{1 \leqslant s \leqslant 2^L} \big\| T_{k \cdot 2^L + s} - T_{k \cdot 2^L} \big\|, \tag{8.25}$$

$$\Delta_3 = \max_{0 \leqslant k \leqslant 2^{N-P}} \big\| S_{k \cdot 2^L} - T_{k \cdot 2^L} \big\|, \tag{8.26}$$

where the maximums are taken over integer values of k and s. It is evident that

$$\Delta \leqslant \Delta_1 + \Delta_2 + \Delta_3. \tag{8.27}$$

Define the event

$$A = \Big\{ \omega : \max\big\{ \Delta_1,\, \Delta_2 \big\} < z \Big\}, \qquad \text{where} \quad z = \max\big\{ y \cdot 2^{L-1},\, x/8 \big\} \tag{8.28}$$

(we shall always assume that all considered random vectors are measurable mappings of elementary event $\omega \in \Omega$). For the complementary event we shall use the notation $\overline{A} = \Omega \setminus A$. Recall that all d-dimensional Gaussian distributions belong to classes $\mathcal{A}_d(\tau)$ for all $\tau \geqslant 0$. Applying Lemma 2.7 with $B^2 \leqslant C_2^2 \cdot 2^L$ (see (6.2) and (6.16)) and using (8.20), (8.24), (8.25) and (8.28), we obtain that

$$\begin{aligned} \mathbf{P}\big\{ \overline{A} \big\} = 1 - \mathbf{P}\big\{ A \big\} &\leqslant 2^N \cdot 4d \max\Big\{ \exp\big(-z^2/4dB^2\big),\ \exp\big(-z/4\tau\sqrt{d}\big) \Big\} \\ &\leqslant 2^N \cdot 4d \exp\Big(-\frac{cx}{d^{5/2}\tau} \Big). \end{aligned} \tag{8.29}$$

If $M > N$ then $L = N < M = P$, $\Delta_3 = 0$, $z = x/8$, and, for $\omega \in A$, it is easy to see that $\Delta \leqslant x/4$ (one should use (8.22), (8.23), (8.26), (8.27) and (8.28)). Hence, $\mathbf{P}\big\{ \Delta \geqslant x \big\} \leqslant 1 - \mathbf{P}\big\{ A \big\}$ and one can conclude via (8.29).

Let now $M \leqslant N$. Then, by (8.22), (8.23) and (8.28), we have

$$0 \leqslant L = P \leqslant N, \qquad M \leqslant L, \qquad z \leqslant y \cdot 2^L. \tag{8.30}$$

Note that if $L = P = 0$ then $\Delta = \Delta_3$ and if $L = P = M$ then $z = x/8$ (see (8.22), (8.26), (8.28) and (8.30)). Thus, by (8.27) and (8.28), in both possible cases we have

$$\Delta \leqslant x/4 + \Delta_3, \qquad \text{for} \quad \omega \in A. \tag{8.31}$$

Fix an integer k satisfying $1 \leqslant k \leqslant 2^{N-L}$ and denote for simplicity

$$j = j(k) \overset{\text{def}}{=} k \cdot 2^L. \tag{8.32}$$

By Lemma 6.7 and Remark 6.8, we have

$$\big\| S_{k\cdot 2^L} - T_{k \cdot 2^L} \big\| = \big\| S_j - T_j \big\| \leqslant c_{17} \big\| U_{N,1} - V_{N,1} \big\| + c_{17} \sum_{n=L+1}^{N} \big\| \widetilde{\mathcal{U}}_{n,l_{n,j}} - \widetilde{\mathcal{V}}_{n,l_{n,j}} \big\|, \tag{8.33}$$

where $l_{n,j}$ are integers, defined by $\big(l_{n,j} - 1\big) \cdot 2^n < j \leqslant l_{n,j} \cdot 2^n$.

Recall that $U_{n,l} = S_{l \cdot 2^n} - S_{(l-1)\cdot 2^n}$. By virtue of (8.20), (8.24), (8.28) and (8.30), for $\omega \in A$ we have

$$\big| U_{L,l} \big| \leqslant \big\| U_{L,l} \big\| \leqslant \Delta_1 \leqslant z \leqslant y \cdot 2^L = \frac{c_{49} \cdot 2^L}{d^{3/2}\tau}, \qquad 1 \leqslant l \leqslant 2^{N-L}, \tag{8.34}$$

and, by (5.2), (5.22)–(5.24), $U_{L,l}$ are sums over blocks consisting of 2^L summands. Moreover, $U_{n,l}$, $L+1 \leqslant n \leqslant N$, $1 \leqslant l \leqslant 2^{N-n}$, are sums over blocks containing each 2^n summands. These sums can be represented as a sum of 2^{n-L} sums over blocks containing each 2^L summands and satisfying (8.34). Therefore, for $\omega \in A$, $L+1 \leqslant n \leqslant N$, $1 \leqslant l \leqslant 2^{N-n}$ we have (see (5.24), (6.39), (8.20) and (8.34))

$$\big| U_{n,l} \big| \leqslant \big\| U_{n,l} \big\| \leqslant 2^{n-L} z \leqslant y \cdot 2^n = \frac{c_{49} \cdot 2^n}{d^{3/2}\tau} \leqslant \frac{c_{41} \cdot 2^n}{d^{3/2}\tau}, \tag{8.35}$$

$$\|\widetilde{\mathbf{U}}_{n,l}\| = \left(\|U_{n-1,2l-1}\|^2 + \|U_{n-1,2l}\|^2\right)^{1/2} \leqslant \frac{c_{49}\cdot 2^n}{\sqrt{2}\,d^{3/2}\,\tau}, \tag{8.36}$$

$$\|\mathbf{U}_{n,l}\| = \|\mathbb{A}_{n,l}\,\widetilde{\mathbf{U}}_{n,l}\| \leqslant \|\mathbb{A}_{n,l}\|\cdot\|\widetilde{\mathbf{U}}_{n,l}\| \leqslant \frac{c_{16}^{-1}\,c_{49}\cdot 2^n}{\sqrt{2}\,d^{3/2}\,\tau} \leqslant \frac{c_{41}\cdot 2^n}{d^{3/2}\,\tau}. \tag{8.37}$$

Using (8.20), (8.35) and (8.37), we see that the conditions of Lemma 8.2 are satisfied for τ, $U_{N,1}$ and $\mathbf{U}_{n,l}$, if $\omega\in A$, $\tau\,d^{3/2}\leqslant c_{41}$, $L+1\leqslant n\leqslant N$, $1\leqslant l\leqslant 2^{N-n}$. By (6.39), (8.33), (8.36), (8.37) and by Lemma 8.2, for $\omega\in A$ we have

$$\begin{aligned}\|S_j-T_j\| &\leqslant c\tau\left(d^2+d^{3/2}\cdot 2^{-N}\,\|U_{N,1}\|^2\right)+\sum_{n=L+1}^{N}c\tau\left(d^2+d^{3/2}\cdot 2^{-n}\,\|\mathbf{U}_{n,l_{n,j}}\|^2\right)\\ &\leqslant c\tau\left(d^2(N+1)+d^{3/2}\left(2^{-N}\,\|U_{N,1}\|^2+\sum_{n=L}^{N-1}2^{-n}\left(\|U^{(n)}\|^2+\|U_{(n)}\|^2\right)\right)\right),\end{aligned} \tag{8.38}$$

where

$$U^{(n)} = U_{n,l_{n,j}}, \qquad U_{(n)} = U_{n,\widetilde{l}_{n,j}}, \tag{8.39}$$

and

$$\widetilde{l}_{n-1,j} = \begin{cases} 2\,l_{n,j}, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j}-1,\\ 2\,l_{n,j}-1, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j},\end{cases} \qquad L<n\leqslant N, \tag{8.40}$$

(it is not hard to verify that $l_{n-1,j}$ can be equal either to $2\,l_{n,j}$ or to $2\,l_{n,j}-1$ if $l_{n,j}$ is fixed). According to Remark 5.5, $U^{(n)}$, $L\leqslant n\leqslant N$, are the sums over the blocks of 2^n summands which contains X_j. The sum $U_{(n)}$ does not contain X_j and

$$U^{(n+1)} = U^{(n)} + U_{(n)}, \qquad L\leqslant n<N \tag{8.41}$$

(see (5.24)). The equality (8.41) implies

$$U^{(n)} = U^{(L)} + \sum_{s=0}^{n-L-1} U_{(L+s)}, \qquad L\leqslant n\leqslant N. \tag{8.42}$$

It is important that all summands in the right-hand side of (8.42) are the sums of disjoint blocks of independent summands. Therefore, they are themselves independent. This argument will be used below in similar situations.

Put $\beta = 1/\sqrt{2}$. Then, using (8.42) and the Hölder inequality, one can easily derive that, for $L\leqslant n\leqslant N$,

$$\|U^{(n)}\|^2 \leqslant c_{51}\left(\beta^{-(n-L)}\,\|U^{(L)}\|^2 + \sum_{s=0}^{n-L-1}\beta^{-(n-L-1)+s}\,\|U_{(L+s)}\|^2\right), \tag{8.43}$$

with $c_{51} = \sum\limits_{j=0}^{\infty} \beta^j = \dfrac{\sqrt{2}}{\sqrt{2}-1}$. It is easy to see that

$$\sum_{n=L}^{N} 2^{-n}\,\beta^{-(n-L)}\,\big\|\,U^{(L)}\big\|^2 \leqslant c_{51}\cdot 2^{-L}\,\big\|\,U^{(L)}\big\|^2. \tag{8.44}$$

Moreover,

$$\begin{aligned}\sum_{n=L+1}^{N}\ \sum_{s=0}^{n-L-1} 2^{-n}\,\beta^{-(n-L-1)+s}\,\big\|\,U_{(L+s)}\big\|^2 &\\ = \sum_{s=0}^{N-L-1}\ \sum_{n=L+1+s}^{N} 2^{-n}\,\beta^{-(n-L-1)+s}\,\big\|\,U_{(L+s)}\big\|^2 &\\ \leqslant c_{51} \sum_{s=0}^{N-L-1} 2^{-(L+1+s)}\,\big\|\,U_{(L+s)}\big\|^2. &\end{aligned} \tag{8.45}$$

It is clear that the inequalities (8.43)–(8.45) imply

$$\begin{aligned} 2^{-N}\,\big\|\,U_{N,1}\big\|^2 + \sum_{n=L}^{N-1} 2^{-n}\Big(\big\|\,U^{(n)}\big\|^2 + \big\|\,U_{(n)}\big\|^2\Big) &\\ \leqslant c_{51}\left(\frac{\|\,U^{(L)}\|^2}{2^L} + \sum_{s=0}^{N-L-1}\frac{\|\,U_{(L+s)}\|^2}{2^{L+1+s}}\right) + \sum_{n=L}^{N-1}\frac{\|\,U_{(n)}\|^2}{2^n} &\\ \leqslant c\left(\frac{\|\,U^{(L)}\|^2}{2^L} + \sum_{n=L}^{N-1}\frac{\|\,U_{(n)}\|^2}{2^n}\right). &\end{aligned} \tag{8.46}$$

From (8.38) and (8.46) it follows that for $\omega\in A$ we have

$$\begin{aligned}\big\|\,S_j - T_j\big\| &\leqslant c_{52}\,\tau\left(d^2\,(N+1) + d^{3/2}\left(\frac{\|\,U^{(L)}\|^2}{2^L} + \sum_{n=L}^{N-1}\frac{\|\,U_{(n)}\|^2}{2^n}\right)\right)\\ &\leqslant c_{52}\,d^{5/2}\,\tau\left(N+1+\frac{|\,U^{(L)}|^2}{2^L} + \sum_{n=L}^{N-1}\frac{|\,U_{(n)}|^2}{2^n}\right).\end{aligned} \tag{8.47}$$

Write (for $0\leqslant n\leqslant N,\ 1\leqslant l\leqslant 2^{N-n}$)

$$W_{n,l} = \begin{cases} 2^{-n}\,C_2^{-2}\,\big|\,U_{n,l}\big|^2, & \text{if } \ \big|\,U_{n,l}\big|\leqslant y\cdot 2^n,\\ 0, & \text{otherwise.}\end{cases} \tag{8.48}$$

Let us show that

$$\mathbf{E}\,\exp\big(t\,W_{n,l}\big)\leqslant 2\,d+1, \qquad \text{for} \quad 0\leqslant t\leqslant \frac{1}{8}. \tag{8.49}$$

Indeed, integrating by parts, we obtain

$$\begin{aligned}\mathbf{E}\,\exp\big(tW_{n,l}\big) &= 1+\int_0^{\gamma} t\,e^{t\,u}\,\mathbf{P}\big\{W_{n,l}\geqslant u\big\}\,du \\ &\leqslant 1+\frac{1}{8}\int_0^{\gamma} e^{u/8}\,\mathbf{P}\Big\{\big|U_{n,l}\big|\geqslant C_2\cdot 2^{n/2}\sqrt{u}\Big\}\,du,\end{aligned}\tag{8.50}$$

where $\gamma=y^2\cdot 2^n\,C_2^{-2}$. Taking into account (5.24), (6.2), (6.16), (8.20) and using Lemma 2.7 with $B^2\leqslant C_2^2\cdot 2^n$, we obtain that

$$\begin{aligned}\mathbf{P}\Big\{\big|U_{n,l}\big|\geqslant C_2\cdot 2^{n/2}\sqrt{u}\Big\} &\leqslant 2\,d\,\exp\Big(-\min\Big\{\frac{2^n\,C_2^2\,u}{4\cdot 2^nC_2^2}\,,\,\frac{2^{n/2}\,C_2\sqrt{u}}{4\,\tau}\Big\}\Big) \\ &\leqslant 2\,d\,\exp\Big(-\min\Big\{\frac{u}{4}\,,\,\frac{C_2^2\,u}{4\,y\,\tau}\Big\}\Big)=2\,d\,e^{-u/4},\end{aligned}\tag{8.51}$$

if $0\leqslant u\leqslant\gamma=y^2\cdot 2^n\,C_2^{-2}$. The relation (8.49) immediately follows from (8.50) and (8.51).

The relations (8.34), (8.35) and (8.48) imply that, for $L\leqslant n\leqslant N$, $1\leqslant l\leqslant 2^{N-n}$, $\omega\in A$,

$$2^{-n}\big|U_{n,l}\big|^2=C_2^2\,W_{n,l}.\tag{8.52}$$

Thus, according to (8.39), we can rewrite (8.47) in the form

$$\big\|S_j-T_j\big\|\leqslant c_{52}\,d^{5/2}\,\tau\,\Big(N+1+W^{(L)}+\sum_{n=L}^{N-1}W_{(n)}\Big),\qquad \omega\in A,\tag{8.53}$$

where

$$W^{(L)}=W_{L,l_{L,j}},\qquad W_{(n)}=W_{n,\widetilde{l}_{n,j}}.\tag{8.54}$$

Using that the random variables $W^{(L)}$, $W_{(L)}$, … , $W_{(N-1)}$ are independent and applying (8.26), (8.32), (8.49), (8.53) and (8.54), we obtain

$$\mathbf{E}\,e^{h\Delta_3}\mathbf{1}_A\leqslant 2^N\cdot(2\,d+1)^{N+1}\,e^{(N+1)/8},\qquad\text{where}\quad h=\big(8\,c_{52}\,d^{5/2}\,\tau\big)^{-1}\tag{8.55}$$

and $\mathbf{1}_A$ is the indicator function of the event A. By (8.31),

$$\mathbf{P}\big\{\Delta\geqslant x\big\}\leqslant\mathbf{P}\big\{\overline{A}\big\}+\mathbf{P}\big\{\Delta_3\geqslant 3x/4,\,A\big\}\leqslant\mathbf{P}\big\{\overline{A}\big\}+e^{-3hx/4}\,\mathbf{E}\,e^{h\Delta_3}\mathbf{1}_A.\tag{8.56}$$

Now (8.21) follows via (8.29), (8.55) and (8.56).

Let the quantities ε, $x_0>0$ be defined by the relations

$$\varepsilon=\frac{1}{2\,c_{50}\,d^{5/2}\,\tau}\,,\qquad e^{\varepsilon x_0}=(15\,d)^{N+1}.\tag{8.57}$$

Integrating by parts and using (8.21) and (8.57), we obtain

$$\mathbf{E}\, e^{\varepsilon\Delta} = \int_0^\infty \varepsilon\, e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx + 1,$$

$$\int_0^{x_0} \varepsilon\, e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx \leqslant \int_0^{x_0} \varepsilon\, e^{\varepsilon x}\, dx = e^{\varepsilon x_0} - 1 = (15\, d)^{N+1} - 1,$$

$$\int_{x_0}^\infty \varepsilon\, e^{\varepsilon x}\ \mathbf{P}\left\{\Delta \geqslant x\right\} dx \leqslant \int_{x_0}^\infty \varepsilon\, e^{-\varepsilon(x-x_0)}\, dx = 1,$$

and, hence, $\mathbf{E}\, e^{\varepsilon\Delta} \leqslant (15\, d)^{N+1} + 1 \leqslant (16\, d)^{N+1}$. Together with (8.19) and (8.57), this completes the proof of Theorem 8.1. □

Theorem 8.4. *Let* $\tau \geqslant 1$ *and Conditions* $\mathbf{A}(\tau)$, $\mathbf{B}$, $\mathbf{D}_\Phi$ *and* $\mathbf{C}_\Psi(p)$ *be satisfied with* $p = 0$ *for the dyadic scheme described in Sections* 5–7*. Then there exist absolute positive constants* c_{53}, c_{54} *such that*

$$\mathbf{E}\, \exp\Big(\frac{c_{53}\, \Delta(X,Y)}{d^{5/2}\, \tau} \Big) \leqslant \exp\Big(c_{54}\, (N+1)\, \log^*\, d \Big), \tag{8.58}$$

where $\Delta(X,Y)$ *is defined in* (1.1) *with* $s = 2^N$.

Proof. By Lemma 7.6 and by the condition $\mathbf{A}(\tau)$,

$$\mathcal{L}(\mathbf{U}_{n,k}^j) \in \mathcal{A}_j^*(\,c_{30}\,\tau,\ 4,\ 4), \qquad 1 \leqslant k \leqslant 2^{N-n}, \quad j = 1, \dots, 2\, d, \tag{8.59}$$

for any positive integer n satisfying $\tau^2\, d^3/2^n \leqslant c_{29}$, and $1 \leqslant n \leqslant N$. Put

$$c_{55} = \min\{\, c_{29},\ c_{41}^2\, c_{30}^{-2},\ c_{30}^{-2},\ 1/2\,\}. \tag{8.60}$$

Let the integer M be defined by the relation

$$\frac{d^3\, \tau^2}{c_{55}} \leqslant 2^M < \frac{2\, d^3\, \tau^2}{c_{55}}. \tag{8.61}$$

It is clear that, by (8.60) and (8.61) and since $\tau \geqslant 1$, we have $M \geqslant 1$. Let $M \leqslant N$. Write

$$\begin{aligned} \widetilde{X}_m &= 2^{-M/2}\big(\,X_{(m-1)\cdot 2^M+1} + \dots + X_{m\cdot 2^M}\,\big) = 2^{-M/2}\, U_{M,m}, \\ \widetilde{Y}_m &= 2^{-M/2}\big(\,Y_{(m-1)\cdot 2^M+1} + \dots + Y_{m\cdot 2^M}\,\big) = 2^{-M/2}\, V_{M,m}, \end{aligned} \qquad m = 1, \dots, 2^{N-M}. \tag{8.62}$$

It is easy to see that, according to Lemma 5.3 and Remark 5.2, the random vectors $2^{-M/2}\, U_{n,k}$, $M \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$, are the quantile transformations of the vectors $2^{-M/2}V_{M,1} = \widetilde{Y}_1, \ldots, 2^{-M/2}V_{M,2^{N-M}} = \widetilde{Y}_{2^{N-M}}$ associated with the operators $\mathbb{A}_{n,k}$, $M \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$. One can verify that the corresponding *reduced dyadic scheme* (considered with the change of 2^N by 2^{N-M} and with multiplying of all vectors by $2^{-M/2}$) satisfies the Conditions $\mathbf{B}$ and $\mathbf{D}_\Phi$. For example, Condition $\mathbf{B}_2$ for $\operatorname{cov}\widetilde{X}_m$, $1 \leqslant m \leqslant 2^{N-M}$, obviously follows from (6.3), (6.19) and (8.62).

Let us show that the reduced dyadic scheme satisfies Condition $\mathbf{A}^*(\tau_0)$, where

$$\tau_0 = c_{30}\,\tau \cdot 2^{-M/2}. \tag{8.63}$$

Indeed, using (8.59)–(8.61) and Lemma 3.2, we see that

$$\mathcal{L}\big(2^{-M/2}\,\mathbf{U}^j_{n,k}\big) \in \mathcal{A}^*_j(\tau_0,\, 4,\, 4), \qquad n = M,\ldots,N, \quad 1 \leqslant k \leqslant 2^{N-n}, \quad j = 1,\ldots,2\,d. \tag{8.64}$$

By (6.29), the minimal eigenvalues of covariance operators of the vectors $2^{-M/2}\,\mathbf{U}^j_{n,k}$ from (8.64) cannot be less than 1. Moreover, by (8.60), (8.61) and (8.63) we have $\tau_0\, d^{1/2} \leqslant 1$. Therefore, from (3.4), (3.5) and (8.64) it follows that

$$\mathcal{L}\big(2^{-M/2}\,\mathbf{U}^j_{n,k}\big) \in \overline{\mathcal{A}}_j(\tau_0, 4), \qquad n = M,\ldots,N, \quad 1 \leqslant k \leqslant 2^{N-n}, \quad j = 1,\ldots,2\,d. \tag{8.65}$$

The relations (8.65) provide the validity of the required Condition $\mathbf{A}^*(\tau_0)$.

The relations (8.60), (8.61) and (8.63) imply that $\tau_0\, d^{3/2} \leqslant c_{41}$. Thus, the reduced dyadic scheme satisfies the the conditions of Theorem 8.1 with the change of 2^N by 2^{N-M} and τ by τ_0. By this theorem,

$$\mathbf{E}\,\exp\Big(2^{M/2}\,b\,\Delta(\widetilde{X},\widetilde{Y})\Big) \leqslant \exp\Big(c_{40}\,(N+1)\,\log^* d\Big), \tag{8.66}$$

where

$$\Delta(\widetilde{X},\widetilde{Y}) = \max_{1 \leqslant k \leqslant 2^{N-M}} \Big\| \sum_{i=1}^k \widetilde{X}_i - \sum_{i=1}^k \widetilde{Y}_i \Big\| \tag{8.67}$$

and

$$b = \frac{c_{56}}{d^{5/2}\,\tau} \qquad \text{with some} \quad c_{56} < {}^1\!/_2. \tag{8.68}$$

Note that Theorem 8.1 yields (8.66) with

$$b = \frac{c_{39}}{d^{5/2}\,\tau_0 \cdot 2^{M/2}} = \frac{c_{39}}{c_{30}\, d^{5/2}\,\tau}\,. \tag{8.69}$$

This implies (8.66) with b defined in (8.68) and $c_{56} = \min\big\{\, c_{39}\, c_{30}^{-1},\ {}^1\!/_3 \,\big\}$.

Write, by analogy with the proof of Theorem 8.1,

$$\Delta_1 = \max_{0\leqslant k<2^{N-M}} \max_{1\leqslant s\leqslant 2^M} \left| S_{k\cdot 2^M+s} - S_{k\cdot 2^M} \right|, \tag{8.70}$$

$$\Delta_2 = \max_{0\leqslant k<2^{N-M}} \max_{1\leqslant s\leqslant 2^M} \left| T_{k\cdot 2^M+s} - T_{k\cdot 2^M} \right|, \tag{8.71}$$

$$\Delta_3 = \max_{1\leqslant k\leqslant 2^{N-M}} \left\| S_{k\cdot 2^M} - T_{k\cdot 2^M} \right\|, \tag{8.72}$$

where the maximums are taken over integer values of k and s. It is evident that from (1.1), (5.7), (5.23), (8.70)–(8.72) it follows that

$$\Delta(X,Y) \leqslant d^{1/2}\,\Delta_1 + d^{1/2}\,\Delta_2 + \Delta_3. \tag{8.73}$$

Hence, by the Hölder inequality

$$\mathbf{E}\,\exp\bigl(b\,\Delta(X,Y)/3\bigr) \leqslant \Bigl(\mathbf{E}\,\exp\bigl(b\,d^{1/2}\,\Delta_1\bigr)\,\mathbf{E}\,\exp\bigl(b\,d^{1/2}\,\Delta_2\bigr)\,\mathbf{E}\,\exp\bigl(b\,\Delta_3\bigr)\Bigr)^{1/3}. \tag{8.74}$$

Moreover, the relations (1.1), (8.62), (8.67), (8.72), imply

$$\Delta_3 = 2^{M/2}\,\Delta(\widetilde{X},\widetilde{Y}). \tag{8.75}$$

From (8.68) it follows that $0 < d^{1/2}\,b\tau < {}^1\!/_2$. Recall that all d-dimensional Gaussian distributions belong to all classes $\mathcal{A}_d(\tau)$, $\tau \geqslant 0$. Using Lemma 2.7 with $t = d^{1/2}\,b$, $B^2 \leqslant 2^M\,C_2^2$ (see (6.3) and (6.19)), (8.61), (8.70), (8.71) and (8.68), we have

$$\mathbf{E}\,\exp\bigl(d^{1/2}\,b\Delta_l\bigr) \leqslant 2^{N-M}\cdot 3\,d\,\exp\bigl(4\,d\,b^2\cdot 2^M\,C_2^2\bigr) \leqslant c\,d\cdot 2^N, \qquad l=1,2. \tag{8.76}$$

Now the inequality (8.58) can be easily derived from (8.66), (8.68), (8.74)–(8.76) by a suitable choice of c_{53} and c_{54}.

Let now $M > N$. From (8.61) it follows that $2^N < \dfrac{2\,d^3\,\tau^2}{c_{55}}$. We shall use the rough bound

$$\Delta(X,Y) \leqslant d^{1/2}\,\Delta_4 + d^{1/2}\,\Delta_5, \tag{8.77}$$

where

$$\Delta_4 = \max_{1\leqslant k\leqslant 2^N} \left|S_k\right|, \qquad \Delta_5 = \max_{1\leqslant k\leqslant 2^N} \left|T_k\right|. \tag{8.78}$$

Applying Lemma 2.7 with $t = \dfrac{1}{3\,\tau\,d^{3/2}}$, $B^2 \leqslant 2^N\,C_2^2$ (see (6.3) and (6.19)), we obtain the inequality

$$\mathbf{E}\,\exp\Bigl(t\,\Delta_l\Bigr) \leqslant 3\,d\,\exp\bigl(4t^2\cdot 2^N\,C_2^2\bigr) \leqslant c\,d, \qquad l=4,5. \tag{8.79}$$

Now the statement of Theorem 8.4 in the case $M > N$ can be easily derived from (8.77)–(8.79) by an application of the Hölder inequality. □

9. Estimates in the dyadic scheme in the case of coinciding moments of the first three orders

Lemma 9.1. *Let $\tau \geqslant 1$ and Conditions $\mathbf{A}(\tau)$, $\mathbf{B}$, $\mathbf{D}_\Psi$ and $\mathbf{C}_\Psi(p)$ be satisfied with some p, $1 \leqslant p \leqslant N$ for the dyadic scheme described in Sections 5–7. Then there exist absolute positive constants c_{57}, c_{58} such that*

a) *If $\tau^2 d^3 p/2^{N-p} \leqslant c_{57}$, then*

$$\left\| U_{N,1} - V_{N,1} \right\| \leqslant c_{58}\,\tau^2 \cdot 2^{-N/2} \left(d^{5/2} + d \cdot 2^{-3N/2} \left\| U_{N,1} \right\|^3 \right) \tag{9.1}$$

provided that $\left\| U_{N,1} \right\| \leqslant \dfrac{c_{57} \cdot 2^N}{d^{3/2}\,\tau}$.

b) *If $1 \leqslant n \leqslant N$, $1 \leqslant k \leqslant 2^{N-n}$, $\tau^2 d^3 p/2^{n-p} \leqslant c_{57}$, then*

$$\left\| \widetilde{\mathcal{U}}_{n,k} - \widetilde{\mathcal{V}}_{n,k} \right\| \leqslant \frac{c_{58}\,\tau^2}{2^{n/2}} \left(d^{5/2} + d \cdot 2^{-3n/2} \left\| \mathbf{U}_{n,k} \right\|^3 \right) + \frac{1}{8\,c_{17}} \left\| U_{n,k} - V_{n,k} \right\| \tag{9.2}$$

provided that $\max\left\{ \left\| \mathbf{U}_{n,k} \right\|, \left\| V_{n,k} \right\| \right\} \leqslant \dfrac{c_{57} \cdot 2^n}{d^{3/2}\,\tau}$;

Proof. From Lemma 7.3, from the definition of the distributions $\Psi(\cdot)$ and from Conditions $\mathbf{B}_1$, $\mathbf{D}_\Psi$ and $\mathbf{C}_\Psi(p)$ it follows that the vectors ξ_k, X_k and Y_k have the same moments of the first three orders for each $k = 1, \ldots, 2^N$. Using the linearity and independence arguments, we see that the same can be said about corresponding moments of the random vectors $\mathbf{U} = \left\{ U_{n,k}, \widetilde{\mathbf{U}}_{n,k}, \mathbf{U}_{n,k}, \mathcal{U}_{n,k}, \mathcal{U} \right\}$ and $\mathbf{V} = \left\{ V_{n,k}, \widetilde{\mathbf{V}}_{n,k}, \mathbf{V}_{n,k}, \mathcal{V}_{n,k}, \mathcal{V} \right\}$ from Remark 6.3.

Let us prove at first the statement a). By condition $\mathbf{B}$ (see (6.13) and (6.14)), the coordinates $\mathcal{U}^{(1)}, \ldots, \mathcal{U}^{(d)}$ of the vectors $\mathcal{U} = \mathbb{J}_{N,1} U_{N,1}$ and $\mathcal{U}^j = \overline{\mathbb{P}}_j(\mathcal{U})$, $j = 1, \ldots, d$, are non-correlated and have mean zero and the same variance 2^N. The same is valid for the coordinates $\mathcal{V}^{(1)}, \ldots, \mathcal{V}^{(d)}$ of the vectors $\mathcal{V} = \mathbb{J}_{N,1} V_{N,1}$ and $\mathcal{V}^j = \overline{\mathbb{P}}_j(\mathcal{V})$. By Corollary 7.7,

$$\mathcal{L}(\mathcal{U}^j),\ \mathcal{L}(\mathcal{V}^j) \in \mathcal{A}_j^*(c_{30}\,\tau,\ 4,\ 4), \qquad j = 1, \ldots, d,$$

if $\tau \geqslant 1$ and $\tau^2 d^3 (p+1)/2^{N-p} \leqslant c_{29}$. Thus, the distributions $\mathcal{L}(\mathcal{U}^j)$, $\mathcal{L}(\mathcal{V}^j)$, satisfy in the j-dimensional case the conditions of Lemma 3.5 with $\operatorname{cov} \xi_k = \operatorname{cov} \mathcal{U}^j = 2^N \mathbb{I}_j$, $\mathbb{B} = \operatorname{cov} \mathcal{U}^{j-1} = 2^N \mathbb{I}_{j-1}$ (the last equality for $j \geqslant 2$) and with the change of τ by $c_{30}\,\tau$ provided that $c_{57} \leqslant c_{29}/2$.

Taking into account (5.14), (5.15) and applying Lemmas 3.5, 8.3 in the j-dimensional case, $j = 1, \dots, d,$ we obtain that

$$\left|\mathcal{U}^{(1)} - \mathcal{V}^{(1)}\right| \leqslant \frac{c\tau^2}{2^{N/2}}\left(1 + \frac{\left|\mathcal{U}^{(1)}\right|^3}{2^{3N/2}}\right), \tag{9.3}$$

provided that $\dfrac{\tau}{2^{N/2}} \leqslant c_{59}$, $\left|\mathcal{U}^{(1)}\right| \leqslant \dfrac{c_{59}\cdot 2^N}{\tau}$. Moreover, for $2 \leqslant j \leqslant d$,

$$\left|\mathcal{U}^{(j)} - \mathcal{V}^{(j)}\right| \leqslant \frac{c\tau^2}{2^{N/2}}\left(j^2 + \delta_{j-1}\left(j + \frac{\left|\mathcal{U}^{(j)} - y_{2,j}\right|}{2^{N/2}}\right) + \frac{\left|\mathcal{U}^{(j)} - y_{2,j}\right|^3}{2^{3N/2}}\right) + \left|y_{2,j} - y_{1,j}\right|, \tag{9.4}$$

provided that

$$\frac{\tau j^{3/2}}{2^{N/2}} \leqslant c_{59}, \quad \left|\mathcal{U}^{(j)} - y_{2,j}\right| \leqslant \frac{c_{59}\cdot 2^N}{j\tau}, \quad \frac{\max\{\left|\mathcal{U}^{j-1}\right|, \left|\mathcal{V}^{j-1}\right|\}}{2^{N/2}} \leqslant \frac{c_{59}\cdot 2^{N/2}}{j^{3/2}\tau}, \tag{9.5}$$

where

$$\left|y_{2,j}\right| \leqslant c\,\tau\,\frac{\left\|\mathcal{U}^{j-1}\right\|^2}{2^N}, \qquad \left|y_{2,j} - y_{1,j}\right| \leqslant c\left(\tau^2 2^{-N/2}\,\Gamma_{j-1}^3 + \tau\Lambda_{j-1}\Gamma_{j-1}\right), \tag{9.6}$$

$$\delta_{j-1} = \Gamma_{j-1}^2 + 2^{N/2}\tau^{-1}\Lambda_{j-1}, \tag{9.7}$$

$$\Lambda_{j-1} = \frac{\left\|\mathcal{U}^{j-1} - \mathcal{V}^{j-1}\right\|}{2^{N/2}}, \qquad \Gamma_{j-1} = \frac{\max\{\left\|\mathcal{U}^{j-1}\right\|, \left\|\mathcal{V}^{j-1}\right\|\}}{2^{N/2}}. \tag{9.8}$$

Below we shall use that, obviously,

$$j\cdot 2^{-N}\left\|\mathcal{U}^{j-1}\right\|^2 \leqslant j^2 + j^{1/2}\cdot 2^{-3N/2}\left\|\mathcal{U}^{j-1}\right\|^3, \quad \left\|\mathcal{V}^{j-1}\right\| \leqslant \left\|\mathcal{U}^{j-1}\right\| + \left\|\mathcal{U}^{j-1} - \mathcal{V}^{j-1}\right\|. \tag{9.9}$$

Using (8.9) and (9.3)–(9.9), we see that one can choose c_{60} to be so large and c_{61} to be so small that, for $1 \leqslant j \leqslant d$,

$$\left|\mathcal{U}^{(j)} - \mathcal{V}^{(j)}\right| \leqslant \frac{c_{60}\,\tau^2}{2^{N/2}}\left(d^2 + \frac{d^{1/2}\left\|\mathcal{U}^j\right\|^3}{2^{3N/2}}\right) + \frac{1}{2}\left|\mathcal{U}^{j-1} - \mathcal{V}^{j-1}\right| \tag{9.10}$$

provided that

$$\frac{\tau^2 d^3 p}{2^{N-p}} \leqslant c_{61}, \quad \left|\mathcal{U}^{(j)}\right| \leqslant \frac{c_{61}\cdot 2^N}{d\tau}, \quad \frac{\max\{\left|\mathcal{U}^{j-1}\right|, \left|\mathcal{V}^{j-1}\right|\}}{2^{N/2}} \leqslant \frac{c_{61}\cdot 2^{N/2}}{d^{3/2}\tau}, \tag{9.11}$$

where $\left|\mathcal{U}^0 - \mathcal{V}^0\right| = \left|\mathcal{V}^0\right| \overset{\text{def}}{=} 0.$

Using the induction with respect to j, we shall prove that, for $j = 1, \dots, d,$

$$\left|\mathcal{V}^{(j-1)}\right| \leqslant \frac{c_{61}\cdot 2^N}{d^{3/2}\tau}, \tag{9.12}$$

$$\left|\mathcal{U}^{(j)} - \mathcal{V}^{(j)}\right| \leqslant \frac{c_{62}\,\tau^2}{2\cdot 2^{N/2}}\left(d^2 + \frac{d^{1/2}\left\|\mathcal{U}^j\right\|^3}{2^{3N/2}}\right) + \frac{1}{2}\left|\mathcal{U}^{j-1} - \mathcal{V}^{j-1}\right|, \tag{9.13}$$

provided that

$$\bigl|\mathcal{U}\bigr| \leqslant \frac{c_{63}\cdot 2^N}{d^{3/2}\,\tau}, \qquad \frac{\tau^2\,d^3\,p}{2^{N-p}} \leqslant c_{64}, \tag{9.14}$$

where

$$c_{62}=2\,c_{60}, \quad c_{63}=\min\Bigl\{\frac{c_{61}}{6},\ \frac{1}{\sqrt{c_{62}}}\Bigr\}, \quad c_{64}=\min\Bigl\{c_{61},\Bigl(\frac{c_{63}}{c_{62}}\Bigr)^{2/3}\Bigr\}. \tag{9.15}$$

It is clear that (9.15) implies

$$c_{65} \overset{\text{def}}{=} c_{63}+\frac{c_{62}}{2}\bigl(c_{64}^{3/2}+c_{63}^3\bigr)+\frac{c_{63}}{2}+\frac{c_{61}}{2} \leqslant c_{61}. \tag{9.16}$$

Inequality (9.13) for $j=1$ follows directly from (9.10), (9.15), since the conditions (9.11) are in this case the consequences of (9.14) and (9.15). Inequality (9.12) for $j=1$ is valid automatically.

Suppose now that the inequalities (9.12), (9.13) are proved for $j=1,\dots,s-1$, $2\leqslant s\leqslant d$. We shall prove that these inequalities are valid for $j=s$. Using (8.9), (9.12) (with $j=1,\dots,s-1$), (9.13) (with $j=s-1$) and (9.14)–(9.16), we obtain

$$\bigl|\mathcal{V}^{(s-1)}\bigr| \leqslant \bigl|\mathcal{U}^{(s-1)}\bigr|+\bigl|\mathcal{U}^{(s-1)}-\mathcal{V}^{(s-1)}\bigr| \leqslant \frac{c_{65}\cdot 2^N}{d^{3/2}\tau} \leqslant \frac{c_{61}\cdot 2^N}{d^{3/2}\tau}. \tag{9.17}$$

Furthermore, it is clear that the inequality (9.13) for $j=s$ can be easily derived from (9.10), (9.15) since the validity of (9.11) follows in this case from (9.12) (with $j=1,\dots,s-1$), (9.14), (9.15), (9.17), (5.17), (5.18) and (8.9). This completes the proof of (9.12), (9.13) for $j=1,\dots,d$.

Taking into account (5.13), (5.17), (6.23) and (8.9) we see that (9.13) (considered for all $j=1,\dots,d$) implies

$$\bigl|\mathcal{U}-\mathcal{V}\bigr| \leqslant \frac{c_{62}\,\tau^2}{2^{N/2}}\Bigl(d^2+\frac{d^{1/2}\,\|\mathcal{U}\|^3}{2^{3N/2}}\Bigr)$$

and

$$\bigl|\mathcal{U}\bigr| \leqslant \bigl\|\mathcal{U}\bigr\| \leqslant C_1^{-1}\bigl\|U_{N,1}\bigr\|, \qquad \bigl\|U_{N,1}-V_{N,1}\bigr\| \leqslant C_2\,d^{1/2}\bigl|\mathcal{U}-\mathcal{V}\bigr|.$$

This completes the proof of assertion a) provided that c_{58} is chosen to be sufficienly large and c_{57} sufficienly small.

Let us prove now the assertion b). By condition **B** (see (6.28) and (6.29)), the coordinates $\mathbf{U}_{n,k}^{(1)},\dots,\mathbf{U}_{n,k}^{(2d)}$ of the random vectors $\mathbf{U}_{n,k}=\mathbb{A}_{n,k}\,\widetilde{\mathbf{U}}_{n,k}$ and $\mathbf{U}_{n,k}^j=\overline{\mathbb{P}}_j(\mathbf{U}_{n,k})$, $j=1,\dots,2\,d$, are non-correlated and have mean zero and the same variance 2^n. The same can be said about the coordinates $\mathbf{V}_{n,k}^{(1)},\dots,\mathbf{V}_{n,k}^{(2d)}$ of the vectors $\mathbf{V}_{n,k}=\mathbb{A}_{n,k}\,\widetilde{\mathbf{V}}_{n,k}$ and $\mathbf{V}_{n,k}^j=\overline{\mathbb{P}}_j(\mathbf{V}_{n,k})$. By Lemma 7.6,

$$\mathcal{L}(\mathbf{U}_{n,k}^j),\ \mathcal{L}(\mathbf{V}_{n,k}^j)\in\mathcal{A}_j^*(c_{30}\,\tau,\ 4,\ 4), \qquad j=1,\dots,2\,d,$$

provided that $\tau \geqslant 1$ and $\tau^2 d^3 (p+1)/2^{n-p} \leqslant c_{29}$. Thus, the distributions $\mathcal{L}(\mathbf{U}^j_{n,k})$, $\mathcal{L}(\mathbf{V}^j_{n,k})$, $j = d+1, \ldots, 2d$, satisfy in the j-dimensional situation the conditions of Lemma 3.5 with $\operatorname{cov} \xi_k = \operatorname{cov} \mathbf{U}^j_{n,k} = 2^n \mathbb{I}_j$, $\mathbb{B} = \operatorname{cov} \mathbf{U}^{j-1}_{n,k} = 2^n \mathbb{I}_{j-1}$ and with the change of τ by $c_{30}\tau$ provided that $c_{57} \leqslant c_{29}/2$.

Taking into account (5.20) and applying Lemmas 3.5, 8.3 in the j-dimensional case, $j = d+1, \ldots, 2d$, we obtain that

$$\bigl|\mathbf{U}^{(j)}_{n,k} - \mathbf{V}^{(j)}_{n,k}\bigr| \leqslant \frac{c\tau^2}{2^{n/2}} \left(j^2 + \delta_{j-1}\left(j + \frac{|\mathbf{U}^{(j)}_{n,k} - y_{2,j}|}{2^{n/2}}\right) + \frac{|\mathbf{U}^{(j)}_{n,k} - y_{2,j}|^3}{2^{3n/2}}\right) + |y_{2,j} - y_{1,j}|, \tag{9.18}$$

provided that

$$\frac{\tau j^{3/2}}{2^{n/2}} \leqslant c_{66}, \quad \bigl|\mathbf{U}^{(j)}_{n,k} - y_{2,j}\bigr| \leqslant \frac{c_{66}\cdot 2^n}{j\tau}, \quad \frac{\max\{|\mathbf{U}^{j-1}_{n,k}|, |\mathbf{V}^{j-1}_{n,k}|\}}{2^{n/2}} \leqslant \frac{c_{66}\cdot 2^{n/2}}{j^{3/2}\tau}, \tag{9.19}$$

where

$$|y_{2,j}| \leqslant c\tau \frac{\|\mathbf{U}^{j-1}_{n,k}\|^2}{2^n}, \qquad |y_{2,j} - y_{1,j}| \leqslant c\left(\tau^2 2^{-n/2}\Gamma^3_{j-1} + \tau\Lambda_{j-1}\Gamma_{j-1}\right), \tag{9.20}$$

$$\delta_{j-1} = \Gamma^2_{j-1} + 2^{n/2}\tau^{-1}\Lambda_{j-1}, \tag{9.21}$$

$$\Lambda_{j-1} = \frac{\|\mathbf{U}^{j-1}_{n,k} - \mathbf{V}^{j-1}_{n,k}\|}{2^{n/2}}, \qquad \Gamma_{j-1} = \frac{\max\{\|\mathbf{U}^{j-1}_{n,k}\|, \|\mathbf{V}^{j-1}_{n,k}\|\}}{2^{n/2}}. \tag{9.22}$$

Obviously,

$$j\cdot 2^{-n}\bigl\|\mathbf{U}^{j-1}_{n,k}\bigr\|^2 \leqslant j^2 + j^{1/2}\cdot 2^{-3n/2}\bigl\|\mathbf{U}^{j-1}_{n,k}\bigr\|^3, \quad \bigl\|\mathbf{V}^{j-1}_{n,k}\bigr\| \leqslant \bigl\|\mathbf{U}^{j-1}_{n,k}\bigr\| + \bigl\|\mathbf{U}^{j-1}_{n,k} - \mathbf{V}^{j-1}_{n,k}\bigr\|. \tag{9.23}$$

Using (8.16) and (9.18)–(9.23), we see that one can choose c_{67} to be so large and c_{68} to be so small that, for $j = d+1, \ldots, 2d$,

$$\bigl|\mathbf{U}^{(j)}_{n,k} - \mathbf{V}^{(j)}_{n,k}\bigr| \leqslant \frac{c_{67}\tau^2}{2^{n/2}}\left(d^2 + \frac{d^{1/2}\|\mathbf{U}^j_{n,k}\|^3}{2^{3n/2}}\right) + \frac{c_{69}}{\sqrt{d}}\bigl|\mathbf{U}^{j-1}_{n,k} - \mathbf{V}^{j-1}_{n,k}\bigr|, \tag{9.24}$$

where

$$c_{69} = \frac{\min\{C_1 c_{17}^{-1}, 1\}}{16}, \tag{9.25}$$

provided that

$$\frac{\tau^2 d^3 p}{2^{n-p}} \leqslant c_{68}, \quad \bigl|\mathbf{U}^{(j)}_{n,k}\bigr| \leqslant \frac{c_{68}\cdot 2^n}{d\tau}, \quad \frac{\max\{|\mathbf{U}^{j-1}_{n,k}|, |\mathbf{V}^{j-1}_{n,k}|\}}{2^{n/2}} \leqslant \frac{c_{68}\cdot 2^{n/2}}{d^{3/2}\tau}. \tag{9.26}$$

Using the induction with respect to j, we shall prove that, for $j = d+1, \dots, 2d$,

$$\left|\mathbf{V}_{n,k}^{(j-1)}\right| \leqslant \frac{c_{68} \cdot 2^n}{d^{3/2}\tau}, \tag{9.27}$$

$$\left|\mathbf{U}_{n,k}^{(j)} - \mathbf{V}_{n,k}^{(j)}\right| \leqslant \frac{c_{70}\,\tau^2}{2 \cdot 2^{n/2}} \left(d^2 + \frac{d^{1/2}\,\|\mathbf{U}_{n,k}^{j}\|^3}{2^{3n/2}}\right) + \frac{c_{69}}{\sqrt{d}}\left|\mathbf{U}_{n,k}^{j-1} - \mathbf{V}_{n,k}^{j-1}\right|, \tag{9.28}$$

provided that

$$\max\Big\{\left|\mathbf{U}_{n,k}\right|, \left|\mathcal{V}_{n,k}\right|\Big\} \leqslant \frac{c_{71} \cdot 2^n}{d^{3/2}\,\tau}, \qquad \frac{\tau^2\,d^3\,p}{2^{n-p}} \leqslant c_{72}, \tag{9.29}$$

where

$$c_{70} = 2\,c_{67}, \quad c_{71} = \min\left\{\frac{c_{68}}{4}, \frac{1}{\sqrt{c_{70}}}\right\}, \quad c_{72} = \min\left\{c_{68}, \left(\frac{c_{71}}{c_{70}}\right)^{2/3}\right\}. \tag{9.30}$$

It is clear that (9.30) implies

$$c_{73} \stackrel{\mathrm{def}}{=} c_{71} + \frac{c_{70}}{2}\,(c_{72}^{3/2} + c_{71}^3) + \frac{c_{71}}{8} + \frac{c_{68}}{16} \leqslant c_{68}. \tag{9.31}$$

Inequality (9.27) for $j = d+1$ immediately follows from (9.29) and (9.30) since $\mathbf{V}_{n,k}^{(d)} = \mathbb{P}_d\,\mathcal{V}_{n,k}$ (see (5.9), (6.7) and (6.8)). Inequality (9.28) for $j = d+1$ follows directly from (9.24) and (9.30), since the conditions (9.26) are in this case the consequences of (5.21), (6.8), (9.29) and (9.30).

Assume now that the inequalities (9.27), (9.28) are proved for $j = d+1, \dots, s-1$, $d+2 \leqslant s \leqslant 2d$. We shall prove that these inequalities are valid for $j = s$. Using (5.21), (6.8), (8.16), (9.25), (9.27) (with $j = d+1, \dots, s-1$), (9.28) (with $j = s-1$) and (9.29)–(9.31), we obtain

$$\left|\mathbf{V}_{n,k}^{(s-1)}\right| \leqslant \left|\mathbf{U}_{n,k}^{(s-1)}\right| + \left|\mathbf{U}_{n,k}^{(s-1)} - \mathbf{V}_{n,k}^{(s-1)}\right| \leqslant \frac{c_{73} \cdot 2^n}{d^{3/2}\tau} \leqslant \frac{c_{68} \cdot 2^n}{d^{3/2}\tau}. \tag{9.32}$$

Furthermore, it is clear that the inequality (9.28) for $j = s$ can be easily derived from (9.24), (9.30) since the validity of (9.26) follows in this case from (5.21), (8.16), (9.27) (with $j = d+1, \dots, s-1$), (9.29), (9.30) and (9.32). This completes the proof of (9.27), (9.28) for $j = d+1, \dots, 2d$.

Taking into account (5.21), (5.24), (6.4), (6.5), (6.7), (6.8), (6.23) and (9.25), we see that (9.28) (considered for all $j = d+1, \dots, 2d$) implies

$$\left|\widetilde{\mathcal{U}}_{n,k} - \widetilde{\mathcal{V}}_{n,k}\right| \leqslant \frac{c_{70}\,\tau^2}{2 \cdot 2^{n/2}} \left(d^2 + \frac{d^{1/2}\,\|\mathbf{U}_{n,k}\|^3}{2^{3n/2}}\right) + \frac{1}{16}\left|\widetilde{\mathcal{U}}_{n,k} - \widetilde{\mathcal{V}}_{n,k}\right| + \frac{C_1\,c_{17}^{-1}}{16\sqrt{d}}\left|\mathcal{U}_{n,k} - \mathcal{V}_{n,k}\right|$$

and

$$\left|\mathcal{V}_{n,k}\right| \leqslant C_1^{-1}\left\|V_{n,k}\right\|, \qquad \left|\mathcal{U}_{n,k} - \mathcal{V}_{n,k}\right| \leqslant C_1^{-1}\left\|U_{n,k} - V_{n,k}\right\|.$$

Together with trivial bound $\|x\| \leqslant \sqrt{d}\,|x|$, $x \in \mathbb{R}^d$, this completes the proof of assertion b) provided that c_{58} is chosen to be sufficiently large and c_{57} sufficiently small. □

10. Proof of Theorem 1.5

Let u, v, w be arbitrary numbers satisfying the relations

$$\frac{5}{8} \leqslant v < \frac{2}{3}\,, \qquad u = 1 - v, \qquad w = 2 - 3v. \tag{10.1}$$

It is clear that

$$\frac{1}{3} < u \leqslant \frac{3}{8}\,, \qquad 0 < w \leqslant \frac{1}{8}\,, \qquad 3u - w = 1. \tag{10.2}$$

Below the symbols $a_5, \dots, a_{11}$ will be used for denoting positive quantities depending only on v (and, of course, on C_1 and C_2). The letter a will denote different positive quantities depending only on v, C_1 and C_2 when we do not need to fix their numerical values. It is clear that, if v, C_1 and C_2 are fixed, $a, a_5, \dots, a_{11}$ are absolute constants. Write

$$c_{74} = \min\{c_{57}, 1\},$$
$$a_5 = \min\left\{\frac{c_{74}}{4}\,, \frac{c_{74}\,c_{16}}{4}\,, \frac{1}{(18\,c_{58}\,c_{17})^{1/2}}\,, \frac{c_{16}^{3/2}}{(65\,c_{58}\,c_{17})^{1/2}}\,, \frac{1-2^{-w}}{2}\,, C_2^2\right\}, \tag{10.3}$$
$$a_6 = \min\{c_{57},\, a_5^2,\, 1\}.$$

Recall that (see Lemma 6.7)

$$c_{17} \geqslant 1. \tag{10.4}$$

Let

$$\tau \geqslant 1, \qquad \lambda \geqslant \max\{C_2, 1\}\sqrt{d}. \tag{10.5}$$

Define p as the minimal integer satisfying the inequalities

$$p \geqslant 4, \qquad \frac{a_5 \cdot 2^{vp}}{d^{3/2}\,\tau} \geqslant 4\,\lambda. \tag{10.6}$$

Denote by M the maximal integer such that

$$\frac{\tau^2\,d^3\,p}{2^{M-p}} \geqslant a_6\ . \tag{10.7}$$

Let N be a positive integer such that

$$\frac{\tau^2\,d^3\,p}{2^{N-p}} < a_6\ , \qquad \frac{N-M}{p} \quad \text{is integer.} \tag{10.8}$$

Obviously, the relations (10.3), (10.5)–(10.8) imply

$$N > M \geqslant p, \qquad N \geqslant 2p. \tag{10.9}$$

Let $\widetilde{\xi}_1, \ldots, \widetilde{\xi}_{2^N}$ be random vectors with distributions $\mathcal{L}(\widetilde{\xi}_k) \in \mathcal{A}_d(\tau)$, $\mathbf{E}\,\widetilde{\xi}_k = 0$, $\operatorname{cov}\widetilde{\xi}_k = \mathbb{B}_k$, $k = 1, \ldots, 2^N$. Assume that the covariance operators $\mathbb{B}_k$ satisfy the conditions (6.2). Let $\widetilde{\eta}_k$ be random vectors which have the same moments of the first three orders as the vectors $\widetilde{\xi}_k$, and

$$\mathcal{L}(\widetilde{\eta}_k) \in \mathcal{A}_d(\tau), \qquad \mathbf{P}\Big\{ \big\| \widetilde{\eta}_k \big\| \leqslant \lambda \Big\} = 1, \qquad k = 1, \ldots, 2^N. \tag{10.10}$$

Suppose that the independent random vectors $\widetilde{Y}_k$, $k = 1, \ldots, 2^N$, with

$$\mathcal{L}(\widetilde{Y}_k) = \Psi\big(\mathcal{L}(\widetilde{\eta}_k)\big) \tag{10.11}$$

are already constructed (the distributions $\Psi(\cdot)$ were defined in Section 7). According to Lemma 7.3, $\widetilde{Y}_k$ and $\widetilde{\xi}_k$ have the same moments of the first three orders. Our aim is to construct independent random vectors $\widetilde{X}_k$, $k = 1, \ldots, 2^N$, with $\mathcal{L}(\widetilde{X}_k) = \mathcal{L}(\widetilde{\xi}_k)$ so that the sequential sums of $\widetilde{X}_k$ and $\widetilde{Y}_k$ would be close with large probability. To this end we define the random vectors

$$\Big\{ Y_k^{[N-lp]},\ Z_{(k-1)\cdot 2^{lp}+1}^{[N-lp]},\ \ 1 \leqslant k \leqslant 2^{N-lp},\ \ 0 \leqslant l \leqslant (N-M-p)/p \Big\} \tag{10.12}$$

with the help of a recurrent procedure described below. In one-dimensional case a very similar procedure was used by Sakhanenko (1984). Note also that the arguments of this section are close to those of Zaitsev (1995, Section 13) and Zaitsev (1998, Section 8).

At first, we set $Y_k^{[N]} = \widetilde{Y}_k$, $k = 1, \ldots, 2^N$. After this, using the dyadic procedure described in Sections 5–7 and 9 and satisfying Condition $\mathbf{B}$, we define the independent random vectors $Z_k^{[N]}$, $k = 1, \ldots, 2^N$, with distributions $\mathcal{L}\big(Z_r^{[N]}\big) = \Psi\big(\mathcal{L}(\widetilde{\eta}_r)\big)$ for $r = (m-1)\cdot 2^p + 1$, $m = 1, \ldots, 2^{N-p}$. For $r = 1, \ldots, 2^N$, $r \neq (m-1)\cdot 2^p + 1$, $m = 1, \ldots, 2^{N-p}$, we take $\mathcal{L}\big(Z_r^{[N]}\big) = \mathcal{L}(\widetilde{\xi}_r)$ and define $\widetilde{X}_r = Z_r^{[N]}$. Thus, the amount of vectors $\widetilde{X}_k$, which are not constructed, is decreased in 2^p times. If $N = M + p$, the first step of applying this procedure is simultaneously the last one, since the random vectors (10.12) with $l = (N-M-p)/p = 0$ are already constructed.

Suppose now that $N > M + p$. Then we repeat this procedure, working only with vectors which have the distributions $\Psi\big(\mathcal{L}(\widetilde{\eta}_r)\big)$ and diminishing their number after each step in 2^p times.

Namely, if the vectors $\Big\{ Y_k^{[N-lp+p]},\ Z_{(k-1)\cdot 2^{lp-p}+1}^{[N-lp+p]},\ 1 \leqslant k \leqslant 2^{N-lp+p} \Big\}$ are already defined for an integer l and $1 \leqslant l \leqslant (N-M-p)/p$, we set $Y_k^{[N-lp]} = Z_{(k-1)\cdot 2^{lp}+1}^{[N-lp+p]}$, $k = 1, \ldots, 2^{N-lp}$, and, using again the dyadic procedure described in Sections 5–7 and 9 and satisfying Condition $\mathbf{B}$ (with $N-lp$ instead of N), construct the independent random vectors $Z^{[N-lp]}_{(k-1)\cdot 2^{lp}+1}$, $k=1,\dots,2^{N-lp}$, with distributions written out below. If $r=(m-1)\cdot 2^{lp+p}+1$, $m=1,\dots,2^{N-lp-p}$, we choose $\mathcal{L}\Bigl(Z^{[N-lp]}_r\Bigr)=\Psi\bigl(\mathcal{L}(\widetilde{\eta}_r)\bigr)$. In the case if $r=(s-1)\cdot 2^{lp}+1$, $s=1,\dots,2^{N-lp}$, and $r\neq(m-1)\cdot 2^{lp+p}+1$, $m=1,\dots,2^{N-lp-p}$, we take $\mathcal{L}\Bigl(Z^{[N-lp]}_r\Bigr)=\mathcal{L}(\widetilde{\xi}_r)$ and define $\widetilde{X}_r=Z^{[N-lp]}_r$.

It is easy to see that for each l, $0\leqslant l\leqslant(N-M-p)/p$, the dyadic scheme satisfies Conditions $\mathbf{A}(\tau)$, $\mathbf{B}$, $\mathbf{D}_\Psi$ and $\mathbf{C}_\Psi(p)$ (with $N-lp$ instead of N). In the sequel this allows us to apply Lemma 9.1.

After the last step (with $l=(N-M-p)/p$) all independent random vectors $\widetilde{X}_k$, $k=1,\dots,2^N$, are constructed except $\widetilde{X}_r$ having indices of the form $r=(m-1)\cdot 2^{N-M}+1$, $m=1,\dots,2^M$. Instead of this we have constructed the vectors with distributions $\mathcal{L}\Bigl(Z^{[M+p]}_r\Bigr)=\Psi\bigl(\mathcal{L}(\widetilde{\eta}_r)\bigr)$ for $r=(m-1)\cdot 2^{N-M}+1$, $m=1,\dots,2^M$. Together with $\widetilde{X}_r$ these vectors form a collection of jointly independent vectors. This can be easily verified by induction, if we take into account that, in each step, the new collection of independent vectors is constructed as a function of the vectors which are independent of all already constructed vectors $\widetilde{X}_r$.

Now we can add the missing elements to the sequence of already constructed $\widetilde{X}_r$. For this we define the random vectors $Y^*_k=Y^{[M]}_k=Z^{[M+p]}_{(k-1)\cdot 2^{N-M}+1}$, $k=1,\dots,2^M$, and, using again the procedure described in Sections 5–7, construct the independent random vectors X^*_k, $k=1,\dots,2^M$, with $\mathcal{L}(X^*_k)=\mathcal{L}(\widetilde{\xi}_{(k-1)\cdot 2^{N-M}+1})$. Then we set $\widetilde{X}_{(k-1)\cdot 2^{N-M}+1}=X^*_k$. It is clear that the vectors X^*_k are jointly independent and they do not depend on $\widetilde{X}_r$ constructed earlier, since, by construction, X^*_k are functions of Y^*_k. This completes the procedure of constructing $\widetilde{X}_r$, $r=1,\dots,2^N$.

By analogy with arguments of Sakhanenko (1984), we shall prove that the procedure described above gives good approximation, using an induction with respect to N. It is clear that the procedure of constructing the independent random vectors $X_{*,k}=\widetilde{X}_{(k-1)\cdot 2^p+1}$, $k=1,\dots,2^{N-p}$, coincides with the procedure of constructing of $\widetilde{X}_k$, $k=1,\dots,2^N$. The only difference is in the change of N by $N-p$. Thus, below we shall assume that we have suitable approximation in the case when we construct 2^{N-p} summands and prove the analogous statement in the case of 2^N summands.

We shall use the notation which was already used in Sections 5–7 and 9 with $Y_k=Y^{[N]}_k$, $X_k=Z^{[N]}_k$, $k=1,\dots,2^N$. Let j be an integer satisfying $1\leqslant j\leqslant 2^N$. Write $S=S_j$, $T=T_j$ (see (5.7), (5.23)). Sometimes we shall omit for simplicity the index j, keeping in mind that it is fixed. By virtue of (6.2), (10.5), (10.10), (10.11), by construction and by the definition of distributions $\Psi(\cdot)$ (see (7.12)–(7.14)) we have

$$\mathbf{P}\left\{ \| Y_k\| \leqslant 4\lambda \right\} = 1, \qquad k = 1, \dots, 2^N. \tag{10.13}$$

It is clear that from (5.7), (10.13) it follows that

$$\mathbf{P}\left\{ \|T\| \leqslant 4\lambda \cdot 2^N \right\} = 1. \tag{10.14}$$

For fixed integer s, $0 \leqslant s \leqslant N$, define the integer $k_s = k_s(j)$ by the relation

$$(k_s - 1) \cdot 2^s < j \leqslant k_s \cdot 2^s. \tag{10.15}$$

Now we can represent the sums S and T in the form

$$S = S^{(s)} - S^{\{s\}}, \qquad T = T^{(s)} - T^{\{s\}}, \tag{10.16}$$

where

$$\begin{aligned} S^{(s)} &= S_{k_s \cdot 2^s}, \qquad & S^{\{s\}} &= S^{(s)} - S, \\ T^{(s)} &= T_{k_s \cdot 2^s}, \qquad & T^{\{s\}} &= T^{(s)} - T. \end{aligned} \tag{10.17}$$

According to Lemma 6.7,

$$\left\| S^{(s)} - T^{(s)} \right\| \leqslant c_{17} \left\| U_{N,1} - V_{N,1} \right\| + c_{17} \sum_{n=s+1}^{N} \left\| \widetilde{\mathcal{U}}^{(n)} - \widetilde{\mathcal{V}}^{(n)} \right\|, \tag{10.18}$$

where $\widetilde{\mathcal{U}}^{(n)} = \widetilde{\mathcal{U}}_{n,l_{n,j}}$, $\widetilde{\mathcal{V}}^{(n)} = \widetilde{\mathcal{V}}_{n,l_{n,j}}$, $\big(l_{n,j} - 1\big) \cdot 2^n < j \leqslant l_{n,j} \cdot 2^n$ (see (5.34)). Write as well, by analogy with the proof of Theorem 8.1,

$$\begin{aligned} U^{(n)} &= U_{n,l_{n,j}}, \qquad & V^{(n)} &= V_{n,l_{n,j}}, \qquad & U_{(n)} &= U_{n,\widetilde{l}_{n,j}} \\ \mathbf{U}^{(n)} &= \mathbf{U}_{n,l_{n,j}}, \qquad & \widetilde{\mathbf{U}}^{(n)} &= \widetilde{\mathbf{U}}_{n,l_{n,j}}, \qquad & \mathbb{A}^{(n)} &= \mathbb{A}_{n,l_{n,j}}, \end{aligned} \tag{10.19}$$

where

$$\widetilde{l}_{n-1,j} = \begin{cases} 2\,l_{n,j}, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j} - 1, \\ 2\,l_{n,j} - 1, & \text{if} \quad l_{n-1,j} = 2\,l_{n,j}, \end{cases} \qquad 0 \leqslant n \leqslant N. \tag{10.20}$$

As before, according to Remark 5.5, $l_{n,j}$ is the integer such that $U^{(n)}$, the sum over a block of 2^n summands, contains the summand X_j and, consequently, the summand $X_{k_s \cdot 2^s}$, for $s < n \leqslant N$ (see (10.15)). The sum $U_{(n)}$ does not contain X_j and, by (5.24), (6.38), (10.19),

$$\begin{aligned} U^{(n+1)} &= U^{(n)} + U_{(n)}, \qquad & \left\| \widetilde{\mathbf{U}}^{(n+1)} \right\|^2 &= \left\| U^{(n)} \right\|^2 + \left\| U_{(n)} \right\|^2, \\ \mathbf{U}^{(n+1)} &= \mathbb{A}^{(n+1)}\, \widetilde{\mathbf{U}}^{(n+1)}, \qquad & \left\| \mathbb{A}^{(n+1)} \right\| &\leqslant c_{16}^{-1}, \end{aligned} \qquad 0 \leqslant n < N. \tag{10.21}$$

From (10.21) it follows that

$$\| U_{(n)} \| = \| U^{(n+1)} - U^{(n)} \| \leqslant \| U^{(n+1)} \| + \| U^{(n)} \| \tag{10.22}$$

and, hence,

$$\begin{aligned} &\| \widetilde{\mathbf{U}}^{(n+1)} \| \leqslant 2 \max\left\{ \| U^{(n)} \|,\ \| U_{(n)} \| \right\} \leqslant 2 \| U^{(n+1)} \| + 2 \| U^{(n)} \|, \\ &\| \mathbf{U}^{(n+1)} \| \leqslant 2 c_{16}^{-1} \left(\| U^{(n+1)} \| + \| U^{(n)} \| \right). \end{aligned} \tag{10.23}$$

It is evident that $U^{(N)} = U_{N,1}$, $V^{(N)} = V_{N,1}$. Thus, from (10.16), (10.18) it follows that

$$\| S - T \| \leqslant c_{17} \left\| U^{(N)} - V^{(N)} \right\| + c_{17} \sum_{n=s+1}^{N} \left\| \widetilde{\mathcal{U}}^{(n)} - \widetilde{\mathcal{V}}^{(n)} \right\| + \| S^{\{s\}} \| + \| T^{\{s\}} \|. \tag{10.24}$$

Introduce the random integer $\mu = \mu_j$ by $\mu = N+1$, if at least one of the following inequalities is valid:

$$\frac{\tau^2 d^3 p}{2^{N-p}} > a_6, \qquad \| S^{\{N\}} \| > \frac{2 a_5 \cdot 2^N}{d^{3/2} \tau}, \qquad \| U^{(N)} \| > \frac{2 a_5 \cdot 2^N}{d^{3/2} \tau} \tag{10.25}$$

and by

$$\begin{aligned} \mu = \min\Big\{ m :\ p \leqslant m \leqslant N,\ \ &\frac{\tau^2 d^3 p}{2^{m-p}} \leqslant a_6; \\ &\| S^{\{n\}} \| \leqslant \frac{2 a_5 \cdot 2^n}{d^{3/2} \tau},\ \ \| U^{(n)} \| \leqslant \frac{2 a_5 \cdot 2^{um+vn}}{d^{3/2} \tau},\ n = m, \dots, N \Big\} \end{aligned} \tag{10.26}$$

otherwise. Here the numbers u, v are defined by (10.1). Note that the first of inequalities (10.25) is not satisfied in our conditions (see (10.8)). The same can be said about inequalities (10.41), (10.45).

Write

$$Y_{*,k} = Y_k^{[N-p]} = Z_{(k-1)\cdot 2^p + 1}^{[N]}, \qquad k = 1, \dots, 2^{N-p}. \tag{10.27}$$

By construction and by (6.2), (7.12)–(7.14), (10.5), (10.10), (10.11) (cf. (10.13)), we have

$$\mathbf{P}\left\{ \| Y_{*,k} \| \leqslant 4\lambda \right\} = 1, \qquad k = 1, \dots, 2^{N-p}. \tag{10.28}$$

Let $i = i(j)$ be the largest integer such that $(i-1) \cdot 2^p + 1 \leqslant j$. We shall denote by $V_*^{(n)}$, $n = 0, \dots, N-p$, the sum of a block of 2^n summands $Y_{*,k}$ which contains the summand $Y_{*,i}$. Let

$$T_* = T_{*,i} = Y_{*,1} + \dots + Y_{*,i}. \tag{10.29}$$

Recall that, by construction,

$$\bigcup_{k=1}^{2^{N-p}} \Big\{\big\{\widetilde{X}_r = Z_r^{[N]},\ r=(k-1)\cdot 2^p+2,\dots,k\cdot 2^p\big\}\cup\big\{Y_{*,k}=Z^{[N]}_{(k-1)\cdot 2^p+1}\big\}\Big\} \tag{10.30}$$

is a collection of random vectors, which are jointly independent. Set

$$U_0^{(n)} = U^{(n)} - V_*^{(n-p)}, \qquad n=p,\dots,N. \tag{10.31}$$

It can be easily verified that the sum $U_0^{(n)}$ consists of the summands of the sum $U^{(n)}$ with indices which can not be represented in the form $r=(k-1)\cdot 2^p+1$ with integer k. In other words, the sum $U_0^{(n)}$ is the sum of contained in $U^{(n)}$ summands $Z_r^{[N]}$ which are chosen as $\widetilde{X}_r$. For the proof of this fact one should take into account that, by construction, the numbers j and $(i-1)\cdot 2^p+1$ are contained in the same block of indices $\big\{m:\ m=(i-1)\cdot 2^p+1,\dots,i\cdot 2^p\big\}$. Recall that, according to our construction, after the first step, when the independent random vectors (10.30) are constructed, we define the rest of the random vectors $\widetilde{X}_r$ as the (d-dimensional) functions of the random vectors $Y_{*,k}$, $k=1,\dots,2^{N-p}$. Write

$$\begin{aligned} X_{*,k} &= \widetilde{X}_{(k-1)\cdot 2^p+1}, \qquad k=1,\dots,2^{N-p};\\ S_* = S_{*,i} &= X_{*,1}+\dots+X_{*,i}\,, \qquad S_0 = S - T_*\,. \end{aligned} \tag{10.32}$$

By analogy with (10.16), (10.17), for $s=p,\dots,N$ we can represent the sums S_*, S_0 and T_* in the form

$$S_0 = S_0^{(s)} - S_0^{\{s\}}, \qquad T_* = T_*^{(s-p)} - T_*^{\{s-p\}}, \tag{10.33}$$

where

$$\begin{aligned} T_*^{(s-p)} &= T_{*,k_s\cdot 2^{s-p}}, & T_*^{\{s-p\}} &= T_*^{(s-p)} - T_*,\\ S_0^{(s)} &= S^{(s)} - T_*^{(s-p)}, & S_0^{\{s\}} &= S_0^{(s)} - S_0 = S^{\{s\}} - T_*^{\{s-p\}}, \end{aligned} \tag{10.34}$$

and the numbers k_s are defined by (10.15).

From the definition of numbers i,j,k_s, $s=0,\dots,N$, it follows that

$$\begin{aligned} 0 &\leqslant k_s\cdot 2^s - j < 2^s, & s&=0,\dots,N;\\ 0 &\leqslant k_s\cdot 2^{s-p} - i < 2^{s-p}, & s&=p,\dots,N. \end{aligned} \tag{10.35}$$

It is easy to see that from (5.7), (10.13), (10.16), (10.17), (10.28), (10.29), (10.34), (10.35) one can derive that, with probability one,

$$\begin{aligned} \big\|T^{\{s\}}\big\| &\leqslant 4\lambda\cdot 2^s, & s&=0,\dots,N;\\ \big\|T_*^{\{s-p\}}\big\| &\leqslant 4\lambda\cdot 2^{s-p}, & s&=p,\dots,N. \end{aligned} \tag{10.36}$$

Similarly, using (10.28) and the definition of $V_*^{(n)}$, we see that

$$\left\| V_*^{(n)} \right\| \leqslant 4\lambda \cdot 2^n \quad \text{a.s.}, \qquad n = 0, \dots, N-p. \tag{10.37}$$

Define also

$$\begin{aligned} \widetilde{S} = \sum_{k=1}^{j} \widetilde{X}_k, \qquad \delta_j(N) &= \left\| \sum_{k=1}^{j} \widetilde{Y}_k - \sum_{k=1}^{j} \widetilde{X}_k \right\| = \left\| T - \widetilde{S} \right\|, \\ \delta_i(N-p) &= \left\| \sum_{k=1}^{i} Y_{*,k} - \sum_{k=1}^{i} X_{*,k} \right\| = \left\| T_* - S_* \right\|, \\ \psi_j(N) &= \left\| \sum_{k=1}^{j} \widetilde{Y}_k - \sum_{k=1}^{j} Z_k^{[N]} \right\| = \left\| T - S \right\| \end{aligned} \tag{10.38}$$

(see (5.7), (5.23), (10.29), (10.32)). It is clear that from (10.27), (10.29), (10.30), (10.32), (10.38) and from the definition of i and S it follows that

$$S = \sum_{k=1}^{j} Z_k^{[N]} = S_0 + T_*, \qquad \widetilde{S} = S_0 + S_*, \qquad S - \widetilde{S} = T_* - S_*. \tag{10.39}$$

It is easy to see that the relations (10.38), (10.39) imply

$$\delta_j(N) \leqslant \psi_j(N) + \delta_i(N-p). \tag{10.40}$$

Introduce the random integer $\beta_j(N)$ by $\beta_j(N) = N+1$, if at least one of the following inequalities is valid:

$$\frac{\tau^2 d^3 p}{2^{N-p}} > a_6, \qquad \left\| V^{(N)} \right\| > \frac{4a_5 \cdot 2^N}{d^{3/2}\tau} \tag{10.41}$$

and by

$$\begin{aligned} \beta_j(N) = \min\Big\{ m : \; p \leqslant m \leqslant N, \quad \frac{\tau^2 d^3 p}{2^{m-p}} \leqslant a_6, \\ \left\| V^{(n)} \right\| \leqslant \frac{4a_5 \cdot 2^{um+vn}}{d^{3/2}\tau}, \quad n = m, \dots, N \Big\} \end{aligned} \tag{10.42}$$

otherwise.

Similarly, taking into account (10.9), we define the random integer $\beta_i(N-p)$ by the equality $\beta_i(N-p) = N-p+1$, if at least one of the following inequalities is valid:

$$\frac{\tau^2 d^3 p}{2^{N-2p}} > a_6, \qquad \left\| V_*^{(N-p)} \right\| > \frac{4a_5 \cdot 2^{N-p}}{d^{3/2}\tau} \tag{10.43}$$

and by

$$\beta_i(N-p) = \min\Big\{ m:\ p \leqslant m \leqslant N-p,\quad \frac{\tau^2 d^3 p}{2^{m-p}} \leqslant a_6, \\ \big\| V_*^{(n)} \big\| \leqslant \frac{4\, a_5 \cdot 2^{um+vn}}{d^{3/2}\,\tau},\quad n = m, \ldots, N-p \Big\} \tag{10.44}$$

otherwise.

Finally, define the random integer $\nu = \nu_j$ by $\nu = N+1$, if at least one of the following inequalities is valid:

$$\frac{\tau^2 d^3 p}{2^{N-p}} > a_6, \qquad \big\| U_0^{(N)} \big\| > \frac{a_5 \cdot 2^N}{d^{3/2}\,\tau}, \qquad \big\| S_0^{\{N\}} \big\| > \frac{a_5 \cdot 2^N}{d^{3/2}\,\tau} \tag{10.45}$$

and by

$$\nu = \min\Big\{ m:\ p \leqslant m \leqslant N,\quad \frac{\tau^2 d^3 p}{2^{m-p}} \leqslant a_6; \\ \big\| S_0^{\{n\}} \big\| \leqslant \frac{a_5 \cdot 2^n}{d^{3/2}\,\tau},\quad \big\| U_0^{(n)} \big\| \leqslant \frac{a_5 \cdot 2^{um+vn}}{d^{3/2}\,\tau},\ n = m, \ldots, N \Big\} \tag{10.46}$$

otherwise.

Put now

$$\rho_j(N) = 5\,\lambda \cdot 2^{\beta_j(N)}, \qquad \rho_i(N-p) = 5\,\lambda \cdot 2^{\beta_i(N-p)}, \tag{10.47}$$

$$\gamma_j(N) = \begin{cases} 10\,\lambda \cdot 2^{\nu}, & \text{if} \quad \nu \leqslant N, \\ U + 22\,\lambda \cdot 2^N, & \text{if} \quad \nu = N+1, \end{cases} \tag{10.48}$$

where

$$U = \big\| U_0^{(N)} \big\| + \big\| S_0^{\{N\}} \big\|. \tag{10.49}$$

Lemma 10.1. *The following inequalities are valid*:

$$\delta_j(N) + \rho_j(N) \leqslant \delta_i(N-p) + \rho_i(N-p) + \gamma_j(N) \tag{10.50}$$

and, for each $\varepsilon > 0$,

$$\mathbf{E} \exp\Big(\varepsilon\big(\delta_j(N) + \rho_j(N)\big)\Big) \leqslant \mathbf{E} \exp\Big(\varepsilon\big(\delta_i(N-p) + \rho_i(N-p)\big)\Big)\, \mathbf{E} \exp\big(\varepsilon\, \gamma_j(N)\big). \tag{10.51}$$

Proof. Assume that $\nu \leqslant N$ and, taking into account (10.9), define the integer valued random variables

$$\beta = \beta_i(N-p) - 1 \leqslant N-p, \qquad \varkappa = \max\big\{ \beta,\ \nu \big\} \leqslant N. \tag{10.52}$$

The relations (10.1), (10.6), (10.36) imply

$$\left\| T_*^{\{s-p\}} \right\| \leqslant \frac{a_5 \cdot 2^s}{d^{3/2}\,\tau} \quad \text{a.s.}, \qquad s = p, \dots, N. \tag{10.53}$$

Using (10.34), (10.46), (10.52), (10.53), we obtain

$$\left\| S^{\{n\}} \right\| \leqslant \left\| S_0^{\{n\}} \right\| + \left\| T_*^{\{n-p\}} \right\| \leqslant \frac{2\,a_5 \cdot 2^n}{d^{3/2}\,\tau}, \qquad n = \varkappa, \dots, N, \tag{10.54}$$

(here and below we use that (10.46), (10.52) imply the inequality $\varkappa \geqslant p$). By virtue of (10.1), (10.6), (10.37), (10.52), for $\beta \leqslant n \leqslant \beta + p$, $n \geqslant p$, we have

$$\left\| V_*^{(n-p)} \right\| \leqslant 4\lambda \cdot 2^{n-p} \leqslant \frac{a_5 \cdot 2^{u\beta+vn}}{d^{3/2}\,\tau} \leqslant \frac{a_5 \cdot 2^{u\varkappa+vn}}{d^{3/2}\,\tau}. \tag{10.55}$$

If $\beta + p + 1 \leqslant n \leqslant N$, then, applying (10.1), (10.2), (10.6), (10.44), (10.52), we obtain

$$\left\| V_*^{(n-p)} \right\| \leqslant \frac{4\,a_5 \cdot 2^{u(\beta+1)+v(n-p)}}{d^{3/2}\,\tau} \leqslant \frac{a_5 \cdot 2^{u\varkappa+vn}}{d^{3/2}\,\tau} \tag{10.56}$$

(we use that the inequalities $p \geqslant 4$, (10.1), (10.2) imply $2^{vp} \geqslant 2^{5p/8} > 4{\cdot}2^{3/8} \geqslant 4{\cdot}2^u$). From (10.31), (10.46), (10.52), (10.55), (10.56) it follows that

$$\begin{aligned} \left\| U^{(n)} \right\| \leqslant \left\| U_0^{(n)} \right\| + \left\| V_*^{(n-p)} \right\| &\leqslant \frac{a_5 \cdot 2^{u\varkappa+vn}}{d^{3/2}\,\tau} + \frac{a_5 \cdot 2^{u\varkappa+vn}}{d^{3/2}\,\tau} \\ &= \frac{2\,a_5 \cdot 2^{u\varkappa+vn}}{d^{3/2}\,\tau}, \qquad n = \varkappa, \dots, N. \end{aligned} \tag{10.57}$$

Taking into account (10.26), (10.46), (10.52), (10.54), (10.57), we see that

$$\mu \leqslant \varkappa \leqslant N. \tag{10.58}$$

Suppose now that an integer n satisfies the inequalities

$$\mu \leqslant n \leqslant N. \tag{10.59}$$

Then, taking into account (10.1)–(10.3), (10.6), (10.26), (10.59), we see that

$$\tau^2\, d^3 \leqslant a_6 \cdot 2^{\mu} \leqslant a_5^2 \cdot 2^{2u\mu+(2v-1)n} \leqslant a_5^2 \cdot 2^{2u\mu+(2v-1)N}. \tag{10.60}$$

Using (10.1), (10.3), (10.26), (10.59), we obtain the inequalities

$$\left\| U^{(n)} \right\| \leqslant \frac{2\,a_5 \cdot 2^{u\mu+vn}}{d^{3/2}\,\tau} \leqslant \frac{c_{74} \cdot 2^n}{d^{3/2}\,\tau} \leqslant \frac{c_{57} \cdot 2^n}{d^{3/2}\,\tau} \tag{10.61}$$

Taking here $n = N$ and using (10.1)–(10.4), (10.19), (10.26), (10.59)–(10.61) and Lemma 9.1(a), we obtain

$$\begin{aligned}\big\|U^{(N)} - V^{(N)}\big\| &\leqslant c_{58}\,\tau^2\cdot 2^{-N/2}\left(d^{5/2} + d\cdot 2^{-3N/2}\big\|U^{(N)}\big\|^3\right)\\ &\leqslant c_{58}\,d^{5/2}\,\tau^2\left(2^{-N/2} + \frac{8\,a_5^3\cdot 2^{3u\mu-wN}}{d^5\,\tau^3}\right)\\ &\leqslant \frac{9\,a_5^3\,c_{58}\cdot 2^{3u\mu-wN}}{d^2\,\tau} \leqslant \frac{a_5\cdot 2^{3u\mu-wN}}{2\,c_{17}\,d^2\,\tau} \leqslant \frac{a_5\cdot 2^{3u\mu-wN}}{2\,d^2\,\tau}.\end{aligned} \tag{10.62}$$

From (10.1), (10.59), (10.61), (10.62) it follows that

$$\big\|V^{(N)}\big\| \leqslant \big\|U^{(N)}\big\| + \big\|V^{(N)} - U^{(N)}\big\| \leqslant \frac{4\,a_5\cdot 2^{u\mu+vN}}{d^{3/2}\,\tau}. \tag{10.63}$$

Similarly, if we suppose that

$$\mu \leqslant n \leqslant N-1, \tag{10.64}$$

then from (10.1)–(10.3), (10.23), (10.26), (10.64) one can easily derive that

$$\big\|\mathbf{U}^{(n+1)}\big\| \leqslant \frac{4\,c_{16}^{-1}\,a_5\cdot 2^{u\mu+v(n+1)}}{d^{3/2}\,\tau} \leqslant \frac{c_{57}\cdot 2^{n+1}}{d^{3/2}\,\tau}. \tag{10.65}$$

Using the induction with respect to n, we shall prove that, if $\mu < N$, then, for $n = N-1,\, N-2,\, \dots,\, \mu$,

$$\big\|\widetilde{\mathcal{U}}^{(n+1)} - \widetilde{\mathcal{V}}^{(n+1)}\big\| \leqslant \frac{2\,a_5\cdot 2^{3u\mu-w(n+1)}}{c_{17}\,d^2\,\tau}, \tag{10.66}$$

$$\big\|U^{(n)} - V^{(n)}\big\| \leqslant \frac{2\,a_5\cdot 2^{3u\mu-wn}}{d^2\,\tau}, \tag{10.67}$$

$$\big\|V^{(n)}\big\| \leqslant \frac{4\,a_5\cdot 2^{u\mu+vn}}{d^{3/2}\,\tau}. \tag{10.68}$$

For $n = N$ the relations (10.67), (10.68) are already proved (see (10.62) and (10.63)).

Assume that the inequalities (10.67), (10.68) are valid for $n = s+1,\, \dots,\, N$ and the inequality (10.66) holds for $n = s+1,\, \dots,\, N-1$, $\mu \leqslant s \leqslant N-1$. It suffices to prove the relations (10.66)–(10.68) for $n = s$.

Using (10.1)–(10.3), (10.19), (10.26), (10.60), (10.64), (10.65), (10.67), (10.68) and Lemma 9.1(b), we obtain that in this case

$$\begin{aligned}\big\|\widetilde{\mathcal{U}}^{(s+1)} - \widetilde{\mathcal{V}}^{(s+1)}\big\| &\leqslant \frac{c_{58}\,d^{5/2}\,\tau^2}{2^{(s+1)/2}}\left(1 + \frac{\big\|\mathbf{U}^{(s+1)}\big\|^3}{2^{3(s+1)/2}\,d^{1/2}}\right) + \frac{\big\|U^{(s+1)} - V^{(s+1)}\big\|}{8\,c_{17}}\\ &\leqslant c_{58}\,d^{5/2}\,\tau^2\left(2^{-(s+1)/2} + \frac{64\,a_5^3\cdot 2^{3u\mu-w(s+1)}}{c_{16}^3\,d^5\,\tau^3}\right) + \frac{a_5\cdot 2^{3u\mu-w(s+1)}}{4\,c_{17}\,d^2\,\tau}\\ &\leqslant \frac{65\,c_{58}\,a_5^3\cdot 2^{3u\mu-w(s+1)}}{c_{16}^3\,d^2\,\tau} + \frac{a_5\cdot 2^{3u\mu-w(s+1)}}{4\,c_{17}\,d^2\,\tau}\\ &\leqslant \frac{2\,a_5\cdot 2^{3u\mu-w(s+1)}}{c_{17}\,d^2\,\tau}.\end{aligned} \tag{10.69}$$

The fulfilment of the conditions of Lemma 9.1(b) can be easily deduced from the relations (10.1), (10.19), (10.26), (10.64), (10.65), (10.68) (with $n = s+1$) and from the definitions of a_5, a_6, c_{74} (see (10.3)). The bound (10.69) yields (10.66) for $n = s$.

Applying Lemma 6.7 (see (6.41), (6.43) and (6.56)) and using (10.1), (10.2), (10.61), (10.62), (10.66) (with $n = s, \dots, N-1$), we get

$$\left\| U^{(s)} - V^{(s)} \right\| \leqslant \frac{c_{17} \left\| U^{(N)} - V^{(N)} \right\|}{2^{N-s}} + \sum_{k=s+1}^{N} \frac{c_{17} \left\| \widetilde{\mathcal{U}}^{(k)} - \widetilde{\mathcal{V}}^{(k)} \right\|}{2^{k-s}} \leqslant \frac{2\, a_5 \cdot 2^{3u\mu - ws}}{d^2\, \tau}$$

and, consequently,

$$\left\| V^{(s)} \right\| \leqslant \left\| U^{(s)} \right\| + \left\| U^{(s)} - V^{(s)} \right\| \leqslant \frac{2\, a_5 \cdot 2^{u\mu + vs}}{d^{3/2}\, \tau} + \frac{2\, a_5 \cdot 2^{3u\mu - ws}}{d^2\, \tau} \leqslant \frac{4\, a_5 \cdot 2^{u\mu + vs}}{d^{3/2}\, \tau}.$$

This completes the proof of (10.66)–(10.68) for $n = s$.

It is evident that from (10.26), (10.42), (10.63), (10.68) it follows that

$$\beta_j(N) \leqslant \mu. \tag{10.70}$$

It is clear that the relations (10.52), (10.58), (10.70) imply

$$2^{\beta_j(N)} + 2^{\mu} \leqslant 2 \cdot 2^{\mu} \leqslant 2 \cdot 2^{\varkappa} \leqslant 2^{\beta_i(N-p)} + 2 \cdot 2^{\nu}. \tag{10.71}$$

Using (10.2), (10.3), (10.5), (10.24) (with $s = \mu$), (10.26), (10.36), (10.38), (10.62), (10.66), we see that

$$\begin{aligned} \psi_j(N) &\leqslant \frac{a_5 \cdot 2^{3u\mu - wN}}{2\, d^2\, \tau} + \sum_{n=\mu+1}^{N} \frac{2\, a_5 \cdot 2^{3u\mu - wn}}{d^2\, \tau} + \frac{2\, a_5 \cdot 2^{\mu}}{d^{3/2}\, \tau} + 4\,\lambda \cdot 2^{\mu} \\ &\leqslant \bigl(2\,(1 - 2^{-w})^{-1}\, a_5 + 4\,\lambda\bigr) \cdot 2^{\mu} \leqslant 5\,\lambda \cdot 2^{\mu}. \end{aligned} \tag{10.72}$$

The relations (10.40), (10.47), (10.71), (10.72) imply

$$\delta_j(N) + \rho_j(N) \leqslant \delta_i(N-p) + 5\,\lambda \left(2^{\mu} + 2^{\beta_j(N)}\right) \leqslant \delta_i(N-p) + \rho_i(N-p) + 10\,\lambda \cdot 2^{\nu}. \tag{10.73}$$

From (10.48), (10.73) it is easy to derive the inequality (10.50) in the case $\nu \leqslant N$.

Let now $\nu = N+1$. From (10.15), (10.17), (10.31) it follows that

$$S^{(N)} = U^{(N)} = U_0^{(N)} + V_*^{(N-p)}. \tag{10.74}$$

Using (10.16), (10.38), we see that

$$\psi_j(N) \leqslant \left\| S^{(N)} \right\| + \left\| S^{\{N\}} \right\| + \left\| T \right\|. \tag{10.75}$$

By virtue of (10.34) we have

$$\| S^{\{N\}} \| \leqslant \| S_0^{\{N\}} \| + \| T_*^{\{N-p\}} \|. \tag{10.76}$$

It is clear that the relations (10.14), (10.36), (10.37), (10.49), (10.74)–(10.76) imply

$$\psi_j(N) \leqslant U + 12\,\lambda \cdot 2^N. \tag{10.77}$$

From (10.47) and from definition of $\beta_j(N)$ it follows that $\beta_j(N) \leqslant N+1$ and

$$\rho_j(N) \leqslant 10\,\lambda \cdot 2^N. \tag{10.78}$$

The relations (10.40), (10.48), (10.77), (10.78) imply

$$\delta_j(N) + \rho_j(N) \leqslant \delta_i(N-p) + U + 22\,\lambda \cdot 2^N = \delta_i(N-p) + \gamma_j(N). \tag{10.79}$$

Thus, the inequality (10.50) is proved in the case $\nu = N+1$.

By construction (see (10.29)–(10.34), (10.38), (10.43)–(10.48)), the random variable $\gamma_j(N)$ is a function of

$$\bigcup_{k=1}^{2^{N-p}} \Big\{ \widetilde{X}_r = Z_r^{[N]},\ r = (k-1)\cdot 2^p + 2, \dots, k \cdot 2^p \Big\} \tag{10.80}$$

and the random variables $\delta_i(N-p)$, $\rho_i(N-p)$ are functions of

$$\Big\{ Y_{*,k} = Z_{(k-1)\cdot 2^p + 1}^{[N]},\ k = 1, \dots, 2^{N-p} \Big\}. \tag{10.81}$$

Since the sets (10.80), (10.81) are disjoint sets consisting of jointly independent vectors (see (10.30)), the random variables $\gamma_j(N)$ and $\delta_i(N-p) + \rho_i(N-p)$ are independent. Therefore, the inequality (10.51) follows from (10.50) and the proof of Lemma 10.1 is completed. $\square$

In this section we shall denote by ε the quantity

$$\varepsilon = \frac{a_5^2}{512\,\lambda\, C_2^2\, d^4\, \tau^2}\,. \tag{10.82}$$

Lemma 10.2. *There exists a positive quantity a_7 depending only on v and such that the following inequality is valid*:

$$\mathbf{E}\,\exp\bigl(\varepsilon\,\gamma_j(N)\bigr)\leqslant\exp\Bigl(a_7\,p\,(\lambda\,d^{3/2}\,\tau)^{1/v}\Bigr). \tag{10.83}$$

Proof. Denote by

$$\zeta=\mathbf{I}\bigl\{\,\nu=N+1\,\bigr\} \tag{10.84}$$

the indicator function of the event $\bigl\{\,\nu=N+1\,\bigr\}$ (see (10.45)). Using (10.48) and the Hölder inequality, we obtain

$$\mathbf{E}\,\zeta\,\exp\bigl(\varepsilon\,\gamma_j(N)\bigr)\leqslant\bigl(\mathbf{E}\,\zeta^2\bigr)^{1/2}\,\exp\bigl(22\,\varepsilon\,\lambda\cdot 2^N\bigr)\bigl(\mathbf{E}\,e^{2\varepsilon U}\bigr)^{1/2}. \tag{10.85}$$

From the definitions of $U_0^{(n)}$, $S_0^{\{n\}}$ it follows that $U_0^{(N)}$, $S_0^{\{N\}}$ are sums of not more than 2^N independent random vectors with distributions $\mathcal{L}(\widetilde{X}_k)\in\mathcal{A}_d(\tau)$ and covariance operators $\operatorname{cov}\widetilde{X}_k=\mathbb{B}_k$ satisfying the condition (6.2) (see (10.31), (10.34)). Therefore, the relations (6.2), (6.16), (10.3), (10.8), (10.45), (10.84) and Lemma 2.7 with $B^2\leqslant C_2^2\cdot 2^N$ imply

$$\begin{aligned}\mathbf{E}\,\zeta^2&\leqslant\mathbf{P}\Bigl\{\bigl|U_0^{(N)}\bigr|>\frac{a_5\cdot 2^N}{d^2\,\tau}\Bigr\}+\mathbf{P}\Bigl\{\bigl|S_0^{\{N\}}\bigr|>\frac{a_5\cdot 2^N}{d^2\,\tau}\Bigr\}\\&\leqslant 4\,d\,\exp\Bigl(-\min\Bigl\{\frac{a_5^2\cdot 2^N}{4\,C_2^2\,d^4\,\tau^2},\ \frac{a_5\cdot 2^N}{4\,d^2\,\tau^2}\Bigr\}\Bigr)\\&=4\,d\,\exp\Bigl(-\frac{a_5^2\cdot 2^N}{4\,C_2^2\,d^4\,\tau^2}\Bigr).\end{aligned} \tag{10.86}$$

It is clear that (see (10.49))

$$\mathbf{E}\,e^{2\varepsilon U}\leqslant\Bigl(\mathbf{E}\,\exp\bigl(4\varepsilon\,\bigl\|U_0^{(N)}\bigr\|\bigr)\,\mathbf{E}\,\exp\bigl(4\varepsilon\,\bigl\|S_0^{\{N\}}\bigr\|\bigr)\Bigr)^{1/2}. \tag{10.87}$$

By virtue of (10.3), (10.5), (10.82) we have

$$4\,d^{1/2}\,\varepsilon\,\tau<1. \tag{10.88}$$

Using (10.88) and applying Lemma 2.7 with $B^2\leqslant C_2^2\cdot 2^N$, $h=4\,d^{1/2}\varepsilon$, we see that

$$\begin{aligned}\mathbf{E}\,\exp\bigl(4\varepsilon\,\bigl\|U_0^{(N)}\bigr\|\bigr)&\leqslant 2\,d\,\exp\bigl(16\,C_2^2\,d\varepsilon^2\cdot 2^N\bigr),\\ \mathbf{E}\,\exp\bigl(4\varepsilon\,\bigl\|S_0^{\{N\}}\bigr\|\bigr)&\leqslant 2\,d\,\exp\bigl(16\,C_2^2\,d\varepsilon^2\cdot 2^N\bigr).\end{aligned} \tag{10.89}$$

From (10.1), (10.3), (10.5), (10.6), (10.8), (10.82), (10.85)–(10.87), (10.89) it follows that

$$\begin{aligned}\mathbf{E}\,\zeta\,\exp\big(\varepsilon\,\gamma_j(N)\big) &\leqslant c\,d\,\exp\Big(-\frac{a_5^2\cdot 2^N}{8\,C_2^2\,d^4\,\tau^2}+22\,\varepsilon\,\lambda\cdot 2^N+8\,C_2^2\,d\varepsilon^2\cdot 2^N\Big)\\ &\leqslant c\,d\,\exp\Big(-\frac{a_5^2\cdot 2^N}{16\,C_2^2\,d^4\,\tau^2}\Big)\leqslant c\,d\,\exp\Big(-\frac{p\cdot 2^p}{16\,C_2^2\,d}\Big)\\ &\leqslant c\,d\,\exp\Big(-\frac{d^{1/2}\,\tau\,\lambda}{a_5\,C_2^2}\Big)\leqslant c. \end{aligned}\tag{10.90}$$

Denote now by m_0 the minimal integer such that

$$\frac{\tau^2\,d^3\,p}{2^{m_0-p}}<a_6\,. \tag{10.91}$$

It is evident that (10.3), (10.7)–(10.9), (10.91) imply $p\leqslant m_0=M+1\leqslant N$. Using the definition of the random variable ν (see (10.46)), we get

$$\mathbf{P}\big\{\nu=m+1\big\}\leqslant\sum_{n=m}^{N}\bigg(\mathbf{P}\Big\{\big\|U_0^{(n)}\big\|\geqslant\frac{a_5\cdot 2^{um+vn}}{d^{3/2}\,\tau}\Big\}+\mathbf{P}\Big\{\big\|S_0^{\{n\}}\big\|\geqslant\frac{a_5\cdot 2^{n}}{d^{3/2}\,\tau}\Big\}\bigg),\qquad m=m_0,\dots,N-1. \tag{10.92}$$

From the definitions of vectors $U_0^{(n)}, S_0^{\{n\}}$ it follows that they are sums of not more than 2^n independent random vectors with distributions $\mathcal{L}(\widetilde{X}_k)\in\mathcal{A}_d(\tau)$ and covariance operators $\operatorname{cov}\widetilde{X}_k=\mathbb{B}_k$ satisfying the condition (6.2) (see (10.31), (10.34)). Therefore, for $m\geqslant m_0$, the relations (6.2), (6.16), (10.1), (10.3), (10.91) and Lemma 2.7 with $B^2\leqslant C_2^2\cdot 2^n$ imply

$$\begin{aligned}\mathbf{P}\Big\{\big\|U_0^{(n)}\big\|\geqslant\frac{a_5\cdot 2^{um+vn}}{d^{3/2}\,\tau}\Big\}&\leqslant 2\,d\,\exp\Big(-\min\Big\{\frac{a_5^2\cdot 2^{2um+2vn}}{4\,C_2^2\,d^4\,\tau^2\cdot 2^n},\ \frac{a_5\cdot 2^{um+vn}}{4\,d^2\,\tau^2}\Big\}\Big)\\ &=2\,d\,\exp\Big(-\frac{a_5^2\cdot 2^{2um+(2v-1)n}}{4\,C_2^2\,d^4\,\tau^2}\Big),\qquad n=m,\dots,N,\end{aligned}\tag{10.93}$$

and, analogously,

$$\mathbf{P}\Big\{\big\|S_0^{\{n\}}\big\|\geqslant\frac{a_5\cdot 2^{n}}{d^{3/2}\,\tau}\Big\}\leqslant 2\,d\,\exp\Big(-\frac{a_5^2\cdot 2^{n}}{4\,C_2^2\,d^4\,\tau^2}\Big),\qquad n=m,\dots,N. \tag{10.94}$$

Taking into account (10.48), (10.84), we see that

$$\mathbf{E}\,(1-\zeta)\,\exp\big(\varepsilon\,\gamma_j(N)\big)\leqslant\exp\big(10\lambda\varepsilon\cdot 2^{m_0}\big)+\sum_{m=m_0+1}^{N}e^{10\,\lambda\,\varepsilon\cdot 2^m}\,\mathbf{P}\big\{\nu=m\big\}. \tag{10.95}$$

From (10.1), (10.3), (10.5), (10.82), (10.91) and from the definition of p (see (10.6)) it follows that

$$10\lambda\varepsilon\cdot 2^{m_0} \leqslant \frac{10\,a_5^2}{512\,C_2^2\,d^4\,\tau^2}\cdot\frac{d^3\,\tau^2\,p\cdot 2^{p+1}}{a_6} \leqslant ap\cdot 2^p \leqslant ap\,(\lambda\,d^{3/2}\,\tau)^{1/v}. \tag{10.96}$$

Using (10.1)–(10.3), (10.5), (10.6), (10.82), (10.91)–(10.94), we obtain

$$\begin{aligned}\sum_{m=m_0+1}^{N} e^{10\,\lambda\,\varepsilon\cdot 2^m}\,\mathbf{P}\left\{\nu=m\right\} &\leqslant \sum_{m=m_0+1}^{N}\ \sum_{n=m-1}^{N} 4\,d\,\exp\Big(\frac{10\,a_5^2\cdot 2^m}{512\,C_2^2\,d^4\,\tau^2}-\frac{a_5^2\cdot 2^{2um+(2v-1)n}}{8\,C_2^2\,d^4\,\tau^2}\Big)\\
&\leqslant 4\,d\sum_{m=m_0+1}^{N}\ \sum_{n=m-1}^{N}\ \exp\Big(-\frac{a_5^2\cdot 2^{2um+(2v-1)n}}{16\,C_2^2\,d^4\,\tau^2}\Big)\\
&\leqslant 4\,d\sum_{m=m_0+1}^{\infty} c\,\log^*(a_5^{-1}\,d\,\tau)\,\exp\Big(-\frac{a_5^2\cdot 2^m}{32\,C_2^2\,d^4\,\tau^2}\Big)\\
&\leqslant c\,d\,\big(\log^*(a_5^{-1}\,d\,\tau)\big)^2\,\exp\Big(-\frac{a_5^2\cdot 2^{m_0}}{16\,C_2^2\,d^4\,\tau^2}\Big)\\
&\leqslant c\,d\,\big(\log^*(a_5^{-1}\,d\,\tau)\big)^2\,\exp\Big(-\frac{p\cdot 2^p}{16\,C_2^2\,d}\Big)\\
&\leqslant c\,d\,\big(\log^*(a_5^{-1}\,d\,\tau)\big)^2\,\exp\Big(-\frac{d^{1/2}\,\tau\,\lambda}{C_2^2\,a_5}\Big)\leqslant c.\end{aligned} \tag{10.97}$$

The inequality (10.83) follows from (10.5), (10.6), (10.90), (10.95)–(10.97). □

Lemma 10.3. *Assume now (instead of* (10.8) *and* (10.9)*) that* $N = M$. *Then there exists a positive quantity* a_8 *depending only on* v *and such that the following inequality is valid*:

$$\mathbf{E}\,\exp\Big(\varepsilon\,\big(\rho_j(M)+\delta_j(M)\big)\Big) \leqslant \exp\big(a_8\,p\,(\lambda\,d^{3/2}\,\tau)^{1/v}\big), \qquad j=1,\ldots,2^M. \tag{10.98}$$

Proof. Similarly to (10.78), one can verify that $\rho_j(M) \leqslant 10\lambda\cdot 2^M$. Therefore, using (10.38), we obtain

$$\mathbf{E}\,\exp\Big(\varepsilon\,\big(\rho_j(M)+\delta_j(M)\big)\Big) \leqslant \exp\big(10\,\varepsilon\,\lambda\cdot 2^M\big)\,\mathbf{E}\,e^{\varepsilon W}, \tag{10.99}$$

where $W = \|\widetilde{S}\| + \|T\|$. From the definitions of $\widetilde{S}$, $T = T_j$ (see (5.7), (10.38)) it follows that they are sums of not more than 2^M independent random vectors $\widetilde{X}_k$ and $Y_k^{[M]}$ respectively. By analogy with (10.14), we have $\mathbf{P}\left\{\|T\| \leqslant 4\lambda\cdot 2^M\right\} = 1$. Hence,

$$\mathbf{E}\,e^{\varepsilon W} \leqslant \mathbf{E}\,e^{\varepsilon\|\widetilde{S}\|}\,\exp\big(4\varepsilon\,\lambda\cdot 2^M\big). \tag{10.100}$$

Recall that $\operatorname{cov} \widetilde{X}_k = \mathbb{B}_k$, $\mathcal{L}(\widetilde{X}_k) \in \mathcal{A}_d(\tau)$. Taking into account (10.88) and applying Lemma 2.7 with $B^2 \leqslant C_2^2 \cdot 2^M$, $h = d^{1/2}\varepsilon$, we see that

$$\mathbf{E}\, e^{\varepsilon \|\widetilde{S}\|} \leqslant 2\,d\, \exp\bigl(C_2^2\, d\varepsilon^2 \cdot 2^M\bigr). \tag{10.101}$$

From (10.1), (10.3), (10.5)–(10.7), (10.82), (10.99)–(10.101) it follows that

$$\begin{aligned}\mathbf{E}\, \exp\Bigl(\varepsilon\,\bigl(\rho_j(M) + \delta_j(M)\bigr)\Bigr) &\leqslant 2\,d\, \exp\Bigl(14\,\varepsilon\,\lambda \cdot 2^M + C_2^2\, d\varepsilon^2 \cdot 2^M\Bigr)\\ &\leqslant 2\,d\, \exp\Bigl(\frac{15\,a_5^2 \cdot 2^M}{512\,C_2^2\, d^4\,\tau^2}\Bigr) \leqslant 2\,d\, \exp\bigl(a\,p \cdot 2^p\bigr)\\ &\leqslant 2\,d\, \exp\bigl(a\,p\,(\lambda\, d^{3/2}\,\tau)^{1/v}\bigr). \end{aligned}\tag{10.102}$$

It is clear that (10.1), (10.5), (10.6), (10.102) imply (10.98). □

Write

$$\begin{aligned} b(N) &= \max_{1\leqslant m\leqslant 2^N} \mathbf{E}\, \exp\Bigl(\varepsilon\bigl(\delta_m(N) + \rho_m(N)\bigr)\Bigr),\\ b(N-p) &= \max_{1\leqslant m\leqslant 2^{N-p}} \mathbf{E}\, \exp\Bigl(\varepsilon\bigl(\delta_m(N-p) + \rho_m(N-p)\bigr)\Bigr)\\ r(N) &= \max_{1\leqslant m\leqslant 2^N} \mathbf{E}\, \exp\bigl(\varepsilon\,\gamma_m(N)\bigr). \end{aligned}\tag{10.103}$$

One can verify that the definitions of $b(N)$, $b(N-p)$ are good coordinated: if we begin with 2^{N_0} vectors, where $N_0 = N - p$, then $b(N_0) = b(N-p)$ (see (10.27), (10.38)). Since j, $1 \leqslant j \leqslant 2^N$, is arbitrary, we can easily derive from (10.51), (10.103) and Lemma 10.2 that

$$b(N) \leqslant b(N-p)\, r(N) \leqslant b(N-p)\, \exp\Bigl(a_7\, p\,(\lambda\, d^{3/2}\,\tau)^{1/v}\Bigr). \tag{10.104}$$

Continuing this procedure and taking into account (10.7)–(10.9), we obtain

$$b(N) \leqslant b(M)\, \exp\Bigl(a_7\,(N-M)\,(\lambda\, d^{3/2}\,\tau)^{1/v}\Bigr). \tag{10.105}$$

From (10.9), (10.98), (10.103), (10.105) it follows that

$$b(N) \leqslant \exp\Bigl(a_9\, N\,(\lambda\, d^{3/2}\,\tau)^{1/v}\Bigr) \tag{10.106}$$

with $a_9 = \max\{a_7, a_8\}$. Using (10.38), (10.103), (10.106), we obtain

$$\begin{aligned}\mathbf{E}\, \exp\Bigl(\varepsilon\,\Delta(\widetilde{X}, \widetilde{Y})\Bigr) &\leqslant \sum_{m=1}^{2^N} \mathbf{E}\, \exp\Bigl(\varepsilon\,\delta_m(N)\Bigr) \leqslant \sum_{m=1}^{2^N} \mathbf{E}\, \exp\Bigl(\varepsilon\bigl(\delta_m(N) + \rho_m(N)\bigr)\Bigr)\\ &\leqslant 2^N\, \exp\Bigl(a_9\, N\,(\lambda\, d^{3/2}\,\tau)^{1/v}\Bigr) \leqslant \exp\Bigl(a\, N\,(\lambda\, d^{3/2}\,\tau)^{1/v}\Bigr). \end{aligned}\tag{10.107}$$

Assume now, instead of (10.8), that $N \leqslant M$ and let the vectors $\widetilde{X}_k$, $k = 1, \dots, 2^M$, be constructed with the help of the vectors $\widetilde{Y}_k$, $k = 1, \dots, 2^M$, in the same way as the vectors $X_k^* = \widetilde{X}_{(k-1)\cdot 2^{N-M}+1}$, $k = 1, \dots, 2^M$, was constructed earlier with the help of the vectors $Y_k^* = Z^{[M+p]}_{(k-1)\cdot 2^{N-M}+1}$, $k = 1, \dots, 2^M$. For $k = 2^N + 1, \dots, 2^M$ we can take $\mathcal{L}(\widetilde{X}_k) = \mathcal{L}(\widetilde{X}_1)$, $\mathcal{L}(\widetilde{Y}_k) = \mathcal{L}(\widetilde{Y}_1)$. Now, if we denote

$$\Delta_N(\widetilde{X}, \widetilde{Y}) = \max_{1 \leqslant j \leqslant 2^N} \left\| \sum_{k=1}^{j} \widetilde{Y}_k - \sum_{k=1}^{j} \widetilde{X}_k \right\|, \tag{10.108}$$

then, using (10.38), (10.98), (10.103), (10.108), we obtain

$$\begin{aligned} \mathbf{E} \exp\Big(\varepsilon\, \Delta_N(\widetilde{X}, \widetilde{Y})\Big) &\leqslant \sum_{m=1}^{2^N} \mathbf{E} \exp\Big(\varepsilon\, \delta_m(M)\Big) \leqslant 2^N \exp\Big(a_8\, p\, (\lambda\, d^{3/2}\, \tau)^{1/v}\Big) \\ &\leqslant \exp\Big(a\, (N+p)\, (\lambda\, d^{3/2}\, \tau)^{1/v}\Big). \end{aligned} \tag{10.109}$$

Proof of Theorem 1.5. Without loss of generality we suppose that $0 < \alpha \leqslant 1/5$ and $n = 2^N$. Suppose that independent Gaussian vectors $Y_1, \dots, Y_{2^N}$ are already constructed. According to Theorem 8.4, we can construct the independent random vectors $\widetilde{Y}_1, \dots, \widetilde{Y}_{2^N}$ so that $\mathcal{L}(\widetilde{Y}_k) = \Psi\big(\mathcal{L}(\widetilde{\eta}_k)\big)$, $k = 1, \dots, n$, and

$$\mathbf{E} \exp\Big(\frac{c_{53}\, \Delta(\widetilde{Y}, Y)}{d^{5/2}\, \tau}\Big) \leqslant \exp\Big(c_{54}\, (N+1)\, \log^* d\Big). \tag{10.110}$$

Then we construct the vectors $\widetilde{X}_1, \dots, \widetilde{X}_{2^N}$ as functions of $\widetilde{Y}_1, \dots, \widetilde{Y}_{2^N}$ as it was described above in this section.

Taking $v = \dfrac{2}{3+\alpha}$, we see that v satisfies (10.1) and

$$v^{-1} = \frac{3+\alpha}{2}. \tag{10.111}$$

The relations (10.1), (10.3), (10.111) show that $a, a_5, \dots, a_9$ can be considered as quantities depending only on α, C_1 and C_2.

Let M be the maximal integer satisfying the inequality (10.7) with a_6, p defined by (10.3), (10.6). Define N as the minimal integer such that $n \leqslant 2^N$, if $n \leqslant 2^M$, and as the minimal integer satisfying the relations $n \leqslant 2^N$ and (10.8), if $n > 2^M$. In both cases we have

$$N \leqslant c\, \log n + p. \tag{10.112}$$

It is evident that from (10.1), (10.5), (10.6), (10.111) it follows that

$$p \leqslant a_{10}\, \log^*(\lambda\, d\, \tau) \leqslant a_{11}\, (\lambda\, d^{3/2}\, \tau)^{\alpha/2}, \tag{10.113}$$

where a_{10}, a_{11} depend only on α. The statement of Theorem 1.5 follows now from (10.5), (10.82), (10.107)–(10.113). □

11. Proof of Theorem 1.4

Theorem 11.1. *Suppose that* $\alpha > 0$, $A > 0$, *and* $\xi_1, \dots, \xi_s$ *are 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* $i_0 = 0$, $i_1, \dots, i_n = s$ *satisfying the following conditions. Write*

$$\widetilde{\xi}_k = A\left(\xi_{i_{k-1}+1} + \dots + \xi_{i_k}\right), \qquad k = 1, \dots, n, \tag{11.1}$$

and suppose that (for all $k = 1, \dots, n$*) the covariance operators* $\operatorname{cov}\widetilde{\xi}_k = \mathbb{B}_k$ *satisfy Condition* $\mathbf{B}_2$ *and* $\mathcal{L}(\widetilde{\xi}_k) \in \mathcal{A}_d(C_1)$. *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* $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathbf{E}\,Y_j = 0$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $j = 1, \dots, s$, *and*

$$\mathbf{E}\,\exp\Big(\frac{a_{12}\,A\,\Delta(X,Y)}{d^{9/2}\,\log^* d}\Big) \leqslant \exp\left(a_{13}\,d^{3+\alpha}\,\log^* n\right), \tag{11.2}$$

where a_{12}, a_{13} *are positive quantities depending only on* α, C_1, C_2. *Moreover, if all moments of the third order of the vectors* $\widetilde{\xi}_k$ *are equal to zero, then then this assertion is valid without the factor* $\log^* d$ *in the denominator of the fraction in the left-hand side of* (11.2).

For the proof of Theorem 11.1 we need the following Lemma 11.2 which is proved in Zaitsev (1995, Lemma 14.1) or (1998, Lemma 9.2), see as well the proof of Theorem 1.6 in Zaitsev (1998).

Lemma 11.2. *Let* $\mathcal{L}(\xi) \in \mathcal{A}_d(1)$, $\mathbf{E}\,\xi = 0$ *and* $\operatorname{cov}\xi = \mathbb{I}_d$. *Then there exists a random vector* η *and absolute positive constants* $c_{75}, c_{76} \geqslant 1$ *such that*

$$\begin{gathered}\mathbf{E}\,\eta = \mathbf{E}\,\xi = 0, \qquad \operatorname{cov}\eta = \operatorname{cov}\xi = \mathbb{I}_d,\\ \mathbf{P}\left\{|\eta| \leqslant \rho\right\} = 1, \qquad \rho = c_{75}\,(1 + \log d), \qquad \mathcal{L}(\eta) \in \mathcal{A}_d(c_{76}),\end{gathered} \tag{11.3}$$

and, for all $u, v, w \in \mathbb{R}^d$,

$$\mathbf{E}\,\langle \eta, u\rangle\langle \eta, v\rangle\langle \eta, w\rangle = \mathbf{E}\,\langle \xi, u\rangle\langle \xi, v\rangle\langle \xi, w\rangle. \tag{11.4}$$

Moreover, if all moments of the third order of the vector ξ *are equal to zero, then then the same assertion is valid with* $\rho = c_{75}$.

Proof of Theorem 11.1. Consider the random vectors $\zeta_k = \mathbb{B}_k^{-1/2}\,\widetilde{\xi}_k$. Obviously, $\operatorname{cov}\zeta_k = \mathbb{I}_d$ and Condition $\mathbf{B}_2$ implies that $\left\|\mathbb{B}_k^{-1/2}\right\| \leqslant C_1^{-1}$, $\left\|\mathbb{B}_k^{1/2}\right\| \leqslant C_2$ (see (6.15), (6.18) and (6.21)). Therefore, by Lemma 2.5, we have $\mathcal{L}(\zeta_k) \in \mathcal{A}_d(1)$. Using

Lemma 11.2, we see that there exist random vectors η_k, which have the same moments of the first three orders as ζ_k, $\mathcal{L}(\eta_k) \in \mathcal{A}_d(c_{76})$, $\mathbf{P}\big\{\, \|\,\eta_k\| \leqslant \rho\sqrt{d}\big\} = 1$, where $\rho = c_{75}\,(1+\log d)$. Put $\widetilde{\eta}_k = \mathbb{B}_k^{1/2}\,\eta_k$. The random vectors $\widetilde{\eta}_k$ have the same moments of the first three orders as $\widetilde{\xi}_k$, $\mathbf{P}\big\{\, \|\,\widetilde{\eta}_k\| \leqslant C_2\,\rho\sqrt{d}\big\} = 1$, and, by Lemma 2.5, $\mathcal{L}(\widetilde{\eta}_k) \in \mathcal{A}_d(c_{76}\,C_2)$. Thus, the random vectors $\widetilde{\xi}_1, \dots, \widetilde{\xi}_n$ satisfy the conditions of Theorem 1.5 with

$$\tau = \max\big\{\,1,\ C_1,\ c_{76}\,C_2\,\big\}, \qquad \lambda = \max\big\{\,C_2,\ 1,\ C_2\,c_{75}\,(1+\log d)\,\big\}\sqrt{d}. \tag{11.5}$$

Therefore, for any $\beta > 0$ one can construct on a probability space a sequence of independent random vectors $\widetilde{X}_1, \dots, \widetilde{X}_n$ and a sequence of independent Gaussian random vectors $\widetilde{Y}_1, \dots, \widetilde{Y}_n$ so that $\mathcal{L}(\widetilde{X}_k) = \mathcal{L}(\widetilde{\xi}_k)$, $\operatorname{cov}\widetilde{Y}_k = \operatorname{cov}\widetilde{X}_k$, $\mathbf{E}\,\widetilde{Y}_k = 0$, $k = 1, \dots, n$, and

$$\mathbf{E}\,\exp\big(b\,\Delta(\widetilde{X}, \widetilde{Y})\big) \leqslant \exp\big(a_{15}\,(\lambda\,d^{3/2}\,\tau)^{3/2+\beta}\log^* n\big), \tag{11.6}$$

where

$$b = \frac{a_{14}}{\lambda\,d^4\,\tau^2} \tag{11.7}$$

and a_{14}, a_{15} are quantities depending only on β, C_1, C_2. Using the Berkes–Philipp lemma (1979, Lemma A1, p. 53) and (11.1), we can suppose without loss of generality that

$$\begin{aligned} \widetilde{X}_k &= A\,\big(X_{i_{k-1}+1} + \dots + X_{i_k}\big), \\ \widetilde{Y}_k &= A\,\big(Y_{i_{k-1}+1} + \dots + Y_{i_k}\big), \end{aligned} \qquad k = 1, \dots, n, \tag{11.8}$$

where $\big\{\,X_1, \dots, X_s\big\}$ and $\big\{\,Y_1, \dots, Y_s\big\}$ are two collections of mutually independent random vectors with $\mathcal{L}(X_j) = \mathcal{L}(\xi_j)$, $\mathcal{L}(Y_j) \in \mathfrak{G}_d$, $\operatorname{cov} Y_j = \operatorname{cov} X_j$, $\mathbf{E}\,Y_j = 0$, for $j = 1, \dots, s$.

Write

$$\begin{aligned} S_j &= X_1 + \dots + X_j, \\ T_j &= Y_1 + \dots + Y_j, \end{aligned} \qquad j = 1, \dots, s; \qquad S_0 = T_0 = 0; \tag{11.9}$$

and

$$\Delta_1 = \max_{0\leqslant k<n}\ \max_{1\leqslant m\leqslant i_{k+1}-i_k} \big\|\,X_{i_k+1} + \dots + X_{i_k+m}\big\|, \tag{11.10}$$

$$\Delta_2 = \max_{0\leqslant k<n}\ \max_{1\leqslant m\leqslant i_{k+1}-i_k} \big\|\,Y_{i_k+1} + \dots + Y_{i_k+m}\big\|, \tag{11.11}$$

$$\Delta_3 = \max_{1\leqslant k\leqslant n} \big\|\,S_{i_k} - T_{i_k}\big\|. \tag{11.12}$$

It is evident that from (1.1), (11.9)–(11.12) it follows that

$$\Delta(X, Y) \leqslant \Delta_1 + \Delta_2 + \Delta_3. \tag{11.13}$$

Moreover, the relations (1.1), (11.8), (11.9), (11.12) imply

$$\Delta_3 = A^{-1}\,\Delta(\widetilde{X},\widetilde{Y}). \tag{11.14}$$

It is clear that we can take $a_{14} < {}^1\!/_2$. From (11.7) and from the conditions of Theorem 11.1 it follows that $0 < d^{1/2}\,b\,\tau < {}^1\!/_2$. Recall that Gaussian distributions from $\mathfrak{G}_d$ belong to all classes $\mathcal{A}_d(\tau)$. Using (11.5), Condition $\mathbf{B}_2$ and Lemma 2.7 with $t = d^{1/2}\,b$, $B^2 \leqslant C_2^2$, we have

$$\mathbf{E}\,e^{bA\Delta_l} \leqslant n\cdot 3\,de, \qquad l = 1, 2. \tag{11.15}$$

Now the statement of Theorem 11.1 follows easily from the relations (11.5)–(11.7) and (11.13)–(11.15) by an application of the Hölder inequality to the exponential moment $\mathbf{E}\,\exp\bigl(bA\,\Delta(X,Y)/3\bigr)$.

The case when all moments of the third order of the vectors $\widetilde{\xi}_k$ are equal to zero can be considered in a similar way (cf. Lemma 11.2). □

Proof of Theorem 1.4. Without loss of generality we assume that $\tau^2 = 2^M\,C_1^2$, and $l = 2^N$, where M and N are positive integers and $M \leqslant N$. Consider the random vectors

$$\widetilde{\xi}_k = 2^{-M/2}\,\bigl(\zeta_{(k-1)\cdot 2^M+1} + \cdots + \zeta_{k\cdot 2^M}\bigr), \qquad k = 1, \ldots, n = 2^{N-M}.$$

These vectors satisfy the conditions of Theorem 11.1 with $A = D\cdot 2^{-M/2} = C_1\,D/\tau$, $n = 2^{N-M} = C_1^2\,l/\tau^2$. It remains to apply this theorem. □

Proof of Theorem 1.2. It is clear that it suffices to consider the case $\tau = 1$. Otherwise one should use the result for $\tau = 1$ to the random variables ξ_j/τ, $j = 1, \ldots, s$ (obviously, $\mathcal{L}\bigl(\xi_j/\tau\bigr) \in \mathcal{S}_1(1)$). Thus, let $\tau = 1$. Then, by Lemma 2.4, $\mathcal{L}(\xi_j) \in \mathcal{B}_1(1)$. Hence, by Lemma 2.3(a), $\mathcal{L}(\xi_j) \in \mathcal{A}_1(c_2)$ and $\mathbf{E}\,\xi_j^2 \leqslant 12$.

There exists an absolute constant c_{77} such that the inequality (1.3) is valid for $B^2 \leqslant c_{77}$ with some sufficiently small absolute constant c_1 if we construct the variables $X_1, \ldots, X_s$ and $Y_1, \ldots, Y_s$ arbitrarily, for example, to be jointly independent. Write $S = X_1 + \cdots + X_s$. By the completeness of the classes $\mathcal{A}_d(\tau)$ with respect to convolution, we have $\mathcal{L}(S) \in \mathcal{A}_1(c_2)$. The rest of the proof in the considered case uses Lemma 2.7 and the Kolmogorov and Hölder inequalities. It is elementary, repeats the arguments of Sakhanenko (1984, Section 3.4) and therefore is omitted. In this case the inequality (1.3) only reflects the fact that the compared random sequences are both sufficiently small.

Let now $B^2 \geqslant c_{77}$. Then one can show that there exists an absolute constant c_{78} and a strictly increasing sequence of non-negative integers $m_0 = 0,\ m_1, \ldots,\ m_l = s$ such that

$$c_{78} \leqslant \mathbf{E}\,\zeta_p^2 \leqslant 24, \qquad \mathcal{L}(\zeta_p) \in \mathcal{A}_1(c_2),$$

where

$$\zeta_p = \xi_{m_{p-1}+1} + \xi_{m_{p-1}+2} + \cdots + \xi_{m_p}, \qquad p = 1, \ldots, l.$$

It remains to use Theorem 1.4 with $d = D = 1$, $C_1 = c_{78}$, $C_2 = 24$ and $\tau = c_2$ taking into account that $B^2 \geqslant c_{78}l$. One should also note that, using the Hölder inequality, we can transfer the large factors under exponential sign in the right-hand side of inequalities of type (1.8) to the denominator of the fraction in the left-hand side. □

Acknowledgment. The main part of this research was done in Bielefeld University. I would like to thank Professor Friedrich Götze for his hospitality during my stay in Bielefeld and for useful discussions.

References

Berger, E., *Fast sichere Approximation von Partialsummen unabhängiger und stationärer ergodischer Folgen von Zufallsvectoren*, Dissertation (1982), Universität Göttingen.

Berkes, I., and Philipp, W., *Approximation theorems for independent and weakly dependent random vectors*, Ann. Probab. **7** (1979), 29–54.

Borovkov, A. A., *On the rate of convergence in the invariance principle*, Theor. Probab. Appl. **18** (1973), no. 2, 207–225.

Csörgő, M., and Révész, P., *A new method to prove Strassen type laws of invariance principle.* I; II, Z. Wahrscheinlichkeitstheor. verw. Geb. **31** (1975), 255–259; 261–269.

Csörgő, M., and Révész, P., *Strong approximations in probability and statistics*, Academic Press, New York, 1981.

Csörgő, S., and Hall, P., *The Komlós–Major–Tusnády approximations and their applications*, Austral. J. Statist. **26** (1984), no. 2, 189–218.

Einmahl, U., *A useful estimate in the multidimensional invariance principle*, Probab. Theor. Rel. Fields **76** (1987a), no. 1, 81–101.

Einmahl, U., *Strong invariance principles for partial sums of independent random vectors*, Ann. Probab. **15** (1987b), 1419–1440.

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

Gihman, I. I., and Skorokhod, A. V., *Theory of random processes.* I, Springer, Berlin–Heidelberg–New York, 1974.

Götze, F., and Zaitsev A. Yu., *Multidimensional Hungarian construction for vectors with almost Gaussian smooth distributions*, Preprint 97-071 SFB 343, Universität Bielefeld (1997), 1–29.

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

Massart, P., *Strong approximation for multivariate empirical and related processes, via KMT construction*, Ann. Probab. **17** (1989), no. 1, 266–291.

Philipp, W., *Almost sure invariance principles for sums of B-valued random variables*, Lect. Notes in Math. **709** (1979), 171–193.

Prokhorov, Yu. V., *Convergence of random processes and limit theorem of probability theory*, Theor. Probab. Appl. **1** (1956), no. 2, 157–214.

Rosenblatt, M., *Remarks on a multivariate transformation*, Ann. Math. Statist. **23** (1952), 470–472.

Rudzkis, R., *Probabilities of large deviations of random vectors*, Lithuanian Math. J. **23** (1983), no. 1, 113–120.

Sakhanenko, A. I., *Rate of convergence in the invariance principles for variables with exponential moments that are not identically distributed*, In: Trudy Inst. Mat. SO AN SSSR, vol. 3, Nauka, Novosibirsk, 1984, pp. 4–49. (Russian)

Saulis, L., *Large deviations for random vectors for certain classes of sets.* I, Lithuanian Math. J. **23** (1983), no. 3, 308–317.

Sazonov, V. V., *Normal approximation – some recent advances*, Lect. Notes in Math. **879** (1981).

Shao, Qi-Man, *Strong approximation theorems for independent random variables and their applications*, J. Multivar. Anal. **52** (1995), no. 1, 107–130.

Skorokhod, A. V., *Studies in the theory of random processes*, Univ. Kiev, Kiev, 1961 (Russian); Engl. transl., Addison–Wesley, Reading, Mass., 1965.

Statulevičius, V. A., *On large deviations*, Z. Wahrscheinlichkeitstheor. verw. Geb. **6** (1966), no. 2, 133–144.

Yurinskii, V. V., *On the error of the Gaussian approximation to the probability of a ball*, unpublished manuscript.

Zaitsev, A. Yu., *On the Gaussian approximation of convolutions under satisfaction of multidimensional analogs of the Bernstein inequality conditions*, LOMI Preprint P-9-84, Leningrad (1984), 1–54. (Russian)

Zaitsev, A. Yu., *Estimates of the Lévy–Prokhorov distance in the multivariate central limit theorem for random variables with finite exponential moments*, Theor. Probab. Appl. **31** (1986), no. 2, 203–220.

Zaitsev, A. Yu., *On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein inequality conditions*, Probab. Theor. Rel. Fields **74** (1987), no. 4, 535–566.

Zaitsev, A. Yu., *On the connection between two classes of probability distributions*, In: Rings and modulus. Limit theorems of probability theory. Vol. 2, Leningrad University, 1988, pp. 153–158. (Russian)

Zaitsev, A. Yu., *Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments.*, Preprint 95-055 SFB 343, Universität Bielefeld (1995), 1–112.

Zaitsev, A. Yu., *Estimates for quantiles of smooth conditional distributions and multidimensional invariance principle*, Siberian Math. J. **37** (1996), no. 4, 807–831. (Russian)

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

Andrei Zaitsev

St. Petersburg Branch of Steklov Mathematical Institute,

Fontanka 27, St. Petersburg 191011, Russia

E-mail address: zaitsev@pdmi.ras.ru