A Nonstandard Hungarian Construction for Partial Sums

Ion Grama[1] and Michael Nussbaum
Institute of Mathematics, Kishinev, Moldova and Weierstrass Institute, Berlin, Germany

June 10, 1997

Abstract

We develop a Hungarian construction for the partial sum process of independent, non-identically distributed random variables. The process is indexed by functions f from a functional class $\mathcal{H}$, but the supremum over $f \in \mathcal{H}$ is taken outside the probability. This nonstandard form is a prerequisite for the functional Komlós-Major-Tusnády inequality in the space of bounded functionals $l^\infty(\mathcal{H})$, but contrary to the latter it essentially preserves the classical $n^{-1/2}\log n$ approximation rate over large functional classes $\mathcal{H}$ such as the Hölder ball of smoothness $1/2$. The nonstandard form has a specific statistical application in the asymptotic equivalence theory for experiments.

1 Introduction

Let X_i, $i = 1, ..., n$ be a sequence of independent random variables with zero means and finite variances. Let $\mathcal{H}$ be a class of real valued functions on the unit interval $[0,1]$ and $t_i = \frac{i}{n}$, $i = 1, ..., n$. The *partial sum process indexed by functions* is the process

$$X^{(n)}(f) = n^{-1/2} \sum_{i=1}^{n} f(t_i) X_i,\ f \in \mathcal{H}.$$

Suppose $f \in \mathcal{H}$ are uniformly bounded; then $X^{(n)}$ may be regarded as a random element with values in $l^\infty(\mathcal{H})$- the space of real valued functionals on $\mathcal{H}$, equipped with the sup-norm $\left\|X^{(n)}\right\|_{\mathcal{H}} = \sup_{f\in\mathcal{H}} \left|X^{(n)}(f)\right|$. The class $\mathcal{H}$ is Donsker if $X^{(n)}$ converges weakly in $l^\infty(\mathcal{H})$ to a Gaussian process. We are interested in associated coupling results, i. e. in finding versions of $X^{(n)}$ and of this Gaussian process on a common probability space which are close as random variables. The standard coupling results of the type "nearby variables with nearby laws" (cf.

---

[1] Research supported by the Deutsche Forschungsgemeinschaft and by the Deutsche Akademische Austauschdienst.

1990 *Mathematics Subject Classification.* Primary 60F17, 60F99 ; Secondary 62G07

Key words and phrases. Komlós-Major-Tusnády inequality, partial sum process, non-identically distributed variables, function classes, asymptotic equivalence of statistical experiments.

Dudley [4], sec. 11.6) naturally refer to the metric $\|\cdot\|_{\mathcal{H}}$. For an appropriate version of $X^{(n)}$ ($\mathbb{X}^{(n)}$, say) and of a Gaussian process $\mathbb{N}^{(n)}$ we might then obtain that

$$P^*\left(\left\|\mathbb{X}^{(n)} - \mathbb{N}^{(n)}\right\|_{\mathcal{H}} > x\right) \to 0 \tag{1}$$

where P^* refers to outer probability on the common probability space (cf. van der Vaart and Wellner [20], 1.9.3, 1. 10. 4).

We are here interested in a different type of coupling. We are looking for versions $\mathbb{X}^{(n)}$, $\mathbb{N}^{(n)}$ such that

$$\sup_{f\in\mathcal{H}} P\left(\left|\mathbb{X}^{(n)}(f) - \mathbb{N}^{(n)}(f)\right| > x\right) \to 0 \tag{2}$$

such that in addition exponential bounds of the Komlós, Major and Tusnády type are valid. Note that (2) is weaker than (1) since the supremum is taken outside the probability, but there is the further exponential bound requirement. More specifically, we are interested in a construction involving also a rate sequence $r_n \to 0$ such that

$$\sup_{f\in\mathcal{H}} P\left(\left|\mathbb{X}^{(n)}(f) - \mathbb{N}^{(n)}(f)\right| > x\, r_n\right) \le c_0 \exp\{-c_1 x\} \tag{3}$$

Here c_0, c_1 are constants depending on the class $\mathcal{H}$.

The classical results of Komlós, Major and Tusnády [11] and [12] refer to a sup inside the probability for a function class $\mathcal{H} = \mathcal{H}_0$, where $\mathcal{H}_0$ is the class of indicators $f(t) = \mathbf{1}(t \le s)$, $s \in [0,1]$. The following bound was established: for $r_n = n^{-1/2}\log n$

$$P\left(\left\|\mathbb{X}^{(n)} - \mathbb{N}^{(n)}\right\|_{\mathcal{H}_0} > x\, r_n\right) \le c_0 \exp\{-c_1 x\}, \tag{4}$$

provided $X_1, ..., X_n$ is a sequence of i.i.d. r.v.'s fulfilling the Cramér condition

$$E\exp\{tX_i\} < \infty, \quad |t| \le t_0, \quad i = 1, ..., n, \tag{5}$$

where c_0, c_1 are constants depending on the common distribution of the X_i. Note that r_n in (4) can be interpreted as a rate of convergence in the CLT over $l^\infty(\mathcal{H}_0)$. *The main motivation for our paper is that an extension of (4) to larger functional classes $\mathcal{H}$ in general implies a substantial loss of approximation rate r_n* (cp. Koltchinskii ([10], theorem 11.1). Our aim is a construction where the almost $n^{-1/2}$-rate of the original KMT result is preserved despite the passage to large functional classes $\mathcal{H}$ like Lipschitz classes.

Couplings of the type (3) have first been obtained by Koltchinskii ([10], theorem 3.5) and Rio [18] for the empirical process of i.i.d. random variables, as intermediate results. They can be extended to a full functional KMT result, i.e. to a coupling in $l^\infty(\mathcal{H})$ with exponential bounds, but a reduced approximation rate r_n determined by the size of $\mathcal{H}$ in terms of entropy conditions. We thus carry over Koltchinskii's theorem 3.5 from the empirical to the partial sum process, but under very general conditions: the distributions of X_i are allowed to be *nonidentical* and *nonsmooth.* That setting substantially complicates the task of a Hungarian construction. We can rely on the powerful methodology of Sakhanenko [19] who established the classical coupling (4) for the nonidentical and nonsmooth case. We stress however that for the functional version (3) we need to perform the construction entirely anew. Our results relate to Sakhanenko's [19] as Koltchinskii's theorem 3.5 relates to Komlós, Major and Tusnády [11] and [12].

Our further discussion can be arranged into three points.

A. Statistical motivation. The Komlós-Major-Tusnády approximation has found an application recently in the asymptotic theory of statistical experiments. In [14] the classical KMT inequality for the empirical process was used to establish that a nonparametric experiment of i.i.d. observation on an interval can be approximated, in the sense of Le Cam's deficiency distance, by a sequence of signal estimation problems in Gaussian white noise. The two sequences of experiments are then asymptotically equivalent for all purposes of statistical decision with bounded loss. This appears as a generalization of Le Cam's theory of local asymptotic normality, applicable to ill-posed problems like density estimation.

The Hungarian construction had been applied in statistics before, mostly for results on strong approximation of particular density and regression estimators (cf. Csörgő and Révész [3]). It is typical for these results that the *supremum inside the probability* is needed; for such an application of the functional KMT cf. Rio [18]. However for asymptotic equivalence of experiments, it became apparent that it is sufficient to have a coupling like (3) with the "supremum outside the probability". Applying theorem 3.5 from Koltchinskii [10], it became possible in [15] to extend the scope of asymptotic equivalence, for the density estimation problem, down to the limit of smoothness $1/2$. Analogously, the present result can be used for establishing asymptotic equivalence of smooth nongaussian regression models to a sequence of Gaussian experiments, cf [7].

B. Nonidentical and nonsmooth distributions. The assumption of identically distributed r.v.'s substantially restricts the scope of application of the classical KMT inequality for partial sums. However this assumption happens to be an essential point in the original proof by Komlós, Major and Tusnády and also in much of the subsequent work. The original bound was extended and improved by many authors. Multidimensional versions were proved by Einmahl [5] and Zaitsev [21] with a supremum over the class of indicators $\mathcal{H}_0$. A transparent proof of the original result was given by Bretagnolle and Massart [1]. We would like to mention the series of papers by Massart [13] and Rio [16], [17]. They treat the case of $\mathbf{R}^k$-valued r.v.'s X_i, indexed in $\mathbb{Z}_+^d$ with a supremum taken over classes $\mathcal{H}$ of indicator functions $f = \mathbf{1}_S$ of Borel sets S satisfying some regularity conditions. Condition (5) is also relaxed to moment assumptions, but identical distributions are still assumed.

Although there are no formal restrictions on the distributions of X_i when performing a Hungarian construction, it is not possible to get useful quantile inequalities if the r.v.'s X_i are non-identically and non-smoothly distributed. (Recall that in the coupling of a r.v. with a Gaussian via the two distribution functions, a quantile inequality refers to the distance of the two random variables, cf. section 4.) This can be argued in the following way (see Sakhanenko [19]). Let us consider the sum $S = X_1 + ... + X_n$, where X_i takes values $\pm(1 + 2^{-i})$. Then we can identify each realization X_i by knowing only S. In the dyadic Hungarian scheme, the conditional distribution of $X_1 + ... + X_{[n/2]}$ given S is considered and used for coupling with a Gaussian random variable. However this distribution is now degenerate and hence not useful for coupling. This problem does not appear in the i.i.d. case, due to the exchangeability of the X_i. Quantile inequalities are an essential ingredient in the results of Komlós et al. [11] and [12].

We take a way to overcome this difficulty proposed by Sakhanenko [19]. In his original paper Sakhanenko treats the case of independent non-identically distributed r.v.'s for a class of intervals $\mathcal{H} = \mathcal{H}_0$. The particularly simple structure of functions in the set $\mathcal{H}_0$ simplifies the problem substantially. Here we consider the problem in another setting: $\mathcal{H} = \mathcal{H}(\frac{1}{2}, L)$, where $\mathcal{H}(\frac{1}{2}, L)$ is a Hölder ball with exponent $\frac{1}{2}$ and the sup is outside the probability, i.e. we give an exponential bound for the quantity (3) uniformly in f over the set of functions $\mathcal{H}(\frac{1}{2}, L)$. This setting makes the problem more complicated. In particular this is related to the fact that the pairs $(\widetilde{X}_i, \widetilde{W}_i)$, $i = 1, ..., n$, of r.v's $\widetilde{X}_i \stackrel{d}{=} X_i$ and $\widetilde{W}_i \stackrel{d}{=} W_i$, $i = 1, ..., n$ constructed on the same probability space by the KMT method are no longer independent.

C. Coupling from marginals. A weaker coupling of $\mathbb{X}^{(n)}$ and $\mathbb{N}^{(n)}$ can be obtained as follows. Assume for a moment that the X_i are uniformly bounded: $|X_i| \leq L$, $i = 1, \ldots, n$ and also that $\|f\|_\infty \leq L$, $f \in \mathcal{H}$. Take a finite collection of functions $\mathcal{H}_{00} = (f_j)_{j=1,\ldots,d} \subset \mathcal{H}$ and consider $Z_i = (f(t_i)X_i)_{f \in \mathcal{H}_{00}}$ as random vectors in $\mathbf{R}^d$. Reasoning as in Fact 2.2 of Einmahl and Mason [6] (using the result of Zaitsev [22] on the Prokhorov distance between the law of $\sum_{i=1}^n Z_i$ and a Gaussian law) we infer that for all such $\mathcal{H}_{00}$ there are versions $\mathbb{X}^{(n)}$, $\mathbb{N}^{(n)}$ such that

$$P\left(\max_{f \in \mathcal{H}_{00}} n^{1/2} \left|\mathbb{X}^{(n)}(f) - \mathbb{N}^{(n)}(f)\right| \geq x\right) \leq c_0(d) \exp(-c_1(d) x L^{-1}),\ x \geq 0. \tag{6}$$

This yields (3) with rate $r_n = n^{-1/2}$ for every finite class $\mathcal{H}_{00} \subset \mathcal{H}$ of size d, but with constants $c_0(d), c_1(d)$ depending on d. Hence any attempt to construct $\mathbb{X}^{(n)}$ and $\mathbb{N}^{(n)}$ on the full class $\mathcal{H}$ from (6) is bound to entail a substantial loss in rate r_n, but laws of the iterated logarithm can be established in this way (cf. Einmahl and Mason [6]). Thus, to obtain (3) for $r_n = n^{-1/2} \log^2 n$ and a full Hölder class $\mathcal{H}(\frac{1}{2}, L)$, the shortcut via (6) appears not feasible, and we revert to a direct KMT-type construction.

In order to make the proof more transparent we prefer to give a non-optimal (up to a logarithmic term) result, but we believe that it is possible to get the optimal rate by using the very delicate technique of the paper [19]. The main idea is, roughly speaking, to consider some *smoothed* sequences of r.v.'s instead of the initial *unsmoothed* sequence $X_1, ..., X_n$, and to apply the KMT construction for the smoothed sequences. This we perform by substituting normal r.v.'s N_i for the original r.v.'s X_i, for even indices $i = 2k$ in the initial sequence. Thus we are able to construct one half of our sequence and combine it with a Haar expansion of the function f. For the other half we apply the same argument, which leads to a recursive procedure. It turns out that this kind of smoothing is enough to obtain "good" quantile inequalities although it gives rise to an additional $\log n$ term. On the other hand the usual smoothing technique (of each r.v. X_i individually) fails. Unfortunately even the above smoothing procedure applied with normal r.v.'s is not sufficient to obtain the best power for $\log n$ in the KMT inequality for non-identically distributed r.v.'s. An optimal approach is developed in the paper of Sakhanenko [19] and uses r.v.'s constructed in a special way instead of normals. Roughly speaking it corresponds to taking into consideration the higher terms in an asymptotic expansion for the probabilities of large deviations, which dramatically complicates the problem. For more details we refer the reader to this beautiful paper.

Nevertheless we would like to point out that the additional $\log n$ term which appears in our KMT result does not affect the eventual applications that we have in mind, i. e. the asymptotic equivalence of sequences of nonparametric statistical experiments. We also believe that that a stronger version of this result (with a sup inside the probability) might be of use for constructing efficient kernel estimators in nonparametric models. But such an extension is beyond of the scope of the paper.

2 Notations and main results

Suppose that on the probability space $(\Omega', \mathcal{F}', P')$ we are given a sequence of independent r.v.'s $X_1, ..., X_n$ such that for all $i = 1, ..., n$

$$E'X_i = 0, \quad \gamma_n C_{\min} \leq E'X_i^2 \leq C_{\max}\gamma_n \tag{7}$$

with some constants $0 < C_{\min} < C_{\max} < \infty$, where γ_n is a sequence of real numbers $0 < \gamma_n \leq 1$, $n \geq 1$. Hereafter E' is the expectation under the measure P'. Assume also that the following condition due to Sakhanenko [19]

$$\lambda_0 E' |X_i|^3 \exp\{\lambda_0 |X_i|\} \leq E'X_i^2 \tag{8}$$

holds true for $i = 1, ..., n$ with some constant $\lambda_0 > 0$.

Along with this assume that on another probability space $(\Omega, \mathcal{F}, P)$ we are given a sequence of independent normal r.v.'s $N_1, ..., N_n$ such that

$$EN_i = 0, \quad EN_i^2 = E'X_i^2, \tag{9}$$

for $i = 1, ..., n$. Hereafter E is the expectation under the measure P.

Let $\mathcal{H}(\frac{1}{2}, L)$ be the Hölder ball with exponent $\frac{1}{2}$, i.e. the set of real valued functions f defined on the unit interval $[0,1]$ and satisfying the following conditions

$$|f(x) - f(y)| \leq L |x - y|^{1/2},$$

where $L > 0$ and

$$\|f\|_\infty \leq L/2.$$

Let $t_i = 1/n$, $i = 1, ..., n$ be a uniform grid in the unit interval $[0,1]$. The notation $Y \stackrel{d}{=} X$ for random variables means equality in distribution.

Theorem 1 *A sequence of independent r.v.'s $\widetilde{X}_1, ..., \widetilde{X}_n$ can be constructed on the probability space $(\Omega, \mathcal{F}, P)$ such that $\widetilde{X}_i \stackrel{d}{=} X_i$, $i = 1, ..., n$ and*

$$\sup_{f \in \mathcal{H}(\frac{1}{2}, L)} P\left(\left| \sum_{i=1}^n f(t_i)(\widetilde{X}_i - N_i) \right| > x \log^2 n \right) \leq c_0 \exp\{-c_1 x\}, \quad x \geq 0,$$

where c_0, c_1 are constants depending only on $C_{\min}$, $C_{\max}$, λ_0, L.

Remark 1 *In the above theorem the r.v.'s X_i, $i = 1, ..., n$ are not supposed to be identically distributed nor to have smooth distributions, although the result is new even in the case of i.i.d. r.v.'s.*

Remark 2 *In the notation of (3), we have a rate $r_n = n^{-1/2} \log^2 n$.*

Remark 3 *The lower bound in condition (7) cannot be relaxed under the uniform design $t_i = 1/n$, $i = 1, ..., n$. We conjecture that $E'X_i^2$ could be arbitrarily small if, for instance, the design is chosen to be*

$$t_i = \frac{B_i^2}{B_n^2}, \quad B_i^2 = \sum_{k=1}^i E'X_i^2, \quad i = 1, ..., n,$$

but this does not follow directly from our proof.

Theorem 1 can be formulated in an equivalent but a little bit more compact form.

Theorem 2 *A sequence of independent r.v.'s* $\widetilde{X}_1,...,\widetilde{X}_n$ *can be constructed on the probability space* $(\Omega,\mathcal{F},P)$ *such that* $\widetilde{X}_i \stackrel{d}{=} X_i$, $i=1,...,n$ *and*

$$\sup_{f\in\mathcal{H}(\frac{1}{2},L)} E\exp\left\{c_0\frac{1}{\log^2 n}\left|\sum_{i=1}^n f(t_i)(\widetilde{X}_i - N_i)\right|\right\} \le c_1,$$

where c_0 *and* c_1 *are constants depending only on* $C_{\min}$, $C_{\max}$, λ_0, L.

It is easy to see that Theorems 1 and 2 follow from Theorem 3 below.

Theorem 3 *A sequence of independent r.v.'s* $\widetilde{X}_1,...,\widetilde{X}_n$ *can be constructed on the probability space* $(\Omega,\mathcal{F},P)$ *such that* $\widetilde{X}_i \stackrel{d}{=} X_i$, $i=1,...,n$ *and for any* t *satisfying* $|t|\le c_0$

$$\sup_{f\in\mathcal{H}(\frac{1}{2},L)} E\exp\left\{t\frac{1}{\log^2 n}\left(\sum_{i=1}^n f(t_i)(\widetilde{X}_i - N_i)\right)\right\} \le \exp\left\{t^2 c_1\right\},$$

where c_0 *and* c_1 *are constants depending only on* $C_{\min}$, $C_{\max}$, λ_0, L.

The proof of Theorem 3 is given in Section 6.

Now we turn to a particular case of the above results. Assume that the sequence of independent r.v.'s $X_1,...,X_n$ is such that for all $i=1,...,n$

$$E'X_i = 0, \quad C_{\min} \le E'X_i^2 \le C_{\max} \tag{10}$$

for some constants $0 < C_{\min} < C_{\max} < \infty$. Assume also that the following Cramér condition

$$E'\exp\{C_0|X_i|\} \le C_1 \tag{11}$$

holds true for $i=1,...,n$ with some constants $C_0>0$ and $1<C_1<\infty$.

We establish that Sakhanenko's condition (8) holds true under (10) and (11). This follows from the next almost obvious assertion.

Proposition 4 *Let* X_i, $i=1,...,n$ *be r.v.'s satisfying (10) and (11). Then (8) is also satisfied with some* $\lambda_0>0$.

Proof. Assume w. l. o. g. that $C_0\le 1$ and put $\lambda_0 = C_0^3\min\left\{1/2, C_1^{-1}C_{\min}/48\right\}$. Since $y^3\le 6\exp\{y\}$, we have for any $y\ge 0$, with $t=C_0/2$

$$\begin{aligned}\lambda_0 E'|X_i|^3\exp\{\lambda_0|X_i|\} &= \lambda_0 t^{-3}E'|tX_i|^3\exp\{\lambda_0|X_i|\} \le 48\lambda_0 C_0^{-3}E'\exp\{C_0|X_i|\} \\ &\le 48\lambda_0 C_0^{-3}C_1 \le 48\lambda_0 C_0^{-3}C_1 C_{\min}^{-1}E'X_i^2 \le E'X_i^2. \blacksquare\end{aligned}$$

3 Haar expansion

We will make use of some elementary facts on Haar expansions (see for instance Kashin and Saakjan [9]).

The Fourier-Haar basis on the interval $[0, 1]$ is introduced as follows. Consider the dyadic system of partitions by setting

$$s_{k,j} = j2^{-k},$$

for $j = 1, ..., 2^k$ and

$$\Delta_{k,1} = [0, s_{k,1}], \quad \Delta_{k,j} = (s_{k,j-1}, s_{k,j}], \tag{12}$$

for $j = 2, ..., 2^k$, where $k \geq 0$. Define Haar functions via indicators $1(\Delta_{k,j})$

$$h_0 = 1(\Delta_{0,1}), \quad h_{k,j} = 2^{k/2}(1(\Delta_{k+1,2j-1}) - 1(\Delta_{k+1,2j})),$$

for $j = 1, ..., 2^k$ and $k \geq 0$.

If f is a function from $\mathcal{L}_2([0, 1])$ then the following Haar expansion

$$f = c_0(f)h_0 + \sum_{k=0}^{\infty} \sum_{j=1}^{2^k} c_{k,j}(f)h_{k,j},$$

holds true with Fourier-Haar coefficients

$$c_0(f) = \int_0^1 f(u)h_0(u)du, \quad c_{k,j}(f) = \int_0^1 f(u)h_{k,j}(u)\,du, \tag{13}$$

for $j = 1, ..., 2^k$ and $k \geq 0$. Along with this, consider the truncated Haar expansion

$$f_m = c_0(f)h_0 + \sum_{k=0}^{m-1} \sum_{j=1}^{2^k} c_{k,j}(f)h_{k,j}, \tag{14}$$

for some $m \geq 1$.

Proposition 5 *For $f \in \mathcal{H}(\frac{1}{2}, L)$ we have*

$$|c_0(f)| \leq L/2, \quad |c_{k,j}(f)| \leq 2^{-3/2}L2^{-k},$$

for $k = 0, 1, ...$ and $j = 1, ..., 2^k$.

Proof. It is easy to see that

$$\begin{aligned} c_{k,j} &= 2^{k/2}\left(\int_{\Delta_{k+1,2j-1}} f(u)\,du - \int_{\Delta_{k+1,2j}} f(u)\,du\right), \\ &= 2^{k/2} \int_{\Delta_{k+1,2j-1}} (f(u) - f(u + 2^{-(k+1)}))\,du. \end{aligned}$$

Since f is in the Hölder ball $\mathcal{H}(\frac{1}{2}, L)$ we get

$$\begin{aligned} |c_{k,j}| &\leq 2^{k/2} \sup_{u \in \Delta_{k+1,2j-1}} |f(u) - f(u + 2^{-(k+1)})| \int_{\Delta_{k+1,2j-1}} du \\ &\leq 2^{k/2}L2^{-(k+1)/2}2^{-(k+1)} \leq 2^{-3/2}L2^{-k}. \blacksquare \end{aligned}$$

Now we give an estimate for the uniform distance between f and f_m.

Proposition 6 *For $f \in \mathcal{H}(\frac{1}{2}, L)$ we have*

$$\sup_{0\le t\le 1} |f(t) - f_m(t)| \le L 2^{-m/2}.$$

Proof. It is easy to check (see for instance Kashin and Saakjan [9], p. 81) that whenever $t \in \Delta_{m,j}$

$$f_m(t) = 2^m \int_{\Delta_{m,j}} f(s)\, ds,$$

for $j = 1, ..., 2^m$, which gives us $f_m(t) = f(\widetilde{t}_{m,j})$, with some $\widetilde{t}_{m,j} \in \Delta_{m,j}$. Since $f(t)$ is in the Hölder ball $\mathcal{H}(\frac{1}{2}, L)$, we obtain for any $j = 1, ..., 2^m$ and $t \in \Delta_{m,j}$

$$|f(t) - f_m(t)| = |f(t) - f_m(\widetilde{t}_{m,j})| \le L|t - \widetilde{t}_{m,j}|^{1/2} \le L2^{-m}. \blacksquare$$

4 Background on quantile transforms

Assume that on the probability space $(\Omega', \mathcal{F}', P')$ we are given a sequence of independent r.v.'s $X_1, ..., X_n$ which satisfies for any $i = 1, ..., n$ the conditions

$$E' X_i = 0, \tag{15}$$

and

$$E' |X_i|^4 \exp\{\lambda X_i\} < \infty, \tag{16}$$

for some $\lambda > 0$. For any h satisfying $|h| \le \lambda$ introduce the r.v.'s with *conjugate* distributions, i.e. the r.v.'s $X_i(h)$, $i = 1, ..., n$ whose distributions are Cramér transforms

$$P(X_i(h) \le x) = \frac{1}{E' \exp\{hX_i\}} \int_{-\infty}^{x} e^{hy} dF_i(y), \quad x \in \mathbf{R}^1,$$

of distributions $F_i(y) = P(X_i \le y)$, $y \in \mathbf{R}^1$, $i = 1, ..., n$. Put for brevity

$$\begin{aligned} B(h)^2 &= \sum_{i=1}^{n} E' X_i(h)^2, \quad B^2 = B(0)^2 = \sum_{i=1}^{n} E' X_i^2 \\ L(h) &= \sum_{i=1}^{n} E' |X_i(h)|^3, \quad T(h) = \frac{1}{4} \frac{B(h)^2}{L(h)}. \end{aligned}$$

Let $X = X_1 + ... + X_n$ and $X(h) = X_1(h) + ... + X_n(h)$. The characteristic function of $X(h)$ is

$$\varphi(t, h) = E' \exp\left\{\sqrt{-1}\, tX(h)\right\} = \frac{E' \exp\left\{(h + \sqrt{-1}\, t)X\right\}}{E' \exp\{hX\}}.$$

For any $r > 0$ denote

$$\begin{aligned} U(r) &= \sup_{|h|\le r} \int_{|t|>T(h)} |\varphi(t, h)| dt, \\ K(r) &= \sup_{|h|\le r} \sum_{i=1}^{n} E' |X_i(h)|^4. \end{aligned}$$

Definition 1 *Let $r > 0$. The r.v. X is said to be in the class $\mathcal{D}_0(r)$ if $0 < E'X^2 < \infty$ and there are independent r.v.'s $X_1, ..., X_n$, satisfying conditions (15) and (16) with $\lambda = r$, such that $X = X_1 + ... + X_n$ and $r^2U(r) \le B^{-3}$, $4r^2K(r) \le B^2$.*

We now introduce the *quantile transform* and the associated basic inequality (see Lemma 7). Let X be an arbitrary r.v. on a probability space $(\Omega', \mathcal{F}', P')$ and N be a normal r.v. on another probability space $(\Omega, \mathcal{F}, P)$ with distribution functions $F(x)$ and $\Phi(x)$ respectively. Note that the r.v. $U = \Phi(N)$ is distributed uniformly on $[0, 1]$.

Definition 2 *A r.v. $\widetilde{X}$ on $(\Omega, \mathcal{F}, P)$ is said to be the quantile transform of the r.v. N if it satisfies the equation*

$$F(\widetilde{X}) = \Phi(N) = U.$$

It is easy to see that a solution $\widetilde{X}$ exists if and only if F is continuous, and in that case $\widetilde{X}$ is unique a. s. and has distribution function F. Note that for X in $\mathcal{D}_0(r)$, F is continuous.

The following assertion is due to Sakhanenko [19] (see Lemma 1, p. 32).

Lemma 7 *In addition to the above suppose that $X \in \mathcal{D}_0(1)$ and N are such that $E'X^2 = EN^2 = B^2$. Then*

$$\left|\widetilde{X} - N\right| \le c_1 \left\{1 + \frac{\widetilde{X}^2}{B^2}\right\},$$

provided $|\widetilde{X}| \le B^2/2$ and $B \ge 4$. Here c_1 is an absolute constant.

Let us now introduce the *conditional quantile transform* and the associated basic inequality (Lemma 8 below).

Let X_1, X_2 be independent r.v.'s on a probability space $(\Omega', \mathcal{F}', P')$ and N_1, N_2 be independent normal r.v.'s on another probability space $(\Omega, \mathcal{F}, P)$. Put $X_0 = X_1 + X_2$ and $N_0 = N_1 + N_2$. Suppose that we have constructed a r.v. $\widetilde{X}_0$ with the same distribution as X_0, which depends only on N_0 and on some random vector W. Suppose that the r.v.'s N_1 and N_2 do not depend on W. Let $F(x|y)$ be the conditional distribution function of the r.v. X_1 w.r.t. X_0 and $\Phi(x|y)$ be the conditional distribution function of the r.v. N_1 w.r.t. N_0.

Definition 3 *A r.v. $\widetilde{X}_1$ is said to be a quantile transform of N_1 conditionally w.r.t. $\widetilde{X}_0$ and N_0 if it satisfies the equation*

$$F(\widetilde{X}_1|\widetilde{X}_0) = \Phi(N_1|N_0) = U.$$

Remark 4 *The r.v.'s $\widetilde{X}_1$ and $\widetilde{X}_2 \equiv \widetilde{X}_0 - \widetilde{X}_1$ are independent and $\widetilde{X}_1 \stackrel{d}{=} X_1$, $\widetilde{X}_2 \stackrel{d}{=} X_2$. Moreover the r.v.'s $\widetilde{X}_1$ and $\widetilde{X}_2$ are functions of the r.v.'s $\widetilde{X}_0$, N_1 and N_2 only.*

Proof. Consider the r.v. $U = \Phi(N_1|N_0)$. It is easy to see that the distribution of U given $N_0 = y$, for any real y, is uniform on $[0, 1]$. This means that the r.v.'s U and N_0 are independent. Since N_1 and N_2 do not depend on W, we conclude that t U does not depend on N_0 and W. But $\widetilde{X}_0$ is a functionof N_0 and W only. Hence U and $\widetilde{X}_0$ are also independent.

Next, since the uniform r.v. U does not depend on $\widetilde{X}_0$, we can easily check that the distribution of $\widetilde{X}_1$ given $\widetilde{X}_0 = y$, for any real y, is exactly $F(\cdot|y)$. Taking into account that $\widetilde{X}_0 \stackrel{d}{=} X_0$, we conclude the two-dimensional distributions of the pairs $(\widetilde{X}_1, \widetilde{X}_0)$ and (X_1, X_0) coincide. From this we obtain in particular that $\widetilde{X}_1$ and $\widetilde{X}_2$ are independent r.v.'s and that $\widetilde{X}_1 \stackrel{d}{=} X_1$, $\widetilde{X}_2 \stackrel{d}{=} X_2$. Moreover it is obvious by construction that $\widetilde{X}_1$ and $\widetilde{X}_2$ are functions of $\widetilde{X}_0$, N_1 and N_2 only. ■

Remark 5 *We point out that the r.v.* $\widetilde{X}_2 = \widetilde{X}_0 - \widetilde{X}_1$ *defined above is in fact the quantile transform of the r.v.* N_2 *conditionally w.r.t.* X_0 *and* N_0.

The following assertion is due to Sakhanenko [19] (see Lemma 3, p. 32).

Lemma 8 *In addition to the above suppose that* $X_1, X_2 \in \mathcal{D}_0(1)$ *and* N_1, N_2 *are such that* $E'X_1 = EN_1$ *and* $E'X_2 = EN_2$. *Put*

$$\alpha_i = \frac{E'X_i^2}{E'X_0^2}, \quad B^2 = \frac{E'X_1^2 E'X_2^2}{E'X_0^2}.$$

Then for $i = 1, 2$

$$\left|\widetilde{X}_i - N_i - \alpha_i\left(\widetilde{X}_0 - N_0\right)\right| \le c_2\left\{1 + \frac{\widetilde{X}_i^2}{B^2} + \frac{\widetilde{X}_0^2}{B^2}\right\},$$

provided $|\widetilde{X}_i| \le B^2/6$, $|\widetilde{X}_0| \le B^2/6$ *and* $B \ge 4$. *Here* c_2 *is an absolute constant.*

The following remark is easy to check, so the proof is left to the reader.

Remark 6 *If* $X \in \mathcal{D}_0(r)$ *for some* $r > 0$, *then* $rX \in \mathcal{D}_0(1)$.

Remark 7 *If the r.v.'s* X *and* X_1, X_2 *in Lemmas 7 and 8 are in the class* $\mathcal{D}_0(r)$ *for some* $r > 0$, *then the assertions of Lemmas 7 and 8 hold true with* $c_1(r) = r^{-1}c_1$ *and* $c_2(r) = r^{-1}c_2$ *replacing* c_1 *and* c_2 *respectively.*

In the sequel we will give some sufficient conditions for a r.v. X to be in the class $\mathcal{D}_0(r)$.
Assume that on the probability space $(\Omega', \mathcal{F}', P')$ we are given a sequence of independent r.v.'s $X_1, ..., X_n$ which for $i = 1, ..., n$ satisfies the conditions

$$E'X_i = 0, \quad E'X_i^2 \le C_{\max} \tag{17}$$

and

$$\lambda_0 E' |X_i|^3 \exp\{\lambda_0 |X_i|\} \le E'X_i^2, \tag{18}$$

for some $\lambda_0 > 0$. Put $X = X_1 + ... + X_n$.

Proposition 9 *Assume that for any* $\varepsilon \in (0, 1]$ *and any* $\delta \in (0, 1]$ *there is a constant* c_1 *depending only on* ε, δ, λ_0 *and* $C_{\max}$ *such that*

$$\varepsilon^2 \sup_{|h| \le \varepsilon} \int_{|t| > \delta} |\varphi(t, h)|\, dt \le c_1 \left(E'X^2\right)^{-3/2}.$$

Then $X \in \mathcal{D}_0(r)$ *for some constant* $r > 0$ *depending on* λ_0 *and* $C_{\max}$.

Proof. Condition $0 < E'X^2 < \infty$ follows from the independence of $X_1, ..., X_n$ and (17). Condition (18) with $\lambda = \lambda_0/2$ follows from (18) and the inequality $y \leq \exp\{y\}$, $y \geq 0$:

$$E' |X_i|^4 \exp\{\lambda_0 X_i/2\} \leq \frac{1}{\lambda_0} E' |X_i|^3 \exp\{\lambda_0 |X_i|\} \leq \frac{1}{\lambda_0} E' X_i^2 < \infty.$$

Now we proceed to check condition $r^2 U(r) \leq B^{-3}$. First we will show that $T(h) \geq \delta_0$ if $|h| \leq \varepsilon_0$ for some ε_0 and δ_0 depending on λ_0 and $C_{\max}$. To this end put $\psi_i(h) = E' \exp\{hX_i\}$ for $|h| \leq \lambda_0$. A three term Taylor expansion yields for $|h| \leq \lambda_0$

$$\psi_i(h) = 1 + \frac{1}{2} h^2 E' X_i^2 + \frac{1}{6} h^3 E' X_i^3 \exp\{\theta h X_i\},$$

with $0 \leq \theta \leq 1$. Hence by (18) and (17) we have for $|h| \leq \min\left\{\lambda_0, C_{\max}^{-1/2}\right\}$

$$\psi_i(h) \leq 1 + \frac{2}{3} h^2 E' X_i^2 \leq 1 + \frac{2}{3} h^2 C_{\max} \leq \frac{5}{3}. \tag{19}$$

In the same way we show that for $|h| \leq \lambda_0/2$

$$\frac{1}{2} E' X_i \leq \psi_i''(h) \leq \frac{3}{2} E' X_i. \tag{20}$$

For this we note that by Taylor expansion

$$\psi_i''(h) = \psi_i''(0) + h\psi_i'''(\theta h) = E' X_i^2 + hE' X_i^3 \exp\{\theta h X_i\},$$

where $0 \leq \theta \leq 1$, and make use of the inequalities (18) and (17). Inequalities (19) and (20) yield

$$E' X_i(h)^2 = \frac{\psi_i''(h)}{\psi_i(h)} \geq \frac{3}{10} E' X_i^2.$$

Hence for $|h| \leq \varepsilon_0 = \min\left\{\lambda_0/2, C_{\max}^{-1/2}\right\}$

$$B(h)^2 = \sum_{i=1}^{n} E' X_i(h)^2 \geq \frac{3}{10} \sum_{i=1}^{n} E' X_i^2 = \frac{3}{10} B^2. \tag{21}$$

On the other hand by (18) for $|h| \leq \lambda_0$

$$L(h) = \sum_{i=1}^{n} E' |X_i(h)|^3 \leq \frac{1}{\lambda_0} \sum_{i=1}^{n} E' X_i^2 = \frac{1}{\lambda_0} B^2. \tag{22}$$

Put $\delta_0 = \min\left\{\frac{3}{40}\lambda_0, 1\right\}$. Then from (21) and (22) we have for $|h| \leq \varepsilon_0$

$$T(h) = \frac{1}{4} \frac{B(h)^2}{L(h)} \geq \frac{3}{40} \lambda_0 \geq \delta_0.$$

This bound implies

$$U(\varepsilon_0) \leq \sup_{|h| \leq \varepsilon_0} \int_{|t| > \delta_0} |\varphi(t, h)| \, dt \equiv U^*(\varepsilon_0, \delta_0).$$

By the assumptions of the proposition for these ε_0 and δ_0 there is a constant c_1 depending only on ε_0, δ_0 and λ_0 such that $\varepsilon_0^2 U^*(\varepsilon_0, \delta_0) \leq c_1 (E' X^2)^{-3/2} = c_1 B^{-3}$. From the last inequality we get that for $r \leq r_1 = \varepsilon_0 \min \left\{ 1, c_1^{-1/2} \right\}$

$$r^2 U(r) \leq c_1^{-1} \varepsilon_0^2 U(\varepsilon_0) \leq c_1^{-1} \varepsilon_0^2 U^*(\varepsilon_0, \delta_0) \leq B^{-3}.$$

It remains only to check the condition $4r^2 K(r) \leq B^2$; this can easily be obtained from (18) and the inequality $y \leq \exp\{y\}$, $y \geq 0$, if we take $r \leq r_2 = \lambda_0 / 4$:

$$\begin{aligned} 4r^2 K(r) &\leq 4r^2 \sum_{i=1}^n E' |X_i|^4 \exp\{r |X_i|\} \\ &\leq 4r \sum_{i=1}^n E' |X_i|^3 \exp\{2r |X_i|\} \\ &\leq \lambda_0 \sum_{i=1}^n E' |X_i|^3 \exp\{\lambda_0 |X_i|\} \\ &\leq \sum_{i=1}^n E' X_i^2 = B^2. \end{aligned}$$

Now the assertion follows if we put $r = \min\{r_1, r_2\}$. ■

5 A construction for non-identically distributed r.v.'s.

In this section we assume that we are given a sequence of independent r.v.'s $X_i, i = 1, ..., n$ satisfying the relations (7) and (8) for all $i = 1, ..., n$. We will construct this sequence on the same probability space with a sequence of independent normal r.v.'s $N_i, i = 1, ..., n$ satisfying (9) so that they are *as close as possible.* More precisely, the construction is performed so that the quantile inequalities in Section 4 are applicable. The sequences obtained are dependent.

5.1 Some notations

Put $M = [\log_2 n]$. It is clear that $2^M \leq n < 2^{M+1}$. Introduce uniform design points $t_i = \frac{i}{n}$, $i = 1, ..., n$ on the unit interval $[0, 1]$. For any fixed $m = 1, ..., M$ define

$$J_m = \left\{ j : 1 \leq j 2^{M-m} \leq n \right\}.$$

Denote the number of elements in J_m by n_m, i.e. $n_m = \# J_m$. For any $m = 0, ..., M$ and $j \in J_m$ put for brevity $\langle m, j \rangle = j 2^{M-m}$. Let $t_j^m = t_{\langle m,j \rangle}$ and $X_j^m = X_{\langle m,j \rangle}$ for $j \in J_m$.

Put $\beta_m(0) = 0$ and $\beta_m(s) = E'(X_j^m)^2$ if $s \in (t_{j-1}^m, t_j^m]$, $j \in J_m$. If $t_{n_m}^m < 1$ then define $\beta_m(s) = E'(X_{n_m}^m)^2$ for $s \in (t_{n_m}^n, 1]$. Introduce the increasing function $b_m : [0, 1] \to [0, 1]$ as follows:

$$b_m(t) = \frac{\int_0^t \beta_m(s) ds}{\int_0^1 \beta_m(s) ds}, \quad t \in [0, 1].$$

Let $a_m(t)$ be the inverse of $b_m(t)$, i.e.

$$a_m(t) = \inf\{s \in [0, 1] : b_m(s) > t\}. \tag{23}$$

It is easy to see that condition (7) implies that both $b_m(t)$ and $a_m(t)$ are Lipschitz functions: for any $t_1, t_2 \in [0,1]$ we have

$$|b_m(t_2) - b_m(t_1)| \leq L_{\max} |t_2 - t_1|$$

and

$$|a_m(t_2) - a_m(t_1)| \leq L_{\max} |t_2 - t_1| \tag{24}$$

with $L_{\max} = C_{\max}/C_{\min}$.

Consider the dyadic scheme of partitions $\Delta_{k,j}$, $j = 1, ..., 2^k$, $k = 0, ...M$ of the interval $[0,1]$ as defined by (12). For any $m = 0, ..., M$ denote by $I^m_{k,j}$ the set of indexes $i \in J_m$ for which $b_m(t^m_i)$ falls into $\Delta_{k,j}$, i.e.

$$I^m_{k,j} = \{i \in J_m : b_m(t^m_i) \in \Delta_{k,j}\}, \quad j = 1, ..., 2^k, \quad k = 0, ..., m.$$

Since $\Delta_{k,j} = \Delta_{k+1,2j-1} \cup \Delta_{k+1,2j}$ and $\Delta_{k+1,2j-1} \cap \Delta_{k+1,2j} = \emptyset$, it is clear that $I^m_{k,j} = I^m_{k+1,2j-1} + I^m_{k+1,2j-1}$ for $j = 1, ..., 2^k$. In particular $J_M = I^m_{0,1}$.

It is not hard to see that each set $I^m_{k,j}$ contains at least one element, if we choose n to be large enough.

Note that we have introduced the above sets of indices such that the sequence $X^m_j = X_{\langle m,j \rangle}$, $j \in J_m$ is split into blocks with "almost" the same variances. This turns out to be one of the crucial points in the proof of our results, as we will see later. Indeed, if we set now

$$X^m_{k,j} = \sum_{i \in I^m_{k.j}} X^m_j, \tag{25}$$

then it is easy to see that the following holds true.

Proposition 10 *For any $k = 0, ..., 1$ and $j = 1, ..., 2^k$ we have*

$$\left| E'(X^m_{k,2j-1})^2 - E'(X^m_{k,2j})^2 \right| \leq 2C_{\max}\gamma_n.$$

Proof. Let $\lambda(\Delta_{k,j})$ be the length of the interval $\Delta_{k,j}$ and $\Delta^*_{k,j} = a_m(\Delta_{k,j})$ be the image of $\Delta_{k,j}$ by the map $a_m(t)$. Then

$$\lambda(\Delta_{k,j}) = \frac{\int_{\Delta^*_{k,j}} \beta_m(s) ds}{\int_0^1 \beta_m(s) ds}. \tag{26}$$

By the definition of $\beta_m(s)$ and (7) we have

$$\int_{\Delta^*_{k,j}} \beta_m(s) ds = h_m \sum_{i \in I^m_{k,j}} E' \left(X^m_i\right)^2 + \theta h_m C_{\max}\gamma_n,$$

where $h_m = t^m_i - t^m_{i-1} = 2^{M-m}/n$ and $|\theta| \leq 1$. This and (26) imply

$$h_m^{-1} \int_0^1 \beta_m(s) ds \lambda(\Delta_{k,j}) = \sum_{i \in I^m_{k,j}} E' \left(X^m_i\right)^2 + \theta C_{\max}\gamma_n,$$

from which we easily obtain the assertion if we note that

$$E' \left(X^m_{k,j}\right)^2 = \sum_{i \in I^m_{k,j}} E' \left(X^m_i\right)^2,$$

due to the independence of the r.v.'s X^m_i, $i \in I^m_{k,j}$. ■

5.2 The construction

Recall at this moment that we are given just two sequences of independent r.v.'s: X_i, $i = 1, ..., n$ on the probability space $(\Omega', \mathcal{F}', P')$ and N_i, $i = 1, ..., n$ on the probability space $(\Omega, \mathcal{F}, P)$. We would like to construct on the probability space $(\Omega, \mathcal{F}, P)$ a sequence of independent r.v.'s $\widetilde{X}_i$, $i = 1, ..., n$ such that each $\widetilde{X}_i$ has the same distribution as X_i. We now describe an appropriate version of the Komlós-Major-Tusnády construction.

- *KMT procedure.* Let $\xi_{m,j}$, $j = 1, ..., 2^m$ be a sequence of independent r.v.'s defined on a probability space $(\Omega', \mathcal{F}', P')$ and $\eta_{m,j}$ $j = 1, ..., 2^m$ be a sequence of independent normal r.v.'s on the probability space $(\Omega, \mathcal{F}, P)$. We want to construct a sequence independent r.v.'s $\widetilde{\xi}_{m,j}$, $j = 1, ..., 2^m$ on $(\Omega, \mathcal{F}, P)$, such that $\widetilde{\xi}_{m,j} \overset{d}{=} \xi_{m,j}$, $j = 1, ..., 2^m$. Put $\xi_{k,j} = \xi_{k+1,2j-1} + \xi_{k+1,2j}$, $\eta_{k,j} = \eta_{k+1,2j-1} + \eta_{k+1,2j}$, $j = 1, ..., 2^k$, for $k = 0, ..., m-1$. First define $\widetilde{\xi}_{0,1}$, to be the quantile transform of $\eta_{0,1}$ (see Section 4). Supposing that for some $k = 0, ..., m-1$ we have already constructed the r.v.'s $\widetilde{\xi}_{k,j}$, $j = 1, ..., 2^k$, let $\widetilde{\xi}_{k+1,2j-1}$ be the quantile transform of $\eta_{k+1,2j-1}$ conditional w.r.t. $\widetilde{\xi}_{k,j}$ and $\eta_{k,j}$, for $j = 1, ..., 2^k$ (see Section 4). Finally, let $\widetilde{\xi}_{k+1,2j} = \widetilde{\xi}_{k,j} - \widetilde{\xi}_{k+1,2j-1}$, $j = 1, ..., 2^k$, this completing the KMT procedure.

The following lemma, due to Komlos et al. [11], [12] and Sakhanenko [19], shows that $\widetilde{\xi}_{m,j}$, $j = 1, ..., 2^m$ is the required sequence.

Lemma 11 *For any $k = 0, ..., m$ the r.v.'s $\widetilde{\xi}_{k,j}$, $j = 1, ..., 2^k$ are independent and $\widetilde{\xi}_{k,j} \overset{d}{=} \xi_{k,j}$, $j = 1, ..., 2^k$. Moreover the r.v.'s $\widetilde{\xi}_{k,j}$, $j = 1, ..., 2^k$ are functions of the sequence $\eta_{k,j}$, $j = 1, ..., 2^k$ only.*

Proof. The case with $k = 1$ follows from the Remark 4. Assume that the assertion holds true for some $k \geq 1$ and let us prove it for $k + 1 \leq m$.

For the sake of brevity put $\mathbb{X}_k = \{\widetilde{\xi}_{k,j} : j = 1, ..., 2^k\}$ and $\mathbb{Y}_k = \left\{\eta_{k,j} : j = 1, ..., 2^k\right\}$. Only the independence needs to be checked, the equality in distributions being obvious by Remark 4. Again by Remark 4 it follows that the r.v.'s $\mathbb{X}_{k+1}$ are functions of the r.v.'s $\mathbb{Y}_{k+1}$ only. Note that each pair $(\widetilde{\xi}_{k+1,2j-1}, \widetilde{\xi}_{k+1,2j})$ is a function of the r.v.'s $\widetilde{\xi}_{k,j}$ and $U_{k,j} = \Phi(\eta_{k+1,2j-1}|\eta_{k,j})$ only, while the r.v. $U_{k,j}$ does not depend on the r.v.'s $\mathbb{Y}_k$ and, in particular, on the r.v. $\widetilde{\xi}_{k,j}$. Hence for any real x, y

$$P\left(\widetilde{\xi}_{k+1,2j-1} \leq x,\ \widetilde{\xi}_{k+1,2j} \leq y\,|\mathbb{Y}_k\right) = P\left(\widetilde{\xi}_{k+1,2j-1} \leq x,\ \widetilde{\xi}_{k+1,2j} \leq y\,|\widetilde{\xi}_{k,j}\right). \tag{27}$$

Since the r.v.'s $\mathbb{X}_{k+1}$ are independent conditionally w.r.t. $\mathbb{Y}_k$, taking into account (27) we obtain

$$\begin{aligned} P\left(\prod_{j=1}^{2^{k+1}}\left\{\widetilde{\xi}_{k+1,j} \leq x_j\right\}\right) &= E\prod_{j=1}^{2^k} P\left(\left\{\widetilde{\xi}_{k+1,2j-1} \leq x_{2j-1},\ \widetilde{\xi}_{k+1,2j} \leq x_{2j}\right\}|\mathbb{Y}_k\right) \\ &= E\prod_{j=1}^{2^k} P\left(\left\{\widetilde{\xi}_{k+1,2j-1} \leq x_{2j-1},\ \widetilde{\xi}_{k+1,2j} \leq x_{2j}\right\}|\widetilde{\xi}_{k,j}\right), \end{aligned}$$

for any reals x_j, $j = 1, ..., 2^{k+1}$. By the induction assumption the r.v.'s $\widetilde{\xi}_{k,j}$, $j = 1, ..., 2^k$ are independent, hence

$$P\left(\prod_{j=1}^{2^{k+1}}\left\{\widetilde{\xi}_{k+1,j}\le x_j\right\}\right)=\prod_{j=1}^{2^k}P\left(\widetilde{\xi}_{k+1,2j-1}\le x_{2j-1},\ \widetilde{\xi}_{k+1,2j}\le x_{2j}\right).$$

To complete the proof it suffices to make use of the independence of r.v.'s $\widetilde{\xi}_{k+1,2j-1}$ and $\widetilde{\xi}_{k+1,2j}$. ∎

It turns out that these properties are enough for proving a KMT result if the "indexing" functions belong to the class of indicators. However for proving our functional version of the KMT approximation we need more properties of this construction. To formulate them we introduce the following notations:

$$\begin{aligned}
\beta_{k+1,2j-1} &= \left(\frac{E\eta^2_{k+1,2j}}{E\eta^2_{k+1,2j-1}}\right)^{1/2}, \quad \beta_{k+1,2j}=\left(\frac{E\eta^2_{k+1,2j-1}}{E\eta^2_{k+1,2j}}\right)^{1/2},\\
\widetilde{V}_{k,j} &= \beta_{k+1,2j-1}\eta_{k+1,2j-1}-\beta_{k+1,2j}\eta_{k+1,2j},\\
Z_{k,j} &= \widetilde{\xi}_{k+1,2j-1}-\widetilde{\xi}_{k+1,2j}, \quad \zeta_{k,j}=Z_{k,j}-\widetilde{V}_{k,j}.
\end{aligned}$$

Lemma 12 *For any $k=0,...,m$ the r.v.'s $\zeta_{k,j}$, $j=1,...,2^k$ are independent.*

Proof. The proof is similar to that of Lemma 11. We also keep the same notations. First we note that the r.v.'s $\widetilde{V}_{k,j}$ and $\eta_{k,j}=\eta_{k+1,2j-1}+\eta_{k+1,2j}$ are independent since they are normal and uncorrelated. Obviously each r.v. $\zeta_{k,j}$ is a function of the r.v.'s $\widetilde{\xi}_{k,j}$, $U_{k,j}$ and $\widetilde{V}_{k,j}$ only. Also $U_{k,j}$ and $\widetilde{V}_{k,j}$ do not depend on the r.v.'s $\mathbb{Y}_k$ and, in particular, on the r.v. $\widetilde{\xi}_{k,j}$. Hence for any real x

$$P\left(\zeta_{k,j}\le x|\mathbb{Y}_k\right)=P(\zeta_{k,j}\le x|\widetilde{\xi}_{k,j}).$$

Since r.v.'s $\zeta_{k,j}$ $j=1,...,2^k$ are independent conditionally w.r.t. $\mathbb{Y}_k$

$$P\left(\prod_{j=1}^{2^k}\{\zeta_{k,j}\le x_j\}\right)=E\prod_{j=1}^{2^k}P\left(\zeta_{k,j}\le x_j|\mathbb{Y}_k\right)=E\prod_{j=1}^{2^k}P(\zeta_{k,j}\le x_j|\widetilde{\xi}_{k,j}),$$

for any reals x_j, $j=1,...,2^k$. Now we make use of the independence of the r.v.'s $\widetilde{\xi}_{k,j}$, $j=1,...,2^k$ to obtain the assertion. ∎.

In the sequel we shall need also an auxiliary procedure which is not as powerful as the KMT construction, but permits us to construct *somehow* the components inside an already constructed *arbitrary* sum of independent r.v.'s. Below we present one of the possible methods.

• *Auxiliary construction.* We start from an arbitrary sequence of independent r.v.'s $\xi_1,...,\xi_n$ given on $(\Omega',\mathcal{F}',P')$. Put $S_n=\xi_1+...+\xi_n$. Suppose that on another probability space $(\Omega,\mathcal{F},P)$ we have constructed only the r.v. $\widetilde{S}_n\stackrel{d}{=}S_n$, which corresponds to the sum S_n and we wish to construct its components, i.e. the independent r.v.'s $\widetilde{\xi}_1,...,\widetilde{\xi}_n$ such that $\widetilde{\xi}_1\stackrel{d}{=}\xi_n$ and $\widetilde{S}_n=\widetilde{\xi}_1+...+\widetilde{\xi}_n$. As a prerequisite we assume that on the probability space $(\Omega,\mathcal{F},P)$ we are given independent normal r.v.'s $\eta_1,...,\eta_n$. First we define $\widetilde{\xi}_1$ to be the quantile transform of η_1 conditional w.r.t. $\widetilde{S}_n$ and U_n, where $U_n=\eta_1+...+\eta_n$. If for some $k<n$

the r.v. $\widetilde{\xi}_1, ..., \widetilde{\xi}_{k-1}$ are already constructed, we define $\widetilde{\xi}_k$ to be the quantile transform of η_k conditional w.r.t. $\widetilde{S}_n - \widetilde{\xi}_1 - ... - \widetilde{\xi}_{k-1}$ and $U_n - \eta_1 - ... - \eta_{k-1}$. Finally for $k = n$ we put $\widetilde{\xi}_n = \widetilde{S}_n - \widetilde{\xi}_1 - ... - \widetilde{\xi}_{n-1}$, this completing our procedure.

The easy proof of the following assertion is left to the reader.

Remark 8 *The r.v.'s $\widetilde{\xi}_1, ..., \widetilde{\xi}_n$ just constructed are independent and such that $\widetilde{\xi}_i \stackrel{d}{=} \xi_i$, $i = 1, ..., n$. Moreover the r.v.'s $\widetilde{\xi}_1, ..., \widetilde{\xi}_n$ are functions of the r.v.'s $\eta_1, ..., \eta_n$ and $\widetilde{S}_n$ only.*

The KMT procedure described above allows us to make use of the quantile inequalities proved in Section 4. It should be pointed out however, that in order to get *precise* quantile inequalities, one has to assume the r.v.'s $\xi_{m,i}$, $i \in J_m$ to be in the class $\mathcal{D}_0(r)$ for some $r > 0$ (see Section 4 for more details) or to be identically distributed (see Komlós etc. [11], [12]). We will avoid such type of assumptions by using a construction which goes back to the paper of Sakhanenko [19]. The idea is to substitute the initial sequence with some smoothed sequences and to apply the above KMT procedure to them. We proceed to describe formally this construction. Consider the product probability space $(\Omega'', \mathcal{F}'', P'') = (\Omega', \mathcal{F}', P') \times (\Omega, \mathcal{F}, P)$ on which sequences X_i, $i = 1, ..., n$ and N_i, $i = 1, ..., n$ are independent.

- *M-th step.* For any $i \in J_M$ put

$$X_i^M = X_i, \quad W_i^M = N_i \tag{28}$$

and

$$Y_i^M = \begin{cases} X_i^M, & \text{if } i \text{ is odd,} \\ W_i^M, & \text{if } i \text{ is even.} \end{cases} \tag{29}$$

The meaning of these notations is: X_i^M, $i \in J_M$ is the sequence X_i, $i \in J_M$ which we wish to construct on $(\Omega, \mathcal{F}, P)$, W_i^M, $i \in J_M$ is the corresponding sequence of normal r.v.'s given on $(\Omega, \mathcal{F}, P)$ and Y_i^M, $i \in J_M$ is the smoothed sequence which we will construct at this step. First we split the sequences Y_i^M, $i \in J_M$ and W_i^M, $i \in J_M$ into blocks as follows: for $j = 1, ..., 2^k$ and $k = 0, ..., M$ put

$$Y_{k,j}^M = \sum_{i \in I_{k,j}^M} Y_i^M, \quad W_{k,j}^M = \sum_{i \in I_{k,j}^M} W_i^M.$$

Then obviously for $j = 1, ..., 2^k$ and $k = 0, ..., M-1$

$$\begin{aligned} Y_{k,j}^M &= Y_{k+1,2j-1}^M + Y_{k+1,2j}^M, \\ W_{k,j}^M &= W_{k+1,2j-1}^M + W_{k+1,2j}^M. \end{aligned}$$

We are now prepared to apply the KMT procedure as described above with $m = M$, $\xi_{m,i} = Y_{M,j}^M$ and $\eta_{m,j} = W_{M,j}^M$, $j = 1, ..., 2^M$ to construct a sequence of independent r.v.'s $\widetilde{Y}_{M,j}^M$, $j = 1, ..., 2^M$ such that $\widetilde{Y}_{M,j}^M \stackrel{d}{=} Y_{M,j}^M$, $j = 1, ..., 2^M$. For this let $\widetilde{Y}_{0,1}^M$ be the quantile transform of $W_{0,1}^M$. Having defined $\widetilde{Y}_{k,j}^M$, $j = 1, ..., 2^k$ for some $k = 0, ..., M-1$, let $\widetilde{Y}_{k+1,2j-1}^M$ be the quantile transform of $W_{k+1,2j-1}^M$ conditional w.r.t. $\widetilde{Y}_{k,j}^M$ and $W_{k,j}^M$, for $j = 1, ..., 2^k$. For even indexes $2j$ we put $\widetilde{Y}_{k+1,2j}^M = \widetilde{Y}_{k,j}^M - \widetilde{Y}_{k+1,2j-1}^M$, $j = 1, ..., 2^k$. By Lemma 11 the r.v.'s $\widetilde{Y}_{M,j}^M$, $j = 1, ..., 2^M$ are independent and such that

$$\widetilde{Y}_{M,j}^M \stackrel{d}{=} Y_{M,j}^M, \quad j = 1, ..., 2^M.$$

It remains to construct the components inside each sum $\widetilde{Y}^M_{M,j}$. For this we make use of the auxiliary procedure for an arbitrary sequence presented above. For each fixed $j = 1, ..., 2^M$ this construction provides a sequence of r.v.'s $\widetilde{Y}^M_i$, $i \in I^M_{M,j}$ such that

$$\widetilde{Y}^M_{M,j} = \sum_{i \in I^M_{M,j}} \widetilde{Y}^M_i$$

and

$$\widetilde{Y}^M_i \stackrel{d}{=} Y^M_i, \quad i \in I^M_{M,j}. \tag{30}$$

Moreover each r.v. $\widetilde{Y}^M_i$ is a function of $\widetilde{Y}^M_{M,j}$ and W^M_i, $i \in I^M_{M,j}$ only. This completes the initial step of our construction.

Let us remark that actually we have constructed only half of the initial sequence, that is we have constructed the r.v.'s X_i only for odd $i \in J_M$. In order to construct the second part of the sequence, we will repeat the same procedure. More generally, we proceed now to describe formally the m-th step of our construction.

- *m-th step.* Suppose that for some $m = M-1, ..., 0$ we have already constructed the r.v.'s Y^{m+1}_i, W^{m+1}_i and $\widetilde{Y}^{m+1}_i$, $i \in J_{m+1}$. Then we define X^m_i, W^m_i and Y^m_i, for $i \in J_m$, as follows

$$X^m_i = X^{m+1}_{2i}, \quad W^m_i = \widetilde{Y}^{m+1}_{2i}, \tag{31}$$

and

$$Y^m_i = \begin{cases} X^m_i, & \text{if} \quad i \quad \text{is odd}, \\ W^m_i, & \text{if} \quad i \quad \text{is even}. \end{cases} \tag{32}$$

The meaning of these notations is: X^m_i, $i \in J_m$ is the part of the sequence X_i, $i \in J_M$ which is not yet constructed, W^m_i, $i \in J_m$ is the corresponding sequence of normal r.v.'s given on $(\Omega, \mathcal{F}, P)$ and Y^m_i, $i \in J_m$ is the smoothed sequence which we will construct at this step. First we split the sequences Y^m_i, $j \in J_m$ and W^m_i, $j \in J_m$ into blocks as follows: for $j = 1, ..., 2^k$ and $k = 0, ..., m$ put

$$Y^m_{k,j} = \sum_{i \in I^m_{k,j}} Y^m_i, \quad W^m_{k,j} = \sum_{i \in I^m_{k,j}} W^m_i.$$

Then obviously for $j = 1, ..., 2^k$ and $k = 0, ..., m-1$

$$\begin{aligned} Y^m_{k,j} &= Y^m_{k+1,2j-1} + Y^m_{k+1,2j}, \\ W^m_{k,j} &= W^m_{k+1,2j-1} + W^m_{k+1,2j}. \end{aligned}$$

We will apply the KMT procedure with $\xi_{m,j} = Y^m_{m,j}$, and $\eta_{m,j} = W^m_{m,j}$, $j = 1, ..., 2^m$ to construct a sequence of independent r.v.'s $\widetilde{Y}^m_{m,j}$, $j = 1, ..., 2^m$ satisfying $\widetilde{Y}^m_{m,j} \stackrel{d}{=} Y^m_{m,j}$, $j = 1, ..., 2^m$. Let $\widetilde{Y}^m_{0,1}$ be the quantile transform of $W^m_{0,1}$. Having defined $\widetilde{Y}^m_{k,j}$, $j = 1, ..., 2^k$ for some $k = 0, ..., m-1$, let $\widetilde{Y}^m_{k+1,2j-1}$ be the quantile transform of $W^m_{k+1,2j-1}$ conditionally w.r.t. $\widetilde{Y}^m_{k,j}$ and $W^m_{k,j}$, for $j = 1, ..., 2^k$. For even indexes $2j$ we put $\widetilde{Y}^m_{k+1,2j} = \widetilde{Y}^m_{k,j} - \widetilde{Y}^m_{k+1,2j-1}$, $j = 1, ..., 2^k$. By Lemma 11 we have that the r.v.'s $\widetilde{Y}^m_{m,j}$, $j = 1, ..., 2^m$ are independent and such that

$$\widetilde{Y}^m_{m,j} \stackrel{d}{=} Y^m_{m,j}, \quad j = 1, ..., 2^m.$$

It remains to construct the components inside each sum $\widetilde{Y}_{m,j}^m$, $j = 1, ..., 2^m$. Again we make use of the auxiliary construction described above. For each fixed $j = 1, ..., 2^m$ it provides a sequence of r.v.'s $\widetilde{Y}_i^m$, $i \in I_{m,j}^m$ such that

$$\widetilde{Y}_{m,j}^m = \sum_{i \in I_{m,j}^m} \widetilde{Y}_i^m$$

and

$$\widetilde{Y}_i^m \stackrel{d}{=} Y_i^m, \quad i \in I_{m,j}^m. \tag{33}$$

Moreover each r.v. $\widetilde{Y}_i^m$ is a function of $\widetilde{Y}_{m,j}^m$ and W_i^m, $i \in I_{m,j}^m$ only. This completes the m-th step of our construction.

5.3 Some useful properties

Let us discuss some properties of the r.v.'s introduced above. In analogy to X_j^m (see Section 5.1) set $N_j^m = N_{\langle m,j \rangle}$, where $m = 0, ..., M$, $j \in J_m$ and $\langle m,j \rangle = j2^{M-m}$.

Proposition 13 *For all $m = 0, ..., M$ and $i \in J_m$*

$$W_i^m \stackrel{d}{=} N_i.$$

Proof. Indeed by (28) $W_j^M = N_j$, $j \in J_M$ and by (31) $W_j^m = \widetilde{Y}_{2j}^{m+1} \stackrel{d}{=} W_{2j}^{m+1}$, if $m = 1, ..., M-1$. The last equality in distribution is due to (33), (30) and (32), (29) for $i = 2j$ even. A simple recursion argument completes the proof. ■

Proposition 14 *For all $m = 0, ..., M$ the r.v.'s $\widetilde{Y}_i^m$, $i \in J_m$ are independent and such that for any $i \in J_m$*

$$\widetilde{Y}_i^m \stackrel{d}{=} \begin{cases} X_i^m, & \text{if } i \text{ is odd}, \\ N_i^m, & \text{if } i \text{ is even}. \end{cases}$$

Proof. The independence follows from the Lemma 11. Next, it follows easily from (30) and (29) in the case $m = M$, and from (33) and (32) for $m = 0, ..., M-1$ that $\widetilde{Y}_i^m \stackrel{d}{=} Y_i^m = W_i^m$ if i is even and $\widetilde{Y}_i^m \stackrel{d}{=} Y_i^m = X_i^m$ if i is odd, $i \in J_m$. It remains only to apply the previous Proposition 13. ■

Proposition 15 *The vectors $\left\{ \widetilde{Y}_i^m : i \in J_m, \, i\text{-odd} \right\}$, $m = M, ..., 0$ are independent.*

Proof. It is easy to see that according to Lemma 11 and Remark 8 the r.v.'s $\widetilde{Y}_i^m$, $i \in J_m$ are functions of the normal r.v.'s W_i^m, $i \in J_m$ only. By (31) this implies that the r.v. $\widetilde{Y}_i^m$, $i \in J_m$ depend only on the vector $\left\{ \widetilde{Y}_k^{m+1} : k \in J_{m+1}, \, k\text{-even} \right\}$. Since $\widetilde{Y}_k^{m+1}$, $k \in J_{m+1}$ is a sequence of independent r.v.'s, this means that the vector $\left\{ \widetilde{Y}_i^m : i \in J_m \right\}$ does not depend on the vector $\left\{ \widetilde{Y}_k^{m+1} : k \in J_{m+1}, \; k\text{-odd} \right\}$. Applying this recursively we get that $\left\{ \widetilde{Y}_k^m : k \in J_m, \; k\text{-odd} \right\}$, $m = M, ..., 0$ is a sequence of independent random vectors. ■

The desired sequence $\widetilde{X}_i,\ i=1,...,n$ can be constructed on the probability space $(\Omega,\mathcal{F},P)$ in the following way. For any $i=1,...,n$ let (m,j) be the unique pair such that $i=j2^{M-m}$, where $0\le m\le M$ and j is odd in J_m. Then we set

$$\widetilde{X}_i=\widetilde{Y}_j^m. \tag{34}$$

Proposition 16 *The r.v.'s $\widetilde{X}_i,\ i=1,...,n$ are independent and such that*

$$\widetilde{X}_i\overset{d}{=}X_i,\quad i=1,...,n.$$

Proof. It is clear that

$$\left\{\widetilde{X}_i:i=1,...,n\right\}=\cup_{k=0}^m\left\{\widetilde{Y}_j^m:j\in J_m,\ j\text{ -odd}\right\}.$$

Let J_m^1 be the set of odd numbers in J_m. Note that for any fixed $m\in\{0,...,M\}$ by Proposition 14 the r.v.'s $\widetilde{Y}_j^m,\ j\in J_m^1$ are independent and such that $\widetilde{Y}_j^m\overset{d}{=}X_j^m,\ j\in J_m^1$. The assertion is immediate if we note that the sequences $\widetilde{Y}_j^m,\ j\in J_m^1$ are independent for different m by Proposition 15. ■

The following elementary representation is essential in the proof of our results.

Proposition 17 *For any function $f(t):[0,1]\to\mathbf{R}^1$*

$$\sum_{i=1}^n f(t_i)\left(\widetilde{X}_i-N_i\right)=\sum_{m=0}^M\sum_{i\in J_m}f(t_i^m)\left(\widetilde{Y}_i^m-W_i^m\right).$$

Proof. Put for brevity

$$\widetilde{X}_i^m=\widetilde{X}_{\langle m,j\rangle}, \tag{35}$$

where $\langle m,j\rangle=j2^{M-m},\ j=1,...,2^m,\ m=0,...,M$ and

$$S^m=\sum_{i\in J_m}f(t_i^m)\left(\widetilde{X}_i^m-W_i^m\right).$$

We will show that for any $m=M,...,0$

$$S^m=\sum_{i\in J_m^1}f(t_i^m)\left(\widetilde{Y}_i^m-W_i^m\right)+S^{m-1}, \tag{36}$$

where $S_{-1}=0$. Fix an $m\in\{0,...,M\}$. By (35) and (34) $\widetilde{X}_i^m=\widetilde{Y}_i^m$ for any $i\in J_m^1$. Hence

$$S^m=\sum_{i\in J_m^1}f(t_i^m)\left(\widetilde{Y}_i^m-W_i^m\right)+\sum_{i\in J_m^2}f(t_i^m)\left(\widetilde{X}_i^m-\widetilde{Y}_i^m\right), \tag{37}$$

where J_m^2 is the set of even indexes in J_m. It is easy to see that $i=2j\in J_m^2$ if and only if $j\in J_{m-1}$ and that $t_i^m=t_j^{m-1}$. Moreover for $i=2j\in J_m^2$ we have by (35) and (34) $\widetilde{X}_i^m=\widetilde{X}_j^{m-1}$, while by (31) $\widetilde{Y}_i^m=W_j^{m-1}$. Then the last sum on the right-hand side of (37) equals S^{m-1}, this proving (36). Next, since by (35) $\widetilde{X}_i=\widetilde{X}_i^M$ and by (28) $N_i=W_i^M$ for $i\in J_M=\{1,...,n\}$, it is obvious that

$$\sum_{i=1}^n f(t_i^m)\left(\widetilde{X}_i-N_i\right)=\sum_{i\in J_M}f(t_i^m)\left(\widetilde{X}_i^M-W_i^M\right)=S^M. \tag{38}$$

The assertion of the lemma follows from (38) and (36). ■

5.4 Quantile inequalities

In analogy to (25) put for $m = 0, ..., M,\ k = 0, ..., m$ and $j \in J_k$

$$\widetilde{Y}_{k,j}^m = \sum_{i \in I_{k,j}^m} \widetilde{Y}_i^m. \tag{39}$$

The following lemma shows that the r.v.'s $\widetilde{Y}_{k,j}^m$, $j \in J_k$ are smooth enough to allow application of the quantile inequalities in Section 4.

Lemma 18 *For $m = 0, ..., M,\ k = 0, ..., m-1,\ j = 1, ..., 2^k$ the r.v. $\widetilde{Y}_{k,j}^m$ is in the class $\mathcal{D}_0(r)$, for some constant $r > 0$ depending on $C_{\min}, C_{\max}, \lambda_0$.*

Proof. We check the conditions of Proposition 9. Toward this end fix m, k, j as in the condition of the lemma and note that

$$\zeta_0 \equiv \widetilde{Y}_{k,j}^m = \sum_{i \in I_{k,j}^m} \widetilde{Y}_i^m = \sum_{i \in I_1} \widetilde{Y}_i^m + \sum_{i \in I_2} \widetilde{Y}_i^m \equiv \zeta_1 + \zeta_2,$$

where I_1 and I_2 are the sets of all odd and even indexes in $I_{k,j}^m$ respectively. By Proposition 14 we have $\widetilde{Y}_i^m \stackrel{d}{=} N_i$ for any $i \in I_2$. Thus the r.v. ζ_2 is actually a sum of independent normal r.v.'s. Note that condition $k \le m-1$ assures that for n large enough the set $I_{k,j}^m$ has at least two elements, from which we conclude that I_2 has at least one element. Next taking into account (7) and the obvious inequality $\#I_2 \ge \frac{1}{3}\#I_{k,j}^m$ we get

$$E\zeta_2^2 \ge C_{\min}\gamma_n\#I_2 \ge \frac{C_{\min}}{3}\gamma_n\#I_{k,j}^m \ge c_1 E\zeta_0^2,$$

with $c_1 = C_{\min}/(3C_{\max})$. For $|h| \le \lambda_0$ and $t \in R$ let

$$f_{\zeta_i,h}(t) = E\exp\{(h + \sqrt{-1}\,t)\zeta_i\}/E\exp\{h\zeta_i\}$$

be the conjugate characteristic function of the r.v. ζ_i, $i = 0, 1, 2$. Since ζ_1 and ζ_2 are independent and ζ_2 is normal

$$\begin{aligned} |f_{\zeta_0,h}(t)| &= |f_{\zeta_1,h}(t)\ f_{\zeta_2,h}(t)| \le |f_{\zeta_2,h}(t)| \\ &\le \exp\{-\frac{t^2}{2}E\zeta_2^2\} \le \exp\{-\frac{t^2}{2}c_1E\zeta_0^2\}, \end{aligned}$$

for $|h| \le \lambda_0$, $t \in \mathbf{R}^1$. With this bound we have for any $\delta > 0$

$$\int_{|t|>\delta} |f_{\zeta_0,h}(t)|dt \le \int_{|t|>\delta} \exp\{-\frac{t^2}{2}c_1E\zeta_0^2\}dt \le c_2(E\zeta_0^2)^{-3/2},$$

where c_2 is a constant depending only on $C_{\min}$, $C_{\max}$ and δ. This proves that the condition of Proposition 9 is satisfied. It remains only to show that conditions (17) and (18) are satisfied. This follows from (7) and (8) as soon as $\widetilde{Y}_i^m \stackrel{d}{=} X_i$ or $\widetilde{Y}_i^m \stackrel{d}{=} N_i$ for any $i \in I_{k,j}^m \subseteq J_m$ by Proposition 14. Here we also make use of the elementary fact that Sakhanenko's condition (18) holds true for any normal r.v. . ■

For any $m=1,...,M,\ k=0,...,m$ and $j\in J_k$ put

$$\begin{aligned} S^m_{k,j} &= \widetilde{Y}^m_{k,j} - W^m_{k,j}, \\ B^m_{k,j} &= E\left(\widetilde{Y}^m_{k,j}\right)^2, \\ \widetilde{Y}^{m,*}_{k,j} &= \widetilde{Y}^m_{k,j}\mathbf{1}\left(\left|\widetilde{Y}^m_{k,j}\right| \le B^m_{k,j}\right). \end{aligned} \tag{40}$$

Introduce the sets

$$G^m_{0,1} = \left\{\left|\widetilde{Y}^m_{0,1}\right| \le c_0 B^m_{0,1}\right\}$$

and

$$G^m_{k,j} = \left\{\left|\widetilde{Y}^m_{k,j}\right| \le c_0 B^m_{k,j},\ \left|W^m_{k,j}\right| \le c_0 B^m_{k,j}\right\}, \tag{41}$$

for $k=1,....,m$, where $c_0 = C_{\min}/(72C_{\max})$.

The following quantile inequalities are crucial in the proof of our results.

Lemma 19 *On the set* $G_{0,1}$

$$\left|S^m_{0,1}\right| \le c_1\left\{1 + \frac{\left(\widetilde{Y}^{m,*}_{0,1}\right)^2}{B^m_{0,1}}\right\},$$

where c_1 *is a positive constant depending only on* $C_{\max}$, $C_{\min}$, λ_0.

Proof. It is enough to note that by Lemma 18 the r.v. $\widetilde{Y}^m_{0,1}$ is in the class $\mathcal{D}_0(r)$ for some $r>0$ depending on $C_{\max}$, $C_{\min}$, λ_0 and to apply Remark 6 and then Lemma 7 with $\widetilde{X} = \widetilde{Y}^m_{0,1}$, $X = Y^m_{0,1}$ and $N = W^m_{0,1}$. ■

Lemma 20 *Let* $m=0,...,M,\ k=1,...,m-1,\ j\in J_{k-1}$. *On the set* $G^m_{k,2j-1}\cap G^m_{k,2j}$

$$\left|S^m_{k,2j-1} - S^m_{k,2j}\right| \le c_1\left\{1 + \frac{\left(\widetilde{Y}^{m,*}_{k,2j-1}\right)^2}{B^m_{k,2j-1}} + \frac{\left(\widetilde{Y}^{m,*}_{k,2j}\right)^2}{B^m_{k,2j}}\right\},$$

where c_1 *is a positive constant depending only on* $C_{\max}$, $C_{\min}$, λ_0.

Proof. Fix m, k, and j as in the condition of the lemma. We will make use of Lemma 8 with

$$\widetilde{X}_1 = \widetilde{Y}^m_{k,2j-1}, \quad \widetilde{X}_2 = \widetilde{Y}^m_{k,2j}, \quad \widetilde{X}_0 = \widetilde{Y}^m_{k-1,j}, \tag{42}$$

and

$$N_1 = W^m_{k,2j-1}, \quad N_2 = W^m_{k,2j}, \quad N_0 = W^m_{k-1,j}. \tag{43}$$

By Lemma 18 r.v.'s $\widetilde{X}_0$, $\widetilde{X}_1$ and $\widetilde{X}_2$ are in the class $\mathcal{D}_0(r)$ for some $r>0$ depending only on $C_{\min}$, $C_{\max}$, λ_0.

By Proposition 14 we have $B^m_{k,l} = \sum_{i\in I^m_{k,l}} E(X^m_i)^2$ for $l\in J_m$. Note that the set $I^m_{k,l}$ contains at least one element, i.e $n^m_{k,l} = \# I^m_{k,l} \ge 1$. Then by (7) $B^m_{k,l} \ge C_{\min}\gamma_n$ for $l\in J_m$. By

Proposition 10 we get $B^m_{k,2j-1} \le B^m_{k,2j} + 2C_{\max}\gamma_n \le c_3 B^m_{k,2j}$, where $c_3 = 3C_{\max}/C_{\min}$. In the same way we get $B^m_{k,2j} \le c_3 B^m_{k,2j-1}$. Using these inequalities we arrive at

$$B^2 = \frac{B^m_{k,2j-1}B^m_{k,2j}}{B^m_{k,2j-1}+B^m_{k,2j}} \ge \frac{1}{6}\frac{C_{\min}}{C_{\max}} \max\{B^m_{k,2j-1}, B^m_{k,2j}\}. \tag{44}$$

Now we can check the conditions of Lemma 8. Indeed, by (44) on the set $J_{k,2j-1} \cap J_{k,2j}$

$$|\widetilde{X}_i| \le B^2/12$$

for $i = 1, 2$ and thus

$$|\widetilde{X}_0| \le |\widetilde{X}_1| + |\widetilde{X}_2| \le B^2/6.$$

Now Lemma 8 and Remark 6 imply

$$\left|S^m_{k,2j-1} - S^m_{k,2j-1}\right| \le |\alpha_1 - \alpha_2| \left|\widetilde{X}_0 - N_0\right| + c_2 \left\{1 + B^{-2}\left(\widetilde{X}_1^2 + \widetilde{X}_2^2 + \widetilde{X}_0^2\right)\right\}, \tag{45}$$

where $\alpha_i = EX_i^2/EX_0^2$ and c_2 is some constant depending only on $C_{\min}$, $C_{\max}$, λ_0. From Proposition 10 it follows that

$$|\alpha_1 - \alpha_2| = \frac{\left|E'X_1^2 - E'X_2^2\right|}{E'X_0^2} \le \frac{2C_{\max}\gamma_n}{B^m_{k+1,j}} \le \frac{2C_{\max}}{B^m_{k+1,j}}, \tag{46}$$

while on the set $G^m_{k,2j-1} \cap G^m_{k,2j}$ we have

$$\left|\widetilde{X}_0 - N_0\right| \le \left|\widetilde{X}_1 - N_1\right| + \left|\widetilde{X}_2 - N_2\right| \le 2c_0 B^m_{k+1,j}. \tag{47}$$

The assertion of the lemma can easily be obtained from (45), (44), (46) and (47). ■

6 Proof of the main results

6.1 Auxiliary statements

We keep the same notations as in the previous section.

Lemma 21 *Let $m = 0, ..., M$, $k = 0, ..., m$ and $j \in J_k$. For any $c_0 > 0$ there is a constant c_1 depending only on λ_0 and c_0 such that*

$$P\left(\left|\widetilde{Y}^m_{k,j}\right| > c_0 B^m_{k,j}\right) \le 2\exp\left\{-c_1 B^m_{k,j}\right\}$$

and

$$P\left(\left|W^m_{k,j}\right| > c_0 B^m_{k,j}\right) \le 2\exp\left\{-c_1 B^m_{k,j}\right\}.$$

Proof. By the Chebyshev inequality we have with $t > 0$

$$P\left(\widetilde{Y}^m_{k,j} > tB^m_{k,j}\right) \le \exp\left\{-tB^m_{k,j}\right\} E\exp\{tY^m_{k,j}\}. \tag{48}$$

Since $\widetilde{Y}^m_{k,j}$ is a sum of independent r.v.'s $\widetilde{Y}^m_i$, $i \in I^m_{k,j}$ by (39), by (8) and Lemma 25 we obtain for $|t| \le \lambda_0/2$

$$E\exp\left\{t\widetilde{Y}^m_{k,j}\right\} = \prod_{i\in I^m_{k,j}} E\exp\left\{t\widetilde{Y}^m_i\right\} \le \prod_{i\in I^m_{k,j}} \exp\left\{t^2 E(\widetilde{Y}^m_i)^2\right\} = \exp\left\{t^2 B^m_{k,j}\right\}.$$

Choosing $t = t_0 = \min\{c_0/(2c_2), \lambda_0/2\}$ and inserting this bound into (48), we obtain

$$E\left(\widetilde{Y}_{k,j}^m > c_0 B_{k,j}^m\right) \le \exp\left\{-t_0(c_0 - t_0 c_2) B_{k,j}^m\right\} \le \exp\left\{-c_1 B_{k,j}^m\right\},$$

where $c_1 = t_0 c_0/4$. In the same way one can show that

$$E(\widetilde{Y}_{k,j}^m < -c_0 B_{k,j}^m) \le \exp\{-c_1 B_{k,j}^m\},$$

which together with the previous bound proves the first assertion of the lemma. The second claim is straightforward since the r.v. $W_{k,j}^m$ is normal. ■

Recall that the r.v's $S_{k,j}^m$ are defined by (40).

Lemma 22 *There are two positive constants c_0 and c_1 depending only on $C_{\min}$, $C_{\max}$ and λ_0 such that for any $m = 0, ..., M,\ k = 0, ..., m-1,\ j = 1, ..., 2^k$*

$$E\exp\left\{c_0 \left|S_{k,j}^m\right|\right\} \le c_1.$$

Proof. Fix m, k and j as in the condition of the lemma. Let $t = \lambda_0/4$, with λ_0 from the condition (8). It is easy to see that

$$E\exp\left\{t\left|S_{k,j}^m\right|\right\} = Q_1 + Q_2,$$

where

$$\begin{aligned} Q_1 &= E\exp\left\{t\left|S_{k,j}^m\right|\right\}\mathbf{1}\left(G_{k,2j-1}^{m,c} \cup G_{k,2j-1}^{m,c}\right), \\ Q_2 &= E\exp\left\{t\left|S_{k,j}^m\right|\right\}\mathbf{1}\left(G_{k,2j-1}^{m} \cap G_{k,2j-1}^{m}\right), \end{aligned} \tag{49}$$

the set $G_{k,l}^m$ being defined by (41) and $G_{k,l}^{m,c}$ being the complement of the set $G_{k,l}^m$, $l = 2j-1, 2j$.

First we give an estimate for Q_1. Applying the Hölder inequality we get from (49)

$$Q_1 \le \left(\exp\left\{2t\left|S_{k,j}^m\right|\right\}\right)^{1/2}\left(P\left(G_{k,2j-1}^{m,c}\right)^{1/2} + P\left(G_{k,2j}^{m,c}\right)^{1/2}\right). \tag{50}$$

By Lemma 21 we have with $l = 2j-1, 2j$

$$P\left(G_{k,l}^{m,c}\right) = P\left(\left|\widetilde{Y}_{k,l}^m\right| > c_2 B_{k,l}^m\right) + P\left(\left|W_{k,l}^m\right| > c_2 B_{k,l}^m\right) \le 2\exp\left\{-c_3 B_{k,l}^m\right\}, \tag{51}$$

where $c_2 = C_{\min}/(76 C_{\max})$ and c_3 depends only on c_2 and λ_0.

On the other hand from (40) and from the Hölder inequality we get

$$E\exp\left\{2t|S_{k,j}^m|\right\} \le \left(E\exp\left\{4t|\widetilde{Y}_{k,j}^m)|\right\} E\exp\left\{4t|W_{k,j}^m|\right\}\right)^{1/2}. \tag{52}$$

Since $\widetilde{Y}_{k,j}^m$ is exactly the sum of the independent r.v.'s $\widetilde{Y}_i^m$, $i \in I_{k,j}^m$,

$$\begin{aligned} E\exp\left\{4t|\widetilde{Y}_{k,j}^m|\right\} &\le E\exp\left\{4t\widetilde{Y}_{k,j}^m\right\} + E\exp\left\{-4t\widetilde{Y}_{k,j}^m\right\} \\ &\le \prod_{i\in I_{k,j}^m} E\exp\left\{4t\widetilde{Y}_i^m\right\} + \prod_{i\in I_{k,j}^m} E\exp\left\{-4t\widetilde{Y}_i^m\right\}. \end{aligned}$$

Taking into account (8) and $4t \leq \lambda_0/3$ by Lemma 25 we arrive at

$$E \exp\left\{4t|X_{k,j}^m|\right\} \leq 2 \prod_{i \in J_{k,j}^m} E \exp\left\{16t^2 E(X_i^m)^2\right\} \leq 2 \exp\left\{16t^2 B_{k,j}^m\right\}. \tag{53}$$

A similar bound holds for the second expectation in the right-hand side of (52), namely

$$E \exp\left\{4t|W_{k,j}^m|\right\} \leq 2 \exp\left\{16t^2 B_{k,j}^m\right\}. \tag{54}$$

Inserting the inequalities (54) and (53) into (52) and then (52) and (51) into (50) and choosing t to satisfy the inequality $t \leq t_0 = \min\{c_3/32, \lambda_0/4\}$, we arrive at

$$Q_1 \leq 4 \exp\left\{\left(8t^2 - \frac{1}{2}c_3\right) B_{k,j}^m\right\} \leq 4 \exp\left\{-\frac{1}{4}c_3 B_{k,j}^m\right\} \leq 4. \tag{55}$$

Now we proceed to give a bound for Q_2. By virtue of Lemma 20 we have that on the set $G_{k,2j-1}^m \cap G_{k,2j}^m$ with some constant and c_4 depending only on $C_{\max}$, $C_{\min}$, λ_0

$$|S_{k,j}^m| \leq c_4 \left\{1 + U_{k,2j-1}^m + U_{k,2j}^m\right\}, \tag{56}$$

for $k \geq 1$, where we denote $U_{k,l}^m = (\widetilde{Y}_{k,l}^{m,*})^2 / B_{k,l}^m$, $l = 2j-1, 2j$. Similarly, by Lemma 19 we have

$$|S_{0,1}^m| \leq c_5(1 + U_{0,1}^m),$$

where $U_{0,1}^m = (\widetilde{Y}_{0,1}^{m,*})^2 / B_{0,1}^m$ and c_5 is a constant depending only on $C_{\max}$, $C_{\min}$, λ_0.

If $k \geq 1$, then according to (56) and by the Hölder inequality

$$\begin{aligned} Q_2 &\leq E \exp\left\{tc_4\left(1 + U_{k,2j-1}^m + U_{k,2j}^m\right)\right\} \\ &\leq \exp\{tc_4\}\left(E \exp\left\{2tc_4 U_{k,2j-1}^m\right\}\right)^{1/2}\left(E \exp\left\{2tc_4 U_{k,2j}^m\right\}\right)^{1/2}. \end{aligned} \tag{57}$$

By Lemma 27 for some constant c_6 depending only on λ_0 we have

$$E \exp\left\{c_6 U_{k,2j-1}^m\right\} \leq (1 + 2/c_6) \tag{58}$$

and a similar bound holds true for $U_{k,2j}^m$. If we take t to be such that $tc_4 \leq \min\{c_6, t_0\}$, then from (57) and (58) we obtain

$$Q_2 \leq \exp\{c_6\}(1 + 2/c_6)^2. \tag{59}$$

The case with $k = 0$ is similar. Combining the estimates for Q_1 and Q_2 given by (55) and (59) we obtain the assertion of the lemma. ■

6.2 Proof of Theorem 3

For the sake of brevity put

$$S_n(f) = \sum_{i=1}^n f(t_i)\left(\widetilde{X}_i - N_i\right).$$

We have to show is that there are two positive constants c_0 and c_1, depending only on $C_{\min}$, $C_{\max}$, λ_0, L, such that for any t satisfying $|t| \leq c_0$

$$E \exp\{t \frac{1}{\log^2 n} S_n(f)\} \leq \exp\left\{t^2 c_1\right\}. \tag{60}$$

Toward this end let $M = [\log_2 n]$ and note that according to Proposition 17

$$S_n(f) = \sum_{m=0}^{M} S^m,$$

where

$$S^m = \sum_{i \in J_m} f(t_i^m) \left(\widetilde{Y}_i^m - W_i^m \right).$$

By the Hölder inequality

$$E \exp \left\{ t \frac{1}{\log^2 n} S_n(f) \right\} \leq \prod_{m=0}^{M} \left(E \exp \left\{ t(M+1) \frac{1}{\log^2 n} S^m \right\} \right)^{1/(M+1)}. \tag{61}$$

Put for brevity

$$u_n = (M+1)/\log^2 n. \tag{62}$$

Obviously $u_n \leq 1$ for n such that $\log n \geq 2$.

It is easy to see that inequality (60) will follow from (61) if we prove that constants c_0 and c_1 can be chosen so that for any t satisfying $|t| \leq c_0$

$$E \exp \{ t u_n S^m \} \leq \exp \left\{ t^2 c_1 \right\}, \tag{63}$$

for $m = 0, ..., M$. In the sequel we will give a proof of (63).

First we consider the case $m = 0, 1$. If we choose the constant c_0 to be $c_0 = \lambda_0/(6L)$, then it is easy to see that $|2tu_nL| \leq \lambda_0/3$ and thus by Lemma 25 we have for $i \in J_m$

$$E \exp\{2tu_n L \widetilde{Y}_i^m\} \leq \exp \left\{ 4t^2 L^2 E(\widetilde{Y}_i^m)^2 \right\}. \tag{64}$$

An analogous bound holds true for the r.v.'s W_i^m, $i \in J_m$:

$$E \exp\{2tu_n L W_i^m\} \leq \exp \left\{ 2t^2 L^2 E(W_i^m)^2 \right\}. \tag{65}$$

By the Hölder inequality

$$E \exp \{ t u_n S^m \} \leq \left(E \exp\{2tu_n \sum_{j \in J_m} f(t_i^m) \widetilde{Y}_i^m\} E \exp\{2tu_n \sum_{j \in J_m} f(t_i^m) W_i^m\} \right)^{1/2}. \tag{66}$$

Using the independence of the r.v.'s $\widetilde{Y}_i^m$, $j \in J_m$ and W_i^m, $j \in J_m$ and the inequality $\|f\|_\infty \leq L/2$, we obtain from (64), (65) and (66)

$$E \exp \{ t u_n S^m \} \leq \exp \left\{ \frac{3}{4} t^2 L^2 \sum_{i \in J_m} E (X_i^m)^2 \right\}. \tag{67}$$

Since for $m = 0, 1$ the set J_m has cardinality less than 3, by (7) we have

$$\sum_{i \in J_m} E (X_i^m)^2 \leq \# J_m C_{\max} \leq 3 C_{\max}. \tag{68}$$

Hence (63) follows from (67) and (68), provided $m = 0, 1$.

For the case $m \geq 2$ introduce the function $g(s) = f(a(s))$, $s \in [0,1]$, where $a(s)$ is defined by (23). Put for brevity $s_i^m = b(t_i^m)$, $i \in J_m$. Then for the sum S^m we get the following representation

$$S^m = \sum_{i \in J_m} g\left(s_i^m\right)\left(\widetilde{Y}_i^m - W_i^m\right).$$

Let g_{m-1} be the truncated Haar expansion of g for $m \geq 2$ (see (14):

$$g_{m-1} = c_0(g)h_0 + \sum_{k=0}^{m-2} 2^{k/2} \sum_{j=1}^{2^k} c_{k,j}(g)h_{k,j}, \tag{69}$$

where $c_0(g)$ and $c_{k,j}(g)$ are the corresponding Fourier-Haar coefficients defined by (13) with g replacing f. Then obviously

$$S^m = S_1^m + S_2^m,$$

where

$$\begin{aligned} S_1^m &= \sum_{i \in J_m} \left(g\left(s_i^m\right) - g_{m-1}\left(s_i^m\right)\right)\left(\widetilde{Y}_i^m - W_i^m\right), \qquad (70) \\ S_2^m &= \sum_{i \in J_m} g_{m-1}\left(s_i^m\right)\left(\widetilde{Y}_i^m - W_i^m\right). \end{aligned}$$

By the Hölder inequality

$$E\exp\left\{tu_n S^m\right\} \leq \left(E\exp\left\{2tu_n S_1^m\right\} E\exp\left\{2tu_n S_2^m\right\}\right)^{1/2}. \tag{71}$$

Now the inequality (63) for $m \geq 2$ follows from Propositions 23 and 24 below. This will complete the proof of Theorem 3.

First we prove the following

Proposition 23 *There exist two positive constants c_0 and c_1, depending only on $C_{\max}$, $C_{\min}$, λ_0, L, such that for any t satisfying $|t| \leq c_0$*

$$E\exp\{tu_n S_1^m\} \leq \exp\left\{t^2 c_1\right\}.$$

Proof. Since by (24) the function $a(s)$ is Lipschitz and $f \in \mathcal{H}(\frac{1}{2}, L)$, it is easy to see that the function $g(s) = f(a(s))$ is also in a Hölder ball $\mathcal{H}(\frac{1}{2}, L_0)$, but with another constant L_0 depending on $C_{\max}$, $C_{\min}$ and L. By Hölder's inequality

$$E\exp\{tu_n S_1^m\} \leq \left(E\exp\left\{\sum_{i \in J_m} \rho_i \widetilde{Y}_i^m\right\} E\exp\left\{-\sum_{i \in J_m} \rho_i W_i^m\right\}\right)^{1/2}, \tag{72}$$

where $\rho_i = 2tu_n\left(g\left(s_i^m\right) - g_{m-1}\left(s_i^m\right)\right)$ and $|t| \leq c_0 = \lambda_0/(6L_0)$. Note that by Proposition 6 $\|g - g_{m-1}\|_\infty \leq L_0 2^{-(m-1)/2}$. Therefore for $|t| \leq c_0$

$$|\rho_i| \leq 2|t|\, u_n L_0 2^{-(m-1)/2} \leq 2|t|\, L_0 2^{-(m-1)/2} \leq \lambda_0/3.$$

Then according to Lemma 25 we have for $i \in J_m$

$$E\exp\left\{\rho_i \widetilde{Y}_i^m\right\} \leq \exp\left\{\rho_i^2 E(\widetilde{Y}_i^m)^2\right\} \leq \exp\left\{8t^2 L_0^2 2^{-m} E(X_i^m)^2\right\}. \tag{73}$$

An analogous bound holds true for the normal r.v.'s W_i^m, $i \in J_m$:

$$E \exp\left\{\rho_i W_i^m\right\} \le \exp\left\{4t^2 L_0^2 2^{-m} E(X_i^m)^2\right\}. \tag{74}$$

Taking into account that $\widetilde{Y}_i^m$, $i \in J_m$ and W_i^m, $i \in J_m$ are sequences of independent r.v.'s and inserting (73) and (74) into (72) we obtain

$$E \exp\{t u_n S_1^m\} \le \exp\left\{6t^2 L_0^2 2^{-m} \sum_{i \in J_m} E(X_i^m)^2\right\}. \tag{75}$$

Now we remark that $\# J_m \le 2^{m+1}$. Then by (7)

$$\sum_{i \in J_m} E(X_i^m)^2 \le \# J_m C_{\max} \le 2^{m+1} C_{\max}. \tag{76}$$

Inserting (76) into (75) we obtain the assertion. ■

Now we will produce a bound for the second expectation on the right-hand side of (71).

Proposition 24 *There exist two constants c_0 and c_1, depending only on $C_{\min}$, $C_{\max}$, λ_0, L, such that for any t satisfying $|t| \le c_0$*

$$E \exp\{t u_n S_2^m\} \le \exp\left\{t^2 c_1\right\}.$$

Proof. From (70) and (69) we obtain

$$S_2^m = c_0(g) S_{0,1}^m + \sum_{k=0}^{m-2} 2^{k/2} \sum_{j=1}^{2^k} c_{k,j}(g)\{S_{k+1,2j-1}^m - S_{k+1,2j}^m\},$$

where $S_{k,j}^m$ are defined by (40). Since the function $g(s)$ is in the Hölder ball with a Hölder constant L_0 depending on $C_{\max}$, $C_{\min}$, L, according to Proposition 5 we have the following bounds for the Fourier-Haar coefficients:

$$c_0(g) \le L_0/2, \quad |c_{k,j}(g)| \le 2^{-3/2} L_0 2^{-k}, \tag{77}$$

for $k = 0, ..., m-2$. Note also that by Lemma 22 there are two constants c_2 and c_3, depending only on $C_{\max}$, $C_{\min}$, λ_0, such that

$$E \exp\left\{c_2 \left|S_{k,j}^m\right|\right\} \le c_3, \tag{78}$$

for $j = 1, ..., 2^k$ and $k = 0, ..., m-1$.

Put $c_0 = c_2/(8L_0)$. By Hölder's inequality we have for any t satisfying $|t| \le c_0$

$$E \exp\left\{t u_n S_2^m\right\} \le \left(E \exp\left\{t m u_n c_0(g) S_{0,1}^m\right\} \prod_{k=0}^{m-2} E \exp\left\{t m u_n \xi_k\right\}\right)^{1/m},$$

where for the sake of brevity we denote

$$\xi_k = 2^{k/2} \sum_{j=1}^{2^k} c_{k,j}(g)\{S_{k+1,2j-1}^m - S_{k+1,2j}^m\}, \tag{79}$$

for $k = 0, ..., m-2$. The statement of the proposition will be proved if we show that for t satisfying $|t| \le c_0$

$$E \exp\left\{tmu_n c_0(g) S_{0,1}^m\right\} \le \exp\left\{t^2 L_0^2 c_4\right\} \tag{80}$$

and

$$E \exp\left\{tmu_n \xi_k\right\} \le \exp\left\{t^2 L_0^2 c_5\right\}, \tag{81}$$

for some constants c_4 and c_5 depending only on $C_{\max}$, $C_{\min}$, λ_0. We shall assume m and k fixed from now on.

It is easy to show (80). For this we note that by (77) and (62) for $|t| \le c_0$

$$|tmu_n c_0(g)| \le |t|\, m(M+1)L_0/(2\log^2 n) \le |t|\, L_0 \le c_2. \tag{82}$$

Then (82), (78) with $k = 0$, $j = 1$ and Lemma 26 imply that the inequality (80) holds true with $c_4 = 4c_3/c_2^2$.

The proof of (81) is a bit more intricate. The main trouble is that the r.v.'s

$$\zeta_j = S_{k+1,2j-1}^m - S_{k+1,2j}^m, \quad j = 1, ..., 2^k \tag{83}$$

are dependent and so we cannot make use of the product structure of the exponent $\exp\{t\xi_k\}$ directly. However Proposition 10 ensures that the components of the sum ξ_k (see (79)) are *almost* independent, this allowing to exploit the product structure in an implicit way. The main idea is to introduce the r.v.'s

$$\widetilde{V}_j = \beta_{2j-1} W_{k+1,2j-1}^m - \beta_{2j} W_{k+1,2j}^m,$$

where

$$\beta_{2j-1} = \left(\frac{B_{2j}}{B_{2j-1}}\right)^{1/2}, \quad \beta_{2j} = \left(\frac{B_{2j-1}}{B_{2j}}\right)^{1/2},$$

with $B_{2j-1} = B_{k+1,2j-1}^m$ and $B_{2j} = B_{k+1,2j}^m$. The r.v. $\widetilde{V}_j$ can be easily seen to be independent of the r.v. $W_{k+1,2j-1}^m - W_{k+1,2j}^m = W_{k,j}^m$, for $j = 1, ..., 2^k$. Set also for brevity

$$Z_j = \widetilde{Y}_{k+1,2j-1}^m - \widetilde{Y}_{k+1,2j}^m, \quad V_j = W_{k+1,2j-1}^m - W_{k+1,2j}^m,$$

for $j = 1, ..., 2^k$. By Lemma 12 the r.v.'s

$$\widetilde{\zeta}_j = Z_j - \widetilde{V}_j, \quad j = 1, ..., 2^k$$

are independent. It is obvious that

$$\zeta_j = \widetilde{\zeta}_j + \widetilde{V}_j - V_j. \tag{84}$$

From (79), (83) and (84) we obtain

$$\xi_k = \xi_k^1 + \xi_k^2,$$

where

$$\begin{aligned} \xi_k^1 &= 2^{k/2} \sum_{j=1}^{2^k} c_{k,j}(g) \widetilde{\zeta}_j, \\ \xi_k^2 &= 2^{k/2} \sum_{j=1}^{2^k} c_{k,j}(g) \left(\widetilde{V}_j - V_j\right). \end{aligned} \tag{85}$$

By Hölder's inequality

$$E\exp\{tmu_n\xi_k\} \le \left(E\exp\left\{2tmu_n\xi_k^1\right\} E\exp\left\{2tmu_n\xi_k^2\right\}\right)^{1/2}. \tag{86}$$

Now we proceed to estimate the first expectation in the right-hand side of (86). We make use of the independence of r.v.'s $\widetilde{\zeta}_j$, $j = 1, ..., 2^k$ to get

$$E\exp\left\{2tmu_n\xi_k^1\right\} \le \prod_{j=1}^{2^k} E\exp\left\{tr_j\widetilde{\zeta}_j\right\}, \tag{87}$$

where $r_j = 2mu_n 2^{k/2} c_{k,j}(g)$. We will show that for $j = 1, ..., 2^k$

$$E\exp\left\{tr_j\widetilde{\zeta}_j\right\} \le \exp\left\{t^2 2^{-k} L_0^2 c_6\right\}, \tag{88}$$

with some constant c_6. First by (84) and by Hölder's inequality we have

$$E\exp\left\{tr_j\widetilde{\zeta}_j\right\} \le \left(E\exp\{2tr_j\zeta_j\} E\exp\left\{2tr_j\left(V_j - \widetilde{V}_j\right)\right\}\right)^{1/2}. \tag{89}$$

By (83) and by Hölder's again inequality we get

$$E\exp\{2tr_j\zeta_j\} \le \left(E\exp\left\{4tr_j S_{k+1,2j-1}^m\right\} E\exp\left\{4tr_j S_{k+1,2j}^m\right\}\right)^{1/2}. \tag{90}$$

Note that by (77) and (62)

$$|4tr_j| \le \left|8tmu_n 2^{k/2} c_{k,j}(g)\right| \le 4|t| L_0 2^{-k/2} \le c_2. \tag{91}$$

From (78), (91) and Lemma 26 we obtain for $l = 2j-1, 2j$ and $k \le m-2$

$$E\exp\left\{4tr_j S_{k+1,l}^m\right\} \le \exp\left\{t^2 2^{-k} L_0^2 c_7\right\}, \tag{92}$$

with the constant $c_7 = 4c_3/c_2^2$. Inserting (92) into (90) we obtain

$$E\exp\{2tr_j\zeta_j\} \le \exp\left\{t^2 2^{-k} L_0^2 c_7\right\}. \tag{93}$$

Thus we have estimated the first expectation in the right-hand side of (89). To estimate the second one, we note that in view of the independence of the normal r.v.'s $W_{k+1,2j-1}^k$ and $W_{k+1,2j}^k$

$$E\exp\left\{2tr_j\left(V_j - \widetilde{V}_j\right)\right\} = \exp\left\{4t^2 r_j^2\left(\sqrt{B_{2j-1}} - \sqrt{B_{2j}}\right)^2\right\}. \tag{94}$$

Because of the elementary inequality $(\sqrt{a} - \sqrt{b})^2 \le |a-b|$, with $a, b \ge 0$, and of Proposition 10

$$\left(\sqrt{B_{2j-1}} - \sqrt{B_{2j}}\right)^2 \le |B_{2j-1} - B_{2j}| \le 2C_{\max}.$$

Inserting this bound into (94) and using (91) one obtains

$$E\exp\left\{2tr_j\left(V_j - \widetilde{V}_j\right)\right\} \le \exp\left\{t^2 2^{-k} L_0^2 c_8\right\}, \tag{95}$$

with some constant c_8 depending only on $C_{\max}$. From (93), (95) and (89) we obtain the inequality (88). Inserting in turn (88) into (87) we arrive at the bound

$$E\exp\left\{2tmu_n\xi_k^1\right\}\leq\exp\left\{t^2L_0c_6\right\}. \tag{96}$$

Thus we have estimated the first expectation in the right-hand side of (86). It remains to estimate the second one. Since the r.v.'s $\widetilde{V}_j - V_j$, $j = 1, ..., 2^k$ are independent, (85) and (95) we obtain

$$E\exp\left\{2tmu_n\xi_k^2\right\}\leq\exp\left\{t^2L_0^2c_8\right\}. \tag{97}$$

The inequalities (96) and (97) imply (81), this completing the proof of the proposition. ∎

7 Appendix

In the proofs we made use of the following simple auxiliary results.

Lemma 25 *Let ξ be a real valued r.v. such that $E\xi = 0$, $0 < E\xi^2 < \infty$ and for which Sakhanenko's condition*

$$\lambda_0 E\left|\xi\right|^3\exp\{\lambda_0|\xi|\}\leq E\xi^2$$

holds true for some $\lambda_0 > 0$. Then for all $|t| \leq \lambda_0/3$

$$E\exp\{t\xi\}\leq\exp\left\{t^2E\xi^2\right\}.$$

Proof. Let $\mu(t) = E\exp(t\xi)$ and $\psi(t) = \log\mu(t)$ be the moment and cumulant generating functions respectively. The conditions of the lemma imply that $\mu(t) \leq c_1$ for any real $|t| \leq \lambda_0/3$, and using a three term Taylor expansion we get with $0 \leq \nu \leq 1$

$$\psi(t)=\psi(0)+\psi'(0)t+\psi''(0)\frac{t^2}{2}+\psi'''(\nu t)\frac{t^3}{6}.$$

Note that $\psi(0) = 0$, $\psi'(0) = 0$, $\psi''(0) = E\xi^2$ and $\mu(t) \geq 1$ by Jensen's inequality, while for the third derivative we have for any real s satisfying $|s| \leq \lambda_0/3$

$$\psi'''(s)=\mu'''(s)\mu(s)^{-1}-3\mu''(s)\mu'(s)\mu(s)^{-2}+2\mu'(s)^3\mu(s)^{-3}.$$

Using Hölder's inequality and $\mu(s) \geq 1$ we arrive at the bound

$$\left|\psi'''(s)\right|\leq 6E|\xi|^3\exp(\lambda_0|\xi|).$$

Since $|t| \leq \lambda_0/3$, by Sakhanenko's condition

$$0\leq\psi(t)\leq\frac{t^2}{2}E\xi^2+t^3E|\xi|^3\exp(\lambda_0|\xi|)\leq t^2E\xi^2. \quad ∎$$

Lemma 26 *Let ξ be a real valued r.v. such that $E\xi = 0$ and*

$$E\exp\{\lambda_0|\xi|\}\leq c_1,$$

for some $\lambda_0 \geq 0$ and $c_1 \geq 1$. Then for all $|t| \leq \lambda_0/2$

$$E\exp\{t\xi\}\leq\exp\{c_2t^2\},$$

where $c_2 = 4c_1/\lambda_0^2$.

Proof. The argument is similar to Lemma 25. We use the same notations. A two term Taylor expansion yields with $0 \le \nu \le 1$

$$\psi(t) = \psi(0) + \psi'(0)t + \psi''(\nu t)\frac{t^2}{2}.$$

Since $x^2 \le 2\exp(|x|)$ for any real x, for any s satisfying $|s| \le \lambda_0/2$ we have

$$\begin{aligned} 0 &\le \psi''(s) = \mu(s)^{-2}\{E\xi^2\exp(s\xi) - (E\xi\exp(s\xi))^2\} \\ &\le E\xi^2\exp(s\xi) \le E\xi^2\exp(\frac{\lambda_0}{2}|\xi|) \le 8\frac{c_1}{\lambda_0^2}. \end{aligned}$$

Consequently

$$0 \le \psi(t) = \psi''(\nu t)\frac{t^2}{2} \le 4\frac{c_1}{\lambda_0^2}t^2. \ \blacksquare$$

Lemma 27 *Let ξ_i, $i = 1, ..., n$ be a sequence of independent r.v.'s such that for all $i = 1, ..., n$ $E\xi_i = 0$, $0 < E\xi_i^2 < \infty$ and*

$$E|\xi_i|^3\exp\{\lambda_0|\xi_i|\} \le E\xi_i^2,$$

for some positive constant λ_0. Put $S_n = \xi_1 + ... + \xi_n$ and $S_n^* = S_n\mathbf{1}\left(|S_n| \le B_n^2\right)$. Then

$$E\exp\left\{c_1(S_n^*/B_n)^2\right\} \le 1 + 2/c_1,$$

where $c_1 = \frac{1}{4}\min\{\lambda_0/3, 1/2\}$.

Proof. Denote

$$F(x) = P\left((S_n^*/B_n)^2 > x\right).$$

We will prove first that

$$F(x) \le 2\exp\{-c_2 x\}, \quad x \ge 0, \tag{98}$$

where $c_2 = 2c_1$. For this we note that

$$F(x) = P\left(S_n^*/B_n > \sqrt{x}\right) + P\left(S_n^*/B_n < -\sqrt{x}\right).$$

It suffices to estimate only the first probability in the right-hand side of the above equality, the second one being handled in the same way. If $x > B_n^2$, then

$$P\left(S_n^*/B_n > \sqrt{x}\right) = 0,$$

thus there is nothing to prove in this case. Let $x \le B_n$. Then denoting $t = 2c_2\sqrt{x}$ one obtains

$$\begin{aligned} P\left(S_n^* > \sqrt{x}\right) &\le P\left(S_n > \sqrt{x}\right) \le \exp\left\{-t\sqrt{x}\right\}E\exp\{tS_n/B_n\} \\ &= \exp\left\{-t\sqrt{x}\right\}\prod_{i=1}^n E\exp\{t\xi_i/B_n\}. \end{aligned} \tag{99}$$

Note that $t/B_n = 2c_2\sqrt{x}/B_n \le 2c_2 \le \lambda_0/3$. Then by Lemma 25

$$E\exp\{t\xi_i/B_n\} \le \exp\left\{t^2E\xi_i^2/B_n^2\right\}.$$

Inserting this into (99) we get

$$\begin{aligned} P\left(S_n^*/B_n > \sqrt{x}\right) &\leq \exp\left\{-t\sqrt{x}\right\} \prod_{i=1}^{n} \exp\left\{t^2 E\xi_i^2/B_n^2\right\} \\ &= \exp\left\{-t\sqrt{x} + t^2\right\} \leq \exp\left\{-c_2 x\right\}. \end{aligned}$$

this proving (98). Integrating by parts we have

$$\begin{aligned} E\exp\left\{c_1 (S_n^*)^2/B_n\right\} &= \int_0^\infty \exp\{c_1 x\} dF(x) \\ &= 1 + \int_0^\infty F(x)\exp\{c_1 x\} dx \\ &\leq 1 + 2\int_0^\infty \exp\{c_1 x - c_2 x\} dx \\ &\leq 1 + 2/c_1. \blacksquare \end{aligned}$$

References

[1] Bretagnolle, J., Massart, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Probab.* **17**, 239-256.

[2] Brown, L. D. and Low, M. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** (6)

[3] Csörgő , M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

[4] Dudley, R. (1989). *Real Analysis and Probability.* Wadsworth and Brooks/Cole, Pacific Grove, Ca.

[5] Einmahl, U. (1989). Extensions on results of Komlós, Major and Tusnády to the multivariate Case. *J. Multivariate Anal.* 28, 20-68.

[6] Einmahl, U. and Mason, D. M. (1995). Gaussian approximation of local empirical processes indexed by functions. *Probab. Theory Relat. Fields,* to appear.

[7] Grama, I. and Nussbaum, M. (1996). Asymptotic equivalence for nonparametric generalized linear models. Preprint No. 289, Weierstrass Institute, Berlin.

[8] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag, New York etc.

[9] Kashin B. S., and Saakjan, A. A. (1984). *Orthogonal Series.* Nauka, Moscow (in Russian).

[10] Koltchinskii, V. I. (1994). Komlós-Major-Tusnády approximation for the general empirical process and Haar expansions of classes of functions. *J. Theoretical Probab.* **7**, 73-118.

[11] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv's and the sample df. I *Z. Wahrsch. verw. Gebiete* **32**, 111-131.

[12] Komlós, J., Major, P. and Tusnády, G. (1976). An approximation of partial sums of independent rv's and the sample df. II *Z. Wahrsch. verw. Gebiete* **34**, 33-58.

[13] Massart, P. (1989). Strong approximation for multivariate empirical and related processes, via KMT constructions. *Ann. Probab.* **17**, 266-291.

[14] Nussbaum, M. (1993). Asymptotic equivalence of density estimation and white noise. Preprint No.35, Institute of Applied Analysis and Stochastics, Berlin.

[15] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and white noise. *Ann. Statist.* **24** 2399-2430

[16] Rio, E. (1993). Strong approximation for set-indexed partial sum Processes, via K.M.T. constructions I. *Ann. Probab.* **21**, 759-790.

[17] Rio, E. (1993). Strong approximation for set-indexed partial sum Processes, via K.M.T. constructions II. *Ann. Probab.* **21**, 1706-1727.

[18] Rio, E. (1994). Local invariance principles and their applications to density estimation. *Probab. Theory Relat. Fields* **98**, 21-45.

[19] Sakhanenko, A. (1984). The rate of convergence in the invariance principle for non-identically distributed variables with exponential moments. *Limit theorems for sums of random variables: Trudy Inst. Matem., Sibirsk. Otdel. AN SSSR.* Vol. 3, 3-49 (in Russian).

[20] Van der Vaart, A. and Wellner, J. (1996). *Weak Convergence and Empirical Processes.* Springer, New York.

[21] Zaitsev, A. (1996). Multidimensional version of the results of Komlós, Major and Tusnády for vectors with finite exponential moments. *Unpublished manuscript.*

[22] Zaitsev, A. (1987). On the Gaussian approximation of convolutions under multidimensional analogues of S. N. Bernstein's inequality conditions. *Probab. Theory Relat. Fields* **74**, 534-566

Institute of Mathematics
Academy of Sciences
Academiei str. 5
Kishinev 277028, Moldova
e-mail 16grama@mathem.moldova.su

Weierstrass Institute
Mohrenstr. 39
D-10117 Berlin, Germany
e-mail nussbaum@wias-berlin.de