Z. Wahrscheinlichkeitstheorie verw. Gebiete
34, 33–58 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

An Approximation of Partial Sums of Independent RV's, and the Sample DF. II

J. Komlós, P. Major and G. Tusnády
Mathematical Institute of the Hungarian Academy of Sciences
1053 Budapest, Reáltonada u. 13-15, Hungary

Given a d.f. $F(x)$ with the property

$$\int x F(dx) = 0, \quad \int x^2 F(dx) = 1, \quad \int e^{tx} F(dx) < \infty$$

for $|t| < t_0$, we construct a sequence $X_1, X_2, \dots$ of i.i.d.r.v.'s with d.f. $F(x)$ and $Y_1, Y_2, \dots$ with standard normal distribution in such a way that the sequences $S_n = X_1 + \dots + X_n$, $T_n = Y_1 + \dots + Y_n$, $n = 1, 2, \dots$ satisfy the relation $|S_n - T_n| = O(\log n)$ a.s. Under some mild conditions the best possible normal approximation of the process S_n will also be given in the case when F has not a moment generating function.

1. Introduction

Given a d.f. $F(x)$, $\int x F(dx) = 0$, $\int x^2 F(dx) = 1$, let us consider a sequence $X_1, X_2, \dots$ of i.i.d.r.v.'s with distribution F and a sequence $Y_1, Y_2, \dots$ of i.i.d.r.v.'s with standard normal distribution. Set $S_n = X_1 + \dots + X_n$, $T_n = Y_1 + \dots + Y_n$, $n = 1, 2, \dots$ Our aim is to construct the sequences X_n and Y_n in such a way that for all n, S_n and T_n are as near to each other as possible. This problem was solved under some special conditions in Theorem 1 in part I. Our first aim is to show that the regularity conditions (i) or (ii) in Theorem 1 in part I were superfluous and the following generalization holds true:

Theorem 1. *If $\int e^{tx} F(dx) < \infty$ for $|t| \leq t_0$, the sequences $X_1, X_2, \dots$ and Y_1, $Y_2, \dots$ can be constructed in such a way that for all $x > 0$ and every n*

$$P\left(\max_{k \leq n} |S_k - T_k| > C \log n + x\right) < K e^{-\lambda x}, \tag{1.1}$$

where C, K, λ depend only on F, and λ can be taken as large as desired by choosing C large enough. Consequently, $|S_n - T_n| = O(\log n)$ a.s.

The proof of Theorem 1 will be based on the following two special cases:

Theorem 1 A. *If F satisfies the conditions of Theorem 1, and, in addition, it has an absolutely continuous component, i.e. $F(x) = pF_1(x) + (1-p) F_2(x)$ where*

$0<p\leqq 1$, F_1 and F_2 are d.f.'s, F_1 is absolutely continuous, then the sequences X_n and Y_n can be constructed in such a way that (1.1) *holds.*

Theorem 1B. *If F is concentrated on a finite interval, then an appropriate construction of the pairs (X_n, Y_n) satisfies* (1.1).

Let us remark that if F has not a finite moment generating function in any neighbourhood of the origin, then

$$\limsup_{n\to\infty} \frac{|S_n - T_n|}{\log n} = \infty \text{ a.s.}$$

for any construction of the sequences X_n and Y_n. This can be read out from the first part of the proof of Theorem 2 in part I. This means that the conditions of Theorem 1 are necessary.

One may also be interested in the case when F has not a moment generating function. If F has some finite moments, the following result gives information about the closeness of S_n to a normal sequence.

Theorem 2. *Let $\int |x|^r F(dx) < \infty$ for some $r>3$. Then for an appropriate construction*

$$|S_n - T_n| = o(n^{1/r}) \qquad a.s.$$

For $r \geqq 4$ this theorem is an improvement of Strassen's who proved

$$O(n^{\frac{1}{4}} (\log n)^{\frac{1}{2}} (\log\log n)^{\frac{1}{4}})$$

instead of $o(n^{1/r})$.

Now we generalize Theorem 2 in the following way: Let $H(x)>0$, $x>0$ be a monotone increasing, continuous function having the following properties

$$\text{(i)} \quad \frac{H(x)}{x^{3+\delta}} \text{ is monotone increasing for some } \delta>0 \text{ and all } x>x_0, \tag{1.2}$$

$$\text{(ii)} \quad \frac{\log H(x)}{x} \quad \text{is monotone decreasing for } x>x_0.$$

Theorem 3. *Let $H(x)$ satisfy (i) and (ii). Define K_n by the equation $H(K_n)=n$. If $\int H(|x|)\, F(dx) < \infty$, then there exists a construction of $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$, and a constant $C>0$ such that*

$$P\left(\limsup \frac{|S_n - T_n|}{K_n} \leqq C\right) = 1. \tag{1.3}$$

Remarks. One can choose a sequence $\varepsilon_n \to 0$ such that K_n can be substituted by K_n', the solution of the equation $H(K_n') = \varepsilon_n \cdot n$. Indeed, consider a function $f(x)$ tending to ∞ such that $\int H(|x|)\, f(|x|)\, F(dx) < \infty$.

Apply Theorem 3 substituting $H(x)$ by $H_1(x) = H(x) f(x)$. Thus Theorem 2 is a special case of Theorem 3: one has only to choose $H(x)=x^r$, $r>3$ and apply Theorem 3 and the Remark, to obtain Theorem 2. Taking $H(x)=e^{tx}$ in Theorem 3, we get the statement $|S_n - T_n| = O(\log n)$ in Theorem 1.

Theorem 3 can be reformulated in the following more attractive way: If K_n is a sequence of positive numbers, $\frac{K_n}{\log n}$ is monotone increasing, and $\frac{K_n}{n^{\frac{1}{3}-\delta}}$ is monotone decreasing for $n>n_0$, then the condition

$$\sum_{n=1}^{\infty} P(|X_n|>K_n)<\infty$$

implies that $|S_n-T_n|=O(K_n)$ a.s. for an appropriate construction. I.e. the difference $|S_n-T_n|$ doesn't grow any faster than the individual terms $|X_n|$ themselves.

Theorem 3 is sharp in the following sense:

If $\sum_{n=1}^{\infty} P(|X_n|>K'_n)=\infty$, then $|S_n-S_{n-1}|>K'_n$ occurs infinitely often (a.s.); now $K_n\geqq c\cdot\log n$ implies that $|T_n-T_{n-1}|>\frac{1}{2}K'_n$ only finitely many times, whence we get

$$P\left(\limsup_{n\to\infty}\frac{|S_n-T_n|}{K'_n}\geqq\frac{1}{4}\right)=1.$$

In case of Theorem 2, K'_n can be chosen $K'_n=\varepsilon(n)\cdot n^{1/r}$, where $\varepsilon(n)\to 0$ as slowly as one likes if F is chosen appropriately. Thus, the estimation $o(n^{1/r})$ in Theorem 2 is the best possible.

Let $x>K_n$. One may ask, what is the probability of the event

$$\sup_{k\leqq n}|S_k-T_k|>x$$

for an appropriate construction. The following theorem gives an answer to this question.

Theorem 4. *Let $H(x)$ and $F(x)$ satisfy the conditions of Theorem 3. Then for any x, $K_n<x<C_1\sqrt{n\log n}$ (more generally $x>K_n$, $x^2/\log H(x)<c_1 n$) there exist two finite sequences $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ such that*

$$P(\sup_{k\leqq n}|S_k-T_k|>x)\leqq C_2\frac{n}{H(ax)},$$

where C_1, C_2 and a are positive constants depending only on F.

Just as in the remark following Theorem 3, one sees that the numerator n in Theorem 4 can be replaced by $o(n)$.

We remark that Theorem 1B is the only place in the two papers, where the fact that $X_1, X_2, \ldots$ are identically distributed is essentially used. Thus, for variables $X_1, X_2, \ldots$ whose distributions have continuous components (uniformly in a certain sense), our dyadic construction may also be appropriate.

We also mention that in the paper [1] we discussed some applications of this paper, e.g. for the rate of convergence of S_n-T_n in the Prohorov distance.

2. The Proof of Theorem 1

Theorem 1 differs from Theorem 1 in part I only in that we do not require the existence of a smooth density function. Let us remark that this additional assumption was needed only in the proof of Lemma 1, or more precisely in the following formula about conditional distributions

$$P(S_n-2S_{n/2}>nx\,|\,S_n=ny)$$
$$=[1-\phi(\sqrt{n}\,x)]\exp\{O(nx^3+nx^2|y|+x+|y|+1/\sqrt{n})\}$$

if $0\leqq x<\varepsilon$, $|y|<\varepsilon$.

We proved this relation by integrating the conditional density function. One would hope that upon finding another proof, one might get rid of the condition requiring the existence of a density function. This would mean that the proof of Theorem 1 in part I applies to our new theorem without any modification. This hope however proves illusory. The next example shows that the above-mentioned expansion does not hold if we assume only the existence of the moment generating function.

Example. Let the r.v. X have a distribution concentrated on a sequence x_0, $x_1, \ldots$ with the following properties

a) $n<|x_n|<n+1, \quad n=0,1,\ldots;$

b) $x_0, x_1, \ldots$ are linearly independent i.e. for any finite linear combination $\sum k_i x_i$ with rational coefficients, $\sum k_i x_i=0$ iff $k_i=0$ for all i.

Let $p_n=P(X=x_n)$ satisfy $c\cdot\exp(-n)<p_n<\exp(-n)$, and assume that $EX=0$ and $EX^2=1$. Consider a sequence $X_1, X_2, \ldots$ of i.i.d.r.v.'s having the same distribution as X.

A little consideration shows that on a set A_n of positive probability the following relations hold with appropriate constants C and p (C and p do not depend on n):

(i) $|S_{n-1}|<C\sqrt{n}$

(ii) $\sum_{i=1}^{n/2} X_{k(i)}<C\sqrt{n}$ for at least $p\binom{n-1}{n/2}$

subsets $\{k(1),\ldots,k(n/2)\}$ of the set $\{1,2,\ldots,n-1\}$.

Let α be a sufficiently large constant, $m=[\alpha C\sqrt{n}]$, and let $a=S_{n-1}(\omega)$ for an arbitrary $\omega\in A_n$. Then

$$P(S_n-2S_{n/2}>C(\alpha-4)\sqrt{n}\,|\,S_n=a+x_m)\geqq p/2$$

instead of $\phi(\alpha-4)\cdot(1+O(1/\sqrt{n}))$.

That is why we slightly modify the original construction, and the original proof applies for the new construction with some alterations. A natural idea is to smooth the random variables by adding small normal variables. By this smoothing we want to ensure that the density functions of the new variables satisfy the central limit theorem. On the other hand the smoothing r.v.'s must be negligible. We show that both requirements can be satisfied if F has an absolute continuous component. This is the idea of the proof of Theorem 1A.

In the proof of Theorem 1B we use a somewhat different method. Knowing the values $X_1, X_2, \ldots, X_n$ we rearrange them in such a way that their joint distribution be the prescribed one, and $S_{n/2}$ be near $T_{n/2}$. This rearrangement will be done by using Lemma 3 which is of independent interest.

Before proving Theorems 1A and 1B we show how these two particular cases (the "smooth case" and the "bounded case") imply Theorem 1. For this we need the following

Lemma 1. *Given the distribution functions F_1, F_2 and G_1, G_2, let $S_1^{(i)}, S_2^{(i)}, \ldots$ resp. $T_1^{(i)}, T_2^{(i)}, \ldots$ be the partial sums of i.i.d.r.v.'s with d.f. F_i resp. G_i, $i=1,2$. For any $0 \leqq p \leqq 1$ there are two sequences $S_1, S_2, \ldots$ and $T_1, T_2, \ldots$ which are the partial sums of i.i.d.r.v.'s with d.f. $pF_1+(1-p)F_2$ resp. $pG_1+(1-p)G_2$, and satisfy the inequality*

$$P\big(\sup_{k \leqq n} |S_k - T_k| > a+b\big) \leqq P\big(\sup_{k \leqq n} |S_k^{(1)} - T_k^{(1)}| > a\big) + P\big(\sup_{k \leqq n} |S_k^{(2)} - T_k^{(2)}| > b\big)$$

for all $a>0$, $b>0$ and any n.

Proof. We may suppose that the sequences $S_1^{(1)}, S_2^{(1)}, \ldots$ $T_1^{(1)}, T_2^{(1)}, \ldots$are independent of the sequences $S_1^{(2)}, S_2^{(2)}, \ldots$ $T_1^{(2)}, T_2^{(2)}$. Let $\varepsilon_1, \varepsilon_2, \ldots$ be a sequence of i.i.d.r.v.'s independent of all $S^{(i)}$ and $T^{(i)}$ with the following distribution

$$P(\varepsilon_k=0)=1-P(\varepsilon_k=1)=p.$$

Define

$$\nu(n)=\sum_{k=1}^{n} \varepsilon_k, \qquad S_n = S_{\nu(n)}^{(1)} + S_{n-\nu(n)}^{(2)},$$

and $T_n = T_{\nu(n)}^{(1)} + T_{n-\nu(n)}^{(2)}$.

Then the S_n-s and T_n-s have the prescribed distribution and satisfy the lemma, because

$$|S_k - T_k| \leqq \sup_{j \leqq n} |S_j^{(1)} - T_j^{(1)}| + \sup_{j \leqq n} S_j^{(2)} - T_j^{(2)}|,$$

for all $k \leqq n$.

Proof of Theorem 1. Write F in the form $F=pF_1+(1-p)F_2$, where F_1, F_2 are d.f.-s, $p>0$, and F_1 is concentrated on a finite interval. Applying Theorem 1B one can construct two sequences $S_1^{(1)}, S_2^{(1)}, \ldots$ and $T_1^{(1)}, T_2^{(1)}, \ldots$ with the following properties

a) $S_1^{(1)}, S_2^{(1)}, \ldots$ are the partial sums of i.i.d.r.v.'s with d.f. F_1, $T_1^{(1)}, T_2^{(1)}, \ldots$ are the partial sums of i.i.d.r.v.'s with normal distribution with expectation $m=\int x F_1(dx)$ and variance $\sigma^2=\int x^2 F_1(dx)-m^2$.

b) $P\big(\sup_{k \leqq n} |S_k^{(1)} - T_k^{(1)}| > C\log n + x\big) < K e^{-\lambda x}$.

Applying Lemma 1 one gets two sequences $S_1, S_2, \ldots$ and $U_1, U_2, \ldots$ which are partial sums of i.i.d.r.v.'s with d.f. F, resp. $p\phi(m,\sigma)+(1-p)F_2$ and satisfy relation b). This new distribution has an absolute continuous component already, therefore by Theorem 1A one can construct two sequences $U_1', U_2', \ldots$ and T_1, $T_2, \ldots$ which are partial sums of i.i.d.v.'s with d.f. $p\phi(m,\sigma)+(1-p)F_2$ resp. with standard normal distribution, and satisfy relation b). It can be shown by standard measure theoretical arguments that $U_1, U_2, \ldots$ and $U_1', U_2', \ldots$ can be chosen the same sequence. This last remark completes the proof.

3. The Proof of Theorem 1 A

We need the following lemma:

Lemma 2. *Let $X_1, X_2, \ldots$ be a sequence of i.i.d.r.v.'s, $EX_1=0$, $EX_1^2=1$,*

$$R(t)=E\exp(tX_1)<\infty$$

for $|t|\leqq t_0$. Let F, the distribution function of X_i, have an absolute continuous component. Let $\eta_1, \eta_2, \ldots$ be a sequence of i.i.d.r.v.'s with standard normal distribution, and let σ_n be a sequence satisfying $1/n^2>\sigma_n^2>q^n$ with an appropriate $0<q<1$ (q may depend on F). Let the η's be independent of the X's. Introducing the notations

$$S_n=\sum_{i=1}^n X_i+\sigma_n\sum_{i=1}^n \eta_i, \qquad F_n(x)=P(S_n<x), \qquad f_n(x)=d/dx\,F_n(x),$$

one has the following formulas

(a) $1-F_n(nx)=[1-\phi(\sqrt{n}\,x)]\exp[nx^3\lambda(x)+O(x+n^{-\frac12})]$ *for* $0\leqq x\leqq\eta$,

(a') $F_n(-nx)=\phi(-\sqrt{n}\,x)\exp[-nx^3\lambda(-x)+O(x+n^{-\frac12})]$ *for* $0\leqq x\leqq\eta$,

(b) $f_n(nx)=n^{-\frac12}\varphi(\sqrt{n}\,x)\exp[nx^3\lambda(x)+O(|x|+n^{-\frac12})]$ *for* $|x|\leqq\eta$,

(c) $f_n(ny)\leqq Cf_n(nx)\exp[-tn(y-x)]$ *if* $|x|\leqq\eta$, $xy>0$, $\dfrac{R'(t)}{R(t)}=x$,

where C is an appropriate constant, $\phi(x)$ is the standard normal distribution function, $\varphi(x)$ the standard normal density function, $\lambda(x)$ is analytic in $|x|\leqq\eta$ and depends only on F, η is an appropriately fixed number and $O(t)$ is uniform in the interval $|t|\leqq\eta$.

Proof. The proof is similar to the classical proof of large deviation theorems. We shall prove only relations (b) and (c), (a) can be proved similarly.

Define the following conjugated distributions

$$V^{(1)}(dx)=e^{tx}\frac{F(dx)}{R(t)}, \qquad V^{(2)}(dx)=e^{tx}\phi\left(\frac{dx}{\sigma_n}\right)e^{-t^2\sigma_n^2/2}$$

$$V(x)=V^{(1)}(x)*V^{(2)}(x), \qquad V_n(x)=V(x)^{*n}.$$

Then we have

$$F_n(dx)=e^{-tx-nt^2\frac{\sigma_n^2}{2}}R^n(t)\,V_n(dx). \tag{3.1}$$

Denote

$$\psi(t)=\log R(t), \qquad v(x)=\frac{d}{dx}V(x), \qquad v_n(x)=\frac{d}{dx}V_n(x).$$

A little calculation shows that

$$M=\int x\,V(dx)=\psi'(t)+t\sigma_n^2,$$
$$D^2=\int x^2\,V(dx)-M^2=\psi''(t)+\sigma_n^2.$$

We need an estimation similar to the Berry-Essen inequality for $v_n(x)$. Therefore we apply the following Lemma A (see [4]).

Lemma A. *Let* $Y_1, Y_2, \ldots$ *be independent r.v.'s,* $EY_i=0$, *satisfying the conditions*

(i) $\sum_{i=1}^{n} EY_i^2 \geqq ng, \quad \sum_{i=1}^{n} E|Y_i|^3 \leqq nG.$

(ii) *The characteristic functions* $f_j(s)=E\exp(isY_j)$ *satisfy the relation*

$$\int_{|s|>\varepsilon} \prod_{j=1}^{n} |f_j(s)|\, ds = O(1/n), \qquad n\to\infty,$$

for a fixed ε, $0<\varepsilon<g/24G$. Then for every sufficiently large n there is a continuous density function $q(x)$ of the r.v.'s

$$\left[\sum_{j=1}^{n} EY_j^2\right]^{-\frac{1}{2}} \sum_{j=1}^{n} Y_j \quad \text{and} \quad \sup_x |q(x)-\varphi(x)| < \frac{c}{\sqrt{n}}.$$

Let us remark that Lemma A holds for triangular arrays, too. We want to apply this lemma for i.i.d.r.v.'s $Y_1, \ldots, Y_n$ with distribution function $V(x-M)$. Condition (i) holds obviously, we have to check only condition (ii).

We have

$$|f_j(s)| = |\int e^{isx}\, dV(x-M)| = |\int e^{isx} V_1(dx)|\, e^{-\frac{\sigma_n^2 s^2}{2}}$$

It is sufficient to prove that

$$|\int e^{isx} V_1(dx)| \leqq \bar{q} < 1, \qquad \text{if } |s|>\varepsilon \quad \text{and} \quad |t|<t_0. \tag{3.2}$$

Indeed, one gets then

$$\int_{|s|>\varepsilon} \prod_{i=1}^{n} |f_j(s)|\, ds \leqq \bar{q}^n \int_{|s|>\varepsilon} \exp(-n\sigma^2 s^2/2)\, ds < \bar{q}^n \frac{\sqrt{\pi}}{\sqrt{2}\sigma_n\sqrt{n}} < q'^n,$$

where $q'<1$, if q is chosen in such a way that $q>\bar{q}$. (3.2) means that

$$|\int e^{tx+isx} F(dx)| \leqq \bar{q} R(t).$$

$F(x)$ can be decomposed into the form $pF_1(x)+(1-p)F_2(x)$, where F_1, F_2 are distribution functions, $p>0$ and $F_1(x)$ is absolutely continuous. Therefore

$$|\int e^{isx+tx} F_1(dx)| < \alpha \int e^{tx} F_1(dx)$$

with some $\alpha<1$ if $|t|<t_0$ and $|s|>\varepsilon$, and this implies relation (3.2).

To see this last inequality one has only to remark that

$$\lim_{s\to\infty} \int e^{isx+tx} F_1(dx) \to 0 \qquad \text{uniformly if } |t|\leqq t_0,$$

and

$$|\int e^{isx+tx} F_1(dx)| < \int e^{tx} F_1(dx) \qquad \text{if } s \neq 0.$$

Thus using Lemma A one gets

$$\left| v_n(x) - \frac{1}{\sqrt{n}D}\, \varphi\left(\frac{x-nM}{\sqrt{n}D}\right) \right| < \frac{C}{\sqrt{n}}$$

and formula (3.1) can be rewritten as

$$f_n(nx)=R^n(t)\,e^{-txn}\frac{1}{\sqrt{n\psi''(t)}}\left[\varphi\left(\frac{nx-n\psi'(t)}{\sqrt{n\psi''(t)}}\right)+O\left(\frac{1}{\sqrt{n}}\right)\right].$$

Choosing t as the solution of the equation $x=\psi'(t)$ one obtains (b) in the usual way.

Considering an arbitrary y, $xy>0$, and choosing t again as the solution of the equation $x=\psi'(t)$, one also gets relation (c). Now we turn to the

Proof of Theorem 1A. Let us consider a sequence $X_1, X_2, \ldots$ of i.i.d.r.v.'s with d.f. F, and a sequence $\eta_1, \eta_2, \ldots$ of i.i.d.r.v.'s with standard normal distribution. Let the η-s be independent of the X-s. Define

$$X_n'=X_n+\frac{1}{4^j}\eta_n, \qquad \text{if } 2^j<n\leqq 2^{j+1},$$

and

$$S_n'=\sum_{i=1}^{n} X_i', \qquad n=1,2,\ldots$$

We shall approximate S_n' by normal T_n in the same way as in Theorem 1 of part I. The only difference is that since the X_i-s in different blocks have different distributions, we have to change slightly the quantile transformation.

To describe this alteration, let us define the r.v.'s U_j, $\tilde{U}_j$, V_j, $\tilde{V}_j$, U_{jk}, $\tilde{U}_{nk}$, V_{jk}, $\tilde{V}_{nk}$ the same way as in Theorem 1 of part I, only substituting S_j by S_j' everywhere. Now we can define the functions

$$F_j(x)=P(\tilde{U}_j<x), \qquad F_{jk}(x\,|\,y)=P(\tilde{U}_{jk}<x\,|\,U_{jk}=y),$$

$$G_j(t)=\sup\{x: F_j(x)\leqq t\}, \qquad G_{jk}(t\,|\,y)=\sup\{x: F_{jk}(x\,|\,y)\leqq t\}.$$

We remark again that originally only the r.v.'s Y_i-s, T_i-s and V_i-s are given, but the functions F-s and G-s are known. Now the construction of the U-s is the same as in part I with the following modifications:

$$U_0=G_0(\phi(V_0)),$$

$$\tilde{U}_j=G_j(\phi(2^{-j/2}\tilde{V}_j)),$$

$$\tilde{U}_{jk}=G_{jk}(\phi(2^{-j/2}\tilde{V}_{jk})\,|\,U_{jk}),$$

where the G-s are the functions we have just defined. Now we claim that writing S_n' instead of S_n, relation (1.1) holds true for this construction. The same proof applies as in Theorem 1 of part I. One has only to check Lemma 1 of part I. Since the proof uses the formulas proved in Lemma 1, one can prove that in our case

$$|\tilde{U}_j-\tilde{V}_j|<C_1\cdot 2^{-j}\tilde{U}_j^2+C_2, \qquad \text{if } |\tilde{U}_j|<\varepsilon\cdot 2^j,$$

$$|\tilde{U}_{jk}-\tilde{V}_{jk}|<C_1\cdot 2^{-j}(\tilde{U}_{jk}^2+U_{jk}^2)+C_2, \qquad \text{if } |U_{jk}|<\varepsilon\cdot 2^j,\ |\tilde{U}_{jk}|<\varepsilon\cdot 2^j$$

and $k<2^{\alpha\cdot 2^j}$ with some $\alpha>0$. We need this new condition because it guarantees that the variance of the smoothing r.v.'s is in the range prescribed by the conditions of Lemma 2. This new condition, however, causes no trouble, since it holds for those pairs (j,k) for which we have to apply the above formula in the proof.

The r.v.'s X'_n can, by definition, be decomposed into the form $X'_n = X_n + \sigma_n \eta_n$, where $X_1, X_2, \ldots$ is a sequence of i.i.d.r.v.'s with d.f. F, $\eta_1, \eta_2, \ldots$ are independent $N(0,1)$ r.v.'s $\sigma_n = 4^{-j}$ if $2^j < n \leqq 2^{j+1}$, and the X-s and η-s are independent. Putting $S_n = \sum_{i=1}^{n} X_i$ the following remark completes the proof:

$$P\Big(\sup_{1\leqq m\leqq 2^N} |S_m - S'_m| > x\Big) = P\left(\sup_{1\leqq m\leqq 2^N} \left|\sum_{i=1}^{m} \sigma_i \eta_i\right| > x\right) < \exp(-x).$$

4. Proof of Theorem 1 B

Before proving this theorem we formulate an auxiliary statement.

We are given $2N$ real numbers $x_1, \ldots, x_{2N}$ satisfying

$$\max|x_i| \leqq K \quad \text{and} \quad \sigma^2 = \sum (x_i - \bar{x})^2 > cN, \qquad \bar{x} = \frac{1}{2N}\sum x_i. \tag{4.1}$$

Consider a random permutation π of the indices i, where each permutation of the numbers $(1, 2, \ldots, 2N)$ is chosen with the same probability $\frac{1}{(2N)!}$.

We are concerned with the random sum

$$\tilde{U} = (x_{\pi(1)} + \cdots + x_{\pi(N)}) - (x_{\pi(N+1)} + \cdots + x_{\pi(2N)}) = S_1 - S_2 = 2S_1,$$

where

$$S_1 = \sum_{i=1}^{N} (x_{\pi(i)} - \bar{x}), \qquad S_2 = \sum_{i=N+1}^{2N} (x_{\pi(i)} - \bar{x}).$$

We prove the following central limit theorem:

Lemma 3. *Under the assumption* (4.1) *we have*

a) $$P\left(S_1 < -\frac{x}{2}\sqrt{N}\right) = P(\tilde{U} < -x\sqrt{N}) = \phi\left(-\frac{x}{\sigma/\sqrt{N}}\right) \exp O\left(\frac{x^3+1}{\sqrt{N}}\right)$$

b) $$P\left(S_1 > \frac{x}{2}\sqrt{N}\right) = P(\tilde{U} > x\sqrt{N}) = \left(1 - \phi\left(\frac{x}{\sigma/\sqrt{N}}\right)\right) \exp O\left(\frac{x^3+1}{\sqrt{N}}\right)$$

for all $0 \leqq x \leqq \varepsilon\sqrt{N}$, *with* $O(\cdot)$ *uniform for these values of* x; ε *and the constant involved in* $O(\cdot)$ *depend on* K *and* c *(and not on* N *and* x_i*).*

The following lemma is a particular case of Petrov's central limit theorem for non-identically distributed r.v.'s (actually one needs to modify the proof a little) [3]:

Lemma B. *If* $\xi_1, \ldots, \xi_N$ *are random variables,*

$$E\xi_i = 0, \quad |\xi_i| \leqq K_i \quad \text{a.s.} \ (i = 1, \ldots, N),$$

$$\frac{B_N}{N} = \frac{1}{N}\sum E\xi_i^2 > c_1, \qquad \frac{1}{N}\sum K_i^3 < c_2,$$

then

$$P(S_N < -x\sqrt{B_N}) = \phi(-x)\exp O\left(\frac{x^3+1}{\sqrt{N}}\right),$$

$$P(S_N > x\sqrt{B_N}) = (1-\phi(x))\exp O\left(\frac{x^3+1}{\sqrt{N}}\right)$$

for all $0 \leqq x \leqq \varepsilon_1\sqrt{N}$, *with* $O(\cdot)$ *uniform for these values of* x; ε_1 *and the constant involved in* $O(\cdot)$ *depend on* c_1 *and* c_2 *(and not on* N *and* K_i*).*

Proof of Lemma 3. We prove only b), the proof of part a) is similar.

We may assume $\sum x_i = 0$.

Let $r_1, \ldots, r_N$ be i.i.d.r.v.'s taking the values ± 1 with probability 1/2 and let the r_i-s be independent of π. Consider the random variable

$$\bar{U} = \sum_{i=1}^{N} (x_{\pi(i)} - x_{\pi(N+i)})\, r_i .$$

It is clear that $\bar{U}$ has the same distribution as $\tilde{U}$. Thus, we have to estimate the probability $P(\bar{U} > x\sqrt{N})$. For given $\pi = (\pi(1), \ldots, \pi(2N))$ we can apply Lemma B for the variables $\xi_i = (x_{\pi(i)} - x_{\pi(N+i)})\, r_i$ if only

$$\sigma_\pi^2 = \sum_{i=1}^{N} (x_{\pi(i)} - x_{\pi(N+i)})^2 > c_1 N .$$

Put $c_1 = \frac{c}{2}$. For such a π, thus,

$$P(\bar{U} > x\sqrt{N} \mid \pi) = \left(1 - \phi\left(\frac{x}{\sigma_\pi/\sqrt{N}}\right)\right) \exp O\left(\frac{x^3+1}{\sqrt{N}}\right)$$

for all $0 \leqq x \leqq \varepsilon_1\sqrt{N}$, where $O(\cdot)$ is uniform for these values of x and all those π.

Now

$$P(\bar{U} > x\sqrt{N}) = EP(\bar{U} > x\sqrt{N} \mid \pi) = \frac{1}{(2N)!} \sum_{\pi:\, \sigma_\pi^2 \leqq c_1 N} P(\bar{U} > x\sqrt{N} \mid \pi)$$
$$+ \frac{1}{(2N)!} \sum_{\pi:\, \sigma_\pi^2 > c_1 N} \left[1 - \phi\left(\frac{x}{\sigma_\pi/\sqrt{N}}\right)\right]$$
$$\cdot \exp O\left(\frac{x^3+1}{\sqrt{N}}\right), \qquad 0 \leqq x \leqq \varepsilon_1\sqrt{N}. \tag{4.2}$$

Since

$$E\sigma_\pi^2 = \frac{1}{(2N)!} \sum_{\pi} \sum_{i=1}^{N} (x_{\pi(i)} - x_{\pi(N+i)})^2 = \sigma^2 \frac{2N}{2N-1},$$

it is natural to expect that the above distribution—being near the mixture of normal distributions of variance σ_π^2/N—is near a normal distribution of variance

σ^2/N. Thus, we will estimate the ratio

$$\frac{P(\bar{U}>x\sqrt{N})}{1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)}.$$

We will show that

$$\frac{1}{(2N)!}\sum_{\pi:\,\sigma_\pi^2\leqq c_1N} P(\bar{U}>x\sqrt{N}\,|\,\pi)\leqq\frac{1}{(2N)!}\sum_{\pi:\,\sigma_\pi^2\leqq c_1N}1=O(e^{-\alpha N}),\tag{4.3}$$

and thus, if $\varepsilon<\varepsilon_1$ is small enough $\left(\text{whence } 1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)>e^{-\frac{\alpha}{2}N}, 0\leqq x\leqq\varepsilon\sqrt{N}\right)$, we will have

$$0\leqq\frac{\frac{1}{(2N)!}\sum_{\sigma_\pi^2\leqq c_1N}P(\bar{U}>x\sqrt{N}\,|\,\pi)}{1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)}\leqq e^{-\frac{\alpha}{2}N}=O\left(\frac{x^3+1}{\sqrt{N}}\right),\quad 0\leqq x\leqq\varepsilon\sqrt{N}.$$

Since the terms $O(\cdot)$ on the right-hand side of (4.2) are uniform for those π, for which $\sigma_\pi^2>c_1N$, to establish the estimation

$$P(\bar{U}>x\sqrt{N})\Big/\left(1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)\right)=\exp O\left(\frac{x^3+1}{\sqrt{N}}\right)$$

it remains to show that

$$\Delta=\frac{1}{(2N)!}\sum_{\sigma_\pi^2>c_1N}\frac{1-\phi\left(\frac{x}{\sigma_\pi/\sqrt{N}}\right)}{1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)}=\exp O\left(\frac{x^3+1}{\sqrt{N}}\right),\quad 0\leqq x\leqq\varepsilon\sqrt{N}.$$

We actually prove the stronger estimation

$$\Delta=\exp O\left(\frac{x^4+1}{N}\right),\quad 0\leqq x\leqq\varepsilon\sqrt{N}.\tag{4.4}$$

We use the inequality

$$\exp\left\{\frac{\varphi(t)}{1-\phi(t)}(y-t)+A_2(y-t)^2\right\}\leqq\frac{1-\phi(t)}{1-\phi(y)}\leqq\exp\left\{\frac{\varphi(t)}{1-\phi(t)}(y-t)+A_1(y-t)^2\right\}\tag{4.5}$$

for all non-negative t and y, where $0<A_1<A_2$ are universal constants.

Since $c<\frac{\sigma^2}{N}<2K^2$ and $\frac{\sigma_\pi^2}{N}<4K^2$, in case $c_1<\frac{\sigma_\pi^2}{N}$ (4.5) implies

$$\frac{1-\phi\left(\frac{x}{\sigma_\pi/\sqrt{N}}\right)}{1-\phi\left(\frac{x}{\sigma/\sqrt{N}}\right)}=\exp\left\{\frac{1}{x^2}g\left(\frac{x}{\sigma/\sqrt{N}}\right)\left(\frac{\sigma_\pi^2}{N}-\frac{\sigma^2}{N}\right)+O\left((x^2+1)\left(\frac{\sigma_\pi^2-\sigma^2}{N}\right)^2\right)\right\},\tag{4.6}$$

where

$$g(u)=\frac{\varphi(u)\,u^3}{2(1-\phi(u))}.$$

The right-hand side of (4.6) is dominated by

$$\exp(c'\,x^2)\leqq\exp(c'\,\varepsilon^2\,N),$$

and

$$\frac{1}{(2N)!}\sum_{\sigma_\pi^2\leqq c_1 N} 1=O(e^{-\alpha N}),$$

thus choosing ε so small that $c'\,\varepsilon^2<\frac{\alpha}{2}$, we see that our task (in addition to prove (4.3)) is to establish the relation

$$I=E\exp\left\{\frac{1}{x^2}\,g\left(\frac{x}{\sigma/\sqrt{N}}\right)\frac{2\eta}{\sqrt{N}}+\eta^2\,O\left(\frac{x^2+1}{N}\right)\right\}=\exp O\left(\frac{x^4+1}{N}\right), \tag{4.7}$$

where

$$\eta=\frac{\sigma_\pi^2-\sigma^2}{2\sqrt{N}}=-\frac{\sum_{i=1}^{N} x_{\pi(i)}\,x_{\pi(N+i)}}{\sqrt{N}}.$$

One expects that the variable η is approximately normal, and thus

$$E\,e^{t\eta}<e^{Ct^2} \tag{4.8}$$

and $E\,e^{u\eta^2}<e^{Cu}$. Taking

$$t=\frac{2}{x^2}\,g\left(\frac{x}{\sigma/\sqrt{N}}\right)\frac{1}{\sqrt{N}}=O\left(\frac{x^2+1}{\sqrt{N}}\right),$$

and $u=\frac{x^2+1}{N}$ these inequalities imply that

$$I^2\leqq E\exp\{2\,t\,\eta\}\cdot E\exp\{2\,C\,u\,\eta^2\}=\exp O\left(\frac{x^4+1}{N}\right)$$

$$I\geqq E\exp\{+t\,\eta-C\,u\,\eta^2\}\geqq E^2\exp\{t\,\eta\}\,[E\exp\{t\,\eta+C\,u\,\eta^2\}]^{-1}$$

$$\geqq E^2\exp\{t\,\eta\}\cdot[E\exp\{2\,t\,\eta\}\,E\exp\{2\,C\,u\,\eta^2\}]^{-\frac{1}{2}}=\exp O\left(\frac{x^4+1}{N}\right),$$

that is (4.7). The inequality (4.8), however can be expected to hold only for $|t|>\frac{\alpha}{\sqrt{N}}$, since

$$E\,\eta=\frac{\sum x_i^2}{2\sqrt{N}(2N-1)}>\frac{c}{2}\,\frac{1}{\sqrt{N}}.$$

In fact, we show that

$$E\,e^{t\eta} < e^{Ct^2} \quad \text{for } \frac{1}{\sqrt{N}} \leqq |t| \leqq \frac{1}{16K^2}\sqrt{N} \tag{4.9}$$

and

$$E\,e^{u\eta^2} < e^{Cu} \quad \text{for } 0 \leqq u \leqq u_0, \tag{4.10}$$

where u_0 depends on C and K only. For $0 \leqq t \leqq \frac{1}{\sqrt{N}}$, i.e. $0 \leqq x \leqq L$ we have

$$\begin{aligned} &E\exp\left\{\frac{2}{x^2}\, g\left(\frac{x}{\sigma/\sqrt{N}}\right)\frac{\sigma_\pi^2-\sigma^2}{N}\right\} \\ &\quad \leqq E\left(1+\frac{\sigma_\pi^2-\sigma^2}{N}\,\frac{2}{x^2}\, g\left(\frac{x}{\sigma/\sqrt{N}}\right)+L'\left(\frac{\sigma_\pi^2-\sigma^2}{N}\right)^2\right) \\ &\quad \leqq 1=\frac{L_1}{\sqrt{N}}\,E\eta+L'\frac{4}{N}\,E\eta^2=\exp O\left(\frac{1}{N}\right), \end{aligned}$$

since by (4.10)

$$E\eta^2 \leqq \frac{1}{u_0}\,E\,e^{u_0\eta^2} < \frac{1}{u_0}\,e^{Cu_0} < \infty.$$

This estimation and (4.10) imply (4.7) again. Thus, we have to establish (4.9), (4.10) and (4.3).

First we estimate the conditional expectation

$$E(e^{t\eta}\,|\,\pi(1), \ldots, \pi(N)).$$

Write $b_1, \ldots, b_N$ for $x_{\pi(1)}, \ldots, x_{\pi(N)}$, and $(y_1, \ldots, y_N)$ for a random permutation of the set

$$D=\{x_1, \ldots, x_{2N}\}/\{b_1, \ldots, b_N\}.$$

Put $\bar{b}=\frac{1}{N}(b_1+\cdots+b_N)$, and $a_i=\bar{b}-b_i$.

Then,

$$\eta=-\frac{1}{\sqrt{N}}\sum_{i=1}^{N} b_i y_i=-\frac{(\sum b_i)(\sum y_i)}{N^{\frac{3}{2}}}+\frac{1}{\sqrt{N}}\sum a_i y_i=\frac{(\sum b_i)^2}{N^{\frac{3}{2}}}+\frac{1}{\sqrt{N}}\sum a_i y_i.$$

Thus,

$$E(e^{t\eta}\,|\,\pi(1), \ldots, \pi(N))=e^{\frac{t}{M}(\sum b_i)^2}\,E e^{\frac{t}{\sqrt{N}}\sum a_i y_i},$$

where in the expectation on the right-hand side the numbers a_i are given, $\sum a_i=0$, $\max|a_i| \leqq 2K$, and the N-tuple $(y_1, \ldots, y_N)$ runs over all the $N!$ permutations of the elements of D and $M=N^{3/2}$. We apply the following

Lemma 4. *If $a_1, \dots, a_N$ are real numbers, $\sum a_i = 0$, and the N-tuple $(y_1, \dots, y_N)$ runs through all permutations of the numbers $(d_1, \dots, d_N)$, then for any real t*

$$Ee^{t\Sigma a_i y_i} = \frac{1}{N!} \sum_{(y_1, \dots, y_N) = \operatorname{perm}(d_1, \dots, d_N)} e^{t\Sigma q_i y_i} \leqq e^{t^2 (\max a_i^2) \Sigma d_i^2},$$

and consequently

$$Ee^{u\frac{(\Sigma a_i y_i)^2}{D_N}} \leqq e^{C_0 u}$$

for $0 \leqq u \leqq u_0 < \frac{1}{4}$, where $D_N = (\max a_i^2) \sum d_i^2$, and the constant C_0 depends only on u_0.

By Lemma 4

$$Ee^{\frac{t}{\sqrt{N}}\Sigma a_i y_i} \leqq e^{\frac{t^2}{N}(\max a_i^2)\Sigma y_i^2} \leqq e^{8K^4 t^2}.$$

Hence, again with $M = N^{3/2}$

$$Ee^{t\eta} \leqq e^{8K^4t^2} Ee^{\frac{t}{M}(\Sigma b_i)^2} = e^{8K^4t^2} Ee^{\frac{t}{4M}\tilde{U}^2} \leqq e^{8K^4t^2} e^{\frac{C_0|t|}{M}\Sigma x_i^2} \leqq e^{8K^4t^2 + 2K^2 C_0 \frac{|t|}{\sqrt{N}}}$$

for $|t| \leqq \frac{1}{16K^2}\sqrt{N}$. If, in addition, $|t| \geqq \frac{1}{\sqrt{N}}$, then we get (4.9).

For (4.10) it is sufficient to show that $Ee^{u_0\eta^2} < \infty$, which follows immediately from (4.9) and the fact $\eta^2 \leqq K^4 N$.

Now we prove (4.3), i.e. the relation

$$P(\sigma_\pi^2 \leqq c_1 N) = O(e^{-\alpha N}).$$

$$\sigma_\pi^2 \leqq c_1 N = \frac{c}{2} N \text{ implies}$$

$$2\sum x_{\pi(i)} x_{\pi(N+i)} \geqq \frac{c}{2} N, \quad \text{i.e. } \eta \leqq -\frac{c}{4}\sqrt{N}.$$

Thus, (4.9) implies (4.3). (Note that in the proof of (4.9) we have not used (4.3).)

For the proof of Lemma 4 we use the following

Lemma 5. *Let $\mathbf{a} = (a_1, \dots, a_N)$ be an N-dimensional vector with $\sum a_i = 0$, $\max |a_i| = K$. Then $\mathbf{a}$ is a finite linear combination*

$$\mathbf{a} = \sum \alpha_k \mathbf{R}_k, \quad \alpha_k > 0, \quad \sum \alpha_k = K,$$

where each vector $\mathbf{R}_k$ has half the coordinates 1 and half (-1) if N is even, and $\frac{N-1}{1}$ 1's, $\frac{N-1}{2}$ (-1)'s and one 0 if N is odd.

Lemma 5 is a consequence of the fact that in the convex set $\{\mathbf{a}: \sum a_i = 0, \max |a_i| = 1\}$ the above mentioned vectors $\mathbf{R}$ are the extreme points.

We prove Lemma 4 first in the case $\mathbf{a} = \mathbf{R}$, i.e.when $|a_i| = 1$ for all i except for at most one (if N is odd). Since we can disregard this coordinate, we may assume (in the case $\mathbf{a} = \mathbf{R}$) that N is even, $N = 2M$. The same way as above we consider the random variable $S_2 = \sum (y_i - y_{M+i}) r_i$ instead of the variable $S_1 = \sum (y_i - y_{M+i})$, where $r_1, \dots, r_M$ are i.i.d.r.v.'s taking the values ± 1 with probability $\frac{1}{2}$, and are independent of the random permutation. S_2 and S_1 have the same distribution. Thus (if π denotes the random permutation)

$$Ee^{tS_1}=Ee^{tS_2}=EE(e^{tS_2}|\pi)=E\prod_i E(e^{t(y_i-y_{M+i})r_i}|\pi)$$

$$=E\prod_i \operatorname{ch}[(y_i-y_{M+i})t]\leqq E\prod_i e^{t^2(y_i^2+y_{M+i}^2)}=Ee^{t^2\Sigma y_i^2}=e^{t^2\Sigma d_i^2}.$$

To prove the lemma in the case of a general **a**, we apply Lemma 5 and the following Hölder-type inequality

$$|\int f_1^{p_1}\dots f_k^{p_k}|\leqq(\int|f_1|)^{p_1}\dots(\int|f_k|)^{p_k},\qquad p_i\geqq 0,\qquad \sum p_i=1.$$

Hence,

$$Ee^{t(\mathbf{a},\,y)}=Ee^{\sum_k \alpha_k t(\mathbf{R}_k,\,y)}\leqq\prod_k\left(Ee^{t(\Sigma\alpha_i)(\mathbf{R}_k,\,y)}\right)^{\frac{\alpha_k}{\Sigma\alpha_i}}$$

$$\leqq\prod_k\left(e^{t^2(\Sigma\alpha_i)^2\Sigma d_i^2}\right)^{\frac{\alpha_k}{\Sigma\alpha_i}}=e^{t^2\max a_i^2)\,\Sigma d_i^2}.$$

Thus Lemma 4, and hence Lemma 3 is proved.

Now we turn to the

Proof of Theorem 1B. Let $X_1',\dots,X_N'$, $N=2^n$ be i.i.d.r.v.'s with distribution function $F(x)$. Let $X_1^*,\dots,X_N^*$ be their order statistics, and

$$S_N=\sum_{i=1}^N X_i'=\sum_{i=1}^N X_i^*.$$

Let T_N be a normally distributed random variable with expectation 0 and variance N, such that

$$|S_N-T_N|\leqq C_1\frac{S_N^2}{N}+C_2\qquad\text{if } |S_N|<\varepsilon N,$$

where C_1, C_2 and ε depend only on F (and not on N). The existence of such a T_N is guaranteed by formula (2.6) of Lemma 1 in part I (note that in the proof of formula (2.6) of Lemma 1 in part I we used only the conditions

$$\int xF(dx)=0,\qquad\int x^2F(dx)=1,\qquad\int e^{tx}F(dx)<\infty,\quad |t|<t_0,$$

and did not use the smoothness conditions (i) or (ii) – they were needed only for proving (2.7), i.e. for handling the conditional distributions).

Let $\tilde V_{m,k}(m=1,2,\dots,n;k=0,1,\dots;(k+1)2^m\leqq N)$ be independent normally distributed random variables, independent of $\{X_1',\dots,X_N',T_N\}$, for which

$$E\tilde V_{m,k}=0,\qquad E\tilde V_{m,k}^2=2^m.$$

We denote T_N also by $V_{n,0}$, and define the random variables $Y_1,\dots,Y_N$ by the formulae

$$T_0=0,\qquad T_j=\sum_{i=1}^j Y_i,\quad j=1,\dots,N,$$

$$V_{j,k}=T_{(k+1)2^j}-T_{k\cdot 2^j},\quad j=0,1,\dots,n;k=0,1,\dots;(k+1)2^j\leqq N,$$

$$\tilde V_{m,k}=V_{m-1,2k}-V_{m-1,2k+1},\qquad m=1,2,\dots,n;k=0,1,\dots;(k+1)2^m\leqq N.$$

It is easy to see (since the variables $\tilde{V}_{m,k}$ and $V_{n,0}$ are given) that the variables $Y_1, \ldots, Y_N$ are uniquely defined by these formulae, and form a sequence of i.i.d. standard normal random variables.

Our aim is to define the variables $X_1, \ldots, X_N$ in such a way that $X_1^*, \ldots, X_N^*$ be the order statistics of $X_1, \ldots, X_N$ (hence $S_N = X_1 + \cdots + X_N$), and

$$P\left(\sup_{k \leq N} |S_k - T_k| > C \log N + x\right) < K e^{-\lambda x} \tag{4.11}$$

for all $x > 0$. (Putting independent blocks of length 2^{2^n} together, one easily gets the infinite analogue of (4.11), i.e. Theorem 1 B.)

Taking

$$S_0 = 0, \quad S_j = \sum_{i=1}^{j} X_i, \quad j = 1, \ldots, N$$

$$U_{j,k} = U_{(k+1)2^j} - U_{k \cdot 2^j}, \quad j = 0, 1, \ldots, n; k = 0, 1, \ldots; (k+1) 2^j \leqq N$$

$$\tilde{U}_{m,k} = U_{m-1,2k} - U_{m-1,2k+1}, \quad m = 1, 2, \ldots, n; k = 0, 1, \ldots; (k+1) 2^m \leqq N,$$

we proceed the same "dyadic" way as in part I, i.e. given $S_N = U_{n,0}$ we define first the variable

$$\tilde{U}_{n,0} \left(= \sum_{i=1}^{N/2} X_i - \sum_{i=N/2+1}^{N} X_i \right),$$

which (together with S_N) determines $U_{n-1,0} (= S_{N/2})$ and $U_{n-1,1} (= S_N - S_{N/2})$, and then keep going on the same way down to the individual terms.

We will obtain the estimation

$$|\tilde{U}_{m,k} - \tilde{V}_{m,k}| \leqq C(2^{-m} \tilde{U}_{m,k}^2 + 2^{-m} U_{m,k}^2 + 2^{-m} W_{m,k}^2 + 1), \tag{4.12}$$

if only

$$|\tilde{U}_{m,k}| < \varepsilon \cdot 2^m, \quad |U_{m,k}| < \varepsilon \cdot 2^m \quad \text{and} \quad |W_{m,k}| < \varepsilon \cdot 2^m, \tag{4.13}$$

where

$$W_{m,k} = \sum_{i=k \cdot 2^m + 1}^{(k+1) 2^m} (X_i^2 - 1).$$

(4.12) is analogous to (2.7) of Lemma 1 in part I. The only modification in the dyadic construction here is that now given $S_N = U_{n,0}$ *and* $X_1^*, \ldots, X_N^*$ (i.e. the set of the values $\{X_1, \ldots, X_N\}$, but not their particular order), we want to define

$$\tilde{U}_{n,0} = S_{N/2} - (S_N - S_{N/2})$$

and the set of values $\{X_1, \ldots, X_{N/2}\}$ (but not yet their order), and so on at each step. Since the N-tuple $(X_1, \ldots, X_N)$ is defined if we specify a permutation $\pi = (\pi_1, \ldots, \pi_N)$, by taking

$$X_i = X_{\pi_i}^*, \quad i = 1, \ldots, N,$$

what we do is to define π by assigning first the two halves

$$H_{1,1} = \{\pi_1, \ldots, \pi_{N/2}\}, \quad H_{1,2} = \{\pi_{N/2+1}, \ldots, \pi_N\}$$

and then keep going on halving them until we get the whole permutation. If we want to ensure that the obtained variables $X_1, \ldots, X_N$ are independent and distributed according to F, all we have to check is that for the random permutation $\pi(\omega)$

$$P(\pi=(p_1, \ldots, p_N))=\frac{1}{N!} \tag{4.14}$$

for any permutation $(p_1, \ldots, p_N)$ of the first N integers.

Now we define the halving $(H_{1,1}, H_{1,2})$ as follows:

To every subset $H=\{h_1, \ldots, h_{N/2}\}$ of $\{1, \ldots, N\}$ with $N/2$ elements we assign the number

$$\tilde{U}_H=\sum_{i\in H} X_i^* - \sum_{i\notin H} X_i^*.$$

Thus we get $A=\binom{N}{N/2}$ numbers and we order them in increasing order of magnitude,

$$\tilde{U}_{H_1} \leqq \cdots \leqq \tilde{U}_{H_A} \quad \text{(if two sums are equal, we order them arbitrarily).}$$

Let

$$I_1=(-\infty, a_1), \quad I_2=[a_1, a_2), \ldots, I_{A-1}=[a_{A-2}, a_{A-1}), \quad I_A=[a_{A-1}, \infty)$$

be the disjoint intervals, for which

$$\int_{I_k} \frac{e^{-t^2/2}\,dt}{\sqrt{2\pi}}=\frac{1}{A}, \quad k=1, \ldots, A.$$

Given ω we define

$$H_{1,1}=H_k \quad (\text{and } H_{1,2}=\{1, \ldots, N\}/H_k)$$

if $\tilde{V}_{n,0}\in I_k$. In other words, $\tilde{U}_{n,0}$ is defined as a conditional quantile transform of $\tilde{V}_{n,0}$:

$$\tilde{U}_{n,0}=G_n(\phi(2^{-n/2}\,\tilde{V}_{n,0})|X_1^*, \ldots, X_N^*),$$

where we define $G_k(t|y_1, \ldots, y_{2^k})$ as follows: Let $\xi_1, \ldots, \xi_{2^k}$ be i.i.d.r.v.'s with d.f. F, and $\xi_1^*, \ldots, \xi_{2^k}^*$ be the order statistics.

Put

$$F_k(x|y_1, \ldots, y_{2^k})=P\left(\sum_{i=1}^{2^{k-1}} \xi_i - \sum_{i=2^{k-1}+1}^{2^k} \xi_i < x \,\middle|\, X_1^*, \ldots, X_{2^k}^*\right),$$

and let G_k be the inverse of F_k:

$$G_k(t|y_1, \ldots, y_{2^k})=\sup\{x\colon F_k(x|y_1, \ldots, y_{2^k})\leqq t\}.$$

Having $(H_{1,1}, H_{1,2})$ we define the halving $(H_{2,1}, H_{2,2})$ of $H_{1,1}$ the same way using $\tilde{V}_{n-1,0}$ instead of $\tilde{V}_{n,0}$, and the halving $(H_{2,3}, H_{2,4})$ of $H_{1,2}$ using $\tilde{V}_{n-1,1}$, etc.

Thus, each $\tilde{U}_{m,k}$ is actually defined as a conditional quantile transform of $\tilde{V}_{m,k}$, e.g.

$$\tilde{U}_{n-1,0} = G_{n-1}(\phi(2^{-\frac{n-1}{2}} \tilde{V}_{n-1,0}) | X^*_{1,1}, \ldots, X^*_{1,N/2}),$$

where $X^*_{1,1} \leqq \cdots \leqq X^*_{1,N/2}$ are the ordered $N/2$-tuple formed from the set $\{X^*_i : i \in H_{1,1}\}$. Since $H_{1,1} = H_{1,1}(\omega)$ takes all possible "values" $\left(\binom{N}{N/2} \text{ subsets}\right)$ with equal probability and this holds at each step for the constructed random halving, further these halvings were independent due to the independence of the variables $\tilde{V}_{m,k}$ and that of $\tilde{V}_{m,k}$ and X^*_i, the obtained random permutation obeys (4.14).

Now we prove (4.12). As a consequence of (4.13), condition (4.1) holds, and thus we may apply Lemma 3. Using inequality (4.5) and condition (4.1) we may rewrite part b) of Lemma 3 as follows:

$$\begin{aligned} &1 - F_m(x | x_1, \ldots, x_{2m}) \\ &\quad = (1-\phi(x)) \times \exp\left\{ \frac{x\varphi(x)}{1-\phi(x)} \, \frac{\sigma^2 \cdot 2^{-m} - 1}{2} \right. \\ &\qquad \left. + O((x^2+1)(\sigma^2 \cdot 2^{-m} - 1)^2) + O((x^3+1)2^{-m/2}) \right\}, \\ &0 \leqq x \leqq \varepsilon \cdot 2^{m/2}. \end{aligned}$$

Writing $1 - F_m(x | x_1, \ldots, x_{2m}) = 1 - \phi(x+\delta)$, applying (4.5) again and using the identity

$$\sigma^2 \cdot 2^{-m} - 1 = 2^{-m} \sum_{i=1}^{2^m} (x_i^2 - 1) - \left(2^{-m} \sum_{i=1}^{2^m} x_i\right)^2,$$

we get for δ:

$$\begin{aligned} |\delta \cdot 2^{m/2}| \leqq\ & x(2^{-m/2} |\textstyle\sum (x_i^2 - 1)| + 2^{-\frac{3m}{2}} (\sum x_i)^2) \\ & + C_1(x+1)(2^{-\frac{3m}{2}} (\textstyle\sum (x_i^2-1))^2 + 2^{-\frac{7m}{2}} (\sum x_i^4)) + C_2(x^2+1). \end{aligned}$$

Using the inequalities

$$|\textstyle\sum x_i| \leqq K \cdot 2^m, \qquad |\sum (x_i^2 - 1)| \leqq (K^2+1) 2^m,$$

we get

$$|\delta \cdot 2^{m/2}| \leqq C \left[x^2 + \left(2^{-m/2} \sum_{i=1}^{2^m} x_i\right)^2 + \left(2^{-m/2} \sum_{i=1}^{2^m} (x_i^2 - 1)\right)^2 + 1 \right].$$

The same inequality can be proved on the negative semi-axis, i.e. for

$$F_m(-x | x_1, \ldots, x_{2m}),$$

and thus we have proved (4.12) for $k=0$.

Since the joint distributions of $\{\tilde{U}_{m,k}, \tilde{V}_{m,k}, U_{m,k}, W_{m,k}\}$ are the same for $k = 0, 1, \ldots$, (4.12) is proved for all m, k.

To show that (4.12) implies (4.11), we may proceed the same way as in part I (at the very end of the proof of Theorem 1). The only change is that in estimating Δ_4 we have to replace the sum

$$\sum_{j=M+1}^{n} 2^{-j}(\tilde{U}_{j,0}^2+U_{j,0}^2+1) \quad \text{by} \quad \sum_{j=M+1}^{n} 2^{-j}(\tilde{U}_{j,0}^2+U_{j,0}^2+W_{j,0}^2+1)$$

and introduce the variable $\tilde{W}_j=W_{j+1,0}-W_{j,0}$. Then the same estimation can be carried out for $W_{j,0}$ and $\tilde{W}_j$ as for U_j and $\tilde{U}_j$, namely

$$\sum_{j=M}^{n-1} 2^{-j} W_{j,0}^2 \leqq \frac{2}{(\sqrt{2}-1)^2}\left(2^{-M} W_{M,0}^2+\sum_{j=M}^{n-2} 2^{-j}\tilde{W}_j^2\right),$$

whence we boil down to the same estimation

$$2^{n-M}\cdot P\left(2C_3\sum_{j=1}^{n}\tau_j'+NC_2>\frac{x}{4}\right)\leqq\gamma_4\, e^{\alpha_4 N-\beta_4 x},$$

where

$$\tau_j'=\begin{cases}2^{-j}\tilde{W}_j^2 & \text{if } |\tilde{W}_j|<\varepsilon\cdot 2^j\\ 0 & \text{otherwise},\end{cases}$$

the τ_j'-s are independent, and $E e^{t\tau_j'}<\infty$ for $0<t<\delta$, since the variables (X_i^2-1) have expectation 0 and are bounded.

5. The Proof of Theorem 3

The line of the proof of Theorem 3 is the same as that of Theorem 1. Theorem 1 B and the argument of proving Theorem 1 show that it is sufficient to prove Theorem 3 with the additional assumption that F has an absolute continuous component. Therefore we assume this in the sequel, and we try to apply the construction of Theorem 1 A. The main problem constitutes in finding good asymptotics for the appearing conditional and unconditional distributions. To this end we select the following method: we truncate the r.v.'s X_i in such a way that the outer part has negligible influence on the difference S_n-T_n. We approximate these truncated random variables by normal r.v.'s. They have bounded moment generating functions, therefore one can apply the technique of conjugated distributions. But the range where the moment generating function behaves nicely, depends on the level of the truncation, and hence also on the index of the truncated random variables. The length of this range tends to 0 as n tends to infinity. That is why we get weaker approximations if the moment generating function does not exist.

Since we may have different truncations for different n, the truncation has an influence on the form of the asymptotic formulas. We must however check that the $O(\cdot)$ in the appearing formulas are uniform not only in x but also in n. The numbers $C, C_1, \ldots$ will denote appropriate constants in the sequel. The same letter may denote different constants in different formulas.

Lemma 6 helps to find the appropriate level of truncation.

Lemma 6. *Given a monotone increasing function satisfying* (i) *in formula* (1.2) *consider a sequence* $X_1, X_2, \ldots$ *of i.i.d.r.v.'s such that* $EX_1=0$, $EX_1^2=1$, $EH(|X_1|)<\infty$. *Define the number* K_n *by the equation* $H(K_n)=n$, *and the r.v.'s* $\tilde{X}_1, \tilde{X}_2, \ldots$ *in the following way*

$$\tilde{X}_m=\begin{cases} X_m & \text{if } |X_m|<K_{2^n} \text{ (if only } 2^n\leqq m<2^{n+1}) \\ 0 & \text{otherwise.}\end{cases}$$

Let $\bar{X}_k=\dfrac{\tilde{X}_k-E\tilde{X}_k}{D\tilde{X}_k}$, *and put* $\tilde{S}_n=\sum_{k=1}^{n}\tilde{X}_k$, $\bar{S}_n=\sum_{k=1}^{n}\bar{X}_k$. *Then we have the following statements:*

a) $X_k(\omega)=\tilde{X}_k(\omega)$ *if* $k>k_0(\omega)$

b) $\dfrac{S_n-\bar{S}_n}{K_n}\to 0$ *a.s.*

Proof. a) $\sum P(X_m\neq\tilde{X}_m)\leqq\sum P(H(|X_m|)\geqq\frac{1}{2}m)<\infty$.

b) We prove the following somewhat stronger result $\sum\dfrac{X_m-\bar{X}_m}{\bar{K}_m}$ is convergent, where $\bar{K}_m=K_{2^n}$, if only $2^n\leqq m<2^{n+1}$. This is, indeed, a stronger statement, since $K_m\leqq K_{2m}\leqq 2K_m$ by virtue of (i) in (1.2). It is sufficient to check that the series

$$\sum\frac{E(\tilde{X}_m-\bar{X}_m)}{\bar{K}_m} \quad \text{and} \quad \sum\frac{D^2(\tilde{X}_m-\bar{X}_m)}{\bar{K}_m^2}$$

are convergent. Since $E(\tilde{X}_m-\bar{X}_m)=E\tilde{X}_m$,

$$D^2(\tilde{X}_m-\bar{X}_m)=(1-D\tilde{X}_m)^2\leqq 1-D^2\tilde{X}_m$$

we have to prove that

$$\sum 2^n\frac{E\tilde{X}_{2^n}}{K_{2^n}}$$

is convergent and

$$\sum 2^n\frac{1-E\tilde{X}_{2^n}^2}{K_{2^n}^2}<\infty.$$

Now

$$\begin{aligned}\sum_{n=1}^{\infty}2^n\frac{1-E\tilde{X}_{2^n}^2}{K_{2^n}^2}&=\sum_{n=1}^{\infty}\sum_{j=n}^{\infty}\frac{2^n}{K_{2^n}^2}\int_{\{2^j\leqq H(|x|)<2^{j+1}\}}x^2F(dx)\\ &\leqq\sum_{n=1}^{\infty}\frac{2^n}{K_{2^n}^2}\sum_{j=n}^{\infty}K_{2^{j+1}}^2P(2^j\leqq H(|X_1|)<2^{j+1})\\ &=\sum_{j=1}^{\infty}P(2^j\leqq H(|X_1|)<2^{j+1})K_{2^{j+1}}^2\sum_{n=1}^{j}\frac{2^n}{K_{2^n}^2}.\end{aligned}$$

This sum is finite since

$$K_{2^{j+1}}^2\sum_{n=1}^{j}\frac{2^n}{K_{2^n}^2}<C\cdot 2^j$$

which, in turn, follows from the following estimation: by the monotonicity of $H(x)\,x^{-3-\delta}$ one gets for $n \leqq j$

$$\frac{2^{j+1}}{K_{2^{j+1}}^{3+\delta}} \geqq \frac{2^n}{K_{2^n}^{3+\delta}},$$

thus

$$\left(\frac{K_{2^{j+1}}}{K_{2^n}}\right)^2 \leqq 2^{(j+1-n)\frac{2}{3+\delta}}.$$

The other sum can be estimated similarly.

Lemma 7 is the analogue of Lemma 2.

Lemma 7. *Let $X_1, \ldots, X_n$ be i.i.d.r.v.'s, $EX_1=0$, $D^2 X_1<\infty$, $EH(|X_1|)<\infty$, where $H(x)$ is a monotone increasing, positive, continuous function satisfying* (1.2). *Let the distribution function of X_1 have an absolute continuous component. Let us be given a sequence $\eta_1, \ldots, \eta_n$ of i.i.d.r.v.'s with normal distribution, $E\eta_1=0$, $E\eta_1^2=\frac{1}{n^2}$, and let the η-s be independent of the X-s.*

Define the number K_n by the equation $H(K_n)=n$, the number u_n by $u_n=c'\frac{\log n}{K_n}$, and the random variables

$$\tilde{X}_i=\begin{cases} X_i & \text{if } |X_i|<K_n \\ 0 & \text{otherwise}, \end{cases}$$

$$\bar{X}_i=\frac{\tilde{X}_i-E\tilde{X}_i}{D\tilde{X}_i}, \qquad i=1, \ldots, n,$$

and

$$S_k=\sum_{i=1}^{k}(\bar{X}_i+\eta_i), \qquad k=1, \ldots, n.$$

Put

$$F_k^{(n)}(x)=P(S_k<x) \quad \text{and} \quad f_k^{(n)}(x)=\frac{d}{dx}F_k^{(n)}(x).$$

Then we have the following relations for $C_0 \log n<k \leqq n$:

a) $1-F_k^{(n)}(kx)=[1-\phi(k^{\frac{1}{2}}x)]\exp[kx^3\lambda_n(x)+O(x+k^{-\frac{1}{2}})] \qquad 0 \leqq x \leqq cu_n$

a′) $F_k^{(n)}(-kx)=\phi(-k^{\frac{1}{2}}x)\exp[-kx^3\lambda_n(-x)+O(x+k^{-\frac{1}{2}})] \qquad 0 \leqq x \leqq cu_n$

b) $f_k^{(n)}(kx)=\varphi(k^{\frac{1}{2}}x)\exp[kx^3\lambda_n(x)+O(|x|+k^{-\frac{1}{2}})]$

in the interval $|X| \leqq cu_n$ and $O(\cdot)$ is uniform in n, k, and X. Further, one has

$$f_k^{(n)}(ky) \leqq C_1 f_k^{(n)}(kx)\, e^{-tk(y-x)}$$

if $|X| \leqq u_n$, $xy>0$, and t is the solution of the equation $\frac{R_n'(t)}{R_n(t)}=x$, where

$$R_n(t)=E\exp(t\bar{X}_1).$$

Proof. The proof if similar to that of Lemma 2. The only difference is that we have to consider the conjugated distributions

$$V(dx) = e^{tx} \frac{\bar{F}(dx)}{E \exp(t\bar{X}_1)},$$

($\bar{F}(x)$ is the distribution function of $\bar{X}_1$) only in the range $|t| \leqq t_0 = c' \frac{\log n}{K_n}$. c' depends only on n and we get the expansion for

$$|X| \leqq \frac{R_n'(t_0)}{R_n(t_0)} = c \frac{\log n}{K_n}.$$

The restriction of the range of t guarantees that $\int_{-K_n}^{K_n} |x|^3 e^{tx} F(dx) < EH(|X_1|) < \infty$, and thus the conditions of Lemma A are satisfied.

A little calculation shows that $|\lambda_n(x)| < c_1$, if $|x| \leqq c \frac{\log n}{n}$. Thus relation (a) of Lemma 7 can be rewritten as

$$1 - F_k^{(n)}(kx) = [1 - \phi(k^{\frac{1}{2}} x)] \exp O(kx^3 + x + k^{-\frac{1}{2}})$$

$$\text{if } 0 \leqq x \leqq c \frac{\log n}{K_n} \quad c_0 \log n < k \leqq n. \tag{5.1}$$

A similar expansion for conditional distributions will be proved in Lemma 8.

Lemma 8. *Let $S_1, \ldots, S_n$ be the same as in the previous lemma. Let m be an even integer $C \log n < m \leqq n$ and define*

$$F_m^{(n)}(x|y) = P(2S_{m/2} - S_m < x | S_m = y).$$

The following asymptotic expansion is valid:

$$1 - F_m^{(n)}(mx|my) = [1 - \phi(m^{\frac{1}{2}} x)] \exp O(mx^3 + mx^2 |y| + |y| + m^{-\frac{1}{2}}),$$

$$F_m^{(n)}(-mx|my) = \phi(-m^{\frac{1}{2}} x) \exp O(mx^3 + mx^2 |y| + |y| + m^{-\frac{1}{2}})$$

in the range

$$0 \leqq x \leqq c \frac{\log n}{K_n}, \qquad |y| < \frac{\log n}{K_n}.$$

$O(\cdot)$ is uniform in x, y, m and n.

Proof. Put $u_n = c' \frac{\log n}{K_n}$. Then the estimation

$$I_1 = F_m^{(n)}(mu_n|my) - F_m^{(n)}(mx|my) = \int_x^{u_n} m f_m^{(n)}(mt|my)\, dt$$

$$= (1 - \phi(m^{\frac{1}{2}} x)) \exp O(mx^3 + mx^2 |y| + |y| + m^{-\frac{1}{2}}) \qquad 0 \leqq x \leqq u_n, \ |y| \leqq u_n$$

can be proved similarly to the estimation of I_1 in Lemma 1 of part I

$$\left(f_m^{(n)}(x|y) = \frac{d}{dx} F_m^{(n)}(x|y) \right).$$

One must, however, be a little cautious at the following step:

$$f_m^{(n)}(mx|my)=m^{-\frac{1}{2}}\varphi(m^{\frac{1}{2}}x)\exp[m\mu_n(x,y)+O(|x|+|y|+m^{-\frac{1}{2}})]$$
$$=m^{-\frac{1}{2}}\varphi(m^{\frac{1}{2}}x)\exp O(m|x|^3+mx^2|y|+|y|+m^{-\frac{1}{2}})$$

where

$$\mu_n(x,y)=\tfrac{1}{2}(x+y)^3\lambda_n(x+y)+\tfrac{1}{2}(y-x)^3\lambda_n(y-x)-y^3\lambda_n(y)$$
$$\text{if } |x|<u_n,\ |y|<u_n.$$

Here applying the Taylor expansion of $\lambda_n(y\pm x)$ around the point y up to two terms, one has to show that the $O(\cdot)$ of the above formula can be chosen independently of n.

To this aim it is sufficient to prove that

$$|\lambda_n(t)|<C_1,\qquad |\lambda_n'(t)|<C_1\frac{K_n}{\log n},\qquad |\lambda_n''(t)|<C_1\left(\frac{K_n}{\log n}\right)^2$$

in the interval $|t|<2u_n$. Expressing $\lambda_n(t)$, $\lambda_n'(t)$ and $\lambda_n''(t)$ by $R_n(t)$ and its derivatives, this statement reduces to the following ($R_n^{(i)}$ denotes the i-th derivative of R_n)

$$|R_n^{(i)}(t)|<C_2,\qquad i=1,2,3,$$
$$|R_n^{IV}(t)|<C_2\frac{K_n}{\log n},\qquad |R_n^{V}(t)|<C_2\left(\frac{K_n}{\log n}\right)^2\qquad \text{if } |t|<2u_n.$$

To check these inequalities note that

$$|x|^3\log H(|x|)\,e^{tx}<H(|x|)\qquad \text{if } |t|<c'\frac{\log n}{K_n},\quad |x|<K_n$$

and

$$\int_{-K_n}^{K_n}x^4e^{tx}F(dx)=\int_{-K_n}^{K_n}\frac{|x|}{\log H(|x|)}|x|^3\log H(|x|)\,e^{tx}F(dx)$$
$$\leqq\frac{K_n}{\log H(K_n)}\int_{-K_n}^{K_n}H(|x|)\,F(dx)\leqq C\frac{K_n}{\log n}.$$

The other inequalities can be proved similarly. On the other hand

$$I_2=1-F_m^{(n)}(mu_n|my)=\int_{u_n}^{\infty}f_m^{(n)}(ms|my)\,ds,$$

$$f_m^{(n)}(ms|my)\leqq C\exp(-dmt_0s),$$

where t_0 is defined by

$$\frac{R'(t_0)}{R(t_0)}=\frac{c'}{2}\,\frac{\log n}{K_n}.$$

Thus choosing the appearing constants appropriately, I_2 becomes negligible, compared to $1-F_m^{(n)}(mx|my)$.

The Proof of Theorem 3. Using Theorem 1B and the reasoning in proving Theorem 1, one may suppose that F has an absolutely continuous component.

Given a sequence $X_1, X_2, \ldots$ of i.i.d.r.v.'s with d.f. F define the sequences $\tilde{X}_n$, $\tilde{S}_n$, $\bar{X}_n$, $\bar{S}_n$ as in Lemma 6. Let us remark that it is sufficient to prove formula (1.3) substituting S_n by $\bar{S}_n$. Indeed, because of Lemma 6

$$\frac{\bar{S}_n - \tilde{S}_n}{K_n} \to 0 \quad \text{a.s.},$$

and defining the S_n-s in such a way that their truncations be the above defined $\tilde{S}_n$, we have $|S_n(\omega) - \tilde{S}_n(\omega)| \leqq K(\omega)$.

Let $\eta_1, \eta_2, \ldots$ be a sequence of i.i.d.r.v.'s with standard normal distribution. Let the η-s be independent of the $\bar{X}_i$-s, too.

Define

$$X_n' = \bar{X}_n + 4^{-j} \eta_n \quad \text{if } 2^j \leqq n < 2^{j+1},$$

and

$$S_n' = \sum_{i=1}^{n} X_i', \quad n = 1, 2, \ldots$$

Now it is sufficient to construct the variables S_n' and T_n in such a way that (1.3) hold.

Given the sequence $T_1, T_2, \ldots$ we construct the sequence $S_1', S_2', \ldots$ the same way as in Theorem 1A. Naturally the definition of the F-s will be substituted by the distribution and conditional distribution functions of the just defined S_n'-s.

Even the proof of Theorem 1A applies with slight modifications.

The following relations imply the desired result:

(i) $\limsup \dfrac{|S_{2^n}' - T_{2^n}|}{K_{2^n}} \leqq C.$

(ii) There exist appropriate constants $\alpha > 0$, $\beta > 0$ such that for any n, m, k satisfying the relations $k = 2^r$,

$$\frac{\alpha}{2} \frac{K_{2^n}^2}{n} < k \leqq \alpha \frac{K_{2^n}^2}{n}, \qquad 2^n \leqq m < m + k < 2^{n+1},$$

the following inequalities hold

$$P\Big(\sup_{m \leqq j < m+k} |S_j' - S_m'| > 2\beta K_{2^n}\Big) < \exp(-n),$$

$$P\Big(\sup_{m \leqq j < m+k} |T_j - T_m| > 2\beta K_{2^n}\Big) < \exp(-n).$$

(iii) Putting

$$\tilde{T}_j = \frac{2^{n+1} - j}{2^n} T_{2^n} + \frac{j - 2^n}{2^n} T_{2^{n+1}},$$

$$\tilde{S}_j = \frac{2^{n+1} - j}{2^n} S_{2^n}' + \frac{j - 2^n}{2^n} S_{2^{n+1}}', \qquad 2^n \leqq j \leqq 2^{n+1},$$

$$Z_j = |(S_j' - \tilde{S}_j) - (T_j - \tilde{T}_j)|,$$

we have

$$P(Z_s > \beta \cdot K_{2^n}) < \exp(-n)$$

if s, $2^n \leqq s \leqq 2^{n+1}$, is of the form $s = 2^n + ak$, a is integer, and k is the same as in (ii).

(i) guarantees that the S_{2^n}-s and T_{2^n}-s, (iii) guarantees that the S_{2^n+ck}-s and T_{2^n+ck}-s and finally (ii) guarantees that all S_n and T_n are sufficiently near each other.

In the proof of (i), (ii) and (iii) we use the notations of Theorem 1A and of Theorem 1 in part I.

Proof of (i).

Define

$$S_{2^n}^{(1)} = \begin{cases} S'_{2^n} - S'_{2^{n-1}} & \text{if } |S'_{2^n} - S'_{2^{n-1}}| < y_n \cdot 2^n \\ 0 & \text{otherwise} \end{cases}$$

where $y_n = c\dfrac{n}{K_{2^n}}$, and

$$S_{2^n}^{(2)} = (S'_{2^n} - S'_{2^{n-1}}) - S_{2^n}^{(1)}.$$

Using formula 3.1 one can prove that

$$P(S_{2^n}^{(2)} \neq 0) < \exp(-2^n c y_n^2) < q^n$$

with some $q < 1$, and

$$|S_{2^n}^{(1)} - (T_{2^n} - T_{2^{n-1}})| < C_1 \frac{(S_{2^n}^{(1)})^2}{2^n} + C_2$$

(this is the analogue of Lemma 1 in part I). Further, $S_{2^n}^{(1)}$ has a moment-generating function for $|t| < t_0$. These relations imply (i).

Proof of (ii).

$$P\Big(\sup_{m \leqq j < m+k} |S'_j - S'_m| > \beta K_{2^n}\Big) \leqq 2P\left(|S_{k+m} - S_m| > \frac{\beta}{2} K_{2^n}\right) < \exp(-n)$$

by (3.1) if α, β are chosen appropriately.

The other relation can be proved similarly.

Proof of (iii).

Similarly to the proof of Theorem 1 and Lemma 1 in part I one obtains

$$Z_j \leqq \sum_{i=r+1}^{n} |\tilde{U}_{i,k(i)} - \tilde{V}_{i,k(i)}|$$

and

$$|\tilde{U}_{i,l} - \tilde{V}_{i,l}| < C_1 \cdot 2^{-i}(\tilde{U}_{i,l}^2) + C_2$$

if

$$i > r, \quad |U_{i,l}| < c \cdot 2^i \frac{n}{K_{2^n}}, \quad |\tilde{U}_{i,l}| < c \cdot 2^i \frac{n}{K_{2^n}}.$$

At the proof of the last step one needs Lemma 8.

The estimations

$$P\left(|U_{i,l}|\geqq c\cdot 2^i\frac{n}{K_{2^n}}\right)<\exp(-n),\qquad P\left(|\tilde{U}_{i,l}|\geqq c\cdot 2^i\frac{n}{K_{2^n}}\right)<\exp(-n)$$

hold true because of (5.1) if α and thus r is chosen large enough.

The remaining part reduces, just as in Theorem 1 of part I, to the estimation

$$P\left(\sum_{j=r}^{n}\tau_j>cK_n\right)<\exp(-n),$$

where

$$\tau_j=\begin{cases}(S'_{2^n+2^j}-S'_{2^n+2^{j-1}})^2 & \text{if } |S'_{2^n+2^j}-S'_{2^n+2^{j-1}}|<c\cdot 2^j\dfrac{n}{K_{2^n}}\\ 0 & \text{otherwise}\end{cases}$$

This estimation is valid since the τ-s have a finite moment-generating function, and thus the proof is finished.

Let us remark that Theorem 2 holds true also in the case when $E|X_1|^r<\infty$ for $2<r\leqq 3$, but the proof is different (though easier). We turn to this question in a subsequent paper.

The Proof of Theorem 4. The construction and the proof of Theorem 4 is similar to that of Theorem 3. The only difference is that now we truncate the r.v.'s $X_1,\ldots,X_n$ at the level ax. Thus Lemma 7 and 8 hold in the range

$$|x|<c\frac{\log H(ax)}{ax}.$$

New problems do not arise, however, since we have to investigate S_k-S_j only for $k-j>\dfrac{ax^2}{c\log H(ax)}$. This construction approximates the X_k-s by normal variables, with expectation $m=\int_{-ax}^{ax}tF(dt)$ and variance $\sigma^2=\int_{-ax}^{ax}t^2F(dt)-m^2$.

A little calculation shows that

$$m=O\left(\frac{x}{H(ax)}\right),\qquad \sigma^2=1-O\left(\frac{x^2}{H(ax)}\right)$$

and therefore the $N(m,\sigma)$ variables are near enough to their standardization.

References

1. Komlós, J., Major, P., Tusnády, G.: Weak convergence and embedding. Trans. Coll. Limit Theorems. 149–165, Keszthely 1974
2. Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent RV's and the sample DF.I. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 111–131 (1975)
3. Petrov, V.V.: On sums of independent random variables, p. 270 and 283, (in Russian). Moscow: Nauka 1972
4. Petrov, V.V.: A local theorem for densities of sums of independent random variables (in Russian). Teor. Verojatnost. i Primenen. **3**, 349–357 (1956)

Received September 29, 1974