ON THE UNIVERSAL A.S. CENTRAL LIMIT THEOREM

SIEGFRIED HÖRMANN

Institute of Statistics, Graz University of Technology, Steyrergasse 17/IV, 8010 Graz, Austria

email: shoermann@tugraz.at

Abstract. Let (X_k) be a sequence of independent r.v.'s such that for some measurable functions $g_k : \mathbb{R}^k \to \mathbb{R}$ a weak limit theorem of the form

$$g_k(X_1, \dots, X_k) \xrightarrow{\mathcal{L}} G$$

holds with some distribution function G. By a general result of Berkes and Csáki ("universal ASCLT"), under mild technical conditions the strong analogue

$$\frac{1}{D_N} \sum_{k=1}^{N} d_k I\{g_k(X_1, \dots, X_k) \le x\} \to G(x) \quad \text{a.s.}$$

is also valid, where (d_k) is a logarithmic weight sequence and $D_N = \sum_{k=1}^{N} d_k$. In this paper we extend the last result for a very large class of weight sequences (d_k), leading to considerably sharper results. We show that logarithmic weights, used traditionally in a.s. central limit theory, are far from optimal and the theory remains valid with averaging procedures much closer to, in some cases even identical with, ordinary averages.

Keywords: Almost sure limit theory, summation methods.

2000 Mathematics Subject Classification: 60F15, 60F05.

1. Introduction

Let $X_1, X_2, \dots$ be i.i.d. random variables with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$ and let $S_k = X_1 + \dots + X_k$. By the simplest version of the almost sure central limit theorem (ASCLT) we have for any fixed $x \in \mathbb{R}$

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k=1}^N \frac{1}{k} I\left\{\frac{S_k}{\sqrt{k}} \le x\right\} = \Phi(x) \quad \text{a.s.,} \tag{1}$$

where $\Phi(x)$ denotes the standard normal distribution function. Relation (1) was proved independently by Brosamler [10], Fisher [17] and Schatte [32] under different degrees of generality. Fisher [18] and Lacey and Philipp [23] were the first who showed (1) under assuming only finite variances. Clearly, (1) is a weighted strong law of large numbers for the strongly dependent r.v.'s $I\{S_k/\sqrt{k} \le x\}$; note that the ordinary SLLN with Cesàro instead of logarithmic averages does not hold even for $x = 0$, as is seen immediately from the arc sine law. In the past decade several papers investigated logarithmic limit theorems of the type (1), and relation (1) has been extended in various directions. Ibragimov and Lifshits [22] proved that if (X_k) is a sequence of independent r.v.'s satisfying

$$\frac{S_k}{a_k} - b_k \xrightarrow{\mathcal{L}} H$$

for some distribution function H, then under mild moment conditions on S_k we have for any bounded continuous function $f : \mathbb{R} \to \mathbb{R}$

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^N d_k f\left(\frac{S_k}{a_k} - b_k\right) = \int_{-\infty}^{\infty} f(x)\, dH(x) \quad \text{a.s.,} \tag{2}$$

where

$$d_k = \log(a_{k+1}/a_k), \qquad D_N = \log a_N. \tag{3}$$

In the case where the sequence (X_k) satisfies the Lindeberg condition this relation has been first proved by Atlagh [2] with $a_k = \text{Var} S_k$ and H the standard normal distribution function. If (a_k) is nondecreasing, the last relation can be written equivalently as

$$\lim_{N\to\infty} \frac{1}{\log a_N} \int_1^{\log a_N} \frac{1}{t} f\left(\frac{S_t}{a_t} - b_t\right) dt = \int_{-\infty}^{\infty} f(x)\, dH(x) \quad \text{a.s.,}$$

where S_t, a_t, b_t are the continuous time extensions of S_k, a_k, b_k, taking a constant value on the intervals $[a_k, a_{k+1})$. Thus while the weights d_k in (3) depend formally on the norming sequence (a_k), the employed averaging procedure is simply logarithmic averaging on the natural time scale determined by the norming sequence (a_k). For various further extensions and related "logarithmic" limit theorems see the survey papers Berkes [5] and Atlagh and Weber [4].

Most results in the ASCLT literature concern partial sum behavior, i.e. deal with sums of the type (2), but there exist a few extensions for nonlinear functionals of independent r.v.'s as well. Marcus and Rosen [24], Csáki and Földes [13] and Horváth and Khoshnevisan [21] obtained analogues of (1) for local times and Fahrner and Stadtmüller [16] and Cheng et al. [12] proved a similar result for extreme order statistics. In [12] and [16] it is shown that if $X_1, X_2, \dots$ are i.i.d. r.v.'s such that for some numerical sequences (a_k), (b_k) we have

$$(\max_{i\le k} X_i - a_k)/b_k \xrightarrow{\mathcal{L}} G$$

with a nondegenerate distribution G, then

(4)
$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k\le N} \frac{1}{k} I\left\{\frac{\max_{i\le k} X_i - a_k}{b_k} < x\right\} = G(x) \qquad \text{a.s. for any } x\in\mathbb{R}.$$

This result is generalized by Csáki and Gonchigdanzan [14] to stationary Gaussian sequences under certain dependence assumptions. For further examples of nonlinear extensions of (2), see Antonini and Weber [1] and Mörters [25], [26]. In view of these results the question arises which distributional limit theorems have similar a.s. versions and Berkes and Csáki [6] proved the surprising result that *every* weak limit theorem for independent random variables, subject to minor technical assumptions, has an a.s. "logarithmic" version. Specifically, they proved the following result:

Theorem A. *Let $X_1, X_2, \dots$ be independent random variables satisfying the weak limit theorem*

(5)
$$g_l(X_1, X_2, \dots, X_l) \xrightarrow{\mathcal{L}} G,$$

where $g_l : \mathbb{R}^l \to \mathbb{R}$ $(l = 1, 2 \dots)$ are measurable functions and G is a distribution function. Assume that for each $1 \le k < l$ there exists a measurable function $g_{k,l} : \mathbb{R}^{l-k} \to \mathbb{R}$ such that

(6)
$$\mathbf{E}\left(|g_l(X_1, \dots, X_l) - g_{k,l}(X_{k+1}, \dots, X_l)| \wedge 1\right) \le A(c_k/c_l)$$

with a constant $A > 0$ and a positive, nondecreasing sequence (c_k) satisfying $c_k \to \infty$, $c_{k+1}/c_k = O(1)$. Then we have

(7)
$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k\le N} d_k I\left\{f_k(X_1, \dots, X_k) < x\right\} = G(x) \qquad \text{a.s. for any } x \in C_G,$$

where

$$d_k = \log(c_{k+1}/c_k), \quad D_N = \log c_N$$

and C_G *denotes the set of continuity points of* G.

The simplest form (1) of the ASCLT is obtained by letting

$$g_l = (X_1 + \ldots + X_l)/\sqrt{l}, \qquad g_{k,l} = (X_{k+1} + \ldots + X_l)/\sqrt{l},$$

where (X_n) is an i.i.d. sequence with $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$. It is easy to see that in this case (6) is satisfied with $c_k = k^\alpha$ for some $\alpha > 0$ and thus (7) holds with $d_k \sim \text{const}\cdot 1/k$. Similarly, the limit theorem (4) is obtained with

$$g_l(X_1, \ldots, X_l) = (\max_{1\le i\le l} X_i - a_l)/b_l, \quad g_{k,l}(X_{k+1}, \ldots, X_l) = (\max_{k+1\le i\le l} X_i - a_l)/b_l$$

and in this case (6) is satisfied with $c_k = k$, leading again to $d_k \sim \text{const}\cdot 1/k$. In both cases (6) means that changing the first $o(l)$ terms of the sequence $X_1, \ldots, X_l$ has very little effect on the normed partial sum $(X_1 + \cdots + X_l)/\sqrt{l}$ respectively the extremal statistics

$$(\max_{1\le i\le l} X_i - a_l)/b_l.$$

An example for a limit theorem with a different c_k is the Darling-Erdős theorem stating

$$a_l\left(\max_{i\le l}\frac{S_i}{\sqrt{i}} - b_l\right) \xrightarrow{\mathcal{L}} e^{-e^{-x}} \tag{8}$$

for suitable (a_l) and (b_l), where S_l are partial sums of independent r.v.'s with mean 0, variance 1 and uniformly bounded $(2+\delta)$-th moments. (See Darling and Erdős [15] for $\delta = 1$ and Shorack [33] for $\delta > 0$.) Here (6) holds with $c_k = \log k$, and (6) means that changing the first $l^{o(1)}$ terms of the sequence $X_1, \ldots, X_l$ will change the left hand side of (8) only unessentially. Note that in this case the initial segment of $X_1, \ldots, X_l$ not influencing the value of $g_l(X_1, \ldots, X_l)$ is much shorter: the dependence of g_l on its initial variables became more sensitive. As a consequence, we get a different weight sequence in (7): instead of $d_k = 1/k$ we get now $d_k = 1/(k\log k)$, $D_N = \log\log N$, i.e. in the a.s. version of the Darling-Erdős limit theorem we have loglog averages. We thus see that the more sensitively the functional $g_l(x_1, \ldots, x_l)$ depends on its initial variables, the smaller weight sequence (d_k) in (7) is obtained. Another version of Theorem A with logarithmic weights can be derived from distributional limit theorems for central order statistics (cf. Stadtmüller [34], Peng and Qi [28]). Here the verification of (6) causes substantial difficulties. A simplified approach for this problem can be found in Hörmann [20] where it also shown that Césaro averaging fails. For a detailed discussion of Theorem A and several examples as well as more refined versions of the latter result we refer to Berkes and Csáki [6].

The characteristic feature of a.s. central limit theory is the use of logarithmic averages; with ordinary averages even the simplest relation (1) becomes false. Clearly, the dependence of the sequence $S_k/\sqrt{k}$ is too strong for the indicators $I\{S_k/\sqrt{k} \le x\}$ to satisfy the ordinary law of large numbers. However, logarithmic averaging is not the only one providing a.s. convergence to $\Phi(x)$ in (1). Peligrad and Révész [27] showed that letting

$$d_k = (\log k)^\alpha/k, \qquad D_N = \sum_{k=1}^N d_k, \qquad \alpha > -1 \tag{9}$$

we have

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^N d_k I\left\{\frac{S_k}{\sqrt{k}} \le x\right\} = \Phi(x) \quad \text{a.s.} \tag{10}$$

and Berkes and Csáki [6] showed that (10) holds also if

$$d_k = \exp((\log k)^\alpha)/k, \qquad D_N = \sum_{k=1}^N d_k, \qquad 0 \le \alpha < 1/2. \tag{11}$$

For nonstandard weights in the ASCLT for i.i.d. r.v.'s see also Weber [35]; for not identically distributed r.v.'s see Rodzik and Rychlik [29], [30]. To compare these results with (1), we recall some well known facts from analysis. Given a positive sequence $\mathbf{D} = (d_1, d_2, \ldots)$ with $D_N = \sum_{k=1}^N d_k \to \infty$, we say that a sequence (x_k) is $\mathbf{D}$-summable to x if

$$\lim_{N\to\infty} D_N^{-1} \sum_{k=1}^N d_k x_k = x.$$

By a result of Hardy (see [11, p. 35]), if $\mathbf{D}$ and $\mathbf{D}^*$ are summation procedures with $D_N^* = \mathcal{O}(D_N)$, then under minor technical assumptions, the summation $\mathbf{D}^*$ is stronger than $\mathbf{D}$, i.e., if a sequence (x_k) is $\mathbf{D}$-summable to x, then it is also $\mathbf{D}^*$-summable to x. Moreover, by a result of Zygmund (see [11, p. 35]) if $D_N^\alpha \le D_N^* \le D_N^\beta$ $(N \ge N_0)$ for some $\alpha > 0, \beta > 0$, then $\mathbf{D}$ and $\mathbf{D}^*$ are equivalent, and if $D_N^* = \mathcal{O}(D_N^\varepsilon)$ for any $\varepsilon > 0$, then $\mathbf{D}^*$ is strictly stronger than $\mathbf{D}$. These results show that the larger the norming sequence (D_N) in (10) is, the stronger the relation (10) becomes and thus the strongest form of the ASCLT is given for the largest possible norming sequence (D_N). In view of Zygmund's theorem, the summation defined by the sequence in (9) is actually equivalent to logarithmic averaging and the summation defined by

$$d_k = k^{-\alpha}, \qquad D_N = \sum_{k=1}^N d_k, \qquad \alpha < 1$$

is equivalent to ordinary averaging. On the other hand, the summation procedures defined by (11) lie strictly between log and Cesàro averaging and are pairwise non-equivalent for different α's. Note also that for $\alpha = 0$, resp. $\alpha = 1$ (11) reduces to logarithmic, resp. ordinary averaging and thus the result of Berkes and Csáki shows that while the ASCLT fails with ordinary averaging, we can still go at least "halfway" from logarithmic to ordinary averaging. The same phenomenon occurs in the general nonlinear Theorem A, as one can easily see. These remarks show that logarithmic averaging, used almost exclusively in a.s. central limit theory, is only one of many suitable averaging processes and the optimal, "true" version of the theorems is obtained with the largest possible weight sequence (d_k). In the case of the ordinary central limit theorem for i.i.d. random variables, this optimal (d_k) was determined in Hörmann [19] up to an unknown constant. Specifically, we proved that if (d_k) satisfies the condition

$$d_k = O\left(\frac{D_k}{k(\log\log k)^\alpha}\right) \tag{12}$$

with $\alpha > 3$, then not only the SLLN (10), but the corresponding central limit theorem and LIL also hold; on the other hand, this becomes false if we assume (12) with $\alpha < 1$. This shows that the critical sequence for the ASCLT (10) is

$$D_N = \exp(\log N/(\log\log N)^\alpha) \tag{13}$$

for some (unknown) $1 \le \alpha \le 3$. For $\alpha = 0$ this gives ordinary averaging and thus, in some sense, the critical sequence (13) is much closer to ordinary than log averaging. The purpose of the present paper is to study the analogous problem in the case of the general weak limit theorem (5). Due to the general character of (5), the large deviation and strong approximation techniques used in [19] do not seem to be applicable in this case, preventing us from getting the same complete characterization as in the case of the CLT. However, we will show that a slightly stronger version of the Kolmogorov condition (12) remains sufficient even for the most general limit theorem (5). Specifically, we shall prove the following result:

Theorem 1. *Let $X_1, X_2, \dots$ be independent r.v.'s such that for some measurable functions $g_l : \mathbb{R}^l \to \mathbb{R}$ the weak limit theorem*

$$g_l(X_1, \dots, X_l) \xrightarrow{\mathcal{L}} G \tag{14}$$

holds with some distribution function G. Assume further that for some functions $g_{k,l} : \mathbb{R}^{l-k} \to \mathbb{R}$, $1 \le k < l$, we have

$$\mathbf{E}\left(|g_l(X_1, \dots, X_l) - g_{k,l}(X_{k+1}, \dots, X_l)| \wedge 1\right) \le A(c_k/c_l)^\beta \tag{15}$$

for some $A,\ \beta > 0$ *and some positive nondecreasing sequence* (c_k) *with* $c_k \to \infty$*. Finally set* $d_k^* = \log(c_{k+1}/c_k)$ *and assume that the Kolmogorov type condition*

$$d_k = \mathcal{O}\left(d_k^* \frac{D_k}{(\log D_k)^\rho} \right) \tag{16}$$

is satisfied for some $\rho > 0$ *and in addition*

$$d_k \gg d_k^*, \qquad d_k/(d_k^* c_k^\beta) \quad \textit{is nonincreasing} \tag{17}$$

and

$$d_k D_k / (\log D_k)^\rho = \mathcal{O}(1) \tag{18}$$

hold. Then if f *is a bounded Lipschitz 1 function on the real line or is the indicator function of a Borel set* A *with* $G(\partial A) = 0$*, we have*

$$\lim_{N\to\infty} D_N^{-1} \sum_{k=1}^{N} d_k f\left(g_k(X_1, \ldots, X_k)\right) = \int_{-\infty}^{\infty} f(x)\, dG(x) \quad \text{a.s.} \tag{19}$$

In view of the result of Hardy quoted above, the validity of relation (19) automatically extends to smaller weight sequences (d_k), and thus we are interested in finding large weight sequences such that (19) is valid. Since (19) holds for $d_k = d_k^*$, the first assumption $d_k \gg d_k^*$ in (17) is a natural one. The second condition of (17) and relation (18) clearly limit the speed of growth of D_k from above. In view of relation (33) below, we have $\sum_{k=1}^n d_k^* c_k^\beta \sim \text{const} \cdot c_k^\beta$, and thus the second condition in (17) implies $D_k \ll c_k^\beta$. On the time scale determined by the c_k, the weights d_k^* give logarithmic averaging, while $D_k = c_k^\beta$ is equivalent, by the theorem of Zygmund, with (rescaled) ordinary Cesàro averaging belonging to $D_k = c_k$. Thus the second relation of (17) limits D_k above by Cesàro averaging and the same holds with relation (18), which implies that the increments of $D_k^2 (\log D_k)^{-\rho}$ are bounded and thus $D_k \ll k^{1/2} (\log k)^{\rho/2}$. Restricting the summation procedures in (19) above by Cesàro averaging is quite natural, since ordinary averaging is usually too large even in the simplest versions of the ASCLT. However, Example B in the next section will show that in certain non-i.i.d. situations one can actually use even the Cesàro weights in the general strong limit theorem (19).

The crucial condition in Theorem 1 is (16). It tells us how far we can move from the natural logarithmic weights (d_k^*) towards a larger weight sequence (d_k). As we will see, in typical cases (16) permits choosing (d_k) much closer to the corresponding Cesàro averaging than to logarithmic weights. Several examples will be given in Section 2.

The following theorem covers the case when the functionals g_k depend not on an independent sequence (X_k), but on a more general stochastic process $\{X_t, t \geq 0\}$ with independent increments.

Theorem 2. *Let $\{X_t, t \geq 0\}$ be a stochastic process with independent increments and assume $X_0 = 0$. Let η_l, $l = 1, 2, \ldots$ be r.v.'s such that η_l is measurable with respect to $\sigma\{X(t), t \leq l\}$ and assume that the weak limit theorem*

$$\eta_l \xrightarrow{\mathcal{L}} G$$

holds with some distribution function G. Assume finally that for every $1 \leq k < l$ there exists a $\sigma\{X(t) - X(s), k \leq s \leq t \leq l\}$-measurable random variable $\eta_{k,l}$ such that

$$\mathbf{E}\left(|\eta_l - \eta_{k,l}| \wedge 1\right) \leq A(c_k/c_l)^{\beta} \quad (1 \leq k < l), \tag{20}$$

with A, $\beta > 0$ and some nondecreasing sequence (c_k) with $c_k \to \infty$. If (16)-(18) are satisfied with $d_k^* = \log(c_{k+1}/c_k)$, then for any bounded Lipschitz 1 function f on the real line or for the indicator function f of any Borel set $A \subset \mathbb{R}$ with $G(\partial A) = 0$, we have

$$\lim_{N\to\infty} D_N^{-1} \sum_{k=1}^{N} d_k f(\eta_k) = \int_{-\infty}^{\infty} f(x)\, dG(x) \quad \text{a.s.} \tag{21}$$

2. Examples

In this section we will give some applications of Theorems 1 and 2. In each case we will specify the natural weights d_k^* and the corresponding best weight sequence our theorems provide. In what follows, f denotes a bounded Lipschitz 1 function or the indicator function of a Borel set $A \subset \mathbb{R}$ with $G(\partial A) = 0$, where G is some distribution function occurring as a weak limit of the form (14) in the given examples.

A. *Partial sums of i.i.d. r.v.'s.*

Let $X_1, X_2, \ldots$ be i.i.d. r.v.'s and let S_l denote the l-th partial sum. Assume that there exist numerical sequences (a_l) and (b_l) such that a weak limit theorem of the form

$$g_l(X_1, \ldots, X_l) := \frac{S_l}{a_l} - b_l \xrightarrow{\mathcal{L}} G$$

holds with some (possibly degenerate) distribution function G. Assume further that $\sup_{l\geq 1} \mathbf{E}|S_l/a_l - b_l|^{\nu} < \infty$ for some $\nu > 0$. We choose

$$g_{k,l}(X_{k+1}, \ldots, X_l) = \frac{S_l - S_k}{a_l} - b_l \quad (1 \leq k \leq l).$$

A standard argument in a.s. central limit theory shows that there exist positive constants C and β such that

$$\mathbf{E}(|g_l(X_1,\ldots,X_l) - g_{k,l}(X_{k+1},\ldots,X_l)| \wedge 1) \le C(a_k/a_l)^\beta.$$

For example, if $\mathbf{E}X_1 = 0$ and $\mathbf{E}X_1^2 = 1$, then $a_l = \sqrt{l}$ and Theorem 1 applies with $d_k^* = 1/k$ and with G denoting the standard normal distribution function. If X_k are i.i.d. r.v.'s belonging to the domain of attraction of a stable distribution G, then (a_l) is regularly varying with exponent $1/\alpha$ for some $0 < \alpha \le 2$ and from the representation theorem for regularly varying functions (cf. Bingham et al. [9, p. 12]) we obtain easily that $a_k/a_l \le C'(k/l)^{\beta'}$. Hence in this case the natural weights d_k^* are again the logarithmic ones. On the other hand, it is easily checked that for

$$D_N = \exp((\log N)^\alpha), \quad d_k = D_k - D_{k-1}, \quad 0 < \alpha < 1, \tag{22}$$

conditions (16)-(18) are satisfied, and thus Theorem 1 yields the stronger result

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^N d_k f\left(\frac{S_k}{a_k} - b_k\right) = \int_{-\infty}^{\infty} f(x)\, dG(x) \quad \text{a.s.}$$

B. *Sums of not identically distributed r.v.'s.*

Let X_k be independent r.v.'s with $\mathbf{E}X_k = 0$ and $\mathbf{E}X_k^2 = s_{k+1}^2 - s_k^2$, where $s_k^2 = e^{(\log k)^{1+\varepsilon}}$, $\varepsilon > 0$, and let S_k denote the k-th partial sum. Assume that the sequence (X_k) satisfies the Lindeberg condition and let g_l and $g_{k,l}$ be the same as in the last example with $a_l = s_l$ and $b_l = 0$. Then (15) is satisfied with $c_k = s_k$ and thus we get the natural weights $d_k^* = (\log k)^\varepsilon k^{-1}$. From a summability point of view, the summation method defined by (d_k^*) is equivalent to log summation. However, Theorem 1 shows the surprising fact that in this case, even Cesàro means work in (19). Indeed, define $d_k = 1/\sqrt{k}$. Again conditions (16)-(18) are easily checked. By Zygmund's theorem the summation defined by $d_k = 1/\sqrt{k}$ is equivalent to Cesàro summation. Hence in this example we get

$$\frac{1}{N} \sum_{k=1}^N f\left(\frac{S_k}{s_k}\right) \to \int_{-\infty}^{\infty} f(x)\, d\Phi(x) \quad \text{a.s.} \tag{23}$$

Clearly, the faster the sequence (s_k) grows, the more influence X_k has on the partial sum S_k and the smaller is the dependence between the $f(S_k/s_k)$. For example if $s_k^2 = \log k$ then the dependence between the $f(S_k/s_k)$ is much stronger than in the standard case $s_k^2 = k$ and Example 1 in Berkes and Dehling [7, p. 1649] shows that the ASCLT fails with $d_k = k^{-1}$. In this case the natural weights are $d_k^* = 1/(k \log k)$ and the norming sequence is $D_N^* = \log\log N$. But an application of Theorem 1 shows that the ASCLT actually holds in this example with the larger sequence $D_N = \exp((\log\log N)^\alpha)$ for any $0 < \alpha < 1$.

C. *Subsequences.*

As an immediate consequence of the first example in **B**, we get an almost sure central limit theorem for i.i.d. r.v.'s along subsequences, using Cesàro summation. Let X_k be i.i.d. r.v.'s with $\mathbf{E}X_1 = 0$ and $\mathbf{E}X_1^2 = 1$ and set $n_k = [s_k^2]$, where $s_k^2 = \exp((\log k)^{1+\varepsilon})$ with $\varepsilon > 0$. Define $Y_k = S_{n_k} - S_{n_{k-1}}$. Clearly Y_k are independent, $\mathbf{E}Y_k = 0$, $\mathbf{E}Y_k^2 = n_k - n_{k-1}$ and by the central limit theorem $S_{n_k}/\sqrt{n_k} \xrightarrow{\mathcal{L}} N(0,1)$. Hence we have by (23)

$$\frac{1}{N}\sum_{k=1}^{N} f\left(\frac{S_{n_k}}{\sqrt{n_k}}\right) \to \int_{-\infty}^{\infty} f(x)\, d\Phi(x) \quad \text{a.s.} \tag{24}$$

Schatte [32] and Atlagh and Weber [3] showed a similar result for $n_k = [c^k]$ with $c > 1$. Of course, the faster n_k grows, the less is the dependence between the partial sums S_{n_k} and the weaker summation methods apply. Rychlik and Szuster [31] showed that (24) also holds for $n_k = [c^{k^\alpha}]$ with $c > 1$ and $\alpha > 0$. The last result shows that we can weaken the growth rates used in [32], [3] and [31] of (n_k) substantially. Note also that the growth condition for the subsequence (n_k) is sharp in some sense. Choosing $\varepsilon = 0$ gives $n_k = k$ and then (24) fails (cf. [32, Theorem 1]).

D. *Sample extremes.*

Let $X_1, X_2, \dots$ be i.i.d. r.v.'s. Assume further that there are numerical sequences (a_l) and (b_l) such that for some distribution function G a weak limit theorem of the form

$$g_l(X_1, \dots, X_l) := \frac{\max_{i \le l} X_i - a_l}{b_l} \xrightarrow{\mathcal{L}} G$$

is valid. Then setting

$$g_{k,l}(x_{k+1}, \dots, x_l) = \frac{\max_{k < i \le l} x_i - a_l}{b_l}$$

we get (cf. [6, p. 122])

$$\mathbf{E}(|g_l(X_1, \dots, X_l) - g_{k,l}(X_{k+1}, \dots, X_l)| \wedge 1) \le \frac{k}{l}.$$

Thus Theorem A yields

$$\frac{1}{D_N^*}\sum_{k=1}^N d_k^* f\left(\frac{\max_{i\le l} X_i - a_l}{b_l}\right) \to \int_{-\infty}^{\infty} f(x)\,dG(x) \quad \text{a.s.,} \tag{25}$$

where the natural weights d_k^* are the logarithmic ones. As observed in Fahrner and Stadtmüller [16] and in Cheng et. al. [12], if we replace logarithmic means in (25) by the ordinary Cesàro means, the result fails. However, Theorem 1 shows that (25) is still true when the summation method (D_N^*) is replaced by (D_N) defined in (22).

E. *The Darling-Erdős theorem.*

Let $\mathbb{W}$ be a Wiener process and

$$\eta_l = a_l\left(\sup_{1\le t\le l}\frac{\mathbb{W}(t)}{\sqrt{t}} - b_l\right)$$

where

$$a_l = (2\log\log l)^{1/2} \quad \text{and} \quad b_l = a_l + \frac{\log\log\log l - \log 4\pi}{2a_l} \quad (l \ge 3).$$

By the Darling-Erdős theorem (see [15]) we have $\eta_l \xrightarrow{\mathcal{L}} G$ where $G(x) = e^{-e^{-x}}$. Let $c_l = \exp(\sqrt{\log\log l})$, $A_l = \exp(\log l/\exp(\sqrt{\log\log l}))$ and

$$\eta_{k,l} = \begin{cases} a_l\left(\sup_{A_l^2\le t\le l}\frac{\mathbb{W}(t)-\mathbb{W}(k)}{\sqrt{t}} - b_l\right) & \text{if } k \le A_l \\ 0 & \text{otherwise.}\end{cases}$$

In Berkes and Csáki [6] it is shown that

$$\mathbf{E}(|\eta_l - \eta_{k,l}| \wedge 1) \le 4(c_k/c_l)^{1/2} \quad (k \le l). \tag{26}$$

Also, η_l is measurable with respect to $\sigma\{\mathbb{W}(t), t \le l\}$ and $\eta_{k,l}$ is measurable with respect to $\sigma\{\mathbb{W}(t') - \mathbb{W}(t) : k \le t \le t' \le l\}$. Thus Theorem 5 in Berkes and Csáki [6] implies

$$\lim_{N\to\infty}\frac{1}{D_N}\sum_{k=1}^N d_k f\left(a_k\left(\sup_{1\le t\le k}\frac{\mathbb{W}(k)}{\sqrt{k}} - b_k\right)\right) = \int_{-\infty}^{\infty} f(x)\,dG(x) \quad \text{a.s.,} \tag{27}$$

where

$$d_k = \frac{1}{k\log k\sqrt{\log\log k}}, \qquad D_N = \sqrt{\log\log N}. \tag{28}$$

By Zygmund's theorem, the averaging procedure determined by the weight sequence in (28) is equivalent to loglog averaging, i.e. (27) holds also with

$$d_k = \frac{1}{k \log k}, \qquad D_N = \log\log N.$$

On the other hand, using relation (26) with $c_l = \exp(\sqrt{\log\log l})$, Theorem 2 implies (27) with the considerably larger weights

$$d_k \sim \frac{1}{k \log k} e^{(\log\log k)^\alpha}, \quad D_N \sim e^{(\log\log N)^\alpha} \quad 0 < \alpha < 1/2.$$

More generally, consider a nondecreasing, unbounded function $h : \mathbb{R}^+ \to \mathbb{R}^+$ with $1 \le h(x) \le x$ and set

$$\eta_k^{(h)} = a_k^{(h)} \left(\sup_{k/h(k) \le t \le k} \frac{\mathbb{W}(t)}{\sqrt{t}} - b_k^{(h)} \right)$$

where

$$a_k^{(h)} = (2 \log\log h(k))^{1/2} \quad \text{and} \quad b_k^{(h)} = a_k^{(h)} + \frac{\log\log\log h(k) - \log 4\pi}{2 a_k^{(h)}}$$

Then a slightly more general form of the Darling-Erdős theorem yields

$$\eta_k^{(h)} \xrightarrow{\mathcal{L}} G. \tag{29}$$

An a.s. version of the last theorem was obtained in Berkes and Weber [8] who showed that

$$\frac{1}{D_N} \sum_{k=1}^{N} d_k f(\eta_k^{(h)}) \to \int_{-\infty}^{\infty} f(x)\, dG(x) \quad a.s., \tag{30}$$

where the weights d_k depend on the function h: the slower h grows, the stronger summation method is required in (30). The dependence of the weights d_k on h is rather involved, so we consider a simple special case, e.g. $h(x) = e^{(\log x)^\alpha}$, $0 < \alpha < 1$. Letting

$$\eta_{k,l}^{(h)} = \begin{cases} a_l^{(h)} \left(\sup_{R_l \le t \le l} \frac{\mathbb{W}(t) - \mathbb{W}(k)}{\sqrt{t}} - b_l^{(h)} \right) & \text{if } k \le R_l \\ 0 & \text{otherwise,} \end{cases}$$

with $R_l = l/h(l)$, in Berkes and Weber [8] it is shown that the analogue of (26) for $\eta_l^{(h)}$, $\eta_{k,l}^{(h)}$ holds with $c_k = e^{2(\log k)^{1-\alpha}}$. We can now apply Theorem 2 and get (30) with the natural weights $d_k^* = \frac{1}{k(\log k)^\alpha}$. By Zygmund's theorem this summation is equivalent to log summation, which is exactly the result derived in [8]. But Theorem 2 shows that we can use also the summation defined by $D_N = e^{(\log N)^\gamma}$ with $0 < \gamma < 1 - \alpha$, giving a stronger result.

In our previous considerations we dealt with the Darling-Erdős theorem for the Wiener process, but with a simple invariance argument like in [6], [8], [33], we can extend the results for partial sums of independent random variables with mean 0, variance 1 and uniformly bounded $(2+\delta)$-th moments.

3. Proofs

In what follows, we will give the proof of Theorem 1; the proof of Theorem 2 is very similar. We first prove some preparatory lemmas. Let f be a Lipschitz 1 function with $|f| \le 1$ and put for $1 \le k < l$

$$\begin{aligned} \xi_l &:= f\left(g_l(X_1, \ldots, X_l)\right) - \mathbf{E} f\left(g_l(X_1, \ldots, X_l)\right) \\ \xi_{k,l} &:= f\left(g_{k,l}(X_{k+1}, \ldots, X_l)\right) - \mathbf{E} f\left(g_{k,l}(X_{k+1}, \ldots, X_l)\right), \end{aligned} \tag{31}$$

where (X_k) is a sequence of independent random variables and $g_l : \mathbb{R}^l \to \mathbb{R}$ and $g_{k,l} : \mathbb{R}^{l-k} \to \mathbb{R}$ are measurable functions. Here, and in the sequel, the constants c, α, etc. depend only on the sequences (X_k), (g_l), $(g_{k,l})$ and f. Constants like C_p, E_p will depend on the parameter p as well.

Lemma 1. *Define ξ_l and $\xi_{k,l}$ as in (31) and assume that (15) holds. If (d_k) is a numerical sequence satisfying (17), then we have for any $k \le m \le n$ and $p \in \mathbb{N}$*

$$\mathbf{E}\left|\sum_{l=m}^{n} d_l(\xi_l - \xi_{k,l})\right|^p \le E_p \left(\frac{c_k}{c_m}\right)^\beta \left(\sum_{l=m}^{n} d_l c_l^{-\beta} \left(\sum_{k=1}^{l} d_k c_k^\beta\right)\right)^{p/2},$$

where

$$E_p = K^p p^{p/2},$$

with some constant K.

Proof. We set $Q(l) = Q(k,l) = \xi_l - \xi_{k,l}$. Trivially $|Q(l)| \le 4$ and thus

$$\begin{aligned} \mathbf{E}|Q(l)|^p &\le 4^{p-1}\mathbf{E}|Q(l)| \le 2 \cdot 4^{p-1}\mathbf{E}|f(g_l(X_1, \ldots, X_l)) - f(g_{k,l}(X_{k+1}, \ldots, X_l))| \\ &\le \text{const} \cdot 4^p \mathbf{E}\left(|g_l(X_1, \ldots, X_l) - g_{k,l}(X_{k+1}, \ldots, X_l)| \wedge 1\right). \end{aligned}$$

By (15) we get thus

$$\mathbf{E}|Q(l)|^p \le K_0 \cdot 4^p (c_k/c_l)^\beta,$$

with some $K_0 > 0$. Using the Hölder and the Cauchy-Schwarz inequality we get, setting $d_l^* = \log(c_{l+1}/c_l)$,

$$\begin{aligned}
\mathbf{E}\left|\sum_{l=m}^{n} d_l(\xi_l - \xi_{k,l})\right|^p &\leq \sum_{l_1=m}^{n}\cdots\sum_{l_p=m}^{n} d_{l_1}\cdots d_{l_p}\left(\mathbf{E}|Q(l_1)|^p\cdots\mathbf{E}|Q(l_p)|^p\right)^{1/p} \\
&\leq K_0\cdot 4^p c_k^\beta \sum_{l_1=m}^{n}\cdots\sum_{l_p=m}^{n} d_{l_1}\cdots d_{l_p}\, c_{l_1}^{-\beta/p}\cdots c_{l_p}^{-\beta/p} \\
&= K_0\cdot 4^p c_k^\beta \left(\sum_{l=m}^{n} d_l c_l^{-\beta/p}\right)^p \\
&\leq K_0\cdot 4^p c_k^\beta \left(\sum_{l=m}^{n} d_l^2 (d_l^*)^{-1}\right)^{p/2}\left(\sum_{l=m}^{n} c_l^{-\frac{2\beta}{p}}\log(c_{l+1}/c_l)\right)^{p/2}.
\end{aligned}$$

Since (c_k) is nondecreasing, for any $\gamma > 0$ we have, setting $u_k = \log c_k$,

$$\sum_{k=1}^{n-1} e^{\gamma u_k}(u_{k+1}-u_k) \leq \int_{u_1}^{u_n} e^{\gamma x}\,dx \leq \sum_{k=1}^{n-1} e^{\gamma u_{k+1}}(u_{k+1}-u_k). \tag{32}$$

By (17) and (18) we have $d_k^* = \log(c_{k+1}/c_k) \to 0$, i.e. $c_{k+1}/c_k \to 1$ and thus (32) yields

$$\sum_{k=1}^{n} c_k^\gamma \log(c_{k+1}/c_k) \sim \frac{1}{\gamma} c_n^\gamma \quad (\gamma > 0,\, n \to \infty). \tag{33}$$

A similar argument yields for $\gamma < 0$

$$\sum_{k=n}^{\infty} c_k^\gamma \log(c_{k+1}/c_k) \sim \frac{1}{|\gamma|} c_n^\gamma \quad (\gamma < 0,\, n \to \infty). \tag{34}$$

Hence (17) and (33) imply that

$$\sum_{k=1}^{l} d_k c_k^\beta \gg \frac{d_l}{d_l^* c_l^\beta}\sum_{k=1}^{l} d_k^* c_k^{2\beta} \gg \frac{d_l}{d_l^*} c_l^\beta.$$

Thus there is some $K_1 > 0$ with

$$\sum_{l=m}^{n} d_l^2 (d_l^*)^{-1} \leq K_1 \sum_{l=m}^{n} d_l c_l^{-\beta}\left(\sum_{k=1}^{l} d_k c_k^\beta\right).$$

On the other hand, (34) implies that

$$\sum_{l=m}^{n} c_l^{-\frac{2\beta}{p}} \log(c_{l+1}/c_l) \leq K_2 \cdot \frac{p}{2\beta} \cdot c_m^{-\frac{2\beta}{p}},$$

with $K_2 > 0$. This completes the proof of Lemma 1. □

The crucial step of the proof of Theorem 1 is the following moment inequality:

Lemma 2. *Assume that $X_1, X_2, \dots$ are independent r.v.'s and assume that (15)-(18) hold. Then for every $p \in \mathbb{N}$ we have*

$$\mathbf{E}\left|\sum_{k=1}^{N} d_k \left\{f(g_k(X_1,\dots,X_k)) - \mathbf{E}f\left(g_k(X_1,\dots,X_k)\right)\right\}\right|^p$$

(35)
$$\le C_p \left(\sum_{1\le k\le l\le N} d_k d_l \left(\frac{c_k}{c_l}\right)^{\beta}\right)^{p/2},$$

where $C_p > 0$ is a constant.

Proof. Again we use notation (31) and set

$$V_{m,n} := \sum_{l=m}^{n} d_l c_l^{-\beta} \left(\sum_{k=1}^{l} d_k c_k^{\beta}\right) \qquad (1 \le m \le n).$$

Further we put

$$C_p = (4\gamma)^{p^2}. \tag{36}$$

We show that if the number γ is chosen large enough, then

$$\mathbf{E}\left|\sum_{k=m}^{n} d_k \xi_k\right|^p \le C_p (V_{m,n})^{p/2} \quad \text{for all} \;\; 1 \le m \le n. \tag{37}$$

Since $V_{1,N}$ equals the double sum on the right hand side of (35), this will prove Lemma 2. We use induction on p.

First observe that ξ_k and $\xi_{k,l}$ are independent for $1 \le k < l$. Therefore from $|\xi_k| \le 2$ and the Lipschitz 1 continuity of f we get

$$\begin{aligned} |\mathbf{E}\xi_k\xi_l| &= |\mathbf{E}\xi_k(\xi_l - \xi_{k,l})| \le 2\mathbf{E}|\xi_l - \xi_{k,l}| \\ &\le K_3 \mathbf{E}(|g_l(X_1,\dots,X_l) - g_{k,l}(X_{k+1},\dots,X_l)| \wedge 1), \end{aligned}$$

for some $K_3 > 0$. Together with (15) we get

$$\begin{aligned} \mathbf{E}\left(\sum_{k=m}^{n} d_k \xi_k\right)^2 &\le 2 \sum_{m\le k\le l\le n} d_k d_l |\mathbf{E}\xi_k\xi_l| \\ &\le 2K_3 A \sum_{m\le k\le l\le n} d_k d_l \left(\frac{c_k}{c_l}\right)^{\beta} \le 2K_3 A V_{m,n}. \end{aligned}$$

Hence if we choose γ so large that $(4\gamma)^4 \geq 2K_3A$, then (37) holds for $p=2$.

Assume now that (37) is true for $p-1 \geq 2$. From $d_k \gg d_k^* = \log(c_{k+1}/c_k)$ and (33) it follows that there is a positive constant K_4 such that

$$\sum_{k=1}^{l} d_k c_k^{\beta} \geq K_4 c_l^{\beta}.$$

Now choose γ so large that the C_p defined in (36) satisfies $C_p > (2/K_4)^p \gamma^{p/2}$. Then using $|\xi_l| \leq 2$ we get for $V_{m,n} \leq \gamma$

$$\left|\sum_{l=m}^{n} d_l \xi_l\right| \leq (2/K_4) \sum_{l=m}^{n} d_l c_l^{-\beta} \left(\sum_{k=1}^{l} d_k c_k^{\beta}\right) = (2/K_4) V_{m,n} \leq (2/K_4)\, \gamma^{1/2} V_{m,n}^{1/2}.$$

Hence in the case $V_{m,n} \leq \gamma$ relation (37) is valid. We now show that if $X \geq \gamma$ is arbitrary and (37) holds for $V_{m,n} \leq X$, then it will also hold for $V_{m,n} \leq 3X/2$. As the validity of (37) is already verified for $V_{m,n} \leq \gamma$, this will show that (37) holds for any value of $V_{m,n}$, and this will complete the induction step.

Assume $V_{m,n} \leq 3X/2$ and set

$$S_1 + S_2 := \sum_{k=m}^{w} d_k \xi_k + \sum_{k=w+1}^{n} d_k \xi_k \qquad (m \leq w \leq n).$$

Put further

$$T_2 := \sum_{k=w+1}^{n} d_k \xi_{w,k}.$$

For a fixed m and n we choose w in such a way that

$$V_{m,w} \leq X, \qquad V_{w+1,n} \leq X \qquad \text{and} \qquad \frac{V_{w+1,n}}{V_{m,w}} = \lambda \in [1/2, 1].$$

To see that this is possible, we have to show that for every choice of $1 \leq m < n$ with $V_{m,n} \geq X$ there is some $w \in \{m+1, \ldots, n\}$ with

$$\frac{1}{2} V_{m,n} \leq V_{m,w} \leq \frac{2}{3} V_{m,n}. \tag{38}$$

We define $w := \min\{k \geq m : (1/2)V_{m,n} \leq V_{m,k}\}$. Then (38) will follow if we show that the increment $V_{m,w} - V_{m,w-1} \leq 1/6 V_{m,n}$. But from (16) we get

$$\begin{aligned} V_{m,w} - V_{m,w-1} &= d_w c_w^{-\beta} \left(\sum_{k=1}^{w} d_k c_k^{\beta}\right) \ll d_w c_w^{-\beta} \frac{D_w}{(\log D_w)^{\rho}} \sum_{k=1}^{w} d_k^* c_k^{\beta} \\ &\ll d_w D_w / (\log D_w)^{\rho}, \end{aligned}$$

where we used again (33) in the last step. Since by assumption (18) the last term is bounded, and since $V_{m,n} \geq X \geq \gamma$ the result follows for sufficiently large γ.

Now we prove that

$$\mathbf{E}|S_1 + S_2|^p \leq C_p (V_{m,n})^{p/2}.$$

To do so, we need some simple inequalities.

From the mean value theorem we get

$$|S_2^j - T_2^j| \leq j\,|S_2 - T_2|(|S_2|^{j-1} + |T_2|^{j-1}) \qquad (j \geq 1). \tag{39}$$

Using Lemma 1 and the assumption that (c_k) is nondecreasing we get for all $j \geq 1$

$$\mathbf{E}|S_2 - T_2|^j \leq E_j (V_{w+1,n})^{j/2}.$$

We also have

$$\mathbf{E}|S_1|^j \leq C_j (V_{m,w})^{j/2} \qquad (1 \leq j \leq p) \tag{40}$$

and

$$\mathbf{E}|S_2|^j \leq C_j (V_{w+1,n})^{j/2} \leq C_j \lambda^{j/2} (V_{m,w})^{j/2} \qquad (1 \leq j \leq p). \tag{41}$$

For $1 \leq j \leq p-1$ the last two inequalities are valid by the induction hypothesis, and for $j = p$ they follow from the validity of (37) for $V_{m,n} \leq X$. Hence Minkowski's inequality yields

$$\mathbf{E}|T_2|^j \leq 2^j C_j \lambda^{j/2} (V_{m,w})^{j/2} \qquad (1 \leq j \leq p). \tag{42}$$

Finally combining the Hölder inequality with the latter results shows for $j = 1, 2, \ldots, p-1$

$$\begin{aligned} \mathbf{E}|S_1|^j |S_2 - T_2| |S_2|^{p-j-1} &\leq (\mathbf{E}|S_1|^p)^{j/p} \, (\mathbf{E}|S_2 - T_2|^p)^{1/p} \, (\mathbf{E}|S_2|^p)^{(p-j-1)/p} \\ &\leq C_p^{(p-1)/p} E_p^{1/p} \lambda^{(p-j)/2} (V_{m,w})^{p/2}. \qquad (43) \end{aligned}$$

The last inequality remains valid, with an extra factor 2^{p-j-1} on the right hand side, if $|S_2|^{p-j-1}$ on the left hand side is replaced by $|T_2|^{p-j-1}$. Since S_1 and T_2 are independent, we get by the binomial formula and the triangle inequality

$$\begin{aligned} \mathbf{E}\,|S_1 + S_2|^p &\leq \mathbf{E}|S_1|^p + \mathbf{E}|S_2|^p \\ &+ \sum_{j=1}^{p-1} \binom{p}{j} (\mathbf{E}|S_1|^j |S_2^{p-j} - T_2^{p-j}| + \mathbf{E}|S_1|^j \, \mathbf{E}|T_2|^{p-j}) \,. \end{aligned}$$

We substitute (39) and (40)-(43) (using also the analogue of (43) with $|T_2|^{p-j-1}$) in the above inequality and get

$$\mathbf{E}|S_1+S_2|^p \le C_p(V_{m,w})^{p/2}\Bigg(1+\lambda^{p/2} \quad + \quad C_p^{-1/p}\,E_p^{1/p}\sum_{j=1}^{p-1} 2^{\,p-j}\binom{p}{j}(p-j)\lambda^{(p-j)/2}$$

$$+ \quad C_p^{-1}\sum_{j=1}^{p-1} 2^{p-j}\lambda^{(p-j)/2}\binom{p}{j}C_j\,C_{p-j}\Bigg)\ .$$

Now

$$C_p^{-1/p}E_p^{1/p} \le K\cdot p^{1/2}(4\gamma)^{-p}, \qquad C_jC_{p-j}/C_p \le (4\gamma)^{-p} \quad (1\le j\le p-1)$$

and thus by $\lambda \le 1$

$$C_p^{-1/p}\,E_p^{1/p}\sum_{j=1}^{p-1} 2^{\,p-j}\binom{p}{j}(p-j)\lambda^{(p-j)/2} \le \text{const}\cdot p^{\,3/2}\gamma^{\,-p}$$

and

$$C_p^{-1}\sum_{j=1}^{p-1} 2^{p-j}\lambda^{(p-j)/2}\binom{p}{j}C_j\,C_{p-j} \le \text{const}\cdot\gamma^{\,-p}.$$

Since $\lambda \ge 1/2$ we see that for a large enough γ the relation $\mathbf{E}|S_1+S_2|^p \le C_p\,(1+\lambda)^{p/2}\,(V_{m,w})^{p/2} = C_p(V_{m,n})^{p/2}$ is true. Thus we proved the validity of (37) for $V_{m,n} \le 3X/2$ and the proof of Lemma 2 is completed. □

The following lemma estimates the double sum appearing on the right hand side of (35).

Lemma 3. *Assume that (16) holds, then for any* $\beta > 0$ *and any* $\eta < \rho$ *we have*

$$\sum_{1\le k\le l\le N} d_kd_l\left(\frac{c_k}{c_l}\right)^{\beta} = \mathcal{O}\left(\frac{D_N^2}{(\log D_N)^{\eta}}\right).$$

Proof. By the monotonicity of (c_k) we have

$$\sum_{1\le k\le l\le N} d_kd_l\left(\frac{c_k}{c_l}\right)^{\beta}$$

$$\le \sum_{\substack{1\le l\le N\\ c_k\le c_l/(\log D_N)^{\rho/\beta}}} d_kd_l\left(\frac{c_k}{c_l}\right)^{\beta} + \sum_{\substack{1\le l\le N\\ c_l/(\log D_N)^{\rho/\beta}<c_k\le c_l}} d_kd_l =: \sigma_N+\tau_N.$$

Clearly $\sigma_N \le D_N^2(\log D_N)^{-\rho}$ and by (16)

$$\tau_N \ll \sum_{1\le l\le N} d_l \sum_{c_l/(\log D_N)^{\rho/\beta}<c_k\le c_l} d_k^* \frac{D_k}{(\log D_k)^\rho}$$

$$\ll \frac{D_N}{(\log D_N)^\rho} \sum_{1\le l\le N} d_l \sum_{c_l/(\log D_N)^{\rho/\beta}<c_k\le c_l} \log(c_{k+1}/c_k) \ll \frac{D_N^2}{(\log D_N)^\rho} \log\log D_N.$$

□

Proof of Theorem 1. We use notation (31). From Lemmas 2-3 and the Markov inequality we derive for any $\varepsilon > 0$, $p \in \mathbb{N}$,

$$P\left(\left|\sum_{k=1}^N d_k\xi_k\right| > \varepsilon D_N\right) \le c(p,\varepsilon)(\log D_N)^{-p\eta/2} \quad \text{for} \quad N \ge N_0.$$

By (18) we get $d_k \to 0$ and consequently $D_{N+1}/D_N \to 1$. Thus we can choose an increasing sequence (N_j) of positive integers such that $D_{N_j} \sim \exp(\sqrt{j})$. Hence choosing $p > 4/\eta$ and using the Borel-Cantelli lemma we get

$$\lim_{j\to\infty} \frac{1}{D_{N_j}} \sum_{k=1}^{N_j} d_k\xi_k = 0 \quad \text{a.s.}$$

For $N_j \le N < N_{j+1}$ we have by $|\xi_k| \le 2$

$$\frac{1}{D_N}\left|\sum_{k=1}^N d_k\xi_k\right| \le \frac{1}{D_{N_j}}\left|\sum_{k=1}^{N_j} d_k\xi_k\right| + 2\left(\frac{D_{N_{j+1}}}{D_{N_j}} - 1\right).$$

Since $D_{N_{j+1}}/D_{N_j} \to 1$, the convergence of the subsequence implies that the whole sequence converges a.s.

We have proved Theorem 1 for all bounded Lipschitz functions f. The result for indicator functions follows by routine approximation arguments, similar e.g. to those in [23]. □

Acknowledgement. I am grateful to Prof. István Berkes for several inspiring discussions and his valuable comments.

References

[1] R.G. Antonini and M. Weber, The intersective ASCLT, *Stochastic Anal. Appl.*, 22 (2004), 1009–1025.

[2] M. Atlagh, Théorèm central limite presque sûr et loi du logarithme itéré pour de sommes de variables alétoires indépendantes, *C.R. Acad. Sci. Paris Sér. I Math.*, 316 (1993), 929–933.

[3] M. Atlagh and M. Weber, Un théorème central limite presque sûr relatif à des sous-suites, *C. R. Acad. Sci. Paris Sér. I Math.*, 315 (1992), 203–206.

[4] M. Atlagh and M. Weber, Le théorème central limite presque sûr, *Expo. Math.*, 18 (2000), 97–126.

[5] I. Berkes. Results and problems related to the pointwise central limit theorem, In *Asymptotic methods in probability and statistics (Ottawa, ON, 1997)*, pp. 59–96. North-Holland, Amsterdam, 1998.

[6] I. Berkes and E. Csáki, A universal result in almost sure central limit theory, *Stochastic Process. Appl.*, 94 (2001), 105–134.

[7] I. Berkes and H. Dehling, Some limit theorems in log density, *The Annals of Probability*, 21 (1993), 1640–1670.

[8] I. Berkes and M. Weber, Almost sure versions of the Darling-Erdős theorem, *Statist. Probab. Letters*, 76 (2006), 280–290.

[9] N.H. Bingham, C.M. Goldie, J.L. Teugels, *Regular Variation*, volume 27, Encyclopedia of Mathematics and its Applications, Cambridge, 1987.

[10] G. A. Brosamler, An almost everywhere central limit theorem, *Math. Proc. Cambridge Philos. Soc.*, 104 (1988), 561–574.

[11] K. Chandrasekharan and S. Minakshisundaram, *Typical Means*, Oxford University Press, 1952.

[12] S. Cheng, L. Peng and Y. Qi, Almost sure convergence in extreme value theory, *Math. Nachr.* 190 (1998), 43–50.

[13] E. Csáki and A. Földes, On the logarithmic average of additive functionals, *Statist. Probab. Lett.*, 22 (1995), 261–268.

[14] E. Csáki and H. Gonchigdanzan, Almost sure limit theorems for the maximum of stationary Gaussian sequences, *Statist. Probab. Lett.*, 58 (2002), 195–203.

[15] D. A. Darling and P. Erdős, A limit theorem for the maximum of normalized sums of independent random variables, *Duke Math. J.*, 23 (1956), 143–155.

[16] I. Fahrner and U. Stadtmüller, On almost sure max-limit theorems, *Statist. Probab. Lett.*, 37 (1998), 229–236.

[17] A. Fisher, Convex-invariant means and a pathwise central limit theorem, *Adv. in Math.*, 63 (1987), 213–246.

[18] A. Fisher, A pathwise central limit theorem for random walks, *Preprint*, (1989).

[19] S. Hörmann, Critical behavior in almost sure central limit theory, *J. Theoret. Probab.*, accepted for publication.

[20] S. Hörmann, A note on the almost sure convergence of central order statistics, *Probab. Math. Stat.*, 25 (2006), 317–329.

[21] L. Horváth and D. Khoshnevisan, Weight functions and pathwise local central limit theorems, *Stochastic Process. Appl.*, 59 (1995), 105–123.

[22] I. A. Ibragimov and M. A. Lifshits, On limit theorems of "almost sure" type, *Teor. Veroyatnost. i Primenen.*, 44 (1999), 328–350.

[23] M. T. Lacey and W. Philipp, A note on the almost sure central limit theorem, *Statist. Probab. Lett.*, 9 (1990), 201–205.

[24] M. Marcus and J. Rosen, Logarithmic averages for the local times of recurrent random walks and Lévy processes, *Stochastic Process. Appl.*, 59 (1995), 175–184.

[25] P. Mörters, Almost sure Kallianpur-Robbins laws for Brownian motion in the plane, *Probab. Theory Related Fields*, 118 (2000), 49–64.

[26] P. Mörters, A pathwise version of Spitzer's law, In *Limit theorems in probability and statistics, Vol. II (Balatonlelle, 1999)*, pp. 427–436. János Bolyai Math. Soc., Budapest, 2002.

[27] M. Peligrad and P. Révész, On the almost sure central limit theorem, In *Almost everywhere convergence, II (Evanston, IL, 1989)*, pp. 209–225. Academic Press, Boston, MA, 1991.

[28] L. Peng and Y. Qi, Almost sure convergence of the distributional limit theorem for order statistis, *Prob. Math. Stat.*, 23 (2003), 217–228.

[29] B. Rodzik and Z. Rychlik, An almost sure central limit theorem for independent random variables, *Ann. Inst. H. Poincaré Probab. Statist.*, 30 (1994), 1–11.

[30] B. Rodzik and Z. Rychlik, On the central limit theorem for independent random variables with almost sure convergence, *Probab. Math. Statist.*, 16 (1996), 299–309.

[31] Z. Rychlik and K. Szuster, On strong versions of the central limit theorem, *Statist. Probab. Lett.*, 61 (2003), 347–357.

[32] P. Schatte, On strong versions of the central limit theorem, *Math. Nachr.*, 137 (1988), 249–256.

[33] G.R. Shorack, Extension of the Darling and Erdős theorem on the maximum of normalized sums, *Ann. Probab.*, 7 (1979), 1092–1096.

[34] U. Stadtmüller, Almost sure versions of distributional limit theorems for certain order statistics, *Stat. Probab. Lett.*, 58 (2002), 413–426.

[35] M. Weber, Un théorème central limite presque sûr à moments généralisés pour les rotations irrationnelles, *Manuscripta Math.*, 101 (2000), 175–190.