J Theor Probab
DOI 10.1007/s10959-007-0080-3

Critical Behavior in Almost Sure Central Limit Theory

Siegfried Hörmann

Received: 30 November 2005 / Revised: 18 October 2006

Abstract Let $X_1, X_2, \dots$ be i.i.d. random variables with $EX_1 = 0$, $EX_1^2 = 1$ and let $S_k = X_1 + \cdots + X_k$. We study the a.s. convergence of the weighted averages

$$D_N^{-1} \sum_{k=1}^{N} d_k I\left\{ \frac{S_k}{\sqrt{k}} \le x \right\},$$

where (d_k) is a positive sequence with $D_N = \sum_{k=1}^{N} d_k \to \infty$. By the a.s. central limit theorem, the above averages converge a.s. to $\Phi(x)$ if $d_k = 1/k$ (logarithmic averages) but diverge if $d_k = 1$ (ordinary averages). Under regularity conditions, we give a fairly complete solution of the problem for what sequences (d_k) the weighted averages above converge, resp. the corresponding LIL and CLT hold. Our results show that logarithmic averaging, despite its prominent role in a.s. central limit theory, is far from optimal and considerably stronger results can be obtained using summation methods near ordinary (Cesàro) summation.

Keywords Almost sure central limit theorem · Summation methods · Law of the iterated logarithm

1 Introduction and Results

Let $X_1, X_2, \dots$ be i.i.d. random variables with $EX_1 = 0$, $EX_1^2 = 1$ and let $S_k = X_1 + \cdots + X_k$. The simplest version of the almost sure central limit theorem (ASCLT)

S. Hörmann (✉)
Institute of Statistics, Graz University of Technology, Steyrergasse 17/IV, 8010 Graz, Austria
e-mail: shoermann@tugraz.at

states that

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{k=1}^{N} \frac{1}{k} I\left\{\frac{S_k}{\sqrt{k}} \le x\right\} = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2}\, dt \quad \text{a.s.} \tag{1.1}$$

for every fixed $x \in \mathbb{R}$. This result was proved by Brosamler [7] and Schatte [17] under some additional moment conditions and by Fisher [9] and Lacey and Philipp [11] assuming only finite variances. (Actually, (1.1) was known to Lévy [12, p. 270] but he did not specify conditions and gave no proof.) In recent years, many authors investigated limit theorems of this type and several variants and extensions of (1.1) have been obtained. We refer to Atlagh and Weber [1] and Berkes [3] for surveys of the field.

A characteristic feature of the theory is the use of logarithmic averages in (1.1), and from the arc sine law it follows that with ordinary averages relation (1.1) fails even for $x = 0$. Why logarithmic averages work here is best seen from the Wiener analogue of (1.1), i.e.

$$\lim_{N\to\infty} \frac{1}{\log N} \int_{1}^{N} \frac{1}{t} I\left\{\frac{W(t)}{\sqrt{t}} \le x\right\} dt = (2\pi)^{-1/2} \int_{-\infty}^{x} e^{-t^2/2}\, dt \quad \text{a.s. for all } x.$$

After the transformation $t = e^u$ this reduces to the ergodic theorem for the Ornstein-Uhlenbeck process $e^{-u/2}W(e^u)$. Via a strong approximation argument, this also proves the ASCLT (1.1) under moment conditions only slightly stronger than $EX_1^2 < \infty$. Despite the simplicity of this argument, it is important to note that logarithmic summation is not the only possible summation that leads to a.s. convergence to $\Phi(x)$ in (1.1). Peligrad and Révész [14] showed that

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^{N} d_k I\left\{\frac{S_k}{\sqrt{k}} \le x\right\} = \Phi(x) \quad \text{a.s.,} \tag{1.2}$$

holds if

$$d_k = \frac{(\log k)^\alpha}{k}, \qquad D_n = \sum_{k=1}^{n} d_k \quad (\alpha > -1) \tag{1.3}$$

and Berkes and Csáki [4] showed that (1.2) holds even if

$$d_k = \frac{\exp((\log k)^\alpha)}{k} \quad (0 \le \alpha < 1/2). \tag{1.4}$$

For summation methods in a.s. limit theory which are different from log-summation we also refer to Becker-Kern [2] and Weber [19]. An extension of these results is given in Hörmann [10] where we proved that the summation defined in (1.4) applies for any $\alpha \in [0, 1)$. Furthermore in [4] it is observed that (1.2) is valid for any $d_k \le 1/k$ with $\sum d_k = \infty$. Thus the weights $d_k = 1/k$ play no special role in (1.2); there is a large class of possible weight sequences, and several smaller and larger sequences work equally well. To understand this phenomenon better, let us recall some results from classical summability theory. Given a positive sequence $\mathbf{D} = (d_k)$ with $D_n = \sum_{k=1}^{n} d_k \to \infty$, we say that a sequence (x_n) 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 [8, p. 35]), if $\mathbf{D}$ and $\mathbf{D}^*$ are summation procedures with $D_n^* = O(D_n)$, then under minor technical assumptions, the summation $\mathbf{D}^*$ is stronger than $\mathbf{D}$, i.e. if a sequence (x_n) is $\mathbf{D}$-summable to x, then it is also $\mathbf{D}^*$-summable to x. Moreover, by a result of Zygmund (see [8, 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 $D_n^* = O(D_n^\varepsilon)$ for every $\varepsilon > 0$, then $\mathbf{D}^*$ is strictly stronger than $\mathbf{D}$. For example, logarithmic summation, defined by $d_n = 1/n$ is stronger than ordinary (Cesàro) summation defined by $d_n = 1$ and weaker than loglog summation defined by $d_n = 1/(n \log n)$. On the other hand, all summation methods defined by

$$d_n = (\log n)^\alpha / n, \quad \alpha > -1$$

are equivalent to logarithmic summation and all summation methods defined by

$$d_n = n^\alpha, \quad \alpha > -1$$

are equivalent to Cesàro summation. These remarks show that relation (1.2) with the sequence in (1.3) is, despite their formal difference, actually equivalent to the case $d_k = 1/k$ and show also that the sequences in (1.4) define summation procedures which are pairwise nonequivalent and also nonequivalent with logarithmic averaging. The result of Hardy also shows that by increasing the weight sequence (d_k) in (1.2), the result becomes stronger. Thus the strongest, "true" form of the a.s. central limit theorem is the one with the largest weight sequence (d_k). This summation method (d_k) is unknown and its determination will be the objective of the present paper. We will also study the optimal weight sequences (d_k) in refined versions of (1.2), for example in the CLT and LIL corresponding to the strong law (1.2). Our main result will show that, under certain regularity conditions on (d_k), relation (1.2) and the corresponding CLT and LIL hold provided

$$d_n = O\left(\frac{D_n}{n(\log\log n)^\alpha}\right) \tag{1.5}$$

for $\alpha > 3$ and this becomes false if $\alpha < 1$. Here and in the sequel we write $\log\log x$ when we mean $\log(\max\{\log x, e\})$. Thus the optimal weight condition for (1.2) and the corresponding CLT and LIL is relation (1.5) with some $1 \le \alpha \le 3$, whose value remains unknown. Condition (1.5) is an asymptotic negligibility condition resembling Kolmogorov's classical condition for the LIL, except the factor n in the denominator on the right hand side, which is due to the strong dependence of the sequence $(S_n/\sqrt{n})$ and which forces the summation D_n to grow considerably slower than the norming sequence in the classical LIL for independent r.v.'s. (Note that a similar effect of strong dependence leads to unusual coefficients in the LIL for lacunary trigonometric series, see Takahashi [18].) In terms of the norming factor D_n, our results show that (1.2) and the corresponding LIL are valid if

$$D_n = \exp(\log n/(\log\log n)^\alpha)$$

for $\alpha > 3$, but not if $\alpha < 1$. The CLT remains valid for $\alpha > 1$. Recalling that $D_n = \log n$ resp. $D_n = n = \exp(\log n)$ correspond to logarithmic, resp. Cesàro averaging, the last relation shows the surprising fact that the critical weight sequence in the ASCLT is, in some sense, much closer to Cesàro than to logarithmic averaging. Thus, despite the prominent role log averaging plays in a.s. central limit theory, its true significance is secondary.

Incidentally, the methods of our paper will also lead to optimal conditions for the "stochastic" version of the ASCLT, i.e. when we require relation (1.2) in probability. Theorem 5 at the end of this section will show that the stochastic ASCLT holds if

$$d_n = o(D_n/n). \tag{1.6}$$

Replacing o by O in (1.6), the result becomes false, as the example $d_n = 1$ shows. Theorem 5 states that Relation (1.6) holds e.g. if

$$D_n = \exp(\log n/\omega(n)), \tag{1.7}$$

where $\omega(n) \to \infty$, $\log n/\omega(n) \nearrow \infty$ and $\omega(n)$ satisfies mild regularity conditions. By Zygmund's theorem quoted above, the summation method defined by (1.7) is stronger than Cesàro summation (corresponding to $\omega(n) = 1$) and conversely, any summation method stronger than Cesàro summation and satisfying suitable regularity conditions has the above representation. Thus if we are interested in convergence in probability in (1.2), the summation method can be pushed arbitrary close to Cesàro summation.

We turn now to formulating our results in detail. In order to specify regularity conditions for the summation procedure **D**, recall first that for each $\varepsilon > 0$, the sequence $D_n = n^\varepsilon$ defines a summation equivalent to Cesàro summation, and hence this D_n is already too large for (1.2). Thus without losing much generality, we can assume that $D_n = O(n^\varepsilon)$ for each $\varepsilon > 0$. All slowly varying sequences D_n satisfy this condition and confining our attention to this class will not put additional restrictions on the speed of growth of D_n, since for every sequence D_n satisfying $D_n = O(n^\varepsilon)$ for all $\varepsilon > 0$ there exists a slowly varying D_n^* such that $D_n = o(D_n^*)$ for $n \to \infty$ (cf. Bingham et al. [10, Theorem 2.3.6]). Hence in searching for the largest possible norming sequence D_n in ASCLT theory, we may assume that D_n is slowly varying. By the theory of regular variation, D_n can be represented in the form

$$D_n = c_n \exp\left(\int_A^n \varepsilon(u)/u\, du\right) \quad (n \geq A), \tag{1.8}$$

where $A > 0$, $c_n \to c \in (0, \infty)$, and $\varepsilon(x) \to 0$ for $x \to \infty$. Our final technical assumption on D_n will require that $c_n = 1$, ε is non-increasing, slowly varying and obeys the condition

$$\varepsilon(x)/\varepsilon(x^2) = O(1) \quad (x \to \infty). \tag{1.9}$$

These conditions are stronger than necessary and could be easily weakened, but they will simplify our calculations considerably, and, as before, put no extra restrictions on the speed of growth of D_n. Regarding (1.9), note that $\varepsilon(x)$ must tend to 0 very slowly in order that $D_n \to \infty$. E.g., $\varepsilon(x) = (\log x)^{-(1+\eta)}$, $\eta > 0$ implies already that the exponent in (1.8) is bounded. However, this $\varepsilon(x)$ still satisfies (1.9). For the same reason, the assumption of slow variation of ε imposes only a regularity condition for $\varepsilon(x)$, but it puts no restriction on its speed of decrease.

The previously discussed technical conditions are summarized in the following

Definition A summation method **D** belongs to the class $\mathcal{W}$ if $D_n \to \infty$ and

$$D_n = \exp\left(\int_A^n \varepsilon(u)/u\, du\right) \quad (n \geq A), \tag{1.10}$$

where $\varepsilon(x)$ is non-increasing, slowly varying, tends to 0 for $x \to \infty$, and satisfies (1.9).

From the mean value theorem it follows that

$$d_n \sim D_n \frac{\varepsilon(n)}{n} \quad (n \to \infty)$$

($a_n \sim b_n$ means that $a_n/b_n \to 1$) and thus $d_n = L(n)/n$ where L is slowly varying. We mention a few examples.

(a) $D_N = (\log N)^\gamma$, $\gamma > 0$;
(b) $D_N = \exp((\log N)^\beta)$, $0 < \beta < 1$;
(c) $D_N = \exp(\log N/(\log\log N)^\alpha)$, $\alpha > 0$.

Let $\mathcal{L}$ denote the class of bounded Lipschitz 1 functions on $\mathbb{R}$. By a standard observation in a.s. central limit theory (see e.g. Lacey and Philipp [11]), relation (1.2) follows if

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^N d_k f\left(\frac{S_k}{\sqrt{k}}\right) = \int_{-\infty}^{\infty} f(t)d\Phi(t) \quad \text{a.s.} \tag{1.11}$$

for every $f \in \mathcal{L}$. For this reason, we will work in our paper with the version of the ASCLT of the type (1.11).

We are ready now to formulate our results. In the sequel $\lambda(B)$ denotes the Lebesgue measure of some Borel set B. Finally we allude once more to the convention $\log\log x = \log(\max\{\log x, e\})$.

Theorem 1 *Let $X_1, X_2, \ldots$ be i.i.d. random variables satisfying $EX_1 = 0$ and $EX_1^2 = 1$ and put $S_n = X_1 + \cdots + X_n$. Assume that* $\mathbf{D} \in \mathcal{W}$, *the relation*

$$d_k = O\left(\frac{D_k}{k(\log\log k)^\alpha}\right) \tag{1.12}$$

holds for some $\alpha > 3$ and

$$(kd_k) \text{ is non-decreasing.} \tag{1.13}$$

Then we have for every $f \in \mathcal{L}$ *or* f *an indicator function of a Borel set* A *with* $\lambda(\partial A) = 0$

$$\lim_{N\to\infty} \frac{1}{D_N} \sum_{k=1}^{N} d_k \left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right) = 0 \quad a.s. \tag{1.14}$$

Theorem 2 *Let* $X_1, X_2, \ldots$ *be i.i.d. random variables satisfying* $EX_1 = 0$ *and* $EX_1^2 = 1$ *and put* $S_n = X_1 + \cdots + X_n$. *Assume that* $\mathbf{D} \in \mathcal{W}$, (1.12) *holds for some* $\alpha > 1$ *and*

$$\liminf_{k\to\infty} kd_k > 0. \tag{1.15}$$

Then we have for every non-constant $f \in \mathcal{L}$

$$\lambda_N^{-1/2} \sum_{k=1}^{N} d_k \left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right) \xrightarrow{d} \mathcal{N}, \tag{1.16}$$

where

$$\lambda_N := \mathrm{Var}\left(\sum_{k=1}^{N} d_k f\left(\frac{S_k}{\sqrt{k}}\right)\right) \qquad (N \geq 1), \tag{1.17}$$

and $\mathcal{N}$ *is a standard normal r.v.*

Both relations (1.13) and (1.15) imply that $d_k \geq C/k$. In Theorem 1 this is indeed the interesting case, since relation (1.14) is known to hold for all $d_k \leq 1/k$ with $\sum d_k = \infty$, see Berkes and Csáki [4]. Regarding Theorem 2, we will show in Lemma 4 and Lemma 5 that if $\mathbf{D} \in \mathcal{W}$, then the order of magnitude of λ_N is $\sum_{k=1}^{N} kd_k^2$. Hence (1.15) implies that $\lambda_N \to \infty$. Without (1.15), Theorem 2 fails: for example, if $d_k = k^{-1}(\log k)^{-a}$ with $1/2 < a \leq 1$ (which is in $\mathcal{W}$), then $\limsup_N \lambda_N < \infty$ and thus the sum in (1.16) remains bounded in probability.

In the case $D_N = \log N$ Theorem 2 was proved by Berkes and Horváth [5]. Moreover, the conditions of Theorem 2 are satisfied in Example (a) if $\gamma \geq 1$, in Example (b) for all $0 < \beta < 1$ and in Example (c) if $\alpha > 1$.

Theorem 3 *Let* $X_1, X_2, \ldots$ *be i.i.d. random variables satisfying* $EX_1 = 0$ *and* $EX_1^2 = 1$ *and put* $S_n = X_1 + \cdots + X_n$. *Let* λ_N $(N \geq 1)$ *be defined as in* (1.17). *Assume that* $\mathbf{D} \in \mathcal{W}$ *satisfies* (1.15) *and relation* (1.12) *holds for some* $\alpha > 3$. *Then we have for every non-constant* $f \in \mathcal{L}$

$$\limsup_{N\to\infty} (2\lambda_N \log\log \lambda_N)^{-1/2} \sum_{k=1}^{N} d_k \left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right) = 1 \quad a.s. \tag{1.18}$$

The conditions of Theorem 3 are satisfied in Example (a) if $\gamma \geq 1$, in Example (b) for all $0 < \beta < 1$ and in Example (c) if $\alpha > 3$.

Our theorems show that under regularity assumptions on $\mathbf{D}$, the strong law (1.14), the CLT (1.16) and the LIL (1.18) are all valid provided the Kolmogorov type condition (1.12) holds for $\alpha > 3$. The next theorem shows that except for the numerical value of α, condition (1.12) is sharp.

Theorem 4 *For every $0 < \alpha < 1$ there exists a summation procedure $\mathbf{D} \in \mathcal{W}$ satisfying (1.12) and (1.13) such that the LIL (1.18) fails.*

In fact, this is the case if $\mathbf{D}$ is defined by

$$D_N = \exp(\log N/(\log\log N)^\alpha). \tag{1.19}$$

Thus the sequence (1.19) is critical in the theory: for $\alpha > 3$ it implies all of (1.14), (1.16), (1.18) and this becomes false if $\alpha < 1$. What happens for $1 \le \alpha \le 3$ remains open.

By Hardy's minoration principle, if (1.2) holds with a weight sequence (d_k), then, under certain regularity conditions, it will also hold for all smaller weight sequences (d_k^*). Hardy assumed that $D_n^* = \psi(D_n)$, where $\psi(x) \le x$ is an elementary function composed of rational, exponential and logarithmic functions. A much larger class of ψ's was constructed by Hirst [8, p. 37]. It seems likely that for sequences $d_k \ge 1/k$ an analogous minoration principle holds for the CLT (1.16) and the LIL (1.18), but this remains open. As the remarks after Theorem 2 show, without $d_k \ge 1/k$ this minoration principle is not valid.

Our final result gives a sharp condition for relation (1.14) to hold in probability.

Theorem 5 *Let $X_1, X_2, \dots$ be i.i.d. random variables satisfying $EX_1 = 0$ and $EX_1^2 = 1$ and put $S_n = X_1 + \cdots + X_n$. Assume that*

$$d_n = o(D_n/n). \tag{1.20}$$

Then we have for every $f \in \mathcal{L}$ or f an indicator function of a Borel set A with $\lambda(\partial A) = 0$

$$\frac{1}{D_N}\sum_{k=1}^N d_k\left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right) \to 0 \quad \text{in probability.}$$

If o in (1.20) is replaced with O the proposition fails.

In comparison to (1.12) we can omit the extra factor $(\log\log n)^\alpha$ in the denominator for the stochastic version of the ASCLT. As the example $d_n = 1$ (Cesáro summation) shows, we cannot replace o with O in (1.20). For example relation (1.20) holds for

$$D_n = \exp(\log n/\omega(n))$$

if $w(x)$ is some differentiable function satisfying $w(x) \to \infty$ (in order that D_n is not equivalent to Cesáro averaging) and $\log x/w(x) \nearrow \infty$ (in order that D_n defines a summation method) and

$$x\omega'(x)/\omega(x) = O((\log x)^{-1}). \tag{1.21}$$

The last condition is satisfied e.g. if $\omega(x)=\log_k x$ $(k=2,3,\ldots)$ where $\log_k$ denotes k times iterated logarithm. It is also easy to see that (1.21) permits arbitrary slow increase of ω.

2 Auxiliary Lemmas

Lyapunov's classical CLT condition or Kolmogorov's condition for the LIL provide the corresponding limit theorems in terms of specific moment assumptions. Via a blocking technique the proofs of our theorems will make use of these results. Hence an accurate study of the variances respectively higher moments of the processes under investigation is important. The core of this section are Lemma 4 and Lemma 5. In what follows, let f be a bounded Lipschitz 1 function on $\mathbb{R}$; without loss of generality we assume $|f|\le 1$. All the constants occurring in the following lemmas may depend on f, $\mathbf{D}$ and the sequence $X_1, X_2, \ldots$; we will make no mention of this fact in the sequel. Constants like C_p, $C(\varepsilon)$, etc. may depend also on the parameters indicated. The relation $a_n \ll b_n$ will mean $|a_n/b_n|=O(1)$.

Lemma 1 *Assume $f\in\mathcal{L}$. Then there is a constant C such that for all $1\le k\le l$*

$$\left|\operatorname{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)\right| \le C\left(\frac{k}{l}\right)^{1/2}.$$

Lemma 1 is a standard tool in a.s. central limit theory. For completeness, we give the short

Proof Clearly

$$\operatorname{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)=\operatorname{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)-f\left(\frac{S_l-S_k}{\sqrt{l}}\right)\right).$$

Hence the Lipschitz continuity of f, $|f|\le 1$ and the Cauchy-Schwarz inequality give

$$\begin{aligned}&\left|\operatorname{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)\right|\\ &\quad\le 2E\left|f\left(\frac{S_l}{\sqrt{l}}\right)-f\left(\frac{S_l-S_k}{\sqrt{l}}\right)\right|\le CE\left|\frac{S_k}{\sqrt{l}}\right|\le C\sqrt{\frac{k}{l}}. \qquad \square \quad (2.1)\end{aligned}$$

Lemma 2 *Let $f\in\mathcal{L}$, f non-constant. Then there exist an integer $m\ge 1$, a real $c>0$ and for every $\varepsilon>0$ an $A=A(\varepsilon)$ such that*

$$\operatorname{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right)\ge c\left(\frac{k}{l}\right)^{m/2} \quad \textit{for } A\le k<l,\ k/l\ge\varepsilon^{2/m}.$$

Proof Consider a Wiener process $\{W_t, t\ge 0\}$. From Rozanov [16, 182 f.] we get

$$\operatorname{Cov}\left(f\left(\frac{W_k}{\sqrt{k}}\right), f\left(\frac{W_l}{\sqrt{l}}\right)\right)=\sum_{\nu=1}^{\infty}\frac{\rho^\nu}{\nu!}\alpha_\nu^2 \quad (1\le k\le l), \tag{2.2}$$

where α_ν are the coefficients of the Hermite expansion of $g := f - Ef(W_1)$, i.e.

$$g(x) = \sum_{\nu=1}^{\infty} \frac{\alpha_\nu}{\nu!} H_\nu(x),$$

$$H_\nu(x) = (-1)^\nu e^{x^2/2} \frac{d^\nu}{dx^\nu} e^{-x^2/2}$$

and $\rho = \sqrt{k/l}$ is the correlation between $W_k/\sqrt{k}$ and $W_l/\sqrt{l}$ for $k \le l$. Since we exclude the trivial case where f is constant, there is a $\nu \ge 1$ such that $\alpha_\nu^2 > 0$. Let m be the smallest of these integers. By (2.2) there exists a $c_1 > 0$ such that

$$\mathrm{Cov}\left(f\left(\frac{W_k}{\sqrt{k}}\right), f\left(\frac{W_l}{\sqrt{l}}\right)\right) \ge c_1 \left(\frac{k}{l}\right)^{m/2} \quad (1 \le k \le l). \tag{2.3}$$

To estimate the covariance in the general case we use an invariance principle of Major [13] which implies that we can define $X_1, X_2, \ldots$ on a new probability space together with a Wiener process $\{W_t, t \ge 0\}$ such that

$$(S_n - W_n)/\sqrt{n} \xrightarrow{P} 0.$$

Set

$$c_{k,l} := \mathrm{Cov}\left(f\left(\frac{S_k}{\sqrt{k}}\right), f\left(\frac{S_l}{\sqrt{l}}\right)\right), \qquad c_{k,l}^* := \mathrm{Cov}\left(f\left(\frac{W_k}{\sqrt{k}}\right), f\left(\frac{W_l}{\sqrt{l}}\right)\right).$$

Since f is a bounded Lipschitz function, it is easy to see that

$$|c_{k,l}^* - c_{k,l}| \to 0 \quad \text{for } \min\{k, l\} \to \infty.$$

Hence for every $\varepsilon > 0$ there is an $A = A(\varepsilon)$ such that

$$|c_{k,l} - c_{k,l}^*| \le \varepsilon^2 \quad \text{for } \min\{k, l\} \ge A. \tag{2.4}$$

By (2.3) we have

$$c_{k,l}^* \ge c_1 \varepsilon \quad \text{if } k/l \ge \varepsilon^{2/m}, \ 1 \le k \le l,$$

and together with (2.4) and again (2.3) this yields

$$c_{k,l} \ge c_{k,l}^* \left(1 - \frac{\varepsilon}{c_1}\right) \ge (c_1 - \varepsilon)\left(\frac{k}{l}\right)^{m/2} \quad \text{if } k/l \ge \varepsilon^{2/m}, \ A \le k \le l.$$

Since we may assume that $\varepsilon < c_1/2$ this proves the assertion. □

To simplify the notation, in the sequel we will write for $1 \le k \le l$

$$\xi_l := f\left(\frac{S_l}{\sqrt{l}}\right) - Ef\left(\frac{S_l}{\sqrt{l}}\right) \quad \text{and} \quad \xi_{k,l} := f\left(\frac{S_l - S_k}{\sqrt{l}}\right) - Ef\left(\frac{S_l - S_k}{\sqrt{l}}\right).$$

Lemma 3 *If $k \le m \le n$ and (d_l) is an arbitrary sequence of positive numbers, then we have for every $p \in \mathbb{N}$*

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

where

$$E_p = \text{const} \cdot 4^p(2p)^{p/2}.$$

Proof We set $Q(l) = Q(k,l) = \xi_l - \xi_{k,l}$. By $f \in \mathcal{L}$, $|f| \le 1$ and (2.1) we have

$$E|Q(l)|^p \le 4^{p-1}E|Q(l)| \le 4^{p-1} \cdot 2E\left|f\left(\frac{S_l}{\sqrt{l}}\right) - f\left(\frac{S_l - S_k}{\sqrt{l}}\right)\right| \le 4^{p-1}C\sqrt{\frac{k}{l}}.$$

Using the Hölder inequality we obtain

$$\begin{aligned}
E\left|\sum_{l=m}^{n} d_l(\xi_l - \xi_{k,l})\right|^p &\le \sum_{l_1=m}^{n} \cdots \sum_{l_p=m}^{n} d_{l_1} \cdots d_{l_p}(E|Q(l_1)|^p \cdots E|Q(l_p)|^p)^{1/p} \\
&\le 4^{p-1}Ck^{1/2}\sum_{l_1=m}^{n} \cdots \sum_{l_p=m}^{n} d_{l_1} \cdots d_{l_p} l_1^{-\frac{1}{2p}} \cdots l_p^{-\frac{1}{2p}} \\
&= 4^{p-1}Ck^{1/2}\left(\sum_{l=m}^{n} d_l l^{-\frac{1}{2p}}\right)^p \\
&\le 4^{p-1}Ck^{1/2}\left(\sum_{l=m}^{n} d_l^2 l\right)^{p/2}\left(\sum_{l=m}^{n} l^{-\frac{1}{p}-1}\right)^{p/2}.
\end{aligned}$$

Clearly for $m \ge 2$

$$\left(\sum_{l=m}^{n} l^{-\frac{1}{p}-1}\right)^{p/2} \le \left(\int_{m-1}^{\infty} l^{-\frac{1}{p}-1} dl\right)^{p/2} \le \left(\frac{1}{m-1}\right)^{1/2} p^{p/2},$$

and similarly, for $m = 1$ one obtains

$$\left(\sum_{l=1}^{n} l^{-\frac{1}{p}-1}\right)^{p/2} \le (1+p)^{p/2} \le (2p)^{p/2}.$$

□

Lemma 4 *Let $\mathbf{D} = (d_k)$ be a summation method with $d_k = L(k)/k$, $k \ge 1$, where $L(k) \gg 1$ and $L(k)$ is slowly varying at infinity. Then for every $f \in \mathcal{L}$ and every $p \in \mathbb{N}$*

$$E\left|\sum_{k=m}^{n} d_k\left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right)\right|^p \le C_p\left(\sum_{k=m}^{n} k d_k^2\right)^{p/2}, \tag{2.5}$$

where $C_p > 0$ is a constant.

Proof We define

$$V_{m,n} := \sum_{k=m}^{n} k d_k^2,$$

and show that

$$C_p = (4\gamma)^{p^2} \quad (p = 1, 2, \ldots)$$

will work for some fixed and large enough γ. We use induction on p. By Lemma 1 we get

$$E\left(\sum_{k=m}^{n} d_k \xi_k\right)^2 \leq 2 \sum_{m \leq k \leq l \leq n} d_k d_l \, |E\xi_k \xi_l| \leq 2C \sum_{m \leq l \leq n} d_l l^{-1/2} \sum_{1 \leq k \leq l} d_k k^{1/2}. \quad (2.6)$$

Note that for a slowly varying function L we have

$$\sum_{k=1}^{N} L(k) k^{\rho} \sim \frac{1}{\rho + 1} L(N) N^{\rho+1} \quad \text{if } \rho > -1, \quad (2.7)$$

(see e.g. [6, Corollary 1.7.3]) and therefore by the definition of d_k

$$\sum_{k=1}^{l} d_k k^{1/2} = O(d_l l^{3/2}). \quad (2.8)$$

Hence from (2.6) and (2.8) we see that (2.5) holds for $p = 2$ if γ is chosen sufficiently large. Using the Cauchy Schwarz inequality this proves also the case $p = 1$. Note that we did not use $L(k) \gg 1$ for the calculation of second moments. If $p = 1, 2$ relation (2.5) is e.g. valid for $d_k = (k \log k)^{-1}$.

Assume now that (2.5) is true for $p - 1 \geq 2$. By $L(k) \gg 1$ we can assume without loss of generality that $L(k) \geq 1$. This means $k d_k \geq 1$ and since $|\xi_k| \leq 2$ we get for $V_{m,n} \leq \gamma$

$$\left|\sum_{k=m}^{n} d_k \xi_k\right| \leq 2 \sum_{k=m}^{n} k d_k^2 \leq 2\gamma^{1/2} (V_{m,n})^{1/2}.$$

Hence in the case $V_{m,n} \leq \gamma$ (2.5) holds for every $p \geq 1$, since $C_p > (4\gamma)^{p/2}$. We show now that if $X \geq \gamma$ is arbitrary and (2.5) holds for $V_{m,n} \leq X$, then it will also hold for $V_{m,n} \leq 3X/2$. As the validity of (2.5) is already verified for $V_{m,n} \leq \gamma$, this will show that (2.5) holds for any value of $V_{m,n}$, and this will complete the induction step.

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

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

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

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

This is possible since by the remark after the definition of the class $\mathcal{W}$ in the Introduction, $d_k = L(k)/k$ with L slowly varying and thus $kd_k^2 \to 0$. Now we prove that

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

To do so, we need some simple inequalities.

From the mean value theorem we get

$$|S_2^j - T_2^j| \le j|S_2 - T_2|(|S_2|^{j-1} + |T_2|^{j-1}) \quad (j \ge 1). \tag{2.9}$$

Using Lemma 3 we get for all $j \ge 1$

$$E|S_2 - T_2|^j \le E_j (V_{w+1,n})^{j/2}. \tag{2.10}$$

For $1 \le j \le p$ we have

$$E|S_1|^j \le C_j (V_{m,w})^{j/2}, \tag{2.11}$$

and

$$E|S_2|^j \le C_j (V_{w+1,n})^{j/2} \le C_j \lambda^{j/2} (V_{m,w})^{j/2}. \tag{2.12}$$

If $1 \le j \le p-1$ the last two inequalities are valid by the induction hypothesis, and for $j = p$ they follow from $V_{m,w} \le X$, $V_{w+1,n} \le X$. Combining the latter results with the Minkowski's inequality yields

$$E|T_2|^j \le 2^j C_j \lambda^{j/2} (V_{m,w})^{j/2} \quad (1 \le j \le p). \tag{2.13}$$

Finally the Hölder inequality and (2.10–2.12) show for $j = 1, 2, \ldots, p-1$

$$\begin{aligned} E|S_1|^j |S_2 - T_2||S_2|^{p-j-1} &\le (E|S_1|^p)^{j/p} (E|S_2 - T_2|^p)^{1/p} (E|S_2|^p)^{(p-j-1)/p} \\ &\le C_p^{(p-1)/p} E_p^{1/p} \lambda^{(p-j)/2} (V_{m,w})^{p/2}. \end{aligned} \tag{2.14}$$

It follows from (2.13) that 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

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

We substitute (2.9–2.14) (also using the analogue of (2.14) with $|T_2|^{p-j-1}$) in the above inequality and get

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

Now

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

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

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

Since $\lambda\ge 1/2$ we see that for a large enough γ the relation $E|S_1+S_2|^p\le C_p(1+\lambda)^{p/2}(V_{m,w})^{p/2}=C_p(V_{m,n})^{p/2}$ is true. Thus we proved the validity of (2.5) for $V_{m,n}\le 3X/2$ and the proof of Lemma 4 is completed. □

Lemma 5 *Let* $\mathbf{D}=(d_k)$ *be a summation method with* $d_k=L(k)/k$, $k\ge 1$, *where* $L(k)\gg 1$ *and* $L(k)$ *is slowly varying at infinity. Then for every non-constant* $f\in\mathcal{L}$ *we have*

$$\operatorname{Var}\left(\sum_{k=1}^{N}d_kf\left(\frac{S_k}{\sqrt{k}}\right)\right)\gg\sum_{k=1}^{N}kd_k^2. \tag{2.15}$$

Proof Let $0<\varepsilon<1$ to be chosen later and $A(\varepsilon)$ and m the same as in Lemma 2. Set $\delta=\varepsilon^{2/m}$. Clearly it suffices to prove Lemma 5 with the summations in (2.15) started with $k=A$ instead of $k=1$. Now

$$\operatorname{Var}\left(\sum_{k=A}^{N}d_kf\left(\frac{S_k}{\sqrt{k}}\right)\right)$$
$$=\sum_{A\le k\le N}d_k^2E\xi_k^2+2\sum_{\substack{A\le i<k\le N\\ i/k<\delta}}d_id_kE\xi_i\xi_k+2\sum_{\substack{A\le i<k\le N\\ i/k\ge\delta}}d_id_kE\xi_i\xi_k$$
$$=:S^{(1)}+S^{(2)}+S^{(3)}.$$

Clearly $\sum_{k=1}^{\infty} d_k^2 < \infty$, and since the ξ_k are uniformly bounded, it follows that $S^{(1)} = O(1)$. From $d_k = L(k)/k$ with a slowly varying L we get by (2.7) and the definition of slow variation that for every $\gamma > 0$

$$\sum_{1 \le i < \delta k} d_i i^{\gamma} \sim \frac{1}{\gamma} L(\delta k)(\delta k)^{\gamma} \sim \frac{1}{\gamma} \delta^{\gamma} L(k) k^{\gamma} \quad (k \to \infty).$$

Hence for $k \ge k_0(\delta)$ we have

$$\delta^{\gamma} \frac{1}{2\gamma} k^{\gamma+1} d_k \le \sum_{1 \le i < \delta k} d_i i^{\gamma} \le \delta^{\gamma} \frac{2}{\gamma} k^{\gamma+1} d_k. \tag{2.16}$$

From Lemma 1 and (2.16) we infer

$$\begin{aligned}
|S^{(2)}| &\le 2C \sum_{\substack{1 \le i < k \le N \\ i/k < \delta}} d_i d_k \left(\frac{i}{k}\right)^{1/2} = 2C \sum_{1 \le k \le N} d_k k^{-1/2} \sum_{1 \le i < \delta k} d_i i^{1/2} \\
&\le 2C \sum_{1 \le k \le k_0} d_k k^{-1/2} \sum_{1 \le i < \delta k} d_i i^{1/2} + 8C\delta^{1/2} \sum_{1 \le k \le N} k d_k^2 \\
&= 8C\delta^{1/2} \sum_{1 \le k \le N} k d_k^2 + R_1,
\end{aligned}$$

where $R_1 = R_1(\delta)$. Similarly we get from (2.16)

$$\sum_{\substack{1 \le i < k \le N \\ i/k < \delta}} d_i d_k \left(\frac{i}{k}\right)^{m/2} \le \frac{4}{m} \delta^{m/2} \sum_{1 \le k \le N} k d_k^2 + R_2 \tag{2.17}$$

and

$$\sum_{A \le i < k \le N} d_i d_k \left(\frac{i}{k}\right)^{m/2} \ge \frac{1}{m} \sum_{k=1}^{N} k d_k^2 - R_3, \tag{2.18}$$

where $R_2, R_3 > 0$ depend on δ. Now Lemma 2 gives

$$\begin{aligned}
S^{(3)} &\ge 2c \sum_{\substack{A \le i < k \le N \\ i/k \ge \delta}} d_i d_k \left(\frac{i}{k}\right)^{m/2} \\
&\ge 2c \left(\sum_{A \le i < k \le N} d_i d_k \left(\frac{i}{k}\right)^{m/2} - \sum_{\substack{1 \le i < k \le N \\ i/k < \delta}} d_i d_k \left(\frac{i}{k}\right)^{m/2} \right).
\end{aligned}$$

Note finally that $L(k) \gg 1$ implies $\sum_{k \ge 1} k d_k^2 = \infty$. Choosing ε small will also make δ small, and thus combining the estimate for $S^{(3)}$ with (2.17–2.18) and using the estimates for $S^{(1)}$, $S^{(2)}$ proves the lemma. □

Lemma 6 *Let* $\mathbf{D} = (d_k)$ *be a summation method with* $d_k = O(1)$. *If for some* $\alpha > 0$

$$d_N = O\left(\frac{D_N}{N(\log\log N)^\alpha}\right),$$

it follows that

$$\log \frac{D_N}{D_M} \ll (\log\log M)^{-\alpha} \log \frac{N}{M} \quad (M_0 \le M < N).$$

Proof From $D_k \to \infty$ and $d_k = O(1)$ we conclude that $D_{k+1}/D_k \to 1$ and

$$\log \frac{D_{k+1}}{D_k} \ll \frac{d_{k+1}}{D_k} \ll \frac{1}{k(\log\log k)^\alpha}.$$

Hence $\sum_{k=M}^{N-1} \log \frac{D_{k+1}}{D_k} \ll (\log\log M)^{-\alpha} \sum_{k=M}^{N-1} \frac{1}{k}$. □

3 Proofs

The method to prove our theorems is based on a blocking technique. We partition $\mathbb{N}$ into disjoint blocks: $\mathbb{N} = A_1 \cup B_1 \cup A_2 \cup B_2 \cup \cdots$, where

$$A_j = \{2^{p'_j} + 1, \ldots, 2^{q'_j}\} \quad \text{and} \quad B_j = \{2^{p_j} + 1, \ldots, 2^{q_j}\} \quad (j \ge 1).$$

We set $q_0 = 1$ and for some $r > 0$ (to be chosen later) define the exponents as

$$p'_j = q_{j-1}, \qquad q'_j = p'_j + [12 \log j], \qquad p_j = q'_j, \qquad q_j = p_j + [12(\log j)^{1+r}].$$

Obviously the length of the blocks B_j will grow much faster than the length of A_j and the block A_j precedes B_j on the real line. Set

$$M_j = 2^{p_j}, \qquad N_j = 2^{q_j}, \qquad M'_j = 2^{p'_j}, \qquad N'_j = 2^{q'_j} \quad (j \ge 1).$$

Recalling the definition of ξ_k and $\xi_{k,l}$ we define now

$$Z_j := \sum_{k \in B_j} d_k \xi_k, \qquad Z_j^* := \sum_{k \in B_j} d_k \xi_{N_{j-1},k} \quad (j \ge 1),$$

$$R_j := \sum_{k \in A_j} d_k \xi_k, \qquad R_j^* := \sum_{k \in A_j} d_k \xi_{N'_{j-1},k} \quad (j \ge 1).$$

To visualize this, the sum Z_j is

$$\sum_{k=M_j+1}^{N_j} d_k \left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right),$$

and Z_j^* is obtained from Z_j by replacing S_k in the previous sum by $S_k - S_{N_{j-1}}$, so that Z_j^* involves only X_k's with $N_{j-1} < k \leq N_j$. It follows that the Z_j^* are independent random variables which we will use as approximations for the original random variables Z_j. Similarly, the R_j^* are independent r.v.'s. We will call the Z_j and Z_j^* 'long block sums', R_j and R_j^* 'short block sums'. We stress again that the short block sum R_j precedes the long block sum Z_j on the real line. In a first step we derive limit theorems for the sequences (Z_j^*) and (R_j^*) via classical theorems for independent random variables. We will see that the contribution of the R_j and the error we make by replacing Z_j by Z_j^* will be small, so that our results carry over to the sequence $(Z_j + R_j)$. Observing that

$$\sum_{j=1}^{n}(Z_j + R_j) = \sum_{k=1}^{N_n} d_k\left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right) \tag{3.1}$$

shows that the desired results hold along the subsequence (N_j). In order to permit fast growing summation methods D_k and to provide that the limit relations hold along the subsequence (N_j), we had to choose "large" blocks. However, in a final step we have to control the fluctuation between the subsequence (N_j), which forces us to use "smaller" blocks. Hence in this stage of the proof the key for optimal results lies in an optimal choice of the block lengths.

In order to give upper bounds for the moments of Z_j, Z_j^*, R_j and R_j^* we introduce the notation

$$V_j = V_j(\mathbf{D}) := \sum_{k\in B_j} kd_k^2 \quad \text{and} \quad U_j = U_j(\mathbf{D}) := \sum_{k\in A_j} kd_k^2. \tag{3.2}$$

For $p \in \mathbb{N}$ we get immediately from Lemma 3 and Lemma 4

$$\begin{aligned} E|Z_j - Z_j^*|^p &\leq 2^{(q_{j-1}-p_j)/2} E_p V_j^{p/2}, \\ E|R_j - R_j^*|^p &\leq 2^{(q'_{j-1}-p'_j)/2} E_p U_j^{p/2}, \end{aligned} \tag{3.3}$$

and

$$E|Z_j|^p \leq C_p V_j^{p/2} \quad \text{and} \quad E|R_j|^p \leq C_p U_j^{p/2}. \tag{3.4}$$

Proof of Theorem 2 From Lyapunov's theorem (cf. Petrov [15, Theorem 4.9]) it follows that the sequence (Z_n^*) obeys the CLT if we show that

$$\frac{\sum_{j=1}^{n} E|Z_j^*|^4}{\mathrm{Var}^2(\sum_{j=1}^{n} Z_j^*)} \to 0. \tag{3.5}$$

Hence we need estimates for the expectation of the perturbed random variables $(\sum_{j=1}^{n} Z_j^*)^2$ and $\sum_{j=1}^{n}(Z_j^*)^4$. A look at (3.3) and (3.4) makes it evident that we have to study the order of magnitude of U_j and V_j given in (3.2). In order to handle this issue we use the representation of D_n in (1.10). It implies that

$$\sum_{k=m}^{n} kd_k^2 \asymp \int_m^n \left[\exp\left(2\int_A^x \varepsilon(u)/u\,du\right)\right]\frac{\varepsilon^2(x)}{x}\,dx \quad (m, n \to \infty). \tag{3.6}$$

(For numerical sequences (a_n) and (b_n) let $a_n \asymp b_n$ mean $a_n \ll b_n$ and $b_n \ll a_n$.) The monotonicity of the terms in the integrand of (3.6) and condition (1.9) yield

$$U_j \ll \int_{M_j'}^{N_j'} \left[\exp\left(2\int_A^x \varepsilon(u)/u\,du\right)\right]\frac{\varepsilon^2(x)}{x}\,dx \ll \varepsilon^2(M_j')D_{N_j'}^2 \log j,$$

where we used

$$\int_{M_j'}^{N_j'} 1/x\,dx \sim 12\log j.$$

Similarly we get

$$V_j \gg \varepsilon^2(N_j)D_{M_j}^2(\log j)^{1+r}.$$

A simple calculation shows that

$$p_j \sim q_j \sim p_j' \sim q_j' \sim 12j(\log j)^{1+r}, \tag{3.7}$$

and thus condition (1.9) implies that the ratio $\varepsilon(M_j')/\varepsilon(N_j)$ remains bounded for $j \to \infty$ and since $N_j' = M_j$ this shows that $U_j/V_j \to 0$. By condition (1.15) we clearly have $\sum_{j=1}^n V_j \to \infty$ and hence we obtain from the Minkowski inequality, Lemma 4, the independence of R_j^* and (3.3–3.4)

$$\begin{aligned}
&\left|\left\|\sum_{j=1}^n Z_j^*\right\|_2 - \left\|\sum_{j=1}^n (Z_j + R_j)\right\|_2\right| \\
&\quad \le \left\|\sum_{j=1}^n (Z_j^* - (Z_j + R_j))\right\|_2 \\
&\quad \le \left\|\sum_{j=1}^n R_j^*\right\|_2 + \sum_{j=1}^n \|Z_j - Z_j^*\|_2 + \sum_{j=1}^n \|R_j - R_j^*\|_2 \\
&\quad \ll \left(\sum_{j=1}^n U_j\right)^{1/2} + \sum_{j=1}^n j^{-3\log 2}V_j^{1/2} = o\left(\left(\sum_{j=1}^n V_j\right)^{1/2}\right).
\end{aligned}$$

(Here $\|\cdot\|_2$ denotes the L_2 norm.) We can now compare the variances λ_N defined in (1.17) to $\mathrm{Var}(\sum_{j=1}^n Z_j^*)$. Combining the latter estimate with Lemma 4 and Lemma 5 shows that

$$\mathrm{Var}\left(\sum_{j=1}^n Z_j^*\right) \sim \lambda_{N_n} \quad \text{and} \quad \lambda_{N_n} \asymp \sum_{k=1}^{N_n} kd_k^2 \asymp \sum_{j=1}^n V_j. \tag{3.8}$$

From (3.3) and (3.4) we get that

$$\sum_{j=1}^{n} E|Z_j^*|^4 \ll \sum_{j=1}^{n} V_j^2 \le \max\{V_1, \ldots, V_n\} \sum_{j=1}^{n} V_j \quad (n \to \infty).$$

On the other hand (3.8) gives $\mathrm{Var}(\sum_{j=1}^n Z_j^*) \gg \sum_{j=1}^n V_j$. Thus

$$\frac{\sum_{j=1}^n E|Z_j^*|^4}{\mathrm{Var}^2(\sum_{j=1}^n Z_j^*)} \ll \frac{\max_{1\le j\le n} V_j}{\sum_{j=1}^n V_j}. \tag{3.9}$$

From (3.6) we infer

$$\sum_{k=1}^{N} kd_k^2 \gg \varepsilon(N) D_N^2 \quad \text{and} \quad V_n \ll \varepsilon(M_n) D_{N_n}^2 \int_{M_n}^{N_n} \varepsilon(x)/x\, dx. \tag{3.10}$$

Using again (3.8), (3.10), (1.9), (1.10), the explicit formulas for M_n, N_n (see (3.7)) and (1.12) combined with Lemma 6 we derive

$$\begin{aligned} \frac{V_n}{\sum_{j=1}^n V_j} &\ll V_n \left(\sum_{k=1}^{N_n} kd_k^2 \right)^{-1} \ll \frac{\varepsilon(M_n)}{\varepsilon(N_n)} \int_{M_n}^{N_n} \varepsilon(x)/x\, dx \\ &\ll \int_{M_n}^{N_n} \varepsilon(x)/x\, dx \ll \log \frac{D_{N_n}}{D_{M_n}} \ll (\log n)^{(1+r-\alpha)}. \end{aligned}$$

Since $\alpha > 1$ in (1.12), we can chose $0 < r < \alpha - 1$, which shows

$$V_n/(V_1 + \cdots + V_n) \to 0 \quad (n \to \infty). \tag{3.11}$$

Together with (3.9) this proves (3.5), i.e. the central limit theorem holds for the random variables (Z_j^*).

In the next step we show that $\sum_{j=1}^n (Z_j + R_j)$ and $\sum_{j=1}^n Z_j^*$ are "close" to each other. This shows that the CLT is valid for the sequence $(Z_j + R_j)$ as well. Observe that by (3.3)

$$\sum_{j=1}^{\infty} \left(\frac{E(Z_j - Z_j^*)^2}{\sum_{1\le l\le j} V_l} \right)^{1/2} < \infty,$$

and consequently

$$\sum_{j=1}^{\infty} \frac{|Z_j - Z_j^*|}{(\sum_{1\le l\le j} V_l)^{1/2}} < \infty \quad \text{a.s.}$$

Thus by the Kronecker lemma and (3.8) we have

$$\frac{1}{\lambda_{N_n}^{1/2}} \sum_{j=1}^{n} (Z_j - Z_j^*) \to 0 \quad \text{a.s.} \tag{3.12}$$

By the same arguments it follows easily that the last relation holds with $(R_j - R_j^*)$ instead of $(Z_j - Z_j^*)$. Further (3.3–3.4), the independence of R_j^* and the Cauchy-Schwarz inequality give

$$E\left|\sum_{j=1}^{n} R_j^*\right| \ll \left(\sum_{j=1}^{n} U_j\right)^{1/2}.$$

We have already shown that $U_j/V_j \to 0$. Hence the Markov inequality and (3.8) yield

$$P\left(\left|\sum_{j=1}^{n} R_j^*\right| > \varepsilon \lambda_{N_n}^{1/2}\right) = o(1) \quad \text{for every } \varepsilon > 0.$$

Using $Z_j + R_j = Z_j^* + R_j^* + (Z_j - Z_j^*) + (R_j - R_j^*)$ and recalling (3.1) we proved that

$$\lambda_{N_n}^{-1/2} \sum_{k=1}^{N_n} d_k \xi_k \xrightarrow{d} \mathcal{N}.$$

In order to finish the proof of Theorem 2 it suffices now to show that

$$\lim_j \sup_{N \in (N_{j-1}, N_j]} E|T_N - T_{N_j}| = 0, \tag{3.13}$$

where

$$T_N = \lambda_N^{-1/2} \sum_{k=1}^{N} d_k \xi_k.$$

By Minkowski's inequality and Lemma 4 we get for $N_{j-1} \le N \le N_j$

$$|\lambda_{N_j}^{1/2} - \lambda_N^{1/2}| \le \mathrm{Var}^{1/2}\left(\sum_{k=N+1}^{N_j} d_k \xi_k\right) \le (4\gamma)^2 (U_j + V_j)^{1/2}. \tag{3.14}$$

Therefore we have

$$\begin{aligned} E|T_N - T_{N_j}| &\le |\lambda_N^{-1/2} - \lambda_{N_j}^{-1/2}| E\left|\sum_{k=1}^{N} d_k \xi_k\right| + \lambda_{N_j}^{-1/2} E\left|\sum_{k=N+1}^{N_j} d_k \xi_k\right| \\ &\le \lambda_{N_j}^{-1/2}\left(|\lambda_{N_j}^{1/2} - \lambda_N^{1/2}| + \mathrm{Var}^{1/2}\left(\sum_{k=N+1}^{N_j} d_k \xi_k\right)\right) \\ &\le 2(4\gamma)^2 \lambda_{N_j}^{-1/2} (U_j + V_j)^{1/2}. \end{aligned}$$

Applying (3.8) and (3.11) we can show (3.13). This finishes the proof of Theorem 2. □

Proof of Theorem 3 Trivially $|Z_j^*| \le 2(D_{N_j} - D_{M_j})$. In order to apply Kolmogorov's law of the iterated logarithm (cf. [15, p. 239]) to the sequence (Z_j^*), it suffices to verify that

$$D_{N_j} - D_{M_j} = o\left(\left(\frac{s_j^2}{\log\log s_j^2}\right)^{1/2}\right),$$

where $s_j^2 = \mathrm{Var}(\sum_{l=1}^j Z_l^*)$. First note that by the representation (1.10)

$$\begin{aligned} D_{N_j} - D_{M_j} &= D_{N_j}\left(1 - \exp\left(-\int_{M_j}^{N_j} \epsilon(u)/u\,du\right)\right) \\ &\le D_{N_j} \int_{M_j}^{N_j} \epsilon(u)/u\,du. \end{aligned} \tag{3.15}$$

Relation (3.8) and the first statement of (3.10) show that

$$D_{N_j}^2 \varepsilon(N_j) \ll s_j^2 \ll D_{N_j}^2 \ll N_j. \tag{3.16}$$

Using (1.9), (3.15), (3.16) and the explicit formulas for M_n, N_n we derive

$$\begin{aligned} &(D_{N_j} - D_{M_j})\left(\frac{s_j^2}{\log\log s_j^2}\right)^{-1/2} \\ &\quad \ll \frac{D_{N_j} - D_{M_j}}{D_{N_j}}\left(\frac{\log\log N_j}{\varepsilon(N_j)}\right)^{1/2} \ll \left(\frac{\log\log N_j}{\varepsilon(N_j)}\right)^{1/2} \int_{M_j}^{N_j} \varepsilon(u)/u\,du \\ &\quad \ll (\log\log N_j)^{1/2} \int_{M_j}^{N_j} \varepsilon^{1/2}(u)/u\,du. \end{aligned} \tag{3.17}$$

From Lemma 6, (1.10) and the Cauchy-Schwarz inequality it follows that

$$\begin{aligned} \int_{M_j}^{N_j} \varepsilon^{1/2}(u)/u\,du &\le \left(\int_{M_j}^{N_j} \varepsilon(u)/u\,du\right)^{1/2} \left(\int_{M_j}^{N_j} 1/u\,du\right)^{1/2} \\ &\ll \left(\log \frac{D_{N_j}}{D_{M_j}}\right)^{1/2} \left(\log \frac{N_j}{M_j}\right)^{1/2} \\ &\ll \left(\log \frac{N_j}{M_j}\right)(\log\log M_j)^{-\alpha/2}. \end{aligned} \tag{3.18}$$

By the definition of M_j and N_j we see that the last expression of (3.17) is bounded by const$\cdot(\log j)^{3/2-\alpha/2+r}$. This tends to zero if we choose $0 < r < (\alpha - 3)/2$. Remember that by (3.8) $\lambda_{N_j} \sim s_j^2$. Thus setting $L_N = (2\lambda_N \log\log \lambda_N)^{1/2}$ we get from Kolmogorov's law of the iterated logarithm the relation

$$\limsup_{j\to\infty} L_{N_j}^{-1} \sum_{l=1}^{j} Z_l^* = 1 \quad \text{a.s.}$$

and by (3.12) this holds also if Z_l^* is replaced by Z_l:

$$\limsup_{j\to\infty} L_{N_j}^{-1}\sum_{l=1}^{j} Z_l = 1 \quad \text{a.s.}$$

Similar arguments show that the short block sums are negligible, i.e.

$$L_{N_j}^{-1}\sum_{l=1}^{j} R_l \to 0 \quad \text{a.s.}$$

Hence by (3.1) the LIL (1.18) is true along the subsequence (N_j).

Finally we shall show that the maximal fluctuation between the subsequence (N_j) is small enough for our purposes. For $N_{j-1} \le N < N_j$ a trivial estimate (using $|\xi_k| \le 2$) shows

$$\begin{aligned}
&\left| L_N^{-1}\sum_{k=1}^{N} d_k\xi_k - L_{N_{j-1}}^{-1}\sum_{k=1}^{N_{j-1}} d_k\xi_k \right| \\
&\quad \le \left| \frac{L_{N_{j-1}}}{L_N} - 1 \right| L_{N_{j-1}}^{-1}\left| \sum_{k=1}^{N_{j-1}} d_k\xi_k \right| + 2L_N^{-1}(D_{N_j} - D_{N_{j-1}}).
\end{aligned}$$

We have already proved that $L_{N_{j-1}}^{-1}|\sum_{k=1}^{N_{j-1}} d_k\xi_k| = O(1)$ a.s. Hence Theorem 3 will follow from

$$\lim_{j\to\infty} \sup_{N_{j-1}\le N<N_j} \left| \frac{L_{N_{j-1}}}{L_N} - 1 \right| = 0 \tag{3.19}$$

and

$$\limsup_{j\to\infty} \sup_{N_{j-1}\le N<N_j} L_N^{-1}(D_{N_j} - D_{N_{j-1}}) = 0. \tag{3.20}$$

Relations (3.8), (3.11) and (3.14) imply that

$$\sup_{N_{j-1}\le N<N_j} |\lambda_{N_j}^{1/2} - \lambda_N^{1/2}| = o(\lambda_{N_j}^{1/2}) \quad \text{as } j \to \infty, \tag{3.21}$$

which immediately yields (3.19). To show (3.20) we first observe that by (3.8), (3.10), (3.21) and (1.9) we have for $N_{j-1} \le N < N_j$

$$L_N^{-1}(D_{N_j} - D_{N_{j-1}}) \le \lambda_N^{-1/2}(D_{N_j} - D_{N_{j-1}}) \ll D_{N_{j-1}}^{-1}\varepsilon^{-1/2}(N_j)(D_{N_j} - D_{N_{j-1}}).$$

Using the same argument as in (3.15) we get an upper bound for $D_{N_j} - D_{N_{j-1}}$. Thus by (1.9) and observing that $N_j \le N_{j-1}^2$ for large enough j we obtain

$$L_N^{-1}(D_{N_j} - D_{N_{j-1}}) \ll \frac{\varepsilon^{-1/2}(N_j)D_{N_j}}{D_{N_{j-1}}} \int_{N_{j-1}}^{N_j} \varepsilon(u)/u\,du$$
$$\ll \frac{D_{N_j}}{D_{N_{j-1}}} \int_{N_{j-1}}^{N_j} \varepsilon^{1/2}(u)/u\,du.$$

The same arguments as in (3.18), Lemma 6 and (3.7) give

$$\begin{aligned} &L_N^{-1}(D_{N_j} - D_{N_{j-1}}) \\ &\quad \ll \frac{D_{N_j}}{D_{N_{j-1}}}\left(\log \frac{N_j}{N_{j-1}}\right)(\log\log N_{j-1})^{-\alpha/2} \\ &\quad \ll \left(\frac{N_j}{N_{j-1}}\right)^{(\log\log N_{j-1})^{-\alpha}}\left(\log \frac{N_j}{N_{j-1}}\right)(\log\log N_{j-1})^{-\alpha/2} \\ &\quad \ll \exp(\mathrm{const}\cdot(\log j)^{1+r-\alpha})(\log j)^{1+r-\alpha/2}. \end{aligned}$$

This tends to zero if r is chosen in $(0, \alpha/2-1)$. □

Proof of Theorem 1 By Lemma 4 we have (actually under weaker conditions than we assumed in Theorem 1)

$$\lambda_N \ll \sum_{k=1}^N k d_k^2.$$

The assumption (kd_k) is non-decreasing implies

$$\sum_{k=1}^N k d_k^2 \le N d_N D_N, \quad N = 1, 2, \ldots.$$

Hence by relation (1.12) we conclude that

$$(\lambda_N \log\log \lambda_N)^{1/2} \ll (N d_N D_N \log\log N)^{1/2}$$
$$\ll D_N(\log\log N)^{(1-\alpha)/2},$$

which means for $\alpha > 1$ $(\lambda_N \log\log \lambda_N)^{1/2} = o(D_N)$, and thus for $f \in \mathcal{L}$ Theorem 1 follows from Theorem 3. The proof for indicator functions follows from routine approximation arguments as e.g. those in [11]. □

Proof of Theorem 5 We will show that

$$\mathrm{Var}\left(\frac{1}{D_N}\sum_{k=1}^N d_k\left(f\left(\frac{S_k}{\sqrt{k}}\right) - Ef\left(\frac{S_k}{\sqrt{k}}\right)\right)\right) = \lambda_N/D_N^2 \to 0.$$

By (1.20) we have for every $\varepsilon > 0$ an $N_0(\varepsilon)$, such that for all $N \geq N_0$

$$\sup_{k \leq N} \frac{kd_k}{D_N} < \varepsilon. \tag{3.22}$$

By (2.6) and (3.22) we have for $N \geq N_0$

$$\lambda_N/D_N^2 \ll \frac{1}{D_N} \sum_{1 \leq l \leq N} d_l l^{-1/2} \sum_{1 \leq k \leq l} \frac{kd_k}{D_N} k^{-1/2} \leq 2\varepsilon.$$

□

Consider the summation $D_N = \log N$. By Lemma 4 and Lemma 5 we infer $\lambda_N \asymp D_N$. On the other hand if $D_N = N$, then it's possible to show (using Lemma 1 and Lemma 2) that $\lambda_N \asymp D_N^2$. Trivially, a necessary condition for the LIL is $\lambda_N \log\log \lambda_N \leq D_N^2$. If the numerical value α of the weight sequence (D_N) defined in (1.19) falls below the critical value 1 the last condition is violated. The proof of Theorem 4 is based on this simple observation.

Proof of Theorem 4 Let

$$D_N = \exp(\log N/(\log\log N)^\alpha), \quad \text{where } 0 < \alpha < 1. \tag{3.23}$$

A simple calculation shows that the representation (1.10) holds with $c = 1$ and

$$\varepsilon(x) = (\log\log x)^{-\alpha} - \alpha(\log\log x)^{-\alpha-1} \sim (\log\log x)^{-\alpha} \tag{3.24}$$

and thus this $\mathbf{D}$ is in $\mathcal{W}$. Also, by (3.23) and the mean value theorem

$$d_N = D_N - D_{N-1} \sim \frac{1}{N} \exp(\log N/(\log\log N)^\alpha)(\log\log N)^{-\alpha}$$

and thus relations (1.12) and (1.13) are valid. Finally, Lemma 5, (3.10) and (3.24) yield

$$\lambda_N \gg \sum_{k=1}^N kd_k^2 \gg \varepsilon(N) D_N^2 \gg D_N^2 (\log\log N)^{-\alpha}$$

whence

$$(\lambda_N \log\log \lambda_N)^{1/2} \gg D_N(\log\log N)^{(1-\alpha)/2}.$$

Since $|\sum_{k=1}^N d_k \xi_k| \leq 2D_N$, relation (1.18) cannot hold if $\alpha < 1$. □

Acknowledgements I am indebted to Prof. István Berkes for his great support and for inducting me into this topic. Also I would like to thank Peter Becker-Kern for valuable discussions and several comments on my manuscript.

References

1. Atlagh, M., Weber, M.: Le théorème central limite presque sûr. Expo. Math. **18**, 97–126 (2000)

2. Becker-Kern, P.: Almost sure limit theorems of mantissa type for semistable domains of attraction. Acta Math. Hung. **114**, 301–336 (2007)
3. Berkes, I.: 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)
4. Berkes, I., Csáki, E.: A universal result in almost sure central limit theory. Stoch. Process. Appl. **94**, 105–134 (2001)
5. Berkes, I., Horváth, L.: Almost sure invariance principles for logarithmic averages. Studia Sci. Math. Hung. **33**, 1–24 (1997)
6. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation. Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge (1987)
7. Brosamler, G.A.: An almost everywhere central limit theorem. Math. Proc. Camb. Philos. Soc. **104**, 561–574 (1988)
8. Chandrasekharan, K., Minakshisundaram, S.: Typical Means. Oxford University Press (1952)
9. Fisher, A.: A pathwise central limit theorem for random walks. Preprint (1989)
10. Hörmann, S.: An extension of almost sure central limit theory. Stat. Probab. Lett. **76**, 191–202 (2006)
11. Lacey, M.T., Philipp, W.: A note on the almost sure central limit theorem. Stat. Probab. Lett. **9**, 201–205 (1990)
12. Lévy, P.: Théorie de l'addition des variables alétoires, 12th edn. Jacques Gabay (2003)
13. Major, P.: An improvement of Strassen's invariance principle. Ann. Probab. **7**, 55–61 (1979)
14. Peligrad, M., Révész, P.: On the almost sure central limit theorem. In: Almost Everywhere Convergence, II (Evanston, IL, 1989), pp. 209–225. Academic, Boston (1991)
15. Petrov, V.: Limit Theorems of Probability Theory. Oxford Science Publications, Oxford (1995)
16. Rozanov, Y.A.: Stationary Random Processes. Holden-Day, San Francisco (1967)
17. Schatte, P.: On strong versions of the central limit theorem. Math. Nachr. **137**, 249–256 (1988)
18. Takahashi, S.: On the law of the iterated logarithm for lacunary trigonometric series. Tohoku Math. J. **24**, 319–329 (1972)
19. Weber, M.: Un théorème central limite presque sûr à moments généralisés pour les rotations irrationnelles. Manuscripta Math. **101**, 175–190 (2000)