On the Extremal Theory of Continued Fractions

Alina Bazarova · István Berkes · Lajos Horváth

Received: 28 November 2013 / Revised: 8 September 2014 / Published online: 10 October 2014
© Springer Science+Business Media New York 2014

Abstract Letting $x = [a_1(x), a_2(x), \ldots]$ denote the continued fraction expansion of an irrational number $x \in (0, 1)$, Khinchin proved that $S_n(x) = \sum_{k=1}^{n} a_k(x) \sim \frac{1}{\log 2} n \log n$ in measure, but not for almost every $x$. Diamond and Vaaler showed that, removing the largest term from $S_n(x)$, the previous asymptotics will hold almost everywhere, this shows the crucial influence of the extreme terms of $S_n(x)$ on the sum. In this paper we determine, for $d_n \to \infty$ and $d_n/n \to 0$, the precise asymptotics of the sum of the $d_n$ largest terms of $S_n(x)$ and show that the sum of the remaining terms has an asymptotically Gaussian distribution.

Keywords Continued fraction expansion · Extreme elements · Mixing random variables · Central limit theorem

Mathematics Subject Classification (2010) Primary 11K50 · 60F05 · 60G70

Alina Bazarova received research support from Austrian Science Fund (FWF) Grant W1230.
István Berkes received research support from Austrian Science Fund (FWF) Grant P24302-N18 and from OTKA Grant K 108615.
Lajos Horváth received research support from NSF Grant DMS-13-05858.

A. Bazarova · I. Berkes (✉)
Institute of Statistics, Graz University of Technology, Kopernikusgasse 24, 8010 Graz, Austria
e-mail: berkes@tugraz.at
A. Bazarova
e-mail: albazarova@gmail.com

L. Horváth
Department of Mathematics, University of Utah, Salt Lake City, UT 84112-0090, USA
e-mail: horvath@math.utah.edu
1 Introduction

For an irrational number \( x \in (0, 1) \) let

\[
x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\ddots}}}.
\]

be the continued fraction expansion of \( x \). Clearly,

\[ a_1(x) = \lfloor 1/x \rfloor, \quad a_{n+1}(x) = a_1(T^n x), \quad n \geq 1, \]

where the transformation \( T \) is defined by \( Tx = \{1/x\}, x \in (0, 1) \setminus \mathbb{Q} \); here \( \lfloor \cdot \rfloor \) and \( \{\cdot\} \) denote integral resp. fractional part, and \( \mathbb{Q} \) denotes the set of rational numbers. Let

\[
\mu(E) = \frac{1}{\log 2} \int_E \frac{1}{1 + x} dx
\]

be the Gauss measure on the class \( B \) of Borel subsets of \( (0, 1) \). It is known (see, e.g., [3]) that \( T \) is an ergodic transformation preserving the Gauss measure and thus, with respect to the probability space \(( (0, 1), B, \mu ) \), \( \{a_n(x), n \geq 1\} \) is a stationary ergodic sequence. Clearly, the set \( \{a_1 = k\} \) is the interval \((1/(k + 1), 1/k)\), and thus

\[
\mu\{a_1 = k\} = \frac{1}{\log 2} \int_{1/(k+1)}^{1/k} \frac{1}{1 + x} dx = \frac{1}{\log 2} \log \left\{ 1 + \frac{1}{k(k + 2)} \right\} \sim \frac{1}{\log 2} \frac{1}{k^2}.
\]

(We say that \( a_k \sim b_k \) if \( \lim_{k \to \infty} a_k/b_k = 1 \).) Thus, by the ergodic theorem, we have for any function \( F : \mathbb{N} \to \mathbb{R} \)

\[
\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} F(a_k(x)) = \frac{1}{\log 2} \sum_{j=1}^{\infty} F(j) \log \left\{ 1 + \frac{1}{j(j + 2)} \right\} \quad \text{a.e.} \quad (1.1)
\]

provided that the series on the right-hand side converges absolutely.

The sequence \( \{a_k(x), k \geq 1\} \) has remarkable mixing properties. Gauss noted that the distribution of \( T^n x = [a_{n+1}(x), a_{n+2}(x), \ldots] \) with respect to the uniform measure in \((0, 1)\) converges to \( \mu \) and asked for the speed of convergence. (For a discussion, see [3], pp. 49–50 or [17], p. 552.) Kusmin [19] showed that the convergence speed is \( O(e^{-\lambda \sqrt{k}}) \), and Lévy [21] improved this to \( O(e^{-\lambda k}) \). Lévy’s result implies that the sequence \( \{a_k(x), k \geq 1\} \) is \( \psi \)-mixing with exponential rate, i.e., for all \( A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+n}^\infty, k \geq 1, n \geq 1 \), we have

\[
|\mu(A \cap B) - \mu(A)\mu(B)| \leq \psi(n)\mu(A)\mu(B) \quad (1.2)
\]
where $\psi(n) = C e^{-\lambda n}$ with positive absolute constants $C, \lambda$, and $\mathcal{F}_r^s$ denotes the $\sigma$-field generated by the variables $\{a_k(x), \ r \leq k \leq s\}$.

Letting $E$ denote expectation with respect to $\mu$, we have $Ea_1 = \infty$ and correspondingly for $F(x) = x$ the right-hand side of (1.1) is $+\infty$. Thus the partial sums $\sum_{k=1}^{N} a_k(x)$ grow faster than $N$. Lévy [22] proved that

$$\frac{1}{N} \sum_{k=1}^{N} a_k(x) - \frac{\log N}{\log 2} \xrightarrow{d} G,$$

where $\xrightarrow{d}$ means convergence in distribution in the probability space $((0, 1), B, \mu)$, and $G$ is a stable distribution with characteristic function

$$\exp \left(-\frac{\pi |t|}{2 \log 2} - \frac{it \log |t|}{2 \log 2} - \frac{i \kappa t}{2 \log 2} \right),$$

where $\kappa = 0.577 \ldots$ is the Euler–Mascheroni constant. See also Theorem 2, pp. 159–160 of Heinrich [14], where a remainder term estimate for the convergence in (1.3) is obtained. This implies that

$$\lim_{N \to \infty} \frac{1}{N \log N} \sum_{k=1}^{N} a_k(x) = \frac{1}{\log 2} \quad \text{in measure},$$

a result obtained earlier by Khinchin [18]. Khinchin also noted that (1.5) cannot hold almost everywhere. Diamond and Vaaler [9] showed that the obstacle to a.e. convergence in (1.5) is the occurrence of one single large term in the sum $\sum_{k=1}^{N} a_k(x)$ and established an a.e. analogue of (1.5) by excluding the largest summand. They proved namely

$$\lim_{N \to \infty} \frac{1}{N \log N} S_N^{(1)}(x) = \frac{1}{\log 2} \quad \text{for almost all } x$$

where $S_N^{(d)}(x)$ denotes the sum $\sum_{k=1}^{N} a_k(x)$ after discarding its $d$ largest summands. The proof shows that (1.6) remains valid if $S_N^{(1)}$ is replaced by $S_N^{(d)}$ for any fixed $d \geq 2$ and discarding more terms improves the rate of a.e. convergence in (1.6). An analogous result for the St. Petersburg game was proved by Csörgő and Simons [7]. For further analogies between continued fraction digits and the St. Petersburg game, we refer to Vardi [32]. In view of these facts, it is natural to ask what happens if from the sum $S_N = \sum_{k=1}^{N} a_k(x)$ we remove $d = d_N$ terms, where

$$d_N \to \infty, \quad d_N/N \to 0$$

so that the number of discarded terms is ‘large’, but is still negligible compared with $N$. The purpose of this paper was to answer this question. Let
m(t) = \frac{1}{\log 2} \sum_{1 \leq k \leq t} k \log \left(1 + \frac{1}{k(k+2)}\right), \quad t \geq 1. \quad (1.8)

We will prove the following result.

**Theorem 1.1** Let \( d = d_N \) satisfy (1.7). Then we have

\[
\frac{S_N^{(d)} - N m(\eta_{d,N})}{N / \sqrt{d}} \xrightarrow{d} N\left(0, (\log 2)^{-1}\right)
\]  

(1.9)

where \( \eta_{d,N} \) denotes the \( d \)-th largest of \( a_1, \ldots, a_N \) and \( N(\mu, \sigma^2) \) denotes the normal distribution with mean \( \mu \) and variance \( \sigma^2 \).

Theorem 1.1 reduces the asymptotic study of \( S_N^{(d)} \) to that of \( \eta_{d,N} \), which is a much simpler problem. We will show in (3.15) that \( \eta_{d,N} \sim N/d \) in probability and since \( m(t) \sim (\log 2)^{-1} \log t \) as \( t \to \infty \), Theorem 1.1 can be rewritten equivalently as

\[
S_N^{(d)} = N m(\eta_{d,N}) + (N / \sqrt{d}) \xi_N
= (1 + o_P(1)) \frac{1}{\log 2} N \log(N/d) + (N / \sqrt{d}) \xi_N,
\]  

(1.10)

where \( \xi_N \xrightarrow{d} N(0, 1 / \log 2) \). Here and in the sequel, \( p \) will denote convergence in probability and \( o_P(1) \) a quantity converging to 0 in probability. Relation (1.10) shows that \( N m(\eta_{d,N}) \) is the main term in an asymptotic expansion of \( S_N^{(d)} \). As a comparison, write Lévy’s limit theorem (1.3) in the form

\[
S_N = \frac{1}{\log 2} N \log N + N \xi^*,
\]  

(1.11)

where \( \xi^*_N \) converges in distribution to the Cauchy variable with characteristic function (1.4). In addition to the change of the order of magnitude of \( S_N \) caused by removing the \( d \) largest terms, note that the Cauchy fluctuations of \( S_N \) around \( \frac{1}{\log 2} N \log N \) described by (1.11) changed to Gaussian fluctuations around \( N m(\eta_{d,N}) \) in (1.10). An immediate consequence of relation (1.10) is

\[
\frac{S_N^{(d)}}{N \log(N/d)} \xrightarrow{p} \frac{1}{\log 2}
\]

under (1.7). If \( d \) grows slower than any power of \( N \), i.e., \( \log d / \log N \to 0 \), then the last relation implies

\[
\frac{1}{N \log N} \frac{S_N^{(d)}}{N} \xrightarrow{p} \frac{1}{\log 2}.
\]
Thus, in this case, the order of magnitude of $S_N^{(d)}$ is the same as that of the complete sum $S_N$, i.e., the contribution of the $d$ largest terms of $S_N$ is still negligible compared with the whole sum. If $d \sim N^\gamma$ for some $0 < \gamma < 1$, then

$$\frac{1}{N \log N} S_N^{(d)} \xrightarrow{P} \frac{1 - \gamma}{\log 2}.$$ 

We thus see that the removal of a small portion of extreme elements of $S_N$ changes the asymptotic order of magnitude of the sum; hence the role of large elements in $S_N$ is very substantial.

In case of i.i.d. variables in the domain of attraction of a stable law with parameter $0 < \alpha < 2$, the effect of the extremal terms on the partial sums is well known. For positive variables, Darling [8] showed (see also Arov and Bobrov [1]) that under some additional regularity assumptions, the ratio of the sum and its largest term has a non-degenerate limit distribution if $0 < \alpha < 1$, and this holds also for $1 < \alpha < 2$ provided we center the partial sum by its mean. The case $\alpha = 1$ is critical and is not covered in [1,8]. The sequence $\{a_k(x), k \geq 1\}$ in the continued fraction expansion corresponds to this case, except that the variables $a_k$ are weakly dependent. Theorem 1.1 and its corollaries above show that the contribution of the $d$ largest terms of $S_N$ is negligible (in probability) compared with the total sum $S_N$ if and only if $\log d / \log N \to 0$. In particular, this holds for $d = 1$, i.e., in the case of the largest term. In the i.i.d. case, Csörgö et al. [6] also showed that removing the $d$ largest and $d$ smallest elements from the partial sum, where (1.7) holds, the remaining sum $S_N^{(d)}$ becomes asymptotically normal. Our Theorem 1.1 is a dependent analogue of this result for continued fractions.

There is a large literature on the metric properties of continued fractions and using the exponential $\psi$-mixing property of the transformation $T$ above, many classical limit theorems for partial sums of independent random variables have been extended to continued fractions. We refer to Doeblin [10], Gordin and Reznik [13], Ibragimov [15], Iosifescu [16, 17], Philipp [23, 25], Philipp and Stockelberg [26], Samur [27, 28], Stackelberg [29], Szewczak [30] and the references therein. Using the extremal theory of dependent processes, (see, e.g., Leadbetter and Rootzen [20]), asymptotic properties of the (individual) extremes of $(a_1(x), \ldots, a_n(x))$ can be established; limit theorems for the largest digit were obtained by Galambos [11,12], Philipp [24]. Note that an analogue of Theorem 1.1 for a different, less natural trimming of the partial quotients $a_j$ was obtained in Philipp [25].

In Sect. 2, we will prove Theorem 1.1 in a probabilistic form and we will change the notation accordingly.

**Theorem 1.2** Let $\{X_j, j \geq 1\}$ be a strictly stationary sequence of positive, integer valued random variables with

$$P(X_1 = k) \sim c_0 k^{-2} \quad \text{as} \quad k \to \infty \quad (1.12)$$

for some constant $c_0 > 0$. Assume that $\{X_j, j \geq 1\}$ satisfies the $\psi$-mixing condition (1.2) with $\psi(n) \leq Ce^{-\lambda n}$ for some $C > 0, \lambda > 0$. Let $\eta_{d,n}$ denote the $d$-th largest of $X_1, \ldots, X_n$ and assume that $d = d_n$ satisfies (1.7). Let $m(t) = E X_1 I\{X_1 \leq t\}$ and
Then
\[ \frac{1}{A_n} \sum_{i=1}^{[nt]} \left( X_i I\{X_i \leq \eta_{d,n}\} - m(\eta_{d,n}) \right) \xrightarrow{D[0,1]} W(t), \] (1.14)

where \( W \) is the Wiener process.

**Remark 1.1** If \((X_n)\) is a sequence of positive random variables such that with probability one \(X_1, X_2, \ldots \) are different, then the sum \( \sum_{i=1}^{[nt]} X_i I\{X_i \leq \eta_{d,n}\} \) in (1.14) is obtained from \( \sum_{i=1}^{[nt]} X_i \) by removing the \( d-1 \) largest terms, and thus, the conclusion of Theorem 1.2 for \( t = 1 \) reduces to that of Theorem 1.1. However, for integer valued variables \( X_n, \eta_{d,n} \) can appear in the sequence \((X_1, \ldots, X_n)\) more than once and in this case the number of terms of the sum \( \sum_{i=1}^{[nt]} X_i \) exceeding \( \eta_{d,n} \) can be smaller than \( d-1 \) and can actually be random. Thus, in a formal sense, Theorem 1.1 is not a special case of Theorem 1.2. However, using a simple perturbation argument, Theorem 1.1 will be deduced from Theorem 1.2.

Let
\[ U_n(t, s) = \sum_{i=1}^{[nt]} (X_i I\{X_i \leq s(n/d)\} - EX_i I\{X_i \leq s(n/d)\}) \quad (t \geq 0, s \geq 0). \]

We will derive Theorem 1.2 from the following two-dimensional limit theorem.

**Theorem 1.3** Under the assumptions of Theorem 1.2, we have
\[ \frac{1}{A_n} U_n(t, s) \rightsquigarrow W(t, s) \text{ weakly in } D([0, 1] \times [1/2, 3/2]), \] (1.15)

where \( \{W(t, s), \ t \geq 0, s \geq 0\} \) is a two-parameter Wiener process.

As we already noted, under (1.7) we have
\[ \frac{\eta_{d,n}}{n/d} \rightarrow 1 \quad \text{in probability}. \]

Since the limit process \( W(t, s) \) in (1.15) has continuous trajectories a.s., Theorem 1.3 and Billingsley [4], pp. 144–145 imply that
\[ \frac{1}{A_n} U_n(t, \eta_{d,n}/(n/d)) \xrightarrow{D[0,1]} W(t, 1) \]

which is exactly the functional central limit theorem (1.14).

In conclusion we note that Theorem 1.2 and Theorem 1.3 remain valid assuming a suitable polynomial \( \psi \)-mixing rate instead of the exponential rate. However, as this requires extensive changes in the arguments and we do not know of any practically interesting examples for \( \psi \)-mixing sequences with polynomial rate, we omit the details.
2 Some Lemmas

In the rest of the paper, \((X_k)\) denotes a sequence of random variables satisfying the conditions of Theorem 1.2 and \(d = d_n\) denotes a sequence of positive integers satisfying (1.7). Moreover, \(c_0\) denotes the constant in (1.12). Given a process \(Y(s, t)\) defined on a rectangle \(H = [a, b] \times [a', b']\), let \(Y(H)\) denote the increment of \(Y\) over \(H\).

**Lemma 2.1** Let \(\{Y_n(t, s), n \geq 1\}\) be processes defined on a rectangle \([a, b] \times [a', b'] \subset [0, \infty)^2\) and assume that for some \(\gamma > 0\)
\[E|Y_n(B)|^\gamma |Y_n(C)|^\gamma \leq \mu(B)\mu(C), \tag{2.1}\]
where \(\mu\) denotes area and \(B\) and \(C\) are rectangles of the form \([t_1, t_2] \times [s_1, s_2]\) having one common edge, but otherwise disjoint. Then the sequence \(\{Y_n(t, s), n \geq 1\}\) is tight.

If every \(X_n(t, s)\) is piecewise constant in \(t\), i.e., there exists a finite set \(H_n \subset [a, b]\) such that \(X_n(t, s)\) is constant on the left closed intervals determined by the elements of \(H_n \cup \{a\} \cup \{b\}\), then it suffices to verify (2.1) for rectangles \([t_1, t_2] \times [s_1, s_2]\) where \(t_1, t_2 \in H_n\).

This is a special case of a general tightness condition in Bickel and Wichura [2].

**Lemma 2.2** Let \(X, Y\) be integrable random variables such that \(X\) is measurable with respect to \(\sigma(X_1, \ldots, X_k)\) and \(Y\) is measurable with respect to \(\sigma(X_k+n, X_{k+n+1}, \ldots)\). Then \(XY\) is also integrable and
\[|E[XY] - EXEY| \leq \psi(n)E|X|E|Y|,\]
This follows from Theorem 3.10 in Bradley [5], p. 75.

**Lemma 2.3** Let \(G_k\) denote the \(\sigma\)-field generated by \(X_k\), let \(n_1 < \cdots < n_r\) be positive integers and let \(Y_1, \ldots, Y_r\) be bounded r.v.'s such that \(Y_j\) is \(G_{n_j}\) measurable \((j = 1, 2, \ldots, r)\). Then
\[|E[Y_1 \cdots Y_r]| \leq C_r E|Y_1| \cdots E|Y_r|,\]
where \(C_r = (1 + \psi(1))^r\).

**Proof** This is immediate by induction upon observing that by the previous lemma we have for any \(1 \leq j \leq r - 1\)
\[E|Y_1 \cdots Y_{j+1}| \leq E|Y_1 \cdots Y_j| E|Y_{j+1}| + \psi(1) E|Y_1 \cdots Y_j| E|Y_{j+1}| = (1 + \psi(1)) E|Y_1 \cdots Y_j| E|Y_{j+1}|.\]

**Lemma 2.4** For any \(T \geq 3\) we have
\[EX_1 I\{X_1 \leq T\} \leq C_1 \log T, \quad EX_1^4 I\{X_1 \leq T\} \leq C_1 T^3. \tag{2.2}\]
Moreover, for any fixed $0 \leq s_1 < s_2$, we have
\[
\mathbb{E} X_1^2 I\{s_1(n/d) < X_1 \leq s_2(n/d)\} \sim c_0(s_2 - s_1)(n/d) \quad \text{as } n \to \infty \tag{2.3}
\]
and for any fixed $0 < s_1 < s_2$ and sufficiently large $n$
\[
\mathbb{E} X_1 I\{s_1(n/d) < X_1 \leq s_2(n/d)\} \leq C_2(s_2 - s_1)/s_1. \tag{2.4}
\]
Here $C_1$, $C_2$ are positive constants depending only on the sequence $(X_k)$.

This is immediate from (1.12).

**Lemma 2.5** Let
\[
X_{k,n}^{(s_1,s_2)} = X_k I\{s_1(n/d) < X_k \leq s_2(n/d)\} - \mathbb{E} X_k I\{s_1(n/d) < X_k \leq s_2(n/d)\}.
\]
Then for any fixed $0 \leq t_1 < t_2 \leq 1$, $0 \leq s_1 < s_2 < \infty$, we have
\[
\mathbb{E} \left( \sum_{k=nt_1+1}^{nt_2} X_{k,n}^{(s_1,s_2)} \right)^2 \sim c_0(n^2/d)(t_2 - t_1)(s_2 - s_1) \quad \text{as } n \to \infty \tag{2.5}
\]
provided $nt_1, nt_2$ are integers. Moreover,
\[
\mathbb{E} \left( \sum_{i=nt_1+1}^{nt_2} X_{i,n}^{(s_1,s_2)} \right) \left( \sum_{j=nt_1'+1}^{nt_2'} X_{j,n}^{(s_1',s_2')} \right) = o(n^2/d) \quad \text{as } n \to \infty \tag{2.6}
\]
provided $0 \leq t_1 < t_2 \leq 1$, $0 \leq t_1' < t_2' \leq 1$, $0 \leq s_1 < s_2 < \infty$, $0 \leq s_1' < s_2' < \infty$, $nt_1, nt_2, nt_1', nt_2'$ are integers and the intervals $(nt_1, nt_2)$ and $(nt_1', nt_2')$ are identical or disjoint and the same holds for the intervals $(s_1, s_2)$ and $(s_1', s_2')$, but identity cannot hold at both places.

**Proof** We have
\[
\mathbb{E} \left( \sum_{k=nt_1+1}^{nt_2} X_{k,n}^{(s_1,s_2)} \right)^2 = n(t_2 - t_1)\mathbb{E} \left( X_{1,n}^{(s_1,s_2)} \right)^2 + R
\]
where
\[
R = 2 \sum_{j=2}^{nt_2 - nt_1} (nt_2 - nt_1 - j + 1)\mathbb{E} \left( X_{1,n}^{(s_1,s_2)} X_{j,n}^{(s_1,s_2)} \right).
\]
Using Lemmas 2.2 and 2.4, we get, using $n/d \to \infty$,
\[
\mathbb{E} \left( X_{1,n}^{(s_1,s_2)} \right)^2
\]
\begin{align*}
\mathbb{E}(X_1 I\{s_1(n/d) \leq X_1 \leq s_2(n/d)\})^2 & = \mathbb{E}(X_1 I\{s_1(n/d) < X_1 \leq s_2(n/d)\})^2 - \mathbb{E}^2(X_1 I\{s_1(n/d) < X_1 < s_2(n/d)\}) \\
& = c_0(1 + o(1))(n/d)(s_2 - s_1) + O(\log^2(n/d)) \sim c_0(n/d)(s_2 - s_1)
\end{align*}

and

\[ |R| \leq 2n \sum_{j=2}^{nt_2-nt_1} \psi(j - 1) \left( \mathbb{E}|X_{1,n}^{(s_1, s_2)}| \right)^2 \leq C_3 n \log^2(n/d) \sum_{j=1}^{\infty} e^{-\lambda j} = o(n^2/d), \]

proving (2.5).

To prove (2.6), consider a generic term

\begin{align*}
\mathbb{E}X_{i,j}^{(s_1, s_2)} X_{j,n}^{(s_1', s_2')}
& = \mathbb{E}X_i X_j I\{s_1(n/d) < X_i \leq s_2(n/d)\}I\{s_1'(n/d) < X_j \leq s_2'(n/d)\} \\
& - \mathbb{E}X_i I\{s_1(n/d) < X_i \leq s_2(n/d)\}\mathbb{E}X_j I\{s_1'(n/d) < X_j \leq s_2'(n/d)\}
\end{align*}

(2.7)

of the left hand side of (2.6). Fix \( r \geq 0 \) and sum those covariances in (2.7) where \( j - i = r \) and \( nt_1 + 1 \leq i \leq nt_2, nt_1' + 1 \leq j \leq nt_2' \). Clearly, the case \( r = 0 \) can occur only if \((nt_1, nt_2) = (nt_1', nt_2')\), but in this case, by the assumptions of the lemma, \((s_1, s_2)\) and \((s_1', s_2')\) must be disjoint and thus the product of the two indicators in the second line of (2.7) is 0. Thus, by the first statement of Lemma 2.4, the product expectation in the first line of (2.7) is \( O(\log^2(n/d)) \) and since the number of such terms in the expansion of (2.6) is at most \( n \), the contribution of such terms in the sum in (2.6) is at most \( O(n \log^2(n/d)) = o(n^2/d) \) by \( n/d \to \infty \). For \( r \geq 1 \), the covariance in (2.7) is at most \( \psi(r) O(\log^2(n/d)) \) by Lemma 2.2 and the first statement of Lemma 2.4 and since for fixed \( r \) the number of pairs \((i, j)\) is at most \( n \), the contribution of all such terms for all \( r \geq 1 \) is at most \( C n \log^2(n/d) \sum_{r=1}^{\infty} \psi(r) = O(n \log^2(n/d)) = o(n^2/d) \), proving (2.6).

The following central limit theorem for \( \phi \)-mixing sequences is due to Utev [31].

**Lemma 2.6** Let \( \{x_{nk}, 1 \leq k \leq n, n \geq 1\} \) be a triangular array of random variables with zero mean and finite variances. Assume that the array is \( \phi \)-mixing, i.e.,

\[ \phi(k) := \sup_{n \in \mathbb{N}, n > k} \max_{1 \leq m \leq n-k} \sup_{A \in \mathcal{F}_{1,m}, B \in \mathcal{F}_{m+1,k,n}, P(A) > 0} |P(B|A) - P(B)| \rightarrow 0 \]

as \( k \to \infty \),

(2.8)

where \( \mathcal{F}_{a,b}^{(n)} \) denotes the \( \sigma \)-algebra generated by the r.v.’s \( \{x_{nk}, a \leq k \leq b\} \). Assume further that the rows of the array are strictly stationary and

\[ \lim_{n \to \infty} \mathbb{E}\left( \sum_{k=1}^{n} x_{nk} \right)^2 = \sigma^2 < \infty \]

(2.9)
and that the Lindeberg condition
\[
\lim_{n \to \infty} n \mathbb{E} \left[ x_{n1}^2 I(|x_{n1}| \geq \varepsilon) \right] = 0 \quad \text{for all } \varepsilon > 0 \quad (2.10)
\]
holds. Then
\[
\sum_{k=1}^{n} x_{nk} \xrightarrow{d} N(0, \sigma^2) \quad \text{as } n \to \infty.
\]

3 Proof of Theorem 1.3

Put
\[
Q_n = \frac{1}{A_n} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j} U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])
\]
and
\[
Z = \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j} W([t_{m-1}, t_m] \times [s_{j-1}, s_j])
\]
for all \( M \geq 1, J \geq 1 \), real coefficients \( \mu_{m,j}, 0 = s_0 < s_1 < s_2 < \cdots < s_J < \infty, 0 = t_0 < t_1 < \cdots < t_M = 1 \). Clearly, \( Z \) is a normal random variable with mean zero and
\[
\mathbb{E} Z^2 = \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^2 (t_m - t_{m-1})(s_j - s_{j-1}). \quad (3.1)
\]
We claim that
\[
Q_n \xrightarrow{d} Z \quad \text{for all considered values of } M, J, \mu_{m,j}, t_m, s_j. \quad (3.2)
\]
Since the processes \( U_n \) and \( W \) are equal to 0 on the boundary of the first quadrant, we have
\[
U_n(t_m, s_j) = \sum_{p=1}^{m} \sum_{q=1}^{j} U_n([t_{p-1}, t_p] \times [s_{q-1}, s_q])
\]
and the same relation holds for \( W \). Thus (3.2) implies
\[
\frac{1}{A_n} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^* U_n(t_m, s_j) \xrightarrow{d} \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^* W(t_m, s_j)
\]
for arbitrary real coefficients $\mu_{m,j}^*$ and this, by the Cramér–Wold device, implies the convergence of the finite-dimensional distributions in Theorem 1.3.

Clearly, $U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])$ equals

$$
\sum_{k=\lfloor nt_{m-1}\rfloor + 1}^{nt_m} X_k I[s_{j-1}(n/d) < X_k \leq s_j(n/d)] - \mathbb{E} X_k I[s_{j-1}(n/d) < X_k \leq s_j(n/d)]
$$

and thus relation (3.2) is equivalent to

$$
\frac{1}{A_n} \sum_{k=1}^{n} (z_{nk} - \mathbb{E} z_{nk}) \xrightarrow{d} N(0, \mathbb{E} Z^2),
$$

where

$$
z_{nk} = \sum_{j=1}^{J} \mu_{m,j} X_k I[s_{j-1}(n/d) < X_k \leq s_j(n/d)],
$$

$$
[nt_{m-1}] + 1 \leq k \leq [nt_m].
$$

Since the terms of the sum in (3.4) are random variables with disjoint support, by relation (2.3) of Lemma 2.4 we have

$$
\mathbb{E} z_{nk}^2 = (1 + o_n(1))c_0(n/d) \sum_{j=1}^{J} \mu_{m,j}^2 (s_j - s_{j-1}),
$$

$$
[nt_{m-1}] + 1 \leq k \leq [nt_m].
$$

Consequently, letting

$$
B_m = c_0 \sum_{j=1}^{J} \mu_{m,j}^2 (s_j - s_{j-1}),
$$

we get

$$
\text{Var } z_{nk} \leq (1 + o_n(1))(n/d) B_m, \quad [nt_{m-1}] + 1 \leq k \leq [nt_m].
$$

Further, Lemma 2.5 implies for $n \to \infty$

$$
\mathbb{E} U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])^2 = (1 + o_n(1))c_0(t_m - t_{m-1})(s_j - s_{j-1})(n^2/d)
$$

and

$$
\mathbb{E} U_n([t_{m_1-1}, t_{m_1}] \times [s_{j_1-1}, s_{j_1}]) U_n([t_{m_2-1}, t_{m_2}] \times [s_{j_2-1}, s_{j_2}]) = o_n(n^2/d)
$$

\( \square \) Springer
provided the pairs \((m_1, j_1)\) and \((m_2, j_2)\) are different. Thus

\[
E \left( \sum_{k=1}^{n} (z_{nk} - E z_{nk}) \right)^2 = E \left( \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j} U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j]) \right)^2
\]

\[
= \sum_{m_1,m_2=1}^{M} \sum_{j_1,j_2=1}^{J} \mu_{m_1,j_1} \mu_{m_2,j_2} E \left[ U_n([t_{m_1-1}, t_{m_1}] \times [s_{j_1-1}, s_{j_1}]) U_n([t_{m_2-1}, t_{m_2}] \times [s_{j_2-1}, s_{j_2}]) \right]
\]

\[
\sim c_0(n^2/d) \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^2 (t_m - t_{m-1})(s_j - s_{j-1}) = c_0(n^2/d) E Z^2 = A_n^2 E Z^2. \tag{3.7}
\]

Also, by (3.4) and the second relation of (2.2) we have

\[
\| z_{nk} \|_4 \leq \sum_{j=1}^{J} |\mu_{m,j}| \left( E X_k^4 I \{ s_{j-1}(n/d) < X_k \leq s_j(n/d) \} \right)^{1/4}
\]

\[
\leq C_1^{1/4} (n/d)^{3/4} \sum_{j=1}^{J} |\mu_{m,j}| s_j^{3/4}
\]

and consequently

\[
\| z_{nk} - E z_{nk} \|_4 \leq 2C_1^{1/4} (n/d)^{3/4} \sum_{j=1}^{J} |\mu_{m,j}| s_j^{3/4}. \tag{3.8}
\]

We apply now Lemma 2.6 for the triangular array

\[
x_{nk} = (z_{nk} - E z_{nk})/A_n, \quad 1 \leq k \leq n, n \geq 1. \tag{3.9}
\]

Since \( \{ X_j, j \geq 1 \} \) is \( \psi \)-mixing with exponential rate, the array (3.9) satisfies the \( \phi \)-mixing condition (2.8). Relation (3.7) shows that (2.9) holds with \( \sigma^2 = E Z^2 \). Finally, (3.8), (1.13) and \( d = d_n \rightarrow \infty \) show that the array (3.9) satisfies the Lyapunov condition

\[
\lim_{n \rightarrow \infty} n E x_{n1}^4 = 0 \tag{3.10}
\]

which implies the Lindeberg condition (2.10). Thus Lemma 2.6 applies and the central limit theorem (3.3) follows.

Next, we prove tightness in Theorem 1.3. Let

\[
B_{11} = [t_1, t] \times [s_1, s], \quad B_{12} = [t_1, t] \times [s, s_2], \quad B_{21} = [t, t_2] \times [s_1, s]
\]
where \(0 \leq t_1 < t < t_2 \leq 1, 1/2 \leq s_1 < s < s_2 \leq 3/2\). In view of Lemma 2.1, it suffices to show that

\[
E \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{ij}) \right|^2 \leq C^* \mu(B_{11}) \mu(B_{ij}), \tag{3.11}
\]

holds for each \(ij \in \{12, 21\}\) with some constant \(C^* > 0\). Moreover, since \(U_n(t, s)\) is constant on intervals \(k/n \leq t < (k + 1)/n\), by the last statement of Lemma 2.1 we may assume that \(nt, nt_1\) and \(nt_2\) are all integers. To prove (3.11), we introduce the notations

\[
X_i^{(1)} = X_i I\{s_1(n/d) < X_i \leq s(n/d)\}, \quad m_i^{(1)} = EX_i^{(1)},
\]

\[
X_i^{(2)} = X_i I\{s(n/d) < X_i \leq s_2(n/d)\}, \quad m_i^{(2)} = EX_i^{(2)}.
\]

Using Lemmas 2.3 and 2.5 and (1.13), we get

\[
E \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{21}) \right|^2
= E \left( \frac{1}{A_n} \sum_{i=nt_1+1}^{nt} (X_i^{(1)} - m_i^{(1)}) \right)^2 \left( \frac{1}{A_n} \sum_{i=nt_1+1}^{nt_2} (X_i^{(1)} - m_i^{(1)}) \right)^2
\leq (1 + \psi(1))^2 \frac{1}{A_n^4} E \left( \sum_{i=nt_1+1}^{nt} (X_i^{(1)} - m_i^{(1)}) \right)^2 E \left( \sum_{i=nt_1+1}^{nt_2} (X_i^{(1)} - m_i^{(1)}) \right)^2
\leq C_8(t - t_1)(t_2 - t)(s - s_1)^2 = C_8 \mu(B_{11}) \mu(B_{21})
\]

for \(n \geq n_0\). On the other hand,

\[
E \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{12}) \right|^2
= \frac{1}{A_n^4} E \left( \sum_{i=nt_1+1}^{nt} (X_i^{(1)} - m_i^{(1)}) \right)^2 \left( \sum_{i=nt_1+1}^{nt} (X_i^{(2)} - m_i^{(2)}) \right)^2
= \frac{1}{A_n^4} E \left( \sum_{i=nt_1+1}^{nt} Y_i^{(1)} \right)^2 \left( \sum_{i=nt_1+1}^{nt} Y_i^{(2)} \right)^2, \tag{3.12}
\]

where we put

\[
Y_i^{(1)} = X_i^{(1)} - m_i^{(1)}, \quad Y_i^{(2)} = X_i^{(2)} - m_i^{(2)}.
\]
The expression in the third line of (3.12) equals the sum of all expressions

\[ A_n^{-4} \mathbb{E}(Y^{(1)}_i Y^{(1)}_j Y^{(2)}_k Y^{(2)}_{\ell}), \]  

(3.13)

where \( nt_1 + 1 \leq i, j, k, \ell \leq nt \). The following facts can be verified by elementary calculations using Lemmas 2.2–2.4:

(a) \( \mathbb{E}|Y^{(1)}_i| \ll s - s_1, \quad \mathbb{E}|Y^{(2)}_i| \ll s_2 - s, \quad \mathbb{E}|Y^{(1)}_i Y^{(2)}_i| \ll (s - s_1)(s_2 - s) \)

(b) \( \mathbb{E}(Y^{(1)}_i)^2 \ll (n/d)(s - s_1), \quad \mathbb{E}(Y^{(2)}_i)^2 \ll (n/d)(s_2 - s) \),

(c) \( \mathbb{E}(Y^{(1)}_i)^2|Y^{(2)}_i| \ll (n/d)(s - s_1)(s_2 - s), \quad \mathbb{E}|Y^{(1)}_i| |(Y^{(2)}_i)^2 \ll (n/d)(s - s_1)(s_2 - s), \)

(d) \( \mathbb{E}(Y^{(1)}_i)^2(Y^{(2)}_i)^2 \ll (n/d)(s - s_1)(s_2 - s), \)

where \( \ll \) means the same as the \( O \) notation, with an implied constant depending on the sequence \( (X_n) \). We prove relation (d), the proof of (a), (b), (c) is similar (and simpler). We have

\[ \mathbb{E}(Y^{(1)}_i)^2(Y^{(2)}_i)^2 = \mathbb{E}\left[ (X^{(1)}_i - m^{(1)}_i)^2(X^{(2)}_i - m^{(2)}_i)^2 \right] \]

\[ = \mathbb{E}(X^{(1)}_i)^2(X^{(2)}_i)^2 - 2m^{(2)}_i\mathbb{E}(X^{(1)}_i)^2X^{(2)}_i + (m^{(2)}_i)^2\mathbb{E}(X^{(1)}_i)^2 - 2m^{(1)}_i\mathbb{E}X^{(1)}_i(X^{(2)}_i)^2 + 4m^{(1)}_i m^{(2)}_i\mathbb{E}X^{(1)}_i X^{(2)}_i - 2m^{(1)}_i (m^{(2)}_i)^2\mathbb{E}X^{(1)}_i + (m^{(1)}_i)^2\mathbb{E}(X^{(2)}_i)^2 \]

\[ - 2(m^{(1)}_i)^2m^{(2)}_i\mathbb{E}(X^{(2)}_i) + (m^{(1)}_i)^2(m^{(2)}_i)^2. \]

Clearly, \( X^{(1)}_i \) and \( X^{(2)}_i \) are supported on different sets and thus \( X^{(1)}_i X^{(2)}_i = 0 \). Thus, among the nine terms above, the first, second, fourth and fifth are equal to 0. Also, the second and third statement of Lemma 2.4 imply, in view of \( 1/2 \leq s_1 < s < s_2 \leq 3/2, \)

\[ m^{(1)}_i = \mathbb{E}X^{(1)}_i \ll s - s_1, \quad m^{(2)}_i = \mathbb{E}X^{(2)}_i \ll s_2 - s, \]

\[ \mathbb{E}(X^{(1)}_i)^2 \ll (s - s_1)(n/d), \quad \mathbb{E}(X^{(2)}_i)^2 \ll (s_2 - s)(n/d) \]

for \( n \geq n_0 \). This shows that the remaining five terms of the nine-term sum above are \( \ll (n/d)(s - s_1)(s_2 - s) \), proving statement (d) above. Statements (a), (b) and (c) can be proved similarly.

We can now estimate the expressions in (3.13). We will distinguish four cases according as \( i, j, k, \ell \) are all different, or the number of different ones among them is 1, 2 or 3. Consider first the case when \( i, j, k, \ell \) are all different, say \( i < j < k < \ell \); let \( r = j - i \). Applying Lemma 2.2 with \( X = Y^{(1)}_i, \ Y = Y^{(1)}_j Y^{(2)}_k Y^{(2)}_{\ell} \) and using that \( EX = 0 \), we get that the absolute value of the expectation in (3.13) is bounded by

\[ A_n^{-4} \psi(r) \mathbb{E}|X| \mathbb{E}|Y| \leq CA_n^{-4} \psi(r) \mathbb{E}|Y^{(1)}_i| \mathbb{E}|Y^{(1)}_j| \mathbb{E}|Y^{(2)}_k| \mathbb{E}|Y^{(2)}_{\ell}| \]

\[ \leq CA_n^{-4} \psi(r)(s - s_1)^2(s_2 - s)^2, \]

where we used Lemma 2.3 to estimate \( \mathbb{E}|Y| \) and relation (a) above. Here, and in the rest of the tightness proof, \( C \) denotes (possibly different) constants depending only
on the sequence \((X_n)\). Arguing similarly, but splitting the four-term product in (3.13) after the third term, we get the same bound, except that \(\psi(r)\) gets replaced by \(\psi(r')\), where \(r' = \ell - k\). Thus, the absolute value of the expression in (3.13) is at most

\[
CA_n^{-4}\psi(r)^{1/2}\psi(r')^{1/2}(s-s_1)^2(s_2-s)^2.
\]

Fixing the pair \((i, \ell)\) and summing for \((j, k)\) means summing for \((r, r')\), and since \(\sum_{n=1}^{\infty} \psi(n)^{1/2} < \infty\) and the pair \((i, \ell)\) can be chosen by at most \((nt-nt_1)^2\) different ways, it follows that the contribution of all terms (3.13) with \(i < j < k < \ell\) is at most

\[
CA_n^{-4}(nt-nt_1)^2(s-s_1)^2(s_2-s)^2 \leq C(d^2/n^2)(t-t_1)^2(s-s_1)^2(s_2-s)^2
\]

\[
\leq C(t-t_1)^2(s-s_1)(s_2-s) = C\mu(B_{11})\mu(B_{12}),
\]

using (1.13) and \(d/n \rightarrow 0\). The contribution of terms (3.13) where \(i, j, k, \ell\) are different, but their order is different can be estimated similarly.

Next, we consider the case when \(i = j = k = \ell\). In this case, the expression (3.13) becomes \(A_n^{-4}E(Y_{i(1)}^2)(Y_{i(2)}^2)^2\), which by the estimate in (d) above is at most \(CA_n^{-4}(n/d)(s-s_1)(s_2-s)\). Since the number of choices for \(i\) is \(nt-nt_1 \leq (nt-nt_1)^2\), the contribution of all such expressions is bounded by

\[
CA_n^{-4}(n/d)(s-s_1)(s_2-s)(nt-nt_1)^2 \leq C(d/n)(s-s_1)(s_2-s)(t-t_1)^2
\]

\[
\leq C\mu(B_{11})\mu(B_{12}),
\]

using again (1.13) and \(d/n \rightarrow 0\).

Assume now that among \(i, j, k, \ell\), there are two different ones, i.e., these numbers are pairwise equal or three are equal and the fourth is different. Starting with the case of two pairs, assume, e.g., that \(i = j\) and \(k = l\), but \(i \neq k\). In this case, the expression (3.13) becomes \(A_n^{-4}E(Y_{i(1)}^2)(Y_{k(2)}^2)\) which, in view of Lemma 2.3 and the estimate in (b) above is at most

\[
CA_n^{-4}(n/d)^2(s-s_1)(s_2-s).
\]

Since the number of choices for the pair \((i, k)\) is at most \((nt-nt_1)^2\), using (1.13) it follows that the total contribution of all such terms (3.13) is at most

\[
CA_n^{-4}(n/d)^2(s-s_1)(s_2-s)(nt-nt_1)^2 \leq C(s-s_1)(s_2-s)(t-t_1)^2
\]

\[
= C\mu(B_{11})\mu(B_{12}).
\]

If \(i = k, j = l\) and \(i \neq j\), then the expression (3.13) becomes \(A_n^{-4}E|Y_{i(1)}Y_{i(2)}|E|Y_{j(1)}Y_{j(2)}|\), which by Lemma 2.3 and the estimate in (a) above is bounded by

\[
CA_n^{-4}E|Y_{i(1)}Y_{i(2)}||E|Y_{j(1)}Y_{j(2)}| \leq CA_n^{-4}(s-s_1)^2(s_2-s)^2.
\]
Since the number of pairs \((i, j)\) is \(\leq (nt - nt_1)^2\), the contribution of such terms is at most

\[
CA_n^{-4}(s - s_1)^2(s_2 - s)^2(nt - nt_1)^2 \leq C(s - s_1)(s_2 - s)(t - t_1)^2
\]

\[
= C\mu(B_{11})\mu(B_{12}).
\]

Assume now that from the indices \(i, j, k, l\) three are equal and the fourth one is different. Letting, e.g., \(i = j = k\) and \(i \neq \ell\), the expression (3.13) becomes \(A_n^{-4}E(Y_i^{(1)})^2Y_i^{(2)}Y_\ell^{(2)}\) which is, by Lemma 2.3 and the estimates (a) and (c) above is bounded by

\[
CA_n^{-4}E(Y_i^{(1)})^2|Y_i^{(2)}|E|Y_\ell^{(2)}| \leq CA_n^{-4}(n/d)(s - s_1)(s_2 - s)^2.
\]

Since the number of pairs \((i, \ell)\) is \(\leq (nt - nt_1)^2\), the total contribution of such terms is at most \(C\mu(B_{11})\mu(B_{12})\).

Finally, if the number of different indices among \(i, j, k, l\) is 3, e.g., if \(i = j < k < \ell\), then the expression (3.13) becomes \(A_n^{-4}E(Y_i^{(1)})^2Y_k^{(2)}Y_\ell^{(2)}\), which by using \(EY_\ell^{(2)} = 0\), Lemma 2.2, Lemma 2.3 and estimates (a) and (b) above, can be estimated by

\[
CA_n^{-4}\psi(r)E(Y_i^{(1)})^2E|Y_k^{(2)}|E|Y_\ell^{(2)}| \leq CA_n^{-4}\psi(r)(n/d)(s - s_1)(s_2 - s)^2,
\]

where \(r = \ell - k\). Since for fixed \(r\) the number of triples \((i, k, \ell)\) with \(\ell - k = r\) is at most \((nt - nt_1)^2\), the contribution of such terms (3.13) is at most

\[
CA_n^{-4}\psi(r)(n/d)(s - s_1)(s_2 - s)^2(nt - nt_1)^2 \leq C\psi(r)(s - s_1)(s_2 - s)(t - t_1)^2
\]

and summing for \(r\) we get again \(\leq C\mu(B_{11})\mu(B_{12})\). The other cases (e.g., \(i < j = k < \ell\)) can be treated similarly, and the proof of tightness in Theorem 1.3 is completed. This also completes the proof of the theorem.

We prove now, as claimed after Theorem 1.3 that

\[
\frac{\eta_{d,n}}{n/d} \overset{p}{\rightarrow} 1 \quad (3.15)
\]

for \(d = d_n \rightarrow \infty, d_n/n \rightarrow 0\). Fix \(n \geq 1, 1/2 < t < 2\) and let us denote \(T_k = I\{X_k \geq tn/d\}, 1 \leq k \leq n\). Then by Lemma 2.2 and (1.12), we get

\[
|E T_1 T_k - E T_1 E T_k| \leq \psi(k - 1)E T_1 E T_k \leq C_9 \exp(-\lambda k)(d/n)^2
\]

and thus, setting \(\overline{T}_k = T_k - E T_k\), we conclude that

\[
E \left(\sum_{k=1}^{n} \overline{T}_k\right)^2 = nE \overline{T}_1^2 + 2 \sum_{k=2}^{n} (n - k + 1)E \overline{T}_1 \overline{T}_k
\]
\[ \leq n \left( \mathbb{E} \tilde{T}_1^2 + 2 \sum_{k=2}^{n} |\mathbb{E} \tilde{T}_1 \tilde{T}_k| \right) \]
\[ \leq n \left( \mathbb{E} T_1^2 + C_{10} (d/n)^2 \sum_{k=2}^{n} \exp(-\lambda k) \right) \]
\[ \leq n \left( \mathbb{E} T_1 + C_{11} (d/n)^2 \right) \]
\[ \leq C_{12} d. \]

Hence, Markov’s inequality and \( d = d_n \to \infty \) imply for any \( \varepsilon > 0 \)

\[ P \left\{ \sum_{k=1}^{n} \tilde{T}_k \geq \varepsilon d \right\} \to 0, \]

and since \( \mathbb{E} \tilde{T}_k = \mathbb{E} T_1 \sim d/(nt) \) by (1.12), it follows that

\[ \#\{k \leq n : X_k \geq tn/d\} = \sum_{k=1}^{n} I\{X_k \geq tn/d\} \sim d/t \quad \text{in probability as } n \to \infty. \]

Thus, for fixed \( t > 1 \) and \( n \) large, with probability tending to 1 the number of \( X_k \)'s, \( 1 \leq k \leq n \) exceeding \( tn/d \) is smaller than \( d \) and thus \( \eta_{d,n} \leq tn/d \). Similarly, for \( t < 1 \) and \( n \) large, with probability tending to 1 we have \( \eta_{d,n} \geq tn/d \), and thus, (3.15) is proved.

**Proof of Remark 1.1** Let \((X_n)\) be a sequence satisfying the assumptions of Theorem 1.2, and put \( X'_n = X_n + 4^{-n} \). Letting \( \eta'_{d,n} \) denote the \( d \)-th largest of \( X'_1, \ldots, X'_n \) and \( S^{(r)}_n \) and \( S^{(r)}'_n \) denote the sums \( \sum_{k=1}^{n} X_k, \sum_{k=1}^{n} X'_k \) after removing their \( r \) largest terms, it is easily seen that

\[ |S^{(r)}_n - S^{(r)}'_n| \leq 2 \quad \text{for any } r \geq 1 \]  \hspace{1cm} (3.16)

and

\[ n|m(\eta_{d,n}) - m(\eta'_{d,n})| = O_p(1). \]  \hspace{1cm} (3.17)

Clearly, relation (1.12) will fail for the perturbed sequence \((X'_n)\), but as inspection shows, all the lemmas in the proof of Theorem 1.2 and the subsequent arguments remain valid, so conclusion (1.14) of the theorem remains valid if we replace \( X_i \) by \( X'_i \) and \( \eta_{d,n} \) by \( \eta'_{d,n} \). Since the \( X_n \) are integer valued, with probability one, all the \( X'_j \), \( j = 1, 2, \ldots \) are different, and thus, the sum of the \( X'_j \)'s, \( 1 \leq j \leq n \) not exceeding \( \eta'_{d,n} \) equals \( S^{(d-1)}_n \). Thus we have

\[ \frac{S^{(d-1)}_n - nm(\eta'_{d,n})}{n/\sqrt{d}} \overset{d}{\rightarrow} N(0, c_0). \]  \hspace{1cm} (3.18)
In view of (3.16) and (3.17), we can drop the primes in (3.18) and since $S_n^{(d-1)} - S_n^{(d)} = \eta_{d,n} = O_P(n/d)$ by (3.15), the conclusion of Theorem 1.1 follows.

**Acknowledgments**  We would like to thank the referee for her/his remarks leading to a substantial improvement of the paper.

**References**

1. Arov, D.Z., Bobrov, A.A.: The extreme terms of a sample and their role in the sum of independent variables. Theor. Probab. Appl. 5, 377–396 (1960)
2. Bickel, P.J., Wichura, M.J.: Convergence criteria for multiparameter stochastic processes and some applications. Ann. Math. Stat. 42, 1656–1670 (1971)
3. Billingsley, P.: Ergodic Theory and Information. Wiley, New York (1965)
4. Billingsley, P.: Convergence of Probability Measures. Wiley, New York (1968)
5. Bradley, R.: Introduction to Strong Mixing Conditions, vol. I. Kendrick Press, Heber City (2007)
6. Csörgő, S., Horváth, L., Mason, D.: What portion of the sample makes a partial sum asymptotically stable or normal? Z. Wahrschein. verw. Gebiete 72, 1–16 (1986)
7. Csörgő, S., Simons, G.: A strong law of large numbers for trimmed sums, with applications to generalized St. Petersburg games. Stat. Probab. Lett. 26, 65–73 (1996)
8. Darling, D.: The influence of the maximum term in the addition of independent random variables. Trans. Am. Math. Soc. 73, 95–107 (1952)
9. Diamond, H., Vaaler, J.: Estimates for partial sums of continued fraction partial quotients. Pac. J. Math. 122, 73–82 (1989)
10. Doeblin, W.: Remarques sur la théorie métrique des fractions continues. Compos. Math. 7, 353–371 (1940)
11. Galambos, J.: The distribution of the largest coefficient in continued fraction expansions. Q. J. Math. Oxf. Ser. 23, 147–151 (1972)
12. Galambos, J.: An iterated logarithm type theorem for the largest coefficient in continued fractions. Acta Arith. 25, 359–364 (1973/74)
13. Gordin, M.I., Reznik, M.H.: The law of the iterated logarithm for the denominators of continued fractions. Vestn. Leningr. Univ. 25, 28–33 (1970). (In Russian)
14. Heinrich, L.: Rates of convergence in stable limit theorems for sums of exponentially $\psi$-mixing random variables with an application to metric theory of continued fractions. Math. Nachr. 131, 149–165 (1987)
15. Ibragimov, I.A.: A theorem from the metric theory of continued fractions. Vestn. Leningr. Univ. 1, 13–24 (1960). (In Russian)
16. Iosifescu, M.: A Poisson law for $\psi$-mixing sequences establishing the truth of a Doeblin’s statement. Rev. Roum. Math. Pures Appl. 22, 1441–1447 (1977)
17. Iosifescu, M.: A survey of the metric theory of continued fractions, fifty years after Doeblin’s 1940 paper. In: Grigelionis, B., et al. (eds.) Probability Theory and Mathematical Statistics, vol. 1, pp. 550–572. Mokslas, Vilnius (1990)
18. Khinchin, A.J.: Metrische Kettenbruchprobleme. Compos. Math. 1, 361–382 (1935)
19. Kusmin, R.: Sur un problème de Gauss. Atti Congr. Int. Bol. 6, 83–89 (1928)
20. Leadbetter, M., Rootzén, H.: Extremal theory for stochastic processes. Ann. Probab. 16, 431–478 (1988)
21. Lévy, P.: Sur les lois de probabilité dont dépendent les quotients complets et incomplets d’une fraction continue. Bull. Sci. Math. Fr. 57, 178–194 (1929)
22. Lévy, P.: Fractions continues aléatoires. Rend. Circ. Mat. Palermo 1, 170–208 (1952)
23. Philipp, W.: Some metrical theorems in number theory II. Duke Math. J. 37, 447–458 (1970)
24. Philipp, W.: A conjecture of Erdős on continued fractions. Acta Arith. 28, 379–386 (1975/76)
25. Philipp, W.: Limit theorems for partial quotients of continued fractions. Mon. Math. 105, 195–206 (1988)
26. Philipp, W., Stackelberg, O.: Zwei Grenzwertssätze für Kettenbrüche. Math. Ann. 181, 152–156 (1969)
27. Samur, J.: On some limit theorems for continued fractions. Trans. Am. Math. Soc. 316, 53–79 (1989)
28. Samur, J.: Some remarks on a probability limit theorem for continued fractions. Trans. Am. Math. Soc. 348, 1411–1428 (1996)
29. Stackelberg, O.P.: On the law of the iterated logarithm for continued fractions. Duke Math. J. 33, 801–819 (1966)
30. Szewczak, Z.S.: On limit theorems for continued fractions. J. Theor. Probab. 22, 239–255 (2009)
31. Utev, S.A.: On the central limit theorem for $\varphi$-mixing arrays of random variables. Theory Probab. Appl. 35, 131–139 (1990)
32. Vardi, I.: The St. Petersburg game and continued fractions. C. R. Acad. Sci. Paris Ser. I Math. 324, 913–918 (1997)