Alina Bazarova

Asymptotic properties of trimmed sums and their applications in Analysis and Statistics

DISSERTATION

zur Erlangung des akademischen Grades einer
Doktorin der technischen Wissenschaften

Doktoratsstudium der Technischen
Wissenschaften im Rahmen der Doktoratsschule
"Mathematik und Wissenschaftliches Rechnen"

Technische Universität Graz

Betreuer:
Univ.-Prof. Dr. István Berkes

Institut für Statistik

Graz, April 2014

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am ....

(Unterschrift)

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely marked all material which has been quotes either literally or by content from the used sources.

....

date (signature)

Contents

Abstract

Random variables with infinite moments is not a new concept in Probability Theory. Furthermore, it was shown that quite often one can observe suchlike distributions in real life. For example (see [37], [72], [73]) distributions of stock and commodity returns are quite often heavy-tailed with infinite variance. Needless to say that at this moment in time there exists a big variety of methods to deal with heavy-tailed distributions. This thesis concentrates on the one which was developed in 1960s, that is the trimming theory. The main idea of this method is to remove elements with extreme values from the sample. We develop a new approach to the Central Limit Theorem for modulus trimmed sums. This approach gives a simpler proof of the CLT for independent random variables with heavy tails and also provides a possibility of extending the CLT to the case of dependent random variables. We immediately demonstrate application of this last result in Analysis by using it for the case of continued fractions. Further we establish functional trimmed CLT for AR(1) processes with stable (heavy-tailed) innovations. This result is used to get the asymptotics of the CUSUM process. Finally, for the same type of samples we develop two types of test-statistics and run Monte-Carlo simulations, which in the end give satisfactory results.

Acknowledgement

I am very grateful to Prof. István Berkes and Prof. Lajos Horváth, persons without whom this thesis would never exist. I also would like to thank my collegues from the Institute of Statistics for friendly atmosphere, good coffee and many interesting experiences.

I dedicate this document to my mother, who was the one to encourage me to take this step in life.

Introduction

Nowadays, eighty years since Kolmogorov laid the modern axiomatic foundations of probability theory, one can state that a great deal of achievements has been made, a plenty of probabilistic models has been described and implemented into real life. However, as it turns out, there is still a lot of space left for new results, even in the topics which seem to be fairly well explored.

Sequences of random variables with infinite moments have been in the zone of consideration since the very start of development of probability theory. But in 1960's a new way of dealing with suchlike sequences started to emerge, i.e. trimming theory. The main idea of this theory is, given a sample of random variables, to remove the largest (one or several) elements from it. This approach was motivated by the fact that for certain distributions the reason one can not handle them in the same way as the ones which have finite, for example second moments, is essentially a few large terms (see Figure 1).

Trimming of random variables with heavy tails is what binds all three chapters of this thesis together. Therefore before we move to the description of each chapter in particular we will give a brief overview on the history and development of trimming theory.

When talking about trimming of the sample $X_1, \dots, X_n$ of random variables one can consider two models:

(i) We order this sample in a way that $X_1^{(n)} \leq \cdots \leq X_n^{(n)}$. Then the sum $S_n^{(r)} = \sum_{k=r+1}^{n-r} X_k$ is called ordinary trimmed sum.

(ii) We order the sample by its moduli, i.e. $|X_{n,1}| \leq \cdots \leq |X_{n,n}|$. Then a sum $S_n^{(r)} = \sum_{k=r+1}^{n} X_k$ is called modulus trimmed sum.

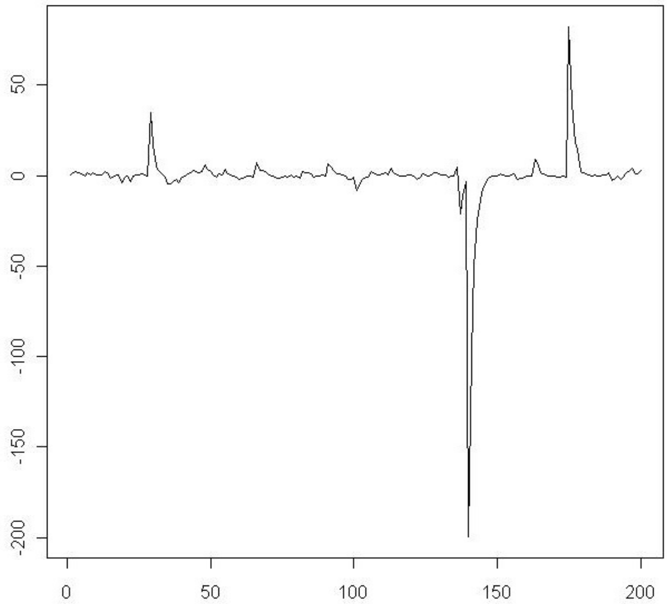

Figure 1: Simulated sample of AR(1) process with stable innovations

One can immediately see, and commonly known results confirm it, that there are substantial differences between two models. However before we mention any classical results in this theory within each model we introduce a subdivision according to the amount of the removed elements r_n

(i) Light trimming, when $r_n = r$ is a constant which does not depend on the size of a sample.

(ii) Moderate trimming, when $r_n/n \to 0$ as $n \to \infty$.

(iii) Heavy trimming, when $r_n \sim \alpha n$, where α is a constant from the interval $(0, 1)$.

Kesten and Maller studied the effects of both ordinary and modulus trimming on various forms of convergence and divergence of the sample sum of i.i.d. random variables (see e.g. [56], [57], [58]), provided that the amount of trimmed variables is fixed (i.e. under assumption of light trimming). Kesten shown that this kind of trimming does not lead to a change behavior of the sample sum (see [55]).

Stigler in 1973 (see [90]) summarized the limit behavior of heavily trimmed sums (in case of ordinary trimming) by giving necessary and sufficient conditions of its asymptotic normality. Note that the sufficient condition was well known before that (see Huber [50], Stigler [90]). Stigler shown that the asymptotic normality is guaranteed if and only if the quantiles corresponding to the proportions of trimming are uniquely defined for the underlying distribution. In case this condition is not satisfied one observes convergence to more complicated process, parameters of which also depend on the proportions of trimming (and quantiles of the distribution).

Maller (see [71]) and Mori (see [74]) shown that modulus trimmed sum is asymptotically normal if and only if the underlying distribution is in the domain of attraction of the normal. The case of underlying distribution in the domain of attraction of a stable law was investigated by Arov and Bobrov (see [3]).

The case of moderate ordinary trimming for samples of random variables in a domain of attraction of a stable law was resolved by Csörgő, Horváth and Mason in 1986(see [26]). They shown that if one removes r_n largest and r_n smallest elements from the nth partial sum, provided r_n satisfies the conditions of moderate trimming, one can center and normalize it in a way that it would weakly converge to standard normal distribution.

However the case of modulus trimming turned out to differ from the ordinary one quite fundamentally. In 1987 Griffin and Pruitt (see [44]) found necessary and sufficient conditions for asymptotic normality of a modulus trimmed sum only in the case when underlying distribution function is symmetric, but not in general. Later in 1989 (see Griffin and Pruitt [45]) they shown also that if the sum with r_n biggest and r_n smallest terms removed is asymptotically normal, then trimming out elements with $2r_n$ largest modulus will also lead to asymptotic normality. However this holds again only in case of symmetric distribution function. For the asymmetric case counterexamples were provided.

Concerning the non-symmetric distribution function a remarkable result was achieved by Berkes and Horváth in 2012 (see [14]). They shown that the validity of the Central Limit Theorem for modulus trimmed sum depends sensitively on the speed of convergence of the tail ratio of the underlying distribution and the asymptotical amount of the trimmed variables.

Assuming that to this point we have given enough insight in the history of the trimming theory we now move into descriptions of the chapters.

Chapter 1:

First chapter deals with modulus trimmed sums of independent identically distributed random variables in the domain of attraction of a stable law. The main result is the CLT for modulus trimmed sums with random centering factor. We show that allowing a random centering, the modulus trimmed CLT holds under exactly the same conditions as the one for ordinary trimming.

Under additional assumptions modulus trimmed CLT was proved in Berkes, Horváth and Schauer (see [16]) using a fairly complicated argument. The argument which is provided in chapter 1 is simpler and allows extension to the case of dependent variables which is going to be discussed in chapter 2.

The main idea is to deduce the required CLT from the two-dimensional limit result by using continuity of the limit process and classical results from Kiefer and Billingsley (see [60], [19] respectively).

This result can also be used to detect change point in the mean or location in the case of observations without second moments, as it provides a limit relation for CUSUM process where r_n observations with biggest moduli are excluded from the sample.

Chapter 2:

In chapter 2 we provide a number-theoretical application of the modulus trimmed CLT for dependent random variables. We consider continued fraction expansion of an irrational number from the interval $(0, 1)$.

It is known that the sequence of continued fraction expansion of x is a stationary ergodic sequence with respect to Gauss measure μ on the class of Borel subsets of $(0, 1)$ (see e.g. Billingsley, [18]), where

$$\mu(E) = \frac{1}{\log 2} \int\limits_E \frac{1}{1+t} dt.$$

As it turned out the sequence of continued fraction expansion digits possesses quite remarkable dependence properties, that is $\psi-$mixing. Gauss noted that the distribution of each term of the sequence with respect to the uniform measure in

$(0, 1)$ converges to μ and asked for the speed of convergence. Almost a hundred years later Kusmin (see [63]) showed that the convergence speed is sub-exponential and a year later Lévy (see [66]) improved it to an exponential rate. Lévy's result also implies that the sequence is $\psi-$mixing with an exponential rate.

Khinchin (see [59]) studied asymptotic behavior of partial sums of the continued fractions expansion digits. He provided a result for convergence of normalized partial sums in measure and noted that in this case convergence type can not be improved to almost everywhere convergence.

Later Diamond and Vaaler (see [35]) shown that convergence type actually can be improved to a.e. However to do it one needs to exclude the largest summand from the partial sum.

In chapter 2 we provide a CLT for trimmed sums of partial quotients, where the amount of trimmed elements satisfies the condition of moderate trimming. We also prove analogous theorem in probabilistic form. Same way as in chapter 1 it is derived from a corresponding two-dimensional limit theorem.

Chapter 3:

In chapter 3 we discuss important applications of trimming in statistics. As an example the detection of possible changes in location model is considered. Under the null hypothesis the location parameter is constant for all elements of the sample and under alternative certain amount of changes occurs.

A very popular way to test one hypothesis against the other (see Csörgő and Horváth, [26], Aue and Horváth, [6]) is based on the CUSUM process.

It is well known that for i.i.d. random variables with finite second moment appropriately normalized CUSUM process converges weakly to Brownian bridge process. However in case of random variables in the domain of attraction of a stable law Aue, Berkes and Horváth (see [4]) shown that the limiting process is different. So far not much is known about this process, therefore Berkes, Horváth and Schauer (see [16]) suggested to apply trimming procedure to the CUSUM process.As a result, it turned out that trimmed CUSUM process for i.i.d. variables with heavy tails also converges weakly to Brownian bridge process.

CUSUM process is also used in the cases of dependent random variables. However most of the time it is assumed that the variables have high moments and the dependence in the sequence is usually quite weak (see e.g. Aue and Horváth [6]). Fama and Mandelbrot (see [37], [72], [73]) shown that heavy-tailed distributions have applications in economics, therefore investigation of them is of great importance.

In chapter 3 we study sequences of trimmed AR(1) variables with heavy tails. We formulate a limit theorem for CUSUM process based on this sequence.

Further we develop two types of tests to detect changes in the location parameters. We study two types of statistics. The maximally selected CUSUM process we estimate the long run variance by kernel estimators. We also propose ratio statistics which do not depend on the long run variances. Monte Carlo simulations illustrate the the limit results can be used even in case of small and moderate sample sizes.

Chapter 1

On the Central Limit Theorem for modulus trimmed sums

1.1 Introduction

For convenience in the beginning of the chapter we shall review and discuss some classical and new results on the subject. Some of them have already been mentioned in the introduction. Here we are giving a more detailed overview of the results which are directly connected to those ones of this chapter.

Let $X_1, X_2, \ldots$ be independent, identically distributed random variables in the domain of attraction of a stable law G with parameter $0 < \alpha < 2$. That is, assume that the partial sums $S_n = \sum_{k=1}^{n} X_k$ satisfy

$$(S_n - b_n)/a_n \xrightarrow{d} G \tag{1.1}$$

with suitable norming and centering sequences $\{a_n\}$, $\{b_n\}$. The necessary and sufficient condition for (1.1) is that F, the distribution function of X_1, satisfies

$$1 - F(x) + F(-x) = x^{-\alpha} L(x), \qquad x > 0 \tag{1.2}$$

and

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} \to p, \qquad \frac{F(-x)}{1 - F(x) + F(-x)} \to q \qquad (x \to \infty) \tag{1.3}$$

where L is a function slowly varying at ∞ and $p, q \geq 0$, $p + q = 1$. We use the definition of a slowly varying function given by Feller (1971,[38]).

Definition 1. *We say that the function L varies slowly at infinity if it satisfies the following condition*

$$\frac{L(tx)}{L(t)} \to 1 \qquad as \quad t \to \infty$$

for every $x > 0$.

Similarly,

Definition 2. *We say that the function L varies slowly at zero if it satisfies the following condition*

$$\frac{L(tx)}{L(t)} \to 1 \qquad as \quad t \to 0$$

for every $x > 0$.

Further whenever we mention slowly varying function without specification we mean that the function varies slowly at infinity. Immediately we give several statements for slowly varying functions from Bingham, Goldie and Teugels (1989, [20]) which will be extensively used in the further proofs.

Theorem 1 (Bingham, Goldie and Teugels, 1989)**.** *If L is slowly varying, $L(tx)/L(t) \to 1$ uniformly in x on each compact set $(0, \infty)$.*

Theorem 2 (Karamata's theorem; direct half)**.** *Let L vary slowly and be locally bounded on $[X, \infty)$. then*

(i) for any $\sigma \geq -1$,

$$x^{\sigma+1} L(x) / \int\limits_X^x t^\sigma L(t) dt \to \sigma + 1 \qquad (x \to \infty)$$

(ii) for any $\sigma < -1$ (and for $\sigma = -1$ if $\int^{\infty} L(t)dt < \infty$)

$$x^{\sigma+1}L(x)/\int\limits_{x}^{x} t^{\infty}L(t)dt \to -(\sigma+1) \qquad (x \to \infty).$$

Proposition 1 (Bingham, Goldie and Teugels, 1989). *Following statements hold:*

(i) If L varies slowly $\log(L(x))/\log x \to 0$ as $x \to \infty$.

(ii) If L varies slowly, so does $(L(x))^{\alpha}$ for every $\alpha \in \mathbb{R}$.

(iii) If L_1, L_2 vary slowly, so does $L_1(x)L_2(x)$, $L_1(x)+L_2(x)$ and (if $L_2(x) \to \infty$ as $x \to \infty$) $L_1(L_2(x))$.

(iv) If $L_1, \ldots, L_k$ vary slowly and $r(x_1, \ldots, x_k)$ is a rational function with positive coefficients, $r(L_1(x), \ldots, L_k(x))$ varies slowly.

(v) If L varies slowly and $\alpha > 0$,

$$x^{\alpha}L(x) \to \infty, \qquad x^{-\alpha}L(x) \to 0, \quad (x \to \infty).$$

Proposition 2 (Bingham, Goldie and Teugels, 1989). *If L is slowly varying, X is so large that $L(x)$ is locally bounded on $[X, \infty)$, and $\alpha > -1$, then*

$$\int\limits_{X}^{x} t^{\alpha}L(t)dt \sim \{x^{\alpha+1}L(x)\}/(\alpha+1) \qquad (x \to \infty).$$

Proposition 3 (Bingham, Goldie and Teugels, 1989). *Let F satisfy* (1.2),(1.3). *Then for $V(x) = \int_{-x}^{x} t^2F(t)dt$ the following holds*

$$\frac{x^2\{1 - F(x) + F(-x)\}}{V(x)} \to \frac{2-\alpha}{\alpha} \qquad (x \to \infty).$$

Another proposition regarding the properties of slowly varying functions in the stable distributions was given in the paper by Csörgő, Csörgő, Horváth and Mason ([27], 1986).We will need it in further arguments as well.

Proposition 4 (Csörgő, Csörgő, Horváth and Mason, 1986). *There exist a function l slowly varying at infinity, an $0 < \alpha < 2$ and a $0 \le p \le 1$ such that the conditions* (1.2), (1.3) *hold if and only if for some function L slowly varying near zero we have*

$$K(1-u) = u^{-1/\alpha} L(u), \tag{1.4}$$

$$\lim_{u \downarrow 0} Q_1(1-u)/K(1-u) = q^{1/\alpha}, \tag{1.5}$$

and

$$\lim_{u \downarrow 0} Q_2(1-u)/K(1-u) = p^{1/\alpha}, \tag{1.6}$$

where Q is the quantile function of our sample, $Q_1(u) = (-Q(1-u) \vee 0)$, $Q_2(u) = Q(u) \vee 0$, α and p are the same as in (1.2) *and* (1.3)

In contrast to the case of finite variances, the contribution of extremal terms in the partial sums S_n is not negligible and dropping a single term can change the asymptotic behavior of the sum. Let $X_{n,1} \le X_{n,2} \le \ldots \le X_{n,n}$ be the order statistics of $(X_1, X_2, \ldots, X_n)$ and put for $d \ge 1$

$$S_n^{(d)} = \sum_{j=d+1}^{n-d} X_{n,j}. \tag{1.7}$$

For fixed d, LePage et al. ([65]) determined the asymptotic distribution of the trimmed sum $S_n^{(d)}$ by using the following elementary property of order statistics: let G be a function such that $F = 1 - G$, where F is the common distribution function of our sequence $X_1, X_2, \ldots$; then there is a sequence $E_1, \ldots, E_2$ of exponential random variables with unit mean for which

$$(X_{n,1}, \ldots, X_{n,n}) =_d [G^{-1}(\Gamma_1/\Gamma_{n+1}, \ldots, \Gamma_n/\Gamma_{n+1})], \qquad n \ge 1$$

where

$$\Gamma_k = E_1 + \ldots E_k \qquad k \ge 1$$

and

$$G^{-1}(u) = \inf\{y : G(y) \le u\}, \qquad 0 < u < 1.$$

This property of the order statistics clarified in which way several big summands determine the limiting distribution.

Later Csörgő, Horváth and Mason ([26], 1986) proved the following under

$$d_n \to \infty, \quad d_n/n \to 0 \tag{1.8}$$

Theorem 3 (Csörgő, Horváth and Mason, 1986). *Let $X_1, X_2 \dots, X_n$ be independent and identically distributed random variables (rv's) with common distribution function F, and $X_{1,n} < \cdots < X_{n,n}$ denote the respective order statistics. Assume F is in the domain of attraction of a stable law. Then for any sequence of positive integers d, such that $1 \le d_n \le n - d_n < n$ and $d_n \to \infty$ and $d_n/n \to 0$ as $n \to \infty$,*

$$A_n(d_n)\left\{\sum_{i=d_n+1}^{n-d_n} X_{i,n} - C_n(d_n)\right\} \xrightarrow{\mathcal{D}} N(0,1)$$

where $A_n(d_n) = 1/(n^{1/2}\sigma(d_n/n))$ with

$$\sigma^2(s) = \int_1^{1-s}\int_1^{1-s} (u \wedge v - uv)dQ(u)dQ(v)$$

$$Q(u) = \inf\{x : F(x) \ge u\}, 0 < u < 1.$$

and

$$C_n(d_n) = n \int_{d_n/n}^{1-d_n/n} Q(u)du.$$

These results give a remarkable picture on the partial sum behavior of i.i.d. sequences in the domain of attraction of a non-normal stable law. They show that the contribution of d_n extremal terms under (2.6) already gives the stable limit distribution of the total partial sum S_n and the contribution of the remaining elements will be an asymptotically normal variable with magnitude negligible compared with S_n.

The previous results describe the effects of the extremal elements of an i.i.d. sample on their partial sum. Note, however, that other kinds of trimming lead to different phenomena. For $1 \le d \le n$ let $\eta_{d,n}$ denote the d-th largest of $|X_1|, \dots, |X_n|$ and let

$$^{(d)}S_n = \sum_{k=1}^{n} X_k I\{|X_k| \le \eta_{d,n}\}. \tag{1.9}$$

If the distribution of X_1 is continuous, then $|X_1|, |X_2|, \dots$ are different with probability 1, and thus $^{(d)}S_n$ coincides with the usual modulus trimmed sum obtained by discarding from S_n the $d-1$ elements with the largest moduli. Griffin and Pruitt ([44], 1987) showed that if X_1 has a symmetric distribution, then $^{(d_n)}S_n$ is asymptotically normal, i.e. provided necessary and sufficient conditions, for any $d_n \to \infty$, $d_n/n \to 0$, but this is generally false in the nonsymmetric case.

The purpose of this chapter is to describe the asymptotic distribution of $^{(d_n)}S_n$ in the general case. Put

$$H(t) = P(|X| \ge t) \quad \text{and} \quad m(t) = EXI\{|X| \le t\},$$

and let $H^{-1}(t) = \inf\{x : H(x) \le t\}$ $(0 < t < 1)$ denote the generalized inverse of H. Our main result is the following.

Theorem 4. *Let $X_1, X_2, \dots$ be i.i.d. random variables with distribution function F satisfying (1.2), (1.3) and assume that (2.6) holds. Then we have*

$$\frac{1}{A_n} \sum_{i=1}^{[nt]} \left(X_i I\{|X_i| \le \eta_{d,n}\} - m(\eta_{d,n}) \right) \xrightarrow{\mathcal{D}[0,1]} W(t) \tag{1.10}$$

where

$$A_n^2 = \frac{\alpha}{2-\alpha} d (H^{-1}(d/n))^2 \tag{1.11}$$

and W is the Wiener process.

Theorem 20 shows that allowing a random centering factor, the modulus trimmed CLT holds for continuous i.i.d. variables under exactly the same conditions as under ordinary trimming. If F is not continuous, the sample $(X_1, \ldots, X_n)$ may contain equal elements with positive probability; according to the definition in Griffin and Pruitt ([44], 1987), 'ties' between elements with equal moduli are broken according to the order in which the variables occur in $(X_1, \ldots, X_n)$. But no matter how we break the ties, it may happen that from a set of sample elements with equal moduli some are discarded and others are not, which is rather unnatural from the statistical point of view, since trimming is mainly used to improve the performance of statistical procedures by removing large elements from the sample. The definition of ${}^{(d)}S_n$ in (1.9) resolves this difficulty and leads to satisfactory asymptotic results in the general case.

Theorem 20 enables one to give, among others, change point tests for heavy tailed processes, while the standard CUSUM test fails under infinite variances. A fairly precise characterization of the modulus trimmed CLT with nonrandom centering and norming factors was given in Berkes and Horváth ([14], 2012). For example one of the results of that chapter is the following

Theorem 5 (Berkes and Horváth, 2012)**.** *Assume that*

$$1 - F(x) + F(-x) = x^{-\alpha} L(x), \tag{1.12}$$

and the following condition for density $f(x)$ *is satisfied*

$$0 < \liminf_{x \to \infty} \frac{xf(x)}{1 - F(x)} \leq \limsup_{x \to \infty} \frac{xf(x)}{1 - F(x)} < \infty \tag{1.13}$$

Let

$$\frac{1 - F(x)}{1 - F(x) + F(-x)} = p + O(x^{-\beta}), \quad asx \to \infty, \tag{1.14}$$

where $0 < p < 1, 0 < \beta < 2, \beta > 0$. *Let*

$$\gamma_1 = \frac{\beta}{\max(1 + 2\alpha, \beta + \alpha/2)}, \quad \gamma_2 = \frac{\beta}{\beta + \alpha/4}.$$

Under the condition

$$d_n \to \infty, \text{ and } d_n = O(n^\gamma) \text{ with } \gamma < \gamma_1 \tag{1.15}$$

we have the central limit theorem

$$\frac{1}{A_n}\{T_{n,d} - B_n\} \xrightarrow{\mathcal{D}} N(0,1)$$

where $A_n > 0$ *and* B_n *are numerical sequences. On the other hand, for any* $0 < \alpha < 2$, $\beta > 0$ *there exists a distribution function* F *satisfying (1.12), (1.13) and (1.14) for some* $0 < p < 1$ *such that with the choice*

An interesting fact established in this paper is that paradoxally, as the upper bound on the power in (1.15) looks somewhat unnatural, the increase in the amount of trimmed terms does not lead to the improvement of CLT behaviour. In the same paper it was shown what happens if the expression on the left hand side of (1.14) is constant for $x > x_0$ or if its convergence to p is slower than any polynomial.

Under additional technical assumptions on the distribution function of X_1 and on the growth speed of d_n, i.e. if $\lim_{n\to\infty} d(n)/(\log n)^{7+\varepsilon} = \infty$ with some $\varepsilon > 0$, Theorem 20 was proved in Berkes, Horváth and Schauer ([16], 2011) with a fairly complicated argument. The proof of Theorem 20 is much simpler and extends to dependent samples as well, as we will show in chapter 2. Let

$$\hat{A}_n^2 = \sum_{i=1}^{n} X_i^2 I\{|X_i| \le \eta_{d,n}\} - \frac{1}{n}\left(\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right)^2.$$

Berkes, Horváth and Schauer ([14], 2011) showed that under the conditions of Theorem 20 we have that

$$\hat{A}_n / A_n \xrightarrow{P} 1$$

and therefore Theorem 20 yields

$$\frac{1}{\hat{A}_n}\left(\sum_{i=1}^{[nt]} X_i I\{|X_i| \le \eta_{d,n}\} - \frac{[nt]}{n}\sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{d,n}\}\right) \xrightarrow{\mathcal{D}[0,1]} B(t),$$

where $B(t) = W(t) - tW(1)$ denotes a Brownian bridge. Hence standard CUSUM techniques can be used to detect changes in the mean and/or location when in the case of observations without second moments, observations with modulus larger than $\eta_{d,n}$ are excluded from the sample.

Let

$$U_n(t,s) = \sum_{i=1}^{[nt]} \left(X_i I\{|X_i| \le sH^{-1}(d/n)\} - EX_i I\{|X_i| \le sH^{-1}(d/n)\} \right) \qquad (s \ge 0, t \ge 0).$$

We will deduce (2.15) from the following two-dimensional limit theorem.

Theorem 6. *Let $X_1, X_2, \ldots$ be i.i.d. random variables with distribution function F satisfying (1.2), (1.3) and assume that (2.6) holds. Then*

$$\frac{1}{A_n} U_n(t,s) \longrightarrow W(t, s^{2-\alpha}) \quad \text{weakly in } \mathcal{D}([0,1] \times [1/2, 3/2]),$$

where A_n is defined by (2.14) and $\{W(x,y),\, x \ge 0, y \ge 0\}$ is a two-parameter Wiener process.

Note that by Kiefer ([60], 1972) we have

$$\frac{\eta_{d,n}}{H^{-1}(d/n)} \xrightarrow{P} 1. \tag{1.16}$$

Since the limit process in Theorem 31 has continuous trajectories a.s., Billingsley's argument on the random change of time (see [19], 1968), p. 144-145 we use (1.16) for a substitution and hence

$$\frac{1}{A_n} U_n(t, \eta_{d,n}/H^{-1}(d/n)) \xrightarrow{\mathcal{D}[0,1]} W(t,1)$$

which is exactly the functional CLT in (2.15), since $W(t,1)$ is a Wiener process. Thus Theorem 20 is a consequence of Theorem 31.

1.2 Proof of the results

For the proof of the tightness we will need the theorem stated by Bickel and Wichura ([17], 1971)

Theorem 7 (Bickel and Wichura, 1971)**.** *Let $T = T_1 \times \cdots \times T_q$ for some positive q, where for each $1 \le i \le q$ T_i is a subset of $[0,1]$. Let $(X_t))t \in T$ be a stochastic process. Call block B a subset of T if $B = \prod_{i=1}^{q}(s_i, t_i]$ with all s_i, t_i, $i = 1,\dots,q$ in T. Then define $X(B)$ as follows*

$$X(B) = \sum_{\varepsilon_1=0,1} \cdots \sum_{\varepsilon_q=0,1} (-1)^{q-\sum_p \varepsilon_p} X(s_1 + \varepsilon_1(t_1 - s_1), \dots, s_q + \varepsilon_q(t_q - s_q))$$

pth face of the block B is $\prod_{\rho \ne p}(s_p, t_p]$.Disjoint Blocks B and C are p-neighbours if they abut and have he same p-face. For each pair of neighbouring blocks B and C define

$$m(B, C) = \min\{|X(B)|, |X(C)|\}. \tag{1.17}$$

Now let $\{X_n(t), t \in [T\}$ be processes vanishing along the lower boundary of T and there exist constants $\beta > 1$, $\gamma > 0$ and a finite nonnegative measure μ on T with continuous marginals such that for each n

$$P\{m(B, C) \ge \lambda\} \le \lambda^{-\gamma}(\mu(B \cap C))^{\beta}, \tag{1.18}$$

then the sequence of processes $\{X_n(t)\}$ is tight.

One can easily see that 1.18 follows by Chebyshev's inequality from its moment version, that is

$$E(|X(B)|^{\gamma_1}|X(C)|^{\gamma_2}) \le (\mu(B))^{\beta^1}(\mu(C))^{\beta_2}, \tag{1.19}$$

where $\gamma = \gamma_1 + \gamma_2$ and $\beta = \beta_1 + \beta_2$. However for our purposes we need these statements only in two-dimensional case. Given a process $Y(s, t)$ defined on a rectangle $H = [a, b] \times [c, d]$, let $Y(H)$ denote the increment of Y over H, i.e.

$Y(H) = Y(a,c) - Y(a,d) + Y(b,d) - Y(b,c)$. Using 1.19 and reformulating theorem 7 for only two dimensions we get the following lemma.

Lemma 1. *Let* $\{Y_n(t,s), n \geq 1\}$ *be processes defined on a rectangle* $[a,b] \times [c,d] \subset [0,\infty)^2$ *and assume that for some* $\gamma > 0$

$$E|Y_n(B)|^{\gamma}|Y_n(C)|^{\gamma} \leq \mu(B)\mu(C), \tag{1.20}$$

where μ *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* $Y_n(t,s)$ *is piecewise constant in* t*, i.e. there exists a finite set* $H_n \subset [a,b]$ *such that* $Y_n(t,s)$ *is constant in the left closed intervals determined by the elements of* $H_n \cup \{a\} \cup \{b\}$*, then it suffices to verify (1.20) for rectangles* $[t_1,t_2] \times [s_1,s_2]$ *where* $t_1, t_2 \in H_n$*.*

As is shown in Proposition 4, the conditions of Theorem 31 imply that $H^{-1}(t) = t^{-1/\alpha}\ell(t)$ $(0 < t < 1)$, where ℓ is slowly varying at 0. Then by (2.14) we have

$$A_n^2 \sim \frac{\alpha}{2-\alpha} d\,(n/d)^{2/\alpha}\ell^2(d/n) \quad \text{as } n \to \infty \tag{1.21}$$

where $a_n \sim b_n$ means $a_n/b_n \to 1$ as $n \to \infty$.

Lemma 2. *If the conditions of Theorem 31 are satisfied, then for any* $p \geq 2$ *and any fixed* $0 \leq a < b < \infty$ *we have*

$$E|X_1|^p I\{aH^{-1}(d/n) < |X_1| \leq bH^{-1}(d/n)\} \sim \frac{\alpha}{p-\alpha}(b^{p-\alpha} - a^{p-\alpha})\ell^p(d/n)(n/d)^{(p-\alpha)/\alpha} \tag{1.22}$$

as $n \to \infty$*. Also, if* $b > 0$*, then*

$$E|X_1|I\{|X_1| \leq bH^{-1}(d/n)\} = \begin{cases} O((n/d)^{(1-\alpha)/\alpha}\ell(d/n)) & \text{if } \alpha < 1, \\ O((n/d)^{\varepsilon}) & \text{if } \alpha = 1, \\ O(1) & \text{if } \alpha > 1 \end{cases} \tag{1.23}$$

for any $\varepsilon > 0$*.*

Proof. Assume first $p \geq 2$, $0 < a < b < \infty$. Clearly the left hand side of (1.22) equals

$$-\int_{aH^{-1}(d/n)}^{bH^{-1}(d/n)} t^p dH(t) = \int_{H(bH^{-1}(d/n))}^{H(aH^{-1}(d/n))} H^{-1}(u)^p du. \tag{1.24}$$

(Note that H is non-increasing and thus the left hand side of (1.24) is nonnegative.) Since H is regularly varying with exponent $-\alpha$, we have

$$H(aH^{-1}(d/n)) \sim a^{-\alpha}(d/n), \qquad H(bH^{-1}(d/n)) \sim b^{-\alpha}(d/n) \qquad \text{as } n \to \infty.$$

Thus using the uniform convergence theorem for regularly varying functions (see Theorem 1; note that we actually need the analogous result for regular variation at 0), we see that for $n \to \infty$ we have, uniformly for all u in the interval of integration of the second integral in (1.24),

$$H^{-1}(u) = u^{-1/\alpha}\ell(u) \sim u^{-1/\alpha}\ell(d/n).$$

Thus the integral equals

$$(1 + o(1)) \int_{(1+o(1))b^{-\alpha}(d/n)}^{(1+o(1))a^{-\alpha}(d/n)} u^{-p/\alpha}\ell^p(d/n)\, du, \tag{1.25}$$

which yields the right hand side of (1.22) after a simple calculation, since $p \neq \alpha$. If $a = 0$, then the upper limit in the integral on the right hand side of (1.24) and thus also in (1.25) becomes $H(0) = 1$ and by using Theorem 2 we get the right hand side of (1.22) with $a = 0$.

In the case of (1.23), instead of the integral in (1.25) we get

$$\int_{(1+o(1))b^{-\alpha}(d/n)}^{1} u^{-1/\alpha}\ell(u)\, du. \tag{1.26}$$

By Proposition 1(i) we have $\ell(u) = O(u^{-\varepsilon})$ as $u \to 0$ for any $\varepsilon > 0$ which shows that for $\alpha > 1$ the integral $\int_0^1 u^{-1/\alpha}\ell(u)\, du$ converges and thus the expression (1.26) is $O(1)$. Using the same estimate for $\ell(u)$ for $\alpha = 1$ we get the second bound in (1.23). Finally, for $\alpha < 1$ Theorem 2 yields the first bound in (1.23), completing the proof of Lemma 2.

Before we start the proof of Theorem 31 we will remind to the reader the fundamental result from classical pobability theory, that is Cramér-Wold theorem.

Theorem 8 (Cramér-Wold)**.** *Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be a sequence of random $k-$vectors. Then $\boldsymbol{X}_n \to d\boldsymbol{X}$ if and only if $\boldsymbol{a}^T\boldsymbol{X}_n \to d\boldsymbol{a}^T\boldsymbol{X}$ for all $\boldsymbol{a} \in \mathbb{R}^k$.*

Proof of Theorem 31. Let $\Gamma(t,s)$ denote the limit process in Theorem 31 and 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

$$Z = \sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}\Gamma([t_{m-1},t_m]\times[s_{j-1},s_j])$$

for all $M \geq 1, J \geq 1$, real coefficients $\mu_{m,j}$, $1/2 \leq s_1 < s_2 < \ldots < s_J \leq 3/2$, $0 < t_1 < \ldots < t_M = 1$, $t_0 = s_0 = 0$. Clearly, Z is a centered normal r.v. and

$$EZ^2 = \sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}^2(s_j^{2-\alpha} - s_{j-1}^{2-\alpha})(t_m - t_{m-1}). \tag{1.27}$$

We claim that

$$Q_n \xrightarrow{d} Z \qquad \text{for all considered values of } M, J, \mu_{m,j}, t_m, s_j. \tag{1.28}$$

Since the processes U_n and Γ are equal to 0 on the boundary of the first quadrant, we have

$$U_n(t_m,s_j) = \sum_{m=1}^{M}\sum_{j=1}^{J}U_n([t_{m-1},t_m]\times[s_{j-1},s_j])$$

and the same relation holds for Γ. Thus (2.24) implies

$$\frac{1}{A_n}\sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}^*U_n(t_m,s_j) \xrightarrow{d} \sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}^*\Gamma(t_m,s_j)$$

for arbitrary real coefficients $\mu^*_{m,j}$ and this, by the Cramér-Wold device (theorem 8), implies the convergence of the finite-dimensional distributions in Theorem 31.

In view of the definition of $U_n(t,s)$ in Section 1 we have

$$U_n([t_{m-1},t_m]\times[s_{j-1},s_j]) = \sum_{i=[nt_{m-1}]+1}^{[nt_m]} (v_{i,j} - Ev_{i,j})$$

where

$$v_{i,j} = X_i I\{s_{j-1}H^{-1}(d/n) < |X_i| \le s_j H^{-1}(d/n)\}.$$

Thus relation (2.24) can be written equivalently as

$$\frac{1}{A_n}\sum_{k=1}^{n}(z_{k,n} - Ez_{k,n}) \xrightarrow{d} N(0, EZ^2) \qquad \text{as } n\to\infty, \tag{1.29}$$

where

$$z_{k,n} = \sum_{j=1}^{J} \mu_{m,j} X_k I\{s_{j-1}H^{-1}(d/n) < |X_k| \le s_j H^{-1}(d/n)\}, \quad [nt_{m-1}]+1 \le k \le [nt_m].$$

Since the terms in the last sum are random variables with disjoint support, we get from Lemma 2

$$Ez^2_{k,n} = (1+o_n(1))\frac{\alpha}{2-\alpha}(n/d)^{(2-\alpha)/\alpha}\ell^2(d/n)\sum_{j=1}^{J}\mu^2_{m,j}(s_j^{2-\alpha}-s_{j-1}^{2-\alpha}), \quad [nt_{m-1}]+1 \le k \le [nt_m]$$

and similarly

$$Ez^4_{k,n} = (1+o_n(1))\frac{\alpha}{4-\alpha}(n/d)^{(4-\alpha)/\alpha}\ell^4(d/n)\sum_{j=1}^{J}\mu^4_{m,j}(s_j^{4-\alpha}-s_{j-1}^{4-\alpha}), \quad [nt_{m-1}]+1 \le k \le [nt_m].$$

Thus using $d = d_n \to \infty$ we get by a simple calculation

$$\lim_{n\to\infty}\frac{\sum_{k=1}^{n} Ez^4_{k,n}}{\left(\sum_{k=1}^{n} Ez^2_{k,n}\right)^2} = 0. \tag{1.30}$$

On the other hand, the previous asymptotics for $Ez^2_{k,n}$ and the statement of Lemma 2 for $p=1$ imply

$$E^2|z_{k,n}| = o_n(1)Ez^2_{k,n}, \qquad 1 \le k \le n$$

and thus by Minkowski's inequality

$$E|z_{k,n}-Ez_{k,n}|^2 = (1+o_n(1))Ez_{k,n}^2, \qquad E|z_{k,n}-Ez_{k,n}|^4 = (1+o_n(1))Ez_{k,n}^4. \quad (1.31)$$

Thus (1.30) remains valid if we replace $z_{k,n}$ with $z_{k,n} - Ez_{k,n}$. Further by (1.21) and (2.23)

$$\sum_{k=1}^{n} Ez_{k,n}^2 = (1+o_n(1))\frac{\alpha}{2-\alpha}n(n/d)^{(2-\alpha)/\alpha}\ell^2(d/n)\sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}^2(s_j^{2-\alpha}-s_{j-1}^{2-\alpha})(t_m-t_{m-1})$$
$$= (1+o_n(1))A_n^2 EZ^2.$$

The last relation, together with (1.30), (1.31) and Ljapunov's CLT for triangular arrays, implies (2.25).

Next we prove tightness in Theorem 31. Consider two pairs of sets $B_{11} = [t_1, t] \times [s_1, s]$, $B_{12} = [t_1, t] \times [s, s_2]$ and $B_{11} = [t_1, t] \times [s_1, s]$, $B_{21} = [t, t_2] \times [s_1, s]$, where $t_1 < t < t_2$, $s_1 < s < s_2$. In view of Lemma 1, it suffices to show that

$$E\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{ij})\right|^2 \le C\mu(B_{11})\mu(B_{ij}), \quad (1.32)$$

holds for each $ij \in \{12, 21\}$ with some constant $C > 0$. Moreover, since $U_n(t, s)$ is constant on intervals $k/n \le t < (k+1)/n$, by the last statement of Lemma 1 we may assume that nt, nt_1 and nt_2 are all integers. Using the independence of the X_j's, relation (1.21), Lemma 2 and the fact that the function $x^{2-\alpha}$ has a bounded derivative on $[1/2, 3/2]$, we get

$$E\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{21})\right|^2$$
$$= E\left(\frac{1}{A_n}\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2$$
$$\times\left(\frac{1}{A_n}\sum_{i=nt+1}^{nt_2}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2$$
$$= E\left(\frac{1}{A_n}\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n) < |X_i| \le sH^{-1}(d/n)\} - m_i\right)\right)^2$$

$$\times E\left(\frac{1}{A_n}\sum_{i=nt+1}^{nt_2}\left(X_iI\{s_1H^{-1}(d/n)<|X_i|\le sH^{-1}(d/n)\}-m_i\right)\right)^2$$

$$\le \frac{1}{A_n^4}\left(\sum_{i=nt_1+1}^{nt}EX_i^2I\{s_1H^{-1}(d/n)<|X_i|\le sH^{-1}(d/n)\}\right) \tag{1.33}$$

$$\times\left(\sum_{i=nt+1}^{nt_2}EX_i^2I\{s_1H^{-1}(d/n)<|X_i|\le sH^{-1}(d/n)\}\right)$$

$$\le C_1(t-t_1)(t_2-t)(s^{2-\alpha}-s_1^{2-\alpha})^2\le C_2(t-t_1)(t_2-t)(s-s_1)^2$$

$$= C_2\mu(B_{11})\mu(B_{21}),$$

where

$$m_i=m_i(s_1,s)=EX_iI\{s_1H^{-1}(d/n)<|X_i|\le sH^{-1}(d/n)\}$$

and C_1, C_2 are positive constants. On the other hand,

$$E\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{12})\right|^2$$

$$=\frac{1}{A_n^4}E\left(\sum_{i=nt_1+1}^{nt}\left(X_iI\{s_1H^{-1}(d/n)<|X_i|\le sH^{-1}(d/n)\}-m_i^{(s_1,s)}\right)\right)^2$$

$$\times\left(\sum_{i=nt_1+1}^{nt}\left(X_iI\{sH^{-1}(d/n)<|X_i|\le s_2H^{-1}(d/n)\}-m_i^{(s,s_2)}\right)\right)^2$$

$$=\frac{1}{A_n^4}E\left(\sum_{i=nt_1+1}^{nt}(X_i^{(s_1,s)}-m_i^{(s_1,s)})\right)^2\left(\sum_{i=nt_1+1}^{nt}(X_i^{(s,s_2)}-m_i^{(s,s_2)})\right)^2 \tag{1.34}$$

where we put

$$X_i^{(u,v)}=X_iI\{uH^{-1}(d/n)<|X_i|\le vH^{-1}(d/n)\},\quad m_i^{(u,v)}=EX_i^{(u,v)}.$$

Expanding the product expectation in (2.33), we get the sum of all expressions

$$E(X_i^{(s_1,s)}-m_i^{(s_1,s)})(X_j^{(s_1,s)}-m_j^{(s_1,s)})(X_k^{(s,s_2)}-m_k^{(s,s_2)})(X_\ell^{(s,s_2)}-m_\ell^{(s,s_2)}), \tag{1.35}$$

where $nt_1+1\le i,j,k,\ell\le nt$. By the independence of the X_ν's, the product expectation in (1.35) equals 0 if one of the i,j,k,ℓ differs from the other three.

Thus it suffices to estimate the contribution of the terms where i, j, k, ℓ are pairwise equal, or all are equal. Assume first that $i = j$, $k = \ell$ and $i \neq k$; the other cases $i = k, j = \ell$, $i \neq j$ and $i = \ell, j = k$, $i \neq j$ can be handled similarly as the case $i = j = k = \ell$ below. Then X_i and X_k are independent, and thus using Lemma 2, the product expectation (1.35) becomes

$$E\left[(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2(X_k^{(s,s_2)} - m_k^{(s,s_2)})^2\right] = E(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2 E(X_k^{(s,s_2)} - m_k^{(s,s_2)})^2 \tag{1.36}$$

$$\leq E(X_i^{(s_1,s)})^2 E(X_k^{(s,s_2)})^2 \sim \frac{\alpha^2}{(2-\alpha)^2}(s^{2-\alpha} - s_1^{2-\alpha})(s_2^{2-\alpha} - s^{2-\alpha})\ell^4(d/n)(n/d)^{(4-2\alpha)/\alpha}$$

$$\leq C_3(s - s_1)(s_2 - s)\ell^4(d/n)(n/d)^{(4-2\alpha)/\alpha}.$$

The number of such pairs (i, k) is at most $(nt - nt_1)^2$ and thus dividing by A_n^4 and using (1.21) we get that the contribution of such terms (1.35) is not greater than

$$C_4(t - t_1)^2(s - s_1)(s_2 - s) = C_4\mu(B_{11})\mu(B_{12}).$$

Consider now the case $i = j = k = \ell$. In this case (1.35) becomes, expanding and introducing new letters to lighten the notations,

$$\begin{aligned} E&\left[(X_i^{(s_1,s)} - m_i^{(s_1,s)})^2(X_i^{(s,s_2)} - m_i^{(s,s_2)})^2\right] = E(\xi - m^{(1)})^2(\eta - m^{(2)})^2 \\ &= E\xi^2\eta^2 - 2m^{(2)}E\xi^2\eta + (m^{(2)})^2E\xi^2 - 2m^{(1)}E\xi\eta^2 + 4m^{(1)}m^{(2)}E\xi\eta \\ &- 2m^{(1)}(m^{(2)})^2E\xi + (m^{(1)})^2E\eta^2 - 2(m^{(1)})^2m^{(2)}E\eta + (m^{(1)})^2(m^{(2)})^2, \end{aligned} \tag{1.37}$$

where

$$\xi = X_i^{(s_1,s)},\ \eta = X_i^{(s,s_2)},\ m^{(1)} = E\xi,\ m^{(2)} = E\eta.$$

Clearly ξ and η have disjoint support and thus $\xi\eta = 0$, showing that the first, second, fourth and fifth term of the last sum in (1.37) are equal to 0. Thus the sum equals

$$(m^{(2)})^2E\xi^2 - 2m^{(1)}(m^{(2)})^2E\xi + (m^{(1)})^2E\eta^2 - 2(m^{(1)})^2m^{(2)}E\eta + (m^{(1)})^2(m^{(2)})^2$$

$$= (m^{(2)})^2 E\xi^2 - 2(m^{(1)})^2(m^{(2)})^2 + (m^{(1)})^2 E\eta^2 - 2(m^{(1)})^2(m^{(2)})^2 + (m^{(1)})^2(m^{(2)})^2.$$

By the Cauchy-Schwarz inequality we have $(m^{(1)})^2 \le E\xi^2$, $(m^{(2)})^2 \le E\eta^2$ and thus the absolute value of the last sum is at most $7E(\eta^2)E(\xi^2)$, which, apart from the coefficient, is exactly the third expression in (1.36), leading to the same estimate as there. The number of choices for i in (1.37) is $nt - nt_1 \le (nt - nt_1)^2$, so for the contribution of all terms in (1.37) we get the same estimate as for (1.36), i.e. $C_5\mu(B_{11})\mu(B_{12})$. Thus we proved (2.32) for $B_{ij} = B_{12}$ and the proof of Theorem 31 is completed.□

Chapter 2

On the extremal theory of the continued fractions

2.1 Introduction

In this chapter we extend the results of the previous chapter to the dependent samples and mainly discuss its applications to the theory of continued fractions. Therefore in this chapter we will give a brief overview on the main results in this theory in connection to our problem.

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

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

where the transformation $T : (0,1) \to [0,1)$ is defined by $Tx = \{1/x\}$; here $[\cdot]$ and $\{\cdot\}$ denote integral resp. fractional part. Let

$$\mu(E) = \frac{1}{\log 2} \int\limits_E \frac{1}{1+x} dx$$

be the Gauss measure on the class $\mathcal{B}$ of Borel subsets of $(0,1)$. It is known (see e.g. Billingsley ([18], 1965) that T is an ergodic transformation preserving the Gauss measure and thus with respect to the probability space $((0,1),\mathcal{B},\mu)$, $\{a_n(x), n \geq 1\}$ is a stationary ergodic sequence. Clearly, the set $\{a_1 = k\}$ is the interval $(1/(k+1), 1/k]$ and thus

$$\mu\{a_1 = k\} = \frac{1}{\log 2} \int\limits_{1/(k+1)}^{1/k} \frac{1}{1+x} dx = \frac{1}{\log 2} \log\left\{1 + \frac{1}{k(k+2)}\right\} \sim \frac{1}{\log 2}\frac{1}{k^2}.$$

(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 $F: \mathbb{N} \to \mathbb{R}$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} F(a_k(x)) = \frac{1}{\log 2}\sum_{j=1}^{\infty} F(j)\log\left\{1 + \frac{1}{j(j+2)}\right\} \qquad \text{a.e.} \tag{2.1}$$

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 a_k with respect to the uniform measure in $(0,1)$ converges to μ and asked for the speed of convergence. Kusmin ([63], 1928) showed that the convergence speed is $O(e^{-\lambda\sqrt{k}})$ and Lévy ([66], 1929) improved this to $O(e^{-\lambda k})$. Lévy's result implies that the sequence $\{a_k(x), k \geq 1\}$ is ψ-mixing with exponential rate, i.e.

$$\sup_{A\in\mathcal{F}_1^k, B\in\mathcal{F}_{k+n}^\infty} |\mu(A\cap B) - \mu(A)\mu(B)| \leq Ce^{-\lambda n}\mu(A)\mu(B)$$

with positive absolute constants C, λ, where $\mathcal{F}_r^s$ denotes the σ-field generated by the variables $\{a_k(x),\, r \leq k \leq s\}$.

Letting E denote expectation with respect to μ, we have $Ea_1 = \infty$ and correspondingly for $F(x) = x$ the right hand side of (2.1) is $+\infty$. Thus the partial sums $\sum_{k=1}^N a_k(x)$ grow faster than N. Lévy ([68], 1952) proved that

$$\frac{1}{N}\sum_{k=1}^{N} a_k(x) - \frac{\log N}{\log 2} \xrightarrow{d} G, \tag{2.2}$$

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

$$\exp\left(-it\log|t| - \frac{\pi|t|}{2\log 2}\right). \tag{2.3}$$

Remainder term estimates for the convergence in (2.2) were obtained by Heinrich in the following theorem.

Theorem 9. *For the continued fraction expansion $(a_n)_{n\in\mathbb{N}}$ by the Gaussian measure μ we have*

$$\left|\mu\left(\left\{\omega : \frac{1}{n}\sum_{k=1}^{n} a_k(\omega) - \frac{\ln n - \kappa}{\ln 2} - G_{1,1}\left(x, \frac{\pi}{2\ln 2}\right)\right\}\right)\right| \leq C_0\frac{(\ln n)^2}{n},$$

where C_0 does not depend on x and n and κ stands for Euler's constant being equal to

$$\lim_{n\to\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{n}-\ln n\right) = 0.57722..$$

This implies that

$$\lim_{N\to\infty}\frac{1}{N\log N}\sum_{k=1}^{N} a_k(x) = \frac{1}{\log 2} \qquad \text{in measure,} \tag{2.4}$$

a result obtained earlier by Khinchin ([59], 1935). Khinchin also noted that (2.4) cannot hold almost everywhere. Diamond and Vaaler ([35], 1986) showed that the obstacle to a.e. convergence in (2.4) 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 (2.4) by excluding the largest summand.

They proved namely

$$\lim_{N\to\infty}\frac{1}{N\log N}S_N^{(1)}(x) = \frac{1}{\log 2} \qquad \text{for almost all } x \tag{2.5}$$

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 (2.5) 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 (2.5).

Interestingly, similar results were obtained for St. Petersburg game. This is a game of chance in which a fair coin s tossed at each stage. The first time a tail appears (assume it happened at kth stage), the game ends and the player wins 2^k units of money. Obviously this game can be modeled by a random variable with infinite expectation.

An analogous to (2.5) result for the St. Petersburg game was proved by Csörgő and Simons ([28], 1996). For further analogies between continued fraction digits and the St. Petersburg game we refer to Vardi ([96], 1997). 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

$$d_N \to \infty, \qquad d_N/N \to 0 \tag{2.6}$$

so that the number of discarded terms is 'large', but is still negligible compared with N. The purpose of this chapter is to answer this question. Let

$$m(t) = \frac{1}{\log 2} \sum_{1 \le k \le t} k \log \left(1 + \frac{1}{k(k+2)}\right), \qquad t \ge 1. \tag{2.7}$$

We will prove the following result.

Theorem 10. *Let $d = d_N$ satisfy (2.6). Then we have*

$$\frac{S_N^{(d)} - Nm(\eta_{d,N})}{N/\sqrt{d}} \xrightarrow{d} N\left(0, (\log 2)^{-1}\right) \tag{2.8}$$

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 μ and variance σ^2.

Theorem 10 reduces the asymptotic study of $S_N^{(d)}$ to that of $\eta_{d,N}$, which is a much simpler problem. We will show in (2.36) that $\eta_{d,N} \sim N/d$ in probability and since $m(t) \sim (\log 2)^{-1} \log t$ as $t \to \infty$, Theorem 10 can be rewritten equivalently as

$$S_N^{(d)} = Nm(\eta_{d,N}) + (N/\sqrt{d})\zeta_N = (1 + o_P(1))\frac{1}{\log 2} N \log(N/d) + (N/\sqrt{d})\zeta_N, \tag{2.9}$$

where $\zeta_N \xrightarrow{d} N(0, 1/\log 2)$. Here and in the sequel, $\xrightarrow{P}$ will denote convergence in probability and $o_P(1)$ a quantity converging to 0 in probability. Relation (2.9) shows that $Nm(\eta_{d,N})$ is the main term in an asymptotic expansion of $S_N^{(d)}$. As a comparison, write Lévy's limit theorem (2.2) in the form

$$S_N = \frac{1}{\log 2} N \log N + N\zeta_N^*, \tag{2.10}$$

where ζ_N^* converges in distribution to the Cauchy variable with characteristic function (2.3). 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 (2.10) changed to Gaussian fluctuations around $Nm(\eta_{d,N})$ in (2.9). An immediate consequence of relation (2.9) is

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

under (2.6). If d grows slower than any power of N, i.e. $\log d / \log N \to 0$, then the last relation implies

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

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 to 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 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 ([29], 1952) obtained the following results (see also Arov and Bobrov [3], 1960).

Theorem 11 (Darling, 1952)**.** *Let $X_i \geq 0$ have an exponent $1 < \alpha < 2$ and such that S_n has a limiting stable distribution, then*

$$\lim_{n\to\infty} P\left\{S_n < yX_n^*\right\} = G(y)$$

where $G(y)$ has the characteristic function

$$\int_0^\infty e^{ity} dG(y) = \frac{e^{it}}{1 - \gamma\int_0^1 (e^{it\alpha} - 1)\frac{d\alpha}{\alpha^{\gamma+1}}}$$

Theorem 12. *[Darling, 1952] Let $X_i \geq 0$ have an exponent $1 < \alpha < 2$ and such that S_n has a limiting stable distribution, and let a sequence $\{c_n\}$ be determined by the relation $f(c_n) = 1/n$. Then $\mu = E(X_i)$ exists, and if $a_n = n/c_n$, we have*

$$\lim_{n\to\infty} P\left\{\frac{S_n}{X_n^*} < a_n x\right\} = 1 - e^{-(x/\mu)^\gamma}$$

The case $\alpha = 1$ is critical and is not covered neither in [3], nor in [29]. 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 10 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, Csörgő, Horváth and Mason (see Theorem 3) also showed that removing the d largest and d smallest elements from the partial sum, where (2.6) holds, the remaining sum $S_N^{(d)}$ becomes asymptotically normal. Our Theorem 10 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 ψ-mixing property of the transformation T above, many classical limit theorems for partial sums of independent random variables have been extended to continued fractions.

Series of remarkable results were established by Doeblin (see [36], 1940). Stackelberg (see [89], 1966) shown that the law of iterated logarithm holds for the sequence $\log a_i(x)$. Doeblin's result concerning the law of iterated logarithm was later generilized by Philipp and Stackelberg ([80], 1969). Gordin and Reznick ([42], 1970) provided the law of iterated logarithm for the denominators of continued fractions $q_n(x)$. They shown that

$$\limsup\{|\log q_n(x) - na|/(b\sqrt{2n\log\log n})\} = 1$$

for $n \to \infty$, where $a = \pi^2/(12\log 2)$ and $b > 0$ is a constant.

Ibragimov (see [51], 1960) established existence of the constants a and $\sigma > 0$ such that

$$\lim \operatorname{mes} \mathcal{E}_{x\in(0,1)} \left[\frac{\log q_n(x) - na}{\sigma\sqrt{n}} < z\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-u^2/2} du.$$

Philipp ([79], 1988) shown that partial quotients a_i cannot satisfy a strong law of large numbers for any sequence $\{\sigma(n), n \geq 1\}$, such that $\sigma(n)/n$ is non-decreasing. In particular,

$$\lim_{N\to\infty} \frac{S_N}{\sigma(N)} = 0 \quad a.s.$$

or

$$\limsup \frac{S_N}{\sigma(N)} = \infty \quad a.s.,$$

according as $\sum \frac{1}{\sigma(N)} < \infty$ or $= \infty$. Iosifescu in his works ([53], 1977) using the methods of Galambos ([39], 1972) shown that the number $m_N(x, y)$ of the partial quatients $a_j(x)$, $1 \leq j \leq N$ such that $a_j(N) > yN$ with respect to measure Q. Last one is supposed to be absolutely continuous with respect to the Gauss measure. Later he gave a large survey on the results in metric theory of contiued fractions, where he takes Doeblin's paper ([36], 1940) as a starting point ([54], 1990).

Samur ([86], 1989) applied his preceding results for mixing random variables ([85], 1984) to the sequences of partial quotients and obtained certain functional limit theorems for quantities directly connected to the continued fraction expansion. Later he shown that in case the variance of the partial sums is strictly positive, result due to Philipp and Stout ([81], 1975) can be applied to the sequence $\{u_j\}_{j\geq}$, where u_j is defined as follows

$$x_j(x) = [a_j(x), a_{j+1}(x)\dots],$$

$$y_j(x) = [a_j(x), a_{j-1}(x)\dots a_1(x)],$$

$$u_1 = x_1, \quad u_j = x_j + \frac{1}{y_{j-1}},$$

giving the final result

Theorem 13 (Samur, 1996). *Assume that $f : [0, \infty) \to \mathbb{R}$ is either*

$$f = I_{[b,\infty)} \qquad for\ some\ b > 1$$

or a function satisfying

$$|f(x) - f(y)| \leq K|x - y| \quad (x, y > 1) \quad for\ some\ K > 0$$

and

$$\int_1^\infty |f(x)|^{2+\delta} x^{-2} dx < \infty \qquad for\ some\ \delta > 0.$$

$$Var(\sum_{j=1}^n f(u_j)) = n\sigma^2 + O(1) \quad as\ n \to \infty. \tag{2.11}$$

Define the partial sums process $\{S(t) : t \in [0, \infty)\}$ by

$$S(t) = \sum_{1\leq j\leq t} (f(u_j) - m), \qquad j \geq 0$$

with $m = \int_1^\infty f(x)\frac{1}{\log 2}\left(I_{[1,2]}(x)\frac{1}{x}\left(1 - \frac{1}{x}\right) + I_{[2,\infty)}(x)\frac{1}{x^2}\right) dx$.

If the constant σ^2 in (2.11) is strictly positive then the almost sure invariance principle holds for $\{S(t)\}$, that is, there exists a probability space and processes $\{S^*(t) : t \in [0, \infty)\}$, $\{X(t) : t \in [0, \infty)\}$ defined on it such that

(i) $\{S(t)\}$ and $\{S^*(t)\}$ has the same distribution,

(ii) $\{X(t)\}$ *is a standard Brownian motion,*

(iii) $|S^*(t) - X(\sigma t)| = O(t^{\frac{1}{2}-\varepsilon})$ *almost surely as* $t \to \infty$ *for some* $\varepsilon > 0$*(the constant implied by* O *being random).*

The same result holds for x_j, y_j *but with other values of* m.

Some of the most recent results were provided by Szewczak ([92], 2009). These concern the functionals of partial quotients themselve.

Theorem 14 (Szewczak, 2009)**.** *Let* $c_n \to \infty$ *be a sequence of positive numbers and* f *be a Borel function. In order that there exist a sequence* $\{b_n\}$ *such that*

$$c_n^{-1}(\sum_{i=1}^{n} f(a_k) - b_n) \to P0$$

it is necessary and sufficient that simultaneously

$$nP[|f(a_1)| > c_n] \to \infty, \quad \frac{n}{c_n^2} E[f^2(a_1) I_{[|f(a_1)| \le c_n]}] \to 0.$$

If the latter conditions are satisfied we can set $b_n = nE[f(a_1)I_{[|f(a_1)| \le c_n]}$

Another theorem by Szewczak is related to the domain of attraction of the normal law.

Theorem 15 (Szewczak, 2009)**.** *Let* f *be a Borel function.In order that there exist sequences* $\{c_n\}$ *and* $\{d_n\}$ *such that*

$$\lim_{n\to\infty} \mathcal{L}(c_n^{-1}(\sum_{i=1}^{n} f(a_k) - d_n)) = \mathcal{N}(0,1)$$

it is necessary and sufficient that the function $E[f^2(a_1)I_{[|f(a_1)| \le x]}]$ *is slowly varying. If the latter condition is satisfied we can set* $d_n = nE[f(a_1)]$.

Clearly it is very hard to list here all the remarkable results in the metric theory of continued fractions. For a more complete picture we refer the reader to the references of already mentioned works.

Using the extremal theory of dependent processes, (see e.g. Leadbetter and Rootzen [64], 1988), asymptotic properties of the (individual) extremes of $(a_1(x), \dots, a_n(x))$ can be established. And as the extreme values are of a particular interest to us we are giving here the results on the asymptotics of the largest digits, obtained by Galambos and Philipp in the 70s. Galambos (see [39], 1972) established the following limit theorem which is closely connected with Philipp's result on the mixing stochastic processes ([76], 1969)

Theorem 16 (Galambos, 1972). *Put $a_n = a_n(x)$ and let*

$$L_N = \max(a_1, a_2 \dots, a_N).$$

Then

$$\lim_{n\to\infty} P\left(\frac{L_N}{N} < \frac{y}{\log 2}\right) = \exp\left(\frac{1}{-y}\right). \tag{2.12}$$

Later Galambos ([40], 1974/75) provided the result for almost everywhere convergence.

Theorem 17 (Galambos, 1974/75). *For almost all x in $(0,1)$ (with respect to Lebesgue measure),*

$$\limsup_{N\to\infty} \frac{\log L_N - \log N}{\log\log N} = 1$$

and

$$\liminf_{N\to\infty} \frac{\log L_N - \log N}{\log\log N} = 0,$$

where L_N is the same as in theorem 16

Slightly later Philipp ([78], 1975/76) provided series of asymptotic results on the limit behavior of the largest digit of continued fraction expansion. He improved the result of theorem 16 introducing more precise asymptotics of the limit, namely replacing $o(1)$ on the right hand side of the expression (2.12) by $O(\exp(-(\log N)^{\delta}))$. He also proved the following Erdös's conjecture

Theorem 18 (Philipp, 1975/76)**.** *For allmost all* x

$$\liminf_{N\to\infty} \frac{L_N(x)\log\log N}{N} = \frac{1}{\log 2}$$

In the same work Philipp also gave a refinement of the theorem 18 by replacing the iterated logarithm by an arbitrary sequence ψ_N, satisfying certain constraints

Theorem 19 (Philipp, 1975/76)**.** *Let* ψ_N *be nonincreasing such that* $\psi_N N$ *is non-decreasing. Then*

$$L_N(x) \le \frac{\psi_N N}{\log 2}$$

finitely often or infinitely often for almost all x *according as*

$$\sum e^{-1/\psi_n} \frac{n}{\log\log n}$$

converges or diverges.

However, given all the variety of the results we have mentioned, no results on trimmed sums $S_N^{(d)}$ are known for continued fractions or, for that matter, for any dependent sequence of random variables. In Section 2.2, we will prove Theorem 10 in a probabilistic form and we will change the notation accordingly.

Theorem 20. *Let* $\{X_j,\, j \ge 1\}$ *be a strictly stationary sequence of positive, integer valued random variables with*

$$P(X_1 = k) \sim c_0 k^{-2} \quad as \quad k \to \infty \tag{2.13}$$

for some constant $c_0 > 0$*. Assume that* $\{X_j,\, j \ge 1\}$ *is* ψ*-mixing with rate* $\psi(n) = Ce^{-\lambda n}$ *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 (2.6). Let* $m(t) = EX_1 I\{X_1 \le t\}$ *and*

$$A_n = \sqrt{c_0}\, n/\sqrt{d}. \tag{2.14}$$

Then

$$\frac{1}{A_n}\sum_{i=1}^{[nt]} (X_i I\{X_i \le \eta_{d,n}\} - m(\eta_{d,n})) \overset{\mathcal{D}[0,1]}{\longrightarrow} W(t), \tag{2.15}$$

where W *is the Wiener process.*

Remark 2.1.1. If (X_n) is a sequence of positive random variables such that with probability one $X_1, X_2, \dots$ are different, then the sum $\sum_{i=1}^{[nt]} X_i I\{X_i \le \eta_{d,n}\}$ in (2.15) is obtained from $\sum_{i=1}^{[nt]} X_i$ by removing the $d-1$ largest terms and thus the conclusion of Theorem 20 for $t = 1$ reduces to that of Theorem 10. However, for integer valued variables X_n, $\eta_{d,n}$ can appear in the sequence $(X_1, \dots, X_n)$ more than once and in this case the number of terms of the sum $\sum_{i=1}^{[nt]} X_i$ exceeding $\eta_{n,d}$ can be smaller than $d-1$ and can actually be random. Thus, in a formal sense, Theorem 10 is not a special case of Theorem 20. However, using a simple perturbation argument Theorem 10 will be deduced from Theorem 20.

Let

$$U_n(t,s) = \sum_{i=1}^{[nt]} (X_i I\{X_i \le s(n/d)\} - EX_i I\{X_i \le s(n/d)\}) \qquad (t \ge 0, s \ge 0).$$

We will derive Theorem 20 from the following two-dimensional limit theorem.

Theorem 21. *Under the assumptions of Theorem 20 we have*

$$\frac{1}{A_n} U_n(t,s) \longrightarrow W(t,s) \quad \textit{weakly in } \mathcal{D}[0,1] \times \mathcal{D}[1/2, 3/2], \tag{2.16}$$

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

As we already noted, under (2.6) we have

$$\frac{\eta_{d,n}}{n/d} \longrightarrow 1 \qquad \text{in probability.}$$

Since the limit process $W(t,s)$ in (2.16) has continuous trajectories a.s., Theorem 21 and Billingsley's argument on the random change of time ([19], p. 144-145, 1968) imply that

$$\frac{1}{A_n} U_n(t, \eta_{d,n}/(n/d)) \overset{\mathcal{D}[0,1]}{\longrightarrow} W(t,1)$$

which is exactly the functional central limit theorem (2.15).

2.2 Preliminary lemmas

In the rest of the chapter 2 $(X_k)_{k\geq 1}$ denotes a sequence of random variables satisfying the conditions of Theorem 20 and $d = d_n$ denotes a sequence of positive integers satisfying (2.6). Moreover, c_0 denotes the constant in (2.13). 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, defined as in the beginning of section 1.2.

For our argument in this chapter we again will need Bickel and Wichura's theorem, to be more precise its special case which is given by lemma 1 in section 1.2.

For the purpose of dealing with dependents we would need also a correlation inequlity (see Bradley, [21], 2007)

Theorem 22. *Suppose $\mathcal{A}$ and $\mathcal{B}$ are $\sigma-$fields such that $\psi(\mathcal{A},\mathcal{B}) < \infty$. If $X \in \mathcal{L}^1(\mathcal{A})$ and $Y \in \mathcal{L}^1(\mathcal{B})$, then $E|XY| < \infty$, and*

$$|EXY - EXEY| \leq \psi(\mathcal{A},\mathcal{B})||X||_1||Y||_1.$$

However in our argument we are going to use not theorem 22 but the lemma below which easily follows from the latter one.

Lemma 3. *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

$$|EXY - EXEY| \le \psi(n)E|X|E|Y|.$$

Lemma 4. *Let $\mathcal{G}_k$ denote the σ-field generated by X_k, let $n_1 < \ldots < n_r$ be positive integers and let $Y_1, \ldots, Y_r$ be bounded r.v.'s such that Y_j is $\mathcal{G}_{n_j}$ measurable ($j = 1, 2, \ldots, r$). Then*

$$E|Y_1 \cdots Y_r| \le C_r E|Y_1| \cdots E|Y_r|,$$

where $C_r = (1 + \psi(1))^r$.

Proof. This is immediate by induction upon observing that by the previous lemma we have for any $1 \le j \le r-1$

$$E|Y_1 \cdots Y_{j+1}| \le E|Y_1 \cdots Y_j|E|Y_{j+1}| + \psi(1)E|Y_1 \cdots Y_j|E|Y_{j+1}| = (1+\psi(1))E|Y_1 \cdots Y_j|E|Y_{j+1}|.$$

Lemma 5. *For any $T \ge 3$ we have*

$$EX_1 I\{X_1 \le T\} \le C_1 \log T. \tag{2.17}$$

Moreover, for any fixed $0 \le s_1 < s_2$ we have

$$EX_1^2 I\{s_1(n/d) < X_1 \le s_2(n/d)\} \sim c_0(s_2 - s_1)(n/d) \qquad \text{as } n \to \infty \tag{2.18}$$

and for any fixed $0 < s_1 < s_2$ and sufficiently large n

$$EX_1 I\{s_1(n/d) < X_1 \le s_2(n/d)\} \le C_2(s_2 - s_1)/s_1. \tag{2.19}$$

Here C_1, C_2 are positive constants depending only on the sequence (X_k).

This is immediate from (2.13).

Lemma 6. *Let*

$$X_{k,n}^{(s_1,s_2)} = X_k I\{s_1(n/d) < X_k \le s_2(n/d)\} - EX_k I\{s_1(n/d) < X_k \le s_2(n/d)\}.$$

Then for any fixed $0 \le t_1 < t_2 \le 1$, $0 \le s_1 < s_2 < \infty$ *we have*

$$E\left(\sum_{k=nt_1+1}^{nt_2} X_{k,n}^{(s_1,s_2)}\right)^2 \sim c_0(n^2/d)(t_2-t_1)(s_2-s_1) \quad as \quad n\to\infty \tag{2.20}$$

provided nt_1, nt_2 *are integers. Moreover,*

$$E\left(\sum_{i=nt_1+1}^{nt_2} X_{i,n}^{(s_1,s_2)}\right)\left(\sum_{j=nt_1'+1}^{nt_2'} X_{j,n}^{(s_1',s_2')}\right) = o(n^2/d) \quad as \quad n\to\infty \tag{2.21}$$

provided $0 \le t_1 < t_2 \le 1$, $0 \le t_1' < t_2' \le 1$, $0 \le s_1 < s_2 < \infty$, $0 \le s_1' < s_2' < \infty$, 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

$$E\left(\sum_{k=nt_1+1}^{nt_2} X_{k,n}^{(s_1,s_2)}\right)^2 = n(t_2-t_1)E\left(X_{1,n}^{(s_1,s_2)}\right)^2 + R$$

where

$$R = 2\sum_{j=2}^{nt_2-nt_1}(nt_2-nt_1-j+1)E\left(X_{1,n}^{(s_1,s_2)}X_{j,n}^{(s_1,s_2)}\right).$$

Using Lemmas 3 and 5 we get, using $n/d\to\infty$,

$$E\left(X_{1,n}^{(s_1,s_2)}\right)^2 = E\left(X_1 I\{s_1(n/d) < X_1 \le s_2(n/d)\}\right)^2 - E^2\left(X_1 I\{s_1(n/d) < X_1 \le s_2(n/d)\}\right)$$
$$= c_0(1+o(1))(n/d)(s_2-s_1) + O(\log^2(n/d)) \sim c_0(n/d)(s_2-s_1)$$

and

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

proving (2.20).

To prove (2.21), consider a generic term

$$EX_{i,n}^{(s_1,s_2)}X_{j,n}^{(s_1',s_2')}$$

$$= EX_iX_jI\{s_1(n/d) < X_i \le s_2(n/d)\}I\{s_1'(n/d) < X_j \le s_2'(n/d)\} \quad (2.22)$$
$$- EX_iI\{s_1(n/d) < X_i \le s_2(n/d)\}EX_jI\{s_1'(n/d) < X_j \le s_2'(n/d)\}$$

of the left hand side of (2.21). Fix $r \ge 0$ and sum those covariances in (2.22) where $j-i=r$ and $nt_1+1 \le i \le nt_2$, $nt_1'+1 \le j \le nt_2'$. 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 (2.22) is 0. Thus by the first statement of Lemma 5 the product expectation in the first line of (2.22) is $O(\log^2(n/d))$ and since the number of such terms in the expansion of (2.21) is at most n, the contribution of such terms in the sum in (2.21) is at most $O(n\log^2(n/d)) = o(n^2/d)$ by $n/d \to \infty$. For $r \ge 1$ the covariance in (2.22) is at most $\psi(r)O(\log^2(n/d))$ by Lemma 3 and the first statement of Lemma 5 and since for fixed r the number of pairs (i,j) is at most n, the contribution of all such terms for all $r \ge 1$ is at most $Cn\log^2(n/d)\sum_{r=1}^{\infty}\psi(r) = O(n\log^2(n/d)) = o(n^2/d)$, proving (2.21).

The central limit theorem for ψ-mixing sequences follows after applying a simple blocking argument and theorem 1.7.3 from [53].

2.3 Proof of the main results

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 $M \ge 1, J \ge 1$, real coefficients $\mu_{m,j}$, $0 = s_0 < s_1 < s_2 < \ldots < s_J < \infty$, $0 = t_0 < t_1 < \ldots < t_M = 1$. Clearly, Z is a normal random variable with mean zero and

$$EZ^2 = \sum_{m=1}^{M} \sum_{j=1}^{J} \mu_{m,j}^2 (t_m - t_{m-1})(s_j - s_{j-1}). \tag{2.23}$$

We claim that

$$Q_n \xrightarrow{d} Z \qquad \text{for all considered values of } M, J, \mu_{m,j}, t_m, s_j. \tag{2.24}$$

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 (2.24) 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 21.

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

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

and thus relation (2.24) is equivalent to

$$\frac{1}{A_n} \sum_{k=1}^{n} (z_{nk} - E z_{nk}) \xrightarrow{d} N(0, EZ^2), \tag{2.25}$$

where

$$z_{nk} = \sum_{j=1}^{J} \mu_{m,j} X_k I\{s_{j-1}(n/d) < X_k \leq s_j(n/d)\}, \qquad [nt_{m-1}] + 1 \leq k \leq [nt_m]. \tag{2.26}$$

Since the terms of the sum in (2.26) are random variables with disjoint support, by the second relation of Lemma 5 we have

$$E z_{nk}^2 = (1 + o_n(1)) c_0(n/d) \sum_{j=1}^{J} \mu_{m,j}^2 (s_j - s_{j-1}), \qquad [nt_{m-1}] + 1 \leq k \leq [nt_m].$$

Also, by the first relation of Lemma 5 we have

$$E|z_{nk}| = O\left(\log(n/d)\right),$$

where the constant in the O depends also on the $\mu_{m,j}, s_j$. Consequently, letting

$$B_m = c_0 \sum_{j=1}^{J} \mu_{m,j}^2 (s_j - s_{j-1}), \tag{2.27}$$

we get

$$\operatorname{Var} z_{nk}$$
$$= (1+o_n(1))c_0(n/d)\sum_{j=1}^{J}\mu_{m,j}^2(s_j - s_{j-1}) = (1+o_n(1))(n/d)B_m, \qquad [nt_{m-1}]+1 \le k \le [nt_m]. \tag{2.28}$$

Further, Lemma 6 implies for $n \to \infty$

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

and

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

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

$$E\left(\sum_{k=1}^{n}(z_{nk} - Ez_{nk})\right)^2 = E\left(\sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}U_n([t_{m-1}, t_m] \times [s_{j-1}, s_j])\right)^2$$
$$= \sum_{m_1,m_2=1}^{M}\sum_{j_1,j_2=1}^{J}\mu_{m_1,j_1}\mu_{m_2,j_2}E\left[U_n([t_{m_1-1}, t_{m_1}] \times [s_{j_1-1}, s_{j_1}])U_n([t_{m_2-1}, t_{m_2}] \times [s_{j_2-1}, s_{j_2}])\right] \tag{2.29}$$
$$\sim c_0(n^2/d)\sum_{m=1}^{M}\sum_{j=1}^{J}\mu_{m,j}^2(t_m - t_{m-1})(s_j - s_{j-1}) = c_0(n^2/d)\,EZ^2 = A_n^2\,EZ^2.$$

Let

$$B = \max_{1\le m\le M} B_m, \qquad B' = \min_{1\le m\le M} B_m$$

where B_m is defined by (2.27). Clearly, without loss of generality we can assume that for each $1 \leq m \leq M$ at least one $\mu_{m,j}$ differs from 0, and thus B, B' are positive numbers depending on the numbers $c_0, s_j, \mu_{m,j}$. Now let

$$x_{nk} = \frac{z_{nk} - Ez_{nk}}{\sqrt{Bn/d}}. \tag{2.30}$$

Let us also note that by the first relation of Lemma 5, we have

$$\sum_{k=u+1}^{u+r} E|z_{nk} - Ez_{nk}| \leq C_4 r \log(n/d) \tag{2.31}$$

where C_4 is a positive constant depending on the $\mu_{m,j}, s_j$. Using (2.31), (2.29) and correlation inequality for $\psi-$mixing one can easily show that CLT also holds for the sequence (2.30) by applying to it the same tools as in the end of section 2.2.

Next we prove tightness in Theorem 21. Let

$$B_{11} = [t_1, t] \times [s_1, s], \qquad B_{12} = [t_1, t] \times [s, s_2], \qquad B_{21} = [t, t_2] \times [s_1, s].$$

where $0 \leq t_1 < t < t_2 \leq 1$, $1/2 \leq s_1 < s < s_2 \leq 3/2$. In view of Theorem 7, it suffices to show that

$$E \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{ij}) \right|^2 \leq C^* \mu(B_{11}) \mu(B_{ij}), \tag{2.32}$$

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 Theorem 7 we may assume that nt, nt_1 and nt_2 are all integers. To prove (2.32), we introduce the notations

$$X_i^{(1)} = X_i I\{s_1(n/d) < X_i \leq s(n/d)\}, \qquad m_i^{(1)} = EX_i^{(1)},$$
$$X_i^{(2)} = X_i I\{s(n/d) < X_i \leq s_2(n/d)\}, \qquad m_i^{(2)} = EX_i^{(2)}.$$

Using Lemmas 4 and 6 and (2.14) we get

$$E \left| \frac{1}{A_n} U_n(B_{11}) \right|^2 \left| \frac{1}{A_n} U_n(B_{21}) \right|^2$$

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

for $n \ge n_0$. On the other hand,

$$E\left|\frac{1}{A_n}U_n(B_{11})\right|^2\left|\frac{1}{A_n}U_n(B_{12})\right|^2$$
$$= \frac{1}{A_n^4}E\left(\sum_{i=nt_1+1}^{nt}(X_i^{(1)}-m_i^{(1)})\right)^2\left(\sum_{i=nt_1+1}^{nt}(X_i^{(2)}-m_i^{(2)})\right)^2$$
$$= \frac{1}{A_n^4}E\left(\sum_{i=nt_1+1}^{nt}Y_i^{(1)}\right)^2\left(\sum_{i=nt_1+1}^{nt}Y_i^{(2)}\right)^2, \tag{2.33}$$

where we put

$$Y_i^{(1)} = X_i^{(1)} - m_i^{(1)}, \qquad Y_i^{(2)} = X_i^{(2)} - m_i^{(2)}.$$

The expression in the third line of (2.33) equals the sum of all expressions

$$A_n^{-4}\,E(Y_i^{(1)}Y_j^{(1)}Y_k^{(2)}Y_\ell^{(2)}), \tag{2.34}$$

where $nt_1+1 \le i,j,k,\ell \le nt$. The following facts can be verified by elementary calculations using Lemmas 3–5:

(a) $E|Y_i^{(1)}| \ll s-s_1$, $E|Y_i^{(2)}| \ll s_2-s$, $E|Y_i^{(1)}Y_i^{(2)}| \ll (s-s_1)(s_2-s)$

(b) $E(Y_i^{(1)})^2 \ll (n/d)(s-s_1)$, $E(Y_i^{(2)})^2 \ll (n/d)(s_2-s)$,

(c) $E(Y_i^{(1)})^2|Y_i^{(2)}| \ll (n/d)(s-s_1)(s_2-s)$, $E|Y_i^{(1)}|(Y_i^{(2)})^2 \ll (n/d)(s-s_1)(s_2-s)$,

(d) $E(Y_i^{(1)})^2(Y_i^{(2)})^2 \ll (n/d)(s-s_1)(s_2-s)$,

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

$$E(Y_i^{(1)})^2(Y_i^{(2)})^2 = E\left[(X_i^{(1)} - m_i^{(1)})^2(X_i^{(2)} - m_i^{(2)})^2\right]$$
$$= E(X_i^{(1)})^2(X_i^{(2)})^2 - 2m_i^{(2)}E(X_i^{(1)})^2X_i^{(2)} + (m_i^{(2)})^2E(X_i^{(1)})^2 - 2m_i^{(1)}EX_i^{(1)}(X_i^{(2)})^2$$
$$+ 4m_i^{(1)}m_i^{(2)}EX_i^{(1)}X_i^{(2)} - 2m_i^{(1)}(m_i^{(2)})^2EX_i^{(1)} + (m_i^{(1)})^2E(X_i^{(2)})^2 - 2(m_i^{(1)})^2m_i^{(2)}EX_i^{(2)}$$
$$+ (m_i^{(1)})^2(m_i^{(2)})^2.$$

Clearly $X_i^{(1)}$ and $X_i^{(2)}$ are supported on different sets and thus $X_i^{(1)}X_i^{(2)} = 0$. Thus among the 9 terms above, the first, second, fourth and fifth are equal to 0. Also, the second and third statement of Lemma 5 imply, in view of $1/2 \le s_1 < s < s_2 \le 3/2$,

$$m_i^{(1)} = EX_1^{(1)} \ll s - s_1, \qquad m_i^{(2)} = EX_i^{(2)} \ll s_2 - s$$
$$E(X_i^{(1)})^2 \ll (s - s_1)(n/d), \qquad E(X_i^{(2)})^2 \ll (s_2 - s)(n/d)$$

for $n \ge n_0$. This shows that the remaining five terms of the 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 (2.34). We will distinguish four cases according as i, j, k, ℓ are all different, or the number of different ones among them is 1, 2 or 3. Consider first the case when i, j, k, l are all different, say $i < j < k < \ell$; let $r = j - i$. Applying Lemma 3 with $X = Y_i^{(1)}$, $Y = Y_j^{(1)}Y_k^{(2)}Y_\ell^{(2)}$ and using that $EX = 0$, we get that the absolute value of the expression (2.34) is bounded by

$$A_n^{-4}\psi(r)E|X|E|Y| \le CA_n^{-4}\psi(r)E|Y_i^{(1)}|E|Y_j^{(1)}|E|Y_k^{(2)}|E|Y_\ell^{(2)}| \le CA_n^{-4}\psi(r)(s-s_1)^2(s_2-s)^2,$$

where we used Lemma 4 to estimate $E|Y|$ 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 (2.34) 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 (2.34) 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,ℓ) 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,ℓ) can be chosen by at most $(nt-nt_1)^2$ different ways, it follows that the contribution of all terms (2.34) with $i<j<k<\ell$ is at most

$$\begin{aligned} &CA_n^{-4}(nt-nt_1)^2(s-s_1)^2(s_2-s)^2 \le C(d^2/n^2)(t-t_1)^2(s-s_1)^2(s_2-s)^2 \\ &\le C(t-t_1)^2(s-s_1)(s_2-s) = C\mu(B_{11})\mu(B_{12}), \end{aligned}$$

using (2.14) and $d/n \to 0$. The contribution of terms (2.34) where i,j,k,ℓ 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 (2.34) becomes $A_n^{-4}E(Y_i^{(1)})^2E(Y_i^{(2)})^2$, 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 $nt-nt_1 \le (nt-nt_1)^2$, the contribution of all such expressions is bounded by

$$CA_n^{-4}(n/d)(s-s_1)(s_2-s)(nt-nt_1)^2 \le C(d/n)(s-s_1)(s_2-s)(t-t_1)^2 \le C\mu(B_{11})\mu(B_{12}),$$

using again (2.14) and $d/n \to 0$.

Assume now that among i,j,k,ℓ 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 $i \ne k$. In this case the expression (2.34) becomes $A_n^{-4}E(Y_i^{(1)})^2(Y_k^{(2)})^2$ which, in view of Lemma 4 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 (2.14) it follows that the total contribution of all such terms (2.34) is at most

$$CA_n^{-4}(n/d)^2(s-s_1)(s_2-s)(nt-nt_1)^2 \le C(s-s_1)(s_2-s)(t-t_1)^2 = C\mu(B_{11})\mu(B_{12}).$$

If $i=k, j=l$ and $i\neq j$, then the expression (2.34) becomes $A_n^{-4}EY_i^{(1)}Y_i^{(2)}Y_j^{(1)}Y_j^{(2)}$ which by Lemma 4 and the estimate in (a) above is bounded by

$$CA_n^{-4}E|Y_i^{(1)}Y_i^{(2)}|E|Y_j^{(1)}Y_j^{(2)}| \le CA_n^{-4}(s-s_1)^2(s_2-s)^2.$$

Since the number of pairs (i,j) is $\le (nt-nt_1)^2$, the contribution of such terms is at most

$$CA_n^{-4}(s-s_1)^2(s_2-s)^2(nt-nt_1)^2 \le C(s-s_1)(s_2-s)(t-t_1)^2 = C\mu(B_{11})\mu(B_{12}).$$

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 (2.34) becomes $A_n^{-4}E(Y_i^{(1)})^2Y_i^{(2)}Y_\ell^{(2)}$ which is, by Lemma 4 and the estimates (a) and (c) above is bounded by

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

Since the number of pairs (i,ℓ) is $\le (nt-nt_1)^2$, 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 (2.34) becomes $A_n^{-4}E(Y_i^{(1)})^2Y_k^{(2)}Y_\ell^{(2)}$ which by using $EY_\ell^{(2)}=0$, Lemma 3, Lemma 4 and estimates (a) and (b) above, can be estimated by

$$CA_n^{-4}\psi(r)E(Y_i^{(1)})^2E|Y_k^{(2)}|E|Y_\ell^{(2)}| \le CA_n^{-4}\psi(r)(n/d)(s-s_1)(s_2-s)^2, \qquad (2.35)$$

where $r=\ell-k$. Since for fixed r the number of triples (i,k,ℓ) with $\ell-k=r$ is at most $(nt-nt_1)^2$, the contribution of such terms (2.34) is at most

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

and summing for r we get again $\leq C\mu(B_{11})\mu(B_{12})$. The other cases (e.g. $i < j = k < \ell$, etc.) can be treated similarly and the proof of tightness in Theorem 21 is completed. This also completes the proof of the theorem.

We prove now, as claimed after Theorem 21, that

$$\frac{\eta_{d,n}}{n/d} \xrightarrow{P} 1 \tag{2.36}$$

for $d = d_n \to \infty$, $d_n/n \to 0$. Fix $n \geq 1$, $1/2 < t < 2$ and let $T_k = I\{X_k \geq tn/d\}, 1 \leq k \leq n$. Then by Lemma 3 and (2.13) we get

$$|ET_1T_k - ET_1ET_k| \leq \psi(k-1)ET_1ET_k \leq C_9 \exp(-\lambda k)(d/n)^2$$

and thus setting $\bar{T}_k = T_k - ET_k$ we conclude that

$$\begin{aligned}
E\left(\sum_{k=1}^{n} \bar{T}_k\right)^2 &= nE\bar{T}_1^2 + 2\sum_{k=2}^{n}(n-k+1)E\bar{T}_1\bar{T}_k \\
&\leq n\left(E\bar{T}_1^2 + 2\sum_{k=2}^{n}|E\bar{T}_1\bar{T}_k|\right) \\
&\leq n\left(ET_1^2 + C_{10}(d/n)^2\sum_{k=2}^{n}\exp(-\lambda k)\right) \\
&\leq n\left(ET_1 + C_{11}(d/n)^2\right) \\
&\leq C_{12}d.
\end{aligned}$$

Hence Markov's inequality and $d = d_n \to \infty$ imply for any $\varepsilon > 0$

$$P\left\{\sum_{k=1}^{n}\bar{T}_k \geq \varepsilon d\right\} \longrightarrow 0,$$

and since $ET_k = ET_1 \sim d/(nt)$ by (2.13), 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 (2.36) is proved.

Proof of Remark 1.1. Let (X_n) be a sequence satisfying the assumptions of Theorem 20, and put $X'_n = X_n + 4^{-n}$. Letting $\eta'_{d,n}$ denote the d-th largest of $X'_1, \ldots, X'_n$ and $S_n^{(r)}$ and $S_n'^{(r)}$ 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

$$|S_n^{(r)} - S_n'^{(r)}| \leq 2 \qquad \text{for any } r \geq 1 \tag{2.37}$$

and

$$n|m(\eta_{d,n}) - m(\eta'_{d,n})| = O_P(1). \tag{2.38}$$

Clearly, relation (2.13) will fail for the perturbed sequence (X'_n), but as inspection shows, all the lemmas in the proof of Theorem 20 and the subsequent arguments remain valid, so conclusion (2.15) 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 $S_n'^{(d-1)}$. Thus we have

$$\frac{S_n'^{(d-1)} - nm(\eta'_{d,n})}{n/\sqrt{d}} \xrightarrow{d} N(0, c_0). \tag{2.39}$$

In view of (2.37) and (2.38), we can drop the primes in (2.39) and since $S_n^{(d-1)} - S_n^{(d)} = \eta_{d,n} = O_P(n/d)$ by (2.36), the conclusion of Theorem 10 follows.□

Chapter 3

Trimmed stable AR(1) processes

3.1 Introduction

In this chapter we concentrate on the applications of trimming in statistics. We provide theoretical base for AR(1) processes and then illustrate it with simulations. As usual we start with a brief historical overview.

Let $X_1, X_2, \dots,$ be independent, identically distributed random variables in the domain of attraction of a stable law with index $0 < \alpha < 2$. Lévy ([67], 1935) and Darling (see theorems 11, 12) noted that the order of magnitude of the sum $S_n = \sum_{k=1}^{n} X_k$ is the same as that of its largest term and the contribution of a fixed, but large number of extremal terms is essentially responsible for the distribution of S_n. The asymptotic distribution of the trimmed sum $S_n^{(d)}$ obtained from S_n by discarding the d smallest and d largest summands was determined by LePage, Woodrofe and Zinn ([65], 1981) as we have already mentione in section 1.1 and Csörgő, Horváth and Mason (see theorem 3) proved that for $d(n) \to \infty$, $d(n)/n \to 0$ the trimmed sum $S_n^{(d)}$ satisfies the central limit theorem. Arov and Bobrov ([3], 1960), Mori ([74], 1984), Hall ([46], 1978), Teugels ([94], 1981), Griffin and Pruitt ([44], 1987, [45] 1989) and Kesten ([55], 1993) considered a different type of trimming of the sample. Let now $\eta_{n,d}$ denote the d–th largest element of

$|X_1|, \ldots, |X_n|$. These authors were interested in the asymptotic behavior of the modulus trimmed sum ${}^{(d)}S_n = \sum_{k=1}^{n} X_k I\{|X_k| \leq \eta_{n,d}\}$, i.e. when from the sum we remove the d elements with the largest absolute values.

In case of light trimming, that is when the amount of removed alements d_n is constant an important result for both modulus and ordinary trimming was obtained by Kesten (see [55], 1993). This result shows that light trimming simply does not lead to change of limit behavior.

Theorem 23 (Kesten,1993)**.** *Let $X_1, X_2, \ldots$ be i.i.d. random variables. If for some fixed d and sequences of constants a_n, b_n with $b_n \to \infty$ one has*

$$\frac{{}^{(d)}S_n - a_n}{b_n} \text{ converges in distribution as } n \to \infty \tag{3.1}$$

or

$$\frac{S_n^{(d)} - a_n}{b_n} \text{ converges in distribution as } n \to \infty, \tag{3.2}$$

then also

$$\frac{S_n - a_n}{b_n} \text{ converges in distribution as } n \to \infty. \tag{3.3}$$

However the case of moderate trimming ($d_n \to \infty$, $d_n/n \to 0$) appears to be more complicated.

Arov and Bobrov ([3], 1960) stated the following asymptotics for partial sums S_n given that condtions (1.2),(1.3) are satisfied.

Theorem 24 (Arov and Bobrov, 1960)**.** *Let $X_1, \ldots, X_n$ be independent random variables with the same distribution function, satisfying (1.2),(1.3). Then using the above notation for $\alpha = 0$*

$$S_n = \eta_{n,1} + \cdots + \eta_{n,d} + o(1)\eta_{n,d},$$

for $0 < \alpha < 1$

$$S_n = \eta_{n,1} + \cdots + \eta_{n,d_n} + \left(\frac{\alpha}{1-\alpha} + o(1)\right) d_n \eta_{n,d_n}, \quad d_n = o(n/\log n),\ d_n \to \infty,$$

for $1 < \alpha < 2$

$$S_n = na + o(1)\eta_{n,d_n}, \qquad a = \int_{-\infty}^{\infty} x dF(x),\ d_n = o(n/\log n),\ d_n \to \infty.$$

Hall ([46], 1978) found the limiting characteristic function of a suitebly normalized trimmed sum ${}^{(d)}S_n$ and stated the folloewing.

Theorem 25 (Hall, 1978)**.** *Define the constants* b_n *by*

$$b_n = \begin{cases} 0 & if\ 0 < \alpha < 1, \\ na_n E\sin(X_1/a_n) & if\ \alpha = 1, \\ nEX_1 & if\ 1 < \alpha < 2. \end{cases} \tag{3.4}$$

where a_n *is defined by*

$$a_n = \inf\{x : F(-x) + 1 - F(x) \le 1/n\},$$

and F satisfies (1.2),(1.3). Then as $n \to \infty$ *a random variable* $({}^{(d)}S_n - a_n)/b_n$ *with characteristic function given by*

$$\eta_k(t) = [(k-1)!]^{-1}\zeta_\alpha(t)\int_0^\infty u^{k-1}\exp\{-\mu_\alpha(t,u)\}du,$$

where $\zeta_\alpha \equiv 1$ *if* $\alpha < 1$, $\zeta_\alpha \equiv \psi_\alpha$ *if* $1 \le \alpha < 2$,

$$\mu_\alpha(t,u) = u[p\exp(itu^{-1/\alpha}) + q\exp\{itu^{-1/\alpha}\}] - it\int_0^{u^{-1/\alpha}} z^{-\alpha}(pe^{itz} - qe^{-itz})$$

for $0 < \alpha < 1$ *and*

$$\mu_\alpha(t,u) = u[p\exp(itu^{-1/\alpha}) + q\exp\{itu^{-1/\alpha}\}] + it\int_0^{u^{-1/\alpha}} z^{-\alpha}(pe^{itz} - qe^{-itz})$$

for $1 \le \alpha < 2$

ψ_α *is a characteristic function of a stable distribution and therefore*

$$\log\psi_\alpha(t) = -|t|^\alpha\left[\int_0^\infty x^{-\alpha}\sin x dx\right](1 - sgn(t)i(p-q)\tan(\alpha\pi/2)) \tag{3.5}$$

for $\alpha \neq 1$ *and*

$$\log \psi_1(t) = -|t|(\frac{1}{2}\pi + sgn(t)i(p-q)\log|t|). \tag{3.6}$$

Teugels (see [94], 1981) was studying the distribution of ratio ${}^{(d)}S_n/|\eta_{n,d}|$. Under conditions (1.2), (1.3) he provided the limiting characteristic function of ${}^{(d)}S_n/|\eta_{n,d}|$ and as a corollary established the limit result

Theorem 26 (Teugels, 1981)**.** *There exist* d_n *and* d'_n *tending to infinity such that*

$$\frac{{}^{(d)}S_n}{d_n|\eta_{n,d}|} \xrightarrow{P} \gamma \tag{3.7}$$

$$\frac{{}^{(d)}S_n/|\eta_{n,d}| - \gamma d'_n}{(\beta d'_n)^{1/2}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \tag{3.8}$$

where

$$\gamma = \frac{\alpha}{1-\alpha}(p-q) \qquad \beta = \frac{\alpha}{2-\alpha} + \left(\frac{\alpha(p-q)}{\alpha-1}\right)$$

In the same work Teugels provided the constraints on d_n and d'_n such that (3.7) and (3.8) hold.

As we have already mentined in chapter 1 Griffin and Pruitt ([44], 1987) proved that the trimmed central limit theorem of Csörgő, Horváth and Mason (theorem 3) remains valid for modulus trimmed sums provided the distribution of X_1 is symmetric, but it generally fails for nonsymmetric variables and it can happen that ${}^{(d)}S_n$ is asymptotically normal for some $d(n)$, but not for another $d'(n) \geq d(n)$. Sufficient conditions for the asymptotic normality of ${}^{(d)}S_n$ in the nonsymmetric case were given by Berkes and Horváth ([14], 2012, see e.g. theorem 5). On the other hand, Berkes, Horvath and Schauer ([16], 2011) showed that if $d(n) \to \infty$, $d(n)/n \to 0$, a functional central limit theorem always holds for ${}^{(d)}S_n$ with a random centering factor.

Trimming also has important applications in statistics. As an example, we consider the detection of possible changes in the location model

$$X_j = c_j + e_j, \quad 1 \leq j \leq n, \tag{3.9}$$

where $e_1, \ldots, e_n$ are random errors. Under the null hypothesis of stability, the location parameter is constant, i.e.

$$H_0: \quad c_1 = c_2 = \ldots = c_n.$$

If H_0 holds, then

$$X_j = c + e_j, \quad 1 \le j \le n, \tag{3.10}$$

with some constant c. Under the alternative there are r changes:

$$H_A: \text{ there is } r \ge 1 \text{ and } 1 < k_1 < k_2 < \ldots < k_r < n \text{ such that}$$
$$c_1 = \ldots = c_{k_1-1} \neq c_{k_1} = c_{k_1+1} = \ldots = c_{k_2-1} \neq c_{k_2} = c_{k_2+1} = \ldots$$
$$= c_{k_r-1} \neq c_{k_r} = \ldots = c_n.$$

The most popular methods to test H_0 against H_A (cf. Csörgő and Horváth [25], 1998 and Aue and Horváth [5], 2012) are based on the CUSUM process

$$U_n(x) = \sum_{i=1}^{\lfloor nx \rfloor} X_i - \frac{\lfloor nx \rfloor}{n} \sum_{i=1}^{n} X_i, \tag{3.11}$$

where $\lfloor \cdot \rfloor$ denotes the integer part. Clearly, if H_0 is true, then $U_n(t)$ does not depend on the common but unknown location parameter c_1. It is well known if $X_1, \ldots, X_n$ are independent and identically distributed random variables with a finite second moment, then

$$\frac{1}{(n\text{var}(X_1))^{1/2}} U_n(x) \xrightarrow{\mathcal{D}[0,1]} B(x),$$

where $B(x)$ is a Brownian bridge and $\xrightarrow{\mathcal{D}[0,1]}$ means weak convergence in the space $\mathcal{D}[0,1]$ of cadlag functions equipped with the Skorokhod J_1 topology (cf. Billingsley [19], 1968) which is defined as follows.

Definition 3. *Let Λ denote the class of strictly increasing, continuous mappings of $[0,1]$ onto itself. If $\lambda \in \Lambda$, then $\lambda(0) = 0$ and $\lambda(1) = 1$. For x and y in the* *space of cadlag functions D, define $d(x,y)$ to be the infimum of those positive ε for which there exists in Λ a λ such that*

$$\sup_t |\lambda t - t| \le \varepsilon$$

and

$$\sup_t |x(t) - y(\lambda t)| \le \varepsilon.$$

Metric d defines Skorokhod J_1 topology.

Assuming that $X_1, X_2, \ldots, X_n$ are independent and identically distributed random variables in the domain of attraction of a stable law of index $\alpha \in (0,2)$, Aue, Berkes and Horváth ([4], 2008) showed that

$$\frac{1}{n^{1/\alpha}\hat{L}(n)} U_n(x) \stackrel{\mathcal{D}[0,1]}{\longrightarrow} B_\alpha(x),$$

where $\hat{L}$ is a slowly varying function at ∞ and $B_\alpha(x)$ is an α–stable bridge. (The α–stable bridge is defined as $B_\alpha(x) = W_\alpha(x) - xW_\alpha(1)$, where W_α is a Lévy α–stable motion.) Since nothing is known on the distributions of the functionals of α–stable bridges, Berkes, Horváth and Schauer ([16], 2011) suggested the trimmed CUSUM process

$$T_{n,d}(x) = \sum_{i=1}^{\lfloor nx \rfloor} X_i I\{|X_i| \le \eta_{n,d}\} - \frac{\lfloor nx \rfloor}{n} \sum_{i=1}^{n} X_i I\{|X_i| \le \eta_{n,d}\}. \tag{3.12}$$

Assuming that the X_i's are independent and identically distributed and are in the domain of attraction of a stable law, they proved

$$\frac{1}{\sigma_n} T_{n,d}(x) \stackrel{\mathcal{D}[0,1]}{\longrightarrow} B(x), \tag{3.13}$$

where

$$\sigma_n^2 = \frac{\alpha}{2-\alpha}(H^{-1}(d/n))^2 d,$$

$B(t)$ is a Brownian bridge and H^{-1} denotes the generalized inverse of H, the survival function of X_1. The CUSUM process has also been widely used in case of dependent variables but it is nearly always assumed that the observations have high moments and the dependence in the sequence is weak. For a review we refer to Aue and Horváth ([5], 2012). However, very few papers consider the instability of time series models with heavy tails.

Fama ([37], 1965) and Mandelbrot ([72], 1963, [73], 1967) pointed out that the distributions of commodity and stock returns are often heavy tailed with possible infinite variance. We can easily the similarities between the simulated values for AR(1) process (see Figure 1) and the actuall values obtained during the study on the stock return on the swedish index (see Figure 3.1).

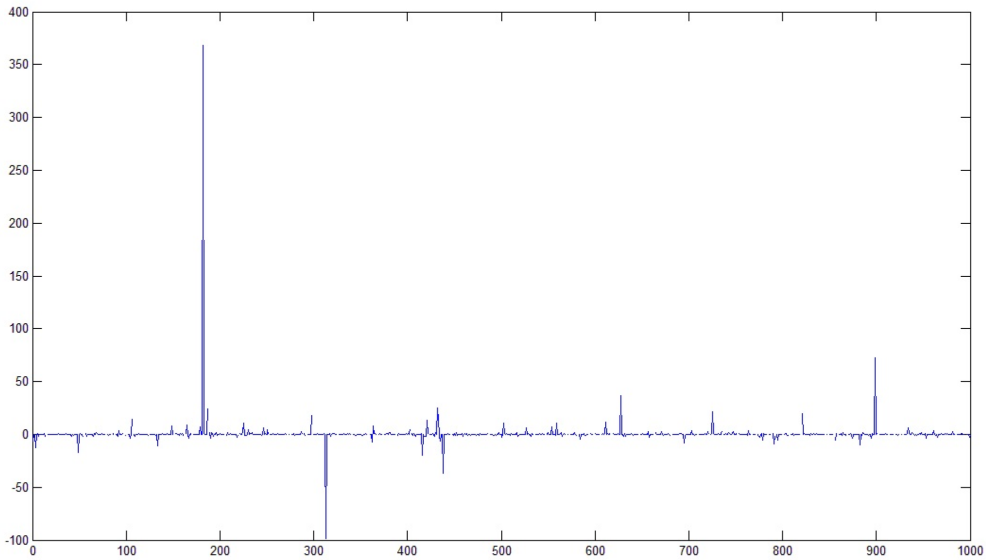

Figure 3.1: Empirical returns obtained during the study on the swedish stock market

Fama's and Mandelbrot's research started the investigation of time series models where the marginal distributions have regularly varying tails. Davis and Resnick ([31], 1985, [32], 1986) investigated the properties of moving averages with regularly varying tails and obtained non–Gaussian limits for the sample covariances and correlations. Here we are giving some important results

Theorem 27 (Davis and Resnick, 1985). *Let* $\{Y_k, -\infty < k < \infty\}$ *be a sequence of i.i.d. variables, such that (1.2), (1.3) are satisfied with* $0 < \alpha < 2$. *Define* $X_n = \sum_{j=0}^{\infty} c_j Y_{n-j}$. *Let* $\{a_n\}$ *be a sequence of numbers such that*

$$nP(|X_1| > a_n x) \to x^{-\alpha} \quad for\ all\ x > 0,$$

and let

$$\sum_{j=0}^{\infty} |c_j|^{\delta} < \infty \quad for\ some\ \delta < \alpha,\ \delta \leq 1.$$

Then

$$\frac{S_n - n\sum_{j=0}^{\infty} c_j b_n}{a_n} \Rightarrow S \quad in\ \mathbb{R}$$

where S *has a stable distribution with index* α.

Theorem 28 (Davis and Resnick, 1986). *Let* $\{Y_k, -\infty < k < \infty\}$ *be a sequence of i.i.d. variables, such that (1.2), (1.3) are satisfied with* $2 \leq \alpha < 4$. *Define* $X_n = \sum_{j=-\infty}^{\infty} c_j Y_{n-j}$ *so that* $\sum_{-\infty}^{\infty} |c_j| < \infty$.

$$a_n = \inf\{x : P(|Y_1| > x) \leq n^{-1}\},$$

$\hat{\gamma}(h)$ *is a sample covariance function of* $(X_1, ..., X_n)$, *and* $\gamma(h) = cov(X_n, X_{n+h})$. *If* $EY_k = 0$, *then for any positive integer* l

$$\left(\frac{n(\hat{\gamma}(h) - b_{h,n})}{a_n^2}, 0 \leq h \leq l\right) \Rightarrow S\left(\sum_j c_j^2, \sum_j c_j c_{j+1}, \ldots, \sum_j c_j c_{j+1}\right),$$

where S *is a stable random variable with index* $\alpha/2$ *and* $b_{h,n} = \sum_{i=-\infty}^{\infty} c_i c_{i+h} EY_1^2 I_{\{|Y_1 \leq a_n|\}}$, $0 \leq h \leq l$. *Moreover, if* $2 < \alpha < 4$,*then*

$$\left(\frac{n(\hat{\gamma}(h) - \gamma(h)}{a_n^2}, 0 \leq h \leq l\right) \Rightarrow \left(S - \frac{\alpha}{2-\alpha}\right)\frac{(\gamma(0), \ldots, \gamma(l))}{\sigma^2},$$

where $\sigma^2 = varX_n$.

These results were extended to heavy tailed ARCH by Davis and Mikosch ([30], 1998). The empirical periodogram of the form

$$I_{n,X} = |J_{n,X}|^2 = |b_n^{-1}\sum_{t=1}^{n} X_t \exp(-i2\pi xt)|^2 =$$
$$\left(b_n^{-1}\sum_{t=1}^{n} X_t \cos(2\pi xt)\right)^2 + \left(b_n^{-1}\sum_{t=1}^{n} X_t \sin(2\pi xt)\right)^2, \quad x \in [0, 0.5], \quad (3.14)$$

where $\{X_t\}$ is a heavy-tailed random sequence and b_n is a sequence of appropriate norming factors was studied by Mikosch, Resnick and Samorodnitsky ([74], 2000). They shown that in case $\alpha < 1$ maxima is a weakly converges to a certain distribution.

Theorem 29 (Mikosch, Resnick and Samorodnitsky, 2000)**.** *Let $\{X_t\}$ be a sequence of i.i.d. heavy-tailed random variables stisfying (1.2),(1.3) for $\alpha \in (0, 1)$. Then*

$$\max_{x\in[0,0.5]} I_{n,X}(x) \Rightarrow Y_\alpha^2$$

holds, where Y_α has a stable distribution with parmeter α.

For the case of $1 \le \alpha < 2$ they shown that the sequence of maxima is not tight and determined that

$$\frac{\max I_{n,X}}{\ln n}^{(2/\alpha)(\alpha-1)}, \qquad n \ge 2, \quad \alpha \in (1, 2)$$

and

$$\frac{\max I_{n,X}}{(\ln\ln n)^2}, \qquad n > e, \alpha = 1$$

are tight.

Andrews, Calder and Davis ([1], 2009) studied autoregressive processes with stable innovations. They gave a nondegenerate limiting distribution for maximum-likelihood parameters of the autoregressive model equation and the peremeters of the stable noise distribution.

In this chapter we study trimmed sums of AR(1) sequences with heavy tails. Let e_i be a $\sigma(\varepsilon_j, j \le i)$ measurable solution of

$$e_i = \rho e_{i-1} + \varepsilon_i \quad -\infty < i < \infty. \tag{3.15}$$

We assume throughout this chapter that

$$\varepsilon_j, -\infty < j < \infty \text{ are independent and identically distributed,} \tag{3.16}$$

$$\begin{gathered}\varepsilon_0 \text{ belongs to the domain of attraction of a stable} \\ \text{random variable } \xi^{(\alpha)} \text{ with parameter } 0 < \alpha < 2,\end{gathered} \tag{3.17}$$

and

$$\varepsilon_0 \text{ is symmetric when } \alpha = 1. \tag{3.18}$$

Assumption (3.17) means that

$$\left(\sum_{j=1}^{n} \varepsilon_j - a_n\right) \Big/ b_n \xrightarrow{\mathcal{D}} \xi^{(\alpha)} \tag{3.19}$$

for some numerical sequences a_n and b_n. The necessary and sufficient condition for this is

$$\lim_{t\to\infty} \frac{P\{\varepsilon_0 > t\}}{L_*(t)t^{-\alpha}} = p \quad \text{and} \quad \lim_{t\to\infty} \frac{P\{\varepsilon_0 \le -t\}}{L_*(t)t^{-\alpha}} = q \tag{3.20}$$

for some numbers $p \ge 0$, $q \ge 0$, $p + q = 1$, where L_* is a slowly varying function at ∞. It is known that (3.15) has a unique stationary non–anticipative solution if and only if

$$-1 < \rho < 1. \tag{3.21}$$

Under assumptions (3.16)–(3.21), $\{e_j\}$ is a stationary sequence and $E|e_0|^\kappa < \infty$ for all $0 < \kappa < \alpha$ but $E|e_0|^\kappa = \infty$ for all $\kappa > \alpha$.

AR(1) processes with stable innovations were considered by Chan and Tran ([23], 1989). They established strong consistency of the ordinary least squares estimator b_n of ρ in case $\rho = 1$. Chan and Tran also shown that the limiting distribution of b_n has a form of a functional of Lévy process.

Chan ([22], 1990) established the asymptotic distribution of least squares estimator for the case when ρ is close to 1. This case was also covered by Aue and Horváth ([5], 2007) in their work on structural breaks in time sreies models. Zhang and Chan ([98], 2010) who investigated the model $X_t = \rho_n X_{t-1} + \varepsilon_{t-1}$ in our notation with $\rho_n = 1 - y/n$, where $y > 0$ is a constant. They shown that the distribution of maximum likelihood estimator of ρ_n and θ are mixtures of a stable procces and Gaussian processes. Here θ is the parameter of the characteristic function of ε_t.

The convergence of the finite dimensional distributions of $U_n(x)$ in the AR(1) case is an immediate consequence of Phillips and Solo ([82], 1992) stable limit representation. Let $\overset{fdd}{\longrightarrow}$ denote the convergence of the finite dimensional distributions. If (3.10)–(3.18) and (3.21) hold, then we have

$$\frac{1-\rho}{n^{1/\alpha} L_*(n)} U_n(x) \overset{\text{fdd}}{\longrightarrow} B_\alpha(x), \tag{3.22}$$

where $B_\alpha(x), 0 \le x \le 1$ is an α–stable bridge and L_* is defined in (3.20). It has been pointed out by Avram and Taqqu ([7], 1986, [8], 1992) that the fdd convergence in (3.22) cannot be replaced with weak convergence in $\mathcal{D}[0,1]$. However, Avram and Taqqu ([8], 1992) proved that $U_n(x)$ converges in the weak–M_1 sense. The definition of M_1-topology is given below (see also Basrak, Krizmanič and Segers, [9], 2012).

Definition 4. *For $x \in D[0,1]$ the completed graph of x is the set*

$$\Gamma_x = \{(t,z) \in [0,1] \times \mathbb{R} : z = \lambda x(t-) + (1-\lambda) x(t) \ for \ some \ \lambda \in [0,1]\}$$

We define order on the graph Γ_x by saying that $(t_1, z_1) \le (t_2, z_2)$ if either $t_1 \le t_2$ or $t_1 = t_2$ and $|x(t_1-) - z_1| \le |x(t_2-) - z_2|$. A parametric representation of the completed graph Γ_x is a continuous nondecreasing function (r,u) mapping $[0,1]$

onto Γ_x *with* r *being the time component and* u *being the spatial component. Let* $\Pi(x)$ *denote the set of parametric representations of the graph* Γ_x*. For* x_1, x_2 *define*

$$d_{M_1}(x_1, x_2) = \inf\{||r_1 - r_2||_{[0,1]} \vee ||u_1 - u_2||_{[0,1]} : (r_i, u_i) \in \Pi(x_i), i = 1, 2\},$$

where $||x(t)||_{[0,1]} = \sup\{x(t) : t \in [0,1]\}$*. Metric* d_{M_1} *induces* M_1*-topology.*

Avram and Taqqu's result holds under the following technical conditions (which are necessary for the case $\alpha > 1$).

Let $\alpha > 1$, $\mathbf{c} = \{c_i, -\infty < i < \infty\}$, $\sum_{i=-\infty}^{\infty} |c_i|^\delta$, where $0 < \delta < \alpha$ and for some $0 < \eta < \alpha - 1$

$$\lim_{n\to\infty} (\ln n)^{1+\alpha+\eta} s(\alpha - \eta, \mathbf{c}^{>n}) = 0.$$

Here $s(\alpha - \eta, \mathbf{c}) = \left(\sum_{i=-\infty}^{\infty} |c_i|^\delta\right) \left(\sum_{i=-\infty}^{\infty} |c_i|\right)^{\alpha-\eta-\delta}$, where $\alpha - \eta > 1 > \delta$ and $\mathbf{c}^{>n} = \{c_i^{>n}, -\infty < i < \infty\}$, where

$$c_i^{>n} = \begin{cases} c_i & \text{if } |i| > n, \\ 0 & \text{otherwise.} \end{cases}$$

Some of these conditions were removed by Tyran–Kamińska ([95], 2010). For further results on this subject we refer to Basrak, Krizmanič and Segers ([9], 2012) who provided a new limit theorem for point processes which they use to derive a functional limit theorem for weak convergence of dependent sequence with infinite variance in M_1 topology.

We formulate now our main results. On the truncation parameter $d = d(n)$ we will assume

$$\lim_{n\to\infty} d(n)/n = 0 \tag{3.23}$$

and

$$d(n) \geq n^\delta \quad \text{with some} \quad 0 < \delta < 1. \tag{3.24}$$

Let $F(x) = P\{X_0 \leq x\}$, $H(x) = P\{|X_0| > x\}$ and let $H^{-1}(t)$ be the (generalized) inverse of H. Our last condition will be used to establish the weak law of large numbers for $\eta_{n,d}$. We assume that ε_0 has a density function $p(t)$ which satisfies

$$\int_{-\infty}^{\infty} |p(t+s) - p(t)|dt \leq C|s| \quad \text{with some} \quad C. \tag{3.25}$$

Let

$$A_n = d^{1/2} H^{-1}(d/n) \tag{3.26}$$

and

$$m(t) = EX_1 I\{|X_1| \leq t\}. \tag{3.27}$$

Theorem 30. *If (3.10)–(3.18) and (3.21)–(3.25) hold, then we have that*

$$\left(\frac{2-\alpha}{\alpha}\right)^{1/2} \left(\frac{1-\rho}{1+\rho}\right)^{1/2} \frac{1}{A_n} \sum_{k=1}^{n} [X_k I\{|X_k| \leq \eta_{n,d}\} - m(\eta_{n,d})] \stackrel{\mathcal{D}[0,1]}{\longrightarrow} W(x),$$

where $W(x)$ is a Wiener process.

The result in Theorem 30 uses the the random centering factor $m(\eta_{n,d})$. This factor is characteristic for the asymptotic distribution of the modulus trimmed partial sums process, as first observed in Berkes, Horváth and Schauer ([14], 2011). Since a random translation of the terms in the CUSUM process cancels out, the next result is an immediate consequence of Theorem 30.

Theorem 31. *If (3.10)–(3.18) and (3.21)–(3.25) hold, then we have that*

$$\left(\frac{2-\alpha}{\alpha}\right)^{1/2} \left(\frac{1-\rho}{1+\rho}\right)^{1/2} \frac{T_{n,d}(x)}{A_n} \stackrel{\mathcal{D}[0,1]}{\longrightarrow} B(x),$$

where $B(x)$ is a Brownian bridge.

Statistical applications of Theorem 31 require the estimation of the norming factor from the observations. The estimation of this term will be studied in sections 3.5, 3.6.

3.2 Preliminary results

The proofs of Theorems 30 and 31 are based on several technical lemmas.

We can and will assume without loss of generality that

$$E\varepsilon_0 = 0, \text{ if } 1 < \alpha < 2. \tag{3.28}$$

Under these conditions, in (3.19) we can choose $a_n = 0$ and b_n can be chosen any sequence satisfying

$$\frac{n}{b_n^\alpha} L_*(b_n) \to 1. \tag{3.29}$$

According to the result of Cline ([24], 1983) (cf. also Davis and Resnick ([32], 1986)), $H(x)$, the survival function of $|X_0|$ satisfies

$$H(x) = x^{-\alpha} L(x), \tag{3.30}$$

where $L(x)$ is a slowly varying function at ∞ and

$$\lim_{x\to\infty} \frac{H(x)}{P\{|\varepsilon_0| > x\}} = \lim_{x\to\infty} \frac{L(x)}{L_*(x)} = \frac{1}{1-|\rho|^\alpha}. \tag{3.31}$$

Let

$$u_{k,n}(t) = X_k I\{|X_k| \le tH^{-1}(d/n)\} \quad \text{and} \quad m_n(t) = E[X_0 I\{|X_0| \le tH^{-1}(d/n)\}].$$

The main goal of this section is to get bounds for $Eu_0(t)u_k(s)$ and $\operatorname{cov}(u_0(t), u_k(s))$.

Lemma 7. *We assume that (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold. Let* $\mathbf{Y}^{(k)} = (X_0, X_k)$ *and let* $\mathbf{Y}_i^{(k)}, i = 1, 2, \ldots$ *be independent and identically distributed copies of* $\mathbf{Y}^{(k)}$. *Then*

$$\frac{\mathbf{Y}_1^{(k)} + \ldots + \mathbf{Y}_n^{(k)}}{n^{1/\alpha} L_*(n)} \xrightarrow{\mathcal{D}} \mathbf{Z}^{(k)} \quad \text{as} \quad n \to \infty,$$

where $\mathbf{Z}^{(k)} = (Z_1^{(k)}, Z_2^{(k)})$ *with*

$$Z_1^{(k)} = \sum_{\ell=0}^{\infty} \rho^{\ell} \xi_{-\ell}^{(\alpha)} \quad \textit{and} \quad Z_2^{(k)} = \sum_{\ell=0}^{\infty} \rho^{\ell} \xi_{k-\ell}^{(\alpha)}$$

and $\xi_{\ell}^{(\alpha)}, -\infty < \ell < \infty$ *are independent and identically distributed copies of* $\xi^{(\alpha)}$.

Proof. It follows from (3.15) that

$$X_k - c = \sum_{\ell=0}^{\infty} \rho^{\ell} \varepsilon_{k-\ell} = \sum_{\ell=0}^{k-1} \rho^{\ell} \varepsilon_{k-\ell} + \rho^k X_0, \quad 1 \leq k < \infty. \tag{3.32}$$

Let $\varepsilon_{\ell}^{(i)}, -\infty < \ell < \infty, i = 1, 2, \ldots$ be independent and identically distributed copies of ε_0. Clearly

$$\mathbf{Y}_i^{(k)} = (Y_{i,1}^{(k)}, Y_{i,2}^{(k)}) \quad \text{with} \quad Y_{i,1}^{(k)} = \sum_{\ell=0}^{\infty} \rho^{\ell} \varepsilon_{-\ell}^{(i)} \quad \text{and} \quad Y_{i,2}^{(k)} = \sum_{\ell=0}^{\infty} \rho^{\ell} \varepsilon_{k-\ell}^{(i)}$$

are independent and identically distributed copies of $\mathbf{Y}^{(k)}$. Elementary algebra gives

$$\sum_{i=1}^{n} Y_{i,1}^{(k)} = \sum_{\ell=0}^{\infty} \rho^{\ell} \sum_{i=1}^{n} \varepsilon_{-\ell}^{(i)} \quad \text{and} \quad \sum_{i=1}^{n} Y_{i,2}^{(k)} = \sum_{\ell=0}^{k-1} \rho^{\ell} \sum_{i=1}^{n} \varepsilon_{k-\ell}^{(i)} + \rho^k \sum_{\ell=0}^{\infty} \rho^{\ell} \sum_{i=1}^{n} \varepsilon_{-\ell}^{(i)}.$$

For every $L \geq 0$ by (3.19) we have that (recall that under our conditions the centering factors a_n in (3.19) can be chosen 0)

$$\frac{1}{b_n} \left(\sum_{i=1}^{n} \varepsilon_{\ell}^{(i)}, -L \leq \ell \leq L \right) \xrightarrow{\mathcal{D}} \left(\xi_{\ell}^{(\alpha)}, -L \leq \ell \leq L \right),$$

where $\xi_{\ell}^{(\alpha)}, -\infty < \ell < \infty$ are independent and identically distributed copies of $\xi^{(\alpha)}$.

Theorem 32 (Acosta and Giné, [34], 1979)**.** *Let* ρ *be a non-degenerate stable probability measure of order* $\alpha \in (0, 2]$ *on* B *and let* X *belong to the domain of attraction of* ρ *with norming constants* $\{a_n\} \in \mathbb{R}_+$ *and centering constants* $\{b_n\} \in B$,*. Then for every* $\beta \in (0, \alpha)$

$$\lim_n E||S_n/a_n - b_n||^{\beta} = \int ||x||^{\beta} d\rho(x).$$

Let $0 < \kappa < \alpha$. It follows from theorem32 that

$$E\left|\frac{1}{b_n}\sum_{i=1}^{n}\varepsilon_\ell^{(i)}\right|^\kappa \leq C_1,$$

and therefore for every $x > 0$ we have that

$$\lim_{L\to\infty}\limsup_{n\to\infty} P\left\{\sum_{\ell=L+1}^{\infty}\rho^\ell\left|\frac{1}{b_n}\sum_{i=1}^{n}\varepsilon_\ell^{(i)}\right| > x\right\} = 0$$

and similarly

$$\lim_{L\to\infty} P\left\{\sum_{\ell=L+1}^{\infty}\rho^\ell|\xi_\ell^{(\alpha)}| > x\right\} = 0.$$

This completes the proof of the lemma. □

Let $\mathbf{i}$ denote the imaginary unit.

Lemma 8. *Let* $\mathbf{Y}$ *be a stable vector variable with characteristic function* $\psi(s,t)$*. Then there exists a measure* ν *on the Borel sets of* R^2 *such that for some* $\mathcal{C}_1, \mathcal{C}_2$ *and any* $\gamma > 0$

$$\psi(s,t) = \exp\left\{\mathbf{i}(\mathcal{C}_1 s + \mathcal{C}_2 t) + \int\limits_{|\mathbf{u}|>\gamma}(e^{\mathbf{i}(su_1+tu_2)} - 1)\nu(du_1, du_2) + \int\limits_{0<|\mathbf{u}|\leq\gamma}(e^{\mathbf{i}(su_1+tu_2)} - 1 - \mathbf{i}(su_1+tu_2))\nu(du_1, du_2)\right\},$$

where $\mathbf{u} = (u_1, u_2)$.

The result can be found, for example, in Gikhman and Skorohod ([88], Chapter 5, 1969). ν is called the Lévy measure in the canonical representation of the characteristic function of $\mathbf{Y}$. The stable vectors in our paper will be centered, i.e. $c_1 = c_2 = 0$.

Lemma 9. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, then we have*

$$\lim_{T\to\infty}\frac{T^{\alpha-2}}{L_*(T)}EX_0 I\{|X_0| \leq vT\}X_k I\{|X_k| \leq wT\} = \frac{\alpha}{2-\alpha}\frac{\rho^k}{1-|\rho|^\alpha}(\min(v, w|\rho|^{-k}))^{2-\alpha}.$$

Proof.

Theorem 33 (Resnick and Greenwood,[84], 1979). *Let* $\{Y_n\} = \{Y_n^{(1)}, Y_n^{(2)}\}$ *be i.i.d. sectors in* $\mathbb{R}^2$ *and let* $X(t)$ *be a Lévy process with Lévy measure* ν*, defined by* $\nu \circ \tau = \tilde{\nu}$*,* $\tau x = ((signx_1)|x_1|^{1/\alpha_1}, (signx_2)|x_2|^{1/\alpha_2})$*, and* $\tilde{\nu}$ *is a measure of a certain form. The following are equivalent: there exist* $a_n^{(i)} > 0$*,* $b_n \in \mathbb{R}^2$ *such that*

(i) $(S_n^{(1)}/a_n^{(1)}, S_n^{(2)}/a_n^{(2)}) - b_n \Rightarrow X(1)$ *in* $\mathbb{R}^2$

(ii) $(S_{[n\cdot]}^{(1)}/a_n^{(1)}, S_{[n\cdot]}^{(2)}/a_n^{(2)}) - (\cdot)b_n \Rightarrow X(\cdot)$ *in* $\mathbb{R}^2$

(iii) Let $Y_{ni} = (Y_i^{(1)}/a_n^{(1)}, Y_i^{(2)}/a_n^{(2)})$ *for all* $A \in \mathfrak{B}(\mathbb{R} - \{0\})$ *such that* $\nu(\partial A) = 0$*,* $\nu(A) < \infty$ *we have*

$$\lim_{n\to\infty} nP(Y_{n1} \in A) = \nu(A),$$

where ν *is Lévy measure of* X*, and* $\mathfrak{B}$ *denotes borel sigma-algebra.*

It follows from Theorem 33 that

$$\lim_{n\to\infty} nP\left\{\frac{(X_0, X_k)}{b_n} \in A\right\} = \nu(A), \tag{3.33}$$

where b_n is defined in (3.29) and A is any Borel set of R^2, not containing $(0,0)$, $\nu(A) < \infty$ and the ν–measure of the boundary of A is 0. Since $nL_*(b_n)/b_n^\alpha \to 1$, with the choice of $n = \lfloor T^\alpha/L_*(T)\rfloor$ we get from (3.33) that

$$\lim_{T\to\infty} \frac{T^\alpha}{L_*(T)} P\{(X_0, X_k)/T \in \mathbf{A}\} = \nu(\mathbf{A}), \tag{3.34}$$

where ν is the Lévy measure in the canonical representation of the characteristic function of $\mathbf{Z}^{(k)}$. By elementary arguments we conclude from (3.34)

$$\lim_{T\to\infty} \frac{T^{\alpha-2}}{L_*(T)} EX_0 I\{|X_0| \le vT\} X_k I\{|X_k| \le wT\} = \int_{-v}^{v}\int_{-w}^{w} xy\nu(dx, dy).$$

Since $\xi^{(\alpha)}$ is a stable random variable, its characteristic function can be written as $\exp(-\psi(t))$ and with this notation we get

$$E\exp(\mathbf{i}(sZ_1^{(k)}+tZ_2^{(k)}))=\exp\left(-\sum_{\ell=0}^{\infty}\psi(s\rho^\ell+t\rho^{k+\ell})-\sum_{\ell=0}^{k-1}\psi(t\rho^\ell)\right).$$

If $\hat{\nu}_\ell$ denotes the Lévy measure associated with the characteristic function $\exp(-\psi(s\rho^\ell+t\rho^{k+\ell}))$ and $\tilde{\nu}_\ell$ corresponds to $\exp(-\psi(t\rho^\ell))$, then we have

$$\nu(\mathbf{A})=\sum_{\ell=0}^{\infty}\hat{\nu}_\ell(\mathbf{A})+\sum_{\ell=0}^{k-1}\tilde{\nu}_\ell(\mathbf{A}).$$

Hence

$$\int_{-v}^{v}\int_{-w}^{w}xy\nu(dx,dy)=\sum_{\ell=0}^{\infty}\int_{-v}^{v}\int_{-w}^{w}xy\hat{\nu}_\ell(dx,dy).$$

Next we note that there is a positive constant A^* such that

$$\lim_{x\to\infty}\frac{P\{|\xi^{(\alpha)}|>x\}}{x^{-\alpha}}=A^*$$

and therefore by Bingham, Goldie and Teugels ([20], p. 346, 1989) we obtain that

$$\lim_{x\to\infty}\frac{E(\xi^{(\alpha)})^2I\{|\xi^{(\alpha)}|\le x\}}{x^2P\{|\xi^{(\alpha)}|>x\}}=\frac{\alpha}{2-\alpha}$$

resulting in

$$\lim_{x\to\infty}\frac{E(\xi^{(\alpha)})^2I\{|\xi^{(\alpha)}|\le x\}}{x^{2-\alpha}}=A^*\frac{\alpha}{2-\alpha}.$$

The last relation implies

$$\lim_{T\to\infty}T^{\alpha-2}E\left[\rho^{2\ell+k}(\xi^{(\alpha)})^2I\{|\xi^{(\alpha)}|\le T\min(v|\rho|^{-\ell},w|\rho|^{-(\ell+k)})\}\right]$$
$$=A^*\frac{\alpha}{2-\alpha}\rho^{2\ell+k}(\min(v|\rho|^{-\ell},w|\rho|^{-(\ell+k)}))^{2-\alpha}.$$

We note that $\exp(-\psi(s\rho^\ell+t\rho^{k+\ell}))$ is the characteristic function of the vector $(\rho^\ell\xi^{(\alpha)},\rho^{k+\ell}\xi^{(\alpha)})$, so repeating the arguments leading to (3.33) and (3.34) for this vector instead of (X_0,X_k) we get

$$\lim_{T\to\infty}\rho^{k+2\ell}\frac{T^{\alpha-2}}{A^*}E\xi^{(\alpha)}I\{|\rho^\ell\xi^{(\alpha)}|\le vT\}\xi^{(\alpha)}I\{|\rho^{\ell+k}\xi^{(\alpha)}|\le wT\}=\int_{-v}^{v}\int_{-w}^{w}xy\hat{\nu}_\ell(dx,dy),$$

and therefore

$$\int_{-v}^{v}\int_{-w}^{w} xy\hat{\nu}_\ell(dx,dy) = \frac{\alpha}{2-\alpha}\rho^k|\rho|^{\alpha\ell}(\min(v,w|\rho|^{-k}))^{2-\alpha}.$$

Summing for $\ell = 0,1,\ldots$, we get Lemma 9. □

Lemma 10. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, then for every* $k = 0,1,2,\ldots$

$$\lim_{n\to\infty}\frac{nE(u_{0,n}(s)-m_n(s))(u_{k,n}(t)-m_n(t))}{A_n^2} = \frac{\alpha}{2-\alpha}\rho^k(\min(s,t|\rho|^{-k}))^{2-\alpha}. \quad (3.35)$$

Proof. If $1<\alpha<2$, then

$$\lim_{n\to\infty} m_n(t) = EX_0 \quad \text{for any } \; t>0. \quad (3.36)$$

If $0<\alpha<1$, then

$$|m_n(t)| \le \int_{-tH^{-1}(d/n)}^{tH^{-1}(d/n)} |x|dF(x) \quad (3.37)$$

$$= -\int_0^{tH^{-1}(d/n)} xdH(x) = -xH(x)\Big|_0^{tH^{-1}(d/n)} + \int_0^{tH^{-1}(d/n)} H(x)dx.$$

By (3.30) and Bingham, Goldie and Teugels ([20], p. 26, 1989) we have for $0<\alpha<1$

$$\lim_{y\to\infty}\frac{\int_0^y H(x)dx}{yH(y)/(1-\alpha)} = 1, \quad (3.38)$$

and therefore

$$m_n(t) = O\left(H^{-1}(d/n)\frac{d}{n}\right). \quad (3.39)$$

If $\alpha = 1$, by assumption e_0 is symmetric, so under (3.10) we have that $X_1 = e_1 + c_1$ and therefore

$$m_n(t) = O(1) + E[e_0 I\{|X_0| \le tH^{-1}(d/n)\}] \quad (3.40)$$

$$= O(1) + \int_{tH^{-1}(d/n)-c_1}^{tH^{-1}(d/n)+c_1} x dP\{e_1 \le x\}$$

$$= O\left(H^{-1}(d/n)\frac{d}{n}\right) + \int_{tH^{-1}(d/n)-c_1}^{tH^{-1}(d/n)+c_1} P\{e_1 \le x\} dx$$

$$= O\left(H^{-1}(d/n)\frac{d}{n}\log H^{-1}(d/n)\right).$$

Thus we get from (3.36)–(3.40) for all $0 < \alpha < 2$ that

$$\frac{n m_n(s) m_n(t)}{A_n^2} \to 0. \tag{3.41}$$

Lemma 9 yields

$$\lim_{n\to\infty} \frac{n}{A_n^2} \frac{L(H^{-1}(d/n))}{L_*(H^{-1}(d/n))} E X_0 I\{|X_0| \le sH^{-1}(d/n)\} X_k I\{|X_k| \le tH^{-1}(d/n)\}$$
$$= \frac{\alpha}{2-\alpha} \frac{\rho^k}{1-|\rho|^\alpha} (\min(s, t|\rho|^{-k}))^{2-\alpha}.$$

By (3.31) we have

$$\lim_{n\to\infty} \frac{L(H^{-1}(d/n))}{L_*(H^{-1}(d/n))} = \frac{1}{1-|\rho|^\alpha},$$

which completes the proof of the lemma. □

Lemma 11. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, we have for all* $1/2 \le s \le t \le 3/2$ *and* $0 \le x \le 1$ *that*

$$\lim_{n\to\infty} \frac{1}{A_n^2} E\left(\sum_{k=1}^{\lfloor nx \rfloor}(u_{k,n}(s) - m_n(s))\right)\left(\sum_{k=1}^{\lfloor nx \rfloor}(u_{k,n}(t) - m_n(t))\right)$$
$$= x\frac{\alpha}{2-\alpha}\left(s^{2-\alpha} + \sum_{k=1}^{\infty}\rho^k[(\min(s, t|\rho|^{-k})^{2-\alpha} + \min(t, s|\rho|^{-k})^{2-\alpha}]\right).$$

Proof. We note that

$$E\Bigg(\sum_{k=1}^{\lfloor nx \rfloor}(u_{k,n}(s) - m_n(s))\Bigg)\Bigg(\sum_{k=1}^{\lfloor nx \rfloor}(u_{k,n}(t) - m_n(t))\Bigg)$$

$$= \lfloor nx \rfloor E(u_{0,n}(s) - m_n(s))(u_{0,n}(t) - m_n(t))$$
$$+ \sum_{k=1}^{\lfloor nx \rfloor -1} (\lfloor nx \rfloor - k) E(u_{0,n}(s) - m_n(s))(u_{k,n}(t) - m_n(t))$$
$$+ \sum_{k=1}^{\lfloor nx \rfloor -1} (\lfloor nx \rfloor - k) E(u_{0,n}(t) - m_n(t))(u_{k,n}(s) - m_n(s)).$$

Let

$$e_k^* = \sum_{\ell=0}^{k-1} \rho^\ell \varepsilon_{k-\ell} \quad \text{and} \quad X_k^* = c_1 + e_k^*. \tag{3.42}$$

It follows from Cline([24], 1983) that there is a constant C_1 such that

$$P\{|X_k^*| > x\} \le C_1 x^{-\alpha} L(x) \quad \text{for all } k \text{ and } 0 \le x < \infty. \tag{3.43}$$

Clearly as in (3.32),

$$X_k - X_k^* = e_k - e_k^* = \sum_{\ell=k}^{\infty} \rho^\ell \varepsilon_{k-\ell} = \sum_{j=0}^{\infty} \rho^{k+j} \varepsilon_{-j} = \rho^k (X_0 - c_1). \tag{3.44}$$

Next we write

$$|E(u_{0,n}(s) - m_n(s))(u_{k,n}(t) - m_n(t))|$$
$$= |E u_{0,n}(t) u_{0,n}(s) - m_n(t) m_n(s)|$$
$$\le |E(X_{0,n}(X_k - X_k^*) I\{|X_0| \le s H^{-1}(d/n)\} I\{|X_k| \le t H^{-1}(d/n)\}|$$
$$+ |E(X_0 X_k^* I\{|X_0| \le s H^{-1}(d/n)\} I\{|X_k| \le t H^{-1}(d/n)\} - m_n(s) m_n(t)|$$
$$\le A_{1,k,n} + A_{2,k,n} + A_{3,k,n}$$

with

$$A_{1,k,n} = E|X_0(X_k - X_k^*) I\{|X_0| \le s H^{-1}(d/n)\} I\{|X_k| \le t H^{-1}(d/n)\}|,$$
$$A_{2,k,n} = E\big[|X_0 X_k^*| I\{|X_0| \le s H^{-1}(d/n)\}$$
$$\times \big| I\{|X_k| \le t H^{-1}(d/n)\} - I\{|X_k^*| \le t H^{-1}(d/n)\}\big|\big]$$

and

$$A_{3,k,n} = |E(X_0 X_k^*) I\{|X_0| \le s H^{-1}(d/n)\} I\{|X_k^*| \le t H^{-1}(d/n)\} - m_n(s) m_n(t)|.$$

Using (3.41) and (3.44) we conclude

$$\begin{aligned} A_{1,k,n} &\le |\rho|^k E|X_0||X_0 - c_1|I\{|X_0| \le sH^{-1}(d/n)\} \\ &\le C_2|\rho|^k (H^{-1}(d/n))^2 d/n \end{aligned} \tag{3.45}$$

with some constant C_2. Next we note that

$$\begin{aligned} A_{2,k,n} \le E\big[&|X_0 X_k^*| I\{|X_0| \le sH^{-1}(d/n)\} \\ &\times I\{tH^{-1}(d/n) - |\rho|^k|X_0| \le |X_k^*| \le tH^{-1}(d/n)\}\big] \\ &+ E\big[|X_0 X_k^*| I\{|X_0| \le sH^{-1}(d/n)\} \\ &\times I\{tH^{-1}(d/n) \le |X_k^*| \le tH^{-1}(d/n) + |\rho|^k|X_0|\}\big] \\ = A_{2,k,n}^{(1)} &+ A_{2,k,n}^{(2)}. \end{aligned} \tag{3.46}$$

Using the independence of X_0 and X_k^* we get

$$\begin{aligned} A_{2,k,n}^{(1)} &\le E|X_0| I\{|X_0| \le sH^{-1}(d/n)\} \\ &\times E|X_k^*| I\{tH^{-1}(d/n) - |\rho|^k H^{-1}(d/n) \le |X_k^*| \le tH^{-1}(d/n)\}. \end{aligned}$$

By (3.43) we have that

$$\begin{aligned} &E|X_k^*| I\{tH^{-1}(d/n) - |\rho|^k H^{-1}(d/n) \le |X_k^*| \le tH^{-1}(d/n)\} \\ &= -xP\{|X_k^*| > x\}\Big|_{tH^{-1}(d/n)-|\rho|^k H^{-1}(d/n)}^{tH^{-1}(d/n)\}} + \int_{tH^{-1}(d/n)-|\rho|^k H^{-1}(d/n)}^{tH^{-1}(d/n)\}} P\{|X_k^*| > x\}dx \\ &\le \int_{tH^{-1}(d/n)-|\rho|^k H^{-1}(d/n)}^{tH^{-1}(d/n)\}} P\{|X_k^*| > x\}dx \\ &\le C_3|\rho|^k H^{-1}(d/n) d/n, \end{aligned} \tag{3.47}$$

where C_3 is a constant. Hence, on account of (3.36), (3.39) and (3.40) we obtain that with some constant C_4

$$A_{2,k,n}^{(1)} \le C_4 \rho^k (H^{-1}(d/n))^2 d/n$$

and similarly

$$A^{(2)}_{2,k,n} \le C_4 \rho^k (H^{-1}(d/n))^2 d/n,$$

resulting in

$$A_{2,k,n} \le C_5 \rho^k (H^{-1}(d/n))^2 d/n. \tag{3.48}$$

Using again the independence of X_0 and X_k^* we get

$$A_{3,k,n} = |m_n(s)||EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} - m_n(t)|.$$

It is easy to see that

$$\begin{aligned} &EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} \\ &\quad = EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} I\{|X_0| > |\rho|^{-k/2} H^{-1}(d/n)\} \\ &\qquad + EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} I\{|X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\} \end{aligned}$$

and by the independence of X_0 and X_k^* and (3.43) we have

$$\begin{aligned} |EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} I\{|X_0| > |\rho|^{-k/2} H^{-1}(d/n)\}| &\le C_5 |m_n(t)| H(|\rho|^{-k/2} H^{-1}(d/n)) \\ &\le C_6 |m_n(t)| |\rho|^{k\alpha/2} d/n. \end{aligned}$$

Next we note that

$$\begin{aligned} &\big| E\big[X_k^* I\{|X_k^*| \le tH^{-1}(d/n), |X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\}\big] \\ &\qquad - E\big[(X_k^* + \rho^k (X_0 - c_1)) I\{|X_k^* + \rho^k (X_0 - c_1))| \le tH^{-1}(d/n), \\ &\qquad\qquad |X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\}\big]\big| \\ &\quad \le |\rho|^k E\big[|X_0 - c_1| I\{|X_k^* + \rho^k (X_0 - c_1))| \le tH^{-1}(d/n), \\ &\qquad\qquad |X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\}\big] \\ &\quad + E\big[|X_k^*| |I\{|X_k^*| \le tH^{-1}(d/n), |X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\} \\ &\qquad - I\{|X_k^* + \rho^k (X_0 - c_1))| \le tH^{-1}(d/n), |X_0| \le |\rho|^{-k/2} H^{-1}(d/n)\}|\big] \end{aligned}$$

$$\leq |\rho|^k(|\rho|^{-k/2}H^{-1}(d/n)+|c_1|)$$
$$+E|X_k^*|I\{(t-|\rho|^{k/2})H^{-1}(d/n)-|c_1||\rho|^k \leq |X_k^*| \leq tH^{-1}(d/n)\}$$
$$+E|X_k^*|I\{tH^{-1}(d/n) \leq |X_k^*| \leq (t+|\rho|^{-k/2})H^{-1}(d/n)+|c_1||\rho|^k\}$$
$$\leq C_7(|\rho|^{k/2}H^{-1}(d/n)+|\rho|^kH^{-1}(d/n)d/n)$$

by (3.47). Similarly

$$|EX_kI\{|X_k| \leq tH^{-1}(d/n)\} - EX_kI\{|X_k| \leq tH^{-1}(d/n), |X_0| \leq |\rho|^{-k/2}H^{-1}(d/n)\}|$$
$$\leq C_8(|\rho|^{k/2}H^{-1}(d/n)+|\rho|^kH^{-1}(d/n)d/n).$$

Hence

$$A_{3,k,n} \leq C_9|\rho|^{\tau k}(H^{-1}(d/n))^2 d/n, \quad \text{where} \quad \tau = \min\{1,\alpha\}/2. \tag{3.49}$$

Putting together (3.45), (3.48) and (3.49) we get that

$$\lim_{K\to\infty}\limsup_{n\to\infty}\frac{1}{A_n^2}\sum_{k=K}^{\lfloor nx\rfloor-1}|(\lfloor nx\rfloor-k)E(u_{0,n}(s)-m_n(s))(u_{k,n}(t)-m_n(t))| = 0. \tag{3.50}$$

The lemma now follows from Lemma 10 and (3.50). □

3.3 A weak convergence result

Define the two–parameter process

$$L_n(t,x) = \frac{1}{A_n}\sum_{i=1}^{\lfloor nx\rfloor}(X_iI\{|X_i| \leq tH^{-1}(d/n)\} - m_n(t)),$$

for $0 \leq x \leq 1, 1/2 \leq t \leq 3/2$. First we show the tightness of $L_n(t)$.

Lemma 12 (Berkes, Horváth, Ling, Schauer, [13], 2011). *Let* $\{\zeta_i(s), 0 \leq s \leq 1, i \geq 1\}$ *be non-decreasing processes in* $\mathcal{D}[0,1]$, *let* $\zeta(s)$, $0 \leq s \leq 1$, *be a non-decreasing function and define*

$$K_n(t,s) = \frac{1}{n^{1/2}}\sum_{i=1}^{[nt]}(\zeta_i(s)-\zeta(s)).$$

If there exist a $\tau > 2$, $C > 0$ *and a sequence* a_n *such that* $a_n\sqrt{n} \to 0$ *and*

$$E(K_n(t_2,s) - K_n(t_1,s))^6 \le C|t_2 - t_1|^\tau \ if\ |t_2 - t_1| \ge a_n,$$

$$E(K_n(t,s_2) - K_n(t,s_1))^6 \le C|s_2 - s_1|^\tau \ if\ |s_2 - s_1| \ge a_n,$$

and

$$n^{1/2} \sup_{|s_2 - s_1| \le a_n} |\zeta(s_2) - \zeta(s_1)| \to 0,$$

then $K_n(t,s)$ *is tight.*

The proof is based on a generalization of Lemma 12 in Berkes, Horváth, Ling, Schauer ([13], 2011). We introduce

$$X_{i,1} = \max(X_i, 0), \qquad X_{i,2} = \min(X_i, 0)$$

and

$$m_{n,1}(t) = EX_{0,1}I\{|X_0| \le tH^{-1}(d/n)\}, \qquad m_{n,2}(t) = EX_{0,2}I\{|X_0| \le tH^{-1}(d/n)\}.$$

Similarly to $L_n(t,x)$, we define

$$L_{n,1}(t,x) = \frac{1}{A_n}\sum_{i=1}^{\lfloor nx \rfloor}(X_{i,1}I\{|X_i| \le tH^{-1}(d/n)\} - m_{n,1}(t)),$$

and $L_{n,2}(t,x)$ is defined in a similar fashion. Clearly, if both $L_{n,1}$ and $L_{n,2}$ are tight, then $L_n(t,x)$ is tight as well. We prove only tightness of $L_{n,1}$, the same argument can be used in case of $L_{n,2}$. Let

$$g_n = \frac{1}{d^{1/2}\log\log n}.$$

Lemma 13. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, then*

$$m_{n,1}(t) \ \text{ is a non-decreasing function on } [1/2, 3/2], \tag{3.51}$$

$$\frac{n}{A_n} \sup_{|t_2 - t_1| \le g_n} |m_{n,1}(t_2) - m_{n,1}(t_1)| \to 0, \quad n \to \infty, \tag{3.52}$$

$$E|L_{n,1}(t_2,x)-L_{n,1}(t_1,x)|^6 \le C_1|t_2-t_1|^\tau, \quad \textit{if } |t_2-t_1|\ge g_n, \tag{3.53}$$

and

$$E|L_{n,1}(t,x_2)-L_{n,1}(t,x_1)|^6 \le C_1|x_2-x_1|^\tau, \quad \textit{if } |x_2-x_1|\ge g_n, \tag{3.54}$$

with some $\tau > 2$ and constant C_1.

Proof. The definition of $m_{n,1}(t)$ implies immediately (3.51).

By the definition of $m_{n,1}(t)$ we have for all $1/2 \le t_1 \le t_2 \le 3/2$ that

$$\begin{aligned} 0 &\le m_{n,1}(t_2)-m_{n,1}(t_1) = EX_{0,1}(I\{t_1H^{-1}(d/n) < |X_0| \le t_2H^{-1}(d/n)\}) \\ &\le \int_{t_1H^{-1}(d/n)}^{t_2H^{-1}(d/n)} xdH(x) \\ &\le C_2\Big(|t_2H^{-1}(d/n)H(t_2H^{-1}(d/n)) - t_1H^{-1}(d/n)H(t_1H^{-1}(d/n))| \\ &\qquad + |t_2-t_1|H^{-1}(d/n)H(t_1H^{-1}(d/n))\Big) \\ &\le C_3|t_2-t_1|\frac{d}{n}H^{-1}(d/n) \end{aligned}$$

on account of integration by parts and (3.30), establishing (3.52).

Next we introduce

$$Y_i = \sum_{k=0}^{\lfloor K\log n\rfloor} \rho^k \varepsilon_{i-k} + c_1, \qquad Y_{i,1} = \max(Y_i,0) \tag{3.55}$$

and $\xi_i = \eta_i - E\eta_i$ with

$$\eta_i = \eta_i(t_1,t_2) = Y_{i,1}I\{t_1H^{-1}(d/n) < |Y_i| \le t_2H^{-1}(d/n)\}.$$

Since $E|\varepsilon_0|^{\alpha/2} < \infty$, using Markov's inequality we see that for every $\beta > 0$ there is a constant $K = K(\beta)$ such that

$$\left| E(L_{n,1}(t_2, x) - L_{n,1}(t_1, x))^6 - \frac{1}{A_n^6} \sum_{1 \le i_1, \ldots, i_6 \le \lfloor nx \rfloor} E\xi_{i_1} \ldots \xi_{i_6} \right| \le C_5 n^{-\beta}. \qquad (3.56)$$

We note that by definition, $\{\xi_i\}$ is a stationary, $\lfloor K \log n \rfloor$–dependent sequence with zero mean. Let us divide the indices $i_1, \ldots, i_6$ into groups so that the difference between the indices within a group are less than $\lfloor K \log n \rfloor$ and between groups is larger than $\lfloor K \log n \rfloor$. Clearly $E\xi_{i_1} \ldots \xi_{i_6} = 0$, if there is at least one group containing a single element. So it suffices to consider the cases when all groups contain at least two elements. This allows the cases of one single group with 6 elements (D_1), two groups with 3+3 (D_2) or 4+2 (D_3) elements and finally 3 groups with 2 elements in each (D_4). If there is only one group, then via Hölder's inequality we have

$$|E\xi_{i_1} \ldots \xi_{i_6}| \le E|\xi_0|^6 \le 2^6 (E|\eta_0|^6 + |E\eta_0|^6)$$

Since the cardinality of D_1 is bounded by constant times $n(\log n)^5$ we conclude

$$\begin{aligned} &\left| \frac{1}{A_n^6} \sum_{D_1} E\xi_{i_1} \ldots \xi_{i_6} \right| \\ &\quad \le C_6 \Bigg(\frac{n(\log n)^5}{A_n^6} [EX_0^6 I\{t_1 H^{-1}(d/n) \le |X_0| \le t_2 H^{-1}(d/n)\} \\ &\qquad\qquad + (EX_0 I\{t_1 H^{-1}(d/n) \le |X_0| \le t_2 H^{-1}(d/n)\})^6] + n^{-\beta} \Bigg). \end{aligned}$$

Integration by parts and (3.30) yield

$$EX_0^6 I\{t_1 H^{-1}(d/n) \le |X_0| \le t_2 H^{-1}(d/n)\} \le C_7 |t_2 - t_1| \frac{d}{n} (H^{-1}(d/n))^6,$$

resulting in

$$\left| \frac{1}{A_n^6} \sum_{D_1} E\xi_{i_1} \ldots \xi_{i_6} \right| \le C_8 \left(\frac{(\log n)^5}{d^2} |t_2 - t_1| + n^{-\beta} \right).$$

Using again the $\lfloor K \log n \rfloor$ dependence of $\{\xi_i\}$ and the fact that the cardinality of D_2 is constant times $n^2(\log n)^4$ we conclude via Hölder's inequality

$$\begin{aligned}
&\left|\frac{1}{A_n^6}\sum_{D_2} E\xi_{i_1}\ldots\xi_{i_6}\right| \\
&= \left|\frac{1}{A_n^6}\sum_{D_2} E\xi_{i_1}\xi_{i_2}\xi_{i_3}E\xi_{i_4}\xi_{i_5}\xi_{i_6}\right| \\
&\le C_8\bigg(\frac{n^2(\log n)^4}{A_n^6}[EX_0^3 I\{t_1H^{-1}(d/n) \le |X_0| \le t_2H^{-1}(d/n)\} \\
&\qquad\qquad + (EX_0 I\{t_1H^{-1}(d/n) \le |X_0| \le t_2H^{-1}(d/n)\})^3]^2 + n^{-\beta}\bigg) \\
&\le C_9\left(\frac{(\log n)^4}{d}(t_2-t_1)^2 + n^{-\beta}\right).
\end{aligned}$$

Similar arguments give

$$\left|\frac{1}{A_n^6}\sum_{D_3} E\xi_{i_1}\ldots\xi_{i_6}\right| \le C_{10}\left(\frac{(\log n)^4}{d}(t_2-t_1)^2 + n^{-\beta}\right).$$

Following the proof of Lemma 11 we obtain

$$\begin{aligned}
\left|\frac{1}{A_n^6}\sum_{D_4} E\xi_{i_1}\ldots\xi_{i_6}\right| &\le C_{11}\left(\frac{1}{A_n^6}\left(n\sum_{i=0}^{\infty}\xi_0\xi_i\right)^3 + n^{-\beta}\right) \\
&\le C_{11}\left(|t_2-t_1|^3 + n^{-\beta}\right).
\end{aligned}$$

Putting together our estimates and using the choice of g_n we conclude for all $|t_2 - t_1| \ge g_n$

$$\begin{aligned}
E(L_{n,1}(t_2,x) - L_{n,1}(t_2,x))^6 &\le C_{12}\left(\frac{(\log n)^5}{d^2}|t_2-t_1| + \frac{(\log n)^4}{d}|t_2-t_1|^2 + |t_2-t_1|^3 + n^{-\beta}\right) \\
&\le C_{13}|t_2-t_1|^\tau
\end{aligned}$$

with any $2 < \tau \le 3$ on account of assumption (3.24). Hence the proof of (3.53) is complete.

The proof of (3.54) goes along the lines of the arguments used to establish (3.53) and therefore it is omitted. □

Lemma 14. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, then* $L_n(t,x)$ *is tight in*
$\mathcal{D}([1/2,3/2]\times[0,1])$.

Proof. It follows from a minor modification of Lemma 5.1 in Berkes, Horváth, Ling and Schauer ([13], 2011) that both $L_{n,1}$ and $L_{n,2}$ are tight. Since $L_n = L_{n,1} + L_{n,2}$, the result is proven. □

Next we consider the convergence of the finite dimensional distributions. It is based in the following lemma:

Lemma 15. *We assume that (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold. Let* $N = \lfloor(\log n)^\gamma\rfloor$ *with some* $\gamma > 0$. *Then*

$$E\left(\sum_{i=1}^{N}(X_i I\{|X_i|\le tH^{-1}(d/n)\} - E[X_i I\{|X_i|\le tH^{-1}(d/n)\}]\right)^4 \tag{3.57}$$
$$\le C_{13}\left(N(\log N)^3(H^{-1}(d/n))^4\frac{d}{n} + N^2(H^{-1}(d/n))^4\left(\frac{d}{n}\right)^2\right)$$

with some constant C_{13} *and*

$$\lim_{n\to\infty}\frac{Nn}{A_n^2}E\left(\sum_{k=1}^{N}(u_{k,n}(s)-m_n(s))\right)\left(\sum_{k=1}^{N}(u_{k,n}(t)-m_n(t))\right) \tag{3.58}$$
$$=\frac{\alpha}{2-\alpha}\left(s^{2-\alpha}+\sum_{k=1}^{\infty}\rho^k[(\min(s,t|\rho|^{-k})^{2-\alpha}+\min(t,s|\rho|^{-k})^{2-\alpha}]\right).$$

Proof. We recall the definition of ξ_i from the proof of Lemma 13. For any $\beta > 0$, choosing K in the definition of Y_i in (3.55) we get that

$$E\left(\sum_{i=1}^{N}(X_i I\{|X_i|\le tH^{-1}(d/n)\} - E[X_i I\{|X_i|\le tH^{-1}(d/n)\}]\right)^4 \le C_{14}\left(E\left(\sum_{i=1}^{N}\xi_i\right)^4 + n^{-\beta}\right).$$

We write

$$E\left(\sum_{i=1}^{N}\xi_i\right)^4 = \sum_{i_1,\ldots,i_4}^{N} E\xi_{i_1}\ldots\xi_{i_4}.$$

We note again that the $\{\xi_i\}$ is a stationary $K\log n$ dependent sequence with 0 mean. Let us divide the indices $i_1,\ldots,i_4$ into blocks so that the difference between the indices within a block is less than $K\log n$ and between blocks is larger than $K\log n$. Clearly $E\xi_{i_1}\ldots\xi_{i_4}=0$, if there is at least one block containing only a single element. So we need to consider the cases of one single block with 4 elements (D_1) and two blocks with 2+2 elements (D_2). The number of the elements in D_1 is not greater than constant times $N(\log N)^3$ and as we showed in the proof of Lemma 13

$$E\xi_0^4 \le C_{14}\left((H^{-1}(d/n))^4\frac{d}{n}+n^{-\beta}\right),$$

assuming that K in (3.55) is sufficiently large. Hence

$$\left|\sum_{D_1}^{N} E\xi_{i_1}\ldots\xi_{i_4}\right| \le C_{15}\left(N(\log N)^3(H^{-1}(d/n))^4\frac{d}{n}+n^{-\beta}\right).$$

As in the proof of Lemma 13 we get that

$$\left|\sum_{D_2}^{N} E\xi_{i_1}\ldots\xi_{i_4}\right| \le C_{16}N^2\left(\sum_{i=0}|E\xi_0\xi_i|\right)^2$$

and

$$\sum_{i=0}^{\infty}|E\xi_0\xi_i| \le \left(C_{17}(H^{-1}(d/n))^2\frac{d}{n}+n^{-\beta}\right),$$

completing the proof of (3.57). The proof of (3.58) goes along the lines of the arguments used to establish Lemma 11. □

Lemma 16. *If (3.10)–(3.18), (3.21)–(3.24) and (3.28) hold, then*

$$L_n(t,x) \quad\longrightarrow\quad \Gamma(t,x) \quad \textit{weakly in} \quad \mathcal{D}([1/2,3/2])\times[0,1]),$$

where $\Gamma(t,x)$ *is a Gaussian process with* $E\Gamma(t,x)=0$ *and*

$$E\Gamma(t,x)\Gamma(s,y)$$
$$=\min(x,y)\frac{\alpha}{2-\alpha}\left((\min(s,t))^{2-\alpha}+\sum_{k=1}^{\infty}\rho^k[(\min(s,t|\rho|^{-k})^{2-\alpha}+\min(t,s|\rho|^{-k})^{2-\alpha}]\right).$$

Proof. By Lemma 14, the process $L_n(t,x)$ is tight, so we need only to show the convergence of the finite dimensional distributions. By the Cramér–Wold device it is sufficient to prove the asymptotic normality of

$$Q_n = \sum_{j=1}^{J}\sum_{\ell=0}^{L}\mu_{j,\ell}(L_n(t_j,x_{\ell+1}) - L_n(t_j,x_\ell))$$

for all J, L, real coefficients $\mu_{j,\ell}, 1/2 \le t_j \le 3/2, 1 \le j \le J$, and $0 = x_0 < x_1 < \ldots < x_L < x_{L+1} = 1$. We recall the definition of X_k^* from the proof of Lemma 11 (cf. (3.44)) and define

$$\bar{L}_n(t,x) = \frac{1}{A_n}\sum_{i=1}^{\lfloor nx\rfloor}(X_k^* I\{|X_k^*| \le tH^{-1}(d/n)\} - EX_k^* I\{|X_k^*| \le tH^{-1}(d/n)\}).$$

Choosing K large enough in the definition of X_k^*, we get from the arguments used in the proof of Lemmas 11, 13 and 15 that

$$E(L_n(t,x) - \bar{L}_n(t,x))^2 \to 0.$$

So we need to establish only the asymptotic normality of

$$\bar{Q}_n = \sum_{\ell=0}^{L}\sum_{j=1}^{J}\mu_{j,\ell}(\bar{L}_n(t_j,x_{\ell+1}) - \bar{L}_n(t_j,x_\ell)).$$

Let

$$z_{k,\ell} = \sum_{j=1}^{J}\mu_{j,\ell}(X_k^* I\{|X_k^*| \le t_j H^{-1}(d/n)\} - E[X_k^* I\{|X_k^*| \le t_j H^{-1}(d/n)\}]).$$

Since for all ℓ

$$E\left(\frac{1}{A_n}\sum_{k=1}^{\lfloor K\log n\rfloor} z_{k,\ell}\right)^2 \to 0,$$

by stationarity and the $\lfloor K\log n\rfloor$–dependence of $z_{k,\ell}$ for any ℓ we get that the variables

$$\frac{1}{A_n}\sum_{k=\lfloor nx_\ell\rfloor+1}^{\lfloor nx_{\ell+1}\rfloor} z_{k,\ell}, \quad 1 \le \ell \le L \text{ are asymprotically independent.}$$

By stationarity we have

$$\frac{1}{A_n}\sum_{k=\lfloor nx_\ell\rfloor+1}^{\lfloor nx_{\ell+1}\rfloor} z_{k,\ell} \stackrel{\mathcal{D}}{=} \frac{1}{A_n}\sum_{k=1}^{\lfloor nx_{\ell+1}\rfloor-\lfloor nx_\ell\rfloor} z_{k,\ell}.$$

Let us divide the integers of $[1,\lfloor nx_{\ell+1}\rfloor-\lfloor nx_\ell\rfloor]$ into consecutive blocks $R_1,V_1,R_2,V_2,\ldots,R_s,V_s$ such that for $1\le i\le s-1$, R_i contains $\lfloor(\log n)^\gamma\rfloor$ integers, V_i contains $\lfloor K\log n\rfloor$ integers, the last two blocks might contain less elements. Let

$$\zeta_{i,1}=\sum_{k\in R_i} z_{k,\ell} \quad\text{and}\quad \zeta_{i,2}=\sum_{k\in V_i} z_{k,\ell}.$$

Due to the $\lfloor K\log n\rfloor$ dependence and stationarity, the variables $\zeta_{i,2}, 1\le i<s$ are independent and identically distributed and the proof of Lemma 11 shows that

$$E\left(\frac{1}{A_n}\sum_{i=1}^{s}\zeta_{i,2}\right)^2\to 0.$$

Using Lemma 15 we get that

$$E\zeta_{i,1}^2\ge C_{18}(\log n)^\gamma(H^{-1}(d/n))^2 d/n$$

and

$$E\zeta_{i,1}^2\le C_{19}\left((\log n)^\gamma(\log\log n)^3(H^{-1}(d/n))^4\frac{d}{n}+(\log n)^{2\gamma}(H^{-1}(d/n))^4\left(\frac{d}{n}\right)^2\right).$$

Since s is proportional to $n/(\log n)^\gamma$, a simple calculation yields

$$\frac{\sum_{i=1}^{s}E\zeta_{i,1}^4}{\left(\sum_{i=1}^{s}E\zeta_{i,1}^2\right)^2}\to 0,$$

Thus the central limit theorem with Lyapunov's remainder term (cf. Petrov [83], p. 154, 1995, p. 154) implies the asymptotic normality of
$\sum_{1\le k\le\lfloor nx_{\ell+1}\rfloor-\lfloor nx_\ell\rfloor} z_{k,\ell}$. This completes the proof of Lemma 16. □

3.4 Proof of the main results

To prove the results of this chapter we will need the following theorems.

Theorem 34 (Withers,[97], 1981)**.** *Let $\{Z_j\}$ be independent r.v.s. on $\mathbb{R}$ with densities $\{p_j(x)\}$ satisfying $\gamma = \max_j E|Z_j|^\delta < \infty$ for some $\delta > 0$ and*

$$\max_j \int |p_j(x) - p_j(x+y)|dx \le C|y|, \text{ where } \quad C < \infty.$$

Suppose that

$$EZ_j \equiv 0 \quad if\ \delta \ge 1$$

and

$$varZ_j \equiv 1 \quad if\ \delta \ge 2,$$

$\{g_k\}$ are the complex numbers, satisfying

$$G_t = S_t(\min(1,\delta))^{\max(1,\delta)} \to \ ast \to \infty,$$

where

$$S_t(\delta) = \sum_{\nu=t}^{\infty} |g_\nu|^\delta$$

and

$$\sum_0^\infty g_k z^k \ne 0 \quad for\ |z| \le 1.$$

(a) Suppose that $g_k = O(k^{-\nu})$ where $\nu > \frac{3}{2}$. For $\{X_t\}$ defined as above, if $\nu + 1/2 > \delta > 2/(\nu - 1)$, then

$$\alpha(k) = O(k^{-\varepsilon}) \quad with\ \varepsilon = (\delta(\nu-1) - 2)/(\delta+1).$$

Under the slightly stronger moment condition $\delta = \nu + \frac{1}{2}$, $\alpha(k) = O(k^{\frac{3}{2}-\nu}(\ln k)^{1/2})$.

(b) Suppose that $g_k = O(e^{-\nu k})$ where $\nu > 0$. Then for $\{X_t\}$, $\alpha(k) = O(e^{-\nu\lambda k})$, where $\lambda = \delta(1+\delta)^{-1}$.

(c) Also, (a) and (b) are true for process

$$Y_i = \sum_{i=0}^{\infty} h_i Z_{t-i}$$

where $\{h_i\}$ *are given by* $\sum_0^\infty h_i z^i = P(z) \sum_0^\infty g_k z^k$ *and* $P(z)$ *is any polynomial.*

Similar result can be found in Gorodetskii([43], 1977).

Theorem 35 (Davydov, [33], 1968)**.** *Let* X_n *satisfy the strong mixing condition. Let r.v.* ξ *be measurable in respect to* $\mathfrak{M}_0^k$ *and* η *be measurable in respect to* $\mathfrak{M}_{+\infty}^{n+k}$*. Let* $E|\xi|^p < \infty$ *and* $|E|\eta|^q| < \infty$*,* $1/p + 1/q < 1$*. Then*

$$|E\xi\eta - E\xi E\eta| \leq 12(\alpha(k))^{1-1/p-1/q}(E|\xi|^p)^{1/p}(E|\eta|^q)^{1/q}.$$

We need the weak law of large numbers for $\eta_{d,n}$.

Lemma 17. *If (3.10)–(3.18) and (3.21)–(3.25) hold, then we have*

$$\frac{\eta_{d,n}}{H^{-1}(d/n)} \xrightarrow{P} 1$$

Proof. Using theorem 34 and Gorodetskii ([43], 1977) we get that X_k is a strongly mixing stationary sequence with mixing rate $\alpha(k) \leq C_1 \exp(-\lambda k)$ for some $C_1 > 0$ and $\lambda > 0$. Fix $1/2 < t < 2$ and let $T_k = I\{|X_k| \geq tH^{-1}(d/n)\}, 1 \leq k \leq n$. Clearly, $ET_k = P\{|X_k| \geq tH^{-1}(d/n)\} = H(tH^{-1}(d/n))$ and due to the the regular variation of H, $ET_k/(d/n) \to t^{-\alpha}$, as $n \to \infty$. On the other hand, by the correlation inequality of theorem 35 we get for any $p > 2$ that

$$\begin{aligned}
|ET_0T_k - ET_0ET_k| &\leq (\alpha(k))^{(p-1)/p}(ET_0^p)^{1/p}(ET_k^p)^{1/p} \\
&\leq C_1 \exp(-\lambda k(p-1)/p)(ET_0^p)^{2/p} \\
&= C_1 \exp(-\lambda k(p-1)/p)(ET_0)^{2/p} \\
&\leq C_2 \exp(-\lambda k(p-1)/p)(d/n)^{2/p}.
\end{aligned}$$

Hence setting $\bar{T}_k = T_k - ET_k$ we conclude that

$$\begin{aligned} E\left(\sum_{k=1}^{n} \bar{T}_k\right)^2 &= nE\bar{T}_0^2 + 2\sum_{k=1}^{n-1}(n-k)E\bar{T}_0\bar{T}_k \\ &\le n\left(E\bar{T}_k^2 + 2\sum_{k=1}^{n-1}|E\bar{T}_0\bar{T}_k|\right) \\ &\le n\left(ET_0^2 + C_3\sum_{k=1}^{n-1}\exp(-\lambda k(p-1)/p)(d/n)^{2/p}\right) \\ &\le n\left(ET_0 + C_5(d/n)^{2/p}\right) \\ &\le n(d/n)^{2/p}. \end{aligned}$$

Thus by Markov's inequality we have that

$$P\left\{\sum_{k=1}^{n} \bar{T}_k \ge d^{2/p}\right\} \le C_6 n^{(p-2)/p}/d^{2/p} \to 0,$$

provided that $d/n^{(2-p)/p} \to 0$. Since $d \ge n^{\delta}$, choosing p near 2, it follows that

$$\sum_{k=1}^{n} T_k = t^{-\alpha}d(1+o_P(1)) + o_P(d^{2/p}) = t^{-\alpha}d(1+o_P(1)).$$

In other words,

$$\frac{1}{d}\#\{k \le n : \ |X_k| \ge tH^{-1}(d/n)\} \xrightarrow{P} t^{-\alpha}, \quad \text{as} \quad n \to \infty.$$

This shows that

$$\lim_{n\to\infty} P\{\eta_{n,d} \ge tH^{-1}(d/n)\} = 1 \quad \text{for} \quad t < 1$$

and

$$\lim_{n\to\infty} P\{\eta_{n,d} \ge tH^{-1}(d/n)\} = 0 \quad \text{for} \quad t > 1,$$

completing the proof of Lemma 17. □

Proof of Theorem 30. We note that $\Gamma(t,x)$ is a continuous process. Hence combining Lemmas 16 and 17 we conclude

$$L_n(\eta_{d,n}/H^{-1}(d/n), x) \quad \xrightarrow{\mathcal{D}[0,1]} \quad \Gamma(1,x).$$

It is easy to see that

$$\{\Gamma(1,x), 0 \le x \le 1\} \stackrel{\mathcal{D}}{=} \left\{ \left(\frac{\alpha}{2-\alpha} \frac{1+\rho}{1-\rho} \right)^{1/2} W(x), 0 \le x \le 1 \right\},$$

where $W(x)$ is a Wiener process, which completes the proof. □

Proof of Theorem 31. Since

$$\frac{1}{A_n} T_{n,d}(x) = L_n(\eta_{d,n}/H^{-1}(d/n), x) - \frac{\lfloor nx \rfloor}{n} L_n(\eta_{d,n}/H^{-1}(d/n), 1),$$

Theorem 30 yields

$$\frac{1}{A_n} T_{n,d}(x) \quad \stackrel{\mathcal{D}[0,1]}{\longrightarrow} \quad \left(\frac{\alpha}{2-\alpha} \frac{1+\rho}{1-\rho} \right)^{1/2} (W(x) - xW(1)).$$

By definition, $B(x) = W(x) - xW(1), 0 \le x \le 1$ is a Brownian bridge, so the proof of Theorem 31 is complete. □

3.5 Estimation of the long run variance

The weak convergence in Theorem 31 can be used to construct tests to detect possible changes in the location parameter in model (3.9). However, the normalizing sequence depends heavily on unknown parameters and they should be replaced with consistent estimators. We discuss this approach in this section. We show in section 3.6 that ratio statistics can also be used so we can avoid the estimation of the long run variances.

The limit result in Theorem 31 is the same as one gets for the CUSUM process in case of weakly dependent stationary variables (cf. Aue and Horváth ([5], 2013)). Hence we interpret the normalizing sequence as the long run variance of the sum of the trimmed variables. Based on this interpretation we suggest Bartlett type estimators as the normalization.

The Bartlett estimator computed from the trimmed variables $X_i^* = X_i I\{|X_i| \leq \eta_{n,d}\}$ is given by

$$\hat{s}_n^2 = \hat{\gamma}_0 + 2\sum_{j=1}^{n-1} \omega_j \left(\frac{j}{h(n)}\right) \hat{\gamma}_j,$$

where

$$\hat{\gamma}_j = \frac{1}{n}\sum_{i=1}^{n-j}(X_i^* - \bar{X}_n^*)(X_{i+j}^* - \bar{X}_n^*), \quad \bar{X}_n^* = \frac{1}{n}\sum_{i=1}^{n} X_i^*,$$

$\omega(\cdot)$ is the kernel and $h(\cdot)$ is the length of the window. We assume that $\omega(\cdot)$ and $h(\cdot)$ satisfy the following standard assumptions:

$$\omega(0) = 1, \tag{3.59}$$

$$\omega(t) = 0 \quad \text{if} \quad t > a \quad \text{with some} \quad a > 0, \tag{3.60}$$

$$\omega(\cdot) \quad \text{is a Lipschitz function,} \tag{3.61}$$

$$\hat{\omega}(\cdot), \text{ the Fourier transform of } \omega(\cdot), \text{ is also Lipschitz and integrable} \tag{3.62}$$

and

$$h(n) \to \infty \quad \text{and} \quad h(n)/n \to \infty \quad \text{as} \quad n \to \infty. \tag{3.63}$$

For functions satisfying (3.59)–(3.62) we refer to Taniguchi and Kakizawa ([93], 2000). Following the methods in Liu and Wu ([70], 2010) and Horváth and Reeder ([49], 2012), the following weak law of large numbers can be established under H_0:

$$\frac{n\hat{s}_n^2}{A_n^2(1+\rho)\alpha/((1-\rho)(2-\alpha))} \xrightarrow{P} 1, \quad \text{as} \quad n \to \infty. \tag{3.64}$$

The next result is an immediate consequence of Theorem 31 and (3.64).

Corollary 1. *If H_0, (3.15)-(3.18), (3.21)-(3.25) and* (3.64) *hold, then we have that*

$$\frac{T_n(x)}{n^{1/2}\hat{s}_n} \xrightarrow{\mathcal{D}[0,1]} B(x),$$

where $B(x)$ is a Brownian bridge.

It follows immediately that under the no change null hypothesis

$$\hat{\mathcal{Q}}_n = \sup_{0\le x\le 1} \frac{|T_n(x)|}{n^{1/2}\hat{s}_n} \xrightarrow{\mathcal{D}} \sup_{0\le x\le 1} |B(x)|.$$

Simulations show that $\hat{s}_n$ performs well under H_0 but it overestimates the norming sequence under the alternative. Hence $\hat{\mathcal{Q}}_n$ has little power. The estimation of the long–run variance when a change occurs has been addressed in the literature. We follow the approach of Antoch, Huškova and Praškova ([2], 1997), who provided estimators for the long run variance which are asymptotically consistent under the H_0 as well as under the one change alternative. Let x_0 denote the smallest value in $[0,1]$ where $|T_n(x)|$ reaches its maximum and let $\tilde{k} = \lfloor x_0 n \rfloor$. The modified Bartlett estimator is defined as

$$\tilde{s}_n^2 = \hat{\gamma}_0' + 2\sum_{j=1}^{n-1} \omega\left(\frac{j}{h(n)}\right)\tilde{\gamma}_j,$$

where

$$\tilde{\gamma}_j = \frac{1}{n-j}\sum_{\ell=1}^{n-j} \iota_\ell \iota_{\ell+j}, \qquad \iota_\ell = X_\ell^* - \frac{1}{\hat{k}}\sum_{\ell=1}^{\hat{k}} X_\ell^*, \quad \ell = 1,...,\hat{k},$$

$$\iota_\ell = X_\ell^* - \frac{1}{n-\hat{k}}\sum_{\ell=\hat{k}+1}^{n} X_\ell^*, \quad \ell = \hat{k}+1,...,n.$$

Combining the proofs in Antoch, Huškova and Praškova ([2], 1997) with Liu and Wu ([70], 2010) and Horváth and Reeder ([49], 2012) one can verify that

$$\frac{n\tilde{s}_n^2}{A_n^2(1+\rho)\alpha/((1-\rho)(2-\alpha))} \xrightarrow{P} 1, \quad \text{as} \quad n\to\infty \tag{3.65}$$

under H_0 as well as under the one change alternative H_A. Due to (3.65) we immediately have the following result:

Corollary 2. *If H_0, (3.15)–(3.18), (3.21)–(3.25) and* (3.65) *hold, then we have that*

$$\frac{T_n(x)}{n^{1/2}\tilde{s}_n} \xrightarrow{\mathcal{D}[0,1]} B(x),$$

where $B(x)$ is a Brownian bridge.

We suggest testing procedures based on

$$\tilde{\mathcal{Q}}_n = \frac{1}{n^{1/2}\tilde{s}_n} \sup_{0\le x\le 1} |T_n(x)|.$$

It follows immediately from Corollary 2 that under H_0

$$\tilde{\mathcal{Q}}_n \xrightarrow{\mathcal{D}} \sup_{0\le x\le 1} |B(x)|. \tag{3.66}$$

First we study experimentally the rate of convergence in Theorem 31. In this section we assume that the innovations ε_i in (3.15)–(3.19) have the common distribution function

$$F(t) = \begin{cases} q(1-t)^{-3/2}, & \text{if} \quad -\infty < t \le 0, \\ 1 - p(1-t)^{-3/2}, & \text{if} \quad 0 < t < \infty, \end{cases}$$

where $p \ge 0$, $q \ge 0$ and $p + q = 1$. We present the results for the case of $\rho = p = q = 1/2$ based on 10^5 repetitions. We simulated the elements of an autoregressive sample $(e_1, \ldots, e_n)$ from the recursion (3.15) starting with some initial value and with a burn in period of 500, i.e. the first 500 generated variables were discarded and the next n give the sample $(e_1, \ldots, e_n)$. Thus $(e_1, \ldots, e_n)$ are from the stationary solution of (3.15). We trimmed the sample using $d(n) = \lfloor n^{0.45} \rfloor$ and computed

$$\mathcal{Q}_n = \left(\frac{2-\alpha}{\alpha}\right)^{1/2} \left(\frac{1-\rho}{1+\rho}\right)^{1/2} \frac{1}{A_n} \sup_{0\le x\le 1} |T_n(x)|.$$

Under H_0 we have

$$\mathcal{Q}_n \xrightarrow{\mathcal{D}} \sup_{0\le x\le 1} |B(x)|.$$

The critical values in Table 3.1 provide information on the rate of convergence in Theorem 31.

Figures 3.2 and 3.3 show the empirical power of the test for H_0 against H_A based on the statistic $\mathcal{Q}_n$ for a change at time $k^* = n/4$ and $n/2$ and when the

n	400	600	800	1000	∞
	1.29	1.32	1.33	1.34	1.36

Table 3.1: Simulated 95% percentiles of the distribution of $\mathcal{Q}_n$ under H_0

location changes from 0 to $c \in \{-3, -2.9, ..., 2.9, 3\}$ and the level of significance is 0.05. We used the asymptotic critical value 1.36. Comparing Figures 3.2 and 3.3 we see that we have higher power when the change occurs in the middle of the data at $k^* = n/2$. We provided these results to illustrate the behaviour of functionals of T_n without introducing further noise due to the estimation of the norming sequence.

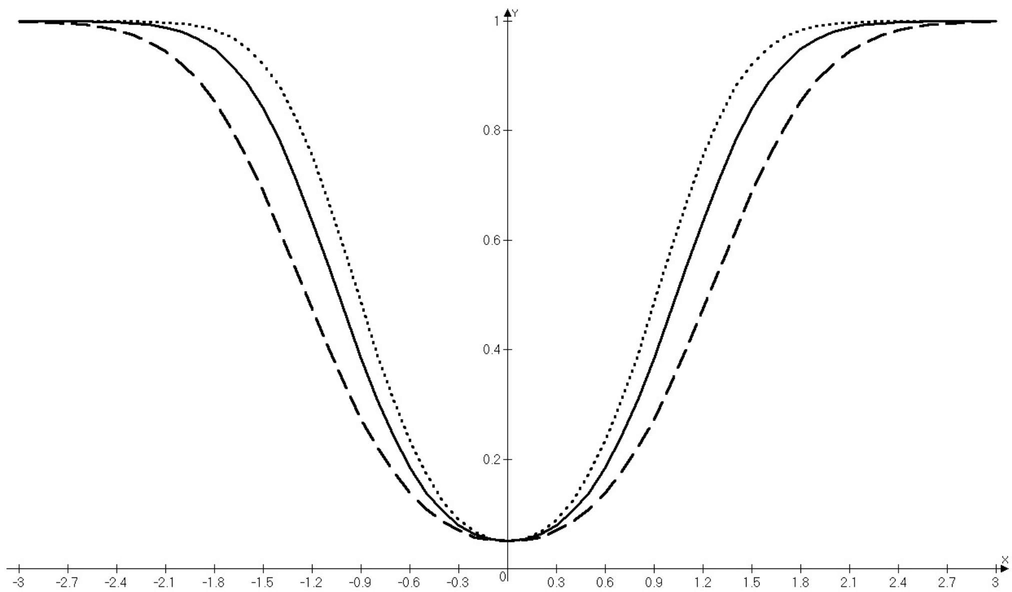

Figure 3.2: Empirical power for $\mathcal{Q}_n$ with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/2$

Next we study the applicability of (3.66) in case of small and moderate sample

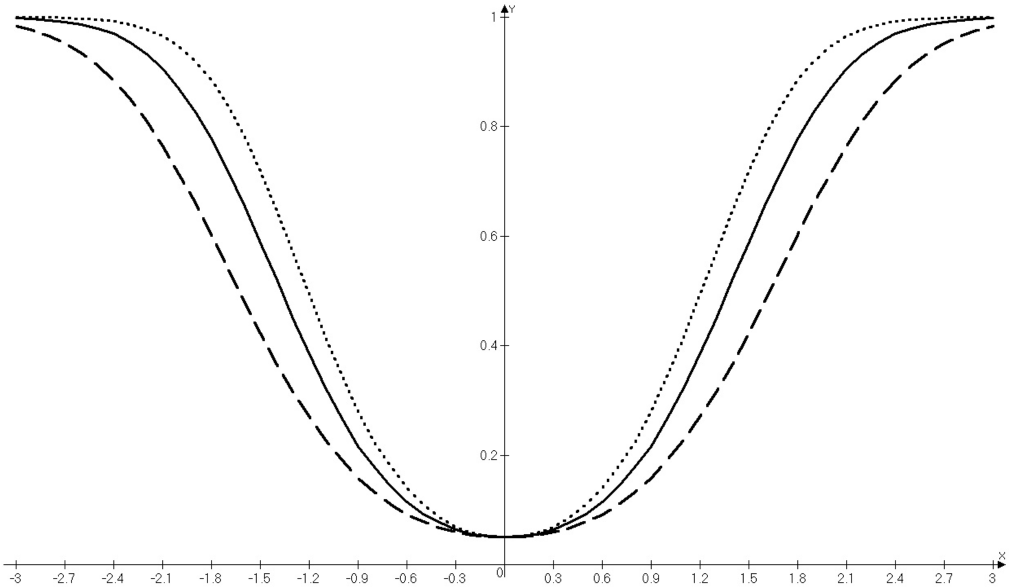

Figure 3.3: Empirical power for $\mathcal{Q}_n$ with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/4$

sizes. We used $h(n) = n^{1/2}$ as the window and the flat top kernel

$$\omega(t) = \begin{cases} 1 & 0 \leq t \leq .1 \\ 1.1 - |t| & .1 \leq t \leq 1.1 \\ 0 & t \geq 1.1 \end{cases}$$

Figures 3.4 and 3.5 show the empirical power of the test for H_0 against H_A based

n	400	600	800	1000	∞
	1.57	1.52	1.50	1.49	1.36

Table 3.2: Simulated 95% percentiles of the distribution of $\tilde{\mathcal{Q}}_n$ under H_0

on the statistic $\tilde{\mathcal{Q}}_n$ for a change at time $k^* = n/4$ and $n/2$ and when the location changes from 0 to $c \in \{-3, -2.9, ..., 2.9, 3\}$ and the level of significance is 0.05. We used the asymptotic critical value 1.36. Comparing Figures 3.4 and 3.5 we see that we have again higher power when the change occurs in the middle of the data at

$k_1 = n/2$.

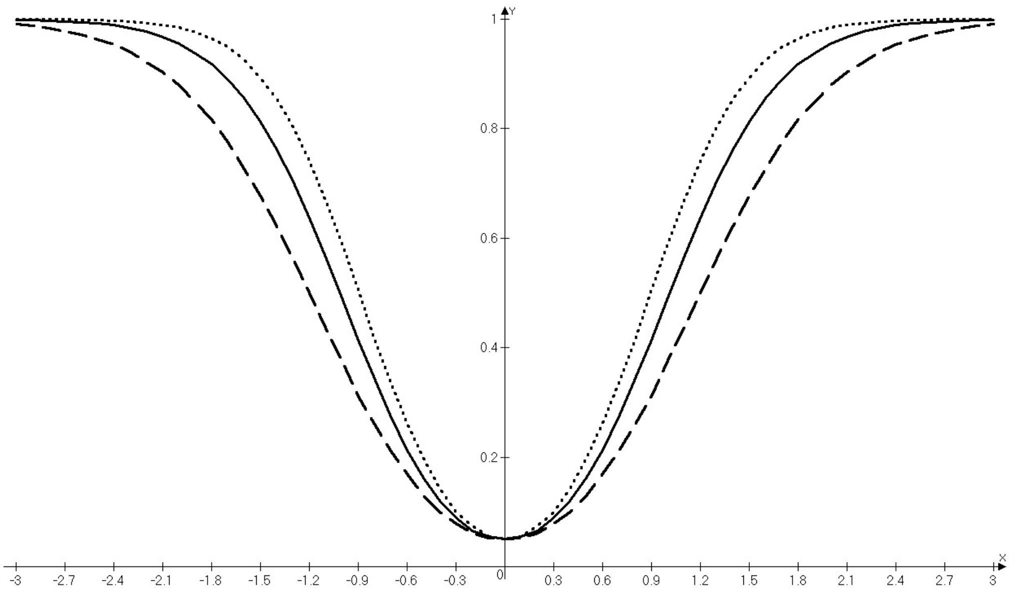

Figure 3.4: Empirical power for $\tilde{\mathcal{Q}}_n$ with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/2$

Figure 3.6 shows how the power of the test behaves depending on the value of $d = n^\epsilon$, $\epsilon \in \{0.3, 0.35, 0.42, 0.45, 0.5\}$ for $n = 400$. The bigger the d is, the better is the power curve.

3.6 Ratio statistics

The statistics $\hat{Q}_n$ as well as $\tilde{Q}_n$ are very sensitive to the behaviour of $\hat{s}_n$ and $\tilde{s}_n$. As we pointed out, $\hat{s}_n$ is the right norming only under H_0. The sequence $\tilde{Q}_n$ works under H_0 and under the one change alternative, but it could break down if multiple changes occur under the alternative. Even if the Bartlett type estimator is the asymptotically correct norming factor, the rate of convergence can be slow. Also,

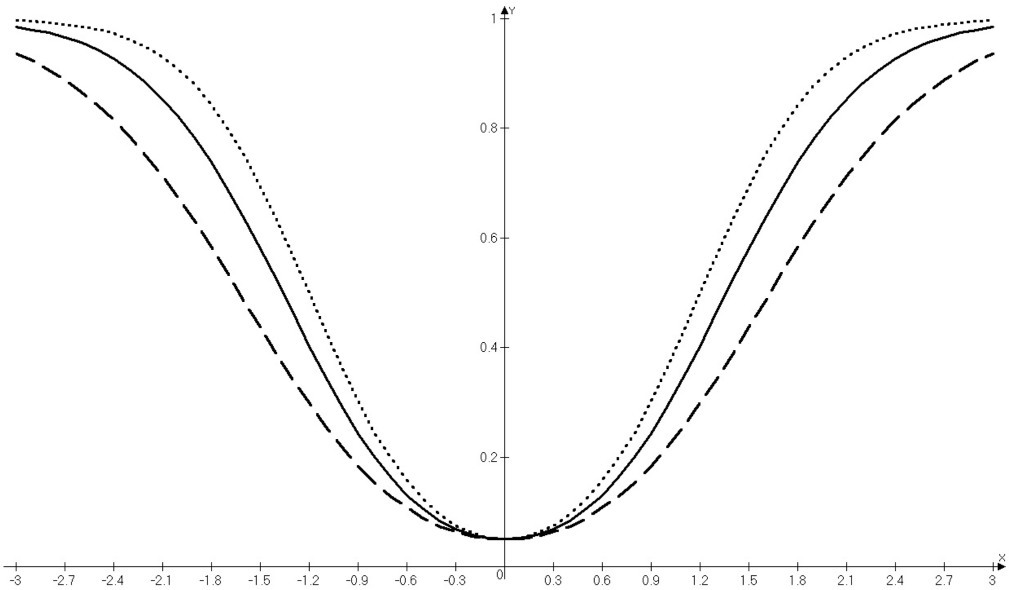

Figure 3.5: Empirical power for $\tilde{\mathcal{Q}}_n$ with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/4$

these estimators are very sensitive to the choice of the window $h = h(n)$. Following the work of Kim ([61], 2000) (cf. also Kim, Belair-Franch and Amador ([62], 2002)) and Leybourne and Taylor ([69], 2006), Horváth, L., Horváth, Zs. and Huškova ([48], 2008) proposed ratio type statistics of functionals of CUSUM processes. We adapt their approach to the trimmed CUSUM process. Let $0 < \delta < 1$ and define

$$Z_n = \max_{n\delta \le k \le n - n\delta} \frac{Z_{n,1}(k)}{Z_{n,2}(k)},$$

where

$$Z_{n,1}(k) = \max_{1 \le i \le k} \left| \sum_{j=1}^{i} (X_j I\{|X_j| \le \eta_{n,d}\} - (1/k) \sum_{j=1}^{k} (X_j I\{|X_j| \le \eta_{n,d})\}) \right|$$

and

$$Z_{n,2}(k) = \max_{k < i \le n} \left| \sum_{j=i}^{n} (X_j I\{|X_j| \le \eta_{n,d}\} - (1/(n-k)) \sum_{j=k+1}^{n} (X_j I\{|X_j| \le \eta_{n,d})\}) \right|.$$

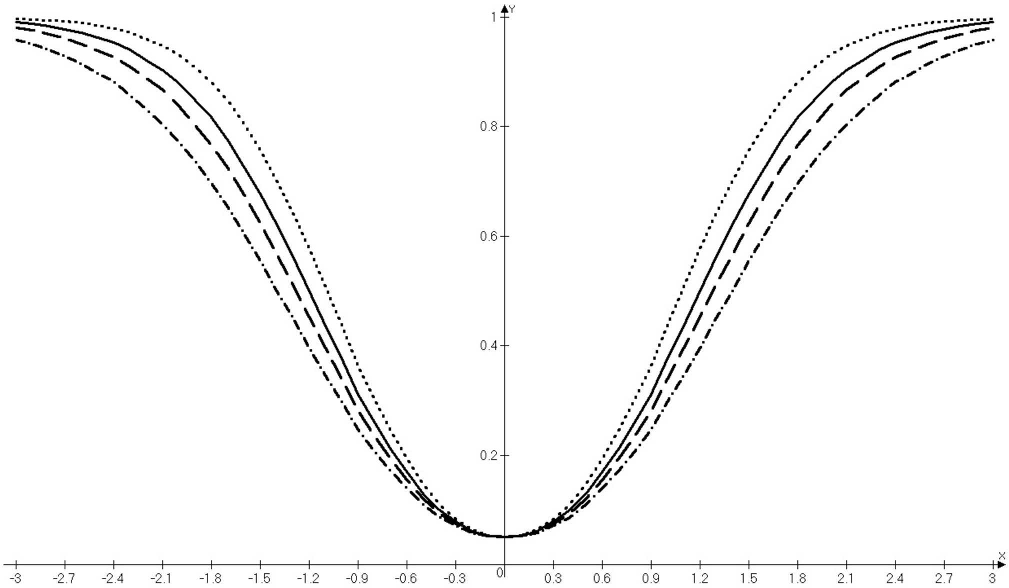

Figure 3.6: Empirical power curves for $\tilde{\mathcal{Q}}_n$ with significance level 0.05 for $d = n^{\epsilon}$, $\epsilon = 0.35$ (dash-dotted), $\epsilon = 0.42$ (dashed), $\epsilon = 0.45$ (solid), $\epsilon = 0.5$ (dotted) with $n = 400$, $k_1 = n/2$

Roughly speaking, we split the data into two subsets at k, compute the maximum of the CUSUM in both subsamples and compare these maxima. To state the limit distribution of Z_n under the null hypothesis, we need to introduce

$$z_1(t) = \sup_{0 \le s \le t} |W(s) - (s/t)W(t)|$$

and

$$z_2(t) = \sup_{t \le s \le 1} |W^*(s) - ((1-s)/(1-t))W^*(t)|,$$

where $W^*(t) = W(1) - W(t)$. The following result is an immediate consequence of Theorem 31.

Theorem 36. *If H_0, (3.15)–(3.18) and (3.21)–(3.25) hold, then we have that*

$$Z_n \quad \xrightarrow{\mathcal{D}} \quad \sup_{\delta \le t \le 1-\delta} \frac{z_1(t)}{z_2(t)}. \tag{3.67}$$

We reject the no change null hypothesis if Z_n is large. Using Monte Carlo simulations, it is easy to obtain the distribution function of the limit in (3.67). Selected critical values can be found in Horváth L., Horváth, Zs. and Huškova ([48], 2008), where some probabilistic properties of the limit are also discussed.

n	400	600	800	1000	5000
	5.90	5.67	5.49	5.43	5.03

Table 3.3: Simulated 95% percentiles of the distribution of Z_n under H_0

Below we study the finite sample behaviour of Z_n. Table 3.3 contains simulated significance levels when $\delta = .2$, $n = 400, 600, 800, 1,000$ and $n = 5,000$. (Since the distribution function of the limit in (3.67) is unknown, we used $n = 5,000$ for the limit distribution.)

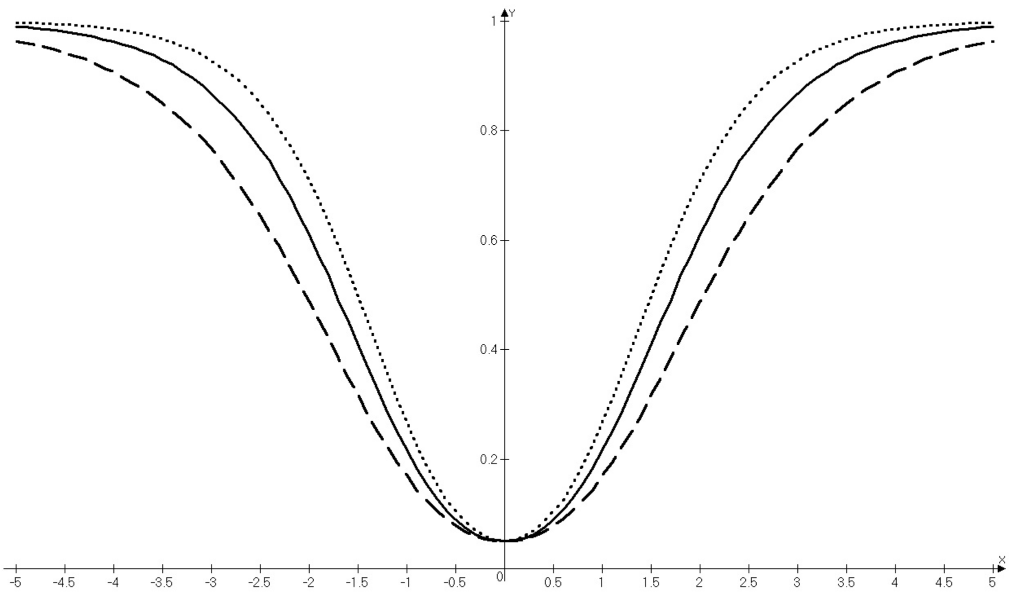

Figure 3.7: Empirical power curves for Z_n with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/2$

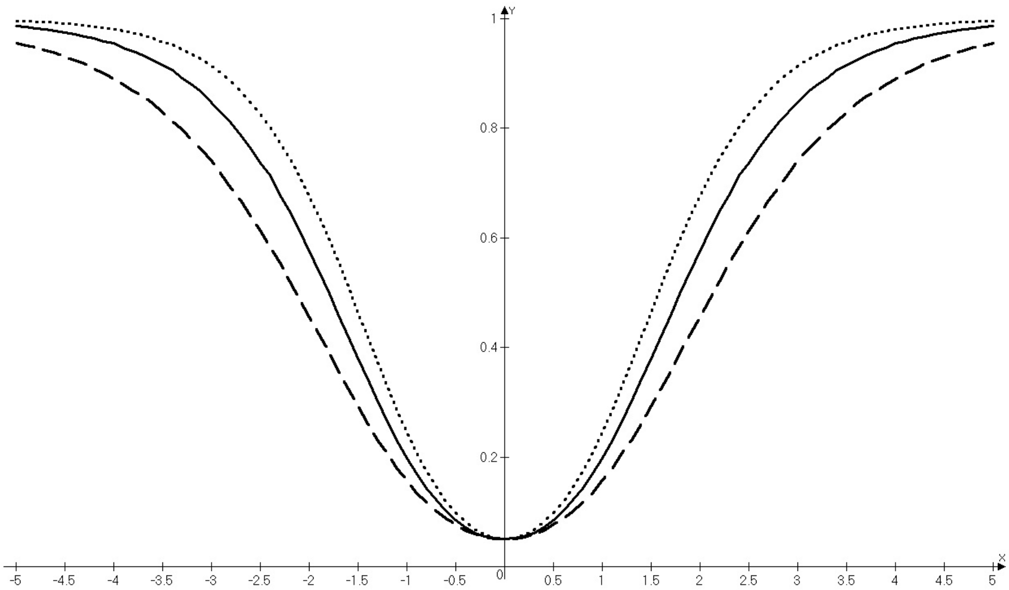

Figure 3.8: Empirical power curves for Z_n with significance level 0.05, $n = 400$ (dashed), $n = 600$ (solid) and $n = 800$ (dotted) with $k_1 = n/4$

Figures 3.7 and 3.8 contain the empirical power curves of the test for H_0 against H_A based on the statistic Z_n for a change at time $k^* = n/4$ and $n/2$ and when the location changes from 0 to $c \in \{-5, -4.9, ..., 4.9, 5\}$ and the level of significance is 0.05. We used critical values from Table 3.3. Figure 3.9 shows how the power of the test behaves depending on the value of $d = n^\epsilon$, $\epsilon \in \{0.3, 0.35, 0.42, 0.45, 0.5\}$ for $n = 400$. The bigger the d is, the better is the power curve.

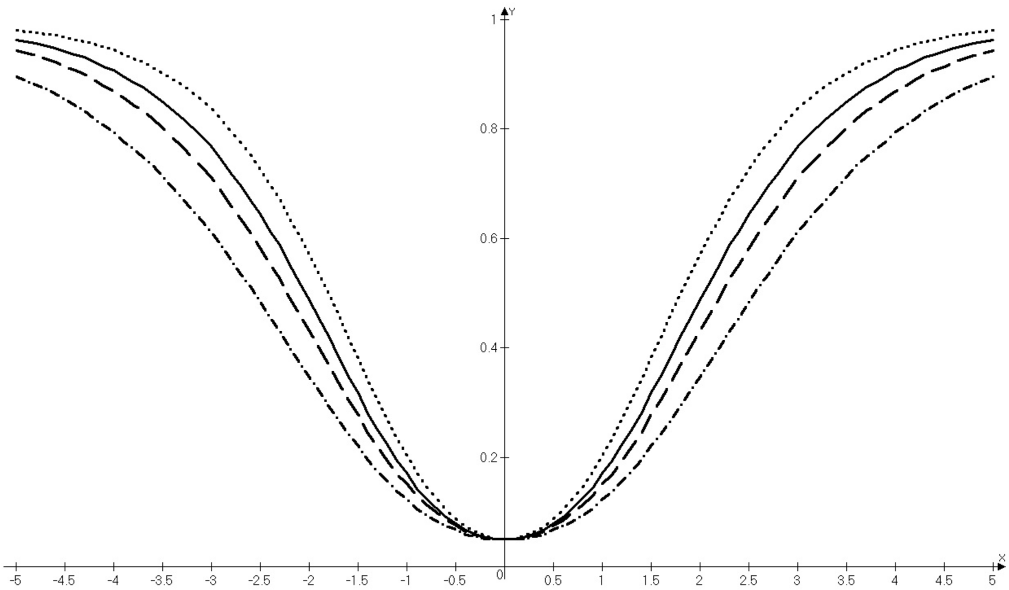

Figure 3.9: Empirical power curves for Z_n with significance level 0.05 for $d = n^\epsilon$, $\epsilon = 0.35$ (dash-dotted), $\epsilon = 0.42$ (dashed), $\epsilon = 0.45$ (solid), $\epsilon = 0.5$ (dotted) with $n = 400$, $k_1 = n/2$

Bibliography

[1] Andrews, B., Calder, M. and Davis, R.A.: Maximum likelihood estimation for α–stable autoregressive processes. *Annals of Statistics* **37**(2009), 1946–1982.

[2] Antoch, J., Hušková, M. and Prašková, Z.: Effect of dependence on statistics for determination of change. *Journal of Statistical Planning Inference* **60**(1997), 291-310.

[3] Arov, D.Z. and Bobrov, A.A.: The extreme members of a sample and their role in the sum of independent variables. Theor. Probability Appl. 5 (1960), 377–396.

[4] Aue, A., Berkes, I. and Horváth, L.: Selection from a stable box. *Bernoulli* **14**(2008), 125–139.

[5] Aue, A. and Horváth, L.: A limit theorem for mildly explosive autoregression with stable errors. *Econometric Theory* **23**(2007), 201–220.

[6] Aue, A. and Horváth, L.: Stuctural breaks in time series. *Journal of Time Series Analysis* **34**(2013), 1–16.

[7] Avram, F.S. and Taqqu, M.S.: Weak convergence of moving averages with infinite variance. In: *Dependence in Probability and Statistics:A Survey of Recent Results* (E. Eberlain and M.S. Taqqu, eds.), 399-425, Birkhäuser, Boston, 1986.

[8] Avram, F.S. and Taqqu, M.S.: Weak convergence of moving averages in the α–stable domain of attraction. *Annals of Probability* **20**(1992), 483–503.

[9] Basrak, B., Krizmanič, D. and Segers, J.: A functional limit theorem for partial sums of dependent random variables with infinite variance. *Annals of Probability* To appear, 2012.

[10] Bazarova, A., Berkes, I. and Horváth, L.: Change-point detection with stable AR(1) errors: , 2013, In press.

[11] Bazarova, A., Berkes, I. and Horváth, L.: Trimmed stable AR(1) processes. *Stochastic Processes and Their Applications.*, 2013, In press.

[12] Bazarova, A., Berkes, I. and Horváth, L.: On the central limit theorem for modulus trimmed sums: *Statistics and Probability Letters*, 2013, In press.

[13] Bazarova, A., Berkes, I. and Horváth, L.: On the extremal theory of continued fractions: *Submitted for publication.*

[14] Berkes, I. and Horváth, L., 2012. The central limit theorem for sums of trimmed variables with heavy tails. Stochastic Process. Appl. **122**(2012), 449–465.

[15] Berkes, I., Horváth, L., Ling, S. and Schauer, J.: Testing for structural change of AR model to threshold AR model. *J. Time Series Analysis* **32**(2011), 547–565.

[16] Berkes, I., Horváth, L. and Schauer, J., 2011. Asymptotics of trimmed CUSUM statistics. Bernoulli 17, 1344–1367.

[17] Bickel, P. J. and Wichura, M. J.: Convergence criteria for multiparameter stochastic processes and some applications. Ann. Math. Statist. 42 (1971), 1656–1670.

[18] Billingsley, P.: Ergodic theory and information. Wiley, 1965.

[19] Billingsley, P.: Convergence of Probability Measures. Wiley, New York 1968.

[20] Bingham, N.H., Goldie, C. M., Teugels, J.L., 1987. Regular variation. Cambridge University Press.

[21] Bradley, R.: Introduction to strong mixing conditions, Vol I. Kendrick Press, 2007.

[22] Chan, N.H.: Inference for near–integrated time series with infinite variance. *Journal of the American Statistical Association* **85**(1990), 1069–1074.

[23] Chan, N.H. and Tran, L.T.: On the first–order autoregressive processes with infinite variance. *Econometric Theory* **5**(1989), 354–362.

[24] Cline, D.: *Estimations and Linear Prediction for Regression, Autoregression and ARMA with Infinite Variance Data.* PhD Thesis, Colorado State University, Fort Collins, 1983.

[25] Csörgő, M. and Horváth, L.: *Limit Theorems in Change-Point Analysis.* Wiley, 1998.

[26] Csörgő, S., Horváth, L. and Mason, D.: What portion of the sample makes a partial sum asymptotically stable or normal ? Z. Wahrschein. verw. Gebiete 72 (1986), 1–16.

[27] Csörgő, S., Csörgő, M., Horváth, L. and Mason, D., 1986. Normal and stable convergence of integral functions of the empirical distribution function. Ann. Probability 14, 86–118.

[28] Csörgő, S. and Simons, G.: A strong law of large numbers for trimmed sums, with applications to generalized St. Petersburg games. Statist. Prob. Letters 26 (1996), 65–73.

[29] Darling, D.: The influence of the maximum term in the addition of independent random variables. Trans. Amer. Math. Soc. 73 (1952), 95–107.

[30] Davis, R.A. and Mikosch, T.: The sample autocorrelations of heavy–tailed processes with applications to ARCH. *Annals of Statistics* **26**(1998), 2049–2080.

[31] Davis, R. and Resnick, S.: Limit theory for moving averages of random variables with regularly varying tail probabilities. *Annals of Probability* **13**(1985), 179–195.

[32] Davis, R. and Resnick, S.: Limit theory for the sample covariance and correlation functions of moving averages. *Annals of Statistics* **14**(1986), 533–558.

[33] Davydov, Y.A.: Convergence of distributions generated by stationary stochastic processes. *Theor. Probab. Appl.* **13**(1968), 691–696.

[34] de Acosta, A. and Giné, E.: Convergence of moments and related functionals in the general central limit theorem in Banach spaces. *Z. Wahrsch. Verw. Gebiete* **48**(1979), 213-231.

[35] Diamond, H. and Vaaler, J.: Estimates for partial sums of continued fraction partial quotients. Pacific J. Math. 122 (1986), 73–82.

[36] Doeblin, W.: Remarques sur la théorie métrique des fractions continues. Comp. Math. 7 (1940), 353–371.

[37] Fama, E.: The behavior of stock market price. *J. of Business* **38**(1965), 34–105

[38] Feller, W., 1971. An introduction to probability theory and its applications, Vol II, Second Edition. Wiley, New York.

[39] Galambos, J.: The distribution of the largest coefficient in continued fraction expansions. Quart. J. Math. Oxford Ser. 23 (1972), 147–151.

[40] Galambos, J.: An iterated logarithm type theorem for the largest coefficient in continued fractions. Acta Arith. **25** (1973/74), 359–364.

[41] Gikhman, I. and Skorohod, A. V.: *Introduction to the Theory of Random Processes.* Saunders, Philadelphia.

[42] Gordin, M.I. and Reznik, M.H.: The law of the iterated logarithm for the denominators of continued fractions. Vestnik Leningrad Univ. 25 (1970), 28–33. (In Russian)

[43] Gorodetskii, V.V.: On the strong mixing property for linear sequences. *Theor. Probability Appl.* **22**(1977), 411–413.

[44] Griffin, P. S. and Pruitt, W. E.: The central Math. Proc. Cambridge Philos. Soc., 102 (1987), 329–349.

[45] Griffin, P. and Pruitt, W.E.: Asymptotic normality and subsequential limits of trimmed sums. *Annals of Probability* **17**(1989), 1186–1219.

[46] Hall, P.: On the extreme terms of a sample from the domain of attraction of a stable law. *Journal of the London Mathematical Society* **18**(1978), 181–191.

[47] Heinrich, L.: Rates of convergence in stable limit theorems for sums of exponentially ψ- mixing random variables with an application to metric theory of continued fractions. Math. Nachr. 131 (1987), 149–165.

[48] Horváth, L., Horváth, Zs. and Hušková, M.: Ratio test for change point detection. In: *IMS Collection* **1**(2008), 293–304.

[49] Horváth, L. and Reeder, R.: Detecting changes in functional linear models. *Journal of Multivariate Analysis* **111**(2012), 310–334.

[50] Huber, P.J. : Théorie de l'Inférence Statistique Robuste. Les Presses de l'Université de Montréal (1969).

[51] Ibragimov, I.A.: A theorem from the metric theory of continued fractions. Vestnik Leningrad Univ. 1 (1960), 13–24. (In Russian)

[52] Ibragimov, I.A., Linnik, Yu.V.: *Independent and stationary sequences of random variables.* Wolters-Noordhoff Publishing, Groningen.

[53] Iosifescu, M.: A Poisson law for ψ-mixing sequences establishing the truth of a Doeblin's statement. Rev. Roumaine Math. Pures Appl. 22 (1977), 1441–1447.

[54] Iosifescu, M.: A survey of the metric theory of continued fractions, fifty years after Doeblins 1940 pa- per. In: Probability Theory and Mathematical Statistics, vol. I, pp. 550–572. Vilnius, Mokslas (1990)

[55] Kesten, H.: Convergence in distribution of lightly trimmed and untrimmed sums are equivalent. *Math. Proc. Cambridge Phil. Soc.* **113**(1993), 615–638.

[56] Kesten, H., Maller, R.A.: Ratios of trimmed sums and order statistics. *Annals of Probability* **20**(1992), 1805–1842.

[57] Kesten, H., Maller, R.A.: Infinite limits and infinite limit points of random walks and trimmed sums. *Annals of Probability* **22**(1994), 1473–1513.

[58] Kesten, H., Maller, R.A.: The effect of trimming on the strong law of large numbers. *Proc. Lon. Math. Soc.* **(3) 71**, no. 2 (1995), 441–480

[59] Khinchin, A. J.: Metrische Kettenbruchprobleme. Compositio Math. 1 (1935), 361–382.

[60] Kiefer, J., 1972. Iterated logarithm analogues for sample quantiles when $p_n \to 0$. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, pp. 227–244. University of California Press.

[61] Kim, J.-Y.: Detection of change in persistence of a linear time series. *Journal of Econometrics* **95**(2000), 97–116.

[62] Kim, J.-Y., Belaire–Franch, J, and Amador, R.: Corrigendum to "Detection of change in persistence of a linear time series". *Journal of Econometrics* **109**(2002), 389–392.

[63] Kusmin, R.: Sur un problème de Gauss. Atti Congr. Intern. Bologne 1928. 6, 83–89.

[64] Leadbetter, M. and Rootzén, H.: Extremal theory for stochastic processes. Ann. Probab. 16 (1988), 431–478.

[65] LePage, R., Woodrofe, M. and Zinn, J., 1981. Convergence to a stable distribution via order statistics. Ann. Prob. 9, 624–632.

[66] Lévy, P.: Sur les lois dont dépendent les quotients complets et incomplets d'une fraction continue. Bull. Sci. Math. France 57 (1929), 178–194.

[67] Lévy, P.: Propriétés asymptotiques des sommes de variables aléatoires indṕendentes en enchaines. *Journal de Mathématiques*, **14** (1935) 347–402.

[68] Lévy, P.: Fractions continues aléatoires. Rend. Circ. Mat. Palermo 1 (1952), 170–208.

[69] Leybourne, S. and Taylor, A.: Persistence change tests and shifting stable autoregressions. *Economics Letters* **91**(2006), 44–49.

[70] Liu, W. and Wu, W.B.: Asymptotics of spectral density estimates. *Econometric Theory* **26**(2010), 1218–1245.

[71] Maller, R.A.: Asymptotic normality of lightly trimmed meansa converse. *Math. Proc. Cambridge Phil. Soc.* **92**(1982), 535–545.

[72] Mandelbrot, B.B.: The variation of certain speculative price. *J. of Business* **36**(1963), 394–419.

[73] Mandelbrot, B.B.: The variation of other speculative prices. *j. of Business* **40**(1967), 393–413.

[74] Mikosch, T., Resnick, S. and Samorodnitsky, G.: The maximum of the periodogram for a heavy–tailed sequence. *Annals of Probability* **28**(2000), 885–908.

[75] Mori, T.: On the limit distributions of lightly trimmed sums. *Math. Proc. Cambridge Phil. Soc.* **96**(1984), 507–516.

[76] Philipp, W.: The remainder in the central limit theorem for mixing stochastic processes. Ann. Math. Statist. 40 (1969), 601–609.

[77] Philipp, W.: Some metrical theorems in number theory II. Duke Math. J. 37 (1970), 447–458.

[78] Philipp, W.: A conjecture of Erdős on continued fractions. Acta Arith. 28 (1975/76), 379–386.

[79] Philipp, W.: Limit theorems for partial quotients of continued fractions. Monatsh. Math. 105 (1988), 195–206.

[80] Philipp, W. and Stackelberg. O.: Zwei Grenzwertssätze für Kettenbrüche. Math. Ann. 181 (1969), 152–156.

[81] Phillip, W. and Stout, W. : Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* (**161**) 1–140.

[82] Phillips, P.C.B. and Solo, V.: Asymptotics for linear processes. *Annals of Statistics* **20**(1992), 971–1001.

[83] Petrov, V.V.: *Limit Theormes of Probability Theory.* Clarendon Press, Oxford.

[84] Resnick, S. and Greenwood, P.: A bivariate stable characterization and domains of attraction. *J. Multivariate Analysis* **9**(1979), 206–221.

[85] Samur, J.: Convergence of sums of mixing triangular arrays of random vectors with stationary rows. *The Annals of Probability* **12**(1984), 390–426.

[86] Samur, J.: On some limit theorems for continued fractions. Trans. Amer. Math. Soc. 316 (1989), 53–79.

[87] Samur, J.: Some remarks on a probability limit theorem for continued fractions. Trans. Amer. Math. Soc. 348 (1996), 1411–1428.

[88] Skorokhod, A.V.: Limit theorems for stochastic processes. *Theory of Probability and Its Applications* **1**(1956), 261–290

[89] Stackelberg, O.P.: On the law of the iterated logarithm for continued fractions. *Duke Math. J.* **33**(1966), 801–819.

[90] Stigler, S.M.: The asymptotic distribution of the trimmed mean. *The Annals of Statistics* **1**(1973), 472–477.

[91] Stigler, S.M.: Linear functions of order statistics with smooth weight functions. University of Wisconsin Mathematical Research Center T. R. # 1211(1972)

[92] Szewczak, Z.S.: On limit theorems for continued fractions. *J. Theor. Probab.* **22**(2009), 239–255.

[93] Taniguchi, M. and Kakizawa, Y.: *Asymptotic Theory of Statistical Inference for Time Series.* Springer, 2000.

[94] Teugels, J.L.: Limit theorems on order statistics. *Annals of Probability* **9**(1981), 868–880.

[95] Tyran–Kamińska, M.: Functional limit theorems for linear processes in the domain of attraction of stable laws. *Statistics and Probability Letters* **80**(2010), 975–981.

[96] Vardi, I.: The St. Petersburg game and continued fractions. C.R. Acad. Sci. Paris Ser. I. Math. 324 (1997), 913–918.

[97] Withers, C.S.: Conditions for linear processes to be strong–mixing. *Z. Wahrsch. verw. Gebiete* **57**(1981), 477-480.

[98] Zhang, R–M. and Chan, N.H.: Maximum likelihood estimation for nearly non–stationary stable autoregressive processes. *J. Time Series Analysis* (2011).