Estimation of cluster functionals for regularly varying time series: sliding blocks estimators

Youssouph Cissokho and Rafal Kulik

Department of Mathematics and Statistics, University of Ottawa,
150 Louis Pasteur Private, Ottawa ON K1N 6N5
e-mail: yciss079@uottawa.ca*;* rkulik@uottawa.ca

Abstract: Cluster indices describe extremal behaviour of stationary time series. We consider their sliding blocks estimators. Using a modern theory of multivariate, regularly varying time series, we obtain central limit theorems under conditions that can be easily verified for a large class of models. In particular, we show that in the Peaks-Over-Threshold framework, sliding and disjoint blocks estimators have the same limiting variance.

Keywords and phrases: Regularly varying time series, extremes, cluster index, extremal index.

Contents

1. Introduction

Consider a stationary, regularly varying $\mathbb{R}^d$-valued time series $\boldsymbol{X} = \{\boldsymbol{X}_j, j \in \mathbb{Z}\}$. We are interested in a thorough understanding of its extremal behaviour. A classical approach to this problem is to calculate the *extremal index*. If $|\cdot|$ is an arbitrary norm on $\mathbb{R}^d$, then the extremal index θ (if exists) of $\{|\boldsymbol{X}_j|, j \in \mathbb{Z}\}$ is defined as a parameter in the limiting distribution of the maxima. With Q being the quantile function of $|\boldsymbol{X}_0|$ and $a_n = Q(1 - 1/n)$ we have

$$\lim_{n \to \infty} \mathbb{P}(a_n^{-1} \max_{j=1,\ldots,n} \{|\boldsymbol{X}_1|, \ldots, |\boldsymbol{X}_n|\} \le x) = \exp(-\theta x^{-\alpha}) , \quad x > 0 .$$

The parameter $\theta \in (0, 1]$ indicates the amount of clustering, with $\theta = 1$ (the case of extremal independence) meaning no-clustering of large values.

The extremal index is just one parameter that describes clustering of extremes. A related object is the cluster size distribution. It is the probability mass function of the number of exceedances over a large threshold within a given cluster. Both the extremal index and the cluster size distribution stem

from an application of a suitable functional to a cluster. This leads to a more general concept of cluster indices.

Informally speaking, a cluster is a triangular array $(\boldsymbol{X}_1/u_n, \ldots, \boldsymbol{X}_{r_n}/u_n)$ with $r_n, u_n \to \infty$ that converges in distribution in a certain sense. Cluster indices are obtained by applying the appropriate functional H to the cluster. The functionals are defined on $(\mathbb{R}^d)^{\mathbb{Z}}$, the space of $\mathbb{R}^d$-valued sequences, and are such that their values do not depend on coordinates that are equal to zero. More precisely, for $\boldsymbol{X} = \{\boldsymbol{X}_j, j \in \mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}}$ and $i \leq j \in \mathbb{Z}$, we denote $\boldsymbol{X}_{i,j} = (\boldsymbol{X}_i, \ldots, \boldsymbol{X}_j) \in (\mathbb{R}^d)^{(j-i+1)}$. Then, we identify $H(\boldsymbol{X}_{i,j})$ with $H((\boldsymbol{0}, \boldsymbol{X}_{i,j}, \boldsymbol{0}))$, where $\boldsymbol{0} \in (\mathbb{R}^d)^{\mathbb{Z}}$ is the zero sequence. Such functionals H will be called *cluster functionals*.

Let $|\cdot|$ be an arbitrary norm on $\mathbb{R}^d$ and $\{u_n\}$, $\{r_n\}$ be such that

$$\lim_{n \to \infty} u_n = \lim_{n \to \infty} r_n = \lim_{n \to \infty} n\mathbb{P}(|\boldsymbol{X}_0| > u_n) = \infty \ ,$$

$$\lim_{n \to \infty} r_n/n = \lim_{n \to \infty} r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n) = 0 \ . \qquad (\mathcal{R}(r_n, u_n))$$

Given a cluster functional H on $(\mathbb{R}^d)^{\mathbb{Z}}$, we want to estimate the limiting quantity

$$\boldsymbol{\nu}^*(H) = \lim_{n \to \infty} \boldsymbol{\nu}^*_{n,r_n}(H) = \lim_{n \to \infty} \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \ . \qquad (1.1)$$

To guarantee existence of the limit we will require additional anticlustering assumptions on the time series $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$. For $\boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}}$ define $\boldsymbol{x}^* = \sup_{j \in \mathbb{Z}} |\boldsymbol{x}_j|$. The cluster indices of interest are, among others:

- the extremal index obtained with $H_1(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$, $\boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}}$;
- the cluster size distribution obtained with

$$H_2(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j \in \mathbb{Z}} \mathbb{1}\{|\boldsymbol{x}_j| > 1\} = m\right\} , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}} , \quad m \in \mathbb{N} ;$$
$$(1.2)$$

- the stop-loss index of a univariate time series obtained with

$$H_3(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j \in \mathbb{Z}} (x_j - 1)_+ > \eta\right\} , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}} , \quad \eta > 0 ; \quad (1.3)$$

- the large deviation index of a univariate time series obtained with

$$H_4(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\} , \quad K(\boldsymbol{x}) = \left(\sum_{j \in \mathbb{Z}} x_j\right)_+ , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}} ;$$
$$(1.4)$$

- the ruin index of a univariate time series obtained with

$$H_5(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\} \, , \quad K(\boldsymbol{x}) = \sup_{i \in \mathbb{Z}} \left(\sum_{j \leq i} x_j \right)_+ \, , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}} \, .$$

$$(1.5)$$

As mentioned above, the extremal index is the classical quantity that arises in the extreme value theory for dependent sequences. The cluster size distribution is again a well-known object and was studied in [Hsi91] and [DR10]. The large deviation index was studied under the name *cluster index* in [MW13, MW14]. It quantifies the effect of dependence in large deviations results. The remaining cluster indices seem to be new.

It has to be pointed out that the aforementioned cluster indices describe clustering of extremes in blocks of an increasing size r_n. In this general framework, we can also consider summation functionals

$$H_\phi(\boldsymbol{x}) = \sum_{j \in \mathbb{Z}} \phi(\boldsymbol{x}_j, \ldots, \boldsymbol{x}_{j+h})$$

with $\phi : \mathbb{R}^{d(h+1)} \to \mathbb{R}$. Such summation functionals yield *tail array sums*, which in turn give finite-dimensional extremal characteristics. A suitable choice of ϕ leads to *extremogram* (see [DM09]) or the distribution of the spectral tail process (see [DSW15]). See also Section 10.4.3 in [KS20] for more examples. However, as we will indicate below, the tail array sums are not interesting in the context of the present paper.

Several methods of estimation of the limit $\boldsymbol{\nu}^*(H)$ in (1.1) may be employed. The natural one is to consider a statistics based on disjoint blocks of size r_n, cf. [DR10] and [KS20],

$$\widetilde{\boldsymbol{\nu}}^*_{n,r_n}(H) := \frac{1}{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/u_n) \, ,$$

where $m_n = [n/r_n]$. It is proven in the aforementioned references that the appropriately scaled and centered estimator is asymptotically normal with the limiting variance given by $\boldsymbol{\nu}^*(H^2)$ (see [KS20, Chapter 10] for the expression for the limiting variance). The data-based estimator is constructed as follows. Let $k_n \to \infty$ be a sequence of integers and define u_n by $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$. Let $|\boldsymbol{X}|_{(n:1)} \leq \cdots \leq |\boldsymbol{X}|_{(n:n)}$ be order statistics from $|\boldsymbol{X}_1|, \ldots, |\boldsymbol{X}_n|$. Define

$$\widehat{\boldsymbol{\nu}}^*_{n,r_n}(H) := \frac{1}{k_n} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/|\boldsymbol{X}|_{(n:n-k_n)}) \, . \tag{1.6}$$

Although some special cases were considered (estimation of the extremal index in [Hsi91] and [SW94]; tail array sums in [RLdH98]), the general theory was developed in [DR10]. The summary of the theory for the disjoint blocks estimators can be found in [KS20, Chapter 10], where consistency and the central limit theorems are established. The limiting variance of the disjoint blocks estimator

can be represented as

$$\boldsymbol{\nu}^*(\{H - \boldsymbol{\nu}^*(H)\mathcal{E}\}^2) \, , \tag{1.7}$$

where $\mathcal{E}(\boldsymbol{x}) = \sum_{j \in \mathbb{Z}} \mathbb{1}\{|\boldsymbol{x}_j| > 1\}$. This result was established (implicitly) in [DR10], but the form of the limiting variance is again given in [KS20, Chapter 10].

In this paper we consider the sliding blocks statistics

$$\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(H) := \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) \, , \tag{1.8}$$

where $q_n = n - r_n - 1$ and the corresponding estimator defined in terms of order statistics:

$$\widehat{\boldsymbol{\mu}}^*_{n,r_n}(H) = \frac{1}{r_n k_n} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/|\boldsymbol{X}|_{(n:n-k_n)}\right) \, . \tag{1.9}$$

The sliding blocks estimators have been studied for some specific functionals H, however there has been no unified theory available. Recently, [DN20] used the framework of [DR10] and showed that the limiting variance of the sliding blocks estimator never exceeds that of the disjoint blocks estimator. In case of the extremal index, both variances were proven to be equal.

The goal of this paper is to obtain the asymptotic normality of the sliding blocks estimators. Our focus is on providing the conditions that can be easily verified for a variety of time series models. At the same time, we will show that the limiting variance of both disjoint and sliding blocks estimators is the same. To achieve our goal, we combine [DR10] approach with the modern theory of stationary, regularly varying time series. We note in passing that in case of the tail array sums, the sliding blocks estimators reduce to disjoint blocks one. As such, the limiting theory is already known, see [RLdH98] and Chapter 10 in [KS20]. In order to proceed, in Section 2 we fix the notation, recall the notion of the tail process associated to a stationary regularly varying time series; and introduce the cluster indices.

Next, we need to answer a non-trivial question: *When does the limit $\boldsymbol{\nu}^*(H)$ exist?*. For this, Section 3 deals with convergence of cluster measures and cluster indices $\boldsymbol{\nu}^*(H)$ appear as the limit. Existence of the limit requires an anticlustering assumption. In conjunction with a particular choice of functionals, we will be in position to give specific examples of cluster indices. The contents of this section is based on [KS20, Chapter 6]. Some results stem from [MW14, MW16] and [BPS18].

The main result is Theorem 4.3. We prove the central limit theorem for the data-based sliding blocks estimator (1.9) under easy to verify assumptions. Those conditions can be verified for a variety of models: regularly varying functions of Markov chains, infinite order moving averages, max-stable processes. See [KSW19] and [KS20, Part III].

The most important (and somehow surprising) conclusion is that both sliding (1.9) and disjoint (1.6) blocks estimators yield the same variance. This is in

agreement with the result for the extremal index in [DN20]. On the other hand, it seems to be a contradiction with other available results; see [BS18b], [BS18a] or [ZVB20]. The main difference is that we obtain our asymptotic results in the Peaks-over-Threshold framework, while the latter papers deal with the block maxima framework. We discuss different results in both frameworks in Section 5.

In Section 6 we illustrate the asymptotic theory by a small simulation study for simple time series models, AR(1) and ARCH(1). Interestingly, although estimators of the extremal index perform better in case of a stronger dependence (which is not surprising, see e.g. simulation studies in [RSF09]), we have the opposite situation for the stop-loss index.

All proofs are included in Section 7.

2. Preliminaries

In this section we fix the notation and introduce the relevant classes of functions. In Section 2.3 we recall the notion of the tail and the spectral tail process (cf. [BS09]). In Section 2.4 we define cluster indices; see [KS20, Chapter 5] for a detailed introduction.

2.1. Notation

Let $|\cdot|$ be a norm on $\mathbb{R}^d$. For a sequence $\boldsymbol{x} = \{\boldsymbol{x}_j, j \in \mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}}$ and $i \le j \in \mathbb{Z} \cup \{-\infty, \infty\}$ we denote $\boldsymbol{x}_{i,j} = (\boldsymbol{x}_i, \ldots, \boldsymbol{x}_j) \in (\mathbb{R}^d)^{j-i+1}$, $\boldsymbol{x}_{i,j}^* = \max_{i \le l \le j} |\boldsymbol{x}_l|$ and $\boldsymbol{x}^* = \sup_{j \in \mathbb{Z}} |\boldsymbol{x}_j|$. By $\boldsymbol{0}$ we denote the zero sequence; its dimension can be different in each of its occurrences.

By $\ell_0(\mathbb{R}^d)$ we denote the set of $\mathbb{R}^d$-valued sequences which tend to zero at infinity. Likewise, $\ell_1(\mathbb{R}^d)$ consists of sequences such that $\sum_{j \in \mathbb{Z}} |\boldsymbol{x}_j| < \infty$.

We will use the blocking method. If $\boldsymbol{X}$ is a time series of interest, then $(\boldsymbol{X}_1^{\dagger}, \ldots, \boldsymbol{X}_n^{\dagger})$ is a pseudo-sample such that the blocks $(\boldsymbol{X}_{(i-1)r_n+1}^{\dagger}, \ldots, \boldsymbol{X}_{ir_n}^{\dagger})$, $i = 1, \ldots, m_n = [n/r_n]$, are mutually independent with the same distribution as the original block $(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{r_n})$.

2.2. Classes of functions

Functionals H are defined on $\ell_0(\mathbb{R}^d)$ with the convention $H(\boldsymbol{x}_{i,j}) = H((\boldsymbol{0}, \boldsymbol{x}_{i,j}, \boldsymbol{0}))$. In particular, the map $\mathcal{E}$ is defined on $\ell_0(\mathbb{R}^d)$ by $\mathcal{E}(\boldsymbol{x}) = \sum_{j \in \mathbb{Z}} \mathbb{1}\{|\boldsymbol{x}_j| > 1\}$. For $s > 0$, the function $H_s : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ is defined by $H_s(\boldsymbol{x}) = H(\boldsymbol{x}/s)$. We consider the following classes:

- $\mathcal{L}$ is the class of bounded real-valued functions defined on $(\mathbb{R}^d)^{\mathbb{Z}}$ that are either Lipschitz continuous with respect to the uniform norm or almost surely continuous with respect to the distribution of the tail process $\boldsymbol{Y}$. This class includes functions like $\mathbb{1}\{\boldsymbol{x}^* > 1\}$, $\mathbb{1}\left\{\sum_{j \in \mathbb{Z}} |\boldsymbol{x}_j| > 1\right\}$. See Remark 6.1.6 in [KS20].

- $\mathcal{A} \subset \mathcal{L}$ is the class of shift-invariant functionals with support separated from $\mathbf{0}$. In particular, for $H \in \mathcal{A}$, $H(\mathbf{0}) = 0$. The class $\mathcal{A}$ includes $\mathbb{1}\{\boldsymbol{x}^* > 1\}$.
- $\mathcal{K}$ is the class of shift-invariant functionals $K : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ defined on $\ell_1(\mathbb{R}^d)$ such that $K(\mathbf{0}) = 0$ and which are Lipschitz continuous with constant L_K, i.e.

$$|K(\boldsymbol{x}) - K(\boldsymbol{y})| \le L_K \sum_{j \in \mathbb{Z}} |\boldsymbol{x}_j - \boldsymbol{y}_j| \ , \quad \boldsymbol{x}, \boldsymbol{y} \in \ell_1(\mathbb{R}^d) \ . \tag{2.1}$$

- $\mathcal{B} \subset \mathcal{L}$ is the class of functionals H of the form $H = \mathbb{1}\{K > 1\}$, where $K \in \mathcal{K}$. Functionals in $\mathcal{B}$ may have support which is not separated from $\mathbf{0}$. The typical example is $H(\boldsymbol{x}) = \mathbb{1}\left\{\sum_j |\boldsymbol{x}_j| > 1\right\}$; note that $H \notin \mathcal{A}$.

2.3. Tail and spectral tail process

Let $\boldsymbol{X} = \{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying time series with values in $\mathbb{R}^d$ and tail index α. In particular,

$$\lim_{x \to \infty} \frac{\mathbb{P}(|\boldsymbol{X}_0| > tx)}{\mathbb{P}(|\boldsymbol{X}_0| > x)} = t^{-\alpha}$$

for all $t > 0$. Then, there exists a sequence $\boldsymbol{Y} = \{\boldsymbol{Y}_j, j \in \mathbb{Z}\}$ such that

$$\mathbb{P}(x^{-1}(\boldsymbol{X}_i, \ldots, \boldsymbol{X}_j) \in \cdot \mid |\boldsymbol{X}_0| > x) \text{ converges weakly to } \mathbb{P}((\boldsymbol{Y}_i, \ldots, \boldsymbol{Y}_j) \in \cdot)$$

as $x \to \infty$ for all $i \le j \in \mathbb{Z}$. We call $\boldsymbol{Y}$ the tail process. See [BS09]. We note that, in particular, $|\boldsymbol{Y}_0|$ has Pareto distribution with the density $\alpha x^{-\alpha-1}$, $x > 1$. As such, it follows automatically that $\boldsymbol{Y}^* = \sup_{j \in \mathbb{Z}} |\boldsymbol{Y}_j| > 1$. Equivalently, viewing $\boldsymbol{X}$ and $\boldsymbol{Y}$ as random elements with values in $(\mathbb{R}^d)^{\mathbb{Z}}$, we have for every bounded or non-negative functional H on $(\mathbb{R}^d)^{\mathbb{Z}}$, continuous with respect to the product topology,

$$\lim_{x \to \infty} \frac{\mathbb{E}[H(x^{-1}\boldsymbol{X})\mathbb{1}\{|\boldsymbol{X}_0| > x\}]}{\mathbb{P}(|\boldsymbol{X}_0| > x)} = \mathbb{E}[H(\boldsymbol{Y})] \ .$$

Define $\boldsymbol{\Theta}_j = \boldsymbol{Y}_j/|\boldsymbol{Y}_0|$, $j \in \mathbb{Z}$. The sequence $\boldsymbol{\Theta} = \{\boldsymbol{\Theta}_j, j \in \mathbb{Z}\}$ is called the spectral tail process. The random variable $|\boldsymbol{Y}_0|$ has the Pareto distribution with index α and is independent from $\boldsymbol{\Theta}$. Hence for a non-negative measurable function $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$,

$$\mathbb{E}[H(\boldsymbol{Y})] = \int_1^{\infty} \mathbb{E}[H(r\boldsymbol{\Theta})]\alpha r^{-\alpha-1}\mathrm{d}r \ . \tag{2.2}$$

2.4. Cluster measure and cluster indices

Consider the infargmax functional $\mathcal{C}_0$ defined on $(\mathbb{R}^d)^{\mathbb{Z}}$ by $\mathcal{C}_0(\boldsymbol{y}) = \inf\{j : \boldsymbol{y}_{-\infty,j}^* = \boldsymbol{y}^*\}$, with the convention that $\inf \emptyset = +\infty$. If $\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$ then we can define

$$\vartheta = \mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) = 0) \ . \tag{2.3}$$

In fact, $\mathcal{C}_0$ can be replaced with any anchoring map (see [PS18] and [KS20, Theorem 5.5.3]), but we do not pursue it here. For the purpose of this paper it is sufficient to note that we can write alternatively

$$\vartheta = \mathbb{P}\left(\sup_{j \leq -1} |\boldsymbol{Y}_j| \leq 1\right) . \tag{2.4}$$

The relationship between (2.3) and (2.4) is certainly not obvious. The proof is given in Section 7.1. Equation (2.4) emphasizes a special role of the event $\{\sup_{j \leq -1} |\boldsymbol{Y}_j| \leq 1\}$ and with its help ϑ can be recognized as the (candidate) extremal index. It becomes the usual extremal index under additional mixing and anticlustering conditions.

Definition 2.1. *Let $\boldsymbol{Y}$ and $\boldsymbol{\Theta}$ be the tail process and the spectral tail process, respectively, such that $\mathbb{P}(\lim_{|j| \to \infty} \boldsymbol{Y}_j = \boldsymbol{0}) = 1$. The **cluster measure** is the measure $\boldsymbol{\nu}^*$ on $\ell_0(\mathbb{R}^d)$ defined by*

$$\boldsymbol{\nu}^* = \vartheta \int_0^\infty \mathbb{E}[\delta_{r\boldsymbol{\Theta}} \mathbb{1}\{\mathcal{C}_0(\boldsymbol{\Theta}) = 0\}]\alpha r^{-\alpha-1}\mathrm{d}r . \tag{2.5}$$

The measure $\boldsymbol{\nu}^*$ is boundedly finite on $(\mathbb{R}^d)^{\mathbb{Z}} \setminus \{\boldsymbol{0}\}$, puts no mass at $\boldsymbol{0}$ and is α-homogeneous. Furthermore, the cluster measure can be expressed in terms of another sequence.

Definition 2.2. *Assume that $\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$. The conditional spectral tail process $\boldsymbol{Q}$ is a random sequence with the distribution of $(\boldsymbol{Y}^*)^{-1}\boldsymbol{Y}$ conditionally on $\mathcal{C}_0(\boldsymbol{Y}) = 0$.*

The sequence $\boldsymbol{Q}$ appeared implicitly in the seminal paper [DH95]. See also [BS09], [PS18, Definition 3.5] and [KS20, Chapter 5]. An abstract setting is considered in [DHS18].

Note that $\mathcal{C}_0(\boldsymbol{Y}) = 0$ if and only if $\mathcal{C}_0(\boldsymbol{\Theta}) = 0$. Then also $\boldsymbol{Y}^* = |\boldsymbol{Y}_0|$. Thus, (2.5) and the definition of $\boldsymbol{Q}$ give for a bounded or non-negative measurable function H on $\ell_0(\mathbb{R}^d)$,

$$\boldsymbol{\nu}^*(H) = \vartheta \int_0^\infty \mathbb{E}[H(r\boldsymbol{Q})]\alpha r^{-\alpha-1}\mathrm{d}r = \vartheta \int_0^\infty \mathbb{E}[H(r\boldsymbol{\Theta})\mathbb{1}\{\mathcal{C}_0(\boldsymbol{\Theta}) = 0\}]\alpha r^{-\alpha-1}\mathrm{d}r . \tag{2.6}$$

If moreover H is such that $H(\boldsymbol{y}) = 0$ if $\boldsymbol{y}^* \leq \epsilon$ for one $\epsilon > 0$, then

$$\boldsymbol{\nu}^*(H) = \epsilon^{-\alpha}\mathbb{E}[H(\epsilon\boldsymbol{Y})\mathbb{1}\{\mathcal{C}_0(\boldsymbol{Y}) = 0\}] = \epsilon^{-\alpha}\mathbb{E}[H(\epsilon\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\}] . \tag{2.7}$$

Note that with $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$ and recalling that $\boldsymbol{Y}^* > 1$, (2.7) reduces to (2.4). As such, functionals from $\mathcal{A}$ will have typically the representation given in (2.7). On the other hand, for functionals from $\mathcal{B}$ we are not able to conclude the representation (2.7), however, the general form (2.6) is still valid, possibly under additional conditions. See Section 3.3. Comparing (2.5) or (2.7) with (2.2) we can see that the $\boldsymbol{\nu}^*(H)$ does not agree with $\mathbb{E}[H(\boldsymbol{Y})]$. The additional indicator comes essentially from the conditioning on the location of the maximum of the sequence $\boldsymbol{Y}$.

Definition 2.3 (Cluster index)**.** *We will call $\boldsymbol{\nu}^*(H)$ the cluster index associated to the functional H.*

3. Convergence of cluster measure

Recall $\mathcal{R}(r_n, u_n)$. Define the measures $\boldsymbol{\nu}^*_{n,r_n}$, $n \geq 1$, on $\ell_0(\mathbb{R}^d)$ as follows:

$$\boldsymbol{\nu}^*_{n,r_n} = \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[\delta_{u_n^{-1}\boldsymbol{X}_{1,r_n}}\right] .$$

We are interested in convergence of $\boldsymbol{\nu}^*_{n,r_n}$ to $\boldsymbol{\nu}^*$. The results of this section are extracted from [KS20, Chapter 6]. See also [PS18] and [BPS18].

3.1. Anticlustering condition

For each fixed $r \in \mathbb{N}$, the distribution of $u_n^{-1}\boldsymbol{X}_{-r,r}$ conditionally on $|\boldsymbol{X}_0| > u_n$ converges weakly to the distribution of $\boldsymbol{Y}_{-r,r}$. In order to let r tend to infinity, we must embed all these finite vectors into one space of sequences. By adding zeroes on each side of the vectors $u_n^{-1}\boldsymbol{X}_{-r,r}$ and $\boldsymbol{Y}_{-r,r}$ we identify them with elements of the space $\ell_0(\mathbb{R}^d)$. Then $\boldsymbol{Y}_{-r,r}$ converges (as $r \to \infty$) to $\boldsymbol{Y}$ in $\ell_0(\mathbb{R}^d)$ if (and only if) $\boldsymbol{Y} \in \ell_0(\mathbb{R}^d)$ almost surely.

However, this is not enough for statistical purposes and we consider the following definition.

Definition 3.1 ([DH95], Condition 2.8)**.** *Condition $\mathcal{AC}(r_n, u_n)$ holds if for all $x, y > 0$,*

$$\lim_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\left(\max_{k \leq |j| \leq r_n} |\boldsymbol{X}_j| > u_n x \mid |\boldsymbol{X}_0| > u_n y\right) = 0 . \qquad (\mathcal{AC}(r_n, u_n))$$

Condition $\mathcal{AC}(r_n, u_n)$ is referred to as the anticlustering condition. It is fulfilled by many models, including geometrically ergodic Markov chains, short-memory linear or max-stable processes. $\mathcal{AC}(r_n, u_n)$ implies that $\boldsymbol{Y} \in \ell_0(\mathbb{R}^d)$. Its main consequence is the following result.

Proposition 3.2 ([BS09], Proposition 4.2; [KS20], Theorem 6.1.4)**.** *Let $H \in \mathcal{L}$. If Condition $\mathcal{AC}(r_n, u_n)$ holds, then*

$$\lim_{n \to \infty} \mathbb{E}[H(u_n^{-1}\boldsymbol{X}_{-r_n,r_n}) \mid |\boldsymbol{X}_0| > u_n] = \mathbb{E}[H(\boldsymbol{Y})] .$$

Condition $\mathcal{AC}(r_n, u_n)$ holds for sequence of i.i.d. random variables whenever $\lim_{n \to \infty} r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n) = 0$, which can be recognized as on the restrictions imposed in $\mathcal{R}(r_n, u_n)$

3.2. Vague convergence of cluster measure

We now investigate the unconditional convergence of $u_n^{-1}\boldsymbol{X}_{1,r_n}$. Contrary to Proposition 3.2, where an extreme value was imposed at time 0, a large value in

the cluster can happen at any time. Moreover, the convergence of $\boldsymbol{\nu}^*_{n,r_n}(H)$ to $\boldsymbol{\nu}^*(H)$ may hold only for shift-invariant functionals H. Therefore, we need the following definition.

Definition 3.3. *The space $\tilde{\ell}_0(\mathbb{R}^d)$ is the space of equivalence classes of $\ell_0(\mathbb{R}^d)$ endowed with the equivalence relation $\sim$ defined by*

$$\boldsymbol{x} \sim \boldsymbol{y} \Longleftrightarrow \exists j \in \mathbb{Z} \,, \ B^j\boldsymbol{x} = \boldsymbol{y} \,,$$

where B is the backshift operator.

The proof of the next result is given in Section 7.

Proposition 3.4. *Let condition $\mathcal{AC}(r_n, u_n)$ hold. The sequence of measures $\boldsymbol{\nu}^*_{n,r_n}$, $n \geq 1$ converges vaguely$^\#$ on $\tilde{\ell}_0(\mathbb{R}^d) \setminus \{\boldsymbol{0}\}$ to $\boldsymbol{\nu}^*$, that is, for all $H \in \mathcal{A}$,*

$$\lim_{n\to\infty} \boldsymbol{\nu}^*_{n,r_n}(H) = \lim_{n\to\infty} \frac{\mathbb{E}[H(u_n^{-1}\boldsymbol{X}_{1,r_n})]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \boldsymbol{\nu}^*(H) \,.$$

The immediate consequence is the following limit (cf. (2.3)):

$$\lim_{n\to\infty} \frac{\mathbb{P}(\boldsymbol{X}^*_{1,r_n} > u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \vartheta \,.$$

Since $H_2, H_3 \in \mathcal{A}$ (cf. (1.2)-(1.3)), we can introduce the following cluster indices.

Example 3.5 (Cluster size distribution). If $\mathcal{AC}(r_n, u_n)$ holds, Proposition 3.4 yields

$$\lim_{n\to\infty} \mathbb{P}\left(\sum_{j=1}^{r_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\} = m \mid \boldsymbol{X}^*_{1,r_n} > u_n\right)$$

$$= \mathbb{P}\left(\sum_{j\in\mathbb{Z}} \mathbb{1}\{|\boldsymbol{Y}_j| > 1\} = m \mid \boldsymbol{Y}^*_{-\infty,-1} \leq 1\right) =: \pi(m) \,.$$

⊞

Example 3.6 (Stop-loss index). Consider a univariate time series. Define the stop-loss index:

$$\theta_{\text{stoploss}}(\eta) = \lim_{n\to\infty} \frac{\mathbb{P}\left(\sum_{j=1}^{r_n}(X_j - u_n)_+ > \eta u_n\right)}{r_n\mathbb{P}(X_0 > u_n)}$$

$$= \mathbb{P}\left(\sum_{j=0}^{\infty}(Y_j - 1)_+ > \eta, \boldsymbol{Y}^*_{-\infty,-1} \leq 1\right) \,.$$

This index seems to be new.

⊞

3.3. Indicator functionals not vanishing around zero

Proposition 3.4 entails convergence of $\boldsymbol{\nu}^*_{n,r_n}(H)$ for $H \in \mathcal{A}$. For functionals which are not defined on the whole space $\ell_0(\mathbb{R}^d)$, such as H_4 and H_5 from (1.4)-(1.5), we need an additional assumption on Asymptotic Negligibility of Small Jumps.

Definition 3.7. *Condition* ANSJB(r_n, u_n) *holds if for all $\eta > 0$,*

$$\lim_{\epsilon \to 0} \limsup_{n \to \infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j| \mathbb{1}\{|\boldsymbol{X}_j| \leq \epsilon u_n\} > \eta u_n)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = 0 . \qquad (\text{ANSJB}(r_n, u_n))$$

The proofs of the next two results are given in Section 7.

Lemma 3.8. *If $\mathcal{AC}(r_n, u_n)$ and ANSJB(r_n, u_n) hold, then*

$$\lim_{n \to \infty} \frac{\mathbb{P}(\sum_{i=1}^{r_n} |\boldsymbol{X}_j| > u_n)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \mathbb{E}\left[\left(\sum_{j \in \mathbb{Z}} |\boldsymbol{Q}_j|\right)^\alpha\right] < \infty .$$

Proposition 3.9. *Assume that $\mathcal{AC}(r_n, u_n)$ and ANSJB(r_n, u_n) hold. Then for $K \in \mathcal{K}$,*

$$\boldsymbol{\nu}^*(\mathbb{1}\{K > 1\}) = \lim_{n \to \infty} \frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \vartheta \int_0^\infty \mathbb{P}(K(z\boldsymbol{Q}) > 1)\alpha z^{-\alpha-1}\mathrm{d}z$$

and the latter integral is finite.

If K is a 1-homogeneous satisfying the assumptions of Proposition 3.9, then

$$\boldsymbol{\nu}^*(\mathbb{1}\{K > 1\}) = \vartheta \mathbb{E}[K_+^\alpha(\boldsymbol{Q})] = \mathbb{E}[K_+^\alpha(\boldsymbol{\Theta}_{0,\infty}) - K_+^\alpha(\boldsymbol{\Theta}_{1,\infty})] .$$

Example 3.10 (Large deviations index). Let $\{X_j, j \in \mathbb{Z}\}$ be an univariate time series. The functional H_4 defined in (1.4) yields the large deviations index:

$$\theta_{\text{largedev}} = \lim_{n \to \infty} \frac{\mathbb{P}\left(\left(\sum_{j=1}^{r_n} X_j\right)_+ > u_n\right)}{r_n \mathbb{P}(|X_0| > u_n)} = \mathbb{E}\left[\left(\sum_{j=0}^\infty \Theta_j\right)^\alpha_+ - \left(\sum_{j=1}^\infty \Theta_j\right)^\alpha_+\right] .$$

The index θ_{largedev}, under the name *cluster index*, was introduced in [MW16]. ⊞

Example 3.11 (Ruin index). Take H_5 defined in (1.5). Proposition 3.9 gives

$$\theta_{\text{ruin}} = \lim_{n \to \infty} \frac{\mathbb{P}(\max_{1 \leq j \leq r_n} \sum_{i=1}^j X_i > u_n)}{r_n \mathbb{P}(|X_0| > u_n)} = \vartheta \mathbb{E}\left[\sup_{i \in \mathbb{Z}} \left(\sum_{j \leq i} Q_j\right)^\alpha_+\right] .$$

⊞

Remark 3.12. At this point we would like to point out the following. Consider $H \in \mathcal{A}$ to be an indicator functional. If moreover H is such that $H(\boldsymbol{y}) = 0$ if $\boldsymbol{y}^* \leq 1$, then thanks to (2.7),

$$\boldsymbol{\nu}^*(H) = \mathbb{E}[H(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\}] \in (0,1] \ .$$

This is the situation for the extremal index and the functionals from Examples 3.5 and 3.6. On the other hand, if H does not vanish around zero, then at the first place we need additional conditions to guarantee that $\boldsymbol{\nu}^*(H) < \infty$ (e.g. ANSJB(r_n, u_n)). Second, there is no restriction on the values of the cluster index.

4. Central limit theorem for blocks estimators

4.1. Sliding blocks estimators

Let $q_n = n - r_n + 1$. Thanks to Proposition 3.4 and Proposition 3.9, we have for $H \in \mathcal{A} \cup \mathcal{B}$,

$$\lim_{n \to \infty} \mathbb{E}\left[\frac{\sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right)}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}\right]$$
$$= \lim_{n \to \infty} \frac{\mathbb{E}\left[H\left(\boldsymbol{X}_{1,r_n}/u_n\right)\right]}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \lim_{n \to \infty} \boldsymbol{\nu}^*_{n,r_n}(H) = \boldsymbol{\nu}^*(H) \ .$$

This indicates that a consistent pseudo-estimator of $\boldsymbol{\nu}^*(H)$ can be defined as

$$\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(H) := \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) \ . \qquad (4.1)$$

The above estimator is not feasible, since it involves an unspecified sequence $\{u_n\}$ and the tail of $|\boldsymbol{X}_0|$. Thus, in (4.1) we replace $q_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$ with its empirical estimate $\sum_{j=1}^{q_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\}$ to obtain a quasi-feasible estimator

$$\widehat{\widetilde{\boldsymbol{\mu}}}^*_{n,r_n}(H) = \frac{1}{r_n} \frac{1}{\sum_{j=1}^{q_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\}} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) \ .$$

Likewise, let k_n be an intermediate sequence of integers, i.e. $\lim_{n \to \infty} k_n = \infty$, $\lim_{n \to \infty} k_n/n = 0$. Define u_n by $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$. Replacing u_n with $|\boldsymbol{X}|_{(n:n-k_n)}$ and noting that (assuming for simplicity that there are not ties in the data)

$$\sum_{j=1}^{n} \mathbb{1}\left\{|\boldsymbol{X}_j| > |\boldsymbol{X}|_{(n:n-k_n)}\right\} = k_n \ ,$$

we obtain a feasible estimator of $\boldsymbol{\nu}^*(H)$ given in (1.9).

4.2. Weak dependence assumptions

For asymptotic normality, we need to strengthen the anticlustering condition $\mathcal{AC}(r_n, u_n)$.

Definition 4.1. *Condition $\mathcal{S}(r_n, u_n)$ holds if for all $s, t > 0$*

$$\lim_{m \to \infty} \limsup_{n \to \infty} \frac{1}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=m}^{r_n} \mathbb{P}(|\boldsymbol{X}_0| > u_n s, |\boldsymbol{X}_j| > u_n t) = 0 . \quad (\mathcal{S}(r_n, u_n))$$

This condition implies that $\sum_{j \in \mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_j| > 1) < \infty$. The latter series appears explicitly in the statement for the limiting variance.

Dependence in $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ will be controlled by the β-mixing rates $\{\beta_n\}$. Recall $\mathcal{R}(r_n, u_n)$. Let $\{\ell_n\}$ be a sequence of integers such that $\lim_{n \to \infty} \ell_n = \infty$ and $\lim_{n \to \infty} \ell_n / r_n = 0$.

Definition 4.2. *Condition $\beta(r_n)$ holds if:*

1. *$\beta_j = O(j^{-\nu})$, $\nu > 1$ and $\lim_{n \to \infty} r_n^{1+\nu}/n = +\infty$; and*
2. *there exists $\delta > 0$ such that $\lim_{n \to \infty} r_n^{\nu-\delta} \mathbb{P}(|\boldsymbol{X}_0| > u_n) = +\infty$.*

From the basic assumptions on the time series, we have $\lim_{n \to \infty} r_n / n = 0$. Thus, ν has to be big enough. The above mixing condition is clearly satisfied for time series with geometric mixing rates since then ν can be chosen arbitrarily large.

4.3. Main result

Let $\mathbb{G}$ be the Gaussian process on $L^2(\boldsymbol{\nu}^*)$ with covariance

$$\operatorname{cov}(\mathbb{G}(H), \mathbb{G}(\widetilde{H})) = \boldsymbol{\nu}^*(H\widetilde{H}) .$$

Recall that for a functional $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ and $s > 0$ we define $H_s(\boldsymbol{x}) = H(\boldsymbol{x}/s)$.

The main result of this paper is Theorem 4.3, the asymptotic normality of the appropriately normalized estimator $\widehat{\boldsymbol{\mu}}^*_{n,r_n}(H)$. The limiting variance agrees with the one for the disjoint blocks estimator; cf. [DR10] and [KS20, Chapter 10].

Theorem 4.3. *Let $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying $\mathbb{R}^d$-valued time series. Assume that $\mathcal{R}(r_n, u_n)$, $\beta(r_n)$, $\mathcal{S}(r_n, u_n)$ hold. Fix $0 < s_0 < 1 < t_0 < \infty$. Let $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ be a shift-invariant measurable map such that the class $\{H_s : s \in [s_0, t_0]\}$ is linearly ordered. Assume moreover that*

$$\lim_{n \to \infty} \sqrt{k_n} \sup_{s \in [s_0, t_0]} |\mathbb{E}[\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(\mathcal{E}_s)] - \boldsymbol{\nu}^*(\mathcal{E}_s)| = 0 , \quad (4.2a)$$

$$\lim_{n \to \infty} \sqrt{k_n} \sup_{s \in [s_0, t_0]} |\mathbb{E}[\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(H_s)] - \boldsymbol{\nu}^*(H_s)| = 0 . \quad (4.2b)$$

If $H \in \mathcal{A}$, then

$$\sqrt{k_n}\left\{\widehat{\boldsymbol{\mu}}^*_{n,r_n}(H) - \boldsymbol{\nu}^*(H)\right\} \xrightarrow{\mathrm{d}} \mathbb{G}(H - \boldsymbol{\nu}^*(H)\mathcal{E}) . \tag{4.3}$$

If moreover ANSJB(r_n, u_n) *is satisfied, then* (4.3) *holds for* $H \in \mathcal{B}$.

Remark 4.4. The limiting distribution is centered Gaussian with variance (cf. Lemma 7.22):

$$\boldsymbol{\nu}^*(\{H - \boldsymbol{\nu}^*(H)\mathcal{E}\}^2) = \boldsymbol{\nu}^*(H^2) - 2\boldsymbol{\nu}^*(H)\boldsymbol{\nu}^*(H\mathcal{E}) + (\boldsymbol{\nu}^*(H))^2\boldsymbol{\nu}^*(\mathcal{E}^2)$$

$$= \boldsymbol{\nu}^*(H^2) - 2\boldsymbol{\nu}^*(H)\mathbb{E}[H(\boldsymbol{Y})] + (\boldsymbol{\nu}^*(H))^2 \sum_{j \in \mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_j| > 1) .$$

Thus, in view of (1.7), the limiting variance for the sliding blocks estimator agrees with the one for the disjoint blocks estimator. $\oplus$

Remark 4.5. The linear ordering of the class $\{H_s : s \in [s_0, t_0]\}$ may seem to be too restrictive. However, all the cluster functionals that, from the authors perspective, seem to be of interest in the context of sliding blocks estimators, have this property. This includes the functionals H_1, H_2, H_3, H_4, H_5 considered in the Introduction. The linear ordering can be replaced with an assumption that the function class is of VC-type (see [DR10]), or can be approximated by VC-classes. See [BBKS20, Lemma A.3], [DK20] and [KS20, Appendix C.4]. This generalization is obvious, but does not bring anything to the contents of the paper, since the random entropy assumption has to be checked for each case separately.

4.4. Examples

Example 4.6 (Extremal index). For $H(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x}^* > 1\}$ we have $\boldsymbol{\nu}^*(H) = \boldsymbol{\nu}^*(H^2) = \vartheta$. The data-based estimator (1.9) is asymptotically normal with mean zero and the limiting variance is

$$\boldsymbol{\nu}^*(\{H - \boldsymbol{\nu}^*(H)\mathcal{E}\}^2) = \vartheta^2 \sum_{j \in \mathbb{Z}} \mathbb{P}(Y_j > 1) - \vartheta .$$

See Section 5 for a discussion on the existing results. $\boxplus$

Example 4.7 (Cluster size distribution). Consider the situation from Example 3.5. The limiting distribution is centered normal with the variance

$$\pi(m)\left(1 - 2\mathbb{P}(\mathcal{E}(\boldsymbol{Y}) = m)\right) + \pi^2(m) \sum_{j \in \mathbb{Z}} \mathbb{P}(Y_j > 1) .$$

$\boxplus$

Example 4.8 (Stop-loss index). Consider the stop-loss index $\theta_{\mathrm{stoploss}}(\eta)$ introduced in Example 3.6. The limiting distribution is centered normal with the variance

$$\theta_{\mathrm{stoploss}}(\eta)\left(1 - 2\mathbb{P}\left(\sum\nolimits_{j \in \mathbb{Z}}(Y_j - 1)_+ > \eta\right)\right) + \theta^2_{\mathrm{stoploss}}(\eta) \sum_{j \in \mathbb{Z}} \mathbb{P}(Y_j > 1) .$$

$\boxplus$

Example 4.9 (Large deviations index). We continue with the situation from Example 3.10. The limiting distribution is centered Gaussian with variance

$$\theta_{\text{largedev}} - 2\theta_{\text{largedev}}\mathbb{P}\left(\left(\sum_{j\in\mathbb{Z}} Y_j\right)_+ > 1\right) + \theta_{\text{largedev}}^2 \sum_{j\in\mathbb{Z}}\mathbb{P}(|Y_j| > 1) .$$

$$\boxplus$$

5. Comments and extensions

5.1. Existing results

We discuss the existing results. For the sake of clarify, we consider univariate, non-negative, regularly varying time series with the marginal distribution F.

PoT approach. In [DN20] the authors study asymptotic normality of the sliding blocks estimators in a general setting. They show that the limiting variance of such estimators does not exceed the one for the disjoint blocks estimators. For the extremal index they found the variances to be equal. As in this paper, they use the threshold u_n such as in $\mathcal{R}(r_n, u_n)$. The results in [DR10] and [KS20, Chapters 9-10] (disjoint blocks) as well as in [DN20] and in the current paper fit into Peaks-Over-Threshold (PoT) framework.

In particular, consider the disjoint blocks estimator of the extremal index,

$$\widetilde{\vartheta}_{n,1} = \widetilde{\vartheta}_{n,1}(x) = \frac{\sum_{i=1}^{m_n} \mathbb{1}\left\{\boldsymbol{X}^*_{(i-1)r_n+1, ir_n} > x\right\}}{\sum_{j=1}^{m_n r_n} \mathbb{1}\{X_j > x\}} .$$

In [KS20, Example 10.4.2] we calculated the limiting variance of $\widetilde{\vartheta}_{n,1}(u_n)$ to be $\sigma_1^2 = -\vartheta + \vartheta^2 \sum_{j\in\mathbb{Z}}\mathbb{P}(Y_j > 1)$. This is in agreement with Corollary 4.6 in [Hsi91] (where the variance σ_1^2 is given in a complicated form). We can see that σ_1^2 agrees with the one limiting variance for the sliding blocks estimator in Example 4.6. The blocks estimator $\widetilde{\vartheta}_{n,1}(u_n)$ is also considered in [SW94] and [WN98].

Block maxima framework. One can also use the threshold c_{r_n} given by

$$r_n\overline{F}(c_{r_n}) \to 1 . \tag{5.1}$$

We are not aware of the asymptotic theory for $\widetilde{\vartheta}_{n,1}(c_{r_n})$. However, using [RSF09, Theorem 4.2] and the delta method we can compare the variances of $\widetilde{\vartheta}_{n,1}(u_n)$ and $\widetilde{\vartheta}_{n,1}(c_{r_n})$:

$$\sigma_1^2 = -\vartheta + \vartheta^2 \sum_{j\in\mathbb{Z}}\mathbb{P}(Y_j > 1) \ \text{ vs. } \sigma_3^2 := \mathrm{e}^{-\vartheta}(1 - \mathrm{e}^{-\vartheta}) - 2\vartheta\mathrm{e}^{-\vartheta} + \vartheta^2 \sum_{j\in\mathbb{Z}}\mathbb{P}(Y_j > 1) .$$

Thus, the estimator $\widetilde{\vartheta}_{n,1}(u_n)$ has a smaller variance than $\widetilde{\vartheta}_{n,1}(c_{r_n})$.

In the following discussion, we will use the threshold (5.1). In [RSF09] the authors consider another disjoint blocks estimator of the extremal index, motivated by the approximation $\log(1-x) \sim x$ ($x \to 0$). Also, the corresponding sliding blocks estimator is considered. It is shown that the sliding blocks one yields a smaller asymptotic variance.

In [BS18b, BS18a] the authors estimate the parameters (α, σ) of the Fréchet distribution stemming from the limiting behaviour of the maxima. Disjoint blocks yield a larger variance than sliding blocks. Similarly, in [BB18] the authors use the blocking method to estimate the extremal index and again the sliding block estimator is more efficient.

The estimator $\widetilde{\vartheta}_{n,1}(c_{r_n})$ as well as the ones in [RSF09] and [BS18b, BS18a] can be thought of as the application of the block maxima method. Indeed, the threshold c_{r_n} is the normalizing sequence for the limiting distribution of maxima. In the context of the latter two papers, $\boldsymbol{X}^*_{1,r_n}/\sigma_{r_n}$ converges in distribution to a standard Fréchet random variable with tail index α (denoted by Z). On the other hand, for $\xi \in (0,1)$, the pair

$$(\boldsymbol{X}^*_{1,r_n}/\sigma_{r_n}, \boldsymbol{X}^*_{1+[\xi r_n],r_n+[\xi r_n]}/\sigma_{r_n})$$

converges in distribution to a dependent random vector (Z_1, Z_2) with Fréchet marginals and parametrized by $\xi \in (0,1)$. See [BS18a, Lemma 5.1]. Consider now for $f : \mathbb{R} \to \mathbb{R}$

$$\mathbb{G}_n^{(BS)}(f) = \sqrt{m_n} \left\{ m_n^{-1} \sum_{j=1}^{m_n} f\left(\frac{\boldsymbol{X}^*_{(j-1)r_n+1,jr_n}}{\sigma_{r_n}}\right) - \mathbb{E}[f(Z)] \right\},$$

$$\mathbb{F}_n^{(BS)}(f) = \sqrt{m_n} \left\{ q_n^{-1} \sum_{i=1}^{q_n} f\left(\frac{\boldsymbol{X}^*_{i,i+r_n-1}}{\sigma_{r_n}}\right) - \mathbb{E}[f(Z)] \right\}.$$

The aforementioned convergence gives the limiting variance. For the disjoint blocks empirical process the limiting variance is $\mathrm{var}(f(Z))$, while for the sliding blocks one it becomes (cf. Lemma 5.3 in [BS18a])

$$C(f) := 2 \int_0^1 \mathrm{cov}_\xi(f(Z_1), f(Z_2)) \mathrm{d}\xi .$$

In the context of our paper, if we choose $f(z) = \mathbb{1}\{z > 1\}$, then we can evaluate:

$$\mathrm{var}(f(Z)) = \exp(-1) - \exp(-2) > C(f) = 2\exp(-1) - 4\exp(-2) .$$

In the PoT framework considered in our paper, both the disjoint blocks and the sliding blocks empirical processes yield the limiting variance $\boldsymbol{\nu}^*(H)$.

In summary:

- **The PoT method, as proven in this paper, gives the same limiting behaviour for both disjoint and sliding blocks estimators.**

- The situation seems to be different in case of the block maxima method, at least for the inference problems considered up to date.
- One can argue that the blocks maxima method is restricted to estimation of the parameters of the limiting distribution of maxima (the tail index, the extremal index) and is rather hard to see how the method can be employed to other cluster indices.

5.2. Bias

The sliding block estimators of cluster functionals are subjected to bias

$$\mathbb{E}[\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(H)] - \boldsymbol{\nu}^*(H) \ .$$

The bias vanishes asymptotically thanks to the assumption (4.2b). The latter assumption imposes some restrictions on k_n, r_n. Classically, e.g. in case of the Hill estimator of the tail index, the bias is controlled by the second order condition. In the present context we know nothing about how to control bias, except that *there exist sequences k_n and r_n such that* (4.2b) *holds.* On a positive side, from a point of comparing disjoint and sliding blocks estimators, the theoretical bias is obviously the same for both.

5.3. Open questions

- For the sliding blocks estimators, obtain consistency under minimal conditions (that is, without relying on β-mixing). In [KS20, Chapter 10] we obtain consistency of the disjoint blocks estimators for time series that can be approximated by m-dependent sequences, including long memory ones.
- Extend Theorem 4.3 to unbounded functionals H. The method of the proof presented in the paper should be applicable, however, some substantial modifications may be needed. Certainly, more restrictive conditions will need to be implemented.
- In view of the behaviour of $\widetilde{\vartheta}_{n,1}(u_n)$ and $\widetilde{\vartheta}_{n,1}(c_{r_n})$, it would be interesting to know if (whenever possible) the PoT method always gives a smaller variance than the block maxima ones.

6. Simulation study

We conducted some simulations in order to study the finite sample performance of the sliding and disjoint blocks estimators for selected cluster indices.

6.1. Stationary AR process

We start with a simple AR(1) process. For this process we have explicit formulas for all cluster indices. Samples of size $n = 1000$ are generated from AR(1) with $\alpha = 4$ and $\rho = 0.5, 0.9$. Simulations for the classical extremal index are compared to simulations for the stop-loss index.

Extremal index. For AR(1) with $\rho > 0$ the extremal index is $\theta = 1 - \rho^\alpha$; cf. [KS20, p. 396].

- We start with the Hill plots in Figure 1. There, for one simulated data set we compare a performance of both disjoint and sliding blocks estimators. For weak dependence ($\rho = 0.5$), both sliding and disjoint blocks estimators under-estimate the extremal index while for strong dependence ($\rho = 0.9$), the results are stable around the true extremal index for a small number of order statistics. Bigger values of k introduce more bias. In any case, the performance of both estimators is comparable.
- Table 1 includes the results for Monte Carlo simulation for the extremal index based on disjoint and sliding blocks, with the block size $r_n = 7, \ 8, \ 9, \ 10$. We used $k = 5\%$ and 10% order statistics. We note that for the strong dependence ($\rho = 0.9$), the estimation is acceptable for the small block sizes and small k for both disjoint and sliding estimators. A larger block size r_n and/or larger number of order statistics k results in a biased estimation. For weak dependence ($\rho = 0.5$), the results are heavily biased for all considered parameters. We note that both disjoint and sliding blocks estimators yield almost the same variances, which is in agreement with the theoretical results obtained in the paper.
- The box plots and histograms in Figure 2 and Figure 3 are based again on Monte Carlo simulations. The following parameters are used: $\rho = 0.9$, $\alpha = 4$ and the block size $r_n = 7$ along with $k = 5\%$ and 10%. We notice again that $\rho = 0.9$, $r_n = 7$ and $k = 5\%$ yield acceptable results. However, small ρ yields a lot of bias.

In summary, **in case of the extremal index, both disjoint and sliding blocks estimators yield similar results (as suggested by theory). Stronger dependence implies better performance. Typically estimators suffer from a bias.** As such, bias-reduction techniques should be investigated.

We note also that the fact that stronger dependence yields smaller variability of the estimators is not surprising, cf. e.g. Figure 5 in [RSF09].

Stop-loss index. For AR(1) with $\rho > 0$ the formula for the stop-loss index is given in [KS20, p. 619]:

$$\theta_{\text{stop-loss}}(S) = (1 - \rho^\alpha)\mathbb{P}\left(\sum_{j=0}^{\infty}(\rho^j Y_0 - 1)_+ > S\right), \qquad (6.1)$$

where Y_0 is a Pareto random variable with the parameter α.

- At the first step we use the formula (6.1) and performed the Monte-Carlo simulation to obtain the approximate value of the stop-loss index.
- Figure 4 displays Hill plots for the stop-loss index. The Hill plots indicate a similar performance of both disjoint and sliding blocks estimators. We note that, unlike in the extremal index case or in the classical case of the

TABLE 1

The median and the variance (in brackets) of disjoint and sliding blocks estimators for the extremal index. Data are simulated from AR(1) with $\alpha = 4$, $\rho = 0.5$ (thus, $\theta = 0.94$), and $\rho = 0.9$ (thus $\theta = 0.34$). Block size $r_n = 7$, 8, 9, 10. The number of order statistics is $k = 5\%$ and 10% for a sample $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

	$\rho = 0.9$, Extremal Index=0.34				$\rho = 0.5$, Extremal Index= 0.94			
$(k~\%)$	$k = 5$		$k = 10$		$k = 5$		$k = 10$	
$r_n = 7$								
Disjoint bl	**0.34**	(0.05)	**0.31**	(0.03)	**0.68**	(0.05)	**0.58**	(0.03)
Sliding bl	**0.35**	(0.04)	**0.31**	(0.03)	**0.68**	(0.04)	**0.58**	(0.03)
$r_n = 8$								
Disjoint bl	**0.32**	(0.05)	**0.29**	(0.03)	**0.67**	(0.05)	**0.56**	(0.03)
Sliding bl	**0.33**	(0.04)	**0.29**	(0.03)	**0.67**	(0.04)	**0.56**	(0.03)
$r_n = 9$								
Disjoint bl	**0.32**	(0.05)	**0.28**	(0.03)	**0.66**	(0.05)	**0.53**	(0.03)
Sliding bl	**0.32**	(0.04)	**0.28**	(0.03)	**0.65**	(0.05)	**0.53**	(0.03)
$r_n = 10$								
Disjoint bl	**0.30**	(0.05)	**0.26**	(0.03)	**0.64**	(0.05)	**0.52**	(0.03)
Sliding bl	**0.30**	(0.04)	**0.26**	(0.03)	**0.63**	(0.05)	**0.52**	(0.03)

Hill estimator of the tail index, one needs to take a much bigger number of the order statistics. We do not have clear explanation for this.

- With this in mind, we performed simulation studies for $k = 50\%$ and $k = 70\%$. We noticed then (see Table 2) that, as opposed to the extremal index, the weaker dependence ($\rho = 0.5$) yields a good estimation for any given block size, while for the strong dependence ($\rho = 0.9$), the simulation results are rather poor. This may be quite intuitive, since the stop-loss functional is based on *sums* of large values. In any case, both sliding and disjoint blocks estimators yield comparable results.

In summary, **in case of the stop-loss index, both disjoint and sliding blocks estimators yield similar results (as suggested by theory). Weaker dependence implies better performance. One should use a high number of order statistics.**

6.2. Stationary ARCH process

We consider a stationary ARCH(1) process defined by $X_j^2 = \sqrt{\beta + \lambda X_{j-1}^2} Z_j$, where $\{Z_j, j \in \mathbb{Z}\}$ are i.i.d standard normal random variables. For $\lambda = 0.9$ the extremal index is $\theta = 0.612$ (see [EKM97, p. 480]).
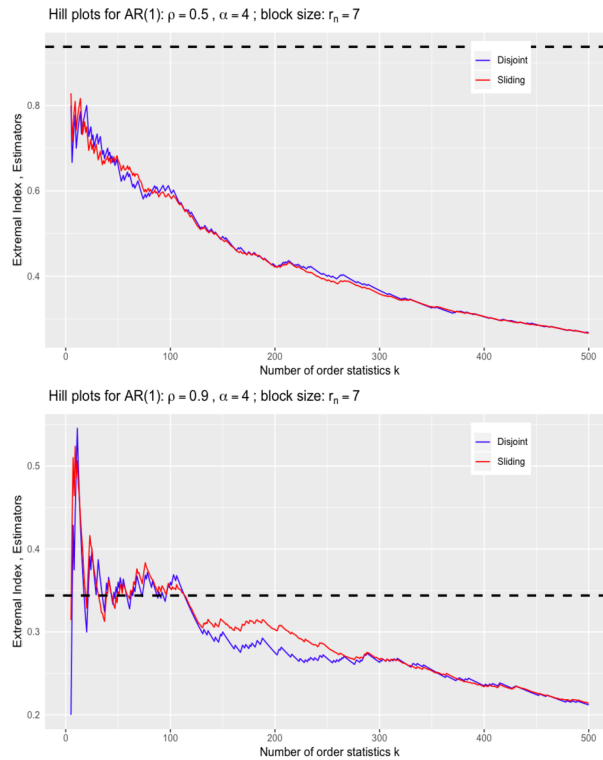
FIG 1. *Hill plots of sliding and disjoint blocks estimators for the extremal index plotted against the number of order statistics k. Data are simulated from AR(1) with $\rho = 0.5$ and $\theta = 0.94$ (left panel), $\rho = 0.9$ and $\theta = 0.34$ (right panel), $\alpha = 4$, block size $r_n = 7$. Dotted lines indicated the true value of the extremal index.*

- We start with Hill plots in Figure 5. The plots again illustrate little difference between the disjoint and sliding blocks estimators.
- Monte Carlo results are included in Table 3 and visualised as boxplots in Figure 6.

Again, **both disjoint and sliding blocks estimators yield similar results (as suggested by the theory)**.

6.3. Other cluster indices

We conducted simulation studies for other cluster indices, in particular for those from class $\mathcal{B}$. Here the results are rather not very promising. One of the reasons is the following observation. Recall (see Remark 3.12) that unlike in class $\mathcal{A}$, the cluster indices for functional from class $\mathcal{B}$ can have arbitrary values. At the same time, the largest possible value of the disjoint blocks estimator is m_n/k_n.
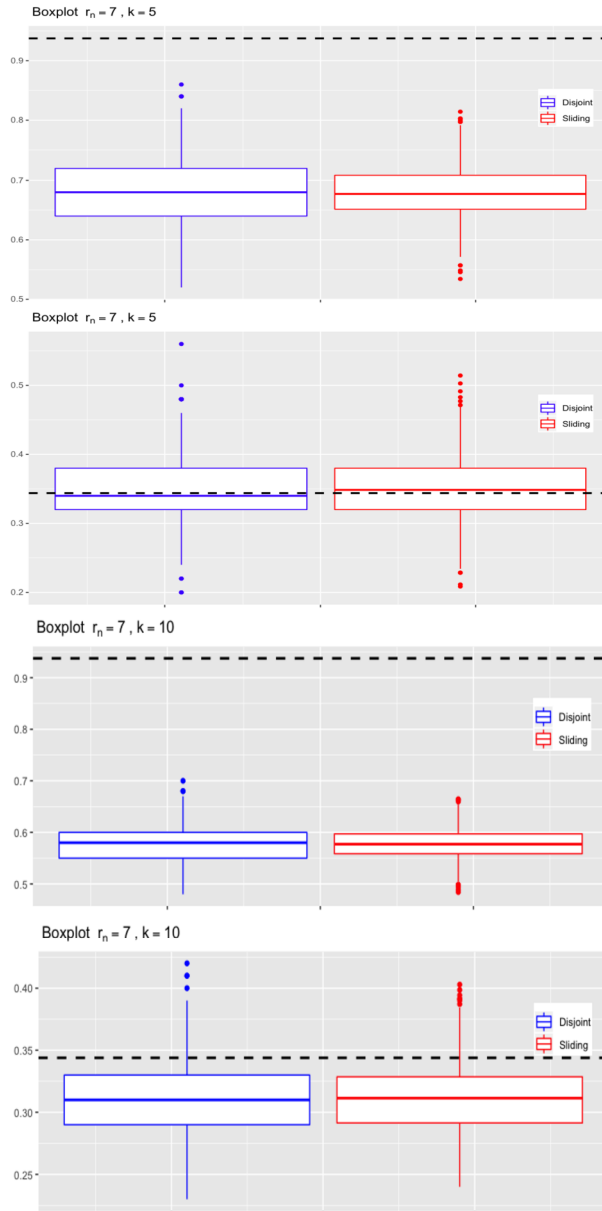
FIG 2. *Monte Carlo simulations of sliding and disjoint blocks estimators for the extremal index. Data are simulated from AR(1) with $\rho = 0.5$ and $\theta = 0.94$ (left panel), $\rho = 0.9$ and $\theta = 0.34$ (right panel), $\alpha = 4$, block size $r_n = 7$ and the number of order statistics $k = 5\%$ and 10%. Dotted lines indicated the true value of the extremal index.*

In theory, this ratio is

$$\frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$
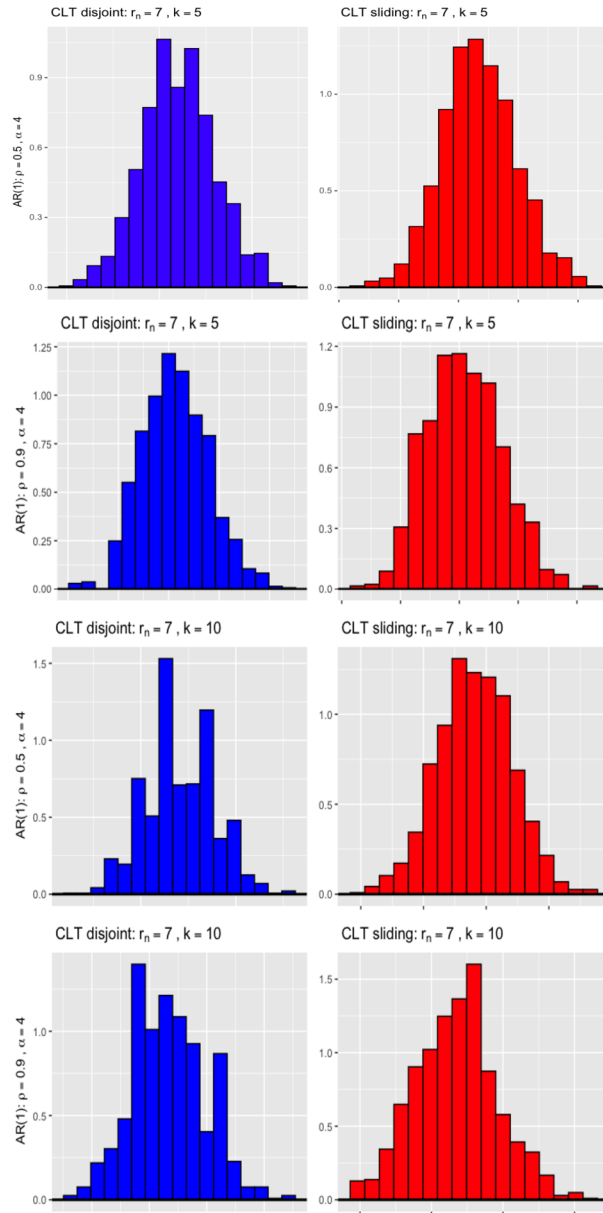
Fig 3. *Monte Carlo simulations of disjoint (blue) and sliding (red) blocks estimators for the extremal index. Data are simulated from AR(1) with $\rho = 0.5$ (left panel), $\rho = 0.9$ (right panel), $\alpha = 4$, block size $r_n = 7$ and the order statistics $k = 5\%$ and 10%.*

and hence diverges to infinity thanks to $\mathcal{R}(r_n, u_n)$. However, for finite samples, the ratio stays bigger than one for very few values of k_n. As such, we believe

FIG 4. *Hill plots of disjoint and sliding blocks estimators for the stop-loss index plotted against the number of order statistics k. Data are simulated from AR(1) with: $\rho = 0.5$, $S = 0.3$ (top left), $\rho = 0.5$, $S = 0.7$ (bottom left), $\rho = 0.5$, $S = 0.7$ (top right) and $\rho = 0.9$, $S = 0,7$, (bottom right), $\alpha = 4$, block size $r_n = 7$, 9.*

Fig 5. *Hill plots of disjoint and sliding blocks estimators for extremal index plotted against the number of order statistics k. Data are simulated from ARCH(1) with:* $\lambda = 0.9$.
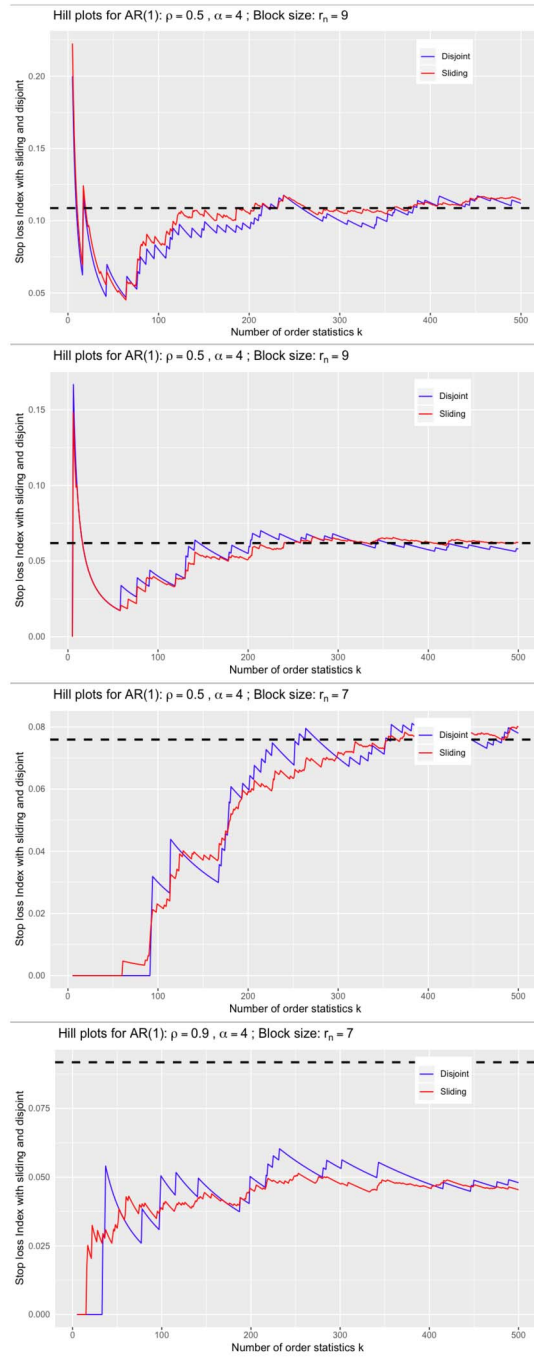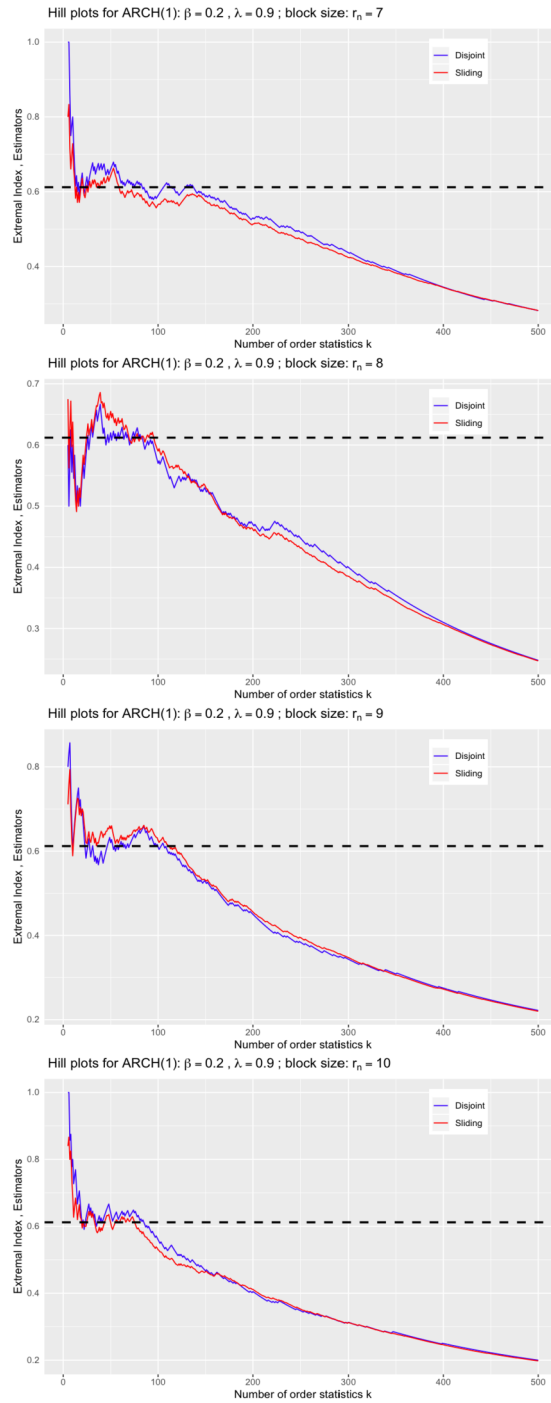
TABLE 2

The median and the variance (in brackets) of disjoint and sliding blocks estimators for stop-loss index with $S = 0.7$. Data are simulated from AR(1) with $\alpha = 4$, $\rho = 0.5$, $\rho = 0.9$. The block size is $r_n = 7$, 8, 9, 10. The number of order statistics is $k = 50\%$ and 70% for a sample $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

	$\rho = 0.9$, Stop-loss Index=0.096				$\rho = 0.5$, Stop-loss Index= 0.11			
(k %)	$k = 50$		$k = 70$		$k = 50$		$k = 70$	
$r_n = 7$								
Disjoint bl	**0.0320**	(0.009)	**0.0414**	(0.008)	**0.0960**	(0.009)	**0.1071**	(0.007)
Sliding bl	**0.0314**	(0.008)	**0.0418**	(0.008)	**0.0954**	(0.008)	**0.1067**	(0.006)
$r_n = 8$								
Disjoint bl	**0.0340**	(0.008)	**0.0443**	(0.007)	**0.0980**	(0.009)	**0.1071**	(0.006)
Sliding bl	**0.0332**	(0.008)	**0.0436**	(0.007)	**0.0968**	(0.007)	**0.1055**	(0.005)
$r_n = 9$								
Disjoint bl	**0.0360**	(0.008)	**0.0457**	(0.007)	**0.0960**	(0.008)	**0.1043**	(0.006)
Sliding bl	**0.0349**	(0.007)	**0.0460**	(0.007)	**0.0967**	(0.007)	**0.1038**	(0.004)
$r_n = 10$								
Disjoint bl	**0.0320**	(0.007)	**0.0429**	(0.007)	**0.0980**	(0.007)	**0.1014**	(0.005)
Sliding bl	**0.0314**	(0.007)	**0.0424**	(0.006)	**0.0974**	(0.006)	**0.1009**	(0.004)

that alternative methods of estimation of e.g. the large deviation index have to be implemented.

7. Proofs

In Section 7.1 we prove the equivalence between (2.3) and (2.4). In Section 7.3 we show that (1.1) holds for $H \in \mathcal{A} \cup \mathcal{B}$. The proofs in that section stem from [KS20]. The results from Section 7.3 are extended in Section 7.4 to covariance of clusters. In Section 7.5 we introduce the empirical process of sliding blocks and state its functional convergence. The proof of the latter is separated into several parts. First, in Section 7.6 we derive the limiting covariance of the empirical process of sliding blocks. Next, in Section 7.7 we prove the finite-dimensional convergence. Asymptotic continuity is dealt with in Section 7.8. We conclude the proof in Section 7.9.

7.1. Representations of the (candidate) extremal index

We first quote the time-change formula (see [BS09], [KS20, Theorem 5.3.1]). Let B be the backshift operator on $(\mathbb{R}^d)^{\mathbb{Z}}$ defined by $(B\boldsymbol{x})_i = \boldsymbol{x}_{i-1}$, $i \in \mathbb{Z}$.

TABLE 3
The median and the variance (in brackets) of disjoint and sliding blocks estimators for the extremal index in ARCH(1) model with $\lambda = 0.9$. The block size is $r_n = 7,\ 8,\ 9,\ 10$. The number of order statistics is $k = 5\%$ and 10% for a sample $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

| (k %) | Extremal Index=0.612 | | | |
	$k = 5$		$k = 10$	
$r_n = 7$				
Disjoint bl	**0.680**	(0.06)	**0.620**	(0.04)
Sliding bl	**0.670**	(0.06)	**0.620**	(0.03)
$r_n = 8$				
Disjoint bl	**0.660**	(0.06)	**0.600**	(0.04)
Sliding bl	**0.648**	(0.06)	**0.593**	(0.03)
$r_n = 9$				
Disjoint bl	**0.640**	(0.06)	**0.590**	(0.04)
Sliding bl	**0.631**	(0.05)	**0.567**	(0.03)
$r_n = 10$				
Disjoint bl	**0.620**	(0.06)	**0.550**	(0.04)
Sliding bl	**0.616**	(0.05)	**0.546**	(0.03)
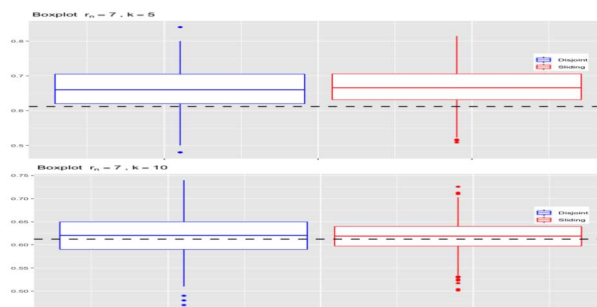
FIG 6. *Monte Carlo simulations of sliding and disjoint blocks estimators for the extremal index. Data are simulated from ARCH(1) with $\lambda = 0.9$, block size $r_n = 7$ and the number of order statistics $k = 5\%$ (left panel) and 10% (right panel). Dotted lines indicated the true value of the extremal index.*

Lemma 7.1. *Let $\mathbf{Y}$ be the tail process. Let H be a bounded or non-negative measurable functional on $(\mathbb{R}^d)^{\mathbb{Z}}$. Then for all $i \in \mathbb{Z}$,*

$$\mathbb{E}[H(B^i \mathbf{Y})\mathbb{1}\{|\mathbf{Y}_{-i}| > 1\}] = \mathbb{E}[H(\mathbf{Y})\mathbb{1}\{|\mathbf{Y}_i| > 1\}].$$

Lemma 7.2. *Assume that* $\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$. *Then*

$$\vartheta = \mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) = 0) = \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \leq 1) .$$

Proof. Recall that $\mathcal{C}_0(\boldsymbol{y}) = \inf\{j : \boldsymbol{y}^*_{-\infty,j} = \boldsymbol{y}^*\}$. From the definition of the tail process we have $\boldsymbol{Y}^* > 1$. Thus

$$\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) = 0) = \mathbb{P}\left(\sup_{j \leq -1} |\boldsymbol{Y}_j| < |\boldsymbol{Y}_0|, \sup_{j \geq 1} |\boldsymbol{Y}_j| \leq |\boldsymbol{Y}_0|, \boldsymbol{Y}^* > 1\right) .$$

By the assumption $\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$, the maximum $\boldsymbol{Y}^*$ is achieved at some $|\boldsymbol{Y}_i|$, $i \in \mathbb{Z}$. We split the event $\{\boldsymbol{Y}^* > 1\}$ according to the first exceedence over 1 and then apply the time-change formula:

$$\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) = 0) = \sum_{i \in \mathbb{Z}} \mathbb{P}\left(\boldsymbol{Y}^*_{-\infty,-1} < |\boldsymbol{Y}_0|, \boldsymbol{Y}^*_{1,\infty} \leq |\boldsymbol{Y}_0|, \boldsymbol{Y}^*_{-\infty,i-1} \leq 1, |\boldsymbol{Y}_i| > 1\right)$$

$$= \sum_{i \in \mathbb{Z}} \mathbb{P}\left(\boldsymbol{Y}^*_{-\infty,-i-1} < |\boldsymbol{Y}_{-i}|, \boldsymbol{Y}^*_{-i+1,\infty} \leq |\boldsymbol{Y}_{-i}|, \boldsymbol{Y}^*_{-\infty,-1} \leq 1, |\boldsymbol{Y}_{-i}| > 1\right)$$

$$= \sum_{i \in \mathbb{Z}} \mathbb{P}(C_i \cap \{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\})$$

with
$$C_i = \{\boldsymbol{Y}^*_{-\infty,-i-1} < |\boldsymbol{Y}_{-i}|, \boldsymbol{Y}^*_{-i+1,\infty} \leq |\boldsymbol{Y}_{-i}|, |\boldsymbol{Y}_{-i}| > 1\} .$$

The events C_i are disjoint, their union gives $\{\boldsymbol{Y}^* > 1\}$ and the latter event holds with probability one. Thus

$$\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) = 0) = \mathbb{P}(\{\boldsymbol{Y}^* > 1\} \cap \{\boldsymbol{Y}^*_{-\infty,-1} \leq 1\}) = \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \leq 1) . \qquad \square$$

The next lemma shows that there are other possible representations for ϑ.

Lemma 7.3. *Assume that* $\mathbb{P}(\mathcal{C}_0(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$. *Then*

$$\vartheta = \mathbb{P}(\boldsymbol{Y}^*_{1,\infty} \leq 1) .$$

Proof. We use again $\boldsymbol{Y}^* > 1$ and the fact that the maximum $\boldsymbol{Y}^*$ is achieved at some $|\boldsymbol{Y}_i|$. This time we split the event $\{\boldsymbol{Y}^* > 1\}$ according to the last exceedence over 1, and then apply the time-change formula:

$$\mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \leq 1) = \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \leq 1, \boldsymbol{Y}^* > 1)$$

$$= \sum_{i \in \mathbb{Z}} \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-1} \leq 1, |\boldsymbol{Y}_i| > 1, \boldsymbol{Y}^*_{i+1,\infty} \leq 1)$$

$$= \sum_{i \in \mathbb{Z}} \mathbb{P}(\boldsymbol{Y}^*_{-\infty,-i-1} \leq 1, |\boldsymbol{Y}_{-i}| > 1, \boldsymbol{Y}^*_{1,\infty} \leq 1)$$

$$= \mathbb{P}(\boldsymbol{Y}^* > 1, \boldsymbol{Y}^*_{1,\infty} \leq 1) = \mathbb{P}(\boldsymbol{Y}^*_{1,\infty} \leq 1) . \qquad \square$$

In fact, we can replace $\mathcal{C}_0$ with any anchoring map; see [KS20, Theorem 5.5.3] for more details.

7.2. Consequences of the mixing assumption

Since ℓ_n can be chosen as $r_n^{1-\delta}$ ($\delta > 0$) with δ arbitrarily close to zero, $\beta(r_n)$ gives:

$$\lim_{n\to\infty} \frac{n}{r_n}\beta_{r_n} = 0 \, , \tag{7.1a}$$

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=\ell_n}^{\infty} \beta_j = 0 \, , \tag{7.1b}$$

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{\infty} \beta_{jr_n} = 0 \, . \tag{7.1c}$$

We recall the covariance inequality for bounded, beta-mixing random variables (in fact, the inequality holds for α-mixing). Let $\beta(\mathcal{F}_1, \mathcal{F}_2)$ be the β-mixing coefficient between two sigma fields. Then ([Ibr62])

$$|\mathrm{cov}(H(Z_1), H(Z_2))| \le \mathrm{cst} \ \|H\|_\infty \|\widetilde{H}\|_\infty \beta(\sigma(Z_1), \sigma(Z_2)) \, . \tag{7.2}$$

In (7.2) the constant cst does not depend on $H, \widetilde{H}$.

7.3. Convergence of cluster measure

Proof of Proposition 3.4. Since H has a support separated from zero, there exists $\epsilon > 0$ such that $H(\boldsymbol{x}) = 0$ if $\boldsymbol{x}^* \le \epsilon$. Applying its shift invariance and the stationarity, we obtain

$$\boldsymbol{\nu}_{n,r_n}^*(H)$$

$$= \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{r_n} \mathbb{E}\left[H(u_n^{-1}\boldsymbol{X}_{1,r_n})\mathbb{1}\big\{\boldsymbol{X}_{1,j-1}^* \le u_n\epsilon\big\}\mathbb{1}\big\{|\boldsymbol{X}_j| > u_n\epsilon\big\}\right]$$

$$= \frac{\mathbb{P}(|\boldsymbol{X}_0| > u_n\epsilon)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \frac{1}{r_n} \sum_{j=1}^{r_n} \mathbb{E}\left[H(u_n^{-1}\boldsymbol{X}_{1-j,r_n-j})\mathbb{1}\big\{\boldsymbol{X}_{1-j,-1}^* \le u_n\epsilon\big\} \mid |\boldsymbol{X}_0| > u_n\epsilon\right]$$

$$= \frac{\mathbb{P}(|\boldsymbol{X}_0| > u_n\epsilon)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \int_0^1 g_n(v)\mathrm{d}v \, ,$$

with

$$g_n(v) = \mathbb{E}\left[H(u_n^{-1}\boldsymbol{X}_{1-[r_nv],r_n-[r_nv]})\mathbb{1}\big\{\boldsymbol{X}_{1-[r_nv],-1}^* \le u_n\epsilon\big\} \mid |\boldsymbol{X}_0| > u_n\epsilon\right] \, .$$

By Proposition 3.2, $\lim_{n\to\infty} g_n(v) = \mathbb{E}[H(\epsilon\boldsymbol{Y})\mathbb{1}\big\{\boldsymbol{Y}_{-\infty,-1}^* \le 1\big\}]$ for each $v \in (0,1)$. Moreover, the sequence g_n is uniformly bounded, thus by dominated convergence, regular variation of $|\boldsymbol{X}_0|$ and (2.7), we obtain

$$\lim_{n\to\infty} \boldsymbol{\nu}_{n,r_n}^*(H) = \epsilon^{-\alpha}\mathbb{E}[H(\epsilon\boldsymbol{Y})\mathbb{1}\big\{\boldsymbol{Y}_{-\infty,-1}^* \le 1\big\}] = \boldsymbol{\nu}^*(H) \, . \qquad \square$$

Proof of Lemma 3.8. By Proposition 3.4 and (2.6), we have

$$\lim_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\, \mathbb{1}\{|\boldsymbol{X}_j| > \epsilon u_n\} > u_n)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$= \vartheta \int_0^\infty \mathbb{P}\left(z\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\, \mathbb{1}\{z\,|\boldsymbol{Q}_j| > \epsilon\} > 1 \right) \alpha z^{-\alpha-1}\mathrm{d}z\,. \quad (7.3)$$

By monotone convergence, the right hand side converges to $\vartheta\mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\right)^\alpha\right]$ as $\epsilon \to 0$.

Consider the function

$$g(\zeta) = \vartheta \int_0^\infty \mathbb{P}\left(z\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\mathbb{1}\{z|\boldsymbol{Q}_j| > \zeta\} > 1 \right) \alpha z^{-\alpha-1}\mathrm{d}z\,.$$

It increases when ζ decreases to zero and its limit is $\vartheta\mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\right)^\alpha\right]$. To prove that this quantity is finite, it suffices to prove that the function g is bounded. Fix $\epsilon > 0$ and $\eta \in (0,1)$. By ANSJB(r_n, u_n), there exists ζ such that

$$\limsup_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\, \mathbb{1}\{|\boldsymbol{X}_j| \le \zeta u_n\} > \eta u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \le \epsilon\,.$$

Fix $\zeta' < \zeta$. Starting from (7.3) and applying ANSJB(r_n, u_n), we obtain

$$0 \le g(\zeta') = \vartheta \int_0^\infty \mathbb{P}\left(z\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\mathbb{1}\{z|\boldsymbol{Q}_j| > \zeta'\} > 1 \right) \alpha z^{-\alpha-1}\mathrm{d}z$$

$$= \lim_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\mathbb{1}\{|\boldsymbol{X}_j| > u_n\zeta'\} > u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$= \lim_{n\to\infty} \frac{\mathbb{P}\left(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\mathbb{1}\{|\boldsymbol{X}_j| > u_n\zeta\} + |\boldsymbol{X}_j|\mathbb{1}\{u_n\zeta \ge |\boldsymbol{X}_j| > \epsilon u_n\zeta'\} > u_n \right)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\le \limsup_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\, \mathbb{1}\{|\boldsymbol{X}_j| \le u_n\zeta\} > \eta u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$+ \lim_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\, \mathbb{1}\{|\boldsymbol{X}_i| > u_n\zeta\} > (1-\eta)u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\le \epsilon + \vartheta \int_0^\infty \mathbb{P}\left(z\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\, \mathbb{1}\{z\,|\boldsymbol{Q}_j| > \zeta\} > 1 - \eta \right) \alpha z^{-\alpha-1}\mathrm{d}z \le \epsilon + \vartheta\zeta^{-\alpha}\,.$$

The latter bound holds since the probability inside the integral is zero if $z \le \zeta$ since $|\boldsymbol{Q}_j| \le 1$ for all j. This proves that the function g is bounded in a neighbourhood of zero as claimed.

By Condition ANSJB(r_n, u_n), we finally obtain

$$\lim_{n\to\infty} \frac{\mathbb{P}\left(\sum_{j=1}^{r_n} |\boldsymbol{X}_j| > u_n\right)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \lim_{\epsilon\to 0}\lim_{n\to\infty} \frac{\mathbb{P}\left(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\mathbb{1}\{|\boldsymbol{X}_j| > u_n\epsilon\} > u_n\right)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$= \lim_{\epsilon\to 0} \vartheta \int_0^\infty \mathbb{P}\left(z\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\mathbb{1}\{z|\boldsymbol{Q}_j| > \epsilon\} > \epsilon\right)\alpha z^{-\alpha-1}\mathrm{d}z = \vartheta\mathbb{E}\left[\left(\sum_{j\in\mathbb{Z}} |\boldsymbol{Q}_j|\right)^\alpha\right].$$

$\square$

Proof of Proposition 3.9. For $\epsilon > 0$, we define the truncation operator T_ϵ by

$$T_\epsilon(\boldsymbol{x}) = \{\boldsymbol{x}_j\mathbb{1}_{\{|\boldsymbol{x}_j|>\epsilon\}}, j \in \mathbb{Z}\} . \tag{7.4}$$

The operator T_ϵ is continuous with respect to the uniform norm at every $\boldsymbol{x} \in \ell_0$ such that $|\boldsymbol{x}_j| \neq \epsilon$ for all $j \in \mathbb{Z}$.

Fix $\eta \in (0,1)$ and $\zeta > 0$. Let L_K be as in (2.1) and choose $\epsilon > 0$ such that

$$\limsup_{n\to\infty} \frac{\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\,\mathbb{1}\{|\boldsymbol{X}_j| \leq \epsilon u_n\} > \eta u_n/L_K)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \leq \zeta .$$

Set $K_\epsilon = K \circ T_\epsilon$. Applying assumption (2.1), we obtain

$$\frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\leq \frac{\mathbb{P}(K_\epsilon(\boldsymbol{X}_{1,r_n}/u_n) > 1-\eta)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} + \frac{\mathbb{P}(|K(\boldsymbol{X}_{1,r_n}/u_n) - K_\epsilon(\boldsymbol{X}_{1,r_n}/u_n)| > \eta)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\leq \frac{\mathbb{P}(K_\epsilon(\boldsymbol{X}_{1,r_n}/u_n) > 1-\eta)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} + \frac{\mathbb{P}(\sum_{i=1}^{r_n} |\boldsymbol{X}_j|\,\mathbb{1}\{|\boldsymbol{X}_j| \leq \epsilon u_n\} > \eta u_n/\mathrm{cst})}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} .$$

Applying Proposition 3.4 to K_ϵ along with the representation (2.6), this yields

$$\limsup_{n\to\infty} \frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \leq \limsup_{n\to\infty} \frac{\mathbb{P}(K_\epsilon(\boldsymbol{X}_{1,r_n}/u_n) > 1-\eta)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} + \zeta$$

$$= \int_0^\infty \mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > 1-\eta)\alpha z^{-\alpha-1}\mathrm{d}z + \zeta .$$

Similarly,

$$\liminf_{n\to\infty} \frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \geq \int_0^\infty \mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > 1+\eta)\alpha z^{-\alpha-1}\mathrm{d}z - \zeta .$$

Since $K(\boldsymbol{0}) = 0$, (2.1) implies that $|K(\boldsymbol{x})| \leq \mathrm{cst}\sum_{j\in\mathbb{Z}} |\boldsymbol{x}_j|$, thus for all $y > 0$,

$$\mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > y) \leq \mathbb{P}\left(\sum_{j\in\mathbb{Z}} z\,|\boldsymbol{Q}_j| > y/\mathrm{cst}\right)$$

and the latter quantity is integrable (as a function of z) with respect to $\alpha z^{-\alpha-1}\mathrm{d}z$ in view of ANSJB(r_n, u_n) and Lemma 3.8. By bounded convergence, this yields

$$\lim_{\epsilon \to 0} \int_0^\infty \mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > y)\alpha z^{-\alpha-1}\mathrm{d}z = \int_0^\infty \mathbb{P}(K(z\boldsymbol{Q}) > y)\alpha z^{-\alpha-1}\mathrm{d}z \ .$$

Altogether, we obtain

$$\int_0^\infty \mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > 1 + \eta)\alpha z^{-\alpha-1}\mathrm{d}z - \zeta \leq \liminf_{n\to\infty} \frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\leq \limsup_{n\to\infty} \frac{\mathbb{P}(K(\boldsymbol{X}_{1,r_n}/u_n) > 1)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \leq \int_0^\infty \mathbb{P}(K_\epsilon(z\boldsymbol{Q}) > 1 - \eta)\alpha z^{-\alpha-1}\mathrm{d}z + \zeta \ .$$

Since ζ and η are arbitrary, this finishes the proof. $\qquad\square$

7.4. Covariance of clusters

We consider the limit

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\mathrm{cov}\left(H(\boldsymbol{X}_{1,r_n}/u_n), \widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n)\right)$$

for different choices of h, possibly depending on n. Under the conditions of Proposition 3.4, if moreover $\mathcal{R}(r_n, u_n)$ holds, the above limit is the same as

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n)\right] \ .$$

Thus, we impose $\mathcal{R}(r_n, u_n)$ and switch freely between $\mathbb{E}$ and cov whenever suitable.

7.4.1. Uniform convergence of cluster measure

In Propositions 3.4 and 3.9 we proved (1.1) for $H \in \mathcal{A} \cup \mathcal{B}$. We note further that if (1.1) holds for H and $\widetilde{H}$, then it also holds for any linear combination of both functions. To deal with asymptotic normality, we need (1.1) to hold uniformly over a subclass of functions. With this in mind, we introduce two additional classes of functions. First, we recall that for a class $\mathcal{G}$ of functions $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ its envelope is

$$\boldsymbol{G}(\boldsymbol{x}) = \sup_{H \in \mathcal{G}} |H(\boldsymbol{x})| \ , \quad \boldsymbol{x} \in (\mathbb{R}^d)^{\mathbb{Z}} \ .$$

Definition 7.4. $\widetilde{\mathcal{A}} \subseteq \mathrm{span}(\mathcal{A})$ *(resp. $\widetilde{\mathcal{B}} \subseteq \mathrm{span}(\mathcal{B})$) is a class of functions with a finite envelope such that*

$$\lim_{n\to\infty} \sup_{H \in \widetilde{\mathcal{A}}} \boldsymbol{\nu}^*_{n,r_n}(|H|) < \infty \qquad (7.5)$$

(resp. $\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{B}}} \boldsymbol{\nu}^*_{n,r_n}(|H|) < \infty$) and that for each H there exist functions $K^H_n : (\mathbb{R}^d)^{\ell_n} \to \mathbb{R}_+$ such that

$$\left| H\left(\frac{\boldsymbol{X}_{1,r_n}}{u_n}\right) - H\left(\frac{\boldsymbol{X}_{1,r_n-\ell_n}}{u_n}\right) \right| \le K^H_n(\boldsymbol{X}_{r_n-\ell_n+1,r_n}) \,,$$

$$\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}} \frac{\mathbb{E}\left[K^H_n(\boldsymbol{X}_{1,\ell_n})\right]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = 0 \,. \tag{7.6}$$

Remark 7.5. The uniform convergence condition (7.5) strengthens the statement of Proposition 3.4. Conditions (7.5) and (7.6) are needed for asymptotic equicontinuity of empirical cluster process to be introduced below. $\qquad \oplus$

Remark 7.6. We note that

$$\lim_{n\to\infty} \frac{\mathbb{E}\left[K^H_n(\boldsymbol{X}_{1,\ell_n})\right]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = 0$$

for each $H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$. Let us verify it for $H \in \mathcal{B}$. We have

$$|\mathbb{1}\{K(\boldsymbol{x}_{1,r_n}) > 1\} - \mathbb{1}\{K(\boldsymbol{x}_{1,r_n-\ell_n}) > 1\}|$$
$$= \mathbb{1}\{K(\boldsymbol{x}_{1,r_n}) > 1\}\mathbb{1}\{K(\boldsymbol{x}_{1,r_n-\ell_n}) \le 1\} + \mathbb{1}\{K(\boldsymbol{x}_{1,r_n}) \le 1\}\mathbb{1}\{K(\boldsymbol{x}_{1,r_n-\ell_n}) > 1\} \,.$$

We consider the first pair of indicators in the last line. The events $\{K(\boldsymbol{x}_{1,r_n}) > 1\}$ and $\{K(\boldsymbol{x}_{1,r_n-\ell_n}) \le 1\}$ imply that there exists $s > 0$ such that $K(\boldsymbol{x}_{1,r_n}) - K(\boldsymbol{x}_{1,r_n-\ell_n}) > s$. Applying the same reasoning to the second pair of indicators, we have

$$|\mathbb{1}\{K(\boldsymbol{x}_{1,r_n}) > 1\} - \mathbb{1}\{K(\boldsymbol{x}_{1,r_n-\ell_n}) > 1\}| \le 2\mathbb{1}\left\{\text{cst} \sum_{j=r_n-\ell_n+1}^{r_n} |\boldsymbol{x}_j| > s\right\} \,.$$

Since $\ell_n = o(r_n)$

$$\mathbb{P}\left(\sum_{j=r_n-\ell_n+1}^{r_n} |\boldsymbol{X}_j| > su_n\right) = O(\ell_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)) = (r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)) \,.$$

In summary, (7.6) holds if the envelope function is in $\widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$. $\qquad \oplus$

Remark 7.7. Let $\delta > 0$. If H is bounded then

$$\boldsymbol{\nu}^*_{n,r_n}(|H|^{1+\delta}) \le \|H\|^\delta_\infty \boldsymbol{\nu}^*_{n,r_n}(|H|)$$

and by the assumptions on the classes $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$,

$$\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}\cup\widetilde{\mathcal{B}}} \boldsymbol{\nu}^*_{n,r_n}(|H|^{1+\delta}) < \infty \,.$$

$\qquad \oplus$

Remark 7.8. Assume that $\mathcal{AC}(r_n, u_n)$ holds. Fix $0 < s_0 < t_0 < \infty$. Let $H \in \mathcal{A}$ and recall that $H_s(\boldsymbol{x}) = H(\boldsymbol{x}/s)$. Assume that $\widetilde{\mathcal{A}} := \{H_s, s \in [s_0, t_0]\}$ is linearly ordered. Note that $\widetilde{\mathcal{A}} \subset \mathcal{A}$. The envelope is $|H_{s_0}| \vee |H_{t_0}| \in \widetilde{\mathcal{A}}$ hence (7.6) holds. Moreover, $\sup_{s \in [s_0, t_0]} \boldsymbol{\nu}^*_{n, r_n}(|H_s|)$ is achieved at s_0 or t_0. Likewise,

$$\lim_{n \to \infty} \sup_{s, t \in [s_0, t_0]} \boldsymbol{\nu}^*_{n, r_n}(|H_s - H_t|) < \infty \ .$$

The same applies to $H \in \mathcal{B}$ if additionally $\mathrm{ANSJB}(r_n, u_n)$ holds. $\qquad \oplus$

7.4.2. Conditional convergence

We consider conditional convergence of functions $H, \widetilde{H}$ acting on overlapping blocks.

Lemma 7.9. *Assume that $\mathcal{AC}(r_n, u_n)$ holds. Let $h < r_n$, $H, \widetilde{H} \in \mathcal{L}$ and $\widetilde{H}(\boldsymbol{0}) = 0$. Then*

$$\lim_{n \to \infty} \mathbb{E}[H(\boldsymbol{X}_{1, r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h, r_n+h}/u_n) \mid |\boldsymbol{X}_0| > u_n]$$
$$= \begin{cases} \mathbb{E}[H(\boldsymbol{Y}_{1, \infty})\widetilde{H}(\boldsymbol{Y}_{1+h, \infty})] \ , & \text{if } h \text{ fixed} \ , \\ 0 \ , & \text{if } h = h_n \to \infty \ . \end{cases}$$

and

$$\lim_{n \to \infty} \mathbb{E}[H(\boldsymbol{X}_{-r_n, r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{-r_n+h, r_n+h}/u_n) \mid |\boldsymbol{X}_0| > u_n] = \mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})] \ .$$

Proof. Since $H, \widetilde{H}$ are bounded, the first expectation of interest is dominated by

$$\|H\|_\infty \|\widetilde{H}\|_\infty \mathbb{P}(\boldsymbol{X}^*_{1+h, r_n+h} > u_n \mid |\boldsymbol{X}_0| > u_n) \ .$$

Thus, the statement for $h = h_n \to \infty$ follows immediately from $\mathcal{AC}(r_n, u_n)$ (cf. the argument in the proof of [KS20, Theorem 6.1.4]).

Now, let h be fixed. Fix r. Since $H, \widetilde{H}$ are bounded Lipschitz continuous, we have by Proposition 3.2,

$$\lim_{n \to \infty} \mathbb{E}[H(\boldsymbol{X}_{1, r}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h, r+h}/u_n) \mid |\boldsymbol{X}_0| > u_n] = \mathbb{E}[H(\boldsymbol{Y}_{1, r})\widetilde{H}(\boldsymbol{Y}_{1+h, r+h})] \ ,$$

$$\lim_{n \to \infty} \mathbb{E}[H(\boldsymbol{X}_{-r, r}/u_n)\widetilde{H}(\boldsymbol{X}_{-r+h, r+h}/u_n) \mid |\boldsymbol{X}_0| > u_n]$$
$$= \mathbb{E}[H(\boldsymbol{Y}_{-r, r})\widetilde{H}(\boldsymbol{Y}_{-r+h, r+h})] \ .$$

Since the tail process tends to zero under condition $\mathcal{AC}(r_n, u_n)$, it also holds that

$$\lim_{r \to \infty} \mathbb{E}[H(\boldsymbol{Y}_{1, r})\widetilde{H}(\boldsymbol{Y}_{1+h, r+h})] = \mathbb{E}[H(\boldsymbol{Y}_{1, \infty})\widetilde{H}(\boldsymbol{Y}_{1+h, \infty})] \ ,$$
$$\lim_{r \to \infty} \mathbb{E}[H(\boldsymbol{Y}_{-r, r})\widetilde{H}(\boldsymbol{Y}_{-r+h, r+h})] = \mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})] = \mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})] \ .$$

Indeed, considering the first statement only we have

$$
\begin{aligned}
&\lim_{r\to\infty} \left| \mathbb{E}[H(\boldsymbol{Y}_{1,r})\widetilde{H}(\boldsymbol{Y}_{1+h,r+h})] - \mathbb{E}[H(\boldsymbol{Y}_{1,\infty})\widetilde{H}(\boldsymbol{Y}_{1+h,\infty})] \right| \\
&\leq \lim_{r\to\infty} \mathbb{E}[|H(\boldsymbol{Y}_{1,r}) - H(\boldsymbol{Y}_{1,\infty})| \, |\widetilde{H}|(\boldsymbol{Y}_{1+h,r+h})] \\
&\quad + \lim_{r\to\infty} \mathbb{E}\left[|H|(\boldsymbol{Y}_{1,\infty}) \left| \widetilde{H}(\boldsymbol{Y}_{1+h,r+h}) - \widetilde{H}(\boldsymbol{Y}_{1+h,\infty}) \right| \right] \\
&\leq \|\widetilde{H}\|_\infty \lim_{r\to\infty} \left\{ \mathbb{E}\left[|H(\boldsymbol{Y}_{1,r}) - H(\boldsymbol{Y}_{1,\infty})|\right] + \mathbb{E}\left[\left| \widetilde{H}(\boldsymbol{Y}_{1+h,r+h}) - \widetilde{H}(\boldsymbol{Y}_{1+h,\infty}) \right| \right] \right\}
\end{aligned}
$$

and the limit is zero. To conclude, we only need to apply the triangular argument, that is to prove that

$$
\lim_{r\to\infty} \limsup_{n\to\infty} \\
\left| \mathbb{E}[H(\boldsymbol{X}_{1,r}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r+h}/u_n) - H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n) \mid |\boldsymbol{X}_0| > u_n] \right|
$$

is zero. Using again the fact that $H, \widetilde{H}$ are bounded, the conditional expectation is dominated by

$$
\begin{aligned}
&\|\widetilde{H}\|_\infty \left| \mathbb{E}[H(\boldsymbol{X}_{1,r}/u_n) - H(\boldsymbol{X}_{1,r_n}/u_n) \mid |\boldsymbol{X}_0| > u_n] \right| \\
&+ \|H\|_\infty \left| \mathbb{E}[\widetilde{H}(\boldsymbol{X}_{1+h,r+h}/u_n) - \widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n) \mid |\boldsymbol{X}_0| > u_n] \right| . \qquad (7.7)
\end{aligned}
$$

Fix $\epsilon > 0$. Since H is Lipschitz continuous, applying condition $\mathcal{AC}(r_n, u_n)$ yields

$$
\begin{aligned}
&\lim_{r\to\infty} \limsup_{n\to\infty} \left| \mathbb{E}[H(\boldsymbol{X}_{1,r}/u_n) - H(\boldsymbol{X}_{1,r_n}/u_n) \mid |\boldsymbol{X}_0| > u_n] \right| \\
&\qquad\qquad \leq \mathrm{cst} \left\{ \epsilon + \lim_{r\to\infty} \limsup_{n\to\infty} \mathbb{P}(\boldsymbol{X}_{r,r_n}^* > \epsilon u_n \mid |\boldsymbol{X}_0| > u_n) \right\} = \mathrm{cst} \times \epsilon .
\end{aligned}
$$

The same argument applies to (7.7). Since ϵ is arbitrary, this concludes the proof. $\square$

7.4.3. Covariance of clusters: Disjoint blocks

The first result is straightforward under the beta-mixing conditions.

Lemma 7.10 (Disjoint blocks I). *Assume that* $\mathcal{AC}(r_n, u_n)$, $\mathcal{R}(r_n, u_n)$, (7.1b) *hold. Then*

$$
\lim_{n\to\infty} \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sup \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+[\xi' r_n], r_n+[\xi' r_n]}/u_n) \right] = 0 ,
$$

where the sup *is taken over* $\xi' > 1$ *and* $H, \widetilde{H} \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$.

Proof of Lemma 7.10. Let $H, \widetilde{H} \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$. Then, using (7.6),

$$
\left| \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+[\xi' r_n], r_n+[\xi' r_n]}/u_n) \right] \right|
$$

$$\leq \left| \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n-\ell_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+[\xi'r_n],r_n+[\xi'r_n]}/u_n) \right] \right| + \|\widetilde{H}\|_\infty \mathbb{E}\left[K_n^H(\boldsymbol{X}_{1,\ell_n}) \right]$$

and the latter term is $o(r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n))$, uniformly over $\widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$ by the assumption.

Using (7.2) and (7.1b), we have

$$\frac{\left| \operatorname{cov}\left(H(\boldsymbol{X}_{1,r_n-\ell_n}/u_n), \widetilde{H}(\boldsymbol{X}_{1+[\xi'r_n],r_n+[\xi'r_n]}/u_n) \right) \right|}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$\leq \|H\|_\infty \|\widetilde{H}\|_\infty \frac{\beta_{\ell_n+[(\xi'-1)r_n]}}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

and the latter is $o(1)$ uniformly over the class of functions. $\square$

We extend the above result to the excess functional $\mathcal{E}_s(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{|\boldsymbol{x}_j| > s\}$.

Lemma 7.11 (Disjoint blocks II). *Assume that* $\mathcal{AC}(r_n, u_n)$, $\mathcal{R}(r_n, u_n)$, (7.1b) *hold. Then*

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sup \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\mathcal{E}_s(\boldsymbol{X}_{1+[\xi'r_n],r_n+[\xi'r_n]}/u_n) \right] = 0 \ ,$$

where the sup *is taken over* $\xi' > 1$, $H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$ *and* $s \in [s_0, t_0]$.

Proof. Recall that $\ell_n = o(r_n)$. Split the sum $\sum_{j=[\xi'r_n]+1}^{r_n+[\xi'r_n]}$ into $\sum_{j=[\xi'r_n]+1}^{r_n+\ell_n}$ and $\sum_{j=r_n+\ell_n+1}^{r_n+[\xi'r_n]}$.

For the first sum we have

$$\frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=[\xi'r_n]+1}^{r_n+\ell_n} \left| \mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{|\boldsymbol{X}_j| > u_ns\}] \right|$$

$$\leq \|H\|_\infty \frac{\ell_n\mathbb{P}(|\boldsymbol{X}_0| > u_ns)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = o(1)$$

uniformly over the class of functions and over s.

Using (7.2) we have

$$\frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=r_n+\ell_n+1}^{r_n+[\xi'r_n]} \left| \operatorname{cov}(H(\boldsymbol{X}_{1,r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_j| > u_ns\}) \right|$$

$$\leq \frac{\|H\|_\infty}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=r_n+\ell_n+1}^{\infty} \beta_{j-r_n} \ .$$

We finish the proof using the mixing assumption (7.1b). $\square$

7.4.4. Covariance of clusters: Overlapping blocks

We consider three cases separately: a) $H, \widetilde{H} \in \mathcal{A}$ (Proposition 7.12); b) $H, \widetilde{H} \in \mathcal{B}$ (Proposition 7.13); c) the excess functional (Proposition 7.14).

Proposition 7.12 (Overlapping blocks I)**.** *Assume that* $\mathcal{AC}(r_n, u_n)$ *and* $\mathcal{R}(r_n, u_n)$ *hold. Let* $h < r_n$ *and* $\xi \in (0, 1)$*. For* $H, \widetilde{H} \in \mathcal{A}$ *we have*

$$\lim_{n \to \infty} \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n)\right]$$
$$= \begin{cases} \boldsymbol{\nu}^*(H\widetilde{H})\,, & \text{if } h/r_n \to 0\,, \\ (1-\xi)\boldsymbol{\nu}^*(H\widetilde{H})\,, & \text{if } h = h_n = [\xi r_n] \end{cases} \quad . \tag{7.8}$$

Proof of Proposition 7.12. Note that if $H, \widetilde{H} \in \mathcal{A}$, then $H\widetilde{H} \in \mathcal{A}$ (but it does not mean that we can apply Lemma 7.9 since here the functions are applied to different blocks).

Since $H, \widetilde{H}$ vanish around $\boldsymbol{0}$, there exists $\epsilon > 0$ such that $H(\boldsymbol{x}_{1,r_n})\widetilde{H}(\boldsymbol{x}_{1,r_n}) = 0$ whenever $\boldsymbol{x}_{1,r_n}^* < \epsilon$. Then, splitting the event $\{\boldsymbol{X}_{1,r_n}^* > u_n\}$ and using stationarity we write the expression of interest as

$$\frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n)\mathbb{1}\{\boldsymbol{X}_{1,r_n}^* > u_n\epsilon\}\right]$$
$$= \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$
$$\times \sum_{j=1}^{r_n} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h,r_n+h}/u_n)\mathbb{1}\{\boldsymbol{X}_{1,j-1}^* \le u_n\epsilon\}\mathbb{1}\{|\boldsymbol{X}_j| > u_n\epsilon\}\right]$$
$$= \frac{1}{r_n} \frac{\mathbb{P}(|\boldsymbol{X}_0| > \epsilon u_n)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$
$$\times \sum_{j=1}^{r_n} \mathbb{E}\left[H(\boldsymbol{X}_{1-j,r_n-j}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h-j,r_n+h-j}/u_n)\mathbb{1}\{\boldsymbol{X}_{1-j,-1}^* \le u_n\epsilon\} \mid |\boldsymbol{X}_0|\right.$$
$$\left. > u_n\epsilon\right]\,.$$

To shorten the exposition, let $A_j = \{\boldsymbol{X}_{1-j,-1}^* \le u_n\epsilon\}$. We write the last expression as $\int_0^1 g_n(v)\mathrm{d}v$ with $g_n(v)$ defined by

$$\mathbb{E}[H(\boldsymbol{X}_{1-[r_n v],r_n-[r_n v]}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h-[r_n v],r_n+h-[r_n v]}/u_n)\mathbb{1}\{A_{[r_n v]}\} \mid |\boldsymbol{X}_0| > u_n\epsilon]\,.$$

If $h = o(r_n)$, then using the second part of Lemma 7.9 we get

$$\lim_{n \to \infty} g_n(v) = \epsilon^{-\alpha} \mathbb{E}[H(\epsilon\boldsymbol{Y})\widetilde{H}(\epsilon\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}_{-\infty,-1}^* \le 1\}] = \boldsymbol{\nu}^*(H\widetilde{H})$$

independently of $v \in (0, 1)$. If $h = [\xi r_n]$, $\xi \in (0, 1)$, then we split the integral.

If $\xi > v$, then we use boundedness of both $H, \widetilde{H}$ and the fact that $\widetilde{H}$ vanishes around $\mathbf{0}$. Thanks to the anticlustering condition $\mathcal{AC}(r_n, u_n)$, we have as $n \to \infty$,

$$\mathbb{E}[H(\boldsymbol{X}_{1-[r_nv], r_n-[r_nv]}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h-[r_nv], r_n+h-[r_nv]}/u_n)\mathbb{1}\{A_{[r_nv]}\} \mid |\boldsymbol{X}_0| > u_n\epsilon]$$

$$\leq \mathrm{cst}\mathbb{P}\left(\boldsymbol{X}^*_{1+[\xi r_n]-[r_nv], r_n+[\xi r_n]-[r_nv]} > u_n\epsilon \mid |\boldsymbol{X}_0| > u_n\epsilon\right)$$

$$\leq \mathrm{cst}\mathbb{P}\left(\boldsymbol{X}^*_{[r_n(\xi-v)], 3r_n} > u_n\epsilon \mid |\boldsymbol{X}_0| > u_n\epsilon\right) \to 0 .$$

If $\xi \leq v$, then we apply the second part of Lemma 7.9:

$$\mathbb{E}[H(\boldsymbol{X}_{1-[r_nv], r_n-[r_nv]}/u_n)\widetilde{H}(\boldsymbol{X}_{1+h-[r_nv], r_n+h-[r_nv]}/u_n)\mathbb{1}\{A_{[r_nv]}\} \mid |\boldsymbol{X}_0| > u_n\epsilon]$$

$$\to \epsilon^{-\alpha}\mathbb{E}[H(\epsilon\boldsymbol{Y})\widetilde{H}(\epsilon\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty, -1} \leq 1\}] = \boldsymbol{\nu}^*(H\widetilde{H}) , \quad n \to \infty .$$

Since the sequence $\{g_n\}$ is uniformly bounded, we have

$$\lim_{n\to\infty} \int_0^\xi g_n(v)\mathrm{d}v + \lim_{n\to\infty} \int_\xi^1 g_n(v)\mathrm{d}v = 0 + (1-\xi)\boldsymbol{\nu}^*(H\widetilde{H}) . \qquad \square$$

Proposition 7.13 (Overlapping blocks II). *Assume that* $\mathcal{AC}(r_n, u_n)$, $\mathcal{R}(r_n, u_n)$ *and* $\mathrm{ANSJB}(r_n, u_n)$ *hold. Let* $h < r_n$ *and* $\xi \in (0, 1)$. *Then* (7.8) *holds for* $H, \widetilde{H} \in \mathcal{B}$.

Proof of Proposition 7.13. We mimic the proof of Proposition 3.9 (refer to that proof for the notation). Set $K^\epsilon = K \circ T^\epsilon$, $\widetilde{K}^\epsilon = \widetilde{K} \circ T^\epsilon$. Let $\eta \in (0, 1)$. Note that $H^\epsilon_\pm := \mathbb{1}\{K^\epsilon > 1 \pm \eta\} \in \mathcal{A}$, $\widetilde{H}^\epsilon_\pm := \mathbb{1}\{\widetilde{K}^\epsilon > 1 \pm \eta\} \in \mathcal{A}$ and hence $H^\epsilon \widetilde{H}^\epsilon \in \mathcal{A}$; see the comment at the beginning of the proof of Proposition 7.12.

Fix $\eta \in (0, 1)$ and $\zeta > 0$. Let $L_K, L_{\widetilde{K}}$ be as in (2.1) and choose $\epsilon > 0$ such that

$$\limsup_{n\to\infty} \frac{2\mathbb{P}(\sum_{j=1}^{r_n} |\boldsymbol{X}_j|\,\mathbb{1}\{|\boldsymbol{X}_j| \leq \epsilon u_n\} > \eta u_n/(L_K \vee L_{\widetilde{K}}))}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \leq \zeta .$$

This is allowed thanks to $\mathrm{ANSJB}(r_n, u_n)$. We have

$$\frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\mathbb{P}\left(K(\boldsymbol{X}_{1,r_n}/u_n) > 1, \widetilde{K}(\boldsymbol{X}_{1+h, r_n+h}/u_n) > 1\right)$$

$$\leq \frac{2\mathbb{P}(\sum_{i=1}^{r_n} |\boldsymbol{X}_j|\,\mathbb{1}\{|\boldsymbol{X}_j| \leq \epsilon u_n\} > \eta u_n/\mathrm{cst})}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$+ \frac{\mathbb{E}[H^\epsilon_-(\boldsymbol{X}_{1,r_n}/u_n)\widetilde{H}^\epsilon_-(\boldsymbol{X}_{1+h, r_n+h}/u_n)]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} .$$

Application of Proposition 7.12 gives

$$\lim_{n\to\infty} \frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\mathbb{P}\left(K(\boldsymbol{X}_{1,r_n}/u_n) > 1, \widetilde{K}(\boldsymbol{X}_{1+h, r_n+h}/u_n) > 1\right)$$

$$\leq \zeta + \begin{cases} \boldsymbol{\nu}^*(H^\epsilon_-(\boldsymbol{Y})\widetilde{H}^\epsilon_-(\boldsymbol{Y})) , & \text{if } h/r_n \to 0 , \\ (1-\xi)\boldsymbol{\nu}^*(H^\epsilon_-(\boldsymbol{Y})\widetilde{H}^\epsilon_-(\boldsymbol{Y})) , & \text{if } h = h_n = [\xi r_n] \end{cases} .$$

Similarly, we obtain the lower bound with $H^\epsilon_+, \widetilde{H}^\epsilon_+$ instead of $H^\epsilon_-, \widetilde{H}^\epsilon_-$ and $-\zeta$ instead of $+\zeta$. Since ζ is arbitrary, the proof is concluded by letting $\epsilon \to 0$. This follows the same argument as in the proof of Proposition 3.9. $\qquad\square$

Proposition 7.14 (Overlapping blocks III)**.** *Assume that* $\mathcal{AC}(r_n, u_n)$, $\mathcal{R}(r_n, u_n)$ *hold. Let* $h < r_n$. *For* $H \in \mathcal{L}$ *we have*

$$\lim_{n\to\infty} \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\mathcal{E}_s(\boldsymbol{X}_{1+h,r_n+h}/u_n)\right]$$

$$= \begin{cases} s^{-\alpha}\mathbb{E}[H(s\boldsymbol{Y})] , & \text{if } h/r_n \to 0 , \\ s^{-\alpha}(1-\xi)\mathbb{E}[H(s\boldsymbol{Y})] , & \text{if } h = h_n = \xi r_n \end{cases} .$$

Proof of Proposition 7.14. We have for $h < r_n$,

$$\frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)\mathcal{E}_s(\boldsymbol{X}_{1+h,r_n+h}/u_n)\right]$$

$$= \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=h+1}^{r_n+h} \mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)\mathbb{1}\{|\boldsymbol{X}_j| > u_n s\}]$$

$$= \frac{1}{r_n} \frac{\mathbb{P}(|\boldsymbol{X}_0| > u_n s)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=h+1}^{r_n+h} \mathbb{E}[H(\boldsymbol{X}_{1-j,r_n-j}/u_n) \mid |\boldsymbol{X}_0| > u_n s] .$$

We write the last expression as

$$\frac{\mathbb{P}(|\boldsymbol{X}_0| > u_n s)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \int_{h/r_n}^{1+h/r_n} g_n(v)\mathrm{d}v$$

with (omitting the dependence on s)

$$g_n(v) = \mathbb{E}[H(s\boldsymbol{X}_{1-[r_n v],r_n-[r_n v]}/(u_n s)) \mid |\boldsymbol{X}_0| > u_n s] .$$

Since H is bounded, $\mathcal{AC}(r_n, u_n)$ and Proposition 3.2 give

$$\lim_{n\to\infty} g_n(v) = \begin{cases} \mathbb{E}[H(s\boldsymbol{Y})] & \text{if } v \in (0,1) , \\ 0 & \text{if } v > 1 . \end{cases}$$

We split

$$\int_{h/r_n}^{1+h/r_n} g_n(v)\mathrm{d}v = \int_{h/r_n}^{1} g_n(v)\mathrm{d}v + \int_{1}^{1+h/r_n} g_n(v)\mathrm{d}v .$$

Since the sequence $\{g_n\}$ is uniformly bounded, for any $h < r_n$ the second integral above converges to zero as $n \to \infty$. If $h = o(r_n)$ and since there is no problem at $v = 0$ with $g_n(v)$, then the first integral converges to

$$\int_0^1 \mathbb{E}[H(s\boldsymbol{Y})]\mathrm{d}v = \mathbb{E}[H(s\boldsymbol{Y})] .$$

Likewise, when $h = [\xi r_n]$ then the first integral converges to

$$\int_\xi^1 \mathbb{E}[H(s\boldsymbol{Y})]\mathrm{d}v = (1 - \xi)\mathbb{E}[H(s\boldsymbol{Y})] . \qquad \square$$

7.5. Empirical cluster process of sliding blocks

Recall that for $s > 0$, $H_s(\boldsymbol{x}) = H(\boldsymbol{x}/s)$. In order to deal with asymptotic normality of sliding blocks estimators, we study the empirical process

$$\mathbb{F}_n(H_s) := \sqrt{k_n}\left\{\widetilde{\boldsymbol{\mu}}_{n,r_n}^*(H_s) - \boldsymbol{\nu}^*(H_s)\right\}$$
$$= \sqrt{k_n}\left\{\frac{\sum_{i=0}^{q_n-1} H_s\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right)}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} - s^{-\alpha}\boldsymbol{\nu}^*(H)\right\} .$$

The process $\mathbb{F}_n(H_s)$ is viewed as a random element with values in $\mathbb{D}([s_0, t_0])$.

Theorem 7.15. *Let $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying $\mathbb{R}^d$-valued time series. Assume that $\mathcal{R}(r_n, u_n)$, $\beta(r_n)$ and $\mathcal{AC}(r_n, u_n)$ hold. Let $H \in \mathcal{A}$ be such that the class $\{H_s : s \in [s_0, t_0]\}$ is linearly ordered and (4.2b) holds.*

Then $\mathbb{F}_n(H.)$ converges weakly in $(\mathbb{D}([s_0, t_0]), J_1)$ to a Gaussian process with the covariance $\boldsymbol{\nu}^*(H_s H_t)$.

If moreover $\mathrm{ANSJB}(r_n, u_n)$ is satisfied, then the convergence holds for $H \in \mathcal{B}$.

If additionally $\mathcal{S}(r_n, u_n)$ and (4.2a) are satisfied, then the processes $\mathbb{F}_n(H.)$ and $\mathbb{F}_n(\mathcal{E}.)$ converge jointly.

7.5.1. Tail empirical process

Consider the following tail empirical process:

$$\widetilde{\mathbb{T}}_n(s) = \sqrt{k_n}\left\{T_n(s) - s^{-\alpha}\right\} = \sqrt{k_n}\left\{\frac{\sum_{j=1}^{q_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n s\}}{q_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} - s^{-\alpha}\right\} , \quad s > 0 .$$

Note that this is the classical tail empirical process based on the random variables $|\boldsymbol{X}_j|$, $j \geq 1$, with the only one difference: q_n replaces n. We argue that this process can be obtained (approximately) as the empirical process of sliding blocks. Indeed,

$$\widetilde{\boldsymbol{\mu}}_{n,r_n}^*(\mathcal{E}_s)$$

$$= \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}\left\{\sum_{j=1}^{r_n} j + r_n \sum_{j=r_n+1}^{q_n} + \sum_{j=q_n+1}^{n}(n - j)\right\}\mathbb{1}\{|\boldsymbol{X}_j| > u_n s\} .$$

The difference between $\widetilde{\boldsymbol{\mu}}_{n,r_n}^*(\mathcal{E}_s)$ and $T_n(s)$ is

$$A := \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}\left\{\sum_{j=1}^{r_n}(r_n - j) - \sum_{j=q_n+1}^{n}(n - j)\right\}\mathbb{1}\{|\boldsymbol{X}_j| > u_n s\} .$$

We have $\sum_{j=1}^{r_n}(r_n - j) \leq r_n^2$ and $\sum_{j=q_n+1}^{n}(n-j) \leq r_n^2$, thus under $\mathcal{R}(r_n, u_n)$:

$$\lim_{n\to\infty} \sqrt{k_n}\mathbb{E}[|A|] \leq \text{cst} \lim_{n\to\infty} \sqrt{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\frac{r_n}{q_n}$$

$$= \text{cst} \lim_{n\to\infty} \sqrt{\frac{r_n}{n}}\sqrt{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} = 0 \ .$$

This implies that $\mathbb{F}_n(\mathcal{E}_s)$ and $\widetilde{\mathbb{T}}_n(s)$ are asymptotically equivalent in the sense that they yield the same process $\mathbb{F}(\mathcal{E}_s)$ as the distributional limit.

7.6. Covariance of the empirical process of sliding blocks

Proposition 7.16. *Assume that $\mathcal{AC}(r_n, u_n)$ and $\mathcal{R}(r_n, u_n)$ are satisfied. Let*

- *$H, \widetilde{H} \in \widetilde{\mathcal{A}}$, or*
- *$H, \widetilde{H} \in \widetilde{\mathcal{B}}$ and $\text{ANSJB}(r_n, u_n)$ holds.*

If (7.1b) *and* (7.1c) *hold then*

$$\lim_{n\to\infty} \text{cov}(\mathbb{F}_n(H), \mathbb{F}_n(\widetilde{H})) = \boldsymbol{\nu}^*(H\widetilde{H}) \ . \tag{7.9}$$

If (7.1b) *holds then*

$$\lim_{n\to\infty} \text{cov}(\mathbb{F}_n(H), \mathbb{F}_n(\mathcal{E}_s)) = \boldsymbol{\nu}^*(H\mathcal{E}_s) = \mathbb{E}[H(s\boldsymbol{Y})] \ . \tag{7.10}$$

Remark 7.17.   • The second equality in (7.10) follows from Lemma 7.22.
    • In view of the discussion in Section 7.5.1, (7.10) can be re-phrased as

$$\lim_{n\to\infty} \text{cov}(\mathbb{F}_n(H), \widetilde{\mathbb{T}}_n(s)) = \boldsymbol{\nu}^*(H\mathcal{E}_s) = \mathbb{E}[H(s\boldsymbol{Y})] \ .$$

$$\oplus$$

7.6.1. Bounds for integral representation

Before we proceed with the proof, we define

$$g_n(\xi; H) = \mathbb{E}\left[H(\boldsymbol{X}_{1,r_n}/u_n)H(\boldsymbol{X}_{1+[r_n\xi],[r_n\xi]+r_n}/u_n)\right] \ , \quad \xi > 0$$

and

$$\widetilde{g}_n(\xi; H) = \frac{g_n(\xi; H)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \ .$$

For $\xi = 0$, using Remark 7.7 we immediately obtain under $\mathcal{AC}(r_n, u_n)$:

$$\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}\cup\widetilde{\mathcal{B}}} \widetilde{g}_n(0; H) = \lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}\cup\widetilde{\mathcal{B}}} \boldsymbol{\nu}^*_{n,r_n}(H^2) < \infty \ . \tag{7.11}$$

Furthermore, for $j = 1, 2, 3, \ldots$,

$$\frac{1}{r_n}\sum_{i=(j-1)r_n}^{jr_n-1} \widetilde{g}_n(i/r_n; H) = \int_{j-1}^{j} \widetilde{g}_n(\xi; H)\mathrm{d}\xi \ .$$

For $j = 1$ we will need the precise behaviour of this integral and we will handle it using Propositions 7.12 and 7.13. For $j \geq 2$ the integral vanishes with a given rate.

Lemma 7.18. *Assume that $\mathcal{AC}(r_n, u_n)$ holds.*

- *If* (7.1b) *holds then for any finite M,*

$$\lim_{n \to \infty} \sup_{H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}} \int_1^M \widetilde{g}_n(\xi; H) \mathrm{d}\xi = 0$$

- *For $j \geq 3$,*

$$\sup_{H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}} \int_{j-1}^j \widetilde{g}_n(\xi; H) \mathrm{d}\xi \leq \mathrm{cst} \, \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \beta_{(j-2)r_n} \, .$$

Proof. For the first part we apply Lemma 7.10 and the dominated convergence:

$$\sup_{H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}} \sup_{\xi \in (1,2)} |\widetilde{g}_n(\xi; H)| \leq \sup_{H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}} \|H\|_\infty \boldsymbol{\nu}^*_{n,r_n}(|H|) \leq \mathrm{cst} \sup_{H \in \widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}} \boldsymbol{\nu}^*_{n,r_n}(|H|) < \infty \, .$$

For the second part, we use (7.2) and the fact that $\widetilde{\mathcal{A}} \cup \widetilde{\mathcal{B}}$ has a finite envelope. $\qquad \square$

7.6.2. Representation for covariance between blocks

Recall that $q_n = n - r_n + 1$. Evaluation of the covariance of the empirical process of sliding blocks will use consecutive disjoint blocks of indices of size r_n:

$$J_j = \{(j-1)r_n, \ldots, jr_n - 1\} \, , \quad j = 1, \ldots, m_n = [q_n/r_n] \, .$$

Clearly, $\bigcup_{j=1}^{m_n} J_j = \{0, \ldots, n - r_n\}$. We will assume for simplicity that q_n/r_n is an integer.

Write

$$\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) = \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{m_n} \Psi_j(H)$$

with

$$\Psi_j(H) = \sum_{i \in J_j} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) \, . \tag{7.12}$$

Note that the indices of the random vectors $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{2r_n-1}$ used in the construction of Ψ_1 overlap with the indices of $\boldsymbol{X}_{r_n+1}, \ldots, \boldsymbol{X}_{3r_n-1}$ used to define Ψ_2, but do not overlap with the indices used in the definition of Ψ_3. Likewise, the indices used in the definition of Ψ_2 overlap with those in Ψ_3, but not with any other term Ψ_j, $j \geq 4$. This partially explains where does a contribution to

the limiting variance come from: from the dependence within each block J_j and cross dependence between J_j and two neighbouring blocks.

For $j \geq 1$ we have

$$\frac{\mathbb{E}[\Psi_1(H)\Psi_{j+1}(H)]}{r_n^3\mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$= \frac{1}{r_n^3\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\mathbb{E}\left[\sum_{h=0}^{r_n-1} H\left(\boldsymbol{X}_{h+1,h+r_n}/u_n\right)\sum_{i=jr_n}^{(j+1)r_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right)\right]$$

$$= \frac{1}{r_n}\sum_{i=(j-1)r_n}^{jr_n}\left(\frac{i}{r_n} - (j-1)\right)\widetilde{g}_n(i/r_n; H)$$

$$+ \frac{1}{r_n}\sum_{i=jr_n+1}^{(j+1)r_n}\left((j+1) - \frac{i}{r_n}\right)\widetilde{g}_n(i/r_n; H)$$

and hence

$$\frac{\mathbb{E}[|\Psi_1(H)\Psi_{j+1}(H)|]}{r_n^3\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \leq \frac{1}{r_n}\sum_{i=(j-1)r_n}^{(j+1)r_n}|\widetilde{g}_n(i/r_n; H)| \leq \int_{j-1}^{j+1}|\widetilde{g}_n(\xi; H)|\mathrm{d}\xi. \quad (7.13)$$

7.6.3. Proof of Proposition 7.16, Eq. (7.9)

Proof. Note that (since $q_n \sim n$)

$$\frac{k_n m_n}{(q_n r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n))^2} = \frac{q_n n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}{r_n\left(q_n r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)\right)^2} \sim \frac{1}{r_n^3\mathbb{P}(|\boldsymbol{X}_0| > u_n)}. \quad (7.14)$$

Write $\mathrm{var}(\mathbb{F}_n(H))$ as

$$\frac{k_n m_n}{(q_n r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n))^2}\mathrm{cov}\left(\Psi_2(H), \Psi_1(H) + \Psi_2(H) + \Psi_3(H)\right) + A_n(H) \quad (7.15)$$

with the reminder $A_n(H)$ given by

$$A_n(H) := -2\frac{k_n}{(q_n r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n))^2}\mathrm{cov}(\Psi_1(H), \Psi_2(H)) \quad (7.16)$$

$$+ 2\frac{1 + o(1)}{r_n^3\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\sum_{j=2}^{m_n-1}\left(1 - \frac{j}{m_n}\right)\{\mathrm{cov}\left(\Psi_1(H), \Psi_{1+j}(H)\right)\} \quad (7.17)$$

$$=: A_{n,1}(H) + (1 + o(1))B_n(H).$$

If we show that the leading term on the right-hand side of (7.15) converges to a finite limit, then automatically $\lim_{n\to\infty} A_{n,1}(H) = 0$ (since $m_n \to \infty$). Thus, the reminder $A_n(H)$ will be negligible if we show that

$$\lim_{n\to\infty} B_n(H) = 0. \quad (7.18)$$

We will start by analysing the first term in (7.15). Set

$$R_n(H) = \frac{1}{r_n} \sum_{i=r_n}^{2r_n-1} (2 - i/r_n) \widetilde{g}_n(i/r_n; H) .$$

Since (7.1b) holds, the application of the first part of Lemma 7.18 gives

$$\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}\cup\widetilde{\mathcal{B}}} R_n(H) = 0 . \tag{7.19}$$

Write the first term in (7.15) as (cf. (7.14))

$$\frac{1+o(1)}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \left\{ r_n g_n(0; H) + 2r_n \sum_{i=1}^{r_n-1} g_n(i/r_n; H) + 2r_n R_n \right\}$$

$$= (1+o(1)) \left\{ \frac{1}{r_n} \widetilde{g}_n(0; H) + 2\frac{1}{r_n} \sum_{i=1}^{r_n-1} \widetilde{g}_n(i/r_n; H) + 2R_n(H) \right\} . \tag{7.20}$$

Then, using (7.11), (7.18) and (7.19), we have

$$\lim_{n\to\infty} \operatorname{var}(\mathbb{F}_n(H)) = 2 \lim_{n\to\infty} \int_0^1 \widetilde{g}_n(\xi; H)\mathrm{d}\xi ,$$

Applying Propositions 7.12 and 7.13 (the case $h = [\xi r_n]$), we have

$$\lim_{n\to\infty} \operatorname{var}(\mathbb{F}_n(H)) = 2\boldsymbol{\nu}^*(H^2) \int_0^1 (1 - \xi)\mathrm{d}\xi = \boldsymbol{\nu}^*(H^2) .$$

To conclude the proof, we show (7.18) in the following lemma. $\qquad\square$

Lemma 7.19. *Assume that* (7.1b)-(7.1c) *hold. Then*

$$\lim_{n\to\infty} \sup_{H\in\widetilde{\mathcal{A}}\cup\widetilde{\mathcal{B}}} B_n(H) = 0 . \tag{7.21}$$

Proof of Lemma 7.19. Using (7.13) we have

$$|B_n(H)| \leq \frac{\mathbb{E}[|\Psi_1(H)\Psi_3(H)|]}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} + \frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=3}^{m_n-1} |\mathbb{E}[|\Psi_1(H)\Psi_{1+j}(H)|]$$

$$\leq \int_1^3 |\widetilde{g}_n(\xi; H)|\mathrm{d}\xi + \sum_{j=3}^{m_n-1} \int_{j-1}^{j+1} |\widetilde{g}_n(\xi; H)|\mathrm{d}\xi .$$

The first term is $o(1)$ uniformly over the class of functions (cf. the first part of Lemma 7.18). Using the second part of Lemma 7.18 we bound

$$\sum_{j=3}^{m_n-1} \int_{j-1}^{j+1} |\widetilde{g}_n(\xi; H)|\mathrm{d}\xi \leq \operatorname{cst} \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{\infty} \beta_{jr_n} .$$

We finish the proof by applying the mixing assumption (7.1c). $\qquad\square$

7.6.4. Proof of Proposition 7.16, Eq. (7.10)

Proof. We write (recall that $q_n \sim n$)

$$
k_n \left(\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \right)^2 \mathrm{cov} \left(\sum_{i=0}^{q_n-1} H(\boldsymbol{X}_{i+1,i+r_n}/u_n), \sum_{j=1}^{q_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\} \right)
$$

$$
\sim \frac{k_n}{n^2 r_n \mathbb{P}^2(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} \sum_{j=1}^{q_n} \mathrm{cov}\left(H(\boldsymbol{X}_{i+1,i+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_j| > u_n\} \right)
$$

$$
= \frac{1}{n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} \sum_{j=1}^{q_n} \mathrm{cov}\left(H(\boldsymbol{X}_{i-j+1,i-j+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right) .
$$

Split the inner sum into two pieces, $\sum_{j=1}^{i}$ and $\sum_{j=i+1}^{q_n}$, in the first one replace j with $h = i - j$, in the second one replace j with $h = j - i$ to get

$$
\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} \sum_{j=1}^{i} \mathrm{cov}\left(H(\boldsymbol{X}_{i-j+1,i-j+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
+ \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} \sum_{j=i+1}^{q_n} \mathrm{cov}\left(H(\boldsymbol{X}_{i-j+1,i-j+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
= \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=1}^{q_n-1} \sum_{h=0}^{i-1} \mathrm{cov}\left(H(\boldsymbol{X}_{h+1,h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
+ \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} \sum_{h=1}^{q_n-i} \mathrm{cov}\left(H(\boldsymbol{X}_{-h+1,-h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right) .
$$

This gives further

$$
\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=0}^{q_n-2} \sum_{i=h+1}^{q_n-1} \mathrm{cov}\left(H(\boldsymbol{X}_{h+1,h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
+ \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=1}^{q_n-2} \sum_{i=1}^{q_n-h} \mathrm{cov}\left(H(\boldsymbol{X}_{-h+1,-h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
= \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=0}^{q_n-1} (1 - h/q_n) \mathrm{cov}\left(H(\boldsymbol{X}_{h+1,h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right)
$$

$$
\tag{7.22}
$$

$$
+ \frac{1}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=0}^{q_n-1} (1 - h/q_n) \mathrm{cov}\left(H(\boldsymbol{X}_{-h+1,-h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\} \right) .
$$

$$
\tag{7.23}
$$

We show that the term in (7.22) is negligible, while the one in (7.23) yields the limit. We split the term in (7.22) into two pieces, according to $h \le r_n$ and

$h > r_n$. Then the first part is bounded by

$$\frac{1}{r_n} \sum_{h=0}^{r_n} \mathbb{E}\left[|H(\boldsymbol{X}_{h+1,h+r_n}/u_n)|\,|\,|\boldsymbol{X}_0| > u_n\right] = \int_0^1 g_n(v)\mathrm{d}v$$

with

$$g_n(v) = \mathbb{E}\left[|H(\boldsymbol{X}_{[r_n v]+1,[r_n v]+r_n}/u_n)|\,|\,|\boldsymbol{X}_0| > u_n\right] .$$

Under $\mathcal{AC}(r_n, u_n)$, $g_n(v) \to 0$ (cf. the first part of Lemma 7.9 with $H \equiv 1$ and $\widetilde{H} = |H|$). Likewise, since $r_n/q_n \to 0$,

$$\frac{1}{r_n q_n} \sum_{h=0}^{r_n} h\mathbb{E}\left[|H(\boldsymbol{X}_{h+1,h+r_n}/u_n)|\,|\,|\boldsymbol{X}_0| > u_n\right] = q_n^{-1} \int_0^1 [r_n v]g_n(v)\mathrm{d}v \to 0 .$$

Furthermore, applying (7.2),

$$\frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=r_n+1}^{q_n-1} (1 - h/q_n)|\mathrm{cov}\left(H(\boldsymbol{X}_{h+1,h+r_n}/u_n), \mathbb{1}\{|\boldsymbol{X}_0| > u_n\}\right)|$$

$$\leq \frac{\|H\|_\infty}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{h=1}^{n} \beta_{h+r_n}$$

and the latter term vanishes by (7.1b). In summary, (7.22) is negligible.

For the term in (7.23) we write (recall that we can replace cov with $\mathbb{E}$ thanks to $\mathcal{R}(r_n, u_n)$)

$$\frac{1}{r_n} \sum_{h=0}^{r_n} \mathbb{E}\left[H(\boldsymbol{X}_{-h+1,-h+r_n}/u_n)\,|\,|\boldsymbol{X}_0| > u_n\right] = \int_0^1 g_n(v)\mathrm{d}v$$

with

$$g_n(v) = \mathbb{E}\left[H(\boldsymbol{X}_{-[r_n v]+1,-[r_n v]+r_n}/u_n)\,|\,|\boldsymbol{X}_0| > u_n\right] .$$

By Proposition 3.2, $g_n(v) \to \mathbb{E}[H(\boldsymbol{Y})]$ for each v. $\qquad\square$

7.7. Proof of Theorem 7.15 - fidi convergence

Recall that $q_n = n - r_n + 1$ and recall the disjoint blocks of size r_n:

$$J_j = \{(j-1)r_n, \ldots, jr_n - 1\} , \quad j = 1, \ldots, m_n = [q_n/r_n] .$$

These blocks were chosen to calculate the limiting covariance of the process $\mathbb{F}_n$. However, they are not appropriate for a proof of the central limit theorem. We need to introduce a large-small blocks decomposition. For this purpose let z_n be a sequence of integers such that $z_n \to \infty$ and

$$\lim_{n\to\infty} \frac{z_n}{\sqrt{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}} = 0 . \tag{7.24}$$

Set
$$\widetilde{m}_n = \left[\frac{q_n}{(z_n + 2)r_n} \right]$$

and assume for simplicity that $\widetilde{m}_n$ is an integer. Since $z_n \to \infty$, we have $\widetilde{m}_n = o(m_n)$. For $j = 1, \ldots, \widetilde{m}_n$ define now large and small blocks as follows:

$$
\begin{aligned}
L_1 &= \{0, \ldots, z_n r_n - 1\} , \quad S_1 = \{z_n r_n \ldots, z_n r_n + 2r_n - 1\} , \\
L_2 &= \{z_n r_n + 2r_n, \ldots, 2z_n r_n + 2r_n - 1\} , \\
S_2 &= \{2z_n r_n + 2r_n \ldots, 2z_n r_n + 4r_n - 1\} , \\
L_j &= \{(j-1)z_n r_n + 2(j-1)r_n, \ldots, jz_n r_n + 2(j-1)r_n - 1\} , \\
S_j &= \{jz_n r_n + 2(j-1)r_n, \ldots, jz_n r_n + 2jr_n - 1\} .
\end{aligned}
$$

The block L_1 is obtained by merging z_n consecutive blocks $J_1, \ldots, J_{z_n}$ of size r_n. Likewise, $S_1 = J_{z_n+1} \cup J_{z_n+2}$. Therefore, the large block of size $z_n r_n$ is followed by the small block of size $2r_n$, which in turn is followed by the large block of size $z_n r_n$ and so on. All together,

$$\bigcup_{j=1}^{\widetilde{m}_n} (L_j \cup S_j) = \{0, \ldots, n - r_n\} .$$

Write

$$\sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1, i+r_n}/u_n\right) = \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(l)}(H) + \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(s)}(H) , \qquad (7.25)$$

where now

$$\Psi_j^{(l)}(H) = \sum_{i \in L_j} H\left(\boldsymbol{X}_{i+1, i+r_n}/u_n\right) , \quad \Psi_j^{(s)}(H) = \sum_{i \in S_j} H\left(\boldsymbol{X}_{i+1, i+r_n}/u_n\right) .$$

With such the decomposition, $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_{z_n r_n + r_n - 1}$ used in the definition of $\Psi_1^{(l)}(H)$ are separated by $r_n + 2$ from the random variables that define $\Psi_2^{(l)}(H)$. The mixing condition (7.1a) allows us to replace $\boldsymbol{X}$ with the independent blocks process, that is, we can treat the random variables $\Psi_j^{(l)}(H)$, $j = 1, \ldots, \widetilde{m}_n$, as independent. The same applies to $\Psi_j^{(s)}(H)$.

Set

$$\mathbb{Z}_n(H) = \sum_{j=1}^{\widetilde{m}_n} \{Z_{n,j}(H) - \mathbb{E}[Z_{n,j}(H)]\} =: \sum_{j=1}^{\widetilde{m}_n} \bar{Z}_{n,j}(H) \qquad (7.26)$$

with

$$Z_{n,j}(H) = \frac{\sqrt{k_n}}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \Psi_j^{(l)}(H) . \qquad (7.27)$$

The next steps are standard.

- First, we show that the limiting variance of the large blocks process $\mathbb{Z}_n$ is the same as that of the process $\mathbb{F}_n$;
- Next, we show that the small blocks process (the second term in (7.25)) is negligible;
- Finally, we will verify the Lindeberg condition for the large blocks process.

Variance of the large blocks. We have (using the assumed independence of $\Psi_j^{(l)}(H)$)

$$
k_n \text{var} \left(\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(l)}(H) \right) = \frac{k_n \widetilde{m}_n}{(q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n))^2} \text{var}(\Psi_1^{(l)}(H))
$$

$$
\sim \frac{1}{z_n r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \text{var} \left(\sum_{i=0}^{z_n r_n - 1} H\left(\boldsymbol{X}_{i+1,i+r_n} / u_n \right) \right)
$$

$$
= \frac{1}{z_n r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \text{var} \left(\sum_{j=1}^{z_n} \Psi_j(H) \right) , \tag{7.28}
$$

where in the last line we decomposed the block $L_1 = \{0, \dots, z_n r_n - 1\}$ into z_n disjoint blocks $J_1, \dots, J_{z_n}$, used the notation (7.12), the asymptotics (7.14) and $\widetilde{m}_n \sim m_n / z_n$.

The next steps are a repetition of the proof of Proposition 7.16, with the appropriate adjustments. The term in (7.28) becomes

$$
\frac{\text{var}\left(\Psi_1(H)\right)}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} + 2 \frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{z_n-1} \left(1 - \frac{j}{z_n} \right) \text{cov}(\Psi_1(H), \Psi_{1+j}(H))
$$

and as in (7.15) we can write it as

$$
\frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \left\{ \text{cov}(\Psi_2(H), \Psi_1(H) + \Psi_2(H) + \Psi_3(H)) \right\} + \widetilde{A}_n(H) \tag{7.29}
$$

with the reminder $\widetilde{A}_n(H)$ given this time by (cf. (7.16)-(7.17))

$$
\widetilde{A}_n(H) := -2 \frac{1}{z_n} \frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \text{cov}(\Psi_1(H), \Psi_2(H))
$$

$$
+ \frac{2}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=2}^{z_n-1} \left(1 - \frac{j}{z_n} \right) \left\{ \text{cov}\left(\Psi_1(H), \Psi_{1+j}(H)\right) \right\}
$$

$$
=: \widetilde{A}_{n,1}(H) + \widetilde{B}_n(H) . \tag{7.30}
$$

The reminder is negligible by the same argument as before. Indeed, we note that $\widetilde{B}_n(H)$ is just $B_n(H)$ from (7.17) with m_n replaced with z_n. The dependence on m_n vanishes in the final stage of the proof of Lemma 7.19. The leading term in (7.29) is the same as in the proof of Proposition 7.16; cf. (7.15).

In summary, the variance of the large block process is

$$\lim_{n \to \infty} \mathrm{var} \left(\mathbb{Z}_n(H) \right) = \boldsymbol{\nu}^*(H^2) \ .$$

Variance of the small blocks. We have (using again the assumed independence of $\Psi_j^{(s)}(H)$ thanks to the beta-mixing)

$$k_n \mathrm{var} \left(\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(s)}(H) \right) \sim \frac{1}{z_n r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathrm{var}(\Psi_1^{(s)}(H)) \ .$$

Since $\Psi_1^{(s)}(H)$ is just $\Psi_1(H)$ defined in (7.12), we have

$$k_n \mathrm{var} \left(\frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(s)}(H) \right) = O(1/z_n) = o(1) \ .$$

Lindeberg condition for $\mathbb{Z}_n(H)$. We need to show that for all $\eta > 0$,

$$\lim_{n \to \infty} \widetilde{m}_n \mathbb{E} \left[Z_{n,1}^2(H) \mathbb{1}\{|Z_{n,1}(H)| > \eta\} \right] = 0 \ . \tag{7.31}$$

Since H is bounded, then by (7.24),

$$|Z_{n,1}(H)| \le \frac{\sqrt{k_n} z_n r_n}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \|H\|_\infty \sim \frac{z_n}{\sqrt{n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}} \|H\|_\infty = o(1) \ .$$

Thus, the indicator in (7.31) becomes zero for large n.

Lindeberg condition for $\mathbb{Z}_n(\mathcal{E})$. The functional $\mathcal{E}$ is not bounded and we will prove the Lindeberg condition under $\mathcal{S}(r_n, u_n)$. Write

$$\widetilde{w}_n = \frac{\sqrt{k_n}}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

so that

$$Z_{n,1}(\mathcal{E}) = \widetilde{w}_n \sum_{i=0}^{z_n r_n - 1} \mathcal{E} \left(\boldsymbol{X}_{i+1, i+r_n}/u_n \right) = \widetilde{w}_n \sum_{i=0}^{z_n r_n - 1} \sum_{j=i+1}^{i+r_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\}$$

$$= \widetilde{w}_n \left\{ \sum_{j=1}^{r_n} \sum_{i=0}^{j-1} + \sum_{j=r_n+1}^{r_n(z_n+1)} \sum_{i=j-r_n}^{j-1} \right\} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\} \le \widetilde{w}_n r_n \sum_{j=1}^{r_n(z_n+1)} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\}$$

$$\le \frac{\sqrt{k_n}}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} r_n \sum_{j=1}^{2r_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\}$$

$$= \frac{1 + o(1)}{\sqrt{n \mathbb{P}(|\boldsymbol{X}_0| > u_n)}} \sum_{j=1}^{2r_n} \mathbb{1}\{|\boldsymbol{X}_j| > u_n\} \ .$$

The last term can be recognized as one (scaled) block of size $2r_n$ of the tail empirical process $\widetilde{\mathbb{T}}_n(s)$. [KSW19, Lemma 3.6] (see also [KS20, Lemma 9.2.8]) gives

$$\lim_{n\to\infty} m_n \mathbb{E}\left[Z_{n,1}^2(\mathcal{E})\mathbb{1}\{|Z_{n,1}(\mathcal{E})| > \eta\}\right] = 0 \ .$$

If moreover $\mathcal{R}(r_n, u_n)$ holds then

$$\lim_{n\to\infty} m_n \mathbb{E}\left[\bar{Z}_{n,1}^2(\mathcal{E})\mathbb{1}\{|\bar{Z}_{n,1}(\mathcal{E})| > \eta\}\right] = 0 \ .$$

Since $\widetilde{m}_n = o(m_n)$, we obtain the Lindeberg condition for $\mathbb{Z}_n(\mathcal{E})$.

7.8. Proof of Theorem 7.15 - asymptotic equicontinuity

We need the following lemma which is an adapted version of Theorem 2.11.1 in [vdVW96]. Let $\mathbb{Z}_n$ be the empirical process indexed by a semi-metric space $(\mathcal{G}, \rho)$, defined by

$$\mathbb{Z}_n(f) = \sum_{j=1}^{\widetilde{m}_n} \{Z_{n,j}(f) - \mathbb{E}[Z_{n,j}(f)]\} \ ,$$

where $\{Z_{n,j}, n \geq 1\}$, $j = 1, \ldots, \widetilde{m}_n$, are i.i.d. separable, stochastic processes and $\widetilde{m}_n$ is a sequence of integers such that $\widetilde{m}_n \to \infty$. Define the random semi-metric d_n on $\mathcal{G}$ by

$$d_n^2(f, g) = \sum_{j=1}^{\widetilde{m}_n} \{Z_{n,j}(f) - Z_{n,j}(g)\}^2 \ , f, g \in \mathcal{G} \ .$$

Lemma 7.20. *Assume that $(\mathcal{G}, \rho)$ is totally bounded. Assume moreover that:*

(i) For all $\eta > 0$,

$$\lim_{n\to\infty} \widetilde{m}_n \mathbb{E}[\|Z_{n,1}\|_{\mathcal{G}}^2 \mathbb{1}\{\|Z_{n,1}\|_{\mathcal{G}}^2 > \eta\}] = 0 \ .$$

(ii) For every sequence $\{\delta_n\}$ which decreases to zero,

$$\lim_{n\to\infty} \sup_{\substack{f,g\in\mathcal{G} \\ \rho(f,g)\leq\delta_n}} \mathbb{E}[d_n^2(f,g)] = 0 \ . \tag{7.32}$$

(iii) There exists a measurable majorant $N^*(\mathcal{G}, d_n, \epsilon)$ of the covering number $N(\mathcal{G}, d_n, \epsilon)$ such that for every sequence $\{\delta_n\}$ which decreases to zero,

$$\int_0^{\delta_n} \sqrt{\log N^*(\mathcal{G}, d_n, \epsilon)}\mathrm{d}\epsilon \overset{\mathbb{P}}{\longrightarrow} 0 \ . \tag{7.33}$$

Then $\{\mathbb{Z}_n, n \geq 1\}$ is asymptotically ρ-equicontinuous, i.e. for each $\eta > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}\left(\sup_{\substack{f,g\in\mathcal{G} \\ \rho(f,g)<\delta}} |\mathbb{Z}_n(f) - \mathbb{Z}_n(g)| > \eta \right) = 0 \ .$$

Remark 7.21. The separability assumption is not in [vdVW96]. It implies measurability of $\|Z_{n,1}\|_{\mathcal{G}}$. Furthermore, the separability also implies that for all $\delta > 0$, $n \in \mathbb{N}$, $(e_j)_{1 \le j \le \widetilde{m}_n} \in \{-1, 0, 1\}^{\widetilde{m}_n}$ and $i \in \{1, 2\}$, the supremum

$$\sup_{\substack{f,g \in \mathcal{G} \\ \rho(f,g) < \delta}} \left| \sum_{j=1}^{\widetilde{m}_n} e_j \left(Z_{n,j}(f) - Z_{n,j}(g) \right)^i \right| = \sup_{\substack{f,g \in \mathcal{G}_0 \\ \rho(f,g) < \delta}} \left| \sum_{j=1}^{\widetilde{m}_n} e_j \left(Z_{n,j}(f) - Z_{n,j}(g) \right)^i \right|$$

is measurable, which is an assumption of [vdVW96]. $\qquad\qquad\oplus$

7.8.1. Asymptotic equicontinuity of the empirical process of sliding blocks

Recall the big-blocks process $\mathbb{Z}_n(H)$ (cf. (7.26)-(7.27)). Recall also that thanks to β-mixing we can consider random variables $\Psi_j^{(l)}(H)$, $j = 1, \ldots, \widetilde{m}_n$ to be independent. We need to prove asymptotic equicontinuity of $\mathbb{Z}_n(H)$ indexed by the class $\mathcal{G} = \{H_s, s \in [s_0, t_0]\}$ equipped with the metric $\boldsymbol{\rho}^*(H, \widetilde{H}) = \boldsymbol{\nu}^*(\{H - \widetilde{H}\}^2)$. The same argument can be used to prove asymptotic equicontinuity for the small blocks process. This yields asymptotic equicontinuity of $\mathbb{F}_n(H.)$. We note further that asymptotic continuity of $\mathbb{F}_n(\mathcal{E}.)$ follows from [KSW19].

- The Lindeberg condition (i) of Lemma 7.20 holds because the class $\mathcal{G}$ is linearly ordered and by applying (7.31).
- Since $\mathcal{G}$ is linearly ordered, the random entropy condition (7.33) of Lemma 7.20 holds.
- Define the random metric

$$d_n^2(H, \widetilde{H}) = \sum_{j=1}^{\widetilde{m}_n} (Z_{n,j}(H) - Z_{n,j}(\widetilde{H}))^2 .$$

We need to evaluate $\mathbb{E}[d_n^2(H_s, H_t)]$:

$$\begin{aligned}
&\mathbb{E}[d_n^2(H_s, H_t)] \\
&= \frac{k_n \widetilde{m}_n}{(q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n))^2} \mathbb{E}\left[\left(\sum_{i=0}^{z_n r_n - 1} \{ H_s \left(\boldsymbol{X}_{i+1, i+r_n}/u_n \right) \right. \right. \qquad (7.34) \\
&\qquad \left. \left. - H_t \left(\boldsymbol{X}_{i+1, i+r_n}/u_n \right) \} \right)^2 \right] \\
&\sim \frac{1}{z_n r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[\left(\sum_{j=1}^{z_n} \{ \Psi_j(H_s) - \Psi_j(H_t) \} \right)^2 \right], \qquad (7.35)
\end{aligned}$$

where in the last line we decomposed the block L_1 into z_n disjoint blocks $J_1, \ldots, J_{z_n}$, used the notation (7.12), the asymptotics (7.14) and $\widetilde{m}_n \sim m_n/z_n$; cf. (7.28).

The term in (7.34) becomes

$$\frac{\mathbb{E}[(\Psi_1(H_s) - \Psi_1(H_t))^2]}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)}$$

$$+ \frac{2}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=1}^{z_n-1} \left(1 - \frac{j}{z_n}\right) \mathbb{E}[\{\Psi_1(H_s) - \Psi_1(H_t)\}$$

$$\times \{\Psi_{1+j}(H_s) - \Psi_{1+j}(H_t)\}]$$

and as in (7.15) we can write it as

$$\frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}\left[\{\Psi_2(H_s) - \Psi_2(H_t)\} \sum_{j=1}^{2} \{\Psi_j(H_s) - \Psi_j(H_t)\}\right] + \widetilde{A}_n(H, s, t)$$

(7.36)

with the reminder (cf. (7.30))

$$\widetilde{A}_n(H, s, t) := -2\frac{1}{z_n} \frac{1}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}[\{\Psi_1(H_s) - \Psi_1(H_t)\}\{\Psi_2(H_s) - \Psi_2(H_t)\}]$$

$$+ \frac{2}{r_n^3 \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{j=2}^{z_n-1} \left(1 - \frac{j}{z_n}\right) \mathbb{E}\left[\{\Psi_1(H_s) - \Psi_1(H_t)\}\right.$$

$$\times \{\Psi_{j+1}(H_s) - \Psi_{j+1}(H_t)\}]$$

$$= \widetilde{A}_{n,1}(H_s - H_t) + \widetilde{B}_n(H_s - H_t) .$$

Remark 7.8 applies and hence by Lemma 7.19,

$$\lim_{n\to\infty} \sup_{s \in [s_0, t_0]} \widetilde{B}_n(H_s - H_t) = 0 .$$

The leading term in (7.36) is decomposed as (cf. (7.20))

$$\frac{1}{r_n} \widetilde{g}_n(0; H_s - H_t) + 2\frac{1}{r_n} \sum_{i=1}^{r_n-1} \widetilde{g}_n(i/r_n; H_s - H_t) + 2R_n(H_s - H_t) .$$

Again, Remark 7.8 applies and (7.19) gives

$$\lim_{n\to\infty} \sup_{s \in [s_0, t_0]} R_n(H_s - H_t) = 0 .$$

It remains to show that for every sequence $\{\delta_n\}$ decreasing to zero,

$$\lim_{n\to\infty} \sup_{\substack{s,t \in [s_0, t_0] \\ |s-t| \le \delta_n}} \frac{1}{r_n} \sum_{i=1}^{r_n-1} \widetilde{g}_n(i/r_n; H_s - H_t)$$

$$= \lim_{n\to\infty} \sup_{\substack{s,t \in [s_0, t_0] \\ |s-t| \le \delta_n}} \int_0^1 \widetilde{g}_n(\xi, H_s - H_t) \mathrm{d}\xi = 0 .$$

Because of the monotonicity

$$|\widetilde{g}_n(\xi, H_s - H_t)| \leq 2 \sup_{s \in [s_0, t_0]} |H_s| \boldsymbol{\nu}_{n,r_n}^*(|H_s - H_t|)$$

$$\leq 2 \max\{|H_{s_0}|, |H_{t_0}|\} |\boldsymbol{\nu}_{n,r_n}^*(H_s) - \boldsymbol{\nu}_{n,r_n}^*(H_t)| .$$

The convergence of $\boldsymbol{\nu}_{n,r_n}^*(H_s)$ to $s^{-\alpha}\boldsymbol{\nu}^*(H^2)$ is uniform on $[s_0, t_0]$. Thus, for $s, t \in [s_0, t_0]$,

$$\left|\boldsymbol{\nu}_{n,r_n}^*(H_s) - \boldsymbol{\nu}_{n,r_n}^*(H_t)\right| \leq 2 \sup_{s_0 \leq u \leq t_0} \left|\boldsymbol{\nu}_{n,r_n}^*(H_u) - \boldsymbol{\nu}^*(H_u)\right| + \boldsymbol{\nu}^*(H)\{s^{-\alpha} - t^{-\alpha}\} .$$

Fix $\eta > 0$. For large enough n, the uniform convergence yields

$$\sup_{\substack{s_0 \leq s, t \leq t_0 \\ |s-t| \leq \delta_n}} \left|\boldsymbol{\nu}_{n,r_n}^*(H_s) - \boldsymbol{\nu}_{n,r_n}^*(H_t)\right|$$

$$\leq \eta + \boldsymbol{\nu}^*(H) \sup_{\substack{s_0 \leq s, t \leq t_0 \\ |s-t| \leq \delta_n}} \{s^{-\alpha} - t^{-\alpha}\} \leq \eta + \alpha s_0^{-\alpha-1} \delta_n \boldsymbol{\nu}^*(H) .$$

This proves that (7.32) holds. The conditions of Lemma 7.20 hold, thus the sequence $\mathbb{Z}_n$ is asymptotically equicontinuous.

7.9. Proof of Theorem 4.3

Write $\zeta_n = |\boldsymbol{X}|_{(n:n-k_n)}/u_n$. Since $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$, we have the relationship $\widehat{\boldsymbol{\mu}}_{n,r_n}^*(H) = \widetilde{\boldsymbol{\mu}}_{n,r_n}^*(H_{\zeta_n})$ (cf. (1.8)-(1.9)). Therefore,

$$\sqrt{k_n} \left\{\widehat{\boldsymbol{\mu}}_{n,r_n}^*(H) - \boldsymbol{\nu}^*(H)\right\} = \mathbb{F}_n(H_{\zeta_n}) + \sqrt{k_n} \left\{\boldsymbol{\nu}^*(H_{\zeta_n}) - \boldsymbol{\nu}^*(H)\right\} . \quad (7.37)$$

Step 1. Theorem 7.15 gives local uniform convergence of $\{\mathbb{F}_n(H_s), s \in [s_0, t_0]\}$ to a continuous Gaussian process $\mathbb{G}$. At the same time, convergence of $\{\mathbb{F}_n(\mathcal{E}_s), s \in [s_0, t_0]\}$ yields $\zeta_n \xrightarrow{\mathbb{P}} 1$, jointly with $\mathbb{F}_n(H_s)$. Therefore, $\mathbb{F}_n(H_{\zeta_n}) \xrightarrow{\mathrm{d}} \mathbb{G}(H)$.

Step 2. Using Vervaat's theorem, we have, jointly with the previous convergence, $\sqrt{k}(\zeta_n^{-\alpha} - 1) \xrightarrow{\mathrm{d}} \mathbb{G}(\mathcal{E})$. Therefore, by the homogeneity of $\boldsymbol{\nu}^*$,

$$\sqrt{k} \left\{\boldsymbol{\nu}^*(H_{\zeta_n}) - \boldsymbol{\nu}^*(H)\right\} = \boldsymbol{\nu}^*(H)\sqrt{k}(\zeta_n^{-\alpha} - 1) \xrightarrow{\mathrm{d}} -\boldsymbol{\nu}^*(H)\mathbb{G}(\mathcal{E}) .$$

Since the convergences hold jointly, we conclude the result.

7.10. Auxiliary results

Lemma 7.22 (Problems 5.24 and 5.25 in [KS20]). *Assume that* $\mathbb{P}(\lim_{|j| \to \infty} |\boldsymbol{Y}_j| = 0) = 1$ *and let* H, H' *be bounded functionals on* $(\mathbb{R}^d)^{\mathbb{Z}}$ *such that* $H'(\boldsymbol{x}) = 0$ *if* $\boldsymbol{x}^* \leq 1$ *and* $\mathbb{E}[|H(\boldsymbol{Y})||H'(\boldsymbol{Y}_{0,\infty}) - H'(\boldsymbol{Y}_{1,\infty})|] < \infty$. *Then*

$$\boldsymbol{\nu}^*(HH') = \mathbb{E}[H(\boldsymbol{Y})\{H'(\boldsymbol{Y}_{0,\infty}) - H'(\boldsymbol{Y}_{1,\infty})\}] ,$$

$$\boldsymbol{\nu}^*(H\mathcal{E}) = \mathbb{E}[H(\boldsymbol{Y})] , \quad \boldsymbol{\nu}^*(\mathcal{E}) = 1 , \quad \boldsymbol{\nu}^*(\mathcal{E}^2) = \sum_{j \in \mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_j| > 1) .$$

Proof. Applying (2.7) and the time change formula (see [PS18, Lemma 2.2]), we obtain

$$
\begin{aligned}
\boldsymbol{\nu}^*(HH') &= \mathbb{E}[H(\boldsymbol{Y})H'(\boldsymbol{Y})\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] \\
&= \mathbb{E}[|H(\boldsymbol{Y})||H'(\boldsymbol{Y}_{0,\infty})|\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] \\
&\le \sum_{j=0}^{\infty} \mathbb{E}[|H(\boldsymbol{Y})||H'(\boldsymbol{Y}_{j,\infty}) - H'(\boldsymbol{Y}_{j+1,\infty})|\mathbb{1}\{|\boldsymbol{Y}_j| > 1\}\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-1} \le 1\}] \\
&= \sum_{j=0}^{\infty} \mathbb{E}[H(\boldsymbol{Y})|H'(\boldsymbol{Y}_{0,\infty}) - H'(\boldsymbol{Y}_{1,\infty})|\mathbb{1}\{|\boldsymbol{Y}_{-j}| > 1\}\mathbb{1}\{\boldsymbol{Y}^*_{-\infty,-j-1} \le 1\}] \\
&= \mathbb{E}[H(\boldsymbol{Y})|H'(\boldsymbol{Y}_{0,\infty}) - H'(\boldsymbol{Y}_{1,\infty})|] < \infty \, .
\end{aligned}
$$

This proves that $\boldsymbol{\nu}^*(HH') < \infty$. Hence, we can switch the expectation with the summation and the first result follows. The second statement follows by noting that $\mathcal{E}(\boldsymbol{Y}_{0,\infty}) - \mathcal{E}(\boldsymbol{Y}_{1,\infty}) = 1$ almost surely. $\square$

Lemma 7.23 (Example 6.2.2 and Problem 6.7 in [KS20])**.** *Let* $\pi(m)$, $m \ge 0$, *be the limiting cluster size distribution. Assume that* $\mathbb{P}(\lim_{|j|\to\infty} |\boldsymbol{Y}_j| = 0) = 1$. *Then*

$$
\sum_{m=1}^{\infty} m\pi(m) = \vartheta^{-1} \, , \quad \sum_{m=1}^{\infty} m^2\pi(m) = \vartheta^{-1} \sum_{j\in\mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_j| > 1) \, .
$$

Proof. For the first statement, applying (2.7) and Lemma 7.22, we have,

$$
\begin{aligned}
\sum_{m=1}^{\infty} m\pi(m) &= \sum_{m=1}^{\infty} m\mathbb{P}(\mathcal{E}(\boldsymbol{Y}) = m \mid \mathcal{C}_0(\boldsymbol{Y}) = 0) \\
&= \mathbb{E}[\mathcal{E}(\boldsymbol{Y}) \mid \mathcal{C}_0(\boldsymbol{Y}) = 0] = \vartheta^{-1}\boldsymbol{\nu}^*(\mathcal{E}) = \vartheta^{-1} \, .
\end{aligned}
$$

Likewise,

$$
\sum_{m=1}^{\infty} m^2\pi(m) = \mathbb{E}[\mathcal{E}^2(\boldsymbol{Y}) \mid \mathcal{C}_0(\boldsymbol{Y}) = 0] = \vartheta^{-1}\boldsymbol{\nu}^*(\mathcal{E}^2) = \vartheta^{-1} \sum_{j\in\mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_j| > 1) \, .
$$

$\square$

References

[BB18] Betina Berghaus and Axel Bücher. Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Annals of Statistics*, 46(5):2307–2335, 2018. MR3845019

[BBKS20] Clemonell Bilayi-Biakana, Rafał Kulik, and Philippe Soulier. Statistical inference for heavy tailed series with extremal independence. *Extremes*, 23(1):1–33, 2020. MR4064603

[BPS18] Bojan Basrak, Hrvoje Planinić, and Philippe Soulier. An invariance principle for sums and record times of regularly varying stationary sequences. *Probability Theory and Related Fields*, 172(3-4):869–914, 2018. MR3877549

[BS09]   Bojan Basrak and Johan Segers. Regularly varying multivariate time series. *Stochastic Processes and their Applications*, 119(4):1055–1080, 2009. MR2508565

[BS18a] Axel Bücher and Johan Segers. Inference for heavy tailed stationary time series based on sliding blocks. *Electronic Journal of Statistics*, 12(1):1098–1125, 2018. MR3780041

[BS18b] Axel Bücher and Johan Segers. Maximum likelihood estimation for the fréchet distribution based on block maxima extracted from a time series. *Bernoulli*, 24(2):1427–1462, 2018. MR3706798

[DM09]   Richard A. Davis and Thomas Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 38A:977–1009, 2009. Probability, statistics and seismology. MR2597580

[DH95]   Richard A. Davis and Tailen Hsing. Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Annals of Probability*, 23(2):879–917, 1995. MR1334176

[DHS18] Clément Dombry, Enkelejd Hashorva, and Philippe Soulier. Tail measure and spectral tail process of regularly varying time series. *Annals of Applied Probability*, 28(6):3884–3921, 2018. MR3861829

[DK20]   Holger   Drees   and   Miran   Knezevic.   Peak-over-threshold   estimators   for   spectral   tail   processes: Random   vs   deterministic   thresholds.   *Extremes*,   2020.   DOI: https://doi.org/10.1007/s10687-019-00367-x. MR4129561

[DN20]   Holger Drees and Sebastian Neblung. Asymptotics for sliding blocks estimators of rare events. arXiv:2003.01016, 2020.

[DR10]   Holger Drees and Holger Rootzén. Limit theorems for empirical processes of cluster functionals. *Annals of Statistics*, 38(4):2145–2186, 2010. MR2676886

[DSW15] Holger Drees, Johan Segers, and Michał Warchoł. Statistics for tail processes of Markov chains. *Extremes*, 18(3):369–402, 2015. MR3377814

[EKM97] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer–Verlag, 1997. MR1458613

[Hsi91] Tailen Hsing. Estimating the parameters of rare events. *Stochastic Processes and their Applications*, 37(1):117–139, 1991. MR1091698

[Ibr62] I. A. Ibragimov. Some limit theorems for stationary processes. *Theor. Probab. Appl.*, 7:349–382, 1962. MR0148125

[KS20]   Rafał Kulik and Philippe Soulier. *Heavy tailed time series*. Springer, 2020. MR4174389

[KSW19] Rafał Kulik, Philippe Soulier, and Olivier Wintenberger. The tail empirical process of regularly varying functions of geometrically ergodic markov chains. *Stochastic Processes and their Applications*,

129(1):4209–4238, 2019. MR4013860

[MW13] Thomas Mikosch and Olivier Wintenberger. Precise large deviations for dependent regularly varying sequences. *Probability Theory and Related Fields*, 156(3-4):851–887, 2013. MR3078288

[MW14] Thomas Mikosch and Olivier Wintenberger. The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probability Theory and Related Fields*, 159(1-2):157–196, 2014. MR3201920

[MW16] Thomas Mikosch and Olivier Wintenberger. A large deviations approach to limit theorem for heavy-tailed time series. *Probability Theory and Related Fields*, 166(1-2):233–269, 2016. MR3547739

[PS18] Hrvoje Planinić and Philippe Soulier. The tail process revisited. *Extremes*, 21(4):551–579, 2018. MR3874760

[RLdH98] Holger Rootzén, Ross M. Leadbetter, and Laurens de Haan. On the distribution of tail array sums for strongly mixing stationary sequences. *Annals of Applied Probability*, 8(3):868–885, 1998. MR1627799

[RSF09] Christian Y. Robert, Johan Segers, and Christopher A. T. Ferro. A sliding blocks estimator for the extremal index. *Electronic Journal of Statistics*, 3:993–1020, 2009. MR2540849

[SW94] Richard L. Smith and Ishay Weissman. Estimating the extremal index. *Journal of the Royal Statistical Society. Series B. Methodological*, 56(3):515–528, 1994. MR1278224

[vdVW96] Aad W. van der Vaart and Jon A. Wellner. *Weak convergence and empirical processes*. Springer, New York, 1996. MR1385671

[WN98] Ishay Weissman and S. Yu. Novak. On blocks and runs estimators of the extremal index. *Journal of Stat. Planning and Inference*, 66(2):281–288, 1998. MR1614480

[ZVB20] Nan Zou, Stanislav Volgushev, and Axel Bücher. Multiple block sizes and overlapping blocks for multivariate time series extremes. arXiv:1907.09477, 2020. MR4206679