Electronic Journal of Statistics
Vol. 16 (2022) 3561–3607
ISSN: 1935-7524
https://doi.org/10.1214/22-EJS2026

Estimation of cluster functionals for regularly varying time series: Runs 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 runs estimators of cluster indices. 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 blocks and runs estimators have the same limiting variance.

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

Received September 2021.

Contents

1 Introduction . . . 3562
2 Preliminaries . . . 3567
2.1 Notation . . . 3567
2.2 Classes of functions . . . 3567
2.3 Tail and spectral tail process . . . 3568
2.4 Anchoring maps . . . 3568
2.5 Cluster measure and cluster indices . . . 3569
2.6 Convergence of cluster measure . . . 3571
2.6.1 Anticlustering conditions . . . 3571
2.6.2 Vague convergence of cluster measure . . . 3572
2.6.3 Indicator functionals not vanishing around zero . . . 3572
3 Central limit theorem for runs estimators . . . 3572
3.1 Runs estimator . . . 3573
3.2 Mixing assumptions . . . 3574
3.3 Limiting covariances . . . 3575
3.3.1 Runs statistics . . . 3575
3.3.2 Runs and disjoint blocks statistics . . . 3575
3.4 Central limit theorem . . . 3575
3.4.1 Comments on the conditions . . . 3576
4 Simulation study . . . 3577
4.1 Stationary AR process . . . 3577

4.2 Stationary ARCH process 3579
4.3 Summary 3580
5 Proofs 3581
5.1 Mixing 3582
5.2 The anticlustering conditions 3582
5.3 Conditional convergence 3583
5.4 Limiting covariances 3587
5.5 Empirical cluster process of runs statistics 3591
5.6 Proof of Theorem 3.6 3592
5.7 Proof of Theorem 5.6 – fidi convergence 3592
5.8 Proof of Theorem 5.6 – asymptotic equicontinuity 3596
5.8.1 Asymptotic equicontinuity of the empirical process of sliding blocks 3597
5.9 Random entropy 3603
Supplementary Material 3604
References 3604

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 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

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

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. If $\theta = 0$, then the limiting distribution is degenerated. This situation is commonly referred to as *long memory in extremes*. In this case, a different normalization of order $o(a_n)$ is required. See [Sam2016]. The case $\theta = 0$ is excluded from our studies. We would like to point out that there is a notion of the multivariate extremal index; see [BGTS09, Section 10.5.2].

The extremal index is just one parameter that describes clustering of extremes. Informally speaking, it arises as the limit

$$\lim_{n\to\infty} \frac{\mathbb{E}[H((\boldsymbol{X}_1, \dots, \boldsymbol{X}_{r_n})/u_n)]}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \ ,$$

for the particular choice of function $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$, and a suitable choice of the scaling sequence $u_n \to \infty$ and the block size $r_n \to \infty$. (Formally speaking, $(\boldsymbol{X}_1, \dots, \boldsymbol{X}_{r_n})$ is a random element of $(\mathbb{R}^d)^{r_n}$, while the domain of H is $(\mathbb{R}^d)^{\mathbb{Z}}$. This inconsistency will be explained later).

In particular, the extremal index is achieved by applying a suitable functional to a cluster:

$$H((\boldsymbol{X}_1,\dots,\boldsymbol{X}_{r_n})/u_n) = \mathbb{1}\{\max\{|\boldsymbol{X}_1|,\dots,|\boldsymbol{X}_{r_n}|\} > u_n\}\ .$$

That is,

$$\theta = \lim_{n\to\infty} \frac{\mathbb{P}(\max\{|\boldsymbol{X}_1|,\dots,|\boldsymbol{X}_{r_n}|\} > u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\ . \tag{1.1}$$

Informally speaking, a cluster is a triangular array $(\boldsymbol{X}_1/u_n,\dots,\boldsymbol{X}_{r_n}/u_n)$ 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 $i \le j \in \mathbb{Z}$, we denote $\boldsymbol{X}_{i,j} = (\boldsymbol{X}_i,\dots,\boldsymbol{X}_j) \in (\mathbb{R}^d)^{(j-i+1)}$. 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

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

The sequence u_n will play a role of a threshold (that is, only the values of $|\boldsymbol{X}_t|$ that exceed u_n play a role), while r_n will be a block size in the formulation of the runs estimators. Given a cluster functional H on $(\mathbb{R}^d)^{\mathbb{Z}}$, we want to estimate the limiting quantity

$$\boldsymbol{\nu}^*(H) = \lim_{n\to\infty} \boldsymbol{\nu}^*_{n,r_n}(H) = \lim_{n\to\infty} \frac{\mathbb{E}[H(\boldsymbol{X}_{1,r_n}/u_n)]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\ . \tag{1.2}$$

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

$$H_{2,m}(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j\in\mathbb{Z}} \mathbb{1}\big\{|\boldsymbol{x}_j| > 1\big\} = m\right\}\ ,\quad \boldsymbol{x} = \{\boldsymbol{x}_j, j\in\mathbb{Z}\} \in (\mathbb{R}^d)^{\mathbb{Z}}\ , m\in\mathbb{N}\ ; \tag{1.3}$$

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

$$H_{3,\eta}(\boldsymbol{x}) = \mathbb{1}\left\{\sum_{j\in\mathbb{Z}} (x_j - 1)_+ > \eta\right\}\ ,\quad \boldsymbol{x} = \{\boldsymbol{x}_j, j\in\mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}}\ ,\quad \eta > 0\ ; \tag{1.4}$$

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

$$H_4(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\} \ , \quad K(\boldsymbol{x}) = \left(\sum_{j\in\mathbb{Z}} x_j\right)_+ \ , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j\in\mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}} \ ; \tag{1.5}$$

- the ruin index of a univariate time series obtained with

$$H_5(\boldsymbol{x}) = \mathbb{1}\{K(\boldsymbol{x}) > 1\} \ , \quad K(\boldsymbol{x}) = \sup_{i\in\mathbb{Z}} \left(\sum_{j\leq i} x_j\right)_+ \ , \quad \boldsymbol{x} = \{\boldsymbol{x}_j, j\in\mathbb{Z}\} \in \mathbb{R}^{\mathbb{Z}}. \tag{1.6}$$

As indicated above, the extremal index is the classical quantity that arises in the extreme value theory for dependent sequences. Similarly, the cluster size distribution has been 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.

Several methods of estimation of the limit $\boldsymbol{\nu}^*(H)$ in (1.2) 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 $m_n = [n/r_n]$ is the number of disjoint blocks. 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

$$\widehat{\boldsymbol{\nu}}^*_{n,r_n}(H) := \frac{1}{k_n} \sum_{i=1}^{m_n} H(\boldsymbol{X}_{(i-1)r_n+1,ir_n}/|\boldsymbol{X}|_{(n:n-k_n)}) \ . \tag{1.7}$$

The general asymptotic theory for disjoint blocks estimators was developed in [DR10]. See also [KS20, Chapter 10]. The limiting variance of the disjoint blocks estimator can be represented as

$$\boldsymbol{\nu}^*(\{H - \boldsymbol{\nu}^*(H)\mathcal{E}\}^2) \ , \tag{1.8}$$

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].

Another approach to estimation of $\boldsymbol{\nu}^*(H)$ is to consider the sliding blocks statistics

$$\widetilde{\boldsymbol{\mu}}^*_{n,r_n}(H) := \frac{1}{q_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/u_n\right) \tag{1.9}$$

and the corresponding estimator defined in terms of order statistics:

$$\widehat{\boldsymbol{\mu}}^*_{n,r_n}(H) = \frac{1}{r_n k_n} \sum_{i=0}^{q_n-1} H\left(\boldsymbol{X}_{i+1,i+r_n}/|\boldsymbol{X}|_{(n:n-k_n)}\right) . \tag{1.10}$$

Here, $q_n = n - r_n - 1$ is the number of sliding blocks. In [DN21] the authors 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. In [CK21] it was shown that the limiting variances for both disjoint and sliding blocks estimators agree and are given by the expression in (1.8) for an arbitrary choice of H. We note at this point that the methodology used in [DR10, DN21, KS20, CK21] fits into Peak Over Threshold (PoT) framework. On the other hand, in the Block Maxima (BM) framework, sliding blocks estimators yield typically smaller variance; see [BS18b, BS18a]. As of this moment, there is no thorough explanation of these phenomena and no formal comparison between PoT and BM framework. See [FdH15] for some partial results and [BZ21] for a recent review.

In the present paper we are interested in the so-called *runs estimators*. In the context of the extremal index, this approach goes back to [WN98] and stems from the following representation of the extremal index:

$$\theta = \lim_{n\to\infty} \mathbb{P}(\max\{|\boldsymbol{X}_1|, \dots, |\boldsymbol{X}_{r_n}|\} \le u_n \mid |\boldsymbol{X}_0| > u_n) . \tag{1.11}$$

We note that

$$\begin{aligned}&\mathbb{P}(\max\{|\boldsymbol{X}_1|, \dots, |\boldsymbol{X}_{r_n}|\} \le u_n \mid |\boldsymbol{X}_0| > u_n)\\ &= \frac{1}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \mathbb{E}[\mathbb{1}\{\mathcal{C}((\boldsymbol{X}_0, \dots, \boldsymbol{X}_{r_n})/u_n) = 0\}\mathbb{1}\{|\boldsymbol{X}_0| > u_n\}] ,\end{aligned}$$

where

$$\mathcal{C}(\boldsymbol{x}) = \sup\{j : |\boldsymbol{x}_j| > 1\}$$

gives the position of the last exceedence above 1 in a particular block. Recall again the convention $\mathcal{C}((\boldsymbol{X}_0, \dots, \boldsymbol{X}_{r_n})/u_n) = \mathcal{C}((\boldsymbol{0}, \boldsymbol{X}_0, \dots, \boldsymbol{X}_{r_n})/u_n)$. The latter expression provides an alternative representation for the extremal index. Then, $\mathcal{C}$ is an example of so-called *anchoring map*. Special cases of anchoring maps were considered in [Has18] and [BP18], while in [KS20] their connection to cluster indices $\boldsymbol{\nu}^*(H)$ was thoroughly investigated. It turns out that with an arbitrary choice of the anchoring map $\mathcal{C}$ we have

$$\boldsymbol{\nu}^*(H) = \mathbb{E}[H^{\mathcal{C}}(\boldsymbol{Y})] ,$$

where

$$H^{\mathcal{C}}(\boldsymbol{x}) = H(\boldsymbol{x})\mathbb{1}\{\mathcal{C}(\boldsymbol{x}) = 0\}\mathbb{1}\{|\boldsymbol{x}_0| > 1\}$$

and $\boldsymbol{Y}$ is the tail process, a distributional limit (as $x \to \infty$) of $\boldsymbol{X}/x$ given $|\boldsymbol{X}_0| > x$; see Section 2.3 for the precise definition.

This motivates the following runs statistics:

$$\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \frac{1}{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \sum_{i=r_n+1}^{n-r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) . \tag{1.12}$$

Indeed, under the appropriate conditions, Proposition 2.7 gives

$$\lim_{n\to\infty} \mathbb{E}[\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})] = \mathbb{E}[H^{\mathcal{C}}(\boldsymbol{Y})] = \boldsymbol{\nu}^*(H) .$$

The data-based runs estimator is then

$$\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \frac{1}{k_n} \sum_{i=r_n+1}^{n-r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/|\boldsymbol{X}|_{(n:n-k)}\right) .$$

The main result of this paper is Theorem 3.6, the asymptotic normality of the appropriately normalized estimator $\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$. We show, in particular, that the limiting variance agrees with the one for the disjoint blocks and sliding blocks estimators; cf. [DR10], [KS20, Chapter 10], [CK21]. This is also in accordance with the result of [DN21] for the extremal index.

Furthermore, we prove that we cannot achieve variance reduction by considering a linear combination of runs estimators with a different choice of anchoring maps $\mathcal{C}$ and $\tilde{\mathcal{C}}$. Indeed, it turns out that $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$ and $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\tilde{\mathcal{C}}})$ are totally dependent in the limit. We note in passing that even though general ideas of proofs are similar to those of [CK21], however, technicalities are significantly different. Differences stem primarily from conditioning on $\{|\boldsymbol{X}_j| > u_n\}$ used in case of the runs estimators.

Thus, from the theoretical point of view the limiting behaviour of all (disjoint blocks, sliding blocks, runs) estimators is the same. However, for finite samples a bias has to be taken into account. We note first that the theoretical finite-sample bias for both disjoint and sliding blocks estimators is the same. This can be also seen in extensive simulation studies in [CK21]. On the other hand, we were not able to get an useful formula for the bias in the runs estimator case. As such we relied on simulations. It turns out that runs estimator are typically heavily biased when estimation of the extremal index is concerned. However, the runs estimators may have an advantage when other cluster indices are considered.

The paper is structured as follows. Section 2 contains definitions, notation and preliminary results on convergence of clusters. It is primarily based on [KS20, Chapters 5 and 6], with some results from [BS09], [BPS18], [PS18]. Section 3 defines runs pseudo-estimators and estimators. The main result of the paper is the central limit theorem for runs estimators in Theorem 3.6. We note again that the limiting variance agrees with the one for disjoint and sliding blocks estimators. A small simulation study is performed in Section 4. All the proofs are contained in Section 5.

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]). Section 2.4 introduces anchoring maps (cf. [BP18], [Has18]). In Section 2.5 we define cluster indices. We refer to [KS20, Chapter 5] for more details. In Section 2.6 we discuss convergence of the cluster measure, following [KS20, Chapter 6].

The most important conclusion of these preliminaries is a representation of the cluster index $\boldsymbol{\nu}^*(H)$ (cf. (1.2)) as $\mathbb{E}[H^{\mathcal{C}}(\boldsymbol{Y})]$, with $H^{\mathcal{C}}$ defined in (2.7) and $\boldsymbol{Y}$ being the tail process. Also, Proposition 2.7 on conditional weak convergence and Propositions 2.8 and 2.10 on unconditional weak convergence play a central role in the rest of the paper.

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 $\boldsymbol{x}_{i,j} = (\boldsymbol{x}_i, \dots, \boldsymbol{x}_j) \in (\mathbb{R}^d)^{j-i+1}$, $\boldsymbol{x}^*_{i,j} = \max_{i \le l \le j} |\boldsymbol{x}_l|$ 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 $\sum_{j\in\mathbb{Z}} |\boldsymbol{x}_j| < \infty$.

2.2. Classes of functions

Functionals H are defined on $\ell_0(\mathbb{R}^d)$ with the convention $H(\boldsymbol{x}_{i,j}) = H((\boldsymbol{0}, \boldsymbol{x}_{i,j}, \boldsymbol{0}))$. 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 $\boldsymbol{0}$. In particular, for $H \in \mathcal{A}$, $H(\boldsymbol{0}) = 0$. 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 $K(\boldsymbol{0}) = 0$ and which are Lipschitz continuous with constant L_K, i.e.

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

- $\mathcal{B} \subset \mathcal{L}$ 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 $H \notin \mathcal{A}$.

We will also need the map $\mathcal{E}$ is defined on $\ell_0(\mathbb{R}^d)$ by $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\left\{|\boldsymbol{x}_j| > 1\right\}$. Note that $\mathcal{E}$ is shift-invariant, with the support separated from zero, but is not bounded.

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,

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

The spectral tail process $\{\boldsymbol{\Theta}_j, j \in \mathbb{Z}\}$ is defined by $\boldsymbol{\Theta} = |\boldsymbol{Y}_0|^{-1}\boldsymbol{Y}$ and is independent of the tail process $\boldsymbol{Y}$.

2.4. Anchoring maps

Definition 2.1 (Anchoring map). *A measurable map $\mathcal{C} : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{Z}\cup\{-\infty, \infty\}$ is called an anchoring map if the following two properties hold:*

An(i): *$\mathcal{C}(\boldsymbol{x}) = j$ implies $|\boldsymbol{x}_j| \ge |\boldsymbol{x}_0| \wedge 1$;*
An(ii): *$\mathcal{C}(B\boldsymbol{x}) = \mathcal{C}(\boldsymbol{x}) + 1$, where B is a backsift operator.*

Three basic examples of anchoring maps are:

- The infargmax functional: $\mathcal{C}^{(0)}(\boldsymbol{y}) = \inf\{j : \boldsymbol{y}^*_{-\infty,j} = \boldsymbol{y}^*\}$;
- The first exceedence above one: $\mathcal{C}^{(1)}(\boldsymbol{x}) = \inf\{j : |\boldsymbol{x}_j| > 1\}$;
- The last exceedence above one: $\mathcal{C}^{(2)}(\boldsymbol{x}) = \sup\{j : |\boldsymbol{x}_j| > 1\}$.

In what follows we use the convention $\inf\emptyset = +\infty$. We note that $\mathcal{C}^{(0)}$ is 0-homogeneous, that is, $\mathcal{C}_s^{(0)}(\boldsymbol{x}) = \mathcal{C}^{(0)}(\boldsymbol{x})$ for all $s > 0$, while $\mathcal{C}_s^{(1)}(\boldsymbol{x}) = \mathcal{C}^{(1)}(\boldsymbol{x}/s)$ and $\mathcal{C}_s^{(2)}(\boldsymbol{x}) = \mathcal{C}^{(2)}(\boldsymbol{x}/s)$ are increasing and decreasing in s, respectively, but they are not 0-homogeneous. This will play a role in the proofs.

A special importance is given to the time index 0. In particular,

- If $\mathcal{C}^{(0)}(\boldsymbol{x}) = 0$, then $\boldsymbol{x}^*_{-\infty,-1} < |\boldsymbol{x}_0|$ and $\boldsymbol{x}^*_{1,\infty} \le |\boldsymbol{x}_0|$;
- If $\mathcal{C}^{(1)}(\boldsymbol{x}) = 0$, then $\boldsymbol{x}^*_{-\infty,-1} \le 1$ and $|\boldsymbol{x}_0| > 1$;
- If $\mathcal{C}^{(2)}(\boldsymbol{x}) = 0$, then $\boldsymbol{x}^*_{1,\infty} \le 1$ and $|\boldsymbol{x}_0| > 1$.

Applying an anchoring map to a finite block, say $\boldsymbol{x}_{-r,r}$ with $r \in \mathbb{N}$, is equivalent to applying it $\boldsymbol{x} = (\boldsymbol{0}, \boldsymbol{x}_{-r,r}, \boldsymbol{0})$. For example, $\mathcal{C}^{(0)}(\boldsymbol{x}_{-r,r}) = 0$ means that $\boldsymbol{x}^*_{-r,-1} < |\boldsymbol{x}_0|$ and $\boldsymbol{x}^*_{1,r} \le |\boldsymbol{x}_0|$. This in turn implies also that $\mathcal{C}^{(0)}(\boldsymbol{x}_{-s,s}) = 0$ for $0 < s < r$. Similarly,

$$\mathcal{C}(\boldsymbol{x}_{-r,r}) = 0 \Rightarrow \mathcal{C}(\boldsymbol{x}_{-s,s}) = 0\,, \quad 0 < s < r \tag{2.1}$$

for $\mathcal{C} = \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$. However, we do not know if this property (used explicitly in the proofs) holds for any anchoring map. Furthermore, there are crucial monotonicity and homogeneity properties, indicated above, used explicitly in the proofs (of tightness). As such, in the paper we focus on the three anchoring maps introduced above. We will clearly indicate which results hold for an arbitrary anchoring map, and which are specific to any of $\mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$.

Furthermore, note that An(ii) gives that

$$\mathcal{C}(\boldsymbol{x}_{h-r,h+r}) = h \Leftrightarrow \mathcal{C}(B^{-h}\boldsymbol{x}_{-r,r}) = 0\,. \tag{2.2}$$

Indeed, consider for example $\mathcal{C}^{(0)}$. Then $\mathcal{C}^{(0)}(\boldsymbol{x}_{h-r,h+r}) = h$ means that $\boldsymbol{x}^*_{h-r,h-1} < |\boldsymbol{x}_h|$ and $\boldsymbol{x}_{h+1,h+r} \le |\boldsymbol{x}_h|$. Set $\tilde{\boldsymbol{x}} = B^{-h}\boldsymbol{x}$, so that $\tilde{\boldsymbol{x}}_{-r} = \boldsymbol{x}_{h-r}$, $\tilde{\boldsymbol{x}}_0 = \boldsymbol{x}_h$ and $\tilde{\boldsymbol{x}}_r = \boldsymbol{x}_{h+r}$. Thus, $\tilde{\boldsymbol{x}}_{-r,-1} < |\tilde{\boldsymbol{x}}_0|$ and $\tilde{\boldsymbol{x}}_{1,r} \le |\tilde{\boldsymbol{x}}_0|$. This in turn means that $\mathcal{C}^{(0)}(\widetilde{\boldsymbol{x}}_{-r,r}) = \mathcal{C}^{(0)}(B^{-h}\boldsymbol{x}_{-r,r}) = 0$. The similar argument applies to the other two anchoring maps.

2.5. Cluster measure and cluster indices

Let $\mathcal{C}$ be an anchoring map. If $\mathbb{P}(\mathcal{C}(\boldsymbol{Y}) \notin \mathbb{Z}) = 0$ then we can define

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

We want to emphasize that ϑ does not depend on the choice of the anchoring map (see [PS18] and [KS20, Corollary 5.5.4]). In particular,

$$\vartheta = \mathbb{P}(\mathcal{C}^{(1)}(\boldsymbol{Y}) = 0) = \mathbb{P}\left(\sup_{j\le -1}|\boldsymbol{Y}_j| \le 1\right) = \mathbb{P}(\mathcal{C}^{(2)}(\boldsymbol{Y}) = 0) = \mathbb{P}\left(\sup_{j\ge 1}|\boldsymbol{Y}_j| \le 1\right).$$

The above identity follows from the time-change formula, see [CK21, Section 7.1]. Therefore, ϑ can be recognized as the (candidate) extremal index. It becomes the usual extremal index under additional mixing and anticlustering conditions (cf. Section 7.5 in [KS20]).

Recall that $\mathcal{E}(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \mathbb{1}\{|\boldsymbol{x}_j| > 1\}$. The property An(i) of the anchoring maps implies

$$\sum_{h\in\mathbb{Z}} \mathbb{P}(\mathcal{C}(\boldsymbol{Y}) = h) \le \sum_{h\in\mathbb{Z}} \mathbb{P}(|\boldsymbol{Y}_h| > 1) \ . \tag{2.4}$$

By [KS20, Lemma 9.2.3] the latter series is finite if an appropriate anticlustering condition holds (see $\mathcal{S}(r_n, u_n)$ to be introduced later on).

Definition 2.2 (Cluster measure)**.** *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*

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

where δ is the Dirac measure.

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

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 for example that $\mathcal{C}^{(0)}(\boldsymbol{Y}) = 0$ gives $\boldsymbol{Y}^* = |\boldsymbol{Y}_0|$. Thus, (2.5) and the definition of $\boldsymbol{Q}$ give for a bounded or non-negative measurable function H on $\ell_0(\mathbb{R}^d)$ (see Definition 5.4.11 in [KS20]),

$$\boldsymbol{\nu}^*(H) = \vartheta \int_0^\infty \mathbb{E}[H(r\boldsymbol{Q})] \alpha r^{-\alpha-1} \mathrm{d}r \ .$$

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

$$\boldsymbol{\nu}^*(H) = \epsilon^{-\alpha} \mathbb{E}[H(\epsilon \boldsymbol{Y}) \mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}] \ . \tag{2.6}$$

For a shift-invariant $H : (\mathbb{R}^d)^{\mathbb{Z}} \to \mathbb{R}$ and an anchoring map $\mathcal{C}$ define

$$H^{\mathcal{C}}(\boldsymbol{x}) = H(\boldsymbol{x}) \mathbb{1}\{\mathcal{C}(\boldsymbol{x}) = 0\} \mathbb{1}\{|\boldsymbol{x}_0| > 1\} \ . \tag{2.7}$$

Thus, since $|\boldsymbol{Y}_0| > 1$, if H is such that $H(\boldsymbol{y}) = 0$ whenever $\boldsymbol{y}^* \le 1$, then (2.6) gives

$$\boldsymbol{\nu}^*(H) = \mathbb{E}[H^{\mathcal{C}}(\boldsymbol{Y})] \ . \tag{2.8}$$

Note that the $\boldsymbol{\nu}^*(H)$ does not agree with $\mathbb{E}[H(\boldsymbol{Y})]$.

Definition 2.4 (Cluster index)**.** *We will call $\boldsymbol{\nu}^*(H)$ the cluster index associated to the functional H.*

2.6. Convergence of cluster measure

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].

2.6.1. Anticlustering conditions

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 $\boldsymbol{Y}_{-r,r}$ (see Section 2.3). 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 2.5 ([DH95], Condition 2.8)**.** *Condition $\mathcal{AC}(r_n, u_n)$ holds if for all $x, y > 0$,*

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

Condition $\mathcal{AC}(r_n, u_n)$ is referred to as the anticlustering condition. It holds for i.i.d. regularly varying sequences if $\lim_{n\to\infty} r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n) = 0$. Note that the latter condition is a part of the $\mathcal{R}(r_n, u_n)$ assumption. It is also fulfilled by many models, including geometrically ergodic Markov chains, short-memory linear or max-stable processes. $\mathcal{AC}(r_n, u_n)$ implies that $\boldsymbol{Y} \in \ell_0(\mathbb{R}^d)$. See [KSW19] and [KS20].

A stronger version of the anticlustering condition reads as follows.

Definition 2.6. *Condition $\mathcal{S}(r_n, u_n)$ holds if for all $s, t > 0$*

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

The main consequence of the anticlustering condition $\mathcal{AC}(r_n, u_n)$ is the following result.

Proposition 2.7 ([BS09], Proposition 4.2; [KS20], Theorem 6.1.4)**.** *Let $H \in \mathcal{L}$. If Condition $\mathcal{AC}(r_n, u_n)$ holds, then*

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

2.6.2. Vague convergence of cluster measure

We now state the unconditional convergence of $u_n^{-1}\boldsymbol{X}_{1,r_n}$. Contrary to Proposition 2.7, 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. The following result is a re-formulation of Theorem 6.2.5 in [KS20].

Proposition 2.8. *Let condition $\mathcal{AC}(r_n,u_n)$ hold. 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)):

$$\lim_{n\to\infty}\frac{\mathbb{P}(\boldsymbol{X}^*_{1,r_n}>u_n)}{r_n\mathbb{P}(|\boldsymbol{X}_0|>u_n)}=\vartheta\ . \tag{2.9}$$

2.6.3. Indicator functionals not vanishing around zero

Proposition 2.8 entails convergence of $\boldsymbol{\nu}^*_{n,r_n}(H)$ for $H\in\mathcal{A}$. For functionals which are not defined on the whole space $\ell_0(\mathbb{R}^d)$ we need an additional assumption on Asymptotic Negligibility of Small Jumps.

Definition 2.9. *Condition ANSJB(r_n,u_n) holds if for all $\eta>0$,*

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

Proposition 2.10 (Theorem 6.2.16 in [KS20])**.** *Assume that $\mathcal{AC}(r_n,u_n)$ and ANSJB(r_n,u_n) hold. Then for $K\in\mathcal{K}$,*

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

3. Central limit theorem for runs estimators

In this section we introduce and study runs estimators of cluster indices. A pseudo-estimator is defined in (3.4). Its limiting covariance (for different anchoring maps) is studied in Lemma 3.3. In particular, for two different anchoring maps, the runs statistics are totally dependent. As a consequence we cannot reduce the limiting variance for the estimation of $\boldsymbol{\nu}^*(H)$ by considering linear combinations of the runs statistics. In Lemma 3.4 we consider covariance between runs and disjoint blocks estimators. Again, we obtain total dependence in the limit. The main result of the paper is the central limit theorem for runs estimators; see Theorem 3.6. The limiting variance agrees with the one for the disjoint blocks and sliding blocks estimators.

3.1. Runs estimator

To introduce runs estimators recall that (cf. (2.2))

$$\begin{aligned}
&H^{\mathcal{C}}\left(\boldsymbol{X}_{(j-1)r_n+h,(j+1)r_n+h}/u_n\right) = H^{\mathcal{C}}(B^{-h-jr_n}\boldsymbol{X}_{-r_n,r_n}/u_n) \\
&= H(B^{-h-jr_n}\boldsymbol{X}_{-r_n,r_n}/u_n)\mathbb{1}\left\{\mathcal{C}(B^{-h-jr_n}\boldsymbol{X}_{-r_n,r_n}/u_n)=0\right\}\mathbb{1}\left\{\left|B^{-h-jr_n}\boldsymbol{X}_0\right|>u_n\right\} \\
&= H\left(\boldsymbol{X}_{(j-1)r_n+h,(j+1)r_n+h}/u_n\right)\times \\
&\qquad \mathbb{1}\left\{\mathcal{C}(\boldsymbol{X}_{(j-1)r_n+h,(j+1)r_n+h}/u_n)=h+jr_n\right\}\mathbb{1}\left\{|\boldsymbol{X}_{h+jr_n}|>u_n\right\}. \qquad (3.1)
\end{aligned}$$

Set $q_n = n - r_n$ and $m_n = n/r_n$. Without loss of generality assume that m_n is an integer. Consider disjoint blocks

$$J_j := \{jr_n+1,\dots,(j+1)r_n\}\,,\quad j=0,\dots,m_n-1\,. \qquad (3.2)$$

The union of these blocks gives $\{1,\dots,n\}$. We assume we have data $\boldsymbol{X}_{1-r_n},\dots,\boldsymbol{X}_{n+r_n}$. For $j=0,\dots,m_n-1$ define

$$\begin{aligned}
H^{\mathcal{C}}_{n,j} &= \sum_{i=jr_n+1}^{(j+1)r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) = \sum_{i=jr_n+1}^{(j+1)r_n} H^{\mathcal{C}}\left(B^{-i}\boldsymbol{X}_{-r_n,r_n}/u_n\right) \\
&= \sum_{i=jr_n+1}^{(j+1)r_n} H\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)=i\}\mathbb{1}\{|\boldsymbol{X}_i|>u_n\}\,.
\end{aligned} \qquad (3.3)$$

Each $H^{\mathcal{C}}_{n,j}$ is a function of the block $\boldsymbol{X}_{(j-1)r_n+1,\dots,(j+2)r_n}$ of size $3r_n$. The number j in the notation $H^{\mathcal{C}}_{n,j}$ indicates that the indicator $|\boldsymbol{X}_i|>u_n$ is applied with $i\in J_j$.

We consider a random process

$$\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \frac{1}{n\mathbb{P}(|\boldsymbol{X}_0|>u_n)}\sum_{i=1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) \qquad (3.4)$$

that can be decomposed as

$$\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \frac{1}{n\mathbb{P}(|\boldsymbol{X}_0|>u_n)}\sum_{j=0}^{m_n-1} H^{\mathcal{C}}_{n,j}\,.$$

If the anticlustering condition $\mathcal{AC}(r_n,u_n)$ holds, then using stationarity, definition (2.7) of $H^{\mathcal{C}}$, Proposition 2.7 and (2.8) we have

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}[\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})] &= \lim_{n\to\infty}\frac{1}{\mathbb{P}(|\boldsymbol{X}_0|>u_n)}\mathbb{E}\left[H^{\mathcal{C}}(\boldsymbol{X}_{-r_n,r_n}/u_n)\right] \\
&= \mathbb{E}[H^{\mathcal{C}}(\boldsymbol{Y})] = \boldsymbol{\nu}^*(H)\,.
\end{aligned}$$

Now, let k_n be a sequence of integers (depending on n) such that $k_n \to \infty$ and $k_n/n \to 0$. Define u_n by $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$ and replace u_n in

$H^{\mathcal{C}}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n)$ with (k_n+1)th order statistics $|\boldsymbol{X}|_{(n:n-k_n)}$ to get the runs estimator:

$$\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \frac{1}{k_n}\sum_{i=1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/|\boldsymbol{X}|_{(n:n-k_n)}\right) . \tag{3.5}$$

In what follows we will use interchangeably k_n and $n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$, whatever is more suitable.

3.2. Mixing assumptions

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 3.1. *Condition $\beta'(r_n)$ holds if:*

$$\lim_{n\to\infty}\left\{\frac{n}{r_n} + \frac{n}{k_n}\right\}\beta_{r_n} = 0 , \tag{3.6a}$$

$$\lim_{n\to\infty}\frac{1}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\sum_{i=r_n+1}^{\infty}\beta_i = \lim_{n\to\infty}\frac{n}{k_n}\sum_{i=r_n+1}^{\infty}\beta_i = 0 , \tag{3.6b}$$

$$\lim_{n\to\infty}\frac{1}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\sum_{j=1}^{\infty}\beta_{jr_n} = \lim_{n\to\infty}\frac{n}{r_n k_n}\sum_{j=1}^{\infty}\beta_{jr_n} = 0 . \tag{3.6c}$$

Remark 3.2. We note first that $\mathcal{R}(r_n, u_n)$ gives $r_n k_n/n \to 0$.

- Assume first that $\beta_j = O(j^{-\gamma})$, $\gamma > 1$. Then $\beta'(r_n)$ reduces to $n/r_n^{1+\gamma} + n/(r_n^{\gamma}k_n) \to 0$ and $(nr_n^{1-\gamma})/k_n \to 0$. Choose $r_n = n^{\delta_1}$, $k_n = n^{\delta_2}$, $\delta_1 < \delta_2$, $\delta_1 + \delta_2 < 1$. Then all the conditions hold if $\delta_1 > 1/(1+\gamma)$ and $\delta_2 > 1 + \delta_1(1-\gamma)$. In other words, there are little restrictions on r_n and k_n if γ is big enough, that is the dependence in the time series is weak. We note further that the choice $r_n = (\log n)^{\delta}$, $\delta > 0$, is not allowed. Examples of processes with beta-mixing coefficients $\beta_j = O(j^{-\gamma})$ include: max-moving averages (Theorem 13.5.5 in [KS20]); infinite order moving averages with regularly varying innovations (Lemma 15.3.1 in [KS20]).
- Assume that $\beta_j = O(\exp(-\gamma j))$, $\gamma \geq 1$. Then we can choose $k_n = n^{\delta_2}$, $\delta_2 > 0$, and $r_n = (\log n)^{\delta_1}$, $\delta_1 \geq 1$. Examples of processes with beta-mixing coefficients $\beta_j = O(\exp(-\gamma j))$ include: subordinated Gaussian max-stable processes (Example 13.5.4 in [KS20]); geometrically ergodic Markov chains (Section 14.3 in [KS20]).

3.3. Limiting covariances

3.3.1. Runs statistics

The first result deals with covariance of the process $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}$ defined in (3.4).

Lemma 3.3. *Assume $\mathcal{R}(r_n,u_n)$, $\mathcal{AC}(r_n,u_n)$, $\mathcal{S}(r_n,u_n)$ and (3.6b) hold. Let $H,\widetilde{H}\in\mathcal{L}$ and $\mathcal{C},\widetilde{\mathcal{C}}$ be any of the anchoring maps $\mathcal{C}^{(0)},\mathcal{C}^{(1)},\mathcal{C}^{(2)}$. Then*

$$\lim_{n\to\infty} n\mathbb{P}(|\boldsymbol{X}_0|>u_n)\mathrm{cov}\left(\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}),\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(\widetilde{H}^{\widetilde{\mathcal{C}}})\right)=\boldsymbol{\nu}^*(H\widetilde{H})\;. \tag{3.7}$$

We note that the limit does not depend on the choice of the anchoring maps. In other words, for two different anchoring maps, $\mathcal{C}$ and $\widetilde{\mathcal{C}}$, the runs statistics $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$ and $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\widetilde{\mathcal{C}}})$ are totally dependent. As a consequence we cannot reduce the limiting variance for the estimation of $\boldsymbol{\nu}^*(H)$ by considering a linear combination of $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$ and $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\widetilde{\mathcal{C}}})$.

3.3.2. Runs and disjoint blocks statistics

We analyse covariance between $\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$ defined in (3.4) and the disjoint blocks statistics

$$\frac{1}{n\mathbb{P}(|\boldsymbol{X}_0|>u_n)}\sum_{j=0}^{m_n-1}H\left(\boldsymbol{X}_{jr_n+1,(j+1)r_n}/u_n\right)=\widetilde{\boldsymbol{\nu}}^*_{n,r_n}(H)\;. \tag{3.8}$$

The disjoint blocks statistics are considered in [KS20, Chapter 10].

Lemma 3.4. *Assume $\mathcal{R}(r_n,u_n)$, $\mathcal{AC}(r_n,u_n)$, $\mathcal{S}(r_n,u_n)$ and (3.6c) hold. Let $H,\widetilde{H}\in\mathcal{L}$, $\widetilde{H}(\boldsymbol{0})=0$ and $\mathcal{C}$ be any of the anchoring maps $\mathcal{C}^{(0)},\mathcal{C}^{(1)},\mathcal{C}^{(2)}$. Then*

$$\lim_{n\to\infty} n\mathbb{P}(|\boldsymbol{X}_0|>u_n)\mathrm{cov}\left(\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}),\widetilde{\boldsymbol{\nu}}^*_{n,r_n}(\widetilde{H})\right)=\boldsymbol{\nu}^*(H\widetilde{H})\;. \tag{3.9}$$

Again, irrespectively of the choice of the anchoring map $\mathcal{C}$, the runs and disjoint blocks statistics are totally dependent and we cannot reduce the limiting variance by considering their linear combinations.

3.4. Central limit theorem

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

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

Recall that for a functional $H:(\mathbb{R}^d)^{\mathbb{Z}}\to\mathbb{R}_+$ and $s>0$ we define $H_s(\boldsymbol{x})=H(\boldsymbol{x}/s)$. Also, recall that $\mathcal{E}(\boldsymbol{x})=\sum_{j\in\mathbb{Z}}\mathbb{1}\{|\boldsymbol{x}_j|>1\}$.

Consider the class

$$\mathcal{G}=\{H_s^{\mathcal{C}},s\in[s_0,t_0]\}=\{H(\boldsymbol{x}/s)\mathbb{1}\{\mathcal{C}(\boldsymbol{x}/s)=0\}\mathbb{1}\{|\boldsymbol{x}_0|>s\},\;\; s\in[s_0,t_0]\}.$$

We need the following assumption on its random entropy.

Assumption 3.5. *There exists a random metric d_n on $\mathcal{G}$ and 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,*

$$\int_0^{\delta_n} \sqrt{\log N^*(\mathcal{G}, d_n, \epsilon)} \mathrm{d}\epsilon \xrightarrow{\mathbb{P}} 0 \; . \tag{3.10}$$

The main result of this paper is Theorem 3.6, the asymptotic normality of the appropriately normalized estimator $\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$. The limiting variance agrees with the one for the disjoint blocks and sliding blocks estimators; cf. [DR10], [KS20, Chapter 10], [CK21].

Theorem 3.6. *Let $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying $\mathbb{R}^d$-valued time series. Assume that $\mathcal{R}(r_n, u_n)$, $\beta'(r_n)$, $\mathcal{S}(r_n, u_n)$ and*

$$\lim_{n\to\infty} \frac{r_n}{\sqrt{k_n}} = \lim_{n\to\infty} \frac{r_n}{\sqrt{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}} = 0 \tag{3.11}$$

hold. Suppose that Assumption 3.5 is satisfied. Fix $0 < s_0 < 1 < t_0 < \infty$. Assume moreover that for $\mathcal{C} = \mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$,

$$\lim_{n\to\infty} \sqrt{k_n} \sup_{s\in[s_0,t_0]} \left| \frac{\mathbb{P}(|\boldsymbol{X}_0| > u_n s)}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)} - s^{-\alpha} \right| = 0 \; , \tag{3.12a}$$

$$\lim_{n\to\infty} \sqrt{k_n} \sup_{s\in[s_0,t_0]} |\mathbb{E}[\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H_s^{\mathcal{C}})] - \boldsymbol{\nu}^*(H_s)| = 0 \; . \tag{3.12b}$$

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

$$\sqrt{k_n} \left\{ \widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) - \boldsymbol{\nu}^*(H) \right\} \xrightarrow{\mathrm{d}} \mathbb{G}(H - \boldsymbol{\nu}^*(H)\mathcal{E}) \; . \tag{3.13}$$

If moreover ANSJB(r_n, u_n) is satisfied, then (3.13) holds for $H \in \mathcal{B}$.

3.4.1. Comments on the conditions

One chooses typically $k_n = n^\epsilon$ with some $\epsilon \in (0, 1)$. We note that $\beta'(r_n)$ holds if e.g. $\beta_n = O(n^{-\delta})$ with $\delta > 1$ big enough or if β_n decays logarithmically. In the latter case, we typically choose $r_n = (\log n)^{1+\delta}$ with some $\delta > 0$. Recalling the choice of k_n we can see that (3.11) is not a very stringent assumption.

Furthermore, (3.12a) controls the bias in the tail empirical process and can be related to the classical second order assumptions.

Assumption 3.5 controls the size of the class $\mathcal{G}$. We are not able to provide a general set of conditions under which this condition is satisfied, however, we will verify it for virtually all functionals H that appeared in the paper. See Section 5.9.

4. Simulation study

We conducted extensive simulations in order to study the finite sample performance of the runs estimators for selected cluster indices. We compare their performance with the disjoint and sliding blocks estimators (see [CK21]). Recall that the limiting variances are the same for all estimators. We do not have theoretical formulas for bias. We note that the bias for disjoint and sliding blocks estimators is the same, but different for runs estimators. We present an extensive discussion regarding the choice of the size of the block r_n and the number of order statistics k.

We include the most important findings below. An extensive discussion, supported by a number of graphs, is included in the supplementary material [CK22].

4.1. Stationary AR process

We start with a simple AR(1) process defined by $X_{j+1} = \rho X_j + Z_{j+1}$, where $\rho \in (0,1)$ and $\{Z_j, j \in \mathbb{Z}\}$ is a sequence of i.i.d. regularly varying random variables with tail index $\alpha > 0$. For this process we have the explicit formulas for all cluster indices. Samples of size $n = 1000$ are generated from AR(1) with $\alpha = 4$ and $\rho = 0.5, 0.9$. We perform simulations for the classical extremal index as well as for the stop-loss index.

The following parameters are used:

- Blocks sizes $r_n = 6,\ 10,\ 20,\ 30$;
- Number of order statistics $k = 6\%,\ 10\%,\ 20\%,\ 30\%$ of the sample size n.

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

- The estimators for the extremal index perform the best in case of a) strong dependence; b) small number of order statistics; c) small block sizes. This is illustrated in Figure 1, where the following parameters are used: block size $r_n = 6$ and order statistics $k = 6\%,\ 10\%,\ 20\%,\ 30\%$ of the sample size $n = 1000$. We note that in case of strong dependence ($\rho = 0.9$), the sliding and disjoint blocks estimators perform well for small values of r_n and small values of k. For larger values of r_n and k, blocks estimators became heavily biased. Runs estimators are heavily biased irrespective of the choice of the block size and the order statistics. For weak dependence ($\rho = 0.5$), the results are bad for virtually all the estimators regardless of the considered parameters, primarily due to a strong bias. This is illustrated in Figures 1–8 in the Supplementary file.
- In this spirit, taking the "best" set of parameters suggested by the graphs, in Table 1 we included the results for Monte Carlo simulation for the extremal index based on disjoint blocks, sliding blocks and runs estimators with the block sizes $r_n = 5, 6, 7, 8$. We used the number of order statistics $k = 5\%$ and 10% of the sample size n. We note again that in case of strong

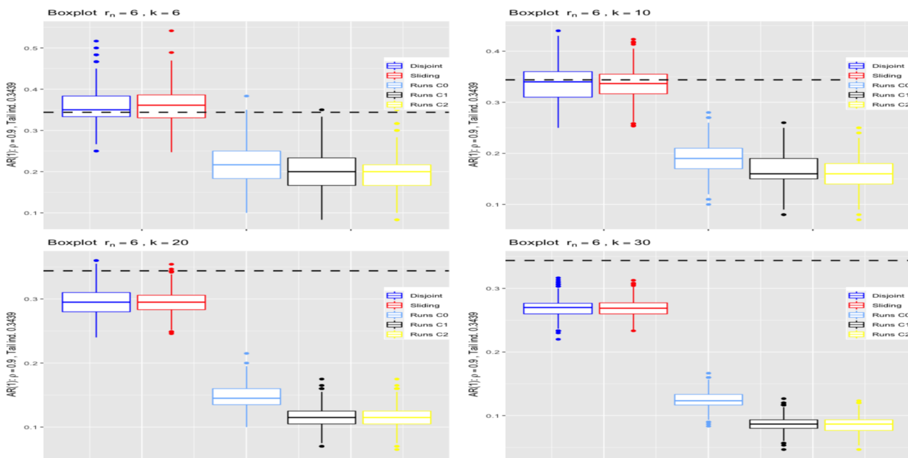

FIG 1. Monte Carlo simulations for a sample $n = 1000$ based on $N = 1000$ replicates of disjoint, sliding blocks and runs estimators ($C^{(0)}$, $C^{(1)}$ and $C^{(2)}$) for the extremal index. Data are simulated from AR(1) with $\alpha = 4$, $\rho = 0.9$ (thus, $\theta = 0.34$) The block size $r_n = 6$ and the number of order statistics $k = 6\%$ (top left), $k = 10\%$ (top right), $k = 20\%$ (bottom left), $k = 30\%$ (bottom right) of the sample size n. Dotted lines indicated the true value of the cluster index.

dependence ($\rho = 0.9$), the sliding and disjoint blocks estimators outperform runs estimators for all considered parameters. For weak dependence ($\rho = 0.5$), the results are heavily biased for all considered parameters.

- In either case (weak and strong dependence) the variability of all estimators is approximately the same, supporting the theoretical findings of this paper.

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

$$\theta_{\text{stop-loss}}(S) = (1-\rho^{\alpha})\mathbb{P}\left(\sum_{j=0}^{\infty}(\rho^j Y_0 - 1)_+ > S\right) , \tag{4.1}$$

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

- At the first step we used the formula (4.1) and performed the Monte-Carlo simulation to obtain the approximate value of the stop-loss index.
- With this in mind, we performed simulation studies for values of r_n and k defined above. As noted in [CK21], the stop-loss index estimation requires a higher number of order statistics.
- For the weak dependence ($\rho = 0.5$), we notice (see Figure 2) that, as opposed to the extremal index, disjoint and sliding blocks as well as the runs estimator $C^{(0)}$ perform well regardless of the choice of the size of the block (with a very good performance for a wide range of the block sizes r_n). In some cases, the estimator $C^{(0)}$ outperforms disjoint and sliding blocks. The estimators $C^{(1)}$, $C^{(2)}$ are heavily biased most of the time.

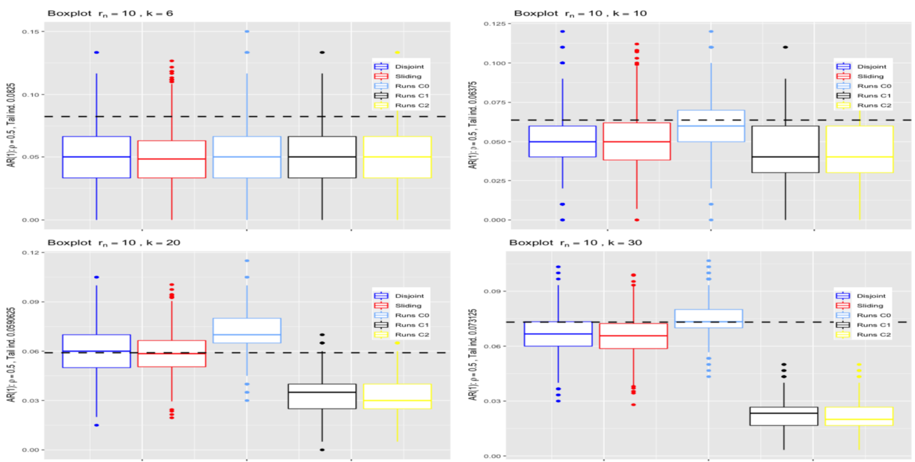

FIG 2. Monte Carlo simulations for a sample $n = 1000$ based on $N = 1000$ replicates of disjoint, sliding blocks and runs estimators ($C^{(0)}$, $C^{(1)}$ and $C^{(2)}$) for stop-loss index with $S = 0.9$. Data are simulated from AR(1) with $\rho = 0.5$, $r_n = 10$ and the number of order statistics $k = 6\%$ (top left), $k = 10\%$ (top right), $k = 20\%$ (bottom left), $k = 30\%$ (bottom right) of the sample size n. Dotted lines indicated the true value of the cluster index.

- For the strong dependence ($\rho = 0.9$), the simulation results are rather poor for all the estimators irrespective of the chosen parameters due to the bias. This may be quite intuitive, since the stop-loss functional is based on *sums* of large values. This is illustrated in Figures 9-16 in the Supplementary file.
- In this spirit, in Table 2 we included the results for Monte Carlo simulation for the stop-loss index based on disjoint blocks, sliding blocks and runs estimators $C^{(0)}$ with the block sizes $r_n = 6, 10, 20, 30$ (to show stability across the wide range of block sizes) and $k = 10\%, 40\%$ (to show good performance for a large number of order statistics).
- The variability of the runs estimators is not quite in line with the theoretical results of the paper.

4.2. ***Stationary ARCH process***

We consider a stationary ARCH(1) process defined by $X_j = \sqrt{\beta + \lambda X_{j-1}^2} Z_j$, 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]).

- Monte Carlo results are included in Table 3. In this case both disjoint and sliding blocks estimators yield better results as compared to runs. This is primarily due to bias.
- The variability of all estimators is approximately the same, in line with the theoretical results of the paper.

Table 1

The median and the variance (in brackets) of disjoint, sliding blocks and runs (with the anchoring maps $C^{(0)}, C^{(1)}$ and $C^{(2)}$) 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 sizes are $r_n = 5,\ 6,\ 7,\ 8$. The number of order statistics are $k = 5\%,\ 10\%$ of the sample size n with $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

	$\rho = 0.9$, Extremal Index=0.34				$\rho = 0.5$, Extremal Index= 0.94			
(k %)	$k = 5$		$k = 10$		$k = 5$		$k = 10$	
$r_n = 5$								
Disjoint bl	**0.380**	(0.0499)	**0.350**	(0.0334)	**0.680**	(0.0513)	**0.620**	(0.0355)
Sliding bl	**0.376**	(0.0441)	**0.348**	(0.0293)	**0.692**	(0.0429)	**0.622**	(0.0285)
Runs $C^{(0)}$	**0.240**	(0.0488)	**0.180**	(0.0314)	**0.560**	(0.0442)	**0.450**	(0.0342)
Runs $C^{(1)}$	**0.220**	(0.0498)	**0.150**	(0.0318)	**0.540**	(0.0563)	**0.420**	(0.0357)
Runs $C^{(2)}$	**0.220**	(0.0499)	**0.150**	(0.0320)	**0.540**	(0.0564)	**0.420**	(0.0357)
$r_n = 6$								
Disjoint bl	**0.340**	(0.0498)	**0.310**	(0.0313)	**0.660**	(0.0540)	**0.580**	(0.0351)
Sliding bl	**0.340**	(0.0441)	**0.317**	(0.0275)	**0.670**	(0.0447)	**0.539**	(0.0283)
Runs $C^{(0)}$	**0.200**	(0.0468)	**0.170**	(0.0273)	**0.540**	(0.0555)	**0.420**	(0.0340)
Runs $C^{(1)}$	**0.180**	(0.0480)	**0.140**	(0.0288)	**0.520**	(0.0581)	**0.390**	(0.0361)
Runs $C^{(2)}$	**0.180**	(0.0479)	**0.140**	(0.0288)	**0.520**	(0.0577)	**0.390**	(0.0362)
$r_n = 7$								
Disjoint bl	**0.320**	(0.0515)	**0.290**	(0.0314)	**0.640**	(0.0524)	**0.550**	(0.0343)
Sliding bl	**0.317**	(0.0458)	**0.291**	(0.0281)	**0.649**	(0.0442)	**0.558**	(0.0281)
Runs $C^{(0)}$	**0.180**	(0.0465)	**0.160**	(0.0270)	**0.520**	(0.0537)	**0.400**	(0.0300)
Runs $C^{(1)}$	**0.180**	(0.0483)	**0.140**	(0.0275)	**0.500**	(0.0557)	**0.370**	(0.0316)
Runs $C^{(2)}$	**0.180**	(0.0482)	**0.140**	(0.0273)	**0.500**	(0.0560)	**0.370**	(0.0317)
$r_n = 8$								
Disjoint bl	**0.300**	(0.0479)	**0.270**	(0.0318)	**0.640**	(0.0537)	**0.540**	(0.0340)
Sliding bl	**0.303**	(0.0425)	**0.273**	(0.0284)	**0.633**	(0.0451)	**0.535**	(0.0281)
Runs $C^{(0)}$	**0.180**	(0.0437)	**0.150**	(0.0273)	**0.500**	(0.0517)	**0.370**	(0.0298)
Runs $C^{(1)}$	**0.180**	(0.0441)	**0.130**	(0.0281)	**0.480**	(0.0539)	**0.340**	(0.0323)
Runs $C^{(2)}$	**0.180**	(0.0440)	**0.130**	(0.0281)	**0.480**	(0.0539)	**0.340**	(0.0322)

- The estimation works relatively well for a) small values of k; b) small block sizes ($r_n = 6, 10$).

4.3. Summary

Our findings are summarized here:

- All estimators (blocks and runs) have virtually same variance, which is in line with the theoretical results.
- For the extremal index, runs estimators are inferior as compared to blocks estimators. This is primarily due to bias.
- For the extremal index, blocks estimators perform well in case of a) strong dependence; b) small number of order statistics; c) small block sizes.
- For the stop-loss index, both blocks and runs estimator $C^{(0)}$ are superior, in comparison to $C^{(1)}$, $C^{(2)}$; the latter being heavily biased in most cases.
- For the stop-loss index, blocks estimators anc $C^{(0)}$ perform well in case of a) weak dependence; b) small number of order statistics; c) wide range of block sizes.

TABLE 2
The median and the variance (in brackets) of disjoint, sliding blocks and runs ($C^{(0)}$) estimators for stop-loss index with $S = 0.9$. Data are simulated from AR(1) with $\alpha = 4$, $\rho = 0.5,\ 0.9$. The block sizes are $r_n = 6,\ 10,\ 20,\ 30$. The number of order statistics are $k = 10\%,\ 40\%$ of the sample size n with $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

	$\rho = 0.9$, Stop-loss Index=0.085				$\rho = 0.5$, Stop-loss Index= 0.078			
(k %)	$k = 10$		$k = 40$		$k = 10$		$k = 40$	
$r_n = 6$								
Disjoint bl	**0.000**	(0.0118)	**0.008**	(0.0073)	**0.030**	(0.0192)	**0.055**	(0.0105)
Sliding bl	**0.000**	(0.0110)	**0.009**	(0.0069)	**0.032**	(0.0177)	**0.055**	(0.0093)
Runs $C^{(0)}$	**0.010**	(0.0106)	**0.020**	(0.0061)	**0.040**	(0.0192)	**0.070**	(0.0084)
Runs $C^{(1)}$	**0.000**	(0.0084)	**0.003**	(0.0032)	**0.030**	(0.0184)	**0.020**	(0.0070)
Runs $C^{(2)}$	**0.000**	(0.0014)	**0.000**	(0.0000)	**0.030**	(0.0177)	**0.015**	(0.0068)
$r_n = 10$								
Disjoint bl	**0.000**	(0.0122)	**0.018**	(0.0073)	**0.040**	(0.0196)	**0.068**	(0.0089)
Sliding bl	**0.006**	(0.0108)	**0.017**	(0.0067)	**0.049**	(0.0179)	**0.067**	(0.0074)
Runs $C^{(0)}$	**0.010**	(0.0110)	**0.022**	(0.0049)	**0.050**	(0.0191)	**0.070**	(0.0071)
Runs $C^{(1)}$	**0.010**	(0.0097)	**0.005**	(0.0038)	**0.030**	(0.0183)	**0.010**	(0.0054)
Runs $C^{(2)}$	**0.000**	(0.0082)	**0.000**	(0.0024)	**0.030**	(0.0180)	**0.010**	(0.0049)
$r_n = 20$								
Disjoint bl	**0.010**	(0.0114)	**0.023**	(0.0054)	**0.050**	(0.0188)	**0.073**	(0.0067)
Sliding bl	**0.010**	(0.0106)	**0.024**	(0.0046)	**0.049**	(0.0168)	**0.073**	(0.0051)
Runs $C^{(0)}$	**0.010**	(0.0102)	**0.023**	(0.0039)	**0.060**	(0.0169)	**0.053**	(0.0048)
Runs $C^{(1)}$	**0.010**	(0.0098)	**0.005**	(0.0037)	**0.030**	(0.0163)	**0.000**	(0.0020)
Runs $C^{(2)}$	**0.010**	(0.0098)	**0.003**	(0.0033)	**0.020**	(0.0165)	**0.000**	(0.0020)
$r_n = 30$								
Disjoint bl	**0.010**	(0.0106)	**0.023**	(0.0046)	**0.060**	(0.0188)	**0.065**	(0.0043)
Sliding bl	**0.011**	(0.0092)	**0.024**	(0.0038)	**0.056**	(0.0160)	**0.066**	(0.0032)
Runs $C^{(0)}$	**0.010**	(0.0096)	**0.020**	(0.0037)	**0.075**	(0.0161)	**0.035**	(0.0040)
Runs $C^{(1)}$	**0.010**	(0.0094)	**0.003**	(0.0035)	**0.020**	(0.0142)	**0.000**	(0.0006)
Runs $C^{(2)}$	**0.010**	(0.0094)	**0.003**	(0.0029)	**0.020**	(0.0142)	**0.000**	(0.0006)

- Furthermore, it is rather not feasible that linear combinations of the blocks and runs will reduces the bias.

5. Proofs

In Section 5.3 we prove several lemmas on conditional convergence when anchoring maps are involved. One needs to distinguish between finite blocks (when the conditional convergence follows basically from the conditional convergence to the tail process) and growing blocks (when the anticlustering condition is needed).

In Section 5.4 we prove the asymptotic behaviour of the covariances of runs estimators, that is we prove Lemma 3.3 and Lemma 3.4. Section 5.5 deals with the empirical cluster process of runs statistics. The functional central limit theorem (Theorem 5.6) established there yields immediately the central limit theorem for runs estimators. See Section 5.6. A long proof of Theorem 5.6 is given in Sections 5.7 and 5.8. Finally, in Section 5.9 we discuss the random entropy assumption.

TABLE 3
The median and the variance (in brackets) of disjoint, sliding blocks and runs ($C^{(0)}$, $C^{(1)}$ and $C^{(2)}$) estimators for the extremal index in ARCH(1) model with $\lambda = 0.9$. The block size is $r_n = 6,\ 10,\ 20,\ 30$. The number of order statistics is $k = 6\%$, 10%, 20% and 30% of the sample size n with $n = 1000$ based on $N = 1000$ Monte Carlo simulations.

	Extremal Index=0.612							
(k %)	$k=6$		$k=10$		$k=20$		$k=30$	
$r_n = 6$								
Disjoint bl	**0.650**	(0.0544)	**0.620**	(0.0340)	**0.550**	(0.0216)	**0.470**	(0.0127)
Sliding bl	**0.647**	(0.0479)	**0.625**	(0.0338)	**0.551**	(0.0175)	**0.472**	(0.0096)
Runs $C^{(0)}$	**0.483**	(0.0573)	**0.440**	(0.0392)	**0.320**	(0.0185)	**0.233**	(0.0118)
Runs $C^{(1)}$	**0.467**	(0.0594)	**0.410**	(0.0399)	**0.245**	(0.0183)	**0.1167**	(0.0122)
Runs $C^{(2)}$	**0.467**	(0.0594)	**0.410**	(0.0401)	**0.245**	(0.0183)	**0.1167**	(0.0121)
$r_n = 10$								
Disjoint bl	**0.567**	(0.0563)	**0.530**	(0.0370)	**0.410**	(0.0157)	**0.317**	(0.0060)
Sliding bl	**0.568**	(0.0497)	**0.526**	(0.0314)	**0.410**	(0.0120)	**0.314**	(0.0041)
Runs $C^{(0)}$	**0.417**	(0.0493)	**0.340**	(0.0305)	**0.210**	(0.0134)	**0.143**	(0.0085)
Runs $C^{(1)}$	**0.400**	(0.0493)	**0.300**	(0.0308)	**0.115**	(0.0145)	**0.0013**	(0.0082)
Runs $C^{(2)}$	**0.400**	(0.0499)	**0.300**	(0.0309)	**0.115**	(0.0146)	**0.0030**	(0.0081)
$r_n = 20$								
Disjoint bl	**0.450**	(0.0448)	**0.380**	(0.0244)	**0.240**	(0.0054)	**0.167**	(0.0009)
Sliding bl	**0.450**	(0.0340)	**0.374**	(0.0196)	**0.266**	(0.0036)	**0.163**	(0.0005)
Runs $C^{(0)}$	**0.300**	(0.0349)	**0.210**	(0.0201)	**0.110**	(0.0093)	**0.073**	(0.0062)
Runs $C^{(1)}$	**0.250**	(0.0351)	**0.130**	(0.0210)	**0.0015**	(0.0083)	**0.000**	(0.0021)
Runs $C^{(2)}$	**0.250**	(0.0356)	**0.130**	(0.0209)	**0.015**	(0.0084)	**0.000**	(0.0021)
$r_n = 30$								
Disjoint bl	**0.383**	(0.0352)	**0.290**	(0.0159)	**0.165**	(0.0018)	**0.110**	(0.0001)
Sliding bl	**0.373**	(0.0291)	**0.284**	(0.0118)	**0.160**	(0.0010)	**0.010**	(0.0000)
Runs $C^{(0)}$	**0.217**	(0.0280)	**0.140**	(0.0154)	**0.07**	(0.008)	**0.047**	(0.0005)
Runs $C^{(1)}$	**0.150**	(0.0300)	**0.060**	(0.0172)	**0.000**	(0.0036)	**0.000**	(0.0004)
Runs $C^{(2)}$	**0.150**	(0.0300)	**0.060**	(0.0170)	**0.000**	(0.0036)	**0.000**	(0.0004)

5.1. Mixing

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 (cf. [Ibr62])

$$|\mathrm{cov}(H(Z_1), H(Z_2))| \le \mathrm{cst}\ \|H\|_\infty \|\widetilde{H}\|_\infty \beta(\sigma(Z_1), \sigma(Z_2)) \ . \tag{5.1}$$

In (5.1) the constant cst does not depend on $H, \widetilde{H}$.

5.2. The anticlustering conditions

Lemma 5.1. *Assume that $\mathcal{R}(r_n, u_n)$ and $\mathcal{AC}(r_n, u_n)$ hold. If $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ is beta-mixing and $\lim_{n\to\infty} \beta_{r_n}/\mathbb{P}(|\boldsymbol{X}_0| > u_n) = 0$, then holds $\mathcal{AC}(r_n + h_n, u_n)$ with any $h_n \le r_n/2$.*

Proof. We have

$$\mathbb{P}\left(\max_{m\le|j|\le r_n+h_n} |\boldsymbol{X}_j| > u_n x \mid |\boldsymbol{X}_0| > u_n y\right)$$

$$\leq 2\mathbb{P}\left(\max_{m\leq|j|\leq r_n}|\boldsymbol{X}_j|>u_nx \mid |\boldsymbol{X}_0|>u_ny\right)$$
$$+ 2\mathbb{P}\left(\max_{r_n+1\leq|j|\leq r_n+h_n}|\boldsymbol{X}_j|>u_nx \mid |\boldsymbol{X}_0|>u_ny\right) .$$

The second last term vanishes as $n\to\infty$ and then $m\to\infty$, on account of $\mathcal{AC}(r_n,u_n)$. Applying (5.1) to the last term we get

$$\begin{aligned}
&\mathbb{P}\left(\max_{r_n+1\leq|j|\leq r_n+h_n}|\boldsymbol{X}_j|>u_nx \mid |\boldsymbol{X}_0|>u_ny\right)\\
&\leq \mathbb{P}\left(\max_{r_n+1\leq|j|\leq r_n+h_n}|\boldsymbol{X}_j|>u_nx\right)+\frac{\beta_{r_n}}{\mathbb{P}(|\boldsymbol{X}_0|>u_ny)}\\
&= \mathbb{P}\left(\max_{1\leq j\leq 2h_n}|\boldsymbol{X}_j|>u_nx\right)+o(1) ,
\end{aligned}$$

where the latter $o(1)$ holds when $n\to\infty$ by the assumptions. The first term in the last line vanishes as $n\to\infty$. Indeed, since $h_n\leq r_n/2$, the assumed $\mathcal{AC}(r_n,u_n)$ implies $\mathcal{AC}(2h_n,u_n)$. Then (2.9) gives

$$\lim_{n\to\infty}\frac{\mathbb{P}\left(\max_{1\leq j\leq 2h_n}|\boldsymbol{X}_j|>u_nx\right)}{2h_n\mathbb{P}(|\boldsymbol{X}_0|>u_n)}=x^{-\alpha}\vartheta\in(0,\infty) .$$

Since $\mathcal{R}(r_n,u_n)$ holds, $\lim_{n\to\infty} r_n\mathbb{P}(|\boldsymbol{X}_0|>u_n)=0$. This implies that

$$\lim_{n\to\infty}h_n\mathbb{P}(|\boldsymbol{X}_0|>u_n)=0$$

and hence $\lim_{n\to\infty}\mathbb{P}\left(\max_{1\leq j\leq 2h_n}|\boldsymbol{X}_j|>u_nx\right)=0$. This finished the proof. □

5.3. Conditional convergence

Lemma 5.2. *Let $\mathcal{C}=\mathcal{C}^{(0)},\mathcal{C}^{(1)},\mathcal{C}^{(2)}$. Then for $h\in\mathbb{Z}$,*

$$\lim_{n\to\infty}\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r,h+r}/u_n)=h \mid |\boldsymbol{X}_0|>u_n)=\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{h-r,h+r})=h) .$$

Proof. We will do the proof for $\mathcal{C}^{(0)}$ only. By the definition of the tail process

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}(\mathcal{C}^{(0)}(\boldsymbol{X}_{h-r,h+r}/u_n)=h \mid |\boldsymbol{X}_0|>u_n)\\
&=\lim_{n\to\infty}\mathbb{P}(\boldsymbol{X}^*_{h-r,h-1}/u_n<|\boldsymbol{X}_h|/u_n, \boldsymbol{X}^*_{h+1,h+r}/u_n\leq|\boldsymbol{X}_h|/u_n \mid |\boldsymbol{X}_0|>u_n)\\
&=\mathbb{P}(\boldsymbol{Y}^*_{h-r,h-1}<|\boldsymbol{Y}_h|, \boldsymbol{Y}^*_{h+1,h+r}\leq|\boldsymbol{Y}_h|)=\mathbb{P}(\mathcal{C}^{(0)}(\boldsymbol{Y}_{h-r,h+r})=h) .
\end{aligned}$$
□

Lemma 5.3. *Let $\mathcal{C},\widetilde{\mathcal{C}}$ be any of the anchoring maps $\mathcal{C}^{(0)},\mathcal{C}^{(1)},\mathcal{C}^{(2)}$. Then for $h\in\mathbb{Z}$,*

$$\begin{aligned}
&\lim_{n\to\infty}\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{-r,r}/u_n)=0,\widetilde{\mathcal{C}}(\boldsymbol{X}_{h-r,h+r}/u_n)=h \mid |\boldsymbol{X}_0|>u_n)\\
&=\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{-r,r})=0,\widetilde{\mathcal{C}}(\boldsymbol{Y}_{h-r,h+r})=h) .
\end{aligned} \tag{5.2}$$

Proof. We verify the statement for one combination of the anchoring maps only. For $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$, we have

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}(\mathcal{C}^{(1)}(\boldsymbol{X}_{h-r,h+r}/u_n) = h, \mathcal{C}^{(2)}(\boldsymbol{X}_{-r,r}/u_n) = 0 \mid |\boldsymbol{X}_0| > u_n) \\
&= \mathbb{P}(\boldsymbol{X}^*_{h-r,h-1} \le u_n, |\boldsymbol{X}_h| > u_n, \boldsymbol{X}^*_{1,r} \le u_n, |\boldsymbol{X}_0| > u_n \mid |\boldsymbol{X}_0| > u_n) \\
&= \mathbb{P}(\boldsymbol{Y}^*_{h-r,h-1} \le 1, |\boldsymbol{Y}_h| > 1, \boldsymbol{Y}^*_{1,r} \le 1, |\boldsymbol{Y}_0| > 1) \; . \\
&= \mathbb{P}(\mathcal{C}^{(1)}(\boldsymbol{Y}_{h-r,h+r}) = h, \mathcal{C}^{(2)}(\boldsymbol{Y}_{-r,r}) = 0) \; . \qquad \square
\end{aligned}$$

Recall the definition of $H^{\mathcal{C}}$ in (3.1). Let $h_n \le r_n/2$ be a sequence of integers diverging to infinity. For bounded $H, \widetilde{H}$, a direct application of $\mathcal{AC}(r_n, u_n)$ gives

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right) \widetilde{H}^{\widetilde{\mathcal{C}}}\left(\boldsymbol{X}_{h_n-r_n,h_n+r_n}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right] \\
&\le \|H\| \|\widetilde{H}\| \lim_{n\to\infty} \mathbb{P}(|\boldsymbol{X}_{h_n}| > u_n \mid |\boldsymbol{X}_0| > u_n) = 0 \; .
\end{aligned}$$

Likewise, if $H, \widetilde{H}$ are bounded and $\widetilde{H} \in \mathcal{L}$ is such that $\widetilde{H}(\mathbf{0}) = 0$, then the Lipschitz continuity of $\widetilde{H}$, $\mathcal{AC}(r_n, u_n)$ and Lemma 5.1 imply

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right) \widetilde{H}\left(\boldsymbol{X}_{h_n,h_n+r_n}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right] \\
&\le \|H\| \lim_{n\to\infty} \mathbb{E}\left[\widetilde{H}\left(\boldsymbol{X}_{h_n,h_n+r_n}/u_n\right) - \widetilde{H}(\mathbf{0}) \mid |\boldsymbol{X}_0| > u_n\right] \\
&\le \|H\| \|\widetilde{H}\| \lim_{n\to\infty} \mathbb{P}\left(\boldsymbol{X}^*_{h_n,h_n+r_n} > u_n \mid |\boldsymbol{X}_0| > u_n\right) = 0 \; . \qquad (5.3)
\end{aligned}$$

The statement in (5.3) is also valid if $\widetilde{H}\left(\boldsymbol{X}_{h_n,h_n+r_n}/u_n\right)$ is replaced with $\widetilde{H}\left(\boldsymbol{X}_{-h_n,-h_n-r_n}/u_n\right)$.

On the other hand, for fixed h we have the following lemma that extends Lemma 5.3 from fixed r to $r_n \to \infty$. For this, we need to assume additionally that $\mathcal{AC}(r_n, u_n)$ holds.

Lemma 5.4. *Assume the conditions of Lemma 5.1 are satisfied. Let $H, \widetilde{H} \in \mathcal{L}$ and $\mathcal{C}, \widetilde{\mathcal{C}}$ be any of the anchoring maps $\mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$. Then*

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right) \widetilde{H}^{\widetilde{\mathcal{C}}}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right] \\
&= \mathbb{E}\left[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\mathbb{1}\left\{\widetilde{\mathcal{C}}(\boldsymbol{Y}) = h\right\}\right] =: \mathcal{I}(H, \widetilde{H}, \mathcal{C}, \widetilde{\mathcal{C}}; h) \; . \qquad (5.4)
\end{aligned}$$

Before we prove the above lemma, we make several comments. First, as a corollary we obtain

$$\lim_{n\to\infty} \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right] = \boldsymbol{\nu}^*(H) \qquad (5.5)$$

and

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{C}(\boldsymbol{X}_{-r_n,r_n}/u_n) = 0 \mid |\boldsymbol{X}_0| > u_n) = \boldsymbol{\nu}^*(1) = 1 \; . \qquad (5.6)$$

Indeed, if we take $\widetilde{H} \equiv 1$, $\mathcal{C} = \widetilde{\mathcal{C}}$ and $h = 0$, then by (2.8),

$$\begin{aligned}\mathcal{I}(H, 1, \mathcal{C}, \mathcal{C}; 0) &= \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\right] \\ &= \mathbb{E}\left[H(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\mathbb{1}\{|\boldsymbol{Y}_0| > 1\}\right] = \boldsymbol{\nu}^*(H) \,.\end{aligned}$$

Since the definition of $\boldsymbol{\nu}^*$ does not depend on the anchoring map, we have $\mathcal{I}(H, \widetilde{H}, \mathcal{C}, \widetilde{\mathcal{C}}; 0) = \boldsymbol{\nu}^*(H\widetilde{H})$ for any $\mathcal{C}, \widetilde{\mathcal{C}}$. Since the value of any anchoring map is uniquely determined, we conclude immediately that $\mathcal{I}(H, \widetilde{H}, \mathcal{C}, \mathcal{C}; h) = 0$ for $h \neq 0$. Furthermore,

$$\begin{aligned}&\sum_{h\in\mathbb{Z}} \mathbb{E}\left[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\mathbb{1}\left\{\widetilde{\mathcal{C}}(\boldsymbol{Y}) = h\right\}\right] \\ &= \mathbb{E}\left[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\mathbb{1}\left\{\widetilde{\mathcal{C}}(\boldsymbol{Y}) \in \mathbb{Z}\right\}\right] = \boldsymbol{\nu}^*(H\widetilde{H}) \,.\end{aligned}$$

This implies that for arbitrary anchoring maps $\mathcal{C}$, $\widetilde{\mathcal{C}}$,

$$\mathcal{I}(H, \widetilde{H}, \mathcal{C}, \widetilde{\mathcal{C}}; h) = 0 \,, \quad h \neq 0 \,. \tag{5.7}$$

Proof of Lemma 5.4. In [CK21] we proved a version of the lemma without anchoring maps included. Since $\boldsymbol{x} \to \mathbb{1}\{\mathcal{C}(\boldsymbol{x}) = 0\}$ is not Lipschitz continuous, Lemma 6.6 in [CK21] is not directly applicable. As such, we will focus on the anchoring maps only, assuming $H = \widetilde{H} \equiv 1$.

In the first step we prove that for all $h \in \mathbb{Z}$

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n) = h \mid |\boldsymbol{X}_0| > u_n) = \mathbb{P}(\mathcal{C}(\boldsymbol{Y}) = h) \,. \tag{5.8}$$

We already know that (cf. Lemma 5.2)

$$\lim_{n\to\infty} \mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r,h+r}/u_n) = h \mid |\boldsymbol{X}_0| > u_n) = \mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{h-r,h+r}) = h) \,.$$

Since $r_n \to \infty$ and r is fixed we can assume $0 < r < r_n$. Now, for $\mathcal{C} = \mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$ the value of

$$\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}_{h-r,h+r}) = h\} - \mathbb{1}\{\mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n}) = h\}$$

is non zero if and only if $\mathbb{P}(h + r < \mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n}) \leq h + r_n) > 0$ or $\mathbb{P}(h - r_n \leq \mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n}) < h - r) > 0$. Indeed, take for simplicity $h = 0$. If $\mathcal{C}^{(1)}(\boldsymbol{Y}_{-r,r}) = 0$ and $\mathcal{C}^{(1)}(\boldsymbol{Y}_{-r_n,r_n}) \neq 0$, then $\boldsymbol{Y}^*_{-r,-1} \leq 1$, $|\boldsymbol{Y}_0| > 1$ and then $\boldsymbol{Y}^*_{-r_n,-r-1} > 1$, while $\mathcal{C}^{(1)}(\boldsymbol{Y}_{-r,r}) \neq 0$ and $\mathcal{C}^{(1)}(\boldsymbol{Y}_{-r_n,r_n}) = 0$ cannot happen. The same reasoning applied to the other anchoring maps.

Coming back to the general case of h, the first property of the anchoring map implies that $|\boldsymbol{Y}_j| > 1$ for some $j \in \{h+r+1, \ldots, h+r_n\} \cup \{h-r_n, \ldots, h-r-1\}$. Since we let $r, r_n \to \infty$, we can assume that $h < r$. Thus, using the property An(i) of the anchoring maps,

$$\lim_{r\to\infty} \lim_{n\to\infty} |\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{h-r,h+r}) = h) - \mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n}) = h)|$$

$$\leq \lim_{r\to\infty}\lim_{n\to\infty}\mathbb{P}\left(\max\left\{\max_{h+r\leq j\leq h+r_n}|\boldsymbol{Y}_j|, \max_{h-r_n\leq j\leq h-r}|\boldsymbol{Y}_j|\right\}>1\right)=0 \quad (5.9)$$

since $\mathcal{AC}(r_n,u_n)$ implies $\boldsymbol{Y}_j\to\boldsymbol{0}$ almost surely as $|j|\to\infty$. Also, the vanishing property of $\boldsymbol{Y}_j$ and the property An(i) of the anchoring map imply that

$$\lim_{n\to\infty}\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n})=h)=\mathbb{P}(\mathcal{C}(\boldsymbol{Y})=h)\ .$$

Similarly,

$$|\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)=h\mid|\boldsymbol{X}_0|>u_n)-\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r,h+r}/u_n)=h\mid|\boldsymbol{X}_0|>u_n)|$$

is non zero if and only if $\mathbb{P}(h+r<\mathcal{C}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)\leq h+r_n)>0$ or

$\mathbb{P}(h-r_n\leq\mathcal{C}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)<h-r)>0$. The first property of the anchoring map implies that $|\boldsymbol{X}_j|>|\boldsymbol{X}_0|\wedge u_n$ for some $j\in\{h+r+1,\dots,h+r_n\}\cup\{h-r_n,\dots,h-r-1\}$. Again, we can assume that $h<r$. Keeping in mind the conditioning we have:

$$\begin{aligned}&\left|\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)=h\mid|\boldsymbol{X}_0|>u_n)-\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{h-r,h+r}/u_n)=h\mid|\boldsymbol{X}_0|>u_n)\right|\\&\leq\mathbb{P}\left(\max\left\{\max_{h+r\leq j\leq h+r_n}|\boldsymbol{X}_j|, \max_{h-r_n\leq j\leq h-r}|\boldsymbol{X}_j|\right\}>u_n\mid|\boldsymbol{X}_0|>u_n\right)\ . \quad (5.10)\end{aligned}$$

By Lemma 5.1, the latter expression vanishes by letting first $n\to\infty$ and then $r\to\infty$. This finishes the proof of (5.8).

Now, we will prove

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{P}(\mathcal{C}(\boldsymbol{X}_{-r_n,+r_n}/u_n)=0,\widetilde{\mathcal{C}}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)=h\mid|\boldsymbol{X}_0|>u_n)\\&=\mathbb{P}(\mathcal{C}(\boldsymbol{Y})=0,\widetilde{\mathcal{C}}(\boldsymbol{Y})=h)\ . \quad (5.11)\end{aligned}$$

In view of Lemma 5.3, (5.11) holds with r_n replaced with r. Now, the idea is to reduce the bivariate case to the univariate.

Note first that for the anchoring maps considered here, the event $A_1:=\{\mathcal{C}(\boldsymbol{Y}_{h-r_n,h+r_n})=h\}$ is included in $A_2:=\{\mathcal{C}(\boldsymbol{Y}_{h-r,h+r})=h\}$. We also note that for any event B and any pair of ordered events A_1,A_2 we have

$$|\mathbb{P}(A_1\cap B)-\mathbb{P}(A_2\cap B)|\leq|\mathbb{P}(A_1)-\mathbb{P}(A_2)|\ .$$

Thus, we can bound

$$\left|\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{-r,r})=0,\widetilde{\mathcal{C}}(\boldsymbol{Y}_{h-r,h+r})=h)-\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{-r_n,r_n})=0,\widetilde{\mathcal{C}}(\boldsymbol{Y}_{h-r_n,h+r_n})=h)\right| \quad (5.12)$$

by

$$\begin{aligned}&|\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{-r,r})=0)-\mathbb{P}(\mathcal{C}(\boldsymbol{Y}_{-r_n,r_n})=0)|\\&+\left|\mathbb{P}(\widetilde{\mathcal{C}}(\boldsymbol{Y}_{h-r,h+r})=h)-\mathbb{P}(\widetilde{\mathcal{C}}(\boldsymbol{Y}_{h-r_n,h+r_n})=h)\right|\end{aligned}$$

and we use the first step to conclude that the expression in (5.12) vanishes by letting first $n \to \infty$ and then $r \to \infty$.

Therefore, (5.9) can be extended to the bivariate case. The same argument allows to extend (5.10) to the bivariate case. In summary, the proof of (5.11) is finished. □

In the next lemma, we analyse the conditional convergence for the product of $H^{\mathcal{C}}$ and $\widetilde{H}$. Its proof is almost the same as above and hence it is omitted.

Lemma 5.5. *Assume that $\mathcal{AC}(r_n, u_n)$ holds. Let $H, \widetilde{H} \in \mathcal{L}$, $\widetilde{H}(\mathbf{0}) = 0$ and $\mathcal{C}$ be any of the anchoring maps $\mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$. Then, for $h, h' \geq 0$,*

$$\lim_{n\to\infty} \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{h-r_n,h'+r_n}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right]$$
$$= \mathbb{E}\left[H\left(\boldsymbol{Y}\right)\widetilde{H}\left(\boldsymbol{Y}\right)\mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\}\right] = \boldsymbol{\nu}^*(H\widetilde{H}) \ . \tag{5.13}$$

5.4. Limiting covariances

The goal of this section is to prove Lemmas 3.3 and 3.4. Two situations will arise when dealing with the covariances:

- Situation 1: we will deal with $\sum_{h=-r_n}^{r_n} \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)]$, where

$$\lim_{n\to\infty} \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)] = \mathbb{E}[c_h(\boldsymbol{Y})] \ , \quad \sum_{h\in\mathbb{Z}} \mathbb{E}[|c_h(\boldsymbol{Y})|] < \infty \ .$$

  We will fix an integer $r > 0$; the convergence of $\sum_{h=-r}^{r} \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)]$ to $\sum_{h=-r}^{r} \mathbb{E}[c_h(\boldsymbol{Y})]$ will follow. The reminder $\sum_{|h|>r} \mathbb{E}[c_h(\boldsymbol{Y})]$ is negligible (as $r \to \infty$) by the summability assumption, while $\sum_{h>|r_n|} \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)]$ will be treated by the anticlustering condition $\mathcal{S}(r_n, u_n)$.
- Situation 2: we will deal with $r_n^{-1} \sum_{h=1}^{r_n} \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)] = \int_0^1 g_n(\xi)\mathrm{d}\xi$, where $g_n(h) = \mathbb{E}[c_{h,n}(\boldsymbol{X}/u_n)]$ and $g_n(\xi) \to g(\xi)$ as $n \to \infty$. Bounded convergence argument will be applied.

In what follows, to shorten the notation, we set $v_n = \mathbb{P}(|\boldsymbol{X}_0| > u_n)$.

Proof of Lemma 3.3. Recall that

$$H^{\mathcal{C}}_{n,j} = \sum_{i=jr_n+1}^{(j+1)r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) \ .$$

The covariance of the scaled statistics is

$$nv_n \mathrm{cov}\left(\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}), \widetilde{\boldsymbol{\xi}}^*_{n,r_n}(\widetilde{H}^{\widetilde{\mathcal{C}}})\right) = \frac{1}{r_n v_n}\mathrm{cov}\left(H^{\mathcal{C}}_{n,0}, \widetilde{H}^{\widetilde{\mathcal{C}}}_{n,0}\right)$$
$$+ \frac{1}{r_n v_n} \sum_{j=1}^{m_n-1}\left(1 - \frac{j}{m_n}\right)\left\{\mathrm{cov}(H^{\mathcal{C}}_{n,0}, \widetilde{H}^{\widetilde{\mathcal{C}}}_{n,j}) + \mathrm{cov}(\widetilde{H}^{\widetilde{\mathcal{C}}}_{n,0}, H^{\mathcal{C}}_{n,j})\right\} \ . \tag{5.14}$$

Using $\mathcal{R}(r_n,u_n)$, we will show that $\operatorname{cov}\left(H^{\mathcal{C}}_{n,0},\widetilde{H}^{\widetilde{\mathcal{C}}}_{n,0}\right)$ is determined by

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{r_nv_n}\operatorname{cov}\left(H^{\mathcal{C}}_{n,0},\widetilde{H}^{\widetilde{\mathcal{C}}}_{n,0}\right)\\
&=\lim_{n\to\infty}\frac{1}{v_n}\sum_{h=-r_n}^{r_n}\left(1-\frac{|h|}{r_n}\right)\mathbb{E}[H^{\mathcal{C}}\left(\boldsymbol{X}_{0,2r_n}/u_n\right)\widetilde{H}^{\widetilde{\mathcal{C}}}\left(\boldsymbol{X}_{h,h+2r_n}/u_n\right)]\\
&=\boldsymbol{\nu}^*(H\widetilde{H})\ . \qquad (5.15)
\end{aligned}$$

We are in the Situation 1. For fixed r, using (5.4) and (5.7) we have

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{v_n}\sum_{h=-r}^{r}\mathbb{E}[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}^{\widetilde{\mathcal{C}}}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right)]\\
&=\sum_{h=-r}^{r}\mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y})=0\}\mathbb{1}\left\{\widetilde{\mathcal{C}}(\boldsymbol{Y})=h\right\}]=\sum_{h=-r}^{r}\mathcal{I}(H,\widetilde{H},\mathcal{C},\widetilde{\mathcal{C}},h)=\\
&=\mathcal{I}(H,\widetilde{H},\mathcal{C},\widetilde{\mathcal{C}},0)=\mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y})\mathbb{1}\{\mathcal{C}(\boldsymbol{Y})=0\}]=\boldsymbol{\nu}^*(H\widetilde{H})\ . \qquad (5.16)
\end{aligned}$$

The value above does not depend on r. Moreover,

$$\begin{aligned}
&\frac{1}{v_n}\sum_{r<|h|\le r_n}\mathbb{E}[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}^{\widetilde{\mathcal{C}}}\left(\boldsymbol{X}_{-r_n+h,r_n+h}/u_n\right)]\\
&\le\|H\|\|\widetilde{H}\|\frac{1}{v_n}\sum_{r<|h|\le r_n}\mathbb{P}(|\boldsymbol{X}_0|>u_n,|\boldsymbol{X}_h|>u_n)\ . \qquad (5.17)
\end{aligned}$$

Letting $n\to\infty$ and then $r\to\infty$, we finish the proof of (5.15) by applying $\mathcal{S}(r_n,u_n)$.

Now, we deal with the term in (5.14). For $j\ge1$,

$$\begin{aligned}
&\frac{1}{r_nv_n}\left|\operatorname{cov}(H^{\mathcal{C}}_{n,0},\widetilde{H}^{\widetilde{\mathcal{C}}}_{n,j})\right|\\
&=\frac{1}{r_nv_n}\left|\operatorname{cov}\left(\sum_{h=1}^{r_n}H^{\mathcal{C}}\left(B^{-h}\boldsymbol{X}_{-r_n,r_n}/u_n\right),\sum_{i=1}^{r_n}\widetilde{H}^{\widetilde{\mathcal{C}}}\left(B^{-i}\boldsymbol{X}_{(j-1)r_n,(j+1)r_n}/u_n\right)\right)\right|\\
&\le\sum_{h=(j-1)r_n+1}^{jr_n}\left(\frac{h}{r_n}-(j-1)\right)|g_n(h)|+\sum_{h=jr_n+1}^{(j+1)r_n}\left((j+1)-\frac{h}{r_n}\right)|g_n(h)|\\
&\le\sum_{h=(j-1)r_n+1}^{(j+1)r_n}|g_n(h)|=:I_j \qquad (5.18)
\end{aligned}$$

with

$$g_n(h)=\frac{1}{v_n}\operatorname{cov}(H^{\mathcal{C}}(\boldsymbol{X}_{-r_n,r_n}/u_n),\widetilde{H}^{\widetilde{\mathcal{C}}}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n))\ .$$

For $h > 2r_n$ we have by (5.1),

$$|g_n(h)| \le \frac{\|H\|_\infty \|\widetilde{H}\|_\infty}{v_n} \beta_{h-2r_n} \ . \tag{5.19}$$

Thus,

$$\frac{1}{r_n v_n} \sum_{j=4}^{m_n-1} \left| \mathrm{cov}(H^{\mathcal{C}}_{n,0}, \widetilde{H}^{\widetilde{\mathcal{C}}}_{n,j}) \right| \le \frac{\|H\|_\infty \|\widetilde{H}\|_\infty}{v_n} \sum_{j=4}^{m_n-1} \sum_{h=(j-1)r_n+1}^{(j+1)r_n} \beta_{h-2r_n}$$
$$\le 2 \frac{\|H\|_\infty \|\widetilde{H}\|_\infty}{v_n} \sum_{h=3r_n+1}^{\infty} \beta_{h-2r_n} = O(1) \frac{1}{v_n} \sum_{i=r_n+1}^{\infty} \beta_i = o(1)$$

by the assumption (3.6b).

The terms that correspond to $j = 1, 2, 3$ in (5.14) have to be dealt with separately. We are again in the Situation 1. We have

$$I_1 + I_2 + I_3 \le 2 \sum_{h=1}^{4r_n} |g_n(h)| = 2 \left\{ \sum_{h=1}^{r} + \sum_{i=r+1}^{4r_n} \right\} |g_n(h)| \ .$$

Both parts are negligible. Indeed, as in (5.16),

$$\lim_{n\to\infty} \sum_{h=1}^{r} |g_n(h)| \le \lim_{n\to\infty} \frac{1}{v_n} \sum_{h=1}^{r} \mathbb{E}[H^{\mathcal{C}}(\boldsymbol{X}_{-r_n, r_n}/u_n)\, \widetilde{H}^{\widetilde{\mathcal{C}}}(\boldsymbol{X}_{h-r_n, h+r_n}/u_n)]$$
$$= \sum_{h=1}^{r} \mathbb{E}[H(\boldsymbol{Y})\widetilde{H}(\boldsymbol{Y}) \mathbb{1}\{\mathcal{C}(\boldsymbol{Y}) = 0\} \mathbb{1}\left\{\widetilde{\mathcal{C}}(\boldsymbol{Y}) = h\right\}] = \sum_{h=1}^{r} \mathcal{I}(H, \widetilde{H}, \mathcal{C}, \widetilde{\mathcal{C}}, h)$$

and by (5.7) the last term vanishes.

For the term $\sum_{h=r+1}^{4r_n}$ we apply $\mathcal{S}(r_n, u_n)$; see the argument used in (5.17).

This finishes the proof of the lemma. □

Proof of Lemma 3.4. Recall that

$$\widetilde{H}_j = H\left(\boldsymbol{X}_{jr_n+1,(j+1)r_n}/u_n\right) \ .$$

Here, $\widetilde{H}_j$ is a function of the jth block $\boldsymbol{X}_{jr_n+1,(j+1)r_n}$, $j = 0, \dots, m_n - 1$. Since $H^{\mathcal{C}}_{n,j}$, $j = 0, \dots, m_n - 1$, is a function of the block $\boldsymbol{X}_{(j-1)r_n+1,\dots,(j+2)r_n}$ (recall that we assumed that we have data $\boldsymbol{X}_{1-r_n}, \dots, \boldsymbol{X}_{n+r_n}$), for $|q| \ge 3$,

$$\mathrm{cov}(H^{\mathcal{C}}_{n,j}, \widetilde{H}_{j+q}) \le \|H\| \|\widetilde{H}\| \beta_{(|q|-2)r_n} \ ; \tag{5.20}$$

cf. (5.1). We have

$$k_n \mathrm{cov}\left(\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}), \widetilde{\boldsymbol{\nu}}^*_{n,r_n}(\widetilde{H})\right) = \frac{1}{r_n v_n} \mathrm{cov}\left(H^{\mathcal{C}}_{n,0}, \widetilde{H}_0\right)$$

$$+\frac{1}{r_n v_n}\sum_{j=1}^{m_n-1}\left(1-\frac{j}{m_n}\right)\left\{\operatorname{cov}(H^{\mathcal{C}}_{n,0},\widetilde{H}_j)+\operatorname{cov}(\widetilde{H}_0,H^{\mathcal{C}}_{n,j})\right\}\ . \tag{5.21}$$

We analyse cov $\left(H^{\mathcal{C}}_{n,0},\widetilde{H}_0\right)$. We are in the Situation 2:

$$\begin{aligned}
&\frac{1}{r_n v_n}\mathbb{E}\left[H^{\mathcal{C}}_{n,0}\widetilde{H}_0\right]\\
&=\frac{1}{r_n v_n}\sum_{i=1}^{r_n}\mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{1,r_n}/u_n\right)\right]\\
&=\frac{1}{r_n v_n}\sum_{i=1}^{r_n}\mathbb{E}\Big[H\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{i-r_n,i+r_n}\right)=0\}.\\
&\qquad\qquad\mathbb{1}\{|\boldsymbol{X}_i|>u_n\}\widetilde{H}\left(\boldsymbol{X}_{1,r_n}/u_n\right)\Big]\\
&=\frac{1}{r_n v_n}\sum_{i=1}^{r_n}\mathbb{E}\Big[H\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}\right)=0\}\\
&\qquad\qquad\mathbb{1}\{|\boldsymbol{X}_0|>u_n\}\widetilde{H}\left(\boldsymbol{X}_{1-i,r_n-i}/u_n\right)\Big]\\
&=\frac{1}{r_n}\sum_{i=1}^{r_n}\mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{1-i,r_n-i}/u_n\right)\mid|\boldsymbol{X}_0|>u_n\right]=\int_0^1 h_{n,0}(\xi)\mathrm{d}\xi
\end{aligned}$$

with

$$h_{n,0}(\xi)=\mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{1-[\xi r_n],r_n-[\xi r_n]}/u_n\right)\mid|\boldsymbol{X}_0|>u_n\right],\ \xi\in(0,1)\ .$$

Note that the third equality follows by stationarity. By (5.13), for each $\xi\in(0,1)$, $h_{n,0}(\xi)\to\boldsymbol{\nu}^*(H\widetilde{H})$. Furthermore, the sequence $\{h_{n,0},n\geq 1\}$ is uniformly bounded in n and ξ. Thus, with help of $\mathcal{R}(r_n,u_n)$,

$$\lim_{n\to\infty}\frac{1}{r_n v_n}\operatorname{cov}(H^{\mathcal{C}}_{n,0},\widetilde{H}_0)=\lim_{n\to\infty}\frac{1}{r_n v_n}\mathbb{E}[H^{\mathcal{C}}_{n,0}\widetilde{H}_0]=\boldsymbol{\nu}^*(H\widetilde{H})\ .$$

The other covariances vanish. Indeed, we analyse $\operatorname{cov}(H^{\mathcal{C}}_{n,0},\widetilde{H}_j)$, $j\geq 1$. We have, using again the stationarity as above,

$$\begin{aligned}
&\frac{1}{r_n v_n}\mathbb{E}[H^{\mathcal{C}}_{n,0},\widetilde{H}_j]\\
&=\frac{1}{r_n v_n}\sum_{i=1}^{r_n}\mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{jr_n+1,(j+1)r_n}/u_n\right)\right]\\
&=\frac{1}{r_n v_n}\sum_{i=1}^{r_n}\mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{jr_n+1-i,(j+1)r_n-i}/u_n\right)\right]\\
&=\int_0^1 h_{n,j}(\xi)\mathrm{d}\xi
\end{aligned}$$

with a function $h_{n,j}$ defined on $(0,1)$ by

$$h_{n,j}(\xi) = \mathbb{E}\left[H^{\mathcal{C}}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\widetilde{H}\left(\boldsymbol{X}_{jr_n-[\xi r_n]+1,(j+1)r_n-[\xi r_n]}/u_n\right) \mid |\boldsymbol{X}_0| > u_n\right] .$$

Until now we proceeded as in the case $j=0$ above. However, now we use (5.3). For each $\xi \in (0,1)$, $jr_n - [\xi r_n] \to +\infty$. Hence, $h_{n,j}(\xi) \to 0$. Bounded convergence and $\mathcal{R}(r_n, u_n)$ give

$$\lim_{n\to\infty} \frac{1}{r_n v_n} \operatorname{cov}(H^{\mathcal{C}}_{n,0}, \widetilde{H}_j) = 0 . \tag{5.22}$$

The same idea applies to $\operatorname{cov}(\widetilde{H}_0, H^{\mathcal{C}}_{n,j})$, $j \geq 1$:

$$\lim_{n\to\infty} \frac{1}{r_n v_n} \operatorname{cov}(\widetilde{H}_0, H^{\mathcal{C}}_{n,j}) = 0 . \tag{5.23}$$

Now, by (5.22)-(5.23), the terms that correspond to $j = 1, 2$ in (5.21) vanish, while (5.20) and (3.6c) give

$$\begin{aligned} &\frac{1}{r_n v_n} \sum_{j=3}^{m_n-1} \left\{\operatorname{cov}(H^{\mathcal{C}}_{n,0}, \widetilde{H}_j) + \operatorname{cov}(\widetilde{H}_0, H^{\mathcal{C}}_{n,j})\right\} \\ &= O(1) \frac{1}{r_n v_n} \sum_{j=1}^{\infty} \beta_{jr_n} = o(1) . \end{aligned}$$

□

5.5. *Empirical cluster process of runs statistics*

Recall that

$$H^{\mathcal{C}}(\boldsymbol{x}) = H(\boldsymbol{x})\mathbb{1}\{\mathcal{C}(\boldsymbol{x}) = 0\}\mathbb{1}\{|\boldsymbol{x}_0| > 1\} .$$

Define

$$H^{\mathcal{C}}_s(\boldsymbol{x}) = H(\boldsymbol{x}/s)\mathbb{1}\{\mathcal{C}(\boldsymbol{x}/s) = 0\}\mathbb{1}\{|\boldsymbol{x}_0| > s\} . \tag{5.24}$$

Recall that $0 < s_0 < 1 < t_0 < \infty$. Recall also that $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$. Define also the classical tail empirical process by

$$\mathbb{T}_n(s) = \sqrt{k_n}\left\{\frac{\sum_{j=1}^n \mathbb{1}\{|\boldsymbol{X}_j| > s u_n\}}{k_n} - s^{-\alpha}\right\} , \quad s \in [s_0, t_0] .$$

In order to deal with asymptotic normality of runs estimators, we study the empirical process

$$\begin{aligned} \mathbb{F}_n(H^{\mathcal{C}}_s) &:= \sqrt{k_n}\left\{\widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}_s) - \boldsymbol{\nu}^*(H_s)\right\} \\ &= \sqrt{k_n}\left\{\frac{\sum_{i=1}^n H^{\mathcal{C}}_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)}{k_n} - s^{-\alpha}\boldsymbol{\nu}^*(H)\right\} . \end{aligned}$$

The process $\mathbb{F}_n(H^{\mathcal{C}}_s)$ is viewed as a random element with values in $\mathbb{D}([s_0, t_0])$. The next result is crucial to establish convergence of runs estimators.

Theorem 5.6. *Let $\{\boldsymbol{X}_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying $\mathbb{R}^d$-valued time series. Assume that $\mathcal{R}(r_n, u_n)$, $\beta'(r_n)$, $\mathcal{S}(r_n, u_n)$, (3.11) and (3.12b) hold. Suppose that Assumption 3.5 is satisfied.*

Then $\mathbb{F}_n(H^{\mathcal{C}}_{\cdot})$ converges weakly in $(\mathbb{D}([s_0, t_0]), J_1)$ to a Gaussian process $\mathbb{G}(H_{\cdot})$ with the covariance $\boldsymbol{\nu}^*(H_s H_t)$. If moreover ANSJB$(r_n, u_n)$ is satisfied, then the convergence holds for $H \in \mathcal{B}$. If additionally (3.12a) is satisfied, then the processes $\mathbb{F}_n(H^{\mathcal{C}}_{\cdot})$ and $\mathbb{T}_n(\cdot)$ converge jointly to $(\mathbb{G}(H_{\cdot}), \mathbb{G}(\mathcal{E}_{\cdot}))$.

Remark 5.7. The proof of Theorem 5.6 consists of two parts: fidi convergence (Section 5.7) and tightness (Section 5.8). We note that once the behaviour of covariances is established (see Lemmas 3.3 and 3.4) the fidi portion of the proof can be in principle omitted thanks to Theorem 2.1 and 2.3 in [DN21] (see the proof of Proposition B.4 there). This would require some modifications in the assumptions. As such, we keep the proof of fidi convergence for completeness.

5.6. Proof of Theorem 3.6

Write $\psi_n = |\boldsymbol{X}|_{(n:n-k_n)}/u_n$. Since $k_n = n\mathbb{P}(|\boldsymbol{X}_0| > u_n)$, we can rewrite $\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}})$ as $\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) = \widetilde{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}_{\psi_n})$ (cf. (3.4)-(3.5)). Therefore,

$$\sqrt{k_n}\left\{\widehat{\boldsymbol{\xi}}^*_{n,r_n}(H^{\mathcal{C}}) - \boldsymbol{\nu}^*(H)\right\} = \mathbb{F}_n(H^{\mathcal{C}}_{\psi_n}) + \sqrt{k_n}\left\{\boldsymbol{\nu}^*(H_{\psi_n}) - \boldsymbol{\nu}^*(H)\right\} .$$

We have local uniform convergence of $\{\mathbb{F}_n(H^{\mathcal{C}}_s),\ s \in [s_0, t_0]\}$ to a continuous Gaussian process $\mathbb{G}$ thanks to Theorem 5.6. Moreover, the convergence of $\{\mathbb{T}_n(\cdot),\ s \in [s_0, t_0]\}$ yields $\psi_n \xrightarrow{\mathrm{d}} 1$, jointly with $\mathbb{F}_n(H^{\mathcal{C}}_s)$. Therefore,

$\mathbb{F}_n(H^{\mathcal{C}}_{\psi_n}) \xrightarrow{\mathrm{d}} \mathbb{G}(H)$. Using Vervaat's theorem, we have, jointly with the previous convergence, $\sqrt{k_n}(\psi_n^{-\alpha} - 1) \xrightarrow{\mathrm{d}} -\mathbb{G}(\mathcal{E})$. Therefore, by the homogeneity of $\boldsymbol{\nu}^*$,

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

Since the convergence hold jointly, we conclude the result.

5.7. Proof of Theorem 5.6 – fidi convergence

Recall the disjoint blocks of size r_n (cf. (3.2)):

$$J_j := \{jr_n + 1, \ldots, (j+1)r_n\} , \quad j = 0, \ldots, m_n - 1 .$$

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

$$\lim_{n\to\infty} z_n r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n) = \lim_{n\to\infty} \frac{z_n r_n}{\sqrt{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}} = \lim_{n\to\infty} \frac{z_n r_n}{\sqrt{k_n}} = 0 . \quad (5.25)$$

This is possible thanks to the assumptions $\mathcal{R}(r_n, u_n)$ and (3.11). We note that this assumption is needed for the Lindeberg condition only. Set

$$\widetilde{m}_n = \frac{q_n}{(z_n+3)r_n} = \frac{n-r_n}{(z_n+3)r_n} \sim \frac{n}{z_n r_n}$$

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:

$$\begin{aligned} &L_1 = \{1, \ldots, z_n r_n\}\ , \quad S_1 = \{z_n r_n + 1 \ldots, z_n r_n + 3r_n\}\ , \\ &L_j = \{(j-1)z_n r_n + 3(j-1)r_n + 1, \ldots, j z_n r_n + 3(j-1)r_n\}\ , \\ &S_j = \{j z_n r_n + 3(j-1)r_n + 1, \ldots, j z_n r_n + 3jr_n\}\ . \end{aligned}$$

The block L_1 is obtained by merging z_n consecutive blocks $J_0, \ldots, J_{z_n-1}$ of size r_n. Likewise, $S_1 = J_{z_n} \cup J_{z_n+1} \cup J_{z_n+2}$. Therefore, the large block of size $z_n r_n$ is followed by the small block of size $3r_n$, which in turn is followed by the large block of size $z_n r_n$ and so on. All together,

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

Write

$$\begin{aligned} &\sum_{i=1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) \qquad (5.26) \\ &= \sum_{i=1}^{q_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) + \sum_{i=q_n+1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) \\ &= \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(l)}(H^{\mathcal{C}}) + \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(s)}(H^{\mathcal{C}}) + W_n\ , \qquad (5.27) \end{aligned}$$

where now

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

and

$$W_n = \sum_{i=q_n+1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) = \sum_{i=n-r_n+1}^{n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\ .$$

With such the decomposition, $\boldsymbol{X}_{1-r_n}, \ldots, \boldsymbol{X}_{z_n r_n + r_n}$ used in the definition of $\Psi_1^{(l)}(H^{\mathcal{C}})$ are separated by at least r_n from the random variables that define $\Psi_2^{(l)}(H^{\mathcal{C}})$. The mixing condition (3.6a) allows us to replace $\boldsymbol{X}$ with the independent blocks process, that is, we can treat the random variables $\Psi_j^{(l)}(H^{\mathcal{C}})$, $j = 1, \ldots, \widetilde{m}_n$, as independent. The same applies to $\Psi_j^{(s)}(H^{\mathcal{C}})$.

Set

$$\mathbb{Z}_n(H^{\mathcal{C}}) = \sum_{j=1}^{\widetilde{m}_n} \left\{ Z_{n,j}(H^{\mathcal{C}}) - \mathbb{E}[Z_{n,j}(H^{\mathcal{C}})] \right\} =: \sum_{j=1}^{\widetilde{m}_n} \bar{Z}_{n,j}(H^{\mathcal{C}}) \tag{5.28}$$

with

$$Z_{n,j}(H^{\mathcal{C}}) = \frac{1}{\sqrt{k_n}} \Psi_j^{(l)}(H^{\mathcal{C}}) \ . \tag{5.29}$$

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 scaled second term in (5.26)) is negligible;
- We show that the boundary term W_n is also negligible;
- Finally, we will verify the Lindeberg condition for the large blocks process.

Again, to shorten the notation, we set $v_n = \mathbb{P}(|\boldsymbol{X}_0| > u_n)$. Hence, $k_n = nv_n$.

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

$$\begin{aligned} &\operatorname{var}\left(\frac{1}{\sqrt{k_n}} \sum_{j=1}^{\widetilde{m}_n} \Psi_j^{(l)}(H^{\mathcal{C}}) \right) = \frac{\widetilde{m}_n}{k_n} \operatorname{var}(\Psi_1^{(l)}(H^{\mathcal{C}})) \\ &\sim \frac{1}{z_n r_n v_n} \operatorname{var}\left(\sum_{i=1}^{z_n r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) \right) = \frac{1}{z_n r_n v_n} \operatorname{var}\left(\sum_{j=0}^{z_n-1} H^{\mathcal{C}}_{n,j} \right) , \end{aligned} \tag{5.30}$$

where $H^{\mathcal{C}}_{n,j}$ is defined in (3.3) and where in the last line we decomposed the block $L_1 = \{1, \ldots, z_n r_n\}$ into z_n disjoint blocks $J_0, \ldots, J_{z_n-1}$ and $\widetilde{m}_n \sim m_n/z_n$. The next steps follow easily from (5.14) with m_n replaced by z_n.

The term in (5.30) becomes

$$\frac{\operatorname{var}\left(H^{\mathcal{C}}_{n,0}\right)}{r_n v_n} + \frac{2}{r_n v_n} \sum_{j=1}^{z_n-1} \left(1 - \frac{j}{z_n}\right) \operatorname{cov}(H^{\mathcal{C}}_{n,0}, H^{\mathcal{C}}_{n,j}) \ . \tag{5.31}$$

It follows immediately from (5.15) that the limit of the first term above is

$$\lim_{n\to\infty} \frac{\operatorname{var}\left(H^{\mathcal{C}}_{n,0}\right)}{r_n \mathbb{P}(|\boldsymbol{X}_0| > u_n)} = \boldsymbol{\nu}^*(H^2) \ . \tag{5.32}$$

Now, for the second term in (5.31) we adapt the proof of Lemma 3.3 from m_n to z_n.

As in (5.18), for $j \geq 1$,

$$\frac{1}{r_n v_n}\left|\mathrm{cov}(H^{\mathcal{C}}_{n,0}, H^{\mathcal{C}}_{n,j})\right| \leq \sum_{h=(j-1)r_n+1}^{(j+1)r_n} |g_n(h)| =: I_j$$

with (this time)

$$g_n(h) = \frac{1}{v_n}\mathrm{cov}(H^{\mathcal{C}}(\boldsymbol{X}_{-r_n,r_n}), H^{\mathcal{C}}(\boldsymbol{X}_{h-r_n,h+r_n})) \ .$$

For $h > 2r_n$, similarly to (5.19), we have by (5.1),

$$|g_n(h)| \leq \frac{\|H\|_\infty^2}{v_n}\beta_{h-2r_n} \ .$$

Thus,

$$\frac{1}{r_n v_n}\sum_{j=4}^{z_n-1}|\mathrm{cov}(H^{\mathcal{C}}_{n,0}, H^{\mathcal{C}}_{n,j})| \leq \frac{\|H\|_\infty^2}{v_n}\sum_{j=4}^{z_n-1}\sum_{h=(j-1)r_n+1}^{(j+1)r_n}\beta_{h-2r_n}$$
$$\leq 2\frac{\|H\|_\infty^2}{v_n}\sum_{h=3r_n+1}^{\infty}\beta_{h-2r_n} = O(1)\frac{1}{v_n}\sum_{i=r_n+1}^{\infty}\beta_i = o(1)$$

by the assumption (3.6b). The terms that correspond to $j = 1, 2, 3$ in (5.14) are negligible.

In summary, we showed that

$$\lim_{n\to\infty}\mathrm{var}\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^{\widetilde{m}_n}\Psi_j^{(l)}(H^{\mathcal{C}})\right) = \boldsymbol{\nu}^*(H^2) \ .$$

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

$$\mathrm{var}\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^{\widetilde{m}_n}\Psi_j^{(s)}(H^{\mathcal{C}})\right) = \frac{\widetilde{m}_n}{k_n}\mathrm{var}(\Psi_1^{(s)}(H^{\mathcal{C}})) \sim \frac{1}{z_n r_n v_n}\mathrm{var}(\Psi_1^{(s)}(H^{\mathcal{C}})) \ .$$

Since the size of $\Psi_1^{(s)}(H^{\mathcal{C}})$ is 3 times the size of $H^{\mathcal{C}}_{n,1}$ defined in (3.3), we have by (5.15)

$$\mathrm{var}\left(\frac{1}{\sqrt{k_n}}\sum_{j=1}^{\widetilde{m}_n}\Psi_j^{(s)}(H^{\mathcal{C}})\right) \sim \frac{1}{z_n r_n v_n}r_n v_n \boldsymbol{\nu}^*(H^2) = O(1/z_n) = o(1) \ .$$

Variance of the boundary term W_n. We have (cf. (3.3))

$$\operatorname{var}\left(\frac{1}{\sqrt{k_n}}W_n\right) = \operatorname{var}\left(\frac{1}{\sqrt{k_n}}\sum_{i=1}^{r_n} H^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\right)$$
$$= \frac{\operatorname{var}(H^{\mathcal{C}}_{n,0})}{k_n} = \frac{\operatorname{var}(H^{\mathcal{C}}_{n,0})}{nv_n} .$$

The latter term vanishes when $n \to \infty$, using (5.32) and $r_n/n \to 0$.

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

$$\lim_{n\to\infty} \widetilde{m}_n \mathbb{E}\left[Z^2_{n,1}(H^{\mathcal{C}})\mathbb{1}\left\{|Z_{n,1}(H^{\mathcal{C}})| > \eta\right\}\right] = 0 . \tag{5.33}$$

Since H is bounded, then by (5.25),

$$|Z_{n,1}(H^{\mathcal{C}})| \le \frac{\sqrt{k_n}z_n r_n}{nv_n}\|H\|_\infty \sim \frac{z_n r_n}{\sqrt{nv_n}}\|H\|_\infty = o(1) . \tag{5.34}$$

Thus, the indicator in (5.33) becomes zero for large n.

5.8. *Proof of Theorem 5.6 – 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

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

where $\{Z_{n,j}, n \ge 1\}$, $j = 1, \dots, \widetilde{m}_n$, 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 5.8. *Assume that $(\mathcal{G}, \rho)$ is totally bounded. Assume moreover that:*

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

$$\lim_{n\to\infty} \widetilde{m}_n \mathbb{E}[\|Z_{n,1}\|^2_{\mathcal{G}}\mathbb{1}\{\|Z_{n,1}\|^2_{\mathcal{G}} > \eta\}] = 0 . \tag{5.35}$$

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

$$\lim_{n\to\infty} \sup_{\substack{f,g\in\mathcal{G}\\ \rho(f,g)\le\delta_n}} \mathbb{E}[d_n^2(f,g)] = 0 . \tag{5.36}$$

(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,

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

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

5.8.1. *Asymptotic equicontinuity of the empirical process of sliding blocks*

Recall the big-blocks process $\mathbb{Z}_n(H^{\mathcal{C}})$ (cf. (5.28)-(5.29)). Recall also that thanks to the β-mixing we can consider random variables $\Psi_j^{(l)}(H^{\mathcal{C}})$, $j = 1, \ldots, \widetilde{m}_n$ to be independent. Recall that $H_s^{\mathcal{C}}$ is defined in (5.24). We need to prove the asymptotic equicontinuity of $\mathbb{Z}_n(H_s^{\mathcal{C}})$ indexed by the class $\mathcal{G} = \{H_s^{\mathcal{C}}, s \in [s_0, t_0]\}$ equipped with the metric $\boldsymbol{\rho}^*(H_s^{\mathcal{C}}, H_t^{\mathcal{C}}) = \boldsymbol{\nu}^*(\{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}^2)$. The same argument can be used to prove the asymptotic equicontinuity for the small blocks process. This yields asymptotic equicontinuity of $\mathbb{F}_n(H_s^{\mathcal{C}})$.

In what follows, the proof of the Lindeberg-type condition (5.35) is easy. The proof of (5.36) is quite involved.

Thanks to Assumption 3.5, the condition (3.10) is satisfied. Its validity is discussed in Section 5.9.

Lindeberg condition: Proof of (5.35). We re-write (5.34) as follows:

$$\sup_{s \in [s_0, t_0]} |Z_{n,1}(H_s^{\mathcal{C}})| \leq \frac{\sqrt{k_n} z_n r_n}{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)} \|H\|_\infty \leq \frac{z_n r_n}{\sqrt{n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}} \sup_{s \in [s_0, t_0]} \|H_s\|_\infty \,,$$

Since the class $\{H_s : s \in [s_0, t_0]\}$ is linearly ordered, $\sup_{s \in [s_0, t_0]} \|H_s\|_\infty$ is achieved either at $s = s_0$ or $s = t_0$. Hence, the Lindeberg condition (i) of

Lemma 5.8 holds by (5.33).

Asymptotic continuity of random semi-metric: Proof of (5.36). The proof is rather long and technical. Again, to shorten the notation we set $v_n = \mathbb{P}(|\boldsymbol{X}_0| > u_n)$.

Define the random metric

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

Let (cf. (3.3))

$$H_{s,n,j}^{\mathcal{C}} = \sum_{i=jr_n+1}^{(j+1)r_n} H\left(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n)\right) \mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n)\right) = i\} \mathbb{1}\left\{|\boldsymbol{X}_i| > su_n\right\} .$$

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

$$\begin{aligned}
&\mathbb{E}[d_n^2(H_s^{\mathcal{C}}, H_t^{\mathcal{C}})] \\
&= \frac{k_n \widetilde{m}_n}{(nv_n)^2} \mathbb{E}\left[\left(\sum_{i=1}^{z_n r_n} \left\{H_s^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right) - H_t^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\right\}\right)^2\right] \\
&\sim \frac{1}{z_n r_n v_n} \mathbb{E}\left[\left(\sum_{j=0}^{z_n-1} \left\{H_{s,n,j}^{\mathcal{C}} - H_{t,n,j}^{\mathcal{C}}\right\}\right)^2\right] , \qquad (5.37)
\end{aligned}$$

where in the last line we decomposed the block L_1 into z_n disjoint blocks $J_0, \dots, J_{z_n-1}$, $\widetilde{m}_n \sim m_n/z_n$; cf. (5.30). The term in (5.37) becomes

$$\begin{aligned}
&\frac{\mathbb{E}[\left\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\right\}^2]}{r_n v_n} \\
&+ 2\frac{1}{r_n v_n} \sum_{j=1}^{z_n-1} \left(1 - \frac{j}{z_n}\right) \mathbb{E}[\left\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\right\} \left\{H_{s,n,j}^{\mathcal{C}} - H_{t,n,j}^{\mathcal{C}}\right\}]
\end{aligned}$$

The above lines correspond to (5.14) with m_n replaced by z_n.

We are going to prove two statements:

$$\lim_{n\to\infty} \sup_{\substack{s_0 \le s,t \le t_0 \\ |s-t| \le \delta_n}} \frac{\mathbb{E}[\left\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\right\}^2]}{r_n v_n} = 0 \qquad (5.38)$$

and

$$\lim_{n\to\infty} \sup_{\substack{s_0 \le s,t \le t_0 \\ |s-t| \le \delta_n}} \frac{1}{r_n v_n} \sum_{j=1}^{z_n-1} \left(1 - \frac{j}{z_n}\right) \mathbb{E}[\left\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\right\} \left\{H_{s,n,j}^{\mathcal{C}} - H_{t,n,j}^{\mathcal{C}}\right\}] = 0 . \qquad (5.39)$$

Proof of (5.38). We will write $\{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}(\boldsymbol{x})$ for $H_s^{\mathcal{C}}(\boldsymbol{x}) - H_t^{\mathcal{C}}(\boldsymbol{x})$.

Similarly to (5.15),

$$\frac{\mathbb{E}[\left\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\right\}^2]}{r_n v_n}$$

$$\leq \|H\| \frac{1}{v_n} \sum_{h=-r_n}^{r_n} \left| \mathbb{E}[\{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}(\boldsymbol{X}_{-r_n,r_n}/u_n) \times \{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)] \right|$$

$$=: \|H\| \sum_{h=-r_n}^{r_n} |g_n(h, H_s^{\mathcal{C}} - H_t^{\mathcal{C}})| \tag{5.40}$$

with

$$|g_n(h,G)| = \left| \frac{1}{v_n} \mathbb{E}[G(\boldsymbol{X}_{-r_n,r_n}/u_n) G(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)] \right| . \tag{5.41}$$

Using the definition (5.24) of $H_s^{\mathcal{C}}$, the fact that $s,t \geq s_0$ and since H is bounded, we immediately get

$$|g_n(h, H_s^{\mathcal{C}} - H_t^{\mathcal{C}})| \leq \frac{4}{v_n} \|H\|^2 \mathbb{P}(|\boldsymbol{X}_0| > s_0 u_n, |\boldsymbol{X}_h| > s_0 u_n) . \tag{5.42}$$

To get a more precise bound that involves the difference $s-t$ we need to consider two cases. The reason for this is that we need to keep the absolute value in (5.41) outside of the expectation. As such, computations below are quite technically involved.

To shorten our displays, we introduce the notation

$$\mathcal{J}(i,s) := \mathbb{1}\{|\boldsymbol{X}_i| > s u_n\} . \tag{5.43}$$

Case 1. Assume here that $\mathcal{C}$ is 0-homogeneous. Then for any i,

$$\begin{aligned} &H_s^{\mathcal{C}}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) \\ &= H(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n)) \mathbb{1}\{\mathcal{C}(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n)) = i\} \mathcal{J}(i,s) \\ &= H_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) \mathbb{1}\{\mathcal{C}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) = i\} \mathcal{J}(i,s) ; \end{aligned} \tag{5.44}$$

(we keep u_n in the argument of $\mathcal{C}$, although it can be omitted). What is important in this decomposition is that we can control monotonicity (withe respect to s) of each term.

Then

$$\begin{aligned} &H_s^{\mathcal{C}}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) - H_t^{\mathcal{C}}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) = \\ &= \mathbb{1}\{\mathcal{C}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) = i\} H_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) \Big(\mathcal{J}(i,s) - \mathcal{J}(i,t)\Big) \\ &\quad + \mathbb{1}\{\mathcal{C}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) = i\} \mathcal{J}(i,t) \Big(H_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) - H_t(\boldsymbol{X}_{i-r_n,i+r_n}/u_n)\Big) \\ &=: T_1(i) + T_2(i) \end{aligned}$$

and hence

$$\left| \mathbb{E}\left[\{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}(\boldsymbol{X}_{-r_n,r_n}/u_n) \{H_s^{\mathcal{C}} - H_t^{\mathcal{C}}\}(\boldsymbol{X}_{h-r_n,h+r_n}/u_n) \right] \right|$$

is bounded by the sum of four nonnegative terms $W_{11} + W_{22} + W_{12} + W_{21}$ that we are going to define below. The general idea is that we will obtain rough bounds in terms of the difference $\Big(\mathcal{J}(i,s) - \mathcal{J}(i,t)\Big)$, except of one case which involves the product $T_2(0)T_2(h)$. Coming back to the definitions of $W_{i,j}$, the double sub-index $\{12\}$ of W indicates that W_{12} is related to multiplying $T_1(0)$ by $T_2(h)$; W_{21} means we multiply $T_2(0)$ and $T_1(h)$):

$$W_{11} := \|H\|^2 \mathbb{E}\left[(\mathcal{J}(0,s) - \mathcal{J}(0,t)) (\mathcal{J}(h,s) - \mathcal{J}(h,t)) \right] ,$$

(above, the indicators of the anchoring map are omitted);

$$W_{22} := \mathbb{E}\left[\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)=0\}\mathcal{J}(0,t)\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\right. \\ \left.\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right)\right] ,$$

(above, $\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right)=h\}$ and $\mathcal{J}(h,t)$ are omitted);

$$W_{12} := \pm\|H\|\mathbb{E}\left[\Big(\mathcal{J}(0,s)-\mathcal{J}(0,t)\Big)\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right)\right] ,$$

(both indicators of the anchoring maps and $\mathcal{J}(h,t)$ are omitted);

$$W_{21} := \pm\|H\|\mathbb{E}\left[\Big(\mathcal{J}(h,s)-\mathcal{J}(h,t)\Big)\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\right] ,$$

(both indicators of the anchoring maps and $\mathcal{J}(0,t)$ are omitted).

Note that the right-hand side of both W_{11}, W_{22} is nonnegative thanks to the monotonicity and $s<t$, while for the right-hand side of W_{12}, W_{21} we need to put $\pm$, since the sign of the expressions there depends on whether the map $s\to H_s$ is decreasing or increasing.

Recall that $v_n=\mathbb{P}(|\boldsymbol{X}_0|>u_n)$. Set also $v_n(s)=\mathbb{P}(|\boldsymbol{X}_0|>su_n)$, $s>0$. Thus,

$$W_{11} \le \|H\|^2\left(v_n(s)-v_n(t)\right) , \tag{5.45}$$

$$W_{22} \le \pm 2\|H\|\mathbb{E}\left[\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)=0\}\mathcal{J}(0,t)\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\right] , \tag{5.46}$$

$$W_{12}+W_{21} \le 4\|H\|^2\left(v_n(s)-v_n(t)\right) . \tag{5.47}$$

The bound on W_{22} (with $+$) is obvious if $H_s-H_t\ge 0$ (thus, $s\to H_s$ is decreasing), while in an increasing case of $s\to H_s$ we use the following observation: if $a,b<0$, $|b|<c$, then $ab\le -ac$ (yielding $-$ on the right-hand side of (5.46)). The bound on $W_{12}+W_{21}$ follows from the same reasoning.

In summary, with $g_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})$ defined in (5.41), we have

$$|g_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})| \le 5\|H\|^2\frac{v_n(s)-v_n(t)}{v_n} \tag{5.48}$$

$$\pm 2\|H\|\frac{\mathbb{E}\left[\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)=0\}\mathbb{1}\left\{|\boldsymbol{X}_0|>s_0u_n\right\}\left\{H_s-H_t\right\}\left(\boldsymbol{X}_{-r_n,r_n}/u_n\right)\right]}{v_n} , \tag{5.49}$$

where again the presence of $\pm$ depends on the sign of H_s-H_t.

We can ignore the scaling factor $2\|H\|$ in (5.49) and write it as (recall that the anchoring map is 0-homogeneous)

$$\begin{aligned}
&\mathbb{E}[H_{s/s_0}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)=0\}\mid|\boldsymbol{X}_0|>s_0u_n]\frac{v_n(s_0)}{v_n} \\
&\quad-\mathbb{E}[H_{t/s_0}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)=0\}\mid|\boldsymbol{X}_0|>s_0u_n]\frac{v_n(s_0)}{v_n} \\
&=\Big(\mu_{n,r_n}(s)-\mu_{n,r_n}(t)\Big)\frac{v_n(s_0)}{v_n}=:\Big(\tilde{\mu}_{n,r_n}(s)-\tilde{\mu}_{n,r_n}(t)\Big) ,
\end{aligned}$$

with

$$\mu_{n,r_n}(\cdot)=\mathbb{E}[H_{\cdot/s_0}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)\mathbb{1}\{\mathcal{C}\left(\boldsymbol{X}_{-r_n,r_n}/(s_0u_n)\right)=0\}\mid|\boldsymbol{X}_0|>s_0u_n]$$

and

$$\tilde{\mu}_{n,r_n}(s) = \mu_{n,r_n}(s)\frac{v_n(s_0)}{v_n} \,.$$

Thanks to (5.5), $\lim_{n\to\infty}\mu_{n,r_n}(s) = \boldsymbol{\nu}^*(H_{s/s_0})$. Thanks to the monotonicity of $s \to H_S$ and homogeneity of $\boldsymbol{\nu}^*$, the convergence of $\tilde{\mu}_{n,r_n}(s)$ to

$$s_0^{-\alpha}\boldsymbol{\nu}^*(H_{s/s_0}) = s_0^{-\alpha}(s/s_0)^{-\alpha}\boldsymbol{\nu}^*(H) = s^{-\alpha}\boldsymbol{\nu}^*(H)$$

is uniform on $[s_0, t_0]$. Thus, for $s, t \in [s_0, t_0]$,

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

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

$$\begin{aligned}\sup_{\substack{s_0\leq s,t\leq t_0\\|s-t|\leq\delta_n}}|\tilde{\mu}_{n,r_n}(s) - \tilde{\mu}_{n,r_n}(t)| &\leq \eta + \boldsymbol{\nu}^*(H)\sup_{\substack{s_0\leq s,t\leq t_0\\|s-t|\leq\delta_n}}\{s^{-\alpha} - t^{-\alpha}\}\\ &\leq \eta + \alpha s_0^{-\alpha-1}\delta_n\boldsymbol{\nu}^*(H)\,. \end{aligned} \tag{5.50}$$

The uniform convergence also yields that the term in (5.48) is bounded by $\eta + \alpha s_0^{-\alpha-1}\delta_n$. This, together with (5.50), gives

$$|g_n(h, H_s^{\mathcal{C}} - H_t^{\mathcal{C}})| \leq \eta + \text{cst}\ \delta_n \tag{5.51}$$

with a generic constant cst.

Fix an integer r. Using (5.42) and (5.51) we have

$$\begin{aligned}&\sup_{\substack{s_0\leq s,t\leq t_0\\|s-t|\leq\delta_n}}\frac{\mathbb{E}[\{H_{s,n,0}^{\mathcal{C}} - H_{t,n,0}^{\mathcal{C}}\}^2]}{r_n\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\\ &\leq \sup_{\substack{s_0\leq s,t\leq t_0\\|s-t|\leq\delta_n}}\sum_{h=-r}^{r}|g_n(h, H_s^{\mathcal{C}} - H_t^{\mathcal{C}})| + \sup_{\substack{s_0\leq s,t\leq t_0\\|s-t|\leq\delta_n}}\sum_{|h|=r}^{r_n}|g_n(h, H_s^{\mathcal{C}} - H_t^{\mathcal{C}})|\\ &\leq \text{cst}\ r(\eta+\delta_n) + \frac{4}{\mathbb{P}(|\boldsymbol{X}_0| > u_n)}\sum_{|h|=r}^{r_n}\mathbb{P}(|\boldsymbol{X}_0| > s_0u_n, |\boldsymbol{X}_h| > s_0u_n)\,.\end{aligned}$$

Applying the anticlustering conditions $\mathcal{S}(r_n, u_n)$ to the second term, letting $\delta_n \to 0$, since η is arbitrary, this proves (5.38).

Case 2. Now, we consider the anchoring maps $\mathcal{C}^{(1)}$ and $\mathcal{C}^{(2)}$ which are not 0-homogeneous. Note that for $j = 1, 2$ we can write

$$\begin{aligned}&H_s^{\mathcal{C}^{(j)}}(\boldsymbol{X}_{i-r_n,i+r_n}/u_n)\\ &= H(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n))\mathbb{1}\Big\{\mathcal{C}^{(j)}(\boldsymbol{X}_{i-r_n,i+r_n}/(su_n)) = i\Big\}\mathcal{J}(i,s)\\ &= H_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n)\mathcal{J}(i,s)F_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n)\ ,\end{aligned}$$

where

$$F_s(\boldsymbol{X}_{i-r_n,i+r_n}/u_n) = \mathbb{1}\{\boldsymbol{X}^*_{i-r_n,i-1} \leq su_n\}$$

or

$$F_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)=\mathbb{1}\left\{\boldsymbol{X}^*_{i+1,i+r_n}\le su_n\right\}$$

in case $j=1$ and $j=2$, respectively. Note that regardless of the monotonicity of the map $s\to\mathcal{C}_s^{(j)}$, the map $s\to F_s$ is always non-decreasing. Then (5.44) gives, for any $i\in\mathbb{N}$, and $\mathcal{C}=\mathcal{C}^{(1)},\mathcal{C}^{(2)}$,

$$\begin{aligned}
&H_s^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)-H_t^{\mathcal{C}}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)=\\
&=H_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)F_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\Big(\mathcal{J}(i,s)-\mathcal{J}(i,t)\Big)\\
&\quad+\mathcal{J}(i,t)F_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\left(H_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)-H_t\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\right)\\
&\quad+\mathcal{J}(i,t)H_t\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\Big(F_s\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)-F_t\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\Big)\\
&=:T_1(i)+T_2(i)+T_3(i)\ .
\end{aligned}$$

Again, in this decomposition we can control monotonicity of each term. The argument now is very similar to that of Case 1. Hence, it is omitted.

Therefore, (5.38) is proved for $\mathcal{C}^{(j)}$, $j=0,1,2$.

Proof of (5.39). We proceed similarly. Recall that for $j\ge 1$ (cf. (5.18)),

$$\begin{aligned}
&\frac{1}{r_nv_n}\mathbb{E}\Big[\Big\{H^{\mathcal{C}}_{s,n,0}-H^{\mathcal{C}}_{t,n,0}\Big\}\Big\{H^{\mathcal{C}}_{s,n,j}-H^{\mathcal{C}}_{t,n,j}\Big\}\Big]\\
&=\frac{1}{r_nv_n}\mathbb{E}\left[\sum_{h=1}^{r_n}\Big\{H_s^{\mathcal{C}}-H_t^{\mathcal{C}}\Big\}\left(\boldsymbol{X}_{h-r_n,h+r_n}/u_n\right)\times\right. &(5.52)\\
&\qquad\qquad\left.\sum_{i=jr_n+1}^{(j+1)r_n}\Big\{H_s^{\mathcal{C}}-H_t^{\mathcal{C}}\Big\}\left(\boldsymbol{X}_{i-r_n,i+r_n}/u_n\right)\right]\\
&=\sum_{h=(j-1)r_n+1}^{jr_n}\left(\frac{h}{r_n}-(j-1)\right)g_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})+ &(5.53)\\
&\qquad\qquad\sum_{h=jr_n+1}^{(j+1)r_n}\left((j+1)-\frac{h}{r_n}\right)g_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})\\
&\le\sum_{h=(j-1)r_n+1}^{(j+1)r_n}g_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})=:I_j(s,t) &(5.54)
\end{aligned}$$

with the same g_n as in (5.41).

Write $g_n(h,G)$ as

$$\begin{aligned}
&\frac{1}{v_n}\mathrm{cov}[G(\boldsymbol{X}_{-r_n,r_n}/u_n),G(\boldsymbol{X}_{h-r_n,h+r_n}/u_n)]+\frac{1}{v_n}\mathbb{E}^2[G(\boldsymbol{X}_{-r_n,r_n}/u_n)]\\
&=:\widetilde{g}_n(h,G)+\frac{1}{v_n}\mathbb{E}^2[G(\boldsymbol{X}_{-r_n,r_n}/u_n)]\ .
\end{aligned}$$

For $h>2r_n$ we have by (5.1),

$$|\widetilde{g}_n(h,H_s^{\mathcal{C}}-H_t^{\mathcal{C}})|\le\frac{\mathrm{cst}}{v_n}\beta_{h-2r_n}\ . \tag{5.55}$$

Thus,

$$\begin{aligned}
&\frac{1}{r_n v_n}\sum_{j=4}^{z_n-1}\left(1-\frac{j}{z_n}\right)\mathbb{E}[\left(H^{\mathcal{C}}_{s,n,0}-H^{\mathcal{C}}_{t,n,0}\right)\left(H^{\mathcal{C}}_{s,n,j}-H^{\mathcal{C}}_{t,n,j}\right)]\\
&\le \frac{\mathrm{cst}}{v_n}\sum_{j=4}^{z_n-1}\sum_{h=(j-1)r_n+1}^{(j+1)r_n}\beta_{h-2r_n}+\frac{\mathrm{cst}}{v_n}z_n r_n\mathbb{E}^2[\{H^{\mathcal{C}}_s-H^{\mathcal{C}}_t\}(\boldsymbol{X}_{-r_n,r_n}/u_n)]\\
&\le \frac{\mathrm{cst}}{v_n}\sum_{h=3r_n+1}^{\infty}\beta_{h-2r_n}+\mathrm{cst}z_n r_n\mathbb{P}(|\boldsymbol{X}_0|>s_0u_n)\\
&=\frac{\mathrm{cst}}{v_n}\sum_{i=r_n+1}^{\infty}\beta_i+o(1)=o(1)
\end{aligned} \tag{5.56}$$

uniformly in $s,t\in[s_0,t_0]$. In the last line we applied the assumption (3.6b), and the assumption (5.25).

The terms that correspond to $j=1,2,3$ in (5.52) have to be dealt with separately. We note that $I_1(s,t)=\sum_{h=1}^{r_n} g_n(h,H^{\mathcal{C}}_s-H^{\mathcal{C}}_t)$ is bounded by the term in (5.40). Hence,

$$\lim_{n\to\infty}\sup_{\substack{s_0\le s,t\le t_0\\|s-t|\le\delta_n}} I_1(s,t)=0\,. \tag{5.57}$$

Next, using (5.42) and $\mathcal{S}(r_n,u_n)$,

$$\begin{aligned}
&I_2(s,t)+I_3(s,t)\\
&\le \mathrm{cst}\sum_{h=r_n+1}^{4r_n} g_n(h)\le \mathrm{cst}\sum_{h=r_n+1}^{4r_n}\mathbb{P}(|\boldsymbol{X}_0|>s_0u_n,|\boldsymbol{X}_h|>s_0u_n)=0\,,
\end{aligned} \tag{5.58}$$

uniformly in $s,t\in[s_0,t_0]$.

Combination of (5.56), (5.57), (5.58) finishes the proof of (5.39).

This, together with (5.38) concluded the proof of (5.36).

5.9. Random entropy

In this section we discuss validity of Assumption 3.5. We cannot check this condition for arbitrary functionals H and anchoring maps $\mathcal{C}$, however, we will see that the conditions is satisfied for most relevant cases considered in the paper.

Recall the class

$$\mathcal{G}=\{H^{\mathcal{C}}_s,s\in[s_0,t_0]\}=\{H(\boldsymbol{x}/s)\mathbb{1}\{\mathcal{C}(\boldsymbol{x}/s)=0\}\mathbb{1}\{|\boldsymbol{x}_0|>s\},\ \ s\in[s_0,t_0]\}.$$

We start first with H of the form

$$H_s(\boldsymbol{x})=\mathbb{1}\{K(\boldsymbol{x})>s\}\,, \tag{5.59}$$

where $K:\mathbb{R}^{\mathbb{Z}}\to\mathbb{R}$. This is the case of the functionals that lead to the extremal index, the large deviation index and the ruin index.

Since $\mathcal{C}^{(0)}$ is 0-homogeneous, it does not play a role in calculating the class entropy. Then

$$H_s(\boldsymbol{x})\mathbb{1}\{|\boldsymbol{x}_0|>s\}=\mathbb{1}\{\min\{K(\boldsymbol{x}),|\boldsymbol{x}_0|\}>s\}\,.$$

Hence, the map $s \to H_s(\boldsymbol{x})\mathbb{1}\left\{|\boldsymbol{x}_0| > s\right\}$ is decreasing. Therefore, $\mathrm{VC}(\mathcal{G}) = 2$.

As for $\mathcal{C}^{(1)}$ we have

$$\mathbb{1}\left\{\mathcal{C}^{(1)}(\boldsymbol{x}/s) = 0\right\} = \mathbb{1}\left\{\boldsymbol{x}^*_{-\infty,-1} \le s, |\boldsymbol{x}_0| > s\right\} .$$

Thus,

$$H(\boldsymbol{x}/s)\mathbb{1}\{\mathcal{C}(\boldsymbol{x}/s) = 0\}\mathbb{1}\left\{|\boldsymbol{x}_0| > s\right\} = \mathbb{1}\left\{\min\{K(\boldsymbol{x}), |\boldsymbol{x}_0|\} > s\right\}\mathbb{1}\left\{\boldsymbol{x}^*_{-\infty,-1} \le s\right\} .$$

Now, the class

$$\mathcal{F} = \{(-\infty, s) \times (s, +\infty) : s \in \mathbb{R}\}$$

has the VC-index 3. By [KS20, Example C.4.14] the class $\mathcal{G}$ has VC-index at most 3.

Similarly, for $\mathcal{C}^{(2)}$ we have

$$\mathbb{1}\left\{\mathcal{C}^{(2)}(\boldsymbol{x}/s) = 0\right\} = \mathbb{1}\left\{\boldsymbol{x}^*_{1,\infty} \le s, |\boldsymbol{x}_0| > s\right\} .$$

Thus,

$$H(\boldsymbol{x}/s)\mathbb{1}\{\mathcal{C}(\boldsymbol{x}/s) = 0\}\mathbb{1}\left\{|\boldsymbol{x}_0| > s\right\} = \mathbb{1}\left\{\min\{K(\boldsymbol{x}), |\boldsymbol{x}_0|\} > s\right\}\mathbb{1}\left\{\boldsymbol{x}^*_{1,\infty} \le s\right\}$$

and again the class $\mathcal{G}$ has VC-index at most 3.

In summary, for functionals H given in (5.59) and the anchoring maps $\mathcal{C}^{(0)}, \mathcal{C}^{(1)}, \mathcal{C}^{(2)}$ the class $\mathcal{G}$ has the VC-index at most 3 and hence the random entropy Assumption 3.5 is satisfied.

Now, assume that the map $s \to H_s$ is decreasing. This is the case of (again) the extremal index, the large deviation index and the ruin index. This is also the case of the stop-loss index and the cluster size distribution. If we choose $\mathcal{C} = \mathcal{C}^{(0)}$, since $\mathcal{C}^{(0)}$ is 0-homogeneous, the maps $s \to H_s^{\mathcal{C}}$ is also decreasing. Thus, the VC-index of $\mathcal{G}$ is at most 2. The random entropy condition is satisfied.

Supplementary Material

Supplementary material for "Estimation of cluster functionals for regularly varying time series: Runs estimators"
(doi: 10.1214/22-EJS2026SUPP; .pdf). Simulation studies

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

[BP18] Bojan Basrak and Hrvoje Planinić. A note on vague convergence of measures. *Statistics and Probability Letters*, 153:180–186, 2019. MR3979308

[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

[BGTS09] Jan Beirlant and Yuri Goegebeur and Jozef Teugels and Johan Segers. *Statistics of extremes* John Wiley & Sons, Ltd., Chichester, 2004. MR2108013

[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

[BZ21] Axel Bücher and Chen Zhou. A horse racing between the block maxima method and the peak–over–threshold approach. *Statistical Science*, 36(3), 360-378, 2021. MR4293095

[CK21] Youssouph Cissokho and Rafal Kulik. Estimation of cluster functionals for regularly varying time series: sliding blocks estimators. *Electronic Journal of Statistics*, 15(1): 2777–2831, 2021. MR4280158

[CK22] Youssouph Cissokho and Rafal Kulik. Supplementary material for "Estimation of cluster functionals for regularly varying time series: runs estimators". *Electronic Journal of Statistics*, 2022. DOI: https://doi.org/10.1214/22-EJS2026SUPP.

[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

[DN21] Holger Drees and Sebastian Neblung. Asymptotics for sliding blocks estimators of rare events. *Bernoulli* 27(2), 1239–1269, 2021. MR4255233

[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

[FdH15] Ana Ferreira, Laurens de Haan. On the block maxima method in extreme value theory: PWM estimators *Annals of Statistics*, 43(1): 276–298, 2015. MR3285607

[Has18] Enkelejd Hashorva. Representations of max-stable processes via exponential tilting, *Stochastic Processes and their Applications*, 128(9): 2952–2978, 2018. MR3834849

[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

[Sam2016] Gennady Samorodnitsky. *Stochastic processes and long range dependence* Springer, 2016. MR3561100

[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

[ZVB21] Nan Zou, Stanislav Volgushev, and Axel Bücher. Multiple block sizes and overlapping blocks for multivariate time series extremes. *Annals of Statistics* 49(1), 295–320, 2021. MR4206679