Bernoulli **27**(2), 2021, 1239–1269
https://doi.org/10.3150/20-BEJ1272

Asymptotics for sliding blocks estimators of rare events

HOLGER DREES* and SEBASTIAN NEBLUNG†

Department of Mathematics, SPST, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany. E-mail: *drees@math.uni-hamburg.de; †sebastian.neblung@uni-hamburg.de

Drees and Rootzén (*Ann. Statist.* **38** (2010) 2145–2186) have established limit theorems for a general class of empirical processes of statistics that are useful for the extreme value analysis of time series, but do not apply to statistics of sliding blocks, including so-called runs estimators. We generalize these results to empirical processes which cover both the class considered by Drees and Rootzén (*Ann. Statist.* **38** (2010) 2145–2186) and processes of sliding blocks statistics. Using this approach, one can analyze different types of statistics in a unified framework. We show that statistics based on sliding blocks are asymptotically normal with an asymptotic variance which, under rather mild conditions, is smaller than or equal to the asymptotic variance of the corresponding estimator based on disjoint blocks. Finally, the general theory is applied to three well-known estimators of the extremal index. It turns out that they all have the same limit distribution, a fact which has so far been overlooked in the literature.

Keywords: Asymptotic efficiency; empirical processes; extremal index; extreme value analysis; sliding vs disjoint blocks; time series; uniform central limit theorems

1. Introduction

The analysis of the serial dependence between large observations is crucial for a thorough understanding of the extreme value behavior of stationary time series. In the peaks over threshold (POT) approach, estimators of the dependence structure can usually be defined blockwise. To be more specific, assume that, starting from a stationary $\mathbb{R}^d$-valued time series $(X_t)_{1\le t\le n}$, random variables (rv's) $X_{n,i}$ are defined, that in some sense capture its extreme value behavior. The most common example is $X_{n,i} := (X_i/u_n)\mathbb{1}_{(u_n,\infty)}(\|X_i\|)$ for some threshold u_n and some norm $\|\cdot\|$ on $\mathbb{R}^d$, but for certain applications $X_{n,i}$ may also depend on observations in the neighborhood of extreme observations. We consider statistics $g(Y_{n,j})$ of blocks

$$Y_{n,j} := (X_{n,j},\dots,X_{n,j+s_n-1}) \tag{1.1}$$

of (possibly increasing) length s_n, starting with the jth rv. Estimators and test statistics of interest can then be defined in terms of averages of such blocks statistics. For example, the well-known blocks estimator of the extremal index (roughly speaking, the reciprocal of the mean size of a cluster of extreme values) is of this type; see Section 3 for details. Other examples are the empirical extremogram analyzed by Davis and Mikosch [5], forward and backward estimators of the distribution of the spectral tail process of a regularly varying time series examined by Drees et al. [11] and Davis et al. [4], and the estimator of the cluster size distribution proposed by Hsing [14].

Here one may average either statistics $g(Y_{n,is_n+1})$, $0\le i\le \lfloor n/s_n\rfloor -1$, of disjoint blocks or statistics $g(Y_{n,i})$, $1\le i\le n-s_n+1$, of overlapping sliding blocks. It has been suggested in the literature that the latter approach may often be more efficient; see, for example, Beirlant et al. [1], page 390, for such a statement about blocks estimators of the extremal index. However, the asymptotic performance of both approaches has been compared only for a couple of estimators, while general results showing the

1350-7265

superiority of the sliding blocks estimators are not yet known in the POT setting. Robert et al. [20] proved that for a different type of estimators of the extremal index the version using sliding blocks has a strictly smaller asymptotic variance than the one based on disjoint blocks, while the bias is asymptotically the same. In a block maxima setting, Zou et al. [26] proved that under quite general conditions an estimator of the extreme value copula of multivariate stationary time series is more efficient if it is based on sliding rather than disjoint blocks. The same observation has been made by Bücher and Segers [3] for the maximum likelihood estimator of the parameters of a Fréchet distribution based on maxima of sliding or disjoint blocks, respectively, of a stationary time series with marginal distribution in the maximum domain of attraction of this Fréchet distribution.

Drees and Rootzén [10] provided a general framework to analyze the asymptotic behavior of statistics which are based on averages of functionals of disjoint blocks from an absolutely regular time series. Sufficient conditions for convergence of the empirical process of so-called cluster functionals established there proved to be a powerful tool for establishing asymptotic normality of a wide range of estimators; see, for example, Drees [7], Davis et al. [4], and Drees and Knezevic [8]. Unfortunately, the setting considered by Drees and Rootzén [10] is too restrictive to accommodate empirical processes based on sliding blocks.

The first aim of the present paper is thus to establish results on the convergence of empirical processes of the type

$$\bar{Z}_n(g) := \frac{1}{\sqrt{p_n} b_n(g)} \sum_{j=1}^{n-s_n+1} \big(g(Y_{n,j}) - Eg(Y_{n,j}) \big), \quad g \in \mathcal{G}.$$

Here $Y_{n,j}$ is defined by (1.1) for some row-wise stationary triangular array $(X_{n,i})_{1 \le i \le n, n \in \mathbb{N}}$, $\mathcal{G}$ is a set of functionals defined on vectors of arbitrary length that vanish if applied to a null vector and $\sqrt{p_n} b_n(g)$ is a normalizing sequence which will be introduced in Section 2. We are mainly interested in the case when $X_{n,i}$ are suitably standardized extremes. In particular, we will assume $P\{\exists g \in \mathcal{G} : g(Y_{n,1}) \neq 0\} \to 0$. It is worth mentioning, though, that our general results can be applied to other statistics of rare events (cf. Drees and Rootzén [10], Ex. 3.5).

The second aim is to compare the performance of estimators derived from $\bar{Z}_n(g)$ with their analogs based on disjoint blocks. To this end, we will prove convergence of certain empirical processes in an abstract unifying framework which encompasses both the aforementioned setting to deal with sliding blocks processes $\bar{Z}_n$ and the setting discussed by Drees and Rootzén [10]. This way one may derive the asymptotic normality of functionals of sliding resp. disjoint blocks under similar conditions, and the expressions obtained for their asymptotic variances become comparable. It will be shown that indeed, under weak conditions, the asymptotic variance of an estimator using sliding blocks statistics is never greater than the asymptotic variance of its counterpart based on disjoint blocks.

Sometimes block based extreme value statistics are motivated by the interpretation that all large values in such a block form a cluster of extremes. In another interpretation, all large values which are not separated in time by a certain number of smaller values form a cluster. This leads to so-called runs estimators, the best-known example of which is the estimator of the extremal index, proposed by Hsing [15]. Such runs estimators can be considered as a special type of sliding blocks estimators and can thus be analyzed with the techniques developed in this paper under comparable conditions as estimators based on disjoint blocks. It turns out that both types of estimators of the extremal index have the same asymptotic variance. While the asymptotic normality of both estimators has already been proved by Weissman and Novak [25], the equality of their asymptotic variances has been overlooked, because the variances were expressed differently. In addition, we establish the asymptotic normality of the direct sliding blocks analog to the disjoint blocks estimator. Under mild conditions, this estimator has the same asymptotic variance, too. This application demonstrates that, by analyzing different estimators of the same parameter in a unifying framework, one may gain new insights.

The paper is organized as follows. In Section 2, we first establish sufficient conditions for the convergence of empirical processes of sliding blocks statistics. Table 1 provides an overview of several sequences of real and integer numbers arising in this context. In Section 2.1, the asymptotic variances of estimators using sliding and disjoint blocks, respectively, are compared. In Section 3, the general theory is applied to three estimators of the extremal index. Process convergence in the general abstract setting is presented in Appendix A, while all proofs are collected in Appendix B. Refinements to some of the results of this paper and detailed sufficient conditions for the asymptotic normality of statistics considered in Section 2.1 are presented in a Supplement [9].

Throughout the paper, $(E, \|\cdot\|)$ denotes a complete normed vector space and $E_\cup := \bigcup_{n\in\mathbb{N}} E^n$ the set of vectors of arbitrary length with E-valued components. $\mathbb{N}$ denotes the natural numbers excluding 0. For any doubly indexed sequence $Q_{n,i}$, $1 \le i \le m_n$, of random variables that are identically distributed, Q_n denotes a generic random variable with the same distribution as $Q_{n,1}$. Outer probabilities are denoted by P^*, outer expectations by E^*. Weak convergence is indicated by $\xrightarrow{w}$, while $\xrightarrow{P}$ denotes convergence in probability and $\xrightarrow{P^*}$ convergence in outer probability. The positive part of any $x \in \mathbb{R}$ is denoted by $x^+ := \max(x, 0)$.

2. Empirical processes of sliding blocks statistics

Throughout this section, we assume that $(X_{n,i})_{1\le i\le n, n\in\mathbb{N}}$ is a triangular array of row-wise stationary E-valued random variables. First, we establish conditions under which an empirical process of sliding blocks statistics of the type

$$\bar{Z}_n(g) := \frac{1}{\sqrt{p_n}\, b_n(g)} \sum_{j=1}^{n-s_n+1} \big(g(Y_{n,j}) - Eg(Y_{n,j})\big), \quad g \in \mathcal{G}, \tag{2.1}$$

converges to a Gaussian process in the space $\ell^\infty(\mathcal{G})$ of bounded functions on $\mathcal{G}$, endowed with the supremum norm. The normalizing sequence $\sqrt{p_n}\, b_n(g) \to \infty$ is discussed below.

To this end, we will apply the general abstract results presented in Appendix A to

$$V_{n,i}(g) := \frac{1}{b_n(g)} \sum_{j=1}^{r_n} g(Y_{n,(i-1)r_n+j}), \tag{2.2}$$

where r_n denotes a sequence that grows faster than s_n but slower than n. Furthermore, r_n is chosen such that it is unlikely to have any extreme value in a sequence of r_n consecutive observations. More precisely, we assume

$$p_n := P\big\{\exists g \in \mathcal{G} : V_n(g) \neq 0\big\} \to 0 \tag{2.3}$$

as $n \to \infty$, where V_n has the same distribution as any $V_{n,i}$. The set $\{\exists g \in \mathcal{G} : V_n(g) \neq 0\}$ is measurable under the following condition:

(D0) The processes V_n, $n \in \mathbb{N}$, are separable.

Condition (D0) helps to avoid measurability problems; in particular, it is fulfilled if $\mathcal{G}$ is finite. Note that $\bar{Z}_n$ can be approximated by

$$
\begin{aligned}
Z_n(g) &:= \frac{1}{\sqrt{p_n} b_n(g)} \sum_{j=1}^{m_n r_n} \big(g(Y_{n,j}) - Eg(Y_{n,j})\big) \\
&= \frac{1}{\sqrt{p_n}} \sum_{i=1}^{m_n} \big(V_{n,i}(g) - EV_{n,i}(g)\big), \quad g \in \mathcal{G},
\end{aligned} \tag{2.4}
$$

with $m_n := \lfloor (n - s_n + 1)/r_n \rfloor$. We will see below that under suitable conditions the last $n - s_n + 1 - m_n r_n < r_n$ summands in definition (2.1) of $\bar{Z}_n$ are asymptotically negligible.

We will prove process convergence using the well-known "big blocks, small blocks" technique where each $Y_{n,j}$ takes over the role of a single observation and r_n is the length of the big blocks. In addition, we need to choose the length l_n of the smaller blocks which must not be smaller than s_n, so that $Y_{n,j}$ and $Y_{n,j+l_n}$ do not overlap. Moreover, we assume that the dependence between observations separated in time by $l_n - s_n$ vanishes asymptotically. The strength of dependence will be measured by the mixing coefficients

$$
\beta_{n,k}^X := \sup_{1 \le l \le n-k-1} E\Big[\sup_{B \in \mathcal{B}_{n,l+k+1}^n} \big|P\big(B|\mathcal{B}_{n,1}^l\big) - P(B)\big|\Big], \tag{2.5}
$$

where $\mathcal{B}_{n,i}^j$ denotes the σ-field generated by $(X_{n,l})_{i \le l \le j}$. To summarize, we require the following conditions on the observational scheme, the different sequences and the function class:

- **(A1)** $(X_{n,i})_{1 \le i \le n}$ is stationary for all $n \in \mathbb{N}$.
- **(A2)** The sequences $l_n, r_n, s_n \in \mathbb{N}$, p_n defined in (2.3), and $b_n(g) > 0$, $g \in \mathcal{G}$, satisfy $s_n \le l_n = o(r_n)$, $r_n = o(n)$, $p_n \to 0$ and $r_n = o(\sqrt{p_n} \inf_{g \in \mathcal{G}} b_n(g))$.
- **(MX)** $m_n \beta_{n,l_n - s_n}^X \to 0$ for $m_n := \lfloor (n - s_n + 1)/r_n \rfloor$.

An overview of the sequences and their interpretations can be found in Table 1. Finally, to ensure the convergence of the finite dimensional marginal distributions (fidis) of $\bar{Z}_n$, we assume:

- **(C)** There exists a function $c : \mathcal{G}^2 \to \mathbb{R}$ such that

$$
\frac{m_n}{p_n} \operatorname{Cov}\big(V_n(g), V_n(h)\big) \to c(g,h), \quad \forall g, h \in \mathcal{G}.
$$

Our first result deals with the convergence of the fidis if $\mathcal{G}$ is uniformly bounded.

Theorem 2.1. *Suppose* $g_{\max} = \sup_{g \in \mathcal{G}} |g|$ *is bounded and measurable and the conditions* (A1), (A2), (D0) *and* (MX) *are met. Moreover, assume*

$$
E\Bigg[\bigg(\sum_{j=1}^{r_n} \mathbb{1}_{\{g(Y_{n,j}) \neq 0\}}\bigg)^2\Bigg] = O\bigg(\frac{p_n b_n^2(g)}{m_n}\bigg), \quad \forall g \in \mathcal{G}. \tag{2.6}
$$

Then

$$
\sup_{g \in \mathcal{G}} \big|Z_n(g) - \bar{Z}_n(g)\big| \xrightarrow{P^*} 0. \tag{2.7}
$$

Table 1. Overview of sequences occurring in Sections 2 and 2.1

	Interpretation	$\to$	Main constraints	Typ. behavior	First use
n	number of observations	∞			page 1239
s_n	length of sliding blocks			$s_n \to \infty$	(1.1)
r_n	length of big block	∞	in Section 2: $r_n = \mathrm{o}(n)$, $r_n = \mathrm{o}(\sqrt{p_n}\inf_{g\in\mathcal{G}} b_n(g))$ in Section 3: $r_n v_n \to 0$, $r_n = \mathrm{o}(\sqrt{n v_n})$		(2.2)
l_n	length of small block	∞	$s_n \le l_n = \mathrm{o}(r_n)$		(A2)
m_n	number of big blocks	∞	$m_n \asymp n/r_n$		(2.4)
u_n	threshold for X to be large	∞			page 1239
v_n	$P\{X_{n,1} \ne 0\}$	0	$n v_n \to \infty$		page 1245
p_n	$P\{\exists 1 \le i \le r_n : X_{n,i} \ne 0\}$	0	$r_n = \mathrm{o}(\sqrt{p_n}\inf_{g\in\mathcal{G}} b_n(g))$	$p_n \asymp r_n v_n$	(2.3)
$b_n(g)$	normalizing constant	∞	$\sqrt{p_n} b_n(g) \to \infty$	$b_n(g) \asymp \sqrt{m_n}$ or $b_n(g) \asymp \sqrt{m_n} s_n$	(2.1)
a_n	normalization in Section 2.1			$a_n \asymp 1$	page 1245

If, in addition, (C) *is fulfilled, then the fidis of each of the empirical processes* $(Z_n(g))_{g\in\mathcal{G}}$ *and* $(\bar{Z}_n(g))_{g\in\mathcal{G}}$ *converge weakly to the fidis of a Gaussian process* $(Z(g))_{g\in\mathcal{G}}$ *with covariance function* c.

The following criterion is often useful to verify condition (2.6):

(S) For all $g \in \mathcal{G}$ and $n \in \mathbb{N}$

$$\sum_{k=1}^{r_n} P\{g(Y_{n,1}) \ne 0, g(Y_{n,k}) \ne 0\} = \mathrm{O}\left(\frac{p_n b_n^2(g)}{n}\right).$$

Lemma 2.2. *If condition* (S) *is satisfied, then* (2.6) *holds.*

For instance, in Section 3 we consider the bounded functions

$$g_1(x_1, \dots, x_s) := \mathbb{1}_{\{\max_{1\le i\le s} x_i > 1\}}, \qquad g_2(x_1, \dots, x_s) := \mathbb{1}_{\{x_1 > 1\}}$$

to analyze the sliding blocks estimator of the extremal index. Here appropriate normalizing sequences are $b_n(g_1) = \sqrt{m_n} s_n$ and $b_n(g_2) = \sqrt{m_n}$. Note that already in this rather simple example, the normalizing sequences converge at a different rate for different functions. Indeed, it is somewhat archetypical that the event $g(Y_{n,1}) \ne 0$ either depends on all observations of the block $Y_{n,1}$ (as for $g = g_1$), or it only depends on a single fixed observation $X_{n,i}$ (as for $g = g_2$); usually the normalizing factor $b_n(g)$ is larger by the factor s_n in the former case.

To ensure asymptotic equicontinuity or tightness of the processes $(Z_n(g))_{g\in\mathcal{G}}$ and $(\bar{Z}_n(g))_{g\in\mathcal{G}}$, and thus process convergence if the conditions of Theorem 2.1 are fulfilled, we need the following additional conditions.

(D1) There exists a semi-metric ρ on $\mathcal{G}$ such that $\mathcal{G}$ is totally bounded (i.e., for all $\epsilon > 0$, it can be covered by finitely many balls with radius ϵ w.r.t. ρ) and

$$\lim_{\delta\downarrow 0} \limsup_{n\to\infty} \sup_{g,h\in\mathcal{G},\rho(g,h)<\delta} \frac{m_n}{p_n} E\big[\big(V_n(g) - V_n(h)\big)^2\big] = 0.$$

(D2)

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\int_0^\delta \sqrt{\log N_{[\cdot]}\big(\epsilon,\mathcal{G},L_2^n\big)}\,d\epsilon=0,$$

where $N_{[\cdot]}(\epsilon,\mathcal{G},L_2^n)$ denotes the ϵ-bracketing number of $\mathcal{G}$ w.r.t. L_2^n, that is, the smallest number N_ϵ such that for each $n\in\mathbb{N}$ there exists a partition $(\mathcal{G}_{n,k}^\epsilon)_{1\le k\le N_\epsilon}$ of $\mathcal{G}$ satisfying

$$\frac{m_n}{p_n}E^*\Big[\sup_{g,h\in\mathcal{G}_{n,k}^\epsilon}\big(V_n(g)-V_n(h)\big)^2\Big]\le\epsilon^2,\quad \forall 1\le k\le N_\epsilon.$$

(D3) Denote by $N(\epsilon,\mathcal{G},d_n)$ the ϵ-covering number of $\mathcal{G}$ w.r.t. the random semi-metric

$$d_n(g,h)=\left(\frac{1}{p_n}\sum_{i=1}^{m_n}\big(V_{n,i}^*(g)-V_{n,i}^*(h)\big)^2\right)^{1/2}$$

with $V_{n,i}^*$, $1\le i\le m_n$, independent copies of $V_{n,1}$, i.e. $N(\epsilon,\mathcal{G},d_n)$ is the smallest number of balls with respect to d_n with radius ϵ which is needed to cover $\mathcal{G}$. We assume

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}P^*\left\{\int_0^\delta\sqrt{\log\big(N(\epsilon,\mathcal{G},d_n)\big)}\,d\epsilon>\tau\right\}=0,\quad \forall\tau>0.$$

Roughly speaking, condition (D1) ensures the continuity of the process w.r.t. ρ while (D2) and (D3) ensure that the parameter set $\mathcal{G}$ is not too complex. In particular, condition (D3) is satisfied if $\mathcal{G}$ is a VC-class (cf. Drees and Rootzén [10], Remark 2.11).

Theorem 2.3. *Suppose the conditions of Theorem 2.1 are satisfied. If, in addition, one of the following sets of conditions*

(i) (D1) *and* (D2), *or*
(ii) (D1) *and* (D3)

is fulfilled, then each of the empirical processes $(Z_n(g))_{g\in\mathcal{G}}$ *and* $(\bar{Z}_n(g))_{g\in\mathcal{G}}$ *converge weakly to a Gaussian process with covariance function* c.

So far, we have only discussed the case of bounded functions g. This assumption can be dropped if the moment condition (2.6) is strengthened.

Theorem 2.4.

(i) *Suppose all conditions of Theorem 2.1 except for the boundedness of* $g_{\max}$ *and* (2.6) *are met. In addition, we assume* $m_n l_n P\{V_n(|g|)\neq 0\}=\mathrm{o}(r_n b_n^2(g)p_n)$ *for all* $g\in\mathcal{G}$ *and*

$$E\left[\left(\sum_{i=1}^{r_n}\big|g(Y_{n,i})\big|\right)^{2+\delta}\right]=\mathrm{O}\left(\frac{p_n b_n^2(g)}{m_n}\right),\quad \forall g\in\mathcal{G},\tag{2.8}$$

for some $\delta>0$. *Then the fidis of* $(Z_n(g))_{g\in\mathcal{G}}$ *and of* $(\bar{Z}_n(g))_{g\in\mathcal{G}}$ *converge to the fidis of the Gaussian process* $(Z(g))_{g\in\mathcal{G}}$ *defined in Theorem 2.1.*

(ii) *If, in addition,* $b_n(g)=b_n$ *is the same for all* $g\in\mathcal{G}$, (2.8) *holds for* $g=g_{\max}$ *and the conditions* (i) *or* (ii) *of Theorem 2.3 are fulfilled, then the processes* $(Z_n(g))_{g\in\mathcal{G}}$ *and* $(\bar{Z}_n(g))_{g\in\mathcal{G}}$ *converge weakly to* $(Z(g))_{g\in\mathcal{G}}$ *uniformly.*

Note that usually $P\{V_n(|g|) \neq 0\} = O(p_n)$; in particular this holds true if g has fixed sign. Then the condition $m_n l_n P\{V_n(|g|) \neq 0\} = \mathrm{o}(r_n b_n^2(g) p_n)$ is fulfilled for the typical behavior of the sequences outlined in Table 1. As mentioned above, usually it suffices to consider just two different normalizing sequences, say $b_{n,1}$ and $b_{n,2}$. In this case, one may apply Theorem 2.4 separately to $(Z_n(g))_{g\in\mathcal{G}_i}$ for $i \in \{1,2\}$ with $\mathcal{G}_i := \{g \in \mathcal{G} | b_n(g) = b_{n,i}, \forall n \in \mathbb{N}\}$ to conclude that both processes are asymptotically tight. This in turn implies the asymptotic tightness of $(Z_n(g))_{g\in\mathcal{G}}$ and thus, in view of part (i), its convergence to $(Z(g))_{g\in\mathcal{G}}$. Hence, in fact the extra condition on b_n in part (ii) does not further restrict the setting in the vast majority of applications.

2.1. Sliding vs. disjoint blocks statistics

The previous section was devoted to general limit theorems for sliding blocks statistics. In this section, we want to compare the asymptotic variance of a sliding blocks statistic for a single functional g with that of the corresponding disjoint blocks statistic. Here we use a different parametrization of the normalizing constants, partly because the probability p_n used in the normalization above refers to the whole process and seems inappropriate in the present context, partly to facilitate the comparison of the asymptotic variances. More precisely, we consider the sliding blocks statistic and its disjoint blocks analog

$$T_n^s(g) := \frac{1}{n v_n s_n a_n} \sum_{i=1}^{n-s_n+1} g(Y_{n,i}),$$

$$T_n^d(g) := \frac{1}{n v_n a_n} \sum_{i=1}^{\lfloor n/s_n \rfloor} g(Y_{n,(i-1)s_n+1}),$$

with $v_n := P(X_{n,1} \neq 0) \to 0$. We assume that a_n is chosen such that $E(T_n^s(g))$ converges in $\mathbb{R}$, i.e. that there exists some $\xi \in \mathbb{R}$ such that

$$E\big[T_n^s(g)\big] = \frac{1}{s_n v_n a_n} E\big[g(Y_n)\big] \frac{n - s_n + 1}{n} \to \xi. \tag{2.9}$$

Then also $E(T_n^d(g))$ tends to ξ. Moreover, the difference between both expectations is asymptotically negligible if

$$\big|E\big[T_n^d(g) - T_n^s(g)\big]\big| = \frac{1}{s_n v_n a_n} \big|E\big[g(Y_n)\big]\big| \cdot \left| \frac{s_n}{n} \left\lfloor \frac{n}{s_n} \right\rfloor - \frac{n - s_n + 1}{n} \right| = \mathrm{O}(s_n/n)$$

is of smaller order than $(n v_n)^{-1/2}$ (cf. (2.11), (2.12)), which in particular holds under the basic condition $s_n v_n \to 0$. In that case, $T_n^s(g)$ will be a more efficient estimator than $T_n^d(g)$ if its asymptotic variance is smaller.

Applying Theorem 2.1 with $b_n(g) = \sqrt{n v_n / p_n} a_n s_n$, under suitable conditions including the convergence

$$c^{(s)} = \lim_{n\to\infty} \frac{1}{r_n v_n s_n^2 a_n^2} \operatorname{Var}\left(\sum_{i=1}^{r_n} g(Y_{n,i}) \right) \in (0,\infty), \tag{2.10}$$

one can prove the asymptotic normality of the sliding blocks statistics

$$\sqrt{n v_n}\big(T_n^s(g) - E\big[T_n^s(g)\big]\big) \xrightarrow{w} \mathcal{N}\big(0, c^{(s)}\big). \tag{2.11}$$

To establish an analogous result for the statistic based on disjoint blocks, one applies Theorem A.1 to $V_{n,i}(g) = \sqrt{p_n/(nv_n a_n^2)} \sum_{j=1}^{r_n/s_n} g(Y_{n,(j-1)s_n+(i-1)r_n+1})$, $1 \le i \le m_n$. Recall that the sequence r_n is only needed in the proofs which use the "big blocks, small blocks" technique, that is, it has no operational meaning, but it must be chosen such that the conditions of Theorem 2.1 resp. Theorem A.1 are met. For example, suppose that for a given sequence $(s_n)_{n\in\mathbb{N}}$, $(r_n)_{n\in\mathbb{N}}$ is a sequence such that the assumptions of Theorem 2.1 are satisfied. Let $r_n^* := \lfloor r_n/s_n \rfloor s_n \sim r_n$, so that $l_n = \mathrm{o}(r_n^*)$, $r_n^* = \mathrm{o}(n)$ and $m_n^* := \lfloor (n - s_n + 1)/r_n^* \rfloor \sim m_n$. Moreover, the proof of Theorem 2.1 (cf. (B.4)) shows that for

$$V_{n,1}^*(g) := \frac{1}{b_n(g)} \sum_{j=1}^{r_n^*} g(Y_{n,j}), \quad \text{and} \quad p_n^* := P\{\exists g \in \mathcal{G} : V_{n,1}^*(g) \neq 0\},$$

one has

$$E\big[\big(V_{n,1}^*(g) - V_{n,1}(g)\big)^2\big] = E\left[\left(\frac{1}{b_n(g)} \sum_{j=r_n^*+1}^{r_n} g(Y_{n,j})\right)^2\right] = \mathrm{o}\left(\frac{p_n}{m_n}\right),$$

$$\big|p_n^* - p_n\big| \le s_n v_n.$$

Hence, if $p_n \asymp r_n v_n$ (which holds true for all known examples), $p_n^* \sim p_n$ and the conditions of Theorem 2.1 are still fulfilled if one replaces r_n with r_n^*. One may argue similarly in the setting of Theorem 2.4.

We may thus assume w.l.o.g. that r_n is a multiple of s_n, where the multiplicity depends on n. Note that r_n/s_n must tend to ∞ if Theorem 2.1 shall be applied. We then obtain

$$\sqrt{nv_n}\big(T_n^d(g) - E\big[T_n^d(g)\big]\big) \xrightarrow{w} \mathcal{N}\big(0, c^{(d)}\big), \tag{2.12}$$

with

$$c^{(d)} = \lim_{n\to\infty} \frac{1}{r_n v_n a_n^2} \operatorname{Var}\left(\sum_{i=1}^{r_n/s_n} g(Y_{n,is_n+1})\right). \tag{2.13}$$

See the Supplement [9] for details about the conditions under which (2.11) and (2.12) hold. Alternatively, one could prove the asymptotic normality of $T_n^d(g)$ using Theorem 2.3 of Drees and Rootzén [10] with r_n replaced by s_n, but the above representation of the asymptotic variance $c^{(d)}$ simplifies the comparison with $c^{(s)}$. The following theorem shows that the asymptotic variance of the sliding blocks statistic is never greater than that of the disjoint blocks statistic.

Theorem 2.5. *If conditions* (A1), (2.10) *and* (2.13) *hold, and $r_n/s_n \in \mathbb{N}$ for all $n \in \mathbb{N}$, then $c^{(s)} \le c^{(d)}$.*

Indeed, one can even prove a multivariate version of this theorem: under suitable conditions the asymptotic covariance matrix of a vector of sliding blocks statistics $(T_n^s(g_i))_{1\le i\le I}$ is smaller w.r.t. the Loewner order than the corresponding matrix of the disjoint blocks statistics (see Supplement [9]).

Usually, the probability v_n that a single observation $X_{n,1}$ does not vanish is unknown, whereas the normalizing constant a_n may depend on g, but not on the unknown distribution of $X_{n,1}$. In what follows, we thus analyze versions of our statistics where v_n is replaced with a simple empirical estimator. This results in the estimators

$$\tilde{T}_n^s(g) := \frac{nv_n T_n^s(g)}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_{n,i}\neq 0\}}} = \frac{\frac{1}{s_n a_n} \sum_{i=1}^{n-s_n+1} g(Y_{n,i})}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_{n,i}\neq 0\}}},$$

$$\tilde{T}_n^d(g) := \frac{n v_n T_n^d(g)}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_{n,i}\neq 0\}}} = \frac{\frac{1}{a_n}\sum_{i=1}^{\lfloor n/s_n \rfloor} g(Y_{n,(i-1)s_n+1})}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_{n,i}\neq 0\}}}$$

of ξ. In order to prove convergence of these estimators, one needs the joint convergence of the numerator and denominator. This can again be concluded from Theorem 2.1 or Theorem A.1, respectively, now applied with $\mathcal{G} = \{g, h\}$ and $h(x_1, \dots, x_s) = \mathbb{1}_{\{x_1\neq 0\}}$. Similarly as before, one obtains

$$\sqrt{n v_n}\begin{pmatrix} T_n^\sharp(g) - E\big[T_n^\sharp(g)\big] \\ \frac{1}{n v_n}\sum_{i=1}^{n-s_n+1}(\mathbb{1}_{\{X_{n,i}\neq 0\}} - v_n)\end{pmatrix} \xrightarrow{w} \mathcal{N}_2\left(0, \begin{pmatrix} c^{(\sharp)} & c^{(\sharp,v)} \\ c^{(\sharp,v)} & c^{(v)}\end{pmatrix}\right),$$

where $\sharp$ stands either for d or s and

$$c^{(s,v)} := \lim_{n\to\infty} \frac{1}{r_n v_n s_n a_n}\operatorname{Cov}\left(\sum_{i=1}^{r_n} g(Y_{n,i}), \sum_{i=1}^{r_n}\mathbb{1}_{\{X_{n,i}\neq 0\}}\right),$$

$$c^{(d,v)} := \lim_{n\to\infty} \frac{1}{r_n v_n a_n}\operatorname{Cov}\left(\sum_{j=1}^{r_n/s_n} g(Y_{n,(j-1)s_n+1}), \sum_{i=1}^{r_n}\mathbb{1}_{\{X_{n,i}\neq 0\}}\right),$$

$$c^{(v)} := \lim_{n\to\infty} \frac{1}{r_n v_n} E\left[\left(\sum_{i=1}^{r_n}\mathbb{1}_{\{X_{n,i}\neq 0\}}\right)^2\right].$$

Note that the same result holds if $\sum_{i=1}^{n-s_n+1}$ is replaced with $\sum_{i=1}^{n}$ (cf. (2.7)).

By some standard continuous mapping argument (see Supplement [9]), one may conclude

$$\sqrt{n v_n}\big(\tilde{T}_n^\sharp - \xi\big) \xrightarrow{w} \mathcal{N}\big(0, \tilde{c}^{(\sharp)}\big)$$

with $\tilde{c}^{(\sharp)} := c^{(\sharp)} + \xi^2 c^{(v)} - 2\xi c^{(\sharp,v)}$, provided the bias of the estimator is negligible, that is $E[g(Y_n)]/s_n v_n a_n - \xi = \mathrm{o}((n v_n)^{-1/2})$.

It turns out that under rather mild conditions again the asymptotic variance of the estimator using sliding blocks is not greater than that of the disjoint blocks estimator, if the function g has constant sign.

Theorem 2.6. *Suppose the conditions of Theorem 2.5 are satisfied, (2.9) holds, the function g is bounded and does not change its sign, $s_n = \mathrm{o}(r_n a_n)$ and $s_n v_n \to 0$. If, in addition, there exists a sequence $k_n = \mathrm{o}(r_n a_n)$ of natural numbers such that the β-mixing coefficients defined in (2.5) satisfy $\sum_{i=k_n}^{r_n} \beta_{n,i}^X = \mathrm{o}(r_n v_n a_n)$, then $\tilde{c}^{(s)} \leq \tilde{c}^{(d)}$.*

In fact, it can be shown that $\tilde{c}^{(d)} - \tilde{c}^{(s)} = c^{(d)} - c^{(s)}$. In the most common case that the mixing coefficients decrease exponentially fast and $\log n = \mathrm{o}(r_n a_n)$, the sequence $k_n = \lfloor c \log n \rfloor$ with sufficiently large constant $c > 0$ fulfills the conditions of Theorem 2.6.

3. Estimating the extremal index

In this section, we apply the general theory presented in Section 2 and Appendix A to analyze the asymptotic behavior of three estimators for the extremal index of a real-valued stationary time series

$(X_t)_{t\in\mathbb{Z}}$. If for all thresholds $u_n(\tau)$ such that $nP\{X_0 > u_n(\tau)\} \to \tau$ for some $\tau > 0$ one has

$$\lim_{n\to\infty} P\Big\{\max_{1\le i\le n} X_i \le u_n(\tau)\Big\} = e^{-\theta\tau},$$

then θ is said to be the *extremal index* of the time series (Leadbetter [16]). The extremal index always lies in [0, 1]. In what follows, we exclude the degenerate case $\theta = 0$ and assume $\theta > 0$.

The estimation of this extremal index has been much discussed in the literature, see, for example, Smith and Weissman [22], Ferro and Segers [13], Süveges [23], Robert et al. [20], Berghaus et al. [2], among others. We examine two of the most popular estimators, the blocks and the runs estimator, and a variant of the former. Throughout this section, we use the notation $M_{i,j} := \max(X_i, \dots, X_j)$.

If the extremal index exists then, under weak additional conditions,

$$\frac{P\{M_{1,k_n} > u_n\}}{k_n P\{X_1 > u_n\}} \to \theta \tag{3.1}$$

for sequences $k_n \to \infty$ and u_n such that $k_n P\{X_1 > u_n\} \to 0$. In particular, this holds if $\beta^X_{n,l_n}/(k_n v_n) \to 0$ for some $l_n = \mathrm{o}(k_n)$ (cf. Leadbetter [16], Theorem 3.4). If one replaces the unknown probabilities by empirical ones, using disjoint blocks to estimate the numerator for $k_n = s_n$, one arrives at the following estimator proposed by Hsing [14]:

$$\hat{\theta}^d_n := \frac{\sum_{i=1}^{\lfloor n/s_n\rfloor} \mathbb{1}_{\{M_{(i-1)s_n+1,is_n} > u_n\}}}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_i > u_n\}}}.$$

He proved asymptotic normality of this blocks estimator under some tailor-made conditions. As suggested in Section 10.3.4 of Beirlant et al. [1], alternatively one may use sliding blocks, which leads to

$$\hat{\theta}^s_n := \frac{\frac{1}{s_n}\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{M_{i,i+s_n-1} > u_n\}}}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_i > u_n\}}}.$$

The so-called runs estimator of θ is based on the following characterization of the extremal index:

$$P(M_{2,k_n} \le u_n \mid X_1 > u_n) \to \theta, \tag{3.2}$$

which was first proven by O'Brien [17] under suitable conditions. Again, by replacing the unknown probabilities for $k_n = s_n$ by empirical counterparts, one arrives at

$$\hat{\theta}^r_n := \frac{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_i > u_n, M_{i+1,i+s_n-1} \le u_n\}}}{\sum_{i=1}^{n-s_n+1} \mathbb{1}_{\{X_i > u_n\}}}.$$

This runs estimator was suggested by Hsing [15]. Its asymptotic normality was first established in Weissman and Novak [25] who also proved the asymptotic normality of $\hat{\theta}^d_n$ under somewhat simpler conditions than Hsing [14]. For a very specific model, Weissman and Novak [25] showed that the asymptotic variances of both estimators are the same, but they did not realize that this is indeed true under quite general structural assumptions, as we will show below.

To establish asymptotic normality of these estimators, we need the following conditions:

($\boldsymbol{\theta}$1) For $v_n := P\{X_1 > u_n\} \to 0$, one has $nv_n \to \infty$ and $s_n \to \infty$. In addition, there exists a sequence $(r_n)_{n\in\mathbb{N}}$ such that $s_n = \mathrm{o}(r_n)$, $r_n v_n \to 0$, $r_n = \mathrm{o}(\sqrt{nv_n})$ and $(n/r_n)\beta^X_{n,s_n-1} \to 0$.

($\boldsymbol{\theta}$2) $c := \lim_{n\to\infty} \frac{1}{r_n v_n} E[(\sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}})^2]$ exists in $[0, \infty)$.
($\boldsymbol{\theta}$P) For all $n \in \mathbb{N}$ and $k \in \mathbb{N}$ there exists $e_n(k)$ such that

$$e_n(k) \geq P(X_k > u_n | X_0 > u_n)$$

and $\lim_{n\to\infty} \sum_{k=1}^{r_n} e_n(k) = \sum_{k=1}^{\infty} \lim_{n\to\infty} e_n(k) < \infty$.

By Pratt's lemma (Pratt [19]), condition (θP) enables us to exchange sums and limits in the calculation of variance and covariance. Moreover, under (θ1) and (θP), both (3.1) and (3.2) hold for all $k_n \leq r_n$ such that $k_n \to \infty$. This follows from Theorem 1 and Corollary 2 of Segers [21] in combination with the aforementioned result on convergence (3.1).

The limit c is the asymptotic variance of the estimator for $v_n = P\{X_i > u_n\}$. If (θP) holds and the positive part $(X_t^+)_{t\in\mathbb{Z}}$ of the time series is regular varying, then c can be represented in terms of its tail process $(W_t)_{t\in\mathbb{Z}}$ (see Supplement [9]), that is, (θ2) holds with

$$\begin{aligned} c &= 1 + \lim_{n\to\infty} \sum_{k=1}^{r_n - 1} \left(1 - \frac{k}{r_n}\right)\big(P(X_k > u_n | X_0 > u_n) + P(X_0 > u_n | X_{-k} > u_n)\big) \\ &= 1 + 2\sum_{k=1}^{\infty} P\{W_k > 1\}. \end{aligned}$$

Alternatively, one may use the representation $c = \sum_{k\in\mathbb{Z}} P\{W_k > 1\}$.

In addition, we have to assume that convergence (3.1) for $k_n = s_n$ and convergence (3.2), respectively, is sufficiently fast to ensure that the bias of the block based estimators or runs estimators, respectively, is asymptotically negligible:

($\mathbf{B}_b$) $\frac{P\{M_{1,s_n} > u_n\}}{s_n v_n} - \theta = \mathrm{o}((n v_n)^{-1/2})$.
($\mathbf{B}_r$) $P(M_{2,s_n} \leq u_n | X_1 > u_n) - \theta = \mathrm{o}((n v_n)^{-1/2})$.

The following result shows that under our conditions all three estimator have the same limit distribution.

Theorem 3.1. *If the conditions* (θ1), (θ2) *and* (θP) *are satisfied*, *then*

$$\sqrt{n v_n}\big(\hat{\theta}_n^\sharp - \theta\big) \xrightarrow{w} \mathcal{N}\big(0, \theta(\theta c - 1)\big),$$

provided (B_b) *holds when* $\sharp$ *stands for 'd' or 's'*, *and* (B_r) *holds when* $\sharp$ *stands for 'r'*.

In practice, usually the threshold u_n is replaced with some data driven choice $\hat{u}_n$, like an intermediate order statistic of the observed time series. By the techniques developed in Drees and Knezevic [8], one may prove that these versions of the estimators of the extremal index asymptotically behave the same, provided $\hat{u}_n / u_n \xrightarrow{P} 1$ and the time series $(X_t^+)_{t\in\mathbb{Z}}$ is regular varying. To this end, the results about the convergence of the fidis are not sufficient any more, but the full process convergence is needed. The precise results and their proofs are given in the Supplement [9].

Appendix A: Functional limit theorems in an abstract setting

In this section, we prove abstract limit theorems for empirical processes which imply both the limit theorems 2.1, 2.3 and 2.4 for statistics of sliding blocks and the limit theorems established by Drees and

Rootzén [10]. As in Section 2, we consider a triangular array $(X_{n,i})_{1\le i\le n,n\in\mathbb{N}}$ of row-wise stationary E-valued random variables. Fix sequences $r_n = \mathrm{o}(n)$ and $s_n = \mathrm{o}(r_n)$ of natural numbers. In what follows, $V_{n,i}(g)$ are real-valued random variables that are measurable w.r.t. $(X_{n,(i-1)r_n+1}, \dots, X_{n,ir_n+s_n-1})$, for all $1 \le i \le m_n$ and $g \in \mathcal{G}$, which are assumed to form a stationary sequence of processes. We are interested in the weak convergence of

$$Z_n(g) := \frac{1}{\sqrt{p_n}} \sum_{i=1}^{m_n} \bigl(V_{n,i}(g) - E\bigl[V_{n,i}(g)\bigr]\bigr), \quad g \in \mathcal{G},$$

where $m_n := \lfloor (n - s_n + 1)/r_n \rfloor$ and $p_n := P\{\exists g \in \mathcal{G} : V_n(g) \neq 0\} \to 0$ is assumed.

The choice $V_{n,i}(g) = m_n^{-1/2} \sum_{j=1}^{r_n} g(X_{n,(i-1)r_n+j})$ leads to the generalized tail array sums examined in Section 3 of Drees and Rootzén [10]. Sums of more general statistics of disjoint blocks can be analyzed using $V_{n,i}(g) = \sum_{j=0}^{r_n/s_n-1} g(Y_{n,(i-1)r_n+js_n+1})/d_n(g)$ for suitable normalizing sequences $d_n(g)$ (assuming that r_n is a multiple of s_n), while the choice (2.2) yields sums of statistics of sliding blocks.

In an abstract version of the "big blocks, small blocks" approach, we approximate $V_{n,i}$ by stationary sequences of random processes $\tilde{V}_{n,i}$ that are asymptotically independent. For example, $V_{n,i}(g) = m_n^{-1/2} \sum_{j=1}^{r_n} g(X_{n,(i-1)r_n+j})$ can be approximated by $\tilde{V}_{n,i}(g) = m_n^{-1/2} \sum_{j=1}^{r_n-l_n} g(X_{n,(i-1)r_n+j})$ for a suitable sequence $l_n = \mathrm{o}(r_n)$.

We now list the conditions used to establish convergence of the finite dimensional marginal distributions (fidis) of Z_n.

- **(A)** $(X_{n,i})_{1\le i\le n}$ is stationary for all $n \in \mathbb{N}$ and the sequences $s_n, r_n \in \mathbb{N}$ satisfy $s_n = \mathrm{o}(r_n)$ and $r_n = \mathrm{o}(n)$.
- **(V)** For all $n \in \mathbb{N}$, $1 \le i \le m_n = \lfloor (n - s_n + 1)/r_n \rfloor$, $V_{n,i}$ and $\tilde{V}_{n,i}$ are real valued processes indexed by $\mathcal{G}$ that are measurable w.r.t. $(X_{n,(i-1)r_n+1}, \dots, X_{n,ir_n+s_n-1})$, and $(V_{n,i}, \tilde{V}_{n,i})_{1\le i\le m_n}$ is stationary.
- **(M$\tilde{\mathbf{V}}$)** $m_n \beta^{\tilde{V}}_{n,0} \to 0$.
- **(MX$_k$)** $m_n \beta^X_{n,(k-1)r_n-s_n} \to 0$.
- **($\boldsymbol{\Delta}$)** $\Delta_n := V_n - \tilde{V}_n$ satisfies
  - (i) $E[(\Delta_n(g) - E[\Delta_n(g)])^2 \mathbb{1}_{\{|\Delta_n(g)-E[\Delta_n(g)]|\le\sqrt{p_n}\}}] = \mathrm{o}(p_n/m_n)$, $\forall g \in \mathcal{G}$,
  - (ii) $P\{|\Delta_n(g) - E[\Delta_n(g)]| > \sqrt{p_n}\} = \mathrm{o}(1/m_n)$, $\forall g \in \mathcal{G}$,
  - (iii) For some $\tau > 0$
    $E[(\Delta_n(g) - E[\Delta_n(g)]) \mathbb{1}_{\{|\Delta_n(g)-E[\Delta_n(g)]|\le\tau\sqrt{p_n}\}}] = \mathrm{o}(\sqrt{p_n}/m_n)$, $\forall g \in \mathcal{G}$.
- **(L)** $E[(V_n(g) - E[V_n(g)])^2 \mathbb{1}_{\{|V_n(g)-E[V_n(g)]|>\epsilon\sqrt{p_n}\}}] = \mathrm{o}(p_n/m_n)$, $\forall g \in \mathcal{G}, \epsilon > 0$.

In addition, Conditions (C) and (D0) stated in Section 2 are needed. Condition (Δ) ensures that the approximation of $V_{n,i}$ by $\tilde{V}_{n,i}$ is sufficiently accurate. It is always fulfilled if

$$E\bigl[\bigl(\Delta_n(g)\bigr)^2\bigr] = \mathrm{o}(p_n/m_n), \quad \forall g \in \mathcal{G}. \tag{A.1}$$

The mixing conditions (MX$_k$) and (M$\tilde{\mathrm{V}}$) enable us to replace the summands by independent copies, while (C) and the Lindeberg condition (L) imply convergence of the sum of independent copies of $V_n(g)$.

Theorem A.1. *Suppose the conditions* (A), (V), (M$\tilde{\mathrm{V}}$), (MX$_k$) *for some* $k \in \mathbb{N}$, $k \ge 2$, (Δ), (L), (D0) *and* (C) *are satisfied. Then the fidis of the empirical process* $(Z_n(g))_{g\in\mathcal{G}}$ *converge weakly to the fidis of a Gaussian process with covariance function* c.

To conclude convergence of the processes $(Z_n(g))_{g\in\mathcal{G}}$, we have to show that they are asymptotically tight or asymptotically equicontinuous. To this end, we need (D1)–(D3) from Section 2 and the following conditions:

(B) $E[|V_n(g)|^2] < \infty$ for all $g \in \mathcal{G}$, and $V_n(\mathcal{G}) := \sup_{g\in\mathcal{G}} |V_n(g)| < \infty$.
(L1) $E^*[V_n(\mathcal{G})\mathbb{1}_{\{V_n(\mathcal{G})>\epsilon\sqrt{p_n}\}}] = \mathrm{o}(\sqrt{p_n}/m_n)$, $\forall\epsilon > 0$.

Condition (L1) follows from the following condition of Lindeberg type, that also implies (L) (see Supplement [9]):

(L2) $E^*[(V_n(\mathcal{G}))^2\mathbb{1}_{\{V_n(\mathcal{G})>\epsilon\sqrt{p_n}\}}] = \mathrm{o}(p_n/m_n)$, $\forall\epsilon > 0$.

Theorem A.2.

(i) *If the conditions* (A), (V), (MX$_k$) *for some* $k \in \mathbb{N}$, $k \geq 2$, (B), (L1), (D0), (D1) *and* (D2) *are satisfied, then the processes* $(Z_n(g))_{g\in\mathcal{G}}$ *are asymptotically tight.*
(ii) *If the conditions* (A), (V), (MX$_k$) *for some* $k \in \mathbb{N}$, $k \geq 2$, (B), (L2), (D0), (D1) *and* (D3) *are satisfied, then the processes* $(Z_n(g))_{g\in\mathcal{G}}$ *are asymptotically equicontinuous.*

Hence, the processes converge to a Gaussian process with covariance function c if, in addition, the assumptions of Theorem A.1 are fulfilled.

Appendix B: Proofs

B.1. Proofs of Appendix A

We first show that for the proof of convergence of the fidis it suffices to consider independent copies of $V_{n,i}$.

Lemma B.1. *Suppose the conditions* (A), (Δ), (M$\tilde{\mathrm{V}}$) *and* (MX$_k$) *for some* $k \in \mathbb{N}$, $k \geq 2$, *are satisfied. Let*

$$Z_n^*(g) := \frac{1}{\sqrt{p_n}}\sum_{i=1}^{m_n}\big(V_{n,i}^*(g) - E\big[V_{n,i}^*(g)\big]\big), \quad g \in \mathcal{G},$$

where $V_{n,i}^*$ *are independent copies of* $V_{n,i}$, $1 \leq i \leq m_n$. *Then the fidis of* $(Z_n(g))_{g\in\mathcal{G}}$ *converge weakly if and only if the fidis of* $(Z_n^*(g))_{g\in\mathcal{G}}$ *converge, and if so, the limits coincide.*

Proof of Lemma B.1. Let $\Delta_{n,i} := V_{n,i} - \tilde{V}_{n,i}$ and $\Delta_{n,i}^*$ be independent copies of $\Delta_{n,i}$, $1 \leq i \leq m_n$. For the k for which (MX$_k$) is satisfied and for all $i \in \{1,\dots,k\}$, condition (Δ) and Theorem 1 of Section IX.1 of Petrov [18] yield

$$\frac{1}{\sqrt{p_n}}\sum_{j=1}^{m_{n,k,i}}\big(\Delta_{n,jk-i}^*(g) - E\big[\Delta_{n,jk-i}^*(g)\big]\big) = \mathrm{o}_P(1), \quad \forall g \in \mathcal{G}, \tag{B.1}$$

where $m_{n,k,i} := \lfloor (m_n + i)/k\rfloor \leq m_n$.

Recall that $\Delta_{n,jk-i}$ is measurable w.r.t. $(X_{n,(jk-i-1)r_n+1},\dots,X_{n,(jk-i)r_n+s_n-1})$. For different j, these blocks are separated by at least $(k-1)r_n - s_n$ observations. Hence, by (MX$_k$) and Lemma 2 of Eberlein [12], the total variation distance between the joint distribution of $\Delta_{n,jk-i}$, $1 \le j \le m_{n,k,i}$, and that of $\Delta^*_{n,jk-i}$, $1 \le j \le m_{n,k,i}$, converges to 0:

$$\left\| P^{(\Delta^*_{n,jk-i})_{1\le j\le m_{n,k,i}}} - P^{(\Delta_{n,jk-i})_{1\le j\le m_{n,k,i}}} \right\|_{\mathrm{TV}} \le m_{n,k,i}\beta^X_{n,(k-1)r_n-s_n} \to 0. \tag{B.2}$$

Hence (B.1) holds with $\Delta_{n,jk-i}$ instead of $\Delta^*_{n,jk-i}$. Summing over all $i \in \{0,\dots,k-1\}$ leads to

$$\frac{1}{\sqrt{p_n}}\sum_{j=1}^{m_n}\bigl(\Delta_{n,j}(g) - E\Delta_{n,j}(g)\bigr) = o_P(1), \quad \forall g \in \mathcal{G}. \tag{B.3}$$

Thus the fidis of $\tilde{Z}_n$ defined by

$$\tilde{Z}_n(g) := \frac{1}{\sqrt{p_n}}\sum_{i=1}^{m_n}\bigl(\tilde{V}_{n,i}(g) - E\tilde{V}_{n,i}(g)\bigr) = Z_n(g) - \frac{1}{\sqrt{p_n}}\sum_{j=1}^{m_n}\bigl(\Delta_{n,j}(g) - E\Delta_{n,j}(g)\bigr), \quad g \in \mathcal{G},$$

converge if and only if the fidis of Z_n converge, and the limits coincide if they exist.

Now, by assumption (M$\tilde{\mathrm{V}}$) and again the inequality by Eberlein [12],

$$\left\| P^{(\tilde{V}^*_{n,j})_{1\le j\le m_n}} - P^{(\tilde{V}_{n,j})_{1\le j\le m_n}} \right\|_{\mathrm{TV}} \le m_n\beta^{\tilde{V}}_{n,0} \to 0,$$

where $\tilde{V}^*_{n,i}$ are iid copies of $\tilde{V}_{n,i}$. Hence, the fidis of $\tilde{V}_n$ converge if and only if the fidis of

$$\tilde{Z}^*_n(g) := \frac{1}{\sqrt{p_n}}\sum_{i=1}^{m_n}\bigl(\tilde{V}^*_{n,i}(g) - E\tilde{V}^*_{n,i}(g)\bigr), \quad g \in \mathcal{G},$$

converge. Finally using the analog to (B.3) with $\Delta^*_{n,j}$ instead of $\Delta_{n,j}$, we arrive at the assertion. □

Proof of Theorem A.1. In view of the assumptions (L) and (C), the multivariate central limit theorem by Lindeberg–Feller yield convergence of the fidis of $(Z^*_n(g))_{g\in\mathcal{G}}$. The assertion thus follows from Lemma B.1. □

Proof of Theorem A.2. It suffices to prove that, for all $i \in \{0,\dots,k-1\}$, the processes

$$Z^{(i)}_n(g) = \frac{1}{\sqrt{p_n}}\sum_{j=1}^{m_{n,k,i}}\bigl(V_{n,kj-i}(g) - EV_{n,kj-i}(g)\bigr), \quad g \in \mathcal{G}$$

(with $m_{n,k,i} := \lfloor (m_n+i)/k \rfloor$) are asymptotically tight or asymptotically equicontinuous, respectively, since these properties carry over to their sum Z_n.

By the same arguments as in the proof of Lemma B.1 (cf. (B.2)), we may conclude

$$\left\| P^{(V^*_{n,jk-i})_{1\le j\le m_{n,k,i}}} - P^{(V_{n,jk-i})_{1\le i\le m_{n,k,i}}} \right\|_{\mathrm{TV}} \le m_{n,k,i}\beta^X_{n,(k-1)r_n-s_n} \to 0,$$

where $V^*_{n,jk-i}$, $1 \le j \le m_{n,k,i}$ are independent copies of $V_{n,1}$.

Therefore, it suffices to prove asymptotic tightness or asymptotic equicontinuity, respectively, of

$$Z^{(i)*}_n(g) = \frac{1}{\sqrt{p_n}}\sum_{j=1}^{m_{n,k,i}}\bigl(V^*_{n,kj-i}(g) - E\bigl[V^*_{n,kj-i}(g)\bigr]\bigr), \quad g \in \mathcal{G}.$$

This, however, follows under the given conditions (B), (L1), (D1) and (D2) from Theorem 2.11.9, and under the given conditions (B), (L2), (D0), (D1) and (D3) from Theorem 2.11.1 of Van der Vaart and Wellner [24]. Note that the measurability condition of the latter theorem is automatically fulfilled if the processes are separable. □

B.2. Proofs of Section 2

Proof of Theorem 2.1. First check that, by assumption (A2),

$$E^*\Big[\sup_{g\in\mathcal{G}}\big(Z_n(g)-\bar{Z}_n(g)\big)^2\Big] = E^*\Bigg[\sup_{g\in\mathcal{G}}\Bigg(\frac{1}{\sqrt{p_n}b_n(g)}\sum_{j=r_nm_n+1}^{n-s_n}\big(g(Y_{n,j})-E\big(g(Y_{n,j})\big)\big)\Bigg)^2\Bigg]$$
$$\le \frac{\|g_{\max}\|_\infty^2 r_n^2}{p_n \inf_{g\in\mathcal{G}} b_n(g)^2} \to 0,$$

which implies (2.7).

To prove convergence of the fidis, we apply Theorem A.1 to $V_{n,i}$ defined by (2.2) and $\tilde{V}_{n,i}(g) = (b_n(g))^{-1}\sum_{j=1}^{r_n-l_n} g(Y_{n,(i-1)r_n+j})$, for which condition (V) is obvious. The conditions (M$\tilde{\mathrm{V}}$) and (MX$_2$) follow readily from (MX) and $\ell_n = \mathrm{o}(r_n)$.

For the above choices, we obtain

$$\Delta_{n,1}(g) = V_{n,1}(g) - \tilde{V}_{n,1}(g) = \frac{1}{b_n(g)}\sum_{j=r_n-l_n+1}^{r_n} g(Y_{n,j}) \overset{d}{=} \frac{1}{b_n(g)}\sum_{j=1}^{l_n} g(Y_{n,j}).$$

Using the arguments of the proof of Cor. 3.6 of Drees and Rootzén [10] with $X_{n,i}$ replaced by $g(Y_{n,i})$ (cf. also the proof of Theorem 2.4), we see that

$$\begin{aligned} E\big[\big(\Delta_n(g)\big)^2\big] &\le \frac{1}{b_n(g)^2}\|g_{\max}\|_\infty^2 E\Bigg[\Bigg(\sum_{j=1}^{l_n}\mathbb{1}_{\{g(Y_{n,j})\neq 0\}}\Bigg)^2\Bigg] \\ &= \mathrm{O}\Bigg(\frac{l_n}{r_nb_n(g)^2}E\Bigg(\sum_{j=1}^{r_n}\mathbb{1}_{\{g(Y_{n,j})\neq 0\}}\Bigg)^2\Bigg) \\ &= \mathrm{o}\bigg(\frac{p_n}{m_n}\bigg). \end{aligned} \tag{B.4}$$

Hence, Condition (A.1) is fulfilled, which in turn implies Condition (Δ).

Since $g_{\max}$ is bounded and $\inf_{g\in\mathcal{G}} b_n(g) > 0$, we have

$$V_n(\mathcal{G}) = \sup_{g\in\mathcal{G}}\frac{1}{b_n(g)}\sum_{j=1}^{r_n} g(Y_{n,(i-1)r_n+j}) \le r_n\|g_{\max}\|_\infty \frac{1}{\inf_{g\in\mathcal{G}} b_n(g)} < \infty. \tag{B.5}$$

Because of $r_n = o(\sqrt{p_n} \inf_{g \in \mathcal{G}} b_n(g))$, for all $\epsilon > 0$, eventually $V_n(\mathcal{G}) \le \sqrt{p_n}\epsilon$, so that Condition (L2) (and thus (L), too) is trivial. Now the assertion follows from Theorem A.1. □

Proof of Lemma 2.2. The stationarity assumption (A1) and condition (S) imply

$$\begin{aligned} E\left(\sum_{j=1}^{r_n} \mathbb{1}_{\{g(Y_{n,j})\neq 0\}}\right)^2 &= \sum_{i=1}^{r_n}\sum_{j=1}^{r_n} E[\mathbb{1}_{\{g(Y_{n,i})\neq 0\}}\mathbb{1}_{\{g(Y_{n,j})\neq 0\}}] \\ &\le 2r_n \sum_{k=1}^{r_n}\left(1-\frac{k-1}{r_n}\right) P\{g(Y_{n,1})\neq 0, g(Y_{n,k})\neq 0\} \\ &= O\left(\frac{p_n b_n(g)^2}{m_n}\right). \end{aligned}$$

□

Proof of Theorem 2.3. Because of (2.7), the convergence of Z_n and the convergence of $\bar{Z}_n$ are equivalent. To prove the former, we apply Theorem A.2 to the processes $V_{n,i}$ and $\tilde{V}_{n,i}$ defined in the proof of Theorem 2.1. Since the conditions (V), (MX$_2$) and (L2) have already been verified there and the (Di)-conditions, $i \in \{0, 1, 2, 3\}$, are explicitly assumed to hold in Theorem 2.3, it remains to show that (B) holds. This, however, is obvious from (B.5). □

Proof of Theorem 2.4. We again apply Theorem A.1 to establish fidi-convergence of $(Z_n(g))_{g\in\mathcal{G}}$. Only the conditions ($\Delta$) and (L) must be verified, because the remaining conditions follow as in the proof of Theorem 2.1.

By the Hölder inequality, the generalized Markov inequality and (2.8), for all $g \in \mathcal{G}$, we obtain

$$\begin{aligned} & E\big[(V_n(g))^2 \mathbb{1}_{\{|V_n(g)|>\sqrt{p_n}\epsilon\}}\big] \\ &= \frac{1}{b_n^2(g)} E\left[\left(\sum_{i=1}^{r_n} g(Y_{n,i})\right)^2 \mathbb{1}_{\{|\sum_{i=1}^{r_n} g(Y_{n,i})|>\sqrt{p_n}b_n(g)\epsilon\}}\right] \\ &\le \frac{1}{b_n^2(g)}\left(E\left[\left|\sum_{i=1}^{r_n} g(Y_{n,i})\right|^{2+\delta}\right]\right)^{2/(2+\delta)} \big(E[\mathbb{1}_{\{|\sum_{i=1}^{r_n} g(Y_{n,i})|>\sqrt{p_n}b_n(g)\epsilon\}}]\big)^{\delta/(2+\delta)} \\ &\le \frac{1}{b_n^2(g)}\left(E\left[\left|\sum_{i=1}^{r_n} g(Y_{n,i})\right|^{2+\delta}\right]\right)^{2/(2+\delta)} \left(\frac{E[|\sum_{i=1}^{r_n} g(Y_{n,i})|^{2+\delta}]}{(\sqrt{p_n}b_n(g)\epsilon)^{2+\delta}}\right)^{\delta/(2+\delta)} \\ &= O\left(\frac{1}{b_n^2(g)}\cdot\frac{p_n b_n^2(g)}{m_n}\cdot\frac{1}{(\sqrt{p_n}b_n(g))^{\delta}}\right) = o\left(\frac{p_n}{m_n}\right), \end{aligned}$$

because $\sqrt{p_n}b_n(g) \to \infty$ by assumption (A2). It is easily seen (cf. Section 7 in the Supplement [9]) that this bound implies condition (L).

Furthermore,

$$E\left[\left(\sum_{i=1}^{r_n}\left|g(Y_{n,i})\right|\right)^2\right] \geq \sum_{j=1}^{\lfloor r_n/l_n\rfloor} E\left[\left(\sum_{i=1}^{l_n}\left|g(Y_{n,(j-1)l_n+i})\right|\right)^2\right] = \lfloor r_n/l_n\rfloor E\left[\left(\sum_{i=1}^{l_n}\left|g(Y_{n,i})\right|\right)^2\right]$$

and thus, by (2.8),

$$\begin{aligned}
E\big(\Delta_n(g)^2\big) &\leq \frac{1}{b_n^2(g)} E\left[\left(\sum_{i=1}^{l_n}\left|g(Y_{n,i})\right|\right)^2\right] \\
&\leq \frac{1}{b_n^2(g)\lfloor r_n/l_n\rfloor} E\left[\left(\sum_{i=1}^{r_n}\left|g(Y_{n,i})\right|\right)^2\right] \\
&\leq \frac{1}{b_n^2(g)\lfloor r_n/l_n\rfloor} E\left[\left(\sum_{i=1}^{r_n}\left|g(Y_{n,i})\right|\right)^{2+\delta} + \mathbb{1}_{\{\sum_{i=1}^{r_n}|g(Y_{n,i})|\neq 0\}}\right] \\
&= \mathrm{O}\left(\frac{l_n}{r_n b_n^2(g)}\left(\frac{p_n b_n^2(g)}{m_n} + P\big\{V_n\big(|g|\big)\neq 0\big\}\right)\right) = \mathrm{o}\left(\frac{p_n}{m_n}\right),
\end{aligned}$$

where in the last step we have used the assumption $m_n l_n P\{V_n(|g|)\neq 0\} = \mathrm{o}(r_n b_n^2(g) p_n)$ for all $g\in\mathcal{G}$. Hence, condition (A.1) holds, which in turn implies (Δ). Now, the convergence of the fidis of $(Z_n(g))_{g\in\mathcal{G}}$ follows from Theorem A.1.

Similarly,

$$\begin{aligned}
E\big(\big(\bar{Z}_n(g) - Z_n(g)\big)^2\big) &= \mathrm{Var}\left[\frac{1}{\sqrt{p_n}b_n(g)} \sum_{j=r_n m_n+1}^{n-s_n} g(Y_{n,j})\right] \\
&\leq \frac{1}{p_n b_n^2(g)} E\left[\left(\sum_{j=r_n m_n+1}^{n-s_n}\left|g(Y_{n,j})\right|\right)^2\right] \\
&= \mathrm{O}\left(\frac{1}{m_n} + \frac{P\{V_n(|g|)\neq 0\}}{p_n b_n^2(g)}\right) \to 0,
\end{aligned}$$

because $p_n b_n^2(g)\to\infty$ by assumption (A2), so that the fidi-convergence of $(\bar{Z}_n(g))_{g\in\mathcal{G}}$ follows, too.

Under the conditions of part (ii), the above calculations with $g_{\max}$ instead of g yield (2.7) as well as

$$E^*\big[\big(V_n(\mathcal{G})\big)^2\mathbb{1}_{\{V_n(\mathcal{G})>\sqrt{p_n}\epsilon\}}\big] = \frac{1}{b_n^2} E\left[\left(\sum_{i=1}^{r_n} g_{\max}(Y_{n,i})\right)^2 \mathbb{1}_{\{\sum_{i=1}^{r_n} g_{\max}(Y_{n,i})>\sqrt{p_n}b_n\epsilon\}}\right] = \mathrm{o}\left(\frac{p_n}{m_n}\right),$$

that is, (L2). Since Condition (B) is obvious, the assertion follows from Theorem A.2. $\square$

B.3. Proofs of Section 2.1

Proof of Theorem 2.5. We compare the pre-asymptotic variances which converge to $c^{(d)}$ and $c^{(s)}$, respectively. Check that, by stationarity,

$$\begin{aligned}
&\frac{1}{r_n v_n a_n^2}\operatorname{Var}\Bigg(\sum_{i=1}^{r_n/s_n} g(Y_{n,is_n+1})\Bigg)\\
&\quad=\frac{1}{r_n v_n a_n^2}E\Bigg[\sum_{i=1}^{r_n/s_n}\sum_{j=1}^{r_n/s_n} g(Y_{n,js_n+1})g(Y_{n,is_n+1})\Bigg]-\frac{1}{r_n v_n a_n^2}\bigg(\frac{r_n}{s_n}E\big[g(Y_{n,0})\big]\bigg)^2\\
&\quad=\frac{1}{r_n v_n a_n^2}\sum_{k=-r_n/s_n+1}^{r_n/s_n-1}\bigg(\frac{r_n}{s_n}-|k|\bigg)E\big[g(Y_{n,ks_n})g(Y_{n,0})\big]-\frac{r_n E[g(Y_{n,0})]^2}{s_n^2 v_n a_n^2}\\
&\quad=\frac{1}{s_n v_n a_n^2}\sum_{l=-r_n+1}^{r_n-1}\mathbb{1}_{\{l \bmod s_n=0\}}\bigg(1-\frac{|l|}{r_n}\bigg)E\big[g(Y_{n,l})g(Y_{n,0})\big]-\frac{r_n E[g(Y_{n,0})]^2}{s_n^2 v_n a_n^2}.
\end{aligned}$$

Similarly

$$\begin{aligned}
&\frac{1}{r_n v_n s_n^2 a_n^2}\operatorname{Var}\Bigg(\sum_{i=1}^{r_n} g(Y_{n,i})\Bigg)\\
&\quad=\frac{1}{v_n s_n^2 a_n^2}\sum_{k=-r_n+1}^{r_n-1}\bigg(1-\frac{|k|}{r_n}\bigg)E\big[g(Y_{n,0})g(Y_{n,k})\big]-\frac{r_n E[g(Y_{n,0})]^2}{v_n s_n^2 a_n^2}.
\end{aligned}$$

In view of (2.10) and (2.13), it suffices to show that the difference between these pre-asymptotic variances

$$\frac{1}{s_n v_n a_n^2}\Bigg(\sum_{k=-r_n+1}^{r_n-1}\bigg(1-\frac{|k|}{r_n}\bigg)\gamma_n(k)E\big[g(Y_{n,0})g(Y_{n,k})\big]\Bigg)$$

is non-negative. Here

$$\gamma_n(k)=\begin{cases}1-\dfrac{1}{s_n}, & \text{if } k \bmod s_n=0,\\ -\dfrac{1}{s_n}, & \text{if } k \bmod s_n\neq 0,\end{cases}$$

for $k\in\mathbb{Z}$. To this end, we take up an idea by Zou et al. [26], proof of Lemma A.10.

Let U_n be uniformly distributed on $\{0,\dots,s_n-1\}$ and independent of $(X_{n,i})_{1\le i\le n}$. Define

$$\phi_{n,k}=\begin{cases}\dfrac{s_n-1}{\sqrt{s_n}}, & \text{if } k \bmod s_n=U_n,\\ -\dfrac{1}{\sqrt{s_n}}, & \text{else,}\end{cases}$$

for $k \in \mathbb{Z}$. If $(h \bmod s_n) = 0$ then

$$E[\phi_{n,k}\phi_{n,k+h}] = \frac{1}{s_n} \cdot \frac{(s_n-1)^2}{s_n} + \frac{s_n-1}{s_n} \cdot \frac{1}{s_n} = 1 - \frac{1}{s_n},$$

whereas for $(h \bmod s_n) \neq 0$

$$E[\phi_{n,k}\phi_{n,k+h}] = \frac{2}{s_n} \cdot \frac{s_n-1}{\sqrt{s_n}} \cdot \frac{-1}{\sqrt{s_n}} + \frac{s_n-2}{s_n} \cdot \frac{1}{s_n} = -\frac{1}{s_n}.$$

Thus, $E[\phi_{n,k}\phi_{n,k+h}] = \gamma_n(h)$ and

$$\begin{aligned} E\big[\phi_{n,j}\phi_{n,i} g(Y_{n,i}) g(Y_{n,j})\big] &= E[\phi_{n,j}\phi_{n,i}] E\big[g(Y_{n,i}) g(Y_{n,j})\big] \\ &= \gamma_n\big(|i-j|\big) E\big[g(Y_{n,0}) g(Y_{n,|i-j|})\big] \end{aligned}$$

for all $i, j \in \{1, \ldots, r_n\}$, since U_n and $(X_1, \ldots, X_n)$ are independent. Similarly as above, we conclude

$$\begin{aligned} 0 &\le \frac{1}{r_n} E\left[\left(\sum_{j=1}^{r_n} \phi_{n,j} g(Y_{n,j})\right)^2\right] = \frac{1}{r_n} \sum_{j=1}^{r_n} \sum_{i=1}^{r_n} \gamma_n\big(|i-j|\big) E\big[g(Y_{n,0}) g(Y_{n,|i-j|})\big] \\ &= \sum_{k=-r_n+1}^{r_n-1} \left(1 - \frac{|k|}{r_n}\right) \gamma_n\big(|k|\big) E\big[g(Y_{n,0}) g(Y_{n,k})\big], \end{aligned}$$

which proves the assertion. □

Proof of Theorem 2.6. W.l.o.g. we assume $g \ge 0$ which implies $\xi \ge 0$. Since $\tilde{c}^{(d)} - \tilde{c}^{(s)} = c^{(d)} - c^{(s)} - 2\xi(c^{(d,v)} - c^{(s,v)})$, in view of Theorem 2.5 it suffices to show that $c^{(d,v)} \le c^{(s,v)}$. Using the row-wise stationarity of the triangular scheme, the asymptotic covariance $c^{(s,v)}$ can be calculated as the limit of

$$\begin{aligned} &\frac{1}{r_n v_n s_n a_n} \operatorname{Cov}\left(\sum_{j=1}^{r_n} g(Y_{n,j}), \sum_{i=1}^{r_n} \mathbb{1}_{\{X_{n,i} \neq 0\}}\right) \\ &\quad = \frac{1}{r_n v_n s_n a_n} \sum_{j=1}^{r_n} E\left[g(Y_{n,j}) \sum_{i=1}^{r_n} \mathbb{1}_{\{X_{n,i} \neq 0\}}\right] - \frac{1}{r_n v_n s_n a_n} \cdot r_n E g(Y_{n,1}) \cdot r_n v_n \\ &\quad = \frac{1}{r_n v_n s_n a_n} \sum_{j=1}^{r_n} E\left[g(Y_{n,1}) \sum_{i=2-j}^{r_n-j+1} \mathbb{1}_{\{X_{n,i} \neq 0\}}\right] + \frac{r_n E g(Y_{n,1})}{s_n a_n}. \end{aligned} \tag{B.6}$$

Likewise, $c^{(d,v)}$ is the limit of

$$\frac{1}{r_n v_n a_n}\operatorname{Cov}\left(\sum_{k=1}^{r_n/s_n} g(Y_{n,(k-1)s_n+1}), \sum_{i=1}^{r_n} \mathbb{1}_{\{X_{n,i}\neq 0\}}\right)$$

$$= \frac{1}{r_n v_n a_n}\sum_{k=1}^{r_n/s_n} E\left[g(Y_{n,1})\sum_{i=1-(k-1)s_n}^{r_n-(k-1)s_n} \mathbb{1}_{\{X_{n,i}\neq 0\}}\right] + \frac{r_n Eg(Y_{n,1})}{s_n a_n}$$

$$= \frac{1}{r_n v_n s_n a_n}\sum_{j=1}^{r_n} E\left[g(Y_{n,1})\sum_{i=1-\lfloor\frac{j-1}{s_n}\rfloor s_n}^{r_n-\lfloor\frac{j-1}{s_n}\rfloor s_n} \mathbb{1}_{\{X_{n,i}\neq 0\}}\right] + \frac{r_n Eg(Y_{n,1})}{s_n a_n}. \tag{B.7}$$

It remains to show that the limit superior of the following difference between both right-hand sides of (B.7) and (B.6) is not positive. To this end, note that

$$\frac{1}{r_n v_n s_n a_n}\sum_{j=1}^{r_n} E\left[g(Y_{n,1})\left(\sum_{i=1-\lfloor\frac{j-1}{s_n}\rfloor s_n}^{r_n-\lfloor\frac{j-1}{s_n}\rfloor s_n} \mathbb{1}_{\{X_{n,i}\neq 0\}} - \sum_{i=2-j}^{r_n-j+1} \mathbb{1}_{\{X_{n,i}\neq 0\}}\right)\right]$$

$$\le \frac{1}{r_n v_n s_n a_n}\sum_{j=2}^{r_n}\sum_{i=r_n-j+2}^{r_n-\lfloor\frac{j-1}{s_n}\rfloor s_n} E\big(g(Y_{n,1})\mathbb{1}_{\{X_{n,i}\neq 0\}}\big)$$

$$= \frac{1}{r_n v_n s_n a_n}\sum_{i=2}^{r_n}\sum_{j=r_n-i+2}^{(\lfloor\frac{r_n-i}{s_n}\rfloor+1)s_n} E\big(g(Y_{n,1})\mathbb{1}_{\{X_{n,i}\neq 0\}}\big)$$

$$\le \frac{1}{r_n v_n a_n}\sum_{i=2}^{r_n} E\big(g(Y_{n,1})\mathbb{1}_{\{X_{n,i}\neq 0\}}\big). \tag{B.8}$$

Note that $Eg(Y_{n,1}) = \mathrm{O}(s_n a_n v_n)$ by (2.9). Using

$$\begin{aligned} E\big(g(Y_{n,1})\mathbb{1}_{\{X_{n,i}\neq 0\}}\big) &\le Eg(Y_{n,1})P\{X_{n,i}\neq 0\} + 2\|g\|_\infty \beta^X_{n,i-s_n-1} \\ &= \mathrm{O}\big(s_n a_n v_n^2\big) + 2\|g\|_\infty \beta^X_{n,i-s_n-1} \end{aligned}$$

for $i > s_n + k_n$ (see Doukhan [6], Section 1.2, Lemma 3 and Section 1.1, Prop. 1) and $E(g(Y_{n,1})\mathbb{1}_{\{X_{n,i}\neq 0\}}) \le \|g\|_\infty v_n$ for $i \le s_n + k_n$, we conclude that (B.8) is bounded by

$$\frac{s_n + k_n}{r_n a_n}\|g\|_\infty + \mathrm{O}(s_n v_n) + \frac{2\|g\|_\infty}{r_n v_n a_n}\sum_{l=k_n}^{r_n} \beta^X_{n,l}$$

which tends to 0 under the given conditions. (In fact, similarly one can establish a lower bound on the difference between the pre-asymptotic covariances which shows that the difference tends to 0.) □

B.4. Proofs of Section 3

If $(\theta 1)$ and $(\theta \mathrm{P})$ hold for some sequence r_n, then the former is obviously fulfilled by $r_n^* := \lfloor r_n/s_n \rfloor s_n$, too, and $(\theta \mathrm{P})$ remains true because of

$$\sum_{k=r_n^*+1}^{r_n} P(X_k > u_n | X_0 > u_n) \le \frac{s_n}{v_n}\big(v_n^2 + \beta_{n,r_n^*}^X\big) \le r_n v_n + \frac{n}{r_n}\beta_{n,s_n}\frac{r_n^2}{n v_n} \to 0.$$

Moreover, the arguments given in Section 2.1 show that the limit c in $(\theta 2)$ does not change if we replace r_n with r_n^*. Thus, w.l.o.g. we may assume that r_n/s_n is a natural number (tending to ∞) for all $n \in \mathbb{N}$.

For all three estimators, we first prove joint convergence of a bivariate vector with components related to the numerator and the denominator, respectively, using the general theory developed in Section 2 and Appendix A.

We start with analyzing the disjoint blocks estimator using Theorem A.1. For $i \in \{1, \dots, m_n\}$ with $m_n = \lfloor (n - s_n + 1)/r_n \rfloor$, let

$$V_{n,i}^d := \frac{1}{\sqrt{m_n}} \sum_{j=1}^{r_n/s_n} \mathbb{1}_{\{M_{(i-1)r_n+(j-1)s_n+1,(i-1)r_n+js_n} > u_n\}},$$

$$\tilde{V}_{n,i}^d := \frac{1}{\sqrt{m_n}} \sum_{j=1}^{r_n/s_n-1} \mathbb{1}_{\{M_{(i-1)r_n+(j-1)s_n+1,(i-1)r_n+js_n} > u_n\}},$$

$$V_{n,i}^c := \frac{1}{\sqrt{m_n}} \sum_{j=1}^{r_n} \mathbb{1}_{\{X_{(i-1)r_n+j} > u_n\}},$$

$$\tilde{V}_{n,i}^c := \frac{1}{\sqrt{m_n}} \sum_{j=1}^{r_n-s_n} \mathbb{1}_{\{X_{(i-1)r_n+j} > u_n\}}.$$

Let $p_n = P\{M_{1,r_n} > u_n\}$. Recall that, under the conditions $(\theta 1)$ and $(\theta \mathrm{P})$, (3.1) holds for all $k_n \to \infty$, $k_n \le r_n$, which in turn yields

$$p_n = r_n v_n\big(\theta + \mathrm{o}(1)\big), \qquad P\{M_{1,s_n} > u_n\} = s_n v_n\big(\theta + \mathrm{o}(1)\big), \tag{B.9}$$

with $v_n = P\{X_1 > u_n\}$.

Proposition B.2. *If the conditions* $(\theta 1)$, $(\theta 2)$ *and* $(\theta \mathrm{P})$ *are satisfied*, *then*

$$\begin{pmatrix} Z_n^d \\ Z_n^c \end{pmatrix} := \begin{pmatrix} \frac{1}{\sqrt{p_n}} \sum_{i=1}^{m_n} \big(V_{n,i}^d - E\big[V_{n,i}^d\big]\big) \\ \frac{1}{\sqrt{p_n}} \sum_{i=1}^{m_n} \big(V_{n,i}^c - E\big[V_{n,i}^c\big]\big) \end{pmatrix} \xrightarrow{w} \begin{pmatrix} Z^d \\ Z^c \end{pmatrix} \sim \mathcal{N}_2\left(0, \begin{pmatrix} 1 & 1/\theta \\ 1/\theta & c/\theta \end{pmatrix}\right).$$

Proof. The conditions (A), (V), $(\mathrm{M}\tilde{\mathrm{V}})$ and (MX_2) follow readily from $(\theta 1)$. It thus suffices to verify the conditions (Δ), (L) (which can be checked separately for $V_{n,i}^d$ and $V_{n,i}^c$) and (C), in order to conclude the assertion from Theorem A.1.

Check that

$$\Delta_n^d := V_{n,1}^d - \tilde{V}_{n,1}^d = \frac{1}{\sqrt{m_n}} \mathbb{1}_{\{M_{r_n-s_n+1,r_n} > u_n\}} \stackrel{d}{=} \frac{1}{\sqrt{m_n}} \mathbb{1}_{\{M_{1,s_n} > u_n\}}.$$

Now (3.1) and $s_n = \mathrm{o}(r_n)$ imply (A.1), and thus (Δ), for $V_{n,i}^d$:

$$\frac{m_n}{p_n} E\big[\big(\Delta_n^d\big)^2\big] \le \frac{P\{M_{1,s_n} > u_n\}}{P\{M_{1,r_n} > u_n\}} = \frac{P\{M_{1,s_n} > u_n\}}{s_n P\{X_1 > u_n\}} \cdot \frac{r_n P\{X_1 > u_n\}}{P\{M_{1,r_n} > u_n\}} \cdot \frac{s_n}{r_n} \to 0.$$

Condition (L) for $V_{n,i}^d$ follows immediately from $V_{n,i}^d \le m_n^{-1/2} r_n/s_n = \mathrm{O}(r_n/(s_n\sqrt{nv_n})\ \sqrt{r_n v_n}) = \mathrm{o}(\sqrt{p_n})$, because of (B.9) and ($\theta$1).

Since $V_{n,1}^c$ is a sliding blocks statistic with $X_{n,i} := X_i/u_n$, bounded function $h(x_1,\dots,x_s) = 1_{(1,\infty)}(x_1)$ and $b_n = \sqrt{m_n}$, the proof of Theorem 2.1 shows that (Δ) and (L) hold if $r_n = \mathrm{o}(\sqrt{p_n} b_n) = \mathrm{o}(\sqrt{r_n v_n m_n}) = \mathrm{o}(\sqrt{n v_n})$ and condition (2.6) is satisfied; both are immediate consequences of our assumptions (θ1) and (θ2).

It remains to show convergence (C) of the covariance matrix. To this end, first note that by stationarity one has uniformly for all $1 \le \ell \le r_n - s_n$

$$\begin{aligned}
\sum_{j=\ell+s_n+1}^{r_n} P\{M_{\ell+1,\ell+s_n} > u_n, X_j > u_n\} &\le \sum_{j=s_n+1}^{r_n} P\{M_{1,s_n} > u_n, X_j > u_n\} \\
&\le \sum_{i=1}^{s_n} \sum_{j=s_n+1}^{r_n} P\{X_i > u_n, X_j > u_n\} \\
&= s_n v_n \sum_{k=1}^{r_n} \min\Big(1, \frac{k}{s_n}, \frac{r_n-k}{s_n}\Big) P(X_k > u_n | X_0 > u_n) \\
&= \mathrm{o}(s_n v_n). \qquad \text{(B.10)}
\end{aligned}$$

In the last step we have used Pratt's lemma (Pratt [19]) according to which, under condition (θP), the limit of the last sum can be calculated as the infinite sum of the limit of each summand, which all equal 0, because $k/s_n \to 0$. Likewise,

$$\sum_{j=1}^{\ell} P\{M_{\ell+1,\ell+s_n} > u_n, X_j > u_n\} \le \sum_{i=1}^{s_n} \sum_{j=s_n-r_n+1}^{0} P\{X_i > u_n, X_j > u_n\} = \mathrm{o}(s_n v_n) \qquad \text{(B.11)}$$

uniformly for $1 \le \ell \le r_n - s_n$.

By stationarity and (B.9),

$$\begin{aligned}
\frac{m_n}{p_n} \operatorname{Var}\big(V_n^d\big) &= \frac{r_n}{s_n p_n} P\{M_{1,s_n} > u_n\}\big(1 - P\{M_{1,s_n} > u_n\}\big) \\
&\quad + \frac{2}{p_n} \sum_{1 \le i < j \le r_n/s_n} \operatorname{Cov}\big(\mathbb{1}_{\{M_{(i-1)s_n+1,is_n} > u_n\}}, \mathbb{1}_{\{M_{(j-1)s_n+1,js_n} > u_n\}}\big)
\end{aligned}$$

$$= \bigl(1 + \mathrm{o}(1)\bigr) + \frac{2}{p_n} \sum_{1 \le i < j \le r_n/s_n} P\{M_{(i-1)s_n+1, is_n} > u_n, M_{(j-1)s_n+1, js_n} > u_n\} + \mathrm{O}\biggl(\frac{1}{p_n}\Bigl(\frac{r_n}{s_n}\Bigr)^2 (s_n v_n)^2\biggr).$$

In view of (B.10), the second term can be bounded by

$$\frac{2}{p_n} \sum_{i=1}^{r_n/s_n - 1} \sum_{k=is_n+1}^{r_n} P\{M_{(i-1)s_n+1, is_n} > u_n, X_k > u_n\} = \mathrm{o}\Bigl(\frac{r_n v_n}{p_n}\Bigr) = \mathrm{o}(1).$$

Since $(r_n/s_n)^2(s_n v_n)^2/p_n = \mathrm{O}(r_n v_n) \to 0$ by (B.9) and (θ1), we conclude

$$\frac{m_n}{p_n} \operatorname{Var}\bigl(V_n^d\bigr) \to 1.$$

Next check that, by (B.9) and (θ2),

$$\begin{aligned}
\frac{m_n}{p_n} \operatorname{Var}\bigl(V_{n,1}^c\bigr) &= \frac{1}{p_n} \operatorname{Var}\Biggl(\sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}}\Biggr) \\
&= \frac{r_n v_n}{p_n} \cdot \frac{1}{r_n v_n} E\Biggl[\Biggl(\sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}}\Biggr)^2\Biggr] - \frac{1}{p_n}(r_n v_n)^2 \\
&= \bigl(1/\theta + \mathrm{o}(1)\bigr)\bigl(c + \mathrm{o}(1)\bigr) + \mathrm{O}(r_n v_n) \\
&\to c/\theta.
\end{aligned} \tag{B.12}$$

Finally, again by (B.9), (B.10) and (B.11),

$$\begin{aligned}
&\frac{m_n}{p_n} \operatorname{Cov}\bigl(V_{n,1}^d, V_{n,1}^c\bigr) \\
&\quad = \frac{1}{p_n}\Biggl(\sum_{i=1}^{r_n/s_n} \sum_{j=1}^{r_n} P\{M_{(i-1)s_n+1, is_n} > u_n, X_j > u_n\} - \frac{r_n}{s_n} P\{M_{1,s_n} > u_n\} r_n v_n\Biggr) \\
&\quad = \frac{1}{p_n} \sum_{i=1}^{r_n/s_n} \Biggl(s_n v_n + \sum_{j=1}^{(i-1)s_n} P\{M_{(i-1)s_n+1, is_n} > u_n, X_j > u_n\} \\
&\qquad\qquad + \sum_{j=is_n+1}^{r_n} P\{M_{(i-1)s_n+1, is_n} > u_n, X_j > u_n\}\Biggr) + \mathrm{O}(r_n v_n) \\
&\quad = \frac{1}{p_n} \sum_{i=1}^{r_n/s_n} \bigl(s_n v_n + \mathrm{o}(s_n v_n)\bigr) + \mathrm{O}(r_n v_n) \\
&\quad \to 1/\theta.
\end{aligned}$$

$\square$

Next, we turn to the sliding blocks estimator. Numerator and denominator can be written in terms of the process $\bar{Z}_n$ in (2.1) based on $X_{n,i} := X_i/u_n$ and the following bounded functions:

$$g(x_1,\dots,x_s) := \mathbb{1}_{\{\max_{1\le i\le s} x_i>1\}}, \qquad h(x_1,\dots,x_s) := \mathbb{1}_{\{x_1>1\}}.$$

As normalizing sequences we choose $b_n(g) = \sqrt{nv_n/p_n^s}s_n$ and $b_n(h) = \sqrt{nv_n/p_n^s}$ with $p_n^s := P\{M_{1,r_n+s_n-1} > u_n\} = r_n v_n \theta(1+\mathrm{o}(1))$, by (3.1).

Proposition B.3. *If the conditions* (θ1), (θ2) *and* (θP) *are satisfied*, *then*

$$\begin{pmatrix} \bar{Z}_n(g) \\ \bar{Z}_n(h) \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{nv_n}s_n} \displaystyle\sum_{i=1}^{n-s_n+1} (\mathbb{1}_{\{M_{i,i+s_n-1}>u_n\}} - p_n) \\ \dfrac{1}{\sqrt{nv_n}} \displaystyle\sum_{i=1}^{n-s_n+1} (\mathbb{1}_{\{X_i>u_n\}} - v_n) \end{pmatrix}$$
$$\xrightarrow{w} \begin{pmatrix} Z(g) \\ Z(h) \end{pmatrix} \sim \mathcal{N}_2\left(0, \begin{pmatrix} \theta & 1 \\ 1 & c \end{pmatrix}\right).$$

Proof. We are going to apply Theorem 2.1. Condition (A1) is obvious, and (A2) with $l_n = 2s_n - 1$ and (MX) easily follow from (θ1). Condition (2.6) for the functional h is immediate from (θ2) (see proof of Proposition B.2). To check it for g, we employ Lemma 2.2. First note that $p_n^s b_n(g)^2/n = s_n^2 v_n$. Moreover, by stationarity of the time series,

$$\begin{aligned}
&\frac{1}{s_n^2 v_n} \sum_{k=1}^{r_n} P\{M_{1,s_n} > u_n, M_{k,k+s_n-1} > u_n\} \\
&\quad \le \frac{1}{s_n^2 v_n} \sum_{k=1}^{r_n}\sum_{i=1}^{s_n}\sum_{j=k}^{k+s_n-1} P\{X_i > u_n, X_j > u_n\} \\
&\quad \le \frac{1}{s_n v_n} \sum_{i=1}^{s_n}\left(\sum_{j=1}^{s_n} P\{X_i > u_n, X_j > u_n\} + \sum_{j=s_n+1}^{r_n+s_n-1} P\{X_i > u_n, X_j > u_n\}\right) \\
&\quad \le 1 + 2\sum_{k=1}^{s_n-1} P(X_k > u_n | X_0 > u_n) + \sum_{k=1}^{r_n+s_n-2} P(X_k > u_n | X_0 > u_n).
\end{aligned}$$

Therefore, condition (S) follows from (θP) and

$$\sum_{k=r_n+1}^{r_n+s_n-2} P(X_k > u_n | X_0 > u_n) \le \frac{s_n}{v_n}\left(v_n^2 + \beta_{n,r_n}^X\right) = \mathrm{o}\left(s_n v_n + \frac{n}{r_n}\beta_{n,r_n}^X\right) \to 0.$$

Then, condition (2.6) for g follows from Lemma 2.2. It remains to prove convergence (C) of the standardized covariance matrix. For the variance pertaining to g and the covariance, this is done in

Lemma B.5 (iii) and (iv). The convergence

$$\frac{m_n}{p_n^s}\operatorname{Var}\bigl(V_n(h)\bigr)=\frac{1+\mathrm{o}(1)}{r_n v_n}\operatorname{Var}\left(\sum_{j=1}^{r_n}\mathbb{1}_{\{X_j>u_n\}}\right)\to c$$

has been shown in (B.12). □

Finally, we examine the statistics pertaining to the runs estimator, again using Theorem 2.1. Here we consider $X_{n,i}$, and the functions h defined above and

$$f(x_1,\dots,x_s)=\mathbb{1}_{\{x_1>1,\max_{2\le i\le s}x_i\le 1\}}.$$

The normalization is chosen as $b_n:=\sqrt{nv_n/p_n}$ for both functions f and h.

Proposition B.4. *If the conditions* $(\theta 1)$, $(\theta 2)$ *and* $(\theta \mathrm{P})$ *are satisfied, then*

$$\begin{pmatrix}\bar{Z}_n(f)\\ \bar{Z}_n(h)\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{nv_n}}\sum_{i=1}^{n-s_n+1}\bigl(\mathbb{1}_{\{X_i>u_n,M_{i+1,i+s_n-1}\le u_n\}}-P\{X_1>u_n,M_{2,s_n}\le u_n\}\bigr)\\ \frac{1}{\sqrt{nv_n}}\sum_{i=1}^{n-s_n+1}(\mathbb{1}_{\{X_i>u_n\}}-v_n)\end{pmatrix}$$
$$\xrightarrow{w}\begin{pmatrix}\bar{Z}(f)\\ \bar{Z}(h)\end{pmatrix}\sim\mathcal{N}_2\left(0,\begin{pmatrix}\theta & 1\\ 1 & c\end{pmatrix}\right).$$

Proof of Proposition B.4. Conditions (A1), (A2), (MX) and (2.6) for the functional h have already been checked in the proof of Proposition B.3. Condition (2.6) for f follows readily, because $f(x)\neq 0$ implies $h(x)\neq 0$. While condition (C) for $\operatorname{Var}(V_n(h))$ has been verified in the proof of Proposition B.3, it is established for $\operatorname{Var}(V_n(f))$ and $\operatorname{Cov}(V_n(f),V_n(h))$ in Lemma B.5 (i) and (ii). Thus, the assertion follows from Theorem 2.1. □

Now Theorem 3.1 easily follows from the above propositions by a continuous mapping argument.

Proof of Theorem 3.1. Since the arguments are basically the same for all three estimators, we give the details only for the disjoint blocks estimator. In view of $E[V_n^c]=m_n^{-1/2}r_n v_n$, $E[V_n^d]=m_n^{-1/2}(r_n/s_n)P\{M_{1,s_n}>u_n\}$ and $p_n^{1/2}m_n^{-1/2}(r_n v_n)^{-1}=(\theta/(nv_n))^{1/2}(1+\mathrm{o}(1))=\mathrm{o}(1)$ (by (B.9) and $(\theta 1)$), direct calculations show that

$$\begin{aligned}\sqrt{nv_n}\bigl(\hat{\theta}_n^d-\theta\bigr)&=\sqrt{nv_n}\left(\frac{\sum_{i=1}^{m_n}V_{n,i}^d}{\sum_{i=1}^{m_n}V_{n,i}^c}-\theta\right)\\ &=\sqrt{nv_n}\cdot\frac{\sqrt{p_n}(Z_n^d-\theta Z_n^c)+m_n(E[V_n^d]-\theta E[V_n^c])}{m_nE[V_n^c]+\sqrt{p_n}Z_n^c}\\ &=\sqrt{\frac{nv_np_n}{m_n(r_nv_n)^2}}\cdot\frac{Z_n^d-\theta Z_n^c+\sqrt{m_n/p_n}r_nv_n(P\{M_{1,s_n}>u_n\}/(s_nv_n)-\theta)}{1+\sqrt{p_n/m_n}(r_nv_n)^{-1}Z_n^c}\end{aligned}$$

$$= \sqrt{\theta}\big(1+\mathrm{o}(1)\big)\frac{Z_n^d - \theta Z_n^c + \mathrm{O}(\sqrt{n v_n})(P\{M_{1,s_n} > u_n\}/(s_n v_n) - \theta)}{1 + \mathrm{o}_P(1)}$$
$$\to \sqrt{\theta}\big(Z^d - \theta Z^c\big),$$

where in the last step we have used Proposition B.2 and the bias condition (B_b). The limit random variable is centered and normally distributed with variance $\theta(1 - 2\theta(1/\theta) + \theta^2(c/\theta)) = \theta(\theta c - 1)$. □

Lemma B.5. *If the conditions* (θ1), (θ2) *and* (θP) *are met*, *then*

(i) $\lim_{n\to\infty} \frac{1}{r_n v_n} \mathrm{Var}(\sum_{i=1}^{r_n} \mathbb{1}_{\{X_i > u_n, M_{i+1,i+s_n-1} \le u_n\}}) = \theta$,
(ii) $\lim_{n\to\infty} \frac{1}{r_n v_n} \mathrm{Cov}(\sum_{i=1}^{r_n} \mathbb{1}_{\{X_i > u_n\}}, \sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n, M_{j+1,j+s_n-1} \le u_n\}}) = 1$,
(iii) $\lim_{n\to\infty} \frac{1}{r_n s_n v_n} \mathrm{Cov}(\sum_{i=1}^{r_n} \mathbb{1}_{\{M_{i,i+s_n-1} > u_n\}}, \sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n\}}) = 1$,
(iv) $\lim_{n\to\infty} \frac{1}{r_n s_n^2 v_n} \mathrm{Var}(\sum_{i=1}^{r_n} \mathbb{1}_{\{M_{i,i+s_n-1} > u_n\}}) = \theta$.

Proof. To prove assertion (i), check that by stationarity

$$\begin{aligned}
&\frac{1}{r_n v_n}\mathrm{Var}\Bigg(\sum_{i=1}^{r_n} \mathbb{1}_{\{X_i > u_n, M_{i+1,i+s_n-1} \le u_n\}}\Bigg)\\
&\quad = \frac{1}{r_n v_n}\sum_{i=1}^{r_n}\sum_{j=1}^{r_n} P\{X_i > u_n, M_{i+1,i+s_n-1} \le u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n\}\\
&\qquad - r_n v_n\big(P(M_{2,s_n} \le u_n | X_1 > u_n)\big)^2\\
&\quad = P(M_{2,s_n} \le u_n | X_1 > u_n) + \mathrm{O}(r_n v_n)\\
&\qquad + \frac{2}{r_n v_n}\sum_{i=1}^{r_n - s_n}\sum_{j=i+s_n}^{r_n} P\{X_i > u_n, M_{i+1,i+s_n-1} \le u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n\},
\end{aligned}$$

where in the last step we have used that the probability in the sum equals 0 if $1 \le |i-j| < s_n$. The last term is bounded by

$$\frac{2}{r_n v_n}\sum_{i=1}^{r_n - s_n}\sum_{j=i+s_n}^{r_n} P\{X_i > u_n, X_j > u_n\} \le 2\sum_{k=s_n-1}^{r_n} P(X_k > u_n | X_0 > u_n) \tag{B.13}$$

and hence it tends to 0 by Pratt's lemma and (θP). Now (3.2) and $r_n v_n \to 0$ yields the convergence of the normalized variance to θ.

Next, we consider (ii). Similarly as above, stationarity implies

$$\begin{aligned}
&\frac{1}{r_n v_n}\mathrm{Cov}\Bigg(\sum_{i=1}^{r_n} \mathbb{1}_{\{X_i > u_n\}}, \sum_{j=1}^{r_n} \mathbb{1}_{\{X_j > u_n, M_{j+1,j+s_n-1} \le u_n\}}\Bigg)\\
&\quad = P(M_{2,s_n} \le u_n | X_1 > u_n)\\
&\qquad + \frac{1}{r_n v_n}\sum_{i=1}^{r_n-1}\sum_{j=i+1}^{r_n} P\{X_i > u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n\}
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{r_n v_n}\sum_{i=s_n+1}^{r_n}\sum_{j=1}^{i-s_n} P\{X_i > u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n\} + \mathrm{O}(r_n v_n) \\
&=: I + II + III + \mathrm{O}(r_n v_n),
\end{aligned}
$$

where $I \to \theta$ by (3.2). Term III can be bounded by $(r_n v_n)^{-1}\sum_{j=1}^{r_n-s_n}\sum_{i=j+s_n}^{r_n} P\{X_i > u_n, X_j > u_n\}$, which tends to 0 by (B.13). Moreover,

$$
\begin{aligned}
II = \frac{1}{r_n v_n}\sum_{i=1}^{r_n-1}\sum_{j=i+1}^{r_n} \big(&P\{X_i > u_n, X_j > u_n, M_{j+1,r_n+s_n-1} \le u_n\} \\
&+ P\{X_i > u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n, M_{j+s_n,r_n+s_n-1} > u_n\}\big).
\end{aligned}
$$

If first j is interpreted as the last instance of an exceedance in $\{i+1, \dots, r_n + s_n - 1\}$ and then i as the last instance of an exceedance in $\{1, \dots, r_n - 1\}$, then one obtains

$$
\begin{aligned}
&\frac{1}{r_n v_n}\sum_{i=1}^{r_n-1}\sum_{j=i+1}^{r_n} P\{X_i > u_n, X_j > u_n, M_{j+1,r_n+s_n-1} \le u_n\} \\
&\quad = \frac{1}{r_n v_n}\sum_{i=1}^{r_n-1} P\{X_i > u_n, M_{i+1,r_n+s_n-1} > u_n\} \\
&\quad = \frac{(r_n-1)v_n}{r_n v_n} - \frac{1}{r_n v_n}\sum_{i=1}^{r_n-1} P\{X_i > u_n, M_{i+1,r_n+s_n-1} \le u_n\} \\
&\quad = 1 + \mathrm{o}(1) - \frac{1}{r_n v_n} P\{M_{1,r_n-1} > u_n, M_{r_n,r_n+s_n-1} \le u_n\} \\
&\quad \to 1 - \theta,
\end{aligned}
$$

because of (B.9) and $P\{M_{r_n,r_n+s_n-1} > u_n\} \le s_n v_n = \mathrm{o}(r_n v_n)$. Furthermore,

$$
\begin{aligned}
&\frac{1}{r_n v_n}\sum_{i=1}^{r_n-1}\sum_{j=i+1}^{r_n} P\{X_i > u_n, X_j > u_n, M_{j+1,j+s_n-1} \le u_n, M_{j+s_n,r_n+s_n-1} > u_n\} \\
&\le \frac{1}{r_n v_n}\sum_{i=1}^{r_n-1}\sum_{j=i+1}^{r_n} \big(P\{X_i > u_n, X_j > u_n\} P\{M_{j+s_n,r_n+s_n-1} > u_n\} + \beta^X_{n,s_n-1}\big) \\
&\le r_n v_n \sum_{k=1}^{r_n} P(X_k > u_n \mid X_0 > u_n) + \frac{r_n}{v_n}\beta^X_{n,s_n-1} \\
&\to 0,
\end{aligned}
$$

by (θ1) and (θP). To sum up, $II \to 1 - \theta$, which concludes the proof of (ii).

In view of (B.10) and (B.11), the standardized covariance in (iii) equals

$$\begin{aligned}
&\frac{1}{r_n s_n v_n}\operatorname{Cov}\Bigg(\sum_{i=1}^{r_n}\mathbb{1}_{\{M_{i,i+s_n-1}>u_n\}},\sum_{j=1}^{r_n}\mathbb{1}_{\{X_j>u_n\}}\Bigg)\\
&=\frac{1}{r_n s_n v_n}\sum_{i=1}^{r_n}\sum_{j=1}^{r_n}P\{M_{i,i+s_n-1}>u_n,X_j>u_n\}-\frac{r_n}{s_n}P\{M_{1,s_n}>u_n\}\\
&=\frac{1}{r_n s_n v_n}\Bigg(\sum_{i=1}^{r_n}\sum_{j=i}^{\min(i+s_n-1,r_n)}P\{X_j>u_n\}+\mathrm{o}(r_n s_n v_n)\Bigg)+\mathrm{O}(r_n v_n)\\
&=\frac{1}{r_n s_n}\bigg((r_n-s_n+1)s_n+\frac{s_n(s_n-1)}{2}\bigg)+\mathrm{o}(1)\\
&\to 1.
\end{aligned}$$

Finally, we turn to (iv). Stationarity implies

$$\begin{aligned}
&\operatorname{Var}\Bigg(\sum_{i=1}^{r_n}\mathbb{1}_{\{M_{i,i+s_n-1}>u_n\}}\Bigg)\\
&=\sum_{i=1}^{r_n}\sum_{j=1}^{r_n}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}-\big(r_nP\{M_{1,s_n}>u_n\}\big)^2\\
&=2\sum_{i=1}^{r_n}\sum_{j=i}^{r_n}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}-r_nP\{M_{1,s_n}>u_n\}+\mathrm{O}\big((r_ns_nv_n)^2\big)\\
&=2\Bigg[\sum_{i=1}^{r_n-3s_n}\sum_{j=i}^{i+s_n-1}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}\\
&\quad+\sum_{i=r_n-3s_n+1}^{r_n}\sum_{j=i}^{r_n}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}\\
&\quad+\sum_{i=1}^{r_n-3s_n}\sum_{j=i+s_n}^{r_n-s_n}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}\\
&\quad+\sum_{i=1}^{r_n-3s_n}\sum_{j=r_n-s_n+1}^{r_n}P\{M_{i,i+s_n-1}>u_n,M_{j,j+s_n-1}>u_n\}\Bigg]+\mathrm{o}\big(r_ns_n^2v_n\big)\\
&=:2[I+II+III+IV]+\mathrm{o}\big(r_ns_n^2v_n\big).
\end{aligned}$$

Term II is of the order $s_n^2s_nv_n=\mathrm{o}(r_ns_n^2v_n)$. Term III can be bounded by

$$\sum_{i=1}^{r_n-3s_n}\sum_{j=i+s_n}^{r_n-s_n}\sum_{k=j}^{j+s_n-1}P\{M_{i,i+s_n-1}>u_n,X_k>u_n\}$$

$$\leq s_n \sum_{i=1}^{r_n-3s_n} \sum_{k=i+s_n}^{r_n} P\{M_{i,i+s_n-1} > u_n, X_k > u_n\}$$
$$= \mathrm{o}\big(r_n s_n^2 v_n\big)$$

by (B.10). Moreover, by (θ1),

$$IV \leq \sum_{i=1}^{r_n-3s_n} \sum_{j=r_n-s_n+1}^{r_n} \big(P\{M_{i,i+s_n-1} > u_n\} \cdot P\{M_{j,j+s_n-1} > u_n\} + \beta^X_{n,s_n-1}\big)$$
$$= \mathrm{O}\big(r_n s_n \big((s_n v_n)^2 + \beta^X_{n,s_n-1}\big)\big) = \mathrm{o}\big(r_n s_n^2 v_n\big)$$

because $r_n s_n = r_n^2 s_n / r_n = \mathrm{o}(n v_n s_n / r_n) = \mathrm{o}(n/r_n)$.

It remains to be shown that

$$\frac{I}{r_n s_n^2 v_n} = \frac{1+\mathrm{o}(1)}{s_n^2 v_n} \sum_{k=1}^{s_n} P\{M_{1,s_n} > u_n, M_{k,k+s_n-1} > u_n\} \to \frac{\theta}{2}.$$

Distinguish according to the last exceedance in $\{1, \ldots, s_n\}$ to conclude

$$\sum_{k=1}^{s_n} P\{M_{1,s_n} > u_n, M_{k,k+s_n-1} > u_n\}$$
$$= \sum_{k=1}^{s_n} \sum_{j=1}^{s_n} P\{X_j > u_n, M_{j+1,s_n} \leq u_n, M_{k,k+s_n-1} > u_n\}$$
$$= \sum_{k=1}^{s_n} \sum_{j=k}^{s_n} P\{X_j > u_n, M_{j+1,s_n} \leq u_n\} + \mathrm{O}\left(\sum_{k=1}^{s_n} \sum_{j=1}^{k-1} P\{X_j > u_n, M_{k,k+s_n-1} > u_n\}\right)$$
$$= \sum_{j=1}^{s_n} j P\{X_j > u_n, M_{j+1,s_n} \leq u_n\} + \mathrm{o}\big(s_n^2 v_n\big)$$
$$= \sum_{j=1}^{s_n} j P\{X_1 > u_n, M_{2,s_n-j+1} \leq u_n\} + \mathrm{o}\big(s_n^2 v_n\big),$$

where in the penultimate step we have employed (B.11). The last sum can be bounded from below by

$$\sum_{j=1}^{s_n} j P\{X_1 > u_n, M_{2,s_n} \leq u_n\} = \frac{s_n(s_n+1)}{2} v_n P(M_{2,s_n} \leq u_n | X_1 > u_n) = \frac{s_n^2 v_n}{2} \theta\big(1+\mathrm{o}(1)\big)$$

because of (3.2). Similarly, for any sequence $t_n = \mathrm{o}(s_n)$ tending to ∞, (3.2) yields the asymptotic behavior of the following upper bound

$$\sum_{j=1}^{s_n-t_n} j P\{X_1 > u_n, M_{2,t_n} \leq u_n\} + t_n s_n v_n = \frac{s_n^2 v_n}{2} \theta\big(1+\mathrm{o}(1)\big).$$

Hence, the sum divided by $s_n^2 v_n$ must tend to $\theta/2$, which concludes the proof. □

Acknowledgements

We thank two anonymous referees for their valuable comments and suggestions which lead to a substantial improvement of the presentation.

Supplementary Material

Supplement to "Asymptotics for sliding blocks estimators of rare events" (DOI: 10.3150/20-BEJ1272SUPP; .pdf). The Suplelemnt contains detailed sufficient conditions for the asymptotic normality of statistics considered in Section 2.1, refinements to Theorem 2.5 and Theorem 3.1, and a discussion of the relationship between the conditions (L), (L1) and (L2).

References

[1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels, J. (2004). *Statistics of Extremes*: *Theory and Applications*. *Wiley Series in Probability and Statistics*. Chichester: Wiley. MR2108013 https://doi.org/10.1002/0470012382

[2] Berghaus, B. and Bücher, A. (2018). Weak convergence of a pseudo maximum likelihood estimator for the extremal index. *Ann*. *Statist*. **46** 2307–2335. MR3845019 https://doi.org/10.1214/17-AOS1621

[3] Bücher, A. and Segers, J. (2018). Inference for heavy tailed stationary time series based on sliding blocks. *Electron*. *J*. *Stat*. **12** 1098–1125. MR3780041 https://doi.org/10.1214/18-EJS1415

[4] Davis, R.A., Drees, H., Segers, J. and Warchoł, M. (2018). Inference on the tail process with application to financial time series modeling. *J*. *Econometrics* **205** 508–525. MR3813529 https://doi.org/10.1016/j.jeconom.2018.01.009

[5] Davis, R.A. and Mikosch, T. (2009). The extremogram: A correlogram for extreme events. *Bernoulli* **15** 977–1009. MR2597580 https://doi.org/10.3150/09-BEJ213

[6] Doukhan, P. (1994). *Mixing*: *Properties and Examples*. *Lecture Notes in Statistics* **85**. New York: Springer. MR1312160 https://doi.org/10.1007/978-1-4612-2642-0

[7] Drees, H. (2015). Bootstrapping empirical processes of cluster functionals with application to extremograms. Preprint. Available at arXiv:1511.00420.

[8] Drees, H. and Knežević, M. (2020). Peak-over-threshold estimators for spectral tail processes: Random vs deterministic thresholds. *Extremes* **23** 465–491. MR4129561 https://doi.org/10.1007/s10687-019-00367-x

[9] Drees, H. and Neblung, S. (2021). Supplement to "Asymptotics for sliding blocks estimators of rare events." https://doi.org/10.3150/20-BEJ1272SUPP

[10] Drees, H. and Rootzén, H. (2010). Limit theorems for empirical processes of cluster functionals. *Ann*. *Statist*. **38** 2145–2186. MR2676886 https://doi.org/10.1214/09-AOS788

[11] Drees, H., Segers, J. and Warchoł, M. (2015). Statistics for tail processes of Markov chains. *Extremes* **18** 369–402. MR3377814 https://doi.org/10.1007/s10687-015-0217-1

[12] Eberlein, E. (1984). Weak convergence of partial sums of absolutely regular sequences. *Statist*. *Probab*. *Lett*. **2** 291–293. MR0777842 https://doi.org/10.1016/0167-7152(84)90067-1

[13] Ferro, C.A.T. and Segers, J. (2003). Inference for clusters of extreme values. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **65** 545–556. MR1983763 https://doi.org/10.1111/1467-9868.00401

[14] Hsing, T. (1991). Estimating the parameters of rare events. *Stochastic Process*. *Appl*. **37** 117–139. MR1091698 https://doi.org/10.1016/0304-4149(91)90064-J

[15] Hsing, T. (1993). Extremal index estimation for a weakly dependent stationary sequence. *Ann*. *Statist*. **21** 2043–2071. MR1245780 https://doi.org/10.1214/aos/1176349409

[16] Leadbetter, M.R. (1983). Extremes and local dependence in stationary sequences. *Z*. *Wahrsch*. *Verw*. *Gebiete* **65** 291–306. MR0722133 https://doi.org/10.1007/BF00532484

[17] O'Brien, G.L. (1987). Extreme values for stationary and Markov sequences. *Ann*. *Probab*. **15** 281–291. MR0877604

[18] Petrov, V.V. (1975). *Sums of Independent Random Variables*. New York: Springer. MR0388499

[19] Pratt, J.W. (1960). On interchanging limits and integrals. *Ann. Math. Stat.* **31** 74–77. MR0123673 https://doi.org/10.1214/aoms/1177705988

[20] Robert, C.Y., Segers, J. and Ferro, C.A.T. (2009). A sliding blocks estimator for the extremal index. *Electron. J. Stat.* **3** 993–1020. MR2540849 https://doi.org/10.1214/08-EJS345

[21] Segers, J. (2003). Functionals of clusters of extremes. *Adv. in Appl. Probab.* **35** 1028–1045. MR2014268 https://doi.org/10.1239/aap/1067436333

[22] Smith, R.L. and Weissman, I. (1994). Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B* **56** 515–528. MR1278224

[23] Süveges, M. (2007). Likelihood estimation of the extremal index. *Extremes* **10** 41–55. MR2407640 https://doi.org/10.1007/s10687-007-0034-2

[24] Van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics. Springer Series in Statistics*. New York: Springer. MR1385671 https://doi.org/10.1007/978-1-4757-2545-2

[25] Weissman, I. and Novak, S.Y. (1998). On blocks and runs estimators of the extremal index. *J. Statist. Plann. Inference* **66** 281–288. MR1614480 https://doi.org/10.1016/S0378-3758(97)00095-5

[26] Zou, N., Volgushev, S. and Bücher, A. (2019). Multiple block sizes and overlapping blocks for multivariate time series extremes. *Ann. Statist.* To appear. Available at arXiv:1907.09477.

Received April 2020 and revised September 2020