Weighted weak convergence of the sequential tail empirical process for heteroscedastic time series with an application to extreme value index estimation

Tobias Jennessen & Axel Bücher

Extremes (2024) 27:163–184
https://doi.org/10.1007/s10687-023-00476-8

Weighted weak convergence of the sequential tail empirical process for heteroscedastic time series with an application to extreme value index estimation

Tobias Jennessen[1] · **Axel Bücher**[1]

Received: 14 March 2023 / Revised: 7 August 2023 / Accepted: 11 August 2023 /
Published online: 24 August 2023

Abstract
The sequential tail empirical process is analyzed in a stochastic model allowing for serially dependent observations and heteroscedasticity of extremes in the sense of Einmahl et al. (J. R. Stat. Soc. Ser. B. Stat. Methodol. **78**(1), 31–51, 2016). Weighted weak convergence of the sequential tail empirical process is established. As an application, a central limit theorem for an estimator of the extreme value index is proven.

Keywords Sequential tail empirical process · Weighted weak convergence · Extreme value index · Non-stationary extremes · Regular varying time series

AMS 2000 Subject Classifications 62G32 · 62M10 · 62G20

1 Introduction

Classical extreme value statistics focuses on analyzing the extreme behavior of a set of independent and identically distributed (i.i.d.) random variables. However, this assumption is often not valid in practical situations where data are collected over time. In such cases, the observations may show serial dependence or they may be drawn from a distribution that changes continuously as time progresses.

The model developed by Einmahl et al. (2016) and extended in Bücher and Jennessen (2022) to the case of serially dependent observations allows for the consideration of non-stationary time series observations. In the latter reference, selected statistical procedures for various target parameters of interest were proposed and analyzed asymptotically. For that purpose, the authors have shown, as a crucial intermediate step, weak convergence of the sequential tail empirical process (STEP) $\mathbb{F}_n$ to some Gaussian limit $\mathbb{F}$, see Sections 2 and 3 below for details. While this result may be useful for the asymptotic analysis of various

✉ Axel Bücher
axel.buecher@hhu.de

[1] Heinrich-Heine-Universität Düsseldorf , Mathematisches Institut, Düsseldorf, Germany

statistical procedures, it may not be sufficiently informative for some others. For instance, in the serially independent case considered in Einmahl et al. (2016), the analysis of the Hill estimator required a result on weighted weak convergence of the STEP; see also Drees (2000) for similar results and discussions in the non-sequential, stationary time series case. Weighted weak convergence is indeed more informative than non-weighted convergence: since both the STEP $\mathbb{F}_n$ and its weak limit $\mathbb{F}$ are close to zero in a neighbourhood of zero, convergence of $\mathbb{F}_n/q$ to $\mathbb{F}/q$ for some suitable weight function q with $\lim_{x\downarrow 0} q(x) = 0$ entails more information on the behavior of $\mathbb{F}_n$ in that neighbourhood (which essentially concerns the most extreme observations). It is the main purpose of this paper to derive such a weighted weak convergence result in the serially dependent heteroscedastic case, thereby extending Einmahl et al. (2016) to the serially dependent case and Drees (2000) to the sequential, heteroscedastic case. As an application, we illustrate how the result can be used to deduce asymptotic normality of the Hill estimator for the extreme value index.

The remaining parts of this paper are organized as follows: in Section 2, the model assumptions needed to prove the asymptotic results are summarized and discussed, and a location-scale model meeting these assumptions is introduced. Section 3 is concerned with the weighted weak convergence of the (simple) STEP. In Section 4, a central limit theorem for the Hill estimator of the extreme value index is presented. The quality of the normal approximation is illustrated by means of Monte Carlo simulation in Section 5. Finally, all proofs are postponed to Section 6.

Throughout, all convergences are for $n \to \infty$ if not mentioned otherwise. Weak convergence is denoted by $\rightsquigarrow$. The left-continuous generalized inverse of some increasing function H is denoted by $H^{-1}(p) = \inf\{x \in \mathbb{R} : H(x) > p\}$.

2 Model assumptions

We work under the following model from Bücher and Jennessen (2022), which is an extension of the model from Einmahl et al. (2016) to the serially dependent case: for sample size n and at time points $i \in \{1, \dots, n\}$, we observe possibly dependent random variables $X_1^{(n)}, \dots, X_n^{(n)}$ with continuous cumulative distribution functions (c.d.f.s) $F_{n,1}, \dots, F_{n,n}$. We assume that all these distribution functions share a common right endpoint $x^* = \sup\{x \in \mathbb{R} : F_{n,i}(x) < 1\}$, and that there exists some continuous reference c.d.f. F with the same right endpoint x^* that is strictly increasing on its support and some positive function c on [0, 1] such that

$$\lim_{x\uparrow x^*} \frac{1 - F_{n,i}(x)}{1 - F(x)} = c\left(\frac{i}{n}\right). \tag{2.1}$$

The function c is referred to as the *scedasis function*, which we additionally assume to be a bounded and continuous probability density function. The case where $c \equiv 1$ corresponds to *homogeneous extremes*, while the opposite is referred to as *heteroscedastic extremes*. The integrated scedasis function is denoted by

$$C(s) := \int_0^s c(x)\,\mathrm{d}x, \quad s \in [0,1].$$

Serial dependence is allowed for as follows: for each $n \in \mathbb{N}$, the unobservable sample $U_1^{(n)}, \dots, U_n^{(n)}$ with $U_i^{(n)} = F_{n,i}(X_i^{(n)})$ is assumed to be an excerpt from a strictly stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ whose distribution does not depend on n. The dynamics of the extremes of the latter series will later be captured by the concept of regular variation (Basrak and Segers, 2009), see Condition (B1) below for details.

The simple sequential tail empirical process (simple STEP) $\mathbb{S}_n$ and the sequential tail empirical process (STEP) $\mathbb{F}_n$ with parameter $k \in \mathbb{N}$ are defined, for $(s,x) \in [0,1] \times [0,\infty)$, as

$$\mathbb{S}_n(s,x) = \sqrt{k}\Big\{ \frac{1}{k} \sum_{i=1}^{[ns]} \mathbf{1}\{U_i^{(n)} > 1 - \tfrac{k}{n} c(\tfrac{i}{n})x\} - xC(s) \Big\},$$

$$\mathbb{F}_n(s,x) = \sqrt{k}\Big\{ \frac{1}{k} \sum_{i=1}^{[ns]} \mathbf{1}\Big\{X_i^{(n)} > V\Big(\frac{n}{kx}\Big)\Big\} - xC(s) \Big\},$$

where $V = (\frac{1}{1-F})^{-1}$. As usual when discussing asymptotics for extremes, $k = k_n$ is assumed to be an increasing integer sequence satisfying $k \to \infty$ and $k = o(n)$ as $n \to \infty$.

Our main result, which is Proposition 3.4 below, claims weighted weak convergence of $\mathbb{F}_n$, thereby extending Proposition 6.2 in Bücher and Jennessen (2022). For that purpose, we need several additional regularity conditions. Let $L \geq 1$ be some arbitrary but fixed constant (we will consider weak convergence uniformly for $x \in [0,L]$). Set $c_\infty(L) = 1 + L\|c\|_\infty$, where $\|\cdot\|_\infty$ denotes the sup norm of a real-valued function.

(B0) **Basic assumptions.** The model assumptions formulated at the beginning of this section are met.

(B1) **Multivariate regular variation.** For each $n \in \mathbb{N}$, $U_1^{(n)}, \dots, U_n^{(n)}$ is an excerpt from a strictly stationary time series $(U_t^{(n)})_{t\in\mathbb{Z}}$ whose marginal stationary distribution is standard uniform on (0, 1). The processes $(U_t^{(n)})_{t\in\mathbb{Z}}$ are all equal in law; denote a generic version by $(U_t)_{t\in\mathbb{Z}}$. The process $Z_t = 1/(1-U_t)$ (note that Z_t is standard Pareto) is stationary and regularly varying, necessarily with index $\alpha = 1$ (Basrak and Segers, 2009).

(B2) **Regularity of *c*.** The function c is Hölder-continuous of order 1/2, that is, there exists $K_c > 0$ such that

$$|c(s) - c(s')| \leq K_c |s - s'|^{1/2} \qquad \forall\, s, s' \in [0,1].$$

(B3) **Blocking sequences and Beta-mixing.** There exist sequences $1 < \ell_n < r = r_n < n$, both converging to infinity as $n \to \infty$ and satisfying $\ell_n = o(r)$, $r = o(\sqrt{k} \vee \frac{n}{k})$, such that the beta-mixing coefficients of $(U_t)_{t\in\mathbb{Z}}$ satisfy

$$\beta(n) = o(1), \qquad \frac{n}{r}\beta(\ell_n) = o(1).$$

Moreover, the sequence r satisfies

$$r = o(k^{1/2} \log^{-5/2}(k)). \tag{2.2}$$

(B4) **Moment bound on the number of extreme observations.** There exists $\delta > 0$ such that

$$\mathrm{E}\Big[\Big\{ \sum_{s=1}^{r} \mathbf{1}(U_s > 1 - \tfrac{k}{n} c_\infty(L)) \Big\}^{2+\delta}\Big] = O(r\tfrac{k}{n}).$$

(B5) **Moment bound on extreme increments.** There exists a constant K, such that, for all sufficiently large n,

$$\mathrm{E}\Big[\Big\{ \sum_{s=1}^{r} \mathbf{1}\big(1 - \tfrac{k}{n} c(\tfrac{(j-1)r+s}{n})x \geq U_s > 1 - \tfrac{k}{n} c(\tfrac{(j-1)r+s}{n})y\big) \Big\}^{2}\Big] \leq K\frac{rk}{n}(y-x)$$

for all $j = 1, \dots, \lfloor n/r \rfloor$ and $0 \leq x \leq y \leq c_\infty(L)$.

(B6) **Second order condition.** There exists a positive, eventually decreasing function A with $\lim_{t\to\infty} A(t) = 0$ such that, as $x \uparrow x^*$,

$$\sup_{n\in\mathbb{N}} \max_{1\leq i\leq n} \left| \frac{1 - F_{n,i}(x)}{1 - F(x)} - c\big(\tfrac{i}{n}\big) \right| = O\Big(A\Big(\frac{1}{1-F(x)}\Big)\Big).$$

Note that Conditions (B0)–(B4), (B6) and a weaker version of (B5) have also been imposed in Bücher and Jennessen (2022). It is worth noting that Conditions (B4) and (B5) (and only these) depend on the constant $L \geq 1$.

Condition (B1) allows to control the serial dependence within the observed time series via *tail processes* (Basrak and Segers, 2009). More precisely, by Theorem 2.1 in Basrak and Segers (2009), regular variation of $(Z_t)_{t\in\mathbb{Z}}$ is equivalent to the fact that there exists a process $(Y_t)_{t\in\mathbb{N}_0}$ (the *tail process*) with Y_0 standard Pareto such that, for every $\ell \in \mathbb{N}$ and as $x \to \infty$,

$$\mathrm{P}(x^{-1}(Z_0, \dots, Z_\ell) \in \cdot \mid Z_0 > x) \rightsquigarrow \mathrm{P}((Y_0, \dots, Y_\ell) \in \cdot), \tag{2.3}$$

where, necessarily, $Y_j \geq 0$ for $j \geq 1$. Further, by Theorem 2 and its subsequent discussion in Segers (2003), Y_j is absolutely continuous on $(0, \infty)$ and may have an atom at 0.

Condition (B2) has also been imposed in Einmahl et al. (2016). Since $k = o(n)$, it implies that

$$\lim_{n\to\infty} \sup_{s\in[0,1]} \sqrt{k}\left| \frac{1}{n} \sum_{i=1}^{\lfloor ns \rfloor} c(\tfrac{i}{n}) - C(s) \right| = 0,$$

which will imply that there is no asymptotic bias in our main result below.

The conditions in (B3) and (B4) are essentially conditions imposed in Example 3.8 in Drees and Rootzén (2010) for deriving weak convergence of the standard non-sequential univariate tail empirical process. The assumption (2.2) on r is very similar to the one in Drees (2000). Note that the sequence ℓ_n in (B3) plays the role of a small-block length in a big-block-small-block technique, while $r - \ell_n$ is the length of a corresponding big-block.

In the case of a constant scedasis function, Condition (B5) has also been made in Drees (2000) and is further discussed in Rootzén (2009), see, e.g., Drees (2000) for solutions of stochastic recurrence equations. In the current heteroscedastic case, a slightly weaker version has been imposed in Bücher and Jennessen (2022). The condition can for instance be shown to hold for M-dependent sequences for any $M \in \mathbb{N}$.

Finally, Condition (B6) is a second-order condition on the speed of convergence in (2.1); it was also used in Einmahl et al. (2016).

Example 2.1 Consider the location-scale model defined by

$$X_i^{(n)} = \sigma\big(\tfrac{i}{n}\big) W_i + \mu\big(\tfrac{i}{n}\big), \quad i = 1, \dots, n,$$

where $(W_t)_{t\in\mathbb{Z}}$ is a strictly stationary time series (see below for an explicit example) with c.d.f. F and where $\sigma : [0,1] \to (0,\infty)$ is Hölder-continuous of order 1/2 and $\mu : [0,1] \to \mathbb{R}$ is arbitrary. We then have

$$F_{n,i}(x) = F\Bigg(\frac{x - \mu(\frac{i}{n})}{\sigma(\frac{i}{n})} \Bigg), \quad x \in \mathbb{R},$$

and $U_i^{(n)} = F_{n,i}(X_i^{(n)}) = F(W_i)$, $i = 1, \dots, n$, such that $U_1^{(n)}, \dots, U_n^{(n)}$ is an excerpt from a strictly stationary time series, with marginal distribution given by the uniform distribution on [0, 1].

Next, as a special case, let $(W_t)_{t\in\mathbb{Z}}$ be an M-dependent process for some $M \in \mathbb{N}$, i.e., $\{W_t : t \le s\}$ and $\{W_t : t > s + M\}$ are independent for all $s \in \mathbb{Z}$, with c.d.f. $F(x) = \exp(-1/x)$. Then Condition (B1) follows from Example 5.2.7 in Kulik and Soulier (2020) and Lemma 2.1 in Drees et al. (2015), and (B6) was shown in Bücher and Jennessen (2022). Further, one can easily show that Conditions (B3)–(B5) are fulfilled. In particular, this model includes moving-maximum processes of the form

$$W_t = \max_{j=0,\dots,q} a_j V_{t-j}, \quad t \in \mathbb{Z},$$

where $a_j > 0$ and $V_t, t \in \mathbb{Z}$, are independent and Frechét-distributed. General moving-maximum models have been studied and applied in Zhang and Smith (2001, 2010), Hall et al. (2002), Ferreira (2012), among others.

3 Weighted weak convergence of the STEP

The subsequent two propositions have been shown in Bücher and Jennessen (2022).

Proposition 3.1 (Proposition 6.1 in Bücher and Jennessen 2022). *Suppose that Conditions* (*B*0)–(*B*3) *hold. Fix some constant* $L \geq 1$ *and suppose that Conditions* (*B*4) *and* (*B*5) *hold for L. Then, as* $n \to \infty$,

$$\mathbb{S}_n \rightsquigarrow \mathbb{S} \quad \text{in} \quad (\ell^\infty([0,1] \times [0,L]), \|\cdot\|_\infty),$$

where $\mathbb{S}$ *denotes a tight, centered Gaussian process on* $[0,1] \times [0,L]$ *with covariance given by*

$$\mathfrak{c}((s,x),(s',x')) = C(s \wedge s')\sigma^2(x,x'),$$

where

$$\sigma^2(x,x') = d_0(x,x') + \sum_{h=1}^{\infty} \big(d_h(x,x') + d_h(x',x)\big)$$

with, recalling the tail process $(Y_t)_{t \in \mathbb{N}_0}$ *associated with* $(Z_t)_{t \in \mathbb{Z}}$ *from* (2.3),

$$d_h(x,x') = \mathrm{P}\Big(Y_0 > \frac{1}{x}, Y_h > \frac{1}{x'}\Big).$$

It is part of the assertion that the above series is convergent.

Proposition 3.2 (*Proposition* 6.2 *in Bücher and Jennessen* 2022). *Suppose that Conditions* (*B*0)–(*B*3) *and* (*B*6) *hold. Fix some constant* $L \geq 1$ *and suppose that Conditions* (*B*4) *and* (*B*5) *hold for L. If k satisfies* $\sqrt{k}A(\frac{n}{Lk}) \to 0$ *as* $n \to \infty$, *then*

$$\sup_{(s,x)\in[0,1]\times[0,L]} |\mathbb{F}_n(s,x) - \mathbb{S}_n(s,x)| = o_{\mathrm{P}}(1).$$

As a consequence,

$$\mathbb{F}_n \rightsquigarrow \mathbb{S} \quad \text{in} \quad (\ell^\infty([0,1] \times [0,L]), \|\cdot\|_\infty).$$

In the following, the above results are extended to allow for weighted weak convergence. The respective proofs are presented in Section 6.

Proposition 3.3 *Suppose that Conditions* (*B*0)–(*B*3) *hold. Fix some constant* $L \geq 1$ *and suppose that Conditions* (*B*4) *and* (*B*5) *hold for L. Then, for any* $\mu \in [0, 1/4]$,

$$\left\{ \frac{\mathbb{S}_n(s,x)}{q(x)} \right\}_{(s,x)\in[0,1]\times[0,L]} \rightsquigarrow \left\{ \frac{\mathbb{S}(s,x)}{q(x)} \right\}_{(s,x)\in[0,1]\times[0,L]}$$

in $(\ell^\infty([0,1] \times [0,L]), \|\cdot\|_\infty)$, *where* $q(x) = x^\mu$. *The above convergence also holds for* $\mu \in (1/4, 1/2)$ *provided the sequence r additionally satisfies* $r = o(k^{(3-6\mu)/\{4(1-\mu)\}})$.

Proposition 3.4 *Suppose that Conditions* ($B0$)–($B3$) *and* ($B6$) *hold. Fix some constant* $L \geq 1$ *and suppose that Conditions* ($B4$) *and* ($B5$) *hold for some* $L' > L$. *Let* k *satisfy* $\sqrt{k}A(\frac{n}{Lk}) \to 0$ *as* $n \to \infty$. *Then, for any* $\mu \in [0, 1/4]$,

$$\sup_{(s,x)\in[0,1]\times[0,L]} \left| \frac{\mathbb{F}_n(s,x)}{q(x)} - \frac{\mathbb{S}_n(s,x)}{q(x)} \right| = o_{\mathrm{P}}(1),$$

where $q(x) = x^\mu$. *As a consequence*,

$$\left\{ \frac{\mathbb{F}_n(s,x)}{q(x)} \right\}_{(s,x)\in[0,1]\times[0,L]} \rightsquigarrow \left\{ \frac{\mathbb{S}(s,x)}{q(x)} \right\}_{(s,x)\in[0,1]\times[0,L]}$$

in $(\ell^\infty([0,1]\times[0,L]), \|\cdot\|_\infty)$. *The above convergences also hold for* $\mu \in (1/4, 1/2)$ *provided the sequence* r *additionally satisfies* $r = o(k^{(3-6\mu)/\{4(1-\mu)\}})$.

The additional assumption on the sequence r in the case $\mu \in (1/4, 1/2)$ has also been used in Drees (2000, condition (2.3)). It also implies our condition (2.2) in (B3).

4 Estimation of the extreme value index

The results from the previous section can be used to derive weak convergence of the Hill estimator, see also Einmahl et al. (2016) for a similar result in the serially independent case. For that purpose, we must additionally assume that F belongs to the domain of attraction of a generalized extreme value distribution. Thus, there is a real number γ, called the extreme value index, and a positive scale function σ such that, for all $x > 0$,

$$\lim_{t\to\infty} \frac{V(tx) - V(t)}{\sigma(t)} = \frac{x^\gamma - 1}{\gamma},$$

where again $V = (\frac{1}{1-F})^{-1}$. Setting $V_{n,i} = (\frac{1}{1-F_{n,i}})^{-1}$ it can further be shown by (2.1) that

$$\lim_{t\to\infty} \frac{V_{n,i}(tx) - V_{n,i}(t)}{\sigma(t)c^\gamma(i/n)} = \frac{x^\gamma - 1}{\gamma}$$

such that all $F_{n,i}$ have the same extreme value index γ (Einmahl et al. 2016, page 32).

We only consider the heavy-tailed case $\gamma > 0$, which implies that $x^* = \infty$ and that the above limit relations can be simplified to

$$\lim_{t\to\infty} \frac{V(tx)}{V(t)} = x^\gamma \quad \text{and} \quad \lim_{t\to\infty} \frac{V_{n,i}(tx)}{V(t)c^\gamma(i/n)} = x^\gamma.$$

Our aim is to consistently estimate the extreme value index $\gamma > 0$. To this end, we will show that the classical Hill estimator can be applied and prove a corresponding central limit theorem. The proof will be based on the weighted weak convergence result of the STEP in Proposition 3.4.

Consider the order statistic $X_{n,1} \leq \ldots \leq X_{n,n}$ of $X_1^{(n)}, \ldots, X_n^{(n)}$. The classical Hill estimator is given by

$$\hat{\gamma}_n = \frac{1}{k} \sum_{j=1}^{k} \log X_{n,n-j+1} - \log X_{n,n-k}.$$

For our asymptotic result, we need the subsequent second-order condition, which has also been used in Einmahl et al. (2016).

(B7) There exists a function B, eventually being positive or negative, and some $\rho < 0$ such that $\lim_{t\to\infty} B(t) = 0$ and for any $x > 0$,

$$\lim_{t\to\infty} \frac{V(tx)/V(t) - x^\gamma}{B(t)} = x^\gamma \frac{x^\rho - 1}{\rho}.$$

Theorem 4.1 *Suppose that Conditions* (*B*0)–(*B*3), (*B*6) *and* (*B*7) *hold. Further, suppose that Conditions* (*B*4) *and* (*B*5) *hold for some* $L > 2$. *If k satisfies*

$$\sqrt{k}A(\tfrac{n}{Lk}) \to 0 \quad \text{and} \quad \sqrt{k}B(\tfrac{n}{k}) \to 0,$$

then, as $n \to \infty$,

$$\sqrt{k}(\hat{\gamma}_n - \gamma) \rightsquigarrow \mathcal{N}(0, \gamma^2\sigma^2(1,1)),$$

where σ^2 *is defined in Proposition* 3.1.

Example 4.2 Let us continue with the location-scale model in Example 2.1. In particular, we now assume $(W_t)_{t\in\mathbb{Z}}$ to be a max-autoregressive process (ARMAX) defined by the recursion

$$W_t = \max\{\lambda W_{t-1}, (1-\lambda)V_t\}, \quad t \in \mathbb{Z}, \tag{4.1}$$

where $\lambda \in [0, 1)$ and $(V_t)_{t\in\mathbb{Z}}$ is an i.i.d. sequence of Fréchet(1)-distributed random variables with c.d.f. $F(x) = \exp(-1/x)$ for $x > 0$. The corresponding stationary solution is

$$W_t = \max_{j\geq 0}(1-\lambda)\lambda^j V_{t-j}$$

which is again Fréchet(1)-distributed. It was shown in Bücher and Jennessen (2022) that the scedasis function c is given by σ. Further note that $\gamma = 1$ and

$$\lim_{t\to\infty} t\big(V(tx)/V(t) - x\big) = \frac{x-1}{2}$$

such that Condition (B7) is satisfied.

We are going to calculate the asymptotic variance in Theorem 4.1 explicitly. By Theorem 13.5.5 in Kulik and Soulier (2020) the spectral tail process $(\tilde{\Theta}_t)_t$ of $(W_t)_t$ exists and for $t \in \mathbb{N}_0$ it is of the form $\tilde{\Theta}_t = \lambda^t$. Recall that $(Y_t)_t$ denotes the tail process of $(Z_t)_t$ where $Z_t = 1/\{1 - F(W_t)\}$. Lemma 2.1 in Drees et al. (2015) implies that $Y_t = \lambda^t Y_0$, $t \in \mathbb{N}_0$, where Y_0 is standard Pareto-distributed. As a consequence, we obtain

$$d_h(1,1) = \mathrm{P}(Y_0 > 1, Y_h > 1) = \mathrm{P}(Y_0 > 1, \lambda^h Y_0 > 1) = \lambda^h, \quad h \in \mathbb{N}_0,$$

such that

$$\sigma^2(1,1) = d_0(1,1) + 2\sum_{h=1}^{\infty} d_h(1,1) = \frac{1+\lambda}{1-\lambda}.$$

According to Theorem 4.1 we arrive at $\sqrt{k}(\hat{\gamma}_n - 1) \rightsquigarrow \mathcal{N}(0, \frac{1+\lambda}{1-\lambda})$ where the limiting variance equals 1 in the case of independent variables, i.e. for $\lambda = 0$, and is strictly greater than 1 for $\lambda > 0$.

5 Simulation study

A small simulation study is carried out to analyze the normal approximation of the Hill estimator $\hat{\gamma}_n$ for the extreme value index in finite samples. Results are presented for the scale model considered in Examples 2.1 and 4.2. Precisely, consider

$$X_i^{(n)} = c\left(\tfrac{i}{n}\right) W_i, \quad i = 1, \dots, n,$$

where $(W_t)_{t\in\mathbb{Z}}$ is an ARMAX process with model parameter $\lambda \in [0, 1)$ as defined in (4.1). Set

$$c(s) = (0.5 + 2s)\mathbf{1}(s \in [0, 0.5]) + (2.5 - 2s)\mathbf{1}(s \in (0.5, 1]).$$

Note that the ARMAX model with $\lambda = 0$ corresponds to the case that the observations are independent. We call this case simply the independent model.

In the subsequent simulation study, the parameter λ of the ARMAX process is set to $\lambda = 0, 0.25, 0.5$. In each case, different sample sizes $n \in \{2000, 4000, 8000\}$ are considered and the performance of the Hill estimator is assessed based on $N = 1000$ simulation runs each. Recall that k denotes the parameter of the Hill estimator.

By Example 4.2 we know that $(\frac{1-\lambda}{1+\lambda}k)^{1/2}(\hat{\gamma}_n - 1) \approx \mathcal{N}(0, 1)$ in distribution for large n. Figures 1, 2, and 3 present histograms of $(\frac{1-\lambda}{1+\lambda}k)^{1/2}(\hat{\gamma}_n - 1)$ for values of $k \in \{50, 100, 200\}$, respectively. One can see that the approximation of the normal distribution seems to become more accurate as both n and k increase. Further, the approximation gets better for smaller values of λ which is to be expected since the temporal dependence of the underlying ARMAX-process decreases with decreasing λ.

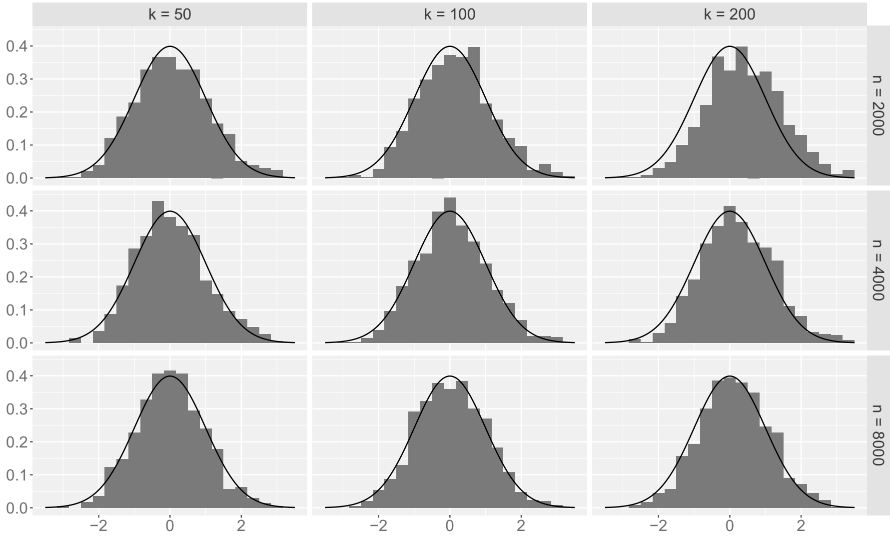

Fig. 1 Histograms of $\sqrt{\frac{1-\lambda}{1+\lambda}}k(\hat{\gamma}_n - 1)$ for $\lambda = 0$ and for values of $k \in \{50, 100, 200\}$ and $n \in \{2000, 4000, 8000\}$, compared to the density of the standard normal distribution

6 Proofs

Proof of Proposition 3.3 The finite-dimensional distributions converge by Proposition 3.1. It remains to show asymptotic tightness. First, let us rewrite $\mathbb{S}_n$ as in the proof of Proposition 6.1 in Bücher and Jennessen (2022). For $i \in \{1, \ldots, n\}$ and $n \in \mathbb{N}$, define

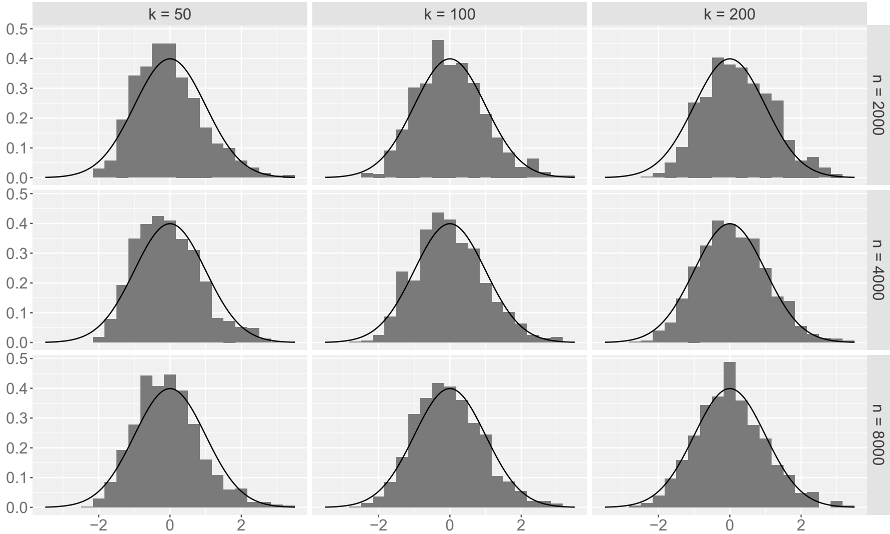

Fig. 2 Histograms of $\sqrt{\frac{1-\lambda}{1+\lambda}}k(\hat{\gamma}_n - 1)$ for $\lambda = 0.25$ and for values of $k \in \{50, 100, 200\}$ and $n \in \{2000, 4000, 8000\}$, compared to the density of the standard normal distribution

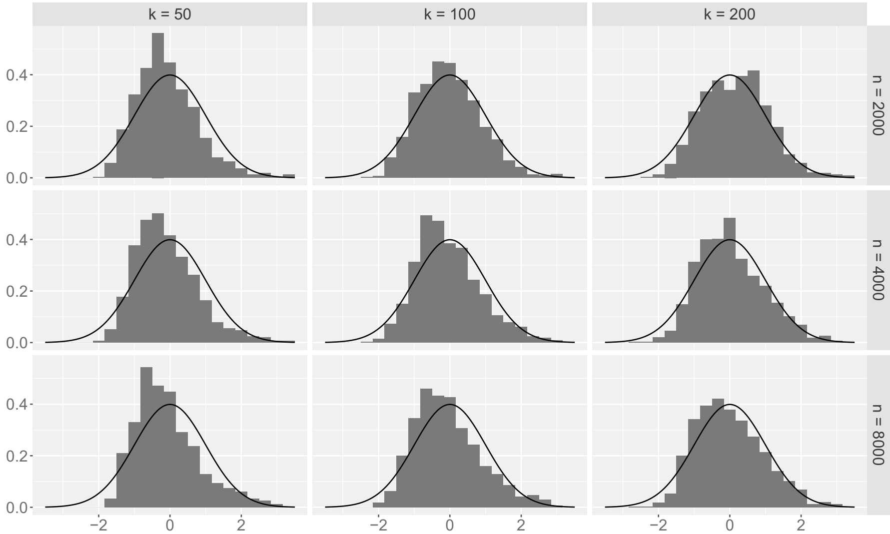

Fig. 3 Histograms of $\sqrt{\frac{1-\lambda}{1+\lambda}k}(\hat{\gamma}_n - 1)$ for $\lambda = 0.5$ and for values of $k \in \{50, 100, 200\}$ and $n \in \{2000, 4000, 8000\}$, compared to the density of the standard normal distribution

$$X'_{n,i} = \left(\frac{U_i^{(n)} - (1 - \frac{k}{n}c_\infty(L))}{\frac{k}{n}}\right)_+ = \max\left(\frac{U_i^{(n)} - (1 - \frac{k}{n}c_\infty(L))}{\frac{k}{n}}, 0\right).$$

Recall that $1 < r < n$ denotes an integer sequence converging to infinity such that $r = o(n)$ as $n \to \infty$. Let $Y_{n,j}$ denote the jth block of r consecutive values of $X'_{n,1}, \dots, X'_{n,n}$, i.e.,

$$Y_{n,j} = (X'_{n,i})_{i \in I_j}, \quad I_j = \{(j-1)r + 1, \dots, jr\}, \quad j = 1, \dots, m = \lfloor n/r \rfloor.$$

The proof of Proposition 6.1 in Bücher and Jennessen (2022) (leading to equation (8.5)) reveals that

$$\left\{\frac{\mathbb{S}_n(s,x)}{q(x)}\right\}_{(s,x)} = \left\{\frac{\mathbb{Z}_n(s,x)}{q(x)}\right\}_{(s,x)} + o_P(1) \quad \text{in} \quad (\ell^\infty([0,1] \times [0,L]), ||\cdot||_\infty), \tag{6.1}$$

where

$$\mathbb{Z}_n(s,x) = \frac{1}{\sqrt{k}} \sum_{j=1}^{m} \left\{f_{j,n,s,x}(Y_{n,j}) - \mathrm{E}[f_{j,n,s,x}(Y_{n,j})]\right\}$$

and

$$f_{j,n,s,x}(y_1, \dots, y_\ell) = \mathbf{1}(j \le \lfloor sm \rfloor) g_{j,n,x}(y_1, \dots, y_\ell)$$

with

$$g_{j,n,x}(y_1,\ldots,y_\ell)=\sum_{i=1}^{\ell}\mathbf{1}(y_i>c_\infty(L)-c(\tfrac{(j-1)r+i}{n})x),\qquad \ell\in\mathbb{N}.$$

As a consequence, it suffices to show asymptotic tightness of $\mathbb{Z}_n/q$. For this purpose, let $(Y_{n,j}^*)_{1\le j\le m}$ denote an i.i.d. sequence with $Y_{n,1}^*$ being equal in distribution to $Y_{n,1}$ and let $\mathbb{Z}_n^*$ be defined as $\mathbb{Z}_n$ but in terms of $Y_{n,j}^*$. Decomposing $\mathbb{Z}_n^*=\mathbb{Z}_n^{\mathrm{even},*}+\mathbb{Z}_n^{\mathrm{odd},*}$ into sums over even and odd numbered blocks, the same arguments as in the proof in the above reference imply that asymptotic tightness of $\mathbb{Z}_n/q$ follows from asymptotic tightness of $\mathbb{Z}_n^{\mathrm{even},*}/q$ and $\mathbb{Z}_n^{\mathrm{odd},*}/q$. Instead of these two processes, we consider $\mathbb{Z}_n^*/q$ to reduce the notational complexity.

Clearly, it is sufficient to prove weak convergence of $\mathbb{Z}_n^*/q$ on $[0,1]\times[0,L]$. In view of the functional weak convergence of $\mathbb{Z}_n^*/q$ on any fixed interval $[0,1]\times[\delta,L]$ (a consequence of the proof of Proposition 6.1 in Bücher and Jennessen, 2022) and by Theorem 25.5 in Billingsley (1995), it is sufficient to show that, for any $\varepsilon>0$,

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\mathrm{P}\Big(\sup_{s\in[0,1]}\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(s,x)|}{q(x)}\ge\varepsilon\Big)=0 \tag{6.2}$$

$$\lim_{\delta\downarrow 0}\mathrm{P}\Big(\sup_{s\in[0,1]}\sup_{0\le x\le\delta}\frac{|\mathbb{Z}^*(s,x)|}{q(x)}\ge\varepsilon\Big)=0$$

Viewing $\mathbb{Z}_n^*(s,x)/q(x)$ as an element of the complete and separable space $D([0,1]^2)$ equipped with the Skorohod-metric, it is in fact sufficient to prove only (6.2) (see Theorem 2 in Dehling et al., 2009). In order to show (6.2) we first prove that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{s\in[0,1]}\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(s,x)|}{q(x)}\ge\varepsilon\Big)=0 \tag{6.3}$$

and we show later that (6.3) implies (6.2). For the proof of (6.3) we follow ideas from Drees (2000) and Shao and Yu (1996). Let $\varepsilon_j=\varepsilon q(\delta 2^{-j})=\varepsilon(\delta 2^{-j})^\mu$. Then

$$\begin{aligned}\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(s,x)|}{q(x)}\ge\varepsilon\Big)&\le\sum_{j=1}^{\infty}\mathrm{P}\Big(\sup_{\delta 2^{-j}<x\le\delta 2^{-j+1}}\frac{|\mathbb{Z}_n^*(s,x)|}{q(x)}\ge\varepsilon\Big)\\&\le\sum_{j=1}^{\infty}\mathrm{P}\Big(\sup_{\delta 2^{-j}<x\le\delta 2^{-j+1}}|\mathbb{Z}_n^*(s,x)|\ge\varepsilon q(\delta 2^{-j})\Big)\\&\le\sum_{j=1}^{\infty}\mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}}|\mathbb{Z}_n^*(s,x)|\ge\varepsilon_j\Big).\end{aligned}$$

Split the above sum according to whether $j\in G_n$ or $j\in H_n$, where

$$G_n = \{j \in \mathbb{N} : \sqrt{k}\delta 2^{-j+1} \le \varepsilon_j/2\},$$

$$H_n = \{j \in \mathbb{N} : \sqrt{k}\delta 2^{-j+1} > \varepsilon_j/2\}.$$

To prove (6.3), it suffices to show that

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{s\in[0,1]}\sum_{j\in G_n} \mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge \varepsilon_j\Big) = 0, \tag{6.4}$$

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\sup_{s\in[0,1]}\sum_{j\in H_n} \mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge \varepsilon_j\Big) = 0. \tag{6.5}$$

Let us start by showing (6.4). For $j \in G_n$, by (6.10) and (6.8) in Lemma 6.1 below and Markov's inequality,

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge 2\varepsilon_j\Big) &\le \mathrm{P}\Big(|\mathbb{Z}_n^*(s,\delta 2^{-j+1})| + \sqrt{k}\delta 2^{-j+2} \ge 2\varepsilon_j\Big) \\
&\le \mathrm{P}\Big(|\mathbb{Z}_n^*(s,\delta 2^{-j+1})| \ge \varepsilon_j\Big) \\
&\le \frac{\delta 2^{-j+1}}{\varepsilon_j^2} = \frac{2}{\varepsilon^2}(\delta 2^{-j})^{1-2\mu}
\end{aligned}$$

for sufficiently large n. Note that the upper bound is uniform in $s \in [0,1]$. Hence,

$$\limsup_{n\to\infty}\sup_{s\in[0,1]}\sum_{j\in G_n} \mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge 2\varepsilon_j\Big) \lesssim \delta^{1-2\mu}\sum_{j=1}^{\infty}(2^{1-2\mu})^{-j},$$

and this expression goes to zero for $\delta \downarrow 0$. We have shown (6.4).

It remains to prove (6.5). For that purpose, let

$$\Delta_{n,j} = \frac{\varepsilon_j}{8\sqrt{k}}$$

and start by observing the bound

$$\begin{aligned}
\mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge \varepsilon_j\Big) &\le \mathrm{P}\Big(\max_{1\le i\le\Delta_{n,j}^{-1}\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})| \ge \varepsilon_j/2\Big) \\
&\quad + \mathrm{P}\Big(\max_{0\le i\le\Delta_{n,j}^{-1}\delta 2^{-j+1}}\ \sup_{x'\in(i\Delta_{n,j},(i+1)\Delta_{n,j}]} |\mathbb{Z}_n^*(s,i\Delta_{n,j}) - \mathbb{Z}_n^*(s,x')| \ge \varepsilon_j/2\Big).
\end{aligned}$$

By (6.10) in Lemma 6.1, the second probability on the right-hand side may be further bounded by

$$\mathrm{P}\Big(\max_{0\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})-\mathbb{Z}_n^*(s,(i+1)\Delta_{n,j})| + 2\sqrt{k}\Delta_{n,j} \ge \varepsilon_j/2\Big)$$
$$= \mathrm{P}\Big(\max_{0\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})-\mathbb{Z}_n^*(s,(i+1)\Delta_{n,j})| \ge \varepsilon_j/4\Big)$$
$$\le \mathrm{P}\Big(\max_{1\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+2}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})| \ge \varepsilon_j/8\Big)$$

for sufficiently large n. Hence, each summand in (6.5) can be bounded by

$$\mathrm{P}\Big(\sup_{0<x\le\delta 2^{-j+1}} |\mathbb{Z}_n^*(s,x)| \ge \varepsilon_j\Big) \le 2\mathrm{P}\Big(\max_{1\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+2}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})| \ge \varepsilon_j/8\Big).$$

Now, for any $1 \le i' < i \le \Delta_{n,j}^{-1}\delta 2^{-j+2}$, by (6.9) in Lemma 6.1 below,

$$\sup_{s\in[0,1]} \mathrm{E}|\mathbb{Z}_n^*(s,i\Delta_{n,j})-\mathbb{Z}_n^*(s,i'\Delta_{n,j})|^4 \lesssim \big(|i-i'|\Delta_{n,j}\big)^2 + \tfrac{r^2}{k}\big(|i-i'|\Delta_{n,j}\big).$$

Next, apply the main result in Móricz (1982), with $\gamma = 4, f(b,m) = m\Delta_{n,j}$ and $\varphi(t,m) = (m\Delta_{n,j} + r^2/k)^{1/4}$, which can be applied since

$$\mathbb{Z}_n^*(s,i\Delta_{n,j}) = \sum_{\ell=1}^{i} \mathbb{Z}_n^*(s,\ell\Delta_{n,j}) - \mathbb{Z}_n^*(s,(\ell-1)\Delta_{n,j}).$$

We obtain that

$$\sup_{s\in[0,1]} \mathrm{E}\Big[\max_{1\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+2}} |\mathbb{Z}_n^*(s,i\Delta_{n,j})|^4\Big]$$
$$\lesssim \delta 2^{-j+2}\Bigg\{\sum_{l=0}^{\lfloor \log_2(\lfloor \Delta_{n,j}^{-1}\delta 2^{-j+2}\rfloor)\rfloor-1} \Big(\delta 2^{-j-l+1} + \frac{r^2}{k}\Big)^{1/4}\Bigg\}^4$$
$$\le \delta 2^{-j+2}\Bigg\{\delta 2^{-j+1}\Big(\sum_{l=0}^{\infty} 2^{-l/4}\Big)^4 + \frac{r^2}{k}\log_2^4(\Delta_{n,j}^{-1}\delta 2^{-j+2})\Bigg\}$$
$$\lesssim \delta^2 2^{-2j} + \delta 2^{-j}\frac{r^2}{k}\log_2^4(\Delta_{n,j}^{-1}\delta 2^{-j+2}),$$

where we used subadditivity of $x \mapsto x^{1/4}$. Then, by Markov's inequality we obtain

$$\sup_{s\in[0,1]} \mathrm{P}\Big(\max_{1\le i\le \Delta_{n,j}^{-1}\delta 2^{-j+2}} |\mathbb{Z}_n^*(1,i\Delta_{n,j})| > \varepsilon_j/8\Big) \lesssim \varepsilon_j^{-4}\delta^2 2^{-2j}$$
$$+ \varepsilon_j^{-4}\delta 2^{-j}\frac{r^2}{k}\log_2^4(\Delta_{n,j}^{-1}\delta 2^{-j+2}).$$

Now, by the definition of $\varepsilon_j = \varepsilon(\delta 2^{-j})^{\mu}$,

$$\limsup_{n\to\infty}\sum_{j\in H_n} \varepsilon_j^{-4}\delta^2 2^{-2j} = \delta^{2-4\mu}\varepsilon^{-4}\limsup_{n\to\infty}\sum_{j\in H_n} 2^{-j(2-4\mu)} \lesssim \delta^{2-4\mu}\varepsilon^{-4},$$

which converges to zero for $\delta \downarrow 0$, since $\mu < 1/2$. Hence, in order to prove (6.5), it is sufficient to show that

$$\lim_{n\to\infty} \frac{r^2}{k} \sum_{j\in H_n} \varepsilon_j^{-4} \delta 2^{-j} \log_2^4(\Delta_{n,j}^{-1}\delta 2^{-j+2}) = 0 \tag{6.6}$$

for any $\delta > 0$.

For that purpose let $j_n = \max(H_n) = \max\{j \in \mathbb{N} : \sqrt{k}\delta 2^{-j+1} > \varepsilon_j/2\}$ be the maximum of H_n and note that

$$j_n = \lceil \log_2(4\delta^{1-\mu}\varepsilon^{-1}\sqrt{k})/(1-\mu)\rceil - 1 \le \log_2(4\delta^{1-\mu}\varepsilon^{-1}\sqrt{k})/(1-\mu).$$

Further, by the definition of $\Delta_{n,j} = \varepsilon_j/(8\sqrt{k}) = \varepsilon(\delta 2^{-j})^{\mu}/(8\sqrt{k})$,

$$\begin{aligned}
\log_2(\Delta_{n,j}^{-1}\delta 2^{-j+2}) &= \log_2(32\varepsilon^{-1}\delta^{1-\mu}\sqrt{k}2^{-j(1-\mu)}) \\
&= \log_2(32\varepsilon^{-1}\delta^{1-\mu}\sqrt{k}2^{(j_n-j)(1-\mu)}2^{-j_n(1-\mu)}) \\
&\le \log_2(32\varepsilon^{-1}\delta^{1-\mu}\sqrt{k}2^{(j_n-j)(1-\mu)}2^{-\log_2(4\delta^{1-\mu}\varepsilon^{-1}\sqrt{k})}) \\
&= \log_2(32\varepsilon^{-1}\delta^{1-\mu}\sqrt{k}2^{(j_n-j)(1-\mu)}(4\delta^{1-\mu}\varepsilon^{-1}\sqrt{k})^{-1}) \\
&= 3 + (j_n - j)(1-\mu)
\end{aligned}$$

for all $j \in H_n$. As a consequence, since $(a+b)^4 \lesssim a^4 + b^4$,

$$\begin{aligned}
\frac{r^2}{k}\sum_{j\in H_n} \varepsilon_j^{-4}\delta 2^{-j}\log_2^4(\Delta_{n,j}^{-1}\delta 2^{-j+2}) &\lesssim \frac{r^2}{k}\sum_{j\in H_n}\varepsilon_j^{-4}\delta 2^{-j}\{1 + (j_n - j)^4\} \\
&\lesssim \frac{r^2}{k}\sum_{j=1}^{j_n} 2^{-j(1-4\mu)}\{1 + (j_n - j)^4\}.
\end{aligned} \tag{6.7}$$

For $\mu \in [0, 1/4]$ we obtain that the last expression can be bounded by

$$\frac{r^2}{k} j_n^5 = O\Big(\frac{r^2}{k}\log_2^5(k)\Big).$$

which converges to zero by (2.2) from Condition (B3). Now consider $\mu \in (1/4, 1/2)$, such that $4\mu - 1 > 0$. Then, the sum on the right-hand side of (6.7) can be written as

$$\frac{r^2}{k}2^{(4\mu-1)j_n}\sum_{j=0}^{j_n-1} 2^{-(4\mu-1)j}(1 + j^4) \lesssim \frac{r^2}{k}2^{(4\mu-1)j_n} = O\Big(\frac{r^2}{k}k^{\frac{4\mu-1}{2(1-\mu)}}\Big) = O\Big(r^2 k^{\frac{6\mu-3}{2(1-\mu)}}\Big)$$

which converges to zero by assumption, eventually proving (6.6), and thus also (6.3).

Now, to prove (6.2), we invoke Ottaviani's inequality, see Proposition A.1.1 in van der Vaart and Wellner (1996) (note that $\mathbb{Z}_n^*$ is based on independent blocks). We obtain,

$$\mathrm{P}\Big(\sup_{s\in[0,1]}\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(s,x)|}{q(x)}\ge 2\varepsilon\Big)=\mathrm{P}\Big(\max_{j=1}^{m}\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(j/m,x)|}{q(x)}\ge 2\varepsilon\Big)$$
$$\le\frac{\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(1,x)|}{q(x)}\ge\varepsilon\Big)}{1-\max_{j=1}^{m}\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(1,x)-\mathbb{Z}_n^*(j/m,x)|}{q(x)}\ge\varepsilon\Big)}.$$

The denominator is bounded below by

$$1-\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(1,x)|}{q(x)}\ge\varepsilon/2\Big)-\max_{j=1}^{m}\mathrm{P}\Big(\sup_{0\le x\le\delta}\frac{|\mathbb{Z}_n^*(j/m,x)|}{q(x)}\ge\varepsilon/2\Big).$$

By (6.3), we may choose $\delta_0>0$ such that the limes inferior of this expression for $n\to\infty$ is larger than 1/2, for all $\delta<\delta_0$. Hence, (6.2) is an immediate consequence of (6.3).

Lemma 6.1 *Suppose the assumptions of Proposition 3.3 are met. Then, for all* $x,x'\in[0,1]$,

$$\sup_{s\in[0,1]}\mathrm{E}\{\mathbb{Z}_n^*(s,x)-\mathbb{Z}_n^*(s,x')\}^2\lesssim|x-x'|,\tag{6.8}$$

$$\sup_{s\in[0,1]}\mathrm{E}\{\mathbb{Z}_n^*(s,x)-\mathbb{Z}_n^*(s,x')\}^4\lesssim|x-x'|^2+\frac{r^2}{k}|x-x'|.\tag{6.9}$$

Furthermore, there exists $n_0\in\mathbb{N}$ *such that for all* $n\ge n_0,\ s\in[0,1],\delta\in[0,1]$, $0\le x\le x'\le x+\delta\le 1$,

$$|\mathbb{Z}_n^*(s,x)-\mathbb{Z}_n^*(s,x')|\le|\mathbb{Z}_n^*(s,x+\delta)-\mathbb{Z}_n^*(s,x)|+2\sqrt{k}\delta.\tag{6.10}$$

Proof Note that $\mathrm{E}|X-\mathrm{E}X|^\ell\lesssim\mathrm{E}|X|^\ell$ for any $\ell\ge 1$ and any random variable $X\in L^\ell$, whence it is sufficient to bound the non-centred increments of $\mathbb{Z}_n^*$. Recall the notation

$$G_j(x)=g_{j,n,x}(Y_{n,j})=\textstyle\sum_{i\in J_{n,j}}\mathbf{1}(X'_{n,i}>c_\infty-c(\frac{i}{n})x).$$

Without loss of generality, let $x\le x'$. Then, by Condition (B5),

$$\begin{aligned}\mathrm{E}\big|G_j(x)-G_j(x')\big|^2&=\mathrm{E}\Big\{\sum_{i\in J_{n,j}}\mathbf{1}(c_\infty-c(\tfrac{i}{n})x\ge X'_{n,i}>c_\infty-c(\tfrac{i}{n})x')\Big\}^2\\&\lesssim\frac{k}{n}(x'-x)\sum_{i\in J_{n,j}}c(\tfrac{i}{n})\lesssim\frac{rk}{n}(x'-x).\end{aligned}\tag{6.11}$$

This implies (6.8):

$$\mathrm{E}\{\mathbb{Z}_n^*(s,x)-\mathbb{Z}_n^*(s,x')\}^2\lesssim\frac{m}{k}\frac{rk}{n}|x-x'|\le|x-x'|,$$

where '$\lesssim$' is uniform in $s \in [0,1]$.

For the proof of (6.9) note that

$$\mathrm{E}\big|G_j(x) - G_j(x')\big|^4 \le r^2 \mathrm{E}\big|G_j(x) - G_j(x')\big|^2.$$

Then, again by inequality (6.11) and by Rosenthal's inequality, see, e.g., Ibragimov and Sharakhmetov (2001),

$$\begin{aligned}
&\sup_{s\in[0,1]} \mathrm{E}\{\mathbb{Z}_n^*(s,x) - \mathbb{Z}_n^*(s,x')\}^4 \\
&\lesssim \frac{1}{k^2}\left\{ \sum_{j=1}^{m} \mathrm{E}\big|G_j(x) - G_j(x')\big|^4 + \left(\sum_{j=1}^{m} \mathrm{E}\big|G_j(x) - G_j(x')\big|^2 \right)^2 \right\} \\
&\lesssim \frac{mr^3}{nk}|x - x'| + \frac{m^2 r^2}{n^2}|x - x'|^2,
\end{aligned}$$

which implies (6.9) since $m \le n/r$.

Finally, for sufficiently large n, and for $0 \le x < x' \le x + \delta \le 1$, we may write

$$\begin{aligned}
&\mathbb{Z}_n^*(s,x') - \mathbb{Z}_n^*(s,x) \\
&= \frac{1}{\sqrt{k}} \sum_{j=1}^{\lfloor sm \rfloor} \sum_{i\in J_{n,j}} \mathbf{1}(c_\infty - c(\tfrac{i}{n})x \ge X_{n,i}^* > c_\infty - c(\tfrac{i}{n})x') - \tfrac{k}{n} c(\tfrac{i}{n})(x' - x) \\
&\le \mathbb{Z}_n^*(s, x+\delta) - \mathbb{Z}_n^*(s,x) + \frac{1}{\sqrt{k}} \sum_{j=1}^{\lfloor sm \rfloor} \sum_{i\in J_{n,j}} \tfrac{k}{n} c(\tfrac{i}{n})(x + \delta - x') \\
&\le \mathbb{Z}_n^*(s, x+\delta) - \mathbb{Z}_n^*(s,x) + 2\frac{mrk}{n\sqrt{k}}\delta \\
&\le |\mathbb{Z}_n^*(s, x+\delta) - \mathbb{Z}_n^*(s,x)| + 2\sqrt{k}\delta.
\end{aligned}$$

A similar inequality from below proves (6.10).

Proof of Proposition 3.4 Let $\mu \in [0, 1/2)$ and $(s,x) \in [0,1] \times [0,L]$. In the proof of Proposition 6.2 in Bücher and Jennessen (2022) it was shown that under Condition (B6) there exists some $\tau > 0$ such that, for n large enough,

$$\mathbb{S}_n(s, x(1-\delta_n)) - \sqrt{k}\delta_n x C(s) \le \mathbb{F}_n(s,x) \le \mathbb{S}_n(s, x(1+\delta_n)) + \sqrt{k}\delta_n x C(s)$$

almost surely, where $\delta_n = \frac{\tau}{c_{\min}} A(\frac{n}{kL})$. Therefore,

$$\begin{aligned}
\left|\frac{\mathbb{F}_n(s,x)}{q(x)} - \frac{\mathbb{S}_n(s,x)}{q(x)}\right| &\le \left|\frac{\mathbb{S}_n(s,x(1+\delta_n))}{q(x)} - \frac{\mathbb{S}_n(s,x)}{q(x)}\right| \\
&\quad + \left|\frac{\mathbb{S}_n(s,x(1-\delta_n))}{q(x)} - \frac{\mathbb{S}_n(s,x)}{q(x)}\right| + 2\sqrt{k}\delta_n x^{1-\mu} C(s).
\end{aligned}$$

Note that $\sqrt{k}\delta_n = o(1)$ by assumption. We obtain that it suffices to show

$$\sup_{(s,x)\in[0,1]\times[0,L]} \Big| \frac{\mathbb{S}_n(s, x(1 \pm \delta_n) - \mathbb{S}_n(s,x)}{q(x)} \Big| = o_{\mathrm{P}}(1). \tag{6.12}$$

Let $\theta \in (0, L]$ and set $\varepsilon' = (L' - L)/L > 0$ such that $L' = (1 + \varepsilon')L$. Then, for n large enough such that $\delta_n < \varepsilon'$, the above term can be bounded by

$$\begin{aligned} &3 \sup_{(s,x)\in[0,1]\times[0,2\theta]} \Big| \frac{\mathbb{S}_n(s,x)}{q(x)} \Big| + \frac{1}{q(\theta)} \sup_{(s,x)\in[0,1]\times[\theta,L]} \Big| \mathbb{S}_n(s, x(1 \pm \delta_n)) - \mathbb{S}_n(s,x) \Big| \\ \le\ &3 \sup_{(s,x)\in[0,1]\times[0,2\theta]} \Big| \frac{\mathbb{S}_n(s,x)}{q(x)} \Big| + \frac{1}{q(\theta)} \sup_{\substack{(s,x),(s,y)\in[0,1]\times[\theta(1-\varepsilon'),L(1+\varepsilon')]\,: \\ |x-y|<L\delta_n}} \Big| \mathbb{S}_n(s,x) - \mathbb{S}_n(s,y) \Big|. \end{aligned}$$

The second summand is asymptotically negligible as shown in Eq. (8.19) in the proof of Proposition 6.2 in Bücher and Jennessen (2022). Regarding the first summand note that $\mathbb{S}_n$ can be substituted by $\mathbb{Z}_n^*$ from the proof of Proposition 3.3 according to Eq. (6.1) and the subsequent arguments. For the resulting term we have, for any $\varepsilon > 0$,

$$\lim_{\theta\downarrow 0} \limsup_{n\to\infty} \mathrm{P}\Big(\sup_{(s,x)\in[0,1]\times[0,\theta]} \Big| \frac{\mathbb{Z}_n^*(s,x)}{q(x)} \Big| > \varepsilon \Big) = 0$$

by Eq. (6.2). Together, we have shown (6.12).

Proof of Theorem 4.1 Let $\mu \in (0, 1/4]$ and $s_0 \in [2^{-\gamma}, 1)$. As in Example 5.1.5 in de Haan and Ferreira (2006) we can write $\sqrt{k}(\hat\gamma_n - \gamma) = A_n + B_n$, where

$$\begin{aligned} A_n &= \textstyle\int_{X_{n,n-k}/V(\frac{n}{k})}^{1} \sqrt{k}\, \frac{n}{k} \Big[1 - \hat F_n\big(sV(\tfrac{n}{k})\big) \Big] s^{-1}\, \mathrm{d}s, \\ B_n &= \textstyle\int_1^\infty \sqrt{k} \Big\{ \frac{n}{k} \Big[1 - \hat F_n\big(sV(\tfrac{n}{k})\big) \Big] - s^{-1/\gamma} \Big\} s^{-1}\, \mathrm{d}s, \end{aligned}$$

and $\hat F_n$ denotes the empirical cumulative distribution function of $X_1^{(n)}, \dots, X_n^{(n)}$.

First, note that

$$B_n = \int_1^\infty s^{-(1+\mu/\gamma)} s^{\mu/\gamma} \sqrt{k} \Big\{ \frac{n}{k} \Big[1 - \hat F_n\big(sV(\tfrac{n}{k})\big) \Big] - s^{-1/\gamma} \Big\}\, \mathrm{d}s$$

weakly converges to

$$B = \int_1^\infty s^{-1} \mathbb{S}(1, s^{-1/\gamma})\, \mathrm{d}s = \gamma \int_0^1 x^{-1} \mathbb{S}(1, x)\, \mathrm{d}x$$

by Lemma 6.2 and the Continuous Mapping Theorem. Regarding the term A_n, decompose $A_n = A_{n1} + A_{n2}$ with

$$\begin{aligned} A_{n1} &= \textstyle\int_{X_{n,n-k}/V(\frac{n}{k})}^{1} \sqrt{k}\, s^{-1/\gamma - 1}\, \mathrm{d}s, \\ A_{n2} &= \textstyle\int_{X_{n,n-k}/V(\frac{n}{k})}^{1} \sqrt{k}\, \Big\{ \frac{n}{k} \Big[1 - \hat F_n\big(sV(\tfrac{n}{k})\big) \Big] - s^{-1/\gamma} \Big\} s^{-1}\, \mathrm{d}s. \end{aligned}$$

We start with treating A_{n1}. Setting $\mu = 0$ and $s = t^{-\gamma}$ in Lemma 6.2 implies that $D_n(t) := \sqrt{k}(E_n(t) - t), t \le t_0 := s_0^{-1/\gamma}$, where

$$E_n(t) = \frac{1}{k} \sum_{j=1}^{n} \mathbf{1}\big(X_j^{(n)} > t^{-\gamma} V(\tfrac{n}{k})\big),$$

satisfies $\{D_n(t)\}_{t\in[0,t_0]} \rightsquigarrow \{\mathbb{S}(1,t)\}_{t\in[0,t_0]}$ in $(\ell^\infty([0,t_0]), \|\cdot\|_\infty)$. Note that $E_n^{-1}(t) = (X_{n,n-\lfloor kt\rfloor}/V(\frac{n}{k}))^{-1/\gamma}$ such that the functional delta-method applied to the inverse map (Theorem 3.9.4 in van der Vaart and Wellner, 1996) implies

$$\left\{ \sqrt{k}\left\{ \Big(\frac{X_{n,n-\lfloor kt\rfloor}}{V(\frac{n}{k})}\Big)^{-1/\gamma} - t \right\} + D_n(t) \right\}_{t\in[0,t_0]} = o_{\mathrm{P}}(1).$$

Thus, we obtain

$$A_{n1} = \gamma\sqrt{k}\left\{ \Big(\frac{X_{n,n-k}}{V(\frac{n}{k})}\Big)^{-1/\gamma} - 1 \right\} = -\gamma\, D_n(1) + o_{\mathrm{P}}(1) \rightsquigarrow -\gamma\, \mathbb{S}(1,1).$$

By the above convergence we also know that $X_{n,n-k}/V(\frac{n}{k}) \xrightarrow{\mathbb{P}} 1$ and along with Lemma 6.2 we immediately obtain $A_{n2} = o_{\mathrm{P}}(1)$.

Altogether, it follows that

$$\sqrt{k}(\hat{\gamma}_n - \gamma) \rightsquigarrow \gamma \int_0^1 x^{-1}\mathbb{S}(1,x)\,\mathrm{d}x - \gamma\, \mathbb{S}(1,1).$$

Since $\mathbb{S}$ is a centered Gaussian process the above limit is normally distributed with zero mean and variance (recall the definition of σ^2 in Proposition 3.1)

$$\begin{aligned}
\tau^2 = \gamma^2\Big[&\int_0^1\int_0^1 x^{-1}y^{-1}\mathrm{Cov}(\mathbb{S}(1,x),\mathbb{S}(1,y))\,\mathrm{d}x\,\mathrm{d}y \\
&- 2\int_0^1 x^{-1}\mathrm{Cov}(\mathbb{S}(1,x),\mathbb{S}(1,1))\,\mathrm{d}x + \mathrm{Var}(\mathbb{S}(1,1))\Big] \\
= \gamma^2\Big[&\int_0^1\int_0^1 x^{-1}y^{-1}\sigma^2(x,y)\,\mathrm{d}x\,\mathrm{d}y - 2\int_0^1 x^{-1}\sigma^2(x,1)\,\mathrm{d}x + \sigma^2(1,1)\Big].
\end{aligned}$$

By properties of the tail process $(Y_t)_t$ we further know $\sigma^2(ax,ay) = a\sigma^2(x,y)$ for $a > 0$. Straightforward calculations yield $\int_0^1\int_0^1 x^{-1}y^{-1}\sigma^2(x,y)\,\mathrm{d}x\mathrm{d}y = 2\int_0^1 x^{-1}\sigma^2(x,1)\,\mathrm{d}x$, implying $\tau^2 = \gamma^2\sigma^2(1,1)$.

Lemma 6.2 *Let $\mu \in [0,1/2)$ and $s_0 > 0$. Suppose that the assumptions from Proposition 3.4 are satisfied for some $L > \lceil s_0^{-1/\gamma}\rceil$, and that Condition (B7) holds. Then, as $n \to \infty$,*

$$\left\{ s^{\mu/\gamma}\sqrt{k}\left\{ \frac{n}{k}\left[1-\hat{F}_n\left(sV(\tfrac{n}{k})\right)\right] - s^{-1/\gamma} \right\} \right\}_{s\geq s_0} \rightsquigarrow \left\{ s^{\mu/\gamma}\ \mathbb{S}(1, s^{-1/\gamma}) \right\}_{s\geq s_0}$$

in $\left(\ell^\infty([s_0,\infty)), \|\cdot\|_\infty\right)$, *where* $\hat{F}_n$ *denotes the empirical cumulative distribution function of* $X_1^{(n)}, \dots, X_n^{(n)}$.

Proof Let

$$x_n(s) = \frac{n}{k}\left[1 - F\left(sV(\tfrac{n}{k})\right)\right], \quad s \in (0,\infty),$$

and note that

$$\mathbb{F}_n(1, x_n(s)) = \sqrt{k}\left\{ \frac{n}{k}\left[1-\hat{F}_n\left(sV(\tfrac{n}{k})\right)\right] - x_n(s) \right\}$$

which implies

$$s^{\mu/\gamma}\sqrt{k}\left\{ \frac{n}{k}\left[1-\hat{F}_n\left(sV(\tfrac{n}{k})\right)\right] - s^{-1/\gamma} \right\} = s^{\mu/\gamma}\mathbb{F}_n(1, x_n(s)) + \sqrt{k}(x_n(s) - s^{-1/\gamma})s^{\mu/\gamma}. \tag{6.13}$$

As shown in the proof of Corollary 3 in Einmahl et al. (2016) we have, under Condition (B7),

$$\sup_{s\geq s_0}\left| \frac{x_n(s) - s^{-1/\gamma}}{B(n/k)s^{-1/\gamma}} \right| = O(1),$$

such that

$$\sup_{s\geq s_0}\sqrt{k}\left|x_n(s) - s^{-1/\gamma}\right| s^{1/\gamma} = O(\sqrt{k}B(n/k)) = o(1), \tag{6.14}$$

by assumption. Consequently, by Eq. (6.13) it suffices to show that

$$\left\{ s^{\mu/\gamma}\mathbb{F}_n(1, x_n(s)) \right\}_{s\geq s_0} \rightsquigarrow \left\{ s^{\mu/\gamma}\mathbb{S}(1, s^{-1/\gamma}) \right\}_{s\geq s_0} \tag{6.15}$$

in $\left(\ell^\infty([s_0,\infty)), \|\cdot\|_\infty\right)$. To this, bound

$$\sup_{s\geq s_0}\left| s^{\mu/\gamma}\mathbb{F}_n(1, x_n(s)) - s^{\mu/\gamma}\mathbb{S}_n(1, s^{-1/\gamma}) \right| \leq A_n + B_n$$

where

$$A_n = \sup_{s\geq s_0}\left| \frac{\mathbb{F}_n(1, x_n(s)) - \mathbb{S}_n(1, x_n(s))}{q(s^{-1/\gamma})} \right|, \quad B_n = \sup_{s\geq s_0}\left| \frac{\mathbb{S}_n(1, x_n(s)) - \mathbb{S}_n(1, s^{-1/\gamma})}{q(s^{-1/\gamma})} \right|.$$

We will prove that both terms are asymptotically negligible such that Proposition 3.3 implies the assertion in (6.15).

First, note that by Eq. (6.14) we know that for any $\varepsilon > 0$ and n large enough, $\sup_{s\geq s_0} |x_n(s) - s^{-1/\gamma}| < \delta_n := k^{-1/2}\varepsilon$. We may then show $B_n = o_{\mathrm{P}}(1)$ similar to the proof of Eq. (6.12). Further, for sufficiently large n,

$$A_n \leq \sup_{s\geq s_0} |s^{1/\gamma} x_n(s)|^{\mu} \sup_{x\in[0,L]} \left| \frac{\mathbb{F}_n(1,x) - \mathbb{S}_n(1,x)}{q(x)} \right|$$

where the first supremum is bounded by Eq. (6.14) and the second supremum is asymptotically negligible according to Proposition 3.4, yielding $A_n = o_{\mathrm{P}}(1)$.

Acknowledgements The authors are grateful to Chen Zhou for helpful discussions on early stages of this project. The authors are also grateful to an unknown referee, an associate editor and an editor for their constructive comments on an earlier version of this article which helped to improve the presentation.

Author contribution The proofs of Proposition 3.3 and Lemma 6.1 were initially worked out by AB, and later improved by both authors. All other mathematical results are due to TJ, who is also responsible for the simulation study. The first draft of the manuscript was written by TJ, which was later improved by both authors.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability The simulation study was performed in R. The code is available upon request from the authors.

Declarations

Ethical approval Not applicable.

Conflict of interest The authors declare that they have no conflict of interest.

References

Basrak, B., Segers, J.: Regularly varying multivariate time series. Stoch. Process. Appl. **119**(4), 1055–1080 (2009)

Billingsley, P.: Probability and measure (Third ed.). Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., New York. A Wiley-Interscience Publication (1995)

Bücher, A., Jennessen, T.: Statistics for heteroscedastic time series extremes. Preprint at http://arxiv.org/abs/2204.09534. To appear in *Bernoulli* (2022)

de Haan, L., Ferreira, A.: Extreme value theory: An introduction. Springer (2006)

Dehling, H., Durieu, O., Volny, D.: New techniques for empirical processes of dependent data. Stoch. Process. Appl. **119**(10), 3699–3718 (2009)

Drees, H.: Weighted approximations of tail processes for β-mixing random variables. Ann. Appl. Probab. **10**(4), 1274–1301 (2000)

Drees, H., Rootzén, H.: Limit theorems for empirical processes of cluster functionals. Ann. Statist. **38**(4), 2145–2186 (2010)

Drees, H., Segers, J., Warchoł, M.: Statistics for tail processes of markov chains. Extremes **18**(3), 369–402 (2015)

Einmahl, J.H.J., de Haan, L., Zhou, C.: Statistics of heteroscedastic extremes. J. R. Stat. Soc. Ser. B. Stat. Methodol. **78**(1), 31–51 (2016)

Ferreira, M.: On the extremal behavior of a Pareto process: an alternative for ARMAX modeling. Kybernetika (Prague) **48**(1), 31–49 (2012)

Hall, P., Peng, L., Yao, Q.: Moving-maximum models for extrema of time series. J. Statist. Plann. Inference **103**(1–2), 51–63 (2002)

Ibragimov, R., Sharakhmetov, S.: The exact constant in the Rosenthal inequality for random variables with mean zero. Teor. Veroyatnost. i Primenen. **46**(1), 134–138 (2001)

Kulik, R., Soulier, P.: Heavy-tailed time series. Springer series in operations research and financial engineering, Springer, New York (2020)

Móricz, F.: A general moment inequality for the maximum of partial sums of single series. Acta Sci. Math. (Szeged) **44**(1–2), 67–75 (1982)

Rootzén, H.: Weak convergence of the tail empirical process for dependent sequences. Stoch. Process. Appl. **119**(2), 468–490 (2009)

Segers, J.: Functionals of clusters of extremes. Adv. Appl. Probab. **35**(4), 1028–1045 (2003)

Shao, Q.-M., Yu, H.: Weak convergence for weighted empirical processes of dependent sequences. Ann. Probab. **24**(4), 2098–2127 (1996)

van der Vaart, A.W., Wellner, J.A.: Weak convergence and empirical processes - Springer series in statistics. Springer, New York (1996)

Zhang, Z., Smith, R.L.: Modeling financial time series data as moving maxima processes. Manuscript, UNC (2001)

Zhang, Z., Smith, R.L.: On the estimation and application of max-stable processes. J. Statist. Plann. Inference **140**(5), 1135–1153 (2010)