Bernoulli **27**(2), 2021, 1409–1439
https://doi.org/10.3150/20-BEJ1279

Bootstrapping Hill estimator and tail array sums for regularly varying time series

CARSTEN JENTSCH[1] and RAFAŁ KULIK[2]

[1]*Faculty of Statistics, TU Dortmund University, D-44221 Dortmund, Germany*
[2]*Department of Mathematics and Statistics, University of Ottawa. E-mail: rkulik@uottawa.ca*

In the extreme value analysis of stationary regularly varying time series, tail array sums form a broad class of statistics suitable to analyze their extremal behavior. This class includes for example, the Hill estimator or estimators of the extremogram and the tail dependence coefficient.

General asymptotic theory for tail array sums has been developed by Rootzén et al. (*Ann. Appl. Probab.* **8** (1998) 868–885) under mixing conditions and in Kulik et al. (*Stochastic Process. Appl.* **129** (2019) 4209–4238) for functions of geometrically ergodic Markov chains. A more general framework of cluster functionals is presented in Drees and Rootzén (*Ann. Statist.* **38** (2010) 2145–2186).

However, the resulting limiting distributions turn out to be very complex and cumbersome to estimate as they usually depend on the whole extremal dependence structure of the time series. Hence, a suitable bootstrap procedure is desired, but available bootstrap consistency results for tail array sums are scarce. In this paper, following Drees (Drees (2015)), we consider a multiplier block bootstrap to estimate the limiting distribution of tail array sums. We prove that, conditionally on the data, an *appropriately* constructed multiplier block bootstrap statistic converges to the correct limiting distribution. Interestingly, in contrast, it turns out that an apparently natural, but naïve application of the multiplier block bootstrap scheme does not yield the correct limit.

In simulations, we provide numerical evidence of our theoretical findings and illustrate the superiority of the proposed multiplier block bootstrap over some obvious competitors. The proposed bootstrap scheme proves to be computationally efficient in comparison to other approaches.

Keywords: Heavy tails; Hill estimator; multiplier bootstrap; regular variation; stationary time series; tail empirical process; tail array sums

1. Introduction

The common framework that allows meaningful asymptotic theory to study the extremal behavior of stationary time series is based on the concept of regular variation. Throughout this paper, let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying univariate time series with marginal distribution function $F(x) = \mathbb{P}(X_0 \le x)$, tail function $\bar{F}(x) = \mathbb{P}(X_0 > x) = 1 - F(x)$ and tail index $\alpha > 0$. This means that for each integer $h \ge 0$, there exists a non-zero Radon measure $\boldsymbol{\nu}_{0,h}$ on $\overline{\mathbb{R}}^{h+1} \setminus \{\mathbf{0}\}$ such that $\boldsymbol{\nu}_{0,h}(\overline{\mathbb{R}}^{h+1} \setminus \mathbb{R}^{h+1}) = 0$ and

$$\lim_{x\to\infty} \frac{\mathbb{P}((X_0, \dots, X_h) \in xA)}{\mathbb{P}(X_0 > x)} = \boldsymbol{\nu}_{0,h}(A), \tag{1.1}$$

for all relatively compact sets $A \subset \overline{\mathbb{R}}^{h+1} \setminus \{\mathbf{0}\}$ satisfying $\boldsymbol{\nu}_{0,h}(\partial A) = 0$. The measure $\boldsymbol{\nu}_{0,h}$, called the exponent measure of $(X_0, \dots, X_h)$, is homogeneous with index $-\alpha$, $\alpha > 0$, that is, $\boldsymbol{\nu}_{0,h}(tA) = t^{-\alpha}\boldsymbol{\nu}_{0,h}(A)$. The choice of the denominator $\mathbb{P}(X_0 > x)$ entails that $\boldsymbol{\nu}_{0,h}((1,\infty) \times \mathbb{R}^h) = 1$ that is, that the right tail of the stationary distribution is not trivial and that X_0 satisfies the so-called balanced tail condition.

1350-7265

1.1. Tail array sums: Random vs. deterministic levels

Tail array sums form a broad class of statistics suitable to analyze the extremal behavior of stationary, regularly varying time series. This class includes the Hill estimator for the tail index α and estimators of the tail dependence coefficient. Based on the concept of regular variation, such (population) quantities, can be represented in a unified manner. For $i \le j$ we denote $\boldsymbol{x}_{i,j} = (x_i, \dots, x_j)$ and let $\phi : \mathbb{R}^{h+1} \to \mathbb{R}$ be such that $\phi(x_0, \dots, x_h) = 0$ whenever $\max\{|x_0|, \dots, |x_h|\} < \epsilon$ for some $\epsilon > 0$. Then, the general goal is to estimate the quantity

$$\boldsymbol{\nu}_{0,h}(\phi) := \lim_{x\to\infty} \frac{\mathbb{E}[\phi(\boldsymbol{X}_{0,h}/x)]}{\bar{F}(x)} = \int_{\mathbb{R}^{h+1}} \phi(\boldsymbol{x}_{0,h})\boldsymbol{\nu}_{0,h}(\mathrm{d}\boldsymbol{x}_{0,h}). \tag{1.2}$$

The above limit exists and is finite due to regular variation, as long as the appropriate moment condition holds (cf. the definition of the space $\mathcal{L}_q$ below). Let $X_{n:1} \le \dots \le X_{n:n}$ be the order statistics from the sample $X_1, \dots, X_n$ and $k = k_n \to \infty$ such that $k/n \to 0$ as $n \to \infty$. Then, to estimate the quantity $\boldsymbol{\nu}_{0,h}(\phi)$, we use the estimator

$$\widehat{M}_n(\phi) := \frac{1}{k}\sum_{j=1}^{n-h} \phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right) = \frac{1}{k}\sum_{j=1}^{n-h} \phi\left(\frac{(X_j, \dots, X_{j+h})}{X_{n:n-k}}\right). \tag{1.3}$$

We will call $\widehat{M}_n(\phi)$ a *tail array sum with random level*. Further, let us denote its associated process by

$$\widehat{\mathbb{M}}_n(\phi) = \sqrt{k}\{\widehat{M}_n(\phi) - \boldsymbol{\nu}_{0,h}(\phi)\} = \sqrt{k}\left\{\frac{1}{k}\sum_{j=1}^{n-h} \phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right) - \boldsymbol{\nu}_{0,h}(\phi)\right\}. \tag{1.4}$$

The above representation of $\widehat{M}_n(\phi)$ is very flexible and allows a unified treatment of many practically relevant statistics from extreme value statistics. By choosing different functions ϕ, different estimation problems can be addressed.

Example 1 (Special cases of tail array sums).

(i) Take $h = 0$ and $\phi(\boldsymbol{x}_{0,h}) = \log(x_0)\mathbb{1}\{x_0 > 1\}$, then $\boldsymbol{\nu}_{0,h}(\phi) = 1/\alpha$ and $\widehat{M}_n(\phi)$ is the aforementioned Hill estimator.

(ii) Take $h \ge 1$ and $\phi(\boldsymbol{x}_{0,h}) = \mathbb{1}\{x_0 > 1, x_h > 1\}$. Then $\boldsymbol{\nu}_{0,h}(\phi) = \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x)$, which is the (upper) tail dependence coefficient between X_0 and X_h. In turn, this is a special case of the extremogram considered in Davis and Mikosch [4], and Drees [9].

(iii) Take $\phi(\boldsymbol{x}_{0,h}) = x_h\mathbb{1}\{x_0 > 1\}$. Then $\boldsymbol{\nu}_{0,h}(\phi) = \lim_{x\to\infty} \mathbb{E}[(X_h/x) \mid X_0 > x]$. We note that for $h = 0$ the quantity $\mathbb{E}[X_0 \mid X_0 > x]$ for $x = F^{\leftarrow}(1-p)$, where $F^{\leftarrow}$ is the generalized inverse of F, is known in the risk management literature as the expected shortfall.

(iv) Take $h \ge 1$ and $\phi(\boldsymbol{x}_{0,h}) = \mathbb{1}\{|x_0| > 1, x_h \le y\}$. Then $\boldsymbol{\nu}_{0,h}(\phi) = p^{-1}\lim_{x\to\infty} \mathbb{P}(X_h \le xy \mid |X_0| > x) = p^{-1}\mathbb{P}(Y_h \le y)$, where $p = \lim_{x\to\infty} \mathbb{P}(X_0 > x)/\mathbb{P}(|X_0| > x)$ and $\{Y_j, j \in \mathbb{Z}\}$ is the tail process.

(v) Take $h \ge 1$ and $\phi(\boldsymbol{x}_{0,h}) = \mathbb{1}\{|x_0| > 1, x_h \le y|x_0|\}$. Then $\boldsymbol{\nu}_{0,h}(\phi) = p^{-1}\lim_{x\to\infty} \mathbb{P}(X_h \le y|X_0| \mid |X_0| > x) = p^{-1}\mathbb{P}(\Theta_h \le y)$, where $\{\Theta_j, j \in \mathbb{Z}\}$ is the tail spectral process. The problem of estimation of the tail spectral process was tackled in Drees et al. [12], Davis et al. [3].

Limiting theory for tail array sums can be concluded, under some weak dependence conditions, from Kulik et al. [18], utilizing the general theory from Drees and Rootzén [11]. The particular case of the extremogram was considered in Davis and Mikosch [4] and Drees et al. [12]. To be able to derive asymptotic theory for tail array sums with random levels $\widehat{M}_n(\phi)$ one needs to consider first the related quantity

$$\widetilde{M}_n(\phi) := \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^{n-h} \phi\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right), \tag{1.5}$$

where $\{u_n\}$ is a deterministic sequence with $u_n \to \infty$ as $n \to \infty$ such that $n\bar{F}(u_n) \to \infty$. We call $\widetilde{M}_n(\phi)$ a *tail array sum with deterministic level*. Such tail array sums were considered in Rootzén et al. [20]. Note that this version of a tail array sum requires the knowledge of $\bar{F}$ and the choice of the deterministic sequence $\{u_n\}$, whereas $\widehat{M}_n(\phi)$ only requires the choice of k.

It is crucial to compare the limiting distribution of two versions $\widehat{M}_n(\phi)$ and $\widetilde{M}_n(\phi)$. First, we point out that the limiting distributions for the estimator $\widehat{M}_n(\phi)$ defined in (1.3) with the random level $X_{n:n-k}$ and the estimator $\widetilde{M}_n(\phi)$ defined in (1.5) with deterministic level u_n typically differ. Second, the limiting distribution of $\widehat{M}_n(\phi)$ is concluded from that of (1.5) by considering functional convergence of $\widetilde{M}_n(\phi_s)$, where $\phi_s(\boldsymbol{x}_{0,h}) = \phi(\boldsymbol{x}_{0,h}/s)$, $0 < s_0 < s < t_0 < \infty$.

1.2. Bootstrapping for tail array sums: A naïve idea

Although a general asymptotic theory has been established under different weak dependence conditions, such theoretical limiting results for tail array sums $\widehat{M}_n(\phi)$ are not directly applicable for the construction of confidence intervals. This is because the limiting variances usually have a very complicated form that depends on the whole extremal dependence structure of the time series. To overcome this issue, a suitable bootstrap procedure is required to estimate the limiting distribution of the tail array sums. Although, for stationary regularly varying time series, there are some theoretical results on (multiplier) bootstrapping for *infeasible* statistics of the form (1.5) (cf. Davis et al. [5], Davis et al. [3]), to the best of our knowledge there are no results on bootstrapping the *feasible* and hence practically more relevant estimators of the form (1.3) except for the sample extremogram Drees [9] and for the Hill estimator and spectral tail processes Drees and Knezevic [10].

Hence, the main goal of this paper is to propose a suitable bootstrap scheme for the feasible general tail array sums $\widehat{M}_n(\phi)$ defined in (1.3) and prove its consistency for stationary, regularly varying time series under mild regularity conditions. As used already by Drees [9] for the extremogram, we employ a multiplier block bootstrap for the broad class of tail array sums in this paper and derive corresponding bootstrap consistency results. Precisely, let $\{\xi_j, j \in \mathbb{Z}\}$ be a sequence of i.i.d. centered random variables with unit variance, independent of the sequence $\{X_j, j \in \mathbb{Z}\}$. Let r_n be the block length such that $r_n \to \infty$ and define $m_n = (n-h)/r_n$. Without loss of generality, we assume that m_n is integer-valued. Let $k = k_n$ be a sequence of integers such that $k \to \infty$, $k/n \to 0$ and define u_n by $k = n\bar{F}(u_n)$. Then we can rewrite $\widehat{M}_n(\phi)$ as follows

$$\widehat{M}_n(\phi) = \frac{1}{k} \sum_{i=1}^{m_n} \sum_{j=(i-1)r_n+1}^{ir_n} \phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right) = \frac{1}{k} \sum_{i=1}^{m_n} \widehat{\Phi}_{n,i}, \tag{1.6}$$

with an obvious notation for $\widehat{\Phi}_{n,i}$, $i = 1, \dots, m_n$. Consequently, a natural, but naïve multiplier bootstrap version of the associated process $\widehat{\mathbb{M}}_n(\phi)$ defined in (1.4) is

$$\widehat{\mathbb{M}}_{n,\xi}(\phi) = \sqrt{k}\left(\frac{1}{k} \sum_{i=1}^{m_n} (1+\xi_i)\widehat{\Phi}_{n,i} - \frac{1}{k} \sum_{i=1}^{m_n} \widehat{\Phi}_{n,i}\right) = \frac{1}{\sqrt{k}} \sum_{i=1}^{m_n} \xi_i \widehat{\Phi}_{n,i}. \tag{1.7}$$

However, as we will argue below, such a multiplier bootstrap turns out to be not consistent for the limiting distribution of the tail array sums $\widehat{M}_n(\phi)$ as it does not lead to the correct limiting distribution.

To understand how to construct a suitable multiplier bootstrap procedure for tail array sums $\widehat{M}_n(\phi)$ that leads to the correct limiting distribution and is asymptotically consistent, it is crucial to consider its relationship to $\widetilde{M}_n(\phi)$ in more detail. In the following, we shed some light on this relationship for a more simplified special case of tail array sums. Precisely, we consider the Hill estimator from Example 1(i) to estimate the tail index.

1.3. The Hill estimator - a deep but simple example

Estimation of the tail index α of a regularly varying distribution is one of the most important problems in statistical extreme value analysis and the Hill estimator is a commonly used method to do so. A distribution P of a random variable X is said to have (right) tail index α if

$$\boldsymbol{\nu}_{0,0}\big((u,\infty)\big) = \lim_{x\to\infty} \frac{P(X > ux)}{P(X > x)} = u^{-\alpha}. \tag{1.8}$$

Note that (1.8) is obtained from the definition of regular variation in (1.1) by setting $A = (u,\infty)$ and $h = 0$.

In the literature, two practically relevant variants of the Hill estimator have been defined. Additionally, we consider another practically useless variant of the Hill estimator for academic reasons to understand why and how the naïve application of a multiplier block bootstrap in (1.7) fails and how this can be cured. For the first variant let, as before, $u_n \to \infty$ and $n\bar{F}(u_n) \to \infty$ as $n \to \infty$ be a deterministic sequence. Then, the Hill estimator for $\gamma := 1/\alpha$ with *deterministic levels and random normalization* is defined by

$$\widehat{\gamma}_{u_n} := \frac{1}{\sum_{j=1}^n \mathbb{1}\{X_j > u_n\}} \sum_{j=1}^n \log\left(\frac{X_j}{u_n}\right) \mathbb{1}\{X_j > u_n\}. \tag{1.9}$$

Alternatively, recalling that $k = n\bar{F}(u_n)$, replacing u_n in (1.9) by the order statistic $X_{n:n-k}$, which gives (assuming for simplicity that there are no ties in the sequence $\sum_{j=1}^n \mathbb{1}\{X_j > X_{n:n-k}\} = k$), leads to the second variant of the Hill estimator with *random levels and deterministic normalization* defined by

$$\widehat{\gamma}_k := \frac{1}{\sum_{j=1}^n \mathbb{1}\{X_j > X_{n:n-k}\}} \sum_{j=1}^n \log\left(\frac{X_j}{X_{n:n-k}}\right) \mathbb{1}\{X_j > X_{n:n-k}\} \tag{1.10}$$

$$= \frac{1}{k} \sum_{j=0}^{k-1} \log\left(\frac{X_{n:n-j}}{X_{n:n-k}}\right). \tag{1.11}$$

The third variant is the Hill estimator with *deterministic levels and deterministic normalization* defined by

$$\widetilde{\gamma}_{u_n} := \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^n \log\left(\frac{X_j}{u_n}\right) 1\{X_j > u_n\}. \tag{1.12}$$

The latter is a mixture of the other two variants and is defined only for academic reasons. By setting $h = 0$ and $\phi(\boldsymbol{x}_{0,h}) = \log(x_0)\mathbb{1}\{x_0 > 1\}$, the second variant $\widehat{\gamma}_k$ of the Hill estimator in (1.11) and its third variant $\widetilde{\gamma}_{u_n}$ in (1.12) are obtained as special cases from (1.3) and (1.5), respectively. Note again that the first and the second variants are practically relevant, as they require only the choice of the deterministic sequence u_n for $\widehat{\gamma}_{u_n}$ or the choice of k for $\widehat{\gamma}_k$, whereas the third variant requires not only the choice of u_n, but also the knowledge of $\bar{F}$.

The asymptotic theory for the Hill estimators is well-known; see, for example, de Haan and Ferreira [6] in the i.i.d. case or Drees [7] in the weakly dependent case. Under appropriate conditions, the limiting distribution is centered normal with a variance given by a complicated infinite series that depends on the whole extremal dependence structure of the time series. To approximate the limiting distribution by using a suitable resampling scheme, Drees and Rootzén [11], Example 3.4, propose a block bootstrap procedure for the Hill estimator with deterministic levels and random normalization (that is, $\widehat{\gamma}_{u_n}$) and prove its consistency.

In the i.i.d. case, the complicated limiting variances of all variants of the Hill estimator simplify and allow for rather straightforward asymptotic inference. Precisely, under suitable regularity conditions, we have

$$\sqrt{k}(\widehat{\gamma}_{u_n} - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2) \quad \text{and} \quad \sqrt{k}(\widehat{\gamma}_k - \gamma) \xrightarrow{d} \mathcal{N}(0, \gamma^2), \tag{1.13}$$

but

$$\sqrt{k}(\widetilde{\gamma}_{u_n} - \gamma) \xrightarrow{d} \mathcal{N}(0, 2\gamma^2), \tag{1.14}$$

such that the limiting variances of the first two practically relevant variants of the Hill estimator share the *same limiting distribution*, but this *differs* from the limiting distribution of the third variant by a factor of 2. A comparison of (1.9) and (1.11) reveals that randomness in (1.11) is contained only in the summands, but in (1.9) randomness is contained in these summands *and* in the summands in the denominator. Nevertheless, for both versions, we get the same limiting distributions in (1.13). This important observation will help us to construct a proper version of a mulitplier block bootstrap for tail array sums with random levels in Section 3.

To the best of our knowledge and to our surprise, there are very limited results on bootstrapping the Hill estimator with random levels $\widehat{\gamma}_k$. In the i.i.d. case we are aware of two results. The first one is Loukrati [19], Theorem 5.3.1. There, the author establishes convergence of bootstrapped tail empirical processes and obtains a result for the Hill estimator $\widehat{\gamma}_k$ with help of integral functionals. The second one is Theorem 3.3 in Groen [15], where the author approximates bootstrapped order statistics by appropriately constructed sequences of Brownian motions and concludes a result for the Hill estimator $\widehat{\gamma}_k$. Proofs of these results rely on the i.i.d. structure of the underlying sequence. In Drees and Knezevic [10], Theorem 2.3, the authors consider multiplier bootstrap for the Hill estimator with deterministic threshold and random normalization, corresponding to (1.9).

As foreshadowed in Section 1.2, a natural, but naïve approach of a multiplier bootstrap (actually a common wild bootstrap after setting $r_n = 1$ for the i.i.d. case), will generally fail to replicate the limiting distribution of tail array sums with random levels. Precisely, for the Hill estimator in the i.i.d. case, the naïve multiplier bootstrap process from (1.7) gives

$$\sqrt{k}(\widehat{\gamma}_{k,\xi} - \widehat{\gamma}_k) := \sqrt{k}\left(\frac{1}{k}\sum_{j=0}^{k-1}(1+\xi_j)\log\left(\frac{X_{n:n-j}}{X_{n:n-k}}\right) - \frac{1}{k}\sum_{j=0}^{k-1}\log\left(\frac{X_{n:n-j}}{X_{n:n-k}}\right)\right) \tag{1.15}$$

$$= \frac{1}{\sqrt{k}}\sum_{j=0}^{k-1}\xi_j \log\left(\frac{X_{n:n-j}}{X_{n:n-k}}\right). \tag{1.16}$$

As a special case of our general theory for multiplier bootstrapping of tail array sums in Section 2 of this paper, we will conclude that

$$\sqrt{k}(\widehat{\gamma}_{k,\xi} - \widehat{\gamma}_k) \xrightarrow{d} \mathcal{N}\big(0, 2\gamma^2\big) \quad \text{in probability.}$$

Two crucial findings are in order. First, the naive multiplier bootstrap does not replicate the distribution of $\widehat{\gamma}_k$ as desired, but that of $\widetilde{\gamma}_{u_n}$. Second, as the limiting distributions of $\widehat{\gamma}_{u_n}$ and of $\widehat{\gamma}_k$ coincide, the idea is to construct a suitable multiplier bootstrap that aims at replicating the random structure of $\widehat{\gamma}_{u_n}$ to properly replicate the limiting distribution of $\widehat{\gamma}_k$. It turns out that a remedy to cure the bootstrap inconsistency is to construct a multiplier bootstrap that additionally incorporates a proper randomization of k in the denominator of (1.11).

1.4. Outline of the paper

The paper is structured as follows. In Section 2, we quote from Kulik et al. [18] and gather all relevant asymptotic results for $\widetilde{M}_n(\phi)$ and $\widehat{M}_n(\phi)$. These results are a benchmark for the multiplier block bootstrap procedures with and without randomizing k, as considered in this paper. In Section 3, we state our main result on the multiplier bootstrap consistency in Theorem 3.1. Its proof follows in principle the lines of Drees [9], with appropriate modifications by incorporating some techniques developed in Kulik et al. [18]. In Section 4, we perform extensive simulations studies. There, we confirm our theoretical findings that the randomization of k is indeed necessary. Furthermore, we conclude that the properly implemented multiplier bootstrap outperforms i.i.d. or block bootstraps, respectively. Additionally, the multiplier bootstrap is computationally considerably less demanding than i.i.d. or block bootstraps as it avoids the sorting step for the bootstrap samples.

2. Convergence of the tail array sums

In this section, we gather all relevant limiting results for tail array sums. For simplicity of the presentation, we focus on time series that can be expressed as regularly varying functions of Markov chains as proposed by Kulik et al. [18]. For this broad class of time series, the conditions for validity of the central limit theorem for the estimators (1.3) and (1.5) are rather straightforward as proven in Kulik et al. [18]. However, corresponding limiting results can be established also under different sets of assumptions as for example, those used in Drees and Rootzén [11]. We point out that our results on the multiplier bootstrap discussed in Section 3 are extendable to a more general class, as long as the appropriate non-bootstrap functional central limit theorem holds; see Drees and Rootzén [11]. That functional CLT requires existence of the limiting covariance along with additional moment assumptions, among others. As the approach based on Markov chains allows a formulation in terms of simple and easy to verify conditions, we prefer to use this approach in the following and throughout the paper.

The set-up is as in Kulik et al. [18]. We assume that $\{X_j, j \in \mathbb{N}\}$ is a function of a stationary Markov chain $\{\mathbb{Y}_j, j \in \mathbb{N}\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in a measurable space $(E, \mathcal{E})$. That is, there exists a measurable real valued function $g : E \to \mathbb{R}$ such that $X_j = g(\mathbb{Y}_j)$.

Assumption 1.

(i) The Markov chain $\{\mathbb{Y}_j, j \in \mathbb{Z}\}$ is strictly stationary under $\mathbb{P}$.
(ii) The sequence $\{X_j = g(\mathbb{Y}_j), j \in \mathbb{Z}\}$ is regularly varying with tail index $\alpha > 0$.

(iii) There exist a measurable function $V : E \to [1, \infty)$, $\gamma \in (0, 1)$ and $b > 0$ such that for all $y \in E$,

$$\mathbb{E}\big[V(\mathbb{Y}_1) \mid \mathbb{Y}_0 = y\big] \le \gamma V(y) + b. \tag{2.1}$$

(iv) There exist an integer $m \ge 1$ and $x_0 \ge 1$ such that for all $x \ge x_0$, there exists a probability measure ν on $(E, \mathcal{E})$ and $\epsilon > 0$ such that, for all $y \in \{V \le x\}$ and all measurable sets $B \in \mathcal{E}$,

$$\mathbb{P}(\mathbb{Y}_m \in B \mid \mathbb{Y}_0 = y) \ge \epsilon \nu(B). \tag{2.2}$$

(v) There exist $q_0 \in (0, \alpha)$ and a constant $c > 0$ such that

$$|g|^{q_0} \le cV. \tag{2.3}$$

(vi) For every $s > 0$,

$$\limsup_{n\to\infty} \frac{1}{u_n^{q_0} \bar{F}(u_n)} \mathbb{E}\big[V(\mathbb{Y}_0)\mathbb{1}\big\{\big|g(\mathbb{Y}_0)\big| > u_n s\big\}\big] < \infty. \tag{2.4}$$

Various important examples of Markov chain time series that fulfill Assumption 1 can be found in Kulik et al. [18]. These include for example, ARMA(p, q) processes, solutions to stochastic recurrence equations or different threshold models.

Let $|\cdot|$ be an arbitrary norm on $\mathbb{R}^{h+1}$. Throughout the paper, we will write $\boldsymbol{x}_{a,b}$ for $(x_a, \dots, x_b)$, $a \le b \in \mathbb{Z}$, for any sequence $\boldsymbol{x} = \{x_j, j \in \mathbb{Z}\}$. For $q \ge 0$, let $\mathcal{L}_q$ be the space of measurable real-valued functions ϕ defined on $\mathbb{R}^{h+1}$ such that

(i) there exists a constant $\epsilon > 0$ such that $|\phi(\boldsymbol{x})| \le \epsilon^{-1}(|\boldsymbol{x}|^q \vee 1)\mathbb{1}\{|\boldsymbol{x}| > \epsilon\}$ for $\boldsymbol{x} \in \mathbb{R}^{h+1}$;
(ii) for all $j \ge 0$, the function $\boldsymbol{x}_{0,j+h} \mapsto \phi(\boldsymbol{x}_{j,j+h})$ is almost surely continuous with respect to $\boldsymbol{\nu}_{0,j+h}$.

Clearly, the case $q = 0$ corresponds to bounded functions. The purpose of the bound in (i) is two-fold. First, it will guarantee existence of moments, second, it implies that the function vanishes around zero.

If Assumption 1 holds and $\phi, \phi' \in \mathcal{L}_{q_0/2}$ for $q_0 \in (0, \alpha)$, then by Kulik et al. [18], Lemma 2.2, the following quantities are well-defined:

$$\sigma^2(\phi) = \int_{\mathbb{R}^{h+1}} \phi^2(\boldsymbol{x})\boldsymbol{\nu}_{0,h}(\mathrm{d}\boldsymbol{x}) + 2\sum_{j=1}^{\infty} \int_{\mathbb{R}^{j+h+1}} \phi(\boldsymbol{x}_{0,h})\phi(\boldsymbol{x}_{j,j+h})\boldsymbol{\nu}_{0,j+h}(\mathrm{d}\boldsymbol{x}), \tag{2.5}$$

$$C(\phi, \phi') = \frac{1}{2}\big\{\sigma^2(\phi + \phi') - \sigma^2(\phi) - \sigma^2(\phi')\big\}. \tag{2.6}$$

Define

$$\mathbb{M}_n(\phi) = \sqrt{n\bar{F}(u_n)}\big\{\widetilde{M}_n(\phi) - \mathbb{E}\big[\widetilde{M}_n(\phi)\big]\big\}. \tag{2.7}$$

Let $\mathbb{M}$ be a Gaussian process indexed by $\mathcal{L}_q$ with covariance function C. We quote the following result that establishes finite dimensional convergence of $\mathbb{M}_n$ on $\mathcal{L}_q$ with $q < q_0/2$. Note that the result in Theorem 2.1 corresponds to the convergence of finite dimensional distributions of tail array sums with deterministic levels defined in (1.5).

Theorem 2.1 (Kulik et al. [18], Theorem 2.3). *Let Assumption 1 hold and let $\{u_n\}$ be an increasing sequence such that*

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} n\bar{F}(u_n) = +\infty. \tag{2.8}$$

Assume there exists $\eta > 0$ such that

$$\lim_{n\to\infty} \log^{1+\eta}(n)\bar{F}(u_n) = 0. \tag{2.9}$$

Assume moreover that either ϕ is bounded or there exists $\delta > 0$ such that $q(2+\delta) \le q_0$ and

$$\lim_{n\to\infty} \frac{\log^{1+\eta}(n)}{\{n\bar{F}(u_n)\}^{\delta/2}} = 0. \tag{2.10}$$

Then $\mathbb{M}_n \stackrel{\text{fi.di.}}{\longrightarrow} \mathbb{M}$ on $\mathcal{L}_q$.

For statistical purposes, we need to consider the process $\mathbb{M}_n$ indexed by a subclass $\mathcal{G} \subset \mathcal{L}_q$ of functions and convergence of $\mathbb{M}_n$ to $\mathbb{M}$ must be strengthened to weak convergence in $\ell_\infty(\mathcal{G})$, in particular in order to replace the deterministic threshold u_n by an appropriate sequence of order statistics. The general theory of weak convergence in $\ell_\infty(\mathcal{G})$ is developed in van der Vaart and Wellner [21] and Giné and Nickl [14] and was adapted in full generality in the context of cluster statistics in Drees and Rootzén [11], while in the present context it was considered in Kulik et al. [18]. We quote the following result from Kulik et al. [18], Theorem 2.4, to which we also refer for some notions related to empirical processes and classes of functions. To proceed, let ρ_h be the pseudometric defined on $\mathcal{L}_q$ by

$$\rho_h^2(\phi,\psi) = \boldsymbol{\nu}_{0,h}\big((\phi-\psi)^2\big).$$

Note that ρ_h is well defined under the assumptions of Theorem 2.1 which imply $q < q_0/2$. Choose $r_n = \log^{1+\eta}(n)$ such that (2.9) and (2.10) hold. Set $m_n = (n-h)/r_n$ and assume for simplicity that m_n is an integer. Then, as $r_n \to \infty$, we have also $m_n \to \infty$. Consider m_n non-overlapping blocks $\{j = (i-1)r_n + 1, \dots, ir_n\}$, $i = 1, \dots, m_n$. Define the random pseudometric d_n on $\mathcal{G}$ by

$$d_n^2(\phi,\phi') = \frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \left\{ \sum_{j=(i-1)r_n+1}^{ir_n} \left(\phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right) - \phi'\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right)\right)\right\}^2, \quad \phi,\phi' \in \mathcal{G}.$$

Let $N(\epsilon,\mathcal{G},d_n)$ be the minimum number of balls in the pseudometric d_n needed to cover $\mathcal{G}$.

Theorem 2.2. *Let the assumptions of Theorem 2.1 hold and let $\mathcal{G} \subset \mathcal{L}_q$. Assume moreover that*

(i) *$\mathcal{G}$ is pointwise separable and linearly ordered*;
(ii) *the envelope function $\Phi_{\mathcal{G}} = \sup_{\phi\in\mathcal{G}} |\phi|$ is in $\mathcal{L}_q$*;
(iii) *$(\mathcal{G},\rho_h)$ is totally bounded*;
(iv) *for every sequence $\{\delta_n\}$ which decreases to zero,*

$$\limsup_{n\to\infty} \sup_{\substack{\phi,\psi\in\mathcal{G}\\ \rho_h(\phi,\psi)\le\delta_n}} \frac{\mathbb{E}[\{\phi(u_n^{-1}\boldsymbol{X}_{0,h}) - \psi(u_n^{-1}\boldsymbol{X}_{0,h})\}^2]}{\bar{F}(u_n)} = 0; \tag{2.11}$$

Then $\mathbb{M}_n \Rightarrow \mathbb{M}$ in $\ell_\infty(\mathcal{G})$.

Now, we turn to the practically more relevant tail array sums with random levels defined in (1.3). We set $k = n\bar{F}(u_n)$ and define the processes $\widehat{\mathbb{M}}_n$ and $\widehat{\mathbb{M}}$ on $\mathcal{L}_q$ by

$$\widehat{\mathbb{M}}_n(\phi) = \sqrt{k}\left\{\frac{1}{k}\sum_{j=1}^{n}\phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right) - \boldsymbol{\nu}_{0,h}(\phi)\right\},$$

$$\widehat{\mathbb{M}}(\phi) = \mathbb{M}(\phi) - \boldsymbol{\nu}_{0,h}(\phi)\mathbb{M}\big(\mathbb{1}_{(1,\infty)\times\mathbb{R}^h}\big).$$

Remark 2.3. For each ϕ, $\widehat{\mathbb{M}}(\phi)$ has variance $\sigma^2(\phi - \boldsymbol{\nu}_{0,h}(\phi)\mathbb{1}_{(1,\infty)\times\mathbb{R}^h})$ in contrast to $\mathbb{M}(\phi)$ which has variance $\sigma^2(\phi)$. Hence, the distributions of $\widehat{\mathbb{M}}(\phi)$ and $\mathbb{M}(\phi)$ differ.

The following results establish central limit theorems for the estimator $\widehat{M}_n(\phi)$ of $\boldsymbol{\nu}_{0,h}(\phi)$. Recall that for $\phi : \mathbb{R}^{h+1} \to \mathbb{R}$ and $s > 0$ the function ϕ_s is defined by $\phi_s(\boldsymbol{x}) = \phi(\boldsymbol{x}/s)$.

Corollary 2.4 (Kulik et al. [18], Corollary 2.6). *Let the assumptions of Theorem 2.1 hold. Let* $0 < s_0 < 1 < t_0 < \infty$ *and let* $\mathcal{G}_0^* \subset \mathcal{L}_q$. *Define* $\mathcal{G}^* = \{\phi_s, \phi \in \mathcal{G}_0^*, s \in [s_0, t_0]\}$. *If* $\mathcal{G}^*$ *satisfies the assumptions* (*i*)–(*iv*) *of Theorem 2.2 and*

$$\lim_{n\to\infty}\sqrt{k}\sup_{s_0\le s\le t_0}\left|\frac{\bar{F}(u_n s)}{\bar{F}(u_n)} - s^{-\alpha}\right| = 0, \tag{2.12}$$

$$\lim_{n\to\infty}\sqrt{k}\sup_{s_0\le s\le t_0}\sup_{\phi\in\mathcal{G}_0^*}\left|\frac{\mathbb{E}[\phi(\boldsymbol{X}_{0,h}/(u_n s))]}{\bar{F}(u_n)} - s^{-\alpha}\boldsymbol{\nu}_{0,h}(\phi)\right| = 0 \tag{2.13}$$

hold, then $\widehat{\mathbb{M}}_n \Rightarrow \widehat{\mathbb{M}}$ *on* $\ell_\infty(\mathcal{G}_0^*)$.

We note that (2.12), (2.13) are the bias conditions that cannot be dispensed of. They hold for *some* sequences k and u_n. The first one can be handled by the classical univariate second order assumption applied to F; see de Haan and Ferreira [6]. The second one, in the context of time series, has to be treated on case-by-case basis, by considering specific functions ϕ and particular time series models. We do not pursue this direction here. Indeed, as we will see below, the bootstrap procedures are valid without any additional bias conditions.

In Kulik et al. [18] the authors applied the above result to different functions ϕ to obtain central limit theorems for quantities related to extremal behaviour of time series such as conditional tail expectation or distribution of the spectral tail process. Another classical application is estimation of the tail index of F. In all these applications, the limiting variance is given by a complicated, infinite sum. Therefore, we want to employ resampling techniques to conduct statistical inference for tail array sums. In the following section, we propose a multiplier block bootstrap that is in particular appealing in terms of low computational burden and high theoretical tractability.

3. Convergence of multiplier tail array sums

In this section, we consider a multiplier bootstrap procedure for tail array sums with random levels $\widehat{\mathbb{M}}_n(\phi)$. Recall that the latter can be represented in two ways:

$$\widehat{\mathbb{M}}_n(\phi) = \sqrt{k}\left(\frac{\sum_{j=1}^{n-h}\phi(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}})}{k} - \boldsymbol{\nu}_{0,h}(\phi)\right) \tag{3.1}$$

$$= \sqrt{k}\left(\frac{\sum_{j=1}^{n-h}\phi(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}})}{\sum_{j=1}^{n}\mathbb{1}_{\{X_j>X_{n:n-k}\}}} - \boldsymbol{\nu}_{0,h}(\phi)\right). \tag{3.2}$$

The goal is to create a multiplier bootstrap analogue to $\widehat{\mathbb{M}}_n(\phi)$ that mimics the limiting distribution of the latter. Let $\{\xi_i, i \geq 1\}$ be a sequence of i.i.d. random variables such that $\mathbb{E}[\xi_0] = 0$ and $\mathbb{E}[\xi_0^2] = 1$. In view of the two representations (3.1) and (3.2), two options appear to be natural to apply a multiplier (block) bootstrap for $\widehat{\mathbb{M}}_n(\phi)$:

(i) Consider the first representation of $\widehat{\mathbb{M}}_n(\phi)$ in (3.1), split the sum $\sum_{j=1}^{n-h}\phi(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}})$ into m_n blocks of size r_n, apply the multipliers ξ_i to each of the blocks, while keeping k in the denominator.

(ii) Consider the second representation of $\widehat{\mathbb{M}}_n(\phi)$ in (3.2), split both sums $\sum_{j=1}^{n-h}\phi(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}})$ and $\sum_{j=1}^{n}\mathbb{1}_{\{X_j>X_{n:n-k}\}}$ into m_n blocks of size r_n, align the corresponding blocks and apply the multipliers ξ_i to each of the joint, i.e. two-dimensional, blocks.

To proceed, set

$$\widehat{\Phi}_{n,i} := \sum_{j=(i-1)r_n+1}^{ir_n}\phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right), \qquad \widehat{\Upsilon}_{n,i} = \sum_{j=(i-1)r_n+1}^{ir_n}\mathbb{1}\{X_j > X_{n:n-k}\}. \tag{3.3}$$

Define the multiplier bootstrap version of $\widehat{\mathbb{M}}_n(\phi)$ corresponding to option (i) by

$$\widehat{\mathbb{M}}_{n,\xi}(\phi) = \sqrt{k}\left(\frac{\sum_{i=1}^{m_n}(1+\xi_i)\widehat{\Phi}_{n,i}}{k} - \frac{\sum_{i=1}^{m_n}\widehat{\Phi}_{n,i}}{k}\right) = \frac{\sum_{i=1}^{m_n}\xi_i\widehat{\Phi}_{n,i}}{\sqrt{k}} \tag{3.4}$$

and to option (ii) by

$$\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\phi) = \sqrt{k}\left(\frac{\sum_{i=1}^{m_n}(1+\xi_i)\widehat{\Phi}_{n,i}}{\sum_{i=1}^{m_n}(1+\xi_i)\widehat{\Upsilon}_{n,i}} - \frac{\sum_{i=1}^{m_n}\widehat{\Phi}_{n,i}}{\sum_{i=1}^{m_n}\widehat{\Upsilon}_{n,i}}\right). \tag{3.5}$$

For a metric space $(\mathcal{X}, \rho_{\mathcal{X}})$ we denote

$$\mathbb{BL}(\mathcal{X}) = \left\{\Psi : \mathcal{X} \to \mathbb{R} : \sup_{x\in\mathcal{X}}\left|\Psi(x)\right| \leq 1, \left|\Psi(x) - \Psi(y)\right| \leq \rho_{\mathcal{X}}(x, y), x, y \in \mathcal{X}\right\}. \tag{3.6}$$

This space metrizes weak convergence in $\mathcal{X}$; see van der Vaart and Wellner [21], p. 73. The main result of this section is the following theorem which shows (conditional) weak convergence of $\widehat{\widehat{\mathbb{M}}}_{n,\xi}$ and $\widehat{\mathbb{M}}_{n,\xi}$ as processes indexed by the class from Corollary 2.4 In what follows, E_ξ denotes the conditional expectation (given $X_1, \ldots, X_n$).

Theorem 3.1. *Let the assumptions of Theorem 2.1 hold and that either*

- ϕ *is bounded and*

$$\lim_{n\to\infty}\frac{r_n^2}{n\bar{F}(u_n)} = 0 \tag{3.7}$$

holds; *or*

- $\phi \in \mathcal{L}_q$, $q < q_0/4$, *and*

$$\lim_{n\to\infty} \frac{r_n^3}{n\bar{F}(u_n)} = 0 \tag{3.8}$$

holds.

Let $0 < s_0 < 1 < t_0 < \infty$ *and let* $\mathcal{G}_0^* \subset \mathcal{L}_q$. *Define* $\mathcal{G}^* = \{\phi_s, \phi \in \mathcal{G}_0^*, s \in [s_0, t_0]\}$. *If* $\mathcal{G}^*$ *satisfies the assumptions* (*i*)–(*iv*) *of Theorem* 2.2 *and* (2.12)–(2.13) *hold, then*

$$\sup_{\Psi\in\mathbb{BL}(\ell_\infty(\mathcal{G}_0^*))} \left|\mathbb{E}_\xi\left[\Psi(\widehat{\widehat{\mathbb{M}}}_{n,\xi})\right] - \mathbb{E}\left[\Psi(\widehat{\mathbb{M}})\right]\right| \to 0$$

and

$$\sup_{\Psi\in\mathbb{BL}(\ell_\infty(\mathcal{G}_0^*))} \left|\mathbb{E}_\xi\left[\Psi(\widehat{\mathbb{M}}_{n,\xi})\right] - \mathbb{E}\left[\Psi(\mathbb{M})\right]\right| \to 0,$$

in probability.

Remark 3.2. The corresponding results for the functions $\psi_{s,y}(x_0, \ldots, x_h) = \mathbb{1}\{|x_0| > s, x_h \le y|x_0|\}$ and the class $\{\psi_{s,y}, s \in [s_0, t_0], y \in \mathbb{R}\}$ were obtained in Drees and Knezevic [10], Theorem 2.3.

Remark 3.3. We note that the limiting distributions for the multiplier statistics $\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\phi)$ and $\widehat{\mathbb{M}}_n(\phi)$ coincide. On the other hand, $\widehat{\mathbb{M}}_{n,\xi}(\phi)$ does not recover the desired limit $\widehat{\mathbb{M}}_n(\phi)$.

Remark 3.4. In the i.i.d. case the Hill estimator $\widehat{\gamma}_k$ in (1.11) has the limiting variance α^{-2}, while the bootstrapped Hill estimator $\widehat{\gamma}_{k,\xi}$ in (1.15) is approximated by a normal random variable with variance $2\alpha^{-2}$. Note that the latter limiting variance is that of the Hill estimator with deterministic level and deterministic normalization defined in (1.12). Hence, $\widehat{\gamma}_{k,\xi}$ obviously fails to mimic the randomness in $\widehat{\mathbb{M}}_n(\phi)$ induced from random levels, whereas $\widehat{\widehat{\gamma}}_{k,\xi}$ (defined analogously to $\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\phi)$) succeeds in mimicking the randomness induced from random normalization leading to the correct limiting distribution. The phenomenon that a multiplier bootstrap may lead to a wrong limiting distribution has already been observed in Bücher and Dette [2], where it is applied to tail copulas. The problem is due to a replacement of marginal cumulative distribution functions by empirical distribution functions, which is comparable to the issue of replacing deterministic levels by random levels.

Remark 3.5. Let $\mathcal{I}_s(x_0, \ldots, x_h) = \mathbb{1}\{x_0 > s\}$ and define the class $\mathcal{G}' = \{\mathcal{I}_s, s \in [s_0, t_0]\}$. Note that $\mathcal{G}'$ satisfies the assumptions of Theorem 2.2 (cf. Kulik et al. [18]). This will have implications for a joint convergence below; see, for example, (3.13). In turn the joint convergence is needed to replace the deterministic threshold with the order statistics.

3.1. Two examples of tail array sums

For the general estimator introduced in (1.3), that is,

$$\widehat{M}_n(\phi) := \frac{1}{k}\sum_{j=1}^{n-h} \phi\left(\frac{\boldsymbol{X}_{j,j+h}}{X_{n:n-k}}\right),$$

where $X_{n:1} \leq \cdots \leq X_{n:n}$ are the order statistics from the sample $X_1, \ldots, X_n$ and $k = k_n \to \infty$ is such that $k/n \to 0$, we choose two specifications; compare Example 1 in Section 1.

3.1.1. *Hill estimator*

First, with $h = 0$ and $\phi(\boldsymbol{x}_{0,h}) = \log(x_0)\mathbb{1}\{x_0 > 1\}$, we get the well-known Hill estimator

$$\widehat{\gamma}_k := \frac{1}{k}\sum_{j=1}^{n} \log\left(\frac{X_j}{X_{n:n-k}}\right)\mathbb{1}\left\{\frac{X_j}{X_{n:n-k}} > 1\right\} = \frac{1}{k}\sum_{j=0}^{k-1} \log\left(\frac{X_{n:n-j}}{X_{n:n-k}}\right), \tag{3.9}$$

as an estimator for $M(\phi) = 1/\alpha =: \gamma$. Clearly, the class $\{\phi_s, s \in [s_0, t_0]\}$ is linearly ordered. Under the appropriate conditions (as those in Theorem 2.1) the limiting distribution of the appropriately scaled Hill estimator is normal with the variance given by an infinite series; see, for example, the stochastic recurrence example in Drees [7] or the linear process in Drees [8]. In the language of the so-called tail process $\{Y_j, j \in \mathbb{Z}\}$ (see Example 1(iv)), the distributional limit of the sequence $\{X_j, j \in \mathbb{Z}\}$ given that $|X_0|$ is large (see, e.g., Basrak and Segers [1] for the precise statement), the limiting variance is given by (cf. Theorem 9.5.2 in Kulik and Soulier [17])

$$\alpha^{-2}\sum_{j\in\mathbb{Z}} \mathbb{P}(Y_j > 1 \mid Y_0 > 1).$$

In the easy case of an AR(1) process $\{X_j, j \in \mathbb{Z}\}$ with coefficient $\rho \in (0, 1)$, the limiting variance of the Hill estimator of α^{-1} given in (3.9) can be stated explicitly (cf. Drees [8]) and is given by

$$\alpha^{-2}\left(\frac{1+\rho^\alpha}{1-\rho^\alpha}\right).$$

3.1.2. *Extremogram*

Second, with $h \geq 1$ and $\phi(\boldsymbol{x}_{0,h}) = \mathbb{1}\{x_0 > 1, x_h > 1\}$, we obtain the extremogram at lag h as considered in Davis and Mikosch [4]. Again, the class $\{\phi_s, s \in [s_0, t_0]\}$ is linearly ordered. The sample extremogram is

$$\widehat{M}_n(\phi) := \frac{1}{k}\sum_{j=1}^{n-h} \mathbb{1}\left\{\frac{X_j}{X_{n:n-k}} > 1, \frac{X_{j+h}}{X_{n:n-k}} > 1\right\}.$$

It is an estimator for $M(\phi) = \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x)$ which is known as the (upper) tail dependence coefficient between X_0 and X_h.

To the best of our knowledge, there is no explicit formula available in the literature for the limiting distribution of an appropriately normalized estimator of the extremogram, but we can argue that the limiting variance can be again represented as an infinite series in terms of the tail process (note that Davis and Mikosch [4] consider the estimator with deterministic levels, hence different limiting distribution).

3.2. Proof of Theorem 3.1

We proceed in three steps. We consider a multiplier process with deterministic levels (Proposition 3.6), multiplier process with random normalization (Proposition 3.12) and finally multiplier process with random levels that concludes the proof of Theorem 3.1.

3.2.1. *Multiplier process with deterministic levels*

Recall that $\phi_s(\cdot) = \phi(\cdot/s)$. Similarly to (3.3), let

$$\Phi_{n,i}(s) = \sum_{j=(i-1)r_n+1}^{ir_n} \phi_s\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right), \qquad \Psi_{n,i}(s;y) = \sum_{j=(i-1)r_n+1}^{ir_n} \psi_{s,y}\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right),$$

$$\Upsilon_{n,i}(s) = \sum_{j=(i-1)r_n+1}^{ir_n} \mathbb{1}\{X_j > u_n s\}.$$

Recall that $m_n = (n-h)/r_n$ is assumed to be an integer, and define the multiplier process by

$$\mathbb{M}_{n,\xi}(\phi_s) = \frac{1}{\sqrt{n\bar{F}(u_n)}}\left(\frac{\sum_{i=1}^{m_n} \xi_i\{\Phi_{n,i}(s) - \mathbb{E}[\Phi_{n,i}(s)]\}}{n\bar{F}(u_n)}\right). \tag{3.10}$$

The analogous expression holds for $\mathbb{M}_{n,\xi}(\psi_{s,y})$. For the special case of function $\phi(\boldsymbol{x}_{0,h}) = \mathbb{1}\{x_0 > 1\}$ and $\mathcal{I}_s(\boldsymbol{x}_{0,h}) = \mathbb{1}\{x_0 > s\}$ we consider

$$\mathbb{M}_{n,\xi}(\mathcal{I}_s) = \sqrt{n\bar{F}(u_n)}\left(\frac{\sum_{i=1}^{m_n} \xi_i\{\Upsilon_{n,i}(s) - \mathbb{E}[\Upsilon_{n,i}(s)]\}}{n\bar{F}(u_n)}\right). \tag{3.11}$$

We note that for all $\phi \in \mathcal{L}_q$,

$$\frac{\mathbb{E}[\Phi_{n,1}(s)]}{r_n\bar{F}(u_n)} \to \boldsymbol{\nu}_{0,h}(\phi_s), \qquad \frac{\mathbb{E}[\Upsilon_{n,1}(s)]}{r_n\bar{F}(u_n)} \to s^{-\alpha}. \tag{3.12}$$

By Theorem 2.2, Remark 3.5 and the bias conditions (2.13), we have (joint) weak convergence in $\ell_\infty(\mathcal{G}^*)$ of the processes

$$\sqrt{n\bar{F}(u_n)}\left\{\frac{\sum_{i=1}^{m_n} \Phi_{n,i}(s)}{n\bar{F}(u_n)} - \boldsymbol{\nu}_{0,h}(\phi_s)\right\}, \qquad \sqrt{n\bar{F}(u_n)}\left\{\frac{\sum_{i=1}^{m_n} \Psi_{n,i}(s;y)}{n\bar{F}(u_n)} - \boldsymbol{\nu}_{0,h}(\psi_{s,y})\right\} \tag{3.13}$$

and

$$\sqrt{n\bar{F}(u_n)}\left\{\frac{\sum_{i=1}^{m_n} \Upsilon_{n,i}(s)}{n\bar{F}(u_n)} - s^{-\alpha}\right\}. \tag{3.14}$$

Vervaat lemma implies, in particular, that $X_{n:n-k}/u_n \xrightarrow{p} 1$.

The first result is about weak convergence of the multiplier process $\mathbb{M}_{n,\xi}$, indexed by the class $\mathcal{G} = \mathcal{G}^*$.

Proposition 3.6. *Let $\mathcal{G}^*$ be as in Theorem 3.1. Then*

$$\sup_{\Psi \in \mathbb{BL}(\ell_\infty(\mathcal{G}^*))} \left|\mathbb{E}_\xi\left[\Psi(\mathbb{M}_{n,\xi})\right] - \mathbb{E}\left[\Psi(\mathbb{M})\right]\right| \to 0,$$

as $n \to \infty$ in probability.

Remark 3.7. We point out that in Theorem 3.1 the sup is taken over $\mathbb{BL}(\ell_\infty(\mathcal{G}_0^*))$, while here over a larger class $\mathbb{BL}(\ell_\infty(\mathcal{G}^*))$. The smaller class $\mathcal{G}_0^*$ will show up when substituting u_n with order statistics.

Proof. As usual, we proceed by proving finite dimensional convergence and tightness.

Let $I_{n,i} = \{(i-1)r_n + 1, \dots, ir_n\}$. Let $\{X_j^\dagger, j \in \mathbb{Z}\}$ be a sequence such that the blocks $\{X_i^\dagger, i \in I_{n,i}\}$, $i = 1, \dots, m_n$, are independent, but have the same distribution as the blocks $\{X_i, i \in I_{n,i}\}$, $i = 1, \dots, m_n$ of the original sequence. Let $\Phi_{n,i}^\dagger(s)$ be the statistics defined in the same way as $\Phi_{n,i}(s)$, but based on the i.i.d. blocks. Clearly, $\mathbb{E}[\Phi_{n,i}^\dagger(s)] = \mathbb{E}[\Phi_{n,i}(s)]$. Set

$$\mathbb{M}_{n,\xi}^\dagger(\phi_s) = \sqrt{n\bar{F}(u_n)}\left(\frac{\sum_{i=1}^{m_n} \xi_i\{\Phi_{n,i}^\dagger(s) - \mathbb{E}[\Phi_{n,i}(s)]\}}{n\bar{F}(u_n)}\right). \tag{3.15}$$

With help of beta-mixing, arguing as in Drees and Rootzén [11] and Kulik et al. [18], in order to prove finite dimensional convergence of $\mathbb{M}_{n,\xi}$, it suffices to consider the process $\mathbb{M}_{n,\xi}^\dagger$. With help of Assumption 1 we obtain some convergence results in Lemma 3.8. These are used to prove fidi convergence (conditional) in Lemma 3.10 and tightness (unconditional) in Lemma 3.11. The proof is concluded by showing the conditional weak convergence. □

3.2.2. *Fidi convergence* (*conditional*)

Here, we provide calculations for $\Phi_{n,i}(s)$ only. Computations for $\Psi_{n,i}(s, y)$ are analogous.

Lemma 3.8. *Let Assumption 1 hold. Then for all $\phi \in \mathcal{L}_q$, $q < q_0/2$ and $s, t \in [s_0, t_0]$,*

$$\lim_{n\to\infty} \frac{1}{r_n\bar{F}(u_n)} \operatorname{cov}\big(\Phi_{n,1}(s), \Phi_{n,1}(t)\big) = \lim_{n\to\infty} \frac{1}{r_n\bar{F}(u_n)} \mathbb{E}\big[\Phi_{n,1}(s)\Phi_{n,1}(t)\big] = C(\phi_s, \phi_t) \tag{3.16}$$

and for all $\eta > 0$

$$\lim_{n\to\infty} \frac{m_n}{n\bar{F}(u_n)} \mathbb{E}\big[\big\{\Phi_{n,1}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2 \mathbb{1}\big\{\big|\Phi_{n,1}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big| > \eta\sqrt{n\bar{F}(u_n)}\big\}\big]$$
$$= \lim_{n\to\infty} \frac{m_n}{n\bar{F}(u_n)} \mathbb{E}\big[\big\{\Phi_{n,1}(s)\big\}^2 \mathbb{1}\big\{\big|\Phi_{n,1}(s)\big| > \eta\sqrt{n\bar{F}(u_n)}\big\}\big] = 0. \tag{3.17}$$

Proof. The first and the second statements are proven in Kulik et al. [18] (Lemmas 3.3, 3.6 and 3.7). □

Lemma 3.9. *Let Assumption 1 hold. If either*

- *ϕ is bounded and (3.7) holds; or*
- *$\phi \in \mathcal{L}_q$, $q < q_0/4$, and (3.8) holds,*

then

$$\frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \Phi_{n,i}^\dagger(s)\Phi_{n,i}^\dagger(t) \xrightarrow{p} C(\phi_s, \phi_t). \tag{3.18}$$

Proof. We note first that (cf. Kulik et al. [18])

$$\frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \mathbb{E}\big[\Phi_{n,i}^\dagger(s)\Phi_{n,i}^\dagger(t)\big] = \frac{m_n}{n\bar{F}(u_n)} \mathbb{E}\big[\Phi_{n,1}^\dagger(s)\Phi_{n,1}^\dagger(t)\big]$$
$$= \frac{1}{r_n\bar{F}(u_n)} \operatorname{cov}\big(\Phi_{n,1}(s), \Phi_{n,1}(t)\big) + o(1) \to C(\phi_s, \phi_t)$$

by (3.16). Thanks to independence between the blocks,

$$\begin{aligned}
&\operatorname{var}\left(\frac{1}{n\bar F(u_n)}\sum_{i=1}^{m_n}\Phi^\dagger_{n,i}(s)\Phi^\dagger_{n,i}(t)\right)\\
&\quad=\frac{m_n}{n^2\bar F^2(u_n)}\operatorname{var}\big(\Phi^\dagger_{n,1}(s)\Phi^\dagger_{n,1}(t)\big)\\
&\quad=\frac{m_n}{n^2\bar F^2(u_n)}\mathbb{E}\big[\big(\Phi^\dagger_{n,1}(s)\Phi^\dagger_{n,1}(t)\big)^2\big]-\frac{m_n}{n^2\bar F^2(u_n)}\mathbb{E}\big[\big(\Phi^\dagger_{n,1}(s)\big)^2\big]\mathbb{E}\big[\big(\Phi^\dagger_{n,1}(t)\big)^2\big].
\end{aligned}$$

Lemma 3.3 in Kulik et al. [18] yields $\mathbb{E}[(\Phi^\dagger_{n,1}(s))^2]=O(r_n\bar F(u_n))$ and hence the last term in the display above is of order

$$\frac{m_n}{n^2\bar F^2(u_n)}r_n^2\bar F^2(u_n)=\frac{n}{r_n n^2}r_n^2=r_n/n=o(1).$$

Now, if ϕ is bounded, then

$$\mathbb{E}\big[\big(\Phi^\dagger_{n,1}(s)\Phi^\dagger_{n,1}(t)\big)^2\big]\le\|\phi\|_\infty r_n^2\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_t\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right)\right)^2\right]$$

and therefore the first term in the display for the variance is of order

$$\frac{m_n}{n^2\bar F^2(u_n)}r_n^2 O\big(r_n\bar F(u_n)\big)=\frac{O(r_n^2)}{n\bar F(u_n)}=o(1).$$

In case of unbounded functions, we proceed as follows:

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_s\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right)\right)^2\left(\sum_{j=1}^{r_n}\phi_t\left(\frac{\boldsymbol{X}_{j,j+h}}{u_n}\right)\right)^2\right]\\
&\quad=:\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_{s,j}\right)^2\left(\sum_{j=1}^{r_n}\phi_{t,j}\right)^2\right]=\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_{s,j}^2+\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{s,j}\phi_{s,j'}\right)\left(\sum_{j=1}^{r_n}\phi_{t,j}^2+\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{t,j}\phi_{t,j'}\right)\right]\\
&\quad=\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_{s,j}^2\right)\left(\sum_{i=1}^{r_n}\phi_{t,i}^2\right)\right]+\mathbb{E}\left[\left(\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{s,j}\phi_{s,j'}\right)\left(\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{t,j}\phi_{t,j'}\right)\right]\\
&\qquad+\mathbb{E}\left[\left(\sum_{j=1}^{r_n}\phi_{s,j}^2\right)\left(\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{t,j}\phi_{t,j'}\right)\right]+\mathbb{E}\left[\left(\sum_{i=1}^{r_n}\phi_{t,i}^2\right)\left(\sum_{\substack{j,j'=1\\ j\ne j'}}^{r_n}\phi_{s,j}\phi_{s,j'}\right)\right]=:I_1+I_2+I_3+I_4.
\end{aligned}$$

For I_1 we have

$$\frac{m_n}{n^2\bar F^2(u_n)}I_1=o(1)\frac{1}{r_n\bar F(u_n)}I_1.$$

Application of (3.16) gives $I_1 = O(r_n \bar{F}(u_n))$. Therefore, the term that corresponds to I_1 vanishes. For I_2 we have

$$\frac{m_n}{n^2 \bar{F}^2(u_n)} I_2 \le \frac{m_n}{n^2 \bar{F}^2(u_n)} \sum_{\substack{i,i'=1 \\ i \ne i'}}^{r_n} \sum_{\substack{j,j'=1 \\ j \ne j'}}^{r_n} \mathbb{E}[\phi_{s,i}\phi_{s,i'}\phi_{t,j}\phi_{t,j'}].$$

Considering the case of distinct indices only we have

$$\frac{m_n}{n^2 \bar{F}^2(u_n)} I_2 \le \frac{m_n}{n^2 \bar{F}^2(u_n)} r_n^4 \mathbb{E}^{1/2}\big[\phi_{s,1}^4\big]\mathbb{E}^{1/2}\big[\phi_{t,1}^4\big] = \frac{n}{r_n n^2 \bar{F}^2(u_n)} r_n^4 O\big(\bar{F}(u_n)\big) = O\left(\frac{r_n^3}{n\bar{F}(u_n)}\right).$$

The terms I_3 and I_4 are treated in analogous way. □

The next result shows that the finite dimensional distributions of the multiplier process $\mathbb{M}_{n,\xi}^{\dagger}$ converge to the same limit as those of $\mathbb{M}_n$. The lemma below is a modified version of the result from Kosorok [16], since we believe that the original result has a gap in its proof. We state it just in a one dimensional case, but extension to a vector $(s_1, \ldots, s_d)$ is straightforward.

Lemma 3.10. *Assume that* (3.18) *and* (3.17) *hold. Then for each* $\phi \in \mathcal{L}_q$ *and* $s \in [s_0, t_0]$,

$$\sup_{\Psi \in \mathbb{BL}(\mathbb{R})} \big|\mathbb{E}_\xi\big[\Psi\big(\mathbb{M}_{n,\xi}^{\dagger}(\phi_s)\big)\big] - \mathbb{E}\big[\Psi\big(\mathbb{M}(\phi_s)\big)\big]\big| \to 0$$

as $n \to \infty$ *in probability.*

Proof. Note that (3.17) and the Markov inequality imply that

$$\frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \big\{\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2 \mathbb{1}\big\{\big|\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big| > \epsilon\sqrt{n\bar{F}(u_n)}\big\} \to 0, \quad (3.19)$$

in probability. If we assume for simplicity that ξ_i's are bounded, then this automatically implies that

$$\frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \xi_i^2 \big\{\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2 \mathbb{1}\big\{|\xi_i|\big|\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big| > \epsilon\sqrt{n\bar{F}(u_n)}\big\} \to 0, \quad (3.20)$$

in probability. In case of unbounded ξ_i's one needs to apply a truncation argument. For arbitrary $A > 0$, we have

$$\begin{aligned} &\frac{1}{n\bar{F}(u_n)} \mathbb{E}_\xi\left[\sum_{i=1}^{m_n} \xi_i^2 \big\{\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2 \mathbb{1}\big\{|\xi_i|\big|\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big| > \epsilon\sqrt{n\bar{F}(u_n)}\big\}\right] \\ &\le \frac{A^2}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \big\{\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2 \mathbb{1}\big\{\big|\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big| > \epsilon\sqrt{n\bar{F}(u_n)}/A\big\} \\ &\quad + \mathbb{E}\big[\xi_1^2 \mathbb{1}\big\{|\xi_1| > A\big\}\big] \frac{1}{n\bar{F}(u_n)} \sum_{i=1}^{m_n} \big\{\Phi_{n,i}^{\dagger}(s) - \mathbb{E}\big[\Phi_{n,1}(s)\big]\big\}^2. \end{aligned}$$

The first expression in the last line converges in probability to zero as $n \to \infty$ by (3.19). By (3.18), the second part is bounded in probability in n and vanishes by letting $A \to \infty$.

Furthermore, again by (3.18),

$$\frac{1}{n\bar{F}(u_n)}\mathbb{E}_\xi\left[\sum_{i=1}^{m_n}\xi_i^2\left\{\Phi_{n,i}^\dagger(s)-\mathbb{E}\left[\Phi_{n,i}^\dagger(s)\right]\right\}^2\right]$$
$$=\mathbb{E}\left[\xi_1^2\right]\frac{1}{n\bar{F}(u_n)}\sum_{i=1}^{m_n}\left\{\Phi_{n,i}^\dagger(s)-\mathbb{E}\left[\Phi_{n,i}^\dagger(s)\right]\right\}^2\xrightarrow{p}C(\phi_s,\phi_s). \tag{3.21}$$

Thus, for each subsequence n' there exists a further subsequence n'' such that the convergence in (3.20)–(3.21) holds almost surely on subsequences n''. Therefore, we can use the Lindeberg central limit theorem to conclude the result. □

3.2.3. *Tightness* (*unconditional*)

Consider the process

$$\mathbb{Z}_{n,\xi}^\dagger(\phi_s)=\frac{1}{\sqrt{n\bar{F}(u_n)}}\sum_{i=1}^{m_n}\xi_i\Phi_{n,i}^\dagger(s),\quad \phi\in\mathcal{G}_0^*, s\in[s_0,t_0].$$

Note that

$$\frac{1}{\sqrt{n\bar{F}(u_n)}}\sum_{i=1}^{m_n}\xi_i\mathbb{E}\left[\Phi_{n,i}(s)\right]=\frac{\mathbb{E}[\Phi_{n,1}(s)]}{r_n\bar{F}(u_n)}\underbrace{\frac{\sqrt{m_nr_n}}{\sqrt{n}}\sqrt{r_n\bar{F}(u_n)}\frac{1}{\sqrt{m_n}}\sum_{i=1}^{m_n}\xi_i}_{=:a_n}\xrightarrow{p}0$$

by the central limit theorem applied to ξ_i's, (2.9) and (3.12). Thus, the processes $\mathbb{M}_{n,\xi}^\dagger$ and $\mathbb{Z}_{n,\xi}^\dagger$ have the same limiting finite dimensional distributions. Furthermore,

$$\sup_{\Psi\in\mathbb{BL}(\ell_\infty(\mathcal{G}^*))}\left|\mathbb{E}_\xi\left[\Psi\left(\mathbb{M}_{n,\xi}^\dagger\right)\right]-\mathbb{E}_\xi\left[\Psi\left(\mathbb{Z}_{n,\xi}^\dagger\right)\right]\right|\le|a_n|\sup_{\phi\in\mathcal{G}_0^*}\sup_{s\in[s_0,t_0]}\frac{\mathbb{E}[\Phi_{n,1}(s)]}{r_n\bar{F}(u_n)}\xrightarrow{p}0,$$

by (3.12) and since we assumed that the envelope function belongs to $\mathcal{L}_q$. Therefore, the processes $\mathbb{M}_{n,\xi}^\dagger$ and $\mathbb{Z}_{n,\xi}^\dagger$ have the same weak limits. Thus, it suffices to prove tightness of $\mathbb{Z}_{n,\xi}^\dagger$.

Lemma 3.11. *Let $\mathcal{G}^*$ be as in Theorem 3.1. Then the process $\mathbb{Z}_{n,\xi}^\dagger$ is tight.*

Proof. We mimic the proof of Theorem 2.4 in Kulik et al. [18], with the appropriate modifications to accommodate the multipliers.

Let $\ell_0(\mathbb{R}^{h+1})$ be the set of $\mathbb{R}^{h+1}$-valued sequences $\boldsymbol{x}=\{\boldsymbol{x}_j, j\in\mathbb{Z}\}$ such that $\lim_{|j|\to\infty}|\boldsymbol{x}_j|=0$. Let $\mathcal{H}_q$ be the set of functions f defined on $\ell_0(\mathbb{R}^{h+1})$ for which there exists $\phi\in\mathcal{L}_q$ such that

$$f(\boldsymbol{x})=\sum_{j\in\mathbb{Z}}\phi(\boldsymbol{x}_j),\quad \boldsymbol{x}\in\ell_0\left(\mathbb{R}^{h+1}\right). \tag{3.22}$$

Since functions in $\mathcal{L}_q$ vanish in a neighborhood of zero, the series has finitely many non zero terms and the function ϕ is uniquely determined by f and will be denoted ϕ^f. We define a pseudometric ρ on $\mathcal{H}_q$ by

$$\rho^2(f,g) = \rho_h^2(\phi^f, \phi^g) = \boldsymbol{\nu}_{0,h}\big(\{\phi^f - \phi^g\}^2\big). \tag{3.23}$$

Recall that $\mathcal{G}_0^* \subset \mathcal{L}_q$, $\mathcal{G}^* = \{\phi_s, \phi \in \mathcal{G}_0^*, s \in [s_0, t_0]\} \subset \mathcal{L}_q$. Define the subclass $\mathcal{F}$ of $\mathcal{H}_q$ which corresponds to the subclass $\mathcal{G}^*$ of $\mathcal{L}_q$, that is $\mathcal{F} = \{f : f(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \phi_s(\boldsymbol{x}_j),\ \phi \in \mathcal{G}_0^*,\ s \in [s_0, t_0]\}$. For $f \in \mathcal{F}$ define $Tf : \mathbb{R} \times \ell_0(\mathbb{R}^{h+1}) \to \mathbb{R}$ by $Tf(t, \boldsymbol{x}) = tf(\boldsymbol{x})$. Set $\mathbb{X}_{n,i}^\dagger = (X_{(i-1)r_n+1}^\dagger, \dots, X_{ir_n+h}^\dagger)/u_n$, $i = 1, \dots, m_n$ and let $\mathbb{X}_{n,1}$ has the same distribution as $\mathbb{X}_{n,1}^\dagger$. Define the process

$$\mathbb{Z}_n^\dagger(Tf) = \frac{1}{\sqrt{n\bar{F}(u_n)}} \sum_{i=1}^{m_n} Tf\big(\xi_i, \mathbb{X}_{n,i}^\dagger\big), \quad f \in \mathcal{F}.$$

We note that

$$\mathbb{Z}_{n,\xi}^\dagger\big(\phi_s^f\big) = \mathbb{Z}_n^\dagger(Tf_s), \qquad f_s(\boldsymbol{x}) = \sum_{j\in\mathbb{Z}} \phi_s(\boldsymbol{x}_j).$$

We apply Theorem A.2 to the process $\mathbb{Z}_n^\dagger$ indexed by $T\mathcal{F} = \{Tf : f \in \mathcal{F}\}$, equipped with the semi-metric $\rho_T(Tf, Tg) = \mathbb{E}[\xi_1^2]\rho(f,g) = \rho(f,g)$.

- The pointwise separability of $\mathcal{G}^*$ implies that $\mathcal{F}$ and hence $T\mathcal{F}$ are also pointwise separable.
- $(\mathcal{F}, \rho)$ is totally bounded since $(\mathcal{G}^*, \rho_h)$ is totally bounded by assumption. This implies that $(T\mathcal{F}, \rho_T)$ is totally bounded.

Next:

(i) For the envelope $|T\mathcal{F}|$ of $T\mathcal{F}$ we have

$$\begin{aligned}
&\frac{m_n}{n\bar{F}(u_n)} \mathbb{E}\big[|T\mathcal{F}|^2(\xi_1, \mathbb{X}_{n,1}) \mathbb{1}\big\{|T\mathcal{F}|(\xi_1, \mathbb{X}_{n,1}) > \epsilon\sqrt{r_n\bar{F}(u_n)}\big\}\big] \\
&\quad \le A_1^2 \frac{1}{r_n\bar{F}(u_n)} \mathbb{E}\big[\Phi_{\mathcal{G}^*}^2(\mathbb{X}_{n,1}) \mathbb{1}\big\{\big|\Phi_{\mathcal{G}^*}(\mathbb{X}_{n,1})\big| > \epsilon\sqrt{r_n\bar{F}(u_n)}/A\big\}\big] \\
&\qquad + \mathbb{E}\big[\xi_1^2 \mathbb{1}\big\{|\xi_1| > A\big\}\big] \frac{\mathbb{E}[\Phi_{\mathcal{G}^*}^2(\mathbb{X}_{n,1})]}{r_n\bar{F}(u_n)}.
\end{aligned} \tag{3.24}$$

We assumed that $\Phi_{\mathcal{G}^*} = \sup_{\phi\in\mathcal{G}^*} |\phi|$ is in $\mathcal{L}_q$. Therefore, the term in the second last line converges to zero as $n \to \infty$ for each A by (3.17). Also,

$$\frac{1}{r_n\bar{F}(u_n)} \mathbb{E}\big[\Phi_{\mathcal{G}^*}^2(\mathbb{X}_{n,1})\big] = \frac{1}{r_n\bar{F}(u_n)} \mathbb{E}\Big[\sup_{\phi\in\mathcal{G}_0^*} \sup_{s\in[s_0,t_0]} \Phi_{n,1}^2(s)\Big].$$

Again, since we assumed that the envelope function is in $\mathcal{L}_q$, (3.16) and $\mathbb{E}[\xi_1^2] < \infty$ imply that the term in (3.24) vanishes by letting $n \to \infty$ and then $A \to \infty$. Therefore, the Lindeberg condition (A.2) holds.

(ii) Condition (A.3) follows directly from Kulik et al. [18].

(iii) Define $(Tf)_+ : \mathbb{R}_+ \times \ell_0(\mathbb{R}^{h+1}) \to \mathbb{R}$ by $(Tf)_+(t, \boldsymbol{x}) = tf(\boldsymbol{x})$ and analogously $(Tf)_- : \mathbb{R}_- \times \ell_0(\mathbb{R}^{h+1}) \to \mathbb{R}$. Then $T\mathcal{F} = T_+\mathcal{F} \cup T_-\mathcal{F}$, where $T_+\mathcal{F} = \{(Tf)_+ : f \in \mathcal{F}\}$ and analogously $T_-\mathcal{F}$. Both $T_+\mathcal{F}$ and $T_-\mathcal{F}$ are linearly ordered and hence are VC-subgraph classes. Thus, $T\mathcal{F}$ is a finite union of VC-subgraph classes and Lemma A.3 applies. □

3.2.4. *Weak convergence* (*conditional*)

To conclude the proof of Proposition 3.6, we need to justify that the conditional fidi convergence of Lemma 3.10 and the unconditional tightness of Lemma 3.11 yield the conditional weak convergence. The argument is rather standard, but we provide it for completeness in Appendix B.

3.2.5. *Multiplier process with random normalization*

Next, we consider

$$\widetilde{\mathbb{M}}_{n,\xi}(\phi_s) = \sqrt{n\bar{F}(u_n)} \left\{ \frac{\sum_{i=1}^{m_n}(1+\xi_i)\Phi_{n,i}(s)}{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon_{n,i}(s)} - \frac{\sum_{i=1}^{m_n}\Phi_{n,i}(s)}{\sum_{i=1}^{m_n}\Upsilon_{n,i}(s)} \right\}.$$

Proposition 3.12. *Let $\mathcal{G}^*$ be as in Theorem 3.1. Then*

$$\widetilde{\mathbb{M}}_{n,\xi}(\phi_s) \quad \Rightarrow \quad s^\alpha \mathbb{M}(\phi_s) - s^\alpha \boldsymbol{\nu}_{0,h}(\phi)\mathbb{M}\big(\mathbb{1}_{(s,\infty)\times\mathbb{R}^h}\big)$$

in $\ell_\infty(\mathcal{G}^*)$ in probability.

Proof. It suffices to consider the independent blocks process $\widetilde{\mathbb{M}}^\dagger_{n,\xi}(\phi_s)$ that we write as

$$\sqrt{n\bar{F}(u_n)}\frac{\sum_{i=1}^{m_n}\xi_i\Phi^\dagger_{n,i}(s)}{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon^\dagger_{n,i}(s)} - \sqrt{n\bar{F}(u_n)}\frac{\sum_{i=1}^{m_n}\xi_i\Upsilon^\dagger_{n,i}(s)}{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon^\dagger_{n,i}(s)}\frac{\sum_{i=1}^{m_n}\Phi^\dagger_{n,i}(s)}{\sum_{i=1}^{m_n}\Upsilon^\dagger_{n,i}(s)}. \tag{3.25}$$

Now, we consider both parts in (3.25) separately and want to write them in terms of $\mathbb{M}_{n,\xi}(\phi_s)$ defined in (3.10).

The first term in (3.25) is

$$\frac{n\bar{F}(u_n)}{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon^\dagger_{n,i}(s)}\mathbb{Z}^\dagger_{n,\xi}(\phi_s).$$

We have $\mathbb{Z}^\dagger_{n,\xi}(\phi_s) \Rightarrow \mathbb{M}(\phi_s)$ in $\ell_\infty(\mathcal{G}^*)$ in probability by Proposition 3.6. Writing (cf. (3.11) and (3.12))

$$\frac{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon^\dagger_{n,i}(s)}{n\bar{F}(u_n)} = \frac{1}{\sqrt{n\bar{F}(u_n)}}\mathbb{Z}^\dagger_{n,\xi}(\mathcal{I}_s) + \frac{\sum_{i=1}^{m_n}\Upsilon^\dagger_{n,i}(s)}{n\bar{F}(u_n)}$$

and noting that $\sum_{i=1}^{m_n}\Upsilon^\dagger_{n,i}(s)/(n\bar{F}(u_n)) \xrightarrow{p} s^{-\alpha}$ yields that the first term in (3.25) converges to $s^\alpha\mathbb{M}(\phi_s)$ in $\ell_\infty(\mathcal{G}^*)$ in probability.

For the second term in (3.25), by (3.13) and (3.14) and since $\boldsymbol{\nu}_{0,h}(\phi_s) = s^{-\alpha}\boldsymbol{\nu}_{0,h}(\phi)$, the ratio $\sum_{i=1}^{m_n}\Phi^\dagger_{n,i}(s)/\sum_{i=1}^{m_n}\Upsilon^\dagger_{n,i}(s)$ converges in probability to $\boldsymbol{\nu}_{0,h}(\phi)$. We recognize

$$\sqrt{n\bar{F}(u_n)}\frac{\sum_{i=1}^{m_n}\xi_i\Upsilon^\dagger_{n,i}(s)}{\sum_{i=1}^{m_n}(1+\xi_i)\Upsilon^\dagger_{n,i}(s)}$$

as the one in (3.25), with Φ replaced with Υ.

Bearing in mind that all convergences hold jointly (cf. Remark 3.5), combination of the convergence of the first and the second term yields the first result. □

3.2.6. *Multiplier process with random levels - conclusion of the proof of Theorem 3.1*

Proof. Set $\zeta_n = X_{n:n-k}/u_n$ and recall that $k = n\bar{F}(u_n)$. For the first statement note that $\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\phi) = \widetilde{\mathbb{M}}_{n,\xi}(\phi_{\zeta_n})$ and $\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\psi_{1,y}) = \widetilde{\mathbb{M}}_{n,\xi}(\psi_{\zeta_n,y})$. The result follows from Proposition 3.12 and $\zeta_n \xrightarrow{p} 1$.

For the second statement note that $\Phi_{n,i} = \Phi_{n,i}(\zeta_n)$ and write

$$\widehat{\mathbb{M}}_{n,\xi}(\phi) = \frac{1}{\sqrt{k}}\left(\sum_{i=1}^{m_n} \xi_i \left\{\Phi_{n,i}(\zeta_n) - \mathbb{E}_{\zeta_n}\left[\Phi_{n,i}(\zeta_n)\right]\right\}\right) + \mathbb{E}_{\zeta_n}\left[\Phi_{n,1}(\zeta_n)\right]\frac{1}{\sqrt{k}}\sum_{i=1}^{m_n} \xi_i, \tag{3.26}$$

where $\mathbb{E}_{\zeta_n}[\Phi_{n,1}(\zeta_n)]$ denotes a random quantity obtained by evaluating the function $s \mapsto \mathbb{E}[\Phi_{n,1}(s)]$ at $s = \zeta_n$. Thus, using (3.12),

$$\widehat{\mathbb{M}}_{n,\xi}(\phi) = \mathbb{M}_{n,\xi}(\phi_{\zeta_n}) + \underbrace{\frac{\mathbb{E}_{\zeta_n}[\Phi_{n,1}(\zeta_n)]}{r_n\bar{F}(u_n)}}_{=O_P(1)} \frac{r_n\bar{F}(u_n)\sqrt{m_n}}{\sqrt{n\bar{F}(u_n)}} \underbrace{\frac{\sum_{i=1}^{m_n}\xi_i}{\sqrt{m_n}}}_{=O_P(1)} = \mathbb{M}_{n,\xi}(\phi_{\zeta_n}) + O_P(1)\sqrt{r_n\bar{F}(u_n)}.$$

Proposition 3.6 and the assumption $r_n\bar{F}(u_n) \to 0$ finish the proof for the class $\mathcal{G}_0^*$. □

4. Simulation setup

In this section, we illustrate the performance of different bootstrap procedures and asymptotic approaches to construct 95%-confidence intervals for tail array sums by means of coverage rates. We consider bootstrap strategies that appear to be more or less natural for the purpose of conducting statistical inference for tail array sums. If available, we also use asymptotic approaches based on central limit theorems. Precisely, we apply

(i) a multiplier (block) bootstrap *without* randomizing k corresponding to $\widehat{\mathbb{M}}_{n,\xi}(\phi)$ in (3.4);
(ii) a multiplier (block) bootstrap *with* randomizing k corresponding to $\widehat{\widehat{\mathbb{M}}}_{n,\xi}(\phi)$ in (3.5);
(iii) an i.i.d. (or moving block) bootstrap.

We note that the latter bootstrap approach is included for numerical comparison, but it has not been justified theoretically.

To compare the performance of these bootstrap approaches, we consider two prominent members of the family of tail array sums. Precisely, we address the Hill estimator and the extremogram; see Section 3.1. We consider several data generating procedures. First, we will use i.i.d. data as a benchmark for our simulation studies, before we consider time series data. We will consider the realization $X_1, \ldots, X_n$ of one of the following models:

(I) $\{X_j\}$ i.i.d. with Pareto distribution with index $\alpha = 4$;
(II) $\{X_j\}$ i.i.d. with t-distribution with $\alpha = 4$ degrees of freedom;
(III) AR(1) model $\{X_j\}$ with $X_j = \rho X_{j-1} + \varepsilon_j$, where ε_j are i.i.d. t-distributed with $\alpha = 4$ degrees of freedom and AR coefficient $\rho = 0.8$;
(IV) ARCH(1) model $\{X_j\}$ with $X_j = \sqrt{\beta + \lambda X_{j-1}^2}\varepsilon_j$, where ε_j are i.i.d. standard normal, $\beta = 1.9 * 10^{-5}$ and $\lambda = 0.7$.

In models (I)–(III), we used $\alpha = 4$ leading to tail index also equal to 4. In the latter case of model (IV), we used parameters that lead to tail index $\alpha = 3.18$; see Embrechts et al. [13], Table 8.4.8.

For each data generating process, we conducted Monte Carlo studies for two different sample sizes $n \in \{500, 1000\}$. In the time series case, where either a multiplier block bootstrap or a moving block bootstrap is required to capture the dependence structure, we used block lengths $L = 20$ for $n = 500$ and $L = 30$ for $n = 1000$, respectively. Further simulation results (not reported in this paper) indicate that a too small block length does not capture the dependence structure sufficiently well leading to a performance loss, which is also observed for block length chosen too large due to less variation in the bootstrap sample. In each setting, we used $M = 1000$ Monte Carlo samples and $B = 1000$ bootstrap replications.

In general, the purpose of the following simulation studies is two-fold. First, we want to illustrate our theoretical findings of the previous sections. In particular, these include (a) consistency for the multiplier (block) bootstrap *with* randomizing k and (b) inconsistency for the multiplier (block) bootstrap *without* randomizing k. Second, we want to study (c) the potential of finite sample performance of the multiplier bootstrap in comparison to a naïve application of an i.i.d. or moving block bootstrap. Additionally, in cases where feasible (i.e., Hill estimator for i.i.d. and for AR(1) model), we compare these results also with asymptotic confidence intervals based on normal approximation from central limit theorems.

The main findings of our simulation studies are as follows:

- As suggested by the theory, the multiplier bootstrap *without* randomization of k does not yield proper confidence intervals. More precisely, the confidence intervals turn out to be systematically too wide.
- The multiplier bootstrap *with* randomization of k performs at least as good as the i.i.d. or moving block bootstrap and outperforms the latter in most of the cases.
- The computational burden of i.i.d. and moving block bootstraps is considerably higher than of a multiplier bootstrap. This is explained by the computationally demanding sorting routine that has to be applied only once for the original sample for the multiplier bootstraps, but additionally also for each bootstrap sample for i.i.d. and moving block bootstrap.

In view of the findings raised above, we recommend using the multiplier bootstraps with randomizing k for statistical inference of tail array sums with random levels.

4.1. Bootstrap performance: Hill estimator

We recall that in the i.i.d. case, the asymptotic distribution of the Hill estimator (3.9) of $\gamma = 1/\alpha$ is normal with variance α^{-2}. In the AR(1) case with the positive coefficient, the limiting variance is $\alpha^{-2}(\frac{1+\rho^\alpha}{1-\rho^\alpha})$.

In the left panels of Figures 1 and 2, we show realizations of Hill plots for the Hill estimator of $\gamma = 1/\alpha$ with different bootstrap confidence intervals for models (I) and (II), respectively. Further, based on Monte Carlo simulations, coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) are reported.

Starting with a discussion of Figure 1 showing the results for i.i.d. Pareto-distributed data, it is clearly visible that the Hill estimator $\widehat{\gamma}$ of $\gamma = 1/\alpha$ is doing a good job in estimating the target $\gamma = 0.25$ for a broad range of k's. All kinds of confidence intervals appear to be plausible. All of them are rather near to each other, but a closer look reveals that in particular the multiplier bootstrap *without* randomizing k leads to systematically wider confidence intervals than the other approaches. This observation is supported by the corresponding coverage rates and the mean lengths of the con-

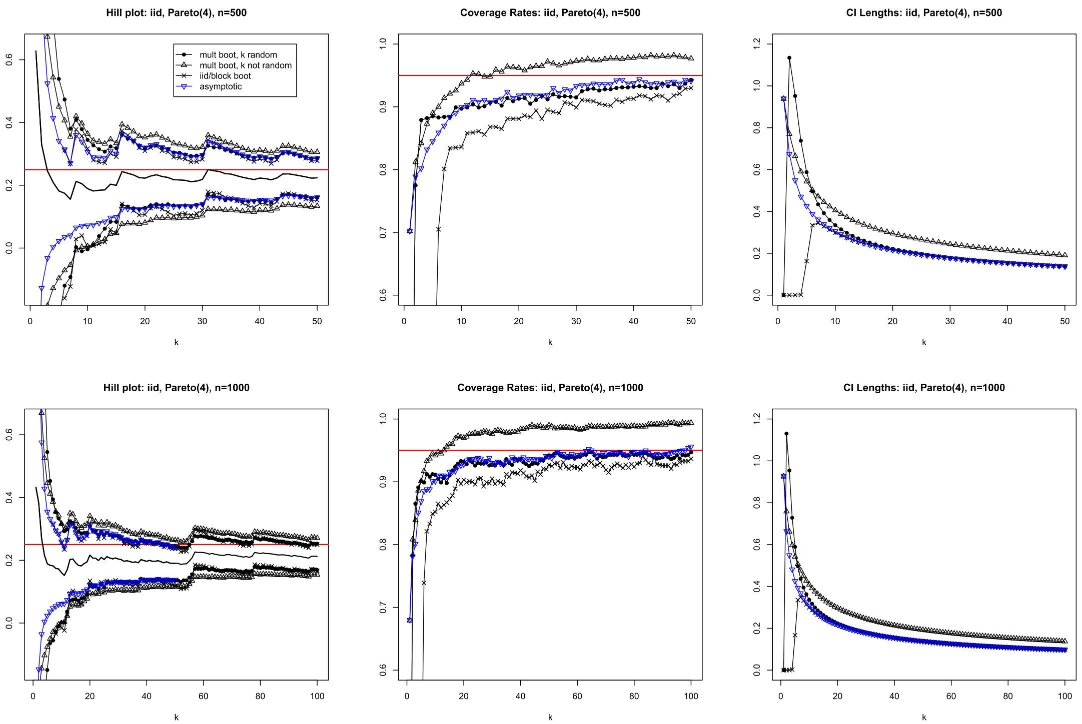

Figure 1. Model (I), i.i.d., Pareto(4): Typical realizations of Hill plots for the Hill estimator with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \leq 0.1 * n$ and with $n = 500$ (upper panels) and $n = 1000$ (lower panels). Results for multiplier bootstrap with randomizing k, multiplier bootstrap without randomizing k, i.i.d. bootstrap and asymptotic confidence intervals are reported. The targets $\gamma = 1/\alpha = 0.25$ and 95%, respectively, are marked with red horizontal lines.

fidence intervals. Whereas all other procedures tend to slightly understate the desired coverage rate of 95% which improves for larger k's, the multiplier bootstrap without randomizing k systematically overstates the 95% coverage rate for all sufficiently large k's. This behavior nicely supports the bootstrap inconsistency as argued in Section 3. A pairwise comparison of multiplier bootstraps with and without randomizing k leads to the conclusion that randomizing k indeed cures this issue leading to (asymptotically) valid results. Interestingly, the confidence intervals produced by a naïve application of an i.i.d. bootstrap, which is computationally a lot more demanding than the multiplier bootstrap which can be efficiently implemented, tend to be systematically too small over the whole range of the k's. Increasing the sample size from $n = 500$ to $n = 1000$ clearly shows an expected and desired pattern. For larger sample size the confidence intervals tend to be narrower and the understatement of the 95% coverage rate is improved as well. This includes also the i.i.d. bootstrap, which, however, is still outperformed by the other valid approaches. Note that for a direct comparison of top and bottom panels on each figure one has to take into account that the range of k's is $\{1, \ldots, 50\}$ for $n = 500$ and $\{1, \ldots, 100\}$ for $n = 1000$. The asymptotic confidence intervals essentially coincide with the valid bootstrap confidence intervals (i.i.d. bootstrap and multiplier bootstrap with randomized k) and show a good performance.

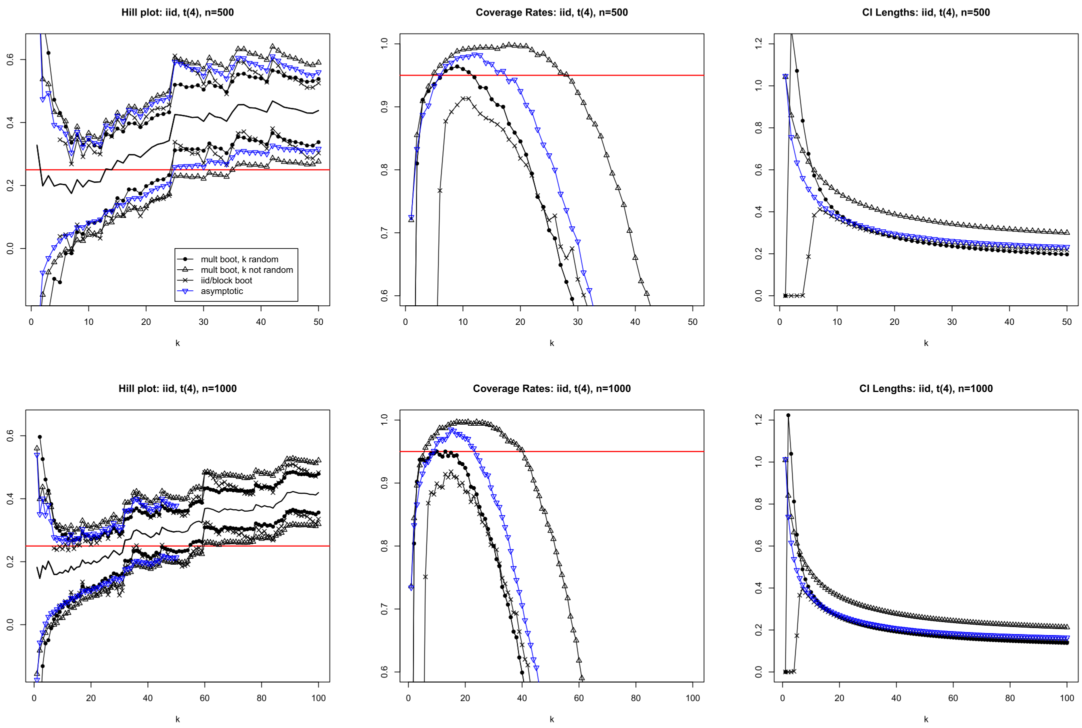

Figure 2. Model (II), i.i.d., t(4): Typical realizations of Hill plots for the Hill estimator with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \leq 0.1 * n$ and with $n = 500$ (upper panels) and $n = 1000$ (lower panels). Results for multiplier bootstrap with randomizing k, multiplier bootstrap without randomizing k, i.i.d. bootstrap and asymptotic confidence intervals are reported. The targets $\gamma = 1/\alpha = 0.25$ and 95%, respectively, are marked with red horizontal lines.

The picture for Figure 2 showing the results for t-distributed data becomes different as the Hill estimator is now doing a good job in estimating the target $\gamma = 0.25$ only for a considerably narrower range of k's. Apparently, the Hill estimators show a pronounced bias for larger k's. Comparing the Hill plots for samples sizes $n = 500$ and $n = 1000$, it seems that the range of k's leading to satisfactory estimation results becomes somewhat larger which is surely plausible. Considering the coverage rates in Figure 2, similar to the plots in Figure 1, the multiplier bootstrap without randomizing k overstates the desired coverage rate of 95% for a broad range of k's, whereas the multiplier bootstrap with randomizing k is close to 95% for some range of k's without overstating it. In contrast to the results for the Pareto distribution shown in Figure 1, the performance of the asymptotic confidence intervals considerably decreases. These intervals tend to be too wide for a broad range of k's. This phenomenon could be explained by the bias when estimating the tail index (which is the main ingredient for the asymptotic confidence interval) from data following a t-distribution. The pattern for the i.i.d. bootstrap is similar to that one observed already in Figure 1. It shows a systematic understatement of the desired coverage rate of 95% which improves with increasing sample size, but it is clearly outperformed by the multiplier bootstrap with randomizing k. In particular, due to the bad performance and the large computational demand in comparison to the multiplier bootstrap with randomizing k, the i.i.d. bootstrap turns out to be not advisable.

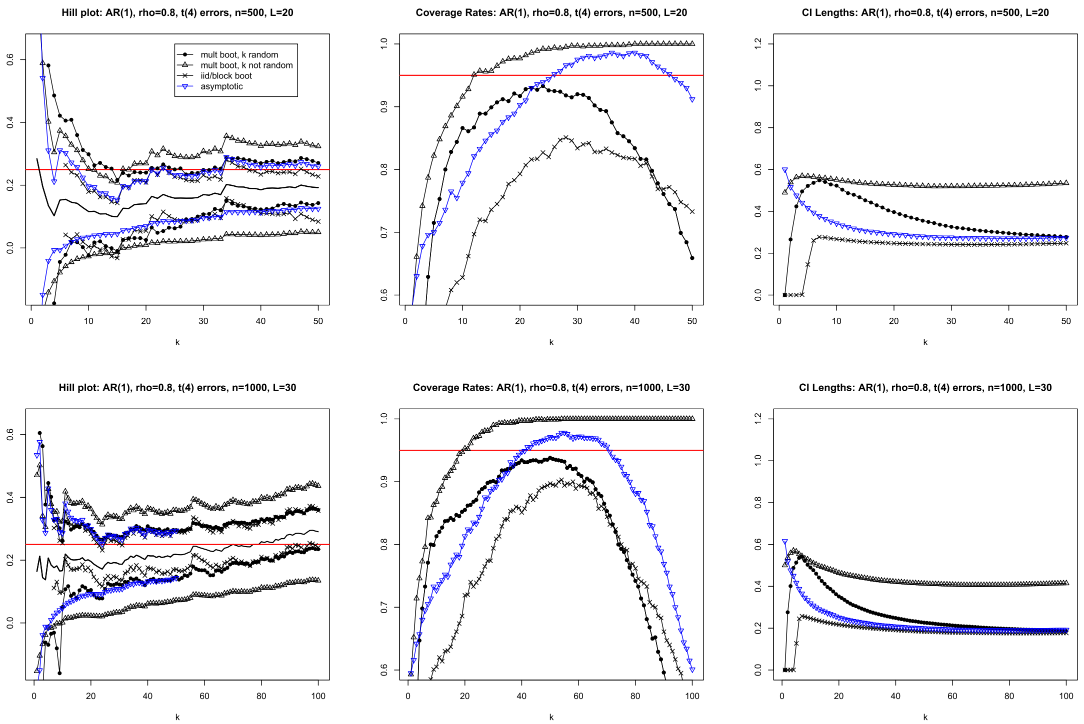

Figure 3. Model (III), AR(1) $X_j = 0.8X_{j-1} + \varepsilon_j$, $\varepsilon_j \sim t(4)$. Typical realization of Hill-plots for the Hill estimator with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \le 0.1 * n$ and with $L = 20$ for $n = 500$ (upper panels) and with $L = 30$ for $n = 1000$ (lower panels). Results for multiplier block bootstrap with randomizing k, multiplier block bootstrap without randomizing k, moving block bootstrap and asymptotic confidence intervals are reported. The targets $\gamma = 1/\alpha = 0.25$ and 95%, respectively, are marked with red horizontal lines.

The general picture for the AR(1) model with t-innovations shown in Figure 3 is similar to the pattern described above. Again the multiplier bootstrap without randomizing k overstates the desired coverage rate of 95% for a broad range of k's, whereas the multiplier bootstrap with randomizing k is close to 95% for some range of k's without overstating it. The results improve for increasing sample size. Also, the intervals from an asymptotic approximation tend to be too wide for a broad range of k's. We used $L = 20$ for $n = 500$ and $L = 30$ for $n = 1000$. We conducted simulations for a broader range of block length, where the results were actually quite stable. We observed a loss in performance for too small block lengths that did not allow to capture the serial dependence in the data. We did not observe a big difference of the results for example, between $L = 20$ and $L = 30$.

For the ARCH(1) case in Figure 4, again the picture is the same as before. Here, it might be advisable to actually use larger block lengths as we observed somewhat better performance when using $L = 30$ for $n = 500$ (not shown here) instead of $L = 20$. We note that for the ARCH(1) model the formula for the asymptotic variance of the Hill estimator is given by a complicated infinite sum and hence the confidence intervals are not displayed here.

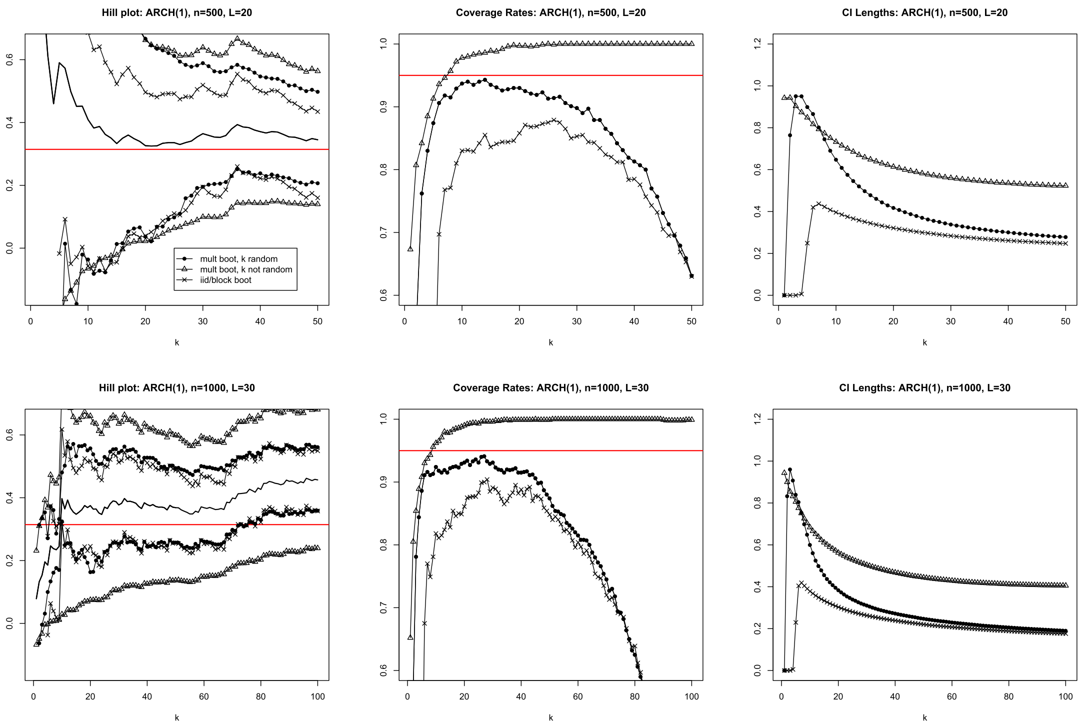

Figure 4. Model (IV), ARCH(1) $X_j = \sqrt{1.9 * 10^{-5} + 0.7X_{j-1}^2}\varepsilon_j$, $\varepsilon_j \sim \mathcal{N}(0,1)$. Typical realization of Hill–plots for the Hill estimator with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \leq 0.1 * n$ and with $L = 20$ for $n = 500$ (upper panels) and with $L = 30$ for $n = 1000$ (lower panels). Results for multiplier block bootstrap with randomizing k, multiplier block bootstrap without randomizing k and block bootstrap confidence intervals are reported. The targets $\gamma = 1/\alpha = 0.31$ and 95%, respectively, are marked with red horizontal lines.

4.2. Bootstrap performance: Extremogram

In the left panels of Figures 5 and 6, we show typical realization of Hill-plots for the extremogram at lags $h = 1, 2, 3$ with different bootstrap confidence intervals for sample sizes $n = 500$ and $n = 1000$, respectively. Based on Monte Carlo simulations, coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) are reported.

The general pattern in Figures 5 and 6 is comparable to those observed already above. The inconsistent dependent multiplier bootstrap without randomizing k produces too large intervals and the actual coverage rate overstates systematically the target of 95%. The consistent dependent multiplier bootstrap with randomizing k performs quite well and gets close to the target for a broad range of k's. However, the performance decreases when increasing the lag of the extremogram. This phenomenon was expected as, due to the rather small sample sizes considered here, quite few summands actually do not vanish. The block bootstrap suffers from the same problem observed already above for the Hill estimator. It produces confidence intervals that are far too small leading to considerable understatement of the coverage rates.

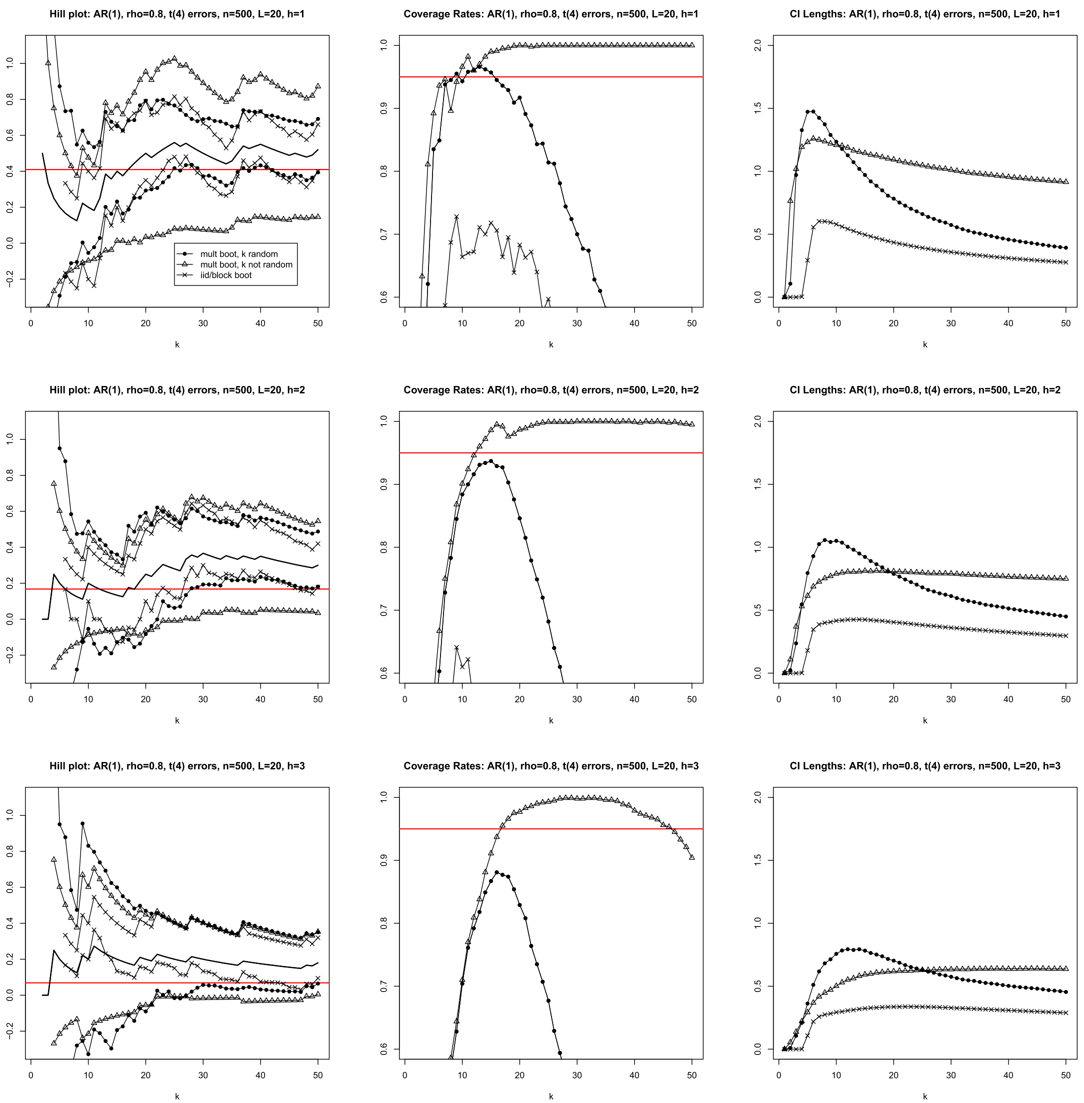

Figure 5. Model (III), AR(1) $X_j = 0.8X_{j-1} + \varepsilon_j$, $\varepsilon_j \sim t(4)$. Typical realization of Hill-plots for extremograms with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \leq 0.1 * n$ and $n = 500$ at lags $h \in \{1, 2, 3\}$ (from top to bottom). Results for multiplier block bootstrap with randomizing k, multiplier block bootstrap without randomizing k and block bootstrap are reported. The targets $\gamma = 0.8^{4*h}$, $h = 1, 2, 3$ (from top to bottom) and 95%, respectively, are marked with red horizontal lines.

5. Conclusion

In this paper, we studied validity of the multiplier block bootstrap for tail array sums based on regularly varying time series. We showed that the *natural*, but naïve block-multiplier procedure does not yield a

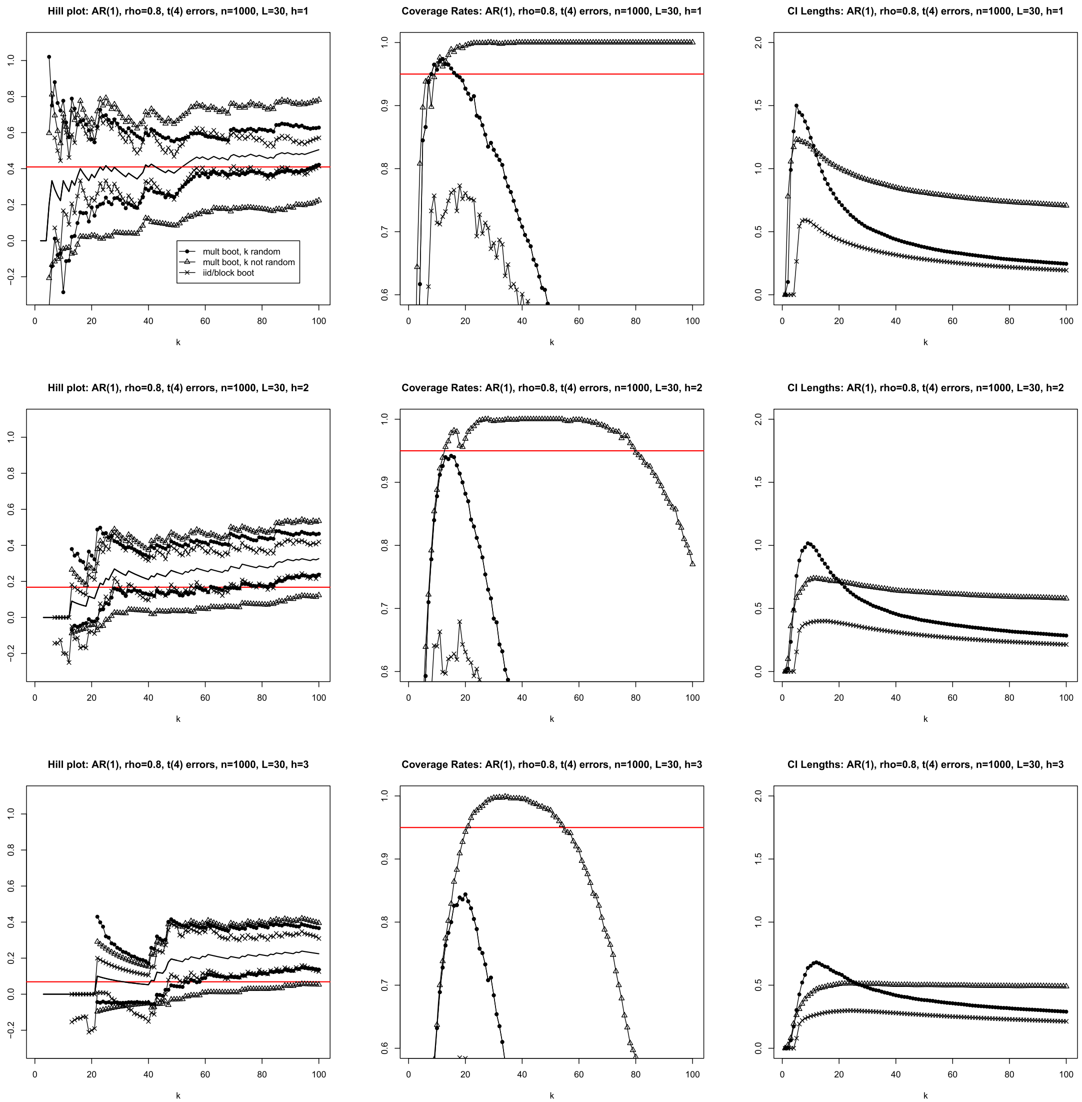

Figure 6. Model (III), AR(1) $X_j = 0.8X_{j-1} + \varepsilon_j$, $\varepsilon_j \sim t(4)$. Typical realization of Hill-plots for extremograms with different bootstrap confidence intervals (left panels) and coverage rates (center panels) and the resulting mean lengths of confidence intervals (right panels) from Monte Carlo studies for $k \leq 0.1 * n$ and $n = 1000$ at lags $h = 1, 2, 3$ (from top to bottom). Results for multiplier block bootstrap with randomizing k, multiplier block bootstrap without randomizing k and block bootstrap are reported. The targets $\gamma = 0.8^{4*h}$, $h = 1, 2, 3$ (from top to bottom) and 95%, respectively, are marked with red horizontal lines.

valid asymptotic distribution. This is fixed by the proper randomization of the number of order statistics k. We also justified the use of the multiplier block bootstrap numerically.

A validity of the moving block bootstrap for tail array sums with random levels remains open.

We also do not discuss the issue of choosing k and r_n. For geometrically ergodic Markov chains the blocks can be chosen as $r_n = \log(n)$.

Appendix A: Convergence in ℓ_∞

Theorem A.1 (Giné and Nickl [14], Theorem 3.7.23). *Let $\{\mathbb{Z}_n, n \in \mathbb{N}\}$, be a sequence of processes with values in $\ell_\infty(\mathcal{F})$. Then the following statements are equivalent.*

(i) *The finite dimensional distributions of the processes $\mathbb{Z}_n$ converge in law and there exists a pseudometric ρ on $\mathcal{F}$ such that $(\mathcal{F}, \rho)$ is totally bounded and for all $\epsilon > 0$,*

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\mathbb{P}^*\Big(\sup_{\rho(f,g)<\delta}\big|\mathbb{Z}_n(f)-\mathbb{Z}_n(g)\big|>\epsilon\Big)=0. \tag{A.1}$$

(ii) *There exists a process $\mathbb{Z}$ whose law is a tight Borel probability measure on $\ell_\infty(\mathcal{F})$ and such that $\mathbb{Z}_n \Rightarrow \mathbb{Z}$ in $\ell_\infty(\mathcal{F})$.*

Moreover, if (i) holds, then the process $\mathbb{Z}$ in (ii) has a version with bounded uniformly continuous paths for ρ.

The following result provides a sufficient condition for (A.1) above. Let $\{X_{n,i}, 1 \le i \le m_n\}$, $n \ge 1$, be an array of row-wise i.i.d. random elements in a measurable space $(\mathsf{X}, \mathcal{X})$ and define $Z_{n,i}(f) = f(X_{n,i})$, $f \in \mathcal{F}$. Let a_n be a non decreasing sequence and $\mathcal{F}$ be a set of measurable functions defined on X. Define the random pseudometric d_n on $\mathcal{F}$ by

$$d_n^2(f,g)=\frac{1}{a_n^2}\sum_{i=1}^{m_n}\big\{f(X_{n,i})-g(X_{n,i})\big\}^2, \quad f,g\in\mathcal{F}.$$

Let $N(\epsilon, \mathcal{F}, d_n)$ be the minimum number of balls in the pseudometric d_n needed to cover $\mathcal{F}$. Let $\mathbb{Z}_n$ be the empirical process defined by

$$\mathbb{Z}_n(f)=\frac{1}{a_n}\sum_{i=1}^{m_n}\big\{f(X_{n,i})-\mathbb{E}\big[f(X_{n,i})\big]\big\}, \quad f\in\mathcal{F}.$$

Define finally the sup-norm $\|H\|_{\mathcal{F}} = \sup_{f\in\mathcal{F}} |H(f)|$ for any functional H on $\mathcal{F}$. If $\mathcal{F}$ is a pseudometric space and H is measurable on $\mathcal{F}$ then the separability of $\mathcal{F}$ implies that $\|H\|_{\mathcal{F}}$ is measurable.

Theorem A.2 (Adapted from van der Vaart and Wellner [21], Theorem 2.11.1). *Assume that the pseudometric space $\mathcal{F}$ is totally bounded and pointwise separable.*

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

$$\lim_{n\to\infty} a_n^{-2} m_n \mathbb{E}\big[\|Z_{n,1}\|_{\mathcal{F}}^2 \mathbb{1}\big\{\|Z_{n,1}\|_{\mathcal{F}} > \eta a_n\big\}\big]=0. \tag{A.2}$$

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

$$\lim_{n\to\infty}\sup_{\substack{f,g\in\mathcal{F}\\ \rho(f,g)\le\delta_n}}\mathbb{E}\big[d_n^2(f,g)\big]=0, \tag{A.3}$$

$$\int_0^{\delta_n}\sqrt{\log N(\epsilon,\mathcal{F},d_n)}\,\mathrm{d}\epsilon \xrightarrow{\mathbb{P}} 0. \tag{A.4}$$

Then $\mathbb{Z}_n$ is asymptotically ρ-equicontinuous, that is, (A.1) holds.

A sufficient condition for (A.4) is provided by Giné and Nickl [14], Theorem 3.7.37.

Lemma A.3. *If $\mathcal{F}$ is linearly ordered or if $\mathcal{F}$ is a VC-subgraph class of functions or a union of such classes, then* (A.4) *holds.*

Appendix B: Conditional weak convergence

Let μ be the law of the limiting process $\{\mathbb{M}(\phi_s), s \in [s_0, t_0]\}$ considered as a random element in $\ell_\infty(\mathcal{G}^*)$, where $\mathcal{G}^* = \{\phi_s, \phi \in \mathcal{G}_0^*, s \in [s_0, t_0]\}$ is introduced in Theorem 3.1. Let $\mathcal{F}_n^X$ be the sigma-field generated by $(X_1, \dots, X_n)$ and let μ_n be the conditional law of $\{\mathbb{M}_{n,\xi}^\dagger(\phi_s), s \in [s_0, t_0]\}$ given $\mathcal{F}_n^X$ (see Equation (3.15) for the definition of $\mathbb{M}_{n,\xi}^\dagger$). We will prove that

$$d_{\mathbb{BL}}(\mu_n, \mu) \xrightarrow{p} 0, \tag{B.1}$$

where $d_{\mathbb{BL}}$ is the uniform distance on $\mathbb{BL}(\ell_\infty(\mathcal{G}^*))$. For this, we will use a triangular argument.

Recall that each function ϕ_s from $\mathcal{G}^*$ induces the summation functional $f_s \in \mathcal{F}$; cf. (3.22). For $f_s(\boldsymbol{x}) = \sum_j \phi_s(\boldsymbol{x}_j) \in \mathcal{F}$ we recall the metric ρ from (3.23): $\rho^2(f_s, f_t) = \boldsymbol{\nu}_{0,h}(\{\phi_s - \phi_t\}^2)$.

Next, for $\delta > 0$, let

$$N(\delta) = \inf\{j \in \mathbb{N} : s_0 + j\delta/(\boldsymbol{\nu}_{0,h}(\phi^2)\alpha s_0^{-\alpha-1}) > t_0\}.$$

Define

$$t_j = s_0 + j\delta/(\boldsymbol{\nu}_{0,h}(\phi^2)\alpha s_0^{-\alpha-1}), \quad j = 0, \dots, N(\delta) - 1, t_{N(\delta)} = t_0.$$

The α-homogeneity of $\boldsymbol{\nu}_{0,h}$ yields for $s \ge s_0$ such that $s + \delta \le t_0$

$$\begin{aligned} \rho^2(f_s, f_{s+\delta}) &= \boldsymbol{\nu}_{0,h}(\{\phi_s - \phi_{s+\delta}\}^2) \le |\boldsymbol{\nu}_{0,h}(\phi_s^2) - \boldsymbol{\nu}_{0,h}(\phi_{s+\delta}^2)| \\ &= \boldsymbol{\nu}_{0,h}(\phi^2)\{s^{-\alpha} - (s+\delta)^{-\alpha}\} \le \boldsymbol{\nu}_{0,h}(\phi^2)\alpha s_0^{-\alpha-1}\delta. \end{aligned}$$

Thus,

$$\sup_{s,t \in [t_j, t_{j+1}]} \rho^2(f_s, f_{s+\delta}) \le \delta. \tag{B.2}$$

We define the process $\mathbb{M}_{n,\xi}^\delta$ indexed by $\mathcal{G}^*$ by

$$\mathbb{M}_{n,\xi}^\delta(\phi_s) = \mathbb{M}_{n,\xi}^\dagger(\phi_{t_j}), \quad s \in [t_j, t_{j+1}), j = 0, \dots, N(\delta) - 1,$$

and $\mathbb{M}_{n,\xi}^\delta(\phi_{t_0}) = \mathbb{M}_{n,\xi}^\dagger(\phi_{t_0})$. Similarly, we define

$$\mathbb{M}^\delta(\phi_s) = \mathbb{M}(\phi_{t_j}), \quad s \in [t_j, t_{j+1}), j = 0, \dots, N(\delta) - 1,$$

and $\mathbb{M}^\delta(\phi_{t_0}) = \mathbb{M}(\phi_{t_0})$. These processes are random elements in $\ell_\infty(\mathcal{G}^*)$. Let μ_n^δ be the conditional law of $\mathbb{M}_{n,\xi}^\delta$ given $\mathcal{F}_n^X$ and μ^δ be the (unconditional) law of $\mathbb{M}^\delta$.

We prove that for all $\epsilon > 0$,

$$\lim_{\delta\to 0} \mathbb{P}\big(d_{\mathbb{BL}}\big(\mu^{\delta}, \mu\big) > \epsilon\big) = 0, \tag{B.3a}$$

$$\lim_{n\to\infty} \mathbb{P}\big(d_{\mathbb{BL}}\big(\mu_n^{\delta}, \mu^{\delta}\big) > \epsilon\big) = 0, \tag{B.3b}$$

$$\lim_{\delta\to 0} \mathbb{P}\Big(\limsup_{n\to\infty} d_{\mathbb{BL}}\big(\mu_n^{\delta}, \mu_n\big) > \epsilon\Big) = 0. \tag{B.3c}$$

(i) The process $\mathbb{M}(\phi_s)$, $s \in [s_0, t_0]$ is almost surely uniformly continuous, thus (B.3a) holds.

(ii) For a fixed $\delta > 0$, the processes $\mathbb{M}_{n,\xi}^{\delta}$ and $\mathbb{M}^{\delta}$ are random step functions defined on the same fixed grid. Thus, (B.3b) is equivalent to the joint weak convergence of $\mathbb{M}_{n,\xi}^{\delta}(\phi_{t_i})$ to $\mathbb{M}^{\delta}(\phi_{t_i})$, $i = 0, \dots, N(\delta)$. The latter in turn follows from Lemma 3.10 along with the comment before it.

(iii) By the definition of the distance $d_{\mathbb{BL}}$ and Markov inequality it suffices to prove that

$$\lim_{n\to\infty} \mathbb{E}\Big[\sup_{\Psi\in\mathbb{BL}(\ell_\infty(\mathcal{G}^*))} \big|\mathbb{E}_{\xi}\big[\Psi\big(\mathbb{M}_{n,\xi}^{\delta}\big)\big] - \mathbb{E}_{\xi}\big[\Psi\big(\mathbb{M}_{n,\xi}^{\dagger}\big)\big]\big|\Big] = 0.$$

The expression on the last-hand side is bounded by

$$\mathbb{E}\Big[\sup_{s_0\le s\le t_0} \big|\mathbb{M}_{n,\xi}^{\delta}(\phi_s) - \mathbb{M}_{n,\xi}(\phi_s)\big| \wedge 2\Big] \le \mathbb{E}\Big[\sup_{\substack{s_0\le s,t\le t_0\\ \nu_{0,h}(\phi_s,\phi_t)\le\delta}} \big|\mathbb{M}_{n,\xi}^{\dagger}(\phi_s) - \mathbb{M}_{n,\xi}^{\dagger}(\phi_t)\big| \wedge 2\Big].$$

By Lemma 3.11, we have

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \mathbb{P}\Big(\sup_{\substack{s,t\in[s_0,t_0]\\ \nu_{0,h}(\phi_s,\phi_t)<\delta}} \big|\mathbb{M}_{n,\xi}^{\dagger}(\phi_s) - \mathbb{M}_{n,\xi}^{\dagger}(\phi_t)\big| > \epsilon\Big) = 0,$$

which implies

$$\lim_{\delta\to 0}\limsup_{n\to\infty} \mathbb{E}\Big[\sup_{\substack{s,t\in[s_0,t_0]\\ \nu_{0,h}(\phi_s,\phi_t)<\delta}} \big|\mathbb{M}_{n,\xi}^{\dagger}(\phi_s) - \mathbb{M}_{n,\xi}^{\dagger}(\phi_t)\big| \wedge 2\Big] = 0.$$

This proves (B.3c).

Thus, we have proved (B.3a), (B.3b) and (B.3c). By the triangular argument, this concludes the proof.

Acknowledgements

Large parts of this research were conducted while Carsten Jentsch was visiting the University of Ottawa funded by the Ontario/Baden-Württemberg Faculty Mobility Program and during a Research-in-Pairs stay of Carsten Jentsch and Rafal Kulik at the MFO Oberwolfach. Further, parts of the research of Carsten Jentsch were financially supported by the German Research Foundation DFG via the Collaborative Research Center SFB 884 (Project B6) and the Baden-Württemberg-Stiftung via the Eliteprogram for Postdocs.

References

[1] Basrak, B. and Segers, J. (2009). Regularly varying multivariate time series. *Stochastic Process. Appl.* **119** 1055–1080. MR2508565 https://doi.org/10.1016/j.spa.2008.05.004

[2] Bücher, A. and Dette, H. (2013). Multiplier bootstrap of tail copulas with applications. *Bernoulli* **19** 1655–1687. MR3129029 https://doi.org/10.3150/12-BEJ425

[3] 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

[4] 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

[5] Davis, R.A., Mikosch, T. and Cribben, I. (2012). Towards estimating extremal serial dependence via the bootstrapped extremogram. *J. Econometrics* **170** 142–152. MR2955945 https://doi.org/10.1016/j.jeconom.2012.04.003

[6] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3

[7] Drees, H. (2000). Weighted approximations of tail processes for β-mixing random variables. *Ann. Appl. Probab.* **10** 1274–1301. MR1810875 https://doi.org/10.1214/aoap/1019487617

[8] Drees, H. (2008). Some aspects of extreme value statistics under serial dependence. *Extremes* **11** 35–53. MR2420196 https://doi.org/10.1007/s10687-007-0051-1

[9] Drees, H. (2015). Bootstrapping empirical processes of cluster functionals with application to extremograms. Preprint.

[10] 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

[11] 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

[12] 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

[13] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events: For Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. MR1458613 https://doi.org/10.1007/978-3-642-33483-2

[14] Giné, E. and Nickl, R. (2016). *Mathematical Foundations of Infinite-Dimensional Statistical Models. Cambridge Series in Statistical and Probabilistic Mathematics*. New York: Cambridge Univ. Press. MR3588285 https://doi.org/10.1017/CBO9781107337862

[15] Groen, T. (2015). Bootstrapping extreme value statistics. M.Sc. thesis, Erasmus Univ., Rotterdam.

[16] Kosorok, M.R. (2003). Bootstraps of sums of independent but not identically distributed stochastic processes. *J. Multivariate Anal.* **84** 299–318. MR1965224 https://doi.org/10.1016/S0047-259X(02)00040-4

[17] Kulik, R. and Soulier, P. (2020). *Heavy Tailed Time Series*. Berlin: Springer.

[18] Kulik, R., Soulier, P. and Wintenberger, O. (2019). The tail empirical process of regularly varying functions of geometrically ergodic Markov chains. *Stochastic Process. Appl.* **129** 4209–4238. MR4013860 https://doi.org/10.1016/j.spa.2018.11.014

[19] Loukrati, H. (2018). Tail empirical processes: Limit theorems and bootstrap techniques, with applications to risk measures. Ph.D. thesis, Univ. Ottawa.

[20] Rootzén, H., Leadbetter, M.R. and de Haan, L. (1998). On the distribution of tail array sums for strongly mixing stationary sequences. *Ann. Appl. Probab.* **8** 868–885. MR1627799 https://doi.org/10.1214/aoap/1028903454

[21] 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

Received October 2019 and revised September 2020