



**HAL**  
open science

## Marginal Expected Shortfall Risk Measure for Time Series

Yuri Goegebeur, Armelle Guillou, Jing Qin

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou, Jing Qin. Marginal Expected Shortfall Risk Measure for Time Series. 2025. <hal-04630073v4>

**HAL Id: hal-04630073**

**<https://hal.science/hal-04630073v4>**

Preprint submitted on 26 Aug 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Marginal Expected Shortfall Risk Measure for Time Series

Yuri Goegebeur<sup>(1)</sup>, Armelle Guillou<sup>(2)</sup> and Jing Qin<sup>(1)</sup>

<sup>(1)</sup> Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

<sup>(2)</sup> Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

## Abstract

We consider the estimation of the marginal expected shortfall  $\mathbb{E}(X_h|Y_0 > U_Y(1/p))$  at extreme levels, when  $((X_t, Y_t))_{t \in \mathbb{Z}}$  is a strictly stationary  $\beta$ -mixing time series with marginal distributions of Pareto-type,  $U_Y$  is the tail quantile function associated to  $Y_t$ ,  $h$  is a positive integer and  $p \in (0, 1)$  is such that  $p \rightarrow 0$ . We propose an estimator for this risk measure based on a Weissman-type construction. First, in case of a non-negative time series, we establish the weak convergence of our estimator by using empirical processes arguments combined with the cluster method of Drees and Rootzén (2010). Then, we extend our result to the case of real-valued time series by using the decomposition of the original time series into the positive and negative parts, and we also propose a bootstrap procedure. The performance of our estimator is illustrated on a simulation experiment. Finally, the method is applied on river flow data.

**Keywords:** time series, mixing, extreme value theory, marginal expected shortfall, empirical process, bootstrap.

## 1 Introduction

In this paper, we investigate the asymptotic behaviour of tail parameters based on extreme values in case of dependent observations. These types of results are relevant for real-life time series where the assumption of independence seems inappropriate, e.g., in finance or on internet traffic datasets, among others. In particular, we study the well-known Marginal Expected Shortfall (MES) risk measure defined as

$$\theta_{h,p} = \mathbb{E}(X_h|Y_0 > U_Y(1/p)), p \in (0, 1), h = 1, 2, \dots, \quad (1)$$

in case where  $(Z_t)_{t \in \mathbb{Z}} := ((X_t, Y_t))_{t \in \mathbb{Z}}$  is a strictly stationary time series with marginal distributions of Pareto-type, i.e., with both survival functions satisfying

$$\bar{F}_\bullet(x) = x^{-\frac{1}{\gamma_\bullet}} \ell_\bullet(x), \quad (2)$$

where  $\bullet$  denotes either  $X$  or  $Y$ ,  $\gamma_\bullet > 0$  is the extreme value index, and  $\ell_\bullet$  is a slowly varying function at infinity, i.e., a measurable positive function such that

$$\lim_{x \rightarrow \infty} \frac{\ell_\bullet(\lambda x)}{\ell_\bullet(x)} = 1, \forall \lambda > 0.$$

In that case we say that  $\overline{F}_\bullet$  is regularly varying at infinity with index  $-1/\gamma_\bullet$ , denoted  $RV_{-1/\gamma_\bullet}$ . This is equivalent to assume that the tail quantile function  $U_\bullet(x) = \inf\{z : F_\bullet(z) \geq 1 - 1/x\}$ ,  $x > 1$ , is also regularly varying at infinity but this time with index  $\gamma_\bullet$ . Note that in the application considered in this paper, this assumption of heavy-tailed marginal distributions is credible, and thus it justifies the above Pareto-type models.

In (1) the parameter  $p$  is small, in the sense  $p \leq 1/n$ , where  $n$  is the number of observations available for the estimation. In that case, we say that we extrapolate outside the data range.

The marginal expected shortfall is an important factor when measuring the systemic risk of financial institutions. It has been introduced by Acharya et al. (2010) and studied extensively in the literature but almost only for independent observations, e.g., by Cai et al. (2015), Di Bernardino and Prieur (2018), or in a regression context, e.g., by Goegebeur et al. (2023). To the best of our knowledge, the only contribution dealing with MES for a strictly stationary bivariate time series is given in Davison et al. (2023), but in case of a MES defined without time lags, i.e., based on  $X$  and  $Y$  observations obtained at the same time instance. Although  $X$  and  $Y$  in the latter paper are dependent processes, our study cannot be trivially considered as the special case where  $(X, Y)$  in Davison et al. (2023) is replaced by  $(X_h, Y_0)$ , unless redoing completely their paper under this new framework and changing all their assumptions, in particular their Condition D (iii) and (iv).

The estimation of our risk measure  $\theta_{h,p}$  has natural applications in finance, but not only. Indeed, we will illustrate our methodology in environmental science, where, for instance, one wants to get an idea of how the flow of a river changes at two different locations. If at time zero of a location  $Y$ , i.e.,  $Y_0$ , the river flow is extreme (above some high quantile), we might be interested in its value some time, say  $h$ , later at a different location, denoted  $X_h$ , along the river. This is helpful since river flows at two locations (e.g., upstream and downstream of a given site) obviously show dependence, but they are also different, due to, e.g., gains from tributary streams and/or groundwater.

Our contribution in this paper consists to estimate the multivariate risk measure  $\theta_{h,p}$  and to establish its main asymptotic properties by using empirical processes arguments combined with the cluster method of Drees and Rootzén (2010). Recently, Goegebeur et al. (2025) studied the estimation of the conditional tail expectation (CTE) with time lags, defined as  $\mathbb{E}(X_h | X_0 > U_X(1/p))$ , in the framework of a strictly stationary time series. The main differences between the present work and Goegebeur et al. (2025) are as follows. Firstly,  $\theta_{h,p}$  considered here is a bivariate risk measure, which needs to be estimated from a strictly stationary bivariate time series  $((X_t, Y_t))_{t \in \mathbb{Z}}$ , while the CTE of Goegebeur et al. (2025) is essentially univariate, and is estimated from a strictly stationary time series  $(X_t)_{t \in \mathbb{Z}}$ . Secondly, Goegebeur et al. (2025) used a multivariate regular variation approach, a method that we do not wish to use here to avoid the implicit assumption that  $X$  and  $Y$  have the same extreme value index. See, e.g., Padoan et al. (2023) where a slightly different version of the MES was considered, also in case of heavy-tailed distributions, but with the same indices. In the present paper, we work under a more general, and commonly accepted, assumption on the convergence of the dependence structure involving the concept of the tail copula function (Assumption  $(\mathcal{R})$  below),

and allowing for different extreme value indices for  $X$  and  $Y$ . These two differences complicate our study considerably as we have to deal with the dependence between  $X_h$  and  $Y_0$ .

The remainder of the paper is organized as follows. In Section 2, we consider the case of a bivariate nonnegative time series and derive a first order approximation for  $\theta_{h,p}$  which motivates the construction of our estimator based on a Weissman-type idea. Then, under suitable assumptions, we establish the weak convergence of the estimator towards a Gaussian process. Finally, in Section 3, we extend our approach to the case of a bivariate real-valued time series by using the decomposition of the original time series into the positive and negative parts and we derive the same type of convergence for the final estimator under additional assumptions. As this result is difficult to use in practice, due to the complicated covariance structure of the limiting process, we also adjust the multiplier block bootstrap initially introduced by Drees (2015) to our context. The finite sample properties of our estimator are evaluated with a simulation study in Section 4, and the practical applicability is showcased on river flow data in Section 5. All the proofs are postponed to Section 6.

## 2 Nonnegative time series

Our aim in this section is to establish the weak convergence of our risk measure (1) in case where both  $(X_t)_{t \in \mathbb{Z}}$  and  $(Y_t)_{t \in \mathbb{Z}}$  are assumed nonnegative. It is assumed that  $F_X$  and  $F_Y$  are continuous on  $(0, \infty)$ . Since our measure involves the two variables, an assumption on the right-hand upper tail dependence of the pair  $(X_h, Y_0)$  is required. Let  $R_r(x_h, y_0) := r \mathbb{P}(\bar{F}_X(X_h) \leq \frac{x_h}{r}, \bar{F}_Y(Y_0) \leq \frac{y_0}{r})$ . This assumption is stated as follows.

**Assumption ( $\mathcal{R}$ )** *There exists a function  $R$  such that for all  $(x_h, y_0) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}$ , we have*

$$\lim_{r \rightarrow \infty} R_r(x_h, y_0) = R(x_h, y_0).$$

The next lemma provides a theoretical approximation for  $\theta_{h,p}$ .

**Lemma 2.1** *Assume that  $(Z_t)_{t \in \mathbb{Z}}$  is a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ ,  $F_X$  being of Pareto-type with  $\gamma_X < 1$ . Then, for  $p \searrow 0$  we have under Assumption ( $\mathcal{R}$ ):*

$$\frac{\theta_{h,p}}{U_X(1/p)} \longrightarrow \int_0^\infty R(z^{-1/\gamma_X}, 1) dz.$$

Note that this lemma is similar to Lemma 2.1 in Goegebeur et al. (2025), but our parameter of interest  $\theta_{h,p}$  now involves both variables  $X$  and  $Y$ . Therefore, Lemma 2.1 above requires some adjustments of the assumptions imposed in Lemma 2.1 of Goegebeur et al. (2025), among them, Assumption ( $\mathcal{R}$ ). Despite these, the proof of the lemma is similar to that of Lemma 2.1 in Goegebeur et al. (2025), and we omit it for brevity from the paper. Assuming that the integral in the limit is positive, and since this lemma is valid for any  $p \searrow 0$ , if we consider an

intermediate sequence  $k$ , i.e., a sequence such that  $k = k_n \rightarrow \infty$  and  $k/n \rightarrow 0$ , we have

$$\frac{\theta_{h,p}/U_X(1/p)}{\theta_{h,k/n}/U_X(n/k)} \rightarrow 1,$$

from which we deduce by a Weissman-type construction (see Weissman, 1978)

$$\theta_{h,p} \simeq \frac{U_X(1/p)}{U_X(n/k)} \theta_{h,k/n} \simeq \left(\frac{k}{np}\right)^{\gamma_X} \theta_{h,k/n}$$

since the distribution of  $X$  is of Pareto-type. If  $\hat{\gamma}_{X,k}$  and  $\bar{\theta}_{h,k/n}$  denote estimators for  $\gamma_X$  and  $\theta_{h,k/n}$ , respectively, then we obtain our final estimator for  $\theta_{h,p}$  given by

$$\hat{\theta}_{h,p} = \left(\frac{k}{np}\right)^{\hat{\gamma}_{X,k}} \bar{\theta}_{h,k/n}.$$

To estimate  $\gamma_X$ , we will use the Hill estimator (Hill, 1975) defined as

$$\hat{\gamma}_{X,k} = \frac{1}{k} \sum_{i=1}^k \log X_{(n-i+1)} - \log X_{(n-k)}, \quad (3)$$

where  $X_{(1)} \leq \dots \leq X_{(n)}$  are the order statistics of  $X_1, \dots, X_n$ . Regarding the estimator of the parameter of interest at an intermediate level, namely  $\theta_{h,k/n}$ , we refer to the following subsection, where we also establish its weak convergence.

## 2.1 Estimator in the intermediate case

The natural estimator for  $\theta_{h,k/n}$  is

$$\bar{\theta}_{h,k/n} := \frac{1}{k} \sum_{i=1}^n X_{i+h} \mathbb{1}_{\{Y_i > Y_{(n-k)}\}},$$

where  $\mathbb{1}_A$  denotes the indicator function of the set  $A$ .

To handle the randomness of the threshold  $Y_{(n-k)}$ , we introduce the statistic

$$\bar{\theta}_n(s) := \frac{1}{k} \sum_{i=1}^n \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k)} > s\right\}},$$

where  $s > 0$ . This gives then

$$\bar{\theta}_{h,k/n} = U_X\left(\frac{n}{k}\right) \bar{\theta}_n\left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right).$$

In particular, we will study the asymptotic behavior of the process related to  $\bar{\theta}_n(s)$ , with  $s \in [1 - \varepsilon, 1 + \varepsilon]$ ,  $\varepsilon \in (0, 1)$ , using the theory for weak convergence of empirical processes of

cluster functionals introduced by Drees and Rootzén (2010).

To this aim, we consider

$$Z_{n,i} = (Z_{n,i,1}, Z_{n,i,2}) := \left( \frac{X_{i+h}}{U_X(n/k)}, \frac{Y_i}{U_Y(n/k)} \right) \mathbb{1}_{\left\{ \frac{X_{i+h}}{U_X(n/k)} > 1-\varepsilon \text{ or } \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}},$$

and denote

$$\Phi := \{ \phi_s, \phi_{D_X, s}, \phi_{D_Y, s}, \phi_{H, s} : s \in [1 - \varepsilon, 1 + \varepsilon] \},$$

where, for  $z = (z_1, z_2)$ ,

$$\begin{cases} \phi_s(z) := z_1 \mathbb{1}_{\{z_2 > s\}} \\ \phi_{D_X, s}(z) := \mathbb{1}_{\{z_1 > s\}} \\ \phi_{D_Y, s}(z) := \mathbb{1}_{\{z_2 > s\}} \\ \phi_{H, s}(z) := \log\left(\frac{z_1}{s}\right) \mathbb{1}_{\{z_1 > s\}} =: \log^+\left(\frac{z_1}{s}\right). \end{cases}$$

The weak convergence of all the estimators that we have to handle in order to study  $\hat{\theta}_{h,p}$ , namely  $X_{(n-k)}$ ,  $Y_{(n-k)}$ ,  $\hat{\gamma}_{X,k}$  and  $\bar{\theta}_{h,k/n}$ , can be obtained through the weak convergence of statistics of the form  $\sum_{i=1}^n \Psi(Z_{n,i})$ , for  $\Psi \in \Phi$ . We can then resort to Drees and Rootzén (2010) who provide conditions according to which

$$\frac{1}{\sqrt{nv_n}} \sum_{i=1}^n \{ \Psi(Z_{n,i}) - \mathbb{E}(\Psi(Z_{n,1})) \},$$

where  $v_n := \mathbb{P}(Z_{n,1} \neq (0, 0))$ , converges towards a centered Gaussian process.

For instance, returning to  $\bar{\theta}_n(s)$ , note that  $\bar{\theta}_n(s) = \frac{1}{k} \sum_{i=1}^n \phi_s(Z_{n,i})$ , for  $s \in [1 - \varepsilon, 1 + \varepsilon]$ .

Under our framework and assuming that  $F_X$  and  $F_Y$  are continuous on  $(0, \infty)$ , we have

$$\begin{aligned} \frac{n}{k} v_n &= \frac{n}{k} \mathbb{P} \left( \frac{X_{1+h}}{U_X(n/k)} > 1 - \varepsilon \text{ or } \frac{Y_1}{U_Y(n/k)} > 1 - \varepsilon \right) \\ &= \frac{\bar{F}_X((1 - \varepsilon)U_X(n/k))}{\bar{F}_X(U_X(n/k))} + \frac{\bar{F}_Y((1 - \varepsilon)U_Y(n/k))}{\bar{F}_Y(U_Y(n/k))} \\ &\quad - \frac{n}{k} \mathbb{P} \left( \bar{F}_X(X_{1+h}) < \frac{k}{n} \frac{\bar{F}_X((1 - \varepsilon)U_X(n/k))}{\bar{F}_X(U_X(n/k))}, \bar{F}_Y(Y_1) < \frac{k}{n} \frac{\bar{F}_Y((1 - \varepsilon)U_Y(n/k))}{\bar{F}_Y(U_Y(n/k))} \right) \\ &\longrightarrow (1 - \varepsilon)^{-\frac{1}{\gamma_X}} + (1 - \varepsilon)^{-\frac{1}{\gamma_Y}} - R \left( (1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}} \right) = L \left( (1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}} \right), \quad (4) \end{aligned}$$

where  $L$  is the stable tail dependence function, see, e.g., Huang (1992) and Drees and Huang (1998). This ensures in particular that  $v_n \rightarrow 0$ .

Let us now use Drees and Rootzén (2010) to establish sufficient conditions for the process  $\mathcal{Z}_n$  defined as

$$\begin{aligned} \mathcal{Z}_n(\Psi) &= \sqrt{k} \left\{ \frac{1}{k} \sum_{i=1}^n \Psi(Z_{n,i}) - \frac{n}{k} \mathbb{E}(\Psi(Z_{n,1})) \right\} \\ &= \sqrt{\frac{n}{k} v_n} \frac{1}{\sqrt{nv_n}} \sum_{i=1}^n \{ \Psi(Z_{n,i}) - \mathbb{E}(\Psi(Z_{n,1})) \}, \end{aligned}$$

to converge towards a Gaussian process in the space of bounded functions indexed by  $\Phi$ , denoted  $\ell^\infty(\Phi)$ . This requires some assumptions, listed below. The strength of the temporal dependence between the random vectors  $Z_{n,1}, \dots, Z_{n,n}$  is evaluated through the  $\beta$ -mixing coefficient

$$\beta_{n,k} := \sup_{1 \leq \ell \leq n-k-1} \mathbb{E} \left\{ \sup_{B \in \mathcal{B}_{n,\ell+k+1}^\ell} \left| \mathbb{P}(B | \mathcal{B}_{n,1}^\ell) - \mathbb{P}(B) \right| \right\},$$

where  $\mathcal{B}_{n,i}^j$  denotes the  $\sigma$ -field generated by  $(Z_{n,\ell})_{i \leq \ell \leq j}$ . We assume that there exist sequences of integers  $(r_n)_{n \in \mathbb{N}}$  and  $(\ell_n)_{n \in \mathbb{N}}$ , constituting the basic elements of the so-called ‘‘big blocks, small blocks’’ technique, such that  $r_n, \ell_n \rightarrow \infty$  and the following conditions hold:

- (A) As  $n \rightarrow \infty$ , we have  $\ell_n = o(r_n)$ ,  $r_n = o(\sqrt{nv_n})$ ,  $r_n v_n \rightarrow 0$  and  $\beta_{n,\ell_n} \frac{n}{r_n} \rightarrow 0$ ;  
(B) For all  $0 \leq i \leq r_n - 1$ , there exists

$$\begin{aligned} s_n(i) \geq & \mathbb{E} \left[ \max \left\{ \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}}, \log^+ \frac{X_h}{(1-\varepsilon)U_X(n/k)}, \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}}, \mathbb{1}_{\left\{ \frac{X_h}{U_X(n/k)} > 1-\varepsilon \right\}} \right\} \right. \\ & \times \max \left\{ \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}}, \log^+ \frac{X_{i+h}}{(1-\varepsilon)U_X(n/k)}, \right. \\ & \left. \left. \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}}, \mathbb{1}_{\left\{ \frac{X_{i+h}}{U_X(n/k)} > 1-\varepsilon \right\}} \right\} \left| \frac{X_h}{U_X(n/k)} > 1-\varepsilon \text{ or } \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right], \end{aligned} \quad (5)$$

such that  $s_\infty(i) = \lim_{n \rightarrow \infty} s_n(i)$  exists, and  $\lim_{n \rightarrow \infty} \sum_{i=0}^{r_n-1} s_n(i) = \sum_{i=0}^{\infty} s_\infty(i) < \infty$ . Moreover there exists  $\delta > 0$  such that, as  $n \rightarrow \infty$

$$\begin{aligned} & \sum_{i=0}^{r_n-1} \left\{ \mathbb{E} \left[ \left( \max \left\{ \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}}, \log^+ \frac{X_h}{(1-\varepsilon)U_X(n/k)} \right\} \right. \right. \right. \\ & \times \max \left\{ \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}}, \log^+ \frac{X_{i+h}}{(1-\varepsilon)U_X(n/k)} \right\} \left. \left. \left. \right)^{1+\delta} \left| \frac{X_h}{U_X(n/k)} > 1-\varepsilon \text{ or } \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right] \right\}^{\frac{1}{1+\delta}} \\ & = O(1). \end{aligned} \quad (6)$$

Condition (A) imposes some restrictions on the rate of convergence towards infinity of the different sequences involved, i.e., on  $\ell_n, r_n$  and  $k$ . If we consider a time series whose  $\beta$ -mixing coefficient decreases geometrically, i.e., such that  $\beta_{n,k} = O(\eta^k)$  for some  $\eta \in (0, 1)$ , then one may choose  $\ell_n = O(\log n)$  and Condition (A) is fulfilled for a suitably chosen  $r_n$  if  $(\log n)^2 = o(k)$  and  $k = o(n/\log n)$ . These two conditions on  $k$  are trivially satisfied if the intermediate sequence  $k$  is of the form  $k = \lfloor n^a \rfloor$ , as is the case in our main theorem (Theorem 3.1).

Regarding Condition (B), the situation is more complicated since it cannot be shown that this condition is satisfied in general by assuming only a geometrically  $\beta$ -mixing time series, but it must be checked on a case-by-case basis, depending on the explicit model of the time series considered, whether or not this is the case. We refer the reader to Davis et al. (2018) or Drees and Knežević (2020) where some techniques are discussed in order to verify similar but simpler versions of Condition (B) for some suitable heavy tailed linear time series or solutions to stochastic

recurrence equations. Note that, by using the Cauchy-Schwarz inequality, the conditional expectations in (5) will be finite if  $\gamma_X < 1/2$ , while for (6) they will be finite if  $\gamma_X < 1/(2(1 + \delta))$ . The latter restriction is an usual condition required for weak convergence of the MES in the time series context, see Davison et al. (2023, Theorem 4.1).

Additionally, to compute the covariance structure of the limiting process of  $\mathcal{Z}_n$ , we need an assumption on the right-hand upper tail dependence of the vector  $(X_h, Y_0, X_{i+h}, Y_i)$  stated as follows.

**Assumption ( $\tilde{\mathcal{R}}$ )** *There exists a function  $\tilde{R}_i$  such that for all  $(x_h, y_0, x_{i+h}, y_i) \in [0, \infty]^4 \setminus \{(\infty, \infty, \infty, \infty)\}$ , we have*

$$\lim_{r \rightarrow \infty} r \mathbb{P} \left( \bar{F}_X(X_h) \leq \frac{x_h}{r}, \bar{F}_Y(Y_0) \leq \frac{y_0}{r}, \bar{F}_X(X_{i+h}) \leq \frac{x_{i+h}}{r}, \bar{F}_Y(Y_i) \leq \frac{y_i}{r} \right) = \tilde{R}_i(x_h, y_0, x_{i+h}, y_i),$$

for all  $i = 1, 2, \dots$

Clearly, if Assumption ( $\tilde{\mathcal{R}}$ ) is satisfied this is also the case for Assumption ( $\mathcal{R}$ ) since we have the link  $R(x_h, y_0) = \tilde{R}_i(x_h, y_0, \infty, \infty)$ .

The next theorem states the weak convergence, denoted by  $\rightsquigarrow$ , of the process  $\mathcal{Z}_n$  indexed by  $\Phi$ .

**Theorem 2.1** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ , both of Pareto-type with  $\gamma_X < 1$ . Assume also that Assumption ( $\tilde{\mathcal{R}}$ ) and Conditions **(A)** and **(B)** are fulfilled. Then, for any  $\varepsilon \in (0, 1)$ , we have, for  $\Psi \in \Phi$ ,*

$$\mathcal{Z}_n(\Psi) \rightsquigarrow \sqrt{L \left( (1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}} \right)} \mathcal{Z}_\varepsilon(\Psi) =: \mathcal{Z}(\Psi),$$

in  $\ell^\infty(\Phi)$ , where  $\mathcal{Z}_\varepsilon$  is a centered Gaussian process with covariance function given by

$$\text{Cov}(\mathcal{Z}_\varepsilon(\Psi_1), \mathcal{Z}_\varepsilon(\Psi_2)) = \mathbb{E}[\Psi_1(\mathcal{W}_0)\Psi_2(\mathcal{W}_0)] + \sum_{i=1}^{\infty} [\mathbb{E}(\Psi_1(\mathcal{W}_0)\Psi_2(\mathcal{W}_i)) + \mathbb{E}(\Psi_1(\mathcal{W}_i)\Psi_2(\mathcal{W}_0))],$$

for  $\Psi_1, \Psi_2 \in \Phi$ , with

$$\mathcal{W}_i := \left( W_{i+h} \mathbb{1}_{\{W_{i+h} > 1 - \varepsilon \text{ or } \tilde{W}_i > 1 - \varepsilon\}}, \tilde{W}_i \mathbb{1}_{\{W_{i+h} > 1 - \varepsilon \text{ or } \tilde{W}_i > 1 - \varepsilon\}} \right),$$

where  $(W_h, \tilde{W}_0, W_{i+h}, \tilde{W}_i)$  is a random vector of nonnegative components whose survival distribution function is given by

$$\begin{aligned} \mathbb{P} \left( W_h > x_h, \tilde{W}_0 > y_0, W_{i+h} > x_{i+h}, \tilde{W}_i > y_i \right) &= \frac{1}{L \left( (1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}} \right)} \\ &\times \left\{ \tilde{R}_i \left( [\max(x_h, 1 - \varepsilon)]^{-\frac{1}{\gamma_X}}, y_0^{-\frac{1}{\gamma_Y}}, x_{i+h}^{-\frac{1}{\gamma_X}}, y_i^{-\frac{1}{\gamma_Y}} \right) + \tilde{R}_i \left( x_h^{-\frac{1}{\gamma_X}}, [\max(y_0, 1 - \varepsilon)]^{-\frac{1}{\gamma_Y}}, x_{i+h}^{-\frac{1}{\gamma_X}}, y_i^{-\frac{1}{\gamma_Y}} \right) \right. \\ &\left. - \tilde{R}_i \left( [\max(x_h, 1 - \varepsilon)]^{-\frac{1}{\gamma_X}}, [\max(y_0, 1 - \varepsilon)]^{-\frac{1}{\gamma_Y}}, x_{i+h}^{-\frac{1}{\gamma_X}}, y_i^{-\frac{1}{\gamma_Y}} \right) \right\}. \end{aligned} \quad (7)$$

Before looking at the weak convergence of our estimator  $\bar{\theta}_{h,k/n}$ , we need first to study the random threshold  $Y_{(n-k)}$ . To this aim, we need to make precise the slowly varying function  $\ell_\bullet$  appearing in model (2). This gives a Hall-type model, similar to the Hall class of distribution functions (Hall, 1982, Hall and Welsh, 1985), but slightly more general.

**Assumption ( $\mathcal{D}_\bullet$ )** *The survival functions of  $X$  and  $Y$  satisfy*

$$\bar{F}_\bullet(x) = A_\bullet x^{-1/\gamma_\bullet} \left( 1 + \frac{1}{\gamma_\bullet} \delta_\bullet(x) \right),$$

where  $A_\bullet > 0$ ,  $\gamma_\bullet > 0$ , and  $|\delta_\bullet(\cdot)| \in RV_{-\beta_\bullet}$ ,  $\beta_\bullet > 0$ .

This assumption is very general, satisfied by almost all the usual Pareto-type distributions: Fréchet, student  $t$ ,  $F$ , generalized Pareto and Burr distributions. In particular, the distributions that satisfy this assumption also satisfy the second-order condition in univariate extreme value statistics (see, e.g., Theorem 2.3.9 in de Haan and Ferreira, 2006). Under Assumption ( $\mathcal{D}_\bullet$ ),  $U_\bullet$  satisfies

$$U_\bullet(x) = A_\bullet^{\gamma_\bullet} x^{\gamma_\bullet} (1 + a_\bullet(x)),$$

where  $a_\bullet(x) := \delta_\bullet(U_\bullet(x))(1 + o(1))$ , and thus  $|a_\bullet| \in RV_{-\beta_\bullet \gamma_\bullet}$ .

**Corollary 2.1** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ , satisfying Assumptions ( $\mathcal{D}_X$ ) and ( $\mathcal{D}_Y$ ) with  $\gamma_X < 1$ . Assume also that Assumption ( $\tilde{\mathcal{R}}$ ) and Conditions (**A**) and (**B**) are fulfilled. Then if  $\sqrt{k}\delta_X(U_X(n/k)) \rightarrow 0$  and  $\sqrt{k}\delta_Y(U_Y(n/k)) \rightarrow 0$ , we have*

$$\sqrt{k} \left\{ \frac{X_{(n-k)}}{U_X(n/k)} - 1 \right\} \rightsquigarrow \gamma_X \mathcal{Z}(\phi_{D_X,1}),$$

and

$$\sqrt{k} \left\{ \frac{Y_{(n-k)}}{U_Y(n/k)} - 1 \right\} \rightsquigarrow \gamma_Y \mathcal{Z}(\phi_{D_Y,1}).$$

The proof of this corollary is similar to the proof of Corollary 2.1 in Goegebeur et al. (2025). Note that if interest is only in obtaining a result for the intermediate quantile of a univariate time series, then this can be achieved under simpler assumptions.

To obtain the weak convergence of our estimator  $\bar{\theta}_{h,k/n}$ , we need to strengthen the convergence of  $R_r$  towards its limit  $R$  by specifying the uniform rate of convergence.

**Assumption ( $\mathcal{S}$ )** *There exist  $\kappa > \gamma_X$  and  $\tau < 0$  such that, as  $r \rightarrow \infty$*

$$\sup_{\substack{x_h > 0 \\ y_0 \in [\frac{1}{2}, 2]}} \frac{|R_r(x_h, y_0) - R(x_h, y_0)|}{x_h^\kappa \wedge 1} = O(r^\tau).$$

Note that the assumptions in Lemma 2.1 allow the case where  $(X_h, Y_0)$  is asymptotically upper tail independent, corresponding to the case  $R(x_h, y_0) = 0$  for all  $x_h, y_0 \geq 0$ , and thus a null limit in this lemma. On the contrary, the uniformity requirement in Assumption ( $\mathcal{S}$ ) excludes that

case.

We have now all the ingredients to state the weak convergence of our estimator in the intermediate case.

**Theorem 2.2** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ , satisfying Assumptions  $(\mathcal{D}_X)$  and  $(\mathcal{D}_Y)$  with  $\gamma_X < 1$ . Assume also that Assumptions  $(\tilde{\mathcal{R}})$  and  $(\mathcal{S})$  together with Conditions  $(\mathbf{A})$  and  $(\mathbf{B})$  are fulfilled. Then, if  $k = \lfloor n^a \rfloor$  with*

$$0 < a < \min \left( \frac{2\gamma_X^2(\beta_X - \varepsilon_X)}{2\gamma_X^2(\beta_X - \varepsilon_X) + 1 + \beta_X\gamma_X + \xi\gamma_X}, \frac{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y)}{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y) + 1}, \frac{-2\tau}{1 - 2\tau} \right),$$

for  $\varepsilon_X \in (0, \beta_X)$ ,  $\varepsilon_Y \in (0, \beta_Y)$  and  $\xi \in (0, \beta_X + (1 - \gamma_X)/\gamma_X)$ , we have

$$\sqrt{k} \left( \frac{\bar{\theta}_{h,k/n}}{\theta_{h,k/n}} - 1 \right) \rightsquigarrow \frac{\mathcal{Z}(\phi_1)}{\int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz} - (1 - \gamma_X) \mathcal{Z}(\phi_{D_Y,1}).$$

Let us now turn to our main objective, namely the study of  $\hat{\theta}_{h,p}$  when  $p \leq 1/n$ . This requires extreme value arguments to extrapolate beyond the data range.

## 2.2 Estimator under extrapolation

To study the asymptotic properties of the estimator  $\hat{\theta}_{h,p}$  we need first to show the weak convergence of  $\hat{\gamma}_{X,k}$  defined in (3).

Resnick and Stărică (1997) showed the asymptotic normality of the Hill estimator for some specific models, but in our case, we need to establish this weak convergence in terms of the stochastic process  $\mathcal{Z}$  appearing in Theorem 2.1 in order to combine this result with Theorem 2.2. This is the aim of the next corollary whose proof is similar to the one of Theorem 2.3 in Goegebeur et al. (2025).

**Corollary 2.2** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ , satisfying Assumptions  $(\mathcal{D}_X)$  and  $(\mathcal{D}_Y)$  with  $\gamma_X < 1$ . Assume also that Assumption  $(\tilde{\mathcal{R}})$  and Conditions  $(\mathbf{A})$  and  $(\mathbf{B})$  are fulfilled. Then, if  $\sqrt{k} \delta_X(U_X(n/k)) \rightarrow 0$ , we have*

$$\sqrt{k} (\hat{\gamma}_{X,k} - \gamma_X) \rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_X,1}).$$

We can now state the weak convergence of  $\hat{\theta}_{h,p}$  properly normalized.

**Theorem 2.3** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a nonnegative strictly stationary time series, with continuous marginal distribution functions  $F_X$  and  $F_Y$  on  $(0, \infty)$ , satisfying Assumptions  $(\mathcal{D}_X)$  and  $(\mathcal{D}_Y)$  with  $\gamma_X < 1$ . Assume also that Assumptions  $(\tilde{\mathcal{R}})$  and  $(\mathcal{S})$  together with Conditions  $(\mathbf{A})$  and  $(\mathbf{B})$  are fulfilled. Then, if  $k = \lfloor n^a \rfloor$  with*

$$0 < a < \min \left( \frac{2\gamma_X^2(\beta_X - \varepsilon_X)}{2\gamma_X^2(\beta_X - \varepsilon_X) + 1 + \beta_X\gamma_X + \xi\gamma_X}, \frac{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y)}{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y) + 1}, \frac{-2\tau}{1 - 2\tau} \right),$$

for  $\varepsilon_X \in (0, \beta_X)$ ,  $\varepsilon_Y \in (0, \beta_Y)$  and  $\xi \in (0, \beta_X + (1 - \gamma_X)/\gamma_X)$ , we have for  $p$  satisfying  $\frac{k}{np} \rightarrow \infty$  and  $\frac{\log k/(np)}{\sqrt{k}} \rightarrow 0$

$$\frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}}{\theta_{h,p}} - 1 \right) \rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_X,1}).$$

Using the covariance structure from Theorem 2.1, we can show that the limiting distribution in Theorem 2.3 is  $\mathcal{N}(0, \sigma^2)$  with

$$\begin{aligned} \sigma^2 := & \gamma_X^2 + 2 \sum_{i=1}^{\infty} \left\{ \gamma_X^2 \tilde{R}_i(1, \infty, 1, \infty) + \int_1^{\infty} \int_1^{\infty} \tilde{R}_i \left( z_1^{-\frac{1}{\gamma_X}}, \infty, z_2^{-\frac{1}{\gamma_X}}, \infty \right) \frac{dz_1}{z_1} \frac{dz_2}{z_2} \right. \\ & \left. - \gamma_X \int_1^{\infty} \left[ \tilde{R}_i \left( z^{-\frac{1}{\gamma_X}}, \infty, 1, \infty \right) + \tilde{R}_i \left( 1, \infty, z^{-\frac{1}{\gamma_X}}, \infty \right) \right] \frac{dz}{z} \right\}. \end{aligned}$$

### 3 Real-valued time series

We come back to our initial aim, i.e., estimation of  $\theta_{h,p}$  on the original time series. The idea is to decompose  $(X_t)_{t \in \mathbb{Z}}$ , respectively  $(Y_t)_{t \in \mathbb{Z}}$ , into the positive and negative parts:  $X_t := X_t^+ - X_t^-$  where  $X_t^+ := \max(X_t, 0)$  and  $X_t^- := -\min(X_t, 0)$ , and similar expressions for  $Y_t^+$  and  $Y_t^-$ . Define

$$\theta_{h,p}^+ := \mathbb{E} \left[ X_h^+ \mid Y_0^+ > U_{Y^+} \left( \frac{1}{p} \right) \right]$$

and

$$\hat{\theta}_{h,p}^+ := \left( \frac{k}{np} \right)^{\hat{\gamma}_{X^+,k}} \frac{1}{k} \sum_{i=1}^n X_{i+h}^+ \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}}.$$

The idea is then to proceed into two steps, as follows:

**First step:** We use  $\hat{\theta}_{h,p}^+$  to estimate  $\theta_{h,p}^+$  by applying Theorem 2.3 and thus we obtain

$$\frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}^+}{\theta_{h,p}^+} - 1 \right) \rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_X,1}).$$

**Second step:** We use the decomposition

$$\frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}}{\theta_{h,p}} - 1 \right) = \frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}^+}{\theta_{h,p}^+} - 1 \right) \frac{\theta_{h,p}^+}{\theta_{h,p}} + \frac{\sqrt{k}}{\log k/(np)} \left( \frac{\theta_{h,p}^+}{\theta_{h,p}} - 1 \right),$$

and we prove that, under suitable assumptions, the second term of the right-hand side in the above display is negligible.

To this aim, we need to assume that the lower tail of the marginal distribution function  $F_X$  is regularly varying at infinity with an index  $\alpha$  larger than  $1/\gamma_X$ , i.e., for  $x > 0$ ,

$$\mathbb{P}(X_0 < -x) = x^{-\alpha} \ell(x) \text{ with } \alpha > 1/\gamma_X, \quad (8)$$

where  $\ell$  is a slowly varying function at infinity.

Our main result is summarized in the next theorem.

**Theorem 3.1** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a strictly stationary time series, with marginal distribution functions  $F_X$  and  $F_Y$  continuous on  $(0, \infty)$ . Suppose that Assumptions  $(\tilde{\mathcal{R}})$  and  $(\mathcal{S})$  are satisfied on the original time series, together with Assumptions  $(\mathcal{D}_Y)$ ,  $(\mathcal{D}_X)$  with  $\gamma_X < 1$ , and that  $F_X$  satisfies (8). Then, if Conditions **(A)** and **(B)** are fulfilled for  $((X_t^+, Y_t^+)_{t \in \mathbb{Z}}$ , and if  $k = \lfloor n^\alpha \rfloor$  with*

$$0 < a < \min \left( \frac{2\gamma_X^2(\beta_X - \varepsilon_X)}{2\gamma_X^2(\beta_X - \varepsilon_X) + 1 + \beta_X\gamma_X + \xi\gamma_X}, \frac{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y)}{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y) + 1}, \frac{-2\tau}{1 - 2\tau} \right),$$

for  $\varepsilon_X \in (0, \beta_X)$ ,  $\varepsilon_Y \in (0, \beta_Y)$  and  $\xi \in (0, \beta_X + (1 - \gamma_X)/\gamma_X)$ , we have for  $p$  satisfying  $\frac{k}{np} \rightarrow \infty$ ,  $\frac{\log k/(np)}{\sqrt{k}} \rightarrow 0$  and  $k = O(p^{2\tau(1-\gamma_X)})$

$$\frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}^+}{\theta_{h,p}} - 1 \right) \rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_X,1}).$$

As mentioned above, assumption (8), implies that the lower tail of the distribution of  $X$  is lighter than the upper tail. In case the lower tail is dominating, i.e.,  $\alpha < 1/\gamma_X$ , then there is more risk there, and we can apply Theorem 3.1 to the time series  $(-X_t, Y_t)_{t \in \mathbb{Z}}$ , while in case  $\alpha = 1/\gamma_X$ , we can apply the theory of Section 2 to each tail separately.

The limiting distribution of Theorem 3.1 is difficult to use in practice, e.g., for the construction of confidence intervals for  $\theta_{h,p}$ , due to the complicated covariance function of the process  $\mathcal{Z}$  (see Theorem 2.1). To address this hurdle we will adjust the multiplier block bootstrap method, introduced in Drees (2015), to our context. The multiplier block bootstrap method consists of splitting the original sample into  $m_n = \lfloor n/r_n \rfloor$  blocks of length  $r_n$  and multiplying the cluster functionals of each block by a random variable. More precisely, let  $(\xi_j)_{j \in \mathbb{N}}$  be independent and identically distributed (i.i.d.) random variables such that  $\mathbb{E}(\xi_1) = 0$ ,  $\text{Var}(\xi_1) = 1$ , and which are independent of  $(Z_t)_{t \in \mathbb{Z}}$ . Then, denoting  $I_j := \{(j-1)r_n + 1, \dots, j r_n\}$ ,  $1 \leq j \leq m_n$ , the multiplier block bootstrap version of the estimator  $\hat{\theta}_{h,p}^+$  is defined as

$$\hat{\theta}_{h,p}^* = \left( \frac{k}{np} \right)^{\hat{\gamma}_{X^+,k}^*} \bar{\theta}_{h,k/n}^{+,*},$$

where

$$\begin{aligned} \hat{\gamma}_{X^+,k}^* &:= \frac{\sum_{j=1}^{m_n} (1 + \xi_j) \sum_{i \in I_j} \log \frac{X_i^+}{X_{(n-k)}^+} \mathbb{1}_{\{X_i^+ > X_{(n-k)}^+\}}}{\sum_{j=1}^{m_n} (1 + \xi_j) \sum_{i \in I_j} \mathbb{1}_{\{X_i^+ > X_{(n-k)}^+\}}} \\ \bar{\theta}_{h,k/n}^{+,*} &:= \frac{\sum_{j=1}^{m_n} (1 + \xi_j) \sum_{i \in I_j} X_{i+h}^+ \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}}}{\sum_{j=1}^{m_n} (1 + \xi_j) \sum_{i \in I_j} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}}}. \end{aligned}$$

The consistency of the multiplier block bootstrap procedure is stated in the following theorem. Let  $\mathbb{P}_\xi$  denote the probability with respect to  $\xi := (\xi_j)_{j \in \mathbb{N}}$ , i.e., the conditional probability given  $(Z_{n,i})_{1 \leq i \leq n}$ .

**Theorem 3.2** *Let  $(Z_t)_{t \in \mathbb{Z}}$  be a strictly stationary time series, with marginal distribution functions  $F_X$  and  $F_Y$  continuous on  $(0, \infty)$ . Let  $(\xi_j)_{j \in \mathbb{N}}$  be i.i.d. random variables with  $\mathbb{E}(\xi_1) = 0$ ,  $\text{Var}(\xi_1) = 1$  and independent of  $(Z_t)_{t \in \mathbb{Z}}$ . Suppose that Assumptions  $(\tilde{\mathcal{R}})$  and  $(\mathcal{S})$  are satisfied on the original time series, together with Assumptions  $(\mathcal{D}_Y)$ ,  $(\mathcal{D}_X)$  with  $\gamma_X < 1$ , and that  $F_X$  satisfies (8). Then, if Conditions **(A)** and **(B)** are fulfilled for  $((X_t^+, Y_t^+))_{t \in \mathbb{Z}}$  and if  $k = \lfloor n^a \rfloor$  with*

$$0 < a < \min \left( \frac{2\gamma_X^2(\beta_X - \varepsilon_X)}{2\gamma_X^2(\beta_X - \varepsilon_X) + 1 + \beta_X\gamma_X + \xi\gamma_X}, \frac{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y)}{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y) + 1}, \frac{-2\tau}{1 - 2\tau} \right),$$

for  $\varepsilon_X \in (0, \beta_X)$ ,  $\varepsilon_Y \in (0, \beta_Y)$  and  $\xi \in (0, \beta_X + (1 - \gamma_X)/\gamma_X)$ , we have for  $p$  satisfying  $\frac{k}{np} \rightarrow \infty$ ,  $\frac{\log k/(np)}{\sqrt{k}} \rightarrow 0$  and  $k = O(p^{2\tau(1-\gamma_X)})$

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}_\xi \left( \frac{\sqrt{k}}{\log \frac{k}{np}} \left( \frac{\hat{\theta}_{h,p}^*}{\hat{\theta}_{h,p}^+} - 1 \right) \leq t \right) - \mathbb{P} \left( \frac{\sqrt{k}}{\log \frac{k}{np}} \left( \frac{\hat{\theta}_{h,p}^+}{\theta_{h,p}} - 1 \right) \leq t \right) \right| \rightarrow 0,$$

in probability.

The result of Theorem 3.2 can be used to construct confidence intervals for  $\theta_{h,p}$ . Let  $q_l$  and  $q_u$  denote the lower and upper  $\alpha/2$  empirical quantiles in  $B$  bootstrap replicates of  $\hat{\theta}_{h,p}^*$ . Then the interval

$$\left[ \frac{(\hat{\theta}_{h,p}^+)^2}{q_u}, \frac{(\hat{\theta}_{h,p}^+)^2}{q_l} \right],$$

is a confidence interval for  $\theta_{h,p}$  with approximate coverage probability  $1 - \alpha$ .

## 4 Simulation results

In this section we evaluate the finite sample behavior of the proposed estimators with a simulation experiment. We consider the following models

**Model 1.** We generate  $(X_t, Y_t)$ ,  $t = 1, 2, \dots$ , as a Markovian time series, where  $(X_t, Y_t)$  has a joint distribution determined by a Gumbel copula given by

$$C_\zeta(u_1, u_2) = \exp \left( - \left[ (-\log u_1)^\zeta + (-\log u_2)^\zeta \right]^{1/\zeta} \right),$$

where  $\zeta \in [1, \infty)$ ,  $X_t$  follows a Burr( $\lambda, \eta$ ) distribution, with distribution function

$$F_X(x) = 1 - (1 + x^\eta)^{-\lambda}, \quad x > 0,$$

and  $\lambda, \eta > 0$ , and  $Y_t$  a Fréchet( $\varphi$ ) distribution, with distribution function

$$F_Y(x) = e^{-x^{-\varphi}}, \quad x > 0,$$

and  $\varphi > 0$ . The dependence between  $X_{t+1}$  and  $Y_t$  is modelled by a  $t$  copula function

$$C_{\nu, \mu}(u_1, u_2) = \int_{-\infty}^{F_\nu^{-1}(u_1)} \int_{-\infty}^{F_\nu^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\mu^2}} \left(1 + \frac{v_1^2 - 2\mu v_1 v_2 + v_2^2}{\nu(1-\mu^2)}\right)^{-(\nu+2)/2} dv_2 dv_1,$$

where  $\nu > 0$  denotes the degrees of freedom,  $\mu \in (-1, 1)$ , and  $F_\nu$  is the  $t$  distribution function with  $\nu$  degrees of freedom. Model 1 satisfies Assumption  $(\mathcal{D}_X)$  with  $\gamma_X = 1/(\lambda\eta)$  and  $\beta_X = \eta$ , and  $(\mathcal{D}_Y)$  with  $\gamma_Y = 1/\varphi$  and  $\beta_Y = \varphi$ . For the marginal distributions we consider the parameter settings  $(\lambda, \eta) = (0.5, 10), (1, 5), (2, 2.5)$  and  $\varphi = 10$ , which gives  $\gamma_X = 0.2, \gamma_Y = 0.1$  and the second order parameter  $\rho_X := -\beta_X\gamma_X$  equal to  $-2, -1$  and  $-0.5$ , respectively. For the copula functions we set  $(\zeta, \nu, \mu) = (2, 4, 0.5)$  and  $(10, 4, 0.8)$ .

**Model 2.** The time series is a AR(1) process with regularly varying innovations:

$$Z_t = AZ_{t-1} + B_t, \quad t \in \mathbb{Z},$$

where  $A$  is a deterministic  $\mathbb{R}^{2 \times 2}$ -matrix and  $B_t \in \mathbb{R}^2, t \in \mathbb{Z}$ , are independent and identically distributed multivariate regularly varying with positive index; see Janssen and Segers (2014) for further details. We set

$$A = \begin{bmatrix} 0.2 & 0.4 \\ 0.4 & 0.2 \end{bmatrix},$$

and  $B_t$  follows a bivariate  $t$  distribution with  $\nu = 4$ , and  $\mu = 0.3$  and  $0.8$ . According to Meinguet and Segers (2010), in this case the time series  $(Z_t)_{t \in \mathbb{Z}}$  is regularly varying with index  $1/\nu$ , and moreover it can be shown to satisfy Assumption  $(\mathcal{R})$ .

From each of the above models we simulate 1000 datasets with size  $n = 2000$ .

In Figures 1 and 2 we show for Model 1 the boxplots of  $\hat{\theta}_{h,p}/\theta_{h,p}$ , obtained in the 1000 replications, with  $p = 1/n$  at some values of  $k$ , and for  $h = 1$  (left) and  $h = 3$  (right), with  $(\zeta, \nu, \mu) = (2, 4, 0.5)$  and  $(10, 4, 0.8)$ , respectively. The rows of these figures correspond with the values of the second order parameter  $\rho_X$ , namely  $\rho_X = -2$  (top),  $-1$  (middle) and  $-0.5$  (bottom). The theoretical value of  $\theta_{h,p}$  is approximated by Monte-Carlo averaging, using 500 samples of size  $10^8$ . Overall, we see that the ratios are typically close to the value one at the smaller values of  $k$ , but there are differences according to the distribution parameters. Also, the estimation becomes more difficult with increasing values for  $h$ . For fixed marginal distributions, increasing the strength of the dependence improves the estimation results, as expected, as is clear from comparing Figures 1 and 2. For a fixed value of the copula dependence parameter, we see that the estimation results depend on the marginal distribution that was used. As is clear from Corollary 2.2 and Theorem 2.3, the limiting distribution of  $\hat{\theta}_{h,p}/\theta_{h,p}$ , properly normalised, is that of the estimator for  $\gamma_X$ . As for the latter, the bias is typically determined by the second

order parameter  $\rho_X$ , with larger values of  $|\rho_X|$  leading to smaller bias. This can also be observed in Figures 1 and 2 where the bias increases from the top to the bottom row.

Figure 3 shows the corresponding results for Model 2 with  $\mu = 0.3$  (top) and  $\mu = 0.8$  (bottom). Note that in Model 2, the assumption (8) is not satisfied, since the distribution of  $X_t$  has a lower and upper tail of equal tail heaviness. In that case, we can work separately on the two tails, but for brevity we only work on the positive part, and hence we consider the ratio  $\hat{\theta}_{h,p}^+/\theta_{h,p}^+$ . Overall, we can draw conclusions similar to those for Model 1: increasing the strength of the dependence improves the estimation results, and with increasing  $h$  the estimation becomes more difficult.

In Figure 4 we show the empirical coverage probabilities of approximate 95% multiplier block bootstrap confidence intervals for  $\theta_{h,p}$  where  $p = 1/n$ ,  $h = 1$  (black solid line) and  $h = 3$  (blue dashed line) as a function of  $k$ . Here we use  $r_n = 100$ ,  $(\xi_j)_{j \in \mathbb{N}}$  i.i.d.  $t_5$  (rescaled to unit variance) and  $B = 1000$ . For brevity, we only show the bootstrap results for the most challenging estimation cases, namely  $\rho_X = -0.5$  in Model 1 and  $\mu = 0.3$  in Model 2. As expected, the  $k$  range where the coverage probabilities are close to 95% corresponds to the  $k$  values where the boxplots of the estimates are centered close to the value one. Remark that the large variabilities of  $\hat{\theta}_{h,p}$  in Figure 2, are translated in wide confidence intervals and thus a longer stable range close to 95% of the coverage probabilities. In the stable parts of the plots of coverage probabilities,  $h = 1$  and  $h = 3$  show similar performance. Here a common value of  $r_n$  was chosen for all models, but a data driven method to determine  $r_n$  could possibly improve the results.

## 5 Illustration on river flow data

In hydrology, the river flow rate is an important parameter when assessing the risk of flooding. We consider a dataset of river flow rates, measured in  $m^3/s$ , at two stations along the river Coquet in Northumberland (UK), namely Morwick and Rothbury, where Morwick is downstream compared to Rothbury, covering the period September 1, 2023 till June 30, 2024, leading to  $n = 304$ . The dataset can be freely downloaded from the website <https://environment.data.gov.uk/hydrology/explore>. For our analysis we take as the  $X$ -variable the flow rate at Morwick and as  $Y$ -variable the flow rate at Rothbury. In this case,  $X$  and  $Y$  are nonnegative variables, and hence we can apply the methodology of Section 2. Concerning stationarity of the time series we note the following. Firstly, we did not include data for the summer months July and August, since from time series plots of the historical data it was clear that these months had lower flow rates, possibly due to lower precipitation in mid summer. Secondly, we assessed the stationarity of the retained data statistically by the augmented Dickey-Fuller test, which lead to  $p$ -values below 0.05 for values of the lag parameter up to 50, indicating stationarity. In Figure 5 we show the time plot of the flow rates at both locations. These plots clearly indicate the occurrence of severe peaks in the river flow rates, hinting at heavy-tailed underlying distribution. This is confirmed by the Pareto-quantile plots, shown in Figure 6, top row, as these plots show a linear behaviour for the extreme datapoints, supporting the assumption of underlying Pareto-type distributions. In Figure 6, bottom row, we show the corresponding Hill estimates of  $\gamma_X$  and  $\gamma_Y$  as a function of  $k$ . These plots show a stable pattern for  $k$  in the range 20-30 for Morwick and

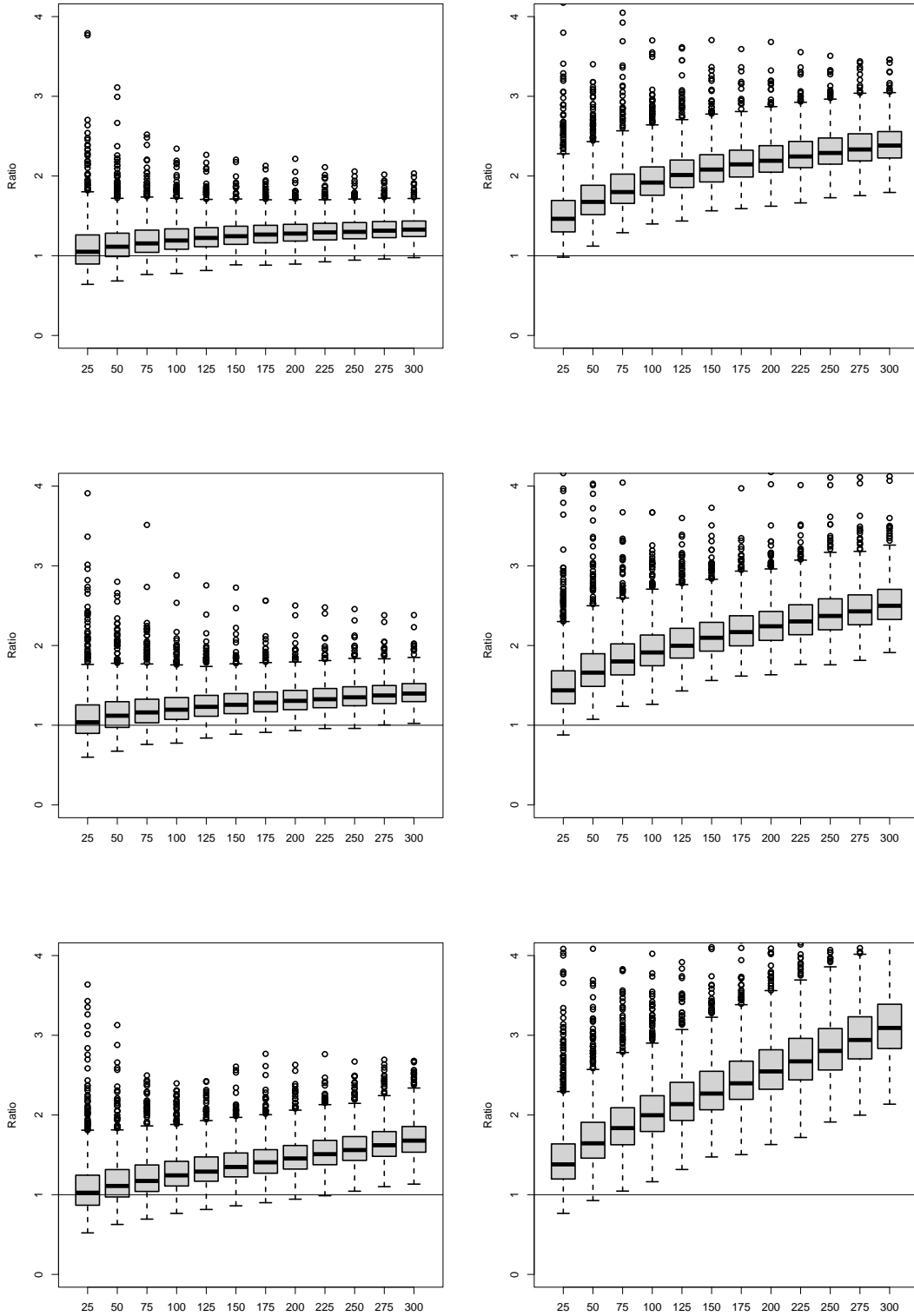


Figure 1: Model 1,  $(\zeta, \nu, \mu) = (2, 4, 0.5)$ . Boxplots of  $\hat{\theta}_{h,p}/\theta_{h,p}$  with  $p = 1/n$ , and for  $h = 1$  (left) and  $h = 3$  (right) as a function of  $k$ . Top row  $\rho_X = -2$ , middle row  $\rho_X = -1$  and bottom row  $\rho_X = -0.5$ .

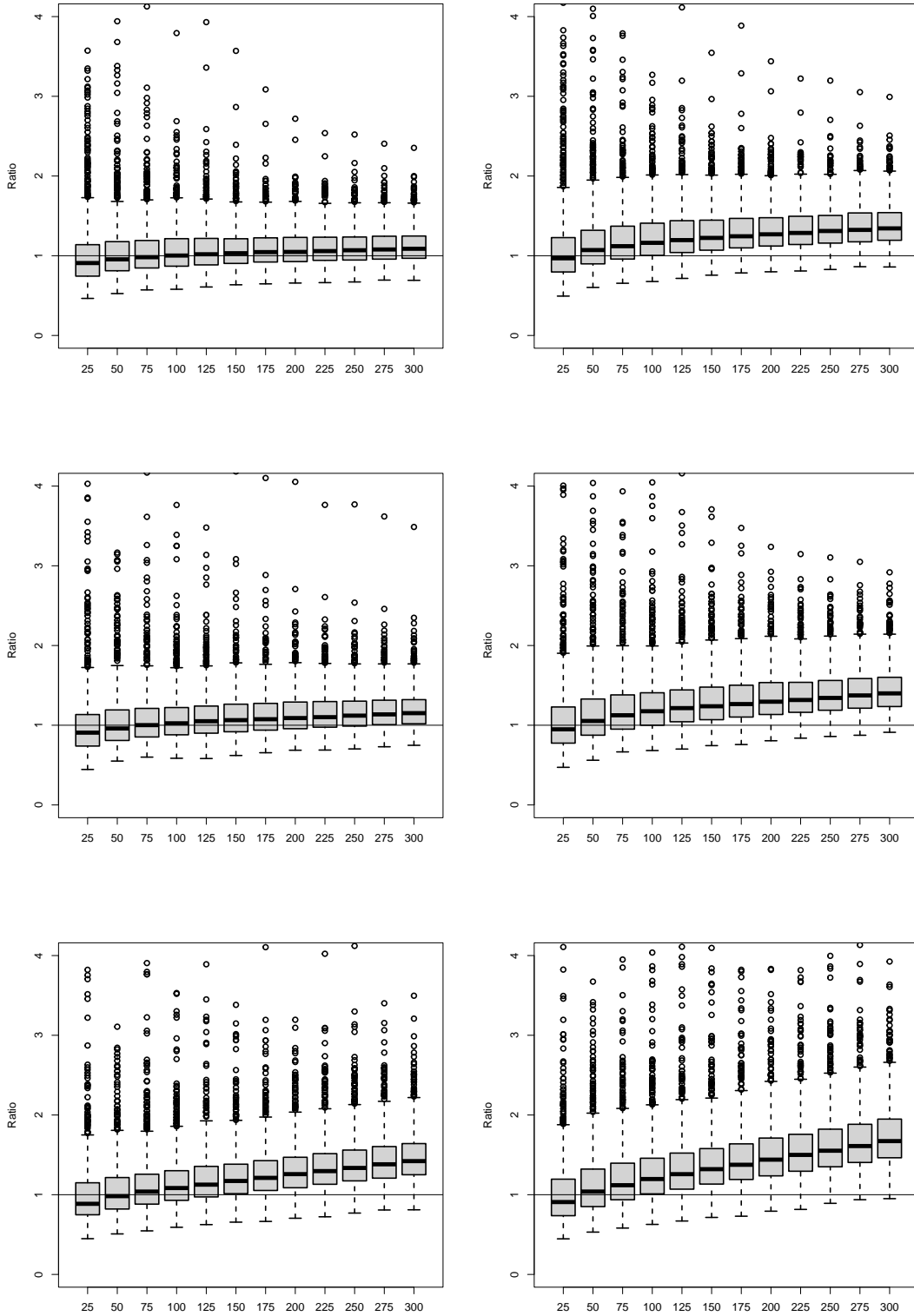


Figure 2: Model 1,  $(\zeta, \nu, \mu) = (10, 4, 0.8)$ . <sup>16</sup>Boxplots of  $\hat{\theta}_{h,p}/\theta_{h,p}$  with  $p = 1/n$ , for  $h = 1$  (left) and  $h = 3$  (right) as a function of  $k$ . Top row  $\rho_X = -2$ , middle row  $\rho_X = -1$  and bottom row  $\rho_X = -0.5$ .

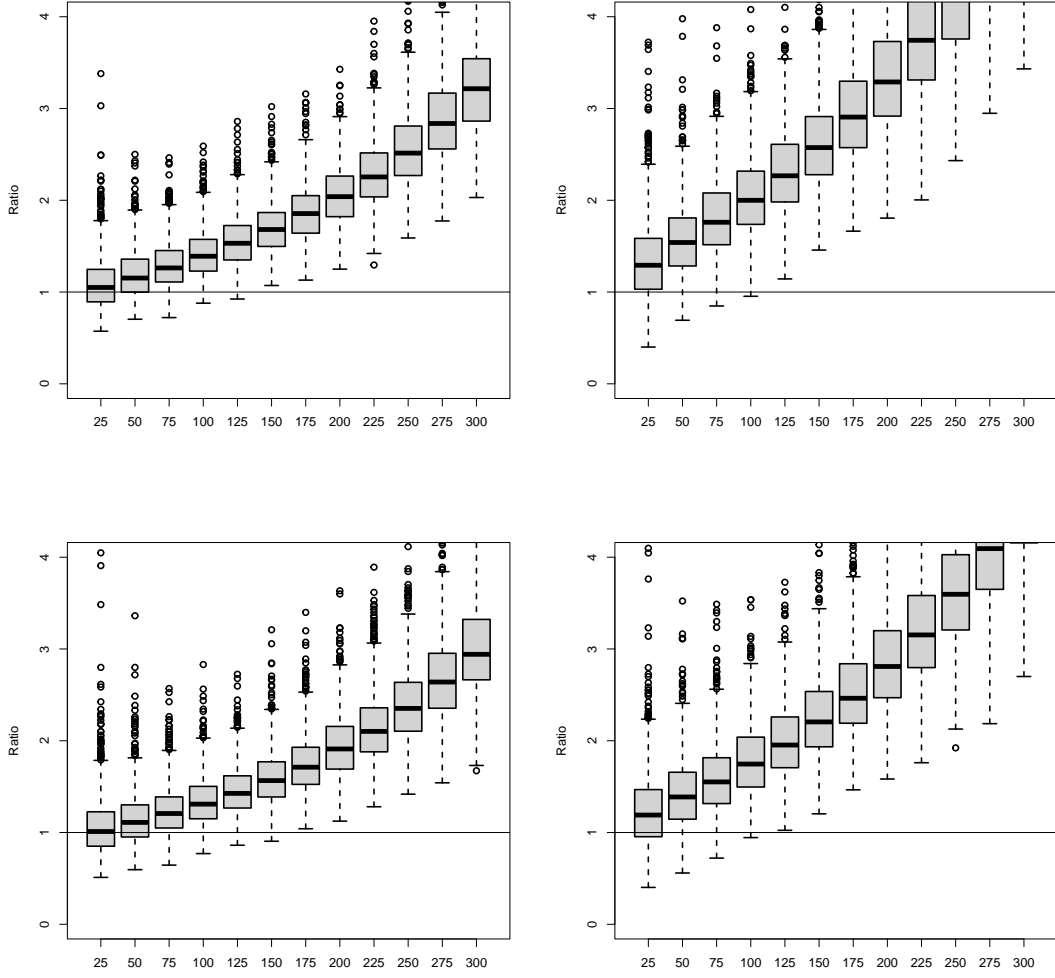


Figure 3: Model 2. Boxplots of  $\hat{\theta}_{h,p}^+/\theta_{h,p}^+$  with  $p = 1/n$ , for  $h = 1$  (left) and  $h = 3$  (right) as a function of  $k$ . Top row  $\mu = 0.3$ , bottom row  $\mu = 0.8$ .

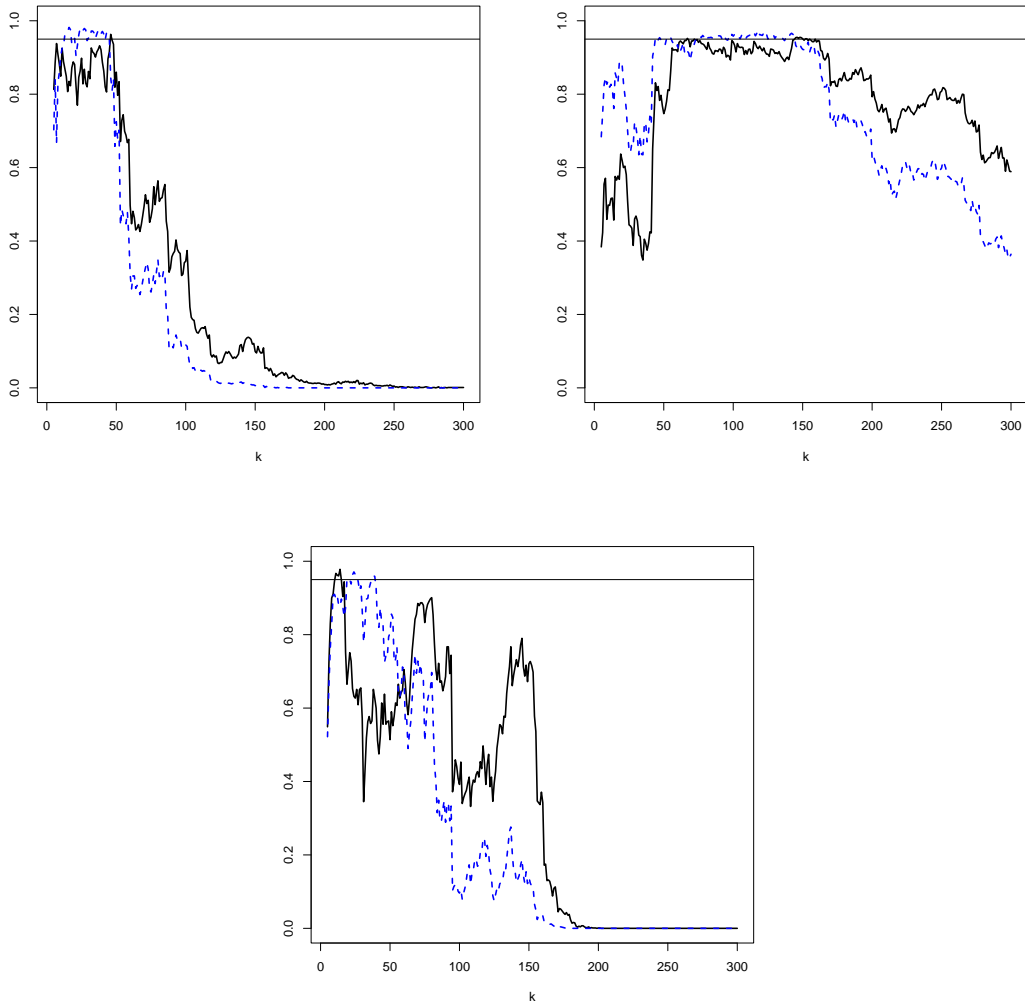


Figure 4: Empirical coverage probabilities of 95% multiplier block bootstrap confidence intervals for  $\theta_{h,p}$  with  $p = 1/n$ ,  $n = 2000$ ,  $r_n = 100$ , and  $h = 1$  (black solid line) and  $h = 3$  (blue dashed line), as a function of  $k$  for Model 1,  $(\zeta, \nu, \mu) = (2, 4, 0.5)$  and  $\rho_X = -0.5$  (top left), Model 1,  $(\zeta, \nu, \mu) = (10, 4, 0.8)$  and  $\rho_X = -0.5$  (top right), Model 2,  $\mu = 0.3$  (bottom).

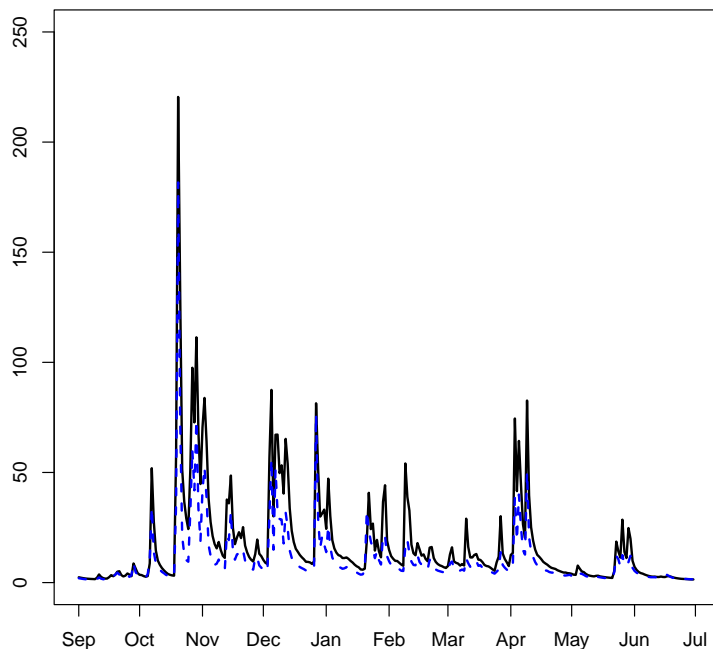


Figure 5: River flow data. Time series plot of the flow rates at Morwick (black solid line) and Rothbury (blue dashed line).

10-30 for Rothbury, approximately, indicating a tail index estimate of about 0.4 at both locations. In order to get an idea about the extreme dependence between  $X_h$  and  $Y_0$ , we constructed scatterplots of  $(X_{t+h}, Y_t)$  for  $h = 1, 2, 3$ , see the top row of Figure 7, where there is clearly a dependence in the largest observations, and which seems to decrease with  $h$ , which is further confirmed by the plots of an estimate of  $R(1, 1)$  (see, e.g., Schmidt and Stadtmüller, 2006) as a function of  $k$  given in the bottom row of the figure. Finally, in Figure 8, we plot  $\hat{\theta}_{h,p}$  along with approximate 95% multiplier block bootstrap confidence intervals as a function of  $k$  for  $h = 1$  (left) and 3 (right), and with  $p = 1/n$ . Overall, the estimates are more or less stable for  $k$  in the range 20-30, and they decrease with  $h$ . At  $k = 30$  we have  $\hat{\theta}_{1,p} = 249.89$  with confidence interval  $[110.68, 881.19]$  while  $\hat{\theta}_{3,p} = 180.99$  with confidence interval  $[94.96, 737.83]$ .

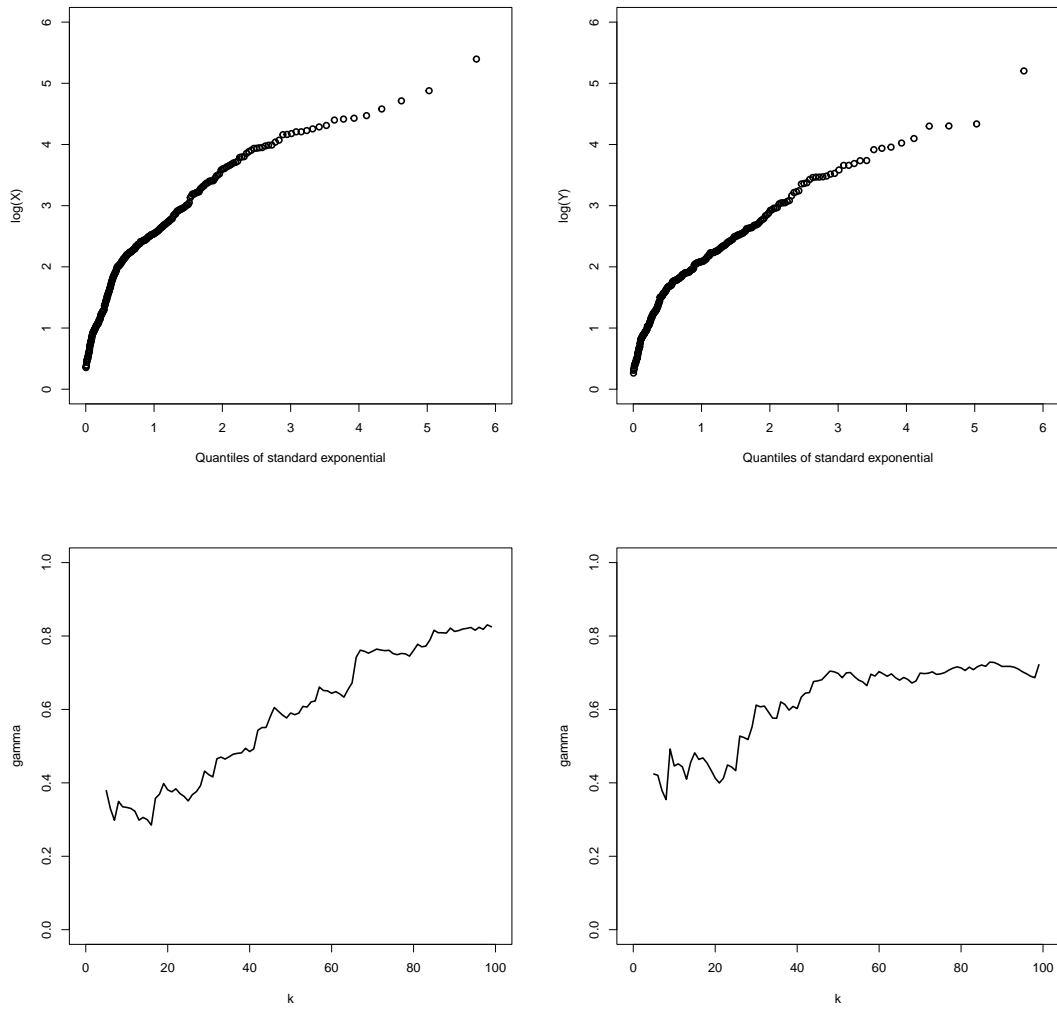


Figure 6: River flow data. Pareto quantile plot of the river flow rate at Morwick (top left) and Rothbury (top right), and  $\hat{\gamma}_{X,k}$  (bottom left) and  $\hat{\gamma}_{Y,k}$  (bottom right) as a function of  $k$ .

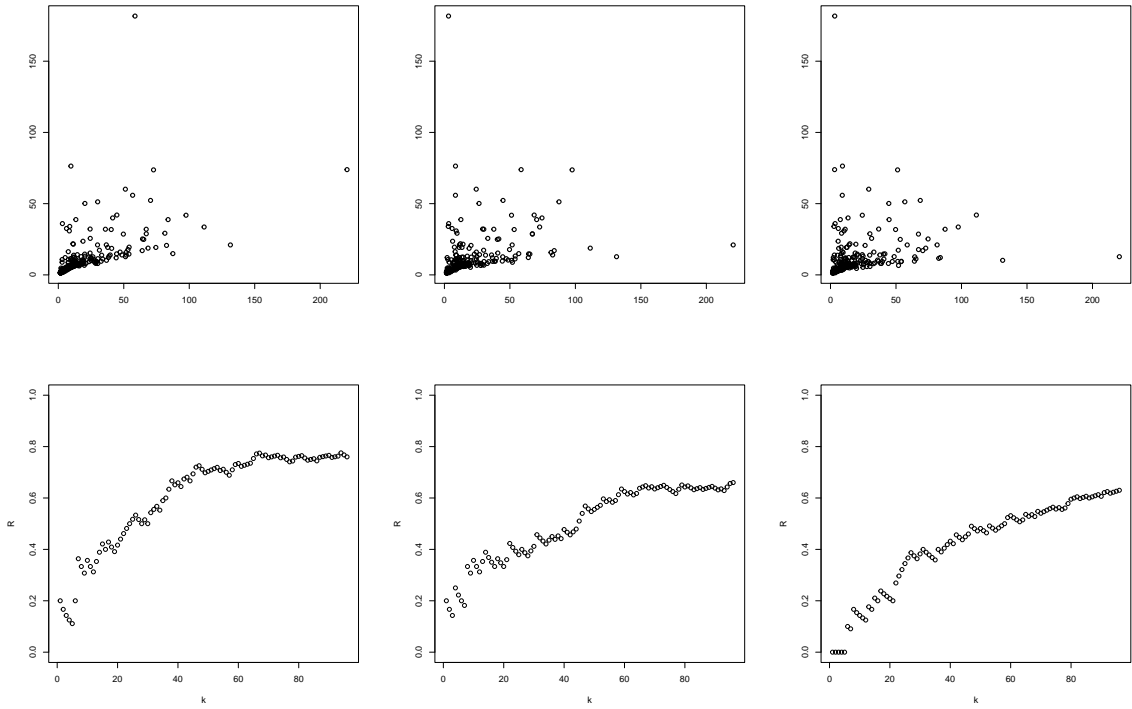


Figure 7: River flow data. Top row: scatterplots of  $(X_{t+h}, Y_t)$  for  $h = 1, 2, 3$ , from left to right. Bottom row:  $\hat{R}(1,1)$  as a function of  $k$  for  $h = 1, 2, 3$ , from left to right.

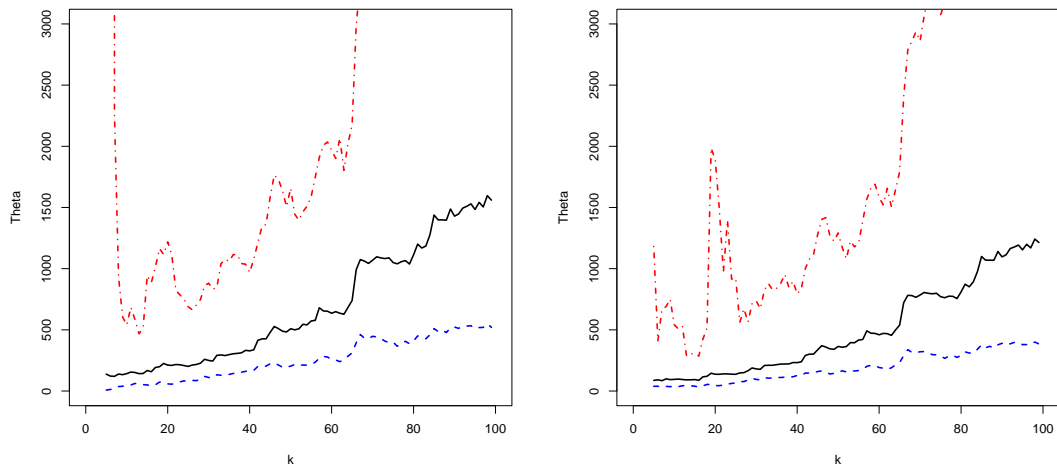


Figure 8: River flow data.  $\hat{\theta}_{h,p}$  along with approximate 95% multiplier block bootstrap confidence intervals as a function of  $k$  with  $h = 1$  (left) and  $3$  (right), and  $p = 1/n$ .

## 6 Proofs

### 6.1 Proof of Theorem 2.1

To prove this theorem, it remains to show that

$$\mathcal{Z}_{n,\varepsilon}(\Psi) := \frac{1}{\sqrt{nv_n}} \sum_{i=1}^n \{\Psi(Z_{n,i}) - \mathbb{E}(\Psi(Z_{n,1}))\} \rightsquigarrow \mathcal{Z}_\varepsilon(\Psi), \quad (9)$$

in  $\ell^\infty(\Phi)$ .

First we state and verify the assumptions from Drees and Rootzén (2010) for the class  $\Phi_\varepsilon := \{\phi_s : s \in [1 - \varepsilon, 1 + \varepsilon]\}$ . Let  $[0, \infty)_\cup := \bigcup_{r \in \mathbb{N}} [0, \infty)^{2r}$ , the set of all vectors with non-negative entries and arbitrary length, and  $\mathcal{F}_\varepsilon := \{f_s; s \in [1 - \varepsilon, 1 + \varepsilon]\}$  with

$$f_s(z) := \sum_{i=1}^r \phi_s(z_i),$$

for  $z = (z_1, \dots, z_r) \in [0, \infty)^{2r}$  and  $r \in \mathbb{N}$ . So  $\mathcal{F}_\varepsilon$  contains arbitrary long sums with summands coming from  $\Phi_\varepsilon$ .

**Assumption (D1):**  $\mathbb{E} \left[ \left\{ \sum_{i=1}^{r_n} \phi_s(Z_{n,i}) \right\}^2 \right]$  has to be finite for all  $s \in [1 - \varepsilon, 1 + \varepsilon]$  and all  $n \geq 1$ , and the envelope function

$$F_\varepsilon(z) := \sup_{f \in \mathcal{F}_\varepsilon} f(z),$$

has to be finite for all  $z \in [0, \infty)_\cup$ .

**Proof of Assumption (D1).** We have

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \phi_s(Z_{n,i}) \right)^2 \right] \\ & \leq \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \right)^2 \right] \\ & = \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left[ \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \frac{X_{j+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_j}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \right] \\ & \leq 2 r_n \sum_{i=0}^{r_n-1} \mathbb{E} \left[ \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \right] \\ & \leq 2 r_n v_n \sum_{i=0}^{r_n-1} \mathbb{E} \left[ \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right\}} \mid \frac{X_h}{U_X(n/k)} > 1 - \varepsilon \text{ or } \frac{Y_0}{\bar{U}_Y(n/k)} > 1 - \varepsilon \right] \\ & \leq 2 r_n v_n \sum_{i=0}^{r_n-1} s_n(i) \\ & = O(r_n v_n) = o(1), \end{aligned} \quad (10)$$

by Conditions **(A)** and **(B)**. Now  $\phi_s(z) \leq \phi_{1-\varepsilon}(z)$ ,  $\forall z$  and  $\forall s \in [1 - \varepsilon, 1 + \varepsilon]$ . This ends the proof of Assumption  $(D1)$ .  $\square$

**Assumption  $(D2')$ :** We have,  $\forall \tilde{\varepsilon} > 0$

$$\mathbb{E} \left[ \left\{ \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \right\}^2 \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] = o(r_n v_n).$$

**Proof of Assumption  $(D2')$ .** We have, for  $\tilde{\varepsilon} > 0$ , using Hölder's and Markov's inequalities combined with (10)

$$\begin{aligned} & \mathbb{E} \left[ \left\{ \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \right\}^2 \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] \\ &= \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left[ \frac{X_{i+h}}{U_X(n/k)} \frac{X_{j+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon, \frac{Y_j}{U_Y(n/k)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] \\ &\leq \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \left\{ \mathbb{E} \left[ \left( \frac{X_{i+h}}{U_X(n/k)} \frac{X_{j+h}}{U_X(n/k)} \right)^{1+\delta} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon, \frac{Y_j}{U_Y(n/k)} > 1-\varepsilon \right\}} \right] \right\}^{\frac{1}{1+\delta}} \\ &\quad \times \left\{ \mathbb{P} \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} > \tilde{\varepsilon} \sqrt{n v_n} \right) \right\}^{\frac{\delta}{1+\delta}} \\ &\leq 2 r_n \sum_{i=0}^{r_n-1} \left\{ \mathbb{E} \left[ \left( \frac{X_{i+h}}{U_X(n/k)} \frac{X_h}{U_X(n/k)} \right)^{1+\delta} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon, \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}} \right] \right\}^{\frac{1}{1+\delta}} \\ &\quad \times \left\{ \frac{1}{\tilde{\varepsilon}^2 n v_n} \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \right)^2 \right] \right\}^{\frac{\delta}{1+\delta}} \\ &\leq 2 r_n v_n^{\frac{1}{1+\delta}} \times O \left( \left[ \frac{1}{n v_n} r_n v_n \right]^{\frac{\delta}{1+\delta}} \right) \\ &\times \sum_{i=0}^{r_n-1} \left\{ \mathbb{E} \left[ \left( \frac{X_{i+h}}{U_X(n/k)} \frac{X_h}{U_X(n/k)} \right)^{1+\delta} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \mid \frac{X_h}{U_X(n/k)} > 1-\varepsilon \text{ or } \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right] \right\}^{\frac{1}{1+\delta}} \\ &= O \left( r_n v_n \left( \frac{r_n}{n v_n} \right)^{\frac{\delta}{1+\delta}} \right) \text{ by Condition **(B)**} \\ &= O \left( r_n v_n \left( \frac{r_n}{\sqrt{n v_n}} \right)^{\frac{\delta}{1+\delta}} \left( \frac{1}{\sqrt{n v_n}} \right)^{\frac{\delta}{1+\delta}} \right) \\ &= o(r_n v_n) \text{ by Condition **(A)**.} \end{aligned}$$

This achieves the proof of Assumption  $(D2')$ .  $\square$

**Assumption (D3):** There exists a semi-metric  $\rho$  such that

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \sup_{\substack{1-\varepsilon \leq s, t \leq 1+\varepsilon \\ \rho(s, t) < \delta}} \frac{1}{r_n v_n} \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > s \right\}} - \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > t \right\}} \right)^2 \right] = 0.$$

**Proof of Assumption (D3).** Let  $s, t \in [1 - \varepsilon, 1 + \varepsilon]$ , and, without loss of generality, assume  $s < t$ . Then

$$\begin{aligned} & \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > s \right\}} - \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > t \right\}} \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ s < \frac{Y_i}{U_Y(n/k)} \leq t \right\}} \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ s < \frac{Y_i}{U_Y(n/k)} \leq t \right\}} \right)^2 \right] \\ &= \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left( \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \frac{X_{j+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_j}{U_Y(n/k)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ s < \frac{Y_i}{U_Y(n/k)} \leq t \right\}} \mathbb{1}_{\left\{ s < \frac{Y_j}{U_Y(n/k)} \leq t \right\}} \right) \\ &\leq \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \left\{ \mathbb{E} \left[ \left( \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \frac{X_{j+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_j}{U_Y(n/k)} > 1-\varepsilon \right\}} \right)^{1+\delta} \right] \right\}^{\frac{1}{1+\delta}} \\ &\quad \times \left\{ \mathbb{P} \left( s < \frac{Y_i}{U_Y(n/k)} \leq t, s < \frac{Y_j}{U_Y(n/k)} \leq t \right) \right\}^{\frac{\delta}{1+\delta}} \text{ by Hölder's inequality} \\ &\leq 2 r_n v_n^{\frac{1}{1+\delta}} \left[ \overline{F}_Y(s U_Y(n/k)) - \overline{F}_Y(t U_Y(n/k)) \right]^{\frac{\delta}{1+\delta}} \\ &\times \sum_{i=0}^{r_n-1} \left\{ \mathbb{E} \left[ \left( \frac{X_h}{U_X(n/k)} \frac{X_{i+h}}{U_X(n/k)} \right)^{1+\delta} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > 1-\varepsilon \right\}} \mid \frac{X_h}{U_X(n/k)} > 1-\varepsilon \text{ or } \frac{Y_0}{U_Y(n/k)} > 1-\varepsilon \right] \right\}^{\frac{1}{1+\delta}}. \end{aligned}$$

Denoting by  $C$  a positive constant, independent of  $s$  and  $t$ , that can change from one line to

another, this implies, by Condition **(B)**, that

$$\begin{aligned}
& \frac{1}{r_n v_n} \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > s \right\}} - \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > t \right\}} \right)^2 \right] \\
& \leq C \left[ \frac{\bar{F}_Y((1-\varepsilon)U_Y(n/k))}{v_n} \right]^{\frac{\delta}{1+\delta}} \left[ \frac{\bar{F}_Y(sU_Y(n/k)) - \bar{F}_Y(tU_Y(n/k))}{\bar{F}_Y((1-\varepsilon)U_Y(n/k))} \right]^{\frac{\delta}{1+\delta}} \\
& \leq C \left[ \frac{\bar{F}_Y(sU_Y(n/k)) - \bar{F}_Y(tU_Y(n/k))}{\bar{F}_Y((1-\varepsilon)U_Y(n/k))} \right]^{\frac{\delta}{1+\delta}} \\
& \leq C \left\{ \left| \frac{\bar{F}_Y\left(\frac{s}{1-\varepsilon}(1-\varepsilon)U_Y(n/k)\right)}{\bar{F}_Y((1-\varepsilon)U_Y(n/k))} - \left(\frac{s}{1-\varepsilon}\right)^{-\frac{1}{\gamma_Y}} \right| + \left| \frac{\bar{F}_Y\left(\frac{t}{1-\varepsilon}(1-\varepsilon)U_Y(n/k)\right)}{\bar{F}_Y((1-\varepsilon)U_Y(n/k))} - \left(\frac{t}{1-\varepsilon}\right)^{-\frac{1}{\gamma_Y}} \right| \right. \\
& \quad \left. + \left| \left(\frac{s}{1-\varepsilon}\right)^{-\frac{1}{\gamma_Y}} - \left(\frac{t}{1-\varepsilon}\right)^{-\frac{1}{\gamma_Y}} \right| \right]^{\frac{\delta}{1+\delta}} \\
& =: C \{Q_{1,n} + Q_{2,n} + Q_{3,n}\}^{\frac{\delta}{1+\delta}}.
\end{aligned}$$

Note that, by regular variation,  $Q_{1,n}$  and  $Q_{2,n}$  tend to 0 uniformly for  $s, t \in [1-\varepsilon, 1+\varepsilon]$ , see Theorem B.1.4 in de Haan and Ferreira (2006), and

$$Q_{3,n} \leq C \left| s^{-\frac{1}{\gamma_Y}} - t^{-\frac{1}{\gamma_Y}} \right| =: C \rho(s, t).$$

We deduce then that

$$\limsup_{n \rightarrow \infty} \sup_{\substack{1-\varepsilon \leq s \leq t \leq 1+\varepsilon \\ \rho(s,t) < \tilde{\delta}}} \frac{1}{r_n v_n} \mathbb{E} \left[ \left( \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > s \right\}} - \sum_{i=1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > t \right\}} \right)^2 \right] \leq C \tilde{\delta}^{\frac{\delta}{1+\delta}}.$$

This achieves the proof of Assumption **(D3)**.  $\square$

**Assumption (D5):** Let  $\tilde{Z}_{n,j} := (\tilde{Z}_{n,j,i})_{1 \leq i \leq r_n}$  be independent copies of  $(Z_{n,i})_{1 \leq i \leq r_n}$ , for  $1 \leq j \leq \lfloor n/(2r_n) \rfloor$ . The map

$$\sup_{\substack{1-\varepsilon \leq s, t \leq 1+\varepsilon \\ \rho(s,t) < \tilde{\delta}}} \sum_{j=1}^{\lfloor n/(2r_n) \rfloor} e_j \left\{ \sum_{i=1}^{r_n} \left[ \phi_s(\tilde{Z}_{n,j,i}) - \phi_t(\tilde{Z}_{n,j,i}) \right] \right\}^k,$$

is measurable,  $\forall \tilde{\delta} > 0, n \in \mathbb{N}, (e_j)_{1 \leq j \leq \lfloor n/(2r_n) \rfloor} \in \{-1, 0, 1\}^{\lfloor n/(2r_n) \rfloor}$  and  $k \in \{1, 2\}$ .

The proof is completely similar to the one in Goegebeur et al. (2025).  $\square$

**Assumption (D6'):** The envelope function  $F_\varepsilon$  is measurable with

$$\mathbb{E} \left[ \left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) \right\}^2 \right] = O(r_n v_n)$$

and

$$\int_0^1 \sup_{Q \in \mathcal{Q}} \sqrt{\log \mathcal{N} \left( \zeta \left( \int F_\varepsilon^2 dQ \right)^{1/2}, \mathcal{F}_\varepsilon, L_2(Q) \right)} d\zeta < \infty,$$

where  $\mathcal{N}$  denotes the covering numbers and  $Q$  ranges over the set of discrete probability measures  $\mathcal{Q}$ .

The proof is completely similar to the one in Goegebeur et al. (2025).  $\square$

We also need to verify the two following conditions of Theorem 2.10 in Drees and Rootzén (2010), ensuring that all finite-dimensional distributions converge:

**Assumption (C1):** We have

$$\begin{aligned} \mathbb{E} \left[ (\Delta_n(s) - \mathbb{E}\Delta_n(s))^2 \mathbb{1}_{\{|\Delta_n(s) - \mathbb{E}\Delta_n(s)| \leq \sqrt{n v_n}\}} \right] &= o(r_n v_n), \\ \mathbb{P} (|\Delta_n(s) - \mathbb{E}\Delta_n(s)| > \sqrt{n v_n}) &= o\left(\frac{r_n}{n}\right), \end{aligned}$$

for all  $s \in [1 - \varepsilon, 1 + \varepsilon]$ , where

$$\Delta_n(s) := \sum_{i=r_n - \ell_n + 1}^{r_n} \frac{X_{i+h}}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k)} > s \right\}}.$$

The proof is completely similar to the one in Goegebeur et al. (2025).  $\square$

**Assumption (C3):** We have for all  $s, t \in [1 - \varepsilon, 1 + \varepsilon]$

$$\begin{aligned} \frac{1}{r_n v_n} \text{Cov} \left( \sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \\ \longrightarrow \mathbb{E} [\phi_s(\mathcal{W}_0) \phi_t(\mathcal{W}_0)] + \sum_{i=1}^{\infty} [\mathbb{E} (\phi_s(\mathcal{W}_0) \phi_t(\mathcal{W}_i)) + \mathbb{E} (\phi_s(\mathcal{W}_i) \phi_t(\mathcal{W}_0))]. \end{aligned}$$

**Proof of Assumption (C3).** We have

$$\begin{aligned}
& \frac{1}{r_n v_n} \text{Cov} \left( \sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \\
&= \frac{1}{r_n v_n} \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \{ \mathbb{E} [\phi_s(Z_{n,i}) \phi_t(Z_{n,j})] - \mathbb{E} [\phi_s(Z_{n,0})] \mathbb{E} [\phi_t(Z_{n,0})] \} \\
&= \frac{1}{r_n v_n} \left\{ \sum_{i=1}^{r_n-1} (r_n - i) \{ \mathbb{E} [\phi_s(Z_{n,0}) \phi_t(Z_{n,i})] + \mathbb{E} [\phi_s(Z_{n,i}) \phi_t(Z_{n,0})] \} \right. \\
&\quad \left. + r_n \mathbb{E} [\phi_s(Z_{n,0}) \phi_t(Z_{n,0})] - r_n^2 \mathbb{E} [\phi_s(Z_{n,0})] \mathbb{E} [\phi_t(Z_{n,0})] \right\} \\
&= \sum_{i=1}^{r_n-1} \left( 1 - \frac{i}{r_n} \right) \{ \mathbb{E} [\phi_s(Z_{n,0}) \phi_t(Z_{n,i}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
&\quad + \mathbb{E} [\phi_s(Z_{n,i}) \phi_t(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \} \\
&\quad + \mathbb{E} [\phi_s(Z_{n,0}) \phi_t(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
&\quad - r_n v_n \mathbb{E} [\phi_s(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
&\quad \times \mathbb{E} [\phi_t(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)].
\end{aligned}$$

Now, we look at

$$\begin{aligned}
& \mathbb{E} [\phi_s(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
&= \frac{1}{v_n} \mathbb{E} \left( \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > s \right\}} \right) \\
&= \frac{1}{L((1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}})} (1 + o(1)) \int_0^\infty \frac{n}{k} \mathbb{P} \left( \frac{X_h}{U_X(n/k)} > x, \frac{Y_0}{U_Y(n/k)} > s \right) dx \\
&= \frac{1}{L((1 - \varepsilon)^{-\frac{1}{\gamma_X}}, (1 - \varepsilon)^{-\frac{1}{\gamma_Y}})} (1 + o(1)) \int_0^\infty R_{n/k} \left( \frac{\bar{F}_X(x U_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y(s U_Y(n/k))}{\bar{F}_Y(U_Y(n/k))} \right) dx.
\end{aligned}$$

We will use the dominated convergence theorem. First, note that using Assumption  $(\mathcal{R})$  combined with the Pareto-type models (2) for  $X$  and  $Y$ , we have

$$\lim_{n \rightarrow \infty} R_{n/k} \left( \frac{\bar{F}_X(x U_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y(s U_Y(n/k))}{\bar{F}_Y(U_Y(n/k))} \right) = R \left( x^{-\frac{1}{\gamma_X}}, s^{-\frac{1}{\gamma_Y}} \right).$$

Now, by Proposition B.1.9.5 in de Haan and Ferreira (2006), for  $C_1, C_2, \tilde{\varepsilon}, \tilde{\delta} > 0$  and  $n$  large

$$\begin{aligned}
& R_{n/k} \left( \frac{\overline{F}_X(x U_X(n/k))}{\overline{F}_X(U_X(n/k))}, \frac{\overline{F}_Y(s U_Y(n/k))}{\overline{F}_Y(U_Y(n/k))} \right) \\
& \leq \min \left\{ \frac{\overline{F}_X(x U_X(n/k))}{\overline{F}_X(U_X(n/k))}, \frac{\overline{F}_Y(s U_Y(n/k))}{\overline{F}_Y(U_Y(n/k))} \right\} \\
& \leq \frac{\overline{F}_Y(s U_Y(n/k))}{\overline{F}_Y(U_Y(n/k))} \mathbb{1}_{\{x \leq \frac{U_X(1/\overline{F}_Y(s U_Y(n/k)))}{U_X(n/k)}\}} \\
& \quad + \frac{\overline{F}_X(x U_X(n/k))}{\overline{F}_X(U_X(n/k))} \mathbb{1}_{\{x > \frac{U_X(1/\overline{F}_Y(s U_Y(n/k)))}{U_X(n/k)}\}} \\
& \leq (1 + \tilde{\varepsilon}) s^{-\frac{1}{\gamma_Y} \pm \tilde{\delta}} \mathbb{1}_{\{x \leq \frac{U_X(1/\overline{F}_Y(s U_Y(n/k)))}{U_X(n/k)}\}} \\
& \quad + (1 + \tilde{\varepsilon}) x^{-\frac{1}{\gamma_X} \pm \tilde{\delta}} \mathbb{1}_{\{x > \frac{U_X(1/\overline{F}_Y(s U_Y(n/k)))}{U_X(n/k)}\}} \\
& \leq (1 + \tilde{\varepsilon}) s^{-\frac{1}{\gamma_Y} \pm \tilde{\delta}} \mathbb{1}_{\{x \leq C_1 s^{\frac{\gamma_X}{\gamma_Y} \pm \tilde{\delta}}\}} + (1 + \tilde{\varepsilon}) x^{-\frac{1}{\gamma_X} \pm \tilde{\delta}} \mathbb{1}_{\{x > C_2 s^{\frac{\gamma_X}{\gamma_Y} \pm \tilde{\delta}}\}} \\
& =: g(x),
\end{aligned}$$

where  $x^{\pm \xi} = x^\xi \mathbb{1}_{\{x \geq 1\}} + x^{-\xi} \mathbb{1}_{\{x < 1\}}$ . This function  $g$  is integrable on  $[0, \infty)$  by selecting  $\tilde{\delta} \in \left(0, \frac{1}{\gamma_X} - 1\right)$ , since  $\gamma_X < 1$ .

Consequently

$$\mathbb{E}[\phi_s(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] < \infty,$$

from which we deduce that

$$\begin{aligned}
& \frac{1}{r_n v_n} \text{Cov} \left( \sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \\
& = \sum_{i=1}^{r_n-1} \left( 1 - \frac{i}{r_n} \right) \{ \mathbb{E}[\phi_s(Z_{n,0}) \phi_t(Z_{n,i}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
& \quad + \mathbb{E}[\phi_s(Z_{n,i}) \phi_t(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \} \\
& \quad + \mathbb{E}[\phi_s(Z_{n,0}) \phi_t(Z_{n,0}) | X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k)] \\
& \quad + O(r_n v_n).
\end{aligned}$$

Now, under our Assumption  $(\tilde{\mathcal{R}})$ , denoting by  $\mathcal{L}(V)$  the distribution of a vector  $V$ , we have

$$\begin{aligned}
& \mathcal{L} \left( \left( \frac{X_h}{U_X(n/k)}, \frac{Y_0}{U_Y(n/k)}, \frac{X_{i+h}}{U_X(n/k)}, \frac{Y_i}{U_Y(n/k)} \right) \middle| X_h > (1 - \varepsilon)U_X(n/k) \text{ or } Y_0 > (1 - \varepsilon)U_Y(n/k) \right) \\
& \rightsquigarrow \mathcal{L} \left( (W_h, \tilde{W}_0, W_{i+h}, \tilde{W}_i) \right),
\end{aligned}$$

where  $(W_h, \widetilde{W}_0, W_{i+h}, \widetilde{W}_i)$  has a survival distribution function given by (7). Thus

$$\begin{aligned} & \mathcal{L}\left((Z_{n,0}, Z_{n,i}) \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right) \\ \rightsquigarrow & \mathcal{L}\left(W_h \mathbb{1}_{\{W_h > 1-\varepsilon \text{ or } \widetilde{W}_0 > 1-\varepsilon\}}, \widetilde{W}_0 \mathbb{1}_{\{W_h > 1-\varepsilon \text{ or } \widetilde{W}_0 > 1-\varepsilon\}}, \right. \\ & \left. W_{i+h} \mathbb{1}_{\{W_{i+h} > 1-\varepsilon \text{ or } \widetilde{W}_i > 1-\varepsilon\}}, \widetilde{W}_i \mathbb{1}_{\{W_{i+h} > 1-\varepsilon \text{ or } \widetilde{W}_i > 1-\varepsilon\}}\right) \\ = & \mathcal{L}(\mathcal{W}_0, \mathcal{W}_i). \end{aligned}$$

This implies that

$$\mathcal{L}\left((\phi_s(Z_{n,0})\phi_t(Z_{n,i})) \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right) \rightsquigarrow \mathcal{L}(\phi_s(\mathcal{W}_0)\phi_t(\mathcal{W}_i)).$$

Using now the corollary of Theorem 25.12 in Billingsley (1995, p. 338), since by Condition **(B)** for some  $\delta > 0$

$$\sup_n \mathbb{E}\left[\{\phi_s(Z_{n,0})\phi_t(Z_{n,i})\}^{1+\delta} \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right] < \infty,$$

we have

$$\mathbb{E}\left[\phi_s(Z_{n,0})\phi_t(Z_{n,i}) \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right] \longrightarrow \mathbb{E}[\phi_s(\mathcal{W}_0)\phi_t(\mathcal{W}_i)].$$

Now, remark that

$$\begin{aligned} 0 & \leq \mathbb{E}\left[\phi_s(Z_{n,0})\phi_t(Z_{n,i}) \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right] \\ & \leq \mathbb{E}\left[\phi_{1-\varepsilon}(Z_{n,0})\phi_{1-\varepsilon}(Z_{n,i}) \middle| X_h > (1-\varepsilon)U_X(n/k) \text{ or } Y_0 > (1-\varepsilon)U_Y(n/k)\right] \\ & \leq s_n(i). \end{aligned}$$

Again by Condition **(B)**, we can apply Pratt's lemma (see Pratt, 1960) in order to deduce the convergence of Assumption (C3).  $\square$

This achieves the weak convergence of (9) in  $\ell^\infty(\Phi_\varepsilon)$ .

We now briefly comment on the weak convergence in  $\ell^\infty(\Phi)$ . Similarly to the above we need to verify the assumptions in Drees and Rootzén (2010), but now for  $\Phi = \Phi_\varepsilon \cup \Phi_{D,H,\varepsilon}$  where  $\Phi_{D,H,\varepsilon} := \{\phi_{D_X,s}, \phi_{D_Y,s}, \phi_{H,s} : s \in [1-\varepsilon, 1+\varepsilon]\}$ . Remember that the convergence (9) for the function class  $\Phi_{D,H,\varepsilon}$  was already established in Drees and Knežević (2020). Assumption (C1) needs to hold for each individual element of a function class, so it is satisfied for  $\Phi$ . For Assumption (C3) we only need to compute the covariances between the functions from  $\Phi_\varepsilon$  and those from  $\Phi_{D,H,\varepsilon}$ , but this is readily done with arguments similar to the above. Concerning the asymptotic equicontinuity (the verification of Assumptions (D1), (D2'), (D3), (D5) and (D6')), one can follow the argument on p. 48 in Knežević (2020), according to which it is sufficient to verify the asymptotic equicontinuity for the subclasses  $\Phi_\varepsilon$  and  $\Phi_{D,H,\varepsilon}$ . As we have this, the convergence (9) in  $\ell^\infty(\Phi)$  follows.  $\square$

## 6.2 Proof of Theorem 2.2

We have the decomposition

$$\begin{aligned} \sqrt{k} \left( \frac{\bar{\theta}_{h,k/n}}{\theta_{h,k/n}} - 1 \right) &= \frac{U_X(n/k)}{\theta_{h,k/n}} \left\{ \sqrt{k} \left[ \bar{\theta}_n \left( \frac{Y_{(n-k)}}{U_Y(n/k)} \right) - \mathbb{E} \left( \bar{\theta}_n \left( \frac{Y_{(n-k)}}{U_Y(n/k)} \right) \right) \right] \right. \\ &\quad \left. + \sqrt{k} \left[ \mathbb{E} \left( \bar{\theta}_n \left( \frac{Y_{(n-k)}}{U_Y(n/k)} \right) \right) - \mathbb{E} (\bar{\theta}_n(1)) \right] \right\} \\ &=: \frac{U_X(n/k)}{\theta_{h,k/n}} \{Q_{4,n} + Q_{5,n}\}. \end{aligned}$$

**Study of  $Q_{4,n}$ .** By Theorem 2.1 and Corollary 2.1, we have that  $\left( \mathcal{Z}_n, \sqrt{k} \left( \frac{Y_{(n-k)}}{U_Y(n/k)} - 1 \right) \right) \rightsquigarrow (\mathcal{Z}, \gamma_Y \mathcal{Z}(\phi_{D_Y,1}))$ . Using Skorohod's theorem there exists also versions for which the convergence is almost surely. We will apply the Skorohod theorem throughout the proof, though without a change in notation.

For  $n$  large

$$|Q_{4,n} - \mathcal{Z}(\phi_1)| \leq \sup_{s \in [1-\varepsilon, 1+\varepsilon]} |\mathcal{Z}_n(\phi_s) - \mathcal{Z}(\phi_s)| + \left| \mathcal{Z} \left( \phi_{\frac{Y_{(n-k)}}{U_Y(n/k)}} \right) - \mathcal{Z}(\phi_1) \right|,$$

and hence  $|Q_{4,n} - \mathcal{Z}(\phi_1)| \rightarrow 0$  almost surely as  $n \rightarrow \infty$ .

**Study of  $Q_{5,n}$ .** We have

$$\begin{aligned} \mathbb{E}(\bar{\theta}_n(s)) &= \frac{n}{k} \mathbb{E} \left( \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > s \right\}} \right) \\ &= \int_0^\infty \frac{n}{k} \mathbb{P} \left( \frac{X_h}{U_X(n/k)} \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k)} > s \right\}} > x \right) dx \\ &= \int_0^\infty \frac{n}{k} \mathbb{P} \left( \frac{X_h}{U_X(n/k)} > x, \frac{Y_0}{U_Y(n/k)} > s \right) dx \\ &= \int_0^\infty R_{n/k} \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y(sU_Y(n/k))}{\bar{F}_Y(U_Y(n/k))} \right) dx. \end{aligned}$$

We have on the set  $\left\{ \frac{Y_{(n-k)}}{U_Y(n/k)} \in [1 - \varepsilon, 1 + \varepsilon] \right\}$ :

$$\begin{aligned}
Q_{5,n} &= \sqrt{k} \left\{ \int_0^\infty R_{n/k} \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \right) dx \right. \\
&\quad \left. - \int_0^\infty R_{n/k} \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1 \right) dx \right\} \\
&= \sqrt{k} \int_0^\infty \left[ R \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \right) - R \left( x^{-\frac{1}{\gamma_X}}, 1 \right) \right] dx \\
&\quad + \sqrt{k} \int_0^\infty \left[ R_{n/k} \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \right) \right. \\
&\quad \left. - R \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \right) \right] dx \\
&\quad - \sqrt{k} \int_0^\infty \left[ R \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1 \right) - R \left( x^{-\frac{1}{\gamma_X}}, 1 \right) \right] dx \\
&\quad - \sqrt{k} \int_0^\infty \left[ R_{n/k} \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1 \right) - R \left( \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1 \right) \right] dx.
\end{aligned}$$

Using the homogeneity of the function  $R$  on the first term in the right-hand side display, i.e.,  $R(ta, tb) = tR(a, b)$ , we have

$$\begin{aligned}
Q_{5,n} &= \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{\frac{1}{\gamma_Y}} \\
&\quad \times \sqrt{k} \int_0^\infty \left\{ R\left(\left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}, \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}}\right) \right. \\
&\quad \quad \left. - R\left(x^{-\frac{1}{\gamma_X}}, \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}}\right)\right\} dx \\
&+ \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{\frac{1}{\gamma_Y}} \sqrt{k} \int_0^\infty \left[ R\left(x^{-\frac{1}{\gamma_X}}, \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}}\right) - R\left(x^{-\frac{1}{\gamma_X}}, 1\right) \right] dx \\
&+ \sqrt{k} \left[ \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{\frac{1}{\gamma_Y}} - 1 \right] \int_0^\infty R\left(x^{-\frac{1}{\gamma_X}}, 1\right) dx \\
&+ \sqrt{k} \int_0^\infty \left[ R_{n/k}\left(\frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))}\right) \right. \\
&\quad \left. - R\left(\frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))}\right) \right] dx \\
&- \sqrt{k} \int_0^\infty \left[ R\left(\frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1\right) - R\left(x^{-\frac{1}{\gamma_X}}, 1\right) \right] dx \\
&- \sqrt{k} \int_0^\infty \left[ R_{n/k}\left(\frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1\right) - R\left(\frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))}, 1\right) \right] dx \Big\} \\
&=: \sum_{i=1}^6 Q_{5,n}^{(i)}.
\end{aligned}$$

**Term  $Q_{5,n}^{(3)}$ .** Under Assumption  $(\mathcal{D}_Y)$ , we have

$$\begin{aligned}
&\sqrt{k} \left[ \frac{\bar{F}_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\bar{F}_Y(U_Y(n/k))} \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{\frac{1}{\gamma_Y}} - 1 \right] \\
&= \frac{\sqrt{k} \delta_Y(U_Y(n/k))}{\gamma_Y + \delta_Y(U_Y(n/k))} \left[ \frac{\delta_Y\left(\frac{Y_{(n-k)}}{U_Y(n/k)}U_Y(n/k)\right)}{\delta_Y(U_Y(n/k))} - 1 \right] = o_{\mathbb{P}}(1),
\end{aligned}$$

by the uniformity of the convergence in regular variation and the fact that  $\sqrt{k}\delta_Y(U_Y(n/k)) \rightarrow 0$  as  $n \rightarrow 0$ . Thus  $Q_{5,n}^{(3)} = o_{\mathbb{P}}(1)$ .

**Terms  $Q_{5,n}^{(1)}$  and  $Q_{5,n}^{(5)}$ .** The aim is to show that

$$\sup_{s \in [1-\varepsilon, 1+\varepsilon]} \left| \sqrt{k} \int_0^\infty \left[ R\left(s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(sU_Y(n/k))}, s^{-\frac{1}{\gamma_Y}}\right) - R\left(x^{-\frac{1}{\gamma_X}}, s^{-\frac{1}{\gamma_Y}}\right) \right] dx \right| = o(1). \quad (11)$$

If this convergence is true, clearly  $Q_{5,n}^{(1)} = o_{\mathbb{P}}(1)$  and  $Q_{5,n}^{(5)} = o_{\mathbb{P}}(1)$ .

To prove this convergence, we use the Lipschitz property of  $R$ , i.e., for  $T_n = \left(\frac{n}{k}\right)^b$ ,  $0 < b < 1$ ,

$$\begin{aligned}
& \left| \sqrt{k} \int_0^\infty \left[ R \left( s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(x U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))}, s^{-\frac{1}{\gamma_Y}} \right) - R \left( x^{-\frac{1}{\gamma_X}}, s^{-\frac{1}{\gamma_Y}} \right) \right] dx \right| \\
&= \left| \sqrt{k} \int_0^\infty \left[ R \left( s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))}, s^{-\frac{1}{\gamma_Y}} \right) - R \left( z, s^{-\frac{1}{\gamma_Y}} \right) \right] dz^{-\gamma_X} \right| \\
&\leq -\sqrt{k} \int_0^{T_n} \left| R \left( s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))}, s^{-\frac{1}{\gamma_Y}} \right) - R \left( z, s^{-\frac{1}{\gamma_Y}} \right) \right| dz^{-\gamma_X} \\
&\quad -\sqrt{k} \int_{T_n}^\infty \left| R \left( s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))}, s^{-\frac{1}{\gamma_Y}} \right) - R \left( z, s^{-\frac{1}{\gamma_Y}} \right) \right| dz^{-\gamma_X} \\
&\leq -\sqrt{k} \int_0^{T_n} \left| s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))} - z \right| dz^{-\gamma_X} \\
&\quad + 2\sqrt{k} T_n^{-\gamma_X} \sup_{\substack{x>0 \\ y \in [(1+\varepsilon)^{-1/\gamma_Y}, (1-\varepsilon)^{-1/\gamma_Y}]}} R(x, y).
\end{aligned}$$

From Assumptions  $(\mathcal{D}_X)$  and  $(\mathcal{D}_Y)$  and Proposition B.1.10 in de Haan and Ferreira (2006), we obtain for  $\xi, z > 0$  and  $n$  large

$$\begin{aligned}
& \left| s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))} - z \right| \\
&= \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k)) s^{1/\gamma_Y}} \left| \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} - z \frac{\bar{F}_Y(s U_Y(n/k)) s^{1/\gamma_Y}}{\bar{F}_Y(U_Y(n/k))} \right| \\
&\leq \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k)) s^{1/\gamma_Y}} \left\{ \left| \frac{\bar{F}_X(z^{-\gamma_X} U_X(n/k))}{\bar{F}_X(U_X(n/k))} - z \right| + z \left| \frac{\bar{F}_Y(s U_Y(n/k)) s^{1/\gamma_Y}}{\bar{F}_Y(U_Y(n/k))} - 1 \right| \right\} \\
&\leq \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k)) s^{1/\gamma_Y}} \left\{ C \left| \delta_X \left( U_X \left( \frac{n}{k} \right) \right) \right| z \left[ (z^{-\gamma_X})^{-\beta_X \pm \xi} + z^{\gamma_X \beta_X} + 1 \right] + C \left| \delta_Y \left( U_Y \left( \frac{n}{k} \right) \right) \right| z \right\},
\end{aligned}$$

where  $C$  does not depend on  $s$ .

This implies that, for  $\xi \in (0, \beta_X + (1 - \gamma_X)/\gamma_X)$  and  $n$  large

$$\begin{aligned}
& \sup_{s \in [1-\varepsilon, 1+\varepsilon]} \left| \sqrt{k} \int_0^\infty \left[ R \left( s^{-\frac{1}{\gamma_Y}} \frac{\bar{F}_X(x U_X(n/k))}{\bar{F}_X(U_X(n/k))} \frac{\bar{F}_Y(U_Y(n/k))}{\bar{F}_Y(s U_Y(n/k))}, s^{-\frac{1}{\gamma_Y}} \right) - R \left( x^{-\frac{1}{\gamma_X}}, s^{-\frac{1}{\gamma_Y}} \right) \right] dx \right| \\
&\leq C \sqrt{k} \left| \delta_X \left( U_X \left( \frac{n}{k} \right) \right) \right| T_n^{1-\gamma_X + \gamma_X \beta_X + \gamma_X \xi} + C \sqrt{k} \left| \delta_Y \left( U_Y \left( \frac{n}{k} \right) \right) \right| T_n^{1-\gamma_X} \\
&\quad + 2\sqrt{k} T_n^{-\gamma_X} \sup_{\substack{x>0 \\ y \in [(1+\varepsilon)^{-1/\gamma_Y}, (1-\varepsilon)^{-1/\gamma_Y}]}} R(x, y) \\
&= o(1),
\end{aligned}$$

if and only if

$$\begin{cases} \sqrt{k} T_n^{-\gamma_X} & \longrightarrow 0 \\ \sqrt{k} |\delta_X(U_X(\frac{n}{k}))| T_n^{1-\gamma_X+\gamma_X\beta_X+\gamma_X\xi} & \longrightarrow 0 \\ \sqrt{k} |\delta_Y(U_Y(\frac{n}{k}))| T_n^{1-\gamma_X} & \longrightarrow 0. \end{cases} \quad (12)$$

If we choose  $k = \lfloor n^a \rfloor$ , with  $a \in (0, 1)$ , then (12) is satisfied by choosing

$$\frac{a}{2\gamma_X(1-a)} < b < \min\left(\frac{2\gamma_X(\beta_X - \varepsilon_X)(1-a) - a}{2(1-a)(1-\gamma_X + \gamma_X\beta_X + \xi\gamma_X)}, \frac{2\gamma_Y(\beta_Y - \varepsilon_Y)(1-a) - a}{2(1-a)(1-\gamma_X)}\right),$$

with  $\varepsilon_X \in (0, \beta_X)$  and  $\varepsilon_Y \in (0, \beta_Y)$ . As the lower bound for  $b$  must be below the upper bounds, we get the following restriction on  $a$ :

$$a < \min\left(\frac{2\gamma_X^2(\beta_X - \varepsilon_X)}{2\gamma_X^2(\beta_X - \varepsilon_X) + 1 + \gamma_X\beta_X + \xi\gamma_X}, \frac{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y)}{2\gamma_X\gamma_Y(\beta_Y - \varepsilon_Y) + 1}\right).$$

**Term  $Q_{5,n}^{(2)}$ .** Using the homogeneity of  $R$ , we have

$$\begin{aligned} & \sqrt{k} \int_0^\infty \left[ R\left(x^{-\frac{1}{\gamma_X}}, \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}}\right) - R\left(x^{-\frac{1}{\gamma_X}}, 1\right) \right] dx \\ &= \sqrt{k} \int_0^\infty \left[ \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1}{\gamma_Y}} R\left(\left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{\frac{1}{\gamma_Y}} x^{-\frac{1}{\gamma_X}}, 1\right) - R\left(x^{-\frac{1}{\gamma_X}}, 1\right) \right] dx \\ &= \sqrt{k} \left( \left(\frac{Y_{(n-k)}}{U_Y(n/k)}\right)^{-\frac{1-\gamma_X}{\gamma_Y}} - 1 \right) \int_0^\infty R\left(x^{-\frac{1}{\gamma_X}}, 1\right) dx \\ &= -\frac{1-\gamma_X}{\gamma_Y} \sqrt{k} \left(\frac{Y_{(n-k)}}{U_Y(n/k)} - 1\right) \int_0^\infty R\left(x^{-\frac{1}{\gamma_X}}, 1\right) dx + o_{\mathbb{P}}(1), \end{aligned}$$

from which we deduce that

$$Q_{5,n}^{(2)} = -(1-\gamma_X) \mathcal{Z}(\phi_{D_Y,1}) \int_0^\infty R\left(x^{-\frac{1}{\gamma_X}}, 1\right) dx + o_{\mathbb{P}}(1),$$

by Corollary 2.1.

**Terms  $Q_{5,n}^{(4)}$  and  $Q_{5,n}^{(6)}$ .** We use Assumption (S) and obtain for  $n$  large and with arbitrary large probability

$$\begin{aligned} |Q_{5,n}^{(4)}| &\leq \sqrt{k} \sup_{\substack{x_h > 0 \\ y_0 \in [\frac{1}{2}, 2]}} \frac{|R_{n/k}(x_h, y_0) - R(x_h, y_0)|}{x_h^\kappa \wedge 1} \int_0^\infty \left\{ \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))} \right\}^\kappa \wedge 1 dx \\ &= O\left(\sqrt{k} \left(\frac{n}{k}\right)^\tau \left(1 + \int_1^\infty \left\{ \frac{\bar{F}_X(xU_X(n/k))}{\bar{F}_X(U_X(n/k))} \right\}^\kappa dx\right)\right) \\ &= O\left(\sqrt{k} \left(\frac{n}{k}\right)^\tau\right), \end{aligned}$$

since by Proposition B.1.9.5 in de Haan and Ferreira (2006) the integral is finite. A similar argument yields  $Q_{5,n}^{(6)} = O\left(\sqrt{k} \left(\frac{n}{k}\right)^\tau\right)$ . This condition implies a new restriction on  $a$ , that is  $a < -\frac{2\tau}{1-2\tau}$ . Overall, we have

$$Q_{5,n} = -(1 - \gamma_X) \mathcal{Z}(\phi_{D_Y,1}) \int_0^\infty R\left(x^{-\frac{1}{\gamma_X}}, 1\right) dx + o_{\mathbb{P}}(1),$$

from which Theorem 2.2 follows.  $\square$

### 6.3 Proof of Theorem 2.3

We have the decomposition

$$\begin{aligned} \frac{\hat{\theta}_{h,p}}{\theta_{h,p}} - 1 &= \left(\frac{k}{np}\right)^{\hat{\gamma}_{X,k}} \frac{\bar{\theta}_{h,k/n}}{\theta_{h,p}} - 1 \\ &= \underbrace{\left\{ \left(\frac{k}{np}\right)^{\hat{\gamma}_{X,k} - \gamma_X} \right\}}_{Q_{6,n}} \underbrace{\left\{ \frac{\bar{\theta}_{h,k/n}}{\theta_{h,k/n}} \right\}}_{Q_{7,n}} \underbrace{\left\{ \frac{\left(\frac{k}{np}\right)^{\gamma_X} \theta_{h,k/n}}{\theta_{h,p}} \right\}}_{Q_{8,n}} - 1 \\ &= \left\{ \left(\frac{k}{np}\right)^{\hat{\gamma}_{X,k} - \gamma_X} - 1 \right\} Q_{7,n} Q_{8,n} + \left\{ \frac{\bar{\theta}_{h,k/n}}{\theta_{h,k/n}} - 1 \right\} Q_{8,n} + \left\{ \frac{\left(\frac{k}{np}\right)^{\gamma_X} \theta_{h,k/n}}{\theta_{h,p}} - 1 \right\}. \end{aligned}$$

Since  $\frac{\log k/(np)}{\sqrt{k}} \rightarrow 0$ , this implies that

$$\begin{aligned} \frac{\sqrt{k}}{\log \frac{k}{np}} \left( \frac{\hat{\theta}_{h,p}}{\theta_{h,p}} - 1 \right) &= \sqrt{k} (\hat{\gamma}_{X,k} - \gamma_X) (1 + o_{\mathbb{P}}(1)) Q_{7,n} Q_{8,n} \\ &\quad + \sqrt{k} \left\{ \frac{\bar{\theta}_{h,k/n}}{\theta_{h,k/n}} - 1 \right\} \frac{Q_{8,n}}{\log \frac{k}{np}} \\ &\quad + \sqrt{k} \left\{ \frac{\theta_{h,k/n}/U_X(n/k)}{\theta_{h,p}/U_X(1/p)} - 1 \right\} \frac{1}{\log \frac{k}{np}} \frac{U_X(n/k)}{U_X(1/p)} \left(\frac{k}{np}\right)^{\gamma_X} \\ &\quad + \sqrt{k} \left\{ \frac{U_X(n/k)}{U_X(1/p)} \left(\frac{k}{np}\right)^{\gamma_X} - 1 \right\} \frac{1}{\log \frac{k}{np}}. \end{aligned}$$

Under Assumption  $(\mathcal{D}_X)$ , we have

$$\sqrt{k} \left\{ \frac{U_X(n/k)}{U_X(1/p)} \left(\frac{k}{np}\right)^{\gamma_X} - 1 \right\} \frac{1}{\log \frac{k}{np}} = \frac{1}{1 + a_X(1/p)} \left\{ \sqrt{k} a_X \left(\frac{n}{k}\right) - \sqrt{k} a_X \left(\frac{1}{p}\right) \right\} \frac{1}{\log \frac{k}{np}} = o(1),$$

since  $a_X$  is regularly varying with index  $-\beta_X \gamma_X$ ,  $\sqrt{k} \delta_X(U_X(n/k)) \rightarrow 0$  and  $\frac{k}{np} \rightarrow \infty$ .

Now, we have

$$\begin{aligned}
\frac{\theta_{h,k/n}}{U_X(n/k)} &= \int_0^\infty R_{n/k} \left( \frac{n}{k} \bar{F}_X(z U_X(n/k)), 1 \right) dz \\
&= \int_0^\infty R \left( z^{-1/\gamma_X}, 1 \right) dz \\
&\quad + \int_0^\infty \left[ R_{n/k} \left( \frac{n}{k} \bar{F}_X(z U_X(n/k)), 1 \right) - R \left( \frac{n}{k} \bar{F}_X(z U_X(n/k)), 1 \right) \right] dz \\
&\quad + \int_0^\infty \left[ R \left( \frac{n}{k} \bar{F}_X(z U_X(n/k)), 1 \right) - R(z^{-1/\gamma_X}, 1) \right] dz \\
&=: \int_0^\infty R \left( z^{-1/\gamma_X}, 1 \right) dz + Q_{9,n} + Q_{10,n}.
\end{aligned}$$

Similarly as for  $Q_{5,n}^{(6)}$ , we have

$$|Q_{9,n}| = O \left( \left( \frac{n}{k} \right)^\tau \right),$$

and, using (11),  $Q_{10,n} = o(1/\sqrt{k})$ . Thus, under our assumptions

$$\frac{\theta_{h,k/n}}{U_X(n/k)} = \int_0^\infty R \left( z^{-1/\gamma_X}, 1 \right) dz + o \left( \frac{1}{\sqrt{k}} \right),$$

which can also be shown to hold for  $\theta_{h,p}/U_X(1/p)$ . This yields

$$\sqrt{k} \left\{ \frac{\theta_{h,k/n}/U_X(n/k)}{\theta_{h,p}/U_X(1/p)} - 1 \right\} = o(1).$$

Finally, Theorem 2.2 and Corollary 2.2 achieve the proof of Theorem 2.3.  $\square$

## 6.4 Proof of Theorem 3.1

For the first step, we need to check the assumptions of Theorem 2.3. First we need to show that Assumption  $(\tilde{R})$  is satisfied for  $(X_h^+, Y_0^+, X_{i+h}^+, Y_i^+)$ . This is the case since for  $r$  large

$$\begin{aligned}
&r \mathbb{P} \left( \bar{F}_{X^+}(X_h^+) \leq \frac{x_h}{r}, \bar{F}_{Y^+}(Y_0^+) \leq \frac{y_0}{r}, \bar{F}_{X^+}(X_{i+h}^+) \leq \frac{x_{i+h}}{r}, \bar{F}_{Y^+}(Y_i^+) \leq \frac{y_i}{r} \right) \\
&= r \mathbb{P} \left( X_h^+ \geq U_{X^+} \left( \frac{r}{x_h} \right), Y_0^+ \geq U_{Y^+} \left( \frac{r}{y_0} \right), X_{i+h}^+ \geq U_{X^+} \left( \frac{r}{x_{i+h}} \right), Y_i^+ \geq U_{Y^+} \left( \frac{r}{y_i} \right) \right) \\
&= r \mathbb{P} \left( X_h \geq U_X \left( \frac{r}{x_h} \right), Y_0 \geq U_Y \left( \frac{r}{y_0} \right), X_{i+h} \geq U_X \left( \frac{r}{x_{i+h}} \right), Y_i \geq U_Y \left( \frac{r}{y_i} \right) \right) \\
&= r \mathbb{P} \left( \bar{F}_X(X_h) \leq \frac{x_h}{r}, \bar{F}_Y(Y_0) \leq \frac{y_0}{r}, \bar{F}_X(X_{i+h}) \leq \frac{x_{i+h}}{r}, \bar{F}_Y(Y_i) \leq \frac{y_i}{r} \right) \\
&\rightarrow \tilde{R}_i(x_h, y_0, x_{i+h}, y_i).
\end{aligned}$$

Now, we have to show that Assumption  $(\mathcal{S})$  is also satisfied for  $(X_h^+, Y_0^+)$ . Again this is the case since for  $0 < x_h < r\{1 - F_X(0)\}$ , we have  $\{1 - F_{X^+}(X_h^+) \leq \frac{x_h}{r}\} = \{1 - F_X(X_h^+) \leq \frac{x_h}{r}\} =$

$\{1 - F_X(X_h) \leq \frac{x_h}{r}\}$ , and otherwise, if  $x_h \geq r\{1 - F_X(0)\}$ , clearly  $1 - F_{X^+}(X_h^+) \leq \frac{x_h}{r}$ . All the assumptions of Theorem 2.3 are satisfied, thus step 1 is established.

For the second step, remark that for  $p$  small

$$\begin{aligned} \frac{\theta_{h,p}}{\theta_{h,p}^+} &= \frac{\mathbb{E}[X_h | Y_0 > U_Y(1/p)]}{\mathbb{E}[X_h^+ | Y_0^+ > U_{Y^+}(1/p)]} \\ &= \frac{\mathbb{E}[X_h | Y_0 > U_Y(1/p)]}{\mathbb{E}[X_h^+ | Y_0 > U_Y(1/p)]} \\ &= 1 - \frac{\mathbb{E}[X_h^- | Y_0 > U_Y(1/p)]}{\mathbb{E}[X_h^+ | Y_0 > U_Y(1/p)]}. \end{aligned}$$

Now, remark that, using Hölder's inequality

$$\begin{aligned} \mathbb{E}[X_h^- | Y_0 > U_Y(1/p)] &= \frac{1}{p} \mathbb{E}[X_h^- \mathbb{1}_{\{Y_0 > U_Y(1/p)\}}] \\ &= \frac{1}{p} \mathbb{E}[X_h^- \mathbb{1}_{\{X_h < 0, Y_0 > U_Y(1/p)\}}] \\ &\leq \frac{1}{p} \left\{ \mathbb{E} \left[ (X_h^-)^{1/\gamma_X} \right] \right\}^{\gamma_X} [\mathbb{P}(X_h < 0, Y_0 > U_Y(1/p))]^{1-\gamma_X} \\ &\leq C \frac{1}{p} [\mathbb{P}(X_h < 0, Y_0 > U_Y(1/p))]^{1-\gamma_X}, \end{aligned}$$

by (8).

Also, we have, for  $p$  small,

$$\begin{aligned} \frac{1}{p} \mathbb{P}(X_h < 0, Y_0 > U_Y(1/p)) &= 1 - \frac{1}{p} \mathbb{P}(X_h \geq 0, Y_0 > U_Y(1/p)) \\ &= 1 - \frac{1}{p} \mathbb{P}(\bar{F}_X(X_h) \leq \bar{F}_X(0), \bar{F}_Y(Y_0) < p) \\ &= 1 - R_{1/p} \left( \frac{\bar{F}_X(0)}{p}, 1 \right) \\ &\leq \left| 1 - R \left( \frac{\bar{F}_X(0)}{p}, 1 \right) \right| \\ &\quad + \left| R_{1/p} \left( \frac{\bar{F}_X(0)}{p}, 1 \right) - R \left( \frac{\bar{F}_X(0)}{p}, 1 \right) \right| \\ &= O(p^{-\tau}), \end{aligned}$$

by Assumption  $(\mathcal{S})$  since  $R_{\bar{F}_X(0)/p} \left( \frac{\bar{F}_X(0)}{p}, 1 \right) = 1$ . This yields

$$\mathbb{P}(X_h < 0, Y_0 > U_Y(1/p)) = O(p^{1-\tau}),$$

and hence

$$\mathbb{E}[X_h^- | Y_0 > U_Y(1/p)] = O\left(p^{-1+(1-\tau)(1-\gamma_X)}\right).$$

Finally, using Lemma 2.1, we have for  $p$  small

$$\frac{\theta_{h,p}}{\theta_{h,p}^+} = 1 + \frac{O(p^{-1+(1-\tau)(1-\gamma_X)})}{U_{X^+}(1/p)} = 1 + O(p^{-\tau(1-\gamma_X)}) = 1 + O\left(\frac{1}{\sqrt{k}}\right),$$

from which we deduce that

$$\frac{\sqrt{k}}{\log \frac{k}{np}} \left( \frac{\theta_{h,p}^+}{\theta_{h,p}} - 1 \right) = o(1).$$

Combining the two steps, Theorem 3.1 follows.  $\square$

## 6.5 Proof of Theorem 3.2

For simplicity of the arguments, we assume that  $n = m_n r_n$ .

First we consider the multiplier block bootstrap version of the process  $\mathcal{Z}_{n,\varepsilon}$ , defined as

$$\mathcal{Z}_{n,\varepsilon,\xi}^*(\Psi) = \sqrt{\frac{1}{n v_n}} \sum_{j=1}^{m_n} \xi_j \sum_{i \in I_j} \{\Psi(Z_{n,i}) - \mathbb{E}\Psi(Z_{n,i})\}.$$

Let  $\mathbb{E}_\xi$  denote the expectation with respect to  $\xi := (\xi_j)_{j \in \mathbb{N}}$ , i.e., the conditional expectation given  $(Z_{n,i})_{1 \leq i \leq n}$ . Under the assumptions of Theorem 3.2 we have that, for any  $\varepsilon \in (0, 1)$

$$\mathcal{Z}_{n,\varepsilon,\xi}^*(\Psi) \rightsquigarrow \mathcal{Z}_\varepsilon(\Psi) \tag{13}$$

in  $\ell^\infty(\Phi)$  and

$$\sup_{g \in BL_1(\ell^\infty(\Phi))} |\mathbb{E}_\xi g(\mathcal{Z}_{n,\varepsilon,\xi}^*) - \mathbb{E}g(\mathcal{Z}_\varepsilon)| \longrightarrow 0, \tag{14}$$

in probability, where

$$BL_1(\ell^\infty(\Phi)) := \left\{ g : \ell^\infty(\Phi) \rightarrow \mathbb{R} : \|g\|_\infty := \sup_{z \in \ell^\infty(\Phi)} |g(z)| \leq 1, \right. \\ \left. |g(z_1) - g(z_2)| \leq \sup_{\psi \in \Phi} |z_1(\psi) - z_2(\psi)| \text{ for all } z_1, z_2 \in \ell^\infty(\Phi) \right\}.$$

The first statement can be obtained in similar way as the proof of Proposition 2.5 in Drees (2015), and the second one from the proof of Theorem 2.3 in Drees (2015).

Now let

$$S_{Y^+,n} := \frac{Y_{(n-k)}^+}{U_{Y^+}(\frac{n}{k})}, \\ \mathbb{E}_{D_{Y^+,n}}(s) := \mathbb{E} \left[ \mathbb{1}_{\{Y_0^+ > s U_{Y^+}(\frac{n}{k})\}} \left| X_h^+ > (1-\varepsilon) U_{X^+}(\frac{n}{k}) \text{ or } Y_0^+ > (1-\varepsilon) U_{Y^+}(\frac{n}{k}) \right. \right], \\ \mathbb{E}_n(s) := \mathbb{E} \left[ \frac{X_h^+}{U_{X^+}(\frac{n}{k})} \mathbb{1}_{\{Y_0^+ > s U_{Y^+}(\frac{n}{k})\}} \left| X_h^+ > (1-\varepsilon) U_{X^+}(\frac{n}{k}) \text{ or } Y_0^+ > (1-\varepsilon) U_{Y^+}(\frac{n}{k}) \right. \right].$$

We start to study the bootstrap version of our estimator of the parameter of interest at an intermediate level, namely  $\bar{\theta}_{h,k/n}^+$ , where the latter notation is used for  $\bar{\theta}_{h,k/n}$  applied to  $((X_t^+, Y_t^+))_{t \in \mathbb{Z}}$ . Note that

$$\begin{aligned} \sqrt{k} \left( \frac{\bar{\theta}_{h,k/n}^{+,*}}{\bar{\theta}_{h,k/n}^+} - 1 \right) &= \frac{U_{X^+(\frac{n}{k})}}{\bar{\theta}_{h,k/n}^+} \sqrt{k} \left( \frac{\bar{\theta}_{h,k/n}^{+,*}}{U_{X^+(\frac{n}{k})}} - \frac{\bar{\theta}_{h,k/n}^+}{U_{X^+(\frac{n}{k})}} \right) \\ &= \frac{U_{X^+(\frac{n}{k})}}{\bar{\theta}_{h,k/n}^+} \frac{\frac{1}{\sqrt{k}} \left\{ \sum_{j=1}^{m_n} \xi_j \sum_{i \in I_j} \frac{X_{i+h}^+}{U_{X^+(\frac{n}{k})}} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}} - \frac{\bar{\theta}_{h,k/n}^+}{U_{X^+(\frac{n}{k})}} \sum_{j=1}^{m_n} \xi_j \sum_{i \in I_j} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}} \right\}}{\frac{1}{k} \sum_{j=1}^{m_n} (1 + \xi_j) \sum_{i \in I_j} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}}} \\ &=: \frac{U_{X^+(\frac{n}{k})}}{\bar{\theta}_{h,k/n}^+} \frac{N}{D}. \end{aligned}$$

Using Theorem 2.2 and Lemma 2.1, we have

$$\frac{\bar{\theta}_{h,k/n}^+}{U_{X^+(\frac{n}{k})}} \xrightarrow{\mathbb{P}} \int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz.$$

Also, since  $m_n^{-1/2} \sum_{j=1}^{m_n} \xi_j = O_{\mathbb{P}}(1)$ , we have, from the convergences

$$\begin{aligned} S_{Y^+,n} &\xrightarrow{\mathbb{P}} 1 \\ \mathbb{E}_{D_{Y^+,n}}(S_{Y^+,n}) &\xrightarrow{\mathbb{P}} \frac{1}{L \left( (1-\varepsilon)^{-\frac{1}{\gamma_X}}, (1-\varepsilon)^{-\frac{1}{\gamma_Y}} \right)} \\ \mathbb{E}_n(S_{Y^+,n}) &\xrightarrow{\mathbb{P}} \frac{1}{L \left( (1-\varepsilon)^{-\frac{1}{\gamma_X}}, (1-\varepsilon)^{-\frac{1}{\gamma_Y}} \right)} \int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz, \end{aligned}$$

and (13)

$$\begin{aligned} N &= \sqrt{\frac{n v_n}{k}} \left\{ \frac{1}{\sqrt{n v_n}} \sum_{j=1}^{m_n} \xi_j \sum_{i \in I_j} \frac{X_{i+h}^+}{U_{X^+(\frac{n}{k})}} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}} - \frac{\bar{\theta}_{h,k/n}^+}{U_{X^+(\frac{n}{k})}} \frac{1}{\sqrt{n v_n}} \sum_{j=1}^{m_n} \xi_j \sum_{i \in I_j} \mathbb{1}_{\{Y_i^+ > Y_{(n-k)}^+\}} \right\} \\ &= \sqrt{L \left( (1-\varepsilon)^{-\frac{1}{\gamma_X}}, (1-\varepsilon)^{-\frac{1}{\gamma_Y}} \right)} \left\{ \mathcal{Z}_{n,\varepsilon,\xi}^*(\phi_{S_{Y^+,n}}) - \mathcal{Z}_{n,\varepsilon,\xi}^*(\phi_{D_{Y^+,n}}) \int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz \right\} + o_{\mathbb{P}}(1) \\ &\rightsquigarrow \mathcal{Z}(\phi_1) - \mathcal{Z}(\phi_{D_{Y^+,1}}) \int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz, \end{aligned}$$

and

$$D = \frac{n v_n}{k} \mathbb{E}_{D_{Y^+,n}}(S_{Y^+,n}) \left\{ 1 + \frac{1}{m_n} \sum_{j=1}^{m_n} \xi_j \right\} + o_{\mathbb{P}}(1) \xrightarrow{\mathbb{P}} 1.$$

Consequently

$$\sqrt{k} \left( \frac{\bar{\theta}_{h,k/n}^{+,*}}{\bar{\theta}_{h,k/n}^+} - 1 \right) \rightsquigarrow \frac{\mathcal{Z}(\phi_1)}{\int_0^\infty R \left( z^{-\frac{1}{\gamma_X}}, 1 \right) dz} - \mathcal{Z}(\phi_{D_{Y^+,1}}). \quad (15)$$

As already shown in the proof of Theorem 4.2 in Goegebeur et al. (2025), we have

$$\sqrt{k} \left( \hat{\gamma}_{X^+,k}^* - \hat{\gamma}_{X^+,k} \right) \rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_{X^+},1}). \quad (16)$$

Finally, using (15) and (16), we deduce that

$$\begin{aligned} V_n^* &:= \frac{\sqrt{k}}{\log k/(np)} \left( \frac{\hat{\theta}_{h,p}^*}{\hat{\theta}_{h,p}^+} - 1 \right) = \frac{\sqrt{k}}{\log k/(np)} \left( \left( \frac{k}{np} \right)^{\hat{\gamma}_{X^+,k}^* - \hat{\gamma}_{X^+,k}} \frac{\bar{\theta}_{h,k/n}^{+,*}}{\bar{\theta}_{h,k/n}^+} - 1 \right) \\ &= \frac{\sqrt{k}}{\log k/(np)} \left\{ \left( \frac{k}{np} \right)^{\hat{\gamma}_{X^+,k}^* - \hat{\gamma}_{X^+,k}} - 1 \right\} \frac{\bar{\theta}_{h,k/n}^{+,*}}{\bar{\theta}_{h,k/n}^+} + \frac{\sqrt{k}}{\log k/(np)} \left( \frac{\bar{\theta}_{h,k/n}^{+,*}}{\bar{\theta}_{h,k/n}^+} - 1 \right) \\ &= \sqrt{k} \left( \hat{\gamma}_{X^+,k}^* - \hat{\gamma}_{X^+,k} \right) + o_{\mathbb{P}}(1) \\ &\rightsquigarrow \mathcal{Z}(\phi_{H,1}) - \gamma_X \mathcal{Z}(\phi_{D_{X^+},1}) =: V. \end{aligned}$$

By (14) we get then also

$$\sup_{g \in BL_1(\mathbb{R})} |\mathbb{E}_\xi(g(V_n^*)) - \mathbb{E}(g(V))| \xrightarrow{\mathbb{P}} 0,$$

where  $BL_1(\mathbb{R}) := \{g : \mathbb{R} \rightarrow \mathbb{R} : \sup_{z \in \mathbb{R}} |g(z)| \leq 1, |g(z_1) - g(z_2)| \leq |z_1 - z_2| \text{ for all } z_1, z_2 \in \mathbb{R}\}$ , from which Theorem 3.2 follows.  $\square$

## Acknowledgement

The authors sincerely thank the referees, the Associate Editor and the Editor for their helpful comments on our paper. The research of Armelle Guillou was supported by the CNRS grant International Emerging Action IEA-111. Computation/simulation for the work described in this paper was supported by the DeIC National HPC Centre, SDU.

## References

- [1] Acharya, V., Pedersen, L., Philippon, T. and Richardson, M. (2010). Measuring systemic risk, FRB of Cleveland Working Paper No. 10-02, <https://ssrn.com/abstract=1595075>.
- [2] Billingsley, P. (1995). Probability and measure, third edition, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York.
- [3] Cai, J.J., Einmahl, J.H.J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme, *Journal of the Royal Statistical Society Series B*, 77, 417–442.
- [4] Davis, R.A., Drees, H., Segers, J. and Warchol, M. (2018). Inference on the tail process with application to financial time series modeling, *Journal of Econometrics*, 205, 508–525.

- [5] Davison, A.C., Padoan, S.A. and Stupfler, G. (2023). Tail risk inference via expectiles in heavy-tailed time series, *Journal of Business & Economic Statistics*, 41, 876–889.
- [6] Di Bernardino, E. and Prieur, C. (2018). Estimation of the multivariate conditional tail expectation for extreme risk levels: Illustration on environmental data sets, *Environmetrics*, 29, 1–22.
- [7] Drees, H. (2015). Bootstrapping empirical processes of cluster functionals with application to extremograms, <https://arxiv.org/abs/1511.00420>.
- [8] Drees, H. and Huang, X. (1998). Best attainable rates of convergence for estimators of the stable tail dependence function, *Journal of Multivariate Analysis*, 64, 25–47.
- [9] Drees, H. and Knežević, M. (2020). Peak-over-threshold estimators for spectral tail processes: random vs deterministic thresholds, *Extremes*, 23, 465–491.
- [10] Drees, H. and Rootzén, H. (2010). Limit theorems for empirical processes of cluster functionals, *The Annals of Statistics*, 38, 4, 2145–2186.
- [11] Goegebeur, Y., Guillou, A., Ho, N.K.L. and Qin, J. (2023). A Weissman-type estimator of the conditional marginal expected shortfall, *Econometrics and Statistics*, 27, 173–196.
- [12] Goegebeur, Y., Guillou, A. and Qin, J. (2025). A conditional tail expectation type risk measure for time series, *Journal of Time Series Analysis*, <https://doi.org/10.1111/jtsa.12853>.
- [13] de Haan, L. and Ferreira, A. (2006). *Extreme value theory. An introduction*, Springer.
- [14] Hall, P. (1982). On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society Series B*, 44, 37–42.
- [15] Hall, P. and Welsh, A.H. (1985). Adaptive estimates of parameters of regular variation, *The Annals of Statistics*, 13, 331–341.
- [16] Hill B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 5, 1163–1174.
- [17] Huang, X. (1992). *Statistics of bivariate extremes*, PhD Thesis, Erasmus University Rotterdam, Tinbergen Institute Research series No. 22.
- [18] Janssen, A. and Segers, J. (2014). Markov tail chains, *Journal of Applied probability*, 51, 1133–1153.
- [19] Knežević, M. (2020). *Analyse extremer Abhängigkeiten bei Zeitreihen*, <https://ediss.sub.uni-hamburg.de/bitstream/ediss/8697/1/Dissertation.pdf>.
- [20] Meinguet, T. and Segers, J. (2010). Regularly varying time series in Banach spaces, <http://arxiv.org/abs/1001.3262>.

- [21] Padoan, S.A., Rizzelli, S. and Schiavone, M. (2023). Marginal expected shortfall inference under multivariate regular variation, <https://arxiv.org/abs/2304.07578>.
- [22] Pratt, J.W. (1960). On interchanging limits and integrals, *The Annals of Mathematical Statistics*, 31, 74–77.
- [23] Resnick, S. and Stărică, C. (1997). Asymptotic behavior of Hill’s estimator for autoregressive data, *Communications in Statistics - Stochastic Models*, 13, 703–721.
- [24] Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence, *Scandinavian Journal of Statistics*, 33, 307–335.
- [25] Weissman, I. (1978). Estimation of parameters and large quantiles based on the  $k$  largest observations, *Journal of the American Statistical Association*, 73, 812–815.