



**HAL**  
open science

# Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles

Clément Albert, Anne Dutfoy, Stéphane Girard

► **To cite this version:**

Clément Albert, Anne Dutfoy, Stéphane Girard. Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles. *Extremes*, 2020, 23, pp.349-380. 10.1007/s10687-019-00370-2 . hal-01692544v4

**HAL Id: hal-01692544**

**<https://hal.science/hal-01692544v4>**

Submitted on 14 Feb 2019

**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.

# Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles

Clément Albert<sup>(1)</sup>, Anne Dutfoy<sup>(2)</sup> and Stéphane Girard<sup>(1, \*)</sup>

<sup>(1)</sup> *Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France*

<sup>(2)</sup> *EDF R&D dept. Périclès, 91120 Palaiseau, France*

## Abstract

We investigate the asymptotic behavior of the (relative) extrapolation error associated with some estimators of extreme quantiles based on extreme-value theory. It is shown that the extrapolation error can be interpreted as the remainder of a first order Taylor expansion. Conditions are then provided such that this error tends to zero as the sample size increases. Interestingly, in case of the so-called Exponential Tail estimator, these conditions lead to a subdivision of Gumbel maximum domain of attraction into three subsets. In contrast, the extrapolation error associated with Weissman estimator has a common behavior over the whole Fréchet maximum domain of attraction. First order equivalents of the extrapolation error are then derived showing that Weissman estimator may lead to smaller extrapolation errors than the Exponential Tail estimator on some subsets of Gumbel maximum domain of attraction. The accuracy of the equivalents is illustrated numerically and an application on real data is also provided.

**Keywords:** Extrapolation error, Extreme quantiles, Extreme-value theory.

**AMS 2000 subject classification:** 62G32, 62G20.

## 1 Introduction

The starting point of this work is the study of the asymptotic behavior of the Exponential Tail (ET) estimator, a nonparametric estimator of the extreme quantiles from an unknown distribution. Theoretical developments can be found in [6] while numerical aspects are investigated in [12]. Given a  $n$ -sample  $X_1, \dots, X_n$  from a cumulative distribution function  $F$  with associated survival distribution function  $\bar{F}$ , an extreme quantile is a  $(1 - p_n)$ th quantile  $q(p_n)$  of  $F$  essentially larger than the maximal observation, *i.e.* such that  $\bar{F}(q(p_n)) = p_n$  with  $np_n \rightarrow 0$  as  $n \rightarrow \infty$ . The estimation of extreme quantiles requires specific methods. Among them, the Peaks Over Threshold (POT) method relies on an approximation of the distribution of excesses over a given threshold [25]. More precisely, let  $u_n$  be a deterministic threshold such that  $\bar{F}(u_n) = \alpha_n$

---

\*Corresponding author, [Stephane.Girard@inria.fr](mailto:Stephane.Girard@inria.fr)

or equivalently  $u_n = q(\alpha_n)$  with  $\alpha_n \rightarrow 0$  and  $n\alpha_n \rightarrow \infty$  as  $n \rightarrow \infty$ . The excesses above  $u_n$  are defined as  $Y_i = X_i - u_n$  for all  $X_i > u_n$ . The survival distribution function of an excess is given by  $\bar{F}_{u_n}(x) = \bar{F}(u_n + x)/\bar{F}(u_n)$ . Pickands theorem [24] states that, under mild conditions,  $\bar{F}_{u_n}$  can be approximated by a Generalized Pareto Distribution (GPD). As a consequence, the extreme quantile  $q(p_n)$  can be in turn approximated by the deterministic term

$$\tilde{q}_{\text{GPD}}(p_n; \alpha_n) = q(\alpha_n) + \frac{\sigma_n}{\gamma_n} \left[ \left( \frac{\alpha_n}{p_n} \right)^{\gamma_n} - 1 \right], \quad (1)$$

where  $\sigma_n$  and  $\gamma_n$  are respectively the scale and shape parameters of the GPD distribution. Then, the POT method consists in estimating these two unknown parameters. The ET method corresponds to the important particular case where  $F$  belongs to Gumbel Maximum Domain of Attraction, MDA(Gumbel). In such a situation,  $\gamma_n = 0$  and the GPD distribution reduces to an Exponential distribution with scale parameter  $\sigma_n$ . Thus, approximation (1) can be rewritten as

$$\tilde{q}_{\text{ET}}(p_n; \alpha_n) = q(\alpha_n) + \sigma_n \log(\alpha_n/p_n) \quad (2)$$

and the associated estimator [6] is

$$\hat{q}_{\text{ET}}(p_n; \alpha_n) = \hat{q}(\alpha_n) + \hat{\sigma}_n \log(\alpha_n/p_n)$$

where  $\hat{q}(\alpha_n) = X_{n-k_n+1,n}$  with  $k_n = \lfloor n\alpha_n \rfloor$  and

$$\hat{\sigma}_n = \frac{1}{k_n} \sum_{i=1}^{k_n} X_{n-i+1,n} - X_{n-k_n+1,n}.$$

Let us recall that  $X_{1,n} \leq \dots \leq X_{n,n}$  denote the order statistics associated with  $X_1, \dots, X_n$ . The error  $(q(p_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n))$  can be expanded as a sum of two terms:

$$q(p_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n) = (\tilde{q}_{\text{ET}}(p_n; \alpha_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n)) + (q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n)),$$

the first one being a random estimation error

$$\tilde{q}_{\text{ET}}(p_n; \alpha_n) - \hat{q}_{\text{ET}}(p_n; \alpha_n) = q(\alpha_n) - \hat{q}(\alpha_n) + (\sigma_n - \hat{\sigma}_n) \log(\alpha_n/p_n) \quad (3)$$

and the second one being a deterministic extrapolation error

$$q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n) = q(p_n) - q(\alpha_n) - \sigma_n \log(\alpha_n/p_n). \quad (4)$$

The asymptotic behavior of the estimation error (3) is driven by the asymptotic distributions of  $\hat{q}(\alpha_n)$  and  $\hat{\sigma}_n$  established for instance in [13] or [8, Theorem 2.4.1 and Theorem 3.4.2].

In this paper, we focus on the asymptotic behavior of the extrapolation error (4). Indeed, in view of (2), the ET method extrapolates in the distribution tail from  $q(\alpha_n)$  to  $q(p_n)$  thanks to an additive correction proportional to  $\log(\alpha_n/p_n)$ . Our goal is thus to quantify to what extent this extrapolation can be performed in a consistent way. More specifically, we provide conditions on the pair  $(p_n, \alpha_n)$  such that the relative extrapolation error

$$\varepsilon_{\text{ET}}(p_n; \alpha_n) := (q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n))/q(p_n) \quad (5)$$

tends to zero as  $n \rightarrow \infty$ . These conditions depend on the underlying distribution function  $F$  and they lead to a subdivision of MDA(Gumbel) into three sub-domains depending on the restrictions they impose on the extrapolation range. Related works include [7, 20] who exhibited penultimate approximations for  $F^n$  together with convergence rates for distributions in MDA(Gumbel). These results were extended to other maximum domains of attraction in [21, 22] while penultimate approximations were established for the distribution of the excesses [27]. The relative extrapolation error induced by the approximation of  $\bar{F}_{u_n}$  by the survival distribution function of a GPD is studied in [4].

Here, similarly to [4], we focus on the approximation of quantiles rather than approximations of distribution functions. Let us also highlight that these investigations are not limited to the ET method. To illustrate this, let us introduce  $x(n) = \log(1/\alpha_n)$ ,  $y(n) = \log(1/p_n)$  and  $\varphi(\cdot) = (\bar{F})^{-1}(1/\exp(\cdot))$ . The extrapolation error (4) can thus be interpreted as the remainder of a first order Taylor expansion:

$$q(p_n) - \tilde{q}_{\text{ET}}(p_n; \alpha_n) = \varphi(y(n)) - \varphi(x(n)) - \sigma_n(y(n) - x(n)) \text{ where } \sigma_n = \varphi'(x(n)). \quad (6)$$

We shall show that Weissman estimator [26] dedicated to MDA(Fréchet) can also enter this framework thanks to adapted definitions of functions  $x$ ,  $y$  and  $\varphi$ . In this case, the necessary and sufficient conditions on the extrapolation range are automatically fulfilled for most distributions in MDA(Fréchet) which is a very different situation from MDA(Gumbel). It is also shown that, in some sub-domains of MDA(Gumbel), Weissman approximation is better than (or equivalent to) the ET one even though Weissman estimator was not initially designed for this framework.

The paper is organized as follows: The asymptotic behavior of the remainder associated with the first order Taylor expansion (6) is investigated in Section 2. The applications to ET and Weissman approximations are detailed in Section 3 and Section 4 respectively. Some numerical illustrations are presented in Section 5 and an application to real data is proposed in Section 6. Proofs are postponed to Section 7 and auxiliary results can be found in the Appendix.

## 2 Theoretical framework

The following functions are introduced.

**(A1)**  $x$  and  $y$  are two functions  $\mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $0 < x(t) \leq y(t)$  for  $t$  large enough,  $x(t) \rightarrow \infty$  as  $t \rightarrow \infty$  and  $0 < \liminf_{t \rightarrow \infty} x(t)/y(t) \leq \limsup_{t \rightarrow \infty} x(t)/y(t) \leq 1$ .

**(A2)**  $\varphi$  is a twice differentiable, increasing function.

Motivated by (5) and (6), we introduce

$$\Delta(t) = \frac{\varphi(y(t)) - \varphi(x(t)) - (y(t) - x(t))\varphi'(x(t))}{\varphi(y(t))}, \quad (7)$$

for all  $t > 0$ . The goal of this section is to establish necessary and sufficient conditions on  $\delta(t) := (y(t) - x(t))/y(t)$  so that  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$  in the setting  $0 \leq \liminf \delta(t) \leq \limsup \delta(t) < 1$  of

assumption **(A1)**. The following two functions are of the utmost importance in this study:

$$K_1(s) = \frac{s\varphi'(s)}{\varphi(s)}, \quad K_2(s) = \frac{s^2\varphi''(s)}{\varphi(s)}, \quad s \geq 0.$$

The study of  $\Delta$  relies on the assumption that  $K_1$  is regularly-varying at infinity with index  $\theta_1 \leq 1$ . This property is denoted for short by

**(A3)**  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 \leq 1$

and means that  $K_1$  is ultimately positive such that  $K_1(ts)/K_1(s) \rightarrow t^{\theta_1}$  as  $s \rightarrow \infty$  for all  $t > 0$ . We refer to [5] for a general account on regular variation theory. This assumption is discussed in Section 3 and Section 4 while applying this general framework to the particular cases of ET and Weissman estimators. Finally, a monotonicity assumption is also considered:

**(A4)**  $K_1'$  is ultimately monotone.

Under **(A4)**,  $K_1$  is also ultimately monotone and therefore the limits of  $K_1(s)$  and  $K_2(s)$  when  $s \rightarrow \infty$  exist in  $\bar{\mathbb{R}}$ . The following notations are thus introduced:

$$\lim_{s \rightarrow \infty} K_1(s) = \ell_1 \in \bar{\mathbb{R}}_+ \quad \text{and} \quad \lim_{s \rightarrow \infty} K_2(s) = \ell_2 \in \bar{\mathbb{R}}.$$

We are now in position to state our first main result:

**Proposition 1 (Role of  $\ell_1$  for  $\Delta \rightarrow 0$ )** *Suppose **(A1)**–**(A4)** hold.*

- (i) *If  $\ell_1 \in \{0, 1\}$  then  $\ell_2 = 0$  and  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ .*
- (ii) *If  $\ell_1 \in (0, \infty) \setminus \{1\}$  then  $\ell_2 \in (0, \infty)$  and  $\Delta(t) \rightarrow 0$  if and only if  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ .*
- (iii) *If  $\ell_1 = \infty$  then  $|\ell_2| = \infty$  and  $\Delta(t) \rightarrow 0$  if and only if  $\delta^2(t)K_2(y(t)) \rightarrow 0$  as  $t \rightarrow \infty$ .*

Three cases appear. If  $\ell_1 \in \{0, 1\}$  then  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$  as soon as **(A1)** holds. If  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$  then a necessary and sufficient condition for  $\Delta(t) \rightarrow 0$  is  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ . If  $\ell_1 = \infty$  then the necessary and sufficient condition for  $\Delta(t) \rightarrow 0$  is  $\delta^2(t)K_2(y(t)) \rightarrow 0$  as  $t \rightarrow \infty$ . Clearly, this condition implies  $\delta(t) \rightarrow 0$  since, in this situation,  $|\ell_2| = \infty$ .

Finally, letting  $c(a, b) = \int_0^1 (1 - au)^{b-2} u du$ ,  $a \geq 0$ ,  $b \geq 0$ , first order approximations of  $\Delta$  can be provided in each situation.

**Proposition 2 (First order approximations of  $\Delta$ )** *Suppose **(A1)**–**(A4)** hold.*

- (i) *Assume  $\ell_1 \in \{0, 1\}$  (and thus  $\ell_2 = 0$ ). If  $\ell_1 = 1$ , let us suppose that there exists  $\theta_2 \leq 0$  such that  $|K_2| \in RV_{\theta_2}$ .*

- (a) *If  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ , then*

$$\Delta(t) \sim \frac{1}{2} \delta^2(t) K_2(x(t)) \quad \text{as } t \rightarrow \infty.$$

(b) If  $\delta(t) \rightarrow \delta_\infty \in (0, 1)$  as  $t \rightarrow \infty$ , then

$$\Delta(t) \sim \delta_\infty^2 (1 - \delta_\infty)^{-\theta_2} c(\delta_\infty, \ell_1 + \theta_2) K_2(x(t)) \text{ as } t \rightarrow \infty.$$

(ii) Assume  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$ .

(a) If  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ , then

$$\Delta(t) \sim \frac{\ell_1(\ell_1 - 1)}{2} \delta^2(t) \text{ as } t \rightarrow \infty.$$

(b) If  $\delta(t) \rightarrow \delta_\infty \in (0, 1)$  as  $t \rightarrow \infty$ , then

$$\Delta(t) \rightarrow \delta_\infty^2 c(\delta_\infty, \ell_1) \ell_1 (\ell_1 - 1) \text{ as } t \rightarrow \infty.$$

(iii) Assume  $\ell_1 = \infty$ .

(a) If  $\delta(t) K_1(y(t)) \rightarrow 0$  as  $t \rightarrow \infty$ , then

$$\Delta(t) = \frac{1}{2} \delta^2(t) K_1^2(x(t)) \sim \frac{1}{2} \delta^2(t) K_2(x(t)) \text{ as } t \rightarrow \infty.$$

(b) If  $\delta(t) K_1(y(t)) \rightarrow a \in (0, \infty]$  as  $t \rightarrow \infty$ , then

$$\Delta(t) \rightarrow \int_0^a u \exp(-u) du \text{ as } t \rightarrow \infty.$$

In situation (i) where  $\ell_1 \in \{0, 1\}$ ,  $\Delta \rightarrow 0$  in both cases  $\delta \rightarrow 0$  and  $\delta \rightarrow \delta_\infty \neq 0$ , and the convergence is the fastest in the case  $\delta \rightarrow 0$ . In situation (ii) where  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$ ,  $\Delta$  is asymptotically proportional to  $\delta^2$ . In situation (iii) where  $\ell_1 = \infty$ ,  $\Delta \rightarrow 0$  is the only case where  $\delta K_1(y) \rightarrow 0$  and  $\Delta$  is asymptotically proportional to  $(\delta K_1(x))^2$  or equivalently to  $\delta^2 K_2(x)$ .

**Remark 1** When  $\delta \rightarrow 0$ , the first order approximations provided in (i), (ii) and (iii) can be rewritten in an unified way as

$$\Delta(t) \sim \frac{1}{2} \delta^2(t) K_2(x(t)).$$

This opens the door to the estimation of  $\Delta(t)$  via the estimation of  $K_2(x(t))$ , see also (8) in the ET framework and Section 6.

### 3 Application to the ET approximation

Recall that  $y(n) = \log(1/p_n)$ ,  $x(n) = \log(1/\alpha_n)$  with  $0 < p_n \leq \alpha_n < 1$ . Introduce

$$\tau_n = \frac{\log(1/p_n)}{\log(n)} \text{ and } \tau'_n = \frac{\log(1/\alpha_n)}{\log(n)}$$

so that  $p_n = n^{-\tau_n}$ ,  $\alpha_n = n^{-\tau'_n}$ ,  $\tau'_n \leq \tau_n$  and  $\delta(n) = (y(n) - x(n))/y(n) = 1 - \tau'_n/\tau_n$ . In the sequel,  $F$  is assumed to be increasing and twice differentiable and the cumulative hazard rate function is denoted by  $H(\cdot) = -\log \bar{F}(\cdot)$ . Following the ideas of Section 1, we let  $\varphi(\cdot) = (\bar{F})^{-1}(1/\exp(\cdot)) =$

$H^{-1}(\cdot)$  so that  $\varepsilon_{\text{ET}}(p_n; \alpha_n) = \Delta(n)$ . In this context, the assumption  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 \in \mathbb{R}$  is a sufficient condition for  $\log H^{-1}$  is extended regularly varying, see [8, Section B.2] for details on extended regular variation. This assumption has been introduced and discussed in [9, 10, 11]. The next result describes the tail behavior of  $F$  according to the sign of  $\theta_1$ . We refer to [10, Theorem 1] for a characterization under the weaker assumption of extended regular variation.

**Proposition 3 (Characterizations, ET framework)** *Suppose  $F$  is increasing, twice differentiable and  $K_1'$  is ultimately monotone. Let  $x^* := \sup\{x : F(x) < 1\}$  be the endpoint of  $F$ .*

- (i) *If  $H \in RV_{\beta}$ ,  $\beta > 0$ , then  $K_1 \in RV_0$  and  $\ell_1 = 1/\beta$ .*
- (ii)  *$K_1 \in RV_{\theta_1}$ ,  $\theta_1 > 0$  (and thus  $\ell_1 = \infty$ ) if and only if  $x^* = \infty$  and  $H(\exp \cdot) \in RV_{1/\theta_1}$ .*
- (iii)  *$K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 0$  (and thus  $\ell_1 = 0$ ) if and only if  $x^* < \infty$  and  $H(x^*(1 - 1/\cdot)) \in RV_{-1/\theta_1}$ .*

In the case (i) where  $H$  is regularly varying with index  $\beta > 0$ , necessarily  $\theta_1 = 0$  and  $F$  is referred to as a Weibull tail-distribution, see for instance [3, 16, 19]. Such distributions encompass Gaussian, Gamma, Exponential and strict Weibull distributions. In the case (ii) where  $H(\exp \cdot)$  is regularly varying,  $F$  is called a log-Weibull tail-distribution, see [2, 14, 18], the most popular example being the lognormal distribution. The case (iii) corresponds to distributions with a Weibull tail behavior in the neighborhood of a finite endpoint.

Besides, let us highlight that the domain of attraction associated with  $F$  depends on the position of  $\theta_1$  with respect to 1. Note that [10, Proposition 1] provides a similar classification under the weaker assumption of extended regular variation.

**Proposition 4 (Domains of attraction, ET framework)** *Suppose  $F$  is increasing, twice differentiable and  $K_1'$  is ultimately monotone.*

- (i) *If  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 1$ , then  $F \in \text{MDA}(\text{Gumbel})$ .*
- (ii) *If  $F \in \text{MDA}(\text{Fréchet})$  then  $K_1 \in RV_1$ .*
- (iii) *If  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 > 1$ , then  $F$  does not belong to any domain of attraction.*

These results justify the assumption  $\theta_1 \leq 1$  introduced in **(A3)**:  $\text{MDA}(\text{Gumbel})$  is associated with  $\theta_1 < 1$  while  $\text{MDA}(\text{Fréchet})$  is associated with  $\theta_1 = 1$ . However, there is no perfect one-to-one correspondence as illustrated by the following two examples:

- Consider the distribution defined by  $H_a^{-1}(x) = \exp \int_1^x \exp(-\log(t)^a) dt$ ,  $x \geq 1$ ,  $a > 1$ . From [8, Corollary 1.1.10], this distribution belongs to  $\text{MDA}(\text{Gumbel})$  while  $K_1(x) = x \exp(-(\log x)^a)$  is not regularly varying.
- Consider the distribution defined by  $H^{-1}(x) = \exp(x \log x)$ ,  $x \geq 1$ . From [8, Corollary 1.2.10], this distribution does not belong to  $\text{MDA}(\text{Fréchet})$  while  $K_1(x) \sim x \log x$  is regularly varying with index  $\theta_1 = 1$ .

The situation  $\theta_1 > 1$  which does not correspond to any domain of attraction is sometimes referred to as super-heavy tails, see [2] or [5, Section 8.8] for further developments on this topic. Applying Proposition 1 to the ET framework yields:

**Theorem 1 (Necessary and sufficient conditions on  $(\alpha_n, p_n)$  for  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$ )**

Suppose  $F$  is increasing, twice differentiable and **(A3)**, **(A4)** hold. Let  $0 < p_n \leq \alpha_n < 1$  such that  $\limsup \delta(n) < 1$  or equivalently  $\limsup \log(1/p_n)/\log(1/\alpha_n) < \infty$ .

(i) If  $\ell_1 \in \{0, 1\}$  then  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$ .

(ii) If  $\ell_1 \in (0, \infty) \setminus \{1\}$  then  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$  if and only if  $\delta(n) \rightarrow 0$  as  $n \rightarrow \infty$ .

(iii) If  $\ell_1 = \infty$  then  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$  if and only if  $\delta^2(n)K_2(\tau_n \log n) \rightarrow 0$  as  $n \rightarrow \infty$ .

If, moreover,  $n\alpha_n \rightarrow \infty$  then  $\delta(n) \rightarrow 0$  implies  $\limsup \tau_n \leq 1$  i.e.  $\limsup \log(1/p_n)/\log(n) \leq 1$ .

First, if  $F \in \text{MDA}(\text{Fréchet})$  then  $\theta_1 = 1$  in view of Proposition 4(ii) and thus  $\ell_1 = \infty$ . From Theorem 1(iii), it is possible to extrapolate even though the ET method has not been designed for this situation:  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$  under the restriction on  $(\alpha_n, p_n)$  that  $\delta^2(n)K_2(\tau_n \log n) \rightarrow 0$  as  $n \rightarrow \infty$ . Second, it appears that, from the extrapolation error point of view, three sub-domains of  $\text{MDA}(\text{Gumbel})$  can be exhibited:

- $\text{MDA}_1(\text{Gumbel})$  defined by  $\ell_1 \in \{0, 1\}$  and where the relative extrapolation error tends to zero as soon as  $\limsup \log(1/p_n)/\log(1/\alpha_n) < \infty$ . As illustrated by Proposition 3(iii), the case  $\ell_1 = 0$  includes distributions with a finite endpoint. The case  $\ell_1 = 1$  encompasses Weibull tail-distributions with shape parameter  $\beta = 1$  (Proposition 3(i)), i.e. close to the Exponential distribution (e.g. the Gamma distribution) as well as the class E, see [7].
- $\text{MDA}_2(\text{Gumbel})$  defined by  $\ell_1 \in (0, \infty) \setminus \{1\}$  and where the relative extrapolation error tends to zero for extreme quantiles close to the maximal observation in the sense that  $\limsup \log(1/p_n)/\log(n) \leq 1$  as  $n \rightarrow \infty$ . Extreme orders such as  $p_n = n^{-\tau}$ ,  $\tau > 1$  are thus not permitted. As illustrated by Proposition 3(i), this situation encompasses Weibull tail-distributions with shape parameter  $\beta \neq 1$  i.e. far from the Exponential distribution (the Gaussian distribution for instance).
- $\text{MDA}_3(\text{Gumbel})$  defined by  $\ell_1 = \infty$  and where the relative extrapolation error tends to zero under strong restrictions on the order  $p_n$  of the extreme quantile:  $\log(1/p_n)/\log(n) = 1 + o(|K_2(\tau_n \log n)|^{1/2})$  as  $n \rightarrow \infty$ . As illustrated by Proposition 3(ii), this case corresponds to log-Weibull tail-distributions (including the lognormal distribution).

We refer to Table 1 for examples of distributions in each sub-domain. Note that these three sub-domains do not cover the whole  $\text{MDA}(\text{Gumbel})$  since they require the existence of  $\ell_1$  and thus  $K_1$ . To conclude this part, one may obtain first order approximations of the relative extrapolation error  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  thanks to Proposition 2. The results are collected in Theorem 2 below. Remark that the assumption  $|K_2|$  is regularly varying is needed only in the case  $\ell_1 = 1$ , since, in other situations it is a consequence of **(A3)**, see Lemma 3.



**Theorem 2 (First order approximations of  $\varepsilon_{ET}(p_n; \alpha_n)$ )** *Suppose the assumptions of Theorem 1 hold.*

(i) *Assume  $F \in MDA_1(\text{Gumbel})$ . If  $\ell_1 = 1$ , assume there exists  $\theta_2 \leq 0$  such that  $|K_2| \in RV_{\theta_2}$ .*

(a) *If  $\delta(n) \rightarrow 0$  then  $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{1}{2}\delta^2(n)K_2(\tau'_n \log n)$ .*

(b) *If  $\delta(n) \rightarrow \delta_\infty \in (0, 1)$  then  $\varepsilon_{ET}(p_n; \alpha_n) \sim \delta_\infty^2(1 - \delta_\infty)^{-\theta_2}c(\delta_\infty, \ell_1 + \theta_2)K_2(\tau'_n \log n)$ .*

(ii) *Assume  $F \in MDA_2(\text{Gumbel})$*

(a) *If  $\delta(n) \rightarrow 0$  then  $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{\ell_1(\ell_1-1)}{2}\delta^2(n)$ .*

(b) *If  $\delta(n) \rightarrow \delta_\infty \in (0, 1)$  then  $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow \delta_\infty^2\ell_1(\ell_1 - 1)c(\delta_\infty, \ell_1)$ .*

(iii) *Assume  $F \in MDA_3(\text{Gumbel})$*

(a) *If  $\delta(n)K_1(\log n) \rightarrow 0$  then  $\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{1}{2}\delta^2(n)K_2(\tau'_n \log n)$ .*

(b) *If  $\delta(n)K_1(\log n) \rightarrow a \in (0, \infty]$  then  $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow \int_0^a u \exp(-u)du$ .*

Before commenting the asymptotic behavior of  $\varepsilon_{ET}(p_n; \alpha_n)$ , let us compare our results with [4].

**Remark 2** *The asymptotic equivalents provided by [4, Theorem 2] can be compared to our results. However, let us point out that [4, Theorem 2] only holds in the case where  $\delta(n) \rightarrow 0$  as  $n \rightarrow \infty$  and for the particular case of “Weibull type distributions” implying in particular that  $\ell_1 \neq 0$ . It can be shown that the asymptotic equivalents provided by [4], Theorem 2.1, Theorem 2.2 and Theorem 2.3 coincide with the ones of Theorem 2(ii)-(a), Theorem 2(i)-(a) and Theorem 2(iii)-(a) respectively, up to a typo in the statement of [4, Theorem 2.2].*

The only situation where  $\delta(n) \rightarrow \delta_\infty \neq 0$  and  $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$  occurs for  $F \in MDA_1(\text{Gumbel})$ . In this particular case, it is possible to choose extreme orders such that  $p_n = n^{-\tau}$ ,  $\tau > 1$ , and the relative extrapolation error tends to zero at a logarithmic rate. As expected, in the three situations (i,ii,iii)-(a) where  $\delta(n) \rightarrow 0$  and  $\varepsilon_{ET}(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$ , the convergence is the fastest in  $MDA_1(\text{Gumbel})$  and the slowest in  $MDA_3(\text{Gumbel})$ . Let us also highlight that the rate of convergence is independent from the distribution in  $MDA_2(\text{Gumbel})$ . As already pointed out in Remark 1, in all three cases, the equivalent provided by Theorem 2(i,ii,iii)-(a) can be rewritten in an unified way as

$$\varepsilon_{ET}(p_n; \alpha_n) \sim \frac{1}{2}\delta^2(n)K_2(\tau'_n \log n) \text{ as } n \rightarrow \infty, \quad (8)$$

which can thus be estimated from real data, see Section 6. Before that, to illustrate these results, let us focus on the distributions introduced in Table 1. Clearly, in all six cases,  $F \in MDA(\text{Gumbel})$ ,  $K_1$  and  $|K_2|$  are regularly varying so that the assumptions of Theorem 2 are fulfilled. Let us consider the case where  $p_n = 1/(n \log n)$  and  $\alpha_n = (\log n)/n$  leading to

$$\tau_n = 1 + \frac{\log \log n}{\log n}, \tau'_n = 1 - \frac{\log \log n}{\log n} \text{ and } \delta(n) \sim 2 \frac{\log \log n}{\log n} \text{ as } n \rightarrow \infty. \quad (9)$$

Let us stress that  $\delta(n) \rightarrow 0$  and  $\delta(n)K_1(\log n) \rightarrow 0$  so that Theorem 2(i,ii,iii)-(a) holds and  $\varepsilon_{\text{ET}}(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$  for all six distributions. The associated first order approximations of  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  are provided in Table 2 (second column). In most cases the convergence of the relative extrapolation error to zero is rather slow. The log-Weibull( $\beta > 1$ ) distribution corresponds to the worst case, since arbitrary low rates of convergence can be obtained by letting  $\beta \xrightarrow{>} 1$ . At the opposite, the Finite endpoint( $\beta > 0$ ) distribution is the most favorable case, letting  $\beta \xrightarrow{>} 0$  could lead to arbitrary high logarithmic rates of convergence.

As a conclusion, the extrapolation abilities of the ET method are poor. To overcome this limitation, two main approaches are usually considered. The first one is to focus on a subset of distributions, for instance Weibull tail-distributions in  $\text{MDA}_2(\text{Gumbel})$ , where adapted estimators can outperform the ET method, see [15] for an illustration. The second one is to rely on new assumptions on the distribution tail, such as the log-generalized Weibull tail limit [1, 11].

## 4 Application to Weissman approximation

When  $F \in \text{MDA}(\text{Fréchet})$ ,  $\gamma_n > 0$  and the GPD approximation (1) can be simplified by letting  $\sigma_n = \gamma_n q(\alpha_n)$ , see [8, Theorem 1.2.5], leading to

$$\tilde{q}_{\text{W}}(p_n; \alpha_n) = q(\alpha_n) \left( \frac{\alpha_n}{p_n} \right)^{\gamma_n}, \quad (10)$$

which is called Weissman approximation in the sequel. Weissman estimator [26] is then obtained by replacing the intermediate quantile  $q(\alpha_n)$  and the tail index  $\gamma_n$  by appropriate estimators:

$$\hat{q}_{\text{W}}(p_n; \alpha_n) = \hat{q}(\alpha_n) \left( \frac{\alpha_n}{p_n} \right)^{\hat{\gamma}_n}.$$

The most common choices are  $\hat{q}(\alpha_n) = X_{n-k_n+1,n}$ , see Section 1, and Hill estimator [23]:

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

Taking the logarithm of (10) yields

$$\log q(p_n) - \log \tilde{q}_{\text{W}}(p_n; \alpha_n) = \log q(p_n) - \log q(\alpha_n) - \gamma_n \log(\alpha_n/p_n)$$

and thus, similarly to the ET case (4), the extrapolation error can be interpreted as a first order Taylor remainder. To this end, recall that  $y(n) = \log(1/p_n)$ ,  $x(n) = \log(1/\alpha_n)$  with  $0 < p_n < \alpha_n < 1$  and introduce  $\varphi(\cdot) = \log(\bar{F})^{-1}(1/\exp(\cdot)) = \log H^{-1}(\cdot) = \log U(\exp \cdot)$  where  $U$  is the tail quantile function, so that

$$\log q(p_n) - \log \tilde{q}_{\text{W}}(p_n; \alpha_n) = \varphi(y(n)) - \varphi(x(n)) - \gamma_n(y(n) - x(n)) \text{ where } \gamma_n = \varphi'(x(n)).$$

The quantity of interest is

$$\varepsilon_{\text{W}}(p_n; \alpha_n) := (q(p_n) - \tilde{q}_{\text{W}}(p_n; \alpha_n))/q(p_n) = 1 - \exp(-\Delta(n) \log q(p_n)), \quad (11)$$

where  $\Delta(n)$  is defined in (7). The next result provides a characterization of the tail behavior of  $F$  according to the limit  $\ell_1$ .

**Proposition 5 (Characterizations, Weissman framework)** *Suppose  $F$  is increasing, twice differentiable and  $K_1'$  is ultimately monotone.*

- (i) *If  $\exp H \in RV_{1/\gamma}$ ,  $\gamma > 0$ , then  $K_1 \in RV_0$  and  $\ell_1 = 1$ .*
- (ii) *If  $H \in RV_\beta$ ,  $\beta > 0$ , then  $K_1 \in RV_0$  and  $\ell_1 = 0$ .*
- (iii) *If  $H(\exp(\cdot)) \in RV_\beta$ ,  $\beta > 0$  then  $K_1 \in RV_0$  and  $\ell_1 = 1/\beta$ .*

In the case (i) where  $\exp H$  is regularly varying with positive index,  $F$  is referred to as a Pareto tail-distribution. Burr, Cauchy, Fréchet, Pareto, Student distributions are the most famous ones. The cases (ii) and (iii) correspond respectively to Weibull and log-Weibull tail-distributions, see Proposition 3(i,ii). Besides, let us highlight that, in the Weissman framework, the domain of attraction associated with  $F$  depends on the position of  $\theta_1$  with respect to 0:

**Proposition 6 (Domains of attraction, Weissman framework)** *Suppose  $F$  is increasing, twice differentiable and  $K_1'$  is ultimately monotone.*

- (i) *If  $F \in MDA(\text{Fréchet})$  then  $K_1 \in RV_0$  and  $\ell_1 = 1$ .*
- (ii) *If  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 > 0$ , then  $F$  does not belong to any domain of attraction.*
- (iii) *If  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 0$ , then  $F$  is not a proper cumulative distribution function.*

Let us first note that, in case (i), there is no perfect one-to-one correspondence between  $K_1 \in RV_0$  and  $F \in MDA(\text{Fréchet})$  as illustrated by the following example. Consider the distribution defined by  $H^{-1}(x) = \exp(x \log x)$ ,  $x \geq 1$ . From [8, Corollary 1.2.10], this distribution does not belong to  $MDA(\text{Fréchet})$  while  $K_1(x) = 1 + (x \log x)^{-1}$  is verifying  $\ell_1 = 1$  and is thus regularly varying with index  $\theta_1 = 0$ .

Second, in view of Proposition 5 and 6, the only case of interest is  $\theta_1 = 0$ . The asymptotic behavior of  $\varepsilon_W(p_n; \alpha_n)$  is thus investigated in the three situations where  $K_1 \in RV_0$  described by Proposition 5: Pareto / Weibull / log-Weibull tail-distributions. The next two results are derived by applying Proposition 2 to the Weissman framework.

**Theorem 3 (Necessary and sufficient conditions on  $(\alpha_n, p_n)$  for  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$ )**

*Suppose  $F$  is increasing, twice differentiable and **(A4)** holds. Let  $0 < p_n \leq \alpha_n < 1$  such that  $\limsup \delta(n) < 1$  or equivalently  $\limsup \log(1/p_n)/\log(1/\alpha_n) < \infty$ .*

- (i) *Suppose  $F \in MDA(\text{Fréchet})$  with tail index  $\gamma > 0$ . Let  $L(t) := t^{-\gamma}U(t)$ ,  $\eta(t) := tL'(t)/L(t)$ ,  $t > 0$  and assume  $|\eta| \in RV_\rho$  with  $\rho < 0$ . If  $\delta(n) \rightarrow \delta_\infty \in [0, 1)$  then  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$ .*
- (ii) *Weibull tail-distributions. Suppose  $H \in RV_\beta$ ,  $\beta > 0$ . Then,  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  if and only if  $\delta(n) \rightarrow 0$  as  $n \rightarrow \infty$ .*
- (iii) *Log-Weibull tail-distributions. Suppose  $H(\exp \cdot) \in RV_\beta$ ,  $\beta > 0$  and  $\beta \neq 1$ . Then,  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  if and only if  $\delta^2(n) \log q(p_n) \rightarrow 0$  as  $n \rightarrow \infty$ .*

In the situation (i) where  $F \in \text{MDA}(\text{Fréchet})$ , the function  $L$  is slowly-varying [5] and  $\eta$  is called the auxiliary function associated with  $L$ . The assumption  $|\eta| \in RV_\rho$ ,  $\rho < 0$ , is recurrent in extreme-value statistics to control the bias of estimators,  $\rho$  being known as the second-order parameter, see *e.g.* [17]. This assumption holds for most heavy-tailed distributions such as Burr, Cauchy, Fréchet, Pareto or Student distributions. Let us also remark that one can choose extreme orders such that  $p_n = n^{-\tau}$ ,  $\tau > 1$  as in  $\text{MDA}_1(\text{Gumbel})$ , see Theorem 2(i)-(b), and still obtain  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Theorem 3(ii,iii) also shows that  $\varepsilon_W(p_n; \alpha_n)$  asymptotically vanishes provided  $\delta(n) \rightarrow 0$  in case of Weibull distributions or provided  $\delta^2(n) \log q(p_n) \rightarrow 0$  in case of log-Weibull tail-distributions even though they do not belong to  $\text{MDA}(\text{Fréchet})$ .

**Theorem 4 (First order approximations of  $\varepsilon_W(p_n; \alpha_n)$ )**

(i) Suppose the assumptions of Theorem 3(i) hold.

If  $\delta(n) \rightarrow \delta_\infty \in [0, 1)$  then  $\varepsilon_W(p_n; \alpha_n) \sim -\frac{1}{1-\delta_\infty} \delta(n) \log(1/\alpha_n) \eta(1/\alpha_n)$ .

(ii) Suppose the assumptions of Theorem 3(ii) hold.

(a) If  $\delta(n) \rightarrow 0$  then  $\varepsilon_W(p_n; \alpha_n) \sim -\frac{1}{2\beta} \delta^2(n)$ .

(b) If  $\delta(n) \rightarrow \delta_\infty \in (0, 1)$  then  $\varepsilon_W(p_n; \alpha_n) \rightarrow 1 - \exp\left(\frac{1}{\beta} \frac{\delta_\infty^2}{1-\delta_\infty}\right)$ .

(iii) Suppose the assumptions of Theorem 3(iii) hold.

(a) If  $\delta^2(n) \log q(p_n) \rightarrow 0$  then  $\varepsilon_W(p_n; \alpha_n) \sim \frac{1-\beta}{2\beta^2} \delta^2(n) \log q(p_n)$ .

(b) If  $\delta^2(n) \log q(p_n) \rightarrow a \in (0, \infty)$  then  $\varepsilon_W(p_n; \alpha_n) \rightarrow 1 - \exp\left(-\frac{1-\beta}{2\beta^2} a\right)$ .

If  $F \in \text{MDA}(\text{Fréchet})$ , case (i), one can choose extreme orders such that  $p_n = n^{-\tau}$ ,  $\tau > 1$  leading to polynomial extrapolation errors which is coherent with usual convergence rates, see for instance [8, Section 3.2].

**Remark 3** These conclusions can also be found in [4, Theorem 1] where it is established that  $\varepsilon_W(p_n; \alpha_n) \sim c \eta(1/p_n)$ , for some explicit constant  $c \in \mathbb{R}$ , under a second order assumption on  $F \in \text{MDA}(\text{Fréchet})$ .

Up to our knowledge, situations (ii) and (iii) have not been considered so far. They can be illustrated similarly to Section 3 by considering  $p_n = 1/(n \log n)$  and  $\alpha_n = (\log n)/n$ . These choices entail  $\delta(n) \rightarrow 0$  and  $\delta^2(n) \log q(p_n) \rightarrow 0$ , see (9), so that Theorem 4(ii,iii)-(a) can be applied and  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  as  $n \rightarrow \infty$  for the last five distributions of Table 1. The first order approximations of  $\varepsilon_W(p_n; \alpha_n)$  are provided in Table 2 (third column). Surprisingly, in  $\text{MDA}_2(\text{Gumbel})$  and  $\text{MDA}_3(\text{Gumbel})$ , the convergence of  $\varepsilon_W(p_n; \alpha_n)$  to zero is equivalent to, or even faster than, the convergence of  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$ . In such cases, Weissman approximation is better than the ET one even though Weissman estimator was not initially designed for these frameworks. This confirms the conclusion drawn in Section 3: the extrapolation abilities of the ET method are poor, even when compared to *a priori* ill-adapted competitors. Finally, since  $\varepsilon_W(p_n; \alpha_n) < 0$  while  $\varepsilon_{\text{ET}}(p_n; \alpha_n) > 0$  for log-Weibull tail-distributions, there is a hope to build extrapolation methods achieving a compromise between ET and Weissman approximations leading to smaller errors.

## 5 Numerical illustrations

First, the quality of the first order approximations of the ET relative extrapolation error given in Table 2 is assessed graphically. Recall that these results are obtained by applying Theorem 2 to sequences  $(\tau_n)$  and  $(\tau'_n)$  given in (9) and distributions described in Table 1: Finite endpoint( $\beta = 5$ ), Gamma( $a = 0.1$ ), Weibull( $\beta = 5$ ), Gaussian, log-Weibull( $\beta = 3$ ) and lognormal( $\sigma = 0.5$ ). The exact relative extrapolation error  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  as well as the corresponding first order approximation provided by Theorem 2 are computed as functions of  $\log n$ . The results are displayed on Figures 1–3. It appears that, for all six distributions, the relative extrapolation error converges towards zero as predicted by Theorem 2, even though the convergence can be very slow in  $\text{MDA}_3(\text{Gumbel})$ , see Figure 3. In all cases, the asymptotic sign of  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  is coherent with the first order equivalent given in Table 2 (second column): Positive for Gamma( $a < 1$ ), log-Weibull( $\beta > 1$ ) and lognormal distributions, negative for Finite endpoint( $\beta > 0$ ), Weibull( $\beta > 1$ ) and Gaussian distributions. Finally, the first order equivalent provides a reasonable approximation of the error behavior in all situations.

Second, Figure 4 displays the relative extrapolation error  $\varepsilon_{\text{W}}(p_n; \alpha_n)$  associated with Weissman estimator together with its first order approximation in  $\text{MDA}(\text{Fréchet})$  provided by Theorem 4(i) as a function of  $\log n$ . These results are obtained by choosing sequences  $p_n = n^{-5/4}$  and  $\alpha_n = n^{-3/4}$  such that  $\delta(n) = 2/5$  and by considering a Burr distribution defined by  $U(t) := (t^{-\rho} - 1)^{-1/\rho}$ ,  $t \geq 1$ ,  $\rho < 0$ , with tail index  $\gamma = 1$  and auxiliary function  $\eta(t) = 1/(t^{-\rho} - 1)$ . Clearly,  $\eta$  is regularly varying with index  $\rho$ . In both cases  $\rho = -1/3$  (top) and  $\rho = -1/4$  (bottom), it appears that the relative extrapolation error converges to zero even though  $\delta(n)$  is constant. This graphical assessment is in agreement with Theorem 4(i). As expected, both errors are negative since the auxiliary function  $\eta$  is positive. It also appears that, the smaller  $\rho$ , the faster the convergence is. This is in accordance with  $\eta \in \text{RV}_\rho$ . Finally, the first order equivalent also provides a reasonable approximation of the error behavior in the Burr case.

## 6 Application to real data

The goal of this section is to illustrate how the first order approximations of the relative extrapolation error provided by Theorem 2 and Theorem 4 can be used to assess the extrapolation range associated with ET or Weissman methods. Focusing on the ET framework and letting  $\alpha_n := k_n/n$  with  $k_n \rightarrow \infty$  and  $k_n/n \rightarrow 0$ , Remark 3 or equivalently (8) yields

$$\varepsilon_{\text{ET}}(p_n; \alpha_n) \sim \frac{1}{2} \delta^2(n) K_2(\log(n/k_n))$$

when  $\delta(n) \rightarrow 0$ . In view of (20) in the proof of Lemma 3, we thus introduce

$$\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n) := \frac{1}{2} \delta^2(n) \hat{K}_1(\log(n/k_n)) \left( \hat{K}_1(\log(n/k_n)) + \hat{\theta}_1 - 1 \right),$$

where  $\hat{\theta}_1$  and  $\hat{K}_1(\log(n/k_n))$  are suitable estimators of  $\theta_1$  and  $K_1(\log(n/k_n))$  respectively. We refer to [11, Equations (19,20)] and [1, Equations (7,8)] for examples.

The finite-sample behavior of  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$  is first illustrated on simulated samples from the Gamma( $a = 0.1$ ) distribution, where  $p_n = 1/(n \log n)$  and  $\alpha_n = 40 \log(n)/n$ . Figure 5 displays the behavior of  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$  built on [11, Equations (19,20)] and averaged over  $N = 100$  replications as a function of  $n \in \{10^3, \dots, 10^6\}$  compared to the true relative error  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  and its first order approximation. It appears that  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$  shows pretty good results in terms of bias, even for moderate values of  $n$ . This justifies the use of  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$  on the considered real dataset, that we shall graphically demonstrate to be approximately Gamma distributed.

The dataset under consideration is a set of wind daily measures (in m/s) at Reims (France) from 01/01/1981 to 04/30/2011. For seasonality reasons, only the months from October to March are considered, resulting in  $n = 5,371$  measures. Figure 6 displays two estimators of  $\theta_1$  as functions of the number of exceedances  $k_n$ : The proposal introduced in [1, Equation (7)] together with its 95% asymptotic confidence interval provided in [1, Theorem 3] and the proposal associated with [11, Equation (19)]. It appears that the two estimates are similar, and, for  $k_n \geq 200$ , the value  $\theta_1 = 0$  cannot be excluded since it belongs to all the 95% asymptotic confidence intervals. Moreover, the quantile-quantile plot on Figure 7 (empirical quantiles vs Gamma quantiles) displays a strong linear trend. These two graphical assessments point towards the same conclusion: It makes sense to estimate the relative extrapolation error induced by the ET method using  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$ . Letting  $k = 40 \log n$  (which is in the stability range of  $\hat{\theta}_1$ , see Figure 6) yields  $|\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)| \approx \delta^2(n) \times 1\%$  which is accordance with our previous conclusions: The ET method is able to extrapolate far into the tails in case of Gamma-like distributions.

## 7 Proofs of main results

**Proof of Proposition 1.** (i) If  $\ell_1 \in \{0, 1\}$  then Lemma 3(i,ii) shows that  $\ell_2 = 0$ . Lemma 5(i) concludes the proof.

(ii) If  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$  then Lemma 3(iii) entails that  $\ell_2$  is finite and non zero. Lemma 5(i,ii) concludes the proof.

(iii) If  $\ell_1 = \infty$  then  $K_2$  is regularly varying from Lemma 3(iv) and thus  $K_2(x(t))$  and  $K_2(y(t))$  are of the same order as  $t \rightarrow \infty$  under **(A1)**. Lemma 5(i,ii) concludes the proof. ■

**Proof of Proposition 2.** (i) If  $\ell_1 \in \{0, 1\}$  and  $\delta(t) \rightarrow \delta_\infty \in [0, 1)$  as  $t \rightarrow \infty$ , then the result is a consequence of Lemma 4(i) and of  $K_2(y(t)) \sim (1 - \delta_\infty)^{-\theta_2} K_2(x(t))$  since  $|K_2| \in RV_{\theta_2}$ .

(ii) Assume  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$ . Lemma 3(iii) entails  $\ell_2 = \ell_1(\ell_1 - 1)$ , and Lemma 4(i) yields

$$\Delta(t) \sim \delta^2(t) \int_0^1 K_2(y(t)(1 - \delta(t)u))(1 - \delta(t)u)^{\ell_1 - 2} u du.$$

When  $\delta(t) \rightarrow \delta_\infty \in [0, 1)$  as  $t \rightarrow \infty$ , Lebesgue's dominated convergence theorem entails

$$\int_0^1 K_2(y(t)(1 - \delta(t)u))(1 - \delta(t)u)^{\ell_1 - 2} u du \rightarrow \ell_1(\ell_1 - 1) \int_0^1 (1 - \delta_\infty u)^{\ell_1 - 2} u du,$$

and the result is proved.

(iii) Assume  $\ell_1 = \infty$ . Then Lemma 3(iv) entails that  $K_2(x) \sim K_1^2(x)$  as  $x \rightarrow \infty$ . Consequently,

(21) in the proof of Lemma 4 and Lebesgue's dominated convergence theorem yield

$$\begin{aligned}\Delta(t) &\sim \delta^2(t) \int_0^1 \frac{K_1^2(y(t)(1 - \delta(t)u))}{(1 - \delta(t)u)^2} \exp(K_1(y(t))L_{\theta_1}(1 - \delta(t)u)(1 + o(1))) \, udu \\ &\sim \delta^2(t)K_1^2(y(t)) \int_0^1 (1 - \delta(t)u)^{2\theta_1-2} \exp(K_1(y(t))(L_{\theta_1}(1 - \delta(t)u)(1 + o(1)))) \, udu,\end{aligned}$$

in view of the regular variation property **(A3)**. Two main situations are considered:

1. If  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ , then  $L_{\theta_1}(1 - \delta(t)u) \sim -\delta(t)u$ . Letting  $A(t) = \delta(t)K_1(y(t))$ , it follows

$$\Delta(t) \sim A^2(t) \int_0^1 \exp(-A(t)u(1 + o(1))) \, udu \sim \Phi(A(t)(1 + o(1)))A^2(t),$$

with  $\Phi(\cdot) = \Psi_1(\cdot; 1)$ , see Lemma 1. Three sub-cases arise: (a) If  $A(t) \rightarrow 0$  as  $t \rightarrow \infty$ , then  $\Phi(A(t)) \rightarrow 1/2$  in view of Lemma 1(i) and

$$\Delta(t) \sim \frac{1}{2}\delta^2(t)K_1^2(y(t)) \sim \frac{1}{2}\delta^2(t)K_1^2(x(t))$$

since  $K_1$  is regularly varying and  $x(t) \sim y(t)$  when  $\delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ . (b) If  $A(t) \rightarrow a \in (0, \infty)$  then  $\Delta(t) \rightarrow a^2\Phi(a) = \int_0^a u \exp(-u) \, du$  as  $t \rightarrow \infty$  in view of the continuity of  $\Phi$ , see Lemma 1(i). If  $A(t) \rightarrow \infty$ , then  $\Phi(A(t)) \sim 1/A^2(t)$  from Lemma 1(ii) and therefore  $\Delta(t) \rightarrow 1 = \int_0^\infty u \exp(-u) \, du$  as  $t \rightarrow \infty$ .

2. If  $\delta(t) \rightarrow \delta_\infty \in (0, 1)$ , then  $A(t) \rightarrow \infty$  as  $t \rightarrow \infty$ . Two successive change of variables yield

$$\begin{aligned}\Delta(t) &\sim \delta_\infty^2 K_1^2(y(t)) \int_0^1 (1 - \delta_\infty u)^{2\theta_1-2} \exp(K_1(y(t))L_{\theta_1}(1 - \delta(t)u)(1 + o(1))) \, udu \\ &\sim K_1^2(y(t)) \int_{1-\delta_\infty}^1 (1 - v)v^{2\theta_1-2} \exp(K_1(y(t))L_{\theta_1}(v)(1 + o(1))) \, dv \\ &\sim K_1^2(y(t)) \int_{L_{\theta_1}(1-\delta_\infty)}^0 (L_{\theta_1}^{-1}(w))^{\theta_1-1} (1 - L_{\theta_1}^{-1}(w)) \exp(K_1(y(t))w(1 + o(1))) \, dw.\end{aligned}$$

Let us introduce  $\xi(w) = (L_{\theta_1}^{-1}(w))^{\theta_1-1} (1 - L_{\theta_1}^{-1}(w))$  for all  $w \in [L_{\theta_1}(1 - \delta_\infty), 0]$ . Routine calculations show that  $\xi(0) = 0$  and  $\xi'(0) = -1$ . A second order Taylor expansion thus yields  $\xi(w) = -w + w^2\xi''(\eta_w)/2$  with  $\eta_w \in [w, 0] \subset [L_{\theta_1}(1 - \delta_\infty), 0]$ . Replacing, we get

$$\begin{aligned}\Delta(t) &= -K_1^2(y(t)) \int_{L_{\theta_1}(1-\delta_\infty)}^0 w \exp(K_1(y(t))w(1 + o(1))) \, dw(1 + o(1)) + R(t) \\ &= K_1^2(y(t))\Psi_1(K_1(y(t))(1 + o(1)); -L_{\theta_1}(1 - \delta_\infty)) + R(t),\end{aligned}$$

where  $\Psi_1$  is defined in Lemma 1 and

$$R(t) = \frac{1}{2}K_1^2(y(t)) \int_{L_{\theta_1}(1-\delta_\infty)}^0 w^2\xi''(\eta_w) \exp(K_1(y(t))w(1 + o(1))) \, dw(1 + o(1)).$$

Remarking that  $|\xi''|$  is bounded on compact sets, there exists  $M > 0$  such that

$$|R(t)| \leq MK_1^2(y(t))\Psi_2(K_1(y(t))(1 + o(1)); -L_{\theta_1}(1 - \delta_\infty)),$$

where  $\Psi_2$  is defined in Lemma 1. As a consequence of Lemma 1(ii),  $R(t) = O(1/K_1(y(t)))$  and  $\Delta(t) \rightarrow 1$  as  $t \rightarrow \infty$ . ■

**Proof of Proposition 3.** Proposition 3(i) (resp. (ii), (iii)) is a straightforward consequence of Lemma 2(i) (resp. (ii), (iii)), with  $\varphi = H^{-1}$  in the ET framework. ■

**Proof of Proposition 4.** (i) Assume  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 1$  and let  $U(\cdot) = H^{-1}(\log \cdot)$  be the tail quantile function. For all  $x > 0$  and  $t > 0$ , consider

$$\frac{U'(tx)}{U'(t)} = \frac{1}{x} \frac{(H^{-1})'(\log tx)}{(H^{-1})'(\log t)} = \frac{1}{x} \frac{\log t}{\log tx} \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} \frac{K_1(\log tx)}{K_1(\log t)} \sim \frac{1}{x} \frac{H^{-1}(\log tx)}{H^{-1}(\log t)}$$

as  $t \rightarrow \infty$  since  $K_1 \in RV_{\theta_1}$  and thus  $K_1(\log \cdot)/\log(\cdot) \in RV_0$ . Besides,

$$\frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left(\int_{\log t}^{\log tx} (\log H^{-1})'(u) du\right) = \exp\left(\log x \int_0^1 \frac{K_1(\log t + v \log x)}{\log t + v \log x} dv\right),$$

and the regular variation property of  $K_1$  implies that

$$\frac{K_1(\log t + v \log x)}{\log t + v \log x} = \frac{K_1(\log t)}{\log t} (1 + o(1))$$

as  $t \rightarrow \infty$  uniformly locally on  $v \in [0, 1]$ . It follows that

$$\frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left(\log x \frac{K_1(\log t)}{\log t} (1 + o(1))\right) \rightarrow 1 \quad (12)$$

as  $t \rightarrow \infty$  since  $K \in RV_{\theta_1}$  with  $\theta_1 < 1$ . As a conclusion,  $U'(tx)/U'(t) \rightarrow 1/x$  as  $t \rightarrow \infty$  for all  $x > 0$  and thus  $U' \in RV_{-1}$ . This implies that  $F \in \text{MDA}(\text{Gumbel})$ , see [8, Corollary 1.1.10].

(ii) Assume  $F \in \text{MDA}(\text{Fréchet})$ . From [8, Corollary 1.2.10], there exists  $\gamma > 0$  such that  $U \in RV_\gamma$ . Since  $H^{-1}(\cdot) = U(\exp \cdot)$ , it follows that

$$K_1(x) = x \frac{\exp(x) U'(\exp x)}{U(\exp x)} \sim \gamma x$$

as  $x \rightarrow \infty$  from the monotone density theorem [5, Theorem 1.7.2]. It is thus clear that  $K_1 \in RV_1$ .

(iii) Assume  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 > 1$ . First, Proposition 3(ii) implies that  $x^* = \infty$  and thus  $F \notin \text{MDA}(\text{Weibull})$ . Second, Proposition 4(ii) shows that  $F \in \text{MDA}(\text{Fréchet})$  entails  $K_1 \in RV_1$ . It is thus clear that  $F \notin \text{MDA}(\text{Fréchet})$ . Finally, it remains to show that  $F \notin \text{MDA}(\text{Gumbel})$ . To this end, consider for all  $x > 0$  and  $t \rightarrow \infty$ ,

$$\frac{U(tx)}{U(t)} = \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left(\log x \frac{K_1(\log t)}{\log t} (1 + o(1))\right)$$

from (12). Recalling that  $\theta_1 > 1$ , it is then clear that  $K_1(\log t)/(\log t) \rightarrow \infty$  as  $t \rightarrow \infty$  and therefore  $U(tx)/U(t) \rightarrow 0$  if  $x < 1$  while  $U(tx)/U(t) \rightarrow \infty$  if  $x > 1$ . Finally [8, Lemma 1.2.9] shows that  $F \notin \text{MDA}(\text{Gumbel})$  since  $U(tx)/U(t)$  does not converge to 1 as  $t \rightarrow \infty$ . ■

**Proof of Theorem 1.** The proof of (i)–(iii) is a direct application of Proposition 1 since **(A1)** is fulfilled under the assumptions  $0 < p_n \leq \alpha_n < 1$  and  $\limsup \delta(n) < 1$ . If  $\delta(n) \rightarrow 0$  and  $n\alpha_n \rightarrow \infty$  as  $n \rightarrow \infty$  then, for all  $A > 0$ ,  $(1 - \tau'_n) \log(n) \geq A$  for  $n$  large enough. Thus

$$\tau_n \leq \frac{\tau_n}{\tau'_n} \left(1 - \frac{A}{\log n}\right) = \frac{1}{1 - \delta(n)} \left(1 - \frac{A}{\log n}\right)$$

and consequently  $\limsup \tau_n \leq 1$ . ■



**Proof of Theorem 2.** The result is a consequence of Proposition 2. ■

**Proof of Proposition 5.** The proof of (i) (resp (ii), (iii)) relies on the application of Lemma 2 (iv) (resp (v), (i)) together with the fact that  $\varphi = \log H^{-1}$  in the Weissman framework. ■

**Proof of Proposition 6.** (i) Assume  $F \in \text{MDA}(\text{Fréchet})$ . From [8, Corollary 1.2.1],  $U \in RV_\gamma$  for some  $\gamma > 0$  which can be rewritten as  $\exp H \in RV_{1/\gamma}$ . Proposition 5(i) proves the result.

(ii) Assume  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 > 0$  or equivalently that  $(\log \varphi)' \in RV_{\theta_1-1}$ . Then  $\log \varphi \in RV_{\theta_1}$  from [5, Theorem 1.5.11]. Consequently,  $\log \log U(\exp(\cdot)) \in RV_{\theta_1}$  and therefore  $U$  is not regularly varying. Lemma 1.2.9 and Corollary 1.2.10 in [8] conclude the proof.

(iii) From (ii),  $\theta_1 < 0$  implies that  $U$  is ultimately decreasing and the conclusion follows. ■

**Proof of Theorem 3.** (i) The proof relies on the application of Lemma 4(i) with  $\ell_1 = 1$ :

$$\Delta(n) \sim \delta^2(n) \int_0^1 K_2(y(n)(1-\delta(n)u)) (1-\delta(n)u)^{-1} u du.$$

Besides,  $\varphi''(t) = \exp(t)\eta'(\exp(t))$  and consequently, as  $t \rightarrow \infty$ ,

$$K_2(t) \sim \frac{1}{\gamma} t \exp(t)\eta'(\exp(t)) \sim \frac{\rho}{\gamma} t \eta(\exp(t)),$$

since  $|\eta| \in RV_\rho$  implies  $x\eta'(x)/\eta(x) \rightarrow \rho$  as  $x \rightarrow \infty$ . It follows, when  $\delta(n) \rightarrow \delta_\infty \in [0, 1)$ ,

$$\Delta(n) \sim y(n)\delta^2(n) \frac{\rho}{\gamma} \int_0^1 u \eta\left(e^{y(n)(1-\delta(n)u)}\right) du.$$

Since  $|\eta| \in RV_\rho$ , from Potter's bounds [8, Proposition B.1.9], there exists  $0 < \epsilon < |\rho|$  such that

$$(1-\epsilon)e^{y(n)\delta(n)(1-u)(\rho-\epsilon)} \leq \frac{|\eta|\left(e^{y(n)(1-\delta(n)u)}\right)}{|\eta|\left(e^{x(n)}\right)} \leq (1+\epsilon)e^{y(n)\delta(n)(1-u)(\rho+\epsilon)}.$$

Recalling that  $\eta$  is ultimately monotone with a constant sign yields

$$\Delta(n) \sim \frac{\rho}{\gamma} \eta\left(e^{x(n)}\right) I_n y(n) \delta^2(n),$$

where  $I_n^- \leq I_n \leq I_n^+$  with

$$I_n^- = (1-\epsilon) \int_0^1 u e^{y(n)\delta(n)(1-u)(\rho-\epsilon)} du \text{ and } I_n^+ = (1+\epsilon) \int_0^1 u e^{y(n)\delta(n)(1-u)(\rho+\epsilon)} du.$$

Straightforward calculations show that  $y(n)\delta(n)x \int_0^1 u e^{y(n)\delta(n)(1-u)x} du \rightarrow -1$ , as  $n \rightarrow \infty$  for all  $x < 0$ , since  $y(n)\delta(n) \rightarrow \infty$  in view of  $\alpha_n/p_n \rightarrow \infty$  as  $n \rightarrow \infty$ .

Consequently,  $I_n^- \sim (1-\epsilon)/(y(n)\delta(n)(\epsilon-\rho))$ ,  $I_n^+ \sim (1+\epsilon)/(y(n)\delta(n)(-\epsilon-\rho))$  and thus

$$\frac{(1-\epsilon)}{(\epsilon-\rho)}(1+o(1)) \leq I_n y(n) \delta(n) \leq \frac{(1+\epsilon)}{(-\epsilon-\rho)}(1+o(1)).$$

Letting  $\epsilon \rightarrow 0$  entails  $I_n y(n) \delta(n) \rightarrow -1/\rho$  as  $n \rightarrow \infty$  and thus

$$\Delta(n) \sim -\frac{1}{\gamma} \delta(n) \eta(e^{x(n)}).$$

Remarking that  $\log q(p_n) = \varphi(y(n)) \sim \gamma y(n) \sim \gamma x(n)/(1 - \delta_\infty)$  and taking account of (11) yield

$$\varepsilon_W(p_n; \alpha_n) = 1 - \exp\left(-\frac{\gamma}{1 - \delta_\infty} \Delta(n) x(n) (1 + o(1))\right), \quad (13)$$

when  $\delta(n) \rightarrow \delta_\infty \in [0, 1)$ . Finally, since  $|\eta| \in RV_\rho$ ,  $\rho < 0$ ,  $\Delta(n) x(n) \sim -\frac{1}{\gamma} \delta(n) x(n) \eta(e^{x(n)}) \rightarrow 0$  as  $n \rightarrow \infty$  and the conclusion follows.

(ii) From (11),  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  if and only if  $\Delta(n) \log q(p_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Besides, in view of Proposition 5(ii),  $\theta_1 = 0$  and  $\ell_1 = 0$  leading to  $K_2(t) \sim -K_1(t)$  as  $t \rightarrow \infty$  and  $\theta_2 = 0$  from Lemma 3(i). Proposition 2(i) thus yields

$$\Delta(n) \sim \frac{1}{2} \delta^2(n) K_2(x(n)) \text{ if } \delta(n) \rightarrow 0 \text{ as } n \rightarrow \infty, \quad (14)$$

$$\Delta(n) \sim \frac{\delta_\infty^2}{1 - \delta_\infty} K_2(x(n)) \text{ if } \delta(n) \rightarrow \delta_\infty \in (0, 1) \text{ as } n \rightarrow \infty, \quad (15)$$

since  $c(\delta_\infty, 0) = 1/(1 - \delta_\infty)$ . In view of the regular variations of  $H^{-1}$ , the key quantity is

$$\begin{aligned} K_1(x(n)) \log q(p_n) &\sim \frac{1}{\beta} \frac{\varphi(y(n))}{\varphi(x(n))} \sim \frac{1}{\beta} \frac{\log H^{-1}(y(n))}{\log H^{-1}(x(n))} \sim \frac{1}{\beta} \frac{\log y(n)}{\log x(n)} \\ &= \frac{1}{\beta} \left(1 - \frac{\log(1 - \delta(n))}{\log x(n)}\right) \rightarrow \frac{1}{\beta} \end{aligned} \quad (16)$$

as  $n \rightarrow \infty$  since  $\limsup \delta(n) < 1$ . As a consequence,  $\Delta(n) \log q(p_n) \rightarrow 0$  if and only if  $\delta(n) \rightarrow 0$  and the result follows.

(iii) The proof is similar to (ii). From (11),  $\varepsilon_W(p_n; \alpha_n) \rightarrow 0$  if and only if  $\Delta(n) \log q(p_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Besides, in view of Proposition 5(iii),  $\theta_1 = 0$  and  $\ell_1 = 1/\beta \neq 1$  leading to  $K_2(t) \rightarrow (1 - \beta)/\beta^2$  as  $t \rightarrow \infty$  and  $\theta_2 = 0$  from Lemma 3(iii). Proposition 2(ii) thus shows that  $\delta(n) \rightarrow 0$  is a necessary condition for  $\Delta(n) \rightarrow 0$  and, in that case,  $\Delta(n) \sim \frac{1-\beta}{2\beta^2} \delta^2(n)$ . ■

**Proof of Theorem 4.** (i) Equation (13) in the proof of Theorem 3(i) states that

$$\varepsilon_W(p_n; \alpha_n) = 1 - \exp\left(-\frac{\gamma}{1 - \delta_\infty} \Delta(n) x(n) (1 + o(1))\right).$$

with  $\Delta(n) x(n) \sim -\frac{1}{\gamma} \delta(n) x(n) \eta(e^{x(n)}) \rightarrow 0$  as  $n \rightarrow \infty$  and thus

$$\varepsilon_W(p_n; \alpha_n) \sim -\frac{1}{1 - \delta_\infty} x(n) \delta(n) \eta(e^{x(n)}) \sim -\frac{1}{1 - \delta_\infty} \log(1/\alpha_n) \delta(n) \eta(1/\alpha_n).$$

(ii) Equations (14)–(16) in the proof of Theorem 3(ii) yield

$$\Delta(n) \log q(p_n) \sim -\frac{1}{2\beta} \delta^2(n) \text{ if } \delta(n) \rightarrow 0 \text{ as } n \rightarrow \infty,$$

$$\Delta(n) \log q(p_n) \rightarrow -\frac{1}{\beta} \frac{\delta_\infty^2}{1 - \delta_\infty} \text{ if } \delta(n) \rightarrow \delta_\infty \in (0, 1) \text{ as } n \rightarrow \infty.$$

The conclusion follows from (11).

(iii) The result is a direct consequence of (11) and  $\Delta(n) \sim \frac{1-\beta}{2\beta^2} \delta^2(n)$ . ■

## References

- [1] C. Albert, A. Dutfoy, L. Gardes, and S. Girard. An extreme quantile estimator for the log-generalized Weibull-tail model. *Econometrics and Statistics*, to appear, 2019.
- [2] I. Alves, L. de Haan, and C. Neves. A test procedure for detecting super-heavy tails. *Journal of Statistical Planning and Inference*, 139(2):213–227, 2009.
- [3] J. Beirlant, M. Broniatowski, J. Teugels, and P. Vynckier. The mean residual life function at great age: Applications to tail estimation. *Journal of Statistical Planning and Inference*, 45(1-2):21–48, 1995.
- [4] J. Beirlant, J-P. Raoult, and R. Worms. On the relative approximation error of the generalized Pareto approximation for a high quantile. *Extremes*, 13:335–360, 2003.
- [5] N.H. Bingham, C.M. Goldie, and J.L. Teugels. *Regular Variation*, volume 27 of *Encyclopedia of Mathematics and its application*. Cambridge University Press, 1987.
- [6] L. Breiman, C.J. Stone, and C. Kooperberg. Robust confidence bounds for extreme upper quantiles. *Journal of Statistical Computation and Simulation*, 37:127–149, 1990.
- [7] J. Cohen. Convergence rates for the ultimate and penultimate approximations in extreme-value theory. *Advances in Applied Probability*, 14(4):833–854, 1982.
- [8] L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007.
- [9] C. de Valk. Approximation and estimation of very small probabilities of multivariate extreme events. *Extremes*, 19(4):687–717, 2016.
- [10] C. de Valk. Approximation of high quantiles from intermediate quantiles. *Extremes*, 19(4):661–686, 2016.
- [11] C. de Valk and J.-J. Cai. A high quantile estimator based on the log-generalized Weibull tail limit. *Econometrics and Statistics*, 6:107–128, 2018.
- [12] J. Diebolt, M.-A. El-Aroui, V. Durbec, and B. Villain. Estimation of extreme quantiles: Empirical tools for methods assessment and comparison. *International Journal of Reliability, Quality and Safety Engineering*, 7(01):75–94, 2000.
- [13] J. Diebolt and S. Girard. A note on the asymptotic normality of the ET method for extreme quantile estimation. *Statistics and Probability Letters*, 62(4):397–406, 2003.
- [14] J. El Methni, L. Gardes, S. Girard, and A. Guillou. Estimation of extreme quantiles from heavy and light tailed distributions. *Journal of Statistical Planning and Inference*, 142(10):2735–2747, 2012.

- [15] L. Gardes and S. Girard. Estimating extreme quantiles of Weibull tail-distributions. *Communication in Statistics - Theory and Methods*, 34:1065–1080, 2005.
- [16] L. Gardes and S. Girard. Estimation of the Weibull tail-coefficient with linear combination of upper order statistics. *Journal of Statistical Planning and Inference*, 138(5):1416–1427, 2008.
- [17] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.
- [18] L. Gardes, S. Girard, and A. Guillou. Weibull tail-distributions revisited: a new look at some tail estimators. *Journal of Statistical Planning and Inference*, 141(1):429–444, 2011.
- [19] Y. Goegebeur, J. Beirlant, and T. De Wet. Generalized kernel estimators for the Weibull-tail coefficient. *Communications in Statistics-Theory and Methods*, 39(20):3695–3716, 2010.
- [20] M.I. Gomes. Penultimate limiting forms in extreme value theory. *Annals of the Institute of Statistical Mathematics*, 36(1):71–85, 1984.
- [21] M.I. Gomes and L. de Haan. Approximation by penultimate extreme value distributions. *Extremes*, 2(1):71–85, 1999.
- [22] M.I. Gomes and D.D. Pestana. Nonstandard domains of attraction and rates of convergence. In M.L. Puri, J.P. Vilaplana, and W. Wertz, editors, *New Perspectives in Theoretical and Applied Statistics*, volume 141 of *Wiley series in probability and mathematical statistics*, pages 467–477. Wiley, 1987.
- [23] B. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975.
- [24] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3:119–131, 1975.
- [25] R.L. Smith. Estimating tails of probability distributions. *The Annals of Statistics*, 15(3):1174–1207, 1987.
- [26] I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.
- [27] R. Worms. Penultimate approximation for the distribution of the excesses. *ESAIM: Probability and Statistics*, 6:21–31, 2002.

**Acknowledgments.** This work benefited from the support of the Chair Stress Test, Risk Management and Financial Steering, led by the French Ecole polytechnique and its Foundation and sponsored by BNP Paribas. This work was also supported by the French National Research Agency in the framework of the Investissements d’Avenir program (ANR-15-IDEX-02). The authors thank two anonymous referees and the associate editor for their helpful suggestions and comments which contributed to an improved presentation of the results of this paper.

## Appendix: Auxiliary results

We begin with an elementary result whose proof is straightforward.

**Lemma 1** For all  $(a, b, t) \in \mathbb{R}_+^3$ , let  $\Psi_a(t; b) = \int_0^b u^a \exp(-tu) du$ .

(i)  $\Psi_a(\cdot; b)$  is continuous, non-increasing on  $\mathbb{R}_+$ ,  $\Psi_a(0; b) = \frac{b^{a+1}}{a+1}$  and  $\Psi_a(t; b) \rightarrow 0$  as  $t \rightarrow \infty$ .

(ii)  $\Psi_1(t, b) \sim 1/t^2$  and  $\Psi_2(t, b) \sim 1/t^3$  as  $t \rightarrow \infty$ .

The next lemma establishes some links between the regular variation properties of  $\varphi$  and  $K_1$ .

**Lemma 2** Assume **(A2)** and **(A4)** hold.

(i) If  $\varphi \in RV_{1/\beta}$ ,  $\beta > 0$  then  $K_1 \in RV_0$  and  $\ell_1 = 1/\beta$ .

(ii) Let  $\beta > 0$ . Then,  $\log \varphi \in RV_\beta$  if and only if  $K_1 \in RV_\beta$ .

(iii) Let  $\varphi_\infty := \lim_{t \rightarrow \infty} \varphi(t) \in (0, \infty]$  and  $\theta_1 < 0$ . Then,  $\varphi_\infty < \infty$  and  $1 - \varphi/\varphi_\infty \in RV_{\theta_1}$  if and only if  $K_1 \in RV_{\theta_1}$ .

(iv) If  $\exp \varphi(\log(\cdot)) \in RV_\gamma$ ,  $\gamma > 0$  then  $K_1 \in RV_0$  and  $\ell_1 = 1$ .

(v) If  $\exp \varphi \in RV_{1/\beta}$ ,  $\beta > 0$  then  $K_1 \in RV_0$  and  $\ell_1 = 0$ .

**Proof.** Recall that  $K_1(x) = x(\log \varphi)'(x)$ .

(i) If  $\varphi \in RV_{1/\beta}$ ,  $\beta > 0$  then the monotone density theorem [5, Theorem 1.7.2] yields  $\varphi(x) \sim \beta x \varphi'(x)$  or equivalently  $K_1(x) \rightarrow 1/\beta$  as  $x \rightarrow \infty$ . It follows that  $\ell_1 = 1/\beta$  and  $K_1 \in RV_0$ .

(ii,  $\implies$ ) Let us assume that  $\log \varphi \in RV_\beta$ ,  $\beta > 0$ . Then, the monotone density theorem implies  $(\log \varphi)' \in RV_{\beta-1}$  i.e.  $K_1 \in RV_\beta$ .

(ii,  $\impliedby$ ) Conversely, assume  $K_1 \in RV_\beta$ ,  $\beta > 0$ . Then, necessarily  $\ell_1 = \infty$ . From [5, Theorem 1.5.8], we have for all  $x_0$  sufficiently large,

$$\log \varphi(x) - \log \varphi(x_0) = \int_{x_0}^x (\log \varphi(t))' dt = \int_{x_0}^x \frac{K_1(t)}{t} dt \sim \frac{1}{\beta} K_1(x), \quad (17)$$

as  $x \rightarrow \infty$ . It is thus clear that  $\log \varphi \in RV_\beta$ .

(iii,  $\implies$ ) Let us assume that  $\varphi_\infty < \infty$ ,  $\varphi(\cdot) = \varphi_\infty(1 - h(\cdot))$  where  $h \in RV_{\theta_1}$ ,  $\theta_1 < 0$ . Straightforward calculations and the monotone density theorem lead to

$$\log \varphi(x) = \log \varphi_\infty + \log(1 - h(x)) \quad \text{and} \quad K_1(x) = \frac{xh'(x)}{h(x) - 1} \sim -\theta_1 h(x).$$

As a conclusion,  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 0$ .

(iii,  $\impliedby$ ) Conversely, assume  $K_1 \in RV_{\theta_1}$ ,  $\theta_1 < 0$ . Thus  $(\log \varphi)' \in RV_{\theta_1-1}$  and [5, Theorem 1.5.8] yields first, for all  $x$  sufficiently large,

$$\log \varphi_\infty - \log \varphi(x) = \int_x^\infty (\log \varphi)'(t) dt < \infty \quad (18)$$

and thus  $\varphi_\infty < \infty$ . Second, one also has

$$\frac{K_1(x)}{\int_x^\infty (\log \varphi)'(t) dt} \rightarrow -\theta_1 \quad (19)$$

as  $x \rightarrow \infty$ . Combining the two above results (18), (19) yields  $K_1(x)/(\log \varphi_\infty - \log \varphi(x)) \rightarrow -\theta_1$  as  $x \rightarrow \infty$  and consequently

$$\varphi(x) = \varphi_\infty \exp\left(\frac{1}{\theta_1} K_1(x)(1 + o(1))\right) = \varphi_\infty \left(1 + \frac{1}{\theta_1} K_1(x)(1 + o(1))\right)$$

since  $K_1(x) \rightarrow 0$  as  $x \rightarrow \infty$ .

(iv) Assume  $\exp \varphi(\log(\cdot)) \in RV_\gamma$ ,  $\gamma > 0$ . The monotone density theorem implies  $\varphi'(x) \rightarrow \gamma$  as  $x \rightarrow \infty$ . Thus  $K_1(x) \rightarrow 1$  as  $x \rightarrow \infty$  and therefore  $K_1 \in RV_0$ .

(v) Assume  $\exp \varphi \in RV_{1/\beta}$ ,  $\beta > 0$ . From the monotone density theorem,  $x\varphi'(x) \rightarrow 1/\beta$  as  $x \rightarrow \infty$ . Thus  $\varphi \in RV_0$  implies  $K_1 \in RV_0$  and  $\varphi(x) \rightarrow \infty$  as  $x \rightarrow \infty$  yields  $\ell_1 = 0$ .  $\blacksquare$

Lemma 3 shows that  $K_1 \in RV_{\theta_1}$  implies  $|K_2| \in RV_{\theta_2}$  when  $\ell_1 \neq 1$ . In the situation where  $\ell_1 = 1$ , the logistic distribution defined by  $H^{-1}(x) = \log(\exp(x) - 1)$ ,  $x > 0$  is a case where  $-K_2(x) \sim x \exp(-x)$  is not regularly varying as  $x \rightarrow \infty$ .

**Lemma 3** *Assume (A2)–(A4) hold.*

(i) *If  $\ell_1 = 0$  then  $\theta_1 \leq 0$ ,  $\ell_2 = 0$ ,  $-K_2 \in RV_{\theta_1}$  and  $K_2(x) \sim (\theta_1 - 1)K_1(x)$  as  $x \rightarrow \infty$ .*

(ii) *If  $\ell_1 = 1$  then  $\theta_1 = 0$  and  $\ell_2 = 0$ .*

(iii) *If  $0 < \ell_1 < \infty$  and  $\ell_1 \neq 1$  then  $\theta_1 = 0$ ,  $\ell_2 = \ell_1(\ell_1 - 1) \neq 0$  and  $|K_2| \in RV_0$ .*

(iv) *If  $\ell_1 = \infty$  then  $\theta_1 \geq 0$ ,  $\ell_2 = \infty$ ,  $K_2 \in RV_{2\theta_1}$  and  $K_2(x) \sim K_1^2(x)$  as  $x \rightarrow \infty$ .*

**Proof.** The proof relies on the following four facts: First, for all  $x \in \mathbb{R}$ ,

$$\frac{K_2(x)}{K_1^2(x)} = 1 + \frac{1}{K_1(x)} \left( \frac{xK_1'(x)}{K_1(x)} - 1 \right). \quad (20)$$

Second,  $xK_1'(x)/K_1(x) \rightarrow \theta_1$  as  $x \rightarrow \infty$  from the monotone density theorem [5, Theorem 1.7.2]. Third, it straightforwardly follows that  $\ell_2 = \ell_1(\ell_1 + \theta_1 - 1)$ . Finally, for all positive function  $K$ ,  $K(x) \rightarrow c > 0$  as  $x \rightarrow \infty$  implies  $K \in RV_0$ .  $\blacksquare$

The next lemma establishes the links between  $\delta$  and  $\Delta$  through  $K_1$  and  $K_2$ .

**Lemma 4** *Suppose (A1)–(A4) hold.*

(i) *Suppose  $\ell_1 < \infty$ . Then, as  $t \rightarrow \infty$ ,*

$$\Delta(t) \sim \delta^2(t) \int_0^1 K_2(y(t)(1 - \delta(t)u))(1 - \delta(t)u)^{\ell_1 - 2} u du.$$

(ii) *Suppose  $\ell_1 \neq 1$ . Then, for all  $t > 0$ :*

$$\begin{aligned} |\Delta(t)| &\leq \max(|K_2(y(t))|, |K_2(x(t))|) \frac{\delta^2(t)}{(1 - \delta(t))^2} \Phi(\delta(t)K_1(y(t))(1 + o(1))) \text{ and} \\ |\Delta(t)| &\geq \min(|K_2(y(t))|, |K_2(x(t))|) \delta^2(t) \Phi\left(\delta(t)K_1(y(t))(1 - \delta(t))^{\theta_1 - 1}(1 + o(1))\right), \end{aligned}$$

where  $\Phi(s) = \Psi_1(s; 1) = \int_0^1 u \exp(-us) du$  for all  $s \geq 0$ .

**Proof.** (i) Under **(A2)**, a second order Taylor expansion with integral remainder yields

$$\begin{aligned}\Delta(t) &= \int_{x(t)}^{y(t)} \frac{K_2(s)}{s^2} \frac{\varphi(s)}{\varphi(y(t))} (y(t) - s) ds \\ &= \delta^2(t) \int_0^1 \frac{K_2(y(t)(1 - \delta(t)u))}{(1 - \delta(t)u)^2} \frac{\varphi(y(t)(1 - \delta(t)u))}{\varphi(y(t))} u du,\end{aligned}$$

thanks to the change of variable  $u = (y(t) - s)/(y(t) - x(t))$ . Besides,

$$\frac{\varphi(y(t)(1 - \delta(t)u))}{\varphi(y(t))} = \exp\left(\int_{y(t)}^{y(t)(1 - \delta(t)u)} (\log \varphi(s))' ds\right) = \exp\left(K_1(y(t)) \int_1^{1 - \delta(t)u} \frac{K_1(vy(t))}{K_1(y(t))} \frac{dv}{v}\right).$$

Since  $1 - \delta(t)u \in [1 - \delta(t), 1]$ , **(A3)** yields  $K_1(vy(t))/K_1(y(t)) \rightarrow v^{\theta_1}$  uniformly locally as  $t \rightarrow \infty$  and consequently  $y(t) \rightarrow \infty$ . Condition **(A1)** then leads to

$$\frac{\varphi(y(t)(1 - \delta(t)u))}{\varphi(y(t))} = \exp(K_1(y(t))L_{\theta_1}(1 - \delta(t)u)(1 + o(1))),$$

where  $L_{\theta_1}(x) = \int_1^x u^{\theta_1 - 1} du$  for all  $x \in \mathbb{R}$ . It thus follows that

$$\Delta(t) = \delta^2(t) \int_0^1 \frac{K_2(y(t)(1 - \delta(t)u))}{(1 - \delta(t)u)^2} \exp(K_1(y(t))L_{\theta_1}(1 - \delta(t)u)(1 + o(1))) u du. \quad (21)$$

If  $\ell_1 = 0$  then

$$\Delta(t) \sim \delta^2(t) \int_0^1 K_2(y(t)(1 - \delta(t)u))(1 - \delta(t)u)^{-2} u du.$$

In the situation where  $0 < \ell_1 < \infty$ , Lemma 3(iii) entails  $\theta_1 = 0$  and (21) yields

$$\Delta(t) \sim \delta^2(t) \int_0^1 K_2(y(t)(1 - \delta(t)u))(1 - \delta(t)u)^{\ell_1 - 2 + o(1)} u du,$$

the first part of the result is proved.

(ii) From Lemma 3, when  $\ell_1 \neq 1$  the sign of  $K_2$  is ultimately constant so that (21) entails

$$|\Delta(t)| = \delta^2(t) \int_0^1 \frac{|K_2(y(t)(1 - \delta(t)u))|}{(1 - \delta(t)u)^2} \exp(K_1(y(t))L_{\theta_1}(1 - \delta(t)u)(1 + o(1))) u du.$$

Let us remark that, for all  $u \in [0, 1]$  and  $\theta_1 \leq 1$ , one has  $1 - \delta(t) \leq 1 - \delta(t)u \leq 1$  and

$$-(1 - \delta(t))^{\theta_1 - 1} \delta(t)u \leq L_{\theta_1}(1 - \delta(t)u) \leq -\delta(t)u.$$

It is thus clear that

$$\begin{aligned}|\Delta(t)| &\leq \frac{\delta^2(t)}{(1 - \delta(t))^2} \int_0^1 |K_2(y(t)(1 - \delta(t)u))| \exp(-\delta(t)K_1(y(t))u(1 + o(1))) u du, \\ |\Delta(t)| &\geq \delta^2(t) \int_0^1 |K_2(y(t)(1 - \delta(t)u))| \exp\left(-\delta(t)K_1(y(t))(1 - \delta(t))^{\theta_1 - 1} u(1 + o(1))\right) u du.\end{aligned}$$

Besides, Lemma 3 entails that  $|K_2|$  is regularly varying when  $\ell_1 \neq 1$ . Therefore,  $|K_2|$  is ultimately monotone and it follows that, for  $t$  large enough,  $m(t) \leq |K_2(y(t)(1 - \delta(t)u))| \leq M(t)$ , where  $m(t) := \min(|K_2(y(t))|, |K_2(x(t))|)$  and  $M(t) := \max(|K_2(y(t))|, |K_2(x(t))|)$ , leading to

$$\begin{aligned}|\Delta(t)| &\leq M(t) \frac{\delta^2(t)}{(1 - \delta(t))^2} \int_0^1 u \exp(-\delta(t)K_1(y(t))u(1 + o(1))) du \text{ and} \\ |\Delta(t)| &\geq m(t) \delta^2(t) \int_0^1 u \exp\left(-\delta(t)K_1(y(t))(1 - \delta(t))^{\theta_1 - 1} u(1 + o(1))\right) du.\end{aligned}$$

Introducing for all  $s \geq 0$ ,  $\Phi(s) = \int_0^1 u \exp(-us) du$ , the above bounds can be rewritten as

$$\begin{aligned} |\Delta(t)| &\leq M(t) \frac{\delta^2(t)}{(1-\delta(t))^2} \Phi(\delta(t)K_1(y(t))(1+o(1))) \text{ and} \\ |\Delta(t)| &\geq m(t)\delta^2(t)\Phi\left(\delta(t)K_1(y(t))(1-\delta(t))^{\theta_1-1}(1+o(1))\right), \end{aligned}$$

which concludes the proof.  $\blacksquare$

As a consequence of the above result, a sufficient condition as well as a necessary condition can be established under **(A1)** such that  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ .

**Lemma 5** *Suppose (A1)–(A4) hold.*

(i) *If  $\delta^2(t) \max(|K_2(y(t))|, |K_2(x(t))|) \rightarrow 0$  then  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$ .*

(ii) *If  $\Delta(t) \rightarrow 0$  then  $\delta^2(t) \min(|K_2(y(t))|, |K_2(x(t))|) \rightarrow 0$  as  $t \rightarrow \infty$ .*

**Proof.** Let us first note that when  $\ell_1 = 1$  then  $\ell_2 = 0$  from Lemma 3(ii). It is thus clear in view of Lemma 4(i) that  $\Delta(t) \rightarrow 0$  as  $t \rightarrow \infty$  under **(A1)**. In the following, we thus focus on the case where  $\ell_1 \neq 1$ . Lemma 3 entails that  $|K_2|$  is regularly varying since  $\ell_1 \neq 1$ . Therefore,  $|K_2|$  is ultimately monotone. Let us focus on the situation where  $|K_2|$  is ultimately non decreasing and introduce  $A(t) = \delta(t)K_1(y(t))$  for all  $t > 0$ .

(i) Assume that  $\delta^2(t)|K_2(y(t))| \rightarrow 0$  as  $t \rightarrow \infty$ . From Lemma 1(i),  $0 \leq \Phi(s) \leq 1/2$  for all  $s \geq 0$  and thus Lemma 4(ii) entails

$$|\Delta(t)| \leq \frac{\delta^2(t)|K_2(y(t))|}{2(1-\delta(t))^2} \rightarrow 0 \quad (22)$$

as  $t \rightarrow \infty$  in view of **(A1)**.

(ii) From Lemma 4(ii), one has

$$|\Delta(t)| \geq |K_2(x(t))|\delta^2(t)\Phi\left(A(t)(1-\delta(t))^{\theta_1-1}(1+o(1))\right) \geq |K_2(x(t))|\delta^2(t)\Phi(cA(t))$$

for  $t$  large enough and some  $c > 0$  since  $\Phi$  is non-increasing, see Lemma 1(i). For all  $s \geq 0$ , let  $\psi(s) = \int_0^s x \exp(-x) dx = s^2\Phi(s)$ . Consider  $s_0 \geq c(3-2\theta_1)$  with  $\theta_1 \leq 1$  and remark that  $\Phi(s) \geq \Phi(s_0)$  for all  $0 \leq s \leq s_0$  and  $\psi(s) \geq \psi(s_0)$  for all  $s \geq s_0$ . As a consequence, for all  $s > 0$ ,

$$\Phi(s) \geq \frac{\psi(s_0)}{s_0^2} \mathbb{I}\{s \leq s_0\} + \frac{\psi(s_0)}{s^2} \mathbb{I}\{s \geq s_0\},$$

and thus

$$\begin{aligned} |\Delta(t)| &\geq \frac{\psi(s_0)}{s_0^2} |K_2(x(t))|\delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} + \frac{\psi(s_0)}{c^2} \frac{|K_2(x(t))|}{K_1^2(y(t))} \mathbb{I}\{A(t) \geq s_0/c\} \\ &\geq \frac{\psi(s_0)}{s_0^2} |K_2(x(t))|\delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} + \frac{\psi(s_0)}{c^2} \frac{|K_2(x(t))|}{K_1^2(x(t))} \frac{K_1^2(x(t))}{K_1^2(y(t))} \mathbb{I}\{A(t) \geq s_0/c\}. \end{aligned} \quad (23)$$

Since  $K_1 \in RV_{\theta_1}$ ,  $K_1(x(t))/K_1(y(t)) \sim (1-\delta(t))^{\theta_1} \geq c' > 0$  as  $t \rightarrow \infty$  in view of **(A1)** and

$$|\Delta(t)| \geq \frac{\psi(s_0)}{s_0^2} |K_2(x(t))|\delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} + \psi(s_0) \left(\frac{c'}{c}\right)^2 \frac{|K_2(x(t))|}{K_1^2(x(t))} \mathbb{I}\{A(t) \geq s_0/c\}.$$



Remarking that (20) in the proof of Lemma 3 implies that, for  $t$  large enough,

$$\frac{K_2(x(t))}{K_1^2(x(t))} = 1 + \frac{1}{K_1(x(t))} \left( \frac{x(t)K_1'(x(t))}{K_1(x(t))} - 1 \right) = 1 + \frac{\delta(t)}{A(t)} (\theta_1 - 1 + o(1))$$

which yields when  $A(t) \geq s_0/c$ ,

$$\left| \frac{K_2(x(t))}{K_1^2(x(t))} - 1 \right| \leq \frac{c\delta(t)}{s_0} |\theta_1 - 1 + o(1)| \leq \frac{c}{s_0} (3/2 - \theta_1) \leq \frac{1}{2}.$$

It thus follows that

$$\frac{|K_2(x(t))|}{K_1^2(x(t))} \mathbb{I}\{A(t) \geq s_0/c\} \geq \frac{1}{2} \mathbb{I}\{A(t) \geq s_0/c\}$$

and therefore,

$$|\Delta(t)| \geq \frac{\psi(s_0)}{s_0^2} |K_2(x(t))| \delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} + \frac{\psi(s_0)}{2} \left( \frac{c'}{c} \right)^2 \mathbb{I}\{A(t) \geq s_0/c\}.$$

As a conclusion,  $|\Delta(t)| \rightarrow 0$  implies  $|K_2(x(t))| \delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} \rightarrow 0$  and  $\mathbb{I}\{A(t) \geq s_0/c\} \rightarrow 0$  as  $t \rightarrow \infty$ . Consequently,  $A(t) \leq s_0/c$  eventually and  $\delta^2(t) K_2(x(t)) \rightarrow 0$  as  $t \rightarrow \infty$ .

Let us now consider the situation where  $|K_2|$  is ultimately non increasing.

(i) The proof is similar, the upper bound (22) is replaced by

$$|\Delta(t)| \leq \frac{\delta^2(t) |K_2(x(t))|}{2(1 - \delta(t))^2}.$$

(ii) The lower bound (23) is replaced by

$$|\Delta(t)| \geq \frac{\psi(s_0)}{s_0^2} |K_2(y(t))| \delta^2(t) \mathbb{I}\{A(t) \leq s_0/c\} + \frac{\psi(s_0)}{c^2} \frac{|K_2(y(t))|}{K_1^2(y(t))} \mathbb{I}\{A(t) \geq s_0/c\}$$

and the end of the proof is similar. ■

	$\bar{F}(x)$	$\theta_1$	$\theta_2$	$K_1(x)$	$K_2(x)$	$\ell_1$
<b>MDA<sub>1</sub>(Gumbel)</b>						
Finite endpoint ( $\beta > 0$ )	$\exp\left(-(-\log x)^{-\beta}\right)$ $x \in (0, 1)$	$-1/\beta$	$-1/\beta$	$\frac{1}{\beta}x^{-1/\beta}$	$-\frac{1+\beta}{\beta^2}x^{-1/\beta}(1+o(1))$	0
Gamma ( $a > 0$ )	$\frac{1}{\Gamma(a)} \int_x^\infty t^{a-1}e^{-t}dt$ $x \geq 0$	0	-1	$1+o(1)$	$\frac{1-a}{x}(1+o(1))$	1
<b>MDA<sub>2</sub>(Gumbel)</b>						
Weibull ( $\beta \neq 1$ )	$\exp(-x^\beta)$ $x \geq 0$	0	0	$\frac{1}{\beta}$	$\frac{1-\beta}{\beta^2}$	$1/\beta$
Gaussian	$\frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left(-\frac{t^2}{2}\right) dt$	0	0	$\frac{1}{2}+o(1)$	$-\frac{1}{4}+o(1)$	$1/2$
<b>MDA<sub>3</sub>(Gumbel)</b>						
Log-Weibull ( $\beta > 1$ )	$\exp(-(\log x)^\beta)$ $x \geq 1$	$1/\beta$	$2/\beta$	$\frac{1}{\beta}x^{1/\beta}$	$\frac{1}{\beta^2}x^{2/\beta}(1+o(1))$	$+\infty$
Lognormal ( $\sigma > 0$ )	$\frac{1}{\sigma\sqrt{2\pi}} \int_x^\infty \frac{1}{t} \exp\left(-\frac{(\log t)^2}{2\sigma^2}\right) dt$ $x \geq 0$	$1/2$	1	$\frac{\sigma}{\sqrt{2}}x^{1/2}(1+o(1))$	$\frac{\sigma^2}{2}x(1+o(1))$	$+\infty$

Table 1: Examples of distributions in MDA(Gumbel), ET framework:  $\varphi = H^{-1}$ .

Distribution	$\varepsilon_{\text{ET}}(p_n; \alpha_n)$	$\varepsilon_{\text{W}}(p_n; \alpha_n)$
<b>MDA<sub>1</sub>(Gumbel)</b>		
Finite endpoint( $\beta > 0$ )	$-\frac{2(1+\beta)}{\beta^2} \frac{(\log \log n)^2}{(\log n)^{2+1/\beta}}$	
Gamma( $a > 0$ )	$2(1-a) \frac{(\log \log n)^2}{(\log n)^3}$	$-2 \frac{(\log \log n)^2}{(\log n)^2}$
<b>MDA<sub>2</sub>(Gumbel)</b>		
Weibull( $\beta \neq 1$ )	$\frac{2(1-\beta)}{\beta^2} \frac{(\log \log n)^2}{(\log n)^2}$	$-\frac{2}{\beta} \frac{(\log \log n)^2}{(\log n)^2}$
Gaussian	$-\frac{1}{2} \frac{(\log \log n)^2}{(\log n)^2}$	$-\frac{(\log \log n)^2}{(\log n)^2}$
<b>MDA<sub>3</sub>(Gumbel)</b>		
Log-Weibull( $\beta > 1$ )	$\frac{2}{\beta^2} \frac{(\log \log n)^2}{(\log n)^{2-2/\beta}}$	$\frac{2(1-\beta)}{\beta^2} \frac{(\log \log n)^2}{(\log n)^{2-1/\beta}}$
Lognormal	$\sigma^2 \frac{(\log \log n)^2}{\log n}$	$-\frac{\sqrt{2\sigma^2}}{2} \frac{(\log \log n)^2}{(\log n)^{3/2}}$

Table 2: First order approximations of  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  and  $\varepsilon_{\text{W}}(p_n; \alpha_n)$  with  $p_n = 1/(n \log n)$  and  $\alpha_n = (\log n)/n$  associated with the distributions described in Table 1.

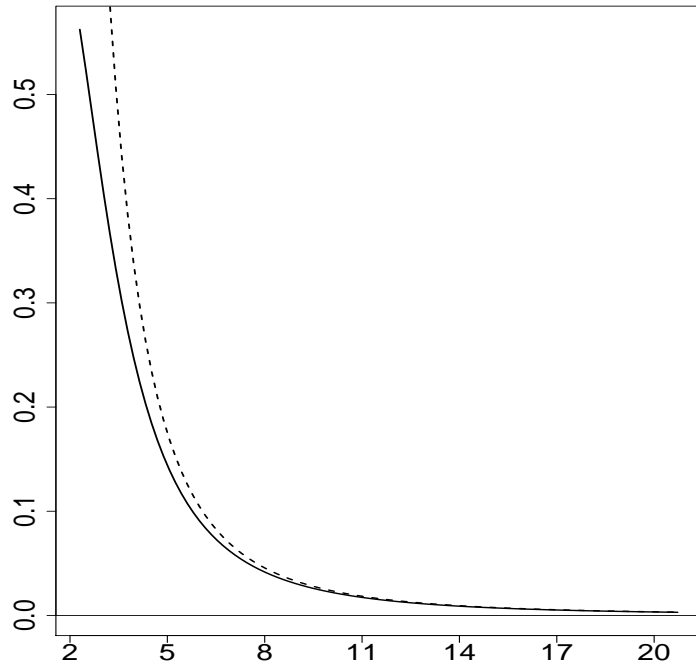
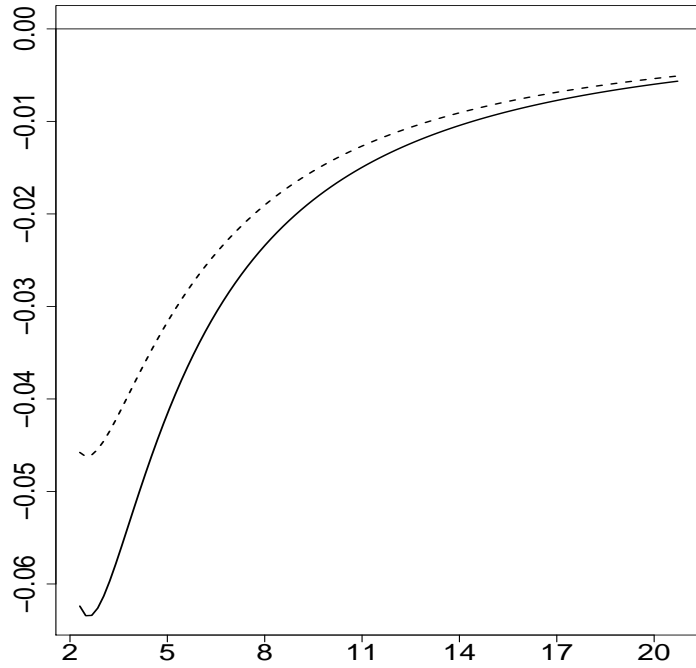


Figure 1: Extrapolation in  $MDA_1(\text{Gumbel})$ . Vertically: Relative extrapolation error  $\varepsilon_{ET}(p_n; \alpha_n)$  (solid line) and its first order approximation  $\frac{1}{2}\delta^2(n)K_2(\log n)$  (dashed line) provided by Theorem 2(i)-(a). Horizontally:  $\log n$ . Top: Finite endpoint( $\beta = 5$ ) distribution, bottom: Gamma( $a = 0.1$ ) distribution.

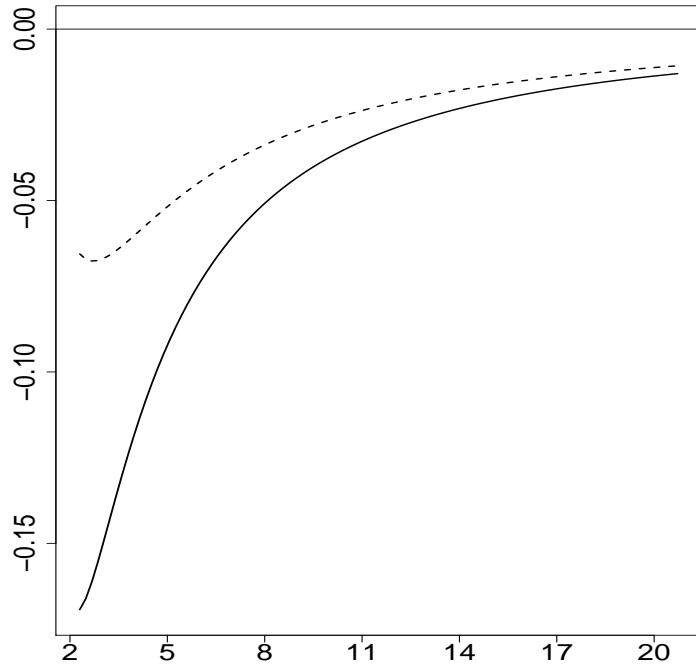
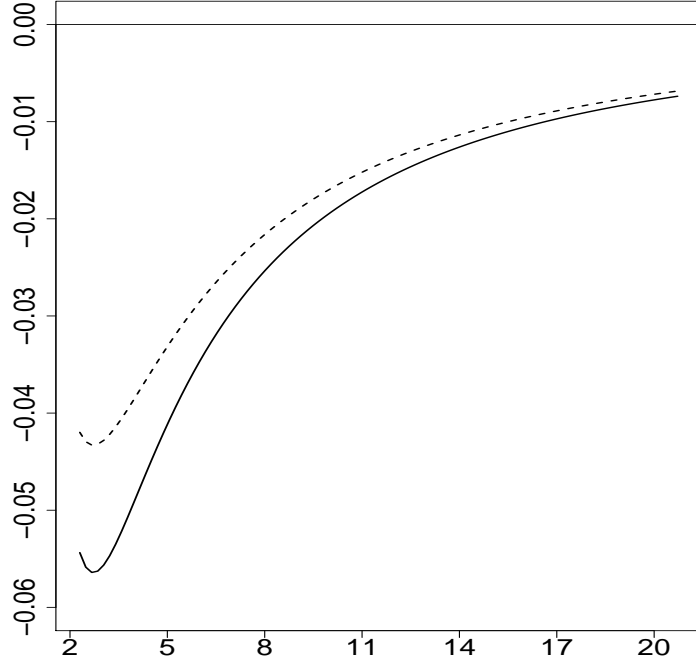


Figure 2: Extrapolation in  $\text{MDA}_2(\text{Gumbel})$ . Vertically: Relative extrapolation error  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  (solid line) and its first order approximation  $\frac{\ell_1(\ell_1-1)}{2}\delta^2(n)$  (dashed line) provided by Theorem 2(ii)-(a). Horizontally:  $\log n$ . Top: Weibull( $\beta = 5$ ) distribution, bottom: Gaussian distribution.

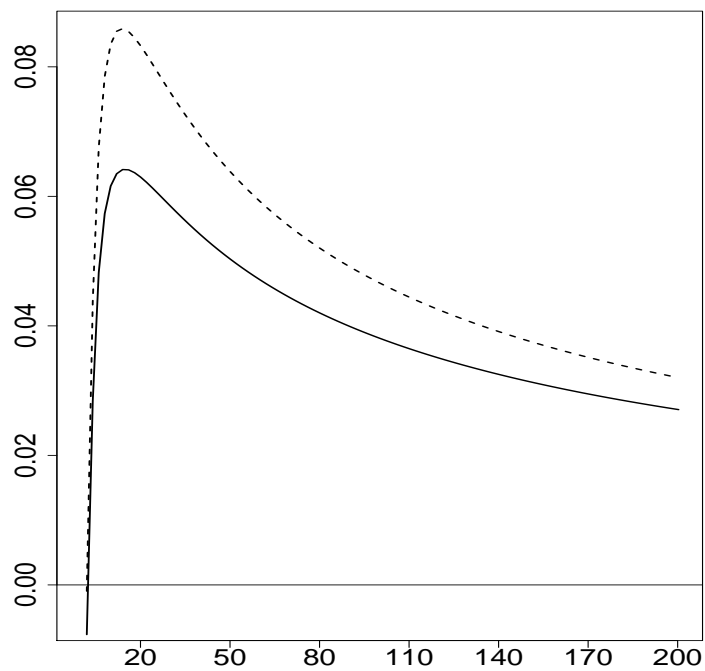
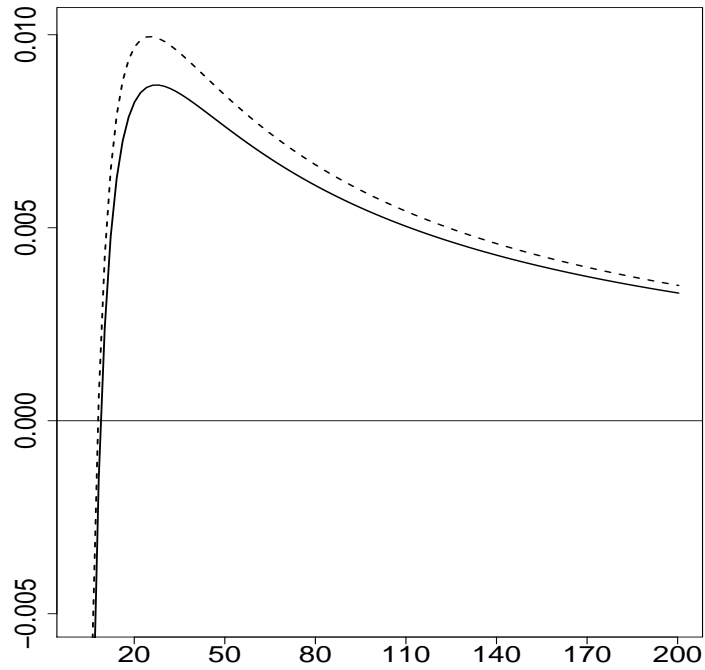


Figure 3: Extrapolation in  $\text{MDA}_3(\text{Gumbel})$ . Vertically: Relative extrapolation error  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  (solid line) and its first order approximation  $\frac{1}{2}\delta^2(n)K_2(\log n)$  (dashed line) provided by Theorem 2(iii)-(a). Horizontally:  $\log n$ . Top: log-Weibull( $\beta = 3$ ) distribution, bottom: lognormal( $\sigma = 0.5$ ) distribution.

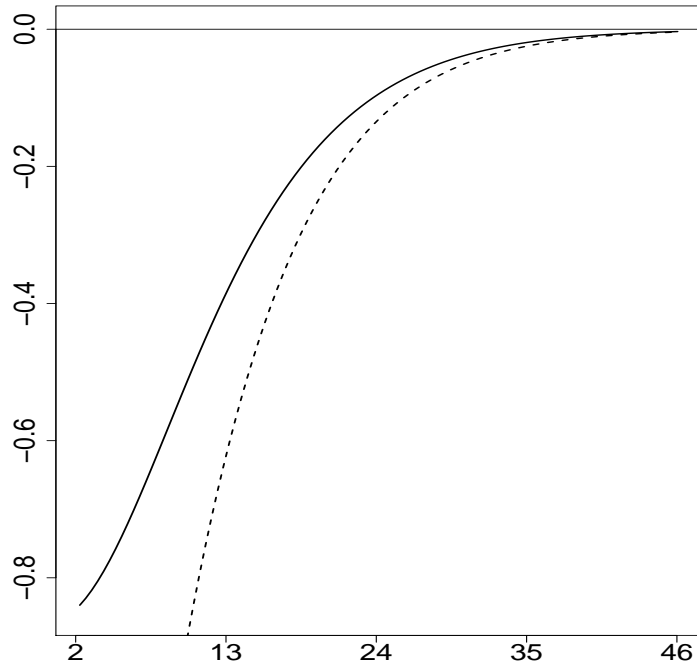
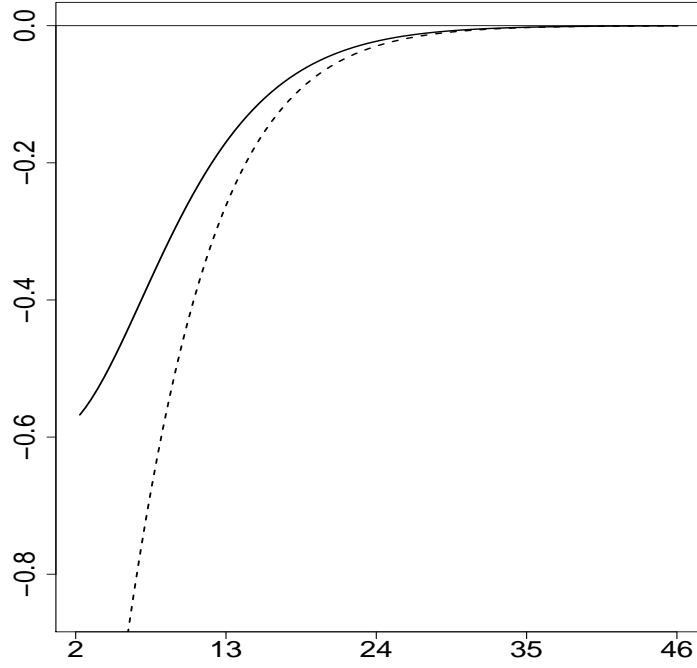


Figure 4: Extrapolation in MDA(Fréchet). Vertically: Relative extrapolation error  $\varepsilon_{\mathbb{W}}(p_n; \alpha_n)$  (solid line) and its first order approximation  $-\frac{1}{1-\delta_\infty}\delta(n)\log(1/\alpha_n)\eta(1/\alpha_n)$  (dashed line) provided by Theorem 4(i). Horizontally:  $\log n$ . Top: Burr( $\rho = -1/3$ ) distribution (see Section 5), bottom: Burr( $\rho = -1/4$ ) distribution.

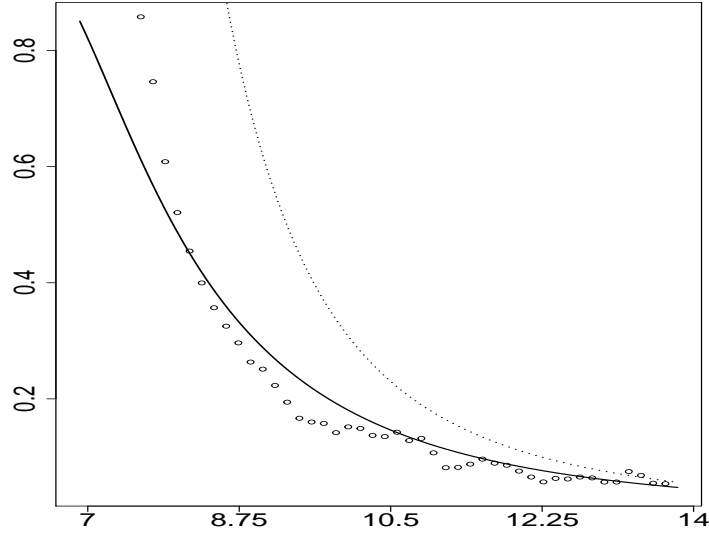


Figure 5: Relative extrapolation error for a Gamma( $a = 0.1$ ) distribution. Vertically:  $\varepsilon_{\text{ET}}(p_n; \alpha_n)$  (solid line), first order approximation  $\frac{1}{2}\delta^2(n)K_2(\log(n/k_n))$  (dashed line) and estimation  $\hat{\varepsilon}_{\text{ET}}(p_n; \alpha_n)$  built on [11, Equations (19,20)] and averaged over  $N = 100$  replications (dotted line). Horizontally:  $\log n$ .

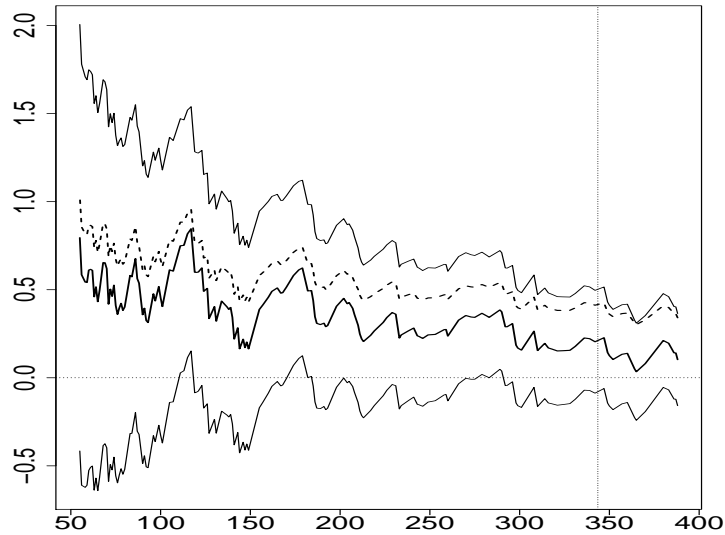


Figure 6: Estimations of  $\theta_1$  as functions of  $k_n$  on the wind speeds dataset. Full line:  $\hat{\theta}_1$  defined in [1, Equation (7)] together with its 95% asymptotic confidence interval. Dashed line:  $\hat{\theta}_1$  proposed by [11, Equation (19)]. Vertical dotted line:  $k_n = 40 \log n \approx 344$ . Horizontal dotted line :  $\theta_1 = 0$ .



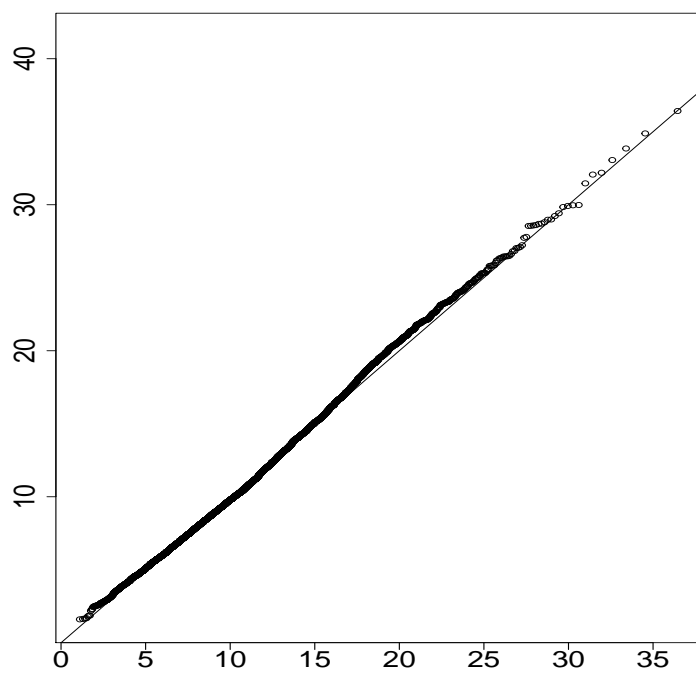


Figure 7: Quantile-quantile plot.  $\circ$ : empirical quantiles of the wind speeds dataset (vertically) vs quantiles from a Gamma distribution (horizontally). Continuous line: diagonal line.