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

Clément Albert, Anne Dutfoy, Stéphane Girard

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

Clément Albert[(1)], Anne Dutfoy[(2)] and Stéphane Girard[(1, *)]

[(1)] *Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France*

[(2)] *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. Necessary and sufficient 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 constrast, 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 and their accuracy is illustrated numerically.

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 [5] while numerical aspects are investigated in [11]. 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 $p_n \leq 1/n$. 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 us introduce a deterministic threshold u_n such that $\bar{F}(u_n) = \alpha_n$ or equivalently $u_n = q(\alpha_n)$ with $\alpha_n \to 0$ and $n\alpha_n > 1$ as $n \to \infty$. 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

*Corresponding author, `Stephane.Girard@inria.fr`

by $\bar{F}_{u_n}(x) = \bar{F}(u_n + x)/\bar{F}(u_n)$. Pickands theorem [14, 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

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

where σ_n and γ_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 σ_n. Thus, approximation (1) can be rewritten as

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

and the associated estimator [5] is

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

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 $X_{1,n} \le \ldots \le X_{n,n}$ denote the order statistics associated with $X_1, \ldots, X_n$. The error $(q(p_n) - \hat{q}_{\mathrm{ET}}(p_n; \alpha_n))$ can be expanded as a sum of two terms:

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

the first one being a random estimation error

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

and the second one being a deterministic extrapolation error

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

The asymptotic behavior of the estimation error (3) is driven by the asymptotic distributions of $\hat{q}(\alpha_n)$ and $\hat{\sigma}_n$ which can be found for instance in [12] or in [7], Theorem 2.4.1 and Theorem 3.4.2 respectively.

In this paper, we focus on the asymptotic behavior of the extrapolation error (4). Indeed, in view of (2), it appears that 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 necessary and sufficient conditions on the pair (p_n, α_n) such that the relative extrapolation error

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

tends to zero as $n \to \infty$. 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 [6, 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 a the survival distribution function of a GPD is studied in [3].

Here, similarly to [3], 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:

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

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 φ. 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).

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. As a conclusion, some numerical illustrations are presented in Section 5. Proofs are postponed to Section 6 and auxiliary results can be found in the Appendix.

2 Theoretical framework

The following functions are introduced.

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

(A2) φ is a twice differentiable, increasing function.

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

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

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 $\Delta(t) \to 0$ as $t \to \infty$. The following two functions are of the utmost importance in this study:

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

The study of Δ 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

$$\frac{K_1(ts)}{K_1(s)} \to t^{\theta_1} \text{ as } s \to \infty \text{ for all } t > 0.$$

We refer to [4] 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 $s \to \infty$ exist in $\bar{\mathbb{R}}$. The following notations are thus introduced:

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

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

Proposition 1 (Role of ℓ_1 for $\Delta \to 0$) *Suppose* **(A1)**–**(A4)** *hold.*

(i) If $\ell_1 \in \{0,1\}$ then $\ell_2 = 0$ and $\Delta(t) \to 0$ as $t \to \infty$.

(ii) If $\ell_1 \in (0,\infty) \setminus \{1\}$ then $\ell_2 \in (0,\infty)$ and $\Delta(t) \to 0$ if and only if $\delta(t) \to 0$ as $t \to \infty$.

(iii) If $\ell_1 = \infty$ then $|\ell_2| = \infty$ and $\Delta(t) \to 0$ if and only if $\delta^2(t)K_2(y(t)) \to 0$ as $t \to \infty$.

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

Proposition 2 (First order approximations of Δ) *Suppose* **(A1)**–**(A4)** *hold.*

(i) Assume $\ell_1 \in \{0,1\}$ (and thus $\ell_2 = 0$).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

3 Application to the ET approximation

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

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

so that $p_n = n^{-\tau_n}$, $\tau_n \geq 1$, $\alpha_n = n^{-\tau_n'}$, $\tau_n' \leq 1$ 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 $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n) = \Delta(n)$. In this context, the assumption $K_1 \in RV_{\theta_1}$, $\theta_1 \in \mathbb{R}$ is a sufficient condition for H^{-1} is extended regularly varying, see [7], Section B.2 for details on extended regular variation. This assumption has been introduced and discussed by Cees de Valk *et. al.* in a series of papers [8, 9, 10]. The next result provides a characterization of the tail behavior of F according to the sign of θ_1. We refer to [9], Theorem 1 for a characterization under the weaker assumption of extended regular variation.

Proposition 3 (Characterizations)
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 [2, 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 [1, 13, 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 θ_1 with respect to 1. Note that [9], Proposition 1 provides a similar classification under the weaker assumption of extended regular variation.

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

(i) If $K_1 \in RV_{\theta_1}$, $\theta_1 < 1$, then $F \in$MDA(Gumbel).

(ii) If $F \in$ MDA(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)**: MDA(Gumbel) is associated with $\theta_1 < 1$ while MDA(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 [7], Corollary 1.1.10, this distribution belongs to MDA(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 [7], Corollary 1.2.10, this distribution does not belong to MDA(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 [1] or [4], Section 8.8 for further developments on this topic. Applying Proposition 1 to the ET framework yields:

Theorem 1 (Necessary and sufficient conditions on (α_n, p_n) for $\varepsilon_{\mathbf{ET}}(p_n; \alpha_n) \to 0$) *Suppose F is increasing, twice differentiable and* **(A3)**, **(A4)** *hold. Let $0 < p_n \le 1/n \le \alpha_n < 1$ such that $\limsup \delta_n < 1$.*

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

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

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

First, if $F \in$MDA(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 be designed for this situation: $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n) \to 0$ under the restriction on (α_n, p_n) that $(\tau_n - \tau_n')^2 K_2(\log n) \to 0$ as $n \to \infty$. Second, it appears that, from the extrapolation error point of view, three sub-domains of MDA(Gumbel) can be exhibited:

- MDA_1(Gumbel) defined by $\ell_1 \in \{0, 1\}$ and where the relative extrapolation error tends to zero without restriction on the order p_n of the extreme quantile. 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 (the Gamma distribution for instance) as well as the class E defined in [6].

- MDA_2(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 $\log(p_n) \sim \log(1/n)$ as $n \to \infty$. 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).

- MDA_3(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: $\log(p_n)/\log(1/n) = 1 + o(|K_2(\log n)|^{1/2})$ as $n \to \infty$. As illustrated by Proposition 3(ii), this case corresponds to log-Weibull tail-distributions (including the lognormal distribution for instance).

We refer to Table 1 for examples of distributions in each sub-domain. Note that these three sub-domains do not cover the whole MDA(Gumbel) since they require the existence of ℓ_1 and thus K_1. To conclude this part, one may obtain first order approximations of the relative extrapolation error $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ 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 2.

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

(i) Assume $F \in$ MDA$_1$(Gumbel). If $\ell_1 = 1$, let us suppose that there exists $\theta_2 \leq 0$ such that $|K_2| \in RV_{\theta_2}$.

(a) If $\delta(n) \to 0$ then $\varepsilon_{ET}(p_n;\alpha_n) \sim \frac{1}{2}(\tau_n - \tau_n')^2 K_2(\log n) \sim \frac{1}{2}\delta^2(n) K_2(\log n)$.

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

(ii) Assume $F \in$ MDA$_2$(Gumbel)

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

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

(iii) Assume $F \in$ MDA$_3$(Gumbel)

(a) If $\delta(n) K_1(\log n) \to 0$ then $\varepsilon_{ET}(p_n;\alpha_n) \sim \frac{1}{2}(\tau_n - \tau_n')^2 K_1^2(\log n) \sim \frac{1}{2}\delta^2(n) K_2(\log n)$.

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

Before commenting the asymptotic behavior of $\varepsilon_{\mathrm{ET}}(p_n;\alpha_n)$, we would like to compare our results with [3].

Remark 1 *The extrapolation error associated with the GPD approximation has been studied in [3]. Focusing on MDA(Gumbel), the asymptotic equivalents provided by [3], Theorem 2 can be compared to our results. However, we would like to stress that [3], Theorem 2 only holds in the case where $\delta(n) \to 0$ as $n \to \infty$ 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 [3], Theorem 2.1 and Theorem 2.3 coincide with the ones of Theorem 2(i,a) and Theorem 2(iii,a) respectively. However, the first order approximation of $\varepsilon_{ET}(p_n;\alpha_n)$ stated in [3], Theorem 2.2 do not coincide with Theorem 2(ii,a) and seems to be wrong. This can be easily checked on the distribution defined by $H^{-1}(x) = x + 1/x$ where $\varepsilon_{ET}(p_n;\alpha_n) \sim (\delta(n)/\log n)^2$ as $n \to \infty$.*

The only situation where $\delta(n) \to \delta_\infty \neq 0$ and $\varepsilon_{\mathrm{ET}}(p_n;\alpha_n) \to 0$ as $n \to \infty$ occurs for $F \in$ MDA$_1$(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 $\delta(n) \to 0$ and $\varepsilon_{\mathrm{ET}}(p_n;\alpha_n) \to 0$ as $n \to \infty$, the convergence is the fastest in MDA$_1$(Gumbel) and the slowest in MDA$_3$(Gumbel). Let us also highlight that the rate of convergence is independent from the distribution in MDA$_2$(Gumbel). To illustrate these results, let us focus on the distributions introduced in Table 1. Clearly, in all six cases, $F \in$ DA(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

$$\tau_n = 1 + \frac{\log\log n}{\log n}\ , \tau_n' = 1 - \frac{\log\log n}{\log n} \text{ and } \delta(n) \sim 2\frac{\log\log n}{\log n}, \tag{8}$$

as $n \to \infty$. Let us stress that these choices entail $\delta(n) \to 0$ and $\delta(n)K_1(\log n) \to 0$ as $n \to \infty$ so that Theorem 2(i,ii,iii)-(a) can be applied and $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n) \to 0$ as $n \to \infty$ for all six distributions. The associated first order approximations of $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ are provided in Table 2. It appears that, 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 $\beta \overset{>}{\to} 1$. At the opposite, the Finite endpoint($\beta > 0$) distribution is the most favorable case, letting $\beta \overset{>}{\to} 0$ could lead to arbitrary high logarithmic rates of convergence.

As a conclusion, in MDA(Gumbel), the extrapolation abilites 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 MDA_2(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 [10].

4 Application to Weissman approximation

When $F \in$ MDA(Fréchet), $\gamma_n > 0$ and the GPD approximation (1) can be simplified by letting $\sigma_n = \gamma_n q(\alpha_n)$, see [7], Theorem 1.2.5, leading to

$$\tilde{q}_{\mathrm{W}}(p_n; \alpha_n) = q(\alpha_n) \left(\frac{\alpha_n}{p_n} \right)^{\gamma_n}, \tag{9}$$

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 γ_n by appropriate estimators:

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

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 (9) yields

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

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 $0 < p_n \le 1/n \le \alpha_n < 1$ and introduce $\varphi(\cdot) = \log(\bar{F})^{-1}(1/\exp(\cdot)) = \log U(\exp \cdot)$ where U is the tail quantile function, so that

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

The quantity of interest is

$$\varepsilon_{\mathrm{W}}(p_n;\alpha_n) := (q(p_n) - \tilde{q}_{\mathrm{W}}(p_n;\alpha_n))/q(p_n) = 1 - \exp(-\Delta(n)\log q(p_n)), \tag{10}$$

where $\Delta(n)$ is defined in (7). Here, the property $K_1 \in RV_{\theta_1}$ is a direct consequence of the assumption $F \in \mathrm{MDA}(\text{Fréchet})$. Indeed, from [7], Corollary 1.2.1, $F \in \mathrm{MDA}(\text{Fréchet})$ is equivalent to $U \in RV_\gamma$ for some $\gamma > 0$ which can be rewritten as

$$U(t) = t^\gamma L(t), \text{ with } L \in RV_0 \text{ and } \gamma > 0. \tag{11}$$

The function L is said to be slowly-varying [4]. Classical properties of slowly-varying functions yield, as $t \to \infty$,

$$\begin{aligned} \varphi(t) &= \gamma t + \log L(\exp t) \sim \gamma t, \\ \varphi'(t) &= \gamma + \eta(\exp t) \to \gamma, \end{aligned} \tag{12}$$

where $\eta(t) = tL'(t)/L(t)$ is called the auxiliary function associated with L. It follows that $K_1(t) \to 1$ as $t \to \infty$ and thus $\ell_1 = 1$ and $K_1 \in RV_0$. This means that only the case (i) of Proposition 1 and Proposition 2 has to be considered. In particular $\Delta(n) \to 0$ as $n \to \infty$ without further assumption, which is a very different situation from Section 3:

Theorem 3 (First order approximation of $\varepsilon_{\mathbf{W}}(p_n;\alpha_n)$)
Let $0 < p_n \le 1/n \le \alpha_n < 1$ such that $\limsup \delta_n < 1$. Suppose F is increasing, twice differentiable and (11) holds. Let $\eta(t) = tL'(t)/L(t)$ be the auxiliary function associated with L. Suppose K_1', η are asymptotically monotone and $|\eta| \in RV_\rho$ with $\rho < 0$.

(i) If $\delta(n) \to \delta_\infty \in (0,1)$ then $\varepsilon_W(p_n;\alpha_n) \to 0$ as $n \to \infty$.

(ii) Moreover, as $n \to \infty$,

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

The assumption $|\eta| \in RV_\rho$, $\rho < 0$, is recurrent in extreme-value statistics to control the bias of estimators, ρ being known as the second-order parameter, see *e.g.* [17]. This assumption holds for most heavy-tailed distributions such as Burr, 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 $\mathrm{MDA}_1(\text{Gumbel})$, see Theorem 2(i)-(b), and still obtain $\varepsilon_{\mathrm{W}}(p_n;\alpha_n) \to 0$ as $n \to \infty$. However, here, the relative extrapolation error converges to zero at a polynomial rate, depending on ρ.

5 Numerical illustrations

First, the quality of the first order approximations associated with the ET relative extrapolation error given in Table 2 is assessed graphically. Recall that these results are obtained by applying Theorem 2 to sequences (τ_n) and (τ_n') given in (8) 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 $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ 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 DA_3(Gumbel), see Figure 3. In all cases, the asymptotic sign of $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ is coherent with the first order equivalent given in Table 2: 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.

To conclude, Figure 4 displays the exact relative extrapolation error $\varepsilon_{\mathrm{W}}(p_n; \alpha_n)$ associated with Weissman estimator together with its corresponding first order approximation provided by Theorem 3(ii) 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 $U(t) := (t^{1/k} - 1)^k$, $t \geq 1$, $k > 0$, with extreme-value index $\gamma = 1$ and auxiliary function $\eta(t) = 1/(t^{1/k} - 1)$. Clearly, η is regularly varying with index $\rho = -1/k$. In both cases $k = 3$ (top) and $k = 4$ (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 3(i). As expected, both errors are negative since the auxiliary function η is positive. It also appears that, the smaller k, the faster the convergence is. This is in accordance with $\eta \in RV_{-1/k}$. Finally, the first order equivalent also provides a reasonable approximation of the error behavior in the Burr case.

6 Proofs of main results

Proof of Proposition 1. (i) If $\ell_1 = 0$ then Lemma 2(i) shows that $\ell_2 = 0$. If $\ell_1 = 1$ then, from Lemma 2(ii), $\ell_2 = 0$. Lemma 5(i) concludes the proof.
(ii) If $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$ then Lemma 2(iii) entails that ℓ_2 is finite and non zero. Lemma 5(i,ii) concludes the proof.
(iii) If $\ell_1 = \infty$ then $K_2(x) \sim K_1^2(x)$ as $x \to \infty$, see Lemma 2(iv). Besides, K_2 is regularly varying of order $2\theta_1$ and thus $K_2(x(t))$ and $K_2(y(t))$ are of the same order as $t \to \infty$ under **(A1)**. Lemma 5(i,ii) concludes the proof. ■

Proof of Proposition 2. (i) If $\ell_1 \in \{0, 1\}$ and $\delta(t) \to \delta_\infty \in [0, 1)$ as $t \to \infty$, then the result is a straightforward consequence of Lemma 4.
(ii) Assume $0 < \ell_1 < \infty$ and $\ell_1 \neq 1$. Then Lemma 2(iii) entails $\ell_2 = \ell_1(\ell_1 - 1)$, and Lemma 4 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 $\delta(t) \to \delta_\infty \in [0,1)$ as $t \to \infty$, Lebesgue's dominated convergence theorem entails

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

and the result is proved.

(iii) Assume $\ell_1 = \infty$. Then Lemma 2(iv) entails that $K_2(x) \sim K_1^2(x)$ as $x \to \infty$. As a consequence, Lemma 3(i) and Lebesgue's dominated convergence theorem yield

$$\Delta(t) \sim \delta^2(t)\int_0^1 \frac{K_1^2(y(t)(1-\delta(t)u))}{(1-\delta(t)u)^2}\exp\left(K_1(y(t))L_{\theta_1}(1-\delta(t)u)(1+o(1))\right)udu.$$

The regular variation property **(A3)** entails

$$\Delta(t) \sim \delta^2(t)K_1^2(y(t))\int_0^1 (1-\delta(t)u)^{2\theta_1-2}\exp\left(K_1(y(t))(L_{\theta_1}(1-\delta(t)u)(1+o(1))\right)udu.$$

Two main situations are considered:

1. If $\delta(t) \to 0$ as $t \to \infty$, then $L_{\theta_1}(1-\delta(t)u) = -\delta(t)u(1+o(1))$. 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 $A(t) \to 0$ as $t \to \infty$, then $\Phi(A(t)) \to 1/2$ in view of Lemma 1(i) and

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

(b) If $A(t) \to a \in (0,\infty)$ then $\Delta(t) \to a^2\Phi(a) = \int_0^a u\exp(-u)du$ as $t \to \infty$ in view of the continuity of Φ, see Lemma 1(i). If $A(t) \to \infty$, then $\Phi(A(t)) \sim 1/A^2(t)$ from Lemma 1(ii) and therefore $\Delta(t) \to 1 = \int_0^\infty u\exp(-u)du$ as $t \to \infty$.

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

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

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

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

where Ψ_1 is defined in Lemma 1 and

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

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

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

where Ψ_2 is defined in Lemma 1. As a consequence of Lemma 1(ii), $R(t) = O(1/K_1(y(t)))$ and $\Delta(t) \to 1$ as $t \to \infty$. Let us remark that this case is similar to the situation $\delta(t) \to 0$ and $A(t) \to \infty$. ■

Proof of Proposition 3. (i) If $H \in RV_\beta$, $\beta > 0$ then the monotone density theorem ([4], Proposition 1.7.2) yields

$$H(t) \sim \frac{1}{\beta}tH'(t) \text{ as } t \to \infty.$$

Letting $x = H(t)$, we have

$$x \sim \frac{1}{\beta}\frac{H^{-1}(x)}{(H^{-1})'(x)}$$

or equivalently $K_1(x) \to 1/\beta$ as $x \to \infty$. It follows that $\ell_1 = 1/\beta$ and $K_1 \in RV_0$.
(ii, $\Longleftarrow$) Let us assume that $H(\exp\cdot) \in RV_{1/\theta_1}$, $\theta_1 > 0$. Then, $\log H^{-1} \in RV_{\theta_1}$ and the monotone density theorem ([4], Proposition 1.7.2) implies $(\log H^{-1})' \in RV_{\theta_1 - 1}$ *i.e.* $K_1 \in RV_{\theta_1}$.
(ii, $\Longrightarrow$) Conversely, assume $K_1 \in RV_{\theta_1}$, $\theta_1 > 0$. Then, necessarily $\ell_1 = \infty$. From [4], Theorem 1.5.8, we have for all x_0 sufficiently large,

$$\log H^{-1}(x) - \log H^{-1}(x_0) = \int_{x_0}^x (\log H^{-1}(t))'dt = \int_{x_0}^x \frac{K_1(t)}{t}dt = \sim \frac{1}{\theta_1}K_1(x), \tag{13}$$

as $x \to \infty$. It is thus clear that $\log H^{-1} \in RV_{\theta_1}$ and therefore $H(\exp\cdot) \in RV_{1/\theta_1}$.
(iii, $\Longleftarrow$) Let us assume that $x^* < \infty$ and $h(\cdot) := H(x^*(1 - 1/\cdot)) \in RV_{-1/\theta_1}$, $\theta_1 < 0$. Consequently, $H^{-1}(\cdot) = x^*(1 - 1/h^{-1}(\cdot))$ where $h^{-1} \in RV_{-\theta_1}$ and $h^{-1}(x) \to \infty$ as $x \to \infty$. Straightforward calculations and the monotone density theorem ([4], Proposition 1.7.2) lead to

$$\begin{aligned}
\log H^{-1}(x) &= \log x^* + \log\left(1 - \frac{1}{h^{-1}(x)}\right) \\
(\log H^{-1})'(x) &= \frac{(h^{-1})'(x)}{h^{-1}(x)(h^{-1}(x)-1)} \\
K_1(x) &= \frac{x(h^{-1})'(x)}{h^{-1}(x)(h^{-1}(x)-1)} \sim -\frac{\theta_1}{h^{-1}(x)},
\end{aligned}$$

and therefore $K_1 \in RV_{\theta_1}$, $\theta_1 < 0$.
(iii, $\Longrightarrow$) Conversely, assume $K_1 \in RV_{\theta_1}$, $\theta_1 < 0$. Thus $(\log H^{-1})' \in RV_{\theta_1 - 1}$ and [4], Theorem 1.5.8 yields first, for all x sufficiently large,

$$\log x^* - \log H^{-1}(x) = \int_x^\infty (\log H^{-1})'(t)dt < \infty$$

and thus $x^* < \infty$. Second, one also has

$$\frac{K_1(x)}{\int_x^\infty (\log H^{-1})'(t)dt} \to -\theta_1$$

as $x \to \infty$. Combining the two above results yield

$$\frac{K_1(x)}{\log x^* - \log H^{-1}(x)} = -\theta_1(1+o(1))$$

and consequently

$$H^{-1}(x) = x^* \exp\left(\frac{1}{\theta_1}K_1(x)(1+o(1))\right) = x^*\left(1+\frac{1}{\theta_1}K_1(x)(1+o(1))\right)$$

since $K_1(x) \to 0$ as $x \to \infty$. Applying [4], Theorem 1.5.12 yields $H(x^*(1-1/\cdot)) \in RV_{-1/\theta_1}$ and concludes the proof. ■

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

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

Since $K_1 \in RV_{\theta_1}$ and the logarithm is a slowly-varying function, $K_1(\log \cdot) \in RV_0$ and thus

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

as $t \to \infty$. Besides,

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

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

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

as $t \to \infty$ since $K \in RV_{\theta_1}$ with $\theta_1 < 1$. As a conclusion, $U'(tx)/U'(t) \to 1/x$ as $t \to \infty$ for all $x > 0$ and thus $U' \in RV_{-1}$. This implies that $F \in \text{MDA(Gumbel)}$, see [7], Corollary 1.1.10.
(ii) Assume $F \in \text{MDA(Fréchet)}$. From [7], Corollary 1.2.10, there exists $\gamma > 0$ such that the tail quantile function $U \in RV_\gamma$. 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 $x \to \infty$ from the monotone density theorem [4], 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 $F \notin \text{MDA(Weibull)}$. Second, Proposition 4(ii) shows that $F \in \text{MDA(Fréchet)}$ entails $K_1 \in RV_1$. It is thus clear that $F \notin \text{MDA(Fréchet)}$. Finally, it remains to show that $F \notin \text{MDA(Gumbel)}$. To this end, consider for all $x > 0$ and $t \to \infty$,

$$\frac{U(tx)}{U(t)} = \frac{H^{-1}(\log tx)}{H^{-1}(\log t)} = \exp\left\{ \frac{1}{\theta_1} \left(K_1(\log tx) - K_1(\log t)\right)(1 + o(1))\right)\right\}$$

from (13) in the proof of Proposition 3(ii, $\Longrightarrow$). A first order Taylor expansion yields

$$\frac{U(tx)}{U(t)} = \exp\left\{ \frac{\log x}{\theta_1} K_1'(\log t + \eta \log x)(1 + o(1))\right\} = \exp\left\{ \frac{\log x}{\theta_1} K_1'(\log t)(1 + o(1))\right\}$$

where $\eta \in (0, 1)$ since $K_1' \in RV_{\theta_1 - 1}$. Recalling that $\theta_1 > 1$, it is then clear that $K_1'(\log t) \to \infty$ as $t \to \infty$ and therefore $U(tx)/U(t) \to 0$ as $t \to \infty$ if $x < 1$ while $U(tx)/U(t) \to \infty$ as $t \to \infty$ if $x > 1$. Finally [7], Lemma 1.2.9 shows that $F \notin \text{MDA(Gumbel)}$ since $U(tx)/U(t)$ does not converge to 1 as $t \to \infty$. ■

Proof of Theorem 1. The proof relies on the application of Proposition 1. Condition **(A1)** is fulfilled under the assumptions $0 < p_n \le 1/n \le \alpha_n < 1$ and $\limsup \delta_n < 1$.
(i) is a straightforward consequence of Proposition 1(i).
(ii) is based on the remark that $\delta(n) \to 0$ if and only if $\tau_n \to 1$ and $\tau_n' \to 1$ since, by assumption, $\tau_n' \le 1 \le \tau_n$.
(iii) Since $\ell_2 = \infty$, $\delta^2(n) K_2(\tau_n \log n) \to 0$ implies $\delta(n) \to 0$ and thus $\tau_n \to 1$ and $\tau_n' \to 1$ in view of the above remark. Thus, $\delta^2(n) \sim (\tau_n - \tau_n')^2$ as $n \to \infty$. Besides, Lemma 2(iv) entails that K_2 is regularly varying when $\ell_2 = \infty$. As a consequence, $K_2(\tau_n \log n) \sim K_2(\log n)$ as $n \to \infty$ and the result is proved. ■

Proof of Theorem 2. The proof relies on the application of Proposition 2.
(i) is a consequence of Proposition 2(i). Let us highlight that, when $\delta(n) \to 0$ then $x(n) \sim y(n)$ and thus $K_2(x_n) \sim K_2(y_n) \sim K_2(\log n)$ in view of the property $|K_2|$ is regularly varying. Moreover, $\delta^2(n) \sim (\tau_n - \tau_n')^2$ as already seen in the proof of Theorem 1.
(ii) is a straightforward consequence of Proposition 2(ii).
(iii) is a consequence of Proposition 2(iii). If $\delta(n) K_1(\log n) \to a \in [0, \infty)$ then, necessarily, $\delta(n) \to 0$ and thus $y(n) \sim \log n$ leading to $K_1(y(n)) \sim K_1(\log n)$. In the case where $\delta(n) K_1(\log n) \to \infty$, one still has $1 \le \liminf \tau_n \le \limsup \tau_n < \infty$ and thus $K_1(y(n))$ and $K_1(\log n)$ are asymptotically of the same order, the result is proved. ■

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

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

From (12), $\varphi''(t) = \exp(t)\eta'(\exp(t))$ and consequently

$$K_2(t) \sim \frac{1}{\gamma} t \exp(t)\eta'(\exp(t)) \text{ as } t \to \infty.$$

Moreover, η asymptotically monotone and $|\eta| \in RV_\rho$ imply $x\eta'(x)/\eta(x) \to \rho$ as $x \to \infty$, leading to, as $t \to \infty$,

$$K_2(t) \sim \frac{\rho}{\gamma} t\eta(\exp(t)).$$

It follows, when $\delta(n) \to \delta_\infty \in (0,1)$,

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

Since $|\eta| \in RV_\rho$, Potter's bounds (see for instance [7], Proposition B.1.9 (5.)) entail that there exists $0 < \epsilon < |\rho|$ such that

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

Recalling that η is asymptotically monotone with a constant sign yields

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

where $I_n^- \le I_n \le I_n^+$ and

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

Straightforward calculations show that, for all $x < 0$,

$$\int_0^1 u e^{y(n)\delta(n)(1-u)x} du \sim -\frac{1}{y(n)\delta_\infty x},$$

as $n \to \infty$, since $\delta(n) \to \delta_\infty \in (0,1)$ and thus $y(n)\delta(n) \to \infty$ as $n \to \infty$.
Consequently, $I_n^- \sim (1-\epsilon)/(y(n)\delta_\infty(\epsilon-\rho))$, $I_n^+ \sim (1+\epsilon)/(y(n)\delta_\infty(-\epsilon-\rho))$ and thus $I_n y(n) \in \left[\frac{(1-\epsilon)}{\delta_\infty(\epsilon-\rho)}(1+o(1)); \frac{(1+\epsilon)}{\delta_\infty(-\epsilon-\rho)}(1+o(1))\right]$. Letting $\epsilon \to 0$ entails $I_n y(n) \to -1/(\rho\delta_\infty)$ as $n \to \infty$ and thus

$$\Delta(n) \sim -\frac{\delta_\infty}{\gamma}\eta\left(e^{x(n)}\right).$$

Remarking that $\log q(p_n) = \varphi(y(n)) \sim \gamma y(n) \sim \dfrac{\gamma}{1-\delta_\infty} x(n)$ and taking account of (10) yield, when $\delta(n) \to \delta_\infty \in (0,1)$,

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

Finally, since $|\eta| \in RV_\rho$, $\rho < 0$, $\Delta(n)x(n) \sim -\dfrac{\delta_\infty}{\gamma}x(n)\eta(e^{x(n)}) \to 0$ as $n \to \infty$ and thus

$$\varepsilon_{\mathrm{W}}(p_n;\alpha_n) \sim -\frac{\delta_\infty}{1-\delta_\infty}x(n)\eta(e^{x(n)}) \sim -\frac{\delta_\infty}{1-\delta_\infty}\log(1/\alpha_n)\eta\left(1/\alpha_n\right),$$

which concludes the proof of (i) and (ii). ■

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] J. Cohen. Convergence rates for the ultimate and pentultimate approximations in extreme-value theory. *Advances in Applied Probability*, 14(4):833–854, 1982.

[7] L. de Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Science & Business Media, 2007.

[8] C. de Valk. Approximation and estimation of very small probabilities of multivariate extreme events. *Extremes*, 19(4):687–717, 2016.

[9] C. de Valk. Approximation of high quantiles from intermediate quantiles. *Extremes*, 19(4):661–686, 2016.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] J. Galambos. *The asymptotic theory of extreme order statistics*. R.E. Krieger publishing compagny, 1987.

[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: New Perspecties in Theoretical and Applied Statistics*, pages 467–477, 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.

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.

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* $\mathbb{R}_+$*,* $\Psi_a(0; b) = b^{a+1}/(a+1)$ *and* $\Psi_a(t; b) \to 0$ *as* $t \to \infty$*.*

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

The proof is straightforward. Lemma 2 below shows that $K_1 \in RV_{\theta_1}$ implies $|K_2| \in RV_{\theta_2}$ when $\ell_1 \neq 1$. In the case 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 $x \to \infty$.

Lemma 2 *Assume* **(A3)**, **(A4)** *hold.*

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

(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* $K_2(t) \sim K_1^2(t)$ *as* $t \to \infty$*.*

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

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

or, equivalently,

$$\frac{K_2(x)}{K_1^2(x)} = 1 + \frac{1}{K_1(x)}\left(\frac{xK_1'(x)}{K_1(x)} - 1\right). \tag{14}$$

Second, $xK_1'(x)/K_1(x) \to \theta_1$ as $x \to \infty$ from the monotone density theorem ([4], Proposition 1.7.2). Third, it straightforwardly follows that $\ell_2 = \ell_1(\ell_1 + \theta_1 - 1)$. Finally, for all positive function K, $K(x) \to c > 0$ as $x \to \infty$ implies $K \in RV_0$. ■

The next lemma establishes the links between δ and Δ through K_1 and K_2.

Lemma 3 *Suppose* **(A1)**–**(A4)** *hold.*

(i) For all $t > 0$*:*

$$\Delta(t) = \delta^2(t)\int_0^1 \frac{K_2(y(t)(1-\delta(t)u))}{(1-\delta(t)u)^2}\exp\left(K_1(y(t))L_{\theta_1}(1-\delta(t)u)(1+o(1))\right)udu,$$

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

(ii) If, moreover, $\ell_1 \neq 1$, then, for all $t > 0$:

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

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

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

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

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

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

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

It thus follows that

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

and the first part of the result is proved.
(ii) From Lemma 2, when $\ell_1 \neq 1$ the sign of K_2 is ultimately constant so that

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

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

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

Besides, Lemma 2 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

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

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

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

which concludes the proof. ■

In the case where $\ell_1 < \infty$, the asymptotic equivalent provided in Lemma 3(i) can be simplified as follows:

Lemma 4 *Suppose* **(A1)**–**(A4)** *hold and* $\ell_1 < \infty$. *Then,*

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

as $t \to \infty$.

Proof. If $\ell_1 = 0$ then Lemma 3(i) yields

$$
\Delta(t) = \delta^2(t)\int_0^1 K_2(y(t)(1-\delta(t)u))(1-\delta(t)u)^{-2}udu(1+o(1)).
$$

In the situation where $0 < \ell_1 < \infty$, Lemma 2(iii) entails $\theta_1 = 0$ and Lemma 3(i) yields

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

and the result is proved. ■

As a consequence of the two above results, a sufficient condition as well as a necessary condition can be established such that $\Delta(t) \to 0$ as $t \to \infty$.

Lemma 5 *Suppose* **(A1)**–**(A4)** *hold.*

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

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

Proof. Let us first note that when $\ell_1 = 1$ then $\ell_2 = 0$ from Lemma 2(ii). It is thus clear in view of Lemma 4 that $\Delta(t) \to 0$ as $t \to \infty$ without restriction on $\delta(t)$. In the following, we thus focus on the case where $\ell_1 \neq 1$. Lemma 2 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 $\delta^2(t)|K_2(y(t))| \to 0$ as $t \to \infty$. From Lemma 1(i), $0 \leq \Phi(s) \leq 1/2$ for all $s \geq 0$ and thus Lemma 3(ii) entails

$$|\Delta(t)| \leq \frac{\delta^2(t)|K_2(y(t))|}{2(1-\delta(t))^2} \to 0 \tag{15}$$

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

(ii) From Lemma 3(ii), one has

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

for t large enough and some $c > 0$ since Φ 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

$$\begin{aligned}
|\Delta(t)| &\geq \frac{\psi(s_0)}{s_0^2}|K_2(x(t))|\delta^2(t)\mathbb{I}\{A(t) \leq s_0/c\} \\
&+ \frac{\psi(s_0)}{c^2}\frac{|K_2(x(t))|}{K_1^2(y(t))}\mathbb{I}\{A_n \geq s_0/c\} \\
&\geq \frac{\psi(s_0)}{s_0^2}|K_2(x(t))|\delta^2(t)\mathbb{I}\{A(t) \leq s_0/c\} \\
&+ \frac{\psi(s_0)}{c^2}\frac{|K_2(x(t))|}{K_1^2(x(t))}\frac{K_1^2(x(t))}{K_1^2(y(t))}\mathbb{I}\{A(t) \geq s_0/c\}.
\end{aligned} \tag{16}$$

Since K_1 is regularly varying, $K_1(x(t))/K_1(y(t)) \sim (1-\delta(t))^{\theta_1} \geq c' > 0$ as $t \to \infty$ 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, (14) in the proof of Lemma 2 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, $|\Delta(t)| \to 0$ implies $|K_2(x(t))|\delta^2(t)\mathbb{I}\{A(t) \leq s_0/c\} \to 0$ and $\mathbb{I}\{A(t) \geq s_0/c\} \to 0$ as $t \to \infty$. Consequently, $A(t) \leq s_0/c$ eventually and $\delta^2(t)K_2(x(t)) \to 0$ as $t \to \infty$.
Let us now consider the situation where $|K_2|$ is ultimately non increasing.
(i) The proof is similar, the upper bound (15) is replaced by

$$|\Delta(t)| \leq \frac{\delta^2(t)|K_2(x(t))|}{2(1-\delta(t))^2}. \tag{17}$$

(ii) The lower bound (16) 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. ■

	$\bar{F}(x)$	θ_1	θ_2	$K_1(x)$	$K_2(x)$	ℓ_1
DA$_1$(Gumbel)						
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
DA$_2$(Gumbel)						
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	$\frac{1}{2}+o(1)$	$-\frac{1}{4}+o(1)$	$1/2$
DA$_3$(Gumbel)						
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 DA(Gumbel).

Distribution	First order approximation of $\varepsilon_{\mathrm{ET}}(p_n;\alpha_n)$
DA$_1$(Gumbel)	
Finite endpoint($\beta>0$)	$-\dfrac{2(1+\beta)}{\beta^2}\dfrac{(\log\log n)^2}{(\log n)^{2+1/\beta}}$
Gamma($a>0$)	$2(1-a)\dfrac{(\log\log n)^2}{(\log n)^3}$
DA$_2$(Gumbel)	
Weibull($\beta\neq 1$)	$\dfrac{2(1-\beta)}{\beta^2}\dfrac{(\log\log n)^2}{(\log n)^2}$
Gaussian	$-\dfrac{1}{2}\dfrac{(\log\log n)^2}{(\log n)^2}$
DA$_3$(Gumbel)	
Log-Weibull($\beta>1$)	$\dfrac{2}{\beta^2}\dfrac{(\log\log n)^2}{(\log n)^{2-2/\beta}}$
Lognormal	$\sigma^2\dfrac{(\log\log n)^2}{\log n}$

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

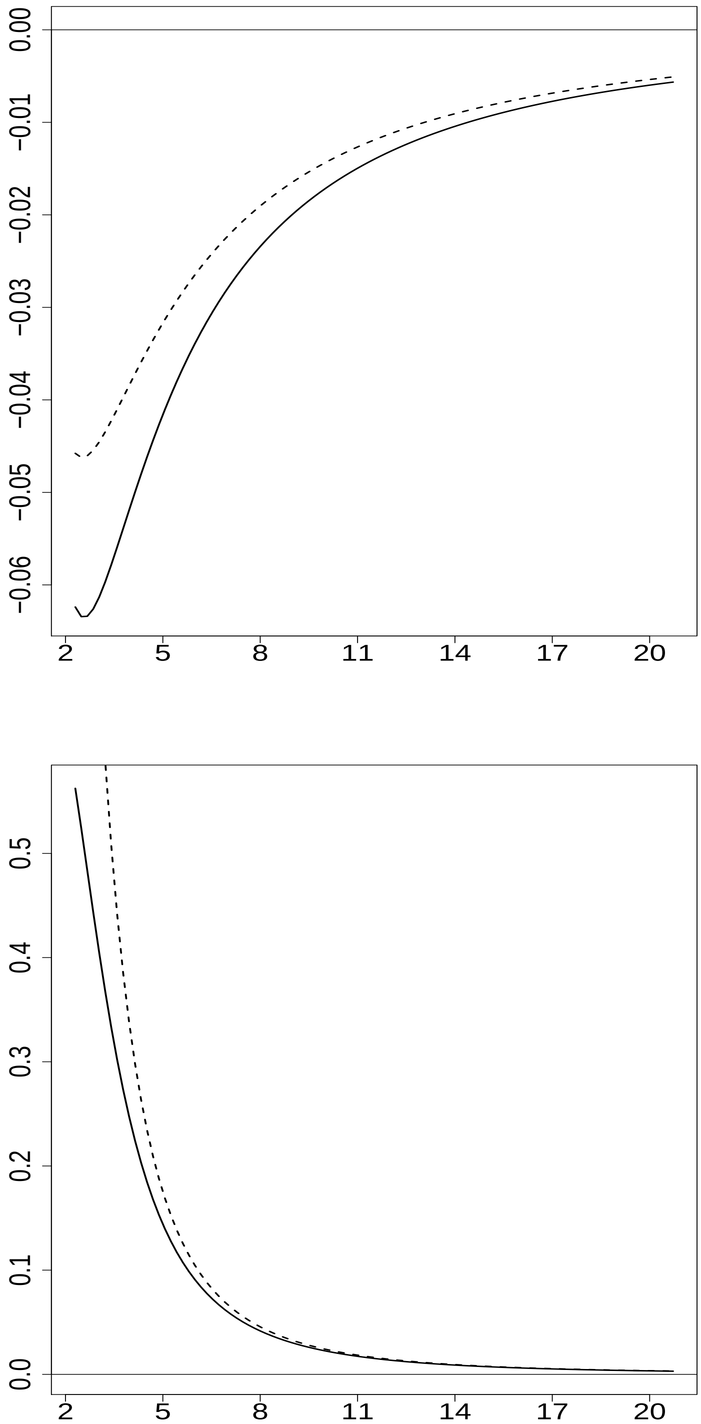

Figure 1: Extrapolation error in DA_1(Gumbel). Vertically: Extrapolation error $\varepsilon_{\mathrm{ET}}(p_n;\alpha_n)$ (solid line) and its first order approximation $\frac{1}{2}\eta_n^2 K_2(\log n)$ (dashed line) provided by Theorem 2(i)-(a). Horizontally: $\log n$. Top: Finite endpoint($\beta = 5$) distribution, bottom: Gamma($a = 0.1$) distribution.

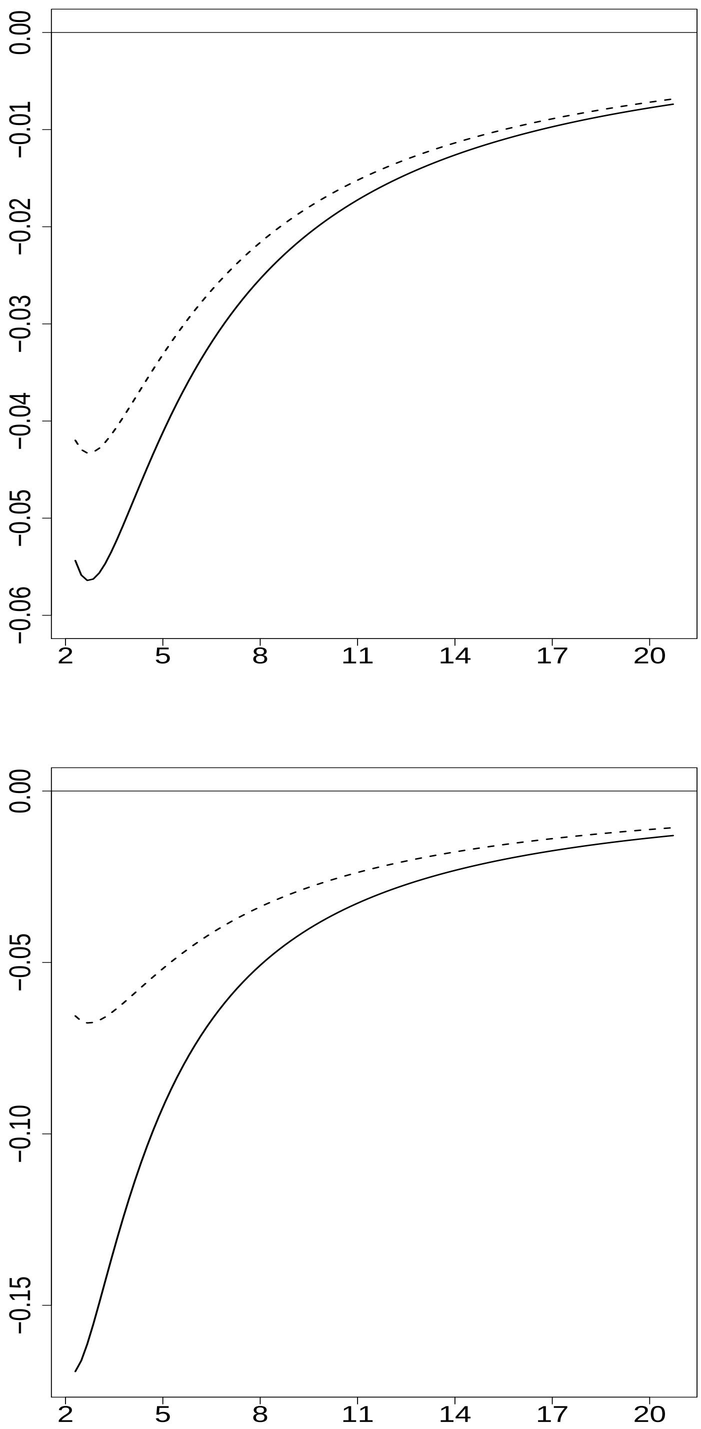

Figure 2: Extrapolation error in DA_2(Gumbel). Vertically: Extrapolation error $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ (solid line) and its first order approximation $\frac{\ell_1(\ell_1-1)}{2}\eta_n^2$ (dashed line) provided by Theorem 2(ii)-(a). Horizontally: $\log n$. Top: Weibull($\beta = 5$) distribution, bottom: Gaussian distribution.

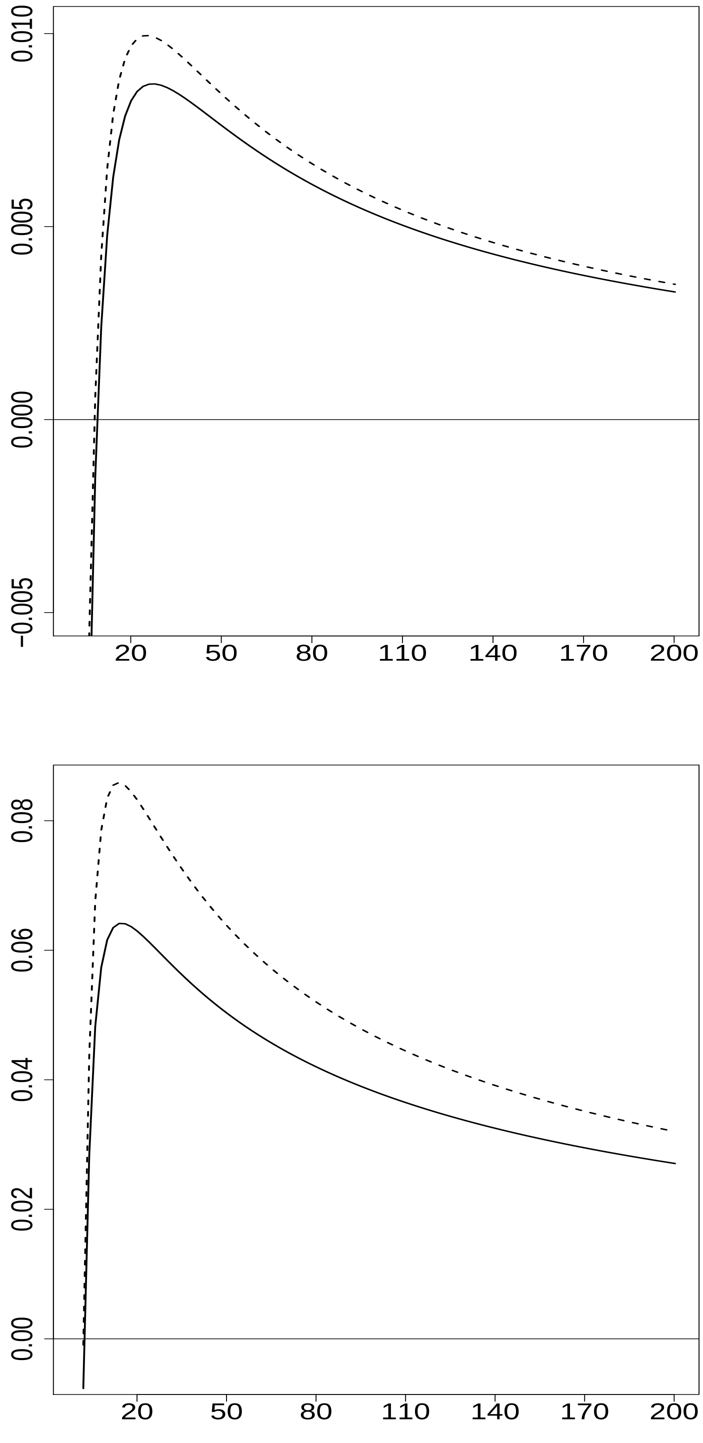

Figure 3: Extrapolation error in DA_3(Gumbel). Vertically: Extrapolation error $\varepsilon_{\mathrm{ET}}(p_n; \alpha_n)$ (solid line) and its first order approximation $\frac{1}{2}\eta_n^2 K_2(\log n)$ (dashed line) provided by Theorem 2(iii)-(a). Horizontally: $\log n$. Top: log-Weibull($\beta = 3$) distribution, bottom: lognormal($\sigma = 0.5$) distribution.

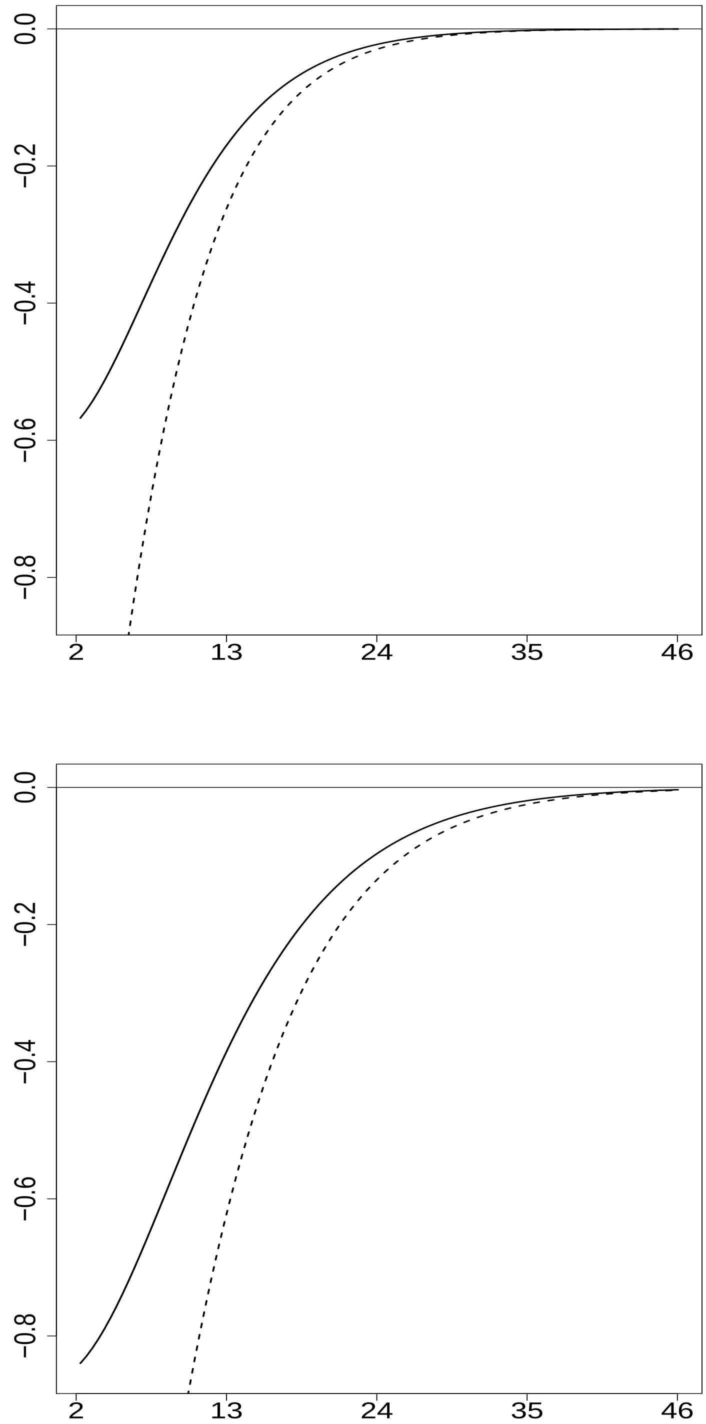

Figure 4: Extrapolation error in DA(Fréchet). Vertically: Extrapolation error $\varepsilon_{\mathrm{W}}(p_n;\alpha_n)$ (solid line) and its first order approximation $-\dfrac{\delta_\infty}{1-\delta_\infty}\log(1/\alpha_n)\eta(1/\alpha_n)$ (dashed line) provided by Theorem 3(ii). Horizontally: $\log n$. Top: Burr($k=3$) distribution (see Section 5), bottom: Burr($k=4$) distribution.