Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint

Approximations to the Tail Empirical Distribution Function with Application to Testing Extreme Value Conditions

Holger Drees, Laurens de Haan and Li Deyuan

Preprint No. 74

Saarbrücken 2002

Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Approximations to the Tail Empirical Distribution Function with Application to Testing Extreme Value Conditions

Holger Drees

Saarland University
Department of Mathematics
P.O. Box 151 150
66041 Saarbrücken
Germany
E-Mail: drees@num.uni-sb.de

Laurens de Haan

Department of Econometrics
Faculty of Economics
Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
E-Mail: ldhaan@few.eur.nl

Li Deyuan

Department of Econometrics
Faculty of Economics
Erasmus University Rotterdam
P.O. Box 1738
3000 DR Rotterdam
The Netherlands
E-Mail: dli@few.eur.nl

submitted: November 8, 2002

Preprint No. 74

Saarbrücken 2002

Edited by
FR 6.1 – Mathematik
Im Stadtwald
D–66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Abstract

A weighted approximation to the tail empirical distribution function is derived which is suitable for applications in extreme value statistics. The approximation is used to develop a Cramér-von Mises type test for the extreme value conditions. A useful auxiliary result is a tail approximation to the distribution function.

AMS 2000 Subject Classification: primary 62G30, 62G32; secondary 62G10
Key words: domain of attraction, limit distribution, maximum likelihood, tail empirical distribution function, weighted approximation

1 Introduction

Since Doob's (1949) and Donsker's (1952) path breaking work, it is well known that empirical processes do not only offer a concise way of describing observed independent and identically (i.i.d.) random variables (r.v.'s), but that the asymptotics of whole classes of statistics can be easily derived from limit theorems for empirical processes. If one is interested in the extremal behavior of the underlying distribution function (d.f.) then tail empirical processes are the objects of choice. Classical weighted approximations to the uniform tail empirical process can be found in Csörgő and Horváth (1993); see also Einmahl (1997) for further results of that type.
De Haan and Resnick (1998) and Resnick and Stărică (1997) used unweighted approximations to the tail empirical process to establish the asymptotic normality of the Hill estimator and its ramifications. To this end, however, one needs additional arguments to deal with the largest observations. To avoid such technical difficulties, one may use uniform approximations w.r.t. a weighted supremum norm. For tail empirical quantile functions such approximations were established by Drees (1998). It is the main aim of the present paper to derive analogous approximations to the tail empirical d.f.
When analyzing the tail behavior of the empirical d.f. for not necessarily uniformly distributed r.v.'s, one clearly needs regularity conditions on the tail behavior of the parent d.f. Since we have applications in extreme value statistics in mind, d.f.'s in the domain of attraction of an extreme value distribution suggest themselves. Thus, throughout the paper, we assume that i.i.d. r.v.'s X_i, $1 \le i \le n$, with d.f. F are observed such that

$$\lim_{n\to\infty} P\Big\{a_n^{-1}(\max_{1\le i\le n} X_i - b_n) \le x\Big\} = G_\gamma(x)$$

for all $x \in \mathbb{R}$, with some normalizing constants $a_n > 0$ and $b_n \in \mathbb{R}$, that is, $F \in D(G_\gamma)$. Here

$$G_\gamma(x) := \exp\big(-(1+\gamma x)^{-1/\gamma}\big) \tag{1.1}$$

for all $x \in \mathbb{R}$ such that $1 + \gamma x > 0$, and $\gamma \in \mathbb{R}$ is the so-called extreme value index. For $\gamma = 0$, the right-hand side of (1.1) is defined as $\exp(-\mathrm{e}^{-x})$.
This extreme value condition can be rephrased in the following way:

$$\lim_{t\to\infty} t\bar{F}(\tilde{a}(t)x + \tilde{b}(t)) = (1+\gamma x)^{-1/\gamma} \tag{1.2}$$

for all x with $1 + \gamma x > 0$. Here $\bar{F} := 1 - F$, $\tilde{a}$ is some positive normalizing function and $\tilde{b}(t) := U(t)$ with

$$U(t) := \left(\frac{1}{1-F}\right)^{\leftarrow}(t) = F^{\leftarrow}\Big(1 - \frac{1}{t}\Big)$$

and $F^{\leftarrow}$ denoting the generalized inverse of F. The right-hand side of (1.2) equals 1 minus the generalized Pareto d.f. in von Mises representation. A straightforward inversion shows that (1.2) is equivalent to

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{\tilde{a}(t)} = \frac{x^\gamma - 1}{\gamma} \tag{1.3}$$

for all $x > 0$.
In the sequel, we also need a second order refinement of (1.3) that specifies the speed of convergence. More concretely, we assume

$$\lim_{t\to\infty} \frac{\dfrac{U(tx)-U(t)}{\tilde{a}(t)} - \dfrac{x^\gamma - 1}{\gamma}}{\tilde{A}(t)} = \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma - 1}{\gamma}\right) := H_{\gamma,\rho}(x) \tag{1.4}$$

for all $x > 0$ and some $\rho \leq 0$. De Haan and Stadtmüller (1996) proved that any non-trivial limit must be of the type $H_{\gamma,\rho}$ and that $|\tilde{A}|$ is necessarily ρ–varying. Moreover, they proved that (1.4) is equivalent to a second order refinement of (1.2):

$$\lim_{t\to\infty} \frac{t\bar{F}\big(\tilde{a}(t)x + \tilde{b}(t)\big) - (1+\gamma x)^{-1/\gamma}}{\tilde{A}(t)} = (1+\gamma x)^{-1-1/\gamma} H_{\gamma,\rho}\big((1+\gamma x)^{-1/\gamma}\big) \tag{1.5}$$

for all x with $1 + \gamma x > 0$.
Now let $k_n \in \mathbb{N}$, $n \in \mathbb{N}$, be an intermediate sequence, that is

$$\lim_{n\to\infty} k_n = \infty, \quad \lim_{n\to\infty} k_n/n = 0.$$

The pertaining tail empirical d.f. is the process

$$x \mapsto \frac{n}{k_n} \bar{F}_n\Big(a\Big(\frac{n}{k_n}\Big)x + b\Big(\frac{n}{k_n}\Big)\Big), \quad x \in \mathbb{R}$$

where $\bar{F}_n := 1 - F_n$ with F_n denoting the empirical d.f. defined by

$$F_n(x) := \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i \le x\}}, \quad x \in \mathbb{R},$$

and a and b are suitable modifications of the normalizing functions $\tilde{a}$ and $\tilde{b}$ to be specified later (see Lemma 2.1 below).

In view of (1.2) and (1.5), the following normalization of the tail d.f. suggests itself:

$$\sqrt{k_n}\Big(\frac{n}{k_n} \bar{F}_n\Big(a\Big(\frac{n}{k_n}\Big)x + b\Big(\frac{n}{k_n}\Big)\Big) - (1+\gamma x)^{-1/\gamma}\Big), \quad x \in \mathbb{R}. \tag{1.6}$$

It is one of the goals of the paper to establish a weighted approximation to (1.6). It will turn out in Section 2 that, unlike the analogous approximation to the tail empirical quantile function, the asymptotic behavior of (1.6) is qualitatively different in the case $\gamma = \rho = 0$ from all other cases.

This approximation will then be used to devise a test for the extreme value condition $F \in D(G_\gamma)$ for some unspecified $\gamma > -1/2$. To this end, note that, according to the approximation to (1.6), a suitable distance between the tail empirical d.f. and the limiting Pareto d.f. will be small if $F \in D(G_\gamma)$, while this should not be expected if F does not belong to some domain of attraction. However, to construct a test statistic, first one must replace the unknown parameters γ, $a(n/k_n)$ and $b(n/k_n)$ by suitable estimators and prove an approximation to the resulting counterpart to (1.6). From this, we derive the limit distribution of a test statistic of Cramér-von Mises type in Section 2.

One important step in the proof of the approximation to (1.6) is an analogous result when $\bar{F}_n$ is replaced with $\bar{F}$, which we establish in Section 3. Section 4 contains the proof of the approximation to (1.6), while the proofs of the limit theorems for the tail empirical d.f. with estimated parameters and for the resulting test statistic are given in Section 5. The paper concludes with a simulation study about the limiting distribution of the test statistic and the size of the test for finite samples.

2 Main results

If i.i.d. uniformly distributed r.v.'s U_i are observed, then (1.3) holds with $\tilde{a}(t) = 1/t$ and $\gamma = -1$. For this particular case, Einmahl (1997, Corollary

3.3) gave a weighted approximation to the normalized tail empirical d.f. (1.6): there exist versions of the uniform tail empirical d.f.

$$U_n(t) := \frac{1}{n}\sum_{i=1}^{n} I_{\{U_i \le t\}}, \qquad t \in \mathbb{R},$$

(again denoted by U_n) and a sequence of Brownian motions W_n such that

$$\sup_{t>0} t^{-1/2}\mathrm{e}^{-\epsilon|\log t|}\left|\sqrt{k_n}\Big(\frac{n}{k_n}U_n\Big(\frac{k_n}{n}t\Big) - t\Big) - W_n(t)\right| \xrightarrow{P} 0 \tag{2.1}$$

as $n \to \infty$ for all intermediate sequences k_n, $n \in \mathbb{N}$.
By the well-known quantile transformation, $\bar{F}_n(\cdot)$ has the same distribution as $U_n(\bar{F}(\cdot))$. Hence, by (2.1), for suitable versions of $\bar{F}_n$

$$\sup_{\{x: z_n(x)>0\}} (z_n(x))^{-1/2}\mathrm{e}^{-\epsilon|\log z_n(x)|}$$
$$\left|\sqrt{k_n}\Big[\frac{n}{k_n}\bar{F}_n\Big(\tilde{a}\big(\frac{n}{k_n}\big)x + \tilde{b}\big(\frac{n}{k_n}\big)\Big) - z_n(x)\Big] - W_n(z_n(x))\right| \xrightarrow{P} 0 \tag{2.2}$$

with

$$z_n(x) := \frac{n}{k_n}\bar{F}\Big(\tilde{a}\big(\frac{n}{k_n}\big)x + \tilde{b}\big(\frac{n}{k_n}\big)\Big).$$

In view of (1.2), one may conjecture that (2.2) still holds if $z_n(x)$ is replaced with $(1+\gamma x)^{-1/\gamma}$. However, for this to be justified one must replace the normalizing functions $\tilde{a}$ and $\tilde{b}$ with suitable modifications such that (1.2) holds in a certain uniform sense.
In the sequel, we will focus on distributions which satisfy the second order conditions (1.4) or (1.5), since for these one can calculate the difference between (1.6) and its first order approximation $W_n((1+\gamma x)^{-1/\gamma})$ and, as a consequence, give simple conditions on k_n such that a weighted approximation to (1.6) is valid. In that case, the suitable normalizing functions a and b were determined by Drees (1998) and Cheng and Jiang (2001):

Lemma 2.1. *Suppose the second order condition (1.4) holds. Then there exist a function A, satisfying $A(t) \sim \tilde{A}(t)$ as $t \to \infty$, and for all $\epsilon > 0$ a constant $t_\epsilon > 0$ such that for all t and x with* $\min(t, tx) \ge t_\epsilon$

$$x^{-(\gamma+\rho)}\mathrm{e}^{-\epsilon|\log x|}\left|\frac{\dfrac{U(tx)-b(t)}{a(t)} - \dfrac{x^\gamma - 1}{\gamma}}{A(t)} - K_{\gamma,\rho}(x)\right| < \epsilon. \tag{2.3}$$

Here

$$a(t) := \begin{cases} ct^\gamma & \text{if} \quad \rho < 0, \\ \gamma U(t) & \text{if} \quad \rho = 0, \ \gamma > 0, \\ -\gamma(U(\infty) - U(t)) & \text{if} \quad \rho = 0, \ \gamma < 0, \\ U^{**}(t) + U^*(t) & \text{if} \quad \rho = 0, \ \gamma = 0, \end{cases}$$

$$b(t) := \begin{cases} U(t) - a(t)A(t)/(\gamma+\rho) & \text{if } \gamma + \rho \neq 0, \ \rho < 0, \\ U(t) & \text{else,} \end{cases}$$

and

$$K_{\gamma,\rho}(x) := \begin{cases} \frac{1}{\gamma+\rho} x^{\gamma+\rho} & \text{if} \quad \rho < 0, \gamma + \rho \neq 0, \\ \log x & \text{if} \quad \rho < 0, \gamma + \rho = 0, \\ \frac{1}{\gamma} x^\gamma \log x & \text{if} \quad \rho = 0 \neq \gamma, \\ \frac{1}{2} \log^2 x & \text{if} \quad \rho = 0 = \gamma, \end{cases}$$

with $c := \lim_{t\to\infty} t^{-\gamma}\tilde{a}(t)$ *(which exists in that case), and for any integrable function* g *the function* g^* *is defined by*

$$g^*(t) := g(t) - \frac{1}{t}\int_0^t g(u)dt.$$

In the sequel, we denote the right endpoint of the support of the generalized Pareto d.f. with extreme value index γ by

$$\frac{1}{(-\gamma) \vee 0} = \begin{cases} -1/\gamma & \text{if} \quad \gamma < 0, \\ \infty & \text{if} \quad \gamma \geq 0, \end{cases}$$

and its left endpoint by

$$-\frac{1}{\gamma \vee 0} = \begin{cases} -\infty & \text{if} \quad \gamma \leq 0, \\ -1/\gamma & \text{if} \quad \gamma > 0. \end{cases}$$

Then we have the following main result.

Theorem 2.1. *Suppose that the second order condition (1.4) holds for some* $\gamma \in \mathbb{R}$ *and* $\rho \leq 0$. *Let* k_n *be an intermediate sequence such that* $\sqrt{k_n}A(n/k_n)$, $n \in \mathbb{N}$, *is bounded and choose* a, b *and* A *as in Lemma 2.1. Then there exist versions of* F_n *and a sequence of Brownian motions* W_n *such that for all* $x_0 > -1/(\gamma \vee 0)$

(i)

$$\sup_{x_0 \le x < 1/((-\gamma)\vee 0)} \left((1+\gamma x)^{-1/\gamma}\right)^{-1/2+\epsilon} \cdot$$
$$\cdot \Big| \sqrt{k_n}\Big[\frac{n}{k_n}\bar{F}_n\Big(a\big(\frac{n}{k_n}\big)x + b\big(\frac{n}{k_n}\big)\Big) - (1+\gamma x)^{-1/\gamma}\Big]$$
$$- W_n\big((1+\gamma x)^{-1/\gamma}\big) - \sqrt{k_n}A\big(\frac{n}{k_n}\big)(1+\gamma x)^{-1/\gamma-1}K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big)\Big|$$
$$\xrightarrow{P} 0$$

if $\gamma \neq 0$ or $\rho < 0$, and

(ii)

$$\sup_{x_0 \le x < \infty} \Big(\max\Big(\mathrm{e}^{-x}, \frac{n}{k_n}\bar{F}\Big(a\big(\frac{n}{k_n}\big)x + b\big(\frac{n}{k_n}\big)\Big)\Big)\Big)^{-1/2+\epsilon} \cdot$$
$$\cdot \Big| \sqrt{k_n}\Big[\frac{n}{k_n}\bar{F}_n\Big(a\big(\frac{n}{k_n}\big)x + b\big(\frac{n}{k_n}\big)\Big) - \mathrm{e}^{-x}\Big]$$
$$- W_n(\mathrm{e}^{-x}) - \sqrt{k_n}A\big(\frac{n}{k_n}\big)\mathrm{e}^{-x}\frac{x^2}{2}\Big| \xrightarrow{P} 0$$

if $\gamma = \rho = 0$.

Remark 2.1. *If, in particular, $\sqrt{k_n}A(n/k_n)$ tends to 0, then the bias term $\sqrt{k_n}A(n/k_n)\ (1+\gamma x)^{-1/\gamma-1}K_{\gamma,\rho}((1+\gamma x)^{1/\gamma})$ is asymptotically negligible. In order for this statement to be true, it is sufficient to assume that the left-hand side of (1.5) remains bounded (rather than the present limit requirement) provided that k_n tends to infinity sufficiently slowly.*

The assertion in Theorem 2.1(ii) is wrong if the maximum of e^{-x} and n/k_n $\bar{F}(a(n/k_n)x + b(n/k_n))$ is replaced with just one of these two terms. Hence the asymptotic behavior of the tail empirical d.f. in the case $\gamma = \rho = 0$ is qualitatively different from the behavior in the case (i). This is due to the fact that in the case $\gamma \neq 0$ or $\rho < 0$ the tail behavior of F is essentially determined by the parameters γ and ρ, while in the case $\gamma = \rho = 0$ tail behaviors as diverse as $\bar{F}(x) \sim \exp(-\log^2 x)$, $\bar{F}(x) \sim \exp(-\sqrt{x})$ and $\bar{F}(x) \sim \exp(-x^2)$, say, are possible (cf. Example 3.1).
Nevertheless, also in the case $\gamma = \rho = 0$ results similar to the one in case (i) hold if $\max(\mathrm{e}^{-x},\ n/k_n\bar{F}(a(n/k_n)x + b(n/k_n)))$ is replaced with some weight function converging to ∞ much slower than e^{-x} as x tends to ∞:

Corollary 2.1. *Under the conditions of Theorem 2.1 with $\gamma = \rho = 0$ one* *has for all* $\tau > 0$

$$\sup_{x_0 \le x < \infty} \max(1, x^\tau) \Big| \sqrt{k_n} \Big[\frac{n}{k_n} \bar{F}_n \Big(a\big(\frac{n}{k_n}\big) x + b\big(\frac{n}{k_n}\big) \Big) - \mathrm{e}^{-x} \Big] - W_n(\mathrm{e}^{-x}) - \sqrt{k_n} A\big(\frac{n}{k_n}\big) \mathrm{e}^{-x} \frac{x^2}{2} \Big| \xrightarrow{P} 0.$$

The proofs of Theorem 2.1 and Corollary 2.1 are given in section 4. According to these results, the standardized tail empirical d.f.

$$x \mapsto \sqrt{k_n} \Big(\frac{n}{k_n} \bar{F}_n \Big(a\big(\frac{n}{k_n}\big) \frac{x^{-\gamma} - 1}{\gamma} + b\big(\frac{n}{k_n}\big) \Big) - x \Big), \quad x \in (0, 1],$$

converges to a Brownian motion plus a bias term if k_n tends to ∞ not too fast. This may be used to construct a test for $F \in D(G_\gamma)$. However, to this end, first the unknown parameters $\gamma, a(n/k_n)$ and $b(n/k_n)$ must be replaced with suitable estimators. The following result is an analog to Theorem 2.1(i) and Corollary 2.1 for the process with estimated parameters in the case $\gamma > -1/2$.

Proposition 2.1. *Suppose that the conditions of Theorem 2.1 are satisfied for some* $\gamma > -1/2$. *Let* $\hat{\gamma}_n$, $\hat{a}(n/k_n)$ *and* $\hat{b}(n/k_n)$ *be estimators such that*

$$\sqrt{k_n} \Big(\hat{\gamma}_n - \gamma, \ \frac{\hat{a}(n/k_n)}{a(n/k_n)} - 1, \ \frac{\hat{b}(n/k_n) - b(n/k_n)}{a(n/k_n)} \Big) - \big(\Gamma(W_n), \ \alpha(W_n), \ \beta(W_n) \big) \xrightarrow{P} 0 \qquad (2.4)$$

for some measurable real-valued functionals Γ, α *and* β *of the Brownian motions* W_n *used in Theorem 2.1. Then, for the versions of* $\bar{F}_n$ *used in Theorem 2.1 and every* $\epsilon > 0$ *and* $\tau > 0$, *one has*

$$\sup_{0 < x \le 1} h(x) \Big| \sqrt{k_n} \Big[\frac{n}{k_n} \bar{F}_n \Big(\hat{a}\big(\frac{n}{k_n}\big) \frac{x^{-\hat{\gamma}} - 1}{\hat{\gamma}} + \hat{b}\big(\frac{n}{k_n}\big) \Big) - x \Big] - W_n(x) - L_n^{(\gamma)}(x) - \sqrt{k_n} A\big(\frac{n}{k_n}\big) x^{\gamma+1} K_{\gamma,\rho}\big(\frac{1}{x}\big) \Big| \xrightarrow{P} 0 \qquad (2.5)$$

with

$$L_n^{(\gamma)}(x) := \begin{cases} \frac{1}{\gamma} x \big(\frac{1}{\gamma} \Gamma(W_n) - \alpha(W_n) \big) + \frac{1}{\gamma} \Gamma(W_n) x \log x \\ \quad - \frac{1}{\gamma} x^{1+\gamma} \big(\gamma \beta(W_n) + \frac{1}{\gamma} \Gamma(W_n) - \alpha(W_n) \big) & \text{if } \ \gamma \neq 0, \\ x \big(-\beta(W_n) - \frac{1}{2} \Gamma(W_n) \log^2 x + \alpha(W_n) \log x \big) & \text{if } \ \gamma = 0, \end{cases}$$

and

$$h(x) = \begin{cases} x^{-1/2+\epsilon} & \text{if} \quad \gamma \neq 0 \text{ or } \rho < 0, \\ (1 + |\log x|)^{\tau} & \text{if} \quad \gamma = \rho = 0. \end{cases}$$

Remark 2.2. *(i) If $\gamma < -1/2$, a rate of convergence of $k_n^{-1/2}$ for the estimators in (2.4) is not sufficient to ensure the approximation (2.5). To see this, note that in this case $\hat{b}(n/k_n) - b(n/k_n)$ is of larger order than $k_n^{-1/2}(n/k_n)^{\gamma-\epsilon}$ and hence also of larger order than the difference between the i_nth largest order statistic and the right endpoint $F^{\leftarrow}(1)$ for some sequence $i_n \to \infty$ not too fast, leading, for small $x > 0$, to a non-negligible difference between $\bar{F}_n(a(n/k_n)(x^{-\gamma} - 1)/\gamma + b(n/k_n))$ and the corresponding expression with estimated parameters.*

(ii) Typically the functionals Γ, α and β depend on the underlying d.f. F only through γ if the estimators $\hat{\gamma}_n$, $\hat{a}(n/k_n)$ and $\hat{b}(n/k_n)$ use only the largest $k_n + 1$ order statistics and $\sqrt{k_n}A(n/k_n) \to 0$. This justifies the notation $L_n^{(\gamma)}$ for the limiting function occurring in (2.5) in that case. However, if $\sqrt{k_n}A(n/k_n) \to c > 0$ then $L_n^{(\gamma)}$ will also depend on c; for simplicity, we ignore this dependence in the notation.

Example 2.1. *In Proposition 2.1 one may use the so-called maximum likelihood estimator in a generalized Pareto model discussed by Smith (1987). Since the excesses $X_{n-i+1,n} - X_{n-k_n,n}$, $1 \le i \le k_n$ over the random threshold $X_{n-k_n,n}$ are approximately distributed according to a generalized Pareto distribution with shape parameter γ and scale parameter $\sigma_n := a(n/k_n)$ if $F \in D(G_\gamma)$ and k_n is not too big, γ and σ_n are estimated by the pertaining maximum likelihood estimators $\hat{\gamma}_n$ and $\hat{\sigma}_n$ in an exact generalized Pareto model for the excesses. They can be calculated as the solutions to the equations*

$$\begin{aligned} \frac{1}{k}\sum_{i=1}^{k} \log\Big(1 + \frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})\Big) &= \gamma \\ \frac{1}{k}\sum_{i=1}^{k} \frac{1}{1 + \frac{\gamma}{\sigma}(X_{n-i+1,n} - X_{n-k,n})} &= \frac{1}{\gamma+1}. \end{aligned}$$

In Theorem 2.1 of Drees et al. (2002) it is proved that $\hat{\gamma}_n$, $\hat{a}(n/k_n) := \hat{\sigma}_n$ and

$\hat{b}(n/k_n) := X_{n-k_n,n}$ satisfy (2.4) with

$$\begin{aligned}
\Gamma(W_n) &= -\frac{(\gamma+1)^2}{\gamma}\big((2\gamma+1)S_n - R_n\big) + (\gamma+1)W_n(1),\\
\alpha(W_n) &= -\frac{\gamma+1}{\gamma}\big(R_n - (\gamma+1)(2\gamma+1)S_n\big) - (\gamma+2)W_n(1),\\
\beta(W_n) &= W_n(1),
\end{aligned}$$

where

$$\begin{aligned}
R_n &:= \int_0^1 t^{-1}W_n(t)\,dt,\\
S_n &:= \int_0^1 t^{\gamma-1}W_n(t)\,dt,
\end{aligned}$$

provided $\sqrt{k_n}A(n/k_n) \to 0$; if $\sqrt{k_n}A(n/k_n) \to c > 0$ then additional bias terms enter the formulas. As usual, for $\gamma = 0$, these expressions are to be interpreted as their limits as γ tends to 0, that is,

$$\begin{aligned}
\Gamma(W_n) &= -\int_0^1 (2+\log t)t^{-1}W_n(t)\,dt + W_n(1),\\
\alpha(W_n) &= \int_0^1 (3+\log t)t^{-1}W_n(t)\,dt - 2W_n(1),\\
\beta(W_n) &= W_n(1).
\end{aligned}$$

(Applying Vervaat's (1972) lemma to the approximation to the tail empirical distribution function given in Theorem 2.1, restricted to a compact interval bounded away from 0, and then using a Taylor expansion of $t \mapsto (t^{-\gamma}-1)/\gamma$ shows that the Brownian motions used by Drees et al. (2002) are indeed the Brownian motions used in Proposition 2.1 multiplied with -1.)
Hence one may apply Proposition 2.1 to obtain the asymptotics of the tail empirical distribution function with estimated parameters.

It is not difficult to devise tests for $F \in D(G_\gamma)$ with $\gamma > -1/2$ using approximation (2.5). Here we consider tests of Cramér-von Mises type based on the weighted L_2-statistic

$$T_n := \int_0^1 \left[\frac{n}{k_n}\bar{F}_n\left(\hat{a}(\frac{n}{k_n})\frac{x^{-\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}(\frac{n}{k_n})\right) - x\right]^2 x^{\eta-2}\,dx \qquad (2.6)$$

with suitable $\eta > 0$. The critical values for the corresponding test rejecting the null hypothesis if k_nT_n is too large can be calculated using the following limit theorem.

Theorem 2.2. *Under the conditions of Proposition 2.1 with $\sqrt{k_n}A(n/k_n) \to 0$ one has*

$$k_n T_n - \int_0^1 \Big(W_n(x) + L_n^{(\gamma)}(x)\Big)^2 x^{\eta-2}\, dx \xrightarrow{P} 0 \tag{2.7}$$

for all $\eta > 0$ if $\gamma \neq 0$ or $\rho < 0$, and all $\eta \geq 1$ if $\gamma = \rho = 0$.

Since the continuous distribution of $\int_0^1 (W_n(x) + L_n^{(\gamma)}(x))^2 x^{\eta-2} dx$ does not depend on n, for fixed $\gamma > -1/2$ its quantiles $Q_{p,\gamma}$ defined by $P\{\int_0^1 (W_n(x) + L_n^{(\gamma)}(x))^2 x^{\eta-2} dx \leq Q_{p,\gamma}\} = p$ can be easily obtained by simulations (see Section 6). Then the one-sided test rejecting $F \in D(G_\gamma)$ if $k_n T_n > Q_{1-\bar\alpha,\gamma}$ has asymptotic size $\bar\alpha \in (0,1)$.

If one wants to test $F \in D(G_\gamma)$ for an arbitrary unknown $\gamma > -1/2$, one may use the test rejecting the null hypothesis if $k_n T_n > Q_{1-\bar\alpha,\tilde\gamma_n}$ for some estimator $\tilde\gamma_n$ which is consistent for γ if $F \in D(G_\gamma)$. If the functionals Γ, α and β determining the limit distributions of $\hat\gamma_n$, $\hat a(n/k_n)$ and $\hat b(n/k_n)$ are continuous functions of γ (like the ones obtained in Example 2.1), then also $L_n^{(\gamma)}(x)$ and hence the quantiles $Q_{p,\gamma}$ are continuous functions of γ. Thus the test has asymptotic size $\bar\alpha$.

However, recall that, in fact, for (2.7) to hold we have not merely assumed that $F \in D(G_\gamma)$ but also that the second order condition (1.4) holds and, for the particular k_n used in the definition of the test statistic T_n, in addition we have assumed that $A(t) \to 0$ sufficiently fast such that $\sqrt{k_n}A(n/k_n) \to 0$. Hence, we actually test only the subset of the hypothesis $F \in D(G_\gamma)$ described by these additional assumptions.

A test for a similar hypothesis, but based on the tail empirical quantile function instead of the tail empirical distribution function, has been discussed by Dietrich et al. (2001). That test does not require $\gamma > -1/2$ but, on the other hand, $U(\infty) > 0$ and a slightly different second order condition were assumed.

The test based on the statistic $k_n T_n$ becomes particularly simple if Γ, α and β are the zero functional, that is, the estimators $\hat\gamma_n$, $\hat a(n/k_n)$ and $\hat b(n/k_n)$ converge at a faster rate than $k_n^{-1/2}$. This can be achieved by using suitable estimators based on m_n largest order statistics with $k_n = o(m_n)$ and $\sqrt{m_n}A(n/m_n) \to 0$. (For example, γ may be estimated by the estimator given in Example 2.1 with m_n instead of k_n, and $b(n/k_n)$ by a quantile estimator of the type described in de Haan and Rootzén (1993).) In that case the limit distribution $\int_0^1 W_n^2(x) x^{\eta-2} dx$ of the test statistic $k_n T_n$ does not depend on γ, so that no consistent estimator $\tilde\gamma_n$ for γ is needed. However, this approach has two disadvantages. Firstly, in practice it is often not an easy task to choose k_n such that the bias is negligible (i.e. $\sqrt{k_n}A(n/k_n) \to 0$). It is even more delicate to choose two numbers k_n and m_n such that k_n is much smaller than m_n but not too small and, at the same time, the bias of the estimators of the parameters is still not dominating when these are based on m_n order statistics. Secondly, while this approach may lead to a test whose actual size is closer to the nominal value $\bar{\alpha}$, the power of the test will probably higher if one choose a larger value for k_n, e.g. $k_n = m_n$, because the larger k_n the larger will typically be the test statistic $k_n T_n$ if the tail empirical d.f. is not well approximated by a generalized Pareto d.f. For these reasons, in the simulation study we will focus on the case where the tail empirical d.f. and the estimators $\hat{\gamma}_n$, $\hat{a}(n/k_n)$ and $\hat{b}(n/k_n)$ are based on the same number of largest order statistics.

3 Tail Approximation to the Distribution Function

A substantial part of the proof of Theorem 2.1 consists of proving an approximation to the tail of the (deterministic) distribution function.
For all $c, \delta > 0$ define sets

$$D_{t,\rho} := D_{t,\rho,\delta,c} := \begin{cases} \{x : t\bar{F}(a(t)x + b(t)) \le ct^{-\delta+1}\} & \text{if } \rho < 0, \\ \{x : t\bar{F}(a(t)x + b(t)) \le |A(t)|^{-c}\} & \text{if } \rho = 0. \end{cases}$$

Check that, in particular, eventually $[x_0, \infty) \subset D_{t,\rho}$ for all $x_0 > -1/(\gamma \vee 0)$.

Proposition 3.1. *Suppose that the second order relation (1.4) holds for some $\gamma \in \mathbb{R}$ and $\rho \le 0$. For $\epsilon > 0$, define*

$$w_t(x) := \begin{cases} \big(t\bar{F}(a(t)x + b(t))\big)^{\rho-1} \cdot \\ \quad \cdot \exp\big(-\epsilon|\log(t\bar{F}(a(t)x + b(t)))|\big) & \text{if } \gamma \ne 0 \text{ or } \rho \ne 0, \\ \min\Big((t\bar{F}(a(t)x + b(t)))^{-1} \cdot \\ \quad \cdot \exp\big(-\epsilon|\log(t\bar{F}(a(t)x + b(t)))|\big), \mathrm{e}^{x-\epsilon|x|}\Big) & \text{if } \gamma = \rho = 0. \end{cases}$$

Then, for all $\epsilon, \delta, c > 0$,

$$sup_{x \in D_{t,\rho}} w_t(x) \Big| \frac{t\bar{F}(a(t)x + b(t)) - (1 + \gamma x)^{-1/\gamma}}{A(t)} - \big(t\bar{F}(a(t)x + b(t))\big)^{1+\gamma} K_{\gamma,\rho}\Big(\frac{1}{t\bar{F}(a(t)x + b(t))}\Big)\Big| \to 0.$$

Moreover, we establish an analogous result where $t\bar{F}(a(t)x+b(t))$ is replaced with $(1+\gamma x)^{-1/\gamma}$. To this end, let

$$\tilde{D}_{t,\rho} := \tilde{D}_{t,\rho,\delta,c} := \begin{cases} \{x : (1+\gamma x)^{-1/\gamma} \le ct^{-\delta+1}\} & \text{if } \rho < 0, \\ \{x : (1+\gamma x)^{-1/\gamma} \le |A(t)|^{-c}\} & \text{if } \rho = 0, \end{cases}$$

and, for $\gamma \neq 0$ or $\rho < 0$,

$$\tilde{w}_t(x) := \big((1+\gamma x)^{-1/\gamma}\big)^{\rho-1} \exp\big(-\epsilon|\log((1+\gamma x)^{-1/\gamma})|\big).$$

Proposition 3.2. *If the second order relation (1.4) holds for some $\gamma \in \mathbb{R}$ and $\rho \le 0$, then*

$$\sup_{x\in D_{t,\rho}} w_t(x)\Big|\frac{t\bar{F}(a(t)x+b(t)) - (1+\gamma x)^{-1/\gamma}}{A(t)} \\ -(1+\gamma x)^{-1/\gamma-1} K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big)\Big| \to 0.$$

Moreover, if $\gamma \neq 0$ or $\rho < 0$, then

$$\sup_{x\in \tilde{D}_{t,\rho}} \tilde{w}_t(x)\Big|\frac{t\bar{F}(a(t)x+b(t)) - (1+\gamma x)^{-1/\gamma}}{A(t)} \\ -(1+\gamma x)^{-1/\gamma-1} K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big)\Big| \to 0,$$

and for $\gamma = \rho = 0$

$$\sup_{x\in \tilde{D}_{t,0}} w_t(x)\Big|\frac{t\bar{F}(a(t)x+b(t)) - \mathrm{e}^{-x}}{A(t)} - \mathrm{e}^{-x}\frac{x^2}{2}\Big| \to 0$$

for all $\delta, c > 0$.

At first glance, it is somewhat surprising that the results look differently in the case $\gamma = \rho = 0$ in that one needs a more complicated weight function, namely the minimum of a function of the standardized tail d.f. $t\bar{F}(a(t)x+b(t))$ and the corresponding function of the limiting exponential d.f. The following example shows that indeed the straightforward analog to the assertion in the case $\gamma \neq 0$ or $\rho < 0$ does not hold, because, in the case $\gamma = \rho = 0$, these two functions may behave quite differently for large x, despite the fact that for fixed x the former converges to the latter.

Example 3.1. *Here we give an example of a d.f. satisfying (1.4) such that*

$$\sup_{\{x: x > c\log|A(t)|\}} \mathrm{e}^{x-\epsilon|x|}\Big|\frac{t\bar{F}\big(a(t)x+b(t)\big) - \mathrm{e}^{-x}}{A(t)} - \mathrm{e}^{-x}\frac{x^2}{2}\Big| \qquad (3.1)$$

does not tend to 0 for any $c, \epsilon > 0$.
Let $F(x) := 1 - \mathrm{e}^{-\sqrt{x}}$, $x > 0$, *and* $a(t) := 2\log t$, $b(t) := \log^2 t$, $A(t) := 1/\log t$. *Then* $U(x) = \log^2 x$ *satisfies the second order condition (1.4):*

$$\frac{1}{A(t)}\Big(\frac{U(tx) - U(t)}{a(t)} - \log x\Big) \to \frac{\log^2 x}{2}.$$

Moreover

$$\begin{aligned} t\bar{F}\big(a(t)x + b(t)\big) &= t\exp\Big(-\sqrt{2x\log t + \log^2 t}\Big) \\ &= \exp\Big(-\log t\big(\sqrt{1 + 2x/\log t} - 1\big)\Big). \end{aligned}$$

Hence, for $x = x(t) = \lambda(t)\log t/2$ *with* $\lambda(t) \to \infty$ *as* $t \to \infty$, *one obtains*

$$\begin{aligned} t\bar{F}\big(a(t)x + b(t)\big) &= \exp\Big(-\log t\sqrt{\lambda(t)}(1 + o(1))\Big), \\ \mathrm{e}^{-x}\frac{x^2}{2} &= \frac{1}{8}\exp\Big(2(\log\log t + \log\lambda(t)) - \frac{1}{2}\lambda(t)\log t\Big) \\ &= o(t\bar{F}(a(t)x + b(t))), \\ \mathrm{e}^{-x} &= o\big(t\bar{F}(a(t)x + b(t))\big), \end{aligned}$$

so that

$$\frac{t\bar{F}\big(a(t)x + b(t)\big) - \mathrm{e}^{-x}}{A(t)} - \mathrm{e}^{-x}\frac{x^2}{2} = \frac{t\bar{F}\big(a(t)x + b(t)\big)}{A(t)}(1 + o(1)).$$

However, this contradicts the convergence of (3.1) to 0 as $t \to \infty$*:*

$$\begin{aligned} &(\mathrm{e}^{-x})^{-1+\epsilon}\Big|\frac{t\bar{F}(a(t)x + b(t)) - \mathrm{e}^{-x}}{A(t)} - \mathrm{e}^{-x}\frac{x^2}{2}\Big| \\ &= (\mathrm{e}^{-x})^{-1+\epsilon}\frac{t\bar{F}(a(t)x + b(t))}{A(t)}(1 + o(1)) \\ &= \exp\Big(\frac{1-\epsilon}{2}\lambda(t)\log t - \sqrt{\lambda(t)}\log t(1 + o(1))\Big)\cdot\frac{1 + o(1)}{A(t)} \\ &\to \infty. \end{aligned}$$

Likewise one can show that $F(x) = 1 - \mathrm{e}^{-x^2}$, $x > 0$, *satisfies the second order condition (1.4) but that*

$$\sup_{x\in D_{t,0}}\big(t\bar{F}(a(t)x + b(t))\big)^{-1}\exp\big(-\epsilon|\log(t\bar{F}(a(t)x + b(t)))|\big)\cdot \Big|\frac{t\bar{F}\big(a(t)x + b(t)\big) - \mathrm{e}^{-x}}{A(t)} - \mathrm{e}^{-x}\frac{x^2}{2}\Big| \to \infty.$$

Before proving the propositions, we need an auxiliary lemma. Let

$$q_t(x) := \frac{U(tx)-b(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}.$$

Lemma 3.1. *For each $\epsilon > 0$, there exists $\tilde{t}_\epsilon > 0$ such that for $t \geq \tilde{t}_\epsilon$*

$$\sup_{x \geq \tilde{t}_\epsilon/t} x^{-(\gamma+\rho)} \mathrm{e}^{-\epsilon|\log x|} |q_t(x)| = O(A(t)).$$

Proof: We focus on the case $\gamma = \rho = 0$; the assertion can be proved by the same arguments in the other cases. From Lemma 2.1 we know that, for each $\delta > 0$, there exists t_δ such that for $t, tx \geq t_\delta$

$$\mathrm{e}^{-\epsilon|\log x|}|q_t(x)| \leq \mathrm{e}^{-\epsilon|\log x|}|A(t)|\Big(\frac{\log^2 x}{2} + \delta \mathrm{e}^{\delta|\log x|}\Big).$$

Choose $\delta < \epsilon$ and $\tilde{t}_\epsilon = t_\delta$ to obtain the assertion, since $\sup_{x>0} \mathrm{e}^{-\epsilon|\log x|}\log^2 x < \infty$. □

Let

$$B_{t,\rho} := B_{t,\rho,\delta,c} \begin{cases} [ct^{\delta-1}, \infty) & \text{if } \rho < 0, \\ \{y : |\log y| \leq c|\log|A(t)||\} & \text{if } \rho = 0, \end{cases}$$

with $\delta, c > 0$.

Corollary 3.1. *For all $c, \delta > 0$,*

$$\sup_{x \in B_{t,\rho}} x^{-\gamma}|q_t(x)| \to 0$$

as $t \to \infty$.

Proof:

For $\rho < 0$, choose $\epsilon \leq |\rho|$ in Lemma 3.1 to obtain

$$\sup_{x\geq 1} x^{-\gamma}|q_t(x)| \leq \sup_{x \geq 1} x^{-(\gamma+\rho)}\mathrm{e}^{-\epsilon|\log x|}|q_t(x)| = O(A(t)) = o(1).$$

For all $c, \delta, t_\epsilon > 0$, eventually $ct^{\delta-1}$ is greater than $\tilde{t}_\epsilon/t$. Hence, by Lemma 3.1,

$$\sup_{ct^{\delta-1}\leq x<1} x^{-\gamma}|q_t(x)| \leq O(A(t)) \cdot \sup_{ct^{\delta-1}\leq x<1} x^{\rho-\epsilon} = O\big(A(t)\cdot t^{(\delta-1)(\rho-\epsilon)}\big) \to 0$$

if $(\delta-1)(\rho-\epsilon) < -\rho$ (which is satisfied for sufficient small $\epsilon > 0$), since $A(t)$ is ρ-varying and hence $A(t) = o(t^{\eta+\rho})$ for all $\eta > 0$.

In the case $\rho = 0$, one has for all $\epsilon \in (0, 1/c)$

$$\sup_{x \in B_{t,\rho}} x^{-\gamma} |q_t(x)| \le O(A(t)) \cdot \sup_{x \in B_{t,\rho}} \mathrm{e}^{\epsilon |\log x|} = O\big(A(t) \mathrm{e}^{\epsilon c |\log |A(t)||}\big) \to 0.$$

□

Proof of Proposition 3.1. :

For simplicity assume that F is eventually strictly increasing. (For more general F, the assertion follows by standard extra arguments using the second order condition (1.5).) Let $g(x) := (1 + \gamma x)^{-1/\gamma}$ and

$$y := \frac{1}{t\bar{F}(a(t)x + b(t))},$$

which implies that $x = (U(ty) - b(t))/a(t)$. Then $g'(x) = -(g(x))^{\gamma+1}$ and $g''(x) = (\gamma + 1)(g(x))^{2\gamma+1}$, and so

$$\begin{aligned}
& t\bar{F}\big(a(t)x + b(t)\big) - (1 + \gamma x)^{-1/\gamma} \\
&= -\bigg(\Big(1 + \gamma \frac{U(ty) - b(t)}{a(t)}\Big)^{-1/\gamma} - \Big(1 + \gamma \frac{y^\gamma - 1}{\gamma}\Big)^{-1/\gamma}\bigg) \\
&= -\bigg(g\Big(\frac{U(ty) - b(t)}{a(t)}\Big) - g\Big(\frac{y^\gamma - 1}{\gamma}\Big)\bigg) \\
&= q_t(y) \bigg(-g'\Big(\frac{y^\gamma - 1}{\gamma}\Big)\bigg) - \int_0^{q_t(y)} \int_0^s g''\Big(\frac{y^\gamma - 1}{\gamma} + u\Big)\, du\, ds \\
&= q_t(y) y^{-\gamma-1} - \int_0^{q_t(y)} \int_0^s (1 + \gamma) \bigg(1 + \gamma\Big(\frac{y^\gamma - 1}{\gamma} + u\Big)\bigg)^{-1/\gamma - 2} du\, ds
\end{aligned}$$

with $(1 + \gamma x)^{-1/\gamma - j} := \mathrm{e}^{-x}$ for $\gamma = 0$ and $j = 1, 2$.
Since $(1 + \gamma((y^\gamma - 1)/\gamma + u))^{-1/\gamma-2}$ lies between $(1 + \gamma(y^\gamma - 1)/\gamma)^{-1/\gamma-2} = y^{-1-2\gamma}$ and $(1 + \gamma((y^\gamma - 1)/\gamma + q_t(y)))^{-1/\gamma-2} = y^{-1-2\gamma}(1 + \gamma y^{-\gamma} q_t(y))^{-1/\gamma-2}$, Corollary 3.1 yields

$$\big|t\bar{F}(a(t)x + b(t)) - (1 + \gamma x)^{-1/\gamma} - q_t(y) y^{-1-\gamma}\big| \le 2|1 + \gamma| y^{-1-2\gamma} q_t^2(y) \quad (3.2)$$

for all $y \in B_{t,\rho}$ and sufficiently large t.
Since $ty \to \infty$ uniformly for $y \in B_{t,\rho}$, (2.3), Lemma 3.1 and Corollary 3.1

imply

$$\sup_{y\in B_{t,\rho}} w_t(x)\Big|\frac{t\bar F\big(a(t)x+b(t)\big)-(1+\gamma x)^{-1/\gamma}}{A(t)} - \big(t\bar F(a(t)x+b(t))\big)^{1+\gamma}K_{\gamma,\rho}\big(\frac{1}{t\bar F(a(t)x+b(t))}\big)\Big|$$

$$\le \sup_{y\in B_{t,\rho}} y^{1-\rho}\mathrm{e}^{-\epsilon|\log y|}\Big(\Big|\frac{q_t(y)y^{-(1+\gamma)}}{A(t)} - y^{-(1+\gamma)}K_{\gamma,\rho}(y)\Big| + 2|1+\gamma|\frac{y^{-(1+2\gamma)}}{|A(t)|}q_t^2(y)\Big)$$

$$\le \sup_{y\in B_{t,\rho}} y^{-(\gamma+\rho)}\mathrm{e}^{-\epsilon|\log y|}\Big|\frac{q_t(y)}{A(t)} - K_{\gamma,\rho}(y)\Big| + 2|1+\gamma|\sup_{y\in B_{t,\rho}} y^{-(\gamma+\rho)}\mathrm{e}^{-\epsilon|\log y|}\frac{|q_t(y)|}{|A(t)|}\sup_{y\in B_{t,\rho}} y^{-\gamma}|q_t(y)|$$

$$\to 0.$$

Because $x \in D_{t,\rho,\delta,c}$ is equivalent to $y \in B_{t,\rho,\delta,1/c}$ if $\rho < 0$, the assertion is proved in that case.
In the case $\rho = 0$, it remains to prove that, for sufficiently large c,

$$\sup_{\{x:1/y<|A(t)|^c\}} w_t(x)\Big|\frac{y^{-1}-(1+\gamma x)^{-1/\gamma}}{A(t)} - y^{-(1+\gamma)}K_{\gamma,0}(y)\Big| \to 0.$$

Note that

$$w_t(x)\Big|y^{-(1+\gamma)}K_{\gamma,0}(y)\Big| = O\Big(\mathrm{e}^{-\epsilon|\log y|}\log^2 y\Big) = o(1)$$

uniformly for $1/y < |A(t)|^c$. Moreover, for $c > 1/\epsilon$,

$$w_t(x)y^{-1} \le \mathrm{e}^{-\epsilon|\log y|} \le |A(t)|^{c\epsilon} = o(A(t))$$

for all x such that $1/y < |A(t)|^c$.
Therefore, it suffices to verify that

$$\sup_{\{x:\ t\bar F(a(t)x+b(t))<|A(t)|^c\}} w_t(x)(1+\gamma x)^{-1/\gamma} = o(A(t)). \tag{3.3}$$

To this end, we distinguish 3 cases.
First suppose $\gamma > 0$. Then $(1+\gamma x)U(t) = a(t)x+b(t) \to \infty$ uniformly for all x such that $1/y = t\bar F(a(t)x+b(t)) < |A(t)|^c \to 0$. By the Potter bounds (see Bingham et al. (1987), Theorem 1.5.6)

$$y^{-1} = t\bar F\big(a(t)x+b(t)\big) = \frac{\bar F\big((1+\gamma x)U(t)\big)}{\bar F(U(t))} \ge \frac{1}{2}(1+\gamma x)^{-1/(\gamma(1-\epsilon/2))} \tag{3.4}$$

for sufficient large t. Hence the left side of (3.3) is bounded by

$$\sup_{\{x:1/y<|A(t)|^c\}} y^{1-\epsilon}\big(2y^{-1}\big)^{1-\epsilon/2} \le 2|A(t)|^{\epsilon c/2} = o(A(t))$$

when we choose $c > 2/\epsilon$.
Likewise, for $\gamma < 0$, one has

$$\begin{aligned} y^{-1} &= t\bar{F}\big(a(t)x+b(t)\big) = \frac{\bar{F}\big(U(\infty)-(1+\gamma x)(U(\infty)-U(t))\big)}{\bar{F}\big(U(\infty)-(U(\infty)-U(t))\big)} \\ &\ge \frac{1}{2}(1+\gamma x)^{-1/(\gamma(1-\epsilon/2))} \end{aligned} \tag{3.5}$$

and one can argue like in the case $\gamma > 0$.
Finally, if $\gamma = 0$ then the left side of (3.3) is bounded by

$$\sup_{\{x:\ t\bar{F}(a(t)x_t+b(t))<|A(t)|^c\}} \mathrm{e}^{-\epsilon x} = \mathrm{e}^{-\epsilon x_t}$$

with $x_t = \inf\{x : t\bar{F}(a(t)x+b(t)) \le |A(t)|^c\}$. According to (3.2), Lemma 3.1, and Corollary 3.1, one has eventually

$$\begin{aligned} \mathrm{e}^{-x_t} &\le |A(t)|^c + \big|q_t\big(|A(t)|^{-c}\big)\big||A(t)|^c + 2q_t^2\big(|A(t)|^{-c}\big)|A(t)|^c \\ &= |A(t)|^c\Big(1+\big|q_t\big(|A(t)|^{-c}\big)\big| + 2q_t^2\big(|A(t)|^{-c}\big)\Big) \\ &= |A(t)|^c\Big(1+O\big(|A(t)|^{1-\epsilon c}\big)\Big) \\ &= O\big(|A(t)|^{c(1-\epsilon)}\big) \end{aligned}$$

which implies that $\mathrm{e}^{-\epsilon x_t} = O(|A(t)|^{\epsilon c(1-\epsilon)}) = o(A(t))$ for $c > 2/\epsilon$ and $\epsilon < 1/2$.
The proof of Proposition 3.1 is complete. □

Proof of Proposition 3.2:

Recall the definition $y := 1/\big(t\bar{F}(a(t)x+b(t))\big)$. We consider three cases.

Case(i): $\rho < 0$.
Inequality (3.2) and Corollary 3.1 imply

$$\sup_{x\in D_{t,\rho}} \Big|y(1+\gamma x)^{-1/\gamma}-1\Big| \le \sup_{y\ge ct^{\delta-1}} y^{-\gamma}|q_t(y)| + 2|1+\gamma| \sup_{y\ge ct^{\delta-1}} (y^{-\gamma}q_t(y))^2 \to 0. \tag{3.6}$$

Hence, for $\gamma+\rho \ne 0$, by the definition of $K_{\gamma,\rho}$

$$\begin{aligned} &\sup_{x\in D_{t,\rho}} w_t(x)\Big|(1+\gamma x)^{-(1+1/\gamma)}K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big) - y^{-(1+\gamma)}K_{\gamma,\rho}(y)\Big| \\ &= \sup_{x\in D_{t,\rho}} \mathrm{e}^{-\epsilon|\log y|}\frac{1}{|\gamma+\rho|}\Big|\big(y(1+\gamma x)^{-1/\gamma}\big)^{1-\rho}-1\Big| \\ &\to 0. \end{aligned} \tag{3.7}$$

If $\gamma + \rho = 0$, then the left-hand side of (3.7) equals

$$\sup_{x\in D_{t,\rho}} \mathrm{e}^{-\epsilon|\log y|}\Big|\big(y(1+\gamma x)^{-1/\gamma}\big)^{1+\gamma}\log\big(y(1+\gamma x)^{-1/\gamma}\big) \\ +\Big(\big(y(1+\gamma x)^{-1/\gamma}\big)^{1+\gamma}-1\Big)(-\log y)\Big| \to 0.$$

Now the first assertion is immediate from Proposition 3.1. In view of (3.6), $\tilde{w}_t(x)/w_t(x)$ tends to 1 uniformly for $x \in D_{t,\rho}$. Moreover, $(1+\gamma x)^{-1/\gamma} \le ct^{-\delta+1}$ implies $t\bar{F}(a(t)x+b(t)) \le 2ct^{-\delta+1}$ for sufficient large t. Thus the second assertion follows immediate from the first.

Case (ii): $\rho = 0,\ \gamma \neq 0$.
Define

$$\begin{aligned} D^1_{t,0} &:= \big\{x : |A(t)|^c \le t\bar{F}(a(t)x+b(t)) \le |A(t)|^{-c}\big\} \\ &= \big\{x : \big|\log\big(t\bar{F}(a(t)x+b(t))\big)\big| \le c|\log|A(t)||\big\}, \\ D^2_{t,0} &:= \big\{x : t\bar{F}(a(t)x+b(t)) \le |A(t)|^c\big\}, \end{aligned}$$

so that $D_{t,0} = D^1_{t,0} \cup D^2_{t,0}$. As in the first case, (3.2) and Corollary 3.1 imply

$$\sup_{x\in D^1_{t,0}}\Big|y(1+\gamma x)^{-1/\gamma}-1\Big| \le \sup_{y\in B_{t,0}} y^{-\gamma}|q_t(y)| + 2|1+\gamma|\sup_{y\in B_{t,0}}(y^{-\gamma}q_t(y))^2 \to 0. \tag{3.8}$$

Hence

$$\begin{aligned} &\sup_{x\in D^1_{t,0}} w_t(x)\Big|(1+\gamma x)^{-(1+1/\gamma)}K_{\gamma,0}\big((1+\gamma x)^{1/\gamma}\big) - y^{-(1+\gamma)}K_{\gamma,0}(y)\Big| \\ &= \frac{1}{|\gamma|}\sup_{x\in D^1_{t,0}} \mathrm{e}^{-\epsilon|\log y|}\Big|y(1+\gamma x)^{-1/\gamma}\log\big(y(1+\gamma x)^{-1/\gamma}\big) \\ &\qquad + \big(y(1+\gamma x)^{-1/\gamma}-1\big)(-\log y)\Big| \\ &\to 0. \end{aligned} \tag{3.9}$$

Note that

$$\sup_{x\in D^2_{t,0}} w_t(x)\Big|y^{-(1+\gamma)}K_{\gamma,0}(y)\Big| = \sup_{x\in D^2_{t,0}}\frac{1}{|\gamma|}\mathrm{e}^{-\epsilon|\log y|}|\log y| \to 0. \tag{3.10}$$

Therefore, it remains to verify that

$$\begin{aligned} &\sup_{x\in D^2_{t,0}} w_t(x)\Big|(1+\gamma x)^{-1/\gamma-1}K_{\gamma,0}\big((1+\gamma x)^{1/\gamma}\big)\Big| \\ &= \frac{1}{|\gamma|}\sup_{x\in D^2_{t,0}} w_t(x)(1+\gamma x)^{-1/\gamma}\big|\log\big((1+\gamma x)^{-1/\gamma}\big)\big| = o(1). \end{aligned} \tag{3.11}$$

For $\gamma > 0$, (3.4) shows that the right-hand side of (3.11) is bounded by

$$\begin{aligned}
&\frac{1}{|\gamma|}\sup_{x\in D^2_{t,0}} w_t(x)\big(2y^{-1}\big)^{1-\epsilon/2}\big|\log\big(2y^{-1}\big)\big| \\
&\leq \frac{1}{|\gamma|}\sup_{y\geq |A(t)|^{-c}} y^{1-\epsilon}\big(2y^{-1}\big)^{1-\epsilon/2}\big|\log\big(2y^{-1}\big)\big| \to 0,
\end{aligned}$$

and hence (3.11). In the case $\gamma < 0$, we can argue likewise. So the first assertion is immediate from (3.9)–(3.11) and Proposition 3.1.
For the second assertion, it suffices to prove that $\{x : (1+\gamma x)^{-1/\gamma} \leq |A(t)|^{-c}\} \subset \{x : t\bar{F}(a(t)x+b(t)) \leq |A(t)|^{-2c}\}$ eventually, and that $\sup_{x\in D_{t,0}} \tilde{w}_t(x)/w_t(x)$ is bounded when in the definition of w_t we replace ϵ with $\epsilon/2$.
From (3.8), we have $\sup_{x\in D^1_{t,0}} \tilde{w}_t(x)/w_t(x) \to 1$, so we must check whether $\sup_{x\in D^2_{t,0}} \tilde{w}_t(x)/w_t(x)$ is bounded.
We only discuss the case $\gamma > 0$, since the arguments are similar for $\gamma < 0$. By the Potter bounds, for all $\eta > 0$ and sufficiently large t,

$$2(1+\gamma x)^{-(1-\eta)/\gamma} \geq y^{-1} = t\bar{F}\big(a(t)x+b(t)\big) \geq \frac{1}{2}(1+\gamma x)^{-1/(\gamma(1-\eta))}$$

uniformly for all $x \in D^2_{t,0}$ (cf. (3.4)). Thus

$$\begin{aligned}
&\sup_{x\in D^2_{t,0}} \frac{\tilde{w}_t(x)}{w_t(x)} = \sup_{x\in D^2_{t,0}} \frac{y^{\epsilon/2-1}}{(1+\gamma x)^{-(1-\epsilon)/\gamma}} \\
&\leq \sup_{x\in D^2_{t,0}} 2^{1-\epsilon/2}\frac{(1+\gamma x)^{-1/\gamma}}{(1+\gamma x)^{-(1-\epsilon)/\gamma}} \leq 2\sup_{x\in D^2_{t,0}} (1+\gamma x)^{-\epsilon/\gamma} \\
&\leq 2\sup_{x\in D^2_{t,0}} (2y^{-1})^{\epsilon(1-\epsilon/2)} \to 0.
\end{aligned}$$

Thus $\sup_{x\in D_{t,0}} \tilde{w}_t(x)/w_t(x)$ is bounded for $\gamma > 0$.
Next we verify $\{x : (1+\gamma x)^{-1/\gamma} \leq |A(t)|^{-c}\} \subset \{x : t\bar{F}(a(t)x+b(t)) \leq |A(t)|^{-2c}\}$. To this end, define $x_t := \inf\{x \,|\, t\bar{F}(a(t)x+b(t)) \leq |A(t)|^{-2c}\}$. Then by the analog to (3.8), $(1+\gamma x_t)^{-1/\gamma} \sim t\bar{F}(a(t)x_t+b(t)) = |A(t)|^{-2c}$. Hence for x satisfying $(1+\gamma x)^{-1/\gamma} \leq |A(t)|^{-c}$, we have for sufficient large t, $(1+\gamma x)^{-1/\gamma} < (1+\gamma x_t)^{-1/\gamma}$, which implies $x > x_t$, and $t\bar{F}(a(t)x+b(t)) < t\bar{F}(a(t)x_t+b(t)) = |A(t)|^{-2c}$. Hence we obtain $\{x : (1+\gamma x)^{-1/\gamma} \leq |A(t)|^{-c}\} \subset \{x : t\bar{F}(a(t)x+b(t)) \leq |A(t)|^{-2c}\}$, and the proof of the second assertion is complete.

Case (iii): $\gamma = \rho = 0$.
In the very same way as for $\rho < 0$, we obtain for all $d > 0$

$$\sup_{\{x:|\log y|\leq d|\log|A(t)||\}} |y\mathrm{e}^{-x} - 1| \to 0. \tag{3.12}$$

Thus

$$\begin{aligned}
&\sup_{x\in D^1_{t,0}} w_t(x)\Big|\mathrm{e}^{-x}x^2-\frac{\log^2 y}{y}\Big|\\
&\le \sup_{x\in D^1_{t,0}} \mathrm{e}^{-\epsilon|\log y|}\Big|y\mathrm{e}^{-x}\log\left(y\mathrm{e}^{-x}\right)\log(\mathrm{e}^{-x}/y)+(y\mathrm{e}^{-x}-1)\log^2 y\Big|\\
&\to 0.
\end{aligned}$$

Moreover, in view of (3.12) with $d=2c$, eventually $-\log y < c\log|A(t)|$ implies $-x<c\log|A(t)|/2$. Hence

$$\begin{aligned}
&\sup_{x\in D^2_{t,0}} w_t(x)\Big|\mathrm{e}^{-x}x^2-\frac{\log^2 y}{y}\Big|\\
&\le \sup_{x>-c\log|A(t)|/2}\mathrm{e}^{-\epsilon|x|}x^2+\sup_{x>-c\log|A(t)|/2}\mathrm{e}^{-\epsilon|\log y|}\log^2 y\to 0.
\end{aligned}$$

Again the first assertion follows from Proposition 3.1.
Finally, in view of (3.12), $\mathrm{e}^{-x}<|A(t)|^{-c}$ implies $1/y<|A(t)|^{-2c}$ for sufficiently large t, so that the second assertion is obvious.
The proof of Proposition 3.2 is complete. □

4 Tail Approximation to the Empirical Distribution Function

For the proof of Theorem 2.1, we need two additional Lemmas.

Lemma 4.1. *Suppose* $x_0>-1/(\gamma\vee 0)$.
(i) If $\rho<0$*, then*

$$\sup_{x_0\le x<1/((-\gamma)\vee 0)}\Big|\frac{(1+\gamma x)^{-1/\gamma}}{t\bar F(a(t)x+b(t))}-1\Big|\ \to\ 0.$$

(ii) If $\rho=0$ *and* $\gamma\ne 0$*, then for all* $\eta>0$

$$\sup_{x_0\le x<1/((-\gamma)\vee 0)}\frac{t\bar F\big(a(t)x+b(t)\big)-(1-\gamma x)^{-1/\gamma}}{\Big((1+\gamma x)^{-1/\gamma}\Big)^{1-\eta}}\ \to\ 0$$

as $t\to\infty$ *so that*

$$\sup_{x_0\le x<1/((-\gamma)\vee 0)}\frac{t\bar F\big(a(t)x+b(t)\big)}{\Big((1+\gamma x)^{-1/\gamma}\Big)^{1-\eta}}\qquad \textit{is bounded.}$$

(iii) If $\gamma = \rho = 0$, then for all $\eta, c > 0$

$$\sup_{x_0 \le x < -c \log |A(t)|} \frac{t\bar{F}\big(a(t)x + b(t)\big) - e^{-x}}{e^{-(1-\eta)x}} \to 0$$

as $t \to \infty$ so that

$$\sup_{x_0 \le x < -c \log |A(t)|} \frac{t\bar{F}\big(a(t)x + b(t)\big)}{e^{-(1-\eta)x}} \qquad \textit{is bounded.}$$

Proof:

(i) By (3.6), one has for all $\delta \in (0,1)$ and $c > 0$

$$\Big[x_0, \frac{1}{(-\gamma) \vee 0}\Big) \subset \Big\{x : (1+\gamma x)^{-1/\gamma} \le \frac{c}{2} t^{-\delta+1}\Big\} \subset D_{t,\rho}$$

for sufficiently large t. Hence, again by (3.6),

$$\sup_{x_0 \le x < 1/((-\gamma)\vee 0)} \Big| \frac{(1+\gamma x)^{-1/\gamma}}{t\bar{F}\big(a(t)x + b(t)\big)} - 1 \Big| \to 0.$$

(ii) By similar arguments as in (i), one concludes $[x_0, 1/((-\gamma) \vee 0)) \subset D_{t,0}$. Hence Proposition 3.2 with $\epsilon = \eta$ implies

$$\begin{aligned}
&\sup_{x_0 \le x < 1/((-\gamma)\vee 0))} \big((1+\gamma x)^{-1/\gamma}\big)^{\eta-1} \Big| \frac{t\bar{F}\big(a(t)x + b(t)\big) - (1+\gamma x)^{-1/\gamma}}{A(t)} \\
&\qquad\qquad - (1+\gamma x)^{-1/\gamma-1} K_{\gamma,0}\big((1+\gamma x)^{1/\gamma}\big)\Big| \\
&= \frac{1}{A(t)} \Big| \frac{t\bar{F}\big(a(t)x + b(t)\big) - (1+\gamma x)^{-1/\gamma}}{\big((1+\gamma x)^{-1/\gamma}\big)^{1-\eta}} \\
&\qquad\qquad - A(t)\big((1+\gamma x)^{-1/\gamma}\big)^{\gamma+\eta} K_{\gamma,0}\big((1+\gamma x)^{1/\gamma}\big)\Big| \\
&\to 0.
\end{aligned}$$

Because $A(t) \to 0$ and $(1+\gamma x)^{-1/\gamma}$ is bounded for $x \ge x_0$, the assertions are immediate from the definition of $K_{\gamma,0}$.

(iii) The proof is similar to the one of (ii). Note that (3.12) shows that $w_t(x)/e^{(1-\epsilon)x} \to 1$ uniformly for $x_0 \le x < -c \log |A(t)|$. □

Lemma 4.2. *Let W denote a Brownian motion.*
(i) If $\gamma \ne 0$ or $\rho < 0$, then

$$\begin{aligned}
\sup_{x_0 \le x < 1/((-\gamma)\vee 0)} \big((1+\gamma x)^{-1/\gamma}\big)^{-1/2+\epsilon} \Big| &W\Big(\frac{n}{k_n}\bar{F}\big(a(\frac{n}{k_n})x + b(\frac{n}{k_n})\big)\Big) \\
&- W\big((1+\gamma x)^{-1/\gamma}\big)\Big| \to 0 \quad \textit{a.s.}
\end{aligned}$$

as $n \to \infty$.
(ii) If $\rho = \gamma = 0$, *then*

$$\sup_{x_0 \le x} \Big(\max \big(\mathrm{e}^{-x}, \frac{n}{k_n} \bar{F}\big(a(\frac{n}{k_n})x + b(\frac{n}{k_n})\big)\big) \Big)^{-\frac{1}{2}+\epsilon} \cdot$$
$$\cdot \Big| W\Big(\frac{n}{k_n}\bar{F}\big(a(\frac{n}{k_n})x + b(\frac{n}{k_n})\big)\Big) - W\big((1+\gamma x)^{-1/\gamma}\big) \Big| \to 0 \quad a.s.$$

Proof:

(i) Let $s := (1+\gamma x)^{-1/\gamma}$ and $u(n,s) := (n/k_n)\bar{F}(a(n/k_n)x + b(n/k_n)) - s$. Then $x_0 \le x < 1/((-\gamma) \vee 0)$ implies $0 < s \le s_0$, where s_0 is a constant depending on x_0. So we only need to prove

$$\sup_{0<s\le s_0} s^{-1/2+\epsilon} \big| W\big(u(n,s)+s\big) - W(s)\big| \to 0 \quad a.s. \tag{4.1}$$

From Lemma 4.1, one can easily conclude that $u(n,s) = o(s^{1-\eta})$ uniformly for $s \in (0, s_0]$, so that in the sequel we may assume $u(n,s) \le s^{1-\eta}$.
For all $0 < a < 1$

$$\sup_{0<s\le a} s^{-1/2+\epsilon} |W(u(n,s)+s) - W(s)|$$
$$\le \sup_{0<s\le a} s^{-1/2+\epsilon}\big(s+u(n,s)\big)^{1/2-\epsilon/2} \sup_{0<s\le a} \left| \frac{W(u(n,s)+s)}{\big(s+u(n,s)\big)^{(1-\epsilon)/2}} \right| + \sup_{0<s\le a} \left| \frac{W(s)}{s^{1/2-\epsilon}} \right|.$$

Since

$$\lim_{a\to 0} \sup_{0<s\le a} s^{-1/2+\epsilon}\big(s+u(n,s)\big)^{(1-\epsilon)/2} = \lim_{a\to 0} \sup_{0<s\le a} \Big(s^{\epsilon/(1-\epsilon)} + \frac{u(n,s)}{s^{1-\epsilon/(1-\epsilon)}} \Big)^{(1-\epsilon)/2} = 0,$$

the law of iterated logarithm yields

$$\lim_{a\to 0} \sup_{0<s\le a} s^{-1/2+\epsilon} |W(u(n,s)+s) - W(s)| = 0 \quad a.s.$$

On the other hand, by the continuity of W, for all fixed $a > 0$

$$\lim_{n\to\infty} \sup_{a<s\le s_0} s^{-1/2+\epsilon} |W(u(n,s)+s) - W(s)| = 0 \quad a.s.$$

Therefore one obtains (4.1) by a standard diagonal argument.
(ii) We consider $x \in [x_0, -c\log|A(n/k_n)|)$ and $x \in [-c\log|A(n/k_n)|, \infty)$ separately.

As in the proof of (i), one may conclude from Lemma 4.1 that

$$\sup_{x_0\le x<-c\log|A(n/k_n)|}(\mathrm{e}^{-x})^{-\frac12+\epsilon}\Big|W\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)-W(\mathrm{e}^{-x})\Big|\to 0\quad a.s. \tag{4.2}$$

Since $\mathrm{e}^{-x}\to 0$ and $(n/k_n)\bar F(a(n/k_n)x+b(n/k_n))\to 0$ uniformly for $x\ge -c\log|A(n/k_n)|$, we get

$$\begin{aligned}
&\sup_{x\ge -c\log|A(n/k_n)|}\Big(\max\big(\mathrm{e}^{-x},\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\big)\Big)^{-1/2+\epsilon}\cdot\\
&\qquad\qquad\cdot\Big|W\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)-W(\mathrm{e}^{-x})\Big|\\
\le\;&\sup_{x\ge -c\log|A(n/k_n)|}\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)^{-1/2+\epsilon}\Big|W\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)\Big|\\
&+\sup_{x\ge -c\log|A(n/k_n)|}(\mathrm{e}^{-x})^{-1/2+\epsilon}|W(\mathrm{e}^{-x})|\\
\to\;& 0\quad a.s.
\end{aligned} \tag{4.3}$$

by the law of the iterated logarithm. A combination of (4.2) and (4.3) proves the assertion. □

Proof of Theorem 2.1:

We confine ourselves to the case $\gamma\ne 0$ or $\rho<0$, because the other case can be treated similarly.

Define

$$\epsilon^*=\begin{cases}\epsilon & \text{if } \rho<0,\\ \epsilon/2 & \text{if } \rho=0\ne\gamma,\end{cases}$$

and

$$\begin{aligned}
I:=\;&\big((1+\gamma x)^{-1/\gamma}\big)^{-1/2+\epsilon}\,\Big|\sqrt{k_n}\Big[\frac{n}{k_n}\bar F_n\Big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\Big)-(1+\gamma x)^{-1/\gamma}\Big]\\
&-W_n\big((1+\gamma x)^{-1/\gamma}\big)-\sqrt{k_n}A(\frac{n}{k_n})(1+\gamma x)^{-1/\gamma-1}K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big)\Big|\\
\le\;&\frac{\big((1+\gamma x)^{-1/\gamma}\big)^{-1/2+\epsilon}}{\big(\frac{n}{k_n}\bar F(a(\frac{n}{k_n})x+b(\frac{n}{k_n}))\big)^{-1/2+\epsilon^*}}\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)^{-1/2+\epsilon^*}\cdot\\
&\cdot\;\Big|\sqrt{k_n}\Big[\frac{n}{k_n}\bar F_n\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)-\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big]\\
&\quad-W_n\Big(\frac{n}{k_n}\bar F\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)\Big|
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\big((1+\gamma x)^{-1/\gamma}\big)^{-1/2+\epsilon}}{\tilde{w}_t(x)}\tilde{w}_t(x)\sqrt{k_n}A\Big(\frac{n}{k_n}\Big)\cdot\\
&\cdot\Bigg|\frac{\frac{n}{k_n}\bar{F}\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)-(1+\gamma x)^{-1/\gamma}}{A(\frac{n}{k_n})}\\
&\qquad\qquad -(1+\gamma x)^{-1/\gamma-1}K_{\gamma,\rho}\big((1+\gamma x)^{1/\gamma}\big)\Bigg|\\
&+\big((1+\gamma x)^{-1/\gamma}\big)^{-1/2+\epsilon}\Bigg|W_n\Big(\frac{n}{k_n}\bar{F}\big(a(\frac{n}{k_n})x+b(\frac{n}{k_n})\big)\Big)-W_n\big((1+\gamma x)^{-1/\gamma}\big)\Bigg|\\
&:=I_1+I_2+I_3.
\end{aligned}
$$

By (2.2) and Lemma 4.1 one has $\sup_{x_0\le x<1/((-\gamma)\vee 0)} I_1 \xrightarrow{P} 0$. From Proposition 3.2 and the fact that $((1+\gamma x)^{-1/\gamma})^{-1/2+\epsilon}/\tilde{w}_t(x)$ is bounded uniformly for $x_0 \le x < 1/((-\gamma)\vee 0)$, it follows that $\sup_{x_0\le x<1/((-\gamma)\vee 0)} I_2 \to 0$. Finally Lemma 4.2 shows that $\sup_{x_0\le x<1/((-\gamma)\vee 0)} I_3 \xrightarrow{P} 0$. □

Proof of Corollary 2.1:

Because of Theorem 2.1(ii) and $\max(1,x^\tau)=o(\mathrm{e}^{(1/2-\epsilon)x})$ as $x\to\infty$ for all $\tau>0$ and $\epsilon\in(0,1/2)$, it suffices to prove that

$$\sup_{x_0\le x<\infty}\Big(\frac{n}{k_n}\bar{F}\Big(a\big(\frac{n}{k_n}\big)x+b\big(\frac{n}{k_n}\big)\Big)\Big)^{1/2-\epsilon}\max(1,x^\tau)=O(1).$$

According to Lemma 2.2 of Resnick (1987), there exists a function $\bar{a}$ such that $a(t)/\bar{a}(t)\to 1$ as $t\to\infty$ and

$$F^t\big(\bar{a}(t)x+b(t)\big)\ge 1-(1+\delta)^3(1+\delta x)^{-1/\delta}$$

for all $\delta>0$, sufficiently large t and $x\ge x_0$. Thus, by the mean value theorem, there exists $\theta_{t,x}\in(0,1)$ such that

$$
\begin{aligned}
t\bar{F}\big(\bar{a}(t)x+b(t)\big) &\le t\Big(1-\Big(1-(1+\delta)^3(1+\delta x)^{-1/\delta}\Big)^{1/t}\Big)\\
&= (1+\delta)^3(1+\delta x)^{-1/\delta}\Big(1-\theta_{t,x}(1+\delta)^3(1+\delta x)^{-1/\delta}\Big)^{1/t-1}\\
&\le 2(1+\delta x)^{-1/\delta}
\end{aligned}
$$

if $x\ge 0$ and $\delta>0$ is sufficiently small. Since by the locally uniform convergence in (1.2)

$$\sup_{x_0\le x<0}\Big(\frac{n}{k_n}\bar{F}\Big(a\big(\frac{n}{k_n}\big)x+b\big(\frac{n}{k_n}\big)\Big)\Big)^{1/2-\epsilon}\max(1,x^\tau)=O(1),$$

it follows that

$$\begin{aligned}
&\sup_{x_0\le x<\infty}\Big(\frac{n}{k_n}\bar F\Big(a\big(\frac{n}{k_n}\big)x+b\big(\frac{n}{k_n}\big)\Big)\Big)^{1/2-\epsilon}\max(1,x^\tau)\\
&= O(1)+2\sup_{0\le x<\infty}\Big(1+\frac{\delta}{2}x\Big)^{-1/\delta}\max(1,x^\tau)\\
&= O(1)
\end{aligned}$$

if δ is chosen smaller than $1/\tau$. □

5 Tail Empirical Process With Estimated Parameters: Proofs

In this section we prove the approximation to the tail empirical process with estimated parameters stated in Proposition 2.1 and the limit theorem 2.2 for the test statistic T_n. To this end, we need a sequence of lemmas.
Define

$$\begin{aligned}
A_{n,k_n} &:= \frac{\hat a(n/k_n)}{a(n/k_n)},\\
B_{n,k_n} &:= \frac{\hat b(n/k_n)-b(n/k_n)}{a(n/k_n)},\\
y_n(x) &:= \Big(1+\gamma\big(B_{n,k_n}+A_{n,k_n}\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}\big)\Big)^{-1/\gamma}.
\end{aligned}$$

Recall from (2.4) that

$$\begin{aligned}
A_{n,k_n} &= 1+k_n^{-1/2}\alpha(W_n)+o_P(k_n^{-1/2}),\\
B_{n,k_n} &= k_n^{-1/2}\beta(W_n)+o_P(k_n^{-1/2}),\\
\hat\gamma_n &= \gamma+k_n^{-1/2}\Gamma(W_n)+o_P(k_n^{-1/2}).
\end{aligned}\tag{5.1}$$

Lemma 5.1. *Suppose (5.1) holds. Let $\lambda_n>0$ be such that $\lambda_n\to 0$, and $k_n^{-1/2}\lambda_n^\gamma\to 0$ if $\gamma<0$, or $k_n^{-1/2}\log^2\lambda_n\to 0$ if $\gamma=0$.*
(i) If $\gamma>0$ then, for all $\epsilon>0$, $x^{-1/2+\epsilon}\big(\sqrt{k_n}(y_n(x)-x)-L_n^{(\gamma)}(x)\big)\xrightarrow{P}0$, and $x^{\epsilon-1}(y_n(x)-x)\xrightarrow{P}0$ as $n\to\infty$ uniformly for $x\in(0,1]$.
(ii) If $-1/2<\gamma\le 0$ then, for all $\epsilon>0$, $x^{-1/2+\epsilon}\big(\sqrt{k_n}(y_n(x)-x)-L_n^{(\gamma)}(x)\big)\xrightarrow{P}0$ and $(y_n(x)-x)/x\xrightarrow{P}0$ as $n\to\infty$ uniformly for $x\in[\lambda_n,1]$.

Proof: For $\gamma \neq 0$, define $\delta_n := 1 + \gamma B_{n,k_n} - A_{n,k_n}\gamma/\hat{\gamma}_n$, and $\Delta_n := \Delta_{n,x} := \delta_n \hat{\gamma}_n/(\gamma A_{n,k_n} x^{-\hat{\gamma}_n})$, so that $\delta_n = O_P(k_n^{-1/2})$.
(i) By the mean value theorem there exist $\theta_{n,x} \in (0,1]$ such that

$$
\begin{aligned}
y_n(x) &= \Big(1 + \gamma\big(B_{n,k_n} + A_{n,k_n}\frac{x^{-\hat{\gamma}_n} - 1}{\hat{\gamma}_n}\big)\Big)^{-1/\gamma} \\
&= \Big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} x^{-\hat{\gamma}_n} + \delta_n\Big)^{-1/\gamma} \\
&= \Big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} x^{-\hat{\gamma}_n}\Big)^{-1/\gamma} (1+\Delta_n)^{-1/\gamma} \\
&= \big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} x^{-\hat{\gamma}_n}\big)^{-1/\gamma} - \frac{1}{\gamma}\Big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} x^{-\hat{\gamma}_n}(1+\theta_{n,x}\Delta_n)\Big)^{-1/\gamma-1}\delta_n \\
&= \big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} x^{-\hat{\gamma}_n}\big)^{-1/\gamma} - \frac{1}{\gamma} x^{\hat{\gamma}_n(1/\gamma+1)}\delta_n(1+o_P(1)) \qquad (5.2)
\end{aligned}
$$

where the $o_P(1)$-term tends to 0 uniformly for $x \in (0,1]$. Hence again by the mean value theorem and (5.1), for some $\theta_{n,x} \in (0,1)$,

$$
\begin{aligned}
& y_n(x) - x \\
&= \Big(\big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n}\big)^{-1/\gamma} - 1\Big) x^{\hat{\gamma}_n/\gamma} + (x^{\hat{\gamma}_n/\gamma} - x) - \frac{1}{\gamma} x^{\hat{\gamma}_n(1/\gamma+1)}\delta_n(1+o_p(1)) \\
&= -\frac{1}{\gamma}(1+o_p(1))\big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} - 1\big) x^{\hat{\gamma}_n/\gamma} + x^{1+\theta_{n,x}(\hat{\gamma}_n/\gamma-1)} \log x \big(\frac{\hat{\gamma}_n}{\gamma} - 1\big) \\
&\quad - \frac{1}{\gamma} x^{\hat{\gamma}_n(1/\gamma+1)}\delta_n(1+o_p(1)) \\
&= \frac{1}{\gamma}(1+o_p(1)) x^{\hat{\gamma}_n/\gamma}\Big(\frac{\hat{\gamma}_n - \gamma}{\hat{\gamma}_n} A_{n,k_n} - (A_{n,k_n} - 1)\Big) \\
&\quad + x^{1+\theta_{n,x}(\hat{\gamma}_n/\gamma-1)} \log x \frac{\hat{\gamma}_n - \gamma}{\gamma} \qquad (5.3) \\
&\quad - \frac{1}{\gamma} x^{\hat{\gamma}_n(1/\gamma+1)}\Big(\gamma B_{n,k_n} + \frac{1}{\hat{\gamma}_n}(\hat{\gamma}_n - \gamma) - \frac{\gamma}{\hat{\gamma}_n}(A_{n,k_n} - 1)\Big)(1+o_p(1)).
\end{aligned}
$$

Now the first assertion is a straightforward consequence of (5.1). For example,

$$
\begin{aligned}
& x^{-1/2+\epsilon}\sqrt{k_n}\frac{1}{\gamma}(1+o_P(1)) x^{\hat{\gamma}_n/\gamma}\Big(\frac{\hat{\gamma}_n - \gamma}{\hat{\gamma}_n} A_{n,k_n} - (A_{n,k_n} - 1)\Big) \\
&= \frac{1}{\gamma} x^{-1/2+\epsilon} \exp\big((\hat{\gamma}_n/\gamma - 1)\log x\big) x \Big(\sqrt{k_n}\frac{\hat{\gamma}_n - \gamma}{\hat{\gamma}_n} A_{n,k_n} - \sqrt{k_n}(A_{n,k_n} - 1)\Big) \\
&= \frac{1}{\gamma} x^{-1/2+\epsilon} x \Big(\frac{\Gamma(W_n)}{\gamma} - \alpha(W_n)\Big)(1+o_P(1))
\end{aligned}
$$

uniformly for $x \in (0,1]$.

Moreover, in view of (5.4),

$$\begin{aligned}
&x^{\epsilon-1}(y_n(x)-x)\\
&=-\frac{1}{\gamma}(1+o_p(1))\big(\frac{\gamma}{\hat\gamma_n}A_{n,k_n}-1\big)x^{\hat\gamma_n/\gamma-1+\epsilon}+x^{\epsilon+\theta_{n,x}(\hat\gamma_n/\gamma-1)}\log x\big(\frac{\hat\gamma_n}{\gamma}-1\big)\\
&\qquad\qquad -\frac{1}{\gamma}x^{\hat\gamma_n-1+\epsilon+\hat\gamma_n/\gamma}\delta_n(1+o_p(1))\\
&\xrightarrow{P} 0
\end{aligned}$$

as $n\to\infty$ uniformly for $x\in(0,1]$.

(ii) First we consider the case $\gamma=0$. Then

$$\begin{aligned}
&y_n(x)-x\\
&=\exp\Big(-\big(B_{n,k_n}+A_{n,k_n}\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}\big)\Big)-x\\
&=x\Big(\exp\Big(-\big(B_{n,k_n}+A_{n,k_n}\big(\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}+\log x\big)-(A_{n,k_n}-1)\log x\big)\Big)-1\Big).
\end{aligned}\tag{5.4}$$

A Taylor expansion of $\gamma\mapsto x^{-\gamma}$ together with (5.1) yields

$$\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}=-\log x+\frac{1}{2}\hat\gamma_n\log^2 x+O_P(k_n^{-3/2}\log^3 x).\tag{5.5}$$

It follows that

$$\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}+\log x\xrightarrow{P}0$$

as $n\to\infty$ uniformly for $x\in[\lambda_n,1]$, since then $k_n^{-1/2}\log^2 x\le k_n^{-1/2}\log^2\lambda_n\to 0$. Hence

$$B_{n,k_n}+A_{n,k_n}\big(\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}+\log x\big)-(A_{n,k_n}-1)\log x\xrightarrow{P}0,$$

so that by (5.1) and (5.5) and the series representation of the exponential function

$$\begin{aligned}
&x^{-1/2+\epsilon}\sqrt{k_n}(y_n(x)-x)\\
&\quad=\quad -x^{-1/2+\epsilon}x\Big(\sqrt{k_n}B_{n,k_n}+\sqrt{k_n}A_{n,k_n}\big(\frac{x^{-\hat\gamma_n}-1}{\hat\gamma_n}+\log x\big)\\
&\qquad\qquad -\sqrt{k_n}(A_{n,k_n}-1)\log x\Big)(1+o_P(1))\\
&\quad=\quad -x^{-1/2+\epsilon}x\Big(\beta(W_n)+\frac{1}{2}\Gamma(W_n)\log^2 x-\alpha(W_n)\log x\Big)(1+o_P(1))
\end{aligned}$$

uniformly for $x \in [\lambda_n, 1]$, that is, the first assertion.
Likewise one concludes from (5.4), (5.1) and (5.5) that $(y_n(x) - x)/x$ tends to 0 uniformly for $x \in [\lambda_n, 1]$.

Next assume $-1/2 < \gamma < 0$.
Because $\delta_n = O_P(k_n^{-1/2})$ and, by the definition of λ_n and (5.1),

$$k_n^{-1/2} x^{\hat\gamma_n} \le k_n^{-1/2} \lambda_n^{\hat\gamma_n} = o\Big(\exp\big(\log \lambda_n(\hat\gamma_n - \gamma)\big)\Big) = o_P(1),$$

$\Delta_n \to 0$ in probability uniformly for $x \in [\lambda_n, 1]$. Therefore, the first assertion can be established as in the case $\gamma > 0$.
Furthermore, according to (5.4),

$$\begin{aligned}
&\frac{y_n(x) - x}{x} \\
&= -\frac{1}{\gamma}(1 + o_p(1))\big(\frac{\gamma}{\hat\gamma_n} A_{n,k_n} - 1\big) x^{\hat\gamma_n/\gamma - 1} + x^{\theta_{n,x}(\hat\gamma_n/\gamma - 1)} \log x \big(\frac{\hat\gamma_n}{\gamma} - 1\big) \\
&\quad - \frac{1}{\gamma} x^{\hat\gamma_n + \hat\gamma_n/\gamma - 1} \delta_n (1 + o_p(1)) \\
&= -\frac{1}{\gamma}\big(\frac{\gamma}{\hat\gamma_n} A_{n,k_n} - 1\big) \exp\Big(\frac{\sqrt{k_n}(\hat\gamma_n - \gamma)}{\gamma} \cdot \frac{\log x}{\sqrt{k_n}}\Big)(1 + o_p(1)) \\
&\quad + \frac{\log x}{\sqrt{k_n}} \frac{\sqrt{k_n}(\hat\gamma_n - \gamma)}{\gamma} \exp\Big(\frac{\theta_{n,x}\sqrt{k_n}(\hat\gamma_n - \gamma)}{\gamma} \frac{\log x}{\sqrt{k_n}}\Big) \\
&\quad - \frac{1}{\gamma} \exp\Big(\frac{\sqrt{k_n}(\hat\gamma_n - \gamma)}{\gamma} \frac{\log x}{\sqrt{k_n}} + \sqrt{k_n}(\hat\gamma_n - \gamma) \frac{\log x}{\sqrt{k_n}}\Big) \frac{x^\gamma}{\sqrt{k_n}} \cdot \sqrt{k_n}\delta_n \\
&\xrightarrow{P} 0
\end{aligned}$$

as $n \to \infty$ uniformly for $x \in [\lambda_n, 1]$ by the choice of λ_n. □

Lemma 5.2. *Under the conditions of Lemma 5.1 one has for all $\epsilon > 0$:*
(i) If $\gamma > 0$, then $x^{-1/2+\epsilon}\big(W_n(y_n(x)) - W_n(x)\big) \xrightarrow{P} 0$ as $n \to \infty$ uniformly for $x \in (0, 1]$.
(ii) If $-1/2 < \gamma \le 0$, then $x^{-1/2+\epsilon}\big(W_n(y_n(x)) - W_n(x)\big) \xrightarrow{P} 0$ as $n \to \infty$ uniformly for $x \in [\lambda_n, 1]$.

Proof:

Let $u_n > 0$, $n \in \mathbb{N}$, be an arbitrary sequence converging to 0. According to Lemma 5.1 and the law of iterated logarithm

$$x^{-1/2+\epsilon}\big(W_n(y_n(x)) - W_n(x)\big) = \frac{(y_n(x))^{1/2-\epsilon/2}}{x^{1/2-\epsilon}} \frac{W_n(y_n(x))}{(y_n(x))^{1/2-\epsilon/2}} - \frac{W_n(x)}{x^{1/2-\epsilon}} \xrightarrow{P} 0$$

uniformly for $x \in (0, u_n]$ if $\gamma > 0$, and uniformly for $x \in [\lambda_n, u_n]$ if $-1/2 < \gamma \leq 0$. Since, due to the continuity of W_n and Lemma 5.1,

$$\sup_{u \leq x \leq 1} x^{-1/2+\epsilon} \big| W_n(y_n(x)) - W_n(x) \big| \xrightarrow{P} 0$$

for all $u \in (0, 1]$, the assertion follows by a standard diagonal argument. □

Lemma 5.3. *Under the conditions of Lemma 5.1 one has for all $\epsilon > 0$:*
(i) For $\gamma > 0$

$$x^{-1/2+\epsilon} \Big((y_n(x))^{\gamma+1} K_{\gamma,\rho}\big(\frac{1}{y_n(x)}\big) - x^{\gamma+1} K_{\gamma,\rho}\big(\frac{1}{x}\big) \Big) \xrightarrow{P} 0$$

as $n \to \infty$ uniformly for $x \in (0, 1]$.
(ii) For $-1/2 < \gamma \leq 0$

$$x^{-1/2+\epsilon} \Big((y_n(x))^{\gamma+1} K_{\gamma,\rho}\big(\frac{1}{y_n(x)}\big) - x^{\gamma+1} K_{\gamma,\rho}\big(\frac{1}{x}\big) \Big) \xrightarrow{P} 0$$

as $n \to \infty$ uniformly for $x \in [\lambda_n, 1]$.

Proof:

(i) We only consider the case $\gamma > 0 = \rho$; the assertion can be proved similarly in the case $\gamma > 0 > \rho$. Equation (5.2) implies

$$\begin{aligned} \log \frac{y_n(x)}{x} &= \log \Big(\big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} \big)^{-1/\gamma} x^{\hat{\gamma}_n/\gamma - 1} \Big) (1 + o_P(1)) \\ &= \Big(\big(\frac{\hat{\gamma}_n}{\gamma} - 1 \big) \log x - \frac{1}{\gamma} \log \big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} \big) \Big) (1 + o_P(1)) \end{aligned}$$

uniformly for $x \in (0, 1]$. Hence, by the definition of $K_{\gamma,0}$ and Lemma 5.1(i),

$$\begin{aligned} & x^{-1/2+\epsilon} \Big((y_n(x))^{\gamma+1} K_{\gamma,0}\big(\frac{1}{y_n(x)}\big) - x^{\gamma+1} K_{\gamma,0}\big(\frac{1}{x}\big) \Big) \\ &= x^{-1/2+\epsilon} \Big(- \frac{y_n(x) \log(y_n(x))}{\gamma} + \frac{x \log x}{\gamma} \Big) \\ &= -\frac{1}{\gamma} \Big(x^{-1/2+\epsilon} y_n(x) \log \frac{y_n(x)}{x} + x^{-1/2+\epsilon} (y_n(x) - x) \log x \Big) \\ &= -\frac{1}{\gamma} \Bigg(x^{\epsilon-1} y_n(x) x^{1/2} \log x \big(\frac{\hat{\gamma}_n}{\gamma} - 1 \big) (1 + o_P(1)) \\ &\quad - \frac{1}{\gamma} x^{\epsilon-1} y_n(x) x^{1/2} \log \big(\frac{\gamma}{\hat{\gamma}_n} A_{n,k_n} \big) (1 + o_P(1)) + x^{\epsilon-1} (y_n(x) - x) x^{1/2} \log x \Bigg) \\ &\xrightarrow{P} 0 \end{aligned}$$

as $n \to \infty$ uniformly for $x \in (0,1]$.
(ii) In the case $\gamma = 0 > \rho$, according to the definition of $K_{0,\rho}$, Lemma 5.1(ii) and the mean value theorem, there exists $\theta_{n,x} \in (0,1)$ such that

$$\begin{aligned}
& x^{-1/2+\epsilon}\Big(y_n(x)K_{0,\rho}(\frac{1}{y_n(x)}) - xK_{0,\rho}(\frac{1}{x})\Big) \\
= \;& x^{-1/2+\epsilon}\Big(\frac{(y_n(x))^{1-\rho}}{\rho} - \frac{x^{1-\rho}}{\rho}\Big) \\
= \;& \frac{1-\rho}{\rho} x^{1/2+\epsilon}\frac{y_n(x)-x}{x}\big(x+\theta_{n,x}(y_n(x)-x)\big)^{-\rho} \\
\stackrel{P}{\to} \;& 0
\end{aligned}$$

as $n \to \infty$ uniformly for $x \in [\lambda_n, 1]$.
In the other cases the assertion can be proved likewise. □

Remark 5.1. *The part (ii) of Lemma 5.1 with weight function $x^{\epsilon-1-\gamma}$ instead of $x^{-1/2+\epsilon}$, and of the Lemmas 5.2 and 5.3 also hold true for $-1 < \gamma \le 0$.*

Lemma 5.4. *Suppose $p_n \to 0$, $np_n \to 0$, and $k_n^{-1/2}\log^2(np_n) \to 0$ as $n \to \infty$. Define*

$$\hat{x}_{p_n} := \frac{(\frac{k_n}{np_n})^{\hat{\gamma}_n} - 1}{\hat{\gamma}_n}\hat{a}(\frac{n}{k_n}) + \hat{b}(\frac{n}{k_n}).$$

Then, under the conditions of Proposition 2.1 for $-\frac{1}{2} < \gamma \le 0$, $P\{\hat{x}_{p_n} \le X_{n,n}\} \to 0$ as $n \to \infty$.

Proof: According to Theorem 1 of de Haan and Stadtmüller (1996), one has

$$\frac{\frac{a(tx)}{x^\gamma a(t)} - 1}{A(t)} \to \frac{x^\rho - 1}{\rho}$$

as $t \to \infty$. By similar arguments as used by Drees (1998) and Cheng and Jiang (2001) it follows that, for all $0 < \epsilon < \frac{1}{2}$, there exists $t_\epsilon > 0$ such that for all $t \ge t_\epsilon$ and $x \ge 1$

$$\left|\frac{\frac{a(tx)}{x^\gamma a(t)} - 1}{A(t)} - \frac{x^\rho - 1}{\rho}\right| \le \epsilon x^{\rho+\epsilon}.$$

Hence

$$\frac{a(n)}{k_n^\gamma a(n/k_n)} = 1 + A\Big(\frac{n}{k_n}\Big)\frac{k_n^\rho - 1}{\rho} + o\Big(A\Big(\frac{n}{k_n}\Big)k_n^{\rho+\epsilon}\Big) \to 1 \qquad (5.6)$$

because $\rho \le 0$ and $\sqrt{k_n}A(n/k_n) = O(1)$.

Now, we distinguish two cases.

Case (i): $-1/2 < \gamma < 0$.
Then

$$\begin{aligned}
&\frac{\hat{x}_{p_n} - X_{n,n}}{a(n/k_n)} \\
&= -\frac{1}{\gamma}\Big(\frac{\hat{a}(n/k_n)}{a(n/k_n)} - 1\Big) + \frac{1}{\hat{\gamma}_n}\frac{\hat{a}(n/k_n)}{a(n/k_n)}\Big(\frac{k_n}{np_n}\Big)^{\hat{\gamma}_n} + \frac{\hat{a}(n/k_n)}{a(n/k_n)}\Big(\frac{1}{\gamma} - \frac{1}{\hat{\gamma}_n}\Big) \\
&\quad + \frac{\hat{b}(n/k_n) - b(n/k_n)}{a(n/k_n)} - \Big(\frac{b(n) - b(n/k_n)}{a(n/k)} + \frac{1}{\gamma}\Big) \\
&\quad - \frac{X_{n,n} - b(n)}{a(n)} \cdot \frac{a(n)}{a(n/k_n)} \\
&=: T_1 + T_2 + T_3 + T_4 - T_5 - T_6.
\end{aligned}$$

Assumption (5.1) implies $T_1 + T_3 + T_4 = O_P(k_n^{-1/2}) = o_p(k_n^\gamma)$ and

$$T_2 = O_P\bigg(\Big(\frac{k_n}{np_n}\Big)^\gamma \exp\Big((\hat{\gamma}_n - \gamma)\log\frac{k_n}{np_n}\Big)\bigg) = O_P\bigg(\Big(\frac{k_n}{np_n}\Big)^\gamma\bigg) = o_P(k_n^\gamma)$$

because $np_n \to 0$ and $k_n^{-1/2}\log(np_n) \to 0$.
Since, in view of (5.6) and the definition of $b(n)$,

$$\frac{U(n) - b(n)}{a(n/k_n)} = \frac{a(n)}{a(n/k_n)} \cdot \frac{A(n)}{\gamma + \rho} 1\{\rho < 0\} = o(k_n^\gamma),$$

approximation (2.3) yields

$$T_5 = \frac{k_n^\gamma - 1}{\gamma} + o\Big(k_n^{\gamma+\rho+\epsilon} A\big(\frac{n}{k_n}\big)\Big) + o(k_n^\gamma) + \frac{1}{\gamma} = \frac{k_n^\gamma}{\gamma} + o(k_n^\gamma).$$

Finally, $k_n^{-\gamma} T_6$ converges to G_γ in distribution because of $F \in D(G_\gamma)$ and (5.6).
Summing up, one obtains

$$\frac{\hat{x}_{p_n} - X_{n,n}}{k_n^\gamma a(n/k_n)} \xrightarrow{d} -\Big(M + \frac{1}{\gamma}\Big)$$

for a G_γ-distributed r.v. M. Now the assertion follows from the fact that $-(M + 1/\gamma) > 0$ a.s.

Case (ii): $\gamma = 0$.
By similar arguments as in the first case one obtains

$$\begin{aligned}
&\frac{\hat{x}_{p_n} - X_{n,n}}{a(n/k_n)} \\
&= \left(\frac{\left(\frac{k_n}{np_n}\right)^{\hat{\gamma}_n} - 1}{\hat{\gamma}_n} - \log\frac{k_n}{np_n}\right)\frac{\hat{a}(n/k_n)}{a(n/k_n)} + \left(\frac{\hat{a}(n/k_n)}{a(n/k_n)} - 1\right)\log k_n \\
&\quad + \frac{\hat{a}(n/k_n)}{a(n/k_n)}\log\frac{1}{np_n} + \frac{\hat{b}(n/k_n) - b(n/k_n)}{a(n/k_n)} \\
&\quad - \left(\frac{b(n) - b(n/k_n)}{a(n/k)} - \log k_n\right) - \frac{X_{n,n} - b(n)}{a(n)} \cdot \frac{a(n)}{a(n/k_n)} \\
&= o_P(1) + o_P(1) + \log\frac{1}{np_n}(1 + o_P(1)) + O_P(k_n^{-1/2}) + o(1) + O_P(1) \\
&= \log\frac{1}{np_n}(1 + o_P(1)) \\
&\stackrel{P}{\to} \infty
\end{aligned}$$

from which the assertion is obvious. □

Proof of Proposition 2.1:

Recall the definition

$$y_n(x) := \left(1 + \gamma\left(\frac{\hat{b}(\frac{n}{k_n}) - b(\frac{n}{k_n})}{a(\frac{n}{k_n})} + \frac{\hat{a}(\frac{n}{k_n})}{a(\frac{n}{k_n})}\frac{x^{-\hat{\gamma}_n} - 1}{\hat{\gamma}_n}\right)\right)^{-\frac{1}{\gamma}}.$$

Observe that

$$\begin{aligned}
I :&= x^{-1/2+\epsilon}\left(\sqrt{k_n}\left[\frac{n}{k_n}\bar{F}_n\left(\hat{a}\left(\frac{n}{k_n}\right)\frac{x^{-\hat{\gamma}_n} - 1}{\hat{\gamma}_n} + \hat{b}\left(\frac{n}{k_n}\right)\right) - x\right]\right. \\
&\qquad \left. - W_n(x) - L_n^{(\gamma)}(x) - \sqrt{k_n}A\left(\frac{n}{k_n}\right)x^{\gamma+1}K_{\gamma,\rho}\left(\frac{1}{x}\right)\right) \\
&= \frac{x^{-1/2+\epsilon}}{(y_n(x))^{-1/2+\epsilon/2}}(y_n(x))^{-1/2+\epsilon/2}. \\
&\qquad \cdot \left(\sqrt{k_n}\left[\frac{n}{k_n}\bar{F}_n\left(a\left(\frac{n}{k_n}\right)\frac{(y_n(x))^{-\gamma} - 1}{\gamma} + b\left(\frac{n}{k_n}\right)\right) - y_n(x)\right]\right. \\
&\qquad \left. - W_n(y_n(x)) - \sqrt{k_n}A\left(\frac{n}{k_n}\right)(y_n(x))^{\gamma+1}K_{\gamma,\rho}\left(\frac{1}{y_n(x)}\right)\right)
\end{aligned}$$

$$
\begin{aligned}
&+ x^{-1/2+\epsilon}\Big(\sqrt{k_n}(y_n(x)-x) - L_n^{(\gamma)}(x)\Big)\\
&+ x^{-1/2+\epsilon}\Big(W_n(y_n(x)) - W_n(x)\Big)\\
&+ x^{-1/2+\epsilon}\Big(\sqrt{k_n}A\big(\frac{n}{k_n}\big)(y_n(x))^{\gamma+1}K_{\gamma,\rho}\big(\frac{1}{y_n(x)}\big) - \sqrt{k_n}A\big(\frac{n}{k_n}\big)x^{\gamma+1}K_{\gamma,\rho}\big(\frac{1}{x}\big)\Big)\\
&:= I_1 + I_2 + I_3 + I_4
\end{aligned}
$$

Now we distinguish three cases.

Case (i): $\gamma > 0$.
By Lemma 5.1(i), $\sup_{x\in(0,1]} x^{-1/2+\epsilon}/(y_n(x))^{-1/2+\epsilon/2}$ is stochastically bounded. Combining this with Theorem 2.1, we obtain $\sup_{x\in(0,1]} |I_1| \to 0$ in probability as $n \to \infty$. An application of Lemma 5.1(i), Lemma 5.2(i), and Lemma 5.3(i) gives

$$
\sup_{x\in(0,1]} |I_2| \xrightarrow{d} 0, \qquad \sup_{x\in(0,1]} |I_3| \xrightarrow{P} 0, \qquad \sup_{x\in(0,1]} |I_4| \xrightarrow{P} 0,
$$

respectively. Hence $\sup_{x\in(0,1]} |I| \to 0$ in probability as $n \to \infty$.

Case (ii) $-1/2 < \gamma < 0$, or $\gamma = 0$ and $\rho < 0$.
Let $\lambda_n := 1/(k_n \log k_n)$. Obviously $\lambda_n \to 0$, $k_n^{-1/2}\lambda_n^{\gamma} \to 0$ and $k_n^{-1/2}\log^2\lambda_n \to 0$ as $n \to \infty$, so that Lemmas 5.1, 5.2 and 5.3 apply. Like in case (i), we obtain $\sup_{x\in(\lambda_n,1]} |I| \to 0$ in probability as $n \to \infty$.
It remains to prove that $\sup_{x\in(0,\lambda_n]} |I| \to 0$ in probability. To this end, let $p_n := 1/(n \log k_n)$, so that $np_n \to 0$ and $k_n^{-1/2}\log^2(np_n) \to 0$ as $n \to \infty$. Thus, for $x \in (0, \lambda_n]$,

$$
\begin{aligned}
z_n(x) := \hat{a}\big(\frac{n}{k_n}\big)\frac{x^{-\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}\big(\frac{n}{k_n}\big) \quad &\geq \quad \hat{a}\big(\frac{n}{k_n}\big)\frac{\lambda_n^{-\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}\big(\frac{n}{k_n}\big)\\
&= \quad \hat{a}\big(\frac{n}{k_n}\big)\frac{\big(\frac{k_n}{np_n}\big)^{\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}\big(\frac{n}{k_n}\big),
\end{aligned}
$$

so that by Lemma 5.4,

$$
P\Big\{z_n(x) < X_{n,n} \quad \text{for some } x \in (0,\lambda_n]\Big\} \to 0.
$$

Let

$$
\tau_n := \sup_{x\in(0,\lambda_n]} x^{-1/2+\epsilon}\frac{n}{\sqrt{k_n}}\bar{F}_n\Big(\hat{a}\big(\frac{n}{k_n}\big)\frac{x^{-\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}\big(\frac{n}{k_n}\big)\Big).
$$

By the definition of $\bar{F}_n$, $z_n(x) < X_{n,n}$ for some $x \in (0, \lambda_n]$ implies $\tau \neq 0$. Therefore,

$$
P\{\tau_n \neq 0\} \to 0 \tag{5.7}
$$

as $n \to \infty$.
Furthermore, it is easy to check that

$$\begin{aligned} x^{-1/2+\epsilon}\sqrt{k_n}x \to 0, \qquad & x^{-1/2+\epsilon}W_n(x) \xrightarrow{P} 0, \\ x^{-1/2+\epsilon}L_n^{(\gamma)}(x) \xrightarrow{P} 0, \quad & x^{-1/2+\epsilon}\sqrt{k_n}A\Big(\frac{n}{k_n}\Big)x^{\gamma+1}K_{\gamma,\rho}\Big(\frac{1}{x}\Big) \to 0 \end{aligned} \tag{5.8}$$

uniformly for $x \in (0, \lambda_n]$ as $n \to \infty$. For example, the second convergence is an immediate consequence of the law of the iterated logarithm, and in the case $-1/2 < \gamma < 0$

$$\begin{aligned} & \sup_{x\in(0,\lambda_n]} x^{-1/2+\epsilon}|L_n^{(\gamma)}(x)| \\ \le & \sup_{x\in(0,\lambda_n]} \frac{1}{|\gamma|}x^{1/2+\epsilon}\Big|\frac{1}{\gamma}\Gamma(W_n) - \alpha(W_n)\Big| + \sup_{x\in(0,\lambda_n]} \frac{1}{|\gamma|}|\Gamma(W_n)|x^{1/2+\epsilon}\log x \\ & + \sup_{x\in(0,\lambda_n]} \frac{1}{|\gamma|}x^{1/2+\gamma+\epsilon}\Big|\gamma\beta(W_n) + \frac{1}{\gamma}\Gamma(W_n) - \alpha(W_n)\Big| \\ \xrightarrow{P} & \; 0. \end{aligned}$$

In view of (5.7) and (5.8), the assertion $\sup_{x\in(0,\lambda_n]}|I| \to 0$ in probability is immediate.

Case (iii): $\gamma = \rho = 0$.
According to Lemma 5.1, $y_n(x)/x \to 1$ in probability uniformly for $x \in [\lambda_n, 1]$ with $\lambda_n := 1/(k_n \log k_n)$, so that

$$\frac{(1+|\log x|)^\tau}{(1+|\log y_n(x)|)^\tau} = \Big(\frac{1+|\log x|}{1+|\log x| + o_P(1)}\Big)^\tau = O_P(1)$$

uniformly for $x \in [\lambda_n, 1]$. Therefore, one can argue as in case (ii) (using Corollary 2.1 instead of Theorem 2.1) to establish the assertion. □

Proof of Theorem 2.2:

By Proposition 2.1 one has

$$\begin{aligned} & \Big(\sqrt{k_n}\Big[\frac{n}{k_n}\bar{F}_n\Big(\hat{a}\Big(\frac{n}{k_n}\Big)\frac{x^{-\hat{\gamma}_n}-1}{\hat{\gamma}_n} + \hat{b}\Big(\frac{n}{k_n}\Big)\Big) - x\Big]\Big)^2 \\ = & \Big(W_n(x) + L_n^{(\gamma)}(x) + \sqrt{k_n}A\Big(\frac{n}{k_n}\Big)x^{\gamma+1}K_{\gamma,\rho}\Big(\frac{1}{x}\Big) + \frac{o_p(1)}{h(x)}\Big)^2 \end{aligned} \tag{5.9}$$

Using the law of iterated logarithm, it is readily checked that

$$\begin{aligned}\int_0^1 \left(W_n(x) + L_n^{(\gamma)}\right)^2 x^{\eta-2}\, dx &= O_P(1) \\ \int_0^1 \left(x^{\gamma+1} K_{\gamma,\rho}\left(\frac{1}{x}\right)\right)^2 x^{\eta-2}\, dx &< \infty \\ \int_0^1 \frac{x^{\eta-2}}{h^2(x)}\, dx &< \infty\end{aligned}$$

for $\eta > 0$, and $\eta \geq 1$ if $\gamma = \rho = 0$. Hence the assertion is an immediate consequence of (5.9) and $\sqrt{k_n} A(n/k_n) \to 0$. □

6 Simulations

First we want to determine the limiting distribution of the test statistic $k_n T_n$ defined by (2.6), where we use the maximum likelihood estimator $\hat{\gamma}_n$, $\hat{a}(n/k_n)$ and $\hat{b}(n/k_n)$ described in Example 2.1. Here we have chosen $\eta = 1$, thus giving maximal weight to deviations in the extreme tail region that is possible in the framework of Theorem 2.2 for all values of $\gamma > -1/2$.
To simulate $\int_0^1 (W_n(x) + L_n^{(\gamma)}(x))^2 x^{-1}\, dx$, the Brownian motion W_n on the unit interval is simulated on a grid with 50 000 points. Then the integral is approximated by a Riemann sum for the extreme value indices $\gamma = 2, 1.5, 1, 0.5, 0.25, 0, -0.25, -0.375$ and -0.5. Note that for $\gamma < -0.5$ the term $L_n^{(\gamma)}$ is not defined since the integral $S_n = \int_0^1 t^{\gamma-1} W_n(t)\, dt$ defined in Example 2.1 may not exist. The empirical quantiles of the integral statistic obtained in 20 000 runs are reported in Table 6. It is not surprising that the extreme upper quantiles increase rapidly as $\gamma < 0$ decreases, since $|S_n| \to \infty$ in probability as $\gamma \downarrow -1/2$, and thus the limit distribution of $k_n T_n$ converges weakly to ∞, too.
Next we investigate the finite sample behavior of the test described in Section 2, that rejects the hypothesis that $F \in D(G_\gamma)$ for some $\gamma > -1/2$ if $k_n T_n$ exceeds $\hat{Q}_{1-\bar{\alpha},\tilde{\gamma}_n}$. Here we use the maximum likelihood estimator for γ also as the pilot estimator, that is, $\tilde{\gamma}_n = \hat{\gamma}_n$. Since we have approximately determined the quantiles $Q_{p,\gamma}$ only for 9 different values of γ, we use linear interpolation to approximate the quantiles for intermediate values of γ, that is, for $\tilde{\gamma}_n \in [\gamma_1, \gamma_2]$ we define

$$\hat{Q}_{p,\tilde{\gamma}_n} = Q_{p,\gamma_1} + \frac{\tilde{\gamma}_n - \gamma_1}{\gamma_2 - \gamma_1}(Q_{p,\gamma_2} - Q_{p,\gamma_1})$$

where Q_{p,γ_i} denote the quantiles given in Table 6. Moreover, we define $\hat{Q}_{p,\tilde{\gamma}_n} := Q_{p,2}$ if $\tilde{\gamma}_n > 2$.

$\gamma =$ p	2	1.5	1	0.5	0.25	0	-0.25	-0.375	-0.5
0.995	0.545	0.513	0.507	0.525	0.553	0.621	0.672	0.739	0.909
0.99	0.477	0.462	0.459	0.474	0.494	0.554	0.604	0.667	0.795
0.975	0.408	0.389	0.383	0.390	0.409	0.459	0.510	0.558	0.657
0.95	0.349	0.337	0.330	0.337	0.355	0.390	0.431	0.468	0.552
0.9	0.289	0.281	0.278	0.285	0.295	0.318	0.355	0.381	0.444
0.8	0.231	0.227	0.224	0.229	0.239	0.254	0.280	0.299	0.343
0.7	0.197	0.193	0.191	0.195	0.201	0.213	0.235	0.253	0.286
0.6	0.171	0.168	0.166	0.169	0.175	0.185	0.204	0.217	0.243
0.5	0.151	0.148	0.147	0.149	0.154	0.162	0.178	0.189	0.211
0.4	0.132	0.131	0.130	0.132	0.136	0.144	0.157	0.164	0.183
0.3	0.116	0.114	0.114	0.116	0.120	0.126	0.135	0.144	0.158
0.2	0.100	0.099	0.098	0.100	0.103	0.108	0.116	0.122	0.134
0.1	0.083	0.082	0.081	0.082	0.085	0.089	0.095	0.099	0.106
0.05	0.071	0.070	0.070	0.071	0.073	0.078	0.080	0.083	0.090
0.025	0.062	0.062	0.062	0.063	0.064	0.068	0.071	0.073	0.078
0.01	0.053	0.054	0.054	0.055	0.056	0.059	0.060	0.062	0.067
0.005	0.048	0.049	0.049	0.050	0.051	0.052	0.054	0.055	0.060

Table 1: Quantiles $Q_{p,\gamma}$ of the limit distribution of $k_n T_n$.

As usually in extreme value theory, the choice of the number k_n of order statistics used for the inference is a crucial point. Here we consider $k_n = 25, 50$ and 75 for sample size $n = 200$, and $k_n = 25, 50, \ldots, 200$ for sample size $n = 1000$.

1000 samples were drawn from each of the following distribution functions belonging to the domain of attraction of G_γ for some $\gamma > -1/2$:

- Cauchy distribution $\quad (\gamma = 1,\ \rho = -2)$:

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan x, \qquad x \in \mathbb{R}.$$

- Burr(β, τ, λ) distribution $\quad (\gamma = 1/(\tau\lambda),\ \rho = -1/\lambda)$:

$$F(x) = 1 - \left(\frac{\beta}{\beta + x^\tau}\right)^\lambda, \qquad x > 0,$$

  with $(\beta, \tau, \lambda) = (1, 2, 2)$.

- Extreme Value distribution $EV(\gamma) \quad (\gamma \in \mathbb{R}, \rho = -1)$:

$$F(x) = \exp\big(-(1 + \gamma x)^{-1/\gamma}\big), \quad 1 + \gamma x > 0.$$

- Weibull(λ, τ) distribution $(\gamma = 0, \rho = 0)$:

$$F(x) = 1 - \exp(-\lambda x^\tau), \quad x > 0$$

with $(\lambda, \tau) = (1, 0.5)$.

- Reversed Burr(β, τ, λ) distribution $(\gamma = -1/(\tau\lambda),\ \rho = -1/\lambda)$:

$$F(x) = 1 - \left(\frac{\beta}{\beta + (x_+ - x)^{-\tau}}\right)^\lambda, \quad x < x_+$$

with $(\beta, \tau, \lambda) = (1, 4, 1)$ and $x_+ = 1$.

In some simulations either there exists no solution to the likelihood equations, or the maximum likelihood estimate of γ is less than $-1/2$, so that the test cannot be applied. The relative frequency of simulations in which this happened are given in the Tables 6–6; for all other values of k_n not mentioned in these tables, the test could be performed in all simulations.
For the reversed Burr distribution, one gets estimates of γ less than $-1/2$ in at least 1% of the simulations for all values of k_n and in about one third of all simulations if $n = 200$, while for all other distributions this happened only if a small proportion of the data is used for the inference. It is clear that the problem of pilot estimates of γ being smaller than $-1/2$ becomes more and more acute as the true extreme value index approaches $-1/2$; this is particularly true for small sample sizes.
In the Tables 6 and 6 the empirical size of the test with nominal size $\bar{\alpha} = 0.05$ is reported, that is, the relative frequency of simulations in which the hypothesis is rejected. These frequencies are based only on those simulations in which the test could actually be applied. The overall impression is that the empirical size of the test is quite close to the nominal value for a wide range of values of k_n. Hence, as far as the size is concerned, the test is rather insensitive to the choice of the proportion of the data used for testing, although for very small k_n the test seems a bit too conservative. This conclusion is also supported by Figure 1 that displays the empirical size of the test for the Cauchy distribution and sample size $n = 1000$.
At first glance, it might be surprising that, unlike estimators of γ, the test behaves equally well for small and large values of $|\rho|$. However, recall that for the actual size to be close to the nominal value it is not important how accurate the estimators are but only how precise the Gaussian approximation for the tail empirical distribution function with estimated parameters is. While the rate of convergence of estimators of the extreme value index deteriorates as ρ tends to 0, this is not necessarily true for the accuracy of the normal approximation.

To sum up, if $F \in D(G_\gamma)$ with γ not too close to $-1/2$, then the empirical size of the test is close to the nominal value for a wide range of values of k_n.

Acknowledgment: Part of the work of the first two authors was done while visiting the Stochastics Center at Chalmers University Gothenburg. Grateful acknowledgement is made for hospitality particularly to Holger Rootzén. During that time Holger Drees was supported by the European Union TMR grant ERB-FMRX-CT960095. His work was also partly supported by the Netherlands Organization for Scientific Research through the Netherlands Mathematical Research Foundation and by the Heisenberg program of the DFG.

References

[1] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation.* Cambridge University Press, Cambridge.

[2] Cheng, S. and Jiang, C. (2001). The Edgeworth expansion for distributions of extreme values. *Sci. China* **44**, 427-437.

[3] Csörgő, M. and Horváth, L. (1993). *Weighted approximations in probability and statistics.* Wiley, Chichester.

[4] Dietrich, D., de Haan, L. and Hüsler, J. (2001). Testing extreme value conditions. To appear in *Extremes.*

[5] Donsker, M.D. (1952). Justification and extension of Doob's heuristic approach to the Komogorov-Smirnov theorems. *Ann. Math. Statistics* **23**, 277–281.

[6] Doob, J.L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statistics* **20**, 393–403.

[7] Drees, H. (1998). On smooth statistical tail functionals. *Scand. J. Statist.* **25**, 187-210.

[8] Drees, H., Ferreira, A. and de Haan, L. (2002). On maximum likelihood estimation of the extreme value index. *Preprint, submitted.*

[9] Einmahl, J.H.J. (1997). Poisson and Gaussian approximation of weighted local empirical processes. *Stoch. Proc. Appl.* **70**, 31-58.

[10] de Haan, L. and Resnick, S.I. (1998). On the asymptotic normality of the Hill estimator. *Stochastic Models* **14**, 849–867.

k_n	Cauchy $\gamma = 1$, $\rho = -2$	Burr(1,2,2) $\gamma = 0.25$, $\rho = -0.5$	EV(0.25) $\gamma = 0.25$, $\rho = -1$	EV(0) $\gamma = 0$, $\rho = -1$	Weibull(1,0.5) $\gamma = 0$, $\rho = 0$	Rev. Burr(1,4,1) $\gamma = -0.25$ $\rho = -1$
25	0.000	0.001	0.004	0.008	0.001	0.096
50	0.000	0.000	0.000	0.000	0.000	0.008
75	0.000	0.000	0.000	0.000	0.000	0.001

Table 2: Relative frequency of simulations in which no maximum likelihood estimate was found for sample size $n = 200$

k_n	Cauchy	Burr(1,2,2)	EV(0.25)	EV(0)	Weibull(1,0.5)	R-Buur(1,4,1)
25	0.000	0.001	0.026	0.104	0.017	0.340
50	0.000	0.000	0.000	0.021	0.000	0.332
75	0.000	0.000	0.000	0.009	0.000	0.374

Table 3: Relative frequency of simulations in which $\hat{\gamma}_n < -0.5$ for sample size $n = 200$

k_n	Cauchy	Burr(1,2,2)	EV(0.25)	EV(0)	Weibull(1,0.5)	Rev. Burr(1,4,1)
25	0.000	0.000	0.003	0.019	0.006	0.067

Table 4: Relative frequency of simulations in which no maximum likelihood estimate was found for sample size $n = 1000$

k_n	Cauchy	Burr(1,2,2)	EV(0.25)	EV(0)	Weibull(1,0.5)
25	0.002	0.004	0.014	0.081	0.029
50	0.000	0.000	0.001	0.024	0.000
75	0.000	0.000	0.000	0.004	0.000

k_n	25	50	75	100	125	150	175	200
Rev. Burr(1,4,1)	0.293	0.175	0.082	0.057	0.038	0.019	0.018	0.016

Table 5: Relative frequency of simulations in which $\hat{\gamma}_n < -0.5$ for sample size $n = 1000$

k_n	Cauchy	Burr(1,2,2)	EV(0.25)	EV(0)	Weibull(1,0.5)	Rev. Burr(1,4,1)
25	0.042	0.042	0.041	0.037	0.049	0.016
50	0.056	0.044	0.042	0.035	0.060	0.027
75	0.098	0.045	0.054	0.043	0.072	0.059

Table 6: Empirical size the one-sided test with nominal size $\bar{\alpha} = 0.05$ for sample size $n = 200$.

k_n	Cauchy	Burr(1,2,2)	EV(0.25)	EV(0)	Weibull(1,0.5)	Rev. Burr(1,4,1)
25	0.040	0.043	0.031	0.030	0.033	0.020
50	0.040	0.045	0.033	0.033	0.031	0.017
75	0.043	0.047	0.038	0.044	0.047	0.022
100	0.044	0.049	0.042	0.051	0.042	0.042
125	0.043	0.049	0.053	0.038	0.047	0.044
150	0.044	0.057	0.054	0.041	0.061	0.039
175	0.062	0.055	0.051	0.037	0.069	0.042
200	0.059	0.050	0.040	0.040	0.079	0.051

Table 7: Empirical size of the one-sided test with nominal size $\bar{\alpha} = 0.05$ for sample size $n = 1000$.

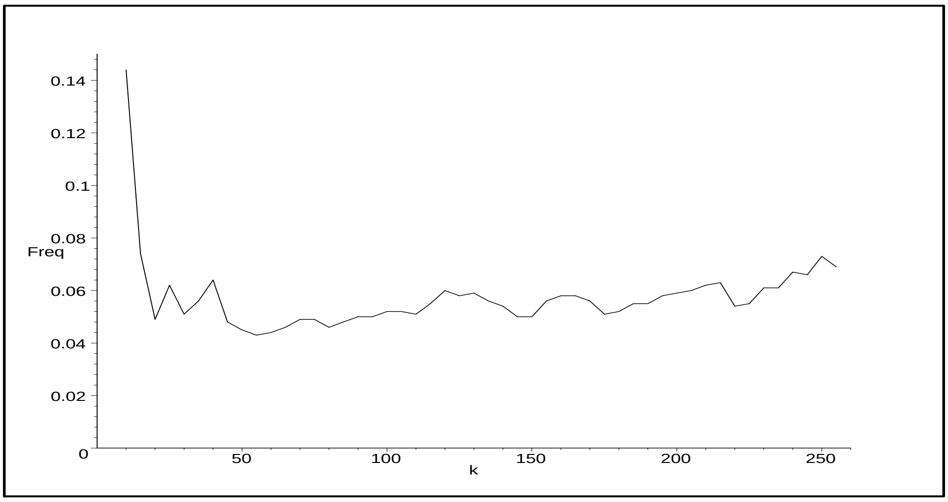

Figure 1: Empirical size of the one-sided test with nominal size $\bar{\alpha} = 0.05$ as a function of k_n for Cauchy samples of size $n = 1000$.

[11] de Haan, L. and Rootzén, H. (1993). On the estimation of high quantiles. *J. Statist. Planning Inference* **35**, 1-13.

[12] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc. Ser. A* **61**, 381-395.

[13] Resnick, S.I. (1987). *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[14] Resnick, S.I. and Stărică, C. (1997). Smoothing the Hill estimator. *Adv. Appl. Probab.* **29**, 271–293.

[15] Smith, R.L. (1987). Estimating tails of probability distributions. *Ann. Statist.* **15**, 1174-1207.

[16] Vervaat, W. (1972). Functional central limit theorems for processes with positive drift and their inverses. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **23**, 245–253.