An extreme quantile estimator for the log-generalized Weibull-tail model

Clément Albert[(1)], Anne Dutfoy[(2)], Laurent Gardes[(3)] 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*

[(3)] *Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes, 67084 Strasbourg Cedex, France*

Abstract

We propose a new estimator for extreme quantiles under the log-generalized Weibull-tail model, introduced by Cees de Valk. This model relies on a new regular variation condition which, in some situations, permits to extrapolate further into the tails than the classical assumption in extreme-value theory. The asymptotic normality of the estimator is established and its finite sample properties are illustrated both on simulated and real datasets.

Keywords: Extreme quantiles, Extreme-value theory, Extended regular variation.

AMS 2000 subject classification: 62G32, 62G20.

1 Introduction

Let X be a random variable with distribution function $F(\cdot) = \mathbb{P}(X \leq \cdot)$ and survival function $S := 1 - F$. Starting from a $n-$sample from X, our goal is to estimate extreme quantiles from S of level $1 - \beta_n$ with $\beta_n \to 0$ as $n \to \infty$. Recall that a quantile of level $1 - \beta$ is given by $Q(\beta) := \inf\{y;\ S(y) \leq \beta\}$. The rate of convergence of β_n to zero drives the difficulty of the estimation problem. Indeed, if $n\beta_n \to 0$ as $n \to \infty$, then $Q(\beta_n)$ is asymptotically almost surely larger than the sample maxima. In finance or assurance contexts, an extreme quantile is interpreted as the Value-at-Risk associated with an extreme loss, see [10, 19] for links between extreme-value theory and risk theory. In environmental applications, an extreme quantile coincides with the return level associated with an exceptional climatic event (extreme rainfalls [6], extreme wind velocities [16], extreme wave heights [18], river peak flows [17],...).

Dedicated methods have been designed to address the estimation of extreme quantiles, see [9, Chapter 6] or [15, Chapter 4], for an overview. Most of them rely on an extended regular variation assumption on the function Q. Recently, an alternative method has been initiated by Cees de Valk in a series of papers [21, 22], the goal being to estimate "more" extreme quantiles *i.e.* associated with sequences β_n tending to zero at a faster rate than in the previously mentioned

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

approaches [9, 15]. The idea is to put the extended regular variation assumption on the function $V(\cdot) := \ln Q(1/\exp \cdot)$ rather than on $Q(\cdot)$, see Paragraph 1.1 for technical details and Paragraph 1.2 for examples. Dedicated estimation methods are introduced in [23]. The goal of this work is to contribute to the popularity of this model by proposing alternative estimators, which are more efficient in some situations.

1.1 Tail model

Let X be a random variable with survival function S. For the sake of simplicity, we assume in what follows that $S(1) = 1$ *i.e.* $\mathbb{P}(X \geq 1) = 1$. The tail model considered in this work is given by

$$S(x) = \exp[-V^{\leftarrow}(\ln x)], \tag{1}$$

where $V^{\leftarrow}(\cdot) := \inf\{y;\ V(y) \geq \cdot\}$ is the generalized inverse of $V(\cdot) := \ln Q(1/\exp \cdot)$ with Q the quantile function. The function V is supposed to be of extended regular variation with index $\theta \in \mathbb{R}$. More specifically, there exists a positive function a (called the auxiliary function) such that, for all $t > 0$

$$\lim_{y\to\infty} \frac{V(ty) - V(y)}{a(y)} = \int_1^t u^{\theta-1} du =: L_\theta(t). \tag{2}$$

The class of extended regularly varying functions is denoted by $\mathcal{ERV}(\theta)$. Model (1) is referred to as the "log-generalized Weibull-tail model" [21, 22, 23]. From [15, Corollary 1.1.10], a sufficient condition for (2) is

(A1) V is differentiable with derivative V' satisfying

$$\lim_{x\to\infty} \frac{V'(tx)}{V'(x)} = t^{\theta-1}. \tag{3}$$

Such a function V' is said to be regularly varying with index $\theta - 1$ and this property is denoted by $V' \in \mathcal{RV}(\theta - 1)$. We refer to [5] for a general account on regular variation theory. Moreover, under **(A1)**, a possible choice in (2) is $a(y) = yV'(y)$.

1.2 Properties and examples

Condition **(A1)** generalizes the tail model introduced in [8, 12] where it is assumed that the function V in (2) is asymptotically proportional to L_τ for some $\tau \in [0, 1]$. One can then easily show that such a tail parameter τ coincides with the index θ of extended regular variation in the situation where $\theta \in [0, 1]$. In terms of Maximum Domain of Attraction (MDA), the following result has been established in [2, Proposition 4]:

Lemma 1 *Assume F is twice differentiable.*

(i) If **(A1)** *holds with $\theta < 1$ then $F \in$ MDA(Gumbel).*

(ii) If $F \in$ MDA(Fréchet) then **(A1)** *holds with $\theta = 1$.*

(iii) If **(A1)** *holds with $\theta > 1$ then F does not belong to any MDA.*

It thus appears that model **(A1)** with $\theta \leq 1$ is of particular interest since it is associated with most distributions in MDA(Gumbel) $\cup$ MDA(Fréchet). The situation $\theta > 1$ which does not correspond to any domain of attraction is sometimes referred to as super-heavy tails, see for instance [3]. The following examples are taken from [2, Proposition 3]:

Example 1 *Let $x^* := \sup\{x \geq 1,\, F(x) < 1\}$ be the endpoint of F. Then, under some monotonicity assumptions:*

(i) If $V^{\leftarrow}(\ln \cdot) \in \mathcal{RV}(1/\beta)$, $\beta > 0$, then **(A1)** *holds with $\theta = 0$. In this case, F is referred to as a Weibull tail-distribution, see for instance [4, 11, 14]. Such distributions encompass Gaussian, Gamma, Exponential and strict Weibull distributions.*

(ii) $V^{\leftarrow} \in \mathcal{RV}(1/\beta)$, $0 < \beta < 1$ if and only if **(A1)** *holds with $\theta = \beta > 0$. Here, F is called a log-Weibull tail-distribution, see [3, 8, 12], the most popular example being the lognormal distribution.*

(iii) $1 \leq x^* < \infty$ and $V^{\leftarrow}(\ln x^* + \ln(1 - 1/\cdot)) \in \mathcal{RV}_{-1/\beta}$, $\beta < 0$ if and only if **(A1)** *holds with $\theta = \beta < 0$. This case corresponds to distributions with a Weibull tail behavior in the neighborhood of a finite endpoint.*

We also refer to Table 1 for examples of distributions corresponding to the three above families: $\theta = 0$, $\theta > 0$ and $\theta < 0$.

1.3 Outline

The inference aspects associated with model (1) are examined in Section 2: Estimators for extreme quantiles are introduced as well as estimators for the extended regular variation index θ and the auxiliary function a. The asymptotic distributions of these estimators are established in Section 3. Their finite sample performance are investigated in Section 4 on simulated data and compared to the proposals introduced in [23]. Finally, an illustration on real data is presented in Section 5. Proofs are postponed to Section 6.

2 Inference

Let $X_1, \dots, X_n$ be n independent copies of a random variable X distributed as in (1). The associated ordered statistics are denoted by $X_{1,n} \leq \dots \leq X_{n,n}$ throughout the paper. Starting from this random sample, we focus on the estimation of extreme quantiles *i.e.* $Q(u) := S^{\leftarrow}(u) = \exp[V(\ln(1/u))]$ when $u \to 0$. Two situations for the level u are considered.

Intermediate case. If $u = \alpha_n$ where α_n is an intermediate level satisfying $\alpha_n \to 0$ and $n\alpha_n \to \infty$ as $n \to \infty$, a natural estimator is obtained by replacing Q by its empirical counterpart $\hat{Q}_n$. More precisely, $Q(\alpha_n)$ is estimated by $\hat{Q}_n(\alpha_n) = X_{n-\lfloor n\alpha_n \rfloor, n}$.

Extreme case. If $u = \beta_n$ where β_n is an extreme level such that $n\beta_n \to c \geq 0$ as $n \to \infty$, a simple order statistics cannot be used. Extrapolation beyond the sample should be performed.

Starting from an intermediate level $\alpha_n := k_n/n$ where $k_n \to \infty$ and $k_n/n \to 0$, we propose to estimate $Q(\beta_n)$ by

$$\check{Q}_n(\beta_n) := \hat{Q}_n\left(\frac{k_n}{n}\right)\exp\left[\hat{a}_n[\ln(n/k_n)]L_{\hat{\theta}_n}\left(\frac{\ln\beta_n}{\ln(k_n/n)}\right)\right], \tag{4}$$

where $\hat{\theta}_n$ and $\hat{a}_n[\ln(n/k_n)]$ are suitable estimators of θ and $a[\ln(n/k_n)]$. The rationale behind (4) is based on (2) which basically means that for α close to 0 and for all $t > 0$,

$$\ln Q(t\alpha) \approx \ln Q(\alpha) + a[\ln(1/\alpha)]L_\theta\left(1 + \frac{\ln(t)}{\ln(\alpha)}\right).$$

Estimator (4) is then obtained by taking $\alpha = k_n/n$ and $t = n\beta_n/k_n$ and by replacing the unknown quantities $Q(k_n/n)$, $a[\ln(n/k_n)]$ and θ by their corresponding estimators. Since k_n/n is an intermediate level, $Q(k_n/n)$ is estimated by $\hat{Q}_n(k_n/n) = X_{n-k_n,n}$.

Parameters estimation. Let us now propose new estimators of θ and $a[\ln(n/k_n)]$. To this end, for $j \in \{1, 2\}$, consider the statistic

$$M_n^{(j)} := \frac{1}{k_n}\sum_{i=0}^{k_n-1}\left[\ln_2(X_{n-i,n}) - \ln_2(X_{n-k_n,n})\right]^j,$$

where $\ln_2 := \ln\ln$, as well as the functions

$$\mu_b(t,\zeta) := \int_0^1\left[L_\zeta\left(1 + \frac{\ln(1/s)}{t}\right)\right]^b ds \text{ and } \Psi_t(\zeta) := \frac{\mu_1^2(t,\zeta)}{\mu_2(t,\zeta)},$$

defined for $t > 0$, $b \in \mathbb{N}\setminus\{0\}$ and $\zeta < 1$. Let us mention that $\mu_1(t, 0) = \mathrm{e}^t E_1(t)$ where $E_1(t) := \int_t^\infty u^{-1}\mathrm{e}^{-u}du$ is the exponential integral, see for instance [1, eq 5.1.1]. Furthermore, it can be shown (see Lemma 5) that Ψ_t is a decreasing function, at least for t large enough, and thus its generalized inverse $\Psi_t^{\leftarrow}$ is well defined for t large enough. The following statistics are then introduced:

$$\hat{\theta}_{n,+}^{(M)} := \frac{M_n^{(1)}}{\mu_1[\ln(n/k_n), 0]}, \tag{5}$$

$$\hat{\theta}_{n,-}^{(M)} := \Psi_{\ln(n/k_n)}^{\leftarrow}\left(\frac{[M_n^{(1)}]^2}{M_n^{(2)}}\right), \tag{6}$$

$$\hat{\theta}_n^{(M)} := \hat{\theta}_{n,+}^{(M)} + \hat{\theta}_{n,-}^{(M)}, \tag{7}$$

$$\hat{a}_n^{(M)}[\ln(n/k_n)] := \frac{\ln X_{n-k_n,n}}{\mu_1[\ln(n/k_n), \hat{\theta}_{n,-}^{(M)}]}M_n^{(1)}. \tag{8}$$

We conclude this section by giving the main ideas leading to the estimators (7) and (8) of respectively θ and $a[\ln(n/k_n)]$. The estimator (7) is similar in spirit to the moment estimator introduced in [7]. Its construction is based on the following two results. Letting $\theta_+ := \theta \vee 0$ and $\theta_- := \theta \wedge 0$, for any increasing function $V \in \mathcal{ERV}(\theta)$,

$$\lim_{x\to\infty}\frac{V(x)}{a(x)}\ln\frac{V(tx)}{V(x)} = L_{\theta_-}(t), \tag{9}$$

locally uniformly in $(0, \infty)$, see [15, Lemma B.3.16]. Moreover, one has (see for instance [15, Eq. 3.5.5]),

$$\lim_{x \to \infty} \frac{a(x)}{V(x)} = \theta_+.$$

Plugging $x := \ln(1/\alpha)$ and $t := 1 + \ln(s)/\ln(\alpha)$ in (9) yields the approximation

$$\ln_2 Q(s\alpha) - \ln_2 Q(\alpha) \approx \theta_+ L_0 \left(1 + \frac{\ln s}{\ln \alpha}\right),$$

as $\alpha \to 0$ and for all $s \in (0,1)$. Integrating with respect to s on $(0,1)$ leads to

$$\int_0^1 [\ln_2 Q(s\alpha) - \ln_2 Q(\alpha)]\, ds \Big/ \int_0^1 L_0 \left(1 + \frac{\ln s}{\ln \alpha}\right) ds \approx \theta_+.$$

Considering $\alpha = k_n/n$ where k_n is an intermediate sequence such that $k_n \to \infty$ and $k_n/n \to 0$ and replacing Q by its empirical estimator $\hat{Q}_n$ lead to the estimator (5) of θ_+. Similarly, remark that (9) leads to the approximation

$$\left\{\int_0^1 [\ln_2 Q(s\alpha) - \ln_2 Q(\alpha)]\, ds\right\}^2 \Big/ \int_0^1 [\ln_2 Q(s\alpha) - \ln_2 Q(\alpha)]^2\, ds \approx \Psi_{\ln(1/\alpha)}(\theta_-),$$

as $\alpha \to 0$. Replacing again in the previous approximation α by k_n/n and Q by its empirical counterpart suggests to estimate θ_- by (6). Finally, estimator (8) is obtained by remarking that, from (9)

$$\frac{\ln Q(\alpha)}{a[\ln(1/\alpha)]} \int_0^1 \ln \frac{\ln Q(s\alpha)}{\ln Q(\alpha)} ds \approx \mu_1[\ln(1/\alpha), \theta_-],$$

for α close to 0. Replacing α by k_n/n, Q by $\hat{Q}_n$ and θ_- by $\hat{\theta}_{n,-}^{(M)}$ gives (8).

3 Main results

3.1 Quantile estimation: Intermediate case

Let us first focus on the asymptotic behavior of the quantile estimator in the intermediate case.

Theorem 1 *Under model (1), assume that* **(A1)** *holds. For all intermediate level α_n (i.e. such that $\alpha_n \to 0$ and $n\alpha_n \to \infty$ as $n \to \infty$), one has*

$$\frac{(n\alpha_n)^{1/2} \ln(1/\alpha_n)}{a[\ln(1/\alpha_n)]} \ln\left(\frac{\hat{Q}_n(\alpha_n)}{Q(\alpha_n)}\right) \xrightarrow{d} \mathcal{N}(0,1).$$

First, remark that introducing $k_n = \lfloor n\alpha_n \rfloor$ and choosing $a(t) = tV'(t)$ (see Section 1), the above asymptotic normality result can be rewritten as

$$\frac{k_n^{1/2}}{V'[\ln(n/k_n)]} \ln\left(\frac{\hat{Q}_n(k_n/n)}{Q(k_n/n)}\right) \xrightarrow{d} \mathcal{N}(0,1).$$

If, moreover,

$$k_n^{1/2}/V'[\ln(n/k_n)] \to \infty \text{ as } n \to \infty, \tag{10}$$

then

$$\frac{k_n^{1/2}}{(n/k_n)U'(n/k_n)}\left(\hat{Q}_n(k_n/n)-Q(k_n/n)\right)\xrightarrow{d}\mathcal{N}(0,1),$$

where $U(\cdot)=Q(1/\cdot)$ is the tail quantile function. This result coincides with [15, Theorem 2.2.1] established under a von Mises' condition for the maximum domain of attraction of an extreme-value distribution. Clearly, (10) holds when $\theta<1$ since, in this case, $V'(t)\to 0$ as $t\to\infty$. Moreover, if $F\in$ MDA(Fréchet) then $\theta=1$ from Lemma 1(ii) and $U\in\mathcal{RV}(\gamma)$ for some $\gamma>0$. It thus follows that $V'(\ln t)=tU'(t)/U(t)\to\gamma$ as $t\to\infty$ and (10) is verified. The case $\theta>1$ is not relevant here, since, in this case, F does not belong to any domain of attraction, see Lemma 1(iii).
Second, under additional conditions,

$$\frac{\hat{a}_n^{(M)}[\ln(1/\alpha_n)]}{a[\ln(1/\alpha_n)]}\xrightarrow{\mathbb{P}}1,$$

see Theorem 4 below, and thus

$$\frac{(n\alpha_n)^{1/2}\ln(1/\alpha_n)}{\hat{a}_n^{(M)}[\ln(1/\alpha_n)]}\ln\left(\frac{\hat{Q}_n(\alpha_n)}{Q(\alpha_n)}\right)\xrightarrow{d}\mathcal{N}(0,1)$$

which provides a way for constructing asymptotic confidence interval for intermediate quantiles $Q(\alpha_n)$ based on $\hat{Q}_n(\alpha_n)$. Letting u_ζ the $(1+\zeta)/2$th quantile from a standard Gaussian distribution,

$$\left[\hat{Q}_n(\alpha_n)\exp\left(-\frac{\hat{a}_n^{(M)}[\ln(1/\alpha_n)]}{(n\alpha_n)^{1/2}\ln(1/\alpha_n)}u_\zeta\right);\hat{Q}_n(\alpha_n)\exp\left(\frac{\hat{a}_n^{(M)}[\ln(1/\alpha_n)]}{(n\alpha_n)^{1/2}\ln(1/\alpha_n)}u_\zeta\right)\right]$$

is a confidence interval for $Q(\alpha_n)$ of asymptotic level ζ.

3.2 Quantile estimation: Extreme case

Our next goal is to establish the asymptotic normality of $\check{Q}_n(\beta_n)$ for an extreme level β_n satisfying $n\beta_n\to c\geq 0$. A second-order condition is needed on $V\in\mathcal{ERV}(\theta)$ to control the rate of convergence in (2):

(A2) There exist a function $\tilde{A}$ with $\tilde{A}(x)\to 0$ as $x\to\infty$ and $\rho<0$ such that for all $t>0$,

$$\lim_{x\to\infty}\frac{1}{\tilde{A}(x)}\left[\frac{V(tx)-V(x)}{a(x)}-L_\theta(t)\right]=H_{\theta,\rho}(t):=\int_1^t s^{\theta-1}L_\rho(s)ds.$$

locally uniformly for $t>0$.

Note that **(A2)** also provides the rate of convergence in (9). Indeed, from [15, Lemma B.3.16], condition **(A2)** with $\theta\neq\rho$ entails that there exists a function A with $A(x)\to 0$ as $x\to\infty$ such that

$$\lim_{x\to\infty}\frac{1}{A(x)}\left[\frac{V(x)}{a(x)}\ln\frac{V(tx)}{V(x)}-L_{\theta_-}(t)\right]=H_{\theta_-,\rho'}(t). \tag{11}$$

The function $|A|$ is regularly varying with index $\rho'\leq 0$ where, according to [15, Lemma B.3.16],

$$\rho'=\begin{cases}\rho & \text{if } \theta<\rho,\\ \theta & \text{if } \rho<\theta\leq 0,\\ -\theta & \text{if } (0<\theta<-\rho \text{ and } l\neq 0),\\ \rho & \text{if } (0<\theta<-\rho \text{ and } l=0) \text{ or } (\theta\geq-\rho).\end{cases} \tag{12}$$

with, for $\theta > 0$,

$$l := \lim_{x\to\infty}\left[V(x) - \frac{a(x)}{\theta}\right].$$

Let us also introduce the positive function B defined by $B(x) := \max(|\tilde{A}(x)|, |A(x)|/x)$. It is easily checked that B is regularly varying with index $\max(\rho, \rho' - 1)$. We are now in position to establish the asymptotic distribution of $\check{Q}_n(\beta_n)$ for general estimators of θ and $a[\ln(n/k_n)]$ satisfying the condition:

(A3) There exist a sequence $\sigma_n \to 0$ and a random vector (B, Θ, Λ) such that

$$\sigma_n^{-1}\left\{\frac{\ln X_{n-k_n,n} - \ln Q(k_n/n)}{a[\ln(n/k_n)]H_{\theta,0}(d_n)}, \hat{\theta}_n - \theta, \frac{L_\theta(d_n)}{H_{\theta,0}(d_n)}\left(\frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right)\right\} \xrightarrow{d} (B, \Theta, \Lambda),$$

where $d_n := \ln(1/\beta_n)/\ln(n/k_n)$.

Theorem 2 *Under model (1), assume conditions* **(A2)**, **(A3)** *hold. Let (k_n) and (β_n) be two sequences such that $n\beta_n \to c \geq 0$, $k_n \to \infty$, $k_n/n \to 0$, $d_n \to d \in [1, \infty]$, $\sigma_n \ln(d_n) \to 0$ and $\sigma_n^{-1}\tilde{A}[\ln(n/k_n)] \to 0$ as $n \to \infty$. Then,*

$$\frac{\sigma_n^{-1}}{a[\ln(n/k_n)]H_{\theta,0}(d_n)} \ln\left(\frac{\check{Q}_n(\beta_n)}{Q(\beta_n)}\right) \xrightarrow{d} B + \Theta + \lambda.$$

Under the conditions of Theorem 2, three situations can arise for the extreme quantile level β_n. The first one is when $d_n \to 1$ which corresponds to the least extreme case. This condition is achieved for instance when $n\beta_n \to c > 0$. In this situation,

$$H_{\theta,0}(d_n) \overset{d_n\to 1}{\sim} (d_n - 1)^2/2 \to 0. \tag{13}$$

The second case corresponds to the situation where $d_n \to d \in (1, \infty)$. Here,

$$H_{\theta,0}(d_n) \overset{d_n\to d}{\longrightarrow} H_{\theta,0}(d) > 0. \tag{14}$$

Note that for these two situations, $\sigma_n \ln(d_n) \to 0$ is a consequence of the assumption $\sigma_n \to 0$ as $n \to \infty$. Finally, the most extreme case occurs when $d_n \to \infty$ leading to

$$H_{\theta,0}(d_n) \overset{d_n\to\infty}{\sim} \begin{cases} d_n^\theta \ln(d_n)/\theta & \text{if } \theta > 0, \\ \ln^2(d_n)/2 & \text{if } \theta = 0, \\ 1/\theta^2 & \text{if } \theta < 0. \end{cases} \tag{15}$$

As expected, the rate of convergence in Theorem 2 is getting worse when the quantile level β_n is getting more extreme. Let us also highlight that, when $\theta < 0$, the rates of convergence in situations $d_n \to d > 1$ and $d_n \to \infty$ are of the same order.

To conclude this section, let us give the following consistency result.

Proposition 1 *Under the conditions of Theorem 2,*

$$\frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)]} \xrightarrow{\mathbb{P}} 1 \text{ and } \frac{H_{\hat{\theta}_n,0}(d_n)}{H_{\theta,0}(d_n)} \xrightarrow{\mathbb{P}} 1. \tag{16}$$

and therefore

$$\frac{\sigma_n^{-1}}{\hat{a}_n[\ln(n/k_n)]H_{\hat{\theta}_n,0}(d_n)} \ln\left(\frac{\check{Q}_n(\beta_n)}{Q(\beta_n)}\right) \xrightarrow{d} B + \Theta + \lambda.$$

Proposition 1 can be used to construct asymptotic confidence intervals for extreme quantiles $Q(\beta_n)$ based on $\check{Q}_n(\beta_n)$, see (18).

3.3 Parameters estimation

First, the asymptotic distribution of the estimator of θ proposed in (7) is provided.

Theorem 3 *Under model (1), assume that condition* **(A2)** *holds with $\theta \neq \rho$. Let (k_n) be a sequence such that $k_n/\ln^2(n) \to \infty$, $k_n/n \to 0$ and $k_n A^2[\ln(n/k_n)]/\ln^2(n/k_n) \to 0$ as $n \to \infty$. Then,*

$$\frac{k_n^{1/2}}{\ln(n/k_n)}\left(\hat{\theta}_n^{(M)} - \theta\right) \xrightarrow{d} \mathcal{N}(0,1).$$

Note that from [13, Lemma 1], condition $k_n A^2[\ln(n/k_n)]/\ln^2(n/k_n) \to 0$ implies $\ln(k_n)/\ln(n) \to 0$ as $n \to \infty$ and thus $\ln(n/k_n) \sim \ln(n)$. Second, the asymptotic distribution of the estimator of $a[\ln(n/k_n)]$ proposed in (8) is established in the following theorem.

Theorem 4 *Under model (1), assume that condition* **(A2)** *holds with $\theta \neq \rho$. Let (k_n) be a sequence such that $k_n/\ln^2(n) \to \infty$, $k_n/n \to 0$ and $k_n B^2[\ln(n/k_n)] \to 0$ as $n \to \infty$. Then,*

$$k_n^{1/2}\left(\frac{\hat{a}_n^{(M)}[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right) \xrightarrow{d} \mathcal{N}(0,2).$$

Note that, if $\rho > -1$ and $k_n \tilde{A}^2[\ln(n)] \to 0$, then $k_n/\ln^2(n) \to 0$. Hence, Theorem 4 does not apply when $\rho \in (-1,0)$. Let us stress that this limitation also appears in [8, Theorem 1]. As a straightforward consequence of Theorems 1 – 4, the asymptotic normality of the extreme quantile estimator $\check{Q}_n^{(M)}(\beta_n)$ is obtained by considering $\hat{\theta}_n = \hat{\theta}_n^{(M)}$ and $\hat{a}_n[\ln(n/k_n)] = \hat{a}_n^{(M)}[\ln(n/k_n)]$ in (4).

Corollary 1 *Under model (1), assume that* **(A2)** *holds with $\theta \neq \rho$. Let (k_n) and (β_n) be two sequences such that $n\beta_n \to c \geq 0$, $k_n/n \to 0$, $k_n B^2[\ln(n/k_n)] \to 0$, $d_n \to d \in [1,\infty]$ and $[\ln(n)\max(1,\ln(d_n))]^2/k_n \to 0$ as $n \to \infty$. Then,*

$$\frac{k_n^{1/2}/\ln(n/k_n)}{a[\ln(n/k_n)]H_{\theta,0}(d_n)} \ln\left(\frac{\check{Q}_n^{(M)}(\beta_n)}{Q(\beta_n)}\right) \xrightarrow{d} \mathcal{N}(0,1).$$

It appears that $\ln(n/k_n)a[\ln(n/k_n)]H_{\theta,0}(d_n)/k_n^{1/2} \to 0$ is a sufficient condition to ensure that $\check{Q}_n^{(M)}(\beta_n)$ is a relatively consistent estimator of $Q(\beta_n)$, *i.e.* such that $\check{Q}_n^{(M)}(\beta_n)/Q(\beta_n) \xrightarrow{\mathbb{P}} 1$. Recalling that $\ln(n/k_n) \sim \ln n$ as $n \to \infty$, that $B \in \mathcal{RV}(\max(\rho,\rho'-1))$ and $a \in \mathcal{RV}(\theta)$, we end up with a set of three conditions on the sequences (k_n) and (β_n): $k_n B^2(\ln n) \to 0$, $[\ln(n)\max(1,\ln(d_n))]^2/k_n \to 0$ and $[\ln(n)a(\ln n)H_{\theta,0}(d_n)]^2/k_n \to 0$ as $n \to \infty$. Let us illustrate how these conditions may limit the extrapolation range β_n depending on the index θ of extended regular variation in three typical situations:

- Let $\beta_n = c/n$, $c \in (0,1)$. Here $d_n \to 1$ as $n \to \infty$, this is the least extreme case considered in Subsection 3.2, and, in view of (13), $H_{\theta,0}(d_n) \sim [\ln(k_n)/\ln(n)]^2/2$. Two constraints arise on the distribution parameters: $\rho \leq -1$ and $\theta \leq \min(1-\rho, 2-\rho')$. The first one, $\rho \leq -1$, was already imposed by Theorem 4. The second one is fulfilled as soon as $\theta \leq 2$ including MDA(Fréchet), see Lemma 1(ii), Finite endpoint, Weibull-tail, log-Weibull tail distributions (see Example 1) and some super-heavy tail distributions.

- Let $\beta_n = n^{-\tau}$, $d > 1$. Here, $d_n = \tau$, this is the second extreme case considered in Subsection 3.2, and, as a particular case of (14), $H_{\theta,0}(d_n)$ is constant. The constraints are: $\rho \leq -1$ and $\theta \leq -\max(1+\rho,\rho')$. The condition on θ is fulfilled by Finite endpoint (Example 1(iii)), Weibull-tail distributions (Example 1(i)) and some log-Weibull tail distributions (Example 1(ii)). In MDA(Fréchet), $\theta = 1$ and thus the condition on the second order parameters is strengthened: $\rho \leq -2$ and $\rho' \leq -1$.

- Let $\beta_n = \exp(-cn)$, $c > 0$. Here $d_n \to \infty$ as $n \to \infty$, this is the most extreme case considered in Subsection 3.2. In view of (15), three subcases have to be considered. If $\theta < 0$ then $H_{\theta,0}(d_n)$ is asymptotically constant and the conditions are $\rho \leq -2$ and $\rho' \leq -1$. If $\theta = 0$ then necessarily $\rho' = 0$ in view of (12), $H_{\theta,0}(d_n) \sim (\ln n)^2/2$ and it is not possible to find sequences satisfying the constraints. If $\theta > 0$ then $H_{\theta,0}(d_n) \sim (c^\theta/\theta) n^\theta (\ln n)^{1-\theta}$ and it is not possible either to find sequences satisfying the constraints.

In the second case where $\beta_n = n^{-\tau}$, $\tau > 1$, it is possible to compare the asymptotic standard deviation of $\ln \check{Q}_n^{(M)}(\beta_n)$, denoted by σ_n, to the one associated with the estimator introduced in [23], denoted by σ_n'. Our Corollary 1 and [23, Corollary 2] yield:

$$\begin{aligned} \sigma_n &\sim H_{\theta,0}(\tau) k_n^{-1/2} (\ln n) a(\ln n), \\ \sigma_n' &\sim (L_\theta^2(\tau) + H_{\theta,0}^2(\tau))^{1/2} k_n^{-1/2} (\ln n) a(\ln n). \end{aligned}$$

As a consequence, the asymptotic standard deviations are equivalent up to a multiplicative constant:

$$\frac{\sigma_n}{\sigma_n'} \to \left(1 + \frac{L_\theta^2(\tau)}{H_{\theta,0}^2(\tau)}\right)^{-1/2} =: \Lambda_\theta(\tau) \leq 1 \text{ as } n \to \infty. \tag{17}$$

The behavior of $\Lambda_\theta(\tau)$ with respect to θ and τ is illustrated on Figure 1. It appears that $\Lambda_\theta(\tau)$ is an increasing function of τ and θ. As expected $\Lambda_\theta(\tau) \leq 1$ meaning that $\check{Q}_n^{(M)}(\beta_n)$ is asymptotically more efficient than [23]'s competitor, all the more so as θ is small.
Finally, in view of Proposition 1, the unknown quantities $H_{\theta,0}(d_n)$ and $a[\ln(n/k_n)]$ can be replaced by their corresponding estimators $H_{\hat\theta_n^{(M)},0}(d_n)$ and $\hat a_n^{(M)}[\ln(n/k_n)]$ without changing the asymptotic distribution in Corollary 1. As mentioned before, the obtained result can then lead to asymptotic confidence intervals. Letting u_ζ the $(1+\zeta)/2$th quantile from a standard Gaussian distribution,

$$\check{Q}_n^{(M)}(\beta_n) \left[\exp\left(-\frac{\hat a_n^{(M)}[\ln(n/k_n)] H_{\hat\theta_n^{(M)},0}(d_n)}{k_n^{1/2}/\ln(n/k_n)} u_\zeta \right); \exp\left(\frac{\hat a_n^{(M)}[\ln(n/k_n)] H_{\hat\theta_n^{(M)},0}(d_n)}{k_n^{1/2}/\ln(n/k_n)} u_\zeta \right) \right] \tag{18}$$

is a confidence interval for $Q(\beta_n)$ of asymptotic level ζ.

4 Validation on simulations

The finite-sample behavior of the quantile estimator $\check{Q}_n^{[1]}(\beta_n) := \check{Q}_n(\beta_n)$ defined in (4) is investigated on $N = 500$ simulated random samples of size $n = 5000$, in the case where $\beta_n = n^{-2} = 4.10^{-8}$.

Estimators. Three competitors are considered:

1. The first one, $\check{Q}_n^{[2]}(\beta_n)$ is deduced from (4) by letting $\hat{\theta}_{n,-}^{(M)} := 0$ in $\hat{a}_n^{(M)}[\ln(n/k_n)]$ and $\hat{\theta}_n^{(M)}$, see (8) and (7). The resulting estimator $\check{Q}_n^{[2]}(\beta_n)$ should perform well for estimating extreme quantiles from distributions with associated $\theta \geq 0$.

2. Similarly, the second one is also obtained by letting $\hat{\theta}_n^{(M)} := 0$ in (4) and $\hat{\theta}_{n,-}^{(M)} := 0$ in (8). We thus obtain:

$$\check{Q}_n^{[3]}(\beta_n) := X_{n-k_n,n} \exp\left[\hat{a}_n[\ln(n/k_n)] \ln\left(\frac{\ln \beta_n}{\ln(k_n/n)}\right)\right],$$

   which is exactly the estimator dedicated to extreme quantiles from Weibull-tail distributions introduced in [13]. It should perform well for estimating extreme quantiles from distributions with associated $\theta = 0$.

3. Finally, the third estimator was introduced in [23]:

$$\check{Q}_n^{[4]}(\beta_n) := X_{n-\ell_n,n} \exp\left[\hat{a}_{\ell_n,n}^{[4]} L_{\hat{\theta}_{k_n,n}^{[4]}}\left(\frac{\ln(1/\beta_n)}{\nu_{\ell_n+1,n}}\right)\right],$$

   with

$$\begin{aligned} \hat{a}_{\ell_n,n}^{[4]} &:= \frac{\hat{\gamma}_{\ell_n,n}}{\frac{1}{\ell_n}\sum_{j=1}^{\ell_n} L_{\hat{\theta}_{k_n,n}^{[4]}}\left(\frac{\nu_{j,n}}{\nu_{\ell_n+1,n}}\right)}, \\ \hat{\theta}_{k_n,n}^{[4]} &:= 1 + \frac{\sum_{i=1}^{k_n-1}(\ln \hat{\gamma}_{i,n} - \ln \hat{\gamma}_{k_n,n})}{\sum_{i=1}^{k_n-1}(\ln \nu_{i+1,n} - \ln \nu_{k_n+1,n})}, \\ \hat{\gamma}_{i,n} &:= \frac{1}{i}\sum_{j=1}^{i}(\ln X_{n-j+1,n} - \ln X_{n-i,n}), \end{aligned}$$

   where $\nu_{i,n} := \sum_{j=i}^{n} j^{-1}$ and $\ell_n = k_n \nu_{k_n+1,n}^2$.

Distribution functions. The estimators are compared on the 8 distributions described in Table 1: Gamma($a = 1.5$, s), Weibull($k = 0.5$, λ_1), Gaussian(μ_1, $\sigma = 1$), Lognormal(μ_2, $\sigma = 1$), Burr(λ_2, $c = 0.5$, $k = 0.5$), Pareto-like, super heavy-tail and finite endpoint (x^*). Note that the Pareto-like distribution is taken from [23]. The position parameters μ_1, μ_2 as well as the scaling parameters s, λ_1, λ_2 and the finite endpoint x^* are chosen such that the simulated data points are all larger than 1.

Results. The log ratio errors $\check{\nu}_n^{[q]} := \ln\left(\check{Q}_n^{[q]}(\beta_n)/Q_n(\beta_n)\right)$ are computed for all 4 estimators ($q = 1, \dots, 4$), for each of the 500 datasets from the 8 distributions. The bias of each estimator is then estimated (on a logarithmic scale) by averaging the $\check{\nu}_n^{[q]}$ over the $N = 500$ replications. Similarly, the mean-squared error (MSE) is evaluated (on a logarithmic scale) by averaging the squared $\check{\nu}_n^{[q]}$ over the $N = 500$ replications.

The resulting bias and MSE associated are displayed on Figures 2–5 as functions of k_n. In terms of bias, it appears that $\check{Q}_n^{[1]}(\beta_n)$ show pretty good results with a small bias over a large range of k_n values for Gamma, Weibull, Gaussian, Lognormal, super heavy-tail and finite endpoint distributions. The bias behavior of $\check{Q}_n^{[1]}(\beta_n)$ is less satisfying on Burr and Pareto-like distributions ($\theta = 1$ in both cases) where $\check{Q}_n^{[4]}(\beta_n)$ is the best in terms of bias stability. From the MSE point of view, $\check{Q}_n^{[1]}(\beta_n)$ achieves better performances than $\check{Q}_n^{[4]}(\beta_n)$ on almost all distributions except the Pareto-like where the results are similar and the Burr distribution where $\check{Q}_n^{[4]}(\beta_n)$ is better than $\check{Q}_n^{[1]}(\beta_n)$.

Let us also note that assuming $\theta = 0$ improves the results only on the strict Weibull distribution, the results of $\check{Q}_n^{[3]}(\beta_n)$ being disappointing for other Weibull tail-distributions such as Gaussian or Gamma. Similarly, assuming $\theta > 0$ improves the results only on the Gamma distribution, the results of $\check{Q}_n^{[2]}(\beta_n)$ are not convincing on other distributions. This phenomenon indicates that $\hat{\theta}_{n,-}^{(M)}$ is useful even in case where $\theta > 0$, since it may temper the positive bias associated with $\hat{\theta}_{n,+}^{(M)}$.

5 Illustration on real data

In this section, the extreme quantile estimators $\check{Q}_n^{[1]}(\beta_n)$ and $\check{Q}_n^{[4]}(\beta_n)$ are compared on the average daily river flows (in m^3/s) of the Rhône river (France). The dataset covers the period 1915–2013, and for stationarity reasons, only the winter and spring seasons were considered (from December, 1st to May, 31st), leading to $n = 18043$ measures. We focus on the extreme quantile $Q(\beta_n)$ with $\beta_n = 5.5 \times 10^{-6}$ which is exceeded with a frequency of 10^{-3} per year. Figure 6 displays the index estimates $\hat{\theta}_n^{(M)}$ and $\hat{\theta}_{k_n,n}^{[4]}$ as well as the estimates $\check{Q}_n^{[1]}(\beta_n)$ and $\check{Q}_n^{[4]}(\beta_n)$ of the extreme quantile together with their corresponding 95% asymptotic confidence intervals.

In both cases, the index estimates seem fairly stable as a function of k_n, suggesting a positive value for $\theta \in [0.3, 0.4]$ associated with a log-Weibull tail-distribution. The behavior of extreme quantile estimates $\check{Q}_n^{[1]}(\beta_n)$ and $\check{Q}_n^{[4]}(\beta_n)$ are also similar, $\check{Q}_n^{[1]}(\beta_n)$ being more stable with respect to k_n than $\check{Q}_n^{[4]}(\beta_n)$. The first estimator $\check{Q}_n^{[1]}(\beta_n)$ points towards a constant value $Q(\beta_n) \approx 10,000 m^3/s$ while the second one $\check{Q}_n^{[4]}(\beta_n)$ exhibits a trend from 8000 to $12000 m^3/s$ as k_n vary from 100 to 2000. At the opposite, the widths of the 95% asymptotic confidence intervals associated with both estimators are significantly different. Indeed, the interval associated with $\check{Q}_n^{[1]}(\beta_n)$ is 10 times narrower than the one associated with $\check{Q}_n^{[4]}(\beta_n)$. This result is in accordance with (17) since here $\tau \simeq 1.24$ yields $\Lambda_\theta(\tau) \simeq 0.1$ for a large range of θ values, see Figure 1.

References

[1] Abramowitz, M. and Stegun, I.A. (1965). *Handbook of mathematical functions with formulas, graphs and mathematical tables*, Dover Book on Advanced Mathematics, New York.

[2] Albert, C., Dutfoy, A. and Girard, S. (2018). Asymptotic behavior of the extrapolation error associated with the estimation of extreme quantiles, `https://hal.inria.fr/hal-01692544`.

[3] Alves, I., de Haan, L. and Neves, C. (2009). A test procedure for detecting super-heavy tails, *Journal of Statistical Planning and Inference*, **139**(2), 213–227.

[4] Beirlant, J., Broniatowski, M., Teugels, J. and Vynckier, P. (1995). The mean residual life function at great age: Applications to tail estimation, *Journal of Statistical Planning and Inference*, **45**(1-2), 21–48.

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

[6] Coles, S., Pericchi, L.R., and Sisson, S. (2003). A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, **273**(1-4), 35–50.

[7] Dekkers, A., Einmhal, J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, **17**(4), 1833–1855.

[8] El Methni, J., Gardes, L., Girard, S. and Guillou, A. (2012). Estimation of extreme quantiles from heavy and light tailed distributions, *Journal of Statistical Planning and Inference*, **142**(10), 2735–2747.

[9] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*, Springer.

[10] Embrechts, P. (2000). *Extremes and integrated risk management*, Risk Books.

[11] Gardes, L. and Girard, S. (2008). Estimation of the Weibull tail-coefficient with linear combination of upper order statistics, *Journal of Statistical Planning and Inference*, **138**(5), 1416–1427.

[12] Gardes, L., Girard, S. and Guillou, A. (2011). Weibull tail-distributions revisited: a new look at some tail estimators, *Journal of Statistical Planning and Inference*, **141**(1), 429–444.

[13] Gardes, L. and Girard, S. (2006). Comparison of Weibull tail-coefficients estimators, *REVSTAT - Statistical Journal*, **4**, 163–188.

[14] Goegebeur, Y., Beirlant, J. and De Wet, T. (2010). Generalized kernel estimators for the Weibull-tail coefficient, *Communications in Statistics-Theory and Methods*, **39**(20), 3695–3716.

[15] de Haan, L., and Ferreira, A. (2006). *Extreme Value Theory: An introduction*, Springer Series in Operations Research and Financial Engineering, Springer.

[16] Jagger, T.H. and Elsner, J.B. (2006). Climatology models for extreme hurricane winds near the United States. *Journal of Climate*, **19**(13), 3220–3236.

[17] Katz, R W., Parlange, M.B. and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in water resources*, **25**(8-12), 1287–1304.

[18] Muir, L.R. and El-Shaarawi, A.H. (1986). On the calculation of extreme wave heights: a review. *Ocean Engineering*, **13**(1), 93–118.

[19] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative risk management: concepts, techniques, and tools*, Princeton university press.

[20] Smirnov, N.V. (1949). Limit distributions for the terms of a variational series, *Trudy Matematicheskogo Instituta im. V.A. Steklova*, **25**, 3–60.

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

[22] de Valk, C. (2016). Approximation and estimation of very small probabilities of multivariate extreme events, *Extremes*, **19**(4), 686–717.

[23] de Valk, C., and Cai, J.-J. (2017). A high quantile estimator based on the log-generalized Weibull tail limit, *Econometrics and Statistics*, `https://doi.org/10.1016/j.ecosta.2017.03.001`

Acknowledgements

The authors would like to thank Cees de Valk for fruitful discussions.

6 Appendix: Proofs

6.1 Preliminary lemmas

We first give a general tool for establishing the convergence in distribution of random vectors.

Lemma 2 *For $p \in \mathbb{N} \setminus \{0\}$ and $n \in \mathbb{N}$, let $W_n := (W_{n,1}, \dots, W_{n,p})^\top$ and $W := (W_1, \dots, W_p)^\top$ be two random vectors in $\mathbb{R}^p$. If there exist a sequence $\sigma_n \to 0$ and $\lambda := (\lambda_1, \dots, \lambda_p)^\top \in \mathbb{R}^p$ such that $\sigma_n^{-1}(W_n - \lambda) \xrightarrow{d} W$ then, for all $q \in \mathbb{N} \setminus \{0\}$ and all continuously differentiable functions $\varphi_1, \dots, \varphi_q$ from $\mathbb{R}^p$ to $\mathbb{R}$,*

$$\sigma_n^{-1} \left[(\varphi_1(W_n), \dots, \varphi_q(W_n))^\top - (\varphi_1(\lambda), \dots, \varphi_q(\lambda))^\top \right] \xrightarrow{d} \left(W^\top \nabla \varphi_1(\lambda), \dots, W^\top \nabla \varphi_q(\lambda) \right),$$

where, for all $i \in \{1, \dots, q\}$, $\nabla \varphi_i(\lambda)$ is the gradient of φ_i evaluated at point λ.

The following lemma is the cornerstone for establishing the asymptotic normality of the quantile estimator in the intermediate case.

Lemma 3 *Let $Z_1, \dots, Z_n$ be n independent copies of a random variable Z. Denote by S_Z the survival function of Z and by $Q_Z = S_Z^{\leftarrow}$ the associated quantile function. Assume Q_Z is differentiable and that $-Q'_Z(1/\cdot)$ is regularly varying. Then, for all sequence (α_n) such that $\alpha_n \to 0$ and $n\alpha_n \to \infty$,*

$$\frac{n^{1/2}}{\alpha_n^{1/2} Q'_Z(\alpha_n)} \left(Z_{n-\lfloor n\alpha_n \rfloor, n} - Q_Z(\alpha_n) \right) \xrightarrow{d} \mathcal{N}(0,1).$$

An elementary result on ordered statistics from standard uniform random variables is provided below.

Lemma 4 *Let $U_1, \dots, U_n$ be independent standard uniform variables. For all intermediate sequence (k_n), i.e. such that $k_n \to \infty$ and $k_n/n \to 0$, one has*

(i) $U_{k_n+1,n} \xrightarrow{\mathbb{P}} 0$.

(ii) Let $\{F_1,\dots,F_{k_n}\}$ and $\{E_1,\dots,E_n\}$ be two independent samples of independent standard exponential random variables. Then,

$$\left\{\frac{\ln(U_{i+1,n}/U_{k_n+1,n})}{\ln(U_{k_n+1,n})},\ i=0,\dots,k_n-1\right\} \stackrel{d}{=} \left\{\frac{F_{k_n-i,k_n}}{E_{n-k_n,n}},\ i=0,\dots,k_n-1\right\},$$

where $\{F_1,\dots,F_{k_n}\}$ are independent from $E_{n-k_n,n}$.

(iii)

$$\max_{i\in\{0,\dots,k_n-1\}} \frac{\ln(U_{i+1,n}/U_{k_n+1,n})}{\ln(U_{k_n+1,n})} \xrightarrow{\mathbb{P}} 0.$$

Let us introduce some additional notations. For $J\in\mathbb{N}\setminus\{0\}$, $\zeta:=(\zeta_1,\dots,\zeta_J)^\top\in(-\infty,1)^J$ and $t>0$, consider the functions

$$S_n(\zeta):=\frac{1}{k_n}\sum_{i=0}^{k_n-1}\prod_{j=1}^{J} L_{\zeta_j}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right) \text{ and } \mu(t,\zeta):=\int_0^1\prod_{j=1}^{J} L_{\zeta_j}\left(1-\frac{\ln s}{t}\right)ds$$

and remark that, for $J=1$ and $\zeta<1$, $\mu(t,\zeta)=\mu_1(t,\zeta)$ and, for $J=2$, $\mu[t,(\zeta,\zeta)]=\mu_2(t,\zeta)$. Let us also recall that, from Section 2, $\Psi_t(\zeta)=\mu_1^2(t,\zeta)/\mu_2(t,\zeta)$ for all $t>0$ and $\zeta<1$. The next result is of analytical nature. It provides first-order asymptotic expansions as $t\to\infty$ for functions $\mu(t,\zeta)$ and $\Psi_t(\zeta)$ locally uniformly on ζ.

Lemma 5 *(i) Let $J\in\mathbb{N}\setminus\{0\}$. For all hyper-rectangle $\mathcal{R}_J\subset(-\infty,1)^J$, one has*

$$\lim_{t\to\infty}\sup_{\zeta\in\mathcal{R}_J}\left|t^J\mu(t,\zeta)-J!\right|=0.$$

(ii) Denoting by Ψ_t' the first derivative of Ψ_t, one has, for all closed interval $I\subset(-\infty,1)$,

$$\lim_{t\to\infty}\sup_{\zeta\in I}\left|t\Psi_t'(\zeta)+\frac{1}{2}\right|=0.$$

As a consequence of Lemma 5(ii), the function Ψ_t is decreasing at least for t large enough. Lemma 6 below states Law of Large Numbers type results dedicated to particular triangular arrays of random variables.

Lemma 6 *Let (t_m) be a sequence such that $\log(m)/t_m\to 0$ as $m\to\infty$ and let $F_1,\dots,F_m$ be independent copies of a standard exponential random variable.*
(i) For all $\delta>0$ and $\zeta\in(-\infty,1)^J$ with $J\in\mathbb{N}\setminus\{0\}$, one has

$$\frac{t_m^{J-1}}{m}\sum_{i=1}^{m}F_i\left(1+\frac{F_i}{t_m}\right)^{\zeta_1-1}\prod_{j=2}^{J}L_{\zeta_j}\left(1+\frac{F_i}{\delta t_m}\right)\xrightarrow{\mathbb{P}} J!/\delta^{J-1}.$$

(ii) For all $\delta>0$, $(\xi_1,\dots,\xi_4)\in(-\infty,1)^4$ with $\xi_3>\xi_4$ and $J_i\in\mathbb{N}$, $i\in\{1,2,3\}$, one has for $J=J_1+J_2+2J_3$ that

$$\frac{t_m^J}{m}\sum_{i=1}^{m}L_{\xi_1}^{J_1}\left(1+\frac{F_i}{\delta t_m}\right)L_{\xi_2}^{J_2}\left(1+\frac{F_i}{\delta t_m}\right)\left[L_{\xi_3}\left(1+\frac{F_i}{\delta t_m}\right)-L_{\xi_4}\left(1+\frac{F_i}{\delta t_m}\right)\right]^{J_3}\xrightarrow{\mathbb{P}}\frac{J!}{\delta^J}\left(\frac{\xi_3-\xi_4}{2}\right)^{J_3}.$$

Finally, Lemmas 7 and 8 are the key tools for establishing the joint asymptotic normality of the random pair $(M_n^{(1)},M_n^{(2)})$.

Lemma 7 *Let* (k_n) *be an intermediate sequence such that* $k_n \to \infty$ *and* $\ln(k_n)/\ln(n) \to 0$ *as* $n \to \infty$. *For* $J_1, J_2 \in \mathbb{N} \setminus \{0\}$ *and for all* $\zeta^{(1)} \in (-\infty, 1)^{J_1}$, $\zeta^{(2)} \in (-\infty, 1)^{J_2}$, *the random vector*

$$k_n^{1/2} \left\{ \frac{S_n(\zeta^{(1)})}{\mu(\ln(n/k_n), \zeta^{(1)})} - 1, \frac{S_n(\zeta^{(2)})}{\mu(\ln(n/k_n), \zeta^{(2)})} - 1 \right\}$$

converges in distribution to a centered Gaussian random vector with covariance matrix

$$\begin{pmatrix} (2J_1)!/(J_1!)^2 - 1 & (J_1+J_2)!/(J_1!J_2!) - 1 \\ (J_1+J_2)!/(J_1!J_2!) - 1 & (2J_2)!/(J_2!)^2 - 1 \end{pmatrix}.$$

Lemma 8 *Let* (k_n) *be an intermediate sequence such that* $k_n \to \infty$ *and* $\ln(k_n)/\ln(n) \to 0$ *as* $n \to \infty$. *For all* $(\xi_1, \ldots, \xi_4) \in (-\infty, 1)^4$ *with* $\xi_3 > \xi_4$ *and* $J_i \in \mathbb{N}$, $i \in \{1, 2, 3\}$, *one has for* $J = J_1 + J_2 + 2J_3$ *that*

$$\frac{[\ln(n/k_n)]^J}{k_n} \sum_{i=0}^{k_n-1} L_{\xi_1}^{J_1}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right) L_{\xi_2}^{J_2}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right) \left[L_{\xi_3}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right) - L_{\xi_4}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right)\right]^{J_3}$$

converges in probability to $J!\left(\frac{\xi_3-\xi_4}{2}\right)^{J_3}$.

6.2 Proofs of main results

Proof of Theorem 1 – Let $\{Z_i := \ln(X_i),\ i = 1, \ldots, n\}$. These random variables are independent with common quantile function $Q_Z(u) = V[\ln(1/u)]$, $u \in (0,1)$. Under **(A1)**, Q_Z is differentiable with first derivative verifying $-Q_Z'(1/x) = xV'[\ln(x)] \in \mathcal{RV}(1)$. One can thus apply Lemma 3 to obtain

$$\frac{(n\alpha_n)^{1/2}}{V'[\ln(\alpha_n^{-1})]} \left(Z_{n-\lfloor n\alpha_n \rfloor,n} - Q_Z(\alpha_n)\right) \xrightarrow{d} \mathcal{N}(0,1).$$

Now, since $a(x) \sim xV'(x)$ as $x \to \infty$ in view of [15, Corollary 1.1.10], it follows that

$$\frac{(n\alpha_n)^{1/2}\ln(\alpha_n^{-1})}{a[\ln(\alpha_n^{-1})]} \left(Z_{n-\lfloor n\alpha_n \rfloor,n} - Q_Z(\alpha_n)\right) \xrightarrow{d} \mathcal{N}(0,1).$$

The result is then proved by remarking that $Z_{n-\lfloor n\alpha_n \rfloor,n} = \ln(X_{n-\lfloor n\alpha_n \rfloor,n})$ and $Q_Z = \ln Q$. ■

Proof of Proposition 1 – Let us first show that

$$\frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)]} \xrightarrow{\mathbb{P}} 1.$$

In view of

$$\sigma_n^{-1} \frac{L_\theta(d_n)}{H_{\theta,0}(d_n)} \left(\frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right) \xrightarrow{d} \Lambda,$$

it is sufficient to prove that $\sigma_n^{-1} L_\theta(d_n)/H_{\theta,0}(d_n) \to \infty$ as $n \to \infty$.

Let us first assume that $d_n \to 1$ as $n \to \infty$. Since $L_\theta(1+u) \sim u$ and $H_{\theta,0}(1+u) \sim u^2/2$ as $u \to 0$, $L_\theta(d_n)/H_{\theta,0}(d_n) \sim 2/(d_n - 1)$ and $\sigma_n^{-1} L_\theta(d_n)/H_{\theta,0}(d_n) \to \infty$ as $n \to \infty$. Second, if $d_n \to d \in (1, \infty)$ then $L_\theta(d_n)/H_{\theta,0}(d_n) \to L_\theta(d)/H_{\theta,0}(d) > 0$ and the result is proved. Finally, if $d_n \to \infty$, remarking that, as $t \to \infty$,

$$\frac{L_\theta(t)}{H_{\theta,0}(t)} \sim \begin{cases} 1/\ln(t) & \text{if } \theta > 0, \\ 2/\ln(t) & \text{if } \theta = 0, \\ -\theta & \text{if } \theta < 0 \end{cases} \tag{19}$$

implies $\sigma_n^{-1} L_\theta(d_n)/H_{\theta,0}(d_n) \to \infty$ by assumption.

Let us now prove the second part of Proposition 1. The following equality holds:

$$H_{\hat{\theta}_n,0}(d_n) - H_{\theta,0}(d_n) = (\hat{\theta}_n - \theta) \int_1^{d_n} s^{\theta-1} \ln^2(s) \frac{\exp[(\hat{\theta}_n - \theta)\ln(s)] - 1}{(\hat{\theta}_n - \theta)\ln(s)} ds. \tag{20}$$

Since for all $s \in (1, d_n)$, $|(\hat{\theta}_n - \theta)\ln(s)| \leq |\hat{\theta}_n - \theta| \ln(d_n) = O_{\mathbb{P}}(\sigma_n \ln d_n) = o_{\mathbb{P}}(1)$ by assumption, it is easy to check that

$$H_{\hat{\theta}_n,0}(d_n) - H_{\theta,0}(d_n) = (\hat{\theta}_n - \theta) \int_1^{d_n} s^{\theta-1} \ln^2(s) ds (1 + o_{\mathbb{P}}(1)),$$

or equivalently,

$$\frac{H_{\hat{\theta}_n,0}(d_n)}{H_{\theta,0}(d_n)} - 1 = \sigma_n^{-1}(\hat{\theta}_n - \theta) \times \frac{\sigma_n}{H_{\theta,0}(d_n)} \int_1^{d_n} s^{\theta-1} \ln^2(s) ds (1 + o_{\mathbb{P}}(1)).$$

The three situations $d_n \to 1$, $d_n \to d > 1$ and $d_n \to \infty$ are again considered separately. First, since

$$\int_1^{1+u} s^{\theta-1} \ln^2(s) ds \sim \frac{u^3}{3} \text{ and } H_{\theta,0}(u) \sim \frac{u^2}{2},$$

as $u \to 0$, one has for $d_n \to 1$ that

$$\frac{H_{\hat{\theta}_n,0}(d_n)}{H_{\theta,0}(d_n)} - 1 \sim \sigma_n^{-1}(\hat{\theta}_n - \theta) \times \frac{2}{3} \sigma_n (d_n - 1) \xrightarrow{\mathbb{P}} 0.$$

The case $d_n \to d$ is straightforward. Finally, when $d_n \to \infty$,

$$\frac{1}{H_{\theta,0}(d_n)} \int_1^{d_n} s^{\theta-1} \ln^2(s) ds \sim \begin{cases} \ln(d_n) & \text{if } \theta > 0, \\ 3\ln(d_n)/2 & \text{if } \theta = 0, \\ -2/\theta & \text{if } \theta < 0. \end{cases}$$

Collecting conditions $\sigma_n \ln(d_n) \to 0$ and $\sigma_n^{-1}(\hat{\theta}_n - \theta) \xrightarrow{d} \Theta$ concludes the proof. ■

Proof of Theorem 2 – Let us start with the expansion:

$$\frac{\sigma_n^{-1}}{a[\ln(n/k_n)] H_{\theta,0}(d_n)} \ln \frac{\check{Q}_n(\beta_n)}{Q(\beta_n)} = T_{1,n} + T_{2,n} + T_{3,n} + T_{4,n},$$

with

$$\begin{aligned}
T_{1,n} &= \frac{\sigma_n^{-1}}{a[\ln(n/k_n)] H_{\theta,0}(d_n)} \left[\ln X_{n-k_n,n} - \ln Q(k_n/n)\right], \\
T_{2,n} &= \frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)] H_{\theta,0}(d_n)} \sigma_n^{-1} \left[L_{\hat{\theta}_n}(d_n) - L_\theta(d_n)\right], \\
T_{3,n} &= \frac{L_\theta(d_n)}{H_{\theta,0}(d_n)} \sigma_n^{-1} \left[\frac{\hat{a}_n[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right], \\
T_{4,n} &= \frac{\sigma_n^{-1}}{H_{\theta,0}(d_n)} \left[\frac{\ln Q(k_n/n) - \ln Q(\beta_n)}{a[\ln(n/k_n)]} + L_\theta(d_n)\right].
\end{aligned}$$

Clearly, under **(A3)**, $T_{1,n} \xrightarrow{d} B$ and $T_{3,n} \xrightarrow{d} \Lambda$. Next, remark that Proposition 1 entails that the asymptotic distribution of $T_{2,n}$ is the same as the one of

$$\frac{\sigma_n^{-1}}{H_{\theta,0}(d_n)} \left[L_{\hat{\theta}_n}(d_n) - L_\theta(d_n) \right].$$

Furthermore, similarly to (20) in the proof of Proposition 1, one can show that $L_{\hat{\theta}_n}(d_n) - L_\theta(d_n) = (\hat{\theta}_n - \theta) H_{\theta,0}(d_n)(1 + o_{\mathbb{P}}(1))$. As a consequence, $T_{2,n} \xrightarrow{d} \Theta$. Finally, since $\ln Q(\alpha) = V[\ln(1/\alpha)]$, it follows that

$$H_{\theta,0}(d_n)\sigma_n T_{4,n} = \frac{V[\ln(n/k_n)] - V[\ln(\beta_n)]}{a[\ln(n/k_n)]} + L_\theta(d_n).$$

Let us consider separately the three cases $d_n \to 1$, $d_n \to d > 1$ and $d_n \to \infty$.
First, if $d_n \to 1$, the second order condition **(A2)** entails $H_{\theta,0}(d_n)\sigma_n T_{4,n} \sim H_{\theta,\rho}(d_n)\tilde{A}[\ln(n/k_n)]$. Since for all $\rho \leq 0$, $H_{\theta,\rho}(1+u) \sim H_{\theta,0}(1+u) \sim u^2/2$ as $u \to 0$, it follows that $T_{4,n} \sim \sigma_n^{-1}\tilde{A}[\ln(n/k_n)] = o(1)$ by assumption.
Next, if $d_n \to d > 1$, conditions **(A2)** and $\sigma_n^{-1}\tilde{A}[\ln(n/k_n)] \to 0$ imply that $T_{4,n} \to 0$ as $n \to \infty$.
Finally, when $d_n \to \infty$, [15, Lemma 4.3.5] entails that

$$T_{4,n} = \mathcal{O}\left\{ \frac{L_\theta(d_n)}{H_{\theta,0}(d_n)} \sigma_n^{-1} \tilde{A}[\ln(n/k_n)] \right\} = o(1),$$

using (19). To conclude, if $d_n \to d \in [1, \infty]$, $T_{4,n} \to 0$ as $n \to \infty$ and the result is proved. ■

Proof of Theorem 3 – For $i = 1, \ldots, n$, let $U_i := S(X_i)$ so that $\{U_1, \ldots, U_n\}$ is distributed a set of independent standard uniform random variables. Let $\delta > 0$ and $r(\cdot) := a(\cdot)/V(\cdot)$. For $s \in (\alpha^\delta, 1)$, let us plug $x := \ln(1/\alpha)$ and $t := 1 + \ln s / \ln \alpha$ in (11). Consequently, as $\alpha \to 0$,

$$r^{-1}[\ln(1/\alpha)] \ln \frac{\ln Q(s\alpha)}{\ln Q(\alpha)} = L_{\theta_-}\left(1 + \frac{\ln s}{\ln \alpha}\right) + A[\ln(1/\alpha)] H_{\theta_-,\rho'}\left(1 + \frac{\ln s}{\ln \alpha}\right)(1 + o(1)), \tag{21}$$

uniformly in $s \in (\alpha^\delta, 1)$. As a consequence of Lemma 4(i, ii), one may apply (21) with α replaced by $U_{k_n+1,n}$ and s replaced by $U_{i+1,n}/U_{k_n+1,n}$ to get, for n large enough,

$$\begin{aligned} r^{-1}(\ln U_{k_n+1,n}^{-1}) \frac{M_n^{(1)}}{\mu_1[\ln(n/k_n), \theta_-]} - 1 &= \frac{S_n(\theta_-)}{\mu_1[\ln(n/k_n), \theta_-]} - 1 \\ &+ \frac{A(\ln U_{k_n+1,n}^{-1})}{k_n \mu_1[\ln(n/k_n), \theta_-]} \sum_{i=0}^{k_n-1} H_{\theta_-,\rho'}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right)(1 + o_{\mathbb{P}}(1)). \end{aligned}$$

Similarly, we have

$$\begin{aligned} r^{-2}(\ln U_{k_n+1,n}^{-1}) \frac{M_n^{(2)}}{\mu_2[\ln(n/k_n), \theta_-]} - 1 &= \frac{S_n[(\theta_-, \theta_-)]}{\mu_2[\ln(n/k_n), \theta_-]} - 1 \\ + \frac{2A(\ln U_{k_n+1,n}^{-1})}{k_n \mu_2[\ln(n/k_n), \theta_-]} &\sum_{i=0}^{k_n-1} L_{\theta_-}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right) H_{\theta_-,\rho'}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right)(1 + o_{\mathbb{P}}(1)) \\ + \frac{A^2(\ln U_{k_n+1,n}^{-1})}{k_n \mu_2[\ln(n/k_n), \theta_-]} &\sum_{i=0}^{k_n-1} H^2_{\theta_-,\rho'}\left(\frac{\ln U_{i+1,n}}{\ln U_{k_n+1,n}}\right)(1 + o_{\mathbb{P}}(1)). \end{aligned}$$

From Rényi's representation, $\ln(1/U_{k_n+1,n})/\ln(n/k_n) \xrightarrow{\mathbb{P}} 1$ and since $|A|$ is regularly varying, it follows that

$$\left|\frac{A[\ln(1/U_{k_n+1,n})]}{A[\ln(n/k_n)]}\right| \xrightarrow{\mathbb{P}} 1 \tag{22}$$

as $n \to \infty$. Now, since for all $t > 0$,

$$H_{\theta_-,\rho'}(t) = \begin{cases} 1/\rho'[L_{\theta_-+\rho'}(t) - L_{\theta_-}(t)] & \text{if } \rho' \neq 0, \\ L_{\theta_-}(t)L_0(t) - \theta_-^{-1}[L_{\theta_-}(t) - L_0(t)] & \text{if } \rho' = 0 \text{ and } \theta_- \neq 0, \\ L_0^2(t)/2 & \text{if } \rho' = \theta_- = 0, \end{cases}$$

Lemma 5(i), Lemma 7, Lemma 8 and condition $k_n A^2[\ln(n/k_n)]/\ln^2(n/k_n) \to 0$ yield

$$r^{-1}(\ln U_{k_n+1,n}^{-1})\frac{M_n^{(1)}}{\mu_1(\ln(n/k_n),\theta_-)} - 1 = \frac{S_n(\theta_-)}{\mu_1(\ln(n/k_n),\theta_-)} - 1 + o_{\mathbb{P}}\left(k_n^{-1/2}\right),$$

and

$$r^{-2}(\ln U_{k_n+1,n}^{-1})\frac{M_n^{(2)}}{\mu_2[\ln(n/k_n),\theta_-]} - 1 = \frac{S_n[(\theta_-,\theta_-)]}{\mu_2[\ln(n/k_n),\theta_-]} - 1 + o_{\mathbb{P}}\left(k_n^{-1/2}\right).$$

Note that Lemma 7 can be applied since $k_n A^2[\ln(n/k_n)]/\ln^2(n/k_n) \to 0$ implies $\ln(k_n)/\ln(n) \to 0$, see [13, Lemma 1]. As a consequence, the random vector

$$k_n^{1/2}\left(r^{-1}(\ln U_{k_n+1,n}^{-1})\frac{M_n^{(1)}}{\mu(\ln(n/k_n),\theta_-)} - 1, r^{-2}(\ln U_{k_n+1,n}^{-1})\frac{M_n^{(2)}}{\mu(\ln(n/k_n),(\theta_-,\theta_-))} - 1\right) \tag{23}$$

converges in distribution to a centered Gaussian random vector (P_1, P_2) with covariance matrix

$$\begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix}.$$

Let us investigate the asymptotic distribution of $k_n^{1/2}(\hat{\theta}_{n,+}^{(M)} - \theta_+)$ where we recall that

$$\hat{\theta}_{n,+}^{(M)} = \frac{M_n^{(1)}}{\mu_1(\ln(n/k_n),0)},$$

see (5). From [15, Eq. 3.5.13], $r(\cdot)$ is regularly varying and thus $\ln(1/U_{k_n+1,n})/\ln(n/k_n) \xrightarrow{\mathbb{P}} 1$ implies $r[\ln(1/U_{k_n+1,n})]/r[\ln(n/k_n)] \xrightarrow{\mathbb{P}} 1$. Since $\mu_1(\ln(n/k_n),\theta_-) = \mu_1(\ln(n/k_n),0)$ for $\theta > 0$ and $\mu_1(\ln(n/k_n),\theta_-) \sim \mu_1(\ln(n/k_n),0)$ for $\theta \leq 0$ from Lemma 5(i), convergence in distribution (23) yields in both cases

$$k_n^{1/2}\left(\hat{\theta}_{n,+}^{(M)} - \theta_+\right) = r[\ln(n/k_n)]P_{1,n} + k_n^{1/2}\left\{r[\ln(1/U_{k_n+1,n})] - \theta_+\right\}(1 + o(1)),$$

where $P_{1,n} \xrightarrow{d} P_1$. Now, [15, Eq. B.3.46] ensures that $[r(x) - \theta_+]/A(x) \to \lambda \in \mathbb{R}$ as $x \to \infty$ and thus, taking into account of (22),

$$k_n^{1/2}\left(\hat{\theta}_{n,+}^{(M)} - \theta_+\right) = r[\ln(n/k_n)]P_{1,n} + \mathcal{O}_{\mathbb{P}}(k_n^{1/2}A[\ln(n/k_n)]) = \theta_+ P_{1,n} + \mathcal{O}_{\mathbb{P}}(k_n^{1/2}A[\ln(n/k_n)]), \tag{24}$$

since $r[\ln(n/k_n)] \to \theta_+$ as $n \to \infty$. Now, using convergence in distribution (23) and a Taylor expansion yield

$$\frac{1}{\Psi_{\ln(n/k_n)}(\theta_-)}\frac{[M_n^{(1)}]^2}{M_n^{(2)}} = 1 + k_n^{-1/2}(2P_{1,n} - P_{2,n}) + o_{\mathbb{P}}(k_n^{-1/2}),$$

where $(P_{1,n}, P_{2,n}) \xrightarrow{d} (P_1, P_2)$. From Lemma 5(i), $\Psi_{\ln(n/k_n)}(\theta_-) \to 1/2$ as $n \to \infty$, and thus

$$2k_n^{1/2} \left(\frac{[M_n^{(1)}]^2}{M_n^{(2)}} - \Psi_{\ln(n/k_n)}(\theta_-) \right) \xrightarrow{d} 2P_1 - P_2, \tag{25}$$

where it is easily seen that $2P_1 - P_2 \sim \mathcal{N}(0,1)$. Now let $\sigma_n := k_n^{-1/2} \ln(n/k_n) \to 0$. For all $z \in \mathbb{R}$ and n large enough,

$$\mathbb{P}\left[\sigma_n^{-1} \left(\hat{\theta}_{n,-}^{(M)} - \theta_-\right) \le z\right] = \mathbb{P}\left[\frac{[M_n^{(1)}]^2}{M_n^{(2)}} \ge \Psi_{\ln(n/k_n)}\left(\theta_- + \sigma_n z\right)\right],$$

since for n large enough, $\Psi_{\ln(n/k_n)}$ is decreasing. Hence,

$$\mathbb{P}\left[\sigma_n^{-1} \left(\hat{\theta}_{n,-}^{(M)} - \theta_-\right) \le z\right] = \mathbb{P}\left[2k_n^{1/2} \left(\frac{[M_n^{(1)}]^2}{M_n^{(2)}} - \Psi_{\ln(n/k_n)}(\theta_-)\right) \ge z_{n,k_n}\right],$$

with $z_{n,k_n} := 2k_n^{1/2}[\Psi_{\ln(n/k_n)}\left(\theta_- + \sigma_n z\right) - \Psi_{\ln(n/k_n)}\left(\theta_-\right)]$. The mean-value theorem entails that

$$z_{n,k_n} = 2k_n^{1/2} \sigma_n z \Psi'_{\ln(n/k_n)}\left(\theta_- + \tau_n \sigma_n z\right),$$

where $\tau_n \in (0,1)$. We thus have that $z_{n,k_n} \to -z$ as $n \to \infty$ from Lemma 5(ii) and replacing σ_n by its expression. Taking into account of convergence (25) leads to

$$\frac{k_n^{1/2}}{\ln(n/k_n)} \left(\hat{\theta}_{n,-}^{(M)} - \theta_-\right) \xrightarrow{d} 2P_1 - P_2 \sim \mathcal{N}(0,1). \tag{26}$$

Collecting (24) and (26) concludes the proof. ■

Proof of Theorem 4 – Keeping in mind the notations introduced in the proof of Theorem 3, the following expansion holds

$$\frac{\hat{a}_n^{(M)}[\ln(n/k_n)]}{a[\ln(n/k_n)]} = \mathcal{F}_{1,n} \times \mathcal{F}_{2,n} \times \mathcal{F}_{3,n},$$

with

$$\mathcal{F}_{1,n} := \frac{r^{-1}[\ln(1/U_{k_n+1,n})] M_n^{(1)}}{\mu_1[\ln(n/k_n), \theta_-]}, \ \mathcal{F}_{2,n} := \frac{a[\ln(1/U_{k_n+1,n})]}{a[\ln(n/k_n)]} \text{ and } \mathcal{F}_{3,n} := \frac{\mu_1[\ln(n/k_n), \theta_-]}{\mu_1[\ln(n/k_n), \hat{\theta}_{n,-}^{(M)}]}.$$

First, (23) entails that

$$k_n^{1/2}(\mathcal{F}_{1,n} - 1) \xrightarrow{d} P_1. \tag{27}$$

Let us now consider $\mathcal{F}_{2,n}$. From [15, Theorem 2.3.6 and Corollary 2.3.5], there exist a function a_0 with, as $t \to \infty$

$$\frac{a_0(t)}{a(t)} = 1 + \mathcal{O}[\tilde{A}(t)]$$

and a function A_0 with $A_0(t) = \mathcal{O}[\tilde{A}(t)]$ as $t \to \infty$ such that, for all $\varepsilon > 0$, $\delta > 0$ and n large enough,

$$A_0^{-1}[\ln(n/k_n)] \left[\frac{a_0[\ln(1/U_{k_n+1,n})]}{a_0[\ln(n/k_n)]} - \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^\theta\right] = \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^\theta L_\rho \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right) + R_n,$$

where

$$|R_n| \le \varepsilon \max\left\{ \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta+\rho+\delta}, \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta+\rho-\delta} \right\}.$$

Hence, since $|\tilde{A}|$ is a regularly varying function and $nU_{k_n+1,n}/k_n \xrightarrow{\mathbb{P}} 1$,

$$\begin{aligned} \mathcal{F}_{2,n} &= \left\{ \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta} + \mathcal{O}\{\tilde{A}[\ln(n/k_n)]\} \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta} L_\rho \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right) \right. \\ &+ \left. \mathcal{O}\{\tilde{A}[\ln(n/k_n)]R_n\} \right\} \{1 + \mathcal{O}\{\tilde{A}[\ln(n/k_n)]\}\}. \end{aligned} \tag{28}$$

Let us now consider the expansion

$$k_n^{1/2}(\mathcal{F}_{2,n} - 1) = k_n^{1/2}\left[\mathcal{F}_{2,n} - \left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta}\right] + k_n^{1/2}\left[\left(\frac{\ln(1/U_{k_n+1,n})}{\ln(n/k_n)}\right)^{\theta} - 1\right] =: T_{1,n} + T_{2,n}.$$

Since $\ln(U_{k_n+1,n})/\ln(k_n/n) \xrightarrow{\mathbb{P}} 1$, (28) entails that

$$T_{1,n} = \mathcal{O}\{k_n^{1/2}\tilde{A}[\ln(n/k_n)]\} = o_{\mathbb{P}}(1),$$

by assumption. Next, Lemma 3 yields $\xi_n := k_n^{1/2}[\ln(1/U_{k_n+1,n}) - \ln(n/k_n)] \xrightarrow{d} \mathcal{N}(0,1)$ and thus

$$T_{2,n} = k_n^{1/2}\left[\left(1 + \frac{\xi_n}{k_n^{1/2}\ln(n/k_n)}\right)^{\theta} - 1\right] = o_{\mathbb{P}}(1).$$

To sum up, we have shown that

$$k_n^{1/2}(\mathcal{F}_{2,n} - 1) \xrightarrow{\mathbb{P}} 0. \tag{29}$$

Let us finally focus on $\mathcal{F}_{3,n}$. The mean-value theorem entails that

$$\mu_1\left[\ln(n/k_n), \hat{\theta}_{n,-}^{(M)}\right] - \mu_1\left[\ln(n/k_n), \theta_-\right] = (\hat{\theta}_{n,-}^{(M)} - \theta_-)\dot{\mu}\left[\ln(n/k_n), \theta_{n,-}^*\right],$$

where $\theta_{n,-}^* = \theta_- + \tau(\hat{\theta}_{n,-}^{(M)} - \theta_-)$ for some random value $\tau \in (0,1)$ and

$$\dot{\mu}(t,x) = \frac{\partial}{\partial x}\mu(t,x).$$

It has been shown in the proof of Lemma 5(ii) that $\mathcal{I}_1(t,x) = t^2\dot{\mu}(t,x) \to 1$ as $t \to \infty$, uniformly on all closed interval included in $(-\infty, 1)$. Hence, under the assumptions of Theorem 4, $\theta_{n,-}^* \xrightarrow{\mathbb{P}} \theta_-$ from Theorem 3 and

$$\begin{aligned} & k_n^{1/2}\ln(n/k_n)\left\{\mu_1\left[\ln(n/k_n), \hat{\theta}_{n,-}^{(M)}\right] - \mu_1\left[\ln(n/k_n), \theta_-\right]\right\} \\ =\ & [\ln(n/k_n)]^2\dot{\mu}\left[\ln(n/k_n), \theta_{n,-}^*\right]\frac{k_n^{1/2}}{\ln(n/k_n)}\left(\hat{\theta}_{n,-}^{(M)} - \theta_-\right) \xrightarrow{d} 2P_1 - P2 \end{aligned}$$

from (26). Lemma 5(i) yields

$$k_n^{1/2}(\mathcal{F}_{3,n} - 1) \xrightarrow{d} P_2 - 2P_1. \tag{30}$$

Collecting (27), (29) and (30), Lemma 2 leads to

$$k_n^{1/2}\left[\frac{\hat{a}_n^{(M)}[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right] \xrightarrow{d} P_2 - P_1 \sim \mathcal{N}(0,2),$$

which is the desired result. ■

Proof of Corollary 1 – It is sufficient to show that condition **(A3)** is satisfied by the estimators $\hat{\theta}_n^{(M)}$ and $\hat{a}_n^{(M)}[\ln(n/k_n)]$ with $\sigma_n := k_n^{-1/2}\ln(n/k_n)$ and $(B, \Theta, \Lambda) = (0, \Theta, 0)$ where Θ follows a standard Gaussian distribution. First, from Theorem 1,

$$\frac{k_n^{1/2}}{\ln(n/k_n)} \frac{\ln(X_{n-k_n,n}) - \ln Q(k_n/n)}{a[\ln(n/k_n)]H_{\theta,0}(d_n)} = \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\ln^2(n/k_n)H_{\theta,0}(d_n)}\right).$$

Clearly, when $d_n \to d \in (1, \infty]$, $\ln^2(n/k_n)H_{\theta,0}(d_n) \to \infty$. When $d_n \to 1$,

$$\ln^2\left(\frac{n}{k_n}\right) H_{\theta,0}(d_n) \sim \frac{1}{2}\ln^2\left(\frac{n}{k_n}\right)(d_n - 1)^2 = \ln^2\left(\frac{k_n}{n\beta_n}\right) \to \infty, \tag{31}$$

since $n\beta_n \to c \geq 0$. As a consequence, if $d_n \to d \in [1, \infty]$,

$$\sigma_n^{-1}\frac{\ln(X_{n-k_n,n}) - \ln Q(k_n/n)}{a[\ln(n/k_n)]H_{\theta,0}(d_n)} \xrightarrow{\mathbb{P}} 0 \tag{32}$$

Next, Theorem 3 entails that

$$\sigma_n^{-1}\left(\hat{\theta}_n^{(M)} - \theta\right) \xrightarrow{d} \mathcal{N}(0, 1). \tag{33}$$

Finally, from Theorem 4,

$$\frac{k_n^{1/2}}{\ln(n/k_n)} \frac{L_\theta(d_n)}{H_{\theta,0}(d_n)}\left(\frac{\hat{a}_n^{(M)}[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right) = \mathcal{O}_{\mathbb{P}}\left(\frac{L_\theta(d_n)}{\ln(n/k_n)H_{\theta,0}(d_n)}\right).$$

If $d_n \to 1$, since $H_{\theta,0}(d_n)/L_\theta(d_n) \sim (d_n - 1)/2$,

$$\frac{L_\theta(d_n)}{\ln(n/k_n)H_{\theta,0}(d_n)} \to 0, \tag{34}$$

as shown in (31). When $d_n \to d > 1$, it is clear that (34) holds. Finally, when $d_n \to \infty$, (19) entails that $L_\theta(d_n)/H_{\theta,0}(d_n) \to -\theta_-$ and thus (34) also holds. To sum up, when $d_n \to d \in [1, \infty]$,

$$\sigma_n^{-1}\frac{L_\theta(d_n)}{H_{\theta,0}(d_n)}\left(\frac{\hat{a}_n^{(M)}[\ln(n/k_n)]}{a[\ln(n/k_n)]} - 1\right) \xrightarrow{\mathbb{P}} 0, \tag{35}$$

and the conclusion follows from (32), (33) and (35). ∎

6.3 Proofs of auxiliary results

Proof of Lemma 2 – Let $\varphi : \mathbb{R}^p \mapsto \mathbb{R}$ be a function of class $\mathcal{C}^1$. It suffices to show that

$$\sigma_n^{-1}(\varphi(W_n) - \varphi(\lambda)) \xrightarrow{d} W^\top \nabla\varphi(\lambda).$$

Conclusion of the proof will be then straightforward by applying the Cramér-Wold device. The multivariate version of the mean-value theorem leads to

$$\sigma_n^{-1}(\varphi(W_n) - \varphi(\lambda)) = \sigma_n^{-1}(W_n - \lambda)^\top \nabla\varphi(\lambda_n^*),$$

where $\lambda_n^* := (\lambda_{n,1}^*, \ldots, \lambda_{n,p}^*)^\top$ with for all $i \in \{1, \ldots, p\}$, $\lambda_{n,i}^* = \lambda_i + \tau_i(W_{n,i} - \lambda_i)$ where $\tau_i \in (0, 1)$. By assumption, $\lambda_{n,i}^* \xrightarrow{\mathbb{P}} \lambda_i$ and the continuous mapping theorem entails that $\nabla\varphi(\lambda_n^*) \xrightarrow{\mathbb{P}} \nabla\varphi(\lambda)$ and the proof is completed. ∎

Proof of Lemma 3 – We start with a result due to Smirnov [20] and that can be found for instance in [15, Lemma 2.2.3]. Let α_n be a sequence such that $\alpha_n \to 0$ and $n\alpha_n \to \infty$. If $U_1, \dots, U_n$ are independent random variables from a standard uniform distribution,

$$\frac{n^{1/2}}{\alpha_n^{1/2}} \left(U_{\lfloor n\alpha_n \rfloor+1,n} - \alpha_n \right) \xrightarrow{d} \mathcal{N}(0,1). \tag{36}$$

Since $Z_{n-\lfloor n\alpha_n \rfloor,n} \stackrel{d}{=} Q_Z(U_{\lfloor n\alpha_n \rfloor+1,n})$, we want to show that

$$\frac{n^{1/2}}{\alpha_n^{1/2} Q_Z'(\alpha_n)} \left(Q_Z(U_{\lfloor n\alpha_n \rfloor+1,n}) - Q_Z(\alpha_n) \right) \xrightarrow{d} \mathcal{N}(0,1).$$

Since Q_Z is a differentiable function, the mean-value theorem leads to

$$Q_Z(U_{\lfloor n\alpha_n \rfloor+1,n}) - Q_Z(\alpha_n) = \left(U_{\lfloor n\alpha_n \rfloor+1,n} - \alpha_n \right) Q_Z'(\alpha_n^*),$$

where $\alpha_n^* := \alpha_n + \tau(U_{\lfloor n\alpha_n \rfloor+1,n} - \alpha_n)$ with $\tau \in (0,1)$. From (36) and since $n\alpha_n \to \infty$,

$$\frac{U_{\lfloor n\alpha_n \rfloor+1,n} - \alpha_n}{\alpha_n} \xrightarrow{\mathbb{P}} 0,$$

and thus $\alpha_n^* = \alpha_n(1 + o_{\mathbb{P}}(1))$. Since $-Q_Z'(1/\cdot)$ is regularly varying,

$$\frac{n^{1/2}}{\alpha_n^{1/2} Q_Z'(\alpha_n)} \left(Q_Z(U_{\lfloor n\alpha_n \rfloor+1,n}) - Q_Z(\alpha_n) \right) = \frac{n^{1/2}}{\alpha_n^{1/2}} \left(U_{\lfloor n\alpha_n \rfloor+1,n} - \alpha_n \right) (1 + o_{\mathbb{P}}(1)) \xrightarrow{d} \mathcal{N}(0,1),$$

and the proof is completed. ■

Proof of Lemma 4 – (i) The proof is based on Rényi's representation of standard uniform ordered statistics:

$$U_{k_n+1,n} \stackrel{d}{=} \frac{T_{k_n+1}}{T_{n+1}},$$

where for $j \in \mathbb{N} \setminus \{0\}$, T_j is the sum of j independent standard exponential random variables. The law of large numbers shows that $U_{k_n+1,n} \stackrel{\mathbb{P}}{\sim} k_n/n$ and the conclusion follows.
(ii) Remarking that

$$\left\{ \frac{\ln(U_{i+1,n}/U_{k_n+1,n})}{\ln(U_{k_n+1,n})},\ i = 0, \dots, k_n - 1 \right\} \stackrel{d}{=} \left\{ \frac{E_{n-i,n} - E_{n-k_n,n}}{E_{n-k_n,n}},\ i = 0, \dots, k_n - 1 \right\},$$

the result is then a consequence of the following Rényi's representation:

$$\{E_{j,n},\ j = 1, \dots, n\} \stackrel{d}{=} \left\{ \sum_{r=1}^{j} \frac{F_r}{n - r + 1},\ j = 1, \dots, n \right\}.$$

(iii) It is clear that

$$0 \leq \max_{i \in \{0,\dots,k_n-1\}} \frac{F_{k_n-i,k_n}}{E_{n-k_n,n}} \leq \frac{F_{k_n,k_n}}{E_{n-k_n,n}}.$$

Using the facts that $k_n^{1/2}(E_{n-k_n,n} - \ln(n/k_n)) \xrightarrow{d} \mathcal{N}(0,1)$ and that $F_{k_n,k_n} - \ln k_n$ converges in distribution to a Gumbel random variable entails

$$\frac{F_{k_n,k_n}}{E_{n-k_n,n}} \stackrel{\mathbb{P}}{\sim} \frac{\ln k_n}{\ln(n/k_n)} \to 0,$$

and the conclusion follows. ■

Proof of Lemma 5 – (i) For $j = 1, \ldots, J$, let

$$R_j(t,s) := tL_{\zeta_j}\left(1 + \frac{\ln(1/s)}{t}\right) - \ln(1/s).$$

Since $2|R_j(t,s)| \leq (1-\zeta_j)\ln^2(1/s)/t$, denoting by $\underline{\zeta} := \min\{\zeta_1, \ldots, \zeta_J\}$, one has

$$-\frac{1-\underline{\zeta}}{2}\frac{\ln^2(1/s)}{t} \leq R_j(t,s) \leq \frac{1-\underline{\zeta}}{2}\frac{\ln^2(1/s)}{t}.$$

Hence

$$\int_0^1 \left[\ln(1/s) - \frac{1-\underline{\zeta}}{2}\frac{\ln^2(1/s)}{t}\right]^J ds \leq t^J \mu(t,\zeta) \leq \int_0^1 \left[\ln(1/s) + \frac{1-\underline{\zeta}}{2}\frac{\ln^2(1/s)}{t}\right]^J ds.$$

As a consequence,

$$t^J\mu(t,\zeta) - J! \geq \sum_{j=0}^{J-1}(-1)^{J-j}(2J-j)!C_J^j\left(\frac{1-\underline{\zeta}}{2}\right)^{J-j}\frac{1}{t^{J-j}} \to 0,$$

uniformly for any hyper-rectangle included in $(-\infty, 1)^J$. Similarly,

$$t^J\mu(t,\zeta) - J! \leq \sum_{j=0}^{J-1}(2J-j)!C_J^j\left(\frac{1-\underline{\zeta}}{2}\right)^{J-j}\frac{1}{t^{J-j}} \to 0,$$

uniformly locally and the proof is completed.

(ii) It is easily seen that

$$t[t^2\mu_2(t,x)]^2\Psi_t'(x) = 2[t\mu_1(t,x)][t^2\mu_2(t,x)]\mathcal{I}_1(t,x) - [t\mu_1(t,x)]^2\mathcal{I}_2(t,x),$$

with

$$\begin{aligned}
\mathcal{I}_1(t,x) &= t^2\dot{\mu}_1(t,x) = \int_0^1 t^2\dot{L}_x\left(1+\frac{\ln(1/s)}{t}\right)ds,\\
\mathcal{I}_2(t,x) &= t^2\dot{\mu}_2(t,x) = 2\int_0^1 t^2\dot{L}_x\left(1+\frac{\ln(1/s)}{t}\right)tL_x\left(1+\frac{\ln(1/s)}{t}\right)ds,
\end{aligned}$$

and where the following notations have been introduced

$$\dot{L}_x(u) := \frac{\partial}{\partial x}L_x(u) \text{ and } \dot{\mu}_b(t,x) := \frac{\partial}{\partial x}\mu_b(t,x),\ b \in \{1,2\}.$$

The first step consists in studying the quantities $L_x(1+u)$ and $\dot{L}_x(1+u)$ for $u \geq 0$ and $x < 1$. A Taylor expansion leads to

$$L_x(1+u) = u + \frac{x-1}{2}u^2 + R_x(u), \tag{37}$$

where

$$0 \leq R_x(u) \leq \frac{(x-1)(x-2)}{6}u^3. \tag{38}$$

Next, an integration by part entails

$$\dot{L}_x\,(1+u) = \frac{1}{x}\left[\ln(1+u)(1+u)^x - L_x(1+u)\right]. \tag{39}$$

Let us note that when $x = 0$,

$$\dot{L}_0\left(1+u\right) = \lim_{x\to 0}\dot{L}_x\left(1+u\right) = \frac{1}{2}\ln^2(1+u).$$

Using (37), (39) and remarking that $R_x(u) = \ln(1+u) - u + u^2/2$ yield

$$\dot{L}_x\left(1+u\right) = \frac{u^2}{2} + \bar{R}_x(u), \tag{40}$$

where

$$\begin{aligned}\bar{R}_x(u) &= \frac{x-2}{2}u^3 - \frac{x-1}{4}u^4 + u[R_x(u)+R_0(u)] - \frac{1}{2}R_x(u)u^2 + \frac{R_0(u)-R_x(u)}{x} \\ &+ \frac{x-1}{2}R_0(u)u^2 + R_x(u)R_0(u).\end{aligned} \tag{41}$$

Taking account of

$$\lim_{x\to 0}\frac{R_0(u)-R_x(u)}{x} = -\dot{K}_0(1+u) + \frac{u^2}{2} = -\frac{1}{2}\ln^2(1+u) + \frac{u^2}{2},$$

it follows that

$$\bar{R}_0(u) = \lim_{x\to 0}\bar{R}_x(u) = \dot{L}_0(1+u) - \frac{u^2}{2} = \frac{u^4}{8} + \frac{R_0^2(u)}{2} - \frac{u^3}{2} + uR_0(u) - \frac{u^2R_0(u)}{2}.$$

The second step is to focus on the integral $\mathcal{I}_1(t,x)$. From (40), it can be rewritten as

$$\mathcal{I}_1(t,x) = 1 + \int_0^1 t^2\bar{R}_x\left(\frac{\ln(1/s)}{t}\right)ds.$$

Now, using (41),

$$\begin{aligned}\int_0^1 t^2\bar{R}_x\left(\frac{\ln(1/s)}{t}\right)ds &= \frac{3(x-2)}{t} - \frac{6(x-1)}{t^2} + \int_0^1 \ln(1/s)t\left[R_x\left(\frac{\ln(1/s)}{t}\right) + R_0\left(\frac{\ln(1/s)}{t}\right)\right]ds \\ &- \frac{1}{2}\int_0^1 \ln^2(1/s)R_x\left(\frac{\ln(1/s)}{t}\right)ds + \frac{1}{x}\int_0^1 t^2\left[R_0\left(\frac{\ln(1/s)}{t}\right) - R_x\left(\frac{\ln(1/s)}{t}\right)\right]ds \\ &+ \frac{x-1}{2}\int_0^1 \ln^2(1/s)R_0\left(\frac{\ln(1/s)}{t}\right)ds + \int_0^1 t^2R_x\left(\frac{\ln(1/s)}{t}\right)R_0\left(\frac{\ln(1/s)}{t}\right)ds.\end{aligned}$$

It is clear that the first two terms converge to 0 as $t\to\infty$ uniformly on $x\in I$. Considering the third term, one has by (38) that

$$0 \le \int_0^1 \ln(1/s)t\left[R_x\left(\frac{\ln(1/s)}{t}\right) + R_0\left(\frac{\ln(1/s)}{t}\right)\right]ds \le \frac{4(x-1)(x-2)}{t^2} + \frac{8}{t^2}.$$

As a consequence, the third term also converges to 0 as $t\to\infty$ uniformly on $x\in I$. A similar proof can be done for the fifth, sixth and seventh terms. Considering the fourth term, let us remark that the function $x\to x^{-1}[R_0(u)-R_x(u)]$ is decreasing for all $u>0$. Thus, for all $x\in I =: [x_1, x_2]$,

$$\mathcal{J}_{x_2}(t) \le \frac{1}{x}\int_0^1 t^2\left[R_0\left(\frac{\ln(1/s)}{t}\right) - R_x\left(\frac{\ln(1/s)}{t}\right)\right]ds \le \mathcal{J}_{x_1}(t)$$

where we have introduced

$$\mathcal{J}_x(t) := \frac{1}{x}\int_0^1 t^2\left[R_0\left(\frac{\ln(1/s)}{t}\right) - R_x\left(\frac{\ln(1/s)}{t}\right)\right]ds$$

for $x \neq 0$ and

$$\mathcal{J}_0(t) := \lim_{x\to 0} \mathcal{J}_x(t) = -\frac{1}{2}\int_0^1 t^2 \ln^2\left(1+\frac{\ln(1/s)}{t}\right) ds + 1.$$

From (38), one can show that $\mathcal{J}_x(t) \to 0$ as $t \to \infty$ uniformly on $x \in I$. Finally $\mathcal{I}_1(t,x) \to 1$ as $t \to \infty$, uniformly on $x \in I$. A similar proof can be done to show that $\mathcal{I}_2(t,x) \to 6$ as $t \to \infty$, uniformly on $x \in I$. Moreover, Lemma 5(i) implies that $t[t^2\mu_2(t,x)]^2\Psi_t'(x) \to -2$ and that $[t^2\mu_2(t,x)]^2 \to 4$ as $t \to \infty$ uniformly on $x \in I$. The result is thus proved: $\Psi_t'(x) \to -1/2$ as $t \to \infty$ uniformly on $x \in I$. ■

Proof of Lemma 6 – (i) For $i = 1, \ldots, m$, let

$$Y_{m,i} := t_m^{J-1} F_i \left(1+\frac{F_i}{t_m}\right)^{\zeta_1 - 1} \prod_{j=2}^{J} L_{\zeta_j}\left(1+\frac{F_i}{\delta t_m}\right).$$

Let $\underline{\zeta} := \min\{\xi_1, \xi_2\}$. The following inequalities hold:

$$\frac{F_i}{\delta t_m} + \frac{\underline{\zeta}-1}{2}\frac{F_i^2}{\delta^2 t_m^2} \le \frac{F_i}{\delta t_m} + \frac{\zeta_j - 1}{2}\frac{F_i^2}{\delta^2 t_m^2} \le L_{\zeta_j}\left(1+\frac{F_i}{\delta t_m}\right) \le \frac{F_i}{\delta t_m}$$

Since $\max\{F_1, \ldots, F_m\} - \ln(m)$ converges in distribution to a Gumbel random variable,

$$0 \le \max_{1\le i\le m} \frac{F_i}{t_m} = \frac{\ln(m)}{t_m}\left(1 + \mathcal{O}_{\mathbb{P}}\left(\frac{1}{\ln(m)}\right)\right) = o_{\mathbb{P}}(1), \tag{42}$$

by assumption. It follows that uniformly in $i \in \{1, \ldots, m\}$,

$$Y_{m,i} = \delta^{1-J} F_i^J \left(1 + o_{\mathbb{P}}(1)\right).$$

The law of large numbers entails that

$$\frac{1}{m}\sum_{i=1}^m F_i^J \xrightarrow{\mathbb{P}} \mathbb{E}(F_1^J) = J!,$$

and the conclusion follows.
(ii) Since for all $\xi < 1$ and $u > 0$,

$$L_\xi(1+u) = u - \frac{\xi - 1}{2}u^2 + R_\xi(u),$$

with $0 \le R_\xi(u) \le (\xi-1)(\xi-2)u^3/6$ and taking into account of (42), it follows that uniformly in $i \in \{1, \ldots, m\}$

$$L_{\xi_3}\left(1+\frac{F_i}{\delta t_m}\right) - L_{\xi_4}\left(1+\frac{F_i}{\delta t_m}\right) = \frac{\xi_3 - \xi_4}{2}\left(\frac{F_i}{\delta t_m}\right)^2 (1 + o_{\mathbb{P}}(1)).$$

The rest of the proof follows the same lines as the one of (i) and is thus omitted. ■

Proof of Lemma 7 – Using the Cramér-Wold device, it suffices to obtain the asymptotic distribution of

$$T_n := k_n^{1/2}\left\{\beta_1 \frac{S_n(\zeta^{(1)})}{\mu(\ln(n/k_n), \zeta^{(1)})} + \beta_2 \frac{S_n(\zeta^{(2)})}{\mu(\ln(n/k_n), \zeta^{(2)})} - (\beta_1 + \beta_2)\right\},$$

where $(\beta_1,\beta_2)\in\mathbb{R}^2$. Let us introduce the random processes indexed by $t>0$

$$W_{i,n}(t) := \frac{\beta_1}{\mu(\ln(n/k_n),\zeta^{(1)})}\prod_{j=1}^{J_1} L_{\zeta_j^{(1)}}\left(1+\frac{F_i}{t}\right) + \frac{\beta_2}{\mu(\ln(n/k_n),\zeta^{(2)})}\prod_{j=1}^{J_2} L_{\zeta_j^{(2)}}\left(1+\frac{F_i}{t}\right),$$

for $i=1,\dots,k_n$ and where $F_1,\dots,F_{k_n}$ are independent standard exponential random variables. Lemma 4(ii) yields

$$T_n \stackrel{d}{=} k_n^{-1/2}\sum_{i=1}^{k_n}\left\{W_{i,n}(E_{n-k_n,n}) - \mathbb{E}\left[W_{i,n}(\ln(n/k_n))\right]\right\}.$$

where $E_{n-k_n,n}$ is the ordered statistic associated to a sample $E_1,\dots,E_n$ of standard exponential random values independent of $F_1,\dots,F_{k_n}$. Let us consider the following expansion $T_n =: T_{n,1}+T_{n,2}$ with

$$T_{n,1} := k_n^{-1/2}\sum_{i=1}^{k_n}\bar{W}_{i,n}(\ln(n/k_n))$$

where $\bar{W}_{i,n}(\ln(n/k_n)) := W_{i,n}(\ln(n/k_n)) - \mathbb{E}\left[W_{i,n}(\ln(n/k_n))\right]$ and

$$T_{n,2} := k_n^{-1/2}\sum_{i=1}^{k_n}\left\{W_{i,n}(E_{n-k_n,n}) - W_{i,n}(\ln(n/k_n))\right\}.$$

The asymptotic normality of the random term $T_{n,1}$ is obtained by Lyapunov's theorem. Let us observe that

$$s_n^2 := \operatorname{Var}\left(\sum_{i=1}^{k_n}\bar{W}_{i,n}(\ln(n/k_n))\right) = k_n\left\{\mathbb{E}\left[W_{1,n}^2(\ln(n/k_n))\right] - (\beta_1+\beta_2)^2\right\}.$$

Straightforward calculations then lead to

$$\begin{aligned}\mathbb{E}\left[W_{1,n}^2(\ln(n/k_n))\right] &= \beta_1^2\frac{\mu(\ln(n/k_n),(\zeta^{(1)},\zeta^{(1)}))}{\mu^2(\ln(n/k_n),\zeta^{(1)})} + \beta_2^2\frac{\mu(\ln(n/k_n),(\zeta^{(2)},\zeta^{(2)}))}{\mu^2(\ln(n/k_n),\zeta^{(2)})} \\ &+ 2\beta_1\beta_2\frac{\mu(\ln(n/k_n),(\zeta^{(1)},\zeta^{(2)}))}{\mu(\ln(n/k_n),\zeta^{(1)})\mu(\ln(n/k_n),\zeta^{(2)})}.\end{aligned}$$

As a direct consequence of Lemma 5(i), one has

$$\lim_{n\to\infty}\mathbb{E}\left[W_{1,n}^2(\ln(n/k_n))\right] = \frac{(2J_1)!}{(J_1!)^2}\beta_1^2 + \frac{(2J_2)!}{(J_2!)^2}\beta_2^2 + 2\frac{(J_1+J_2)!}{J_1!J_2!}\beta_1\beta_2, \tag{43}$$

and thus $s_n^2\sim c(\beta_1,\beta_2)k_n$ as $n\to\infty$ where the constant $c(\beta_1,\beta_2)$ is given by

$$c(\beta_1,\beta_2) := [(2J_1)!/(J_1!)^2-1]\beta_1^2 + [(2J_2)!/(J_2!)^2-1]\beta_2^2 + 2[(J_1+J_2)!/(J_1!J_2!)-1]\beta_1\beta_2.$$

Let us now check Lyapunov's condition *i.e.* that

$$\frac{1}{k_n^2}\sum_{i=1}^{k_n}\mathbb{E}\left[\bar{W}_{i,n}^4(\ln(n/k_n))\right] = \frac{1}{k_n}\mathbb{E}\left[\bar{W}_{1,n}^4(\ln(n/k_n))\right]\to 0, \tag{44}$$

as $n\to\infty$. By similar arguments as the ones leading to (43), one can show that

$$\mathbb{E}\left[\bar{W}_{1,n}^4(\ln(n/k_n))\right] = \sum_{l=1}^{4}(-1)^l C_4^l\,\mathbb{E}\left[W_{1,n}^l(\ln(n/k_n))\right]\mathbb{E}^{4-l}\left[W_{1,n}(\ln(n/k_n))\right]$$

converges to a constant as $n \to \infty$ and thus (44) holds. As a conclusion

$$T_{n,1} \xrightarrow{d} \mathcal{N}\left(0, c(\beta_1, \beta_2)\right). \tag{45}$$

It remains to prove that $T_{n,2} \xrightarrow{\mathbb{P}} 0$. For $i = 1, \ldots, k_n$, let $\dot{W}_{i,n}(\cdot)$ be the first derivative of the random function $W_{i,n}(\cdot)$. The mean-value theorem entails that

$$W_{i,n}(E_{n-k_n,n}) - W_{i,n}(\ln(n/k_n)) = \dot{W}_{i,n}(E^*_{n,i})[E_{n-k_n,n} - \ln(n/k_n)],$$

where for $i = 1, \ldots, k_n$, $E^*_{n,i} = \ln(n/k_n) + \Theta_{n,i}[E_{n-k_n,n} - \ln(n/k_n)]$ with $\Theta_{n,i}$ a random variable in $(0, 1)$. Recalling that $k_n^{1/2}[E_{n-k_n,n} - \ln(n/k_n)] \xrightarrow{d} \mathcal{N}(0, 1)$, in order to show that $T_{n,2} \xrightarrow{\mathbb{P}} 0$, it suffices to prove that

$$\frac{1}{k_n} \sum_{i=1}^{k_n} \dot{W}_{i,n}(E^*_{n,i}) \xrightarrow{\mathbb{P}} 0. \tag{46}$$

First, simple calculations show that for all $t > 0$,

$$\begin{aligned}
-\frac{1}{k_n} \sum_{i=1}^{k_n} \dot{W}_{i,n}(t) \quad &= \quad \frac{\beta_1}{t^2 \mu(\ln(n/k_n), \zeta^{(1)})} \sum_{l=1}^{J_1} \left[\frac{1}{k_n} \sum_{i=1}^{k_n} F_i \left(1 + \frac{F_i}{t}\right)^{\zeta_l^{(1)} - 1} \prod_{j \neq l} L_{\zeta_j^{(1)}} \left(1 + \frac{F_i}{t}\right) \right] \\
&+ \quad \frac{\beta_2}{t^2 \mu(\ln(n/k_n), \zeta^{(2)})} \sum_{l=1}^{J_2} \left[\frac{1}{k_n} \sum_{i=1}^{k_n} F_i \left(1 + \frac{F_i}{t}\right)^{\zeta_l^{(2)} - 1} \prod_{j \neq l} L_{\zeta_j^{(2)}} \left(1 + \frac{F_i}{t}\right) \right].
\end{aligned}$$

Hence, we have to deal with random terms proportional to $T_{3,n}(t, t)$ where

$$T_{3,n}(t_1, t_2) := \frac{1}{t_2^2 \mu(\ln(n/k_n), \zeta)} \frac{1}{k_n} \sum_{i=1}^{k_n} F_i \left(1 + \frac{F_i}{t_1}\right)^{\zeta_1 - 1} \prod_{j=2}^{J} L_{\zeta_j} \left(1 + \frac{F_i}{t_2}\right)$$

with $J \in \mathbb{N} \setminus \{0\}$ and $\zeta \in \mathbb{R}^J$. To prove (46) let us check that

$$\bar{T}_{3,n} := \frac{1}{[E^*_{n,i}]^2 \mu(\ln(n/k_n), \zeta)} \frac{1}{k_n} \sum_{i=1}^{k_n} F_i \left(1 + \frac{F_i}{E^*_{n,i}}\right)^{\zeta_1 - 1} \prod_{j=2}^{J} L_{\zeta_j} \left(1 + \frac{F_i}{E^*_{n,i}}\right) \xrightarrow{\mathbb{P}} 0. \tag{47}$$

For all $\eta > 0$, let $0 < \varepsilon < \min\{\eta_+, \eta_-\}$ where $\eta_+ = 1 - [(1 + \eta/2)/(1 + \eta)]^{1/2}$ and $\eta_- = [(1 - \eta/2)/(1 - \eta)]^{1/2} - 1$. Let us also introduce the Borel set

$$A_{n,\varepsilon} = \left\{ \left| \frac{E_{n-k_n,n} - \ln(n/k_n)}{\ln(n/k_n)} \right| \leq \varepsilon \right\}. \tag{48}$$

For all $\eta > 0$, remark that

$$\mathbb{P}\left[|\bar{T}_{n,3}| > \eta\right] \leq \mathbb{P}(A^{\mathrm{C}}_{n,\varepsilon}) + \mathbb{P}\left[\{|\bar{T}_{n,3}| > \eta\} \cap A_{n,\varepsilon}\right] \tag{49}$$

and recall that, for $i = 1, \ldots, k_n$, $E^*_{n,i} = \ln(n/k_n) + \Theta_{i,n}[E_{n-k_n,n} - \ln(n/k_n)]$. Since $\Theta_{n,i} \in (0, 1)$ it is clear that under $A_{n,\varepsilon}$, one has

$$(1 - \varepsilon) \ln(n/k_n) \leq E^*_{n,i} \leq (1 + \varepsilon) \ln(n/k_n)$$

for all $i = 1, \ldots, k_n$. Hence

$$T_{n,3}\left((1 - \varepsilon) \ln(n/k_n), (1 + \varepsilon) \ln(n/k_n)\right) \leq \bar{T}_{n,3} \leq T_{n,3}\left((1 + \varepsilon) \ln(n/k_n), (1 - \varepsilon) \ln(n/k_n)\right),$$

and thus

$$\begin{aligned}\mathbb{P}\left[\{|\bar{T}_{n,3}| > \eta\} \cap A_{n,\varepsilon}\right] &\leq \mathbb{P}\left[T_{n,3}\left((1+\varepsilon)\ln(n/k_n), (1-\varepsilon)\ln(n/k_n)\right) > \eta\right] \\ &+ \mathbb{P}\left[T_{n,3}\left((1-\varepsilon)\ln(n/k_n), (1+\varepsilon)\ln(n/k_n)\right) < -\eta\right].\end{aligned}$$

Applying Lemma 6(i) and the fact that from Lemma 5(i),

$$\frac{\ln^{1-J}(n/k_n)}{\ln^2(n/k_n)\mu(\ln(n/k_n),\zeta)} \sim \frac{1}{J!}\frac{1}{\ln(n/k_n)} \to 0,$$

it follows that $T_{n,3}\left((1\pm\varepsilon)\ln(n/k_n), (1\mp\varepsilon)\ln(n/k_n)\right) \xrightarrow{\mathbb{P}} 0$. As a consequence,

$$\mathbb{P}\left[\{|\bar{T}_{n,3}| > \eta\} \cap A_{n,\varepsilon}\right] \to 0 \tag{50}$$

as $n \to \infty$. Furthermore, since $[E_{n-k_n,n} - \ln(n/k_n)]/\ln(n/k_n) = \mathcal{O}_{\mathbb{P}}[k_n^{-1/2}\ln^{-1}(n/k_n)] = o_{\mathbb{P}}(1)$, one has that $\mathbb{P}(A_{n,\varepsilon}) \xrightarrow{\mathbb{P}} 1$ as $n \to \infty$. Collecting this last result, (49) and (50) implies (47) and thus (46) and the conclusion follows. ■

Proof of Lemma 8 − Let $E_1, \dots, E_n, F_1, \dots, F_n$ be a sample of $2n$ independent standard exponential random variables and let $E_{n-k_n,n}$ be the $(n-k_n)$th ordered statistic associated with the sample $E_1, \dots, E_n$. Let us also introduce the random variable defined for all $t > 0$ by

$$W_n(t) := \frac{[\ln(n/k_n)]^J}{k_n}\sum_{i=1}^{k_n} L_{\xi_1}^{J_1}\left(1+\frac{F_i}{t}\right) L_{\xi_2}^{J_2}\left(1+\frac{F_i}{t}\right)\left[L_{\xi_3}\left(1+\frac{F_i}{t}\right) - L_{\xi_4}\left(1+\frac{F_i}{t}\right)\right]^{J_3}.$$

According to Lemma 4(ii), we have to prove that $W_n(E_{n-k_n,n}) \xrightarrow{\mathbb{P}} K := J![(\xi_3-\xi_4)/2]^{J_3}$. For all $\varepsilon > 0$, let us consider the Borel set

$$\mathcal{A}_{n,\varepsilon} := \left\{\left|\frac{E_{n-k_n,n}}{\ln(n/k_n)} - 1\right| \leq \varepsilon\right\},$$

introduced in (48). Remarking that the function $x \mapsto L_{\xi_1}^{J_1}(x)L_{\xi_2}^{J_2}(x)[L_{\xi_3}(x) - L_{\xi_4}(x)]^{J_3}$ is increasing on $(1,\infty)$ leads to

$$\begin{aligned}\mathbb{P}\left\{|W_n(E_{n-k_n,n}) - K| > \varepsilon\right\} &\leq \mathbb{P}\left\{W_n[(1+\varepsilon)\ln(n/k_n)] > K+\varepsilon\right\} \\ &+ \mathbb{P}\left\{W_n[(1-\varepsilon)\ln(n/k_n)] < K-\varepsilon\right\} + 1 - \mathbb{P}(\mathcal{A}_{n,\varepsilon}).\end{aligned} \tag{51}$$

Using the inequality $K + \varepsilon - K/(1+\varepsilon)^J \geq \varepsilon$ yields

$$\mathbb{P}\left\{W_n[(1+\varepsilon)\ln(n/k_n)] > K+\varepsilon\right\} \leq \mathbb{P}\left\{W_n[(1+\varepsilon)\ln(n/k_n)] - \frac{K}{(1+\varepsilon)^J} > \varepsilon\right\}.$$

Since $\ln(k_n)/\ln(n) \to 0$, one can apply Lemma 6(ii) with $m = k_n$ and $t_m = \ln(n/k_n)$ to obtain

$$\lim_{n\to\infty} \mathbb{P}\left\{W_n[(1+\varepsilon)\ln(n/k_n)] > K+\varepsilon\right\} = 0. \tag{52}$$

In the same way, one has

$$\lim_{n\to\infty} \mathbb{P}\left\{W_n[(1-\varepsilon)\ln(n/k_n)] < K-\varepsilon\right\} = 0. \tag{53}$$

Finally, since $E_{n-k_n,n}/\ln(n/k_n) \xrightarrow{\mathbb{P}} 1$, we conclude the proof by collecting (51) to (53). ■

	$\bar{F}(x)$	θ	ρ	ρ'
$\boldsymbol{\theta = 0}$				
Gamma $(a > 0, s > 0)$	$\frac{1}{s^a\Gamma(a)}\int_x^\infty t^{a-1}e^{-t/s}dt$ $x \geq 0$	0	-1	0
Weibull $(k \neq 1, \lambda > 0)$	$e^{-(x/\lambda)^k}$ $x \geq 0$	0	$-\infty$	0
Gaussian $(\mu \in \mathbb{R}, \sigma > 0)$	$\frac{1}{\sigma\sqrt{2\pi}}\int_x^\infty \exp\left(-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right) dt$ $x \in \mathbb{R}$	0	-1	0
$\boldsymbol{\theta > 0}$				
Lognormal $(\mu \in \mathbb{R}, \sigma > 0)$	$\frac{1}{\sigma\sqrt{2\pi}}\int_x^\infty \frac{1}{t}\exp\left(-\frac{(\ln t-\mu)^2}{2\sigma^2}\right) dt$ $x \geq 0$	$1/2$	-1	-1
Burr $(\lambda > 0, c > 0, k > 0)$	$\left[1+\left(\frac{x}{\lambda}\right)^c\right]^{-k}$ $x \geq 0$	1	$-\infty$	-1
Pareto-like	$1/U^{\leftarrow}(x)$, $U(x) = x\left(1+2\ln^2(x)\right)$	1	-1	-1
Super heavy-tail	$e^{-\ln^{1/2}(x)}$ $x \geq 1$	2	$-\infty$	$-\infty$
$\boldsymbol{\theta < 0}$				
Finite endpoint $(x^* > 0)$	$\exp\left(-\frac{1}{\ln x^* - \ln x}\right)$ $x \in (0, x^*)$	-1	$-\infty$	-1

Table 1: Examples of distributions verifying **(A1)** and **(A2)** with associated values of θ, ρ and ρ' (see (12)).

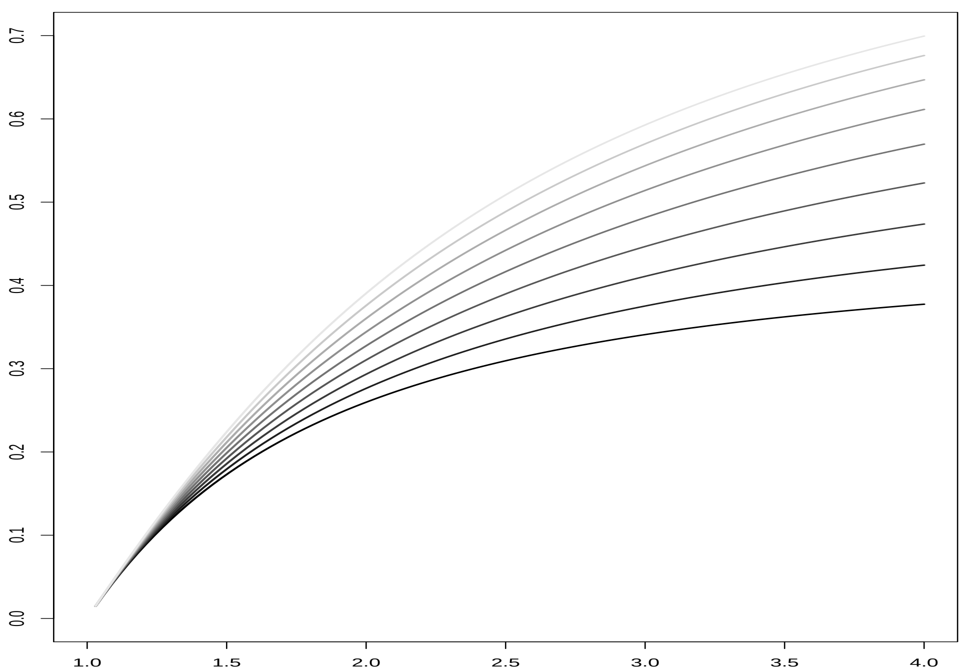

Figure 1: Ratio $\Lambda_\theta(\tau)$ between the asymptotic standard deviations σ_n and σ'_n (see equation (17)) as a function of $\tau \geq 1$ for $\theta \in \{-2, -1.5, \ldots, 2\}$. Dark lines are associated with small values of θ.

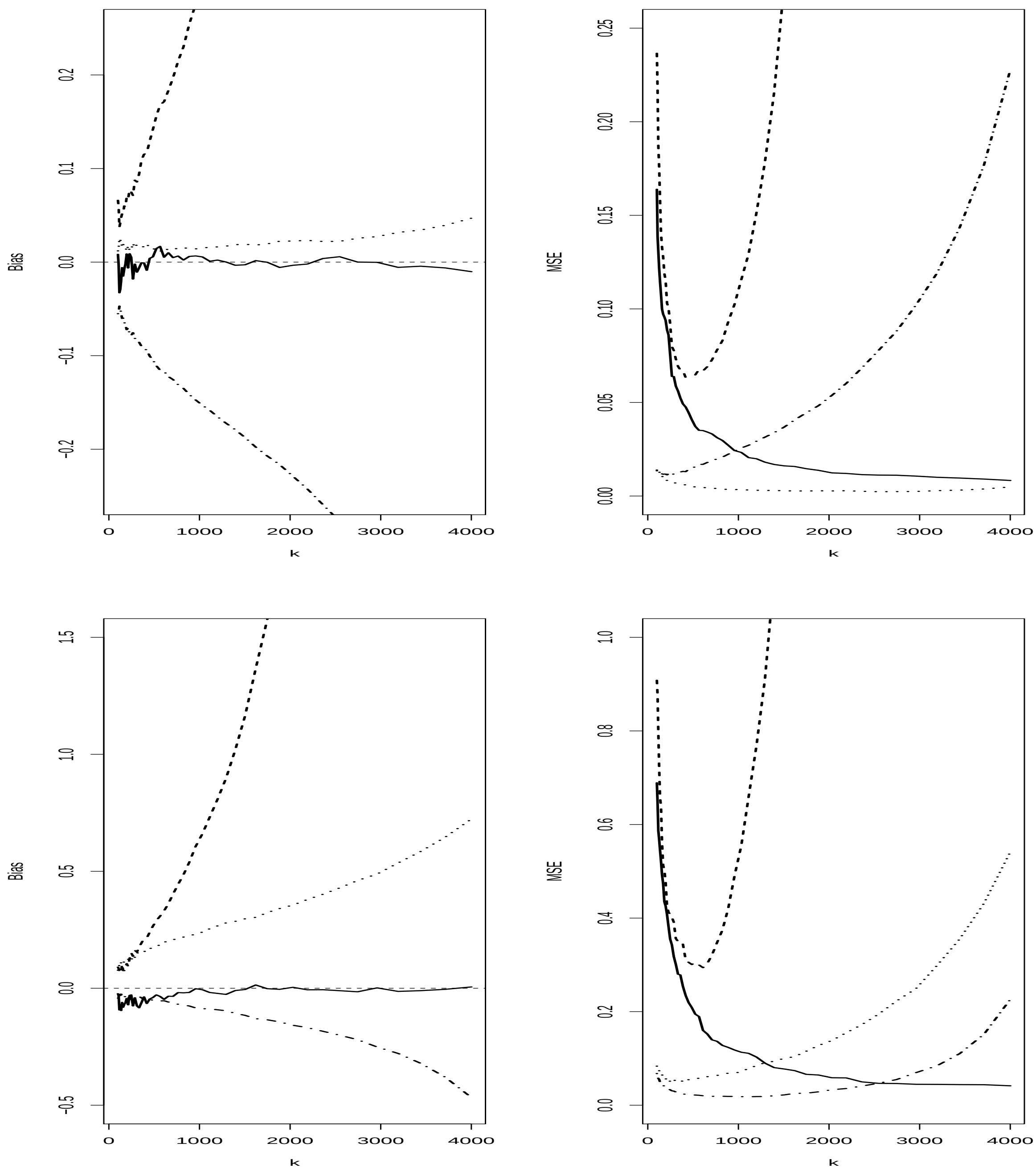

Figure 2: Results on simulated data. Bias (left) and MSE (right) associated with $\check{Q}_n^{[1]}(\beta_n)$ (solid line), $\check{Q}_n^{[2]}(\beta_n)$ (dotted line), $\check{Q}_n^{[3]}(\beta_n)$ (dash-dotted line) and $\check{Q}_n^{[4]}(\beta_n)$ (dashed line) as functions of k_n for $\beta_n = n^{-2}$ and $n = 5000$. Top: Gamma, bottom: Weibull.

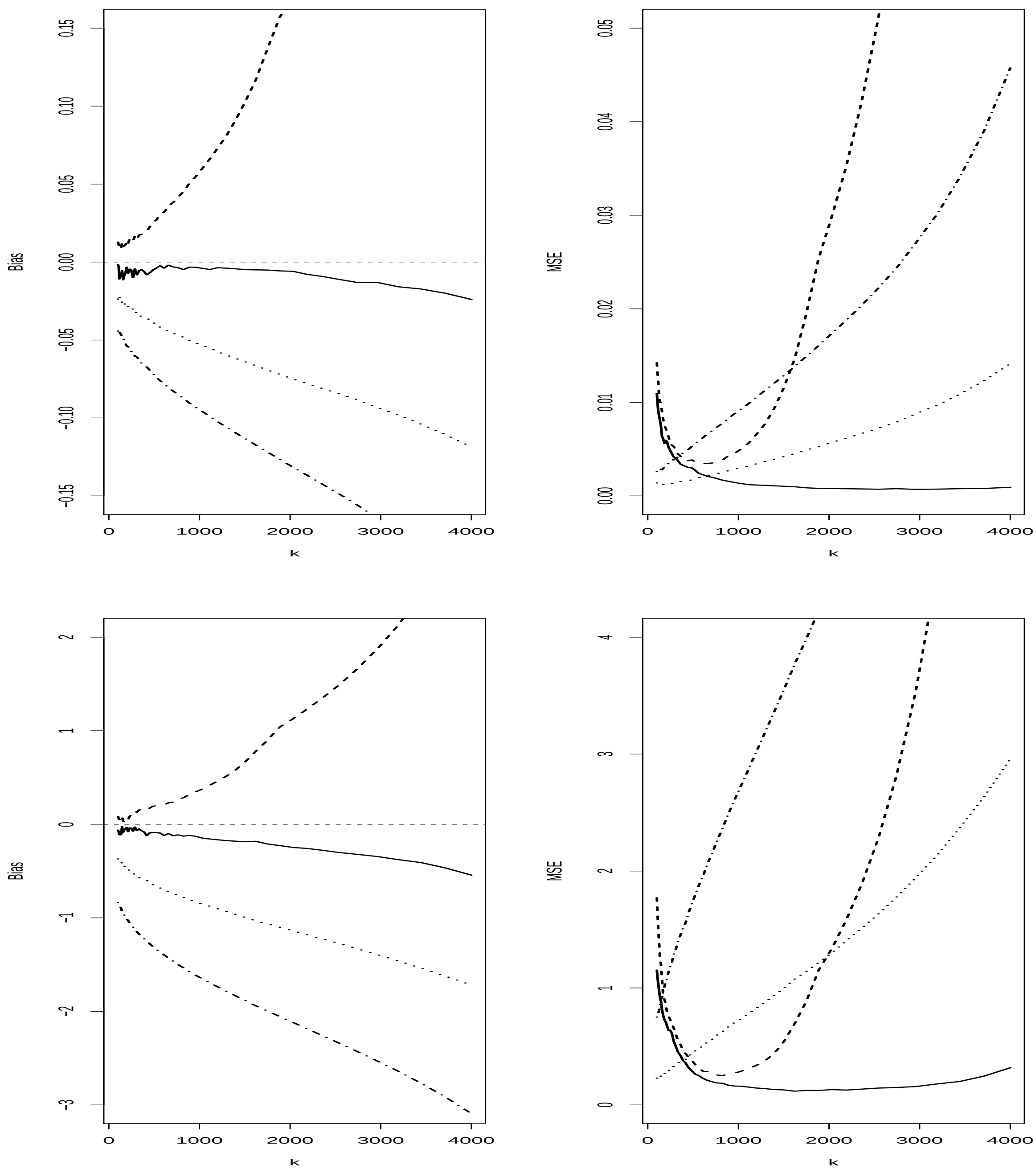

Figure 3: Results on simulated data. Bias (left) and MSE (right) associated with $\check{Q}_n^{[1]}(\beta_n)$ (solid line), $\check{Q}_n^{[2]}(\beta_n)$ (dotted line), $\check{Q}_n^{[3]}(\beta_n)$ (dash-dotted line) and $\check{Q}_n^{[4]}(\beta_n)$ (dashed line) as functions of k_n for $\beta_n = n^{-2}$ and $n = 5000$. Top: Gaussian, bottom: Lognormal.

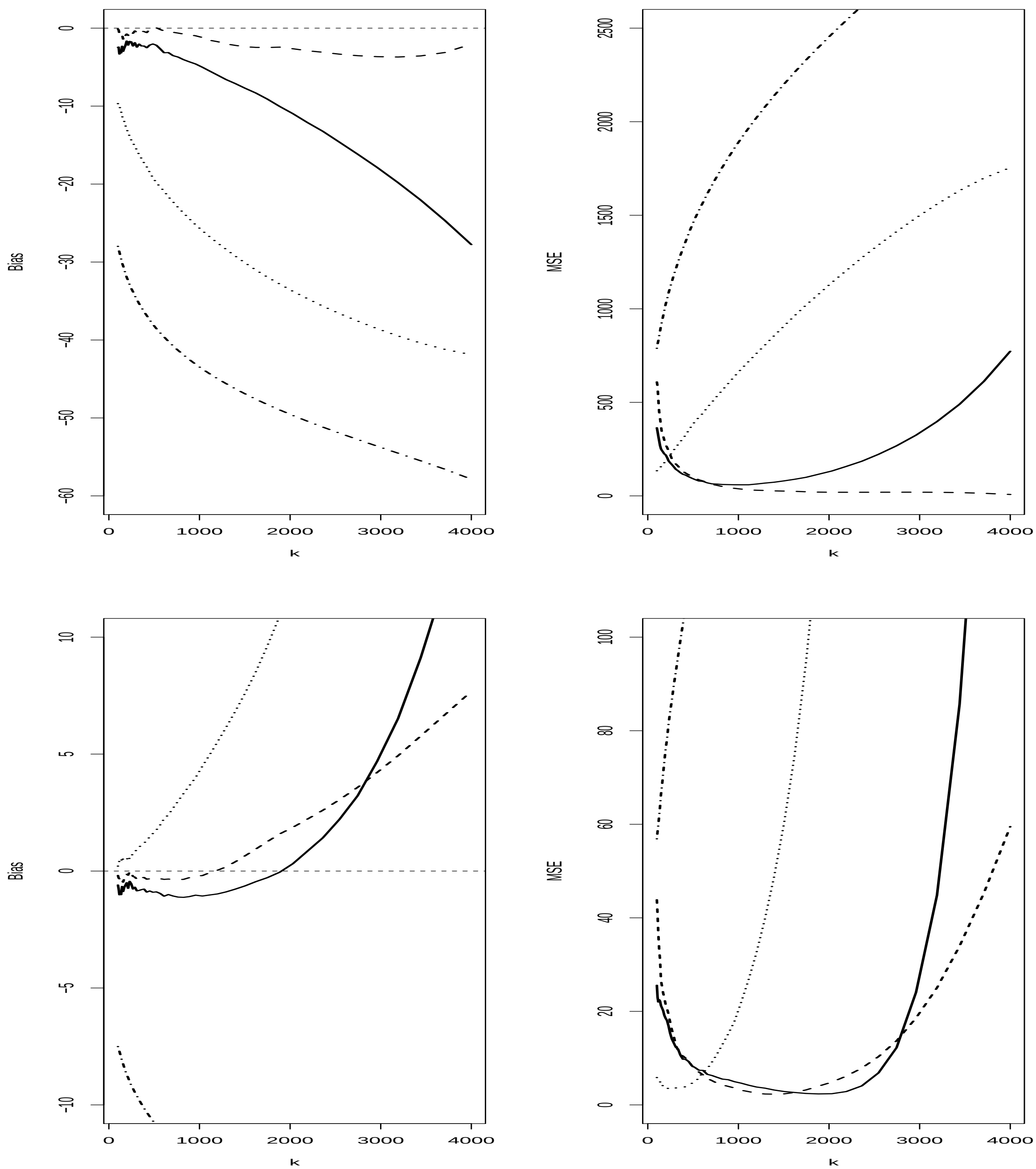

Figure 4: Results on simulated data. Bias (left) and MSE (right) associated with $\check{Q}_n^{[1]}(\beta_n)$ (solid line), $\check{Q}_n^{[2]}(\beta_n)$ (dotted line), $\check{Q}_n^{[3]}(\beta_n)$ (dash-dotted line) and $\check{Q}_n^{[4]}(\beta_n)$ (dashed line) as functions of k_n for $\beta_n = n^{-2}$ and $n = 5000$. Top: Burr, bottom: Pareto-like.

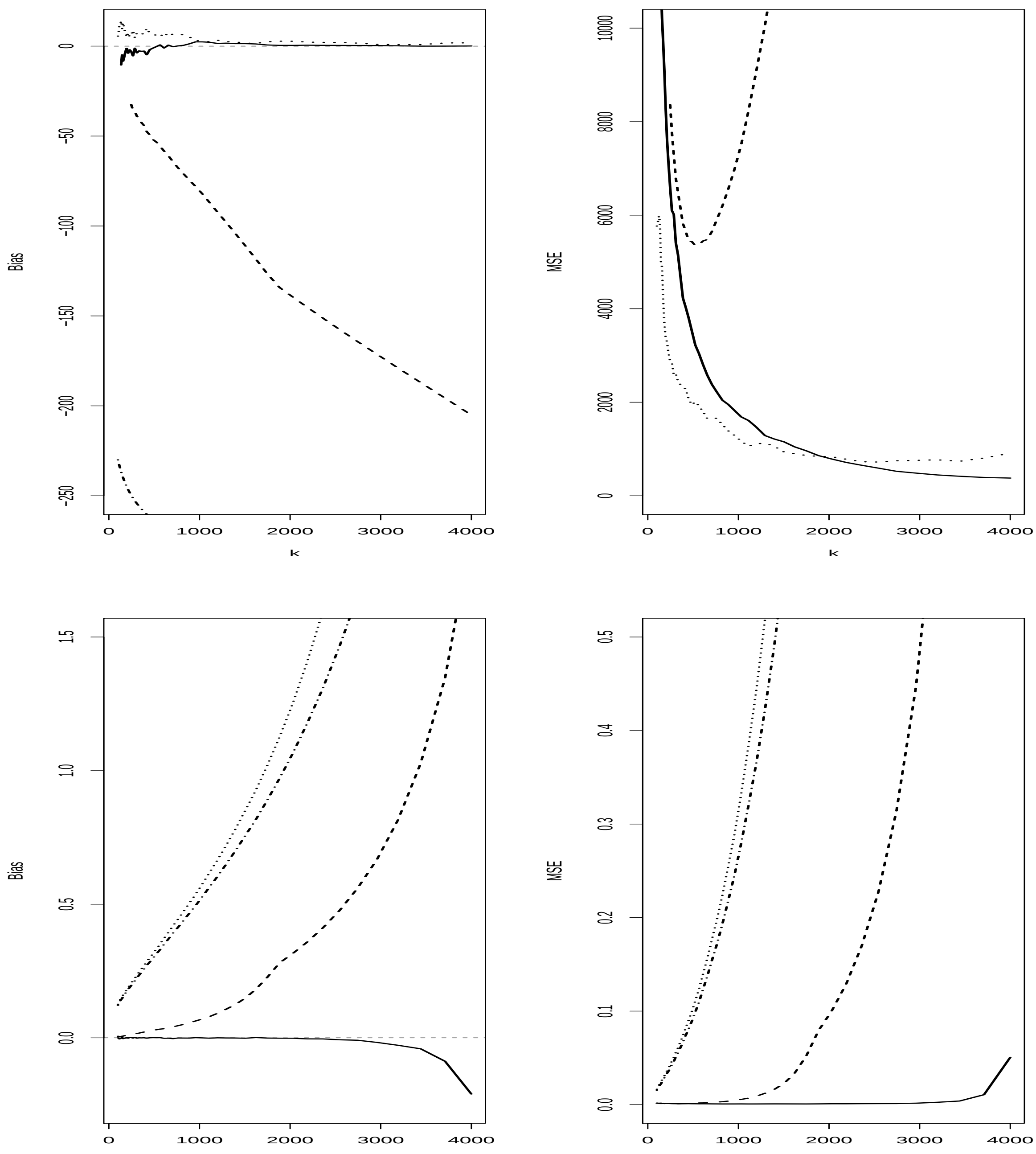

Figure 5: Results on simulated data. Bias (left) and MSE (right) associated with $\check{Q}_n^{[1]}(\beta_n)$ (solid line), $\check{Q}_n^{[2]}(\beta_n)$ (dotted line), $\check{Q}_n^{[3]}(\beta_n)$ (dash-dotted line) and $\check{Q}_n^{[4]}(\beta_n)$ (dashed line) as functions of k_n for $\beta_n = n^{-2}$ and $n = 5000$. Top: super heavy-tail, bottom: finite endpoint.

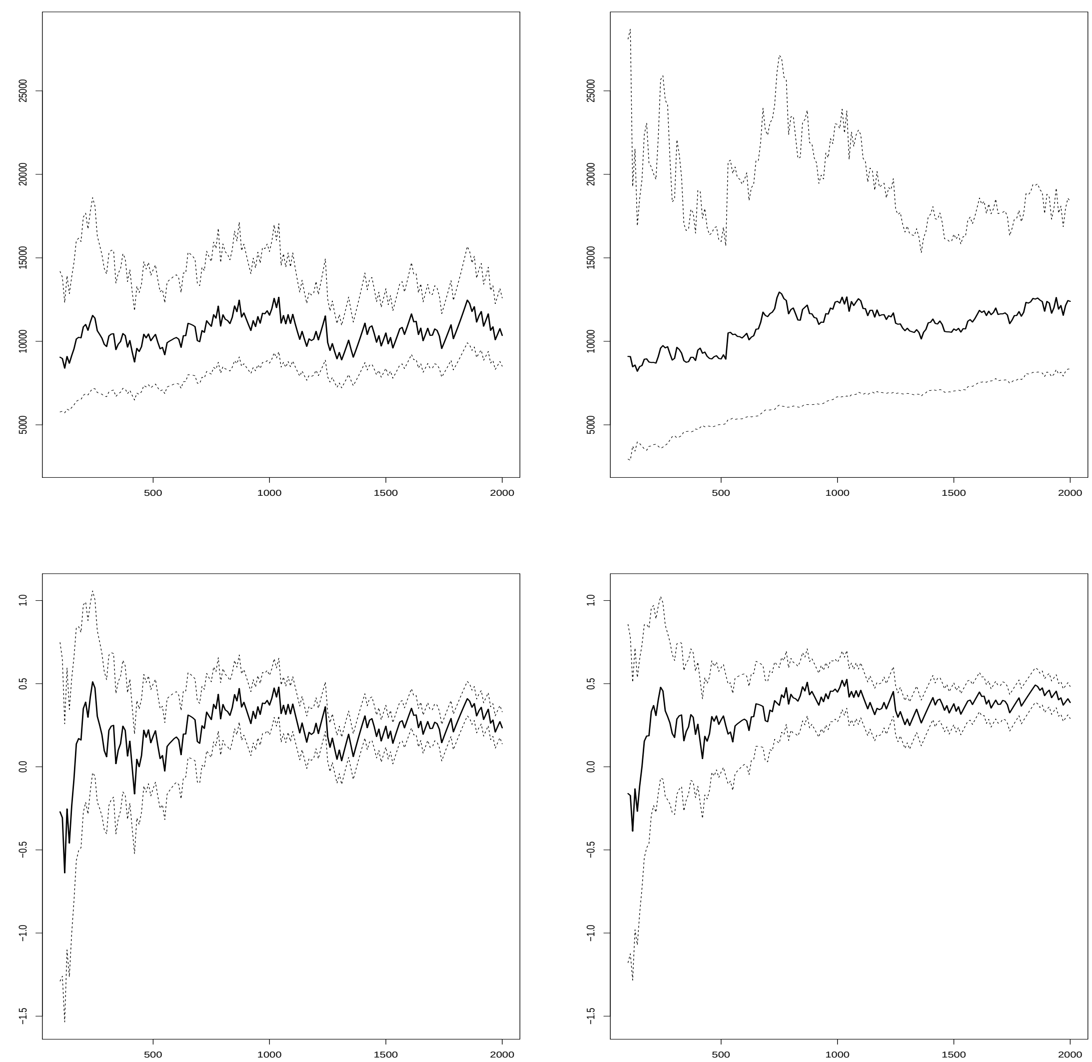

Figure 6: Results on Rhône data. Estimates $\check{Q}_n^{[1]}(\beta_n)$ (top left) and $\check{Q}_n^{[4]}(\beta_n)$ (top right) of $Q(\beta_n)$ with ($\beta_n = 5.5\ 10^{-6}$) and their corresponding index estimates $\hat{\theta}_n^{(M)}$ (bottom left) and $\hat{\theta}_{k_n,n}^{[4]}$ (bottom right) as functions of $k \in \{100, \ldots, 2000\}$. The 95% asymptotic confidence intervals are depicted by dotted lines.