



HAL
open science

Revisiting Weissman estimator for extreme quantiles

Jonathan El Methni, Stéphane Girard

► **To cite this version:**

Jonathan El Methni, Stéphane Girard. Revisiting Weissman estimator for extreme quantiles. 2021.
hal-03266676v1

HAL Id: hal-03266676

<https://inria.hal.science/hal-03266676v1>

Preprint submitted on 22 Jun 2021 (v1), last revised 15 Mar 2022 (v3)

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

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

Revisiting Weissman estimator for extreme quantiles

Jonathan El Methni⁽¹⁾ and Stéphane Girard⁽²⁾

⁽¹⁾ Université de Paris, CNRS, MAP5, UMR 8145, 75006 Paris, France

⁽²⁾ Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France.

Abstract

Weissman extrapolation device for estimating extreme quantiles is based on two estimators: an order statistic to estimate an intermediate quantile and an estimator of the tail index. The common practice is to select the same intermediate sequence for both estimators. In this work, we show how an adapted choice of two different intermediate sequences leads to a reduction of the asymptotic bias associated with the resulting Weissman estimator. Our approach is compared to other bias reduced estimators of extreme quantiles on simulated data.

1 Motivation

Let X_1, \dots, X_n be an independent and identically distributed sample from a cumulative distribution function F supposed to be continuous and strictly increasing and let $X_{1,n} \leq \dots \leq X_{n,n}$ denote the order statistics associated with this sample. We denote by U the associate tail quantile function defined as $U(t) = F^{\leftarrow}(1 - 1/t)$ for all $t > 1$, where $F^{\leftarrow}(\cdot) = \inf\{x \in \mathbb{R}, F(x) > \cdot\}$ denotes the generalized inverse of F . In the following it is assumed that F belongs to the maximum domain attraction of Fréchet, which is equivalent to assuming that U is regularly varying with index $\gamma > 0$ that is

$$\lim_{t \rightarrow \infty} \frac{U(ty)}{U(t)} = y^\gamma, \quad (1)$$

for all $y > 0$. Recall that, equivalently, U can be rewritten as $U(t) = t^\gamma \ell(t)$ where ℓ is a slowly-varying function, *i.e.* a regularly-varying function with index zero. See [5] for more details on regular variation theory. In such a situation, the distribution associated with U is said to be heavy-tailed. The goal is to estimate the extreme quantile $U(1/p_n)$ where $0 < p_n < 1/n$ basing on an intermediate one $U(n/k_n)$ where (k_n) is an intermediate sequence *i.e.* such that $k_n \in \{1, \dots, n-1\}$, $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ as $n \rightarrow \infty$. In view of the regular variation property (1), one has

$$\frac{U(1/p_n)}{U(n/k_n)} \simeq \left(\frac{k_n}{np_n} \right)^\gamma,$$

as $n \rightarrow \infty$. Estimating $U(n/k_n)$ by its empirical counterpart $X_{n-k_n, n}$ and γ by a convenient estimator $\hat{\gamma}_n(k'_n)$ depending on another intermediate sequence (k'_n) yields

$$\hat{q}_n(p_n, k_n, k'_n) = X_{n-k_n, n} \left(\frac{k_n}{np_n} \right)^{\hat{\gamma}_n(k'_n)}. \quad (2)$$

One can for instance use the Hill estimator [22] defined as

$$H(k'_n) = \frac{1}{k'_n} \sum_{i=1}^{k'_n} \log(X_{n-i,n}) - \log(X_{n-k'_n,n}), \quad (3)$$

and choose $k'_n = k_n$ to get the original Weissman estimator [29]:

$$\hat{q}_n(p_n, k_n, k_n) = X_{n-k_n,n} \left(\frac{k_n}{np_n} \right)^{H(k_n)}. \quad (4)$$

The asymptotic properties of $\hat{q}_n(p_n, k_n, k_n)$ are established for instance in [19, Theorem 4.3.8]. In the next section, we show that choosing $k'_n \neq k_n$ can yield better results from an asymptotic point of view.

2 Weissman estimator revisited

2.1 Asymptotic analysis

A classical device in extreme-value analysis for bias quantification is the following second-order condition that refines the initial heavy-tail assumption (1). The tail quantile function U is assumed to be second-order regularly varying with index $\gamma > 0$, second-order parameter $\rho < 0$ and a measurable auxiliary function A having constant sign and converging to 0 at infinity, *i.e.*

$$\lim_{t \rightarrow \infty} \frac{1}{A(t)} \left[\frac{U(ty)}{U(t)} - y^\gamma \right] = y^\gamma \frac{y^\rho - 1}{\rho}, \quad (5)$$

for all $y > 0$. The auxiliary function A drives the bias of most extreme-value estimators. In particular, the asymptotic sign of A determines the sign of the asymptotic bias. Besides, necessarily, $|A|$ is regularly varying with index ρ at infinity, see for instance [19, Theorem 2.3.9]. The larger ρ is, the larger the (absolute) asymptotic bias. Numerous distributions satisfying assumption (5) can be found in [3], see also Table 1 for examples. Our first result is a refinement of [19, Theorem 4.3.8]. It provides an asymptotic normality result for the extreme quantile estimator (2) based on two intermediate sequences (k_n) and (k'_n) .

Theorem 1. *Assume the second-order condition (5) holds. Let (k_n) and (k'_n) be two intermediate sequences and (p_n) a sequence in $(0, 1)$ such that $p_n \rightarrow 0$ as $n \rightarrow \infty$. Let $d_n = k_n/(np_n)$. Suppose, as $n \rightarrow \infty$,*

- (i) $\sqrt{k'_n} A(n/k'_n) \rightarrow \lambda$,
- (ii) $\hat{\gamma}_n(k'_n)$ is an estimator of γ such that $\sqrt{k'_n}(\hat{\gamma}_n(k'_n) - \gamma) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2)$ where $\mu > 0$ and $\sigma > 0$,
- (iii) $d_n \rightarrow \infty$, $(\log d_n)/\sqrt{k'_n} \rightarrow 0$ and $(k'_n/k_n)^\rho / \log d_n \rightarrow c > 0$.

Then, as $n \rightarrow \infty$,

$$\frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(\lambda(\mu + c/\rho), \sigma^2). \quad (6)$$

If, moreover, $(\log d_n)/\sqrt{k'_n} = o(k_n^{-1/4})$, then the following asymptotic expansion holds:

$$\frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1 \right) = \lambda\mu + \lambda(k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n} \right) L(d_n) + \sigma\xi_n + \frac{\sqrt{k'_n/k_n}}{\log d_n} \gamma\xi'_n(1 + o_P(1)),$$

with L a slowly-varying function and where ξ_n and ξ'_n are asymptotically standard Gaussian random variables.

In view of (6), the revisited Weissman estimator inherits its asymptotic distribution from $\hat{\gamma}_n(k'_n)$ with an additional bias component compared to the original Weissman estimator [19, Theorem 4.3.8]. Let us focus for a while on the case where the auxiliary function in (5) is given by

$$A(t) = \beta\gamma t^\rho, \text{ with } \beta \neq 0. \quad (7)$$

This situation arises for instance in the Hall-Welsh class of heavy-tailed distributions [20, 21] defined by

$$U(t) = Ct^\gamma(1 + \gamma\beta t^\rho/\rho + o(t^\rho)), \text{ with } C > 0.$$

It then appears from the proof of Theorem 1 that $L(t) \rightarrow 1$ as $t \rightarrow \infty$ and the asymptotic squared bias of $\hat{q}_n(p_n, k_n, k'_n)$ is given by

$$AB^2(k_n, k'_n, p_n) = (\log d_n)^2 A^2(n/k'_n) \left(\mu + (k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n} \right) \right)^2.$$

The crucial point is that the asymptotic bias can be cancelled by letting

$$k_n^* := k'_n(\rho, k_n, p_n) = k_n \left(-\rho\mu \frac{\log d_n}{1 - d_n^\rho} \right)^{1/\rho}. \quad (8)$$

The next lemma describes the behavior of k_n^* as a function of k_n .

Lemma 1.

- (i) For all $d_n \geq 1$ and $\rho < 0$, k_n^* is an increasing function of k_n and $k_n^* \leq \mu^{1/\rho} k_n$.
- (ii) For all $\rho < 0$, $k_n^* \sim \tau k_n (\log d_n)^{1/\rho}$ as $n \rightarrow \infty$, where $\tau := (-\rho\mu)^{1/\rho}$.
- (iii) If, moreover, $np_n \rightarrow \beta \in (0, 1)$ as $n \rightarrow \infty$, then $k_n^* \sim \tau k_n (\log k_n)^{1/\rho}$.

From (iii), it appears that $k_n^*/k_n \rightarrow 0$ as $n \rightarrow \infty$, meaning that the number of upper order statistics in the tail-index estimator should be asymptotically small compared to k_n . See the next paragraph for a more precise discussion in the case of Hill estimator.

The next results provides the asymptotic distribution of the extreme quantile estimator (2) computed with $k'_n = k_n^*$.

Corollary 1. Assume the second-order condition (5) holds. Let (k_n) be an intermediate sequence and (p_n) a sequence in $(0, 1)$ such that $p_n \rightarrow 0$ as $n \rightarrow \infty$. Let $d_n = k_n/(np_n) \rightarrow \infty$ and define k_n^* as in (8). Let $\hat{\gamma}_n(k_n^*)$ be an estimator of γ such that $\sqrt{k_n^*}(\hat{\gamma}_n(k_n^*) - \gamma) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2)$ where $\mu > 0$ and $\sigma > 0$. Suppose, as $n \rightarrow \infty$,

$$\sqrt{k_n}(\log d_n)^{1/(2\rho)-1} A(n/k_n) \rightarrow \lambda \text{ and } (\log d_n)^{1-1/(2\rho)}/\sqrt{k_n} \rightarrow 0.$$

Then, as $n \rightarrow \infty$,

$$\frac{\sqrt{k_n^*}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k_n^*)}{q(p_n)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, \sigma^2). \quad (9)$$

As expected, the resulting asymptotic Gaussian distribution in (9) is centered, in contrast to (6).

2.2 Examples

Hill estimator. Let us first focus on the case where $\hat{\gamma}_n(k'_n)$ is the Hill estimator (3). The assumptions of the above Theorem are satisfied with $\mu = 1/(1 - \rho)$ and $\sigma = \gamma$, see for instance [19, Theorem 3.2.5], leading to

$$k_n^{\text{H},*} = k_n^{\text{H},*}(\rho, p_n, k_n) = k_n \left(\frac{-\rho \log d_n}{1 - \rho \frac{\log d_n}{1 - d_n^\rho}} \right)^{1/\rho}. \quad (10)$$

Remark that, when $\rho = -\infty$, then $k_n^{\text{H},*}(-\infty, p_n, k_n) = k_n$ and we find back the original Weissman estimator (4). At the opposite, when $\rho \rightarrow 0$, one has $k_n^{\text{H},*}(0, p_n, k_n) = ek_n/\sqrt{d_n} \sim e\sqrt{\beta}\sqrt{k_n}$ as $n \rightarrow \infty$ under the condition $np_n \rightarrow \beta \in (0, 1)$. It thus appears that, in the low bias situation, $k_n^{\text{H},*}$ and k_n should be chosen of the same order. Conversely, in the high bias situation, $k_n^{\text{H},*}$ should be chosen significantly smaller than k_n , see Figure 1 for an illustration in the case $d_n = k_n/2$ and $n = 1,000$. It appears that, the larger ρ is, the smaller $k_n^{\text{H},*}$ is, in order to dampen the extrapolation bias. The practical choice of $k_n^{\text{H},*}$ is discussed in the next paragraph.

Other examples. Similarly, Zipf estimator, based on a least-squares regression on the quantile-quantile plot and proposed simultaneously by [23, 28], fulfills the assumptions of Theorem 1 with $\mu = 1/(1 - \rho)^2$ and $\sigma^2 = 2\gamma^2$, see [2]. Finally, the maximum likelihood estimator, the moment [9] and Pickands estimator [26] all satisfy the assumptions of the above theorem, we refer respectively to [19, Theorem 3.4.2], [19, Theorem 3.5.4] and [19, Theorem 3.3.5] for the associated values of (μ, σ^2) .

In the sequel, we focus on the Hill estimator, which is the tail-index estimator with smallest variance among the above mentioned ones.

2.3 Implementation

In practice, the revisited Weissman estimator is computed using Hill estimator $H(\hat{k}_n^*)$ where $\hat{k}_n^* := k_n^{\text{H},*}(\hat{\rho}, p_n, k_n)$ is an estimation of the intermediate sequence given in (10). Here, we adopt the estimator $\hat{\rho}$ of the second-order parameter introduced in [1, Equation (2.18)] and implemented in the package `evt0` of the R software [24] which features a satisfying behaviour in practice. The resulting revisited Weissman estimator is denoted by $\text{RW}(k_n)$ in the sequel. We refer to [10] for a review on the estimation of ρ in the particular case of heavy-tailed distributions.

This data-driven choice of \hat{k}_n^* is illustrated on Figure 2 together with its consequences on the estimation of the tail-index (left panel) and extreme quantiles (right panel). It appears that, in the low bias situation ($\rho = -2$, top panel), \hat{k}_n^* is automatically limited to the range $\{1, \dots, 252\}$. In the high bias situation ($\rho = -3/4$, bottom panel), \hat{k}_n^* is further limited to the range $\{1, \dots, 118\}$. In both cases, the bias associated with Hill estimator $H(\hat{k}_n^*)$ remains acceptable. As a consequence, the relative mean-squared error associated with the revisited Weissman estimator is small for a wide range of values of k_n , in contrast to the original Weissman estimator. These preliminary numerical experiments are conducted in a more systematic way in the following paragraph.

3 Validation on simulated data

The proposed revisited Weissman estimator is compared on simulated data to the original Weissman estimator and to six other bias-reduced estimators of the extreme quantile.

Experimental design. The comparison is achieved on the following six heavy-tailed distributions: Burr, Fréchet, Fisher, Generalized Pareto Distribution, Inverse Gamma, and Student. All of them satisfy the second-order condition (5) with $A(t) = b\gamma t^\rho$, see Table 1 for their definitions and associated values of γ and ρ . For all distributions, four tail-index values $\gamma \in \{1/8; 1/4; 1/2; 1\}$ are investigated. The choice of the second-order parameter ρ depends on the considered distribution:

- Burr distribution: The second order parameter can be chosen independently from the tail-index, five values are tested $\rho \in \{-1/8; -1/4; -1/2; -1; -2\}$.
- Fréchet distribution: The second order parameter is fixed to $\rho = -1$.
- Fisher, Generalized Pareto Distribution and Inverse Gamma: The second order parameter is fixed to $\rho = -\gamma$.
- Student distribution: The second order parameter is fixed to $\rho = -2\gamma$.

As a result, $4 \times (5 + 5) = 40$ simulating distributions are considered. In each case, we simulate $N = 1,000$ replications of a dataset of $n = 500$ independent realisations from the considered parametric model. Finally, two cases are considered for the order of the extreme quantile: $p_n \in \{1/n; 1/(2n)\}$. Summarizing, this experimental design includes $40 \times 2 = 80$ configurations.

Competitors. Since the original Weissman estimator inherits its asymptotic distribution from $H(k_n)$, the main idea of most of reduced bias estimators of the extreme quantile is to replace Hill estimator $H(k_n)$ in (4) by a bias-reduced version. We shall first consider the Corrected-Hill (CH) [7]:

$$\text{CH}(k_n) = H(k_n) \left(1 - \frac{\hat{\beta}}{1 - \hat{\rho}} \left(\frac{n}{k_n} \right)^{\hat{\rho}} \right), \quad (11)$$

where $\hat{\rho}$ and $\hat{\beta}$ are estimators of the second-order parameters ρ and β , see (7). The associated bias reduced Weissman estimator is studied in [18]. Second, let us consider

$$H_p(k_n) = \begin{cases} \frac{1}{p} \left(1 - \left(\frac{1}{k_n} \sum_{i=1}^{k_n} \left(\frac{X_{n-i,n}}{X_{n-k_n,n}} \right)^p \right)^{-1} \right) & \text{if } p < 1/\gamma \text{ and } p \neq 0, \\ H(k_n) & \text{if } p = 0, \end{cases} \quad (12)$$

the Mean-of-order- p estimator of γ introduced almost simultaneously in [4, 6, 25], where p is some tuning parameter. A bias-reduced version of the previous estimator (12), referred to as Reduced-bias mean-of-order- p (and denoted by CH_p), is considered in [15]:

$$\text{CH}_p(k_n) = H_p(k_n) \left(1 - \frac{\hat{\beta}(1 - pH_p(k_n))}{1 - \hat{\rho} - pH_p(k_n)} \left(\frac{n}{k_n} \right)^{\hat{\rho}} \right), \quad (13)$$

following the same principle as in (11), so that $\text{CH}_0(k_n) = \text{CH}(k_n)$. An alternative bias-reduced version of (12) is proposed in [14] by replacing in the bias correction term of (13) an optimal value of p in terms of asymptotic efficiency. This gives rise to the Partially Reduced-Bias mean-of-order- p (PRB $_p$) estimator defined by

$$\text{PRB}_p(k_n) = H_p(k_n) \left(1 - \frac{\hat{\beta}(1 - \varphi_{\hat{\rho}})}{1 - \hat{\rho} - \varphi_{\hat{\rho}}} \left(\frac{n}{k_n} \right)^{\hat{\rho}} \right), \quad (14)$$

with $\varphi_{\rho} = 1 - \rho/2 - \sqrt{(1 - \rho/2)^2 - 1/2}$. Both CH_p and PRB_p estimators are plugged in the Weissman estimator to obtain extreme quantile estimators, see [16]. In the following, for the sake of simplicity, the extreme-quantile estimators derived from (11), (13) and (14) are denoted by their associated tail-index estimators, namely CH , CH_p and PRB_p . Finally, a Corrected Weissman (CW) estimator is introduced in [18] implementing two bias corrections: a first one consider in the tail-index estimator and a second one in the extrapolation factor:

$$\text{CW}(k_n) = X_{n-k_n, n} \left(\frac{k_n}{np_n} \exp \left(\hat{\beta} \left(\frac{n}{k_n} \right)^{\hat{\rho}} \frac{(k_n/(np_n))^{\hat{\rho}} - 1}{\hat{\rho}} \right) \right)^{\text{CH}(k_n)}.$$

Selection of hyperparameters. All considered extreme-quantile estimators (Weissman, RW, CH, CHP_p , PRB_p and CW) depend on the intermediate sequence k_n . The selection of k_n is a crucial point which has been widely discussed in the extreme-value literature. A standard practice is to pick out a value of k_n in the first stable part of the plot $k_n \mapsto \hat{\gamma}(k_n)$ where $\hat{\gamma}(\cdot)$ is the tail-index estimator of interest, see [19, Chapter 3]. Some attempts at formalizing this procedure can be found in [11, 27] and, more recently, in [12, 13]. Here, we adopt the sample path stability criterion described in [16, Algorithm 4.1]. Besides, CH_p and PRB_p involve an extra parameter p which also has to be selected. Two solutions are possible. First, following [16, p. 1739], one can choose the optimal value of p in terms of efficiency:

$$p^* = \varphi_{\hat{\rho}} / \text{CH}(\hat{k}_0) \quad \text{where} \quad \hat{k}_0 = \min \left(n - 1, \left\lceil \left((1 - \hat{\rho})^2 n^{-2\hat{\rho}} / (-2\hat{\rho}\hat{\beta}^2) \right)^{1/(1-2\hat{\rho})} \right\rceil + 1 \right),$$

which gives rise to two estimators CH_{p^*} and PRB_{p^*} . Second, one may select simultaneously k_n and p using the sample path stability criterion of [16, Algorithm 4.2]. The resulting estimators are still denoted by CH_p and PRB_p . To summarize, 8 extreme quantile estimators are compared in the following: Weissman, RW, CH, CHP_p , PRB_p , CHP_{p^*} , PRB_{p^*} and CW.

Results. The performance of the extreme-quantile estimators is assessed using the Relative mean-squared error (RMSE):

$$\text{RMSE}(\hat{q}_n(p_n)) = \frac{1}{N} \sum_{i=1}^N \left(\frac{\hat{q}_n(p_n)}{q_n(p_n)} - 1 \right)^2,$$

on $N = 1000$ replications of datasets with size $n = 500$. The results are provided in Table 2 ($p_n = 1/n$), Table 4 ($p_n = 1/(2n)$), for the Burr distribution and in Table 3 ($p_n = 1/n$), Table 5 ($p_n = 1/(2n)$) for Fréchet, Fisher, GPD, Inverse Gamma, and Student distributions. Let us first remark that, on the 80 considered situations, the original Weissman estimator never yields the best result, indicating that bias reduction is useful in general. At the opposite, the proposed RW estimator is the most accurate one since it provides the best result 40/80 times. The second main accurate estimator is CH_p which provides the best result 10/80 times.

Appendix: Proofs

Proof of Theorem 1. It follows the same lines as the one of [19, Theorem 4.3.8]. Let us consider the expansion

$$\frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1 \right) = \frac{\sqrt{k'_n}}{\log d_n} \left(\frac{X_{n-k_n, n} d_n^{\hat{\gamma}_n(k'_n)}}{U(1/p_n)} - 1 \right) = \frac{T_{1,n} + T_{2,n} + T_{3,n}}{T_{0,n}},$$

where we have introduced:

$$\begin{aligned} T_{0,n} &= d_n^{-\gamma} \frac{U(1/p_n)}{U(n/k_n)}, \\ T_{1,n} &= \frac{\sqrt{k'_n}}{\log d_n} \left(\frac{X_{n-k_n, n}}{U(n/k_n)} - 1 \right) d_n^{\hat{\gamma}_n(k'_n) - \gamma}, \\ T_{2,n} &= \frac{\sqrt{k'_n}}{\log d_n} \left(d_n^{\hat{\gamma}_n(k'_n) - \gamma} - 1 \right), \\ T_{3,n} &= \frac{\sqrt{k'_n}}{\log d_n} (1 - T_{0,n}). \end{aligned}$$

Let us first focus on $T_{0,n}$. From [19, Theorem 2.3.9], it follows from the second-order condition (5) that, for any $\varepsilon, \delta > 0$, there exists $t_0 > 1$ such that for all $t \geq t_0$ and $x \geq 1$,

$$\left| \frac{1}{A_0(t)} \left(\frac{U(tx)}{U(t)} - x^\gamma \right) - x^\gamma \frac{x^\rho - 1}{\rho} \right| \leq \varepsilon x^{\gamma + \rho + \delta},$$

where A_0 is asymptotically equivalent to A . Letting $x = d_n$ and $t = n/k_n$ then yields

$$\left| \frac{T_{0,n} - 1}{A_0(n/k_n)} - \frac{d_n^\rho - 1}{\rho} \right| \leq \varepsilon d_n^{\rho + \delta},$$

or equivalently,

$$T_{0,n} = 1 + A_0(n/k_n) \left(\frac{d_n^\rho - 1}{\rho} + R_n \right),$$

where $|R_n| \leq \varepsilon d_n^{\rho + \delta}$ for all $\varepsilon > 0$. Now, writing $A(t) = t^\rho \ell(t)$, where ℓ is a slowly-varying function, it follows,

$$A_0(n/k_n) \sim A(n/k_n) = A(n/k'_n) (k'_n/k_n)^\rho \frac{\ell(n/k_n)}{\ell(n/k'_n)} \sim A(n/k'_n) (k'_n/k_n)^\rho \frac{\ell(n/k_n)}{\ell((n/k_n)(\log(d_n)/c)^{1/\rho})},$$

as $n \rightarrow \infty$. From Potter bounds [5, Theorem 1.5.6(i)], for all $B > 1, \eta > 0$ and n large enough,

$$\frac{1}{B} \left(\frac{\log d_n}{c} \right)^{\eta/\rho} \leq \frac{\ell(n/k_n)}{\ell((n/k_n)(\log(d_n)/c)^{1/\rho})} \leq B \left(\frac{\log d_n}{c} \right)^{-\eta/\rho},$$

showing that

$$A_0(n/k_n) \sim A(n/k'_n) (k'_n/k_n)^\rho L(d_n),$$

where L is a slowly-varying function. As a consequence, we obtain

$$T_{0,n} = 1 + (k'_n/k_n)^\rho A(n/k'_n) \left(\frac{d_n^\rho - 1}{\rho} \right) L(d_n). \quad (15)$$

Second, under the assumption $\sqrt{k'_n}(\hat{\gamma}_n(k'_n) - \gamma) \xrightarrow{d} \mathcal{N}(\lambda\mu, \sigma^2)$ as $\sqrt{k'_n}A(n/k'_n) \rightarrow \lambda$, we get

$$d_n^{\hat{\gamma}_n(k'_n) - \gamma} = \exp\left(\frac{(\log d_n)}{\sqrt{k'_n}} \sqrt{k'_n}(\hat{\gamma}_n(k'_n) - \gamma)\right) = \exp\left(\frac{(\log d_n)}{\sqrt{k'_n}}(\lambda\mu + \sigma\xi_n)\right),$$

where $\xi_n \xrightarrow{d} \mathcal{N}(0, 1)$. Recalling that $(\log d_n)/\sqrt{k'_n} \rightarrow 0$ as $n \rightarrow \infty$, the following first order expansion holds

$$d_n^{\hat{\gamma}_n(k'_n) - \gamma} = 1 + \frac{(\log d_n)}{\sqrt{k'_n}}(\lambda\mu + \sigma\xi_n) + O_P\left(\frac{(\log d_n)^2}{k'_n}\right). \quad (16)$$

In particular, $d_n^{\hat{\gamma}_n(k'_n) - \gamma} \xrightarrow{\mathbb{P}} 1$ and therefore,

$$T_{1,n} = \frac{\sqrt{k'_n/k_n}}{\log d_n} \sqrt{k'_n} \left(\frac{X_{n-k_n,n}}{U(n/k_n)} - 1\right) (1 + o_P(1)) = \frac{\sqrt{k'_n/k_n}}{\log d_n} \gamma \xi'_n (1 + o_P(1)), \quad (17)$$

where $\xi'_n \xrightarrow{d} \mathcal{N}(0, 1)$, from [19, Theorem 2.1 and Theorem 2.2]. Third, it immediately follows from (16) that

$$T_{2,n} = \lambda\mu + \sigma\xi_n + O_P\left(\frac{\log d_n}{\sqrt{k'_n}}\right). \quad (18)$$

Finally, in view of (15) and recalling that $\sqrt{k'_n}A(n/k'_n) \rightarrow \lambda$, one has

$$T_{3,n} = \lambda(k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n}\right) L(d_n). \quad (19)$$

Collecting (15), (17), (18), (19) yields

$$\begin{aligned} \frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1\right) &= \lambda\mu + \lambda(k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n}\right) L(d_n) \\ &+ \sigma\xi_n + \frac{\sqrt{k'_n/k_n}}{\log d_n} \gamma \xi'_n (1 + o_P(1)) + O_P\left(\frac{\log d_n}{\sqrt{k'_n}}\right), \end{aligned}$$

since $T_{0,n} = 1 + o(T_{3,n})$. Besides, assumptions $(k'_n/k_n)^\rho/(\log d_n) \rightarrow c$ and $d_n \rightarrow \infty$ as $n \rightarrow \infty$ imply

$$\frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1\right) \xrightarrow{d} \mathcal{N}(\lambda(\mu + c/\rho), \sigma^2),$$

and the first part of the result is proved. If, moreover, $(\log d_n)/\sqrt{k'_n} = o(k_n^{-1/4})$, then

$$\frac{\sqrt{k'_n}}{\log d_n} \left(\frac{\hat{q}_n(p_n, k_n, k'_n)}{q(p_n)} - 1\right) = \lambda\mu + \lambda(k'_n/k_n)^\rho \left(\frac{1 - d_n^\rho}{\rho \log d_n}\right) L(d_n) + \sigma\xi_n + \frac{\sqrt{k'_n/k_n}}{\log d_n} \gamma \xi'_n (1 + o_P(1)),$$

which proves the second part of the result. ■

Proof of Lemma 1. Letting $f(x) = -\rho(\log x)/(1 - x^\rho)$ for all $x \geq 1$ and $\rho < 0$, it is easily shown that f is increasing, $f(1) = 1$ and $f(x) \sim -\rho \log x$ as $x \rightarrow \infty$. Besides, $d_n \sim k_n/\beta$ and thus $\log d_n \sim \log k_n$ as $n \rightarrow \infty$. The conclusion follows. ■

Proof of Corollary 1. It is sufficient to prove that assumptions (i) and (iii) of Theorem 1 hold true which is a consequence of Lemma 1(ii). ■

References

- [1] Fraga Alves, M.I., Gomes, M. I., and de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliae Mathematica*, 60(2), 193–214.
- [2] Beirlant, J., Dierckx, G., and Guillou, A. (2005). Estimation of the extreme-value index and generalized quantile plots. *Bernoulli*, 11(6), 949–970.
- [3] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*, Wiley.
- [4] Beran, J., Schell, D., and Stehlík, M. (2014). The harmonic moment tail index estimator: asymptotic distribution and robustness. *Annals of the Institute of Statistical Mathematics*, 66(1), 193–220.
- [5] Bingham, N. H., Goldie, C. M., and Teugels, J. L. (1989). *Regular variation*, Cambridge university press.
- [6] Brillhante, M. F., Gomes, M. I., and Pestana, D. (2013). A simple generalisation of the Hill estimator. *Computational Statistics and Data Analysis*, 57(1), 518–535.
- [7] Caeiro, F., Gomes, M. I., and Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *Revstat*, 3(2), 113–136.
- [8] Daouia, A., Girard, S., and Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli*, 26(1), 531–556.
- [9] Dekkers, A. L. M., Einmahl, J. H. J., and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, 17, 1833–1855.
- [10] Deme, E., Gardes, L., and Girard, S. (2013). On the estimation of the second order parameter for heavy-tailed distributions, *Revstat*, 11, 277–299.
- [11] Drees, H., de Haan, L., and Resnick, S. (2000). How to make a Hill plot, *Annals of Statistics*, 28, 254–274.
- [12] El Methni, J., and Stupfler, G. (2017). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica*, 27, 907–930.
- [13] El Methni, J., and Stupfler, G. (2018). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions, *Econometrics and Statistics*, 6, 129–148.
- [14] Gomes, M. I., Brillhante, M. F., Caeiro, F., and Pestana, D. (2015). A new partially reduced-bias mean-of-order p class of extreme value index estimators. *Computational Statistics and Data Analysis*, 82, 223–237.

- [15] Gomes, M. I., Brillhante, M. F., and Pestana, D. (2016). New reduced-bias estimators of a positive extreme value index. *Communications in Statistics-Simulation and Computation*, 45(3), 833–862.
- [16] Gomes, M. I., Caeiro, F., Figueiredo, F., Henriques-Rodrigues, L., and Pestana, D. (2020). Reduced-bias and partially reduced-bias mean-of-order- p value-at-risk estimation: a Monte-Carlo comparison and an application. *Journal of Statistical Computation and Simulation*, 90(10), 1735–1752.
- [17] Gomes, M. I., Caeiro, F., Figueiredo, F., Henriques-Rodrigues, L., and Pestana, D. (2020). Corrected-Hill versus partially reduced-bias value-at-risk estimation. *Communications in Statistics-Simulation and Computation*, 49(4), 867–885.
- [18] Gomes, M. I., and Pestana, D. (2007). A sturdy reduced-bias extreme quantile (VaR) estimator. *Journal of the American Statistical Association*, 102(477), 280–292.
- [19] de Haan, L., and Ferreira, A. (2007). Extreme value theory: an introduction. Springer Science and Business Media.
- [20] Hall, P. (1982). On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society: Series B*, 44(1), 37–42.
- [21] Hall, P., and Welsh, A. W. (1985) Adaptive estimates of parameters of regular variation *The Annals of Statistics*, vol. 13, 331–341
- [22] Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3, 1163–1174.
- [23] Kratz, M., and Resnick, S. I. (1996). The QQ-estimator and heavy tails. *Stochastic Models*, 12(4), 699–724.
- [24] Manjunath, B. G., and Caeiro, F. (2013). evt0: Mean of order p , peaks over random threshold Hill and high quantile estimates. *R package version 1.1-3*.
- [25] Paulauskas, V., and Vaičiulis, M. (2013). On an improvement of Hill and some other estimators. *Lithuanian Mathematical Journal*, 53(3), 336–355.
- [26] Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119–131.
- [27] Resnick, S., and Stărică, C. (1997). Smoothing the Hill estimator, *Advances in Applied Probability*, 29, 271–293.
- [28] Schultze, J., and Steinebach, J. (1996). On least squares estimates of an exponential tail coefficient. *Statistics and Risk Modeling*, 14(4), 353–372.
- [29] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364), 812–815.

Acknowledgments

The authors would like to thank Frederico Caeiro for providing the R code associated with Algorithm 4.1 and Algorithm 4.2 from [16].

Distribution (parameters)	Density function	γ	ρ
Generalised Pareto ($\sigma, \xi > 0$)	$\sigma^{-1} \left(1 + \frac{\xi}{\sigma} t\right)^{-1-1/\xi}, t > 0$	ξ	$-\xi$
Burr ($\alpha, \beta > 0$)	$\alpha\beta t^{\alpha-1} (1+t^\alpha)^{-\beta-1}, t > 0$	$1/(\alpha\beta)$	$-1/\beta$
Fréchet ($\alpha > 0$)	$\alpha t^{-\alpha-1} \exp(-t^{-\alpha}), t > 0$	$1/\alpha$	-1
Fisher ($\nu_1, \nu_2 > 0$)	$\frac{(\nu_1/\nu_2)^{\nu_1/2}}{B(\nu_1/2, \nu_2/2)} t^{\nu_1/2-1} \left(1 + \frac{\nu_1}{\nu_2} t\right)^{-(\nu_1+\nu_2)/2}, t > 0$	$2/\nu_2$	$-2/\nu_2$
Inverse Gamma ($\alpha, \beta > 0$)	$\frac{\beta^\alpha}{\Gamma(\alpha)} t^{-\alpha-1} \exp(-\beta/t), t > 0$	$1/\alpha$	$-1/\alpha$
Student ($\nu > 0$)	$\frac{1}{\sqrt{\nu} B(\nu/2, 1/2)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$	$1/\nu$	$-2/\nu$

Table 1: Examples of heavy-tailed distributions satisfying the second-order condition (5) with the associated values of γ and ρ . Here, $B(\cdot, \cdot)$ denotes the Beta function.

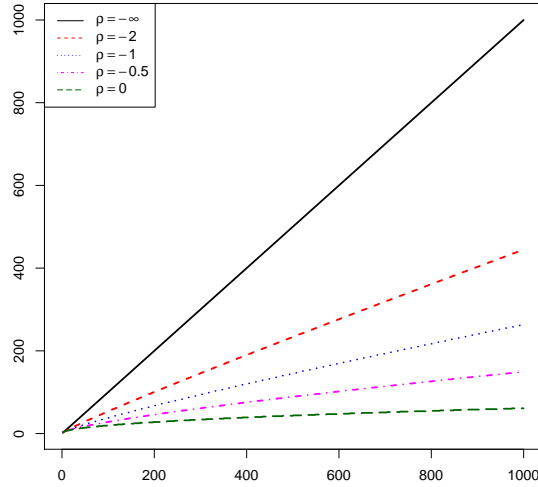


Figure 1: $k_n^{H,*}$ as a function of k_n for $\rho \in \{-\infty, -2, -1, -1/2, 0\}$ when $p_n = 1/(2n)$ i.e. $d_n = k_n/2$ and $n = 1,000$.

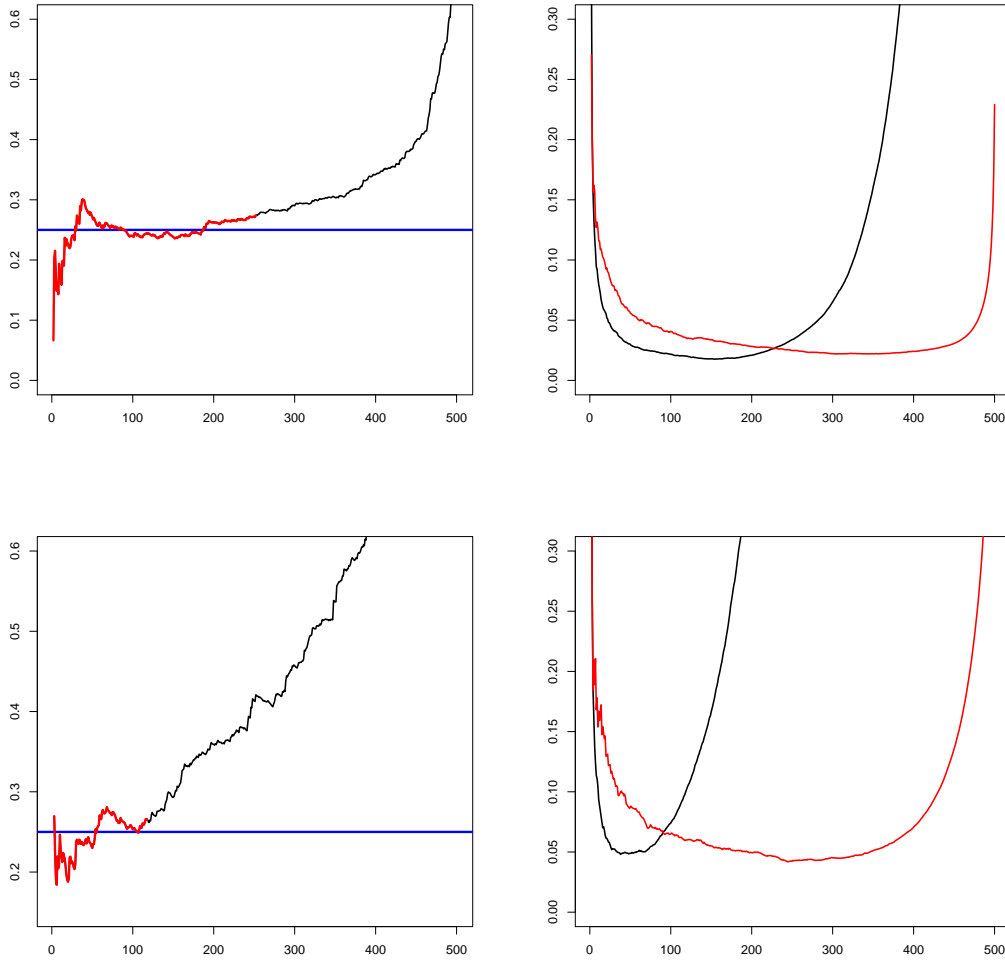


Figure 2: Illustration on a simulated dataset of size $n = 500$ from a Burr distribution with $\gamma = 1/4$. Top : $\rho = -2$ and bottom: $\rho = -3/4$. Left panel: Hill plots associated with $H(k_n)$ (black) and $H(\hat{k}_n^{H,*})$ (red) as functions of k_n . The true value of $\gamma = 1/4$ is depicted by a blue horizontal line. Right panel: RMSEs as functions of k_n associated with Weissman estimator $\hat{q}_n(p_n = 1/(2n), k_n)$ (black) and the revisited version $\hat{q}_n(p_n = 1/(2n), k_n, \hat{k}_n^{H,*})$ (red).

Burr	Weissman	RW	CW	CH	CH _p	PRB _p	CH _p *	PRB _p *
$\gamma = 1/8$								
$\rho = -1/8$	0.6658	0.0590	0.6835	0.2423	0.0178	0.8445	0.1263	0.0237
$\rho = -1/4$	0.1827	0.0224	0.1322	0.0154	0.0058	0.1354	0.0108	0.0054
$\rho = -1/2$	0.0775	0.0100	0.0138	0.0035	0.0048	0.0091	0.0039	0.0073
$\rho = -1$	0.0216	0.0044	0.0067	0.0137	0.0118	0.0148	0.0114	0.0177
$\rho = -2$	0.0420	0.0045	0.0055	0.0111	0.0101	0.0160	0.0096	0.0119
$\gamma = 1/4$								
$\rho = -1/8$	-	0.0389	-	0.6875	0.0800	0.9894	0.4508	0.1434
$\rho = -1/4$	0.7886	0.0131	0.4584	0.0811	0.0240	0.5226	0.0536	0.0242
$\rho = -1/2$	0.2265	0.0104	0.0593	0.0164	0.0239	0.0498	0.0147	0.0319
$\rho = -1$	0.0612	0.0158	0.0287	0.0443	0.0469	0.0723	0.0392	0.0558
$\rho = -2$	0.0195	0.0186	0.0212	0.0359	0.0408	0.0531	0.0322	0.0369
$\gamma = 1/2$								
$\rho = -1/8$	-	0.7514	-	-	0.7702	0.9994	-	0.5554
$\rho = -1/4$	-	0.1055	-	0.2664	0.2233	0.9985	0.2551	0.1150
$\rho = -1/2$	0.9552	0.2893	0.1850	0.0537	0.1006	0.4483	0.0513	0.1442
$\rho = -1$	0.3199	0.0625	0.1267	0.1856	0.1691	0.2741	0.1520	0.2120
$\rho = -2$	0.1448	0.0736	0.0871	0.1387	0.1381	0.1968	0.1205	0.1355
$\gamma = 1$								
$\rho = -1/8$	-	0.9883	-	-	-	-	-	0.9898
$\rho = -1/4$	-	0.9998	-	-	-	-	-	-
$\rho = -1/2$	-	-	-	0.2916	0.9493	0.9986	0.3260	0.4243
$\rho = -1$	-	0.5451	0.8959	0.9399	0.6693	-	0.8286	-
$\rho = -2$	0.7843	0.5035	0.3657	0.4943	0.5485	0.5984	0.4098	0.4552

Table 2: RMSE associated with 8 estimators of the extreme quantile $q(p_n = 1/n)$ on a Burr distribution. The best result is emphasized in red, and the second best result in blue. RMSEs larger than 1 are not reported.

	Weissman	RW	CW	CH	CH _p	PRB _p	CH _p *	PRB _p *
Fréchet ($\rho = -1$)								
$\gamma = 1/8$	0.0088	0.0052	0.0039	0.0083	0.0097	0.0118	0.0072	0.0087
$\gamma = 1/4$	0.0400	0.0165	0.0206	0.0297	0.0341	0.0388	0.0270	0.0305
$\gamma = 1/2$	0.2583	0.0699	0.0831	0.1162	0.1450	0.2047	0.1030	0.1224
$\gamma = 1$	-	0.4469	-	-	0.8632	0.9088	0.4529	0.5251
Fisher ($\rho = -\gamma$)								
$\gamma = 1/8$	0.2559	0.0624	0.7070	0.0081	0.0092	0.0745	0.0084	0.0136
$\gamma = 1/4$	0.5790	0.0584	0.1122	0.0157	0.0272	0.0727	0.0155	0.0396
$\gamma = 1/2$	0.9255	0.1234	0.1973	0.0984	0.2037	0.3109	0.0871	0.1796
$\gamma = 1$	-	0.541	-	0.6597	0.7823	-	0.5993	-
GPD ($\rho = -\gamma$)								
$\gamma = 1/8$	0.6379	0.1075	0.4635	0.1465	0.0163	0.8355	0.1079	0.0276
$\gamma = 1/4$	0.7184	0.2021	0.3716	0.0715	0.0288	0.7292	0.0693	0.0223
$\gamma = 1/2$	-	0.0276	0.2770	0.0597	0.1060	0.4886	0.0627	0.1642
$\gamma = 1$	-	0.4261	-	0.5796	0.832	0.9566	0.5130	0.6658
Inverse Gamma ($\rho = -\gamma$)								
$\gamma = 1/8$	0.1768	0.0193	0.0293	0.0070	0.0097	0.0215	0.0062	0.0131
$\gamma = 1/4$	0.2417	0.0316	0.0325	0.0196	0.0342	0.0770	0.0184	0.0289
$\gamma = 1/2$	0.6804	0.1020	0.1263	0.1396	0.1543	0.4251	0.1155	0.1582
$\gamma = 1$	-	0.6597	0.8210	0.4420	-	0.9985	0.4513	0.5888
Student ($\rho = -2\gamma$)								
$\gamma = 1/8$	0.2741	0.0607	0.2860	0.0914	0.0107	0.4264	0.0471	0.0167
$\gamma = 1/4$	0.3819	0.0991	0.1780	0.0304	0.0202	0.0526	0.0210	0.0149
$\gamma = 1/2$	0.4355	0.2920	0.0494	0.1109	0.1347	0.2383	0.0945	0.1733
$\gamma = 1$	0.7909	0.0906	0.4674	0.5170	0.6569	0.8597	0.5224	0.6633

Table 3: RMSE associated with 8 estimators of the extreme quantile $q(p_n = 1/n)$ on 5 heavy-tailed distributions. The best result is emphasized in red, and the second best result in blue. RMSEs larger than 1 are not reported.

Burr	Weissman	RW	CW	CH	CH _p	PRB _p	CH _p *	PRB _p *
$\gamma = 1/8$								
$\rho = -1/8$	-	0.0591	1.1401	0.4347	0.0595	0.8671	0.2633	0.1226
$\rho = -1/4$	0.3292	0.0158	0.2110	0.0724	0.0088	0.1241	0.0454	0.0188
$\rho = -1/2$	0.0830	0.0172	0.0287	0.0450	0.0510	0.0094	0.0480	0.0590
$\rho = -1$	0.0301	0.0083	0.0064	0.0149	0.0140	0.0188	0.0127	0.0188
$\rho = -2$	0.0096	0.0069	0.0077	0.0135	0.0119	0.0201	0.0114	0.0135
$\gamma = 1/4$								
$\rho = -1/8$	-	0.1515	-	-	0.1582	0.9921	-	0.6497
$\rho = -1/4$	-	0.1022	0.7320	0.2385	0.0365	0.5502	0.1860	0.0493
$\rho = -1/2$	0.4938	0.0252	0.1031	0.0155	0.0270	0.0495	0.0175	0.0237
$\rho = -1$	0.1103	0.0262	0.0424	0.0563	0.0523	0.0802	0.0455	0.0662
$\rho = -2$	0.0275	0.0201	0.0270	0.0459	0.0457	0.0544	0.0406	0.0451
$\gamma = 1/2$								
$\rho = -1/8$	-	0.8734	-	-	0.9123	-	-	-
$\rho = -1/4$	-	0.6167	-	0.8482	0.3754	0.9989	0.8038	0.1772
$\rho = -1/2$	-	0.1236	0.3795	0.0622	0.1188	0.4713	0.0666	0.1321
$\rho = -1$	0.6053	0.1123	0.1871	0.2122	0.1896	0.3101	0.1800	0.2326
$\rho = -2$	0.1849	0.0975	0.1057	0.1520	0.1619	0.1985	0.1395	0.1567
$\gamma = 1$								
$\rho = -1/8$	-	-	-	-	-	-	-	-
$\rho = -1/4$	-	-	-	-	-	-	-	-
$\rho = -1/2$	-	-	-	-	0.9600	-	-	-
$\rho = -1$	-	-	-	-	0.8118	-	-	-
$\rho = -2$	-	0.5719	-	-	0.5721	0.9563	-	-

Table 4: RMSE associated with 8 estimators of the extreme quantile $q(p_n = 1/(2n))$ on a Burr distribution. The best result is emphasized in red, and the second best result in blue. RMSEs larger than 1 are not reported.

	Weissman	RW	CW	CH	CH _p	PRB _p	CH _p *	PRB _p *
Fréchet ($\rho = -1$)								
$\gamma = 1/8$	0.0179	0.0081	0.0096	0.0095	0.0090	0.0157	0.0087	0.0106
$\gamma = 1/4$	0.0567	0.0209	0.0213	0.0343	0.0444	0.0529	0.0312	0.0361
$\gamma = 1/2$	0.3538	0.1081	0.0886	0.1347	0.1826	0.2294	0.1252	0.1413
$\gamma = 1$	0.9850	0.9460	0.9864	0.9877	-	-	0.9870	0.9871
Fisher ($\rho = -\gamma$)								
$\gamma = 1/8$	0.4576	0.0101	0.1336	0.0203	0.0105	0.0631	0.0206	0.0115
$\gamma = 1/4$	0.9311	0.0291	0.1901	0.0232	0.0301	0.0695	0.0273	0.0294
$\gamma = 1/2$	-	0.2334	0.2598	0.0944	0.3213	0.4579	0.0986	0.1703
$\gamma = 1$	-	0.9946	-	-	-	-	-	-
GPD ($\rho = -\gamma$)								
$\gamma = 1/8$	0.9387	0.0809	0.7988	0.3519	0.0328	0.8577	0.8301	0.1088
$\gamma = 1/4$	-	0.2491	0.6321	0.2094	0.0383	0.7582	0.1990	0.0411
$\gamma = 1/2$	-	0.0325	0.5436	0.0824	0.1202	0.5043	0.0948	0.1293
$\gamma = 1$	-	-	-	-	0.8118	-	-	-
Inverse Gamma ($\rho = -\gamma$)								
$\gamma = 1/8$	0.1821	0.0130	0.0404	0.0190	0.0156	0.0750	0.0093	0.0137
$\gamma = 1/4$	0.4147	0.0215	0.0630	0.0254	0.0367	0.0435	0.0226	0.0345
$\gamma = 1/2$	0.8202	0.1955	0.1214	0.1395	0.3718	0.5781	0.1214	0.1689
$\gamma = 1$	-	-	-	-	-	-	0.9215	0.74533
Student ($\rho = -2\gamma$)								
$\gamma = 1/8$	0.5643	0.0103	0.4662	0.2148	0.0180	0.3625	0.1251	0.0663
$\gamma = 1/4$	0.6091	0.0135	0.2795	0.0752	0.0255	0.0560	0.0502	0.0179
$\gamma = 1/2$	0.6488	0.2739	0.0777	0.1120	0.1684	0.2685	0.0983	0.1645
$\gamma = 1$	-	0.5982	-	0.6812	0.8226	-	0.7982	0.9939

Table 5: RMSE associated with 8 estimators of the extreme quantile $q(p_n = 1/(2n))$ on 5 heavy-tailed distributions. The best result is emphasized in red, and the second best result in blue. RMSEs larger than 1 are not reported.