

Weak limits for exploratory plots in the analysis of extremes

BIKRAMJIT DAS¹ and SOUVIK GHOSH²

¹*RiskLab, Department of Mathematics, ETH Zurich, Rämistrasse 101, 8092 Zurich, Switzerland.
E-mail: bikram@math.ethz.ch*

²*Yahoo! Research, 4401 Great America Parkway, Santa Clara, CA 95054, USA.
E-mail: souvik@yahoo-inc.com*

Exploratory data analysis is often used to test the goodness-of-fit of sample observations to specific target distributions. A few such graphical tools have been extensively used to detect subexponential or heavy-tailed behavior in observed data. In this paper we discuss asymptotic limit behavior of two such plotting tools: the quantile–quantile plot and the mean excess plot. The weak consistency of these plots to fixed limit sets in an appropriate topology of \mathbb{R}^2 has been shown in Das and Resnick (*Stoch. Models* **24** (2008) 103–132) and Ghosh and Resnick (*Stochastic Process. Appl.* **120** (2010) 1492–1517). In this paper we find asymptotic distributional limits for these plots when the underlying distributions have regularly varying right-tails. As an application we construct confidence bounds around the plots which enable us to statistically test whether the underlying distribution is heavy-tailed or not.

Keywords: asymptotic theory; confidence bounds; extreme values; ME plot; QQ plot; random set; regular variation

1. Introduction

Statistical analysis of extremes in available data has been very important in varied areas like finance (McNeil, Frey and Embrechts [28]), telecommunication (Maulik, Resnick and Rootzén [27], D’Auria and Resnick [10]), hydrology (Katz, Parlange and Naveau [24]), environmental statistics (Davison and Smith [11], Smith [39]) and many more. Before analyzing features of the data using extreme value analysis, it is imperative that we check whether extreme-value modeling is well suited in the given context; see Drees [17] for a recent survey of exploratory techniques for extremes in an actuarial context. Popular exploratory techniques in this direction have been the mean excess (ME) plots (Davison and Smith [11]) and the quantile–quantile (QQ) plots which are specifically tuned for heavy-tailed data (Kratz and Resnick [25]). A distribution F is *heavy-tailed* if the tail probability $(1 - F)$ is regularly varying (Resnick [34], Chapter 1). It has been shown earlier that under an assumption of heavy tails and with proper normalizations, both plots converge in probability to fixed closed sets (for ME plots, an additional assumption of finiteness of the mean of F is required); see Das and Resnick [9] and Ghosh and Resnick [22]. These results corroborate the use of the QQ plot and the ME plot to test the null hypothesis that the underlying distribution is heavy-tailed. The proximity of the observed plot to the fixed limit set would support the null hypothesis.

Incidentally, one data set leads to just one single plot of each kind. A single plot is often not enough to statistically detect proximity between the plot and the intended fixed limit set; see

the examples in Section 6. Creating appropriate confidence bounds around these plots, though, can help us to test the null hypothesis with some degree of confidence. In this paper we study weak limits of both kinds of plots for heavy-tailed data and use these limits to obtain confidence bounds around them with asymptotic coverage probabilities. The methods used here are general and can be used to find weak limits and confidence bounds for other plots used in the analysis of extremes.

1.1. Plan for this paper

We introduce the two plotting methodologies in Section 1. In Section 2 we set up necessary tools to talk about convergence of random closed sets in \mathbb{R}^2 , since the QQ and ME plots are random closed sets in \mathbb{R}^2 . In Sections 3 and 4, we prove weak convergence of the QQ and the ME plot under the null hypothesis that the underlying distribution F is heavy-tailed. We proceed by expressing both plots as appropriate functionals of the tail empirical measure and then use convergence properties of the tail empirical measure to prove weak convergence of both plots. As an application to obtaining these weak limits, we construct confidence bounds with asymptotic coverage probability for both kinds of plots in Section 5. Finally, in Section 6 we apply the results obtained in the previous sections to simulated and real data sets to exemplify how they perform in practice. We conclude in Section 7 along with a discussion on future directions.

1.2. QQ plots for heavy tails

Suppose we want to test the null hypothesis that observations from a sample are independent and identically distributed (i.i.d.) with some known distribution F . The QQ plot, which is a plot of the empirical quantiles from the data against the distributional quantiles of F , is an intuitive and popular graphical tool for detecting the goodness-of-fit for a sample to the distribution F . If the true distribution of the sample is F , then the QQ plot should converge, in an appropriate sense, to a straight line. Results involving empirical process and quantile process convergences are available in Shorack and Wellner [38], which can be appropriately used to create confidence intervals for QQ plots. The QQ plot we consider is a little different and is specifically designed to check for distributions F where $\bar{F} := 1 - F$ is regularly varying with some index $-1/\xi$, $\xi > 0$, also denoted $\bar{F} \in RV_{-1/\xi}$ (Resnick [34], Chapter 1). For a sample X_1, X_2, \dots, X_n , its decreasing order statistics are denoted by $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ and the QQ plot in this context is defined by

$$Q_n = \left\{ \left(-\log \frac{j}{k}, \log \frac{X_{(j)}}{X_{(k)}} \right) : 1 \leq j \leq k \right\}, \quad k < n.$$

Clearly we concentrate on the top k quantiles of the data justified by the fact that $\bar{F} \in RV_{-1/\xi}$ only provides us with information about the right tail of the data. Under the null hypothesis of $\bar{F} \in RV_{-1/\xi}$ for some $\xi > 0$, Das and Resnick [9] have shown convergence in probability for QQ plots in an appropriate topology of random closed sets when the data is assumed to be an i.i.d. sample.

1.3. ME plots

The ME function of a random variable X is defined as

$$M(u) := E[X - u | X > u], \quad (1.1)$$

provided $EX_+ < \infty$, and is also known as the *mean residual life function*. A natural estimate of $M(u)$ is the empirical ME function $\hat{M}(u)$ defined as

$$\hat{M}(u) = \frac{\sum_{i=1}^n (X_i - u) I_{[X_i > u]}}{\sum_{i=1}^n I_{[X_i > u]}}, \quad u \geq 0. \quad (1.2)$$

The ME plot is the plot of the points $\{(X_{(k)}, \hat{M}(X_{(k)})) : 1 < k \leq n\}$.

The ME plot is often used as a simple graphical test to check if data conform to a generalized Pareto distribution (GPD). The GPD is an important class of distributions and is fundamental for the peaks-over-threshold method used in extreme value analysis (Davison and Smith [11]). The GPD is characterized by its cumulative distribution function $G_{\xi, \beta}$

$$G_{\xi, \beta}(x) = \begin{cases} 1 - (1 + \xi x / \beta)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1 - \exp(-x/\beta) & \text{if } \xi = 0, \end{cases} \quad (1.3)$$

where $\beta > 0$, and $x \geq 0$, when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$, if $\xi < 0$. The parameters ξ and β are referred to as the *shape* and the *scale* parameter, respectively. The GPD in the case $\xi > 0$ corresponds to the classical Pareto law with tail exponent $1/\xi$. For a random variable $X \sim G_{\xi, \beta}$, we have $E(X) < \infty$, if and only if $\xi < 1$, and in this case, the ME function of X is linear in u .

$$M(u) = \frac{\beta}{1 - \xi} + \frac{\xi}{1 - \xi} u, \quad (1.4)$$

where $0 \leq u < \infty$ if $0 \leq \xi < 1$ and $0 \leq u \leq -\beta/\xi$ if $\xi < 0$. In fact, the linearity of the ME function characterizes the GPD class; cf. McNeil, Frey and Embrechts [28] and Embrechts, Klüppelberg and Mikosch [18]. Davison and Smith [11] used this property and suggested that if the ME plot is close to a straight line for high values of the threshold, then there is no evidence against the use of a GPD model. See also Embrechts, Klüppelberg and Mikosch [18] and Hogg and Klugman [23] for the implementation of this plot in practice. Ghosh and Resnick [22] discuss convergence in probability for the high thresholds of suitably normalized ME plots in an appropriate topology of random closed sets when the data is an i.i.d. sample.

The advantage of the ME plot over the QQ plot is that it works when $-\infty < \xi < 1$, whereas the QQ plot works for $\xi > 0$ only. Hence the ME plot can be used whenever the sample is in the maximal domain of attraction of any generalized extreme value distribution with finite mean (Gumbel, Weibull or Fréchet distribution). The QQ plot is restricted to the domain of attraction of Fréchet distribution only. In this paper, though, we restrict to the case when $\xi > 0$, which is the case of maximal domain of attraction of the Fréchet distribution. The disadvantage of the ME plot is that it requires $\xi < 1$ to make proper sense of the result, that is, the underlying distribution should have a finite mean. Still, limits can and have been obtained for the ME plots, even when the distributional mean is not finite; see Ghosh and Resnick [22].

2. Preliminaries

2.1. Topology on closed sets of \mathbb{R}^2

Since we are dealing with plots which are closed sets in \mathbb{R}^2 , it is imperative to understand the topology on closed sets. We denote the collection of all closed (compact) sets in \mathbb{R}^2 by \mathcal{F} (\mathcal{K} , resp.). We consider a hit and miss topology on \mathcal{F} called the Fell topology. The Fell topology is the smallest topology containing the families $\{\mathcal{F}^K, K \text{ compact}\}$ and $\{\mathcal{F}_G, G \text{ open}\}$ where, for any set B ,

$$\mathcal{F}^B = \{F \in \mathcal{F}: F \cap B = \emptyset\} \quad \text{and} \quad \mathcal{F}_B = \{F \in \mathcal{F}: F \cap B \neq \emptyset\}.$$

Hence \mathcal{F}^B and \mathcal{F}_B are collections of closed sets which miss and hit the set B , respectively. This is the reason for which such topologies are called hit and miss topologies. In the Fell topology, a sequence of closed sets $\{F_n\}$ converges to $F \in \mathcal{F}$ if and only if the following two conditions hold:

- $F \in \mathcal{F}_G$ implies there exists $N \geq 1$ such that for all $n \geq N$, $F_n \in \mathcal{F}_G$, for any open set G .
- $F \in \mathcal{F}^K$ implies there exists $N \geq 1$ such that for all $n \geq N$, $F_n \in \mathcal{F}^K$, for any compact set K .

The Fell topology on the closed sets of \mathbb{R}^2 is metrizable (Flachsmeyer [19], Beer [1]) and we indicate convergence in this topology of a sequence (F_n) of closed sets to a limit closed set F by $F_n \rightarrow F$. Often though, it is easier to deal with the following characterization of convergence.

Lemma 2.1. *A sequence $F_n \in \mathcal{F}$ converges to $F \in \mathcal{F}$ in the Fell topology if and only if the following two conditions hold:*

1. For any $t \in F$ there exists $t_n \in F_n$ such that $t_n \rightarrow t$.
2. If, for some subsequence (m_n) , $t_{m_n} \in F_{m_n}$ converges, then $\lim_{n \rightarrow \infty} t_{m_n} \in F$.

See Theorem 1-2-2 in Matheron [26], page 6, for a proof of this lemma.

Let $\sigma_{\mathcal{F}}$ denote the Borel σ -algebra generated by the Fell topology of open sets (not to be confused with open sets in \mathbb{R}^d). A random closed set $X: \Omega \mapsto \mathcal{F}$ is a measurable mapping from $(\Omega, \mathcal{A}, P')$ to $(\mathcal{F}, \sigma_{\mathcal{F}})$. Denote by P the induced probability on $\sigma_{\mathcal{F}}$, that is, $P = P' \circ X^{-1}$.

Since the Fell topology is metrizable, the definition of convergence in probability is obvious. The following result is a well-known and helpful characterization for convergence in probability of random variables, and it holds for random sets as well; see Theorem 6.21 in Molchanov [29], page 92.

Lemma 2.2. *A sequence of random sets (F_n) in \mathcal{F} converges in probability to a random set F if and only if for every subsequence (n') of \mathbb{Z}_+ there exists a further subsequence (n'') of (n') such that $F_{n''} \rightarrow F$ -a.s.*

A sequence of random closed sets $(X_n)_{n \geq 1}$ weakly converges to a random closed set X with distribution P if the corresponding induced probability measures $(P_n)_{n \geq 1}$ converge weakly to P ,

that is,

$$P_n(\mathcal{B}) = P'_n \circ X_n^{-1}(\mathcal{B}) \rightarrow P(\mathcal{B}) = P' \circ X^{-1}(\mathcal{B}), \quad \text{as } n \rightarrow \infty$$

for each $\mathcal{B} \in \sigma_{\mathcal{F}}$ such that $P(\partial\mathcal{B}) = 0$. This is not always straightforward to verify from the definition. The following characterization of weak convergence in terms of sup-measures is very useful; cf. Vervaat [42]. Suppose $h : \mathbb{R}^d \rightarrow \mathbb{R}_+ = [0, \infty)$. For $X \subset \mathbb{R}^d$, define $h(X) = \{h(x) : x \in X\}$, and h^\vee is the sup-measure generated by h defined by

$$h^\vee(X) = \sup\{h(x) : x \in X\}$$

(Molchanov [29], Vervaat [42]). These definitions permit the following characterization (Molchanov [29], page 87).

Lemma 2.3. *A sequence $(X_n)_{n \geq 1}$ of random closed sets converges weakly to a random closed set X if and only if $\mathbb{E}h^\vee(X_n)$ converges to $\mathbb{E}h^\vee(X)$ for every non-negative continuous function $h : \mathbb{R}^d \rightarrow \mathbb{R}$ with a bounded support.*

We often use the following notation: for a $x \in \mathbb{R}$ and a set $A \subset \mathbb{R}^n$, $xA = \{xy : y \in A\}$ and $x + A = \{x + y : y \in A\}$. See Matheron [26] and Molchanov [29] for further details on the theory of random sets.

2.2. Miscellany

Throughout this paper we will take $k := k_n$ to be a sequence increasing to infinity such that $k_n/n \rightarrow 0$. For a distribution function $F(x)$ we write $\bar{F}(x) := 1 - F(x)$ for the tail, and the quantile function is

$$b(u) := F^{\leftarrow} \left(1 - \frac{1}{u} \right) = \inf \left\{ s : F(s) \geq 1 - \frac{1}{u} \right\} = \left(\frac{1}{1 - F} \right)^{\leftarrow} (u).$$

A function $U : (0, \infty) \rightarrow \mathbb{R}_+$ is regularly varying with index $\rho \in \mathbb{R}$, written $U \in RV_\rho$, if

$$\lim_{t \rightarrow \infty} \frac{U(tx)}{U(t)} = x^\rho, \quad x > 0.$$

Regular variation is discussed in several books such as Resnick [33,34], Seneta [37], Geluk and de Haan [20], de Haan [12], de Haan and Ferreira [13], Bingham, Goldie and Teugels [3].

We use $M_+(0, \infty]$ to denote the space of non-negative Radon measures μ on $(0, \infty]$ metrized by the vague metric. Point measures are written as a function of their points $\{x_i, i = 1, \dots, n\}$ by $\sum_{i=1}^n \delta_{x_i}$; see, for example, Resnick [34], Chapter 3.

We will use the following notations to denote different classes of functions: For $0 \leq a < b \leq \infty$,

1. $\mathbb{C}[a, b)$: Continuous functions on $[a, b)$.
2. $\mathbb{D}[a, b)$: Right-continuous functions with finite left limits defined on $[a, b)$.

3. $\mathbb{D}_l[a, b)$: Left-continuous functions with finite right limits defined on $[a, b)$.

$\mathbb{D}[0, 1]$ is complete and separable under a metric $d_0(\cdot)$, which is equivalent to the Skorohod metric $d_S(\cdot)$ (Billingsley [2], page 128), but not under the uniform metric $\|\cdot\|$. As we will see, the limit processes that appear in our analysis below are always continuous. We can check that if x is continuous (in fact, uniformly continuous) in $[0, 1]$, for $x_n \in \mathbb{D}[0, 1]$, $\|x_n - x\| \rightarrow 0$ is equivalent to $d_S(x_n, x) \rightarrow 0$ and hence equivalent to $d_0(x_n, x) \rightarrow 0$ as $n \rightarrow \infty$ (Billingsley [2], page 124). So we use convergence in uniform metric, for our convenience henceforth. For spaces of the form $\mathbb{D}[a, b)$ or $\mathbb{D}_l[a, b)$, we will consider the topology of locally uniform convergence. In some cases we will also consider product spaces of functions, and then the topology will be the product topology. For example, $\mathbb{D}_l^2[1, \infty)$ will denote the class of 2-dimensional functions on $[1, \infty)$ which are left-continuous with right limit. The classes of functions defined on the sets $[a, b]$ or $(a, b]$ will have the obvious notation.

2.3. A useful lemma

The following lemma will be used often in the proofs below. We use “ \Rightarrow ” to denote weak convergence.

Lemma 2.4. *Let $Y_n \in \mathbb{D}_l^2(0, 1]$ be a sequence of random functions and assume the following hold:*

- (i) $Y_n \Rightarrow Y$, where $Y(t)$ has continuous paths with probability 1.
- (ii) *There exists a partition $0 = t_n^{(0)} < t_n^{(1)} < \dots < t_n^{(m_n)} = 1$ such that $Y_n(t)$ is constant on the interval $(t_n^{(i)}, t_n^{(i+1)}]$ for all $0 \leq i < m_n$ with probability 1.*

Then for any $0 < \varepsilon < 1$,

$$\mathcal{Y}_n^\varepsilon := \{Y_n(t_n^{(i)}): 0 < i \leq m_n, t_n^{(i)} \geq \varepsilon\} \Rightarrow \mathcal{Y}^\varepsilon := \{Y(t): \varepsilon \leq t \leq 1\} \quad \text{in } \mathcal{F}. \quad (2.1)$$

Furthermore, if $\lim_{t \downarrow 0, n \rightarrow \infty} |Y_n(t)| = \infty$ with probability 1, then

$$\mathcal{Y}_n := \{Y_n(t_n^{(i)}): 0 < i \leq m_n\} \Rightarrow \mathcal{Y} := \{Y(t): 0 < t \leq 1\} \quad \text{in } \mathcal{F}. \quad (2.2)$$

Proof. Using Lemma 2.3 it suffices to show that

$$\lim_{n \rightarrow \infty} E[h^\vee(\mathcal{Y}_n)] = E[h^\vee(\mathcal{Y})]$$

for any continuous function $h: \mathbb{R}^2 \mapsto \mathbb{R}_+$ with a compact support. So take any such function h . By the Skorohod representation theorem (Billingsley [2], Theorem 6.7, page 70), there exists a probability space (Ω, \mathcal{G}, P) and random elements Y_n^* and Y^* in $\mathbb{D}_l^2(0, 1]$ such that

$$Y_n^* \stackrel{d}{=} Y_n \quad \text{and} \quad Y^* \stackrel{d}{=} Y$$

in the sense of finite dimensional distributions (f.d.d.) and

$$Y_n^* \rightarrow Y^*, \quad P\text{-a.s. in } \mathbb{D}_l^2(0, 1].$$

Now observe that

$$h^\vee(\mathcal{Y}^\varepsilon) = \sup_{x \in \mathcal{Y}^\varepsilon} h(x) \stackrel{d}{=} \sup_{\varepsilon \leq t \leq 1} h(Y^*(t)) \quad \text{and} \quad h^\vee(\mathcal{Y}_n^\varepsilon) = \sup_{x \in \mathcal{Y}_n^\varepsilon} h(x) \stackrel{d}{=} \sup_{\varepsilon \leq t \leq 1} h(Y_n^*(t)).$$

Since $Y^*(t)$ is continuous, we know that $\sup_{\varepsilon \leq t \leq 1} |Y_n^*(t) - Y^*(t)| \rightarrow 0$. Moreover, since h is continuous with a compact support, we get h is uniformly continuous, and hence

$$\sup_{\varepsilon \leq t \leq 1} h(Y_n^*(t)) \rightarrow \sup_{\varepsilon \leq t \leq 1} h(Y^*(t)), \quad P\text{-a.s.} \tag{2.3}$$

As h is bounded, applying the dominated convergence theorem, we get

$$E[h^\vee(\mathcal{Y}_n^\varepsilon)] = E\left[\sup_{\varepsilon \leq t \leq 1} h(Y_n^*(t))\right] \rightarrow E\left[\sup_{\varepsilon \leq t \leq 1} h(Y^*(t))\right] = E[h^\vee(\mathcal{Y}^\varepsilon)],$$

and this proves (2.1).

Since $h : \mathbb{R}^2 \mapsto \mathbb{R}_+$ has a bounded support, we can find $M > 0$ such that $h(x) = 0$ whenever $|x| > M$. If $\lim_{t \downarrow 0, n \rightarrow \infty} |Y_n^*(t)| = \infty$ with probability 1, then almost surely for any $\omega \in \Omega$ we can find $\delta > 0$ and $N \geq 1$ such that $|Y_n^*(t)(\omega)| > M$ for all $\delta \leq t \leq 1$ and $n \geq N$. This implies (using (2.3))

$$\sup_{0 < t \leq 1} h(Y_n^*(t)(\omega)) = \sup_{\varepsilon \leq t \leq 1} h(Y_n^*(t)(\omega)) \rightarrow \sup_{\varepsilon \leq t \leq 1} h(Y^*(t)(\omega)) = \sup_{0 < t \leq 1} h(Y^*(t)(\omega)).$$

The remaining part of the proof of (2.2) can be completed using the same argument used to prove (2.1). □

3. Limit results for the QQ plots

Convergence of empirical processes and quantile processes to functionals of Gaussian processes, usually Brownian motion and Brownian bridges, are quite well known; cf. Shorack and Wellner [38]. We prove similar results for extreme order statistics. We use the weak limit of tail empirical measure and deduce weak convergence of the logarithmic version of the QQ plot of the extreme order statistics as a random set.

The following was proved in Das and Resnick [9]:

Proposition 3.1. *Suppose X_1, \dots, X_n are i.i.d. with common distribution F , and $X_{(1)} \geq X_{(2)} \geq \dots \geq X_{(n)}$ are the order statistics from this sample. If F is strictly increasing and continuous on its support, then*

$$\mathcal{T}_n := \left\{ \left(F^{\leftarrow} \left(\frac{i}{n+1} \right), X_{(n-i+1)} \right) : 1 \leq i \leq n \right\} \xrightarrow{P} \mathcal{T} := \{(x, x) : x \in \text{support}(F)\}$$

in \mathcal{F} .

This proposition though is not enough if one is interested in creating confidence bounds from the data. For that purpose one would need weak convergence results which are widely known in terms of convergence of affine transformations of quantile processes to appropriate Brownian Bridges for a known distribution F ; see Shorack and Wellner [38], Chapter 3, for further details. In the following section, we concentrate on the case where \bar{F} is regularly varying with tail index $-1/\xi$ with $\xi > 0$. The specific form of F is otherwise unknown.

3.1. QQ plots for distributions with regularly varying tails

Now assume that X_1, X_2, \dots, X_n are i.i.d. from a distribution F . Suppose we want to check whether F is heavy-tailed or not. In the sense of testing a hypothesis, our null hypothesis is that $\bar{F} \in RV_{-1/\xi}$ for some $\xi > 0$. Note that we really do not have any specific form for F . We define the following sets:

$$\mathcal{Q}_n = \left\{ \left(-\log \frac{j}{k}, \log \frac{X_{(j)}}{X_{(k)}} \right) : 1 \leq j \leq k \right\}, \quad k < n, \tag{3.1}$$

$$\mathcal{Q} = \{(x, \xi x) : x \geq 0\}. \tag{3.2}$$

The set \mathcal{Q}_n is the logarithmic version of the QQ plot for the first k order statistics from the sample X_1, \dots, X_n .

Das and Resnick [9] proved that under the null hypothesis, $\mathcal{Q}_n \xrightarrow{P} \mathcal{Q}$ in \mathcal{F} as $k, n \rightarrow \infty$ with $k/n \rightarrow 0$. We show below that a distributional convergence can also be obtained in this case.

Assumption 3.2. F satisfies

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right) = 0 \tag{3.3}$$

locally uniformly on $(0, \infty]$ as $k, n, n/k \rightarrow \infty$.

Theorem 3.3. Suppose X_1, \dots, X_n are i.i.d. observations from a distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $\xi > 0$ and Assumption 3.2. Then as $n, k, n/k \rightarrow \infty$

$$\begin{aligned} \mathcal{QN}_n &:= \left\{ \left(-\log \frac{j}{k}, -\xi \log \frac{j}{k} + \sqrt{k} \left(\log \frac{X_{(j)}}{X_{(k)}} + \xi \log \frac{j}{k} \right) \right) : 1 \leq j \leq k \right\} \\ \Rightarrow \mathcal{QN} &:= \left\{ (-\log t, -\xi \log t + \xi t^{-1} B(t)) : 0 < t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned} \tag{3.4}$$

where $B(t)$ is a Brownian Bridge on $[0, 1]$ restricted to $(0, 1]$.

Remark 3.4. The set \mathcal{QN}_n is a suitably normalized version of the QQ plot which allows us to obtain a weak limit. It is important to observe that the format in which we have expressed the result is not standard in the literature as far as weak limits of random variables or functions are concerned. Usual weak limit results will only consider the normalized difference of the random

variable from its mean or its limit in probability. In our setting it is imperative to state the result in the form which we have used. We look at the plot as the probability limit perturbed by the normalized deviation around it; that is, we shift the normalized differences so that we can obtain the distribution of the deviation of the observed points of the QQ plot from its mean position. If we do not make this shift, the weak limit will always hover around the y -axis and will not give the deviation from the actual point in the plot.

Remark 3.5. We have used Assumption 3.2 in order to prove a weak limit for the QQ plots. Without this assumption we can show the convergence of tail empirical measure with unknown centering $\frac{n}{k} \bar{F}(b(n/k)y^{-\xi})$ as in (3.7), but we wish the centering to be y here. To achieve this

$$\lim_{n \rightarrow \infty} \sqrt{k} \left(\frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) - y \right)$$

should exist and have a finite value which we assume to be 0 without loss of any generality. The same theorem can be proved by replacing Assumption 3.2 with the stronger condition of *second order regular variation*; see de Haan and Ferreira [13], de Haan and Stadtmueller [15], de Haan and Peng [14]. Neither Assumption 3.2 nor the second order RV condition is easy to check in practice, albeit we resort to assuming them in order to obtain distributional limits.

Proof of Theorem 3.3. The tail empirical measure defined as

$$v_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \varepsilon_{X_i/b(n/k)}(\cdot) \tag{3.5}$$

is a random element of $M_+(0, \infty]$ where $\varepsilon_x(\cdot)$ puts unit mass at x . By Theorem 4.1 (Resnick [33], page 79), we get that

$$v_n \Rightarrow v \quad \text{in } M_+(0, \infty], \tag{3.6}$$

where $v(y, \infty] = y^{-1/\xi}$, $y > 0$. Furthermore, Theorem 9.1 in Resnick [33], page 292, gives us

$$\sqrt{k} \left(v_n(y^{-\xi}, \infty] - \frac{n}{k} \bar{F}(b(n/k)y^{-\xi}) \right) \Rightarrow W(y) \quad \text{in } \mathbb{D}_l(0, \infty], \tag{3.7}$$

where W is a standard Brownian motion on $[0, \infty)$. Since F satisfies Assumption 3.2, we obtain

$$\sqrt{k} (v_n(y^{-\xi}, \infty] - y) \Rightarrow W(y) \quad \text{in } \mathbb{D}_l(0, \infty]. \tag{3.8}$$

We will use this to find the limiting distribution of

$$\sqrt{k} \left(\log \frac{X_{(\lceil kt \rceil)}}{X_{(k)}} + \xi \log t \right) = \sqrt{k} \log \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} t^\xi \right), \quad 0 < t \leq 1,$$

where for any $z \in \mathbb{R}$, denote by $\lceil z \rceil$, the largest integer less than or equal to z . For $0 < t \leq 1$, let

$$v_n^{\leftarrow}(t) := \inf\{y: v_n(y^{-\xi}, \infty] \geq t\} = \inf \left\{ y: \sum_{i=1}^n \varepsilon_{X_i/b(n/k)}(y^{-\xi}, \infty] \geq kt \right\} = \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi}.$$

Note that we can apply Vervaat's lemma (Resnick [33], Proposition 3.3, page 59) to (3.8) to get

$$\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right) \Rightarrow W(t) \quad \text{in } \mathbb{D}_l(0, 1]. \quad (3.9)$$

Therefore, using the continuous map $f: \mathbb{D}_l(0, 1] \rightarrow \mathbb{D}_l(0, 1]$ with $f(x)(t) = x(t)/t$, we have

$$\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) \Rightarrow \frac{W(t)}{t} \quad \text{in } \mathbb{D}_l(0, 1]. \quad (3.10)$$

Also observe that

$$\begin{aligned} \sqrt{k} \log \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right) &= -\sqrt{k} \xi \log \left[1 - \left(1 - \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right)^{-1/\xi} \right) \right] \\ &= -\sqrt{k} \xi \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) \\ &\quad + o_P \left(\sqrt{k} \xi \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right)^{-1/\xi} - 1 \right) \right). \end{aligned} \quad (3.11)$$

So from (3.10) and (3.11) it follows that

$$\sqrt{k} \log \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \right) \Rightarrow -\xi \frac{W(t)}{t} \quad \text{in } \mathbb{D}_l(0, 1]. \quad (3.12)$$

We again use the continuous mapping theorem with $f: \mathbb{D}_l(0, 1] \rightarrow \mathbb{D}_l(0, 1]$, defined as $f(x)(t) = x(t) - x(1)$, to get the following:

$$\begin{aligned} \sqrt{k} \log \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} t^\xi \right) &= -\sqrt{k} \log \frac{X_{(k)}}{b(n/k)} + \sqrt{k} \log \frac{X_{(\lceil kt \rceil)}}{b(n/k)} t^\xi \\ &\Rightarrow \xi W(1) - \xi \frac{W(t)}{t} \quad \text{in } \mathbb{D}_l(0, 1]. \end{aligned} \quad (3.13)$$

We know that $tW(1) - W(t) \stackrel{d}{=} B(t)$ on $\mathbb{D}_l[0, 1]$, where “ $\stackrel{d}{=}$ ” denotes equality in distribution, and B is a Brownian Bridge on $[0, 1]$. Therefore, it is true on a restriction, and hence

$$\sqrt{k} \left(\log \frac{X_{(\lceil kt \rceil)}}{X_{(k)}} + \xi \log t \right) \Rightarrow \xi t^{-1} B(t) \quad \text{in } \mathbb{D}_l(0, 1].$$

Furthermore, we also get

$$\begin{aligned} S_n(t) &= \left(-\log \frac{\lceil kt \rceil}{k}, -\xi \log \frac{\lceil kt \rceil}{k} + \sqrt{k} \left(\log \frac{X_{(\lceil kt \rceil)}}{X_{(k)}} + \xi \log \frac{\lceil kt \rceil}{k} \right) \right) \\ &\Rightarrow S(t) = \left(-\log t, -\xi \log t + \xi \frac{B(t)}{t} \right) \quad \text{in } \mathbb{D}_l^2(0, 1], \end{aligned} \quad (3.14)$$

using the converging-together lemma (Resnick [33], Proposition 3.1, page 57) and the fact that

$$\sqrt{k} \left(\log \frac{[kt]}{k} - \log t \right) \rightarrow 0$$

locally uniformly on $(0, 1]$. The weak convergence of the set \mathcal{QN}_n follows from Lemma 2.4 once we note that S_n and S in (3.14) satisfy the conditions of Lemma 2.4. \square

4. Limit results for the ME Plots

4.1. Empirical ME function for known distribution F

Suppose X_1, \dots, X_n is an i.i.d. sample from distribution F . Yang [43] studied the properties of the empirical ME function $\hat{M}(u)$ in (1.2) as an estimator of $M(u)$. They showed that $\hat{M}(u)$ is uniformly strongly consistent for $M(u)$: for any $0 < b < \infty$,

$$P \left[\lim_{n \rightarrow \infty} \sup_{0 \leq u \leq b} |\hat{M}(u) - M(u)| = 0 \right] = 1.$$

Yang [43] also proved a weak limit for $\hat{M}(u)$: for any $0 < b < 1$,

$$\sqrt{n} (\hat{M}(F^{\leftarrow}(t)) - M(F^{\leftarrow}(t))) \Rightarrow U(t),$$

where $U(t)$ is a Gaussian process on $[0, b]$ with covariance function

$$\Gamma(s, t) = \frac{(1-t)\sigma^2(t) - t\theta^2(t)}{(1-s)(1-t)^2} \quad \text{for all } 0 \leq s \leq t \leq b$$

with

$$\sigma^2(t) = \text{var}(XI_{[t < F(X) \leq 1]}) \quad \text{and} \quad \theta(t) = E(XI_{[t < F(X) \leq 1]}).$$

Although these properties are stated for the empirical ME function, using Lemma 2.4, it can be shown that the ME plots also exhibit the same features when the distribution F is known.

4.2. ME plot in the regularly varying case

The behavior of $\hat{M}(u)$ near the right end-point of F is not explained in Yang [43]. Here we study the asymptotic properties of the ME plot when the explicit form of the distribution F is not known. Ghosh and Resnick [22] proved the limit in probability of a suitably scaled version of the ME plot under the following null hypothesis:

Theorem 4.1. *If X_1, \dots, X_n are i.i.d. observations with distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $0 < \xi < 1$, then in \mathcal{F} ,*

$$\mathcal{M}_n := \frac{1}{X_{(k)}} \{ (X_{(i)}, \hat{M}(X_{(i)})) : i = 2, \dots, k \} \xrightarrow{P} \mathcal{M} := \left\{ \left(t, \frac{\xi}{1-\xi} t \right) : t \geq 1 \right\}. \quad (4.1)$$

In this paper we obtain the weak limit of the ME plot when the null hypothesis that $\bar{F} \in RV_{-1/\xi}$ for some $\xi > 0$ holds. The limit distribution depends on the value of ξ . We get different limits depending on whether $\xi \leq 1/2$, $1/2 < \xi < 1$ or $\xi \geq 1$.

4.2.1. Case I: $0 < \xi < 1/2$

In this case $\text{var}(X_1) < \infty$ exists and we obtain a Gaussian limit for the suitably normalized ME plots. The following assumption is essential. It is stronger than Assumption 3.2 which was required to obtain the weak limit of the QQ plot. As we discussed in Remark 3.5, it is quite difficult to check this assumption in practice.

Assumption 4.2. *F satisfies Assumption 3.2, and, moreover,*

$$\sqrt{k} \int_1^\infty \left| \frac{n}{k} \bar{F}(b(n/k)y) - y^{-1/\xi} \right| dy \rightarrow 0$$

as $n, k, n/k \rightarrow \infty$.

Theorem 4.3. *Suppose X_1, \dots, X_n are i.i.d. observations from a distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $0 < \xi < 1/2$ and Assumption 4.2 holds. Then for any $0 < \varepsilon < 1$, as $n, k, n/k \rightarrow \infty$,*

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi}, \frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) : i = \lceil \varepsilon k \rceil, \dots, k \right\} \\ \Rightarrow \mathcal{MN} &:= \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1-\xi} t^{-\xi} + \xi t^{-1} \int_0^t y^{-(1+\xi)} B(y) dy \right), \right. \\ &\quad \left. \varepsilon \leq t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned}$$

where $B(t)$ is the standard Brownian bridge on $[0, 1]$ restricted to $(0, 1]$.

Remark 4.4. Similar to Theorem 3.3 we look at the ME plot as the probability limit perturbed by the normalized deviation around it and obtain a weak limit in Theorem 4.3. The assumption that $\xi < 1/2$ is essential. Note that

$$\int_0^t y^{-(1+\xi)} W(y) dy = \int_{t^{-\xi}}^\infty W(u^{-1/\xi}) du = \int_{t^{-\xi}}^\infty \int_0^{y^{-1/\xi}} dW(s) dy = \int_0^t s^{-\xi} dW(s),$$

and it is well known that the integral on the right-hand side exists if and only if $\int_0^t s^{-2\xi} ds < \infty$, for which it is necessary and sufficient to have $\xi < 1/2$; cf. Øksendal [30], Lemma 3.1.5, page 26. This means

$$\int_0^t y^{-(1+\xi)} B(y) dy \stackrel{d}{=} \int_0^t y^{-(1+\xi)} W(y) dy - W(1) \int_0^t y^{-\xi} dy$$

exists if and only if $\xi < 1/2$, and the same is true for the limit \mathcal{MN} .

Proof of Theorem 4.3. Consider a functional form of the ME plot,

$$S_n(t) = (S_n^{(1)}(t), S_n^{(2)}(t)) := \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}}, \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} \right), \quad t \in (0, 1], \tag{4.2}$$

as random elements in $\mathbb{D}_1^2(0, 1]$. Following the proof of Theorem 3.2 in Ghosh and Resnick [22], we know that $S_n(\cdot) \xrightarrow{P} S(\cdot)$ in $\mathbb{D}_1^2(0, 1]$, where

$$S(t) := (S^{(1)}(t), S^{(2)}(t)) = \left(t^{-\xi}, \frac{\xi}{1-\xi} t^{-\xi} \right), \quad t \in (0, 1].$$

Applying Vervaat’s lemma (Resnick [33], Proposition 3.3, page 59) to (3.8), we get

$$\begin{aligned} & \left(\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right), \sqrt{k}(v_n(t^{-\xi}, \infty] - t) \right) \\ & \Rightarrow (-W(t), W(t)) \quad \text{in } \mathbb{D}_1^2(0, \infty). \end{aligned} \tag{4.3}$$

Observe that

$$S_n^{(2)}(t) := \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} = \frac{k}{\lceil kt \rceil - 1} \int_{X_{(\lceil kt \rceil)}/b(n/k)}^{\infty} \hat{v}_n(x, \infty] dx,$$

where

$$\hat{v}_n(\cdot) := \frac{1}{k} \sum_{i=1}^n \varepsilon_{X_i/X_{(k)}}(\cdot). \tag{4.4}$$

Using (4.3) and the converging-together lemma (Resnick [33], Proposition 3.1, page 57), we also have

$$\begin{aligned} & \left(\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right), \sqrt{k}(v_n(t^{-\xi}, \infty] - t), \frac{X_{(k)}}{b(n/k)} \right) \\ & \Rightarrow (-W(t), W(t), 1) \quad \text{in } \mathbb{D}_1^2(0, \infty) \times (0, \infty). \end{aligned} \tag{4.5}$$

Define a map $\hat{T} : \mathbb{D}_1^2(0, \infty) \times (0, \infty) \rightarrow \mathbb{D}_1(0, \infty) \times \mathbb{D}[1, \infty)$ as $\hat{T}(f, g, x)(t, y) = (f(t), g(y^{-1/\xi}x) + y^{-1/\xi}f(1))$. We can check that \hat{T} is continuous at any $(f_0, g_0, x_0) \in \mathbb{C}^2(0, \infty] \times (0, \infty)$ (see Resnick [33], page 83). Hence, by the continuous mapping theorem on (4.5), with

the map \hat{T} , we get that in $\mathbb{D}_l(0, \infty) \times \mathbb{D}[1, \infty)$

$$\begin{aligned}
 J_n(t, y) &:= \left(\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right), \sqrt{k} (\hat{v}_n(y, \infty) - y^{-1/\xi}) \right) \\
 &= \left(\sqrt{k} \left(\left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} \right)^{-1/\xi} - t \right), \right. \\
 &\quad \left. \sqrt{k} \left(\hat{v}_n \left(y \frac{X_{(k)}}{b(n/k)}, \infty \right) - \left(y \frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} \right) \right. \\
 &\quad \left. + y^{-1/\xi} \sqrt{k} \left(\left(\frac{X_{(k)}}{b(n/k)} \right)^{-1/\xi} - 1 \right) \right) \\
 &\Rightarrow (-W(t), W(y^{-1/\xi}) - y^{-1/\xi} W(1)).
 \end{aligned} \tag{4.6}$$

By an application of the functional delta method (van der Vaart and Wellner [41], Theorem 3.9.4) to (4.6), we obtain in $\mathbb{D}^* := \mathbb{D}_l(0, 1] \times \mathbb{D}[1, \infty)$,

$$\begin{aligned}
 J_n^*(t, y) &:= \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} - t^{-\xi} \right), \sqrt{k} (\hat{v}_n(y, \infty) - y^{-1/\xi}) \right) \\
 &\Rightarrow (\xi t^{-(1+\xi)} W(t), W(y^{-1/\xi}) - y^{-1/\xi} W(1)).
 \end{aligned} \tag{4.7}$$

The map $\phi : \mathbb{D}^* \rightarrow \mathbb{D}^*$ given by $\phi(f, g) = (f^{-\xi}, g)$ is Hadamard differentiable at $(f(t), g(y)) = (t, y^{-1/\xi})$, tangentially to $\mathbb{D}_0 := \mathbb{C}(0, 1] \times \mathbb{C}[1, \infty) \subset \mathbb{D}^*$, with the right-hand side of (4.6) being separable and an element of \mathbb{D}_0 . Thus we can apply the functional delta method (van der Vaart and Wellner [41], Theorem 3.9.4) to obtain (4.7).

Consequently, in $\mathbb{D}_l(0, 1] \times \mathbb{D}[1, \infty)$,

$$\begin{aligned}
 H_n(t, y) &:= \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \sqrt{k} (\hat{v}_n(y, \infty) - y^{-1/\xi}) \right) \\
 &= \left(\frac{b(n/k)}{X_{(k)}} \sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{b(n/k)} - t^{-\xi} \right) - \frac{b(n/k)}{X_{(k)}} \sqrt{k} \left(\frac{X_{(k)}}{b(n/k)} - 1 \right) t^{-\xi}, \right. \\
 &\quad \left. \sqrt{k} (\hat{v}_n(y, \infty) - y^{-1/\xi}) \right) \\
 &\Rightarrow (\xi t^{-(1+\xi)} W(t) - \xi t^{-\xi} W(1), W(y^{-1/\xi}) - y^{-1/\xi} W(1)) \\
 &\stackrel{d}{=} (\xi t^{-(1+\xi)} B(t), B(y^{-1/\xi})) =: H(t, y).
 \end{aligned} \tag{4.8}$$

Define, for some $1 \leq K < \infty$, the maps T and T_K from $\mathbb{D}_l(0, 1] \times \mathbb{D}[1, \infty)$ to $\mathbb{D}_l(0, 1] \times \mathbb{D}[1, \infty)$ by

$$T(f, g)(t, y) = \left(f(t), \int_y^\infty g(x) dx \right) \quad \text{and} \quad T_K(f, g)(t, y) = \left(f(t), \int_y^{K \vee y} g(x) dx \right). \tag{4.9}$$

We understand $\int_y^\infty g(x) dx = \infty$ if g is not integrable. Note that, in the Skorohod metric d_S , we get $d_S(T_K(f_n, g_n), T_K(f, g)) \leq d_S(f_n, f) + Kd_S(g_n, g) \rightarrow 0$ where $\{f_n, n \geq 1\}, f \in \mathbb{D}_l(0, 1]$ and $\{g_n, n \geq 1\}, g \in \mathbb{D}_l[1, \infty)$ with $d_S(f_n, f) \rightarrow 0$ and $d_S(g_n, g) \rightarrow 0$ as $n \rightarrow \infty$. So T_K is a continuous mapping. By (4.8) and the continuity of the map T_K , we get that $T_K(H_n) \Rightarrow T_K(H)$. We also claim that, for any $\varepsilon > 0$,

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P[\|T_K(H_n) - T(H_n)\| > \varepsilon] = 0. \tag{4.10}$$

Note that, for any $\varepsilon > 0$,

$$\begin{aligned} & \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P[\|T_K(H_n) - T(H_n)\| > \varepsilon] \\ & \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_K^\infty (\hat{v}_n(x, \infty) - x^{-1/\xi}) dx \right| > \varepsilon\right] \\ & \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_K^\infty \left(v_n\left(x \frac{X(k)}{b(n/k)}, \infty\right) - \left(x \frac{X(k)}{b(n/k)}\right)^{-1/\xi} \right) dx \right| > \varepsilon/2\right] \\ & \quad + \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_K^\infty x^{-1/\xi} \left(\left(\frac{X(k)}{b(n/k)}\right)^{-1/\xi} - 1 \right) dx \right| > \varepsilon/2\right]. \end{aligned}$$

Using (3.9) and the assumption that $\xi < 1/2$, we get

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_K^\infty x^{-1/\xi} \left(\left(\frac{X(k)}{b(n/k)}\right)^{-1/\xi} - 1 \right) dx \right| > \varepsilon/2\right] = 0.$$

Using a change of variable, we obtain

$$\begin{aligned} & \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_K^\infty \left(v_n\left(x \frac{X(k)}{b(n/k)}, \infty\right) - \left(x \frac{X(k)}{b(n/k)}\right)^{-1/\xi} \right) dx \right| > \varepsilon/2\right] \\ & = \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_{K X(k)/b(n/k)}^\infty (v_n(u, \infty) - u^{-1/\xi}) \frac{b(n/k)}{X(k)} du \right| > \varepsilon/2\right]. \end{aligned}$$

Now fix any $\eta > 0$, and note that

$$\lim_{n \rightarrow \infty} P\left[\left| \frac{X(k)}{b(n/k)} - 1 \right| > \eta \text{ or } \left| \frac{b(n/k)}{X(k)} - 1 \right| > \eta\right] = 0.$$

Therefore,

$$\begin{aligned} & \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[\sqrt{k} \left| \int_{K X(k)/b(n/k)}^\infty (v_n(u, \infty) - u^{-1/\xi}) \frac{b(n/k)}{X(k)} du \right| > \varepsilon/2\right] \\ & \leq \lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P\left[(1 + \eta)\sqrt{k} \int_{K(1-\eta)}^\infty |v_n(u, \infty) - u^{-1/\xi}| du > \varepsilon/2\right] + o(1). \end{aligned}$$

Now, since F satisfies Assumption 4.2, it suffices to show that

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\sqrt{k} \int_{K(1-\eta)}^{\infty} \left| v_n(x, \infty) - \frac{n}{k} \bar{F}(b(n/k)x) \right| dx > \frac{\varepsilon}{2(1-\eta)} \right] = 0. \tag{4.11}$$

This can be easily proved using the arguments in the proof of Proposition 9.1 in (Resnick [33], page 296). Observe that, by using the triangle and Chebyshev inequalities,

$$\begin{aligned} & P \left[\sqrt{k} \int_{K(1-\eta)}^{\infty} \left| v_n(x, \infty) - \frac{n}{k} \bar{F}(b(n/k)x) \right| dx > \frac{\varepsilon}{2(1-\eta)} \right] \\ & \leq P \left[\frac{1}{\sqrt{k}} \int_{K(1-\eta)}^{\infty} \sum_{i=1}^n \left| \varepsilon_{X_i/b(n/k)}(x, \infty) - \bar{F}(b(n/k)x) \right| dx > \frac{\varepsilon}{2(1-\eta)} \right] \\ & \leq \left(\frac{\varepsilon}{2(1-\eta)} \right)^{-2} \int_{K(1-\eta)}^{\infty} \frac{n}{k} \text{var}[\varepsilon_{X_i/b(n/k)}(x, \infty)] dx \\ & \leq \left(\frac{\varepsilon}{2(1-\eta)} \right)^{-2} \int_{K(1-\eta)}^{\infty} \frac{n}{k} \bar{F}(b(n/k)x) dx \xrightarrow{n \rightarrow \infty} \left(\frac{\varepsilon}{2(1-\eta)} \right)^{-2} \int_{K(1-\eta)}^{\infty} x^{-1/\xi} dx. \end{aligned}$$

The last limit follows from Karamata’s Theorem; cf. Resnick [33], page 25. Since $\xi < 1$, the integral in the last expression is finite and therefore (4.11), and hence (4.10), holds. From Theorem 3.5 in Resnick [33], page 56, we get $T(H_n) \Rightarrow T(H) = (\xi t^{-(1+\xi)} B(t), \int_y^\infty B(x^{-1/\xi}) dx)$ in $\mathbb{D}_l(0, 1] \times \mathbb{D}[1, \infty)$.

Now consider the random element Y_n in the space $\mathbb{D}_l^2(0, 1] \times \mathbb{D}[1, \infty)$,

$$Y_n(t, y) := \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}}, T(H_n)(t, y) \right).$$

By another application of the converging-together lemma, it is easy to check that $Y_n \Rightarrow Y$, where

$$Y(t, y) = \left(t^{-\xi}, \xi t^{-(1+\xi)} B(t), \int_y^\infty B(x^{-1/\xi}) dx \right).$$

The map $\tilde{T} : \mathbb{D}_l^2(0, 1] \times \mathbb{D}[1, \infty) \rightarrow \mathbb{D}_l^2(0, 1]$ defined by

$$\tilde{T}((f^{(1)}, f^{(2)}), g)(t) = (f^{(2)}(t), g(f^{(1)}(t))) \quad \text{for all } 0 < t \leq 1$$

is continuous at $(f, g) \in \mathbb{C}^2(0, 1] \times \mathbb{C}[1, \infty)$. Therefore

$$\begin{aligned} \tilde{T}(Y_n)(t) &= \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \sqrt{k} \int_{X_{(\lceil kt \rceil)}/X_{(k)}}^\infty (\hat{v}_n(y, \infty) - y^{-1/\xi}) dy \right) \\ &\Rightarrow \left(\xi t^{-(1+\xi)} B(t), \int_{t^{-\xi}}^\infty B(y^{-1/\xi}) dy \right) \quad \text{in } \mathbb{D}_l^2(0, 1]. \end{aligned}$$

This implies

$$\sqrt{k}(S_n(t) - S(t)) \Rightarrow \left(\xi t^{-(1+\xi)} B(t), \frac{1}{t} \int_{t^{-\xi}}^{\infty} B(y^{-1/\xi}) dy \right) \quad \text{in } \mathbb{D}_T^2(0, 1].$$

It is then easy to check that

$$\begin{aligned} & \left(\xi t^{-(1+\xi)} B(t), \frac{1}{t} \int_{t^{-\xi}}^{\infty} B(y^{-1/\xi}) dy \right) \\ & \stackrel{d}{=} \left(\xi t^{-(1+\xi)} B(t), \frac{\xi}{t} \int_0^t y^{-(1+\xi)} B(y) dy \right) \quad \text{in } \mathbb{D}_T^2(0, 1]. \end{aligned}$$

Also observe that

$$\tilde{S}_n(t) := \left(\left(\frac{\lceil kt \rceil}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi} \right) \tag{4.12}$$

$$+ \sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi}, \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{\lceil kt \rceil}{k} \right)^{-\xi} \right) \tag{4.13}$$

$$\begin{aligned} \Rightarrow \tilde{S}(t) & := \left(t^{-\xi}, \frac{\xi}{1-\xi} t^{-\xi} \right) \\ & + \left(\xi t^{-(1+\xi)} B(t), \frac{\xi}{t} \int_0^t y^{-(1+\xi)} B(y) dy \right) \quad \text{in } \mathbb{D}_T^2(0, 1], \end{aligned} \tag{4.14}$$

since

$$\sqrt{k} \left(\left(\frac{\lceil kt \rceil}{k} \right)^{-\xi} - t^{-\xi} \right) \rightarrow 0 \quad \text{as } k \rightarrow \infty$$

locally uniformly on $(0, 1]$. The proof the theorem is completed by applying Lemma 2.4 to \tilde{S}_n and \tilde{S} . □

4.3. Case II: $1/2 < \xi < 1$

When $1/2 < \xi < 1$, the distribution F admits a finite mean but not a finite variance. The ME function, however, exists, and we know the limit in probability of the scaled ME plot from Theorem 4.1.

Assumption 4.5. F satisfies Assumption 3.2, and, moreover,

$$\frac{1}{b(n)} \left(\frac{kb(n/k)}{1-\xi} u^{1-\xi} - C_{ku,n} \right) \rightarrow 0$$

for every $0 < u < 1$. For any $l < n$

$$C_{l,n} := n \int_0^{l/n} F^{\leftarrow}(1-u) \, du. \tag{4.15}$$

Theorem 4.6. Suppose X_1, \dots, X_n are i.i.d. observations from a distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ with $1/2 < \xi < 1$ and Assumption 4.5. Then for any $0 < \varepsilon < 1$

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \left(\sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right) \right\}; \\ &\quad \left. i = \lceil \varepsilon k \rceil, \dots, k \right\} \\ \Rightarrow \mathcal{MN} &:= \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1-\xi} t^{-\xi} + t^{-1} S_{1/\xi} \right), \varepsilon \leq t \leq 1 \right\} \quad \text{in } \mathcal{F}, \end{aligned}$$

where $B(t)$ is the standard Brownian bridge on $[0, 1]$ restricted to $(0, 1]$ and $S_{1/\xi}$ is a stable random variable independent of $B(t)$ with characteristic function

$$E[e^{itS_{1/\xi}}] = \exp \left\{ -\frac{1}{1-\xi} \Gamma \left(2 - \frac{1}{\xi} \right) \cos \frac{\pi}{2\xi} |t|^{1/\xi} \left[1 - i \operatorname{sgn}(t) \tan \frac{\pi}{2\xi} \right] \right\}. \tag{4.16}$$

Remark 4.7. An interesting point to note here is that the two coordinates of the weak limit \mathcal{MN} are independent. The empirical ME function depends on the sum of the order statistics $X_{(1)}, \dots, X_{(k)}$. When $1/2 < \xi < 1$, this sum is dominated by a very few high order statistics, and it turns out that the contribution of $X_{(k)}$ to the suitably normalized $\hat{M}(X_{(k)})$ vanishes in the limit. The proof below formalizes this idea.

This feature is in stark contrast to what happens in the case $0 < \xi < 1/2$. In that case all the top k order statistics have some contribution to $\hat{M}(X_{(k)})$ in the limit. Hence the two coordinates in the limit are obtained from the same Gaussian process and are definitely not independent.

Remark 4.8. Unfortunately, we are unable to obtain a proper weak limit of the ME plot in the case when $\xi = 1/2$. In this case it is known that the weak limit of the suitably normalized sum of the first k order statistics is Gaussian; cf. Csörgo, Haeusler and Mason [6]. So this would be similar to what happens when $0 < \xi < 1/2$, but the problem is that the integral $\int_0^t y^{-2} dB(y)$ does not exist. It is possible to redefine the ME plot in a different way, by leaving out a few of the top order statistics and obtaining a limit in that case, but we did not pursue that direction.

Proof of Theorem 4.6. From Theorem 3 in Csörgo, Horváth and Mason [7] we know that if $l = l_n \rightarrow \infty$ with $l_n/n \rightarrow 0$, then

$$\frac{1}{b(n)} \left(\sum_{i=1}^l X_{(i)} - C_{l,n} \right) \Rightarrow S_{1/\xi}. \tag{4.17}$$

Observe that, by Karamata’s theorem (Resnick [33], Theorem 2.1, page 25),

$$C_{l,n} = n \int_{n/l}^{\infty} b(s)/s^2 ds \sim n \frac{(n/l)b(n/l)}{(n/l)^2(1-\xi)} = \frac{lb(n/l)}{1-\xi}.$$

Choose $l = l_n$ such that $l/k \rightarrow 0$ as $n \rightarrow \infty$. Fix any $0 < u < 1$. Then

$$\begin{aligned} V_n(t) &= (V_n^{(1)}(t), V_n^{(2)}) := \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{1}{b(n)} \left(\sum_{i=1}^l X_{(i)} - C_{l,n} \right) \right) \\ &\Rightarrow (\xi t^{-(1+\xi)} B(t), S_{1/\xi}) \quad \text{in } \mathbb{D}_l[u, 1] \times \mathbb{R}, \end{aligned} \tag{4.18}$$

where $B(t)$ and $S_{1/\xi}$ are as described in the statement of the theorem. The convergence of the coordinates $V_n^{(1)}(t)$ and $V_n^{(2)}$ of $V_n(t)$ follows from (4.8) and (4.17). The asymptotic independence of $V_n^{(1)}(t)$ and $V_n^{(2)}$ is a consequence of Theorem D in Csörgo and Mason [5] or Satz 4 in Rossberg [36]. Using (3.8) we get that

$$\sqrt{k} \left(\frac{1}{kb(n/k)} \sum_{i=\lceil ku \rceil+1}^{\lceil kt \rceil} X_{(i)} - \frac{1}{1-\xi} (t^{1-\xi} - u^{1-\xi}) \right) \Rightarrow \int_u^t W(y) dy \quad \text{in } \mathbb{D}_l[u, 1],$$

and since $\xi > 1/2$, $kb(n/k)/(b(n)\sqrt{k}) \rightarrow 0$, which implies

$$\begin{aligned} U_n^{(2)}(t) &:= \frac{kb(n/k)}{b(n)} \left(\frac{1}{kb(n/k)} \sum_{i=\lceil ku \rceil+1}^{\lceil kt \rceil} X_{(i)} - \frac{1}{1-\xi} (t^{1-\xi} - u^{1-\xi}) \right) \\ &\rightarrow \mathbf{0} \quad \text{in } \mathbb{D}_l[u, 1], \end{aligned} \tag{4.19}$$

where $\mathbf{0} \in \mathbb{D}_l[u, 1]$ denotes the identically zero function. Furthermore, using Theorem 2 in Csörgo, Horváth and Mason [7], we get

$$\frac{1}{\sqrt{k}b(n/k)} \left(\sum_{i=l+1}^{\lceil ku \rceil} X_{(i)} - (C_{ku,n} - C_{l,n}) \right) \Rightarrow N(0, 1)$$

and hence

$$U_n^{(3)} := \frac{1}{b(n)} \left(\sum_{i=l+1}^{\lceil ku \rceil} X_{(i)} - (C_{ku,n} - C_{l,n}) \right) \rightarrow 0. \tag{4.20}$$

Combining (4.18), (4.19) and (4.20) and the converging-together lemma (Resnick [33], Proposition 3.1, page 57), we get an important building block of this proof,

$$\begin{aligned}
 U_n(t) &:= (V_n^{(1)}(t), U_n^{(2)}(t), U_n^{(3)}, V_n^{(2)}) \\
 &\Rightarrow (\xi t^{-(1+\xi)} W(t), \mathbf{0}, 0, S_{1/\xi}) \quad \text{in } \mathbb{D}_l^2[u, 1] \times \mathbb{R}^2.
 \end{aligned}
 \tag{4.21}$$

Next we consider

$$\begin{aligned}
 Z_n(t) &= (Z_n^{(1)}(t), Z_n^{(2)}(t)) := \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} t^{-\xi} \right) \right) \\
 &\in \mathbb{D}_l^2[u, 1]
 \end{aligned}$$

and focus on the second coordinate $Z_n^{(2)}(t)$.

$$\begin{aligned}
 Z_n^{(2)}(t) &= \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} - \frac{\xi}{1-\xi} t^{-\xi} \right) \\
 &= \frac{kb(n/k)}{b(n)} \left(\frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n/k)} - \frac{\xi}{1-\xi} t^{-\xi} \right) + o_P(1) \\
 &= \frac{kb(n/k)}{b(n)} \left(\frac{1}{(\lceil kt \rceil - 1)b(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{X_{(\lceil kt \rceil)}}{b(n/k)} - \frac{\xi}{1-\xi} t^{-\xi} \right) + o_P(1) \\
 &= \frac{kb(n/k)}{b(n)} \left(\frac{1}{(\lceil kt \rceil - 1)b(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{1}{1-\xi} t^{-\xi} \right) + o_P(1) \\
 &= \frac{kb(n/k)}{tb(n)} \left(\frac{1}{kb(n/k)} \sum_{i=1}^{\lceil kt \rceil - 1} X_{(i)} - \frac{1}{1-\xi} t^{1-\xi} \right) + o_P(1) \\
 &= \frac{1}{t} U_n^{(2)}(t) + \frac{1}{t} U_n^{(3)} + \frac{1}{t} V_n^{(2)} + o_P(1),
 \end{aligned}$$

where the last equality holds because of Assumption 4.5. Therefore, we get

$$Z_n(t) \Rightarrow (\xi t^{-(1+\xi)} B(t), t^{-1} S_{1/\xi}) \quad \text{in } \mathbb{D}_l^2[u, 1].$$

Since the above limit holds for every $0 < u < 1$, it holds in $\mathbb{D}_l(0, 1]$ as well. The proof is completed using Lemma 2.4. \square

4.4. Case III: $\xi \geq 1$

In this case, the distribution F need not have a finite mean, and the ME function may not be defined. It definitely does not exist if $\xi > 1$. Still the empirical ME plot can have a limit.

Theorem 4.9. Suppose X_1, \dots, X_n are i.i.d. observations with distribution F satisfying $\bar{F} \in RV_{-1/\xi}$ and Assumption 4.2.

1. If $\xi > 1$ and $n, k, n/k \rightarrow \infty$, then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi} + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{\hat{M}(X_{(i)})}{b(n)/k} \right) : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \{ (\xi t^{-(1+\xi)} B(t), t S_{1/\xi}) : t \geq 1 \} \end{aligned}$$

in \mathcal{F} , where $S_{1/\xi}$ is the positive stable random variable with index $1/\xi$ which satisfies, for $t \in \mathbb{R}$,

$$E[e^{it S_{1/\xi}}] = \exp \left\{ -\Gamma \left(1 - \frac{1}{\xi} \right) \cos \frac{\pi}{2\xi} |t|^{1/\xi} \left[1 - i \operatorname{sgn}(t) \tan \frac{\pi}{2\xi} \right] \right\},$$

and $B(t)$ is a Brownian bridge independent of $S_{1/\xi}$.

2. If $\xi = 1$, and k satisfies $n, k, n/k \rightarrow \infty$, and $kb(n/k)/b(n) \rightarrow 1$, then

$$\begin{aligned} \mathcal{MN}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi} + \sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{\hat{M}(X_{(i)})}{b(n/k)} - \frac{kC_{k,n}^*}{ib(n)} \right) : i = 2, \dots, k \right\} \\ &\Rightarrow \mathcal{MN} := \{ t(t^{-1} B(t), S_1 - 1 - \log t) : t \geq 1 \} \end{aligned}$$

in \mathcal{F} , where

$$C_{k,n}^* = n \int_{1/n}^{k/n} F^{\leftarrow}(1-u) du,$$

S_1 is a positively skewed stable random variable satisfying

$$E[e^{it S_1}] = \exp \left\{ it \int_0^\infty \left(\frac{\sin x}{x^2} - \frac{1}{x(1+x)} \right) dx - |t| \left[\frac{\pi}{2} + i \operatorname{sgn}(t) \log |t| \right] \right\},$$

and $B(t)$ is a Brownian bridge independent of S_1 .

Proof. The theorem is proved in the same fashion as the previous ones. First we prove the weak limit in the functional form of the ME plot, and then we infer the weak limit of the plot as a random set. Define

$$S_n(t) = \begin{cases} \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n)/k} \right) & \text{in part (i)} \\ \left(\sqrt{k} \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}} - t^{-\xi} \right), \frac{\hat{M}(X_{(\lceil kt \rceil)})}{b(n/k)} - \frac{kC_{k,n}^*}{\lceil kt \rceil b(n)} \right) & \text{in part (ii)} \end{cases} \quad \text{for all } 0 < t \leq 1.$$

We have already proved the weak limit of $S_n^{(1)}(t)$ and the weak limit of $S_n^{(2)}(t)$ is proved in Theorem 3.4 in Ghosh and Resnick [22]. The rest of the proof is completed using Lemma 2.4. \square

5. Confidence bounds for the plots

In Sections 3 and 4, we have obtained weak convergence limits for the QQ and ME plots in the Fell topology. Since the limit set in each case is a closed random set, we can compute from the results in Sections 3 and 4, the probability that the random limit set is contained in a fixed set in \mathbb{R}^2 . This leads to creating asymptotic $100(1 - \alpha)\%$ confidence bounds around the plots, given any $0 < \alpha < 1$. The methodology for creating confidence bounds around the plots is explained in details for QQ plots, and the same idea follows for ME plots.

5.1. QQ plots

Under the usual assumptions of Section 3, the QQ plot, \mathcal{Q}_n , as defined in (3.1), consists of $k = k(n) < n$ points in \mathbb{R}^2 . We know that $\mathcal{Q}_n \xrightarrow{P} \mathcal{Q}$, where \mathcal{Q} is a straight line. From Theorem 3.3, we also know that $\mathcal{Q}\mathcal{N}_n$, which is an affine transformation of \mathcal{Q}_n , converges weakly to a random set $\mathcal{Q}\mathcal{N}$ centered around \mathcal{Q} in \mathcal{F} . For fixed $0 < \alpha < 1$, we intend to create a confidence bound around \mathcal{Q}_n which will contain \mathcal{Q} with probability $1 - \alpha$ under the null hypothesis.

The limit distribution for QQ plots obtained in Theorem 3.3 is a linear transformation of $\{t^{-1}B(t): 0 < t \leq 1\}$ where B is a Brownian bridge on $[0, 1]$. So the limit explodes as t comes close to 0, and thus we create confidence bounds under an ε truncation to avoid this. Define

$$\mathcal{Q}_n^\varepsilon := \left\{ \left(-\log \frac{j}{k}, \log \frac{X(j)}{X(k)} \right) : 1 \leq j \leq k \text{ and } \frac{j}{k} \geq \varepsilon \right\}, \quad k < n \tag{5.1}$$

$$\mathcal{Q}^\varepsilon := \{(-\log t, -\xi \log t) : \varepsilon \leq t \leq 1\} = \left\{ (x, \xi x) : 0 \leq x \leq \log \frac{1}{\varepsilon} \right\}. \tag{5.2}$$

Now with similar truncations defined as above, it follows from (3.14) that $\tilde{S}_n^\varepsilon \Rightarrow \tilde{S}^\varepsilon$ in $\mathbb{D}_T^2[\varepsilon, 1]$. This means $\mathcal{Q}\mathcal{N}_n^\varepsilon \Rightarrow \mathcal{Q}\mathcal{N}^\varepsilon$ in \mathcal{F} where $\mathcal{Q}\mathcal{N}^\varepsilon$ and $\mathcal{Q}\mathcal{N}_n^\varepsilon$ are the truncated versions of $\mathcal{Q}\mathcal{N}$ and $\mathcal{Q}\mathcal{N}_n$, respectively, defined in (3.4). Suppose we can calculate $c_{\alpha/2,\varepsilon}$ such that $P(\sup_{\varepsilon \leq t \leq 1} \frac{|B(t)|}{t} \leq c_{\alpha/2,\varepsilon}) = 1 - \alpha$. Then a conservative $100(1 - \alpha)\%$ confidence bound around $\mathcal{Q}_n^\varepsilon$ is given by

$$\mathcal{C}\mathcal{Q}_n^\varepsilon = \mathcal{Q}_n^\varepsilon + \left\{ (0, y) : y \in \xi \left(-\frac{c_{\alpha/2,\varepsilon}}{\sqrt{k}}, \frac{c_{\alpha/2,\varepsilon}}{\sqrt{k}} \right) \right\}. \tag{5.3}$$

It is easy to see that

$$P[\mathcal{Q}\mathcal{N}^\varepsilon \subset \mathcal{C}\mathcal{Q}^\varepsilon] \geq 1 - \alpha,$$

where

$$\mathcal{C}\mathcal{Q}^\varepsilon := \mathcal{Q}^\varepsilon + \left\{ (0, y) : y \in \xi \left(-\frac{c_{\alpha/2,\varepsilon}}{\sqrt{k}}, \frac{c_{\alpha/2,\varepsilon}}{\sqrt{k}} \right) \right\}.$$

An equivalent statement in a different notation is

$$P[\{\rho(x, \mathcal{Q}\mathcal{N}^\varepsilon) = 0, \forall x \in \mathcal{C}\mathcal{Q}^\varepsilon\}] \geq 1 - \alpha,$$

where, for any $x \in \mathbb{R}^2$ and $F \in \mathcal{F}$,

$$\rho(x, F) := \inf\{|x - y|: y \in F\}.$$

From (Molchanov [29], Theorems B.6 and B.13, pages 400–401), we know that if $F_n \rightarrow F$ in \mathcal{F} , then for any compact set $K \subset \mathbb{R}^2$

$$\sup_{x \in K} |\rho(x, F_n) - \rho(x, F)| \rightarrow 0.$$

Since $\mathcal{QN}_n^\varepsilon \Rightarrow \mathcal{QN}^\varepsilon$ in \mathcal{F} , we get

$$\lim_{n \rightarrow \infty} P\{\{\rho(x, \mathcal{QN}_n^\varepsilon) = 0, \forall x \in \mathcal{CQ}^\varepsilon\} \geq 1 - \alpha\}.$$

Hence, $\mathcal{CQ}_n^\varepsilon$ in (5.3) is an asymptotic $100(1 - \alpha)\%$ confidence bound for \mathcal{Q}^ε .

We calculate $P(\sup_{\varepsilon \leq t \leq 1} \frac{|B(t)|}{t} \leq M)$ next in order to complete the construction. Since $W(t) := (1 + t)B(\frac{t}{1+t}), 0 \leq t < \infty$ is a Brownian motion on $[0, \infty)$, we can check that

$$\sup_{\varepsilon \leq t \leq 1} \frac{|B(t)|}{t} = \sup_{t \geq \delta} \frac{|W(t)|}{t}, \tag{5.4}$$

where $\delta = \frac{\varepsilon}{1-\varepsilon}$. In the following theorem we compute the boundary-crossing probability for $\sup_{t \geq \delta} \frac{|W(t)|}{t}$.

Proposition 5.1. *Suppose W is a standard Brownian motion on $[0, \infty)$. Then for all $\delta > 0$ and $M > 0$,*

$$P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) = 4 \sum_{k=1}^{\infty} [\Phi((4k + 1)M\sqrt{\delta}) - \Phi((4k - 1)M\sqrt{\delta})], \tag{5.5}$$

where $\Phi(\cdot)$ denotes the c.d.f. of a standard normal distribution.

Proof. We begin by observing that

$$\begin{aligned} &P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} \leq M\right) \\ &= \int_{s=-M\delta}^{M\delta} P(-Mt - (M\delta + s) \leq W(t) \leq Mt + (M\delta - s), \forall t \geq 0) f_{W(\delta)}(s) ds, \end{aligned} \tag{5.6}$$

where $f_{W(\delta)}$ denotes the density of $W(\delta)$. The right-hand side is obtained by conditioning on $W(\delta) = s$ and using the fact that $\{W(t) - W(\delta): t \geq \delta\}$ is independent of $W(\delta)$ and $\{W(\delta): t \geq \delta\} \stackrel{d}{=} \{W(t): t \geq 0\}$. Now the above boundary (non-)crossing probability of the Brownian motion

can be calculated using Doob [16], equation (4.3), as

$$\begin{aligned}
 &P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) \\
 &= 1 - \int_{s=-M\delta}^{M\delta} \left[1 - \sum_{k=1}^{\infty} (e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k})\right] f_{W(\delta)}(s) \, ds,
 \end{aligned}$$

where

$$\begin{aligned}
 A_k &= [(2k - 1)M]^2\delta - (2k - 1)Ms, & B_k &= [(2k - 1)M]^2\delta + (2k - 1)Ms, \\
 C_k &= 4k^2M^2\delta - 2kMs, & D_k &= 4k^2M^2\delta + 2kMs.
 \end{aligned}$$

Since $W(\delta) \sim N(0, \delta)$, for any $a, b \in \mathbb{R}$, we have

$$\int_{-a}^a e^{bs} f_{W(\delta)}(s) \, ds = e^{b^2\delta/2} \left[\Phi\left(\frac{a - b\delta}{\sqrt{\delta}}\right) - \Phi\left(\frac{-a - b\delta}{\sqrt{\delta}}\right) \right]. \tag{5.7}$$

Now using (5.7), we can compute, for each $k \geq 1$,

$$\begin{aligned}
 \int_{-M\delta}^{M\delta} (e^{-2A_k} + e^{-2B_k}) f_{W(\delta)}(s) \, ds &= 2[\Phi((4k - 1)M\sqrt{\delta}) - \Phi((4k - 3)M\sqrt{\delta})], \\
 \int_{-M\delta}^{M\delta} (e^{-2C_k} + e^{-2D_k}) f_{W(\delta)}(s) \, ds &= 2[\Phi((4k + 1)M\sqrt{\delta}) - \Phi((4k - 1)M\sqrt{\delta})].
 \end{aligned}$$

Therefore we get

$$\begin{aligned}
 &P\left(\sup_{t \geq \delta} \frac{|W(t)|}{t} > M\right) \\
 &= 1 - \int_{s=-M\delta}^{M\delta} \left[1 - \sum_{k=1}^{\infty} (e^{-2A_k} + e^{-2B_k} - e^{-2C_k} - e^{-2D_k})\right] f_Z(s) \, ds \\
 &= 4 \sum_{k=1}^{\infty} [\Phi((4k - 1)M\sqrt{\delta}) - \Phi((4k - 3)M\sqrt{\delta})].
 \end{aligned}$$

□

Remark 5.2. Observe that the confidence bound in (5.3) depends on the value of ξ . While obtaining the width of the band, we replace ξ by its Hill estimate (Resnick [33], page 74). We could use any consistent estimator of ξ and the choice of the estimator does not seem to be important as far as the simulation study is concerned. It is well known that estimating the parameter ξ can often be extremely tricky, see ‘‘Hill–Horror plots’’ in (Resnick [33], page 87). But as far as obtaining confidence bounds is concerned, we can get past that by using a conservative estimate of ξ , that is, a value which we strongly believe is not less than the true value of ξ .

Remark 5.3. It is clear that the probability calculated in Proposition 5.1 is very close to 1 if $M\sqrt{\delta}$ is small. We can approximate the infinite sum in (5.5) by a finite sum whose limit depends on our choice of M and δ . We use Proposition 5.1 to create confidence bands for the QQ plots in the examples in Section 6. Simulation suggests that considering the first 15 terms of the infinite sum is enough to give us approximations correct up to six decimal places.

Remark 5.4. It is possible to join the subsequent points in \mathcal{Q}_n to make a continuous curve $\mathcal{Q}_n^* \in \mathcal{F}$, and we can check that \mathcal{Q}_n^* will converge to the same limit as that of \mathcal{Q}_n as $n \rightarrow \infty$. We mention this result here without proof, which can be completed following Theorem 3.3.

5.2. ME plots

In Section 4 we obtained weak limits for the ME plots, under the assumption that $\bar{F} \in RV_{-1/\xi}$ with $\xi > 0$, where F denotes the underlying distribution. We observed three separate limits in three different cases.

For the case $0 < \xi < 1/2$, where F has a finite second moment, we obtain a limit in terms of functionals of Brownian bridges (see Theorem 4.3). In order to convert this result to obtain confidence bounds, we need to compute boundary-crossing probabilities for these functionals. Analytical solution for such probabilities are available for linear boundaries (Doob [16]) and piecewise linear boundaries (Pötzelberger and Wang [31]) in case of Brownian motion on $[0, \infty)$. Probabilities for nonlinear boundaries, which happens to be our case, are usually approximated using results for piecewise linear boundaries. Instead of such approximations, we resort to Monte Carlo simulation to find appropriate confidence bounds; see Section 6.

For the case $1/2 < \xi < 1$, F has a finite first moment, but its second moment does not exist. The limit distribution for the affinely transformed ME plot consists of a functional of a Brownian bridge in the first component and a Stable distribution in the second component. The feature here is that the normalization required to get the limit depends on $b(n)$ and $b(n/k)$, which in turn depends on the distribution function F and is hence unknown. These can be estimated in practice with $X_{(1)}$ and $X_{(k)}$ respectively. Although, to justify such a procedure we would need to know the joint behavior $(X_{(1)}, X_{(k)}, \sum_{i=1}^k X_{(i)})$ when k, n and $n/k \rightarrow \infty$. Results in Darling [8], Chow and Teugels [4], Resnick [32], Section 4, are quite useful here. Using Theorem 5.3 in Darling [8] we can show that under the assumptions of Theorem 4.6,

$$\begin{aligned} \widetilde{\mathcal{MN}}_n &:= \left\{ \left(\left(\frac{i}{k} \right)^{-\xi}, \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right. \\ &\quad \left. + \left(\sqrt{k} \left(\frac{X_{(i)}}{X_{(k)}} - \left(\frac{i}{k} \right)^{-\xi} \right), \frac{kX_{(k)}}{X_{(1)}} \left(\frac{\hat{M}(X_{(i)})}{X_{(k)}} - \frac{\xi}{1-\xi} \left(\frac{i}{k} \right)^{-\xi} \right) \right) \right\}; \\ &\quad i = 2, \dots, k \} \tag{5.8} \\ \Rightarrow \widetilde{\mathcal{MN}} &:= \left\{ \left(t^{-\xi} + \xi t^{-(1+\xi)} B(t), \frac{\xi}{1-\xi} t^{-\xi} + t^{-1} \tilde{S}_{1/\xi} \right), 0 < t \leq 1 \right\} \text{ in } \mathcal{F}, \end{aligned}$$

where $\tilde{S}_{1/\xi}$ is independent of $B(t)$, and its characteristic function is of the form

$$E[e^{i\lambda\tilde{S}_{1/\xi}}] = e^{i\lambda} \left(1 + \frac{i\lambda}{1-\xi} - \frac{1}{\xi} \int_0^1 (e^{it\lambda} - 1 - it\lambda)t^{-1-1/\xi} dt \right)^{-1}. \tag{5.9}$$

We again resort to Monte Carlo simulation to obtain confidence bounds for the ME plots.

For $\xi \geq 1$, F need not have a finite mean, and the ME plot does not have a non-trivial non-random limit. We obtain weak limits here in Theorem 4.9. Clearly, calculating confidence bounds is not sensible here.

5.2.1. Confidence bound for ME plots

We need to truncate the ME plot near infinity in this case, since the weak limits we obtain (Theorems 4.3 and 4.6) blow up there (relates to t near 0 in the limit \mathcal{MN}_n). According to (4.1), \mathcal{M}_n denotes the ME plot for a sample of size n (with $k < n$ top order statistics under consideration). Define its truncated version

$$\begin{aligned} \mathcal{M}_n^\varepsilon &:= \frac{1}{X_{(k)}} \{ (X_{(i)}, \hat{M}(X_{(i)})) : i = \lceil k\varepsilon \rceil, \dots, k \} \quad \text{and} \\ \mathcal{M}^\varepsilon &:= \left\{ \left(t, \frac{\xi}{1-\xi} t \right) : \varepsilon \leq t \leq 1 \right\}. \end{aligned} \tag{5.10}$$

Then $\mathcal{M}_n^\varepsilon \xrightarrow{P} \mathcal{M}^\varepsilon$.

If $0 < \xi < 1/2$, then using Theorem 4.3 we can give the $(1 - \alpha)100\%$ confidence band for \mathcal{M}^ε as

$$\mathcal{CM}_n^\varepsilon := \mathcal{M}_n^\varepsilon + \left\{ (x, y) : x \in \left(-\frac{c_{\alpha_1/2,\varepsilon}}{\sqrt{k}}, \frac{c_{\alpha_1/2,\varepsilon}}{\sqrt{k}} \right), y \in \left(-\frac{d_{\alpha_2/2,\varepsilon}}{\sqrt{k}}, \frac{d_{\alpha_2/2,\varepsilon}}{\sqrt{k}} \right) \right\}, \tag{5.11}$$

where $\alpha_1, \alpha_2 > 0$ is such that $\alpha = \alpha_1 + \alpha_2$ and

$$\begin{aligned} c_{\alpha,\varepsilon} &= (1 - \alpha)\text{th quantile of } \sup_{\varepsilon \leq t \leq 1} \xi t^{-(1+\xi)} B(t), \\ d_{\alpha,\varepsilon} &= (1 - \alpha)\text{th quantile of } \sup_{\varepsilon \leq t \leq 1} \xi t^{-1} \int_0^t y^{-(1+\xi)} B(y) dy. \end{aligned} \tag{5.12}$$

$\mathcal{CM}_n^\varepsilon$ in (5.11) provides an asymptotic confidence bound around \mathcal{M}^ε with $P(\mathcal{M}^\varepsilon \subset \mathcal{CM}_n^\varepsilon) \geq (1 - \alpha)$ for large n .

If $1/2 < \xi < 1$, then we use Theorem 4.6 and its modified form in (5.8) to give the $(1 - \alpha)100\%$ confidence band for \mathcal{M}^ε as

$$\begin{aligned} \mathcal{CM}_n^\varepsilon &= \left\{ \left(\frac{X_{(\lceil kt \rceil)}}{X_{(k)}}, \frac{\hat{M}(X_{(\lceil kt \rceil)})}{X_{(k)}} \right) \right. \\ &\quad \left. + \left(-\frac{c_{\alpha_1/2,\varepsilon}}{\sqrt{k}}, \frac{c_{\alpha_1/2,\varepsilon}}{\sqrt{k}} \right) \times \left(\frac{X_{(1)}d_{1-\alpha_2/2}}{\lceil kt \rceil X_{(k)}}, \frac{X_{(1)}d_{\alpha_2/2}}{\lceil kt \rceil X_{(k)}} \right) : \varepsilon \leq t \leq 1 \right\}, \end{aligned} \tag{5.13}$$

where

$$d_\alpha = (1 - \alpha)\text{th quantile of } \tilde{S}_{1/\xi} \text{ defined in (5.9).}$$

Here $0 < \alpha_1, \alpha_2 < 1$ are chosen such that $(1 - \alpha) = (1 - \alpha_1)(1 - \alpha_2)$. Since the random components in the first and second components in the limit (Theorem 4.6) are independent, this gives us the right confidence interval so that $P(\mathcal{M}^\varepsilon \subset \mathcal{CM}_n^\varepsilon) \geq 1 - \alpha$. The above quantiles are calculated by Monte Carlo methods for the simulation we report in Section 6.1.2.

Remark 5.5. Throughout the literature of extreme value theory, the top k order statistics where $k = k_n \rightarrow \infty$ and $k/n \rightarrow 0$ as $n \rightarrow \infty$ is considered for inference. The idea is that as the size of data increases we concentrate more on the extreme right-hand tail of the underlying distribution. In practice though, given a data set of fixed size n , albeit large, it is difficult to decide on which value of k to choose. The popular solution is to try out different values of k ; see Embrechts, Klüppelberg and Mikosch [18], Chapter 6, and Resnick [34], Chapter 4, for further discussions on this issue.

In order to obtain confidence bounds for QQ plots and ME plots, along with the problem of choosing k , we also have to choose ε . The choice of ε should be such that, for the purpose of drawing any inference, we leave out the region where data is sparse. In practice, we have to try out different values of ε depending on the size of the data and the choice of k .

Remark 5.6. An important point to note here is that we are suggesting to use the weak limit of the QQ plot to obtain the confidence band. In practice, even if we have a large data set, it will always be finite. A natural question that arises here is what is the rate of convergence in these cases. We do not have the answer at the moment, but all the simulation studies that we have done strongly suggest that this method works well.

6. QQ plot and ME plot in practice

6.1. Simulation

We do a simulation study using the software R to check how well this method of obtaining confidence bounds for the QQ plot and the ME plot works.

6.1.1. QQ plots

We begin with a simple exercise for Pareto distribution with $\xi = 0.25$ ($\bar{F}(x) = x^{-4}$, $x \geq 1$). We simulate a sample of size $n = 50,000$ from this distribution and look at the QQ plot for extremes as defined in (3.1); see Figure 1. The black line denotes the plot \mathcal{Q}_n , and the brown dotted line denotes the true line \mathcal{Q} . We know that \mathcal{Q}_n converges to \mathcal{Q} , and, as we see in the plot, the two lines are close, except for the top-right corner of the plots, which correspond to the very large order statistics. We choose three different values for k : 2000, 1500 and 1000, which are large in absolute terms, but small compared to the sample size n .

Following the discussion in Section 5, we know that the variance of the limiting distribution blows up as we move towards the extreme order statistics (towards the top-right corner)

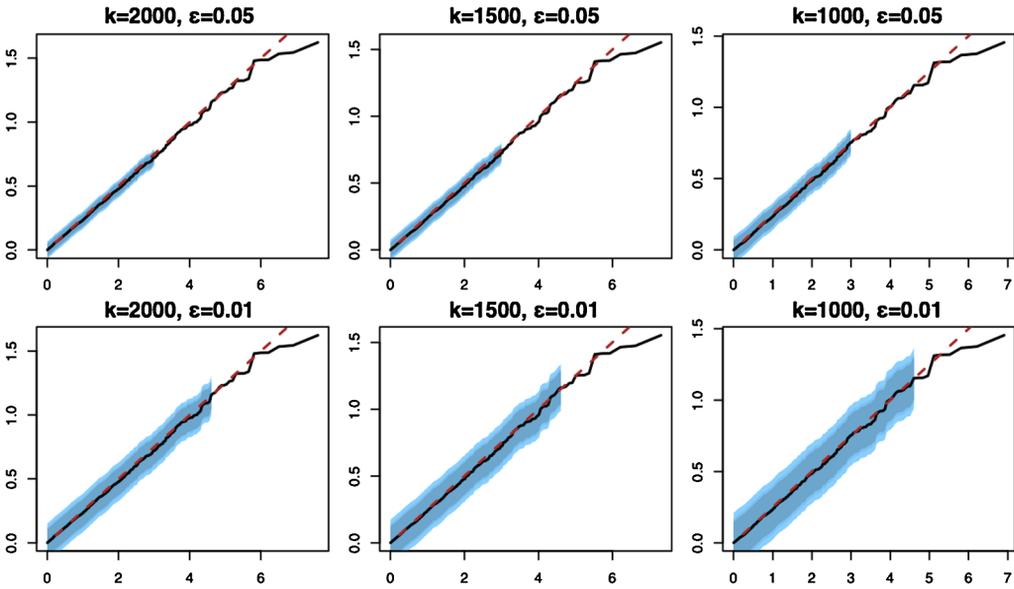


Figure 1. QQ plot for 50,000 i.i.d. Pareto random variables with $\xi = 0.25$.

in the plot. So while obtaining a confidence bound, we truncate at $\lceil \varepsilon k \rceil$ th order statistic for $\varepsilon = 0.05$ and 0.01 . The confidence bounds are obtained for the six cases. The three shades of the colored bands signify the 99%, 95% and the 90% confidence bands for the plot. As is evident in Figure 1, the true line lies within the bound in all the cases. It is also notable that the width of the confidence band increases as k and ε decrease.

Next we do a similar study for a right-skewed stable distribution with $\xi = 2/3$ ($\alpha = 1.5$) and mean 0. We use the same values of n, k and ε . The result is given in Figure 2. Here also we see that the method works well, and the confidence band contains the true line in all the six cases.

We also try a non-standard distribution for which $\bar{F}^{-1}(x) = x^{-1/5}(1 - 10^{-1} \ln x)$, $0 < x \leq 1$. This means that $\bar{F} \in RV_{-4}$, and therefore $\xi = 0.5$. The exact form of \bar{F} is given by

$$\bar{F}(x) = \frac{1}{32} W(2xe^2)^5 x^{-5} \quad \text{for all } x \geq 1, \tag{6.1}$$

where W is the Lambert W function satisfying $W(x)e^{W(x)} = x$ for all $x > 0$. Observe that $W(x) \rightarrow \infty$ as $x \rightarrow \infty$ and $W(x) \leq \log(x)$ for $x > 1$. Furthermore,

$$\frac{\log(x)}{W(x)} = 1 + \frac{\log W(x)}{W(x)} \rightarrow 1 \quad \text{as } x \rightarrow \infty,$$

and hence $W(x)$ is a slowly varying function. This is therefore an example where the slowly varying term contributes significantly to \bar{F} . That was not the case in the Pareto or the stable examples. The result of the simulation is shown in Figure 3. As expected, the choice of k plays

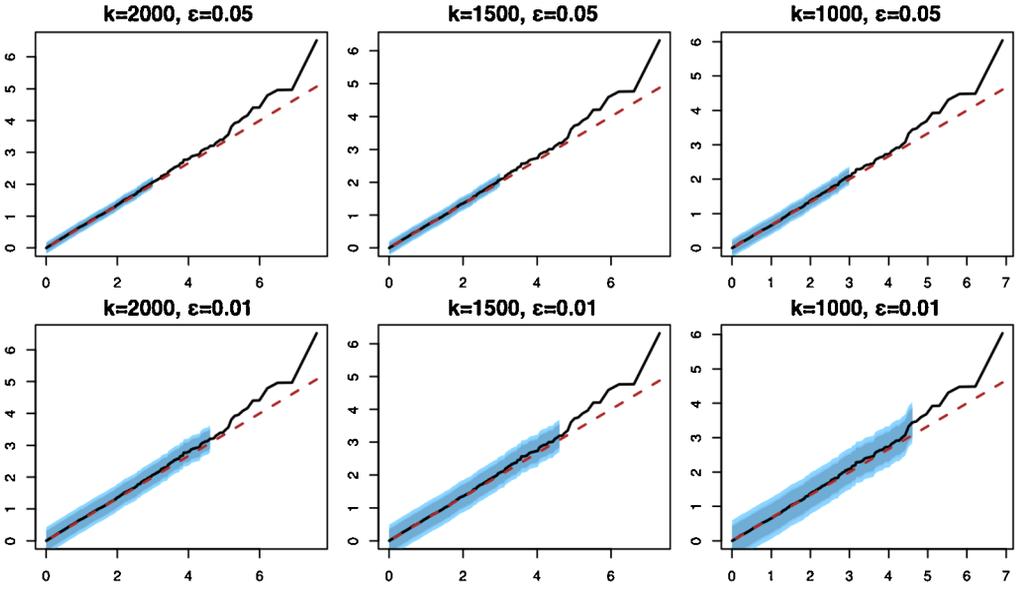


Figure 2. QQ plot for 50,000 i.i.d. right-skewed stable random variables with $\xi = 2/3$.

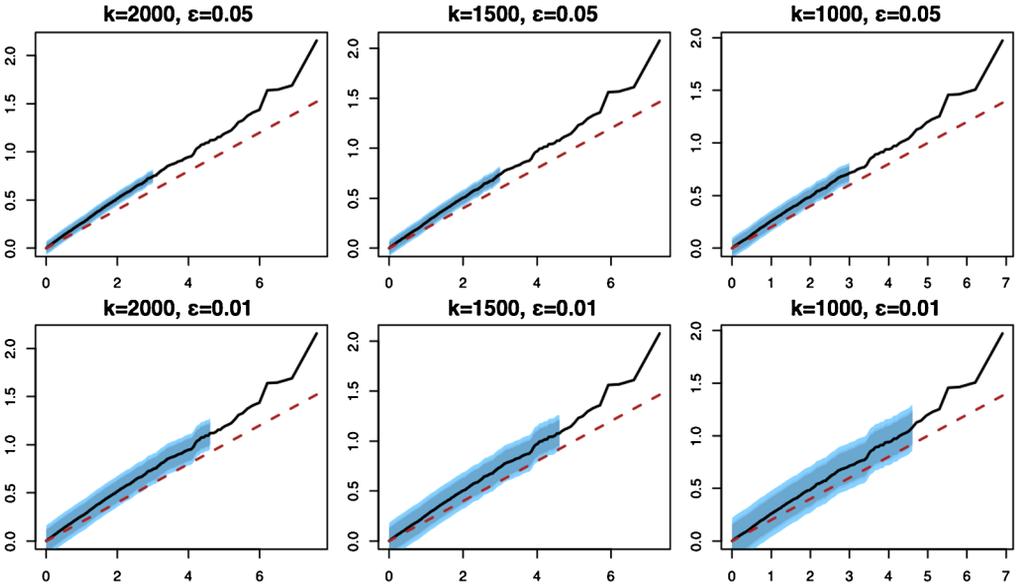


Figure 3. QQ plot for 50,000 i.i.d. random variables with the distribution described in (6.1) ($\xi = 0.2$).

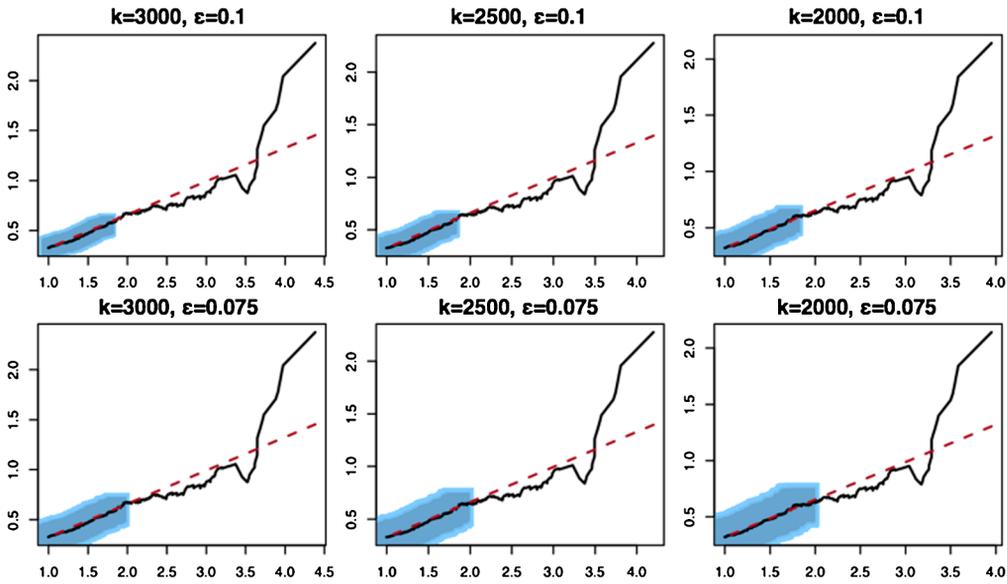


Figure 4. ME plot for 50,000 i.i.d. Pareto random variables with $\xi = 0.25$.

an important role in this case, and we see that the confidence band contains the true line when we choose $k = 1000$ and $\varepsilon = 0.01$. Although not shown in Figure 3, the confidence bands perform better for smaller values of k .

6.1.2. ME plots

Figure 4 shows the ME plot obtained from a data simulated from the Pareto distribution with $\xi = 0.25$. The six plots correspond to different values of k (3000, 2500 and 2000) and ε (0.1 and 0.075). The black line is the observed ME plot, and the brown dotted line denotes the limit in probability. Again, the three shades of the colored bands denote the 99%, 95% and the 90% confidence bands for the plot, respectively. Note that the weak limit is a functional of the Brownian bridge and depends on ξ . We estimate ξ using the Hill estimator and obtain the bounds by simulating 10,000 paths from the weak limit.

A striking feature in all these plots is that they are close to being linear near the bottom-left corner and become quite erratic near top-right corner. The reason behind this phenomenon is that the empirical ME function for high thresholds is the average of the excesses of a small number of upper order statistics. When averaging over few numbers, there is high variability, and therefore this part of the plot appears very nonlinear and is uninformative. Therefore, while obtaining confidence bands it is essential to leave out some of the extreme order statistics. We would also like to point out that, without the confidence bands, it would have been difficult to believe that these plots were obtained from a distribution with tail index 0.25.

A simulation of ME plot for the right skewed stable distribution with $\xi = 2/3$ is shown in Figure 5. We use the band described in (5.13) and estimate the quantiles using simulation. In this

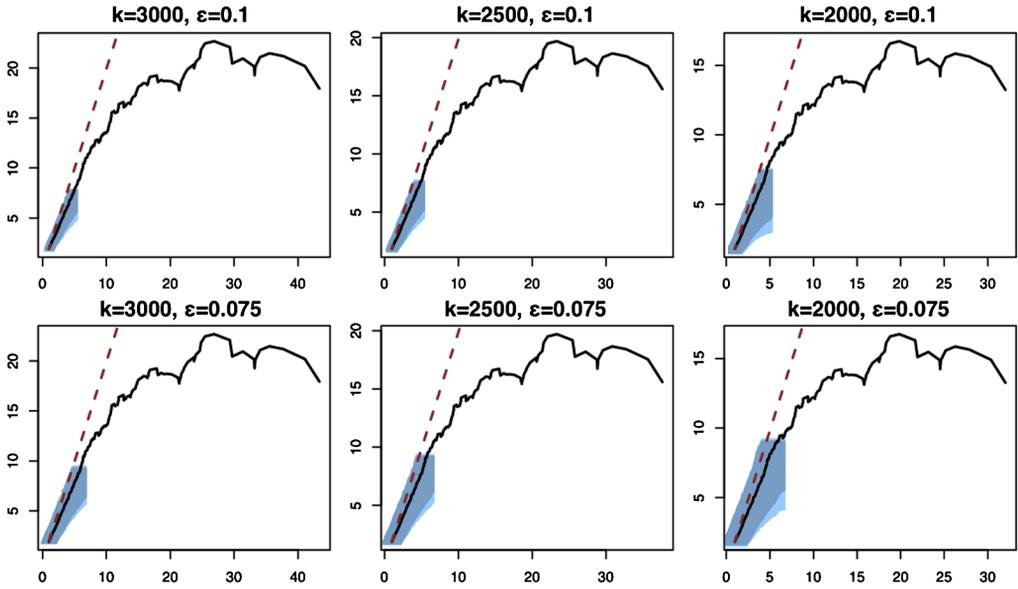


Figure 5. ME plot for 50,000 i.i.d. right-skewed stable random variables with $\xi = 2/3$.

case we only provide the 95% and the 90% confidence band. The 99% confidence band for the stable is very large and using that is not much helpful.

The next simulation is the ME plot for a sample from the distribution function described in (6.1), and the result is given in Figure 6. We use the same values for n, k and ε . We see that this method of getting confidence bands works well in these cases.

6.2. An example with a real data

We study a data set which contains Internet response sizes corresponding to user requests. The sizes are thresholded to be at least 100 KB. The data set consists of 67,287 observations and is part of a bigger set collected in April 2000 at the University of North Carolina at Chapel Hill.

It is often stated that file size data typically exhibits heavy tails, and we observe that is indeed the case here. Figure 7 shows various plots from this data set. The sample variance is of the order of 10^{13} which suggests that the variance is possibly infinite for the underlying distribution (denote by F). This would imply that if \bar{F} is regularly varying for some ξ , then we must have $\xi \geq 1/2$. This is suggested by both the Pickands plot and the Hill plot (Figure 7(b) and (c), resp.). The Hill plot is always above $1/2$ and the Pickands is above $1/2$ for most of the range. But it is difficult to get an estimate of ξ using these two tools since both plots are highly fluctuating and hence inconclusive. We fit a GPD model with the top 2000 order statistics using the command “fit.GPD” in the library “QRMLib.” It gives an estimate 0.6218 of ξ and Figure 7(d) plots the estimated \bar{F} in the log-log scale along with the fitted line.

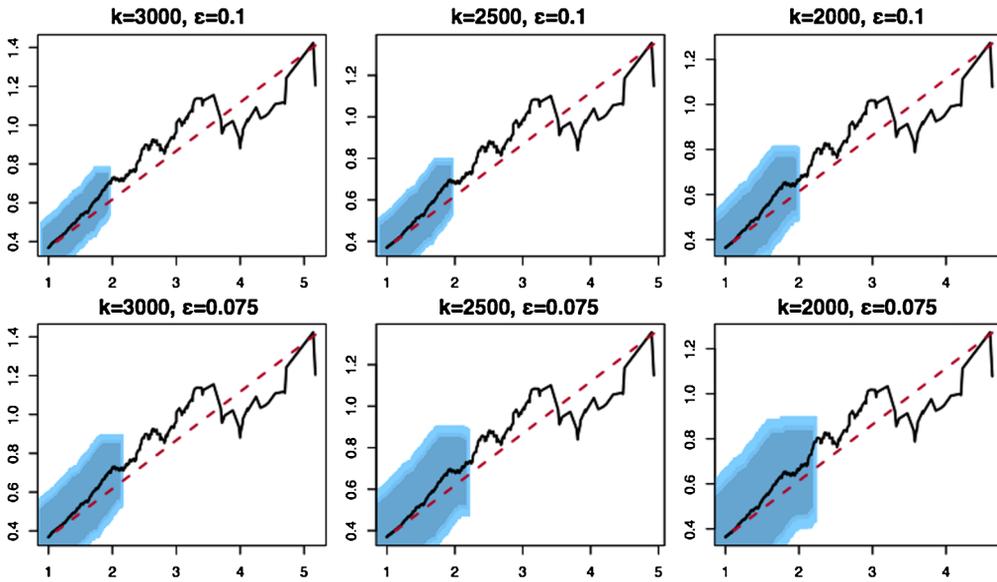


Figure 6. ME plot for 50,000 i.i.d. random variables with the distribution described in (6.1) ($\xi = 0.2$).

We try the QQ plot with data set for $k = 4000$ and 2000 (top 6% and 3% order statistics approximately) and with $\varepsilon = 0.05$ and 0.02 . The plots give an estimate of around 0.62 of ξ . The plots are shown in Figure 7(e)–(h). The ME plots for $k = 5000, 3000$ and $\varepsilon = 0.06, 0.04$ are shown in Figure 7(i)–(l), and they also suggest a similar estimate for ξ .

We observe that, in this example, the different methods of understanding the tail behavior of a data work very well, and all of them are in agreement about the value of ξ . This is not true in many situations, and then it is hard to judge which method one should trust. In those cases it is important to have some more knowledge about the system from which the data was collected, and often that helps in the understanding of the data.

7. Conclusion

Plotting techniques have always been popular as diagnostic tools for goodness-of-fit of observed data, and we believe they will remain so because of their visual and intuitive appeal. In this paper we have concentrated on two such tools used extensively in the extreme-value literature. A weak law of large numbers has been shown previously for both the QQ plots (Das and Resnick [9]) and ME plots (Ghosh and Resnick [22]), considering them as random elements in an appropriate topology. Our contribution in this paper has been to provide distributional limits for them. In the case of QQ plots, we have also provided an explicit expression for confidence bounds (with a truncation to avoid the confidence bounds from blowing up) by using these distributional results. In the case of ME plots we have obtained distributional limits in the cases $0 < \xi < 1/2$,

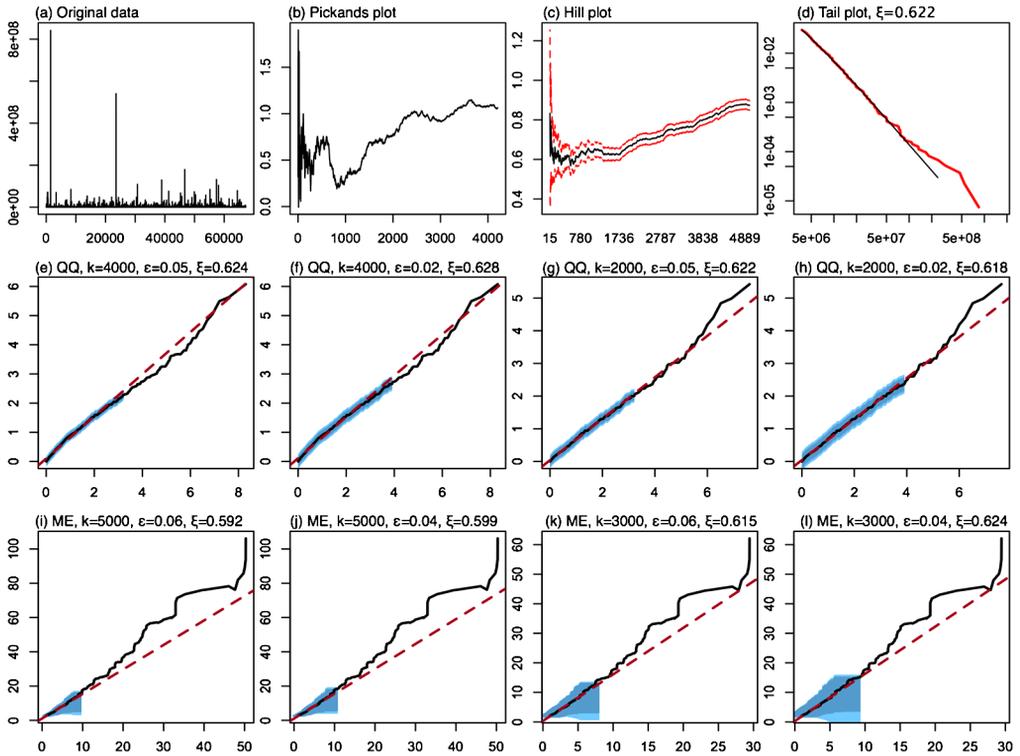


Figure 7. Analysis of the internet response size data.

$1/2 < \xi < 1$ and $\xi \geq 1$ separately where the underlying distribution F is assumed to be regularly varying with index $-1/\xi$. The case $\xi = 1/2$ is still open. We have produced confidence bounds for the ME plots in these cases by Monte Carlo simulation, as explicit expressions for these quantities are not easy to calculate. The explicit expressions would involve boundary-crossing probabilities for a Brownian Bridge with nonlinear boundaries. Boundary-crossing probabilities for Brownian motion can be approximated using piecewise linear boundaries Pötzelberger and Wang [31], but we do not know of a nice approximation for the Brownian Bridge case; hence we resort to simulation. We have illustrated the confidence bounds in both the cases of QQ plots and ME plots with simulated and real data examples in Section 6. The importance of the confidence bounds can be understood very clearly from Figure 4. Here we have a simulated data set of 50,000 points from a Pareto distribution with parameter $\xi = 0.25$. Just looking at the ME plot, it is not at all obvious that this is a heavy-tailed data, whereas when the confidence bounds with the ε -truncation are drawn, the straight line with slope $\xi = 0.25$ remains inside the bounds indicating the true nature of the data.

Since we are using the limiting distribution to obtain the confidence bounds, it is natural to ask what the rate of convergence is. We have observed that this method works well in the simulation

studies that we have done, but we have not answered this theoretically. This is currently a work in progress.

A standing assumption in the results we proved in this paper is that the random variables X_n are i.i.d. We believe that it is possible to obtain similar results under a more general assumption of stationarity and mixing; cf. Rootzén [35]. We intend to look into this further.

We should also note here that often practitioners use the median-excess plot with the implied meaning when $\xi > 1$; that is, the mean for the distribution does not exist (Embrechts, Klüppelberg and Mikosch [18]), but we have not ventured into this kind of plotting tool. We have also not looked into other kinds of plots used in extremes, like the Stărică plot (Stărică [40]) to determine the right k number of upper order statistics, or the Gertensgarbe and Werner plot (Gertensgarbe and Werner [21]), for determining thresholds, over which a data may be assumed to be extreme-valued, or the more popular Hill plot, Pickands plot (Resnick [33]), to detect the right value of the extreme-value parameter. Obtaining results in the same spirit as this paper for these other varieties of plots are a part of intended future research.

Acknowledgements

The authors are thankful to Paul Embrechts (ETH Zurich), Sidney I. Resnick (Cornell University) and Gennady Samorodnitsky (Cornell University) for their detailed comments on a draft of the paper which greatly helped in improving the paper. The authors are also thankful for insightful comments and suggestions from the referees and the associate editor. Bikramjit Das was partially supported by the program IRTG/Pro*Doc. Souvik Ghosh was partially supported by the FRAP program at Columbia University.

References

- [1] Beer, G. (1993). On the Fell topology. *Set-Valued Anal.* **1** 69–80. [MR1230370](#)
- [2] Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley. [MR0233396](#)
- [3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation. Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. [MR0898871](#)
- [4] Chow, T.L. and Teugels, J.L. (1978). The sum and the maximum of iid random variables. In *Proceedings of the Second Prague Symposium on Asymptotic Statistics*. Amsterdam: North-Holland.
- [5] Csörgő, M. and Mason, D.M. (1985). On the asymptotic distribution of weighted uniform empirical and quantile processes in the middle and on the tails. *Stochastic Process. Appl.* **21** 119–132. [MR0834992](#)
- [6] Csörgő, S., Haeusler, E. and Mason, D.M. (1991). The asymptotic distribution of extreme sums. *Ann. Probab.* **19** 783–811. [MR1106286](#)
- [7] Csörgő, S., Horváth, L. and Mason, D.M. (1986). What portion of the sample makes a partial sum asymptotically stable or normal? *Probab. Theory Relat. Fields* **72** 1–16. [MR0835156](#)
- [8] Darling, D.A. (1952). The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.* **73** 95–107. [MR0048726](#)
- [9] Das, B. and Resnick, S.I. (2008). QQ plots, random sets and data from a heavy tailed distribution. *Stoch. Models* **24** 103–132. [MR2384693](#)

- [10] D'Auria, B. and Resnick, S.I. (2006). Data network models of burstiness. *Adv. in Appl. Probab.* **38** 373–404. [MR2264949](#)
- [11] Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds (with discussion and a reply by the authors). *J. Roy. Statist. Soc. Ser. B* **52** 393–442. [MR1086795](#)
- [12] de Haan, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes. Mathematical Centre Tracts* **32**. Amsterdam: Mathematisch Centrum. [MR0286156](#)
- [13] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering*. New York: Springer. [MR2234156](#)
- [14] de Haan, L. and Peng, L. (1998). Comparison of tail index estimators. *Statist. Neerlandica* **52** 60–70. [MR1615558](#)
- [15] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *J. Austral. Math. Soc. Ser. A* **61** 381–395. [MR1420345](#)
- [16] Doob, J.L. (1949). Heuristic approach to the Kolmogorov–Smirnov theorems. *Ann. Math. Statistics* **20** 393–403. [MR0030732](#)
- [17] Drees, H. (2011). Extreme value analysis of actuarial risks: Estimation and model validation. Available at <http://arxiv.org/abs/1103.2872>.
- [18] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance. Applications of Mathematics (New York)* **33**. Berlin: Springer. [MR1458613](#)
- [19] Flachsmeyer, J. (1963). Verschiedene Topologisierungen im Raum der abgeschlossenen Mengen. *Math. Nachr.* **26** 321–337. [MR0174026](#)
- [20] Geluk, J.L. and de Haan, L. (1987). *Regular Variation, Extensions and Tauberian Theorems. CWI Tract* **40**. Amsterdam: Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica. [MR0906871](#)
- [21] Gertensgarbe, F.W. and Werner, P.C. (1989). A method for the statistical definition of extreme-value regions and their application to meteorological time series. *Zeitschrift für Meteorologie* **39** 224–226.
- [22] Ghosh, S. and Resnick, S. (2010). A discussion on mean excess plots. *Stochastic Process. Appl.* **120** 1492–1517. [MR2653263](#)
- [23] Hogg, R.V. and Klugman, S.A. (1984). *Loss Distributions. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. New York: Wiley. [MR0747141](#)
- [24] Katz, R.W., Parlange, M.B. and Naveau, P. (2002). Statistics of extremes in hydrology. *Advances in Water Resources* **25** 1287–1304.
- [25] Kratz, M. and Resnick, S.I. (1996). The QQ-estimator and heavy tails. *Comm. Statist. Stochastic Models* **12** 699–724. [MR1410853](#)
- [26] Matheron, G. (1975). *Random Sets and Integral Geometry. Wiley Series in Probability and Mathematical Statistics*. New York–London–Sydney: Wiley. [MR0385969](#)
- [27] Maulik, K., Resnick, S. and Rootzén, H. (2002). Asymptotic independence and a network traffic model. *J. Appl. Probab.* **39** 671–699. [MR1938164](#)
- [28] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools. Princeton Series in Finance*. Princeton, NJ: Princeton Univ. Press. [MR2175089](#)
- [29] Molchanov, I. (2005). *Theory of Random Sets. Probability and Its Applications (New York)*. London: Springer. [MR2132405](#)
- [30] Øksendal, B. (2003). *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. *Universitext*. Berlin: Springer. [MR2001996](#)
- [31] Pötzelberger, K. and Wang, L. (2001). Boundary crossing probability for Brownian motion. *J. Appl. Probab.* **38** 152–164. [MR1816120](#)
- [32] Resnick, S.I. (1986). Point processes, regular variation and weak convergence. *Adv. in Appl. Probab.* **18** 66–138. [MR0827332](#)

- [33] Resnick, S.I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. New York: Springer. [MR2271424](#)
- [34] Resnick, S.I. (2008). *Extreme Values, Regular Variation and Point Processes*. Springer Series in Operations Research and Financial Engineering. New York: Springer. Reprint of the 1987 original. [MR2364939](#)
- [35] Rootzén, H. (2009). Weak convergence of the tail empirical process for dependent sequences. *Stochastic Process. Appl.* **119** 468–490. [MR2494000](#)
- [36] Rossberg, H.J. (1967). Über das asymptotische Verhalten der Rand- und Zentralglieder einer Variationsreihe. II. *Publ. Math. Debrecen* **14** 83–90. [MR0225454](#)
- [37] Seneta, E. (1976). *Regularly Varying Functions*. Lecture Notes in Mathematics **508**. Berlin: Springer. [MR0453936](#)
- [38] Shorack, G.R. and Wellner, J.A. (1986). *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: Wiley. [MR0838963](#)
- [39] Smith, R.L. (2003). Statistics of extremes, with applications in environment, insurance and finance. In *SemStat: Seminaire Europeen de Statistique, Extreme Values in Finance, Telecommunications, and the Environment* (B. Finkenstadt and H. Rootzén, eds.) 1–78. London: Chapman & Hall.
- [40] Stărică, C. (1999). Multivariate extremes for models with constant conditional correlations. *Journal of Empirical Finance* **6** 515–553.
- [41] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics*. Springer Series in Statistics. New York: Springer. [MR1385671](#)
- [42] Vervaat, W. (1997). Random upper semicontinuous functions and extremal processes. In *Probability and Lattices*. CWI Tract **110** 1–56. Amsterdam: Math. Centrum Centrum Wisk. Inform. [MR1465481](#)
- [43] Yang, G.L. (1978). Estimation of a biometric function. *Ann. Statist.* **6** 112–116. [MR0471233](#)

Received September 2010 and revised August 2011