



HAL
open science

A general estimator for the extreme value index: applications to conditional and heteroscedastic extremes

Laurent Gardes

► **To cite this version:**

Laurent Gardes. A general estimator for the extreme value index: applications to conditional and heteroscedastic extremes. 2015. hal-01075824v2

HAL Id: hal-01075824

<https://hal.science/hal-01075824v2>

Preprint submitted on 6 Jan 2015 (v2), last revised 3 Jun 2015 (v3)

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

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

A general estimator for the extreme value index: applications to conditional and heteroscedastic extremes

Laurent Gardes

Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes,
67084 Strasbourg cedex, France.

Abstract

It is well known that the tail behavior of a survival function is controlled by the so-called extreme value index. The aim of this paper is the estimation of this extreme value index in the case where the observations are not necessarily distributed from the same distribution. A general procedure of estimation is proposed. The idea is to estimate in a consistent way the survival function and to apply a general functional to obtain a consistent estimator for the extreme value index. The procedure of estimation presented in this paper permits to deal with a large set of models such as *conditional extremes* and *heteroscedastic extremes*. The consistency of the obtained estimator is established under general conditions and its finite sample behavior is investigated through a simulation study.

Keywords. Extreme value index, conditional extremes, heteroscedastic extremes, consistency.

AMS Subject Classifications. 62G05; 62G20; 62G30; 62G32.

1 Introduction

In various applications, the behavior of large values (instead of central values) of a random variable Z can be of high interest. For instance, in climatology, Z can represent the temperature or the amount of rain. The study of high values of Z is then a key point to understand the effect of global warming. In actuarial science, the random variable Z can model the claim size and it is of primary interest for insurance companies to estimate the probability of a large value of Z to be exceeded.

Denoting by $S(\cdot)$ the survival function of Z (*i.e.* for all $z \in \mathbb{R}$, $S(z) = \mathbb{P}(Z > z)$), the common departure point to make statistical inference on the tail distribution of $S(\cdot)$ is to assume that $S(\cdot)$ belongs to the maximum domain of attraction of an extreme value distribution (see Fisher and Tippett [14] and Gnedenko [23]). In other words, denoting by $S^{\leftarrow}(u) = \inf\{t, S(t) \leq u\}$ the right-continuous inverse of $S(\cdot)$, it is assumed that there exist a positive auxiliary function $a(\cdot)$ and a parameter $\gamma \in \mathbb{R}$ such that

$$\lim_{\alpha \rightarrow 0} \frac{S^{\leftarrow}(u\alpha) - S^{\leftarrow}(\alpha)}{a(\alpha^{-1})} - L_{\gamma}(1/u) \rightarrow 0, \quad (1)$$

for all $u \in (0, 1]$ where for all $v \geq 1$ and $s \in \mathbb{R}$, $L_s(v) = \int_1^v u^{s-1} du$. According to [26, Definition B.2.3], a survival function satisfying (1) is said to be of extended regular variation. In this paper, the set of extended regularly varying functions is denoted $\mathcal{ERV}(\gamma, a(\cdot))$.

The parameter $\gamma \in \mathbb{R}$ involved in condition (1) is called the extreme value index of $S(\cdot)$. This

parameter controls the decay of the tail distribution. If $\gamma > 0$, $S(\cdot)$ is called an heavy-tailed distribution and has a polynomial decay with an infinite right endpoint. At the opposite, the case $\gamma < 0$ corresponds to distributions with finite right endpoint. Finally, if $\gamma = 0$, $S(\cdot)$ is a light-tailed distribution and has an exponential decay. Obviously, the estimation of the extreme value index γ is often a major step to make statistical inference on the tail distribution in particular to estimate extreme quantile. For this reason, the estimation of the extreme value index has been widely studied in the literature.

1.1 Classical extreme value analysis

In the classical approach, it is assumed that one can observe n independent copies Z_1, \dots, Z_n of a random value Z with survival function $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$. In this framework, there exist numerous consistent estimators of γ , let us recall some of them. For heavy-tailed distributions (*i.e.* when $\gamma > 0$), the most notable estimator is probably the Hill's estimator [28] defined for a sequence $\alpha_n \rightarrow 0$ and such that $n\alpha_n \rightarrow \infty$ as n goes to infinity by

$$\mathcal{H}_n^{(1)}(\alpha_n) = \frac{1}{[n\alpha_n]} \sum_{i=1}^{[n\alpha_n]} \ln \frac{Z_{(n-i+1)}}{Z_{(n-[n\alpha_n])}},$$

where $[x]$ denotes the integer part of x and $Z_{(1)} \leq \dots \leq Z_{(n)}$ are the ordered statistics associated to the sample Z_1, \dots, Z_n . In the general case $\gamma \in \mathbb{R}$, Dekkers *et al.* [8] proposed to extend Hill's estimator by the so-called moment estimator given by

$$\hat{\gamma}_M(\alpha_n) = \mathcal{H}_n^{(1)}(\alpha_n) - 1 - \frac{1}{2} \left(1 - \frac{[\mathcal{H}_n^{(1)}(\alpha_n)]^2}{\mathcal{H}_n^{(2)}(\alpha_n)} \right)^{-1}.$$

One can also cite the estimators proposed by Pickands [32], by Csörgő *et al.* [3] among many others. Note that all the previous mentioned estimators and in fact most of the known estimators of the extreme value index only depends on the empirical tail quantile function $Q_n(\cdot)$ defined for $t \in [0, 1]$ as $Q_n(t) = \hat{S}_n^{\leftarrow}(\alpha_n t)$ with $\hat{S}_n(\cdot)$ the classical empirical estimator of $S(\cdot)$ (see Drees [12]).

1.2 Non classical extreme value analysis

In some situations, one can be interested in the study of the tail distribution of a survival function $S(\cdot)$ but without the possibility to observe an identically distributed (*i.i.d.*) sample from $S(\cdot)$. This problem arises in many different models, we list below some of them (the list is not exhaustive).

In insurance [1], finance [36], climatology [33] to name a few, the variable of interest Y can be often linked to a random covariate $X \in \mathbb{R}^p$. This covariate brings an important additional information on Y . For a fixed value $x_0 \in \mathbb{R}^p$, an important issue is the study of the conditional tail distribution of Y given $X = x_0$. For instance, Y can be the production level and X the quantity of labor (see Daouia *et al.* [4]) and thus a natural question is the estimation of the maximum production level that can be reached for a given quantity of labour. Assuming that the conditional survival function $S(\cdot|x_0)$ of Y given $X = x_0$ belongs to $\mathcal{ERV}(\gamma(x_0), \mathbf{a}(\cdot|x_0))$, a first step is thus the estimation of $\gamma(x_0)$. Unfortunately, it is often impossible to observe an *i.i.d.* sample from $S(\cdot|x_0)$. A more realistic situation is that one observe n independent copies $Z_1 = (X_1, Y_1), \dots, Z_n = (X_n, Y_n)$ of the random vector $Z = (X, Y)$ that must be used to estimate $\gamma(x_0)$. This situation is called in the literature *conditional extremes* and $\gamma(x_0)$ is referred to as the conditional extreme value index at point x_0 . The estimation of the conditional extreme value index has been considered for instance in [5] and [18] with an

estimator inspired from the Refined Pickands estimator proposed by Drees [10]. An adaptation of the moment estimator has been proposed in [24] and [35] and a maximum likelihood approach was considered by Wang and Tsai [38]. In the particular case of a positive conditional extreme value index, Gardes and Stupfler [19] propose an adaptation of the Hill estimator.

Let us also mention the situation where n independent observations $Z_1 = Y_{x_1}, \dots, Z_n = Y_{x_n}$ are recorded where, for all $x \in \mathbb{R}^p$, Y_x is drawn from the survival function $S_x(\cdot)$. Indices $\{x_1, \dots, x_n\}$ represent a deterministic additional information on the variable of interest (for instance the time, the geographical position, ...). Here, for a fixed value $x_0 \in \mathbb{R}^p$, it is assumed that $S_{x_0}(\cdot) \in \mathcal{ERV}(\gamma(x_0), \mathbf{a}_{x_0}(\cdot))$ and one wants to estimate $\gamma(x_0)$ from the non-identically distributed sample Y_{x_1}, \dots, Y_{x_n} . This situation is called in what follows *non identically distributed (i.d.) extremes* and is often considered to study extreme rainfalls at different geographical positions. Assuming that $\gamma(x_0) > 0$, this situation was considered for instance in [17] to estimate extreme rainfall return levels as a function of latitude and longitude and in [16] to study extreme daily flow of a river as a function of time. The estimation of $\gamma(x_0)$ has been addressed by many authors such as Davison and Smith [7], Smith [33], Chavez-Demoulin and Davison [2].

In [13], the authors consider the particular case where survival functions $S_{x_1}(\cdot), \dots, S_{x_n}(\cdot)$ are asymptotically proportional to a survival function $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ and are interested in the estimation of γ . This situation is named by the authors *heteroscedastic extremes*. Einmahl *et al.* [13] show, in the restricted case $\gamma > 0$, that the classical Hill's estimator is still consistent for γ . The model of *heteroscedastic extremes* permits to deal with data presenting a trend in extremes but with a constant shape parameter. A motivating example is the study of extreme daily loss returns of a given financial market where the magnitude of the temporal series, measured by the extreme value index, usually does not depend on time (see [13] for a real data set example).

Finally, one can also think on situations of right censored data or right truncated data where the variable of interest Y with survival function $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ is not fully observed but where the question of estimating γ is still of interest. Such observations can occur for instance in the analysis of lifetime data or reliability data. The estimation of γ under random censoring has been considered by Einmahl *et al.* [11]) and Gomes and Neves [25]. For truncated data, an estimator has been proposed by Gardes and Stupfler [20] in the case of heavy-tailed distributions ($\gamma > 0$).

1.3 Purpose of the paper

Of course, focusing on a particular model, it is always possible to propose specific estimators of γ but the estimation procedure strongly depends on the considered model. For example, the way of estimating γ under random censoring is completely different from the one used in the situation of *conditional extremes*. Up to our knowledge, no tentative to find a common procedure to estimate the extreme value index in a large range of situations (like the ones described in the previous paragraph for instance) has been proposed. This is the purpose of the present paper. More specifically, let $S(\cdot)$ be a survival function in the set $\mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ and let Z_1, \dots, Z_n be random variables non necessarily *i.i.d.* from $S(\cdot)$ but such that there exists a deterministic functional $Q(\cdot)$ for which $\hat{S}_n(\cdot) := Q(\cdot; Z_1, \dots, Z_n)$ is a right-continuous and non-increasing estimator of $S(\cdot)$. The main goal of the present paper is to define a unique and general functional $\mathcal{T}(\cdot)$ such that $\mathcal{T}(\hat{S}_n^{\leftarrow})$ is a consistent estimator of γ . In other words, our idea is to decouple the estimation of the survival function $S(\cdot)$ from that of the extreme value index γ . The advantage is that, for a given model, the estimation of the survival function

is often more easy than the one of γ . For instance, in the situation of *conditional extremes* one can easily think to use the kernel estimator introduced by Nadaraya [30] and Watson [39]. In presence of censored data, the survival function can be easily estimated by the Kaplan-Meier product limit estimator [29] and so on. The way of estimating the extreme value index presented in this paper can thus be used for a large set of models (not restricted the ones mentioned previously).

The rest of the paper is organized as follows. In Section 2, the expression of the general functional $\mathcal{T}(\cdot)$ is introduced and a consistency property on $\hat{S}_n(\cdot)$ is given in order to obtain a consistent estimator of γ . In section 3, our procedure of estimation is illustrated on the three models described briefly in the introduction: *conditional extremes*, *i.d. extremes* and *heteroscedastic extremes*. The finite sample behavior of the proposed estimator is illustrated through a simulation study in Section 4 and a short conclusion is given in Section 5. Section 6 is devoted to the proofs.

2 Estimation of the extreme value index

The framework considered in this paper is the following:

- (F) Let $S(\cdot)$ be a survival function belonging to the set $\mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ where $\gamma \in \mathbb{R}$ and $S(y) = 0$ for all $y < 0$ (i.e. $S(\cdot)$ is associated to a positive random variable). We observe a non necessarily *i.i.d.* sample Z_1, \dots, Z_n such that there exists a deterministic functional $\mathcal{Q}(\cdot)$ for which $\hat{S}_n(\cdot) := \mathcal{Q}(\cdot; Z_1, \dots, Z_n) \in \mathcal{D}(\mathbb{R}^+, [0, 1])$ is an estimator of $S(\cdot)$.

Here, for $E \subset \mathbb{R}$ and $F \subset \mathbb{R}$, $\mathcal{D}(E, F)$ is the set of non-increasing and right-continuous functions from E to F . Note that $\hat{S}_n^{\leftarrow}(\cdot)$, the right-continuous inverse of $S(\cdot)$, belongs to $\mathcal{D}([0, 1], \mathbb{R}^+)$. The simplest model entering in our framework is obviously the situation of an *i.d.d.* sample from $S(\cdot)$ taking for $\hat{S}_n^{\leftarrow}(\cdot)$ the classical empirical estimator. However, as mentioned in the introduction, situation (F) encompasses many others non classical models (*conditional extremes*, *heteroscedastic extremes*, censored or truncated data).

The aim of this section is first to give the definition of the functional $\mathcal{T} : \mathcal{D}([0, 1], \mathbb{R}^+) \mapsto \mathbb{R}$ used throughout this paper. Next, a consistency property on the statistic $\hat{S}_n(\cdot)$ is given in order to ensure the consistency of the extreme value index estimator $\mathcal{T}(\hat{S}_n^{\leftarrow})$.

2.1 Definition of the functional \mathcal{T}

Let $(\eta, \alpha) \in (0, 1)^2$, $\delta \in \mathbb{N}$ and $\varphi(\cdot)$ a positive and bounded function on $[\eta, 1]$. We first introduce the functional $\mathcal{T}^{(\delta)}(\cdot | \eta, \varphi) : \mathcal{D}([0, 1], \mathbb{R}^+) \mapsto \mathbb{R}^+$ defined for all $U(\cdot) \in \mathcal{D}([0, 1], \mathbb{R}^+)$ by

$$\mathcal{T}^{(\delta)}(U | \alpha, \eta, \varphi) := \int_{\eta}^1 \varphi(u) \left(\ln \frac{U(u\alpha)}{U(\alpha)} \right)^{\delta} du \Bigg/ \left(\int_{\eta}^1 \varphi(u) L_0(1/u) du \right)^{\delta},$$

where we recall that for all $v \geq 1$ and $s \in \mathbb{R}$, $L_s(v) = \int_1^v u^{s-1} du$. We also introduce the following function: for $s \leq 0$ let

$$\Psi_{\eta, \varphi}(s) := \left(\int_{\eta}^1 \varphi(u) L_s(1/u) du \right)^2 \Bigg/ \int_{\eta}^1 \varphi(u) L_s^2(1/u) du. \quad (2)$$

It is shown in Lemma 1 that $\Psi_{\eta, \varphi}(\cdot)$ is a decreasing function on $(-\infty, 0]$ and thus its inverse $\Psi_{\eta, \varphi}^{\leftarrow}(\cdot)$ is well defined. The functional $\mathcal{T}(\cdot | \alpha, \eta, \varphi)$ consider in this paper is given for all $U(\cdot) \in \mathcal{D}([0, 1], \mathbb{R}^+)$ by

$$\mathcal{T}(U | \alpha, \eta, \varphi) := \mathcal{T}^{(1)}(U | \alpha, \eta, \varphi) + \Psi_{\eta, \varphi}^{\leftarrow} \left(\max \left\{ \frac{[\mathcal{T}^{(1)}(U | \alpha, \eta, \varphi)]^2}{\mathcal{T}^{(2)}(U | \alpha, \eta, \varphi)}, \Psi_{\eta, \varphi}(0) \right\} \right). \quad (3)$$

Note that taking $\eta = 0$ and $\varphi(\cdot) = 1$ in (3), we found back the functional used to build the moment estimator introduced by Dekkers *et al.* [8] in the classical situation where an *i.i.d.* sample from $S(\cdot)$ is recorded.

2.2 Consistency

We now give a condition on $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ and a consistency property on the statistic $\hat{S}_n(\cdot)$ ensuring that there exist sequences τ_n and α_n converging to 0 such that $\mathcal{T}(\hat{S}_n^\leftarrow | \alpha_n, \eta, \varphi)$ is a consistent estimator of γ with a rate of convergence given by the sequence τ_n . Note that, roughly speaking, α_n represents the percentage of observations among $\{Z_1, \dots, Z_n\}$ used to compute the estimator. The choice of this sequence α_n (or equivalently the choice of $k_n := n\alpha_n$) is a key point in the estimation of the extreme value index. The optimal selection of this sequence is beyond the scope of this paper.

As usual in extreme value theory (see for instance [8], [24], [35]), second-order condition is required in order to precise the rate of convergence in (1). Note that the auxiliary function $a(\cdot)$ in (1) is such that $a(\alpha^{-1})/S^\leftarrow(\alpha) \rightarrow \gamma_+$ as $\alpha \rightarrow 0$ where $(\cdot)_+$ and $(\cdot)_-$ are respectively the positive and negative part functions (see for instance [15, Lemma 3.1]). Furthermore, it is shown in Lemma 2 that convergence (1) is in fact locally uniform. The second order condition used in this paper is given by

(A.1) there exists a closed interval $[\kappa_1, \kappa_2]$ with $0 < \kappa_1 < \eta$ and $\kappa_2 > 1$ such that

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \max \left\{ \sup_{u \in [\kappa_1, \kappa_2]} \left| \frac{S^\leftarrow(u\alpha_n) - S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} - L_\gamma(1/u) \right|, \left| \frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} - \gamma_+ \right| \right\} = 0.$$

Clearly, in order to estimate the extreme value index γ , only the behavior of the statistic $\hat{S}_n(z)$ for large values of z must be controlled. This is done by the following consistency property.

(A.2) For all sequences $y_n(u)$ such that $a^{-1}(\alpha_n^{-1})(y_n(u) - S^\leftarrow(u\alpha_n)) \rightarrow 0$ for all $u \in [\eta, 1]$ as n goes to infinity,

$$\tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n(y_n(u))}{S(y_n(u))} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1).$$

Roughly speaking, condition **(A.2)** means that $\hat{S}_n(y)$ is a uniformly consistent estimator of $S(y)$ on $[S^\leftarrow(\alpha_n), S^\leftarrow(\eta\alpha_n)]$. Note that the case $\eta = 0$ is not allowed since, most of the time, it is impossible to obtain the consistency uniformly on $[S^\leftarrow(\alpha_n), z_S^*]$ where z_S^* is the right endpoint of $S(\cdot)$. Our main result is given below.

Theorem 1. *Under framework **(F)**, if conditions **(A.1)** and **(A.2)** hold then*

$$\left| \mathcal{T}(\hat{S}_n^\leftarrow | \alpha_n, \eta, \varphi) - \gamma \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

The previous result provides an easy way to find a consistent estimator of γ for a large set of models (again, think for instance to the examples given in the introduction). The main advantage of our method is that it reduces the problem of estimating the extreme value index to the one of estimating the associated survival function which is often more simple.

In the next section, we focus on three particular models entering in our framework **(F)**: *conditional extremes*, *non i.d. extremes* and *heteroscedastic extremes*. In each situation, a natural estimator of $S(\cdot)$ satisfying **(A.2)** can be proposed and thus the estimation of the extreme value index can be achieved using our procedure.

3 Applications

3.1 Conditional extremes

This example takes place in a regression context where a positive response variable Y is associated to a random explanatory variable $X \in \mathbb{R}^p$. In what follows, we consider the following model:

- (M.1)** Let $Z_1 = (X_1, Y_1), \dots, Z_n = (X_n, Y_n)$ be n independent copies of a random vector $Z = (X, Y) \in \mathbb{R}^p \times [0, \infty)$. The probability density function of X is denoted by $g(\cdot)$. For a fixed value $x_0 \in \mathbb{R}^p$ such that $g(x_0) > 0$, we suppose that the conditional survival function $S(\cdot|x_0) = \mathbb{P}(Y > \cdot | X = x_0)$ of Y given $X = x_0$ belongs to $\mathcal{ERV}(\gamma(x_0), a(\cdot|x_0))$ where $\gamma(x_0) \in \mathbb{R}$.

The procedure described in the previous section is used to estimate the conditional extreme value index $\gamma(x_0)$ and thus only an estimation of $S(\cdot|x_0)$ is required. We suggest to use the kernel estimator introduced by Nadaraya [30] and Watson [39] and given for all $y \geq 0$ by

$$\hat{S}_n(y|x_0) = \frac{\sum_{i=1}^n \mathbb{I}_{\{Y_i > y\}} K_{H_n}(x_0 - X_i)}{\sum_{i=1}^n K_{H_n}(x_0 - X_i)}, \quad (4)$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function and H_n is a positive-definite matrix controlling the smoothness of the estimator. For the sake of shortness, we have introduced the notation $K_{H_n}(t) := |H_n|^{-1} K(H_n^{-1}t)$, $t \in \mathbb{R}^p$ where $K(\cdot)$ is called the kernel function and, for all square matrix M , $|M|$ denotes the determinant of M . To prove that the kernel estimator satisfies condition **(A.2)**, the following assumptions are introduced. The first one is a regularity assumption on $g(\cdot)$.

- (B.1)** For all $x \in \mathbb{R}^p$ and $x' \in \mathbb{R}^p$, there exists a constant $c_g > 0$ such that $|g(x) - g(x')| \leq c_g \|x - x'\|_\infty$.

Note that the uniform norm was used in condition **(B.1)** but obviously, any norm on \mathbb{R}^p can be also considered. The following condition on the kernel function $K(\cdot)$ is also required:

- (B.2)** $K(\cdot)$ is a bounded density on \mathbb{R}^p with support \mathcal{U}_p , the unit ball of \mathbb{R}^p .

This condition is classical in local estimation (see for instance [5, 24]). Finally, the following notation is required: for a positive-definite matrix M of size p and $c \in \mathbb{R}^p$, $B(c, M) := \{x \in \mathbb{R}^p \mid \|M^{-1}(c - x)\|_\infty \leq 1\}$ denotes the ball of center c and radius M . The next result is a direct application of Theorem 1. Its proof consists in showing that the kernel estimator $\hat{S}_n(\cdot|x_0)$ satisfy condition **(A.2)**.

Corollary 1. *Under model **(M.1)**, let (α_n) be a sequence converging to 0 such that $\sigma_n := (n|H_n|\alpha_n)^{-1/2} \rightarrow 0$ and $\sigma_n^{-1}(\ln(\sigma_n^{-1}))^{-1/2}\|H_n\|_\infty \rightarrow 0$. If condition **(A.1)** holds with sequences $\tau_n := \sigma_n(\ln(\sigma_n^{-1}))^{1/2}$ and α_n and if there exists $\delta > 1$ such that*

$$\sup \left\{ \left| \frac{S(y|x)}{S(y|x_0)} - 1 \right|, x \in B(x_0, H_n), \frac{y}{S^{\leftarrow}(\alpha_n|x_0)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (5)$$

then, under **(B.1)** and **(B.2)**,

$$\left| \mathcal{T} \left(\hat{S}_n^{\leftarrow}(\cdot|x_0) | \alpha_n, \eta, \varphi \right) - \gamma(x_0) \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Note that taking $H_n = h_n I_p$ and $K(\cdot) = \mathbb{I}_{\{\|\cdot\|_\infty \leq 1\}}$ in (4) where h_n is a positive sequence and I_p is the identity matrix of size p and choosing $\eta = 0$ and $\varphi(\cdot) = 1$ in the functional defined in (3) lead to the same estimator as the one proposed by Stupfler [35]. Hence, the expression of the estimator obtained using our procedure can be seen as a generalization of the Stupfler's estimator. A comparison of the two estimators is provided in the simulation

study (see Section 4).

Conditions used in the previous corollary are similar to the ones considered in [5] where the pointwise asymptotic normality of $\hat{S}_n(\cdot|x_0)$ and $\hat{S}_n^{\leftarrow}(\cdot|x_0)$ are established. The expected number of points kept for the estimation is given by $n|H_n|\alpha_n$ and thus conditions $\alpha_n \rightarrow 0$ and $n|H_n|\alpha_n \rightarrow \infty$ are classical in extreme value theory. Finally, condition (5) controls the oscillations of the conditional survival function $S(\cdot|x_0)$. An interesting discussion on this condition can be found in [35].

3.2 Non identically distributed extremes

In this paragraph, the following model is considered:

- (M.2)** Let E be a compact subset of \mathbb{R}^p , $p \in \mathbb{N}^*$. At points $\{x_i \in E, i = 1, \dots, n\}$, we observe positive and independent random variables $Z_1 = Y_{x_1}, \dots, Z_n = Y_{x_n}$ where for all $x \in E$, the survival function of Y_x is given by $S_x(\cdot)$. For a given $x_0 \in \overset{\circ}{E}$ where $\overset{\circ}{E}$ denotes the interior of E , it is assumed that $S_{x_0}(\cdot) \in \mathcal{ERV}(\gamma(\mathbf{x}_0), \mathbf{a}_{\mathbf{x}_0}(\cdot))$.

As mentioned in the introduction, $\{x_1, \dots, x_n\}$ can be seen as a deterministic additional information on the variable of interest. It can be for instance the time or the geographical position. Model **(M.2)** can also be interpreted as a regression model in the fixed design case. Here also, our goal is to use the procedure described in Section 2 to estimate $\gamma(x_0) \in \mathbb{R}$. For the estimation of the survival function $S_{x_0}(\cdot)$, we propose to use the estimator introduced by Stone [34] with Gasser and Müller's weights [21]. For a positive-definite matrix H_n it is given for all $z \geq 0$ by:

$$\hat{S}_{n,x_0}(z) = \sum_{i=1}^n \mathbb{I}_{\{Y_{x_i} > z\}} \int_{A_{n,i}} K_{H_n}(x_0 - t) dt, \quad (6)$$

where $A_{n,i}$ are sets that partition the subset E with $x_i \in A_{n,i}$ and where $K_{H_n}(\cdot)$ is defined as in paragraph 3.1. Estimator (6) seems natural under model **(M.2)** but others estimators for the survival function $S_{x_0}(\cdot)$ can also be considered. For instance, the Nadaraya-Watson estimator defined in (4) still can be used (by replacing X_i by x_i). One can also think on local polynomial estimators (see [37]). A discussion on the advantages and disadvantages of each estimator in the purpose of the estimation of γ would be interesting but it is beyond the scope of this paper.

To ensure that condition **(A.2)** is satisfied by the estimator $\hat{S}_{n,x_0}(\cdot)$, we suppose as before that the kernel function $K(\cdot)$ satisfy **(B.2)**. In addition, let us introduced the following notation: for $i \in \{1, \dots, n\}$, let $\mathcal{V}_{n,i}$ the volume of $A_{n,i}$ and let $\bar{\mathcal{V}}_n := \max(\mathcal{V}_{n,1}, \dots, \mathcal{V}_{n,n})$. It is assumed that

- (C)** There exists a positive constant $C_{\mathcal{V}}$ such that $n\bar{\mathcal{V}}_n \leq C_{\mathcal{V}}$.

Since E is a compact subset, this condition is reasonable and classical in nonparametric regression for the fixed design case (see for instance [22]). Asymptotic property of the extreme value index estimator $\mathcal{T}(\hat{S}_{n,x_0}^{\leftarrow}|\alpha_n, \eta, \varphi)$ is established in the next result.

Corollary 2. *Under model **(M.2)**, let (α_n) be a sequence converging to 0 and H_n a sequence of matrix converging to the zero matrix such that $\sigma_n := (n|H_n|\alpha_n)^{-1/2} \rightarrow 0$. If condition **(A.1)** hold with sequences $\tau_n := \sigma_n(\ln(\sigma_n^{-1}))^{1/2}$ and α_n and if there exist $\delta > 1$ such that*

$$\sup \left\{ \left| \frac{S_x(z)}{S_{x_0}(z)} - 1 \right|, x \in B(x_0, 2H_n), \frac{z}{S_{x_0}^{\leftarrow}(\alpha_n)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (7)$$

then, under **(B.2)** and **(C)**,

$$\left| \mathcal{T} \left(\hat{S}_{n,x_0}^{\leftarrow} | \alpha_n, \eta, \varphi \right) - \gamma(x_0) \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

The conditions in Corollary 2 are very similar to the ones of Corollary 1. The only difference is that condition $\sigma_n^{-1}(\ln(\sigma_n^{-1}))^{-1/2}\|H_n\|_\infty \rightarrow 0$ is not required here since this condition was used in Corollary 1 to ensure the consistency of the probability density estimator.

3.3 Heteroscedastic extremes

As in Einmahl *et al.* [13] and de Haan *et al.* [27], we consider the following model called *Heteroscedastic extremes*.

(M.3) Let $-\infty < b_1 < b_2 < \infty$. We observe at every points $\{x_i \in [b_1, b_2], i = 1, \dots, n\}$ positive and independent random variables $Z_1 = Y_{x_1}, \dots, Z_n = Y_{x_n}$. For all $x \in [b_1, b_2]$, the survival function of Y_x is denoted $S_x(\cdot)$. We assume that the right endpoint $z^* = S_x^{\leftarrow}(\cdot)$ does not depend on x and that there exist a survival function $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$ and a continuous positive function $c(\cdot)$ defined on $[b_1, b_2]$ such that:

$$\lim_{z \rightarrow z^*} \sup_{x \in [b_1, b_2]} \left| \frac{S_x(z)}{S(z)} - c(x) \right| = 0 \text{ with } \int_a^b c(s) ds < \infty. \quad (8)$$

Note that the survival function $S(\cdot)$ and the function $c(\cdot)$ in (8) are not uniquely defined. To ensure uniqueness, we impose from now on that $\int_{b_1}^{b_2} c(x) dx = b_2 - b_1$. Under **(M.3)**, it is easy to check that survival functions $S_x(\cdot)$, $x \in [b_1, b_2]$ share the same extreme value index γ even if these survival functions are different. More precisely, it is shown in [13] that $S_x(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{c}(x)\mathbf{a}(\cdot))$ for all $x \in [b_1, b_2]$. The function $c(\cdot)$ in (8) is called the *skedasis function*. As mentioned in [13], in the case $\gamma \neq 0$, the skedasis function changes the scale of extremes while in the case $\gamma = 0$, it only impacts the location of extreme. This model can thus be used to study data presenting a trend in extremes with a constant shape parameter. Let us highlight that in [13], the points $\{x_i, i = 1, \dots, n\}$ are assumed to be regularly distributed on $[0, 1]$ (*i.e.* $x_i = i/n$). This assumption can be too restrictive for an application purpose like for instance in hydrology since the times for which a certain non null amount of rain is observed are clearly not regularly distributed. In this paper, letting $x_0 = b_1$ and $x_{n+1} = b_2$, it is only assumed that

(D) there exists $d > 0$ such that $\max\{x_i - x_{i-1}, i = 1, \dots, n+1\} \leq d/n$.

Note that model **(M.3)** can be seen as particular case of **(M.2)** by taking $E = [b_1, b_2]$ and assuming that the survival functions $S_{x_1}(\cdot), \dots, S_{x_n}(\cdot)$ satisfy (8). As a consequence, for every $x_0 \in [b_1, b_2]$, the extreme value index in model **(M.3)** can be estimated by the statistic $\mathcal{T}(\hat{S}_{n,x_0}^{\leftarrow} | \alpha_n, \eta, \varphi)$ where $\hat{S}_{n,x_0}^{\leftarrow}(\cdot)$ is defined in (6). Nevertheless, this estimator is not the best one since under model **(M.3)**, one can use a global estimator of $S(\cdot)$ instead of the local estimator $\hat{S}_{n,x_0}^{\leftarrow}(\cdot)$. More specifically, we propose here to use the estimator

$$\hat{S}_n(z) = \sum_{i=1}^n \frac{x_i - x_{i-1}}{x_n - b_1} \mathbb{I}_{\{Y_{x_i} > z\}}. \quad (9)$$

This survival function estimator is global in the sense that the whole set of observations is used to estimate $S(\cdot)$. This estimator is used in our procedure to estimate γ . As a consequence of Theorem 1, the following result is established.

Corollary 3. *Under model (M.3), let α_n be a sequence converging to 0 such that $\sigma_n := (n\alpha_n)^{-1/2} \rightarrow 0$. If condition (A.1) hold with sequences $\tau_n := \sigma_n(\ln(\sigma_n^{-1}))^{1/2}$ and α_n , if*

$$\sup \{|c(u) - c(u')|, (u, u') \in [b_1, b_2]^2 \text{ with } |u - u'| \leq d/n\} = o(\tau_n). \quad (10)$$

and if there exists $\delta > 1$ such that

$$\sup \left\{ \left| \frac{S_x(z)}{S(z)} - c(x) \right|, x \in [b_1, b_2], \frac{z}{S^{\leftarrow}(\alpha_n)} \in [\delta^{-1}, \delta] \right\} = o(\tau_n), \quad (11)$$

then, under **(D)**,

$$\left| \mathcal{T} \left(\hat{S}_n^{\leftarrow} | \alpha_n, \eta, \varphi \right) - \gamma \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Our procedure thus provides an estimator of the extreme value index in the situation of heteroscedastic extremes. Up to our knowledge, the extreme value index estimation has been considered in this situation only in the paper of Einmahl *et al.* [13] where it is shown that, in the restricted case $\gamma > 0$, the classical Hill's estimator is still consistent.

Conditions (10) and (11) are very similar to the ones used in [13]. Condition (10) is a regularity condition on the function $c(\cdot)$ involved in model **(M.3)**. It is satisfied for instance if the function $c(\cdot)$ is Lipschitz continuous of order at least 1/2. Condition (10) is also a regularity condition but on the function $S_x(z)$ considered as a function of x for large values of z . The sequence α_n represents the proportion of largest observations used in the estimation procedure and thus $k_n := n\alpha_n$ is the number of kept observations. Conditions $\alpha_n \rightarrow 0$ and $n\alpha_n \rightarrow \infty$ (or equivalently $k_n/n \rightarrow 0$ and $k_n \rightarrow \infty$) are standard hypothesis for the estimation of the extreme value index.

4 Simulations

In this paper, a general procedure for the estimation of the extreme value index in a set of various models has been proposed. In order to appreciate the finite sample performance of estimators obtained with our procedure, we focus on two specific models: *heteroscedastic extremes* and *conditional extremes*. More precisely, the situation of *heteroscedastic extremes* is investigated by generating data with the following process:

P1 - For $i = 1, \dots, n$, let $x_i = G(i/n)$ where $G(\cdot)$ is the distribution function of a beta distribution with parameters $a = b = 2$. For a given function $\tilde{c} : [0, 1] \mapsto [1, \infty)$, we generate n independent random variables Y_{x_1}, \dots, Y_{x_n} where for all $x \in [0, 1]$, the survival function of Y_x is one of the three following:

P1-1: Heteroscedastic Fréchet distribution: for $z \geq 0$, $S_x(z) = 1 - \exp(-\tilde{c}(x)/z)$.

P1-2: Heteroscedastic uniform distribution: for $z \in [1 - 1/\tilde{c}(x), 1]$, $S_x(z) = (1 - z)\tilde{c}(x)$.

P1-3: Heteroscedastic Weibull distribution: for $z \geq \ln \tilde{c}(x)$, $S_x(z) = \exp[-(z - \ln \tilde{c}(x))^\theta]$, where the parameter θ belongs to $(0, 1]$.

The data generating by this process satisfy model **(M.3)** with $b_1 = 0$ and $b_2 = 1$ and where the common extreme value index and right endpoint are given by $\gamma = 1$ and $z^* = +\infty$ for **P1-1**, $\gamma = -1$ and $z^* = 1$ for **P1-2** and $\gamma = 0$ and $z^* = +\infty$ for **P1-3**. The function $c(\cdot)$ is given by $\tilde{c}(\cdot)/\int_0^1 c(s)ds$ for processes **P1-1** and **P1-2**. For process **P1-3**, it is easy to check that if $\theta \in (0, 1)$, $c(\cdot) = 1$ and if $\theta = 1$, $c(\cdot) = \tilde{c}(\cdot)/\int_0^1 c(s)ds$.

In the situation of *conditional extremes*, the following generating process for model **(M.2)** is considered.

P2 - Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be n independent copies of a random vector (X, Y) where X is uniformly distributed on $[0, 1]$ and, for a positive function $\gamma(\cdot)$ defined on $[0, 1]$ and a function $\tilde{c} : [0, 1] \mapsto [1, \infty)$, the conditional survival function $S(\cdot|x)$ of Y given $X = x$ is one of the three following:

P2-1 Conditional Fréchet distribution: for $y \geq 0$, $S(y|x) = 1 - \exp(-\tilde{c}(x)y^{-1/\gamma(x)})$.

P2-2: Conditional beta distribution: let $G(\cdot|x)$ be the distribution function of a beta distribution with parameters $a = b = -1/\gamma(x)$, for $y \in [G^{\leftarrow}(1 - 1/\tilde{c}(x)|x), 1]$, $S(y|x) = (1 - G(y|x))\tilde{c}(x)$.

P2-3: Conditional Weibull distribution: for a function $\theta : [0, 1] \mapsto (0, 1]$ and $y \geq \ln c(x)$, $S(y|x) = \exp \left[-(y - \ln \tilde{c}(x))^{\theta(x)} \right]$.

The rest of this section is organized as follows: first, the question of the choice of η and $\varphi(\cdot)$ in the functional $\mathcal{T}(\cdot|\alpha, \eta, \varphi)$ defined in (3) is investigated on data generated by process **P1**. Next, always for data generated by process **P1**, estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \tau, \varphi)$ where $\hat{S}_n(\cdot)$ is defined in (9) is compared to the classical moment estimator proposed by Dekkers *et al.* [8]. Recall that the consistency of moment estimator has been proved only in presence of independent and identically distributed random variables but, using similar techniques as in [13], the consistency must be also true for heteroscedastic extremes (the consistency was proved in [13] only in the situation $\gamma > 0$ for the Hill's estimator). Finally, for data generated by the process **P2**, estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}(\cdot|x_0)|\alpha_n, \tau, \varphi)$ where $\hat{S}_n(\cdot|x_0)$ is defined in (4) is compared to the estimator proposed by Stupfler [35].

4.1 Influence of parameter η and of function φ

The functional $\mathcal{T}(\cdot|\alpha, \eta, \varphi)$ defined in (3) depends on a parameter $\eta \in (0, 1)$ and a bounded function $\varphi(\cdot)$. It is natural to wonder about the impact of both η and $\varphi(\cdot)$ on our estimation procedure. We choose here to illustrate this impact under the model of *heteroscedastic extremes*. More precisely, $N = 500$ samples of size $n \in \{50, 100, 200, 400\}$ are generated under **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$ and, for **P1-3**, with $\theta = 1$. In this paragraph, the value of α_n is fixed to $n^{-1/3}$ and, using the survival estimator $\hat{S}_n(\cdot)$ defined in (9), the empirical mean squared error of the extreme value index estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \tau, \varphi)$ is computed for each value of n , for $\eta \in \{0.005 \times 2^j, j = 0, 1, \dots, 5\}$ and for the two functions $\varphi(\cdot) = 1$ and $\varphi(\cdot) = \ln(1/\cdot)$. Recall that for a given estimator $\hat{\gamma}_n$ of γ , denoting by $\hat{\gamma}_n^{(r)}$, $r = 1, \dots, N$ the values of $\hat{\gamma}_n$ observed on each replications, the empirical mean squared error of $\hat{\gamma}_n$ is given by:

$$\text{EMSE}(\hat{\gamma}_n) = \frac{1}{N} \sum_{r=1}^N (\hat{\gamma}_n^{(r)} - \gamma)^2.$$

Results are collected in Table 1 (for $\varphi(\cdot) = 1$) and Table 2 (for $\varphi(\cdot) = \ln(1/\cdot)$). The bold numbers are the best results obtained for each values of n . Concerning the function $\varphi(\cdot)$, one can see that the choice $\varphi(\cdot) = 1$ provides slightly better results for the Fréchet distribution (*i.e.* process **P1-1**) but, for the two other generating processes, the choice $\varphi(\cdot) = \ln(1/\cdot)$ is clearly better (especially for the uniform distribution (**P1-2**)). For the choice of η , it seems that taking η small provides better results (except for the Fréchet distribution with $\varphi(\cdot) = \ln(1/\cdot)$). Let us also highlight that taking $\eta = 0.02$ for each values of n and each generating process leads to mean squared errors quite close to bold numbers. In conclusion, the choice $\eta = 0.02$ and $\varphi(\cdot) = \ln(1/\cdot)$ seems reasonable and will be used in the rest of this simulation study.

4.2 Simulation under heteroscedastic extremes

We are now interested in the behavior of the estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \tau, \varphi)$ where $\hat{S}_n(\cdot)$ is defined in (9). We generate $N = 500$ samples of size $n = 100$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1$ and $\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$ and, for **P1-3**, with $\theta = 1$. Note that when $\tilde{c}(\cdot) = 1$, the observations are independent and identically distributed. In order to appreciate the effect of the sequence α_n , the estimator is computed for $\alpha_n \in \{n^{-1/a}, a = 2, \dots, 6\}$ while η is fixed to 0.02 and $\varphi(\cdot) = \ln(1/\cdot)$. Estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \tau, \varphi)$ is compared to the moment estimator $\hat{\gamma}_M(\alpha_n)$ by computing the ratio

$$R_1(\alpha_n) := \text{EMSE} \left(\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \tau, \varphi) \right) / \text{EMSE}(\hat{\gamma}_M(\alpha_n)),$$

for each values of α_n and each functions $\tilde{c}(\cdot)$. Clearly, a ratio lower than 1 means that our estimator provides better results (in term of mean squared error) than the moment estimator. Results are presented in Table 3. It appears that, for the Fréchet distribution (process **P1-1**), the moment estimator is slightly better than our estimator but for the two other generating processes, our method provides better results in term of mean squared error. Taking into account results collected in Table 1 and Table 2, it seems that the difference between our estimator and the moment estimator is mainly explained by the use of the function $\varphi(\cdot) = \ln(1/\cdot)$ (roughly speaking the moment estimator corresponds to the case $\eta = 0$ and $\varphi(\cdot) = 1$). One can also notice that the function $\tilde{c}(\cdot)$ has not a strong influence on the estimation of γ .

4.3 Simulation under conditional extremes

Finally, under a *conditional extremes* model, the behavior of estimator $\mathcal{T}(\hat{S}_n^{\leftarrow}(\cdot|x_0)|\alpha_n, \eta, \varphi)$ where $\hat{S}_n(\cdot|x_0)$ is given in (4) is investigated by generating $N = 500$ samples of size $n = 500$ using process **P2**. The function $\tilde{c}(\cdot)$ is taken equal to $\ln(1/\cdot) + 1$ and the conditional extreme value index is given by $\gamma(x) = 2/3 + 1/3 \sin(2\pi x)$ for all $x \in [0, 1]$. For the process **P2-3**, we choose $\theta(\cdot) = \gamma(\cdot)$. As before, the value of η is fixed to 0.02 and $\varphi(\cdot) = \ln(1/\cdot)$. The couple of sequences (α_n, H_n) required to compute the survival function estimator $\hat{S}_n(\cdot|x_0)$ are picked in the set $\{(n^{-1/i}, n^{(1-i)/(ij)}), i, j = 2, \dots, 6\}$. Not that taking (α_n, H_n) in this set ensures that $\ln(nH_n\alpha_n) = (i-1)(j-1)/(ij) \ln(n) \rightarrow \infty$ as $n \rightarrow \infty$. The value of x_0 is fixed to 1/4 for which the maximum of the function $\gamma(\cdot)$ is reached ($\gamma(1/4) = 1$).

We compare $\mathcal{T}(\hat{S}_n^{\leftarrow}(\cdot|x_0)|\alpha_n, \eta, \varphi)$ to the estimator proposed by Stupfler [35] defined as follows. Let $(X_1, Y_1), \dots, (X_n, Y_n)$ be random variables drawn from model **(M.1)**. For $\delta \in \mathbb{N}$, let $k_n(x_0) := \lfloor \alpha_n M(x_0, H_n) \rfloor$ and

$$\mathcal{H}_{n,S}^{(\delta)}(\alpha_n|x_0) := \frac{1}{k_n(x_0)} \sum_{i=1}^{k_n(x_0)} \left(\ln \frac{Y_{(M(x_0, H_n) - i + 1)}^*(x_0)}{Y_{(M(x_0, H_n) - k_n(x_0))}^*(x_0)} \right)^\delta,$$

where

$$M(x_0, H_n) := \sum_{i=1}^n \mathbb{1}_{\{\|X_i - x_0\|_\infty \leq H_n\}},$$

and, given $M(x_0, H_n) = q$, $Y_1^*(x_0), \dots, Y_q^*(x_0)$ are the response variables Y_i whose associated covariate X_i is such that $\|X_i - x_0\|_\infty \leq H_n$. As previously mentioned, $Y_{(1)}^*(x_0) \leq \dots \leq Y_{(q)}^*(x_0)$ are the associated ordered statistics. The estimator introduced in [35] is defined by:

$$\hat{\gamma}_S(\alpha_n, H_n|x_0) := \mathcal{H}_{n,S}^{(1)}(\alpha_n|x_0) - 1 - \frac{1}{2} \left(1 - \frac{[\mathcal{H}_{n,S}^{(1)}(\alpha_n|x_0)]^2}{\mathcal{H}_{n,S}^{(2)}(\alpha_n|x_0)} \right)^{-1}.$$

To make the comparison, we compute for each values of α_n and H_n the ratio

$$R_2(\alpha_n, H_n) := \text{EMSE} \left(\mathcal{T}(\hat{S}_n^{\leftarrow}(\cdot|x_0)|\alpha_n, \eta, \varphi) \right) / \text{EMSE} (\hat{\gamma}_S(\alpha_n, H_n|x_0)).$$

The results are gathered in Table 4. For the conditional Fréchet and Weibull distributions (processes **P2-1** and **P2-3**), the estimator $\hat{\gamma}_S(\alpha_n, H_n|x_0)$ provides, for some couples (α_n, H_n) slightly better results than our estimator. For the conditional beta distribution, our estimator is clearly better for all couples (α_n, H_n) .

5 Conclusion

In this paper, a general procedure to estimate the extreme value index γ associated to a survival function $S(\cdot)$ was proposed. It can be used for a large set of models where observations are not necessarily distributed from $S(\cdot)$. From a theoretical point of view, this paper offers an

easy way to establish the consistency of estimators obtained through our procedure. Three different models have been specifically considered here and part of our future work is to use our procedure to deal with the case of censored and/or truncated data. Another possible pursuit of this work is to establish the asymptotic normality of the estimator and treat the problem of extreme quantile estimation.

6 Proofs

6.1 Preliminary results

The first lemma is dedicated to the function $\Psi_\eta(\cdot)$ defined in (2) and its derivative. Introducing the integral,

$$I_\eta^{(\delta)}(s, t) := \int_\eta^1 \varphi(u) L_s^\delta(1/u) u^{-t} du < \infty,$$

where $(s, t) \in [0, \infty)^2$ and $\delta \in \mathbb{N}$, one can write $\Psi_\eta(s) = [I_\eta^{(1)}(s, 0)]^2 / I_\eta^{(2)}(s, 0)$.

Lemma 1. *For all $\eta \in (0, 1)$, the function $\Psi_\eta(\cdot)$ is decreasing on $(-\infty, 0]$ with $\Psi_\eta(s) \rightarrow I_\eta^{(0)}(0, 0)$ as $s \rightarrow -\infty$. Furthermore, the function $\Psi_\eta(\cdot)$ is continuously differentiable on $(-\infty, 0]$ with $\Psi'_\eta(s) \rightarrow 0^-$ as $s \rightarrow -\infty$ and*

$$\lim_{s \rightarrow 0^-} \Psi'_\eta(s) = \frac{[I_\eta^{(2)}(0, 0)]^2 I_\eta^{(1)}(0, 0) - [I_\eta^{(1)}(0, 0)]^2 I_\eta^{(3)}(0, 0)}{[I_\eta^{(2)}(0, 0)]^2}.$$

Proof – First, remark that

$$\Psi_\eta(s) = \left(\int_\eta^1 \varphi(u) s L_s(1/u) du \right)^2 / \int_\eta^1 \varphi(u) (s L_s(1/u))^2 du.$$

Since $\varphi(\cdot)$ is a positive bounded function, $0 \leq -s L_s(1/u) \leq 1$ for all $s \geq 0$ and $s L_s(1/u) \rightarrow -1$ as $s \rightarrow -\infty$, the dominated convergence entails that

$$I_\eta^{(\delta)}(s, 0) = -\frac{1}{s^\delta} \int_\eta^1 \varphi(u) du (1 + o(1)), \quad (12)$$

as $s \rightarrow 0$ and hence, $\Psi_\eta(s) \rightarrow I_\eta^{(0)}(0, 0)$ as s goes to zero.

We now compute the derivative of the function $\Psi_\eta(\cdot)$. Since $\varphi(\cdot)$ is a positive bounded function, $0 \leq L_s(1/u) \leq \ln(1/\eta)$ and $0 \leq u^{-s} \leq 1$ for all $s \geq 0$, the functions $I_\eta^{(\delta)}(s, 0)$ and $I_\eta^{(\delta)}(0, s)$ are continuous on $s \in (-\infty, 0]$ and one can interchange the derivative and the integral sign. Hence, for $s < 0$,

$$A'(s) := \frac{d}{ds} I_\eta^{(1)}(s, 0) = \frac{1}{s} \left(I_\eta^{(1)}(0, s) - I_\eta^{(1)}(s, 0) \right) = \int_\eta^1 \frac{\varphi(u)}{s} (\ln(1/u) u^{-s} - L_s(1/u)) du. \quad (13)$$

Furthermore, remarking that

$$I_\eta^{(2)}(s, 0) = \frac{2}{s} \left(I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(s, 0) \right),$$

one has for $s < 0$;

$$\begin{aligned} B'(s) &= \frac{d}{ds} I_\eta^{(2)}(s, 0) = \frac{2}{s^2} \left(I_\eta^{(1)}(0, 2s) - I_\eta^{(1)}(0, s) + 2 \left(I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(2s, 0) \right) \right) \\ &= \int_\eta^1 \frac{2\varphi(u)}{s} L_s(1/u) [\ln(1/u) u^{-s} - L_s(1/u)] du. \end{aligned} \quad (14)$$

Since the functions $A'(\cdot)$ and $B'(\cdot)$ are continuous, the derivative

$$\Psi'_\eta(s) = \frac{2A'(s)I_\eta^{(1)}(s, 0)I_\eta^{(2)}(s, 0) - [I_\eta^{(1)}(s, 0)]^2 B'(s)}{[I_\eta^{(2)}(s, 0)]^2},$$

is also a continuous function on $s \in (-\infty, 0)$.

We now compute the limit of $\Psi'_\eta(s)$ as $s \rightarrow 0$. First, we focus on the function $A'(\cdot)$. Since for all $u \in (0, 1]$, $L_s(1/u)$ is a non-decreasing function in $s \in \mathbb{R}$, the following inequalities hold:

$$L_s(1/u) \leq \ln(1/u) \text{ if } s \leq 0 \text{ and } L_s(1/u) \geq \ln(1/u) \text{ if } s \geq 0. \quad (15)$$

A straightforward consequence is that for all $s < 0$, $0 \leq [\ln(1/u)u^{-s} - L_s(1/u)]/s \leq \ln^2(1/u)$. Remarking that $\ln(1/u)u^{-s} - L_s(1/u) \rightarrow \ln^2(1/u)/2$ as $s \rightarrow 0$, the dominated convergence theorem lead to

$$\lim_{s \rightarrow 0} A'(s) = \frac{1}{2} \int_\eta^1 \varphi(u) \ln^2(1/u) du = \frac{1}{2} I_\eta^{(2)}(0, 0). \quad (16)$$

We now focus on the function $B'(\cdot)$. Using the inequalities (15) leads to

$$0 \leq \frac{1}{s} L_s(1/u) [\ln(1/u)u^{-s} - L_s(1/u)] \leq \ln^3(1/u),$$

for all $s < 0$. Remarking that $L_s(1/u) [\ln(1/u)u^{-s} - L_s(1/u)]/s \rightarrow \ln^3(1/u)/2$ as $s \rightarrow 0$ and using the dominated convergence theorem leads to

$$\lim_{s \rightarrow 0} B'(s) = \int_\eta^1 \varphi(u) \ln^3(1/u) du = I_\eta^{(3)}(0, 0). \quad (17)$$

Collecting (16) and (17) and since $I_\eta^{(\delta)}(s, 0)$ and $I_\eta^{(\delta)}(0, s)$ are continuous on $(-\infty, 0]$, the limit of Ψ'_η as $s \rightarrow 0$ is given by $\{[I_\eta^{(2)}(0, 0)]^2 I_\eta^{(1)}(0, 0) - [I_\eta^{(1)}(0, 0)]^2 I_\eta^{(3)}(0, 0)\} / [I_\eta^{(2)}(0, 0)]^2$.

We are now interested in the limit of the derivative of $\Psi_\eta(\cdot)$ as $s \rightarrow -\infty$. Collecting (13) and (14),

$$\Psi'_\eta(s) = \frac{I_\eta^{(1)}(s, 0) \left(2I_\eta^{(1)}(0, s) I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(0, s) I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(0, 2s) I_\eta^{(1)}(s, 0) \right)}{2 \left(I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(s, 0) \right)^2}, \quad (18)$$

for $s < 0$. Using (12) and remarking that $I_\eta^{(1)}(0, s) \rightarrow 0$ as $s \rightarrow -\infty$ shows that $\Psi'_\eta(s) \rightarrow 0$ as $s \rightarrow -\infty$.

It remains to prove that $\psi'_\eta(s) < 0$ for all $s < 0$ or equivalently, in view of (18) that

$$\Delta_\eta(s) := 2I_\eta^{(1)}(0, s) I_\eta^{(1)}(2s, 0) - I_\eta^{(1)}(0, s) I_\eta^{(1)}(s, 0) - I_\eta^{(1)}(0, 2s) I_\eta^{(1)}(s, 0) < 0,$$

since $I_\eta^{(1)}(s, 0) > 0$ for $s \leq 0$. Furthermore, using the fact that $2L_{2s}(1/u) = L_s(1/u)(u^{-s} + 1)$, one has

$$\frac{I_\eta^{(1)}(2s, 0)}{I_\eta^{(1)}(s, 0)} = \frac{1}{2} \left(1 + \frac{I_\eta^{(1)}(s, s)}{I_\eta^{(1)}(s, 0)} \right),$$

Hence,

$$\begin{aligned} \Delta_\eta(s) &= I_\eta^{(1)}(s, 0) I_\eta^{(1)}(0, s) \left(2 \frac{I_\eta^{(1)}(2s, 0)}{I_\eta^{(1)}(s, 0)} - 1 - \frac{I_\eta^{(1)}(0, 2s)}{I_\eta^{(1)}(0, s)} \right) \\ &= I_\eta^{(1)}(s, s) I_\eta^{(1)}(0, s) - I_\eta^{(1)}(s, 0) I_\eta^{(1)}(0, 2s) \\ &= \int_\eta^1 \int_\eta^1 \varphi(u) \varphi(v) L_s(1/u) v^{-s} \ln(1/v) (u^{-s} - v^{-s}) dv du \end{aligned}$$

Next, using the decomposition $\int_\eta^1 = \int_\eta^u + \int_u^1$ and Fubini's Theorem (which can be applied since the involved function is of constant sign),

$$\Delta_\eta(s) = \int_\eta^1 \int_\eta^u \varphi(u) \varphi(v) (u^{-s} - v^{-s}) \tilde{L}(u, v) dv du,$$

where $\tilde{L}(u, v) := L_s(1/u)v^{-s} \ln(1/v) - L_s(1/v)u^{-s} \ln(1/u)$. Remark now that for $v \in (\eta, u)$ and $s < 0$,

$$\begin{aligned} \frac{\partial}{\partial v} \tilde{L}(u, v) &= v^{-(s+1)} (u^{-s} \ln(1/u) - L_s(1/u)(s \ln(1/v) + 1)) \\ &\geq v^{-(s+1)} (\ln(1/u) - L_s(1/u)) > 0. \end{aligned}$$

Hence, for all $u \in (\eta, 1)$ and $v \in (\eta, u)$, $\tilde{L}(u, v) < \tilde{L}(u, u) = 0$ and the proof is complete since $\varphi(u)\varphi(v)(u^{-s} - v^{-s}) > 0$ when $u \in (\eta, 1)$ and $v \in (\eta, u)$. \blacksquare

The next two lemmas are general results on extended regular varying functions. The first result shows that the convergence characterizing a function of extended regular variation is locally uniform.

Lemma 2. *Let $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$. For all $0 < \kappa_1 < \kappa_2 < \infty$,*

$$\lim_{\alpha \rightarrow 0} \sup_{u \in [\kappa_1, \kappa_2]} \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a(\alpha^{-1})} - L_\gamma(1/u) \right| = 0.$$

Proof of Lemma 2 – From [26, Theorem B.2.18], for all $\varepsilon \in (0, 1)$, there exists $\alpha_0(\varepsilon)$ such that for all $\alpha < \alpha_0(\varepsilon)$ and all $u \in [\kappa_1, \kappa_2]$,

$$\Delta_{S^\leftarrow}(\alpha, u) := \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a_0(\alpha^{-1})} - L_\gamma(1/u) \right| \leq \frac{\varepsilon}{\kappa_1} \max\{\kappa_1^{-\gamma}, \kappa_2^{-\gamma}\},$$

where for $y \geq 0$,

$$a_0(y) = \begin{cases} \gamma S^\leftarrow(1/y) & \text{if } \gamma > 0, \\ -\gamma(S^\leftarrow(0) - S^\leftarrow(1/y)) & \text{if } \gamma < 0, \\ S^\leftarrow(1/y) - y^{-1} \int_0^y S^\leftarrow(1/s) ds & \text{if } \gamma = 0. \end{cases}$$

Clearly, $a(y)/a_0(y) \rightarrow 1$ as y goes to infinity. Hence, there exists $\alpha_1(\varepsilon)$ such that for $\alpha < \alpha_1(\varepsilon)$, $|1 - a(\alpha^{-1})/a_0(\alpha^{-1})| \leq \varepsilon$. For $\alpha < \alpha_0(\varepsilon) \wedge \alpha_1(\varepsilon)$, we thus have the inequality

$$\begin{aligned} \Delta_{S^\leftarrow}(\alpha, u) &\leq \frac{a_0(\alpha^{-1})}{a(\alpha^{-1})} \left| \frac{S^\leftarrow(u\alpha) - S^\leftarrow(\alpha)}{a_0(\alpha^{-1})} - L_\gamma(1/u) \right| + |L_\gamma(1/u)| \left| 1 - \frac{a(\alpha^{-1})}{a_0(\alpha^{-1})} \right| \\ &\leq (1 + \varepsilon) \frac{\varepsilon}{\kappa_1} \max\{\kappa_1^{-\gamma}, \kappa_2^{-\gamma}\} + L_\gamma(1/\kappa_1) \varepsilon \leq \left(\frac{2 \max\{\kappa_1^{-\gamma}, \kappa_2^{-\gamma}\}}{\kappa_1} + L_\gamma(1/\kappa_1) \right) \varepsilon, \end{aligned}$$

which concludes the proof. \blacksquare

This second result provides equivalent conditions to the second order condition **(A.1)**.

Lemma 3. *If there exist positive sequences α_n and τ_n converging to 0 as $n \rightarrow \infty$ such that the survival function $S(\cdot)$ satisfies **(A.1)** then,*

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \sup_{u \in [\kappa_1, \kappa_2]} \left| \frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} - L_{\gamma_-}(1/u) \right| = 0, \quad (19)$$

and

$$\lim_{n \rightarrow \infty} \tau_n^{-1} \sup_{v \in [L_\gamma(1/\kappa_2), L_\gamma(1/\kappa_1)]} \left| \frac{\alpha_n}{S(S^\leftarrow(\alpha_n) + v a(\alpha_n^{-1}))} - L_\gamma^\leftarrow(v) \right| = 0. \quad (20)$$

Proof of Lemma 3 – We first focus on (19). Let us introduce the notations

$$\Delta_n(u) := \frac{S^\leftarrow(u\alpha_n) - S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} - L_\gamma(1/u) \text{ and } \bar{\Delta}_n := \sup_{u \in [\kappa_1, \kappa_2]} |\Delta_n(u)|.$$

We start with the following equality:

$$\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} = \frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \left[1 + \frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right].$$

We consider the case $\gamma > 0$. A straightforward calculus leads to

$$\ln \left[1 + \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right] - \ln(u^{-\gamma}) = \ln(1 + D_{n,1}(u)),$$

where

$$\begin{aligned} |D_{n,1}(u)| &= u^\gamma \left| \gamma \Delta_n(u) + \left(\frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} - \gamma \right) (L_\gamma(1/u) + \Delta_n(u)) \right| \\ &\leq \max\{\kappa_1^\gamma, \kappa_2^\gamma\} \left\{ \gamma \bar{\Delta}_n + \left| \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} - \gamma \right| (L_\gamma(1/\kappa_1) + \bar{\Delta}_n) \right\} \rightarrow 0. \end{aligned}$$

Hence, using inequality $|\ln(1+x)| \leq 3|x|/2$ for $x \in [-1/2, 1/2]$, one has for n large enough and $\gamma > 0$

$$\begin{aligned} \tau_n^{-1} \sup_{u \in [\kappa_1, \kappa_2]} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - \ln(1/u) \right| &\leq \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \frac{3\tau_n^{-1}}{2} \sup_{u \in [\kappa_1, \kappa_2]} |D_{n,1}(u)| \\ &+ \ln(1/\kappa_1) \tau_n^{-1} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \gamma - 1 \right|, \quad (21) \end{aligned}$$

which converges to 0 by assumption. Now, assume that $\gamma \leq 0$, since

$$|D_{n,2}(u)| := \left| \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/u) + \Delta_n(u)) \right| \leq \frac{a(\alpha_n^{-1})}{S^{\leftarrow}(\alpha_n)} (L_\gamma(1/\kappa_1) + \bar{\Delta}_n) \rightarrow 0$$

and using the inequality $x(1-x) \leq \ln(1+x) \leq x$ for $x \in [-1/2, 1/2]$, one has:

$$-L_\gamma(1/u)D_{n,2}(u) + \Delta_n(u)(1 - D_{n,2}(u)) \leq \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - L_\gamma(1/u) \leq \Delta_n(u).$$

Since $\tau_n^{-1} \sup |D_{n,2}(u)| \rightarrow 0$ and $\tau_n^{-1} \bar{\Delta}_n \rightarrow 0$, it is clear that

$$\tau_n^{-1} \sup_{u \in [\kappa_1, \kappa_2]} \left| \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - L_\gamma(1/u) \right| \rightarrow 0. \quad (22)$$

Collecting (21) and (22) conclude the proof of (19). The proof of equation (20) is a direct consequence of Vervaat's Lemma (see [26, Lemma A.0.2]) applied to

$$x_n(s) = \frac{S^{\leftarrow}(\alpha_n/s) - S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \text{ and } g(s) = L_\gamma(s),$$

with $\delta_n = \tau_n$. ■

The following lemma is a technical result that will be useful in the proof of Lemma 5. A proof of this result can be found in [20, Lemma 6].

Lemma 4. *Let (X_n) be a sequence of positive real-valued random variables such that for every positive nonrandom sequence δ_n converging to 0, the random sequence $\delta_n X_n$ converges to 0 in probability. Then $X_n = \mathcal{O}_{\mathbb{P}}(1)$.*

The next result takes place in our framework **(F)**. It shows that if for large values of y , $\hat{S}_n(y)$ is a consistent estimator of $S(y)$ then $\hat{S}_n^{\leftarrow}(\alpha)$ is also a consistent estimator of $S^{\leftarrow}(\alpha)$ for small values of α . This result is a cornerstone in the proof of Theorem 1.

Lemma 5. *Under **(F)**, let (α_n) and (τ_n) be sequences converging to 0 as $n \rightarrow \infty$ and assume that the survival function $S(\cdot)$ satisfies condition **(A.1)**. If for all sequences $y_n(u)$ such that $a^{-1}(\alpha_n^{-1})(y_n(u) - S^{\leftarrow}(u\alpha_n)) \rightarrow 0$ for all $u \in [\eta, 1]$, one has for all $\eta \in (0, 1)$,*

$$\tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n(y_n(u))}{S(y_n(u))} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1),$$

then

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(u\alpha_n)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1). \quad (23)$$

From equation (20) in Lemma 3, it is easy to see that uniformly on $u \in [\eta, 1]$, $S(y_n(u)) = u\alpha_n(1 + o(1)) \geq \eta\alpha_n/2 > 0$ for n large enough. Hence, the division by $S(y_n(u))$ is allowed at least for n large enough.

Proof of Lemma 5 – Let us introduce the sequence $m_n := \lfloor \tau_n^{-1} \rfloor$. For $j = 1, \dots, m_n$, let $\theta_n(j) := \eta + (j-1)(1-\eta)/(m_n-1)$. Clearly, for all $u \in [\eta, 1]$, there exists $j_u \in \{1, \dots, m_n-1\}$ such that $\theta_n(j_u) \leq u \leq \theta_n(j_u+1)$. Since $S^\leftarrow(\cdot)$ and $\hat{S}_n^\leftarrow(\cdot)$ are non-increasing and right-continuous functions, it is easy to check that for all $u \in [\eta, 1]$,

$$\begin{aligned} \left| \hat{S}_n^\leftarrow(u\alpha_n) - S^\leftarrow(u\alpha_n) \right| &\leq 2 \left| \hat{S}_n^\leftarrow(\theta_n(j_u)\alpha_n) - S^\leftarrow(\theta_n(j_u)\alpha_n) \right| \\ &\quad + \left| \hat{S}_n^\leftarrow(\theta_n(j_u+1)\alpha_n) - S(\theta_n(j_u+1)\alpha_n) \right| \\ &\quad + 2(S^\leftarrow(\theta_n(j_u)\alpha_n) - S^\leftarrow(\theta_n(j_u+1)\alpha_n)). \end{aligned}$$

Hence,

$$\sup_{u \in [\eta, 1]} \left| \hat{S}_n^\leftarrow(u\alpha_n) - S^\leftarrow(u\alpha_n) \right| \leq 3(T_{n,1} + T_{n,2}),$$

with

$$T_{n,1} := \max_{j=1, \dots, m_n-1} [S^\leftarrow(\theta_n(j)\alpha_n) - S^\leftarrow(\theta_n(j+1)\alpha_n)],$$

and

$$T_{n,2} := \max_{j=1, \dots, m_n} \left| \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) - S^\leftarrow(\theta_n(j)\alpha_n) \right|.$$

Let us first focus on the term $T_{n,1}$. Since $S(\cdot)$ satisfies condition **(A.1)**, a straightforward calculus entails that

$$S^\leftarrow(\theta_n(j)\alpha_n) - S^\leftarrow(\theta_n(j+1)\alpha_n) = a(\alpha_n^{-1}) [L_\gamma(1/\theta_n(j)) - L_\gamma(1/\theta_n(j+1))] + o(\tau_n),$$

where the term $o(\tau_n)$ converges to 0 uniformly on $u \in [\eta, 1]$.

Since the derivative of the function $L_\gamma(1/\cdot)$ is bounded on $[\eta, 1]$, a first order Taylor expansion leads to $S^\leftarrow(\theta_n(j)\alpha_n) - S^\leftarrow(\theta_n(j+1)\alpha_n) = a(\alpha_n^{-1}) (\mathcal{O}(m_n^{-1}) + o(\tau_n))$, uniformly on $u \in [\eta, 1]$ and thus

$$T_{n,1} = a(\alpha_n^{-1}) \mathcal{O}(\tau_n). \quad (24)$$

Let us now consider the term $T_{n,2}$. Our goal is to show that

$$\frac{\tau_n^{-1}}{a(\alpha_n^{-1})} T_{n,2} = \mathcal{O}_{\mathbb{P}}(1). \quad (25)$$

To this end, it suffices, from Lemma 4, to show that for every sequence $\delta_n \rightarrow 0$ and for every $\varepsilon > 0$,

$$p_n(\varepsilon) := \mathbb{P} \left[\frac{\delta_n \tau_n^{-1}}{a(\alpha_n^{-1})} \max_{j=1, \dots, m_n} \left| \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) - S^\leftarrow(\theta_n(j)\alpha_n) \right| > \varepsilon \right] \rightarrow 0.$$

Note that it is sufficient to consider sequences $\delta_n \rightarrow 0$ such that $\delta_n^{-1} \tau_n \rightarrow 0$ as n goes to infinity. Introducing the sequence

$$y_n^\pm(\theta_n(j)) := S^\leftarrow(\theta_n(j)\alpha_n) \pm \varepsilon \tau_n \frac{a(\alpha_n^{-1})}{\delta_n},$$

one has

$$\begin{aligned} p_n(\varepsilon) &\leq \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) > y_n^+(\theta_n(j)) \right\} \cup \left\{ \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) \leq y_n^-(\theta_n(j)) \right\} \right] \\ &\leq 1 - \mathbb{P} \left[\bigcap_{j=1}^{m_n} \left\{ \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) \leq y_n^+(\theta_n(j)) \right\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ \hat{S}_n^\leftarrow(\theta_n(j)\alpha_n) \leq y_n^-(\theta_n(j)) \right\} \right]. \end{aligned}$$

Since $\hat{S}_n^{\leftarrow}(\cdot)$ is a non-increasing and right-continuous function, it is easy to see that

$$p_n(\varepsilon) \leq \mathbb{P} \left[\bigcup_{j=1}^{m_n} \{W_n^+(\theta_n(j)) > b_{n,j}^+(\varepsilon)\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \{W_n^-(\theta_n(j)) \leq b_{n,j}^-(\varepsilon)\} \right],$$

where

$$W_n^\pm(\theta_n(j)) := \tau_n^{-1} \left(\frac{\hat{S}_n(y_n^\pm(\theta_n(j)))}{S(y_n^\pm(\theta_n(j)))} - 1 \right) \text{ and } b_{n,j}^\pm(\varepsilon) := \tau_n^{-1} \left(\frac{\theta_n(j)\alpha_n}{S(y_n^\pm(\theta_n(j)))} - 1 \right).$$

By assumption, $W_n^\pm(u) = \mathcal{O}_{\mathbb{P}}(1)$ uniformly on $u \in [\eta, 1]$. Since $S(\cdot)$ satisfies **(A.1)**,

$$y_n^\pm(\theta_n(j)) = S^{\leftarrow}(\alpha_n) + a(\alpha_n^{-1}) [L_\gamma(1/\theta_n(j)) + \zeta_n^\pm(\varepsilon)],$$

where $\zeta_n^\pm(\varepsilon) = \tau_n(\pm\varepsilon\delta_n^{-1} + o(1))$, the term $o(1)$ converging to 0 uniformly on $u \in [\eta, 1]$. Thus, from Lemma 3, equation (20), one has for all $j \in \{1, \dots, m_n\}$,

$$b_{n,j}^+(\varepsilon) = \tau_n^{-1} [\theta_n(j)L_\gamma^{\leftarrow}(L_\gamma(1/\theta_n(j)) + \zeta_n^+(\varepsilon)) - 1 + o(\tau_n)].$$

Since for $v \in [L_\gamma(1/\kappa_2), L_\gamma(1/\kappa_1)]$ the derivative of $L_\gamma^{\leftarrow}(\cdot)$ is larger than $\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1}$, a Taylor expansion leads to, for n large enough:

$$b_{n,j}^+(\varepsilon) \geq \frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2\delta_n}.$$

Similarly, for n large enough, one can show that

$$b_{n,j}^-(\varepsilon) \leq -\frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2\delta_n}.$$

Hence, $p_n(\varepsilon)$ is smaller or equal than

$$\mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ W_n^+(\theta_n(j)) > \frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2\delta_n} \right\} \right] + \mathbb{P} \left[\bigcup_{j=1}^{m_n} \left\{ W_n^-(\theta_n(j)) \leq -\frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2\delta_n} \right\} \right],$$

which is smaller than

$$\mathbb{P} \left[\sup_{u \in [\eta, 1]} \delta_n |W_n^+(u)| > \frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2} \right] + \mathbb{P} \left[\sup_{u \in [\eta, 1]} \delta_n |W_n^-(u)| \geq -\frac{\kappa_1\varepsilon(\kappa_1^{\gamma-1} \wedge \kappa_2^{\gamma-1})}{2} \right] \rightarrow 0,$$

proving (25). Finally, since $S^{\leftarrow}(\cdot)$ is a non-increasing function and collecting (24) and (25),

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - 1 \right| \leq \frac{3\tau_n^{-1}}{a(\alpha_n^{-1})} (T_{n,1} + T_{n,2}) = \mathcal{O}_{\mathbb{P}}(1),$$

which conclude the proof. \blacksquare

The next lemma establishes a uniform convergence result on processes of the form:

$$\hat{\Phi}_n(u) := \sum_{i=1}^n X_{n,i}(u),$$

where, for $\eta \in (0, 1)$, $\{X_{n,i}(u), u \in [\eta, 1]\}$, $i = 1, \dots, n$ are n independent stochastic processes with $X_{n,i}(\cdot)$ non-decreasing and positive. The expectation of $\hat{\Phi}_n(u)$ is denoted $\mu_n(u)$.

Lemma 6. *Let $\tau_n = (\ln(\mu_n(1))/\mu_n(1))^{1/2}$. If $\mu_n(\eta) \rightarrow \infty$ as n goes to infinity, if there exist positive constants C_X and C_μ such that for all $i \in \{1, \dots, n\}$ and $u \in [\eta, 1]$, $X_{n,i}(u) \leq C_X$, $\mu_n(\eta)/\mu_n(1) \geq C_\mu$ for n large enough and*

$$\sup \left\{ \left| \frac{\mu_n(u)}{\mu_n(u')} - 1 \right|, u \in [\eta, 1] \text{ with } |u - u'| \leq (\mu_n(1))^{-1/2} \right\} = \mathcal{O}(\tau_n), \quad (26)$$

then,

$$\sup_{u \in [\eta, 1]} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Proof of Lemma 6 – Let $C_\varepsilon := (3C_X/C_\mu)^{1/2}$. Using a multiplicative form of the Chernoff's inequality for bounded variables (see for instance [9, Theorem 1.1]), one has for all $u \in [\eta, 1]$

$$\mathbb{P} \left[\tau_n^{-1} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| > C_\varepsilon \right] \leq 2 \exp \left(-\frac{C_\varepsilon^2}{3C_X} \frac{\mu_n(u)}{\mu_n(1)} \ln(\mu_n(1)) \right) \leq \frac{2}{\mu_n(1)}. \quad (27)$$

Let us now introduce the sequence $m_n := \lceil (\mu_n(1))^{1/2} \rceil + 1 \rightarrow \infty$, and, for $j = 1, \dots, m_n$, let $\theta_n(j) := \eta + (j-1)(1-\eta)/(m_n-1) \in [\eta, 1]$. Here $\lceil \cdot \rceil$ is the notation for the ceiling function. Clearly, for all $u \in [\eta, 1]$, there exists $j_u \in \{1, \dots, m_n-1\}$ such that $\theta_n(j_u) \leq u < \theta_n(j_u+1)$ and then, since $X_{n,i}(\cdot)$, $i = 1, \dots, n$ are non-increasing,

$$\begin{aligned} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| &\leq \frac{1}{\mu_n(\eta)} \left[\left| \hat{\Phi}_n(\theta_n(j_u+1)) - \mu_n(\theta_n(j_u+1)) \right| + 2 \left| \hat{\Phi}_n(\theta_n(j_u)) - \mu_n(\theta_n(j_u)) \right| \right. \\ &\quad \left. + 2(\mu_n(\theta_n(j_u+1)) - \mu_n(\theta_n(j_u))) \right], \end{aligned}$$

leading to

$$\sup_{u \in [\eta, 1]} \left| \frac{\hat{\Phi}_n(u)}{\mu_n(u)} - 1 \right| \leq \frac{3}{\mu_n(\eta)} (T_{n,1} + T_{n,2}),$$

with

$$T_{n,1} := \max_{j=1, \dots, m_n-1} (\mu_n(\theta_n(j+1)) - \mu_n(\theta_n(j))) \text{ and } T_{n,2} := \max_{j=1, \dots, m_n} \left| \hat{\Phi}_n(\theta_n(j)) - \mu_n(\theta_n(j)) \right|.$$

Under (26), since for all $j \in \{1, \dots, m_n-1\}$, $\theta_n(j) - \theta_n(j+1) = (m_n-1)^{-1} \leq (\mu_n(1))^{-1/2}$, the following holds for n large enough:

$$T_{n,1} = \mathcal{O} \left((\mu_n(1) \ln(\mu_n(1)))^{1/2} \right). \quad (28)$$

Furthermore, using (27)

$$\begin{aligned} \mathbb{P} \left[(\mu_n(1) \ln(\mu_n(1)))^{-1/2} T_{n,2} > C_\varepsilon \right] &\leq \mathbb{P} \left[\left(\frac{\mu_n(1)}{\ln(\mu_n(1))} \right)^{1/2} \bigcup_{j=1}^{m_n} \left\{ \left| \frac{\hat{\Phi}_n(\theta_n(j))}{\mu_n(\theta_n(j))} - 1 \right| > C_\varepsilon \right\} \right] \\ &\leq 2(\mu_n(1))^{-1/2} \rightarrow 0, \end{aligned}$$

since $\mu_n(b) \rightarrow \infty$. Hence,

$$T_{n,2} = \mathcal{O}_{\mathbb{P}} \left((\mu_n(1) \ln(\mu_n(1)))^{1/2} \right). \quad (29)$$

From (28) and (29), since $\mu_n(1)/\mu_n(\eta) \leq C_\mu^{-1}$, $\tau_n^{-1} T_{n,1}/\mu_n(\eta) = \mathcal{O}(1)$ and $\tau_n^{-1} T_{n,2}/\mu_n(\eta) = \mathcal{O}_{\mathbb{P}}(1)$, which conclude the proof. \blacksquare

6.2 Proofs of main results

Proof of Theorem 1 – We start by showing the following equation:

$$\left| \left(\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \right)^\delta \mathcal{T}^{(\delta)}(\hat{S}_n^{\leftarrow} | \alpha_n, \eta, \varphi) - \int_\eta^1 \varphi(u) L_{\gamma_-}^\delta(1/u) du \Big/ \left(\int_\eta^1 \varphi(u) L_0(1/u) du \right)^\delta \right| = \mathcal{O}_{\mathbb{P}}(\tau_n), \quad (30)$$

Let us first introduce the following notations: for $u \in [\eta, 1]$,

$$\Delta_n(u) = \left(\frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{\hat{S}_n^{\leftarrow}(u\alpha_n)} - 1 \right), \quad R_n(u) = \ln \frac{1 + \Delta_n(u)}{1 + \Delta_n(1)},$$

$$\text{and } b_n(u) = \frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(\alpha_n)} - L_{\gamma_-}(1/u).$$

One has

$$\begin{aligned}
\left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{\hat{S}_n^\leftarrow(u\alpha_n)}{\hat{S}_n^\leftarrow(\alpha_n)} \right)^\delta &= \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \right)^\delta \left(\ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} + R_n(u) \right)^\delta \\
&= \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^\delta \\
&\quad + \sum_{j=0}^{\delta-1} C_\delta^j \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^j \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} R_n(u) \right)^{\delta-j}
\end{aligned}$$

Since from Lemma 5, $\hat{S}_n(\cdot)$ satisfies (23),

$$\bar{\Delta}_n := \sup_{u \in [\eta, 1]} |\Delta_n(u)| = \mathcal{O}_{\mathbb{P}} \left(\frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} \tau_n \right) = o_{\mathbb{P}}(1),$$

and thus, $|R_n(u)| = \mathcal{O}_{\mathbb{P}}(\bar{\Delta}_n) = o_{\mathbb{P}}(1)$. Remark that

$$\sum_{j=0}^{\delta-1} C_\delta^j \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^j \leq \sum_{j=0}^{\delta} C_\delta^j (L_{\gamma_-}(1/u) + b_n(u))^j \leq (1 + L_{\gamma_-}(1/\eta) + \bar{b}_n)^\delta,$$

with $\bar{b}_n = \sup\{|b_n(u)|, u \in [\eta, 1]\}$ which converges to 0 from the first part of Lemma 3. Hence,

$$\sum_{j=0}^{\delta-1} C_\delta^j \left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^j = \mathcal{O}(1),$$

and thus,

$$\left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \right)^\delta \left[\left(\ln \frac{\hat{S}_n^\leftarrow(u\alpha_n)}{\hat{S}_n^\leftarrow(\alpha_n)} \right)^\delta - \left(\ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^\delta \right] = \mathcal{O}_{\mathbb{P}}(\tau_n),$$

uniformly on $u \in [\eta, 1]$. Since

$$\left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{S^\leftarrow(u\alpha_n)}{S^\leftarrow(\alpha_n)} \right)^\delta - L_{\gamma_-}^\delta(1/u) = (L_{\gamma_-}(1/u) + b_n(u))^\delta - L_{\gamma_-}^\delta(1/u) = \mathcal{O}(\bar{b}_n),$$

we have, as a first conclusion that, uniformly on $u \in [\eta, 1]$,

$$\left(\frac{S^\leftarrow(\alpha_n)}{a(\alpha_n^{-1})} \ln \frac{\hat{S}_n^\leftarrow(u\alpha_n)}{\hat{S}_n^\leftarrow(\alpha_n)} \right)^\delta = L_{\gamma_-}^\delta(1/u) + \mathcal{O}_{\mathbb{P}}(\tau_n), \quad (31)$$

since, from Lemma 3, $\tau_n^{-1} \bar{b}_n \rightarrow 0$. Multiplying equation (31) by $\varphi(u)$ and integrating between η and 1 lead to (30).

The rest of the proof is based on the decomposition

$$\begin{aligned}
\mathcal{T}(\hat{S}_n | \alpha_n, \eta, \varphi) - \gamma &= \mathcal{T}^{(1)}(\hat{S}_n | \alpha_n, \eta, \varphi) - \gamma_+ \\
&\quad + \Psi_{\eta, \varphi}^\leftarrow \left(\max \left\{ \frac{[\mathcal{T}^{(1)}(\hat{S}_n | \alpha_n, \eta, \varphi)]^2}{\mathcal{T}^{(2)}(\hat{S}_n | \alpha_n, \eta, \varphi)}, \Psi_{\eta, \varphi}(0) \right\} \right) - \gamma_- =: D_n^+ + D_n^-.
\end{aligned}$$

Let us first consider the term D_n^+ . From (30), one has

$$D_n^+ = \frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} \mathcal{O}_{\mathbb{P}}(\tau_n) + \frac{a(\alpha_n^{-1})}{S^\leftarrow(\alpha_n)} \int_{\eta}^1 \varphi(u) L_{\gamma_-}(1/u) du \Big/ \int_{\eta}^1 \varphi(u) L_0(1/u) du - \gamma_+.$$

Since $S(\cdot)$ satisfies **(A.1)** and remarking that if $\gamma \geq 0$,

$$\int_{\eta}^1 \varphi(u) L_{\gamma_-}(1/u) du \Big/ \int_{\eta}^1 \varphi(u) L_0(1/u) du = 1,$$

it is clear that $D_n^+ = \mathcal{O}_{\mathbb{P}}(\tau_n)$. Now, using again (30), it is easy to check that

$$\frac{[\mathcal{T}^{(1)}(\hat{S}_n|\alpha_n, \eta, \varphi)]^2}{\mathcal{T}^{(2)}(\hat{S}_n|\alpha_n, \eta, \varphi)} = \Psi_{\eta, \varphi}(\gamma_-) (1 + \mathcal{O}_{\mathbb{P}}(\tau_n)).$$

Since $\Psi_{\eta, \varphi}(\cdot)$ is a decreasing function (see Lemma 1) and $\tau_n \rightarrow 0$,

$$\max \left\{ \frac{[\mathcal{T}^{(1)}(\hat{S}_n|\alpha_n, \eta, \varphi)]^2}{\mathcal{T}^{(2)}(\hat{S}_n|\alpha_n, \eta, \varphi)}, \Psi_{\eta, \varphi}(0) \right\} = \Psi_{\eta, \varphi}(\gamma_-) (1 + \mathcal{O}_{\mathbb{P}}(\tau_n)).$$

Finally, since from Lemma 1, the derivative of $\Psi_{\eta, \varphi}^{\leftarrow}(\cdot)$ is bounded in a neighborhood of $\Psi_{\eta, \varphi}(\gamma_-)$,

$$D_n^- = \Psi_{\eta, \varphi}^{\leftarrow} \left(\max \left\{ \frac{[\mathcal{T}^{(1)}(\hat{S}_n|\alpha_n, \eta, \varphi)]^2}{\mathcal{T}^{(2)}(\hat{S}_n|\alpha_n, \eta, \varphi)}, \Psi_{\eta, \varphi}(0) \right\} \right) - \Psi_{\eta, \varphi}^{\leftarrow}(\Psi_{\eta, \varphi}(\gamma_-)) = \mathcal{O}_{\mathbb{P}}(\tau_n),$$

which concludes the proof. \blacksquare

Before proving Corollaries 1, 2 and 3, we establish the following result that can be useful when working under framework **(F)**. Let us assume that n independent random variables Z_1, \dots, Z_n (not necessarily identically distributed) are recorded and let us consider the statistic defined for all $y \geq 0$ and for all $i \in 1, \dots, n$ by

$$\hat{R}_{n,i}(y) := R_{n,i}(y; Z_i),$$

where $R_{n,i}(\cdot)$ is a given deterministic functional. Assume that the statistic defined for all $y \geq 0$ by

$$\hat{S}_n(y) := Q(y; Z_1, \dots, Z_n) = \sum_{i=1}^n \hat{R}_{n,i}(y) \Big/ \sum_{i=1}^n \hat{R}_{n,i}(0), \quad (32)$$

is an estimator of the survival function $S(\cdot) \in \mathcal{ERV}(\gamma, \mathbf{a}(\cdot))$. The following proposition gives sufficient conditions ensuring that $\mathcal{T}(\hat{S}_n^{\leftarrow}|\alpha_n, \eta, \varphi)$ is a consistent estimator of γ .

Proposition 1. *Let (σ_n) and (α_n) be sequences converging to 0 as $n \rightarrow \infty$ and let Z_1, \dots, Z_n be independent random variables. If for all $i = 1, \dots, n$ and for some positive constant C_R , the stochastic process $\hat{R}_{n,i}(\cdot) \in \mathcal{D}(\mathbb{R}^+, \mathbb{R}^+)$ with $\sup\{\hat{R}_{n,i}(y), y \geq 0\} \leq C_R$ almost surely, if $S(\cdot)$ satisfies condition **(A.1)** with sequences $\tau_n := \sigma_n (\ln(\sigma_n^{-1}))^{1/2}$ and α_n and if there exists a constant $r_0 > 0$ such that for all sequence $y_n(u)$ satisfying $a^{-1}(\alpha_n^{-1})(y_n(u) - S^{\leftarrow}(u\alpha_n)) \rightarrow 0$ for all $u \in [\eta, 1]$,*

$$\sup_{u \in [\eta, 1]} \left| \frac{r_0 \alpha_n \sigma_n^2 \mu_n(u)}{S(y_n(u))} - 1 \right| = o(\tau_n) \text{ and } \sup_{u \in [\eta, 1]} \left| \alpha_n \sigma_n^2 \sum_{i=1}^n \hat{R}_{n,i}(0) - r_0 \right| = o_{\mathbb{P}}(\tau_n), \quad (33)$$

where

$$\mu_n(u) := \mathbb{E} \left(\sum_{i=1}^n \hat{R}_{n,i}(y_n(u)) \right),$$

then

$$\left| \mathcal{T} \left(\hat{S}_n^{\leftarrow} | \alpha_n, \eta, \varphi \right) - \gamma \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Proof of Proposition 1 – From Theorem 1, it suffices to show that

$$\frac{S^{\leftarrow}(\alpha_n)}{a(\alpha_n^{-1})} \tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n^{\leftarrow}(u\alpha_n)}{S^{\leftarrow}(u\alpha_n)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1), \quad (34)$$

where $\hat{S}_n(\cdot)$ is given by (32). From Lemma 5, it suffices to prove that for every sequence $y_n(u)$ satisfying $a^{-1}(\alpha_n^{-1})(y_n(u) - S^{\leftarrow}(u\alpha_n)) \rightarrow 0$ for all $u \in [\eta, 1]$

$$\tau_n^{-1} \sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n(y_n(u))}{S(y_n(u))} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1). \quad (35)$$

First, let us consider the numerator of $\hat{S}_n(y_n(u))$ given by:

$$\hat{\Phi}_n(u) := \sum_{i=1}^n \hat{R}_{n,i}(y_n(u)).$$

Recall that, by assumption,

$$\mu_n(u) = \mathbb{E} \left(\hat{\Phi}_n(u) \right) = r_0^{-1} \frac{S(y_n(u))}{\alpha_n \sigma_n^2} (1 + o(\tau_n)),$$

uniformly on $u \in [\eta, 1]$. Since the survival function $S(\cdot)$ satisfies condition **(A.1)**, Lemma 3, equation (20) entails that, uniformly on $u \in [\eta, 1]$

$$S(y_n(u)) = \alpha_n \left[1 / L_\gamma^{\leftarrow} (L_\gamma(1/u) + v_n + o(\tau_n)) + o(\tau_n) \right] = u \alpha_n (1 + o(1)),$$

where $v_n := a^{-1}(\alpha_n^{-1})(y_n(u) - S^{\leftarrow}(u \alpha_n)) \rightarrow 0$. Thus, uniformly on $u \in [\eta, 1]$,

$$\mu_n(u) = r_0^{-1} \sigma_n^{-2} \left[1 / L_\gamma^{\leftarrow} (L_\gamma(1/u) + v_n + o(\tau_n)) + o(\tau_n) \right].$$

As a first conclusion, $(\mu_n(1) / \ln(\mu_n(1)))^{-1/2}$ is asymptotically proportional to τ_n and, for n large enough,

$$\frac{\mu_n(\eta)}{\mu_n(1)} \geq \eta/2. \quad (36)$$

We now show that $\mu_n(u)$ satisfies condition (26) of Lemma 6. For $(u, u') \in [\eta, 1]^2$ such that $|u - u'| \leq (\mu_n(1))^{-1/2}$, since the derivatives of $1/L_\gamma^{\leftarrow}(\cdot)$ and $L_\gamma(1/\cdot)$ are bounded on $[\eta, 1]$, a Taylor expansion leads to $|\mu_n(u) - \mu_n(u')| = \mathcal{O}(\sigma_n^{-2} \tau_n)$ uniformly on $u \in [\eta, 1]$. Thus,

$$\left| \frac{\mu_n(u)}{\mu_n(u')} - 1 \right| \leq \frac{1}{\mu_n(\eta)} |\mu_n(u) - \mu_n(u')| = \mathcal{O}(\tau_n). \quad (37)$$

Hence, collecting (36) and (37), Lemma 6 entails that

$$\sup_{u \in [\eta, 1]} \left| \frac{\hat{\Phi}_n(u) r_0 \alpha_n \sigma_n^2 (1 + o(\tau_n)) - 1}{S(y_n(u))} \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

Finally, since by assumption,

$$\alpha_n \sigma_n^2 \sum_{i=1}^n \hat{R}_{n,i}(0) = r_0 + o_{\mathbb{P}}(\tau_n),$$

uniformly on $u \in [\eta, 1]$, one has

$$\sup_{u \in [\eta, 1]} \left| \frac{\hat{S}_n(y_n(u))}{S(y_n(u))} (1 + o_{\mathbb{P}}(\tau_n)) - 1 \right| = \mathcal{O}_{\mathbb{P}}(\tau_n).$$

It is easy to check that the factor $1 + o_{\mathbb{P}}(\tau_n)$ can be removed proving (35) for every sequence $y_n(u)$ satisfying $a^{-1}(\alpha_n^{-1})(y_n(u) - S^{\leftarrow}(u \alpha_n)) \rightarrow 0$ for all $u \in [\eta, 1]$ and consequently (34) is established. This concludes the proof. \blacksquare

Note that under the assumptions of Proposition 1, $\hat{S}_n(\cdot) \in \mathcal{D}(\mathbb{R}^+, [0, 1])$. Proposition 1 is a key point to prove corollaries 1, 2 and 3 since all the survival function estimators can be written as in (32).

Proof of Corollary 1 – It is easy to check that the estimator $\hat{S}_n(\cdot | x_0)$ defined in (4) is of the form (32) with

$$\hat{R}_{n,i}(y) = |H_n| K_{H_n}(x_0 - X_i) \mathbb{I}_{\{Y_i > y\}} \leq \|K\|_{\infty}, \quad (38)$$

under **(B.2)**. Let $y_n(u)$ be a sequence such that $a^{-1}(\alpha_n^{-1} | x_0)(y_n(u) - S^{\leftarrow}(u \alpha_n | x_0)) \rightarrow 0$ as $n \rightarrow \infty$. Let us first focus on the denominator of $\hat{S}_n(\cdot | x_0)$. Let

$$\hat{g}_n(x_0) = \frac{1}{n} \sum_{i=1}^n K_{H_n}(x_0 - X_i) = (n |H_n|)^{-1} \sum_{i=1}^n \hat{R}_{n,i}(0).$$

Under **(B.1)** and **(B.2)**, since $n|H_n| \rightarrow \infty$, it is well known that

$$\hat{g}_n(x_0) = g(x_0) \left(1 + \mathcal{O}_{\mathbb{P}}(\|H_n\|_{\infty}) + \mathcal{O}_{\mathbb{P}}\left((n|H_n|)^{-1/2}\right) \right),$$

(see Parzen [31] for a proof). Then, since by assumption $\tau_n^{-1}\|H_n\|_{\infty}$ and $\tau_n^{-1}(n|H_n|)^{-1/2}$ converge to 0, one has that $\hat{g}_n(x_0)/g(x_0) = 1 + o_{\mathbb{P}}(\tau_n)$. Thus the second part of condition (33) in Proposition 1 is satisfied. Now, let

$$\hat{\Phi}_n(u|x_0) := \sum_{i=1}^n \hat{R}_{n,i}(y_n(u)).$$

Let $\mu_n(u|x_0) = \mathbb{E}(\hat{\Phi}_n(u|x_0))$. Straightforward calculus leads to:

$$\begin{aligned} \frac{\mu_n(u|x_0)}{n|H_n|g(x_0)S(y_n(u)|x_0)} &= \int_{\mathcal{U}_p} \left(\frac{S(y_n(u)|x_0 - H_n t)}{S(y_n(u)|x_0)} - 1 \right) K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt \\ &+ \int_{\mathcal{U}_p} K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt. \end{aligned}$$

Let us first focus on the second term. Under condition **(B.1)**,

$$\int_{\mathcal{U}_p} K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt - 1 \leq \frac{c_g \|H_n\|_{\infty}}{g(x_0)} \int_{\mathcal{U}_p} K(t) \|t\|_{\infty} dt = \mathcal{O}(\|H_n\|_{\infty}) = o(\tau_n), \quad (39)$$

since $\tau_n^{-1}\|H_n\|_{\infty} \rightarrow 0$. Now, since $S(\cdot|x_0)$ satisfies **(A.1)**, it is easy to check that there exists $\delta > 1$ such that for all $u \in [\eta, 1]$, $y_n(u)/S^{\leftarrow}(\alpha_n|x_0) \in [\delta^{-1}, \delta]$. Hence, from condition (5) and using (39)

$$\int_{\mathcal{U}_p} \left(\frac{S(y_n(u)|x_0 - H_n t)}{S(y_n(u)|x_0)} - 1 \right) K(t) \frac{g(x_0 - H_n t)}{g(x_0)} dt = o(\tau_n). \quad (40)$$

Collecting (39) and (40), the first part of condition (33) is satisfied. Proposition 1 concludes the proof. \blacksquare

Proof of Corollary 2 – It is easy to check that the estimator $\hat{S}_{n,x_0}(\cdot)$ defined in (6) is of the form (32) with

$$\hat{R}_{n,i}(z) = n|H_n| \mathbb{I}_{\{Y_{x_i} > z\}} \int_{A_{n,i}} K_{H_n}(x_0 - t) dt. \quad (41)$$

Under **(B.2)** and **(C)**, $\hat{R}_{n,i}(z) \leq n\|K\|_{\infty} \mathcal{V}_{n,i} \leq C_{\mathcal{V}}\|K\|_{\infty}$ and, since $x_0 \in \hat{E}$,

$$\int_E K_{H_n}(x_0 - t) dt = 1 \text{ and thus } \sum_{i=1}^n \hat{R}_{n,i}(0) = n|H_n|.$$

As a consequence, the second part of condition (33) is satisfied. Let $z_n(u)$ be a sequence such that $a_{x_0}^{-1}(\alpha_n^{-1})(z_n(u) - S_{x_0}^{\leftarrow}(u\alpha_n)) \rightarrow 0$ as $n \rightarrow \infty$. Let

$$\hat{\Phi}_{n,x_0}(u) = \sum_{i=1}^n \hat{R}_{n,i}(z_n(u)),$$

and denote by $\mu_{n,x_0}(u)$ its expectation. Recalling that

$$\sum_{i=1}^n \int_{A_i} K_{H_n}(x_0 - t) dt = 1,$$

it is easy to check that,

$$\frac{\mu_{n,x_0}(u)}{n|H_n|S_{x_0}(z_n(u))} = 1 + \sum_{i=1}^n \int \left(\frac{S_{x_i}(z_n(u))}{S_{x_0}(z_n(u))} - 1 \right) \mathbb{I}_{\{t \in A_i \cap B(x_0, H_n)\}} K_{H_n}(x_0 - t) dt.$$

Since $n\bar{\mathcal{V}}_n \leq C_{\mathcal{V}}$, for all $i \in \{1, \dots, n\}$, $A_i \subset B(x_i, r_{n,p}I_p)$ with $2r_{n,p} := (C_{\mathcal{V}}/n)^{1/p}$. Thus if $A_i \cap B(x_0, H_n) \neq \emptyset$, $x_i \in B(x_0, 2H_n)$. Indeed, if there exists $t \in A_i \cap B(x_0, H_n)$,

$$\|H_n^{-1}(x_i - x_0)\|_{\infty} \leq \|H_n^{-1}(x_i - t)\|_{\infty} + \|H_n^{-1}(t - x_0)\|_{\infty} \leq \|H_n^{-1}\|_{\infty} r_{n,p} + 1.$$

By assumption $n|H_n| \rightarrow \infty$ and since for all positive-definite matrix M of size p , $|M|^{1/p} \leq \|M\|_\infty$, one has for n large enough that $\|H_n^{-1}\|_\infty r_{n,p} \leq 1$ and thus, $\|H_n^{-1}(x_i - x_0)\|_\infty \leq 2$ i.e. $x_{n,i} \in B(x_0, 2H_n)$. Furthermore, since $S_{x_0}(\cdot)$ satisfies condition **(A.1)**, it is easy to check that there exists $\delta > 1$ such that for all $u \in [\eta, 1]$ $z_n(u)/S_{x_0}^{\leftarrow}(\alpha_n) \in [\delta^{-1}, \delta]$. Hence, from condition (7),

$$\mu_{n,x_0}(u) = n|H_n|S_{x_0}(z_n(u))(1 + o(\tau_n)),$$

proving the first part of condition (33). Proposition 1 concludes the proof. \blacksquare

Proof of Corollary 3 – Clearly, the estimator $\hat{S}_n(\cdot)$ given in (9) is of the form (32) with for $i \in \{1, \dots, n\}$

$$\hat{R}_{n,i}(z) = n \frac{x_i - x_{i-1}}{x_n - b_1} \mathbb{1}_{\{Y_{x_i} > z\}} \leq \frac{2d}{b_2 - b_1}, \quad (42)$$

under condition **(D)**. It is easy to check that

$$\sum_{i=1}^n \hat{R}_{n,i}(0) = n,$$

and thus the second part of condition (33) of Proposition 1 is clearly satisfied with $r_0 = 1$. Let $z_n(u)$ be a sequence such that $a^{-1}(\alpha_n^{-1})(z_n(u) - S^{\leftarrow}(u\alpha_n)) \rightarrow 0$ as $n \rightarrow \infty$ and let

$$\hat{\Phi}_n(u) := \sum_{i=1}^n \hat{R}_{n,i}(z_n(u)).$$

Our goal is to provide an expansion of $\mu_n(u) := \mathbb{E}(\hat{\Phi}_n(u))$ in order to check the validity of the first part of condition (33) in Proposition 1. We start with

$$\frac{\mu_n(u)}{nS(z_n(u))} = \sum_{i=1}^n \frac{x_i - x_{i-1}}{x_n - b_1} \left(\frac{S_{x_i}(z_n(u))}{S(z_n(u))} - c(x_i) \right) \quad (43)$$

$$+ \sum_{i=1}^n \frac{x_i - x_{i-1}}{x_n - b_1} c(x_i). \quad (44)$$

Since $S(\cdot)$ satisfies **(A.1)**, it is easy to check that there exists $\delta > 1$ such that for all $u \in [\eta, 1]$, $z_n(u)/S^{\leftarrow}(\alpha_n) \in [\delta^{-1}, \delta]$. Hence, from condition (11), the term (43) is a $o(\tau_n)$ uniformly on $u \in [\eta, 1]$. Furthermore,

$$\begin{aligned} 1 &= \frac{1}{b_2 - b_1} \int_{b_1}^{b_2} c(s) ds = \frac{x_n - b_1}{b_2 - b_1} \sum_{i=1}^n \frac{x_i - x_{i-1}}{x_n - b_1} c(x_i) + c(b_2) \frac{x_n - b_1}{b_2 - b_1} (b_2 - x_n) \\ &+ \frac{1}{b_2 - b_1} \sum_{i=1}^{n+1} \int_{x_{i-1}}^{x_i} (c(s) - c(x_i)) ds. \end{aligned}$$

From conditions **(D)** and (10), since $n\sigma_n \rightarrow \infty$, we deduce from the previous equation that the term (44) is equal to $1 + o(\tau_n)$ uniformly on $u \in [\eta, 1]$. Hence, condition (33) is satisfied and the conclusion follows applying Proposition 1. \blacksquare

References

- [1] Beirlant, J., Goegebeur, Y., Segers, J. and Teugels J. (2004). *Statistics of extremes: Theory and applications*, Wiley.
- [2] Chavez-Demoulin, V. and Davison, A.C. (2005). Generalized additive modelling of sample extremes, *Journal of the Royal Statistical Society, Series C*, **54**, 207–222.
- [3] Csörgő, S., Deheuvels, P. and Mason, D. (1985). Kernel estimates of the tail index of a distribution, *The Annals of Statistics*, **13(3)**, 1050–1077.

- [4] Daouia, A., Florens, J.-P., Simar, L. (2010). Frontier estimation and extreme value theory, *Bernoulli*, **16(4)**, 1039–1063.
- [5] Daouia, A., Gardes, L. and Girard, S. (2013). On kernel smoothing for extremal quantile regression, *Bernoulli*, **19(5B)**, 2557–2589.
- [6] Davison, A.C. and Ramesh, N.I. (2000). Local likelihood smoothing of sample extremes, *Journal of the Royal Statistical Society, Series B*, **62**, 191–208.
- [7] Davison, A.C. and Smith, R.L. (1990). Models for exceedances over high thresholds, *Journal of the Royal Statistical Society, Series B*, **52**, 393–442.
- [8] Dekkers, A. L. M., Einmahl, J. H. J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *The Annals of Statistics*, **17(4)**, 1833–1855.
- [9] Dubhashi, D. and Parconesi, A. (2009). *Concentration of measure for the analysis of randomized algorithms*, Cambridge University Press, New York.
- [10] Drees, H. (1995). Refined Pickands estimators of the extreme value index, *The Annals of Statistics*, **23**, 2059–2080.
- [11] Einmahl, J.H.J, Fils-Villetard, A. and Guillou, A. (2008). Statistics of extremes under random censoring, *Bernoulli*, **14(1)**, 207–227.
- [12] Drees, H. (1998). A general class of estimators of the extreme value index, *Journal of Statistical Planning and Inference*, **66**, 95–112.
- [13] Einmahl, J. H. J., de Haan, L. and Zhou, C. (2014). Statistics of Heteroscedastic Extremes, *Journal of the Royal Statistical Society, Series B*, to appear.
- [14] Fisher, R. A. and Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proceedings of the Cambridge Philosophical Society*, **24**, 180–190.
- [15] Fraga Alves, M. I., Gomes, M. I., de Haan, L. and Neves, C. (2007). A note on second order conditions in extreme value theory: linking general and heavy tail conditions, *REVSTAT - Statistical Journal*, **5(3)**, 285–304.
- [16] Gardes, L. and Girard, S. (2008). A moving window approach for nonparametric estimation of the conditional tail index, *Journal of Multivariate Analysis*, **99**, 2368–2388.
- [17] Gardes, L. and Girard, S. (2010). Conditional extremes from heavy-tailed distributions: an application to the estimation of extreme rainfall return levels, *Extremes*, **13(2)**, 177–204.
- [18] Gardes, L. and Girard, S. (2012). Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics*, **6**, 1715–1744.
- [19] Gardes, L. and Stupfler, G. (2014). Estimation of the conditional tail index using a smoothed local Hill estimator, *Extremes*, **17(1)**, 45–75.
- [20] Gardes, L. and Stupfler, G. (2014). Estimating extreme quantiles under random truncation, *TEST*, DOI: 10.1007/s11749-014-0403-5.
- [21] Gasser, T. and Müller, H. G. (1979). Kernel Estimation of Regression Functions, in *Smoothing Techniques for Curve Estimation*, Lecture Notes in Mathematics 757, eds. T. Gasser and M. Rosenblatt, Heidelberg: Springer-Verlag, 23–68.
- [22] Georgiev, A.A. (1988). Consistent nonparametric multiple regression: the fixed design case, *Journal of Multivariate Analysis*, **25**, 100–110.
- [23] Gnedenko, B. V. (1943). Sur la distribution limite du terme maximum d’une série aléatoire, *Annals of Mathematics*, **44**, 423–453.

- [24] Goegebeur, Y., Guillou, A. and Osmann, M. (2014). A local moment type estimator for the extreme value index in regression with random covariates, *Canadian Journal of Statistics*, **42(3)**, 487–507.
- [25] Gomes, M.I. and Neves, N.M. (2011). Estimation of the extreme value index for randomly censored data, *Biometrical Letters*, **48(1)**, 1–22.
- [26] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer.
- [27] de Haan, L., Klein Tank, A. and Neves, C. (2011). On tail trend detection: modeling relative risk, *arXiv preprint, arXiv:1106.4149*.
- [28] Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, **3(5)**, 1163–1174.
- [29] Kaplan, E.L. and Meier, P. (1958). Non-parametric estimation from incomplete observations, *Journal of the American Statistical Association*, **53**, 457–481.
- [30] Nadaraya, E.A. (1964). On estimating regression, *Theory of Probability and its Application*, **9(1)**, 141–142.
- [31] Parzen, E. (1962). On the estimation of a probability density function and mode, *Annals of Mathematical Statistics*, **33**, 1065–1076.
- [32] Pickands, III J. (1975). Statistical inference using extreme order statistics, *The Annals of Statistics*, **3(1)**, 119–131.
- [33] Smith, R.L. (1989). Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion), *Statistical Science*, **4**, 367–393.
- [34] Stone, C. J. (1977). Consistent nonparametric regression, *Annals of Statistics*, **5**, 595–645.
- [35] Stupfler, G. (2013). A moment estimator for the conditional extreme-value index, *Electronic Journal of Statistics*, **7**, 2298–2343.
- [36] Tsay, R.S. (2002). *Analysis of financial time series*. Wiley, New York.
- [37] Wand, M.P. and Jones, M.C. (1995). *Kernel smoothing*, Chapman and Hall, London.
- [38] Wang, H. and Tsai, C.L. (2009). Tail index regression, *Journal of the American Statistical Association*, **104(487)**, 1233–1240.
- [39] Watson, G. S. (1964). Smooth regression analysis, *Sankhya*, **26(15)**, 175–184.

η	P1-1				P1-2			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.194	0.098	0.056	0.035	0.682	0.320	0.176	0.112
0.01	0.195	0.100	0.057	0.035	0.700	0.330	0.182	0.116
0.02	0.210	0.102	0.059	0.037	0.735	0.349	0.193	0.123
0.04	0.229	0.112	0.067	0.044	0.804	0.390	0.218	0.139
0.08	0.303	0.133	0.093	0.056	0.960	0.478	0.276	0.172
0.16	0.438	0.200	0.158	0.088	1.338	0.699	0.419	0.259

η	P1-3			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.108	0.077	0.044	0.030
0.01	0.112	0.082	0.048	0.034
0.02	0.121	0.091	0.056	0.040
0.04	0.143	0.110	0.068	0.050
0.08	0.201	0.166	0.096	0.069
0.16	0.360	0.292	0.159	0.112

Table 1: Values for different values of η and n of $\text{EMSE}(\mathcal{T}(\hat{S}_n^{\leftarrow} | n^{-1/3}, \eta, \varphi))$ in the case $\varphi(\cdot) = 1$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$.

η	P1-1				P1-2			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.242	0.129	0.074	0.047	0.268	0.151	0.085	0.054
0.01	0.228	0.125	0.073	0.046	0.278	0.158	0.089	0.057
0.02	0.224	0.122	0.070	0.045	0.299	0.171	0.096	0.063
0.04	0.221	0.124	0.069	0.050	0.341	0.197	0.113	0.075
0.08	0.265	0.140	0.090	0.057	0.440	0.256	0.148	0.100
0.16	0.373	0.193	0.134	0.080	0.634	0.404	0.241	0.166

η	P1-3			
	$n = 50$	$n = 100$	$n = 200$	$n = 400$
0.005	0.094	0.063	0.034	0.024
0.01	0.095	0.066	0.038	0.027
0.02	0.100	0.072	0.044	0.031
0.04	0.112	0.082	0.053	0.039
0.08	0.149	0.124	0.076	0.056
0.16	0.273	0.221	0.122	0.097

Table 2: Values for different values of η and n of $\text{EMSE}(\mathcal{T}(\hat{S}_n^{\leftarrow} | n^{-1/3}, \eta, \varphi))$ in the case $\varphi(\cdot) = \ln(1/\cdot)$ using the generating processes **P1-1**, **P1-2** and **P1-3** with $\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$.

α_n	P1-1		P1-2		P1-3	
	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$	$\tilde{c} = 1$	$\tilde{c}(\cdot) = 1 + \ln(1/\cdot)$
$n^{-1/2}$	1.018	0.725	0.195	0.141	0.484	0.534
$n^{-1/3}$	1.211	1.070	0.402	0.352	0.809	0.735
$n^{-1/4}$	1.232	1.151	0.428	0.380	0.835	0.930
$n^{-1/5}$	1.210	1.177	0.400	0.252	0.933	0.865
$n^{-1/6}$	1.274	1.172	0.433	0.284	1.000	0.811

Table 3: Comparison of $\mathcal{T}(\hat{S}_n^{\leftarrow} | \alpha_n, \eta, \varphi)$ and the moment estimator: values of the ratio $R_1(\alpha_n)$ with $\varphi(\cdot) = \ln(1/\cdot)$ and $\tau = 0.02$ using the generating processes **P1-1**, **P1-2** and **P1-3** with different values of α_n and different functions $\tilde{c}(\cdot)$.

$(\alpha_{n,i}, H_{n,i,j})$	P2-1					P2-2				
	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 2$	0.390	0.713	0.653	1.051	1.066	0.055	0.139	0.179	0.312	0.327
$i = 3$	0.666	1.089	1.130	1.078	1.036	0.109	0.288	0.571	0.600	0.698
$i = 4$	0.789	1.079	1.268	1.136	1.132	0.208	0.435	0.617	0.692	0.833
$i = 5$	0.936	1.121	1.352	1.286	1.297	0.295	0.454	0.557	0.596	0.642
$i = 6$	0.885	1.256	1.400	1.371	1.355	0.223	0.470	0.580	0.622	0.686

$(\alpha_{n,i}, H_{n,i,j})$	P2-3				
	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 2$	0.018	0.315	0.423	0.407	0.494
$i = 3$	0.214	0.586	0.802	1.058	0.882
$i = 4$	0.457	0.842	1.065	1.167	1.118
$i = 5$	0.354	1.062	1.355	1.172	1.103
$i = 6$	0.532	1.148	1.300	1.148	1.077

Table 4: Comparison of $\mathcal{T}(\hat{S}_n^{\leftarrow}(\cdot | x_0) | \alpha_n, \eta, \varphi)$ and the estimator $\hat{\gamma}_S(\alpha_n, H_n | x_0)$: values of the ratio $R_2(\alpha_n, H_n)$ with $\varphi(\cdot) = \ln(1/\cdot)$ and $\tau = 0.02$ using the generating processes **P2-1**, **P2-2** and **P2-3** for $(\alpha_{n,i} = n^{-1/i}, H_{n,i,j} = n^{(1-i)/(ij)})$, $i, j = 2, \dots, 6$.