On the estimation of the functional Weibull tail-coefficient

Laurent Gardes, Stephane Girard

▶ **To cite this version:**

On the estimation of the functional Weibull tail-coefficient

Laurent Gardes[(1)] and Stéphane Girard[(2,⋆)]

[(1)] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes,
67084 Strasbourg Cedex, France.
[(2)] Inria Grenoble Rhône-Alpes & LJK, team Mistis, Inovallée, 655, av. de l'Europe, Montbonnot,
38334 Saint-Ismier cedex, France.
[(⋆)] `Stephane.Girard@inria.fr` (corresponding author)

Abstract

We present a nonparametric family of estimators for the tail index of a Weibull tail-distribution when functional covariate is available. Our estimators are based on a kernel estimator of extreme conditional quantiles. Asymptotic normality of the estimators is proved under mild regularity conditions. Their finite sample performances are illustrated both on simulated and real data.

Keywords – Conditional Weibull tail-coefficient, extreme quantiles, nonparametric estimation.

AMS Subject classifications – 62G32, 62G05, 62E20.

1   Introduction

Weibull tail-distributions encompass a variety of light-tailed distributions, such as Weibull, Gaussian, gamma and logistic distributions. Let us recall that a cumulative distribution function F has a Weibull tail if it satisfies the following property: there exists $\theta > 0$ such that for all $t > 0$,

$$\lim_{y \to \infty} \frac{\log(1 - F(ty))}{\log(1 - F(y))} = t^{1/\theta}. \tag{1}$$

The parameter θ is referred to as the Weibull tail-coefficient. A general account on Weibull tail-distributions can be found in [7], see also [6] for an application to the modeling of large claims in non-life insurance. Dedicated methods have been proposed to estimate the Weibull tail-coefficient since the relevant information is only contained in the extreme upper part of the sample denoted hereafter by $Y_1, \dots, Y_n$. A first direction was investigated in [8] where an estimator based on the record values is proposed. Another family of approaches [3, 4, 11, 17] consists of using the k_n upper order statistics $Y_{n-k_n+1,n} \leq \cdots \leq Y_{n,n}$ where $k_n \to \infty$ as $n \to \infty$. Note that, since θ is defined through an asymptotic behavior of the tail, the estimator should only use the extreme-values of the sample and thus the extra condition $k_n/n \to 0$ is required. More specifically, most recent estimators are based on the log-spacings between the k_n upper order statistics [7, 16, 24, 25, 26, 27].

Here, we focus on the situation where some covariate information X is recorded simultaneously with the quantity of interest Y. In the general case, the tail heaviness of Y given X depends on X, and thus the Weibull tail-coefficient is a function $\theta(X)$ of the covariate. When the covariate is finite dimensional, some new tools have been introduced [14, 15] to estimate extreme conditional

quantiles. We refer to [18] for an application to the risk modeling associated with extreme rainfalls. In this case, the selected covariate is the geographical location but other relevant informations could be included such as climatic curves. More generally, covariates may be curves (electricity price/demand curves, medical curves, ...) in many other situations coming from applied sciences, see [10], paragraph 1.2.2. However, the estimation of the Weibull tail-coefficient with functional covariates has not been addressed yet. Our approach relies on the use of $\hat{q}_n$ a functional kernel estimator of conditional quantiles, see [20] for an example. Similarly to the unconditional case, the estimation of $\theta(X)$ is based on the extreme observations of $Y|X$. Therefore, a close study of the asymptotic properties of $\hat{q}_n$ when estimating extreme quantiles is necessary. Two statistical fields are thus involved in this study: nonparametric smoothing techniques adapted to functional data are required in order to deal with the covariate X while extreme-value analysis is used to study the tail behavior of $Y|X$.

The family of nonparametric functional estimators is introduced in Section 2 and its asymptotic normality is established. A particular sub-family of estimators is exhibited in Section 3, their finite sample behavior is illustrated on some simulated data in Section 4 and on a real dataset in Section 5. Proofs are postponed to Section 6.

2 Main result

Let (X_i, Y_i), $i = 1, \ldots, n$, be independent copies of a random pair $(X, Y) \in E \times \mathbb{R}$ where E is an arbitrary space associated with a semi-metric d. Recall that a semi-metric (or pseudometric) may allow the distance between two different points to be zero, see [22], Definition 3.2. The conditional survival function of Y given $X = x \in E$ is denoted by $\bar{F}(y|x) := \mathbb{P}(Y > y|X = x)$ and is supposed to be continuous and strictly decreasing with respect to y. Discussing the existence of regular versions of $\bar{F}(.|.)$ is beyond the scope of this paper. Let us just note that such an existence is insured when (E, d) is a Polish space [30]. The associated conditional cumulative hazard function is defined by $H(y|x) := -\log \bar{F}(y|x)$ and the conditional quantile is therefore given by $q(\alpha|x) := \bar{F}^{-1}(\alpha|x) = H^{-1}(\log(1/\alpha)|x)$, for all $\alpha \in (0,1)$. In this paper, we focus on conditional Weibull tail-distributions. In such a case, analogously to (1), $H(.|x)$ is a regularly varying function with index $1/\theta(x)$, i.e.

$$\lim_{y \to \infty} \frac{H(ty|x)}{H(y|x)} = t^{1/\theta(x)},$$

for all $t > 0$. In this situation, $\theta(.)$ is an unknown positive function of the covariate $x \in E$ referred to as the functional Weibull tail-coefficient. From [9], Theorem 1.5.12, $H^{-1}(.|x)$ is also a regularly varying function with index $\theta(x)$ and thus, there exists a slowly-varying function $\ell(.|x)$ such that

$$q(e^{-y}|x) = H^{-1}(y|x) = y^{\theta(x)}\ell(y|x). \tag{2}$$

Recall that the slowly-varying function $\ell(.|x)$ is such that

$$\lim_{y \to \infty} \frac{\ell(ty|x)}{\ell(y|x)} = 1, \tag{3}$$

for all $t > 0$, see [9] for a general account on regular variation theory. In view of (2), it appears that the functional Weibull tail-coefficient drives the asymptotic behavior of conditional extreme quantiles. The object of interest is thus $\theta(x)$ where x is in some arbitrary semi-metric space (E, d).

2

The most usual case is $E = \mathbb{R}^p$, but our framework also includes the infinite dimensional case, for instance when x is a curve. We propose a family of estimators of $\theta(x)$ based on some properties of the log-spacings of the conditional quantiles: Let $\alpha \in (0,1)$ small enough and $\tau \in (0,1)$,

$$
\begin{aligned}
\log q(\tau\alpha|x) - \log q(\alpha|x) &= \log H^{-1}(-\log(\tau\alpha)|x) - \log H^{-1}(-\log(\alpha)|x) \\
&= \theta(x)(\log_{-2}(\tau\alpha) - \log_{-2}(\alpha)) + \log\left(\frac{\ell(-\log(\tau\alpha)|x)}{\ell(-\log(\alpha)|x)}\right) \\
&\approx \theta(x)(\log_{-2}(\tau\alpha) - \log_{-2}(\alpha)) \approx \theta(x)\frac{\log(1/\tau)}{\log(1/\alpha)},
\end{aligned}
\tag{4}
$$

where $\log_{-2}(.) := \log\log(1/.)$, see Lemma 2 in Section 6 for a more precise asymptotic expansion. Hence, for a decreasing sequence $0 < \tau_J < \cdots < \tau_1 \leq 1$, where J is a positive integer, and for all functions ϕ satisfying the shift and location invariance condition

(A.1) $\phi : \mathbb{R}^J \to \mathbb{R}$ is a twice differentiable function such that $\phi(\eta z) = \eta\phi(z)$, $\phi(\eta u + z) = \phi(z)$ for all $\eta \in \mathbb{R} \setminus \{0\}$, $z \in \mathbb{R}^J$ and where $u = (1, \ldots, 1)^t \in \mathbb{R}^J$,

one has:
$$
\theta(x) \approx \log(1/\alpha)\frac{\phi(\log q(\tau_1\alpha|x), \ldots, \log q(\tau_J\alpha|x))}{\phi(\log(1/\tau_1), \ldots, \log(1/\tau_J))}.
\tag{5}
$$

Thus, the estimation of $\theta(x)$ relies on the estimation of conditional quantiles $q(.|x)$. This problem is addressed using a two-step estimator. First, $\bar{F}(y|x)$ is estimated by the kernel estimator defined for all $(x,y) \in E \times \mathbb{R}$ by

$$
\hat{\bar{F}}_n(y|x) = \sum_{i=1}^{n} K(d(x, X_i)/h_n)\mathbb{I}\{Y_i > y\} \Big/ \sum_{i=1}^{n} K(d(x, X_i)/h_n),
\tag{6}
$$

where $\mathbb{I}\{.\}$ is the indicator function and h_n is a nonrandom sequence such that $h_n \to 0$ as $n \to \infty$ (note that this bandwidth h_n may depend on x). The case of random neighbourhoods is addressed in [31] with k-Nearest Neighbours estimators. The function K is assumed to have a support included in $[0,1]$ such that $C_1 \leq K(t) \leq C_2$ for all $t \in [0,1]$ and for some constants $0 < C_1 < C_2 < \infty$. One may also assume without loss of generality that K integrates to one. In this case, K is called a type I kernel, see [22], Definition 4.1. Let $B(x, h_n)$ be the ball of center x and radius h_n, and introduce $\varphi_x(h_n) := \mathbb{P}(X \in B(x, h_n))$ the small ball probability of X. It is easily shown that the τ-th moment $\mu_x^{(\tau)}(h_n) := \mathbb{E}\{K^\tau(d(x, X)/h_n)\}$ can be controlled for all $\tau > 0$ as

$$
0 < C_1^\tau \varphi_x(h_n) \leq \mu_x^{(\tau)}(h_n) \leq C_2^\tau \varphi_x(h_n).
\tag{7}
$$

The estimation of $\bar{F}(y|x)$ is well-documented in the nonparametric literature, see for instance the seminal papers [33, 37] in the case where $E = \mathbb{R}^p$. The estimator was extended to the infinite dimensional setting for instance in [22], page 56. Its rate of uniform strong consistency is proved by [19] for a fixed value of y. Here, its asymptotic normality is established in Lemma 6 in Section 6 for $y = y_n \to \infty$ as $n \to \infty$, *i.e.* when estimating small tail probabilities from Weibull tail-distributions. Second, the kernel estimators of the conditional quantiles $q(\alpha|x)$ are defined via the generalized inverse of $\hat{\bar{F}}_n(.|x)$:

$$
\hat{q}_n(\alpha|x) = \hat{\bar{F}}_n^{\leftarrow}(\alpha|x) = \inf\{y, \ \hat{\bar{F}}_n(y|x) \leq \alpha\},
\tag{8}
$$

3

for all $\alpha \in (0,1)$. Many authors focused on the asymptotic properties of this estimator for fixed $\alpha \in (0,1)$. Weak and strong consistency are proved respectively in [37] and [23]. Asymptotic normality is shown in [5, 34, 38] when E is finite dimensional and by [20] for a general semi-metric space under dependence assumptions. Here, the asymptotic distribution of (8) is established when estimating extreme quantiles from Weibull tail-distributions, *i.e.* when $\alpha = \alpha_n \to 0$ as $n \to \infty$. This new result is established in Lemma 7, see Section 6.

Basing on (5), the considered family of estimators is then given by

$$\hat{\theta}_n(x) = \log(1/\alpha_n) \frac{\phi(\log \hat{q}_n(\tau_1 \alpha_n | x), \ldots, \log \hat{q}_n(\tau_J \alpha_n | x))}{\phi(\log(1/\tau_1), \ldots, \log(1/\tau_J))}, \tag{9}$$

with $\alpha_n \to 0$ as $n \to \infty$ and where the (extreme) conditional quantiles are estimated by (8). Model (2) is not sufficient to establish the asymptotic distribution of $\hat{\theta}_n(x)$, additional assumptions have to be made on $\ell(.|x)$. We further assume that

(A.2) $\ell(.|x)$ is a normalized and continuously differentiable slowly-varying function.

In such a case, the Karamata representation (see [9], Theorem 1.3.1) of the slowly-varying function can be written as

$$\ell(y|x) = c(x) \exp \left\{ \int_1^y \frac{\varepsilon(u|x)}{u} du \right\}, \tag{10}$$

where $c(x) > 0$ and $\varepsilon(u|x) \to 0$ as $u \to \infty$. The auxiliary function $\varepsilon(.|x)$ is thus continuous and given by $\varepsilon(y|x) = y\ell'(y|x)/\ell(y|x)$. In fact, **(A.2)** is equivalent to assuming (10) and the continuity of $\varepsilon(.|x)$. The function $\varepsilon(.|x)$ plays an important role in extreme-value theory since it drives the speed of convergence in (3) and more generally the bias of extreme-value estimators. Therefore, it may be of interest to specify how it converges to 0:

(A.3) $|\varepsilon(.|x)|$ is regularly varying with index $\rho(x) \leq 0$.

This assumption is used for instance by [1, 29] in the unconditional case and the estimation of the regular variation parameter ρ is addressed. Since the considered estimator involves a smoothing in the x direction, it is necessary to assess the regularity of the conditional survival function with respect to x. To this end, the oscillations are controlled by

$$\Delta \bar{F}(x, \alpha, \zeta, h) \quad := \quad \sup_{(x',\beta) \in B(x,h) \times [\alpha,\zeta]} \left| \frac{\bar{F}(q(\beta|x)|x')}{\bar{F}(q(\beta|x)|x)} - 1 \right| = \sup_{(x',\beta) \in B(x,h) \times [\alpha,\zeta]} \left| \frac{\bar{F}(q(\beta|x)|x')}{\beta} - 1 \right|,$$

where $(\alpha, \zeta) \in (0,1)^2$. Finally, let $v = (\log(1/\tau_1), \ldots, \log(1/\tau_J))^t \in \mathbb{R}^J$ and

$$\Lambda_n(x) = \left(n\alpha_n \frac{(\mu_x^{(1)}(h_n))^2}{\mu_x^{(2)}(h_n)} \right)^{-1/2}.$$

The following result establishes the asymptotic normality of estimator (9).

Theorem 1. *Under model (2), suppose* **(A.1)**–**(A.3)** *hold. Let* $x \in E$ *such that* $\varphi_x(h_n) > 0$ *where* $h_n \to 0$ *as* $n \to \infty$. *If* $\alpha_n \to 0$,

$$\sqrt{n\varphi_x(h_n)\alpha_n}\,\varepsilon(\log(1/\alpha_n)|x) \to \lambda \in \mathbb{R} \tag{11}$$

4

and there exists $\eta > 0$ such that $n\varphi_x(h_n)\alpha_n \to \infty$ and

$$\sqrt{n\varphi_x(h_n)\alpha_n}\{\Delta\bar{F}(x,(1-\eta)\tau_J\alpha_n,(1+\eta)\alpha_n,h_n) \vee 1/\log(1/\alpha_n)\} \to 0 \ as \ n \to \infty, \qquad (12)$$

then, $\Lambda_n^{-1}(x)(\hat{\theta}_n(x) - \theta(x)) \xrightarrow{d} \mathcal{N}(\mu_\phi, \theta^2(x)V_\phi)$ with

$$\mu_\phi = \lambda v^t \nabla \log \phi(v), \quad V_\phi = (\nabla \log \phi(v))^t \Sigma (\nabla \log \phi(v)). \qquad (13)$$

and where $\Sigma_{j,j'} = \tau_{j \wedge j'}^{-1}$ for $(j,j') \in \{1,\ldots,J\}^2$.

As pointed out in [14], $n\varphi_x(h_n)\alpha_n \to \infty$ is a necessary and sufficient condition for the almost sure presence of at least one sample point in the region $B(x,h_n) \times (q(\alpha_n|x),\infty)$ of $E \times \mathbb{R}$. This condition states that one cannot estimate small tail probabilities out of the sample using the kernel estimator (6). Let us also highlight that the asymptotic variance of the estimator is asymptotically or order $1/(n\varphi_x(h_n)\alpha_n)$. The convergence rate thus directly depends on the small ball probability $\varphi_x(h_n)$. To overcome the sensitiveness of the method to dimensionality effects or to the choice of the semi-metric, one can use dimension reduction techniques such as single index models, see for instance [28]. Nevertheless, the theoretical properties of such methods are not yet established in the extreme framework. Condition (12) imposes that the (squared) bias induced by the smoothing should be negligible compared to the variance of the estimator. Let us note that condition (11) is standard in the extreme-value framework. Neglecting the slowly-varying function in the construction (4) of $\hat{\theta}_n(x)$ yields a bias that should be of the same order as the asymptotic standard deviation of the estimator. Moreover, when $\lambda \neq 0$, conditions (11) and (12) yield $\log(1/\alpha_n)\varepsilon(\log(1/\alpha_n)|x) \to \infty$ as $n \to \infty$ which, in turn, implies $\rho(x) > -1$. The next corollary provides some possible choices of α_n and h_n sequences under Hölder conditions.

Corollary 1. *Suppose (2), (A.1)–(A.3) hold. Let $x \in E$ such that $\varphi_x(h_n) > 0$ and $y\varepsilon(y|x) \to \infty$ as $y \to \infty$. Assume there exist positive constants L_θ, L_c et L_ε such that, for all $(x,x') \in E^2$,*

$$\left|\frac{1}{\theta(x)} - \frac{1}{\theta(x')}\right| \leq L_\theta d(x,x'),$$
$$\left|\log c(x) - \log c(x')\right| \leq L_c d(x,x'), \qquad (14)$$
$$\sup_{u \in [1,\bar{y}_n(x)]} \left|\varepsilon(u|x) - \varepsilon(u|x')\right| \leq L_\varepsilon d(x,x'),$$

where $\bar{y}_n(x) := \sup\{H(q(\alpha_n|x)|x'), \ x' \in B(x,h_n)\}$. Suppose there exists $\xi > 0$ small enough such that $\varphi_x^{-1}(1/y)(\log y)^{1+\xi-\rho(x)} \to 0$ as $y \to \infty$. Then, letting $\lambda > 0$,

$$\alpha_n = n^{-1+\xi} \ and \ h_n = \varphi_x^{-1}\left(\lambda(1-\xi)^{2\rho(x)}n^{-\xi}(\varepsilon(\log n|x))^{-2}\right), \qquad (15)$$

the assumptions of Theorem 1 hold and therefore $\Lambda_n^{-1}(x)(\hat{\theta}_n(x) - \theta(x)) \xrightarrow{d} \mathcal{N}(\mu_\phi, \theta^2(x)V_\phi)$.

The key assumption here is $\varphi_x^{-1}(1/y)(\log y)^{1+\xi-\rho(x)} \to 0$ as $y \to \infty$. It can be shown that this condition holds in the finite dimensional setting (when X has a continuous density with respect to the Lebesgue measure), or for fractal-type and some exponential-type processes, see [22], Chapter 13. Finally, the asymptotic variance V_ϕ does not depend on the distribution of $Y|X = x$. It is thus possible to look for functions ϕ and for sequences $\tau_J < \cdots < \tau_1$ minimizing V_ϕ. This problem is addressed in the next section.

3 Example

In this section, we focus on the particular family of functions $\phi^{(p)}(z) = \left(\sum_{j=2}^{J} \beta_j (z_j - z_1)^p\right)^{1/p}$, where $z = (z_1, \ldots, z_J)^t \in \mathbb{R}^J$, $p \in \mathbb{N} \setminus \{0\}$ and for all $j \in \{2, \ldots, J\}$, $\beta_j \in \mathbb{R}$. It is clear that assumption **(A.1)** is satisfied and the corresponding estimator of θ writes:

$$\hat{\theta}_n^{(p)}(x) = \log(1/\alpha_n) \left(\sum_{j=2}^{J} \beta_j \left[\log \hat{q}_n(\tau_j \alpha_n | x) - \log \hat{q}_n(\tau_1 \alpha_n | x)\right]^p \Big/ \sum_{j=2}^{J} \beta_j [\log(\tau_1/\tau_j)]^p\right)^{1/p}. \quad (16)$$

Taking $\beta_2 = \ldots = \beta_J$ leads to an estimator analogous to the one proposed in [35] for unconditional heavy-tailed distributions. If furthermore $p = 1$, (16) corresponds to a functional version of the estimator proposed in [24]:

$$\hat{\theta}_n^{(1)}(x) = \log(1/\alpha_n) \sum_{j=1}^{J} \left[\log \hat{q}_n(\tau_j \alpha_n | x) - \log \hat{q}_n(\alpha_n | x)\right] \Big/ \sum_{j=1}^{J} \log(1/\tau_j) .$$

It can also be read as an adaptation of the kernel tail-index estimator introduced in [14] in a finite dimensional setting for conditional heavy-tailed distributions to conditional Weibull tail-distributions. As a consequence of Theorem 1, the associated asymptotic mean and variance of $\hat{\theta}_n^{(p)}(x)$ defined in (13) are given for an arbitrary vector β by $\mu = \lambda$ and

$$V^{(p)} = \frac{(\eta^{(p)})^t A \Sigma A^t \eta^{(p)}}{(\eta^{(p)})^t A v v^t A^t \eta^{(p)}},$$

where A is the $(J-1) \times J$ matrix defined by $A = [I_{J-1}, -u]$ with I_{J-1} the $(J-1) \times (J-1)$ identity matrix and $\eta^{(p)} = (\beta_j(v_j - v_1), j = 2, \ldots, J)^t$. Let us emphasize that the asymptotic bias μ does not depend neither on p and nor on the weights $\{\beta_j, j = 2, \ldots, J\}$. At the opposite, the asymptotic variance $V^{(p)}$ depends both on p and on the weights but only through the vector $\eta^{(p)}$. It is thus possible to minimize $V^{(p)}$ with respect to $\eta^{(p)}$:

Proposition 1. *The asymptotic variance of $\hat{\theta}_n^{(p)}(x)$ is minimal for $\eta^{(p)}$ proportional to $\eta_{\mathrm{opt}} = (A\Sigma A^t)^{-1} A v$ and is given by*

$$V_{\mathrm{opt}} = \frac{1}{(Av)^t \ (A\Sigma A^t)^{-1} Av}.$$

Let us highlight that V_{opt} is independent of p. Moreover, for a fixed value of J, it is possible to minimize numerically the optimal variance V_{opt} given by Proposition 1 with respect to parameters $0 < \tau_J < \cdots < \tau_1 \leq 1$. The resulting values of V_{opt} are displayed in Table 1. The finite sample performance of the associated estimators are illustrated on simulated data in Section 4.

4 Numerical experiments

In this section, the finite sample performance of the estimator (16) associated with optimal weights is investigated. For $j = 2, \ldots, J$, we thus take $\beta_j = \eta_{\mathrm{opt},j}(v_j - v_1)^{1-p}$ where the vector $\eta_{\mathrm{opt}} \in \mathbb{R}^{J-1}$ is given in Proposition 1. Besides, the sequence $\tau_1, \ldots, \tau_J$ is selected by minimizing the asymptotic variance V_{opt}, see Table 1. Some preliminary simulations showed that the choice of a length $J > 5$

did not improve the results, we thus take in the following $J = 5$. The value of p is taken in the set $\{1, 2, 3\}$. The estimator (6) of the survival function is computed with a modified version of the bi-quadratic kernel given by

$$K(u) = \frac{10}{9} \left(\frac{3}{2} \left(1 - u^2\right)^2 + \frac{1}{10} \right) \mathbb{I}\{|u| \leq 1\}.$$

Note that, as required, K is a type I kernel. The set E considered here is a subset of $L^2([0, 1])$ made of trigonometric functions $\psi_z : [0, 1] \mapsto [0, 1]$, $\psi_z(t) = \cos(2\pi z t)$ with different periods indexed by $z \in [1/10, 1/2]$. Two semi-metrics are considered: $d_1(\psi_z, \psi_{z'}) = \left| \|\psi_z\|_2^2 - \|\psi_{z'}\|_2^2 \right|$ and $d_2(\psi_z, \psi_{z'}) = \|\psi_z - \psi_{z'}\|_2$, for all $(z, z') \in [1/10, 1/2]^2$, where

$$\|\psi_z\|_2^2 = \int_0^1 \psi_z^2(t) dt = \frac{1}{2} \left(1 + \frac{\sin(4\pi z)}{4\pi z} \right).$$

The semi-metric d_2 is built on the classical L_2 norm while d_1 measures some spacing between the periods of the trigonometric functions. The covariate X is chosen randomly on E by considering $X = \psi_Z$ where Z is a uniform random variable on $[1/10, 1/2]$. Some examples of simulated random functions X are depicted on the left panel of Figure 1. For a given function $x \in E$, the generalized inverse of the conditional hazard function $H(.|x)$ is given for $y \geq 0$ by $H^{\leftarrow}(y|x) = y^{\theta(x)} \left(1 - \gamma y^{\rho(x)} \right)$, with $\gamma = 1/10$. Note that in this situation $\ell(y|x) = 1 - \gamma y^{\rho(x)}$ and $\varepsilon(y|x) \sim -\gamma \rho(x) y^{\rho(x)}$, and thus assumption (A.3) holds. As it can be seen in Theorem 1, $\rho(x)$ and $\theta(x)$ tune the difficulty of the problem. For small values of $|\rho(x)|$, approximation (4) becomes unreliable and some bias is expected in the estimation while for large values of $\theta(x)$ the asymptotic variance of $\hat{\theta}_n(x)$ increases. To illustrate the effect of these two parameters, the functions defined on E by $\tilde{\theta}(x) = (20\|x\|_2^2 + 1)^{-1} + 1/40$ and $\tilde{\rho}(x) = 50/(60\|x\|_2^2 + 3) - 5/2$ are introduced and the following three situations are considered: (S.1) $\rho(x) = -1$ and $\theta(x) = \tilde{\theta}(x)$; (S.2) $\rho(x) = \tilde{\rho}(x)$ and $\theta(x) = 1/10$ and (S.3) $\rho(x) = \tilde{\rho}(x)$ and $\theta(x) = \tilde{\theta}(x)$. For each case, $N = 100$ samples of size $n = 1000$ were generated. The selection of the two remaining parameters h_n and α_n is addressed in the next paragraph.

In order to show the advantage of using the covariate information, $\hat{\theta}_n(x)$ is compared to the non-conditional estimator proposed in [24] and defined by:

$$\hat{\theta}_n^{NCE} = \sum_{i=1}^{k_n} \left(\log Y_{n-i+1,n} - \log Y_{n-k_n+1,n} \right) \Bigg/ \sum_{i=1}^{k_n} \left(\log_{-2}(n/i) - \log_{-2}(n/k_n) \right), \qquad (17)$$

with $k_n = 250$ and where $Y_{1,n} \leq \ldots \leq Y_{n,n}$ are the ordered statistics.

Selection of the bandwidth h_n and the sequence α_n. The choice of the smoothing parameter h_n is a recurrent issue in nonparametric estimation. We refer to [32] for a functional version of the cross-validation criterion and to [36] for a Bayesian approach. Besides, the selection of α_n is equivalent to the choice of the number of upper order statistics in the non-conditional extreme-value theory. It is still an open question, even though some techniques have been proposed, see for instance [13] for a bootstrap based method. We propose a procedure to select simultaneously h_n and α_n basing on the following remark. For a fixed x, let $\{Z_1(x, h_n), \ldots, Z_{m_n}(x, h_n)\}$ be the m_n random values Y_i for which $X_i \in B(x, h_n)$. According to the result obtained in [16], equation (6), if the sequences h_n and α_n are well chosen then the rescaled log-spacings

$$\{i \log(m_n/i)(\log Z_{m_n-i+1,m_n}(x, h_n) - \log Z_{m_n-i,m_n}(x, h_n)), \ i = 1, \ldots, \lfloor m_n \alpha_n \rfloor\},$$

should be approximately independent with an exponential distribution of parameter $\theta(x)$. For each value of h in the set $\mathcal{H} = \{h_i, \ i = 1, \ldots, M_h\}$ and α in the set $\mathcal{A} = \{\alpha_j, \ j = 1, \ldots, M_\alpha\}$ a Kolmogorov-Smirnov test is performed and the obtained p-value $KS(h, \alpha)$ is recorded. We then select the optimal sequences h_{opt} and α_{opt} such that

$$(h_{opt}, \alpha_{opt}) = \underset{h \in \mathcal{H}, \ \alpha \in \mathcal{A}}{\arg\max} \ KS(h, \alpha).$$

Let us highlight that this procedure is intrinsically location-adaptive since the selected values h_{opt} and α_{opt} both depend on x. In practice, we take $M_h = M_\alpha = 30$, $\mathcal{A} = \{0.05, \ldots, 0.3\}$, $\mathcal{H} = \{0.03, \ldots, 0.2\}$ for distance d_1 and $\mathcal{H} = \{0.03, \ldots, 1\}$ for distance d_2.

Results. Boxplots of the estimations obtained with the non conditional estimator (17) and the proposed estimator $\hat{\theta}_n^{(p)}(x)$ with $p = 1$ and $p = 3$ are represented on Figure 2 (distance d_1) and Figure 3 (distance d_2). As expected, it appears that the variance is larger for large values of $\theta(x)$ (first row of Figure 2). At the opposite, the bias seems approximately independent of $\rho(x)$ (second row of Figure 2). One can also observe that the choice $p = 3$ yields slightly better results than $p = 1$. The non-conditional estimator (17) provides poor results compared to $\hat{\theta}_n^{(p)}(x)$. It is thus essential to take the covariate information into account. Finally, comparing the results obtained with d_1 and d_2, it appears that the use of d_1 provides better result than d_2. This conclusion could be expected since the trigonometric functions of E only differ from their periods and d_1 does measure some spacing between these periods.

Conclusion. To summarize, the estimator $\hat{\theta}_n^{(p)}(x)$ gives satisfying results when combined with optimal weights (see Proposition 1 and Table 1). It is fully data-driven thanks to an automatic selection procedure of the bandwidth h_n and of the sequence α_n. The estimator is very sensitive with respect to the choice of the semi-metric d. We refer to [22], Chapter 3, for a discussion on the definition of a well-adapted semi-metric to a particular estimation problem. A possible extension of this work would be to implement a bias reduction technique based on the estimation of the index of regular variation in **(A.3)** similarly to [12] in the unconditional setting and for heavy-tailed distributions. However, as stressed in [2], this adaptation should be conducted with great care.

5 Illustration on real data

The behaviour of the proposed Weibull tail-coefficient estimator (with $p = 3$) is illustrated on spectrometric data. These data contain $n = 215$ near infrared spectra of absorbance ($x_i, \ i = 1, \ldots, 215$) observed on finely chopped pieces of meat and discretized at 100 wavelengths. This dataset is often considered in papers dedicated to nonparametric functional data analysis, see for instance [21, 22]. For each spectrometric curve x_i, the percentage of fat content $\tilde{y}_i \in [0, 100]$ is recorded. Since the fat contents are bounded, the following transformed variables are considered: $y_i := \log(100/\tilde{y}_i)$ for all $i \in \{1, \ldots, n\}$. Denoting by $y_{1,n} \leq \ldots \leq y_{n,n}$ the ordered transformed observations and by $x_{(1)}, \ldots, x_{(n)}$ their corresponding spectra, we propose to estimate the functional Weibull tail-coefficients associated with the curves $\{\tilde{x}_\beta, \ \beta = 0, 1/4, 1/2, 3/4, 1\}$ where $\tilde{x}_\beta$ is the mean of the curves $\{x_{(i)}, \ i \in I_\beta\}$ with $I_0 = \{n - 9, \ldots, n\}$, $I_1 = \{1, \ldots, 10\}$ and $I_\beta = \{\lfloor n(1 - \beta) \rfloor - 4, \ldots, \lceil n(1 - \beta) \rceil + 4\}$ for $\beta \in \{1/4, 1/2, 3/4\}$. Here, $\lfloor . \rfloor$ and $\lceil . \rceil$ are the floor and ceiling functions. For example, the curve $\tilde{x}_0$ is obtained by averaging the spectra associated with the 10 largest

transformed variables (or equivalently the 10 smallest fat contents). These curves are represented on the right panel of Figure 1. A natural semi-metric to work with is

$$d^2(x_1, x_2) = \int \{x_1^{(2)}(t) - x_2^{(2)}(t)\}^2 dt, \quad (x_1, x_2) \in E^2,$$

where $x^{(2)}$ denotes the second derivative of x, see [22], Chapter 9. Let us now investigate the goodness-of-fit of model (2) to the observations $\{(x_i, y_i), \ i = 1, \ldots, 215\}$. For each curve $\tilde{x}_\beta$, hyperparameters h_n and α_n are selected following the procedure described in Section 4. Using these selected values, the estimated conditional quantile $\hat{q}_n(\alpha_n|\tilde{x}_\beta)$ is computed for each $\beta \in \{0, 1/4, 1/2, 3/4, 1\}$ and the statistics $\{Z_1, \ldots, Z_{m_\beta}\}$ corresponding to the y_i's such that $x_i \in B(\tilde{x}_\beta, h_n)$ and $y_i > \hat{q}_n(\alpha_n|\tilde{x}_\beta)$ are collected. According to (4), under model (2), the points

$$\left\{ \left(\log_{-2} \hat{\bar{F}}_n(Z_{m_\beta-i+1,m_\beta}|\tilde{x}_\beta) - \log_{-2} \alpha_n \ ; \ \log Z_{m_\beta-i+1,m_\beta} - \log \hat{q}_n(\alpha_n|\tilde{x}_\beta) \right), \ i = 1, \ldots, m_\beta \right\},$$

should approximately lie on a straight line. For each β, these points are represented on Figure 4 (5 upper left panels) and it appears that model (2) fits reasonably well the observations. For each curve $\tilde{x}_\beta$, the estimator of the functional Weibull tail-coefficient is depicted on Figure 4 (bottom right panel). It appears that the smaller the fat content, the smaller the functional Weibull tail-coefficient. In other words, heaviest tails are found for small values of fat contents. In terms of tail behaviour, the sample is thus clearly heterogeneous.

6 Proofs

6.1 Preliminary results

Let us start with three analytical results. The first lemma provides a sufficient condition under which $\bar{F}(.|x)$ preserves the equivalence property between sequences.

Lemma 1. *Suppose (2) and* **(A.2)** *hold. Let* (u_n) *and* (v_n) *be sequences such that* $u_n \to \infty$ *and*

$$H(u_n|x) \left(\frac{v_n}{u_n} - 1 \right) \to 0,$$

as $n \to \infty$. *Then,* $\bar{F}(u_n|x)/\bar{F}(v_n|x) \to 1$ *as* $n \to \infty$.

Proof of Lemma 1. Under **(A.2)**, $H(.|x)$ is continuously differentiable and a first order Taylor expansion yields that there exists w_n between u_n and v_n such that:

$$\bar{F}(u_n|x)/\bar{F}(v_n|x) = \exp\{H(v_n|x) - H(u_n|x)\} = \exp\{(v_n - u_n)H'(w_n|x)\}.$$

Besides, under **(A.2)**, $(H^{-1})'(y|x) = y^{\theta(x)-1}\ell(y|x)(\theta(x)+\varepsilon(y|x))$ and thus $(H^{-1})'(.|x)$ is a regularly varying function with index $\theta(x) - 1$. Moreover, since $H'(.|x) = 1/(H^{-1})'(H(.|x)|x)$, it is clear that $H'(.|x)$ is regularly varying with index $1/\theta(x) - 1$. As a consequence, $w_n \sim u_n$ implies $H'(w_n|x) \sim H'(u_n|x)$ as $n \to \infty$. Thus,

$$(v_n - u_n)H'(w_n|x) \sim H(u_n|x) \left(\frac{v_n}{u_n} - 1 \right) \frac{u_n H'(u_n|x)}{H(u_n|x)} \sim H(u_n|x) \left(\frac{v_n}{u_n} - 1 \right) \frac{1}{\theta(x)},$$

concludes the proof. ∎

9

The second lemma establishes an asymptotic expansion of the log-spacing between two extreme conditional quantiles. It justifies the heuristic approximation used in (4).

Lemma 2. *Suppose (2), (A.2) and (A.3) hold. Let $0 < \tau \leq 1$ and $\alpha_n \to 0$ as $n \to \infty$. Then,*

$$\log q(\tau \alpha_n | x) - \log q(\alpha_n | x) = \frac{\log(1/\tau)}{\log(1/\alpha_n)} \left(\theta(x) + \varepsilon(\log(1/\alpha_n)|x)(1 + o(1)) + O\left(\frac{1}{\log(1/\alpha_n)} \right) \right).$$

Proof of Lemma 2. In view of (2), we have

$$
\begin{aligned}
\Delta_n &:= \log q(\tau \alpha_n | x) - \log q(\alpha_n | x) \\
&= \theta(x)[\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n)] + \log \ell(\log(1/\tau\alpha_n)|x) - \log \ell(\log(1/\alpha_n)|x) \\
&= \theta(x)[\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n)] + \psi(\log(1/\tau\alpha_n)|x) - \psi(\log(1/\alpha_n)|x),
\end{aligned}
$$

where $\psi(.|x) := \log \ell(.|x)$. Under **(A.2)**, $\psi(.|x)$ is continuously differentiable and a first order Taylor expansion thus yields that there exists $\eta_n \in (\tau \alpha_n, \alpha_n)$ such that

$$
\begin{aligned}
\Delta_n &= \theta(x)[\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n)] + \log(1/\tau)\psi'(\log(1/\eta_n)|x) \\
&= \theta(x)[\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n)] + \frac{\log(1/\tau)}{\log(1/\eta_n)}\varepsilon(\log(1/\eta_n)|x),
\end{aligned}
$$

where $\varepsilon(.|x)$ is defined in (10). Since $\log(1/\eta_n)$ is asymptotically equivalent to $\log(1/\alpha_n)$, **(A.3)** implies $\varepsilon(\log(1/\eta_n)|x) = \varepsilon(\log(1/\alpha_n)|x)(1 + o(1))$ so that

$$\Delta_n = \theta(x)[\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n)] + \frac{\log(1/\tau)}{\log(1/\alpha_n)}\varepsilon(\log(1/\alpha_n)|x)(1 + o(1)).$$

Finally, remarking that

$$\log_{-2}(\tau \alpha_n) - \log_{-2}(\alpha_n) = \frac{\log(1/\tau)}{\log(1/\alpha_n)} + O\left(\frac{1}{\log^2(1/\alpha_n)} \right),$$

the conclusion follows. ∎

Lemma 3 controls the oscillations of the conditional survival function under Hölder assumptions.

Lemma 3. *Suppose (2), (14) and (A.2) hold. Let (h_n), (α_n) and (ξ_n) be sequences converging to 0 such that $\alpha_n < \xi_n$ and $h_n \log(\alpha_n) \log_{-2}(\alpha_n) \to 0$ as $n \to \infty$. Then, $\bar{\Delta}F(x, \alpha_n, \xi_n, h_n) = O\left(h_n \log(1/\alpha_n) \log_{-2}(\alpha_n) \right).$*

Proof of Lemma 3. Let $y_n(x) := q(\beta_n | x) \to \infty$ and remark that the oscillation can be rewritten as $\bar{\Delta}F(x, \alpha_n, \xi_n, h_n) = \sup\{\delta_n(x', \beta_n), \ (x', \beta_n) \in B(x, h_n) \times [\alpha_n, \xi_n]\}$ where

$$\delta_n(x', \beta_n) := \left| \frac{\bar{F}(y_n(x)|x')}{\bar{F}(y_n(x)|x)} - 1 \right| = \frac{\bar{F}(y_n(x)|x') \vee \bar{F}(y_n(x)|x)}{\bar{F}(y_n(x)|x)} \left(1 - \frac{\bar{F}(y_n(x)|x') \wedge \bar{F}(y_n(x)|x)}{\bar{F}(y_n(x)|x') \vee \bar{F}(y_n(x)|x)} \right).$$

The inequality $1 - z \leq \log(1/z)$ for all $z \in (0, 1]$ entails

$$\delta_n(x', \beta_n) \leq \frac{\bar{F}(y_n(x)|x') \vee \bar{F}(y_n(x)|x)}{\bar{F}(y_n(x)|x)} \left| \log \frac{\bar{F}(y_n(x)|x')}{\bar{F}(y_n(x)|x)} \right|,$$

10

and, in view of $H(y|x) = -\log \bar{F}(y|x)$ it follows that

$$
\begin{aligned}
\left|\log \frac{\bar{F}(y_n(x)|x')}{\bar{F}(y_n(x)|x)}\right| &= \left(H(y_n(x)|x') \vee H(y_n(x)|x)\right)\left(1 - \frac{H(y_n(x)|x') \wedge H(y_n(x)|x)}{H(y_n(x)|x') \vee H(y_n(x)|x)}\right) \\
&\leq \left(H(y_n(x)|x') \vee H(y_n(x)|x)\right)\left|\log \frac{H(y_n(x)|x')}{H(y_n(x)|x)}\right|.
\end{aligned} \tag{18}
$$

Now, since $H(\cdot|x)$ is strictly increasing, it is easy to check that $H(y|x) = (y/\ell(H(y|x)|x))^{1/\theta(x)}$ for all $(x,y) \in E \times \mathbb{R}$. Thus, under (A.2),

$$
\log H(y_n(x)|x') = \frac{1}{\theta(x')}\left(\log y_n(x) - \log c(x') - \int_1^{H(y_n(x)|x')} \frac{\varepsilon(u|x')}{u}du\right)
$$

and the Hölder assumption on $\varepsilon(u|\cdot)$ yields

$$
\left|\int_1^{H(y_n(x)|x)} \frac{\varepsilon(u|x)}{u}du - \int_1^{H(y_n(x)|x')} \frac{\varepsilon(u|x')}{u}du\right| \leq h_n \log H(y_n(x)|x) + \eta_n\left|\log \frac{H(y_n(x)|x')}{H(y_n(x)|x)}\right|,
$$

where $\eta_n \to 0$ uniformly on $x' \in B(x, h_n)$. Moreover, since $H(\cdot|x)$ is a regularly varying function with index $1/\theta(x)$, it follows that $\theta(x)\log H(y_n(x)|x) = \log y_n(x)(1 + o(1))$. As a first conclusion, the regularity assumptions on $\theta(\cdot)$ and $c(\cdot)$ lead to

$$
\left|\log H(y_n(x)|x') - \log H(y_n(x)|x)\right| = O\left(h_n \log y_n(x)\right).
$$

Let us note that $\log y_n(x) = \theta(x)\log_{-2}(\beta_n) + \log \ell(\log(1/\beta_n)|x) = \theta(x)\log_{-2}(\beta_n)(1 + o(1))$ as $n \to \infty$ and therefore

$$
\left|\log H(y_n(x)|x') - \log H(y_n(x)|x)\right| = \mathcal{O}\left(h_n \log_{-2}(\alpha_n)\right) = o(1).
$$

This result implies that $\left(H(y_n(x)|x') \vee H(y_n(x)|x)\right) \leq 2H(y_n(x)|x) = 2\log(1/\beta_n) \leq 2\log(1/\alpha_n)$, and thus, from (18),

$$
\left|\log \bar{F}(y_n(x)|x') - \log \bar{F}(y_n(x)|x)\right| = O\left(h_n \log(1/\alpha_n)\log_{-2}(\alpha_n)\right) = o(1).
$$

As a consequence, $\bar{F}(y_n(x)|x')$ is uniformly equivalent to $\bar{F}(y_n(x)|x)$ and hence

$$
\delta_n(x', \beta_n) = O\left(h_n \log(1/\alpha_n)\log_{-2}(\alpha_n)\right),
$$

which concludes the proof. ∎

We are now interested in the asymptotic normality of $\hat{\bar{F}}_n(y_n|x)$ when $y_n \to \infty$. First, let us remark that the kernel estimator (6) can be rewritten as $\hat{\bar{F}}_n(y|x) = \hat{\psi}_n(y,x)/\hat{g}_n(x)$ with

$$
\begin{aligned}
\hat{\psi}_n(y,x) &= \frac{1}{n\mu_x^{(1)}(h_n)}\sum_{i=1}^n K(d(x,X_i)/h_n)\mathbb{I}\{Y_i > y\}, \\
\hat{g}_n(x) &= \frac{1}{n\mu_x^{(1)}(h_n)}\sum_{i=1}^n K(d(x,X_i)/h_n).
\end{aligned}
$$

The next lemmas are dedicated to the asymptotic properties of $\hat{g}_n$ and $\hat{\psi}_n$.

Lemma 4. *Let $x \in E$ such that $\varphi_x(h_n) > 0$. If $\varphi_x(h_n) \to 0$ and $n\varphi_x(h_n) \to \infty$ as $n \to \infty$, then $\hat{g}_n(x) = 1 + O_{\mathbb{P}}\left((n\varphi_x(h_n))^{-1/2}\right)$.*

11

Proof of Lemma 4. It is clear that $\mathbb{E}(\hat{g}_n(x)) = 1$. Besides, standard calculations yields

$$n\varphi_x(h_n)\text{var}(\hat{g}_n(x)) = \varphi_x(h_n)\left(\frac{\mu_x^{(2)}(h_n)}{(\mu_x^{(1)}(h_n))^2} - 1\right)$$

and (7) entails $(C_1/C_2)^2 \leq \varphi_x(h_n)\mu_x^{(2)}(h_n)/(\mu_x^{(1)}(h_n))^2 \leq (C_2/C_1)^2$. Condition $\varphi_x(h_n) \to 0$ concludes the proof. ∎

Lemma 5. *Suppose (2) and* **(A.2)** *hold. Let $x \in E$ such that $\varphi_x(h_n) > 0$ where $h_n \to 0$ as $n \to \infty$ and introduce $0 < \tau_J < \cdots < \tau_1 \leq 1$ where J is a positive integer. Consider $\alpha_n \to 0$ such that $n\varphi_x(h_n)\alpha_n \to \infty$ and $\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n) \to 0$ as $n \to \infty$ for some $\eta > 0$. Let $y_{n,j} := q(\tau_j\alpha_n|x)(1 + o(1/\log(1/\alpha_n)))$ for all $j = 1, \ldots, J$. Then,*

(i) $\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) = \bar{F}(y_{n,j}|x)(1 + O(\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n)))$, *for $j = 1, \ldots, J$.*

(ii) *The random vector*

$$\left\{\Lambda_n^{-1}(x)\left(\frac{\hat{\psi}_n(y_{n,j}, x) - \mathbb{E}(\hat{\psi}_n(y_{n,j}, x))}{\bar{F}(y_{n,j}|x)}\right)\right\}_{j=1,\ldots,J}$$

is asymptotically Gaussian, centered, with covariance matrix Σ where $\Sigma_{j,j'} = \tau_{j\wedge j'}^{-1}$ for $(j, j') \in \{1, \ldots, J\}^2$.

Proof of Lemma 5. **(i)** Since the (X_i, Y_i), $i = 1, \ldots, n$ are identically distributed, we have

$$\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) = \frac{1}{\mu_x^{(1)}(h_n)}\mathbb{E}\{K(d(x, X)/h_n)\mathbb{I}\{Y > y_{n,j}\}\} = \frac{1}{\mu_x^{(1)}(h_n)}\mathbb{E}\{K(d(x, X)/h_n)\bar{F}(y_{n,j}|X)\}.$$

Let us now consider

$$
\begin{aligned}
\varepsilon_{n,j} &= \mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) - \bar{F}(y_{n,j}|x) = \frac{1}{\mu_x^{(1)}(h_n)}\mathbb{E}(K(d(x, X)/h_n)(\bar{F}(y_{n,j}|X) - \bar{F}(y_{n,j}|x)))\\
&= \frac{\bar{F}(y_{n,j}|x)}{\mu_x^{(1)}(h_n)}\mathbb{E}\left(K(d(x, X)/h_n)\left(\frac{\bar{F}(y_{n,j}|X)}{\bar{F}(y_{n,j}|x)} - 1\right)\right).
\end{aligned}
$$

In view of Lemma 1, $\bar{F}(y_{n,j}|x) \sim \tau_j\alpha_n$ and thus $\bar{F}(y_{n,j}|x) \in [\tau_J(1-\eta)\alpha_n, (1+\eta)\alpha_n]$ eventually, for all $j = 1, \ldots, J$. Consequently,

$$\left|\frac{\bar{F}(y_{n,j}|X)}{\bar{F}(y_{n,j}|x)} - 1\right|\mathbb{I}\{d(x, X) \leq h_n\} \leq \Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n)$$

eventually and therefore $|\varepsilon_{n,j}| \leq \bar{F}(y_{n,j}|x)\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n)$ which concludes the first part of the proof.

(ii) Let $\beta \neq 0$ in $\mathbb{R}^J$ and consider the random variable

$$\Psi_n := \sum_{j=1}^J \beta_j\left(\frac{\hat{\psi}_n(y_{n,j}, x) - \mathbb{E}(\hat{\psi}_n(y_{n,j}, x))}{\Lambda_n(x)\bar{F}(y_{n,j}|x)}\right) =: \sum_{i=1}^n Z_{i,n}$$

12

where, for $i = 1, \ldots, n$, $Z_{i,n}$ is given by

$$\frac{1}{n\Lambda_n(x)\mu_x^{(1)}(h_n)} \left\{ \sum_{j=1}^{J} \frac{\beta_j K(d(x, X_i)/h_n)\mathbb{I}\{Y_i \geq y_{n,j}\}}{\bar{F}(y_{n,j}|x)} - \mathbb{E}\left(\sum_{j=1}^{J} \frac{\beta_j K(d(x, X_i)/h_n)\mathbb{I}\{Y_i \geq y_{n,j}\}}{\bar{F}(y_{n,j}|x)}\right) \right\}$$

and is a set of centered, independent and identically distributed random variables with variance

$$\text{var}(Z_{i,n}) = \frac{1}{n^2 \Lambda_n^2(x)(\mu_x^{(1)}(h_n))^2} \text{var}\left(\sum_{j=1}^{J} \beta_j K(d(x, X_i)/h_n) \frac{\mathbb{I}\{Y_i \geq y_{n,j}\}}{\bar{F}(y_{n,j}|x)}\right) =: \frac{\alpha_n}{n\mu_x^{(2)}(h_n)} \beta^t B\beta,$$

where B is the $J \times J$ covariance matrix with coefficients defined for $(j, j') \in \{1, \ldots, J\}^2$ by

$$
\begin{aligned}
B_{j,j'} &= \frac{A_{j,j'}}{\bar{F}(y_{n,j}|x)\bar{F}(y_{n,j'}|x)}, \\
A_{j,j'} &= \text{cov}\left(K\left(d(x, X)/h_n\right)\mathbb{I}\{Y \geq y_{n,j}\}, \; K\left(d(x, X)/h_n\right)\mathbb{I}\{Y \geq y_{n,j'}\}\right) \\
&= \mathbb{E}\left(K^2\left(d(x, X)/h_n\right)\mathbb{I}\{Y \geq y_{n,j} \vee y_{n,j'}\}\right) \\
&\quad - \mathbb{E}(K(d(x, X)/h_n)\mathbb{I}\{Y \geq y_{n,j}\})\mathbb{E}(K(d(x, X)/h_n)\mathbb{I}\{Y \geq y_{n,j'}\}),
\end{aligned}
$$

where K^2 is a kernel with the same properties as K. Consequently, the three above expectations are of the same nature. Remarking that eventually $y_{n,j} \vee y_{n,j'} = y_{n,j\vee j'}$, part **(i)** of the proof implies

$$
\begin{aligned}
A_{j,j'} &= \mu_x^{(2)}(h_n)\bar{F}(y_{n,j\vee j'}|x)(1 + O(\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n))) \\
&\quad - (\mu_x^{(1)}(h_n))^2 \bar{F}(y_{n,j}|x)\bar{F}(y_{n,j'}|x)(1 + O(\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n)))
\end{aligned}
$$

leading to

$$
\begin{aligned}
B_{j,j'}\frac{\bar{F}(y_{n,j\wedge j'}|x)}{\mu_x^{(2)}(h_n)} = 1 \;\; &+ \;\; O(\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n)) \\
&- \;\; \frac{(\mu_x^{(1)}(h_n))^2}{\mu_x^{(2)}(h_n)}\bar{F}(y_{n,j\wedge j'}|x)(1 + O(\Delta\bar{F}(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n))).
\end{aligned}
$$

In view of (7), $(\mu_x^{(1)}(h_n))^2/\mu_x^{(2)}(h_n)$ is bounded and thus $\bar{F}(y_{n,j\wedge j'}|x) \to 0$ as $n \to \infty$ yields

$$B_{j,j'} = \frac{\mu_x^{(2)}(h_n)}{\bar{F}(y_{n,j\wedge j'}|x)}(1 + o(1)).$$

Now, Lemma 1 implies that $\bar{F}(y_{n,j\wedge j'}|x) \sim \bar{F}(q(\tau_{j\wedge j'}\alpha_n|x)|x) = \tau_{j\wedge j'}\alpha_n$ entailing

$$B_{j,j'} = \frac{\mu_x^{(2)}(h_n)\Sigma_{j,j'}}{\alpha_n}(1 + o(1)),$$

and therefore, $\text{var}(Z_{i,n}) \sim \beta^t\Sigma\beta/n$, for all $i = 1, \ldots, n$. As a preliminary conclusion, the variance of Ψ_n converges to $\beta^t\Sigma\beta$. Consequently, Lyapunov criterion for the asymptotic normality of sums

of triangular arrays reduces to $\sum_{i=1}^{n} \mathbb{E}|Z_{i,n}|^3 = n\mathbb{E}|Z_{1,n}|^3 \to 0$. Remark that $Z_{1,n}$ is a bounded random variable:

$$
\begin{aligned}
|Z_{1,n}| &\leq \frac{2C_2 \sum_{j=1}^{J} |\beta_j|}{n\Lambda_n(x)\mu_x^{(1)}(h_n)\bar{F}(y_{n,J}|x)} = 2C_2\tau_J^{-1}\frac{\mu_x^{(1)}(h_n)}{\mu_x^{(2)}(h_n)}\sum_{j=1}^{J}|\beta_j|\Lambda_n(x)(1+o(1)) \\
&\leq 2(C_2/C_1)^2\tau_J^{-1}\sum_{j=1}^{J}|\beta_j|\Lambda_n(x)(1+o(1)),
\end{aligned}
$$

in view of (7) and thus,

$$
\begin{aligned}
n\mathbb{E}|Z_{1,n}|^3 &\leq 2(C_2/C_1)^2\tau_J^{-1}\sum_{j=1}^{J}|\beta_j|\Lambda_n(x)n\mathrm{var}(Z_{1,n})(1+o(1)) \\
&= 2(C_2/C_1)^2\tau_J^{-1}\sum_{j=1}^{J}|\beta_j|\beta^t\Sigma\beta\Lambda_n(x)(1+o(1)) \to 0
\end{aligned}
$$

as $n \to \infty$ in view of (7). As a conclusion, Ψ_n converges in distribution to a centered Gaussian random variable with variance $\beta^t\Sigma\beta$ for all $\beta \neq 0$ in $\mathbb{R}^J$. The result is proved. ∎

Lemma 6 below establishes the asymptotic distribution of $\hat{\bar{F}}_n$ in the situation of estimating small tail probabilities. It thus may be compared to [14], Theorem 1 and [15], Proposition 1. In the first mentioned paper, a similar result is established in the finite dimensional setting and under the assumption that the conditional distribution of Y given X is heavy-tailed. In the second paper, the heavy-tailed assumption is replaced by a von-Mises condition covering different tail behaviors but under the strong condition that the hazard function $H(.|x)$ is twice differentiable.

Lemma 6. *Suppose (2) and* **(A.2)** *hold. Let $x \in E$ such that $\varphi_x(h_n) > 0$ where $h_n \to 0$ as $n \to \infty$ and introduce $0 < \tau_J < \cdots < \tau_1 \leq 1$, where J is a positive integer. Consider $\alpha_n \to 0$ such that $n\varphi_x(h_n)\alpha_n \to \infty$ and $n\varphi_x(h_n)\alpha_n(\Delta\bar{F})^2(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n) \to 0$ for some $\eta > 0$. Let $y_{n,j} := q(\tau_j\alpha_n|x)(1 + o(1/\log(1/\alpha_n)))$ for all $j = 1, \ldots, J$. Then, the random vector*

$$
\left\{ \Lambda_n^{-1}(x) \left(\frac{\hat{\bar{F}}_n(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)} - 1 \right) \right\}_{j=1,\ldots,J}
$$

is asymptotically Gaussian, centered, with covariance matrix Σ defined in Lemma 5.

Proof of Lemma 6. Keeping in mind the notations of Lemma 5 and letting

$$
\begin{aligned}
\Delta_{1,n} &= \Lambda_n^{-1}(x)\sum_{j=1}^{J}\beta_j\left(\frac{\hat{\psi}_n(y_{n,j}, x) - \mathbb{E}(\hat{\psi}_n(y_{n,j}, x))}{\bar{F}(y_{n,j}|x)}\right) \\
\Delta_{2,n} &= \Lambda_n^{-1}(x)\sum_{j=1}^{J}\beta_j\left(\frac{\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) - \bar{F}(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)}\right) \\
\Delta_{3,n} &= \left(\sum_{j=1}^{J}\beta_j\right)\Lambda_n^{-1}(x)(\hat{g}_n(x) - 1),
\end{aligned}
$$

14

the following expansion holds

$$\Lambda_n^{-1}(x) \sum_{j=1}^{J} \beta_j \left(\frac{\hat{\bar{F}}_n(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)} - 1 \right) = \frac{\Delta_{1,n} + \Delta_{2,n} - \Delta_{3,n}}{\hat{g}_n(x)}. \tag{19}$$

From Lemma 5(ii), the random term $\Delta_{1,n}$ can be rewritten as

$$\Delta_{1,n} = \sqrt{\beta^t \Sigma \beta} \, \xi_n, \tag{20}$$

where ξ_n converges to a standard Gaussian random variable. The nonrandom term $\Delta_{2,n}$ is controlled with Lemma 5(i):

$$\Delta_{2,n} = O(\Lambda_n^{-1}(x) \Delta \bar{F}(x, (1-\eta)\tau_J \alpha_n, (1+\eta)\alpha_n, h_n)) = o(1). \tag{21}$$

Finally, $\Delta_{3,n}$ is a classical term in kernel density estimation. From Lemma 4,

$$\Delta_{3,n} = O_{\mathbb{P}}(\Lambda_n^{-1}(x)(n\varphi_x(h_n))^{-1/2}) = O_{\mathbb{P}}(\alpha_n)^{1/2} = o_{\mathbb{P}}(1). \tag{22}$$

Collecting (19)–(22), it follows that

$$\hat{g}_n(x) \Lambda_n^{-1}(x) \sum_{j=1}^{J} \beta_j \left(\frac{\hat{\bar{F}}_n(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)} - 1 \right) = \sqrt{\beta^t \Sigma \beta} \, \xi_n + o_{\mathbb{P}}(1).$$

Finally, Lemma 4 entails that $\hat{g}_n(x) \xrightarrow{\mathbb{P}} 1$ which concludes the proof. ∎

The next lemma establishes the asymptotic normality of the kernel estimator of a conditional extreme quantile. This result can be obtained directly from [15], Corollary 1 in the finite dimensional setting, but under the strong condition that the hazard function $H(.|x)$ is twice differentiable.

Lemma 7. *Suppose (2) and* **(A.2)** *hold. Let* $0 < \tau_J < \cdots < \tau_1 \leq 1$ *where* J *is a positive integer and* $x \in E$ *such that* $\varphi_x(h_n) > 0$ *where* $h_n \to 0$ *as* $n \to \infty$. *If* $\alpha_n \to 0$ *and there exists* $\eta > 0$ *such that* $n\varphi_x(h_n)\alpha_n \to \infty$, $n\varphi_x(h_n)\alpha_n(\Delta\bar{F})^2(x, (1-\eta)\tau_J\alpha_n, (1+\eta)\alpha_n, h_n) \to 0$, *then, the random vector*

$$\left\{ \log(1/\alpha_n)\Lambda_n^{-1}(x) \left(\frac{\hat{q}_n(\tau_j \alpha_n|x)}{q(\tau_j \alpha_n|x)} - 1 \right) \right\}_{j=1,\ldots,J}$$

is asymptotically Gaussian, centered, with covariance matrix $\theta^2(x)\Sigma$ *where* Σ *is defined in Lemma 5.*

Proof of Lemma 7. Introduce for $j = 1, \ldots, J$, $\alpha_{n,j} = \tau_j \alpha_n$,

$$\begin{aligned}
\sigma_{n,j}(x) &= \theta(x)q(\alpha_{n,j}|x)\Lambda_n(x)/\log(1/\alpha_n) \\
v_{n,j}(x) &= \alpha_{n,j}^{-1}\Lambda_n^{-1}(x) \\
W_{n,j}(x) &= v_{n,j}(x) \left(\hat{\bar{F}}_n(q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) - \bar{F}(q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) \right) \\
a_{n,j}(x) &= v_{n,j}(x) \left(\alpha_{n,j} - \bar{F}(q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) \right)
\end{aligned}$$

and $z_j \in \mathbb{R}$. We examine the asymptotic behavior of J-variate function defined by

$$\Phi_n(z_1,\ldots,z_J) = \mathbb{P}\left(\bigcap_{j=1}^{J} \left\{ \sigma_{n,j}^{-1}(x)(\hat{q}_n(\alpha_{n,j}|x) - q(\alpha_{n,j}|x)) \leq z_j \right\} \right) = \mathbb{P}\left(\bigcap_{j=1}^{J} \{ W_{n,j}(x) \leq a_{n,j}(x) \} \right).$$

Let us first focus on the nonrandom term $a_{n,j}(x)$. From assumption **(A.2)**, $\bar{F}(.|x)$ is continuously differentiable. Thus, a first order Taylor expansion yields that, for each $j \in \{1, \ldots, J\}$, there exists $\zeta_{n,j} \in (0,1)$ such that

$$\bar{F}(q(\alpha_{n,j}|x)|x) - \bar{F}(q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) = \sigma_{n,j}(x)z_j f(q_{n,j}|x), \tag{23}$$

with $q_{n,j} = q(\alpha_{n,j}|x) + \zeta_{n,j}\sigma_{n,j}(x)z_j$ and where $f(.|x) = -\bar{F}'(.|x)$ is the conditional density function. Applying Lemma 1 yields for $j = 1, \ldots, J$,

$$\frac{\bar{F}(q_{n,j}|x)}{\bar{F}(q(\alpha_{n,j}|x)|x)} = \frac{\bar{F}(q_{n,j}|x)}{\alpha_{n,j}} = 1 + o(1). \tag{24}$$

Moreover, remarking that, as $\alpha \to 0$,

$$f(q(\alpha|x)|x) = \frac{\log(1/\alpha)}{\theta(x)H^{-1}(\log(1/\alpha)|x)}\alpha(1 + o(1)), \tag{25}$$

it appears that $f(q(.|x)|x)$ is a regularly varying function at 0 with index 1 and consequently

$$\lim_{n\to\infty} \frac{f(q_{n,j}|x)}{f(q(\alpha_{n,j}|x)|x)} = 1. \tag{26}$$

Collecting (23), (25) and (26), we end up with

$$a_{n,j}(x) = \theta(x)z_j\alpha_{n,j}^{-1}\frac{q(\alpha_{n,j}|x)}{\log(1/\alpha_n)}f(q(\alpha_{n,j}|x)|x)(1 + o(1)) = z_j(1 + o(1)). \tag{27}$$

Let us now turn to the random term $W_{n,j}(x)$. Defining $y_{n,j} = q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j$ for $j = 1, \ldots, J$, we have $y_{n,j} = q(\alpha_{n,j}|x)(1 + o(1/\log(1/\alpha_n)))$. Hence, from (24),

$$\left\{\frac{\Lambda_n^{-1}(x)}{v_{n,j}(x)\bar{F}(y_{n,j}|x)}W_{n,j}\right\}_{j=1,\ldots,J} = (1 + o(1))\left\{W_{n,j}\right\}_{j=1,\ldots,J}$$

and Lemma 6 shows that the above random vector converges to a centered Gaussian random variable with covariance matrix Σ. Taking account of (27), we obtain that $\Phi_n(z_1, \ldots, z_J)$ converges to the cumulative distribution function of a centered Gaussian distribution with covariance matrix Σ evaluated at $(z_1, \ldots, z_J)$. ∎

6.2 Proofs of main results

$$\log \hat{q}_n(\tau_j\alpha_n|x) = \log q(\alpha_n|x) + \log\left(\frac{q(\tau_j\alpha_n|x)}{q(\alpha_n|x)}\right) + \log\left(\frac{\hat{q}_n(\tau_j\alpha_n|x)}{q(\tau_j\alpha_n|x)}\right). \tag{28}$$

First, Lemma 2 entails that

$$\log q(\tau_j\alpha_n|x) - \log q(\alpha_n|x) = \frac{\log(1/\tau_j)}{\log(1/\alpha_n)}\left(\theta(x) + \varepsilon(\log(1/\alpha_n)|x)(1 + o(1)) + O\left(\frac{1}{\log(1/\alpha_n)}\right)\right), \tag{29}$$

where the $o(.)$ and $O(.)$ are uniform in $j = 1, \ldots, J$. Second, it follows from Lemma 7 that

$$\frac{\hat{q}_n(\tau_j\alpha_n|x)}{q(\tau_j\alpha_n|x)} = 1 + \sigma_n\xi_{n,j}, \tag{30}$$

16

where $(\xi_{n,1}, \ldots, \xi_{n,J})^t$ converges to a centered Gaussian random vector with covariance matrix $\theta^2(x)\Sigma$ and $\sigma_n^{-1} = \Lambda_n^{-1}(x) \log(1/\alpha_n)$. Replacing (29) and (30) in (28) yields

$$
\begin{aligned}
\log \hat{q}_n(\tau_j \alpha_n | x) &= \log q(\alpha_n | x) + \sigma_n \xi_{n,j} + O_{\mathbb{P}}(\sigma_n^2) \\
&+ \frac{\log(1/\tau_j)}{\log(1/\alpha_n)} \left(\theta(x) + \varepsilon(\log(1/\alpha_n)|x)(1 + o(1)) + O\left(\frac{1}{\log(1/\alpha_n)} \right) \right),
\end{aligned}
$$

for all $j = 1, \ldots, J$ and therefore, in view of the shift and scale invariance properties of ϕ, it follows that $\phi\left(\{\log \hat{q}_n(\tau_j \alpha_n | x)\}_{j=1,\ldots,J} \right) \log(1/\alpha_n)$ is equal to

$$
\phi\left(\left\{ v_j \left(\theta(x) + \varepsilon(\log(1/\alpha_n)|x)(1 + o(1)) + O\left(\frac{1}{\log(1/\alpha_n)} \right) \right) + \log(1/\alpha_n)(\sigma_n \xi_{n,j} + O_{\mathbb{P}}(\sigma_n^2)) \right\}_{j=1,\ldots,J} \right).
$$

where $v = \{\log(1/\tau_j)\}_{j=1,\ldots,J}$. A first order Taylor expansion yields:

$$
\begin{aligned}
\phi\left(\{\log \hat{q}_n(\tau_j \alpha_n | x)\}_{j=1,\ldots,J} \right) &= \sum_{j=1}^{J} v_j \left(\frac{\varepsilon(\log(1/\alpha_n)|x)}{\log(1/\alpha_n)}(1 + o(1)) + O\left(\frac{1}{\log^2(1/\alpha_n)} \right) \right) \frac{\partial \phi}{\partial x_j}(\theta(x)v) \\
&+ \frac{\phi(\theta(x)v)}{\log(1/\alpha_n)} + \sigma_n \sum_{j=1}^{J} (\xi_{n,j} + O_{\mathbb{P}}(\sigma_n)) \frac{\partial \phi}{\partial x_j}(\theta(x)v) + \frac{R_n}{\log(1/\alpha_n)},
\end{aligned}
$$

where the remainder is bounded above by

$$
\begin{aligned}
R_n &= O_{\mathbb{P}}\left(\sum_{j=1}^{J} \left| \sigma_n \log(1/\alpha_n)(\xi_{n,j} + O_{\mathbb{P}}(\sigma_n)) + v_j \left(\varepsilon(\log(1/\alpha_n)|x)(1 + o(1)) + O\left(\frac{1}{\log(1/\alpha_n)} \right) \right) \right| \right)^2 \\
&= O_{\mathbb{P}}\left(\sigma_n \log(1/\alpha_n) \left(\sum_{j=1}^{J} |\xi_{n,j}| + O_{\mathbb{P}}(\sigma_n) \right) + O(\varepsilon(\log(1/\alpha_n)|x)) + O\left(\frac{1}{\log(1/\alpha_n)} \right) \right)^2 \\
&= O_{\mathbb{P}}\left(\sigma_n \log(1/\alpha_n) \right)^2
\end{aligned}
$$

since $\Lambda_n^{-1}(x)\varepsilon(\log(1/\alpha_n)|x) \to \lambda$ and $\Lambda_n^{-1}(x)/\log(1/\alpha_n) \to 0$ as $n \to \infty$. It follows that

$$
\sigma_n^{-1} \left(\phi\left(\{\log \hat{q}_n(\tau_j \alpha_n | x)\}_{j=1,\ldots,J} \right) - \frac{\phi(\theta(x)v)}{\log(1/\alpha_n)} \right) = \sum_{j=1}^{J} (\lambda v_j + \xi_{n,j}) \frac{\partial \phi}{\partial x_j}(\theta(x)v) + o_{\mathbb{P}}(1).
$$

Taking into account **(A.1)** and the scale invariance property of $\nabla \phi$, we finally obtain

$$
\Lambda_n^{-1}(x)(\hat{\theta}_n(x) - \theta(x)) = \frac{1}{\phi(v)} \sum_{j=1}^{J} (\lambda v_j + \xi_{n,j}) \frac{\partial \phi}{\partial x_j}(v) + o_{\mathbb{P}}(1)
$$

and the conclusion follows. ∎

Proof of Corollary 1. Recalling that $\varepsilon(.|x)$ is regularly varying with index $\rho(x)$, one has $\varepsilon(\log(1/\alpha_n)|x) = (1-\xi)^{\rho(x)}|\varepsilon(\log n|x)|(1+o(1))$ and thus (11) holds under (15). Besides, Lemma 3 implies that $\Delta\bar{F}(x,(1-\eta)\tau_J\alpha_n,(1+\eta)\alpha_n,h_n) = O(h_n\log(n)\log_{-2}(n))$. As a consequence, since φ_x is ultimately increasing,

$$\sqrt{n\varphi_x(h_n)\alpha_n}\Delta\bar{F}(x,(1-\eta)\tau_J\alpha_n,(1+\eta)\alpha_n,h_n)$$
$$= O\left(\varphi_x^{-1}\left(\lambda(1-\xi)^{2\rho(x)}n^{-\xi}(\varepsilon(\log n|x))^{-2}\right)\log(n)\log_{-2}(n)(\varepsilon(\log n|x))^{-1}\right)$$
$$= O\left(\varphi_x^{-1}\left(n^{-\xi/2}\right)\log(n)\log_{-2}(n)(\varepsilon(\log n|x))^{-1}\right)$$
$$= O\left(\varphi_x^{-1}\left(n^{-\xi/2}\right)(\log(n))^{1-\rho(x)+\xi}\right).$$

The first part of condition (12) is thus satisfied under the assumption $\varphi_x^{-1}(1/y)(\log y)^{1+\xi-\rho(x)} \to 0$ as $y \to \infty$. The second part is a direct consequence of $y\varepsilon(y|x) \to \infty$ as $y \to \infty$. ∎

References

[1] Alves, M.I.F., de Haan, L., Lin, T. (2003). Estimation of the parameter controlling the speed of convergence in extreme value theory. *Mathematical Methods of Statistics*, 12, 155–176.

[2] Asimit, A.V., Li, D., Peng, L. (2010). Pitfalls in using Weibull tailed distributions. *Journal of Statistical Planning and Inference*, 140(7), 2018–2024.

[3] Beirlant, J., Bouquiaux, C., Werker, B. (2006). Semiparametric lower bounds for tail index estimation, *Journal of Statistical Planning and Inference*, 136, 705–729.

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

[5] Berlinet, A., Gannoun, A., Matzner-Lober, E. (2001). Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, 35, 139–169.

[6] Beirlant, J., Teugels, J.L. (1992). Modeling large claims in non-life insurance. *Insurance: Mathematics and Economics*, 11, 17–29.

[7] Beirlant, J., Teugels, J., Vynckier, P. (1996). *Practical analysis of extreme values*. Leuven university press, Leuven.

[8] Berred, M. (1991). Record values and the estimation of the Weibull tail-coefficient. *Comptes Rendus de l'Académie des Sciences*, **T. 312**, Série I, 943–946.

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

[10] Bongiorno, E.B., Salinelli, E., Goia, A., Vieu, P. (2014). *Contributions in infinite-dimensional statistics and related topics*. Società Editrice Esculapio.

[11] Broniatowski, M. (1993). On the estimation of the Weibull tail coefficient. *Journal of Statistical Planning and Inference*, 35, 349–366.

[12] Caeiro, F., Gomes, M.I. (2002). Bias reduction in the estimation of parameters of rare events. *Theory of Stochastic Processes*, 8, 67–76.

[13] Danielsson, J., de Haan, L., Peng, L., de Vries, C.G. (2001). Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, 76(2), 226–248.

[14] Daouia, A., Gardes, L., Girard, S., Lekina, A. (2011). Kernel estimators of extreme level curves. *Test*, 20, 311–333.

[15] Daouia, A., Gardes, L., Girard, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli*, 19, 2557–2589.

[16] Diebolt, J., Gardes, L., Girard, S., Guillou, A. (2008). Bias-reduced estimators of the Weibull tail-coefficient. *Test*, 17, 311–331.

[17] Dierckx, G., Beirlant, J., De Waal, D., Guillou, A. (2009). A new estimation method for Weibull-type tails based on the mean excess function. *Journal of the Statistical Planning and Inference*, 139, 1905–1920.

[18] El Methni, J., Gardes, L., Girard, S. (2014). Nonparametric estimation of extreme risk measures from conditional heavy-tailed distributions. *Scandinavian Journal of Statistics*, 41, 988–1012.

[19] Ferraty, F., Laksaci, A., Tadj, A., Vieu, P. (2010). Rate of uniform consistency for nonparametric estimates with functional variables. *Journal of Statistical Planning and Inference*, 140(2), 335–352.

[20] Ferraty, F., Rabhi, A., Vieu, P. (2005). Conditional quantiles for dependent functional data with application to the climatic *El Nino* Phenomenon. *Sankhya: The Indian Journal of Statistics*, 67(2), 378–398.

[21] Ferraty, F., Vieu, P. (2012). The functional nonparametric model and application to spectrometric data. *Computational Statistics*, **17**, 545–564.

[22] Ferraty, F., Vieu, P. (2006). *Nonparametric functional data analysis*. Springer.

[23] Gannoun, A. (1990). Estimation non paramétrique de la médiane conditionnelle, médianogramme et méthode du noyau. *Publications de l'Institut de Statistique de l'Université de Paris*, XXXXVI, 11–22.

[24] Gardes, L., Girard, S. (2006). Comparison of Weibull tail-coefficient estimators. *REVSTAT - Statistical Journal*, 4(2), 163–188,

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

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

[27] Girard, S. (2004). A Hill type estimate of the Weibull tail-coefficient. *Communication in Statistics - Theory and Methods*, 33(2), 205–234.

[28] Goia, A., Vieu, P. (2014). A partitioned single functional index model. *Computational Statistics*, doi 10.1007/s00180-014-0530-1.

[29] Gomes, M.I., Martins, M.J., Neves, M. (2000). Semi-parametric estimation of the second order parameter, asymptotic and finite sample behaviour. *Extremes*, 3, 207–229.

[30] Jiřina, M. (1959). On regular conditional probabilities. *Czechoslovak Mathematical Journal*, 9, 445–451.

[31] Kudraszow, N.L., Vieu, P. (2013). Uniform consistency of kNN regressors for functional variables. *Statistics and Probability Letters*, 83, 1863–1870.

[32] Rachdi, M., Vieu, P. (2007). Nonparametric regression for functional data: automatic smoothing parameter selection. *Journal of Statistical Planning and Inference*, 137, 2784–2801.

[33] Roussas, G.G. (1969). Nonparametric estimation of the transition distribution function of a Markov process. *Ann. Math. Statist.*, 40, 1386–1400.

[34] Samanta, T. (1989). Non-parametric estimation of conditional quantiles. *Statistics and Probability Letters*, 7, 407–412.

[35] Segers, J. (2001). Residual estimators. *Journal of Statistical Planning and Inference*, 98, 15–27.

[36] Shang, H.L. (2014). Bayesian bandwidth estimation for a functional nonparametric regression model with mixed types of regressors and unknown error density. *Journal of Nonparametric Statistics,* 26, 599–615.

[37] Stone, C.J. (1977). Consistent nonparametric regression (with discussion). *The Annals of Statistics*, 5, 595–645.

[38] Stute, W. (1986). Conditional empirical processes. *The Annals of Statistics*, 14, 638–647.

J	V_{opt}	τ_1	τ_2	τ_3	τ_4	τ_5
2	1.5441	1.0000	0.2032			
3	1.2191	1.0000	0.3615	0.0735		
4	1.1223	1.0000	0.4703	0.1702	0.0346	
5	1.0789	1.0000	0.5486	0.2585	0.0936	0.0190

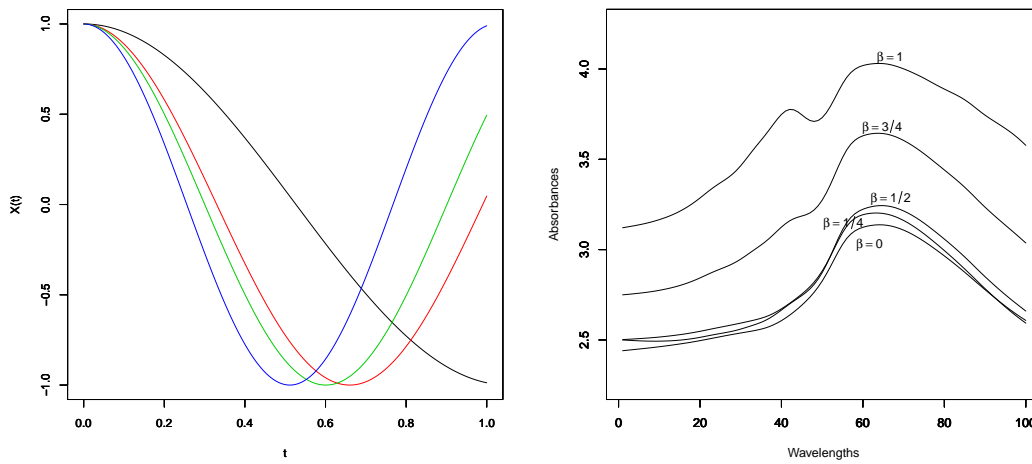Table 1: Optimal values of $(\tau_1, \ldots, \tau_J)$ and associated asymptotic variance V_{opt}.

Figure 1: Samples from the simulated and real datasets. Left: Four simulated random functions $X(.)$. Right: Five spectrometric curves $\tilde{x}_\beta$.
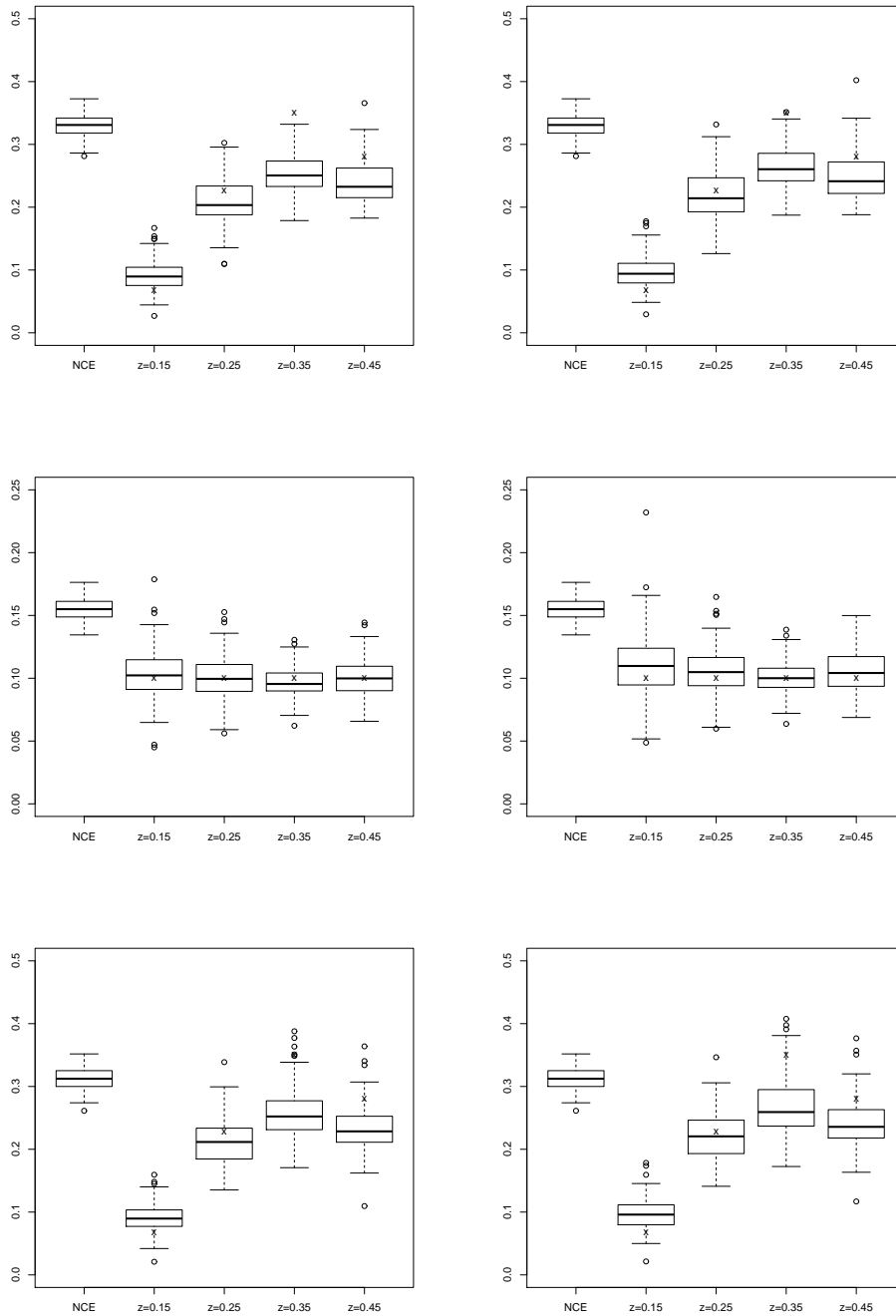
Figure 2: Results obtained with the distance d_1. Boxplots of the $N = 100$ estimations obtained with the non conditional estimator (NCE) and the proposed estimator with $p = 1$ (left) and $p = 3$ (right) for $z \in \{0.15, 0.25, 0.35, 0.45\}$. The first row corresponds to situation (S.1), the second row to (S.2) and the last row to (S.3). The symbol x represents the true value of the functional Weibull tail-coefficient.
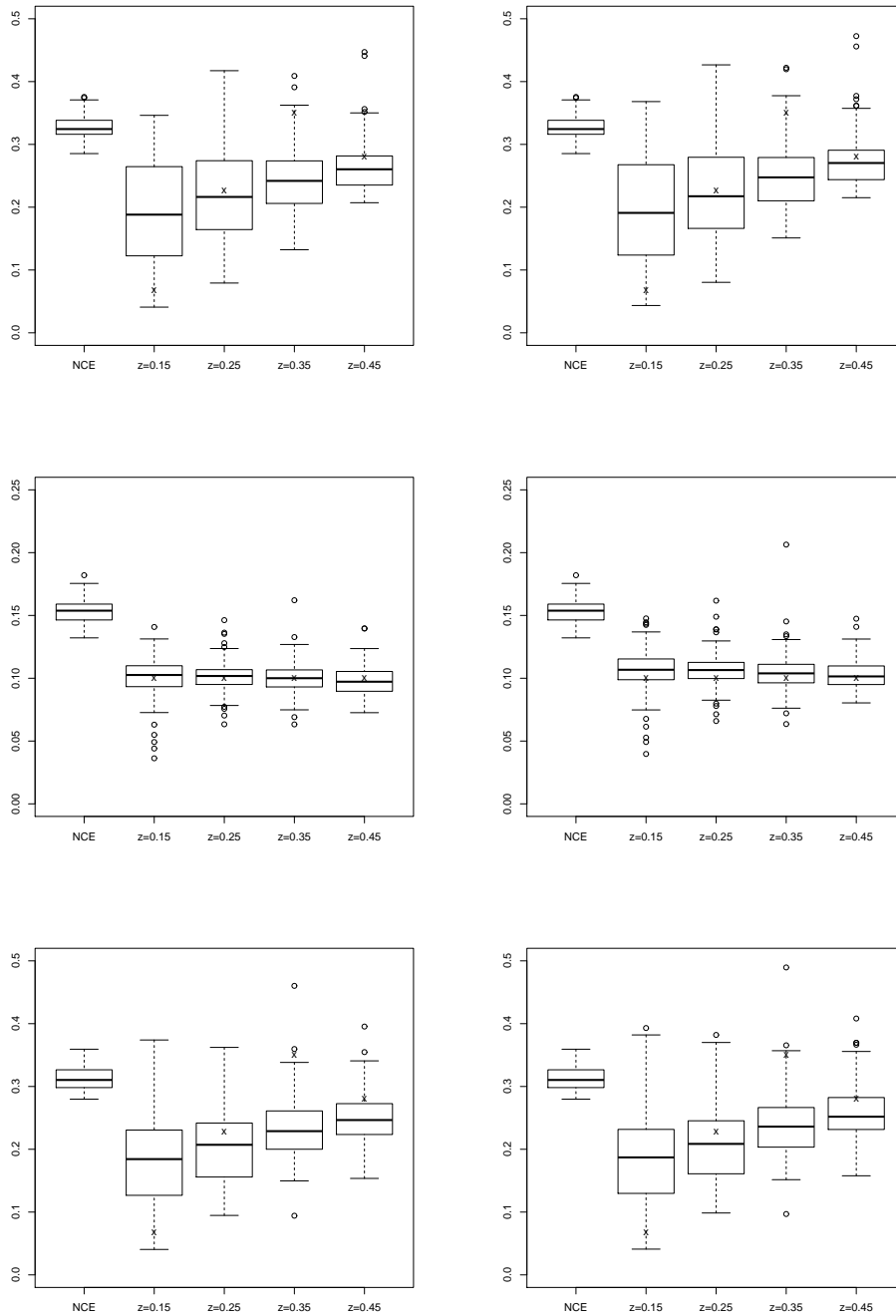
Figure 3: Results obtained with the distance d_2. Boxplots of the $N = 100$ estimations obtained with the non conditional estimator (NCE) and the proposed estimator with $p = 1$ (left) and $p = 3$ (right) for $z \in \{0.15, 0.25, 0.35, 0.45\}$. The first row corresponds to situation (S.1), the second row to (S.2) and the last row to (S.3). The symbol x represents the true value of the functional Weibull tail-coefficient.
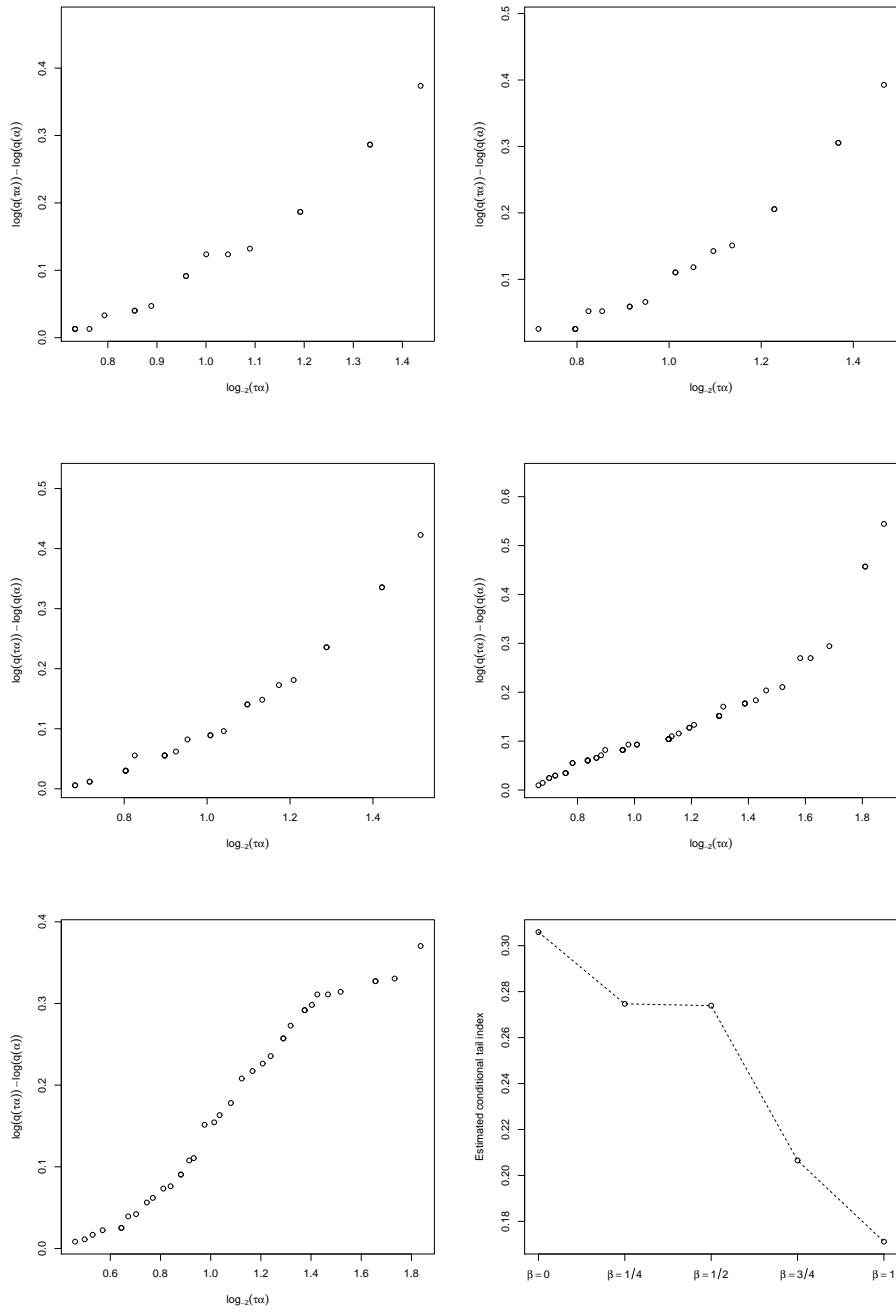
Figure 4: First five graphs: Quantile-quantile plots on the real dataset for the curves $\tilde{x}_\beta$ (left to right and top to bottom). Bottom right: Estimations of the functional Weibull tail-coefficient for the five curves $\tilde{x}_\beta$.