Functional kernel estimators of large conditional quantiles

Laurent Gardes & Stéphane Girard[⋆]

INRIA Rhône-Alpes & LJK, team Mistis, Inovallée, 655, av. de l'Europe,
Montbonnot, 38334 Saint-Ismier cedex, France.
[⋆] Stephane.Girard@inrialpes.fr (corresponding author)

Abstract

We address the estimation of conditional quantiles when the covariate is functional and when the order of the quantiles converges to one as the sample size increases. In a first time, we investigate to what extent these large conditional quantiles can still be estimated through a functional kernel estimator of the conditional survival function. Sufficient conditions on the rate of convergence of their order to one are provided to obtain asymptotically Gaussian distributed estimators. In a second time, basing on these result, a functional Weissman estimator is derived, permitting to estimate large conditional quantiles of arbitrary large order. These results are illustrated on finite sample situations.

Keywords: Conditional quantiles, heavy-tailed distributions, functional kernel estimator, extreme-value theory.
AMS 2000 subject classification: 62G32, 62G30, 62E20.

1 Introduction

Let (X_i, Y_i), $i = 1, \dots, n$ be independent copies of a random pair (X, Y) in $E \times \mathbb{R}$ where E is an infinite dimensional space associated to a semi-metric d. We address the problem of estimating $q(\alpha_n|x) \in \mathbb{R}$ verifying $\mathbb{P}(Y > q(\alpha_n|x)|X = x) = \alpha_n$ where $\alpha_n \to 0$ as $n \to \infty$ and $x \in E$. In such a case, $q(\alpha_n|x)$ is referred to as a large conditional quantile in contrast to classical conditional quantiles (or regression quantiles) for which $\alpha_n = \alpha$ is fixed in $(0, 1)$. While the nonparametric estimation of ordinary regression quantiles has been extensively studied (see for instance [35, 39] or [18], Chapter 5), less attention has been paid to large conditional quantiles despite their potential interest. In climatology, large conditional quantiles may explain how climate change over years might affect extreme temperatures. In the financial econometrics literature, they illustrate the link between extreme hedge fund returns and some measures of risk. Parametric models are introduced in [10, 38] and semi-parametric methods are considered in [2, 31]. Fully non-parametric estimators have been first introduced in [9, 6] through local polynomial and spline models. In both cases, the authors focus on univariate covariates and on the finite sample properties of the estimators. Nonparametric methods based on moving windows and nearest neighbors are introduced respectively in [23, 25] and [24] in the fixed design setting. We also refer to [15], Theorem 3.5.2, for the approximation of the nearest neighbors distribution using the Hellinger distance and to [19] for the study of their asymptotic distribution.

An important literature is devoted to the particular case where the conditional distribution of Y given $X = x$ has a finite endpoint $\varphi(x)$ and when X is a finite dimensional random variable. The function φ is referred to as the frontier and can be estimated from an estimator of the conditional quantile $q(\alpha_n|x)$ with $\alpha_n \to 0$. As an example, a kernel estimator of φ is proposed in [27], the asymptotic normality being proved only when Y given $X = x$ is uniformly distributed on $[0, \varphi(x)]$. We refer to [33] for a review on this topic.

Estimation of unconditional large quantiles is also widely studied since the introduction of Weissman estimator [41] dedicated to heavy-tailed distributions, Weibull-tail estimators [12, 22] dedicated to light-tailed distributions and Dekkers and de Haan estimator [11] adapted to the general case.

In this paper, we focus on the setting where the conditional distribution of Y given $X = x$ has an infinite endpoint and is heavy-tailed, an analytical characterization of this property being given in the next section. In such a case, the frontier function does not exist and $q(\alpha_n|x) \to \infty$ as $\alpha_n \to 0$. Nevertheless, we show, under some conditions, that large regression quantiles $q(\alpha_n|x)$ can still be estimated through a functional kernel estimator of $\mathbb{P}(Y > .|x)$. We provide sufficient conditions on the rate of convergence of α_n to 0 so that our estimator is asymptotically Gaussian distributed. Making use of this, some functional estimators of the conditional tail-index are introduced and a functional Weissman estimator [41] is derived, permitting to estimate large conditional quantiles $q(\beta_n|x)$ where $\beta_n \to 0$ arbitrarily fast.

Assumptions are introduced and discussed in Section 2. Main results are provided in Section 3 and illustrated on simulated data in Section 4. Proofs are postponed to the appendix.

2 Notations and assumptions

The conditional survival function (csf) of Y given $X = x$ is denoted by $\bar{F}(y|x) = \mathbb{P}(Y > y|X = x)$. The functional estimator of $\bar{F}(y|x)$ is defined for all $(x, y) \in E \times \mathbb{R}$ by

$$\hat{\bar{F}}_n(y|x) = \sum_{i=1}^{n} K(d(x, X_i)/h)Q((Y_i - y)/\lambda) \Bigg/ \sum_{i=1}^{n} K(d(x, X_i)/h), \tag{1}$$

with $Q(t) = \int_{-\infty}^{t} Q'(s)ds$ where $K : \mathbb{R}^+ \to \mathbb{R}^+$ and $Q' : \mathbb{R} \to \mathbb{R}^+$ are two kernel functions, and $h = h_n$ and $\lambda = \lambda_n$ are two nonrandom sequences (called window-width) such that $h \to 0$ as

$n \to \infty$. Let us emphasize that the condition $\lambda \to 0$ is not required in this context. This estimator was considered for instance in [18], page 56. In Theorem 1 hereafter, the asymptotic distribution of (1) is established when estimating small tail probabilities, *i.e* when $y = y_n$ goes to infinity with the sample size n. Similarly, the functional estimators of conditional quantiles $q(\alpha|x)$ are defined via the generalized inverse of $\hat{\bar{F}}_n(.|x)$:

$$\hat{q}_n(\alpha|x) = \hat{\bar{F}}_n^{\leftarrow}(\alpha|x) = \inf\{t,\ \hat{\bar{F}}_n(t|x) \leq \alpha\}, \tag{2}$$

for all $\alpha \in (0,1)$. Many authors are interested in this estimator for fixed $\alpha \in (0,1)$. Weak and strong consistency are proved respectively in [39] and [20]. The rate of uniform strong consistency is established by [16] in the functional setting. Asymptotic normality is shown in [3, 36, 40] when E is finite dimensional and by [17] for a general metric space under dependence assumptions. In Theorem 2, the asymptotic distribution of (2) is investigated when estimating large quantiles, *i.e* when $\alpha = \alpha_n$ goes to 0 as the sample size n goes to infinity. The asymptotic behavior of such estimators depends on the nature of the conditional distribution tail. In this paper, we focus on heavy tails. More specifically, we assume that the csf satisfies

$$\textbf{(A.1)}:\ \bar{F}(y|x) = c(x) \exp\left\{ - \int_1^y \left(\frac{1}{\gamma(x)} - \varepsilon(u|x) \right) \frac{du}{u} \right\},$$

where γ is a positive function of the covariate x, c is a positive function and $|\varepsilon(.|x)|$ is continuous and ultimately decreasing to 0. Examples of such distributions are provided in Table 1. **(A.1)** implies that the conditional distribution of Y given $X = x$ is in the Fréchet maximum domain of attraction. In this context, $\gamma(x)$ is referred to as the conditional tail-index since it tunes the tail heaviness of the conditional distribution of Y given $X = x$. More details on extreme-value theory can be found for instance in [14]. Assumption **(A.1)** also yields that $\bar{F}(.|x)$ is regularly varying at infinity with index $-1/\gamma(x)$. *i.e* for all $\zeta > 0$,

$$\lim_{y \to \infty} \frac{\bar{F}(\zeta y|x)}{\bar{F}(y|x)} = \zeta^{-1/\gamma(x)}. \tag{3}$$

We refer to [4] for a general account on regular variation theory. The auxiliary 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. In [1, 28], $|\varepsilon(.|x)|$ is supposed to be regularly varying and the estimation of the corresponding regular variation index is addressed.

Some Lipschitz conditions are also required:

(A.2): There exist κ_ε, κ_c, $\kappa_\gamma > 0$ and $u_0 > 1$ such that for all $(x, x') \in E \times E$ and $u > u_0$,

$$\begin{aligned} |\log c(x) - \log c(x')| &\leq \kappa_c d(x, x'), \\ |\varepsilon(u|x) - \varepsilon(u|x')| &\leq \kappa_\varepsilon d(x, x'), \\ \left| \frac{1}{\gamma(x)} - \frac{1}{\gamma(x')} \right| &\leq \kappa_\gamma d(x, x'). \end{aligned}$$

The last two assumptions are standard in the functional kernel estimation framework.

(A.3): K is a function with support $[0, 1]$ and there exist $0 < C_1 < C_2 < \infty$ such that $C_1 \le K(t) \le C_2$ for all $t \in [0, 1]$.

(A.4): Q' is a probability density function (pdf) with support $[-1, 1]$.

One may also assume without loss of generality that K integrates to one. In this case, K is called a type I kernel, see [18], Definition 4.1. Letting $B(x, h)$ be the ball of center x and radius h, we finally introduce $\varphi_x(h) := \mathbb{P}(X \in B(x, h)$ the small ball probability of X. Under **(A.3)**, the τ-th moment $\mu_x^{(\tau)}(h) := \mathbb{E}\{K^\tau(d(x, X)/h)$ can be controlled for all $\tau > 0$ by Lemma 3 in Appendix. It is shown that $\mu_x^{(\tau)}(h)$ is of the same asymptotic order as $\varphi_x(h)$.

3 Main results

Let us first focus on the estimation of small tail probabilities $\bar{F}(y_n|x)$ when $y_n \to \infty$ as $n \to \infty$. Defining

$$\Lambda_n(x) = \left(n\bar{F}(y_n|x) \frac{(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)} \right)^{-1/2},$$

the following result provides sufficient conditions for the asymptotic normality of $\hat{\bar{F}}_n(y_n|x)$.

Theorem 1 *Suppose* **(A.1)** – **(A.4)** *hold. Let $x \in E$ such that $\varphi_x(h) > 0$ and introduce $y_{n,j} = a_j y_n(1 + o(1))$ for $j = 1, \dots, J$ with $0 < a_1 < a_2 < \dots < a_J$ and where J is a positive integer. If $y_n \to \infty$ such that $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$ and $n\varphi_x(h)\bar{F}(y_n|x)(\lambda/y_n \vee h \log y_n)^2 \to 0$ as $n \to \infty$, then*

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

is asymptotically Gaussian, centered, with covariance matrix $C(x)$ where $C_{j,j'}(x) = a_{j \wedge j'}^{1/\gamma(x)}$ for $(j, j') \in \{1, \dots, J\}^2$.

Note that $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$ is a necessary and sufficient condition for the almost sure presence of at least one sample point in the region $B(x, h) \times (y_n, \infty)$ of $E \times \mathbb{R}$, see Lemma 4 in Appendix. Thus, this natural condition states that one cannot estimate small tail probabilities out of the sample using $\hat{\bar{F}}_n$. Besides, from Lemma 3, $\Lambda_n^{-2}(x)$ is of the same asymptotic order as $n\varphi_x(h)\bar{F}(y_n|x)$ and consequently $\Lambda_n(x) \to 0$ as $n \to \infty$. Theorem 1 thus entails $\hat{\bar{F}}_n(y_{n,j}|x)/\bar{F}(y_{n,j}|x) \xrightarrow{P} 1$ which can be read as a consistency of the estimator. The second condition $n\varphi_x(h)\bar{F}(y_n|x)(\lambda/y_n \vee h \log y_n)^2 \to 0$ imposes to the biases λ/y_n and $h \log y_n$ introduced by the two smoothings to be neglegible compared to the standard deviation $\Lambda_n(x)$ of the estimator. Theorem 1 may be compared to [13] which establishes the asymptotic behavior of the empirical survival function in the unconditional case but without assumption on the distribution. Letting

$$\sigma_n(x) = \left(n\alpha_n \frac{(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)} \right)^{-1/2},$$

the asymptotic normality of $\hat{q}_n(\alpha_n|x)$ when $\alpha_n \to 0$ as $n \to \infty$ can be established under similar conditions.

Theorem 2 *Suppose* **(A.1)** – **(A.4)** *hold. Let $x \in E$ such that $\varphi_x(h) > 0$ and consider a sequence $\tau_1 > \tau_2 > \cdots > \tau_J > 0$ where J is a positive integer. If $\alpha_n \to 0$ such that $\sigma_n(x) \to 0$ and $\sigma_n^{-1}(x)(\lambda/q(\alpha_n|x) \vee h \log \alpha_n) \to 0$ as $n \to \infty$, then*

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

is asymptotically Gaussian, centered, with covariance matrix $\gamma^2(x)\Sigma$ where $\Sigma_{j,j'} = 1/\tau_{j \wedge j'}$ for $(j, j') \in \{1, \dots, J\}^2$.

Remark that **(A.1)** provides an asymptotic expansion of the density function of Y given $X = x$:

$$f(y|x) = \frac{1}{\gamma(x)} \frac{\bar{F}(y|x)}{y}(1 - \varepsilon(y|x)) = \frac{1}{\gamma(x)} \frac{\bar{F}(y|x)}{y}(1 + o(1))$$

as $y \to \infty$. Consequently, Theorem 2 entails that the random vector

$$\left\{ \frac{\mu_x^{(1)}(h)}{(\mu_x^{(2)}(h))^{1/2}} (n\tau_j\alpha_n(1 - \tau_j\alpha_n))^{-1/2} f(q(\tau_j\alpha_n|x)|x) \left(\hat{q}_n(\tau_j\alpha_n|x) - q(\tau_j\alpha_n|x)\right) \right\}_{j=1,\dots,J}$$

is also asymptotically Gaussian and centered. This result coincides with [3], Theorem 6.4 established in the case where $\alpha_n = \alpha$ is fixed in $(0, 1)$ and in a finite dimensional setting. The functional estimator of large quantiles $\hat{q}_n(\alpha_n|x)$ requires the stringent condition $n\varphi_x(h)\alpha_n \to \infty$, since by construction it cannot extrapolate beyond the maximum observation in the ball $B(x, h)$. To overcome this limitation, a functional Weissman estimator [41] can be derived:

$$\hat{q}_n^{\mathrm{W}}(\beta_n|x) = \hat{q}_n(\alpha_n|x)(\alpha_n/\beta_n)^{\hat{\gamma}_n(x)}. \tag{4}$$

Here, $\hat{q}_n(\alpha_n|x)$ is the functional estimator (2) of the large quantile and $\hat{\gamma}_n(x)$ is an estimator of the conditional tail-index $\gamma(x)$. As illustrated in the next theorem, the extrapolation factor $(\alpha_n/\beta_n)^{\hat{\gamma}_n(x)}$ allows to estimate large quantiles of order β_n arbitrary small.

Theorem 3 *Suppose* **(A.1)** – **(A.4)** *hold. Let $x \in E$ and introduce*

- *$\alpha_n \to 0$ such that $\sigma_n(x) \to 0$ and $\sigma_n^{-1}(x)(\lambda/q(\alpha_n|x) \vee h \log \alpha_n \vee \varepsilon(q(\alpha_n|x)|x)) \to 0$ as $n \to \infty$,*
- *(β_n) such that $\beta_n/\alpha_n \to 0$ as $n \to \infty$,*
- *$\hat{\gamma}_n(x)$ such that $\sigma_n^{-1}(x)(\hat{\gamma}_n(x) - \gamma(x)) \xrightarrow{d} \mathcal{N}(0, V(x))$ where $V(x) > 0$.*

Then,

$$\frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)} \left(\frac{\hat{q}_n^{\mathrm{W}}(\beta_n|x)}{q(\beta_n|x)} - 1 \right) \xrightarrow{d} \mathcal{N}(0, V(x)).$$

Note that, when K is the pdf of the uniform distribution, this result is consistent with [25], Theorem 3, obtained in a fixed-design setting.

Let us now focus on the estimation of the conditional tail-index. Let $\alpha_n \to 0$ and consider a sequence $1 = \tau_1 > \tau_2 > \cdots > \tau_J > 0$ where J is a positive integer. Two additional notations are introduced for the sake of simplicity: $u = (1, \ldots, 1)^t \in \mathbb{R}^J$ and $v = (\log(1/\tau_1), \ldots, \log(1/\tau_J))^t \in \mathbb{R}^J$. The following family of estimators is proposed

$$\hat{\gamma}_n^{\phi}(x) = \frac{\phi(\log \hat{q}_n(\tau_1 \alpha_n|x), \ldots, \log \hat{q}_n(\tau_J \alpha_n|x))}{\phi(\log(1/\tau_1), \ldots, \log(1/\tau_J))}, \tag{5}$$

where $\phi : \mathbb{R}^J \to \mathbb{R}$ denotes a twice differentiable function verifying the shift and location invariance conditions

$$\begin{cases} \phi(\theta v) & = & \theta \phi(v) \\ \phi(\eta u + x) & = & \phi(x) \end{cases} \tag{6}$$

for all $\theta > 0$, $\eta \in \mathbb{R}$ and $x \in \mathbb{R}^J$. In the case where $J = 3$, $\tau_1 = 1$, $\tau_2 = 1/2$ and $\tau_3 = 1/4$, the function

$$\phi_{\text{FP}}(x_1, x_2, x_3) = \log\left(\frac{\exp(4x_2) - \exp(4x_1)}{\exp(4x_3) - \exp(4x_2)}\right)$$

leads us to a functional version of Pickands estimator [34]:

$$\hat{\gamma}_n^{\phi_{\text{FP}}}(x) = \frac{1}{\log 2} \log\left(\frac{\hat{q}_n(\alpha_n|x) - \hat{q}_n(2\alpha_n|x)}{\hat{q}_n(2\alpha_n|x) - \hat{q}_n(4\alpha_n|x)}\right).$$

We refer to [26] for a different variant of Pickands estimator in the context where the distribution of Y given $X = x$ has a finite endpoint. Besides, introducing the function $m_p(x_1, \ldots, x_J) = \sum_{j=1}^{J} (x_j - x_1)^p$ for all $p > 0$ and considering $\phi_p(x) = m_p^{1/p}(x)$ gives rise to a functional version of the estimator considered for instance in [37], example (a):

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

As a particular case $\phi_1(x) = m_1(x)$ corresponds to a functional version of the Hill estimator [32]:

$$\hat{\gamma}_n^{\phi_1}(x) = \sum_{j=1}^{J} [\log \hat{q}_n(\tau_j \alpha_n|x) - \log \hat{q}_n(\alpha_n|x)] \Bigg/ \sum_{j=1}^{J} \log(1/\tau_j).$$

More interestingly, if $\{\phi^{(1)}, \ldots, \phi^{(H)}\}$ is a set of H functions satisfying (6) and if $A : \mathbb{R}^H \to \mathbb{R}$ is a homogeneous function of degree 1, then the aggregated function $A(\phi^{(1)}, \ldots, \phi^{(H)})$ also satisfies (6). Generalizations of the functional Hill estimator can then be obtained using $H = 2$, $A_p(x, y) = x^p y^{1-p}$ and defining $\phi_{p,q,r} = A_p(\phi_q, \phi_r) = m_q^{p/q} m_r^{(1-p)/r}$:

$$\hat{\gamma}_n^{\phi_{p,q,r}}(x) = \frac{\left(\sum_{j=1}^{J} [\log \hat{q}_n(\tau_j \alpha_n|x) - \log \hat{q}_n(\alpha_n|x)]^p\right)^{p/q} \left(\sum_{j=1}^{J} [\log(1/\tau_j)]^r\right)^{(p-1)/r}}{\left(\sum_{j=1}^{J} [\log \hat{q}_n(\tau_j \alpha_n|x) - \log \hat{q}_n(\alpha_n|x)]^r\right)^{(p-1)/r} \left(\sum_{j=1}^{J} [\log(1/\tau_j)]^p\right)^{p/q}}.$$

For instance, the estimator introduced by [29], equation (2.2) corresponds to the particular function $\phi_{p,p,1}$ and the estimator of [5] corresponds to $\phi_{p,p\theta,p-1}$.
For an arbitrary function ϕ, the asymptotic normality of $\hat{\gamma}_n^{\phi}(x)$ is a consequence of Theorem 2. The following result permits to establish the asymptotic normality of the above mentioned estimators in an unified way.

Theorem 4 *Under the assumptions of Theorem 2 and if, moreover,* $\sigma_n^{-1}(x)\varepsilon(q(\alpha_n|x)|x) \to 0$ *as* $n \to \infty$*, then,* $\sigma_n^{-1}(x)(\hat{\gamma}_n^{\phi}(x) - \gamma(x))$ *converges to a centered Gaussian random variable with variance*

$$V_{\phi}(x) = \frac{\gamma^2(x)}{\phi^2(v)}(\nabla\phi(\gamma(x)v))^t\Sigma(\nabla\phi(\gamma(x)v)).$$

Let us note that the additional condition $\sigma_n^{-1}(x)\varepsilon(q(\alpha_n|x)|x) \to 0$ is standard in the extreme-value framework: Neglecting the unknown function $\varepsilon(.|x)$ in the construction of $\hat{\gamma}_n^{\phi}(x)$ yields a bias that should be negligible with respect to the standard deviation $\sigma_n(x)$ of the estimator. Finally, combining Theorem 3 and Theorem 4, the asymptotic distribution of the functional large quantile estimator $\hat{q}_n^{\mathrm{W},\phi}(\beta_n|x)$ based on (4) and (5) is readily obtained.

Corollary 1 *Suppose* **(A.1)** – **(A.4)** *hold. Let* $x \in E$ *such that* $\varphi_x(h) > 0$ *and consider a sequence* $1 = \tau_1 > \tau_2 > \cdots > \tau_J > 0$ *where* J *is a positive integer. If*

- $\alpha_n \to 0$*,* $\sigma_n(x) \to 0$ *and* $\sigma_n^{-1}(x)(\lambda/q(\alpha_n|x) \vee h\log\alpha_n \vee \varepsilon(q(\alpha_n|x)|x)) \to 0$ *as* $n \to \infty$*,*
- $\beta_n/\alpha_n \to 0$ *as* $n \to \infty$*,*

then

$$\frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)}\left(\frac{\hat{q}_n^{\mathrm{W},\phi}(\beta_n|x)}{q(\beta_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}(0, V_{\phi}(x)).$$

As an example, in the case of the functional Hill and Pickands estimators, we obtain

$$V_{\phi_1}(x) = \gamma^2(x)\left(\sum_{j=1}^{J}\frac{2(J-j)+1}{\tau_j} - J^2\right)\Bigg/\left(\sum_{j=1}^{J}\log(1/\tau_j)\right)^2.$$

$$V_{\phi_{\mathrm{FP}}}(x) = \frac{\gamma^2(x)(2^{2\gamma(x)+1}+1)}{4(\log 2)^2(2^{\gamma(x)}-1)^2}.$$

Clearly, $V_{\phi_{\mathrm{FP}}}(x)$ is the variance of the classical Pickands estimator, see for instance [30], Theorem 3.3.5.

4 Illustration on simulated data

The finite sample performance is illustrated on $N = 50$ replications of a sample of size $n = 500$ from a random pair (X, Y), where the functional covariate $X \in E = L^2[0,1]$ is defined by $X(t) = \cos(2\pi Zt)$ for all $t \in [0,1]$ where Z is uniformly distributed on $[1/4, 1]$. Some examples of simulated random functions X are depicted on Figure 1. Besides, the conditional distribution of Y given X is a Burr distribution (see Table 1) with parameters $\tau(X) = 2$ and $\lambda(X) = 2/(8\|X\|_2^2 - 3)$ with

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

We focus on the estimation of $q(\beta_n|x)$ with $\beta_n = \log(n)/n$. To this end, the functional Weissman estimator $\hat{q}_n^{\mathrm{W}}(\beta_n|x)$ is used with a piecewise linear kernel $K(t) = (1.9 - 1.8t)\mathbb{I}\{t \in [0,1]\}$ and the triangular kernel Q'. The conditional tail index is estimated by the functional Hill estimator $\hat{\gamma}_n^{\phi_1}$ with $\tau_j = 1/j$ for each $j = 1, \dots, 10$. The choice of the semi-metric d is a recurrent issue in functional estimation [18], Chapter 3. Here, two semi-metrics are considered. The first one is defined for all $(s,t) \in E^2$ by $d_X(s,t) = \|s - t\|_2$ and coincides with the L_2 distance between functions. Remarking that the conditional quantile $q(\alpha_n|X)$ depends only on $\|X\|_2^2$, or equivalently on Z, another interesting semi-metric is $d_Z(s,t) = \left|\|s\|_2^2 - \|t\|_2^2\right|$.

With such choices, the functional Weissman estimator $\hat{q}_n^{\mathrm{W}}(\beta_n|x)$ depends on three parameters α_n, h and λ. The choice 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 [7] for a bootstrap based method. Here, the "optimal" selection of α_n would consist in minimizing the Integrated Mean-Squared Error (IMSE) associated to the estimated conditional extreme-value index $\hat{\gamma}_n^{\phi_1}$:

$$IMSE(\alpha_n) = \frac{1}{N}\sum_{r=1}^{N}\sum_{i=1}^{n}\left(\left(\hat{\gamma}_n^{\phi_1}\right)^{(r)}(X_i) - \gamma(X_i)\right)^2,$$

where $\left(\hat{\gamma}_n^{\phi_1}\right)^{(r)}$ is the estimator $\hat{\gamma}_n^{\phi_1}$ computed on the r-th replication. The "optimal" value of α_n is given by

$$\alpha_n^{opt} = \arg\min\left\{IMSE(\alpha_n),\ \alpha_n = c\frac{\log n}{n},\ c \in \{5, 6, \dots, 20\}\right\}.$$

The results are presented on Figure 2 for the two semi-metrics d_X and d_Z. Clearly, the use of d_Z permits to reach lower IMSE than d_X does. Let us highlight that α_n^{opt} cannot be computed in practical situations where the true function γ is unknown. Nevertheless, it will appear in the following that the estimations are not very sensitive with respect to the choice of α_n. The smoothing parameter h is selected using the cross-validation approach introduced in [42] and implemented for instance in [8, 21]:

$$h^{opt} = \arg\min\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathbb{I}\{Y_i \geq Y_j\} - \hat{\bar{F}}_{n,-i}(Y_j|X_i)\right)^2,\ h \in \mathcal{H}\right\}$$

where $\hat{\bar{F}}_{n,-i}$ is the estimator (depending on h) given in (1) computed from the sample $\{(X_\ell, Y_\ell),\ 1 \leq \ell \leq n,\ \ell \neq i\}$. Here, $\mathcal{H}$ is a regular grid, $\mathcal{H} = \{h_1 \leq h_2 \leq \cdots \leq h_M\}$ with $h_1 = 1/100$, $h_M = 1/10$ and $M = 20$. In our experiments, the choice of the bandwidth λ appeared to be less crucial than the other smoothing parameter h. It could have been selected with the same criteria as previously, but for simplicity reasons, it has been fixed to $\lambda = 0.1$.

In Figure 3, the estimator $\hat{q}_n^{\mathrm{W}}(\beta_n|x)$ is represented as a function of Z. The estimator has been computed for two values of α_n (the "optimal" value α_n^{opt} and an arbitrary value $\alpha_n^{arb} = 15\log(n)/n$) and for the two semi-metrics d_X and d_Z. We have only represented the estimator computed on the replication that gives rise to the median of the L_2-errors $\Delta_d^{(r)}$, $r = 1,\dots,N$ with

$$\Delta_d^{(r)} = \sum_{i=1}^{n} \left((\hat{q}_n^{\mathrm{W}}(\beta_n|X_i))^{(r)} - q(\beta_n|X_i) \right)^2,$$

and where d can be either d_X or d_Z. It appears that the choice of the sequence α_n is not crucial *i.e.* the results obtained with α_n^{opt} are not visually better than these obtained with α_n^{arb}. The choice of the semi-metric seems to be a more challenging issue.

5 Appendix: Proofs

5.1 Preliminary results

The following two lemmas are of analytical nature. The first one is dedicated to the control of the local variations of the csf when the quantity of interest y goes to infinity.

Lemma 1 *Let $x \in E$ and suppose* **(A.1)** *and* **(A.2)** *hold.*
(i) If $y_n \to \infty$ and $h \log y_n \to 0$ as $n \to \infty$, then, for n large enough,

$$\sup_{x' \in B(x,h)} \left| \frac{\bar{F}(y_n|x)}{\bar{F}(y_n|x')} - 1 \right| \le 2(\kappa_c + \kappa_\gamma + \kappa_\varepsilon) h \log y_n.$$

(ii) If $y_n \to \infty$ and $y_n' \to \infty$ as $n \to \infty$, then, for n large enough,

$$\sup_{x' \in B(x,h)} \left| \frac{\bar{F}(y_n'|x')}{\bar{F}(y_n|x')} - 1 \right| \le \left| \left(\frac{y_n}{y_n'} \right)^{2/\gamma(x)} - 1 \right|.$$

Proof. (i) Assumption **(A.1)** yields, for all $x' \in B(x,h)$:

$$\begin{aligned}
\left| \log \left(\frac{\bar{F}(y_n|x)}{\bar{F}(y_n|x')} \right) \right| &\le |\log c(x) - \log c(x')| + \int_1^{y_n} \left(\left| \frac{1}{\gamma(x)} - \frac{1}{\gamma(x')} \right| + |\varepsilon(u|x) - \varepsilon(u|x')| \right) \frac{du}{u} \\
&\le \kappa_c h + \int_1^{y_n} (\kappa_\gamma + \kappa_\varepsilon) h \frac{du}{u} \\
&\le (\kappa_c + \kappa_\gamma + \kappa_\varepsilon) h \log y_n,
\end{aligned}$$

eventually, from **(A.2)**. Thus,

$$\sup_{d(x,x') \le h} \left| \log \left(\frac{\bar{F}(y_n|x)}{\bar{F}(y_n|x')} \right) \right| = O(h \log y_n) \to 0$$

as $n \to \infty$ and taking account of $\log(u+1) \sim u$ as $u \to 0$ gives the result.

(ii) Let us assume for instance $y_n' > y_n$. From **(A.1)** we have

$$\left|\frac{\bar{F}(y_n'|x')}{\bar{F}(y_n|x')} - 1\right| = 1 - \left(\frac{y_n'}{y_n}\right)^{-1/\gamma(x')} \exp\left(\int_{y_n}^{y_n'} \frac{\varepsilon(u|x')}{u} du\right) \leq 1 - \left(\frac{y_n'}{y_n}\right)^{-1/\gamma(x') - |\varepsilon(y_n|x')|}. \quad (7)$$

Now, $x' \in B(x, h)$ and **(A.2)** imply for n large enough that

$$\frac{1}{\gamma(x')} + |\varepsilon(y_n|x')| \leq \frac{1}{\gamma(x)} + (\kappa_\varepsilon + \kappa_\gamma)h + |\varepsilon(y_n|x)| \leq \frac{2}{\gamma(x)}.$$

Replacing in (7), it follows that

$$\left|\frac{\bar{F}(y_n'|x')}{\bar{F}(y_n|x')} - 1\right| \leq 1 - \left(\frac{y_n'}{y_n}\right)^{-2/\gamma(x)}.$$

The case $y_n' \leq y_n$ is similar. ■

The second lemma provides a second order asymptotic expansion of the quantile function. It is proved in [8].

Lemma 2 *Suppose* **(A.1)** *hold.*
(i) Let $0 < \beta_n < \alpha_n$ *with* $\alpha_n \to 0$ *as* $n \to \infty$*. Then,*

$$|\log q(\beta_n|x) - \log q(\alpha_n|x) + \gamma(x)\log(\beta_n/\alpha_n)| = O(\log(\alpha_n/\beta_n)\varepsilon(q(\alpha_n|x)|x)).$$

(ii) If, moreover, $\liminf \beta_n/\alpha_n > 0$*, then*

$$\frac{\beta_n^{\gamma(x)} q(\beta_n|x)}{\alpha_n^{\gamma(x)} q(\alpha_n|x)} = 1 + O(\varepsilon(q(\alpha_n|x)|x)).$$

The following lemma provides a control on the moments $\mu_x^{(\tau)}(h)$ for all $\tau > 0$, the case $\tau = 1$ being studied in [18], Lemma 4.3.

Lemma 3 *Suppose* **(A.3)** *holds. For all* $\tau > 0$ *and* $x \in E$*,*

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

Proof. From **(A.3)**, we have

$$0 < C_1 \mathbb{I}\{t \in [0,1]\} \leq K(t) \leq C_2 \mathbb{I}\{t \in [0,1]\}$$

and thus, for all $\tau > 0$,

$$0 < C_1^\tau \mathbb{I}\{d(x, X) \leq h\} \leq K^\tau(d(x, X)/h) \leq C_2^\tau \mathbb{I}\{d(x, X) \leq h\}.$$

Taking the expectation concludes the proof. ■

The following lemma provides a geometrical interpretation of the condition $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$.

Lemma 4 *Suppose* **(A.1)**, **(A.2)** *hold and let* $y_n \to \infty$ *such that* $h \log y_n \to 0$ *as* $n \to \infty$. *Consider the subset of* $E \times \mathbb{R}$ *defined as* $R_n(x) = B(x,h) \times (y_n, \infty)$ *where* $x \in E$ *is such that* $\varphi_x(h) > 0$. *Then,* $\mathbb{P}(\exists i \in \{1,\dots,n\}, (X_i, Y_i) \in R_n(x)) \to 1$ *as* $n \to \infty$ *if, and only if,* $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$.

Proof. Since (X_i, Y_i), $i = 1,\dots,n$ are independent and identically distributed random variable,

$$\mathbb{P}(\exists i \in \{1,\dots,n\}, (X_i, Y_i) \in R_n(x)) = 1 - (1 - \mathbb{P}((X,Y) \in R_n(x)))^n \tag{8}$$

where

$$\begin{aligned}
\mathbb{P}((X,Y) \in R_n(x))) &= \mathbb{E}(\mathbb{I}\{X \in B(x,h) \cap Y \geq y_n\}) \\
&= \mathbb{E}(\mathbb{I}\{X \in B(x,h)\}\bar{F}(y_n|X)) \\
&= \bar{F}(y_n|x)\varphi_x(h) + \bar{F}(y_n|x)\mathbb{E}\left(\left(\frac{\bar{F}(y_n|X)}{\bar{F}(y_n|x)} - 1\right)\mathbb{I}\{X \in B(x,h)\}\right).
\end{aligned}$$

In view of Lemma 1(i), we have

$$\mathbb{E}\left(\left|\frac{\bar{F}(y_n|X)}{\bar{F}(y_n|x)} - 1\right|\mathbb{I}\{X \in B(x,h)\}\right) \leq 2(\kappa_c + \kappa_\gamma + \kappa_\varepsilon)\varphi_x(h) h \log y_n$$

and therefore

$$\mathbb{P}((X,Y) \in R_n(x)) = \bar{F}(y_n|x)\varphi_x(h)(1 + O(h \log y_n)).$$

Clearly, this probability converges to 0 as $n \to \infty$ and thus (8) can be rewritten as

$$\mathbb{P}(\exists i \in \{1,\dots,n\}, (X_i, Y_i) \in R_n(x)) = 1 - \exp\left(-n\varphi_x(h)\bar{F}(y_n|x)(1 + o(1))\right),$$

which converges to 1 if and only if $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$. ∎

Let us remark that the kernel estimator (1) can be rewritten as $\hat{\bar{F}}_n(y|x) = \hat{\psi}_n(y,x)/\hat{g}_n(x)$ with

$$\begin{aligned}
\hat{\psi}_n(y,x) &= \frac{1}{n\mu_x^{(1)}(h)} \sum_{i=1}^{n} K(d(x,X_i)/h) Q((Y_i - y)/\lambda), \\
\hat{g}_n(x) &= \frac{1}{n\mu_x^{(1)}(h)} \sum_{i=1}^{n} K(d(x,X_i)/h).
\end{aligned}$$

Lemma 5 and Lemma 6 are respectively dedicated to the asymptotic properties of $\hat{g}_n(x)$ and $\hat{\psi}_n(y,x)$.

Lemma 5 *Suppose* **(A.3)** *holds and let* $x \in E$ *such that* $\varphi_x(h) > 0$. *We have:*

(i) $\mathbb{E}(\hat{g}_n(x)) = 1$.

(ii) *If, moreover,* $\varphi_x(h) \to 0$ *as* $h \to 0$ *then*

$$0 < \liminf n\varphi_x(h)\,\mathrm{var}(\hat{g}_n(x)) \leq \limsup n\varphi_x(h)\,\mathrm{var}(\hat{g}_n(x)) < \infty.$$

Therefore, under **(A.3)**, if $\varphi_x(h) \to 0$ and $n\varphi_x(h) \to \infty$ then $\hat{g}_n(x)$ converges to 1 in probability.

Proof. **(i)** is straightforward.

(ii) Standard calculations yields

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

and Lemma 3 entails

$$(C_1/C_2)^2 \le \varphi_x(h)\frac{\mu_x^{(2)}(h)}{(\mu_x^{(1)}(h))^2} \le (C_2/C_1)^2.$$

The condition $\varphi_x(h) \to 0$ concludes the proof. ■

Lemma 6 *Suppose* **(A.1)** – **(A.4)** *hold. Let $x \in E$ such that $\varphi_x(h) > 0$ and introduce $y_{n,j} = a_j y_n(1 + o(1))$ for $j = 1, \dots, J$ with $0 < a_1 < a_2 < \dots < a_J$ and where J is a positive integer. If $y_n \to \infty$ such that $h \log y_n \to 0$, $\lambda/y_n \to 0$ and $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$ as $n \to \infty$, then*

(i) $\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) = \bar{F}(y_{n,j}|x)(1 + O(h \log y_n \vee \lambda/y_n))$, *for $j = 1, \dots, J$.*

(ii) *The random vector*

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

is asymptotically Gaussian, centered, with covariance matrix $C(x)$ where $C_{j,j'}(x) = a_{j\wedge j'}^{1/\gamma(x)}$ for $(j, j') \in \{1, \dots, J\}^2$.

Proof. **(i)** The (X_i, Y_i), $i = 1, \dots, n$ being identically distributed, we have

$$\begin{aligned}\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) &= \frac{1}{\mu_x^{(1)}(h)}\mathbb{E}\{K(d(x, X)/h)Q((Y - y_{n,j})/\lambda)\} \\ &= \frac{1}{\mu_x^{(1)}(h)}\mathbb{E}\{K(d(x, X)/h)\mathbb{E}(Q((Y - y_{n,j})/\lambda)|X)\}\end{aligned}$$

Taking account of **(A.4)**, it follows that

$$\mathbb{E}(Q((Y - y_{n,j})/\lambda)|X) = \bar{F}(y_{n,j}|X) + \int_{-1}^{1} Q'(u)(\bar{F}(y_{n,j} + \lambda u|X) - \bar{F}(y_{n,j}|X))du$$

and thus the bias can be expanded as

$$\mathbb{E}(\hat{\psi}_n(y_{n,j}, x)) - \bar{F}(y_{n,j}|x) =: T_{1,n} + T_{2,n}, \tag{9}$$

where we have defined

$$\begin{aligned}T_{1,n} &= \frac{1}{\mu_x^{(1)}(h)}\mathbb{E}\{K(d(x, X)/h)(\bar{F}(y_{n,j}|X) - \bar{F}(y_{n,j}|x))\}, \\ T_{2,n} &= \frac{1}{\mu_x^{(1)}(h)}\mathbb{E}\left\{K(d(x, X)/h)\bar{F}(y_{n,j}|X)\int_{-1}^{1} Q'(u)\left(\frac{\bar{F}(y_{n,j} + \lambda u|X)}{\bar{F}(y_{n,j}|X)} - 1\right)du\right\}.\end{aligned}$$

Focusing on $T_{1,n}$ and taking account of **(A.3)**, it follows that

$$\begin{aligned}T_{1,n} &= \frac{1}{\mu_x^{(1)}(h)}\mathbb{E}(K(d(x,X)/h)(\bar{F}(y_{n,j}|X) - \bar{F}(y_{n,j}|x))\mathbb{I}\{d(x,X) \le h\}) \\ &= \frac{\bar{F}(y_{n,j}|x)}{\mu_x^{(1)}(h)}\mathbb{E}\left(K(d(x,X)/h)\left(\frac{\bar{F}(y_{n,j}|X)}{\bar{F}(y_{n,j}|x)} - 1\right)\mathbb{I}\{d(x,X) \le h\}\right).\end{aligned}$$

Lemma 1(i) implies that

$$\left|\frac{\bar{F}(y_{n,j}|X)}{\bar{F}(y_{n,j}|x)} - 1\right|\mathbb{I}\{d(x,X) \le h\} \le 2(\kappa_c + \kappa_\gamma + \kappa_\varepsilon)h\log y_{n,j} \le 3(\kappa_c + \kappa_\gamma + \kappa_\varepsilon)h\log y_n,$$

eventually and therefore

$$|T_{1,n}| = \bar{F}(y_{n,j}|x)O(h\log y_n). \tag{10}$$

Let us now consider $T_{2,n}$. From Lemma 1(ii), for all $u \in [-1,1]$, we eventually have

$$\left|\frac{\bar{F}(y_{n,j} + \lambda u|X)}{\bar{F}(y_{n,j}|X)} - 1\right|\mathbb{I}\{d(x,X) \le h\} \le \left|\left(1 + \frac{\lambda u}{y_{n,j}}\right)^{2/\gamma(x)} - 1\right| \le C_{\gamma(x)}\frac{\lambda}{y_{n,j}},$$

since $\lambda/y_n \to 0$ as $n \to \infty$ and where $C_{\gamma(x)}$ is a positive constant. As a consequence,

$$\begin{aligned}|T_{2,n}| &\le C_{\gamma(x)}\frac{\lambda}{y_{n,j}}\frac{1}{\mu_x^{(1)}(h)}\mathbb{E}(K(d(x,X)/h)\bar{F}(y_{n,j}|X)) \\ &= C_{\gamma(x)}\frac{\lambda}{y_{n,j}}(\bar{F}(y_{n,j}|x) + T_{1,n}) = \bar{F}(y_{n,j}|x)O(\lambda/y_n) \end{aligned}\tag{11}$$

in view of (10). Collecting (9), (10) and (11) concludes the first part of the proof.

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

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

where, for all $i = 1,\dots,n$, the random variable $Z_{i,n}$ is defined by

$$\begin{aligned}n\Lambda_n(x)\mu_x^{(1)}(h)Z_{i,n} &= \left\{\sum_{j=1}^{J}\frac{\beta_j K(d(x,X_i)/h)Q((Y_i - y_{n,j})/\lambda)}{\bar{F}(y_{n,j}|x)}\right. \\ &- \left.\mathbb{E}\left(\sum_{j=1}^{J}\frac{\beta_j K(d(x,X_i)/h)Q((Y_i - y_{n,j})/\lambda)}{\bar{F}(y_{n,j}|x)}\right)\right\}.\end{aligned}$$

Clearly, $\{Z_{i,n},\ i = 1,\dots,n\}$ is a set of centered, independent and identically distributed random variables. Let us determine an asymptotic expansion of their variance:

$$\begin{aligned}\mathrm{var}(Z_{i,n}) &= \frac{1}{n^2(\mu_x^{(1)}(h))^2\Lambda_n^2(x)}\mathrm{var}\left(\sum_{j=1}^{J}\beta_j K(d(x,X_i)/h)\frac{Q((Y_i - y_{n,j})/\lambda)}{\bar{F}(y_{n,j}|x)}\right) \\ &= \frac{1}{n^2(\mu_x^{(1)}(h))^2\Lambda_n^2(x)}\beta^t B(x)\beta \\ &= \frac{\bar{F}(y_n|x)}{n\mu_x^{(2)}(h)}\beta^t B(x)\beta, \end{aligned}\tag{12}$$

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

$$
\begin{aligned}
B_{j,j'}(x) &= \frac{A_{j,j'}(x)}{\bar{F}(y_{n,j}|x)\bar{F}(y_{n,j'}|x)}, \\
A_{j,j'}(x) &= \text{cov}\left\{K(d(x,X)/h)Q((Y - y_{n,j})/\lambda),\ K(d(x,X)/h)Q((Y - y_{n,j'})/\lambda)\right\} \\
&= \mathbb{E}\left\{K^2(d(x,X)/h)Q((Y - y_{n,j})/\lambda)Q((Y - y_{n,j'})/\lambda)\right\} \\
&- \mathbb{E}\{K(d(x,X)/h)Q((Y - y_{n,j})/\lambda)\}\mathbb{E}\{K(d(x,X)/h)Q((Y - y_{n,j'})/\lambda)\} \\
&=: T_{3,n} - T_{4,n}.
\end{aligned}
$$

Let us first focus on $T_{3,n}$:

$$T_{3,n} = \mathbb{E}\{K^2(d(x,X)/h)\mathbb{E}(Q((Y - y_{n,j})/\lambda)Q((Y - y_{n,j'})/\lambda)|X)\} \tag{13}$$

and remark that

$$\mathbb{E}(Q((Y - y_{n,j})/\lambda)Q((Y - y'_{n,j})/\lambda)|X) =: \Omega(y_{n,j}, y_{n,j'}) + \Omega(y_{n,j'}, y_{n,j})$$

where we have defined

$$
\begin{aligned}
\Omega(y, z) &= \frac{1}{\lambda}\int_{\mathbb{R}} Q'((t - y)/\lambda)Q((t - z)/\lambda)\bar{F}(t|X)dt \\
&= \int_{-1}^{1} Q'(u)Q(u + (y - z)/\lambda)\bar{F}(y + u\lambda|X)du.
\end{aligned}
$$

Let us consider the case $j < j'$. We thus have $a_j < a_{j'}$ and consequently $(y_{n,j} - y_{n,j'})/\lambda \to -\infty$ as $n \to \infty$. Therefore, for n large enough $u + (y_{n,j} - y_{n,j'})/\lambda < -1$ and $Q(u + (y_{n,j} - y_{n,j'})/\lambda) = 0$. It follows that, eventually $\Omega(y_{n,j}, y_{n,j'}) = 0$. Similarly, for n large enough $Q(u + (y_{n,j'} - y_{n,j})/\lambda) = 1$ and

$$\Omega(y_{n,j'}, y_{n,j}) = \int_{-1}^{1} Q'(u)\bar{F}(y_{n,j'} + u\lambda|X)du.$$

For symmetry reasons, it follows that, for all $j \neq j'$,

$$\mathbb{E}(Q((Y - y_{n,j})/\lambda)Q((Y - y'_{n,j})/\lambda)|X) = \int_{-1}^{1} Q'(u)\bar{F}(y_{n,j\vee j'} + u\lambda|X)du = \mathbb{E}(Q((Y - y_{n,j\vee j'})/\lambda)|X),$$

and replacing in (13) yields

$$T_{3,n} = \mathbb{E}\{K^2(d(x,X)/h)\mathbb{E}(Q((Y - y_{n,j\vee j'})/\lambda)|X)\} = \mathbb{E}\{K^2(d(x,X)/h)Q((Y - y_{n,j\vee j'})/\lambda)\}.$$

Now, since K^2 is a kernel also satisfying assumption **(A.3)**, part **(i)** of the proof implies

$$T_{3,n} = \mu_x^{(2)}(h)\bar{F}(y_{n,j\vee j'}|x)(1 + O(h \log y_n \vee \lambda/y_n)), \tag{14}$$

for all $j \neq j'$. In the case where $j = j'$, by definition,

$$T_{3,n} = \mathbb{E}\{K^2(d(x,X)/h)\mathbb{E}(Q^2((Y - y_{n,j})/\lambda)|X)\}$$

where K^2 is a kernel also satisfying assumption **(A.3)** and where the pdf associated to Q^2 satisfies assumption **(A.4)**. Consequently, (14) also holds for $j = j'$. Second, part **(i)** of the proof implies

$$T_{4,n} = (\mu_x^{(1)}(h))^2 \bar{F}(y_{n,j}|x)\bar{F}(y_{n,j'}|x)(1 + O(h \log y_n \vee \lambda/y_n)).$$

As a consequence,

$$\begin{aligned} A_{j,j'}(x) &= \mu_x^{(2)}(h)\bar{F}(y_{n,j\vee j'}|x)(1 + O(h \log y_n \vee \lambda/y_n)) \\ &- (\mu_x^{(1)}(h))^2 \bar{F}(y_{n,j}|x)\bar{F}(y_{n,j'}|x)(1 + O(h \log y_n \vee \lambda/y_n)) \end{aligned}$$

leading to

$$B_{j,j'}(x) = \frac{\mu_x^{(2)}(h)}{\bar{F}(y_{n,j\wedge j'}|x)} \left(1 + O(h \log y_n \vee \lambda/y_n) - \frac{(\mu_x^{(1)}(h))^2}{\mu_x^{(2)}(h)} \bar{F}(y_{n,j\wedge j'}|x)(1 + O(h \log y_n \vee \lambda/y_n)) \right).$$

In view of Lemma 3, $(\mu_x^{(1)}(h))^2/\mu_x^{(2)}(h)$ is bounded and taking account of $\bar{F}(y_{n,j\wedge j'}|x) \to 0$ as $n \to \infty$ yields

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

Now, from the regular variation property (3), it is easily seen that

$$\bar{F}(y_{n,j\wedge j'}|x) = a_{j\wedge j'}^{-1/\gamma(x)} \bar{F}(y_n|x)(1 + o(1))$$

entailing $B_{j,j'}(x) = C_{j,j'}(x)\mu_x^{(2)}(h)/\bar{F}(y_n|x)(1 + o(1))$. Replacing in (12), it follows that

$$\text{var}(Z_{i,n}) = \frac{\beta^t C(x)\beta}{n}(1 + o(1)),$$

for all $i = 1, \ldots, n$. As a preliminary conclusion, $\text{var}(\Psi_n) \to \beta^t C(x)\beta$ as $n \to \infty$. Consequently, Lyapounov criteria for the asymptotic normality of sums of triangular arrays reduces to $\sum_{i=1}^n \mathbb{E}|Z_{i,n}|^3 = n\mathbb{E}|Z_{1,n}|^3 \to 0$ as $n \to \infty$. Next, remark that $Z_{1,n}$ is a bounded random variable:

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

in view of Lemma 3 and thus,

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

as $n \to \infty$ in view of Lemma 3. As a conclusion, Ψ_n converges in distribution to a centered Gaussian random variable with variance $\beta^t C(x)\beta$ for all $\beta \neq 0$ in $\mathbb{R}^J$. The result is proved. ■

5.2 Proofs of main results

Proof of Theorem 1. Keeping in mind the notations of Lemma 6, the following expansion holds

$$\Lambda_n^{-1}(x)\sum_{j=1}^J \beta_j \left(\frac{\hat{\bar{F}}_n(y_{n,j}|x)}{\bar{F}(y_{n,j}|x)} - 1\right) =: \frac{\Delta_{1,n} + \Delta_{2,n} - \Delta_{3,n}}{\hat{g}_n(x)}, \tag{15}$$

where

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

Let us highlight that assumptions $nh^2\varphi_x(h)\log^2(y_n)\bar{F}(y_n|x) \to 0$ and $n\varphi_x(h)\bar{F}(y_n|x) \to \infty$ imply that $h\log y_n \to 0$ as $n \to \infty$. Thus, from Lemma 6(ii), the random term $\Delta_{1,n}$ can be rewritten as

$$\Delta_{1,n} = \sqrt{\beta^t C(x)\beta}\,\xi_n, \tag{16}$$

where ξ_n converges to a standard Gaussian random variable. The nonrandom term $\Delta_{2,n}$ is controlled with Lemma 6(i):

$$\Delta_{2,n} = O(\Lambda_n^{-1}(x)(h\log y_n \vee \lambda/y_n)) = o(1). \tag{17}$$

Finally, $\Delta_{3,n}$ can be bounded by Lemma 5 and Lemma 3:

$$\Delta_{3,n} = O_P(\Lambda_n^{-1}(x)(n\varphi_x(h))^{-1/2}) = O_P(\bar{F}(y_n|x))^{1/2} = o_P(1). \tag{18}$$

Collecting (15)–(18), it follows that

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

Finally, $\hat{g}_n(x) \xrightarrow{P} 1$ concludes the proof. ■

Proof of Theorem 2. Introduce for $j = 1,\dots,J$,

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

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

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

We first focus on the nonrandom term $a_{n,j}(x)$. Under **(A.1)**, $\bar{F}(.|x)$ is differentiable. Thus, for all $j \in \{1,\ldots,J\}$ there exists $\theta_{n,j} \in (0,1)$ such that

$$\bar{F}(q(\alpha_{n,j}|x)|x) - \bar{F}(q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j|x) = -\sigma_{n,j}(x)z_j\bar{F}'(q_{n,j}|x), \tag{19}$$

where $q_{n,j} = q(\alpha_{n,j}|x) + \theta_{n,j}\sigma_{n,j}(x)z_j$. It is clear that $q(\alpha_{n,j}|x) \to \infty$ and $\sigma_{n,j}(x)/q(\alpha_{n,j}|x) \to 0$ as $n \to \infty$. As a consequence, $q_{n,j} \to \infty$ and thus **(A.1)** entails

$$\lim_{n\to\infty} \frac{q_{n,j}\bar{F}'(q_{n,j}|x)}{\bar{F}(q_{n,j}|x)} = -1/\gamma(x). \tag{20}$$

Moreover, since $q_{n,j} = q(\alpha_{n,j}|x)(1+o(1))$ and $\bar{F}(.|x)$ is regularly varying at infinity, it follows that $\bar{F}(q_{n,j}|x) = \bar{F}(q(\alpha_{n,j}|x)|x)(1+o(1)) = \alpha_{n,j}(1+o(1))$. In view of (19) and (20), we end up with

$$a_{n,j}(x) = \frac{v_{n,j}(x)\sigma_{n,j}(x)\alpha_{n,j}z_j}{\gamma(x)q(\alpha_{n,j}|x)}(1+o(1)) = z_j(1+o(1)). \tag{21}$$

Let us now turn to the random term $W_{n,j}(x)$. Defining $a_j = \tau_j^{-\gamma(x)}$, $y_{n,j} = q(\alpha_{n,j}|x) + \sigma_{n,j}(x)z_j$ for $j = 1,\ldots,J$ and $y_n = q(\alpha_n|x)$, we have $y_{n,j} = q(\alpha_{n,j}|x)(1+o(1)) = a_j y_n(1+o(1))$ since $q(.|x)$ is regularly varying at 0 with index $-\gamma(x)$. Using the same argument, it is easily shown that $\log y_n = -\gamma(x)\log(\alpha_n)(1+o(1))$. As a consequence, Theorem 1 applies and the random vector

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

converges to a centered Gaussian random variable with covariance matrix $C(x)$. Taking account of (21), we obtain that $\Phi_n(z_1,\ldots,z_J)$ converges to the cumulative distribution function of a centered Gaussian distribution with covariance matrix $\gamma^2(x)C(x)$ evaluated at $(z_1,\ldots,z_J)$, which is the desired result. ■

Proof of Theorem 3. The proof is based on the following expansion:

$$\frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)}(\log(\hat{q}_n^{\mathrm{W}}(\beta_n|x)) - \log(q(\beta_n|x))) = \frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)}(Q_{n,1} + Q_{n,2} + Q_{n,3})$$

where we have introduced

$$\begin{aligned}
Q_{n,1} &= \sigma_n^{-1}(x)(\hat{\gamma}_n(x) - \gamma(x)),\\
Q_{n,2} &= \frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)}\log(\hat{q}_n(\alpha_n|x)/q(\alpha_n|x)),\\
Q_{n,3} &= \frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)}(\log q(\alpha_n|x) - \log q(\beta_n|x) + \gamma(x)\log(\alpha_n/\beta_n)).
\end{aligned}$$

First, $Q_{n,1} \xrightarrow{d} \mathcal{N}(0, V(x))$ as a straightforward consequence of the assumptions. Second, Theorem 2 implies that $\hat{q}_n(\alpha_n|x)/q(\alpha_n|x) \xrightarrow{P} 1$ and

$$Q_{n,2} = \frac{\sigma_n^{-1}(x)}{\log(\alpha_n/\beta_n)} \left(\frac{\hat{q}_n(\alpha_n|x)}{q(\alpha_n|x)} - 1 \right) (1 + o_P(1)) = \frac{O_P(1)}{\log(\alpha_n/\beta_n)}.$$

Consequently, $Q_{n,2} \xrightarrow{P} 0$ as $n \to \infty$. Finally, from Lemma 2(i), $Q_{n,3} = O(\sigma_n^{-1}(x)\varepsilon(q(\alpha_n|x)|x))$, which converges to 0 in view of the assumptions. ■

Proof of Theorem 4. The following expansion holds for all $j = 1, \ldots, J$:

$$\log \hat{q}_n(\tau_j \alpha_n|x) = \log q(\alpha_n|x) + \log \left(\frac{q(\tau_j \alpha_n|x)}{q(\alpha_n|x)} \right) + \log \left(\frac{\hat{q}_n(\tau_j \alpha_n|x)}{q(\tau_j \alpha_n|x)} \right). \tag{22}$$

First, Lemma 2(ii) entails that

$$\log \left(\frac{q(\tau_j \alpha_n|x)}{q(\alpha_n|x)} \right) = \gamma(x) \log(1/\tau_j) + O(\varepsilon(q(\alpha_n|x)|x)), \tag{23}$$

where the $O(\varepsilon(q(\alpha_n|x)|x))$ is not necessarily uniform in $j = 1, \ldots, J$. Second, it follows from Theorem 2 that

$$\log \left(\frac{\hat{q}_n(\tau_j \alpha_n|x)}{q(\tau_j \alpha_n|x)} \right) = \sigma_n(x) \xi_{n,j} \tag{24}$$

where $(\xi_{n,1}, \ldots, \xi_{n,J})^t$ converges to a centered Gaussian random vector with covariance matrix $\gamma^2(x)\Sigma$. Replacing (23) and (24) in (22) yields

$$\log \hat{q}_n(\tau_j \alpha_n|x) = \log q(\alpha_n|x) + \gamma(x) \log(1/\tau_j) + \sigma_n(x)\xi_{n,j} + O(\varepsilon(q(\alpha_n|x)|x)),$$

for all $j = 1, \ldots, J$ and therefore, in view of the shift invariance property of ϕ, we have

$$\phi\left(\{\log \hat{q}_n(\tau_j \alpha_n|x)\}_{j=1,\ldots,J}\right) = \phi\left(\{\gamma(x) \log(1/\tau_j) + \sigma_n(x)\xi_{n,j} + O(\varepsilon(q(\alpha_n|x)|x))\}_{j=1,\ldots,J}\right).$$

A first order Taylor expansion yields:

$$\begin{aligned} \phi\left(\{\log \hat{q}_n(\tau_j \alpha_n|x)\}_{j=1,\ldots,J}\right) &= \phi\left(\gamma(x)v\right) + \sum_{j=1}^{J} (\sigma_n(x)\xi_{n,j} + O(\varepsilon(q(\alpha_n|x)|x))) \frac{\partial \phi}{\partial x_j}(\gamma(x)v) \\ &+ O_P\left(\sum_{j=1}^{J} (\sigma_n(x)\xi_{n,j} + O(\varepsilon(q(\alpha_n|x)|x)))^2 \right). \end{aligned}$$

Thus, under the condition $\sigma_n^{-1}(x)\varepsilon(q(\alpha_n|x)|x) \to 0$ as $n \to \infty$, it follows that

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

Taking into account of the scale invariance property of ϕ, we finally obtain

$$\sigma_n^{-1}(x)(\hat{\gamma}_n^{\phi}(x) - \gamma(x)) = \frac{1}{\phi(v)} \sum_{j=1}^{J} \xi_{n,j} \frac{\partial \phi}{\partial x_j}(\gamma(x)v) + o_P(1)$$

and the conclusion follows. ■

References

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

[2] J. Beirlant and Y. Goegebeur. Regression with response distributions of Pareto-type. *Computational Statistics and Data Analysis*, 42:595–619, 2003.

[3] A. Berlinet, A. Gannoun and E. Matzner-Løber. Asymptotic normality of convergent estimates of conditional quantiles. *Statistics*, 35:139–169, 2001.

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

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

[6] V. Chavez-Demoulin and A.C. Davison. Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society, series C.*, 54:207–222, 2005.

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

[8] A. Daouia, L. Gardes, S. Girard and A. Lekina. Kernel estimators of extreme level curves. *Test*, to appear.

[9] A.C. Davison and N.I. Ramesh. Local likelihood smoothing of sample extremes. *Journal of the Royal Statistical Society, series B*, 62:191–208, 2000.

[10] A.C. Davison and R.L. Smith. Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, series B*, 52:393–442, 1990.

[11] A. Dekkers and L. de Haan. On the estimation of the extreme-value index and large quantile estimation, *Annals of Statistics*, 17:1795–1832, 1989.

[12] J. Diebolt, L. Gardes, S. Girard and A. Guillou. Bias-reduced extreme quantile estimators of Weibull tail-distributions. *Journal of Statistical Planning and Inference*, 138(5): 1389–1401, 2008.

[13] J.H.J Einmahl. The empirical distribution function as a tail estimator, *Statistica Neerlandica*, 44:79–82, 1990.

[14] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events*, Springer, 1997.

[15] M. Falk, J. Hüsler and R.D. Reiss. *Laws of small numbers: Extremes and rare events*, 2nd edition, Birkhäuser, 2004.

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

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

[18] F. Ferraty and P. Vieu. *Nonparametric functional data analysis*. Springer, 2006.

[19] A.K. Gangopadhyay. A note on the asymptotic behavior of conditional extremes, *Statistics and Probability Letters*, 25:163–170, 1995.

[20] A. Gannoun. 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, 1990.

[21] A. Gannoun, S. Girard, C. Guinot and J. Saracco. Reference ranges based on nonparametric quantile regression, *Statistics in Medicine*, 21:3119-3135, 2002.

[22] L. Gardes and S. Girard. Estimating extreme quantiles of Weibull tail distributions. *Communications in Statistics-Theory and Methods*, 34(5):1065–1080, 2005.

[23] L. Gardes and S. Girard. A moving window approach for nonparametric estimation of the conditional tail index. *Journal of Multivariate Analysis*, 99:2368–2388, 2008.

[24] L. Gardes and S. Girard. Conditional extremes from heavy-tailed distributions: An application to the estimation of extreme rainfall return levels. *Extremes*, 13(2):177–204, 2010.

[25] L. Gardes, S. Girard and A. Lekina. Functional nonparametric estimation of conditional extreme quantiles. *Journal of Multivariate Analysis*, 101:419–433, 2010.

[26] I. Gijbels and L. Peng. Estimation of a support curve via order statistics. *Extremes*, 3:251–277, 2000.

[27] S. Girard and P. Jacob. Frontier estimation via kernel regression on high power-transformed data. *Journal of Multivariate Analysis*, 99:403–420, 2008.

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

[29] M.I. Gomes and M.J. Martins. Generalizations of the Hill estimator - asymptotic versus finite sample behaviour. *Journal of Statistical Planning and Inference*, 93:161–180, 2001.

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

[31] P. Hall and N. Tajvidi. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Statistical Science*, 15;153–167, 2000.

[32] B.M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3:1163–1174, 1975.

[33] A.P. Korostelev and A.B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New-York, 1993.

[34] J. Pickands. Statistical inference using extreme order statistics. *The Annals of Statistics*, 3:119–131, 1975.

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

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

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

[38] R.L. Smith. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone (with discussion). *Statistical Science*, 4:367–393, 1989.

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

[40] W. Stute. Conditional empirical processes. *The Annals of Statistics*, 14:638–647, 1986.

[41] I. Weissman. Estimation of parameters and large quantiles based on the k largest observations, *Journal of the American Statistical Association*, 73:812–815, 1978.

[42] Q. Yao. Conditional predictive regions for stochastic processes. *Technical report*, University of Kent at Canterbury, 1999.

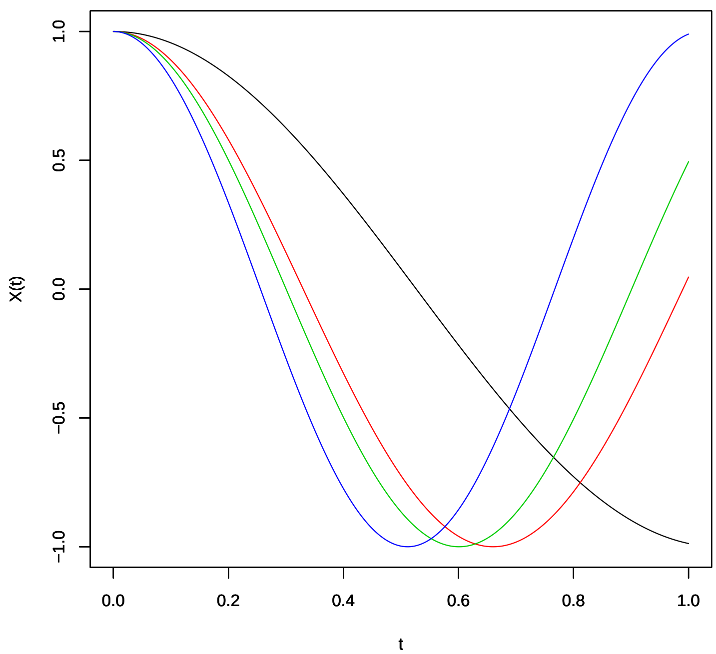

Figure 1: Four realizations of the random function $X(.)$.

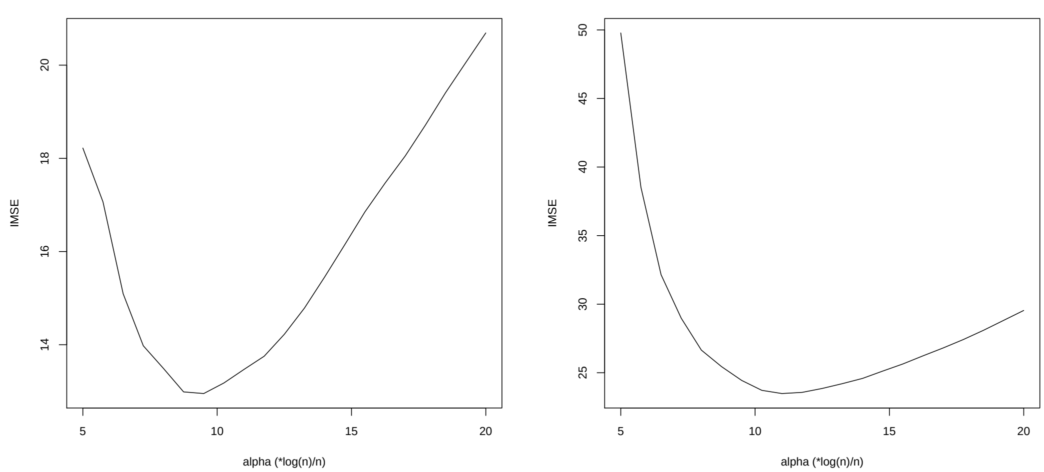

Figure 2: IMSE of $\hat{\gamma}_n^{\phi}$ as a function of α_n. Left: semi-metric d_Z, right: semi-metric d_X.

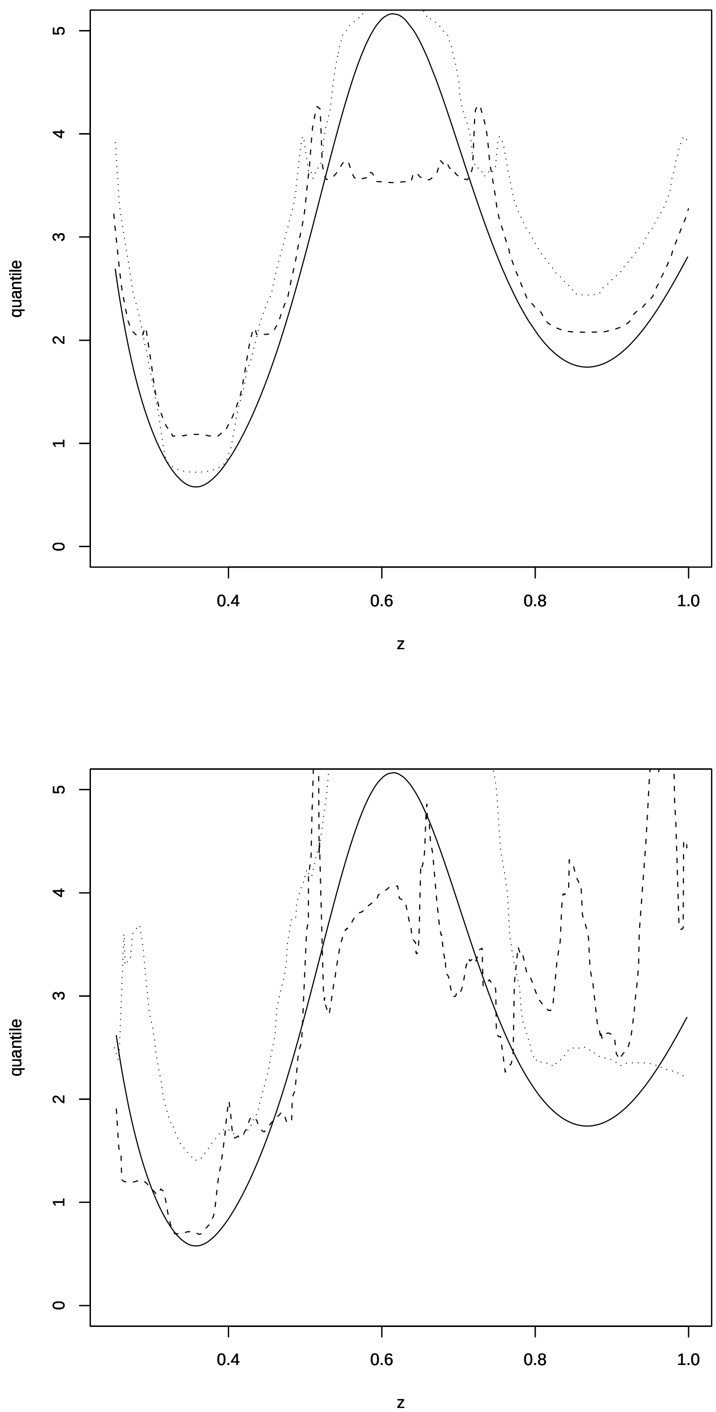

Figure 3: Comparison of the estimated quantile $\hat{q}_n^{\mathrm{W}}(\beta_n|x)$ corresponding to the median error with the true quantile function (continuous line). Horizontally: Z, vertically: quantiles. Two values of α_n are considered: α_n^{opt} (dashed line) and $\alpha_n = 15\log(n)/n$ (dotted line). Top: semi-metric d_Z, bottom: semi-metric d_X.

	$\bar{F}(y\|x)$	$\gamma(x)$	$c(x)$	$\varepsilon(y\|x)$
Pareto	$y^{-\theta(x)}$	$\dfrac{1}{\theta(x)}$	1	0
Cauchy	$\dfrac{1}{\pi}\tan^{-1}(1/y)+\dfrac{1}{2}(1-\text{sign}(y))$	1	$\dfrac{1}{4}$	$\dfrac{2}{3}\dfrac{1}{y^2}(1+o(1))$
Fréchet	$1-\exp(-y^{-\theta(x)})$	$\dfrac{1}{\theta(x)}$	$1-e^{-1}$	$\dfrac{\theta(x)}{2}y^{-\theta(x)}(1+o(1))$
Burr	$(1+y^{\tau(x)})^{-\lambda(x)}$	$\dfrac{1}{\lambda(x)\tau(x)}$	$2^{-\lambda(x)}$	$\lambda(x)\tau(x)y^{-\tau(x)}(1+o(1))$

Table 1: Examples of distributions satisfying **(A.1)**. Their parameters $\theta(x)$, $\tau(x)$ and $\lambda(x)$ are positive.