



**HAL**  
open science

## Conditional marginal expected shortfall

Yuri Goegebeur, Armelle Guillou, Nguyen Khanh Le Ho, Jing Qin

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou, Nguyen Khanh Le Ho, Jing Qin. Conditional marginal expected shortfall. *Extremes*, 2021, 24, pp.797-847. 10.1007/s10687-020-00403-1 . hal-02272392v4

**HAL Id: hal-02272392**

**<https://hal.science/hal-02272392v4>**

Submitted on 16 Dec 2020

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

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

# Conditional marginal expected shortfall

Yuri Goegebeur<sup>(1)</sup>, Armelle Guillou<sup>(2)</sup>, Nguyen Khanh Le Ho<sup>(1)</sup>, Jing Qin<sup>(1)</sup>

<sup>(1)</sup> Department of Mathematics and Computer Science, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

<sup>(2)</sup> Institut Recherche Mathématique Avancée, UMR 7501, Université de Strasbourg et CNRS, 7 rue René Descartes, 67084 Strasbourg cedex, France

## Abstract

In the context of bivariate random variables  $(Y^{(1)}, Y^{(2)})$ , the marginal expected shortfall, defined as  $\mathbb{E}(Y^{(1)}|Y^{(2)} \geq Q_2(1-p))$  for  $p$  small, where  $Q_2$  denotes the quantile function of  $Y^{(2)}$ , is an important risk measure, which finds applications in areas like, e.g., finance and environmental science. Our paper pioneers the statistical modeling of this risk measure when the random variables of main interest  $(Y^{(1)}, Y^{(2)})$  are observed together with a random covariate  $X$ , leading to the concept of the conditional marginal expected shortfall. The asymptotic behavior of an estimator for this conditional marginal expected shortfall is studied for a wide class of conditional bivariate distributions, with heavy-tailed marginal conditional distributions, and where  $p$  tends to zero at an intermediate rate. The finite sample performance is evaluated on a small simulation experiment. The practical applicability of the proposed estimator is illustrated on flood claim data.

## 1 Introduction

In the past years, many risk measures have been introduced in the literature, and these have been used to determine the amount of an asset to be kept in reserve in the financial framework. The most famous of these are the Value-At-Risk (VaR) defined for a random variable  $X$  as the  $p$ -quantile

$$Q(p) := \inf\{x \geq 0 : \mathbb{P}(X \leq x) \geq p\}, \text{ for } p \in (0, 1),$$

and the Conditional Tail Expectation (CTE) defined as

$$CTE_p[X] = \mathbb{E}(X|X > Q(p)), \text{ for } p \in (0, 1).$$

The latter risk measure is more conservative than the VaR for a same level of degree of confidence (see Landsman and Valdez 2003) and it also satisfies the desirable property of being a coherent risk measure as defined by Artzner et al. (1999). For all these reasons, the CTE has been extensively studied and also it has been generalized to the multivariate framework, see, e.g., Cai and Li (2005), Cai et al. (2015), and Di Bernardino and Prieur (2018). More precisely, if  $(Y^{(1)}, Y^{(2)})$  denotes a pair of risk factors, the CTE can be extended into  $\mathbb{E}(Y^{(1)}|Y^{(2)} > Q_2(p))$ , where  $Q_2(p)$  is the  $p$ -quantile of the risk  $Y^{(2)}$ . In such a multivariate context, this risk measure is well-known as the Marginal Expected Shortfall (MES). It was introduced by Acharya et al. (2010), and used to measure the contribution of a financial institution to an overall systemic risk. The ongoing global credit crisis and other former financial crises have demonstrated the vital

aspect of adequate risk measurement. For a financial firm, the MES is defined as its short-run expected equity loss conditional on the market taking a loss greater than its VaR. The MES is very simple to compute and therefore easy for regulators to consider. When estimating this risk measure, one often has the availability of additional information given by covariates, and these are important to take into account in order to obtain more precise estimates. This leads to the concept of conditional marginal expected shortfall.

Our paper is the first contribution dealing with the estimation of the conditional marginal expected shortfall when the random variables of main interest  $(Y^{(1)}, Y^{(2)})$  are recorded together with a random covariate  $X \in \mathbb{R}^d$ . We will denote by  $F_j(\cdot|x)$  the continuous conditional distribution function of  $Y^{(j)}, j = 1, 2$ , given  $X = x$ , and use the notation  $\bar{F}_j(\cdot|x)$  for the conditional survival function and  $U_j(\cdot|x)$  for the associated tail quantile function defined as  $U_j(\cdot|x) = \inf\{y : F_j(y|x) \geq 1 - 1/\cdot\}$ . Also, we will denote by  $f_X$  the density function of the covariate  $X$  and by  $x_0$  a reference position such that  $x_0 \in \text{Int}(S_X)$ , the interior of the support  $S_X \subset \mathbb{R}^d$  of  $f_X$ , which is assumed to be non-empty. Our aim will be to estimate the conditional marginal expected shortfall, given  $X = x_0$ , and defined as

$$\theta_p = \mathbb{E} \left[ Y^{(1)} \mid Y^{(2)} \geq U_2 \left( \frac{1}{p} \mid x_0 \right); x_0 \right],$$

where  $p$  is small. Note that in the financial and actuarial setting where risk measures and in particular MES have been introduced, one is often interested in positive risk factors. Thus, in the sequel, we consider the case where  $Y^{(1)}$  and  $Y^{(2)}$  are positive.

The remainder of the paper is organized as follows. In Section 2, we introduce our estimator for the conditional marginal expected shortfall and we establish its main asymptotic properties. Simulations are provided in Section 3 to illustrate the efficiency of our estimator, while in Section 4 the method is applied to a dataset of flood insurance claims. Some closing remarks are given in Section 5. All the proofs of the results are postponed to Section 6.

## 2 Estimator and asymptotic properties

We assume that  $Y^{(1)}$  and  $Y^{(2)}$  are both positive and follow a conditional Pareto-type model.

**Assumption (D)** For all  $x \in S_X$ , the conditional survival functions of  $Y^{(j)}, j = 1, 2$ , satisfy

$$\bar{F}_j(y|x) = A_j(x) y^{-1/\gamma_j(x)} \left( 1 + \frac{1}{\gamma_j(x)} \delta_j(y|x) \right),$$

where  $A_j(x) > 0$ ,  $\gamma_j(x) > 0$ , and  $|\delta_j(\cdot|x)|$  is normalised regularly varying with index  $-\beta_j(x)$ ,  $\beta_j(x) > 0$ , i.e.,

$$\delta_j(y|x) = B_j(x) \exp \left( \int_1^y \frac{\varepsilon_j(u|x)}{u} du \right),$$

with  $B_j(x) \in \mathbb{R}$  and  $\varepsilon_j(y|x) \rightarrow -\beta_j(x)$  as  $y \rightarrow \infty$ . Moreover, we assume  $y \rightarrow \varepsilon_j(y|x)$  to be a continuous function.

Under Assumption  $(\mathcal{D})$ ,  $F_1(\cdot|x)$  and  $F_2(\cdot|x)$  have density functions. Indeed, straightforward differentiation gives

$$f_j(y|x) = \frac{A_j(x)}{\gamma_j(x)} y^{-1/\gamma_j(x)-1} \left[ 1 + \left( \frac{1}{\gamma_j(x)} - \varepsilon_j(y|x) \right) \delta_j(y|x) \right], j = 1, 2. \quad (2.1)$$

Now, let  $(Y_i^{(1)}, Y_i^{(2)}, X_i)$ ,  $i = 1, \dots, n$ , be independent copies of  $(Y^{(1)}, Y^{(2)}, X)$ . We consider estimating the conditional marginal expected shortfall when  $p \rightarrow 0$  at an intermediate rate, i.e.,  $p = k/n$ , where  $k, n \rightarrow \infty$  such that  $k/n \rightarrow 0$ . A natural idea is then to study

$$\hat{\theta}_n := \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)} \mathbb{1}_{\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}},$$

where  $\hat{U}_2(\cdot|x_0)$  is an estimator for  $U_2(\cdot|x_0)$ , to be introduced later, and  $K_h(\cdot) := K(\cdot/h)/h^d$ , with  $K$  a joint density function on  $\mathbb{R}^d$ ,  $h \equiv h_n$  a positive non-random sequence of bandwidths with  $h \rightarrow 0$  if  $n \rightarrow \infty$ , and  $\mathbb{1}_A$  the indicator function on the event  $A$ . We observe that in our approach the covariate  $X$  is recorded together with the dependent variables  $(Y^{(1)}, Y^{(2)})$ , and hence we are in the situation of a random design. This should be contrasted with the fixed design, where the  $x$ -values are determined beforehand, and where the responses are collected at these specific design points, e.g., in a designed experiment.

To simplify the situation, let us assume for the moment that  $U_2(\cdot|x_0)$  is known and consider

$$\tilde{\theta}_n := \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)} \mathbb{1}_{\{Y_i^{(2)} \geq U_2(n/k|x_0)\}}.$$

Clearly, assuming  $F_1(y|x_0)$  is strictly increasing in  $y$ , we have

$$\begin{aligned} \tilde{\theta}_n &= \int_0^\infty \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(1)} \geq s\}} \mathbb{1}_{\{Y_i^{(2)} \geq U_2(n/k|x_0)\}} ds \\ &= \int_0^\infty \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(1)} \geq s, \bar{F}_2(Y_i^{(2)}|x_0) \leq k/n\}} ds \\ &= \int_0^\infty \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq (k/n) [\frac{n}{k} \bar{F}_1(s|x_0)], \bar{F}_2(Y_i^{(2)}|x_0) \leq k/n\}} ds \\ &= -U_1(n/k|x_0) \int_0^\infty \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq (k/n) s_n(u), \bar{F}_2(Y_i^{(2)}|x_0) \leq k/n\}} du^{-\gamma_1(x_0)}, \end{aligned}$$

where  $s_n(u) := \frac{n}{k} \bar{F}_1(u^{-\gamma_1(x_0)} U_1(n/k|x_0)|x_0)$ . Note that under  $(\mathcal{D})$ , we have  $s_n(u) \rightarrow u$  as  $n \rightarrow \infty$ .

The key statistic to consider is thus, for  $x_0 \in \text{Int}(S_X)$ ,

$$T_n(y_1, y_2|x_0) := \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y_i^{(2)}|x_0) \leq (k/n) y_2\}},$$

where  $y_1, y_2 > 0$ .

As a first main result we study the weak convergence of the process

$$\left\{ \sqrt{kh^d} \frac{T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))}{y_1^\eta}, \quad y_1, y_2 \in (0, T] \right\}, \quad (2.2)$$

for any  $T > 0$ , finite, and  $\eta \in [0, 1/2)$ . This will require some further assumptions.

In order to deal with the regression context,  $f_X$  and the functions appearing in  $F_1(y|x)$  and  $F_2(y|x)$  are assumed to satisfy the following Hölder conditions. Let  $\|\cdot\|$  denote some norm on  $\mathbb{R}^d$ .

**Assumption ( $\mathcal{H}$ )** *There exist positive constants  $M_{f_X}$ ,  $M_{A_j}$ ,  $M_{\gamma_j}$ ,  $M_{B_j}$ ,  $M_{\varepsilon_j}$ ,  $\eta_{f_X}$ ,  $\eta_{A_j}$ ,  $\eta_{\gamma_j}$ ,  $\eta_{B_j}$  and  $\eta_{\varepsilon_j}$ , where  $j = 1, 2$ , such that for all  $x, z \in S_X$ :*

$$\begin{aligned} |f_X(x) - f_X(z)| &\leq M_{f_X} \|x - z\|^{\eta_{f_X}}, \\ |A_j(x) - A_j(z)| &\leq M_{A_j} \|x - z\|^{\eta_{A_j}}, \\ |\gamma_j(x) - \gamma_j(z)| &\leq M_{\gamma_j} \|x - z\|^{\eta_{\gamma_j}}, \\ |B_j(x) - B_j(z)| &\leq M_{B_j} \|x - z\|^{\eta_{B_j}}, \\ \sup_{y \geq 1} |\varepsilon_j(y|x) - \varepsilon_j(y|z)| &\leq M_{\varepsilon_j} \|x - z\|^{\eta_{\varepsilon_j}}. \end{aligned}$$

We also impose a condition on the kernel function  $K$ , which is a standard condition in local estimation.

**Assumption ( $\mathcal{K}$ )**  *$K$  is a bounded density function on  $\mathbb{R}^d$ , with support  $S_K$  included in the unit ball in  $\mathbb{R}^d$ .*

Next, a uniform convergence result is needed for the joint conditional distribution of  $(Y^{(1)}, Y^{(2)})$ . This condition reflects the asymptotic behaviour of the conditional copula. Let  $R_t(y_1, y_2|x) := t\mathbb{P}(\overline{F}_1(Y^{(1)}|x) \leq y_1/t, \overline{F}_2(Y^{(2)}|x) \leq y_2/t | X = x)$ .

**Assumption ( $\mathcal{R}$ )** *For all  $x \in S_X$ , we have as  $t \rightarrow \infty$  and  $h \downarrow 0$  that*

$$R_t(y_1, y_2|x) \rightarrow R(y_1, y_2|x),$$

*uniformly in  $y_1, y_2 \in (0, T]$ , for any  $T > 0$ , and  $x \in B(x_0, h)$ .*

Note that Assumption ( $\mathcal{R}$ ) could also be formulated in an alternative way as follows: for all  $x \in S_X$  and  $(y_1, y_2) \in [0, \infty]^2 \setminus \{(\infty, \infty)\}$ , we have as  $t \rightarrow \infty$  and  $h \downarrow 0$  that

$$R_t(y_1, y_2|x) \rightarrow R(y_1, y_2|x),$$

uniformly in  $x \in B(x_0, h)$ , with  $x \rightarrow R(y_1, y_2|x)$  being a continuous function. This would then imply the uniformity of the convergence in  $(y_1, y_2) \in (0, T]^2$ .

Also note that Assumption ( $\mathcal{R}$ ) is a stronger version of the first order condition in Cai et al. (2015), since in our context the convergence must be uniform in a neighborhood of  $x_0$ , because of the local estimation. We also have the additional Assumptions ( $\mathcal{H}$ ) and ( $\mathcal{K}$ ), due to the

regression context. The latter are though common in the extreme value framework with random covariates. As a final comment, observe that in view of assumption  $(\mathcal{R})$ , our key statistic  $T_n(y_1, y_2|x_0)$  can be linked to the estimation of  $R(y_1, y_2|x_0)$ , though this is not the objective of the present paper. For what concerns the estimation of the conditional extremal dependence structure, we refer to de Carvalho (2016), Escobar-Bach et al. (2018a, b), Castro et al. (2018) and Mhalla et al. (2019).

The weak convergence of (2.2) is then established in the following theorem under the previous assumptions which are quite general, and therefore satisfied by a large class of models, some examples of them being presented in Section 3. Throughout the paper weak convergence is denoted by ‘ $\rightsquigarrow$ ’.

**Theorem 2.1.** *Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$ ,  $(\mathcal{R})$  with  $x \rightarrow R(y_1, y_2|x)$  being a continuous function,  $x_0 \in \text{Int}(S_X)$  with  $f_X(x_0) > 0$ , and  $y \rightarrow F_j(y|x_0)$ ,  $j = 1, 2$ , are strictly increasing. Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$ ,  $kh^d \rightarrow \infty$  and  $h^{\eta_{\gamma_1} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_1} \wedge \eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ . Then for  $\eta \in [0, 1/2)$ , we have,*

$$\sqrt{kh^d} \frac{T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))}{y_1^\eta} \rightsquigarrow \frac{W(y_1, y_2)}{y_1^\eta}, \quad (2.3)$$

in  $D((0, T]^2)$ , for any  $T > 0$ , where  $W(y_1, y_2)$  is a zero centered Gaussian process with covariance function

$$\mathbb{E}(W(y_1, y_2)W(\bar{y}_1, \bar{y}_2)) = \|K\|_2^2 f_X(x_0) R(y_1 \wedge \bar{y}_1, y_2 \wedge \bar{y}_2|x_0).$$

We also introduce the following weak convergence result for a related process. This process will be useful in establishing the asymptotic properties of the quantile estimator  $\hat{U}_2(n/k|x_0)$ . Let

$$\hat{f}_n(x_0) := \frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i)$$

be a classical kernel density estimator.

**Theorem 2.2.** *Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$ , and  $x_0 \in \text{Int}(S_X)$  with  $f_X(x_0) > 0$ . Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$ ,  $kh^d \rightarrow \infty$ ,  $h^{\eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{f_X} \wedge \eta_{A_2}} \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{\gamma_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} |\delta_2(U_2(n/k|x_0)|x_0)| h^{\eta_{B_2}} \rightarrow 0$ , and  $\sqrt{kh^d} |\delta_2(U_2(n/k|x_0)|x_0)| h^{\eta_{\varepsilon_2}} \rightarrow 0$ . Then, we have*

$$\sqrt{kh^d} \left( \frac{T_n(\infty, y_2|x_0)}{\hat{f}_n(x_0)} - y_2 \right) \rightsquigarrow \frac{W(\infty, y_2)}{f_X(x_0)}$$

in  $D((0, T])$ , for any  $T > 0$ , where  $W(\infty, y_2)$  is a zero centered Gaussian process with covariance function

$$\mathbb{E}(W(\infty, y_2)W(\infty, \bar{y}_2)) = \|K\|_2^2 f_X(x_0) (y_2 \wedge \bar{y}_2).$$

The joint weak convergence of the above two processes can be established by showing the joint finite dimensional weak convergence of them, combined with joint tightness. The joint finite dimensional convergence can be established by using the Cramér-Wold device (van der

Vaart 1998, p. 16). This is a standard but tedious calculation which is for brevity omitted from the paper. Note that the joint tightness follows from the individual tightness (similarly to Lemma 1 in Bai and Taqqu 2013).

The main result of this paper is the asymptotic normality of our final estimator for the conditional marginal expected shortfall  $\theta_{k/n}$ , defined as

$$\bar{\theta}_n := \frac{\hat{\theta}_n}{\hat{f}_n(x_0)} = \frac{\frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)} \mathbb{1}_{\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}}}{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i)}. \quad (2.4)$$

Intuitively, we can see this estimator as an extension of the Nadaraya-Watson estimator of  $m(x) := \mathbb{E}[Y|X = x_0]$  (see Nadaraya, 1964; Watson, 1964), defined as

$$\hat{m}(x) := \frac{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)}}{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i)},$$

in the sense that, since in our context the conditioning is on the event  $\{Y^{(2)} \geq U_2(n/k|x_0)\}$ , we need to add the indicator of the event  $\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}$  in (2.4). As a result, the numerator in (2.4) needs to be normalized by the number of excesses above this estimated quantile, i.e.,  $k$ , instead of  $n$ .

In order to obtain the weak convergence of  $\bar{\theta}_n$ , we need to introduce the following second order condition.

**Assumption (S).** *There exist  $\beta > \gamma_1(x_0)$  and  $\tau < 0$  such that, as  $t \rightarrow \infty$  and  $h \downarrow 0$  we have*

$$\sup_{x \in B(x_0, h)} \sup_{0 < y_1 < \infty, \frac{1}{2} \leq y_2 \leq 2} \frac{|R_t(y_1, y_2|x) - R(y_1, y_2|x_0)|}{y_1^\beta \wedge 1} = O(t^\tau).$$

Note that this uniform requirement excludes the case where  $(Y^{(1)}, Y^{(2)})$  are asymptotically upper tail independent given  $X = x_0$ , which corresponds to the case  $R(y_1, y_2|x_0) = 0$ . Assumption (S) is a stronger version of the second order condition in Cai et al. (2015), since the convergence must be additionally uniform in  $x \in B(x_0, h)$  due to the regression context.

**Theorem 2.3.** *Assume (D), (H), (K), (S) with  $x \rightarrow R(y_1, y_2|x)$  being a continuous function, and  $y \rightarrow F_j(y|x_0)$ ,  $j = 1, 2$ , are strictly increasing. Let  $x_0 \in \text{Int}(S_X)$  such that  $f_X(x_0) > 0$ . Consider sequences  $k = \lfloor n^\alpha \ell_1(n) \rfloor$  and  $h = n^{-\Delta} \ell_2(n)$ , where  $\ell_1$  and  $\ell_2$  are slowly varying functions at infinity, with  $\alpha \in (0, 1)$  and*

$$\begin{aligned} & \max \left( \frac{\alpha}{d + 2\gamma_1(x_0)(\eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha}{d + 2(1 - \gamma_1(x_0))(\eta_{A_2} \wedge \eta_{\gamma_2} \wedge \eta_{B_2} \wedge \eta_{\varepsilon_2} \wedge \eta_{f_X})} \right), \\ & \frac{\alpha}{d} - \frac{2(1 - \alpha)\gamma_1^2(x_0)\beta_1(x_0)}{d + d(\beta_1(x_0) + \varepsilon)\gamma_1(x_0)}, \frac{\alpha - 2(1 - \alpha)(\gamma_1(x_0) \wedge (\beta_2(x_0)\gamma_2(x_0)) \wedge (-\tau))}{d} \Big) \\ & < \Delta < \frac{\alpha}{d}. \end{aligned}$$

Then, for  $\gamma_1(x_0) < 1/2$ , we have

$$\sqrt{kh^d} \left( \frac{\bar{\theta}_n}{\theta_{k/n}} - 1 \right) \rightsquigarrow -(1 - \gamma_1(x_0)) \frac{W(\infty, 1)}{f_X(x_0)} + \frac{1}{f_X(x_0)} \frac{\int_0^\infty W(s, 1) ds^{-\gamma_1(x_0)}}{\int_0^\infty R(s, 1|x_0) ds^{-\gamma_1(x_0)}}.$$

The variance of the limiting random variable in Theorem 2.3, denoted  $\mathbb{W}$ , is given by

$$\text{Var}(\mathbb{W}) = \frac{\|K\|_2^2}{f_X(x_0)} \left[ \gamma_1^2(x_0) - 1 - \frac{\int_0^\infty R(s, 1|x_0) ds^{-2\gamma_1(x_0)}}{\left(\int_0^\infty R(s, 1|x_0) ds^{-\gamma_1(x_0)}\right)^2} \right]. \quad (2.5)$$

The conditions on  $k$  and  $h$  in Theorem 2.3 are due to the method of proof of the auxiliary result given in Lemma 6.4 of the Appendix. Also in Cai et al. (2015) one needed a condition on the growth of  $k$ , but in the context without covariates.

### 3 Simulation experiment

In this section we evaluate the finite sample behavior of the proposed estimator with a simulation experiment. Since the topic of estimating the marginal expected shortfall in the regression context is completely new in the literature, we cannot compare our estimator with any alternative estimator. We will compare it with the true value of the models under consideration.

We simulate from the following models:

**Model 1.** We consider the logistic copula model

$$C(u_1, u_2|x) = e^{-[(-\ln u_1)^x + (-\ln u_2)^x]^{1/x}}, \quad u_1, u_2 \in [0, 1], x \geq 2. \quad (3.1)$$

We take  $X \sim U[2, 10]$ , and combine this copula model with Fréchet distributions for  $Y^{(1)}$  and  $Y^{(2)}$ :

$$F_j(y) = e^{-y^{-1/\gamma_j}}, \quad y > 0,$$

$j = 1, 2$ . We set  $\gamma_1 = 0.25$  and  $\gamma_2 = 0.5$ . This model satisfies  $(\mathcal{S})$  with  $R(y_1, y_2|x) = y_1 + y_2 - (y_1^x + y_2^x)^{1/x}$ ,  $\tau = -1$  and  $\beta = 1 - \varepsilon$  for some small  $\varepsilon > 0$ .

**Model 2.** The conditional distribution of  $(Y^{(1)}, Y^{(2)})$  given  $X = x$  is that of

$$(|Z_1|^{\gamma_1(x)}, |Z_2|^{\gamma_2(x)}),$$

where  $(Z_1, Z_2)$  follow a bivariate standard Cauchy distribution with density function

$$f(z_1, z_2) = \frac{1}{2\pi} (1 + z_1^2 + z_2^2)^{-3/2}, \quad (z_1, z_2) \in \mathbb{R}^2.$$

We take  $X \sim U[0, 1]$ , and set

$$\begin{aligned} \gamma_1(x) &= 0.4 [0.1 + \sin(\pi x)] \left[ 1.1 - 0.5e^{-64(x-0.5)^2} \right], \\ \gamma_2(x) &= 0.1 + 0.1x. \end{aligned}$$

This model satisfies  $(\mathcal{S})$  with  $R(y_1, y_2|x) = y_1 + y_2 - \sqrt{y_1^2 + y_2^2}$ ,  $\tau = -1$  and  $\beta = 2$  (see, e.g., Cai et al. 2015, in the context without covariates).



**Model 3.** We consider the logistic copula model from (3.1), with  $X \sim U[2, 10]$ , combined with conditional Burr distributions for  $Y^{(1)}$  and  $Y^{(2)}$ :

$$F_j(y|x) = 1 - \left( \frac{\beta_j}{\beta_j + y^{\tau_j(x)}} \right)^{\lambda_j}, \quad y > 0; \beta_j, \lambda_j, \tau_j(x) > 0,$$

$j = 1, 2$ . We set  $\beta_1 = \beta_2 = 1$ ,  $\lambda_1 = 1$ ,  $\lambda_2 = 0.5$ , and

$$\tau_1(x) = 2e^{0.2x}, \quad \tau_2(x) = 8/\sin(0.3x).$$

Similarly to Model 1, this model satisfies (S).

The marginal conditional distributions in the above models are standard heavy-tailed distributions that satisfy (D), see, e.g., Beirlant et al. (2009). In Model 2, the functions  $\gamma_1(x)$  and  $\gamma_2(x)$  are usual functions in the extreme value framework, see, e.g., Dierckx et al. (2014), Goegebeur et al. (2014). Note that for Model 2 the function  $\gamma_1(x)$  is much more complicated than  $\gamma_2(x)$  since it shows local maxima and minima, whereas  $\gamma_2(x)$  is increasing. For Model 3,  $\gamma_1(x)$  is decreasing while  $\gamma_2(x)$  has a maximum in the interval  $[2, 10]$ . The graphs of these functions are shown in Figure 1 for Model 2 and in Figure 2 for Model 3.

Concerning the kernel function  $K$ , we take the bi-quadratic function

$$K(x) = \frac{15}{16}(1 - x^2)^2 \mathbb{1}_{\{x \in [-1, 1]\}}.$$

To compute our estimator  $\bar{\theta}_n$ , the bandwidth  $h$  need to be chosen. To this aim, we use the cross validation criterion introduced by Yao (1999), and used in an extreme value context by Daouia et al. (2011, 2013) and Escobar-Bach et al. (2018a):

$$h_{cv} := \operatorname{argmin}_{h \in \mathcal{H}} \sum_{i=1}^n \sum_{j=1}^n \left( \mathbb{1}_{\{Y_i^{(2)} \leq Y_j^{(2)}\}} - \hat{F}_{n,h,2,-i}(Y_j^{(2)} | X_i) \right)^2,$$

where  $\mathcal{H}$  is the grid of values defined as  $R_X \times \{0.05, 0.10, \dots, 0.30\}$ , with  $R_X$  the range of the covariate  $X$ , and

$$\hat{F}_{n,h,2,-i}(y|x) := \frac{\sum_{k=1, k \neq i}^n K_h(x - X_k) \mathbb{1}_{\{Y_k^{(2)} \leq y\}}}{\sum_{k=1, k \neq i}^n K_h(x - X_k)}.$$

In Figure 3 we show the boxplots of  $\bar{\theta}_n$  based on 500 replications, at different values of  $x_0$ , for samples of size  $n = 500$  (left) and  $n = 1000$  (right), and for  $k/n = 2\%$  (top row) and  $k/n = 10\%$  (bottom row), along with the true value of  $\theta_{k/n}$ . Figures 4 and 5 are constructed similarly but for Model 2 and Model 3, respectively. The values of  $\theta_{k/n}$  are computed with numerical integration.

From these figures we can draw the following conclusions:

- Overall the estimator  $\bar{\theta}_n$  performs quite well, with the true value  $\theta_{k/n}$  typically located in the central box, but obviously the performance depends on the model and also on the position  $x_0$ . The best results are obtained for Model 1, where the dependence structure depends on  $x$  but the marginal distributions are covariate independent. Model 2 has covariate dependent marginal distributions, but  $R(y_1, y_2|x)$  does not depend on  $x$ , and for Model 3 both the margins and  $R(y_1, y_2|x)$  depend on  $x$ . These models are more challenging than Model 1, but the estimator continues to perform well.

- The estimator behaves as expected in  $n$ , namely, for a fixed  $k/n$  the variance decreases with increasing sample size.
- As a function of  $x$ , the true value  $\theta_{k/n}$  follows the pattern of  $\gamma_1(x)$  rather closely. Similarly, the variance of  $\bar{\theta}_n$  tends to be larger for large values of  $\gamma_1(x)$ .
- The estimator seems to be not too much sensitive on the value of the covariate  $x_0$  in case of Model 1. On the contrary, for Model 2, it depends a lot on the value of the covariate. Note that nearby the local maxima there is a slight underestimation of  $\theta_{k/n}$  which can be explained by the fact that in the local estimation one uses  $Y^{(1)}$  observations from conditional distributions with a smaller tail index than at the reference position  $x_0$ . An opposite behavior is observed at the local minima. Model 3 is in between, with some improvement in the variability of the estimates when the covariate increases, which may be explained by the fact that  $\gamma_1(x)$  decreases in  $x$ .

Next, in Figure 6 we provide some normal quantile plots of  $\sqrt{kh} \ln \bar{\theta}_n / \theta_{k/n}$ , with  $k$  taken as 5% of  $n$  and  $h$  obtained from the above mentioned cross-validation criterion. The rows of Figure 6 correspond with Models 1-3, respectively, while the columns represent the sample sizes,  $n = 500$  and  $n = 1000$ , respectively. For all models and sample sizes, the normal quantile plots show a quite linear pattern, confirming the validity of the normal approximation. Moreover, with increasing  $n$  the normal approximation improves slightly.

## 4 Application to flood insurance claim data

### 4.1 Description of the dataset

In this section we illustrate the practical applicability of the method on a dataset of flood insurance claims. Recently, the Federal Emergency Management Agency (FEMA) has released millions of records from the National Flood Insurance Program (NFIP). In particular, this database contains approximately 2.4 million damage claims dating back to 1978, where for each claim one has information on the date of the flood, location of the property (latitude and longitude), claim amount, and on insurance policy and building characteristics. As such, it provides important information for policymakers, researchers, insurers and prospective homebuyers. The dataset is publicly available on <https://www.fema.gov/media-library/assets/documents/180374>. For our purposes we consider the data from the year 2000 on, and define  $Y_1$  as the sum of the amount paid on the building claim, the content claim and the increased cost of compliance claim,  $Y_2$  is taken as the sum of the insured amount for the building and content, while the covariate  $X$  consists of  $X_1$  : latitude,  $X_2$  : longitude and  $X_3$  : date of loss. Interest is in estimating the expected claim amount conditional on an insured capital that exceeds a high quantile, and for a given location and time. For (re-)insurance companies, accurate modeling and analysis of the upper tail of the claim size distribution is of crucial importance, as extreme claims may pose a major threat to their solvency. Also, the upper tail of the claim size distribution provides important information for pricing (re-)insurance contracts. By taking covariate information into account one can differentiate the risks one is exposed to and hence obtain a more competitive premium calculation.

## 4.2 Preliminary analysis

The estimation method was implemented with the same cross-validation criterion as in the simulation section, including the same choice for  $\mathcal{H}$ , after standardizing the covariates to the interval  $[0, 1]$ . As for the kernel function, we used the bi-quadratic kernel, generalized to the case  $d = 3$ , as follows

$$K_h(x) = \frac{1}{h^3} K\left(\frac{\|x\|}{h}\right),$$

where  $x \in \mathbb{R}^3$ , and  $\|\cdot\|$  denotes the Euclidean norm.

In order to verify the Pareto-type behavior of  $Y^{(1)}$  and  $Y^{(2)}$ , we constructed the local Hill plots of the  $Y^{(1)}$  and  $Y^{(2)}$  data, respectively, for which the  $X$  coordinate is in a neighborhood of (latitude, longitude)=(33.84,-84.45), and date of loss equal to 2018, July, see Figure 7. The location under consideration is in the city of Atlanta. In these plots we show the local Hill estimates  $H_k^{(j)}(x_0) := \frac{1}{k} \sum_{i=1}^k \ln \tilde{Y}_{n_{x_0}-i+1, n_{x_0}}^{(j)} - \ln \tilde{Y}_{n_{x_0}-k, n_{x_0}}^{(j)}$  as a function of  $k$ , where  $\tilde{Y}_{i, n_{x_0}}^{(j)}$ ,  $i = 1, \dots, n_{x_0}$ , are the order statistics of the  $Y^{(j)}$  data for which the  $X$  coordinate belongs to  $B(x_0, h)$ , and  $n_{x_0}$  is the number of observations in  $B(x_0, h)$ . For both  $Y^{(1)}$  and  $Y^{(2)}$  the Hill estimate is clearly positive for the smaller  $k$  values supporting the assumption of underlying conditional Pareto-type distributions. For  $Y^{(1)}$ , total claim amount, the Hill plot shows a stable estimate for  $\gamma_1(x_0)$  of about 0.3 when  $k$  is in the range 50-200. This satisfies the theoretical requirement that  $\gamma_1(x_0) < 0.5$ . For  $Y^{(2)}$ , capital insured, the Hill plot shows some systematic pattern beyond  $k = 150$ , which is due to the occurrence of repeated values for this variable. Despite this, the local Hill plot also suggests an underlying conditional Pareto-type distribution. Similar local Hill plots were obtained at other locations and for other time points. Next, we investigate the asymptotic dependence assumption by plotting an estimate for  $R(1.5, 1.5|x_0)$  as a function of  $k$ , at several values of  $x_0$ . Lemma 6.2 gives an indication that  $T_n(y_1, y_2|x_0)$  estimates  $f_X(x_0)R(y_1, y_2|x_0)$ , but  $T_n(y_1, y_2|x_0)$  is only a pseudo estimator as it depends on the unknown marginal conditional distribution functions  $F_1(\cdot|x_0)$  and  $F_2(\cdot|x_0)$ . To resolve this, we consider an adjustment of the estimator proposed in Escobar-Bach et al. (2018b), which is for the context of estimation of  $L(y_1, y_2|x_0)$ , to the context of estimation of  $R(y_1, y_2|x_0)$ , namely

$$\hat{R}(y_1, y_2|x_0) = \frac{\frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\left\{ \hat{F}_{n,1}(Y_i^{(1)}|X_i) \leq \frac{k}{n} y_1, \hat{F}_{n,2}(Y_i^{(2)}|X_i) \leq \frac{k}{n} y_2 \right\}}}{\hat{f}_n(x_0)}, \quad (4.1)$$

where  $\hat{F}_{n,1}$  is a kernel estimator for  $F_1$ , of the same form as  $\hat{F}_{n,2}$ . In Figure 8 we plot  $\hat{R}(1.5, 1.5|x_0)$  as a function of  $k$ , at  $x_0$  corresponding to the above mentioned location, and for 2009, July 1 (left), and 2017, January 1 (right). Clearly, the displays show a positive estimate for  $R(1.5, 1.5|x_0)$ , which gives evidence of asymptotic dependence of  $Y^{(1)}$  and  $Y^{(2)}$  given  $X = x_0$ .

## 4.3 Estimation of the conditional marginal expected shortfall

We illustrate the estimation of the conditional marginal expected shortfall at the above mentioned location, for the period 2008 till present, and using  $k/n = 1\%$  (solid line) and  $k/n = 10\%$  (dashed-dotted line), respectively, see Figure 9. As expected, the conditional marginal expected shortfall at quantile level  $k/n = 1\%$  shows more variability than the one at level  $k/n = 10\%$ ,

due to the smaller amount of data available to estimate the former, but otherwise they show the same pattern. The plot shows clearly the catastrophic Atlanta flood in 2009, September, resulting from multiple days of prolonged rainfall. The height of the event was on September 20-21 where 10 to 20 inches of rain occurred in less than 24 hours, which led to flash flooding, with flooded river basins remaining swollen for weeks. For this period, the difference between the two levels of the conditional marginal expected shortfall is larger than at other time points included in the analysis, which can be probably explained by the increased frequency of very large damage claims. Note that although the flood event took place at some specific time, its effect is smoothed out over a window due to the local estimation. This is in our setting also partially due to the use of a global bandwidth parameter  $h$ , i.e., a bandwidth that gives a reasonable performance over the whole covariate range. Alternatively, the bandwidth could be selected locally, for the specific  $x_0$  where one wants to do the estimation. One can expect then that at  $x_0$  where the parameter of interest, here  $\theta_p$ , changes quickly, the resulting bandwidth would be smaller. This is a topic for future investigation. Finally, we supplement Figure 9 with pointwise 95% confidence intervals, see Figure 10. The approximate confidence intervals of level  $100(1 - \alpha)\%$  for  $\theta_{k/n}$  are obtained from Theorem 2.3 and given by

$$\left[ \frac{\bar{\theta}_n}{1 + \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\widehat{\text{Var}}(\mathbb{W})}{kh^d}}}, \frac{\bar{\theta}_n}{1 - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\frac{\widehat{\text{Var}}(\mathbb{W})}{kh^d}}} \right],$$

where  $\Phi^{-1}$  denotes the standard normal quantile function and  $\widehat{\text{Var}}(\mathbb{W})$  is an estimate for the asymptotic variance given in (2.5), obtained by using a local Hill estimate for  $\gamma_1(x_0)$  and (4.1) as estimate for  $R(y_1, y_2|x_0)$ . Note that the latter both require an adaptive selection of their respective  $k$ -value, which will be denoted by  $\bar{k}$  to avoid confusion with the  $k$  from  $\theta_{k/n}$ . These adaptive  $\bar{k}$ -values are obtained by plotting the estimates as a function of  $\bar{k}$  whereafter the  $\bar{k}$  is selected by a stability criterion as described in Goegebeur et al. (2019). In Figure 10 we show the approximate pointwise 95% confidence intervals for  $\theta_{k/n}$  with  $k/n = 1\%$  (left) and  $k/n = 10\%$  (right) as a function of time, at the above considered location. Note that the confidence intervals seem reasonable, and are, e.g., wider for  $\theta_{k/n}$  with  $k/n = 1\%$  than for  $k/n = 10\%$ , as expected. At a few  $x_0$  positions we could not obtain a confidence interval, either due to a negative estimate of  $\widehat{\text{Var}}(\mathbb{W})$  or a negative value for  $1 - \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \sqrt{\widehat{\text{Var}}(\mathbb{W})/(kh^d)}$ .

## 5 Closing remarks

This paper is a pioneering contribution to the statistical modeling of the conditional marginal expected shortfall in the asymptotic dependence framework. It also provides a series of interesting open questions which will lead to further investigations, among them:

- The extrapolation outside the  $Y^{(2)}$  data range: due to the conditions  $k, n \rightarrow \infty$  with  $k/n \rightarrow 0$ , the  $Y^{(2)}$  quantile is intermediate, and the estimator  $\bar{\theta}_n$  cannot be used for extrapolation. Clearly this case where  $p < 1/n$  will be not trivial and it will require again an elaborate analysis. A natural idea will be to use a Weissman-type construction (see, Weissman 1978) based on the fact that under Assumptions  $(\mathcal{R})$  and  $(\mathcal{D})$  with  $\gamma_1(x_0) < 1$ ,

one can show (see Cai et al. 2015, Proposition 1) that

$$\lim_{p \rightarrow 0} \frac{\theta_p}{U_1(1/p|x_0)} = - \int_0^\infty R(s, 1|x_0) ds^{-\gamma_1(x_0)},$$

from which the following approximation can be deduced

$$\theta_p \sim \frac{U_1(1/p|x_0)}{U_1(n/k|x_0)} \theta_{\frac{k}{n}} \sim \left( \frac{k}{np} \right)^{\gamma_1(x_0)} \theta_{\frac{k}{n}}. \quad (5.1)$$

To construct an estimator based on (5.1), we need first to define an estimator for  $\gamma_1(x_0)$  and then we need to establish its asymptotic theory in terms of a process, related to the process we have already studied.

- The extension to the case of a real-valued  $Y^{(1)}$  or  $Y^{(2)}$ . In our approach we have assumed that  $Y^{(1)}$  and  $Y^{(2)}$  are non-negative random variables. This was also the case in Joe and Li (2011), and Das and Fasen-Hartmann (2018, 2019) where various properties of the marginal expected shortfall were studied in a framework without covariates. The problem of handling a real valued  $Y^{(1)}$  requires an elaborate analysis, which does not fit in the current framework of our proofs.
- Study of the bias-properties of the proposed estimator. We have studied our estimator under Assumption ( $\mathcal{H}$ ), which allowed us to obtain the order of the various bias-terms, and hence to control them. However, in this approach one cannot make the bias terms explicit. A more precise quantification of the bias terms can be done by replacing Assumption ( $\mathcal{H}$ ) by suitable differentiability conditions on the various parameters (functions of  $x$ ), along with using Taylor series expansions in the theoretical derivations. Alternatively, the estimation of the conditional marginal expected shortfall could also be studied in the framework of local polynomial maximum likelihood estimation. Such estimators are well-known to have an attractive behaviour near the boundary of the support, as documented in Wand and Jones (1995), Fan and Gijbels (1996), and Aerts and Claeskens (1997). In the context of estimation of the conditional marginal expected shortfall, a local polynomial estimator could be obtained by minimisation of

$$S := \sum_{i=1}^n K_h(x_0 - X_i) \left( Y_i^{(1)} - \sum_{j=0}^s \beta_j (X_i - x_0)^j \right)^2 \mathbb{1}_{\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}},$$

where we have for simplicity assumed that  $d = 1$ . In particular, the estimator for the conditional marginal expected shortfall at  $x_0$  would in this case be  $\hat{\theta}_n^{(LP)} = \hat{\beta}_0$ , where  $\hat{\beta}_0$  is the estimator for  $\beta_0$  obtained from solving the above minimisation problem. Interestingly, by taking  $s = 0$ , corresponding to a local constant estimation, one obtains as estimator

$$\check{\theta}_n := \frac{\sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)} \mathbb{1}_{\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}}}{\sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} \geq \hat{U}_2(n/k|x_0)\}}},$$

which is in nature close to the estimator proposed in the present paper.

- High dimensional covariates. Although our theoretical developments allow for an arbitrary dimension  $d$  of the covariate, local estimators are known to have a deteriorating practical performance in higher dimensions due to the so-called curse of dimensionality. With high dimensional covariates, the local estimation could be combined with dimension reduction techniques, e.g., Gardes (2018) and Xu et al. (2020) for dimension reduction in the context of estimation of extreme conditional quantiles.
- The extension to the asymptotic independence framework: in the absence of covariates, this topic has been very recently studied in the literature in, e.g., Das and Fasen-Hartmann (2018) and Cai and Musta (2020).
- Confidence bands. The result of Theorem 2.3 allows to construct confidence intervals for  $\theta_{k/n}$  at a specific  $x_0$ , so pointwise, as illustrated in the real data analysis. A natural question would be to know if a result for  $\bar{\theta}_n$  with a varying  $x_0$ -value in the covariate space can be obtained. This means, in that case, that we inquire about a convergence of  $\bar{\theta}_n$  correctly normalized as a stochastic process in  $x_0$ . However, even in the simpler case of local estimation of conditional tail index  $\gamma(x_0)$  of a Pareto-type tail with random covariates, it remains uncertain whether this type of result is possible. In our context this is highly more complicated than this latter framework, and thus the problem is still open.
- The development of a completely automatic data-driven way for obtaining confidence intervals, along with a study of the corresponding coverage probabilities.

## 6 Appendix

**Lemma 6.1.** *Assume  $(\mathcal{D})$  and  $(\mathcal{H})$  and  $x_0 \in \text{Int}(S_X)$ . Let  $(t_n)_{n \geq 1}$  and  $(h_n)_{n \geq 1}$  be arbitrary sequences satisfying  $t_n \rightarrow \infty$  and  $h_n \rightarrow 0$  such that  $h_n^{\eta_j \wedge \eta_{\varepsilon_j}} \ln t_n \rightarrow 0$ , as  $n \rightarrow \infty$ , and  $0 \leq \eta < 1$ . Then*

$$\left| \frac{t_n \bar{F}_j(U_j(t_n/y|x_0)|x)}{y^\eta} - y^{1-\eta} \right| \rightarrow 0, \text{ as } n \rightarrow \infty,$$

*uniformly in  $y \in (0, T]$  and  $x \in B(x_0, h_n)$ .*

**Lemma 6.2.** *Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$  and  $(\mathcal{R})$  with  $x \rightarrow R(y_1, y_2|x)$  being a continuous function, and  $x_0 \in \text{Int}(S_X)$  such that  $f_X(x_0) > 0$ . Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$  in such a way that  $k/n \rightarrow 0$  and  $h^{\eta_{\gamma_1} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_1} \wedge \eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ . Then, as  $n \rightarrow \infty$*

$$\begin{aligned} \mathbb{E}(T_n(y_1, y_2|x_0)) &\rightarrow f_X(x_0)R(y_1, y_2|x_0), \\ kh^d \text{Var}(T_n(y_1, y_2|x_0)) &\rightarrow \|K\|_2^2 f_X(x_0)R(y_1, y_2|x_0). \end{aligned}$$

The proof of these lemmas and all the subsequent ones are given in Section 6.5.

### 6.1 Proof of Theorem 2.1

To prove the result we will make use of empirical process theory with changing function classes, see for instance van der Vaart and Wellner (1996). To this aim we start by introducing some

notation. Let  $P$  be the distribution measure of  $(Y^{(1)}, Y^{(2)}, X)$ , and denote the expected value under  $P$ , the empirical version and empirical process as follows

$$Pf := \int f dP, \quad \mathbb{P}_n f := \frac{1}{n} \sum_{i=1}^n f(Y_i^{(1)}, Y_i^{(2)}, X_i), \quad \mathbb{G}_n f := \sqrt{n}(\mathbb{P}_n - P)f,$$

for any real-valued measurable function  $f : \mathbb{R}^2 \times \mathbb{R}^d \rightarrow \mathbb{R}$ . For a function class  $\mathcal{F}$ , let  $N_{[]}(\varepsilon, \mathcal{F}, L_2(P))$ , denote the minimal number of  $\varepsilon$ -brackets needed to cover  $\mathcal{F}$ . The bracketing integral is then defined as

$$J_{[]}(\delta, \mathcal{F}, L_2(P)) = \int_0^\delta \sqrt{\ln N_{[]}(\varepsilon, \mathcal{F}, L_2(P))} d\varepsilon.$$

We introduce our sequence of classes  $\mathcal{F}_n$  on  $\mathbb{R}^2 \times \mathbb{R}^d$  as

$$\mathcal{F}_n := \{(u, z) \rightarrow f_{n,y}(u, z), \quad y \in (0, T]^2\}$$

where

$$f_{n,y}(u, z) := \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq (k/n) y_1, \bar{F}_2(u_2|x_0) \leq (k/n) y_2\}}}{y_1^\eta}.$$

Denote also by  $F_n$  an envelope function of the class  $\mathcal{F}_n$ . Now, according to Theorem 19.28 in van der Vaart (1998) the weak convergence of the stochastic process (2.3) follows from the following four conditions. Let  $\rho_{x_0}$  be a semimetric, possibly depending on  $x_0$ , making  $(0, T]^2$  totally bounded. We have to prove that

$$\sup_{\rho_{x_0}(y, \bar{y}) \leq \delta_n} P(f_{n,y} - f_{n,\bar{y}})^2 \longrightarrow 0 \text{ for every } \delta_n \searrow 0, \quad (6.1)$$

$$PF_n^2 = O(1), \quad (6.2)$$

$$PF_n^2 \mathbb{1}_{\{F_n > \varepsilon \sqrt{n}\}} \longrightarrow 0 \text{ for every } \varepsilon > 0, \quad (6.3)$$

$$J_{[]}(\delta_n, \mathcal{F}_n, L_2(P)) \longrightarrow 0 \text{ for every } \delta_n \searrow 0, \quad (6.4)$$

along with the pointwise convergence of the covariance function.

**Proof of condition (6.1).** Let  $\rho_{x_0}(y, \bar{y}) := |y_1 - \bar{y}_1| + |y_2 - \bar{y}_2|$ . Denote  $A_{n,y} := \{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2\}$ . We have then

$$\begin{aligned} P(f_{n,y} - f_{n,\bar{y}})^2 &= \frac{nh^d}{k} \mathbb{E} \left[ K_h^2(x_0 - X) \left( \frac{\mathbb{1}_{A_{n,y}}}{y_1^\eta} - \frac{\mathbb{1}_{A_{n,\bar{y}}}}{\bar{y}_1^\eta} \right)^2 \right] \\ &= \frac{nh^d}{k} \mathbb{E} \left[ K_h^2(x_0 - X) \mathbb{E} \left[ \left( \frac{\mathbb{1}_{A_{n,y}}}{y_1^\eta} - \frac{\mathbb{1}_{A_{n,\bar{y}}}}{\bar{y}_1^\eta} \right)^2 \middle| X \right] \right]. \end{aligned} \quad (6.5)$$

We consider now three cases.

*Case 1:*  $y_1 \wedge \bar{y}_1 \leq \delta_n$ . Assume without loss of generality that  $y_1 \leq \bar{y}_1$ . By expanding the square in the above conditional expectation and using the fact that, e.g.,  $A_{n,y} \subset \{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1\}$ , we obtain the following inequality

$$\mathbb{E} \left[ \left( \frac{\mathbb{1}_{A_{n,y}}}{y_1^\eta} - \frac{\mathbb{1}_{A_{n,\bar{y}}}}{\bar{y}_1^\eta} \right)^2 \middle| X = x \right] \leq \frac{3P(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 | X = x)}{y_1^{2\eta}} + \frac{P(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) \bar{y}_1 | X = x)}{\bar{y}_1^{2\eta}},$$

which, after substituting in (6.5) leads to

$$\begin{aligned} & P(f_{n,y} - f_{n,\bar{y}})^2 \\ & \leq 3 \frac{n}{k} \int_{S_K} K^2(v) \frac{P(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 | X = x_0 - hv)}{y_1^{2\eta}} f_X(x_0 - hv) dv \\ & \quad + \frac{n}{k} \int_{S_K} K^2(v) \frac{P(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) \bar{y}_1 | X = x_0 - hv)}{\bar{y}_1^{2\eta}} f_X(x_0 - hv) dv. \end{aligned}$$

Now note that

$$P(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 | X = x_0 - hv) = \bar{F}_1(U_1(n/(ky_1)|x_0)|x_0 - hv),$$

which, together with the result of Lemma 6.1, motivates the following decomposition

$$\begin{aligned} & P(f_{n,y} - f_{n,\bar{y}})^2 \\ & \leq 3y_1^{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ & \quad + 3 \int_{S_K} K^2(v) \left[ \frac{1}{y_1^{2\eta}} \frac{n}{k} \bar{F}_1(U_1(n/(ky_1)|x_0)|x_0 - hv) - y_1^{1-2\eta} \right] f_X(x_0 - hv) dv \\ & \quad + \bar{y}_1^{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ & \quad + \int_{S_K} K^2(v) \left[ \frac{1}{\bar{y}_1^{2\eta}} \frac{n}{k} \bar{F}_1(U_1(n/(k\bar{y}_1)|x_0)|x_0 - hv) - \bar{y}_1^{1-2\eta} \right] f_X(x_0 - hv) dv. \end{aligned}$$

Using Lemma 6.1 and the fact that  $\rho_{x_0}(y, \bar{y}) \leq \delta_n$  which implies  $\bar{y}_1 \leq 2\delta_n$ , we get

$$P(f_{n,y} - f_{n,\bar{y}})^2 \leq 5\delta_n^{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv + o(1),$$

where the  $o(1)$  term does not depend on  $y_1$  and  $\bar{y}_1$ .

*Case 2:*  $y_1 \wedge \bar{y}_1 > \delta_n$  and  $y_2 \wedge \bar{y}_2 \leq \delta_n$ . Assume without loss of generality that  $y_2 \leq \bar{y}_2$ . Similarly to the approach followed in Case 1, we obtain

$$\mathbb{E} \left[ \left( \frac{\mathbb{1}_{A_{n,y}}}{y_1^\eta} - \frac{\mathbb{1}_{A_{n,\bar{y}}}}{\bar{y}_1^\eta} \right)^2 \middle| X = x \right] \leq \frac{3P(\bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 | X = x)}{(y_1 \wedge \bar{y}_1)^{2\eta}} + \frac{P(\bar{F}_2(Y^{(2)}|x_0) \leq (k/n) \bar{y}_2 | X = x)}{(y_1 \wedge \bar{y}_1)^{2\eta}},$$



and thus

$$\begin{aligned}
& P(f_{n,y} - f_{n,\bar{y}})^2 \\
& \leq \frac{3y_2}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\
& \quad + \frac{3y_2^{2\eta}}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) \left[ \frac{1}{y_2^{2\eta}} \frac{n}{k} \bar{F}_2(U_2(n/(ky_2)|x_0)|x_0 - hv) - y_2^{1-2\eta} \right] f_X(x_0 - hv) dv \\
& \quad + \frac{\bar{y}_2}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\
& \quad + \frac{\bar{y}_2^{2\eta}}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) \left[ \frac{1}{\bar{y}_2^{2\eta}} \frac{n}{k} \bar{F}_2(U_2(n/(k\bar{y}_2)|x_0)|x_0 - hv) - \bar{y}_2^{1-2\eta} \right] f_X(x_0 - hv) dv.
\end{aligned}$$

Again by Lemma 6.1 and using that  $\bar{y}_2 \leq 2\delta_n$  we have that

$$P(f_{n,y} - f_{n,\bar{y}})^2 \leq 5\delta_n^{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv + o(1),$$

where the  $o(1)$  term does not depend on  $y_2$  and  $\bar{y}_2$ .

*Case 3:*  $y_1 \wedge \bar{y}_1 > \delta_n$  and  $y_2 \wedge \bar{y}_2 > \delta_n$ . Let  $y \vee \bar{y}$  denote the vector with the component-wise maxima of  $y$  and  $\bar{y}$ , and similarly  $y \wedge \bar{y}$  is the vector with the component-wise minima of  $y$  and  $\bar{y}$ . Then

$$P(f_{n,y} - f_{n,\bar{y}})^2 \leq \frac{nh^d}{k} \mathbb{E} \left[ K_h^2(x_0 - X) \mathbb{E} \left[ \left( \frac{\mathbb{1}_{A_{n,y \vee \bar{y}}}}{(y_1 \wedge \bar{y}_1)^\eta} - \frac{\mathbb{1}_{A_{n,y \wedge \bar{y}}}}{(y_1 \vee \bar{y}_1)^\eta} \right)^2 \middle| X \right] \right].$$

Note that

$$\left( \frac{\mathbb{1}_{A_{n,y \vee \bar{y}}}}{(y_1 \wedge \bar{y}_1)^\eta} - \frac{\mathbb{1}_{A_{n,y \wedge \bar{y}}}}{(y_1 \vee \bar{y}_1)^\eta} \right)^2 = \left( \frac{1}{y_1^\eta} - \frac{1}{\bar{y}_1^\eta} \right)^2 \mathbb{1}_{A_{n,y \wedge \bar{y}}} + \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} (\mathbb{1}_{A_{n,y \vee \bar{y}}} - \mathbb{1}_{A_{n,y \wedge \bar{y}}}), \quad (6.6)$$

which leads to

$$\begin{aligned}
& P(f_{n,y} - f_{n,\bar{y}})^2 \\
& \leq \frac{(y_1^\eta - \bar{y}_1^\eta)^2}{(y_1 \bar{y}_1)^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) P\left(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 \wedge \bar{y}_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 \wedge \bar{y}_2 \middle| X = x_0 - hv\right) \\
& \quad \times f_X(x_0 - hv) dv \\
& \quad + \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) \left[ P\left(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 \vee \bar{y}_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 \vee \bar{y}_2 \middle| X = x_0 - hv\right) \right. \\
& \quad \left. - P\left(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1 \wedge \bar{y}_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 \wedge \bar{y}_2 \middle| X = x_0 - hv\right) \right] f_X(x_0 - hv) dv \\
& =: Q_{1,n} + Q_{2,n}.
\end{aligned}$$

As for  $Q_{1,n}$ , we easily obtain

$$Q_{1,n} \leq \frac{(y_1^\eta - \bar{y}_1^\eta)^2}{(y_1 \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) \frac{n}{k} \bar{F}_1(U_1(n/(k y_1 \wedge \bar{y}_1)|x_0)|x_0 - hv) f_X(x_0 - hv) dv.$$

Now, by the mean value theorem, applied to  $(y_1^\eta - \bar{y}_1^\eta)^2$ , and a decomposition motivated by Lemma 6.1,

$$\begin{aligned} Q_{1,n} &\leq (y_1 \wedge \bar{y}_1)^{-1-2\eta} (y_1 - \bar{y}_1)^2 \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ &\quad + (y_1 \wedge \bar{y}_1)^{-2} (y_1 - \bar{y}_1)^2 \int_{S_K} K^2(v) \left[ \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \frac{n}{k} \bar{F}_1 \left( U_1(n/(k y_1 \wedge \bar{y}_1) | x_0) \middle| x_0 - hv \right) \right. \\ &\quad \left. - (y_1 \wedge \bar{y}_1)^{1-2\eta} \right] \times f_X(x_0 - hv) dv. \end{aligned}$$

This then gives

$$Q_{1,n} \leq \delta_n^{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv + o(1),$$

where the  $o(1)$  term does not depend on  $y_1$  and  $\bar{y}_1$ .

Concerning  $Q_{2,n}$ , we have the following inequality

$$\begin{aligned} Q_{2,n} &\leq \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) P \left( (k/n) y_1 \wedge \bar{y}_1 \leq \bar{F}_1(Y^{(1)} | x_0) \leq (k/n) y_1 \vee \bar{y}_1 \middle| X = x_0 - hv \right) f_X(x_0 - hv) dv \\ &\quad + \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) P \left( (k/n) y_2 \wedge \bar{y}_2 \leq \bar{F}_2(Y^{(2)} | x_0) \leq (k/n) y_2 \vee \bar{y}_2 \middle| X = x_0 - hv \right) f_X(x_0 - hv) dv \\ &=: Q_{2,1,n} + Q_{2,2,n}. \end{aligned}$$

We only give details about  $Q_{2,1,n}$ , the term  $Q_{2,2,n}$  can be handled analogously. Direct computations give

$$Q_{2,1,n} = \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) \int_{U_1(n/(k(y_1 \wedge \bar{y}_1)) | x_0)}^{U_1(n/(k(y_1 \vee \bar{y}_1)) | x_0)} f_1(y | x_0 - hv) dy f_X(x_0 - hv) dv,$$

and, after substituting  $u = (n/k) \bar{F}_1(y | x_0)$ , we have

$$Q_{2,1,n} = \frac{1}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{S_K} K^2(v) \int_{y_1 \wedge \bar{y}_1}^{y_1 \vee \bar{y}_1} \frac{f_1(U_1(n/(ku) | x_0) | x_0 - hv)}{f_1(U_1(n/(ku) | x_0) | x_0)} du f_X(x_0 - hv) dv.$$

Using (2.1) and arguments similar to those used in the proof of Lemma 6.1 one obtains for  $n$  large and some small  $\kappa > 0$ ,

$$\frac{f_1(U_1(n/(ku) | x_0) | x_0 - hv)}{f_1(U_1(n/(ku) | x_0) | x_0)} \leq C u^{-\kappa},$$

where  $C$  does not depend on  $u$ . Then, for  $n$  large enough,

$$\begin{aligned} Q_{2,1,n} &\leq \frac{C}{(y_1 \wedge \bar{y}_1)^{2\eta}} \int_{y_1 \wedge \bar{y}_1}^{y_1 \vee \bar{y}_1} u^{-\kappa} du \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ &\leq \frac{C}{(y_1 \wedge \bar{y}_1)^{2\eta}} (y_1 \wedge \bar{y}_1)^{-\kappa} (y_1 \vee \bar{y}_1 - y_1 \wedge \bar{y}_1) \\ &\leq C \delta_n^{1-2\eta-\kappa} \\ &= o(1), \end{aligned}$$

for a small  $\kappa \in (0, 1 - 2\eta)$ .

Combining all the above we have verified (6.1).

**Proof of condition (6.2).** A natural envelope function of the class  $\mathcal{F}_n$  is

$$F_n(u, z) := \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq kT/n\}}}{[(n/k) \bar{F}_1(u_1|x_0)]^\eta}.$$

This yields

$$\begin{aligned} PF_n^2 &= \left(\frac{n}{k}\right)^{1-2\eta} h^d \mathbb{E} \left( K_h^2(x_0 - X) \mathbb{E} \left[ \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq kT/n\}}}{(\bar{F}_1(Y^{(1)}|x_0))^{2\eta}} \middle| X \right] \right) \\ &= \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \mathbb{E} \left[ \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq kT/n\}}}{(\bar{F}_1(Y^{(1)}|x_0))^{2\eta}} \middle| X = x_0 - hv \right] f_X(x_0 - hv) dv \\ &= \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{U_1(n/(kT)|x_0)}^\infty \frac{1}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \\ &\quad + \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{U_1(n/(kT)|x_0)}^\infty \frac{1}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} \\ &\quad \quad \times \left\{ \left( \frac{\bar{F}_1(y|x_0 - hv)}{\bar{F}_1(y|x_0)} \right)^{2\eta} - 1 \right\} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \\ &=: Q_{3,n}(T) + Q_{4,n}(T). \end{aligned}$$

Concerning  $Q_{3,n}(T)$  we obtain by direct integration and a slight adjustment of Lemma 6.1, for large  $n$

$$\begin{aligned} Q_{3,n}(T) &= \frac{1}{1-2\eta} \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) [\bar{F}_1(U_1(n/(kT)|x_0)|x_0 - hv)]^{1-2\eta} f_X(x_0 - hv) dv \\ &= \frac{T^{1-2\eta}}{1-2\eta} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ &\quad + \frac{1}{1-2\eta} \int_{S_K} K^2(v) \left[ \left( \frac{n}{k} \bar{F}_1(U_1(n/(kT)|x_0)|x_0 - hv) \right)^{1-2\eta} - T^{1-2\eta} \right] f_X(x_0 - hv) dv \\ &\leq CT^{1-2\eta-\kappa}, \end{aligned} \tag{6.7}$$

for  $\kappa < 1 - 2\eta$ .

Concerning  $Q_{4,n}(T)$ , combining  $(\mathcal{D})$  with  $(\mathcal{H})$  gives the following bound, for  $n$  large and  $y \geq U_1(n/(kT)|x_0)$ ,

$$\begin{aligned} \left| \left( \frac{\bar{F}_1(y|x_0 - hv)}{\bar{F}_1(y|x_0)} \right)^{2\eta} - 1 \right| &\leq C_1 \left( h^{\eta A_1} + y^{C_2 h^{\eta \gamma_1}} h^{\eta \gamma_1} \ln y + |\delta_1(y|x_0)| h^{\eta B_1} \right. \\ &\quad \left. + |\delta_1(y|x_0)| y^{C_3 h^{\eta \varepsilon_1}} h^{\eta \varepsilon_1} \ln y \right). \end{aligned} \tag{6.8}$$

Each of the terms in the right-hand side of the above inequality needs now to be used in  $Q_{4,n}(T)$ , leading to the terms  $Q_{4,j,n}(T)$ ,  $j = 1, \dots, 4$ , studied below. First

$$Q_{4,1,n}(T) :=$$

$$h^{\eta_{A_1}} \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{U_1(n/(kT)|x_0)}^{\infty} \frac{1}{(\overline{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv.$$

This term is clearly of smaller order than  $Q_{3,n}(T)$  studied above and hence  $Q_{4,1,n}(T) = O(1)$ . For the second term in the right-hand side of (6.8) we need to study

$$Q_{4,2,n}(T) :=$$

$$h^{\eta_{\gamma_1}} \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(T)}^{\infty} y^{\xi_{1,n}} \ln y \frac{1}{(\overline{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv$$

where  $t_n(T) := U_1(n/(kT)|x_0)$  and  $\xi_{1,n} := C_2 h^{\eta_{\gamma_1}}$ . Let  $p_n(y) := \xi_{1,n} y^{\xi_{1,n}-1} \ln y + y^{\xi_{1,n}-1}$ . Applying integration by parts on the inner integral gives, for  $n$  large enough,

$$\begin{aligned} Q_{4,2,n}(T) &= \left(\frac{n}{k}\right)^{1-2\eta} \frac{h^{\eta_{\gamma_1}} \ln(t_n(T)) [t_n(T)]^{\xi_{1,n}}}{1-2\eta} \int_{S_K} K^2(v) [\overline{F}_1(t_n(T)|x_0 - hv)]^{1-2\eta} f_X(x_0 - hv) dv \\ &\quad + \left(\frac{n}{k}\right)^{1-2\eta} \frac{h^{\eta_{\gamma_1}}}{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(T)}^{\infty} p_n(y) [\overline{F}_1(y|x_0 - hv)]^{1-2\eta} dy f_X(x_0 - hv) dv \\ &=: Q_{4,2,1,n}(T) + Q_{4,2,2,n}(T). \end{aligned}$$

We obtain, for  $n$  large enough

$$\begin{aligned} Q_{4,2,1,n}(T) &\leq Ch^{\eta_{\gamma_1}} \ln(t_n(T)) [t_n(T)]^{\xi_{1,n}} T^{1-2\eta-\kappa} \\ &= O(1), \end{aligned}$$

since for distributions satisfying  $(\mathcal{D})$  one has that

$$U_1(y|x_0) = (A_1(x_0))^{\gamma_1(x_0)} y^{\gamma_1(x_0)} (1 + a_1(y|x_0)) \quad (6.9)$$

where  $|a_1(\cdot|x_0)|$  is regularly varying with index equal to  $-\gamma_1(x_0)\beta_1(x_0)$ , and by using the fact that  $h^{\eta_{\gamma_1}} \ln(n/k) \rightarrow 0$  as  $n \rightarrow \infty$ .

Now consider  $Q_{4,2,2,n}(T)$ . We have

$$Q_{4,2,2,n}(T) = \frac{h^{\eta_{\gamma_1}} T^{1-2\eta}}{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(T)}^{\infty} p_n(y) \left(\frac{\overline{F}_1(y|x_0 - hv)}{\overline{F}_1(y|x_0)}\right)^{1-2\eta} \left(\frac{\overline{F}_1(y|x_0)}{\overline{F}_1(t_n(T)|x_0)}\right)^{1-2\eta} dy \times f_X(x_0 - hv) dv.$$

For  $n$  large and  $y \geq t_n(T)$ , with  $\xi_{2,n} = Ch^{\eta_{\varepsilon_1}}$ ,

$$\left(\frac{\overline{F}_1(y|x_0 - hv)}{\overline{F}_1(y|x_0)}\right)^{1-2\eta} \leq Cy^{\xi_{1,n}} \left(1 + y^{\xi_{2,n}} h^{\eta_{\varepsilon_1}} \ln y\right).$$

Substituting  $u = y/t_n(T)$  we get

$$\begin{aligned} Q_{4,2,2,n}(T) &\leq Ch^{\eta_{\gamma_1}} T^{1-2\eta} [t_n(T)]^{1+\xi_{1,n}} \\ &\quad \times \int_{S_K} K^2(v) \int_1^{\infty} p_n(t_n(T)u) u^{\xi_{1,n}} (1 + (t_n(T)u)^{\xi_{2,n}} h^{\eta_{\varepsilon_1}} \ln(t_n(T)u)) \left(\frac{\overline{F}_1(t_n(T)u|x_0)}{\overline{F}_1(t_n(T)|x_0)}\right)^{1-2\eta} du \\ &\quad \times f_X(x_0 - hv) dv. \end{aligned}$$

Since  $\bar{F}_1(\cdot|x_0)$  is regularly varying, we can apply the Potter bound (see, e.g., de Haan and Ferreira 2006, Proposition B.1.9), and obtain, for  $n$  large enough and  $0 < \delta < 1/\gamma_1(x_0)$

$$\begin{aligned} Q_{4,2,2,n}(T) &\leq Ch^{\eta_{\gamma_1}} T^{1-2\eta} [t_n(T)]^{2\xi_{1,n}} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\ &\quad \times \int_1^\infty (\xi_{1,n} u^{\xi_{1,n}-1} \ln(t_n(T)) + \xi_{1,n} u^{\xi_{1,n}-1} \ln u + u^{\xi_{1,n}-1}) u^{\xi_{1,n} - (1/\gamma_1(x_0) - \delta)(1-2\eta)} \\ &\quad \times (1 + (t_n(T)u)^{\xi_{2,n}} h^{\eta_{\varepsilon_1}} \ln(t_n(T)u)) du. \end{aligned}$$

After tedious computations one gets

$$\begin{aligned} Q_{4,2,2,n}(T) &\leq CT^{1-2\eta} h^{\eta_{\gamma_1}} [t_n(T)]^{2\xi_{1,n}} \left\{ 1 + h^{\eta_{\gamma_1}} \ln(t_n(T)) + [t_n(T)]^{\xi_{2,n}} h^{\eta_{\varepsilon_1}} \ln(t_n(T)) \right\} \\ &= O(1), \end{aligned}$$

by (6.9) and the fact that  $h^{\eta_{\gamma_1} \wedge \eta_{\varepsilon_1}} \ln(n/k) \rightarrow 0$  as  $n \rightarrow \infty$ . Hence,  $Q_{4,2,n}(T) = O(1)$ .

Finally, the two last terms  $Q_{4,3,n}(T)$  and  $Q_{4,4,n}(T)$  can be dealt with similarly as the two previous ones since

$$\begin{aligned} Q_{4,3,n}(T) &:= h^{\eta_{B_1}} \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(T)}^\infty \frac{|\delta_1(y|x_0)|}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \\ &\leq \left( \sup_{y \geq t_n(T)} |\delta_1(y|x_0)| \right) h^{\eta_{B_1}} \left(\frac{n}{k}\right)^{1-2\eta} \\ &\quad \times \int_{S_K} K^2(v) \int_{t_n(T)}^\infty \frac{1}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \end{aligned} \quad (6.10)$$

and

$$\begin{aligned} Q_{4,4,n}(T) &:= h^{\eta_{\varepsilon_1}} \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(T)}^\infty \frac{|\delta_1(y|x_0)| y^{\xi_{2,n}} \ln y}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \\ &\leq \left( \sup_{y \geq t_n(T)} |\delta_1(y|x_0)| \right) h^{\eta_{\varepsilon_1}} \left(\frac{n}{k}\right)^{1-2\eta} \\ &\quad \times \int_{S_K} K^2(v) \int_{t_n(T)}^\infty \frac{y^{\xi_{2,n}} \ln y}{(\bar{F}_1(y|x_0 - hv))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv. \end{aligned} \quad (6.11)$$

This yields  $Q_{4,3,n}(T) = O(1)$  and  $Q_{4,4,n}(T) = O(1)$ . Combining all these results, we get (6.2).

**Proof of condition (6.3).** To this aim, for any  $\alpha \in (0, 1/\eta - 2)$ , we have

$$\begin{aligned}
PF_n^2 \mathbb{1}_{\{F_n > \varepsilon\sqrt{n}\}} &\leq \frac{1}{\varepsilon^\alpha n^{\alpha/2}} PF_n^{2+\alpha} \\
&= \frac{1}{\varepsilon^\alpha n^{\alpha/2}} \left(\frac{nh^d}{k}\right)^{1+\frac{\alpha}{2}} \mathbb{E} \left( K_h^{2+\alpha}(x_0 - X) \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq kT/n\}}}{[(n/k) \bar{F}_1(Y^{(1)}|x_0)]^{\eta(2+\alpha)}} \right) \\
&= \frac{1}{\varepsilon^\alpha} \frac{1}{(kh^d)^{\alpha/2}} \left(\frac{n}{k}\right)^{1-\eta(2+\alpha)} \int_{S_K} K^{2+\alpha}(v) \mathbb{E} \left( \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq kT/n\}}}{[\bar{F}_1(Y^{(1)}|x_0)]^{\eta(2+\alpha)}} \middle| X = x_0 - hv \right) \\
&\quad \times f_X(x_0 - hv) dv \\
&= \frac{1}{\varepsilon^\alpha} \frac{1}{(kh^d)^{\alpha/2}} \left(\frac{n}{k}\right)^{1-\eta(2+\alpha)} \\
&\quad \times \left\{ \int_{S_K} K^{2+\alpha}(v) \int_{t_n(T)}^\infty \frac{1}{(\bar{F}_1(y|x_0 - hv))^{\eta(2+\alpha)}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \right. \\
&\quad \left. + \int_{S_K} K^{2+\alpha}(v) \int_{t_n(T)}^\infty \frac{1}{(\bar{F}_1(y|x_0 - hv))^{\eta(2+\alpha)}} \right. \\
&\quad \left. \times \left[ \left( \frac{\bar{F}_1(y|x_0 - hv)}{\bar{F}_1(y|x_0)} \right)^{\eta(2+\alpha)} - 1 \right] dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \right\}.
\end{aligned}$$

The terms into brackets can be studied similarly as  $Q_{j,n}(T)$ ,  $j = 3, 4$ , and thus (6.3) is established since  $kh^d \rightarrow \infty$ .

**Proof of condition (6.4).** Without loss of generality assume  $T = 1$  and consider, for  $a, \theta, \tilde{\theta} < 1$ , the classes

$$\begin{aligned}
\mathcal{F}_n^{(1)}(a) &:= \{f_{n,y} \in \mathcal{F}_n : y_1 \leq a\}, \\
\mathcal{F}_n^{(2)}(a) &:= \{f_{n,y} \in \mathcal{F}_n : y_1 > a, y_2 \leq a\}, \\
\mathcal{F}_n(\ell, m) &:= \{f_{n,y} \in \mathcal{F}_n : \theta^{\ell+1} \leq y_1 \leq \theta^\ell, \tilde{\theta}^{m+1} \leq y_2 \leq \tilde{\theta}^m\},
\end{aligned}$$

where  $\ell = 0, \dots, \lfloor \ln a / \ln \theta \rfloor$  and  $m = 0, \dots, \lfloor \ln a / \ln \tilde{\theta} \rfloor$ . We start by showing that  $\mathcal{F}_n^{(1)}(a)$  is an  $\varepsilon$ -bracket, for  $n$  sufficiently large. Clearly

$$\begin{aligned}
0 \leq f_{n,y}(u, z) &\leq \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq (k/n) y_1\}}}{[(n/k) \bar{F}_1(u_1|x_0)]^\eta} \\
&\leq h^{d/2} (n/k)^{1/2-\eta} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq (k/n) a\}}}{(\bar{F}_1(u_1|x_0))^\eta} := u_{1,n}(u, z).
\end{aligned}$$

Then

$$\begin{aligned}
Pu_{1,n}^2 &= \left(\frac{n}{k}\right)^{1-2\eta} \int_{S_K} K^2(v) \int_{t_n(a)}^\infty \frac{1}{(\bar{F}_1(y|x_0))^{2\eta}} dF_1(y|x_0 - hv) f_X(x_0 - hv) dv \\
&= Q_{3,n}(a) + Q_{4,n}(a),
\end{aligned}$$

using the same decomposition as for  $PF_n^2$ . Thus, one can obtain the result from the above analysis of  $Q_{3,n}(T)$  and  $Q_{4,n}(T)$ , taking into account that the various constants involved in these will not depend on  $a$ .

Concerning  $Q_{3,n}(a)$ , according to (6.7), for  $n$  large

$$Q_{3,n}(a) \leq Ca^{1-2\eta-\kappa},$$

where  $C$  does not depend on  $a$ . Now, taking  $a = \varepsilon^{3/(1-2\eta)}$ , for  $n$  large enough and  $\varepsilon$  small we have  $|Q_{3,n}(a)| \leq \varepsilon^2$ .

Concerning  $Q_{4,n}(a)$ , we use the same decomposition as for  $Q_{4,n}(T)$  based on (6.8), which entails that, for  $n$  large enough,  $\varepsilon$  small and some small  $\zeta > 0$

$$\begin{aligned} Q_{4,1,n}(a) &\leq \varepsilon^2, \\ Q_{4,2,1,n}(a) &\leq Ch^{\eta\gamma_1} \ln(t_n(a)) [t_n(a)]^{\xi_{1,n}} a^{1-2\eta-\kappa} \\ &\leq C(1 + |\ln a|) a^{-\zeta} a^{1-2\eta-\kappa} \\ &\leq Ca^{1-2\eta-2\kappa}, \end{aligned}$$

with  $C$  a constant not depending on  $a$ , since from (6.9) and for  $n$  large,

$$h^{\eta\gamma_1} \ln t_n(a) \leq C(1 + |\ln a|).$$

Also, for  $n$  large, and some small  $\zeta > 0$

$$\begin{aligned} Q_{4,2,2,n}(a) &\leq Ca^{1-2\eta} h^{\eta\gamma_1} [t_n(a)]^{2\xi_{1,n}} \left\{ 1 + h^{\eta\gamma_1} \ln(t_n(a)) + [t_n(a)]^{\xi_{2,n}} h^{\eta\varepsilon_1} \ln(t_n(a)) \right\} \\ &\leq Ca^{1-2\eta} h^{\eta\gamma_1} a^{-\zeta} (1 + |\ln a| + a^{-\zeta} (1 + |\ln a|)) \\ &\leq Ca^{1-2\eta-\kappa}, \end{aligned}$$

where  $C$  does not depend on  $a$ . Hence, for  $n$  large and  $\varepsilon$  small we obtain  $Q_{4,2,2,n}(a) \leq \varepsilon^2$ . Using (6.10) and (6.11), we have also  $Q_{4,3,n}(a) \leq \varepsilon^2$  and  $Q_{4,4,n}(a) \leq \varepsilon^2$ . Combining all the terms we get  $Pu_{1,n}^2 \leq \varepsilon^2$  for  $n$  large.

Next consider  $\mathcal{F}_n^{(2)}(a)$ . Then

$$0 \leq f_{n,y}(u, z) \leq \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_2(u_2|x_0) \leq (k/n) a\}}}{a^\eta} =: u_{2,n}(u, z),$$

and

$$\begin{aligned} Pu_{2,n}^2 &= \frac{1}{a^{2\eta}} \frac{n}{k} \int_{S_K} K^2(v) \bar{F}_2 \left( U_2 \left( \frac{n}{ka} \middle| x_0 \right) \middle| x_0 - hv \right) f_X(x_0 - hv) dv \\ &\leq \varepsilon^2, \end{aligned}$$

when  $n$  is large enough and for  $\varepsilon$  small.

Finally, we consider  $\mathcal{F}_n(\ell, m)$ . We obtain the following bounds

$$\begin{aligned} \underline{u}_n(u, z) &:= \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq (k/n) \theta^{\ell+1}, \bar{F}_2(u_2|x_0) \leq (k/n) \tilde{\theta}^{m+1}\}}}{\theta^{\ell\eta}} \leq f_{n,y}(u, z) \leq \\ &\sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \frac{\mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq (k/n) \theta^\ell, \bar{F}_2(u_2|x_0) \leq (k/n) \tilde{\theta}^m\}}}{\theta^{(\ell+1)\eta}} =: \bar{u}_n(u, z). \end{aligned}$$

Then

$$P(\bar{u}_n - \underline{u}_n)^2 = \frac{nh^d}{k} \mathbb{E} \left[ K_h^2(x_0 - X) \left( \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) \theta^\ell, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) \tilde{\theta}^m\}}}{\theta^{(\ell+1)\eta}} - \frac{\mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) \theta^{\ell+1}, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) \tilde{\theta}^{m+1}\}}}{\theta^{\ell\eta}} \right)^2 \right].$$

The difference of the indicator functions can be decomposed as in (6.6), and subsequent calculations follow arguments similar to those used in the verification of (6.1), Case 3. Taking  $\theta = 1 - \varepsilon^3$  and  $\tilde{\theta} = 1 - a$ , gives for  $n$  large enough and  $\varepsilon$  small that  $P(\bar{u}_n - \underline{u}_n)^2 \leq \varepsilon^2$ .

Combining the above, for  $n$  large and  $\varepsilon$  small one obtains that the cover number by bracketing is of the order  $\varepsilon^{-4-3/(1-2\eta)}$ , and hence (6.4) is satisfied.

To conclude the proof, we comment on the pointwise convergence of the covariance function, which is given by  $Pf_{n,y}f_{n,\bar{y}} - Pf_{n,y}Pf_{n,\bar{y}}$ . We have

$$\begin{aligned} Pf_{n,y}f_{n,\bar{y}} &= \frac{\|K\|_2^2}{(y_1\bar{y}_1)^\eta} \frac{n}{k} \mathbb{E} \left[ \frac{1}{\|K\|_2^2 h^d} K^2 \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{A_{n,y \wedge \bar{y}}} \right] \\ &\rightarrow \|K\|_2^2 f_X(x_0) \frac{R(y_1 \wedge \bar{y}_1, y_2 \wedge \bar{y}_2 | x_0)}{(y_1\bar{y}_1)^\eta}, \end{aligned}$$

as  $n \rightarrow \infty$ , by the arguments used in the proof of Lemma 6.2. Also

$$\begin{aligned} Pf_{n,y} &= \sqrt{\frac{kh^d}{n}} \frac{1}{y_1^\eta} \frac{n}{k} \mathbb{E} [K_h(x_0 - X) \mathbb{1}_{A_{n,y}}] \\ &\rightarrow 0, \end{aligned}$$

as  $n \rightarrow \infty$ . ■

## 6.2 Proof of Theorem 2

Recall that

$$T_n(\infty, y_2 | x_0) = \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{\bar{F}_2(Y_i^{(2)} | x_0) \leq \frac{k}{n} y_2\}}.$$

We follow the lines of proof of Theorem 1. We introduce the sequence of classes  $\tilde{\mathcal{F}}_n$  on  $\mathbb{R} \times \mathbb{R}^d$  as

$$\tilde{\mathcal{F}}_n := \{(u, z) \rightarrow \tilde{f}_{n,y}(u, z), y \in (0, T]\}$$

where

$$\tilde{f}_{n,y}(u, z) := \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \mathbb{1}_{\{\bar{F}_2(u | x_0) \leq \frac{k}{n} y\}}.$$



We have to verify the conditions (6.1)-(6.4) in the proof of Theorem 1 for the new functions  $\tilde{f}_{n,y}$ , and with  $\rho_{x_0}(y, \bar{y}) := |y - \bar{y}|$ . Without loss of generality, we may assume that  $y > \bar{y}$ . Thus, we have

$$\begin{aligned}
P \left( \tilde{f}_{n,y} - \tilde{f}_{n,\bar{y}} \right)^2 &= \frac{nh^d}{k} \mathbb{E} \left[ K_h^2(x_0 - X) \left( \mathbb{1}_{\{\bar{F}_2(Y^{(2)}|x_0) \leq \frac{k}{n}y\}} - \mathbb{1}_{\{\bar{F}_2(Y^{(2)}|x_0) \leq \frac{k}{n}\bar{y}\}} \right) \right] \\
&= \frac{n}{k} \int_{S_K} K^2(v) \left[ \bar{F}_2 \left( U_2 \left( \frac{n}{ky} |x_0 \right) \middle| x_0 - hv \right) - \bar{F}_2 \left( U_2 \left( \frac{n}{k\bar{y}} |x_0 \right) \middle| x_0 - hv \right) \right] \\
&\quad \times f_X(x_0 - hv) dv \\
&= (y - \bar{y}) \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\
&\quad + \int_{S_K} K^2(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{ky} |x_0 \right) \middle| x_0 - hv \right) - y \right] f_X(x_0 - hv) dv \\
&\quad - \int_{S_K} K^2(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k\bar{y}} |x_0 \right) \middle| x_0 - hv \right) - \bar{y} \right] f_X(x_0 - hv) dv \\
&\leq \delta_n \int_{S_K} K^2(v) f_X(x_0 - hv) dv + o(1),
\end{aligned}$$

with a  $o(1)$ -term which is uniform in  $y$  and  $\bar{y}$  by Lemma 1. This yields (6.1).

Now, concerning (6.2) we can use the following envelope function of the class  $\tilde{\mathcal{F}}_n$

$$\tilde{F}_n(u, z) := \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \mathbb{1}_{\{\bar{F}_2(u|x_0) \leq \frac{k}{n}T\}}$$

from which we deduce that

$$P\tilde{F}_n^2 = \frac{n}{k} \int_{S_K} K^2(v) \bar{F}_2 \left( U_2 \left( \frac{n}{kT} |x_0 \right) \middle| x_0 - hv \right) f_X(x_0 - hv) dv = O(1).$$

Next condition (6.3) is also a direct consequence of the definition of the envelope since

$$\begin{aligned}
P\tilde{F}_n^2 \mathbb{1}_{\{\tilde{F}_n > \varepsilon\sqrt{n}\}} &\leq \frac{1}{\varepsilon^\alpha n^{\alpha/2}} P\tilde{F}_n^{2+\alpha} \\
&\leq \frac{1}{\varepsilon^\alpha (kh^d)^{\alpha/2}} \frac{n}{k} \int_{S_K} K^{2+\alpha}(v) \bar{F}_2 \left( U_2 \left( \frac{n}{kT} |x_0 \right) \middle| x_0 - hv \right) f_X(x_0 - hv) dv = o(1)
\end{aligned}$$

as soon as  $kh^d \rightarrow \infty$ .

Finally, concerning (6.4), again without loss of generality we assume  $T = 1$  and divide  $[0, 1]$  into  $m$  intervals of length  $1/m$ . Then, for  $y \in [(i-1)/m, i/m]$  we have the bounds

$$\underline{u}_n(u, z) := \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \mathbb{1}_{\{\bar{F}_2(u|x_0) \leq \frac{k}{n} \frac{i-1}{m}\}} \leq \tilde{f}_{n,y}(u, z) \leq \sqrt{\frac{nh^d}{k}} K_h(x_0 - z) \mathbb{1}_{\{\bar{F}_2(u|x_0) \leq \frac{k}{n} \frac{i}{m}\}} =: \bar{u}_n(u, z)$$

from which we deduce that

$$\begin{aligned}
P(\underline{u}_n - \bar{u}_n)^2 &= \frac{1}{m} \int_{S_K} K^2(v) f_X(x_0 - hv) dv \\
&+ \int_{S_K} K^2(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k} \frac{m}{i} \middle| x_0 \right) \middle| x_0 - hv \right) - \frac{i}{m} \right] f_X(x_0 - hv) dv \\
&- \int_{S_K} K^2(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k} \frac{m}{i-1} \middle| x_0 \right) \middle| x_0 - hv \right) - \frac{i-1}{m} \right] f_X(x_0 - hv) dv \\
&\leq \varepsilon^3 \int_{S_K} K^2(v) f_X(x_0 - hv) dv + 2\varepsilon^3
\end{aligned}$$

when  $m = \lceil \frac{1}{\varepsilon^3} \rceil$ . If  $\varepsilon$  is small and  $n$  large, then  $P(\underline{u}_n - \bar{u}_n)^2 \leq \varepsilon^2$ .

The pointwise convergence of the covariance function can be verified with arguments similar to those used in the proof of Theorem 1.

Consequently

$$\sqrt{kh^d} [T_n(\infty, y_2 | x_0) - \mathbb{E}(T_n(\infty, y_2 | x_0))] \rightsquigarrow W(\infty, y_2),$$

in  $D((0, T])$ .

Now, remark that

$$\begin{aligned}
\mathbb{E}(T_n(\infty, y_2 | x_0)) &= y_2 f_X(x_0) + O(h^{\eta_{f_X}}) \\
&+ f_X(x_0) \int_{S_K} K(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right) - y_2 \right] dv \\
&+ \int_{S_K} K(v) \left[ \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right) - y_2 \right] [f_X(x_0 - hv) - f_X(x_0)] dv.
\end{aligned}$$

Following the lines of proof of Lemma 1, we deduce that

$$\left| \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right) - y_2 \right| \leq C \left\{ h^{\eta_{A_2}} + h^{\eta_{\gamma_2}} \ln \frac{n}{k} + |\delta_2(U_2(n/k | x_0) | x_0)| \left( h^{\eta_{B_2}} + h^{\eta_{\varepsilon_2}} \ln \frac{n}{k} \right) \right\}$$

from which we obtain

$$\begin{aligned}
\mathbb{E}(T_n(\infty, y_2 | x_0)) &= y_2 f_X(x_0) + O(h^{\eta_{f_X} \wedge \eta_{A_2}}) + O\left(h^{\eta_{\gamma_2}} \ln \frac{n}{k}\right) + O(|\delta_2(U_2(n/k | x_0) | x_0)| h^{\eta_{B_2}}) \\
&+ O\left(|\delta_2(U_2(n/k | x_0) | x_0)| h^{\eta_{\varepsilon_2}} \ln \frac{n}{k}\right)
\end{aligned}$$

with  $O$ -terms which are uniform in  $y_2 \in (0, T]$ . This implies that, under the assumptions of Theorem 2, we have

$$\sqrt{kh^d} [T_n(\infty, y_2 | x_0) - y_2 f_X(x_0)] \rightsquigarrow W(\infty, y_2), \quad (6.12)$$

in  $D((0, T])$ .

Finally,

$$\sqrt{kh^d} \left( \frac{T_n(\infty, y_2|x_0)}{\hat{f}_n(x_0)} - y_2 \right) = \sqrt{kh^d} \left( \frac{T_n(\infty, y_2|x_0)}{f_X(x_0)} - y_2 \right) - \frac{T_n(\infty, y_2|x_0)}{\hat{f}_n(x_0)f_X(x_0)} \sqrt{\frac{k}{n}} \sqrt{nh^d} \left( \hat{f}_n(x_0) - f_X(x_0) \right),$$

from which Theorem 2 follows.  $\blacksquare$

In the sequel, for convenient representation, all the limiting processes in Theorems 2.1 and 2.2 will be defined on the same probability space, via the Skorohod construction, but it should be kept in mind that they are only in distribution equal to the original processes. The Skorohod representation theorem gives then (with keeping the same notations)

$$\sup_{y_1, y_2 \in (0, T]} \left| \frac{\sqrt{kh^d} [T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))] - W(y_1, y_2)}{y_1^\eta} \right| \rightarrow 0, \text{ a.s.}$$

and

$$\sup_{y_2 \in (0, T]} \left| \sqrt{kh^d} \left( \frac{T_n(\infty, y_2|x_0)}{\hat{f}_n(x_0)} - y_2 \right) - \frac{W(\infty, y_2)}{f_X(x_0)} \right| \rightarrow 0, \text{ a.s.}$$

### 6.3 Convergence result for an auxiliary statistic

In this section we give a convergence result for an auxiliary statistic. In particular, we generalize  $\tilde{\theta}_n$  to  $\tilde{\theta}_n(y_2)$ , defined as

$$\tilde{\theta}_n(y_2) = \frac{1}{k} \sum_{i=1}^n K_h(x_0 - X_i) Y_i^{(1)} \mathbb{1}_{\{Y_i^{(2)} \geq U_2(n/(ky_2)|x_0)\}}.$$

Assuming  $F_1(y|x_0)$  strictly increasing in  $y$ , we have

$$\tilde{\theta}_n(y_2) = -U_1 \left( \frac{n}{k} | x_0 \right) \int_0^\infty T_n(s_n(u), y_2|x_0) du^{-\gamma_1(x_0)}.$$

As motivation for studying  $\tilde{\theta}_n(y_2)$ , note that  $\hat{\theta}_n = \tilde{\theta}_n(\hat{e}_n)$ , where  $\hat{e}_n := \frac{n}{k} \bar{F}_2(\hat{u}_n U_2(\frac{n}{k}|x_0)|x_0)$  with  $\hat{u}_n := \hat{U}_2(\frac{n}{k}|x_0)/U_2(\frac{n}{k}|x_0)$ . To estimate  $U_2(\cdot|x_0)$  we will use  $\hat{U}_2(\cdot|x_0) := \inf\{y : \hat{F}_{n,2}(y|x_0) \geq 1 - 1/\cdot\}$  with

$$\hat{F}_{n,2}(y|x_0) := \frac{\sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} \leq y\}}}{\sum_{i=1}^n K_h(x_0 - X_i)},$$

the empirical kernel estimator of the unknown conditional distribution function of  $Y^{(2)}$  given  $X = x_0$ . See for instance Daouia et al. (2011). The asymptotic behavior of the quantile estimator is given in Lemma 6.6.

**Proposition 6.1.** *Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$ ,  $(\mathcal{R})$  with  $x \rightarrow R(y_1, y_2|x)$  being a continuous function,  $x_0 \in \text{Int}(S_X)$  with  $f_X(x_0) > 0$ , and  $y \rightarrow F_j(y|x_0)$ ,  $j = 1, 2$ , are strictly increasing. Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$ ,  $kh^d \rightarrow \infty$  and  $h^{\eta_{\gamma_1} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_1} \wedge \eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ . Then, for  $\gamma_1(x_0) < 1/2$ , we have*

$$\sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \frac{\sqrt{kh^d}}{U_1(n/k|x_0)} \left[ \tilde{\theta}_n(y_2) - \mathbb{E}(\tilde{\theta}_n(y_2)) \right] + \int_0^\infty W(u, y_2) du^{-\gamma_1(x_0)} \right| \xrightarrow{\mathbb{P}} 0.$$

## Proof of Proposition 6.1

We use the decomposition

$$\sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \frac{\sqrt{kh^d}}{U_1(n/k|x_0)} \left[ \tilde{\theta}_n(y_2) - \mathbb{E}(\tilde{\theta}_n(y_2)) \right] + \int_0^\infty W(u, y_2) du^{-\gamma_1(x_0)} \right| \leq I_1(T) + \sum_{i=2}^4 I_{i,n}(T),$$

where

$$\begin{aligned} I_1(T) &:= \sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \int_T^\infty W(u, y_2) du^{-\gamma_1(x_0)} \right|, \\ I_{2,n}(T) &:= \sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \int_T^\infty \sqrt{kh^d} [T_n(s_n(u), y_2|x_0) - \mathbb{E}(T_n(s_n(u), y_2|x_0))] du^{-\gamma_1(x_0)} \right|, \\ I_{3,n}(T) &:= \sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \int_0^T \left\{ \sqrt{kh^d} [T_n(s_n(u), y_2|x_0) - \mathbb{E}(T_n(s_n(u), y_2|x_0))] - W(s_n(u), y_2) \right\} du^{-\gamma_1(x_0)} \right|, \\ I_{4,n}(T) &:= \sup_{\frac{1}{2} \leq y_2 \leq 2} \left| \int_0^T [W(s_n(u), y_2) - W(u, y_2)] du^{-\gamma_1(x_0)} \right|. \end{aligned}$$

Similarly to the proof of Proposition 2 in Cai et al. (2015), it is sufficient to show that for any  $\varepsilon > 0$ , there exists  $T_0 = T_0(\varepsilon)$  such that

$$\mathbb{P}(I_1(T_0) > \varepsilon) < \varepsilon, \tag{6.13}$$

and  $n_0 = n_0(T_0)$  such that, for any  $n > n_0$

$$\mathbb{P}(I_{j,n}(T_0) > \varepsilon) < \varepsilon, \text{ for } j = 2, 3 \text{ and } 4.$$

Clearly

$$I_1(T) \leq \sup_{u \geq T, \frac{1}{2} \leq y_2 \leq 2} |W(u, y_2)| T^{-\gamma_1(x_0)}.$$

Since a rescaled version of our Gaussian process  $W(\cdot, \cdot)$  gives the one in Cai et al. (2015), according to their Lemma 2, we have  $\sup_{0 < u < \infty, \frac{1}{2} \leq y_2 \leq 2} |W(u, y_2)| < \infty$  with probability one. This implies that there exists  $T_1 = T_1(\varepsilon)$  such that

$$\mathbb{P} \left( \sup_{0 < u < \infty, \frac{1}{2} \leq y_2 \leq 2} |W(u, y_2)| > T_1^{\gamma_1(x_0)} \varepsilon \right) < \varepsilon,$$

from which we deduce that, for any  $T > T_1$

$$\mathbb{P}(I_1(T) > \varepsilon) \leq \mathbb{P} \left( \sup_{0 < u < \infty, \frac{1}{2} \leq y_2 \leq 2} |W(u, y_2)| > T_1^{\gamma_1(x_0)} \varepsilon \right) < \varepsilon.$$

Consequently (6.13) holds for  $T_0 > T_1$ .

We continue with the term  $I_{2,n}(T)$ . We have

$\mathbb{P}(I_{2,n}(T) > \varepsilon)$

$$\begin{aligned}
&\leq \mathbb{P}\left(\sup_{y_1 \geq T, \frac{1}{2} \leq y_2 \leq 2} \left| \sqrt{kh^d} [T_n(s_n(y_1), y_2|x_0) - \mathbb{E}(T_n(s_n(y_1), y_2|x_0))] \right| > \varepsilon T^{\gamma_1(x_0)}\right) \\
&= \mathbb{P}\left(\sup_{y_1 \geq T, \frac{1}{2} \leq y_2 \leq 2} \left| \sum_{i=1}^n \left[ \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X_i}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(Y_i^{(2)}|x_0) \leq \frac{k}{n} y_2\}} \right. \right. \right. \\
&\quad \left. \left. \left. - \mathbb{E} \left( \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(Y^{(2)}|x_0) \leq \frac{k}{n} y_2\}} \right) \right] \right| > \frac{\varepsilon T^{\gamma_1(x_0)} \sqrt{kh^d}}{\|K\|_\infty} \right) \\
&\leq \frac{\|K\|_\infty}{\varepsilon T^{\gamma_1(x_0)} \sqrt{kh^d}} \mathbb{E} \left\{ \sup_{y_1 \geq T, \frac{1}{2} \leq y_2 \leq 2} \left| \sum_{i=1}^n \left[ \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X_i}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y_i^{(1)}|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(Y_i^{(2)}|x_0) \leq \frac{k}{n} y_2\}} \right. \right. \right. \\
&\quad \left. \left. \left. - \mathbb{E} \left( \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(Y^{(2)}|x_0) \leq \frac{k}{n} y_2\}} \right) \right] \right| \right\}.
\end{aligned}$$

Consider the class of functions

$$\begin{aligned}
g_{n,y}(u, z) &:= \frac{K}{\|K\|_\infty} \left( \frac{x_0 - z}{h} \right) \mathbb{1}_{\{\bar{F}_1(u_1|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(u_2|x_0) \leq \frac{k}{n} y_2\}} \\
&\quad - \mathbb{E} \left( \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq \frac{k}{n} s_n(y_1), \bar{F}_2(Y^{(2)}|x_0) \leq \frac{k}{n} y_2\}} \right),
\end{aligned}$$

with  $y_1 \geq T$  and  $1/2 \leq y_2 \leq 2$ , and with envelope function

$$G_n(u, z) := \frac{K}{\|K\|_\infty} \left( \frac{x_0 - z}{h} \right) \mathbb{1}_{\{\bar{F}_2(u_2|x_0) \leq \frac{2k}{n}\}} + \mathbb{E} \left( \frac{K}{\|K\|_\infty} \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{\{\bar{F}_2(Y^{(2)}|x_0) \leq \frac{2k}{n}\}} \right).$$

This class of functions satisfies the conditions of Theorem 7.3 in Wellner (2005) with  $\sigma^2 = O(kh^d/n)$  and  $PG_n^2 = O(kh^d/n)$  for  $n$  large, and thus, for some constant  $C$ ,

$$\mathbb{P}(I_{2,n}(T) > \varepsilon) \leq \frac{C}{\varepsilon T^{\gamma_1(x_0)}}$$

for  $n$  large enough. We have then that for every  $\varepsilon$  there is a  $T = T(\varepsilon)$  such that for  $n$  large enough

$$\mathbb{P}(I_{2,n}(T) > \varepsilon) \leq \varepsilon.$$

Now, to study  $I_{3,n}(T)$ , remark that for any  $T > 0, \exists n_1 = n_1(T) : \forall n > n_1 : s_n(T) < T + 1$ . Hence for  $n > n_1$  and any  $\eta_0 \in (\gamma_1(x_0), 1/2)$  :

$$\begin{aligned}
\mathbb{P}(I_{3,n}(T) > \varepsilon) &\leq \mathbb{P}\left(\sup_{0 < y_1 \leq T+1, \frac{1}{2} \leq y_2 \leq 2} \left| \frac{\sqrt{kh^d} [T_n(y_1, y_2|x_0) - \mathbb{E}(T_n(y_1, y_2|x_0))] - W(y_1, y_2)}{y_1^{\eta_0}} \right| \right. \\
&\quad \left. \times \left| \int_0^T [s_n(u)]^{\eta_0} du^{-\gamma_1(x_0)} \right| > \varepsilon \right).
\end{aligned}$$

According to Lemma 3 in Cai et al. (2015)

$$\left| \int_0^T [s_n(u)]^{\eta_0} du^{-\gamma_1(x_0)} \right| \longrightarrow \frac{\gamma_1(x_0)}{\eta_0 - \gamma_1(x_0)} T^{\eta_0 - \gamma_1(x_0)},$$

which, combining with our Theorem 2.1 and the Skorohod construction, entails that there exists  $n_2(T) > n_1(T)$  such that  $\forall n > n_2(T)$ ,  $\mathbb{P}(I_{3,n}(T) > \varepsilon) < \varepsilon$ .

Finally, concerning  $I_{4,n}(T)$ , we first remark that according to Lemma 2 in Cai et al. (2015), we have for  $\eta_0 \in (\gamma_1(x_0), 1/2)$  and any  $T > 0$ , with probability one,

$$\sup_{0 < y_1 \leq T, \frac{1}{2} \leq y_2 \leq 2} \frac{|W(y_1, y_2)|}{y_1^{\eta_0}} < \infty.$$

Then, applying Lemma 3 in Cai et al. (2015) with  $S = T$ ,  $S_0 = T + 1$  and  $g = W$ , we deduce that there exists  $n_3(T)$  such that for  $n > n_3(T)$  we have  $\mathbb{P}(I_{4,n}(T) > \varepsilon) < \varepsilon$ .

This achieves the proof of Proposition 6.1.  $\blacksquare$

In order to prove Theorem 2.3 we need some auxiliary results. Define for  $u > 0$  and  $v \in S_K$

$$\begin{aligned} \tilde{s}_n(u) &:= \frac{n}{k} \bar{F}_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right), \\ t_n(y_2) &:= \frac{n}{k} \bar{F}_2 \left( U_2 \left( \frac{n}{ky_2} \middle| x_0 \right) \middle| x_0 - hv \right). \end{aligned}$$

**Lemma 6.3.** *Assume (D) and (H) and  $x_0 \in \text{Int}(S_X)$ . Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$  and  $h^{\eta_{\varepsilon_1} \wedge \eta_{\gamma_1}} \ln \frac{n}{k} \rightarrow 0$ . Then, we have, for any  $u \leq T_n \rightarrow \infty$  such that  $kT_n/n \rightarrow 0$  and  $0 < \varepsilon < \beta_1(x_0)$ , that*

$$\begin{aligned} \left| \tilde{s}_n(u) - u \right| &\leq Cu \left\{ h^{\eta_{A_1}} + h^{\eta_{\gamma_1}} \ln \frac{n}{k} + h^{\eta_{\gamma_1}} |\ln u| u^{\pm Ch^{\eta_{\gamma_1}}} \right. \\ &\quad + \left| \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \left[ 1 + u^{\pm Ch^{\eta_{\gamma_1}}} h^{\eta_{\gamma_1}} |\ln u| \right] \\ &\quad \times \left[ u^{\gamma_1(x_0)\beta_1(x_0)} \left( 1 + u^{\pm \gamma_1(x_0)\varepsilon} \right) \left( h^{\eta_{B_1}} + u^{-Ch^{\eta_{\varepsilon_1}}} h^{\eta_{\varepsilon_1}} \left( |\ln u| + \ln \frac{n}{k} \right) \right) \right. \\ &\quad \left. \left. + u^{\gamma_1(x_0)(\beta_1(x_0) \pm \varepsilon)} + \left| u^{\gamma_1(x_0)\beta_1(x_0)} - 1 \right| \right] \right\}, \end{aligned}$$

where  $u^{\pm \bullet}$  means  $u^\bullet$  if  $u$  is greater than 1, and  $u^{-\bullet}$  if  $u$  is smaller than 1.

**Lemma 6.4.** *Assume (D), (H),  $\gamma_1(x_0) < 1$  and  $x_0 \in \text{Int}(S_X)$ . For sequences  $k = \lfloor n^\alpha \ell_1(n) \rfloor$  and  $h = n^{-\Delta} \ell_2(n)$ , where  $\ell_1$  and  $\ell_2$  are slowly varying functions at infinity, with  $\alpha \in (0, 1)$  and*

$$\begin{aligned} \max \left( \frac{\alpha}{d + 2\gamma_1(x_0)(\eta_{A_1} \wedge \eta_{\gamma_1})}, \frac{\alpha}{d + 2(1 - \gamma_1(x_0))(\eta_{A_2} \wedge \eta_{\gamma_2} \wedge \eta_{B_2} \wedge \eta_{\varepsilon_2})}, \right. \\ \left. \frac{\alpha}{d} - \frac{2(1 - \alpha)\gamma_1^2(x_0)\beta_1(x_0)}{d + d(\beta_1(x_0) + \varepsilon)\gamma_1(x_0)}, \frac{\alpha - 2(1 - \alpha)\gamma_1(x_0)}{d} \right) < \Delta < \frac{\alpha}{d}, \end{aligned}$$

one has that

$$\sup_{v \in S_K} \sup_{\frac{1}{2} \leq y_2 \leq 2} \sqrt{kh^d} \left| \int_0^\infty [R(\tilde{s}_n(u), t_n(y_2) | x_0) - R(u, y_2 | x_0)] du^{-\gamma_1(x_0)} \right| \longrightarrow 0$$

and

$$\sup_{\frac{1}{2} \leq y_2 \leq 2} \sqrt{kh^d} \left| \int_0^\infty [R(s_n(u), y_2 | x_0) - R(u, y_2 | x_0)] du^{-\gamma_1(x_0)} \right| \longrightarrow 0.$$

**Lemma 6.5.** Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$ ,  $x_0 \in \text{Int}(S_X)$  with  $f_X(x_0) > 0$  and  $y \rightarrow F_2(y|x_0)$  is strictly increasing. Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$ ,  $kh^d \rightarrow \infty$ ,  $h^{\eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{f_X} \wedge \eta_{A_2}} \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{\gamma_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} |\delta_2(U_2(n/k|x_0)|x_0)| h^{\eta_{B_2}} \rightarrow 0$ , and  $\sqrt{kh^d} |\delta_2(U_2(n/k|x_0)|x_0)| h^{\eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ . Then, for any sequence  $u_n$  satisfying

$$\sqrt{kh^d} \left( \frac{\overline{F}_2(U_2(n/k|x_0)|x_0)}{\overline{F}_2(u_n|x_0)} - 1 \right) \rightarrow c \in \mathbb{R},$$

as  $n \rightarrow \infty$ , we have

$$\sqrt{nh^d \overline{F}_2(u_n|x_0)} \left( \frac{\widehat{F}_{n,2}(u_n|x_0)}{\overline{F}_2(u_n|x_0)} - 1 \right) \rightsquigarrow \frac{W(\infty, 1)}{f_X(x_0)}.$$

**Lemma 6.6.** Assume  $(\mathcal{D})$ ,  $(\mathcal{H})$ ,  $(\mathcal{K})$ ,  $x_0 \in \text{Int}(S_X)$  with  $f_X(x_0) > 0$  and  $y \rightarrow F_2(y|x_0)$  is strictly increasing. Consider sequences  $k \rightarrow \infty$  and  $h \rightarrow 0$  as  $n \rightarrow \infty$ , in such a way that  $k/n \rightarrow 0$ ,  $kh^d \rightarrow \infty$ ,  $h^{\eta_{\varepsilon_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{f_X} \wedge \eta_{A_2}} \rightarrow 0$ ,  $\sqrt{kh^d} h^{\eta_{\gamma_2}} \ln n/k \rightarrow 0$ ,  $\sqrt{kh^d} |\delta_2(U_2(n/k|x_0)|x_0)| \rightarrow 0$ . Then, as  $n \rightarrow \infty$ , we have

$$\sqrt{kh^d} (\widehat{u}_n - 1) \rightsquigarrow \frac{\gamma_2(x_0)W(\infty, 1)}{f_X(x_0)}.$$

## 6.4 Proof of Theorem 2.3

Let  $\mathbb{E}_n(y) := \mathbb{E}(\widehat{\theta}_n(y)/U_1(n/k|x_0))$ . We have the following decomposition:

$$\begin{aligned} \sqrt{kh^d} \left( \frac{\widehat{\theta}_n}{f_X(x_0)\theta_{k/n}} - 1 \right) &= \frac{U_1(n/k|x_0)}{\theta_{k/n}} \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \frac{\widehat{\theta}_n}{U_1(n/k|x_0)} - \mathbb{E}_n(1) \right) \\ &\quad + \frac{U_1(n/k|x_0)}{\theta_{k/n}} \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \mathbb{E}_n(1) - \frac{f_X(x_0)\theta_{k/n}}{U_1(n/k|x_0)} \right) \\ &= \frac{U_1(n/k|x_0)}{\theta_{k/n}} \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \frac{\widehat{\theta}_n(\widehat{e}_n)}{U_1(n/k|x_0)} - \mathbb{E}_n(\widehat{e}_n) \right) \\ &\quad + \frac{U_1(n/k|x_0)}{\theta_{k/n}} \frac{\sqrt{kh^d}}{f_X(x_0)} (\mathbb{E}_n(\widehat{e}_n) - \mathbb{E}_n(1)) \\ &\quad + \frac{U_1(n/k|x_0)}{\theta_{k/n}} \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \mathbb{E}_n(1) - \frac{f_X(x_0)\theta_{k/n}}{U_1(n/k|x_0)} \right) \\ &=: T_1 + T_2 + T_3. \end{aligned}$$

First, remark that the common factor of the three terms,  $U_1(n/k|x_0)/\theta_{k/n}$  can be handled in a similar way as in Proposition 1 in Cai et al. (2015), i.e., as  $n \rightarrow \infty$

$$\frac{U_1(n/k|x_0)}{\theta_{k/n}} \rightarrow \frac{-1}{\int_0^\infty R(s, 1|x_0) ds^{-\gamma_1(x_0)}}.$$

Thus the three terms without this factor need to be studied.

We start with  $T_1$ . Note that

$$\sqrt{kh^d}(\hat{e}_n - 1) = -\frac{f_2(\tilde{u}_n U_2(n/k|x_0)|x_0)U_2(n/k|x_0)}{\bar{F}_2(U_2(n/k|x_0)|x_0)}\sqrt{kh^d}(\hat{u}_n - 1),$$

where  $\tilde{u}_n$  is a random value between  $\hat{u}_n$  and 1. By the continuous mapping theorem we have then

$$\frac{f_2(\tilde{u}_n U_2(n/k|x_0)|x_0)U_2(n/k|x_0)}{\bar{F}_2(U_2(n/k|x_0)|x_0)} \xrightarrow{\mathbb{P}} \frac{1}{\gamma_2(x_0)},$$

and hence by Lemma 6.6

$$\sqrt{kh^d}(\hat{e}_n - 1) \rightsquigarrow -W(\infty, 1)/f_X(x_0). \quad (6.14)$$

This implies that

$$\mathbb{P}\left(|\hat{e}_n - 1| > (kh^d)^{-1/4}\right) \rightarrow 0.$$

Hence, with probability tending to one,

$$\begin{aligned} & \left| \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \frac{\tilde{\theta}_n(\hat{e}_n)}{U_1(n/k|x_0)} - \mathbb{E}_n(\hat{e}_n) \right) + \frac{1}{f_X(x_0)} \int_0^\infty W(s, 1) ds^{-\gamma_1(x_0)} \right| \\ \leq & \sup_{|y-1| \leq (kh^d)^{-1/4}} \left| \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \frac{\tilde{\theta}_n(y)}{U_1(n/k|x_0)} - \mathbb{E}_n(y) \right) + \frac{1}{f_X(x_0)} \int_0^\infty W(s, y) ds^{-\gamma_1(x_0)} \right| \\ & + \frac{1}{f_X(x_0)} \sup_{|y-1| \leq (kh^d)^{-1/4}} \left| \int_0^\infty [W(s, y) - W(s, 1)] ds^{-\gamma_1(x_0)} \right|. \end{aligned}$$

The first term of the right-hand side tends to 0 in probability by our Proposition 6.1, whereas the second term can be handled similarly as in the proof of Proposition 3 in Cai et al. (2015). Consequently

$$T_1 \rightsquigarrow \frac{1}{\int_0^\infty R(s, 1|x_0) ds^{-\gamma_1(x_0)}} \frac{1}{f_X(x_0)} \int_0^\infty W(s, 1) ds^{-\gamma_1(x_0)}. \quad (6.15)$$

Next step consists to look at  $T_2$ . To this aim, remark that for  $y$  equal either to 1 or  $\hat{e}_n$ , we have



$$\begin{aligned}
& \int_0^\infty \mathbb{E}(T_n(s_n(u), y|x_0)) du^{-\gamma_1(x_0)} \\
&= \int_0^\infty \int_{S_K} K(v) R_{\frac{n}{k}}(\tilde{s}_n(u), t_n(y)|x_0 - hv) f_X(x_0 - hv) dv du^{-\gamma_1(x_0)} \\
&= \int_0^\infty \int_{S_K} K(v) R(\tilde{s}_n(u), t_n(y)|x_0) f_X(x_0 - hv) dv du^{-\gamma_1(x_0)} \\
&\quad + \int_0^\infty \int_{S_K} K(v) [R_{\frac{n}{k}}(\tilde{s}_n(u), t_n(y)|x_0 - hv) - R(\tilde{s}_n(u), t_n(y)|x_0)] f_X(x_0 - hv) dv du^{-\gamma_1(x_0)} \\
&= \int_0^\infty R(u, y|x_0) du^{-\gamma_1(x_0)} \int_{S_K} K(v) f_X(x_0 - hv) dv \\
&\quad + \int_{S_K} K(v) \int_0^\infty [R(\tilde{s}_n(u), t_n(y)|x_0) - R(u, y|x_0)] du^{-\gamma_1(x_0)} f_X(x_0 - hv) dv \\
&\quad + \int_0^\infty \int_{S_K} K(v) [R_{\frac{n}{k}}(\tilde{s}_n(u), t_n(y)|x_0 - hv) - R(\tilde{s}_n(u), t_n(y)|x_0)] f_X(x_0 - hv) dv du^{-\gamma_1(x_0)} \\
&=: \tilde{T}_{2,1} + \tilde{T}_{2,2} + \tilde{T}_{2,3}.
\end{aligned}$$

By Lemma 6.4, Assumptions  $(\mathcal{S})$  and  $(\mathcal{H})$  we obtain

$$\begin{aligned}
\tilde{T}_{2,1} &= f_X(x_0) \int_0^\infty R(u, y|x_0) du^{-\gamma_1(x_0)} + O_{\mathbb{P}}(h^{\eta_{f_X}}), \\
\tilde{T}_{2,2} &= o_{\mathbb{P}}\left(\frac{1}{\sqrt{kh^d}}\right), \\
|\tilde{T}_{2,3}| &\leq - \sup_{x \in B(x_0, h)} \sup_{0 < y_1 < \infty, \frac{1}{2} \leq y_2 \leq 2} \frac{|R_{n/k}(y_1, y_2|x) - R(y_1, y_2|x_0)|}{y_1^\beta \wedge 1} \\
&\quad \times \int_{S_K} K(v) \int_0^\infty ([\tilde{s}_n(u)]^\beta \wedge 1) du^{-\gamma_1(x_0)} f_X(x_0 - hv) dv \\
&= O_{\mathbb{P}}\left(\left(\frac{n}{k}\right)^\tau\right).
\end{aligned}$$

Note that the integral appearing in the bound for  $|\tilde{T}_{2,3}|$  is finite for  $n$  large, as  $\tilde{s}_n(u) \leq Cu^{1-\xi}$  for  $u \in (0, 1/2]$ ,  $\xi \in (0, (\beta - \gamma_1(x_0))/\beta)$  and  $n$  large. Consequently, under our assumptions and using the homogeneity of the  $R$ -function and the mean value theorem combining with (6.14), we have

$$\begin{aligned}
& \frac{\sqrt{kh^d}}{f_X(x_0)} (\mathbb{E}_n(\hat{e}_n) - \mathbb{E}_n(1)) \\
&= \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \int_0^\infty \mathbb{E}(T_n(s_n(u), 1|x_0)) du^{-\gamma_1(x_0)} - \int_0^\infty \mathbb{E}(T_n(s_n(u), \hat{e}_n|x_0)) du^{-\gamma_1(x_0)} \right) \\
&= \sqrt{kh^d} \left( \int_0^\infty R(u, 1|x_0) du^{-\gamma_1(x_0)} - \int_0^\infty R(u, \hat{e}_n|x_0) du^{-\gamma_1(x_0)} \right) + o_{\mathbb{P}}(1) \\
&= \sqrt{kh^d} \left( 1 - \hat{e}_n^{1-\gamma_1(x_0)} \right) \int_0^\infty R(u, 1|x_0) du^{-\gamma_1(x_0)} + o_{\mathbb{P}}(1) \\
&\rightsquigarrow (1 - \gamma_1(x_0)) \frac{W(\infty, 1)}{f_X(x_0)} \int_0^\infty R(u, 1|x_0) du^{-\gamma_1(x_0)}.
\end{aligned}$$

This implies that

$$T_2 \rightsquigarrow -(1 - \gamma_1(x_0)) \frac{W(\infty, 1)}{f_X(x_0)}. \quad (6.16)$$

Finally, for  $T_3$  we have,

$$\begin{aligned} & \frac{\sqrt{kh^d}}{f_X(x_0)} \left( \mathbb{E}_n(1) - \frac{f_X(x_0)\theta_{k/n}}{U_1(n/k|x_0)} \right) \\ = & \frac{\sqrt{kh^d}}{f_X(x_0)} \left( - \int_0^\infty \mathbb{E}(T_n(s_n(u), 1|x_0)) du^{-\gamma_1(x_0)} - \frac{f_X(x_0)\theta_{k/n}}{U_1(n/k|x_0)} \right) \\ = & \sqrt{kh^d} \int_0^\infty [R_{n/k}(s_n(u), 1|x_0) - R(u, 1|x_0)] du^{-\gamma_1(x_0)} + o(1) \\ = & \sqrt{kh^d} \int_0^\infty [R_{n/k}(s_n(u), 1|x_0) - R(s_n(u), 1|x_0)] du^{-\gamma_1(x_0)} \\ & + \sqrt{kh^d} \int_0^\infty [R(s_n(u), 1|x_0) - R(u, 1|x_0)] du^{-\gamma_1(x_0)} + o(1) \\ =: & \tilde{T}_{3,1} + \tilde{T}_{3,2} + o(1), \end{aligned}$$

where

$$\begin{aligned} |\tilde{T}_{3,1}| & \leq \sqrt{kh^d} \sup_{x \in B(x_0, h)} \sup_{0 < y_1 < \infty, \frac{1}{2} \leq y_2 \leq 2} \frac{|R_{n/k}(y_1, y_2|x) - R(y_1, y_2|x_0)|}{y_1^\beta \wedge 1} \\ & \quad \times \left| \int_0^\infty ([s_n(u)]^\beta \wedge 1) du^{-\gamma_1(x_0)} \right| \\ & = O\left(\sqrt{kh^d} \left(\frac{n}{k}\right)^\tau\right), \\ \tilde{T}_{3,2} & = o(1). \end{aligned}$$

Overall, we have then

$$T_3 = o(1). \quad (6.17)$$

Combining (6.15), (6.16) and (6.17), and following the argument as at the end of the proof of Theorem 2.2, we can establish the result of Theorem 2.3.  $\blacksquare$

## 6.5 Proofs of the auxiliary results

### Proof of Lemma 1

First note that, by continuity of  $y \rightarrow F_j(y|x)$ ,

$$t_n \bar{F}_j(U_j(t_n/y|x_0)|x) = y \frac{\bar{F}_j(U_j(t_n/y|x_0)|x)}{\bar{F}_j(U_j(t_n/y|x_0)|x_0)}.$$

Then, from condition  $(\mathcal{D})$ , and a straightforward decomposition,

$$\begin{aligned} & \left| \frac{t_n \bar{F}_j(U_j(t_n/y|x_0)|x)}{y^\eta} - y^{1-\eta} \right| \\ & \leq y^{1-\eta} \left\{ \left| \frac{A_j(x)}{A_j(x_0)} - 1 \right| (U_j(t_n/y|x_0))^{1/\gamma_j(x_0)-1/\gamma_j(x)} \frac{1 + \frac{1}{\gamma_j(x)} \delta_j(U_j(t_n/y|x_0)|x)}{1 + \frac{1}{\gamma_j(x_0)} \delta_j(U_j(t_n/y|x_0)|x_0)} \right. \\ & \quad + \left| (U_j(t_n/y|x_0))^{1/\gamma_j(x_0)-1/\gamma_j(x)} - 1 \right| \frac{1 + \frac{1}{\gamma_j(x)} \delta_j(U_j(t_n/y|x_0)|x)}{1 + \frac{1}{\gamma_j(x_0)} \delta_j(U_j(t_n/y|x_0)|x_0)} \\ & \quad \left. + \left| \frac{1 + \frac{1}{\gamma_j(x)} \delta_j(U_j(t_n/y|x_0)|x)}{1 + \frac{1}{\gamma_j(x_0)} \delta_j(U_j(t_n/y|x_0)|x_0)} - 1 \right| \right\}. \end{aligned}$$

Each of the absolute differences in the right-hand side of the above display can be handled by condition  $(\mathcal{H})$ . Obviously, for some constant  $C$ ,

$$\left| \frac{A_j(x)}{A_j(x_0)} - 1 \right| \leq C h_n^{\eta A_j}, \quad \text{for } x \in B(x_0, h_n).$$

Next, using the inequality  $|e^z - 1| \leq e^{|z|}|z|$ , we have, for some constant  $C$  (not necessarily equal to the one introduced above), and  $x \in B(x_0, h_n)$ ,

$$\left| (U_j(t_n/y|x_0))^{1/\gamma_j(x_0)-1/\gamma_j(x)} - 1 \right| \leq e^{C h_n^{\eta \gamma_j} \ln U_j(t_n/y|x_0)} C h_n^{\eta \gamma_j} \ln U_j(t_n/y|x_0).$$

For distributions satisfying  $(\mathcal{D})$ , one easily verifies that

$$U_j(t_n|x_0) = (A_j(x_0))^{\gamma_j(x_0)} t_n^{\gamma_j(x_0)} (1 + a_j(t_n|x_0))$$

where  $|a_j(\cdot|x_0)|$  is regularly varying with index equal to  $-\gamma_j(x_0)\beta_j(x_0)$ . Hence, for some constants  $C_1$  and  $C_2$ , not depending on  $y$ , one gets for  $x \in B(x_0, h_n)$  and  $n$  large,

$$\left| (U_j(t_n/y|x_0))^{1/\gamma_j(x_0)-1/\gamma_j(x)} - 1 \right| \leq C_1 t_n^{C_2 h_n^{\eta \gamma_j}} y^{-C_2 h_n^{\eta \gamma_j}} \left( h_n^{\eta \gamma_j} \ln t_n - h_n^{\eta \gamma_j} \ln y \right).$$

Finally, for  $n$  large,

$$\begin{aligned} & \left| \frac{1 + \frac{1}{\gamma_j(x)} \delta_j(U_j(t_n/y|x_0)|x)}{1 + \frac{1}{\gamma_j(x_0)} \delta_j(U_j(t_n/y|x_0)|x_0)} - 1 \right| \\ & \leq C |\delta_j(U_j(t_n/y|x_0)|x_0)| \left\{ \left| \frac{\delta_j(U_j(t_n/y|x_0)|x)}{\delta_j(U_j(t_n/y|x_0)|x_0)} - 1 \right| + \left| \frac{1}{\gamma_j(x)} - \frac{1}{\gamma_j(x_0)} \right| \right\}. \end{aligned}$$

By the assumptions on  $\delta_j$  we obtain

$$\begin{aligned} \left| \frac{\delta_j(U_j(t_n/y|x_0)|x)}{\delta_j(U_j(t_n/y|x_0)|x_0)} - 1 \right| & \leq \left| \frac{B_j(x)}{B_j(x_0)} - 1 \right| e^{\int_1^{U_j(t_n/y|x_0)} \frac{\varepsilon_j(u|x) - \varepsilon_j(u|x_0)}{u} du} \\ & \quad + \left| e^{\int_1^{U_j(t_n/y|x_0)} \frac{\varepsilon_j(u|x) - \varepsilon_j(u|x_0)}{u} du} - 1 \right|, \end{aligned}$$

and, hence, using  $(\mathcal{H})$ , for  $x \in B(x_0, h_n)$  and  $n$  large,

$$\begin{aligned} & \left| \frac{1 + \frac{1}{\gamma_j(x)} \delta_j(U_j(t_n/y|x_0)|x)}{1 + \frac{1}{\gamma_j(x_0)} \delta_j(U_j(t_n/y|x_0)|x_0)} - 1 \right| \\ & \leq C_1 \left[ h_n^{\eta_{\gamma_j} \wedge \eta_{B_j}} + t_n^{C_2 h_n^{\eta_{\varepsilon_j}}} y^{-C_2 h_n^{\eta_{\varepsilon_j}}} \left( h_n^{\eta_{\varepsilon_j}} \ln t_n - h_n^{\eta_{\varepsilon_j}} \ln y \right) \right]. \end{aligned}$$

Combining the above results establishes the lemma.  $\blacksquare$

### Proof of Lemma 2

We have

$$\begin{aligned} & \mathbb{E}(T_n(y_1, y_2|x_0)) \\ & = \frac{n}{k} \mathbb{E} \left[ K_h(x_0 - X) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2\}} \right] \\ & = \frac{n}{k} \int_{S_K} K(v) \mathbb{P}(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 | X = x_0 - hv) \\ & \quad \times f_X(x_0 - hv) dv \\ & = \int_{S_K} K(v) R(y_1, y_2|x_0 - hv) f_X(x_0 - hv) dv \\ & \quad + \int_{S_K} K(v) \left[ \frac{n}{k} \mathbb{P}(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 | X = x_0 - hv) \right. \\ & \quad \left. - R(y_1, y_2|x_0 - hv) \right] f_X(x_0 - hv) dv \\ & =: T_{1,n} + T_{2,n}. \end{aligned}$$

Concerning  $T_{1,n}$ , by the continuity of  $f_X(x)$  and  $R(y_1, y_2|x)$  at  $x_0$ , we have that  $f_X$  and  $R$  are bounded in a neighborhood of  $x_0$ , and hence, by Lebesgue's dominated convergence theorem

$$T_{1,n} \rightarrow f_X(x_0) R(y_1, y_2|x_0), \text{ as } n \rightarrow \infty.$$

As for  $T_{2,n}$ ,

$$\begin{aligned} |T_{2,n}| & \leq \sup_{v \in S_K} \left| \frac{n}{k} \mathbb{P}(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 | X = x_0 - hv) \right. \\ & \quad \left. - R(y_1, y_2|x_0 - hv) \right| \int_{S_K} K(v) f_X(x_0 - hv) dv, \end{aligned}$$

and note that

$$\begin{aligned} & \mathbb{P}(\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2 | X = x_0 - hv) \\ & = \mathbb{P} \left( \bar{F}_1(Y^{(1)}|x_0 - hv) \leq \frac{k}{n} \frac{n}{k} \bar{F}_1(U_1(n/(ky_1)|x_0)|x_0 - hv), \right. \\ & \quad \left. \bar{F}_2(Y^{(2)}|x_0 - hv) \leq \frac{k}{n} \frac{n}{k} \bar{F}_2(U_2(n/(ky_2)|x_0)|x_0 - hv) | X = x_0 - hv \right). \end{aligned}$$

Then, by the result of Lemma 1 and the uniformity of the convergence in Assumption  $(\mathcal{R})$ , we have that  $T_{2,n} \rightarrow 0$  as  $n \rightarrow \infty$ .

Now, consider the variance. We have

$$\begin{aligned}
& kh^d \text{Var}(T_n(y_1, y_2 | x_0)) \\
&= \frac{nh^d \text{Var}(K_h(x_0 - X) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2\}})}{k} \\
&= \|K\|_2^2 \frac{n}{k} \mathbb{E} \left[ \frac{1}{h^d \|K\|_2^2} K^2 \left( \frac{x_0 - X}{h} \right) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2\}} \right] \\
&\quad - \frac{kh^d}{n} \left\{ \frac{n}{k} \mathbb{E} \left[ K_h(x_0 - X) \mathbb{1}_{\{\bar{F}_1(Y^{(1)}|x_0) \leq (k/n) y_1, \bar{F}_2(Y^{(2)}|x_0) \leq (k/n) y_2\}} \right] \right\}^2,
\end{aligned}$$

from which the result follows. ■

### Proof of Lemma 3

Using Assumption (D), we have

$$\begin{aligned}
\tilde{s}_n(u) &= \frac{\bar{F}_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{\bar{F}_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \\
&= \frac{A_1(x_0 - hv)}{A_1(x_0)} \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \right)^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - hv)}} u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \\
&\quad \times \frac{1 + \frac{1}{\gamma_1(x_0 - hv)} \delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)}.
\end{aligned}$$

This implies that

$$\begin{aligned}
& \left| \tilde{s}_n(u) - u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \right| \\
&\leq u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \left\{ \left| \frac{A_1(x_0 - hv)}{A_1(x_0)} - 1 \right| \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \right)^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - hv)}} \right. \\
&\quad \times \left| \frac{1 + \frac{1}{\gamma_1(x_0 - hv)} \delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \right| \\
&\quad + \left| \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \right)^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - hv)}} - 1 \right| \\
&\quad \times \left| \frac{1 + \frac{1}{\gamma_1(x_0 - hv)} \delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \right| \\
&\quad \left. + \left| \frac{1 + \frac{1}{\gamma_1(x_0 - hv)} \delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{1 + \frac{1}{\gamma_1(x_0)} \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} - 1 \right| \right\} \\
&=: u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \{T_1 + T_2 + T_3\}.
\end{aligned}$$

Using Assumption  $(\mathcal{H})$  and the inequality  $|e^x - 1| \leq |x| e^{|x|}$ , we deduce that, for  $n$  large,

$$\left| \frac{A_1(x_0 - hv)}{A_1(x_0)} - 1 \right| \leq C h^{\eta_{A_1}} \quad (6.18)$$

$$\left| \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \right)^{\frac{1}{\gamma_1(x_0)} - \frac{1}{\gamma_1(x_0 - hv)}} - 1 \right| \leq C h^{\eta_{\gamma_1}} \ln \frac{n}{k}. \quad (6.19)$$

Now, direct computations yield, for  $n$  large,

$$\begin{aligned} T_3 &\leq C \left| \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \left\{ \left| \frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)} - 1 \right| \right. \\ &\quad \times \left| \frac{\delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{\delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \right| \\ &\quad \left. + \left| \frac{\delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 - hv \right)}{\delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} \frac{\delta_1 \left( u^{-\gamma_1(x_0)} U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)}{\delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right)} - 1 \right| \right\}. \end{aligned} \quad (6.20)$$

Using the assumed form for  $\delta_1(y|x)$ ,  $(\mathcal{H})$ , and the uniform bound from Proposition B.1.10 in de Haan and Ferreira (2006) with  $0 < \varepsilon < \beta_1(x_0)$ , we obtain, for  $n$  large, that

$$\begin{aligned} T_3 &\leq C \left| \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \left\{ h^{\eta_{\gamma_1}} + u^{\gamma_1(x_0)\beta_1(x_0)} \left( 1 + u^{\pm\gamma_1(x_0)\varepsilon} \right) \right. \\ &\quad \times \left[ h^{\eta_{B_1}} + u^{-Ch^{\eta_{\varepsilon_1}}} h^{\eta_{\varepsilon_1}} \left( |\ln u| + \ln \frac{n}{k} \right) \right] \\ &\quad \left. + u^{\gamma_1(x_0)(\beta_1(x_0) \pm \varepsilon)} + \left| u^{\gamma_1(x_0)\beta_1(x_0)} - 1 \right| \right\}. \end{aligned} \quad (6.21)$$

Since

$$\begin{aligned} \left| \tilde{s}_n(u) - u \right| &\leq \left| \tilde{s}_n(u) - u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \right| + u \left| u^{\frac{\gamma_1(x_0) - \gamma_1(x_0 - hv)}{\gamma_1(x_0 - hv)}} - 1 \right| \\ &\leq \left| \tilde{s}_n(u) - u^{\frac{\gamma_1(x_0)}{\gamma_1(x_0 - hv)}} \right| + C u^{1 \pm Ch^{\eta_{\gamma_1}}} h^{\eta_{\gamma_1}} |\ln u|, \end{aligned} \quad (6.22)$$

combining (6.18), (6.19), (6.21) with (6.22), Lemma 3 is established.  $\blacksquare$

#### Proof of Lemma 4

We use the following decomposition along with the Lipschitz property of the function  $R$ :

$$\begin{aligned}
& \left| \sqrt{kh^d} \int_0^\infty [R(\tilde{s}_n(u), t_n(y_2)|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\
& \leq \sqrt{kh^d} \left| \int_0^{\delta_n} [R(\tilde{s}_n(u), t_n(y_2)|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\
& \quad + \sqrt{kh^d} \left| \int_{\delta_n}^{T_n} [R(\tilde{s}_n(u), t_n(y_2)|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\
& \quad + \sqrt{kh^d} \left| \int_{T_n}^\infty [R(\tilde{s}_n(u), t_n(y_2)|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\
& \leq \sqrt{kh^d} \left| \int_0^{\delta_n} R(\tilde{s}_n(u), t_n(y_2)|x_0) du^{-\gamma_1(x_0)} \right| \\
& \quad + \sqrt{kh^d} \left| \int_0^{\delta_n} R(u, y_2|x_0) du^{-\gamma_1(x_0)} \right| \\
& \quad - \sqrt{kh^d} \int_{\delta_n}^{T_n} [|\tilde{s}_n(u) - u| + |t_n(y_2) - y_2|] du^{-\gamma_1(x_0)} \\
& \quad + 2 \sup_{u \geq 0, \frac{1}{2} - \zeta \leq y_2 \leq 2 + \zeta} R(u, y_2|x_0) \sqrt{kh^d} T_n^{-\gamma_1(x_0)} \\
& =: T_1 + T_2 + T_3 + T_4,
\end{aligned}$$

for  $\zeta > 0$  small and where  $\delta_n \rightarrow 0$  and  $T_n \rightarrow \infty$ , as  $n \rightarrow \infty$ .

Now, since  $R(y_1, y_2|x_0) \leq y_1 \wedge y_2$ , using Lemma 3, and assuming  $h^{\eta_{\varepsilon_1} \wedge \eta_{\gamma_1}} |\ln \delta_n| \rightarrow 0$ , we obtain after tedious calculations, for  $n$  large,

$$\begin{aligned}
T_1 + T_2 & \leq -2\sqrt{kh^d} \int_0^{\delta_n} u du^{-\gamma_1(x_0)} - \sqrt{kh^d} \int_0^{\delta_n} |\tilde{s}_n(u) - u| du^{-\gamma_1(x_0)} \\
& \leq C \sqrt{kh^d} \delta_n^{1-\gamma_1(x_0)}.
\end{aligned} \tag{6.23}$$

As for  $T_3$ , using again Lemma 3 and following the lines of proof of Lemma 1, we have, for  $n$  large,

$$\begin{aligned}
T_3 & \leq -\sqrt{kh^d} \int_0^{T_n} |\tilde{s}_n(u) - u| du^{-\gamma_1(x_0)} - \sqrt{kh^d} \int_{\delta_n}^{T_n} |t_n(y_2) - y_2| du^{-\gamma_1(x_0)} \\
& \leq C \sqrt{kh^d} T_n^{1-\gamma_1(x_0)} \left\{ h^{\eta_{A_1}} + h^{\eta_{\gamma_1}} \ln \frac{n}{k} + h^{\eta_{\gamma_1}} \ln T_n \right. \\
& \quad \left. + \left| \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| T_n^{(\beta_1(x_0) + \varepsilon)\gamma_1(x_0)} \right\} \\
& \quad + C \sqrt{kh^d} \delta_n^{-\gamma_1(x_0)} \left\{ h^{\eta_{A_2}} + h^{\eta_{\gamma_2}} \ln \frac{n}{k} + \left| \delta_2 \left( U_2 \left( \frac{n}{k} \middle| x_0 \right) \middle| x_0 \right) \right| \right. \\
& \quad \left. \times \left[ h^{\eta_{B_2}} + h^{\eta_{\varepsilon_2}} \ln \frac{n}{k} \right] \right\}
\end{aligned} \tag{6.24}$$

assuming  $h^{\eta_{\varepsilon_1} \wedge \eta_{\gamma_1}} \ln T_n \rightarrow 0$ .

Finally

$$T_4 \leq C \sqrt{kh^d} T_n^{-\gamma_1(x_0)}. \tag{6.25}$$

Take  $\delta_n = h^\xi$  and  $T_n = n^\kappa$ , with  $\xi$  and  $\kappa$  positive numbers, and  $0 < \varepsilon < \beta_1(x_0)$ . Combining (6.23), (6.24) and (6.25), the first part of Lemma 4 follows if the sequences  $\delta_n$  and  $T_n$  are chosen such that

$$\begin{aligned} \alpha - \Delta [d - 2\xi\gamma_1(x_0) + 2(\xi \wedge \eta_{A_2} \wedge \eta_{B_2} \wedge \eta_{\gamma_2} \wedge \eta_{\varepsilon_2})] &< 0, \\ \alpha - \Delta d - 2\kappa\gamma_1(x_0) &< 0, \\ \alpha - \Delta d + 2\kappa(1 - \gamma_1(x_0)) - 2\Delta(\eta_{A_1} \wedge \eta_{\gamma_1}) &< 0, \\ \alpha - \Delta d - 2(1 - \alpha)\gamma_1(x_0)\beta_1(x_0) + 2\kappa[1 + (\beta_1(x_0) + \varepsilon)\gamma_1(x_0) - \gamma_1(x_0)] &< 0. \end{aligned}$$

Note that this is possible if we proceed as follows:

- $\alpha$  and  $\Delta$  are chosen as stated in Lemma 4;
- $\kappa$  is chosen such that

$$\frac{\alpha - \Delta d}{2\gamma_1(x_0)} < \kappa < \min \left( 1 - \alpha, \frac{2\Delta(\eta_{A_1} \wedge \eta_{\gamma_1}) - (\alpha - \Delta d)}{2(1 - \gamma_1(x_0))}, \frac{2(1 - \alpha)\gamma_1(x_0)\beta_1(x_0) - (\alpha - \Delta d)}{2[1 - \gamma_1(x_0) + (\beta_1(x_0) + \varepsilon)\gamma_1(x_0)]} \right);$$

- $\xi$  is chosen such that

$$\frac{\alpha - \Delta d}{2\Delta(1 - \gamma_1(x_0))} < \xi < \eta_{A_2} \wedge \eta_{\gamma_2} \wedge \eta_{B_2} \wedge \eta_{\varepsilon_2}.$$

Note that the choices of  $\kappa$  and  $\xi$  only depend on those of  $\alpha$  and  $\Delta$ .

The second part of Lemma 4 is similar, although simpler. Indeed, a decomposition of the quantity of interest this time into two parts yields

$$\begin{aligned} &\sqrt{kh^d} \left| \int_0^\infty [R(s_n(u), y_2|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\ &\leq \sqrt{kh^d} \left| \int_0^{T_n} [R(s_n(u), y_2|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\ &\quad + \sqrt{kh^d} \left| \int_{T_n}^\infty [R(s_n(u), y_2|x_0) - R(u, y_2|x_0)] du^{-\gamma_1(x_0)} \right| \\ &\leq -\sqrt{kh^d} \int_0^{T_n} |s_n(u) - u| du^{-\gamma_1(x_0)} + 2 \sup_{u \geq 0, \frac{1}{2} \leq y_2 \leq 2} R(u, y_2|x_0) \sqrt{kh^d} T_n^{-\gamma_1(x_0)} \\ &\leq -\sqrt{kh^d} \frac{|\delta_1(U_1(\frac{n}{k}|x_0)|x_0)|}{|\gamma_1(x_0) + \delta_1(U_1(\frac{n}{k}|x_0)|x_0)|} \int_0^{T_n} u \left| \frac{\delta_1(u^{-\gamma_1(x_0)}U_1(\frac{n}{k}|x_0)|x_0)}{\delta_1(U_1(\frac{n}{k}|x_0)|x_0)} - 1 \right| du^{-\gamma_1(x_0)} \\ &\quad + C\sqrt{kh^d} T_n^{-\gamma_1(x_0)} \\ &\leq C\sqrt{kh^d} \left| \delta_1 \left( U_1 \left( \frac{n}{k} \middle| x_0 \right) \right) \right| T_n^{1-\gamma_1(x_0)+(\beta_1(x_0)+\varepsilon)\gamma_1(x_0)} + C\sqrt{kh^d} T_n^{-\gamma_1(x_0)}. \end{aligned}$$

This achieves the proof of Lemma 4. ■



### Proof of Lemma 5

In this proof, as mentioned above, we will use the Skorohod representation with keeping the same notation. First remark that

$$\begin{aligned} & \frac{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} > u_n\}}}{\overline{F}_2(u_n|x_0)} \\ &= \frac{\overline{F}_2(U_2(n/k|x_0)|x_0)}{\overline{F}_2(u_n|x_0)} T_n \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \middle| x_0 \right), \quad a.s. \end{aligned}$$

We have, with  $r_n := \sqrt{nh^d \overline{F}_2(u_n|x_0)}$ ,

$$\begin{aligned} & \left| r_n \left[ \frac{\overline{F}_2(U_2(n/k|x_0)|x_0)}{\overline{F}_2(u_n|x_0)} T_n \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \middle| x_0 \right) - f_X(x_0) \right] - W(\infty, 1) \right| \\ & \leq \left| \sqrt{kh^d} \left[ T_n \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \middle| x_0 \right) - \frac{n}{k} \overline{F}_2(u_n|x_0) f_X(x_0) \right] - W \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \right) \right| \\ & \quad + \sqrt{kh^d} \left| \sqrt{\frac{\overline{F}_2(u_n|x_0)}{\overline{F}_2(U_2(n/k|x_0)|x_0)}} - 1 \right| \left| T_n \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \middle| x_0 \right) - \frac{n}{k} \overline{F}_2(u_n|x_0) f_X(x_0) \right| \\ & \quad + \left| W \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \right) - W(\infty, 1) \right| \\ & \quad + r_n \left| \frac{\overline{F}_2(U_2(n/k|x_0)|x_0)}{\overline{F}_2(u_n|x_0)} - 1 \right| \left| T_n \left( \infty, \frac{n}{k} \overline{F}_2(u_n|x_0) \middle| x_0 \right) - \frac{n}{k} \overline{F}_2(u_n|x_0) f_X(x_0) \right|. \quad (6.26) \end{aligned}$$

From (6.12) combined with the Skorohod construction, we have

$$r_n \left( \frac{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} > u_n\}}}{\overline{F}_2(u_n|x_0)} - f_X(x_0) \right) \rightsquigarrow W(\infty, 1).$$

Finally

$$\begin{aligned} r_n \left( \frac{\widehat{\overline{F}}_{n,2}(u_n|x_0)}{\overline{F}_2(u_n|x_0)} - 1 \right) &= r_n \left( \frac{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} > u_n\}}}{\overline{F}_2(u_n|x_0) f_X(x_0)} - 1 \right) \\ & \quad + r_n \frac{f_X(x_0) - \hat{f}_n(x_0)}{f_X(x_0) \hat{f}_n(x_0)} \frac{\frac{1}{n} \sum_{i=1}^n K_h(x_0 - X_i) \mathbb{1}_{\{Y_i^{(2)} > u_n\}}}{\overline{F}_2(u_n|x_0)} \\ & \rightsquigarrow \frac{W(\infty, 1)}{f_X(x_0)}. \end{aligned}$$

■

### Proof of Lemma 6

To prove the lemma we will use the idea of Wretman (1978), applied to our situation. We have,

for  $z \in \mathbb{R}$ , and  $u_n$  from Lemma 5 taken as  $U_2(n/k|x_0)(1 + z/\sqrt{kh^d})$ , that

$$\begin{aligned} & \mathbb{P}\left(\sqrt{kh^d}(\hat{u}_n - 1) \leq z\right) \\ &= \mathbb{P}\left(\sqrt{nh^d \bar{F}_2(u_n|x_0)} \left(\frac{\hat{F}_{n,2}(u_n|x_0)}{\bar{F}_2(u_n|x_0)} - 1\right)\right. \\ & \quad \left. \leq \sqrt{nh^d \bar{F}_2(u_n|x_0)} \left(\frac{\bar{F}_2(U_2(n/k|x_0)|x_0)}{\bar{F}_2(u_n|x_0)} - 1\right)\right). \end{aligned}$$

We have that in the present context

$$a_n := \sqrt{nh^d \bar{F}_2(u_n|x_0)} \left(\frac{\bar{F}_2(U_2(n/k|x_0)|x_0)}{\bar{F}_2(u_n|x_0)} - 1\right) \rightarrow \frac{z}{\gamma_2(x_0)}.$$

Let  $H_n$  denote the distribution function of  $\sqrt{nh^d \bar{F}_2(u_n|x_0)}(\hat{F}_{n,2}(u_n|x_0)/\bar{F}_2(u_n|x_0) - 1)$ , and  $H$  is the distribution function of  $W(\infty, 1)/f_X(x_0)$ . Then by Lemma 5 and by continuity of  $H$  one has that  $H_n(a_n) \rightarrow H(z/\gamma_2(x_0))$ , as  $n \rightarrow \infty$ , hence the result of the lemma.  $\blacksquare$

## Acknowledgements

The authors would like to thank the referees and Associate Editor for their helpful comments. The research of Armelle Guillou was supported by the French National Research Agency under the grant ANR-19-CE40-0013-01/ExtremReg project and an International Emerging Action (IEA-00179). Computation/simulation for the work described in this paper was supported by the DeIC National HPC Centre, SDU.

## References

- [1] Acharya, V., Pedersen, L., Philippon, T., Richardson, M.: Measuring systemic risk. FRB of Cleveland Working Paper No. 10-02. <https://ssrn.com/abstract=1595075> (2010).
- [2] Aerts, M., Claeskens, G.: Local polynomial estimation in multiparameter likelihood models. *J. Amer. Statist. Assoc.* 92, 1536–1545 (1997).
- [3] Artzner, P., Delbaen, F., Eber, J-M., Heath, D.: Coherent measures of risk. *Math. Finance* 9, 203–228 (1999).
- [4] Bai, S., Taqqu, M.S.: Multivariate limit theorems in the context of long-range dependence. *J. Time Series Anal.* 34, 717–743 (2013).
- [5] Beirlant, J., Joossens, E., Segers, J.: Second-order refined peaks-over-threshold modelling for heavy-tailed distributions. *J. Statist. Plann. Inference* 139, 2800–2815 (2009).
- [6] Cai, J.J., Einmahl, J.H.J., de Haan, L., Zhou, C.: Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 77, 417–442 (2015).

- [7] Cai, J., Li, H.: Conditional tail expectations for multivariate phase-type distributions. *J. Appl. Probab.* 42, 810–825 (2005).
- [8] Cai, J.-J., Musta, E.: Estimation of the marginal expected shortfall under asymptotic independence. *Scand. J. Stat.* 47, 56–83 (2020).
- [9] de Carvalho, M.: Statistics of extremes: challenges and opportunities, in *Extreme Events in Finance: A Handbook of Extreme Value Theory and Its Applications*, ed. Longin, F., Hoboken: Wiley (2016).
- [10] Castro, D., de Carvalho, M., Wadsworth, J. L.: Time-varying extreme value dependence with application to leading European stock markets. *Annals of Applied Statistics* 12, 283–309 (2018).
- [11] Daouia, A., Gardes, L., Girard, S.: On kernel smoothing for extremal quantile regression. *Bernoulli* 19, 2557–2589 (2013).
- [12] Daouia, A., Gardes, L., Girard, S., Lekina, A.: Kernel estimators of extreme level curves. *TEST* 20, 311–333 (2011).
- [13] Das, B., Fasen-Hartmann, V.: Risk contagion under regular variation and asymptotic tail independence. *J. Multivariate Anal.* 165, 194–215 (2018).
- [14] Das, B., Fasen-Hartmann, V.: Conditional excess risk measures and multivariate regular variation. *Stat. Risk Model.* 36, 1–23 (2019).
- [15] Di Bernardino, E., Prieur, C.: Estimation of the multivariate conditional tail expectation for extreme risk levels: Illustration on environmental data sets. *Environmetrics* 29, 7, 1–22 (2018).
- [16] Dierckx, G., Goegebeur, Y., Guillou, A.: Local robust and asymptotically unbiased estimation of conditional Pareto type-tails. *TEST* 23, 330–355 (2014).
- [17] Escobar-Bach, M., Goegebeur, Y., Guillou, A.: Local robust estimation of the Pickands dependence function. *Ann. Statist.* 46, 2806–2843 (2018a).
- [18] Escobar-Bach, M., Goegebeur, Y., Guillou, A.: Local estimation of the conditional stable tail dependence function. *Scand. J. Stat.* 45, 590–617 (2018b)
- [19] Fan, J., Gijbels, I: *Local Polynomial Modelling and Its Applications*. Chapman and Hall (1996).
- [20] Gardes, L.: Tail dimension reduction for extreme quantile estimation. *Extremes* 21, 57–95 (2018).
- [21] Goegebeur, Y., Guillou, A., Osmann, M.: A local moment type estimator for the extreme value index in regression with random covariates. *Canadian J. Statist.* 42, 487–507 (2014).
- [22] Goegebeur, Y., Guillou, A., Qin, J.: Bias-corrected estimation for conditional Pareto-type distributions with random right censoring. *Extremes* 22, 459–498 (2019)
- [23] de Haan, L., Ferreira, A.: *Extreme value theory. An introduction*. Springer (2006).

- [24] Joe, H., Li, H.: Tail risk of multivariate regular variation. *Methodol. Comput. Appl. Probab.* 13, 671–693 (2011).
- [25] Landsman, Z., Valdez, E.: Tail conditional expectations for elliptical distributions. *N. Am. Actuar. J.* 7, 55–71 (2003).
- [26] Mhalla, L., de Carvalho, M., Chavez-Demoulin, V.: Regression type models for extremal dependence. *Scand. J. Stat.* 46, 1141–1167 (2019).
- [27] Nadaraya, E. A.: On Estimating Regression. *Theory Probab. Appl.* 9, 141–142 (1964).
- [28] van der Vaart, A. W.: *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics 3, Cambridge University Press, Cambridge (1998).
- [29] van der Vaart, A. W., Wellner, J. A.: *Weak convergence and empirical processes, with applications to statistics*. Springer Series in Statistics. Springer-Verlag, New York (1996).
- [30] Wand, M., Jones, M.C.: *Kernel Smoothing*. Chapman and Hall (1995).
- [31] Watson, G. S.: Smooth Regression Analysis. *Sankhya: Ser. A* 26, 359–372 (1964).
- [32] Weissman, I.: Estimation of parameters and large quantiles based on the  $k$  largest observations. *J. Amer. Statist. Assoc.* 73, 812–815 (1978).
- [33] Wellner, J.A.: *Empirical Processes: Theory and Applications*. Special topics course. <https://www.stat.washington.edu/jaw/RESEARCH/TALKS/Delft/emp-proc-delft-big.pdf> (2005).
- [34] Wretman, J.: A simple derivation of the asymptotic distribution of a sample quantile. *Scand. J. Stat.* 5, 123–124 (1978).
- [35] Xu, W., Wang, H. J., Li, D.: Extreme quantile estimation based on the tail single-index model. *Statist. Sinica* (to appear).
- [36] Yao, Q.: Conditional predictive regions for stochastic processes. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.2449&rep=rep1&type=pdf> (1999).

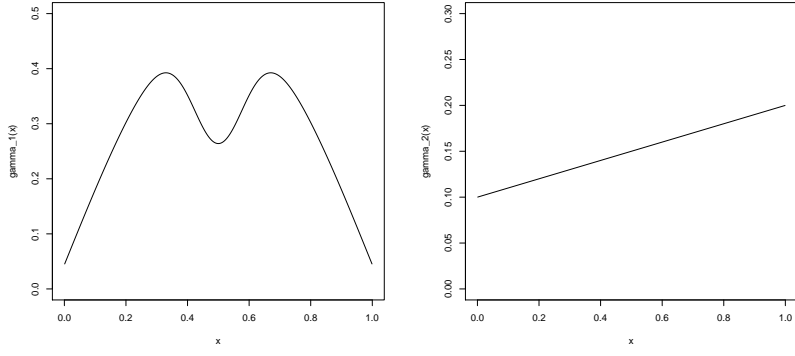


Figure 1: Model 2:  $\gamma_1(x)$  (left) and  $\gamma_2(x)$  (right) as a function of  $x$ .

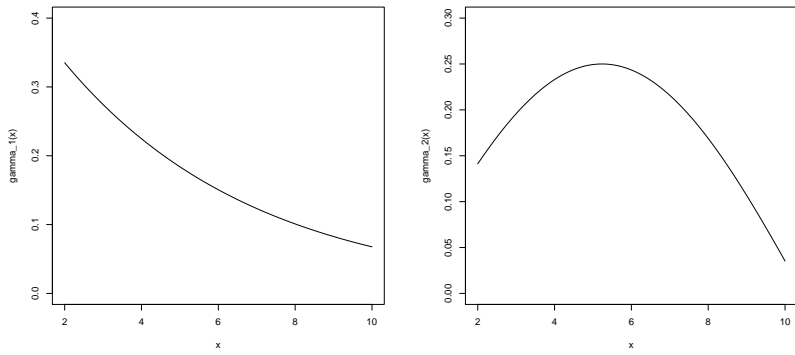


Figure 2: Model 3:  $\gamma_1(x)$  (left) and  $\gamma_2(x)$  (right) as a function of  $x$ .

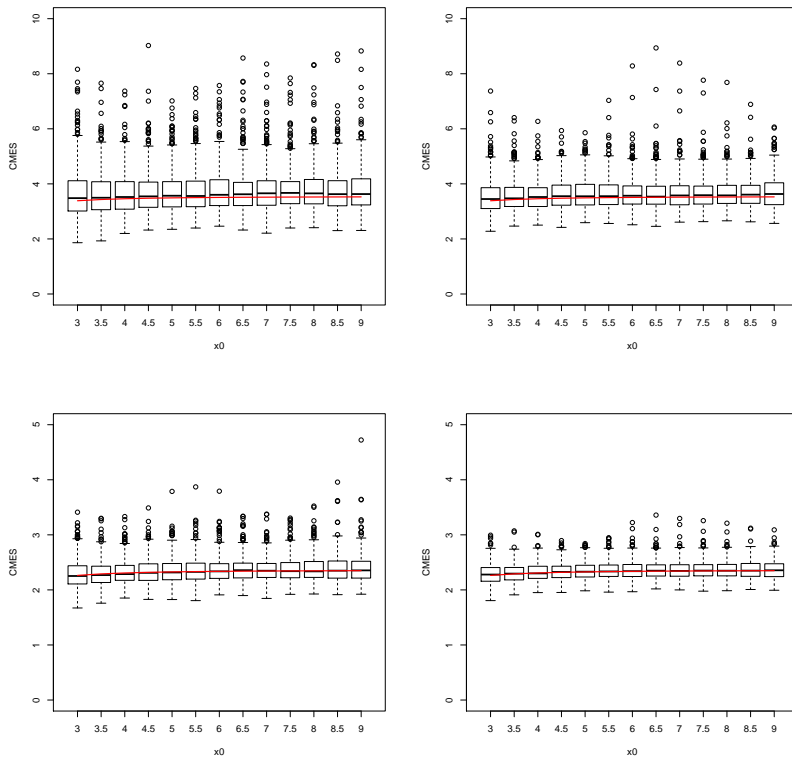


Figure 3: Model 1, boxplots of  $\bar{\theta}_n$  for 500 simulations of size  $n = 500$  (left) and  $n = 1000$  (right) for  $k/n = 2\%$  (top row) and  $k/n = 10\%$  (bottom row). The red curve shows the true value of  $\theta_{k/n}$ .

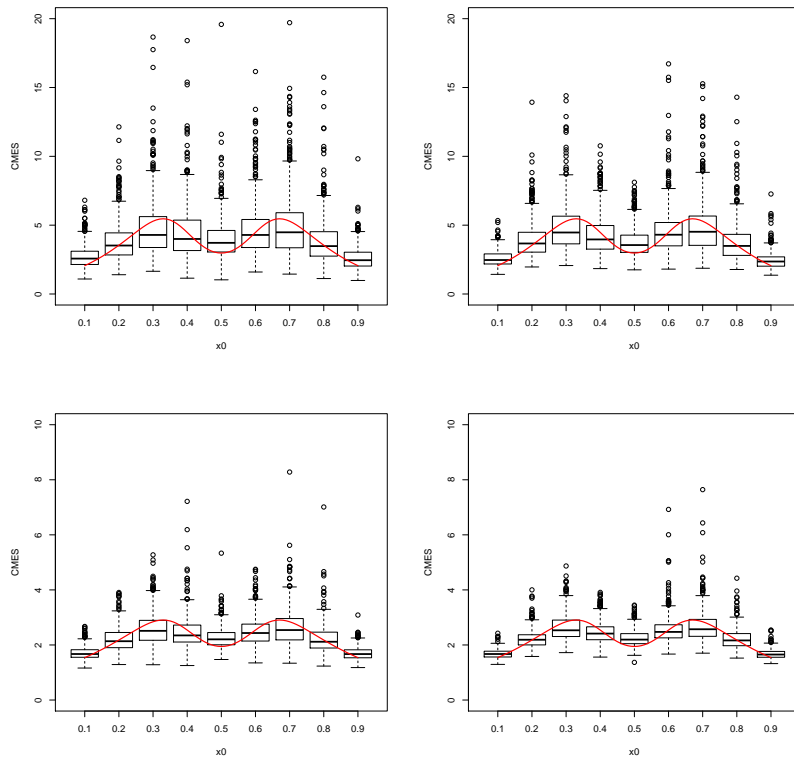


Figure 4: Model 2, boxplots of  $\bar{\theta}_n$  for 500 simulations of size  $n = 500$  (left) and  $n = 1000$  (right) for  $k/n = 2\%$  (top row) and  $k/n = 10\%$  (bottom row). The red curve shows the true value of  $\theta_{k/n}$ .

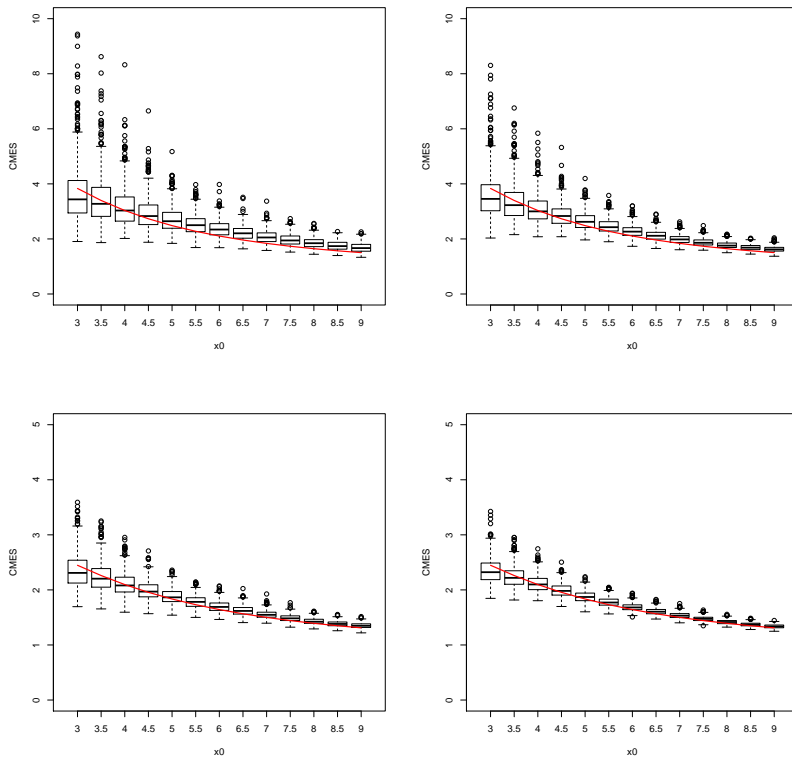


Figure 5: Model 3, boxplots of  $\bar{\theta}_n$  for 500 simulations of size  $n = 500$  (left) and  $n = 1000$  (right) for  $k/n = 2\%$  (top row) and  $k/n = 10\%$  (bottom row). The red curve shows the true value of  $\theta_{k/n}$ .



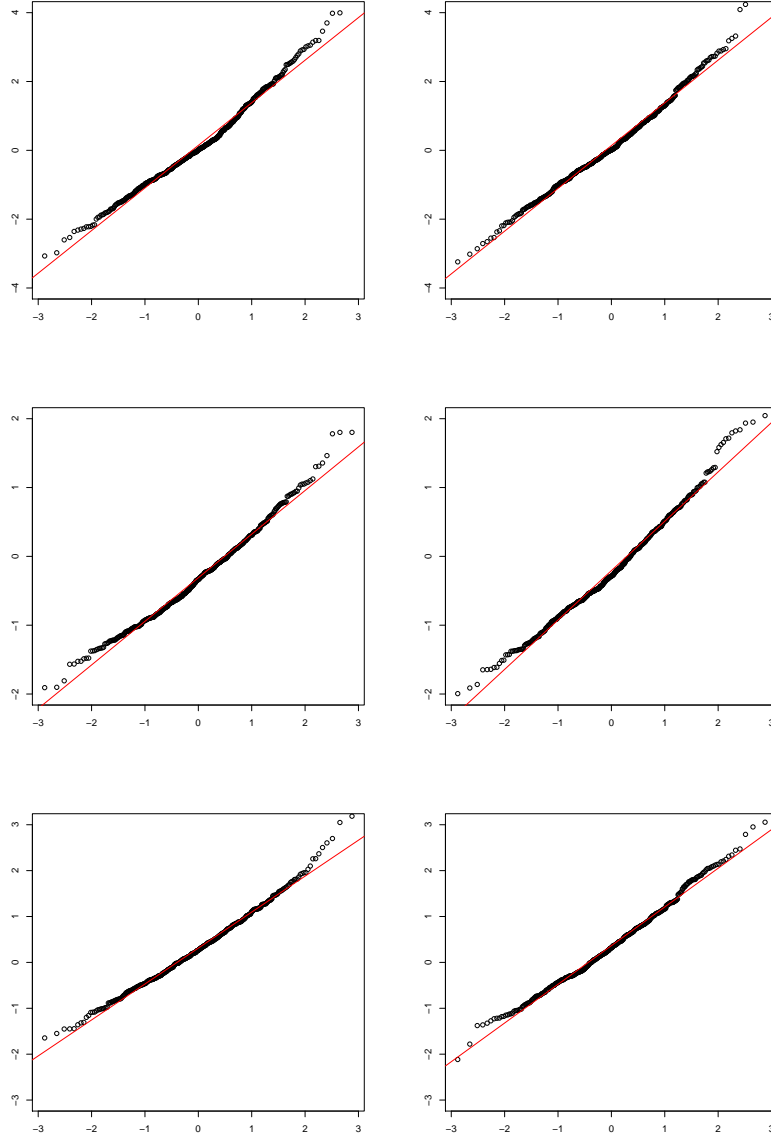


Figure 6: Normal quantile plots of  $\sqrt{kh} \ln \bar{\theta}_n / \theta_{k/n}$ . Top: Model 1,  $x_0 = 3$ , middle: Model 2,  $x_0 = 0.3$ , and bottom: Model 3,  $x_0 = 5$ . The quantile plots are constructed with  $k$  taken as 5% of  $n$ , with  $n = 500$  (left) and  $n = 1000$  (right).

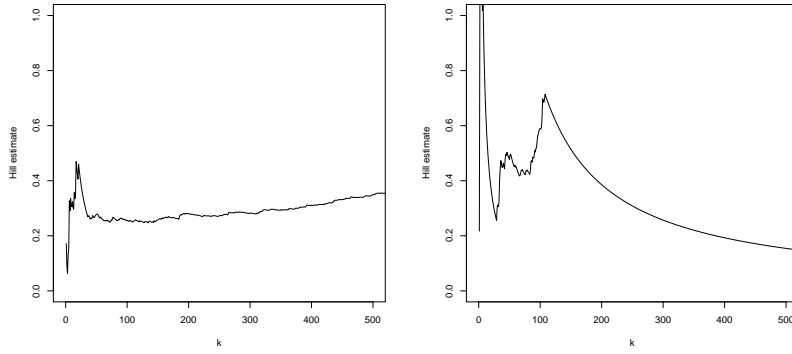


Figure 7: FEMA claim data: local Hill plots for  $Y^{(1)}$ , total claim amount (left), and  $Y^{(2)}$ , capital insured (right).

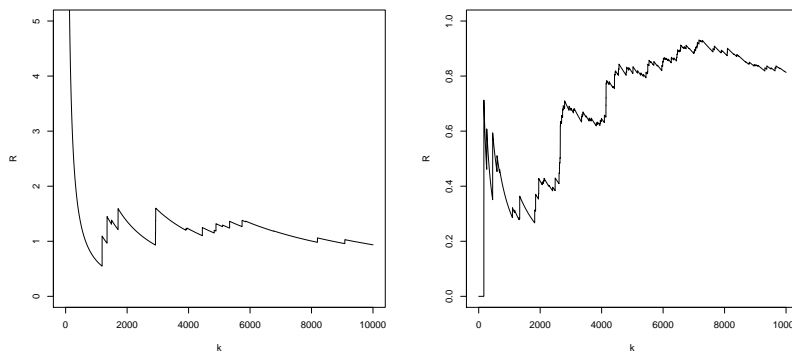


Figure 8: FEMA claim data:  $\hat{R}(1.5, 1.5|x_0)$  as a function of  $k$ , for location (latitude, longitude)=(33.84,-84.45) on 2009, July 1 (left), and 2017, January 1 (right).

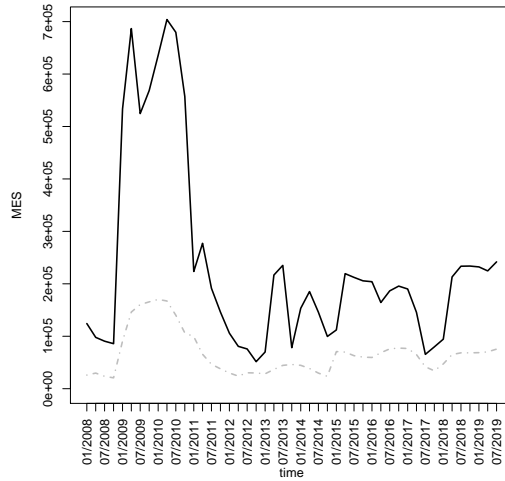


Figure 9: FEMA claim data:  $\bar{\theta}_n$  with  $k/n = 1\%$  (solid line) and  $k/n = 10\%$  (dashed-dotted line), as a function of time, for location (latitude, longitude)=(33.84,-84.45).

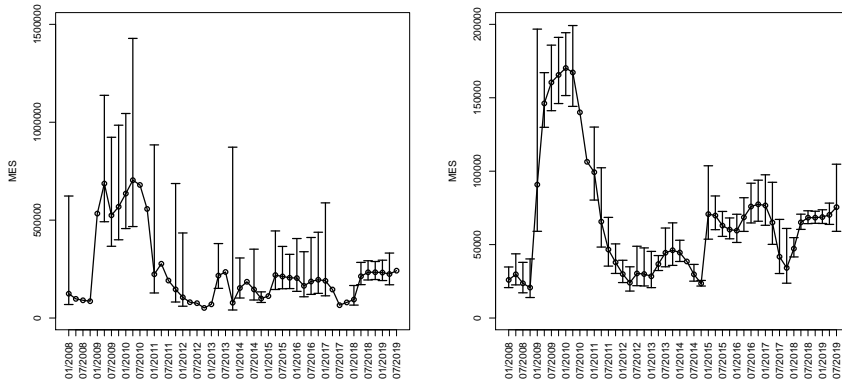


Figure 10: FEMA claim data:  $\bar{\theta}_n$  with  $k/n = 1\%$  (left) and  $k/n = 10\%$  (right), with pointwise 95% confidence intervals, as a function of time, for location (latitude, longitude)=(33.84,-84.45).