



HAL
open science

Extreme conditional quantile estimation for time series

Yuri Goegebeur, Armelle Guillou, Jing Qin

► **To cite this version:**

Yuri Goegebeur, Armelle Guillou, Jing Qin. Extreme conditional quantile estimation for time series. 2026. [⟨hal-04673935v3⟩](#)

HAL Id: hal-04673935

<https://hal.science/hal-04673935v3>

Preprint submitted on 23 Jan 2026

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.



HAL Authorization

Extreme conditional quantile estimation for time series

Yuri Goegebeur⁽¹⁾, Armelle Guillou⁽²⁾, Jing Qin⁽¹⁾

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

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

Abstract

We consider the estimation of an extreme conditional quantile $Q_Y(1-p|x_0)$ for a heavy-tailed distribution in case of a strictly stationary time series $(X_t, Y_t)_{t \in \mathbb{Z}}$. Here, $Q_Y(\cdot|x_0)$ denotes the conditional quantile function of Y given $X = x_0$ and $p \in (0, 1)$ is small, i.e., smaller than $1/n$, where n is the size of the sample on which the estimation is based. In a first step, we propose a conditional extreme value index estimator in that context of time series and we establish its weak convergence using the cluster method. Then, in a second step, by using a Weissman-type construction, we propose an extreme conditional quantile estimator for which we derive also the weak convergence. The performance of our estimators is illustrated on a simulation study and compared to alternative estimators already proposed in the literature. Finally, our methodology is applied on a real dataset of log-returns of the S&P 500 index and market volatility.

Keywords: Time series; Extreme value theory; Extreme conditional quantile; Empirical process.

MSC codes: 62M10; 62G32; 62G20.

1 Introduction

We consider the estimation of an extreme quantile of a heavy-tailed distribution when a random covariate $X \in \mathbb{R}^d$ is recorded simultaneously with the variable of interest $Y \in \mathbb{R}_+$. We denote

by $F_Y(\cdot|x)$ the conditional distribution function of the response variable Y given $X = x$, and by g_X the density function of X . We assume that the conditional response distribution is of Pareto-type, i.e., there exists a positive function $\gamma(x)$ such that $\bar{F}_Y(\cdot|x) := 1 - F_Y(\cdot|x)$ is of the form

$$\bar{F}_Y(y|x) = y^{-\frac{1}{\gamma(x)}} \ell_Y(y|x), \quad y > 0, \quad (1)$$

where $\ell_Y(\cdot|x)$ is a slowly varying function at infinity, i.e., for all $\lambda > 0$

$$\lim_{y \rightarrow \infty} \frac{\ell_Y(\lambda y|x)}{\ell_Y(y|x)} = 1.$$

The tail heaviness of $\bar{F}_Y(y|x)$ is governed by the tail function $\gamma(x)$, where larger values correspond with heavier tails. This Pareto-type model finds many important practical applications, among others, in finance, insurance, telecommunications, ... Since $\gamma(x)$ describes the tail heaviness of the conditional response distribution, it needs to be adequately estimated from the available data. This topic has been extensively studied in the literature in case of an independent and identically distributed (i.i.d.) sample of pairs (X_i, Y_i) , $i = 1, \dots, n$, see, e.g., Wang and Tsai (2009), Daouia et al. (2011), Wang and Li (2013) and Goegebeur et al. (2014), among others. In the case of time series, as far as we know, only the very recent paper of Daouia et al. (2023) deals with this problem by adapting an estimator for $\gamma(x)$ originally proposed by Daouia et al. (2011) in the i.i.d. context and based on a fixed number J of high quantiles.

In our paper, we consider a strictly stationary time series $(X_t, Y_t)_{t \in \mathbb{Z}}$ with conditional distribution function of Y given $X = x$ of Pareto-type. We denote by $U_Y(\cdot|x)$ the associated tail quantile function defined as $U_Y(\cdot|x) = \inf\{y : F_Y(y|x) \geq 1 - 1/\cdot\}$ and our estimation will be done at $X = x_0$, a reference position in the interior of the support of g_X , which is assumed to be non-empty. In Section 2, we start by proposing a conditional extreme value index estimator for time series, and then we seek to estimate an extreme conditional quantile using a Weissman-type construction (see Weissman, 1978). A list of the required assumptions in order to establish the weak convergence of these estimators is also given. Section 3 states these results, which are then validated on a simulation study in Section 4. Finally, our methodology is illustrated in Section 5 on a real dataset of financial returns of the S&P 500 index and market volatility. A

concluding section summarizes the main contributions of the paper. This paper ends with an appendix containing first, the verification of our main conditions of our theorems on the model considered in Section 4, and then, all the proofs of the results.

2 Methodology and assumptions

Let $\|\cdot\|$ be the Euclidean norm on \mathbb{R}^d and $\mathcal{B}_x(r)$ be the $\|\cdot\|$ -ball with center x and radius $r > 0$. Throughout the paper, we will denote by $\underline{0} \in \mathbb{R}^d$, respectively $\underline{2} \in \mathbb{R}^d$, the d -vector whose components are all equal to 0, respectively 2. Hereafter, we will use a kernel type method for our estimations and for this, we consider a kernel K satisfying the following classical hypothesis in the non-parametric framework:

Assumption (\mathcal{K}). *The kernel K is a bounded probability density with support $\mathcal{B}_{\underline{0}}(1)$.*

Our conditional extreme value index estimator is defined as

$$\hat{\gamma}_k(x_0) = \frac{\frac{1}{k h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \left(\log Y_i - \log \hat{U}_Y(n/k|x_0)\right) \mathbb{1}_{\{Y_i > \hat{U}_Y(n/k|x_0)\}}}{\hat{g}_n(x_0)},$$

where $k = k_n$ is an intermediate sequence, i.e., $k \rightarrow \infty$ as $n \rightarrow \infty$ with $k/n \rightarrow 0$, h_n is a positive non-random sequence of bandwidths with $h_n \rightarrow 0$ if $n \rightarrow \infty$, $\hat{g}_n(x_0)$ is the kernel density estimator for g_X defined as

$$\hat{g}_n(x_0) = \frac{1}{n h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right),$$

and $\hat{U}_Y(n/k|x_0)$ is an estimator for $U_Y(n/k|x_0)$. Note that the threshold $\hat{U}_Y(n/k|x_0)$ is local, i.e., it depends on the point x_0 where the estimation is to be made. This estimator $\hat{\gamma}_k(x_0)$ can be viewed as a direct adaptation of the classical Hill estimator (see Hill, 1975) to the regression context.

Although the estimation of a conditional extreme value index is of interest since it gives information on the upper tail heaviness of the distribution, for practical applications it is much more important to estimate an extreme conditional quantile, i.e., $U_Y(1/p|x_0)$, where $p \leq 1/n$, with n the size of the sample on which the estimation is based. Such an estimation requires

extrapolation beyond the observed values of the time series and can be done by a Weissman-type construction based on the approximation

$$U_Y\left(\frac{1}{p}\middle|x_0\right) = U_Y\left(\frac{n}{k}\middle|x_0\right) \frac{U_Y\left(\frac{1}{p}\middle|x_0\right)}{U_Y\left(\frac{n}{k}\middle|x_0\right)} \approx U_Y\left(\frac{n}{k}\middle|x_0\right) \left(\frac{k}{np}\right)^{\gamma(x_0)},$$

motivated by model (1), and from which we deduce that

$$\widehat{U}_Y\left(\frac{1}{p}\middle|x_0\right) = \widehat{U}_Y\left(\frac{n}{k}\middle|x_0\right) \left(\frac{k}{np}\right)^{\widehat{\gamma}_k(x_0)}$$

is an extreme conditional quantile estimator for $U_Y(1/p|x_0)$.

Our first aim is to study the limiting distribution of $\widehat{\gamma}_k(x_0)$ correctly normalized in the time series context, and then to deduce a similar result for $\widehat{U}_Y(1/p|x_0)$. To reach these goals, we propose to use the cluster method of Drees and Rootzén (2010). We define

$$\begin{aligned} Z_{n,i} &= (Z_{n,i,1}, Z_{n,i,2}) = \left(\underline{2} + \frac{x_0 - X_i}{h_n}, \frac{Y_i}{U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{x_0 - X_i}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_i}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}}, \\ \phi_s(z) &= \phi_s(z_1, z_2) = K(z_1 - \underline{2}) \log \frac{z_2}{s} \mathbb{1}_{\{z_2 > s\}}, \\ \psi_s(z) &= \psi_s(z_1, z_2) = K(z_1 - \underline{2}) \mathbb{1}_{\{z_2 > s\}}, \end{aligned}$$

for $s \in [1 - \varepsilon, 1 + \varepsilon]$, $\varepsilon \in (0, 1)$. Here the vector $\underline{2}$ is useful to ensure $Z_{n,i,1} > 0$ for $(x_0 - X_i)/h_n \in \mathcal{B}_{\underline{0}}(1)$, whereas ε allows us to handle the randomness of the threshold $\widehat{U}_Y(n/k|x_0)$. As for the functions ϕ_s and ψ_s , they were introduced to establish the limiting distribution of $\widehat{\gamma}_k(x_0)$ and $\widehat{U}_Y(n/k|x_0)$, respectively.

Let $v_n := \mathbb{P}(Z_{n,1} \neq (0, 0))$. Drees and Rootzén (2010) provided conditions according to which

$$Z_n(H) := \frac{1}{\sqrt{nv_n}} \sum_{i=1}^n \{H(Z_{n,i}) - \mathbb{E}[H(Z_{n,i})]\},$$

converges weakly towards a centered Gaussian process. We will use this result in our framework with $H \in \Phi := \{\phi_s, \psi_s : s \in [1 - \varepsilon, 1 + \varepsilon]\}$.

Let RV_α denote the class of regularly varying functions at infinity with index α , i.e., positive measurable functions f satisfying $f(tx)/f(t) \rightarrow x^\alpha$ as $t \rightarrow \infty$, for all $x > 0$. To establish our

convergence, as usual in extreme value theory, we need to precise the structure of the conditional response distribution given in (1) by imposing a second-order condition. The one used in this paper is the following:

Assumption (D). *The conditional survival function of Y given $X = x$ satisfies*

$$\bar{F}_Y(y|x) = A(x) y^{-1/\gamma(x)} \left(1 + \frac{1}{\gamma(x)} \delta_Y(y|x) \right),$$

where $A(x) > 0$, $\gamma(x) > 0$, and $|\delta_Y(\cdot|x)|$ is normalized regularly varying at infinity with index $-\beta(x)$, $\beta(x) > 0$, i.e.,

$$\delta_Y(y|x) = B(x) \exp \left\{ \int_1^y \frac{\xi_Y(u|x)}{u} du \right\},$$

with $B(x) \in \mathbb{R}$ and $\xi_Y(y|x) \rightarrow -\beta(x)$ as $y \rightarrow \infty$.

Under Assumption (D), $U_Y(\cdot|x)$ satisfies

$$U_Y(y|x) = [A(x)]^{\gamma(x)} y^{\gamma(x)} (1 + a_Y(y|x)),$$

where $a_Y(y|x) := \delta_Y(U_Y(y|x)|x)(1 + o(1))$, and thus $|a_Y(\cdot|x)| \in RV_{-\beta(x)\gamma(x)}$. This Assumption (D) is very general and satisfied by usual distributions among them, Fréchet, student t, F, generalized Pareto and Burr distributions.

Additionally, to deal with the regression context, g_X and the functions appearing in $F_Y(\cdot|x)$ are assumed to satisfy the following Hölder conditions.

Assumption (H). *There exist positive constants $M_g, M_A, M_\gamma, M_B, M_\xi, \eta_g, \eta_A, \eta_\gamma, \eta_B$ and η_ξ , such that for all $x, z \in \mathcal{B}_{x_0}(h)$ for some $h > 0$:*

$$\begin{aligned} |g_X(x) - g_X(z)| &\leq M_g \|x - z\|^{\eta_g}, \\ |A(x) - A(z)| &\leq M_A \|x - z\|^{\eta_A}, \\ |\gamma(x) - \gamma(z)| &\leq M_\gamma \|x - z\|^{\eta_\gamma}, \\ |B(x) - B(z)| &\leq M_B \|x - z\|^{\eta_B}, \\ \sup_{y \geq 1} |\xi_Y(y|x) - \xi_Y(y|z)| &\leq M_\xi \|x - z\|^{\eta_\xi}. \end{aligned}$$

Now, we need to assume a condition according which a joint conditional extreme value of (Y_0, Y_i) cannot be much more likely than a marginal conditional extreme value of Y_0 , uniformly across time and locally uniformly across the covariate space.

Assumption (\mathcal{J}). *There exist $\bar{h}, \bar{y} > 0$ such that*

$$\Omega_{\bar{h}}(\bar{y}|x_0) := \sup_{i \geq 1} \sup_{y_0, y_1 \geq \bar{y}} \sup_{u_0, u_1 \in \mathcal{B}_{x_0}(\bar{h})} \frac{\mathbb{P}(Y_0 > y_0, Y_i > y_1 | X_0 = u_0, X_i = u_1)}{\sqrt{\bar{F}_Y(y_0|u_0)\bar{F}_Y(y_1|u_1)}} < \infty.$$

Also, to ensure that similar values of the covariate cannot occur too often at neighboring time points, we impose the following condition:

Assumption (\mathcal{X}). *There exists an integer $i_0 \geq 1$ such that*

$$\begin{aligned} 1 \leq i < i_0 &\implies \lim_{r \rightarrow 0} r^{-d} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(r), X_i \in \mathcal{B}_{x_0}(r)) = 0, \\ \text{and} &\limsup_{r \rightarrow 0} \sup_{i \geq i_0} r^{-2d} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(r), X_i \in \mathcal{B}_{x_0}(r)) < \infty. \end{aligned}$$

This Assumption (\mathcal{X}) is in particular satisfied if the pair (X_0, X_i) has a density uniformly bounded for all $i \geq 1$ on $\mathcal{B}_{x_0}(r) \times \mathcal{B}_{x_0}(r)$ for some $r > 0$. Note also that Assumption (\mathcal{J}) as well as Assumption (\mathcal{X}) were already introduced in Daouia et al. (2023). We refer to the Supplementary Material of the latter paper for additional comments related to these hypotheses, e.g., examples of models satisfying them.

Now, we come back to our initial aim, namely to show the weak convergence of $Z_n(H)$ towards a Gaussian process in the space of bounded functions indexed by Φ , denoted $\ell^\infty(\Phi)$. Thus we start to compute v_n in our context. Note that

$$\begin{aligned} v_n &= \mathbb{P}\left(\frac{x_0 - X_1}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_1}{U_Y(n/k|x_0)} > 1 - \varepsilon\right) \\ &= \mathbb{E}\left[\mathbb{1}_{\{\frac{x_0 - X}{h_n} \in \mathcal{B}_{\underline{0}}(1)\}} \bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|X)\right] \\ &= \bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0) \left\{ \int_{\mathbb{R}^d} \mathbb{1}_{\{\frac{x_0 - x}{h_n} \in \mathcal{B}_{\underline{0}}(1)\}} g_X(x) dx \right. \\ &\quad \left. + \int_{\mathbb{R}^d} \mathbb{1}_{\{\frac{x_0 - x}{h_n} \in \mathcal{B}_{\underline{0}}(1)\}} \left[\frac{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x)}{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0)} - 1 \right] g_X(x) dx \right\}. \end{aligned}$$

Assuming that $h_n^{\eta\gamma \wedge \eta\varepsilon} \log(n/k) \rightarrow 0$ and $g_X(x_0) > 0$, using Lemma 7.1, we have for n large

$$\begin{aligned} v_n &= (1 - \varepsilon)^{-\frac{1}{\gamma(x_0)}} \frac{k}{n} \int_{\mathbb{R}^d} \mathbb{1}_{\{\frac{x_0-x}{h_n} \in \mathcal{B}_0(1)\}} g_X(x) dx (1 + o(1)) \\ &= (1 - \varepsilon)^{-\frac{1}{\gamma(x_0)}} \frac{k h_n^d}{n} \left\{ V_d g_X(x_0) + \int_{\mathcal{B}_0(1)} [g_X(x_0 - h_n v) - g_X(x_0)] dv \right\} (1 + o(1)) \\ &= (1 - \varepsilon)^{-\frac{1}{\gamma(x_0)}} \frac{k h_n^d}{n} V_d g_X(x_0) (1 + o(1)), \end{aligned}$$

where V_d is the volume of $\mathcal{B}_0(1)$. We have clearly $v_n \rightarrow 0$ as $n \rightarrow \infty$, which means that the theory of Drees and Rootzén (2010) can be applied in our context. Here the strength of the temporal dependence between the random vectors $Z_{n,1}, \dots, Z_{n,n}$ will be evaluated through the β -mixing coefficient

$$\beta_{n,k} := \sup_{1 \leq \ell \leq n-k-1} \mathbb{E} \left\{ \sup_{B \in \mathcal{B}_{n,\ell+k+1}^n} \left| \mathbb{P}(B | \mathcal{B}_{n,1}^\ell) - \mathbb{P}(B) \right| \right\},$$

where $\mathcal{B}_{n,i}^j$ denotes the σ -field generated by $(Z_{n,\ell})_{i \leq \ell \leq j}$. We assume that there exist sequences of integers $(\ell_n)_{n \in \mathbb{N}}$ and $(r_n)_{n \in \mathbb{N}}$, constituting the basic elements of the so-called "big blocks, small blocks" technique, such that $\ell_n, r_n \rightarrow \infty$ and the following holds:

Assumption (A) As $n \rightarrow \infty$, we have $\ell_n = o(r_n)$, $r_n = o(\sqrt{nv_n})$, $r_n h_n^{d/(1+\delta)} = O(1)$ for some $\delta > 0$, and $\beta_{n,\ell_n}(n/r_n) \rightarrow 0$.

Assumption (A) imposes restrictions on the rate at which k tends to infinity and h_n to zero. Sufficient conditions to ensure that a Markov chain is β -mixing can be found in Doukhan (1994, Section 2.4). More specifically, most of the usual time series are geometrically β -mixing under mild conditions, e.g., ARMA, ARCH or GARCH time series. In that case $\beta_{n,k} = O(\eta^k)$ for some $\eta \in (0, 1)$. Then, by choosing $\ell_n = C \log n$ for C sufficiently large, Assumption (A) is fulfilled for a suitably chosen r_n if $(\log n)^2 = o(k h_n^d)$ and $h_n^d = o((\log n)^{-(1+\delta)})$.

3 Main results

The first theorem states the weak convergence, denoted by \rightsquigarrow , of the empirical process Z_n indexed by Φ . This result will allow us to obtain the joint convergence of $\hat{\gamma}_k(x_0)$ and $\hat{U}_Y(n/k|x_0)$,

properly normalized, which are the random quantities appearing in the estimator $\widehat{U}_Y(1/p|x_0)$, though they are also of interest on their own.

Theorem 3.1 *Let $(X_t, Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series, with a conditional distribution function of the response variable Y given $X = x$ satisfying Assumption (\mathcal{D}) . Let k be an intermediate sequence such that $k/n \rightarrow 0$ and $kh_n^d \rightarrow \infty$ as $n \rightarrow \infty$. Consider a bandwidth h_n such that $h_n^{\eta_\gamma \wedge \eta_\xi} \log(n/k) \rightarrow 0$ as $n \rightarrow \infty$. Suppose that Assumptions $(\mathcal{K}), (\mathcal{H}), (\mathcal{J}), (\mathcal{X})$ and (\mathcal{A}) are fulfilled. Then, for x_0 such that $g_X(x_0) > 0$ and for any $\varepsilon \in (0, 1)$, we have,*

$$Z_n \rightsquigarrow Z_\varepsilon,$$

in $\ell^\infty(\Phi)$, where Z_ε is a centered Gaussian process with covariance function given by

$$\text{Cov}(Z_\varepsilon(H_1), Z_\varepsilon(H_2)) = \mathbb{E}[H_1(U + \underline{2}, V) H_2(U + \underline{2}, V)],$$

where U and V are independent, U from a Uniform distribution on $\mathcal{B}_0(1)$ and V from a Pareto distribution with survival distribution function given by $\mathbb{P}(V > t) = (t/(1 - \varepsilon))^{-1/\gamma(x_0)} \mathbb{1}_{\{t > 1 - \varepsilon\}} + \mathbb{1}_{\{t \leq 1 - \varepsilon\}}$.

Unlike recent contributions on extremes in time series without covariates where the covariance of the limiting process is given by a series (see, e.g., Drees et al., 2015, Drees and Knežević, 2020 and Goegebeur et al., 2025), the covariance function of the limiting Gaussian process is here relatively simple. This is due to the Assumptions (\mathcal{J}) and (\mathcal{X}) , which control the occurrence of joint conditional extremes of (Y_0, Y_i) and of similar values of the covariate at neighboring time points, respectively.

As an intermediate result, we need now to define our estimator $\widehat{U}_Y(n/k|x_0)$ and to derive its weak convergence in terms of the Z_ε - process. Consider a kernel estimator of $\overline{F}_Y(y|x_0)$ defined as

$$\widehat{\overline{F}}_Y(y|x_0) = \frac{\frac{1}{nh_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\{Y_i > y\}}}{\widehat{g}_n(x_0)},$$

and let

$$\widehat{U}_Y\left(\frac{n}{k}|x_0\right) = \inf \left\{ y : \widehat{\overline{F}}_Y(y|x_0) \leq \frac{k}{n} \right\}.$$

In order to obtain the limiting distribution of $\widehat{U}_Y(n/k|x_0)$, properly normalized, we will use Theorem 3.1. To this aim, we rewrite $\widehat{U}_Y(n/k|x_0)$ as

$$\widehat{U}_Y\left(\frac{n}{k}|x_0\right) = \inf \left\{ y : -\frac{\frac{1}{k h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > \frac{y}{U_Y(n/k|x_0)}\right\}}}{\widehat{g}_n(x_0)} \geq -1 \right\},$$

and we introduce

$$\check{T}_n(s) := -\frac{\frac{1}{k h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}}{\widehat{g}_n(x_0)}.$$

Note that

$$\check{T}_n^{\leftarrow}(-1) = \frac{\widehat{U}_Y(n/k|x_0)}{U_Y(n/k|x_0)},$$

where the superscript ‘ \leftarrow ’ denotes the generalized inverse of a non-decreasing function. From Theorem 3.1 we get

$$\sqrt{k h_n^d} \left\{ \check{T}_n(s) + s^{-\frac{1}{\gamma(x_0)}} \right\} \rightsquigarrow -\sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_s). \quad (2)$$

Then, by Vervaat’s Lemma (see, e.g., Lemma A.0.2 in de Haan and Ferreira, 2006), we can invert the weak convergence (2) and obtain the result for $\check{T}_n^{\leftarrow}(-1)$.

The above is formalized in the next theorem.

Theorem 3.2 *Let $(X_t, Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series, with a conditional distribution function of the response variable Y given $X = x$ satisfying Assumption (D). Let k be an intermediate sequence such that $k/n \rightarrow 0$ and $k h_n^d \rightarrow \infty$ as $n \rightarrow \infty$. Consider a bandwidth h_n such that $h_n^{\eta_\xi} \log(n/k) \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_A \wedge \eta_g} \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_\gamma} \log(n/k) \rightarrow 0$ and $\sqrt{k h_n^d} |\delta_Y(U_Y(n/k|x_0)|x_0)| \rightarrow 0$ as $n \rightarrow \infty$. Suppose that Assumptions (K), (H), (J), (X) and (A) are fulfilled. Then, for x_0 such that $g_X(x_0) > 0$, we have*

$$\sqrt{k h_n^d} \left\{ \frac{\widehat{U}_Y(n/k|x_0)}{U_Y(n/k|x_0)} - 1 \right\} \rightsquigarrow \gamma(x_0) \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_1).$$

The variance of the limiting distribution of $\widehat{U}_Y(n/k|x_0)$ correctly normalized is $\gamma^2(x_0) \|K\|_2^2 / g_X(x_0)$, which is the same as the variance in the regression context for i.i.d. random variables (see

Lemma 5.1 in Goegebeur et al., 2023). As a further comparison, in the i.i.d. context of Pareto-type distributions without covariates, one has for $\sqrt{k}\{\widehat{U}_Y(n/k)/U_Y(n/k) - 1\}$ an asymptotic variance of γ^2 (see Theorem 2.4.1 in de Haan and Ferreira, 2006).

We have now all the ingredients to state the weak convergence of $\widehat{\gamma}_k(x_0)$ correctly normalized.

Theorem 3.3 *Let $(X_t, Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series, with a conditional distribution function of the response variable Y given $X = x$ satisfying Assumption (D). Let k be an intermediate sequence such that $k/n \rightarrow 0$ and $kh_n^d \rightarrow \infty$ as $n \rightarrow \infty$. Consider a bandwidth h_n such that $h_n^{\eta_\kappa} \log(n/k) \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_A \wedge \eta_g} \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_\gamma} \log(n/k) \rightarrow 0$ and $\sqrt{k h_n^d} |\delta_Y(U_Y(n/k|x_0)|x_0)| \rightarrow 0$ as $n \rightarrow \infty$. Suppose that Assumptions (K), (H), (J), (X) and (A) are fulfilled. Then, for x_0 such that $g_X(x_0) > 0$, we have*

$$\sqrt{k h_n^d} \{\widehat{\gamma}_k(x_0) - \gamma(x_0)\} \rightsquigarrow \sqrt{\frac{V_d}{g_X(x_0) (1 - \varepsilon)^{\frac{1}{\gamma(x_0)}}}} \{Z_\varepsilon(\phi_1) - \gamma(x_0) Z_\varepsilon(\psi_1)\}.$$

Note that the variance of the limiting distribution of $\widehat{\gamma}_k(x_0)$ correctly normalized is $\gamma^2(x_0) \|K\|_2^2 / g_X(x_0)$, which is the same as the variance in the regression context for i.i.d. random variables (see Theorem 2.1 in Goegebeur et al., 2023), while in the univariate i.i.d. framework, the asymptotic variance of the classical Hill estimator is γ^2 . Daouia et al. (2023) introduced, also in a time series context, an alternative estimator for $\gamma(x_0)$, which is in our notation given by

$$\widehat{\gamma}_{1-\frac{k}{n}}^{(J)}(x_0) = \frac{1}{\log(J!)} \sum_{j=1}^J \log \frac{\widehat{U}_Y(\frac{nj}{k}|x_0)}{\widehat{U}_Y(\frac{n}{k}|x_0)},$$

where J is fixed. Note that this estimator is based on a finite number of intermediate conditional quantiles, as was also the case in, e.g., Daouia et al. (2011) for i.i.d. observations. Extending this result to the case where J can tend to infinity with n is not trivial, as it requires a result on the empirical conditional tail quantile process. On the other hand, our approach is based on log-excesses over only a single intermediate conditional quantile, and by using empirical process techniques, we can handle to let $k \rightarrow \infty$ as $n \rightarrow \infty$. Also, our estimator has a smaller asymptotic variance, since for Daouia et al. (2023) one has $J(J-1)(2J-1)\gamma^2(x_0)\|K\|_2^2/\{6[\log(J!)]^2 g_X(x_0)\}$ as asymptotic variance, which is minimal for

$J = 9$ with a corresponding value close to $1.24 \gamma^2(x_0) \|K\|_2^2 / g_X(x_0)$.

Finally, we can give the main theorem of this paper, namely the weak convergence of $\widehat{U}_Y(1/p|x_0)$ correctly normalized.

Theorem 3.4 *Let $(X_t, Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series, with a conditional distribution function of the response variable Y given $X = x$ satisfying Assumption (\mathcal{D}) . Let k be an intermediate sequence such that $k/n \rightarrow 0$ and $kh_n^d \rightarrow \infty$ as $n \rightarrow \infty$. Consider a bandwidth h_n such that $h_n^{\eta_\xi} \log(n/k) \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_A \wedge \eta_g} \rightarrow 0$, $\sqrt{k h_n^d} h_n^{\eta_\gamma} \log(n/k) \rightarrow 0$ and $\sqrt{k h_n^d} |\delta_Y(U_Y(n/k|x_0)|x_0)| \rightarrow 0$ as $n \rightarrow \infty$. Suppose that Assumptions (\mathcal{K}) , (\mathcal{H}) , (\mathcal{J}) , (\mathcal{X}) and (\mathcal{A}) are fulfilled. Consider now a sequence $p = p_n$ such that $k/(np) \rightarrow \infty$ and $\log(k/(np))/\sqrt{k h_n^d} \rightarrow 0$. Then, for x_0 such that $g_X(x_0) > 0$, we have*

$$\frac{\sqrt{k h_n^d}}{\log \frac{k}{np}} \left\{ \frac{\widehat{U}_Y\left(\frac{1}{p}|x_0\right)}{U_Y\left(\frac{1}{p}|x_0\right)} - 1 \right\} \rightsquigarrow \sqrt{\frac{V_d}{g_X(x_0) (1 - \varepsilon)^{\frac{1}{\gamma(x_0)}}}} \{Z_\varepsilon(\phi_1) - \gamma(x_0) Z_\varepsilon(\psi_1)\}.$$

By comparing the result of Theorems 3.3 and 3.4 we see that $\widehat{U}_Y(1/p|x_0)$ inherits the limiting distribution of the estimator $\widehat{\gamma}_k(x_0)$. This was also the case in Daouia et al. (2023), though, as mentioned above, their estimator for $\gamma(x_0)$ is different. In fact, from a closer look at the proof of Theorem 3.4 we see that $\widehat{U}_Y(1/p|x_0)$ will always inherit the limiting distribution of the used estimator for $\gamma(x_0)$, say $\check{\gamma}_k(x_0)$, if it satisfies $\sqrt{k h_n^d} (\check{\gamma}_k(x_0) - \gamma(x_0)) \rightsquigarrow \Gamma$. This was also observed in the i.i.d. context with covariate (Daouia et al., 2011) as well as in the i.i.d. context without covariate (see Theorem 4.3.8 in de Haan and Ferreira, 2006).

4 Simulation study

In this section we evaluate the finite sample performance of our proposed estimators with a simulation experiment. We simulate $N = 500$ samples of size $n = 2000$ from a conditional distribution of Y given $X = x$ which admits the following Burr($\lambda(x), \tau(x)$) distribution function

$$F_Y(y|x) = 1 - \left(1 + y^{\tau(x)}\right)^{-\lambda(x)}, \quad y > 0. \quad (3)$$

This model clearly satisfies Assumption (\mathcal{J}) , see Appendix A, as well as Assumption (\mathcal{D}) with $\gamma(x) = 1/(\lambda(x) \tau(x))$ and $\beta(x) = \tau(x)$. With the usual notation for the second order parameter,

this means that $\rho(x) = -\beta(x) \gamma(x) = -1/\lambda(x)$. In our simulations we use $\lambda(x) = \lambda \in \{0.5, 1, 2\}$, yielding $\rho(x) \in \{-2, -1, -0.5\}$ and

$$\tau(x) = \frac{2}{\lambda(x)} \left\{ (0.1 + \sin(\pi x)) \times \left(1.1 - 0.5 e^{-64(x-0.5)^2} \right) \right\}^{-1}.$$

This function for $\tau(x)$ was also used in the simulation experiments of Daouia et al. (2011), Goegebeur et al. (2014) and Dierckx et al. (2014). The time series $(X_t)_{t \in \mathbb{Z}}$ is generated from a Markovian model with logistic copula function, given by

$$C_\theta(u_1, u_2) = \exp \left(- \left[(-\log u_1)^\theta + (-\log u_2)^\theta \right]^{1/\theta} \right),$$

where $\theta \in [1, \infty)$, so the marginal distribution of X_t is uniform $(0, 1)$, and the dependence between (X_t, X_{t+1}) is governed by this copula function. See, e.g., Drees et al. (2015) and Drees and Knežević (2020) for more details on Markovian time series models. Note also that Assumptions (\mathcal{X}) and (\mathcal{A}) of our main theorems are satisfied, see Appendix A. We take $\theta = 2$ and 8.

We implement our estimators with the bi-quadratic kernel function

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

which satisfies Assumption (\mathcal{K}) . Note that, as is well known in nonparametric local estimation, the choice of the kernel function has not much impact on the results, see, e.g., Section 10.2.2 in Givens and Hoeting (2005).

Concerning the bandwidth parameter h_n , we follow Daouia et al. (2023), where h_n is selected globally based on minimizing the mean integrated squared error of a kernel density estimator for g_X .

For an automated selection of the good k -range for estimation of $\gamma(x_0)$, we propose the following data-driven method, as introduced in Goegebeur et al. (2019):

1. Compute $\hat{\gamma}_k(x_0)$ for $k = 5, \dots, k_{\max}$,
2. compute the standard deviation of the $\hat{\gamma}_k(x_0)$ estimates in a moving block of 50 successive k -values,

3. select the block with the smallest standard deviation,
4. within the block selected in step 3, repeat the steps 2 and 3, now with block size 20 and thereafter block size 10,
5. the estimate for $\gamma_k(x_0)$ is the median of the $\hat{\gamma}_k(x_0)$ in the finally selected block.

In the above, $k_{\max} := \min(2nh_n, m)$, where $m = 800$ for $\lambda = 0.5$, 500 for $\lambda = 1$ and 200 for $\lambda = 2$. The parameter m avoids that in certain cases the k -range gets too big. This is in line with common practice, where in a real data analysis one would plot the estimates of the parameter of interest as a function of k and try to identify the k -range where the estimate is stable, ignoring the part where the estimates diverge from the stable part.

Remark that this selected block for the estimation of $\gamma(x_0)$ will also be used for the estimation of $U_Y(1/p|x_0)$, as this gave for the latter more stable results compared to selecting the k range specifically for $\hat{U}_Y(1/p|x_0)$.

Our selection criteria for h_n and k are only heuristic and hence not optimal, but they allow for an effective data-driven approach on simulated and real data.

In Figure 1, we show the boxplots of $\hat{\gamma}_k(x_0)$ (left) and $\hat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$ (right) with $p = 1/n$ for $x_0 \in \{0.1, 0.2, \dots, 0.9\}$, and with the rows corresponding to the values of the second order parameter $\rho(x) = -2, -1$ and -0.5 , from top to bottom, in case $\theta = 2$. Figure 2 was constructed similarly, but for $\theta = 8$. In the left column we also show $\gamma(x_0)$ in blue. From the simulations we can draw the following conclusions:

- Overall, the estimation of $\gamma(x_0)$ is reasonably good for the cases with $\rho(x) = -2$ and -1 , while for $\rho(x) = -0.5$ there is some bias, but this can be expected when $\rho(x)$ is getting close to zero, since the rate of convergence of $U_Y(ty|x)/U_Y(t|x)$ to $y^{\gamma(x)}$ is then slower. This bias is already present in the simpler i.i.d. case, and also manifests itself here in a more complicated time series setting. In general, for all values of $\rho(x)$ considered in our simulations, $\hat{\gamma}_k(x)$ follows the shape of the true function $\gamma(x)$.

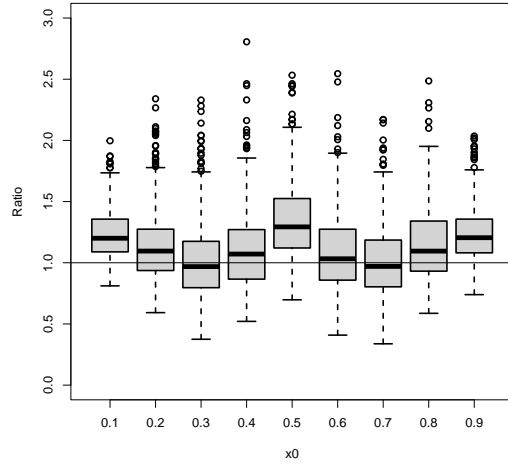
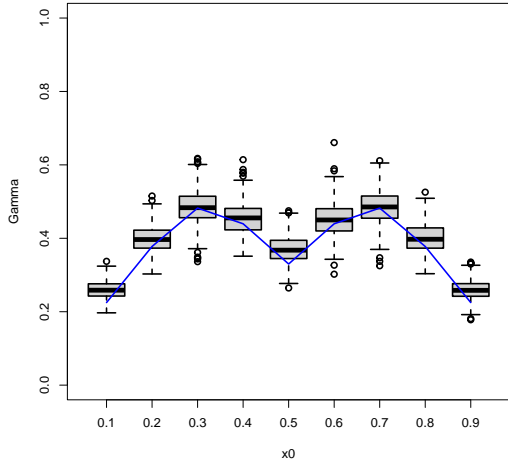
- Estimation of $\gamma(x)$ at the boundary of the X -range is more difficult, but this is common in local estimation.
- Concerning the boxplots of $\widehat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$, we observe that they are reasonably well centered around the value 1, and this for most of the covariate range, indicating a good performance.
- By comparing Figure 1 with Figure 2, we can see that the parameter θ of the copula function has a small impact on the results, both in the estimation of $\gamma(x_0)$ and in that of the quantile $U_Y(1/p|x_0)$.

Now, we compare our estimators with those proposed by Daouia et al. (2023) on the basis of the ratio between the empirical mean squared error (MSE) obtained with our estimators and with their estimators. These ratios are given in Table 1 for $\theta = 2$ and Table 2 for $\theta = 8$. In the case of the estimation of $\gamma(x_0)$, these ratios are often less than 1, indicating that our estimator performs better than that of Daouia et al. (2023), regardless of the values of x_0 , θ and λ . In the case of the quantile $U_Y(1/p|x_0)$, no estimator clearly stands out, as the results depend strongly on the value of x_0 . Therefore, as measure of overall performance we computed the sample mean of the MSE ratios over the considered x_0 (the row labeled 'Mean' in Tables 1 and 2). For the estimation of $\gamma(x_0)$ our estimator outperforms the one of Daouia et al. (2023), while for the estimation of $U_Y(1/p|x_0)$ our estimator is in most cases better, though the gain is less pronounced than for the estimation of $\gamma(x_0)$, indicated by ratios closer to one. However, it should be noted that if x_0 is at the boundary of the X -range, i.e., here 0.1 or 0.9, our estimators of the conditional extreme value index and the conditional extreme quantile are significantly better than those of Daouia et al. (2023), and this, regardless of the value of θ and λ .

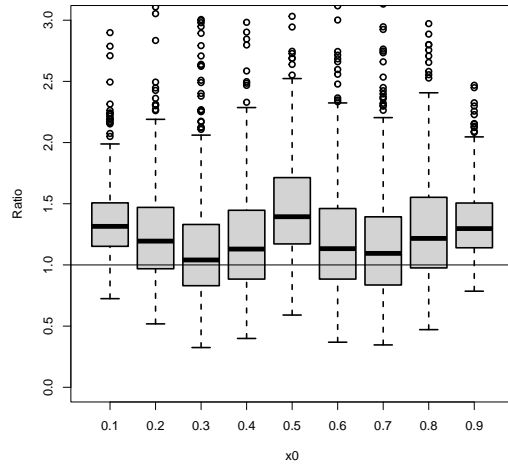
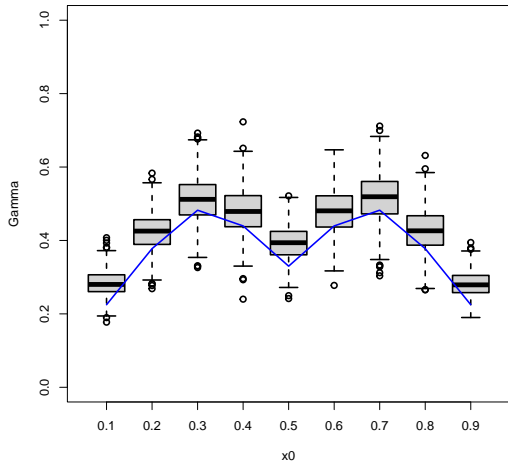
5 Real data analysis

In this section we apply our developed methodology to analyze financial returns of the S&P 500 market index given the market volatility. The dataset contains the daily log-returns of the S&P 500 index (Y) and the market volatility index VIX (X) over the time period January 5, 2021 to January 2, 2026, leading to a sample of size $n = 1255$. The VIX index measures

$$\rho(x) = -2$$



$$\rho(x) = -1$$



$$\rho(x) = -0.5$$

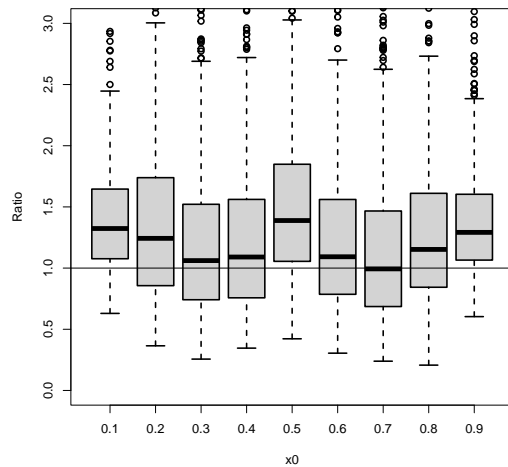
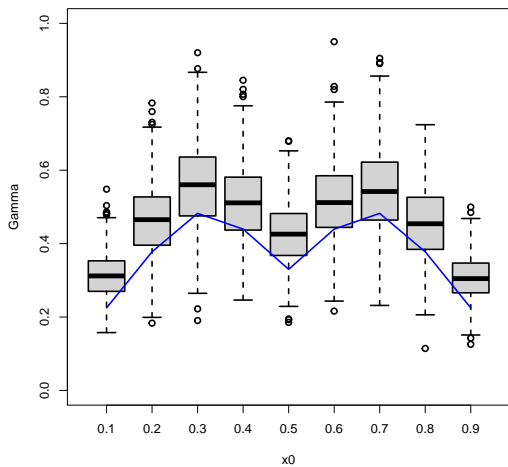
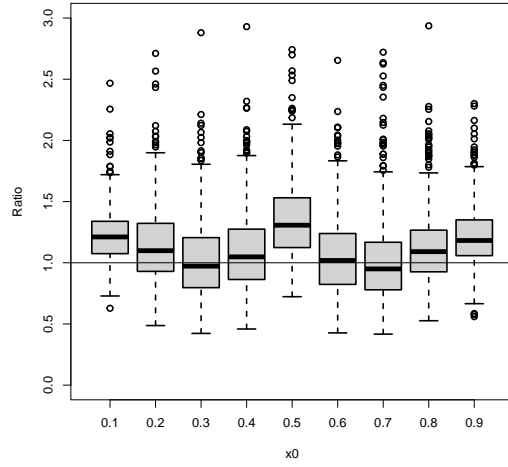
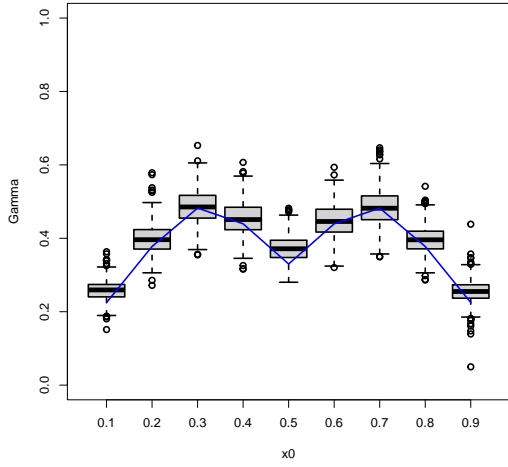
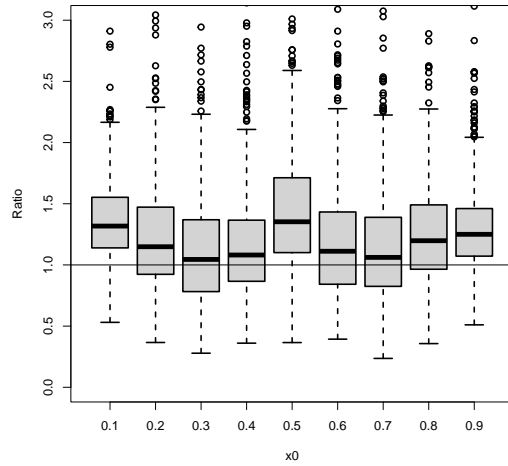
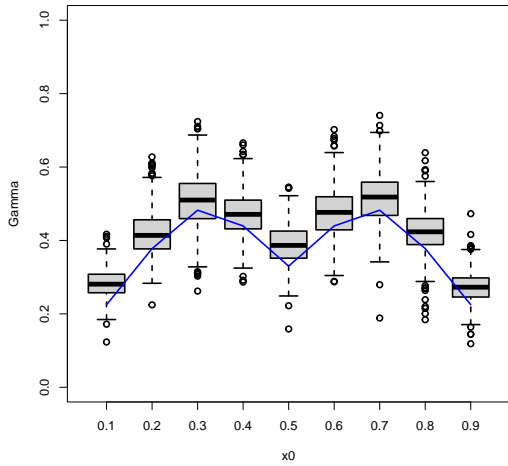


Figure 1: Boxplots of $\hat{\gamma}_k(x_0)$ (left) and $\hat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$ (right) for $x_0 \in \{0.1, 0.2, \dots, 0.9\}$, $p = 1/n$, $\theta = 2$ and where the rows correspond to $\rho(x) = -2$, -1 and -0.5 , from top to bottom. The blue line on the left panels represents the true value of $\gamma(x_0)$.

$$\rho(x) = -2$$



$$\rho(x) = -1$$



$$\rho(x) = -0.5$$

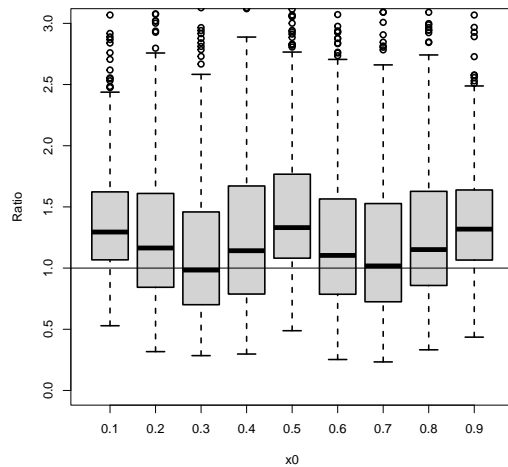
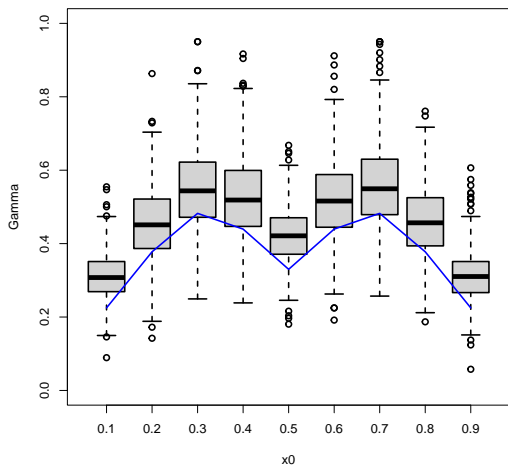


Figure 2: Boxplots of $\hat{\gamma}_k(x_0)$ (left) and $\hat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$ (right) for $x_0 \in \{0.1, 0.2, \dots, 0.9\}$, $p = 1/n$, $\theta = 8$ and where the rows correspond to $\rho(x) = -2$, -1 and -0.5 , from top to bottom. The blue line on the left panels represents the true value of $\gamma(x_0)$.

x_0	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$	
	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$
0.1	0.317	0.245	0.480	0.372	0.495	0.346
0.2	0.888	1.051	0.780	0.823	0.783	0.930
0.3	0.711	1.103	1.157	1.503	1.124	1.706
0.4	0.971	1.408	1.159	1.402	0.799	0.940
0.5	0.510	0.410	0.615	0.430	0.498	0.362
0.6	1.127	1.679	1.112	1.425	0.916	1.508
0.7	0.712	1.069	1.259	1.747	1.122	1.896
0.8	0.939	1.044	0.905	1.076	0.763	0.840
0.9	0.319	0.244	0.451	0.331	0.463	0.311
Mean	0.722	0.917	0.880	1.012	0.774	0.982

Table 1: Ratio of empirical MSE of our estimator versus Daouia et al. (2023) for $\hat{\gamma}_k(x_0)$ and $\hat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$, $\theta = 2$. The row 'Mean' gives the sample mean of the results over the x_0 -positions.

x_0	$\lambda = 0.5$		$\lambda = 1$		$\lambda = 2$	
	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$	$\hat{\gamma}_k(x_0)$	$\hat{U}_Y(1/p x_0)$
0.1	0.344	0.266	0.510	0.380	0.518	0.360
0.2	1.012	1.229	0.872	0.975	0.838	0.779
0.3	0.954	1.403	1.157	1.577	1.008	0.967
0.4	1.041	1.306	1.062	1.166	0.930	1.254
0.5	0.526	0.417	0.615	0.422	0.456	0.286
0.6	0.976	1.361	1.243	1.547	0.940	1.482
0.7	0.736	1.270	1.219	1.533	1.158	2.248
0.8	0.882	0.881	0.835	0.822	0.815	0.534
0.9	0.377	0.294	0.426	0.285	0.515	0.400
Mean	0.761	0.936	0.882	0.967	0.798	0.923

Table 2: Ratio of empirical MSE of our estimator versus Daouia et al. (2023) for $\hat{\gamma}_k(x_0)$ and $\hat{U}_Y(1/p|x_0)/U_Y(1/p|x_0)$, $\theta = 8$. The row 'Mean' gives the sample mean of the results over the x_0 -positions.

the market expectation of near term volatility, and is derived from the S&P 500 options for the 30 days following the measurement date. About half of the log-returns of S&P 500 index are negative, which poses difficulties in the calculation of the local Hill estimator $\hat{\gamma}_k(x_0)$. To overcome this issue, for the analysis we shift the data by subtracting from Y the empirical 10% quantile, so that the majority of the transformed data will be positive. Also, for ease of computation of the local estimator, the covariate X is transformed to the interval $(0, 1)$. All subsequent figures are based on these transformed variables, which are the data used for the analysis. In Figure 3 we show the time plots of the log-returns of S&P 500 (left) and VIX (right). As is common for financial returns, there are periods of high variability, with clusters of successive large losses and gains. The dependence between the two variables is visualized in the scatterplot given in Figure 4. As is clear from the plot, the point cloud is somehow horizontal, i.e., it does not indicate a clear positive or negative association between the variables in the sense of classical linear dependence, but the cloud somehow spreads out with increasing volatility, indicating a more heavy tail of the S&P 500 log-return distribution with increasing values of the VIX. The dependence between the variables is thus of a non-linear nature. In order to validate the Pareto-type assumption for the conditional distribution of Y given X , as given by (1), we make Pareto QQ plots of the log-returns in small neighborhoods of some reference positions x_0 . This is illustrated in Figure 5 for $x_0 = 0.15$ (left) and $x_0 = 0.3$ (right). Obviously, these Pareto QQ plots become linear in the largest observations which supports the assumption that the distribution of Y given X is of Pareto-type.

We now turn to the estimation of $\gamma(x_0)$ and $U_Y(1/p|x_0)$. Our estimators are implemented with the same selection criteria for h_n and k as described in the simulation study resulting in $h_n = 0.058$, and compared with the corresponding estimators from Daouia et al. (2023). We restrict the estimation to X -values up to 0.4, as beyond this value the data is very sparse for local estimation. In the left panel of Figure 6 we show $\hat{\gamma}_k(x_0)$ as a function of x_0 , where the black line represents the estimator from the present paper and the blue dashed line the estimator from Daouia et al. (2023). Globally the estimates of the tail index of the log-returns tend to increase with market volatility, though beyond the value 0.3 for the covariate the estimates obtained with the two methods are quite different. This can be explained by the fact that for a covariate above

0.3 there are fewer data available, making the estimates more variable. Also, the algorithm for selecting k from Daouia et al. (2023) tends to choose larger k -values compared to our method, and generally this leads to a larger bias (typically upwards) of the estimators, though a smaller variance. Our estimate somehow follows the information in the data more closely. The horizontal dashed-dotted line in Figure 6 shows the classical Hill estimator for the log-returns, ignoring the covariate VIX. Obviously, taking the information of a covariate, here VIX, into account is advantageous as it allows to differentiate the tail index, and hence the risk, according to the values of the covariate. Finally, in the right panel of Figure 6 we show the estimates for $\hat{U}_Y(1/p|x_0)$ for $p = 10/n$ along with a scatterplot of the data, where again we compare with Daouia et al. (2023). Overall the extreme quantile estimates tend to increase with the covariate for both estimation methods. Similarly as for the tail index estimates, the differences between the two estimates for the extreme quantile are due to the different algorithms for selecting k used in their implementations. As mentioned above, the Daouia et al. (2023) estimator is typically based on a larger k value, which will here draw the intermediate quantile estimate, i.e., $\hat{U}_Y(n/k|x_0)$, down, which can explain why their estimator is below ours. One can though see that our estimator somehow follows the pattern in the data more closely. The horizontal line in the figure represents a univariate extreme quantile estimate ignoring the covariate. Clearly, this estimate can be seen as some average quantile value, especially when compared to our estimation method.

6 Conclusion

In this article, we focused on estimating an extreme conditional quantile of order $1 - p$ for a strictly stationary time series $(X_t, Y_t)_{t \in \mathbb{Z}}$, when p is smaller than the inverse of the available sample size. More specifically, we assumed that the conditional distribution of Y given $X = x_0$ has a heavy tail, with index $\gamma(x_0)$, and we constructed an estimator using the Weissman method. That means that our extreme conditional quantile estimator depends on an estimator of $\gamma(x_0)$ and an intermediate quantile estimator of order $1 - k/n$, where k is an intermediate sequence. All these estimators are kernel-type and therefore based on observations in a neighborhood of x_0 . We first established the asymptotic properties of the estimators of $\gamma(x_0)$ and the intermediate

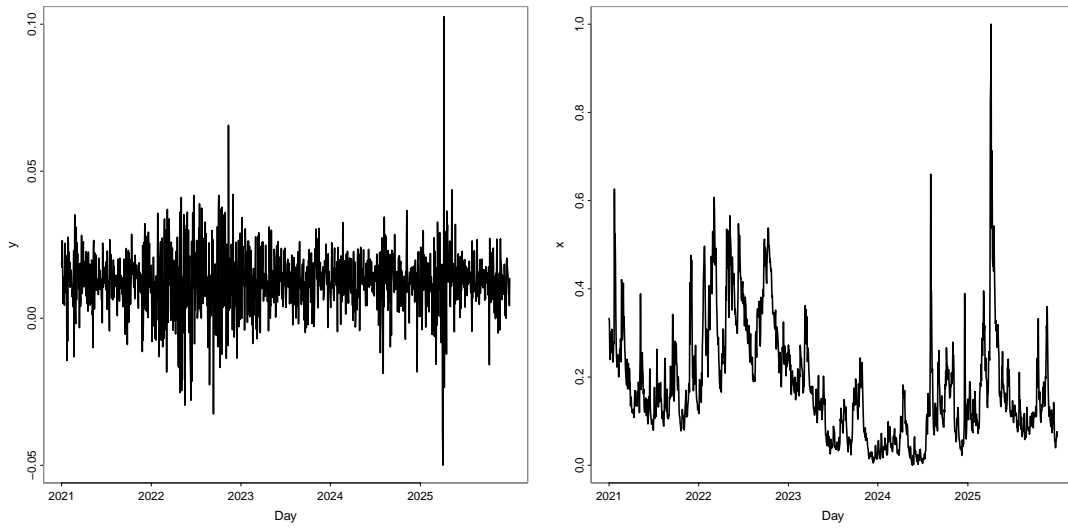


Figure 3: Time series plots of log-returns of S&P 500 (left) and the VIX index (right).

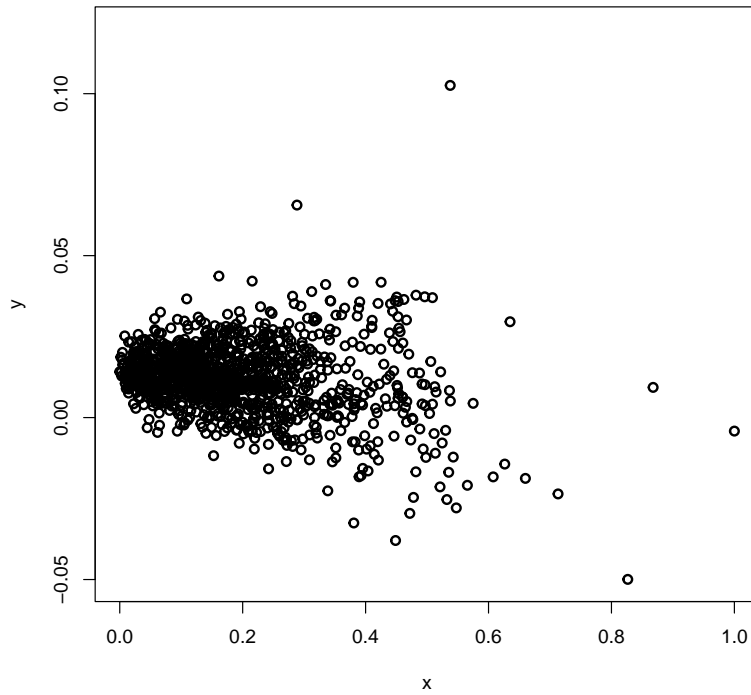


Figure 4: Scatterplot of log-returns of S&P 500 versus VIX.

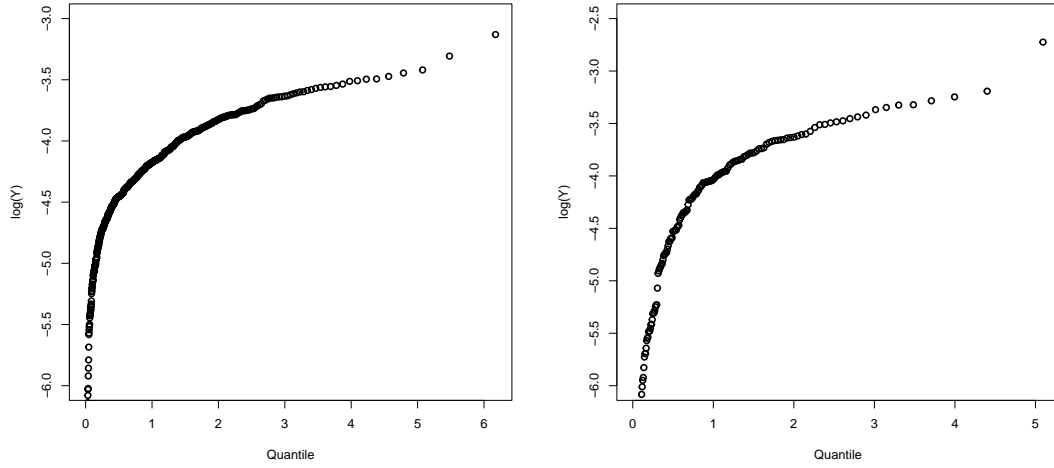


Figure 5: Local Pareto QQ plots of S&P 500 log-returns at $x_0 = 0.15$ (left) and $x_0 = 0.3$ (right).

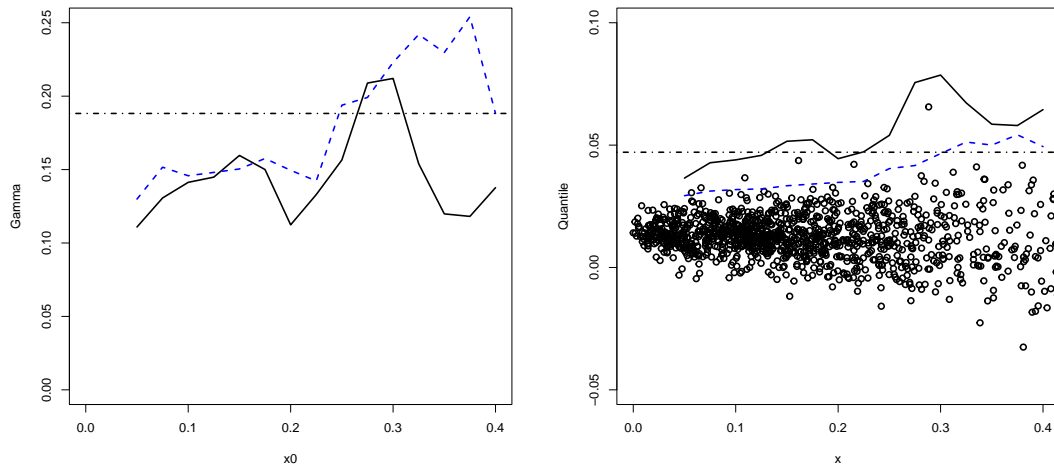


Figure 6: Estimates of $\gamma(x_0)$ (left) and $U_Y(1/p|x_0)$ with $p = 10/n$ along with the scatterplot of the data (right). The estimators proposed in the present paper are represented by the black solid line while the ones from Daouia et al. (2023) are in a blue dashed line. The horizontal dashed-dotted line is the estimate based on a univariate analysis, ignoring the covariate.

quantile, and then deduced those of the parameter of interest, using the cluster method of Drees and Rootzén (2010) under regularity assumptions. The performance of our estimators was illustrated by a simulation study based on a model where $Y|X$ follows a Burr distribution and $(X_t)_{t \in \mathbb{Z}}$ is generated by a Markov process with a Gumbel copula. We showed that this model satisfies the assumptions of our theory and compared our estimators to those proposed by Daouia et al. (2023). We illustrated the methodology on a real dataset where we estimated the tail index and an extreme quantile of the conditional distribution of the log-return of the S&P 500 index given the market volatility measured by the VIX index, highlighting the importance of taking a covariate into account for differentiating risks.

7 Appendix

7.1 Appendix A: Verification of our main assumptions for our model used in Section 4

We consider the model used in Section 4, namely, a conditional distribution of Y given $X = x$ of the form (3) and a time series $(X_t)_{t \in \mathbb{Z}}$ generated from a Markovian model with the logistic copula C_θ . We will show that our main assumptions of our theorems, i.e., Assumptions (\mathcal{J}) , (\mathcal{X}) and (\mathcal{A}) , are satisfied.

We start with Assumption (\mathcal{J}) . Since

$$\frac{\mathbb{P}(Y_0 > y_0, Y_1 > y_1 | X_0 = u_0, X_1 = u_1)}{\sqrt{\overline{F}_Y(y_0|u_0)\overline{F}_Y(y_1|u_1)}} \leq \frac{\min(\overline{F}_Y(y_0|u_0), \overline{F}_Y(y_1|u_1))}{\sqrt{\overline{F}_Y(y_0|u_0)\overline{F}_Y(y_1|u_1)}} \leq 1$$

we have $\Omega_{\bar{h}}(\bar{y}|x_0) \leq 1$. Assumption (\mathcal{J}) is thus satisfied for our model.

Now, concerning Assumption (\mathcal{X}) , we set $i_0 = 1$. In that case, the first part of the latter assumption is trivially satisfied, and thus we only need to show that

$$\limsup_{r \rightarrow 0} \sup_{i \geq 1} r^{-2} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(r), X_i \in \mathcal{B}_{x_0}(r)) < \infty. \quad (4)$$

To prove (4), we proceed as follows. By picking $x_0 \in (0, 1)$ and by choosing ε such that $0 < x_0 - \varepsilon < x_0 + \varepsilon < 1$, the compact set $E := [x_0 - \varepsilon, x_0 + \varepsilon]^2$ is such that $E \subset (0, 1)^2$ and we will

show that

$$\sup_{i \geq 1} \sup_{(u,v) \in E} f_{(X_0, X_i)}(u, v) < \infty, \quad (5)$$

where $f_{(X_0, X_i)}$ denotes the density of the pair (X_0, X_i) .

Now, define

$$\begin{aligned} c_\theta(u, v) &:= \frac{\partial^2}{\partial u \partial v} C_\theta(u, v) \\ &= \frac{C_\theta(u, v)}{u v} \frac{(-\log u)^{\theta-1} (-\log v)^{\theta-1}}{[(-\log u)^\theta + (-\log v)^\theta]^{2-\frac{1}{\theta}}} \left\{ [(-\log u)^\theta + (-\log v)^\theta]^{\frac{1}{\theta}} + \theta - 1 \right\} \end{aligned}$$

the copula density and

$$\begin{aligned} p^{(i)}(u, v) &:= f_{X_i|X_0}(v|u), \\ H(u) &:= (-\log u)^{-1/2} + (-\log u)^\theta, \\ T(v) &:= \sup_{u \in (0,1)} \frac{|c_\theta(u, v)|}{H(u)}. \end{aligned}$$

Note that the function H is integrable with respect to the invariant measure (since $\int_0^1 H(u) du < \infty$) and satisfies the standard requirements for a Lyapunov function as defined by Meyn and Tweedie (1993, page ix), specifically the drift condition for stability. Furthermore, $T(v)$ is finite uniformly on $v \in [x_0 - \varepsilon, x_0 + \varepsilon]$.

According to Chen et al. (2009), the Markov process generated by the logistic copula is geometrically ergodic. Thus, using Meyn and Tweedie (1993, page ix), there exist constants $C > 0$ and $\rho \in (0, 1)$ such that, for $i \geq 2$,

$$\begin{aligned} |f_{(X_0, X_i)}(u, v) - 1| &= \left| \int_0^1 f_{(X_0, X_{i-1}, X_i)}(u, w, v) dw - 1 \right| \\ &= \left| \int_0^1 c_\theta(w, v) p^{(i-1)}(u, w) dw - 1 \right| \\ &= \left| \int_0^1 c_\theta(w, v) \{p^{(i-1)}(u, w) - 1\} dw \right| \\ &= T(v) \left| \int_0^1 \frac{c_\theta(w, v)}{T(v)} \{p^{(i-1)}(u, w) - 1\} dw \right| \\ &\leq T(v) \sup_{h: |h| \leq H} \left| \int_0^1 h(w) \{p^{(i-1)}(u, w) - 1\} dw \right| \\ &\leq C T(v) H(u) \rho^{i-1}. \end{aligned}$$

Thus, we deduce that

$$\begin{aligned} \sup_{i \geq 2} \sup_{(u,v) \in E} f_{(X_0, X_i)}(u, v) &\leq 1 + C \left\{ \sup_{v \in [x_0 - \varepsilon, x_0 + \varepsilon]} T(v) \right\} \left\{ \sup_{u \in [x_0 - \varepsilon, x_0 + \varepsilon]} H(u) \right\} \left\{ \sup_{i \geq 2} \rho^{i-1} \right\} \\ &\leq 1 + \tilde{C} \rho, \end{aligned} \tag{6}$$

for $\tilde{C} > 0$.

Now

$$f_{(X_0, X_1)}(u, v) = c_\theta(u, v)$$

which is continuous on E . Since E is a compact set of $(0, 1)^2$, the function $f_{(X_0, X_1)}(u, v)$ is bounded on E by a finite constant, say C_1 . Combining with (6), this yields

$$\sup_{i \geq 1} \sup_{(u,v) \in E} f_{(X_0, X_i)}(u, v) \leq \max(C_1, 1 + \tilde{C} \rho) =: \check{C},$$

from which (5) follows. Finally, for $r < \varepsilon$, since $[x_0 - r, x_0 + r]^2 \subset E$ we have

$$\mathbb{P}(X_0 \in \mathcal{B}_{x_0}(r), X_i \in \mathcal{B}_{x_0}(r)) = \int_{x_0 - r}^{x_0 + r} \int_{x_0 - r}^{x_0 + r} f_{(X_0, X_i)}(u, v) du dv \leq \check{C} (2r)^2 = 4r^2 \check{C},$$

from which (4) follows.

Note also that Assumption (\mathcal{A}) is satisfied since the Markov process generated by the logistic copula is geometrically ergodic, and thus geometrically β -mixing. Given that the conditional distribution of Y given X is of the form (3) and independent of the past history $\sigma((X_s, Y_s); s < t)$, once X_t is observed, the joint process $(X_t, Y_t)_{t \in \mathbb{Z}}$ inherits the geometrically β -mixing property of $(X_t)_{t \in \mathbb{Z}}$. Our vector $Z_{n,i}$ is defined as a measurable function, say $f_n(X_i, Y_i)$, for each n . By the property of mixing coefficients under measurable transformations, we have $\sigma(Z_{n,\ell}; 1 \leq \ell \leq i) \subset \sigma((X_\ell, Y_\ell); 1 \leq \ell \leq i)$. As a consequence, $\beta_{n,k}$ is bounded by the mixing coefficient of $(X_t, Y_t)_{t \in \mathbb{Z}}$, i.e., of $(X_t)_{t \in \mathbb{Z}}$. Thus, under suitable assumptions on k and h_n , we can find sequences ℓ_n and r_n satisfying Assumption (\mathcal{A}) , see the comment below Assumption (\mathcal{A}) .

7.2 Appendix B: Proofs of the results

The proofs of our main results rely on two auxiliary lemmas presented below, the proofs of which will be given at the end of the paper, namely in Sections 7.2.5 and 7.2.6.

The first lemma allows us to control the relative variation of the conditional survival distribution function evaluated at an extreme level, say y_n , when the covariate x varies uniformly in the ball $\mathcal{B}_{x_0}(h_n)$, namely the variation of the factor $\overline{F}_Y(y_n|x)/\overline{F}_Y(y_n|x_0)$.

Lemma 7.1 *Assume (D) and (H), and let $(y_n)_{n \in \mathbb{N}}$ be a sequence such that $y_n \rightarrow \infty$ and $h_n^{\eta_\gamma \wedge \eta_\xi} \log y_n \rightarrow 0$. Then, uniformly in $x \in \mathcal{B}_{x_0}(h_n)$ and for n large, we have:*

$$\left| \frac{\overline{F}_Y(y_n|x)}{\overline{F}_Y(y_n|x_0)} - 1 \right| \leq C \left(h_n^{\eta_A} + y_n^{Ch_n^{\eta_\gamma}} h_n^{\eta_\gamma} \log y_n + |\delta_Y(y_n|x_0)| h_n^{\eta_B} + |\delta_Y(y_n|x_0)| y_n^{Ch_n^{\eta_\xi}} h_n^{\eta_\xi} \log y_n \right),$$

where C is a positive constant.

The second lemma is more technical and is used in several places in the proof of Theorem 3.1 to establish some assumptions required to apply the cluster method of Drees and Rootzén (2010).

Lemma 7.2 *Let $(X_t, Y_t)_{t \in \mathbb{Z}}$ be a strictly stationary time series, with a conditional distribution function of the response variable Y given $X = x$ satisfying Assumption (D). Let k be an intermediate sequence such that $k/n \rightarrow 0$ and h_n be a bandwidth satisfying $h_n^{\eta_\gamma \wedge \eta_\xi} \log(n/k) \rightarrow 0$ as $n \rightarrow \infty$. Suppose that Assumptions (K), (H), (J), (X) and (A) are fulfilled, and x_0 is such that $g_X(x_0) > 0$. Then*

(i) *For all $0 \leq i \leq r_n - 1$, there are*

$$\begin{aligned} s_n(i) &\geq \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \max \left(\log \frac{Y_0}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \right. \\ &\quad \times \max \left(\log \frac{Y_i}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \\ &\quad \left. \times \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right. \right], \end{aligned}$$

such that $s_\infty(i) = \lim_{n \rightarrow \infty} s_n(i)$ exists, and $\lim_{n \rightarrow \infty} \sum_{i=0}^{r_n-1} s_n(i) = \sum_{i=0}^{\infty} s_\infty(i) < \infty$.

(ii) *Moreover, as $n \rightarrow \infty$*

$$\begin{aligned} &\sum_{i=0}^{r_n-1} \left\{ \mathbb{E} \left[\left(K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \max \left(\log \frac{Y_0}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \right. \right. \right. \\ &\quad \times \max \left(\log \frac{Y_i}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \\ &\quad \left. \left. \left. \times \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \right)^{1+\delta} \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right. \right] \right\}^{\frac{1}{1+\delta}} = O(1). \end{aligned}$$

7.2.1 Proof of Theorem 3.1

We state and verify the assumptions from Drees and Rootzén (2010) for the class $\Phi_\varepsilon := \{\phi_s : s \in [1 - \varepsilon, 1 + \varepsilon]\}$, since the proofs for $\Psi_\varepsilon := \{\psi_s : s \in [1 - \varepsilon, 1 + \varepsilon]\}$ are similar and simpler. Let $[0, \infty)_\cup := \bigcup_{r \in \mathbb{N}} [0, \infty)^{r(d+1)}$, the set of all vectors with nonnegative entries and arbitrary length, and $\mathcal{F}_\varepsilon := \{f_s; s \in [1 - \varepsilon, 1 + \varepsilon]\}$ with

$$f_s(z) := \sum_{i=1}^r \phi_s(z_i),$$

for $z = (z_1, \dots, z_r) \in [0, \infty)^{r(d+1)}$ and $r \in \mathbb{N}$. So \mathcal{F}_ε contains arbitrary long sums with summands coming from Φ_ε .

Assumption (D1): $\mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} \phi_s(Z_{n,i}) \right\}^2 \right]$ has to be finite for all $s \in [1 - \varepsilon, 1 + \varepsilon]$ and all $n \geq 1$, and the envelope function

$$F_\varepsilon(z) := \sup_{f \in \mathcal{F}_\varepsilon} f(z),$$

has to be finite for all $z \in [0, \infty)_\cup$.

Proof of Assumption (D1). Using stationarity of (X_i, Y_i) , Assumptions (\mathcal{K}) , (\mathcal{A}) and Lemma 7.2, we have

$$\begin{aligned} & \mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} \phi_s(Z_{n,i}) \right\}^2 \right] \\ & \leq \mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} K \left(\frac{x_0 - X_i}{h_n} \right) \log \left(\frac{Y_i}{(1 - \varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}} \right\}^2 \right] \\ & = \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left[K \left(\frac{x_0 - X_i}{h_n} \right) K \left(\frac{x_0 - X_j}{h_n} \right) \log \left(\frac{Y_i}{(1 - \varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}} \right. \\ & \quad \left. \times \log \left(\frac{Y_j}{(1 - \varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_j}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}} \right] \\ & \leq 2r_n \sum_{i=0}^{r_n-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \log \left(\frac{Y_0}{(1 - \varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}} \right. \\ & \quad \left. \times \log \left(\frac{Y_i}{(1 - \varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1 - \varepsilon \right\}} \right] \end{aligned}$$

$$\begin{aligned}
&= 2 r_n v_n \sum_{i=0}^{r_n-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \log \left(\frac{Y_0}{(1-\varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \right. \\
&\quad \left. \times \log \left(\frac{Y_i}{(1-\varepsilon)U_Y(n/k|x_0)} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right] \\
&\leq C r_n h_n^d \frac{k}{n} \sum_{i=0}^{r_n-1} s_n(i) = o(1),
\end{aligned}$$

for C a positive constant. Since $\phi_s(z) \leq \phi_{1-\varepsilon}(z) < \infty, \forall z$, Assumption (D1) is established. \square

Assumption (D2'): We have, $\forall \tilde{\varepsilon} > 0$

$$\mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) \right\}^2 \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] = o(r_n v_n).$$

Proof of Assumption (D2'). Using Hölder's and Markov's inequalities, under Assumptions (K), (A) and Lemma 7.2, we have for $\tilde{\varepsilon} > 0$

$$\begin{aligned}
&\mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) \right\}^2 \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] \\
&= \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left[\phi_{1-\varepsilon}(Z_{n,i}) \phi_{1-\varepsilon}(Z_{n,j}) \mathbb{1}_{\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) > \tilde{\varepsilon} \sqrt{n v_n} \right\}} \right] \\
&\leq \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \left[\mathbb{E} \left(\left\{ \phi_{1-\varepsilon}(Z_{n,i}) \phi_{1-\varepsilon}(Z_{n,j}) \right\}^{1+\delta} \right) \right]^{\frac{1}{1+\delta}} \left[\mathbb{P} \left(\sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) > \tilde{\varepsilon} \sqrt{n v_n} \right) \right]^{\frac{\delta}{1+\delta}} \\
&\leq 2 r_n \sum_{i=0}^{r_n-1} \left[\mathbb{E} \left(\left\{ \phi_{1-\varepsilon}(Z_{n,0}) \phi_{1-\varepsilon}(Z_{n,i}) \right\}^{1+\delta} \right) \right]^{\frac{1}{1+\delta}} \left[\frac{1}{\tilde{\varepsilon}^2 n v_n} \mathbb{E} \left(\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) \right\}^2 \right) \right]^{\frac{\delta}{1+\delta}} \\
&\leq 2 r_n v_n^{\frac{1}{1+\delta}} \sum_{i=0}^{r_n-1} \left[\mathbb{E} \left(\left\{ \phi_{1-\varepsilon}(Z_{n,0}) \phi_{1-\varepsilon}(Z_{n,i}) \right\}^{1+\delta} \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right) \right]^{\frac{1}{1+\delta}} \\
&\quad \times O \left(\left(\frac{r_n}{n} \right)^{\frac{\delta}{1+\delta}} \right) \\
&= O \left(r_n v_n \left(\frac{r_n}{\sqrt{n v_n}} \right)^{\frac{\delta}{1+\delta}} \left(\frac{1}{\sqrt{n v_n}} \right)^{\frac{\delta}{1+\delta}} \right) \\
&= o(r_n v_n).
\end{aligned}$$

Note that the last inequality of the above display comes from the proof of our Assumption (D1).

This achieves the proof of Assumption (D2'). \square

Assumption (D3): There exists a semi-metric ρ such that

$$\lim_{\delta \searrow 0} \limsup_{n \rightarrow \infty} \sup_{\substack{1-\varepsilon \leq s, t \leq 1+\varepsilon \\ \rho(s, t) < \delta}} \frac{1}{r_n v_n} \mathbb{E} \left[\left(\sum_{i=1}^{r_n} \phi_s(Z_{n,i}) - \sum_{i=1}^{r_n} \phi_t(Z_{n,i}) \right)^2 \right] = 0.$$

Proof of Assumption (D3). Assume $s, t \in [1 - \varepsilon, 1 + \varepsilon]$ and without loss of generality $s < t$.

Then

$$\begin{aligned} & \mathbb{E} \left[\left(\sum_{i=1}^{r_n} \phi_s(Z_{n,i}) - \sum_{i=1}^{r_n} \phi_t(Z_{n,i}) \right)^2 \right] \\ &= \mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} K \left(\frac{x_0 - X_i}{h_n} \right) \left[\log \frac{Y_i}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_i}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > t\}} \right] \right\}^2 \right] \\ &= \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \mathbb{E} \left[K \left(\frac{x_0 - X_i}{h_n} \right) K \left(\frac{x_0 - X_j}{h_n} \right) \right. \\ &\quad \times \left\{ \log \frac{Y_i}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_i}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > t\}} \right\} \\ &\quad \times \left. \left\{ \log \frac{Y_j}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_j}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_j}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_j}{U_Y(n/k|x_0)} > t\}} \right\} \right] \\ &\leq 2 r_n \sum_{i=0}^{r_n-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \right. \\ &\quad \times \left\{ \log \frac{Y_0}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_0}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > t\}} \right\} \\ &\quad \times \left. \left\{ \log \frac{Y_i}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_i}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > t\}} \right\} \right] \\ &= 2 r_n \sum_{i=0}^{M-1} \mathbb{E} [\dots] + 2 r_n \sum_{i=M}^{r_n-1} \mathbb{E} [\dots] \\ &=: T_{1,n} + T_{2,n}, \end{aligned}$$

for a fixed $M > 0$. We will study these two terms separately.

Term $T_{1,n}$: Note that, since

$$\log \frac{Y_0}{s U_Y(n/k|x_0)} = \log \frac{Y_0}{t U_Y(n/k|x_0)} + \log \frac{t}{s},$$

we have

$$\begin{aligned}
& \log \frac{Y_0}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} - \log \frac{Y_0}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > t\}} \\
&= \log \frac{Y_0}{t U_Y(n/k|x_0)} \mathbb{1}_{\{s < \frac{Y_0}{U_Y(n/k|x_0)} \leq t\}} + \log \frac{t}{s} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} \\
&\leq \log \frac{t}{s} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} \\
&\leq \log \frac{t}{s} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon\}},
\end{aligned}$$

from which it follows by Assumption (\mathcal{K}) that

$$\begin{aligned}
T_{1,n} &\leq 2r_n (\log t - \log s)^2 \sum_{i=0}^{M-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon\}} \right] \\
&\leq 2r_n (\log t - \log s)^2 \\
&\quad \sum_{i=0}^{M-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\{\frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon\}} \right] \\
&\leq C M r_n v_n (\log t - \log s)^2.
\end{aligned}$$

By choosing $\rho(s, t) := (\log t - \log s)^2$, we have for $\rho(s, t) \leq \tilde{\delta}$

$$\frac{T_{1,n}}{r_n v_n} \leq C M \tilde{\delta}. \tag{7}$$

Term $T_{2,n}$: Clearly, we have

$$\begin{aligned}
T_{2,n} &\leq 2r_n \sum_{i=M}^{r_n-1} \left\{ \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \right. \right. \\
&\quad \times \log \frac{Y_0}{s U_Y(n/k|x_0)} \log \frac{Y_i}{s U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > s\}} \left. \right] \\
&\quad + \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \right. \\
&\quad \times \log \frac{Y_0}{t U_Y(n/k|x_0)} \log \frac{Y_i}{t U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > t\}} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > t\}} \left. \right] \left. \right\} \\
&\leq 4r_n \sum_{i=M}^{r_n-1} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \right. \\
&\quad \times \log \frac{Y_0}{(1-\varepsilon) U_Y(n/k|x_0)} \log \frac{Y_i}{(1-\varepsilon) U_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > (1-\varepsilon)\}} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > (1-\varepsilon)\}} \left. \right] \\
&\leq 4r_n v_n \sum_{i=M}^{r_n-1} s_n(i).
\end{aligned}$$

Thus, for $\rho(s, t) \leq \tilde{\delta}$, we have

$$\frac{T_{2,n}}{r_n v_n} \leq 4 \sum_{i=M}^{r_n-1} s_n(i) \longrightarrow 4 \sum_{i=M}^{\infty} s_{\infty}(i) < \infty. \quad (8)$$

By choosing M sufficiently large, the right-hand side of (8) can be made arbitrarily small. Given such M , by choosing $\tilde{\delta}$ small, the right-hand side of (7) can be made arbitrarily small too. \square

Assumption (D5): Let $\tilde{Z}_{n,j} := (\tilde{Z}_{n,j,i})_{1 \leq i \leq r_n}$ be independent copies of $(Z_{n,i})_{1 \leq i \leq r_n}$, for $1 \leq j \leq \lfloor n/(2r_n) \rfloor$. The map

$$\sup_{\substack{1-\varepsilon \leq s, t \leq 1+\varepsilon \\ \rho(s,t) < \tilde{\delta}}} \sum_{j=1}^{\lfloor n/(2r_n) \rfloor} e_j \left\{ \sum_{i=1}^{r_n} \left[\phi_s(\tilde{Z}_{n,j,i}) - \phi_t(\tilde{Z}_{n,j,i}) \right] \right\}^a,$$

is measurable, $\forall \tilde{\delta} > 0, n \in \mathbb{N}, (e_j)_{1 \leq j \leq \lfloor n/(2r_n) \rfloor} \in \{-1, 0, 1\}^{\lfloor n/(2r_n) \rfloor}$ and $a \in \{1, 2\}$.

The proof is completely similar to the one in Goegebeur et al. (2025). \square

Assumption (D6'): The envelope function F_{ε} is measurable with

$$\mathbb{E} \left[\left\{ \sum_{i=1}^{r_n} \phi_{1-\varepsilon}(Z_{n,i}) \right\}^2 \right] = O(r_n v_n)$$

and

$$\int_0^1 \sup_{Q \in \mathcal{Q}} \sqrt{\log \mathcal{N} \left(\zeta \left(\int F_{\varepsilon}^2 dQ \right)^{1/2}, \mathcal{F}_{\varepsilon}, L_2(Q) \right)} d\zeta < \infty,$$

where \mathcal{N} denotes the covering numbers and Q ranges over the set of discrete probability measures \mathcal{Q} .

The proof is completely similar to the one in Goegebeur et al. (2025). \square

We also need to verify the two following conditions of Theorem 2.10 in Drees and Rootzén (2010), ensuring that all finite-dimensional distributions converge:

Assumption (C1): We have

$$\begin{aligned} \mathbb{E} \left[(\Delta_n(s) - \mathbb{E}\Delta_n(s))^2 \mathbb{1}_{\{|\Delta_n(s) - \mathbb{E}\Delta_n(s)| \leq \sqrt{n v_n}\}} \right] &= o(r_n v_n), \\ \mathbb{P} (|\Delta_n(s) - \mathbb{E}\Delta_n(s)| > \sqrt{n v_n}) &= o\left(\frac{r_n}{n}\right), \end{aligned}$$

for all $s \in [1 - \varepsilon, 1 + \varepsilon]$, where

$$\Delta_n(s) := \sum_{i=r_n-\ell_n+1}^{r_n} \phi_s(Z_{n,i}).$$

The proof is completely similar to the one in Goegebeur et al. (2025). \square

Assumption (C3): We have for all $s, t \in [1 - \varepsilon, 1 + \varepsilon]$

$$\frac{1}{r_n v_n} \text{Cov} \left(\sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \longrightarrow \mathbb{E}[\phi_s(U + \underline{2}, V) \phi_t(U + \underline{2}, V)].$$

Proof of Assumption (C3). We have

$$\begin{aligned} & \frac{1}{r_n v_n} \text{Cov} \left(\sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \\ &= \frac{1}{r_n v_n} \sum_{i=1}^{r_n} \sum_{j=1}^{r_n} \{ \mathbb{E}[\phi_s(Z_{n,i}) \phi_t(Z_{n,j})] - \mathbb{E}[\phi_s(Z_{n,0})] \mathbb{E}[\phi_t(Z_{n,0})] \} \\ &= \frac{1}{r_n v_n} \left[\sum_{i=1}^{r_n-1} (r_n - i) \{ \mathbb{E}[\phi_s(Z_{n,0}) \phi_t(Z_{n,i})] + \mathbb{E}[\phi_s(Z_{n,i}) \phi_t(Z_{n,0})] \} \right. \\ & \quad \left. + r_n \mathbb{E}[\phi_s(Z_{n,0}) \phi_t(Z_{n,0})] - r_n^2 \mathbb{E}[\phi_s(Z_{n,0})] \mathbb{E}[\phi_t(Z_{n,0})] \right] \\ &= \sum_{i=1}^{r_n-1} \left(1 - \frac{i}{r_n} \right) \left\{ \mathbb{E} \left[\phi_s(Z_{n,0}) \phi_t(Z_{n,i}) \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right] \right. \\ & \quad \left. + \mathbb{E} \left[\phi_s(Z_{n,i}) \phi_t(Z_{n,0}) \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right] \right\} \\ & \quad + \mathbb{E} \left[\phi_s(Z_{n,0}) \phi_t(Z_{n,0}) \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right] \\ & \quad - r_n v_n \mathbb{E} \left[\phi_s(Z_{n,0}) \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right] \\ & \quad \quad \times \mathbb{E} \left[\phi_t(Z_{n,0}) \middle| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right] \\ &=: T_{3,n} + T_{4,n} + T_{5,n}. \end{aligned}$$

The first step consists to show that $T_{5,n} = O(r_n v_n)$. To this aim, note that

$$\begin{aligned}
& \mathbb{E} \left[\phi_s(Z_{n,0}) \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right. \right] \\
&= \frac{1}{v_n} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) \mathbb{E} \left[\log \frac{Y_0}{sU_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} \middle| X_0 \right] \right] \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} K \left(\frac{x_0 - x}{h_n} \right) \int_{sU_Y(n/k|x_0)}^{\infty} \log \frac{y}{sU_Y(n/k|x_0)} dF_Y(y|x) g_X(x) dx \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} K \left(\frac{x_0 - x}{h_n} \right) \int_{sU_Y(n/k|x_0)}^{\infty} \frac{\bar{F}_Y(y|x)}{y} dy g_X(x) dx \\
&= \frac{h_n^d}{v_n} \int_{\mathcal{B}_{\underline{0}}(1)} K(z) g_X(x_0 - zh_n) dz \int_{sU_Y(n/k|x_0)}^{\infty} \frac{\bar{F}_Y(y|x_0)}{y} dy \\
&\quad + \frac{h_n^d}{v_n} \int_{\mathcal{B}_{\underline{0}}(1)} K(z) \int_1^{\infty} \frac{\bar{F}_Y(sU_Y(n/k|x_0)\omega|x_0)}{\omega} \left[\frac{\bar{F}_Y(sU_Y(n/k|x_0)\omega|x_0 - zh_n)}{\bar{F}_Y(sU_Y(n/k|x_0)\omega|x_0)} - 1 \right] d\omega \\
&\quad \times g_X(x_0 - zh_n) dz \\
&=: T_{5,n}^{(1)} + T_{5,n}^{(2)}.
\end{aligned}$$

Clearly, we have under Assumptions (\mathcal{K}) and (\mathcal{H})

$$\begin{aligned}
T_{5,n}^{(1)} &= \frac{k h_n^d}{n v_n} \left\{ g_X(x_0) + \int_{\mathcal{B}_{\underline{0}}(1)} K(z) [g_X(x_0 - zh_n) - g_X(x_0)] dz \right\} \\
&\quad \times \left\{ \int_1^{\infty} (sz)^{-\frac{1}{\gamma(x_0)}} \frac{1}{z} dz + \int_1^{\infty} \left[\frac{\bar{F}_Y(sU_Y(n/k|x_0)z|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} - (sz)^{-\frac{1}{\gamma(x_0)}} \right] \frac{1}{z} dz \right\} \\
&= O(1),
\end{aligned}$$

by Proposition B.1.10 in de Haan and Ferreira (2006).

Our Lemma 7.1 together with similar arguments yield $T_{5,n}^{(2)} = O(1)$. Thus $T_{5,n} = O(r_n v_n)$.

Next, we study $T_{4,n}$. We have

$$\begin{aligned}
T_{4,n} &= \frac{1}{v_n} \mathbb{E} \left[K^2 \left(\frac{x_0 - X_0}{h_n} \right) \mathbb{E} \left[\log \frac{Y_0}{sU_Y(n/k|x_0)} \log \frac{Y_0}{tU_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > \max(s,t)\}} \middle| X_0 \right] \right] \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} K^2 \left(\frac{x_0 - x}{h_n} \right) \int_{\max(s,t)U_Y(n/k|x_0)}^{\infty} \log \frac{y}{sU_Y(n/k|x_0)} \log \frac{y}{tU_Y(n/k|x_0)} dF_Y(y|x) \\
&\quad \times g_X(x) dx.
\end{aligned}$$

Now, if we look at the inner integral, denoted as $I(x)$, we have

$$\begin{aligned}
I(x) &= - \int_{\max(s,t)U_Y(n/k|x_0)}^{\infty} \left(\log \frac{y}{U_Y(n/k|x_0)} \right)^2 d\bar{F}_Y(y|x) \\
&\quad + \log(st) \int_{\max(s,t)U_Y(n/k|x_0)}^{\infty} \log \frac{y}{U_Y(n/k|x_0)} d\bar{F}_Y(y|x) \\
&\quad - \log s \log t \int_{\max(s,t)U_Y(n/k|x_0)}^{\infty} d\bar{F}_Y(y|x) \\
&= 2 \int_1^{\infty} \frac{\log z}{z} \bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x) dz \\
&\quad + \log \frac{\max(s,t)}{\min(s,t)} \int_1^{\infty} \frac{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x)}{z} dz \\
&= [\max(s,t)]^{-\frac{1}{\gamma(x_0)}} \frac{k}{n} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s,t)}{\min(s,t)} \right\} \\
&\quad + 2 \frac{k}{n} \int_1^{\infty} \left[\frac{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} - (\max(s,t)z)^{-\frac{1}{\gamma(x_0)}} \right] \frac{\log z}{z} dz \\
&\quad + \log \frac{\max(s,t)}{\min(s,t)} \frac{k}{n} \int_1^{\infty} \left[\frac{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} - (\max(s,t)z)^{-\frac{1}{\gamma(x_0)}} \right] \frac{1}{z} dz \\
&\quad + 2 \int_1^{\infty} \bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0) \left[\frac{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x)}{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0)} - 1 \right] \frac{\log z}{z} dz \\
&\quad + \log \frac{\max(s,t)}{\min(s,t)} \\
&\quad \times \int_1^{\infty} \bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0) \left[\frac{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x)}{\bar{F}_Y(\max(s,t)U_Y(n/k|x_0)z|x_0)} - 1 \right] \frac{1}{z} dz \\
&=: [\max(s,t)]^{-\frac{1}{\gamma(x_0)}} \frac{k}{n} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s,t)}{\min(s,t)} \right\} + I_1 + I_2 + I_3(x) + I_4(x).
\end{aligned}$$

Again by Proposition B.1.10 in de Haan and Ferreira (2006), we have $I_1 = o\left(\frac{k}{n}\right)$ and $I_2 = o\left(\frac{k}{n}\right)$.

Lemma 7.1 implies now that

$$\begin{aligned}
T_{4,n} &= \frac{k h_n^d}{n v_n} [\max(s,t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s,t)}{\min(s,t)} \right\} \|K\|_2^2 g_X(x_0) (1 + o(1)) \\
&\quad + \frac{h_n^d}{v_n} \int_{\mathcal{B}_0(1)} K^2(z) I_3(x_0 - h_n z) g_X(x_0 - h_n z) dz \\
&\quad + \frac{h_n^d}{v_n} \int_{\mathcal{B}_0(1)} K^2(z) I_4(x_0 - h_n z) g_X(x_0 - h_n z) dz \\
&= \frac{k h_n^d}{n v_n} [\max(s,t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s,t)}{\min(s,t)} \right\} \|K\|_2^2 g_X(x_0) + o\left(\frac{k h_n^d}{n v_n}\right) \\
&= \frac{(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}{V_d} \|K\|_2^2 [\max(s,t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s,t)}{\min(s,t)} \right\} + o(1).
\end{aligned}$$

Denoting by $g_{(X_0, X_i)}(\cdot, \cdot)$ the density of the pair (X_0, X_i) and by $F_{(Y_0, Y_i)}(\cdot, \cdot | X_0, X_i)$ the distribution function of (Y_0, Y_i) conditionally on (X_0, X_i) , the last step consists to study

$$\begin{aligned}
& \mathbb{E} \left[\phi_s(Z_{n,0}) \phi_t(Z_{n,i}) \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_0(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right. \right] \\
&= \frac{1}{v_n} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \right. \\
&\quad \left. \times \mathbb{E} \left[\log \frac{Y_0}{sU_Y(n/k|x_0)} \log \frac{Y_i}{tU_Y(n/k|x_0)} \mathbb{1}_{\{\frac{Y_0}{U_Y(n/k|x_0)} > s\}} \mathbb{1}_{\{\frac{Y_i}{U_Y(n/k|x_0)} > t\}} \middle| X_0, X_i \right] \right] \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K \left(\frac{x_0 - u_1}{h_n} \right) K \left(\frac{x_0 - u_2}{h_n} \right) \\
&\quad \int_{sU_Y(n/k|x_0)}^{\infty} \int_{tU_Y(n/k|x_0)}^{\infty} \log \frac{v_1}{sU_Y(n/k|x_0)} \log \frac{v_2}{tU_Y(n/k|x_0)} dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) \\
&\quad g_{(X_0, X_i)}(u_1, u_2) du_1 du_2.
\end{aligned}$$

Let $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$. Using Assumption (\mathcal{J}) , the inner integral, denoted $\tilde{I}(u_1, u_2)$, can be bounded, for n large, as follows

$$\begin{aligned}
& \tilde{I}(u_1, u_2) \\
&= \int_{sU_Y(n/k|x_0)}^{\infty} \int_{tU_Y(n/k|x_0)}^{\infty} \left(\int_{sU_Y(n/k|x_0)}^{v_1} \frac{1}{z_1} dz_1 \right) \left(\int_{tU_Y(n/k|x_0)}^{v_2} \frac{1}{z_2} dz_2 \right) dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) \\
&= \int_{sU_Y(n/k|x_0)}^{\infty} \int_{tU_Y(n/k|x_0)}^{\infty} \frac{1}{z_1} \frac{1}{z_2} \mathbb{P}(Y_0 > z_1, Y_i > z_2 | X_0 = u_1, X_i = u_2) dz_2 dz_1 \\
&= \int_{sU_Y(n/k|x_0)}^{\infty} \int_{tU_Y(n/k|x_0)}^{\infty} \frac{1}{z_1} \frac{1}{z_2} \frac{\mathbb{P}(Y_0 > z_1, Y_i > z_2 | X_0 = u_1, X_i = u_2)}{\sqrt{\bar{F}_Y(z_1|u_1) \bar{F}_Y(z_2|u_2)}} \\
&\quad \times \sqrt{\bar{F}_Y(z_1|u_1) \bar{F}_Y(z_2|u_2)} dz_2 dz_1 \\
&\leq \Omega_{\bar{h}}(\bar{z}|x_0) \int_{sU_Y(n/k|x_0)}^{\infty} \int_{tU_Y(n/k|x_0)}^{\infty} \frac{\sqrt{\bar{F}_Y(z_1|x_0)}}{z_1} \frac{\sqrt{\bar{F}_Y(z_2|x_0)}}{z_2} \\
&\quad \times \sqrt{1 + \left| \frac{\bar{F}_Y(z_1|u_1)}{\bar{F}_Y(z_1|x_0)} - 1 \right|} \times \sqrt{1 + \left| \frac{\bar{F}_Y(z_2|u_2)}{\bar{F}_Y(z_2|x_0)} - 1 \right|} dz_2 dz_1 \\
&\leq C \frac{k}{n} \int_1^{\infty} \frac{1}{z_1} \sqrt{\frac{\bar{F}_Y(sU_Y(n/k|x_0) z_1|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)}} \left\{ 1 + \left| \frac{\bar{F}_Y(sU_Y(n/k|x_0) z_1|u_1)}{\bar{F}_Y(sU_Y(n/k|x_0) z_1|x_0)} - 1 \right| \right\} dz_1 \\
&\quad \times \int_1^{\infty} \frac{1}{z_2} \sqrt{\frac{\bar{F}_Y(tU_Y(n/k|x_0) z_2|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)}} \left\{ 1 + \left| \frac{\bar{F}_Y(tU_Y(n/k|x_0) z_2|u_2)}{\bar{F}_Y(tU_Y(n/k|x_0) z_2|x_0)} - 1 \right| \right\} dz_2 \\
&= O\left(\frac{k}{n}\right),
\end{aligned}$$

uniformly in u_1 and u_2 , by our Lemma 7.1 combined with Potter's bounds (see, e.g., Proposition B.1.9(5) in de Haan and Ferreira, 2006) and the fact that for $x \geq 0$, we have $\sqrt{1+x} \leq 1+x$.

Thus

$$\begin{aligned} & \mathbb{E} \left[\phi_s(Z_{n,0}) \phi_t(Z_{n,i}) \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_{\underline{0}}(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1 - \varepsilon \right. \right] \\ &= O \left(\frac{k h_n^{2d}}{n v_n} \right) \int_{\mathcal{B}_{\underline{0}}(1)} \int_{\mathcal{B}_{\underline{0}}(1)} K(z_1) K(z_2) g_{(X_0, X_i)}(x_0 - z_1 h_n, x_0 - z_2 h_n) dz_1 dz_2 \\ &= O \left(\frac{k}{n v_n} \right) \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n)). \end{aligned}$$

If we come back to the covariance, using Assumption (\mathcal{X}) , we have

$$\begin{aligned} & \frac{1}{r_n v_n} \text{Cov} \left(\sum_{i=1}^{r_n} \phi_s(Z_{n,i}), \sum_{j=1}^{r_n} \phi_t(Z_{n,j}) \right) \\ &= \frac{(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}{V_d} \|K\|_2^2 [\max(s, t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s, t)}{\min(s, t)} \right\} + O(r_n v_n) + o(1) \\ &+ O \left(\frac{k}{n v_n} \right) \left\{ h_n^d \sum_{i=1}^{i_0-1} \frac{\mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n))}{h_n^d} \right. \\ &+ \left. h_n^{2d} \sum_{i=i_0}^{r_n-1} \frac{\mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n))}{h_n^{2d}} \right\} \\ &= \frac{(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}{V_d} \|K\|_2^2 [\max(s, t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s, t)}{\min(s, t)} \right\} \\ &+ O(r_n v_n) + o(1) + o \left(\frac{k}{n v_n} h_n^d \right) + O \left(\frac{k}{n v_n} h_n^{2d} r_n \right) \\ &= \frac{(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}{V_d} \|K\|_2^2 [\max(s, t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s, t)}{\min(s, t)} \right\} + o(1), \end{aligned}$$

under our conditions.

Since

$$\begin{aligned} & \mathbb{E}[\phi_s(U + \underline{2}, V) \phi_t(U + \underline{2}, V)] \\ &= \frac{(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}{V_d} \|K\|_2^2 [\max(s, t)]^{-\frac{1}{\gamma(x_0)}} \left\{ 2\gamma^2(x_0) + \gamma(x_0) \log \frac{\max(s, t)}{\min(s, t)} \right\}, \end{aligned}$$

Assumption (C3) is established. □

7.2.2 Proof of Theorem 3.2

First, we start to prove the convergence (2). To this aim, we look at the following decomposition:

$$\begin{aligned}
& \sqrt{k h_n^d} \left\{ \frac{\frac{1}{n h_n^d} \sum_{i=1}^n K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > s \right\}}}{\hat{g}_n(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0)} - 1 \right\} \\
&= \frac{\sqrt{k h_n^d}}{\hat{g}_n(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0)} \\
&\quad \times \left\{ \frac{1}{n h_n^d} \sum_{i=1}^n K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > s \right\}} - \frac{1}{h_n^d} \mathbb{E} \left[K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > s \right\}} \right] \right\} \\
&\quad + \frac{\sqrt{k h_n^d}}{\hat{g}_n(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0)} \\
&\quad \times \left\{ \frac{1}{h_n^d} \mathbb{E} \left[K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > s \right\}} \right] - g_X(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0) \right\} \\
&\quad - \frac{1}{\hat{g}_n(x_0)} \sqrt{\frac{k}{n}} \sqrt{n h_n^d} \{ \hat{g}_n(x_0) - g_X(x_0) \} \\
&=: T_{6,n} + T_{7,n} + T_{8,n}.
\end{aligned}$$

An extension of Example 3.5 in Drees and Rootzén (2010) to the multivariate framework yields

$$T_{8,n} = o_{\mathbb{P}}(1).$$

Now, according to Lemma 7.1, assuming $h_n^{\eta_\gamma \wedge \eta_\varepsilon} \log(n/k) \rightarrow 0$, we have

$$\begin{aligned}
T_{7,n} &= \frac{\sqrt{k h_n^d}}{\hat{g}_n(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0)} \\
&\quad \times \left\{ \frac{1}{h_n^d} \int_{\mathbb{R}^d} K \left(\frac{x_0 - u}{h_n} \right) \bar{F}_Y(sU_Y(n/k|x_0)|u) g_X(u) du - g_X(x_0) \bar{F}_Y(sU_Y(n/k|x_0)|x_0) \right\} \\
&= \frac{\sqrt{k h_n^d}}{\hat{g}_n(x_0)} \left\{ \int_{\mathcal{B}_{\underline{0}}(1)} K(z) \frac{\bar{F}_Y(sU_Y(n/k|x_0)|x_0 - z h_n)}{\bar{F}_Y(sU_Y(n/k|x_0)|x_0)} g_X(x_0 - z h_n) dz - g_X(x_0) \right\} \\
&= \frac{\sqrt{k h_n^d}}{\hat{g}_n(x_0)} \left\{ \int_{\mathcal{B}_{\underline{0}}(1)} K(z) [g_X(x_0 - z h_n) - g_X(x_0)] dz \right. \\
&\quad \left. + \int_{\mathcal{B}_{\underline{0}}(1)} K(z) \left[\frac{\bar{F}_Y(sU_Y(n/k|x_0)|x_0 - z h_n)}{\bar{F}_Y(sU_Y(n/k|x_0)|x_0)} - 1 \right] g_X(x_0 - z h_n) dz \right\}
\end{aligned}$$

$$\begin{aligned}
&= O_{\mathbb{P}}\left(\sqrt{k h_n^d} h_n^{\eta_g}\right) + O_{\mathbb{P}}\left(\sqrt{k h_n^d} h_n^{\eta_A}\right) + O_{\mathbb{P}}\left(\sqrt{k h_n^d} h_n^{\eta_\gamma} \log \frac{n}{k}\right) \\
&\quad + O_{\mathbb{P}}\left(\sqrt{k h_n^d} \left| \delta_Y\left(U_Y\left(\frac{n}{k}|x_0\right)|x_0\right) \right| h_n^{\eta_B}\right) + O_{\mathbb{P}}\left(\sqrt{k h_n^d} \left| \delta_Y\left(U_Y\left(\frac{n}{k}|x_0\right)|x_0\right) \right| h_n^{\eta_\xi} \log \frac{n}{k}\right) \\
&= o_{\mathbb{P}}(1).
\end{aligned}$$

Now, remark that

$$\begin{aligned}
T_{6,n} &= s^{\frac{1}{\gamma(x_0)}} \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} (1 + o_{\mathbb{P}}(1)) \\
&\quad \times \frac{1}{\sqrt{n} v_n} \sum_{i=1}^n \left\{ K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}} - \mathbb{E}\left[K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}\right] \right\} \\
&\rightsquigarrow s^{\frac{1}{\gamma(x_0)}} \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_s),
\end{aligned}$$

by our Theorem 3.1. Combining the above results, we deduce that

$$\sqrt{k h_n^d} \left\{ \frac{\frac{1}{n h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}}{\widehat{g}_n(x_0) \overline{F}_Y(s U_Y(n/k|x_0)|x_0)} - 1 \right\} \rightsquigarrow s^{\frac{1}{\gamma(x_0)}} \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_s). \quad (9)$$

Thus

$$\begin{aligned}
&\sqrt{k h_n^d} \left\{ \frac{\frac{1}{k h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}}{\widehat{g}_n(x_0)} - s^{-\frac{1}{\gamma(x_0)}} \right\} \\
&= \frac{n}{k} \overline{F}_Y(s U_Y(n/k|x_0)|x_0) \sqrt{k h_n^d} \left\{ \frac{\frac{1}{n h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}}{\widehat{g}_n(x_0) \overline{F}_Y(s U_Y(n/k|x_0)|x_0)} \right. \\
&\quad \left. - \frac{k}{n} \frac{s^{-\frac{1}{\gamma(x_0)}}}{\overline{F}_Y(s U_Y(n/k|x_0)|x_0)} \right\} \\
&= s^{-\frac{1}{\gamma(x_0)}} \sqrt{k h_n^d} \left\{ \frac{\frac{1}{n h_n^d} \sum_{i=1}^n K\left(\frac{x_0 - X_i}{h_n}\right) \mathbb{1}_{\left\{\frac{Y_i}{U_Y(n/k|x_0)} > s\right\}}}{\widehat{g}_n(x_0) \overline{F}_Y(s U_Y(n/k|x_0)|x_0)} - 1 \right\} (1 + o(1)) \\
&\quad + O\left(\sqrt{k h_n^d} \delta_Y\left(U_Y\left(\frac{n}{k}|x_0\right)|x_0\right)\right), \quad (10)
\end{aligned}$$

Combining (9) with (10), the convergence (2) follows.

Note that we introduced $\check{T}_n(s)$ since we want to deduce a convergence result similar to (2) but for the quantile. To this aim, we will apply Vervaat's Lemma, and thus we need non-decreasing functions. Since, by the Skorohod construction, but without change in notation,

$$\sqrt{k h_n^d} \left\{ \check{T}_n(s) - \left(-s^{-\frac{1}{\gamma(x_0)}} \right) \right\} \rightarrow -\sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_s), \quad a.s.$$

uniformly on $[1-\varepsilon, 1+\varepsilon]$, and hence, by Vervaat's Lemma

$$\sqrt{k h_n^d} \left\{ \check{T}_n^\leftarrow(t) - (-t)^{-\gamma(x_0)} \right\} \rightarrow \gamma(x_0) (-t)^{-\gamma(x_0)-1} \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\psi_{(-t)^{-\gamma(x_0)}}), \quad a.s.$$

uniformly on $\left[-(1-\varepsilon)^{-\frac{1}{\gamma(x_0)}}, -(1+\varepsilon)^{-\frac{1}{\gamma(x_0)}} \right]$. Now, using the specific value $t = -1$, Theorem 3.2 follows. \square

7.2.3 Proof of Theorem 3.3

We consider the following decomposition:

$$\begin{aligned} & \sqrt{k h_n^d} (\hat{\gamma}_k(x_0) - \gamma(x_0)) \\ &= \frac{1}{\hat{g}_n(x_0)} \sqrt{k h_n^d} \left\{ \frac{1}{k h_n^d} \sum_{i=1}^n K \left(\frac{x_0 - X_i}{h_n} \right) \log \frac{Y_i}{U_Y(n/k|x_0) \check{T}_n^\leftarrow(-1)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > \check{T}_n^\leftarrow(-1) \right\}} \right. \\ & \quad \left. - \frac{n}{k h_n^d} \mathbb{E} \left[K \left(\frac{x_0 - X}{h_n} \right) \log \frac{Y}{U_Y(n/k|x_0) \check{T}_n^\leftarrow(-1)} \mathbb{1}_{\left\{ \frac{Y}{U_Y(n/k|x_0)} > \check{T}_n^\leftarrow(-1) \right\}} \right] \right\} \\ & + \frac{1}{\hat{g}_n(x_0)} \sqrt{k h_n^d} \left\{ \frac{n}{k h_n^d} \mathbb{E} \left[K \left(\frac{x_0 - X}{h_n} \right) \log \frac{Y}{U_Y(n/k|x_0) \check{T}_n^\leftarrow(-1)} \mathbb{1}_{\left\{ \frac{Y}{U_Y(n/k|x_0)} > \check{T}_n^\leftarrow(-1) \right\}} \right] \right. \\ & \quad \left. - \gamma(x_0) g_X(x_0) \right\} \\ & - \frac{\gamma(x_0)}{\hat{g}_n(x_0)} \sqrt{\frac{k}{n}} \sqrt{n h_n^d} \{ \hat{g}_n(x_0) - g_X(x_0) \} \\ & =: T_{9,n} + T_{10,n} + T_{11,n}. \end{aligned}$$

As noticed in the proof of Theorem 3.2, we have $T_{11,n} = o_{\mathbb{P}}(1)$.

Concerning $T_{9,n}$, by Theorem 3.1 and Theorem 3.2, we have

$$\left(Z_n, \sqrt{k h_n^d} \left\{ \check{T}_n^\leftarrow(-1) - 1 \right\} \right) \rightsquigarrow \left(Z_\varepsilon, \gamma(x_0) \sqrt{\frac{V_d}{g_X(x_0)(1-\varepsilon)^{1/\gamma(x_0)}}} Z_\varepsilon(\psi_1) \right).$$

Using Skorohod's theorem, there exists also versions for which the convergence is almost surely. We will apply the Skorohod theorem throughout the proof, though without a change of notation. For n large

$$\left| Z_n \left(\phi_{\check{T}_n^{\leftarrow}(-1)} \right) - Z_\varepsilon(\phi_1) \right| \leq \sup_{s \in [1-\varepsilon, 1+\varepsilon]} |Z_n(\phi_s) - Z_\varepsilon(\phi_s)| + \left| Z_\varepsilon \left(\phi_{\check{T}_n^{\leftarrow}(-1)} \right) - Z_\varepsilon(\phi_1) \right| \longrightarrow 0,$$

almost surely as $n \rightarrow \infty$. This implies that

$$T_{9,n} = \frac{1}{\hat{g}_n(x_0)} \sqrt{\frac{nv_n}{kh_n^d}} Z_n \left(\phi_{\check{T}_n^{\leftarrow}(-1)} \right) \rightsquigarrow \sqrt{\frac{V_d}{g_X(x_0) (1-\varepsilon)^{\frac{1}{\gamma(x_0)}}}} Z_\varepsilon(\phi_1).$$

Concerning $T_{10,n}$, we follow the lines of proof of $T_{5,n}$ and deduce that

$$\begin{aligned} & \frac{n}{k h_n^d} \mathbb{E} \left[K \left(\frac{x_0 - X}{h_n} \right) \log \frac{Y}{sU_Y(n/k|x_0)} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > s \right\}} \right] \\ &= \int_{\mathcal{B}_0(1)} K(z) g_X(x_0 - zh_n) dz \int_1^\infty \frac{\bar{F}_Y(sU_Y(n/k|x_0) y|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} \frac{1}{y} dy \\ & \quad + \int_{\mathcal{B}_0(1)} K(z) \int_1^\infty \frac{\bar{F}_Y(sU_Y(n/k|x_0) y|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} \frac{1}{y} \left[\frac{\bar{F}_Y(sU_Y(n/k|x_0) y|x_0 - zh_n)}{\bar{F}_Y(sU_Y(n/k|x_0) y|x_0)} - 1 \right] dy \\ & \quad g_X(x_0 - zh_n) dz \\ &=: T_{10,n}^{(1)}(s) + T_{10,n}^{(2)}(s). \end{aligned}$$

Thus, under our Assumption (D), we have

$$\begin{aligned} & T_{10,n}^{(1)} \left(\check{T}_n^{\leftarrow}(-1) \right) \\ &= \left\{ g_X(x_0) + \int_{\mathcal{B}_0(1)} K(z) [g_X(x_0 - zh_n) - g_X(x_0)] dz \right\} \left\{ \int_1^\infty \left(\check{T}_n^{\leftarrow}(-1) y \right)^{-\frac{1}{\gamma(x_0)}} \frac{1}{y} dy \right. \\ & \quad \left. + \int_1^\infty \left[\frac{\bar{F}_Y(\check{T}_n^{\leftarrow}(-1) U_Y(n/k|x_0) y|x_0)}{\bar{F}_Y(U_Y(n/k|x_0)|x_0)} - \left(\check{T}_n^{\leftarrow}(-1) y \right)^{-\frac{1}{\gamma(x_0)}} \right] \frac{1}{y} dy \right\} \\ &= \left\{ g_X(x_0) + O(h_n^{\eta_g}) \right\} \left\{ \gamma(x_0) \left(\check{T}_n^{\leftarrow}(-1) \right)^{-\frac{1}{\gamma(x_0)}} \right. \\ & \quad \left. + \int_1^\infty \left[\frac{1 + \frac{1}{\gamma(x_0)} \delta_Y \left(\check{T}_n^{\leftarrow}(-1) U_Y(n/k|x_0) y|x_0 \right)}{1 + \frac{1}{\gamma(x_0)} \delta_Y (U_Y(n/k|x_0)|x_0)} - 1 \right] \left(\check{T}_n^{\leftarrow}(-1) y \right)^{-\frac{1}{\gamma(x_0)}} \frac{1}{y} dy \right\} \end{aligned}$$

$$\begin{aligned}
&= \left\{ g_X(x_0) + O(h_n^{\eta_g}) \right\} \left\{ \gamma(x_0) \left(\check{T}_n^{\leftarrow}(-1) \right)^{-\frac{1}{\gamma(x_0)}} \right. \\
&\quad \left. + \frac{\delta_Y(U_Y(n/k|x_0)|x_0)}{\gamma(x_0) + \delta_Y(U_Y(n/k|x_0)|x_0)} \left(\check{T}_n^{\leftarrow}(-1) \right)^{-\frac{1}{\gamma(x_0)}} \right. \\
&\quad \left. \times \int_1^\infty \left[\frac{\delta_Y(\check{T}_n^{\leftarrow}(-1) U_Y(n/k|x_0) y | x_0)}{\delta_Y(U_Y(n/k|x_0)|x_0)} - 1 \right] y^{-\frac{1}{\gamma(x_0)} - 1} dy \right\} \\
&= \left\{ g_X(x_0) + O(h_n^{\eta_g}) \right\} \left\{ \gamma(x_0) \left(\check{T}_n^{\leftarrow}(-1) \right)^{-\frac{1}{\gamma(x_0)}} + O_{\mathbb{P}} \left(\delta_Y \left(U_Y \left(\frac{n}{k} | x_0 \right) | x_0 \right) \right) \right\} \\
&= \gamma(x_0) g_X(x_0) \left(\check{T}_n^{\leftarrow}(-1) \right)^{-\frac{1}{\gamma(x_0)}} + O(h_n^{\eta_g}) + O_{\mathbb{P}} \left(\delta_Y \left(U_Y \left(\frac{n}{k} | x_0 \right) | x_0 \right) \right),
\end{aligned}$$

by Proposition B.1.10 in de Haan and Ferreira (2006) and Theorem 3.2. Also, using Lemma 7.1, we have

$$T_{10,n}^{(2)} \left(\check{T}_n^{\leftarrow}(-1) \right) = O_{\mathbb{P}}(h_n^{\eta_A}) + O_{\mathbb{P}} \left(h_n^{\eta_\gamma} \log \frac{n}{k} \right) + o_{\mathbb{P}} \left(\delta_Y \left(U_Y \left(\frac{n}{k} | x_0 \right) | x_0 \right) \right).$$

Consequently

$$\begin{aligned}
T_{10,n} &= -\sqrt{k h_n^d} \left\{ \check{T}_n^{\leftarrow}(-1) - 1 \right\} + \frac{1}{\hat{g}_n(x_0)} \sqrt{k h_n^d} T_{10,n}^{(2)} \left(\check{T}_n^{\leftarrow}(-1) \right) + o_{\mathbb{P}}(1) \\
&= -\sqrt{k h_n^d} \left\{ \check{T}_n^{\leftarrow}(-1) - 1 \right\} + o_{\mathbb{P}}(1).
\end{aligned}$$

Combining the above results with Theorem 3.2, the proof of Theorem 3.3 is achieved. \square

7.2.4 Proof of Theorem 3.4

We use the following decomposition:

$$\begin{aligned}
\frac{\sqrt{k h_n^d}}{\log \frac{k}{np}} \left\{ \frac{\widehat{U}_Y \left(\frac{1}{p} \middle| x_0 \right)}{U_Y \left(\frac{1}{p} \middle| x_0 \right)} - 1 \right\} &= \frac{\sqrt{k h_n^d}}{\log \frac{k}{np}} \left\{ \frac{\widehat{U}_Y \left(\frac{n}{k} \middle| x_0 \right)}{U_Y \left(\frac{n}{k} \middle| x_0 \right)} \left(\frac{k}{np} \right)^{\widehat{\gamma}_k(x_0) - \gamma(x_0)} \frac{1 + a_Y \left(\frac{n}{k} \middle| x_0 \right)}{1 + a_Y \left(\frac{1}{p} \middle| x_0 \right)} - 1 \right\} \\
&= \frac{1}{\log \frac{k}{np}} \sqrt{k h_n^d} \left\{ \frac{\widehat{U}_Y \left(\frac{n}{k} \middle| x_0 \right)}{U_Y \left(\frac{n}{k} \middle| x_0 \right)} - 1 \right\} \left(\frac{k}{np} \right)^{\widehat{\gamma}_k(x_0) - \gamma(x_0)} \\
&\quad \times \frac{1 + a_Y \left(\frac{n}{k} \middle| x_0 \right)}{1 + a_Y \left(\frac{1}{p} \middle| x_0 \right)} \\
&\quad + \frac{\sqrt{k h_n^d}}{\log \frac{k}{np}} \left\{ \left(\frac{k}{np} \right)^{\widehat{\gamma}_k(x_0) - \gamma(x_0)} - 1 \right\} \frac{1 + a_Y \left(\frac{n}{k} \middle| x_0 \right)}{1 + a_Y \left(\frac{1}{p} \middle| x_0 \right)} \\
&\quad - \frac{1}{1 + a_Y \left(\frac{1}{p} \middle| x_0 \right)} \frac{\sqrt{k h_n^d} a_Y \left(\frac{n}{k} \middle| x_0 \right)}{\log \frac{k}{np}} \left\{ \frac{a_Y \left(\frac{1}{p} \middle| x_0 \right)}{a_Y \left(\frac{n}{k} \middle| x_0 \right)} - 1 \right\} \\
&=: T_{12,n} + T_{13,n} + T_{14,n}.
\end{aligned}$$

Clearly, by our Theorems 3.2 and 3.3 and the assumptions $k/(np) \rightarrow \infty$ and $\log(k/(np))/\sqrt{k h_n^d} \rightarrow 0$, we have

$$\begin{aligned}
T_{12,n} &= o_{\mathbb{P}}(1) \\
T_{13,n} &= \sqrt{k h_n^d} \{ \widehat{\gamma}_k(x_0) - \gamma(x_0) \} + o_{\mathbb{P}}(1).
\end{aligned}$$

Concerning $T_{14,n}$, we can use the facts that $a_Y(y|x_0) = \delta_Y(U_Y(y|x_0)|x_0)(1+o(1))$ and $|a_Y(\cdot|x_0)| \in RV_{-\beta(x_0)\gamma(x_0)}$ to conclude that $T_{14,n} = o_{\mathbb{P}}(1)$. Combining these results, Theorem 3.4 follows. \square

7.2.5 Proof of Lemma 7.1

Let C be a positive constant whose value can change from one line to another. Using Assumption (\mathcal{D}) , we have

$$\begin{aligned}
\left| \frac{\overline{F}_Y(y_n|x)}{\overline{F}_Y(y_n|x_0)} - 1 \right| &= \left| \frac{A(x)}{A(x_0)} y_n^{\frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)}} \frac{1 + \frac{1}{\gamma(x)} \delta_Y(y_n|x)}{1 + \frac{1}{\gamma(x_0)} \delta_Y(y_n|x_0)} - 1 \right| \\
&\leq \left| \frac{A(x)}{A(x_0)} - 1 \right| y_n^{\frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)}} \left| \frac{1 + \frac{1}{\gamma(x)} \delta_Y(y_n|x)}{1 + \frac{1}{\gamma(x_0)} \delta_Y(y_n|x_0)} \right| \\
&\quad + \left| y_n^{\frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)}} - 1 \right| \left| \frac{1 + \frac{1}{\gamma(x)} \delta_Y(y_n|x)}{1 + \frac{1}{\gamma(x_0)} \delta_Y(y_n|x_0)} \right| \\
&\quad + \left| \frac{1 + \frac{1}{\gamma(x)} \delta_Y(y_n|x)}{1 + \frac{1}{\gamma(x_0)} \delta_Y(y_n|x_0)} - 1 \right|. \tag{11}
\end{aligned}$$

According to Assumption (\mathcal{H}) , we have

$$\left| \frac{A(x)}{A(x_0)} - 1 \right| \leq C h_n^{\eta_A} \text{ for } x \in \mathcal{B}_{x_0}(h_n). \tag{12}$$

Using the inequality $|e^z - 1| \leq |z| e^{|z|}$, we have for $x \in \mathcal{B}_{x_0}(h_n)$ and n large

$$\begin{aligned}
\left| y_n^{\frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)}} - 1 \right| &\leq \left| \frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)} \right| \log y_n e^{|\frac{1}{\gamma(x_0)} - \frac{1}{\gamma(x)}| \log y_n} \\
&\leq C y_n^{C h_n^{\eta_\gamma}} h_n^{\eta_\gamma} \log y_n. \tag{13}
\end{aligned}$$

Finally, for n large

$$\begin{aligned}
&\left| \frac{1 + \frac{1}{\gamma(x)} \delta_Y(y_n|x)}{1 + \frac{1}{\gamma(x_0)} \delta_Y(y_n|x_0)} - 1 \right| \\
&\leq C |\delta_Y(y_n|x_0)| \left\{ \left| \frac{\gamma(x_0)}{\gamma(x)} - 1 \right| \left| \frac{\delta_Y(y_n|x)}{\delta_Y(y_n|x_0)} \right| + \left| \frac{\delta_Y(y_n|x)}{\delta_Y(y_n|x_0)} - 1 \right| \right\} \\
&\leq C |\delta_Y(y_n|x_0)| h_n^{\eta_\gamma} \left| \frac{\delta_Y(y_n|x)}{\delta_Y(y_n|x_0)} \right| + C |\delta_Y(y_n|x_0)| \left| \frac{B(x)}{B(x_0)} - 1 \right| e^{\int_1^{y_n} \frac{\xi_Y(z|x) - \xi_Y(z|x_0)}{z} dz} \\
&\quad + C |\delta_Y(y_n|x_0)| \left| e^{\int_1^{y_n} \frac{\xi_Y(z|x) - \xi_Y(z|x_0)}{z} dz} - 1 \right| \\
&\leq C |\delta_Y(y_n|x_0)| \left\{ h_n^{\eta_\gamma} + h_n^{\eta_B} + y_n^{C h_n^{\eta_\xi}} h_n^{\eta_\xi} \log y_n \right\}. \tag{14}
\end{aligned}$$

Combining (11)-(14), Lemma 7.1 follows. \square

7.2.6 Proof of Lemma 7.2.

We have

$$\begin{aligned}
& \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \max \left(\log \frac{Y_0}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \max \left(\log \frac{Y_i}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \right. \\
& \quad \left. \times \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \left| \frac{x_0 - X_0}{h_n} \in \mathcal{B}_0(1) \text{ and } \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right. \right] \\
&= \frac{1}{v_n} \mathbb{E} \left[K \left(\frac{x_0 - X_0}{h_n} \right) K \left(\frac{x_0 - X_i}{h_n} \right) \mathbb{E} \left\{ \max \left(\log \frac{Y_0}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \right. \right. \\
& \quad \left. \left. \times \max \left(\log \frac{Y_i}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \mathbb{1}_{\left\{ \frac{Y_0}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \mathbb{1}_{\left\{ \frac{Y_i}{U_Y(n/k|x_0)} > 1-\varepsilon \right\}} \left| X_0, X_i \right\} \right] \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K \left(\frac{x_0 - u_1}{h_n} \right) K \left(\frac{x_0 - u_2}{h_n} \right) \\
& \quad \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \max \left(\log \frac{v_1}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \max \left(\log \frac{v_2}{(1-\varepsilon)U_Y(n/k|x_0)}, 1 \right) \\
& \quad \quad \quad dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) g_{(X_0, X_i)}(u_1, u_2) du_1 du_2 \\
&= \frac{1}{v_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K \left(\frac{x_0 - u_1}{h_n} \right) K \left(\frac{x_0 - u_2}{h_n} \right) \sum_{j=1}^4 I_j(u_1, u_2) g_{(X_0, X_i)}(u_1, u_2) du_1 du_2 \tag{15}
\end{aligned}$$

where

$$\begin{aligned}
I_1(u_1, u_2) &:= \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e^{(1-\varepsilon)U_Y(n/k|x_0)}} \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e^{(1-\varepsilon)U_Y(n/k|x_0)}} dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) \\
I_2(u_1, u_2) &:= \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e^{(1-\varepsilon)U_Y(n/k|x_0)}} \int_{e^{(1-\varepsilon)U_Y(n/k|x_0)}}^{\infty} \log \frac{v_2}{(1-\varepsilon)U_Y(n/k|x_0)} dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) \\
I_3(u_1, u_2) &:= \int_{e^{(1-\varepsilon)U_Y(n/k|x_0)}}^{\infty} \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e^{(1-\varepsilon)U_Y(n/k|x_0)}} \log \frac{v_1}{(1-\varepsilon)U_Y(n/k|x_0)} dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2) \\
I_4(u_1, u_2) &:= \int_{e^{(1-\varepsilon)U_Y(n/k|x_0)}}^{\infty} \int_{e^{(1-\varepsilon)U_Y(n/k|x_0)}}^{\infty} \log \frac{v_1}{(1-\varepsilon)U_Y(n/k|x_0)} \log \frac{v_2}{(1-\varepsilon)U_Y(n/k|x_0)} \\
& \quad \quad \quad dF_{(Y_0, Y_i)}(v_1, v_2 | u_1, u_2).
\end{aligned}$$

We will study each of the terms separately.

Term $I_1(u_1, u_2)$. We have

$$\begin{aligned}
I_1(u_1, u_2) &= \mathbb{P}(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > (1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\quad - \mathbb{P}(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > e(1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\quad - \mathbb{P}(Y_0 > e(1 - \varepsilon)U_Y(n/k|x_0), Y_i > (1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\quad + \mathbb{P}(Y_0 > e(1 - \varepsilon)U_Y(n/k|x_0), Y_i > e(1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2).
\end{aligned}$$

These four terms are all similar so we only consider the first one. For n large, by Assumption (\mathcal{J}) and Lemma 7.1, we obtain

$$\begin{aligned}
&\mathbb{P}(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > (1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\leq \Omega_{\bar{h}}(\bar{y}|x_0)\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0) \frac{\sqrt{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|u_1)\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|u_2)}}{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0)} \\
&\leq C \frac{k}{n},
\end{aligned}$$

uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$. Hence

$$I_1(u_1, u_2) \leq C \frac{k}{n},$$

uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$.

Term $I_j(u_1, u_2)$ for $j = 2, 3$. These two terms can be handled similarly, thus we give only the proof for $I_2(u_1, u_2)$. By some straightforward manipulations we have

$$\begin{aligned}
&I_2(u_1, u_2) \\
&= \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e(1-\varepsilon)U_Y(n/k|x_0)} \int_{e(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{v_2} \frac{1}{z_2} dz_2 dF_{(Y_0, Y_i)}(v_1, v_2|u_1, u_2) \\
&= \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{e(1-\varepsilon)U_Y(n/k|x_0)} \int_{\max(e(1-\varepsilon)U_Y(n/k|x_0), z_2)}^{\infty} dF_{(Y_0, Y_i)}(v_1, v_2|u_1, u_2) dz_2 \\
&= \int_{(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} \\
&\quad \times [\mathbb{P}(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > \max(e(1 - \varepsilon)U_Y(n/k|x_0), z_2)|X_0 = u_1, X_i = u_2) \\
&\quad - \mathbb{P}(Y_0 > e(1 - \varepsilon)U_Y(n/k|x_0), Y_i > \max(e(1 - \varepsilon)U_Y(n/k|x_0), z_2)|X_0 = u_1, X_i = u_2)] dz_2
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{P}(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > e(1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\quad - \mathbb{P}(Y_0 > e(1 - \varepsilon)U_Y(n/k|x_0), Y_i > e(1 - \varepsilon)U_Y(n/k|x_0)|X_0 = u_1, X_i = u_2) \\
&\quad + \int_{e(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} P(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > z_2|X_0 = u_1, X_i = u_2) dz_2 \\
&\quad - \int_{e(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} P(Y_0 > e(1 - \varepsilon)U_Y(n/k|x_0), Y_i > z_2|X_0 = u_1, X_i = u_2) dz_2.
\end{aligned}$$

The two probabilities in the right-hand-side of the above display were already studied in $I_1(u_1, u_2)$ and can be dominated by $C k/n$ uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$. The two integrals are similar so we give only the arguments for the first one. Again, for n large, by Assumption (\mathcal{J}) and Lemma 7.1, we obtain

$$\begin{aligned}
&\int_{e(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} P(Y_0 > (1 - \varepsilon)U_Y(n/k|x_0), Y_i > z_2|X_0 = u_1, X_i = u_2) dz_2 \\
&\leq \Omega_{\bar{h}}(\bar{y}|x_0) \sqrt{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0)} \sqrt{\frac{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|u_1)}{\bar{F}_Y((1 - \varepsilon)U_Y(n/k|x_0)|x_0)}} \\
&\quad \times \int_{e(1-\varepsilon)U_Y(n/k|x_0)}^{\infty} \frac{1}{z_2} \sqrt{\bar{F}_Y(z_2|x_0)} \sqrt{\frac{\bar{F}_Y(z_2|u_2)}{\bar{F}_Y(z_2|x_0)}} dz_2 \\
&\leq C \frac{k}{n} \left\{ \int_1^{\infty} \frac{1}{z_2} \sqrt{\frac{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)z_2|x_0)}{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)|x_0)}} dz_2 \right. \\
&\quad \left. + \int_1^{\infty} \frac{1}{z_2} \sqrt{\frac{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)z_2|x_0)}{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)|x_0)}} \left| \frac{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)z_2|u_2)}{\bar{F}_Y(e(1 - \varepsilon)U_Y(n/k|x_0)z_2|x_0)} - 1 \right| dz_2 \right\} \\
&\leq C \frac{k}{n},
\end{aligned}$$

uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$, where the last step follows from the use of Potter bounds (see Proposition B.1.9.(5) in de Haan and Ferreira, 2006) and Lemma 7.1. All by all, we get then

$$I_2(u_1, u_2) \leq C \frac{k}{n},$$

uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$.

Term $I_4(u_1, u_2)$.

The term $I_4(u_1, u_2)$ has been studied in the proof of Assumption $(C3)$, where we show that if Assumption (\mathcal{J}) is satisfied, then $I_4(u_1, u_2) \leq C k/n$ uniformly in $u_1, u_2 \in \mathcal{B}_{x_0}(h_n)$.

Plugging these results related to $I_j(u_1, u_2)$ back into (15), we can choose $s_n(i)$ as

$$s_n(i) = C \frac{k}{n v_n} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n)).$$

This implies that

$$\begin{aligned} \sum_{i=0}^{r_n-1} s_n(i) &= C \frac{k}{n v_n} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n)) \\ &+ C \frac{k h_n^d}{n v_n} \sum_{i=1}^{i_0-1} \frac{1}{h_n^d} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n)) \\ &+ C \frac{k h_n^{2d}}{n v_n} \sum_{i_0}^{r_n-1} \frac{1}{h_n^{2d}} \mathbb{P}(X_0 \in \mathcal{B}_{x_0}(h_n), X_i \in \mathcal{B}_{x_0}(h_n)), \end{aligned}$$

and hence, by Assumption (\mathcal{X}) , we can conclude the first statement of Lemma 7.2.

The second statement of the lemma can be proven with similar arguments and is for brevity omitted here. \square

Acknowledgement

The research of Yuri Goegebeur was supported by a grant of the Novo Nordisk Foundation (NNF25OC0107031). The research of Armelle Guillou was supported by the CNRS grant International Emerging Action IEA-111. Jing Qin is supported by the Novo Nordisk Recruit Grant 2024 (Project number - 3310201). The authors sincerely thank the editor, associate editor and referees for their helpful comments on our paper.

References

- Chen, B.X., Wu, W.B. & Yi, Y. (2009). Efficient estimation of copula-based semiparametric Markov models, *The Annals of Statistics*, **37**, 4214–4253.
- Daouia, A., Gardes, L., Girard, S. & Lekina, A. (2011). Kernel estimators of extreme level curves, *Test*, **20**, 311–333.

Daouia, A., Stupfler, G. & Usseglio-Carleve, A. (2023). Inference for extremal regression with dependent heavy-tailed data, *The Annals of Statistics*, 51(5), 2040–2066.

de Haan, L. & Ferreira, A. (2006). *Extreme value theory, an introduction*. Springer.

Dierckx, G., Goegebeur, Y. & Guillou, A. (2014). Local robust and asymptotically unbiased estimation of conditional Pareto-type tails, *Test*, 23, 330–355.

Doukhan, P. (1994). *Mixing: Properties and Examples*. Lecture Notes in Statistics 85, Springer-Verlag.

Drees, H. & Knežević, M. (2020). Peak-over-threshold estimators for spectral tail processes: random vs deterministic thresholds, *Extremes*, 23, 465–491.

Drees, H. & Rootzén, H. (2010). Limit theorems for empirical processes of cluster functionals, *The Annals of Statistics*, 38(4), 2145–2186.

Drees, H., Segers, J. & Warchol M. (2015). Statistics for tail processes of Markov chains, *Extremes*, 18, 369–402.

Givens, G.H. & Hoeting, J.A. (2005). *Computational Statistics*. Wiley.

Goegebeur, Y., Guillou, A., Ho, N.K.L. & Qin, J. (2023). A Weissman-type estimator of the conditional marginal expected shortfall, *Econometrics and Statistics*, 27, 173–196.

Goegebeur, Y., Guillou, A. & Qin, J. (2019). Bias-corrected estimation for conditional Pareto-type distributions with random right censoring, *Extremes*, 22, 459–498.

Goegebeur, Y., Guillou, A. & Qin, J. (2025). A conditional tail expectation type risk measure

for time series, *Journal of Time Series Analysis*, <https://doi.org/10.1111/jtsa.12853>.

Goegebeur, Y., Guillou, A. & Schorgen, A. (2014). Nonparametric regression estimation of conditional tails - the random covariate case, *Statistics*, 48, 732–755.

Hill, B.M. (1975). A simple general approach to inference about the tail of a distribution, *The Annals of Statistics*, 3(5), 1163–1174.

Meyn, S.P. & Tweedie, R.L. (1993). *Markov chains and stochastic stability*. Cambridge University Press.

Wang, H. & Li, D. (2013). Estimation of extreme conditional quantiles through power transformation, *Journal of the American Statistical Association*, 108, 1062–1074.

Wang, H. & Tsai, C.L. (2009). Tail index regression, *Journal of the American Statistical Association*, 104, 1233–1240.

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