HAL
open science

Nonparametric estimation of the conditional tail copula

Laurent Gardes, Stéphane Girard

To cite this version:

Laurent Gardes, Stéphane Girard. Nonparametric estimation of the conditional tail copula. Journal of Multivariate Analysis, 2015, 137, pp.1-16. 10.1016/j.jmva.2015.01.018 . hal-00964514v3

HAL Id: hal-00964514

https://hal.science/hal-00964514v3

Submitted on 20 Jan 2015

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.

Nonparametric estimation of the conditional tail copula

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

[(1)] Université de Strasbourg & CNRS, IRMA, UMR 7501, 7 rue René Descartes,
67084 Strasbourg cedex, France.
[⋆] Corresponding author, gardes@unistra.fr
[(2)] Team Mistis, Inria Grenoble Rhône-Alpes & Laboratoire Jean Kuntzmann,
655 avenue de l'Europe, Montbonnot, 38334 Saint-Ismier cedex, France.

Abstract

The tail copula is widely used to describe the dependence in the tail of multivariate distributions. In some situations such as risk management, the dependence structure may be linked with some covariate. The tail copula thus depends on this covariate and is referred to as the conditional tail copula. The aim of this paper is to propose a nonparametric estimator of the conditional tail copula and to establish its asymptotic normality. Some illustrations are presented both on simulated and real datasets.

AMS Subject Classifications: 62G05, 62G20, 62G30, 62G32.

Keywords: Extremal dependence, Tail copula, Covariate.

1 Introduction

In various domains such as insurance, finance and hydrology, there is increasing interest in understanding the dependence structure between large values of several variables. For instance, in finance, one can be interested in the link between large values of the market index and large returns of individual stocks, see [2, 30]. The dependence structure of a random vector $Y = (Y_1, \ldots, Y_d)^\top$ with $d \geq 2$ is commonly modeled using a copula. Recall that a copula is a multivariate distribution function with uniformly distributed margins on the interval $[0, 1]$. Let us denote by $\bar{F}_i$ the marginal survival functions of Y, $i = 1, \ldots, d$. From Sklar's theorem [37], if the margins of Y are continuous, then there exists a unique copula S defined on $(0, 1)^d$ such that $\bar{F}(y) := \mathbb{P}(Y \geq y) = S(\bar{F}_1(y_1), \ldots, \bar{F}_d(y_d))$ for all $y = (y_1, \ldots, y_d)^\top \in \mathbb{R}^d$. This function S is therefore given by $S(u) = \bar{F}(q(u))$ for all $u = (u_1, \ldots, u_d)^\top \in (0, 1)^d$, where $q(u) := (\bar{F}_1^{\leftarrow}(u_1), \ldots, \bar{F}_d^{\leftarrow}(u_d))^\top$ is a multivariate quantile. Note that $(a_1, \ldots, a_d) \geq (b_1, \ldots, b_d)$ corresponds to the componentwise inequality $a_i \geq b_i$ for all $i = 1, \ldots, d$. Additionally, $f^{\leftarrow}$ denotes the generalized inverse of a decreasing

function f *i.e*, $f^{\leftarrow}(x) = \inf\{t,\ f(t) \leq x\}$. As mentioned in [2], the tail copula is a useful tool to describe the upper tail dependence structure. It is defined as follows: if there exists a function $\Lambda:\ (0,\infty)^d \to [0,\infty)$ such that

$$\lim_{\alpha\to 0^+} \frac{1}{\alpha} S(\alpha y) = \Lambda(y),\ y \in (0,\infty)^d, \tag{1}$$

then Λ is called the tail copula [34]. Let us highlight that, since the domain of definition of Λ is not limited to the unit hypercube, Λ is not a proper copula. However, many properties of Λ are closely related to copula properties [34, Theorems 1, 2]. Moreover, in view of (1), it appears that Λ is a homogeneous positive function. Besides, S is a multivariate regularly varying function at zero since

$$\lim_{\alpha\to 0^+} \frac{S(\alpha y)}{S(\alpha \mathbf{1})} = W(y),$$

where $W(\cdot) = \Lambda(\cdot)/\Lambda(\mathbf{1})$ and $\mathbf{1} = (1,\ldots,1)^\top \in \mathbb{R}^d$. We refer to [33, Chapter 5] for a definition of multivariate regularly varying functions and to [4] for a general account on regular variation. In practice, the tail copula is often reduced to the tail-dependence coefficient defined by

$$\lambda := \Lambda(\mathbf{1}) = \lim_{\alpha\to 0^+} \frac{1}{\alpha} \mathbb{P}\left\{Y_i \geq \bar{F}_i^{\leftarrow}(\alpha),\ \forall i\right\} = \lim_{\alpha\to 0^+} \mathbb{P}\left\{Y_i \geq \bar{F}_i^{\leftarrow}(\alpha),\ i \neq j \middle| Y_j \geq \bar{F}_j^{\leftarrow}(\alpha)\right\}.$$

It was first introduced by [36] in the bivariate case. The multivariate random vector Y is said to be tail-dependent if $\lambda \in (0,1]$ and tail-independent if $\lambda = 0$. The case $\lambda = 1$ corresponds to complete tail dependence. When $\lambda = 0$, large values of the margins Y_i are unlikely to happen simultaneously. For example, it is well known that a bivariate normal distribution with correlation coefficient different from ± 1 is tail-independent. In finance, λ can be interpreted as the limiting likelihood that a return falls below its Value at Risk, given that another portfolio return has fallen below its Value at Risk at the same level. However, since λ is only a scalar measure of dependence, it may suffer from a loss of information considering the joint behavior of the distribution tails. In contrast, the tail copula S describes the full dependence structure. We also refer to [28] for an application of modeling log returns in crude oil and natural gas futures.

The bivariate case ($d = 2$) has been widely studied in the literature especially when the distribution function of Y belongs to the maximum domain of attraction of a bivariate extreme-value distribution G. In such a case, there exist positive sequences $(a_{1,n})$, $(a_{2,n})$ and sequences $(b_{1,n})$ and $(b_{2,n})$ such that

$$\lim_{n\to\infty} \mathbb{P}\left\{\max_{j=1,\ldots,n} \frac{Y_{1,j} - b_{1,n}}{a_{1,n}} \leq y_1, \max_{j=1,\ldots,n} \frac{Y_{2,j} - b_{2,n}}{a_{2,n}} \leq y_2\right\} = G(y_1, y_2),$$

where $(Y_{1,j}, Y_{2,j})$, $j = 1,\ldots,n$ are independent copies of the random vector $Y = (Y_1, Y_2)$. The bivariate extreme-value distribution G can be written as $G(y_1, y_2) = D(G_1(y_1), G_2(y_2))$, where G_1 and G_2 are the margins of G and D is the bivariate extreme-value copula defined by:

$$D(u,v) = \exp\left\{\ln(uv) A\left(\frac{\ln u}{\ln(uv)}\right)\right\}. \tag{2}$$

The marginal distributions G_1 and G_2 are univariate extreme value distributions *i.e.* for $i \in \{1, 2\}$,

$$G_i(x) = \exp\{-(1 + \gamma_i x)_+^{-1/\gamma_i}\},$$

where $y_+ = \max(0, y)$ and $\gamma_i \in \mathbb{R}$ is called the extreme-value index. The function $A(t)$ is referred to as the Pickands dependence function. It is a convex function lying in a triangular area *i.e.* such that $\max(t, (1-t)) \leq A(t) \leq 1$ for all $t \in [0, 1]$. It can be shown [33, Chapter 5] that for a bivariate distribution function in the maximum domain of attraction of G, the tail copula exists and is given by

$$\Lambda(y_1, y_2) = (y_1 + y_2) \left\{ 1 - A\left(\frac{y_1}{y_1 + y_2} \right) \right\}. \tag{3}$$

Collecting (2) and (3), the bivariate extreme-value copula can be rewritten in terms of the tail copula as

$$D(u, v) = uv \exp\left(\Lambda(-\ln u, -\ln v)\right). \tag{4}$$

As a consequence, in the bivariate case, the tail copula and the Pickands dependence function both fully characterize the extreme dependence structure of Y. Finally, one may equivalently consider the so-called stable tail-dependence function given by

$$\ell(y_1, y_2) = y_1 + y_2 - \Lambda(y_1, y_2) \tag{5}$$

to describe the extreme dependence structure.

In the bivariate case, numerous nonparametric estimators of the Pickands dependence function have been introduced [5, 15, 32, 40] thanks to an interpretation of A in terms of exponential distributions. In [25], it is noted that such estimators may not provide proper dependence functions (convex and lying in the above triangular area) and two new estimators satisfying these constraints are introduced. Similarly, a projection technique is proposed in [16] to force any estimator to be a proper dependence function. The above mentioned estimators are constructed assuming knowledge of the marginal distributions. Rank-based versions of these estimators are proposed for the unknown margins case in [19]. The use of a graphical tool for detecting extreme value dependence is investigated in [9]. The estimation of A within the class of Archimax copulas is studied in [6]. See [3, Chapter 9] for an overview of parametric and nonparametric inference methods for bivariate Pickands dependence functions. We also refer to [23, 24] for extensions to the general multivariate setting. In view of (5), it is clear that estimating the stable tail-dependence function or the tail copula are equivalent problems. First estimators of ℓ were introduced in [27] in the bivariate case. They have been adapted to the estimation of the tail copula [34] and compared to nonparametric estimators based on the empirical copula. Parametric estimators of Λ have also been proposed in the bivariate case [13] and for multivariate elliptical copulas [29].
The case where the random vector Y is recorded simultaneously with a predictive variable $X \in \mathbb{R}^p$ has been less considered. In [39], the authors propose to estimate the tail-dependence coefficient in the regression case using a maximum likelihood approach. In this regression framework, S is

a conditional copula. Its estimation is addressed in [22, 38] with nonparametric techniques and in [1] using semiparametric approaches. The case of functional covariates is studied in [21]. To our knowledge, the estimation of the tail copula in the regression case has not been considered yet despite potential applications in the study of the Chinese stock market [39] or the modeling of financial returns [17].

We thus consider a random vector $(X, Y) \in \mathbb{R}^p \times \mathbb{R}^d$. For $i = 1, \ldots, d$, the conditional survival function of Y_i given $X = x$ is denoted by $\bar{F}_i(\cdot|x) := \mathbb{P}(Y_i \geq \cdot|X = x)$. The dependence structure of Y conditionally on $X = x$ is then characterized by the conditional copula:

$$S(u|x) = \bar{F}\left(q(u|x)|x\right), \; u = (u_1, \ldots, u_d) \in \mathbb{R}^d,$$

where $\bar{F}(\cdot|x) : \mathbb{R}^d \to [0, 1]$ is the conditional survival function of Y given $X = x$ and $q(\cdot|x)$ is the multivariate conditional quantile defined as $q(u|x) := (\bar{F}_1^{\leftarrow}(u_1|x), \ldots, \bar{F}_d^{\leftarrow}(u_d|x))^\top$. Similarly to (1), the conditional tail copula is thus defined by

$$\Lambda(y|x) := \lim_{\alpha \to 0^+} \frac{1}{\alpha} S\left(\alpha y|\, x\right), \; y \in (0, \infty)^d, \tag{6}$$

if this limit exists. In the following, we focus on the case $\Lambda(y|x) > 0$ *i.e.* tail-dependent distributions.

In Section 2, nonparametric estimators of $\Lambda(y|x)$ are introduced for a fixed vector $(x, y) \in \mathbb{R}^p \times (0, \infty)^d$ in both situations where the margins are known and unknown. Their asymptotic properties are established in Section 3. A simulation study is presented in Section 4 and an illustration on real data is proposed in Section 5. Concluding remarks are drawn in Section 6. Finally, proofs of the main results are provided in Section 7.

2 Estimation of the conditional tail copula

Let (X_j, Y_j), $j = 1, \ldots, n$ be independent copies of the random vector (X, Y). The estimator of the conditional tail copula is based on a nonparametric estimator of the conditional copula function of Y given $X = x$. We propose to use a classical kernel estimator. The first step consists of estimating the probability density function g of X by:

$$\hat{g}_h(x) = \frac{1}{nh^p} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h}\right), \; x \in \mathbb{R}^p,$$

where $h = h_{n,x}$ is a non random sequence tending to zero as n goes to infinity and K is a bounded density on $\mathbb{R}^p$ with support $\mathcal{S}$ included in the unit ball of $\mathbb{R}^p$. Note that the sequence h may depend on x. As mentioned in the introduction, in many related papers, the margins are supposed to be known. In this situation, the conditional copula $S(.|x)$ can be estimated by

$$\widetilde{S}_h(u|x) := \frac{1}{nh^p \hat{g}_h(x)} \sum_{j=1}^{n} \mathbb{I}\{Y_j \geq q(u|x)\} K\left(\frac{x - X_j}{h}\right), \; u \in (0, 1)^d, \; x \in \mathbb{R}^p,$$

where $\mathbb{I}\{.\}$ is the indicator function. An estimator of the tail copula is then given by:

$$\tilde{\Lambda}_h(y|x) = \frac{1}{\alpha}\widetilde{S}_h(\alpha y|x), \tag{7}$$

where $y \in (0,\infty)^d$ and $\alpha = \alpha_{n,x}$ is a $(0,1)$-valued sequence tending to zero as n goes to infinity. In this paper, we consider the more realistic case where nothing is known on the margins. In this situation, the estimation of the marginal survival functions of Y given $X = x$ is obviously required. For all $u \in (0,1)^d$, let us denote by

$$\hat{q}(u|x) = (\hat{q}_1(u_1|x), \ldots, \hat{q}_d(u_d|x))^\top$$

an estimator of the vector $q(u|x) \in \mathbb{R}^d$. The following plug-in estimator of the conditional copula is considered:

$$\hat{S}_h(u|x) := \frac{1}{nh^p\hat{g}_h(x)}\sum_{j=1}^{n}\mathbb{I}\{Y_j \geq \hat{q}(u|x)\}K\left(\frac{x - X_j}{h}\right),\ u \in (0,1)^d,\ x \in \mathbb{R}^p$$

and the conditional tail copula is then estimated by:

$$\hat{\Lambda}_h(y|x) = \frac{1}{\alpha}\widehat{S}_h(\alpha y|x). \tag{8}$$

As an example of an estimator of $q(u|x)$, one can propose the statistic

$$\check{q}_k(u|x) := (\hat{\bar{F}}^{\leftarrow}_{k,1}(u_1|x), \ldots, \hat{\bar{F}}^{\leftarrow}_{k,d}(u_d|x))^\top, \tag{9}$$

where $k = k_{n,x}$ is a non random sequence tending to zero as n goes to infinity and, for $i = 1, \ldots, d$, $\hat{\bar{F}}_{k,i}(\cdot|x)$ is the usual kernel estimator (with bandwidth k) of a conditional survival function:

$$\hat{\bar{F}}_{k,i}(z|x) = \frac{1}{nk^p\hat{g}_k(x)}\sum_{j=1}^{n}\mathbb{I}\{Y_{i,j} \geq z\}K\left(\frac{x - X_j}{k}\right).$$

The obtained estimators of the conditional copula and of the tail copula are respectively denoted by $\check{S}_{h,k}(\cdot|x)$ and $\check{\Lambda}_{h,k}(\cdot|x)$. Let us highlight that these estimators depend on two bandwidths h and k. The first one tunes the smoothness of the estimator of the conditional copula, while the second one controls the smoothness of the estimator of the marginal distributions.

3 Main results

In this section, the asymptotic properties of the estimators $\tilde{\Lambda}_h(y|x)$ and $\hat{\Lambda}_h(y|x)$ are established for a fixed vector $(x, y) \in \mathbb{R}^p \times (0,\infty)^d$ such that $\Lambda(y|x) > 0$. The asymptotic behavior of $\check{\Lambda}_{h,k}(\cdot|x)$ is also analyzed in the two situations $h = k$ and $h \neq k$. Some regularity and extreme-value assumptions are required. The first one is a classical second-order condition on the conditional copula function $S(\cdot|x)$, see for instance [34].

H.1 There exist a function $b(\cdot|x)$ with $b(t|x) \to 0$ as $t \to \infty$ and a function $c_\Lambda(\cdot|x)$ such that, for all $y \in (0,\infty)^d$,

$$\lim_{t\to\infty} \frac{\Lambda(y|x) - tS(t^{-1}y|x)}{b(t|x)} = c_\Lambda(y|x) < \infty.$$

Condition **H.1** establishes that the rate of convergence in (6) is driven by a given function $b(\cdot|x)$. Let us recall that, necessarily, $|b|(\cdot|x)$ is a regularly varying function at infinity with index $\rho(x) < 0$, see [14]. In the extreme-value literature, $\rho(x)$ is referred to as the conditional second order parameter. As an example, let us consider the multivariate extreme-value copula defined for all $u = (u_1, \ldots, u_d)^\top$ by

$$\mathbb{P}(Y \leq q(1-u|x)|X = x) = \exp\left\{\kappa(u) A\left(\frac{\ln u_i}{\kappa(u)},\ i = 1, \ldots, d-1 \middle| x\right)\right\}, \tag{10}$$

where

$$\kappa(u) = \sum_{l=1}^{d} \ln u_l,$$

and $A(.|x) :\ \mathbb{R}^{d-1} \to [1/2, 1]$ is the conditional multivariate Pickands dependence function [32]. Clearly, (10) is the extension of the bivariate extreme-value copula (2) to a dimension $d > 2$. It can be shown that, in the dependent case (*i.e.* $A(.|x) \neq 1$) and if $A(.|x)$ is continuously differentiable, then condition **H.1** is satisfied with $\rho(x) = -1$.

A Lipschitz condition on the probability density function g of X is also required. Denoting by $\tilde{d}(x, x')$ the Euclidean distance between $x \in \mathbb{R}^p$ and $x' \in \mathbb{R}^p$, it is assumed that

H.2 There exists a positive constant c_g such that $|g(x) - g(x')| \leq c_g \tilde{d}(x, x')$.

For $x \in \mathbb{R}^p$ and $u \in (0,1)^d$, the oscillation of the conditional survival function $\bar{F}(.|.)$ is controlled in a neighborhood of $q(u|x)$ by the following quantity:

$$\omega_{h,\zeta}(u|x) = \sup\left\{\left|\frac{\bar{F}(q(v|x)|t)}{S(v|x)} - 1\right| \text{ for } t \in B(x,h),\ (1-\zeta)u \leq v \leq (1+\zeta)u\right\}, \tag{11}$$

where $\zeta \in [0,1)$ and $B(x,h)$ is the ball of $\mathbb{R}^p$ centered at x with radius h.

In a first time, conditions **H.1** and **H.2** are sufficient to establish the asymptotic normality of the estimator $\tilde{\Lambda}_h(y|x)$ in the (unrealistic) situation where the margins are known.

Theorem 1. *Suppose* **H.1** *and* **H.2** *hold. Let $x \in \mathbb{R}^p$ and $y \in (0,\infty)^d$ such that $g(x) \neq 0$ and $\Lambda(y|x) \neq 0$. If α in (7) is such that $\alpha \to 0$, $nh^p\alpha \to \infty$ and*

$$nh^p\alpha \max\left(h, \omega_{h,0}(\alpha y|x), b(\alpha^{-1}|x)\right)^2 \to 0 \tag{12}$$

as $n \to \infty$, then,

$$(nh^p\alpha)^{1/2}\left(\tilde{\Lambda}_h(y|x) - \Lambda(y|x)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2 \Lambda(y|x)}{g(x)}\right).$$

As pointed out in [10], $nh^p\alpha \to \infty$ is a necessary and sufficient condition for the almost sure presence of at least one sample point in the region $B(x,h) \times (q(\alpha|x), \infty)$ of $\mathbb{R}^{p+d}$. Condition (12) forces the bias to be negligible compared to the standard-deviation of the estimator. Two main sources of bias appear. The first contribution to the bias is due to the kernel smoothing and to the oscillations of order $\mathcal{O}(h)$ and $\mathcal{O}(\omega_{h,0}(\alpha y|x))$ of g and $\bar{F}(\cdot|\cdot)$ respectively. The second source of bias is due to the approximation of the conditional tail copula $\Lambda(y|x)$ by $\alpha S(\alpha^{-1}y|x)$. This entails a bias of order $\mathcal{O}(b(\alpha^{-1}|x))$, see assumption **H.1**. Let us also highlight that the rate of convergence $(nh^p\alpha)^{1/2}$ strongly decreases as the dimension p of the covariate increases. This phenomena can be interpreted as a theoretical consequence of the "curse of dimensionality" for kernel-smoothing methods. In practice, for moderate sample sizes, such nonparametric methods are limited to the case $p \leq 2$. To overcome this limitation, one can use dimension reduction techniques such as ADE (Average Derivative Estimator), see for instance [26].
In order to deal with the case where the margins are estimated, an extreme-value assumption on the conditional marginal distributions of Y given $X = x$ is required. The following von-Mises condition (see [12], equation (1.1.30)) is considered:

H.3 For $i = 1, \ldots, d$, the marginal survival function $\bar{F}_i(\cdot|x)$ is twice differentiable and such that

$$\lim_{y \uparrow y_{F_i}(x)} \frac{\bar{F}_i(y|x)\bar{F}_i''(y|x)}{(\bar{F}_i')^2(y|x)} = \gamma_i(x) + 1,$$

where $y_{F_i}(x) := \bar{F}_i^{\leftarrow}(0|x) \in (-\infty, \infty]$ is the right endpoint of the $i-$th margin, $\bar{F}_i'(\cdot|x)$ and $\bar{F}_i''(\cdot|x)$ are respectively the first and the second derivatives of $\bar{F}_i(\cdot|x)$.

From [12], Theorem 1.1.8, condition **H.3** entails that the conditional distribution of Y_i given $X = x$ is in the domain of attraction of the univariate extreme-value distribution with conditional extreme-value index $\gamma_i(x) \in \mathbb{R}$ for all $i = 1, \ldots, d$. Note also that from [11], Lemma 2, if condition **H.3** holds and $z_{n,i}(x) \to y_{F_i}(x)$, $t_{n,i}(x) \to t_{0,i}(x)$ as $n \to \infty$ for all $i = 1, \ldots, d$, then there exists a positive auxiliary function $a_i(\cdot|x)$ such that

$$\lim_{n\to\infty} \frac{\bar{F}_i(z_{n,i}(x) + t_{n,i}(x)a_i(z_{n,i}(x)|x)|x)}{\bar{F}_i(z_{n,i}(x)|x)} = (1 + \gamma_i(x)t_{0,i}(x))^{-1/\gamma_i(x)}.$$

Hence, for all sequences $\alpha \to 0$ and $t_{n,i}(x) \to 0$ as $n \to \infty$, choosing $z_{n,i}(x) = \bar{F}_i^{\leftarrow}(\alpha|x) \to y_{F_i}(x)$ as $\alpha \to 0$ yields

$$\lim_{n\to\infty} \alpha^{-1}\bar{F}_i\left(\bar{F}_i^{\leftarrow}(\alpha|x) + t_{n,i}(x)a_i(\bar{F}_i^{\leftarrow}(\alpha|x)|x)\big|\, x\right) = 1. \tag{13}$$

The next assumption allows to control the rate of convergence of the estimator $\hat{q}(\cdot|x)$ towards the true quantile function $q(\cdot|x) = \bar{F}^{\leftarrow}(\cdot|x)$.

H.4 There exists a sequence $\eta_n \to \infty$ such that, for all $y = (y_1, \ldots, y_d)^\top \in (0,\infty)^d$,

$$\eta_n \max_{i=1,\ldots,d} \frac{\bar{F}_i^{\leftarrow}(y_i\alpha|x)}{a_i(\bar{F}_i^{\leftarrow}(y_i\alpha|x)|x)} \left| \frac{\hat{q}_i(y_i\alpha|x)}{\bar{F}_i^{\leftarrow}(y_i\alpha|x)} - 1 \right| = \mathcal{O}_{\mathbb{P}}(1).$$

A condition on the gradient $\nabla S(\cdot|x)$ of the conditional copula function is also introduced. It states that convergence (6) still holds when differentiating.

H.5 The partial derivatives of the conditional copula $S(\cdot|x)$ exist and are continuous on $(0,1)^d$. Furthermore, for $y \in (0,\infty)^d$ and for all vectors $\vartheta(\alpha, y)$ such that $\vartheta(\alpha, y) = \alpha y(1 + o(1))$ as $\alpha \to 0^+$,

$$\lim_{\alpha \to 0^+} y^\top \nabla S(\vartheta(\alpha, y)|x) = \Lambda(y|x).$$

The existence of continuous partial derivatives was shown to be a non restrictive assumption in [35]. **H.5** is for instance satisfied by the multivariate extreme-value copula (10) with $A(\cdot|x) \neq 1$ and when $A(\cdot|x)$ is continuously differentiable on $(0,1)^d$.

We are now in position to state our main result on the asymptotic behavior of $\hat{\Lambda}_h(y|x)$.

Theorem 2. *Suppose* **H.1** - **H.5** *hold. Let $x \in \mathbb{R}^p$ and $y \in (0,\infty)^d$ such that $g(x) \neq 0$ and $\Lambda(y|x) \neq 0$.*
(i) If α in (8) is such that $\alpha \to 0$, $nh^p\alpha\eta_n^{-1} \to \infty$ and, for some $\zeta > 0$,

$$nh^p\alpha \max\left(h, \eta_n^{1/2}\omega_{h,\zeta}(\alpha y|x), b(\alpha^{-1}|x)\right)^2 \to 0, \tag{14}$$

as $n \to \infty$, then

$$\hat{\Lambda}_h(y|x) - \Lambda(y|x) = (nh^p\alpha)^{-1/2}\xi_n + \mathcal{O}_{\mathbb{P}}\left(\eta_n^{-1}\right),$$

where ξ_n converges in distribution to a $\mathcal{N}(0, \|K\|_2^2\Lambda(y|x)/g(x))$ random variable.
(ii) If, moreover, $(nh^p\alpha)^{1/2}\eta_n^{-1} \to 0$, then

$$(nh^p\alpha)^{1/2}(\hat{\Lambda}_h(y|x) - \Lambda(y|x)) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2\Lambda(y|x)}{g(x)}\right).$$

Note that Theorem 2(i) yields $\hat{\Lambda}_h(y|x) \xrightarrow{\mathbb{P}} \Lambda(y|x)$. In situation (ii), condition $(nh^p\alpha)^{1/2}\eta_n^{-1} \to 0$ entails that the estimator of $q(.|x)$ converges faster than the one of the conditional copula $S(.|x)$. Thus, $\hat{\Lambda}_h(y|x)$ has the same asymptotic behavior as $\tilde{\Lambda}_h(y|x)$. In this case, the estimation of the margins does not change the asymptotic distribution.

As an illustration, let us consider the situation where the conditional multivariate quantile $q(u|x)$ is estimated by (9). The regularity of $\bar{F}_i(\cdot|\cdot)$ is controlled by

$$\bar{\omega}_{h,\zeta}^{(i)}(u|x) = \sup\left\{\left|\frac{\bar{F}_i(\bar{F}_i^{\leftarrow}(v_i|x)|t)}{v_i} - 1\right| \text{ for } t \in B(x,h) \text{ and } (1-\zeta)u_i \leq v_i \leq (1+\zeta)u_i\right\}$$

where $u \in (0,1)^d$ and $\zeta \in [0,1)$. Introducing $(e_1, \ldots, e_d)$ the canonical basis of $\mathbb{R}^d$, the above quantity can be rewritten using (11) as $\bar{\omega}_{h,\zeta}^{(i)}(u|x) = \omega_{h,\zeta}((u-1)e_i + \mathbf{1}|x)$. The situations where the same bandwidth or different bandwidths are used to estimate the conditional copula and the margins are considered separately. Let us first focus on the case $h = k$. The asymptotic behavior of $\check{\Lambda}_{h,h}(y|x)$ is a consequence of Theorem 2(i).

Corollary 1. *Suppose* **H.1**, **H.2**, **H.3** *and* **H.5** *hold. Let* $x \in \mathbb{R}^p$ *and* $y \in (0,\infty)^d$ *such that* $g(x) \neq 0$ *and* $\Lambda(y|x) \neq 0$. *If* α *in (8) is such that* $\alpha \to 0$, $nh^p\alpha \to \infty$ *and, for some* $\zeta > 0$,

$$nh^p\alpha \max\left(h, b(\alpha^{-1}|x), \omega_{h,\zeta}^{2/3}(\alpha y|x), \max_{i=1,\dots,d}\left(\bar{\omega}_{h,\zeta}^{(i)}(y_i\alpha|x)\right)\right)^2 \to 0,$$

as $n \to \infty$, *then*

$$\check{\Lambda}_{h,h}(y|x) - \Lambda(y|x) = \mathcal{O}_{\mathbb{P}}\left((nh^p\alpha)^{-1/2}\right).$$

Here, the estimators of $q(\cdot|x)$ and $S(\cdot|x)$ share the same bandwidth and thus the same rate of convergence $(nh^p\alpha)^{-1/2}$. Hence, plugging-in (9) in $\tilde{\Lambda}_h(y|x)$ does not change this rate of convergence. Finally, in the case $h \neq k$, the asymptotic behavior of $\check{\Lambda}_{h,k}(y|x)$ is a consequence of Theorem 2(ii).

Corollary 2. *Suppose* **H.1**, **H.2**, **H.3** *and* **H.5** *hold. Let* $x \in \mathbb{R}^p$ *and* $y \in (0,\infty)^d$ *such that* $g(x) \neq 0$ *and* $\Lambda(y|x) \neq 0$. *If* $h/k \to 0$ *and* α *in (8) is such that* $\alpha \to 0$, $nh^p\alpha \to \infty$, $nk^{p+2}\alpha \to 0$ *and, for some* $\zeta > 0$,

$$nh^p\alpha \max\left(b(\alpha^{-1}|x), \omega_{h,\zeta}^{2/3}(\alpha y|x), \left(\frac{k}{h}\right)^{p/2} \max_{i=1,\dots,d} \bar{\omega}_{h,\zeta}^{(i)}(y_i\alpha|x)\right)^2 \to 0,$$

as $n \to \infty$, *then*

$$(nh^p\alpha)^{1/2}(\check{\Lambda}_{h,k}(y|x) - \Lambda(y|x)) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2\Lambda(y|x)}{g(x)}\right).$$

Note that condition $h/k \to 0$ entails that the rate of convergence of the margin estimators $(nk^p\alpha)^{1/2}$ is faster than the one of the conditional copula estimator $(nh^p\alpha)^{1/2}$. The asymptotic normality of the plug-in estimator can thus be established and the obtained rate of convergence is $(nh^p\alpha)^{1/2}$.

4 Illustration on simulated data

The estimation of the conditional tail copula is investigated on a two-dimensional ($d = 2$) simulated dataset where the marginal distributions are supposed to be unknown. We limit ourselves to a univariate covariate ($p = 1$) and we focus on the estimator $\check{\Lambda}_{h,h}(\cdot|x)$ studied in Corollary 1. The conditional margins are estimated by (9) and the smoothness of both the estimators of the conditional copula and of the marginal distributions is tuned by the same bandwidth h and by the same bi-quadratic kernel given for $x \in \mathbb{R}$ by:

$$K(x) := \frac{15}{16}(1 - x^2)^2\mathbb{I}\{|x| \leq 1\}.$$

The simulated model is described in Paragraph 4.1, implementation details are given in Paragraph 4.2 and results are reported in Paragraph 4.3.

4.1 Model

The covariate $X \in [0,1]$ follows a standard uniform distribution. The conditional distribution of Y_1 given $X = x$ is an exponential distribution with mean parameter $\mu(x) > 0$. Besides, Y_2 given $X = x$ follows a Pareto distribution with tail-index $\gamma(x) > 0$, *i.e.* $\mathbb{P}(Y_2 > y|X = x) = y^{-1/\gamma(x)}$ for all $y \geq 1$. The dependence structure between the margins Y_1 and Y_2 is controlled by the Gumbel copula defined for $(u_1, u_2) \in (0,1)^2$ by:

$$S(u_1, u_2|x) = \exp\left\{-\left[(-\ln(1-u_1))^{\theta(x)} + (-\ln(1-u_2))^{\theta(x)}\right]^{1/\theta(x)}\right\} + u_1 + u_2 - 1,$$

where for all $x \in [0,1]$, $\theta(x) > 1$ tunes the dependence level. If $\theta(x) = 1$, then $S(u_1, u_2|x) = u_1 u_2$ is the independent copula. At the opposite, if $\theta(x) \to +\infty$, then $S(u_1, u_2|x) \to \min\{u_1, u_2\}$ which is the Fréchet upper bound copula for modeling co-monotone random variables. The conditional tail copula is given by:

$$\Lambda(y_1, y_2|x) = y_1 + y_2 - \left(y_1^{\theta(x)} + y_2^{\theta(x)}\right)^{1/\theta(x)}, \tag{15}$$

for all $(y_1, y_2) \in (0,\infty)^2$ and $x \in [0,1]$. The function θ is $\theta(x) = 12x^2 - 12x + 5$ for $x \in [0,1]$, its minimum is $\theta(0.5) = 2$ while its maximum is $\theta(0) = \theta(1) = 5$. Two cases are considered for the functions μ and γ driving the margin distributions:

$$\mu(x) = \gamma(x) = 1 \text{ (Case 1) and } \mu(x) = \gamma(x) + 1/2 = \frac{1}{2}\sin(2\pi x) + 1 \text{ (Case 2)}.$$

Let us note that, in Case 1, the margins do not depend on the covariate. For visualization purposes, only two different values of x are considered: $x = 0.1$ corresponding to a strong dependence and $x = 0.5$ corresponding to a weak dependence. The associated conditional tail copulas are depicted on Figure 1.

4.2 Hyper-parameters selection

The considered conditional tail copula estimator $\check{\Lambda}_{h,h}(.|x)$ depends on two hyper-parameters h and α. These parameters are selected thanks to the homogeneity criterion

$$\mathcal{H}(h,\alpha) := \sum_{t\in\mathcal{T}} \left(\check{\Lambda}_{h,h}(t,t|x) - t\check{\Lambda}_{h,h}(1,1|x)\right)^2 = \frac{1}{\alpha^2}\sum_{t\in\mathcal{T}} \left(\check{S}_{h,h}(t\alpha, t\alpha|x) - t\check{S}_{h,h}(\alpha,\alpha|x)\right)^2,$$

where $\mathcal{T} \subset (0,\infty)$ is a finite set. The selected parameters $h_{sel,x}$ and $\alpha_{sel,x}$ are given by

$$(h_{sel,x}, \alpha_{sel,x}) := \underset{\alpha\in\mathcal{A}, h\in\mathcal{B}}{\arg\min}\, \mathcal{H}(h,\alpha),$$

where $\mathcal{A}$ and $\mathcal{B}$ are finite sets of possible values for α and h respectively. Note that the selected parameters may depend on the value x of the covariate. This selection procedure will be compared

in Section 4.3 to an oracle strategy consisting in taking the values $h_{ora,x}$ and $\alpha_{ora,x}$ that minimize the oracle criterion:

$$\mathcal{H}_{ora}(h,\alpha) := \sum_{t\in\mathcal{T}} \left(\check{\Lambda}_{h,h}(t,t|x) - \Lambda(t,t|x)\right)^2 = \frac{1}{\alpha^2}\sum_{t\in\mathcal{T}} \left(\check{S}_{h,h}(t\alpha,t\alpha|x) - \alpha\Lambda(t,t|x)\right)^2.$$

Of course, this oracle strategy cannot be applied in practical situations since the tail copula $\Lambda(\cdot|x)$ is unknown. The idea behind this homogeneity criterion is that, for well chosen hyper-parameters h and α, $\check{\Lambda}_{h,h}(.|x)$ should provide a good approximation of the conditional tail copula which, among other properties, is an homogeneous function and thus, one can expect the function $\check{\Lambda}_{h,h}(.|x)$ to be approximately homogeneous. At the opposite, a bad choice of h and α can lead to a large difference between $\check{\Lambda}_{h,h}(t,t|x)$ and $t\check{\Lambda}_{h,h}(1,1|x)$ for some values of $t > 0$. In what follows, $\mathcal{T} = \{1/3, 2/3, 1, 4/3, 5/3\}$, $\mathcal{A}$ is the set of 30 regularly spaced values ranging from 0.05 to 0.3 and $\mathcal{B}$ is the set of 30 regularly spaced values ranging from 0.03 to 0.2.

4.3 Results

In this experiment, $N = 100$ independent copies of a sample of size $n = 1000$ are simulated from the model presented in Paragraph 4.1. In order to assess the behavior of the hyper-parameters selection procedure, the distribution over the N replications of the selected hyper-parameters $h_{sel,x}$ and $\alpha_{sel,x}$ is compared with the one of hyper-parameters $h_{ora,x}$ and $\alpha_{ora,x}$ selected by the oracle strategy. The corresponding boxplots are displayed on Figure 2 for Case 1 and $x = 0.1$. The boxplots associated with our procedure and with the oracle strategy are quite similar even if a largest dispersion can be observed for our selection procedure. The median as well as the mean-squared error of the N estimations of $\Lambda(y_1, y_2|x)$ are represented for $(y_1, y_2) \in [0, 3]^2$ respectively on Figures 3 and 4. A numerical summary of Figure 4 is proposed in Table 1 where the numerical values of the mean-squared error associated with the N estimations of $\Lambda(y, y|x)$ are given for $y \in \{1/3, 2/3, 1, 4/3, 5/3, 3\}$. In all cases, the estimations of Figure 3 allow to visually recover the shape of the true conditional tail copula, see Figure 1. Focusing on Figure 4 and Table 1, it appears that the mean-squared error gets larger when the conditional tail copula increases. This phenomena can be explained thanks to Theorem 2(ii): The asymptotic variance is proportional to the conditional tail copula. One can also note that Case 1 (where the margins do not depend on the covariate) yields smaller mean-squared errors than Case 2. Finally, the "boundary effect" due to kernel estimation does not appear in this simulated example. In other situations, one may consider the use of mirror-reflection type estimators [20] or kernel local linear estimators [8].

5 Illustration on real data

In this section, our approach is illustrated on the so-called `pima-indians-diabetes` dataset[1] consisting of nine variables measured on subjects aged from 21 to 81 years. This population was

[1] `http://ftp.ics.uci.edu/pub/machine-learning-databases/pima-indians-diabetes`

studied in order to understand the high prevalence of diabetes in the Pima Indian population. For visualization sake, only two variables are considered: the diastolic blood pressure (DBP, denoted by Y_1) and the body mass index (BMI, denoted by Y_2). It is indeed suspected that high values of DBP and BMI may represent significant risk factors for diabetes [7]. The original dataset contains 768 subjects with missing BMI and/or DBP values for 39 of them. In our illustration, the complete data subset of $n = 729$ subjects is considered. The age is taken as the univariate covariable X and its distribution is depicted on Figure 5. The estimation of the conditional tail copula $\Lambda(y_1, y_2|x)$ permits to precise how the age affects the dependence structure between high values of DBP and BMI. Similarly to Section 4, the conditional tail copula estimator $\check{\Lambda}_{h,h}(.|x)$ is used with the bi-quadratic kernel and the same bandwidth h. Note that both conditional margins are estimated using (9) with a unique bandwidth h even if BMI and DBP have different ranges. Indeed, as a copula function, it can be shown that our conditional tail copula estimator is invariant with respect to strictly monotone transformations and hence, standardizing the variables BMI and DPB would not affect the estimation. We focus on two values for the covariate: $x = 30$ years and $x = 60$ years. This bandwidth h and the sample fraction α are selected using the procedure described in Paragraph 4.2 where $\mathcal{T} = \{1/3, 2/3, 1, 4/3, 5/3\}$, $\mathcal{A}$ is the set of 30 regularly spaced values ranging from 0.07 to 0.5 and $\mathcal{B} = \{1, 2, \ldots, 10\}$. The selected hyper-parameters are given by $(h_{sel,30} = 2, \alpha_{sel,30} = 0.085)$ for $x = 30$ and, for $x = 60$ by $(h_{sel,60} = 9, \alpha_{sel,60} = 0.31)$. The bandwidth selected for $x = 60$ is larger than the one for $x = 30$ which is not surprising since, in view of Figure 5 (left panel), young subjects are more numerous than old ones. The resulting conditional tail copula estimator $\check{\Lambda}_{h,h}(y_1, y_2|x)$ is depicted for $(y_1, y_2) \in [0, 3]^2$ at $x = 30$ years and $x = 60$ years on Figure 6. Clearly, for both values of the age, BMI and DBP are tail dependent. Moreover, it appears that the dependence between large values of BMI and DBP is weaker for young subjects than for old ones. This conclusion was however expected. From (4), one can straightforwardly deduce an estimator of the associated conditional extreme-value distribution: $\check{D}_{h,k}(y_1, y_2|x) = y_1 y_2 \exp\left(\check{\Lambda}_{h,k}(-\ln y_1, -\ln y_2|x)\right)$. This would open the door to the estimation of probabilities of conditional extreme events, such as "BMI and DBP are simultaneously large at a given age".

6 Conclusion

A nonparametric estimator of the conditional tail copula has been proposed so as to capture the extreme dependence of a random vector using an additional information on the variables of interest. Asymptotic properties have been established in the case of known margins and in the more realistic situation of unknown margins. The finite sample properties of the estimator have been illustrated on simulated and real datasets. Our future work would include the theoretical study of the process $x \mapsto \check{\Lambda}_{h,k}(y|x)$ in order to construct confidence regions.

7 Proofs

7.1 Preliminary results

The first lemma is a standard result on the density kernel estimator (see Parzen [31] for a proof).

Lemma 1. *Under* **H.2**, *if* $nh^p \to \infty$ *then for all* $x \in \mathbb{R}^p$ *such that* $g(x) \neq 0$,

$$|\hat{g}_h(x) - g(x)| = \mathcal{O}(h) + \mathcal{O}_{\mathbb{P}}\left((nh^p)^{-1/2}\right).$$

Lemma 2 is dedicated to the random variable $\tilde{S}_h(u|x) - \kappa\tilde{S}_h(v|x)$, where $\kappa \in \{0,1\}$ and $(u,v) \in (0,1)^d \times (0,1)^d$ with $v \leq u$. Clearly, if the margins are known, it can be seen as an estimator of $S(u|x) - \kappa S(v|x)$. The asymptotic normality of this random variable is established in the following result.

Lemma 2. *Suppose* **H.2** *holds and let* $x \in \mathbb{R}^p$ *such that* $g(x) \neq 0$. *Consider two sequences* $u_n \in (0,1)^d$ *and* $v_n \in (0,1)^d$ *converging to zero as* $n \to \infty$ *such that* $nh^p(S(u_n|x) - \kappa S(v_n|x)) \to \infty$. *For some* $\zeta > 0$, *if there exist a sequence* ν_n *such that* $(1-\zeta)\nu_n \leq v_n < u_n < (1+\zeta)\nu_n$ *with*

$$nh^p\left(S(u_n|x) - \kappa S(v_n|x)\right)\max\left(h, \frac{S(u_n|x) + \kappa S(v_n|x)}{S(u_n|x) - \kappa S(v_n|x)} w_{h,\zeta}(\nu_n|x)\right)^2 \to 0,$$

then,

$$\left(nh^p(S(u_n|x) - \kappa S(v_n|x))\right)^{1/2}\left(\frac{\tilde{S}_h(u_n|x) - \kappa\tilde{S}_h(v_n|x)}{S(u_n|x) - \kappa S(v_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2}{g(x)}\right).$$

The proof of this lemma is postponed to the Appendix. Note that if $\kappa = 0$, one can take $u_n = v_n$ and thus Lemma 2 entails that

$$\left(nh^p S(u_n|x)\right)^{1/2}\left(\frac{\tilde{S}_h(u_n|x)}{S(u_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2}{g(x)}\right), \tag{16}$$

and the conditions can be simplified as $nh^p S(u_n|x) \to \infty$ and $nh^p S(u_n|x)\max\left(h, w_{h,\zeta}(\nu_n|x)\right)^2 \to 0$ as $n \to \infty$. Furthermore, in the particular case where $u_n = \mathbf{1} + (\bar{F}_i(z_n|x) - 1)e_i$, $z_n \in (0,1)$, the previous convergence result establishes the asymptotic normality of the kernel estimators of the conditional survival functions $\hat{\bar{F}}_{h,i}(z_n|x)$. Note that this result has also been established in [10, Theorem 1] and [11, Proposition 1]. The case of functional covariates is addressed in [18]. Mimicking the proof of Corollary 1 in [11] thus permits to derive the asymptotic normality of the kernel estimators of the extreme conditional quantiles $\bar{F}_i^{\leftarrow}(\beta_n|x)$.

Lemma 3. *Suppose* **H.2**, **H.3** *hold and let* $x \in \mathbb{R}^p$ *such that* $g(x) \neq 0$. *Consider a sequence* $\beta_n \in (0,1)$ *such that, as* $n \to \infty$, $nh^p\beta_n \to \infty$ *and* $nh^p\beta_n \max(h, (\bar{w}_{h,\zeta}^{(i)}(\beta_n|x)))^2 \to 0$ *for some* $\zeta > 0$. *Then,*

$$(nh^p\beta_n)^{1/2}\frac{\bar{F}_i^{\leftarrow}(\beta_n|x)}{a_i(\bar{F}_i^{\leftarrow}(\beta_n|x)|x)}\left(\frac{\hat{\bar{F}}_{h,i}^{\leftarrow}(\beta_n|x)}{\bar{F}_i^{\leftarrow}(\beta_n|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2}{g(x)}\right).$$

The last lemma will be useful to deal with differences between two indicator functions. It is a straightforward consequence of the monotonicity of the indicator function.

Lemma 4. *Let y, u, v, m and M be vectors of $\mathbb{R}^d$ with, for all $i = 1, \ldots, d$, $M_i > \max(u_i, v_i)$ and $m_i < \min\{u_i, v_i\}$. Then,*

$$\begin{aligned} & |\mathbb{I}\{y_i > u_i,\ i = 1, \ldots, d\} - \mathbb{I}\{y_i > v_i,\ i = 1, \ldots, d\}| \\ \leq\ & \mathbb{I}\{y_i > m_i,\ i = 1, \ldots, d\} - \mathbb{I}\{y_i > M_i,\ i = 1, \ldots, d\}. \end{aligned}$$

7.2 Proof of main results

Proof of Theorem 1 – First, taking $u_n = \alpha y$ in (16), one has under the assumptions of Theorem 1

$$(nh^p\alpha)^{1/2}\left(\frac{\tilde{S}_h(\alpha y|x)}{S(\alpha y|x)} - 1\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\|K\|_2^2}{g(x)\Lambda(y|x)}\right), \tag{17}$$

since $S(\alpha y|x) = \alpha\Lambda(y|x)(1+o(1))$. Furthermore, $(nh^p\alpha)^{1/2}(\tilde{\Lambda}_n(y|x) - \Lambda(y|x)) = T_{1,n} + T_{2,n}$, where

$$\begin{aligned} T_{1,n} &= \left(\frac{nh^p}{\alpha}\right)^{1/2} S(\alpha y|x)\left(\frac{\tilde{S}_h(\alpha y|x)}{S(\alpha y|x)} - 1\right) \\ \text{and } T_{2,n} &= \left(\frac{nh^p}{\alpha}\right)^{1/2} S(\alpha y|x)\left(1 - \frac{\alpha\Lambda(y|x)}{S(\alpha y|x)}\right). \end{aligned}$$

A direct use of (17) entails that

$$T_{1,n} \xrightarrow{d} \mathcal{N}\left(0, \frac{\Lambda(y|x)\|K\|_2^2}{g(x)}\right).$$

Finally, from **H.1**,

$$T_{2,n} = \mathcal{O}\left((nh^p\alpha)^{1/2} b(\alpha^{-1}|x)\right) = o(1),$$

which concludes the proof. ■

Proof of Theorem 2 – We start with the decomposition

$$\hat{\Lambda}_h(y|x) - \Lambda(y|x) = \left(\tilde{\Lambda}_h(y|x) - \Lambda(y|x)\right) + \left(\hat{\Lambda}_h(y|x) - \tilde{\Lambda}_h(y|x)\right).$$

From Theorem 1, the first term of the right-hand side is equal to $(nh^p\alpha)^{-1/2}\xi_n$, where ξ_n converges in distribution to a centered Gaussian variable with variance $\|K\|_2^2\Lambda(y|x)/g(x)$. Let us focus on the second term. From **H.4**, for all $\varepsilon > 0$, there exists a positive constant c_ε such that $\mathbb{P}(\mathcal{A}_n) < \varepsilon/2$ where

$$\mathcal{A}_n = \left\{\eta_n \max_{i=1,\ldots,d} \frac{\bar{F}_i^{\leftarrow}(y_i\alpha|x)}{a_i(\bar{F}_i^{\leftarrow}(y_i\alpha|x)|x)}\left|\frac{\hat{q}_i(y_i\alpha|x)}{\bar{F}_i^{\leftarrow}(y_i\alpha|x)} - 1\right| \geq c_\epsilon\right\}.$$

Furthermore, denoting by $\tilde{q}_n(\vartheta) = (\tilde{q}_{n,1}(\vartheta), \ldots, \tilde{q}_{n,d}(\vartheta))^\top$ with, for $i = 1, \ldots, d$,

$$\tilde{q}_{n,i}(\vartheta) := \bar{F}_i^{\leftarrow}(y_i\alpha|x) + \vartheta\eta_n^{-1}a_i(\bar{F}_i^{\leftarrow}(y_i\alpha|x)|x),$$

Lemma 4 entails that under $\mathcal{A}_n^C$ one has for all $j = 1, \ldots, n$:

$$|\mathbb{I}\{Y_j \geq \hat{q}_n(\alpha y|x)\} - \mathbb{I}\{Y_j \geq q(\alpha y|x)\}| \leq \mathbb{I}\{Y_j \geq \tilde{q}_n(-c_\varepsilon)\} - \mathbb{I}\{Y_j \geq \tilde{q}_n(c_\varepsilon)\}.$$

Thus, under $\mathcal{A}_n^C$,

$$\begin{aligned}\left|\hat{\Lambda}_h(y|x) - \tilde{\Lambda}_h(y|x)\right| &\leq \frac{1}{nh^p\alpha\hat{g}_h(x)}\sum_{j=1}^n |\mathbb{I}\{Y_j \geq \tilde{q}_n(-c_\varepsilon)\} - \mathbb{I}\{Y_j \geq \tilde{q}_n(c_\varepsilon)\}|\, K\left(\frac{x - X_j}{h}\right) \\ &= \alpha^{-1}\left(\tilde{S}_h(\tilde{u}_n(-c_\varepsilon)|x) - \tilde{S}_h(\tilde{u}_n(c_\varepsilon)|x)\right),\end{aligned}$$

with $\tilde{u}_n(\vartheta) = (\bar{F}_1(\tilde{q}_{n,1}(\vartheta)|x), \ldots, \bar{F}_d(\tilde{q}_{n,d}(\vartheta)|x))^\top$. Let us now consider the difference $\tilde{S}_h(\tilde{u}_n(-c_\varepsilon)|x) - \tilde{S}_h(\tilde{u}_n(c_\varepsilon)|x)$ using Lemma 2 with $\kappa = 1$, $u_n = \tilde{u}_n(-c_\varepsilon)$ and $v_n = \tilde{u}_n(c_\varepsilon)$. We thus need to check the following three conditions: there exist $\zeta > 0$ and a sequence ν_n such that

$$(1-\zeta)\nu_n \leq \tilde{u}_n(c_\varepsilon) < \tilde{u}_n(-c_\varepsilon) \leq (1+\zeta)\nu_n, \tag{18}$$

$$nh^p(S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x)) \to \infty, \tag{19}$$

and

$$nh^p\left(S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x)\right)\max\left(h, \frac{S(\tilde{u}_n(-c_\varepsilon)|x) + S(\tilde{u}_n(c_\varepsilon)|x)}{S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x)} w_{h,\zeta}(\nu_n|x)\right)^2 \to 0. \tag{20}$$

To this end, asymptotic equivalents of the sequences $\tilde{u}_n(\pm c_\varepsilon)$ and $S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x)$ are required. First, note that a Taylor expansion yields

$$\begin{aligned}S(\tilde{u}_n(\pm c_\varepsilon)|x) &= S(\alpha y|x) \\ &+ (\tilde{u}_n(\pm c_\varepsilon) - \alpha y)^\top \nabla S\left(\alpha y + \tau(\tilde{u}_n(\pm c_\varepsilon) - \alpha y)\,|x\right)\end{aligned} \tag{21}$$

where $\tau \in (0,1)$. Next, for $i = 1, \ldots, d$, another first order Taylor expansion entails that, for all $i = 1, \ldots, d$,

$$\begin{aligned}\tilde{u}_{n,i}(\pm c_\varepsilon) - \alpha y_i &= \mp F_i^{'}\left(\tilde{q}_{n,i}(\pm\tilde{\tau}_i c_\varepsilon)|\, x\right) c_\varepsilon \eta_n^{-1} a_i(\bar{F}_i^{\leftarrow}(\alpha y_i|x)|x) \\ &= \mp F_i^{'}\left(\bar{F}_i^{\leftarrow}(\tilde{u}_{n,i}(\pm\tilde{\tau}_i c_\varepsilon)|x)\big|\, x\right) c_\varepsilon \eta_n^{-1} a_i(\bar{F}_i^{\leftarrow}(\alpha y_i|x)|x),\end{aligned}$$

for $\tilde{\tau}_i \in (0,1)$. Since for all $i = 1, \ldots, d$, equation (13) entails that $\tilde{u}_{n,i}(\pm\tilde{\tau}_i c_\varepsilon) = y_i\alpha(1 + o(1))$ and since the function $y \mapsto F_i^{'}(\bar{F}_i^{\leftarrow}(y|x)|x)$ is regularly varying at zero with index $\gamma_i(x) + 1$ (see [12], Corollary 1.1.10), one has

$$\tilde{u}_{n,i}(\pm c_\varepsilon) - \alpha y_i = \mp F_i^{'}\left(\bar{F}_i^{\leftarrow}(y_i\alpha|x)|x\right) c_\varepsilon \eta_n^{-1} a_i(\bar{F}_i^{\leftarrow}(\alpha y_i|x)|x)(1 + o(1)).$$

Now, in view of [12], Theorem 1.2.6 and Remark 1.2.7, for $i = 1, \ldots, d$,

$$a_i(y|x) = \frac{\bar{F}_i(y|x)}{F_i^{'}(y|x)}(1 + o(1)) \text{ as } y \to y_{F_i}(x),$$

and thus, for $i = 1, \ldots, d$,

$$\tilde{u}_{n,i}(\pm c_\varepsilon) - \alpha y_i = \mp \alpha c_\varepsilon \eta_n^{-1} y_i (1 + o(1)). \tag{22}$$

Then, condition (18) is satisfied with $\nu_n = \alpha y$. Furthermore, collecting (21) and (22) leads to

$$\begin{aligned} S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x) = \alpha c_\varepsilon \eta_n^{-1} y^\top \quad \Big\{ & \nabla S\left(\alpha y + \tau(\tilde{u}_n(-c_\varepsilon) - \alpha y) \,|x\right) \\ + \quad & \nabla S\left(\alpha y + \tau(\tilde{u}_n(c_\varepsilon) - \alpha y) \,|x\right) \Big\} (1 + o(1)). \end{aligned}$$

Then, under **H.5**,

$$S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x) = 2c_\varepsilon \Lambda(y|x) \alpha \eta_n^{-1} (1 + o(1)), \tag{23}$$

and, in view of $nh^p \alpha \eta_n^{-1} \to \infty$, condition (19) is clearly satisfied. Finally, since $S(\tilde{u}_n(\pm c_\varepsilon)|x) = \alpha \Lambda(y|x)(1 + o(1))$, one has

$$\frac{S(\tilde{u}_n(-c_\varepsilon)|x) + S(\tilde{u}_n(c_\varepsilon)|x)}{S(\tilde{u}_n(-c_\varepsilon)|x) - S(\tilde{u}_n(c_\varepsilon)|x)} = \mathcal{O}_{\mathbb{P}}(\eta_n),$$

and thus, condition (14) implies (20). Lemma 2 can now be used with (23) to obtain

$$\tilde{S}_h(\tilde{u}_n(-c_\varepsilon)|x) - \tilde{S}_h(\tilde{u}_n(c_\varepsilon)|x) = 2c_\varepsilon \Lambda(y|x) \alpha \eta_n^{-1} (1 + o_{\mathbb{P}}(1)).$$

Hence, for all $\varepsilon > 0$, there exists a positive constant c'_ε such that

$$\mathbb{P}\left(\eta_n^{-1} \left| \hat{\Lambda}_h(y|x) - \tilde{\Lambda}_n(y|x) \right| > c'_\varepsilon \,\Big|\, \mathcal{A}_n^C \right) \leq \frac{\varepsilon}{2},$$

and thus

$$\mathbb{P}\left(\eta_n^{-1} \left| \hat{\Lambda}_h(y|x) - \tilde{\Lambda}_n(y|x) \right| > c'_\varepsilon \right) \leq \frac{\varepsilon}{2} + \mathbb{P}(\mathcal{A}_n) \leq \varepsilon,$$

which concludes the proof. ■

Proof of Corollary 1 – Under the assumptions of Lemma 3, the estimator $\hat{q}_n(u|x)$ defined in (9) satisfies condition **H.4** with $\eta_n = (nh^p \alpha)^{1/2}$. Applying Theorem 2(i) concludes the proof. ■

Proof of Corollary 2 – Under the assumptions of Lemma 3, the estimator $\hat{q}_n(u|x)$ defined in (9) satisfies condition **H.4** with $\eta_n = (nk^p \alpha)^{1/2}$. Applying Theorem 2(ii) concludes the proof. ■

References

[1] Abegaz, F., Gijbels, I., and Veraverbeke, N. (2012). Semiparametric estimation of conditional copulas. *Journal of Multivariate Analysis*, **110**, 43–73.

[2] Ané, T. and Kharoubi, C. (2003). Dependence structure and risk measure. *Journal of Business*, **76**, 411–438.

[3] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications.* Wiley.

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

[5] Capéraà, P., Fougères, A.L. and Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme values copulas. *Biometrika*, **84**, 567–577.

[6] Capéraà, P., Fougères, A.L. and Genest, C. (2000). Bivariate distributions with given extreme value attractor. *Journal of Multivariate Analysis*, **72**, 30–49.

[7] Cheng, Y., and De Gooijer, J.G. (2007). On the uth geometric conditional quantile. *Journal of Statistical Planning and Inference*, **137**, 1914–1930.

[8] Cheng, S.C, and Huang, T.M. (2007). Nonparametric estimation of copula functions for dependence modelling. *The Canadial Journal of Statistics*, **35**, 265–282.

[9] Cormier, E., Genest, C. and Nešlehová, J.G. (2014). Using B-splines for nonparametric inference on bivariate extreme-value copulas. *Extremes*, **17**, 633–659

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

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

[12] de Haan, L., and Ferreira, A. (2006). *Extreme value theory: An introduction*, Springer Series in Operations Research and Financial Engineering, Springer.

[13] de Haan, L., Neves, C. and Peng, L. (2008) Parametric tail copula estimation and model testing. *Journal of Multivariate Analysis*, **99**, 1260–1275.

[14] de Haan, L. and Stadtmüller, U. (1996). Generalized regular variation of second order. *Journal of the Australian Mathematical Society*, **61**, 381-395.

[15] Deheuvels, P. (1991). On the limiting behavior of the Pickands estimator for bivariate extreme value distributions. *Statistics and Probability Letters*, **12**, 429–439.

[16] Fils-Villetard, A., Guillou, A. and Segers, J. (2008). Projection estimators of Pickands dependence functions. *Canadian Journal of Statistics*, **36**, 369–382.

[17] Fung, T. and Seneta, E. (2010). Modeling and estimation for bivariate financial returns. *International Statistical Review*, **78**, 117–133.

[18] Gardes, L. and Girard, S. (2013). Functional kernel estimators of large conditional quantiles, *Electronic Journal of Statistics*, **6**, 1715–1744.

[19] Genest, C. and Segers, J. (2009). Rank-based inference for bivariate extreme-value copulas. *The Annals of Statistics*, **37**(5B), 2990-3022.

[20] Gijbels, I. and Mielniczuk, M. (1990). Estimation of the density of a copula function. *Communications in Statistics - Theory and Methods*, **19**, 445–464.

[21] Gijbels, I., Omelka, M., and Veraverbeke, N. (2012). Multivariate and functional covariates and conditional copulas. *Electronic Journal of Statistics*, **6**, 1273–1306.

[22] Gijbels, I., Veraverbeke, N. and Omelka, M. (2011). Conditional copulas, association measures and their applications. *Computational Statistics and Data Analysis*, **55**, 1919–1932.

[23] Gudendorf, G. and Segers, J. (2011). Nonparametric estimation of an extreme-value copula in arbitrary dimensions. *Journal of Multivariate Analysis*, **102**(1), 37-47.

[24] Gudendorf, G. and Segers, J. (2012). Nonparametric estimation of multivariate extreme-value copulas. *Journal of Statistical Planning and Inference*, **142**(12), 3073–3085.

[25] Hall, P. and Tajvidi, N. (2000). Distribution of dependence function estimation for bivariate extreme-value distributions. *Bernoulli*, **6**, 835–844.

[26] Härdle, W., and Stoker, T. (1999). Investigating smooth multiple regression by the method of average derivatives. *Journal of the American Statistical Association*, **84**, 986–995.

[27] Huang, X. (1992). *Statistics of Bivariate Extreme Values*. PhD Thesis, Tinbergen Institute Research Series, Thesis Publishers and Tinbergen Institute.

[28] Jäschke, S., Siburg, K.L. and Stoimenov, P. (2012). Modeling dependence of extreme events in energy markets using tail copulas. *Journal of Energy Markets*, **5**(4), 63–80.

[29] Li, D. and Peng, L. (2009). Goodness-of-fit test for tail copulas modeled by elliptical copulas. *Statistics and Probability Letters*, **79**, 1097–1104.

[30] McNeil, A.J., Frey, R. and Embrechts, P. (2005). *Quantitative risk managements: concepts, techniques, tools*. Princeton University Press, Princeton.

[31] Parzen, E. (1962). On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, **33**, 1065–1076.

[32] Pickands, J. (1981). Multivariate extreme value distribution. *Bulletin of the International Statistical Institute*, 859–878.

[33] Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes*, Springer Verlag, New York.

[34] Schmidt, R. and Stadtmüller, U. (2006). Nonparametric estimation of tail dependence. *The Scandinavian Journal of Statistics*, **33**, 307–335.

[35] Segers, J. (2012). Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. *Bernoulli*, **18**(3), 764–782.

[36] Sibuya, M. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics*, **11**, 195–210.

[37] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Insitut de Statistique de Paris*, **8**, 229–231.

[38] Veraverbeke, N., Omelka, M., and Gijbels, I. (2011). Estimation of a conditional copula and association measures. *Scandinavian Journal of Statistics*, **38**, 766–780.

[39] Zhang, Q., Li, D. and Wang, H. (2013). A note on tail dependence regression. *Journal of Multivariate Analysis*, **120**, 163–172.

[40] Zhang, D., Wells, M.T. and Peng, L. (2008). Nonparametric estimation of the dependence function for a multivariate extreme-value distribution, *Journal of Multivariate Analysis*, **99**, 577–588.

Acknowledgments The authors thank the associate editor and anonymous reviewers for their valuable feedback on this paper.

Appendix

Proof of Lemma 2 – Let us introduce the random variable

$$\tilde{\Psi}_{h,\kappa}(u_n, v_n|x) := \hat{g}_h(x)\left(\tilde{S}_h(u_n|x) - \kappa\tilde{S}_h(v_n|x)\right),$$

where $\kappa \in \{0,1\}$. We first focus on its expectation. Since $(Y_1, \ldots, Y_n)$ are identically distributed, one has

$$\begin{aligned}
\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x)) &= h^{-p}\mathbb{E}\left\{[\mathbb{I}\{Y \geq q(u_n|x)\} - \kappa\mathbb{I}\{Y \geq q(v_n|x)\}]K\left(\frac{x-X}{h}\right)\right\} \\
&= h^{-p}\int_{\mathbb{R}^p}[\bar{F}(q(u_n|x)|t) - \kappa\bar{F}(q(v_n|x)|t)]K\left(\frac{x-t}{h}\right)g(t)dt \\
&= \int_{\mathcal{S}}[\bar{F}(q(u_n|x)|x-hs) - \kappa\bar{F}(q(v_n|x)|x-hs)]K(s)g(x-hs)ds,
\end{aligned}$$

and therefore

$$\begin{aligned}
\left|\frac{\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x))}{S(u_n|x) - \kappa S(v_n|x)} - g(x)\right| &\leq \int_{\mathcal{S}} K(s)|g(x-hs) - g(x)|ds \\
&+ \int_{\mathcal{S}} K(s)\left|\frac{\bar{F}(q(u_n|x)|x-hs) - \kappa\bar{F}(q(v_n|x)|x-hs)}{S(u_n|x) - \kappa S(v_n|x)} - 1\right| g(x-hs)ds.
\end{aligned}$$

Clearly, under **H.2**, the first term of the right hand-side of the previous inequality is a $\mathcal{O}(h)$. Moreover, under **H.2**,

$$\int_{\mathcal{S}} K(s)g(x-hs)ds = g(x)(1+o(1)),$$

and

$$\left|\frac{\bar{F}(q(u_n|x)|x-hs)-\kappa\bar{F}(q(v_n|x)|x-hs)}{S(u_n|x)-\kappa S(v_n|x)}-1\right| \leq \frac{S(u_n|x)+\kappa S(v_n|x)}{S(u_n|x)-\kappa S(v_n|x)}\omega_{h,\zeta}(\nu_n|x).$$

Hence, the second term is a

$$\mathcal{O}\left(\frac{S(u_n|x)+\kappa S(v_n|x)}{S(u_n|x)-\kappa S(v_n|x)}w_{h,\zeta}(\nu_n|x)\right).$$

As a first conclusion, one has that $\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n,v_n|x))$ is equal to:

$$g(x)(S(u_n|x)-\kappa S(v_n|x))\left(1+\mathcal{O}(h)+\mathcal{O}\left(\frac{S(u_n|x)+\kappa S(v_n|x)}{S(u_n|x)-\kappa S(v_n|x)}w_{h,\zeta}(\nu_n|x)\right)\right). \tag{24}$$

Let us now focus on the variance of $\tilde{\Psi}_{h,\kappa}(u_n,v_n|x)$. First, remark that

$$\tilde{\Psi}_{h,\kappa}(u_n,v_n|x)-\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n,v_n|x)) = \sum_{j=1}^{n} Z_{j,n},$$

where

$$Z_{j,n} := \frac{1}{nh^p}\left([\mathbb{I}\{Y_j \geq q(u_n|x)\}-\kappa\mathbb{I}\{Y_j \geq q(v_n|x)\}]K\left(\frac{x-X_j}{h}\right)-h^p\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n,v_n|x))\right),$$

are independent, centered and identically distributed random variables. Remarking that for all $\tau>0$, $[\mathbb{I}\{Y_j \geq q(u_n|x)\}-\kappa\mathbb{I}\{Y_j \geq q(v_n|x)\}]^\tau = \mathbb{I}\{Y_j \geq q(u_n|x)\}-\kappa\mathbb{I}\{Y_j \geq q(v_n|x)\}$, the variance of $\tilde{\Psi}_{h,\kappa}(u_n,v_n|x)$ is equal to:

$$\frac{\|K\|_2^2}{nh^{2p}}\left(\mathbb{E}\left([\mathbb{I}\{Y \geq q(u_n|x)\}-\kappa\mathbb{I}\{Y \geq q(v_n|x)\}]Q\left(\frac{x-X}{h}\right)\right)-\frac{h^{2p}}{\|K\|_2^2}\mathbb{E}^2(\tilde{\Psi}_{h,\kappa}(u_n,v_n|x))\right),$$

where $Q := K^2/\|K\|_2^2$ has the same properties as K. Thus, (24) entails

$$\begin{aligned}\mathrm{Var}(\tilde{\Psi}_h(u_n|x))) &= \frac{\|K\|_2^2 g(x)}{nh^p}(S(u_n|x)-\kappa S(v_n|x)) \\ &\times \left(1+o(1)-\frac{g(x)}{\|K\|_2^2}h^p(S(u_n|x)-\kappa S(v_n|x))(1+o(1))\right),\end{aligned}$$

and, in view of $h^p(S(u_n|x)-\kappa S(v_n|x)) \leq h^p \to 0$, one has

$$\mathrm{Var}(\tilde{\Psi}_h(u_n|x))) = \frac{\|K\|_2^2 g(x)}{nh^p}(S(u_n|x)-\kappa S(v_n|x))(1+o(1)). \tag{25}$$

Lyapounov theorem can be used to establish the asymptotic normality of $\tilde{\Psi}_{h,\kappa}(u_n,v_n|x)$. From (25), the Lyapounov condition reduces to:

$$\lim_{n\to\infty}\left(\frac{nh^p}{S(u_n|x)-\kappa S(v_n|x)}\right)^{1+\delta/2} n\mathbb{E}|Z_{1,n}|^{2+\delta} = 0, \tag{26}$$

for some $\delta > 0$. Hence, if (26) is satisfied, the following convergence holds

$$\left(\frac{nh^p}{S(u_n|x) - \kappa S(v_n|x)}\right)^{1/2} \left(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x) - \mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x))\right) \xrightarrow{d} \mathcal{N}(0, g(x)\|K\|_2^2). \tag{27}$$

To prove (26), remark that for two random variables T_1 and T_2,

$$\mathbb{E}|T_1 + T_2|^{2+\delta} \leq 2^{2+\delta} \max(\mathbb{E}|T_1|^{2+\delta}, \mathbb{E}|T_2|^{2+\delta}).$$

Thus, one has

$$\mathbb{E}|Z_{1,n}|^{2+\delta} \leq \left(\frac{2\|K\|_{2+\delta}}{nh^p}\right)^{2+\delta} \mathbb{E}\left([\mathbb{I}\{Y \geq q(u_n|x)\} - \kappa \mathbb{I}\{Y \geq q(v_n|x)\}]N\left(\frac{x - X}{h}\right)\right),$$

where $N := (K/\|K\|_{2+\delta})^{2+\delta}$ is a bounded density function on $\mathbb{R}^p$ with same support as K. Using (24) yields

$$\mathbb{E}|Z_{1,n}|^{2+\delta} \leq \left(\frac{2\|K\|_{2+\delta}}{nh^p}\right)^{2+\delta} h^p(S(u_n|x) - \kappa S(v_n|x))(1 + o(1)).$$

Hence, condition (26) is satisfied and convergence (27) holds. Finally, remark that

$$(nh^p(S(u_n|x) - \kappa S(v_n|x)))^{1/2} \left(\frac{\tilde{S}_h(u_n|x) - \kappa \tilde{S}_h(v_n|x)}{S(u_n|x) - \kappa S(v_n|x)} - 1\right) = T_{1,n} + T_{2,n} + T_{3,n},$$

where

$$\begin{aligned}
T_{1,n} &:= \frac{(nh^p(S(u_n|x) - \kappa S(v_n|x)))^{1/2}}{\hat{g}_h(x)} \frac{\tilde{\Psi}_{h,\kappa}(u_n, v_n|x) - \mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x))}{S(u_n|x) - \kappa S(v_n|x)}, \\
T_{2,n} &:= \frac{(nh^p(S(u_n|x) - \kappa S(v_n|x)))^{1/2}}{\hat{g}_h(x)} \frac{\mathbb{E}(\tilde{\Psi}_{h,\kappa}(u_n, v_n|x)) - g(x)(S(u_n|x) - \kappa S(v_n|x))}{S(u_n|x) - \kappa S(v_n|x)}, \\
\text{and } T_{3,n} &:= \frac{(nh^p(S(u_n|x) - \kappa S(v_n|x)))^{1/2}}{\hat{g}_h(x)} (g(x) - \hat{g}_h(x)).
\end{aligned}$$

From (27), $\hat{g}_h(x)T_{1,n} \xrightarrow{d} \mathcal{N}(0, g(x)\|K\|_2^2)$. Since from Lemma 1, $\hat{g}_h(x) \xrightarrow{\mathbb{P}} g(x)$, one has that $T_{1,n} \xrightarrow{d} \mathcal{N}(0, \|K\|_2^2/g(x))$. Furthermore, (24) implies that

$$T_{2,n} = \frac{g(x)}{\hat{g}_h(x)}(nh^p(S(u_n|x) - \kappa S(v_n|x)))^{1/2} \left(\mathcal{O}(h) + \mathcal{O}\left(\frac{S(u_n|x) + \kappa S(v_n|x)}{S(u_n|x) - \kappa S(v_n|x)} w_{h,\zeta}(\nu_n|x)\right)\right) = o(1),$$

by assumptions. Finally, Lemma 1 entails that $T_{3,n} = o_{\mathbb{P}}(1)$ which completes the proof. ■

	Case 1						Case 2					
y	1/3	2/3	1	4/3	5/3	3	1/3	2/3	1	4/3	5/3	3
$x = 0.1$	0.04	0.05	0.07	0.10	0.11	0.31	0.07	0.11	0.17	0.22	0.26	0.32
$x = 0.5$	0.06	0.10	0.17	0.21	0.27	0.61	0.11	0.21	0.34	0.43	0.51	0.67

Table 1: Mean-squared errors of the N estimations of $\Lambda(y, y|x)$ for $y \in \{1/3, 2/3, 1, 4/3, 5/3, 3\}$.

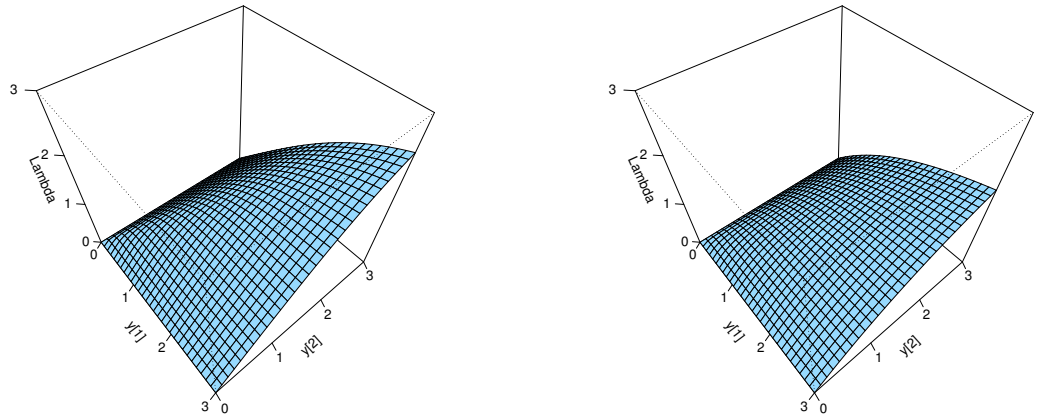

Figure 1: Conditional tail copula $\Lambda(y_1, y_2|x)$ given in (15) for $(y_1, y_2) \in [0, 3]^2$ at $x = 0.1$ (left panel) and $x = 0.5$ (right panel).

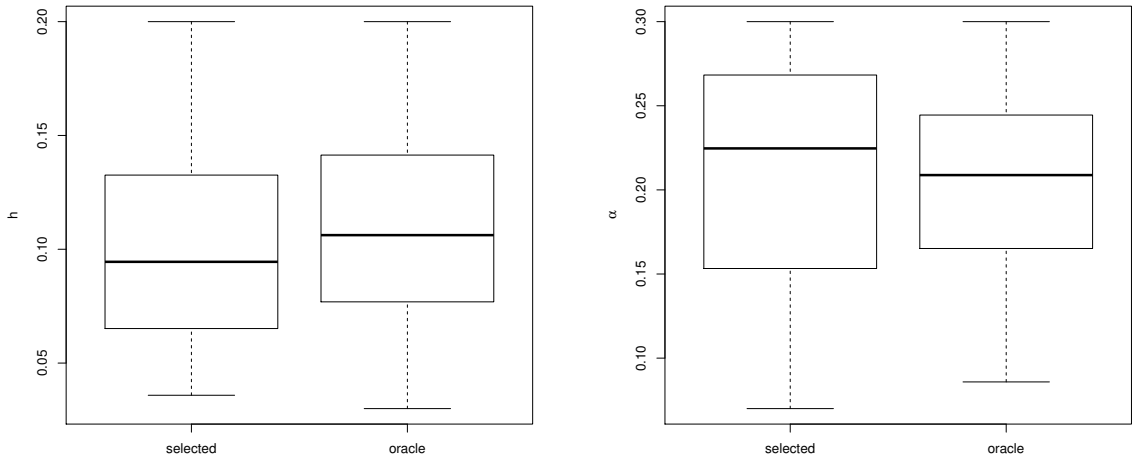

Figure 2: Boxplots of the hyper-parameters h (left) and α (right) selected by our procedure (`selected`) and by the oracle strategy (`oracle`) on simulated data in the Case 1 with $x = 0.1$.

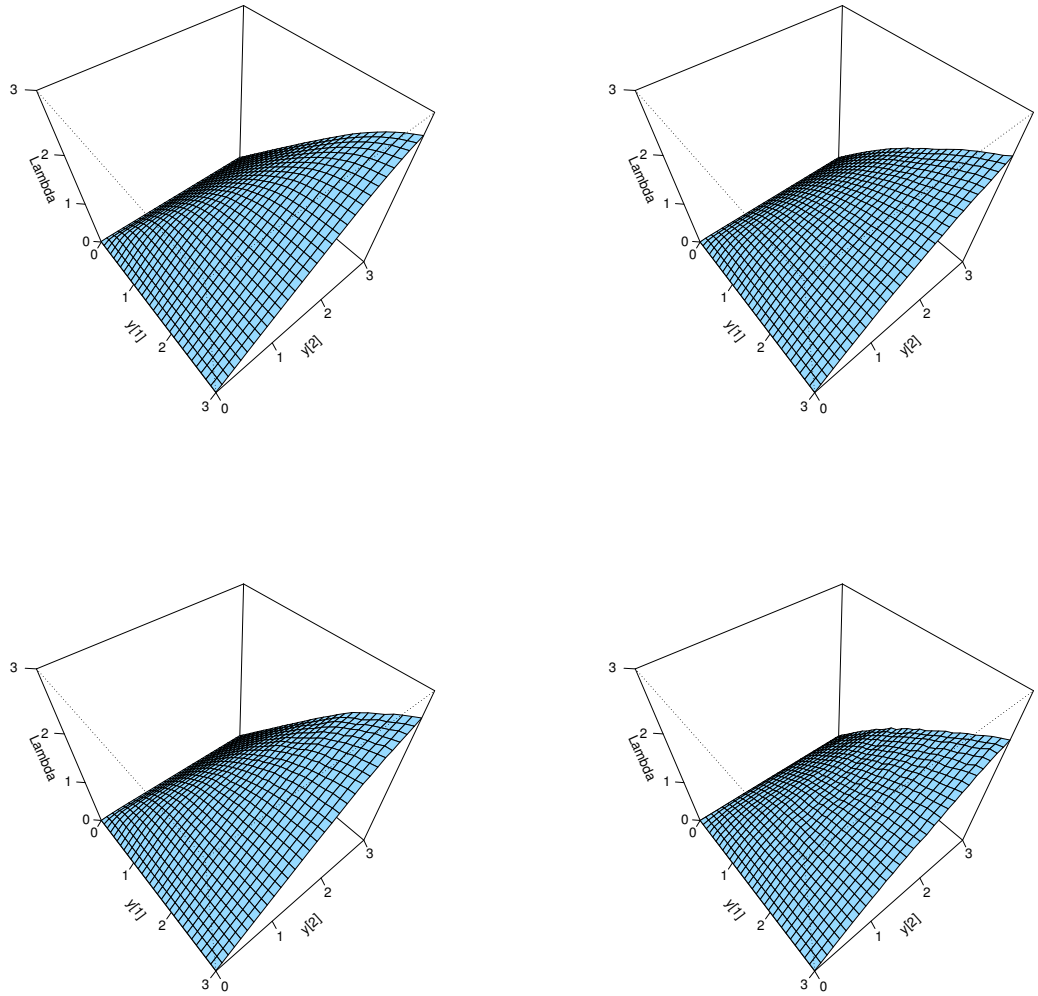

Figure 3: Median of the conditional tail copula estimates computed on simulated data for $(y_1, y_2) \in [0,3]^2$ in Case 1 (top) and Case 2 (bottom) at $x = 0.1$ (left) and $x = 0.5$ (right).

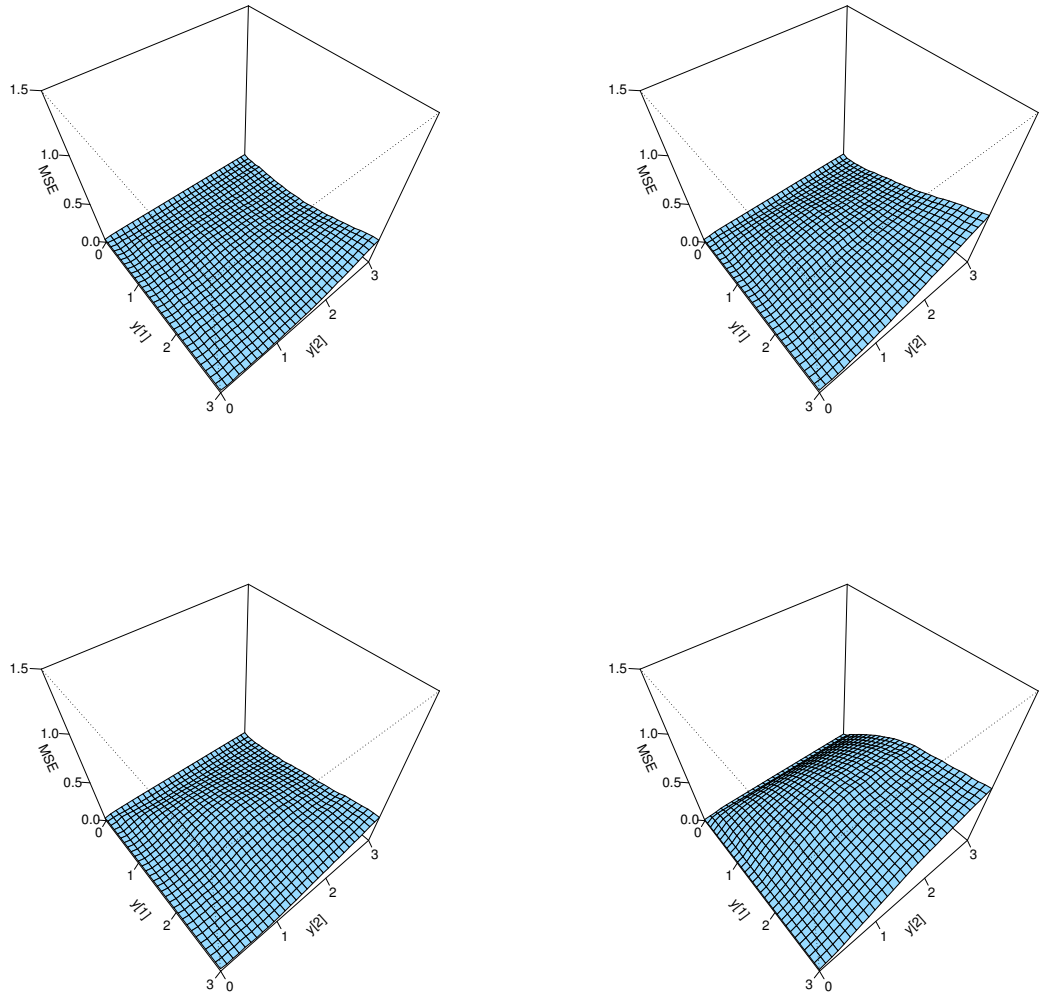

Figure 4: Mean-squared error associated with the conditional tail copula estimates computed on simulated data for $(y_1, y_2) \in [0, 3]^2$ in Case 1 (top) and Case 2 (bottom) at $x = 0.1$ (left) and $x = 0.5$ (right).

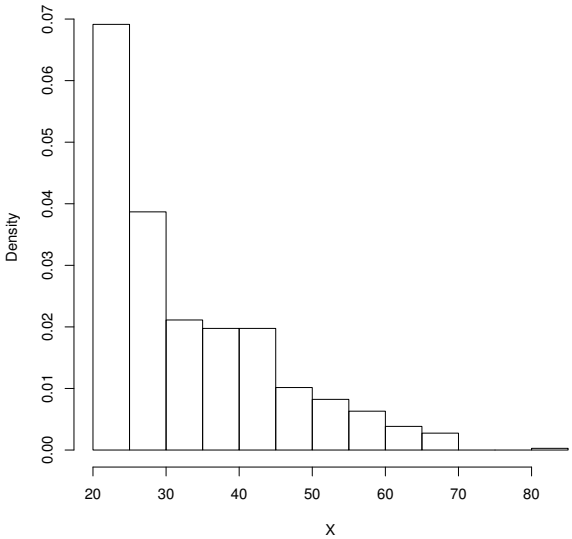

Figure 5: Histogram of the covariate X (age) in the `pima-indians-dataset`.

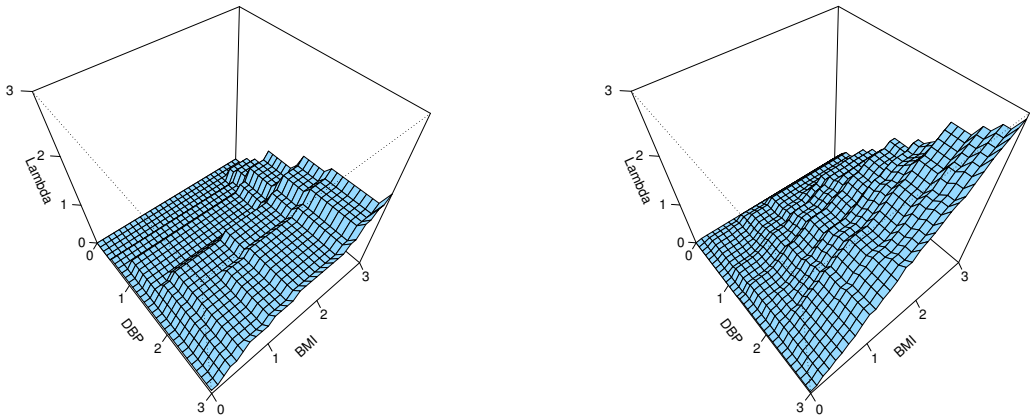

Figure 6: Conditional tail copula estimator $\check{\Lambda}_{h,h}(y_1, y_2|x)$ computed on the `pima-indians-dataset` for $(y_1, y_2) \in [0, 3]^2$ at $x = 30$ years (left panel) and $x = 60$ years (right panel).