Modeling of Censored Bivariate Extremal Events

Enkelejd Hashorva[1], Chengxiu Ling[1] and Zuoxiang Peng[2]

November 1, 2013

Abstract: In this paper we consider the estimation of the coefficient of tail dependence and of small tail probability under a bivariate randomly censoring mechanism. A new class of generalized moment estimators of the coefficient of tail dependence and the estimator of small tail probability are proposed, respectively. Under the bivariate Hall-type conditions, the asymptotic distributions of these estimators are established. Monte Carlo simulations are performed and the new estimators are applied to an insurance data-set.

Key words and phrases: randomly censoring; coefficient of tail dependence; large claims; small tail probability; concomitant order statistics; bivariate Hall-class.

AMS 2000 subject classification: Primary 60G70; Secondary 65C05.

1 Introduction

Modeling dependence structures underlying rare events is a crucial topic for advanced actuarial applications since the misidentification of dependence structures can cause a dramatic risk underestimation (cf. Beirlant et al. (2011) and Haug et al. (2011)). For a given bivariate risk vector (X, Y) with unit Fréchet distributed margins, Ledford and Tawn (1997) proposed the following tail dependence model

$$\mathbb{P}(X > x, Y > y) = x^{-c_1} y^{-c_2} \mathcal{L}(x, y), \quad \text{with } c_1, c_2 > 0. \tag{1.1}$$

Commonly, the parameter $\eta_1 = (c_1 + c_2)^{-1}$ is referred to as the coefficient of tail dependence, whereas $\mathcal{L}$ is a bivariate slowly varying function, i.e., there exists a function g such that for all $x, y > 0$ and $c > 0$

$$\lim_{t \to \infty} \frac{\mathcal{L}(tx, ty)}{\mathcal{L}(t, t)} = g(x, y), \quad \text{with } g(cx, cy) = g(x, y).$$

For various insurance and finance applications it is of interest to estimate both η_1 and $\mathbb{P}(X > x, Y > y)$, see e.g., Embrechts et al. (1997), Peng (1999), Beirlant and Vandewalle (2002), Einmahl et al. (2012), Goegebeur and Guillou (2013), and the references therein. Recently, Beirlant et al. (2011) studied model (1.1) by assuming further that

$$\mathcal{L}(x, y) = g_1(x, y)\Big(1 + x^{p_1} y^{p_2} g_2(x, y)(1 + o(1))\Big), \quad x, y \to \infty, \tag{1.2}$$

with

$$g_i(x, y) = h_i(x/(x + y)), \ i = 1, 2, \quad h_2 \neq 0 \quad \text{and} \quad \tau_1 = p_1 + p_2 < 0.$$

Let next $0 < w < 1, \underline{w} = w/(1 - w)$ and for some positive x define $y = x/\underline{w}$. Then the random variable $Z = \min(X, \underline{w}Y)$, has survival function $\overline{F} = 1 - F$ such that

$$\overline{F}(x) = \mathbb{P}(X > x, Y > x/\underline{w}) = x^{-1/\eta_1} \underline{w}^{c_2} h_1(w)\Big(1 + \underline{w}^{-p_2} h_2(w) x^{\tau_1}(1 + o(1))\Big) \tag{1.3}$$

$$=: C_1 x^{-1/\eta_1}\Big(1 + D_1 x^{\tau_1}(1 + o(1))\Big), \quad x \to \infty,$$

which shows that F belongs to the Hall-class, denoted by $F \in Hall(\eta_1, C_1, D_1, \tau_1)$. In the sequel we say that (X, Y) satisfies the bivariate Hall-type condition with tail dependence parameters $(c_i, h_i, p_i, i = 1, 2)$ if (1.3) holds. By (1.3), the coefficient of tail dependence η_1 and the small tail probability $\mathbb{P}(X > x, Y > y)$ can be estimated on the basis of univariate extreme value techniques, see e.g., Hill (1975), Dekkers et al. (1989), Gomes et al. (2008), Beirlant et al. (2009). However, in many insurance and finance applications complete data are rarely available, and censoring of data is a common phenomenon caused for instance by the existence of deductibles or retention levels. In univariate settings, different estimators of extreme value index under randomly censoring have been proposed, see e.g., Beirlant et al. (2007), Einmahl et al. (2008), Gomes and Neves (2011) and the references therein.

---

[1] Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland

[2] Southwest University, China, email: pzx@swu.edu.cn

So far there are no contributions in the literature dealing with extreme value problems for censored data in higher dimensions. Therefore, the aim of this paper is to establish some new estimators of the coefficient of tail dependence and for the tail probability in the presence of bivariate randomly censoring. Our framework of bivariate randomly censoring is easily explained if we consider two independent bivariate random vectors (X, Y) and $(\widetilde{X}, \widetilde{Y})$. Then the random vector (X, Y) is componentwise randomly censored by $(\widetilde{X}, \widetilde{Y})$, and we will establish our estimators based on samples from (X^*, Y^*) and $(\delta^{(1)}, \delta^{(2)})$ defined by

$$X^* = \min(X, \widetilde{X}), \quad Y^* = \min(Y, \widetilde{Y}), \quad \delta^{(1)} = \mathbb{I}\{X \leq \widetilde{X}\}, \quad \delta^{(2)} = \mathbb{I}\{Y \leq \widetilde{Y}\}, \tag{1.4}$$

with $\mathbb{I}\{\cdot\}$ the indicator function.

The main restrictions are that both (X, Y) and $(\widetilde{X}, \widetilde{Y})$ have unit Fréchet distributed margins and satisfy the bivariate Hall-type conditions with tail dependence parameters $(c_i, h_i, p_i, i = 1, 2)$ and $(\widetilde{c}_i, \widetilde{h}_i, \widetilde{p}_i, i = 1, 2)$. The principal challenge in our framework is that all parameters are assumed to be unknown. Our new estimators are highly flexible with a tuning parameter (see (2.2) and (3.2) below). With a suitable choice of the tuning parameter, under certain extreme value conditions our new estimators are asymptotically normal with zero-mean.

The rest of this paper is organized as follows. Section 2 shall introduce some notation and preliminaries. The main results are presented in Section 3 followed by a section with illustrating examples. Section 5 is dedicated to a small simulation study and a real life data application, whereas the proofs are deferred to Section 6.

2 Notation and Preliminaries

Let (X_i, Y_i) and $(\widetilde{X}_i, \widetilde{Y}_i), i = 1, \ldots, n$ be two independent and identically distributed samples from independent parents (X, Y) and $(\widetilde{X}, \widetilde{Y})$ with unit Fréchet distributed margins. Then by (1.4) the samples $X_i^*, Y_i^*, \delta_i^{(1)}, \delta_i^{(2)}, i = 1, \ldots, n$ are from parents $X^*, Y^*, \delta^{(1)}$ and $\delta^{(2)}$, respectively. Define

$$Z_i^* = \min(X_i^*, Y_i^*), \quad \delta_i^* = \mathbb{I}\{X_i^* \leq Y_i^*\}, \quad \delta_i = \delta_i^{(1)}\delta_i^* + \delta_i^{(2)}(1 - \delta_i^*)$$

for each sample (X_i, Y_i) and $(\widetilde{X}_i, \widetilde{Y}_i)$. Let $Z_{1,n}^* \leq Z_{2,n}^* \leq \ldots \leq Z_{n,n}^*$ be the associated order statistics of Z_i^*. Write $\delta_{[i,n]}$ for the concomitant order statistics with respect to $Z_{i,n}^*$, in other words, $\delta_{[i,n]} = \delta_k$ if $Z_{i,n}^* = Z_{k}^*, i = 1, \ldots, n$. For some intermediate integer sequence $k = k(n)$ satisfying $\lim_{n \to \infty} k(n) = \lim_{n \to \infty} n/k(n) = \infty$ define

$$M_n^{(j)}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\log\frac{Z_{n-i+1,n}^*}{Z_{n-k,n}^*}\right)^j; \quad N_n^{(\beta')}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{Z_{n-i+1,n}^*}{Z_{n-k,n}^*}\right)^{\beta'}; \quad \widehat{p}_n(k) = \frac{1}{k}\sum_{i=1}^{k}\delta_{[n-i+1,n]}, \tag{2.1}$$

where $j = 1, 2$ and $\beta' < 0$. In this paper we propose the following new estimators of η_1

$$\widehat{\eta}_n^{(\alpha)}(k) = \frac{M_n^{(1)}(k) + \alpha\left(1 - \frac{1}{2}\left(1 - \frac{(M_n^{(1)}(k))^2}{M_n^{(2)}(k)}\right)^{-1}\right)}{\widehat{p}_n(k)}, \qquad \text{with } \alpha \in \mathbb{R}, \tag{2.2}$$

which we refer to as the generalized moment estimators since they extend the Hill estimator ($\alpha = 0$) and the moment estimator ($\alpha = 1$) in the absence of censoring. In order to establish their asymptotic distribution, the following two assumptions are needed:

A1. (X, Y) and $(\widetilde{X}, \widetilde{Y})$ are independent and satisfy the bivariate Hall-type conditions with tail dependence parameters $(c_i, h_i, p_i, i = 1, 2)$ and $(\widetilde{c}_i, \widetilde{h}_i, \widetilde{p}_i, i = 1, 2)$, respectively.

A2. For a distribution function (df) $F \in Hall(\eta_1, C_1, D_1, \tau_1)$ defined by (1.3), $\ell_{0,F}(x) = (C_1 D_1 x^{\tau_1})^{-1}(x^{1/\eta_1}\overline{F}(x) - C_1)$ is a normalized regularly varying function (see Bingham et al. (1987), p 15).

Next, we present two lemmas which will be used to prove the main results and to deal with the simulation study. The first one is from Beirlant et al. (2007), p 160.

Lemma 2.1. *Denote by F, G and H the dfs of $Z = \min(X, \underline{w}Y), \widetilde{Z} = \min(\widetilde{X}, \underline{w}\widetilde{Y})$ and $Z^* = \min(X^*, Y^*)$, respectively. If the assumption **A1** holds, then $F \in Hall(\eta_1, C_1, D_1, \tau_1)$ and $G \in Hall(\eta_2, C_2, D_2, \tau_2)$ with*

$$\eta_1 = \frac{1}{c_1 + c_2}, \quad C_1 = h_1(w)\underline{w}^{c_2}, \quad D_1 = h_2(w)\underline{w}^{-p_2}, \quad \eta_2 = \frac{1}{\widetilde{c}_1 + \widetilde{c}_2}, \quad C_2 = \widetilde{h}_1(w)\underline{w}^{\widetilde{c}_2}, \quad D_2 = \widetilde{h}_2(w)\underline{w}^{-\widetilde{p}_2}.$$

Furthermore, if $D_1 + D_2 \neq 0$ for $\tau_1 = \tau_2$, then $H \in Hall(\eta, C, D, \tau)$ with

$$\eta = \frac{\eta_1 \eta_2}{\eta_1 + \eta_2}, \quad C = C_1 C_2, \quad D = D_1 \mathbb{I}\{\tau_1 > \tau_2\} + D_2 \mathbb{I}\{\tau_1 < \tau_2\} + (D_1 + D_2)\mathbb{I}\{\tau_1 = \tau_2\}, \quad \tau = \max(\tau_1, \tau_2).$$

For convenience, we assume $D \neq 0$ throughout this paper, which holds in most applications (cf. Table 1).

Remark 2.2. *Let $U(t) = \inf\{y \in \mathbb{R} : H(y) \geq 1 - 1/t\}$ for $t > 1$. If $H \in Hall(\eta, C, D, \tau)$, then $U(t) = t^\eta C^\eta[1 + \eta D C^\rho t^\rho (1 + o(1))]$. Moreover,*

$$\lim_{t \to \infty} \frac{U(tx)/U(t) - x^\eta}{b(t)} = x^\eta \frac{x^\rho - 1}{\rho} \tag{2.3}$$

holds locally uniformly for all $x > 0$. Here $\rho = \eta\tau < 0$ and $b(t) = \eta(\rho D C^\rho)t^\rho =: \eta\beta t^\rho$ (cf. Lemma 2.4 in Beirlant et al. (2009)).

From Lemma 2.1, Z^* is essentially Z censored independently by $\widetilde{Z}$. To estimate the coefficient of tail dependence η_1, it is sufficient to estimate the uncensored proportion $\mathbb{P}(Z \leq \widetilde{Z})$. Indeed, one may conclude that $\delta = \delta^{(1)}\delta^* + \delta^{(2)}(1 - \delta^*)$ and $\delta = \mathbb{I}\{Z \leq \widetilde{Z}\}$ by straightforward calculations.

We conclude this section with Lemma 2.3 below for the asymptotic expansions of $M_n^{(j)}(k), N_n^{(\tau)}(k)$ and $\widehat{p}_n(k)$. For notational simplicity we denote for $j, j' = 1, 2$

$$\mu_{j,\rho} = \frac{1/(1-\rho)^j - 1}{\rho}, \quad \sigma_j^2 = \Gamma(2j+1) - \Gamma^2(j+1), \quad \sigma_{j,j'} = \frac{\Gamma(j+j'+1) - \Gamma(j+1)\Gamma(j'+1)}{\sigma_j \sigma_{j'}}, \tag{2.4}$$

with $\Gamma(\cdot)$ the Euler Gamma function. If some estimator say $\widehat{\theta}_n$ converges in distribution to a $N(0, 1)$ random variable as $n \to \infty$, we shall abbreviate that as $\widehat{\theta}_n \sim \mathbb{A}s\mathbb{N}(0, 1)$. For two estimators $\widehat{\theta}_n$ and $\widehat{\phi}_n$, denote their asymptotic covariance by $\mathbb{A}s\mathbb{C}(\widehat{\theta}_n, \widehat{\phi}_n)$, and set $\mathbb{A}s\mathbb{V}(\widehat{\theta}_n) = \mathbb{A}s\mathbb{C}(\widehat{\theta}_n, \widehat{\theta}_n)$ for the asymptotic variance. Our notation for equality in distribution, convergence in distribution and convergence in probability are $\overset{d}{=}$, $\overset{d}{\to}$ and $\overset{p}{\to}$, respectively. All the limits are taken as $n \to \infty$ unless otherwise specified.

Lemma 2.3. *Suppose that **A1** and **A2** hold for an intermediate integer sequence $k = k(n)$ and $j, j' = 1, 2$.*

(a). For $P_n^{(j)} \sim \mathbb{A}s\mathbb{N}(0, 1)$ such that $\mathbb{A}s\mathbb{C}(P_n^{(j)}, P_n^{(j')}) = \sigma_{j,j'}$

$$\frac{M_n^{(j)}(k)}{\eta^j \Gamma(j+1)} - 1 \overset{d}{=} \frac{\sigma_j}{\Gamma(j+1)} \frac{P_n^{(j)}}{\sqrt{k}} + \mu_{j,\rho} \frac{b(n/k)}{\eta}(1 + o_p(1)). \tag{2.5}$$

(b). For $K_n \sim \mathbb{A}s\mathbb{N}(0, 1)$ such that $\mathbb{A}s\mathbb{C}(K_n, P_n^{(j)}) = \sqrt{1 - 2\rho}\,\Gamma(j+1)\frac{\mu_{j,\rho}}{\sigma_j}$

$$N_n^{(\tau)}(k) - \frac{1}{1-\rho} \overset{d}{=} \frac{\rho}{(1-\rho)\sqrt{1-2\rho}} \frac{K_n}{\sqrt{k}} + \frac{\rho}{(1-\rho)(1-2\rho)} \frac{b(n/k)}{\eta}(1 + o_p(1)). \tag{2.6}$$

(c). For $J_n \sim \mathbb{A}s\mathbb{N}(0, 1)$ being further asymptotically independent of $P_n^{(j)}$ and K_n

$$\widehat{p}_n(k) \overset{d}{=} \frac{\eta}{\eta_1}\left(1 + \sqrt{\frac{\eta_1}{\eta_2}} \frac{J_n}{\sqrt{k}} + \left(\frac{1}{1-\rho} \frac{b(n/k)}{\eta} - \frac{\eta_1 \tau_1 D_1 (\rho D)^{-\tau_1/\tau}}{1 - \eta_1 \tau_1}\left(\frac{b(n/k)}{\eta}\right)^{\tau_1/\tau}\right)(1 + o_p(1))\right). \tag{2.7}$$

Here $\mu_{j,\rho}, \sigma_j$ and $\sigma_{j,j'}$ are given by (2.4) and

$$K_n = \frac{(1-\rho)\sqrt{1-2\rho}}{\rho} \frac{\sum_{i=1}^k (\xi_i^\rho - 1/(1-\rho))}{\sqrt{k}}, \quad J_n = \frac{\sqrt{\eta_1 \eta_2}}{\eta} \frac{\sum_{i=1}^k \left(\delta_{[n-i+1,n]} - \mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*)\right)}{\sqrt{k}},$$

with $\xi_i, i = 1, \ldots, n$ being identically and independent random variables with common df $F_\xi(x) = 1 - 1/x, x \geq 1$.

3 Main Results

The aim of this section is to establish the asymptotic normality of the estimations of the coefficient of tail dependence η_1 and the tail probability $\mathbb{P}(X > x, Y > y)$. These results are stated in Theorem 3.1 and Theorem 3.3 under the following asymptotic condition

$$\lim_{n \to \infty} \sqrt{k} b(n/k) = \lambda < \infty, \tag{3.1}$$

which has been utilised for instance in Beirlant et al. (2007) where therein $\lambda = 0$ is investigated.

Theorem 3.1. *Under the assumptions* **A1** *and* **A2** *and suppose further the condition (3.1) holds for an intermediate integer sequence $k = k(n)$, then*

$$\sqrt{k}(\widehat{\eta}_n^{(\alpha)}(k) - \eta_1) \overset{d}{\to} \mathcal{Z} \sim N\left(\mu, \sigma^2\right),$$

where $\rho = \eta\tau < 0$ and

$$\mu = \frac{\lambda \eta_1}{\eta^2(1-\rho)} \left(\frac{\alpha\rho}{1-\rho} + \frac{\eta_1 D_1}{D} \mathbb{I}\{\tau_1 \geq \tau_2\} \right), \quad \sigma^2 = \frac{\eta_1^2\left(\alpha^2 + \eta_1\eta\right)}{\eta^2}.$$

Remark 3.2. *(a). For $\lambda = 0$, the results for $\alpha = 0$ and $\alpha = 1$ coincide with those in Beirlant et al. (2007).*

(b). The asymptotic variance is increasing with respect to $|\alpha|$; its minimum value η_1^3/η is obtained at $\alpha = 0$.

The tuning parameter α renders the estimators given in (2.2) to be highly flexible. With a suitable α even when the condition (3.1) holds with $\lambda \neq 0$, the asymptotic bias of $\widehat{\eta}_n^{(\alpha)}(k)$ may become zero.

Next, we consider the estimation of small tail probability

$$p_n = \mathbb{P}(X > x_n, Y > x_n/\underline{w}) = \mathbb{P}(Z > x_n) = \overline{F}(x_n)$$

for given $\underline{w}$ and sufficiently large x_n. The main inspiration for our estimator of small tail probability (denoted by $\widehat{p}_n$ below) comes from the recent contribution Beirlant et al. (2011). We retain the notation and the framework previously introduced, and denote further $\widehat{\tau} = \widehat{\rho}/M_n^{(1)}(k)$ and $\widehat{\tau}_1 = \widehat{\tau}\mathbb{I}\{\tau_1 > \tau_2\}$ with $\widehat{\rho}$ a consistent estimator of ρ, and thus $\widehat{p}_n$ is given by

$$\widehat{p}_n = \widehat{\overline{F}}_n(Z^*_{n-k,n}) \left(\frac{x_n}{Z^*_{n-k,n}} \left(1 + \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)} - \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)} \left(\frac{x_n}{Z^*_{n-k,n}} \right)^{\widehat{\tau}_1} \right) \right)^{-1/\widehat{\eta}_n^{(\alpha)}(k)}, \tag{3.2}$$

with

$$\widehat{b}(n/k) = M_n^{(1)}(k)(1 - 2\widehat{\rho})(1 - \widehat{\rho})^3\widehat{\rho}^{-3} \left(N_n^{(\widehat{\tau})}(k) - \frac{1}{1-\widehat{\rho}} \right), \quad \widehat{\overline{F}}_n(Z^*_{n-k,n}) = \prod_{i=1}^{n-k} \left(1 - \frac{\delta_{[i,n]}}{n-i+1} \right), \tag{3.3}$$

where $\widehat{p}_n(k), N_n^{(\tau)}(k)$ and $\widehat{\eta}_n^{(\alpha)}(k)$ are defined by (2.1) and (2.2), respectively. The estimator of the tail empirical df $\widehat{\overline{F}}_n(Z^*_{n-k,n})$ is from Kaplan and Meier (1958), while $\widehat{\rho}$, the consistent estimator of ρ, can be found in Gomes et al. (2009), de Wet et al. (2012). As in Theorem 5.2 in Beirlant et al. (2009) we assume further the following condition: For an intermediate integer sequence $k = k(n)$

$$\frac{p_n}{\overline{F}(Z^*_{n-k,n})} \overset{p}{\to} 0, \quad \frac{\sqrt{k}}{\log(p_n/\overline{F}(Z^*_{n-k,n}))} \overset{p}{\to} \infty, \quad \text{as } n \to \infty, \tag{3.4}$$

which will be utilised for the derivation of the asymptotic distribution of $\widehat{p}_n$.

Theorem 3.3. *Under the conditions and notation of Theorem 3.1, if further (3.4) holds, then we have the convergence in distribution*

$$\frac{\sqrt{k}(\widehat{p}_n/p_n - 1)}{\log(p_n/\overline{F}(Z^*_{n-k,n}))} \overset{d}{\to} -\frac{\mathcal{Z}}{\eta_1}. \tag{3.5}$$

Remark 3.4. (*a*). *Theorem 3.3 shows the following convergence in probability*

$$\frac{\log\widehat{p}_n - \log p_n}{\log(p_n/\overline{F}(Z^*_{n-k,n}))} \xrightarrow{p} 0.$$

Hence an equivalent statement of (3.5) *is*

$$\frac{\sqrt{k}(\widehat{p}_n/p_n - 1)}{\log(\widehat{p}_n/\widehat{\overline{F}}_n(Z^*_{n-k,n}))} \xrightarrow{d} -\frac{\mathcal{Z}}{\eta_1}.$$

(*b*). *In the absence of censoring,* $\widehat{p}_n$ *simplifies to the Weissman estimator provided that* $\widehat{\tau}_1 \equiv 0$. *In fact, Theorem 5.2 in Beirlant et al. (2009) is a special case of our results.*

Theorem 3.3 shows that under some mild conditions the limit distribution of $\widehat{p}_n$ depends only on the asymptotic distribution of $\widehat{\eta}_n^{(\alpha)}(k)$. If the condition (3.1) holds with $\lambda = 0$ or the asymptotic bias μ of $\widehat{\eta}_n^{(\alpha)}(k)$ is zero with a suitable parameter α, then an asymptotic confidence interval of nominal level $1 - \beta$ is obtained as follows (see Theorem 3.1)

$$\left[\widehat{p}_n \bigg/ \left(1 + \frac{|\log(\widehat{p}_n/\widehat{\overline{F}}_n(Z^*_{n-k,n}))|}{\sqrt{k}}\frac{\widehat{\sigma}}{\widehat{\eta}_n^{(\alpha)}(k)}z_{\beta/2}\right), \widehat{p}_n \bigg/ \left(1 - \frac{|\log(\widehat{p}_n/\widehat{\overline{F}}_n(Z^*_{n-k,n}))|}{\sqrt{k}}\frac{\widehat{\sigma}}{\widehat{\eta}_n^{(\alpha)}(k)}z_{\beta/2}\right)\right],$$

where $\widehat{\sigma}$ and $z_{\beta/2}$ are respectively the estimation of σ and the $(1-\beta/2)$-quantile of the standard normal distribution.

Due to the censoring mechanism and all unknown parameters, common bias-reduction methods for instance Caeiro et al. (2005) and Beirlant et al. (2009) can not be employed. Another approach to reduce bias is to adjust the bias term on the basis of $\widehat{b}(n/k)$ in (3.3), which is however not in the scope of this contribution. From the simulations in Section 5 below, we see that our new estimators with suitable tuning parameter are highly flexible and perform very well.

4 Examples

In the following Q denotes the joint df of (X, Y) with unit Fréchet distributed margins, i.e., $Q_1(x) = Q_2(x) = e^{-1/x}, x > 0$. The corresponding copula of Q with one parameter θ is denoted by $C_\theta(u, v) = Q(Q_1^{-1}(u), Q_2^{-1}(v))$. Our examples below show that $Z = \min(X, \underline{w}Y)$ belongs to $Hall(\eta_1, \tau_1, C_1, D_1)$ with η_1, τ_1, C_1, D_1 listed in Table 1 below.

Example 4.1. *Farlie–Gumbel–Morgenstern* ($FGM(\theta)$). $C_\theta(u, v) = uv + \theta uv(1 - u)(1 - v), \theta \in [-1, 1]$ and further

$$\mathbb{P}(X > x, Y > y) = 1 - e^{-1/x} - e^{-1/y} + e^{-1/x - 1/y}\left(1 + \theta(1 - e^{-1/x})(1 - e^{-1/y})\right), \quad x, y > 0.$$

(*a*). For $\theta = -1/3$

$$\mathbb{P}(X > x, Y > y) = \frac{1}{xy}g_1(x, y)\left(1 + \frac{1}{xy}g_2(x, y)(1 + o(1))\right),$$

with

$$g_1(x, y) = \frac{2}{3}, \quad g_2(x, y) = -\frac{7/6(x/y + y/x) + 9/4}{2}.$$

(*b*). For $\theta = -1$

$$\mathbb{P}(X > x, Y > y) = \frac{1}{xy\sqrt{xy}}g_1(x, y)\left(1 + \frac{1}{\sqrt{xy}}g_2(x, y)(1 + o(1))\right),$$

with

$$g_1(x, y) = \sqrt{\frac{x}{y}} + \sqrt{\frac{y}{x}}, \quad g_2(x, y) = -\frac{7/6(x/y + y/x) + 9/4}{g_1(x, y)}.$$

(*c*). For $\theta \neq -1/3, -1$

$$\mathbb{P}(X > x, Y > y) = \frac{1}{xy} g_1(x,y) \left(1 + g_2(x,y) \frac{1}{\sqrt{xy}} (1 + o(1))\right),$$

with

$$g_1(x,y) = 1 + \theta, \quad g_2(x,y) = -\frac{1 + 3\theta}{2(1+\theta)} \left(\sqrt{\frac{x}{y}} + \sqrt{\frac{y}{x}}\right). \tag{4.1}$$

Example 4.2. *Gumbel copula (Gumbel(θ)).* $C_\theta(u,v) = \exp(-((-\log u)^\theta + (-\log v)^\theta)^{1/\theta}), \theta \in [1, \infty)$, which is not only an extreme value copula but also an Archimedean copula. For all $x, y > 0$

$$\mathbb{P}(X > x, Y > y) = 1 - e^{-1/x} - e^{-1/y} + e^{-(x^{-\theta} + y^{-\theta})^{1/\theta}}.$$

In the following we exclude the case $\theta = 1$ for which X and Y are independent since it is covered by the previous example.

(a). For $\theta = 2$ we have

$$\mathbb{P}(X > x, Y > y) = \frac{1}{\sqrt{xy}} g_1(x,y) \left(1 + \frac{1}{xy} g_2(x,y)(1 + o(1))\right), \quad x, y \to \infty,$$

with

$$g_1(x,y) = \sqrt{\frac{x}{y}} + \sqrt{\frac{y}{x}} - \sqrt{\frac{y}{x} + \frac{x}{y}}, \quad g_2(x,y) = \frac{(x/y)^{3/2} + (y/x)^{3/2} - (x/y + y/x)^{3/2}}{6g_1(x,y)}.$$

(b). If $\theta \notin \{1, 2\}$, then

$$\mathbb{P}(X > x, Y > y) = \frac{1}{\sqrt{xy}} g_1(x,y) \left(1 + \frac{1}{\sqrt{xy}} g_2(x,y)(1 + o(1))\right), \quad x, y \to \infty,$$

with

$$g_1(x,y) = \sqrt{\frac{x}{y}} + \sqrt{\frac{y}{x}} - \sqrt{\frac{y}{x}} \left(1 + \left(\frac{x}{y}\right)^\theta\right)^{1/\theta}, \quad g_2(x,y) = -\frac{x/y + y/x - y/x(1 + (x/y)^\theta)^{2/\theta}}{2g_1(x,y)}.$$

Example 4.3. *Survival Clayton copula (S-clayton(θ)).* Clayton copula C_θ is defined by

$$C_\theta(u,v) = \left(\max(u^{-1/\theta} + v^{-1/\theta} - 1, 0)\right)^{-\theta}, \quad \theta \in [-1, 0) \cup (0, \infty).$$

Its survival copula $\widehat{C}_\theta$ is given by $\widehat{C}_\theta(u,v) = u + v - 1 + C_\theta(1 - u, 1 - v)$. Consequently, if Q has survival copula $\widehat{C}_\theta$ with $\theta > 0$, then as $x, y \to \infty$

$$\mathbb{P}(X > x, Y > y) = 1 - Q_1(x) - Q_2(y) + Q(x,y) = C_\theta(\overline{Q}_1(x), \overline{Q}_2(y))$$
$$= \frac{g_1(x,y)}{\sqrt{xy}} \left(1 + \left(\frac{\mathbb{I}\{\theta = 1\}}{xy} g_2^a(x,y) + \frac{\mathbb{I}\{\theta < 1\}}{\sqrt{xy}} g_2^b(x,y) + \frac{\mathbb{I}\{\theta > 1\}}{(xy)^{1/2\theta}} g_2^c(x,y)\right) (1 + o(1))\right),$$

where

$$g_1(x,y) = \left(\left(\frac{x}{y}\right)^{1/(2\theta)} + \left(\frac{y}{x}\right)^{1/(2\theta)}\right)^{-\theta}, \quad g_2^a(x,y) = -\frac{1}{4},$$

$$g_2^b(x,y) = -\frac{1}{2} \frac{(x/y)^{1/2 - 1/(2\theta)} + (y/x)^{1/2 - 1/(2\theta)}}{(x/y)^{1/(2\theta)} + (y/x)^{1/(2\theta)}}, \quad g_2^c(x,y) = \frac{\theta}{(x/y)^{1/(2\theta)} + (y/x)^{1/(2\theta)}}.$$

Hence if $\theta \in (0, \infty)$, then $\eta_1 = 1$ and $\lim_{x \to \infty} g_1(x,x) \neq 0$, implying its asymptotic dependence (cf. Ledford and Tawn (1997)).

Example 4.4. *Gaussian copula (Gauss(θ)).* $C_\theta(u,v) = \Phi_\theta(\Phi^{-1}(u), \Phi^{-1}(v)), \theta \in (-1, 1)$. The Gaussian copula does not satisfy the assumption **A1**, since $c_1 = c_2 = 1/(1+\theta), p_1 = p_2 = 0, \tau_1 = 0, g_1(x,y)$ is ray independent and $\lim_{x \to \infty} g_1(x,x) = 1$, see Embrechts et al. (1997) and Hashorva (2010, 2012).

Copula	θ	$\eta_1 \in (0,1]$	τ_1	$C_1 > 0$	$D_1 \neq 0$
FGM	$-1/3$	$1/2$	-2	$\frac{2}{3}\underline{w}$	$-\frac{7/6(\underline{w}^2+1)+9/4\underline{w}}{2\underline{w}}$
	-1	$1/3$	-1	$\underline{w}(\underline{w}+1)$	$-\frac{7/6(\underline{w}^2+1)+9\underline{w}/4}{\underline{w}+1}$
	$(-1,-1/3)\cup(-1/3,1]$	$1/2$	-1	$\underline{w}(1+\theta)$	$-\frac{1+3\theta}{2(1+\theta)}\frac{1+\underline{w}}{\sqrt{\underline{w}}}$
Gumbel	1	$1/2$	-1	$\underline{w}$	$-\frac{1}{2}(\underline{w}+1)$
	2	1	-2	$\underline{w}+1-\sqrt{\underline{w}^2+1}$	$\frac{\underline{w}^3+1-(\underline{w}^2+1)^{3/2}}{6C_1}$
	$(1,2)\cup(2,\infty)$	1	-1	$\underline{w}+1-(\underline{w}^\theta+1)^{1/\theta}$	$-\frac{\underline{w}^2+1-(\underline{w}^\theta+1)^{2/\theta}}{2C_1}$
S-clayton	1	1	-2	$\sqrt{\underline{w}}\left(\underline{w}^{1/(2\theta)}+\underline{w}^{-1/(2\theta)}\right)^{-\theta}$	$-\frac{1}{4}\underline{w}$
	$(0,1)$	1	-1		$-\frac{1}{2}\frac{\underline{w}^{1/(2\theta)}+\underline{w}^{1-1/(2\theta)}}{\underline{w}^{1/(2\theta)}+\underline{w}^{-1/(2\theta)}}$
	$(1,\infty)$	1	$-1/\theta$		$\frac{\theta}{\underline{w}^{-1/\theta}+1}$
Gauss	$(-1,1)$	$(1+\theta)/2$	0	$\underline{w}^{1/(1+\theta)}$	$--$

Table 1: Examples $Z = \min(X,\underline{w}Y) \in Hall(\eta_1,C_1,D_1,\tau_1)$ with $(X,Y) \sim C_\theta(Q_1(x),Q_2(y))$ with unit Fréchet margins Q_1,Q_2 and copula C_θ.

5 Simulation Studies and Application to Insurance Data

In this section, we illustrate the finite sample properties of our estimations of the coefficient of tail dependence η_1 and the joint tail probability $\mathbb{P}(X > x, Y > y)$ via Monte Carlo simulations and a real-life example. The estimations of joint tail probability as well as conditional probabilities of the form $\mathbb{P}(X > x | Y > y)$ are proceeded in both simulations and applications.

We first perform small Monte Carlo simulations with 100 samples of size $n = 1000$ from the bivariate random vectors discussed in Section 4. For comparison, we mainly simulate our generalized moment estimators $\widehat{\eta}_n^{(\alpha)}(k)$ (abbreviated as $\widehat{\eta}_n^{(\alpha)}$, the same for other estimators below) with different α, and compare them with $\overline{H}_n^{(c)}$ and $\widetilde{H}_n^{(c)}$ defined by

$$\overline{H}_n^{(c)}(k) = \widehat{\eta}_n^{(0)}(k)\left(1 - \frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right), \quad \widetilde{H}_n^{(c)}(k) = \widehat{\eta}_n^{(0)}(k) - \frac{1}{1-\widehat{\rho}}\frac{\widehat{b}(n/k)}{\widehat{p}_n(k)}, \tag{5.1}$$

where $\widehat{\rho}$ and $\widehat{\beta}$ are respectively the estimators of ρ and β given by Gomes et al. (2009), and $\widehat{p}_n(k)$ and $\widehat{b}(n/k)$ are defined by (2.1) and (3.3), respectively. Caeiro et al. (2005) and Beirlant et al. (2009) studied estimators (5.1) in the absence of censoring, which are rather attractive in the sense of minimal variance reduced bias (MVRB) extreme value index estimators.

As showed below in Table 2 and Figure 1 $\sim$ Figure 4, our new estimators of the coefficient of tail dependence with suitable α are comparable with $\overline{H}_n^{(c)}$ and $\widetilde{H}_n^{(c)}$. For comparison, we also simulate all estimators without censoring, which are respectively superior to those in the presence of censoring. Next we simulate samples from $(X,Y) \sim$ Gumbel(2) censored by $(\widetilde{X},\widetilde{Y}) \sim$ S-clayton(1) with $w = 0.2, 0.5, 0.6$ and 0.8. It turns out that no big discrepancy appears with different w, so we take $w = 0.5$ for the rest of simulations unless otherwise stated. Finally, we make comparisons for the cases $\tau = \tau_1, \tau_2$ and $\tau = 0$. For $\tau_1 > \tau_2$ we take samples from FGM(0.5) censored by Gumbel(2), our estimator with $\alpha = 1$ is comparable to $\widetilde{H}_n^{(c)}$. While for $\tau_2 > \tau_1$, we draw samples from Gumbel(2) censored by FGM(0.5) and S-clayton(2), respectively. Our estimator with $\alpha = 0.5$ (1) for the former (the latter) and $\overline{H}_n^{(c)}$ are better than the others for the case (cf. Table 2 and Figure 3). For $\tau_1 = \tau_2 < 0$, we take samples from Gumbel(2) censored by Gumbel(2) and S-clayton(1), respectively. The simulation shows that all estimators show good finite sample behaviors. As an exceptional case $\tau = 0$ with samples from S-clayton(1) censored by Gauss(0), our estimator with $\alpha = 0.5$ is the best one among all simulated estimators, see Figure 4.

Next, we shall focus on the finite sample behaviors of $\widehat{p}_n$ in (3.2), the estimators of small tail probability $\mathbb{P}(X > x_n, Y > x_n/\underline{w})$, and the estimator of conditional probability $\mathbb{P}(X > x_n | Y > x_n/\underline{w})$, which is obtained by $\widehat{p}_n$ divided by the Kaplan–Meier estimation of survival df of Y (cf. (3.3)). Motivated by Theorem 3.3 and the simulation results above, we mainly consider our tail probability estimators with $\alpha = 0, 0.5, 1$ and $\widehat{\tau} = -1, \widehat{\tau}_1 = 0, -1$ due to the unknown parameters τ_1 and τ_2. To this end, we take samples from Gumbel(2) censored by S-clayton(1) and draw the sample paths in Figure 5 for $\widehat{\tau}_1 = 0$ and Figure 6 for $\widehat{\tau}_1 = -1$, respectively. The results show that our estimators with $\alpha = 1, \widehat{\tau}_1 = 0$ and $\alpha = 0.5, \widehat{\tau}_1 = -1$ are quite stable. For the conditional probability Figure 7 and Figure 8

illustrate the similar finite sample behaviors as those for the estimations of tail probability.

Finally, we consider an application of Loss-ALAE data-set examined by Beirlant et al. (2011). There are 34 data points censored out of 1500 data points. First we transform the original data to be unit Fréchet distributed margins by using the Kaplan-Meier estimation of survival df as follows (cf. Kaplan and Meier (1958))

$$\mathbb{X}_i = -1/\log(1 - \widehat{\overline{F}}_X(X_i^*)), \quad \mathbb{Y}_i = -1/\log(1 - \widehat{\overline{F}}_Y(Y_i^*)),$$

with

$$\widehat{\overline{F}}_X(x) = \prod_{i=1}^{n} \left(1 - \frac{\delta_{[i,n]}^{(1)}}{n-i+1}\mathbb{I}\{X_{i,n}^* \le x\}\right), \quad \widehat{\overline{F}}_Y(x) = \prod_{i=1}^{n} \left(1 - \frac{\delta_{[i,n]}^{(2)}}{n-i+1}\mathbb{I}\{Y_{i,n}^* \le x\}\right).$$

Now we apply our censoring mechanism into the transformed data. Figure 9 draws sample paths of our estimators $\widehat{\eta}_n^{(0)}$ and $\widehat{\eta}_n^{(1)}$ for the coefficient of tail dependence of (Loss, ALAE) with $w = 0.5$, which shows that $\widehat{\eta}_n^{(1)}$ is more stable than $\overline{H}_n^{(c)}$ and $\widetilde{H}_n^{(c)}$. Next, we estimate the tail probability $\mathbb{P}(\text{Loss} > 200000, \text{ALAE} > 100000)$ by our tail probability estimators (3.2) plugged in the estimator (2.2) of the coefficient of tail dependence with $w = 2/3$. Figure 10 shows that our tail probability estimators with $\alpha = 1, \widehat{\tau}_1 = 0$ and $\alpha = 0.5, \widehat{\tau}_1 = -1$ are rather stable and close to the empirical tail probability 0.006. Figure 11 shows similar behaviors for estimations of the conditional tail probability $\mathbb{P}(\text{ALAE} > 100000 | \text{Loss} > 200000)$.
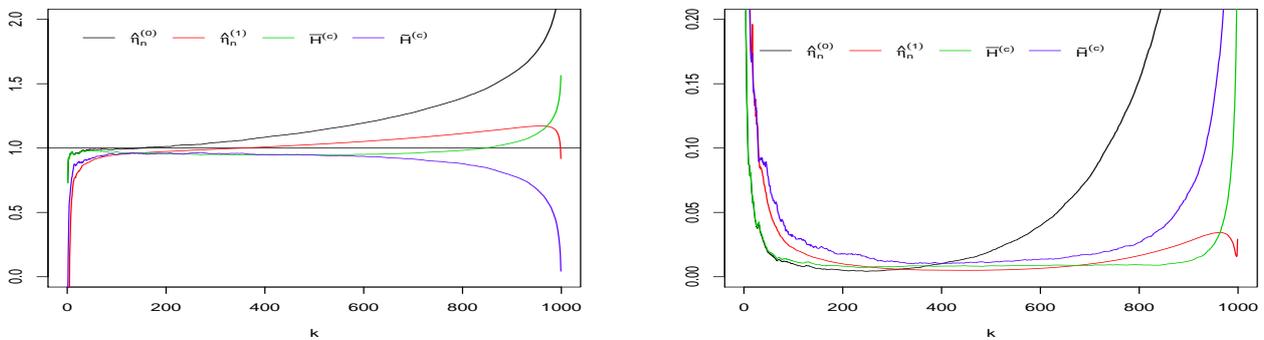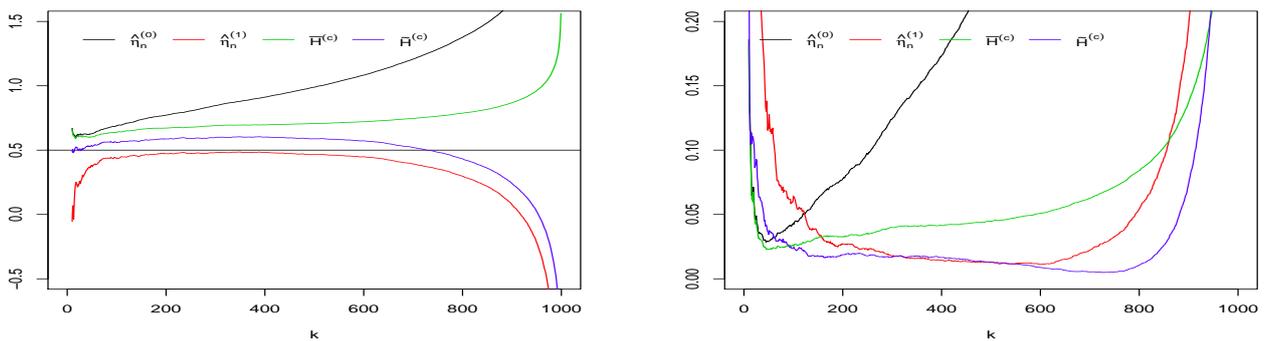

Figure 1: Finite behaviors of mean values (left) and mean squared errors (right) of $\widehat{\eta}_n^{(0)}, \widehat{\eta}_n^{(1)}, \overline{H}_n^{(c)}, \widetilde{H}_n^{(c)}$. Random samples are from $(X, Y) \sim \text{Gumbel}(2)$ without censoring. Here $\eta_1 = 1$.

Figure 2: Finite behaviors of mean values (left) and mean squared errors (right) of $\widehat{\eta}_n^{(0)}, \widehat{\eta}_n^{(1)}, \overline{H}_n^{(c)}, \widetilde{H}_n^{(c)}$. Random samples are from $(X, Y) \sim \text{FGM}(0)$ censored by Gumbel(1). Here $\eta_1 = 1/2$ and $\tau_1 = -1, \tau_2 = -1$.

Conditions	(X,Y)	$(\widetilde{X}, \widetilde{Y})$	$\widehat{\eta}_n^{(0)}$	$\widehat{\eta}_n^{(1)}$	$\overline{\eta}_n^{(0.5)}$	$\overline{H}_n^{(c)}$	$\widetilde{H}_n^{(c)}$	$MSE_{\widehat{\eta}_n^{(0)}}$	$MSE_{\widehat{\eta}_n^{(1)}}$	$MSE_{\widehat{\eta}_n^{(0.5)}}$	$MSE_{\overline{H}_n^{(c)}}$	$MSE_{\widetilde{H}_n^{(c)}}$
$w = 0.5$	Gumbel(2)	*	1.0354	1.0261	1.031	0.9777	0.9771	0.0053	0.0056	0.0048	0.0053	0.0072
$w = 0.2$	Gumbel(2)	S-clayton(1)	1.1093	1.0147	1.0806	1.0724	0.9848	0.0335	0.0154	0.0308	0.0155	0.0172
$= 0.5$	Gumbel(2)	S-clayton(1)	1.0972	1.0023	1.0641	1.0387	1.0371	0.0228	0.0157	0.0184	0.0101	0.0139
$= 0.6$	Gumbel(2)	S-clayton(1)	1.0673	0.9914	1.0553	1.041	0.9788	0.0235	0.0146	0.0194	0.0074	0.0143
$= 0.8$	Gumbel(2)	S-clayton(1)	1.1266	0.9849	1.0651	1.0485	0.9778	0.0320	0.0127	0.0220	0.0116	0.0155
$w = 0.5$	Gumbel(2)	Gumbel(1)	1.1679	0.8863	1.0764	1.0343	0.9581	0.0457	0.0387	0.0191	0.0078	0.0171
$= 0.5$	Gumbel(2)	S-clayton(1)	1.1246	1.0065	1.0805	1.0482	1.0420	0.0267	0.0117	0.0189	0.011	0.0134
$= 0.5$	Gumbel(2)	Gauss(0.5)	1.0935	0.9250	1.0579	1.0182	0.9548	0.0260	0.0230	0.0183	0.0043	0.0171
$\tau = \tau_1$	FGM(0.5)	Gumbel(2)	0.6068	0.5173	0.5538	0.5901	0.5150	0.0235	0.0097	0.0225	0.0201	0.0080
$= \tau_2$	Gumbel(2)	FGM(0.5)	1.1100	0.8290	1.0396	1.0169	0.9240	0.0274	0.0464	0.0103	0.0033	0.0208
$\tau_2 = \tau_1$	Gumbel(2)	Gumbel(2)	1.0671	0.9547	1.0555	1.0157	0.9788	0.0171	0.0123	0.0176	0.0085	0.0132
$=0$	S-clayton(1)	Gauss(0)	1.0867	0.8536	1.0429	1.0052	0.9264	0.0267	0.0522	0.0094	0.0036	0.0262

Table 2: Estimations of the coefficient of tail dependence η_1. The true value of η_1 is 1 except the case $(X,Y) \sim \text{FGM}(0.5)$ with $\eta_1 = 1/2$. FGM(θ), Gumbel(θ), S-clayton(θ) and Gauss(θ) represent respectively the corresponding bivariate models with parameter θ discussed in Section 4 (cf. Table 1).
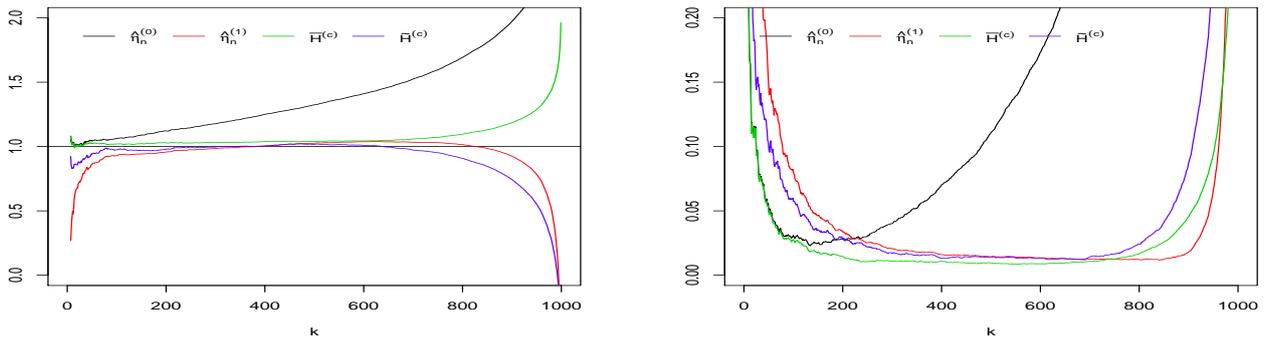
Figure 3: Finite behaviors of mean values (left) and mean squared errors (right) of $\widehat{\eta}_n^{(0)}, \widehat{\eta}_n^{(1)}, \overline{H}_n^{(c)}, \widetilde{H}_n^{(c)}$. Random samples are from $(X, Y) \sim \text{Gumbel}(2)$ censored by S-clayton(2). Here $\eta_1 = 1$ and $\tau_1 = -2, \tau_2 = -1/2$.
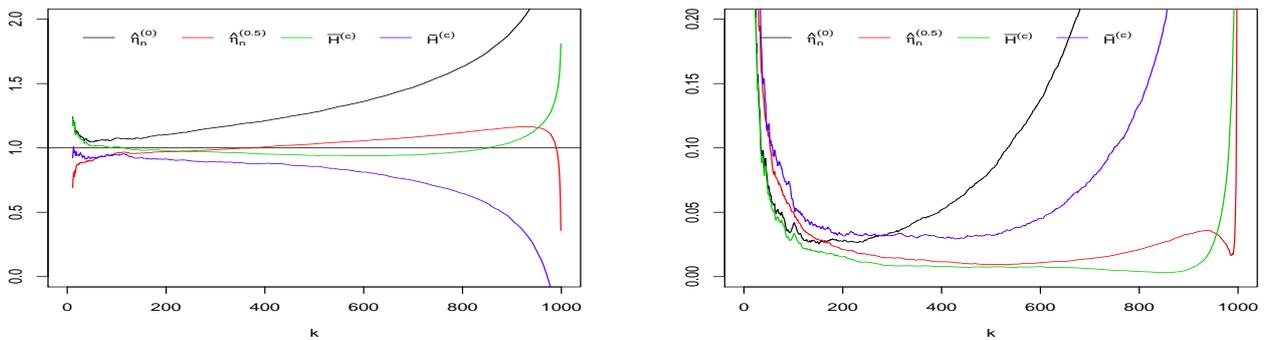

Figure 4: Finite behaviors of mean values (left) and mean squared errors (right) of $\widehat{\eta}_n^{(0)}, \widehat{\eta}_n^{(0.5)}, \overline{H}_n^{(c)}, \widetilde{H}_n^{(c)}$. Random samples are from $(X, Y) \sim \text{S-clayton}(1)$ censored by Gauss(0). Here $\eta_1 = 1$ and $\tau_1 = -2, \tau_2 = 0$.
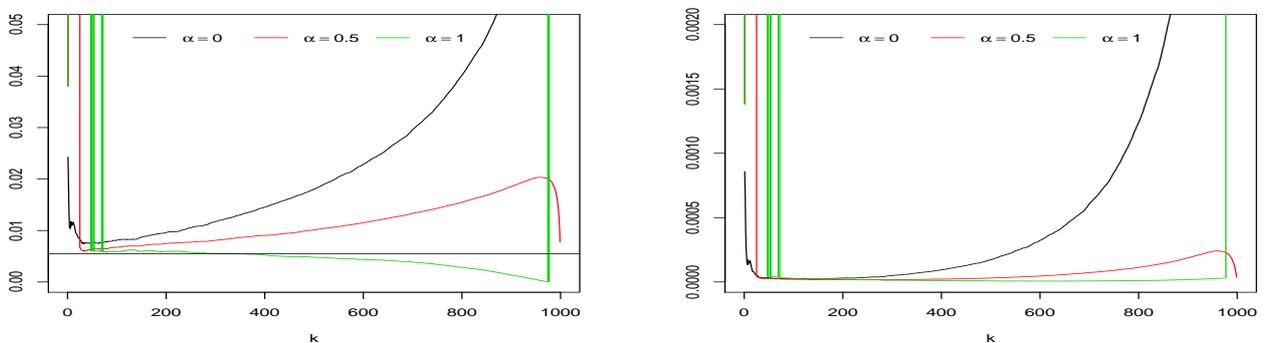
Figure 5: Mean values (left) and mean squared errors (right) of the estimation of $\mathbb{P}(X > 40, Y > 160) = 0.0054$ with $\widehat{\tau}_1 = 0$. Samples are from Gumbel(2) censored by S-clayton(1).
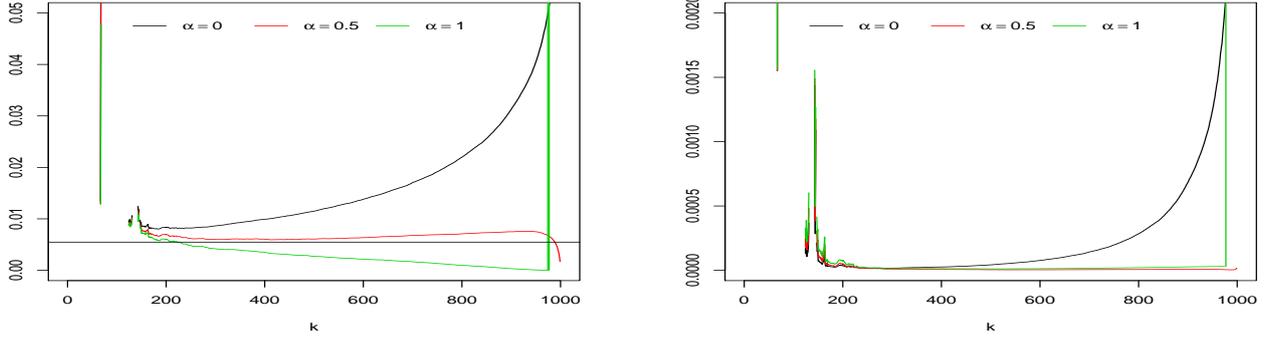
Figure 6: Mean values (left) and mean squared errors (right) of the estimation of $\mathbb{P}(X > 40, Y > 160) = 0.0054$ with $\widehat{\tau}_1 = -1$. Samples are from Gumbel(2) censored by S-clayton(1).
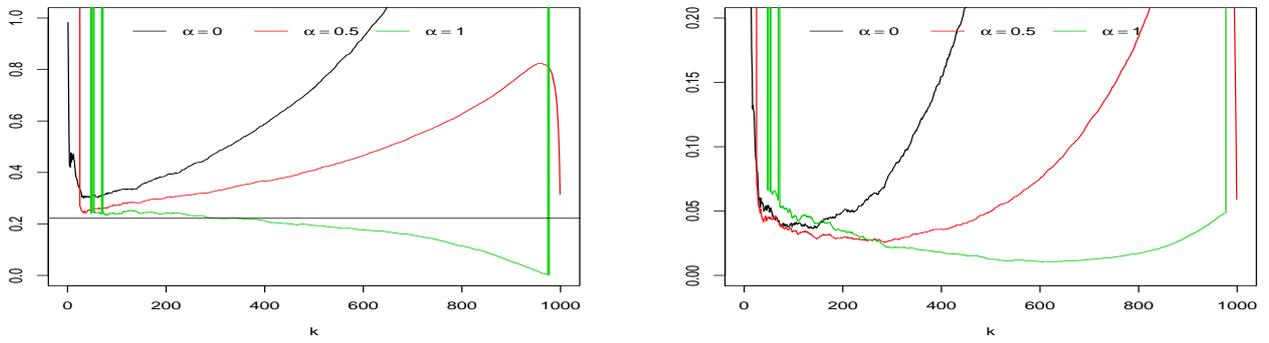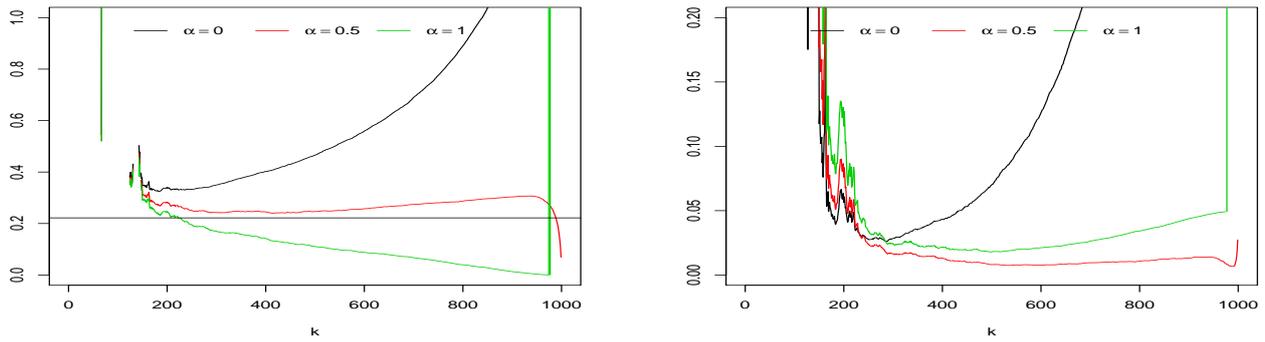

Figure 7: Mean values (left) and mean squared errors (right) of the estimation of $\mathbb{P}(Y > 160|X > 40) = 0.2219$ with $\widehat{\tau}_1 = 0$. Samples are from Gumbel(2) censored by S-clayton(1).

6 Proofs

PROOF OF LEMMA 2.1 By the assumption **A1** we have

$$
\begin{aligned}
\overline{F}(x) &= \mathbb{P}(Z > x) = \mathbb{P}(\min(X, \underline{w}Y) > x) \\
&= \mathbb{P}(X > x, Y > x/\underline{w}) = x^{-c_1} (x/\underline{w})^{-c_2} h_1(w)\Big(1 + x^{p_1} (x/\underline{w})^{p_2} h_2(w)(1 + o(1))\Big) \\
&= x^{-(c_1+c_2)} h_1(w)\underline{w}^{c_2}\Big(1 + x^{p_1+p_2} h_2(w)\underline{w}^{-p_2}(1 + o(1))\Big) \\
&=: x^{-1/\eta_1} C_1\Big(1 + D_1 x^{\tau_1}(1 + o(1))\Big).
\end{aligned}
$$

Consequently, $F \in Hall(\eta_1, C_1, D_1, \tau_1)$. Similarly,

$$
\begin{aligned}
\overline{G}(x) &= \mathbb{P}(\widetilde{Z} > x) = x^{-(\widetilde{c}_1+\widetilde{c}_2)} \widetilde{h}_1(w)\underline{w}^{\widetilde{c}_2}\Big(1 + x^{\widetilde{p}_1+\widetilde{p}_2} \widetilde{h}_2(w)\underline{w}^{-\widetilde{p}_2}(1 + o(1))\Big) \\
&=: x^{-1/\eta_2} C_2\Big(1 + D_2 x^{\tau_2}(1 + o(1))\Big),
\end{aligned}
$$

i.e., $G \in Hall(\eta_2, C_2, D_2, \tau_2)$. Finally, note that $Z^* = \min(X^*, \underline{w}Y^*) = \min(Z, \widetilde{Z})$ and $Z, \widetilde{Z}$ are independent of each other,

$$
\overline{H}(x) = \overline{F}(x)\overline{G}(x) = x^{-(1/\eta_1+1/\eta_2)} C_1 C_2\Big((1 + D_1 x^{\tau_1})(1 + D_2 x^{\tau_2})(1 + o(1))\Big)
$$

Figure 8: Mean values (left) and mean squared errors (right) of the estimation of $\mathbb{P}(Y > 160|X > 40) = 0.2219$ with $\widehat{\tau}_1 = -1$. Samples are from Gumbel(2) censored by S-clayton(1).
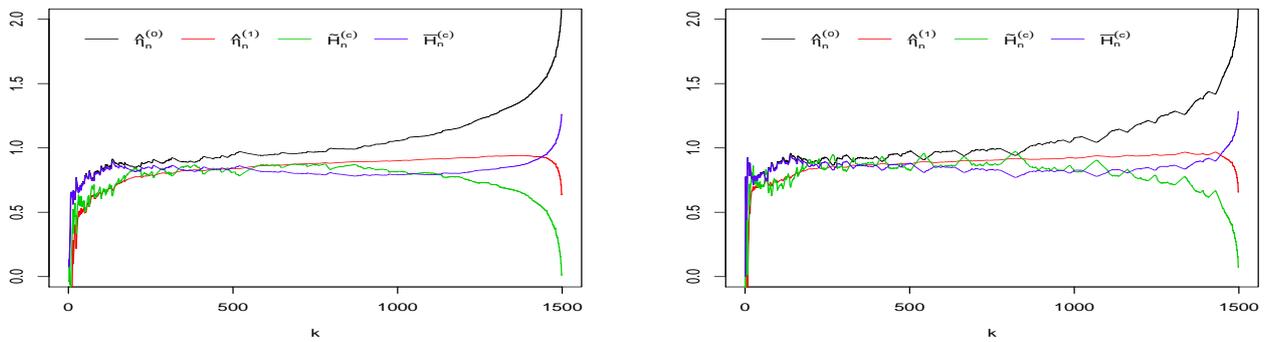

Figure 9: Estimations of the coefficient of tail dependence of (Loss, ALAE) considering the censoring case (left) and neglecting the censoring case (right).
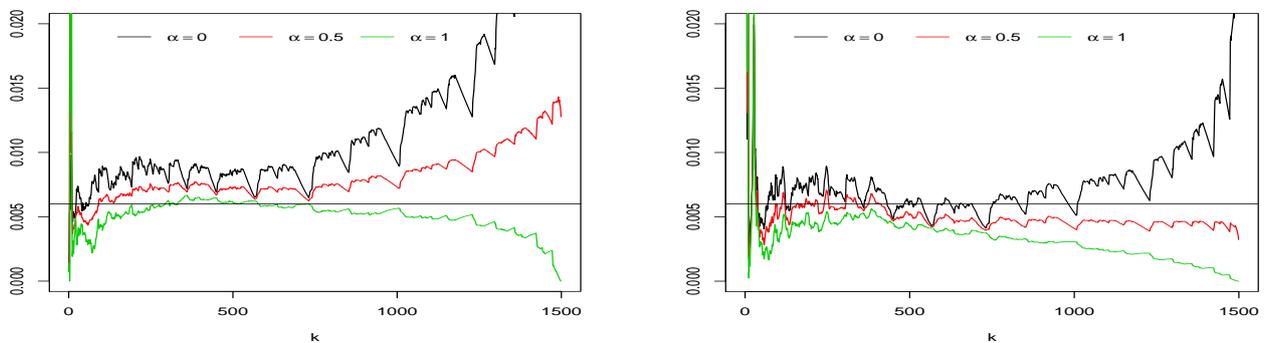

Figure 10: Estimations of the tail probability of $\mathbb{P}(\text{Loss} > 200000, \text{ALAE} > 100000))$ by estimators (3.2) with $\widehat{\tau}_1 = 0$ (left) and $\widehat{\tau}_1 = -1$ (right). The empirical tail probability 0.006 is indicated by the horizontal line.
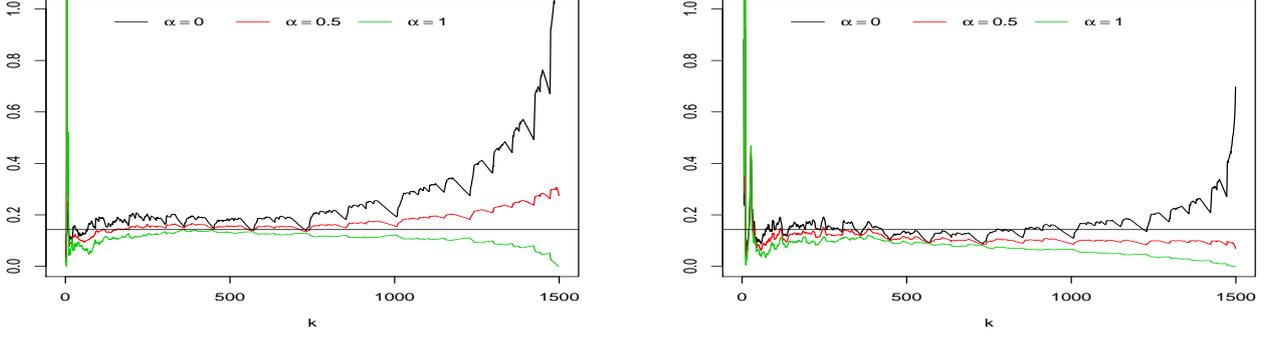
Figure 11: Estimations of the tail probability of $\mathbb{P}(\text{ALAE} > 100000|\text{Loss} > 200000))$ by estimators (3.2) with $\widehat{\tau}_1 = 0$ (left) and $\widehat{\tau}_1 = -1$ (right). The empirical tail probability 0.1428 is indicated by the horizontal line.

$$=: x^{-1/\eta} C\Big(1 + Dx^\tau(1 + o(1))\Big),$$

with

$$D = D_1 \mathbb{I}\{\tau_1 > \tau_2\} + D_2 \mathbb{I}\{\tau_1 < \tau_2\} + (D_1 + D_2)\mathbb{I}\{\tau_1 = \tau_2\} \neq 0, \quad \tau = \max(\tau_1, \tau_2),$$

i.e., $H \in Hall(\eta, C, D, \tau)$ and thus the proof is complete. $\qquad\qquad\square$

PROOF OF LEMMA 2.3 Let $\xi_{1,n} \leq \ldots \leq \xi_{n,n}$ be the associated order statistics of $\{\xi_i\}_{i=1}^n$ from $\xi \sim F_\xi(x) = 1 - 1/x, x \geq 1$. It follows from de Haan and Ferreira (2006) that $(k/n)\xi_{n-k,n} \overset{p}{\to} 1$ and

$$\{U(\xi_{n-i+1,n})\}_{i=1,\ldots,k} \overset{d}{=} \{Z^*_{n-i+1,n}\}_{i=1,\ldots,k}, \quad \{\xi_{n-i+1,n}/\xi_{n-k,n}\}_{i=1,\ldots,k} \overset{d}{=} \{\xi_{k-i+1,k}\}_{i=1,\ldots,k}$$

independent of $\xi_{n-k,n}$. By Lemma 2.1, $H \in Hall(\eta, C, D, \tau)$ and thus H satisfies the second-order condition (2.3), implying that

$$\frac{Z^*_{n-i+1,n}}{Z^*_{n-k,n}} \overset{d}{=} \frac{U(\xi_{n-i+1,n})}{U(\xi_{n-k+1,n})} \overset{d}{=} \xi^\eta_{k-i+1,k}\left(1 + \frac{\xi^\rho_{k-i+1,k} - 1}{\rho}b(n/k)(1 + o_p(1))\right), \quad i = 1, \ldots, k. \qquad (6.1)$$

By the uniform convergence theorem of regular varying function (cf. Theorem 2.3.9 in de Haan and Ferreira (2006))

$$M_n^{(j)}(k) = \frac{1}{k}\sum_{i=1}^k\left(\log\frac{Z^*_{n-i+1,n}}{Z^*_{n-k,n}}\right)^j \overset{d}{=} \frac{1}{k}\sum_{i=1}^k\left(\eta\log\xi_{k-i+1,k} + \frac{\xi^\rho_{k-i+1,k} - 1}{\rho}b(n/k)(1 + o_p(1))\right)^j$$

$$= \frac{\eta^j}{k}\sum_{i=1}^k(\log\xi_{k-i+1,k})^j + \frac{j\eta^{j-1}}{k}\sum_{i=1}^k(\log\xi_{k-i+1,k})^{j-1}\frac{\xi^\rho_{k-i+1,k} - 1}{\rho}b(n/k)(1 + o_p(1))$$

$$= \eta^j\Gamma(j+1) + \frac{\eta^j}{k}\sum_{i=1}^k((\log\xi_i)^j - \Gamma(j+1)) + \eta^j\Gamma(j+1)\frac{1/(1-\rho)^j - 1}{\rho}\frac{b(n/k)}{\eta}(1 + o_p(1))$$

$$= \eta^j\Gamma(j+1)\left(1 + \frac{\sigma_j}{\Gamma(j+1)}\frac{P_n^{(j)}}{\sqrt{k}} + \mu_{j,\rho}\frac{b(n/k)}{\eta}(1 + o_p(1))\right),$$

with $\sigma_j^2 = \Gamma(2j+1) - \Gamma^2(j+1), \mu_{j,\rho} = \frac{1/(1-\rho)^j - 1}{\rho}$ and

$$P_n^{(j)} = \frac{1}{\sigma_j\sqrt{k}}\sum_{i=1}^k((\log\xi_i)^j - \Gamma(j+1)) \sim \mathbb{A}s\mathbb{N}(0,1).$$

Similarly, for $N_n^{(\tau)}(k)$ defined in (2.1), we have

$$N_n^{(\tau)}(k) \overset{d}{=} \frac{1}{k}\sum_{i=1}^k\xi^\rho_{k-i+1,k}\left(1 + \frac{\xi^\rho_{k-i+1,k} - 1}{\rho}b(n/k)(1 + o_p(1))\right)^\tau$$

$$= \frac{1}{k} \sum_{i=1}^{k} \xi_{k-i+1,k}^{\rho} \left(1 + \frac{\xi_{k-i+1,k}^{\rho} - 1}{\eta} b(n/k)(1 + o_p(1)) \right)$$

$$= \frac{1}{1-\rho} + \frac{1}{k} \sum_{i=1}^{k} \left(\xi_i^{\rho} - \frac{1}{1-\rho} \right) + \frac{\rho}{(1-2\rho)(1-\rho)} \frac{b(n/k)}{\eta}(1 + o_p(1))$$

$$= \frac{1}{1-\rho} + \frac{\rho}{(1-\rho)\sqrt{1-2\rho}} \frac{K_n}{\sqrt{k}} + \frac{\rho}{(1-\rho)(1-2\rho)} \frac{b(n/k)}{\eta}(1 + o_p(1)),$$

with $\rho = \eta\tau$ and

$$K_n = \frac{(1-\rho)\sqrt{1-2\rho}}{\rho\sqrt{k}} \sum_{i=1}^{k} \left(\xi_i^{\rho} - \frac{1}{1-\rho} \right) \sim \mathbb{A}s\mathbb{N}(0,1).$$

By the Cramér–Wold device and Liapounov's theorem (cf. Chung (1974), p 200) for $j, j' = 1, 2$,

$$\mathbb{A}s\mathbb{C}(P_n^{(j)}, P_n^{(j')}) = \frac{\Gamma(j+j'+1) - \Gamma(j+1)\Gamma(j'+1)}{\sigma_j \sigma_{j'}}, \quad \mathbb{A}s\mathbb{C}(K_n, P_n^{(j)}) = \sqrt{1-2\rho}\,\Gamma(j+1)\frac{\mu_{j,\rho}}{\sigma_j}.$$

Therefore, it remains to prove the asymptotic distribution expansion of $\widehat{p}_n(k)$ as in (2.7). For that we shall use that

$$\mathbb{E}(\delta | Z^* = t) = \frac{\eta}{\eta_1} \left(1 + \left(\rho D t^{\tau} - \eta_1 \tau_1 D_1 t^{\tau_1} \right)(1 + o(1)) \right) \tag{6.2}$$

for large t, where $\mathbb{E}(\delta | Z^* = t) = \lim_{\varepsilon \to 0} \mathbb{E}(\delta | Z^* \in \Delta_\varepsilon)$ with $\Delta_\varepsilon = (t, t+\varepsilon)$ or $\Delta_\varepsilon = (t-\varepsilon, t)$. In fact, since $F \in Hall(\eta_1, C_1, D_1, \tau_1), H \in Hall(\eta, C, D, \tau)$ and $\ell_{0,F}(t) = (C_1 D_1 t^{\tau_1})^{-1}(t^{1/\eta_1}\overline{F}(t) - C_1)$ is a normalized regularly varying function, as $t \to \infty$

$$\frac{\partial \log \overline{F}(tx)}{\partial x} = -\frac{1}{\eta_1 x} + \frac{D_1 t^{\tau_1} x^{\tau_1 - 1}[\tau_1 \ell_{0,F}(tx) + tx\ell'_{0,F}(tx)]}{1 + D_1(tx)^{\tau_1}\ell_{0,F}(tx)} = \frac{1}{x} \left(\tau_1 D_1(tx)^{\tau_1}(1 + o(1)) - \frac{1}{\eta_1} \right),$$

$$\frac{\overline{H}(tx)}{\overline{H}(t)} = x^{-1/\eta} \left(1 + Dt^{\tau}(x^{\tau} - 1)(1 + o(1)) \right), \quad \frac{\overline{H}(t+\epsilon)}{\overline{H}(t)} = 1 + \frac{\epsilon}{\eta t} \left(\left(\rho D t^{\tau} - \frac{\epsilon}{2t}\left(1 + \frac{1}{\eta} \right) \right)(1 + o(1)) - 1 \right)$$

holds locally uniformly for all $x > 1$ and $\epsilon \in (0,1)$. Recall that $\delta = \delta^{(1)}\delta^* + \delta^{(2)}(1 - \delta^*)$ is the indicator function of the event $\{Z \leq \widetilde{Z}\}$, hence for large t

$$\mathbb{E}(\delta | Z^* = t) = \lim_{\varepsilon \to 0} \mathbb{P}(\delta = 1 | Z^* \in \Delta_\varepsilon)$$

$$= \lim_{\varepsilon \to 0} \frac{\mathbb{P}(Z \leq \widetilde{Z}, Z^* \in \Delta_\varepsilon)}{\mathbb{P}(Z^* \in \Delta_\varepsilon)} = \lim_{\varepsilon \to 0} \frac{\int_{x \in \Delta_\varepsilon} \overline{G}(x)\,\mathrm{d}F(x)}{\mathbb{P}(Z^* \in \Delta_\varepsilon)} = \lim_{\varepsilon \to 0} \frac{\int_{tx \in \Delta_\varepsilon} \overline{H}(tx)\,\mathrm{d}\log\overline{F}(tx)}{-\mathbb{P}(Z^* \in \Delta_\varepsilon)}$$

$$= \lim_{\varepsilon \to 0} \frac{\frac{1}{\varepsilon} \int_{tx \in \Delta_\varepsilon} x^{-1/\eta - 1}\left(1 + Dt^{\tau}(x^{\tau} - 1)(1 + o(1)) \right)\left(\tau_1 D_1 t^{\tau_1} x^{\tau_1}(1 + o(1)) - 1/\eta_1 \right)\mathrm{d}x}{-\mathbb{P}(Z^* \in \Delta_\varepsilon)/(\varepsilon\overline{H}(t))}$$

$$= \frac{\eta}{\eta_1} \left(1 + \rho D t^{\tau}(1 + o(1)) \right)\left(1 - \eta_1 \tau_1 D_1 t^{\tau_1}(1 + o(1)) \right)$$

$$= \frac{\eta}{\eta_1} \left(1 + \left(\rho D t^{\tau} - \eta_1 \tau_1 D_1 t^{\tau_1} \right)(1 + o(1)) \right)$$

thus, (6.2) is proved. Substituting t with $Z_{n-i+1,n}^*$ in (6.2) for $i = 1, \ldots, k$, we have

$$E_n(k) = \frac{1}{k} \sum_{i=1}^{k} \mathbb{E}(\delta | Z^* = Z_{n-i+1,n}^*)$$

$$= \frac{\eta}{\eta_1} \frac{1}{k} \sum_{i=1}^{k} \left(1 + \left(\left(\frac{U(\xi_{n-i+1,n})}{U(\xi_{n-k,n})} \right)^{\tau} \rho D (U(\xi_{n-k,n}))^{\tau} - \left(\frac{U(\xi_{n-i+1,n})}{U(\xi_{n-k,n})} \right)^{\tau_1} \eta_1 \tau_1 D_1 (U(\xi_{n-k,n}))^{\tau_1} \right)(1 + o_p(1)) \right)$$

$$\overset{d}{=} \frac{\eta}{\eta_1} \frac{1}{k} \sum_{i=1}^{k} \left(1 + \left(\xi_{k-i+1,k}^{\rho} \frac{b(n/k)}{\eta} - \xi_{k-i+1,k}^{\eta\tau_1} \frac{\eta_1 \tau_1 D_1}{(\rho D)^{\tau_1/\tau}} \left(\frac{b(n/k)}{\eta} \right)^{\tau_1/\tau} \right)(1 + o_p(1)) \right) \tag{6.3}$$

$$= \frac{\eta}{\eta_1} \left(1 + \left(\frac{1}{1-\rho} \frac{b(n/k)}{\eta} - \frac{1}{1-\eta\tau_1} \frac{\eta_1 \tau_1 D_1}{(\rho D)^{\tau_1/\tau}} \left(\frac{b(n/k)}{\eta} \right)^{\tau_1/\tau} \right)(1 + o_p(1)) \right)$$

$$V_n(k) = \frac{1}{k}\sum_{i=1}^{k}\mathbb{A}s\mathbb{V}(\delta|Z^* = Z_{n-i+1,n}^*) = \frac{1}{k}\sum_{i=1}^{k}(1 - \mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*))\mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*) \xrightarrow{p} \frac{\eta^2}{\eta_1\eta_2}.$$

Note that the concomitant order statistics $\{\delta_{[n-i+1,n]}\}_{i=1,\dots,k}$ is independently Bernoulli distributed with successful probability $\mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*)$, it follows that

$$\widehat{p}_n(k) = E_n(k) + \frac{1}{k}\sum_{i=1}^{k}(\delta_{[n-i+1,n]} - \mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*)) = E_n(k) + \frac{\eta}{\sqrt{\eta_1\eta_2}}\frac{J_n}{\sqrt{k}}$$

$$= \frac{\eta}{\eta_1}\left(1 + \sqrt{\frac{\eta_1}{\eta_2}}\frac{J_n}{\sqrt{k}} + \left(\frac{1}{1-\rho}\frac{b(n/k)}{\eta} - \frac{\eta_1\tau_1 D_1(\rho D)^{-\tau_1/\tau}}{1 - \eta\tau_1}\left(\frac{b(n/k)}{\eta}\right)^{\tau_1/\tau}\right)(1 + o_p(1))\right),$$

with

$$J_n = \frac{\sqrt{V_n(k)}}{\eta/\sqrt{\eta_1\eta_2}}\frac{\sum_{i=1}^{k}\left(\delta_{[n-i+1,n]} - \mathbb{E}(\delta|Z^* = Z_{n-i+1,n}^*)\right)}{\sqrt{kV_n(k)}} \sim \mathbb{A}s\mathbb{N}(0,1).$$

Finally, note that $\sum_{i=1}^{k}\delta_{[n-i+1,n]}$ relates only to $Z_{n-k,n}^*$, independent of $M_n^{(j)}(k)$ and $N_n^{(\tau)}(k)$. On the other hand, by the Cramér–Wold device and Liapounov's theorem (cf. Chung (1974), p 200), $E_n(k)$ is asymptotic independent of $M_n^{(j)}(k)$ and $N_n^{(\tau)}(k)$ (see (6.3)), thus J_n is independent of $M_n^{(j)}(k)$ and $N_n^{(\tau)}(k)$, hence the proof is complete. $\square$

PROOF OF THEOREM 3.1 Lemma 2.3 implies

$$\widehat{\eta}_n^{(\alpha)}(k) - \eta_1 = \frac{1}{\widehat{p}_n(k)}\left(\left(M_n^{(1)}(k) - \eta + \alpha\left(1 - \frac{1}{2}\left(1 - \frac{(M_n^{(1)}(k))^2}{M_n^{(2)}(k)}\right)^{-1}\right)\right) - \eta_1\left(\widehat{p}_n(k) - \frac{\eta}{\eta_1}\right)\right)$$

$$= \frac{\eta_1}{\eta}\left((\eta - 2\alpha)\left(\frac{M_n^{(1)}(k)}{\eta} - 1\right) + \alpha\left(\frac{M_n^{(2)}(k)}{2\eta^2} - 1\right) - \eta_1\left(\widehat{p}_n(k) - \frac{\eta}{\eta_1}\right)\right)(1 + o_p(1))$$

$$= \frac{\eta_1}{\eta}\left(\frac{(\eta - 2\alpha)P_n^{(1)} + \sqrt{5}\alpha P_n^{(2)} - \eta\sqrt{\eta_1/\eta_2}J_n}{\sqrt{k}}\right.$$

$$\left. + \left((\eta - 2\alpha)\mu_{1,\rho} + \alpha\mu_{2,\rho} - \eta\left(\frac{1}{1-\rho} - \frac{\eta_1\tau_1 D_1(\rho D)^{-\tau_1/\tau}}{1 - \eta\tau_1}\left(\frac{b(n/k)}{\eta}\right)^{\tau_1/\tau - 1}\right)(1 + o_p(1))\right)\frac{b(n/k)}{\eta}\right).$$

$$(6.4)$$

In view of condition (3.1) we obtain further

$$\mathbb{A}s\mathbb{V}\left((\eta - 2\alpha)P_n^{(1)} + \sqrt{5}\alpha P_n^{(2)} - \eta\sqrt{\eta_1/\eta_2}J_n\right) = \alpha^2 + \eta_1\eta,$$

$$\frac{\sqrt{k}b(n/k)}{\eta}\left((\eta - 2\alpha)\mu_{1,\rho} + \alpha\mu_{2,\rho} - \eta\left(\frac{1}{1-\rho} - \frac{\eta_1\tau_1 D_1(\rho D)^{-\tau_1/\tau}}{1 - \eta\tau_1}\left(\frac{b(n/k)}{\eta}\right)^{\tau_1/\tau - 1}\right)(1 + o_p(1))\right)$$

$$\xrightarrow{p} \frac{\lambda}{\eta}\left(\frac{\alpha\rho}{(1-\rho)^2} + \frac{\eta_1 D_1}{(1-\rho)D}\mathbb{I}\{\tau_1 \geq \tau_2\}\right)$$

and thus the claim follows. $\square$

PROOF OF THEOREM 3.3 We treat the cases $\tau_1 > \tau_2$ and $\tau_1 \leq \tau_2$ separately. For the case $\tau_1 > \tau_2$, by Lemma 2.1 $\overline{F} \in RV_{-1/\eta_1}$, it follows from the condition (3.4) that

$$y_n = \frac{x_n}{Z_{n-k,n}^*} \xrightarrow{p} \infty. \qquad (6.5)$$

Let $\widetilde{p}_n = \overline{F}(Z_{n-k,n}^*)\overline{G}_{\eta_1,\frac{b(n/k)}{\eta/\eta_1},\tau_1}(y_n)$ with $G_{\eta,\delta,\tau}(x) = 1 - [x(1 + \delta - \delta x^\tau)]^{-1/\eta}$. Rewrite

$$\frac{\sqrt{k}(\log\widehat{p}_n - \log p_n)}{\log(p_n/\overline{F}(Z_{n-k,n}^*))} = \frac{\sqrt{k}(\log\widetilde{p}_n - \log p_n)}{\log(p_n/\overline{F}(Z_{n-k,n}^*))} + \frac{\sqrt{k}(\log\widehat{p}_n - \log\widetilde{p}_n)}{\log(p_n/\overline{F}(Z_{n-k,n}^*))}.$$

We treat the two terms on the right-hand side separately. For the first summand term, it follows from Proposition 2.3 in Beirlant et al. (2009) that

$$\frac{\sqrt{k}(\log\widetilde{p}_n - \log p_n)}{\log(p_n/\overline{F}(Z_{n-k,n}^*))} = \frac{\sqrt{k}}{\log(p_n/\overline{F}(Z_{n-k,n}^*))}\left(\log\overline{G}_{\eta_1,\frac{b(n/k)}{\eta/\eta_1},\tau_1}(y_n) - \log\frac{\overline{F}(y_n Z_{n-k,n}^*)}{\overline{F}(Z_{n-k,n}^*)}\right) \xrightarrow{p} 0.$$

For the second summand term, rewrite

$$\frac{\sqrt{k}(\log\widehat{p}_n - \log\widetilde{p}_n)}{\log(p_n/\overline{F}(Z^*_{n-k,n}))} = \frac{\sqrt{k}}{\log(p_n/\overline{F}(Z^*_{n-k,n}))}\left(\log\frac{\widehat{\overline{F}}_n(Z^*_{n-k,n})}{\overline{F}(Z^*_{n-k,n})} + \log\frac{\overline{G}_{\widehat{\eta}_n^{(\alpha)}(k),\frac{\widehat{b}(n/k)}{\widehat{p}_n(k)},\widehat{\tau}_1}(y_n)}{\overline{G}_{\eta_1,\frac{b(n/k)}{\eta/\eta_1},\tau_1}(y_n)}\right) =: I_n + J_n.$$

It follows from Csörgő (1996) that $I_n \xrightarrow{d} 0$. Next we rewrite J_n as follows

$$J_n = \frac{\sqrt{k}}{\log(p_n/\overline{F}(Z^*_{n-k,n}))}\left(\log\left(y_n\left(1 + \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)} - \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)}y_n^{\widehat{\tau}_1}\right)\right)^{-1/\widehat{\eta}_n^{(\alpha)}(k)} - \log\left(y_n\left(1 + \frac{b(n/k)}{\eta/\eta_1} - \frac{b(n/k)}{\eta/\eta_1}y_n^{\tau_1}\right)\right)^{-1/\eta_1}\right)$$

$$= \frac{\sqrt{k}\log y_n}{\log(p_n/\overline{F}(Z^*_{n-k,n}))}\left(\frac{1}{\eta_1} - \frac{1}{\widehat{\eta}_n^{(\alpha)}(k)}\right) + \frac{\sqrt{k}}{\log(p_n/\overline{F}(Z^*_{n-k,n}))}\left(\frac{1}{\eta_1} - \frac{1}{\widehat{\eta}_n^{(\alpha)}(k)}\right)\log\left(1 + \frac{b(n/k)}{\eta/\eta_1} - \frac{b(n/k)}{\eta/\eta_1}y_n^{\tau_1}\right)$$

$$- \frac{\sqrt{k}}{\widehat{\eta}_n^{(\alpha)}(k)\log(p_n/\overline{F}(Z^*_{n-k,n}))}\left(\log\left(1 + \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)} - \frac{\widehat{b}(n/k)}{\widehat{p}_n(k)}y_n^{\widehat{\tau}_1}\right) - \log\left(1 + \frac{b(n/k)}{\eta/\eta_1} - \frac{b(n/k)}{\eta/\eta_1}y_n^{\tau_1}\right)\right)$$

$$=: J_{1n} + J_{2n} - J_{3n}.$$

It follows from Theorem 3.1 and the fact that $\overline{F} \in RV_{-1/\eta_1}$

$$J_{1n} = \frac{\log y_n}{\log(\overline{F}(y_n Z^*_{n-k,n})/\overline{F}(Z^*_{n-k,n}))}\frac{\sqrt{k}(\widehat{\eta}_n^{(\alpha)}(k) - \eta_1)}{\eta_1\widehat{\eta}_n^{(\alpha)}(k)} \xrightarrow{d} -\frac{\mathcal{Z}}{\eta_1}. \tag{6.6}$$

Similarly, note that $b(n/k) \to 0$ and the condition (3.4) holds, it follows that

$$J_{2n} = \frac{\sqrt{k}(\widehat{\eta}_n^{(\alpha)}(k) - \eta_1)}{\log(p_n/\overline{F}(Z^*_{n-k,n}))}\frac{b(n/k)(1 - y_n^{\tau_1})}{\eta\eta_1}(1 + o_p(1)) \xrightarrow{p} 0. \tag{6.7}$$

We consider next J_{3n}. We have

$$J_{3n} = -\frac{\sqrt{k}}{\log y_n}\left(\frac{\widehat{b}(n/k)}{\widehat{p}_n(k)}(1 - y_n^{\widehat{\tau}_1}) - \frac{b(n/k)}{\eta/\eta_1}(1 - y_n^{\tau_1})\right)(1 + o_p(1)) \xrightarrow{p} 0. \tag{6.8}$$

In view of condition (3.1) and the proved result (6.5), it is sufficient to prove that $\sqrt{k}\widehat{b}(n/k)$ is bounded. Note that $\widehat{\tau}$ is a consistent estimator of τ and

$$\widehat{b}(n/k) \overset{d}{=} \eta(1 - 2\rho)(1 - \rho)^3\rho^{-3}\left(N_n^{(\widehat{\tau})}(k) - \frac{1}{1 - \eta\widehat{\tau}} + \frac{1}{1 - \eta\widehat{\tau}} - \frac{1}{1 - \widehat{\rho}}\right)$$

$$= \eta(1 - 2\rho)(1 - \rho)^3\rho^{-3}\left(N_n^{(\widehat{\tau})}(k) - \frac{1}{1 - \eta\widehat{\tau}} - \frac{\rho}{\eta(1 - \rho)^2}(M_n^{(1)}(k) - \eta)\right)$$

$$= \frac{\eta(1 - 2\rho)(1 - \rho)}{\rho^2}\left(\frac{(1 - \rho)^2}{\rho}\left(N_n^{(\widehat{\tau})}(k) - \frac{1}{1 - \eta\widehat{\tau}}\right) - \left(\frac{M_n^{(1)}(k)}{\eta} - 1\right)\right).$$

Consequently, by Lemma 2.3, $\sqrt{k}\widehat{b}(n/k)$ is bounded. Consequently, combining (6.6), (6.7) and (6.8) the claim follows for the case $\tau_1 > \tau_2$.

For the case $\tau_1 \le \tau_2$, one can follow the line of the proof for the case $\tau_1 > \tau_2$, we leave it to the readers. $\qquad\square$

References

[1] Beirlant, J., Vandewalle, B. (2002) Some comments on the estimation of a dependence index in bivariate extreme value statistics. Statistics and Probability Letters. **60**: 265–278.

[2] Beirlant, J., Guillou A., Dierchx, G., Fils-Villetard, A. (2007) Estimation of the extreme value index and extreme quantiles under censoring. Extremes. **10**: 151–174.

[3] Beirlant, J., Joossens, E., Segers, J. (2009) Second-order refined peaks-over-threshold modeling for heavy-tailed distributions. Journal of Statistical Planning and Inference. **139**: 2800–2815.

[4] Beirlant, J., Guillou A., Toulemonde, G. (2010) Peaks-over-threshold modeling under random censoring. Communications in Statistics - Theory and Methods. **39**: 1158–1179.

[5] Beirlant, J., Dierchx, G., Guillou, A. (2011) Bias-reduced estimators for bivariate tail modeling. Insurance: Mathematics and Economics. **49**: 18–26.

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

[7] Caeiro, F., Gomes, M.I., Pestana, D. (2005) Direct reduction of bias of the classical Hill estimator. Revstat– Statistics Journal. **3**: 113–136.

[8] Chung, K.L. (1974) *A course in Probability Theory,* 2nd ed. Academic, New York.

[9] Csörgő, S. (1996) Universal gaussian approximations under random censoringship. The Annals of Statistics. **24**: 2744–2788.

[10] Dekkers, A.L.M. , Einmahl, J.H.J., de Haan, L. (1989) A moment estimator for the index of an extreme-value distribution. The Annals of Statistics. **17**: 1833–1855.

[11] de Haan, L., Ferreira, A. (2006) *Extreme Value Theory. An Introdution.* Springer, New York.

[12] de Wet, T., Goegebeur, Y., Munch, M.R. (2012) Asymptotically unbiased estimation of the second order tail parameter. Statistics and Probability Letters. **82**: 565–573.

[13] Einmahl, J.H.J, Fils-Villetard, A., Guillou, A. (2008) Statistics of extremes under randomly censoring. Bernoulli. **14**: 207–227.

[14] Einmahl, J.H.J., Krajina, A., Segers, J. (2012) An M-estimator for tail dependence in arbitrary dimensions. The Annals of Statistics. **40**: 1764–1793.

[15] Embrechts, P., Klüppelberg, C., Mikosch, T. (1997) *Modeling Extremal Events for Insurance and Finance.* Springer-Verlag, Berlin.

[16] Goegebeur, Y., Guillou, A. (2013) Asymptotically unbiased estimation of the coefficient of tail dependence. Scandinavian Journal of Statistics. **40**: 174–189.

[17] Gomes, M.I., de Haan, L., Henriques, L. (2008) Tail index estimation for heavy-tailed models: accommodation of bias in weighted log-excesses. Journal of the Royal Statistical Society, Series B. **70**: 31–52.

[18] Gomes, M.I., Mendonça, S., Pestana, D. (2009) Adaptive reduced-bias tail index and value-at-risk estimation. The XIII international conference Applied stochastic models and data analysis. 41–44.

[19] Gomes, M.I., Neves, M.M. (2011) Estimation of the extreme value index for randomly censored data. Biometrical Letters. **48**: 1–22.

[20] Hashorva, E. (2010) On the residual dependence index of elliptical distributions. Statistics and Probability Letters. **80**: 1070–1078.

[21] Hashorva, E. (2012) Exact tail asymptotics in bivariate scale mixture models. Extremes. **15**: 109–128.

[22] Haug, S., Klüppelberg, C., Peng, L. (2011) Statistical models and methods for dependence in insurance data. Journal Korean Statistical Society. **40**: 125–139.

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

[24] Kaplan, E.L., Meier, P. (1958) Non-parametric estimation from incomplete observations. Journal of the American Statistical Association. **53**: 457–481.

[25] Ledford, A.W., Tawn, J.A. (1997) Modeling dependence within joint tail regions. Journal of the Royal Statistical Society. Series B. **59**: 475–499.

[26] Peng, L. (1999) Estimation of the coefficient of tail dependence in bivariate extremes. Statistics and Probability Letters. **43**: 399–409.