Bernoulli **13**(1), 2007, 229–251
DOI: 10.3150/07-BEJ6047

Estimating the tail dependence function of an elliptical distribution

CLAUDIA KLÜPPELBERG[1,*], GABRIEL KUHN[1,**] and LIANG PENG[2]

[1]*Center for Mathematical Sciences, Munich University of Technology, D-85747 Garching, Germany. E-mail: *cklu@ma.tum.de; **gabriel@ma.tum.de*
[2]*School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332-0160, USA. E-mail: peng@math.gatech.edu*

Recently there has been growing interest in applying elliptical distributions to risk management. Under certain conditions, Hult and Lindskog show that a random vector with an elliptical distribution is in the domain of attraction of a multivariate extreme value distribution. In this paper we study two estimators for the tail dependence function, which are based on extreme value theory and the structure of an elliptical distribution. After deriving second-order regular variation estimates and proving asymptotic normality for both estimators, we show that the estimator based on the structure of an elliptical distribution is better than that based on extreme value theory in terms of both asymptotic variance and optimal asymptotic mean squared error. Our simulation study further confirms this.

Keywords: asymptotic normality; elliptical distribution; regular variation; tail dependence function

1. Introduction

Let (X, Y), (X_1, Y_1), (X_2, Y_2), ... be independent random vectors with common distribution function F and continuous marginals F_X and F_Y. Define the *tail dependence function*

$$\lambda(x, y) := \lim_{t \to 0} \frac{1}{t} \Pr(1 - F_X(X) \leqslant tx, 1 - F_Y(Y) \leqslant ty)$$

for $x, y \geqslant 0$. Then $\lambda(1, 1)$ is called the *upper tail dependence coefficient* (see Definition 2.3 of Hult and Linskog 2002), and $l(x, y) := x + y - \lambda(x, y)$ is called the *stable tail dependence function* (Huang 1992: 26). For more details on copulae and tail dependence, see Joe (1997). Assuming that (X, Y) is in the domain of attraction of a bivariate extreme value distribution, there exist several estimators for $l(x, y)$; see Huang (1992), Einmahl *et al.* (1993) and de Haan and Resnick (1993). The optimal rate of convergence for estimating $l(x, y)$ is given by Drees and Huang (1998), and a weighted tail approximation is provided by Einmahl *et al.* (2006). An alternative method for estimating $l(x, y)$ is via estimating the spectral measure; see Einmahl *et al.* (1997, 2001). On modelling the dependence of extremes parametrically, we refer to Tawn (1988) and Ledford and Tawn (1997).

Triggered by financial risk management problems, we observe growing interest in elliptical distributions as natural extensions of the normal family allowing for the modelling

1350–7265

of heavy tails and extreme dependence; see Chapters 3 and 5 of McNeil *et al.* (2005) and Abdous *et al.* (2005). The vector (X, Y) is *elliptically distributed* if

$$(X, Y)^{\mathrm{T}} = \boldsymbol{\mu} + G\mathbf{A}\boldsymbol{U}^{(2)}, \tag{1}$$

where $\boldsymbol{\mu} = (\mu_X, \mu_Y)^{\mathrm{T}}$ is a location vector, $G > 0$ is a random variable, called a *generating variable*, $\mathbf{A} \in \mathbb{R}^{2\times 2}$ is a deterministic matrix with

$$\mathbf{A}\mathbf{A}^{\mathrm{T}} =: \boldsymbol{\Sigma} := \begin{pmatrix} \sigma^2 & \rho\sigma\upsilon \\ \rho\sigma\upsilon & \upsilon^2 \end{pmatrix}$$

and $\mathrm{rank}(\boldsymbol{\Sigma}) = 2$, $\boldsymbol{U}^{(2)}$ is a two-dimensional random vector uniformly distributed on the unit sphere $\mathcal{S}_2 := \{z \in \mathbb{R}^2 : \|z\| = 1\}$, and $\boldsymbol{U}^{(2)}$ is independent of G. Throughout we use the Euclidean norm.

Note that ρ is termed the linear correlation coefficient of $\boldsymbol{\Sigma}$. Under certain conditions, Theorem 4.3 of Hult and Lindskog (2002) shows that regular variation of $\Pr(G > \cdot)$ with index $\alpha > 0$, i.e. $\lim_{t\to\infty} \Pr(G > tx)/\Pr(G > t) = x^{-\alpha}$ for all $x > 0$ (notation: $\Pr(G > \cdot) \in \mathcal{R}_{-\alpha}$), implies regular variation of (X, Y) with the same index $\alpha > 0$; see Section 5.4.2 of Resnick (1987) for the definition of multivariate regular variation. Moreover, if $\Pr(G > \cdot) \in \mathcal{R}_{-\alpha}$, then

$$\lambda(1, 1) = \left(\int_{(\pi/2-\arcsin\rho)/2}^{\pi/2} (\cos\phi)^\alpha \,\mathrm{d}\phi\right)\left(\int_0^{\pi/2} (\cos\phi)^\alpha \,\mathrm{d}\phi\right)^{-1}. \tag{2}$$

Here we are interested in estimating the tail dependence function $\lambda(x, y)$ by assuming that $\Pr(G > \cdot) \in RV_{-\alpha}$ for some $\alpha > 0$. With the help of elliptical distributions, some part of the tail dependence function can be estimated via the whole sample and another part by only employing the data in the tail region of the sample. Therefore, modelling the dependence of multivariate extremes via elliptical distributions avoids the difficulty of the dimensional curse and provides a robust way of dealing with tail structures. Another advantage of employing elliptical distributions in modelling extremes is the simplicity of simulating multivariate extremes. Since $\Pr(G > \cdot) \in RV_{-\alpha}$ implies that (X, Y) is in the domain of attraction of an extreme value distribution, a naive procedure is to apply Huang's estimator by ignoring the structure of the elliptical distribution, i.e.

$$\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{Hu}},n}(x, y) := \frac{1}{k_{\mathrm{Hu}}} \sum_{i=1}^{n} \boldsymbol{I}(X_i \geqslant X_{(n-\lfloor xk_{\mathrm{Hu}}\rfloor,n)},\ Y_i \geqslant Y_{(n-\lfloor yk_{\mathrm{Hu}}\rfloor,n)}), \tag{3}$$

where $X_{(1,n)} \leqslant \ldots \leqslant X_{(n,n)}$ and $Y_{(1,n)} \leqslant \ldots \leqslant Y_{(n,n)}$ denote the order statistics of $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$, respectively, $k_{\mathrm{Hu}} = k_{\mathrm{Hu}}(n) \overset{n\to\infty}{\longrightarrow} \infty$ and $k_{\mathrm{Hu}}/n \overset{n\to\infty}{\longrightarrow} 0$. The same estimator has been analysed by Schmidt and Stadtmüller (2006); see their equation (4.13). The aim of this paper is twofold. Firstly, we suggest a new estimator which exploits the structure of an elliptical distribution similar to (2). Secondly, we aim to determine the optimal number of order statistics to be used in both estimators. The choice will be based on the asymptotic mean squared error of the estimators.

The paper is organized as follows. We first derive an expression for $\lambda(x, y)$ which generalizes equation (2), and then construct a new estimator for $\lambda(x, y)$ via this expression;

see Section 2 for details. After deriving the second-order behaviours for elliptical distributions and the limiting distributions of these two estimators in Section 2, we show that the new estimator is better than the naive empirical estimator from Huang in terms of both asymptotic variance and optimal asymptotic mean squared error in Section 3. More importantly, the optimal choice of the sample fraction for the new estimator is the same as that for Hill's estimator (Hill 1975). That is, all data-driven methods for choosing the optimal sample fraction for Hill's estimator can be applied to our new estimator directly. A simulation study is provided in Section 3 as well. All proofs are summarized in Section 4.

2. Methodology and main results

The following theorem gives an expression for $\lambda(x, y)$ which will be employed to construct an estimator.

Theorem 1. *Suppose* (X, Y), *defined in* (1), *holds with* $\sigma > 0$, $\upsilon > 0$, $|\rho| < 1$ *and* $\Pr(G > \cdot) \in \mathcal{R}_{-\alpha}$ *for some* $\alpha > 0$. *Further, define*

$$g(t) := \arctan((t-\rho)/\sqrt{1-\rho^2}) \in [-\arcsin\rho, \pi/2], \qquad t \geqslant 0.$$

Then, for $x, y \geqslant 0$,

$$\lambda(x, y) = \left(\int_{g((x/y)^{1/\alpha})}^{\pi/2} x(\cos\phi)^\alpha \, d\phi + \int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})} y(\rho\cos\phi + \sqrt{1-\rho}\sin\phi)^\alpha \, d\phi \right) \times \left(\int_{-\pi/2}^{\pi/2} (\cos\phi)^\alpha \, d\phi \right)^{-1}.$$

In order to derive the asymptotic normality of $\hat{\lambda}^{\text{Hu}}_{k_{\text{Hu}},n}(x, y)$, it is known that a second-order condition is needed. Here we seek the relation of the second-order behaviour among the tail dependence function $\lambda(x, y)$, $\sqrt{X^2 + Y^2}$ and G; see the next two theorems for details.

In the setting of (1), assume that there exists $A(t) \to 0$ such that, for all $x > 0$ and some $\beta \leqslant 0$,

$$\lim_{t\to\infty} \frac{\Pr(G > tx)/\Pr(G > t) - x^{-\alpha}}{A(t)} = x^{-\alpha}\frac{x^\beta - 1}{\beta}, \tag{1}$$

where $\beta \leqslant 0$ is called a *second-order regular variation parameter*; see de Haan and Stadtmüller (1996) for more details on second-order regular variation. Additionally, we assume that

$$\lim_{t\to\infty} t^2 A(t) =: a \in [-\infty, \infty]. \tag{2}$$

Since $A \in \mathcal{R}_\beta$, $a = 0$ for $\beta < -2$ and $|a| = \infty$ for $\beta \in (-2, 0]$.

The following two theorems derive the corresponding second-order condition for $\sqrt{X^2+Y^2}$ and the tail dependence function $\lambda(x, y)$. Note that $\rho\cos\phi + \sqrt{1-\rho}\sin\phi = \sin(\phi + \arcsin\rho)$.

Theorem 2. *Assume that the conditions of Theorem 1,* (1) *and* (2) *hold. Further, define, for* $\phi \in (-\pi/2, \pi/2)$,

$$d_1(\phi) = \sigma^2(\cos\phi)^2 + \upsilon^2(\sin(\phi + \arcsin\rho))^2,$$

$$d_2(\phi) = \mu_X\sigma\cos\phi + \mu_Y\upsilon\sin(\phi + \arcsin\rho).$$

Then, for all $x > 0$,

$$\begin{aligned}\lim_{t\to\infty} & \frac{\Pr(\sqrt{X^2+Y^2} > tx)/\Pr(\sqrt{X^2+Y^2} > t) - x^{-\alpha}}{t^{-2} + |A(t)|} \\ &= x^{-\alpha}\left(\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi\right)^{-1}\left\{\frac{a}{1+|a|}\frac{x^\beta - 1}{\beta}\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha-\beta/2}\,\mathrm{d}\phi\right. \\ &\quad \left.+\frac{1}{1+|a|}\frac{\alpha}{2}(x^{-2}-1)\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2-1}[\alpha(d_2(\phi))^2 + d_1(\phi)(\mu_X^2+\mu_Y^2)]\mathrm{d}\phi\right\}. \end{aligned} \tag{3}$$

Also, for all $x > 0$ *and* $V(x) := \inf\{y : \Pr(\sqrt{X^2+Y^2} > y) \leqslant x^{-1}\}$,

$$\begin{aligned}\lim_{t\to\infty} & \frac{V(tx)/V(t) - x^{1/\alpha}}{(F_Y^{\leftarrow}(1-t^{-1}))^{-2} + |A(F_Y^{\leftarrow}(1-t^{-1}))|} \\ &= x^{1/\alpha}\left(\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi\right)^{-1}\left(\frac{\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi}{\upsilon^\alpha\int_0^\pi(\sin\phi)^\alpha\,\mathrm{d}\phi}\right)^{-(2\wedge|\beta|)/\alpha} \\ &\quad \times\left\{\frac{a}{1+|a|}\frac{x^{\beta/\alpha}-1}{\alpha\beta}\left(\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha-\beta/2}\,\mathrm{d}\phi\right)\right. \\ &\quad \left.+\frac{1}{1+|a|}\frac{1}{2}(x^{-2/\alpha}-1)\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2-1}[\alpha(d_2(\phi))^2 + d_1(\phi)(\mu_X^2+\mu_Y^2)]\mathrm{d}\phi\right\} \\ &=: x^{1/\alpha}\mathcal{B}_{(4)}(x), \end{aligned} \tag{4}$$

where $F^{\leftarrow}$ *denotes the generalized inverse of F. In particular, when* $\mu_X = \mu_Y = 0$, *we have, for all* $x > 0$,

$$\lim_{t\to\infty}\frac{V(tx)/V(t)-x^{1/\alpha}}{A(F_Y^{\leftarrow}(1-t^{-1}))}=\upsilon^{-\beta}x^{1/\alpha}\frac{x^{\beta/\alpha}-1}{\alpha\beta}\left(\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi\right)^{-1}$$
$$\times\int_{-\pi}^{\pi}(d_1(\phi))^{(\alpha-\beta/2)}\,\mathrm{d}\phi\left(\frac{\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi}{\int_0^{\pi}(\sin\phi)^{\alpha}\,\mathrm{d}\phi}\right)^{\beta/\alpha}$$
$$=: x^{1/\alpha}\mathcal{B}_{(5)}(x). \tag{5}$$

Theorem 3. *Assume that the conditions of Theorem 1 and* (1) *hold. Define, for* $x<0$,

$$\mathcal{B}_{(6)}(x):=-x\frac{x^{-\beta/\alpha}-1}{\beta}\left(\int_0^{\pi}(\sin\phi)^{\alpha}\,\mathrm{d}\phi\right)^{-1}\left(\int_0^{\pi}(\sin\phi)^{\alpha-\beta}\,\mathrm{d}\phi\right) \tag{6}$$

and

$$\mathcal{S}_2^+:=\{\boldsymbol{z}\in\mathbb{R}^2:\boldsymbol{z}\geqslant\boldsymbol{0}\text{ and }\|\boldsymbol{z}\|=1\}.$$

Then

$$\lim_{t\to 0}\frac{t^{-1}\Pr(F_X(X)\geqslant 1-tx,\,F_Y(Y)\geqslant 1-ty)-\lambda(x,y)}{A(F_Y^{\leftarrow}(1-t))}$$
$$=\upsilon^{-\beta}\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^{\alpha}\,\mathrm{d}\phi\right)^{-1}\left\{\frac{x}{\beta}\int_{g((x/y)^{1/\alpha})}^{\pi/2}[x^{-\beta/\alpha}(\cos\phi)^{\alpha-\beta}-(\cos\phi)^{\alpha}]\mathrm{d}\phi\right.$$
$$+\frac{y}{\beta}\int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})}[y^{-\beta/\alpha}(\sin(\phi+\arcsin\rho))^{\alpha-\beta}-(\sin(\phi+\arcsin\rho))^{\alpha}]\mathrm{d}\phi$$
$$+\mathcal{B}_{(6)}(x)\int_{g((x/y)^{1/\alpha})}^{\pi/2}(\cos\phi)^{\alpha}\,\mathrm{d}\phi+\mathcal{B}_{(6)}(y)\int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})}(\sin(\phi+\arcsin\rho))^{\alpha}\,\mathrm{d}\phi$$
$$\left.-\lambda(x,y)\frac{1}{\beta}\int_{-\pi/2}^{\pi/2}(\cos\phi)^{\alpha}((\cos\phi)^{-\beta}-1)\mathrm{d}\phi\right\}$$
$$=: \mathcal{B}_{(7)}(x,y) \tag{7}$$

holds for all $x,\ y\geqslant 0$ *and uniformly on* $\mathcal{S}_2^+$.

We are now ready to define our new estimator. Set $Z_i=\sqrt{X_i^2+Y_i^2}$, for $i=1,\ldots,n$, and let $Z_{(1,n)}\leqslant\ldots\leqslant Z_{(n,n)}$ denote their order statistics. First we estimate the index α by Hill's estimator, which is defined as

$$\hat{\alpha}^{\mathrm{H}}_{k_{\mathrm{El}},n}:=\left(\frac{1}{k_{\mathrm{El}}}\sum_{i=1}^{k_{\mathrm{El}}}\ln Z_{(n-i+1,n)}-\ln Z_{(n-k_{\mathrm{El}},n)}\right)^{-1},$$

where $k_{\mathrm{El}} = k_{\mathrm{El}}(n) \to \infty$ and $k_{\mathrm{El}}/n \to 0$ as $n \to \infty$. Now let (X, Y) and $(\tilde{X}, \tilde{Y})$ be independently and identically elliptically distributed. Then it follows from Theorem 4.2 of Hult and Lindskog (2002) that $\tau = (2/\pi)\arcsin\rho$, where τ is called *Kendall's tau* and is defined by

$$\tau := \Pr((X - \tilde{X})(Y - \tilde{Y}) > 0) - \Pr((X - \tilde{X})(Y - \tilde{Y}) < 0).$$

As usual, we estimate Kendall's tau by

$$\hat{\tau}_n := \frac{2}{n(n-1)} \sum_{1 \leqslant i \leqslant j \leqslant n} \operatorname{sign}((X_i - X_j)(Y_i - Y_j)),$$

which results in estimating ρ by

$$\hat{\rho}_n = \sin\left(\frac{\pi}{2}\,\hat{\tau}_n\right).$$

Hence, we can estimate $\lambda(x, y)$ by replacing ρ and α in Theorem 1 by $\hat{\rho}_n$ and $\hat{\alpha}^{\mathrm{H}}_{k_{\mathrm{El}},n}$, respectively. Let us denote this estimator by

$$\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x, y). \tag{8}$$

We remark that $\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(1, 1)$ was mentioned in Section 6 of Schmidt (2003), but without further study. The following theorem shows the asymptotic normality of $\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{Hu}},n}(x, y)$ and $\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x, y)$, which allows us to compare these two estimators theoretically.

Theorem 4. *Assume that the conditions of Theorem 1 and* (1) *hold. Suppose* $k_{\mathrm{Hu}} = k_{\mathrm{Hu}}(n) \overset{n\to\infty}{\longrightarrow} \infty$, $k_{\mathrm{Hu}}/n \overset{n\to\infty}{\longrightarrow} 0$ *and*

$$\sqrt{k_{\mathrm{Hu}}}A(F_Y^{\leftarrow}(1 - k_{\mathrm{Hu}}/n)) \overset{n\to\infty}{\longrightarrow}: \mathcal{K}_{\mathrm{Hu}},$$

for $|\mathcal{K}_{\mathrm{Hu}}| < \infty$. *Then, as* $n \to \infty$,

$$\sup_{0 \leqslant x, y \leqslant T}\left|\sqrt{k_{\mathrm{Hu}}}(\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{Hu}},n}(x, y) - \lambda(x, y)) - \mathcal{K}_{\mathrm{Hu}}\mathcal{B}_{(7)}(x, y) - B(x, y)\right| = o_p(1), \tag{9}$$

for every $T > 0$, *where* $\mathcal{B}_{(7)}(x, y)$ *is defined in Theorem 3,*

$$B(x, y) = W(x, y) - \left(1 - \frac{\partial}{\partial x}\lambda(x, y)\right)W(x, 0) - \left(1 - \frac{\partial}{\partial y}\lambda(x, y)\right)W(0, y),$$

and $W(x, y)$ *is a Wiener process with zero mean and covariance structure*

$$\begin{aligned}
&\mathrm{E}(W(x_1, y_1)W(x_2, y_2)) \\
&\quad = x_1 \wedge x_2 + y_1 \wedge y_2 - \lambda(x_1 \wedge x_2, y_1) - \lambda(x_1 \wedge x_2, y_2) - \lambda(x_1, y_1 \wedge y_2) \\
&\qquad - \lambda(x_2, y_1 \wedge y_2) + \lambda(x_1, y_2) + \lambda(x_2, y_1) + \lambda(x_1 \wedge x_2, y_1 \wedge y_2).
\end{aligned}$$

Therefore, for every fixed $x,\ y > 0$,

$$\sqrt{k_{\mathrm{Hu}}}(\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{Hu}},n}(x,\ y)-\lambda(x,\ y))\xrightarrow{d}\mathcal{N}(\mathcal{K}_{\mathrm{Hu}}\mathcal{B}_{(7)}(x,\ y),\ \sigma^2_{\mathrm{Hu}})$$

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

$$\sigma^2_{\mathrm{Hu}}=x\left(\frac{\partial}{\partial x}\lambda(x,\ y)\right)^2+y\left(\frac{\partial}{\partial y}\lambda(x,\ y)\right)^2$$
$$+2\lambda(x,\ y)\left(\frac{1}{2}-\frac{\partial}{\partial x}\lambda(x,\ y)-\frac{\partial}{\partial y}\lambda(x,\ y)+\frac{\partial}{\partial x}\lambda(x,\ y)\frac{\partial}{\partial y}\lambda(x,\ y)\right),\tag{10}$$

$$\frac{\partial}{\partial x}\lambda(x,\ y)=\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^\alpha\,\mathrm{d}\phi\right)^{-1}\int_{g((x/y)^{1/\alpha})}^{\pi/2}(\cos\phi)^\alpha\,\mathrm{d}\phi,\tag{11}$$

$$\frac{\partial}{\partial y}\lambda(x,\ y)=\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^\alpha\,\mathrm{d}\phi\right)^{-1}\int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})}(\sin(\phi+\arcsin\rho))^\alpha\,\mathrm{d}\phi.\tag{12}$$

Theorem 5. *Assume that the conditions of Theorem 1 and* (1) *hold. Further, assume* (2) *holds when* $\boldsymbol{\mu}\neq\mathbf{0}$. *Suppose* $k_{\mathrm{El}}=k_{\mathrm{El}}(n,\ \boldsymbol{\mu})\overset{n\to\infty}{\longrightarrow}\infty$, $k_{\mathrm{El}}/n\overset{n\to\infty}{\longrightarrow}0$ *and*

$$\sqrt{k_{\mathrm{El}}}((F_Y^{\leftarrow}(1-k_{\mathrm{El}}/n))^{-2}+|A(F_Y^{\leftarrow}(1-k_{\mathrm{El}}/n))|)\overset{n\to\infty}{\longrightarrow}:\mathcal{K}_{\mathrm{El}},\qquad\boldsymbol{\mu}\neq\mathbf{0},$$

$$\sqrt{k_{\mathrm{El}}}A(F_Y^{\leftarrow}(1-k_{\mathrm{El}}/n))\overset{n\to\infty}{\longrightarrow}:\mathcal{K}_{\mathrm{El}},\qquad\boldsymbol{\mu}=\mathbf{0},$$

for $|\mathcal{K}_{\mathrm{El}}|<\infty$. *Then, as* $n\to\infty$,

$$\sup_{0\leqslant x,y\leqslant T}|\sqrt{k_{\mathrm{El}}}(\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x,\ y)-\lambda(x,\ y))-\mathcal{B}_{(15)}(x,\ y)Z_0|=o_p(1),\tag{13}$$

where $Z_0\sim\mathcal{N}\ (-\alpha^2\mathcal{K}_{\mathrm{El}}\mathcal{B}_{(14)},\ \alpha^2)$ *with*

$$\mathcal{B}_{(14)}:=\begin{cases}\int_0^1\mathcal{B}_{(4)}(1/s)\mathrm{d}s, & \boldsymbol{\mu}\neq\mathbf{0},\\ \int_0^1\mathcal{B}_{(5)}(1/s)\mathrm{d}s, & \boldsymbol{\mu}=\mathbf{0},\end{cases}\tag{14}$$

$\mathcal{B}_{(4)}(s)$ *and* $\mathcal{B}_{(5)}(s)$ *are defined in Theorem 2 and*

$$\mathcal{B}_{(15)}(x,\ y):=\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^\alpha\,\mathrm{d}\phi\right)^{-1}\left\{\int_{g((x/y)^{1/\alpha})}^{\pi/2}x(\cos\phi)^\alpha\ln(\cos\phi)\mathrm{d}\phi\right.$$
$$+\int_{-\arcsin\rho}^{g((x/y)^{1/\alpha})}y(\sin(\phi+\arcsin\rho))^\alpha\ln(\sin(\phi+\arcsin\rho))\mathrm{d}\phi$$
$$\left.-\lambda(x,\ y)\int_{-\pi/2}^{\pi/2}(\cos\phi)^\alpha\ln(\cos\phi)\mathrm{d}\phi\right\}.\tag{15}$$

Therefore, for every fixed $x,y>0$,

$$\sqrt{k_{\mathrm{El}}}(\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x,\ y) - \lambda(x,\ y)) \xrightarrow{d} \mathcal{N}(-\alpha^2 \mathcal{K}_{\mathrm{El}} \mathcal{B}_{(14)} \mathcal{B}_{(15)}(x,\ y),\ \alpha^2(\mathcal{B}_{(15)}(x,\ y))^2).$$

The next corollary gives the optimal choice of the sample fraction for both estimators. As our criterion we use the *asymptotic mean squared error* of $\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{Hu}},n}$ and $\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}$, denoted by $\mathrm{amse}_{\mathrm{Hu}}(k_{\mathrm{Hu}})$ and $\mathrm{amse}_{\mathrm{El}}(k_{\mathrm{El}})$, respectively.

Corollary 6. *Assume that the conditions of Theorems 4 and 5 hold. Further, suppose that*

$$A(F_Y^{\leftarrow}(1-t)) = b_0 t^{-\beta/\alpha}(1+o(1)),$$

$$(F_Y^{\leftarrow}(1-t))^{-2} + |A(F_Y^{\leftarrow}(1-t))| = b_1 t^{(2\wedge(-\beta))/\alpha}(1+o(1))$$

for some $b_0,\ b_1 > 0$ *as* $t \to 0$, *and define*

$$b_2 t^{-\beta_2/\alpha} := \begin{cases} b_1 t^{(2\wedge(-\beta))/\alpha}, & \boldsymbol{\mu} \neq \mathbf{0}, \\ b_0 t^{-\beta/\alpha}, & \boldsymbol{\mu} = \mathbf{0}. \end{cases}$$

Then

$$\mathrm{amse}_{\mathrm{Hu}}(k_{\mathrm{Hu}}) = \sigma^2_{\mathrm{Hu}} k_{\mathrm{Hu}}^{-1} + (b_0(k_{\mathrm{Hu}}/n)^{-\beta/\alpha}\mathcal{B}_{(7)}(x,\ y))^2$$

and

$$\mathrm{amse}_{\mathrm{El}}(k_{\mathrm{El}}) = (\mathcal{B}_{(15)}(x,\ y))^2(\alpha^2 k_{\mathrm{El}}^{-1} + (\alpha^2 b_2 (k_{\mathrm{El}}/n)^{-\beta_2/\alpha}\mathcal{B}_{(14)})^2).$$

Let $k^{\mathrm{opt}}_{\mathrm{Hu}}$ *and* $k^{\mathrm{opt}}_{\mathrm{El}}$ *denote the optimal sample fraction in the sense of minimizing* $\mathrm{amse}_{\mathrm{Hu}}$ *and* $\mathrm{amse}_{\mathrm{El}}$, *respectively. Then*

$$k^{\mathrm{opt}}_{\mathrm{Hu}} = \left(\frac{-\alpha\sigma^2_{\mathrm{Hu}}}{2\beta b_0^2(\mathcal{B}_{(7)}(x,\ y))^2}\right)^{\alpha/(\alpha-2\beta)} n^{-2\beta/(\alpha-2\beta)},$$

$$k^{\mathrm{opt}}_{\mathrm{El}} = (-2\beta_2\alpha b_2^2(\mathcal{B}_{(14)})^2)^{-\alpha/(\alpha-2\beta_2)} n^{-2\beta_2/(\alpha-2\beta_2)},$$

$$\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}} := \mathrm{amse}_{\mathrm{Hu}}(k^{\mathrm{opt}}_{\mathrm{Hu}}) = n^{2\beta/(\alpha-2\beta)} b_0^{2\alpha/(\alpha-2\beta)}\left(1 - \frac{\alpha}{2\beta}\right) \times ((\sigma^2_{\mathrm{Hu}})^{-\beta/\alpha}\mathcal{B}_{(7)}(x,\ y)\sqrt{-2\beta/\alpha})^{2\alpha/(\alpha-2\beta)},$$

$$\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}} := \mathrm{amse}_{\mathrm{El}}(k^{\mathrm{opt}}_{\mathrm{El}}) = n^{2\beta_2/(\alpha-2\beta_2)} b_2^{2\alpha/(\alpha-2\beta_2)}\left(1 - \frac{\alpha}{2\beta_2}\right) \times \alpha^2(\mathcal{B}_{(15)}(x,\ y))^2(\sqrt{-2\alpha\beta_2}\mathcal{B}_{(14)})^{2\alpha/(\alpha-2\beta_2)}.$$

Remark 1. (i) Note that $k^{\mathrm{opt}}_{\mathrm{El}}$ is independent of x and y, but $k^{\mathrm{opt}}_{\mathrm{Hu}}$ depends on x and y. If $\boldsymbol{\mu} = \mathbf{0}$, both $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}$ and $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}$ depend on n, α, β, ρ, υ, x, y and b_0, $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}$ additionally depends on σ, but the ratio $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}/\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}$ is independent of n and b_0.

(ii) Since the optimal $k_{\mathrm{El}}^{\mathrm{opt}}$ is the same as that for Hill's estimator, when $\boldsymbol{\mu} = \mathbf{0}$, all data-driven methods for choosing the optimal sample fraction for Hill's estimator can be applied to $\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x, y)$ directly.

(iii) The location parameter $\boldsymbol{\mu}$ is the median of (X, Y) and the mean of (X, Y) when $\alpha > 1$. Hence, we could estimate $\boldsymbol{\mu}$ by the sample median, say $\hat{\boldsymbol{\mu}} = (\hat{\mu}_X, \hat{\mu}_Y)$. Therefore, consider the new estimator $\hat{\lambda}^{\mathrm{El}}_{k_{\mathrm{El}},n}(x, y)$ with $Z_i = \sqrt{X_i^2 + Y_i^2}$ replaced by $\sqrt{(X_i - \hat{\mu}_X)^2 + (Y_i - \hat{\mu}_Y)^2}$. Similar to the proofs in Ling and Peng (2004), we can show that Theorem 5 and Corollary 6 hold with $\boldsymbol{\mu} = \mathbf{0}$ for this new estimator. □

3. Comparisons and simulation study

First, we compare σ^2_{Hu} and ${\sigma_{\mathrm{El}}}^2$ given in Theorems 4 and 5. Note that both only depend on α, ρ, x and y. In Figure 1 we plot the ratio $\sigma^2_{\mathrm{El}}(\alpha)/\sigma^2_{\mathrm{Hu}}(\alpha)$ for $x = y = 1$ as a function of α, and each curve corresponds to a different correlation $\rho \in \{0.1, \ldots, 0.9\}$. From Figure 1, we conclude that $\hat{\lambda}^{\mathrm{El}}_{k,n}$ is always better in terms of asymptotic variance.

Second, we compare the two estimators in terms of optimal asymptotic mean squared errors. Since the ratio of the optimal asymptotic mean squared error depends on α, β, $\boldsymbol{\Sigma}$, $\boldsymbol{\mu}$, x, y, we consider elliptical distributions with $\sigma = \upsilon = 1$, $\mu_X = \mu_Y = 0$. In Figure 2 we

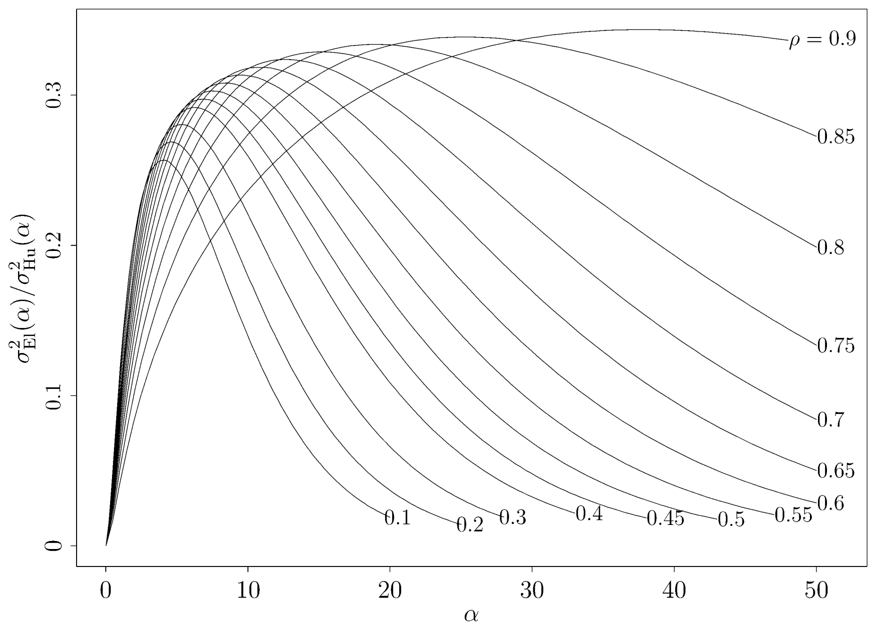

Figure 1. Ratio $\sigma^2_{\mathrm{El}}(\alpha)/\sigma^2_{\mathrm{Hu}}(\alpha)$ for different correlations ρ.

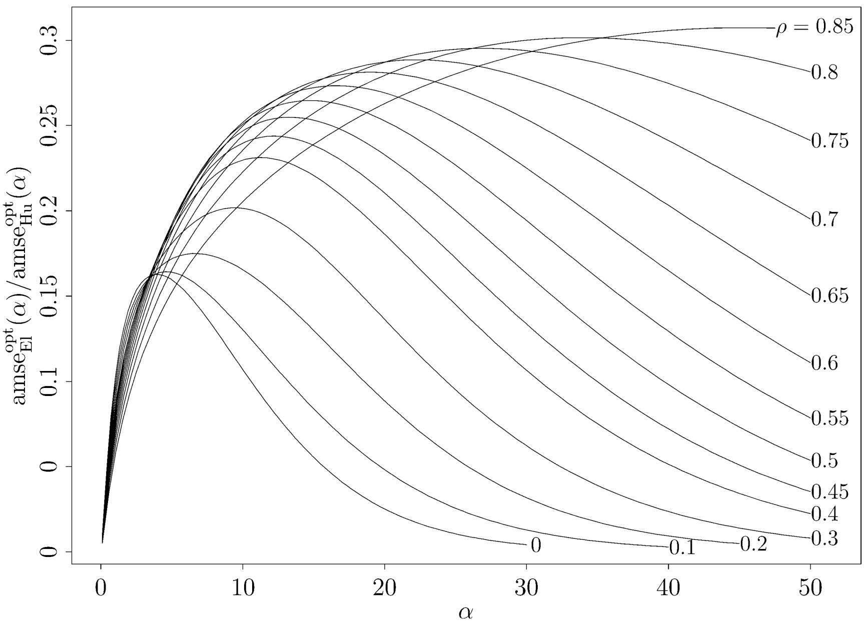

Figure 2. Ratio $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}(\alpha)/\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}(\alpha)$ for different correlations ρ and $\beta=-\alpha$.

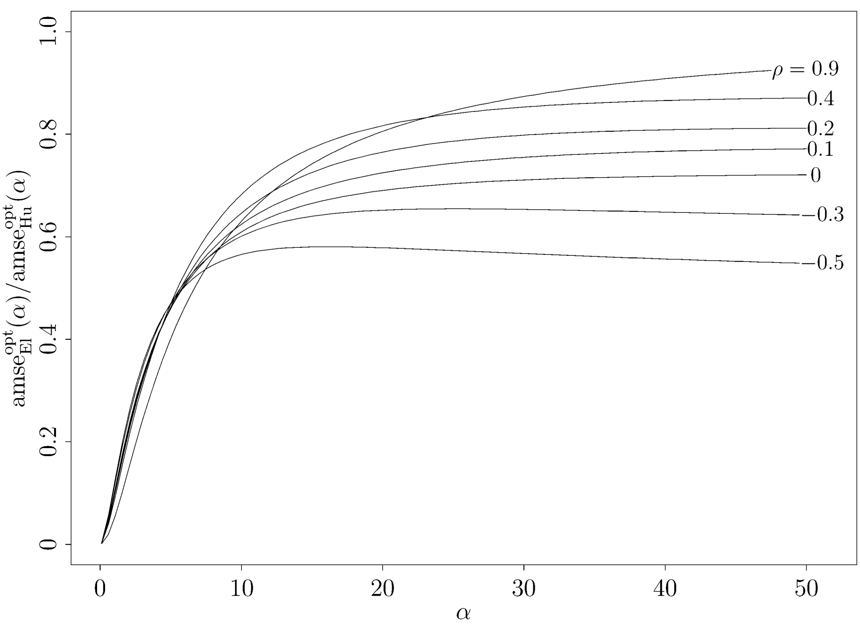

Figure 3. Ratio $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}(\alpha)/\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}(\alpha)$ for different correlations ρ and $\beta=-1$.

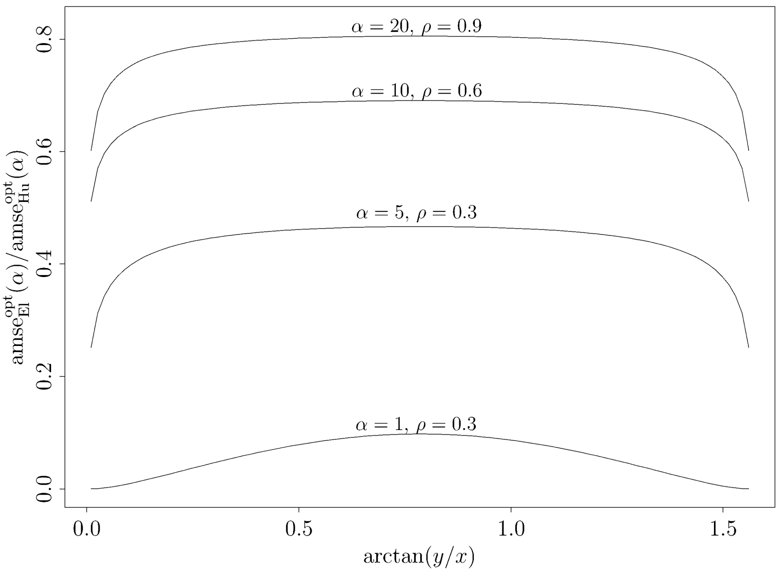

Figure 4. Ratio $\mathrm{amse}_{\mathrm{El}}^{\mathrm{opt}}(\alpha)/\mathrm{amse}_{\mathrm{Hu}}^{\mathrm{opt}}(\alpha)$, $\|(x, y)\| = \sqrt{2}$, for different (α, ρ) and $\beta = -1$.

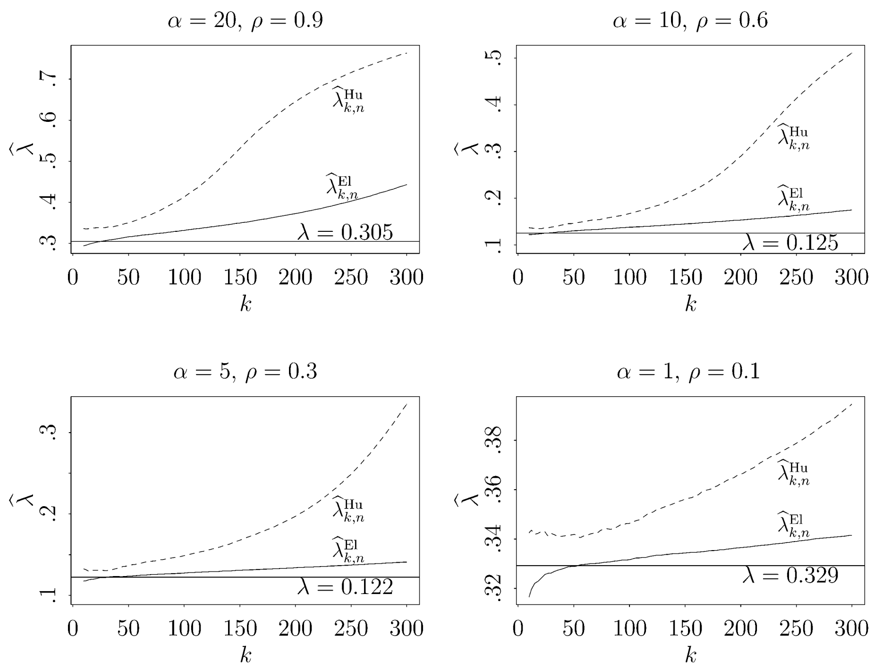

Figure 5. Mean of estimators $\hat{\lambda}_{k,n}^{\mathrm{Hu}}(1, 1)$ and $\hat{\lambda}_{k,n}^{\mathrm{El}}(1, 1)$ for 1000 samples of length $n = 1000$ and different k with $\sigma = \upsilon = 1$, $\boldsymbol{\mu} = \mathbf{0}$, $G \sim$ Fréhet(α), and different pairs (α, ρ).

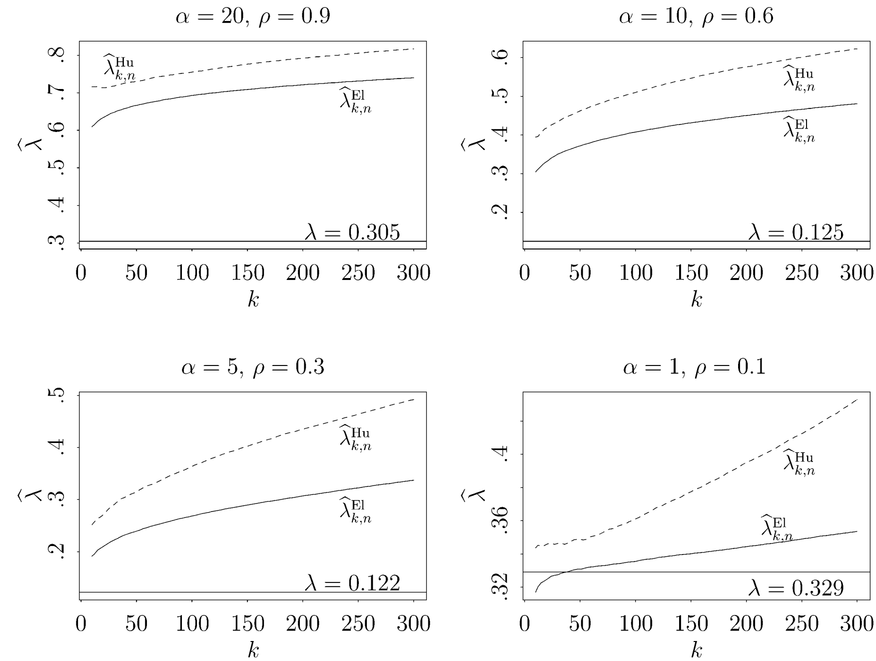

Figure 6. Mean of estimators $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$ and $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ for 1000 samples of length $n = 1000$ and different k with $\sigma = v = 1$, $\boldsymbol{\mu} = \mathbf{0}$, $G \sim$ Pareto(α), and different pairs (α, ρ).

consider $G \sim$ Fréchet(α), i.e. $\Pr(G > x) = \exp(-x^{-\alpha})$, $x > 0$; hence, (1) is satisfied with $\beta = -\alpha$. In Figure 3 we consider $G \sim$ Pareto(α), i.e. $\Pr(G > x) = (1 + x)^{-\alpha}$ for $x > 0$; hence (1) is satisfied with $\beta = -1$. Under the above set-up, the ratio of optimal asymptotic mean squared errors depends only on α, ρ, x, y. Similar to Figure 1, we plot the ratio $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}(\alpha)/\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}(\alpha)$ for $x = y = 1$ as a function of α for different ρs in Figures 2 and 3. We conclude from both figures that $\hat{\lambda}^{\mathrm{El}}_{k,n}$ always performs better than $\hat{\lambda}^{\mathrm{Hu}}_{k,n}$ in terms of optimal asymptotic mean squared errors as well.

Third, we examine the influence of x and y on the ratio of asymptotic mean squared error. We plot the ratio $\mathrm{amse}^{\mathrm{opt}}_{\mathrm{El}}(\alpha)/\mathrm{amse}^{\mathrm{opt}}_{\mathrm{Hu}}(\alpha)$ for $\|(x, y)\| = \sqrt{2}$ and $G \sim$ Pareto(α) in Figure 4, where each curve corresponds to a different pair $(\alpha, \rho) \in \{(20, 0.9),$ $(10, 0.6), (5, 0.3), (1, 0.1)\}$. This figure further confirms that $\hat{\lambda}^{\mathrm{El}}_{k,n}$ always has a smaller optimal asymptotic mean squared error than $\hat{\lambda}^{\mathrm{Hu}}_{k,n}$.

Fourth, we study the finite-sample behaviour of the two estimators $\hat{\lambda}^{\mathrm{El}}_{k,n}(x, y)$ and $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(x, y)$. As above, we consider two elliptical distributions with $\sigma = v = 1$, $\mu_X = \mu_Y = 0$, $G \sim$ Fréchet(α) in Figure 5 and $G \sim$ Pareto(α) in Figure 6. We generate 1000 random samples of size $n = 1000$ from these elliptical distributions with $(\alpha, \rho) \in$ $\{(20, 0.9), (10, 0.6), (5, 0.3), (1, 0.1)\}$, and plot $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ and $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$ against $k = 1,$ $5, 10, \ldots, 300$ for different pairs (α, ρ) where the solid line corresponds to $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ and the dashed line to $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$. This simulation study also confirms the better performance of $\hat{\lambda}^{\mathrm{El}}_{k,n}$.

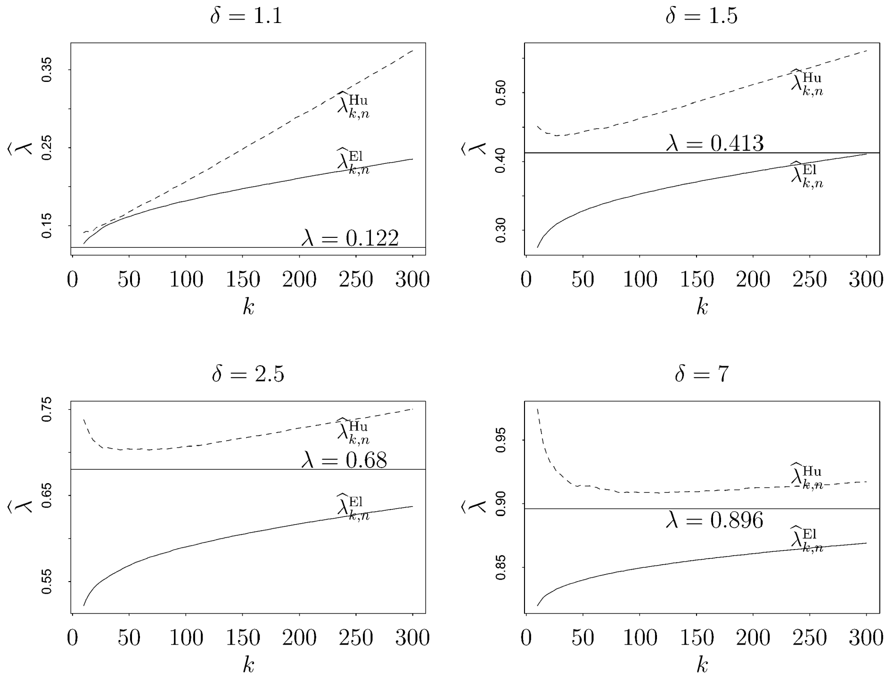

Figure 7. Mean of estimators $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$ and $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ for 1000 samples of length $n = 1000$ and different k with $F_{(X,Y)}(x, y) = C_\delta(F_X(x), F_Y(y))$, where F_X and F_Y are Pareto(5) and C_δ is the Gumbel copula with parameter δ.

Finally, we examine the robustness of the new estimator against the assumption of having an elliptical distribution. Consider the random vector (X, Y) with distribution $F_{(X,Y)}(x, y) = C_\delta(F_X(x), F_Y(y))$, where the marginal distributions F_X and F_Y are Pareto (5) and C_δ is the *Gumbel copula*, i.e. for $\delta \in [1, \infty)$,

$$C_\delta(u, v) := \exp(-((-\ln u)^\delta + (-\ln v)^\delta)^{1/\delta}), \qquad 0 < u, v < 1.$$

Then $F_{(X,Y)}$ is not elliptic. As before, we generate 1000 random samples of size $n = 1000$ with $\delta \in \{1.1, 1.5, 2.5, 7\}$. Note that $\lambda(1, 1) = 2 - 2^{1/\delta}$ (see Theorem 4.12 of Joe 1997). In Figure 7 we plot $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ and $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$ against $k = 1, 5, 10, \ldots, 300$ for the different δ, where the solid line corresponds to $\hat{\lambda}^{\mathrm{El}}_{k,n}(1, 1)$ and the dashed line to $\hat{\lambda}^{\mathrm{Hu}}_{k,n}(1, 1)$. We are not surprised to notice that $\hat{\lambda}^{\mathrm{El}}_{k,n}$ is sensitive to the assumptions of having an elliptical distribution.

4. Proofs

Proof of Theorem 1. Without loss of generality, we assume $\boldsymbol{\mu} = \mathbf{0}$. Let $\Phi \sim \mathrm{unif}(-\pi, \pi)$ be independent of G and let $F_i^{\leftarrow}(x)$ denote the generalized inverse of $F_i(x)$, $i = 1, 2$. Then, by Theorem 4.3 of Hult and Lindskog (2002) and its proof,

$$
\begin{aligned}
&F_X^{\leftarrow}(u) = \frac{\sigma}{v} F_Y^{\leftarrow}(u), && \text{for } 0 < u < 1, \\
&\lim_{t\to\infty}(1 - F_i(tx))/(1 - F_i(t)) = x^{-\alpha}, && \text{for } x > 0 \text{ and } i = 1, 2, . \qquad (16)\\
&(X, Y) \stackrel{\mathrm{d}}{=} (\sigma G \cos\Phi, vG \sin(\arcsin\rho + \Phi))
\end{aligned}
$$

Therefore, for $x, y > 0$

$$
\begin{aligned}
&t^{-1}\Pr(F_X(X) > 1 - tx, F_Y(Y) > 1 - ty) \\
&\quad = t^{-1}\Pr(G\cos\Phi > F_Y^{\leftarrow}(1 - tx)/v, G\sin(\arcsin\rho + \Phi) > F_Y^{\leftarrow}(1 - ty)/v) \\
&\quad = \frac{1}{2\pi t}\int_{-\arcsin\rho}^{\pi/2} \Pr\left(G \geqslant \frac{F_Y^{\leftarrow}(1 - tx)}{v\cos\phi} \vee \frac{F_Y^{\leftarrow}(1 - ty)}{v\sin(\arcsin\rho + \phi)}\right)\mathrm{d}\phi. \qquad (17)
\end{aligned}
$$

Note that

$$
\begin{aligned}
t &= \Pr(X > F_X^{\leftarrow}(1 - t)) = \Pr(G\cos\Phi > F_Y^{\leftarrow}(1 - t)/v) \\
&= \frac{1}{2\pi}\int_{-\pi/2}^{\pi/2} \Pr\left(G > \frac{F_Y^{\leftarrow}(1 - t)}{v\cos\phi}\right)\mathrm{d}\phi
\end{aligned}
$$

and that, for $\phi \in (-\pi/2, \pi/2)$, we have $1 \geqslant \Pr(G > x/\cos\phi)/\Pr(G > x) \stackrel{x\to\infty}{\longrightarrow} (\cos\phi)^\alpha$. Hence, we can apply the dominated convergence theorem and obtain

$$
\frac{1}{\mathcal{B}_{(18)}(t)} := \frac{1}{2\pi t}\Pr(G > F_Y^{\leftarrow}(1 - t)/v) \stackrel{t\to 0}{\longrightarrow} \left(\int_{-\pi/2}^{\pi/2} (\cos\phi)^\alpha\,\mathrm{d}\phi\right)^{-1} =: \frac{1}{\mathcal{B}_{(18)}}. \qquad (18)
$$

Next, we obtain, for $\phi \in (-\arcsin\rho, \pi/2)$,

$$
\frac{F_Y^{\leftarrow}(1 - tx)}{v\cos\phi} \geqslant \frac{F_Y^{\leftarrow}(1 - ty)}{v\sin(\arcsin\rho + \phi)} \Leftrightarrow \frac{F_Y^{\leftarrow}(1 - ty)}{F_Y^{\leftarrow}(1 - tx)} \leqslant \frac{\sin(\arcsin\rho + \phi)}{\cos\phi}.
$$

Note that $\sin(\arcsin\rho + \phi)/\cos\phi$ is strictly increasing, hence its inverse exists and equals $\arctan((\cdot - \rho)/\sqrt{1 - \rho^2})$. Therefore,

$$
\begin{aligned}
&\frac{F_Y^{\leftarrow}(1 - tx)}{v\cos\phi} \geqslant \frac{F_Y^{\leftarrow}(1 - ty)}{v\sin(\arcsin\rho + \phi)} \qquad (19)\\
&\quad \Leftrightarrow \phi \geqslant \arctan\left(\frac{F_Y^{\leftarrow}(1 - ty)/F_Y^{\leftarrow}(1 - tx) - \rho}{\sqrt{1 - \rho^2}}\right) =: g\left(\frac{F_Y^{\leftarrow}(1 - ty)}{F_Y^{\leftarrow}(1 - tx)}\right) =: h(x, y, t).
\end{aligned}
$$

Since $1 - F_Y \in \mathcal{R}_{-\alpha}$, by Proposition 1.7(9) of Geluk and de Haan (1987) we have $F_Y^{\leftarrow}(1 - tx)/F_Y^{\leftarrow}(1 - t) \stackrel{t\to 0}{\longrightarrow} x^{-1/\alpha}$, i.e.

$$
h(x, y, t) \stackrel{t\to 0}{\longrightarrow} g((x/y)^{1/\alpha}). \qquad (20)
$$

It follows from (17) and (19) that

$$t^{-1}\Pr(F_X(X) > 1 - tx,\ F_Y(Y) > 1 - ty)$$

$$= \frac{1}{\mathcal{B}_{(18)}(t)} \int_{h(x,y,t)}^{\pi/2} \frac{\Pr\Big(G > \frac{F_Y^{\leftarrow}(1-t)}{v\cos\phi} \frac{F_Y^{\leftarrow}(1-tx)}{F_Y^{\leftarrow}(1-t)}\Big)}{\Pr(G > F_Y^{\leftarrow}(1-t)/v)} \, \mathrm{d}\phi$$

$$+ \frac{1}{\mathcal{B}_{(18)}(t)} \int_{-\arcsin\rho}^{h(x,y,t)} \frac{\Pr\Big(G > \frac{F_Y^{\leftarrow}(1-t)}{v\sin(\arcsin\rho+\phi)} \frac{F_Y^{\leftarrow}(1-ty)}{F_Y^{\leftarrow}(1-t)}\Big)}{\Pr(G > F_Y^{\leftarrow}(1-t)/v)} \, \mathrm{d}\phi. \tag{21}$$

Hence, the theorem follows from (18), (20) and Potter's inequality; see (1.20) in Geluk and de Haan (1987). □

Proof of Theorem 2. Since $(X, Y) \overset{\mathrm{d}}{=} (\mu_X + \sigma G \cos\Phi,\ \mu_Y + vG\sin(\Phi + \arcsin\rho))$, we have $X^2 + Y^2 \overset{\mathrm{d}}{=} G^2 d_1(\Phi) + 2Gd_2(\Phi) + \mu_X^2 + \mu_Y^2$. Define, for $x > 0$ and $\theta \in (-\pi/2, \pi/2)$,

$$d_3(x, \phi) := \frac{1}{d_1(\phi)}\left(-d_2(\phi) + \sqrt{d_2^2(\phi) - d_1(\phi)(\mu_X^2 + \mu_Y^2 - x^2)}\right).$$

Since $\Pr(X^2 + Y^2 > t^2) = \Pr(G > d_3(t, \Phi))$ holds for large t, we obtain

$$\frac{\Pr(X^2 + Y^2 > t^2x^2)}{\Pr(X^2 + Y^2 > t^2)} - x^{-\alpha}$$

$$= \left(\int_{-\pi}^{\pi} \frac{\Pr(G > d_3(tx, \phi))}{\Pr(G > t)}\, \mathrm{d}\phi\right)\left(\int_{-\pi}^{\pi} \frac{\Pr(G > d_3(t, \phi))}{\Pr(G > t)}\, \mathrm{d}\phi\right)^{-1} - x^{-\alpha}$$

$$= \left\{\int_{-\pi}^{\pi} \left[\frac{\Pr(G > d_3(tx, \phi))}{\Pr(G > t)} - \left(\frac{1}{t} d_3(tx, \phi)\right)^{-\alpha}\right]\mathrm{d}\phi\right.$$

$$+ \int_{-\pi}^{\pi} \left[- x^{-\alpha}\, \frac{\Pr(G > d_3(t, \phi))}{\Pr(G > t)} + x^{-\alpha}\left(\frac{1}{t} d_3(t, \phi)\right)^{-\alpha}\right]\mathrm{d}\phi$$

$$\left.+ \int_{-\pi}^{\pi} \left[\left(\frac{1}{t} d_3(tx, \phi)\right)^{-\alpha} - x^{-\alpha}\left(\frac{1}{t} d_3(t, \phi)\right)^{-\alpha}\right]\mathrm{d}\phi\right\}\left(\int_{-\pi}^{\pi} \frac{\Pr(G > d_3(t, \phi))}{\Pr(G > t)}\, \mathrm{d}\phi\right)^{-1}$$

$$=: (\mathcal{J}_1(t) + \mathcal{J}_2(t) + \mathcal{J}_3(t))\left(\int_{-\pi}^{\pi} \frac{\Pr(G > d_3(t, \phi))}{\Pr(G > t)}\, \mathrm{d}\phi\right)^{-1}.$$

Since $|\rho| < 1$, it is straightforward to check that

$$\lim_{t\to\infty} t^{-1} d_3(t, \phi) = (d_1(\phi))^{-1/2}, \quad 0 < \sup_{-\pi\leqslant\phi\leqslant\pi} d_1(\phi) < \infty, \quad \text{and} \quad \sup_{-\pi\leqslant\phi\leqslant\pi} |d_2(\phi)| < \infty. \tag{22}$$

Hence, similarly to the proof of Theorem 1,

$$\lim_{t\to\infty} \int_{-\pi}^{\pi} \frac{\Pr(G > d_3(t, \phi))}{\Pr(G > t)}\, \mathrm{d}\phi = \int_{-\pi}^{\pi} (d_1(\phi))^{\alpha/2}\, \mathrm{d}\phi. \tag{23}$$

By Lemma 5.2 of Draisma *et al.* (1999), for every $\varepsilon > 0$, there exists $t_0 > 0$ such that for all $t \geqslant t_0$, $d_3(tx, \phi) \geqslant t_0$,

$$\left| \frac{\dfrac{\Pr(G > d_3(tx, \phi))}{\Pr(G > t)} - (t^{-1} d_3(tx, \phi))^{-\alpha}}{A(t)} - (t^{-1} d_3(tx, \phi))^{-\alpha} \frac{(t^{-1} d_3(tx, \phi))^{\beta} - 1}{\beta} \right|$$

$$\leqslant \varepsilon(1 + (t^{-1} d_3(tx, \phi))^{-\alpha} + (t^{-1} d_3(tx, \phi))^{-\alpha+\beta} \exp\{\varepsilon|\ln(t^{-1} d_3(tx, \phi))|\}). \qquad (24)$$

Using (22), for every fixed $x > 0$, we can choose t_0 large enough such that $d_3(tx, \phi) \geqslant t_0$ uniformly for $\phi \in [-\pi, \pi]$. That is, for every fixed $x > 0$, (24) holds uniformly for $\phi \in [-\pi, \pi]$. Therefore, by the dominated convergence theorem and (23), for $x > 0$,

$$\lim_{t\to\infty} \frac{\mathcal{J}_1(t)}{A(t)} = \frac{1}{x^{\alpha}\beta} \int_{-\pi}^{\pi} (x^{\beta}(d_1(\phi))^{(\alpha-\beta)/2} - (d_1(\phi))^{\alpha/2})\mathrm{d}\phi \qquad (25)$$

and

$$\lim_{t\to\infty} \frac{\mathcal{J}_2(t)}{A(t)} = -\frac{1}{x^{\alpha}\beta} \int_{-\pi}^{\pi} ((d_1(\phi))^{(\alpha-\beta)/2} - (d_1(\phi))^{\alpha/2})\mathrm{d}\phi. \qquad (26)$$

It follows from (22) and a Taylor expansion, for $x > 0$, that

$$\mathcal{J}_3(t) = t^{-1}\alpha x^{-\alpha}(x^{-1} - 1)\int_{-\pi}^{\pi} (d_1(\phi))^{\alpha/2-1/2} d_2(\phi)\mathrm{d}\phi + o(t^{-2}) \qquad (27)$$

$$+ t^{-2}\frac{\alpha}{2}x^{-\alpha}(x^{-2} - 1)\int_{-\pi}^{\pi} (d_1(\phi))^{\alpha/2-1}[\alpha(d_2(\phi))^2 + d_1(\phi)(\mu_X^2 + \mu_Y^2)]\mathrm{d}\phi.$$

Recall that $\sin(\phi + \arcsin\rho) = \rho\cos\phi + \sqrt{1-\rho^2}\sin\phi$. Then, splitting the integral into integrals over $[-\pi, -\pi/2)$, $[-\pi/2, 0)$, $[0, \pi/2)$, $[\pi/2, \pi]$ and using the symmetry of sine and cosine, we obtain

$$\int_{-\pi}^{\pi} (d_1(\phi))^{(\alpha-1)/2} d_2(\phi)\mathrm{d}\phi = 0. \qquad (28)$$

Hence (3) follows from (25)–(28). Note that

$$\lim_{t\to\infty} \frac{\Pr(\sqrt{X^2 + Y^2} > t)}{\Pr(G > t)} = \int_{-\pi}^{\pi} (d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi$$

and, since $Y \stackrel{\mathrm{d}}{=} \mu_Y + vG\sin\Phi$ with $\Phi \sim \mathrm{unif}(-\pi, \pi)$,

$$\lim_{t\to\infty} \frac{\Pr(Y > t)}{\Pr(G > t)} = v^{\alpha}\int_0^{\pi} (\sin\phi)^{\alpha}\,\mathrm{d}\phi.$$

Therefore, we have, as $t \to \infty$,

$$V(t) = \inf\left\{ y : \Pr(G > y)\int_{-\pi}^{\pi} (d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi \leqslant t^{-1} \right\}(1 + o(1))$$

and

$$F_Y^{\leftarrow}(1-t^{-1}) = \inf\left\{y : \Pr(G > y)v^\alpha \int_0^\pi (\sin\phi)^\alpha \,\mathrm{d}\phi \leqslant t^{-1}\right\}(1+o(1)).$$

Hence,

$$\lim_{t\to\infty} \frac{V(t)}{F_Y^{\leftarrow}(1-t^{-1})} = \left(\frac{\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi}{v^\alpha \int_0^\pi (\sin\phi)^\alpha\,\mathrm{d}\phi}\right)^{1/\alpha},$$

i.e. since $t^2|A(t)| \overset{t\to\infty}{\longrightarrow} \infty$ for $-2 < \beta \leqslant 0$,

$$\lim_{t\to\infty} \frac{(F_Y^{\leftarrow}(1-t^{-1}))^{-2} + |A(F_Y^{\leftarrow}(1-t^{-1}))|}{(V(t))^{-2} + |A(V(t))|} = \left(\frac{\int_{-\pi}^{\pi}(d_1(\phi))^{\alpha/2}\,\mathrm{d}\phi}{v^\alpha \int_0^\pi (\sin\phi)^\alpha\,\mathrm{d}\phi}\right)^{(2\wedge|\beta|)/\alpha}. \tag{29}$$

Note that, by Taylor expansion,

$$\left(\frac{V(tx)}{V(t)}\right)^{-\alpha} = \frac{1}{x} - \alpha x^{-1-1/\alpha}\left(\frac{V(tx)}{V(t)} - x^{1/\alpha}\right) + o\left(\left|\frac{V(tx)}{V(t)} - x^{1/\alpha}\right|\right). \tag{30}$$

Therefore, replacing t and x in (3) by $V(t)$ and $V(tx)/V(t)$, respectively, and using (29) and (30), we obtain (4). Let $\mu_X = \mu_Y = 0$; then $\mathcal{J}_3(t) = 0$ and we obtain (5). □

Proof of Theorem 3. For the purposes of this proof, we can assume $\mu_X = \mu_Y = 0$ since $\lambda(x, y)$ is independent of margins. We also set $v = 1$ and give the correction at the end of the proof. Using (16) we can also, equivalently, write $Y \overset{\mathrm{d}}{=} G\sin\Phi$, where $\Phi \sim \mathrm{unif}(-\pi, \pi)$ is independent of G. Then write

$$\begin{aligned}\frac{\Pr(Y > tx)}{\Pr(Y > t)} - x^{-\alpha} &= \frac{\int_0^\pi \Pr(G > tx/\sin\phi)\mathrm{d}\phi}{\int_0^\pi \Pr(G > t/\sin\phi)\mathrm{d}\phi} - x^{-\alpha} \\ &= \left(\int_0^\pi \frac{\Pr(G > t/\sin\phi)}{\Pr(G > t)}\,\mathrm{d}\phi\right)^{-1}\left\{\int_0^\pi \left[\frac{\Pr(G > tx/\sin\phi)}{\Pr(G > t)} - \left(\frac{x}{\sin\phi}\right)^{-\alpha}\right]\right. \\ &\qquad \left. -x^{-\alpha}\left[\frac{\Pr(G > t/\sin\phi)}{\Pr(G > t)} - \left(\frac{1}{\sin\phi}\right)^{-\alpha}\right]\mathrm{d}\phi\right\}.\end{aligned}$$

Then, by (1), we have, for $x > 0$,

$$\begin{aligned}&\lim_{t\to\infty} \frac{1}{A(t)}\left(\frac{\Pr(Y > tx)}{\Pr(Y > t)} - x^{-\alpha}\right) \\ &\quad = x^{-\alpha}\frac{x^\beta - 1}{\beta}\left(\int_0^\pi (\sin\phi)^\alpha\,\mathrm{d}\phi\right)^{-1}\left(\int_0^\pi (\sin\phi)^{\alpha-\beta}\mathrm{d}\phi\right).\end{aligned}$$

Replacing t and x in the latter equation by $F_Y^{\leftarrow}(1-s)$ and $F_Y^{\leftarrow}(1-sy)/F_Y^{\leftarrow}(1-s)$, respectively, we obtain, for $y > 0$,

$$\lim_{s\to 0}\frac{1}{A(F_Y^{\leftarrow}(1-s))}\left(\left(\frac{F_Y^{\leftarrow}(1-sy)}{F_Y^{\leftarrow}(1-s)}\right)^{-\alpha}-y\right)=\mathcal{B}_{(6)}(y). \tag{31}$$

Denote $f(t) := F_Y^{\leftarrow}(1-t)$. Then, by (21), we can write

$$\begin{aligned}
&t^{-1}\Pr(F_X(X) > 1-tx,\ F_Y(Y) > 1-ty)\\
&=\frac{1}{\mathcal{B}_{(18)}(t)}\left\{\int_{h(x,y,t)}^{\pi/2}\left[\frac{\Pr(G > \frac{f(t)}{\cos\phi}\frac{f(tx)}{f(t)})}{\Pr(G > f(t))}-\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha}\right]\mathrm{d}\phi\right.\\
&\quad+\int_{h(x,y,t)}^{\pi/2}\left[\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha}-x(\cos\phi)^{\alpha}\right]\mathrm{d}\phi+\int_{h(x,y,t)}^{h(x,y,0)}x(\cos\phi)^{\alpha}\,\mathrm{d}\phi\\
&\quad+\int_{-\arcsin\rho}^{h(x,y,t)}\left[\frac{\Pr(G > \frac{f(t)}{\sin(\arcsin\rho+\phi)}\frac{f(ty)}{f(t)})}{\Pr(G > f(t))}-\left(\frac{f(ty)}{f(t)\sin(\arcsin\rho+\phi)}\right)^{-\alpha}\right]\mathrm{d}\phi \qquad (32)\\
&\quad+\int_{-\arcsin\rho}^{h(x,y,t)}\left[\left(\frac{f(ty)}{f(t)\sin(\arcsin\rho+\phi)}\right)^{-\alpha}-y(\sin(\arcsin\rho+\phi))^{\alpha}\right]\mathrm{d}\phi\\
&\quad+\int_{h(x,y,0)}^{h(x,y,t)}y(\sin(\arcsin\rho+\phi))^{\alpha}\mathrm{d}\phi+\int_{h(x,y,0)}^{\pi/2}x(\cos\phi)^{\alpha}\,\mathrm{d}\phi\\
&\quad\left.+\int_{-\arcsin\rho}^{h(x,y,0)}y(\sin(\arcsin\rho+\phi))^{\alpha}\,\mathrm{d}\phi\right\}=:\frac{1}{\mathcal{B}_{(18)}(t)}\left(\sum_{i=1}^{6}\mathcal{J}_i(t)+\mathcal{J}_7+\mathcal{J}_8\right).
\end{aligned}$$

Note that $1/|\cos\phi| \geqslant 1$ for $\phi \in (-\pi/2, \pi/2)$ and v is given. Using Potter's inequality – see (1.20) in Geluk and de Haan (1987) and Lemma 5.2 of Draisma *et al.* (1999) – for every $\varepsilon > 0$, there exists some small $t_0 > 0$ such that for all $f(t) \geqslant f(t_0)$, $f(tx) \geqslant f(t_0)$ and $\phi \in (-\pi/2, \pi/2)$,

$$\begin{aligned}
&\left|\frac{\Pr\left(G > \frac{f(tx)}{\cos\phi}\right)/\Pr(G > f(t)) - \left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha}}{A(f(t))}-\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha}\frac{\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{\beta}-1}{\beta}\right|\\
&\leqslant\varepsilon\left(1+\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha}+\left(\frac{f(tx)}{f(t)\cos\phi}\right)^{-\alpha+\beta}\exp\left\{\varepsilon\left|\ln\frac{f(tx)}{f(t)\cos\phi}\right|\right\}\right),
\end{aligned} \tag{33}$$

and for all $t \leqslant t_0$ and $tx \leqslant t_0$,

$$(1-\varepsilon)x^{-1/\alpha}\exp\{-\varepsilon|\ln x|\} \leqslant \frac{f(tx)}{f(t)} \leqslant (1+\varepsilon)x^{-1/\alpha}\exp\{\varepsilon|\ln x|\}. \tag{34}$$

Since $f(t) \geqslant t_0$ and $t \leqslant t_0$ imply that $f(tx) \geqslant t_0$ and $tx \leqslant t_0$ for all $0 \leqslant x \leqslant 1$, respectively, by (33), (34), (20) and dominated convergence, we have

$$\lim_{t\to 0}\frac{\mathcal{J}_1(t)}{A(f(t))}=\frac{x}{\beta}\int_{h(x,y,0)}^{\pi/2}[x^{-\beta/\alpha}(\cos\phi)^{\alpha-\beta}-(\cos\phi)^{\alpha}]\mathrm{d}\phi \tag{35}$$

for all $x,\ y\geqslant 0$ and uniformly on $\mathcal{S}_2^+$. Similarly,

$$\lim_{t\to 0}\frac{\mathcal{J}_4(t)}{A(f(t))}=\frac{y}{\beta}\int_{-\arcsin\rho}^{h(x,y,0)}[y^{-\beta/\alpha}(\sin(\phi+\arcsin\rho))^{\alpha-\beta}-(\sin(\phi+\arcsin\rho))^{\alpha}]\mathrm{d}\phi \tag{36}$$

for all $x,\ y\geqslant 0$ and uniformly on $\mathcal{S}_2^+$.

Using (31) and Lemma 5.2 of Draisma *et al.* (1999), for every $\varepsilon>0$, there exists $t_0>0$ such that for all $t\leqslant t_0$ and $tx\leqslant t_0$,

$$\left|\frac{(F_Y^{\leftarrow}(1-tx)/F_Y^{\leftarrow}(1-t))^{-\alpha}-x}{A(F_Y^{\leftarrow}(1-s))}-\mathcal{B}_{(6)}(x)\right|\leqslant\varepsilon(C_1+C_2x+C_3x^{1-\beta/\alpha}\mathrm{e}^{\varepsilon|\ln x|}), \tag{37}$$

where the constants $C_1>0,\ C_2>0,\ C_3>0$ are independent of x and t. Hence, it follows from (20) and (37) that

$$\lim_{t\to 0}\frac{\mathcal{J}_2(t)}{A(F_Y^{\leftarrow}(1-t))}=\mathcal{B}_{(6)}(x)\int_{h(x,y,0)}^{\pi/2}(\cos\phi)^{\alpha}\,\mathrm{d}\phi \tag{38}$$

and

$$\lim_{t\to 0}\frac{\mathcal{J}_5(t)}{A(F_Y^{\leftarrow}(1-t))}=\mathcal{B}_{(6)}(y)\int_{-\arcsin\rho}^{h(x,y,0)}(\sin(\phi+\arcsin\rho))^{\alpha}\,\mathrm{d}\phi \tag{39}$$

for all $x,\ y\geqslant 0$ and uniformly on $\mathcal{S}_2^+$.

Note that

$$\begin{aligned}&\frac{1}{A(f(t))}\left(\left(\frac{f(ty)}{f(tx)}\right)^{-\alpha}-\frac{y}{x}\right)\\&=\frac{1}{A(f(t))}\left[\frac{1}{x}\left(\left(\frac{f(ty)}{f(t)}\right)^{-\alpha}-y\right)-\frac{1}{x}\left(\frac{f(ty)}{f(tx)}\right)^{-\alpha}\left(\left(\frac{f(tx)}{f(t)}\right)^{-\alpha}-x\right)\right]\\&\xrightarrow{t\to 0}\frac{1}{x}\mathcal{B}_{(6)}(y)-\frac{y}{x^2}\mathcal{B}_{(6)}(x).\end{aligned}$$

Apply Potter's bound and Lemma 5.2 of Draisma *et al.* (1999) to both

$$\frac{1}{A(f(t))}\frac{1}{x}\left(\left(\frac{f(ty)}{f(t)}\right)^{-\alpha}-y\right)-\frac{1}{x}\mathcal{B}_{(6)}(y)$$

and

$$\frac{1}{A(f(t))}\frac{1}{x}\left(\frac{f(ty)}{f(tx)}\right)^{-\alpha}\left(\left(\frac{f(tx)}{f(t)}\right)^{-\alpha}-x\right)-\frac{y}{x^2}\mathcal{B}_{(6)}(x).$$

For every $\varepsilon>0$, there exists $t_0>0$ such that for all $t\leqslant t_0,\ tx\leqslant t_0,\ ty\leqslant t_0$,

$$\begin{aligned}
&\left|\frac{1}{A(f(t))}\left(\left(\frac{f(ty)}{f(tx)}\right)^{-\alpha}-\frac{y}{x}\right)-\frac{1}{x}\mathcal{B}_{(6)}(y)+\frac{y}{x^2}\mathcal{B}_{(6)}(x)\right| \\
&\leqslant \frac{1}{x}\varepsilon(C_1+C_2y+C_3y^{1-\beta/\alpha}\mathrm{e}^{\varepsilon|\ln y|}) \\
&\quad +\frac{1}{x}\left(\frac{y}{x}\right)(C_1+C_2x+C_3x^{1-\beta/\alpha}\exp\{\varepsilon|\ln x|)\} \\
&\quad +\frac{1}{x}\left(\frac{y}{x}\right)\exp\{\varepsilon|\ln(y/x)|\}(C_1+C_2x+C_3x^{1-\beta/\alpha}\exp\{\varepsilon|\ln x|)\},
\end{aligned} \tag{40}$$

where constants $C_1>0$, $C_2>0$, $C_3>0$ are independent of t, x, y. Using (40),

$$\begin{aligned}
&\limsup_{z\to 0}|g'(z^{-1/\alpha})z^{2/\alpha}|<\infty, \\
&\limsup_{z\to\infty}|g'(z^{-1/\alpha})|<\infty, \\
&\limsup_{z\to\infty}[\sin(g(z^{-1/\alpha})+\arcsin\rho)]^\alpha z>\infty,
\end{aligned}$$

and, by a Taylor expansion of $g(z^{-1/\alpha})$, we can show that

$$\begin{aligned}
\lim_{t\to 0}\frac{\mathcal{J}_3(t)}{A(f(t))} &= \lim_{t\to 0}\frac{1}{A(f(t))}\int_{g(f(ty)/f(tx))}^{g((x/y)^{1/\alpha})} x(\cos\phi)^\alpha\,\mathrm{d}\phi \\
&= \frac{x}{\alpha}\left[\cos\left(g\left(\left(\frac{x}{y}\right)^{1/\alpha}\right)\right)\right]^\alpha g'\left(\left(\frac{x}{y}\right)^{1/\alpha}\right)\left(\frac{\mathcal{B}_{(6)}(y)}{y}-\frac{\mathcal{B}_{(6)}(x)}{x}\right)\left(\frac{x}{y}\right)^{1/\alpha}
\end{aligned} \tag{41}$$

and

$$\begin{aligned}
\lim_{t\to 0}\frac{\mathcal{J}_6(t)}{A(f(t))} &= \lim_{t\to 0}\frac{1}{A(f(t))}\int_{g((x/y)^{1/\alpha})}^{g(f(ty)/f(tx))} y(\sin(\phi+\arcsin\rho))^\alpha\,\mathrm{d}\phi\} \\
&= -\frac{y}{\alpha}\left[\sin\left(g\left(\left(\frac{x}{y}\right)^{1/\alpha}\right)+\arcsin\rho\right)\right]^\alpha g'\left(\left(\frac{x}{y}\right)^{1/\alpha}\right)\left(\frac{\mathcal{B}_{(6)}(y)}{y}-\frac{\mathcal{B}_{(6)}(x)}{x}\right)\left(\frac{x}{y}\right)^{1/\alpha}
\end{aligned} \tag{42}$$

for all x, $y\geqslant 0$ and uniformly on $\mathcal{S}_2^+$. Since

$$x[\cos(g((x/y)^{1/\alpha}))]^\alpha = y[\sin(g((x/y)^{1/\alpha})+\arcsin\rho)]^\alpha, \tag{43}$$

we obtain $\lim_{t\to 0}(\mathcal{J}_3(t)+\mathcal{J}_6(t))/A(f(t))=0$.

By Theorem 1, $\lambda(x,y)=(\mathcal{J}_7+\mathcal{J}_8)/\mathcal{B}_{(18)}$, hence

$$\lim_{t\to 0}\frac{1}{A(f(t))}\left(\frac{1}{\mathcal{B}_{(18)}(t)}(\mathcal{J}_7+\mathcal{J}_8)-\lambda(x, y)\right)$$
$$=\lim_{t\to 0}\frac{1}{A(f(t))}\left(-\frac{\lambda(x, y)}{\mathcal{B}_{(18)}(t)}(\mathcal{B}_{(18)}(t)-\mathcal{B}_{(18)})\right)$$
$$=-\lambda(x, y)\frac{1}{\beta}\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^{\alpha}((\cos\phi)^{-\beta}-1)\mathrm{d}\phi\right)\left(\int_{-\pi/2}^{\pi/2}(\cos\phi)^{\alpha}\,\mathrm{d}\phi\right)^{-1}, \qquad (44)$$

which obviously holds uniformly on $\mathcal{S}_2^+$ since $\sup_{\mathcal{S}_2^+}\lambda(x, y)<\infty$. Note that

$$\frac{A(F_Y^{\leftarrow}(1-t))}{A(F_{vY}^{\leftarrow}(1-t))}\xrightarrow{t\to 0}v^{-\beta}. \qquad (45)$$

Hence, the theorem follows from (35), (36), (38), (39), (41), (42), (44) and (45). □

Proof of Theorem 4. Similar to the proof of Theorem 2 in Chapter 2 of Huang (1992) with bias taken into account, we have, as $n\to\infty$,

$$\sup_{0\leqslant x,y\leqslant T}|\sqrt{k_{\mathrm{Hu}}}\{x+y-\hat{\lambda}^{\mathrm{Hu}}_{k_{\mathrm{H8}},n}(x, y)-l(x, y)\}-\mathcal{K}_{\mathrm{Hu}}\mathcal{B}_{(7)}(x, y)-B(x, y)|=o_p(1),$$

where

$$B(x, y)=W(x, y)-\left(1-\frac{\partial}{\partial x}\lambda(x, y)\right)W(x, 0)-\left(1-\frac{\partial}{\partial y}\lambda(x, y)\right)W(0, y),$$

and $W(x, y)$ is a Wiener process with zero mean and covariance structure

$$\begin{aligned}\mathrm{E}(W(x_1, y_1)W(x_2, y_2))=\ & l(x_1\wedge x_2, y_1)+l(x_1\wedge x_2, y_2)+l(x_1, y_1\wedge y_2)\\ &+l(x_2, y_1\wedge y_2)-l(x_1, y_2)-l(x_2, y_1)-l(x_1\wedge x_2, y_1\wedge y_2).\end{aligned}$$

Hence (9) follows from $\lambda(x, y)=x+y-l(x, y)$. It is straightforward to check that (10), (11) and (12) hold. Note that the result can also be obtained from Theorem 5 of Schmidt and Stadtmüller (2006) by taking the bias into account. □

Proof of Theorem 5. The result follows directly from

$$\sqrt{k_{\mathrm{El}}}(\hat{\alpha}^{\mathrm{H}}_{k_{\mathrm{El}},n}-\alpha)\xrightarrow{d}\mathcal{N}(-\alpha^2\mathcal{K}_{\mathrm{El}}\mathcal{B}_{(14)}, \alpha^2)$$

as $n\to\infty$ (see Theorem 1 of de Haan and Peng 1998), $\hat{\tau}_n-\tau=o_p(k_{\mathrm{El}}^{-1/2})$ and the delta method for the expression of $\lambda(x, y)$ given in Theorem 1. □

Acknowledgement

This research was supported by the Deutsche Forschungsgemeinschaft (SFB 386, 'Statistical Analysis of Discrete Structures'). Peng's research was also supported by NSF grant DMS-04-03443 and a Humboldt Research Fellowship.

References

Abdous, B., Fougeres, A.L. and Ghoudi, K. (2005) Extreme behaviour of bivariate elliptical distributions. *Canad. J. Statist.*, **33**(3), 317–334.

de Haan, L. and Peng, L. (1998) Comparison of tail index estimators. *Statist. Neerlandica*, **52**(1), 60–70.

de Haan, L. and Resnick, S. (1993) Estimating the limit distribution of multivariate extremes. *Comm. Statist. Stochastic Models*, **9**(2), 275–309.

de Haan, L. and Stadtmüller, U. (1996) Generalized regular variation of second order. *J. Australian Math. Soc. Ser. A*, **61**(3), 381–395.

Draisma, G., de Haan, L., Peng, L. and Pereira, T. (1999) A bootstrap-based method to achieve optimality in estimating the extreme-value index. *Extremes*, **2**(4), 367–404.

Drees, H. and Huang, X. (1998) Best attainable rates of convergence for estimates of the stable tail dependence functions. *J. Multivariate Anal.*, **64**(1), 25–47.

Einmahl, J., de Haan, L. and Huang, X. (1993) Estimating a multidimensional extreme value distribution. *J. Multivariate Anal.*, **47**(1), 35–47.

Einmahl, J., de Haan, L. and Sinha, A. (1997) Estimating the spectral measure of an extreme value distribution. *Stochastic Process. Appl.*, **70**(2), 143–171.

Einmahl, J., de Haan, L. and Piterbarg, V. (2001) Nonparametric estimation of the spectral measure of an extreme value distribution. *Ann. Statist.*, **29**(5), 1401–1423.

Einmahl, J., de Haan, L. and Li, D. (2006) Weighted approximations of tail copula processes with application to testing the multivariate extreme condition. *Ann. Statist.*, **34**(4).

Geluk, J. and de Haan, L. (1987) *Regular Variation, Extensions and Tauberian Theorems*. *CWI Tract* **40**.

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

Huang, X. (1992) *Statistics of bivariate extremes*. Doctoral thesis, Tinbergen Institute Research, Series 22, Erasmus University, Rotterdam.

Hult, H. and Lindskog, F. (2002) Multivariate extremes, aggregation and dependence in elliptical distributions. *Adv. Appl. Probab.*, **34**(3), 587–608.

Joe, H. (1997) *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.

Ledford, A. and Tawn, J. (1997) Modelling dependence within joint tail regions. Statistics for near independence in multivariate extreme values. *J. Royal Statist. Soc. Ser. B*, **59**(2), 475–499.

Ling, S. and Peng, L. (2004) Hill's estimator for the tail index of an ARMA model. *J. Statist. Plann. Inference*, **123**(2), 279–293.

McNeil, A.J., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton, NJ: Princeton University Press.

Resnick, S. (1987) *Extreme Values, Regular Variations and Point Processes*. New York: Springer-Verlag.

Schmidt, R. (2003) Credit risk modelling and estimation via elliptical copulae. In G. Bol, G.

Nakhaeizadeh, S.T. Rachev, T. Ridder and K.-H. Vollmer (eds), *Credit Risk: Measurement, Evaluation and Management*, pp. 267–289. Heidelberg: Physica-Verlag.

Schmidt, R. and Stadtmüller, U. (2006) Nonparametric estimation of tail dependence. *Scand. J. Statist.*, **33**, 307–335.

Tawn, J. (1988) Bivariate extreme value theory: models and estimation. *Biometrika*, **75**(3), 397–415.

Received March 2006 and revised August 2006