CentER

Discussion Paper

No. 2009–29

ESTIMATING EXTREME BIVARIATE QUANTILE REGIONS

By John H.J. Einmahl, Laurens de Haan

April 2009

ISSN 0924-7815

TILBURG UNIVERSITY

Estimating extreme bivariate quantile regions

John H.J. Einmahl
Tilburg University

Laurens de Haan*
Tilburg University, Lisbon University
Erasmus University, Rotterdam

April 21, 2009

Abstract. When simultaneously monitoring two possibly dependent, positive risks one is often interested in quantile regions with very small probability p. These extreme quantile regions contain hardly or no data and therefore statistical inference is difficult. In particular when we want to protect ourselves against a calamity that has not yet occurred, we take $p < 1/n$, with n the sample size. We consider quantile regions of the form $\{(x, y) \in (0, \infty)^2 : f(x, y) \leq \beta\}$, where f, the joint density, is decreasing in both coordinates. Such a region has the property that it consists of the less likely points and hence that its complement is as small as possible. Using extreme value theory, we construct a natural, semiparametric estimator of such a quantile region and prove a refined form of consistency. As an illustration, we compute the estimated quantile regions for simulated data sets.

Running title. Extreme bivariate quantiles.
JEL codes. C13, C14.
AMS 2000 subject classifications. Primary 62G32, 62G05; secondary 60G70, 60F05.
Key words and phrases. Density contour, extreme value, level set, multivariate quantile, rare event, semiparametric estimation, tail dependence.

---

*Partially supported by FCT, Project PTDC/MAT/64924/2006.

1 Introduction

Since there is no natural ordering of the two-dimensional Euclidean space, the concept of a quantile in $\mathbb{R}^2$ is not well defined. If, however, we can find a nice, natural class of nested regions, we are essentially back in the one-dimensional situation and a quantile can be defined to be an appropriate region in this class. Such a class of regions can be generated by the level sets of a function, in particular of the probability density function. When the density has some monotonicity property, these regions or their complements have desirable properties, like connectedness. E.g. for elliptical distributions - like the normal - the boundary of the quantile region is an ellipse. When a density on $(0, \infty)^2$ is monotone in both variables separately, similar quantile regions can be defined. In this paper, we shall consider estimation of such quantiles in the far tail, in a semiparametric setup, using multivariate extreme value theory.

Suppose we simultaneously monitor two possibly dependent, positive risks X and Y. Let the pair (X, Y) have df F with density f on $(0, \infty)^2$. Assume that - outside a square $(0, M]^2$ - the probability density is decreasing in each variable. Denote the probability measure corresponding to f with P. The probability measure on the underlying probability space will be denoted by $\mathbb{P}$. We define quantile regions determined by the levels of f:

$$Q = \{(x, y) \in (0, \infty)^2 : f(x, y) \leq \beta\}.$$

So, for a (small) $p \in (0, 1)$ we try to find a Q of this form such that $PQ = p$. The region $Q^c = \{(x, y) \in (0, \infty)^2 : f(x, y) > \beta\}$ has the property that everywhere on Q^c, f is larger than everywhere on Q, i.e. the quantile region Q is the set of less likely points. As a consequence, Q^c is the region with smallest area such that $PQ^c = 1 - p$.

Now suppose that we have a random sample $(X_1, Y_1), \ldots, (X_n, Y_n)$ from F. Let $p = p_n$ be very small; for the asymptotics think of $np \to c \in [0, \infty)$, so $c = 0$ is possible. In particular when we want to protect ourselves against a calamity that has not yet occurred, we consider the case where $p < 1/n$. The question comes up how to estimate $Q = Q_n$. Such a quantile region Q contains hardly or no data and therefore the estimation is statistically difficult. It is the aim of this paper to propose an estimation procedure for these quantiles Q – connected with a very low probability and in the right upper tail – in the framework of extreme value theory (EVT). The fact that under the EVT condition the tail of a distribution is close to a multivariate generalized Pareto distribution (cf. for example Rootzén and Tajvidi, 2006) of rather simple structure is helpful.

Our results can be applied in, e.g., aviation safety. The Federal Aviation Administration (FAA) needs a system that provides instant assessments of airline performances and that in particular signals those that appear to be extreme. Available is a data set for two possibly dependent key airline performance measures (Incident Rate and Operational Unfavorable Ratio). The bivariate data are positive and higher values correspond to a worse performance. The task now is to identify an extreme risk region desired by the FAA. Our estimator of Q – for very small p – is a very natural extreme risk region and hence could be used for flagging events of extreme aviation risk. See for more details Einmahl, Li and Liu (2009).

The paper is organized as follows. In Section 2 we derive our estimator and present the main asymptotic result. The method is illustrated on simulated data in Section 3 and the proofs are deferred to Section 4.

2 Main Results

We assume throughout that F is in the max-domain of attraction of an extreme-value distribution function G, with *positive* extreme-value indices γ_1, γ_2. The marginal distribution functions of F are denoted by F_1, F_2, respectively. In this case the domain of attraction condition can be written as

$$t\,(1 - F(U_1(t)x^{\gamma_1}, U_2(t)y^{\gamma_2})) \to -\log G(x^{\gamma_1}, y^{\gamma_2}), \quad t \to \infty, \tag{1}$$

on $(0,\infty]^2 \setminus \{(\infty,\infty)\}$, with

$$U_j(t) = F_j^{-1}(1 - 1/t), \quad j = 1, 2;$$

here $-\log G(x^{\gamma_1}, \infty) = -\log G(\infty, x^{\gamma_2}) = 1/x$. This implies the existence of a measure ν, the exponent measure, such that for all Borel sets $A \subset [0,\infty]^2$ that are bounded away from the origin and satisfy $\nu(\partial A) = 0$

$$tP\left(\{(U_1(t)x^{\gamma_1}, U_2(t)y^{\gamma_2}) : (x,y) \in A\}\right) \to \nu(A), \quad t \to \infty. \tag{2}$$

For such an A and $a > 0$, we have $\nu(aA) = \nu(A)/a$. Also, for every $a > 0$ the exponent measure is a finite measure on $[0, \infty]^2 \setminus [0, a]^2$. For more details, see de Haan and Resnick (1977).

We also require the convergence in (1) at the density level. Let g be the density corresponding to the right-hand side of (1), i.e. we have

$$-\log G(x^{\gamma_1}, y^{\gamma_2}) = \nu(\{(u, v) : u > x \text{ or } v > y\}) = \iint_{u>x \text{ or } v>y} g(u, v)dudv.$$

We assume

$$\begin{aligned} q_t(x, y) &:= tU_1(t)U_2(t)f\left(U_1(t)x^{\gamma_1}, U_2(t)y^{\gamma_2}\right) \\ &\to \tfrac{1}{\gamma_1\gamma_2}x^{1-\gamma_1}y^{1-\gamma_2}g(x, y) =: q(x, y), \quad t \to \infty, \end{aligned} \tag{3}$$

on $(0, \infty)^2$. In addition, we assume that

$$\begin{aligned} &f \text{ is decreasing in each coordinate, outside } (0, M]^2 \text{ (for some } M > 0\text{)}, \\ &\text{and that on } (0, M]^2, f \text{ is bounded away from zero.} \end{aligned} \tag{4}$$

Set $A_\lambda = \{(x, y) \in (0, \infty)^2 : x \wedge y \geq \lambda\}$. It follows that the convergence in (3) is uniform on A_λ, for every $\lambda > 0$, since the monotone functions q_t converge pointwise to q on A_λ and the range of values of q on A_λ is bounded.

We mention a few properties of g, q and f. They follow easily from multivariate extreme-value theory and standard arguments. We have

$$g(ax, ay) = a^{-3}g(x, y), \quad x, y > 0, a > 0,$$

i.e. g is homogeneous of degree -3. Also, q is decreasing in each coordinate on $(0, \infty)^2$, g and q are continuous on $(0, \infty)^2$, g and f are positive on $(0, \infty)^2$, and $g(cx, x)$ is decreasing in x, for all $c > 0$.

Condition (1) implies the existence of the spectral measure: a finite measure Ψ on $[0, \pi/2]$ such that

$$-\log G(x^{\gamma_1}, y^{\gamma_2}) = \int_0^{\pi/2} \frac{\cos\theta}{x} \vee \frac{\sin\theta}{y} d\Psi(\theta),$$

with $\int_0^{\pi/2} \cos\theta d\Psi(\theta) = \int_0^{\pi/2} \sin\theta d\Psi(\theta) = 1$; see Corollary 2 in de Haan and Resnick (1977). We have

$$\nu\left(\left\{(u, v) : u^2 + v^2 > r^2, \frac{v}{u} \leq \tan\theta\right\}\right) = \frac{1}{r}\Psi(\theta).$$

(Without confusion, we use Ψ to denote both the spectral measure and its distribution function.) The existence of g implies the existence of $\psi := \Psi'$ on $(0, \pi/2)$ and

$$\psi(\theta) = g(\cos\theta, \sin\theta).$$

It follows that ψ is continuous on $(0, \pi/2)$ and that $\Psi(\{0\}) = \Psi(\{\pi/2\}) = 0$.

Recall

$$Q_n = \{(x, y) \in (0, \infty)^2 : f(x, y) \leq \beta\}$$

where β is taken such that $PQ_n = p$, with $p = p_n$ such that $np \to c \in [0, \infty)$, or slightly weaker $p = O(1/n)$. It is the aim of the paper to estimate Q_n, more precisely we want to show consistency of our estimator $\widehat{Q}_n$ in the following appropriate sense:

$$\frac{P(\widehat{Q}_n \triangle Q_n)}{p} \xrightarrow{\mathbb{P}} 0.$$

(Here $\triangle$ denotes 'symmetric difference': $A \triangle B = A \setminus B \cup B \setminus A$.)

Set

$$S = \{(x, y) : x^{1-\gamma_1} y^{1-\gamma_2} g(x, y) \leq \gamma_1 \gamma_2\},$$

see (3). S is a fixed (i.e. not depending on n) 'basis' for our estimator of Q_n. We will estimate S later and then transform - using in particular p - that estimator into an estimator of Q_n. Throughout

(5) $\qquad k = k_n$ is a sequence of positive integers such that $k \to \infty$ and $k/n \to 0$.

A first step is to find an approximate value of β. In this way, Q_n is approximated by $\bar{Q}_n$ given by

$$\bar{Q}_n = \left\{ (x, y) \in (0, \infty)^2 : f(x, y) \leq \left(\frac{np}{k\nu(S)} \right)^{\gamma_1 + \gamma_2 + 1} \frac{1}{(n/k) U_1(n/k) U_2(n/k)} \right\}.$$

Next, using (3), approximate $\bar{Q}_n$ by a similar expression involving g, the density of the limiting measure, rather than f. Let $z = (x, y)$ and define, in vector notation, the map T_t, $t > 1$, by

$$T_t(z) = U(t) z^\gamma = (U_1(t) x^{\gamma_1}, U_2(t) y^{\gamma_2}),$$

hence (2) can be written as

$$tP(T_t(z) : z \in A) \to \nu(A), \quad t \to \infty. \tag{6}$$

Set

$$\widetilde{Q}_n = T_{n/k}\left(\frac{k\nu(S)}{np}S\right) = U\left(\frac{n}{k}\right)\left(\frac{k\nu(S)}{np}\right)^{\gamma} S^{\gamma}.$$

The obvious step to obtain an estimator of Q_n is now to estimate $\widetilde{Q}_n$, which can be done by estimating $T_{n/k}, \nu(S)$, and in particular S. It is convenient to write S in polar coordinates ($r = \sqrt{x^2+y^2}$, $\theta = \arctan(y/x)$):

$$S = \left\{(x,y) : r \geq \left(\frac{1}{\gamma_1\gamma_2}\psi(\theta)\cos^{1-\gamma_1}\theta \sin^{1-\gamma_2}\theta\right)^{\frac{1}{\gamma_1+\gamma_2+1}},\ \theta \in [0,\pi/2]\right\}. \tag{7}$$

Note that

$$\begin{aligned} \nu(S) &= \iint_{r \geq \left(\frac{1}{\gamma_1\gamma_2}\psi(\theta)\cos^{1-\gamma_1}\theta\sin^{1-\gamma_2}\theta\right)^{\frac{1}{\gamma_1+\gamma_2+1}}} \frac{1}{r^2}\psi(\theta)drd\theta \\ &= \int_0^{\pi/2}\left(\frac{1}{\gamma_1\gamma_2}\psi(\theta)\cos^{1-\gamma_1}\theta\sin^{1-\gamma_2}\theta\right)^{-\frac{1}{\gamma_1+\gamma_2+1}}\psi(\theta)d\theta. \end{aligned} \tag{8}$$

In order to estimate $T_{n/k}, \nu(S)$, and S, it is sufficient to estimate $U_1(n/k), U_2(n/k), \gamma_1, \gamma_2$, and the spectral density ψ. Estimation of the first four is well-known. We estimate $U_1(n/k)$ and $U_2(n/k)$ with the corresponding order statistics $X_{n-k:n}$ and $Y_{n-k:n}$, respectively, so

$$\widehat{U}(n/k) = (\widehat{U}_1(n/k), \widehat{U}_2(n/k)) = (X_{n-k:n}, Y_{n-k:n}).$$

For the two extreme-value indices any $\sqrt{k}$-consistent estimator can be chosen, e.g. the moment estimator in Dekkers, Einmahl and de Haan (1989). The estimator for ψ will be obtained by smoothing a $\sqrt{k}$-consistent estimator $\widehat{\Psi}$ of the spectral measure Ψ, in particular we can choose the maximum empirical likelihood estimator of Ψ in Einmahl and Segers (2009). To be more precise let K be a probability density function being 0 outside $(-1,1)$, symmetric around 0, and of bounded variation. Define

$$\widehat{\psi}(\theta) = \int_{\theta-h}^{\theta+h}\frac{1}{h}K\left(\frac{t-\theta}{h}\right)d\widehat{\Psi}(t), \quad h = h_n > 0.$$

Combining the various estimators we obtain, in vector notation, the following novel estimator of an extreme bivariate quantile region:

$$\widehat{Q}_n = \widehat{U}\left(\frac{n}{k}\right)\left(\frac{k\widehat{\nu(S)}}{np}\right)^{\widehat{\gamma}}\widehat{S}^{\widehat{\gamma}},$$

with (see (7))

$$\widehat{S} = \left\{ (x,y) : r \geq \left(\frac{1}{\widehat{\gamma}_1 \widehat{\gamma}_2} \widehat{\psi}(\theta) \cos^{1-\widehat{\gamma}_1}(\theta \wedge (\pi/2 - h)) \sin^{1-\widehat{\gamma}_2}(\theta \vee h) \right)^{\frac{1}{\widehat{\gamma}_1 + \widehat{\gamma}_2 + 1}} \right\}$$

and (see (8))

$$\widehat{\nu(S)} = \int_0^{\pi/2} \left(\frac{1}{\widehat{\gamma}_1 \widehat{\gamma}_2} \widehat{\psi}(\theta) \cos^{1-\widehat{\gamma}_1}(\theta \wedge (\pi/2 - h)) \sin^{1-\widehat{\gamma}_2}(\theta \vee h) \right)^{-\frac{1}{\widehat{\gamma}_1 + \widehat{\gamma}_2 + 1}} \widehat{\psi}(\theta) d\theta.$$

We are now in a position to present our main result. We need the following marginal second order conditions: for $j = 1, 2$, there exist functions A_j with $\lim_{t\to\infty} A_j(t) = 0$ and constant sign near infinity, such that

$$\lim_{t\to\infty} \frac{\frac{U_j(tx)}{U_j(t)} - x^{\gamma_j}}{A_j(t)} = x^{\gamma_j} \frac{x^{\rho_j} - 1}{\rho_j} \tag{9}$$

for all $x > 0$ and some $\rho_j < 0$.

Theorem *Let $p = O(1/n)$. Assume (1), (3), (4), (5), (9) hold and that $\widehat{\gamma}_1, \widehat{\gamma}_2$ and $\widehat{\Psi}$ are such that $\sqrt{k}(\widehat{\gamma}_j - \gamma_j) = O_p(1), j = 1, 2$, and $\sqrt{k}(\widehat{\Psi} - \Psi)$ converges in distribution on $D[0, \pi/2]$ to a continuous process. Also assume $\inf_{\theta\in(0,\pi/2)} \psi(\theta) > 0$, $lim_{n\to\infty} h = 0$, $\liminf_{n\to\infty} h\sqrt{k} > 0$, and $\lim_{n\to\infty} (\log np)/\sqrt{k} = 0$. Then we have that, as $n \to \infty$,*

$$\frac{P(\widehat{Q}_n \triangle Q_n)}{p} \xrightarrow{\mathbb{P}} 0. \tag{10}$$

Remark 1 The consistency formulation in a ratio setting is appropriate here. Since $p = O(1/n)$, the statement $P(\widehat{Q}_n \triangle Q_n) \xrightarrow{\mathbb{P}} 0$ is pointless: it even holds when taking $\widehat{Q}_n$ the empty set. Actually our result is rather strong, stating that the estimation error is much smaller than the already extremely small p. Observe that it follows from the Theorem that

$$\frac{P(\widehat{Q}_n)}{p} \xrightarrow{\mathbb{P}} 1.$$

Remark 2 In practice it is important that the tuning parameters k used in the estimation of the marginal quantities (γ_j and $U_j, j = 1, 2$) and in the estimation of ψ can be chosen to be different, i.e. we take k_1, k_2 and k_ψ. (E.g., a good value for k_1 can be a bad value for k_2; this depends on ρ_1 and ρ_2 in (9).) The thus adapted estimator becomes

$$\widehat{Q}_n = \left\{ \left(\widehat{U}_1\left(\frac{n}{k_1}\right)\left(\frac{k_1\widehat{\nu(S)}}{np}\right)^{\widehat{\gamma}_1} x^{\widehat{\gamma}_1}, \widehat{U}_2\left(\frac{n}{k_2}\right)\left(\frac{k_2\widehat{\nu(S)}}{np}\right)^{\widehat{\gamma}_2} y^{\widehat{\gamma}_2} \right) : (x, y) \in \widehat{S} \right\}.$$

If we also adapt the conditions of the theorem, in particular if (5) holds for k_1, k_2, k_ψ, $\liminf_{n\to\infty} h\sqrt{k_\psi} > 0$, and $\lim_{n\to\infty}(\log np)/\sqrt{k_j} = 0$, $j = 1, 2$, then (10) remains true.

Remark 3 Note that the estimated quantile region $\widehat{Q}_n$ depends on p in a monotone way: if $p < p'$ then $\widehat{Q}_n(p) \subset \widehat{Q}_n(p')$. It is also a continuous function of p. Hence, starting from a very small $\widehat{Q}_n$ we can enlarge it until it first hits an observation. This observation can then be considered the largest one and it has a "p-value" attached to it. This could be helpful in deciding whether some two-dimensional observation is the most atypical (or: an outlier); see Section 3. Also, by continuing this procedure we can introduce a ranking of the larger observations.

Remark 4 Considering the related paper Einmahl, Li and Liu (2009), the main difference is that in that paper the shape of the quantile region estimator is fixed beforehand to be a quadrant, whereas here the data determine the shape of the quantile region estimator. Also somewhat related are de Haan and Huang (1995) and Joe, Smith and Weissman (1992).

3 Illustration

In this section we illustrate the method on two simulated data sets. We use the adapted estimator of Remark 2.

Consider the bivariate Cauchy distribution on $(0, \infty)^2$ with density

$$f(x, y) = \frac{2}{\pi(1 + x^2 + y^2)^{3/2}}.$$

This is a heavy-tailed density, symmetric in the coordinates x and y and a function of the radius r. We have $\gamma_1 = \gamma_2 = 1$ and $\psi(\theta) = 1$, for $\theta \in (0, \pi/2)$. We simulated a single data set of size 5000 from this distribution and computed the true and estimated quantile regions corresponding to $p = 1/2000, 1/5000$, and 1/10,000, respectively. Observe that for

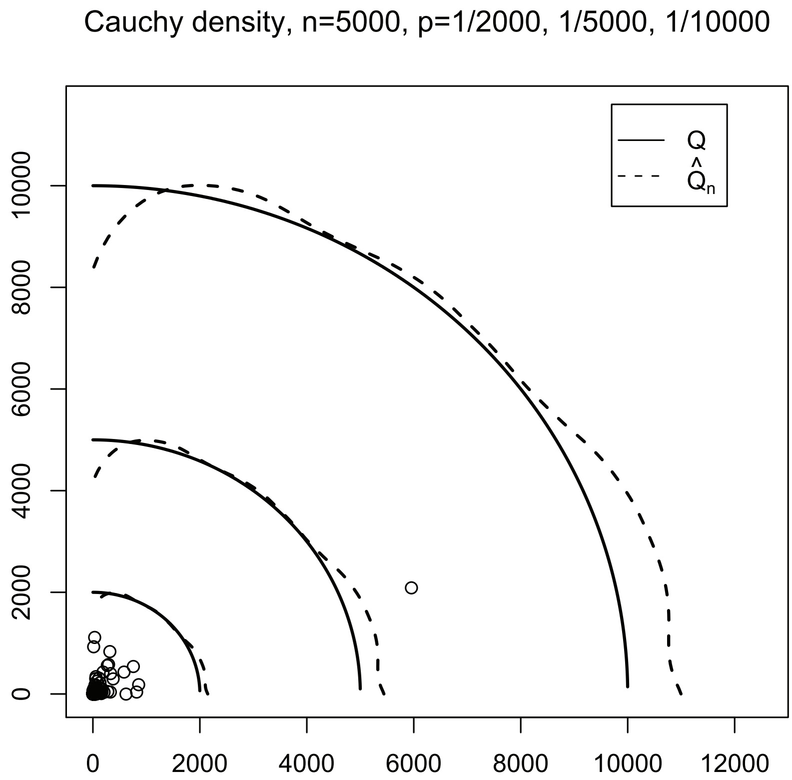

Figure 1: True and estimated quantile regions for $p = 1/2000, 1/5000, 1/10{,}000$ based on a sample of size 5000 from the bivariate Cauchy distribution.

the latter p, np is as small as 0.5. It should be noted that for, e.g., $p = 1/10{,}000$ the (constant) density f on the boundary of the quantile region is less than 10^{-12}. Figure 1 shows excellent behavior of our procedure, in particular considering that the regions ar far away from almost all the data points. We also calculated $P(\widehat{Q}_n)$ for the three p-s. These values only deviate a few percent from p which is a very small error given that the p-s are extremely small.

We will also consider the density on $(0, \infty)^2$ given by

$$f(x, y) = \frac{c}{x^3 + y^4 + 1}, \tag{11}$$

with $c \approx 0.581$. This density is less heavy tailed: $\gamma_1 = 4/5$ and $\gamma_2 = 3/5$. We find

$$\psi(\theta) = \frac{12c}{25} \frac{c_1 c_2}{\left(c_1^3 \cos^{12/5}\theta + c_2^4 \sin^{12/5}\theta\right) \cos^{1/5}\theta \sin^{2/5}\theta}, \quad \theta \in (0, \pi/2),$$

with $c_1 \approx 0.589$ and $c_2 \approx 0.593$. One can generate data from the density f by noting that

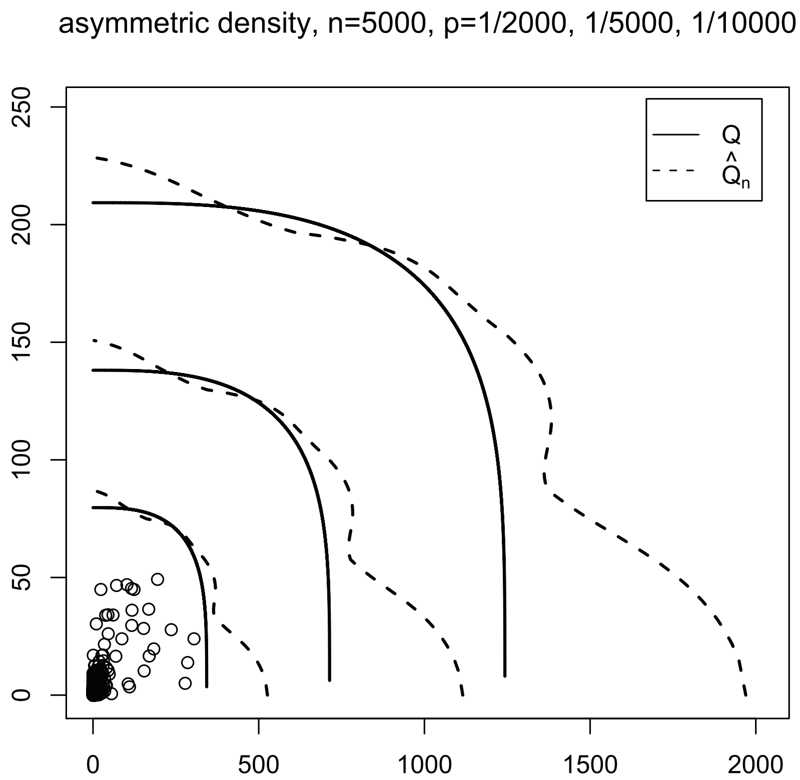

Figure 2: True and estimated quantile regions for $p = 1/2000, 1/5000, 1/10{,}000$ based on a sample of size 5000 from the density f in (11). Note that different scales are used on the two axes.

a corresponding random vector (X, Y) can be represented by two *independent* random variables S and Ξ as follows:

$$(X, Y) = \left(\left(\frac{S}{\Xi + 1} \right)^{1/3}, \left(\frac{S\,\Xi}{\Xi + 1} \right)^{1/4} \right).$$

We simulated a sample of size 5000 from this distribution and computed true and estimated quantile regions corresponding to $p = 1/2000, 1/5000$, and 1/10,000, see Figure 2. For these data the procedure shows the same excellent behavior. The three values of $P(\widehat{Q}_n)$ are now 10–15% too low, a small error given the statistical difficulty of the estimation problem.

4 Proofs

For the proof of the theorem we need several lemmas and propositions. We throughout assume that the conditions of the theorem are in force. We will need the following simple auxiliary result.

Lemma 1 For all Borel sets $A \subset [0,\infty]^2$ that are bounded away from the origin and satisfy $\nu(\partial A) = 0$, we have

$$\frac{\nu(S)}{p} P\left(T_{n/k}\left(\frac{k\nu(S)}{np} A\right)\right) \to \nu(A), \quad n \to \infty. \tag{12}$$

Proof It suffices to prove (12) for sets A of the form $([0,x]\times[0,y])^c$, with $x, y \geq 0, x+y > 0$. Then the left-hand side of (12) is equal to

$$\begin{aligned}
& \frac{\nu(S)}{p}\mathbb{P}\left(X > U_1\left(\frac{n}{k}\right)\left(\frac{k\nu(S)}{np}x\right)^{\gamma_1} \text{ or } Y > U_2\left(\frac{n}{k}\right)\left(\frac{k\nu(S)}{np}y\right)^{\gamma_2}\right) \\
= \; & \frac{\nu(S)}{p}\mathbb{P}\left(X > U_1\left(\frac{\nu(S)}{p}\right)\left[\left(\frac{U_1\left(\frac{n}{k}\right)}{U_1\left(\frac{\nu(S)}{p}\right)}\right)^{1/\gamma_1}\frac{k\nu(S)}{np}x\right]^{\gamma_1}\right. \\
& \left. \text{or } Y > U_2\left(\frac{\nu(S)}{p}\right)\left[\left(\frac{U_2\left(\frac{n}{k}\right)}{U_2\left(\frac{\nu(S)}{p}\right)}\right)^{1/\gamma_2}\frac{k\nu(S)}{np}y\right]^{\gamma_2}\right).
\end{aligned}$$

Using $k\nu(S)/(np) \to \infty$, it follows from the second order conditions (9) and Remark B.3.15 in de Haan and Ferreira (2006, p. 397), that

$$\frac{U_j\left(\frac{n}{k}\frac{k\nu(S)}{np}\right)}{U_j\left(\frac{n}{k}\right)\left(\frac{k\nu(S)}{np}\right)^{\gamma_j}} \to 1, \quad j = 1, 2.$$

Hence by the local uniformity in the limit relation

$$\frac{\nu(S)}{p} P\left(T_{\nu(S)/p}([0,x]\times[0,y])^c\right) \to \nu\left(([0,x]\times[0,y])^c\right),$$

the result follows. □

Our first task is to prove that Q_n and $\widetilde{Q}_n$ are close (Proposition 1). Recall $T_{\nu(S)/p}(z) = U(\nu(S)/p)z^\gamma$ and

$$q_{\nu(S)/p}(z) = \frac{\nu(S)}{p} U_1\left(\frac{\nu(S)}{p}\right) U_2\left(\frac{\nu(S)}{p}\right) f(T_{\nu(S)/p}(z)).$$

Lemma 2 Let $\varepsilon > 0$. Then for large n

$$\bar{Q}_n \subset T_{\nu(S)/p}\left\{z : q_{\nu(S)/p}(z) \le 1+\varepsilon\right\}$$

and

$$\bar{Q}_n \supset T_{\nu(S)/p}\left\{z : q_{\nu(S)/p}(z) \le 1-\varepsilon\right\}.$$

Proof It follows from (9) as in the proof of Lemma 1, that

$$\frac{n}{k} U_1\left(\frac{n}{k}\right) U_2\left(\frac{n}{k}\right)\left(\frac{np}{k\nu(S)}\right)^{-\gamma_1-\gamma_2-1} \sim \frac{\nu(S)}{p} U_1\left(\frac{\nu(S)}{p}\right) U_2\left(\frac{\nu(S)}{p}\right).$$

Hence for large n

$$\bar{Q}_n \subset \left\{u : \frac{\nu(S)}{p} U_1\left(\frac{\nu(S)}{p}\right) U_2\left(\frac{\nu(S)}{p}\right) f(u) \le 1+\varepsilon\right\}.$$

The other inclusion follows in the same way. □

Since the probability density f can be unbounded near the coordinate axes, we want to consider the part of $\bar{Q}_n$ near the axes separately from the part in the middle. Define for $\delta \in (0, \pi/4)$

$$R_\delta = \left\{(x,y) : \frac{y}{x} \wedge \frac{x}{y} \ge \tan\delta\right\}$$

Lemma 3 Let $\varepsilon > 0$ and $c > 0$. Then for large n

$$\left\{(x,y) : q_{\nu(S)/p}(x,y) \le c\right\} \cap R_\delta \subset \{(x,y) : q(x,y) \le c/(1-\varepsilon)\} \cap R_\delta$$

and

$$\left\{(x,y) : q_{\nu(S)/p}(x,y) \le c\right\} \cap R_\delta \supset \{(x,y) : q(x,y) \le c/(1+\varepsilon)\} \cap R_\delta.$$

Proof From the monotonicity of q it follows that there exists a $c' > 0$ such that

$$\left\{(x,y) : q(x,y) \le \frac{c}{1-\varepsilon}\right\} \cap R_\delta \subset \left\{(x,y) : x^2+y^2 > c'\right\} \cap R_\delta.$$

Using again the monotonicity properties of $q_{\nu(S)/p}$ and q we obtain from (3) that $q_{\nu(S)/p} \to q$ uniformly on the latter set. Hence for large n, on that set, $(1-\varepsilon)q(x,y) \le q_{\nu(S)/p}(x,y) \le (1+\varepsilon)q(x,y)$. So on that set, if $q_{\nu(S)/p}(x,y) \le c$, then $q(x,y) \le c/(1-\varepsilon)$ and if $q_{\nu(S)/p}(x,y) > c$, then $q(x,y) > c/(1+\varepsilon)$. The result follows. □

The next lemma can be shown in a similar way as Lemma 2.

Lemma 4 Let $\varepsilon > 0$. Then for large n

$$T_{\nu(S)/p}\left\{(x,y) : q(x,y) \le 1-\varepsilon\right\} \subset \widetilde{Q}_n$$

and

$$T_{\nu(S)/p}\left\{(x,y) : q(x,y) \le 1+\varepsilon, \right\} \supset \widetilde{Q}_n.$$

Write

$$W_\delta = \left\{(u,v) : \left(\frac{v}{U_2\left(\frac{\nu(S)}{p}\right)}\right)^{1/\gamma_2}\left(\frac{U_1\left(\frac{\nu(S)}{p}\right)}{u}\right)^{1/\gamma_1} \wedge \left(\frac{u}{U_1\left(\frac{\nu(S)}{p}\right)}\right)^{1/\gamma_1}\left(\frac{U_2\left(\frac{\nu(S)}{p}\right)}{v}\right)^{1/\gamma_2} \ge \tan\delta\right\}$$

and observe that $T_{\nu(S)/p}R_\delta = W_\delta$.

Lemma 5 We have

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{P(\bar{Q}_n \setminus W_\delta)}{p} = 0 \quad \text{and} \quad \lim_{\delta\downarrow 0}\limsup_{n\to\infty}\frac{P\left(T_{n/k}\left[\frac{k\nu(S)}{np}(S\setminus R_\delta)\right]\right)}{p} = 0.$$

Proof Let (x_0, y_0) be the solution of the equations $q(x,y) = 1+3\varepsilon$ and $y = x\tan\delta$. From the proof of Lemma 3 we have for large n

$$\frac{1+3\varepsilon}{1+\varepsilon}q_{\nu(S)/p}(x_0,y_0) > q(x_0,y_0) = 1+3\varepsilon,$$

i.e. $q_{\nu(S)/p}(x_0,y_0) > 1+\varepsilon$. Because of symmetry we only consider the region near the horizontal axis. From Lemma 2 we see that for large n

$$\begin{aligned}
&\frac{\nu(S)}{p}P(\bar{Q}_n \cap \{(u,v) : (v/U_2(\nu(S)/p))^{1/\gamma_2} < (U_1(\nu(S)/p)/u)^{1/\gamma_1}\tan\delta\}) \\
&\le \frac{\nu(S)}{p}P\left(T_{\nu(S)/p}\left\{(x,y) : q_{\nu(S)/p}(z) \le 1+\varepsilon,\ y < x\tan\delta\right\}\right) \\
&\le \frac{\nu(S)}{p}P\left(T_{\nu(S)/p}\left\{(x,y) : x^2+y^2 > x_0^2,\ y < x\tan\delta\right\}\right),
\end{aligned}$$

which tends to $\nu(\{(x,y) : x^2+y^2 > x_0^2,\ y < x \tan\delta\}) = \frac{1}{x_0}\Psi(\delta)$ as $n\to\infty$, which in turn tends to 0 when $\delta \downarrow 0$. Using Lemma 4, the second statement follows similarly. $\square$

Proposition 1 We have

$$\lim_{n\to\infty} \frac{P(Q_n \triangle \widetilde{Q}_n)}{p} = 0.$$

Proof Write $\bar{Q}_{n,\delta} = \bar{Q}_n \cap W_\delta$ and similarly

$$\widetilde{Q}_{n,\delta} = T_{n/k}\left(\frac{k\nu(S)}{np}(S\cap R_\delta)\right) = T_{n/k}\left(\frac{k\nu(S)}{np}\left\{(x,y) : q(x,y)\le 1, \frac{y}{x}\wedge\frac{x}{y}\ge \tan\delta\right\}\right).$$

Observe that $Q_n \subset \bar{Q}_n$ or $\bar{Q}_n \subset Q_n$. Hence we have

$$\begin{aligned}
P(Q_n\triangle\widetilde{Q}_n) &\le P(Q_n\triangle\bar{Q}_n) + P(\bar{Q}_n\triangle\widetilde{Q}_n)\\
&\le |p - P\bar{Q}_n| + P(\bar{Q}_{n,\delta}\triangle\widetilde{Q}_{n,\delta}) + P(\bar{Q}_n\setminus W_\delta) + P\left(\widetilde{Q}_n \setminus T_{n/k}\left(\frac{k\nu(S)}{np}(S\cap R_\delta)\right)\right)\\
&\le |p - P\bar{Q}_{n,\delta}| + P(\bar{Q}_{n,\delta}\triangle\widetilde{Q}_{n,\delta}) + 2P(\bar{Q}_n\setminus W_\delta) + P\left(\widetilde{Q}_n \setminus T_{n/k}\left(\frac{k\nu(S)}{np}(S\cap R_\delta)\right)\right).
\end{aligned}$$

Let $\varepsilon > 0$. From Lemmas 2-4 and (2) it follows that as $n\to\infty$

$$\begin{aligned}
\frac{\nu(S)}{p}P(\bar{Q}_{n,\delta}\triangle\widetilde{Q}_{n,\delta}) &\le \frac{\nu(S)}{p}P\left(T_{\nu(S)/p}\left\{(x,y) : 1-\varepsilon\le q(x,y)\le 1+\varepsilon, \frac{y}{x}\wedge\frac{x}{y}\ge\tan\frac{\delta}{2}\right\}\right)\\
&\to \nu\left(\left\{(x,y) : 1-\varepsilon\le q(x,y)\le 1+\varepsilon, \frac{y}{x}\wedge\frac{x}{y}\ge\tan\frac{\delta}{2}\right\}\right).
\end{aligned}$$

The latter expression is less than $\left((1+\varepsilon)^{1/(\gamma_1+\gamma_2+1)} - (1-\varepsilon)^{1/(\gamma_1+\gamma_2+1)}\right)\nu(S)$, which in turn tends to 0 when $\varepsilon\downarrow 0$. Hence for all $\delta\in(0,\pi/4)$

$$\lim_{n\to\infty}\frac{P(\bar{Q}_{n,\delta}\triangle\widetilde{Q}_{n,\delta})}{p} = 0. \tag{13}$$

Similarly it follows that

$$\lim_{n\to\infty}\frac{P\bar{Q}_{n,\delta}}{p} = \frac{\nu(S\cap R_\delta)}{\nu(S)}. \tag{14}$$

Now the statement follows from Lemma 5, (13) and letting $\delta\downarrow 0$ in (14). $\square$

Our next task is to prove that $\widetilde{Q}_n$ and $\widehat{Q}_n$ are close (Proposition 3). First we need two results for $\widehat{\psi}$.

Proposition 2 Let $\eta \in (0, \pi/4)$. Then as $n \to \infty$

$$\sup_{\theta \in [\eta, \pi/2-\eta]} |\widehat{\psi}(\theta) - \psi(\theta)| \xrightarrow{\mathbb{P}} 0.$$

Proof Define

$$\psi_n(\theta) = \int_{\theta-h}^{\theta+h} \frac{1}{h} K\left(\frac{t-\theta}{h}\right) d\Psi(t).$$

We have, writing $I = [\eta, \pi/2 - \eta]$,

$$\sup_{\theta \in I} |\psi_n(\theta) - \psi(\theta)| \leq \sup_{\theta \in I} \int_{-h}^{h} |\psi(\theta + t) - \psi(t)| \frac{1}{h} K\left(\frac{t}{h}\right) dt$$

$$\leq \sup_{\theta \in I} \sup_{-h \leq t \leq h} |\psi(\theta + t) - \psi(t)| \to 0, \tag{15}$$

by the uniform continuity of ψ on $[\eta/2, \pi/2 - \eta/2]$.

So it remains to show that

$$\sup_{\theta \in I} |\widehat{\psi}(\theta) - \psi_n(\theta)| \xrightarrow{\mathbb{P}} 0.$$

Observe that

$$\begin{aligned} \widehat{\psi}(\theta) - \psi_n(\theta) &= \int_{\theta-h}^{\theta+h} \frac{1}{h} K\left(\frac{t-\theta}{h}\right) d(\widehat{\Psi}(t) - \Psi(t)) \\ &= -\frac{1}{h} \frac{1}{\sqrt{k}} \int_{-1}^{1} (\alpha_n(\theta + ht) - \alpha_n(\theta)) dK(t), \end{aligned}$$

where $\alpha_n := \sqrt{k}(\widehat{\Psi} - \Psi)$. Hence we have

$$\sup_{\theta \in I} |\widehat{\psi}(\theta) - \psi_n(\theta)| \leq \frac{\int_{-1}^{1} |dK(t)|}{h\sqrt{k}} \sup_{\theta \in I} \sup_{-1 \leq t \leq 1} |\alpha_n(\theta + ht) - \alpha_n(\theta)|. \tag{16}$$

Denote the continuous limiting process of α_n with α. Invoking a Skorohod construction (but keeping the same notation) we see that the right-hand side of (16) is equal to

$$\frac{\int_{-1}^{1} |dK(t)|}{h\sqrt{k}} \left\{ \left[\sup_{\theta \in I} \sup_{-1 \leq t \leq 1} |\alpha(\theta + ht) - \alpha(\theta)| \right] + o_p(1) \right\} = o_p(1),$$

by the uniform continuity of α on $[\eta/2, \pi/2 - \eta/2]$. □

Lemma 6 There exists a $c > 0$ such that with probability tending to one, as $n \to \infty$,

$$\inf_{\theta \in (0, \pi/2)} \widehat{\psi}(\theta) \cos^{1-\widehat{\gamma}_1}(\theta \wedge (\pi/2 - h)) \sin^{1-\widehat{\gamma}_2}(\theta \vee h) > c.$$

Proof By a symmetry argument it suffices to show

$$\inf_{\theta\in(0,\pi/4]} \widehat{\psi}(\theta)\sin^{1-\widehat{\gamma}_2}(\theta\vee h) > c. \tag{17}$$

As in the proof of Proposition 2, we obtain (with the obvious extension of α_n)

$$\sup_{\theta\in(0,\pi/4]} |\widehat{\psi}(\theta)-\psi_n(\theta)| \le \frac{\int_{-1}^{1}|dK(t)|}{h\sqrt{k}} \sup_{\theta\in(0,\pi/4]}\sup_{-1\le t\le 1}|\alpha_n(\theta+ht)-\alpha_n(\theta)| \xrightarrow{\mathbb{P}} 0. \tag{18}$$

Also for $\theta\in(0,\pi/4]$

$$\psi_n(\theta)=\int_{-1}^{1}\psi(\theta+ht)K(t)dt \ge \int_0^1 K(t)dt \inf_{\theta\in(0,\pi/4]}\psi(\theta) = \frac{1}{2}\inf_{\theta\in(0,\pi/4]}\psi(\theta).$$

Hence with probability tending to one $\inf_{\theta\in(0,\pi/4]}\widehat{\psi}(\theta) > c$. This completes the proof of (17) when $\gamma_2>1$. In fact we also have (17) when $\gamma_2=1$, since $\sin^{1-\widehat{\gamma}_2}(\theta\vee h) = \sin^{O_p(1/\sqrt{k})}(\theta\vee h)\xrightarrow{\mathbb{P}} 1$ uniformly for $\theta\in(0,\pi/4]$.

So in the sequel we assume $0<\gamma_2<1$. We obtain from (18)

$$\sup_{\theta\in(0,\pi/4]} |\widehat{\psi}(\theta)-\psi_n(\theta)|\sin^{1-\widehat{\gamma}_2}(\theta\vee h)\xrightarrow{\mathbb{P}} 0.$$

When $h\le\theta\le\pi/4$, with probability tending to one

$$\begin{aligned}
&\sin^{1-\widehat{\gamma}_2}(\theta\vee h)\psi_n(\theta) = \sin^{1-\widehat{\gamma}_2}(\theta\vee h)\int_{\theta-h}^{\theta+h}\frac{1}{h}K\left(\frac{t-\theta}{h}\right)\psi(t)dt\\
&= \sin^{\gamma_2-\widehat{\gamma}_2}(\theta\vee h)\int_{\theta-h}^{\theta+h}\frac{1}{h}K\left(\frac{t-\theta}{h}\right)\psi(t)(\sin t)^{1-\gamma_2}\left(\frac{\sin(\theta\vee h)}{\sin t}\right)^{1-\gamma_2}dt\\
&\ge \sin^{\gamma_2-\widehat{\gamma}_2}(\theta\vee h)\int_{\theta-h}^{\theta+h}\frac{1}{h}K\left(\frac{t-\theta}{h}\right)\psi(t)(\sin t)^{1-\gamma_2}(\cos t)^{2+\gamma_2}\left(\frac{\sin\theta}{\sin(\theta+h)}\right)^{1-\gamma_2}dt\\
&\ge \sin^{\gamma_2-\widehat{\gamma}_2}(\theta\vee h)\int_{\theta-h}^{\theta+h}\frac{1}{h}K\left(\frac{t-\theta}{h}\right)dt\\
&\qquad\cdot\psi\left(\frac{\pi}{4}+h\right)\left(\sin\left(\frac{\pi}{4}+h\right)\right)^{1-\gamma_2}\left(\cos\left(\frac{\pi}{4}+h\right)\right)^{2+\gamma_2}\left(\frac{\sin h}{\sin 2h}\right)^{1-\gamma_2} > c,
\end{aligned}$$

where we have used for the second inequality that $\sin^{1-\gamma_2}t\cos^{2+\gamma_2}t\psi(t)$ is decreasing, which holds since $q(1,\tan t)$ is decreasing.

Finally consider $0<\theta<h$. We again use that $\sin^{1-\gamma_2}t\cos^{2+\gamma_2}t\psi(t)$ is decreasing and that hence $\cos^{2+\gamma_2}t\psi(t)$ is also decreasing ($0\le t\le\pi/4$). We have with probability

tending to one

$$\sin^{1-\widehat{\gamma}_2}(\theta \vee h)\psi_n(\theta) \geq \sin^{1-\widehat{\gamma}_2} h \int_{\theta-h}^{\theta+h} \frac{1}{h} K\left(\frac{t-\theta}{h}\right) \psi(t)(\cos t)^{2+\gamma_2} dt$$

$$\geq \int_{\theta-h}^{\theta+h} \frac{1}{h} K\left(\frac{t-\theta}{h}\right) dt \ \psi(2h)(\cos 2h)^{2+\gamma_2}(\sin h)^{1-\widehat{\gamma}_2}$$

$$= \psi(2h)(\cos 2h)^{2+\gamma_2}(\sin 2h)^{1-\gamma_2}(\sin h)^{\gamma_2-\widehat{\gamma}_2}\left(\frac{\sin h}{\sin 2h}\right)^{1-\gamma_2}$$

$$\geq \psi\left(\frac{\pi}{4}\right)\left(\cos\frac{\pi}{4}\right)^{2+\gamma_2}\left(\sin\frac{\pi}{4}\right)^{1-\gamma_2}(\sin h)^{\gamma_2-\widehat{\gamma}_2}\left(\frac{\sin h}{\sin 2h}\right)^{1-\gamma_2} > c. \qquad \square$$

Lemma 7 As $n \to \infty$

$$\widehat{\nu(S)} \xrightarrow{\mathbb{P}} \nu(S).$$

Proof Recall that

$$\nu(S) = \int_0^{\pi/2} \left(\frac{1}{\gamma_1\gamma_2}\psi(\theta)\cos^{1-\gamma_1}\theta \sin^{1-\gamma_2}\theta\right)^{-\frac{1}{\gamma_1+\gamma_2+1}} \psi(\theta)d\theta.$$

From Proposition 2 it readily follows that for fixed $\eta \in (0, \pi/4)$

$$\int_\eta^{\pi/2-\eta} \left(\frac{1}{\widehat{\gamma}_1\widehat{\gamma}_2}\widehat{\psi}(\theta)\cos^{1-\widehat{\gamma}_1}(\theta \wedge (\pi/2-h))\sin^{1-\widehat{\gamma}_2}(\theta \vee h)\right)^{-\frac{1}{\widehat{\gamma}_1+\widehat{\gamma}_2+1}} \widehat{\psi}(\theta)d\theta$$

$$\xrightarrow{\mathbb{P}} \int_\eta^{\pi/2-\eta} \left(\frac{1}{\gamma_1\gamma_2}\psi(\theta)\cos^{1-\gamma_1}\theta\sin^{1-\gamma_2}\theta\right)^{-\frac{1}{\gamma_1+\gamma_2+1}} \psi(\theta)d\theta,$$

which (since $\nu(S) < \infty$) in turn tends to $\nu(S)$ when $\eta \downarrow 0$.

It remains to consider

$$\int_0^\eta \left(\frac{1}{\widehat{\gamma}_1\widehat{\gamma}_2}\widehat{\psi}(\theta)\cos^{1-\widehat{\gamma}_1}(\theta \wedge (\pi/2-h))\sin^{1-\widehat{\gamma}_2}(\theta \vee h)\right)^{-\frac{1}{\widehat{\gamma}_1+\widehat{\gamma}_2+1}} \widehat{\psi}(\theta)d\theta;$$

the integral on $[\pi/2-\eta, \pi/2]$ can be dealt with in the same way. From Lemma 6 we see, with probability tending to one, that this expression is bounded from above by some constant times $\int_0^\eta \widehat{\psi}(\theta)d\theta$. The latter expression is in turn bounded by

$$\int_0^\eta \int_{\theta-h}^{\theta+h} \frac{1}{h}K\left(\frac{t-\theta}{h}\right) d\widehat{\Psi}(t)d\theta \leq \int_{-h}^{\eta+h}\int_{t-h}^{t+h} \frac{1}{h}K\left(\frac{t-\theta}{h}\right) d\theta d\widehat{\Psi}(t) = \widehat{\Psi}(\eta+h),$$

which converges in probability to $\Psi(\eta)$ when $n \to \infty$. The fact that $\lim_{\eta\downarrow 0}\Psi(\eta) = 0$ completes the proof. $\square$

For $\delta \in (0, \pi/4)$, define $S_\delta = S \cap R_\delta = S \cap \{(x,y) : \delta \le \theta \le \pi/2 - \delta\}$ and $\widehat{S}_\delta = \widehat{S} \cap R_\delta = \widehat{S} \cap \{(x,y) : \delta \le \theta \le \pi/2 - \delta\}$. The following lemma follows easily, using Proposition 2.

Lemma 8 Let $\varepsilon > 0$. Then with probability tending to one, as $n \to \infty$,

$$(1+\varepsilon) S_\delta \subset \widehat{S}_\delta \subset (1-\varepsilon) S_\delta.$$

Recall the definition of $\widetilde{Q}_{n,\delta}$ in the proof of Proposition 1 and write

$$\widehat{Q}_{n,\delta} = \widehat{T}_{n/k}\left(\frac{\widehat{k\nu(S)}}{np} \left\{ (x,y) : x^{1-\widehat{\gamma}_1} y^{1-\widehat{\gamma}_2} \widehat{g}(x,y) \le \widehat{\gamma}_1 \widehat{\gamma}_2, \frac{y}{x} \wedge \frac{x}{y} \ge \tan\delta \right\} \right);$$

here for $z = (x,y)$, $\widehat{T}_{n/k}(z) = \widehat{U}(n/k) z^{\widehat{\gamma}} = (\widehat{U}_1(n/k) x^{\widehat{\gamma}_1}, \widehat{U}_2(n/k) y^{\widehat{\gamma}_2})$ and for $(x,y) = (r\cos\theta, r\sin\theta)$, $\widehat{g}(x,y) = \widehat{\psi}(\theta)/r^3$.

Lemma 9 Let $\delta \in (0, \pi/4)$. Then as $n \to \infty$

$$\frac{P(\widetilde{Q}_{n,\delta} \triangle \widehat{Q}_{n,\delta})}{p} \xrightarrow{\mathbb{P}} 0.$$

Proof We have

$$\begin{aligned} & P(\widetilde{Q}_{n,\delta} \triangle \widehat{Q}_{n,\delta}) \\ \le\ & P\left(T_{n/k}\left(\frac{k\nu(S)}{np} S_\delta \right) \triangle \widehat{T}_{n/k}\left(\frac{k\nu(S)}{np} S_\delta \right) \right) \\ & +P\left(\widehat{T}_{n/k}\left(\frac{k\nu(S)}{np} S_\delta \right) \triangle \widehat{T}_{n/k}\left(\frac{\widehat{k\nu(S)}}{np} \widehat{S}_\delta \right) \right) =: V_1 + V_2. \end{aligned} \tag{19}$$

We consider

$$V_1 = P\left(T_{n/k}\left[\left(\frac{k\nu(S)}{np} S_\delta \right) \triangle \left(\frac{k\nu(S)}{np} \left(\frac{np}{k\nu(S)} T_{n/k}^{-1} \widehat{T}_{n/k} \frac{k\nu(S)}{np} \right) S_\delta \right) \right] \right).$$

Now for $(x,y) \in (0,\infty)^2$,

$$\begin{aligned} & \frac{np}{k\nu(S)} T_{n/k}^{-1} \widehat{T}_{n/k} \frac{k\nu(S)}{np} (x,y) \\ = & \left(\left(\frac{\widehat{U}_1\left(\frac{n}{k}\right)}{U_1\left(\frac{n}{k}\right)} \left(\frac{k\nu(S)}{np} \right)^{\widehat{\gamma}_1 - \gamma_1} \right)^{1/\gamma_1} x^{\widehat{\gamma}_1/\gamma_1}, \left(\frac{\widehat{U}_2\left(\frac{n}{k}\right)}{U_2\left(\frac{n}{k}\right)} \left(\frac{k\nu(S)}{np} \right)^{\widehat{\gamma}_2 - \gamma_2} \right)^{1/\gamma_2} y^{\widehat{\gamma}_2/\gamma_2} \right). \end{aligned}$$

Applying a Skorohod construction (but keeping the same notation) we have, using $(\log np)/\sqrt{k} \to 0$, that with probability 1

$$\left(\left(\frac{\widehat{U}_1\left(\frac{n}{k}\right)}{U_1\left(\frac{n}{k}\right)}\left(\frac{k\nu(S)}{np}\right)^{\widehat{\gamma}_1-\gamma_1}\right)^{1/\gamma_1}, \frac{\widehat{\gamma}_1}{\gamma_1}, \left(\frac{\widehat{U}_2\left(\frac{n}{k}\right)}{U_2\left(\frac{n}{k}\right)}\left(\frac{k\nu(S)}{np}\right)^{\widehat{\gamma}_2-\gamma_2}\right)^{1/\gamma_2}, \frac{\widehat{\gamma}_2}{\gamma_2}\right) \to (1,1,1,1).$$

Writing

$$S^*_{n,\delta} = \frac{np}{k\nu(S)} T^{-1}_{n/k}\widehat{T}_{n/k}\frac{k\nu(S)}{np}S_\delta\,,$$

we obtain that for $\varepsilon \in (0,1)$ and $n \geq 1/\varepsilon$,

$$\frac{V_1}{p} = \frac{1}{p}P\left(T_{n/k}\frac{k\nu(S)}{np}\left(S_\delta \triangle S^*_{n,\delta}\right)\right) \leq \frac{1}{p}P\left(T_{n/k}\frac{k\nu(S)}{np}\bigcup_{m\geq 1/\varepsilon}\left(S_\delta \triangle S^*_{m,\delta}\right)\right).$$

Letting $n \to \infty$, the latter expression tends to $\nu\left(\cup_{m\geq 1/\varepsilon}\left(S_\delta \triangle S^*_{m,\delta}\right)\right)/\nu(S)$ by Lemma 1, which in turn tends to 0 with probability 1, when $\varepsilon \downarrow 0$.

Next consider

$$\frac{V_2}{p} = \frac{1}{p}P\left(\widehat{T}_{n/k}\frac{k}{np}\left(\nu(S)S_\delta \triangle \widehat{\nu(S)}\widehat{S}_\delta\right)\right).$$

Lemmas 7 and 8 imply that with probability tending to one this expression is bounded from above by

$$\begin{aligned}
& \frac{1}{p}P\left(\widehat{T}_{n/k}\frac{k}{np}\left((1-\varepsilon)^2\nu(S)S_\delta\right) \setminus \left((1+\varepsilon)^2\nu(S)S_\delta\right)\right) \\
= \; & \frac{1}{p}P\left(\widehat{T}_{n/k}\frac{k\nu(S)}{np}\left((1-\varepsilon)^2 S_\delta\right) \setminus \left((1+\varepsilon)^2 S_\delta\right)\right) \\
= \; & \frac{1}{p}P\left(T_{n/k}\frac{k\nu(S)}{np}\left(\frac{np}{k\nu(S)}T^{-1}_{n/k}\widehat{T}_{n/k}\frac{k\nu(S)}{np}\right)\left((1-\varepsilon)^2 S_\delta\right) \setminus \left((1+\varepsilon)^2 S_\delta\right)\right)
\end{aligned}$$

Now we are in a similar position as when dealing with V_1/p : letting $n \to \infty$ and next $\varepsilon \downarrow 0$, the latter expression converges to 0 with probability 1.

Combining the results for V_1/p and V_2/p with (19) completes the proof. □

Lemma 10 For every $\varepsilon > 0$ there exists a $\delta > 0$ and an $n_{\varepsilon,\delta}$ such that for $n \geq n_{\varepsilon,\delta}$

$$\mathbb{P}\left(\frac{P\left(\widehat{Q}_n \setminus \widehat{Q}_{n,\delta}\right)}{p} \geq \varepsilon\right) \leq \varepsilon.$$

Proof From Lemma 6 we obtain the existence of a $\widetilde{c}$ such that with probability tending to one, as $n \to \infty$,

$$\widehat{S} \setminus \widehat{S}_\delta \subset \{(x,y) : r \geq \widetilde{c},\ \theta \notin [\delta, \pi/2 - \delta]\} =: Z_{\widetilde{c},\delta}.$$

This implies that with probability tending to one

$$P\left(\widehat{Q}_n \setminus \widehat{Q}_{n,\delta}\right) \leq P\left(\widehat{T}_{n/k} \frac{k\widehat{\nu(S)}}{np} Z_{\widetilde{c},\delta}\right).$$

Hence it suffices to show that there exists a $\delta > 0$ and an $n_{\varepsilon,\delta}$ such that for $n \geq n_{\varepsilon,\delta}$

$$\mathbb{P}\left(\frac{P\left(\widehat{T}_{n/k}\frac{k\widehat{\nu(S)}}{np} Z_{\widetilde{c},\delta}\right)}{p} \geq \varepsilon\right) \leq \varepsilon.$$

This can be proved using similar arguments as in the proof of the previous lemma. □

Proposition 3 We have, as $n \to \infty$,

$$\frac{P(\widetilde{Q}_n \triangle \widehat{Q}_n)}{p} \xrightarrow{\mathbb{P}} 0.$$

Proof The statement follows from Lemmas 9, 5 (second statement) and 10. □

Proof of the Theorem The result follows from combining Propositions 1 and 3. □

Acknowledgements We are very grateful to Andrea Krajina for helping out with the simulations.

References

Dekkers, A.L.M., Einmahl, J.H.J. and de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution, *Ann. Statist.*, **17**, 1833–1855.

Einmahl, J.H.J., Li, J. and Liu, R.Y. (2009). Thresholding events of extreme in simultaneous monitoring of multiple risks. *J. Amer. Statist. Assoc.*, to appear.

Einmahl, J.H.J. and Segers, J. (2009). Maximum empirical likelihood estimation of the spectral measure of an extreme value distribution. *Ann. Statist.*, to appear.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction.* Springer, New York.

de Haan, L. and Huang, X. (1995). Large quantile estimation in a multivariate setting, *J. Multivariate Anal.*, **53**, 247–263.

de Haan, L. and Resnick, S.I. (1977). Limit theory for multivariate sample extremes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **40**, 317–337.

Joe, H., Smith, R. and Weissman, I. (1992). Bivariate threshold methods for extremes, *J. Roy. Statist. Soc. Ser. B*, **54**, 171–183.

Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized Pareto distributions, *Bernoulli*, **12**, 917–930.

Dept. of Econometrics & OR and CentER
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands
j.h.j.einmahl@uvt.nl
ldehaan@ese.eur.nl