



HAL
open science

Modeling and estimation of conditional excesses

Anne-Laure Fougères, Philippe Soulier

► **To cite this version:**

Anne-Laure Fougères, Philippe Soulier. Modeling and estimation of conditional excesses. 2008. hal-00288135v1

HAL Id: hal-00288135

<https://hal.science/hal-00288135v1>

Preprint submitted on 14 Jun 2008 (v1), last revised 13 Sep 2010 (v4)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modeling and estimation of conditional excesses

Anne-Laure Fougères* Philippe Soulier*

June 14, 2008

Abstract

We investigate conditions for the existence of the limiting distribution of a bivariate random vector when one component becomes large. We revisit the existing literature on the topic, and present some new sufficient conditions. We focus on conditions which can be expressed in terms of geometric properties of the level curves of the density of the random vector. When the limiting conditional distribution exists, we propose consistent nonparametric estimators of this distribution and of the related normalizing sequences. We also provide a semi-parametric extrapolation procedure that allows the estimation of conditional probabilities when the conditioning event is extreme. A small simulation study illustrates our results.

1 Introduction

In many practical situations, there is a need of modeling and estimating multivariate extreme events. Extreme means, roughly speaking, that no observations are available in the domain of interest. An increasing number of examples can be found in the literature, in the actuarial and financial context (risk assessment, contagion) or in reliability theory (multivariate failure sets), among others. See for instance the recent books of De Haan and Ferreira (2006), Balkema and Embrechts (2007) (in the financial context) and Resnick (2007) (for teletraffic applications).

Extreme value theory provides an efficient mathematical framework to deal with these problems in the situation where the largest values of the

*Université Paris Ouest-Nanterre

variables of interest tend to occur simultaneously. This situation is referred to as *asymptotic dependence* in extreme value theory. In the opposite case of *asymptotic independence*, the standard theory needs to be refined. To this purpose, building on the ideas of second order regular variation (see for instance (De Haan and Ferreira, 2006, Appendix B)), Resnick (2002); Maulik and Resnick (2004) have introduced the concept of hidden regular variation. Another approach consists in studying, if it exists, the limiting distribution of a random vector conditionally on one component to be large. Formally stated in the bivariate case, this corresponds to assuming that there exist functions m , a and ψ , and a bivariate distribution function (cdf) K such that

$$\lim_{t \rightarrow \infty} \mathbb{P}(X \leq t + \psi(t)x ; Y \leq m(t) + a(t)y \mid X > t) = K(x, y) . \quad (1)$$

This approach was suggested by Heffernan and Tawn (2004) and thoroughly investigated by Heffernan and Resnick (2007). Two different kinds of issues must be addressed.

The first one is to study models under which condition (1) holds. Results of this kind were obtained by Eddy and Gale (1981) and Berman (1992) for spherical distributions, for bivariate elliptic distributions by Abdous et al. (2005) and for multivariate elliptical distributions and related distributions by Hashorva (2006); Hashorva et al. (2007). One essential feature of elliptical distributions is that the level sets of their density are ellipses in the bivariate case, or ellipsoids in general. Such geometric considerations have been deeply investigated and generalized in many directions by Barbe (2003) and Balkema and Embrechts (2007). It must be noted that this geometric approach is incompatible with the customary transformation of the marginal distributions to prescribed ones, be it uniform, Fréchet or Gumbel, since such transformations will generally ruin the geometric properties of the joint density. A specific feature of these models is that the property of asymptotic dependence or independence is related to the nature of the marginal distributions. If they are regularly varying, then the components are asymptotically dependent; if the marginal distributions are rapidly varying, i.e. belong to the maximum domain of attraction of the Gumbel distribution, then the components are asymptotically independent. In addition to the general results referred above, Heffernan and Tawn (2004) showed that Condition (1) holds for a variety of commonly used bivariate distributions.

The second direction of research is to start from condition (1) and to investigate its consequences. This was done in Heffernan and Resnick (2007), in

the framework of regular variation. A straightforward consequence of Condition (1) is that X belongs to the domain of attraction of an extreme value distribution. But, the limiting distribution K is not necessarily a bivariate extreme value distribution. In particular, if X and Y are independent, then choosing $a \equiv 1$ and $m \equiv 0$ in (1) yields a limiting distribution which is the product of the marginal distribution functions of X and Y . Another important finding of Heffernan and Resnick (2007) is that Condition (1) is not robust to nonlinear transformations of the components of the vector.

Having investigated Condition (1), the remaining task is to develop statistical procedures for the estimation of the normalizing functions and the limiting distribution. Heffernan and Resnick (2005) have developed estimation procedures in the related framework of hidden regular variation. In the conditional approach considered here, Heffernan and Tawn (2004) outlined some heuristic procedures.

Our contribution to this problem is twofold. In Section 2, we present bivariate models with rapidly varying marginal distributions that satisfy Condition (1). We recall one result of Balkema and Embrechts (2007) and complement it with a new result (Theorem 3) that also includes several examples that were obtained separately. Since the literature in this domain is not yet widely developed, we hope that we provide a reasonably exhaustive survey of the existing results and methods. In Section 3, we recall the main results of Heffernan and Resnick (2007) and apply them to derive valid statistical procedures. We introduce new estimators of the functions a and m and of the limiting distribution K . These estimators are studied by means of the standard point process techniques that Condition (1), rephrased in the terminology of vague convergence (cf. Resnick (2007)), allows to use. We also present estimators suited to the case of rapidly varying marginal distributions and asymptotic independence. In Section 3.2, we consider non parametric estimators, which can be proved consistent under Condition (1) only (and a technical moment assumption). In Section 3.4, we study semi-parametric estimators, which are of interest when a model is assumed for the tail of the marginal distribution of the conditioning variable, and when extrapolation out of the range of the observations is needed. The paper is concluded by a short numerical illustration in Section 4. The proofs of our results are deferred to the appendix.

2 Models

We start by recalling some definitions that will be used throughout the paper. A random variable with value in $[x_0, x_1)$ or its distribution function H are said to have a *rapidly varying upper tail* if there exists a positive function ψ such that

$$\lim_{t \rightarrow x_1} \frac{1 - H(t + \psi(t)x)}{1 - H(t)} = e^{-x} . \quad (2)$$

It is well known (cf. (De Haan and Ferreira, 2006, Theorem 1.2.5)) that a random variable X is in the max-domain of attraction of the Gumbel distribution if and only if its cdf F has a rapidly varying upper tail. The function ψ is called an *auxiliary function*. It is defined up to asymptotic equivalence and necessarily satisfies $\psi(x) = o(x)$ if $x_1 = \infty$ and $\psi(x) = o(x_1 - x)$ if $x_1 < \infty$. In the sequel for simplicity, we only consider the case $x_1 = \infty$.

We will say that two random variables X and Y are *asymptotically independent* if (X, Y) belongs to the bivariate maximum domain of attraction of an extreme value distribution with independent marginals.

We can now recall the results for an elliptical random vector, i.e. a bivariate random vector (X, Y) that can be expressed as

$$(X, Y) = R(\cos \Theta, \rho \cos \Theta + \sigma \sin \Theta) \quad (3)$$

with $\sigma^2 = 1 - \rho^2$, in terms of a positive random variable R called "radial component" and an "angular" random variable Θ uniformly distributed on $[0, 2\pi[$. The following result was originally proved in the case $\rho = 0$ as a technical lemma under restrictive conditions in Eddy and Gale (1981), then in Berman (1992). The general result was proved in Abdous et al. (2005) in the bivariate case and in Hashorva (2006) in a multivariate setting. Throughout the paper Φ will denote the cdf of the standard normal distribution.

Theorem 1. *Let (X, Y) be an elliptic random vector as defined in (3). If R has a rapidly varying upper tail with auxiliary function ψ , then X and Y have rapidly varying upper tails, are asymptotically independent, and*

$$\lim_{s \rightarrow \infty} \mathbb{P}(X \leq x + \psi(x)\xi, Y \leq \rho x + \sqrt{x\psi(x)}z \mid X > x) = (1 - e^{-\xi})\Phi(z) .$$

If the radial component R has a density h , then the vector (X, Y) has the density f defined by

$$f(x, y) = \frac{h(\sqrt{x^2 + (y - \rho x)^2/\sigma^2})}{\sqrt{x^2 + (y - \rho x)^2/\sigma^2}}.$$

The level lines of the density are homothetic ellipses $x^2 + (y - \rho x)^2/\sigma^2 = c^2$.

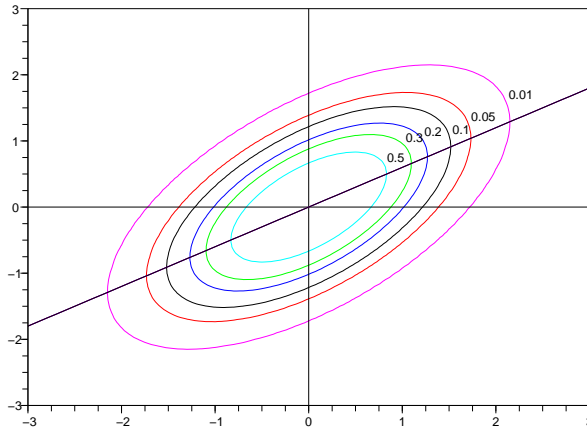


Figure 1: Level lines of the density of an elliptical distribution. The slope of the straight line is $\rho = .6$.

This result can be generalized in two directions: either by weakening the assumptions on the level lines of the density or by extending the representation (3). This will be done in the following two subsections.

2.1 Asymptotically elliptical distributions

In this section, we state a bivariate version of (Balkema and Embrechts, 2007, Theorem 11.1). We first need the following definition (see (Balkema and Embrechts, 2007, Section 11.2)).

Definition 1. A function $L : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ belongs to the class \mathcal{L} if for all $(x, y) \in \mathbb{R}^2$,

$$\lim_{(\xi, \zeta) \rightarrow \infty} \frac{L(x + \xi, y + \zeta)}{L(\xi, \zeta)} = 1.$$

Assumption 1. The random vector (X, Y) has a density f such that

$$f(x, y) = e^{-I(x, y)} L(x, y), \quad (4)$$

where $L \in \mathcal{L}$, and the function I satisfies:

$$\begin{aligned} I(x, y) &= p \circ n(x, y), \\ p(r) &= \int_0^r \frac{ds}{\psi(s)}, \end{aligned} \quad (5)$$

ψ is absolutely continuous with $\lim_{x \rightarrow \infty} \psi'(x) = 0$ and $n : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is 1-homogeneous, n^2 is twice differentiable and the Hessian matrix of n is positive definite.

Theorem 2 (Balkema and Embrechts (2007)). Under Assumption 1, X and Y have both rapidly varying tails with auxiliary function ψ , are asymptotically independent, and there exist real numbers ρ and σ such that

$$\begin{aligned} \lim_{x \rightarrow \infty} \mathbb{P}(X \leq x + \psi(x)\xi ; Y \leq \rho x + \sigma \sqrt{x\psi(x)}z \mid X > x) \\ = (1 - e^{-\xi})\Phi(z). \end{aligned} \quad (6)$$

Remark 1. This result shows that Assumption 1 is a sufficient condition for the limit (1) to hold, with $m(x) = \rho x$, $a(x) = \sigma \sqrt{x\psi(x)}$ and $K(x, y) = (1 - e^{-x})\Phi(y)$. The constants ρ and σ are characterized by the second order expansion of the function n

$$n(1 + x, \rho + \sigma y) = 1 + x + y^2/2 + o(x^2 + y^2).$$

This condition implies that the tangent at the point $(1, \rho)$ to the curve $n(\xi, \zeta) = 1$ is vertical. Note that the level lines of the function n are not those of the density f defined in (4), unless the function L is constant, but, loosely speaking, the level lines of f converge to those of n . Obviously, Theorem 1 is a particular case of Theorem 2.

Example 1. Let h and g be density functions defined on $[0, \infty)$ and $[-\pi/2, \pi/2]$, respectively. The function f defined by

$$f(x, y) = \frac{h(x^2 + (y - \rho x)^2 / (1 - \rho^2))}{\sqrt{x^2 + (y - \rho x)^2 / (1 - \rho^2)}} g \circ \arctan((y - \rho x) / x \sqrt{1 - \rho^2}) \quad (7)$$

is then a bivariate density function on \mathbb{R}^2 . If g is a constant, then f is the density of an elliptical vector. If h can be expressed as in (5) and if

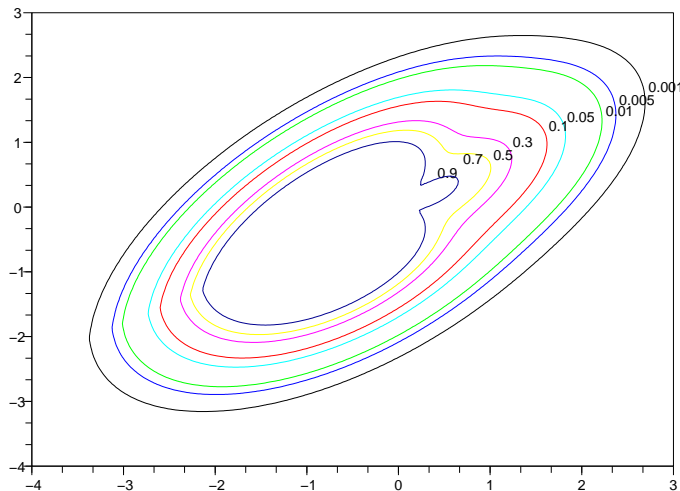


Figure 2: Level lines of the density given by (7)

g is continuous and bounded above and away from zero, then f satisfies Assumption 1. Figure 2 shows the level lines of such a density, with $\rho = .6$, $h(t) = \exp(t^2/2)/\sqrt{2\pi}$ and $g(t) = c\{1 + [\arctan(t)^2 - (\pi/4)^2]^2\}$. The level lines seem to be asymptotically homothetic.

Remark 2. It is important to note that under Assumption 1 the normalizing functions m and a satisfy $a(x) = o(m(x))$, since in the present context $m(x) = \rho x$ and $a(x) = \sigma \sqrt{x\psi(x)}$ with $\psi(x) = o(x)$. This implies that only the local behaviour of the curve $n(\xi, \zeta) = 1$ around the point $(1, \rho)$ matters. In other words, the limit (6) still holds if (X, Y) is conditioned to remain in the cone $\{(\rho - \epsilon)x \leq y \leq (\rho + \epsilon)x\}$ for any arbitrarily small $\epsilon > 0$. This suggests that Assumption 1 must only be checked *locally* to obtain the limit (6).

Example 2 (Mixture of two bivariate Gaussian vectors). Let B be a Bernoulli random variable such that $\mathbb{P}(B = 1) = p \in (0, 1)$. Let X and Z be two i.i.d standard gaussian random variables, $\rho \neq \tau \in [-1, 1]$ and define Y by

$$Y = B(\rho X + \sqrt{1 - \rho^2}Z) + (1 - B)(\tau X + \sqrt{1 - \tau^2}Z). \quad (8)$$

Then Y is a standard Gaussian variable, and (X, Y) is a mixture of two

Gaussian vectors. Figure 3 shows the level curves of the density function of the pair (X, Y) with $p = .4$, $\rho = .8$ and $\tau = -.4$. The density function of (X, Y) does not satisfy Assumption 1, and Theorem 2 cannot be applied. Indeed, applying Theorem 1 to each component of the mixture yields

$$\begin{aligned} & \mathbb{P}(Y - \rho x \leq \sqrt{1 - \rho^2}z \mid X > x) \\ &= p \mathbb{P}(\rho(X - x) + \sqrt{1 - \rho^2}Z \leq \sqrt{1 - \rho^2}z \mid X > x) \\ &+ (1 - p) \mathbb{P}(\tau(X - x) + (\tau - \rho)x + \sqrt{1 - \tau^2}Z \leq \sqrt{1 - \rho^2}z \mid X > x) \\ &\sim p\Phi(z) + (1 - p)\mathbb{1}_{\{\tau < \rho\}}. \end{aligned}$$

Thus the limiting distribution is degenerate, with a positive mass either at $-\infty$ or $+\infty$. However, a proper limiting distribution can be obtained

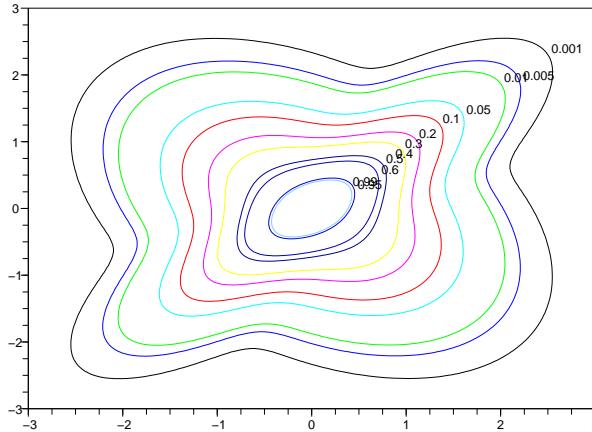


Figure 3: Level lines of the density function of the pair (X, Y) defined in Example 2.

for (X, Y) conditioned to remain in $\mathcal{C} = \{(x, y) \in \mathbb{R}^2 \mid c_1x \leq y \leq c_2x\}$ for $c_1 < c_2$ such that $\rho \in [c_1, c_2]$ and $\tau \notin [c_1, c_2]$. Denote $\mathbb{P}(\cdot \mid (X, Y) \in \mathcal{C})$ by $\mathbb{P}_{\mathcal{C}}(\cdot)$. Then

$$\lim_{x \rightarrow \infty} \mathbb{P}_{\mathcal{C}}(Y \leq \rho x + \sqrt{1 - \rho^2}z \mid X > x) = \Phi(z). \quad (9)$$

To prove this claim, we assume without loss of generality that $\rho = 0$. Then

$$\mathbb{P}_{\mathcal{C}}(Y \leq z \mid X > x) = \frac{\mathbb{P}(Y \leq z; (X, Y) \in \mathcal{C} \mid X > x)}{\mathbb{P}((X, Y) \in \mathcal{C} \mid X > x)}.$$

For fixed z and $x > z/c_2$, it holds that

$$\begin{aligned} & \mathbb{P}(Y \leq z ; (X, Y) \in \mathcal{C} \mid X > x) \\ &= p \mathbb{P}(c_1 X \leq Z \leq z \mid X > x) \\ &+ (1 - p) \mathbb{P}(c_1 X \leq \tau X + \sqrt{1 - \tau^2} Z \leq z \mid X > x) \sim p \Phi(z) . \end{aligned}$$

Since $\rho \in [c_1, c_2]$ and $\tau \notin [c_1, c_2]$, and since $z \leq c_2 x$, it is easily obtained that

$$\begin{aligned} & \lim_{x \rightarrow \infty} \mathbb{P}(c_1 X \leq Z \leq z \mid X > x) = \Phi(z) , \\ & \lim_{x \rightarrow \infty} \mathbb{P}(c_1 X \leq \tau X + \sqrt{1 - \tau^2} Z \leq z \mid X > x) = 0 . \end{aligned}$$

Thus $\lim_{x \rightarrow \infty} \mathbb{P}(Y \leq z ; (X, Y) \in \mathcal{C} \mid X > x) = p \Phi(z)$, which proves (9).

2.2 Distributions with radial representation

In this section, we show that Theorem 1 can be extended from elliptical distributions to more general bivariate distributions that admit a radial representation $R(u(T), v(T))$ where R and T are independent and u and v are more general functions than in the elliptical case. We start by collecting the assumptions that will be needed.

Assumption 2.

A The function $u : [0, 1] \rightarrow [0, 1]$ is continuous, has a unique maximum 1 at a point $t_0 \in (0, 1)$ and has an expansion

$$u(t_0 + t) = 1 - \ell(t) \tag{10}$$

where ℓ is increasing in $[-\epsilon, 0]$, decreasing in $[0, \epsilon]$ for some $\epsilon > 0$ and regularly varying at zero with index $\kappa > 0$. Its inverse ℓ^{\leftarrow} is absolutely continuous and its derivative $(\ell^{\leftarrow})'(s)$ is regularly varying at zero with index $1/\kappa - 1$.

B The function v is strictly increasing in a neighborhood of t_0 , $v(t_0) = \rho$, and the function $t \mapsto v(t_0 + t) - \rho$ is regularly varying with index $\delta \in (0, \kappa)$. Its inverse v^{\leftarrow} is absolutely continuous and its derivative is regularly varying at zero with index $1/\delta - 1$.

Assumption 3. The cdf H is rapidly varying at infinity with auxiliary function ψ , i.e.

$$\lim_{x \rightarrow \infty} \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} = e^{-t},$$

where $\bar{H} = 1 - H$.

Assumption 4. The density function $g : [0, 1] \rightarrow \mathbb{R}^+$ is regularly varying at t_0 with index $\tau > -1$ and locally bounded on $[0, 1] \setminus \{t_0\}$.

Theorem 3. Let R and T be independent random variables such that the cdf H of R satisfies Assumption 3, T admits a density g that satisfies Assumption 4, and the functions u and v satisfy Assumption 2. Define $(X, Y) = R(u(T), v(T))$. Then,

(i) there exists a function k regularly varying at zero with index $(1 + \tau)/\kappa$ such that

$$\mathbb{P}(X > x) \sim k(\psi(x)/x)\bar{H}(x); \quad (11)$$

(ii) there exists a function h regularly varying at zero with index δ/κ such that for all $y \in \mathbb{R}$,

$$\begin{aligned} \lim_{x \rightarrow \infty} \mathbb{P}(X \leq x + \psi(x)\xi, Y \leq \rho x + xh(\psi(x)/x)y \mid X > x) \\ = (1 - e^{-\xi})H_{\delta, \kappa, \tau}(y), \end{aligned}$$

with

$$H_{\delta, \kappa, \tau}(y) = \frac{\int_{-\infty}^y e^{-|s|^{\kappa/\delta}/(\kappa/\delta)} |s|^{(1+\tau)/\delta-1} ds}{\int_{-\infty}^{\infty} e^{-|s|^{\kappa/\delta}/(\kappa/\delta)} |s|^{(1+\tau)/\delta-1} ds}.$$

The proof of this result is in Appendix A. One of its main ingredient is the fact that the tail of R is rapidly varying, which implies that the normalizing function $a(x) = xh(\psi(x)/x)$ is $o(x)$. As a consequence, only the local behaviour of u and v around $(1, \rho)$ matters. This is similar to what was observed under Assumption 1, see Remark 2.

Theorem 3 handles situations where the assumptions of Theorem 2 do not hold. In some cases, the limiting distribution is nevertheless the Gaussian distribution and the normalization is the same as in Theorem 3; see Example 3.

In some other cases, the limiting distribution and the normalisation differ from those that appear in Theorem 2. This can have two types of reasons: the density of T can vanish or be unbounded at zero, or the curvature of the line parameterized by the functions u and v at the point $(u(t_0), v(t_0))$ can be infinite or zero. This is illustrated in Example 4.

It is also important to note that even though Theorem 3 states that X has a rapidly varying upper tail, with the same auxiliary function ψ as R , nothing can be said on the tails of Y . This is in contrast with the situation of Theorem 2.

Example 3. If the density g of the variable T has a positive limit at t_0 , then $\tau = 0$. If u is twice differentiable with $u'(t_0) = 0$ and $u''(t_0) \neq 0$, then $\kappa = 2$ and if $v'(0) > 0$, then $\delta = 1$. If these three conditions hold, the limiting distribution is the standard Gaussian and the normalization is $\sigma\sqrt{x\psi(x)}$. Figure 2 also illustrates this case. As will be shown later, this is actually a particular case of Theorem 2.

Example 4. Hashorva et al. (2007) have introduced a generalisation of the elliptical distributions, which they called L_p -Dirichlet distributions for all $p > 0$. We consider only the case $p > 1$. Instead of being ellipses, the level lines of the density of these distributions have the following equation:

$$|x|^p + \frac{|y - \rho x|^p}{1 - |\rho|^p} = 1,$$

with $\rho \in (-1, 1)$. To simplify the discussion, we consider the case $\rho = 0$. An admissible parametrization is given by $u(t) = (1 - t^p)^{1/p}$ and $v(t) = t$, which yields $\delta = 1$ and $\kappa = p$. Thus, Assumption 1 does not hold except if $p = 2$, which is the elliptical context. If the density of T has a positive limit at zero, then $\tau = 0$ and the cdf of the limiting distribution is then

$$H_{1,p,0}(y) = \frac{\int_{-\infty}^y e^{-|s|^p/p} ds}{2p^{1/p-1} \Gamma(1/p)}.$$

As a last remark in this section, note that the result (11) on the upper tail of X is actually a particular case of a more general result for which we have found no reference, and which might be of interest. We state it here, but omit its proof which is exactly the same as that of (11), and makes use of Lemma 17.

Lemma 4. *Let R be a nonnegative random variable whose cdf H satisfies Assumption 3. Let U be a nonnegative random variable, independent of R ,*

such that $U \leq b < \infty$ a.s. and that admits a density in a neighborhood of b which is regularly varying with index $\tau > -1$. Then RU has a rapidly varying upper tail with auxiliary function $\psi(x/b)$ and

$$\mathbb{P}(RU > x) \sim b^2 \Gamma(\tau + 1) \frac{\psi(x/b)}{x} g(\{1/b + \psi(x/b)/x\}^{-1}) \bar{H}(x/b). \quad (12)$$

2.3 Relation between the two representations

Let X, Y be random variables whose joint density f can be expressed as

$$f(x, y) = g \circ n(x, y), \quad (13)$$

where g is a nonnegative function on \mathbb{R}_+ such that $\int_0^\infty rg(r) dr < \infty$, n is a positively homogeneous function with index 1 and the level line $n(x, y) = 1$ admits the parametrization $t \rightarrow (u(t), v(t))$, $t \in [0, 1]$. The change of variable $x = ru(t)$, $y = rv(t)$ yields:

$$\mathbb{E}[\phi(X, Y)] = \int_0^\infty \int_0^1 \phi(ru(t), rv(t)) rg(r) |u(t)v'(t) - u'(t)v(t)| dr dt.$$

Hence it holds that $(X, Y) = R(u(T), v(T))$, where R has density $c(g)^{-1}rg(r)$, with $c(g) = \int_0^\infty rg(r) dr$ and T has density $c(u, v)^{-1}|uv' - u'v|$ with $c(u, v) = \int_0^1 |u'(t)v(t) - u(t)v'(t)| dt$.

Conversely, if $(X, Y) = R(u(T), v(T))$ where R and T are independent, R has a density h on $[0, \infty)$, and if there exists a point t_0 such that $u'(t_0) = 0$, $u''(t_0) > 0$ and $v'(t_0) > 0$, then the function v/u is invertible on an interval around t_0 . Let ϕ be its inverse, and define $n(x, y) = x/\{u \circ \phi(y/x)\}$. Without loss of generality, we can assume that $u(t_0) = 1$ and denote $v(t_0) = \rho$. Then n is a positively homogeneous function with index 1, and if T admits a density g , then (X, Y) has a density f in a cone $C = \{(x, y) \mid (\rho - \epsilon)x \leq y \leq (\rho + \epsilon)x\}$ around the line $y = \rho x$, defined by

$$f(x, y) = \frac{h(n(x, y))}{n(x, y)} L(x, y),$$

with

$$L(x, y) = \frac{g \circ \phi(y/x)}{|u'v - uv'| \circ \phi(y/x)}.$$

If the density g and the Jacobian $|u'v - uv'|$ are both positive and continuous at t_0 , then the function L is flat. See Lemma 20 for a proof. Thus Assumption 1 locally holds, and Theorem 2 implies Theorem 3 in this context.

3 Estimation

In this section we address the issue of estimating the functions a , m , ψ and the limiting distribution K when (1) holds.

Firstly, we must study the consequences of this convergence. In Subsection 3.1, we recall the results of Heffernan and Resnick (2007) and deduce some results that will be needed to prove the consistency of the estimators defined in Subsection 3.2. These estimators can be applied in a wider framework than that of the examples of Section 2. In Subsection 3.3, we define and study estimators more precisely suited to these examples. In Subsection 3.4, we consider the semi-parametric estimation of the functions a and ψ .

3.1 Assumptions on the distribution of (X, Y)

Assumption 5. *There exists monotone functions a , b , m and ψ such that the measure ν_n defined by*

$$\nu_n(\cdot) = n\mathbb{P} \left(\left\{ \frac{X - b(n)}{\psi \circ b(n)}, \frac{Y - m \circ b(n)}{a \circ b(n)} \right\} \in \cdot \right) \quad (14)$$

converges vaguely to a measure ν on $(-\infty, \infty] \times [-\infty, \infty]$ such that $\nu([0, \infty] \times [-\infty, \infty]) = 1$.

Remark 3. Assumption 5 implies that the marginal distribution of X belongs to a max-domain of attraction. Thus the functions ψ and b are determined by the marginal distribution of X only. The function b can and will be chosen as the left continuous inverse of the survival function of X . The function ψ satisfies

$$\lim_{x \rightarrow x_1} \frac{\psi(x + \psi(x)u)}{\psi(x)} = 1 + \gamma u, \quad (15)$$

where x_1 is the right endpoint of the support of the distribution of X and γ is the extreme value index of X . The first marginal of ν is the generalized Pareto distribution G_γ , defined by $G_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$ if $\gamma \neq 0$ and $G_0(t) = e^{-t}$.

Remark 4. Assumption 5 has little implications on the functions a and m and the second marginal of the measure ν . Let Ψ denote this distribution,

defined by

$$\Psi(z) = \int_0^\infty \int_{-\infty}^z \nu(dx, dy) .$$

If Y is independent of X then Ψ is the distribution of Y , $a \equiv 1$ and $m \equiv 0$. Thus Ψ can be *any* probability distribution. In particular, it is not necessarily an extreme-value distribution.

Remark 5. If the pair (X, Y) satisfies Assumption 5, then so does any affine transformation of (X, Y) . For instance, if X and Y have finite mean and variance, then $((X - \mathbb{E}[X])/\text{var}^{1/2}(X), (Y - \mathbb{E}[Y])/\text{var}^{1/2}(Y))$ also satisfies Assumption 5. But non linear transformations of (X, Y) do not necessarily satisfy the assumption. In particular, the usual (in extreme value theory) transformation of X and Y to random variables with prescribed marginal distributions, is not always possible, as investigated thoroughly in (Heffernan and Resnick, 2007, Section 7). It is never possible in the cases studied in Section 2 where the joint limiting distribution is a product measure. Consequently, we do not make any specific assumption on the marginal distributions of X and Y .

Obviously, the functions a and m are defined up to asymptotic equivalence, i.e. if m' and a' satisfy

$$\lim_{x \rightarrow \infty} \frac{a'(x)}{a(x)} = 1, \quad \lim_{x \rightarrow \infty} \frac{m(x) - m'(x)}{a(x)} = 0,$$

then the measure ν'_n defined as ν_n but with a' and m' instead of a and m converges vaguely to the same limit measure ν . Beyond this trivial remark, the following result summarizes (Heffernan and Resnick, 2007, Proposition 1 and 2) and contains most of what can be inferred from Assumption 5.

Proposition 5. *Under Assumption 5, there exists $\zeta \in \mathbb{R}$ such that the function $a \circ b$ is regularly varying at infinity with index ζ and the function m satisfies*

$$\lim_{t \rightarrow \infty} \frac{m \circ b(tx) - m \circ b(t)}{a \circ b(t)} = J_\zeta(x),$$

with $J_\zeta(x) = (x^\zeta - 1)/\zeta$ if $\zeta \neq 0$ and $J_0(x) = c \log(x)$ for some $c \in \mathbb{R}$, and the convergence is locally uniform on $(0, \infty)$.

Assumption 5 alone is not sufficient for statistical purposes (which typically require bounded convergence arguments), since it has no implication

on the joint moments of X of Y . Therefore the following assumption is introduced.

Assumption 6. *There exists $\zeta^* > 0$, $\kappa^* > 0$ such that*

$$\lim_{n \rightarrow \infty} \int_0^\infty \int_{-\infty}^\infty x^{\zeta^*} |y|^{\kappa^*} \nu_n(dx, dy) = \int_0^\infty \int_{-\infty}^\infty x^{\zeta^*} |y|^{\kappa^*} \nu(dx, dy) . \quad (16)$$

Remark 6. Assumption 6 is satisfied for most of the models studied in section 2. It is shown in Lemma 21 in the Appendix that Assumption 6 holds under the Assumptions of Theorem 3, in the particular case where the density of the angular variable T is bounded, which is the case considered in the examples of Section 3.

Remark 7. Assumption 5 actually implies the convergence (16) with $\kappa^* = 0$ and $\zeta^* < 1/\gamma$, since in that case it is simply a one-dimensional result. See Lemma 16 for a proof.

Remark 8. In applications, it will be assumed that $\kappa^* \geq 2$. The function a and the limiting measure ν are defined up to a change of scale, thus, without loss of generality, we assume henceforth that

$$\int_0^\infty \int_{-\infty}^\infty y^2 \nu(dx, dy) = \int_{-\infty}^\infty y^2 \Psi(dy) = 1 .$$

Remark 9. Condition (16) can be equivalently written as the convergence of conditional moments. For all $\zeta \leq \zeta^*$ and $\kappa \leq \kappa^*$, it holds that

$$\lim_{t \rightarrow \infty} \frac{\mathbb{E}[(X - t)^\zeta |Y - m(t)|^\kappa | X > t]}{\psi^\zeta(t) a^\kappa(t)} = \int_0^\infty \int_{-\infty}^\infty x^\zeta |y|^\kappa \nu(dx, dy) . \quad (17)$$

More generally, for any function g such that $|g(x, y)| \leq Cx^\zeta |y|^\kappa$, it holds that

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathbb{E}[g\{(X - t)/\psi(t), (Y - m(t))/a(t)\} | X > t] \\ = \int_0^\infty \int_{-\infty}^\infty g(x, y) \nu(dx, dy) . \end{aligned} \quad (18)$$

The following result will be needed for statistical purposes, and illustrates the use of the moment Assumption 6. For a sequence (X_i, Y_i) , $1 \leq i \leq n$, let $X_{(n:i)}$ denote the i -th order statistics and $Y_{[n:i]}$ denote its concomitant.

Proposition 6. *If Assumption 5 holds, then for any intermediate sequence k , the random measure $\hat{\nu}_n$ defined by*

$$\hat{\nu}_n = \frac{1}{k} \sum_{k=1}^k \delta \left(\frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})}, \frac{Y_{[n:n-i+1]} - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} \right) \quad (19)$$

converges weakly to ν . For any measurable function g such that $|g(x)| \leq Cx^\zeta$ with $\zeta < 1/\gamma$,

$$\frac{1}{k} \sum_{i=1}^k g \left(\frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})} \right) \rightarrow_P \int_0^\infty g(x) G_\gamma(dx). \quad (20)$$

If moreover Assumption 6 holds, then for any intermediate sequence k and any measurable function g such that $|g(x, y)| \leq Cx^{\zeta^} |y|^{\kappa^*}$,*

$$\begin{aligned} \frac{1}{k} \sum_{i=1}^k g \left(\frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})}, \frac{Y_{[n:n-i+1]} - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} \right) \\ \rightarrow_P \int_0^\infty \int_{-\infty}^\infty g(x, y) \nu(dx, dy). \end{aligned} \quad (21)$$

Remark 10. The convergence (20) is a particular case of (21), but it holds under Assumption 5 only, for the reason already mentioned in Remark 7.

For historical interest, we can also mention the following consequence of Assumption 5, first stated in (Eddy and Gale, 1981, Theorem 6.1) in a restricted case of spherical distributions.

Proposition 7. *Under Assumption 5, $\{Y_{[n:n]} - m(X_{(n:n)})\}/a(X_{(n:n)})$ converges weakly to Ψ . If moreover ν is a product measure, then $\{Y_{[n:n]} - m(X_{(n:n)})\}/a(X_{(n:n)})$ is asymptotically independent of $X_{(n:n)}$.*

Let us finally mention that Davydov and Egorov (2000) obtained functional limit theorems for sums of concomitants corresponding to a number k of order statistics such that $k/n \rightarrow 0$. Their problem differs widely from ours. Their assumptions on the joint distribution of the random pairs are much weaker than Assumption 5, but their results are of a very different nature and it does not seem possible to use them to derive Proposition 6 for instance.

3.2 Nonparametric estimation of a , ψ , m and K

In order to estimate nonparametrically the limiting distribution K , we first need nonparametric estimators of the quantities $\psi(X_{(n:n-k)})$, $m(X_{(n:n-k)})$ and $a(X_{(n:n-k)})$, based on i.i.d. observations $(X_1, Y_1), \dots, (X_n, Y_n)$ with k an intermediate sequence, i.e. such that $k \rightarrow \infty$ and $k/n \rightarrow 0$.

The estimation of $\psi(X_{(n:n-k)})$ is a well known estimation issue, see e.g. Beirlant et al. (2004) or De Haan and Ferreira (2006). If the extreme value index γ of X is less than 1, then ψ can be estimated as the mean residual life. Let $\hat{\gamma}$ be a consistent estimator of γ (see e.g. Beirlant et al. (2004)) and define

$$\hat{\psi}(X_{(n:n-k)}) = \frac{1 - \hat{\gamma}}{k} \sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}. \quad (22)$$

It follows straightforwardly from Proposition 6 that $\hat{\psi}(X_{(n:n-k)})/\psi(X_{(n:n-k)})$ converges weakly to $(1 - \gamma) \int_0^\infty x G_\gamma(dx) = 1$.

If it is assumed (as in Section 3.3 below) that $\gamma = 0$, then the estimator above can be suitably modified:

$$\hat{\psi}(X_{(n:n-k)}) = \frac{1}{k} \sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}. \quad (23)$$

In order to estimate m , define

$$\hat{m}(X_{(n:n-k)}) = \frac{\sum_{i=1}^k Y_{[n:n-i+1]} \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}{\sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}. \quad (24)$$

Proposition 8. *If Assumption 5 holds and Assumption 6 holds with $\zeta^* \geq 1$ and $\kappa^* \geq 1$, then, for any intermediate sequence k , it holds that*

$$\frac{\hat{m}(X_{(n:n-k)}) - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} \rightarrow_P \mu,$$

where $\mu = \int_0^\infty \int_{-\infty}^\infty xy \nu(dx, dy)$.

If moreover $m(x) = \rho x$ and if $\mu = 0$ or $a(x) = o(x)$ then $\hat{m}(X_{(n:n-k)})/X_{(n:n-k)}$ is a consistent estimator of ρ .

Proof. Write

$$\frac{\hat{m}(X_{(n:n-k)}) - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} = \frac{S_n}{T_n},$$

with

$$S_n = \frac{1}{k} \sum_{i=1}^k \frac{Y_{[n:n-i+1]} - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} \frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})}$$

$$T_n = \frac{1}{k} \sum_{i=1}^k \frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})}.$$

Proposition 6 implies that S_n converges weakly to μ and that T_n converges weakly to 1. \square

Remark 11. A sufficient condition for $\mu = 0$ is the symmetry of the measure ν with respect to the second variable. This happens in particular if ν is a product measure, and the distribution Ψ is symmetric.

We now estimate $a(X_{(n:n-k)})$. Many estimators can be defined, each needing an ad hoc moment assumption. The one we have chosen needs $\kappa^* \geq 2$ in Assumption 6. Another estimator could be defined under a weaker assumption, but the present choice is convenient for the estimation of the limiting distribution Ψ . Denote

$$\hat{a}(X_{(n:n-k)}) = \left\{ \frac{1}{k} \sum_{i=1}^k \{Y_{[n:n-i+1]} - \hat{m}(X_{(n:n-k)})\}^2 \right\}^{1/2}.$$

Proposition 9. *If Assumption 5 holds and Assumption 6 holds with $\zeta^* \geq 1$ and $\kappa^* \geq 2$, and if $\mu = 0$, then, for any intermediate sequence k , it holds that*

$$\hat{a}(X_{(n:n-k)})/a(X_{(n:n-k)}) \rightarrow_P \left\{ \int_0^\infty \int_{-\infty}^\infty y^2 \nu(dx, dy) \right\}^{1/2} = 1.$$

Proof. The proof straightforwardly follows from Proposition 6. \square

Remark 12. If $\mu \neq 0$, then $\hat{a}(X_{(n:n-k)})/a(X_{(n:n-k)}) \rightarrow_P \tau$, with

$$\tau^2 = 1 + 2\mu \int_{-\infty}^\infty y \Psi(dy) + \mu^2. \quad (25)$$

We can define the nonparametric estimator of the limiting joint distribution K . Define

$$\hat{K}_n(x, y) = \frac{1}{k} \sum_{i=1}^k \mathbb{1}_{\{X_{(n:n-i+1)} \leq X_{(n:n-k)} + \hat{\psi}(X_{(n:n-k)})x\}} \times \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{m}(X_{(n:n-k)}) + \hat{a}(X_{(n:n-k)})y\}}. \quad (26)$$

We can of course define an estimator of the second marginal Ψ of K . Denote

$$\hat{\Psi}_n(z) = \frac{1}{k} \sum_{i=1}^k \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{m}(X_{(n:n-k)}) + \hat{a}(X_{(n:n-k)})z\}}. \quad (27)$$

Theorem 10. *Under Assumptions 5 and 6 with $\gamma < 1$, $\zeta^* \geq 1$ and $\kappa^* \geq 2$, if $\mu = 0$, and if K is continuous, then, for any intermediate sequence k , $\hat{K}_n(x, y)$ converges weakly to $K(x, y)$.*

Remark 13. Obviously, $\hat{\Psi}_n$ converges to Ψ . If $\mu \neq 0$, then $\hat{\Psi}_n(z)$ converges weakly to $\Psi(\mu + \tau z)$, where τ is defined in (25).

Proof. Denote

$$K_n(x, y) = \hat{\nu}_n([0, x] \times (\infty, y]).$$

By Proposition 6, K_n converges weakly to K , locally uniformly. Moreover, $\hat{K}_n(z) = K_n(\hat{x}_n, \hat{y}_n)$, with

$$\hat{x}_n = \frac{\hat{\psi}(X_{(n:n-k)})}{\psi(X_{(n:n-k)})} x, \quad \hat{y}_n = \frac{\hat{a}(X_{(n:n-k)})}{a(X_{(n:n-k)})} y + \frac{\hat{m}(X_{(n:n-k)}) - m(X_{(n:n-k)})}{a(X_{(n:n-k)})}.$$

Propositions 8 and 9 imply that \hat{x}_n and \hat{y}_n converge weakly to x and y , respectively. Thus $\hat{K}_n(x, y)$ converges to $K(x, y)$, by uniform convergence. \square

3.3 Case of a product measure

In this section, we restrict Assumption 5 to the context corresponding to the examples of Section 2, namely, we make the following additional assumption.

Assumption 7.

The function ψ is an auxiliary function satisfying $\lim_{x \rightarrow \infty} \psi(x)/x = 0$, there exists $\rho \in \mathbb{R}$ such that $m(x) = \rho x$ and the measure ν is of the form

$$\nu((x, \infty) \times [-\infty, y]) = e^{-x} \Psi(y), \quad (28)$$

where Ψ is a distribution function on \mathbb{R} .

The condition $\lim_{x \rightarrow \infty} \psi(x)/x = 0$ implies that the extreme value index of X is 0. (Heffernan and Resnick, 2007, Proposition 2) provides a necessary and sufficient condition for ν to be a product measure.

Proposition 11. *The measure ν is a product measure if and only if $a \circ b$ is slowly varying at infinity and*

$$\lim_{t \rightarrow \infty} \frac{b(tx) - b(t)}{a \circ b(t)} = 0. \quad (29)$$

This result and the same techniques as in the proof of Proposition 6 yield the following corollary.

Corollary 12. *Under Assumptions 5 and 7, $\psi(x) = o(a(x))$, and for all $\tau > 0$,*

$$\lim_{x \rightarrow \infty} \mathbb{E}[(X - x)^\tau / a(x)^\tau \mid X > x] = 0.$$

If moreover Assumption 6 holds, then for $\zeta \in [0, \zeta^*]$ and $\kappa \in [0, \kappa^*]$,

$$\lim_{x \rightarrow \infty} \mathbb{E} \left[\frac{(X - x)^\zeta}{\psi(x)^\zeta} \frac{|Y - \rho X|^\kappa}{a(x)^\kappa} \mid X > x \right] = \Gamma(\zeta + 1) \int_{-\infty}^{\infty} |y|^\kappa \Psi(dy),$$

and

$$\frac{\sum_{i=1}^k \{X_{(n:n-i+1)} - X_{(n:n-k)}\}^\zeta |Y_{[n:n-i+1]} - \rho X_{(n:n-i+1)}|^\kappa}{k \psi^\zeta(X_{(n:n-k)}) a^\kappa(X_{(n:n-k)})}$$

converges weakly to $\Gamma(\zeta + 1) \int_{-\infty}^{\infty} |y|^\kappa \Psi(dy)$.

Under Assumption 7, the previous results allow to define new estimators of ρ , a and the marginal distribution Ψ . Define

$$\hat{\rho} = \frac{\sum_{i=1}^k Y_{[n:n-i+1]} \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}{\sum_{i=1}^k X_{(n:n-i+1)} \{X_{(n:n-i+1)} - X_{(n:n-k)}\}}, \quad (30)$$

$$\check{a}(X_{(n:n-k)}) = \left[\frac{1}{k} \sum_{i=1}^k \{Y_{[n:n-i+1]} - \hat{\rho} X_{(n:n-i+1)}\}^2 \right]^{1/2}, \quad (31)$$

$$\check{\Psi}_n(z) = \frac{1}{k} \sum_{i=1}^k \mathbb{1}_{\{Y_{[n:n-i+1]} \leq \hat{\rho} X_{(n:n-i+1)} + \check{a}(X_{(n:n-k)})z\}}. \quad (32)$$

Proposition 13. *Under Assumptions 5, 6 and 7, for any intermediate sequence k , the two ratios $\hat{\rho}/\rho$ and $\check{a}(X_{(n:n-k)})/a(X_{(n:n-k)})$ converge weakly to 1 and $\check{\Psi}_n$ is a consistent estimator of Ψ .*

Remark 14. The proof of this proposition is a straightforward adaptation of the proof of the previous results and is omitted. Under a second order condition, and with an appropriate choice of the sequence k , following the proof of (Einmahl et al., 1993, Theorem 2), it is possible to prove a functional central limit theorem for $\sqrt{k}(\hat{\Psi}_n - \Psi)$. Such a result allows to apply a Kolmogorov-Smirnov type goodness-of-fit test for a given Ψ . We omit this for brevity.

3.4 Semi-parametric estimation of a and ψ

Nonparametric estimators introduced in Subsection 3.2 yield estimates for a, m and ψ evaluated at some observations $X_{(n:n-k)}$. They lead consequently to approximated values for $a(x)$, $m(x)$ and $\psi(x)$ for such x that are in the range of the observations. Note however that in this case, making use of an empirical estimation is natural, since $\theta(x, y) = \mathbb{P}(Y \leq y \mid X > x)$ can then be estimated by

$$\hat{\theta}_{\text{emp}}(x, y) = \frac{1}{k} \sum_{i=1}^n \mathbb{1}_{\{Y_i \leq y\}} \mathbb{1}_{\{X_i > x\}},$$

for $x = X_{(n:n-k)}$, with k suitably large. Recall that the most interesting situation for using the limit distributions of Theorems 1, 2 or 3 is precisely when x is outside the range of the observations, so that an empirical estimate is no more available. In such a situation, a semi-parametric approach will be needed to extrapolate the functions $a(x)$, $m(x)$ and $\psi(x)$ for values x beyond $X_{(n:n)}$. This requires some modeling restrictions. Driven by an analysis of various examples, we make in this Section the following assumptions:

- (i) the distribution function of X has a rapidly varying tail, with auxiliary function ψ of Weibull type, i.e. $\psi(x) = x^{1-\beta}/(c\beta)$, for $c, \beta > 0$;
- (ii) the pair (X, Y) satisfies (1), and the limiting distribution is a product cdf $K(x, y) = (1 - e^{-x})\Phi(y)$, where Φ is the gaussian cdf;
- (iii) $a(x) = \sigma\sqrt{x\psi(x)}$, for some positive σ ;
- (iv) $m(x) = \rho x$, for some real ρ .

A consequence of the previous hypotheses is that

$$\lim_{x \rightarrow \infty} \theta(x, \rho x + \sigma\sqrt{x\psi(x)}z) = \Phi(z),$$

so that $\theta(x, y)$ can be approximated for x large enough by $\Phi\left(\frac{y - \rho x}{\sigma \sqrt{x\psi(x)}}\right)$. Following Abdous et al. (2008), a semi-parametric estimation of $\psi(x)$ is defined by $\hat{\psi}_n(x) = x^{1-\hat{\beta}_n}/(\hat{c}_n \hat{\beta}_n)$, where $\hat{\beta}_n$ and \hat{c}_n are respectively given by

$$\hat{\beta}_n = \frac{k^{-1} \sum_{i=1}^k \log \log(n/i) - \log \log(n/k)}{k^{-1} \sum_{i=1}^k \log(X_{(n:n-i+1)}) - \log(X_{(n:n-k)})}, \quad (33)$$

and

$$\hat{c}_n = \frac{1}{k} \sum_{i=1}^k \frac{\log(n/i)}{X_{(n:n-i+1)}^{\hat{\beta}_n}}, \quad (34)$$

where k is a user chosen threshold. See Beirlant et al. (1996) for details. Then, the parameter σ is estimated via the nonparametric quantities defined in Section 3.2, namely:

$$\hat{\sigma} = \frac{\check{a}(X_{(n:n-k)})}{\{X_{(n:n-k)} \hat{\psi}_n(X_{(n:n-k)})\}^{1/2}}, \quad (35)$$

in terms of $\check{a}(X_{(n:n-k)})$ and $\hat{\psi}(X_{(n:n-k)})$ respectively defined in (31) and (22). Finally, define

$$\hat{\theta}(x, y) = \Phi \left\{ \frac{y - \hat{\rho}x}{\hat{\sigma} \sqrt{x \hat{\psi}_n(x)}} \right\}. \quad (36)$$

Under the hypotheses listed above (which imply Assumptions 5 and 7) and under the technical Assumption 6, Proposition 13 implies that, for any intermediate sequence k , $\hat{\theta}(x, y)$ is a consistent estimator of $\Phi\left(\frac{y - \rho x}{\sigma \sqrt{x\psi(x)}}\right)$.

The estimation of the conditional quantile function $\theta(x, \cdot)^{\leftarrow}$ for large fixed x can also be done, as in the elliptical case by Abdous et al. (2008)). We skip it here to avoid a lengthening of the paper.

4 Numerical illustration

In this section, we perform a small sample simulation study in order to illustrate in different situations the behavior of the estimators proposed in

the previous section. We assume that the hypotheses of Section 3.4 hold, so that the semi-parametric estimation procedure described there can be used. We consider 3 types of distributions, each of them restricted to the positive quadrant for convenience. These laws are: **(a)** an elliptical distribution (the bivariate normal distribution); **(b)** a distribution with radial representation $R(\cos(T), \sin(T))$, where the r.v. T has a non uniform concave density function, and the radial r.v. R has survival function $\bar{H}(t) = e^{-t^2/2}$; **(c)** a mixture of two gaussian distributions, as considered in Example 2, see equation (8), with $\rho = 0.7$, $\tau = -0.4$ and $p = 0.4$. The first situation is an example of the standard elliptical case, for which estimation results already exist (see Abdous et al. (2008)), whereas the other two cases give two different kinds of extensions: in situation **(b)** the density level lines are “asymptotically elliptic” (cf. Figure 2), and in situation **(c)** the level lines are “locally elliptic” (cf. Figure 3).

In each case, 100 samples of size 1000 are simulated. A proportion of 5% of the observations is used, which are the 50 observations with biggest first component. In the case of the mixture of bivariate Gaussian distributions, we only use observations in a cone (determined graphically) where the estimated density level curves can be reasonably assumed to be elliptic. For each sample, the semi-parametric estimate of $\theta(x, y)$ given in (36) is calculated for three values of x corresponding to the theoretical X -quantiles of order $1 - p$, where $p = 10^{-3}, 10^{-5}, 10^{-7}$, and different values of y chosen to form a grid of possible values of the probability $\theta(x, y)$ from 0 to 1. For each fixed x , this leads to an estimated plot of $p \mapsto \theta(x, y_p)$ where y_p is the theoretical conditional quantile of order p of (X, Y) , given that $X > x$. Figure 4 summarizes the quality of these estimations by showing the median, and the 2.5%- and 97.5%-quantiles of the absolute error $\hat{\theta}(x, y_p) - \theta(x, y_p)$ in terms of $p \in [0, 1]$, for the three fixed values of x specified above.

The estimation results are similarly good for cases **(a)** and **(b)**, see the first two rows of Figure 4. A slight positive bias can be seen for the locally elliptic case **(c)**, which becomes smaller as the conditioning event becomes more extreme (see Row 3, from column 1 to column 3). A common pattern that can also be observed is that the variability of the estimator $\hat{\theta}$ does somewhat increase as x increases, from column 1 to column 3.

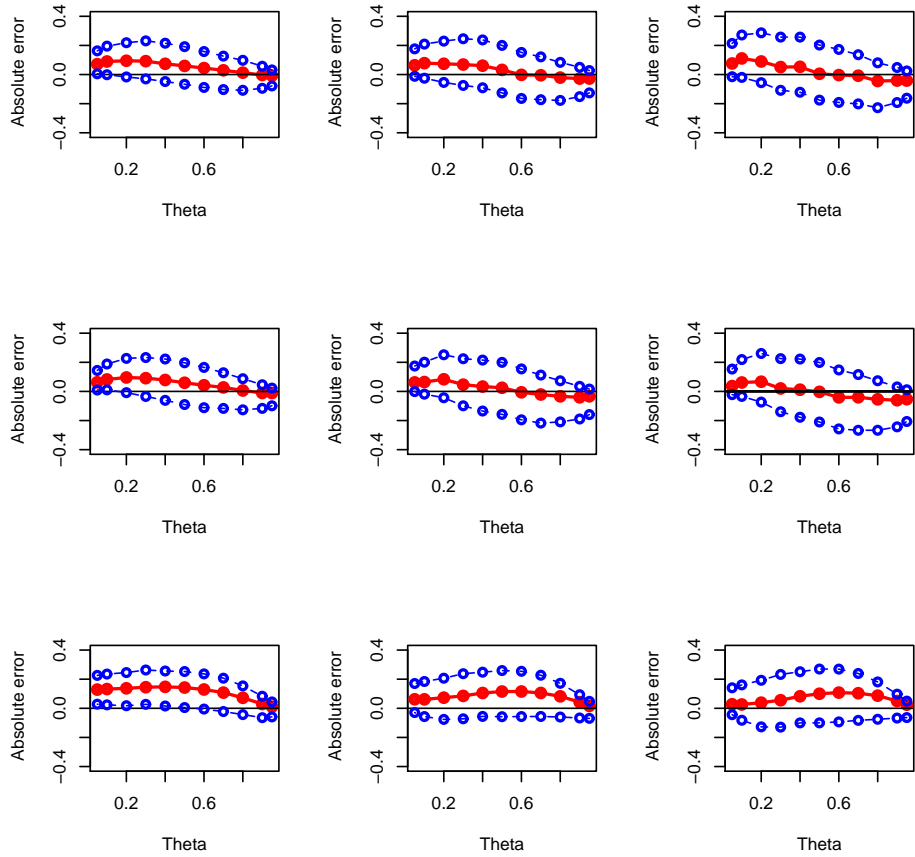


Figure 4: Median, 2.5%- and 97.5%-quantiles of a sample of size 100 of absolute errors $\hat{\theta}(x, y_p) - \theta(x, y_p)$ in terms of $p \in [0, 1]$, where y_p is the theoretical conditional quantile of (X, Y) given that $X > x$. In column 1 (resp. 2, 3) the value x is the theoretical X -quantile of order $1 - 10^{-3}$ (resp. $1 - 10^{-5}$, $1 - 10^{-7}$). The distribution of (X, Y) is **(a)** the bivariate normal distribution on Row 1; **(b)** an asymptotically elliptic distribution on Row 2; **(c)** a locally elliptic distribution on Row 3.

References

- Belkacem Abdous, Anne-Laure Fougères, and Kilani Ghoudi. Extreme behaviour for bivariate elliptical distributions. *Revue Canadienne de Statistiques*, 33(2):1095–1107, 2005.
- Belkacem Abdous, Anne-Laure Fougères, Kilani Ghoudi, and Philippe Soulier. Estimation of bivariate excess probabilities for elliptical models. *Bernoulli*, to appear, 2008.
- Guus Balkema and Paul Embrechts. *High risk scenarios and extremes. A geometric approach*. Zurich Lectures in Advanced Mathematics. Zürich: European Mathematical Society., 2007.
- Philippe Barbe. Approximation of integrals over asymptotic sets with applications to probability and statistics. <http://arxiv.org/abs/math/0312132>, 2003.
- Jan Beirlant, Yuri Goegebeur, Jozef Teugels, and Johan Segers. *Statistics of extremes*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester, 2004.
- Jan Beirlant, Petra Vynckier, and Jozef Teugels. *Practical analysis of extreme values*. Leuven University Press, 1996.
- Simeon M. Berman. *Sojourns and extremes of stochastic processes*. The Wadsworth & Brooks/Cole Statistics/Probability Series. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA, 1992.
- Yu. Davydov and V. Egorov. Functional limit theorems for induced order statistics. *Mathematical Methods of Statistics*, 9(3):297–313, 2000.
- Laurens De Haan and Ana Ferreira. *Extreme value theory. An introduction*. Springer Series in Operations Research and Financial Engineering. New York, NY: Springer., 2006.
- William F. Eddy and James D. Gale. The convex hull of a spherically symmetric sample. *Advances in Applied Probability*, 13(4):751–763, 1981.
- John H. J. Einmahl, Laurens de Haan, and Xin Huang. Estimating a multi-dimensional extreme-value distribution. *Journal of Multivariate Analysis*, 47(1):35–47, 1993.

- Enkelejd Hashorva. Gaussian approximation of conditional elliptic random vectors. *Stoch. Models*, 22(3):441–457, 2006.
- Enkelejd Hashorva, Samuel Kotz, and Alfred Kume. L_p -norm generalised symmetrised Dirichlet distributions. *Albanian Journal of Mathematics*, 1(1):31–56 (electronic), 2007.
- Janet Heffernan and Sidney Resnick. Hidden regular variation and the rank transform. *Advances in Applied Probability*, 37(2):393–414, 2005.
- Janet E. Heffernan and Sidney I. Resnick. Limit laws for random vectors with an extreme component. *Annals of Applied Probability*, 17(2):537–571, 2007.
- Janet E. Heffernan and Jonathan A. Tawn. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B*, 66(3):497–546, 2004.
- Krishanu Maulik and Sidney Resnick. Characterizations and examples of hidden regular variation. *Extremes*, 7(1):31–67 (2005), 2004.
- Sidney Resnick. Hidden regular variation, second order regular variation and asymptotic independence. *Extremes*, 5(4):303–336, 2002.
- Sidney I. Resnick. *Extreme values, regular variation, and point processes*, volume 4 of *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.
- Sidney I. Resnick. *Heavy-Tail Phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

APPENDIX

A Proof of the main results

Proof of Theorem 3. Since u has its maximum at t_0 , there exists $\epsilon > 0$ and $\eta > 0$ such that for all $t \notin [t_0 - \epsilon, t_0 + \epsilon]$, it holds that $u(t) \leq 1 - \eta$. Then,

$$\begin{aligned} \mathbb{P}(X > x ; Y > y) &= \int_0^1 \bar{H} \left(\frac{x}{u(t)} \vee \frac{y}{v(t)} \right) g(t) dt \\ &= \int_{|t-t_0| \leq \epsilon} \bar{H} \left(\frac{x}{u(t)} \vee \frac{y}{v(t)} \right) g(t) dt \\ &\quad + \int_{|t-t_0| > \epsilon} \bar{H} \left(\frac{x}{u(t)} \vee \frac{y}{v(t)} \right) g(t) dt . \end{aligned}$$

Let $r(x)$ denote the last term. Lemma 18 yields that for all $p > 0$,

$$r(x) \leq \bar{H}(x)/(1 - \eta) = o(\{\psi(x)/x\}^p \bar{H}(x)) .$$

Hence $r(x)$ is negligible with respect to the expected order of magnitude of the integral. It suffices to give an equivalent of the integral on the interval $[t_0 - \epsilon, t_0 + \epsilon]$.

By assumption, we can moreover choose ϵ such that the function v/u is continuous and increasing on $[t_0 - \epsilon, t_0 + \epsilon]$, because $\delta < \kappa$, so that $v/u \sim v$ in a neighborhood of t_0 .

If y can be expressed as $y = \rho x + o(x)$, then for large x , it holds that $y/x \in [(v/u)(t_0 - \epsilon), (v/u)(t_0 + \epsilon)]$. If $y/x > \rho = (v/u)(t_0)$, then there exists $t_1 \in [t_0, t_0 + \epsilon]$ such that $(v/u)(t_1) = y/x$. Thus,

$$\begin{aligned} \int_{t_0}^{t_0+\epsilon} \bar{H} \left(\frac{x}{u(t)} \vee \frac{y}{v(t)} \right) g(t) dt \\ = \int_{t_0}^{t_1} \bar{H} \left(\frac{y}{v(t)} \right) g(t) dt + \int_{t_1}^{t_0+\epsilon} \bar{H} \left(\frac{x}{u(t)} \right) g(t) dt . \end{aligned}$$

Let I and J denote the last two integrals, respectively. The successive changes of variables $s = 1/u(t)$ et $s = 1 + z\psi(x)/x$ yield

$$\begin{aligned} J &= \int_{1/u(t_1)}^{1/u(t_0+\epsilon)} \bar{H}(xs) \frac{-(u^\leftarrow)'(1/s)}{s^2} g(u^\leftarrow(1/s)) ds \\ &= \frac{\psi(x)}{x} \int_{x\{1/u(t_1)-1\}/\psi(x)}^{x\{1/u(t_0+\epsilon)-1\}/\psi(x)} \bar{H}(x + \psi(x)z) \\ &\quad \times \frac{-(u^\leftarrow)'(1/\{1 + z\psi(x)/x\})}{1/\{1 + z\psi(x)/x\}^2} g(u^\leftarrow(1/\{1 + z\psi(x)/x\})) . \end{aligned}$$

Let k_1 denote the function defined by

$$k_1(z) = -z(u^\leftarrow)'(1/\{1 + z\psi(x)/x\})g(u^\leftarrow(1/\{1 + z\psi(x)/x\})) .$$

Assumptions 2 and 4 imply that k_1 is regularly varying at zero with index $(1 + \tau)/\kappa$. Denote $\beta = (1 + \tau)/\kappa$. Lemma 17 implies that

$$J \sim k_1(\psi(x)/x) \int_{z_0}^{\infty} e^{-t^{\beta-1}} dt ,$$

where $z_0 = x\{1/u(t_1) - 1\}/\psi(x)$ and if y can be chosen in such a way that z_0 has a finite limit when $x \rightarrow \infty$. We now prove this claim.

Recall that $t_1 = (v/u)^\leftarrow(y/x)$. Assumption 2 implies that if $y/x = \rho + \xi$, then $1 - u(t_1)$ is a regularly varying function of ξ at zero with index κ/δ .

Thus there exists an increasing function h which is regularly varying at zero with index δ/κ such that if $y/x = \rho + h(\psi(x)/x)z$, $z \geq 0$, then $z_0 \sim z^{\kappa/\delta}$. Hence

$$J \sim k_1(\psi(x)/x) \int_{z^{\kappa/\delta}}^{\infty} e^{-t^{\beta-1}} dt , .$$

We now deal with the integral I , still in the case $y/x > \rho$. Noting that $\{y/v(t_1) - x\}/\psi(x) = z_0$, the changes of variables $s = 1/v(t)$ et $s = \{x + \psi(x)z\}/y$ yield

$$\begin{aligned} I &= \int_{1/v(t_1)}^{1/\rho} \bar{H}(ys) \frac{(v^\leftarrow)'(1/s)}{s^2} g(v^\leftarrow(1/s)) ds \\ &= \frac{\psi(x)}{y} \int_{z_0}^{\{y/\rho - x\}/\psi(x)} \bar{H}(x + \psi(x)z) \\ &\quad \times \frac{-(v^\leftarrow)'((y/x)/\{1 + z\psi(x)/x\})}{1/\{1 + z\psi(x)/x\}^2} g(v^\leftarrow((y/x)/\{1 + z\psi(x)/x\})) . \end{aligned}$$

Let the function k_2 be defined

$$k_2(z) = -z(v^\leftarrow)'(\rho/\{1 + z\})f(v^\leftarrow(\rho/\{1 + z\})) .$$

Assumptions 2 and 4 imply that k_2 is regularly varying at zero with index $(1 + \tau)/\delta$. Choosing y as before and applying again Lemma 17 yields

$$I \sim k_2(\psi(x)/x) \int_{z^{\kappa/\delta}}^{\infty} e^{-t^{(1+\tau)/\delta-1}} dt .$$

Since $\delta < \kappa$ by assumption, it holds that $I = o(J)$. The case $y/x < \rho$ can be dealt with similarly and is omitted. \square

Lemma 14. *Let Assumption 5 hold. Let k be an intermediate sequence, i.e. such that $k \rightarrow \infty$ and $k/n \rightarrow 0$. Then*

$$\frac{X_{(n:n-k)} - b(n/k)}{\psi \circ b(n/k)} \xrightarrow{P} 1, \quad (37)$$

$$\frac{\psi(X_{(n:n-k)})}{\psi \circ b(n/k)} \xrightarrow{P} 1, \quad (38)$$

$$\frac{a(X_{(n:n-k)})}{a \circ b(n/k)} \xrightarrow{P} 1, \quad (39)$$

$$\frac{m(X_{(n:n-k)}) - m \circ b(n/k)}{a \circ b(n/k)} = o_P(1). \quad (40)$$

Proof. Denote $Z_i = 1/\bar{F}(X_i)$. As shown in (Resnick, 2007, Proof of Theorem 4.2), $Z_{(n:n-k)}/(n/k)$ converges weakly to 1. By (Resnick, 2007, Theorem 4.1), Assumption 5 implies that the random measure

$$\frac{1}{k} \sum_{i=1}^n \delta_{\{X_i - b(n/k)\}/\psi \circ b(n/k)}$$

converges weakly to a radon measure on $(-\infty, \infty]$. This implies that $\{X_{(n:n-k)} - b(n/k)\}/\psi \circ b(n/k) \xrightarrow{P} 0$. By (De Haan and Ferreira, 2006, Theorem B.2.21), the function $\psi \circ b$ is regularly varying. This property and Lemma 19 imply (38). Denote $\alpha = a \circ b$. Then

$$\frac{a(X_{(n:n-k)})}{a \circ b(n/k)} = \frac{\alpha(Z_{(n:n-k)})}{\alpha(n/k)}.$$

By Proposition 5, α is regularly varying, so (39) holds by Lemma 19. Define now $\beta = m \circ b$. By Proposition 5 and uniform convergence on compact sets, it holds that

$$\begin{aligned} & \frac{m(X_{(n:n-k)}) - m \circ b(n/k)}{a \circ b(n/k)} \\ &= \frac{\beta(Z_{(n:n-k)}) - \beta(n/k)}{\alpha(n/k)} \\ &= \frac{\beta((n/k)(Z_{(n:n-k)}/(n/k))) - \beta(n/k)}{\alpha(n/k)} \xrightarrow{P} K_\zeta(1) = 0, \end{aligned}$$

which proves (40). □

Proof of Proposition 6. Recall the definition of $\hat{\nu}_n$ and define the random measure ν_n by

$$\nu_n = \frac{1}{k} \sum_{i=1}^n \delta_{(\{X_i - b(n/k)\}/\psi \circ b(n/k), \{Y_i - m \circ b(n/k)\}/a \circ b(n/k))} ,$$

$$\hat{\nu}_n = \frac{1}{k} \sum_{k=1}^k \delta \left(\frac{X_{(n:n-i+1)} - X_{(n:n-k)}}{\psi(X_{(n:n-k)})}, \frac{Y_{[n:n-i+1]} - m(X_{(n:n-k)})}{a(X_{(n:n-k)})} \right) .$$

By (Resnick, 2007, Theorem 4.1), Assumption (14) implies that ν_n converges weakly to ν . By continuous mapping, and Lemma 14, $\hat{\nu}_n$ also converges weakly to ν . We now prove (21), and (20) will be proved as a particular case, cf. Remark 10. The convergence of $\hat{\nu}_n$ implies that for all compact set K of $[0, \infty) \times (-\infty, \infty)$ and any bounded function h , it holds that

$$\lim_{n \rightarrow \infty} \iint_K h(x, y) \hat{\nu}_n([x, \infty) \times [y, \infty)) \, dx \, dy$$

$$= \iint_K h(x, y) \nu([x, \infty) \times [y, \infty)) \, dx \, dy .$$

For $M > 0$, define $K = [0, M] \times [0, M]$ and $K^c = [0, \infty) \times (-\infty, \infty) \setminus K$. Let h be a function on $[0, \infty) \times (-\infty, \infty)$ such that $h(x, y) \leq Cx^{\zeta^* - 1}y^{\kappa^* - 1}$. We must prove

$$\limsup_{M \rightarrow \infty} \lim_{n \rightarrow \infty} \iint_{K^c} h(x, y) \hat{\nu}_n([x, \infty) \times [y, \infty)) \, dx \, dy = 0 ,$$

in probability. For $\eta > 0$, define

$$A_\eta = \left\{ \left| \frac{X_{(n:n-k)} - b(n/k)}{\psi \circ b(n/k)} + \left| \frac{\psi(X_{(n:n-k)})}{\psi(b(n/k))} - 1 \right| \right. \right.$$

$$\left. \left. + \left| \frac{a(X_{(n:n-k)})}{a \circ b(n/k)} - 1 \right| + \frac{|m(X_{(n:n-k)}) - m \circ b(n/k)|}{a \circ b(n/k)} \leq \eta \right\} .$$

Denote $R_n = \iint_{K^c} \hat{\nu}_n([x, \infty) \times [y, \infty)) \, dx \, dy$. Then,

$$\mathbb{P}(R_n > \delta) \leq \mathbb{P}(R_n > \delta; A_\eta) + \mathbb{P}(A_\eta^c) .$$

The last term tends to 0 by Lemma 14. Consider now $\hat{\nu}_n \mathbb{1}_{A_\eta}$.

$$\begin{aligned}
& \hat{\nu}_n((x, \infty) \times (y, \infty)) \mathbb{1}_{A_\eta} \\
&= \frac{1}{k} \sum_{k=1}^n \mathbb{1}_{\{X_{(n:n-i+1)} > X_{(n:n-k)} + \psi(X_{(n:n-k)})x\}} \\
&\quad \times \mathbb{1}_{\{Y_{[n:n-i+1]} > m(X_{(n:n-k)}) + a(X_{(n:n-k)})y\}} \mathbb{1}_{A_\eta} \\
&\leq \frac{1}{k} \sum_{k=1}^n \mathbb{1}_{\{X_{(n:n-i+1)} > b(n/k) - \eta\psi(b(n/k)) + \psi(b(n/k))(1-\eta)x\}} \\
&\quad \times \mathbb{1}_{\{Y_{[n:n-i+1]} > m \circ b(n/k) - \eta a \circ b(n/k) + a(b(n/k))(1-\eta)y\}} \mathbb{1}_{A_\eta} \\
&= \nu_n((x(1-\eta) - \eta, \infty) \times (y(1-\eta) - \eta, \infty)) .
\end{aligned}$$

Denote $G_t(x, y) = \mathbb{P}(X > t + \psi(t)x ; Y > m(t) + a(t)y)$ et $t_n = b(n/k)$. Then

$$\mathbb{E}[\hat{\nu}_n((x, \infty) \times (y, \infty)) \mathbb{1}_{A_\eta}] \leq G_{t_n}((x(1-\eta) - \eta, y(1-\eta) - \eta)) .$$

Assumption 5 imply that G_t converges weakly to the function G dfined by $G(x, y) = \nu((x, \infty) \times (y, \infty))$. Assumption 6 (or simply Assumption 5 if $\kappa^* = 0$) and integration by parts imply that for any nonnegative function h such $h(x, y) \leq Cx^{\zeta^*-1}y^{\kappa^*-1}$, it holds that

$$\int_0^\infty \int_{-\infty}^\infty h(x, y) G(x(1-\eta) - \eta, y(1-\eta) - \eta) dx dy < \infty .$$

This implies

$$\begin{aligned}
& \lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \iint_{K^c} h(x, y) \mathbb{E}[\hat{\nu}_n((x, \infty) \times (y, \infty)) \mathbb{1}_{A_\eta}] dx dy \\
&\leq \lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \iint_{K^c} h(x, y) G_{t_n}((x(1-\eta) - \eta, y(1-\eta) - \eta)) dx dy \\
&= \lim_{M \rightarrow \infty} \iint_{K^c} h(x, y) G(x(1-\eta) - \eta, y(1-\eta) - \eta) dx dy = 0 .
\end{aligned}$$

Altogether, we obtain

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(R_n > \delta) = 0 ,$$

wich concludes the proof of Proposition 6. \square

B Lemmas

This following Lemma is a straightforward consequence of Karamata's representation Theorem in the case $\gamma > 0$ and in the case $\gamma = 0$, it can be found in the proof of (Abdous et al., 2005, Theorem 1). We recall the main lines of the proof for the sake of completeness.

Lemma 15. *Let H be a cdf in the domain of attraction of an extreme value distribution with index $\gamma \geq 0$ and with infinite right endpoint. For any $p \in (0, 1/\gamma)$, there exists a constant C such that for all x large enough, and all $t \geq 0$,*

$$\frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \leq C(1+t)^{-p}. \quad (41)$$

Proof. If $\gamma = 0$, the function \bar{H} can be expressed as

$$\bar{H}(x) = c(x) \exp \left\{ - \int_{x_0}^x \frac{ds}{\psi(x)} \right\}$$

where $\lim_{x \rightarrow \infty} c(x) = c \in (0, \infty)$ and $\lim_{x \rightarrow \infty} \psi'(x) = 0$. Thus, for any $\epsilon > 0$ and x large enough, there exists a constant C such that

$$\frac{c(x + \psi(x)t)}{c(x)} \leq C, \quad \frac{\psi(x + \psi(x)t)}{\psi(x)} \leq 1 + \epsilon t.$$

Hence

$$\frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \leq C \exp \left\{ - \int_0^t \frac{ds}{1 + \epsilon s} \right\} = C(1 + \epsilon t)^{-1/\epsilon}.$$

In the case $\gamma > 0$, the auxiliary function is $\psi(x) = \gamma x$ and there exists a slowly varying function L such that $\bar{H}(x) = x^{-1/\gamma} L(x)$, thus

$$\frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} = (1 + \gamma t)^{-1/\gamma} \frac{L(x + \gamma xt)}{L(x)}.$$

Karamata's representation theorem for slowly varying functions yields a bound for the ratio $L(x + \gamma xt)/L(x)$ and the proof is easily concluded. \square

Lemma 16. *Let H be a cdf on $[0, \infty)$ such that there exist $\gamma \geq 0$ and an auxiliary function ψ such that*

$$\lim_{x \rightarrow \infty} \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} = 1 - G_\gamma(t), \quad (42)$$

with $G_\gamma(x) = 1 - (1 + \gamma x)^{-1/\gamma}$. Let $b = (1/\bar{H})^\leftarrow$ and denote $H_n(x) = nH(b(n) + \psi \circ b(n)x)$. Then

$$\lim_{n \rightarrow \infty} \int_0^\infty g(x) H_n(dx) = \int_0^\infty g(x) G_\gamma(dx), \quad (43)$$

for any measurable function g , locally bounded on $[0, \infty)$ and such that $|g(x)| \leq Cx^\beta$ with $\beta < 1/\gamma$.

Proof. For $\eta > 0$ and $A > \eta$, write

$$\begin{aligned} \int_0^\infty g(x) H_n(dx) &= \int_0^\eta g(x) H_n(dx) + \int_\eta^A g(x) H_n(dx) + \int_A^\infty g(x) H_n(dx) \\ &= I_1 + I_2 + I_3. \end{aligned}$$

By uniform convergence on compact sets of $(0, \infty)$, the integral I_2 converges to $\int_\eta^A g(x) G_\gamma(dx)$. Since g is bounded on compact sets, it holds that

$$\sup_{\eta \rightarrow 0} \limsup_{n \rightarrow \infty} I_1 \leq C \lim_{\eta \rightarrow 0} \limsup_{n \rightarrow \infty} H_n(\eta) = C \lim_{\eta \rightarrow 0} G_\gamma(\eta) = 0.$$

By the assumption on g , we get, for some constant c

$$I_3 \leq c \int_A^\infty x^\beta \bar{H}_n(dx) = cA^\beta \bar{H}_n(A) + c \int_A^\infty x^{\beta-1} \bar{H}_n(x) dx.$$

We now use Lemma 15. Thus, for $p \in (\beta, 1/\gamma)$, there exists a constant C such that

$$\limsup_{n \rightarrow \infty} I_3 \leq CA^{\beta-p} + C \int_A^\infty x^{\beta-p-1} dx = O(A^{\beta-p}).$$

Hence $\lim_{A \rightarrow \infty} \limsup_{n \rightarrow \infty} I_3 = 0$. \square

Lemma 17. Let H be a cdf on $[0, \infty)$ that satisfies (42). Let g be a function regularly varying at zero with index $\tau \in (-1, 1/\gamma - 1)$, and locally bounded on $(0, \infty]$. Then

$$\lim_{x \rightarrow \infty} \int_z^\infty \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \frac{g(t\psi(x)/x)}{g(\psi(x)/x)} dt = \int_z^\infty t^\tau \bar{G}_\gamma(t) dt,$$

locally uniformly with respect to $z \geq 0$, with $\bar{G}_\gamma = 1 - G_\gamma$.

Proof of Lemma 17. Denote $\chi(x) = \psi(x)/x$; then $\lim_{x \rightarrow \infty} \chi(x) = 0$. By assumption, $\bar{H}(x + \psi(x)t)/\bar{H}(x)$ converges to $\bar{G}(t)$ and $g(\chi(x)t)/g(x)$ converges to t^τ , and both convergences are uniform on compact sets of $(0, \infty)$. It is thus sufficient to prove that

$$\lim_{A \rightarrow \infty} \limsup_{x \rightarrow \infty} \int_A^\infty \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \frac{g(\chi(x)t)}{g(\chi(x))} dt = 0, \quad (44)$$

$$\lim_{\eta \rightarrow 0} \limsup_{x \rightarrow \infty} \int_0^\eta \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \frac{g(\chi(x)t)}{g(\chi(x))} dt = 0. \quad (45)$$

Let $\epsilon \in (0, 1/\gamma)$ be such that $1/\epsilon - 1 > \tau$. Applying Lemma 15 yields, for large enough x ,

$$\begin{aligned} \int_A^\infty \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} g(\chi(x)t) dt &\leq C \int_A^\infty t^{-1/\epsilon} g(\chi(x)t) dt \\ &= C \chi(x)^{1/\epsilon - 1} \int_{A\chi(x)}^\infty t^{-1/\epsilon} g(t) dt. \end{aligned}$$

Since g is locally bounded on $(0, \infty]$ and $-1/\epsilon + \tau < -1$, Karamata's Theorem implies that there exists a constant C' such that for all A and x , it holds that

$$\int_{A\chi(x)}^\infty t^{-1/\epsilon} g(t) dt \leq C' (\chi(x)A)^{1-1/\epsilon} g(\chi(x)A).$$

Hence

$$\begin{aligned} \limsup_{x \rightarrow \infty} \int_A^\infty \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \frac{g(\chi(x)t)}{g(\chi(x))} dt &\leq CC' A^{1-1/\epsilon} \limsup_{x \rightarrow \infty} \frac{g(\chi(x)A)}{g(\chi(x))} \\ &= CC' A^{1-1/\epsilon + \tau}. \end{aligned}$$

Since $1 - 1/\epsilon + \tau < 0$, this last quantity converges to 0 when A tends to infinity, and this proves (44).

Noting that $\bar{H}(x + \psi(x)t)/\bar{H}(x) \leq 1$ and applying Karamata Theorem, we get that there exists a constant C such that

$$\begin{aligned} \int_0^\eta \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} g(\chi(x)t) dt &\leq \int_0^\eta g(\chi(x)t) dt \\ &= \chi(x)^{-1} \int_0^{\chi(x)\eta} g(t) dt \leq C \eta g(\chi(x)\eta). \end{aligned}$$

Hence

$$\limsup_{x \rightarrow \infty} \int_0^\eta \frac{\bar{H}(x + \psi(x)t)}{\bar{H}(x)} \frac{g(\chi(x)t)}{g(\chi(x))} dt \leq C\eta^{1+\tau}$$

Since $1 + \tau > 0$, this last quantity converges to zero as η tends to zero, and this proves (45). \square

Lemma 18. *Let H be a cdf satisfying Assumption 3. For any $\alpha > 1$ and $p > 0$, there exists a constant C such that for all x , it holds that*

$$\frac{\bar{H}(\alpha x)}{\bar{H}(x)} \leq C(\psi(x)/x)^p .$$

Proof. This follows trivially from (41) by choosing $\epsilon < 1/p$ and by setting $t = (\alpha - 1)x/\psi(x)$. \square

Lemma 19. *Let $\{a_n\}$ and $\{b_n\}$ be two random sequences such that $a_n \rightarrow \infty$ and $a_n/b_n \rightarrow 1$. Let h be regularly varying with index $\zeta \in \mathbb{R}$. Then $h(a_n)/h(b_n) \rightarrow 1$.*

Proof. The result is straightforward if $h(x) = x^\alpha$ for some $\alpha \neq 0$, so it only has to be proved for a slowly varying function h . By Karamata's representation Theorem, a function h slowly varying at infinity can be expressed as

$$h(x) = c(x) \exp \left\{ \int_{x_0}^x \frac{\eta(s)}{s} ds \right\}$$

where $\lim_{x \rightarrow \infty} c(x) = c \in (0, \infty)$ and $\lim_{s \rightarrow \infty} \eta(s) = 0$. Clearly, for the result we want to prove we can assume without loss of generality that the function c is constant. Denote $\epsilon_n = (a_n - b_n)/b_n$. Then $\lim_{n \rightarrow \infty} \epsilon_n = 0$ and

$$\frac{h(a_n)}{h(b_n)} = \exp \left\{ \int_{b_n}^{a_n} \frac{\eta(s)}{s} ds \right\} = \exp \left\{ \int_0^{\epsilon_n} \frac{\eta(b_n + b_n t)}{1+t} dt \right\} .$$

Since $\lim_{n \rightarrow \infty} b_n = \infty$ and $\lim_{s \rightarrow \infty} \eta(s) = 0$, for large enough n , we have $|\eta(b_n + b_n t)| \leq 1$ and

$$\frac{1}{1 + \epsilon_n} = \exp \left\{ - \int_0^{\epsilon_n} \frac{dt}{1+t} \right\} \leq \frac{h(a_n)}{h(b_n)} \leq \exp \left\{ \int_0^{\epsilon_n} \frac{dt}{1+t} \right\} = 1 + \epsilon_n .$$

Since $\lim_{n \rightarrow \infty} \epsilon_n = 0$, this concludes the proof. \square

Lemma 20. *Let ℓ be a continuous function defined on $[0, \infty)$, bounded above and away from zero and with a finite limit at infinity. Define L on $\mathbb{R}_+ \times \mathbb{R}_+$ by $L(x, y) = \ell(y/x)$. Then L belongs to the class \mathcal{L} .*

Proof. Since ℓ is bounded, it suffices to prove that if the limit $\lim_{(\xi, \zeta) \rightarrow \infty} \ell((x + \xi)/(y + \zeta))/\ell(\xi/\zeta)$ exists, then it is equal to 1. Since moreover ℓ is continuous and bounded away from zero, it is enough to consider subsequences and to show that if $\|(\xi_n, \zeta_n)\| \rightarrow \infty$ and if $\lim_{n \rightarrow \infty} \xi_n$ and $\lim_{n \rightarrow \infty} \zeta_n$ both exist, then $\lim_{n \rightarrow \infty} (x + \xi_n)/(y + \zeta_n) = \lim_{n \rightarrow \infty} \xi_n/\zeta_n$. Three cases arise.

(i) If $\lim_{n \rightarrow \infty} \xi_n = \lim_{n \rightarrow \infty} \zeta_n = \infty$, then

$$\frac{x + \xi_n}{y + \zeta_n} = \frac{\xi_n}{\zeta_n} \frac{1 + x/\xi_n}{1 + y/\zeta_n} \sim \xi_n/\zeta_n .$$

(ii) If $\lim_{n \rightarrow \infty} \xi_n < \infty$ et $\lim_{n \rightarrow \infty} \zeta_n = \infty$, then

$$\lim_{n \rightarrow \infty} (x + \xi_n)/(y + \zeta_n) = 0 = \lim_{n \rightarrow \infty} \xi_n/\zeta_n .$$

(iii) If $\lim_{n \rightarrow \infty} \xi_n = \infty$ et $\lim_{n \rightarrow \infty} \zeta_n < \infty$, then

$$\lim_{n \rightarrow \infty} (x + \xi_n)/(y + \zeta_n) = \infty = \lim_{n \rightarrow \infty} \xi_n/\zeta_n .$$

□

Lemma 21. *Assume that $(X, Y) = R(u(T), v(T))$ where R and T are independent, the cdf H of R satisfies Assumption 3, T has a positive bounded density g on $[0, 1]$ and u and v satisfy Assumption 2 with moreover u increasing on $[0, t_0]$, decreasing on $[t_0, 1]$ and (v/u) strictly increasing on $[0, 1]$. Then Assumption 6 holds for any κ^* and ζ^* .*

Proof. Under these assumptions, X has a rapidly varying upper tail and so has finite moments of all order. By Hölder inequality, we need only to prove that for any $q > 0$,

$$\limsup_{x \rightarrow \infty} \mathbb{E}[\{|Y - \rho x|/a(x)\}^q \mid X > x] < \infty ,$$

where $a(x)$ is the normalizing function obtained in Theorem 3. By integration by parts, we have

$$\mathbb{E} \left[\frac{(Y - \rho x)_+^q}{a(x)^q} \mid X > x \right] = q \int_0^\infty \frac{\mathbb{P}(Y > \rho x + a(x)t, X > x)}{\mathbb{P}(X > x)} t^{q-1} dt .$$

Since we have assumed that v/u is increasing, let ϕ denote its left-continuous inverse. Let $z = \phi(\rho + (a(x)/x)t)$. Then

$$\begin{aligned} \mathbb{P}(Y > \rho x + a(x)t, X > x) &= \int_0^1 \mathbb{P}\left(R > \frac{x}{u(s)} \vee \frac{\rho x + a(x)t}{v(s)}\right) g(s) ds \\ &= \int_0^z \mathbb{P}\left(R > \frac{\rho x + a(x)t}{v(s)}\right) g(s) ds + \int_z^1 \mathbb{P}\left(R > \frac{x}{u(s)}\right) g(s) ds \\ &= I_1(x, t) + I_2(x, t). \end{aligned}$$

Since we are only looking for a bound and since we assume that g is bounded, without loss of generality, we can assume that it is constant and equal to 1. We first give a bound for I_2 . If $t > 0$, then $z > t_0$. In that case, we have

$$I_2(x, t) \leq \frac{\psi(x)}{x} \int_{x\{1/u(z)-1\}/\psi(x)}^{\infty} \bar{H}(x + \psi(x)r)(-1/u^-)'(1 + \psi(x)r/x) dr.$$

By Assumption 2, the function $s \rightarrow (1/u^-)'(1 + s)$ is regularly varying at zero with index $1/\kappa - 1$. By (Resnick, 1987, Proposition 0.8), for any $\epsilon > 0$, there exists a constant C such that for large enough x and all $t \geq 1$,

$$\frac{(-1/u^-)'(1 + \psi(x)r/x)}{(-1/u^-)'(1 + \psi(x)/x)} \leq Cr^{1/\kappa + \epsilon - 1}.$$

Besides, we know from the proof of Theorem 3 that

$$\mathbb{P}(X > x) \sim \frac{\psi(x)}{x} (-1/u^-)'(\psi(x)/x) \bar{H}(x).$$

Denote $b_x(t) = x\{1/u(z) - 1\}/\psi(x)$. Applying Lemma 15, we obtain, for any $p > 0$ that

$$\frac{I_2(x, t)}{\mathbb{P}(X > x)} \leq C \int_{b_x(t)}^{\infty} (1+r)^{-p} r^{1/\kappa + \epsilon - 1} dr.$$

We now need a lower bound for $b_x(t)$. Let $k(t) = 1/u \circ \phi(\rho + t) - 1$. Then k is regularly varying at zero with index κ/δ and by definition of the normalizing function $a(x)$, we have $b_x(t) = k(ta(x)/x)/k(a(x)/x)$. Thus, for all $\epsilon > 0$, there exists a constant c such that $b_x(t) \geq ct^{\kappa/\delta + \epsilon}$. Altogether, we obtain, for any $p > 0$ and large enough x ,

$$\frac{I_2(x, t)}{\mathbb{P}(X > x)} \leq C \int_{ct^{\kappa/\delta + \epsilon}}^{\infty} (1+r)^{-p} r^{1/\kappa + \epsilon - 1} dr.$$

Thus $\int_0^{\infty} I_2(x, t)/\mathbb{P}(X > x) dt$ is finite and bounded with respect to x . \square