TAIL ASYMPTOTICS FOR DEPENDENT SUBEXPONENTIAL DIFFERENCES

H. Albrecher [a,b]   S. Asmussen [c]   D. Kortschak [a,c,*]

[a] *Department of Actuarial Science, Faculty of Business and Economics, University of Lausanne, Quartier UNIL-Dorigny, Batiment Extranef, CH-1015 Lausanne, Switzerland.*

[b] *Swiss Finance Institute, Switzerland*

[c] *Department of Mathematical Sciences, Aarhus University, Ny Munkegade, DK-8000 Aarhus C, Denmark*

Abstract

We study the asymptotic behavior of $\mathbb{P}(X - Y > u)$ as $u \to \infty$, where X is subexponential, Y is positive and the random variables X, Y may be dependent. We give criteria under which the subtraction of Y does not change the tail behavior of X. It is also studied under which conditions the comonotonic copula represents the worst-case scenario for the asymptotic behavior in the sense of minimizing the tail of $X - Y$, and an explicit construction of the worst-case copula is provided in the other cases.

1 Introduction

In recent years, there has been quite some progress in understanding the asymptotic effect of dependence on the tail of sums of positive subexponential random variables, see for instance Albrecher et al. [1], Mitra & Resnick [22], Ko & Tang [18], Kortschak & Albrecher [19] and Foss & Richards [14]. In this paper we are interested in the tail asymptotics of differences of random variables, i.e. in

$$\mathbb{P}(X - Y > u)$$

*Supported by the Swiss National Science Foundation Project 200021-124635/1.

for $u \to \infty$, where X is subexponential and the positive random variable Y may have different forms of the tail. If X, Y are independent, this is easy (cf. [6, Lemma 3.2, p. 306]):

$$\mathbb{P}(X - Y > u) \ \sim \ \mathbb{P}(X > u) \tag{1.1}$$

without further conditions. Thus, the problem is dependence.

Since for positive u we have $\mathbb{P}(X - Y > u) = \mathbb{P}(\max(X, 0) - Y > u)$, we can assume w.l.o.g. that X is positive.

There are various areas in which the asymptotics of dependent differences of positive random variables are of interest, for instance random recurrence equations, queueing models and insurance risk models, each in the presence of dependence. In particular, in an insurance context, such a dependent difference can have a natural interpretation as the difference between a claim X and its preceding interarrival time Y, where the random walk structure of the surplus level in the portfolio after a claim occurrence is still preserved (see Albrecher & Teugels [3], Boudreault et al. [10], Asimit & Badescu [4], Li et al. [20] and also Albrecher & Boxma [2] for such and similar dependence structures). In queueing applications similar interpretations are possible.

Asmussen & Biard [7] needed (1.1) for the case where Y is light-tailed. They showed (1.1) essentially when the tail of Y is of smaller magnitude than $\mathrm{e}^{-x^{1/2}}$ and gave a counterexample that (1.1) may not hold with lighter, but still subexponential, tails. The aim of this paper is to provide more general criteria on the dependence between X and Y for the insensitivity to hold and to consider more general distributions of Y. In Section 3 we give a general criterion under which the insensitivity (1.1) holds. Section 4 discusses the role of the mean excess function in this analysis. In Section 5 we discuss the case of light-tailed Y in more detail and provide a substantially simpler construction of a counterexample that $\mathrm{e}^{-x^{1/2}}$ is in fact the critical decay rate of the tail of X, if no dependence structure is specified. This rate is critical in many other contexts and is known as *square-root insensitivity* (e.g. Jelenković et al. [17]). In Section 6 we show (under some regularity conditions) that if there exists a counterexample for the insensitivity (1.1), then the comonotonic copula also provides a counterexample. Yet, the comonotonic copula may not represent the dependence structure that produces the most extreme behavior of $\mathbb{P}(X - Y > u)$. We provide criteria under which the comonotone dependence is indeed the worst case in the sense of minimizing the tail of $X - Y$ and provide an explicit construction of the worst-case copula otherwise. Finally, Section 7 deals with the case of intermediate regularly varying X and relates the present discussion to local limit laws.

2 Preliminaries

In this section we summarize some properties of random variables and classical results that are used later in the paper. For a random variable X with cumulative distribution function $F_X(u)$ denote with $\overline{F}_X(u) = \mathbb{P}(X > u)$ its tail. We say that X is long-tailed if for every constant x

$$\lim_{u\to\infty} \frac{\overline{F}_X(u-x)}{\overline{F}_X(u)} = 1.$$

A nonnegative random variable X is called subexponential if for two independent copies X_1 and X_2 of X it holds that

$$\lim_{u\to\infty} \frac{\mathbb{P}(X_1 + X_2 > u)}{\mathbb{P}(X > u)} = 2.$$

Note that subexponential random variables are long-tailed. A subclass of the subexponential random variables are the regularly varying random variables, for which there exists an index $\alpha > 0$, such that for all $y > 0$

$$\lim_{u\to\infty} \frac{\overline{F}_X(yu)}{\overline{F}_X(u)} = y^{-\alpha}.$$

An extension of regularly varying distributions are distributions that fulfill

$$\lim_{\varepsilon\to 0} \liminf_{u\to\infty} \frac{\mathbb{P}(X > (1+\varepsilon)u)}{\mathbb{P}(X > u)} = 1.$$

This property is known as intermediate regular variation, or also as consistent variation ([11] and [25]). From [13, Th.2.47] it follows that $\overline{F}_X(u)$ is intermediate regularly varying if and only if for any positive function $\delta(u)$ with $\lim_{u\to\infty} \delta(u)/u = 0$

$$\lim_{u\to\infty} \frac{\overline{F}_X(u+\delta(u))}{\overline{F}_X(u)} = 1 \tag{2.1}$$

holds. For a recent survey on heavy-tailed random variables see [13].

Another useful extension of regularly varying distributions is related to extreme value theory (see [24] or [15] for classical references). Let $M_n = \max_{1\le i\le n} X_i$ be the maximum of n independent and identically distributed random variables and assume that there exist constants a_n and b_n and a non-degenerate distribution function $H(x)$ with

$$\lim_{n\to\infty} \mathbb{P}\left((M_n - b_n)a_n \le x\right) = \lim_{n\to\infty} (F_X(a_n x + b_n))^n = H(x). \tag{2.2}$$

Then $H(x)$ is called an extreme value distribution and is known to be of one of the following three types

$$H(x) = \begin{cases} \mathrm{e}^{-x^{-\alpha}}, & x > 0,\ \text{(Fréchet)} \\ \mathrm{e}^{-(-x)^{\alpha}}, & x < 0,\ \text{(Weibull)} \\ \mathrm{e}^{-\mathrm{e}^{-x}}, & x \in \mathbb{R},\ \text{(Gumbel)}, \end{cases}$$

see e.g. [24, Prop. 0.3]. X (or equivalently $\overline{F}_X(x)$) is then said to be in the maximum domain of attraction of the extreme value distribution H). In [24, Ch. 1] it is shown that X is in the maximum domain of attraction of the Fréchet distribution if and only if X is regularly varying. If X is in the maximum domain of attraction of the Weibull distribution then X has a finite right endpoint. Finally, X is in the maximum domain of attraction of the Gumbel distribution if and only if there exists an auxiliary function $e(x)$ such that for all y

$$\lim_{u\to x_r} \frac{\overline{F}_X(u + ye(u))}{\overline{F}_X(u)} = \mathrm{e}^{-y},$$

where $x_r = \inf\{x : F_X(x) = 1\}$ is the right endpoint of X (see also [9, Sec. 3.10]). The function $e(x)$ is unique up to asymptotic equivalence and can be chosen as the mean excess function $e_m(x) = \mathbb{E}(X - x|X > x)$ or – if the density exists – as $1/r(x) = \overline{F}_X(x)/f_X(x)$ (the reciprocal of the hazard rate). The class of distributions in the maximum domain of attraction of the Gumbel distribution contains some subexponential distributions such as lognormal or heavy-tailed Weibull distributions, but also light-tailed distributions like the gamma or the normal.

Since we will consider dependent random variables, it is sometimes useful to decouple the dependence structure from the marginal distributions. Therefore we will use copulas and review now some basic concepts (a standard reference is [23]). A two-dimensional copula $C(u, v)$ is a function that fulfills

- $C(u,0) = 0 = C(0,v)$ for every $u,v \in [0,1]$
- $C(u,1) = u$ and $C(1,v) = v$ for every $u,v \in [0,1]$
- C is 2-increasing, i.e. for every $u_1,u_2,v_1,v_2 \in [0,1]$ with $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2,v_2) - C(u_2,v_1) - C(u_1,v_2) + C(u_1,v_1) \ge 0.$$

Hence, a copula is the joint distribution function of two random variables with uniformly distributed marginal distributions on $[0,1]$. From Sklar's Theorem it follows that there exists a copula C such that the joint distribution function of two random variables X and Y can be expressed as

$$\mathbb{P}(X \le x, Y \le y) = C(F_X(x), F_Y(y)). \tag{2.3}$$

Vice versa, for every copula there exist random variables X and Y with marginal distributions F_X and F_Y such that (2.3) holds. One says that X and Y are dependent according to the copula C. Note that C is invariant under monotone transformations of the marginal distributions. The Frećhet upper bound (or comonotonic) copula $M(u,v) = \min(u,v)$ fulfills

$$C(u,v) \le M(u,v)$$

for all copulas C. Random variables X and Y are said to be comonotonic if they are dependent according to M. For each copula one can define the corresponding survival copula through $\widehat{C}(u,v) = u + v - 1 + C(1-u, 1-v)$, so that

$$\mathbb{P}(X > x, Y > y) = \widehat{C}(\overline{F}_X(x), \overline{F}_Y(y)).$$

Copulas are a useful tool when constructing dependent random variables with given marginal distributions. In this paper we will use the following two methods of contracting copulas (see e.g. [23, Ch. 3]). Denote with $\{J_i\}$ a partition of $[0,1]$ defined here as a collection of closed intervals $J_i = [a_i, b_i]$ that are non-overlapping (except at the endpoints) and $\bigcup J_i = [0,1]$ (one can assign the overlapping points to one of the involved intervals, and then get a partition in the classical sense). For any partition $\{J_i\}$ and for any finite collection of copulas $\{C_i\}$, we define the ordinal sum of $\{C_i\}$ with respect to $\{J_i\}$ as

$$C(u,v) = \begin{cases} a_i + (b_i - a_i) C_i\Big(\frac{u-a_i}{b_i-a_i}, \frac{v-a_i}{b_i-a_i}\Big), & (u,v) \in J_i^2, \\ M(u,v), & \text{otherwise.} \end{cases}$$

Note that when U and V are uniform random variables dependent according to an ordinal sum, then $\mathbb{P}(U \in J_i | V \in J_i) = 1$ and the random vector $(U, V)|(U, V) \in J_i^2$ has uniform marginals on J_i that are dependent according to the copula C_i.

A second type of copulas that will be used in the sequel are so-called straight shuffles of M. Assume we have a copula C, a finite partition $\mathcal{J} = \{J_1, \ldots, J_n\}$ of $[0, 1]$, and a permutation π of $\{1, \ldots, n\}$. The copula C defines a measure on the stripes $J_i \times [0, 1]$ or, equivalently, on stripes of the length $h_i = b_i - a_i$. Now one can reorder these stripes according to the permutation π. So on the stripe $[0, h_{\pi(1)}] \times [0, 1]$ assign the measure which is assigned by C to the stripe $J_{\pi(1)} \times [0, 1]$, on the stripe $[h_{\pi(1)}, h_{\pi(1)} + h_{\pi(2)}] \times [0, 1]$ assign the measure which is assigned to $J_{\pi(2)} \times [0, 1]$ and so on. This defines a new probability measure on $[0, 1] \times [0, 1]$ that (as one easily checks) has again uniform marginal distribution and hence corresponds to a new copula $C_s(\mathcal{J}, \pi)$. We call $C_s(\mathcal{J}, \pi)$ a straight shuffle of M if $C = M$, and then use the notation $M_s(\mathcal{J}, \pi)$. From the discussion after Theorem 3.2.3 in [23] (see also [21]) it follows that every copula can be approximated arbitrary closely by a shuffle w.r.t. supremum norm.

In later sections we will also make use of multivariate extreme value theory, which studies the component-wise maximum of multivariate random variables (the results presented here can e.g. be found in [24, Sec. 5.4] or [15]). Consider the possible limits of

$$\lim_{n \to \infty} \Big[\mathbb{P}\big(X \le a_n x + b_n, Y \le \hat{a}_n y - \hat{b}_n\big) \Big]^n = H(x, y),$$

such that the marginal distributions of H are non-degenerate. Then the marginal distributions of X and Y have to be in the maximum domain of attraction of an extreme value distribution H_X and H_Y, respectively, and there has to exist a copula C_* such that the copula C of X and Y fulfills

$$C_*(u, v) = \lim_{n \to \infty} \left[C(u^{1/n}, v^{1/n}) \right]^n .$$

The copula C is then said to be in the maximum domain of attraction of the extreme value copula C_*. We further have that $H(x, y) = C_*(H_X(x), H_Y(y))$. We now briefly outline the significance of extreme value theory for the purposes needed in later sections. Let for instance $\overline{F}_X$ be regularly varying with index α and $\overline{F}_X(u) \sim \overline{F}_Y(cu)$, then one easily checks that

$$\lim_{t\to\infty} \frac{\mathbb{P}(X > tx \text{ or } Y > tcy)}{\mathbb{P}(X > t)}$$
$$= \lim_{t\to\infty} \frac{1 - C(F_X(tx), F_Y(tcy))}{\overline{F}_X(t)}$$
$$= -\lim_{t\to\infty} \log \left[C\left(\exp\left\{ \overline{F}_X(t) \frac{\log(F_X(tx))}{\overline{F}_X(t)} \right\}, \exp\left\{ \overline{F}_X(t) \frac{\log(F_Y(tcy))}{\overline{F}_X(t)} \right\} \right)^{1/\overline{F}_X(t)} \right]$$
$$= -\log\left(C_*(\mathrm{e}^{-x^{-\alpha}}, \mathrm{e}^{-y^{-\alpha}}) \right).$$

For the last equality we needed that the function on the r.h.s. is continuous. Now for every t the l.h.s. of the equation defines a measure H_t on $[0,\infty]^2$ and the r.h.s. defines a measure H on $[0,\infty]^2 \backslash \{0,0\}$ (the so-called exponential measure). The calculation shows that $H_t \to H$ in the vague sense (i.e. for every set A that is bounded away from $\{0,0\}$ and $H(\partial A) = 0$, we get that $\lim_{t\to\infty} H_t(A) \to H(A)$, where ∂A is the boundary of A). Now for $A = \{(x,y) : x - cy > 1\}$ we have that

$$H_t(A) = \frac{\mathbb{P}(X - Y > t)}{\mathbb{P}(X > t)}.$$

To prove that $H(\partial A) = 0$ is trivial given the special form of H. From the definition of H it is clear that one only needs to consider the case $\alpha = 1$ to get a characterization of H. If we write $x = r\theta$ and $y = r(1-\theta)$, then it follows from [24, Prop. 5.11] that under H the measure μ_r on the radial part is independent of μ_θ on the angular part, μ_r has density r^{-2} and the measure μ_θ satisfies

$$\int_0^1 \theta \, d\mu_\theta = \int_0^1 (1-\theta)\, d\mu_\theta = 1. \tag{2.4}$$

When X is in the maximum domain of attraction of the Gumbel distribution the same steps are applicable.

3 An insensitivity result

From e.g. Foss et al. [13], if a distribution F is long-tailed, this implies that there exists a non-decreasing function δ with $\delta(u) \to \infty$ as $u \to \infty$, such that

$$\overline{F}_X\big(u \pm \delta(u)\big) \sim \overline{F}_X(u) \quad \text{as } u \to \infty\,. \tag{3.1}$$

In the following, we will be interested in choosing $\delta(u)$ as large as possible. The following proposition is essentially equivalent to [14, Proposition 5.1], but we give the proof because of its simplicity and usefulness for later purposes:

Proposition 3.1 *Let $X \geq 0$ be a r.v. with a long-tailed distribution F_X and $Y \geq 0$ a (not necessarily independent) r.v.. Then* (1.1) *holds, provided $\delta(\cdot)$ in* (3.1) *can be chosen with*

$$\mathbb{P}(Y > \delta(u), X > u + \delta(u)) = \mathrm{o}\big(\overline{F}_X(u)\big). \tag{3.2}$$

Proof. Write

$$\mathbb{P}(X - Y > u) \;=\; \mathbb{P}\big(X - Y > u,\, Y \leq \delta(u)\big) \;+\; \mathbb{P}\big(X - Y > u,\, Y > \delta(u)\big)\,.$$

Note that by (3.2)

$$\mathbb{P}\big(X - Y > u,\, Y > \delta(u)\big) \;\leq\; \mathbb{P}\big(X > u + \delta(u), Y > \delta(u)\big) \;=\; \mathrm{o}\big(\overline{F}_X(u)\big)\,.$$

Moreover,

$$\begin{aligned}
\mathbb{P}\big(X - Y > u,\, Y \leq \delta(u)\big) \;&\leq\; \mathbb{P}(X > u) \;=\; \overline{F}_X(u)\,,\\
\mathbb{P}\big(X - Y > u,\, Y \leq \delta(u)\big) \;&\geq\; \mathbb{P}\big(X - \delta(u) > u,\, Y \leq \delta(u)\big)\\
&=\; \mathbb{P}\big(X - \delta(u) > u\big) - \mathbb{P}\big(X - \delta(u) > u,\, Y > \delta(u)\big)\\
&\sim\; \overline{F}_X(u) - \mathrm{o}\big(\overline{F}_X(u)\big)\,.
\end{aligned}$$

Putting these estimates together completes the proof. □

Example 3.2 If X and Y are dependent according to a copula C that is negative quadrant dependent (i.e. $C(u,v) \leq uv$ for $0 \leq u,v \leq 1$) and X is long-tailed, then the assumptions of Proposition 3.1 are fulfilled, in particular

$$\mathbb{P}(Y > \delta(u), X > u + \delta(u)) \leq \mathbb{P}(Y > \delta(u))\mathbb{P}(X > u + \delta(u)) = \mathrm{o}\big(\overline{F}_X(u)\big).$$

Hence (1.1) holds. Note that this criterion does not involve any assumption on the distribution of Y. In terms of the survival copula, a sufficient criterion is $\widehat{C}(u,v) \leq uh(v)$ with $h(v) \to 0$. In terms of distribution functions, this means that for all $x, y \geq 0$

$$\mathbb{P}(X > x, Y > y) \leq \mathbb{P}(X > x)\, h(\mathbb{P}(Y > y))$$

holds. □

Example 3.3 More generally, one can formulate a criterion in terms of stochastic ordering: whenever the pair (X^1, Y^1) fulfills the condition (3.2), then every pair (X^2, Y^2) with the same marginal distributions that is dominated in concordance order (i.e. $\mathbb{P}(X^1 > x, Y^1 > y) \geq \mathbb{P}(X^2 > x, Y^2 > y)$ for all $x > x_0, y > y_0$) also fulfills (3.2). □

4 The role of the mean excess function

Assume that X is regularly varying or in the maximum domain of attraction of the Gumbel distribution with mean excess function $e_m(u)$. Then $\delta(u)$ in (3.1) can be any function with $\delta(u) \to \infty$ and

$$\lim_{u\to\infty} \frac{\delta(u)}{e_m(u)} = 0. \tag{4.1}$$

In a more general setting assume that there exists a function $e(u)$ with

$$\liminf_{u\to\infty} \frac{\mathbb{P}(X - \varepsilon e(u) > u)}{\mathbb{P}(X > u)} < 1$$

for some $\varepsilon > 0$ and

$$\lim_{\varepsilon\to 0} \liminf_{u\to\infty} \frac{\mathbb{P}(X - \varepsilon e(u) > u)}{\mathbb{P}(X > u)} = 1.$$

Then if

$$\lim_{\varepsilon\to 0} \limsup_{u\to\infty} \frac{\mathbb{P}(Y > \varepsilon e(u))}{\mathbb{P}(X > u)} = 0$$

we get by Proposition 3.1 that $\mathbb{P}(X - Y > u) \sim \mathbb{P}(X > u)$.

As we have seen above, for regularly varying distributions or distributions in the maximum domain of attraction of the Gumbel distribution one can choose $e_m(u)$ as the mean excess function (or the reciprocal of the hazard rate $r(u)$). The following result provides another criterion on the distribution of X such that we can still use the mean excess function in (4.1).

Lemma 4.1 *Assume that X is long-tailed with*

$$\overline{F}_X(x) = c(x)\, \exp\Big\{-\int_0^x r^*(t)dt\Big\},$$

where $\lim_{u\to\infty} c(u) = c$, $0 < c < \infty$ *and* $\lim_{u\to\infty} r^*(u) = 0$. *Assume further that there exists an* $\varepsilon_0 > 0$ *such that, uniformly in* $0 < t < \varepsilon_0$,

$$\liminf_{u\to\infty} \frac{r^*\left(u + \frac{t}{r^*(u)}\right)}{r^*(u)} = c_l > 0, \quad \limsup_{u\to\infty} \frac{r^*\left(u + \frac{t}{r^*(u)}\right)}{r^*(u)} = c_u < \infty.$$

Then

$$\limsup_{u\to\infty} \frac{\mathbb{P}\left(X - \varepsilon\frac{1}{r^*(u)} > u\right)}{\mathbb{P}(X > u)} < 1, \quad \lim_{\varepsilon\to 0} \liminf_{u\to\infty} \frac{\mathbb{P}\left(X - \varepsilon\frac{1}{r^*(u)} > u\right)}{\mathbb{P}(X > u)} = 1.$$

Remark 4.2 Note that for an X that fulfills the conditions of Lemma 4.1, the mean excess function $e_m(u)$ satisfies

$$\lim_{u\to\infty} r^*(u) e_m(u) = 1.$$

Proof. We have that

$$\frac{\mathbb{P}\left(X - \varepsilon\frac{1}{r^*(u)} > u\right)}{\mathbb{P}(X > u)} \sim \exp\left(-\int_u^{u+\frac{\varepsilon}{r^*(u)}} r^*(t)dt\right) = \exp\left(-\int_0^{\varepsilon} \frac{r^*\left(u + \frac{t}{r^*(u)}\right)}{r^*(u)} dt\right)$$
$$\lesssim \exp\left(-c_l \int_0^{\varepsilon} dt\right) = \mathrm{e}^{-c_l\varepsilon} < 1$$

(here $f(u) \lesssim g(u)$ means $\limsup_{u\to\infty} f(u)/g(u) \le 1$). Furthermore,

$$\frac{\mathbb{P}\left(X - \varepsilon\frac{1}{r(u)} > u\right)}{\mathbb{P}(X > u)} \sim \exp\left(-\int_0^{\varepsilon} \frac{r^*\left(u + \frac{t}{r^*(u)}\right)}{r^*(u)} dt\right)$$
$$\gtrsim \exp\left(-c_u \int_0^{\varepsilon} dt\right) = \mathrm{e}^{-c_u\varepsilon},$$

from which the result follows. □

Remark 4.3 An example for which the conditions of Lemma 4.1 are not fulfilled is

$$\overline{F}_X(x) = \frac{1}{\log(x)} \text{ for } x \ge e.$$

5 Light-tailed Y

It may be instructive to replace (3.2) by the stronger condition

$$\mathbb{P}\big(Y > \delta(u)\big) = \mathrm{o}\big(\overline{F}_X(u)\big), \tag{5.1}$$

which is now a criterion on the marginal distribution of Y and its comparison to the marginal distribution of X. This gives rise to the following question: If Y is a light-tailed r.v. (i.e. $P(Y > u) = \mathrm{o}(e^{-gu})$ for some $g > 0$), for which long-tailed r.v. X does (1.1) hold across all dependence structures? In this case, condition (5.1) turns into

$$\mathrm{e}^{-g\delta(u)} = \mathrm{o}\big(\overline{F}_X(u)\big),$$

which holds for F_X regularly varying (take $\delta(x) = c\log x$ with c sufficiently large), the lognormal distribution ($\delta(x) = x/\log^2 x$) and the heavy-tailed Weibull with $\overline{F}_X(x) = \mathrm{e}^{-x^\beta}$ with $\beta < 1/2$ ($\delta(x) = x^{1-\beta^*}$ with $\beta < \beta^* < 1$). Thus, the condition covers most standard heavy-tailed distributions except the ones closest to the light-tailed case. Since with independent X, Y and X subexponential, X and $X - Y$ always have the same tail (as discussed in Section 1), one could believe that the condition is just technical. However, it seems to have been observed before that this is not the case, even if we cannot readily provide a precise reference. A counterexample is in Asmussen & Biard [7], and an even simpler construction goes as follows:

Example 5.1 Assume $\mathbb{P}(X > u) \sim \mathrm{e}^{-u^\beta}$ with $0 < \beta < 1$ and let $Y = X^\beta$. Then $\mathbb{P}(Y > u) \sim \mathrm{e}^{-u}$ and hence Y is light tailed. Now

$$\begin{aligned}\mathbb{P}(X - Y > u) &= \mathbb{P}(X > u + X^\beta) \;\leq\; \mathbb{P}(X > u + u^\beta) \;\sim\; \exp\{-(u + u^\beta)^\beta\} \\ &= \exp\{-u^\beta(1 + u^{\beta-1})^\beta\} \;\sim\; \exp\{-u^\beta - \beta u^{2\beta-1}\}\,.\end{aligned}$$

Here $\exp\{-\beta u^{2\beta-1}\} = \mathrm{o}(1)$ if and only if $\beta > 1/2$. □

This counterexample (as well as the one in Asmussen & Biard [7]) involve a comonotonic copula. It is natural to ask whether the comonotonic copula always minimizes the tail of $X - Y$. This is the topic of the next section.

6 The worst-case copula

We will now show under some regularity conditions that if there exists a counterexample for the insensitivity (1.1) to hold, then also the comonotonic copula provides a counterexample:

Lemma 6.1 *Let X and Y be two positive random variables with distribution functions $F_X(x)$ and $F_Y(x)$, respectively. Define*

$$\overline{\gamma}(u) = \sup\{x|F_Y(x-u) < F_X(x), x \geq u\} - u\,,$$
$$\underline{\gamma}(u) = \inf\{x|F_Y(x-u) \geq F_X(x), x \geq u\} - u.$$

If for some $\alpha > 0$, $c > 0$ and all $k > 1$, $\lim_{u\to\infty} \overline{F}_Y(ku)/\overline{F}_Y(u) \leq ck^{-\alpha}$,

$$\lim_{u\to\infty} \frac{\mathbb{P}(X > u + \overline{\gamma}(u))}{\mathbb{P}(X > u)} = 1 \quad \textit{and} \quad \limsup_{u\to\infty} \frac{\mathbb{P}(Y > \overline{\gamma}(u))}{\mathbb{P}(X > u)} < \infty,$$

then

$$\lim_{u\to\infty} \frac{\mathbb{P}(X - Y > u)}{\mathbb{P}(X > u)} = 1.$$

If

$$\liminf_{u\to\infty} \frac{\mathbb{P}(X > u + \underline{\gamma}(u))}{\mathbb{P}(X > u)} < 1,$$

and X and Y are comonotonic, then

$$\liminf_{u\to\infty} \frac{\mathbb{P}(X - Y > u)}{\mathbb{P}(X > u)} < 1.$$

Proof. At first note that $\mathbb{P}(X - Y > u) \leq \mathbb{P}(X > u)$. We have

$$\begin{aligned}
\mathbb{P}(X - Y > u) &= \int_u^\infty \mathbb{P}(Y \leq x - u|X = x)dF_X(x) \\
&= \int_u^\infty \mathbb{P}(Y \leq x - u|X = x)I_{\{F_Y(x-u)<F_X(x)\}}dF_X(x) \\
&\quad + \int_u^\infty \mathbb{P}(Y \leq x - u|X = x)I_{\{F_Y(x-u)\geq F_X(x)\}}dF_X(x).
\end{aligned}$$

To prove the first statement of the Lemma, note that

$$\begin{aligned}
&\int_u^\infty \mathbb{P}(Y \leq x - u|X = x)I_{\{F_Y(x-u)<F_X(x)\}}dF_X(x) \\
&\leq \int_u^\infty I_{\{F_Y(x-u)<F_X(x)\}}dF_X(x) \\
&\leq \int_u^{u+\overline{\gamma}(u)} dF_X(x) = \mathbb{P}(X > u) - \mathbb{P}(X > u + \overline{\gamma}(u)) = o(\mathbb{P}(X > u)).
\end{aligned}$$

For the second integral we have

$$\begin{aligned}
&\int_u^\infty \mathbb{P}(Y \le x-u|X=x) I_{\{F_Y(x-u)\ge F_X(x)\}} dF_X(x)\\
&\ge \int_{u+k\overline{\gamma}(u)}^\infty \mathbb{P}(Y \le x-u|X=x) dF_X(x)\\
&\ge \int_{u+k\overline{\gamma}(u)}^\infty \mathbb{P}(Y \le k\overline{\gamma}(u)|X=x) dF_X(x)\\
&= \mathbb{P}(X > u+k\overline{\gamma}(u)) - \mathbb{P}(X > u+k\overline{\gamma}(u), Y > k\overline{\gamma}(u))\\
&\ge \mathbb{P}(X > u+k\overline{\gamma}(u)) - \mathbb{P}(Y > k\overline{\gamma}(u)).
\end{aligned}$$

Hence there exists a $c_1 > 0$ that does not depend on k, with

$$\frac{\mathbb{P}(Y > k\overline{\gamma}(u))}{\mathbb{P}(X > u)} = \frac{\mathbb{P}(Y > k\overline{\gamma}(u))}{\mathbb{P}(Y > \overline{\gamma}(u))}\frac{\mathbb{P}(Y > \overline{\gamma}(u))}{\mathbb{P}(X > u)} \le c_1 k^{-\alpha}.$$

Since for x_0 with $F_Y(x_0 - u) < F_X(x_0)$ it follows for every $\varepsilon > 0$ that $F_Y((x_0+\varepsilon) - (u+\varepsilon)) < F_X(x_0+\varepsilon)$, we get that

$$\begin{aligned}
\overline{\gamma}(u+\varepsilon) &= \sup\{x|F_Y(x-(u+\varepsilon)) < F_X(x), x \ge u\} - (u+\varepsilon)\\
&\ge \sup\{x|F_Y(x-u) < F_X(x), x \ge u\} + \varepsilon - (u+\varepsilon) = \overline{\gamma}(u)
\end{aligned}$$

and hence $\overline{\gamma}(u)$ is monotonically increasing. Moreover,

$$\begin{aligned}
\liminf_{u\to\infty} \frac{\mathbb{P}(X > u+k\overline{\gamma}(u))}{\mathbb{P}(X > u)} &= \liminf_{u\to\infty} \prod_{l=1}^k \frac{\mathbb{P}(X > u+l\overline{\gamma}(u))}{\mathbb{P}(X > u+(l-1)\overline{\gamma}(u))}\\
&\ge \liminf_{u\to\infty} \prod_{l=1}^k \frac{\mathbb{P}(X > u+(l-1)\overline{\gamma}(u) + \overline{\gamma}(u+(l-1)\overline{\gamma}(u)))}{\mathbb{P}(X > u+(l-1)\overline{\gamma}(u))} = 1,
\end{aligned}$$

from which the first statement follows. For the second, note that for comonotonic X and Y one has

$$\begin{aligned}
\mathbb{P}(X - Y > u) &= \int_u^\infty \mathbb{P}(Y \le X-u|X=x) dF_X(x)\\
&\le \int_u^\infty I_{\{F_Y(x-u)\ge F_X(x)\}} dF_X(x)\\
&\le \int_{u+\underline{\gamma}(u)}^\infty dF_X(x) = \mathbb{P}(X > u+\underline{\gamma}(u)).
\end{aligned}$$

□

Although Lemma 6.1 shows that comonotonic copulas are natural candidates for counterexamples, this does not tell whether the comonotonic copula represents the worst case, i.e. the copula which minimizes $\mathbb{P}(X-Y>u)$ asymptotically for given marginal distributions. To answer that question, let us first consider the case of regularly varying X. In Proposition 7.1 below it will be shown that if $\overline{F}_Y(u)/\overline{F}_X(u) \to 0$, then all copulas provide the same asymptotic properties. On the other hand, if $F_X(x) \geq F_Y(x)$ for X, Y comonotonic, then $\mathbb{P}(X-Y>u)=0$. Hence assume that there exists a $\hat{c}>0$ with

$$\lim_{u\to\infty} \frac{\overline{F}_Y(u)}{\overline{F}_X(u)} = \hat{c}$$

or, equivalently, that there exists a c such that

$$\lim_{u\to\infty} \frac{\overline{F}_Y(cu)}{\overline{F}_X(u)} = 1.$$

We will study the asymptotic behavior of $X-Y$ under the additional condition that

$$\frac{\mathbb{P}(X>xu, Y>ycu)}{\mathbb{P}(X>u)} \to H(x,y),$$

where $H(x,y)$ is not degenerate. Then by extreme value theory it follows that

$$\frac{\mathbb{P}(X-Y>u)}{\mathbb{P}(X>u)} \to H(\{(x,y)|x-cy>1\}).$$

To understand which H minimizes $H(\{(x,y)|x-cy>1\})$, the index of regular variation α of F_X plays a role. When turning to polar coordinates (where we use the sum of components as norm), H can be written as a product of the measure on the radial and angular part. Then the radial measure has density $\alpha r^{-\alpha-1}$ and condition (2.4) is equivalent to (note that we have performed a change of variables)

$$\int_0^1 \theta^\alpha d\mu(\theta) = \int_0^1 (1-\theta)^\alpha d\mu(\theta) = 1.$$

Further note that

$$H(\{(x,y)|x-cy>1\}) = \int_{\frac{c}{1+c}}^1 (\theta - c(1-\theta))^\alpha \, d\mu(\theta). \tag{6.1}$$

Now we can ask which μ^* minimizes (6.1). Consider discrete measures with $\mu(\theta = \theta_i) = p_i$ for $i = 1, \ldots, d$. Then there exists a $\theta_i > 1/2$ ($p_i > 0$) if and only if there exists a $\theta_j < 1/2$ ($p_j > 0$).

Lemma 6.2 *If the measure μ^* that minimizes* (6.1) *assigns positive mass p_i to a $\theta_i \leq \frac{c}{c+1}$, then*

$$\theta_i = \frac{c}{1+c}.$$

Proof. Assume that the result does not hold. Then w.l.o.g. we can assume that $\theta_1 > 1/2$ and $\theta_2 < c/(c+1)$. Define a new measure μ^{**} with $\hat{\theta}_i = \theta_i$ for $i \neq 2$ and $\hat{p}_i = p_i$ for $i > 2$, together with $\hat{\theta}_2 = c/(1+c)$. To ensure that μ is a measure we need

$$p_1\theta_1^\alpha + p_2\theta_2^\alpha = \hat{p}_1\theta_1^\alpha + \hat{p}_2\left(\frac{c}{1+c}\right)^\alpha,$$

$$p_1(1-\theta_1)^\alpha + p_2(1-\theta_2)^\alpha = \hat{p}_1(1-\theta_1)^\alpha + \hat{p}_2\left(\frac{1}{1+c}\right)^\alpha.$$

It follows that

$$\hat{p}_1 = p_1 + p_2\frac{\left(\theta_2\frac{1+c}{c}\right)^\alpha - ((1-\theta_2)(1+c))^\alpha}{\left(\theta_1\frac{1+c}{c}\right)^\alpha - ((1-\theta_1)(1+c))^\alpha} < p_1,$$

where w.l.o.g. we assumed that p_2 is small enough such that $\hat{p}_1 \geq 0$. Thus

$$\int_{\frac{c}{1+c}}^1 (\theta - c(1-\theta))^\alpha \, d\mu^*(\theta) - \int_{\frac{c}{1+c}}^1 (\theta - c(1-\theta))^\alpha \, d\mu^{**}(\theta)$$
$$= (p_1 - \hat{p}_1)\,(\theta_1 - c(1-\theta_1))^\alpha > 0,$$

which is a contradiction to μ^* minimizing (6.1) . ☐

Theorem 6.3 *Assume that $\alpha < 1$. Then μ^* is concentrated on $\theta_1 = 1$ and $\theta_2 = \frac{c}{1+c}$, with $p_1 = 1 - c^\alpha$ and $p_2 = (1+c)^\alpha$.*

Proof. Assume that μ^* assigns positive measure $p_1 > 0$ to $c/(1+c) < \theta_1 < 1$. Then we can define a new measure μ^{**} which is equivalent to μ^* except that we replace θ_1 by 1 and the corresponding probability p_1 by $\hat{p}_1$. Further we add the mass $\hat{p}_0$ to $c/(1+c)$, so that

$$\hat{p}_1 = p_1\left(\theta_1^\alpha - c^\alpha(1-\theta_1)^\alpha\right) > 0$$
$$\hat{p}_0 = p_1(1-\theta_1)^\alpha(1+c)^\alpha.$$

Furthermore,

$$\begin{aligned}
&\int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^*(\theta) - \int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^{**}(\theta) \\
&= p_1 \left(\theta_1 - c(1-\theta_1)\right)^\alpha - \hat{p}_1 \\
&= p_1 \left((\theta_1 - c(1-\theta_1))^\alpha - (\theta_1^\alpha - c^\alpha (1-\theta_1)^\alpha) \right) > 0,
\end{aligned}$$

from which the result follows. □

Theorem 6.4 *Assume that $\alpha > 1$. Then μ^* is concentrated on $\theta_1 = 1/2$.*

Proof. Assume that μ^* assigns positive measure $p_1 > 0$ to $\theta_1 > 1/2$ and $p_2 > 0$ to $\theta_2 < 1/2$, where we assume w.l.o.g. that

$$p_1 \theta_1^\alpha + p_2 \theta_2^\alpha = p_1 (1-\theta_1)^\alpha + p_2 (1-\theta_2)^\alpha.$$

Define the measure μ^{**} with θ_1 and θ_2 replaced by $1/2$ with probability mass $\hat{p}_1 = 2^\alpha (p_1 \theta_1^\alpha + p_2 \theta_2^\alpha)$. We have to distinguish two cases:
a) $\theta_2 > c/(1+c)$: In this case we have to show that

$$\int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^*(\theta) - \int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^{**}(\theta) \geq 0.$$

The l.h.s. equals

$$\begin{aligned}
&p_1 \left(\theta_1 - c(1-\theta_1)\right)^\alpha + p_2 \left(\theta_2 - c(1-\theta_2)\right)^\alpha \\
&\quad - (1-c)^\alpha (p_1 \theta_1^\alpha + p_2 \theta_2^\alpha) \\
&= p_1 \left(\theta_1 - c(1-\theta_1)\right)^\alpha + p_1 \frac{\theta_1^\alpha - (1-\theta_1)^\alpha}{(1-\theta_2)^\alpha - \theta_2^\alpha} \left(\theta_2 - c(1-\theta_2)\right)^\alpha \\
&\quad - p_1 (1-c)^\alpha \left(\theta_1^\alpha + \theta_2^\alpha \frac{\theta_1^\alpha - (1-\theta_1)^\alpha}{(1-\theta_2)^\alpha - \theta_2^\alpha} \right),
\end{aligned}$$

so that we need to show

$$\frac{\left(1 - c\left(\frac{1}{\theta_1} - 1\right)\right)^\alpha - (1-c)^\alpha}{1 - \left(\frac{1}{\theta_1} - 1\right)^\alpha} \geq \frac{\left(1 - c\left(\frac{1}{\theta_2} - 1\right)\right)^\alpha - (1-c)^\alpha}{1 - \left(\frac{1}{\theta_2} - 1\right)^\alpha} \tag{6.2}$$

(cf. the method outlined in Section 2). Since the function

$$\frac{(1-cx)^\alpha - (1-c)^\alpha}{1 - x^\alpha}$$

is decreasing for $x < 1$ and increasing for $x > 1$, we get that we only have to check (6.2) for $\theta_1 = \theta_2 = 1/2$, which holds since

$$\lim_{x \to 1} \frac{(1-cx)^\alpha - (1-c)^\alpha}{1-x^\alpha} = (1-c)^{\alpha-1}.$$

b) $\theta_2 = c/(1+c)$: In this case we have to show that

$$\begin{aligned}
&\int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^*(\theta) - \int_{\frac{c}{1+c}}^{1} (\theta - c(1-\theta))^\alpha \, d\mu^{**}(\theta) \\
&= p_1 \left(\theta_1 - c(1-\theta_1)\right)^\alpha - (1-c)^\alpha \left(p_1 \theta_1^\alpha + p_2 \left(\frac{c}{1+c} \right)^\alpha \right) \\
&= p_1 \left(\theta_1 - c(1-\theta_1)\right)^\alpha - p_1 (1-c)^\alpha \left(\theta_1^\alpha + c^\alpha \frac{\theta_1^\alpha - (1-\theta_1)^\alpha}{1-c^\alpha} \right) \geq 0.
\end{aligned}$$

This is equivalent to showing that

$$\frac{\left(1 - c\left(\frac{1}{\theta_1} - 1\right)\right)^\alpha - (1-c)^\alpha}{1 - \left(\frac{1}{\theta_1} - 1\right)^\alpha} \geq \frac{(1-c)^\alpha c^\alpha}{1-c^\alpha}.$$

Again the left side is minimized for $\theta_1 = 1/2$ and we have to show that

$$(1-c)^{\alpha-1} \geq \frac{(1-c)^\alpha c^\alpha}{1-c^\alpha},$$

which is true for $0 < c < 1$ and $\alpha > 1$. □

Lemma 6.5 *Let X be in the maximum domain of attraction of the Gumbel distribution with auxiliary function $e(x)$. Further assume that there exists a $0 < c < 1$ with*

$$\lim_{u \to \infty} \frac{\mathbb{P}(Y > cu)}{\mathbb{P}(X > u)} = 1$$

and that the copula of (X, Y) is in the maximum domain of attraction of an extreme value copula. Then the copula that asymptotically minimizes $\mathbb{P}(X - Y > u)$ is the comonotonic copula.

Proof. Again we have

$$\frac{\mathbb{P}(X > u + xe(u), Y > cu + yce(u))}{\mathbb{P}(X > u)} \to H(x, y).$$

Here, $H(x, y) = H^*(\mathrm{e}^x, \mathrm{e}^y)$, where under H^*, $R = x + y$ and $\theta = x/(x + y)$ are independent, R has density r^{-2} and the measure μ of θ satisfies

$$\int_0^1 \theta d\mu(\theta) = \int_0^1 1 - \theta d\mu(\theta) = 1.$$

For $b > 0$ we get that

$$\frac{\mathbb{P}(X - Y > (1 - c)u + e(u), X > u - be(u))}{\mathbb{P}(X > u)} \to H(\{(x, y)|x - cy > 1, x > -b\})$$

with

$$H(\{(x, y)|x - cy > 1, x > -b\}) = \int_0^1 \min\left(\mathrm{e}^{-\frac{1}{1-c}}(1 - \theta)\left(\frac{\theta}{1 - \theta}\right)^{\frac{1}{1-c}}, \mathrm{e}^b\right) d\mu(\theta).$$

If $\mu(1) > 0$ and $N > 0$, then as $u \to \infty$

$$\begin{aligned}
&\frac{\mathbb{P}(X - Y > (1 - c)u + e(u), X > u - be(u))}{\mathbb{P}(X > u)} \\
&\gtrsim \frac{\mathbb{P}(X > u - Ne(u)) - \mathbb{P}(X > u - Ne(u)), Y > cu - (N + 2)e(u))}{\mathbb{P}(X > u)} \\
&\sim \mathrm{e}^N - \int_0^1 \min(\theta \mathrm{e}^N, (1 - \theta)\mathrm{e}^{c^{-1}(N+2)}))d\mu(\theta) \\
&\geq \mathrm{e}^N \mu(1) \to \infty,
\end{aligned}$$

as $N \to \infty$. Hence with $b \to \infty$

$$\lim_{u\to\infty} \frac{\mathbb{P}(X - Y > (1 - c)u + e(u))}{\mathbb{P}(X > u)} \geq \mathrm{e}^{-\frac{1}{1-c}} \int_0^1 \mathrm{e}^{-\frac{1}{1-c}}(1 - \theta)\left(\frac{\theta}{1 - \theta}\right)^{\frac{1}{1-c}} d\mu(\theta). \quad (6.3)$$

Note that for X, Y comonotone we can replace $\geq$ by $=$. Finally we have to find the μ that minimizes (6.3). Again, we only consider μ discrete. W.l.o.g we assume that $\theta_1 > 1/2$ and $\theta_2 < 1/2$ with

$$p_1\theta_1 + p_2\theta_2 = p_1(1-\theta_1) + p_2(1-\theta_2) = \frac{p_1+p_2}{2}$$

and we replace θ_1 and θ_2 with $\theta = 1/2$ and $p = p_1 + p_2$. We have to show that

$$p_1(1-\theta_1)\left(\frac{\theta_1}{1-\theta_1}\right)^{\frac{1}{1-c}} + p_2(1-\theta_2)\left(\frac{\theta_2}{1-\theta_2}\right)^{\frac{1}{1-c}} \geq p_1(1-\theta_1) + p_2(1-\theta_2).$$

Since

$$p_2 = p_1\frac{2\theta_1 - 1}{1-2\theta_2},$$

we need to establish that

$$\frac{1-\theta_1}{2\theta_1-1}\left(\left(1+\frac{2\theta_1-1}{1-\theta_1}\right)^{\frac{1}{1-c}}\right) \geq \frac{1-\theta_2}{2\theta_2-1}\left(\left(1+\frac{2\theta_2-1}{1-\theta_2}\right)^{\frac{1}{1-c}}\right)$$

or for $x_i = \frac{2\theta_i - 1}{1-\theta_i}$

$$\frac{(1+x_1)^{\frac{1}{1-c}} - 1}{x_1} \geq \frac{(1+x_2)^{\frac{1}{1-c}} - 1}{x_2},$$

which holds due to $\frac{1}{1-c} > 1$ and $-1 < x_2 < 0 < x_1$. □

Theorem 6.3 shows that when $X \in \mathcal{R}_{-\alpha}$ with index $\alpha < 1$, then comonotonicity does not minimize $\mathbb{P}(X - Y > u)$ asymptotically. On the other hand, Theorem 6.4 suggests that for $\alpha > 1$ comonotonicity does minimize $\mathbb{P}(X - Y > u)$ asymptotically. However, we now show that this is not the case.

As we want to compare the effect of different copulas on the joint distribution of X and Y for fixed marginals F_X and F_Y, define for every copula C the measure $\mathbb{P}_C$ through

$$\mathbb{P}_C(X \leq x, Y \leq y) = C(F_X(x), F_Y(y)).$$

An equivalent formulation for a comonotonic copula minimizing $\mathbb{P}(X-Y>u)$ asymptotically is that for every copula C

$$\liminf_{u\to\infty}\frac{\mathbb{P}_C(X-Y>u)}{\mathbb{P}_M(X-Y>u)}\geq 1. \tag{6.4}$$

In view of Proposition 7.1 shown in the next section, one can assume that for regularly varying X there exists a counterexample for (6.4) if $\overline{F}_X(x)\approx c\overline{F}_Y(x)$ for some $0<c<1$. Therefore we will choose $F_Y(x)=F_X(2x)$, i.e. $2Y\stackrel{d}{=}X$. Further, let X be in the maximum domain of attraction of an extreme value distribution. We will use the following dependence structure.

Definition 6.6 *For a random variable X with distribution function F_X and auxiliary function $e(u)$, define $u_n=u_{n-1}+2e(2u_{n-1})$ for a $u_1>0$ with $F(u_1)>0$, together with a corresponding partition $(J_i)_{n\geq 1}$ of the interval $[0,1]$ $(n\geq 1)$*

$$\begin{aligned} J_1 &= \big[0, F(2u_1)\big)\\ J_{2n} &= \big[F(2u_n), F(2(u_n+e(2u_n)))\big)\\ J_{2n+1} &= \big[F(2(u_n+e(2u_n))), F(2u_{n+1})\big). \end{aligned}$$

Moreover, define a series $(C_n)_{n\geq 1}$ of copulas with

$$C_{2n}(u,v)=uv \quad \text{and} \quad C_{2n+1}(u,v)=\min(u,v).$$

Finally, define the copula $\overline{C}$ as the ordinal sum of the copulas $(C_n)_{n\geq 1}$ with respect to the partition $(J_i)_{n\geq 1}$.

Remark 6.7 *If $2Y\stackrel{d}{=}X$ and X,Y are dependent according to the copula in Definition 6.6, then for $0\leq Y<u_1$ and $u_n+e(2u)\leq Y<u_{n+1}$, we have that $2Y=X$. Furthermore, for $n\geq 1$*

$$\mathbb{P}(X\leq x|u_n\leq Y<u_n+e(2u_n))=\mathbb{P}(X\leq x|2u_n\leq X<2u_n+2e(2u_n)).$$

Proposition 6.8 *Let X be in the maximum domain of attraction of an extreme value distribution and let its density f_X satisfy*

$$\lim_{u\to\infty}\frac{f_X(u+xe(u))}{f_X(u)}=g(x)=\begin{cases}(1+x)^{-\alpha} & \overline{F}_X(x)\in\mathcal{R}_{-\alpha},\ \alpha>0\\ \mathrm{e}^{-x} & X\in MDA(\Lambda)\end{cases}.$$

Further assume that $2Y \stackrel{d}{=} X$ and that X and Y are dependent according to the copula of Definition 6.6. Then

$$\liminf_{u\to\infty} \frac{\mathbb{P}_{\overline{C}}(X-Y>u)}{\mathbb{P}_M(X-Y>u)} < 1.$$

Proof. W.l.o.g we assume that $e(x)$ is monotone. For every n we have

$$\begin{aligned}\mathbb{P}(X-Y>u_n) = \mathbb{P}(X-Y>u_n, Y\le u_n)\\ +\mathbb{P}(X-Y>u_n, u_n<Y\le u_n+e(2u_n))\\ +\mathbb{P}(X-Y>u_n, u_n+e(2u_n)<Y).\end{aligned}$$

Now one can easily check that

$$\mathbb{P}(X-Y>u_n, Y\le u_n)=0$$

and

$$\mathbb{P}(X-Y>u_n, u_n+e(2u_n)<Y)\le \mathbb{P}(Y>u_n+e(2u_n)).$$

On the other hand,

$$\begin{aligned}&\mathbb{P}(X-Y>u_n, u_n<Y\le u_n+e(2u_n))\\ &=\int_{u_n}^{u_n+e(2u_n)} \mathbb{P}\left(X>u_n+y|2u_n<X\le 2(u_n+e(2u_n))\right) f_Y(y)dy\\ &=e(2(u_n))\int_0^1 \mathbb{P}\left(X>2u_n+ye(2u_n)|2u_n<X\le 2(u_n+e(2u_n))\right)\\ &\qquad f_Y(u_n+ye(2u_n))dy.\end{aligned}$$

Note that

$$\begin{aligned}&\mathbb{P}\left(X>2u_n+ye(2u_n)|2u_n<X\le 2(u_n+e(2u_n))\right)\\ &=\frac{\mathbb{P}(X>2u_n+ye(2u_n))-\mathbb{P}(X>2u_n+e(2u_n))}{\mathbb{P}(X>2u_n)-\mathbb{P}(X>2u_n+e(2u_n))}\\ &\to \frac{g(y)-g(1)}{g(0)-g(1)}<1, \quad y>0\end{aligned}$$

for $n\to\infty$. It follows from

$$\frac{f_Y(u_n+ye(2u_n))}{f_Y(u_n)}=\frac{f_X(2u_n+2ye(2u_n))}{f_X(2u_n)}\to g(2y)$$

that

$$\lim_{n\to\infty} \frac{\mathbb{P}(X-Y>u_n, u_n<Y\le u_n+e(2u_n))}{\mathbb{P}(u_n<Y\le u_n+e(2u_n))} < 1$$

and hence

$$\lim_{n\to\infty} \frac{\mathbb{P}(X-Y>u_n)}{\mathbb{P}(Y>u_n)} < 1.$$

□

Example 6.9 As an illustration, consider $\mathbb{P}(X>x)=\mathbb{P}(2Y>x)=1/x$ with $e(x)=x$ and $u_n=5^n$. Figure 1 depicts the plot of $\frac{\mathbb{P}_{\overline{C}}\left(X-Y>\frac{1}{2}10^x\right)}{\mathbb{P}_M\left(X-Y>\frac{1}{2}10^x\right)}$.

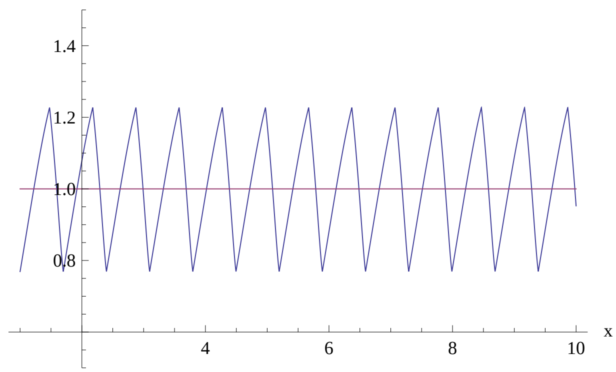

Figure 1: Plot of $\mathbb{P}_{\overline{C}}\left(X-Y>\frac{1}{2}10^x\right)/\mathbb{P}_M\left(X-Y>\frac{1}{2}10^x\right)$

Having seen now that the worst case is not always given by the comonotonic copula, we are now interested in identifying the worst case (given a specific u instead of $u\to\infty$). For that purpose, we will use straight shuffles of M. Since shuffles are dense in the set of copulas we want to find the shuffle that minimizes $\mathbb{P}(X-Y>u)$. For a given F_X, F_Y and u, define

$$g_u(x) = \begin{cases} \inf\{t : F_Y^{-1}(t) \ge F_X^{-1}(x)-u\} & \text{if } F_X^{-1}(x)>u, \\ 0 & \text{otherwise.} \end{cases} \tag{6.5}$$

For uniformly distributed (U_1,U_2) with the same copula C as (X,Y), it is valid that

$$\mathbb{P}(U_2<g_u(U_1)) = \mathbb{P}(X-Y>u).$$

Lemma 6.10 *Let $g(x)$ be an increasing function, such that for all $c \in [-1,1]$ the number of times $g(x) - x - c$ changes sign is finite. Then the shuffle M_s^* that minimizes*

$$\mathbb{P}_{M_s}(U_2 < g(U_1))$$

is of the form $\mathcal{J} = \{[0, x_0], [x_0, 1]\}$ and $\pi = (2, 1)$ for some $0 < x_0 < 1$.

Proof. Let M_s be a shuffle with finite partition $\mathcal{J}$ and permutation π. For $J \in \mathcal{J}$ and $x \in J$, denote by J^π and x^π the interval J (point to which x, respectively) is mapped by the permutation. W.l.o.g we assume that for every $J \in \mathcal{J}$

$$\frac{\mathbb{P}(U_1 \in \{x^\pi : x \in J \ \&\ x < g(x^\pi)\})}{\mathbb{P}(U_1 \in \{x^\pi : x \in J\})} \in \{0, 1\}.$$

Denote with $x_0 = \mathbb{P}_{M_s}(U_2 < g(U_1))$. W.l.o.g we can assume that for every $J \in \mathcal{J}$, $(J \cap [0, x_0]) \in \{\emptyset, J\}$. Further we can split the intervals in the partition $\mathcal{J}$, such that to every interval $J \in \mathcal{J}$ with $\mathbb{P}(U_1 \in \{x^\pi : x \in J \ \&\ x < g(x^\pi)\}) = \mathbb{P}(U_1 \in \{x^\pi : x \in J\})$ we can assign a unique interval $\hat{J}$ with $\hat{J} \cap [0, x_0] = \hat{J}$ and $|J| = |\hat{J}|$. If we change the position of J and $\hat{J}$ in the permutation then $\mathbb{P}(U_2 < g(U_1))$ is the same for both shuffles. Hence we can assume that if $\mathbb{P}(U_1 \in \{x^\pi : x \in J \ \&\ x < g(x^\pi)\}) = \mathbb{P}(U_1 \in \{x^\pi : x \in J\})$, then $J \subset [0, x_0]$. Since $g(x)$ is increasing we can reorder the partitions such that we get the form of M_s^* from which the Lemma follows. □

The worst copula is not unique, as can be seen by the following straightforward result.

Lemma 6.11 *Let $g(x)$ be an increasing function. Let $x_1 = \inf\{x : x \geq g(x)\}$. If $x_1 < 1 - x_0$ for some x_0, then the shuffles $M_s(\{[0, x_0], [x_0, 1]\}, (2, 1))$ and $\hat{M}_s(\{[0, x_1], [x_1, x_1 + x_0], [x_1 + x_0, 1]\}, (1, 3, 2))$ fulfill*

$$\mathbb{P}_{M_s}(U_2 < g(U_1)) \geq \mathbb{P}_{\hat{M}_s}(U_2 < g(U_1)).$$

If $x_1 \geq 1 - x_0$, then

$$\mathbb{P}_{M_s}(U_2 < g(U_1)) \geq \mathbb{P}_{M}(U_2 < g(U_1)).$$

Example 6.12 Let $F_X(x) = 1 - 1/x$, $F_Y(x) = 1 - 1/(2x)$ and $u = 1$. For this case, Figure 2 shows the support of the copula in Lemma 6.10 (bold

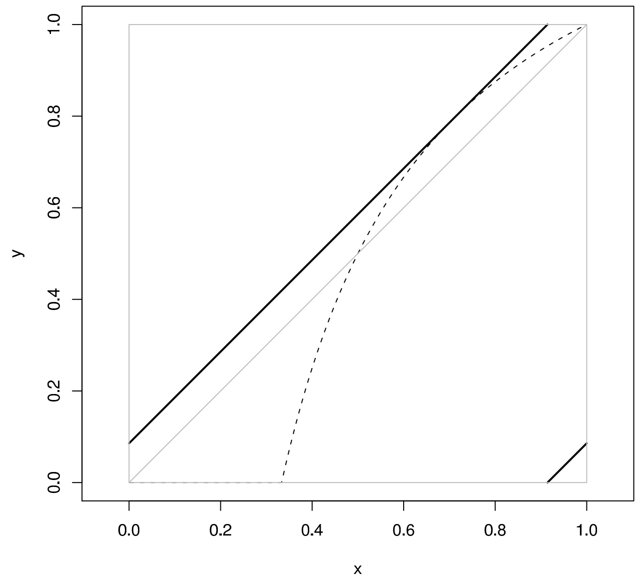

Figure 2: A worst-case copula

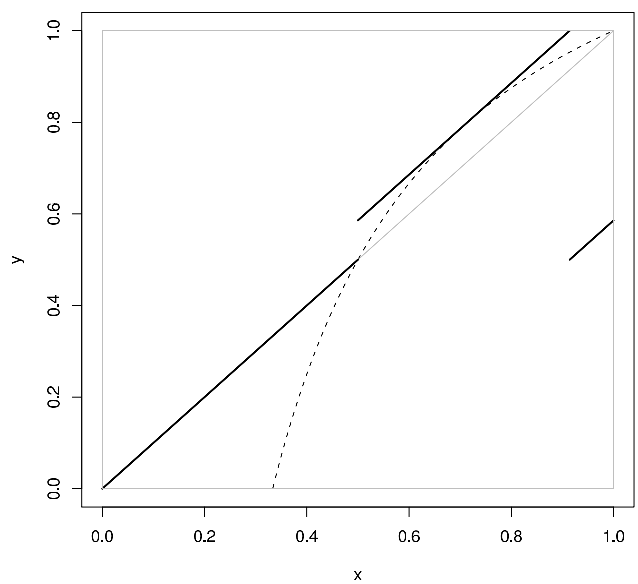

Figure 3: Another worst-case copula

line), where $x_0 \approx 0.086$. In Figure 3, the bold line depicts the support of the copula in Lemma 6.11, where $x_0 \approx 0.086$ and $x_1 = 0.5$. In both plots the dashed line corresponds to the function $g_u(x)$. Here

$$x_0 = x_0^* = \sup_{0 \le x \le 1} g_u(x) - x. \tag{6.6}$$

□

In fact, the choice of $x_0 = x_0^*$ in (6.6) is optimal in general, as can be verified by the following arguments: If $x_0 > x_0^*$, then the line $x + x_0$ corresponding to the interval $[x_0, 1]$ lies above the line $g_u(x)$. Hence we can decrease x_0 to x_0^* so that the line $x + x_0^*$ touches the line $g_u(x)$; certainly $\mathbb{P}_{M_s}(U_2 < g_u(U_1))$ then does not increase. If on the other hand $x_0 < x_0^*$ and x^* is a point with $x_0^* = g_u(x^*) - x^*$, then the monotonicity of $g_u(x)$ implies that the line segment of $x + x_0$ from x^* to $g_u(x^*) - x_0$ lies below $g_u(x)$. Since this line segment has length $g_u(x^*) - x_0 - x^* = x_0^* - x_0$ we see that by using x_0^* instead of x_0 we do not increase the probability of $\mathbb{P}_{M_s}(U_2 < g_u(U_1))$. Further if $x^* > 1/2$ then the line corresponding to the interval $[0, x_0]$ lies below $g_u(x)$. Thus we have proved the following:

Proposition 6.13 *Assume that the conditions of Lemma 6.10 hold and that u is large enough such that x^* with*

$$g_u(x^*) - x^* = \sup_{0 \le x \le 1} g_u(x) - x$$

fulfills $x^* > 1/2$. Then

$$\inf_C \mathbb{P}_C(X - Y > u) = \sup_{0 \le x \le 1} g_u(x) - x.$$

Let us compare this result to the comonotonic copula. To that end, assume that there exists a unique point γ_u such that $g_u(x) - x \le 0$ for $x < \gamma_u$ and $g_u(x) - x > 0$ for $x > \gamma_u$, then $\mathbb{P}_M(X - Y > u) = 1 - \gamma_u$ and

$$\begin{aligned}
&\inf_C \mathbb{P}_C(X - Y > u) \\
&= \mathbb{P}_M(X - Y > u) \sup_{0 \le x \le 1} \frac{g_u(\gamma_u + x(1 - \gamma_u)) - \gamma_u - x(1 - \gamma_u)}{1 - \gamma_u} \\
&= \sup_{0 \le x \le 1} \left(g_u(\gamma_u + x(1 - \gamma(u))) - \gamma_u - x(1 - \gamma_u)\right).
\end{aligned}$$

If the function

$$h_u(x) = \frac{g_u\left(\gamma_u + x(1-\gamma_u)\right) - \gamma_u - x(1-\gamma_u)}{1-\gamma_u}$$

converges for $u \to \infty$ to a function $h_\infty(x)$ with $\sup_{0<x<1} h_\infty(x) = 1$ (i.e. $h_\infty(x) = 1-x$), then for every copula C (6.4) holds. On the other hand, if there exists a sequence u_n with $\lim_{n\to\infty} u_n = \infty$ and $\limsup_{n\to\infty} \sup_{0<x<1} h_{u_n}(x) < 1$ then we can analougously to Proposition 6.8 construct a copula where (6.4) does not hold. The following example shows such a situation where X is Weibull and Y is light tailed.

Example 6.14 Let $F_X(x) = 1 - \mathrm{e}^{-x^\beta}$ $(1/2 < \beta < 1)$ and $F_Z(x) = 1 - \mathrm{e}^{-\frac{(1+\varepsilon)\beta^2}{2\beta-1}x^{2-1/\beta}}$. Define $u_0 = 0$, $u_n = 2^n$ and

$$F_Y(x) = 1 - \mathrm{e}^{-u_n} + \frac{F_Z(x) - F_Z(u_n)}{F_Z(u_{n+1}) - F_Z(u_n)}\left(\mathrm{e}^{-u_n} - \mathrm{e}^{-u_{n+1}}\right), \quad u_n \le x < u_{n+1}.$$

Since for $x > 2$

$$\frac{\overline{F}_Y(x)}{\mathrm{e}^{-x/2}} \le \frac{\overline{F}_y(u_n)}{\mathrm{e}^{-u_{n+1}/2}} = 1$$

we get that Y is light tailed. Further for $u = u_n^{1/\beta} - u_n$ we get that $\gamma_u = (1-e^{-u_n})$ and since $F_Y(x) \le 1-\mathrm{e}^{-x}$ there are no roots of $F_Y(F_X^{-1}(x)-u) = x$ to the left of γ_u. We get that

$$h_u(x) = 1 - x - \frac{\overline{F}_Y((u_n - \log(1-x))^{1/\beta} - u^{1/\beta} + u_n)}{\mathrm{e}^{-u_n}}$$

since for $n \to \infty$

$$(u_n - \log(1-x))^{1/\beta} - u^{1/\beta} + u_n = u_n + (1+o(1))\frac{(-\log(1-x))}{\beta}u_n^{1/\beta-1} \le 2u_n = u_{n+1}.$$

We get that

$$\begin{aligned}
&\frac{\overline{F}_Y((u_n - \log(1-x))^{1/\beta} - u^{1/\beta} + u_n)}{\mathrm{e}^{-u_n}} \\
&= 1 - \frac{\overline{F}_Z\left(u_n + (1+o(1))\frac{(-\log(1-x))}{\beta}u_n^{1/\beta-1}\right) - \overline{F}_Z(u_n)}{\overline{F}_Z(u_{n+1}) - \overline{F}_Z(u_n)}\left(1-e^{-u_n}\right) \\
&\sim \frac{\overline{F}_Z\left(u_n + \frac{(-\log(1-x))}{\beta}u_n^{1/\beta-1}\right)}{\overline{F}_Z(u_n)} \\
&\sim (1-x)^{1+\varepsilon}
\end{aligned}$$

Hence as $n \to \infty$
$$h_{u_n}(x) \to (1-x)(1-(1-x)^{\varepsilon}).$$

7 Intermediate regularly varying X

Proposition 7.1 *If X is intermediate regularly varying and $\overline{F}_Y(u) = o(\overline{F}_X(u))$, then* (1.1) *holds.*

Proof. We have to show that a positive function $\delta(u) = o(u)$ exists that fulfills (5.1) since such a $\delta(u)$ also fulfills (2.1). At first note that for every $c > 0$ there exists a b_c such that

$$\lim_{u\to\infty} \frac{\overline{F}_X(cu)}{\overline{F}_X(u)} \leq b_c.$$

Hence, for every n there exists a $\hat{u}_n$ such that for all $u > \hat{u}_n$

$$\frac{\mathbb{P}(Y > u)}{\mathbb{P}(X > nu)} \leq \frac{1}{n}.$$

Define $u_0 = 0$ and $u_n = \max(n\hat{u}_n, u_{n-1}) + 1$ for $n > 0$. Then for all $u > u_n$

$$\frac{\mathbb{P}(Y > u/n)}{\mathbb{P}(X > u)} \leq \frac{1}{n}.$$

Define
$$\varepsilon(u) = \begin{cases} 1, & u < u_1, \\ \frac{1}{n}, & u_n < u < u_{n+1}. \end{cases}$$

Then for $\delta(u) = \varepsilon(u)u$ we have

$$\lim_{u\to\infty} \frac{\mathbb{P}(Y > \delta(u))}{\mathbb{P}(X > u)} = 0.$$

□

7.1 Approach with local limit laws

Let us now use local limit laws as in Heffernan and Resnick [16] to find the asymptotic behavior of $\mathbb{P}(X - Y > u)$. For that purpose, let either $E = [-\infty, \infty] \times (-\infty, \infty])$ $(e(u)/u \to 0)$ or $E = [-\infty, \infty] \times (-1, \infty])$ $(e(u) = u)$. Further we assume that there exists a measure μ (not equal to zero) for which for every fixed y in $\mathbb{E}$

- $\mu([-\infty, x], (y, \infty])$ is a non-degenerate distribution function in x,
- $\mu([-\infty, x], (y, \infty]) < \infty$, and
- $\displaystyle\lim_{u\to\infty} \frac{\mathbb{P}(Y \leq \beta(u) + x\alpha(u), X > u + ye(u))}{\mathbb{P}(X > u)} = \mu([-\infty, x], (y, \infty])$
  at each continuity point (x, y) of the limit.

Assume that $\alpha(u)/e(u) \to c$ for some constant c, then we have that

$$\lim_{u\to\infty} \frac{\mathbb{P}(X - Y > u - \beta(u))}{\mathbb{P}(X > u)} = \lim_{u\to\infty} \frac{\mathbb{P}\left(\frac{X-u}{e(u)} - \frac{\alpha(u)}{e(u)} \cdot \frac{Y-\beta(u)}{\alpha(u)} > 0, \frac{X-u}{e(u)} > 0\right)}{\mathbb{P}(X > u)}$$
$$= \mu\left(\{(y, x) | x - cy > 0, x > 0\}\right) \leq 1$$

at least if μ is sufficiently continuous. The area we have to measure is depicted in Figure 4.

It follows that

$$\frac{\mathbb{P}(X - Y > u)}{\mathbb{P}(X > u)} \sim \frac{\mathbb{P}(X > u)}{\mathbb{P}(X > u - \beta(u))} \mu\left(\{(y, x) | x - cy > 0, x > 0\}\right).$$

If (1.1) is valid, then we have to assume that $\beta(u)/e(u) \to 0$ and $c = 0$ (i.e. $\alpha(u)/e(u) \to 0$). However note that for every $\varepsilon > 0$

$$\begin{aligned}
&\lim_{u\to\infty} \frac{\mathbb{P}(Y \leq \varepsilon e(u), X > u)}{\mathbb{P}(X > u)} \\
&= \lim_{u\to\infty} \frac{\mathbb{P}\left(Y \leq \beta(u) + \frac{\varepsilon e(u) - \beta(u)}{\alpha(u)} \alpha(u), X > u\right)}{\mathbb{P}(X > u)} \\
&\geq \lim_{u\to\infty} \frac{\mathbb{P}\left(Y \leq \beta(u) + b\alpha(u), X > u\right)}{\mathbb{P}(X > u)} \\
&= \mu([-\infty, b) \times \mu(0, \infty]) \to 1
\end{aligned}$$

as $b \to \infty$. Hence the conditions of Proposition 3.1 are fulfilled, so that we do not need to use local limit law for establishing (1.1).

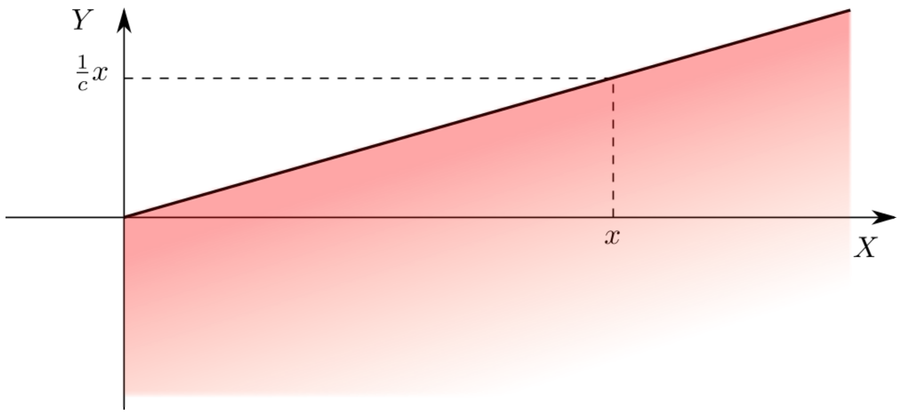

Figure 4: Area to be measured (shaded)

References

[1] H. Albrecher, S. Asmussen & D. Kortschak (2006) Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes* **9**, 107–130.

[2] H. Albrecher & O. J. Boxma (2004) A ruin model with dependence between claim sizes and claim intervals. *Insurance Math. Econom.* **35**, 245–254.

[3] H. Albrecher & J. Teugels (2006) Exponential behavior in the presence of dependence in risk theory. *J. Appl. Prob.* **43**, 257–273.

[4] A. V. Asimit & A. L. Badescu (2010) Extremes on the discounted aggregate claims in a time dependent risk model. *Scand. Actuar. J.*, 93–104.

[5] S. Asmussen (2003) *Applied Probability and Queues* (2nd ed.). Springer-Verlag, New York.

[6] S. Asmussen & H. Albrecher (2010) *Ruin Probabilities* (2nd ed.). World Scientific, Singapore.

[7] S. Asmussen & R. Biard (2011) Ruin probabilities for a regenerative Poisson gap generated risk process. *Europ. Act. J.* **1**, 3–22.

[8] G. Balkema & L. de Haan (1974) Residual life-time at great age. *Ann. Probab.* **2**, 792–804.

[9] N. H. Bingham, C. M. Goldie, and J. L. Teugels. (1989) *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*.2 Cambridge University Press, Cambridge.

[10] M. Boudreault, H. Cossette, D. Landriault & E. Marceau (2006) A risk model with dependence between interclaim arrivals and claim sizes. *Scand. Actuar. J.* 265–285.

[11] D. B. H. Cline (1994) Intermediate regular and Π variation. *Proc. London Math. Soc.* **68**, 594–616.

[12] P. Embrechts, C. Klüppelberg & T. Mikosch (1997) *Modelling Extreme Events for Insurance and Finance*. Springer-Verlag, Berlin.

[13] S. Foss, D. Korshunov & S. Zachary (2011) *An Introduction to Heavy-tailed and Subexponential Distributions*. Springer-Verlag, New York.

[14] S. Foss & A. Richards (2010) On sums of conditionally independent subexponential random variables. *Math. Oper. Res.* **35**, 102–119.

[15] J. Galambos. (1987) *The Asymptotic Theory of Extreme Order Statistics*. Robert E. Krieger Publishing Co. Inc., Melbourne, FL, second edition.

[16] P. Heffernan & S. Resnick (2007) Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.* **17**, 537–571.

[17] P. Jelenković, P. Momcilović & B. Zwart (2004) Reduced load equivalence under subexponentiality. *QUESTA* **46**, 97–112.

[18] B. Ko & Q. Tang (2008) Sums of dependent nonnegative random variables with subexponential tails. *J. Appl. Prob.* **45**, 85–94.

[19] D. Kortschak & H. Albrecher (2009) Asymptotic results for the sum of dependent non-identically distributed random variables. *Methodol. Comput. Appl. Probab.* **11**, 279–306.

[20] J. Li, Q. Tang & R. Wu (2010) Subexponential tails of discounted aggregate claims in a time-dependent renewal risk model. *Adv. Appl Probab.* **42**, 1126–1146.

[21] P. Mikusiński, H. Sherwood, & M. D. Taylor.(1991) Probabilistic interpretations of copulas and their convex sums. In *Advances in Probability Distributions with Given Marginals (Rome, 1990)*, volume 67 of *Math. Appl.*, pp. 95–112. Kluwer Acad. Publ., Dordrecht.

[22] A. Mitra & S. Resnick (2009) Aggregation of rapidly varying risks and asymptotic independence. *Adv. in Appl. Probab.* **41**, 797–828.

[23] R. B. Nelsen (1999) *An Introduction to Copulas*, Springer-Verlag, New York.

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

[25] C. Y. Robert & J. Segers (2008) Tails of random sums of a heavy-tailed number of light-tailed terms. *Insurance Math. Econom.* **43**, 85–92.