Siberian Mathematical Journal, Vol. 53, *No.* 6, *pp.* 965–983, 2012

TAIL ASYMPTOTICS FOR DEPENDENT SUBEXPONENTIAL DIFFERENCES

© **H. Albrecher, S. Asmussen, and D. Kortschak** UDC 519.2

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 and 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 $X - Y$. Some explicit construction of the worst-case copula is provided in other cases.

Keywords: subexponential random variables, differences, dependence, copulas, mean excess function

1. Introduction

In recent years, some progress has been achieved 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 and Resnick [2], Ko and Tang [3], Kortschak and Albrecher [4], and Foss and Richards [5]. In the present paper we are interested in the tail asymptotics of differences of random variables, i.e. in $\mathbb{P}(X - Y > u)$ as $u \to \infty$, where X is subexponential and the positive random variable Y may have different forms of the tail. In case X and Y are independent, this is easy (cf. [6, Lemma 3.2]):

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

without further conditions. Thus, the problem is dependence.

Since $\mathbb{P}(X - Y > u) = \mathbb{P}(\max(X, 0) - Y > u)$ for positive u, we can assume without loss of generality 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 and Teugels [7], Boudreault et al. [8], Asimit and Badescu [9], Li et al. [10] and also Albrecher and Boxma [11] for such and similar dependence structures). Similar interpretations are possible in queueing applications.

Asmussen and Biard in [12] 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 address the case of Y light-tailed 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* (see, e.g., Jelenković et al. [13]).

The authors were supported by the Swiss National Science Foundation (Project 200021–124635/1).

Lausanne, Aarhus, and Lyon. Translated from *Sibirskiĭ Matematicheskiĭ Zhurnal*, Vol. 53, No. 6, pp. 1209–1230, November–December, 2012. Original article submitted September 29, 2011.

0037-4466/12/5306–0965

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 can fail to represent the dependence structure that produces the most extreme behavior of $\mathbb{P}(X - Y > u)$. We provide criteria under which the comonotonic dependence is indeed the worst case in the sense of minimizing the tail of $X - Y$ and provide some 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 the classical results that are used in the paper. For a random variable X with cumulative distribution function, let $F_X(u)$ 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 we have

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

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

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

As an extension of regularly varying distributions we consider the 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 (see [14, 15]). From [16, Theorem 2.47] it follows that $\overline{F}_X(u)$ is intermediate regularly varying if and only if

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

for every positive function $\delta(u)$ with $\lim_{u\to\infty} \delta(u)/u = 0$. For a recent survey on heavy-tailed random variables see [16].

Another useful extension of regularly varying distributions is related to extreme value theory (see [17] or [18] for the 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 nondegenerate distribution function $H(x)$ with

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

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

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

see, e.g., [17, Proposition 0.3]. We say that X (or equivalently $\overline{F}_X(x)$) is in the *maximum attraction domain* of the extreme value distribution H. In [17, Chapter 1] it is shown that X is in the maximum attraction domain of the Fréchet distribution if and only if X is regularly varying. If X is in the maximum attraction domain of the Weibull distribution then X has a finite right endpoint. Finally, X is in the maximum attraction domain 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 [19, Section 3.10]). The function $e(x)$ is unique up to asymptotic equivalence and can be chosen as the mean excess function $e_m(x) = \mathbb{E}(X - x \mid X > x)$ 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 attraction domain 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 marginal distributions. Therefore we will use copulas and review some basic concepts (a standard reference is [20]). A *two-dimensional copula* $C(u, v)$ is a function that fulfills

- $C(u, 0) = 0 = C(0, v)$ for all $u, v \in [0, 1]$;
- $C(u, 1) = u$ and $C(1, v) = v$ for all $u, v \in [0, 1]$;
- C is 2-increasing; i.e., for all $u_1, u_2, v_1, v_2 \in [0, 1]$ 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 each copula there exist random variables X and Y with marginal distributions F_X and F_Y such that (2.3) holds. We say that X and Y are *dependent* according to the copula C. Note that C is invariant under monotonic transformations of the marginal distributions. The Fréchet 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 we 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 for constructing dependent random variables with given marginal distributions. In this paper we will use the following two methods of constructing copulas (see, e.g., [20, Chapter 3]). Denote by $\{J_i\}$ a partition of $[0, 1]$ defined here as a collection of closed intervals $J_i = [a_i, b_i]$ that are nonoverlapping (except at endpoints) and $\bigcup J_i = [0, 1]$ (we can assign the overlapping points to one of the involved intervals and so get a partition in the classical sense). Given a partition $\{J_i\}$ and a 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 if U and V are uniform random variables dependent according to an ordinal sum, then $\mathbb{P}(U \in J_i \mid V \in J_i) = 1$ and the random vector $(U, V) \mid (U, V) \in J_i^2$ has uniform marginals on J_i that are dependent according to the copula C_i.

The second type of copulas to be used in the sequel consists of the so-called straight shuffles of M. Assume that we have a copula C, a finite partition $\mathscr{J} = \{J_1, \dots, J_n\}$ of $[0,1]$, and a permutation π of $\{1, \dots, n\}$. The copula C defines a measure on the strips $J_i \times [0,1]$ or, equivalently, on the strips of the length $h_i = b_i - a_i$. We can now reorder these strips according to the permutation π. So to the strip $[0, h_{\pi(1)}] \times [0,1]$ we assign the measure that is assigned by C to the strip $J_{\pi(1)} \times [0,1]$, to the strip $[h_{\pi(1)}, h_{\pi(1)} + h_{\pi(2)}] \times [0,1]$ we assign the measure that 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 $C_s(\mathscr{J}, \pi)$. We call $C_s(\mathscr{J}, \pi)$ a *straight shuffle* of M if $C = M$, and then use the notation $M_s(\mathscr{J}, \pi)$. From the discussion after Theorem 3.2.3 in [20] (see also [21]) it follows that each copula can be approximated arbitrary closely by a shuffle with respect to the supremum norm.

In the later sections we will also use multivariate extreme value theory which studies the component-wise maximum of multivariate random variables (the results presented here can, for example, be found in [17, Section 5.4] or [18]). Consider the possible limits of

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

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

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

The copula C is then said to lie in the maximum attraction domain of the extreme value copula C_*. Note that $H(x,y) = C_*(H_X(x), H_Y(y))$.

We now briefly outline the significance of extreme value theory for the purposes of the 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 we easily check that

$$\lim_{t\to\infty} \frac{\mathbb{P}(X > tx \text{ or } Y > tcy)}{\mathbb{P}(X > t)} = \lim_{t\to\infty} \frac{1 - C(F_X(tx), F_Y(tcy))}{\overline{F}_X(t)}$$

$$= -\lim_{t\to\infty} \log \left[C\left(\exp\left\{ \overline{F}_X(t) \frac{\log(F_X(tx))}{\overline{F}_X(t)} \right\}, \exp\left\{ \overline{F}_X(t) \frac{\log(F_Y(tcy))}{\overline{F}_X(t)} \right\} \right)^{1/\overline{F}_X(t)} \right]$$

$$= -\log(C_*(\mathrm{e}^{-x^{-\alpha}}, \mathrm{e}^{-y^{-\alpha}})).$$

For the last equality we needed that the function on the right-hand side is continuous. Now for every t the left-hand side of the equation defines a measure H_t on $[0,\infty]^2$ and the right-hand side 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). For $A = \{(x,y) : x - cy > 1\}$ we have now 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 we only need to consider the case $\alpha = 1$ to get some characterization of H. If we write $x = r\theta$ and $y = r(1-\theta)$, then it follows from [17, Proposition 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 attraction domain of the Gumbel distribution the same steps are applicable.

3. An Insensitivity Result

From, e.g., Foss et al. [16] if a distribution F is long-tailed, then there exists a nondecreasing function δ with $\delta(u) \to \infty$ as $u \to \infty$ such that

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

In what follows, we will be interested in choosing $\delta(u)$ as large as possible. The next proposition is essentially equivalent to [5, Proposition 5.1], but we give the proof because of its simplicity and usefulness for the later purposes:

Proposition 3.1. *Let $X \ge 0$ be a real variable with a long-tailed distribution F_X and let $Y \ge 0$ be a (not necessarily independent) random variable. Then (1.1) holds provided that $\delta(\cdot)$ 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. Put

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

Note that by (3.2)

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

Moreover,

$$\begin{gathered}\mathbb{P}(X - Y > u,\, Y \le \delta(u)) \le \mathbb{P}(X > u) = \overline{F}_X(u),\\ \mathbb{P}(X - Y > u,\, Y \le \delta(u)) \ge \mathbb{P}\big(X - \delta(u) > u,\, Y \le \delta(u)\big)\\ = \mathbb{P}(X - \delta(u) > u) - \mathbb{P}(X - \delta(u) > u,\, Y > \delta(u)) \sim \overline{F}_X(u) - \mathrm{o}(\overline{F}_X(u)).\end{gathered}$$

Combining these estimates completes the proof. □

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

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

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 $\widehat{C}(u,v) \le u h(v)$ with $h(v) \to 0$. In terms of distribution functions, this means that $\mathbb{P}(X > x,\, Y > y) \le \mathbb{P}(X > x) h(\mathbb{P}(Y > y))$ for all $x, y \ge 0$. □

Example 3.3. More generally, we can formulate a criterion in terms of stochastic ordering: whenever the pair (X^1, Y^1) fulfills (3.2), then every pair (X^2, Y^2) with the same marginal distributions that is dominated in concordance order (i.e. $\mathbb{P}(X^1 > x, Y^1 > y) \ge \mathbb{P}(X^2 > x, Y^2 > y)$ 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 lies in the maximum attraction domain of the Gumbel distribution with mean excess function $e_m(u)$. Then $\delta(u)$ in (3.1) can be any function satisfying $\delta(u) \to \infty$ 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 attraction domain of the Gumbel distribution we can choose $e(u)$ 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 $\varepsilon_0>0$ such that, uniformly in $0<t<\varepsilon_0$,

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

Then

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

Remark 4.2. Note that for 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

$$\frac{\mathbb{P}\big(X-\varepsilon\frac{1}{r^*(u)}>u\big)}{\mathbb{P}(X>u)}\sim\exp\Big(-\int_u^{u+\frac{\varepsilon}{r^*(u)}}r^*(t)\,dt\Big)$$

$$=\exp\Big(-\int_0^{\varepsilon}\frac{r^*\big(u+\frac{t}{r^*(u)}\big)}{r^*(u)}\,dt\Big)\lesssim\exp\Big(-c_l\int_0^{\varepsilon}dt\Big)=\mathrm{e}^{-c_l\varepsilon}<1$$

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

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

from which the result follows. □

Remark 4.3. The example for which the conditions of Lemma 4.1 are not fulfilled is as follows:

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

5. Y Light-Tailed

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

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

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 random variable (i.e. $\mathbb{P}(Y > \delta(u)) = \mathrm{o}(e^{-gu})$ for some $g > 0$), for which the long-tailed random variable X does (1.1) hold along all dependence structures? In this case condition (5.1) turns into

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

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 $\overline{F}_X(x) = \mathrm{e}^{-x^\beta}$ if $\beta < 1/2$ ($\delta(x) = x^{1-\beta^*}$ if $\beta < \beta^* < 1$). 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), we could believe that the condition is just technical. But 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 and Biard [12], and an even simpler construction goes as follows:

Example 5.1. Assume that $\mathbb{P}(X > u) \sim \mathrm{e}^{-u^\beta}$ 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

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

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

This counterexample (as well as that in Asmussen and Biard [12]) involves 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. Put*

$$\begin{aligned}\overline{\gamma}(u) &= \sup\{x \mid F_Y(x - u) < F_X(x), x \ge u\} - u,\\ \underline{\gamma}(u) &= \inf\{x \mid F_Y(x - u) \ge F_X(x), x \ge u\} - u.\end{aligned}$$

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

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

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. To begin with note that $\mathbb{P}(X-Y>u)\le\mathbb{P}(X>u)$. We have

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

To prove the first statement of the lemma, observe that

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

For the second integral we have

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

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

$$\frac{\mathbb{P}(Y>k\overline{\gamma}(u))}{\mathbb{P}(X>u)}=\frac{\mathbb{P}(Y>k\overline{\gamma}(u))}{\mathbb{P}(Y>\overline{\gamma}(u))}\frac{\mathbb{P}(Y>\overline{\gamma}(u))}{\mathbb{P}(X>u)}\le c_1k^{-\alpha}.$$

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

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

and so $\overline{\gamma}(u)$ 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 X and Y comonotonic we have

$$\mathbb{P}(X-Y>u)=\int_u^\infty \mathbb{P}(Y\le X-u \mid X=x)\,dF_X(x)$$

$$\le \int_u^\infty I_{\{F_Y(x-u)\ge F_X(x)\}}\,dF_X(x)\le \int_{u+\underline{\gamma}(u)}^\infty dF_X(x)=\mathbb{P}(X>u+\underline{\gamma}(u)).\quad \square$$

Although Lemma 6.1 shows that the comonotonic copulas are natural candidates for counterexamples, this does not tell whether the comonotonic copula represents the worst case, i.e., the copula that minimizes $\mathbb{P}(X-Y>u)$ asymptotically for given marginal distributions. To answer the question, let us first consider the case of X regularly varying. 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 $F_X(x)\ge F_Y(x)$ for X and Y comonotonic, then $\mathbb{P}(X-Y>u)=0$. Hence, assume that there exists $\hat{c}>0$ with

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

or, equivalently, there exists 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. From extreme value theory it follows that

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

To understand which H minimizes $H(\{(x,y)\mid x-cy>1\})$, 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 a norm), H can be written as a product of the measures on the radial and angular parts. Then the radial measure has density $\alpha r^{-\alpha-1}$ and (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

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

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

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

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

PROOF. Assume that the result does not hold. Then without loss of generality we can assume that $\theta_1 > 1/2$ and $\theta_2 < c/(c+1)$. Define the 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

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

where without loss of generality we assumed that p_2 is small enough such that $\hat{p}_1 \ge 0$. Thus

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

which is a contradiction to μ^* minimizing (6.1). $\square$

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

PROOF. Assume that μ^* assigns the positive measure $p_1 > 0$ to $c/(1+c) < \theta_1 < 1$. Then we can define the new measure μ^{**} 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

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

Furthermore,

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

from which the result follows. $\square$

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

PROOF. Suppose that μ^* assigns the positive measure $p_1 > 0$ to $\theta_1 > 1/2$ and $p_2 > 0$ to $\theta_2 < 1/2$, where we assume without loss of generality 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 the two cases:

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

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

The left-hand side equals

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

so that we need to show

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

(cf. the method 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 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

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

This is equivalent to showing that

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

Again the left-hand 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 attraction domain of the Gumbel distribution with auxiliary function $e(x)$. Further assume that there exists $0 < c < 1$ 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 attraction domain of an extreme value copula. Then the copula that asymptotically minimizes $\mathbb{P}(X - Y > u)$ is the comonotonic copula.

Proof. Again

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

Here, $H(x, y) = H^*(e^x, e^y)$, where the functions $R = x + y$ and $\theta = x/(x + y)$ corresponding to H^* 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 see that

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

with

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

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

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

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

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

Note that for X and Y comonotonic we can replace $\geq$ by $=$. Finally we have to find μ that minimizes (6.3). Again, we only consider μ discrete. Without loss of generality 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 $-1 < x_2 < 0 < x_1$. $\square$

Theorem 6.3 shows that if $X \in \mathscr{R}_{-\alpha}$ 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. We now show however 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 $\mathbb{P}_C$ as

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

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

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

In view of Proposition 7.1 of the next section, we can assume that for X regularly varying 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. $2Y \overset{d}{=} X$. Further, let X be in the maximum attraction domain 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 $u_1 > 0$ with $F(u_1) > 0$, together with a corresponding partition $(J_i)_{n\ge1}$ of the interval $[0,1]$ $(n \ge 1)$

$$J_1 = [0, F(2u_1)),$$

$$J_{2n} = [F(2u_n), F(2(u_n + e(2u_n)))), \quad J_{2n+1} = [F(2(u_n + e(2u_n))), F(2u_{n+1})).$$

Moreover, define the series $(C_n)_{n\ge1}$ 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 $(C_n)_{n\ge1}$ with respect to the partition $(J_i)_{n\ge1}$.

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

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

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

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

Further assume that $2Y \overset{d}{=} X$ and 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. Without loss of generality we assume that $e(x)$ is monotonic. For every n we have

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

Now we 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,

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

Note that

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

as $n\to\infty$. 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 so

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

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

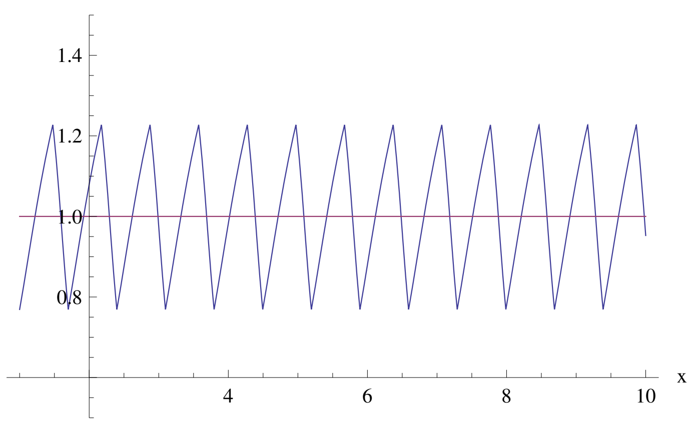

Fig. 1. Plot of $\frac{\mathbb{P}_{\overline{C}}(X-Y>\frac{1}{2}10^x)}{\mathbb{P}_M(X-Y>\frac{1}{2}10^x)}$.

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 $\mathbb{P}(X - Y > u)$. Given F_X, F_Y, and u, define

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

For uniformly distributed (U_1, U_2) with the same copula C as (X, Y), we have

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

Lemma 6.10. *Let $g(x)$ be an increasing function such that for all $c \in [-1, 1]$ the number of the times where $g(x) - x - c$ changes sign is finite. Then the shuffle M_s^* that minimizes $\mathbb{P}_{M_s}(U_2 < g(U_1))$ is of the form $\mathscr{J} = \{[0, x_0], [x_0, 1]\}$ and $\pi = (2, 1)$ for some $0 < x_0 < 1$.*

Proof. Let M_s be a shuffle with finite partition $\mathscr{J}$ and permutation π. For $J \in \mathscr{J}$ and $x \in J$, denote by J^π and x^π the interval J (the point to which x, respectively) is mapped by the permutation. Without loss of generality we assume that for every $J \in \mathscr{J}$

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

Put $x_0 = \mathbb{P}_{M_s}(U_2 < g(U_1))$. Without loss of generality we can assume that for every $J \in \mathscr{J}$, $(J \cap [0, x_0]) \in \{\varnothing, J\}$. Further we can split the intervals in the partition $\mathscr{J}$, such that to every interval $J \in \mathscr{J}$ 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 the unique interval $\widehat{J}$ with $\widehat{J} \cap [0, x_0] = \widehat{J}$ and $|J| = |\widehat{J}|$. If we change the position of J and $\widehat{J}$ 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. $\square$

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

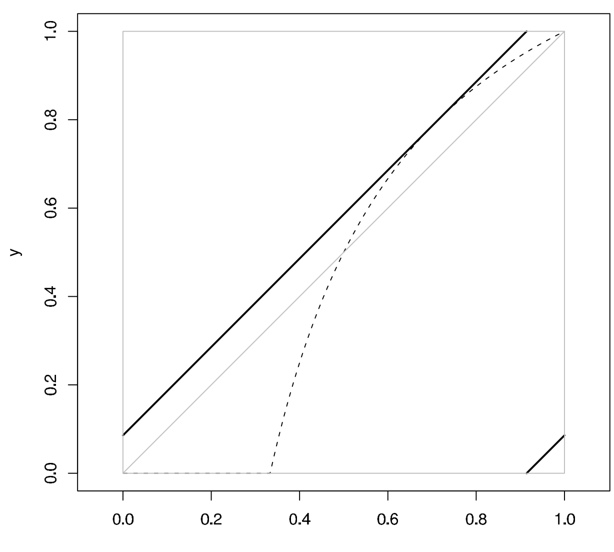

Fig. 2. A worst-case copula.

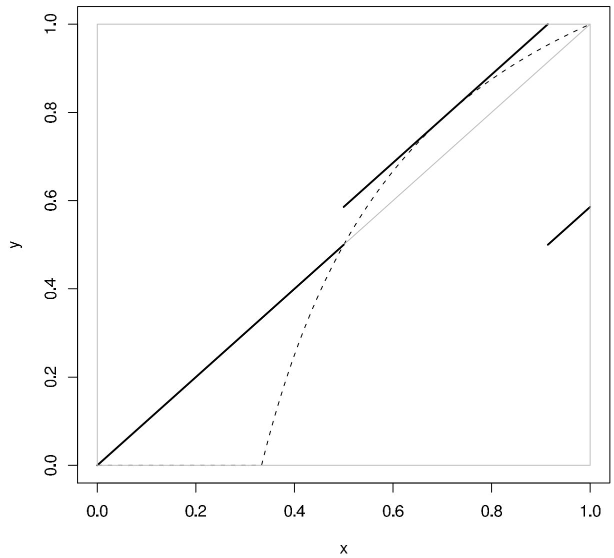

Fig. 3. Another worst-case copula.

Lemma 6.11. *Let $g(x)$ be an increasing function. Put $x_1 = \inf\{x : x \geq g(x)\}$. If $x_1 < 1 - x_0$ for some x_0, then the shuffles $M_s(\{[0, x_0], [x_0, 1]\}, (2, 1))$ and $\widehat{M}_s(\{[0, x_1], [x_1, x_1 + x_0], [x_1 + x_0, 1]\}, (1, 3, 2))$ fulfill $\mathbb{P}_{M_s}(U_2 < g(U_1)) \geq \mathbb{P}_{\widehat{M}_s}(U_2 < g(U_1))$. 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. Assume that $F_X(x) = 1 - 1/x$, $F_Y(x) = 1 - 1/(2x)$ and $u = 1$. For this case, Fig. 2 shows the support of the copula in Lemma 6.10 (the bold line), where $x_0 \approx 0.086$. In Fig. 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 $g_u(x)$. Here

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

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

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

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

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

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

Let us compare this result to the comonotonic copula. To this end, assume that there exists a unique point γ_u such that $g_u(x) - x \leq 0$ 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

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

If the function

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

converges as $u \to \infty$ to the function $h_\infty(x)$ with $\sup_{0<x<1} h_\infty(x) = 1$ (i.e., $h_\infty(x) = 1 - x$), then (6.4) holds for every copula C. 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 by analogy with Proposition 6.8 we can construct some 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

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

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 see 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 have

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

since as $n \to \infty$

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

We get that

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

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

7. X Intermediate Regularly Varying

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 $\delta(u)$ also fulfills (2.1). Note firstly that for every $c > 0$ there exists b_c such that

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

Hence, for every n there exists $\hat{u}_n$ such that $\frac{\mathbb{P}(Y>u)}{\mathbb{P}(X>nu)} \le \frac{1}{n}$ for all $u > \hat{u}_n$. Put $u_0 = 0$ and $u_n = \max(n\hat{u}_n, u_{n-1}) + 1$ for $n > 0$. Then $\frac{\mathbb{P}(Y>u/n)}{\mathbb{P}(X>u)} \le \frac{1}{n}$ for all $u > u_n$. Also 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

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

7.1. Approach via local limit laws. Let us now use local limit laws as in Heffernan and Resnick [22] 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) such that for every fixed y in $\mathbb{E}$

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

Assume that $\alpha(u)/e(u) \to c$ for some constant c. Then

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

at least if μ is sufficiently continuous. The area to be measured is depicted in Fig. 4.

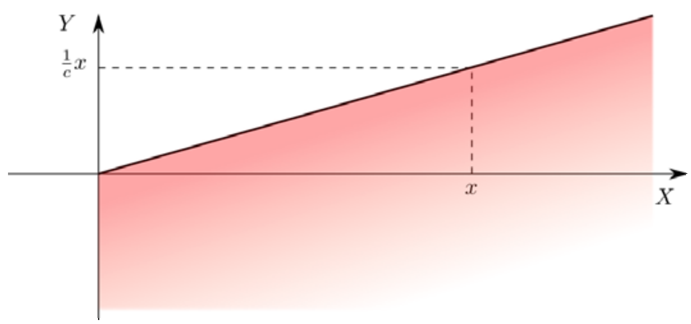

Fig. 4. Area to be measured (shaded).

It follows that

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

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

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

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).

References

1. *Albrecher H., Asmussen S., and Kortschak D.*, "Tail asymptotics for the sum of two heavy-tailed dependent risks," Extremes, **9**, No. 2, 107–130 (2006).
2. *Mitra A. and Resnick S.*, "Aggregation of rapidly varying risks and asymptotic independence," Adv. Appl. Probab., **41**, No. 3, 797–828 (2009).
3. *Ko B. and Tang Q.*, "Sums of dependent nonnegative random variables with subexponential tails," J. Appl. Probab., **45**, No. 1, 85–94 (2008).
4. *Kortschak D. and Albrecher H.*, "Asymptotic results for the sum of dependent non-identically distributed random variables," Methodol. Comput. Appl. Probab., **11**, No. 3, 279–306 (2009).
5. *Foss S. and Richards A.*, "On sums of conditionally independent subexponential random variables," Math. Oper. Res., **35**, No. 1, 102–119 (2010).
6. *Asmussen S. and Albrecher H.*, Ruin Probabilities (2nd ed.), World Scientific, Singapore (2010).
7. *Albrecher H. and Teugels J.*, "Exponential behavior in the presence of dependence in risk theory," J. Appl. Probab., **43**, No. 1, 257–273 (2006).
8. *Boudreault M., Cossette H., Landriault D., and Marceau E.*, "On a risk model with dependence between interclaim arrivals and claim sizes," Scand. Actuar. J., No. 5, 265–285 (2006).
9. *Asimit A. V. and Badescu A. L.*, "Extremes on the discounted aggregate claims in a time dependent risk model," Scand. Actuar. J., No. 2, 93–104 (2010).
10. *Li J., Tang Q., and Wu R.*, "Subexponential tails of discounted aggregate claims in a time-dependent renewal risk model," Adv. Appl Probab., **42**, No. 4, 1126–1146 (2010).
11. *Albrecher H. and Boxma O. J.*, "A ruin model with dependence between claim sizes and claim intervals," Insurance Math. Econom., **35**, No. 2, 245–254 (2004).
12. *Asmussen S. and Biard R.*, "Ruin probabilities for a regenerative Poisson gap generated risk process," Eur. Actuar. J., **1**, No. 1, 3–22 (2011).
13. *Jelenković P., Momcilović P., and Zwart B.*, "Reduced load equivalence under subexponentiality," Queueing Syst., **46**, No. 1–2, 97–112 (2004).
14. *Cline D. B. H.*, "Intermediate regular and Π variation," Proc. London Math. Soc., **68**, No. 3, 594–616 (1994).
15. *Robert C. Y. and Segers J.*, "Tails of random sums of a heavy-tailed number of light-tailed terms," Insurance Math. Econom., **43**, No. 1, 85–92 (2008).
16. *Foss S., Korshunov D., and Zachary S.*, An Introduction to Heavy-Tailed and Subexponential Distributions, Springer-Verlag, New York (2011).
17. *Resnick S. I.*, Extreme Values, Regular Variation, and Point Processes, Springer-Verlag, New York (1987).
18. *Galambos J.*, The Asymptotic Theory of Extreme Order Statistics, Robert E. Krieger Publ. Co. Inc., Melbourne, FL (1987).
19. *Bingham N. H., Goldie C. M., and Teugels J. L.*, Regular Variation, Cambridge Univ. Press, Cambridge (1989) (Encyclopedia of Mathematics and its Applications; V. 27).
20. *Nelsen R. B.*, An Introduction to Copulas, Springer-Verlag, New York (1999).
21. *Mikusiński P., Sherwood H., and Taylor M. D.*, "Probabilistic interpretations of copulas and their convex sums," in: Advances in Probability Distributions with Given Marginals (Rome, 1990), Kluwer Acad. Publ., Dordrecht, 1991, pp. 95–112 (Math. Appl.; V. 67).
22. *Heffernan P. and Resnick S.*, "Limit laws for random vectors with an extreme component," Ann. Appl. Probab., **17**, No. 2, 537–571 (2007).

H. Albrecher
University of Lausanne, Lausanne, Switzerland
E-mail address: `hansjoerg.albrecher@unil.ch`

S. Asmussen
Aarhus University, Aarhus, Denmark
E-mail address: `asmus@imf.au.dk`

D. Kortschak
University of Lausanne, Lausanne, Switzerland
University of Lyon, Lyon, France
E-mail address: `kortschakdominik@gmail.com`