

AGGREGATION OF RAPIDLY VARYING RISKS AND ASYMPTOTIC INDEPENDENCE

ABHIMANYU MITRA^{***} AND
SIDNEY I. RESNICK,^{***} *Cornell University*

Abstract

We study the tail behavior of the distribution of the sum of asymptotically independent risks whose marginal distributions belong to the maximal domain of attraction of the Gumbel distribution. We impose conditions on the distribution of the risks (X, Y) such that $P(X+Y > x) \sim (\text{constant}) P(X > x)$. With the further assumption of nonnegativity of the risks, the result is extended to more than two risks. We note a sufficient condition for a distribution to belong to both the maximal domain of attraction of the Gumbel distribution and the subexponential class. We provide examples of distributions which satisfy our assumptions. The examples include cases where the marginal distributions of X and Y are subexponential and also cases where they are not. In addition, the asymptotic behavior of linear combinations of such risks with positive coefficients is explored, leading to an approximate solution of an optimization problem which is applied to portfolio design.

Keywords: Risk; Gumbel; maximal domain of attraction; asymptotic independence; subexponential

2000 Mathematics Subject Classification: Primary 60E07; 60G51; 60G52; 60G70
Secondary 60F17; 62G30; 90B50; 90C59

1. Introduction

Estimating the probability that a sum of risks, $X + Y$, exceeds a large threshold is important in finance and insurance, and, hence, much applied probability research has been dedicated to this goal. Recent results can be found in Albrecher *et al.* (2006), Klüppelberg and Resnick (2008), Wang and Tang (2006), Asmussen and Rojas-Nandayapa (2008), Alink *et al.* (2004), Embrechts and Puccetti (2006), and Ko and Tang (2008). Approximating this probability helps us evaluate risk measures for investment portfolios as well as estimating credit risk.

The problem is reasonably well understood when risks have regularly varying marginal distributions, but another important large class of risk distributions is the maximal domain of attraction of the Gumbel distribution, denoted by $\text{MDA}(\Lambda)$, where

$$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R},$$

and $\text{MDA}(\Lambda)$ is the class of distributions F for which there exist $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n \rightarrow \infty} n(1 - F(a_n x + b_n)) = \lim_{n \rightarrow \infty} n\bar{F}(a_n x + b_n) = e^{-x}, \quad x \in \mathbb{R} \quad (1.1)$$

Received 29 July 2008; revision received 13 June 2009.

* Postal address: School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY 14853, USA.

Research partially supported by ARO contract W911NF-07-1-0078 at Cornell University.

** Email address: am492@cornell.edu

*** Email address: sir1@cornell.edu

(see Resnick (1987, p. 38)). It is also well known that the risks having distribution in $MDA(\Lambda)$ are rapidly varying, i.e. $-\infty$ -varying (see Resnick (1987, p. 53)). Within the class of risks (X, Y) with marginal distributions $F, G \in MDA(\Lambda)$, results on aggregation of risks are known when X and Y are independent. However, actual risks are often not independent and a somewhat weaker concept called *asymptotic independence* allows risks to be modeled as dependent and is more practical in many modeling situations. Risks X and Y in a maximal domain of attraction are *asymptotically independent* if, for all $\mathbf{x} = (x_1, x_2)$,

$$\lim_{n \rightarrow \infty} H^n(a_n^{(1)}x_1 + b_n^{(1)}, a_n^{(2)}x_2 + b_n^{(2)}) = G_1(x_1)G_2(x_2), \tag{1.2}$$

where H is the joint distribution of X and Y , and both G_1 and G_2 are nondegenerate extreme value distributions (see de Haan and Ferreira (2006, p. 229)). There are also results on aggregation of risks in the absence of asymptotic independence where the analogue of (1.2) holds but with a limit distribution which is not a product; see Klüppelberg and Resnick (2008).

In this paper we consider the case where the risks X and Y are asymptotically independent with marginal distributions $F, G \in MDA(\Lambda)$. We also allow one marginal tail to be lighter and the distribution with the lighter tail does not necessarily belong to the maximal domain of attraction of the Gumbel distribution.

Within the class of vectors (X, Y) satisfying asymptotic independence and marginal distributions $F, G \in MDA(\Lambda)$, two prominent but very distinct behaviors have been observed.

- First, suppose that X and Y are two independent and identically distributed (i.i.d.) risks with common distribution F , which is subexponential and belongs to $MDA(\Lambda)$. Then X and Y are certainly asymptotically independent and

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = 2. \tag{1.3}$$

So one possible behavior is that the sum has a distribution which is tail equivalent to the distribution of a summand.

- Very different tail behavior is exhibited in Theorem 2.10 of Albrecher *et al.* (2006), in which a joint distribution of (X, Y) was given making X and Y asymptotically independent with identical marginal distribution $F \in MDA(\Lambda)$, but

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = \infty.$$

In Section 2 we give a set of conditions on the joint distribution of (X, Y) , guaranteeing behavior of the first sort, namely,

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = 1 + c, \tag{1.4}$$

where $c = \lim_{x \rightarrow \infty} P(Y > x)/P(X > x)$, the limit being assumed to exist. If $c \in (0, \infty)$, our conditions imply that X and Y are asymptotically independent, and that each belongs to the maximal domain of attraction of the Gumbel distribution. When X and Y are identically distributed, (1.3) holds. Under the further assumption of nonnegativity of risks, the result is extended for the case of more than two risks. In Section 3 we provide examples of distributions which satisfy our conditions. The examples include cases where the marginal distributions of X and Y are subexponential and also cases where they are not. We also show one example

which does not satisfy our conditions, but exhibits the tail equivalence between the distribution of the sum and that of the summand. Thus, our conditions are only sufficient. In Section 4 we summarize the asymptotic behavior of finite linear combinations of risks with nonnegative coefficients. In Section 5 we suggest approximate solutions for an optimization problem which is related to portfolio design. The paper closes with concluding remarks and a brief summary of numerical experiments which give a feel for whether asymptotic equivalence is a suitable numerical approximation for exceedance probabilities of aggregated risks.

2. Asymptotic tail probability for aggregated risk

2.1. Asymptotic tail probability for the sum of two random variables

We give conditions guaranteeing (1.4). The constant c satisfies

$$c = \lim_{x \rightarrow \infty} \frac{P(Y > x)}{P(X > x)} \in [0, \infty).$$

When $c \in (0, \infty)$, X and Y are called tail equivalent (see Resnick (1971b)), and then our conditions guarantee that both the marginal distributions $F, G \in \text{MDA}(\Lambda)$ and X and Y are asymptotically independent. When $c = 0$, our result extends to the case where G , the marginal distribution of Y , does not belong to the maximal domain of attraction of the Gumbel distribution and where X and Y need not be asymptotically independent.

2.1.1. *Assumptions.* Suppose that (X, Y) is a pair of random variables satisfying the following set of assumptions.

Assumption 2.1. *The random variable X has a distribution F whose right endpoint x_0 is infinite, that is,*

$$x_0 = \sup\{x : F(x) < 1\} = \infty. \tag{2.1}$$

Furthermore, $F \in \text{MDA}(\Lambda)$ so that (1.1) is satisfied with centering constants $b_n \in \mathbb{R}$ and scaling constants $a_n > 0$. Equivalently (see de Haan (1970, pp. 88, 91) and Resnick (1987, pp. 28, 40–43)), there exists a self-neglecting auxiliary function $f(\cdot)$ whose derivative converges to 0, such that

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(t + xf(t))}{\bar{F}(t)} = e^{-x}. \tag{2.2}$$

Assumption 2.2. *The random variables X and Y have distribution functions F and G such that*

$$\lim_{x \rightarrow \infty} \frac{\bar{G}(x)}{\bar{F}(x)} = c \in [0, \infty).$$

Assumption 2.3. *The conditional distribution of Y , given $X > x$, satisfies, for all $t > 0$,*

$$\lim_{x \rightarrow \infty} P(|Y| > tf(x) \mid X > x) = 0,$$

where $f(x)$ is the auxiliary function corresponding to the distribution of X given in (2.2).

Assumption 2.4. *Symmetrically, assume that, for all $t > 0$,*

$$\lim_{x \rightarrow \infty} P(|X| > tf(x) \mid Y > x) = 0.$$

Assumption 2.5. *For some $L > 0$, suppose that*

$$\lim_{x \rightarrow \infty} \frac{P(Y > Lf(x), X > Lf(x))}{P(X > x)} = 0.$$

2.1.2. *The main result.* Assumptions 2.1–2.5 allow us to conclude that aggregated risks are essentially tail equivalent to individual risks.

Theorem 2.1. *Under Assumptions 2.1–2.5, we have*

$$P(X + Y > x) \sim (1 + c)P(X > x), \quad x \rightarrow \infty.$$

2.1.3. *Comments on the assumptions.* Before giving a proof of Theorem 2.1, we discuss implications of the assumptions.

1. When $F \in \text{MDA}(\Lambda)$, we may choose a_n and b_n appearing in (1.1) as $b_n = b_F(n)$ and $a_n = f(b_n)$. See Resnick (1987, p. 40) or de Haan and Ferreira (2006, p. 10).

2. If $c \in (0, \infty)$ then our assumptions guarantee that both marginal distributions F and G belong to $\text{MDA}(\Lambda)$ and also that (X, Y) are asymptotically independent. From Assumption 2.1, $F \in \text{MDA}(\Lambda)$, and since F and G are tail equivalent, from Resnick (1971b) we obtain $G \in \text{MDA}(\Lambda)$. For asymptotic independence, define

$$b_F(t) = \inf \left\{ s : \frac{1}{1 - F}(s) \geq t \right\} = \left(\frac{1}{1 - F} \right)^{\leftarrow} (t);$$

define $b_G(t)$ similarly. From de Haan and Ferreira (2006, p. 229), if $F, G \in \text{MDA}(\Lambda)$ and

$$\lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_G(t))}{P(X > b_F(t))} = 0, \tag{2.3}$$

then X and Y are asymptotically independent according to (1.2). When $c \in (0, \infty)$, Assumption 2.3 implies (2.3). To verify this, note first that Assumption 2.3 implies that

$$\lim_{x \rightarrow \infty} \frac{P(X > x, Y > x)}{P(X > x)} \leq \lim_{x \rightarrow \infty} \frac{P(X > f(x), Y > x)}{P(X > x)} = 0, \tag{2.4}$$

since $f(x)/x \rightarrow 0$ as $x \rightarrow \infty$ (see Resnick (1987, p. 40)). If $c > 1$ then, for sufficiently large t , $b_F(t) \leq b_G(t)$ and, therefore, using (2.4),

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_G(t))}{P(X > b_F(t))} &\leq \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_F(t))}{P(X > b_F(t))} \\ &= \lim_{t \rightarrow \infty} \frac{P(X > t, Y > t)}{P(X > t)} \\ &= 0, \end{aligned}$$

as required. A similar verification can be constructed for the case $0 < c < 1$. For $c = 1$, $b_F(t) \sim b_G(t)$. Hence,

$$\frac{f(b_F(t))}{b_G(t)} \sim \frac{f(b_F(t))}{b_F(t)} \rightarrow 0.$$

So,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_G(t))}{P(X > b_F(t))} &\leq \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > f(b_F(t)))}{P(X > b_F(t))} \\ &= 0 \quad (\text{by Assumption 2.3 and (2.1)}). \end{aligned}$$

3. The auxiliary function $f(\cdot)$ can be replaced by any asymptotically equivalent function $\tilde{f}(\cdot)$; that is, if $\lim_{x \rightarrow \infty} \tilde{f}(x)/f(x) = 1$, and if Assumptions 2.3–2.5 hold with $f(\cdot)$, they also hold with $\tilde{f}(\cdot)$ replacing $f(\cdot)$ and vice versa. Since the mean excess function

$$e(x) = E(X - x \mid X > x)$$

is asymptotically equivalent to any auxiliary function $f(x)$ (see Embrechts *et al.* (1997, p. 143) and Resnick (1987, p. 48)), $e(x)$ can also be taken as an auxiliary function.

4. If $c = \lim_{x \rightarrow \infty} \tilde{G}(x)/\bar{F}(x) = 0$, we do not need Assumption 2.4 to conclude our result.

5. An easier proof of the result can be given if Assumption 2.5 holds for all $L > 0$. But here we provide an example to show the importance of the weak version of Assumption 2.5.

Example 2.1. Let

$$X = -\log(U), \quad Y = -\log(1 - U), \quad U \sim \text{uniform}(0, 1).$$

It is obvious that in this case both X and Y have distribution exponential(1). So, in this case, the auxiliary function is $f(x) = 1$. Choose L such that $e^{-L} = \frac{3}{4}$ and

$$\begin{aligned} \frac{P(X > Lf(x), Y > Lf(x))}{P(X > x)} &= \frac{P(U < e^{-L}, 1 - U < e^{-L})}{P(X > x)} \\ &= \frac{P(1/4 < U < 3/4)}{P(X > x)} \\ &= \frac{1}{2 P(X > x)} \\ &\rightarrow \infty. \end{aligned}$$

Therefore, this particular choice of L does not satisfy Assumption 2.5. The distribution of (X, Y) is a special case of Example 3.4, below, which discusses certain values of L that do satisfy Assumption 2.5.

6. If, however, both X and Y are nonnegative risks, and Assumption 2.5 is strengthened to hold for all $L > 0$, then Assumptions 2.3 and 2.4 will be automatically satisfied. The proof of this follows from $\lim_{x \rightarrow \infty} f(x)/x = 0$.

7. Similar limit results are found in Lemma 2.7 of Albrecher *et al.* (2006) and Theorem 2.1 of Ko and Tang (2008). They assumed that one of the marginal distributions of the two asymptotically independent variables X and Y , say the distribution of X , is subexponential (i.e. $X \in \mathcal{S}$, where \mathcal{S} is the set of all subexponential distributions), and worked on finding conditions for the tail equivalence of the marginal distribution of X and the sum $X + Y$. Our assumptions are different. We assume that one of the marginal distributions of the two asymptotically independent variables X and Y , say the distribution of X , belongs to the domain of attraction of the Gumbel distribution, i.e. $X \in \text{MDA}(\Lambda)$. We do not assume that the marginal distribution of X is subexponential.

In examples where the marginal distributions of the two asymptotically i.i.d. random variables X and Y belong to the class $\text{MDA}(\Lambda) \cap \mathcal{S}$, an issue is the relative strength of our conditions versus those of Theorem 2.1 of Ko and Tang (2008). We cannot show that either set of conditions implies the other. Below we present an example which satisfies our set of conditions, but does not satisfy the set of conditions given in Theorem 2.1 of Ko and Tang (2008). Thus, our set of conditions is not stronger.

Example 2.2. Suppose that $X = \exp(X_1)$ and $Y = \exp(X_2)$, where (X_1, X_2) is bivariate normal with correlation $\rho \in (0, 1)$. For simplicity, assume that each X_i has mean 0 and variance 1. It is well known that the lognormal distribution belongs to the class $MDA(\Lambda) \cap \mathcal{L}$. In Example 3.5, below, we show that (X, Y) satisfies our set of conditions. Here we show that this example does not satisfy Assumption 2.1 of Ko and Tang (2008), i.e. for all $x^* > 0$,

$$\limsup_{x \rightarrow \infty} \sup_{x^* \leq t \leq x} \frac{P(Y > x - t \mid X = t)}{P(Y > x - t)} = \infty. \tag{2.5}$$

From the exchangeability of X and Y , it is obvious that (2.5) holds even if the role of X and Y is interchanged. Note that

$$\begin{aligned} \sup_{x^* \leq t \leq x} \frac{P(Y > x - t \mid X = t)}{P(Y > x - t)} &= \sup_{x^* \leq t \leq x} \frac{\bar{\Phi}((\log(x - t) - \rho \log t)/\sqrt{1 - \rho^2})}{\bar{\Phi}(\log(x - t))} \\ &\geq \frac{\bar{\Phi}((\log(x/2) - \rho \log(x/2))/\sqrt{1 - \rho^2})}{\bar{\Phi}(\log(x/2))} \\ &= \frac{\bar{\Phi}((1 - \rho) \log(x/2)/\sqrt{1 - \rho^2})}{\bar{\Phi}(\log(x/2))} \\ &\rightarrow \infty. \end{aligned} \tag{2.6}$$

The inequality above follows from choosing x large enough so that $x/2 > x^*$ and setting $t = x/2$. The last convergence follows from the fact that the normal distribution belongs to the class $MDA(\Lambda)$ and, hence, $\bar{\Phi}$ is $-\infty$ -varying (see Resnick (1987, p. 53)). Note that $0 < \rho < 1$ entails $(1 - \rho)/\sqrt{1 - \rho^2} < 1$. Hence, from (2.6), it is obvious that (2.5) holds.

2.1.4. *Proof of Theorem 2.1.* We prove Theorem 2.1 using a proposition and a lemma, which we prove first. Note that we do not need the assumption that the marginal distributions are subexponential, which is a necessary condition in the case where X and Y are independent.

Proposition 2.1. *Under Assumptions 2.1 and 2.3, we have*

$$\lim_{n \rightarrow \infty} P(Y \leq a_n z \mid X > a_n x + b_n) = \mathbf{1}_{\{z > 0\}}, \quad z \neq 0, x \in \mathbb{R},$$

and from Assumptions 2.1 and 2.4, we have

$$\lim_{n \rightarrow \infty} P(X \leq a_n z \mid Y > a_n x + b_n) = \mathbf{1}_{\{z > 0\}}, \quad z \neq 0, x \in \mathbb{R}.$$

Proof. The self-neglecting property of the auxiliary function f , i.e.

$$\lim_{t \rightarrow \infty} \frac{f(t + xf(t))}{f(t)} = 1, \quad x \in \mathbb{R},$$

implies that

$$\lim_{t \rightarrow \infty} P(|Y| > zf(t) \mid X > t) = \lim_{t \rightarrow \infty} P(|Y| > zf(t) \mid X > t + xf(t)). \tag{2.7}$$

Hence, by noting that $a_n = f(b_n)$ and $\lim_{n \rightarrow \infty} b_n = \infty$, the result follows from (2.7). The second part is proved similarly.

Lemma 2.1. (i) *Assumptions 2.1–2.3 imply that the sequence of measures*

$$n P(a_n^{-1}(X - b_n, Y) \in (dx, dy))$$

converges vaguely on $([-M, \infty) \times [-\infty, \infty])$ as $n \rightarrow \infty$ to the limit measure $m_{1,\infty}(dx, dy) = e^{-x} dx \varepsilon_0(dy)$, for some $M > L$ (from Assumption 2.5) such that $-M$ is a continuity point of $(X - b_n)/a_n$ for all n .

(ii) *Assumptions 2.1, 2.2, and 2.4 imply that the sequence of measures*

$$n P(a_n^{-1}(Y - b_n, X) \in (dx, dy))$$

converges vaguely on $([-M, \infty) \times [-\infty, \infty])$ as $n \rightarrow \infty$ to the limit measure $m_{2,\infty}(dx, dy) = ce^{-x} dx \varepsilon_0(dy)$ for some $M > L$ (from Assumption 2.5) such that $-M$ is a continuity point of $(Y - b_n)/a_n$ for all n .

Remark 2.1. Since all the discontinuity points of $(X - b_n)/a_n$ for all n are countable, the choice of such an $M > L$ is not a problem. Moreover, the M in Lemma 2.1(i) and (ii) may be chosen to be the same.

Proof of Lemma 2.1. We consider convergence of the measures evaluated on certain relatively compact regions which guarantee vague convergence.

Region 1: $(x, \infty) \times [-\infty, y]$, $x > -M$, and $y \neq 0$. As $n \rightarrow \infty$,

$$\begin{aligned} n P\left(\frac{X - b_n}{a_n} > x, \frac{Y}{a_n} \leq y\right) &= n P\left(\frac{X - b_n}{a_n} > x\right) P\left(\frac{Y}{a_n} \leq y \mid \frac{X - b_n}{a_n} > x\right) \\ &\rightarrow e^{-x} \mathbf{1}_{\{y > 0\}} \\ &= m_{1,\infty}((x, \infty) \times [-\infty, y]), \end{aligned}$$

by Proposition 2.1.

Region 2: $[-M, x] \times (y, \infty]$, $x > -M$, and $y \neq 0$. Since $-M$ is a continuity point of $(X - b_n)/a_n$ for all n , as $n \rightarrow \infty$,

$$\begin{aligned} n P\left(-M \leq \frac{X - b_n}{a_n} \leq x, \frac{Y}{a_n} > y\right) &= n P\left(-M < \frac{X - b_n}{a_n} \leq x, \frac{Y}{a_n} > y\right) \\ &= n P\left(\frac{X - b_n}{a_n} > -M, \frac{Y}{a_n} > y\right) - n P\left(\frac{X - b_n}{a_n} > x, \frac{Y}{a_n} > y\right) \\ &= n P\left(\frac{X - b_n}{a_n} > -M\right) P\left(\frac{Y}{a_n} > y \mid \frac{X - b_n}{a_n} > -M\right) \\ &\quad - n P\left(\frac{X - b_n}{a_n} > x\right) P\left(\frac{Y}{a_n} > y \mid \frac{X - b_n}{a_n} > x\right) \\ &\rightarrow (e^M - e^{-x}) \mathbf{1}_{\{y < 0\}} \\ &= m_{1,\infty}([-M, x] \times (y, \infty]), \end{aligned}$$

by Assumption 2.1 and Proposition 2.1.

Arguments for convergence on the following regions follow in a similar fashion using Proposition 2.1:

Region 3: $(x, \infty] \times (y, \infty], x > -M, \text{ and } y \neq 0,$

Region 4: $[-M, x] \times [-\infty, y], x > -M, \text{ and } y \neq 0.$

This concludes the proof of vague convergence on part (i).

The proof of part (ii) is similar, only note that if $c = 0,$ we do not need Assumption 2.4. In this case, note that the limit measure $m_{2,\infty}(dx, dy)$ is a zero measure. Also, note that, using Assumptions 2.1 and 2.2, we obtain

$$n P\left(\frac{Y - b_n}{a_n} \geq -M\right) \rightarrow ce^M = 0,$$

which is enough to prove the convergence in this case.

This leads to a formal statement of the main result.

Theorem 2.2. *Under Assumptions 2.1–2.5,*

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = 1 + c.$$

Proof. Choose M to be as in Remark 2.1. We split $P(X + Y > b_n)$ as

$$\begin{aligned} P(X + Y > b_n) &= P(X + Y > b_n, X > b_n - Ma_n) + P(X + Y > b_n, Y > b_n - Ma_n) \\ &\quad - P(X + Y > b_n, X > b_n - Ma_n, Y > b_n - Ma_n) \\ &\quad + P(X + Y > b_n, X \leq b_n - Ma_n, Y \leq b_n - Ma_n). \end{aligned} \tag{2.8}$$

Using Assumption 2.1 and (2.4), we obtain

$$\begin{aligned} &n P(X + Y > b_n, X > b_n - Ma_n, Y > b_n - Ma_n) \\ &\leq n P(X > b_n - Ma_n, Y > b_n - Ma_n) \\ &= n P(X > b_n - Ma_n) \frac{P(X > b_n - Ma_n, Y > b_n - Ma_n)}{P(X > b_n - Ma_n)} \\ &\rightarrow e^M \cdot 0 \\ &= 0, \end{aligned} \tag{2.9}$$

since $b_n - Ma_n \rightarrow \infty.$ Now, consider the convergence of the last term of (2.8) multiplied by $n:$

$$\begin{aligned} P(X + Y > b_n, X \leq b_n - Ma_n, Y \leq b_n - Ma_n) &\leq n P(X > Ma_n, Y > Ma_n) \\ &\sim \frac{n P(X > Mf(b_n), Y > Mf(b_n))}{P(X > b_n)} \\ &\leq \frac{P(X > Lf(b_n), Y > Lf(b_n))}{P(X > b_n)} \\ &\rightarrow 0, \end{aligned} \tag{2.10}$$

by (2.1) and Assumption 2.5.

To deal with the first term of (2.8) multiplied by $n,$ we first define a function $T : [-M, \infty] \times [-\infty, \infty] \mapsto (\infty, \infty]$ by

$$T(x, y) = \begin{cases} x + y & \text{if } y > -\infty, \\ 0 & \text{if } y = -\infty, \end{cases}$$

and, hence,

$$\begin{aligned} n P(X + Y > b_n, X > b_n - Ma_n) \\ = n P(a_n^{-1}(X - b_n, Y) \in T^{\leftarrow}((0, \infty]) \cap [(-M, \infty] \times \{0\})). \end{aligned} \tag{2.11}$$

Note that every set in the space $[-M, \infty] \times [-\infty, \infty]$ is relatively compact, and, hence, so is $T^{\leftarrow}((0, \infty]) \cap [(-M, \infty] \times \{0\}) = S$ (say). Also, since the limit measure $m_{1,\infty}$ is concentrated on $[-M, \infty] \times \{0\}$,

$$m_{1,\infty}(\delta S) = m_{1,\infty}(\delta S \cap [[-M, \infty] \times \{0\}]) = m_{1,\infty}(\{0\} \times \{0\}) = 0. \tag{2.12}$$

Hence, using Lemma 2.1, (2.11), and (2.12), we obtain

$$n P(X + Y > b_n, X > b_n - Ma_n) \rightarrow m_{1,\infty}(S) = 1. \tag{2.13}$$

Similarly,

$$n P(X + Y > b_n, Y > b_n - Ma_n) \rightarrow m_{2,\infty}(S) = c. \tag{2.14}$$

Hence, using (2.8), (2.9), (2.10), (2.13), and (2.14), we obtain

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = \lim_{n \rightarrow \infty} n P(X + Y > b_n) = 1 + c,$$

and this completes the proof.

One immediate application of Theorem 2.2 is to the subexponential family of distributions denoted by \mathcal{S} . The class $MDA(\Lambda) \cap \mathcal{S}$ has been studied in Embrechets *et al.* (1997, p. 149) and several sufficient conditions for belonging to this class are given in Goldie and Resnick (1988). Corollary 2.1, below, gives an additional sufficient condition and follows directly from Theorem 2.2. Example 3.2, below, exhibits a distribution which satisfies the conditions of this corollary.

Corollary 2.1. *Suppose that $F \in MDA(\Lambda)$ with auxiliary function $f(x)$ as described in Assumption 2.1. Suppose also that $\lim_{x \rightarrow \infty} f(x) = \infty$ and that, for some $L > 0$,*

$$\lim_{x \rightarrow \infty} \frac{[\bar{F}(Lf(x))]^2}{\bar{F}(x)} = 0. \tag{2.15}$$

Then, for X and Y i.i.d. with common distribution F , we have, as $x \rightarrow \infty$,

$$P(X + Y > x) \sim 2 P(X > x),$$

and, therefore, if F concentrates on $[0, \infty)$, $F \in \mathcal{S}$.

Following comment 3 in Section 2.1.3, it is enough to check that (2.15) holds for any $\tilde{f}(x)$ satisfying $\tilde{f}(x) \sim f(x)$. Note also that it is natural to add the assumption $f(x) \rightarrow \infty$, since if $F \in MDA(\Lambda) \cap \mathcal{S}$ then necessarily $f(x) \rightarrow \infty$ (see Goldie and Resnick (1988)).

2.2. Asymptotic tail probability for the sum of more than two nonnegative random variables

Suppose that among the risks X_1, X_2, \dots, X_d there is no heavier tail than X_1 in the sense that it is not true that

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_i(x)}{\bar{F}_1(x)} = \infty, \quad i = 2, \dots, d.$$

Assume that X_1 satisfies Assumption 2.1 and that X_1, X_2, \dots, X_d pairwise satisfy Assumptions 2.3 and 2.4 with the auxiliary function $f(\cdot)$ of X_1 . By this we mean for all pairs $1 \leq i \neq j \leq d$, and, for $t > 0$,

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_j > tf(x), X_i > x)}{\mathbb{P}(X_i > x)} = 0,$$

which implies that

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_j > tf(x), X_i > x)}{\mathbb{P}(X_1 > x)} = 0. \tag{2.16}$$

Also, suppose that the risks X_1, X_2, \dots, X_d pairwise satisfy Assumption 2.5 with auxiliary function $f(\cdot)$ of X_1 so that, for $1 \leq i < j \leq d$, there exists some $L_{ij} > 0$ such that either

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{\mathbb{P}(X_i > x)} = 0$$

or

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{\mathbb{P}(X_j > x)} = 0.$$

In either case we have, for $1 \leq i < j \leq d$ and some $L_{ij} > 0$,

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{\mathbb{P}(X_1 > x)} = 0. \tag{2.17}$$

Under the additional assumption of nonnegativity, Theorem 2.2 can be extended to more than two risks.

Corollary 2.2. *Assume that X_1, X_2, \dots, X_d are nonnegative random variables that pairwise satisfy Assumptions 2.3–2.5 with auxiliary function $f(\cdot)$ of X_1 . Moreover, the distribution of X_1 satisfies Assumption 2.1. Suppose that*

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_i > x)}{\mathbb{P}(X_1 > x)} = c_i \in [0, \infty), \quad i = 2, 3, \dots, d. \tag{2.18}$$

Define $S_j = X_1 + X_2 + \dots + X_j$, $1 \leq j \leq d$. We have, for $x \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} n \mathbb{P}(S_d > a_n x + b_n) = \left(1 + \sum_{i=2}^d c_i \right) e^{-x},$$

and, hence,

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}(S_d > x)}{\mathbb{P}(X_1 > x)} = 1 + \sum_{i=2}^d c_i.$$

Remark 2.2. (i) *Asymptotic independence of the random variables.* Suppose that, for all i , $c_i \in (0, \infty)$. Then, for any $1 \leq i \neq j \leq d$, the pair (X_i, X_j) is asymptotically independent by comment 2 in Section 2.1.3. Since the random variables are pairwise asymptotically independent, they are also asymptotically independent (see Resnick (1987, p. 291)).

(ii) *Nonnegativity of the random variables.* The only additional assumption added to the list in Section 2.1.1 is that the random variables are nonnegative.

(iii) *Relaxation.* We have shown in (2.16) and (2.17) that pairwise satisfaction of Assumptions 2.3–2.5 implies that, for $1 \leq i \neq j \leq d$ and $t > 0$,

$$\lim_{x \rightarrow \infty} \frac{P(X_j > tf(x), X_i > x)}{P(X_1 > x)} = 0,$$

and, for $1 \leq i < j \leq d$, there exists an $L_{ij} > 0$ such that

$$\lim_{x \rightarrow \infty} \frac{P(X_j > L_{ij}f(x), X_i > L_{ij}f(x))}{P(X_1 > x)} = 0.$$

We will show that these conditions are actually enough to obtain the desired conclusion.

Proof of Corollary 2.2. We prove the result by induction under relaxation, Remark 2.2(iii). The base case for the induction, $d = 2$, has already been proved in Theorem 2.2, so suppose that the result is true for $d = k \geq 2$. We have

$$\lim_{n \rightarrow \infty} n P(S_k > a_n x + b_n) = \left(1 + \sum_{i=2}^k c_i\right) e^{-x}$$

and

$$\lim_{x \rightarrow \infty} \frac{P(S_k > x)}{P(X_1 > x)} = 1 + \sum_{i=2}^k c_i.$$

Therefore, we have

$$\lim_{x \rightarrow \infty} \frac{P(X_{k+1} > x)}{P(S_k > x)} = \frac{c_{k+1}}{1 + \sum_{i=2}^k c_i} \in [0, \infty). \tag{2.19}$$

We will use Theorem 2.2 with $X = S_k$ and $Y = X_{k+1}$. It remains to check that Assumptions 2.1–2.5 hold for this case. To show that Assumption 2.1 holds, note that S_k is tail equivalent to X_1 and use the fact that $F \in \text{MDA}(\Lambda)$ is closed under tail equivalence. Assumption 2.2 has already been shown to hold in (2.19).

Note that, from the induction hypothesis,

$$P(S_k > x) \sim P\left(\bigcup_{i=1}^k (X_i > x)\right),$$

and, from the positivity of the risks,

$$[S_k > x] \supseteq \bigcup_{i=1}^k [X_i > x].$$

From these two facts, it easily follows that

$$\lim_{x \rightarrow \infty} \frac{P((S_k > x) \cap (\bigcup_{i=1}^k (X_i > x))^c)}{P(S_k > x)} = 0. \tag{2.20}$$

Since S_k and X_1 are tail equivalent, by Resnick (1971b), the auxiliary function $\tilde{f}(\cdot)$ of S_k is asymptotically equal to the auxiliary function $f(\cdot)$ of X_1 . Therefore, given $\varepsilon \in (0, 1)$, there

exists a T such that, for all $x > T$, $\tilde{f}(x) > \varepsilon f(x)$. We now check that Assumption 2.3 holds. For $t > 0$ and $x > T$, using (2.20), as $x \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(|X_{k+1}| > t\tilde{f}(x) \mid S_k > x) &\leq \mathbb{P}(X_{k+1} > t\varepsilon f(x) \mid S_k > x) \\ &= \frac{\mathbb{P}(X_{k+1} > t\varepsilon f(x), S_k > x)}{\mathbb{P}(S_k > x)} \\ &\sim \frac{\mathbb{P}(X_{k+1} > t\varepsilon f(x), S_k > x, \bigcup_{i=1}^k \{X_i > x\})}{\mathbb{P}(S_k > x)} \\ &\leq \frac{\mathbb{P}(X_{k+1} > t\varepsilon f(x), \bigcup_{i=1}^k \{X_i > x\})}{\mathbb{P}(S_k > x)} \\ &\leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > t\varepsilon f(x), X_i > x)}{(1 + \sum_{i=2}^k c_i) \mathbb{P}(X_1 > x)} \\ &\rightarrow 0, \end{aligned}$$

by (2.16).

If $c_{k+1} = 0$, following comment 4 in Section 2.1.3, there is no need to check that Assumption 2.4 holds. So, suppose that $c_{k+1} > 0$. Then, for any $t > 0$, as $x \rightarrow \infty$,

$$\begin{aligned} \mathbb{P}(|S_k| > t\tilde{f}(x) \mid X_{k+1} > x) &\leq \mathbb{P}(S_k > t\varepsilon f(x) \mid X_{k+1} > x) \\ &\leq \sum_{i=1}^k \mathbb{P}\left(X_i > \frac{t\varepsilon f(x)}{k} \mid X_{k+1} > x\right) \\ &= \sum_{i=1}^k \frac{\mathbb{P}(X_i > t\varepsilon f(x)/k, X_{k+1} > x)}{\mathbb{P}(X_1 > x)} \frac{\mathbb{P}(X_1 > x)}{\mathbb{P}(X_{k+1} > x)} \\ &\rightarrow 0. \end{aligned}$$

We know from the statement of Corollary 2.2 that the random variables pairwise satisfy Assumption 2.5 with auxiliary function $f(\cdot)$ of X_1 . Thus, for $1 \leq i < j \leq d$, (2.17) holds. We check Assumption 2.5 with $L = kL_{\max}/\varepsilon$, where $L_{\max} = \max_{1 \leq i \leq k} L_{i,k+1}$ (recall (2.17)). Then, for sufficiently large x , using $\tilde{f}(\cdot)$ as the auxiliary function of S_k ,

$$\begin{aligned} \frac{\mathbb{P}(X_{k+1} > L\tilde{f}(x), S_k > L\tilde{f}(x))}{\mathbb{P}(S_k > x)} &\leq \frac{\mathbb{P}(X_{k+1} > L\varepsilon f(x), S_k > L\varepsilon f(x))}{\mathbb{P}(S_k > x)} \\ &\leq \frac{\mathbb{P}(X_{k+1} > kL_{\max}f(x), \bigcup_{i=1}^k \{X_i > L_{\max}f(x)\})}{\mathbb{P}(S_k > x)} \\ &\leq \frac{\mathbb{P}(X_{k+1} > L_{i,k+1}f(x), \bigcup_{i=1}^k \{X_i > L_{i,k+1}f(x)\})}{\mathbb{P}(S_k > x)} \\ &\leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > L_{i,k+1}f(x), X_i > L_{i,k+1}f(x))}{\mathbb{P}(S_k > x)} \\ &\sim \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > L_{i,k+1}f(x), X_i > L_{i,k+1}f(x))}{(1 + \sum_{i=2}^k c_i) \mathbb{P}(X_1 > x)} \\ &\rightarrow 0, \end{aligned}$$

by (2.17). This completes the induction proof.

3. Examples

In this section we show a few of the many models that satisfy Assumptions 2.1–2.5. In all the examples, both X and Y are nonnegative random variables. It is straightforward to extend these examples to the d -dimensional case and to show that the assumptions of Corollary 2.2 are satisfied.

Our conditions are only sufficient, and we exhibit one example where our conditions do not hold, but the tail equivalence in Theorem 2.2 does hold. Finding a necessary and sufficient condition for the conclusion of Theorem 2.2 is still an open but subtle and difficult issue.

Example 3.1. Suppose that $X_1, X_2,$ and X_3 are i.i.d. with common distribution F , where, for $\alpha > 1$,

$$\bar{F}(x) = \begin{cases} \exp(-(\log x)^\alpha) & \text{if } x > 1, \\ 1 & \text{if } x \leq 1. \end{cases}$$

Define, $X = X_1 \wedge X_2$ and $Y = X_2 \wedge X_3$. It is easy to check that X and Y are identically distributed with common distribution F_1 , where $\bar{F}_1(x) = \exp(-2(\log x)^\alpha), x > 1$. It can be checked that F_1 is a von Mises function, that is, it satisfies

$$\frac{\bar{F}_1 F_1''}{(F_1')^2} \rightarrow -1,$$

a sufficient condition for $F_1 \in \text{MDA}(\Delta)$, and

$$f(x) = \frac{\bar{F}_1(x)}{F_1'(x)} = \frac{x}{2\alpha(\log x)^{\alpha-1}}, \quad x > 1,$$

serves as an auxiliary function (see Resnick (1987, p. 40)). As (2.1) is obvious, we know that Assumption 2.1 is satisfied. Checking that Assumption 2.2 holds is straightforward, so consider Assumption 2.3. Fix $t > 0$, recall that $f(x)/x \rightarrow 0$, and note that, as $x \rightarrow \infty$,

$$\begin{aligned} \frac{\text{P}(X > x, Y > tf(x))}{\text{P}(X > x)} &= \frac{\text{P}(X_1 > x, X_2 > x \vee tf(x), X_3 > tf(x))}{\text{P}(X_1 > x, X_2 > x)} \\ &\sim \frac{\text{P}(X_1 > x, X_2 > x, X_3 > tf(x))}{\text{P}(X_1 > x, X_2 > x)} \\ &= \text{P}(X_3 > tf(x)) \\ &\rightarrow 0, \end{aligned}$$

since $f(x) \rightarrow \infty$. Assumption 2.4 is verified similarly. For Assumption 2.5, we have, for $L = 1$,

$$\begin{aligned} \frac{\text{P}(X > f(x), Y > f(x))}{\text{P}(X > x)} &= \frac{\text{P}(X_1 > f(x), X_2 > f(x), X_3 > f(x))}{\text{P}(X_1 > x, X_2 > x)} \\ &= \frac{\bar{F}(f(x))^3}{\bar{F}(x)^2} \\ &= \exp(-[3(\log f(x))^\alpha - 2(\log x)^\alpha]) \\ &= \exp\left(-2(\log x)^\alpha \left[\frac{3}{2} \left(\frac{\log f(x)}{\log x}\right)^\alpha - 1\right]\right) \\ &= \exp\left(-2(\log x)^\alpha \left[\frac{3}{2} \left(1 - \frac{\log(2\alpha(\log x)^{\alpha-1})}{\log x}\right)^\alpha - 1\right]\right). \end{aligned} \tag{3.1}$$

Since the exponent in (3.1) converges to $-\infty$ as $x \rightarrow \infty$, Assumption 2.5 is satisfied. Hence, the pair (X, Y) satisfies Assumptions 2.1–2.5.

Example 3.2. Suppose that X and Y are i.i.d. with common distribution F , where, for $\alpha > 1$,

$$\bar{F}(x) = \begin{cases} \exp(-(\log x)^\alpha) & \text{if } x > 1, \\ 1 & \text{if } x \leq 1. \end{cases}$$

As in Example 3.1, we can check that the subexponentiality condition, (2.15), holds with $L = 1$ and that, by Corollary 2.1, F is subexponential. Hence,

$$P(X + Y > x) \sim 2P(X > x).$$

Example 3.3. Suppose that $X \sim \text{lognormal}(\mu, \sigma^2)$ and that $Y = e^{2\mu}/X$, so that $X \stackrel{D}{=} Y$, where ‘ $\stackrel{D}{=}$ ’ denotes equality in distribution. We check that Assumptions 2.1–2.5 hold for the pair (X, Y) . The distribution $\text{lognormal}(\mu, \sigma^2)$ belongs to the maximal domain of attraction of the Gumbel distribution and its mean excess function $e(x)$ has the form (see Embrechts *et al.* (1997, pp. 147, 161))

$$e(x) = \frac{\sigma^2 x}{\log x - \mu} (1 + o(1)).$$

As (2.1) is obvious, we know that Assumption 2.1 holds. Following comment 3 in Section 2.1.3 and considering the form of $e(x)$, we take the auxiliary function to be

$$f(x) = \frac{\sigma^2 x}{\log x - \mu}.$$

Since X and Y are identically distributed, Assumption 2.2 holds with $c = 1$. To verify Assumption 2.3, fix $t > 0$ and note that, as $x \rightarrow \infty$,

$$\frac{P(X > x, Y > tf(x))}{P(X > x)} = \frac{P(X > x, e^{2\mu}/X > tf(x))}{P(X > x)} \rightarrow 0,$$

since $f(x) \rightarrow \infty$. Assumption 2.4 is verified similarly. For Assumption 2.5, choosing $L = 1$, we have, as $x \rightarrow \infty$,

$$\frac{P(X > f(x), Y > f(x))}{P(X > x)} = \frac{P(X > f(x), e^{2\mu}/X > f(x))}{P(X > x, Y > x)} \rightarrow 0.$$

Hence, by Theorem 2.2,

$$P(X + Y > x) \sim 2P(X > x).$$

Example 3.4. Example 3.3 is a special case of a more general phenomenon. Suppose that $F \in \text{MDA}(\Lambda)$ with auxiliary function $f(x)$ having the property

$$\liminf_{x \rightarrow \infty} f(x) = \delta > 0. \tag{3.2}$$

Assume that the support of F is a subset of $[0, \infty)$, that $x_0 = \sup\{x : F(x) < 1\} = \infty$, and that $x_1 = \inf\{x : F(x) > 0\} = 0$. Distributions satisfying these conditions include the exponential, gamma, and lognormal. Define $X = F^{\leftarrow}(U)$ and $Y = F^{\leftarrow}(1 - U)$, where $U \sim \text{uniform}(0, 1)$. We check that the pair (X, Y) satisfies Assumptions 2.1–2.5.

Checking that Assumption 2.2 holds is easy since X and Y are identically distributed. To verify Assumption 2.3, fix $t > 0$ and define $\varepsilon_t = F(t\delta/2)$. Since $x_1 = 0$, we have $\varepsilon_t > 0$. Then, for large x making $f(x) > \delta/2$, we have

$$\begin{aligned} \frac{P(X > x, Y > tf(x))}{P(X > x)} &= \frac{P(U > F(x), 1 - U > F(tf(x)))}{P(X > x)} \\ &\leq \frac{P(U > F(x), 1 - U > \varepsilon_t)}{P(X > x)} \\ &= \frac{P(U > F(x), U < 1 - \varepsilon_t)}{P(X > x)} \\ &\rightarrow 0, \end{aligned}$$

since $F(x) \rightarrow 1$ and $x_0 = \infty$. Assumption 2.4 is verified similarly. To verify Assumption 2.5, choose L such that $F(L\delta/2) > \frac{1}{2}$ and, for sufficiently large x ,

$$\begin{aligned} \frac{P(X > Lf(x), Y > Lf(x))}{P(X > x)} &\leq \frac{P(X > L\delta/2, Y > L\delta/2)}{P(X > x)} \\ &= \frac{P(U > F(L\delta/2), 1 - U > F(L\delta/2))}{P(X > x)} \\ &= 0. \end{aligned}$$

Hence, the pair (X, Y) satisfies Assumptions 2.1–2.5 and, by Theorem 2.2,

$$P(X + Y > x) \sim 2P(X > x).$$

In this example, if $\lim_{x \rightarrow \infty} f(x) = \infty$, we do not need the condition that $x_1 = 0$.

Remark 3.1. Note that in Examples 3.3 and 3.4 a comonotonic dependence structure is used.

Example 3.5. Suppose that $X = \exp(X_1)$ and $Y = \exp(X_2)$, where (X_1, X_2) is bivariate normal with correlation $\rho \in [-1, 1)$. For simplicity, assume that each X_i has mean μ and variance $\sigma^2 > 0$. This example is extensively considered in Asmussen and Rojas-Nandayapa (2008). We have already considered the case in which $\rho = -1$ in Example 3.3, so here we consider $\rho \in (-1, 1)$.

Assumptions 2.1 and 2.2 are easily verified. Following the same reasoning as in Example 3.3, we take the auxiliary function to be

$$f(x) = \frac{\sigma^2 x}{\log x - \mu}.$$

Observe that

$$\begin{aligned} \frac{\log f(x) - \mu}{\sigma} &= \frac{\log(\sigma^2 x / (\log x - \mu)) - \mu}{\sigma} \\ &= \frac{\log x - \mu}{\sigma} - \frac{1}{\sigma} \log\left(\frac{\log x - \mu}{\sigma^2}\right) \\ &= \left(\frac{\log x - \mu}{\sigma}\right)(1 + o(1)). \end{aligned} \tag{3.3}$$

For Assumption 2.3, we have, for $t > 0$, as $x \rightarrow \infty$,

$$\begin{aligned} & \frac{P(X > x, Y > tf(x))}{P(X > x)} \\ &= \frac{P(X_1 > \log x, X_2 > \log tf(x))}{P(X_1 > \log x)} \\ &\leq \frac{P(X_1 + X_2 > \log x + \log(tf(x)))}{P(X_1 > \log x)} \\ &= \bar{\Phi}\left(\frac{1}{\sqrt{2\sigma^2(1+\rho)}}(\log x + \log(tf(x)) - 2\mu)\right) \bigg/ \bar{\Phi}\left(\frac{\log x - \mu}{\sigma}\right) \\ &= \bar{\Phi}\left(\frac{1}{\sqrt{2(1+\rho)}}\left(\frac{\log x - \mu}{\sigma} + \frac{\log f(x) - \mu}{\sigma} + \frac{\log t}{\sigma}\right)\right) \bigg/ \bar{\Phi}\left(\frac{\log x - \mu}{\sigma}\right) \\ &= \bar{\Phi}\left(\frac{2}{\sqrt{2(1+\rho)}}\left(\frac{\log x - \mu}{\sigma}\right)(1 + o(1))\right) \bigg/ \bar{\Phi}\left(\frac{\log x - \mu}{\sigma}\right) \\ &\rightarrow 0, \end{aligned}$$

where we have used (3.3) and the fact that $\Phi \in \text{MDA}(\Lambda)$, and, therefore, $\bar{\Phi}$ is $-\infty$ -varying (see Resnick (1987, p. 53)). Note that $\rho < 1$ entails $2/\sqrt{2(1+\rho)} > 1$.

For Assumption 2.5, choose $L = 1$. As $x \rightarrow \infty$, we have, using (3.3),

$$\begin{aligned} \frac{P(X > f(x), Y > f(x))}{P(X > x)} &= \frac{P(X_1 > \log f(x), X_2 > \log f(x))}{P(X_1 > \log x)} \\ &\leq \frac{P(X_1 + X_2 > 2 \log f(x))}{P(X_1 > \log x)} \\ &= \bar{\Phi}\left(\frac{2(\log f(x) - \mu)}{\sqrt{2\sigma^2(1+\rho)}}\right) \bigg/ \bar{\Phi}\left(\frac{\log x - \mu}{\sigma}\right) \\ &= \bar{\Phi}\left(\frac{2}{\sqrt{2(1+\rho)}}\left(\frac{\log x - \mu}{\sigma}\right)(1 + o(1))\right) \bigg/ \bar{\Phi}\left(\frac{\log x - \mu}{\sigma}\right) \\ &\rightarrow 0. \end{aligned}$$

Example 3.6. Let X_1 and X_2 be i.i.d. with common distribution $H \in \text{MDA}(\Lambda)$, whose auxiliary function $f_1(\cdot)$ satisfies (3.2) and whose right endpoint is infinite. Also, suppose that $F \in \text{MDA}(\Lambda)$ with auxiliary function $f_2(\cdot)$, concentrates on $[0, \infty)$, $x_0 = \sup\{x : F(x) < 1\} = \infty$, and that $x_1 = \inf\{x : F(x) > 0\} = 0$. The auxiliary function $f_2(\cdot)$ satisfies (3.2). Define

$$X = F^{\leftarrow}(U) \wedge X_1 \quad \text{and} \quad Y = F^{\leftarrow}(1 - U) \wedge X_2,$$

where U is a uniformly distributed random variable on $(0, 1)$ which is independent of (X_1, X_2) .

From Proposition 1.4 of Resnick (1987, p. 43), the distribution of X belongs to the maximal domain of attraction of the Gumbel distribution with auxiliary function

$$f(x) = \frac{f_1(x)f_2(x)}{f_1(x) + f_2(x)}.$$

Hence,

$$\limsup_{x \rightarrow \infty} \frac{1}{f(x)} \leq \limsup_{x \rightarrow \infty} \frac{1}{f_1(x)} + \limsup_{x \rightarrow \infty} \frac{1}{f_2(x)} < \infty,$$

and, thus,

$$\liminf_{x \rightarrow \infty} f(x) > 0.$$

Also, note that

$$P(X > x) = P(U > F(x), X_1 > x) = P(U > F(x))P(X_1 > x) = \bar{F}(x)\bar{H}(x)$$

and

$$P(Y > x) = P(1 - U > F(x), X_2 > x) = P(1 - U > F(x))P(X_2 > x) = \bar{F}(x)\bar{H}(x).$$

Arguing as in Example 3.4, we can show that the pair (X, Y) satisfies Assumptions 2.1–2.5.

Example 3.7. Here we present an example of a distribution for (X, Y) that does not satisfy Assumptions 2.1–2.5, but for which the asymptotic behavior is the same as in Theorem 2.2. Suppose that X and Y are i.i.d. with common distribution F , where

$$\bar{F}(x) = \exp(-x^\alpha), \quad \alpha \in (0, 1), x > 0.$$

This distribution has been extensively studied in Rootzén (1986) and satisfies $F \in \text{MDA}(\Lambda) \cap \mathcal{S}$. Since it is subexponential,

$$P(X + Y > x) \sim 2P(X > x).$$

However, this distribution does not satisfy Assumption 2.5.

Since F is a von Mises function, we may take the auxiliary function to be

$$f(x) = \frac{\bar{F}(x)}{F'(x)} = \frac{x^{1-\alpha}}{\alpha}.$$

Assumption 2.5 is not satisfied for any $L > 0$, since, for any $L > 0$, as $x \rightarrow \infty$,

$$\begin{aligned} \frac{P(X > Lf(x), Y > Lf(x))}{P(X > x)} &= \frac{[\bar{F}(Lf(x))]^2}{\bar{F}(x)} \\ &= \frac{\exp(-2[Lf(x)]^\alpha)}{\exp(-x^\alpha)} \\ &= \frac{\exp(-2(L/\alpha)^\alpha x^{\alpha(1-\alpha)})}{\exp(-x^\alpha)} \\ &= \exp\left(x^\alpha \left(1 - 2\left(\frac{L}{\alpha}\right)^\alpha x^{-\alpha^2}\right)\right) \\ &\rightarrow \infty. \end{aligned}$$

This also shows that criterion (2.15) for $F \in \mathcal{S}$ is sufficient but not necessary.

4. Linear combinations of random variables with nonnegative coefficients

In this section we study linear combinations of the risks X and Y with nonnegative coefficients. We consider two cases: (i) the distributions of X and Y are tail equivalent, and (ii) the distributions of X and Y lack tail equivalence. We explicitly give the asymptotic tail behavior of the linear combinations of risks in the tail-equivalent case and also in one special case where tail equivalence is absent. We note that we cannot expect similar behavior in the two cases.

4.1. Tail-equivalent cases

4.1.1. *Linear combination of two random variables with nonnegative coefficients.*

Theorem 4.1. *Assume that (U, V) is a pair of random variables which satisfy Assumptions 2.1, 2.3, 2.4, and 2.5. Moreover, assume that Assumption 2.2 holds in the form*

$$\lim_{x \rightarrow \infty} \frac{P(V > x)}{P(U > x)} = c \in (0, \infty). \tag{4.1}$$

Define $\hat{S}_2 = a_1U + a_2V$ and $a_i \geq 0, i = 1, 2$, and set $m_2 = a_1 \vee a_2$. Then, as $x \rightarrow \infty$,

$$P(\hat{S}_2 > x) \sim P\left(U > \frac{x}{m_2}\right) (\mathbf{1}_{\{a_1=m_2\}} + c \mathbf{1}_{\{a_2=m_2\}}).$$

We assume that U and V are tail equivalent, i.e. the constant c cannot be 0 and, hence, both the marginal distributions belong to $MDA(\Lambda)$, the maximal domain of attraction of the Gumbel distribution. If $\lim_{x \rightarrow \infty} P(V > x)/P(U > x) = 0$, the asymptotic behavior of $P(a_1U + a_2V > x)$ as $x \rightarrow \infty$ can be different, as illustrated in the following example.

Example 4.1. Assume that (U, V) are i.i.d. random variables with common distribution F , which satisfy Assumptions 2.1, 2.3, 2.4, and 2.5. Define the two random vectors by $(U_1, V_1) = (U, \frac{1}{5}V)$ and $(U_2, V_2) = (U, \frac{1}{2}V)$. Both (U_1, V_1) and (U_2, V_2) satisfy Assumptions 2.1, 2.3, 2.4, and 2.5. For both pairs, $c = 0$, i.e.

$$\lim_{x \rightarrow \infty} \frac{P(V_1 > x)}{P(U_1 > x)} = 0 \quad \text{and} \quad \lim_{x \rightarrow \infty} \frac{P(V_2 > x)}{P(U_2 > x)} = 0.$$

Since (U, V) satisfies the assumptions of Theorem 4.1, we have, as $x \rightarrow \infty$,

$$P(3U_1 + 10V_1 > x) = P(3U + 2V > x) \sim P(3U > x) = P(3U_1 > x)$$

and

$$P(3U_2 + 10V_2 > x) = P(3U + 5V > x) \sim P(5V > x) = P(10V_2 > x).$$

This example illustrates that we cannot expect Theorem 4.1 to hold for the case in which $c = 0$.

Proof of Theorem 4.1. The case in which $a_1 = a_2$ is resolved by Theorem 2.2 since

$$P(a_1(U + V) > x) = P\left(U + V > \frac{x}{a_1}\right) \sim (1 + c)P\left(U > \frac{x}{a_1}\right).$$

So the interesting cases are $a_1 > a_2$ and $a_1 < a_2$. In the following, we assume that $a_1 > a_2$; the other case is treated similarly.

There is nothing to prove if $a_2 = 0$, so assume that $a_1 > a_2 > 0$, which makes $m_2 = a_1$. It suffices to verify Assumptions 2.1–2.5 for $X = U$ and $Y = a_2V/a_1$. For this definition of X and Y , we have

$$\lim_{x \rightarrow \infty} \frac{P(Y > x)}{P(X > x)} = \lim_{x \rightarrow \infty} \frac{P(a_2V/a_1 > x)}{P(U > x)} = \lim_{x \rightarrow \infty} \frac{P(V > a_1x/a_2)}{P(U > x)} = 0. \tag{4.2}$$

The last equality is true from (4.1) and the fact that the tail of any distribution in $MDA(\Lambda)$ is $-\infty$ -varying (see Resnick (1987, p. 53)). From Theorem 2.2 and (4.2), we obtain, as $x \rightarrow \infty$,

$$\begin{aligned} P(a_1U + a_2V > x) &= P\left(a_1\left(U + \frac{a_2V}{a_1}\right) > x\right) \\ &= P\left(U + \frac{a_2V}{a_1} > \frac{x}{a_1}\right) \\ &= P\left(X + Y > \frac{x}{a_1}\right) \\ &\sim P\left(U > \frac{x}{a_1}\right) \\ &= P\left(U > \frac{x}{m_2}\right)(\mathbf{1}_{\{a_1=m_2\}} + c \mathbf{1}_{\{a_2=m_2\}}). \end{aligned}$$

To complete the proof, Assumptions 2.1–2.5 must be verified for $X = U$ and $Y = a_2V/a_1$. Assumption 2.1 is assumed in the statement of the theorem and Assumption 2.2 was verified in (4.2). To verify Assumption 2.3, note that $U \in MDA(\Lambda)$ and suppose that $f(\cdot)$ is the auxiliary function of the distribution of U . By hypothesis, for $t > 0$,

$$\lim_{x \rightarrow \infty} P(|V| > tf(x) \mid U > x) = 0, \tag{4.3}$$

and, therefore, using (4.3),

$$\lim_{x \rightarrow \infty} P\left(\frac{a_2|V|}{a_1} > tf(x) \mid U > x\right) = \lim_{x \rightarrow \infty} P\left(|V| > \frac{a_1tf(x)}{a_2} \mid U > x\right) = 0.$$

Comment 4 in Section 2.1.3 implies that we do not need to verify Assumption 2.4, so we now check that Assumption 2.5 holds. For this, we have, as $x \rightarrow \infty$,

$$\begin{aligned} \frac{P(a_2V/a_1 > Lf(x), U > Lf(x))}{P(U > x)} &= \frac{P(V > a_1Lf(x)/a_2, U > Lf(x))}{P(U > x)} \\ &\leq \frac{P(V > Lf(x), U > Lf(x))}{P(U > x)} \\ &\rightarrow 0. \end{aligned}$$

This proves the case in which $a_1 > a_2$.

4.1.2. *Linear combination of more than two random variables with nonnegative coefficients.*

Corollary 4.1. *Assume that X_1, X_2, \dots, X_d are nonnegative random variables which pairwise satisfy Assumptions 2.3–2.5. Furthermore, suppose that the distribution of X_1 satisfies Assumption 2.1 and that*

$$\lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = c_i \in (0, \infty), \quad i = 1, 2, \dots, d. \tag{4.4}$$

Set $c_1 = 1$, and define, for $d > 1$, $\hat{S}_d = a_1X_1 + a_2X_2 + \dots + a_dX_d$ for $a_i \geq 0, i = 1, 2, \dots, d$. Also, define

$$m_d = \bigvee_{i=1}^d a_i \quad \text{and} \quad N_d = \sum_{\{1 \leq i \leq d: a_i=m_d\}} c_i.$$

Then

$$P(\hat{S}_d > x) \sim N_d P\left(X_1 > \frac{x}{m_d}\right), \quad x \rightarrow \infty.$$

This result is consistent with the case where X_1, X_2, \dots, X_d are i.i.d. with common distribution in $MDA(\Lambda) \cap \mathcal{F}$; see Davis and Resnick (1988).

The random variables X_1, X_2, \dots, X_d are tail equivalent and pairwise satisfy Assumption 2.3. Therefore, comment 2 in Section 2.1.3 implies pairwise asymptotic independence and, hence, by Resnick (1987, p. 291), X_1, \dots, X_d are asymptotically independent.

In the special case that the random variables are identically distributed, $N_d = |\{1 \leq i \leq d : a_i = m_d\}|$, where $|\cdot|$ is the size of a set.

Remark 4.1. It is possible to prove Corollary 4.1 using Corollary 2.2. However, in the proof it is usually difficult to verify Assumption 2.4. Note that a similar problem is carefully avoided in the proof of Theorem 4.1 through the help of comment 4 in Section 2.1.3. Though a similar comment could also be made for Corollary 2.2, it is notationally inconvenient. So, to avoid this notational difficulty, Theorem 4.1 is used in the proof.

Proof of Corollary 4.1. Proceeding by induction, note that the base case for $d = 2$ is proved in Theorem 4.1. As an induction hypothesis, suppose that the result is true for $d = k$; so, as $x \rightarrow \infty$,

$$P(\hat{S}_k > x) \sim N_k P\left(X_1 > \frac{x}{m_k}\right) \sim \frac{N_k}{c_{k+1}} P\left(X_{k+1} > \frac{x}{m_k}\right).$$

To prove the result for $d = k + 1$, note that

$$m_{k+1} = m_k \vee a_{k+1} \tag{4.5}$$

and

$$N_{k+1} = c_{k+1} \mathbf{1}_{\{a_{k+1}=m_{k+1}\}} + N_k \mathbf{1}_{\{m_k=m_{k+1}\}},$$

so that

$$\frac{N_{k+1}}{c_{k+1}} = \mathbf{1}_{\{a_{k+1}=m_{k+1}\}} + \frac{N_k}{c_{k+1}} \mathbf{1}_{\{m_k=m_{k+1}\}}. \tag{4.6}$$

By the induction hypothesis,

$$\lim_{x \rightarrow \infty} \frac{P(m_k^{-1} \hat{S}_k > x)}{P(X_{k+1} > x)} = \lim_{x \rightarrow \infty} \frac{P(m_k^{-1} \hat{S}_k > x)}{P(X_1 > x)} \frac{P(X_1 > x)}{P(X_{k+1} > x)} = \frac{N_k}{c_{k+1}}. \tag{4.7}$$

If we prove that the assumptions in Theorem 4.1 are valid with $U = X_{k+1}$ and $V = m_k^{-1} \hat{S}_k$, then, Theorem 4.1, (4.5), (4.6), and (4.7) imply that, as $x \rightarrow \infty$,

$$\begin{aligned} P(\hat{S}_{k+1} > x) &= P(a_{k+1} X_{k+1} + m_k \hat{S}_k > x) \\ &\sim \frac{N_{k+1}}{c_{k+1}} P\left(X_{k+1} > \frac{x}{m_{k+1}}\right) \\ &\sim N_{k+1} P\left(X_1 > \frac{x}{m_{k+1}}\right), \end{aligned}$$

and by induction, our result holds for all $d \geq 2$.

Assumption 2.1 is assumed. For (4.1), consider that, on the one hand,

$$N_k = \sum_{\{1 \leq i \leq k: a_i = m_k\}} c_i \geq \bigwedge_{i=1}^k c_i > 0,$$

and on the other hand,

$$N_k = \sum_{\{1 \leq i \leq k: a_i = m_k\}} c_i \leq k \bigvee_{i=1}^k c_i < \infty,$$

and, therefore, the limit in (4.7) satisfies $N_k/c_{k+1} \in (0, \infty)$.

Next, suppose that two random variables U and V are tail equivalent and that both belong to $\text{MDA}(\Delta)$. If $f(\cdot)$ and $\tilde{f}(\cdot)$ are the auxiliary functions of U and V , respectively, then $f(x) \sim \tilde{f}(x)$ as $x \rightarrow \infty$; see Resnick (1971a), (1971b). Since, in the present case, all the random variables are tail equivalent, comment 3 in Section 2.1.3 implies that we can work with the auxiliary function of any one of them, say X_{k+1} . So, X_1, X_2, \dots, X_d pairwise satisfy Assumptions 2.3–2.5 with auxiliary function $f(\cdot)$ of X_{k+1} . That is, for $1 \leq i \neq j \leq d$ and any $t > 0$,

$$\lim_{x \rightarrow \infty} P(X_j > tf(x) \mid X_i > x) = 0 \tag{4.8}$$

and, for $1 \leq i < j \leq d$ and some $L_{ij} > 0$,

$$\frac{P(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{P(X_i > x)} = 0. \tag{4.9}$$

To verify Assumption 2.3, observe that, for $t > 0$, as $x \rightarrow \infty$,

$$\begin{aligned} &P(|m_k^{-1} \hat{S}_k| > tf(x) \mid X_{k+1} > x) \\ &\leq P(a_1 X_1 + a_2 X_2 + \dots + a_k X_k > m_k tf(x) \mid X_{k+1} > x) \\ &\leq \sum_{i=1}^k P\left(X_i > \frac{a_i^{-1} m_k tf(x)}{k} \mid X_{k+1} > x\right) \\ &\leq \sum_{i=1}^k P\left(X_i > \frac{tf(x)}{k} \mid X_{k+1} > x\right) \\ &\rightarrow 0, \end{aligned}$$

by (4.8). For Assumption 2.4, note that

$$\lim_{x \rightarrow \infty} \frac{P(m_k^{-1} \hat{S}_k > x)}{P(X_1 > x)} = N_k, \tag{4.10}$$

and, for $1 \leq i \leq k$,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{P((m_k^{-1} \hat{S}_k > x) \cap (m_k^{-1} a_i X_i > x))}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{P(m_k^{-1} a_i X_i > x)}{P(X_1 > x)} \\ &= c_i \mathbf{1}_{\{a_i = m_k\}}. \end{aligned} \tag{4.11}$$

The first equality uses the assumption that the X_i s are nonnegative. The second equality is true from (4.4) and the fact that the tail of any distribution in the maximal domain of attraction of the Gumbel distribution is $-\infty$ -varying. Now, for $1 \leq i < j \leq k$, using (2.4),

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{\mathbb{P}((m_k^{-1} \hat{S}_k > x) \cap (m_k^{-1} a_i X_i > x) \cap (m_k^{-1} a_j X_j > x))}{\mathbb{P}(X_1 > x)} \\ & \leq \lim_{x \rightarrow \infty} \frac{\mathbb{P}((m_k^{-1} a_i X_i > x) \cap (m_k^{-1} a_j X_j > x))}{\mathbb{P}(X_1 > x)} \\ & \leq \lim_{x \rightarrow \infty} \frac{\mathbb{P}((X_i > x) \cap (X_j > x))}{\mathbb{P}(X_1 > x)} \\ & = 0. \end{aligned}$$

Therefore, using (4.11),

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{\mathbb{P}((m_k^{-1} \hat{S}_k > x) \cap (\bigcup_{i=1}^k (m_k^{-1} a_i X_i > x)))}{\mathbb{P}(X_1 > x)} \\ & = \lim_{x \rightarrow \infty} \frac{\sum_{i=1}^k \mathbb{P}((m_k^{-1} \hat{S}_k > x) \cap (m_k^{-1} a_i X_i > x))}{\mathbb{P}(X_1 > x)} \\ & = N_k. \end{aligned} \tag{4.12}$$

From (4.10) and (4.12), it follows that

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}((m_k^{-1} \hat{S}_k > x) \cap (\bigcup_{i=1}^k (m_k^{-1} a_i X_i > x))^c)}{\mathbb{P}(X_1 > x)} = 0,$$

and this, along with (4.4) and (4.7), gives

$$\lim_{x \rightarrow \infty} \frac{\mathbb{P}((m_k^{-1} \hat{S}_k > x) \cap (\bigcup_{i=1}^k (m_k^{-1} a_i X_i > x))^c)}{\mathbb{P}(m_k^{-1} \hat{S}_k > x)} = 0. \tag{4.13}$$

Now, we check Assumption 2.4. For $t > 0$, as $x \rightarrow \infty$,

$$\begin{aligned} & \mathbb{P}(|X_{k+1}| > tf(x) \mid m_k^{-1} \hat{S}_k > x) \\ & = \frac{\mathbb{P}(X_{k+1} > tf(x), m_k^{-1} \hat{S}_k > x)}{\mathbb{P}(m_k^{-1} \hat{S}_k > x)} \\ & \sim \frac{\mathbb{P}(X_{k+1} > tf(x), m_k^{-1} \hat{S}_k > x, \bigcup_{i=1}^k \{m_k^{-1} a_i X_i > x\})}{\mathbb{P}(m_k^{-1} \hat{S}_k > x)} \\ & \leq \frac{\mathbb{P}(X_{k+1} > tf(x), \bigcup_{i=1}^k \{m_k^{-1} a_i X_i > x\})}{\mathbb{P}(\hat{S}_k > m_k x)} \\ & \leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > tf(x), m_k^{-1} a_i X_i > x)}{\mathbb{P}(\hat{S}_k > m_k x)}, \end{aligned}$$

where we have used (4.13). Using our induction hypothesis, we find that the quantity above is

asymptotically equivalent to

$$\begin{aligned} & \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > tf(x), m_k^{-1} a_i X_i > x)}{N_k \mathbb{P}(X_1 > x)} \\ & \leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > tf(x), X_i > x)}{N_k \mathbb{P}(X_1 > x)} \\ & = \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > tf(x), X_i > x)}{\mathbb{P}(X_i > x)} \frac{\mathbb{P}(X_i > x)}{N_k \mathbb{P}(X_1 > x)} \\ & \rightarrow 0, \end{aligned}$$

by (4.8).

For Assumption 2.5, let $L = kL_{\max}$, where $L_{\max} = \max_{1 \leq i \leq k} L_{i,k+1}$ (recall (4.9)). Then, using (4.7), (4.9), and (4.13), we have

$$\begin{aligned} & \frac{\mathbb{P}(X_{k+1} > kL_{\max} f(x), m_k^{-1} \hat{S}_k > kL_{\max} f(x))}{\mathbb{P}(X_{k+1} > x)} \\ & \sim \frac{\mathbb{P}(X_{k+1} > kL_{\max} f(x), m_k^{-1} \hat{S}_k > x, \bigcup_{i=1}^k \{m_k^{-1} a_i X_i > x\})}{\mathbb{P}(X_{k+1} > x)} \\ & \leq \frac{\mathbb{P}(X_{k+1} > kL_{\max} f(x), \bigcup_{i=1}^k \{m_k^{-1} a_i X_i > L_{\max} f(x)\})}{\mathbb{P}(X_{k+1} > x)} \\ & \leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > L_{i,k+1} f(x), m_k^{-1} a_i X_i > L_{i,k+1} f(x))}{\mathbb{P}(X_{k+1} > x)} \\ & \leq \frac{\sum_{i=1}^k \mathbb{P}(X_{k+1} > L_{i,k+1} f(x), X_i > L_{i,k+1} f(x))}{\mathbb{P}(X_{k+1} > x)} \\ & \rightarrow 0. \end{aligned}$$

4.2. One special case where the distributions are possibly not tail equivalent

Theorem 4.2. Assume that Y_1, Y_2, \dots, Y_d are identically distributed nonnegative random variables. Also, assume that $a_i, \beta_i \geq 0, i = 1, 2, \dots, d$. For $d \geq 1$, define $\hat{S}_d = a_1 Y_1^{\beta_1} + a_2 Y_2^{\beta_2} + \dots + a_d Y_d^{\beta_d}$ and set

$$\beta = \bigvee_{i=1}^d \beta_i, \quad q_d = \bigvee_{\{1 \leq i \leq d: \beta_i = \beta\}} a_i, \quad J_d = |\{1 \leq i \leq d: \beta_i = \beta, a_i = q_d\}|,$$

where $|\cdot|$ denotes the size of the set. Suppose that $q_d Y_1^\beta, q_d Y_2^\beta, \dots, q_d Y_d^\beta$ pairwise satisfy Assumptions 2.3–2.5 and that the distribution of $q_d Y_1^\beta$ satisfies Assumption 2.1, where the auxiliary function $f(x)$ satisfies the additional condition that $f(x) \rightarrow \infty$ as $x \rightarrow \infty$. Then,

$$\mathbb{P}(\hat{S}_d > x) \sim J_d \mathbb{P}\left(Y_1^\beta > \frac{x}{q_d}\right).$$

Remark 4.2. If $\beta_1 > \beta_2$ then $Y_1^{\beta_1}$ and $Y_2^{\beta_2}$ are not tail equivalent. Note that in this case the asymptotic approximation of $\mathbb{P}(a_1 Y_1^{\beta_1} + a_2 Y_2^{\beta_2} > x)$ does not depend on a_2 .

Theorem 4.2 shows a different tail behavior from the tail-equivalent cases, but follows the paradigm that only the heaviest tails matter. The theorem shows that Theorem 1 of

Asmussen and Rojas-Nandayapa (2008) is a special case of a more general phenomenon. Let $(X_1, X_2, \dots, X_d) \sim N(\mathbf{0}, \Sigma)$, where $\mathbf{0} = (0, 0, \dots, 0)$ is the d -dimensional vector of zeros and

$$\Sigma = (\rho_{ij}), \quad \rho_{ii} = 1 \text{ for all } i \text{ and } \rho_{ij} < 1 \text{ for } 1 \leq i < j \leq d.$$

Let $(Y_1, Y_2, \dots, Y_d) \sim (\exp(X_1), \exp(X_2), \dots, \exp(X_d))$. Clearly,

$$a_i Y_i^{\beta_i} \sim \text{lognormal}(\log a_i, \beta_i^2).$$

From Example 3.5, $(q_d Y_1^\beta, q_d Y_2^\beta, \dots, q_d Y_d^\beta)$ satisfies the assumptions of Theorem 4.2, where q_d and β have the same meaning as in Theorem 4.2. Also, $(Z_1, Z_2, \dots, Z_d) = (a_1 Y_1^{\beta_1}, a_2 Y_2^{\beta_2}, \dots, a_n Y_d^{\beta_d})$ satisfies the assumptions of Theorem 1 of Asmussen and Rojas-Nandayapa (2008). The results of that theorem and Theorem 4.2 match.

Proof of Theorem 4.2. Without loss of generality, assume that $\beta_1 = \beta$ and $a_1 = q_d$. Also, assume that $a_i > 0$ for $i = 1, 2, \dots, d$. Define

$$X_i = a_i Y_i^{\beta_i}, \quad i = 1, 2, \dots, d.$$

To start, suppose that, for some $i \in \{2, \dots, d\}$, $\beta_i < \beta$. Then, for large x , $[a_i Y_i^{\beta_i} > x] \subseteq [(q_d/2) Y_1^\beta > x]$, and, hence, for large x ,

$$P(a_i Y_i^{\beta_i} > x) \leq P\left(\frac{q_d}{2} Y_1^\beta > x\right) = P(q_d Y_1^\beta > 2x).$$

Then,

$$c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(a_i Y_i^{\beta_i} > x)}{P(q_d Y_1^\beta > x)} \leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_1^\beta > 2x)}{P(q_d Y_1^\beta > x)} = 0. \tag{4.14}$$

Next, suppose that, for some $i \in \{2, \dots, d\}$, $\beta_i = \beta$ and $a_i < q_d$. Then,

$$c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(a_i Y_i^\beta > x)}{P(q_d Y_1^\beta > x)} = \lim_{x \rightarrow \infty} \frac{P(q_d Y_1^\beta > q_d x/a_i)}{P(q_d Y_1^\beta > x)} = 0. \tag{4.15}$$

In both (4.14) and (4.15), the last equalities are true from the fact that the tail of any distribution in the maximal domain of attraction of the Gumbel distribution is $-\infty$ -varying.

Finally, suppose that, for some $i \in \{2, \dots, d\}$, $\beta_i = \beta$ and $a_i = q_d$. Then,

$$c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(Y_i^\beta > x/q_d)}{P(Y_1^\beta > x/q_d)} = \lim_{x \rightarrow \infty} \frac{P(Y_1^\beta > x/q_d)}{P(Y_1^\beta > x/q_d)} = 1. \tag{4.16}$$

It suffices to check the assumptions in Corollary 2.2 with this set of X_1, X_2, \dots, X_d , since then Corollary 2.2 and (4.14), (4.15), and (4.16) would imply that, as $x \rightarrow \infty$,

$$P(\hat{S}_d > x) \sim \left(1 + \sum_{i=2}^d c_i\right) P(X_1 > x) \sim J_d P(X_1 > x) = J_d P\left(Y_1^\beta > \frac{x}{q_d}\right).$$

Assumption 2.1 is assumed in the statement of the theorem and (2.18) is already shown in (4.14), (4.15), and (4.16). For Assumptions 2.3 and 2.4, proceed as follows. By hypothesis we

know that X_1 belongs to the maximal domain of attraction of the Gumbel distribution. Let $f(\cdot)$ be the auxiliary function corresponding to the distribution of X_1 . By hypothesis we know that, for $t > 0$ and $1 \leq i \neq j \leq d$,

$$\lim_{x \rightarrow \infty} P(q_d Y_j^\beta > t f(x) \mid q_d Y_i^\beta > x) = 0.$$

Using comment 3 in Section 2.1.3, it is enough to show that

$$\lim_{x \rightarrow \infty} \frac{P(X_j > t f(x), X_i > x)}{P(X_1 > x)} = 0,$$

and to see this, note that since $f(x) \rightarrow \infty$, for large x and all $t > 0$, $[X_j > t f(x), X_i > x] \subseteq [q_d Y_j^\beta > t f(x), q_d Y_i^\beta > x]$. Hence,

$$\lim_{x \rightarrow \infty} \frac{P(X_j > t f(x), X_i > x)}{P(X_1 > x)} \leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > t f(x), q_d Y_i^\beta > x)}{P(q_d Y_i^\beta > x)} = 0.$$

For Assumption 2.5, using comment 3 in Section 2.1.3, we show that, for some $L > 0$,

$$\lim_{x \rightarrow \infty} \frac{P(X_j > L f(x), X_i > L f(x))}{P(X_1 > x)} = 0.$$

By hypothesis we know that, for all $1 \leq i < j \leq d$, there exists some $L_{ij} > 0$ such that

$$\lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > L_{ij} f(x), q_d Y_i^\beta > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} = 0. \tag{4.17}$$

Also, note that since $f(x) \rightarrow \infty$, for large x , $[X_j > L_{ij} f(x), X_i > L_{ij} f(x)] \subseteq [q_d Y_j^\beta > L_{ij} f(x), q_d Y_i^\beta > L_{ij} f(x)]$. Hence,

$$\lim_{x \rightarrow \infty} \frac{P(X_j > L_{ij} f(x), X_i > L_{ij} f(x))}{P(X_1 > x)} \leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > L_{ij} f(x), q_d Y_i^\beta > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} = 0$$

by (4.17).

5. An optimization problem

5.1. The problem

Suppose that we have a portfolio consisting of d financial instruments. The risk per unit of the i th instrument is X_i . The goal is to earn a revenue of L dollars. Assume that each unit of the i th instrument earns l_i dollars over the chosen time horizon. Subject to earnings being at least L dollars, how many units of each instrument, a_1, a_2, \dots, a_d , should be used to build the portfolio, so that the probability that the total portfolio risk $a_1 X_1 + a_2 X_2 + \dots + a_d X_d$ exceeds some fixed large threshold x is minimal? Thus, consider the following optimization problem:

$$\min_{\{a_1, \dots, a_d\}} P\left(\sum_{i=1}^d a_i X_i > x\right)$$

such that $a_1 l_1 + a_2 l_2 + \dots + a_d l_d \geq L$ and $a_i \geq 0, i = 1, 2, \dots, d$.

For a more general case, consider the following optimization problem:

$$\min_{\{a_1, \dots, a_d\}} P\left(\sum_{i=1}^d a_i X_i > x\right)$$

such that $h(a_1, a_2, \dots, a_d) \geq L$ and $a_i \geq 0, i = 1, 2, \dots, d$.

5.2. The method

Suppose that X_1, X_2, \dots, X_d satisfy the assumptions of Corollary 4.1. Even with these assumptions, obtaining an exact solution of the optimization problem is difficult. An obvious way to obtain an approximate solution to the optimization problem is to assume that the threshold x is big and use the asymptotic approximation of $P(a_1 X_1 + a_2 X_2 + \dots + a_d X_d > x)$ from Corollary 4.1, hoping that the solution of the resulting optimization problem is close to the actual optimal value. So, using the notation of Corollary 4.1, we solve the following optimization problem:

$$\min_{\{a_1, \dots, a_d\}} N_d P\left(X_1 > \frac{x}{m_d}\right)$$

such that $h(a_1, a_2, \dots, a_d) \geq L$ and $a_i \geq 0, i = 1, 2, \dots, d$.

Suppose that $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_d$ and $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_d$ are two feasible solutions for the given set of constraints. Set

$$\begin{aligned} \hat{m}_d &= \bigvee_{i=1}^d \hat{a}_i, & \hat{N}_d &= \sum_{\{1 \leq i \leq d: \hat{a}_i = \hat{m}_d\}} c_i, \\ \tilde{m}_d &= \bigvee_{i=1}^d \tilde{a}_i, & \tilde{N}_d &= \sum_{\{1 \leq i \leq d: \tilde{a}_i = \tilde{m}_d\}} c_i. \end{aligned}$$

If $\hat{m}_d > \tilde{m}_d$ then, since $P(X_1 \leq x) \in \text{MDA}(\Lambda)$, as $x \rightarrow \infty$,

$$\frac{P(X_1 > x/\hat{m}_d)}{P(X_1 > x/\tilde{m}_d)} \rightarrow \infty.$$

Now, since both $\hat{N}_d, \tilde{N}_d \in [\bigwedge_{i=1}^d c_i, d \bigvee_{i=1}^d c_i]$, we have, as $x \rightarrow \infty$,

$$\frac{\hat{N}_d P(X_1 > x/\hat{m}_d)}{\tilde{N}_d P(X_1 > x/\tilde{m}_d)} \rightarrow \infty.$$

So, we hope that $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_d$ is a better feasible solution for the optimization problem.

Values of a_1, a_2, \dots, a_d which solve the above optimization problem can be computed by solving the following two optimization problems in sequence.

(i) First solve

$$\min_{\{a_1, \dots, a_d\}} m_d = \max\{a_1, a_2, \dots, a_d\}$$

such that $h(a_1, a_2, \dots, a_d) \geq L$ and $a_i \geq 0, i = 1, 2, \dots, d$.

(ii) Suppose that the best choice of a_1, a_2, \dots, a_d gives m as the value of the objective function for the optimization problem in (i). Then we solve

$$\min_{\{a_1, \dots, a_d\}} N_d = \sum_{\{1 \leq i \leq d: a_i = m\}} c_i$$

such that $h(a_1, a_2, \dots, a_d) \geq L, \max\{a_1, a_2, \dots, a_d\} = m, a_i \geq 0, i = 1, 2, \dots, d$.

5.3. A special case

The motivating case is that h is a linear function with positive coefficients of the form

$$h(a_1, a_2, \dots, a_d) = a_1l_1 + a_2l_2 + \dots + a_d l_d.$$

The approximate solution using the asymptotic form of $P(\sum_{i=1}^d a_i X_i > x)$ is

$$a_1 = a_2 = \dots = a_d = \frac{L}{l_1 + l_2 + \dots + l_d}.$$

This leads to $m = L/(l_1 + l_2 + \dots + l_d)$ and $N_d = \sum_{i=1}^d c_i$.

6. Simulation studies

We carried out some simulation studies to check the accuracy of the asymptotic approximation in Theorem 2.2 for fixed large thresholds and also to check how good the approximate solution is for the optimization problem. As expected, in some cases the approximation works well whereas in others it performs poorly, which suggests caution when using the asymptotic results for numerical purposes. Simulation also suggests that the approximate solution of the optimization problem works well in cases where the simulation studies suggest that the approximation is good for fixed large thresholds. One particular model studied, Example 3.5 with $\mu = 0$ and $\sigma = 1$, is noted here to illustrate the point. We varied ρ and observed the asymptotic behavior of the sum of the risks.

6.1. Where is the approximation good?

To test the approximation for $P(X + Y > x)$, we need to find good simulation estimates of the probabilities $P(X + Y > x)$. This, however, is not easy, especially in the case when the marginal distributions of the risks X and Y are subexponential, and is still a topic of current research in the simulation community. The approach usually taken in these cases is conditional Monte Carlo (see Asmussen and Glynn (2007, p. 173)). So, this method is used to compute $P(X + Y > x)$, and the simulation estimates are compared with the theoretical approximations.

The simulation of $P(X + Y > x)$ uses the algorithm suggested in Asmussen and Rojas-Nandayapa (2008) for $\rho \in (-1, 1)$, who also noted the properties of this algorithm. If $\rho = -1$, we have a way to compute the probability exactly. In this case, $X = 1/Y$ almost surely, so in the following manner we compute the required probability:

$$\begin{aligned} P\left(X + \frac{1}{X} > x\right) &= P\left(X > \frac{x + \sqrt{x^2 - 2}}{2}\right) + P\left(X < \frac{x - \sqrt{x^2 - 2}}{2}\right) \\ &= P\left(\log X > \log\left(\frac{x + \sqrt{x^2 - 2}}{2}\right)\right) \\ &\quad + P\left(\log X < \log\left(\frac{x - \sqrt{x^2 - 2}}{2}\right)\right) \\ &= \bar{\Phi}\left(\log\left(\frac{x + \sqrt{x^2 - 2}}{2}\right)\right) + \Phi\left(\log\left(\frac{x - \sqrt{x^2 - 2}}{2}\right)\right). \end{aligned}$$

6.1.1. *Patterns in the results.* For judging the quality of the asymptotic approximation, we focus on the simulation estimate $P(X + Y > x)$ and not the threshold x , since a change of distribution may imply a change in how rare a particular threshold crossing is. So, when comparing the

quality of the asymptotic approximation across different models, it makes more sense to focus on the value of $P(X + Y > x)$, rather than the particular threshold x . When $\rho = -1$, exact calculations suggest that the approximation is extremely good, even when the actual probability $P(X + Y > x)$ is of the order of 10^{-2} . As expected, the asymptotic approximation improves as a function of increasing threshold. When $\rho \in (-1, 1)$, we rely on the simulation estimate as a surrogate for the exact tail probability and compare it with the theoretical approximations.

The results indicate that the closer ρ is to -1 , the better the approximation. For $\rho = -1$, the approximation is good for events with probability of the order of 10^{-2} , and to achieve comparable precision in the relative error when $\rho = 0$, the event has to be much rarer and have a probability of the order of 10^{-10} . For $\rho = 0.9$, the results for different thresholds did not show any convergence pattern. This emphasizes that in practice the numerical approximations should be used with caution. Clearly, for $\rho = 1$, the asymptotic approximation is not correct and $\rho = 0.9$ is expected to behave somewhat like the case in which $\rho = 1$.

Tables 1–4 give representative results. We first give the results for $\rho = -1$ in Table 1, since in this case no simulation is required. The ratio column in Table 1 is defined as

$$\text{ratio} = \frac{\text{actual probability}}{\text{asymptotic approximation}}.$$

For the subsequent tables, the ratio and half-width columns are defined as

$$\text{ratio} = \frac{\text{simulation estimated probability}}{\text{asymptotic approximation}},$$

$$\text{half-width} = \text{half-width of the 95\% confidence interval of the ratio.}$$

In each case, 10^7 observations were used to compute the probability estimates.

TABLE 1: $\rho = -1$.

Threshold	Actual probability	Asymptotic approximation	Ratio
10	0.0219	0.0213	1.0272
16	0.0056	0.0056	1.0121
24	0.0015	0.0015	1.0060
30	6.7365×10^{-4}	6.7091×10^{-4}	1.0041
100	4.1233×10^{-6}	4.1213×10^{-6}	1.0005
1000	4.9238×10^{-12}	4.9238×10^{-12}	1.0000

TABLE 2: $\rho = -0.9$.

Threshold	Simulation estimated probability	Asymptotic approximation	Ratio	Half-width
3	0.3687	0.2719	1.3556	0.0006
5	0.1207	0.1075	1.1227	0.0012
10	0.0221	0.0213	1.0375	0.0026
20	0.0028	0.0027	1.0082	0.0064
30	6.8873×10^{-4}	6.7091×10^{-4}	1.0265	0.0119
40	2.2134×10^{-4}	2.2524×10^{-4}	0.9827	0.0183
50	9.3675×10^{-5}	9.1526×10^{-5}	1.0235	0.0285

TABLE 3: $\rho = 0$.

Threshold	Simulation estimated probability	Asymptotic approximation	Ratio	Half-width
10	0.0338	0.0213	1.5844	0.0033
50	1.0798×10^{-4}	9.1526×10^{-5}	1.1798	0.0002
100	4.5032×10^{-6}	4.1213×10^{-6}	1.0927	0.0001
300	1.2117×10^{-8}	1.1718×10^{-8}	1.0341	0.0000
600	1.6147×10^{-10}	1.5853×10^{-10}	1.0185	0.0122
1000	4.9821×10^{-12}	4.9238×10^{-12}	1.0118	0.0000
2000	1.9620×10^{-14}	2.9310×10^{-14}	1.0106	0.0000

TABLE 4: $\rho = 0.9$.

Threshold	Simulation estimated probability	Asymptotic approximation	Ratio	Half-width
10	0.0521	0.0213	2.4439	0.0088
30	0.0030	6.7091×10^{-4}	4.4081	0.0275
50	5.2652×10^{-4}	9.1526×10^{-5}	5.7527	0.0759
75	1.1217×10^{-4}	1.5781×10^{-5}	7.1077	0.1843
100	3.4333×10^{-5}	4.1213×10^{-6}	8.3307	0.3642

6.2. How good is the portfolio suggestion?

Here, we consider the quality of our approximate solutions for the optimization problem. We choose the same risk model given in Example 3.5, because we have information about which values of ρ lead to good asymptotic approximations. We resort to a naive method for analyzing the optimization. For different (a_1, a_2) , we obtain estimates of $P(a_1X + a_2Y > x)$ through simulation. To obtain the estimates, proceed as follows. For $a_1, a_2 > 0$,

$$\begin{pmatrix} a_1 X \\ a_2 Y \end{pmatrix} = \begin{pmatrix} \exp(\log(a_1) + X_1) \\ \exp(\log(a_2) + X_2) \end{pmatrix}.$$

Now,

$$\begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} = \begin{pmatrix} \log(a_1) + X_1 \\ \log(a_2) + X_2 \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \log(a_1) \\ \log(a_2) \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right), \quad \rho \in [-1, 1].$$

So, again, we are in the framework of Asmussen and Rojas-Nandayapa (2008), and we use the algorithm given in their paper to estimate the rare event probabilities. When either a_1 or a_2 is equal to 0, we can compute the exact probability and, hence, do not need an estimate. We choose (a_1, a_2) in the following way. Let C be the set of all possible (a_1, a_2) which satisfy the constraint. First, a_1 is chosen from the corresponding projection of C with a small grid, and then, for each a_1 , a_2 is determined from the constraint. Let us call this set C^* . For $(a_1, a_2) \in C^*$, $P(a_1X + a_2Y > x)$ is estimated through simulation and then it is observed which (a_1, a_2) gives the minimum estimate of $P(a_1X + a_2Y > x)$. Let $(\tilde{a}_1, \tilde{a}_2)$ be this pair, i.e.

$$P(\tilde{a}_1 X + \tilde{a}_2 Y > x) = \min_{(a_1, a_2) \in C^*} P(a_1 X + a_2 Y > x).$$

Also, let (a_1^*, a_2^*) be the approximate solution of the optimization problem, as noted in the previous section. Relative error of the approximate solution is computed by comparing $P(a_1^*X + a_2^*Y > x)$ with $\min_{(a_1, a_2) \in C^*} P(a_1X + a_2Y > x)$.

6.2.1. *Identifying patterns.* We do not have error estimates for our simulation results. One could consider bootstrapping to obtain such error estimates, but we have not done so. Despite the weaknesses of the naive procedure, the results are interesting.

We note one case with the linear constraint $2a_1 + 3a_2 = 1$. The suggested optimum portfolio based on asymptotic approximation is $(a_1^*, a_2^*) = (0.2, 0.2)$. The cases where $\rho = -0.9, 0, 0.9$ are chosen, the reason being that we know from the results in the earlier section that the asymptotic approximation is good in the case $\rho = -0.9$, reasonable when $\rho = 0$, and rather bad when $\rho = 0.9$. The approximate solution (a_1^*, a_2^*) relies on replacing the original objective function by its asymptotic approximation, and so it is reasonable to expect different accuracies for these three values of ρ , and this turned out to be the case. In the cases $\rho = -0.9$ and $\rho = 0$, we see that \tilde{a}_1 comes close to 0.2 as the threshold x increases. But, in the case of $\rho = 0.9$, no pattern in the convergence of \tilde{a}_1 is observed, which is expected because, for $\rho = 1$, both the risks are actually the same random variable, implying indifference to the choice of $(a_1, a_2) \in C$.

Another remark is that for $\rho = -0.9, 0, 0.9$, the relative errors do not show any convergence pattern. Perhaps to expect otherwise is unrealistic as we are using the minimum of some simulation estimates to compute the relative error. Still, we illustrate through an example the accuracy by comparing with an extreme case where we build the portfolio consisting of only one asset. For $\rho = 0$ and threshold $x = 10$, the extreme cases will yield probabilities 0.2441 and 0.1360. These risk probabilities are quite high compared to that of our suggested optimal portfolio (a_1^*, a_2^*) based on asymptotic approximation, which has risk probability $P(a_1^*X + a_2^*Y > x) = 1.0793 \times 10^{-4}$; also, the minimum of the simulation estimates $P(\tilde{a}_1X + \tilde{a}_2Y > x)$ is of the same order. So, the suggested portfolio (a_1^*, a_2^*) is quite effective in reducing the risk and possibly close to the best one.

The following additional conclusion can be made. In the case $\rho = -0.9$, even when $P(\tilde{a}_1X + \tilde{a}_2Y > x)$ is as big as 0.11, it is quite close to $P(a_1^*X + a_2^*Y > x)$, indicating that the suggested optimal choice (a_1^*, a_2^*) significantly reduces the risk probability. For $\rho = 0$, a comparable statement can be made when the minimum of the probability estimates is of the order of 10^{-2} . However, for $\rho = 0.9$, the relative errors are never close to 0. Interestingly, even for $\rho = 0.9$, $P(\tilde{a}_1X + \tilde{a}_2Y > x)$ and $P(a_1^*X + a_2^*Y > x)$ are almost always of the same order. However, it should be noted at this point that even in the case $\rho = 0.9$, the extreme cases where the portfolio is built on entirely one of the assets, $P(a_1X + a_2Y > x)$ is of a much bigger order than $P(\tilde{a}_1X + \tilde{a}_2Y > x)$. So, in this case, $P(a_1X + a_2Y > x)$ possibly differs considerably from choices where $a_1, a_2 > 0$ and the case where either $a_1 = 0$ or $a_2 = 0$, but does not differ too much among the choices where $(a_1, a_2) \in C, a_1, a_2 > 0$. This fact justifies the intuition mentioned before that the case $\rho = 0.9$ is similar to the case $\rho = 1$.

The results are summarized in Tables 5, 6, and 7 for $\rho = -0.9, 0, 0.9$ and constraint $2a_1 + 3a_2 = 1$. For each fixed ρ , we give

- the threshold x ,
- \tilde{a}_1 , where $(\tilde{a}_1, \tilde{a}_2) \in C^*$ and

$$P(\tilde{a}_1X + \tilde{a}_2Y > x) = \min_{(a_1, a_2) \in C^*} P(a_1X + a_2Y > x),$$

- $E1 = \min_{(a_1, a_2) \in C^*} P(a_1X + a_2Y > x)$,

TABLE 5: $\rho = -0.9$.

Threshold	\tilde{a}_1	$E1$	$E2$	Relative error
1	0.13	0.1097	0.1204	0.0975
3	0.18	0.0067	0.0069	0.0322
5	0.19	0.0013	0.0013	0.0294
10	0.19	1.0299×10^{-4}	1.0592×10^{-4}	0.0284
20	0.21	2.0806×10^{-6}	2.0806×10^{-6}	1.2213×10^{-15}

TABLE 6: $\rho = 0$.

Threshold	\tilde{a}_1	$E1$	$E2$	Relative error
1	0.03	0.1349	0.1723	0.2765
3	0.16	0.0093	0.0101	0.0759
5	0.18	0.0016	0.0017	0.0503
10	0.19	1.0424×10^{-4}	1.0793×10^{-4}	0.0354
20	0.20	4.3888×10^{-6}	4.3888×10^{-6}	0

TABLE 7: $\rho = 0.9$.

Threshold	\tilde{a}_1	$E1$	$E2$	Relative error
1	0.01	0.1360	0.1798	0.3223
3	0.01	0.0140	0.0208	0.4831
5	0.02	0.0033	0.0050	0.5146
10	0.02	2.8357×10^{-4}	4.9475×10^{-4}	0.7447
20	0.04	1.3241×10^{-6}	2.4023×10^{-6}	0.8142

- $E2 = P(a_1^*X + a_2^*Y > x)$,
- the relative error is equal to $(E2 - E1)/E1$.

For each value of ρ , a_1 was chosen with gap 0.01 from the projection of C^* , i.e. we considered $a_1 = 0, 0.01, 0.02, \dots, 0.5$. For each such a_1 , we used 10 000 observations to obtain the estimates of the probability $P(a_1X + a_2Y > x)$.

7. Concluding remarks

An important case for the study of the asymptotic behavior of the sum of risks is the case where the risks are asymptotically independent, identically distributed, and belong to the maximal domain of attraction of the Gumbel distribution. Many commonly occurring risk distributions fall into this category. We have provided sufficient conditions for

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = 2,$$

and extended the conditions to cover the case where the marginal distributions are not the same, and to the case where some risk distributions have lighter tails but the distribution does not belong to the maximal domain of attraction of the Gumbel distribution. We were not able to provide necessary and sufficient conditions for this kind of asymptotic behavior, which is an unresolved problem. It will be interesting to see if it is possible to find a distribution of risks

(X, Y) for which the risks are asymptotically independent, but not independent, identically distributed, and belong to $\text{MDA}(\Delta)$, and the asymptotic behavior of the sum is different than the two cases mentioned in the introduction, viz.

$$\lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} \in \{2, \infty\}.$$

Even for cases where the asymptotic behavior is understood, nothing is known about the rate of convergence in these cases, i.e. a quantitative estimate of how good the approximation $2P(X > x)$ is for the quantity $P(X + Y > x)$ for a large threshold x . Simulation studies indicated that in certain circumstances the approximation is accurate, but in other cases its accuracy is dismal.

We observed in the previous section that when the tail probability approximation is good, the approximate solution of the optimization problem is also accurate, whereas in the other cases this solution has poor accuracy. So, results on the rate of convergence would contribute to understanding the appropriateness of the approximate solutions in different scenarios.

Acknowledgement

An anonymous and conscientious referee provided many insightful and helpful comments.

References

- ALBRECHER, H., ASMUSSEN, S. AND KORTSCHAK, D. (2006). Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes* **9**, 107–130.
- ALINK, S., M. LÖWE, AND WÜTHRICH, M. V. (2004). Diversification of aggregate dependent risks. *Insurance Math. Econom.* **35**, 77–95.
- ASMUSSEN, S. AND GLYNN, P. W. (2007). *Stochastic Simulation: Algorithms and Analysis*. Springer, New York.
- ASMUSSEN, S. AND ROJAS-NANDAYAPA, L. (2008). Asymptotics of sums of lognormal random variables with Gaussian copula. *Statist. Prob. Lett.* **78**, 2709–2714.
- DAVIS, R. AND RESNICK, S. (1988). Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stoch. Process. Appl.* **30**, 41–68.
- DE HAAN, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum, Amsterdam.
- DE HAAN, L. AND FERREIRA, A. (2006). *Extreme Value Theory*. Springer, New York.
- EMBRECHTS, P. AND PUCETTI, G. (2006). Bounds for functions of dependent risks. *Finance Stoch.* **10**, 341–352.
- EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997). *Modelling Extremal Events*. Springer, Berlin.
- GOLDIE, C. M. AND RESNICK, S. (1988). Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Adv. Appl. Prob.* **20**, 706–718.
- KLÜPPELBERG, C. AND RESNICK, S. I. (2008). The Pareto copula, aggregation of risks and the emperor's socks. *J. Appl. Prob.* **45**, 67–84.
- KO, B. AND TANG, Q. (2008). Sums of dependent nonnegative random variables with subexponential tails. *J. Appl. Prob.* **45**, 85–94.
- RESNICK, S. I. (1971a). Products of distribution functions attracted to extreme value laws. *J. Appl. Prob.* **8**, 781–793.
- RESNICK, S. I. (1971b). Tail equivalence and its applications. *J. Appl. Prob.* **8**, 136–156.
- RESNICK, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.
- ROOTZÉN, H. (1986). Extreme value theory for moving average processes. *Ann. Prob.* **14**, 612–652.
- WANG, D. AND TANG, Q. (2006). Tail probabilities of randomly weighted sums of random variables with dominated variation. *Stoch. Models* **22**, 253–272.