

AGGREGATION OF RISKS AND ASYMPTOTIC INDEPENDENCE

ABHIMANYU MITRA AND SIDNEY I. RESNICK

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 $P(X + Y > x) \sim (\text{const})P(X > x)$. With the further assumption of non-negativity 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.

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 are found in Albrecher et al. [2006], Kluppelberg and Resnick [2008], Wang and Tang [2006], Asmussen and Rojas-Nandayapa [2005], Alink et al. [2004], Embrechts and Puccetti [2006], 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 $MDA(\Lambda)$, where

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

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

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

[Resnick, 1987, page 38]. 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 $\mathbf{x} = (x_1, x_2)$,

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

where H is the joint distribution of X and Y and both G_1 and G_2 are non-degenerate extreme value distributions [de Haan and Ferreira, 2006, page 229]. There are also results on aggregation

Key words and phrases. Risk, Gumbel, maximal domain of attraction, asymptotic independence, subexponential. S. I. Resnick and A. Mitra were partially supported by ARO Contract W911NF-07-1-0078 at Cornell University.

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 Kluppelberg and Resnick [2008].

This paper considers the case where the risks X, 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 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.

- (1) First, suppose (X, Y) are two iid risks with common distribution F which is subexponential and $F \in MDA(\Lambda)$. Then X and Y are certainly asymptotically independent and

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

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

- (2) Very different tail behavior is exhibited in Theorem 2.10 of Albrecher et al. [2006], who exhibit a distribution of (X, Y) , with X and Y being asymptotically independent and identically distributed with common 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,

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

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, Y are asymptotically independent and each belongs to the maximal domain of attraction of the Gumbel. When X, Y are identically distributed, (1.3) holds. Under the further assumption of non-negativity 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 yet 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 asymptotic behavior of finite linear combinations of risks with non-negative 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 $c = \lim_{x \rightarrow \infty} P(Y > x)/P(X > x) \in [0, \infty)$. When $c \in (0, \infty)$, X and Y are called tail-equivalent [Resnick, 1971a] and then our conditions guarantee that both the marginal distributions $F, G \in MDA(\Lambda)$ 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, (X, Y) is a pair of random variables satisfying the following set of assumptions.

(1) The random variable X has a distribution F whose right endpoint x_0 is infinite; that is,

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

Further $F \in MDA(\Lambda)$ so that (1.1) is satisfied with centering constants $b_n \in \mathbb{R}$ and scaling constants $a_n > 0$. Equivalently (Resnick [1987, page 28, 40-43], de Haan [1970]) there exists a self-neglecting auxiliary function $f(\cdot)$ with its derivative converging to 0, such that

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

(2) The random variables X and Y have distribution functions F and G such that

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

(3) The conditional distribution of Y given $X > x$, satisfies for all $t > 0$,

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

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

(4) and symmetrically assume for all $t > 0$,

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

(5) For some $L > 0$, suppose

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

2.1.2. *The main result.* The assumptions allow us to conclude aggregated risks are essentially tail equivalent to individual risks.

Theorem 2.1. *Under Assumptions 1–5 in Section 2.1.1, 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.

Remark 2.2. (1) If $c \in (0, \infty)$, then our assumptions guarantee both marginal distributions $F, G \in MDA(\Lambda)$ and also that (X, Y) are asymptotically independent. From Assumption 1, $F \in MDA(\Lambda)$ and since F and G are tail-equivalent, from Resnick [1971a] we get that $G \in MDA(\Lambda)$. For asymptotic independence, define,

$$(2.3) \quad b_F(t) = \inf\{s : \frac{1}{1-F}(s) \geq t\} = \left(\frac{1}{1-F}\right)^{\leftarrow}(t),$$

and similarly $b_G(t)$. From [de Haan and Ferreira, 2006, page 229], if $F, G \in MDA(\Lambda)$ and

$$(2.4) \quad \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_G(t))}{P(X > b_F(t))} = 0,$$

then (X, Y) are asymptotically independent according to (1.2). When $c \in (0, \infty)$, Assumption 3 implies (2.4). To verify this, note first that Assumption 3 implies

$$(2.5) \quad \lim_{x \rightarrow \infty} \frac{P(X > x, Y > x)}{P(X > x)} \leq \lim_{x \rightarrow \infty} \frac{P(X > f(x), Y > x)}{P(X > x)} = 0,$$

since $f(x)/x \rightarrow 0$ as $x \rightarrow \infty$ [Resnick, 1987, page 40]. If $c > 1$, then for sufficiently large t , $b_F(t) \leq b_G(t)$ and therefore, using (2.5),

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

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,

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > b_G(t))}{P(X > b_F(t))} &\leq \lim_{t \rightarrow \infty} \frac{P(X > b_F(t), Y > f(b_F(t)))}{P(X > b_F(t))} \\ &= 0 \quad (\text{by Assumption 3 and (2.1)}). \end{aligned}$$

- (2) 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 3, 4, 5 hold with $f(\cdot)$, they also hold with $\tilde{f}(\cdot)$ replacing $f(\cdot)$ and vice versa. Since the mean excess function

$$e(x) = E(X - x | X > x)$$

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

- (3) If $c = \lim_{x \rightarrow \infty} \bar{G}(x)/\bar{F}(x) = 0$, we do not need Assumption 4 to conclude our result.
(4) When $F \in MDA(\Lambda)$, we may choose a_n, b_n appearing in (1.1) as $b_n = b_F(n)$, $a_n = f(b_n)$. See [Resnick, 1987, page 40] or de Haan and Ferreira [2006].

2.1.4. *Proof of Theorem 2.1.* We prove Theorem 2.1 using a sequence of propositions and lemmas similar to those in Goldie and Resnick [1988b] but for our proof, we do not have X, Y independent nor do we need the assumption of sub-exponentiality.

Proposition 2.3. *Under Assumptions 1 and 3 of Section 2.1.1, we have*

$$(2.6) \quad \lim_{n \rightarrow \infty} P(Y \leq a_n z | X > a_n x + b_n) = 1_{\{z > 0\}}, \quad z \neq 0, x \in \mathbb{R}.$$

Proof. Assume, $z > 0$. Using (2.1) and the self-neglecting property of the auxiliary function f ,

$$(2.7) \quad \lim_{t \rightarrow \infty} \frac{f(t + x f(t))}{f(t)} = 1, \quad x \in \mathbb{R},$$

we conclude,

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

Thus, there exists N such that for $n \geq N$, $\frac{f(b_n)}{f(b_n + x f(b_n))} \geq \epsilon$ for some $\epsilon \in (0, 1)$. Therefore, using assumption 3, as $n \rightarrow \infty$,

$$(2.8) \quad \begin{aligned} &P(Y > z \epsilon f(b_n + x f(b_n)) | X > f(b_n)x + b_n) \\ &\leq P(|Y| > z \epsilon f(b_n + x f(b_n)) | X > f(b_n)x + b_n) \rightarrow 0. \end{aligned}$$

Also, from (2.1) and the fact that $f(t)/t \rightarrow 0$, [Resnick, 1987, page 40], $b_n + xf(b_n) \rightarrow \infty$, as $n \rightarrow \infty$, for $x \in \mathbb{R}$. Choose a_n as $a_n = f(b_n)$ and as $n \rightarrow \infty$,

$$\begin{aligned} P(Y \leq a_n z | X > a_n x + b_n) &= P(Y \leq f(b_n)z | X > f(b_n)x + b_n) \\ &= P(Y \leq f(b_n + xf(b_n))z \frac{f(b_n)}{f(b_n + xf(b_n))} | X > f(b_n)x + b_n) \\ &\geq P(Y \leq z\epsilon f(b_n + xf(b_n)) | X > f(b_n)x + b_n) \\ &= 1 - P(Y > z\epsilon f(b_n + xf(b_n)) | X > f(b_n)x + b_n) \rightarrow 1 \end{aligned}$$

from (2.8). To summarize, as required, we have for $z > 0$,

$$P(Y \leq a_n z | X > a_n x + b_n) \rightarrow 1 = 1_{\{z > 0\}}, \quad n \rightarrow \infty.$$

The argument is similar for $z < 0$. □

Proposition 2.4. *Under Assumptions 1 and 4 of Section 2.1.1, we have*

$$(2.9) \quad \lim_{n \rightarrow \infty} P(X \leq a_n z | Y > a_n x + b_n) = 1_{\{z > 0\}}, \quad z \neq 0, x \in \mathbb{R}.$$

Proof. This proof is similar to the proof of Proposition 2.3. □

Lemma 2.5. *Assumptions 1, 2, 3 and 4 imply that the sequence of measures*

$$nP[a_n^{-1}(X - b_n, Y - b_n, Y, X) \in (dx, dy, dz, dw)]$$

converges vaguely on $([-\infty, \infty]^2 \setminus \{(-\infty, -\infty)\}) \times ([-\infty, \infty]^2)$ as $n \rightarrow \infty$, to the limit measure

$$\begin{aligned} m_\infty(dx, dy, dz, dw) &= \\ &e^{-x} dx \epsilon_{-\infty}(dy) \epsilon_0(dz) \epsilon_\infty(dw) + \epsilon_{-\infty}(dx) c e^{-y} dy \epsilon_\infty(dz) \epsilon_0(dw). \end{aligned}$$

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

REGION 1: $(x, \infty) \times [-\infty, y] \times [-\infty, z] \times (w, \infty)$ $z \neq 0, w \neq 0$. As $n \rightarrow \infty$,

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

by Proposition 2.3. To get the first equality, note that

$$\lim_{n \rightarrow \infty} nP\left[\frac{X - b_n}{a_n} > x, \frac{X}{a_n} \leq w\right] = 0, \quad \lim_{n \rightarrow \infty} nP\left[\frac{Y - b_n}{a_n} > y, \frac{Y}{a_n} \leq z\right] = 0,$$

which uses $\lim_{n \rightarrow \infty} b_n/a_n = \lim_{n \rightarrow \infty} b_n/f(b_n) = \infty$.

REGION 2: $(x, \infty) \times [-\infty, y] \times [-\infty, z] \times [-\infty, w]$, $z \neq 0, w \neq 0$. As $n \rightarrow \infty$,

$$\begin{aligned} nP\left[\frac{X - b_n}{a_n} > x, \frac{Y - b_n}{a_n} \leq y, \frac{Y}{a_n} \leq z, \frac{X}{a_n} \leq w\right] &\leq nP\left[\frac{X - b_n}{a_n} > x, \frac{X}{a_n} \leq w\right] \\ &\rightarrow 0 = m_\infty((x, \infty) \times [-\infty, y] \times [-\infty, z] \times [-\infty, w]). \end{aligned}$$

REGION 3: $[-\infty, x] \times (y, \infty) \times (z, \infty) \times [-\infty, w]$, $z \neq 0, w \neq 0$. The argument for convergence on this region is similar to Region 1, except that we use Proposition 2.4. The proof is omitted, but notice that in the special case of $c = 0$ Assumption 4 is unnecessary since as $n \rightarrow \infty$,

$$\begin{aligned} nP\left[\frac{X - b_n}{a_n} \leq x, \frac{Y - b_n}{a_n} > y, \frac{Y}{a_n} > z, \frac{X}{a_n} \leq w\right] &\leq nP\left[\frac{Y - b_n}{a_n} > y\right] \\ &\rightarrow 0 = m_\infty([-\infty, x] \times (y, \infty) \times (z, \infty) \times [-\infty, w]). \end{aligned}$$

Arguments for convergence on the following regions follow in a similar fashion using (2.5) and Propositions 2.3 and 2.4.

REGION 4: $(x, \infty) \times [-\infty, y] \times (z, \infty) \times (w, \infty)$, $z \neq 0, w \neq 0$,

REGION 5: $[-\infty, x] \times (y, \infty) \times [-\infty, z] \times [\infty, w]$, $z \neq 0, w \neq 0$,

REGION 6: $[-\infty, x] \times (y, \infty) \times (z, \infty) \times (w, \infty)$, $z \neq 0, w \neq 0$,

and convergence on

REGION 7: $(x, \infty) \times (y, \infty) \times [-\infty, \infty]^2$

follows easily from (2.5) since as $n \rightarrow \infty$,

$$\begin{aligned} nP\left[\frac{X - b_n}{a_n} > x, \frac{Y - b_n}{a_n} > y\right] &\leq nP\left[\frac{X - b_n}{a_n} > x \wedge y, \frac{Y - b_n}{a_n} > x \wedge y\right] \\ &\rightarrow 0 = m_\infty((x, \infty) \times (y, \infty) \times [-\infty, \infty]^2). \end{aligned}$$

This concludes the proof of vague convergence. \square

Denote point measures on a nice space \mathbb{E} in the following way. For a subset $A \subset \mathbb{E}$ define

$$\epsilon_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

A Radon point measure is the counting function of the form $\sum_j \epsilon_{x_j}(\cdot)$ for a family of points $\{x_j\}$ and such that any relatively compact region contains a finite number of points. Denote the set of all Radon point measures on the nice space \mathbb{E} by $M_p(\mathbb{E})$.

Suppose, now, $\{(X_k, Y_k)\}$ is an iid sequence where $(X_k, Y_k) \stackrel{d}{=} (X, Y)$. Consider two independent and identically distributed PRM's (Poisson random measures) on $(-\infty, \infty]$ each with mean measure $e^{-x} dx$, which are denoted by $\sum_k \epsilon_{j_k^{(1)}}$ and $\sum_k \epsilon_{j_k^{(2)}}$ and assume that they are independent of $\{(X_k, Y_k)\}$. We use the processes $\sum_k \epsilon_{j_k^{(1)}}$ and $\sum_k \epsilon_{j_k^{(2)} + \log c}$ and the second point process is assumed to be identically 0 if $c = 0$. Then, the following lemma holds true.

Lemma 2.6. *As $n \rightarrow \infty$,*

$$\left(\sum_{k=1}^n \epsilon_{(a_n^{-1}(X_k - b_n, Y_k))}, \sum_{k=1}^n \epsilon_{(a_n^{-1}(Y_k - b_n, X_k))} \right) \Rightarrow \left(\sum_k \epsilon_{(j_k^{(1)}, 0)}, \sum_k \epsilon_{(j_k^{(2)} + \log c, 0)} \right)$$

in $M_p([(-\infty, \infty) \times [-\infty, \infty)]^2)$.

Proof. The proof follows from Lemma 2.5 in the same way as in Lemma 2.2 of Goldie and Resnick [1988b]. \square

From Lemma 2.6 and the vague continuity of addition [Resnick, 1987, page 150] we get

$$\sum_{k=1}^n \left[\epsilon_{(a_n^{-1}(X_k - b_n, Y_k))} + \epsilon_{(a_n^{-1}(Y_k - b_n, X_k))} \right] \Rightarrow \sum_k \left[\epsilon_{(j_k^{(1)}, 0)} + \epsilon_{(j_k^{(2)} + \log c, 0)} \right]$$

in $M_p((-\infty, \infty) \times [-\infty, \infty])$. Restrict the domain of the point measures to the compact set $[-M, \infty] \times [-\infty, \infty]$ for $M > 0$, and the restricted version of the previous convergence is

$$\begin{aligned}
 & \sum_{k=1}^n \left[\epsilon_{(a_n^{-1}(X_k - b_n, Y_k))}(\cdot \cap [-M, \infty] \times [-\infty, \infty]) \right] \\
 & \quad + \epsilon_{(a_n^{-1}(Y_k - b_n, X_k))}(\cdot \cap [-M, \infty] \times [-\infty, \infty]) \\
 (2.10) \quad & \Rightarrow \sum_k \left[\epsilon_{(j_k^{(1)}, 0)} \mathbf{1}_{(j_k^{(1)} \geq -M)} + \epsilon_{(j_k^{(2)} + \log c, 0)} \mathbf{1}_{(j_k^{(2)} + \log c \geq -M)} \right]
 \end{aligned}$$

Define $T : (-\infty, \infty) \times [-\infty, \infty] \mapsto (\infty, \infty]$ by

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

and T induces a transformation $T' : M_p([-M, \infty] \times [-\infty, \infty]) \mapsto M_p((\infty, \infty])$ defined by

$$T' \left(\sum_k \epsilon_{(x_k, y_k)} \right) = \sum_k \epsilon_{T(x_k, y_k)}.$$

From Proposition 2.2 of Davis and Resnick [1988], T' is continuous on

$$\{\mu \in M_p([-M, \infty] \times [-\infty, \infty]) : \mu([-M, \infty] \times \{-\infty\}) = 0\}$$

and hence, T' is almost surely continuous with respect to the limit process in (2.10). So, by applying T' to (2.10) we get

$$\begin{aligned}
 & \sum_{k=1}^n \epsilon_{(a_n^{-1}(X_k + Y_k - b_n))} \left[\mathbf{1}_{((a_n^{-1}(X_k - b_n)) \geq -M)} + \mathbf{1}_{((a_n^{-1}(Y_k - b_n)) \geq -M)} \right] \\
 (2.11) \quad & \Rightarrow \sum_k \left[\epsilon_{j_k^{(1)}} \mathbf{1}_{(j_k^{(1)} \geq -M)} + \epsilon_{j_k^{(2)} + \log c} \mathbf{1}_{(j_k^{(2)} + \log c \geq -M)} \right]
 \end{aligned}$$

This leads to a formal statement of the main result.

Theorem 2.7. *Under the assumptions in Section 2.1.1,*

$$(2.12) \quad \sum_{k=1}^n \epsilon_{(a_n^{-1}(X_k + Y_k - b_n))} \Rightarrow \sum_k \left[\epsilon_{j_k^{(1)}} + \epsilon_{j_k^{(2)} + \log c} \right]$$

in $M_p((-\infty, \infty])$ where the limit is $PRM((1+c)e^{-x}dx)$. Consequently, for $x \in \mathbb{R}$,

$$(2.13) \quad \lim_{n \rightarrow \infty} nP[X + Y > a_n x + b_n] = (1+c)e^{-x}$$

and hence

$$(2.14) \quad \lim_{x \rightarrow \infty} \frac{P(X + Y > x)}{P(X > x)} = (1+c).$$

Proof. If we let $M \rightarrow \infty$, the right side of (2.11) converges almost surely to the right side of (2.12) in the vague metric. Define, $d(\delta, n, M)$ as

$$P \left[\rho \left(\sum_{k=1}^n \epsilon_{(a_n^{-1}(X_k + Y_k - b_n))} (\mathbf{1}_k^{(1)} + \mathbf{1}_k^{(2)}), \sum_{k=1}^n \epsilon_{(a_n^{-1}(X_k + Y_k - b_n))} \right) > \delta \right],$$

where ρ is the vague metric and $1_k^{(1)}$ and $1_k^{(2)}$ are the two indicator functions defined on the left side of (2.11). The result will follow from [Resnick, 2007, page 56] Theorem 3.5, if we show for any $\delta > 0$,

$$\lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} d(\delta, n, M) = 0.$$

For any function $h : (-\infty, \infty] \mapsto [0, \infty)$, denote

$$\tilde{h}(n, M) = \left| \sum_{k=1}^n h(a_n^{-1}(X_k + Y_k - b_n))(1 - 1_k^{(1)} - 1_k^{(2)}) \right|.$$

From the definition of the vague metric, it suffices to show,

$$(2.15) \quad \lim_{M \rightarrow \infty} \limsup_{n \rightarrow \infty} P \left[\tilde{h}(n, M) > \delta \right] = 0$$

for any $h : (-\infty, \infty] \mapsto [0, \infty)$ which is continuous with compact support, say $[\theta, \infty]$. We are free to suppose $M > |\theta|$. The probability in (2.15) is bounded by

$$\begin{aligned} & P \left[\bigcup_{k=1}^n [h(a_n^{-1}(X_k + Y_k - b_n)) > 0, 1 - 1_k^{(1)} - 1_k^{(2)} \neq 0] \right] \\ & \leq nP \left[h(a_n^{-1}(X + Y - b_n)) > 0, 1 - 1_1^{(1)} - 1_1^{(2)} \neq 0 \right] \\ & \leq nP \left[a_n^{-1}(X + Y - b_n) > \theta, 1_1^{(1)} = 1_1^{(2)} = 0 \right] \\ & \quad + nP \left[a_n^{-1}(X + Y - b_n) > \theta, 1_1^{(1)} = 1_1^{(2)} = 1 \right] \\ & = I + II. \end{aligned}$$

By (2.5), as $n \rightarrow \infty$,

$$II \leq nP \left[(a_n^{-1}(X - b_n)) \geq -M, (a_n^{-1}(Y - b_n)) \geq -M \right] \rightarrow 0.$$

We use Assumption 5 to make I go to 0 as $n \rightarrow \infty$ as follows:

$$\begin{aligned} I &= nP \left[a_n^{-1}(X + Y - b_n) > \theta, (a_n^{-1}(X - b_n)) < -M, (a_n^{-1}(Y - b_n)) < -M \right] \\ &\leq nP \left[\frac{X - b_n}{a_n} < -M, \frac{Y - b_n}{a_n} < -M, \frac{X}{a_n} > \theta + M, \frac{Y}{a_n} > \theta + M \right] \\ &\leq nP \left[a_n^{-1}X > \theta + M, a_n^{-1}Y > \theta + M \right] \end{aligned}$$

and choosing M large enough that $M > |\theta| + L$, we get

$$\leq nP \left[a_n^{-1}X > L, a_n^{-1}Y > L \right] = \frac{P[Y > Lf(b(n)), X > Lf(b(n))]}{P[X > b(n)]} \rightarrow 0,$$

by (2.1) and Assumption 5. Hence, (2.15) holds and we conclude our result. \square

One immediate application of Theorem 2.7 is to the subexponential family of distributions denoted \mathcal{S} . The class $MDA(\Lambda) \cap \mathcal{S}$ has been studied in applied probability [Embrechts et al., 1997, page 149] and several sufficient conditions for belonging to this class are given in Goldie and Resnick [1988a]. Corollary 2.8 gives an additional sufficient condition. Example 3.2 exhibits a distribution which satisfies the conditions of this Corollary.

Corollary 2.8. *Suppose, $F \in MDA(\Lambda)$ with auxiliary function $f(x)$ as described in Assumption 1 of Section 2.1.1. Suppose, also, $\lim_{x \rightarrow \infty} f(x) = \infty$, and for some $L > 0$,*

$$(2.16) \quad \lim_{x \rightarrow \infty} \frac{[\bar{F}(Lf(x))]^2}{\bar{F}(x)} = 0.$$

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

$$P[X + Y > x] \sim 2P[X > x],$$

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

Following Remark 2.2(2), it is enough to check (2.16) with any $\tilde{f}(x)$ satisfying $\tilde{f}(x) \sim f(x)$. Note also it is natural to add the assumption $f(x) \rightarrow \infty$, since if $F \in MDA(\Lambda) \cap \mathcal{S}$, then necessarily $f(x) \rightarrow \infty$ [Goldie and Resnick, 1988a].

Proof. Suppose, (X, Y) is a pair of iid random variables with common distribution F . If this pair satisfies the assumptions in Section 2.1.1, then by Theorem 2.7, $F \in \mathcal{S}$. Clearly, this pair satisfies Assumptions 1 and 2 of Section 2.1.1 with $c = 1$. Consider Assumptions 3 and 4. For all $t > 0$, by independence,

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

since $\lim_{x \rightarrow \infty} f(x) = \infty$. For Assumption 5, we take L as in (2.16), and observe that by (2.16),

$$\lim_{x \rightarrow \infty} \frac{P(Y > Lf(x), X > Lf(x))}{P(X > x)} = \lim_{x \rightarrow \infty} \frac{[\bar{F}(Lf(x))]^2}{\bar{F}(x)} = 0.$$

□

2.2. Asymptotic tail probability for the sum of more than two non-negative random variables. Suppose, 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 X_1 satisfies Assumption 1 of Section 2.1.1 and that X_1, X_2, \dots, X_d pairwise satisfy the Assumptions 3 and 4 of Section 2.1.1 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 $t > 0$,

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

which implies

$$(2.17) \quad \lim_{x \rightarrow \infty} \frac{P(X_j > tf(x), X_i > x)}{P(X_1 > x)} = 0.$$

Also, suppose, the risks X_1, X_2, \dots, X_d pairwise satisfy Assumption 5 of Section 2.1.1 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

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

or,

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

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

$$(2.18) \quad \lim_{x \rightarrow \infty} \frac{P(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{P(X_1 > x)} = 0.$$

Under the simple additional assumption of non-negativity of the risks, Theorem 2.7 can be extended to more than two risks.

Corollary 2.9. *Assume, X_1, X_2, \dots, X_d are non-negative random variables which pairwise satisfy Assumptions 3, 4, 5 of Section 2.1.1 with the auxiliary function $f(\cdot)$ of X_1 . Moreover, the distribution of X_1 satisfies Assumption 1 of Section 2.1.1 and suppose*

$$(2.19) \quad \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = c_i \in [0, \infty), \quad i = 2, 3, \dots, d.$$

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

$$(2.20) \quad \lim_{n \rightarrow \infty} nP(S_d > a_n x + b_n) = (1 + \sum_{i=2}^d c_i)e^{-x}$$

and hence

$$(2.21) \quad \lim_{x \rightarrow \infty} \frac{P(S_d > x)}{P(X_1 > x)} = (1 + \sum_{i=2}^d c_i)$$

Remark 2.10. (1) *Asymptotic independence of the random variables:* Suppose, 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 Remark 2.2(1). Since the random variables are pairwise asymptotically independent, they are also asymptotically independent [Resnick, 1987, page 291].

(2) *Non-negativity of random variables:* The only additional assumption added to the list in Section 2.1.1 is that the random variables are non-negative.

(3) *Relaxation:* We have shown in (2.17) and (2.18) that pairwise satisfaction of Assumptions 3, 4, 5 of Section 2.1.1 implies that for $1 \leq i \neq j \leq d$, for $t > 0$,

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

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

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

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

Proof. We prove the result by induction under the relaxation Remark 2.10(3). The base case of the induction for $d = 2$ is already proved in Theorem 2.7, so suppose, the result is true for $d = k \geq 2$ and we have

$$(2.22) \quad \lim_{n \rightarrow \infty} nP(S_k > a_n x + b_n) = (1 + \sum_{i=2}^k c_i)e^{-x}$$

and

$$(2.23) \quad \lim_{x \rightarrow \infty} \frac{P(S_k > x)}{P(X_1 > x)} = 1 + \sum_{i=2}^k c_i$$

Therefore, we have

$$(2.24) \quad \lim_{x \rightarrow \infty} \frac{P(X_{k+1} > x)}{P(S_k > x)} = \frac{c_{k+1}}{1 + \sum_{i=2}^k c_i} \in [0, \infty).$$

Provided we can prove the assumptions in Theorem 2.7 with $X = S_k$ and $Y = X_{k+1}$, we have,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{P(S_{k+1} > x)}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{P(S_{k+1} > x)}{P(S_k > x)} \frac{P(S_k > x)}{P(X_1 > x)} \\ &= \left(1 + \frac{c_{k+1}}{1 + \sum_{i=2}^k c_i}\right) \left(1 + \sum_{i=2}^k c_i\right) = \left(1 + \sum_{i=2}^{k+1} c_i\right) \end{aligned}$$

and hence,

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

It remains to check the Assumptions in Theorem 2.7 with $X = S_k$ and $Y = X_{k+1}$. For Assumption 1, call, $h = \log(1 + \sum_{i=2}^k c_i)$. By the induction hypothesis,

$$\begin{aligned} \lim_{n \rightarrow \infty} nP(S_k > a_n x + b_n) &= \lim_{n \rightarrow \infty} nP(X_1 > a_n x + b_n) \frac{P(S_k > a_n x + b_n)}{P(X_1 > a_n x + b_n)} \\ &= e^{-x} e^h = e^{-(x-h)}, \end{aligned}$$

and therefore,

$$\lim_{n \rightarrow \infty} nP(S_k > a_n(x+h) + b_n) = e^{-x}.$$

Thus,

$$\lim_{n \rightarrow \infty} nP(S_k > a_n x + \tilde{b}_n) = e^{-x}, \quad \tilde{b}_n = b_n + h a_n.$$

Assumption 2 is already checked in (2.24). To check Assumption 3, set $c_1 = 1$ and then, for $1 \leq i \leq k$,

$$(2.25) \quad \frac{P((S_k > x) \cap (X_i > x))}{P(X_1 > x)} = \frac{P(X_i > x)}{P(X_1 > x)} = c_i.$$

Note that the first equality is where the assumption that the X_i 's are non-negative is used. For $1 \leq i \neq j \leq k$,

$$(2.26) \quad \lim_{x \rightarrow \infty} \frac{P((S_k > x) \cap (X_i > x) \cap (X_j > x))}{P(X_1 > x)} \leq \lim_{x \rightarrow \infty} \frac{P((X_i > x) \cap (X_j > x))}{P(X_1 > x)} = 0.$$

The last equality follows from (2.5), (2.17) and (2.19).

Using (2.25) and (2.26) we get,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{P((S_k > x) \cap (\cup_{i=1}^k (X_i > x)))}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{\sum_{i=1}^k P((S_k > x) \cap (X_i > x))}{P(X_1 > x)} \\ (2.27) \quad &= \sum_{i=1}^k c_i = 1 + \sum_{i=2}^k c_i. \end{aligned}$$

From (2.27) it follows that

$$\lim_{x \rightarrow \infty} \frac{P((S_k > x) \cap (\cup_{i=1}^k (X_i > x))^c)}{P(X_1 > x)} = 0,$$

and hence from (2.23),

$$(2.28) \quad \lim_{x \rightarrow \infty} \frac{P((S_k > x) \cap (\cup_{i=1}^k (X_i > x))^c)}{P(S_k > x)} = 0.$$

Since S_k and X_1 are tail equivalent, by Resnick [1971a], the auxiliary function $\tilde{f}(\cdot)$ of S_k is asymptotically equal to the auxiliary function $f(\cdot)$ of X_1 . Therefore, given $\epsilon \in (0, 1)$, there exists T such that for all $x > T$, $\tilde{f}(x) > \epsilon f(x)$. We now are prepared to check Assumption 3. For $t > 0, x > T$, using (2.28), as $x \rightarrow \infty$,

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

by (2.17).

For Assumption 4, if $c_{k+1} = 0$, following Remark 2.2(3), there is no need to check assumption 4. So, suppose, $c_{k+1} > 0$. Then for any $t > 0$, as $x \rightarrow \infty$,

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

For Assumption 5, we know from the assumptions in the statement of Corollary 2.9 that the random variables satisfy Assumption 5 of Section 2.1.1 pairwise with auxiliary function $f(\cdot)$ of X_1 . Thus, for $1 \leq i < j \leq d$, (2.18) holds. We check Assumption 5 with $L = kL_{max}/\epsilon$, where $L_{max} = \max_{1 \leq i \leq k} L_{i,k+1}$ (recall, equation (2.18)). Then, for sufficiently large x , using $\tilde{f}(\cdot)$ as the auxiliary function of S_k ,

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

by (2.18). This completes the induction proof. \square

3. EXAMPLES

This section shows a few of the many models that satisfy the Assumptions in Section 2.1.1. In all examples, both X and Y are non-negative random variables and it is straightforward to extend these examples to the d -dimensional case and show the assumptions of Corollary 2.9 are satisfied.

Our conditions are only sufficient and we exhibit one example where our conditions do not hold, but tail equivalence in Theorem 2.7 holds true. Finding a necessary and sufficient condition for the conclusion of Theorem 2.7 is still an open but difficult issue.

Example 3.1. Suppose X_1, X_2, X_3 are iid 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,

$$X = X_1 \wedge X_2, \quad Y = X_2 \wedge X_3$$

It is easy to check X and Y are identically distributed with the common distribution F_1 , where

$$\bar{F}_1(x) = \exp(-2(\log x)^\alpha), \quad x > 1.$$

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 $F_1 \in MDA(\Lambda)$, 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 [Resnick, 1987, page 40]. Also, (2.1) is obvious and therefore, Assumption 1 of Section 2.1.1 is satisfied. Checking Assumption 2 is straightforward, so consider Assumption 3. Fix $t > 0$, recall $f(x)/x \rightarrow 0$, and note as $x \rightarrow \infty$,

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

since $f(x) \rightarrow \infty$. Assumption 4 is verified the same way. For Assumption 5, we have with $L = 1$,

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

Since the exponent in (3.1) converges to $-\infty$ as $x \rightarrow \infty$, Assumption 5 is satisfied and this pair (X, Y) satisfies the Assumptions in Section 2.1.1.

Example 3.2. Suppose, X and Y are independent and identically distributed 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, one can check the subexponentiality condition (2.16) with $L = 1$ and by Corollary 2.8, F is subexponential. Hence,

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

Example 3.3. Suppose, $X \sim \text{Lognormal}(\mu, \sigma^2)$ and $Y = e^{2\mu}/X$ so that $X \stackrel{d}{=} Y$. We check the Assumptions in Section 2.1.1 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 [Embrechts et al., 1997, page 147, 161]

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

Also, (2.1) is obvious and so, Assumption 1 of Section 2.1.1 is true. Following Remark 2.2(2) and the form of $e(x)$, we may assume the auxiliary function

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

To verify Assumption 3, fix $t > 0$, and note as $x \rightarrow \infty$,

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

since $f(x) \rightarrow \infty$. Assumption 4 is verified similarly. For Assumption 5, choose $L = 1$ and as $x \rightarrow \infty$,

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

We conclude by Theorem 2.7,

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

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

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

Assume as usual $x_0 = \sup\{x : F(x) < 1\} = \infty$ and also that $x_1 = \inf\{x : F(x) > 0\} \leq 0$. Distributions satisfying these conditions include the exponential, gamma, lognormal. Define $X = F^\leftarrow(U)$, and $Y = F^\leftarrow(1 - U)$. This pair (X, Y) satisfies the Assumptions in Section 2.1.1.

Checking Assumption 2 is easy since X and Y are identically distributed. To verify Assumption 3, fix $t > 0$ and define $\epsilon_t = F(\frac{t\delta}{2})$. Since, $x_1 \leq 0$, we have $\epsilon_t > 0$. Then, for large x making $f(x) > \delta/2$, we have

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

since $F(x) \rightarrow 1$, and $x_0 = \infty$. Assumption 4 is similarly verified. To verify Assumption 5, choose L such that $F(\frac{L\delta}{2}) > \frac{1}{2}$ and for x sufficiently large,

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

Hence, (X, Y) satisfy the Assumptions of Section 2.1.1 and by Theorem 2.7,

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

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

Example 3.5. Suppose, $X = \exp(X_1)$, $Y = \exp(X_2)$, where (X_1, X_2) is bivariate normal with correlation $\rho \in [-1, 1)$. For simplicity, assume each X_i has mean μ and variance $\sigma^2 > 0$. This example is extensively considered in Asmussen and Rojas-Nandayapa [2005]. We have already considered the case $\rho = -1$ in Example 3.3, so here we consider $\rho \in (-1, 1)$.

Assumptions 1 and 2 of Section 2.1.1 are easily verified. Following the same reason as in Example 3.3, we take the auxiliary function to be

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

Observe,

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

For Assumption 3, we have for $t > 0$, as $x \rightarrow \infty$,

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

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

For Assumption 5, choose $L = 1$. As $x \rightarrow \infty$, we have using (3.3),

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

Example 3.6. Let X_1 and X_2 be independent and identically distributed with the common distribution $H \in MDA(\Lambda)$, having auxiliary function $f_1(\cdot)$ satisfying (3.2) and infinite right end point. Also, suppose, $F \in MDA(\Lambda)$ with auxiliary function $f_2(\cdot)$, concentrates on $(0, \infty)$ and satisfies the conditions in Example 3.4. Define X and Y as

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

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, page 43], the distribution of X belongs to the maximal domain of attraction of Gumbel 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,

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

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) satisfy the assumptions in Section 2.1.1.

Example 3.7. Here is an example of a distribution for (X, Y) where our assumptions are not satisfied, but the asymptotic behavior is the same as in Theorem 2.7. Suppose, X and Y are iid with common distribution F , where

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

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

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

However, this distribution does not satisfy Assumption 5 of Section 2.1.1.

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

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

For any $L > 0$, as $x \rightarrow \infty$,

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

Hence, Assumption 5 is not satisfied for any $L > 0$. This also shows the criteria (2.16) for $F \in \mathcal{S}$ is sufficient but not necessary.

4. LINEAR COMBINATIONS OF RANDOM VARIABLES WITH NON-NEGATIVE COEFFICIENTS

This section studies linear combinations of risks X, Y with non-negative 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 one cannot expect similar behavior in the two cases.

4.1. Tail-equivalent cases.

4.1.1. Linear Combination of two random variables with non-negative coefficients.

Theorem 4.1. *Assume, (U, V) is a pair of random variables which satisfy Assumptions 1, 3, 4 and 5 of Section 2.1.1. Further assume, that Assumption 2 holds in the form*

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

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 $x \rightarrow \infty$,

$$P(\hat{S}_2 > x) \sim P(U > \frac{x}{m_2}) [1_{\{a_1=m_2\}} + c1_{\{a_2=m_2\}}]$$

We assume 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. 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 4.2. Assume, (U, V) are iid random variables with common distribution F , which satisfy Assumptions 1, 3, 4 and 5 of Section 2.1.1. 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 pairs (U_1, V_1) and (U_2, V_2) satisfy Assumptions 1, 3, 4 and 5 of Section 2.1.1. 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 we cannot expect Theorem 4.1 to hold for the case $c = 0$.

We now turn to the proof of Theorem 4.1.

Proof. The case $a_1 = a_2$ is resolved by Theorem 2.7 since

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

So the interesting cases are $a_1 > a_2$ and $a_1 < a_2$ and for the following, assume $a_1 > a_2$; the case $a_1 < a_2$ is handled similarly.

There is nothing to prove if $a_2 = 0$, so assume $a_1 > a_2 > 0$ which makes $m_2 = a_1$.

It suffices to check the assumptions in Section 2.1.1 for $X = U$ and $Y = a_2V/a_1$. For this definition of X, Y , we have

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

The last equality is true from (4.1) and the fact that the tail of any distribution in $MDA(\Lambda)$ is $-\infty$ -varying [Resnick, 1987, page 53]. From Theorem 2.7 and (4.2), we get, as $x \rightarrow \infty$,

$$\begin{aligned} P(a_1U + a_2V > x) &= P(a_1(U + a_2V/a_1) > x) \\ &= P(U + a_2V/a_1 > x/a_1) = P(X + Y > x/a_1) \\ &\sim P(U > x/a_1) = P(U > \frac{x}{m_2}) [1_{\{a_1=m_2\}} + c1_{\{a_2=m_2\}}]. \end{aligned}$$

To complete the proof, the Assumptions in Section 2.1.1 must be verified for $X = U$ and $Y = a_2V/a_1$. Assumption 1 is assumed in the theorem and Assumption 2 was verified in (4.2). For Assumption 3, note that $U \in MDA(\Lambda)$ and suppose $f(\cdot)$ is the auxiliary function of the distribution of U . By hypothesis, for $t > 0$,

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

and therefore, using (4.3),

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

Remark 2.2(3) implies we do not need to verify Assumption 4, so we turn to checking Assumption 5. 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 $a_1 > a_2$. □

4.1.2. *Linear Combination of more than two random variables with non-negative coefficients.*

Corollary 4.3. *Assume, X_1, X_2, \dots, X_d are non-negative random variables which pairwise satisfy Assumptions 3, 4, 5 of Section 2.1.1. Further suppose the distribution of X_1 satisfies Assumptions 1 of Section 2.1.1 and that*

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

Set $c_1 = 1$ and define for $d > 1$, $\hat{S}_d = a_1 X_1 + a_2 X_2 + \dots + a_d X_d$, 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

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

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

The random variables X_1, X_2, \dots, X_d are tail-equivalent and satisfy Assumption 3 of Section 2.1.1 pairwise. Therefore Remark 2.2(1) implies pairwise asymptotic independence and hence, by [Resnick, 1987, page 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.

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

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

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

$$(4.5) \quad m_{k+1} = m_k \vee a_{k+1},$$

and

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

so that

$$(4.6) \quad \frac{N_{k+1}}{c_{k+1}} = \left[1_{\{a_{k+1} = m_{k+1}\}} + \frac{N_k}{c_{k+1}} 1_{\{m_k = m_{k+1}\}} \right].$$

By the induction hypothesis,

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

If we prove 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, as $x \rightarrow \infty$,

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

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

Assumption 1 is assumed. For (4.1), consider that on the one hand,

$$(4.8) \quad N_k = \sum_{\{1 \leq i \leq k: a_i = m_k\}} c_i \geq \bigwedge_{i=1}^k c_i > 0$$

and on the other,

$$(4.9) \quad N_k = \sum_{\{1 \leq i \leq k: a_i = m_k\}} c_i \leq k \bigvee_{i=1}^k c_i < \infty,$$

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

Next, suppose, two random variables U and V are tail equivalent and both belong to $MDA(\Lambda)$. If $f(\cdot)$, $\tilde{f}(\cdot)$ are the auxiliary functions of U and V respectively, then $f(x) \sim \tilde{f}(x)$, as $x \rightarrow \infty$; see Resnick [1971a,b]. Since, in the present case, all the random variables are tail-equivalent, Remark 2.2(2) implies we can work with the auxiliary function of any one of them, say X_{k+1} . So, X_1, X_2, \dots, X_d satisfy Assumptions 3, 4 and 5 of Section 2.1.1 pairwise with the auxiliary function $f(\cdot)$ of X_{k+1} . That is, for $1 \leq i \neq j \leq d$, and any $t > 0$,

$$(4.10) \quad \lim_{x \rightarrow \infty} P(X_j > tf(x) | X_i > x) = 0$$

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

$$(4.11) \quad \frac{P(X_i > L_{ij}f(x), X_j > L_{ij}f(x))}{P(X_i > x)} = 0.$$

To verify Assumption 3, observe for $t > 0$, that as $x \rightarrow \infty$,

$$\begin{aligned} & P(|m_k^{-1}\hat{S}_k| > tf(x) | X_{k+1} > x) \\ & \leq P(a_1X_1 + a_2X_2 + \dots + a_kX_k > m_k tf(x) | X_{k+1} > x) \\ & \leq \sum_{i=1}^k P(X_i > a_i^{-1}m_k tf(x)/k | X_{k+1} > x) \leq \sum_{i=1}^k P(X_i > tf(x)/k | X_{k+1} > x) \rightarrow 0, \end{aligned}$$

by (4.10). For Assumption 4, note,

$$(4.12) \quad \lim_{x \rightarrow \infty} \frac{P(m_k^{-1}\hat{S}_k > x)}{P(X_1 > x)} = N_k,$$

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

$$(4.13) \quad \begin{aligned} \lim_{x \rightarrow \infty} \frac{P((m_k^{-1}\hat{S}_k > x) \cap (m_k^{-1}a_iX_i > x))}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{P(m_k^{-1}a_iX_i > x)}{P(X_1 > x)} \\ &= c_i 1_{\{a_i = m_k\}}. \end{aligned}$$

The first equality uses the assumption that X_i 's are non-negative. The second equality is true from (4.4) and the fact that the tail of any distribution in the maximal domain of attraction of Gumbel is $-\infty$ -varying [Resnick, 1987, page 53]. Now, for $1 \leq i < j \leq k$, using (2.5),

$$\begin{aligned} & \lim_{x \rightarrow \infty} \frac{P((m_k^{-1}\hat{S}_k > x) \cap (m_k^{-1}a_iX_i > x) \cap (m_k^{-1}a_jX_j > x))}{P(X_1 > x)} \\ & \leq \lim_{x \rightarrow \infty} \frac{P((m_k^{-1}a_iX_i > x) \cap (m_k^{-1}a_jX_j > x))}{P(X_1 > x)} \end{aligned}$$

$$\leq \lim_{x \rightarrow \infty} \frac{P((X_i > x) \cap (X_j > x))}{P(X_1 > x)} = 0.$$

Therefore, using (4.13),

$$(4.14) \quad \begin{aligned} & \lim_{x \rightarrow \infty} \frac{P((m_k^{-1} \hat{S}_k > x) \cap (\cup_{i=1}^k (m_k^{-1} a_i X_i > x)))}{P(X_1 > x)} \\ &= \lim_{x \rightarrow \infty} \frac{\sum_{i=1}^k P((m_k^{-1} \hat{S}_k > x) \cap (m_k^{-1} a_i X_i > x))}{P(X_1 > x)} = N_k. \end{aligned}$$

From (4.12) and (4.14) it follows that

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

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

$$(4.15) \quad \lim_{x \rightarrow \infty} \frac{P((m_k^{-1} \hat{S}_k > x) \cap (\cup_{i=1}^k (m_k^{-1} a_i X_i > x))^c)}{P(m_k^{-1} \hat{S}_k > x)} = 0.$$

Now, we check Assumption 4. For $t > 0$, as $x \rightarrow \infty$,

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

where we have used (4.15). Using our induction hypothesis, we get that the quantity above is asymptotically equivalent to

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

by (4.10).

For assumption 5, let, $L = kL_{max}$, where $L_{max} = \max_{1 \leq i \leq k} L_{i,k+1}$ (recall, equation (4.11)). Then using (4.7), (4.15) and (4.11), we have

$$\begin{aligned} & \frac{P(X_{k+1} > kL_{max} f(x), m_k^{-1} \hat{S}_k > kL_{max} f(x))}{P(X_{k+1} > x)} \\ &\sim \frac{P(X_{k+1} > kL_{max} f(x), m_k^{-1} \hat{S}_k > x, \cup_{i=1}^k \{m_k^{-1} a_i X_i > x\})}{P(X_{k+1} > x)} \end{aligned}$$

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

□

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

Theorem 4.4. *Assume, Y_1, Y_2, \dots, Y_d are identically distributed non-negative random variables. Also, assume $a_i, \beta_i \geq 0$, $i = 1, 2, \dots, d$. 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*

$$\begin{aligned}
\beta &= \bigvee_{i=1}^d \beta_i, & q_d &= \bigvee_{\{1 \leq i \leq d: \beta_i = \beta\}} a_i, \\
J_d &= |\{1 \leq i \leq d : \beta_i = \beta, a_i = q_d\}|
\end{aligned}$$

where $|\cdot|$ denotes the size of the set. Suppose, $q_d Y_1^\beta, q_d Y_2^\beta, \dots, q_d Y_d^\beta$ pairwise satisfy Assumptions 3, 4 and 5 of Section 2.1.1 and that the distribution of $q_d Y_1^\beta$ satisfies Assumption 1 of Section 2.1.1 where the auxiliary function $f(x)$ satisfies the additional condition that $f(x) \rightarrow \infty$, as $x \rightarrow \infty$. Then,

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

Remark 4.5. If $\beta_1 > \beta_2$, then $Y_1^{\beta_1}$ and $Y_2^{\beta_2}$ are not tail-equivalent. Note, in this case, the asymptotic approximation of $P(a_1 Y_1^{\beta_1} + a_2 Y_2^{\beta_2} > x)$ does not depend on a_2 .

Theorem 4.4 shows different tail behavior from the tail-equivalent cases but follows the paradigm that only the heaviest tails matter. The Theorem shows that Theorem 2.1 of Asmussen and Rojas-Nandayapa [2005] as a special case of a more general phenomenon. Let $(X_1, X_2, \dots, X_d) \sim N(\mathbf{0}, \Sigma)$, where

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

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, $(q_d Y_1^\beta, q_d Y_2^\beta, \dots, q_d Y_d^\beta)$ satisfy the assumptions of Theorem 4.4, where q_d, β have the same meaning as in Theorem 4.4. 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 2.1 of Asmussen and Rojas-Nandayapa [2005]. The results of that theorem and Theorem 4.4 match.

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

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

To start, suppose, for some $i \in \{2, \dots, d\}$, $\beta_i < \beta$. Then there exists $c \in (0, \infty)$ such that for $y \geq c$, $a_i y^{\beta_i} < \frac{q_d}{2} y^\beta$. For large $x > \frac{q_d}{2} c^\beta$,

$$P(a_i Y_i^{\beta_i} > x) = P(a_i Y_i^{\beta_i} 1_{\{Y_i > c\}} > x) \leq P(\frac{q_d}{2} Y_i^\beta 1_{\{Y_i > c\}} > x)$$

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

Then,

$$(4.16) \quad c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(a_i Y_i^{\beta_i} > x)}{P(q_d Y_1^\beta > x)} \leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_1^\beta > 2x)}{P(q_d Y_1^\beta > x)} = 0.$$

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

$$(4.17) \quad c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(a_i Y_i^\beta > x)}{P(q_d Y_1^\beta > x)} = \lim_{x \rightarrow \infty} \frac{P(q_d Y_1^\beta > \frac{q_d x}{a_i})}{P(q_d Y_1^\beta > x)} = 0.$$

In both the equations (4.16) and (4.17), the last equalities are true from the fact that the tail of any distribution in the maximal domain of attraction of the Gumbel is $-\infty$ -varying [Resnick, 1987, page 53].

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

$$(4.18) \quad c_i = \lim_{x \rightarrow \infty} \frac{P(X_i > x)}{P(X_1 > x)} = \lim_{x \rightarrow \infty} \frac{P(Y_i^\beta > \frac{x}{q_d})}{P(Y_1^\beta > \frac{x}{q_d})} = \lim_{x \rightarrow \infty} \frac{P(Y_1^\beta > \frac{x}{q_d})}{P(Y_1^\beta > \frac{x}{q_d})} = 1.$$

It suffices to check the assumptions in Corollary 2.9 with this set of X_1, X_2, \dots, X_d , since then Corollary 2.9 and (4.16), (4.17), (4.18) would imply, as $x \rightarrow \infty$,

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

Assumption 1 is assumed in the Theorem statement and (2.19) is already shown in (4.16), (4.17) and (4.18). For assumptions 3 and 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, for $t > 0$, for $1 \leq i \neq j \leq d$,

$$(4.19) \quad \lim_{x \rightarrow \infty} P(q_d Y_j^\beta > tf(x) | q_d Y_i^\beta > x) = 0.$$

Using Remark 2.10(3), it is enough to show

$$(4.20) \quad \lim_{x \rightarrow \infty} \frac{P(X_j > tf(x), X_i > x)}{P(X_1 > x)} = 0,$$

and to see this, note that there exists $c \in (0, \infty)$ such that for $y \geq c$, $a_i y^{\beta_i} \vee a_j y^{\beta_j} < q_d y^\beta$. For large x , $x \wedge tf(x) > q_d c^\beta$. Hence,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{P(X_j > tf(x), X_i > x)}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{P(a_j Y_j^{\beta_j} > tf(x), a_i Y_i^{\beta_i} > x)}{P(q_d Y_i^\beta > x)} \\ &= \lim_{x \rightarrow \infty} \frac{P(a_j Y_j^{\beta_j} 1_{\{Y_j > c\}} > tf(x), a_i Y_i^{\beta_i} 1_{\{Y_i > c\}} > x)}{P(q_d Y_i^\beta > x)} \\ &\leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta 1_{\{Y_j > c\}} > tf(x), q_d Y_i^\beta 1_{\{Y_i > c\}} > x)}{P(q_d Y_i^\beta > x)} \\ &= \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > tf(x), q_d Y_i^\beta > x)}{P(q_d Y_i^\beta > x)} = 0. \end{aligned}$$

For Assumption 5, using Remark 2.10(3), we show, for some $L > 0$,

$$(4.21) \quad \lim_{x \rightarrow \infty} \frac{P(X_j > Lf(x), X_i > Lf(x))}{P(X_1 > x)} = 0.$$

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

$$(4.22) \quad \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > L_{ij} f(x), q_d Y_i^\beta > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} = 0.$$

Also, note that there exists $c \in (0, \infty)$ such that for $y \geq c$, $a_i y^{\beta_i} \vee a_j y^{\beta_j} < q_d y^\beta$. For large x , $L_{ij} f(x) > q_d c^\beta$. Hence,

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{P(X_j > L_{ij} f(x), X_i > L_{ij} f(x))}{P(X_1 > x)} &= \lim_{x \rightarrow \infty} \frac{P(a_j Y_j^{\beta_j} > L_{ij} f(x), a_i Y_i^{\beta_i} > L_{ij} f(x))}{P(q_d Y_1^\beta > x)} \\ &= \lim_{x \rightarrow \infty} \frac{P(a_j Y_j^{\beta_j} 1_{\{Y_j > c\}} > L_{ij} f(x), a_i Y_i^{\beta_i} 1_{\{Y_i > c\}} > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} \\ &\leq \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta 1_{\{Y_j > c\}} > L_{ij} f(x), q_d Y_i^\beta 1_{\{Y_i > c\}} > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} \\ &= \lim_{x \rightarrow \infty} \frac{P(q_d Y_j^\beta > L_{ij} f(x), q_d Y_i^\beta > L_{ij} f(x))}{P(q_d Y_i^\beta > x)} = 0 \end{aligned}$$

by (4.22). □

5. AN OPTIMIZATION PROBLEM

5.1. The problem. Suppose, 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 revenue $\$L$. Assume, each unit of the i -th instrument earns $\$l_i$ over the chosen time horizon. Subject to earnings being at least $\$L$, 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:

$$\begin{aligned} \min_{\{a_1, \dots, a_d\}} & P \left[\sum_{i=1}^d a_i X_i > x \right] \\ \text{s.t.} & \quad a_1 l_1 + a_2 l_2 + \dots + a_d l_d \geq L, \\ & \quad a_i \geq 0, \quad i = 1, 2, \dots, d. \end{aligned}$$

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

$$\begin{aligned} \min_{\{a_1, \dots, a_d\}} & P \left[\sum_{i=1}^d a_i X_i > x \right] \\ \text{s.t.} & \quad h(a_1, a_2, \dots, a_d) \geq L, \\ & \quad a_i \geq 0, \quad i = 1, 2, \dots, d. \end{aligned}$$

5.2. The method. Suppose, X_1, X_2, \dots, X_d satisfy the assumptions of Corollary 4.3. Even with these assumptions, exact solution of the optimization problem is difficult. An obvious way to obtain an approximate solution to the optimization problem is to assume the threshold x is big and use the asymptotic approximation of $P(a_1X_1 + a_2X_2 + \dots + a_dX_d > x)$ from Corollary 4.3, hoping that the solution of the resulting optimization problem is close to the actual optimal value. So, using the notation of Corollary 4.3, we solve the following optimization problem:

$$\begin{aligned} \min_{\{a_1, \dots, a_d\}} \quad & N_d P(X_1 > \frac{x}{m_d}) \\ \text{s.t.} \quad & h(a_1, a_2, \dots, a_d) \geq L, \\ & a_i \geq 0, \quad i = 1, 2, \dots, d. \end{aligned}$$

Suppose $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_d$ 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, $P[X_1 \leq x] \in MDA(\Lambda)$, as $x \rightarrow \infty$,

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

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

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

We conclude $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_d$ is a better feasible solution for the optimization problem.

Thus, 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

$$\begin{aligned} \min_{\{a_1, \dots, a_d\}} \quad & m_d = \max\{a_1, a_2, \dots, a_d\} \\ \text{s.t.} \quad & h(a_1, a_2, \dots, a_d) \geq L, \\ & a_i \geq 0, \quad i = 1, 2, \dots, d. \end{aligned}$$

(ii) Suppose, 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

$$\begin{aligned} \min_{\{a_1, \dots, a_d\}} \quad & N_d = \sum_{\{1 \leq i \leq d: a_i = m\}} c_i \\ \text{s.t.} \quad & h(a_1, a_2, \dots, a_d) \geq L, \\ & \max\{a_1, a_2, \dots, a_d\} = m, \\ & a_i \geq 0 \quad i = 1, 2, \dots, d. \end{aligned}$$

5.3. A special case. The motivating case is that h is a linear function with positive coefficients of the form

$$h(a_1, a_2, \dots, a_d) = a_1 l_1 + a_2 l_2 + \dots + a_d l_d.$$

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

$$a_1 = a_2 = \dots = a_d = L/(l_1 + l_2 + \dots + l_d).$$

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 for fixed large thresholds the accuracy of the asymptotic approximation in Theorem 2.7 and also to check how good is the approximate solution for the optimization problem. As expected, in some cases the approximation works well whereas in others it performs poorly which suggests caution about using the asymptotic results for numerical purposes. Simulation also suggests that the approximate solution of the optimization problem works well in cases where simulation studies suggest that the approximation is good for fixed large thresholds. One particular model studied, Example 3.5 with $\mu = 0, \sigma = 1$, 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 have 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 [Asmussen and Glynn, 2007, page 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 [2005] for $\rho \in (-1, 1)$ who also note 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:

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

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 is a particular threshold crossing. 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

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 |

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 when $\rho = 1$.

The tables give representative results. We first give the results for $\rho = -1$ in Table 1, since in this case no simulation is required. The column ‘Ratio’ in Table 1 is defined as

$$\text{Ratio} = \frac{\text{Actual probability}}{\text{Asymptotic approximation}}.$$

For subsequent tables, the columns ‘Ratio’ and ‘Half-width’ are defined as

$$\text{Ratio} = \frac{\text{Simulation estimated probability}}{\text{Asymptotic approximation}}$$

Half-width = Half-width of the 95% confidence interval of the ratio.

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

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 |

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 approximation. 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 get the estimates proceed as follows: For $a_1, a_2 > 0$

$$\begin{pmatrix} a_1X \\ a_2Y \end{pmatrix} = \begin{pmatrix} \exp\{\log(a_1) + X_1\} \\ \exp\{\log(a_2) + X_2\} \end{pmatrix}$$

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 |

Now,

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

So, again we are in the framework of Asmussen and Rojas-Nandayapa [2005], and we use the estimate given in their paper. 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. $P(\tilde{a}_1X + \tilde{a}_2Y > x) = \min_{(a_1, a_2) \in C^*} P(a_1X + a_2Y > x)$. 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 3 cases where $\rho = -0.9, 0, 0.9$, are chosen, the reason being that we know from the results in 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 of $\rho = -0.9$ 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 which implies indifference to the choice of $(a_1, a_2) \in C$.

Another remark is that in each case of $\rho = -0.9, 0, 0.9$, 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 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 of $\rho = -0.9$, 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 this case of $\rho = 0.9$, 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, possibly $P(a_1X + a_2Y > x)$ 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 as mentioned before that the case $\rho = 0.9$ is similar to case $\rho = 1$. Some of the results are noted in tables below.

Results are summarized in the tables 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

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

- $E1 = \min_{(a_1, a_2) \in C^*} P(a_1X + a_2Y > x)$,
- $E2 = P(a_1^*X + a_2^*Y > x)$,
- the ‘Relative error’ = $\frac{E2-E1}{E1}$.

For each value of ρ , a_1 is 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 10000 observations to obtain the estimates of the probability $P(a_1X + a_2Y > x)$.

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 |

7. CONCLUDING REMARKS

An important case for the study of 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 in 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 tail but the distribution does not belong to the maximal domain of attraction of the Gumbel. We are 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, identically distributed, belong to $MDA(\Lambda)$, and the asymptotic behavior of the sum is different than 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 how good the approximation $2P(X > x)$ is for the quantity $P(X + Y > x)$ for a large threshold x . Simulation studies indicate in certain circumstances the approximation is accurate, but in other cases its accuracy is dismal.

We observed in the previous section that when 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.

REFERENCES

- H. Albrecher, S. Asmussen, and D. Kortschak. Tail asymptotics for the sum of two heavy-tailed dependent risks. *Extremes*, 9(2):107–130, 2006.
- S. Alink, M. Löwe, and M. V. Wüthrich. Diversification of aggregate dependent risks. *Insurance Mathematics and Economics*, 35(1):77–95, 2004.
- S. Asmussen and P. Glynn. *Stochastic Simulation: Algorithms and Analysis*. Stochastic Modeling and Applied Probability. Springer-Verlag, New York, NY, 2007.
- S. Asmussen and L. Rojas-Nandayapa. Sums of dependent lognormal random variables: asymptotics and simulation. *Preprint*, 2005.
- R.A. Davis and S.I. Resnick. Extremes of moving averages of random variables from the domain of attraction of the double exponential distribution. *Stochastic Processes Appl.*, 30(1):41–68, 1988. ISSN 0304-4149.
- L. de Haan. *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematisch Centrum Amsterdam, 1970.
- L. de Haan and A. Ferreira. *Extreme Value Theory: An Introduction*. Springer-Verlag, New York, 2006.
- P. Embrechts and G. Puccetti. Bounds for Functions of Dependent Risks. *Finance and Stochastics*, 10(3):341–352, 2006.
- P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin, 1997.
- C.M. Goldie and S. Resnick. Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Adv. Appl. Prob.*, 20(4):706–718, 1988a.
- C.M. Goldie and S. I. Resnick. Subexponential distribution tails and point processes. *Stochastic Models*, 4(2):361–372, 1988b.
- C. Kluppelberg and S.I. Resnick. The pareto copula, aggregation of risks and the emperor’s socks. *Journal of Applied Probability.*, 45(1):67–84, 2008.
- B. Ko and Q. Tang. Sums of dependent nonnegative random variables with subexponential tails. *Journal of Applied Probability.*, 45(1):85–94, 2008.
- S.I. Resnick. Tail equivalence and its applications. *J. Appl. Probab.*, 8:136–156, 1971a.
- S.I. Resnick. Products of distribution functions attracted to extreme value laws. *J. Appl. Probability*, 8:781–793, 1971b. ISSN 0021-9002.
- S.I. Resnick. *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York, 1987.
- S.I. Resnick. *Heavy Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer-Verlag, New York, 2007. ISBN: 0-387-24272-4.
- H. Rootzén. Extreme value theory for moving average processes. *Ann. Probab*, 14(2):612–652, 1986.
- D. Wang and Q. Tang. Tail probabilities of randomly weighted sums of random variables with dominated variation. *Stochastic Models*, 22(2):253–272, 2006.

ABHIMANYU MITRA, SCHOOL OF OR&IE, CORNELL UNIVERSITY, ITHACA, NY-14853
E-mail address: am492@cornell.edu

SIDNEY I. RESNICK, SCHOOL OF OR&IE, CORNELL UNIVERSITY, ITHACA, NY-14853
E-mail address: sir1@cornell.edu