Extremes
DOI 10.1007/s10687-016-0242-8

Multivariate subexponential distributions and their applications

Gennady Samorodnitsky[1] **· Julian Sun**[2]

Received: 16 September 2015 / Revised: 1 February 2016 / Accepted: 3 February 2016

Abstract We propose a new definition of a multivariate subexponential distribution. We compare this definition with the two existing notions of multivariate subexponentiality, and compute the asymptotic behaviour of the ruin probability in the context of an insurance portfolio, when multivariate subexponentiality holds. Previously such results were available only in the case of multivariate regularly varying claims.

Keywords Heavy tails · Subexponential distribution · Regular variation · Multivariate · Insurance portfolio · Ruin probability

AMS 2000 Subject Classifications Primary 60E05 · 91B30. Secondary 60G70

1 Introduction

Subexponential distributions are commonly viewed as the most general class of heavy tailed distributions. The notion of subexponentiality was introduced by Chistyakov

This research was partially supported by the ARO grant W911NF-12-10385 at Cornell University.

✉ Gennady Samorodnitsky
gs18@cornell.edu

Julian Sun
ys598@cornell.edu

1 School of Operations Research and Information Engineering and Department of Statistical Science, Cornell University, Ithaca, NY 14853, USA

2 School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, USA

Published online: 18 February 2016

(1964) for distributions supported by $[0, \infty)$; if F is such a distribution, and X_1, X_2 are i.i.d. random variables with the law F, then F is subexponential if

$$\lim_{x\to\infty} \frac{P(X_1 + X_2 > x)}{P(X_1 > x)} = 2\,. \tag{1.1}$$

The notion of subexponentiality was later extended to distributions supported by the entire real line (and not only by the positive half-line); see e.g. Willekens (1986). The best known subclass of subexponential distributions is that of regularly varying distributions, but the membership in the class of subexponential distributions does not require power-like tails; we review the basic information on one-dimensional subexponential distributions in Section 2.

The definition (1.1) of subexponential distributions means that the sum of two i.i.d. random variables with a subexponential distribution is large only when one of these random variables is large. The same turns out to be true for the sum of an arbitrary finite number of terms and, in many cases, for the sum of a random number of terms. Theoretically, this leads to the "single large jump" structure of large deviations for random walks with subexponentially distributed steps; see e. g. Foss et al. (2007). In practice, this has turned out to be particularly important in applications to ruin probabilities. In ruin theory the situation where the claim sizes (often assumed to be independent with identical distribution) have a subexponential distribution is usually referred to as the non-Cramér case. The "single large jump" property of subexponential distributions leads to a well known form of the asymptotic behaviour of the ruin probability, and to a particular structure of the surplus path leading to the ruin; see e.g. Embrechts et al. (1997) and Asmussen (2000).

It is desirable to have a notion of a multivariate subexponential distribution. The task is of a clear theoretical interest, and it is of an obvious interest in applications. A typical insurance company, for instance, has multiple insurance portfolios, with dependent claims, so it would be useful if one could build a model in which claims could be said to have a multivariate subexponential distribution. Recall that there exists a well developed notion of a multivariate distribution with regularly varying tails; see e.g. Resnick (2007). In comparison, a notion of a multivariate subexponential distribution has not been developed to nearly the same extent. To the best of our knowledge, a notion of multivariate subexponentiality has been introduced twice, in Cline and Resnick (1992) and in Omey (2006). Both of these papers define a class (or classes) of multivariate distributions that extend the one-dimensional notion of a subexponential distribution in a natural way. They show that their notions of multivariate subexponentiality possess multidimensional analogs of important properties of one-dimensional subexponential distributions. Nonetheless, these notions have not become as widely used as that of, say, a multivariate distribution with regularly varying tails. In this paper we introduce yet another notion of multivariate subexponential distribution. As the reader will observe, this notion is created with ruin probability applications in mind. We hope, therefore, that this notion will turn out to be useful in that area. However, we also hope that the notion we introduce will be found useful in other areas as well.

This paper is organized as follows. In Section 2 we review the basic properties of one-dimensional subexponential distributions, in order to have a benchmark for the properties we would like a multivariate subexponential distribution to have. In Section 3 we discuss the definitions of multivariate subexponentiality of Cline and Resnick (1992) and in Omey (2006). Our notion of multivariate subexponential distributions is introduced in Section 4. Some applications of that notion to multivariate ruin problems are discussed in Section 5.

2 A review of one-dimensional subexponentiality

In this section we review the basic properties of one-dimensional subexponential distributions. We denote the class of such distributions (and random variables with such distributions) by $\mathscr{S}$. Unless stated explicitly, we do not assume anymore that a random variable with a subexponential distribution F is nonnegative; such a random variable (or its distribution) is called subexponential if the nonnegative random variable $X_+ = \max(X, 0)$ is subexponential. Most of the not otherwise attributed facts stated below can be found in Embrechts et al. (1979). We use the standard notation $\bar{F} = 1 - F$ for the tail of a distribution F.

If a distribution $F \in \mathscr{S}$, then F is *long-tailed*: for any $y \in \mathbb{R}$,

$$\lim_{x\to\infty} \frac{\bar{F}(x+y)}{\bar{F}(x)} = 1 \tag{2.1}$$

(implicitly assuming that $\bar{F}(x) > 0$ for all x.) The class of all long-tailed distributions is denoted by $\mathscr{L}$. Note that $\mathscr{S}$ is a proper subset of $\mathscr{L}$; see e.g. Embrechts and Goldie (1980). Furthermore, the class $\mathscr{L}$ of long-tailed distributions is closed under convolutions, while the class $\mathscr{S}$ of subexponential distributions is not, see Leslie (1989).

A distribution F has a regularly varying right tail if there is $\alpha \geq 0$ such that for every $b > 0$

$$\lim_{x\to\infty} \frac{\bar{F}(bx)}{\bar{F}(x)} = b^{-\alpha}\,, \tag{2.2}$$

and the parameter α is the exponent of regular variation. The class of distributions with a regularly varying right tail is denoted by $\mathscr{R}$ (or $\mathscr{R}(\alpha)$ if we wish to emphasize the exponent of regular variation.) Then $\mathscr{R} \subset \mathscr{S}$. If one views $\mathscr{R}$ as the class of distributions with "power-like" right tails, all distributions with "power-like" right tails are subexponential (note that here and in the sequel we are stretching even this informal terminology somewhat since the case $\alpha = 0$ refers to distributions whose tail is slowly varying, i.e. heavier than any power tail). This informal statement, however, should be treated carefully; other classes of distributions can be referred to as having "power-like" right tails, and not all of them form subclasses of $\mathscr{S}$. Indeed, consider the class $\mathscr{D}$ of distributions with *dominated varying tails*, defined by the property

$$\liminf_{x\to\infty} \frac{\bar{F}(2x)}{\bar{F}(x)} > 0\,. \tag{2.3}$$

One could view a distribution $F \in \mathscr{D}$ as having a "power-like" right tail. However, $\mathscr{D} \not\subset \mathscr{S}$. We note, on the other hand, that it is still true that $\mathscr{D} \cap \mathscr{L} \subset \mathscr{S}$; see Goldie (1978).

Many distributions that do not have "power-like" right tails are subexponential as well. Examples include the log-normal distribution, as well as the Weibull distribution with the shape parameter smaller than 1; see e.g. Pitman (1980).

Let $X_1, X_2, \ldots$ be i.i.d. random variables with a subexponential distribution. The defining property (1.1) extends, automatically, to any finite number of terms, i.e.

$$\lim_{x\to\infty} \frac{P(X_1 + \ldots + X_n > x)}{P(X_1 > x)} = n \quad \text{for any } n \geq 1. \tag{2.4}$$

Moreover, the number of terms can also be random. Let N be a random variable independent of the i.i.d. sequence $X_1, X_2, \ldots$ and taking values in the set of nonnegative integers. If

$$E\tau^N < \infty \quad \text{for some } \tau > 1, \tag{2.5}$$

then

$$\lim_{x\to\infty} \frac{P(X_1 + \ldots + X_N > x)}{P(X_1 > x)} = EN\,. \tag{2.6}$$

The classical one-dimensional (Cramér-Lundberg) ruin problem can be described as follows. Suppose that an insurance company has an initial capital $u > 0$. The company receives a stream of premium income at a constant rate $c > 0$ per unit of time. The company has to pay claims that arrive according to a rate λ Poisson process. The claim sizes are assumed to be i.i.d. with a finite mean μ and independent of the arrival process. If $U(t)$ is the capital of the company at time $t \geq 0$, then the ruin probability is defined as the probability the company runs out of money at some point. This probability is, clearly, a function of the initial capital u, and it is often denoted by

$$\psi(u) = P\big(U(t) < 0 \ \text{ for some } t \geq 0\big)\,.$$

The *positive safety loading*, or the *net profit condition*,

$$\rho := \frac{c}{\lambda\mu} - 1 > 0$$

says that, on average, the company receives more in premium income than it spends in claim payments. If the net profit condition fails, then an eventual ruin is certain. If the net profit condition holds, then the ruin probability is a number in (0, 1), and its behaviour for large values of the initial capital u strongly depends on the properties of the distribution F of the claim sizes. Let

$$F_I(x) = \frac{1}{\mu}\int_0^x \bar{F}(y)\,dy, \;\; x \geq 0$$

be the integrated tail distribution. If $F_I \in \mathscr{S}$, then

$$\psi(u) \sim \rho^{-1} F_I(u) \quad \text{as } u \to \infty; \tag{2.7}$$

see Theorem 1.3.6 in (Embrechts et al. 1997).

3 Existing definitions of multivariate subexponentiality

In this section, and in the rest of the paper, we will use the notation F, interchangingly, as denoting a law (so that the notation $F(A)$ for a Borel set A makes sense), or as denoting a distribution function (so that the notation $F(x_1, \ldots, x_d)$ makes sense.) Correspondingly, we will use the notation $\overline{F} = 1 - F$ for the complementary distribution function F on $\mathbb{R}^d$. That is, if F is the law of a random vector $(X^{(1)}, \ldots, X^{(d)})$, then $\overline{F}(x_1, \ldots, x_d) = P\big(X^{(j)} > x_j$ for some $j = 1, \ldots, d\big)$.

The first known to us definition of multivariate subexponential distributions was introduced by Cline and Resnick (1992). They consider distributions supported by the entire d-dimensional space $\mathbb{R}^d$ (and not only by the nonnegative orthant). That paper defines both multivariate subexponential distributions, and multivariate *exponential distributions*. In our discussion here we only consider the subexponential case. The definition is tied to a function $\mathbf{b}(t) = (b_1(t), \ldots, b_d(t))$ such that $b_i(t) \to \infty$ as $t \to \infty$ for $i = 1, \ldots, d$.

One starts with defining the class of long-tailed distributions, i.e. a multivariate analog of the class $\mathscr{L}$ in Eq. 2.1. Let $E = [-\infty, \infty]^d \setminus \{-\boldsymbol{\infty}\}$, and let ν be a finite measure on E concentrated on the purely infinite points, i.e. on $\{-\infty, \infty\}^d \setminus \{-\boldsymbol{\infty}\}$, and such that $\nu(\mathbf{x} \in E : x_i = \infty) > 0$ for each $i = 1, \ldots, d$. Then a probability distribution F is said to belong to the class $\mathscr{L}(\nu; \mathbf{b})$ if, as $t \to \infty$,

$$tF\big(\mathbf{b}(t) + \cdot\big) \xrightarrow{v} \nu \tag{3.1}$$

vaguely in E (see Resnick (1987) for a thorough treatment of vague convergence of measures.) The class of subexponential distributions (with respect to the same function $\mathbf{b}$ and the same measure ν) is defined to be that subset $\mathscr{S}(\nu; \mathbf{b})$ of distributions F in $\mathscr{L}(\nu; \mathbf{b})$ for which

$$tF * F\big(\mathbf{b}(t) + \cdot\big) \xrightarrow{v} 2\nu$$

vaguely in E.

Corollary 2.4 in Cline and Resnick (1992) shows that $F \in \mathscr{S}(\nu; \mathbf{b})$ if and only if $F \in \mathscr{L}(\nu; \mathbf{b})$ and the marginal distribution F_i of F is in the one-dimensional subexponential class $\mathscr{S}$ for each $i = 1, \ldots, d$.

It is shown in Cline and Resnick (1992) that the distributions in $\mathscr{S}(\nu, \mathbf{b})$ possess the natural multivariate extensions of the properties of the one-dimensional subexponential distributions mentioned in Section 2. For example, if $F \in \mathscr{S}(\nu, \mathbf{b})$, then for any $n \geq 1$, $F^{*n} \in \mathscr{S}(n\nu, \mathbf{b})$. More generally, if N is a random variable satisfying (2.5), and $H = \sum_{n=0}^{\infty} P(N = n) F^{*n}$, then $H \in \mathscr{S}(EN\nu, \mathbf{b})$.

The distributions in $\mathscr{S}(\nu, \mathbf{b})$ also possess the right relation with the distributions with multivariate regularly varying tails. It is natural, in this situation, to consider only distributions supported by the nonnegative quadrant $\mathbb{R}_+^d = [0, \infty)^d$. Recall that any distribution F supported by $\mathbb{R}_+^d$ is said to have regularly varying tails if there is a Radon measure μ on $[0, \infty]^d \setminus \{\mathbf{0}\}$ concentrated on finite points, and a function $\mathbf{b}$ as above such that, as $t \to \infty$,

$$tF\big(\mathbf{b}(t)\cdot\big) \xrightarrow{v} \mu \tag{3.2}$$

vaguely in $[0, \infty]^d \setminus \{\mathbf{0}\}$; see Resnick (2007). Note that Eq. 3.2 allows for different scaling in different directions, hence also different marginal exponents of regular variation. This situation is sometimes referred to as *non-standard regular variation*. If we denote by $\mathscr{R}(\mu, \mathbf{b})$ the class of distributions with regularly varying tails satisfying (3.2), then, as shown in Cline and Resnick (1992), $\mathscr{R}(\mu, \mathbf{b}) \subset \mathscr{S}(\nu, \mathbf{b})$ for some ν.

As mentioned above, this definition of multivariate subexponentiality requires, beyond marginal subexponentiality for all components, only the joint long tail property (3.1). This property, together with the nature of the limiting measure, makes this notion somewhat inconvenient in applications, because it is not easy to see how to use it on sets in $\mathbb{R}^d$ that are not "asymptotically rectangular".

Another observation worth making is that in probability theory, many well established multivariate extensions of important one-dimensional notions have a "stability property" with respect to projections on one-dimensional subspaces (i.e., with respect to taking linear combinations of the components.) Specifically, if the distribution of a random vector $(X^{(1)}, \dots, X^{(d)})$ has, say, a property $\mathscr{G}_d$ (the subscript d specifying the dimension in which the property holds), then the distribution of any (non-degenerate) linear combination $\sum_1^d a_i X^{(i)}$ has the property $\mathscr{G}_1$. This is true, for instance, for multivariate regular variation, multivariate Gaussianity, stability and infinite divisibility. Unfortunately, the definition of multivariate subexponentiality by $\mathscr{S}(\nu, \mathbf{b})$ does not have this feature, as the following example shows.

Example 3.1 Consider a 2-dimensional random vector (X, Y) with nonnegative coordinates such that $P(X + Y = 2^n) = 2^{-(n+1)}$ for $n \ge 0$, with the mass distributed uniformly on the simplex $\{(x, y) : x, y \ge 0,\ x + y = 2^n\}$ for each $n \ge 0$. It is elementary to check that $X, Y \in \mathscr{L} \cap \mathscr{D} \subset \mathscr{S}$. Furthermore, for $2^n \le x \le 2^{n+1}$, $n = 0, 1, 2, \dots$ we have

$$P(X > x) = P(Y > x) = 2^{-(n+1)} - \frac{x}{3} 2^{-(2n+1)} = 2P(X > x, Y > x) .$$

If we define a function b by $tP(X > b(t)) = 1$ for $t \ge 2$, then it immediately follows that $(X, Y) \in \mathscr{L}(\nu; \mathbf{b})$ with $\mathbf{b}(t) = (b(t), b(t))$ and

$$\nu = \frac{1}{2}\delta_{(-\infty,\infty)} + \frac{1}{2}\delta_{(\infty,-\infty)} + \frac{1}{2}\delta_{(\infty,\infty)} ,$$

and the result of Cline and Resnick (1992) tells us that $(X, Y) \in \mathscr{S}(\nu; \mathbf{b})$. It is clear, however, that

$$\liminf_{x\to\infty} \frac{P(X + Y > x + 1)}{P(X + Y > x)} = \frac{1}{2} ,$$

so $X + Y$ does not even have a long-tailed, let alone subexponential, distribution.

The second existing definition of multivariate subexponentiality we are aware of is due to Omey (2006). Once again, this definition concentrates on rectangular regions. The paper presents 3 versions of the definition. The versions are similar, and we concentrate only on one of them. Let F be a probability distribution supported by the nonnegative quadrant in $\mathbb{R}^d$. Then one says that $F \in S(\mathbb{R}^d)$ if for all $\mathbf{x} \in (0, \infty]^d$ with $\min(x_i) < \infty$,

$$\lim_{t\to\infty} \frac{\overline{F^{*2}}(t\mathbf{x})}{\overline{F}(t\mathbf{x})} = 2. \tag{3.3}$$

In subsequent papers, Baltrunas et al. (2006) and Omey and Vesilo (2011), the authors studied, under different assumptions, the rate of convergence to the limit in Eq. 3.3.

The definition (3.3), like the definition of Cline and Resnick (1992), has the following property: a distribution $F \in S(\mathbb{R}^d)$ if and only if each marginal distribution

F_i of F is a one-dimensional subexponential distribution, and a multivariate long-tail property holds. In the present case the long-tail property is

$$\lim_{t\to\infty} \frac{\overline{F}(t\mathbf{x}-\mathbf{a})}{\overline{F}(t\mathbf{x})} = 1 \tag{3.4}$$

for each $\mathbf{x} \in (0,\infty]^d$ with $\min(x_i) < \infty$ and each $\mathbf{a} \in [0,\infty)^d$. This follows from Theorem 7 and Corollary 11 in Omey (2006).

The following statement shows that, in fact, the definition (3.3) of multivariate subexponentiality requires *only* marginal subexponentiality of each coordinate. In fact, this statement was already mentioned in Baltrunas et al. (2006) as a by-product of their Proposition 11.

Proposition 3.2 *Let F be a probability distribution supported by the positive quadrant in $\mathbb{R}^d$. Then $F \in S(\mathbb{R}^d)$ if and only if all marginal distributions F_i of F are subexponential in one dimension.*

Proof By choosing $\mathbf{x}$ with only one finite coordinate, we immediately see that if $F \in S(\mathbb{R}^d)$, then $F_i \in \mathscr{S}$ for each $i = 1, \ldots, d$.

In the other direction, we know by the results of Omey (2006), that only the long-tail property (3.4) is needed, in addition to the marginal subexponentiality, to establish that $F \in S(\mathbb{R}^d)$. Therefore, it is enough to check that the long-tail property (3.4) follows from the marginal subexponentiality. In fact, we will show that, if each F_i is long-tailed, i.e. satisfies (2.1), $i = 1, \ldots, d$, then (3.4) holds as well.

Let $\epsilon > 0$. Fix $\mathbf{x} = (x_1, \ldots, x_d) \in (0,\infty)^d$ (allowing some of the components of $\mathbf{x}$ be infinite only leads to a reduction in the dimension), and $\mathbf{a} = (a_1, \ldots, a_d) \in [0,\infty)^d$.

Since $F_i \in \mathscr{L}$, $i = 1, \ldots, d$, for sufficiently large t we have

$$0 \le \overline{F_i}(tx_i - a_i) - \overline{F_i}(tx_i) < \epsilon \overline{F_i}(tx_i)$$

for $i = 1, \ldots, d$. Further, it is clear that

$$0 \le \overline{F}(t\mathbf{x}-\mathbf{a}) - \overline{F}(t\mathbf{x}) \le \sum_{i=1}^{d} \left(\overline{F_i}(tx_i - a_i) - \overline{F_i}(tx_i)\right).$$

Hence for sufficiently large t,

$$\begin{aligned} 0 \le \frac{\overline{F}(t\mathbf{x}-\mathbf{a}) - \overline{F}(t\mathbf{x})}{\overline{F}(t\mathbf{x})} &\le \sum_{i=1}^{d} \frac{\left(\overline{F_i}(tx_i - a_i) - \overline{F_i}(tx_i)\right)}{\overline{F}(t\mathbf{x})} \\ &\le \sum_{i=1}^{d} \frac{\left(\overline{F_i}(tx_i - a_i) - \overline{F_i}(tx_i)\right)}{\overline{F_i}(tx_i)} \\ &< d\epsilon. \end{aligned}$$

Letting $\epsilon \to 0$ gives the desired result. □

Remark 3.3 It is worth noting that the above statement and Corollary 11 in Omey (2006) show that for any probability distribution F supported by the positive quadrant in $\mathbb{R}^d$, such that the marginal distribution F_i of F is subexponential for every $i = 1, \ldots, d$, we have, for all $\mathbf{a} \in [0, \infty)^d$, $\mathbf{x} \in (0, \infty)^d$ and $n \geq 1$,

$$\lim_{t \to \infty} \frac{\overline{F^{*n}}(t\mathbf{x} - \mathbf{a})}{\overline{F}(t\mathbf{x})} = n. \tag{3.5}$$

We also mention that, in the multivariate case, it is possible to obtain in Eq. 3.3 the limit equal to 2 (i.e. corresponding to the subsexponential behaviour) for some directions $\mathbf{x}$ and a limit different from 2 in other directions. This was explored in Omey et al. (2006).

Given Proposition 3.2, using Eq. 3.3 as a definition of multivariate subexponentiality is, therefore, equivalent to merely requiring one-dimensional subexponentiality for each marginal distribution. Such requirement, in particular, cannot guarantee one-dimensional subexponentiality of the linear combinations, as we have seen in Example 3.1. In fact, it was shown in Leslie (1989) that even the sum of independent random variables with subexponential distributions does not need to have a subexponential distribution.

4 Multivariate subexponential distributions

In this section we introduce a new notion of a multivariate subexponential distribution. We approach the task with the multivariate ruin problem in mind. We start with a family $\mathcal{R}$ of open sets in $\mathbb{R}^d$. Recall that a subset A of $\mathbb{R}^d$ is increasing if $\mathbf{x} \in A$ and $\mathbf{a} \in [0, \infty)^d$ imply $\mathbf{x} + \mathbf{a} \in A$. Let

$$\mathcal{R} = \{A \subset \mathbb{R}^d : A \text{ open, increasing, } A^c \text{ convex, } \mathbf{0} \notin \overline{A}\}. \tag{4.1}$$

Remark 4.1 Note that $\mathcal{R}$ is a cone with respect to the multiplication by positive scalars. That is, if $A \in \mathcal{R}$, then $uA \in \mathcal{R}$ for any $u > 0$. Further, half-spaces of the form

$$H = \{\mathbf{x} : a_1 x_1 + \cdots + a_d x_d > b\}, \; b > 0, \; a_1, \ldots, a_d \geq 0 \text{ with } a_1 + \ldots + a_d = 1 \tag{4.2}$$

are members of $\mathcal{R}$.

Remark 4.2 We can write a set $A \in \mathcal{R}$ (in a non-unique way) as $A = \mathbf{b} + G$, with $\mathbf{b} \in (0, \infty)^d$ and $\mathbf{0} \in \partial G$ (with ∂G being the boundary of G). It is clear that the set G is then also increasing. We will adopt this notation in some of the proofs to follow.

To see a connection with the multivariate ruin problem, imagine that for a fixed set $A \in \mathcal{R}$ we view A as the "ruin set" in the sense that if, at any time, the excess of claim amounts over the premia falls in A, then the insurance company is ruined. Note that, in the one-dimensional situation, all sets in $\mathcal{R}$ are of the form $A = (u, \infty)$ with $u > 0$, so the ruin corresponds to the excess of claim amounts over the premia being over the initial capital u. The different shapes of sets in $\mathcal{R}$ can be viewed as allowing different interactions between multiple lines of business. For example, choosing A of the form

$$A = \{\mathbf{x} : x_i > u_i \text{ for some } i = 1, \ldots, d\}, \quad u_1, \ldots, u_d > 0 \tag{4.3}$$

corresponds to completely separate lines of business, where a ruin of one line of business causes the ruin of the company. On the other hand, using as A a half-space of the form (4.2) corresponds to the situation where there is a single overall initial capital b and the proportion of a_i in a shortfall in the ith line of business is charged to the overall capital b. The connections to the ruin problem are discussed more thoroughly in Section 5.

Before we introduce our notion of multivariate subexponentiality, we collect, in the following lemma, certain facts about the family $\mathcal{R}$. Note that part (d) is a general property of convex sets.

Lemma 4.3 *Let $A \in \mathcal{R}$.*

(a) *If $G = A - \mathbf{b}$ for some $\boldsymbol{b} \in \partial A$, then $G^c \supset (-\infty, 0]^d$.*
(b) *If $u_1 > u_2 > 0$ then $u_1 A \subset u_2 A$.*
(c) *There is a set of vectors $I_A \subset \mathbb{R}^d$ such that*

$$A = \left\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{p}^T \boldsymbol{x} > 1 \text{ for some } \boldsymbol{p} \in I_A\right\}.$$

(d) *For any $\mathbf{a}$ there is $u_1 > 0$ such that for all $u > u_1$ we have $(u + u_1)A \subset uA + \mathbf{a} \subset (u - u_1)A$.*
(e) *Let C be any convex set, and $u_1, u_2 > 0$, then $u_1 C + u_2 C = (u_1 + u_2)C$.*

Proof (a) Since G^c is closed, it contains the origin. Since G is increasing, G^c contains the entire quadrant $(-\infty, 0]^d$.

(b) This is an immediate consequence of the fact that A^c is convex and $\mathbf{0} \in A^c$.

(c) Let $\mathbf{x}_0 \in \partial A$. Since A^c is convex, the supporting hyperplane theorem (see e.g. Corollary 11.6.2 in Rockafellar (2015)) tells us that there exists a (not necessarily unique) nonzero vector $\mathbf{p}_{\mathbf{x}_0}$ such that $\mathbf{p}_{\mathbf{x}_0}^T \mathbf{x} \le \mathbf{p}_{\mathbf{x}_0}^T \mathbf{x}_0$ for all $\mathbf{x} \in A^c$. Since $\mathbf{0} \in A^c$, we must have $\mathbf{p}_{\mathbf{x}_0}^T \mathbf{x}_0 \ge 0$. Since A is increasing, the case $\mathbf{p}_{\mathbf{x}_0}^T \mathbf{x}_0 = 0$ is impossible, so $\mathbf{p}_{\mathbf{x}_0}^T \mathbf{x}_0 > 0$.

We scale each $\mathbf{p}_{\mathbf{x}_0}$ so that $\mathbf{p}_{\mathbf{x}_0}^T \mathbf{x}_0 = 1$. Let I_A be the set of all such $\mathbf{p}_{\mathbf{x}_0}$ for all $\mathbf{x}_0 \in \partial A$. Since a closed convex set equals the intersection of the half-spaces bounded by its supporting hyperplanes (see e.g. Corollary 11.5.1 in Rockafellar (2015)), the collection I_A has the required properties.

(d) Since $\mathbf{0} \notin \overline{A}$, the origin is an interior point of A^c. Therefore, we can choose $u_1 > 0$ so large that both $\mathbf{a}/u_1 \in A^c$ and $-\mathbf{a}/u_1 \in A^c$. In order to prove the leftmost inclusion in the statement, we have to show that for any $\mathbf{x} \in A$ and $u > u_1$ we have

$$\frac{u + u_1}{u}\mathbf{x} - \frac{1}{u}\mathbf{a} \in A.$$

Suppose that, to the contrary, the vector above is in A^c. Then, by the convexity of A^c,

$$\mathbf{x} = \frac{u}{u+u_1}\left(\frac{u+u_1}{u}\mathbf{x} - \frac{1}{u}\mathbf{a}\right) + \frac{u_1}{u+u_1}\frac{1}{u_1}\mathbf{a} \in A^c\,,$$

contradicting the assumption that $\mathbf{x} \in A$. Similarly, in order to prove the right-most inclusion in the statement, we have to show that for any $\mathbf{x} \in A$ and $u > u_1$ we have

$$\frac{u}{u-u_1}\mathbf{x} + \frac{1}{u-u_1}\mathbf{a} \in A\,.$$

Suppose that, to the contrary, the vector above is in A^c. Then, by the convexity of A^c,

$$\mathbf{x} = \frac{u-u_1}{u}\left(\frac{u}{u-u_1}\mathbf{x} + \frac{1}{u-u_1}\mathbf{a}\right) + \frac{u_1}{u}\left(-\frac{1}{u_1}\mathbf{a}\right) \in A^c\,,$$

once again contradicting the assumption that $\mathbf{x} \in A$.

(e) Let $\mathbf{x} \in u_1C$ and $\mathbf{y} \in u_2C$, then $\frac{\mathbf{x}}{u_1}, \frac{\mathbf{y}}{u_2} \in C$, and by convexity $\frac{u_1}{u_1+u_2}\frac{\mathbf{x}}{u_1} + \frac{u_2}{u_1+u_2}\frac{\mathbf{y}}{u_2} = \frac{1}{u_1+u_2}(\mathbf{x}+\mathbf{y}) \in C$, so $\mathbf{x}+\mathbf{y} \in (u_1+u_2)C$, implying that $u_1C+u_2C \subset (u_1+u_2)C$. The other direction is obvious.

$\square$

Remark 4.4 Note that the set I_A is not uniquely determined. If, for example, A is the half-space given in Eq. 4.2, then we can use, as I_A, the singleton $\{(a_1/b, \ldots, a_d/b)\}$. If A is the complement of "a corner" as in Eq. 4.3, then one can use as I_A a set with d vectors, the ith vector having the form $\mathbf{e}_i/u_i$, $i = 1, \ldots, d$, where $\mathbf{e_i}$ is the corresponding standard coordinate vector. Note also that, once we have chosen a collection I_A for some $A \in \mathcal{R}$, for any $u > 0$ we can use I_A/u as I_{uA}.

Let F be a probability distribution on $\mathbb{R}^d$ supported by $[0, \infty)^d$. For a fixed $A \in \mathcal{R}$ it follows from part (b) of Lemma 4.3 that the function on $[0, \infty)$ defined by

$$F_A(t) = 1 - F(tA),\ t \geq 0\,,$$

is a probability distribution function on $[0, \infty)$. We will reserve in the sequel the notation Y_A for a generic random variable with the distribution F_A. The following lemma is elementary.

Lemma 4.5 *Suppose that $\boldsymbol{X}$ has distribution F. Then the random variable $Y_A = \sup\{u : \boldsymbol{X} \in uA\}$ has law F_A, and $P(Y_A > t) = P(\boldsymbol{X} \in tA) = P\big(\sup_{\mathbf{p}\in I_A} \mathbf{p}^T\boldsymbol{X} > t\big)$ for $t > 0$.*

Note that the supremum in the definition of Y_A is taken over $u > 0$; for simplicity we omit an explicit statement of that here and in the sequel.

We are now ready to define multivariate subexponentiality.

Definition 4.6 For any $A \in \mathcal{R}$, we say that $F \in \mathscr{S}_A$ if $F_A \in \mathscr{S}$, and we write $\mathscr{S}_{\mathcal{R}} := \cap_{A \in \mathcal{R}} \mathscr{S}_A$.

We view the class $\mathscr{S}_{\mathcal{R}}$ as the class of subexponential distributions. However, for some applications we can use a larger class, such as $\mathscr{S}_A$ for a fixed $A \in \mathcal{R}$, or the intersection of such classes over a subset of $\mathcal{R}$.

Note that by Remark 4.1, if $\mathbf{X}$ is a random vector in $\mathbb{R}^d$ whose distribution is in $\mathscr{S}_{\mathcal{R}}$, then all non-degenerate linear combinations of the components of $\mathbf{X}$ with nonnegative coefficients have one-dimensional subexponential distributions. More generally, we have the following stability property. We say that a linear transformation $T : \mathbb{R}^d \to \mathbb{R}^k$ is increasing if $T\mathbf{x} \in [0, \infty)^k$ for any $\mathbf{x} \in [0, \infty)^d$.

Proposition 4.7 *Let $T : \mathbb{R}^d \to \mathbb{R}^k$ be a linear increasing transformation. If $\mathbf{X}$ is a random vector in $\mathbb{R}^d$ whose distribution is in $\mathscr{S}_{\mathcal{R}}$ (in $\mathbb{R}^d$), then the same is true (in $\mathbb{R}^k$) for the distribution of the random vector $T\mathbf{X}$.*

Proof It suffices to show that for any $A \in \mathcal{R}$ in $\mathbb{R}^k$, the set $T^{-1}A$ is in $\mathcal{R}$ in $\mathbb{R}^d$.

We check that $T^{-1}A$ satisfies each of the conditions in Eq. 4.1. That $T^{-1}A$ is open follows from that fact that T is continuous and A is open. To show that $T^{-1}A$ is increasing, let $\mathbf{x} \in T^{-1}A$, and $\mathbf{a} \geq \mathbf{0}$. In this case $T\mathbf{x} = \mathbf{y}$ for some $\mathbf{y} \in A$, and $T\mathbf{a} = \mathbf{b}$ for some $\mathbf{b} \geq \mathbf{0}$. Hence by linearity and the fact that A is increasing, it follows that $T(\mathbf{x} + \mathbf{a}) = \mathbf{y} + \mathbf{b} \in A$. As for convexity of $(T^{-1}A)^c$, notice that $(T^{-1}A)^c = T^{-1}A^c$. For any $\mathbf{x}, \mathbf{y} \in (T^{-1}A)^c$, then, $T\mathbf{x}, T\mathbf{y} \in A^c$. For $\lambda \in (0, 1)$, $T(\lambda\mathbf{x} + (1 - \lambda)\mathbf{y}) = \lambda T\mathbf{x} + (1 - \lambda)T\mathbf{y} \in A^c$, and convexity of $(T^{-1}A)^c = T^{-1}A^c$ follows. Lastly, notice that $\overline{T^{-1}A} = T^{-1}\overline{A}$, so $\mathbf{0} \in \overline{T^{-1}A}$ leads to the contradiction that $\mathbf{0} \in \overline{A}$. □

Remark 4.8 Note that Proposition 4.7 also holds if instead of the entirety of $\mathscr{S}_{\mathcal{R}}$, we restrict the distributions to those in $\mathscr{S}_{\mathcal{R}_H}$, where $\mathcal{R}_H$ consists only of half-spaces of the form (4.2), and $\mathscr{S}_{\mathcal{R}_H} := \cap_{A \in \mathcal{R}_H} \mathscr{S}_A$. This follows easily from the above proof and the fact that $T^{-1}H$ is still a half-space for any half-space H.

The next lemma is useful and follows quite naturally from Lemma 4.5.

Lemma 4.9 *For any $A \in \mathcal{R}$ and $n \geq 1$,*

$$\overline{(F_A)^{*n}}(t) \geq F^{*n}(tA). \tag{4.4}$$

Proof Let $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ be independent random vectors with distribution F. Let $Y_1, \ldots, Y_n$ be one-dimensional random variables defined by

$$Y_i = \sup\{u : \mathbf{X}^{(i)} \in uA\} = \sup_{\mathbf{p} \in I_A} \mathbf{p}^T \mathbf{X}^{(i)},\ i = 1, \ldots, d\,.$$

By Lemma 4.5, $P(Y_i > t) = \overline{F_A}(t)$. Hence it follows that

$$\begin{aligned} F^{*n}(tA) &= P(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)} \in tA) \\ &= P\big(\sup_{\mathbf{p}\in I_A} \mathbf{p}^T(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)}) > t\big) \\ &\leq P\big(\sup_{\mathbf{p}\in I_A} \mathbf{p}^T\mathbf{X}^{(1)} + \cdots + \sup_{\mathbf{p}\in I_A} \mathbf{p}^T\mathbf{X}^{(n)} > t\big) \\ &= P(Y_1 + \ldots + Y_n > t) \\ &= \overline{(F_A)^{*n}}(t)\,, \end{aligned}$$

as required. □

In spite of this lemma, the two probabilities in Eq. 4.4 are asymptotically equivalent.

Corollary 4.10 *$A \in \mathcal{R}$. Let $\mathbf{X}, \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ be independent random vectors with distribution F. If $F \in \mathscr{S}_A$ for some $A \in \mathcal{R}$, then for all $n \geq 1$,*

$$\lim_{u\to\infty} \frac{P(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)} \in uA)}{P(\mathbf{X} \in uA)} = n. \tag{4.5}$$

Proof It follows from Lemma 4.9 and Eq. 2.4 that only an asymptotic lower bound needs to be established. However, since $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ are all nonnegative, and A is an increasing set, it must be that if $\mathbf{X}^{(1)}+\cdots+\mathbf{X}^{(n)} \in uA^c$, then each $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)} \in uA^c$. Therefore,

$$\begin{aligned} P(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)} \in uA^c) &\leq P(\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)} \in uA^c) \\ &= P(\mathbf{X} \in uA^c)^n. \end{aligned}$$

It follows that

$$\liminf_{u\to\infty} \frac{P(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)} \in uA)}{P(\mathbf{X} \in uA)} \geq \liminf_{u\to\infty} \frac{1 - P(\mathbf{X} \in uA^c)^n}{P(\mathbf{X} \in uA)} = n\,,$$

as required. □

Remark 4.11 We note at this point that the assumption $F \in \mathscr{S}_A$ is NOT equivalent to the assumption that Eq. 4.5 holds for all n. In fact, the latter assumption is weaker. To see that, consider the following example. Let X and Y be two independent nonnegative one-dimensional random variables with subexponential distributions, such that $X + Y$ is not subexponential; recall that such random variables exist, see Leslie (1989). We construct a bivariate random vector $\mathbf{Z}$ by taking a Bernoulli $(1/2)$ random variable B independent of X and Y and setting $\mathbf{Z} = (X, 0)$ if $B = 0$ and $\mathbf{Z} = (0, Y)$ if $B = 1$. Let $A = \{(x, y) : \max(x, y) > 1\}$. Since the marginal distributions of the bivariate distribution of $\mathbf{Z}$ are, obviously, subexponential, we see by Eq. 3.5 that Eq. 4.5 holds for all $n \geq 1$. However, for $u > 0$,

$$\overline{F_A}(u) = \frac{1}{2}P(X > u) + \frac{1}{2}P(Y > u)\,,$$

so the distribution F_A is a mixture of the distributions of X and Y. By Theorem 2 of Embrechts and Goldie (1980), any non-trivial mixture of the distributions of X and Y is subexponential if and only if their convolution is. Since, by construction, that convolution is not subexponential, we conclude that $F_A \notin \mathscr{S}$ and $F \notin \mathscr{S}_A$.

In the next proposition we check that the basic properties of one-dimensional subexponential distributions extend to the multivariate case.

Proposition 4.12 *Let $A \in \mathcal{R}$ and $F \in \mathscr{S}_A$.*

(a) *If G is a distribution on $\mathbb{R}^d$ supported by $[0, \infty)^d$, such that*

$$\lim_{u\to\infty} \frac{F(uA)}{G(uA)} = c > 0,$$

then $G \in \mathscr{S}_A$.

(b) *For any $\mathbf{a} \in \mathbb{R}^d$,*

$$\lim_{u\to\infty} \frac{F(uA + \mathbf{a})}{F(uA)} = 1. \tag{4.6}$$

(c) *Let $\mathbf{X}, \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ be independent random vectors with distribution F. For any $\epsilon > 0$, there exists $K > 0$ such that for all $u > 0$ and $n \geq 1$,*

$$\frac{P(\mathbf{X}^{(1)} + \cdots + \mathbf{X}^{(n)} \in uA)}{P(\mathbf{X} \in uA)} < K(1+\epsilon)^n . \tag{4.7}$$

Proof (a) This is an immediate consequence of the univariate subexponentiality of F_A and the corresponding property of one-dimensional subexponential distributions; see e.g. Lemma 4 in Embrechts et al. (1979).

(b) By part (d) of Lemma 4.3, there exists $u_1 > 0$ such that for all $u > u_1$ we have $(u+u_1)A \subset uA + \mathbf{a} \subset (u-u_1)A$. Therefore,

$$\begin{aligned}\overline{F_A}(u+u_1) &= F((u+u_1)A)\\ &\leq F(uA+\mathbf{a})\\ &\leq F((u-u_1)A) = \overline{F_A}(u-u_1)\,,\end{aligned}$$

and the claim follows from the one-dimensional long tail property of F_A.

(c) The claim follows from Lemma 4.9 and the corresponding one-dimensional bound; see e.g. Lemma 3 in Embrechts et al. (1979).

□

Remark 4.13 In our Definition 4.6 of multivariate subexponentiality one can drop the assumption that a distribution is supported by $[0, \infty)^d$. We can check that both Corollary 4.10 and Proposition 4.12 remain true in this extended case.

Our next step is to show that multivariate regular varying distributions fall within the class $\mathscr{S}_\mathcal{R}$ of multivariate subexponential distributions. The definition of non-standard multivariate regular variation for distributions supported by $[0, \infty)^d$ was given in Eq. 3.2. Presently we would only consider the standard multivariate regular variation, but allow distributions not necessarily restricted to the first quadrant. In this case one assumes that there is a non-zero Radon measure μ on $[-\infty, \infty]^d \setminus \{\mathbf{0}\}$, charging only finite points, and a function b on $(0, \infty)$ increasing to infinity, such that

$$tF\big(b(t)\cdot\big) \xrightarrow{v} \mu \tag{4.8}$$

vaguely on $[-\infty, \infty]^d \setminus \{\mathbf{0}\}$. Recall that the measure μ is called the tail measure of $\mathbf{X}$; it has automatically a scaling property: for some $\alpha > 0$, $\mu(uA) = u^{-\alpha}\mu(A)$ for every $u > 0$ and every Borel set $A \in \mathbb{R}^d$, and the function b in Eq. 4.8 is regularly varying with exponent $1/\alpha$; see Resnick (2007). We say that F (and $\mathbf{X}$) are regularly varying with exponent α and use the notation $F \in MRV(\alpha, \mu)$.

Proposition 4.14 $MRV(\alpha, \mu) \subset \mathscr{S}_{\mathcal{R}}$.

Proof We start by showing that for any $A \in \mathcal{R}$, $\mu(\partial A) = 0$. Since for any $u > 0$,

$$\mu(\partial(uA)) = \mu(u\partial A) = u^{-\alpha}\mu(\partial A)\,,$$

it is enough to show that for any $u > 1$, $\partial(uA) \cap \partial A = \emptyset$ (indeed, $\mu(\partial A) > 0$ would then imply existence of uncountably many disjoint sets of positive measure).

Suppose, to the contrary, that $\partial(uA) \cap \partial A \neq \emptyset$, and let $\mathbf{x} \in \partial(uA) \cap \partial A$. The set I_A in part (c) of Lemma 4.3, has, by construction, the property that $u^{-1}\mathbf{x}$, as an element of $u^{-1}\partial(uA) = \partial A$, satisfies $\mathbf{p}^T u^{-1}\mathbf{x} = 1$ for some $\mathbf{p} \in I_A$. But then $\mathbf{p}^T\mathbf{x} = u > 1$, which says that $\mathbf{x}$ is in A, rather than in ∂A, which is a subset of A^c.

It follows from Eq. 4.8 that for any set $A \in \mathcal{R}$,

$$tP\big(\mathbf{X} \in b(t)A\big) \to \mu(A) \in (0, \infty)$$

as $t \to \infty$. Since the function b is regularly varying with exponent $1/\alpha$, we immediately conclude that the distribution function F_A has a regularly varying tail, hence F_A is subexponential. Because $A \in \mathcal{R}$ is arbitrary, it follows that $F \in \cap_{A\in\mathcal{R}}\mathscr{S}_A = \mathscr{S}_{\mathcal{R}}$. □

We proceed with clarifying the relation between the class $\mathscr{S}_{\mathcal{R}}$ we have introduced in this section and the classes $\mathscr{S}(\nu; \mathbf{b})$ and $S(\mathbb{R}^d)$ of Section 3. We will also provide several examples of distributions that belong to $\mathscr{S}_{\mathcal{R}}$, as well as sufficient conditions for a distribution to be a member of $\mathscr{S}_{\mathcal{R}}$.

Example 3.1, combined with Proposition 4.7, show that neither $\mathscr{S}(\nu; \mathbf{b})$ nor $S(\mathbb{R}^d)$ are subsets of $\mathscr{S}_{\mathcal{R}}$. We will present an example to show that $\mathscr{S}_{\mathcal{R}} \not\subset \mathscr{S}(\nu; \mathbf{b})$.

We start with presenting a sufficient condition for a distribution F to be a member of $\mathscr{S}_{\mathcal{R}}$. We assume for the moment that F is supported by $[0, \infty)^d$.

Let $\mathbf{X} \sim F$ be a nonnegative random vector on $\mathbb{R}^d$ such that $P(\mathbf{X} = \mathbf{0}) = 0$. Denote the L_1 norm (we could have chosen any other norm as well) of $\mathbf{X}$ by

$$W = ||\mathbf{X}||_1 = \sum_{i=1}^{d} X_i\,, \tag{4.9}$$

and the projection of $\mathbf{X}$ onto the d-dimensional unit simplex Δ_d by

$$I = \frac{\mathbf{X}}{||\mathbf{X}||_1} = \frac{\mathbf{X}}{W} \in \Delta_d \,. \tag{4.10}$$

Let ν be the distribution of I over Δ_d, and let $(F_{\boldsymbol{\theta}})_{\boldsymbol{\theta}\in\Delta_d}$ be a set of regular conditional distributions of W given I. Notice that, if the law F of $\mathbf{X}$ in $\mathbb{R}^d$ has a density f with respect to the d-dimensional Lebesgue measure, then a version of $(F_{\boldsymbol{\theta}})_{\boldsymbol{\theta}\in\Delta_d}$ has densities with respect to the one-dimensional Lebesgue measure, given by

$$f_{\boldsymbol{\theta}}(w) = \frac{w^{d-1} f(w\boldsymbol{\theta})}{\int_0^\infty u^{d-1} f(u\boldsymbol{\theta})\, du}, \; w > 0. \tag{4.11}$$

Proposition 4.15 *Suppose* $\mathbf{X}$ *is a random vector on* $\mathbb{R}^d$ *with distribution* F*, supported by* $[0,\infty)^d$*, such that* $P(\mathbf{X} = \mathbf{0}) = 0$*. Suppose the marginal distributions* F_i*,* $i = 1,\ldots,d$ *have dominated varying tails. Further, assume that there is a set of regular conditional distributions* $(F_{\boldsymbol{\theta}})_{\boldsymbol{\theta}\in\Delta_d}$ *of* W *given* I *such that* $F_{\boldsymbol{\theta}} \in \mathscr{L}$ *for each* $\boldsymbol{\theta} \in \Delta_d$ *and for some* $C, t_0 > 0$*,*

$$\frac{\overline{F_{\boldsymbol{\theta}_1}}(2t)}{\overline{F_{\boldsymbol{\theta}_2}}(t)} \le C \tag{4.12}$$

for all $t > t_0$ *and for all* $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \Delta_d$*. Then* $F \in \mathscr{S}_{\mathcal{R}}$*.*

Proof Let $A \in \mathcal{R}$ be fixed. We first check that if each of the marginal distributions have dominated varying tails, then F_A also has a dominated varying tail. Let

$$I = \left\{i = 1,\ldots,d : \; A \cap \{\mathbf{x} : \; x_i \ge 0, \; x_j = 0, \; j \ne i\} \ne \emptyset\right\}.$$

Note that I cannot be empty, since otherwise all positive half-axes would belong to A^c and the latter, by convexity, would cover the entire nonnegative quadrant, contradicting the assumption that A is non-empty. For $i \in I$ we choose $a_i > 0$ such that $(0,\ldots,0,a_i,0,\ldots,0) \in A$. Then $F(A) \ge \overline{F}_i(a_i)$ for each $i \in I$. Furthermore, we claim that

$$A' := A \cap \{\mathbf{x} : \; x_i = 0 \;\text{ for all } i \in I\} = \emptyset\,.$$

Indeed, A', viewed as a set in a smaller-dimensional space, has the same properties as A, and for A' the corresponding set I (again, in a smaller-dimensional space) is empty. As explained above, this forces A' to be empty. That is,

$$\{\mathbf{x} : \; x_i = 0 \;\text{ for all } i \in I\} \subset A^c\,.$$

Since $\mathbf{0}$ is an interior point of A^c, there is $\varepsilon > 0$ such that

$$\{\mathbf{x} : \; x_i \le \varepsilon, \; i \in I, \; x_j = 0, \; j \notin I\} \in A^c\,.$$

By convexity of A^c we deduce that

$$\{\mathbf{x} : \; x_i \le \varepsilon/2 \;\text{ for all } i \in I\} \subset A^c\,.$$

That is, $F(A) \le \sum_{i\in I} \overline{F}_i(\varepsilon/2)$. Therefore,

$$\liminf_{t\to\infty} \frac{\overline{F_A}(2t)}{\overline{F_A}(t)} = \liminf_{t\to\infty} \frac{F(2tA)}{F(tA)} \ge \liminf_{t\to\infty} \frac{\frac{1}{d}\sum_{i\in I} \overline{F_i}(2ta_i)}{\sum_{i\in I} \overline{F_i}(t\varepsilon/2)} > 0,$$

since each F_i has a dominated varying tail.

Since $\mathscr{L} \cap \mathscr{D} \subset \mathscr{S}$, it now suffices to show that $F_A \in \mathscr{L}$. For $\boldsymbol{\theta} \in \Delta_d$, let

$$h_{\boldsymbol{\theta}} = \inf\{w > 0 : w\boldsymbol{\theta} \in A\} > 0\,, \tag{4.13}$$

Note that $h_{\boldsymbol{\theta}}$ is bounded away from 0. Further, we have already proved that $h(\mathbf{e}^{(i)}) < \infty$ for at least one coordinate vector $\mathbf{e}^{(i)}$, $i = 1, \ldots, d$. Since the dominated variation of the marginal tails implies, in particular, that each coordinate of the vector $\mathbf{X}$ is positive with positive probability, we conclude that

$$\nu\big\{\boldsymbol{\theta} \in \Delta_d : h_{\boldsymbol{\theta}} < \infty\big\} > 0\,.$$

We conclude that there is $M > 0$ and a measurable set $B \subset \Delta_d$ with $\delta := \nu(B) > 0$, such that

$$1/M \le h_{\boldsymbol{\theta}} \le M \ \text{ for all } \boldsymbol{\theta} \in B.$$

Note that for $t > 0$,

$$\overline{F}_A(t) = \int_{\Delta_d} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})\,\nu(d\boldsymbol{\theta}). \tag{4.14}$$

Therefore,

$$\frac{\overline{F_A(t)} - \overline{F_A(t+1)}}{\overline{F_A}(t)} = \frac{\int_{\Delta_d} (\overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}}) - \overline{F_{\boldsymbol{\theta}}}((t+1)h_{\boldsymbol{\theta}}))\,\nu(d\boldsymbol{\theta})}{\int_{\Delta_d} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})\,\nu(d\boldsymbol{\theta})}\,, \tag{4.15}$$

and we wish to show that this quantity goes to 0 as $t \to \infty$.

By the assumptions, for any fixed $\boldsymbol{\theta}$, $F_{\boldsymbol{\theta}} \in \mathscr{L}$, hence for any fixed $\boldsymbol{\theta}$ such that $h_{\boldsymbol{\theta}} < \infty$,

$$\lim_{t\to\infty} \frac{\overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}}) - \overline{F_{\boldsymbol{\theta}}}((t+1)h_{\boldsymbol{\theta}})}{\overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})} = 0.$$

Therefore, for a given $\epsilon > 0$, there exists $t_\epsilon > 0$ such that, for all $t > t_\epsilon$, $\nu(S_{t,\epsilon}) < \epsilon$, where

$$S_{t,\epsilon} = \left\{\boldsymbol{\theta} \in \Delta_d : h_{\boldsymbol{\theta}} < \infty \text{ and } \frac{\overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}}) - \overline{F_{\boldsymbol{\theta}}}((t+1)h_{\boldsymbol{\theta}})}{\overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})} > \epsilon\right\}.$$

Let $\epsilon < (\delta/2)^2$. Then $\nu\big(B \cap S^c_{t,\epsilon}\big) > \big(\nu(S_{t,\epsilon})\big)^{1/2}$. By the definition of the set B and by Eq. 4.12, and recalling that $h_{\boldsymbol{\theta}}$ is bounded away from 0, we know that for some $C_1, \tilde{t}_0 > 0$

$$\overline{F_{\boldsymbol{\theta}_1}}(th_{\boldsymbol{\theta}_1}) \le C_1 \overline{F_{\boldsymbol{\theta}_2}}(th_{\boldsymbol{\theta}_2})$$

for any $\boldsymbol{\theta}_1 \in S_{t,\epsilon}$ and any $\boldsymbol{\theta}_2 \in B \cap S_{t,\epsilon}^c$, for all $t > \tilde{t}_0$. Therefore, for $t > t_\epsilon + \tilde{t}_0$,

$$\begin{aligned}\int_{S_{t,\epsilon}} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}}) - \overline{F_{\boldsymbol{\theta}}}((t+1)h_{\boldsymbol{\theta}})\, \nu(d\boldsymbol{\theta}) &\leq \int_{S_{t,\epsilon}} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})\, \nu(d\boldsymbol{\theta}) \\ &< \frac{\nu(S_{t,\epsilon})}{\nu(B \cap S_{t,\epsilon}^c)} C_1 \int_{B \cap S_{t,\epsilon}^c} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})\, \nu(d\boldsymbol{\theta}) \\ &< \epsilon^{1/2} C_1 \int_{\Delta_d} \overline{F_{\boldsymbol{\theta}}}(th_{\boldsymbol{\theta}})\, \nu(d\boldsymbol{\theta}).\end{aligned}$$

Hence, for $t > t_\epsilon + \tilde{t}_0$, the quantity in Eq. 4.15 is bounded above by $\epsilon + \epsilon^{1/2} C_1$. Letting $\epsilon \searrow 0$ gives us the desired result. $\square$

We are now ready to give an example showing that $\mathscr{S}_{\mathcal{R}} \not\subset \mathscr{S}(\nu; \mathbf{b})$.

Example 4.16 Let $0 < |\gamma| \leq 1/12$. It is shown in Cline and Resnick (1992) that a legitimate probability distribution F, supported by $[0, \infty)^2$, satisfies

$$P(X > x, Y > y) = \frac{1 + \gamma \sin(\log(1 + x + y)) \cos(\frac{1}{2}\pi \frac{x-y}{1+x+y})}{1 + x + y}, \; x, y \geq 0\,. \quad (4.16)$$

Then

$$P(X > x) = P(Y > x) \sim x^{-1} \;\text{ as } x \to \infty\,,$$

but $F \notin \mathscr{S}(\nu; \mathbf{b})$ for any ν and $\mathbf{b}$; see Cline and Resnick (1992). Straightforward differentiation gives us the density f of F, and one can check that it satisfies

$$\frac{2 - 4\gamma - 3\gamma\pi - \pi^2/4}{(1 + x + y)^3} \leq f(x, y) \leq \frac{2 + 4\gamma + 3\gamma\pi + \pi^2/4}{(1 + x + y)^3}\,,$$

so by Eq. 4.11, we have

$$a\frac{w}{(1+w)^3} \leq f_{\boldsymbol{\theta}}(w) \leq b\frac{w}{(1+w)^3}, \; w > 0,$$

for some $0 < a < b < \infty$, independent of $\boldsymbol{\theta}$. It is clear that the conditions of Proposition 4.15 are satisfied and, hence, $F \in \mathscr{S}_{\mathcal{R}}$.

Proposition 4.15 gives us a way to check that a multivariate distribution belongs to the class $\mathscr{S}_{\mathcal{R}}$, but it only applies to distributions that have, marginally, dominated varying tails. In the remainder of this section we provide sufficient conditions for membership in $\mathscr{S}_{\mathcal{R}}$ that do not require marginals with dominated varying tails. We start with a motivating example.

Example 4.17 [Rotationally invariant case] Assume that there is a one-dimensional distribution G such that $\overline{F_{\boldsymbol{\theta}}} = \overline{G}$ for all $\boldsymbol{\theta} \in \Delta_d$. Let $A \in \mathcal{R}$, and notice that, in the rotationally invariant case, a random variable Y_A with distribution F_A can be written, in law, as

$$Y_A \stackrel{d}{=} Z H^{-1}, \quad (4.17)$$

with Z and H being independent, Z with the distribution G, and $H = h_\Theta$. Here h is defined by Eq. 4.13, and Θ has the law ν over the simplex Δ_d. Recall that the function h is bounded away from zero, so that the random variable H^{-1} is bounded. If $G \in \mathscr{S}$, then the product in the right hand side is subexponential by Corollary 2.5 in Cline and Samorodnitsky (1994). Hence $F_A \in \mathscr{S}$ for all $A \in \mathcal{R}$, and so $F \in \mathscr{S}_{\mathcal{R}}$.

The rotationally invariant case of Example 4.17 can be slightly extended, without much effort, to the case where there is a bounded positive function $\big(a_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \Delta_d\big)$ such that $F_{\boldsymbol{\theta}}(\cdot) = G(\cdot/a_{\boldsymbol{\theta}})$ for some $G \in \mathscr{S}$. An argument similar to the one in the example shows that we can still conclude that $F \in \mathscr{S}_{\mathcal{R}}$. In order to achieve more than that, we note that the distribution F_A can be represented, by Eq. 4.14, as a mixture of scaled regular conditional distributions. Note also that the product of independent random variables in Eq. 4.17 is just a special case of that mixture, to which we have been able to apply Corollary 2.5 in Cline and Samorodnitsky (1994). It is likely to be possible to extend that result to certain mixtures that are more general than products of independent random variables, and thus to obtain additional criteria for membership in the class $\mathscr{S}_{\mathcal{R}}$. We leave serious extensions of this type to future work. A small extension that still steps away from exact products is below, and it takes a result in Cline and Samorodnitsky (1994) as an ingredient. We formulate the statement in terms of the distribution of a random variable that only in a certain asymptotic sense looks like a product of independent random variables.

Theorem 4.18 *Let $(\Omega_i, \mathcal{F}_i, P_i)$, $i = 1, 2$ be probability spaces. Let Q be a random variable defined on the product probability space. Assume that there are nonnegative random variables X_i, $i = 1, 2$, defined on $(\Omega_1, \mathcal{F}_1, P_1)$ and $(\Omega_2, \mathcal{F}_2, P_2)$ correspondingly, such that X_1 has a subexponential distribution F, and for some $t_0 > 0$ and $C > 0$,*

$$X_1(\omega_1)X_2(\omega_2) - CX_2(\omega_2) \le Q(\omega_1, \omega_2) \le X_1(\omega_1)X_2(\omega_2) + CX_2(\omega_2) \tag{4.18}$$

a.s. on the set $\{Q(\omega_1, \omega_2) > t_0\}$. Suppose $P(X_2 > 0) > 0$, and let G be the distribution of X_2. Suppose that there is a function $a : (0, \infty) \to (0, \infty)$, such that

(1.) $a(t) \nearrow \infty$ *as* $t \to \infty$;
(2.) $\frac{t}{a(t)} \nearrow \infty$ *as* $t \to \infty$;
(3.) $\lim_{t\to\infty} \frac{\overline{F}(t-a(t))}{\overline{F}(t)} = 1$;
(4.) $\lim_{t\to\infty} \frac{\overline{G}(a(t))}{P(X_1X_2>t)} = 0$.

Then Q has a subexponential distribution.

Proof Let H denote the distribution of X_1X_2. It follows by Theorem 2.1 in Cline and Samorodnitsky (1994) that H is subexponential. We show that $P(Q > t) \sim \overline{H}(t)$ as $t \to \infty$. This will imply that Q has a subexponential distribution.

We start by checking that

$$\lim_{t\to\infty} \frac{\overline{H}(t - a(t))}{\overline{H}(t)} = 1, \quad \text{which will imply that } \lim_{t\to\infty} \frac{\overline{H}(t + a(t))}{\overline{H}(t)} = 1\,, \tag{4.19}$$

because by the first limit and monotonicity of a,

$$1 \le \liminf_{t\to\infty} \frac{\overline{H}(t)}{\overline{H}(t+a(t))} \le \limsup_{t\to\infty} \frac{\overline{H}(t)}{\overline{H}(t+a(t))} \le \limsup_{t\to\infty} \frac{\overline{H}(t+a(t)-a(t+a(t)))}{\overline{H}(t+a(t))} = 1\,.$$

To verify the limit, suppose first that $X_2 \ge 1$ a.s., and write

$$P\big(t-a(t) < X_1X_2 \le t\big) \le P_2(X_2 > a(t))$$
$$+ \int_{\Omega_2} P_1(t/X_2(\omega_2) - a(t)/X_2(\omega_2) < X_1 \le t/X_2(\omega_2))\mathbf{1}\big(X_2(\omega_2) \le a(t)\big)\, P_2(d\omega_2)$$

The first term in the right hand side is $o(\overline{H}(t))$ by the assumption (4), while the same is true for the second term by the assumption (3), since by the assumption (2), $a(t)/y \le a(t/y)$ if $y \ge 1$. This proves (4.19) if $X_2 \ge 1$ a.s. and hence, by scaling, if $X_2 \ge \epsilon$ a.s. for some $\epsilon > 0$. An elementary truncation argument then shows that Eq. 4.19 holds if $P(X_2 > 0) > 0$.

Note that for $t > t_0$,

$$\begin{aligned} P(Q > t) &\le P(X_1X_2 + CX_2 > t) \\ &\le \overline{G}(a(t)) + \overline{H}(t - Ca(t))\,. \end{aligned}$$

This implies that $\limsup_{t\to\infty} P(Q > t)/\overline{H}(t) \le 1$. The statement $\liminf_{t\to\infty} P(Q > t)/\overline{H}(t) \ge 1$ can be shown in a similar way. □

Despite a limited scope of the extension given in Theorem 4.18, it allows one to construct a number of examples of multivariate distributions in $\mathscr{S}_{\mathcal{R}}$ by choosing, for example, $\Omega_2 = \Delta_d$ and $X_2(\boldsymbol{\theta}) = 1/h(\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Delta_d$, and selecting a function Q to model additional randomness in the radial direction.

5 Ruin probabilities

As mentioned in the introduction, the notion of subexponentiality we introduced in Section 4 was designed with insurance applications in mind. In this section we describe such an application more explicitly.

Consider a renewal model for the reserves of an insurance company with d lines of business. Suppose that claims arrive according to a renewal process $(N_t)_{t\ge 0}$ given by $N_t = \sup\{n \ge 1 : T_n \le t\}$. The arrival times (T_n) form a renewal sequence

$$T_0 = 0, \quad T_n = Y_1 + \cdots + Y_n \text{ for } n \ge 1, \tag{5.1}$$

where the interarrival times $(Y_i)_{i\ge 1}$ form a sequence of independent and identically distributed positive random variables. We will call a generic interarrival time Y. At the arrival time T_i a random vector-valued claim size $\mathbf{X}^{(i)} = \big(X_1^{(i)}, \ldots, X_d^{(i)}\big)$ is incurred, so that the part of the claim going to the jth line of business is $X_j^{(i)}$. We assume that the claim sizes $(\mathbf{X}^{(i)})$ are i.i.d. random vectors with a finite mean, and we denote their common law by F. We assume further that the claim size process is independent of the renewal process of the claim arrivals. The jth line of business collects premium at the rate of p^j per unit of time. Let $\mathbf{p}$ be the vector of the premium rates, and $\mathbf{X}$ a generic random vector of claim sizes.

Suppose that the company has an initial buffer capital of u, out of which the amount of ub_j is allocated to the jth line of business, $j = 1, 2, \ldots, d$. Here $b_1, \ldots, b_d$ are positive numbers, $b_1 + \cdots + b_d = 1$. Then $u\mathbf{b}$ denotes the vector for the initial capital buffer allocation. With the above notation, the claim surplus process $(\mathbf{S}_t)_{t\geq 0}$ and the risk reserve process $(\mathbf{R}_t)_{t\geq 0}$ are given by

$$\mathbf{S}_t = \sum_{i=1}^{N_t} \mathbf{X}^{(i)} - t\mathbf{p}, \quad \mathbf{R}_t = u\mathbf{b} - \mathbf{S}_t = u\mathbf{b} + t\mathbf{p} - \sum_{i=1}^{N_t} \mathbf{X}^{(i)}, \; t \geq 0 .$$

The company becomes insolvent (ruined) when the risk reserve process hits a certain ruin set $L \subset \mathbb{R}^d$. Equivalently, ruin occurs when the claim surplus process enters the set $u\mathbf{b} - L$. We will assume that the ruin set satisfies the following condition.

Assumption 5.1 The ruin set L is an open decreasing (i.e., $-L$ is increasing) set such that $\mathbf{0} \in \partial L$, satisfying $L = uL$ for $u > 0$, and such that L^c is convex.

Note that this assumption means that the ruin occurs when the claim surplus process enters the set uA, with $A = \mathbf{b} - L \in \mathcal{R}$, as defined in Section 4. In fact, the ruin set L can be viewed as being of the form $-G$, as defined in Remark 4.2. Examples of such ruin sets are, of course, the sets

$$L = \left\{\mathbf{x} : \, x_j < 0 \text{ for some } j = 1, \ldots, d\right\} \text{ and } L = \left\{\mathbf{x} : \, x_1 + \ldots + x_d < 0\right\},$$

discussed in Section 4. A general framework was proposed in Hult and Lindskog (2006). In this framework capital can be transferred between different business lines, but the transfers incur costs, and the solvency set has the form

$$L^c = \left\{ \mathbf{x} : \, \mathbf{x} = \sum_{i \neq j} v_{ij}(\pi_{ij}\mathbf{e}^i - \mathbf{e}^j) + \sum_{i=1}^{d} w_i \mathbf{e}^i, \; v_{ij} \geq 0, \; w_i \geq 0 \right\}, \tag{5.2}$$

where $\mathbf{e}^1, \ldots, \mathbf{e}^d$ are the standard basis vectors, and $\Pi = (\pi_{ij})_{i,j=1}^d$ is a matrix satisfying

(i) $\pi_{ij} \geq 1$ for $i, j \in \{1, \ldots, d\}$,
(ii) $\pi_{ii} = 1$ for $i \in \{1, \ldots, d\}$,
(iii) $\pi_{ij} \leq \pi_{ik}\pi_{kj}$ for $i, j, k \in \{1, \ldots, d\}$.

In the financial literature, a matrix satisfying the above constraints is called a bid-ask matrix. In our context, the entry π_{ij} can be interpreted as the amount of capital that needs to be taken from business line i in order to transfer 1 unit of capital to business line j.

We note that each of the above ruin sets is a cone, i.e. it satisfies $L = uL$ for $u > 0$, as assumed in Assumption 5.1.

We maintain the notation $A = \mathbf{b} - L \in \mathcal{R}$. Note that we can write the ruin probability as

$$\begin{aligned}\psi_{\mathbf{b},L}(u) &= P(\mathbf{R}_t \in L \text{ for some } t \geq 0) \\ &= P\left(\sum_{i=1}^{n} \mathbf{X}^{(i)} - Y_i \mathbf{p} \in uA \text{ for some } n \geq 1\right) \\ &= P\left(\sum_{i=1}^{n} \mathbf{Z}^{(i)} \in uA \text{ for some } n \geq 1\right),\end{aligned} \tag{5.3}$$

where $\mathbf{Z}^{(i)} = \mathbf{X}^{(i)} - Y_i \mathbf{p}$, $i = 1, 2, \ldots$. We let $\mathbf{Z}$ denote a generic element of the sequence $(\mathbf{Z}^{(i)})_{i \geq 1}$. We will assume a positive safety loading, an assumption that takes now the form

$$\mathbf{c} = -\mathbb{E}[\mathbf{Z}] > \mathbf{0},$$

see e.g. Asmussen (2000). The assumption of the finite mean for the claim sizes implies that

$$\theta := \int_0^\infty F\Big([0, \infty)^d + v\mathbf{c}\Big)\, dv < \infty,$$

and we can define a probability measure on $\mathbb{R}^d$, supported by $[0, \infty)^d$, by

$$F^I(\cdot) = \frac{1}{\theta} \int_0^\infty F(\cdot + v\mathbf{c})\, dv. \tag{5.4}$$

Denote

$$H(u) = \int_0^\infty F(uA + v\mathbf{c})\, dv, \; u > 0. \tag{5.5}$$

The following is the main result of this section.

Theorem 5.2 *Suppose that the law F^I is in $\mathscr{S}_A$. Then the ruin probability $\psi_{\mathbf{b},L}$ satisfies*

$$\lim_{u \to \infty} \frac{\psi_{\mathbf{b},L}(u)}{H(u)} = 1. \tag{5.6}$$

Remark 5.3 Notice, for comparison, that in the univariate case, with the ruin set $L = (-\infty, 0)$ (and $b = 1$) we have $A = (1, \infty)$, and

$$H(u) = \int_0^\infty \overline{F}(u + vc)\, dv = \frac{1}{c} \int_u^\infty \overline{F}(v)\, dv.$$

In this case the statement (5.6) agrees with the standard univariate result on subexponential claims; see e.g. Theorem 1.3.8 in Embrechts et al. (1997). If the claim arrival process is Poisson, then this is Eq. 2.7 of Section 2.

Proof of Theorem 5.2 We start by observing that the function H is proportional to the tail of a subexponential distribution F_A^I and, hence, can itself be viewed as the tail of a subexponential distribution. We can and will, for example, simply refer to the "long tail property" of H.

We use the "one big jump" approach to heavy tailed large deviations; see e.g. Zachary (2004), and the first step is to show that

$$\lim_{u\to\infty} \frac{\int_0^\infty P(\mathbf{Z} \in uA + v\mathbf{c})\, dv}{H(u)} = 1\,. \tag{5.7}$$

Indeed, the upper bound in Eq. 5.7 follows from the fact that A is increasing. For the lower bound, notice that, by Fatou's lemma, it is enough to prove that that for each fixed y,

$$\lim_{u\to\infty} \frac{\int_0^\infty F(uA + v\mathbf{c} + y\mathbf{p})\, dv}{H(u)} = 1\,.$$

This, however, follows from the fact that by part (d) of Lemma 4.3, there exists some $u_1 > 0$ such that for $u > u_1$ we have $(u+u_1)A + v\mathbf{c} \subset uA + v\mathbf{c} + y\mathbf{p}$, and the long tail property of H.

We proceed to prove the lower bound in Eq. 5.6. Let $\mathbf{S}_n := \sum_{i=1}^n \mathbf{Z}^{(i)}$, $n = 1, 2, \ldots$. Let ϵ, δ be small positive numbers, by the Weak Law of Large Numbers, we can choose $K = K_{\epsilon,\delta}$ so large that

$$P\big(\mathbf{S}_n > -(K + n(1+\epsilon))\mathbf{c}\big) > 1-\delta,\ n = 1, 2, \ldots\,.$$

Define $M_n = \sup\{u > 0 : \mathbf{S}_i \in uA \text{ for some } 1 \le i \le n\}$ and $M = \sup\{u > 0 : \mathbf{S}_n \in uA \text{ for some } n \ge 1\}$. For $u > 0$,

$$\begin{aligned}
\psi_{\mathbf{b},L}(u) &= P(M > u) = \sum_{n\ge 0} P(M_n \le u,\ \mathbf{S}_{n+1} \in uA)\\
&\ge \sum_{n\ge 0} P\big(M_n \le u,\ \mathbf{S}_n > -(K+n(1+\epsilon))\mathbf{c},\ \mathbf{Z}^{(n+1)} \in uA + (K+n(1+\epsilon))\mathbf{c}\big)\\
&\ge \sum_{n\ge 0} (1-\delta-P(M_n > u))P\big(\mathbf{Z}^{(n+1)} \in uA + K\mathbf{c} + n(1+\epsilon)\mathbf{c}\big)\\
&\ge (1-\delta-P(M>u)) \sum_{n\ge 0} P\big(\mathbf{Z} \in uA + K\mathbf{c} + n(1+\epsilon)\mathbf{c}\big)\,.
\end{aligned}$$

Rearranging, using the monotonicity of A and the fact that $\mathbf{Z}$ has a finite mean, we see that

$$\begin{aligned}
\psi_{\mathbf{b},L}(u) &\ge \frac{(1-\delta)\sum_{n\ge 0} P(\mathbf{Z} \in uA + K\mathbf{c} + n(1+\epsilon)\mathbf{c})}{1 + \sum_{n\ge 0} P(\mathbf{Z} \in uA + K\mathbf{c} + n(1+\epsilon)\mathbf{c})}\\
&\sim \frac{1-\delta}{1+\epsilon} \int_0^\infty P\big(\mathbf{Z} \in uA + K\mathbf{c} + v\mathbf{c}\big)\, dv\\
&\sim \frac{1-\delta}{1+\epsilon} \int_0^\infty P\big(\mathbf{Z} \in uA + v\mathbf{c}\big)\, dv,\ u \to \infty\,.
\end{aligned}$$

Specifically, in the first step above we used the fact that the sum in the denominator is finite and converges to zero as $u \to \infty$ since $\mathbf{Z}$ has a finite mean, and replaced the sum in the numerator by the corresponding integral; the latter step is legitimate since A is monotone. Finally, we changed the variable of integration in the integral. Note also that the last step uses the long tail property of F_A^I. Letting δ, ϵ to 0, we have, thus, obtained the lower bound in Eq. 5.6. We proceed to prove a matching upper bound.

Fix $0 < \epsilon < 1$. For $r > 0$, we define a sequence (τ_n) as follows: we set $\tau_0 = 0$, and

$$\tau_1 = \inf\{n \geq 1 : \mathbf{S}_n \in rA - n(1-\epsilon)\mathbf{c}\} .$$

For $m \geq 2$, we set $\tau_m = \infty$ if $\tau_{m-1} = \infty$. Otherwise, let

$$\tau_m = \tau_{m-1} + \inf\{n \geq 1 : \mathbf{S}_{n+\tau_{m-1}} - \mathbf{S}_{\tau_{m-1}} \in rA - n(1-\epsilon)\mathbf{c}\}.$$

If we let $\gamma = P(\tau_1 < \infty)$, then for any $m \geq 1$, $P(\tau_m < \infty) = \gamma^m$. By the positive safety loading assumption, $\gamma \to 0$ as $r \to \infty$. Note that for $u > 0$,

$$P\big(\tau_1 < \infty,\, \mathbf{S}_{\tau_1} \in uA\big) = \sum_{n\geq 1} P(\tau_1 = n,\, \mathbf{S}_n \in uA) \leq \sum_{n\geq 1} P\big(\mathbf{S}_{n-1} \in rA^c - (n-1)(1-\epsilon)\mathbf{c},\, \mathbf{S}_n \in uA\big) .$$

By part (c) of Lemma 4.3, $\mathbf{S}_n \in uA$ if and only if $\sup_{\mathbf{p}\in I_A} \mathbf{p}^T \mathbf{S}_n > u$. Further,

$$\sup_{\mathbf{p}\in I_A} \mathbf{p}^T \mathbf{S}_n \leq \sup_{\mathbf{p}\in I_A} \mathbf{p}^T \big(\mathbf{S}_{n-1} + (n-1)(1-\epsilon)\mathbf{c}\big) + \sup_{\mathbf{p}\in I_A} \mathbf{p}^T \big(\mathbf{Z}^{(n)} - (n-1)(1-\epsilon)\mathbf{c}\big).$$

Let $u > r$. Recalling Lemma 4.5, if $\mathbf{S}_{n-1} \in rA^c - (n-1)(1-\epsilon)\mathbf{c}$, then $\sup_{\mathbf{p}\in I_A} \mathbf{p}^T \big(\mathbf{S}_{n-1} + (n-1)(1-\epsilon)\mathbf{c}\big) \leq r$, so for $\sup_{\mathbf{p}\in I_A} \mathbf{p}^T \mathbf{S}_n > u$ to hold, it must be the case that $\sup_{\mathbf{p}\in I_A} \mathbf{p}^T \big(\mathbf{Z}^{(n)} - (n-1)(1-\epsilon)\mathbf{c}\big) > u - r$, implying that $\mathbf{Z}^{(n)} \in (u-r)A + (n-1)(1-\epsilon)\mathbf{c}$.

Summing up, we see that, as $u \to \infty$,

$$\begin{aligned} P\big(\tau_1 < \infty,\, \mathbf{S}_{\tau_1} \in uA\big) &\leq \sum_{n\geq 1} P\big(\mathbf{Z}^{(n)} \in (u-r)A + (n-1)(1-\epsilon)\mathbf{c}\big) \\ &\sim \int_0^\infty P(\mathbf{Z} \in (u-r)A + v(1-\epsilon)\mathbf{c})\, dv \\ &\sim \frac{1}{1-\epsilon} H(u-r) . \end{aligned}$$

Letting $\epsilon \to 0$ and using the long tail property of H, we obtain

$$\limsup_{u\to\infty} \frac{P\big(\tau_1 < \infty,\, \mathbf{S}_{\tau_1} \in uA\big)}{H(u)} \leq 1 . \tag{5.8}$$

Let $(\mathbf{V}^{(i)})$ be a sequence of independent identically distributed random vectors whose law is the conditional law of $\mathbf{S}_{\tau_1}$ given that $\tau_1 < \infty$. By Eq. 5.8, there is a distribution B on $[0, \infty)$ such that $\overline{B}(u) \sim \gamma^{-1} H(u)$ as $u \to \infty$ and

$$P\big(\mathbf{V}^{(1)} \in uA\big) \leq \overline{B}(u) \ \text{ for all } u \geq 0.$$

Note, further, that by the definition of the sequence (τ_m), for every $m \geq 0$, on the event $\{\tau_m < \infty\}$, we have, for $1 \leq i < \tau_{m+1}$, $\mathbf{S}_{\tau_m+i} - \mathbf{S}_{\tau_m} \in rA^c - i(1-\epsilon)\mathbf{c} \subset rA^c - (1-\epsilon)\mathbf{c}$. If $\mathbf{S}_{\tau_m} \in (u-r)A^c + (1-\epsilon)\mathbf{c}$, then we have $\mathbf{S}_{\tau_m+i} \in uA^c$. Hence, for the event $\{\mathbf{S}_n \in uA \text{ for some } n\}$ to occur, we must have

$$\mathbf{S}_{\tau_m} \in \big((u-r)A + (1-\epsilon)\mathbf{c}\big) \cup uA \ \text{ for some } m.$$

Therefore, noting that $\big((u-r)A + (1-\epsilon)\mathbf{c}\big) \cup uA \subset (u-r)A$, we can use Lemma 4.9 to obtain

$$\begin{aligned}
\psi_{\mathbf{b},L}(u) &= P(M > u) \le \sum_{m\ge 1} P\big(\mathbf{S}_{\tau_m} \in \big((u-r)A + (1-\epsilon)\mathbf{c}\big) \cup uA\big) \\
&\le \sum_{m\ge 1} P\big(\mathbf{S}_{\tau_m} \in (u-r)A\big) \\
&= \sum_{m\ge 1} \gamma^m P\big(\mathbf{V}^{(1)} + \cdots + \mathbf{V}^{(m)} \in (u-r)A\big) \\
&\le \sum_{m\ge 1} \gamma^m \overline{B^{(m)}}(u-r).
\end{aligned}$$

By the assumption, the H is the tail of a subexponential distribution, and, hence, B is subexponential as well. This implies that

$$\lim_{u\to\infty} \frac{\overline{B^{(m)}}(u)}{\overline{B}(u)} = m\,,$$

and that for any $\epsilon > 0$, there exists $K > 0$ such that for all $u > 0$ and $m \ge 1$,

$$\frac{\overline{B^{(m)}}(u)}{\overline{B}(u)} \le K(1+\epsilon)^m.$$

Since we can make $\gamma > 0$ as small as we wish by choosing r large, we can use the dominated convergence theorem to obtain

$$\limsup_{u\to\infty} \frac{\psi_{\mathbf{b},L}(u)}{\gamma \overline{B}(u-r)} = \sum_{m\ge 1} \gamma^{m-1} m = \frac{1}{(1-\gamma)^2}.$$

Letting $r \to \infty$, which makes $\gamma \to 0$, we have that

$$\limsup_{u\to\infty} \frac{\psi_{\mathbf{b},L}(u)}{H(u)} \le 1\,,$$

which is the required upper bound in Eq. 5.6. $\square$

We finish this section by returning to the special case of multivariate regularly varying claims. Recall that, by Proposition 4.14, the distributions in $MRV(\alpha,\mu)$ are in $\mathscr{S}_{\mathcal{R}}$. The asymptotic behaviour of the ruin probability with the solvency set L^c given by Eq. 5.2, and multivariate regularly varying claims with $\alpha > 1$, was determined by Hult and Lindskog (2006). To state their result, notice that the tail measure of a random vector $\mathbf{X}$ (recall (4.8)) is determined up to a scaling by a positive constant, and a different scaling in the tail measure can be achieved by scaling appropriately the function b in Eq. 4.8. Let us scale the tail measure μ in such a way that it assigns unit mass to the complement of the unit ball in $\mathbb{R}^d$. The norm we choose is unimportant, but for consistency with the notation used elsewhere in the paper, let us use the L_1 norm. With this convention, we can restate (4.8) as

$$\frac{P(\mathbf{X} \in u\cdot)}{P(\|\mathbf{X}\| > u)} \xrightarrow{v} \mu \tag{5.9}$$

vaguely on $[-\infty, \infty]^d \backslash \{\mathbf{0}\}$. To avoid a degenerate situation (and the resulting complications in the notation) we will assume that $\mu\{\mathbf{x} : x_i > 0\} > 0$ for each $i = 1, \dots, d$. It was shown by Hult and Lindskog (2006) that under the assumption (5.9) (and with the solvency set L^c given by (5.2)), the ruin probability satisfies

$$\lim_{u\to\infty} \frac{\psi_{\mathbf{b},L}(u)}{uP(\|\mathbf{X}\| > u)} = \int_0^\infty \mu(\mathbf{b} - L + v\mathbf{c})\, dv. \tag{5.10}$$

We extend the above result to all ruin sets satisfying Assumption 5.1.

Proposition 5.4 *Assume that the ruin set L satisfies Assumption 5.1. If the claim sizes satisfy* (5.9) *with* $\alpha > 1$ *and the non-degeneracy assumption, then* (5.10) *holds.*

Proof By Theorem 5.2, it suffices to show that

$$\lim_{u\to\infty} \frac{\int_0^\infty P(\mathbf{X} \in uA + v\mathbf{c})\, dv}{uP(\|\mathbf{X}\| > u)} = \int_0^\infty \mu(A + v\mathbf{c})\, dv\,,$$

which we proceed to do. By a change of variables,

$$\frac{\int_0^\infty P(\mathbf{X} \in uA + v\mathbf{c})\, dv}{uP(\|\mathbf{X}\| > u)} = \frac{\int_0^\infty P(\mathbf{X} \in u(A + v\mathbf{c}))\, dv}{P(\|\mathbf{X}\| > u)}\,, \tag{5.11}$$

and for every $v > 0$,

$$\frac{P(\mathbf{X} \in u(A + v\mathbf{c}))}{P(\|\mathbf{X}\| > u)} \to \mu(A + v\mathbf{c})$$

as $u \to \infty$. In the last step we use Eq. 5.9, and the fact that the tail measure does not charge the boundary of sets in $\mathcal{R}$, shown in the proof of Proposition 4.14. Therefore, we only need to justify taking the limit inside the integral in Eq. 5.11. However, by the definition of the set A,

$$\frac{P(\mathbf{X} \in u(A + v\mathbf{c}))}{P(\|\mathbf{X}\| > u)} \le \sum_{i=1}^{d} \frac{P(X^{(i)} > ub_i + uvc_i)}{P(X^{(i)} > u)}\,.$$

The non-degeneracy assumption on the measure μ implies that each $X^{(i)}$ is itself regularly varying with exponent α. Therefore, by the Potter bounds, there are finite positive constants C_i, $i = 1, \dots, d$, and a number $\varepsilon \in (0, \alpha - 1)$ such that for all $u \ge 1$,

$$\frac{P(X^{(i)} > ub_i + uvc_i)}{P(X^{(i)} > u)} \le C_i(b_i + vc_i)^{-(\alpha-\varepsilon)},\ i = 1, \dots, d\,.$$

Since the functions in the right hand side are integrable, the dominated convergence theorem applies. □

References

Asmussen, S.: Ruin probabilities. World Scientific Publishing Company (2000)

Baltrunas, A., Omey, E., Gulck, S.V.: Hazard rates and subexponential distributions. Publications de l'Institut Mathématique (Beograd) **80**(94), 29–46 (2006)

Chistyakov, V.: A theorem on sums of independent random variables and its applications to branching random processes. Theory of Probability and its Applications **9**, 640–648 (1964)

Cline, D., Resnick, S.: Multivariate subexponential distributions. Stoch. Process. Appl. **42**, 49–72 (1992)

Cline, D., Samorodnitsky, G.: Subexponentiality of the product of independent random variables. Stoch. Process. Appl. **49**, 75–98 (1994)

Embrechts, P., Goldie, C.: On closure and factorization properties of subexponential distributions. Journal of Australian Mathematical Society, Series A **29**, 243–256 (1980)

Embrechts, P., Goldie, C., Veraverbeke, N.: Subexponentiality and infinite divisibility. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete **49**, 335–347 (1979)

Embrechts, P., Klüppelberg, C., Mikosch, T.: Modelling extremal events for insurance and finance. Springer, Berlin (1997)

Foss, S., Konstantopoulos, T., Zachary, S.: The principle of a single big jump: discrete and continuous time modulated random walks with heavy-tailed increments. J. Theor. Probab. **20**, 581–612 (2007)

Goldie, C.: Subexponential distributions and dominated-variation tails. J. Appl. Probab. **15**, 440–442 (1978)

Hult, H., Lindskog, F.: Heavy-tailed insurance portfolios: buffer capital and ruin probabilities. Technical Report. (2006)

Leslie, J.: On the non-closure under convolution of the subexponential family. J. Appl. Probab. **26**, 58–66 (1989)

Omey, E.: Subexponential distribution functions in $\mathbb{R}^d$. J. Math. Sci. **138**, 5434–5449 (2006)

Omey, E., Mallor, F., Santos, J.: Multivariate subexponential distributions and random sums of random vectors. Adv. Appl. Probab. **38**, 1028–1046 (2006)

Omey, E., Vesilo, R.: The difference between the product and the convolution product of distribution functions in $\mathbb{R}^d$. Publications de l'Institut Mathematique (Beograd) **89**(103), 19–36 (2011)

Pitman, E.: Subexponential distribution functions. Journal of Australian Mathematical Society, Series A **29**, 337–347 (1980)

Resnick, S.: Extreme values regular variation and point processes. Springer, New York (1987)

Resnick, S.: Heavy-tail phenomena: probabilistic and statistical modeling. Springer, New York (2007)

Rockafellar, R.: Convex analysis. Princeton University Press, Princeton (2015)

Willekens, E.: Higher order theory for subexponential distributions. Ph.D. thesis, K.U. Leuven. In Dutch. (1986)

Zachary, S.: A note on Veraverbeke's theorem. Queueing Systems **46**, 9–14 (2004)