MATHEMATICS OF OPERATIONS RESEARCH
Vol. 35, No. 1, February 2010, pp. 102–119
ISSN 0364-765X | EISSN 1526-5471 | 10 | 3501 | 0102

DOI 10.1287/moor.1090.0430

On Sums of Conditionally Independent Subexponential Random Variables

Serguei Foss, Andrew Richards
School of Mathematics and Computer Sciences and the Maxwell Institute for Mathematical Sciences,
Heriot-Watt University, Edinburgh, EH14 4AS, Scotland, UK {s.foss@hw.ac.uk, awr2@hw.ac.uk}

The asymptotic tail behaviour of sums of independent subexponential random variables is well understood, one of the main characteristics being *the principle of the single big jump*. We study the case of dependent subexponential random variables, for both deterministic and random sums, using a fresh approach, by considering conditional independence structures on the random variables. We seek sufficient conditions for the results of the theory with independent random variables to still hold. For a subexponential distribution, we introduce the concept of a boundary class of functions, which we hope will be a useful tool in studying many aspects of subexponential random variables. The examples we give demonstrate a variety of effects owing to the dependence, and are also interesting in their own right.

Key words: heavy tails; subexponentiality; conditional independence
MSC2000 subject classification: Primary: 60G99
OR/MS subject classification: Primary: probability: applications; distributions; stochastic model applications
History: Received June 21, 2008; revised February 19, 2009, April 16, 2009, and July 13, 2009. Published online in *Articles in Advance* December 8, 2009.

1. Introduction. Finding the asymptotic tail behaviour of sums of heavy-tailed random variables is an important problem in finance, insurance, and many other disciplines. The case when the random variables are independent and subexponentially distributed has been extensively studied and is well understood. The key idea is that such a sum will exceed a high threshold because of a single, very large jump; following other authors we shall refer to this as *the principle of the single big jump*. However, for many practical purposes the independence assumption is too restrictive. In recent years, many authors have developed results in this area (see, for example, Albrecher et al. [1], Alink et al. [2], Denuit et al. [8], Goovaerts et al. [12], Kortschak and Albrecher [14], Ko and Tang [15], Laeven et al. [16], Mitra and Resnick [17], Tang [19], Tang and Tsitsiashvili [20], and references therein). Denuit et al. [8] constructed bounds for these sums, but did not consider asymptotics. Goovaerts et al. [12] considered the situation of dependent random variables with regularly varying tails; there have also been results on negative dependence for various classes of subexponential distributions (see, for example Tang [19]) and for dependence structures that are "not too positive" (see Ko and Tang [15]).

Once we drop the requirement of independence, two questions naturally arise. First, *what kind of behaviours can occur as the dependence between the random variables strengthens*? And secondly, *how far beyond the independent case does the principle of the single big jump still hold*? These questions are of real interest, both from theoretical and practical viewpoints.

Albrecher et al. [1] consider the first question for the sum of two dependent random variables. Their approach, as for many authors, is to study the possible effects of the dependence by considering the copula structure. They demonstrate that many possible behaviours naturally occur, and that, in some specific cases the principle of the single big jump is insensitive to the strength of the copula structure. Other papers that concentrate on the copula structure include Alink et al. [2] and Kortschak and Albrecher [14]. Mitra and Resnick [17] investigate random variables belonging to the maximum domain of attraction of the Gumbel distribution and which are asymptotically independent. The results we present contain overlap with all these approaches, but we neither impose a particular dependence structure, nor a particular distribution for the random variables, beyond the necessary constraint that at least one be subexponential.

We wish to consider the second question, and to establish conditions on the strength of the dependence which will preserve the results of the theory established for independent random variables; in particular, the principle of the single big jump. This principle is well known. However, we would like to examine it again from a probabilistic point of view by considering the sum of two identically distributed subexponential random variables X_1, X_2.

$$\begin{aligned}\mathbf{P}(X_1+X_2>x) &= \mathbf{P}(X_1\vee X_2>x)+\mathbf{P}(X_1\vee X_2\le x,\ X_1+X_2>x)\\ &= \mathbf{P}(X_1>x)+\mathbf{P}(X_2>x)-\mathbf{P}(X_1\wedge X_2>x)+\mathbf{P}(X_1\vee X_2\le x,\ X_1+X_2>x)\\ &:= \mathbf{P}(X_1>x)+\mathbf{P}(X_2>x)-P_2(x)+P_1(x), \qquad (1)\end{aligned}$$

where $X_1 \vee X_2 = \max(X_1, X_2)$ and $X_1 \wedge X_2 = \min(X_1, X_2)$. If $P_1(x)$ is negligible compared to $\mathbf{P}(X_1 > x)$, which in the independent case follows from the definition of subexponentiality, we shall say that we have the *principle of the big jump*. If, in addition, $P_2(x)$ is negligible compared to $\mathbf{P}(X_1 > x)$, as again is straightforward in the independent case, then we shall say that we have the *principle of the single big jump*. If the dependence is very strong, for instance if $X_1 = X_2$ a.s. (almost surely), then clearly the principle of the single big jump fails. We shall see in Example 3 a more interesting example where the principle of the big jump holds, but not the principle of the single big jump; but nonetheless a high level is exceeded because of a single big jump a positive fraction of the time.

We consider sums of random variables that are conditionally independent on some sigma algebra. This is a fresh approach to studying the effect of dependence on subexponential sums and allows a great deal of generality (in particular, we need neither specify a particular subclass of subexponential distribution for which our results hold, nor assume the summands are identically distributed, nor specify any particular copula structure). We believe this is a fruitful line of enquiry, both practically and theoretically, as the range of examples we give illustrates.

Clearly, any sequence of random variables can be considered to be conditionally independent by choosing an appropriate sigma algebra on which to condition. This is an obvious observation, and in itself not really helpful. However, there are practical situations where a conditional independence structure arises naturally from the problem. As an example, consider a sequence of identical random variables $X_1, X_2, \ldots, X_n$, each with distribution function F_β depending on some parameter β that is itself drawn from a different distribution. The X_i are independent once β is known: this is a typically Bayesian situation. It is natural to view the X_i as conditionally independent on the sigma algebra generated by β. We suppose the X_i to have subexponential (unconditional) distribution F and ask under what conditions the distribution of the sum follows the principle of the single big jump.

In addition to the assumption of conditional independence, we assume that the distributions of our random variables are asymptotically equivalent to multiples of a given reference subexponential distribution (or more generally we can assume weak equivalence to the reference distribution). This allows us to consider nonidentical random variables. As an example, developed fully in Example 4, which follows ideas in Laeven et al. [16], we consider the problem of calculating the *discounted loss reserve*; this can also be viewed as finding the value of a perpetuity. Let the i.i.d. sequence $X_1, X_2, \ldots, X_n$ denote the net losses in successive years, and the i.i.d sequence $V_1, \ldots, V_n$ denote the corresponding present value discounting factors, where the two sequences are mutually independent. Then $Y_i = X_i \Pi_{j=1}^{i} V_j$ represents the present value of the net loss in year i, and $S_n = \Sigma_{i=1}^{n} Y_i$ is the discounted loss reserve. Conditional on $\sigma(V_1, \ldots, V_n)$ the random variables Y_i are independent. Let the (unconditional) distribution function of Y_i be F_i. We suppose there is a reference subexponential distribution F and finite constants $c_1, \ldots, c_n$, not all zero, such that, for all $i = 1, 2, \ldots, n$,

$$\lim_{x\to\infty} \frac{\bar{F}_i(x)}{\bar{F}(x)} = c_i.$$

We seek conditions on the dependence which will ensure that the principle of the single big jump holds for the discounted loss reserve.

More generally, we want to consider both deterministic sums and randomly stopped sums, where the stopping time τ is independent of the X_i and has light-tailed distribution.

Foss et al. [11] studied time modulated random walks with heavy-tailed increments. In their proofs of two key theorems (Theorems 2.2 and 3.2) they used a coupling argument involving the sum of two conditionally independent random variables that entailed proving a lemma (Lemma A.2) which considered a particular case of conditional independence. The investigation in the present paper considers this problem in much greater generality, whilst retaining the flavour of the simple situation in Foss et al. [11].

The statements of the propositions that we prove in §2 therefore hold no surprises, and indeed, once the conditions under which these propositions hold were determined, the proofs followed relatively straightforwardly with no need for complicated machinery. The interest and effort was in the formulation of the conditions in the first place, which constituted the major intellectual work in this paper, and in finding means by which these conditions could be efficiently checked.

A very useful tool in the study of subexponential distributions is the class of functions which reflect the fact that every subexponential distribution is long-tailed. For a given subexponential distribution F this is the class of functions $\mathscr{h}_F$, such that whenever $h \in \mathscr{h}_F$, h is monotonically tending to infinity, and

$$\lim_{x\to\infty} \frac{\bar{F}(x - h(x))}{\bar{F}(x)} = 1. \tag{2}$$

A quantity that occurs often in the study of such distributions is

$$\int_{h(x)}^{x-h(x)} \bar{F}(x-y)\, F(dy),$$

for any h satisfying (2). The importance of this stems from the fact that this quantity is negligible compared to $\bar{F}(x)$ as $x \to \infty$:

$$\lim_{x\to\infty} \int_{h(x)}^{x-h(x)} \frac{\bar{F}(x-y)}{\bar{F}(x)}\, F(dy) = 0. \tag{3}$$

Intuitively this means that, when considering the probability that the sum of two i.i.d. subexponential random variables exceeds some high level x, the probability that both of them are of "intermediate" size is negligible compared to the probability that exactly one of them exceeds x. Because of the form of the integral, it is convenient only to consider those functions h which satisfy $h(x) < x/2$.

It is clear that if $h_1(x)$ belongs to the class $\hbar_F$, then any $h_2(x) \le h_1(x)$ also satisfies (2). For many long-tailed functions a *boundary class* of functions, $\mathscr{H}$, exists such that the statement $h \in \hbar_F$ is equivalent to $h(x) = o(H(x))$ for every $H \in \mathscr{H}$. This boundary class is particularly useful in dealing with expressions such as that in (3). In these types of expressions we need to find a suitable function h, but the class of functions satisfying the long-tail property (2) is very rich, and finding an appropriate function can be difficult. We note that if we have found an appropriate h_1 satisfying (3), then any other h_2 in the same class such that $h_2(x) > h_1(x)$ will also satisfy (3); we call this *increasing function behaviour*. We show that in cases where the boundary class exists, for any property that exhibits this increasing function behaviour, the property is satisfied for some function $h \in \hbar_F$ if and only if the property is satisfied for all functions in the boundary class. Furthermore, because all functions in the boundary class are weakly equivalent (see §3 for precise definitions and statements) it suffices to verify such a property only for multiples of a single function. We hope that this technical tool will be of use to other researchers.

We give a wide range of examples of collections of random variables, some satisfying the principle of the single big jump, some not, and we suggest that these examples are of independent interest in and of themselves.

The paper is structured as follows. In §2 we formulate our assumptions, then state and prove our main results for conditionally independent nonnegative random variables satisfying the principle of the single big jump, leaving the more general case of real-valued random variables to §5. In §3 we introduce the concept of the Boundary Class for long-tailed distributions, and give some typical examples. In §4 we give examples of conditionally independent subexponential random variables, some of which satisfy the principle of the single big jump, and one of which does not. In §5 we extend our investigation to any real-valued subexponential random variables. This involves imposing an extra condition. We give an example that shows that this condition is nonempty and necessary. Finally, in §6 we collect together the different notation we have used, and also give definitions of the standard classes of distributions (heavy-tailed, long-tailed, subexponential, regularly varying, and so on) that we use in this paper.

2. Main definitions, results, and proofs. A distribution function F supported on the positive half-line is subexponential if and only if

$$\bar{F}^{*2}(x) := \int_0^x \bar{F}(x-y)\, F(dy) + \bar{F}(x) \sim 2\bar{F}(x).$$

It is known (see, for example, Foss and Zachary [10]), and may be easily checked, that a distribution supported on the positive half-line is subexponential if and only if the following two conditions are met:

(i) F is long-tailed. That is, there exists a nondecreasing function $h(x)$, tending to infinity, such that (2) holds.[1] (Examples include: for F regularly varying (see §6.2 for definition), then we can choose $h(x) = x^\delta$, where $0 < \delta < 1$; for F Weibull, with parameter $0 < \beta < 1$, we can choose $h(x) = x^\delta$, where $0 < \delta < 1 - \beta$.)

(ii) For any $h(x) < x/2$ tending monotonically to infinity,

$$\int_{h(x)}^{x-h(x)} \bar{F}(x-y)\, F(dy) = o(\bar{F}(x)). \tag{4}$$

[1] We observe that, given a random variable X with subexponential distribution F, the function h measures how light, compared to X, the distribution of a random variable Y must be so that the tail distribution of the sum $X+Y$ is insensitive to the addition of Y, not assumed to be independent of X. In particular, if $\mathbf{P}(Y > h(x)) = o(\bar{F}(x))$, then $\mathbf{P}(X+Y > x) \sim \mathbf{P}(X > x)$ regardless of how strong any dependence between X and Y is. In the case of regular variation, this comment is originally from Klüppelberg [13].

We work in a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Let X_i, $i = 1, 2, \ldots,$ be nonnegative random variables with distribution function (d.f.) F_i. Let F be a subexponential reference distribution concentrated on the positive half-line and let h be a function satisfying the long-tailed condition (2). Let $\mathcal{G}$ be a σ-algebra, $\mathcal{G} \subset \mathcal{F}$. We make the following assumptions about the dependence structure of the X_i's:

(D1) $X_1, X_2, \ldots$ are conditionally independent given $\mathcal{G}$. That is, for any collection of indices $\{i_1, \ldots, i_r\}$, and any collection of sets $\{B_{i_1}, \ldots, B_{i_r}\}$, all belonging to $\mathcal{F}$, then $\mathbf{P}(X_{i_1} \in B_{i_1}, \ldots, X_{i_r} \in B_{i_r} \mid \mathcal{G}) = \mathbf{P}(X_{i_1} \in B_{i_1} \mid \mathcal{G})\mathbf{P}(X_{i_2} \in B_{i_2} \mid \mathcal{G}) \cdots \mathbf{P}(X_{i_r} \in B_{i_r} \mid \mathcal{G})$.

(D2) For each $i \geq 1$, $\bar{F}_i(x) \sim c_i \bar{F}(x)$, with at least one $c_i \neq 0$; and for all $i \geq 1$, there exists $c > 0$ such that $\bar{F}_i(x) \leq c\bar{F}(x)$ for all $x > 0$.

(D3) For each $i \geq 1$ there exists a nondecreasing function $r(x)$ and an increasing collection of sets $B_i(x) \in \mathcal{G}$, with $B_i(x) \to \Omega$ as $x \to \infty$, such that

$$\mathbf{P}(X_i > x \mid \mathcal{G})\mathbf{1}(B_i(x)) \leq r(x)\bar{F}(x)\mathbf{1}(B_i(x)) \quad \text{almost surely,} \tag{5}$$

and, as $x \to \infty$, uniformly in i,

(i) $\mathbf{P}(\bar{B}_i(h(x))) = o(\bar{F}(x))$,

(ii) $r(x)\bar{F}(h(x)) = o(1)$,

(iii) $r(x)\int_{h(x)}^{x-h(x)} \bar{F}(x-y)\, F(dy) = o(\bar{F}(x))$.

Remark 2.1. In many cases the dependence between the $\{X_i\}$ enables us to choose a common $B(x) = B_i(x)$, for all i. However, we allow for situations where this is not the case. There is no need for a similar generality in choice of the function $r(x)$ because of the uniformity in i. The function $r(x)$ can be chosen so that it is only eventually monotone increasing, and in the case where we are only considering a finite collection of random variables $\{X_i\}$ it is sufficient to show that the chosen function is asymptotically equivalent to a monotone increasing function.

Remark 2.2. If the collection of r.v.'s of interest is finite, then clearly the uniformity in i needed in conditions (D2) and (D3) is guaranteed.

Remark 2.3. If the reference distribution F has a tail that is intermediately regularly varying (see §6.2 for definitions), then it will be shown later that we can check that the conditions (D3) hold for some $h(x)$ satisfying (2) by verifying that the conditions hold when $h(x)$ is replaced by all the functions $H(x) = cx$ where $0 < c < 1/2$.

Remark 2.4. It will sometimes be the case that the random variables $X_1, X_2, \ldots,$ are not identically distributed, and are not all asymptotically equivalent to the reference distribution F. In these cases it is sufficient to require that they are weakly equivalent to F (see §6.1 for the definition of weak equivalence), and that they are subexponentially distributed. The uniformity condition will still be required.

Remark 2.5. The need for and the meaning of the bounding functions $r(x)$ and the bounding sets $B_i(x)$ will become apparent when we give some examples. However, some preliminary comments may assist at this stage.

- To preserve the desired properties from the independent scheme, we need to ensure that the influence of the σ-algebra $\mathcal{G}$ that controls the dependence is not too strong. This we have done by introducing the bounding function $r(x)$ for the ith random variable, which ensures that there are no events in $\mathcal{G}$ that totally predominate if a high level is exceeded. Although $r(x)$ may tend to infinity, it must not do so too quickly.
- Depending on the nature of the interaction of $\mathcal{G}$ with the random variables, there may be events in $\mathcal{G}$ that do overwhelmingly predominate when exceeding a high level; this is not a problem as long as these events are unlikely enough and their probability tends to zero as the level tends to infinity. Within the bounding sets $B_i(x)$ no events in $\mathcal{G}$ predominate, and we then require that the compliments $\bar{B}_i(x)$ decay quickly enough.

We have the following results.

Proposition 2.1. *Let X_i, $i = 1, 2, \ldots,$ satisfy conditions* (D1), (D2), *and* (D3) *for some subexponential F concentrated on the positive half-line and for some $h(x)$ satisfying* (2). *Then*

$$\mathbf{P}(X_1 + \cdots + X_n > x) \sim \sum_{i=1}^{n} \mathbf{P}(X_i > x) \sim \left(\sum_{i=1}^{n} c_i\right)\bar{F}(x).$$

Remark 2.6. Lemma A.2 in Foss et al. [11] follows directly from this proposition.

To use dominated convergence to generalize Proposition 2.1 to random sums, we need the following extension of Kesten's lemma.

LEMMA 2.1. *With the conditions of* (D1), (D2), *and* (D3), *for any* $\epsilon > 0$ *there exist* $V(\epsilon) > 0$ *and* $x_0 = x_0(\epsilon)$ *such that, for any* $x > x_0$ *and* $n \geq 1$,

$$\mathbf{P}(S_n > x) \leq V(\epsilon)(1+\epsilon)^n \bar{F}(x).$$

PROPOSITION 2.2. *If, in addition to the conditions of* (D1), (D2), *and* (D3), τ *is an independent counting random variable such that* $\mathbf{E}(e^{\gamma\tau}) < \infty$ *for some* $\gamma > 0$, *then*

$$\begin{aligned}\mathbf{P}(X_1 + \cdots + X_\tau > x) &\sim \mathbf{E}\left(\sum_{i=1}^{\tau} \mathbf{P}(X_i > x)\right)\\ &\sim \mathbf{E}\left(\sum_{i=1}^{\tau} c_i\right)\bar{F}(x).\end{aligned}$$

Clearly, checking that (D3(ii)) and (D3(iii)) hold is the most laborious part of guaranteeing the conditions for these propositions. Hence we propose a sufficient condition, analogous to the well-known condition for subexponentiality.

PROPOSITION 2.3. *Let* F *be a subexponential distribution concentrated on the positive half-line, let* $h(x)$ *be a function satisfying* (2), *and let* $r(x)$ *be a nondecreasing function. Let* $Q(x) := -\log(\bar{F}(x))$, *the hazard function for* F, *be concave for* $x \geq x_0$, *for some* $x_0 < \infty$. *Let*

$$x r(x) \bar{F}(h(x)) \to 0 \quad \text{as } x \to \infty. \tag{6}$$

Then conditions (D3(ii)) *and* (D3(iii)) *are satisfied.*

Before we can give examples of our conditions in practice, we need to address a specific issue. The conditions depend on being able to choose bounding functions $r(x)$ and bounding sets $B_i(x)$, which themselves depend on our choice of the little-h function satisfying (2). The choice of $h(x)$ is not unique, so the fact that one is unable to find appropriate bounding functions and sets for a particular little-h function does not imply that one cannot find them for some other choice of h. This is the problem we address in the next section.

Now we proceed with the proofs of our results.

PROOF OF PROPOSITION 2.1. First, consider $X_1 + X_2$. Assume, without loss of generality, that $c_1 > 0$. Let Y be a random variable, independent of X_1 and X_2 with distribution function F. We have the inequalities

$$\begin{aligned}\mathbf{P}(X_1 + X_2 > x) \leq{}& \mathbf{P}(X_1 > x - h(x)) + \mathbf{P}(X_2 > x - h(x))\\ &+ \mathbf{P}(h(x) < X_1 \leq x - h(x),\ X_2 > x - X_1),\end{aligned}$$

and

$$\mathbf{P}(X_1 + X_2 > x) \geq \mathbf{P}(X_1 > x) + \mathbf{P}(X_2 > x) - \mathbf{P}(X_1 > x,\ X_2 > x).$$

Now,

$$\begin{aligned}&\mathbf{P}(h(x) < X_1 \leq x - h(x),\ X_2 > x - X_1)\\ &\quad = \mathbf{E}(\mathbf{P}(h(x) < X_1 \leq x - h(x),\ X_2 > x - X_1 \mid \mathcal{G}))\\ &\quad = \mathbf{E}\left(\int_{h(x)}^{x-h(x)} \mathbf{P}(X_1 \in dy \mid \mathcal{G})\mathbf{P}(X_2 > x - y \mid \mathcal{G})(\mathbf{1}(B_2(x-y)) + \mathbf{1}(\bar{B}_2(x-y)))\right)\\ &\quad \leq r(x)\mathbf{E}\left(\int_{h(x)}^{x-h(x)} \mathbf{P}(X_1 \in dy \mid \mathcal{G})\mathbf{P}(Y > x - y)\right) + \mathbf{E}(\mathbf{1}(\bar{B}_2(h(x))))\\ &\quad = r(x)\int_{h(x)}^{x-h(x)} \mathbf{P}(X_1 \in dy)\bar{F}(x-y) + o(\bar{F}(x))\\ &\quad = o(\bar{F}(x)).\end{aligned}$$

Also,

$$\begin{aligned}\mathbf{P}(X_1 > x,\ X_2 > x) &= \mathbf{E}(\mathbf{P}(X_1 > x,\ X_2 > x \mid \mathcal{G})(\mathbf{1}(B_2(x)) + \mathbf{1}(\bar{B}_2(x))))\\ &\leq \mathbf{E}(\mathbf{P}(X_1 > x \mid \mathcal{G})\mathbf{P}(X_2 > x \mid \mathcal{G})\mathbf{1}(B_2(x))) + \mathbf{E}(\mathbf{1}(\bar{B}_2(x)))\\ &\leq r(x)\bar{F}(x)\mathbf{P}(X_1 > x) + o(\bar{F}(x))\\ &= o(\bar{F}(x)).\end{aligned}$$

Hence, $\mathbf{P}(X_1+X_2>x)\sim\mathbf{P}(X_1>x)+\mathbf{P}(X_2>x)$. Since $c_1>0$, then $\mathbf{P}(X_1>x)+\mathbf{P}(X_2>x)\sim(c_1+c_2)\bar{F}(x)$. Then, by induction, we have the desired result. □

Proof of Lemma 2.1. The proof follows the lines of the original proof by Kesten. Again, let Y be a random variable, independent of X_1 and X_2, and with distribution function F. For $x_0\geq 0$, which will be chosen later, and $k\geq 1$, put

$$\alpha_k=\alpha_k(x_0):=\sup_{x>x_0}\frac{\mathbf{P}(S_k>x)}{\bar{F}(x)}.$$

Also observe that

$$\sup_{0<x\leq x_0}\frac{\mathbf{P}(S_k>x)}{\bar{F}(x)}\leq\frac{1}{\bar{F}(x_0)}:=\alpha.$$

Take any $\epsilon>0$. Recall that for all $i>0$, $\bar{F}_i(x)\leq c\bar{F}(x)$, for some $c>0$ and for all $x>0$. Then, for any $n>1$,

$$\begin{aligned}\mathbf{P}(S_n>x)&=\mathbf{P}(S_{n-1}\leq h(x),\ X_n>x-S_{n-1})\\&\quad+\mathbf{P}(h(x)<S_{n-1}\leq x-h(x),\ X_n>x-S_{n-1})\\&\quad+\mathbf{P}(S_{n-1}>x-h(x),\ X_n>x-S_{n-1})\\&\equiv P_1(x)+P_2(x)+P_3(x).\end{aligned}$$

We bound

$$P_1(x)\leq\mathbf{P}(X_n>x-h(x))\leq cL(x_0)\bar{F}(x)$$

and

$$P_3(x)\leq\mathbf{P}(S_{n-1}>x-h(x))\leq\alpha_{n-1}L(x_0)\bar{F}(x)$$

for $x\geq x_0$, where $L(x)=\sup_{y\geq x}(\bar{F}(y-h(y))/(\bar{F}(y)))$.

For $P_2(x)$,

$$\begin{aligned}P_2(x)&=\mathbf{P}(h(x)<S_{n-1}\leq x-h(x),X_n>x-S_{n-1})\\&=\mathbf{E}\left(\int_{h(x)}^{x-h(x)}\mathbf{P}(S_{n-1}\in dy\mid\mathcal{G})\mathbf{P}(X_n>x-y\mid\mathcal{G})(\mathbf{1}(B_n(x-y))+\mathbf{1}(\bar{B}_n(x-y)))\right)\\&\leq\mathbf{E}\left(r(x)\int_{h(x)}^{x-h(x)}\mathbf{P}(S_{n-1}\in dy\mid\mathcal{G})\mathbf{P}(Y>x-y)\right)+\mathbf{P}(\bar{B}_n(h(x)))\\&=r(x)\int_{h(x)}^{x-h(x)}\mathbf{P}(S_{n-1}\in dy)\mathbf{P}(Y>x-y)+\mathbf{P}(\bar{B}_n(h(x)))\\&\leq r(x)\left(\int_{h(x)}^{x-h(x)}\mathbf{P}(Y\in dy)\mathbf{P}(S_{n-1}>x-y)+\mathbf{P}(S_{n-1}>h(x))\mathbf{P}(Y>x-h(x))\right)+\mathbf{P}(\bar{B}_n(h(x)))\\&\leq(\alpha_{n-1}+\alpha)r(x)\left(\int_{h(x)}^{x-h(x)}\mathbf{P}(Y\in dy)\mathbf{P}(Y>x-y)+\mathbf{P}(Y>h(x))\mathbf{P}(Y>x-h(x))\right)+\mathbf{P}(\bar{B}_n(h(x)))\\&=(\alpha_{n-1}+\alpha)\left(r(x)\int_{h(x)}^{x-h(x)}\frac{\bar{F}(x-y)}{\bar{F}(x)}F(dy)+r(x)\bar{F}(h(x))\bar{F}(x-h(x))\right)+\mathbf{P}(\bar{B}_n(h(x))).\end{aligned}$$

We now choose x_0 such that, for all $x\geq x_0$,

$$\begin{gathered}\frac{\bar{F}(x-h(x))}{\bar{F}(x)}\leq L(x_0)\leq 1+\frac{\epsilon}{4},\\ r(x)\int_{h(x)}^{x-h(x)}\frac{\bar{F}(x-y)}{\bar{F}(x)}F(dy)\leq\frac{\epsilon}{4},\\ r(x)\bar{F}(h(x))L(x_0)\leq\frac{\epsilon}{4},\\ \frac{\mathbf{P}(\bar{B}_n(h(x)))}{\bar{F}(x)}\leq 1,\end{gathered}$$

which can be done by virtue of the long-tailedness of F and conditions (D3). We then have that

$$P_2(x) \le \frac{\epsilon}{2}(\alpha_{n-1}+\alpha)\bar{F}(x)+\bar{F}(x).$$

We therefore have

$$\begin{aligned}\mathbf{P}(S_n>x) &\le cL(x_0)\bar{F}(x)+\frac{\epsilon}{2}(\alpha_{n-1}+\alpha)\bar{F}(x)+\bar{F}(x)+\alpha_{n-1}L(x_0)\bar{F}(x)\\ &\le R\bar{F}(x)+\left(1+\frac{3}{4}\epsilon\right)\alpha_{n-1}\bar{F}(x),\end{aligned}$$

for some $0<R<\infty$. Hence,

$$\alpha_n \le R+\left(1+\frac{3}{4}\epsilon\right)\alpha_{n-1}.$$

Then, by induction, we have

$$\begin{aligned}\alpha_n &\le \alpha_1\left(1+\frac{3}{4}\epsilon\right)^{n-1}+R\sum_{r=0}^{n-2}\left(1+\frac{3}{4}\epsilon\right)^r\\ &\le Rn\left(1+\frac{3}{4}\epsilon\right)^{n-1}\\ &\le V(\epsilon)(1+\epsilon)^n,\end{aligned}$$

for some constant $V(\epsilon)$ depending on ϵ.

This completes the proof. □

Proof of Proposition 2.2. The proof follows directly from Proposition 2.1, Lemma 2.1, and the dominated convergence theorem. □

Before proving Proposition 2.3, we prove the following lemma, which was originally used without proof in Denisov et al. [7].

Lemma 2.2. *Let F be long-tailed and concentrated on the positive real line. Then there exists a constant $C>0$ such that for any $b>a>0$,*

$$\int_a^b \bar{F}(x-y)\,F(dy) \le C\int_a^b \bar{F}(x-y)\bar{F}(y)\,dy.$$

Proof of Lemma 2.2. Let $y_0=a$, $s=[b-a]+1$, and $y_i=y_{i-1}+(b-a)/s$, $i=1,2,\dots,s$. Then $y_s=b$. There exists a constant C such that for any $y>0$, $\bar{F}(y)/(\bar{F}(y+1))\le\sqrt{C}<\infty$ since F is long-tailed. Then

$$\begin{aligned}\int_a^b \bar{F}(x-y)\,F(dy) &= \sum_{n=0}^{s-1}\int_{y_n}^{y_{n+1}}\bar{F}(x-y)\,F(dy)\\ &\le \sum_{n=0}^{s-1}\int_{y_n}^{y_{n+1}}\bar{F}(x-y_{n+1})(\bar{F}(y_n)-\bar{F}(y_{n+1}))\,dy\\ &\le \sum_{n=0}^{s-1}\int_{y_n}^{y_{n+1}}\sqrt{C}\bar{F}(x-y)\bar{F}(y_n)\,dy\\ &\le \sum_{n=0}^{s-1}\int_{y_n}^{y_{n+1}}C\bar{F}(x-y)\bar{F}(y)\,dy\\ &= C\int_a^b \bar{F}(x-y)\bar{F}(y)\,dy. \quad \square\end{aligned}$$

Proof of Proposition 2.3. Without loss of generality we may assume that $x_0=0$. Clearly (6) implies that condition (D3(ii)) holds. Since g is concave, the minimum of the sum $g(x-y)+g(y)$ on the interval $[h(x),x-h(x)]$ occurs at the endpoints of the interval. From Lemma 2.2, there exists a constant $C>0$ such that

$$\begin{aligned}\int_{h(x)}^{x-h(x)}\bar{F}(x-y)\,F(dy) &\le C\int_{h(x)}^{x-h(x)}\bar{F}(x-y)\bar{F}(y)\,dy\\ &= C\int_{h(x)}^{x-h(x)}\exp(-(Q(x-y)+Q(y)))\,dy\\ &\le Cx\exp(-(Q(h(x))+Q(x-h(x))))\\ &= Cx\bar{F}(h(x))\bar{F}(x-h(x)),\end{aligned}$$

and so

$$r(x)\int_{h(x)}^{x-h(x)} \frac{\bar F(x-y)}{\bar F(x)}\,F(dy) \le Cxr(x)\bar F(h(x))\frac{\bar F(x-h(x))}{\bar F(x)}$$
$$= o(1).$$

Therefore, condition (D3(iii)) also holds. □

3. The boundary class. The conditions (D3(i)), (D3(ii)), and (D3(iii)) depend on being able to choose a suitable little-h function for the distribution F. It is convenient to define $\hbar_F$ as the class of nondecreasing functions $h(x)$ defined on the positive reals such that $0 < h(x) < x/2$ and $h(x) \to \infty$ as $x \to \infty$ and such that the long-tailed property (2) holds for F. There is a problem in that being unable to find a suitable function $r(x)$ and bounding sets $B_i(x)$ for a particular little-h function does not mean that these objects cannot be found for some other little-h function. The class $\hbar_F$ is usually very rich, and so it may be difficult to find a suitable little-h function. However, if $h_1(x)$ belongs to the appropriate class of functions and satisfies the three conditions in (D3), and h_2, belonging to the same class, is such that $h_2(x) > h_1(x)$ for all $x > x_0$, for some x_0, then $h_2(x)$ also satisfies the three (D3) conditions.

This property of the (D3) conditions, which we refer to as an increasing function property (defined precisely below), allows us to construct a boundary class of functions. Functions in the boundary class do not satisfy the little-h condition, but any function that is asymptotically negligible with respect to any function in the boundary class will be in $\hbar_F$. We show that all functions in the boundary class are weakly equivalent (written $H_1(x) \asymp H_2(x)$; see notations in §6.1). This means that the boundary class can be generated by a single function $H(x)$, and all multiples of $H(x)$.

We will show that if the (D3) conditions are satisfied for all multiples of $H(x)$, then they are satisfied for some $h(x)$ belonging to the little-h class of functions, *without having to find the little-h function*. A generator for the boundary class is usually easy to find, and almost trivial for absolutely continuous distributions.

3.1. Definition and properties. First we define precisely what we mean by an increasing (or decreasing) function property.

Definition 3.1. A property depending on a function h, where h belongs to some class of functions, is said to be an increasing (decreasing) function property if when the property is satisfied for h_1, then it is satisfied by any other function h_2 in the class such that $h_2(x) > h_1(x)$, $(h_2(x) < h_1(x))$, for all $x > x_0$, for some x_0.

We observe that the property of long-tailedness (2) is a decreasing function property. We want to describe its upper boundary, when it exists.

Definition 3.2. Let F be a long-tailed distribution. The *boundary class* (for F), $\mathscr{H}$, consists of all continuous, nondecreasing functions $H(x)$ such that $h(x) \in \hbar_F$ if and only if $h(x) = o(H(x))$.

Remark 3.1. In most cases the boundary class for a long-tailed distribution does exist; however, we note that slowly varying functions do not have a boundary class as all nondecreasing functions $h(x)$ defined on the positive reals such that $0 < h(x) < x/2$ and $h(x) \to \infty$ as $x \to \infty$ can act as little-h functions satisfying (2).

We examine the structure of the boundary class $\mathscr{H}$, and show that all functions in $\mathscr{H}$ are weakly tail equivalent.

Proposition 3.1. *Let $H_1(x)$ belong to the boundary class $\mathscr{H}$. Then $H_2(x) \in \mathscr{H}$ if and only if $H_2(x) \asymp H_1(x)$.*

Proof. Clearly, if $H_2(x) \asymp H_1(x)$, then $H_2(x) \in \mathscr{H}$. So, consider a function $H_2(x)$ for which $\liminf(H_2(x)/(H_1(x))) = 0$. We shall construct a function $h_1(x)$ with the long-tail property (2) which is not $o(H_2(x))$. There exists a sequence, tending to infinity, $0 = x_0 < x_1 < \cdots$, and such that the sequence $\epsilon_n := H_2(x_n)/(H_1(x_n))$ is decreasing, $\epsilon_1 < 1$, and $\lim_{n\to\infty} \epsilon_n = 0$.

Let $h_1(x) = H_2(x)$ for $x < x_1$ and let, for $n \ge 1$ and for $x \in [x_n, x_{n+1})$,

$$h_1(x) = H_2(x_n) + (H_1(x) - H_1(x_n)) \cdot \frac{H_2(x_{n+1}) - H_2(x_n)}{H_1(x_{n+1}) - H_1(x_n)}.$$

Clearly, $h_1(x_n) = H_2(x_n)$ for any $n \ge 1$ and $h_1(x)$ is continuous and nondecreasing. Also, for $x \in [x_n, x_{n+1})$,

$$h_1(x) \le \epsilon_n H_1(x_n) + (H_1(x) - H_1(x_n))\frac{\epsilon_{n+1}H_1(x_{n+1}) - \epsilon_n H_1(x_n)}{H_1(x_{n+1}) - H_1(x_n)} \le \epsilon_n H_1(x),$$

so, $h_1(x) = o(H_1(x))$ as $x \to \infty$. Therefore $h_1(x)$ satisfies the long-tail property, but $\limsup(h_1(x)/(H_2(x))) \ge 1$, so $H_2(x) \notin \mathscr{H}$. We can clearly repeat this argument if $\liminf(H_1(x)/(H_2(x))) = 0$. Hence, if $H_2(x) \in \mathscr{H}$, then $H_2(x) \asymp H_1(x)$. □

The three conditions in (D3) depend on a little-h which satisfies (2). The next proposition will show that if the three conditions (D3) are satisfied by *all* functions in the boundary class $\mathscr{H}$, then they are satisfied by at least one h in the long-tail class. Because all functions in $\mathscr{H}$ are weakly-tail equivalent, it will then be sufficient to show that the conditions (D3) hold for all multiples $\{cH(x);\ c \in \mathbb{R}^+,\ cH(x) < x/2\}$ of any particular function $H \in \mathscr{H}$. We shall then say that $H(x)$ generates the boundary class $\mathscr{H}$.

Before proceeding, we define, for any function $f(x)$,

$$\Pi_f(x) = \sup_i \left(\max\left(\frac{\mathbf{P}(\bar{B}_i(f(x)))}{\bar{F}(x)},\ r(x)\bar{F}(f(x)),\ \int_{f(x)}^{x-f(x)} \frac{\bar{F}(x-y)}{\bar{F}(x)} F(dy) \right)\right).$$

Then the three conditions in (D3) are equivalent to

$$\lim_{x\to\infty} \Pi_h(x) = 0. \tag{7}$$

Proposition 3.2. *Let F be a distribution function having a boundary class $\mathscr{H}$. Then there exists some function $h(x) \in \mathscr{h}_F$ satisfying (7) if and only if (7) holds for every $cH(x)$ in place of $h(x)$, where $c > 0$ and $cH(x) < x/2$, where $H(x)$ is any generator of $\mathscr{H}$.*

Proof. Choose any $H(x) \in \mathscr{H}$, and let $c_n = 2^{-n}$, $n \in \mathbb{N}$.

Define an infinite sequence $0 = y_1 < x_1 < y_2 < \cdots$ recursively, for $r \in \mathbb{N}$, by

$$y_1 = 0,$$
$$x_r = \max\Big(y_r + 1, \sup_{x>0}\{x\colon \Pi_{c_r H}(x) > c_r\}\Big),$$
$$y_{r+1} = \inf_{x > x_r+1}\{x\colon H(x) = 2H(x_r)\}.$$

By construction, this sequence tends to infinity.

For $x \ge 0$, define

$$h(x) = \begin{cases} c_r H(x) & \text{for } x \in [y_r, x_r), \\ c_r H(x_r) & \text{for } x \in [x_r, y_{r+1}). \end{cases}$$

Hence, if (7) holds for all (sufficiently small) multiples of $H(x)$, then it holds for $h(x)$, which, by construction, is $o(H(x))$. Conversely, if (7) holds for some $h(x)$, then it holds for any function $g(x)$ such that $h(x) = o(g(x))$, and hence for all functions in class $\mathscr{H}$. □

Because the properties in (D3(i)), (D3(ii)), and (D3(iii)) are all increasing function properties, rather than finding a function h, it suffices to check them for all multiples of H, for any $H \in \mathscr{H}$, the boundary class. This is a much easier proposition than identifying a suitable little-h function.

Remark 3.2. Let F be a distribution function such that $\bar{F}(x) = f_1(x)f_2(x)$, where each of f_1 and f_2 are long-tailed. Let the boundary class for f_i be $\mathscr{H}_i$, $i = 1, 2$, and generated by $H_i(x)$ respectively. Assume that $H_2(x) = o(H_1(x))$. Then the boundary class for F is $\mathscr{H}_2$. If $f_1(x)$ is slowly varying, then the boundary class for F is again $\mathscr{H}_2$.

Proposition 3.3. *Let F be an absolutely continuous long-tailed distribution function with continuous strictly positive density $f(x)$ and hazard rate $q(x) = f(x)/\bar{F}(x)$. Let $H(x) = 1/q(x)$. Then the boundary class of F is generated by $\{cH(x);\ c \in \mathbb{R}^+,\ cH(x) < x/2\}$.*

Proof. By the long-tailed property (2) as $x \to \infty$ and for appropriate $h(x)$ we know that $\bar{F}(x - h(x)) = \bar{F}(x) + o(\bar{F}(x))$. Hence, $x - h(x) = \bar{F}^{-1}(\bar{F}(x)(1 + o(1)))$, and since $\bar{F}^{-1}$ has a derivative at all points in its domain, $h(x) = o(\bar{F}(x)(-\bar{F}^{-1})'(\bar{F}(x)))$, where the negative sign has been introduced to make the function inside the little-o positive. However, $\bar{F}(x)(-\bar{F}^{-1})'(\bar{F}(x)) = 1/q(x)$.

Conversely, if $h(x) = o(1/q(x))$, then it is easy to show that the long-tailed property (2) holds. □

We give some examples of calculating the boundary class.

3.1. Let $\bar{F}(x) = l(x)x^{-\alpha}$, $x > 1$, where $l(x)$ is slowly varying and $\alpha > 0$. The boundary class for $f_1(x) = l(x)$ is the whole space of functions. For $f_2(x) = x^{-\alpha}$ we have $q_2(x) = \alpha/x$. Hence, the boundary class is generated by $\mathscr{H} = \{cx;\ 0 < c < 1/2\}$.

3.2. Let $\bar{F}(x) = \exp(-\gamma x^{\beta})$, $x > 0$, where $0 < \beta < 1$. Then $q(x) = \gamma\beta x^{-1+\beta}$, and the boundary class is generated by $\mathscr{H} = \{cx^{1-\beta};\ c > 0\}$.

3.3. Let $\bar{F}(x) = f_1(x)\exp(-\gamma(\log(x))^\alpha) := f_1(x)f_2(x)$, $x > 1$, where $\alpha \geq 1$ and $f_2(x) = o(f_1(x))$. We note that this class of functions includes regular variation as in Example 3.1 above, and also log-normal. We need only consider $f_2(x)$. We then find that the boundary class is generated by $\mathcal{H} = \{cx(\log(x))^{1-\alpha};\ c > 0\}$.

As we observe in the footnote in §2, if we have a random variable X distributed with subexponential distribution F with $h(x)$ being any little-h function, and another r.v. X_1 with distribution F_1 such that

$$\bar{F}_1(h(x)) = o(\bar{F}(x)), \tag{8}$$

then $\mathbf{P}(X + X_1 > x) \sim \mathbf{P}(X > x)$, regardless of the dependence structure between X and X_1.

Clearly (8) is an increasing function property. Hence, if we redefine $\Pi_f(x) = \bar{F}_1(f(x))/(\bar{F}(x))$ and then combine Proposition 3.2 with a simple induction we get:

Proposition 3.4. *Let X be a random variable with subexponential distribution F and boundary class $\mathcal{H}$ generated by $H(x)$. Let $X_1, \ldots, X_n$ be r.v.'s with distribution functions $F_1, \ldots, F_n$ such that, for all $i = 1, \ldots, n$ and all $c > 0$, $\bar{F}_i(cH(x)) = o(\bar{F}(x))$, then*

$$\mathbf{P}(X + X_1 + \cdots + X_n > x) \sim \mathbf{P}(X > x),$$

regardless of the dependence structure between the X's.

3.2. The boundary class and auxiliary functions. Our concept of the boundary class is also closely related to the concept of an auxiliary function, introduced by de Haan [6]; see also, for example, Asmussen and Klüppleberg [3], Embrechts et al. [9], Resnick [18]. However, the boundary class can exist when there is no auxiliary function, or when the conditions of the previous proposition are not met, and hence, it is a more general concept.

The concept of an auxiliary function was introduced to characterize distributions that lie in the maximum domain of attraction of the Gumbel extreme value distribution, $MDA(\Lambda)$. A positive function $a(x)$ is an auxiliary function for the distribution function $F(x) \in MDA(\Lambda)$ (with $F(x) < 1\ \forall\, x \in \mathcal{R}$) if and only if

$$\lim_{x\to\infty} \frac{\bar{F}(x + ta(x))}{\bar{F}(x)} = e^{-t}, \quad \forall\, t \in \mathcal{R}. \tag{9}$$

If an auxiliary function exists, all such functions are asymptotically equivalent, possible choices are the reciprocal of the hazard rate and the mean excess function (see, for example, Embrechts et al. [9]) and any auxiliary function is in the boundary class.

The concept of the auxiliary function may be extended. For instance, for regularly varying $\bar{F}(x) \in \mathcal{R}_{-\alpha}$ (see §6.2 for a definition) we have an auxiliary function $a(x)$, which may be taken to be $a(x) = x$, satisfying

$$\lim_{x\to\infty} \frac{\bar{F}(x + ta(x))}{\bar{F}(x)} = (1+t)^{-\alpha}, \quad \forall\, t \in \mathcal{R}, \tag{10}$$

and again any auxiliary function is in the boundary class.

The concept of the boundary class is more general than that of the auxiliary function. We can construct a (subexponential) distribution function belonging to the class of intermediately regularly varying distributions (see §6.2 for a definition) as follows: Let $c_1(x) = 2 + \sin(\log x)$ and $\bar{F}(x) = \max(1, c_1(x)x^{-\alpha})$, $\alpha > 0$. It is straightforward to check that this has boundary class generated by $\mathcal{H} = \{H(x) = cx;\ 0 < c < 1/2\}$ but the limit in (10) does not exist. Indeed, if $h(x) = o(x)$, then $c_1(x + h(x) \sim c_1(x)$, but $\lim_{x\to\infty}(c_1(x + H(x))/(c_1(x)))$ does not exist.

We can also construct a similar example for what we might call an intermediate Weibull distribution. Let $c_2(x) = 2 + \sin(x^\beta)$, $0 < \beta < 1$, and $\bar{F}(x) = \max(1, c_2(x)e^{-x^\beta})$. Then F is a subexponential distribution function, with a well-defined boundary class generated by $\mathcal{H} = \{H(x) = cx^{1-\beta};\ c > 0\}$, but the limit in (9) does not exist.

In both the previous examples, Proposition 3.3 can be used to find the boundary class because of the smoothness of the functions $c_1(x)$ and $c_2(x)$. However, it would be easy to replace these functions with long-tailed functions with the same oscillatory behaviour but without the smoothness. This would not affect the boundary class, but would no longer allow the boundary class to be generated by the reciprocal of the hazard rate.

4. Examples of conditionally independent subexponential random variables.

4.1. Example 1. Let ξ_i, $i = 1, 2, \ldots, n$, be i.i.d. with common distribution function $F_\xi \in R_{-\alpha}$ (for a definition of the class of regularly varying distributions of degree α, see §6.2). Let η be independent of the ξ_i

and have distribution function $F_\eta \in R_{-\beta}$, where $\alpha \neq \beta$. Define $X_i = \xi_i + \eta$ for $i = 1, 2, \ldots, n$, and let the reference distribution be $\bar{F}(x) = x^{-(\alpha\wedge\beta)}$. Then, from well-known properties of independent subexponential random variables, we have, for $i = 1, 2, \ldots, n$,

$$\mathbf{P}(X_i > x) \sim \bar{F}(x).$$

Conditional on the sigma algebra $\mathcal{G} = \sigma(\eta)$, the X_i are independent.

For our reference distribution, the boundary class is $\mathcal{H} = \{cx,\ c > 0\}$. Now, the random variables $\mathbf{P}(X_i > x \mid \mathcal{G})/(\bar{F}(x)) \le 1/(\bar{F}(x))$ are unbounded as $x \to \infty$. If we try to satisfy the condition (5) without using the bounded sets $B(x)$, we need to ensure that, for all $x > 0$, almost surely,

$$\frac{\mathbf{P}(X_i > x \mid \mathcal{G})}{\bar{F}(x)} \le r(x).$$

If we take $r(x) = 1/(\bar{F}(x))$, then condition (D3(ii)) is not satisfied, since for any $c > 0$,

$$r(x)\bar{F}(H(x)) = c^{-(\alpha\wedge\beta)}.$$

Hence, we need to use bounded sets. Let $B(x) = \{\eta \le x/2\}$. This satisfies condition (D3(i)), if and only if $\alpha < \beta$: for any $c > 0$,

$$\frac{\mathbf{P}(\bar{B}(H(x))}{\bar{F}(x)} = \frac{\mathbf{P}(\eta > cx/2)}{\bar{F}(x)} = (2/c)^\beta x^{(\alpha\wedge\beta)-\beta}.$$

Clearly, in the case $\alpha < \beta$, we may take $r(x)$ as a constant, $r(x) = 2^\beta$.

The condition $\alpha < \beta$ agrees with arguments on X_i taken from the standard theory of independent subexponential random variables.

4.2. Example 2. Let η be a random variable with uniform distribution in the interval $(1, 2)$. Conditional on $\mathcal{G} = \sigma(\eta)$, let X_i, $i = 1, 2, \ldots, n$, be i.i.d. with common distribution function

$$\bar{F}_{\xi|\eta}(x) = (1+x)^{-\eta}, \quad x > 0.$$

Routine calculations show that

$$\mathbf{P}(X_i > x) \sim \frac{1}{x\log(1+x)} \equiv \bar{F}(x),$$

where F is our reference distribution. The boundary class is again $\mathcal{H} = \{cx,\ c > 0\}$.

For all $x > 1$ we have, almost surely,

$$\begin{aligned}\frac{\mathbf{P}(X_i > x \mid \mathcal{G})}{\bar{F}(x)} &\le \frac{\mathbf{P}(X_i > x \mid \eta = 1)}{\bar{F}(x)}\\ &= (1+1/x)\log(1+x)\\ &\le r(x) \equiv 2\log(1+x).\end{aligned}$$

Routine calculations show that, for all $0 < c < 1/2$, condition (D3(iii)) is satisfied, and also condition (D3(ii)).

In this example there has been no need to define bounding sets, or equivalently, we can take $B(x) = \Omega$ for all $x > 0$.

4.3. Example 3. In this example we again consider a Bayesian type situation. Let X_i, $i = 1, 2, \ldots, n$, be identically distributed, conditionally independent on parameter β, with conditional distribution F_β given by

$$\bar{F}_\beta(x) = \exp(-\gamma x^\beta), \quad \gamma > 0,$$

where β is drawn from a uniform distribution on (a, b), $0 \le a < 1$, $a < b$. The unconditional distribution of X_i, F_X, is then

$$\bar{F}_X(x) = \frac{1}{(b-a)\log x}(E_1(\gamma x^a) - E_1(\gamma x^b)),$$

where $E_1(x) = \int_x^\infty (e^{-u}/u)\,du$.

We now consider separately the two cases (i) $0 < a < 1$ and (ii) $a = 0$. We start first with case (i).

We find

$$\bar{F}_X(x) \sim \bar{F}(x) := \frac{\exp(-\gamma x^a)}{(b-a)\gamma x^a \log x}.$$

The boundary class $\mathscr{H}$ is the same as for the Weibull distribution,

$$\mathscr{H} = \{cx^{1-a},\ c > 0\}.$$

We shall take $B(x) = \Omega$ for all $x > 0$, and

$$r(x) = \gamma(b-a)x^a \log x.$$

Note that $Q(x) = -\log(\bar{F}(x)) = \log(\gamma(b-a)) + x^a + a\log x + \log\log x$ is convex, and that

$$xr(x)\bar{F}(H(x)) = \frac{x^{1-a^2}\log(x)\exp(-\gamma c^a x^{a(1-a)})}{c^a \log(cx)} \to 0,$$

as $x \to \infty$ for all $c > 0$. Hence, by Lemma 2.3, the conditions in (D3) are met, and the principle of the single big jump holds.

We now consider case (ii), with β distributed uniformly on the interval $(0, b)$. The reference distribution is now

$$\bar{F}(x) = \frac{E_1(1)}{b\log x} := \frac{k}{\log x},$$

where $E_1(1) \approx 0.21938\ldots$.

Because $\bar{F}$ is slowly varying, there is no boundary class, but the class of functions satisfying the long-tailed property (2) is $\{h(x) = O(x)\}$. Therefore, to satisfy (D3(ii)) we need to choose $r(x)$ such that $\lim_{x\to\infty}(r(x)/\log x) = 0$. For the bounded sets $B(x)$, the problems clearly occur near $\beta = 0$, so we may try sets of the form $B(x) = \{\beta \in (a(x), b)\}$, with $a(x) \to 0$ as $x \to \infty$.

To satisfy (5) we need, for each $x > 0$ and for $\beta \in (a(x), b)$,

$$\exp(-\gamma x^{a(x)}) < \exp(-\gamma x^{\beta}) \le \frac{kr(x)}{\log(x)},$$

which cannot be true if both $a(x) \to 0$ and $r(x)/\log x \to 0$ as $x \to \infty$.

Hence, the conditions of (D3) cannot be met.

The question now arises whether, in this case, the principle of the single big jump still holds. The answer is no. To see why, we again consider the representation (1), for the sum of two independent identically distributed subexponential random variables X_1, X_2. For simplicity, we shall consider the case where $b = \gamma = 1$.

Considering the representation in (1), we have

$$P_1(x) = \int_0^1 d\beta \int_0^x \beta u^{\beta-1}\exp(-u^\beta)\,du \int_0^x \beta v^{\beta-1}\exp(-v^\beta)\,dv \mathbf{1}(u+v > x).$$

Making the substitution $u = xy$, $v = xz$, we have

$$\begin{aligned} P_1(x) &= \int_0^1 d\beta x^{2\beta} \int_0^1\int_0^1 \beta^2 y^{\beta-1} z^{\beta-1}\exp(-x^\beta(y^\beta + z^\beta))\mathbf{1}(y+z>1)\,dy\,dz \\ &\le \int_0^1 d\beta x^{2\beta}\exp(-x^\beta)\int_0^1\int_0^1 \beta^2 y^{\beta-1}z^{\beta-1}\mathbf{1}(y+z>1)\,dy\,dz \\ &= \int_0^1 d\beta x^{2\beta}\exp(-x^\beta)J(\beta), \end{aligned}$$

where $J(\beta) = \mathbf{P}(Y_1^{1/\beta} + Y_2^{1/\beta} > 1)$ and $Y_1, Y_2 \sim U(0,1)$ are i.i.d. As $\beta \to 0$, $J(\beta) \to 0$, so for any $\delta > 0$ there exists $\epsilon > 0$ such that $J(t) \le \delta$ for all $t \le \epsilon$. Hence,

$$\begin{aligned} P_1(x) &\le \left(\int_0^\epsilon + \int_\epsilon^1\right) d\beta x^{2\beta}\exp(-x^\beta)J(\beta) \\ &\le \int_0^\epsilon d\beta x^{2\beta}\exp(-x^\beta)\delta + o(\exp(-x^{\epsilon/2})). \end{aligned}$$

However,

$$\begin{aligned}\int_0^{\epsilon} d\beta x^{2\beta}\exp(-x^{\beta}) &\le \int_0^1 d\beta x^{2\beta}\exp(-x^{\beta})\\ &= \frac{1}{\log(x)}\int_1^{\infty} t\exp(-t)\,dt\\ &= \frac{1}{2\log(x)}.\end{aligned}$$

Therefore,

$$P_1(x) = o(\bar{F}(x)),$$

and the principle of the big jump holds. However,

$$P_2(x) = \int_0^1 \exp(-2x^{\beta})\,d\beta = \frac{1}{\log(x)}\int_2^{2x}\frac{e^{-u}}{u}\,du$$

so that

$$\mathbf{P}(X_1 > x,\ X_2 > x) \sim \frac{E_1(1)}{E_1(2)}\mathbf{P}(X_1 > x).$$

Hence, the principle of the *single* big jump does not hold. We note that this result is related to Theorem 2.2 in Albrecher et al. [1].

4.4. Example 4. For $i = 1, 2, \ldots, n$, let $X_i = \xi_i\eta_1\eta_2\cdots\eta_i$, where the $\{\xi_i\}$ are i.i.d, and the $\{\eta_i\}$ are i.i.d. and independent of the $\{\xi_i\}$. Then, conditional on the σ-algebra generated by $\{\eta_1, \ldots, \eta_n\}$, the $\{X_i\}$ are independent. Let the $\{\xi_i\}$ have common distribution function F in the intermediately regularly varying class, $F_\xi := F \in$ IRV, (see §6 for a definition), and let the $\{\eta_i\}$ have common distribution function F that is rapidly varying, $F_{\eta_1} \in \mathcal{R}_{-\infty}$, (see §6 for a definition). This is related to the example given in Laeven et al. [16]. In their example, the $\{\xi_i\}$ were chosen to belong to the class $\mathcal{D}\cap\mathcal{L}$ (again, see §6). We have chosen the slightly smaller class of intermediate regular variation because:

(i) examples which lie in the $\mathcal{D}\cap\mathcal{L}$ class that do not lie in the IRV class are constructed in an artificial manner;

(ii) the IRV class of functions has a common boundary class, and hence, is suitable for general treatment under our methodology.

The boundary class for F is $\mathcal{H} = \{cx,\ 0 < c < 1/2\}$.

By Lemma 6.1 the class $\mathcal{R}_{-\infty}$ is closed under product convolution; hence, for each $i = 1, 2, \ldots, n$ we have X_i is of the form $X_i = \xi_i\eta$ where the d.f. of η, $F_\eta \in \mathcal{R}_{-\infty}$. Then by Lemma 6.2, each X_i has d.f. $\bar{F}_{X_i}(x) \asymp \bar{F}(x)$.

As we noted in Remark 2.3, the results of our propositions follow through with the asymptotic condition in (D2) replaced with weak equivalence.

We now proceed to the construction of the bounding sets, $B(x)$. For (5) to hold, we need to restrict the size of η. By Lemma 6.3 we can choose $\epsilon > 0$ such that $\bar{F}_\eta(x^{1-\epsilon}) = o(\bar{F}(x))$. For such an ϵ we choose $B(x) = \{\eta \le x^{1-\epsilon}\}$. Then, for any $H(x) = cx \in \mathcal{H}$, $0 < c < 1/2$, condition (D3(i)) requires

$$\mathbf{P}(\bar{B}(H(x))) = \bar{F}_\eta((cx)^{1-\epsilon}) = o(\bar{F}_\eta(x^{1-\epsilon})) = o(\bar{F}(x)),$$

as required.

Now consider (5):

$$\mathbf{P}(X_i > x \mid \mathcal{G})\mathbf{1}(B(x)) \le r(x)\bar{F}(x)\mathbf{1}(B(x)).$$

This implies that the choice for $r(x)$ satisfies

$$\frac{\mathbf{P}(\xi_i > x/\eta \mid \eta \le x^{1-\epsilon})}{\bar{F}(x)} \le \frac{\mathbf{P}(\xi_i > x^{\epsilon})}{\bar{F}(x)} = \frac{\bar{F}(x^{\epsilon})}{\bar{F}(x)} \le r(x).$$

Taking $r(x) = \bar{F}(x^{\epsilon})/(\bar{F}(x))$, for any $H(x) = cx \in \mathcal{H}$,

$$r(x)\bar{F}(H(x)) = \frac{\bar{F}(x^{\epsilon})}{\bar{F}(x)}\bar{F}(cx) = o(1),$$

and

$$r(x)\int_{cx}^{(1-c)x}\frac{\bar F(x-y)}{\bar F(x)}F(dy)\le\frac{\bar F(x^\epsilon)\bar F(cx)}{\bar F^2(x)}\bar F(cx)=o(1).$$

Hence, all the conditions of (D3) are met, and the principle of the single big jump holds; that is:

$$\mathbf{P}(X_1+\cdots+X_n>x)\sim\sum_{i=1}^n\bar F_{X_i}(x).$$

REMARK 4.1. If, in addition, F is continuous, then Theorem 3.4(ii) of Cline and Samorodnitsky[5] shows that the restriction $F_\eta\in\mathscr{R}_{-\infty}$ can be eased to $\bar F_\eta=o(\bar F)$.

4.5. Example 5. In this example we consider random variables $X_1,\ldots,X_n$ with lognormal marginals.

First we recall some facts about lognormally distributed random variables. An r.v. $X\sim LN(\mu,\sigma^2)$ if $X=e^Y$ and $Y\sim N(\mu,\sigma^2)$. The distribution function of X is $F_X(x)\sim(\sigma/(\sqrt{2\pi}\log x))\exp(-(1/(2\sigma^2))(\log x-\mu)^2)$. If two r.v.'s $X_1,\sim LN(\mu_1,\sigma_1^2)$, $X_2,\sim LN(\mu_2,\sigma_2^2)$, then X_1 has a heavier tail than X_2, in the sense that $\bar F_{X_2}(x)=o(\bar F_{X_1}(x))$, if and only if either $\sigma_2<\sigma_1$ or both $\sigma_2=\sigma_1$ and $\mu_2<\mu_1$. The boundary class for $F_X(x)$ is generated by $H(x)=x/(\log x)$. We observe that if X_1 has a heavier tail than X_2, then, for all $c>0$, $\bar F_{X_2}(cH(x))=o(\bar F_{X_1}(x))$, and by reference to Proposition 3.4 this suggests we need only consider the dependence structure as it relates to those X_i which have the heaviest tail.

So, first, let the r.v.'s which have the heaviest distribution be $X_1=e^{Y_1},\ldots,X_m=e^{Y_m}$, each distributed with $X_i\sim LN(\mu,\sigma^2)$ with common distribution function F. We specify the dependence structure by assuming that $(Y_1,\ldots,Y_m)\sim MVN((\mu,\ldots,\mu),\Sigma)$ where Σ is of full rank. We perform a factor analysis and write each $Y_i=t_{i1}Z_1+\cdots+t_{ik}Z_k+W_i$, where, for $1\le j\le k$, the Z_j are i.i.d. standard normal and, independently, for $1\le i\le m$, $W_i\sim N(\mu_i,\sigma_i^2)$ are independent normal r.v.'s. Because we place no restriction on k and Σ is of full rank, this factor analysis can always be performed (nonuniquely) such that the W_i are nondegenerate; that is, $\sigma_i>0$ for all $i=1,\ldots,m$.

We take $\mathscr{G}$ to be the sigma algebra generated by $Z_1,\ldots,Z_k$, and conditional on this the X_i are independent. Our reference distribution is F with $\mathbf{P}(X_i>x)\sim\bar F(x)$ for all $i=1,\ldots,m$. Each X_i can be written as $X_i=\psi_i\xi_i$, where $\psi_i=e^{t_{i1}Z_1+\cdots+t_{ik}Z_k}\sim LN(0,s_i^2)$ and $s_i^2=t_{i1}^2+\cdots+t_{ik}^2$ and $\xi_i=e^{W_i}\sim LN(\mu,\sigma_i^2)$. Then, for each i, $s_i^2+\sigma_i^2=\sigma^2$.

We choose bounding sets $B_i(x)=\{\psi_i\le x^\delta\}$, where $1>\delta>\max_i(s_i^2/\sigma^2)$ ensures that $\mathbf{P}(\bar B_i(cH(x)))=o(\bar F(x))$ for all $c>0$.

Then, given $B_i(x)$,

$$\mathbf{P}(X_i>x\mid B_i(x))\le\mathbf{P}(\xi_i>x^{1-\delta})\le\bar F(x^{1-\delta}),$$

for large enough x. So we may take

$$r(x)=\frac{\bar F(x^{1-\delta})}{\bar F(x)}$$

which is monotonically increasing for large enough x.

Because the lognormal reference distribution has a hazard function that is eventually concave, and, for all $c>0$,

$$xr(x)\bar F(cH(x))=x\frac{\bar F(x^{1-\delta})}{\bar F(x)}\bar F\left(\frac{cx}{\log x}\right)\to 0$$

as $x\to\infty$, then we can apply Proposition 2.3 and conditions (D3(ii)) and (D3(iii)) hold. Hence, $\mathbf{P}(X_1+\cdots+X_m>x)\sim m\mathbf{P}(X_1>x)$.

Now we apply Proposition 3.4, with $X_1+\cdots+X_m$ in place of X and $X_{m+1},\ldots,X_n$ as the lighter-tailed r.v.'s and conclude that

$$\mathbf{P}(X_1+\cdots+X_n>x)\sim m\mathbf{P}(X_1>x),$$

and the principle of the single big jump holds.

The example of lognormal random variables with Gaussian copula was studied in Asmussen and Rojas-Nanayapa [4]. Although our results agree with their results, there are some small differences in the assumptions. In Asmussen and Rojas-Nanayapa [4] it is assumed that the whole dependence structure is a Gaussian copula, but in our setup the dependence of the lighter-tailed random variables is not specified. However, we feel that, in practice, this is an unimportant point. More important, we assume that the covariance matrix Σ for the heaviest random variables is of full rank. This is not assumed in Asmussen and Rojas-Nanayapa [4]. The condition in their paper is that each pair $X_i = e^{Y_i}$, $X_j = e^{Y_j}$, where $i \neq j$ and $\mathrm{Var}(Y_i) = \mathrm{Var}(Y_j)$ has correlation between Y_i and Y_j of $\rho_{ij} < 1$. This does not imply that the covariance matrix is of full rank. Indeed, a simple example, for $i = 1, 2, 3$, is given by Y_1, Y_2 i.i.d standard normals, and $Y_3 = (Y_1 + Y_2)/\sqrt{2}$. It is clear that our methodology cannot deal with this example, as only conditioning on the trivial sigma algebra can make these independent. In this sense the result in Asmussen and Rojas-Nanayapa [4] is more general.

5. Real-valued random variables. We wish to extend our investigation beyond nonnegative random variables to conditionally independent subexponential random variables taking real values. To deal with this situation we need to add another condition to those enumerated in §2 at (D1), (D2), and (D3). Again we let F be a reference subexponential distribution, and we let h be a function satisfying (2).

(D4) For each $i, j \geq 1$ we have that

$$\mathbf{P}(X_i > x + h(x), X_j \leq -h(x)) = o(\bar{F}(x)).$$

We then have the following extension of Proposition 2.1.

Proposition 5.1. *Let X_i, $i = 1, 2, \ldots,$ be real-valued random variables satisfying conditions* (D1), (D2), (D3), *and* (D4) *for some subexponential F concentrated on the positive half-line and for some $h(x)$ satisfying* (2). *Then*

$$\mathbf{P}(X_1 + \cdots + X_n > x) \sim \sum_{i=1}^{n} \mathbf{P}(X_i > x) \sim \left(\sum_{i=1}^{n} c_i\right) \bar{F}(x).$$

Proof. The proof follows the general outline of the proof of Proposition 2.1. The derivation of an upper bound for $\mathbf{P}(X_1 + X_2 > x)$ remains as in Proposition 2.1. For the lower bound we have:

$$\begin{aligned}\mathbf{P}(X_1 + X_2 > x) \geq{}& \mathbf{P}(X_1 > x + h(x),\ X_2 > -h(x)) + \mathbf{P}(X_2 > x + h(x),\ X_1 > -h(x)) \\ &- \mathbf{P}(X_1 > x + h(x),\ X_2 > x + h(x)) \\ ={}& \mathbf{P}(X_1 > x) + \mathbf{P}(X_2 > x) + o(\bar{F}(x)),\end{aligned}$$

where we have used the long-tailedness of X_1 and X_2, condition (D4) and $\mathbf{P}(X_1 > x + h(x),\ X_2 > x + h(x)) = o(\bar{F}(x))$ for the same reasons as in Proposition 2.1. Hence $\mathbf{P}(X_1 + X_2 > x) \sim \mathbf{P}(X_1 > x) + \mathbf{P}(X_2 > x)$, and the rest of the proof follows by induction. □

We can see that the only reason for condition (D4) is to deal with the lower bound, and hence no change is needed to show the two following generalisations of Lemma 2.1 and Proposition 2.2.

Lemma 5.1. *We let $\{X_i\}$ be as in Proposition 5.1 and satisfying* (D1), (D2), (D3), *and* (D4). *Then, for any $\epsilon > 0$, there exist $V(\epsilon) > 0$ and $x_0 = x_0(\epsilon)$ such that, for any $x > x_0$ and $n \geq 1$,*

$$\mathbf{P}(S_n > x) \leq V(\epsilon)(1 + \epsilon)^n \bar{F}(x).$$

Proposition 5.2. *If, in addition to the conditions of Lemma 5.1, τ is an independent counting random variable such that $\mathbf{E}(e^{\gamma\tau}) < \infty$ for some $\gamma > 0$, then*

$$\mathbf{P}(X_1 + \cdots + X_\tau > x) \sim \mathbf{E}\left(\sum_{i=1}^{\tau} \mathbf{P}(X_i > x)\right).$$

To show that condition (D4) is both nonempty and necessary we construct an example where it fails to hold and the principle of the single big jump fails.

Example 6. Consider a collection of nonnegative i.i.d random variables $\{Z_i\}_{i\geq 0}$ such that each Z_i has the distribution of a generic independent nonnegative random variable Z, and $\mathbf{P}(Z > x) = 1/x^\alpha$ for $x \geq 1$. Also consider a collection, independent of the Z_i, of nonnegative i.i.d random variables $\{Y_i\}_{i\geq 0}$ such that each Y_i has the distribution of a generic independent nonnegative random variable Y, and $\mathbf{P}(Y > x) = 1/x^\beta$ for $x \geq 1$, where $\alpha > \beta > 1$. For $i \geq 1$ let $X_i = Z_i - Y_i Z_{i-1}$.

First we show that the X_i satisfy conditions (D1), (D2), and (D3).

For any $i \geq 1$ we have that $\mathbf{P}(X_i > x) \sim \bar{F}(x) := 1/x^\alpha$ for $x \geq 1$, and we recall that the boundary class for F is $\mathscr{H} = \{cx;\ 0 < c < 1/2\}$. We take $\mathscr{G} = \sigma(\{Y_i Z_{i-1}\}_{i\geq 1})$, and then, conditional on $\mathscr{G}$, the X_i are independent. To show that condition (D3) is met we need to consider the random variables $\mathbf{P}(X_i > x \mid \mathscr{G})$, so consider

$$\mathbf{P}(X_i > x \mid Y_{i+1} Z_i = w) \leq \mathbf{P}(Z_i > x \mid Y_{i+1} Z_i = w) = \mathbf{P}(Z > x \mid YZ = w).$$

We calculate that

$$\frac{\mathbf{P}(Z > x \mid YZ = w)}{\bar{F}(x)} = \begin{cases} x^{\alpha}\left(\dfrac{x^{-\alpha+\beta} - w^{-\alpha+\beta}}{1 - w^{-\alpha+\beta}}\right) & \text{for } 1 < x \le w, \\ 0 & \text{for } x > c. \end{cases}$$
$$\le x^{\beta} := r(x).$$

Clearly $r(x)\bar{F}(cx) = o(1)$ for all $0 < c < 1/2$.

Also, straightforward estimation shows that $\int_{cx}^{(1-c)x}(\bar{F}(x-y)F(dy)/(\bar{F}(x))) = O(x^{-\alpha})$, and hence condition (D3(iii)) is met. We take $B(x) = \Omega$ for all $x \ge 0$ so that there is nothing to show for (D3(i)).

We now consider condition (D4). For any $i \ge 1$, and any $H(x)$ satisfying (2),

$$\begin{aligned} \mathbf{P}(X_i > x + h(x), X_{i+1} < -h(x)) &= \mathbf{E}(\mathbf{P}(Z_1 - W > x + h(x),\, Z_2 - Y_2 Z_1 < -h(x) \mid W)) \\ &\ge \mathbf{E}(\mathbf{P}(Z_1 > x + h(x) + W,\, Z_2 < Z_1 - h(x) \mid W)) \\ &\ge \mathbf{E}(\mathbf{P}(Z_1 > x + h(x) + W,\, Z_2 < x + W \mid W)) \\ &= \mathbf{E}(\mathbf{P}(Z_1 > x + h(x) + W)\mathbf{P}(Z_2 < x + W \mid W)). \end{aligned}$$

Hence, by Fatou's lemma,

$$\begin{aligned} &\liminf_{x\to\infty} \frac{\mathbf{P}(X_i > x + h(x),\, X_{i+1} < -h(x))}{\bar{F}(x)} \\ &\quad \ge \liminf_{x\to\infty} \mathbf{E}\left(\frac{\mathbf{P}(Z_1 > x + h(x) + W \mid W)}{\mathbf{P}(Z_1 > x)}\mathbf{P}(Z_2 < x + W \mid W)\right) \\ &\quad \ge \mathbf{E}\left(\liminf_{x\to\infty}\left(\frac{\mathbf{P}(Z_1 > x + h(x) + W \mid W)}{\mathbf{P}(Z_1 > x)}\mathbf{P}(Z_2 < x + W \mid W)\right)\right) \\ &\quad = 1, \end{aligned}$$

and so condition (D4) is not met.

Finally, we show that the conclusion of Proposition 5.1 fails in this example:

$$\begin{aligned} \mathbf{P}(X_1 + \cdots + X_n > x) &= \mathbf{P}(Z_n + (1 - Y_n)Z_{n-1} + \cdots + (1 - Y_2)Z_1 - Y_1 Z_0 > x) \\ &\le \mathbf{P}(Z > x). \end{aligned}$$

6. Notations and definitions.

6.1. Notation. For a random variable (r.v.) X with distribution function (d.f.) F we denote its tail distribution by $\mathbf{P}(X > x) := \bar{F}(x)$. For two independent r.v.'s X and Y with d.f.'s F and G we denote the convolution of F and G by $F * G(x) := \int_{-\infty}^{\infty} F(x-y)G(dy) = \mathbf{P}(X + Y \le x)$, and the n-fold convolution $F * \cdots * F := F^{*n}$. In the case of nonnegative random variables we have $F * G(x) := \int_0^x F(x-y)\, G(dy) = \mathbf{P}(X + Y \le x)$.

Throughout, unless stated otherwise, all limit relations are for $x \to \infty$. Let $a(x)$ and $b(x)$ be two positive functions such that

$$l_1 \le \liminf_{x\to\infty} \frac{a(x)}{b(x)} \le \limsup_{x\to\infty} \frac{a(x)}{b(x)} \le l_2.$$

We write $a(x) = O(b(x))$ if $l_2 < \infty$ and $a(x) = o(b(x))$ if $l_2 = 0$. We say that $a(x)$ and $b(x)$ are weakly equivalent, written $a(x) \asymp b(x)$, if both $l_1 > 0$ and $l_2 < \infty$, and that $a(x)$ and $b(x)$ are (strongly) equivalent, written $a(x) \sim b(x)$, if $l_1 = l_2 = 1$.

For any event $A \in \Omega$ we define the indicator function of event A as

$$\mathbf{1}(A) = \begin{cases} 1 & \text{if } A \text{ occurs}, \\ 0 & \text{otherwise}. \end{cases}$$

We use $X \vee Y$ to mean $\max(X, Y)$, and $X \wedge Y$ to mean $\min(X, Y)$.

6.2. Definitions and basic properties. We recall some well-known results on long-tailed and subexponential distributions and their subclasses.

A nonnegative r.v. X with distribution F is heavy-tailed if $\mathbf{E}(e^{\gamma X}) = \infty$ for all $\gamma > 0$, and long-tailed if $\bar{F}(x+1) \sim \bar{F}(x)$. The class of long-tailed distributions, $\mathscr{L}$, is a proper subclass of the heavy-tailed distributions. The distribution F is long-tailed if and only if there exists a positive function h, monotonically increasing to zero and satisfying $h(x) < x$ such that

$$\bar{F}(x - h(x)) \sim \bar{F}(x), \tag{11}$$

An r.v. X with distribution F is subexponential if $\bar{F}^{*2}(x) \sim 2\bar{F}(x)$. This is equivalent to $\bar{F}^{*n}(x) \sim n\bar{F}(x)$ for any $n \geq 1$. The class of subexponential distributions, $\mathscr{S}$, is a proper subclass of the class $\mathscr{L}$ of long-tailed distributions. If X_1 and X_2 are independent and have common d.f. F, then $\mathbf{P}(X_1 + X_2 > x) \sim \mathbf{P}(\max(X_1, X_2) > x)$.

A nonnegative r.v. X with distribution F supported on the positive half-line is subexponential if and only if

(i) F is long-tailed;

(ii) for any $h(x) < x/2$ tending monotonically to infinity,

$$\int_{h(x)}^{x-h(x)} \bar{F}(x-y)\, F(dy) = o(\bar{F}(x)). \tag{12}$$

A positive function l is slowly varying if, for all $\lambda > 0$, $l(\lambda x) \sim l(x)$. A distribution function F belongs to the class of regularly varying distributions of degree α, $\mathscr{R}_{-\alpha}$, if $\bar{F}(x) \sim l(x)x^{-\alpha}$ for some slowly varying function l. A distribution function F belongs to the class of extended regular varying distributions, ERV, if $\liminf_{x\to\infty}(\bar{F}(\lambda x)/\bar{F}(x)) \geq \lambda^{-c}$ for some $c \geq 0$ and all $\lambda \geq 1$. A distribution function F belongs to the class of intermediately regular varying distributions, IRV, also called consistent variation by some authors, if $\lim_{\lambda\downarrow 1}\liminf_{x\to\infty}(\bar{F}(\lambda x)/(\bar{F}(x))) = 1$. A distribution function F belongs to the class of dominatedly regular varying distributions, $\mathscr{D}$, if $\liminf_{x\to\infty}(\bar{F}(\lambda x)/(\bar{F}(x))) \geq 0$ for some $\lambda > 1$. A distribution function F belongs to the class of rapidly varying distributions, $\mathscr{R}_{-\infty}$, if $\lim_{x\to\infty}(\bar{F}(\lambda x)/(\bar{F}(x))) = 0$ for all $\lambda \geq 1$. We have the proper inclusions (see Embrechts et al. [9])

$$\mathscr{R}_{-\alpha} \subset \mathrm{ERV} \subset \mathrm{IRV} \subset \mathscr{D} \cap \mathscr{L} \subset \mathscr{S} \subset \mathscr{L}.$$

The following three lemmas stem from Cline and Samorodnitsky [5], Tang and Tsitsiashvili [20], Laeven et al. [16] and are used in the development of Example 4 in §4.

Lemma 6.1. *The class $\mathscr{R}_{-\infty}$ is closed under product convolution.*

Lemma 6.2. *Let X and Y be two independent positive r.v.'s, and let the distribution function of X, $F_X \in \mathscr{D} \cap \mathscr{L}$, and let that of Y, $F_Y \in \mathscr{R}_{-\infty}$. Let the distribution function of XY be F_{XY}. Then $\bar{F}_{XY}(x) \asymp \bar{F}_X(x)$.*

Lemma 6.3. *If $F \in \mathscr{D}$ and $F_\eta \in \mathscr{R}_{-\infty}$, then there exists $\epsilon > 0$ such that $\bar{F}_\eta(x^{1-\epsilon}) = o(\bar{F}(x))$.*

Acknowledgments. The authors are grateful to Stan Zachary and to Søren Asmussen for many fruitful and insightful conversations. They are also grateful to the associate editor and anonymous referees for many helpful and perceptive comments which have greatly improved the paper.

References

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

[2] Alink, S., M. Löwe, M. Wüthrich. 2004. Diversification of aggregate dependent risks. *Insurance Math. Econom.* **35** 77–95.

[3] Asmussen, S., C. Klüppelberg. 1996. Large deviation results for subexponential tails, with applications to insurance risk. *Stochastic Processes Their Appl.* **64** 103–125.

[4] Asmussen, S., L. Rojas-Nanayapa. 2008. Asymptotics of sums of lognormal random variables with Gaussian copula. *Statist. Probab. Lett.* **78** 2709–2714.

[5] Cline, D., G. Samorodnitsky. 1994. Subexponentiality of the product of independent random variables. *Stochastic Processes Their Appl.* **49** 75–98.

[6] de Haan, L. 1970. *On Regular Variation and Its Application to Weak Convergence of Sample Extremes*. *Math. Centre Tracts*, Vol. 32. Mathematisch Centrum, Amsterdam.

[7] Denisov, D., S. Foss, D. Korshunov. 2004. Tail asymptotics for the supremum of a random walk when the mean is not finite. *Queueing Systems* **46** 15–33.

[8] Denuit, M., C. Genest, E. Marceau. 1999. Stochastic bounds on sums of dependent risks. *Insurance Math. Econom.* **25** 85–104.

[9] Embrechts, P., C. Klüppleberg, T. Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, London.

[10] Foss, S., S. Zachary. 2003. The maximum on a random time interval of a random walk with long-tailed increments and negative drift. *Ann. Appl. Probab.* **13** 37–53.

[11] Foss, S., T. Konstantopoulos, S. Zachary. 2007. The principle of a single big jump: Discrete and continuous time modulated random walks with heavy-tailed increments. *J. Theoret. Probab.* **20** 581–612.

[12] Goovaerts, M., R. Kaas, Q. Tang, R. Vernic. 2004. The tail probability of discounted sums of pareto-like losses in insurance. *Proc. 8th Int. Congress Insurance: Math. Econom., Rome*.

[13] Klüppleberg, C. 1989. Subexponential distributions and characterizations of related classes. *Probab. Theory Related Fields* **82** 259–269.

[14] Kortschak, D., H. Albrecher. 2009. Asymptotic results for the sum of dependent nonidentically distributed random variables. *Methodology Comput. Appl. Probab.* **11** 279–306.

[15] Ko, B., Q. Tang. 2008. Sums of dependent nonnegative random variables with subexponential tails. *J. Appl. Probab.* **45** 85–94.

[16] Laeven, R., M. Goovaerts, T. Hoedemakers. 2005. Some asymptotic results for sums of dependent random variables, with actuarial applications. *Insurance Math. Econom.* **37** 154–172.

[17] Mitra, A., S. Resnick. 2008. Aggregation of risks and asymptotic independence. Eprint arXiv 0807.4200v2, http://arxiv.org/PS_cache/arxiv/pdf/0807/0807.4200v2.pdf.

[18] Resnick, S. 1987. *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

[19] Tang, Q. 2006. Insensitivity to negative dependence of the asymptotic behaviour of precise large deviations. *Electronic J. Probab.* **11** 107–120.

[20] Tang, Q., G. Tsitsiashvili. 2003. Precise estimates for the ruin probability in finite horizon in a discrete-time model with heavy-tailed insurance and financial risks. *Stochastic Processes Their Appl.* **108** 299–325.